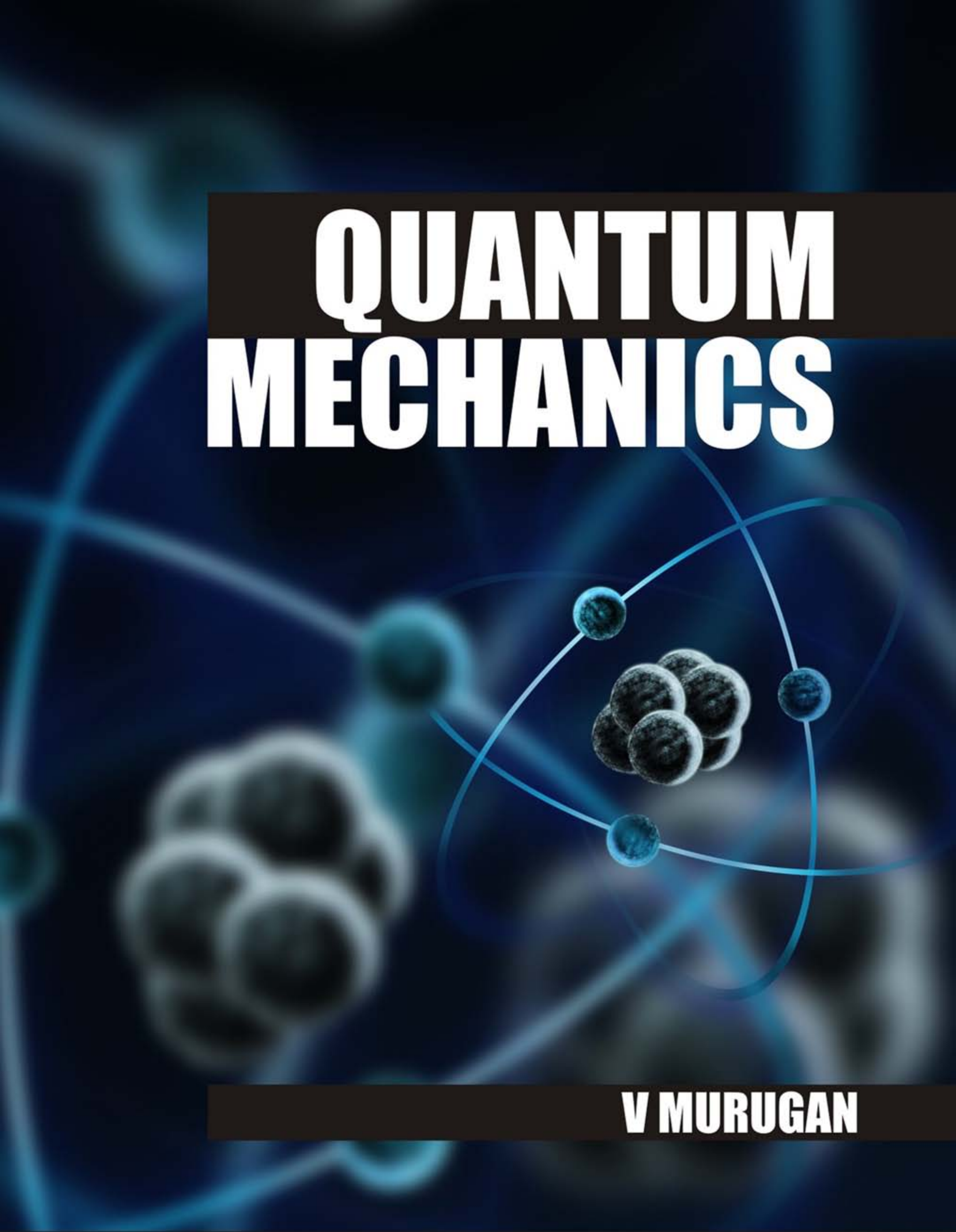


QUANTUM MECHANICS

V MURUGAN

ALWAYS LEARNING

PEARSON

Quantum Mechanics

Quantum Mechanics

V. MURUGAN

Head (Retired)
Department of Physics
Ramakrishna Mission Vivekananda College
Chennai

PEARSON

Chennai • Delhi

ISBN 978-81-317-7362-8
eISBN 978-93-325-3730-9

Head Office: A-8(A), Sector 62, Knowledge Boulevard, 7th Floor, NOIDA 201 309, India
Registered Office:11 Local Shopping Centre, Panchsheel Park, New Delhi 110 017, India

To my mother V. Rajeswari
and
my brother V. Chandrasekaran

Brief Contents

Contents

16 Relativistic Wave Equation 629

Preface

Quantum Mechanics, the new age science,
Has struck wonder since
With the superposition principle indeed
The cat could be half-alive or half-dead.

One more book on quantum mechanics! There are plenty of excellent textbooks on quantum mechanics in the market. Good lecture notes are available on the internet. Then, where is the need for another book on quantum mechanics? Every author feels that there is a lacuna that is not filled by the existing books and believes that his/her book will bridge this gap. Of course, this belief is likely to remain in the realms of wishful thinking.

This book has been written by a teacher, rather than by an expert on quantum mechanics. The justification for writing this book, as also its thrust centres on this point. I taught quantum mechanics for M.Sc students at RKM Vivekananda College, Chennai, for more than two-and-a-half decades. Quantum mechanics is very abstract and mathematical. It requires lengthy articulation on the part of a teacher to clarify the subtleties of various concepts. Also, many students find it difficult to do all the mathematical manipulations themselves. During my lectures, I have often been requested by students as well as teachers to suggest a book where all the calculations can be found in detail. This book makes an attempt to provide all the calculations and explain the subtleties involved in a student-friendly way. However, learning quantum mechanics is not going to be easy even if one finds a book where all the calculations have been worked out in detail. The only way for a student to learn quantum mechanics is by working out all the details himself/herself. There is no shortcut or an alternative method other than working out the details by oneself. The student has to spend sufficient time to reflect on the nature of quantum mechanical ideas. Yet, students do require some help from the teachers or books and the degree of such help depends on the level of the student.

This book has been written with the average student in mind, who wants to understand and appreciate quantum mechanics. Therefore, it may sound too elementary and repetitive at some places. The book presents the mathematical manipulations in full detail while laying considerable emphasis on the conceptual framework of quantum theory.

I have tried to impress upon the reader that the world view generated by quantum mechanics is drastically different from that of classical physics. The best way to appreciate the implications of quantum mechanics is to compare it with classical physics. Such comparisons have been provided throughout the book.

A number of topics such as optical theorem and Kramer’s relation have been presented in more depth than what is generally covered in a textbook at this level. Though these materials are available from other sources, they are not easily accessible to students, in general.

The order in which the various topics are discussed is quite conventional. It is better to first get acquainted with Schrödinger equation and its applications to simple potentials in one dimension. This exposition can be followed by operator formalism. But even in the case of formalism, it is better to do it in the language of wave mechanics. Introducing Dirac's abstract state-vector formalism after the exposure of wave mechanics, helps students to grasp it better.

Classical physics is reviewed in Chapter 1. The evolution of major concepts of classical physics is presented in this chapter, without going into the historical accounts, with a view to prepare the reader to get a proper perspective on quantum mechanics. Feynman's thought experiment on double slits is portrayed in detail. We use the words 'particle' and 'wave' to describe the physical entities in both macroscopic and microscopic world. However, their meanings are different. At the microscopic level, the objects exhibit non-classical particle and non-classical wave property. It is amazing to see how Feynman's thought experiment brings out all the non-classical nature of wave–particle duality and the state of superposition.

In Chapter 2, Schrödinger equation and boundary conditions for the wave functions are developed. The concept of wave packet in the context of uncertainty relation is presented in detail. In Chapter 3, Schrödinger equations for different one-dimensional potentials are solved bringing out the consequences of wave nature in the microscopic world. In particular, it delves into the non-classical aspect of the tunnelling phenomenon.

In Chapter 4, the mathematical background which is required to work on quantum mechanics is provided (in a minimal way). Its major stress is on the exposition of the Hilbert space structure of quantum mechanics, though it is an elementary account without any mathematical rigour. Chapter 5 is devoted to the formalism of quantum mechanics. It brings out the significance of superposition in quantum mechanics and expounds on the uncertainty principle, the concepts of compatible observables and simultaneous measurements. It elucidates the quantum mechanical postulates and shows how all the postulates together imply a radical departure of quantum mechanics from classical physics. The contrast between classical and quantum physics is brought out in detail.

Chapter 6 is on the simple harmonic oscillator. Both wave mechanics and the ladder operator methods are spelt out. The coherent state is explained in detail. Chapter 7 dwells upon the orbital angular momentum. In Chapter 8, Schrödinger equation is applied to various three-dimensional potentials. The square-well potential and the hydrogen atom are presented in detail. There are two appendices: one on associated Laguerre polynomial and another on Kramer's relation.

Chapter 9 takes a close look at Dirac's abstract state vector formalism. The second part of this chapter focuses on symmetry and associated unitary transformation. We choose to work with active transformation. The concept of internal parity is developed at length. Chapter 10 develops angular momentum and deals at length with Clebsh–Gordan coefficients and their evaluation using various methods. The tensor operators have been presented at the introductory level. Chapter 11 stresses on the principle of indistinguishability and the consequences of this principle in statistical mechanics. The density matrix and its properties are analyzed in detail.

Chapter 12 consists of two parts: In the first part, we consider an electron in the presence of constant magnetic field. Next, we consider gauge invariance, which plays a major role in elementary particle physics. However, we restrict ourselves to the gauge invariance of Schrödinger equation only. The Aharonov–Bohm effect is presented. The second part of this chapter deals with two-state problems. There are a number of systems which are unrelated but their state space is two-dimensional. The fact that their state space is two-dimensional Hilbert space causes them to share certain common features irrespective of other details. Fluctuations in ammonia molecule's configuration, magnetic resonance and neutrino oscillation have been considered here as systems illustrating the two-state problem.

Chapter 13 deals with approximations for solving time-independent Schrödinger equation. It examines the time-independent perturbation theory and variational method and describes the WKB approximation and its applications in a detailed manner. Chapter14 considers various types of approximation to solve the time-dependent Schrödinger equation with stress on the time-dependent perturbation theory and its applications. It also discusses the adiabatic theorem and the Berry phase at length. Chapter 15 describes the scattering theory. A major part of this chapter is devoted to partial wave analysis. It also elaborates on the Born approximation.

Chapter 16 is on the relativistic wave equations. It explores Dirac's negative energy sea at length though this idea has been superseded by the quantum field theory. The corrections to Bohr energy levels of hydrogen atom have been worked out using both Klien–Gordan equation and Dirac equation. Solving Dirac equation for hydrogen atom requires considerable manipulations and the chapter presents all calculations in full detail.

I am indebted to Ramakrishna Mission Vivekananda College, Chennai, where I had taught quantum mechanics for about 25 years. Interactions with my former colleagues of physics department were indeed a pleasant experience. I also benefitted a lot from the department library. I take this opportunity to thank my colleague, K. Sethusankar, who was very helpful to me all these years. I am indebted to my former colleagues, R. Sudharsan and J. Segar, who went through the manuscript patiently. Their critical observations made me rewrite a number of sections with better clarity and rigour. I am grateful to R. Sudharsan for proofreading all the chapters painstakingly with patience (The clerihew given above is by Sudharsan). I am obliged to Ganesh Rajkumar for helping me in software-related work. I thank my brothers and sisters for their support. I thank my wife R. G. Shanthi and my daughter M. Nithya Bharathi for their cooperation while I was writing this book.

V. Murugan

1

Evolution of Physics: Classical to Quantum

A brief review of development of key concepts of physics is presented here. There is an evolution of ideas in our understanding of the nature of the physical universe. The conceptual frameworks in physics are continuously evolving. Even in modern physics, this evolution is taking place continuously. Our understanding of an electron has changed a lot in the last hundred years. The concept of photon has undergone a number of changes from the light quantum of Einstein. Tracing such developments in full details will be a big story by itself which is not our purpose. This chapter is restricted to a brief account of conceptual developments with an eye on quantum mechanics. Here, we concentrate on the basic concepts of Newtonian mechanics, development of wave theory of light, and atomic structure of matter. A rapid survey of early quantum theory, namely the theory from 1900 to 1925, is also presented. Finally, the uncertainty principle and wave–particle duality are presented in detail to highlight the radical departure of quantum physics from classical physics.

1.1 CLASSICAL MECHANICS

The central equation of classical mechanics is Newton's equation of motion.

$$m\frac{d^2\mathbf{r}}{dt^2} = \mathbf{f}(\mathbf{r},t)$$

This equation determines the position of an object $\mathbf{r}(t)$ at various instants of time for a given set of forces $\mathbf{f}(\mathbf{r},t)$ acting on the object. $\mathbf{f}(\mathbf{r},t)$ is the known quantity and $\mathbf{r}(t)$ is the unknown quantity in this equation. This is a second-order differential equation. The general solution contains the constants of integration which can be determined from the initial conditions of the object. By initial conditions, we mean the values of the position $\mathbf{r}(0)$ and the momentum $\mathbf{p}(0)$. (Instead of momentum, velocity also could have been chosen). Let us take a simple example to illustrate these ideas. Our example is a simple pendulum executing simple harmonic motion as shown in Fig. 1.1.

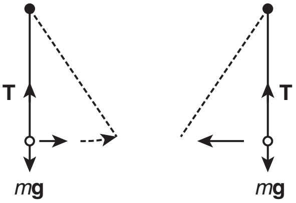


Fig. 1.1 Simple pendulum

The forces acting on the bob are tension **T** and the gravitational force $m\mathbf{g}$ downwards as shown in the Fig. 1.1. So, we have

$$m\frac{d^2\mathbf{r}(t)}{dt^2} = \mathbf{T} + m\mathbf{g}$$

Skipping the details which can be found in any textbook on classical mechanics, let us write the equation for small angles as given below:

$$m\frac{d^2x}{dt^2} = -m\frac{g}{l}x$$

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

The solution of this equation is $x(t) = A\cos\omega t + B\sin\omega t$.
The momentum $p(t)$ is

$$p(t) = -m\,\omega A\sin\omega t + m\omega B\cos\omega t$$

Here A and B are the constants of integration. To determine A and B we should know the initial conditions, namely $x(0) = x_0$, $p(0) = p_0$. From the above equations, we get

$$x(t) = x_0\cos\omega t + \frac{p_0}{m\omega}\sin\omega t$$

The time evaluation of the system is determined by the initial conditions. For a bob at rest in the equilibrium, we can give a gentle kick and make it oscillate. This corresponds to the initial condition $x_0 = 0$ and $p_0 = mv_0 \neq 0$ and the resulting solution is $x(t) = p_0/m\mathrm{w}\sin\mathrm{w}t$. On the other hand, we can pull the bob to an extreme position x_0 and withdraw our hold. This corresponds to the initial condition $x_0 \neq 0$ and $p_0 = 0$ and the solution becomes $x(t) = x_0\cos\omega t$. We can visualize many such situations. The lesson to be learnt from this example is that the forces acting on the object and the initial conditions together determine the particular trajectory of a particle.

Let us leave aside this example now. Newton's equation of motion has been successfully applied to a wide variety of situations. For instance, its success in explaining the planetary motion is remarkable and it is able to account for the motion of the comets well. The motion of all the objects on the earth can be analysed successfully using Newton's equation of motion. All of these applications of Newton's equation of motion share some common features which lead to certain philosophical points of views. Let us summarize them as follows:

1. **Trajectory:** The primary aim or utility of Newton's equation of motion is to determine the position of an object at all instants of time. The function $\mathbf{r}(t)$ is known as trajectory of the object. This is also known as time evolution of the system. To determine the time evolution of a system, we should know all the forces acting on the system with the initial conditions $\mathbf{r}(0)$ and $\mathbf{p}(0)$.
2. **Determinism:** Let us re-interpret the Newton's equation of motion as follows: If we know the forces acting on the system and the initial position and momentum at $t = 0$, ($\mathbf{r}(0)$ and $\mathbf{p}(0)$), then we know the position and momentum of the object at all instants of time later. In other words, the position and momentum at all instants of time ($\mathbf{r}(t)$ and $\mathbf{p}(t)$) are uniquely determined. The forces on the object and the initial conditions together leave no scope for randomness or indeterminism with respect to the motion of the object. So in a sense, the future of the system is determined completely once the force acting to the object and the initial conditions are fixed. This is known as determinism.
   In the case of simple pendulum these ideas can be explained as follows:
   Forces on the object: Tension $\mathbf{T}$ and gravitational force $m\mathbf{g}$.
   Initial conditions: $x(0) = x_0$ and $p(0) = p_0$.
   Position of the bob at any time t with: $x(t) = x(0)\cos\omega t + \frac{p(0)}{m\omega}\sin\omega t$

Therefore, the position of the pendulum cannot be anything other than that dictated by the expression for $x(t)$.

This kind of determinism is lost in quantum mechanics, which will be discussed in later chapters. Just for the sake of completeness, it should be mentioned that recent developments (chaos theory) indicate presence of a new kind of indeterminism even in the case of classical mechanics. However, this indeterminism is different from the indeterminism of quantum mechanics.

3. **Causality:** A related idea is the notion of causality. In a naive sense, for any event, there should be a cause. One expects identical causes to lead to identical effects. In Newtonian mechanics the causes are the forces and the effects are responses like accelerations. Let us express the principle of causality as follows: Identical systems under identical forces with identical initial conditions evolve in time identically. If they do not evolve in time identically then we have only randomness.

   Let us explain this idea again with the example of the simple pendulum. Let us consider thousands of identical simple pendulums. The threads and the bobs are identical. The length of the pendulum in all the cases is 1 m. The forces acting on each bob are the same. All of them are displaced by 5 cm from the equilibrium position at $t = 0$ and then dropped. So the initial conditions are $x_0 = 5$ $p_0 = 0$ for all pendulums. So they are identical in every respect. The time evolution for all the pendulums is determined by $x(t) = 5\cos\pi t$ So for all the pendulums, the positions of the bob are the same at any particular moment.

   Let us now do an experiment in which we measure the position of bobs of such pendulums, at a particular instant of time. Leave aside the experimental details like accuracy of the instruments and other sources of errors. From our earlier analysis of motion of a pendulum, we should get identical experimental results for all such pendulums. For instance, at $t = 1$ sec, we should get only one experimental value namely $x = -5$ cm. This is the essence of causality in classical physics.

   Let us consider a different scenario in which we get a different kind of results. Instead of a single value, our experimental values are a collection of random values for the position of bobs, say ranging from –5 cm to 5 cm. For instance at $t = 1$ sec, suppose the experimental results look like this: $x = -5$ cm for some pendulums, $x = 2.3$ cm for some, and so on. This is a statistical distribution of values of position. What do we conclude from these experimental results? We have to bear in mind that for pendulums obeying Newton's laws of motion, the experimental outcomes should be identical (only one value), provided the pendulums are identical in every respect. Taking this aspect into account, we can arrive at the following conclusions. If they obey Newton's laws motion, then the statistical distribution of experimental results implies that the pendulums are not identical. If we still insist that the pendulums are identical in every respect then they don't obey Newton's law of motion. Perhaps they obey quantum mechanics!

1.2 LIGHT

Fig. 1.2 Expression wave theory of light

The study of nature of light is one of the most fascinating aspects in physics. It has a very long history. We have to start with Newton and Huygens. Geometrical optics was already well known at that time. It relies heavily on the rectilinear propagation of light. There were two theories for light in seventeenth century. They were Newton's corpuscular theory and Huygens wave theory as shown in Fig. 1.2.

In Newton's corpuscular theory, light is assumed to be made of small particles called corpuscles. The explanation of geometrical optics in terms of corpuscles theory is very simple. The concept of wave motion as such is not new at that time. Interference and diffraction are considered to be benchmark for any kind of wave motion. There cannot be a wave motion without these phenomena. The problem for wave theory of light is that interference or diffraction was not observed at that time (apart from Newton's rings). Added to these factors, Newton's monumental achievements lent strong credence to corpuscular theory. The supporters of corpuscular theory included eminent scientists such as Biot, Laplace, Poisson, and others. However, slowly more evidences for wave theory of light were emerging. One of the important developments in this direction is Young's double-slit experiment in 1801. Getting interference fringes by combining light from two sources is not simple. It requires considerable skill. The coherence between the light waves from two sources is very important. The dimension of slit should be very small (fraction of a millimetre). Producing coherent light waves through double-slit arrangement by Young was really ingenious.

The observation of Poisson spot was a major milestone in establishing the wave nature of light. When Fresnel submitted an essay on the wave theory of light in response to a competition announced by French Academy of Sciences to determine the nature of light, Poisson, one of the judges, wanted to disprove the wave nature of light. He deduced from Fresnel's theory that wave nature of light would lead to a bright spot at the centre of the geometrical shadow behind the circular object (see Fig 1.3), which we do not get normally. However, another judge, Arago, performed an experiment for a 2 mm-wide-circular metallic disk and verified the existence of Poisson spot.

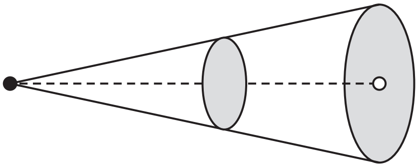


Fig. 1.3 Poisson spot

Many more such evidences for wave nature of light continued to emerge in the period between 1800 and 1900. However, the strong evidence for wave nature of light came from a different route, namely Maxwell's equations. Faraday introduced the concept of fields to describe the interaction between charges and currents. It was not clear, at that time, whether such entities really existed or not. In Faraday's scheme, two charges q_1 and q_2 do not interact directly. The charge q_1 creates a field in space around it, and the interaction is only between this field and charge q_2 (A closer analogy will be the interaction between the sun and the earth. The earth does not directly interact with the sun. It interacts only with the gravitational field of the sun.) These ideas were not seriously taken initially. However, Maxwell took these ideas seriously and developed a set of equations, which are now known as Maxwell's equations to determine electric and magnetic field for a given charge and current densities. One of the important consequences of Maxwell's equations is that electric and magnetic fields can propagate in the form of wave motion even in the absence of charge density and current density. Maxwell found the velocity of this wave and the velocity of the light to be same. So, he identified light as electromagnetic wave, which was proved later experimentally by Hertz.

The immense body of knowledge (both experimental and theoretical) about light led to the following understanding: The interference and diffraction experiments along with Maxwell's equations show that light is a wave. The experimental evidence for wave nature of light was so strong that the wave nature of light was taken as irrefutable. As far as the energy of the light is concerned, it is assumed to be distributed over the wavefront, which continues to expand as it moves away from the source. For instance, 1 Joule of energy from a point source (which may be sodium light) is distributed over spherical surfaces of area $4\pi R^2$, and so when R becomes large, the energy on a small part of the spherical surface can become as low as possible, close to zero. For instance, the energy content of light, due to this source, in 1 cm^2 area of the wavefront at a distance of 5000 km is 3.4×10^{-19} J and at 7000 km is 1.7×10^{-19} J. (Compare this with $h\nu = 3.4 \times 10^{-19}$ J for sodium light.)

In hindsight, let us make the following observation: The source of light used in these experiments is very intense. If one describes light in terms of photons, the intensity of light from a sodium lamp will be of the order $10^{21}\ h\nu$ per second where $h\nu$ for sodium light is 3.4×10^{-19} J. With the technology available at that time, sources of intensity of only this order of magnitude can be produced. However, now the technology has changed a lot. So, today, we can raise a number of questions like whether it is possible to reduce the intensity of light indefinitely close to zero or is there a lowest value of energy beyond which light intensity cannot be reduced. (For instance, is it possible to have the sodium light of intensity less than 3.4×10^{-19} J/sec?). Will low intensity light also produce an interference pattern? If so, will there be any change in the interference pattern? Such questions were not even raised at that time. These are the questions which have become now very important in understanding the quantum theory of light.

1.3 ATOMS

Between 1800 and 1900, a substantial body of knowledge about atoms was generated from the developments in chemistry. Number of phenomena in chemistry can be explained in a simple way using the concept of atoms. The kinetic theory of matter and the statistical mechanics developed by Boltzmann attempted to understand thermodynamics in terms of the properties of atoms.

It was realized that the atoms are so small that it is impossible to observe them directly. All the knowledge of atoms can be obtained only from the studies of properties of macroscopic objects. The standard technique is to propose certain models for the atoms and their interaction, and based on this model, the properties of macroscopic objects are derived which can be experimentally verified.

For instance, in kinetic theory of matter, the equation of state $PV = n\ RT$ is obtained by assuming that the atoms are like billiard balls constantly moving and colliding with each other, obeying Newton's equation of motion. For instance, for diatomic molecules, the ratio of specific heats C_p/C_V is found to be 1.4. This can be easily explained if we assume the shape of a diatomic molecule is similar to a dumbbell.

The recognition of this aspect will help us to appreciate quantum mechanics in a better way. The atoms in classical physics are very small indivisible objects obeying Newton's laws of motion.

1.4 SOME VIEWS BASED ON CLASSICAL PHYSICS

It should be mentioned that the body of knowledge of the physical universe developed from 1600 to 1900 is enormous compared to the knowledge development before 1600. This body of knowledge is more assured and justified. They cannot be thrown away like Aristotle's theoretical understandings. This unprecedented success in physical sciences created many impressions in human mind.

The universe around us exists independent of us. There exists an objective reality whether we know it or not. The universe is knowable, though we do not know a lot. The universe is composed of matter and electromagnetic field filling the space. Matter is composed of atoms which are indivisible. The laws of physics are the same for all time and at all scale. The same laws of physics can be used to explain the motion of objects on earth, motion of stellar objects in macroscopic scale, and the motion of atoms in microscopic scale. All the motions are causally related. Identical objects under identical conditions behave identically. If we measure a property of identical systems in identical state, we should get identical outcomes. As far as measurement of a property of an object is concerned, the classical physics admits uncertainties. There are many sources of uncertainties like faulty instruments.

However, if one raises the question whether these uncertainties can be eliminated, the answer is that in principle there exists no limit to which it can be reduced. This is a hypothetical question and a hypothetical answer. The reader should not mistake these views as some kind of declarations.

Before we conclude this brief review of classical physics, let us reiterate the two important categories of ideas. They are the concepts of particles and the waves.

Particles

The moment we see the word 'particle' it is the image of a tennis ball or a bullet that comes to our mind immediately. They are objects of definite size and shape. They are localized lumps as shown in Fig. 1.4. Their discrete nature is reflected in the fact that they are countable (i.e. one ball, two balls, five balls, etc.).

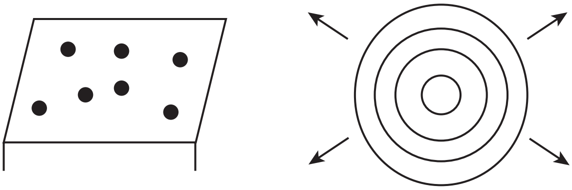


Fig. 1.4 Particles and waves

Waves

The moment we see the word 'wave', it is the image of expanding ripples in a pond of water that comes to our mind. It is an extended object occupying a large volume of space. The benchmark of wave motion is interference and diffraction phenomena.

There is no problem in deciding to which category a particular object belongs to when we see objects like ball or sea waves. We run into a serious problem when we deal with objects like atoms or light which cannot be observed directly but whose nature can be obtained only inferentially.

There were developments which challenged the well-established classical physics of which one of them led to the birth of quantum theory. So, our next task is to trace the developments which led to modern quantum physics. We begin with black body radiation.

1.5 BLACK BODY RADIATION

Any hot body emits electromagnetic radiations. When radiation is incident on an object it absorbs a part, transmits a part and reflects a part of the radiation.

Let a_λ, r_λ, t_λ represent the fractions of light absorbed, reflected and transmitted where $a_\lambda + r_\lambda + t_\lambda = 1$. For a black body, $r_\lambda = 0$, $t_\lambda = 0$, $a_\lambda = 1$. Intuitively, a body which absorbs all the radiation incident on it will look black. However, the definition of a black body as an object for which $a_\lambda = 1$ is more stringent. It is very difficult to produce such a perfect absorber. There was a necessity to find one such object, since Kirchhoff showed that the energy $e_{\lambda B}$, emitted per unit second per unit area from such objects has an universal character. For a long time, how to produce such a perfect black body itself was a challenge since no black-coloured object behaves as a perfect absorber. It was Wien and Lummer in 1895 suggested that a cavity with a small hole is a good model of a perfect black body.

This model of black body has interesting consequences. The cavity and the radiation within the cavity are very much similar to standard example of a thermodynamic system, say, a gas or air in a cylinder with piston. The fact that the radiation and the cavity walls together are in thermodynamic equilibrium leads to some general results which emerge as the consequences of the laws of thermodynamics. Let u_ν be the spectral energy density of the radiation inside the cavity at temperature T. From thermodynamical considerations it can be shown that the spectral energy density is isotropic and homogeneous. It is independent of size and shape of the cavity. It is independent of the nature of the material of the wall. So, the black body radiation inside the cavity itself is a thermodynamic system by itself., It depends only on temperature. So, for black body radiation, we have $u_\nu = u_\nu(T)$. A number of other conclusions can also be reached based on thermodynamics. The total energy density is $u(T)$ defined as $u(T) = \int_0^\infty u_\nu(T)d\nu$. It can be shown that $u(T) = \sigma T^4$. However, there is a limit beyond which thermodynamics alone is not sufficient. We cannot get the spectral distribution function $u_\nu(T)$ from thermodynamics alone. But, even there, we have the following results from thermodynamics alone:

1. The energy density of the black body radiation is maximum at a particular wavelength $\lambda_{\max}$. This depends on the temperature of the black body radiation. It can be established.

$$\lambda_{\max} T = \text{Const}$$

2. The overall functional form of $u_\nu(T)$ is $u_\nu(T) = A\nu^3 f\left(\frac{\nu}{T}\right)$.

   The exact analytical form of the function $u_\nu(T)$ can be obtained only from statistical mechanics by taking into account the nature of radiation and its interaction with matter. There are two distribution functions:

$$\text{Wien distribution function: } u_\nu(T) = \alpha\nu^3 e^{-b\nu/T}$$

$$\text{Rayleigh–Jeans distribution function: } u_\nu(T) = \frac{8\pi\nu^2}{c^3} k_B T$$

Both these expressions worked partially. Wien's distribution worked in the high-frequency range whereas Rayleigh–Jeans' distribution worked in low-frequency range. Note that Rayleigh–Jeans' distribution diverges as ν increases. This is known as ultraviolet catastrophe, a term which was coined by Ehrenfest in 1911.

The correct expression was given by Planck. The Planck's distribution is given by

$$u_\nu(T) = \frac{8\pi h\nu^3}{c^3} \cdot \frac{1}{(e^{h\nu/k_s T} - 1)}$$

This expression becomes Wien's distribution in the high-frequency limit and Rayleigh–Jeans' distributions at low-frequency range. There is an interesting story of how Planck arrived at this expression. In Germany Otto Lummer and Ernst Pringsheim performed the experiment to determine $u_\nu(T)$. Simultaneously Rubens and Kurlbaum determined $u_\nu(T)$ for various temperatures. Rubens and Kurlbaum told Planck their experimental results in a tea party (or in a private lunch hosted by Planck). The results agreed with Wien's distribution and Reyleigh–Jeans' distribution only partially. Planck spent rest of the day to guess the correct expression which he succeeded. He sent his formula (a guess work) to Ruben who two days later informed Planck confirming the fact that Planck's formula fits very well with the experimental data.

Later on, Planck provided the theoretical basis for this formula. We know from Maxwell's electrodynamics that the charges under the accelerated motion can emit electromagnetic radiation. Charges can also be set into accelerated motion by absorbing energy from electromagnetic radiation. So, Planck assumed the cavity walls and the black body radiation together as a system of collection of harmonic oscillators and electromagnetic radiation in thermodynamic equilibrium. For such a system, the average thermal energy density $\rho(\nu,T)$ of the harmonic oscillators and the energy density of the $u_\nu(T)$ black body radiation are related by

$$u_\nu(T) = \frac{8\pi h\nu^3}{c^3}\rho(\nu,T)$$

So, rest of his calculations were reduced to determine the average thermal energy density $\rho(\nu,T)$ of the oscillations to fit his formula for $u_\nu(T)$. His arguments are very involved and relied heavily on thermodynamic relations. He obtained an expression for the entropy of the harmonic oscillators (using guess work again) which will lead to correct expression for $\rho(\nu,T)$. Now he tried to obtain correct expression from basic principles. He used Boltzmann's definition of entropy $S = k_B ln\Omega$, where Ω is the number of microstates. In the process, he was forced to conclude that the required expression for entropy can be obtained provided an oscillator absorbs or emits radiation in terms of integral multiples of energy $h\nu$. So, an oscillator has to absorb or radiate energy $h\nu, 2h\nu, 3h\nu\ldots$. Planck wrote this paper in 1900.

Planck's formula cannot be derived from any fundamental principles of classical physics. The real radical departure from classical physics is restricting the allowed values of the harmonic oscillator to a set of discrete values $h\nu$, $2h\nu$, $3h\nu\ldots$. Within classical physics, this restriction cannot be justified. Planck accepted this result reluctantly. He admitted the inevitability of this result. He did not try to give any physical significance of the discreteness of allowed energy values with regard to nature of either oscillator or electromagnetic field. However, he realized the revolutionary character of restricting the allowed energy values to a set of discrete values.

1.5.1 Derivation of Planck Formula

A more simple way of deriving Planck radiation is to use the lead given by Lord Rayleigh. We provide an outline of this derivation to bring out the enormous differences it has in the spectral distribution function due to discrete and continuous values of allowed energy of a mode.

As mentioned in the previous section, the black body radiation itself is considered as a separate thermodynamic system. The role of the cavity wall is restricted to confine the radiation to a definite volume V. The radiation is assumed to exist as standing waves. Each standing wave is termed as a mode. In classical physics, each mode can occur with energy E_ν ranging from 0 to ∞ (in principle). In fact, as a part of thermal system, the energy of each mode goes on fluctuating. So we can take only the thermal average energy of a normal mode. The average energy density $u_\nu d\nu$ is

$$u_\nu d\nu = N_\nu d\nu \bar{E}_\nu$$

Where $N_\nu d\nu$ is the number of modes between ν and $\nu + d\nu$ and $\bar{E}_\nu$ is the thermal average energy of each mode. From classical electromagnetism, it can be shown that the number of modes $N_\nu d\nu$ is given by

$$N_\nu d\nu = \frac{8\pi V}{c^3}\nu^2 d\nu$$

Here V is the volume of the cavity. Let us discuss the calculation of $\bar{E}_\nu$ in detail, for it brings the differences $\bar{E}_\nu$ due to discrete values and continuous values.

1.5.2 Average Energy of a Mode

From statistical mechanics, we know that for a system in thermal equilibrium, the probability of the system to be in state of energy E is given is $e^{-E/k_BT}/Z$ where Z is the partition function. Using this result for black body radiation, we have the probability of a mode to have energy E is given by

$$P(E) = \frac{e^{-E/k_BT}}{Z}$$

Since E (in classical physics) can take any value from 0 to ∞, Z is given by $Z = \int_0^\infty e^{-E/k_BT}\,dE$

The average energy of the mode is

$$\bar{E}_\nu = \int_0^\infty EP(E)dE = \frac{\int_0^\infty Ee^{-E/k_BT}}{\int_0^\infty e^{-E/k_BT}}$$

$$= k_BT$$

Therefore, the spectral energy density of the black body radiation is

$$u_\nu d\nu = N_\nu d\nu k_BT = \frac{8\pi\nu^2}{c^3}d\nu.k_BT$$

This is Rayleigh distribution. We have not got Planck's distribution. Obviously, we have made some wrong assumptions in this derivation. The expression $N_\nu d\nu$ has been calculated using well-established Maxwell's equations and so its validity is not to be disputed. The other possibility is to consider the expression for $\bar{E}_\nu$. Then, we have to modify the expression for $P(E)$. The key assumption here is that E can take any value from 0 to ∞. Instead of allowing E to take any value from 0 to ∞, let us restrict it to a set of values $h\nu, 2h\nu, 3h\nu \ldots$. Then the integrations in the expressions for the partition function and the average energy value become a summation.

$$\int_0^\infty e^{-E/k_BT}dE \to \sum_n e^{-E_n/k_BT}$$

$$\text{and} \int_0^\infty Ee^{-E/k_BT}dE \to \sum_n E_n e^{-E_n/k_BT}$$

The probability of finding the normal mode in energy $E_n = nhv$ is given by

$$p_n = \frac{e^{-E_n/k_BT}}{Z}$$

where now $Z = \sum_n e^{-E_n/k_BT}$.

The average energy $\bar{E}_\nu$ is given by

$$\bar{E}_\nu = \sum_n p_n E_n = \frac{\sum_{n=0}^\infty E_n e^{-E_n/k_BT}}{\sum_{n=0}^\infty e^{-E_n/k_BT}}$$

$$= -\frac{d}{d\beta}\ln\ \mathrm{Z}$$

Here, $Z = 1 + e^{-\beta h\nu} + e^{-\beta 2h\nu} + \ldots$with $\beta = \dfrac{1}{k_B T}$. This can be evaluated and $\bar{E}_\nu$ is found to be

$$\bar{E}_\nu = \frac{h\nu}{e^{h\nu/k_B T} - 1}$$

$$\therefore u_\nu = \frac{8\pi\nu^2}{c^3} \cdot \frac{h\nu}{(e^{h\nu/k_B T} - 1)}$$

1.6 EINSTEIN'S LIGHT QUANTUM

Einstein, in 1905, in his revolutionary paper titled *On Heuristic Point of View about the Creation and Conversion of Light* goes far beyond Planck. It is a reinterpretation of Planck's work. Einstein makes a bold hypothesis that electromagnetic field itself is a collection of 'energy quanta localized in space which move without being divided and which can be absorbed or emitted as a whole'. Light quanta are very much similar to indivisible atoms. Einstein used the word light quanta. It is in 1926, Lewis coined the word *photon*. It is worth quoting Einstein.

> 'The wave theory of light which operates with continuous function in space has been excellently justified for representation of purely optical phenomena, and it is unlikely ever to be replaced by another theory. One should, however, bear in mind that optical observations refer to the time average and not to the instantaneous values, and notwithstanding the explicit experimental verification of theory of diffraction, reflection, refraction, dispersion, and so on, it is quite conceivable that a theory of light involving the use of continuous function in space will lead to contradiction with experience, if it is applied to phenomena of creation and conversion of light. In fact, it seems to one that the observation on black body radiation, photoluminescence, the production of cathode rays by ultraviolet light can be better understood on the assumption that the energy of light is distributed discontinuously in space. According to the assumption considered here, when a light ray starting from a point is propagated, the energy is not continuously distributed over an ever-increasing volume, but it consists of finite number of energy quanta, localized in space, which move without being divided and which can be absorbed or emitted as a whole'.

This paper is mainly devoted to the question of nature of light. Only in the end, he derives the equation for photoelectric as evidence to this theory of light. He did not reject wave theory of light. It was not a theory going back to Newton's corpuscles of light. Nevertheless, it has an element of visualizing light similar to indivisible atoms. The majority of scientific community did not subscribe to this view of a light at that time. The problem is how to reconcile Einstein's theory with Maxwell's electromagnetic wave theory which has worked very successfully. So, though Einstein's credentials as an outstanding physicist were recognized, many scientists considered this energy of light quantum as a rare aberration on the part of Einstein.

In 1913, Planck Nernst, Rubens and Warburg wrote a recommendation letter proposing Einstein for membership in Prussian Academy and research professorship. While appreciating Einstein for his innovations in other fields, they considered his theory of light quanta as a rare aberration on his part. They lavishly praised Einstein for his contributions to physics but made the following remarks: 'That he may sometimes have missed the target in his speculations as, for example, in his hypothesis of light quanta, cannot be held against him'.

In 1916, Millikan described Einstein's light quantum hypothesis as follows: 'This hypothesis may be called reckless, first because an electromagnetic disturbance which remains localized in space seems to be a violation of very conception of an electromagnetic disturbance, and second it flies in the face of the thoroughly established facts of interference'.

In 1921, Einstein was awarded Nobel Prize for his explanation of photoelectric effect. In 1923, Millikan was awarded Nobel Prize for verifying Einstein's photoelectric equation. Millikan writes as follows: 'I spent ten years of my life testing that 1905 equation of Einstein and contrary to my expectations, I was compelled in 1915, to assert its unambiguous experimental verification in spite of its unreasonableness since it seemed to violate everything that we know about the interference of light.'

Details of photoelectric effect and Einstein's equations are not presented here as they can be found in any introductory text book on modern physics. In 1909, he published the paper on the fluctuations of black body radiation, in which he lays foundation for wave–particle duality. In 1917, he introduced the idea that a light quantum has linear momentum $h\nu/c$.

One of the most significant experiments in establishing the particle property of light is Compton effect. In Compton effect, there is a shift in wavelength of the x-ray when it is scattered by an electron. A simple analysis treats this process as a collision between electron and photon similar to the analysis of collision between billiard balls. Conservation of energy and momentum are used in a way just as we use them in classical mechanics. We have to calculate the total momentum and energy before impact and after impact. The momentum of the photon is taken to be $h\nu/c$. This simple analysis worked very well. After the discovery of Compton effect, the particle property of light was accepted.

1.7 DUAL NATURE OF LIGHT: WAVE–PARTICLE DUALITY

All these experiments show that the nature of light is more complex than it was thought before 1900. From 1800 to 1900, the wave nature of light was established firmly. However, from 1900 onwards, particle aspect of light emerged. They are not the old Newtonian corpuscles. By now, it was established that there were phenomena like interference and diffraction which could be explained only in terms of wave property of light. There are other phenomena like photoelectric effect, Compton effect which can be explained only with particle property of light. One could not settle for either property at the cost of other. This is termed as dual nature of light or wave–particle duality.

1.8 MATTER WAVES

In 1924, de Broglie came with an interesting proposal. He extended the dual nature of light to matter like electrons also. From chemistry and kinetic theory of matter, we get the picture that an atom is a miniature version of a classical tennis ball. With the discovery of electron and nucleons, the indivisible nature of atom is lost. Still, one believes that they will behave like particles. But de Broglie suggested that the matter also behaves like waves. This conclusion is based on philosophical consideration, namely nature loves symmetry. If light exhibits dual nature, why would not other particles in atomic world also exhibit dual nature? He suggested that the wavelength λ of the matter wave is given by

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

The significance of dual nature of atoms or electrons lies in the fact they are also not as simple as they were thought to be before 1900. So, some experiments can be explained only by assuming matter like electrons are waves and some experiments can be explained by assuming that they are particles.

In the case of light, there is no single model of light to describe all the phenomena involving light. In the same way, in the case of electrons also (and for all matter particles), there is no single model that can explain all the phenomena. The wave property of electron was experimentally verified soon. We will discuss the matter waves and the wave–particle duality again at the end of this chapter.

1.9 VECTOR ATOM MODEL

The discussion on the development of atomic structure departing from classical physics has to start with the discovery of electron. Rutherford's α-particle scattering is the key experiment to bring the inadequacies of classical physics in the case of atoms. The experimental results of Rutherford can be explained only in terms of planetary model of an atom. The electrons are point particles revolving around the nucleus. However, this model violates classical physics. The next major development was Bohr model which is based on three basic assumptions:

1. The electron moves in a circular orbit with nucleus at the centre.
2. The orbits of the electrons are restricted to a group of orbits for which the angular momentum of the electron is $l = nh/2\pi$.
3. Radiation is emitted (or absorbed) only when the electron jumps from one orbit to another orbit. The frequency of the emitted radiation is

$$E = h\nu = E_n - E_m$$

This model is very successful in explaining the line spectra of hydrogen atom. The significance of this model lies in the fact that its success is achieved at the cost of classical laws of physics. In classical physics, if an electron moves in a circular orbit, it has to lose energy by continuously emitting radiation making the circular orbits untenable for electrons. However, Bohr simply asserts that the electron does not respect this law. The restriction on the orbits is not established from classical physics. Apart from these violations, all the other calculations are done using classical physics. So, it is some sort of tinkering work with respect to classical physics.

However, this model subsequently developed into vector model, which was quite successful in explaining a number of features of atomic structure. The existence of four quantum numbers, have been established. They are the principal quantum number n, orbital angular momentum quantum number l, azimuthal quantum number m_l, and spin angular momentum quantum number m_s. Vector atom model neatly explains periodic table.

One of the principles which worked well is Bohr's principle of correspondence which establishes the connections between quantum physics and classical physics. In the large limit of a quantum number it is expected that quantum physics will give the same result as that of classical physics. Still, all these models are ad hoc in nature; there is no systematic account to describe all the phenomena at atomic level. It was realized that totally a new mechanics has to be developed instead of tinkering classical physics to describe the atomic structure.

1.10 BIRTH OF QUANTUM MECHANICS

Matrix Mechanics

In 1925, Heisenberg came with a new quantum mechanics. It is now called matrix mechanics. His reasoning was roughly as follows: We cannot determine the position or momentum of electron in an atom exactly. So, the classical description of a particle in atomic world is not possible. Therefore, let us try to formulate a new theory in terms of only observable quantities. So, corresponding to position $x(t)$ in classical physics, Heisenberg produced a new quantity X_{mn} where m and n refer to atomic energy levels E_m and $E_{n.}$ So, X_{mn} is a set of array of numbers which takes into account all the states. Similarly, he replaced classical momentum by a set of array of numbers P_{mn}.

Later this exercise was recognized as associating matrices with the dynamical variables $x(t)$ and $p(t)$. i.e. $x(t) \rightarrow X(t)$ and $p(t) \rightarrow P(t)$ where $X(t)$ and $P(t)$ are matrices. Since X and P are matrices, $XP \neq PX$. In fact, it was shown that $XP - PX = i\hbar$. Since the dynamical variables are replaced by matrices, this mechanics is called matrix mechanics.

Wave Mechanics

In 1926, Schrödinger came with wave mechanics. Unlike Heisenberg's matrix mechanics, Schrödinger's wave mechanics was based on differential equations, a familiar terrain. Both Schrödinger equation and Heisenberg's matrix mechanics gave identical results when applied to problems like harmonic oscillator and hydrogen atom. Yet, they looked very different. Very soon, Schrödinger established the equivalence between wave mechanics and matrix mechanics.

Dirac came up with further clarification. He introduced the concept of abstract operators (q – numbers) and abstract state vectors and showed that matrix mechanics and wave mechanics are special cases of quantum mechanics based on abstract operators and abstract state vectors. The wave mechanics, matrix mechanics, and Dirac's abstract operator method form the main theme of this book.

1.11 HEISENBERG'S UNCERTAINTY PRINCIPLE

In 1927, Heisenberg came up with the uncertainty relation to bring out the limitations of the classical concepts. In particular, Heisenberg argued that the process of observation itself disturbs the system; in other words we cannot measure any quantity without disturbing the system in the process. This led him to the uncertainty relations $\Delta x \Delta p \sim h$ which implies that it is impossible to determine the position and the momentum of an electron simultaneously with any arbitrary precision. This has nothing to do with the limitations of the instrument.

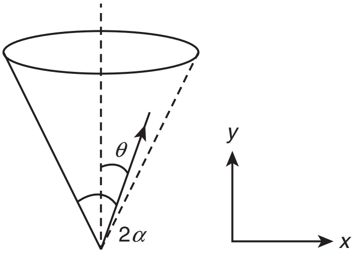


Fig. 1.5 γ-ray microscope

Heisenberg illustrated how the uncertainty relation for position and momentum arises using a hypothetical γ-ray microscope which is schematically shown in Fig. 1.5. The lens and the electron together form a cone of angle $2a$. Let us assume that the electron is at the focus. The γ-ray illuminates the electron and gets scattered into the lens.

From physical optics we know the resolving power of this microscope is

$$\Delta x = \frac{\lambda}{2\sin\alpha}$$

Our eyes do not react to γ-ray and the visible light in the same way. If our eyes do respond in the same way, then we will see a small spot of width Δx and conclude that the electron will be somewhere within this spot. The wave nature of light introduces this uncertainty Δx in the measurement of the position of the electron.

$$\Delta x = \frac{\lambda}{2\sin\alpha}$$

Now, let us study the same phenomenon using the particle property of light to determine the momentum of electron. The γ-ray has a momentum $h\nu/c$ or h/λ. From conservation of momentum, the sum of momenta of γ-ray and the electron are constant. The total momentum along x-axis is

$$P = P_\gamma + P_e$$

If we know P_γ exactly, we can determine P_e exactly. If we know P_γ only approximately, the momentum P_e can also be determined only approximately. If Δp_γ and Δp_e are the uncertainties in the momenta of the photon and electron respectively, then we have

$$\Delta p_e + \Delta p_r = 0$$

$$\text{or} \quad \Delta p_e \sim \Delta p_\gamma .$$

Let us determine the momentum of the photon atom along x-axis. Let the scattered photon enter the microscope at an angle θ as shown in Fig. 1.5. Unfortunately, when a photon is detected in the microscope, all that we can say is that it is a photon scattered at some angle θ between $-\alpha$ and α. Therefore, x-component of the photon momentum is some value between $-h/\lambda \sin\alpha$ and $h/\lambda \sin\alpha$.

$$\therefore \Delta p_\gamma = \frac{2h}{\lambda}\sin\alpha = \Delta p_e$$

$$\Delta x = \frac{\lambda}{2\sin\alpha}$$

$$\Delta x \Delta p_e \sim h$$

Within quantum mechanics, the uncertainty Δx and the uncertainty Δp are defined in a different way. So, their interpretations differ from the above formulation of uncertainties. The uncertainty relations run through the quantum description of all phenomena.

There are other uncertainty relations also. The time–energy uncertainty relation is

$$\Delta E \Delta t \sim h$$

The interpretation of this relation is different from that of position and momentum. Here, Δt is the time taken to measure the energy of the system and ΔE is the resulting uncertainty in energy. The experimental result of energy can be any number from $E - \Delta E/2$ to $E + \Delta E/2$.

The uncertainty principle has serious consequences. One of the consequences is ruling out the exact trajectory of a particle in the sense of classical physics. The idea of exact trajectory requires the knowledge of exact values of position and momentum at the same instant. If these quantities cannot be known without uncertainties there is no way of knowing the trajectory of the particle. So, there is no point in visualizing the orbit of an electron in an atom.

The second important consequence of the uncertainty principle is that it let Bohr to postulate the complementarity principle which states that it is impossible to observe the particle and wave properties simultaneously. The complementarity principle asserts that depending on the design of the experiment, we can observe either wave property or particle property but not both. This is in a sense a veto statement. However, violations of this principle will not contradict the basic principles of quantum mechanics. The current experiments are consistent with principle of complementarity and with the advent of modern technology experiments are being done to test the complementarity principle.

Example 1.1 Determine the de Broglie wavelength of neutron which moves with a velocity 210 m/sec.

Solution:

$$\lambda_{dB} = \frac{\lambda}{p} = \frac{\lambda}{mv} = \frac{6.62\times10^{-34}\,J.\text{sec}}{1.67\times10^{-27}\,\text{kg}.210\,\text{m.s}^{-1}} = 18.8\text{Å}$$

Example 1.2 Consider a double-slit interference experiment for neutrons of de Broglie wavelength λ_{dB}= 18 Å. The width of each slit is 20 μm and they are separated by 100 μm. Determine the spacing between the minima of the intensity in the interference pattern. Give a sketch of the interference pattern. Assume that the screen is at 5m away from the double slit.

Solution:

The width of the slit $a = 20\ \mu\text{m}$
Distance between the slits $d = 100\ \mu\text{m}$
Distance of the screen from the slits $d = 5\ \text{m}$
De Broglie wavelength $\lambda_{dB} = 18\ \text{Å}$
The intensity $I = I_0 \cos^2\beta \left(\frac{\sin\alpha}{\alpha}\right)^2$

Here, $\alpha = \frac{\pi a}{\lambda} \cdot \frac{y}{D}$ and $\beta = \frac{\pi d}{\lambda} \cdot \frac{y}{D}$

$$\alpha = \frac{\pi \times 20 \times 10^{-6}}{18 \times 10^{-10} \times 5} y = \frac{\pi \times 20 \times 10^{-12} \times (y' \mu\text{m})}{18 \times 10^{-10} \times 5} = 0.007 y'$$

$$\beta = \frac{\pi \times 100 \times 10^{-6} \times 10^{-6} (y' \text{ in } \mu\text{m})}{18 \times 10^{-10} \times 5} = 0.03 y'$$

The spacing between the minima is given by $\Delta y = 104.04\ \mu\text{m}$.

The sketch of the intensity distribution is given in Fig. 1.6.

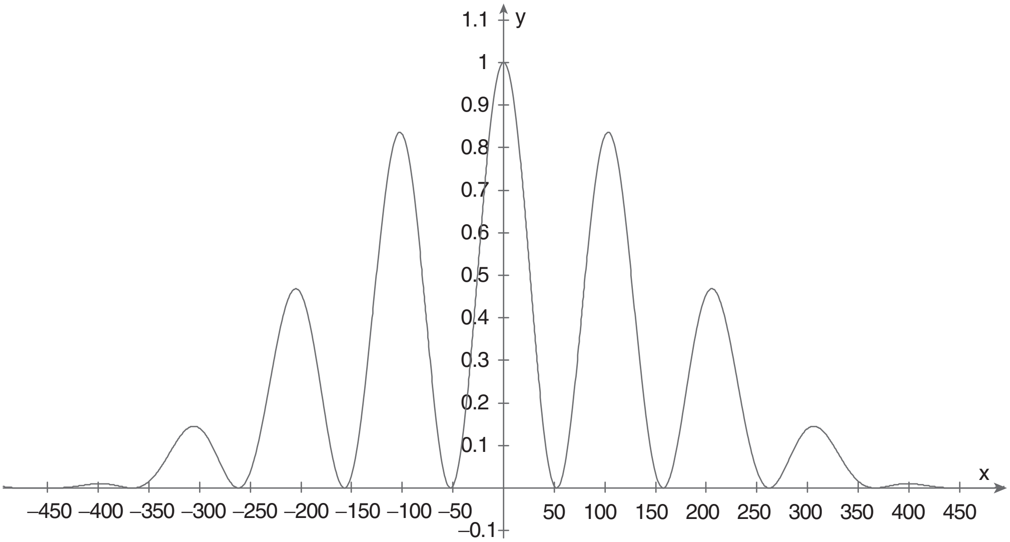


Fig. 1.6 Intensity distributions for double-slit experiments

Example 1.3 A 60 watt bulb emits green light of wavelength $\lambda = 5460$ Å. How many numbers of photons are emitted per second?

Solution: The number of photon emitted per sec is

$$= \frac{\text{power}P.}{h\nu} = \frac{P.\lambda}{\lambda c}$$

$$= \frac{60 \times 5460 \times 10^{-10}}{6.62 \times 16^{-34} \times 3 \times 10^{8}} = 16.5 \times 10^{19}\ \text{photons/sec}$$

Example 1.4 In a double-slit experiment, the distance between the source of light and the screen is 3m. Assuming that the source emits electrons at a constant rate, what is the maximum intensity possible so that only one electron is present at a time?

Solution: Let L be the length between the source and the screen. Light takes a time τ to travel this distance. For only one photon to be present in the apparatus at a time, the time between the emission of two consecutive photons should be greater than or equal to τ. So, the number of photons per second corresponding to this emission interval is

$$N = \frac{1}{\tau} = \frac{c}{L} = \frac{3\times10^8}{3} = 10^8 \text{ photons/sec}$$

1.12 MATTER WAVE: EXPERIMENTAL EVIDENCES

Wave property of light has been established very well between 1800 and 1900. It requires special efforts to observe the particle property of light. The reverse is true in the case of matter. The particle aspect of matter is well established. But it is very difficult to observe wave property of matter. The main problem is that the de Broglie wavelength is very small. It is of the order of few Å or much less than that. It is very difficult to do a single-slit or a double-slit experiment for waves with such a short wavelength. The slits widths should be very small. The preparation of such slits and the determination of the size of the slits are themselves challenging. The earliest experiment to establish wave nature of the electron is Davisson–Germer experiment in 1927. The first double-slit experiment for electrons was performed by Jonsson in 1961. Subsequently double-slit (and single-slit) experiments have been performed for neutron (Broglie wavelength $\lambda_B \simeq 18$ Å), helium atom ($\lambda_B = 0.56$ Å 1.03 Å), carbon-60 ($\lambda_{dB} = 2.5$ pm). Poisson spot has been obtained for deuterium molecules ($\lambda_{dB} = 0.96$ Å). We present on outline of some of these experiments.

In Davisson Germer experiment electron beam incident on a nicked crystal is diffracted in a way similar to x-ray diffraction by a crystal. The experiment is schematically shown in Fig. 1.7. The detector D is free to move around, and the intensity of the scattered waves in different directions can be measured. For a voltage of 54 V, there is a sharp increase in the number of electrons collected by the detectors at an angle 50º. We can use Bragg's condition $d \sin \theta = \lambda$. For a nickel crystal $d = 2.15$ Å, $\theta = 50°$. Therefore, we have

$$\lambda = 2.15 \sin 50° = 1.65 \text{ Å}$$

The de Broglie wavelength λ_{dB} can also be calculated $\lambda_{dB} = h/p = \dfrac{h}{(2meV)^{1/2}}$

For 54 V, we get λ_{dB}=1.67°. This confirms the wave nature of electron.

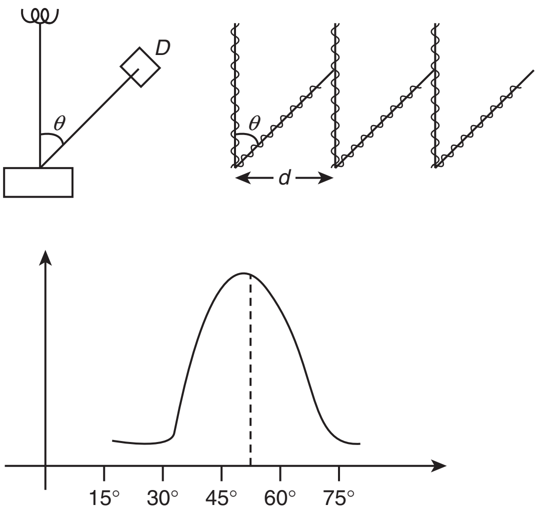


Fig. 1.7 Davisson–Germer experiment

Claus Jönsson's Experiment (1961)

The double-slit experiment for the electrons was performed first time by Claus Jönsson. He got the images of the interference pattern for a multiple slits from a single-slit to five slits. He used 50 keV electrons for which $\lambda_{dB} = 0.05$ Å. The slit width is 0.5 μ and the spacing between the slits are 2.0 μ. This work was *reported in a German journal in 1961 and its translation appeared in American Journal of Physics* in 1974. The interference pattern for single and double slits is shown in Fig. 1.8.

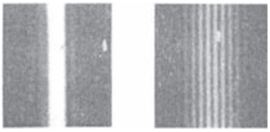
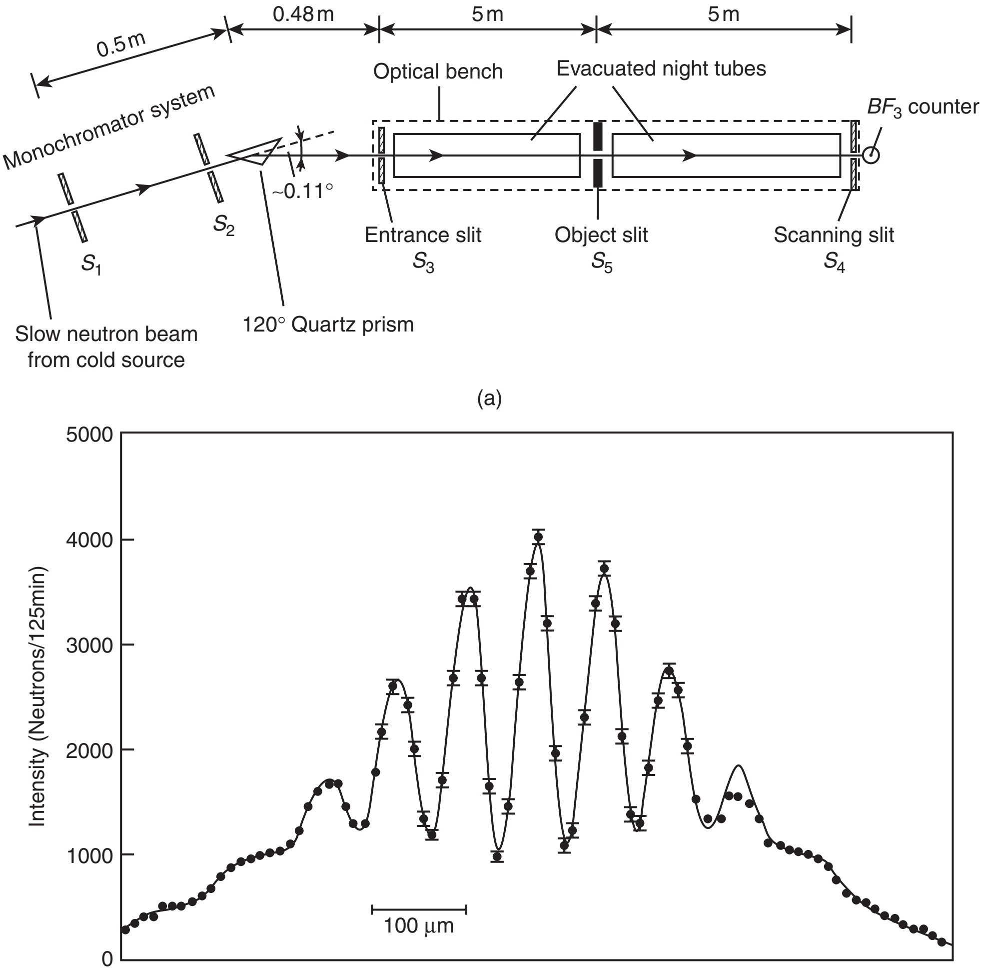


Fig. 1.8 Interference pattern for electrons

Wave Nature of Neutron (1989)

Anton Zeilinger, Roland Gahler, C.A. Shull, Wolfgong Theimer and Walter Mampe reported the diffraction of neutrons by a single and double-slit arrangements. In the single-slit experiment, a slit of width 90 μm was used. The de Broglie wavelength λ_{dB} is 19.26 Å (19.26 ± 0.70 ± 0.02). The detectors were placed 5 m from the single slit. They have also reported for a single slit of width 23 mm. In addition to these experiments, they obtained the double-slit interference pattern for neutrons of wavelength λ_{dB} = 18.45 Å (λ_{dB}= 18.45 ± 1.40 ±.02). The dimensions of the double-slit arrangement is 21.9 – 104.1 – 22.5 mm (left slit-boron wire-right slit). Their experiments are shown in Fig. 1.9.



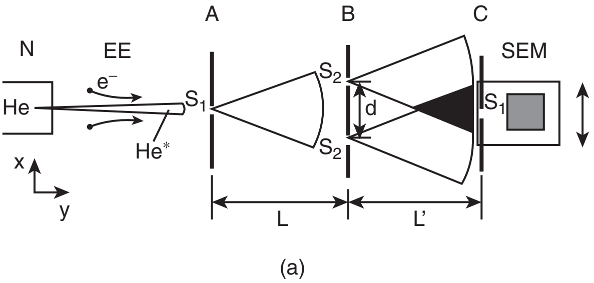
(a)

(b)

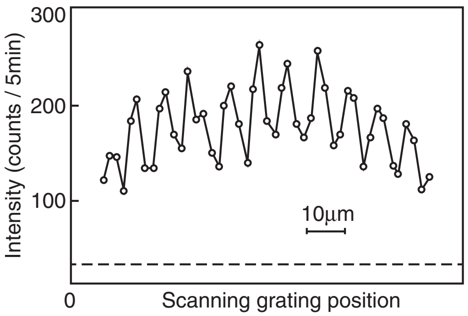
Fig. 1.9 Single-slit and double-slit experiment for neutrons

Wave Nature of Helium (1991)

Double-slit experiment for helium atoms is shown in Fig. 1.10.



(a)



(b)

Fig. 1.10 Double-slit experiment for helium atoms

O. Carnet and J. Mlynek in 1991 performed a double-slit experiment using metastable helium atoms. They allowed helium atom from an oven to pass through 2 μm wide slit in a thin gold foil. The beam from this slit was allowed to incident on a double-slit arrangement. Two slits width 2 μm separated by 8 μm were used to get interference pattern for helium atoms. They did the experiments for helium atoms with different wavelengths. The experimental details are as follows:

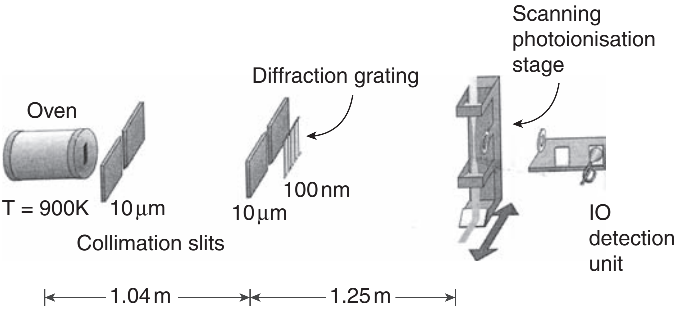
Oven temperature: 295 K de Broglie wavelength $\lambda_{dB} = 0.56$Å

Oven temperature: 83 K de Boglie wavelength $\lambda_{dB} = 1.03$Å

For helium of atoms with $\lambda_{dB} = 0.56$Å, the intensity of atom varies periodically with a period of $\Delta x = 4.5 \pm 0.16$ μm (Fig. 1.10b). The screen is at a distance $L = 64$ cm from the double slits.

Diffraction of C-60 (1999)

Markus Arndt, Lucia Hackermuller and Klaus Hornberger observed the wave nature of buckminsterfullerene C-60 molecules. This molecule is big in size and mass when compared to electron and proton. C-60 is a molecule consisting of 60 carbon atoms and their arrangement will make C-60 look like a football. Its mass is 720 atomic mass units, and its size is 1 nm. The carbon-60 was produced in an oven at 900 K as shown in Fig. 1.11.



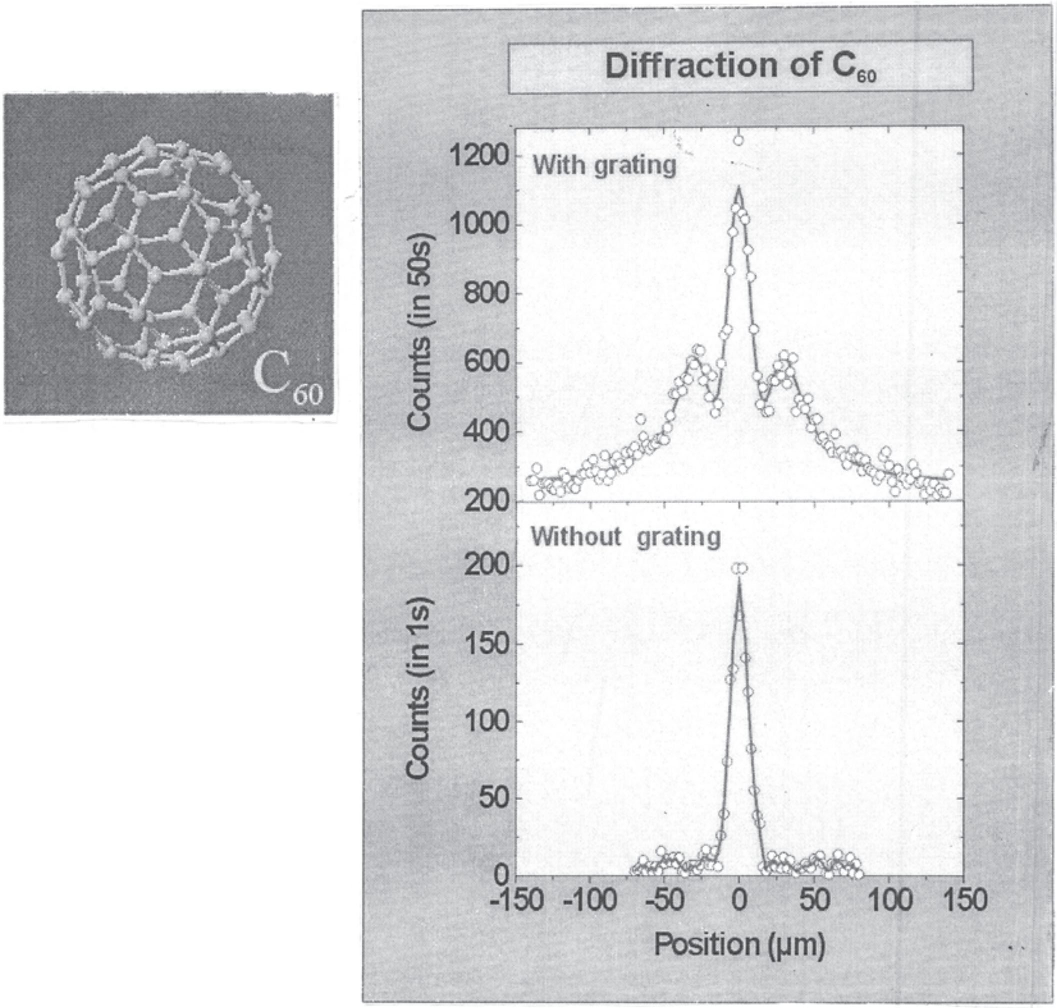
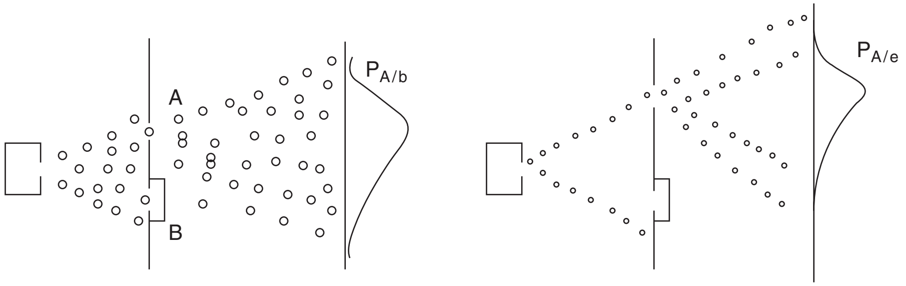


Fig. 1.11 Diffraction of C-60 molecule

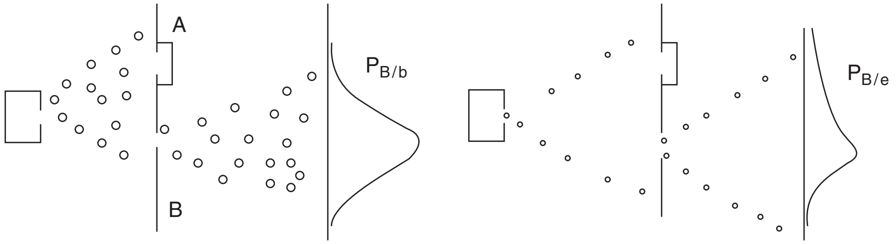
The most probable velocity is 220 m/sec, and the corresponding de Broglie wavelength $\lambda_{dB} = 2.5$ pm (2.5×10^{-12} m). Note that λ_{dB} is 400 times smaller than the size of C-60 molecule. They were allowed to pass through a grating whose grating element is 100 nm and the slit width is 50 nm. The central maximum and the first-order diffraction peaks were observed.

1.13 FEYNMAN'S DOUBLE-SLIT EXPERIMENT: EXPOSITION OF NON-CLASSICAL NATURE OF MICRO WORLD

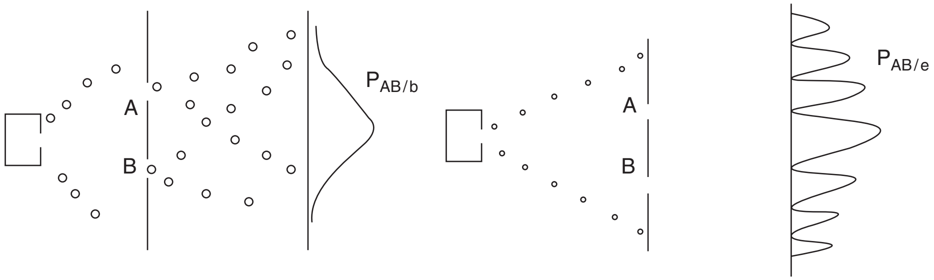
What is the nature of a photon? What is the nature of an electron? It is very difficult to answer such questions for any quantum object. Knowledge about these particles is continuously changing. Without going into more details, we bring the non-classical nature of quantum objects using hypothetical double-slit experiment. We use bullets and electrons as the incident particles as shown in Fig. 1.12. The main purpose of this hypothetical experiment is to highlight some features of quantum mechanics. Only recently experiments have been done using single electron and single photon sources which confirm the quantum mechanical description of the objects like electron or photon.



(i) Experiment # 1



(ii) Experiment # 2



(iii) Experiment # 3

Fig. 1.12 Hypothetical double-slit experiment for bullets and electrons

To compare and contrast the classical physics and quantum physics we describe a set of double-slit experiments using bullets and electrons as the test particles. The bullet represents a classical particle. Let us assume they are indivisible. They have definite volume and mass. These features are described by the term *localized lumps*. When they hit the screen, they hit as localized lumps occupying small volume on the screen.

Experiment 1

There are two slits A and B. They can be closed or kept open. In this experiment, slit A is open and slit B is closed. The bullets are fired in a random way. Correspondingly, the bullets are scattered from the slit A in a random directions. We can count the number of bullets detected at different points (locations) on the screen. Let n_T be the total number of bullets that hit the screen. $n(x)$ is the number of bullets that hit the screen at point x. Then, $P_{A/b}(x) = n(x)/n_T$ gives the probability of a bullet from slit A reaching the screen at x. $P_{A/b}(x)$ is shown in Fig. 1.12.

Now let us conduct a similar experiment with electrons. The moment we say electrons, the immediate picture that comes to our mind is the miniature version of the indivisible bullet. Though the electrons are not visible to us, we imagine that they will be like small (almost point-like) bullets. This follows from the intuitive understanding of an electron. Let $P_{A/e}$ be the probability of finding the electron from the slit A at x on the screen. $P_{A/e}$ is very much similar to $P_{A/b}$.

Experiment 2

The second experiment is literally the same, except one change. Slit A is now closed and slit B is open. The probability distributions $P_{B/b}$ and $P_{B/e}$ for the bullets and electrons look alike.

These two experiments justify our picture of electron, namely they are point-like objects looking like mini bullets.

Experiment 3

The results of the experiment 3 are mysterious. Now both slits A and B are open. The behaviour of bullet agrees with commonsense. Since the bullets are very small, they cannot go through both slits simultaneously. They have to go through either slit A or B. The two possibilities (or alternative as Feynman calls them) are mutually exclusive. Therefore, the probability distribution for the bullets is

$$P_{AB/b} = P_{A/b} + P_{B/b}.$$

Now in the case of electrons, the results are entirely different. This resembles the intensity distribution of water waves. In fact, this distribution is independent of the nature of the wave. The distribution is $I_0^2 \cos^2 \beta[\sin\alpha/\alpha]^2$ where $\alpha = \pi a/\lambda \sin\theta$ and $\beta = \pi d/\lambda \sin\theta$ which can be obtained by superposing waves from slit A and slit B irrespective of the nature of waves (i.e. the waves may be sound waves or water waves or light waves or any other kind of wave). So, we conclude that the electrons behave like waves.

$$P_{AB/e} \neq P_{A/e} + P_{B/e}$$

1.13.1 Non-classical Particle, Non-classical Wave

These three experiments taken together lead to contradictory conclusions about electrons. Judging by the first two experiments, we conclude that an electron is a point particle and judging by the third experiment alone we conclude that an electron is an extended object covering both slits. We don’t come across a macroscopic object which behaves in this way like electron.

Consider the third experiment again. When an electron is detected, it is detected as a point-like particle. We don't detect half of an electron or any fraction of an electron. The detected electrons on the screen form a set of discrete points, or they are always detected as a whole localized lump (full clicks as Feynman describes). So in this respect, they are similar to bullets; but their intensity distribution is similar to waves. So, all that we can conclude is that electrons behave like particles in some respect, but they are not classical particles. In the same way, though they behave like waves to produce interference pattern, this wave is a non-classical wave since the intensity of a classical wave distribution cannot be seen as a collection of localized lumps.

"Which Path" Information and Interference

Let us further investigate the experiment 3 again to bring out the non-classical nature of electron. The basic assumption behind the probability distribution $P_{AB/b}$ for bullets is that a bullet goes through, either slit *A* or slit *B* (but not through both). Do electrons also go through either slit alone? Intuitively, the answer is yes. Since electrons are assumed to be small indivisible objects they cannot go through both the slits. Let us not trust our intuition. Let us decide to test which electron goes through slit *A* and which electron goes through slit *B* before they reach the screen to produce the interference pattern. In essence, we are attempting to determine the trajectory of the electrons. Is the trajectory of electron *S-A-P* (source *S*-slit *A*-point *P* on the screen) or *S-B-P* (source *S*-slit *B*-point *P* on the screen). To answer this question, all that we have to do is to devise an experiment, to watch the motion of electrons. Feynman in his book provides an elaborate discussion on such imaginary experiments. Nothing prevents us from watching the electrons to determine the slit through which it emerges. But, we can do it only at the cost of interference. Using uncertainty principle (see Feynman) it can be shown that interference is now lost. The act of measurement produces a disturbance that changes the result, and hence, interference is lost. The whole exercise leads to an interesting conclusion. There are two alternatives available in this experiment for an electron: going through slit *A* (path *S-A-P*) or going through slit *B* (path *S-B-P*). If only one alternative is allowed, the interference is lost. If both alternatives are available to the electron, interference takes place. In other words, if we know the path of the electrons, the interference is lost. If we don't know the path of the electron, then we have interference.

Knowledge of the path of the electron - No interference pattern
No knowledge of the path of the electron - Interference

With the development of new technology, experiments have been done to test these conclusions. All the experiments confirm this conclusion.

One Electron After Another

Let us now turn to another class of experiments, which brings the non-classical character of a particle in atomic world again.

How do the electrons produce interference pattern? Do the electrons going through slit *A* and the electrons through the slit *B* combine to produce the interference pattern? To answer this question, let us decrease the intensity of the incident beam of electrons. Whatever be the intensity of the incident beam, an electron always exists as a single whole entity (the electron is indivisible). Reduce the intensity to such an extent that we have only one electron at a time in the whole apparatus. So, let us send one electron after another, let us visualize the following scenario. An electron from the source enters the double-slit arrangement and then it reaches the screen. Till then, the second electron is not produced. One can wait for some more time, and then another electron is produced. This procedure is repeated many times. Remember, this is a hypothetical experiment to illustrate the workings of quantum mechanics which leads to the following results. Initially, the electrons will hit the screen in a random fashion. In a strange way, the electrons will not hit the regions of dark fringes. Electrons prefer to hit the regions corresponding to bright fringes. If we wait for a long time, a large number of electrons (through they hit the screen one by one) would have reached the screen. The surprising thing is that the interference fringes are slowly emerging. This is the familiar interference pattern. For electrons from slit *A* to combine with the electrons from slit *B* to produce interference, we should have at least two electrons at a time in the apparatus. But, when the intensity of incident beam of electrons is so weak that only one electron is present at a time in the whole apparatus, such a possibility is ruled out. So, what we observe is that a single electron goes through both slits (in a mysterious way beyond our comprehension) and interferes with itself. So, the interference fringes in the double-slit experiment are not due to interference between two sets of electrons, one set through slit *A* and another set through slit *B*. It is a consequence of each electron interfering with itself.

Real Low-intensity Experiments

Let us now discuss the actual experiments with low-intensity photons and low-intensity electrons. In 1909, Taylor did the experiment, producing diffraction by needle using very low-intensity light. He first determined the diffraction pattern produced by a needle produced by an intense beam of light. Then he decreased the intensity of light. To get a good image in the film, exposure time increases as the intensity of light decreases. He reduced the intensity of light to a level which he estimated that only one quantum of light can be present at a time. He exposed the film for 2000 hours and got the same diffraction pattern obtained for intense beam of light. So, we get the same diffraction pattern whether we send large number of photons at a time or one by one for a very long time.

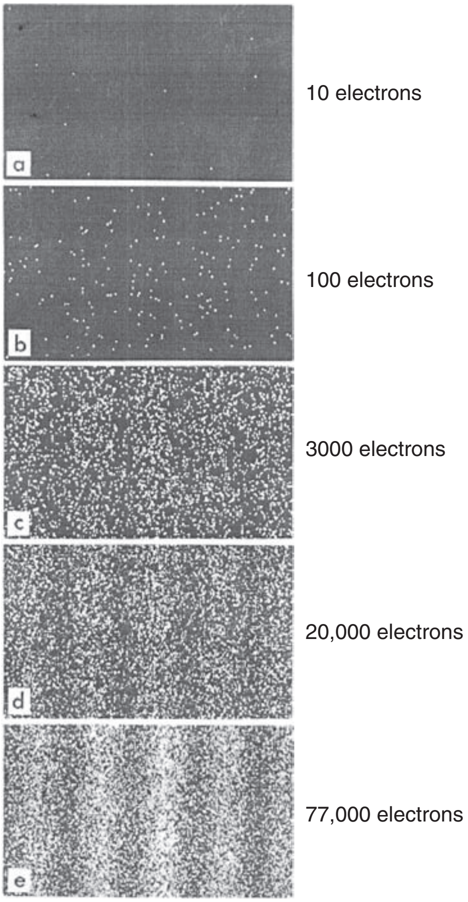


Fig. 1.13 Buildup interference pattern for electrons

In 1974, a research team led by Pier Giorgio Merli performed an experiment with very low-intensity electrons. Akira Tonomura and his team at Hitachi, Japan also performed a double-slit experiment with very low-intensity electron beam such that only one electron is present at a time in the whole apparatus. Without going into the details of the experiment, their results are presented in Fig. 1.13. One can see the randomness in the positions of the electrons initially in as shown in Figs 1.13 (a) and (b). Only in Figs 1.13 (c), (d) and (e) the interference pattern is seen to emerge clearly.

1.13.2 Explanation of Double-slit Experiment by Wave Mechanics

In wave mechanics, the wave function $\psi(\mathbf{r},t)$ describes the state of the system. If one asks 'what is the position of an electron at time?' then the answer from quantum mechanics is very disappointing. There is no way of specifying the position of a particle as done in

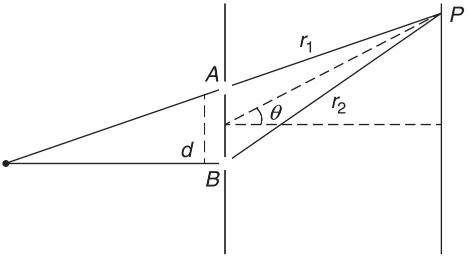


Fig. 1.14 Double-slit experiment

classical mechanics. All that we can tell is the probability of finding the particle at a particular position **r** at the instant *t*. The probability of finding the particle in a volume $d^3\mathbf{r}$ is given by $\psi^*(\mathbf{r},t)\ \psi(\mathbf{r},t)d^3\mathbf{r}$ (This description is true only in non-relativistic quantum mechanics. The position probability is not well defined in relativistic quantum field theory. It is very difficult to define position probability density for an electron or a photon in quantum field theory. We restrict ourselves to non-relativistic quantum mechanics.)

The wave functions ψ_A and ψ_B describe the waves describing the electrons from slit *A* and slit *B*, respectively as shown in Fig. 1.14.

$$\psi_A \simeq f(\theta)e^{i(kr_1-\omega t)}$$

$$\psi_B \simeq f(\theta)e^{i(kr_2-\omega t)}$$

The wave function for an electron on the screen is

$$\begin{aligned}\psi &= \psi_{A\,+}\psi_B \\ &= f(\theta)(e^{i(kr_1-\omega t)}+e^{i(kr_2-\omega t)}) \\ &= f(\theta)e^{i(kr_1-\omega t)}(1+e^{ik(r_2-r_1)}) \\ |\psi|^2 &= 2|f(\theta)|^2\cos^2\left(\frac{\pi d}{\lambda}\sin\theta\right)\end{aligned}$$

This is the expression for classical wave. Now $|\psi|^2$ is interpreted as the probability density for the electron to hit the screen at a pointer **r**. So, the observed intensity distribution of the electrons is consistent with the quantum mechanics. The positions at which the individual electrons hit the screen are random. If $|\psi|^2\,d^3\mathbf{r}$ is zero, it is impossible to find the electrons at these points. These regions correspond to destructive interference. If $|\psi|^2\,d^3\mathbf{r}$ is 1, they correspond to definite events. For other values of probability, the electron may or may not hit at that point. When the number of electrons become large, a definite pattern will emerge.

1.13.3 Superposition: Consequences Unique to Quantum Mechanics

The wave function $\psi = \psi_A + \psi_B$. is superposition of two states ψ_A and ψ_B. Such superpositions are there also in classical wave theory. But, in quantum mechanics, superposition has non-trivial consequences. A simple example will help us to understand the consequences of superposition principle. Let us consider an object which can spin either clockwise or anticlockwise about *z*-axis. Let us assume that no other rotation is possible. The corresponding state functions are given by ψ_{clock} and $\psi_{\text{anticlock.}}$

$$\psi_{\text{clock}} \rightarrow \text{object spinning in clockwise sense}$$

$$\psi_{\text{anticlock}} \rightarrow \text{object spinning in anticlockwise sense}$$

We are familiar with macroscopic objects spinning either clockwise or anticlockwise. So ψ_{clock} and $\psi_{\text{anticlock}}$ represent states which are familiar to us. But, in quantum mechanics, we have one more kind of state.

$$\psi = \psi_{\text{clock}} + \psi_{\text{anticlock}}$$

What is the spin of the object in this state? Does it spin clockwise or anticlockwise? (No other rotation is also permitted.) Quantum mechanics does not provide any answer to this question. (If one wants to dramatize this idea, consider a hypothetical example. Suppose the shape of an object is either square or circular. No other shape is possible. ψ_{square} describes the object in square shape; ψ_{circular} describes the object of the circular shape. Suppose the state of the object is $\psi = \psi_{\text{square}} + \psi_{\text{circular}}$, what is the shape of this object? It is neither square nor circular; other shapes are not allowed. This state looks more mysterious when other possibilities which can be visualized by classical physics are ruled out.)

Now let us come back to the electron in the double-slit experiment. ψ_A represents an electron which goes through slit A. ψ_B represents an electron which goes through slit B. These are similar to our experience with classical objects. We can comprehend these two possibilities for electron though we may not 'see' them. However, what possibility does $\psi = \psi_A + \psi_B$ represent? There is no classical analogue corresponding to this state, allowing us to comprehend any possibilities which are within our experience. We have to proceed with this state as such without raising the question about the path of the electron. However, we can compute the probability of finding the electron at different positions on the screen.

What about the observation process to find out whether the electron goes through slit A? Suppose one finds that the electron has gone through A. This measurement process makes the wave function to collapse from $\psi = \psi_A + \psi_B$ to ψ_A.
That is, detection of electron through slit A: $\psi = \psi_A + \psi_B \xrightarrow{\text{Collapse}} \psi_{A.}$

The corresponding probability density function becomes $|\psi_A|^2$.

$$\text{(i.e.)} \quad P_{\text{A/e}} = |\psi_A|^2$$

The interference is lost.

1.14 PARTICLES AND WAVES FROM QUANTUM FIELD THEORY

Let us conclude this chapter with a brief sketch of description of quantum objects within the framework of quantum field theory. The extension of non-relativistic quantum mechanics to relativistic quantum mechanics brings quantum field theory. There are many reasons to go for quantum field theory. The route for quantum field theory started with the quantization of electromagnetic field. When quantum mechanics was applied to electromagnetic filed, the concept of photons emerges naturally. This scheme has been extended to all the fundamental particles.

Let us briefly explain the framework of quantum field theory. What are the fundamental building blocks of the universe? First, it was thought that the indivisible atoms were the fundamental building blocks. Atom was found to be divisible in late nineteenth century. So, now particles like electrons and photons were considered as the fundamental objects. The advent of quantum field changed this picture. The quantum field theory changed the status of these elementary particles. The fundamental building blocks are now fields which fill the space.

For each elementary particle type, there is an associated underlying field filling the space. So, we have electromagnetic field, electron field, muon field, up quark field, and so on. All of these fields fill the space. These fields are the basic objects of the quantum field theory. What about the particle property and the wave property of photons, electrons, and so on? The particles are seen as the excitations of the respective fields. Their energy and momenta are $h\nu, 2h\nu, 3h\nu \ldots$ and $\hbar k, 2\hbar k, \ldots$. However, they are not like localized lumps in the sense of classical physics. The excitations belong to the entire field.

In that sense, they are like waves. But they are not classical waves. When there is an exchange of energy and momentum between the fields, the exchange takes place at a point in terms of discrete quanta of energy and momentum. So, an electromagnetic field absorbs energy $h\nu, 2h\nu, 3h\nu\ldots$ from the electron field at a point. Similarly, the electromagnetic field emits $h\nu, 2h\nu, 3h\nu\ldots$ at a point when it interacts with electron field.

What is the physical nature of the entity called photon? What is the physical nature of the entity called electron? How do they look like that? There is no simple answer to these questions. The photon of Einstein's conception in 1905 paper (he did not use the term for photon) and the photon in quantum electromagnetic differ in a very subtle way. The theoretical developments and the advances in technology make the knowledge of these quantum objects more and more complex. There is a constant evolution in our understanding of photon and other particles.

We may not have a simple picture about a photon or an electron. But, we have a huge body of knowledge of their behaviour in a wide variety of circumstances. We have built a huge theoretical machinery to predict the possible outcomes which can be verified in laboratories. We have been very successful in manipulating electrons and photons to obtain a vastly superior and reliable technology.

The major part of this book is devoted to non-relativistic quantum mechanics. The electrons are seen as point particles whose dynamics is governed by the laws of quantum mechanics.

EXERCISES

1. Why do you need two initial conditions to describe the motion of an object in classical mechanics?
2. In Hamiltonian formalism, what are the variables for which the initial conditions are needed?
3. Why is it difficult to obtain Poisson spot?
4. Is it necessary for a black body to be black-coloured object?
5. What are the conclusions that can be reached by thermodynamic considerations in the case of black body radiation?
6. How does the emissive power of an arbitrary object differ from that of a black body?
7. What is meant by the statement that a good absorber is a good emitter?
8. What is Einstein's conception of light quantum?
9. What is the principle of complementarity?
10. Can complementarity principle be violated?
11. What is the quantum mechanical explanation of the double-slit experiment?
12. In the double-slit experiment, do the photons going through slit *A* and the slit *B* interfere?
13. What is the significance of superposition principle in quantum mechanics?
14. Consider a single-slit diffraction of neutrons of wavelength $\lambda_{dB} = 18$ Å. The width of the slit is 90 μm. The screen is 5 m away from the slit. Determine the spacings between the minima.
15. A double-slit experiment is done with helium atoms of de Broglie wavelength $\lambda_{dB} = 1$Å. The width of each slit is 2 μm. The distance between the slit is 8 μm. The screen is at 64 cm from the slits. Determine the spacing between the minima in the intensity distribution.

REFERENCES

1. Arago spot, http://en.wikipedia.org/wiki/Arago_spot.
2. Anton Zeilinger, Roland Gőhler, C. G. Shull, Wolfgang Treince and Walter Mampe 1988, 'Single and Double-slit Diffraction of Neutrons', *Rev. of Mod. Phys.*, Vol. 60, p. 1067.

3. O. Carnal and J. Mlynek 1991, ‘Young’s Double-slit Experiment with Atoms—A Simple Atom Interferometer’, *Phys. Rev. Lett*., Vol. 66, p. 2689.
4. Markus Arndt, Olaf Nairz, Julian Voss-Andreae, Claudia Keller, Gerbrand van der Zouw and Anton Zeilinger 1999, ‘Wave-particle Duality of C60 Molecules’, *Nature*, Vol. 401, p. 680. See also Physicaplus (online magazine); www.univie.ac.at/qfp/research/matterwave/cbo/index.html.
5. Thomas Reisinger, Amil A. Patel, Herbert Reingruber, Katrin Fladishcher, Wolfgang E. Ernst, Gianangelo Bracco, Henry I. Smith and Bodil Holst 2009, ‘Poisson Spot with Molecules’, *Phys. Rev. A*., Vol. 79, 053823.
6. Claus Jońsson 1974, ‘Electron Diffraction at Multiple Slits’, *Am. J. Phys*., Vol. 4.
7. A. Tonomura, J. Endo, T. Matsuda, T. Kawasaki and H. Ezawa 1989, ‘Demonstration of Single-electron Buildup of an Interference Pattern’, *Am. J. Phys*., Vol. 57*,* p. 117.
8. Albert Einstein 1905, ‘On a Heuristic Point of View above the Creation and Conversion of Light’ *Ann. der. Phys*., Vol. 17, p. 132, Translation: http://en.wikisource.org/wiki/On_a_Heuristic_Point_of_View_about_the_Creation_and_Conversion_of_Light.
9. S. Wienberg 1977, ‘The Search for Unity: Notes for a History of Quantum Field Theory’, http://fafnir.phyast.pitt.edu/py3765/WienbergQFThistory.pdf
10. Robert A. Millikan 1923, *The Electron and the Light Quanta from the Experimental Point of View*, www.nobelprize*.org/nobel_prizes/physics/laureates/1923/millikan-lecture.pdf*.
11. R.A. Millikan, 1949, *Rev. Mod. Phys*, Vol. 21, 343.
12. A.Pias, 1979, *Rev. Mod. Phys*. Vol. 51, 863.
13. Evelyn Guha 2000, *Basic Thermodynamics*, Narosa Publishing House, New Delhi.

2

Schrödinger Equation

Quantum physics provides the basis to understand the entire universe. The study of quantum mechanics always begins with the non-relativistic quantum mechanics. The entire non-relativistic quantum mechanics centres around Schrödinger equation. The wide range of applications of Schrödinger equation is simply amazing. These include a number of different areas such as atomic physics, chemistry, and solid state physics. Schrödinger developed this equation in 1926.

How do we derive Schrödinger equation? The simple and straight answer is that Schrödinger equation cannot be derived or deduced from any other fundamental principles. However, one can provide a set of arguments leading to the formulation of Schrödinger equation. Each author develops Schrödinger equation in his or her own way. We have to bear in our mind that their derivations are textbook derivations and they do not reflect the way Schrödinger obtained this equation. The best way is to take Schrödinger equation as one of the postulates of quantum mechanics. (See Chapter 5). Here we give a 'derivation' closely following I. Schiff.

2.1 DERIVATION OF SCHRÖDINGER EQUATION

Matter like electrons exhibit wave-like property in some experimental set up and particle-like property in other experimental set up. This is the well-known wave particle duality that has already been discussed in Chapter 1. Now, let us concentrate on the wave nature of matter. If matter behaves like waves then they should have a wave equation. In the case of light, we know that the wave equation is given by Maxwell's equations. In the same way, what is the wave equation for an electron or any other particle in atomic world?

From general considerations, de Broglie suggested that the frequency and wavelength of the matter wave for a free particle is given by,

$$E = h\nu \quad \text{and} \quad \lambda = \frac{h}{p} \tag{2.1}$$

We can rewrite these equations in terms of the angular frequency $\omega = 2\pi\nu$ and the wave number $k = \frac{2\pi}{\lambda}$. We have

$$p = \frac{h}{\lambda} = \frac{h}{2\pi} \cdot \frac{2\pi}{\lambda} = \hbar k \tag{2.2}$$

and

$$E = h\nu = \frac{h}{2\pi} \cdot 2\pi\nu = \hbar\omega \tag{2.3}$$

Equations (2.2) and (2.3) just relate the kinematical particle properties, p and E to the parameters λ and ν describing wave motion. This does not suggest any relation between E and p. For instance, in the case of a non-relativistic free particle, $E = p^2/2m$ while for a relativistic free particle $E = (p^2c^2 + m^2c^4)^{1/2}$. Let us restrict ourselves to non-relativistic case. Then we have,

$$p = \hbar k;\ E = \hbar\omega = \frac{p^2}{2m} = \frac{\hbar^2 k^2}{2m}$$

We have to find a wave equation (i.e. a differential equation) consistent with these relations. Let us start our exercise first, with an attempt to find out an equation for a free particle. The wave function representing a free particle can be taken as a plane wave function. The possible plane wave functions are

$$\psi = \sin(kx - \omega t) \tag{2.4}$$

$$\psi = \cos(kx - \omega t) \tag{2.5}$$

$$\psi = e^{i(kx-\omega t)} \tag{2.6}$$

We should note that the interference and diffraction phenomena are due to superposition of two waves. If ψ_1 and ψ_2 are solutions of the wave equation, the existence of interference phenomenon requires that $\psi = \psi_1 + \psi_2$ should also be a solution. This is possible only for a linear partial differential equation.

Assuming that the wave functions given in Equations (2.4), (2.5), and (2.6) represent a physical situation like an electron beam, we have to look for a linear partial differential equation for which these functions are solutions. Normally, a wave equation is not constructed by starting from a particular solution. For instance, in the case of a string, the wave equation is obtained by applying Newton's second law of motion for the propagation of a disturbance in the string. In the same way, Maxwell's equations are the results of study of various electric and magnetic phenomena. In the case of Schrödinger equation, there is no way of obtaining it by studying a particular set of phenomena based on the force or interaction experienced by the particle. We try to construct an equation for which these functions are solutions. As an exercise in mathematics, in principle, we can obtain a number of partial differential equations for which these functions are solutions. Among these equations, it is the physically observable consequences resulting from particular equation that decides the correct choice of the equation. We have to bear these observations in mind when we derive Schrödinger equation.

Continuing our mathematical exercise of constructing a partial differential equation for which the functions in Equations (2.4), (2.5), and (2.6) are solutions, let us start with the standard well-known wave equation.

$$\frac{1}{\gamma^2}\frac{\partial^2\psi}{\partial t^2} = \frac{\partial^2\psi}{\partial x^2} \tag{2.7}$$

Let us feed the wave functions (2.4), (2.5), and (2.6) in Equation (2.7) and for all the wave functions, we get (use Equations (2.2) and (2.3))

$$\frac{-1}{\gamma^2}\frac{E^2}{\hbar^2}\psi = \frac{-p^2}{\hbar^2}\psi \tag{2.8}$$

or

$$\gamma^2 = \frac{E^2}{p^2} = \frac{p^2}{4m^2} \tag{2.9}$$

Can we accept Equation (2.7) with the choice $\gamma^2 = p^2/4m^2$? We would like to have an equation where the coefficients in the partial differential equations do not depend on the physical quantities such as E or p. This consideration is necessary for the existence of superposition of waves of different E or p. So we would like to look for some other equations for which the coefficients of various terms will contain only parameters such as mass m, change e, and Planck's constant $\hbar$.

Equation (2.8) suggests that our required differential equation should be such that the substitution of the wave functions (2.4), (2.5), and (2.6) in that equation should lead to the equation of the form

$$E\psi = \frac{p^2}{2m}\psi \tag{2.10}$$

The RHS can be obtained from the second-order derivative with respect to x. The LHS is possible with only a first-order derivative with respect to time, provided we consider only exponential function (2.6). Obviously, the wave functions sin $(px - Et)/\hbar$ and cos $(px - Et)/\hbar$ will not lead to such an equation. This suggests the required differential equation of the form

$$\alpha\frac{\partial\psi}{\partial t} = \beta\frac{\partial^2\psi}{\partial x^2} \tag{2.11}$$

On substituting the wave function $\psi = e^{i(kx-\omega t)}$ in Equation (2.11), we get

$$\frac{i\alpha E}{\hbar}\psi = \frac{-\beta p^2}{\hbar^2}\psi \tag{2.12}$$

The simplest choice for α and β are $\alpha = i\hbar$ and $\beta = \frac{-\hbar^2}{2m}$. So, the equation for a free particle is

$$i\hbar\frac{\partial\psi(x,t)}{\partial t} = -\frac{\hbar^2}{2m}\frac{\partial^2\psi}{\partial x^2} \tag{2.13}$$

How do we change this equation when interaction is introduced?

In classical mechanics, the force $F(x, t)$ acting on the particle describes the interaction of the particle with another particle or the external world. Equivalently, the interaction can also be described by using the concept of the potential energy. Let us now choose potential energy $V(x, t)$ as the quantity describing the interaction of the particle with its surrounding. The total energy E is given by

$$E = \frac{p^2}{2m} + V$$

Therefore, we have to modify the equation in such a way that when we put $V = 0$ it reduces to the Equation (2.13). The simplest way to modify the Equation (2.13) is to add $V\psi$ to it. The resulting equation is

$$i\hbar\frac{\partial\psi(x,t)}{\partial t} = -\frac{\hbar^2}{2m}\frac{\partial\psi(x,t)}{\partial x^2} + V(x,t)\psi(x,t) \tag{2.14}$$

This is the well-known Schrödinger equation describing a particle in one dimension. This can be easily extended to three-dimensional space.

$$i\hbar\frac{\partial\psi(\mathbf{r},t)}{\partial t} = -\frac{\hbar^2}{2m}\nabla^2\psi(\mathbf{r},t) + V(\mathbf{r},t)\psi(\mathbf{r},t) \tag{2.15}$$

2.2 HAMILTONIAN OPERATOR

In quantum mechanics, one has to introduce a number of new concepts that cannot be deduced from classical mechanics. One such idea is to associate an operator (may be a differential operator or a matrix operator) to each dynamical variable. How do we deduce experimental results or some numbers in a measurement process using such operators? This will be extensively discussed in Chapter 5. Right now, we can define two operators $\mathbf{p}_{op}$ and H_{op} corresponding to momentum and Hamiltonian. The momentum operator $\mathbf{p}_{op}$ is defined as

$$\mathbf{p}_{op} = -i\hbar\nabla$$

The Hamiltonian operator H_{op} is defined as

$$\begin{aligned} H_{op} &= \frac{\mathbf{p}_{op}\cdot\mathbf{p}_{op}}{2m} + V(\mathbf{r},t) \\ &= \frac{-\hbar^2}{2m}\nabla^2 + V(\mathbf{r},t) \end{aligned}$$

The energy operator E_{op} is defined as

$$E_{op} = i\hbar\frac{\partial}{\partial t} \tag{2.16}$$

Therefore, the Schrödinger equation can be written as

$$i\hbar\frac{\partial\psi}{\partial t}(\mathbf{r},t) = H_{op}\psi(\mathbf{r},t) \tag{2.17}$$

or

$$E_{op}\psi(\mathbf{r},t) = H_{op}\psi(\mathbf{r},t) \tag{2.18}$$

2.3 FREE-PARTICLE SOLUTION

The potential $V(x, t)$ describes the interaction between the particle and its surroundings. For a free particle, $V(x, t) = 0$. Therefore, the Schrödinger equation for a free particle is

$$i\hbar\frac{\partial\psi(x,t)}{\partial t} = -\frac{\hbar^2}{2m}\frac{\partial^2\psi(x,t)}{\partial x^2} \tag{2.19}$$

Let us assume that the solution is of the form

$$\psi(x,t) = Ne^{i(px-Et)/\hbar}$$

(Note that we derived the Schrödinger equation using this function. Now, we work in the reverse way.)
LHS:

$$i\hbar\frac{\partial\psi}{\partial t}(x,t) = i\hbar\left(\frac{-iE}{\hbar}\right)Ne^{i(px-Et)/\hbar} = E\psi$$

RHS:

$$\frac{-\hbar^2}{2m}\frac{\partial^2\psi}{\partial x^2} = \frac{-\hbar^2}{2m}\cdot\left(\frac{ip}{\hbar}\right)^2 Ne^{i(px-Et)/\hbar} = \frac{p^2}{2m}\psi$$

Since $E = p^2/2m$ for a free particle, LHS = RHS

Therefore, the wave function describing a free particle is

$$\psi = Ne^{i(px-Et)/\hbar}$$

In three dimensions, the free-particle Schrödinger equation is

$$i\hbar\frac{\partial\psi}{\partial t}(\mathbf{r},t) = \frac{-\hbar^2}{2m}\nabla^2\psi(\mathbf{r},t) \tag{2.20}$$

and the solution to this equation is

$$\psi(\mathbf{r},t) = Ne^{i(\mathbf{p}\cdot\mathbf{r}-Et)/\hbar} \tag{2.21}$$

2.4 INTERPRETATION OF WAVE FUNCTION $\psi(\mathbf{r}, t)$

Normally, in a partial differential equation for wave motion, the quantity $\psi(\mathbf{r},t)$ stands for the fluctuation of a physical quantity about equilibrium values. For instance, in the case of wave motion in a string, $\psi(\mathbf{r},t)$ represents displacement of the string from the equilibrium as shown in Fig. 2.1.

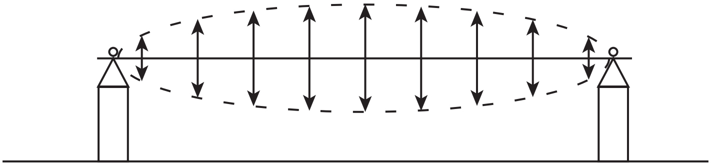


Fig. 2.1 Waves in a string

Similarly, in the case of sound in air, $\psi(\mathbf{r},t)$ represents fluctuations in pressure or density. In the case of a drum (or mirudhangam), $\psi(\mathbf{r},t)$ represents the vibrations of a membrane. In electromagnetic wave, $\psi(\mathbf{r},t)$ represents fluctuations in electric and magnetic fields. The question is, what is the physical property whose fluctuation is represented by $\psi(\mathbf{r},t)$ in Schrödinger equation?

Schrödinger developed this equation in 1926 and was very successful in explaining the energy levels of hydrogen atom. However, very soon it was realized that it is difficult to interpret $\psi(\mathbf{r},t)$ as fluctuation in any physical property. It was Max Born, who gave the probabilistic interpretation of the wave function $\psi(\mathbf{r},t)$. This is the standard interpretation of the wave function $\psi(\mathbf{r},t)$ in non-relativistic quantum mechanics.

Max Born's Interpretation

The wave function $\psi(\mathbf{r},t)$ is now interpreted as the probability amplitude and $\psi^*(\mathbf{r},t)\psi(\mathbf{r},t)d^3\mathbf{r}$ is interpreted as the probability of finding the particle in a volume $d^3\mathbf{r}$ centred around $\mathbf{r}$ at the instant t as shown in Fig. 2.2.

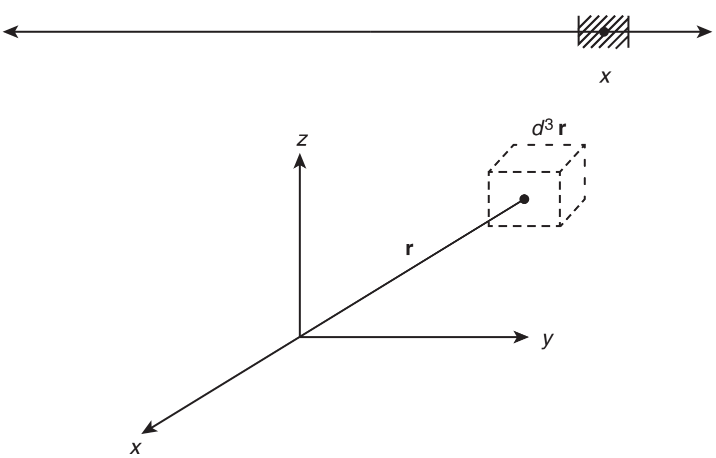


Fig. 2.2 (a) Interval dx at x (b) Volume $d^3\mathbf{r}$ centered at $\mathbf{r}$

In one dimension, $\psi^*(x,t)\psi(x,t)dx$ is interpreted as the probability of finding the particle in the interval dx centred around x.

Note that there are three quantities. The wave function $\psi(\mathbf{r},t)$ is probability amplitude. This can be a real number or a complex number. $\psi^*(\mathbf{r},t)\psi(\mathbf{r},t)$ is the probability density and it is denoted by $\rho(\mathbf{r},t)$ or $\rho(\mathbf{r},t)=\psi^*(\mathbf{r},t)\psi(\mathbf{r},t)$. This has to be a real number. The probability of finding the particle in a volume $d^3\mathbf{r}$ centred around $\mathbf{r}$ is given by

$$dP(\mathbf{r},t)=\rho(\mathbf{r},t)d^3\mathbf{r}=\psi^*(\mathbf{r},t)\psi(\mathbf{r},t)d^3\mathbf{r} \tag{2.22}$$

This has to be a positive real number between 0 and 1.

The probabilistic interpretation assumes that the particle to be indivisible point particle which implies that when a particle is detected it is detected as a whole though it is a random event.

2.5 NORMALIZATION OF WAVE FUNCTION

Let $\rho(x,t)$ be the position probability density for a particle. As mentioned above and shown in Fig. 2.3, the probability of finding the particle in the interval Δx_1 around x_1 is given by

$$P(\text{near } x_1)=\rho(x_1,t)\Delta x_1$$

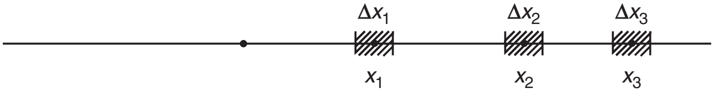


Fig. 2.3 Particle near x_1 or x_2 or x_3

Now, what is the probability of finding the particle near x_1 or x_2? The particle cannot be found to be present simultaneously at both x_1 and x_2. In the theory of probability, these two events are referred to as mutually exclusive events. So the probability of finding the particle within the interval Δx_1 near x_1 or within the interval Δx_2 near x_2 is given by

$$P(\text{near } x_1 \text{ or near } x_2)=\rho(x_1,t)\Delta x_1+\rho(x_2,t)\,\Delta x_2$$

Let us extend this idea further. The probability of finding the particle near x_1 or x_2 or x_3 is given by

$$P\,(\text{near } x_1 \text{ or } x_2 \text{ or } x_3)=\rho(x_1,t)\Delta x_1+\rho(x_2,t)\Delta x_2+\rho(x_3,t)\Delta x_3$$

Continuing this argument further, the probability of finding the particle near x_1 or x_2 or x_3 or x_4... is given by

$$P\,(\text{near } x_1 \text{ or } x_2 \text{ or } x_3 \text{ or } x_4\ldots)=\Sigma\rho(x_i,t)\Delta x_i$$

If there is a particle, it should be found somewhere between $x=-\infty$ and $x=\infty$. In probability theory, this event corresponds to a certain event. Therefore, the probability of finding the particle anywhere between $x=-\infty$ and $x=\infty$ should be equal to 1.

$$\text{(i.e.)}\quad \Sigma\rho(x_i,t)\,\Delta x_i=1$$

In the continuum limit, this result becomes an integral.

$$\text{(i.e.)}\quad \int_{-\infty}^{\infty}\rho(x,t)dx=1 \tag{2.23}$$

$$\int_{-\infty}^{\infty}\psi^*(x,t)\psi(x,t)dx=1 \tag{2.24}$$

In three dimensions, we get

$$\int_{-\infty}^{\infty} \psi^*(\mathbf{r},t)\psi(\mathbf{r},t)d^3\mathbf{r} = 1 \tag{2.25}$$

Equations (2.24) and (2.25) are known as normalization conditions.

Is it possible to normalize any function? The answer is, obviously, that it is possible to normalize only some kind of functions through the condition (2.24) or (2.25). This can be mainly seen as follows:

The normalization condition is

$$\int_{-\infty}^{\infty} \psi^*(x,t)\psi(x,t)dx = \int_{-\infty}^{\infty} |\psi^*(x,t)|^2\, dx = 1.$$

The first thing to be noted is that the integrand $|\psi(x,t)|^2$ is always a positive real number. The general sketch of $|\psi(x,t)|^2$ is likely to be like the curves shown in Fig. 2.4 (a) and (b).

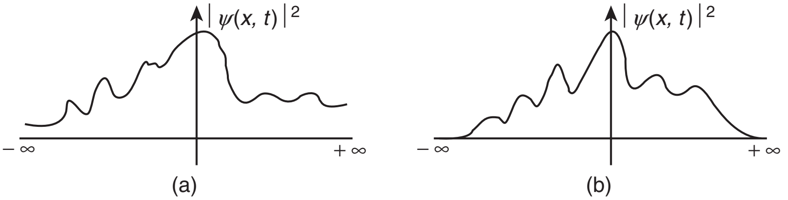


Fig. 2.4 Sketch of $|\psi(x)|^2$

In Fig. 2.4(a), $|\psi(x,t)|^2$ approaches a constant value for large values of x, whereas in Fig. 2.4(b), $|\psi(x,t)|^2$ approaches zero for large values of x. The integral $\int_{-\infty}^{\infty} |\psi(x,t)|^2 dx$ is equal to the area under the curve $|\psi(x,t)|^2$ bound by x–axis extending from $x = -\infty$ to $x = \infty$. If the limits of the integral are finite, then the area under this curve will be always a positive number. In the case of limits of the integral going to infinity, one has to be careful. In this case, the area under this curve can be finite, provided $|\psi(x,t)|^2$ goes to zero at large values of x, that is, $|\psi(x,t)|^2 \to 0$ as $x \to \pm\infty$. In Fig. 2.4(a), $|\psi(x,t)|^2$ approaches a constant in the large x limit. As long as this constant is different from zero (even it may be a very small number) the area under this curve is infinite. On the other hand, in Fig. 2.4(b), $|\psi(x,t)|^2 \to 0$ as $x \to \pm\infty$. The area under this curve is a finite number (may not be equal to 1). More technically, one has to analyse the convergence of the integral $\int_{-\infty}^{\infty} |\psi(x,t)|^2 dx$. For a continuous function $\psi(x, t)$, the integral $\int_{-\infty}^{\infty} |\psi(x,t)|^2 dx$ converges to a finite number provided $\psi(x, t) \to 0$ as $x \to \pm\infty$.

$$\text{(i.e.)} \quad \int_{-\infty}^{\infty} |\psi(x,t)|^2 dx \text{ is a finite number, provided } \psi(x, t) \to 0 \text{ as } x \to \pm\infty \tag{2.26}$$

This finite number need not be equal to 1. It is enough for the integral to converge to a finite number. In such cases, normalization can be done by introducing a suitable normalization constant.

As an example, consider $\psi(x) = e^{-\alpha x^2}$. The normalization integral is $\int_{-\infty}^{\infty} |\psi(x)|^2 dx = \int_{-\infty}^{\infty} e^{-2\alpha x^2} dx = \sqrt{\frac{\pi}{2\alpha}}$. Obviously, $\sqrt{\frac{\pi}{2\alpha}}$ is a finite number, but it is not equal to one. Let us redefine the function $\psi(x)$ as $\psi(x) = \left(\frac{2\alpha}{\pi}\right)^{1/4} e^{-\alpha x^2}$. Now, $\int_{-\infty}^{\infty} |\psi(x)|^2 dx = \left(\frac{2\alpha}{\pi}\right)^{1/2} \left(\frac{\pi}{2\alpha}\right)^{1/2} = 1$.

Example 2.1 Normalize the wave function $\psi(x)$ given by

$$\psi(x) = \frac{N}{(a^2 + x^2)}$$

Solution: The normalization condition is

$$\int_{-\infty}^{\infty} \psi^*(x)\psi(x)dx = \int_{-\infty}^{\infty} \frac{N^2}{(a^2 + x^2)^2} dx = 1.$$

Put $x = a \tan\theta$. Then limits of the integral changes to $\theta = \frac{-\pi}{2}$ and $\theta = \frac{\pi}{2}$.

$$dx = a \sec^2\theta \, d\theta$$

$$\therefore (a^2 + x^2)^2 = (a^2 + a^2 \tan^2\theta)^2 = a^4(1 + \tan^2\theta)^2 = a^4 \sec^4\theta$$

$$\therefore \int_{-\infty}^{\infty} \frac{N^2}{(a^2 + x^2)^2} dx = \int_{-\pi/2}^{\pi/2} \frac{aN^2 \sec^2\theta}{(a^4 \sec^4\theta)} d\theta = \frac{N^2}{a^3} \int_{-\pi/2}^{\pi/2} \cos^2\theta d\theta$$

$$= \frac{N^2}{a^3} . \pi = 1$$

$$\therefore N = \left(\frac{a^3}{\pi}\right)^{1/2}$$

Therefore, the normalized wave function $\psi(x)$ is given by

$$\psi(x) = \left(\frac{a^3}{\pi}\right)^{1/2} \frac{1}{a^2 + x^2}$$

Plane Wave Function and Normalization

Let us now try to normalize the plane wave function $\psi(x,t) = Ne^{i(px - Et)/\hbar}$ through the normalization condition $\int_{-\infty}^{\infty} \psi^*(x,t)\psi(x,t)dx = 1$.

$$\int_{-\infty}^{\infty} \psi^*(x,t)\psi(x,t)dx = N^2 \int_{-\infty}^{\infty} e^{-i(px - Et)/\hbar} e^{i(px - Et)/\hbar} dx$$

$$= N^2 \int_{-\infty}^{\infty} dx = \infty$$

Obviously, the plane wave function is non–normalisable through the normalization condition $\int_{-\infty}^{\infty} \psi^*(x,t)\psi(x,t)dx = 1$.

Later in Section 2.14 we will show that the plane wave function can be normalized by employing different normalization condition.

2.6 SQUARE INTEGRABLE FUNCTIONS

A function $f(x)$ is said to be square integrable if it satisfies the condition:

$$\int_{-\infty}^{\infty} |f(x)^2|\,dx < \infty \text{ or } \int_{-\infty}^{\infty} f^*(x)f(x)\,dx < \infty \tag{2.27}$$

The meaning of the above statement is that the integral $\int_{-\infty}^{\infty} |f(x)^2|\,dx$ converges to a finite number.

$$\text{(i.e.)} \quad \int_{-\infty}^{\infty} |f(x)^2|\,dx = N, \text{ where } N \text{ is a finite number.} \tag{2.28}$$

In three dimensions, $f(\mathbf{r})$ is said to be square integrable if it satisfies the condition:

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |f(\mathbf{r})|^2 d^3\mathbf{r} < \infty \tag{2.29}$$

or

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |f(\mathbf{r})|^2 d^3\mathbf{r} = N, \text{ where } N \text{ is a finite number} \tag{2.30}$$

The fact that the integral $\int_{-\infty}^{\infty} |f(x)^2|\,dx$ should converge to finite number N imposes certain conditions on $f(x)$. All the continuous functions $f(x)$ which go to zero as $x \to \pm\infty$ are square integrable. However, this limiting behaviour is not necessary for a function to be square integrable. Within quantum mechanics, we consider only the kind of square integrable functions which go to zero for large values of x.

So, in quantum mechanics, we choose only one kind of square integrable wave functions $\psi(\mathbf{r},t)$ for which

$$\psi(\mathbf{r},t) \to 0 \text{ as } x \to \pm\infty, y \to \pm\infty, z \to \pm\infty \tag{2.31}$$

Let us reiterate the significance of square integrable functions in quantum mechanics. There are many kinds of solutions to Schrödinger equation. Among these solutions, the only solution relevant to quantum mechanics is square integrable functions. This is the consequence of the interpretation of the wave function as probability amplitude and $\psi^*\psi$ as the probability density for the position of a particle.

There are some set of functions which are widely used in quantum mechanics in spite of the fact that they are not square integrable. One such example is the plane wave function $\psi(\mathbf{r},t) = Ne^{i(px-Et)/\hbar}$ (or $\psi = Ne^{i(\mathbf{p}\cdot\mathbf{r}-Et)/\hbar}$ in three dimensions).

2.7 BOUND AND SCATTERING STATES

What is the meaning of the boundary conditions

$$\psi(\mathbf{r},t) \to 0 \quad \text{as } x \to \pm\infty, y \to \pm\infty, \text{ and } z \to \pm\infty?$$

The fact that $\psi(\mathbf{r},t)$ goes to zero at large $\mathbf{r}$ means that the probability density $\rho(\mathbf{r},t)$ becomes zero at large $\mathbf{r}$ and so the probability of finding the particle is zero at large $\mathbf{r}$ implying that the particles are confined to some finite regions. In other words, such states are bound states in the sense that a particle is bound to remain in a finite region. For example, consider the wave function for an electron obeying the boundary conditions (2.31) in the case of a hydrogen atom. The boundary conditions imply that as one moves away from the nucleus, the probability of finding the electron falls to zero. Obviously, the electron in the hydrogen atom is bound to the nucleus. Therefore, the wave functions which obey the boundary conditions (2.31) are known as the bound state wave functions.

On the other hand, $e^{i(kx-\omega t)}$ represents free-particle solutions. They are not bound to any particular region. It is to be noted that the unbound states need not necessarily mean free-particle states. We can have unbound states even in the presence of potential $V(\mathbf{r})$. It is possible to have solutions which behave like $e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)}$ at large distances. Such a large $\mathbf{r}$ behaviour occurs, for instance, in the scattering theory. Such a wave function is called unbound or scattering–state wave functions.

So, we have two kinds of states in quantum mechanics:

Bound state: $$\psi(\mathbf{r},t) \to 0 \quad \text{as } x \to \pm\infty, y \to \pm\infty, z \to \pm\infty \tag{2.32}$$

Unbound or scattering state: $$\psi(\mathbf{r},t) \to e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)} \quad \text{as } x \to \pm\infty, y \to \pm\infty, z \to \pm\infty \tag{2.33}$$

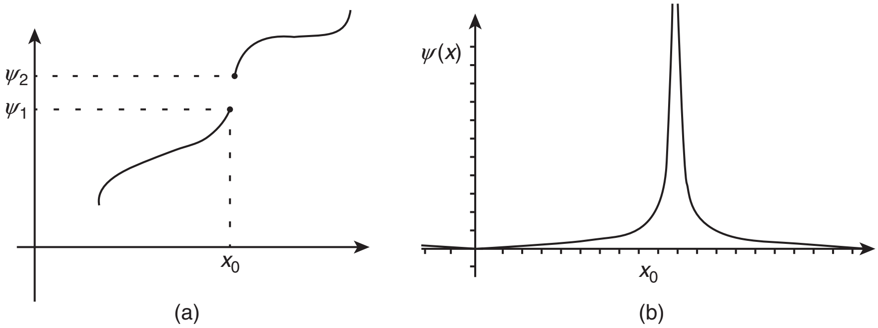
All the square integrable wave functions represent bound state and the non-square integrable wave functions represent unbound state.

2.8 ADMISSIBILITY CONDITIONS ON A WAVE FUNCTION IN QUANTUM MECHANICS

The probability interpretation of ψ and the fact that ψ satisfies a partial differential equation together put some restrictions on $\psi(\mathbf{r},t)$ to accept them as wave functions as allowed in quantum mechanics.

We have already seen that the probability interpretation requires the normalization of the wave function which in turn means that within quantum mechanics, only square integrable functions are accepted. Of course, free particle states and scattering states are exceptions to these requirements.

The probability interpretation further requires that the wave function $\psi(\mathbf{r},t)$ to be finite and single valued and continuous. The fact that $\psi^*(\mathbf{r},t)\psi(\mathbf{r},t)d^3\mathbf{r}$ represents the probability means that this quantity should be a number between 0 and 1. This is possible, provided that $\psi(\mathbf{r},t)$ has a finite and unique value. If $\psi(\mathbf{r},t)$ has a discontinuity at $\mathbf{r}$, as shown in Fig. 2.5, then it will not be well defined at $\mathbf{r}$ and hence it cannot represent probability amplitude.



(a) (b)

Fig. 2.5 Wave functions with discontinuities

In Fig. 2.5(a), the function $\psi(x, t)$ has a finite discontinuity at $x = x_0$: $\psi(x, t)$ does not have a unique value at $x = x_0$. If we go along the lower branch, $\psi(x, t)$ becomes ψ_1 at x_0 and if we go along the upper branch $\psi(x,t)$ becomes ψ_2 at x_0. If one accepts this function as allowed function, then the probability of finding the particle is not unique at x_0. Obviously, we would like to avoid such a situation. On the other hand, in Fig. 2.5(b) $\psi(x, t)$ is discontinuous at $x = x_0$ by becoming infinite making the probability to become infinite. We would like to avoid this kind of situation also. So, the probability interpretation requires that the wave function $\psi(\mathbf{r},t)$ is single valued, finite, and continuous at all points.

This is also consistent with the fact that $\psi(\mathbf{r},t)$ satisfies a partial differential equation. The Schrödinger equation has second–order partial derivatives in x, y, and z and a first–order derivative in time t. This is possible, provided $\psi(\mathbf{r},t)$ and its first–order derivatives are differentiable. This implies the following:

1. $\psi(\mathbf{r},t)$ is single valued, finite, and continuous at all points.
2. $\dfrac{\partial\psi}{\partial x},\dfrac{\partial\psi}{\partial y}$ and $\dfrac{\partial\psi}{\partial z}$ are single valued, finite, and continuous.

The second requirement depends on the behaviour of the potential $V(\mathbf{r},t)$. The potential $V(\mathbf{r},t)$ may, in general, be continuous or discontinuous function. If $V(\mathbf{r},t)$ is a continuous function then the second requirement follows naturally. If the potential $V(\mathbf{r},t)$ has discontinuity, then one has to see whether the function has a finite discontinuity or infinite discontinuity. If $V(\mathbf{r},t)$ has a finite discontinuity, the second–order partial derivatives $\delta^2\psi$(have) must also be discontinuous such that their discontinuities cancel the discontinuity arising from $V(\mathbf{r},t)$. If $V(\mathbf{r},t)$ has infinities like the delta function potential or the potential for a particle in a box, the second requirement breaks down.

Hence, the admissibility conditions for a wave function $\psi(\mathbf{r},t)$ to accept it as an allowed wave function in quantum mechanics are the following:

1. $\psi(\mathbf{r},t)$ is single valued, finite, and continuous at all points (2.34)
2. The first–order derivatives $\dfrac{\partial\psi}{\partial x},\dfrac{\partial\psi}{\partial y}$, and $\dfrac{\partial\psi}{\partial z}$ are single valued, finite, and continuous at all points (2.35)
3. For bound states,

$$\psi(\mathbf{r},t) \to 0 \text{ as } x \to \pm\infty, y \to \pm\infty, z \to \pm\infty \tag{2.36}$$

If $V(\mathbf{r},t)$ becomes infinite at a point, then the second requirement breaks down at that point. These three conditions are also known as boundary conditions on $\psi(\mathbf{r},t)$. To analyse a physical situation using quantum mechanics, we have to solve the Schrödinger equation, subject to the boundary conditions (2.34), (2.35), and (2.36).

2.9 CONSERVATION OF PROBABILITY

Let us reiterate again that in non–relativistic quantum mechanics, the time evolution of the wave function $\psi(\mathbf{r},t)$ is governed by Schrödinger equation:

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

The probability density $\rho(\mathbf{r},t)$ is given by

$$\rho(\mathbf{r},t) = \psi^*(\mathbf{r},t)\psi(\mathbf{r},t)$$

If there exists a particle, it should be found at some point in space.

$$\therefore \int_{\Omega} \rho(\mathbf{r},t)\, d^3\mathbf{r} = 1 \text{ or } \int_{\Omega} \psi^*(\mathbf{r},t)\psi(\mathbf{r},t) d^3\mathbf{r} = 1 \tag{2.37}$$

where Ω is the entire volume of the space.

Note that the LHS involves a function that depends on time, but the RHS is independent of time. The probability of finding a particle at some location in the entire space is always or the total probability is constant in time. So, the total probability $\int_{\Omega} \rho(\mathbf{r},t) d^3\mathbf{r}$ is said to be a conserved quantity. Our task is to prove that the time evolution of $\psi(\mathbf{r},t)$ (as governed by Schrödinger equation) is such that the Equation (2.37) is true at all instants of time. Let us evaluate $\frac{\partial}{\partial t}\int \rho(\mathbf{r},t) d^3\mathbf{r}$.

$$\frac{\partial}{\partial t}\int_{\Omega} \rho(\mathbf{r},t) d^3\mathbf{r} = \frac{\partial}{\partial t}\int_{\Omega} \psi^*(\mathbf{r},t)\psi(\mathbf{r},t) d^3\mathbf{r}$$

$$= \int_{\Omega}\left[\frac{\partial \psi^*}{\partial t}\psi + \psi^*\frac{\partial \psi}{\partial t}\right] d^3\mathbf{r}$$

Using Schrödinger equation and its conjugate, we get

$$\frac{\partial}{\partial t}\int_{\Omega} \rho(\mathbf{r},t) d^3\mathbf{r} = \int_{\Omega}\left\{\left[\frac{1}{i\hbar}\frac{\hbar^2}{2m}\nabla^2\psi^* - V\psi^*\right]\psi + \frac{1}{i\hbar}\left[\frac{-\hbar^2}{2m}\nabla^2\psi + V\psi\right]\psi^*\right\} d^3\mathbf{r}$$

$$= \frac{\hbar}{2im}\int_{\Omega} (\nabla^2\psi^*\psi - \nabla^2\psi\,\psi^*) d^3\mathbf{r}$$

$$= \frac{\hbar}{2im}\int_{\Omega} \nabla\cdot(\nabla\psi^*\psi - \nabla\psi\,\psi^*) d^3\mathbf{r}$$

Convert the volume integral $\int_{\Omega}$ into the surface integral $\int_{\Sigma}$ where Σ is the surface enclosing the volume Ω.

$$\therefore \frac{\partial}{\partial t}\int_{\Omega} \rho(\mathbf{r},t)\, d^3\mathbf{r} = \frac{\hbar}{2im}\int_{\Sigma} \nabla\cdot(\nabla\psi^*\psi - \nabla\psi\,\psi^*)\, d^3\mathbf{r} \tag{2.38}$$

We have to bear in mind that Ω is the whole space and so the values of x, y, and z on the surface Σ is very large and hence ψ and ψ^* are zero on the surface Σ. Therefore, the integrand is zero on the surface Σ, and hence the surface integral is zero.

$$\therefore \frac{\partial}{\partial t}\int_{\Omega} \psi^*(\mathbf{r},t)\, \psi(\mathbf{r},t) d^3\mathbf{r} = 0 \tag{2.39}$$

$\int_{\Omega} \psi^*(\mathbf{r},t)\, \psi(\mathbf{r},t)\, d^3\mathbf{r}$ is independent of time, and hence the total probability is a conserved quantity.

Equation of Continuity

There are many ways to express the conservation of a quantity. One such method is to employ the equation of continuity which is widely used in branches like hydrodynamics or electrodynamics. Hence, let us derive the equation of continuity here to express the conservation of probability.

Schrödinger equation and its conjugate are

$$i\hbar\frac{\partial\psi}{\partial t}=\frac{-\hbar^2}{2m}\nabla^2\psi+V\psi \tag{2.40}$$

$$-i\hbar\frac{\partial\psi^*}{\partial t}=\frac{-\hbar^2}{2m}\nabla^2\psi^*+V\psi^* \tag{2.41}$$

(2.40) × ψ^* – (2.41) × ψ gives

$$i\hbar\psi^*\frac{\partial\psi}{\partial t}+i\hbar\frac{\partial\psi^*}{\partial t}\psi=\frac{-\hbar^2}{2m}\psi^*\nabla^2\psi+\frac{\hbar^2}{2m}\nabla^2\psi^*\psi$$

$$\text{or } i\hbar\frac{\partial}{\partial t}(\psi^*\psi)=\frac{-\hbar^2}{2m}[\psi^*\nabla^2\psi-\nabla^2\psi^*\psi]=\frac{-\hbar^2}{2m}\nabla\cdot[\psi^*\nabla\psi-\nabla\psi^*\psi]$$

$$\begin{aligned}&\frac{\partial}{\partial t}(\psi^*\psi)+\nabla\cdot\frac{\hbar}{2im}[\psi^*\nabla\psi-\nabla\psi^*\psi]=0\\ \text{or}\qquad&\frac{\partial\rho}{\partial t}+\nabla\cdot\mathbf{S}=0\end{aligned} \tag{2.42}$$

provided, we identify $\rho(\mathbf{r},t)$ and $\mathbf{S}(\mathbf{r},t)$ as

$$\rho(\mathbf{r},t)=\psi^*(\mathbf{r},t)\psi(\mathbf{r},t) \tag{2.43}$$

and

$$\mathbf{S}(\mathbf{r},t)=\frac{\hbar}{2im}(\psi^*\nabla\psi-\nabla\psi^*\psi) \tag{2.44}$$

This equation is very much similar to the equation of continuity in electrodynamics, $\partial\rho/\partial t+\nabla\cdot\mathbf{J}=0$, where ρ is the charge density and $\mathbf{J}$ is the current density. $\rho(\mathbf{r},t)$ in Equation (2.44) is identified as the probability density and $\mathbf{S}(\mathbf{r},t)$ is called the probability current density.

Have we derived or deduced Max Born's probability interpretation? It is true that the expressions for $\rho(\mathbf{r},t)$ and $\mathbf{S}(\mathbf{r},t)$ were deduced from Schrödinger equation. This is purely a mathematical exercise. The equation of continuity occurs in other branches of physics also like hydrodynamics and electrodynamics. So, this equation does not necessarily imply any particular interpretation of $\rho(\mathbf{r},t)$.

By comparing with other fields, all that we can do is to interpret $\rho(\mathbf{r},t)$ as the density of some physical property and **S** as corresponding 'current density' for the flow of that physical property. So the question is: If $\rho(\mathbf{r},t)$ is to be interpreted as a density of a physical property, what is that physical property described by $\psi(\mathbf{r},t)$? We cannot answer this question from the knowledge of other similar quantities in other fields. So, Born's interpretation cannot be deduced as a natural consequence of equation of continuity.

However, the equation of continuity restricts the form of the expression for $\rho(\mathbf{r},t)$ in terms of ψ for a given differential equation.

Conservation of Probability and Equation of Continuity

The equation of continuity is

$$\frac{\partial\rho(\mathbf{r},t)}{\partial t}+\nabla\cdot\mathbf{S}=0$$

The interpretation of this equation is simple and very much similar to the interpretation in the case of electromagnetic theory or hydrodynamics. Consider a small volume $\Delta\tau$ in space. Then

$$\frac{\partial\rho(\mathbf{r},t)}{\partial t}\Delta\tau=-\nabla\cdot\mathbf{S}\Delta\tau \tag{2.45}$$

$$\frac{\partial}{\partial t}(\rho(\mathbf{r},t)\Delta\tau) = -\nabla\cdot\mathbf{S}\Delta\tau \tag{2.46}$$

$\rho(\mathbf{r},t)\Delta\tau$ is the total amount of probability in volume $\Delta\tau$ (similar to the quantity of water in a tank of volume $\Delta\tau$). $\partial/\partial t(\rho(\mathbf{r},t)\Delta\tau)$ is the rate of change of amount of probability in volume $\Delta\tau$ (similar to rate of change of quantity of water in the tank). Conservation of probability means that this should be equal to net inflow of probability into the volume $\Delta\tau$ (similar to net inflow of water into the tank). Mathematically, from vector calculus, we know that $(-\nabla\cdot\mathbf{S}\Delta\tau)$ represents the net inflow of probability into the volume $\Delta\tau$. So, the equation of continuity is another way of stating conservation of probability.

Example 2.2 Obtain the probability density and probability current density for free particles which are described by $\phi_{\mathbf{p}}(\mathbf{r},t) = Ne^{i(\mathbf{p}\cdot\mathbf{r}-Et)/\hbar}$.

Solution: The probability density is $\rho(\mathbf{r},t) = \phi_{\mathbf{p}}^{*}(\mathbf{r},t)\phi_{\mathbf{p}}(\mathbf{r},t) = |N|^2$ (2.47)

The current density is

$$\mathbf{S} = \frac{\hbar}{2im}\left[\phi_{\mathbf{p}}^{*}\nabla\phi_{\mathbf{p}} - \nabla\phi_{\mathbf{p}}^{*}\phi_{\mathbf{p}}\right].$$

$$= |N|^2\,\frac{\mathbf{p}}{m} = \rho\frac{\mathbf{p}}{m} \tag{2.48}$$

Note that the expression for $\rho\,\mathbf{p}/m$ is the current density or the flux corresponding to flow of particles in classical physics. Same interpretation is maintained in quantum physics also. For plane waves, **S** represents the flux of the particles if we interpret ρ as number density of the particles. In the case of large number of particles, ρ indeed can be interpreted as number density of particles.

Equation of Continuity in One Dimension

The equation of continuity is of special interest to us, especially when one tries to apply Schrödinger equation for potentials in one dimension.

The equation of continuity is

$$\frac{\partial\rho}{\partial t} = -\frac{\partial S_x}{\partial x} \tag{2.49}$$

Probability current through the interval *ab* is shown in Fig. 2.6.

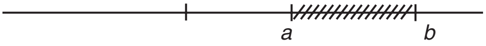


Fig. 2.6 Probability current through the interval *ab*

Integrating both sides of the equation from $x = a$ to $x = b$, we get

$$\int_a^b \frac{\partial\rho}{\partial t}dx = -\int_a^b \frac{\partial S_x}{\partial x}dx$$

or

$$\frac{\partial}{\partial t}\int_a^b \rho(x,t)dx = -\int_a^b \frac{\partial S_x}{\partial x}dx$$

$$= S_x(x = a,\, t) - S_x(x = b,\, t) \tag{2.50}$$

That is, rate of increase of probability within region *ab*

$$= (\text{current entering at } x = a - \text{current leaving at } x = b) \tag{2.51}$$

Suppose $\rho(x, t)$ is independent of time. Then, we have

$$\frac{\partial \rho(x,t)}{\partial t} = 0 \quad \text{or} \quad \frac{\partial}{\partial t}\int_a^b \rho(x,t)dx = 0 \tag{2.52}$$

In such a case, the current entering is the same as the current leaving.

2.10 TIME-INDEPENDENT SCHRÖDINGER EQUATION

Let the potential $V(\mathbf{r},t)$ be independent of time. Then $V(\mathbf{r},t) = V(\mathbf{r})$.

The Schrödinger equation now is $i\hbar\dfrac{\partial \psi(\mathbf{r},t)}{\partial t} = \left[-\dfrac{\hbar^2}{2m}\nabla^2 + V(\mathbf{r})\right]\psi(\mathbf{r},t)$.

Note that the wave function $\psi(\mathbf{r},t)$ is still time-dependent even when the potential is independent of time. However, the time independence of the potential simplifies our analysis and leads to time-independent Schrödinger equation.

Let us write $\psi(\mathbf{r},t)$ as

$$\psi(\mathbf{r},t) = T(t)u(\mathbf{r})$$

Then,

LHS: $$i\hbar\frac{\partial \psi(\mathbf{r},t)}{\partial t} = i\hbar\frac{dT(t)}{dt}u(\mathbf{r})$$

RHS: $$\left[\frac{-\hbar}{2m}\nabla^2 + V(\mathbf{r})\right]\psi(\mathbf{r},t) = T(t)\left[-\frac{h^2}{2m}\nabla^2 u(\mathbf{r}) + V(\mathbf{r})u(\mathbf{r})\right]$$

Equating LHS and RHS, we get

$$i\hbar\, u(\mathbf{r})\frac{dT(t)}{dt} = T(t)\left[-\frac{\hbar^2}{2m}\nabla^2 u(\mathbf{r}) + V(\mathbf{r})u(\mathbf{r})\right]$$

Dividing throughout by $u(\mathbf{r})$ $T(t)$, we get,

$$\frac{i\hbar}{T(t)}\frac{dT(t)}{dt} = \frac{1}{u(\mathbf{r})}\left[-\frac{h^2}{2m}\nabla^2 u(\mathbf{r}) + V(\mathbf{r})u(\mathbf{r})\right] \tag{2.53}$$

Note that in Equation (2.53), LHS is independent of **r** and RHS is independent of time *t* and since they are part of an equation, both LHS and RHS are independent of time *t* and position vector **r**. Therefore, both LHS and RHS should be equal to a constant. Let us choose this constant as *E*. So, we have

$$\frac{i\hbar}{T(t)}\frac{dT(t)}{dt} = \frac{1}{u(\mathbf{r})}\left[-\frac{h^2}{2m}\nabla^2 u(\mathbf{r}) + V(\mathbf{r})u(\mathbf{r})\right] = E$$

$$\frac{i\hbar}{T(t)}\frac{dT(t)}{dt} = E \tag{2.54}$$

and $$\left[-\frac{h^2}{2m}\nabla^2 + V(\mathbf{r})\right]u(\mathbf{r}) = Eu(\mathbf{r}) \tag{2.55}$$

Equation (2.54) can be re–written as

$$\frac{dT}{dt} = -\frac{iE}{\hbar}T \tag{2.56}$$

The solution of this equation is $T(t) = e^{-iEt/\hbar}$.
Therefore the, solution $\psi(\mathbf{r},t)$ is given by

$$\psi(\mathbf{r},t) = e^{-iEt/\hbar}u_E(\mathbf{r}) \tag{2.57}$$

where $u_E(\mathbf{r})$ satisfies the equation

$$\left[-\frac{\hbar^2}{2m}\nabla^2 + V(\mathbf{r})\right]u_E(\mathbf{r}) = E\,u_E(\mathbf{r}). \tag{2.58}$$

This equation is known as time-independent Schrödinger equation. This is the equation that is most widely used in non-relativistic quantum mechanics.

2.11 ENERGY EIGENSTATES AND THEIR PROPERTIES

Let us recall that the Hamiltonian operator $H_{op} = -\frac{\hbar^2}{2m}\nabla^2 + V(\mathbf{r},t)$. Obviously, the time-independent Schrödinger equation is

$$H_{op}\,u_E(\mathbf{r}) = E\,u_E(\mathbf{r}) \tag{2.59}$$

This is an eigenvalue equation. The eigenvalues are energies of the system. For a given Hamiltonian H_{op}, one has to solve this equation to find out the energy eigenvalues and the corresponding energy eigenstates.

Time Evolution of Energy Eigenstates

The wave function $\psi(\mathbf{r},t)$ satisfying the time-dependent Schrödinger equation is given by

$$\psi(\mathbf{r},t) = T(t)u(\mathbf{r}) = e^{-iEt/\hbar}u_E(\mathbf{r}) \tag{2.60}$$

where $u_E(\mathbf{r})$ is the energy eigenstate corresponding to the energy eigenvalue E. Let us assume that the system initially starts in $u_E(\mathbf{r})$.

(i.e.) At $t = 0$, $\psi(\mathbf{r},0) = u_E(\mathbf{r})$

Then $\psi(\mathbf{r},t)$ and $\psi(\mathbf{r},0)$ differ only in phase factor $e^{-iEt/\hbar}$ and therefore they represent the same physical state.

Stationary Energy Eigenstate

The energy eigenstate of the time-independent Hamiltonian is known as stationary energy eigenstate. What is stationary here? For an energy eigenstate, the probability $\rho(\mathbf{r}, t)$ (given by Equation (2.61)) is independent of time. Furthermore, it can be shown (see Chapter 5) that for a collection of systems in which all the systems are in energy eigenstate, the expectation value of a dynamical variable A (provided $\partial A/\partial t$ is zero) is independent of time. These are the reasons why the energy eigenstates of the time-independent Hamiltonian are known as stationary states. We have to recognize that the time independence of the Hamiltonian means that the system is left isolated. Therefore, if a system starts in a particular energy eigenstate and if the system is left isolated it will continue to remain in that same energy eigenstate. To change the state of the system from a particular initial energy eigenstate to another energy eigenstate we need to perturb the system, in which case the Hamiltonian is no longer time-independent. We will see in Chapter 13 that it is the time-dependent Hamiltonian which induces transition from one stationary energy eigenstate to another.

Interpretation of Energy Eigenstates

The interpretation of $u_E(\mathbf{r})$ follows from the interpretation of $\psi(\mathbf{r},t)$. The time-dependent wave function for the stationary energy eigenstate is

$$\psi_E(\mathbf{r},t) = e^{-iEt/\hbar}u_E(\mathbf{r})$$

The probability of finding the particle in a volume $d^3\mathbf{r}$ around $\mathbf{r}$ is given by

$$\psi^*(\mathbf{r},t)\psi(\mathbf{r},t)d^3\mathbf{r} = u_E^*(\mathbf{r})u_E(\mathbf{r})d^3\mathbf{r} \tag{2.61}$$

Therefore, $u_E^*(\mathbf{r})u_E(\mathbf{r})$ is the probability density for the position of the particle.

Normalization of Energy Eigenstates

The normalization condition for $\psi_E(\mathbf{r},t)$ is given by

$$\int \psi_E^*(\mathbf{r},t)\psi_E(\mathbf{r},t)d^3\mathbf{r} = 1$$

$$\int e^{iEt/\hbar}u_E^*(\mathbf{r})e^{-iEt/\hbar}u_E(\mathbf{r})d^3\mathbf{r} = 1$$

or

$$\int u_E^*(\mathbf{r})u_E(\mathbf{r})d^3\mathbf{r} = 1 \tag{2.62}$$

Obviously, the probability interpretation demands that $u_E(\mathbf{r})$ to be a square integrable function.

Boundary Conditions for Energy Eigenstates

The boundary conditions for $u_E(\mathbf{r})$ are literally the same as the boundary conditions for $\psi(\mathbf{r},t)$. Therefore, we have the following:

1. $u_E(\mathbf{r})$ is finite, single valued, and continuous at all points (2.63)
2. $\dfrac{\partial u_E}{\partial x}, \dfrac{\partial u_E}{\partial y}$, and $\dfrac{\partial u_E}{\partial z}$ are finite, single valued, and continuous at all points (2.64)
3. For bound states

$$u_E(\mathbf{r}) \to 0 \text{ as } x \to \pm\infty, y \to \pm\infty, z \to \pm\infty \tag{2.65}$$

The second set of boundary conditions is valid for both continuous potential $V(\mathbf{r})$ and $V(\mathbf{r})$ with finite discontinuities. However, this condition breaks down when the potential $V(\mathbf{r})$ becomes infinite.

Example 2.3 Show that $\partial u_E/\partial x$ is continuous at $\bar{x}$ for a potential $V(x)$, which is discontinuous at $\bar{x}$ with finite discontinuity as shown in Fig. 2.7.

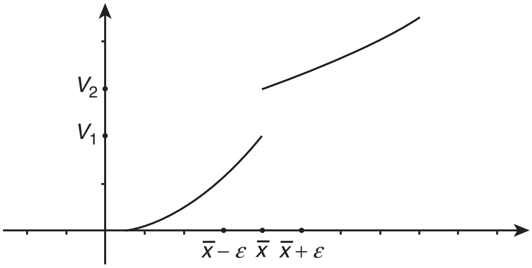


Fig. 2.7 Potential with finite discontinuity

Solution:

$$-\frac{\hbar^2}{2m}\frac{d^2u_E}{dx^2}+V(x)u_E(x)=E\,u_E(x)$$

$$\frac{d^2u_E}{dx^2}=\frac{2m}{\hbar^2}V(x)u_E(x)-\frac{2m}{\hbar^2}E\,u_E(x)$$

Integrating between the limits $\bar{x}-\varepsilon$ to $\bar{x}+\varepsilon$, we get

$$\int_{\bar{x}-\varepsilon}^{\bar{x}+\varepsilon}\frac{d^2u_E}{dx^2}dx=\frac{2m}{\hbar^2}\int_{\bar{x}-\varepsilon}^{\bar{x}+\varepsilon}V(x)u_E(x)dx-\frac{2mE}{\hbar^2}\int_{\bar{x}-\varepsilon}^{\bar{x}+\varepsilon}u_E(x)dx \tag{2.66}$$

Since $u_E(x)$ is continuous at all points, including $x=\bar{x}$, the integral $\int_{\bar{x}-\varepsilon}^{\bar{x}+\varepsilon}u_E(x)dx$ goes to zero in the limit ε going to zero. This can be seen easily as follows. In the range $\bar{x}-\varepsilon<x<\bar{x}+\varepsilon$, we can assume that $u_E(x)$ does not vary much from $u_E(\bar{x})$ and so

$$\int_{\bar{x}-\varepsilon}^{\bar{x}+\varepsilon}u_E(x)dx=\int_{\bar{x}-\varepsilon}^{\bar{x}+\varepsilon}u_E(\bar{x})dx=u_E(\bar{x})\int_{\bar{x}-\varepsilon}^{\bar{x}+\varepsilon}dx=u_E(\bar{x})2\varepsilon$$

Obviously, this integral goes to zero as ε goes to zero. Then, Equation (2.66) becomes

$$\left.\frac{du_E}{dx}\right|_{\bar{x}-\varepsilon}^{\bar{x}+\varepsilon}=\frac{2m}{\hbar^2}u_E(\bar{x})\int_{\bar{x}-\varepsilon}^{\bar{x}+\varepsilon}V(x)dx=\frac{2m}{\hbar^2}u_E(\bar{x})\left(\int_{\bar{x}-\varepsilon}^{\bar{x}}V(x)dx+\int_{\bar{x}}^{\bar{x}+\varepsilon}V(x)dx\right)$$

$$=\frac{2m}{\hbar^2}u_E(\bar{x})\left[V_1\varepsilon+V_2\varepsilon\right]=\frac{2m}{\hbar^2}u_E(\bar{x})\left(V_2+V_1\right)\varepsilon$$

$$\left.\frac{du_E}{dx}\right|_{\bar{x}-\varepsilon}^{\bar{x}+\varepsilon}\to 0\quad\text{as}\quad\varepsilon\to 0$$

Therefore $\frac{du}{dx}$ is continuous at $\bar{x}$

As $\varepsilon\to 0$, RHS tends to be zero.

Note: When V_1 or V_2 becomes infinite, $\frac{du_E}{dx}$ is not continuous at $\bar{x}$.

Example 2.4 Is $\psi(x)=Ne^{-\alpha x}$ an allowed wave function in quantum mechanics?

Solution: We have to study the behaviour of this function for all values of x ranging from $-\infty$ to ∞. The function is well behaved for all values of x. However, it becomes infinite at $x=-\infty$. Therefore, this is not an allowed wave function.

Example 2.5 Normalize the wave function $\psi(x)=Ne^{-\alpha|x|}$.

The wave function is sketched in Fig. 2.8.

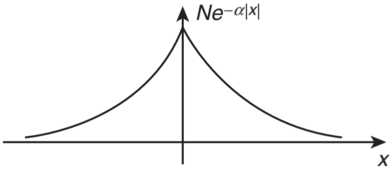


Fig. 2.8

Solution: The wave function $\psi(x) = Ne^{-\alpha|x|}$ can be written in a different but equivalent form as follows:

$$\psi(x) = Ne^{-\alpha|x|} \quad -\infty < x < \infty$$
$$= \begin{cases} Ne^{\alpha x} & -\infty < x < 0 \\ Ne^{-\alpha x} & 0 < x < \infty \end{cases}$$

The second form has the advantage that it can be integrated easily.

$$\int_{-\infty}^{\infty} N^2 e^{-2\alpha|x|} = N^2 \int_{-\infty}^{0} e^{2\alpha x} dx + N^2 \int_{0}^{\infty} e^{-2\alpha x} dx = \frac{N^2}{\alpha} = 1$$
$$\therefore N = \sqrt{\alpha}$$
$$\therefore \psi(x) = \sqrt{\alpha} e^{-\alpha|x|}$$

Example 2.6 Is $\psi(x) = e^{-\alpha|x|}$ an allowed wave function in quantum mechanics?

Solution: We have to check whether the wave function $\psi(x) = e^{-\alpha|x|}$ satisfies the boundary conditions given in (2.63) to (2.65).

The wave function $\psi(x) = e^{-\alpha|x|}$ is single valued and finite for all values of x. So, there is no problem with the behaviour of the $\psi(x)$. Next, let us study the behaviour of the derivative $\partial\psi/\partial x$.

$$\psi(x) = e^{-\alpha|x|} \quad -\infty < x < \infty$$
$$= \begin{cases} e^{\alpha x} & -\infty < x < 0 \\ e^{-\alpha x} & 0 < x < \infty \end{cases}$$

$$\frac{\partial\psi(x)}{\partial x} = \frac{\partial}{\partial x}\left(e^{-\alpha|x|}\right)$$
$$= \begin{cases} \alpha e^{\alpha x} & -\infty < x < 0 \\ -\alpha e^{-\alpha x} & 0 < x < \infty \end{cases}$$

The sketch $\dfrac{\partial\psi}{\partial x}$ is given in Fig. 2.9.

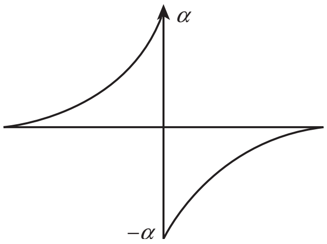


Fig. 2.9

The derivative $\partial\psi/\partial x$ is discontinuous at $x = 0$.

Obviously, $\psi(x) = e^{-\alpha|x|}$ is not an allowed wave function in quantum mechanics.

Example 2.7 Normalize $\psi(x) = e^{-\alpha|x|}\sin\beta x$.

Solution: The wave function $\psi(x)$ is given by

$$\begin{aligned}\psi(x) &= Ne^{-\alpha|x|}\sin\beta x \quad -\infty < x < \infty \\ &= \begin{cases} Ne^{\alpha x}\sin\beta x & -\infty < x < 0 \\ Ne^{-\alpha x}\sin\beta x & 0 < x < \infty \end{cases}\end{aligned}$$

$$\begin{aligned}\therefore \int_{-\infty}^{\infty}\psi^*(x)\psi(x)dx &= N^2\int_{-\infty}^{\infty} e^{-2\alpha|x|}\sin^2\beta x = N^2\left[\int_{-\infty}^{0} e^{2\alpha x}\sin^2\beta x + \int_0^{\infty} e^{-2\alpha x}\sin^2\beta x\right] \\ &= 2N^2\int_0^{\infty} e^{-2\alpha x}\sin^2\beta x\end{aligned}$$

We have changed the variable from x to $-x$ in the first integral to get the second integral.

$$\int_{-\infty}^{0} e^{2\alpha x}\sin^2\beta x\,dx = \int_{\infty}^{0} e^{-2\alpha x}\sin^2(-\beta x)(-dx) = \int_0^{\infty} e^{-2\alpha x}\sin^2\beta x\,dx$$

$$\begin{aligned}\therefore \int_{-\infty}^{\infty}\psi^*(x)\psi(x)dx &= 2N^2\frac{1}{2}\int_0^{\infty} e^{-2\alpha x}[1-\cos 2\beta x]dx = N^2\left\{\frac{1}{2\alpha} - \frac{2\alpha}{4\alpha^2+4\beta^2}\right\} \\ &= \frac{N^2}{2\alpha}\left[1 - \frac{1}{1+\beta^2/\alpha^2}\right]\end{aligned}$$

Example 2.8 Is $\psi(x) = e^{-\alpha|x|}\sin\beta x$ an admissible wave function in quantum mechanics? The sketch of $\psi(x)$ is shown in Fig. 2.10.

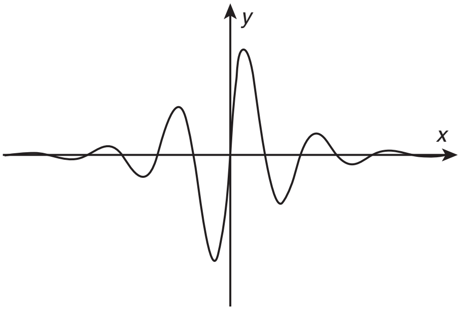


Fig. 2.10

Solution: The wave function $\psi(x)$ is given by

$$\begin{aligned}\psi(x) &= Ne^{-\alpha|x|}\sin\beta x \quad -\infty < x < \infty \\ &= \begin{cases} Ne^{\alpha x}\sin\beta x & -\infty < x < 0 \\ Ne^{-\alpha x}\sin\beta x & 0 < x < \infty \end{cases}\end{aligned}$$

The derivative $\dfrac{\partial\psi}{\partial x}$ is given by

$$
\begin{aligned}
\frac{\partial\psi}{\partial x} &= N\frac{\partial}{\partial x}[e^{-\alpha|x|}\sin\beta x] && -\infty < x < \infty \\
&= \begin{cases} N\dfrac{\partial}{\partial x}(e^{\alpha x}\sin\beta x) & -\infty < x < 0 \\ N\dfrac{\partial}{\partial x}(e^{-\alpha x}\sin\beta x) & 0 < x < \infty \end{cases} \\
&= N\begin{cases} \alpha e^{\alpha x}\sin\beta x + \beta e^{\alpha x}\cos\beta x & -\infty < x < 0 \\ -\alpha e^{-\alpha x}\sin\beta x + \beta e^{-\alpha x}\cos\beta x & 0 < x < \infty \end{cases}
\end{aligned}
$$

Both the functions $e^{\alpha x}\sin\beta x$ and $e^{-\alpha x}\sin\beta x$ and their derivatives are well-behaved and continuous functions. However, $\psi(x) = Ne^{-\alpha|x|}\sin\beta x$ changes from $e^{\alpha x}\sin\beta x$ to $e^{-\alpha x}\sin\beta x$ at $x = 0$. So, one has to verify the behaviour of $\psi(x)$ and $\partial\psi/\partial x$ only at $x = 0$.

At $x = 0$, $e^{\alpha x}\sin\beta x = e^{-\alpha x}\sin\beta x = 0$. This can also be seen from the sketch of $\psi(x)$ given in Fig. 2.10. So, the wave function $\psi(x)$ is single valued, finite, and continuous for the entire range $-\infty < x < \infty$ including $x = 0$.

At $x = 0$, $\dfrac{\partial\psi}{\partial x} \to \beta$ from both left and right.

$$
\begin{aligned}
&\big(\alpha e^{\alpha x}\sin\beta x + \beta e^{\alpha x}\cos\beta x \xrightarrow{x\to 0} \beta \\
&-\alpha e^{-\alpha x}\sin\beta x + \beta e^{-\alpha x}\cos\beta x \xrightarrow{x\to 0} \beta\big)
\end{aligned}
$$

Therefore, $\dfrac{\partial\psi}{\partial x}$ is finite, single valued, continuous in the entire range $-\infty < x < \infty$ including the point $x = 0$.

Therefore, $\psi(x) = Ne^{-\alpha|x|}\sin\beta x$ is an allowed function in quantum mechanics.

2.12 SUPERPOSITION OF ENERGY EIGENSTATES – TIME EVOLUTION

We can construct a new state $\phi(\boldsymbol{r})$ which is a superposition of the energy eigenstates of time-independent Hamiltonian. The wave function $\phi(\boldsymbol{r})$ is given by

$$\phi(\boldsymbol{r}) = \sum_n c_n u_n(\boldsymbol{r})$$

This state is not an energy eigen function of time-independent Hamiltonian. It is interesting to study the time evolution for the state $\psi(\boldsymbol{r}, t)$ under time-independent Hamiltonian for a system whose initial state is $\phi(\boldsymbol{r})$. In other words we have

$$\psi(\boldsymbol{r},0) = \phi(\boldsymbol{r}) = \sum_n c_n u_n(\boldsymbol{r})$$

For any system, $\psi(\boldsymbol{r}, t)$ is determined from the Schrödinger equation

$$i\hbar\frac{\partial\psi(\boldsymbol{r},t)}{\partial t} = H\psi(\boldsymbol{r},t)$$

In the case of time-independent of Hamiltonian, the solution to the above equation can be formally written as

$$\psi(\boldsymbol{r},t) = e^{-iHt/\hbar}\psi(\boldsymbol{r},0) = e^{-iHt/\hbar}\sum_n c_n u_n(\boldsymbol{r}) = \sum_n c_n e^{-iHt/\hbar}u_n(\boldsymbol{r})$$

$$\begin{aligned} e^{-iHt/\hbar}u_n(\boldsymbol{r}) &= \left(1+\left(-\frac{it}{\hbar}\right)H+\left(-\frac{it}{\hbar}\right)^2 H^2+\left(-\frac{it}{\hbar}\right)^3 H^3+\cdots\right)u_n(\boldsymbol{r}) \\ &= \left(1+\left(-\frac{it}{\hbar}\right)E_n+\left(-\frac{it}{\hbar}\right)^2 E_n{}^2+\left(-\frac{it}{\hbar}\right)^3 E_n{}^3+\cdots\right)u_n(\boldsymbol{r}) \\ &= e^{-iE_nt/\hbar}u_n(r) \\ \therefore \quad \psi(\boldsymbol{r},t) &= \sum_n c_n e^{-iE_nt/\hbar}u_n(\boldsymbol{r}) \end{aligned}$$

Neither $\psi(r, 0)$ nor $\psi(r, t)$ is an energy eigen function of time-independent Hamiltonian.

Note that in the case of continuous energy eigen values, we get

$$\psi(\boldsymbol{r},0) = \phi(\boldsymbol{r}) = \int dE\, C(E)\, u_E(\boldsymbol{r})$$

$$\psi(\boldsymbol{r},t) = \int dE\, C(E)\, e^{-iEt/\hbar}u_E(\boldsymbol{r})$$

2.13 MOMENTUM OPERATOR AND ITS EIGENFUNCTIONS

We have discussed many roles played by plane wave functions in quantum mechanics. They are the wave functions describing the free particles. Now, let us see yet another important feature of plane wave functions in quantum mechanics. This time, it plays the role of eigenfunction of the momentum operator. Let us briefly look at the role played by operators and their eigenfunctions in quantum mechanics.

As we have noted in Section 2.3, the description of dynamical variables in classical mechanics and quantum mechanics are different. In classical mechanics, a dynamical variable is a function of position and momentum. They are either simply some numbers or vector quantities. In quantum mechanics, the concept of a dynamical variable is more involved and goes through many new theoretical ideas. Let us mention them briefly here.

1. For each dynamical variable A, we associate an operator A_{op}.
2. Determine the eigenvalues and the eigenfunctions of the operator A_{op}.

   $$A_{op}\, \phi_i = a_i\, \phi_i$$

   The list of all possible eigenvalues are $\{a_1, a_2, a_3, \ldots\}$.
3. The results of measurement of a dynamical variable in any experiment will be one among the list of possible eigenvalues $a_1, a_2\ldots$.

This brief account is presented here for the sake of completeness. A full-fledged discussion on the dynamical variables is presented in Chapter 5.

To discuss the normalization of momentum eigenfunctions, we need one more idea, namely Dirac delta function $\delta(x - x')$.

The Dirac delta function $\delta(x - x')$ is defined as

$$\delta(x-x') = \begin{cases} \infty & \text{at } x = x' \\ 0 & \text{if } x \neq x' \end{cases}$$

such that

$$\int_{-\infty}^{\infty} \delta(x-x')dx = 1 \tag{2.67}$$

This is not a mathematical function in the conventional sense of defining a function. The Dirac delta function and its properties are discussed extensively in Chapter 4. For our purpose, it is sufficient to know the following integral:

$$\int_{-\infty}^{\infty} e^{ikx}\,dk = 2\pi\delta(x) \tag{2.68}$$

2.13.1 Determination of Momentum Eigenvalues

The eigenvalue equation for momentum is

$$p_{op}\phi_p(x) = p\,\phi_p(x) \tag{2.69}$$

Here, p is the momentum eigenvalue and $\phi_p(x)$ is the eigenfunction corresponding to the eigenvalue p. Since the momentum operator in one dimension is $-i\hbar d/dx$, the eigenvalue equation is

$$-i\hbar\frac{d\phi_p(x)}{dx} = p\;\phi_p(x)$$

$$\frac{d\phi_p(x)}{dx} = \frac{ip}{\hbar}\;\phi_p(x) \tag{2.70}$$

This is a linear differential equation, and its solution is given by

$$\phi_p(x) = Ne^{ipx/\hbar} \tag{2.71}$$

In three dimension, the momentum eigenvalue equation is

$$\mathbf{p}_{op}\;\phi_{\mathbf{p}}(\mathbf{r}) = \mathbf{p}\phi_{\mathbf{p}}(\mathbf{r})$$

or

$$-i\hbar\nabla\phi_{\mathbf{p}}(\mathbf{r}) = \mathbf{p}\phi_{\mathbf{p}}(\mathbf{r})$$

The solution to this equation is

$$\phi_{\mathbf{p}}(\mathbf{r}) = Ne^{i\,\mathbf{p}\cdot\mathbf{r}/\hbar} \tag{2.72}$$

In solving the momentum eigenvalue equation, we did not get any condition that restricts it to some select values. Therefore, the eigenvalues p (or $\mathbf{p}$) can take any value from $-\infty$ to ∞.
Therefore, range of possible values of p_x, p_y and p_z are

$$-\infty < p_x < \infty,\ -\infty < p_y < \infty,\ -\infty < p_z < \infty \tag{2.73}$$

The Wave Functions $e^{i\mathbf{p}\cdot\mathbf{r}/\hbar}$ and $e^{i(\mathbf{p}\cdot\mathbf{r}-Et)/\hbar}$

Note that the function $\phi_{\mathbf{p}}(\mathbf{r}) = e^{i\mathbf{p}\cdot\mathbf{r}/\hbar}$ is the spatial part of the function $\psi_{\mathbf{p}}(\mathbf{r},t) = e^{i(\mathbf{p}\cdot\mathbf{r}-Et)/\hbar} = e^{i\mathbf{p}\cdot\mathbf{r}/\hbar}\,e^{-iEt/\hbar}$. We will come across these functions repeatedly in various contexts. These functions occur so frequently in quantum mechanics that it is better to summarize their important features.

1. $\phi_{\mathbf{p}}(\mathbf{r})$ is a free-particle solution of time-independent Schrödinger equation and $\psi_{\mathbf{p}}(\mathbf{r},t)$ is a free-particle solution for time-dependent Schrödinger equation

$$-\frac{\hbar^2}{2m}\nabla^2\phi_{\mathbf{p}}(\mathbf{r}) = \frac{p^2}{2m}\phi_{\mathbf{p}}(\mathbf{r})$$

$$i\hbar\frac{\partial\psi_{\mathbf{p}}(\mathbf{r},t)}{\partial t} = -\frac{\hbar^2}{2m}\nabla^2\psi_{\mathbf{p}}(\mathbf{r},t)$$

2. Both $\phi_{\mathbf{p}}(\mathbf{r}) = Ne^{i\,\mathbf{p}\cdot\mathbf{r}/\hbar}$ and $\psi_{\mathbf{p}}(\mathbf{r},t) = Ne^{i(\mathbf{p}\cdot\mathbf{r}-Et)/\hbar}$ are eigenfunctions of momentum and Hamiltonian operator.

$$\mathbf{p}_{op}\ \phi_{\mathbf{p}}(\mathbf{r}) = \mathbf{p}\,\phi_{\mathbf{p}}(\mathbf{r}) \qquad H_{op}\phi_{\mathbf{p}}(\mathbf{r}) = \frac{p^2}{2m}\phi_{\mathbf{p}}(\mathbf{r})$$

$$\mathbf{p}_{op}\ \psi_{\mathbf{p}}(\mathbf{r}) = \mathbf{p}\,\psi_{\mathbf{p}}(\mathbf{r}) \qquad H_{op}\psi_{\mathbf{p}}(\mathbf{r}) = \frac{p^2}{2m}\psi_{\mathbf{p}}(\mathbf{r})$$

3. Both $\phi_{\mathbf{p}}(\mathbf{r})$ and $\psi_{\mathbf{p}}(\mathbf{r},t)$ are non–integrable. That is,

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} dxdydz\ \phi^*_{\mathbf{p}}(\mathbf{r})\phi_{\mathbf{p}}(\mathbf{r})d^3\mathbf{r} \to \infty$$

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} dxdydz\ \psi^*_{\mathbf{p}}(\mathbf{r},t)\psi_{\mathbf{p}}(\mathbf{r},t)d^3\mathbf{r} \to \infty$$

4. Let us add one more observation. Both $\phi_{\mathbf{p}}(\mathbf{r})$ and $\psi_{\mathbf{p}}(\mathbf{r},t)$ satisfy not only Schrödinger equation but also a number of other equations such as Klien Gordon equation in relativistic quantum mechanics.

2.14 NORMALIZATION OF MOMENTUM EIGENFUNCTION

We have already seen that the plane wave function is not normalisable through the normalization condition $\int \phi^*_{\mathbf{p}}(\mathbf{r})\phi_{\mathbf{p}}(\mathbf{r})d^3\mathbf{r} = 1$. We cannot discard these functions from quantum mechanics. They are the solutions to free-particle Schrödinger equation, and so they describe the state of a free particle. They also happen to be the eigenfunctions of the momentum operator that makes it all the more impossible to do away with plane wave functions from quantum mechanics.

Let us recall that the normalization condition $\int \psi^*(\mathbf{r},t)\psi(\mathbf{r},t)d^3\mathbf{r} = 1$ or $\int \phi^*_{\mathbf{p}}(\mathbf{r})\phi_{\mathbf{p}}(\mathbf{r}) = 1$ is the result of probabilistic interpretation of the wave function. If we cannot employ this normalization condition, can we employ different kind of normalization conditions which will make these wave functions to serve some other purposes in quantum mechanics?

Within quantum mechanics, we employ two methods of normalizations for momentum eigenfunctions. They are called box normalization and delta function normalization.

2.14.1 Box Normalization

To normalize the momentum eigenfunctions, a large volume, preferably a cube of side *L,* is considered. The integration limits are restricted to the size of this box. In addition, we demand the momentum eigenfunction to obey the following boundary condition:

$$\psi(x+L,y,z) = \psi(x,y+L,z) = \psi(x,y,z+L) = \psi(x,y,z) \tag{2.74}$$

This boundary condition is called periodic boundary condition. The normalization is now done through the requirement.

$$\int_{\substack{\text{volume}\\ \text{of the box}}} \phi^*_{\mathbf{p}}(\mathbf{r})\phi_{\mathbf{p}}(\mathbf{r})d^3\mathbf{r} = 1$$

$$\text{i.e.}\quad \int_0^L dx\int_0^L dy\int_0^L dz\ \phi^*_{\mathbf{p}}(\mathbf{r})\phi_{\mathbf{p}}(\mathbf{r})d^3\mathbf{r} = 1 \tag{2.75}$$

What is the consequence of the periodic boundary condition?
Consider the boundary condition

$$\psi(x+L, y, z) = \psi(x, y, z)$$

$$Ne^{i\,p_x(x+L)/\hbar} e^{i\,p_y y/\hbar} e^{i\,p_z z/\hbar} = Ne^{i\,p_x x/\hbar} e^{i\,p_y y/\hbar} e^{i\,p_z z/\hbar}$$

This implies $e^{i\,p_x L/\hbar} = 1$

or $$\frac{p_x L}{\hbar} = 2l\pi, \quad \text{where } l = 0, \pm 1, \pm 2 \ldots$$

$$\therefore p_x = \frac{2l\pi\hbar}{L} = 0, \pm\frac{2\pi\hbar}{L}, \pm\frac{4\pi\hbar}{L}, \pm\frac{6\pi\hbar}{L}, \ldots$$

In the same way, we get

$$p_y = 0, \pm\frac{2\pi\hbar}{L}, \pm\frac{4\pi\hbar}{L}, \pm\frac{6\pi\hbar}{L}, \ldots$$

and $$p_z = 0, \pm\frac{2\pi\hbar}{L}, \pm\frac{4\pi\hbar}{L}, \pm\frac{6\pi\hbar}{L}, \ldots$$

The significant point is that the momentum $\mathbf{p}$ is no longer a continuous variable. It is a set of discrete values.

$$\therefore \mathbf{p} = \frac{2\pi}{L}(l\hat{x} + m\hat{y} + n\hat{z}) \tag{2.76}$$

where l, m, n are set of integers.

(i.e.) l, m, and n: 0, ± 1, ± 2, ± 3,…

The normalization is now done by

$$\int\limits_{\substack{\text{box}\\ \text{volume}}} \phi_{\mathbf{p}}^{*}(\mathbf{r})\phi_{\mathbf{p}}(\mathbf{r})d^3\mathbf{r} = \int_0^L dx \int_0^L dy \int_0^L dz \;\; \phi_{\mathbf{p}}^{*}(\mathbf{r})\phi_{\mathbf{p}}(\mathbf{r})d = 1$$

$$\int_0^L dx \int_0^L dy \int_0^L dz \;\; N^2 e^{-i\mathbf{p}\cdot\mathbf{r}/\hbar} e^{i\mathbf{p}\cdot\mathbf{r}/\hbar} = N^2 L^3 = 1$$

$$\therefore N = \frac{1}{L^{3/2}}$$

The normalized momentum wave function in a box normalization is

$$\phi_{\mathbf{p}}(\mathbf{r}) = \frac{1}{L^{3/2}} e^{i\mathbf{p}.\mathbf{r}/\hbar} = \frac{1}{V^{1/2}} e^{i\mathbf{p}.\mathbf{r}/\hbar} \tag{2.77}$$

where V is the volume of the box, $V = L^3$.
The normalized plane wave function is $\psi_{\mathbf{p}}(\mathbf{r},t) = \frac{1}{L^{3/2}} e^{i(\mathbf{p}.\mathbf{r}-Et)/\hbar} = \frac{1}{V^{1/2}} e^{i(\mathbf{p}.\mathbf{r}-Et)/\hbar}$

Example 2.9 Show that $\int \phi_{\mathbf{p}}^{*}(\mathbf{r})\phi_{\mathbf{p}'}(\mathbf{r})d^3\mathbf{r} = \delta_{\mathbf{pp}'}$ where $\delta_{\mathbf{pp}'}$ is given by $\delta_{\mathbf{pp}'} = \begin{cases} 1 & \text{if } \mathbf{p} = \mathbf{p}' \\ 0 & \text{if } \mathbf{p} \neq \mathbf{p}'. \end{cases}$

Solution: Let $\mathbf{p}$ and $\mathbf{p}'$ be given by

$$\mathbf{p} = \frac{2\pi\hbar}{L}(l\hat{x} + m\hat{y} + n\hat{z})$$

and
$$\mathbf{p}' = \frac{2\pi\hbar}{L}(l'\hat{x} + m'\hat{y} + n'\hat{z})$$

($\mathbf{p} = \mathbf{p}'$ means $l = l'$, $m = m'$ and $n = n'$).

$$\int \phi^*_{\mathbf{p}}(\mathbf{r})\,\phi_{\mathbf{p}'}(\mathbf{r})d^3\mathbf{r} = \frac{1}{L^3}\int e^{-i\mathbf{p}\cdot\mathbf{r}/\hbar}e^{i\mathbf{p}'\cdot\mathbf{r}/\hbar}d^3\mathbf{r} = \frac{1}{L^3}\int_0^L dx \int_0^L dy \int_0^L dz\; e^{i2\pi(l'-l)x/L}\cdot e^{i2\pi(m'-m)y/L}\cdot e^{i2\pi(n'-n)z/L}$$

Consider the x–integral

$$\begin{aligned}\frac{1}{L}\int_0^L dx\; e^{i2\pi(l'-l)x/L} &= \left.\frac{e^{i2\pi(l'-l)x/L}}{2\pi i(l'-l)}\right|_0^L \text{ provided } l' \neq l.\\ &= 0\end{aligned}$$

If $l' = l$, the above integral is $\frac{1}{L}\int_0^L dx\; e^{i2\pi(l-l)x/L} = 1$

$$\therefore \frac{1}{L}\int_0^L dx\; e^{i2\pi(l'-l)x/L} = \delta_{l'l} = \begin{cases}1 & \text{if } l = l'\\ 0 & \text{if } l \neq l'\end{cases}$$

Similarly, we can prove that

$$\frac{1}{L}\int_0^L dy\; e^{i2\pi(m'-m)y/L} = \delta_{m'm} \quad\text{and}\quad \frac{1}{L}\int_0^L dz\; e^{i2\pi(n'-n)z/L} = \delta_{n'n}$$

$$\begin{aligned}\therefore \int \phi^*_{\mathbf{p}}(\mathbf{r})\phi_{\mathbf{p}'}(\mathbf{r})d^3\mathbf{r} &= \delta_{ll'}\cdot\delta_{mm'}\cdot\delta_{nn'}\\ &= \delta_{\mathbf{pp}'}\end{aligned} \tag{2.78}$$

2.14.2 Delta Function Normalization

The momentum eigenfunction is now given by

$$\begin{aligned}\phi_{\mathbf{p}}(\mathbf{r}) &= Ne^{i\mathbf{p}\cdot\mathbf{r}/\hbar}\\ &= Ne^{ip_x x/\hbar}e^{ip_y y/\hbar}e^{ip_z z/\hbar} \qquad -\infty < p_x < \infty,\ -\infty < p_y < \infty,\ -\infty < p_z < \infty\end{aligned}$$

This is in contrast with box normalization where p_x, p_y, and p_z are discrete. There is no necessity for considering an arbitrary box of volume L^3 in this procedure. The delta function normalization must satisfy the condition

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \phi^*_{\mathbf{p}}(\mathbf{r})\phi_{\mathbf{p}'}(\mathbf{r})d^3\mathbf{r} = \delta(\mathbf{p} - \mathbf{p}') \tag{2.79}$$

The reason for this choice of normalization will be explained in Chapter 5.

$$\int \phi^*_{\mathbf{p}}(\mathbf{r})\,\phi_{\mathbf{p}'}(\mathbf{r})d^3\mathbf{r} = \int N^2 e^{-i\mathbf{p}\cdot\mathbf{r}/\hbar} e^{i\mathbf{p}\cdot\mathbf{r}/\hbar} d^3\mathbf{r}$$

$$= N^2 \int_{-\infty}^{\infty} e^{i(p'_x - p_x)x/\hbar} dx \int_{-\infty}^{\infty} e^{i(p'_y - p_y)y/\hbar} dy \int_{-\infty}^{\infty} e^{i(p'_z - p_z)z/\hbar} dz$$

Consider the *x*–integral.

Put
$$x' = \frac{x}{\hbar}$$

$$\therefore\ dx' = \frac{dx}{\hbar}$$

$$\therefore \int_{-\infty}^{\infty} e^{i(p'_x - p_x)\frac{x}{\hbar}} dx = \int_{-\infty}^{\infty} e^{i(p'_x - p_x)dx'} \cdot \hbar\, dx' = 2\pi\hbar\ \delta(p'_x - p_x)$$

In the same way we can show that

$$\therefore \int_{-\infty}^{\infty} e^{i(p'_y - p_y)y/\hbar} dy = 2\pi\hbar\ \delta(p'_y - p_y); \qquad \int_{-\infty}^{\infty} e^{i(p'_z - p_z)z/\hbar} dz = 2\pi\hbar\ \delta(p'_z - p_z).$$

$$\therefore \int \phi^*_{\mathbf{p}}(\mathbf{r})\phi_{\mathbf{p}'}(\mathbf{r})d^3\mathbf{r} = N^2(2\pi\hbar)^3\delta(p'_x - p_x)\delta(p'_y - p_y)\delta(p'_z - p_z)$$

$$= N^2(2\pi\hbar)^3\delta(\mathbf{p}' - \mathbf{p})$$

where $\delta(\mathbf{p}' - \mathbf{p}) = \delta(p'_x - p_x)\delta(p'_y - p_y)\delta(p'_z - p_z)$

$$\therefore N^2(2\pi\hbar)^3\delta(\mathbf{p}' - \mathbf{p}) = \delta(\mathbf{p}' - \mathbf{p})$$

$$\therefore N = \frac{1}{(2\pi\hbar)^{3/2}}$$

The normalized momentum eigenfunction is

$$\phi_{\mathbf{p}}(\mathbf{r}) = \frac{1}{(2\pi\hbar)^{3/2}} e^{i\mathbf{p}\cdot\mathbf{r}/\hbar} \tag{2.80}$$

Example 2.10 Normalize the plane wave function using delta function normalization.

Solution: The normalization of plane wave function is now done by demanding

$$\int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy \int_{-\infty}^{\infty} dz\ \psi^*_{\vec{p}}(\mathbf{r},t)\psi_{\vec{p}}(\mathbf{r},t)d^3\mathbf{r} = \delta(\mathbf{p} - \mathbf{p}')$$

$$\text{(i. e.)} \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy \int_{-\infty}^{\infty} dz\ N^2 e^{-i(\mathbf{p}\cdot\mathbf{r} - E_{\mathbf{p}}t)/\hbar} e^{i(\mathbf{p}\cdot\mathbf{r} - E_{\mathbf{p}}t)/\hbar}$$

$$= N^2 \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy \int_{-\infty}^{\infty} dz\; e^{-i\mathbf{p}\cdot\mathbf{r}/\hbar} e^{i\mathbf{p}'\cdot\mathbf{r}/\hbar} e^{iE_{\mathbf{p}}t/\hbar} e^{-iE_{\mathbf{p}'}t/\hbar}$$

$$= N^2 e^{iE_{\mathbf{p}}t/\hbar} e^{-iE_{\mathbf{p}'}t/\hbar} (2\pi)^3 \delta(\mathbf{p}-\mathbf{p}')$$

$$= N^2 (2\pi)^3 \delta(\mathbf{p}-\mathbf{p}')$$

$$\therefore \psi_{\mathbf{p}}(\mathbf{r},t) = \frac{1}{(2\pi\hbar)^{3/2}} e^{i(\mathbf{p}\cdot\mathbf{r}-E_{\mathbf{p}}t)/\hbar}$$

2.14.3 Purpose of These Normalizations

What is the purpose of these two normalizations? Why do we implement the periodic boundary conditions in box normalization? The answer requires the understanding of concept of 'complete set of orthonormal set of functions' which will be discussed in Chapters 4 and 5. The set of all the eigenfunctions of a dynamical variable should form a complete set. This is one of the requirements in the structure of quantum mechanics. The concept of complete set of basis functions is a very important idea in quantum mechanics.

In the case of box normalization, restricting the limits of integration to a finite volume of the box is not sufficient to ensure the completeness of the basis functions. The periodic boundary conditions ensures that the momentum eigenfunctions form a complete orthonormal set, which means that any arbitrary wave function $\phi(\mathbf{r})$ or $\psi(\mathbf{r},t)$ can be written as

$$\phi(\mathbf{r}) = \sum_{\mathbf{p}} C_{\mathbf{p}}\, e^{i\mathbf{p}\cdot\mathbf{r}/\hbar}$$

or

$$\psi(\mathbf{r},t) = \sum_{\mathbf{p}} C_{\mathbf{p}} e^{i(\mathbf{p}\cdot\mathbf{r}-E_{\mathbf{p}}t)/\hbar} \tag{2.81}$$

Momentum eigenfunctions normalized through delta function normalization also form a complete set, which again implies that any arbitrary function $\psi(\mathbf{r})$ or $\psi(\mathbf{r},t)$ can be written as

$$\phi(\mathbf{r}) = \int C(\mathbf{p}) \frac{1}{(2\pi\hbar)^{3/2}} e^{i\mathbf{p}\cdot\mathbf{r}/\hbar} d^3\mathbf{p}$$

or

$$\psi(\mathbf{r},t) = \int C(\mathbf{p},t) \frac{1}{(2\pi\hbar)^{3/2}} e^{i\mathbf{p}.\mathbf{r}/\hbar} d^3\mathbf{p} \tag{2.82}$$

Note the difference between (2.81) and (2.82). In both the cases, entire set of eigenfunctions are used. However, in (2.81) the summation is over all the discrete values of **p**. In (2.82), since **p** is a continuous, the summation over **p** becomes integration over **p**.

Another Method of Normalization

Sometimes neither the box normalization nor the delta function normalization is used. This is due to the fact that our interest is not the expansion of an arbitrary function in terms of the basis functions.

For instance, in the scattering problems a better choice of normalization, that is, condition is $|N|^2\ p/m = F$ where F is the flux of incident number of particles.

2.15 COORDINATE AND MOMENTUM REPRESENTATION FOR A SYSTEM

Let us consider Equation (2.82) in the context of the Schrödinger. In the second integral for $\psi(\mathbf{r},t)$ the function $C(\mathbf{p},t)$ has to be found using Schrödinger equation and so it depends on the potential $V(\mathbf{r},t)$. However, for free particles, $C(\mathbf{p},t)$ can be easily worked out and it is given by

$$C(\mathbf{p},t) = C(\mathbf{p})e^{-iEt/\hbar} \tag{2.83}$$

Example 2.11 Prove that $C(\mathbf{p},t) = C(\mathbf{p})e^{-iEt/\hbar}$ for free particles, in the expansion

$$\psi(\mathbf{r},t) = \frac{1}{(2\pi\hbar)^{3/2}}\int C(\mathbf{p},t)e^{i\mathbf{p}\cdot\mathbf{r}/\hbar}d^3\mathbf{p}$$

Solution: The Schrödinger equation for free particles is given by

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

But
$$i\hbar\frac{\partial\psi(\mathbf{r},t)}{\partial t} = \frac{1}{(2\pi\hbar)^{3/2}}\int i\hbar\frac{d}{dt}C(\mathbf{p},t)e^{i\mathbf{p}\cdot\mathbf{r}/\hbar}d^3\mathbf{p}$$

and
$$\begin{aligned}\frac{-\hbar^2}{2m}\nabla^2\psi(\mathbf{r},t) &= -\frac{\hbar^2}{2m}\cdot\frac{1}{(2\pi\hbar)^{3/2}}\cdot\nabla^2\int C(\mathbf{p},t)e^{i\mathbf{p}\cdot\mathbf{r}/\hbar}d^3\mathbf{p}\\ &= \frac{1}{(2\pi\hbar)^{3/2}}\int\left(-\frac{\hbar^2}{2m}\right)\left(-\frac{p^2}{\hbar^2}\right)C(\mathbf{p},t)e^{i\mathbf{p}\cdot\mathbf{r}/\hbar}d^3\mathbf{p}\\ &= \frac{1}{(2\pi\hbar)^{3/2}}\int E\ \ C(\mathbf{p},t)e^{i\mathbf{p}\cdot\mathbf{r}/\hbar}d^3\mathbf{p}\end{aligned}$$

Equating these two terms, we get

$$\frac{1}{(2\pi\hbar)^{3/2}}\int i\hbar\frac{dC(\mathbf{p},t)}{dt}e^{i\mathbf{p}\cdot\mathbf{r}/\hbar}d^3\mathbf{p} = \frac{1}{(2\pi\,\hbar)^{3/2}}\int E\ \ C(\mathbf{p},t)e^{i\mathbf{p}\cdot\mathbf{r}/\hbar}d^3\mathbf{p}$$

$$\therefore i\hbar\frac{dC(\mathbf{p},t)}{dt} = E\ \ C(\mathbf{p},t)$$

Solving this equation, we get

$$C(\mathbf{p},t) = C(\mathbf{p})e^{-iEt/\hbar} \tag{2.84}$$

Therefore, for a free particle,

$$\psi(\mathbf{r},t) = \frac{1}{(2\pi\hbar)^{3/2}}\int C(\mathbf{p})e^{i(\mathbf{p}\cdot\mathbf{r}-Et)/\hbar}d^3\mathbf{p} \tag{2.85}$$

Summarizing the above results, we have for free particles

1. $\phi(\mathbf{r}) = \frac{1}{(2\pi\hbar)^{3/2}}\int C(\mathbf{p})e^{i\mathbf{p}\cdot\mathbf{r}/\hbar}d^3\mathbf{p}$ (2.86)
2. $\psi(\mathbf{r},t) = \frac{1}{(2\pi\hbar)^{3/2}}\int C(\mathbf{p})e^{i(\mathbf{p}\cdot\mathbf{r}-Et)/\hbar}d^3\mathbf{p}$ (2.87)

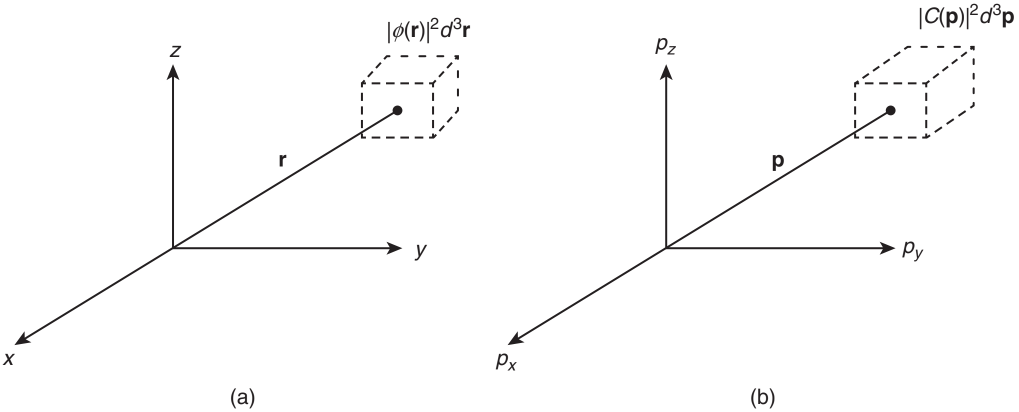
Equation (2.86) is the familiar Fourier transformation equation. $\phi(\mathbf{r})$ and $C(\mathbf{p})$ are Fourier transform pair. In subsequent discussion, we will concentrate on states described by $\phi(\mathbf{r})$ and $C(\mathbf{p})$.

2.15.1 Momentum Space

The wave function $\phi(\mathbf{r})$ is a function of coordinate $\mathbf{r}$. The same fact can be expressed in a different way. The wave function $\phi(\mathbf{r})$ is said to represent the state in the coordinate space. The same thing can be said about $C(\mathbf{p})$. It is a function of $\mathbf{p}$. Now, $C(\mathbf{p})$ is said to represent the same state of the system in momentum space. $C(\mathbf{p})$ is known as momentum wave function of the system.

2.15.2 Interpretation of Momentum Wave Function *C*(p)

The functions $\phi(\mathbf{r})$ and $C(\mathbf{p})$ are the Fourier transform pair. Knowledge of one function determines the other function. That is, $\phi(\mathbf{r})$ determines $C(\mathbf{p})$ and $C(\mathbf{p})$ determines $\phi(\mathbf{r})$. So, if $\phi(\mathbf{r})$ is taken as the wave function, describing the state of the system in coordinate space, then $C(\mathbf{p})$ should also be considered as a function describing the same state in momentum space; The fact that they are Fourier transform pair suggests that $C(\mathbf{p})$ can be interpreted as probability amplitude and $C^*(\mathbf{p})C(\mathbf{p})$ as probability density for momentum. $C^*(\mathbf{p})C(\mathbf{p})d^3\mathbf{p}$ is the probability of finding the momentum value within a volume $d^3\mathbf{p}$ centred around $\mathbf{p}$ as shown in Fig. 2.11.



(a) (b)

Fig. 2.11 Probability density in coordinate space and momentum space

The momentum wave function $C(\mathbf{p})$ should also be normalisable to unity,

(i.e.)
$$\int C^*(\mathbf{p})C(\mathbf{p})d^3\mathbf{p} = 1 \tag{2.88}$$

Hence, $C(\mathbf{p})$ is also a square integrable function. Both $\phi(\mathbf{r})$ and $C(\mathbf{p})$ are probability amplitudes. $\phi(\mathbf{r})$ gives the information on probability of finding the particle at a particular position $\mathbf{r}$ while $C(\mathbf{p})$ gives the information on the probability of finding the particle to have a particular value of momentum $\mathbf{p}$.

Example 2.12 $C(\mathbf{p})$ and $\phi(\mathbf{r})$ are Fourier transform pair. Show that $C(\mathbf{p})$ is square integrable if $\phi(\mathbf{r})$ is square integrable.

Solution: The fact that $\phi(\mathbf{r})$ is square integrable means

$$\int \phi^*(\mathbf{r})\phi(\mathbf{r})d^3\mathbf{r} < \infty$$

Since $\phi(\mathbf{r})$ and $C(\mathbf{p})$ are Fourier transform pair,

$$\phi(\mathbf{r}) = \frac{1}{(2\pi\hbar)^{3/2}}\int C(\mathbf{p})e^{i\mathbf{p}\cdot\mathbf{r}/\hbar}d^3\mathbf{p}$$

$$\therefore \int \phi^*(\mathbf{r})\phi(\mathbf{r})d^3\mathbf{r} = \int\left[\frac{1}{(2\pi\hbar)^{3/2}}\int C^*(\mathbf{p})e^{i\mathbf{p}\cdot\mathbf{r}/\hbar}d^3\mathbf{p}\right]\cdot\left[\frac{1}{(2\pi\hbar)^{3/2}}\int C(\mathbf{p}')e^{-i\mathbf{p}'\cdot\mathbf{r}/\hbar}d^3\mathbf{p}'\right]d^3\mathbf{r}$$

$$= \frac{1}{(2\pi\hbar)^3}\int C^*(\mathbf{p})C(\mathbf{p}')e^{i(\mathbf{p}'-\mathbf{p})\cdot\mathbf{r}/\hbar}d^3\mathbf{p}d^3\mathbf{p}'d^3\mathbf{r}$$

$$= \frac{1}{(2\pi\hbar)^3}\int C^*(\mathbf{p})C(\mathbf{p}')d^3\mathbf{p}d^3\mathbf{p}'\int e^{i(\mathbf{p}'-\mathbf{p})\cdot\mathbf{r}/\hbar}d^3\mathbf{r}$$

$$= \int C^*(\mathbf{p})C(\mathbf{p}')\frac{(2\pi\hbar)^3}{(2\pi\hbar)^3}\delta(\mathbf{p}'-\mathbf{p})d^3\mathbf{p}d^3\mathbf{p}'$$

$$\therefore \int C^*(\mathbf{p})C(\mathbf{p})d^3\mathbf{p} < \infty$$

2.15.3 Widths of Fourier Transform Pair $\phi(\mathbf{r})$ and $C(\mathbf{p})$

The Fourier transform pair has an interesting property. Consider a wave packet $\phi(x)$ and its Fourier transform $C(\mathbf{p})$ for which $|\phi(x)|^2$ and $|C(\mathbf{p})|^2$ are as shown in Fig. 2.12 (they are simply some imaginary sketches).

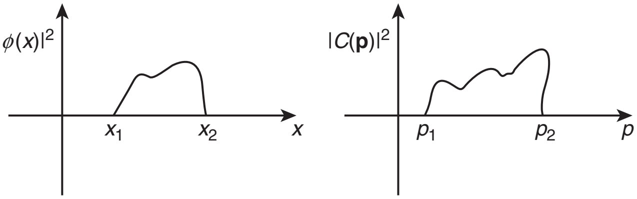


Fig. 2.12 Probability density in coordinate and momentum spaces

$|\phi(x)|^2$ is non–zero or significantly different from zero only in a small region x_1x_2. The Fourier transform $C(\mathbf{p})$ is such that $|C(\mathbf{p})|^2$ also has similar behaviour in the sense that it is non-zero or significantly different from zero only in a small region p_1p_2. We may call $x_2 - x_1 = \Delta x$ as width of $|\phi(x)|^2$ and $p_2 - p_1 = \Delta p$ as width of $|C(\mathbf{p})|^2$. It can be shown that $\Delta x\,\Delta p \sim \hbar$.

It is generally true that for any wave function $\phi(x)$ which is significant only in a small region Δx, its Fourier transform (or, more appropriately, the corresponding momentum wave function) is significant only in a small region Δp such that $\Delta x\,\Delta p \simeq \hbar$

Example 2.13 Find the Fourier transform of $\phi(x) = Ne^{-\frac{\alpha}{2}x^2}$ and sketch the behaviour of $|\phi(x)|^2$ and $|C(\mathbf{p})|^2$.

Solution:

$$\phi(x) = Ne^{-\frac{\alpha}{2}x^2}$$

$$|\phi(x)|^2 = N^2e^{-\alpha x^2}$$

The Fourier transform $C(p)$ of $\phi(x)$ is given by

$$C(p) = \frac{1}{(2\pi\hbar)^{1/2}} \int_{-\infty}^{\infty} N e^{-\alpha x^2/2} e^{-ipx/\hbar} dx = \frac{N}{(2\pi\hbar)^{1/2}} e^{-p^2/2\alpha\hbar^2} \sqrt{\frac{2\pi}{\alpha}}$$

The sketch of $|\phi(x)|^2$ and $|C(\mathbf{p})|^2$ is given in Fig. 2.13.

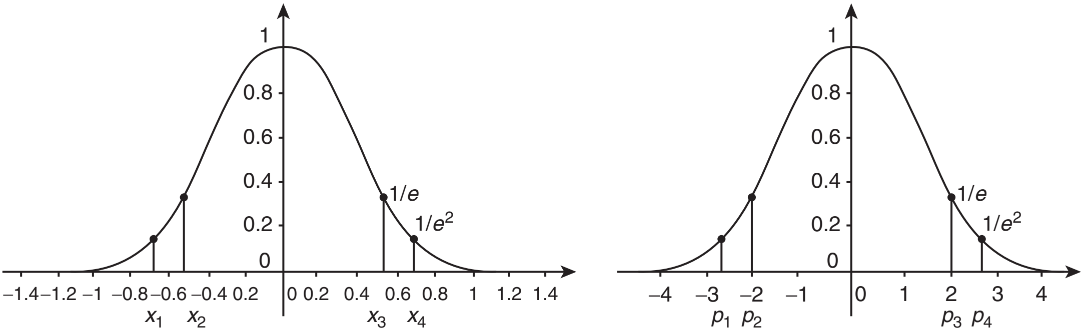


Fig. 2.13 Sketch of $|\phi(x)|^2$ and $|C(\mathbf{p})|^2$ (exp(−4x²) and exp(−p²/4))

What is the 'width' of these two functions?
It depends on how you define width of the function. It is not simple as shown in the Fig. 2.12.

$$|\phi(x)|^2 = \frac{N^2}{e} \quad \text{for} \quad x = \pm\frac{1}{\sqrt{\alpha}} \qquad \text{(Points } x_2 \text{ and } x_3\text{)}$$

$$|C(\mathbf{p})|^2 = \frac{N'^2}{e} \quad \text{for} \quad p = \pm\sqrt{\alpha}\hbar \qquad \text{(Points } p_2 \text{ and } p_3\text{)}$$

If we define Δx and Δp as

$$\Delta x = x_3 - x_2 \quad \text{and} \quad \Delta p = p_3 - p_2, \quad \text{then}$$

$$\Delta x = 2\sqrt{\frac{1}{\alpha}} \quad \text{and} \quad \Delta p = 2\sqrt{\alpha}\hbar$$

$$\therefore\ \Delta x\, \Delta p = 4\hbar$$

On the other hand, if we define

$$\Delta x = x_4 - x_1 \quad \text{and} \quad \Delta p = p_4 - p_1, \quad \text{then}$$

$$\Delta x \Delta p = 2\sqrt{\frac{2}{\alpha}} \cdot 2\sqrt{2\alpha}\ \hbar = 8\hbar$$

The point to understand is that the width of a function is a little arbitrary, but the product of the widths of a function and its Fourier transform is approximately $\hbar$ or a multiple of $\hbar$. Therefore, we can write, for all the reasonable definitions of widths

$$\Delta x\, \Delta p \sim \hbar$$

It is better to understand that Δx and Δp in the above expression are rough estimates. Note that since Δx and Δp are rough estimates, these quantities can also be estimated by considering $\phi(x)$ and $C(p)$ alone.

2.15.4 Uncertainties Δx and Δp

Let us ask the following questions: What is the position of the particle whose state function is $\phi(x)$? What is the momentum of such a particle?

Let us consider a wave packet for which $|\phi(x)|^2$ and $|C(\mathbf{p})|^2$ are as shown in Fig. 2.12. We know that $|\phi(x)|^2$ is the probability density for the position. For the wave packet shown in Fig. 2.12, $|\phi(x)|^2$ is zero for x outside the range x_1x_2 and so the particle will not be found outside this region. Within this region, where will it be found? In quantum mechanics, we can only tell the probability of finding the particle in any position x in this region. This implies that the particle is likely to be found at some point x within this range. We can, therefore, specify the position of the particle to be somewhere in the range bound by x_1 and $x_1 + \Delta x$. So Δx is the uncertainty in the position of the particle. Similarly, we can only specify the momentum of the particle to be some value in the range bound by p_1 and $p_1 + \Delta p$. So Δp is the uncertainty in the momentum of the particle.

Therefore, the Fourier transform properties imply that the product of uncertainties Δp and Δx obey the relation

$$\Delta x \, \Delta p \sim \hbar \tag{2.89}$$

It is to be emphasized again, that Δx and Δp, as they occur here, are rough estimates of the uncertainties from $|\phi(x)|^2$ and $|C(\mathbf{p})|^2$ or $\phi(x)$ and $C(p)$.

A much more rigorous and unambiguous definition of the uncertainties Δx and Δp is provided in Chapter 5. It is defined as follows:

Mean value or average value of

$$x = <x> = \int \psi^*(x,t) x \psi(x,t) dx \tag{2.90}$$

The uncertainty Δx is defined by

$$(\Delta x)^2 = \int \psi^*(x,t)(x - <x>)^2 \psi(x,t) dx \tag{2.91}$$

Similarly, the mean value of p is defined as

$$<p> = \int \psi^*(x,t) p_{op} \psi(x,t) dx \tag{2.92}$$

$$= \int \psi^*(x,t) \left(-i\hbar \frac{\psi}{\psi x} \right) \psi(x,t) dx \tag{2.93}$$

and the uncertainty Δp is defined by

$$(\Delta p)^2 = \int \psi^* (p_{op} - <p>)^2 \psi dx \tag{2.94}$$

With this definition, it will be shown in Chapter 5 that

$$\Delta x \, \Delta p \geq \frac{\hbar}{2} \tag{2.95}$$

It is the rough way of estimating the uncertainty in position and momentum, which is useful in many simple applications. For instance, in the case of hydrogen atom, the size of the hydrogen atom is approximately one Bohr radius which can be taken as the uncertainty in position of the electron. In the case of a particle in a box, we can take the length of the box as the uncertainty in the position.

Example 2.14: Find the momentum wave function for the state described by

$$\phi(x) = \begin{cases} \dfrac{1}{\sqrt{2a}} & -a < x < a \\ 0 & \text{otherwise.} \end{cases}$$

Estimate the uncertainties Δx, Δp and their product.

The sketch of $|\phi(x)|^2$ is given in Fig. 2.14.

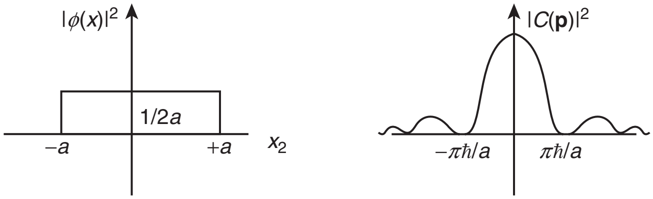


Fig. 2.14 Sketch of $|\phi(x)|^2$ and $|C(p)|^2$

Solution: The momentum wave function $C(p)$ is given by

$$C(p) = \frac{1}{(2\pi\hbar)^{1/2}} \int_{-\infty}^{\infty} \phi(x) e^{-ipx/\hbar} dx = \frac{1}{(2\pi\hbar)^{1/2}} \cdot \frac{1}{\sqrt{2a}} \int_{-a}^{a} e^{-ipx/\hbar} dx$$

$$= \frac{1}{(2\pi\hbar)^{1/2}} \cdot \sqrt{\frac{2}{a}} \cdot \frac{1}{p} \sin \frac{pa}{\hbar}$$

The sketch of $|C(\mathbf{p})|^2$ is given in Fig. 2.14.

From the sketches of $|\phi(x)|^2$ and $|C(\mathbf{p})|^2$, we can estimate the uncertainties Δx and Δp as

$$\Delta x \sim 2a \quad \text{and} \quad \Delta p \sim \frac{2\pi\hbar}{a}$$

$$\therefore \Delta x \, \Delta p \sim 2a \cdot \frac{2\pi\hbar}{a}$$

$$\text{or } \Delta x \, \Delta p \sim 4\pi\hbar.$$

APPENDIX

WAVE PACKET

The concept of wave packet is not unique to quantum mechanics. It is a general property of wave motion like interference and diffraction, which are common to any kind of wave motion. The only requirement is that the wave motion is governed by a linear partial differential equation. We, therefore, first present here some general ideas regarding wave packet and discuss it in the context of quantum mechanics.

Simple Harmonic Wave

The plane wave function $\phi_{\mathbf{p}}(\mathbf{r},t) = Ne^{i(\mathbf{p}\cdot\mathbf{r}-Et)/\hbar}$ is a simple harmonic wave. Normally, it is written as $Ne^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)}$, which may stand for a simple harmonic wave in any phenomenon. This wave is a monochromatic wave extending from $\mathbf{r} = -\infty$ to $+\infty$ ($-\infty < x < \infty$; $-\infty < y < \infty$; $-\infty < z < \infty$). It is an indefinite wave train, and in principle, occupies the entire space. Obviously, this is an idealization that cannot be physically produced.

Phase Velocity

The factor $(\mathbf{k}\cdot\mathbf{r} - \omega t)$ represents phase of simple harmonic wave. It can be written as

$$(\mathbf{k}\cdot\mathbf{r} - \omega t) = k\left(\hat{k}\cdot\mathbf{r} - \frac{\omega}{k}t\right) = k(\hat{k}\cdot\mathbf{r} - \mathrm{v}t)$$

where $\mathrm{v} = \omega/k$ represents the velocity with which the simple harmonic wave or the plane wave front of the simple harmonic wave moves. This is known as phase velocity.

Wave Packet

This is constructed by superposing simple harmonic waves of different wave vector $\mathbf{k}$. Therefore, the wave packet is represented by

$$\psi(\mathbf{r},t) = \frac{1}{(2\pi)^{3/2}}\int A(\mathbf{k})e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)}d^3\mathbf{k}$$

For the sake of clarity, let us restrict our discussion to one dimension. The wave packet is given by

$$\psi(x,t) = \frac{1}{(2\pi)^{1/2}}\int A(k)e^{i(kx-\omega t)}d\mathbf{k}$$

The significance of the wave packet lies in the fact that it is non-zero only in a small range of x values. Outside this range, the value of ψ is zero. A typical wave packet is shown in Fig. 2.15. This is in contrast with a simple harmonic wave that extends from $-\infty$ to ∞.

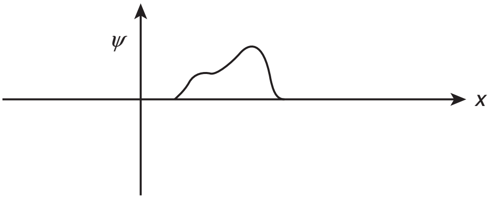


Fig. 2.15 A wave packet

We would like to have a travelling wave packet, travelling with a velocity c as shown in Fig. 2.16. (This is not the velocity of light.) The general functional form of such a wave packet is $f(x - ct)$.

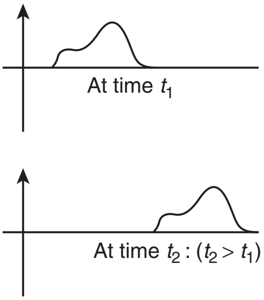


Fig. 2.16 Wave packet at different instants of time

So, we would like to construct a wave packet given by

$$\psi(x,t) = \frac{1}{\sqrt{2\pi}} \int A(k) e^{i(kx-\omega t)} dk = f(x-ct)$$

How do we construct such a wave packet?

The fact that $\omega = \omega(k)$ (i.e. ω is a function of k) helps us to construct such a wave packet. Let us assume that a function $A(k)$ is significant only for values of k near $k = k_0$; that is, as we move away from k_0, $A(k)$ falls rapidly, as shown in Fig. 2.17.

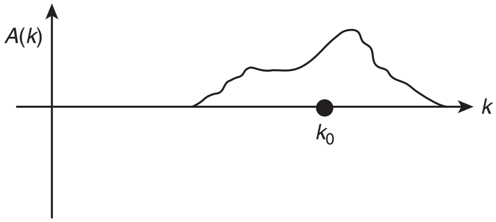


Fig. 2.17 Wave packet centred at $k = k_0$

In such a situation, we can write $\omega(k)$ as

$$\omega(k) = \omega(k_0) + (k-k_0)\left.\frac{d\omega}{dk}\right|_{k=k_0} + \frac{1}{2}(k-k_0)^2 \left.\frac{d^2\omega}{dk^2}\right|_{k_0} + \ldots$$

For the sake of simplicity, let us assume $\omega(k)$ is given by

$$\omega(k) \simeq \omega(k_0) + (k-k_0)\left.\frac{d\omega}{dk}\right|_{k=k_0} = \omega_0 + (k-k_0)\left.\frac{d\omega}{dk}\right|_{k=k_0}$$

$$\therefore \psi(x,t) = \int dk A(k) \exp\left[i\left(kx - \omega_0 t - t(k-k_0)\left.\frac{d\omega}{dk}\right|_{k_0}\right)\right]$$

$$= \int dk A(k) e^{i(k_0 x - \omega_0 t)} \exp\left[i\left((k-k_0)x - t(k-k_0)\left.\frac{d\omega}{dk}\right|_{k_0}\right)\right]$$

$$= e^{i(k_0 x - \omega_0 t)} \int dk A(k) \exp\left[i\left((k-k_0)x - (k-k_0)t\left.\frac{d\omega}{dk}\right|_{k_0}\right)\right]$$

Let us change the integration variable from k to $k' = k - k_0$. Then, $\psi(x, t)$ becomes

$$\begin{aligned}\psi(x,t) &= e^{i(k_0 x-\omega_0 t)}\int dk' A(k')\exp\left[ik'\left(x - t\left.\frac{d\omega}{dk}\right|_{k_0}\right)\right] \\ &= e^{i(k_0 x-\omega_0 t)}F(x-\upsilon_g t)\end{aligned} \tag{2.96}$$

provided we define υ_g as $\upsilon_g = \left.\frac{d\omega}{dk}\right|_{k=k_0}$

This is known as group velocity. That is, this is the velocity with which the wave packet travels. Note that phase velocity is different from group velocity. Phase velocity is given by ω/k whereas group velocity is given by $\upsilon_g = d\omega/dk$.

So far, our discussion about wave packet is general. We have obtained the expression (2.96) based on a number of approximations that may or may not hold in particular cases.

Dispersion of a Wave Packet

Equation (2.96) gives the impression that once a wave packet is created, it will travel as a stable wave packet forever. This is not true. In the derivation of Equation (2.96) we have restricted ourselves to the first order in Taylor's expansion. If we had included other terms also, this result will not hold good.

Normally, as time progresses, different components of the wave packet travel with different phase velocities, and so the wave packet will start dying down. As time progresses, its width will increase while its amplitude (height) will edecrease to zero. Ultimately, the wave packet will disappear. This is known as dispersion.

Wave Packet in Quantum Mechanics

What is the relevance of a wave packet in quantum mechanics? For a free particle, the simple harmonic wave $Ne^{i(\mathbf{p}\cdot\mathbf{r}-Et)/\hbar}$ is a solution to Schrödinger equation. Since Schrödinger equation is a linear partial differential equation, superposition of different simple harmonic waves is also a solution to the Schrödinger equation. This means that a wave packet is also a solution of free particle Schrödinger equation.

Can we use the concept of wave packet to claim that an electron is really a wave? A number of experimental phenomena suggest that it is almost a point–like object. Yet, if one wants to characterize an electron as a wave, it has to be modelled as a wave packet of very small size. The minimal requirement for such a claim is that the wave packet should not disperse into nothing as time progresses.

So, let us investigate the properties of a wave packet obeying Schrödinger equation. We closely follow Greensite here.

Schrödinger equation for free particle is

$$i\hbar\frac{\partial\psi(x,t)}{\partial t} = -\frac{\hbar^2}{2m}\frac{\partial^2\psi(x,t)}{\partial x^2}$$

A plane wave function $\phi_{\mathbf{p}}(x,t) = 1/(2\pi\hbar)^{1/2}\, e^{i(px-Et)/\hbar}$ is a solution to this equation. This is a simple harmonic wave of momentum p that can vary continuously from $-\infty$ to ∞. So we can construct a superposition of all simple harmonic waves whose p values range from $-\infty$ to ∞.

$$\therefore \psi(x, t) = \frac{1}{(2\pi\hbar)^{1/2}}\int C(p)e^{i(px-Et)/\hbar}dp$$

Let us put $t = 0$. Then, we have

$$\psi(x,0) = \frac{1}{(2\pi\hbar)^{1/2}} \int C(p) e^{ipx/\hbar} dp$$

This is the well–known Fourier transform in mathematics. $\psi(x, 0)$ is the Fourier transform of $C(p)$. So the knowledge of $C(p)$ determines the wave packet at $t = 0$. We are interested in the reverse problem: Given a wave packet at $t = 0$, how does it evolve as time progresses? Or given $\psi(x, 0)$, how do you determine $\psi(x, t)$?

The answer is very simple. From the Fourier transformation theory, we know that $C(p)$ can be determined from the relation

$$C(p) = \frac{1}{(2\pi\hbar)^{1/2}} \int \psi(x,0) e^{-ipx/\hbar} dp$$

Once $C(p)$ is determined (from the above equation) it can be used to determine $\psi(x, t)$ using the relation

$$\psi(x,t) = \frac{1}{(2\pi\hbar)^{1/2}} \int C(p) \exp\left[i\left(px - \frac{p^2 t}{2m} \right) / \hbar \right]$$

Subsequent discussion on wave packet can be considered by taking a specific example of a wave packet. The general conclusion will be the same for other wave packets also. The standard textbook example is to choose a Gaussian wave packet as the initial wave packet and determine its evolution.

Example 2.15

The wave packet at $t = 0$ is given by

$$\psi(x,0) = \frac{1}{\left(\sigma^2\pi\right)^{1/4}} e^{-x^2/2\sigma^2} e^{ip_0 x/\hbar}$$

Determine $\psi(x, t)$ which obeys free-particle Schrödinger equation.

Solution: The normalization constant has been chosen such that $\int\limits_{-\infty}^{\infty} |\psi(x,0)|^2 \, dx = 1.$

The following result will be very useful in subsequent calculations.

$$\int\limits_{-\infty}^{\infty} e^{-\alpha x^2 + \beta x} dx = \sqrt{\frac{\pi}{\alpha}} \; e^{\beta^2/4\alpha}$$

$C(p)$ is now given by

$$C(p) = \frac{1}{(2\pi\hbar)^{1/2}} \int\limits_{-\infty}^{\infty} \psi(x,0) e^{-ipx/\hbar} = \frac{N}{(2\pi\hbar)^{1/2}} \int\limits_{-\infty}^{\infty} e^{-x^2/2\sigma^2} e^{-i(p-p_0)x/\hbar} dx$$

$$= \frac{N}{(2\pi\hbar)^{1/2}} \cdot \sqrt{\pi 2\sigma^2} \, e^{-\sigma^2 (p-p_0)^2/2\hbar^2}$$

$$\psi(x,t) = \frac{1}{(2\pi\hbar)^{1/2}} \int\limits_{-\infty}^{\infty} C(p) \exp\left[i\left(\frac{px}{\hbar} - \frac{p^2 t}{2m\hbar} \right) \right]$$

$$= \frac{N}{(2\pi\hbar)} \cdot \sqrt{2\pi\sigma^2} \cdot \int\limits_{-\infty}^{\infty} \exp\left[-\frac{\sigma^2 (p-p_0)^2}{2\hbar^2} + \frac{ipx}{\hbar} - \frac{ip^2 t}{2m\hbar} \right] dp$$

Let us change the integration variable from p to q, where $q = p - p_0$. Then $\psi\left(x, t\right)$ is given by

$$\psi(x, t) = \frac{N}{(2\pi\hbar)}\sqrt{2\pi\sigma^2}\int_{-\infty}^{\infty} \exp\left[-\frac{\sigma^2 q^2}{2\hbar^2} + \frac{i(q+p_0)x}{\hbar} - \frac{i(q+p_0)^2 t}{2m\hbar}\right]dp$$

$$\psi(x,t) = \frac{N}{(2\pi\hbar)}\sqrt{2\pi\sigma^2}\int_{-\infty}^{\infty} \exp\left[-\left(\frac{\sigma^2}{2\hbar^2} + \frac{it}{2m\hbar}\right)q^2 + i\frac{q}{\hbar}\left(x - \frac{p_0}{m}t\right) + i\left(p_0 x - \frac{p_o^2 t}{2m}\right)\frac{1}{\hbar}\right]dp$$

$$\psi(x,t) = \frac{N}{(2\pi\hbar)}\sqrt{2\pi\sigma^2}\exp\left[i\left(p_0 x - \frac{p_o^2 t}{2m}\right)\frac{1}{\hbar}\right]\int_{-\infty}^{\infty} \exp\left[-\left(\frac{\sigma^2}{2\hbar^2} + \frac{it}{2m\hbar}\right)q^2 + i\frac{q}{\hbar}\left(x - \frac{p_0}{m}t\right)\right]dp$$

$$\psi(x,t) = \frac{N}{(2\pi\hbar)}\sqrt{2\pi^2\sigma^2}\,\frac{1}{\left(\frac{\sigma^2}{2\hbar^2} + \frac{it}{2m\hbar}\right)^{1/2}}\exp\left[i\left(p_0 x - \frac{p_o^2 t}{2m}\right)\frac{1}{\hbar}\right]\exp\left[-\frac{(x - p_0 t/m)^2}{4\hbar^2\left(\frac{\sigma^2}{2\hbar^2} + \frac{it}{2m\hbar}\right)}\right]$$

$$\psi(x,t) = \frac{N}{(2\pi\hbar)}\sqrt{2\pi^2\sigma^2}\,\frac{1}{\left(\frac{\sigma^4}{4\hbar^4} + \frac{t^2}{4m^2\hbar^2}\right)^{1/2}}\exp\left[i\left(p_0 x - \frac{p_o^2 t}{2m}\right)\frac{1}{\hbar}\right]\exp\left[-\frac{(x - p_0 t/m)^2\,\frac{\sigma^2}{2\hbar^2}}{4\hbar^2\left(\frac{\sigma^4}{4\hbar^4} + \frac{t^2}{4m^2\hbar^2}\right)}\right]e^{i\theta}$$

This is a Gaussian function whose amplitude decreases continuously as a function of time. In fact, the wave function will become zero for all x for large time t.

Electron is Not a Wave

Let us conclude this section with a simple comment based on the above example. Since an electron exhibits wave properties in some experiments, can we consider electron as a wave? If it is a wave, the small size of the electron implies that it should be a wave packet. So, can we consider electron as a wave packet? The above example implies that if one starts with a wave packet at $t = 0$, the wave packet slowly dies as the time progresses. However, this result is not unique to Gaussian wave packet. It is true for any wave packet obeying free-particle Schrödinger equation. The inability to construct a stable wave packet means that an electron cannot be modelled as a wave packet.

EXERCISES

1. Give arguments leading to the development of Schrodinger equation.
2. Is $\psi(x, t) = A\sin(kx - \omega t)$ a solution to Schrödinger equation?
3. Is it possible to derive Schrödinger equation?
4. What is the physical property whose fluctuations are represented by quantum mechanical wave function $\psi(\mathbf{r}, t)$?
5. What is the interpretation of the wave function $\psi(\mathbf{r}, t)$?
6. Why do we need to normalize the wave function $\psi(\mathbf{r}, t)$?
7. What is meant by conservation of position probability?
8. Why should the position probability for a particle be conserved?
9. Explain how the equation of continuity is a statement of conservation of position probability.
10. Can we normalize the free-particle wave function through the normalization condition $\int \psi^*(\mathbf{r}, t)\psi(\mathbf{r}, t)d^3\mathbf{r} = 1$?

11. What is meant by square integrable wave function?
12. Is $\phi(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}}$ a square integrable function?
13. What are the boundary conditions for the wave functions describing bound state?
14. How do we normalize a plane wave function?
15. Distinguish between box normalization and delta function normalization of plane wave functions.
16. Show that $e^{i\mathbf{p}\cdot\mathbf{r}/\hbar}$ is an eigenfunction of the momentum operator.
17. What is meant by momentum wave function? How do you interpret it?
18. Explain how one estimates the uncertainties from the wave function $\phi(x)$ and its Fourier transforms.
19. Why do you call the eigenstates of time-independent Hamiltonian as stationary eigenstates?
20. Determine the normalization constant of the wave function $\phi(x) = Ne^{-\alpha x^2}$.
21. Show that $\int\limits_{-\infty}^{\infty} e^{-\alpha x^2 - \beta x} dx = e^{\beta^2/4\alpha}\left(\frac{\pi}{\alpha}\right)^{1/2}$.
22. Normalize the wave function $\phi(x) = Ne^{-\alpha|x|}\cos\beta x$.
23. Is $\phi(x) = Ne^{-\alpha|x|}\cos\beta x$ an allowed wave function?

REFERENCES

1. L. I. Schiff 1968, *Quantum Mechanics*, Third Edition, McGraw-Hill International Book Company, New York.
2. S. Gasiorowicz 1974, *Quantum Physics*, John Wiley & Sons, New York.
3. P. M. Mathews and K. Venkatesan 1976, *A Textbook of Quantum Mechanics*, Tata McGraw Hill Publishing Company Limited, New Delhi.
4. J. Greensite, *Lecture Notes on Quantum Mechanics* http://www.physics.sfsu.edu/~greensit/book.pdf.

3

Simple Potentials

Let us recall the Schrödinger equation and the associated boundary conditions which were discussed in Chapter 2. The Schrödinger equation in one dimension is

$$i\hbar\frac{\partial\psi(x,t)}{\partial t} = -\frac{\hbar^2}{2m}\frac{\partial^2\psi(x,t)}{\partial x^2} + V(x,t)\psi(x,t) \tag{3.1}$$

When the potential is independent of time, the solution $\psi(x, t)$ can be written as

$$\psi(x,t) = e^{-iEt/\hbar}u(x) \tag{3.2}$$

where $u(x)$ is the solution of time-independent Schrödinger equation

$$-\frac{\hbar^2}{2m}\frac{d^2u}{dx^2} + V(x)u(x) = Eu(x) \tag{3.3}$$

There are two kinds of solutions: the bound state solution and the scattering state solution. These solutions have to be obtained, subject to the following boundary conditions:

1. $u(x)$ is single valued, finite and continuous. (3.4)
2. The first order derivative $\frac{du}{dx}$ is single valued, finite and continuous. (3.5)
3. For bound state solution,

$$u(x) \to 0 \text{ as } x \to \pm\infty \tag{3.6}$$

For unbound states, the third boundary condition does not hold good.

For bound states, the normalization condition is $\int u^*(x)u(x)dx = 1.$ (3.7)

For bound states, (3.3) is an eigenvalue equation with discrete energy eigenvalues. This equation has to be solved to get the energy eigenvalues E and the corresponding eigenfunctions. In the case of unbound states, the problem is essentially a scattering by a potential. This is very much similar to the case of light crossing a glass slab where the light (wave) is partially reflected and partially transmitted at the interface as shown in Fig. 3.1(i).

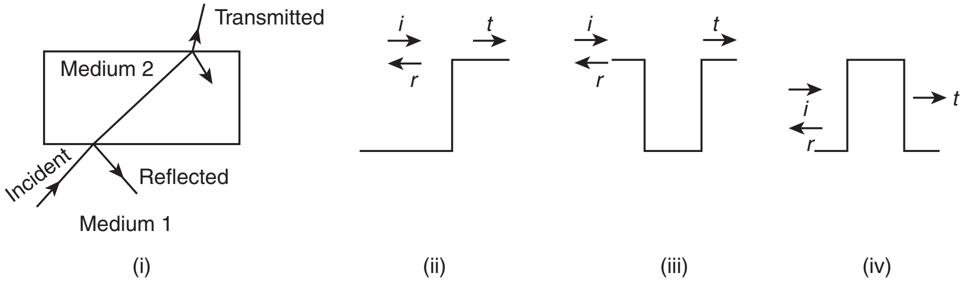


Fig. 3.1 (a) Light waves through slab (b) Scattering of matter waves by different potentials

In the same way, incident matter waves (denoted by *i*) are partially reflected (denoted by *r*) and partially transmitted (denoted by *t*) as shown in Fig. 3.1(b) by different potentials. This implies that some particles of the incident beam of particles are reflected and some particles of the incident beam are transmitted. So, we have to solve the Schrödinger equation for a beam of particles of a given energy *E* to describe the scattering by various potentials. The solution to the Schrödinger equation should describe the incident beam, the reflected beam and the transmitted beam together. The principal task of these problems is to get the reflection and the transmission coefficients for all these scattering processes.

3.1 SCATTERING STATES

3.1.1 Some Preliminaries

Normally, we consider an incident beam of particles moving from left to right. In principle, such an incident beam should be described by a wave packet $f(x-v_g t)$ travelling from left to right, and the time evolution of this wave packet should describe the scattering process. However, for the sake of simplicity, the incident beam is described by the wave function.

$$\psi_{in} = A\,\exp(i(p_{in}x - E_{in}t))/\hbar = A\,e^{i(k_{in}x-\omega_{in}t)} \tag{3.8}$$

This represents a beam of particles of momentum $p_{in} = \hbar k_{in}\hat{\boldsymbol{x}}$ and energy $E_{in} = \hbar^2 k_{in}{}^2/2m$ travelling from left to right. The reflected beam corresponds to a beam of particles moving from right to left, and so they are described by

$$\psi_{ref} = B\,\exp(-i(p_{ref}x + E_{ref}t)/\hbar) = Be^{-i(k_{ref}x+\omega_{ref}t)} \tag{3.9}$$

This represents a beam of particles of momentum $p_{ref} = -\hbar k_{ref}\hat{\boldsymbol{x}}$ and energy $E_{ref} = \hbar^2 k_{ref}{}^2/2m$ travelling left. The transmitted beam corresponds to a beam of particles travelling in the right, which is described by

$$\psi_{tr} = C\,\exp(-i(p_{tr}x - E_{tr}t)/\hbar) = Ce^{-i(k_{tr}x-\omega_{tr}t)} \tag{3.10}$$

This represents a beam of particles of momentum$p_{tr} = \hbar k_{tr}\hat{\boldsymbol{x}}$ and energy $E_{tr} = \dfrac{\hbar^2 k_{tr}{}^2}{2m}$.

For a beam of particles (may be incident or reflected or transmitted), the probability density ρ and the flux of the beam is determined by

$$\rho = \psi^*\psi \quad \text{and} \quad \mathbf{S} = \hat{\boldsymbol{x}}\frac{\hbar}{2im}\left(\psi^*\frac{\partial\psi}{\partial x} - \frac{\partial\psi^*}{\partial x}\psi\right) \tag{3.11}$$

For plane wave functions, we have $\psi = Ne^{i(px-Et)/\hbar}$,

$$\rho = |N|^2\, e^{-i(px-Et)/\hbar} e^{i(px-Et)/\hbar} = |N|^2$$

$$\begin{aligned} \mathrm{S} &= \hat{\boldsymbol{x}}\frac{\hbar}{2im}\left[|N|^2\, e^{-i(px-Et)/\hbar}\,\frac{(ip)}{\hbar}e^{i(px-Et)/\hbar} - |N|^2\, e^{-i(px-Et)/\hbar}\,\frac{(ip)}{\hbar}e^{i(px-Et)/\hbar}\right] \\ &= \hat{\boldsymbol{x}}|N|^2\frac{p}{m} \end{aligned} \tag{3.12}$$

This agrees with the classical notion of flux of particles of velocity v and number density ρ. All the particles within the volume ρv cross the shaded area of unit cross-section in 1 sec.

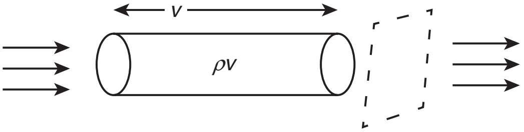


Fig. 3.2 Number of particles crossing unit area is ρv

The reflection coefficient R is defined as

$$R = \frac{\text{Flux of the reflected beam}}{\text{Flux of the incident beam}} \tag{3.13}$$

The transmission coefficient T is defined as

$$T = \frac{\text{flux of the transmitted beam}}{\text{flux of the incident beam}} \tag{3.14}$$

With reference to the wave functions ψ_{in}, ψ_{ref} and ψ_{tr}, given in (3.7), (3.8) and (3.9), we get

$$R = \frac{|B|^2\, \hbar k_{ref}/m}{|A|^2\, \hbar k_{in}/m} = \frac{|B|^2}{|A|^2} \cdot \frac{k_{ref}}{k_{in}} \tag{3.15}$$

$$T = \frac{|C|^2\, \hbar k_{tr}/m}{|A|^2\, \hbar k_{in}/m} = \frac{|C|^2}{|A|^2} \cdot \frac{k_{tr}}{k_{in}} \tag{3.16}$$

Since we are dealing with steady state situation, the factor $e^{-iEt/\hbar}$ or $e^{-i\omega t}$ is normally dropped. So, in subsequent discussions, we will interpret e^{ikx} and e^{-ikx} as follows:

$$e^{ikx} \rightarrow \text{A beam travelling in the positive } x\text{-direction.} \tag{3.17}$$

$$e^{-ikx} \rightarrow \text{A beam travelling in the negative } x\text{-direction.} \tag{3.18}$$

3.1.2 Step Potential $E < V_0$

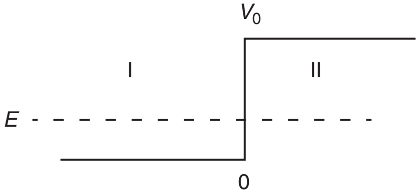


Fig. 3.3 Step potential ($E < V_0$)

The potential $V(x)$ is given by

$$V(x) = \begin{cases} 0 & x < 0 \\ V_0 & x > 0 \end{cases} \tag{3.19}$$

The time-independent Schrödinger equation is $\frac{d^2u}{dx^2} + \frac{2m}{\hbar^2}(E - V(x)) = 0$ (3.20)

Since $V(x)$ is discontinuous at $x = 0$, the Schrödinger equation becomes

$$\frac{d^2u_1}{dx^2} + \frac{2m}{\hbar^2}Eu_1 = 0 \quad x < 0 \tag{3.21}$$

$$\frac{d^2u_{\text{II}}}{dx^2} + \frac{2m}{\hbar^2}(E - V_0)u_{\text{II}} = 0 \quad x > 0 \tag{3.22}$$

This equation has to be solved subject to the boundary conditions (3.4) and (3.5), which lead to the equations

$$u_{\text{I}}(x) = u_{\text{II}}(x) \quad \text{at} \quad x = 0 \tag{3.23}$$

$$\frac{du_{\text{I}}(x)}{dx} = \frac{du_{\text{II}}(x)}{dx} \quad \text{at } x = 0 \tag{3.24}$$

Define k^2 and α^2 as follows:

$$k^2 = \frac{2mE}{\hbar^2} \quad a^2 = \frac{2m}{\hbar^2}(V_0 - E) \tag{3.25}$$

Note that α^2 is a positive quantity. The Schrödinger equation becomes

$$\frac{d^2u_{\text{I}}}{dx^2} + k^2u_{\text{I}} = 0 \qquad x > 0$$

$$\frac{d^2u_{\text{II}}}{dx^2} - \alpha^2u_{\text{II}} = 0 \qquad x < 0$$

The solution to the above equation is

$$u(x) = \begin{cases} u_{\text{I}} = Ae^{ikx} + Be^{-ikx} & x < 0 \\ u_{\text{II}} = Ce^{-\alpha x} + De^{\alpha x} & x > 0 \end{cases}$$

The large x behaviour of $e^{-\alpha x}$ and $e^{-\alpha x}$ are as follows:

$$e^{\alpha x} \to \infty \text{ as } x \to \infty \text{ and } e^{-\alpha x} \to \infty \text{ as } x \to -\infty$$

Therefore, $u_{\text{II}} \to \infty$ as $x \to \infty$. The wave function should be finite at all points. So this is not an acceptable solution. The only way to make this solution acceptable is to drop the term $e^{\alpha x}$ from $u(x)$. This can be done by taking $D = 0$. Therefore, the acceptable solution is

$$u(x) = \begin{cases} u_{\text{I}} = A\,e^{ikx} + Be^{-ikx} & x < 0 \quad (3.26) \\ u_{\text{II}} = Ce^{-\alpha x} & x > 0 \quad (3.27) \end{cases}$$

Substituting u_{I} and u_{II} in the boundary conditions (3.23) and (3.24), we get

$$A + B = C \quad \text{and} \quad ik(A - B) = -\alpha C$$

or

$$1 + \frac{B}{A} = \frac{C}{A} \quad \text{and} \quad 1 - \frac{B}{A} = -\frac{\alpha}{ik}\cdot\frac{C}{A}$$

Solving these two equations, we get

$$\frac{C}{A} = -\frac{2ik}{\alpha - ik}$$

$$\text{and } \frac{B}{A} = -\frac{\alpha + ik}{\alpha - ik}$$

The solution $u(x)$ is

$$u(x) = \begin{cases} A\,e^{ikx} - A\,\dfrac{\alpha + ik}{\alpha - ik}\,e^{-ikx} & x < 0 \\ \dfrac{-2ikA}{\alpha - ik}\,e^{-\alpha x} & x > 0 \end{cases} \tag{3.28}$$

Using (3.25), we get

$$\sqrt{\alpha^2 + k^2} = \sqrt{\frac{2m}{\hbar^2}(V_0 - E + E)} = \sqrt{\frac{2m}{\hbar^2}V_0}$$

Let us write $(\alpha + ik)$ as $(\alpha^2 + k^2)^{1/2} e^{i\theta}$

Then, we have $\dfrac{\alpha + ik}{\alpha - ik} = e^{2i\theta}$ and

$$\frac{1}{\alpha - ik} = \frac{1}{(\alpha^2 + k^2)^{1/2}}\,e^{i\theta}$$

The wave function u_{I} is given by

$$u_{\mathrm{I}} = Ae^{ikx} - Ae^{2i\theta}\,e^{-ikx} = Ae^{i\theta}\,[e^{ikx}e^{-i\theta} - e^{-ikx}e^{i\theta}] = 2\,iAe^{i\theta}\sin\,(kx - \theta)$$

The wave function u_{II} is given by

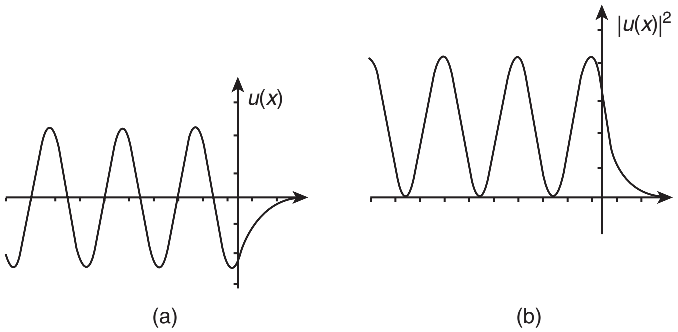
$$u_{\mathrm{II}} = \frac{2iAk}{\sqrt{\alpha^2 + k^2}}\,e^{i\theta}e^{-\alpha x} = -2iA\sin\theta\;e^{i\theta}\;e^{-\alpha x}$$

(Note that $\dfrac{k}{\sqrt{\alpha^2 + k^2}} = \sin\theta$)

The wave function $u(x)$ can be rewritten as

$$u(x) = \begin{cases} N\sin(kx - \theta) & x < 0 \\ -N\sin\theta\;e^{-\alpha x} & x > 0 \end{cases} \tag{3.29}$$

The sketches of the function $u(x)$ and $|u(x)|^2$ are given in Fig. 3.4.



(a) (b)

Fig. 3.4 Sketch of $u(x)$ and $|u(x)|^2$ for step potential

The wave function $u(x)$ as given in (3.26) and (3.27) lends itself to an easy interpretation.
$A\,e^{ikx}$: Wave function describing an incident beam of particles.
Be^{-ikx}: Wave function describing a beam of particles reflected at $x = 0$
$Ce^{-\alpha x}$: A wave function in region II falling to zero rapidly

The wave function (3.29) is the same as (3.28). The distinction between incident beam and the reflected beam is lost in the wave function (3.29). If one is interested only in the probability of finding the particle at some location, this function is sufficient.

Reflection Coefficient

Feeding the expressions Ae^{ikx} and Be^{-ikx} for incident beam and the reflected beam, we get the incident flux J_i and the reflected flux J_r as

$$J_i = |A|^2 \frac{\hbar k}{m}$$

$$J_r = |B|^2 \frac{\hbar k}{m}$$

Therefore, the reflection coefficient R is given by

$$R = \frac{J_r}{J_i} = \frac{|B|^2}{|A|^2} = \left| -\frac{\alpha + ik}{\alpha - ik} \right|^2 = \left| \frac{\sqrt{\alpha^2 + k^2}\, e^{i\theta}}{\sqrt{\alpha^2 + k^2}\, e^{-i\theta}} \right| = 1 \tag{3.30}$$

Obviously, all the incident particles are reflected.

The sketch of the wave function tells more than this ratio. The non-vanishing magnitude of the wave function in region II ($x > 0$) indicates that the incident beam can penetrate small distance in the right, and yet they return back.

Comparison with Classical Physics

Classically, the momentum of the particle is $p = [2m(E - V)]^{1/2}$. Let us recall the assumption that the energy of the particle E is less than V_0. In the region I ($x < 0$), $V = 0$, and so the incident particle has momentum $\sqrt{2mE}$. In the region II ($x > 0$), for a particle with $E < V_0$, $2m(E - V)$ is negative. Therefore, no particle can penetrate the region $x > 0$. This region is a forbidden region for the particle in classical physics. All the particles will reach $x = 0$, where they will be reflected back.

The solution to Schrödinger equation in quantum mechanics also leads to similar results but with some differences. Let us recall that the solution to the Schrödinger equation is

$$u = \begin{cases} A e^{ikx} + B e^{-ikx} & x < 0 \\ C e^{-\alpha x} & x > 0 \end{cases}$$

Remember that $|u(x)|^2$ is the probability density. In the region $x > 0$, the wave function $u(x)$ is not zero, and so there is a definite probability of finding the particle in the classically forbidden region. When we consider a beam of particles, it means that a part of the incident beam penetrates the region $x > 0$ before returning back as the reflected beam of the particles.

Example 3.1 Determine the flux density in the region II for the step potential.

Solution: In the region II, the wave function $u_{\mathrm{II}} = Ce^{-\alpha x}$. The flux density is

$$\begin{aligned} J &= \hat{x} \frac{\hbar}{2im} \left(\psi^* \frac{\partial \psi}{\partial x} - \psi \frac{\partial \psi^*}{\partial x} \right) \\ &= \hat{x} \frac{\hbar}{2m} \left[|C|^2 e^{-\alpha x} (-\alpha) e^{-\alpha x} - |C|^2 e^{-\alpha x} (-\alpha) e^{-\alpha x} \right] \\ &= 0 \end{aligned}$$

3.1.3 Step Potential $E > V_0$

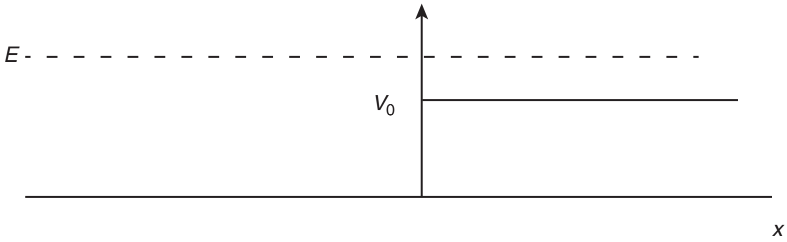


Fig. 3.5 Step potential ($E > V_0$)

Again, consider the potential

$$V(x) = \begin{cases} V_0 & x > 0 \\ 0 & x < 0 \end{cases} \tag{3.31}$$

Now let E be greater than V_0 i.e., $E > V_0$

The Schrödinger equation is

$$\frac{d^2u_{\text{I}}}{dx^2} + \frac{2mE}{\hbar^2}u_{\text{I}} = 0 \qquad x < 0$$

$$\frac{d^2u_{\text{II}}}{dx^2} + \frac{2m}{\hbar^2}(E - V_0)u_{\text{II}} = 0 \quad x > 0$$

Define k^2 and β^2 as

$$k^2 = \frac{2mE}{\hbar^2} \quad \text{and} \quad \beta^2 = \frac{2m}{\hbar^2}(E - V_0)$$

Then Schrödinger equation becomes

$$\frac{d^2u_{\text{I}}}{dx^2} + \alpha^2 u_{\text{I}} = 0 \tag{3.32}$$

$$\frac{d^2u_{\text{II}}}{dx^2} + \beta^2 u_{\text{II}} = 0 \tag{3.33}$$

The solution to Schrödinger equation is

$$u(x) = \begin{cases} u_{\text{I}} = Ae^{ikx} + Be^{-ikx} & x < 0 \\ u_{\text{II}} = Ce^{i\beta x} + De^{-i\beta x} & x < 0 \end{cases}$$

Let us now interpret the various terms in u_{I} and u_{II}.

Ae^{ikx} A wave travelling from left towards $x = 0$, representing the incident beam.

Be^{-ikx} A wave travelling from $x = 0$ towards $x = -\infty$, representing the reflected beam.

$Ce^{i\beta x}$ A wave travelling from $x = 0$ towards $x = \infty$, representing the transmitted beam.

$De^{-i\beta x}$ A wave travelling from $x = \infty$ to $x = 0$. There is no such beam in our problem. So this term should be absent, and therefore, we have to set $D = 0$.

The solution to Schrödinger equation is

$$u(x) = \begin{cases} u_{\text{I}} = Ae^{ikx} + Be^{-ikx} & x < 0 \quad (3.34) \\ u_{\text{II}} = Ce^{i\beta x} & x < 0 \quad (3.35) \end{cases}$$

The boundary condition that the wave function $u(x)$ and its derivative $\frac{du(x)}{dx}$ should be single valued implies

$$u_{\text{I}}(x) = u_{\text{II}}(x) \quad \text{at} \quad x = 0 \tag{3.36}$$

$$\frac{du_{\text{I}}(x)}{dx} = \frac{du_{\text{II}}(x)}{dx} \quad \text{at} \quad x = 0 \tag{3.37}$$

These boundary conditions lead to the following equations:

$$A + B = C \quad \text{and} \quad ik(A - B) = i\beta C$$

Solving these two equations, we get

$$\frac{C}{A} = \frac{2k}{\beta + k} \quad \text{and} \quad \frac{B}{A} = \frac{k - \beta}{k + \beta}$$

Therefore, the wave equations $u(x)$ is given by

$$u(x) = \begin{cases} A\, e^{ikx} + A\,\dfrac{(k - \beta)}{k + \beta} e^{-ikx} & x < 0 \\ A\,\dfrac{2k}{\beta + k} e^{i\beta x} & x > 0 \end{cases} \tag{3.38}$$

Feeding Ae^{ikx} for incident beam, Be^{-ikx} for reflected beam and $Ce^{i\beta x}$ for transmitted beam in the expression for flux, we get

$$\text{Incident flux } J_i = |A|^2 \frac{\hbar k}{m}$$

$$\text{Reflected flux } J_r = |B|^2 \hbar k/m$$

$$\text{and transmitted flux } J_t = |C|^2 \hbar\beta/m$$

The reflection coefficient R is given by

$$R = \frac{J_r}{J_i} = \frac{|B|^2}{|A|^2} = \left(\frac{\beta - k}{\beta + k}\right)^2 \tag{3.39}$$

The transmission coefficient T is given by

$$T = \frac{J_t}{J_i} = \frac{|C|^2}{|A|^2} \cdot \frac{\hbar\beta/m}{\hbar k/m} = \frac{4k\beta}{(k + \beta)^2} \tag{3.40}$$

Note that
$$T + R = \frac{(\beta - k)^2}{(\beta + k)^2} + \frac{(4k\beta)}{(\beta + k)^2} = 1$$

Obviously, this reflects the fact that a single particle is either reflected or transmitted ('either or' in the mutually exclusive sense).

Comparison with Classical Physics

In classical physics, the momentum p of the particle is given by $p = [2m(E - V)]^{1/2}$.

Region I: $x < 0$ $V = 0$ Therefore, the particle moves with a momentum $p = \sqrt{2mE}$.

Region II: $x > 0$ $V = V_0$. The particle continues to move with a momentum $p = \sqrt{2m(E-V)}$.

So the beam of particles of momentum $p = \sqrt{2mE}$ from left reaches $x = 0$ and continues to move with momentum $p = \sqrt{2m(E-V)}$ towards $x = \infty$. The whole beam is transmitted without being reflected.

In quantum mechanics, it is the wave function which describes the system, and $u(x)$ is given by

$$u(x) = \begin{cases} u_{\text{I}} = Ae^{ikx} + A\dfrac{(k-\beta)}{k+\beta}e^{-ikx} & x < 0 \\ u_{\text{II}} = \dfrac{A\,2k}{k+\beta}e^{i\beta x} & x > 0 \end{cases}$$

So, according to quantum mechanics, we get both reflected and transmitted beams. In other words, we get partially reflected and partially transmitted beams.

3.1.4 Square-well Potential – Scattering States

The potential $V(x)$ is given by

$$V(x) = \begin{cases} 0 & |x| > a \\ -V_0 & -a < x < a \end{cases} \tag{3.41}$$

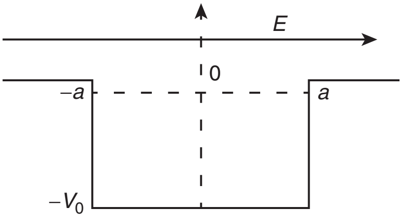


Fig. 3.6 Square-well potential ($E > 0$)

For unbound or scattering states, the energy E should be greater than zero. So $E > 0$.

The Schrödinger equation is

$$\frac{d^2u_{\text{I}}}{dx^2} + \frac{2mE}{\hbar^2}u_{\text{I}} = 0 \qquad x < -a$$

$$\frac{d^2u_{\text{II}}}{dx^2} + \frac{2m}{\hbar^2}(E+V_0)u_{\text{II}} = 0 \qquad -a < x < a$$

$$\frac{d^2u_{\text{III}}}{dx^2} + \frac{2m}{\hbar^2}u_{\text{III}} = 0 \qquad x > a$$

Let us define k^2 and α^2 as

$$k^2 = \frac{2mE}{\hbar^2} \quad \text{and} \quad \alpha^2 = \frac{2m}{\hbar^2}(E+V_0)$$

The Schrödinger equation becomes

$$\frac{d^2u_{\text{I}}}{dx^2} + k^2u_{\text{I}} = 0 \qquad x < -a$$

$$\frac{d^2u_{\text{II}}}{dx^2} + \alpha^2u_{\text{II}} = 0 \qquad -a < x < a$$

$$\frac{d^2u_{\text{III}}}{dx^2} + k^2u_{\text{III}} = 0 \qquad x > a$$

The solutions to these equations are

$$u_{\mathrm{I}} = Ae^{ikx} + Be–^{ikx} \quad x < -a$$
$$u_{\mathrm{II}} = Ce^{i\alpha x} + De^{-i\alpha x} \quad -a < x < a$$
$$u_{\mathrm{III}} = Fe^{ikx} + Ge^{-ikx} \quad x > a$$

Each term of the wave function is interpreted as follows:

Ae^{ikx}	Incident beam travelling from $x = -\infty$ to $x = -a$.
Be^{-ikx}	Beam reflected at $x = -a$ travelling from $x = -a$ to $x = -\infty$.
Ce^{ikx}	Beam travelling from $x = -a$ to $x = a$.
De^{-ikx}	Beam travelling from $x = a$ to $x = -a$.
Fe^{ikx}	Beam travelling from $x = a$ towards $x = \infty$. This represents the transmitted beam.
Ge^{-ikx}	Beam travelling from $x = \infty$ to $x = a$. In our problem, we don't have such a beam. So this term should not be present in the wave function. Therefore, we set $G = 0$.

The constants A, B, C, D and F are determined from the boundary conditions and the normalization condition. The single valuedness of the wave function $u(x)$ and $du(x)/dx$ lead to the following conditions:

1. $u_{\mathrm{I}}(x) = u_{\mathrm{II}}(x) \quad \text{at} \quad x = -a$ (3.42)
2. $\dfrac{du_{\mathrm{I}}(x)}{dx} = \dfrac{du_{\mathrm{II}}(x)}{dx} \quad \text{at} \quad x = -a$ (3.43)
3. $u_{\mathrm{II}}(x) = u_{\mathrm{III}}(x) \quad \text{at} \quad x = a$ (3.44)
4. $\dfrac{du_{\mathrm{II}}(x)}{dx} = \dfrac{du_{\mathrm{III}}(x)}{dx} \quad \text{at} \quad x = a$ (3.45)

These boundary conditions imply

$$Ae^{-ika} + Be^{ika} = Ce^{-i\alpha a} + De^{i\alpha a} \tag{3.46}$$
$$ik\,(Ae^{-ika} - B\,e^{ika}) = i\alpha\,(Ce^{-i\alpha a} - De^{i\alpha a}) \tag{3.47}$$
$$Ce^{i\alpha a} + De^{-i\alpha a} = Fe^{ika} \tag{3.48}$$
$$i\alpha(Ce^{i\alpha a} - De^{-i\alpha a}) = Fik\,e^{ika} \tag{3.49}$$

From (3.48) and (3.49), we get

$$i\alpha\,(Ce^{i\alpha a} - De^{-i\alpha a}) = ik\,(Ce^{i\alpha a} + De^{-i\alpha a})$$

Rearranging the terms, we get

$$D = Ce^{2i\alpha a}\,\frac{(\alpha - k)}{(\alpha + k)} \tag{3.50}$$

Adding (3.46) and ((3.47)/*ik*), we get

$$2Ae^{-ika} = Ce^{-i\alpha a}\left(1 + \frac{\alpha}{k}\right) + De^{i\alpha a}\left(1 - \frac{\alpha}{k}\right) \tag{3.51}$$

Subtracting ((3.47)/*ik*) from (3.46), we get

$$2Be^{ika} = Ce^{-i\alpha a}\left(1 - \frac{\alpha}{k}\right) + De^{i\alpha a}\left(1 + \frac{\alpha}{k}\right) \tag{3.52}$$

Let us now determine $\frac{1}{R}$, where R is the reflection coefficient.

$$\frac{1}{R} = \frac{\text{Incident flux}}{\text{Reflected flux}} = \frac{|A|^2}{|B|^2}$$

Dividing (3.51) by (3.52), we get

$$\frac{A}{B}e^{-2ika} = \frac{Ce^{-i\alpha a}(k+\alpha)+De^{i\alpha a}(k-\alpha)}{Ce^{-i\alpha a}(k-\alpha)+De^{i\alpha a}(k+\alpha)} = \frac{Ce^{-i\alpha a}(k+\alpha)+Ce^{3i\alpha a}\dfrac{(\alpha-k)}{\alpha+k}(k-\alpha)}{Ce^{-i\alpha a}(k-\alpha)+Ce^{3i\alpha a}\dfrac{\alpha-k}{(\alpha+k)}(k+\alpha)}$$

Dividing throughout both numerator and denominator by $Ce^{i\alpha a}$, we get

$$= \frac{e^{-2i\alpha a}(k+\alpha)^2 - e^{2i\alpha a}(\alpha-k)^2}{e^{-2i\alpha a}(k^2-\alpha^2)+e^{2i\alpha a}(\alpha^2-k^2)}$$

$$= \frac{(\cos 2\alpha a - i\sin 2\alpha a)(k^2+\alpha^2+2k\alpha)-(\cos 2\alpha a + i\sin 2\alpha a)(\alpha^2+k^2-2\alpha k)}{2i\sin 2\alpha a\cdot(\alpha^2-k^2)}$$

$$= \frac{2\alpha k\cos 2\alpha a - i\sin 2\alpha a(k^2+\alpha^2)}{i(\alpha^2-k^2)\sin 2\alpha a}$$

$$\frac{1}{R} = \left|\frac{A}{B}e^{-2ika}\right|^2 = \frac{(k^2+\alpha^2)^2\sin^2 2\alpha a + 4\alpha^2k^2\cos^2 2\alpha a}{(\alpha^2-k^2)\sin^2 2\alpha a}$$

$$= \frac{(\alpha^2-k^2)^2\sin^2 2\alpha a + 4\alpha^2k^2}{(\alpha^2-k^2)^2 2\sin^2 2\alpha a} = 1 + \frac{4\alpha^2k^2}{(\alpha^2-k^2)2\sin^2 2\alpha a}$$

$$\therefore R = \frac{1}{\left[1+\dfrac{4\alpha^2k^2}{(\alpha^2-k^2)^2\sin^2 2\alpha a}\right]} \tag{3.53}$$

Transmission Coefficient

The transmission coefficient T is given by

$$T = \frac{\text{Flux of the transmitted particles}}{\text{Flux of the incident particles}}$$

$$= \left|\frac{F}{A}\right|^2$$

Dividing (3.51) by (3.48), we get

$$\frac{A}{F}e^{2ika} = \frac{1}{2k}\cdot\frac{Ce^{-i\alpha a}(\alpha+k)+De^{i\alpha a}(k-\alpha)}{(Ce^{i\alpha a}+De^{-i\alpha a})} = \frac{1}{2k}\left[\frac{Ce^{-i\alpha a}(\alpha+k)+\dfrac{Ce^{3i\alpha a}(k-\alpha)(\alpha-k)}{\alpha+k}}{Ce^{i\alpha a}+Ce^{i\alpha a}\dfrac{\alpha-k}{\alpha+k}}\right]$$

$$= \frac{1}{4\alpha k}[e^{-2i\alpha a}(\alpha+k)^2 - e^{2i\alpha a}(\alpha-k)^2] = \frac{1}{2\alpha k}[2\alpha k\cos 2\alpha a - i(k^2+\alpha^2)\sin 2\alpha a]$$

$$T^{-1} = \left|\frac{A}{F}\right|^2 = \left|\frac{A}{F}e^{-2ika}\right|^2$$

$$= \frac{1}{4\alpha^2k^2}[4\alpha^2k^2\cos^2 2\alpha a + (k^2+\alpha^2)^2\sin^2 2\alpha a]$$

$$= \frac{[1+(\alpha^2-k^2)^2\sin^2 2\alpha a]}{4\alpha^2k^2}$$

$$\therefore T = \frac{1}{\left[1 + \frac{(\alpha^2 - k^2)^2 \sin^2 2\alpha a}{4\alpha^2 k^2}\right]} \tag{3.54}$$

$$\text{Note that } R + T = \frac{1}{\left[1 + \frac{4\alpha^2 k^2}{(\alpha^2 - k^2)\sin^2 2\alpha a}\right]} + \frac{1}{\left[1 + \frac{(\alpha^2 - k^2)\sin^2 2\alpha a}{4\alpha^2 k^2}\right]}$$
$$= 1$$

Let us recall that $\alpha^2 = \frac{2m}{\hbar^2}(E + V_0)$ and $k^2 = \frac{2mE}{\hbar^2}$. Using these definitions, we get

$$R = \frac{1}{\left[1 + \frac{4(E + V_0)}{V_0 \sin^2 2\alpha a}\right]}$$

$$\text{and } T = \left[\frac{1}{1 + \frac{V_0 \sin^2 2\alpha a}{4(E + V_0)}}\right]$$

For large values of E compared to V_0, the transmission coefficient approaches 1. Therefore, when $E >> V_0$, most of the particles are transmitted.

Transmission Resonance

It is interesting to note that the transmission coefficient becomes 1 whenever $2\alpha a = n\pi$. This can be easily seen from (3.54) Note that α^2 and k^2 are very much similar to each other. Since k is a wave number, we can interpret α also as a wave number, provided we define $\alpha = 2\pi/\lambda_e$ where λ_e is the effective de Broglie wavelength of the particle in the region $-a < x < a$. Then we have

$$2\alpha a = n\pi \quad \Rightarrow 2 \cdot \frac{2\pi a}{\lambda_e} = n\pi$$

Note that $2a$ is the width of the potential. Therefore, for transmission resonance,

$$\text{Width of the potential} = \frac{n\lambda_e}{2}$$

So, as we change the energy of the incident particle, the transmission coefficient also changes. However, at some particular values of E, the transmission coefficient becomes 1. In other words, the interaction becomes completely transparent so that all the particles are transmitted. These are energy values such that the de Broglie wavelength λ_e satisfies the relation $\lambda_e = 2/n \cdot 2a$. This can be described as transmission resonance. In fact, this can be taken as a simple model, exhibiting the concept of transmission resonance. Such a transmission resonance was first observed by Ramsauer and Townsend in the scattering of low energy electrons by noble gas atoms. (Note that the analysis of this phenomenon is much more complicated than this simple model.)

3.1.5 Rectangular Potential Barrier $0 < E < V_0$

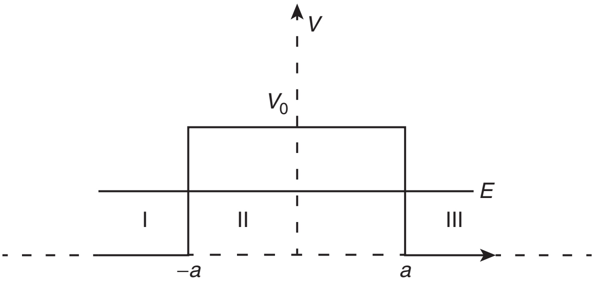


Fig. 3.7 Rectangular potential barrier $0 < E < V_0$

The potential $V(x)$ is given by

$$V(x) = \begin{cases} 0 & x < -a \\ V_0 & -a < x < a \\ 0 & x > a \end{cases} \tag{3.55}$$

The Schrödinger equation is

$$\begin{aligned} \frac{d^2u_\text{I}}{dx^2} + \frac{2mE}{\hbar^2}u_\text{I} &= 0 && x < -a \\ \frac{d^2u_\text{II}}{dx^2} + \frac{2m}{\hbar^2}(E - V_0)u_\text{II} &= 0 && -a < x < a \\ \frac{d^2u_\text{III}}{dx^2} + \frac{2m}{\hbar^2}Eu_\text{III} &= 0 && x > a \end{aligned}$$

Let us assume the energy of the particle to be E, which is less than V_0. i.e., $0 < E < V_0$. Therefore, $(E - V_0)$ is negative. Taking into account of this fact, let us define α^2 and k^2 as

$$k^2 = \frac{2m}{\hbar^2}E \quad \text{and} \quad \alpha^2 = \frac{2m}{\hbar^2}(V_0 - E)$$

Note that α^2 is a positive real number. The Schrödinger equation becomes

$$\begin{aligned} \frac{d^2u_\text{I}}{dx^2} + k^2u_\text{I} &= 0 && x < -a \\ \frac{d^2u_\text{II}}{dx^2} - \alpha^2u_\text{II} &= 0 && -a < x < a \\ \frac{d^2u_\text{III}}{dx^2} + k^2u_\text{III} &= 0 && x > a \end{aligned}$$

The solution to this equation is

$$\begin{aligned} u_\text{I} &= Ae^{ikx} + Be^{-ikx} && x < -a \\ u_\text{II} &= Ce^{\alpha x} + De^{-\alpha x} && -a < x < a \\ u_\text{III} &= Fe^{ikx} + Ge^{-ikx} && x > a \end{aligned}$$

The wave function $u(x)$ is interpreted in the same way as done in the case of square-well potential.

Ae^{ikx}	Incident wave function representing the incident beam of particles travelling from. $-\infty$ to $-a$
Be^{-ikx}	Reflected wave function representing the reflected beam of particles.
$Ce^{\alpha x}$ and $De^{-\alpha x}$	They are well-behaved functions. They are finite and single valued within the range $-a < x < a$. Therefore, they are quantum mechanically allowed wave functions.
Fe^{ikx}	Transmitted wave function representing the transmitted particles.
Ge^{-ikx}	Wave function describing the beam of particles travelling from $x = \infty$ to $x = a$. We are interested in knowing what happens to the incident beam of the particles from the left. We would like to know whether these particles are transmitted beyond $x = a$. So, in this scenario, we should not have a beam travelling from right to left in the region $x > a$. Therefore, we set $G = 0$.

Therefore, the solution is

$$u_{\text{I}} = Ae^{ikx} + Be^{-ikx} \qquad x < -a \tag{3.56}$$

$$u_{\text{II}} = Ce^{\alpha x} + Be^{-\alpha x} \qquad -a < x < a \tag{3.57}$$

$$u_{\text{III}} = Fe^{ikx} \qquad x > a \tag{3.58}$$

The boundary condition that the wave function $u(x)$ and its derivatives $\frac{du(x)}{dx}$ should be (single continuous valued) leads to the following conditions:

$$u_{\text{I}}(x) = u_{\text{II}}(x) \quad \text{at} \quad x - a \tag{3.59}$$

$$\frac{du_{\text{I}}(x)}{dx} = \frac{du_{\text{II}}(x)}{dx} \quad \text{at} \quad x = -a \tag{3.60}$$

$$u_{\text{II}}(x) = u_{\text{III}}(x) \quad \text{at} \quad x = a \tag{3.61}$$

$$\frac{du_{\text{II}}(x)}{dx} = \frac{du_{\text{III}}(x)}{dx} \quad \text{at} \quad x = a \tag{3.62}$$

Substituting u_{I}, u_{II} and u_{III} in these conditions, we get

$$Ae^{-ika} + Be^{ika} = Ce^{-\alpha a} + De^{\alpha a} \tag{3.63}$$

$$ik(Ae^{-ika} - Be^{ika}) = \alpha(Ce^{-\alpha a} - De^{a}) \tag{3.64}$$

$$Ce^{\alpha a} + D\,e^{-\alpha a} = Fe^{ika} \tag{3.65}$$

$$\alpha(Ce^{\alpha a} - De^{-\alpha a}) = F\,ik\,e^{ika} \tag{3.66}$$

Using (3.65) in (3.66), we get

$$Ce^{\alpha a} - De^{-\alpha a} = \frac{ik}{\alpha}(Ce^{\alpha a} + De^{-\alpha a})$$

Rearranging the terms, we get

$$C = De^{-2\alpha a}\frac{(\alpha + ik)}{\alpha - ik} \tag{3.67}$$

Adding (3.63) and ((3.64)/ik), we get

$$2Ae^{-ika} = Ce^{-\alpha a}\left(1 + \frac{\alpha}{ik}\right) + De^{\alpha a}\left(1 - \frac{\alpha}{ik}\right) \tag{3.68}$$

Subtracting ((3.64)/ik) from (3.68) we get,

$$2Be^{ika} = Ce^{-\alpha a}\left(1-\frac{\alpha}{ik}\right)+De^{\alpha a}\left(1+\frac{\alpha}{ik}\right) \tag{3.69}$$

Transmission Coefficient

Let us determine the transmission coefficient T.

From (3.68) and (3.65), we get

$$\frac{2Ae^{-ika}}{Fe^{ika}} = \frac{Ce^{-\alpha a}(ik+\alpha)+De^{\alpha a}(ik-\alpha)}{ik[Ce^{\alpha a}+De^{-\alpha a}]} = \frac{Ce^{-\alpha a}(ik+\alpha)+Ce^{2\alpha a}\dfrac{(\alpha-ik)}{\alpha+ik}e^{\alpha a}(ik-\alpha)}{ik\left[Ce^{\alpha a}+Ce^{2\alpha a}\dfrac{(\alpha-ik)}{(\alpha+ik)}e^{-\alpha a}\right]}$$

Dividing throughout by $Ce^{\alpha a}$, we get

$$\frac{2A\,e^{-ika}}{Fe^{ika}} = \frac{[e^{-2\alpha a}(\alpha+ik)^2-e^{2\alpha a}(\alpha-ik)^2]}{2\alpha ik} = \frac{1}{2\alpha ik}[(e^{-2\alpha a}-e^{2\alpha a})(\alpha^2-k^2)+2i\alpha k(e^{-2\alpha a}+e^{2\alpha a})]$$

$$= \frac{1}{\alpha ik}[2i\alpha k\cosh 2\alpha a-(\alpha^2-k^2)\sinh 2\alpha a]$$

Now, consider $\left|\dfrac{A}{F}\right|^2$

$$\left|\frac{A}{F}\right|^2 = \left|\frac{A}{F}e^{-2ika}\right|^2 = \frac{1}{4\alpha^2k^2}[(\alpha^2-k^2)^2\sinh^2 2\alpha a+4\alpha^2k^2\cosh^2 2\alpha a]$$

$$= \frac{1}{4\alpha^2k^2}[(\alpha^2-k^2)^2\sinh^2 2\alpha a+4\alpha^2k^2(1+\sinh^2 2\alpha a)]$$

$$= \left[1+\frac{(\alpha^2+k^2)^2\sinh^2 2\alpha a}{4\alpha^2k^2}\right]$$

∴ The transmission coefficient T is given by

$$T = \left|\frac{F}{A}\right|^2 = \frac{1}{\left[1+\dfrac{(\alpha^2+k^2)^2\sinh^2 2\alpha a}{4\alpha^2k^2}\right]} \tag{3.70}$$

Similarly, we can calculate the reflection coefficient R, which is given by

$$R = \frac{\text{Flux of the reflected beam}}{\text{Flux of the incident beam}}$$

$$= \left|\frac{B}{A}\right|^2$$

Using (3.68) and (3.69), we get

$$\frac{Ae^{-ika}}{Be^{ika}} = \frac{ik[e^{-2\alpha a}(\alpha+ik)^2-e^{2\alpha a}(\alpha-ik)^2]}{(\alpha^2+k^2)(e^{2\alpha a}-e^{-2\alpha a})}$$

$$= \frac{-(\alpha^2 - k^2)(e^{2\alpha a} - e^{-2\alpha a}) + 2ik\alpha(e^{-2\alpha a} + e^{2\alpha a})}{(\alpha^2 + k^2)(e^{2\alpha a} - e^{-2\alpha a})}$$

$$= \frac{-(\alpha^2 - k^2)\sinh 2\alpha a + 2ik\cosh 2\alpha a}{(\alpha^2 + k^2)\sinh 2\alpha a}$$

$$\left|\frac{A}{B}\right|^2 = \left|\frac{A}{B}e^{2ik\alpha}\right|^2$$

$$= \frac{(\alpha^2 - k^2)^2\sinh^2 2\alpha a + 4k^2\alpha^2\cosh^2 2\alpha a}{(\alpha^2 + k^2)\sinh^2 2\alpha a}$$

$$= \frac{(\alpha^2 - k^2)^2\sinh^2 2\alpha a + 4k^2\alpha^2(1 + \sinh^2 2\alpha a)}{(\alpha^2 + k^2)^2\sinh^2 2\alpha a}$$

$$= \frac{(\alpha^2 + k^2)\sinh^2 2\alpha a + 4\alpha^2 k^2}{(\alpha^2 + k^2)^2\sinh^2 2\alpha a} = \left[1 + \frac{4\alpha^2 k^2}{(\alpha^2 + k^2)^2\sinh^2 2\alpha a}\right]$$

$$\therefore R = \left|\frac{B}{A}\right|^2 = \frac{1}{\left[1 + \frac{4\alpha^2 k^2}{(\alpha^2 + k^2)\sinh^2 2\alpha a}\right]} \tag{3.71}$$

Classically, the transmission of a particle with energy $E < V_0$ is not possible. In region I and III, the energy of the particle is E and the momentum is $p = \sqrt{2mE}$. Such a particle can be present in either region I or region III. However, in region II, $(E - V)$ is negative, which means that we cannot have a particle with $E < V_0$ in region II. In fact, this is a forbidden region in classical physics. A particle with $E < V_0$, coming from left can travel up to $x = -a$ where it will be reflected. Transmission of such a particle to region III from region I is ruled out since it has to cross the forbidden region.

The quantum mechanical results differ from that of classical physics. The solution to Schrödinger equation leads to non-zero transmission coefficient. There is a definite probability that a particle with energy $E < V_0$ travelling from left can be found to be present in region III. In fact, the particle can be found even in region II since the wave function is not zero in region II. This phenomenon is called tunnelling, which is unique to quantum mechanics.

3.1.6 Tunnelling Phenomenon

Tunnelling is not unique to rectangular barrier potentials. It is present in a number of other situations. However, in all such situations, the tunnelling phenomena can be explained only in terms of quantum mechanics. So it is a unique consequence or feature of quantum mechanics. Since the tunnelling phenomenon takes place in a number of other situations, it is worthwhile to consider it in more detail.

Classically, the total energy E, the potential energy $V(x)$ and the momentum p are such that the quantity $2m\,(E - V(x))$ should always be positive. So, in classical physics, we have two kinds of regions:

Classically allowed regions: A region in which $(E - V)$ is positive

Classically forbidden region: A region in which $(E - V)$ is negative

In tunnelling phenomenon, usually we have two classically allowed regions sandwiched by a classically forbidden region or a classically allowed region sandwiched between two classically forbidden regions. The phenomenon in which a particle travelling from one classically allowed region to the other region through a classically forbidden region is known as tunnelling.

In the case of rectangular barrier, in region I and III, the quantity ($E - V$) is positive, and so they are classically allowed regions. In region II, the quantity ($E - V$) is negative and so this is a classically forbidden region.

Can a particle with $E < V_0$, starting from left in the region I, be found in the region III? Classically, this is not possible. The particle will reach $x = -a$, where it will be reflected. To cross region II, it should have energy $E > V_0$.

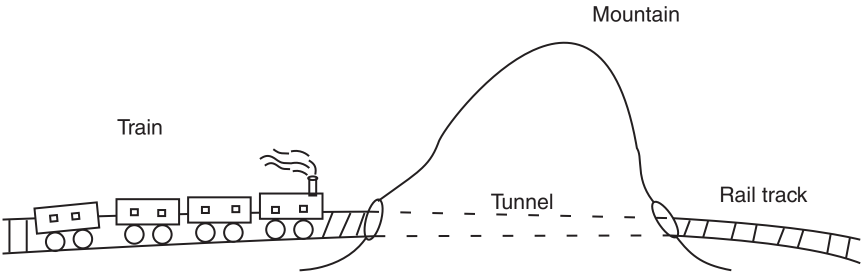


Fig. 3.8 Tunnel through a mountain

This is very much similar to a train in the left side of the mountain as shown in Fig. 3.8. This train cannot reach the right side unless there exists a tunnel through the mountain.

The quantum mechanical analysis is entirely different. In quantum mechanics, the criterion for finding the particle in a region is not determined by the sign of ($E - V$). The presence or absence or impossibility of finding a particle in a specific region is determined by the wave function $\psi(x, t)$. So it is the solution to the Schrödinger equation which determines whether a particle will be found in a region or not. If $\psi \neq 0$ in a region, there is a definite probability of finding the particle there. Only if $\psi = 0$ in a region, the particles cannot be present in that region. Note that in quantum mechanics, the momentum of the particle is not given by $\sqrt{2m(E-V)}$.

A Word of Caution

The analogy between a mountain and the potential barrier should not be extended too far. Mountain is a physical object. The rectangular barrier $V(x)$ as shown in Fig. 3.7 is a mathematical graph of $V(x)$ vs. x, and it is not a pictorial representation of a real physical situation.The real situation may be like the one as shown in Fig. 3.9.

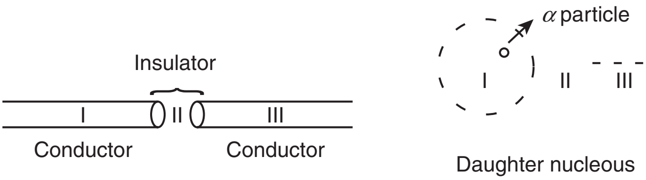


Fig. 3.9 Physical phenomena using tunnelling

In Fig. 3.9 (left figure), the gap between the two conductors may be filled by an insulator. The gap itself is an insulator. Normally, a continuous flow of current cannot be maintained because of the presence of the insulator. However, if the width of the insulator is very small, quantum mechanical tunnelling process allows a continuous flow of current. This concept is used in scanning tunnelling microscope.

Fig. 3.9 (right figure) describes α particle decay process. The α decay is given by ${}^{Z}_{A}X \rightarrow {}^{Z-2}_{A-4}Y + \alpha$. Here X is the parent nucleus and Y is the daughter nucleus. To describe the emission of α particle, X can be thought of as Y within which α particle as an entity is present which tries to escape from Y nucleus. So the volume of the daughter nucleus is region I, a classically allowed region. Region II is immediate outside of the daughter nucleus. This is classically forbidden region. Region III is classically allowed region. The reason for this classification for these regions are explained below. The emission of α particles is possible due to quantum mechanical tunnelling.

Scanning Tunnelling Microscope

In 1981, Gerd Binnig and Heinrich Rohrer constructed a device called scanning tunnelling microscope through which individual atoms can be located or manipulated. The operation of this device is based on tunnelling phenomenon.

Let us start with a simplistic analysis of a metal. The simplest model of a metal is that it is a periodic arrangement of atoms. The valence electrons, which are not attached to the atoms, are confined to the metal but are free to move within the metal. These electrons constitute an 'electron sea'. Application of quantum mechanics to the electron gas leads to the energy levels as shown in Fig. 3.10. All the states whose energy is less than E_F are occupied (strictly speaking, this is true only at zero Kelvin). To lift an electron from the metal, one has to supply a minimum energy ϕ to the electron inside a metal. ϕ is called work function. All these results can be summarized in the potential energy diagram as a function of x as shown in Fig. 3.10 (top left).

The potential for an electron confined to the metal is shown in Fig. 3.10 (top right), which is a step potential similar to Fig. 3.3. (For region inside metal 1, $x < 0$; for region outside the metal 1, $x > 0$.) Generally, one expects an electron inside a metal to be confined within it. The wave function for an

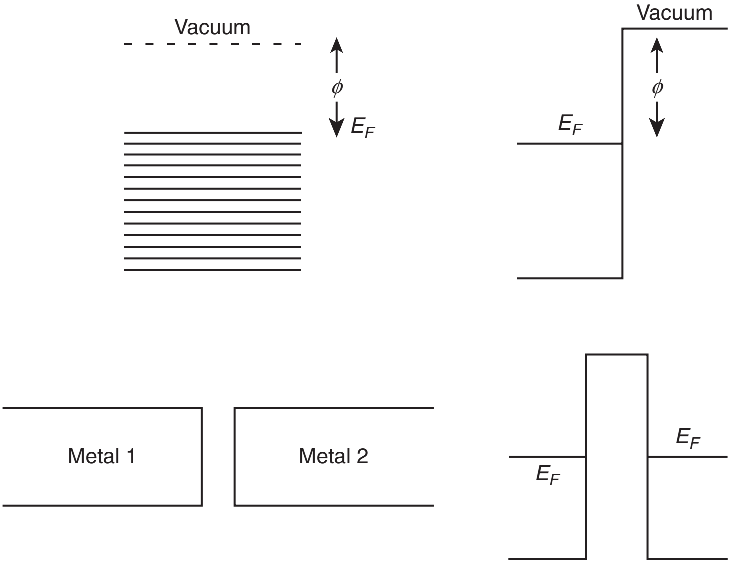


Fig. 3.10 Energy level diagram for two metals with a gap

electron for a step potential tells the other way. There is a definite probability of finding the electron outside the metal in the region $x > 0$. However, these electrons will not leave the metal.

Let us now bring another metal and keep it very close to it as shown in Fig. 3.10 (lower left). Now the potential energy for a system of two metals with a small gap between them is as shown in Fig. 3.10 (lower right). Now there can be a flow of electrons from metal 1 to metal 2 due to tunnelling in spite of the presence of a gap (which acts as an insulator).

The scanning tunnelling microscope is very much similar to it. The schematic diagram of scanning tunnelling microscope is shown in Fig. 3.11. There is a probe, which is a thin needle with sharp tip over the sample metal. The needle never touches the surface. However, there is an electron flow from the sample material to the needle, and hence there is a current which can be measured. The circuit is schematically shown in Fig. 3.11. Note that it is the tunnelling phenomenon which is responsible for the flow of electrons from the sample surface to the probe.

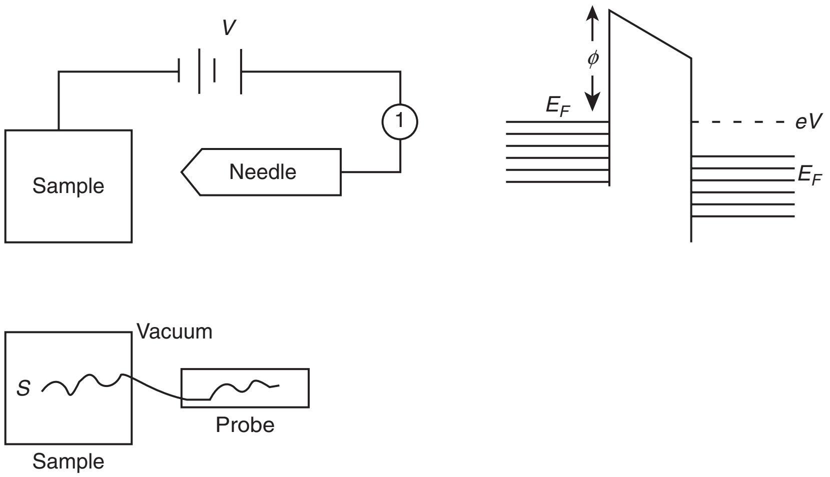


Fig. 3.11 Schematic diagram for scanning tunnelling microscope

α Decay

The α decay is given by

$$ {}^{Z}_{A}X \rightarrow {}^{Z-2}_{A-4}Y + \alpha $$

For instance, the radioactive decay of ${}^{92}_{232}\mathrm{U}$ is given by

$$ {}^{92}_{232}\mathrm{U} \rightarrow \alpha + {}^{90}_{228}\mathrm{Th} $$

It is assumed that an α particle is formed from the nucleons in the parent nucleus ${}^{92}_{232}\mathrm{U}$. Now, the ${}^{92}_{232}\mathrm{U}$ can be considered as a bound system of α particle and ${}^{90}_{228}\mathrm{Th}$, which leads to the following analysis:

The size of a nucleus can be calculated from the relations $r = r_0 A^{1/3}$ where $r_0 = 1.4$ fm. For thorium nucleus, $r = (1.4\text{fm})(228)^{1/3} \simeq 8\,\text{fm}$. Within this distance, the attractive nuclear force is the most dominant which keeps all the nucleons bound together. The nuclear force has very short range, and so we can take it to be absent beyond 8 fm. It is the electrostatic repulsion i.e., dominant force beyond 8 fm.

The electrostatic potential energy $V = \dfrac{Z_1 Z_2}{4\pi \epsilon_0 r}$

$$= \frac{90 \times 2 \times (1.602 \times 10^{-19})^2}{4\pi \times 8.85 \times 10^{-12} r}$$

$$= \frac{260 \text{ MeV}}{r \text{ in Fermi}}$$

All these details are brought out in the potential energy diagram as shown in Fig. 3.12.

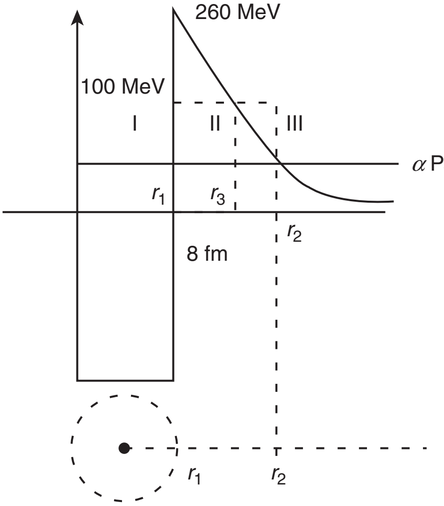


Fig. 3.12 α decay and tunnelling process

Let us do classical physics first. For distances less than r_1, α particles can have 5 MeV. Once it comes out of the nucleus, there is no attractive nuclear force on it, and the only force is electrostatic repulsive force, which is so strong that the α particle has to move with very high velocity. For instance, the velocity of the α particle at r_3 should be such that its energy is 100 MeV. Obviously, a 5 MeV α particle cannot be present here. As the α particle moves away from the nucleus, the repulsive force decreases, and only after r_2 it is possible for an α particles to have 5 MeV. Therefore, regions I and III are classically allowed regions, and region II is classically forbidden region for 5 MeV α particle. Only quantum mechanics allows a 5 MeV α particle present inside the daughter nucleus to escape from the nucleus with the same energy to outside the nucleus. This is again a tunnelling phenomenon. The calculation of the transmission coefficient becomes much more complicated, which will be evaluated using WKB approximation later.

3.2 BOUND STATE SOLUTIONS

3.2.1 Parity

Before taking specific potentials and solving Schrödinger equation, let us introduce an important concept called parity. This is also called space inversion. The concept of parity is presented here in the context of simple one-dimensional potential. This deals with how wave functions $u(x)$ and $u(-x)$ are related. Let us introduce an operator called parity operator P, which is defined as

$$Pu(x) = u(-x) \tag{3.72}$$

Eigenvalues and Eigenfunctions of Parity Operator

The eigenvalue equation of parity operator P is given by

$$P\phi(x) = \lambda\phi(x) \tag{3.73}$$

But
$$P\phi(x) = \phi(-x) \tag{3.74}$$

Consider $P^2\phi(x)$. Using (3.74), we get

$$P^2\phi(x) = PP\phi(x) = P\phi(-x) = \phi(x). \tag{3.75}$$

But, from (3.73), we get

$$P^2\phi(x) = PP\phi(x) = P\lambda\phi(x) = \lambda P\phi(x) = \lambda^2\phi(x) \tag{3.76}$$

Comparing (3.75) and (3.76), we have

$$\lambda^2\phi(x) = \phi(x) \quad \text{or} \quad \lambda^2 = 1$$

$$\therefore\ \lambda = +1, -1. \tag{3.77}$$

Eigenvalue $\lambda = 1$; Even-parity Wave Function

$$P\phi(x) = 1 \cdot \phi(x)$$

But
$$P(x) = \phi(-x).$$

$$\therefore \lambda = 1 \text{ implies } \phi(-x) = \phi(x) \tag{3.78}$$

Such functions are called even functions of x or functions with even parity.

Eigenvalues $\lambda = -1$. Odd-parity Eigenfunction

$$P\phi(x) = (-1)\phi(x)$$

But
$$P\phi(x) = \phi(-x)$$

$$\therefore \phi(-x) = -\phi(x) \tag{3.79}$$

Such functions are called an odd function of x or a function with odd parity.

Example 3.2 Determine the parity of (1) $u(x) = x^2$ (2) $u(x) = \sin\alpha x$.

Solution: The simplest way is to replace x by $-x$ and see how these functions behave.

1. $u(x) = x^2$

$$u(-x) = (-x)^2 = x^2 = u(x)$$

This is an even function or a function with even parity.

2. $u(x) = \sin\alpha x$

$$u(-x) = \sin\alpha(-x) = -\sin\alpha x = -u(x)$$

$\therefore$ This is a function with odd parity.

Example 3.3 Determine the parity of $u(x) = e^{\alpha x}$.

Solution:

$$u(x) = e^{\alpha x}$$
$$u(-x) = e^{-\alpha x}$$

This is not an eigenfunction of parity operator. It is neither odd nor even.

Example 3.4 Determine the parity of $u(x) = A \sin\alpha x + B \cos\alpha x$.

Solution: Note that $\sin\alpha x$ is an odd-parity function and $\cos\alpha x$ is an even-parity function.

i.e., $$\sin \alpha(-x) = -\sin\alpha x \quad \cos\alpha(-x) = \cos\alpha x$$

Let us now consider $u(x) = A \sin\alpha x + B \cos\alpha x$.

$$\therefore u(-x) = -A \sin\alpha x + B \cos\alpha x$$

This is neither even nor odd. Therefore, this is not an eigenfunction of parity operator.

Stationary Eigenstates and Parity

Are the stationary eigestates of a time-independent Hamiltonian parity eigenstates?

The stationary eigenstates are energy eigenstates of the Hamiltonian operator.

$$H_{op}(x) = \frac{p_{op}^2}{2m} + V(x) = -\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + V(x)$$

Changing variable x to $-x$, we get

$$H_{op}(-x) = -\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + V(-x)$$

Is H_{op} itself symmetric with respect to parity transformation? i.e., whether $H_{op}(x)$ and $H_{op}(-x)$ are the same? They may be or may not be the same. It depends on $V(-x)$.

If $$V(-x) = V(x), \quad H_{op}(x) = H_{op}(-x)$$

It can be shown that if $H_{op}(x) = H_{op}(-x)$, the stationary energy eigenstates are also parity eigenstates. i.e., if $V(-x) = V(x)$, then the stationary energy eigenstates are either even or odd function.

For $$V(-x) = V(x), \quad u(-x) = \pm u(x) \tag{3.80}$$

This fact will be useful in determining the energy eigenstates. Note that there are many potentials which are not of the type $V(-x) = V(x)$

3.2.2 Particle in a Box

Let us now choose the potential for a particle in a box in a symmetric way as shown in Fig. 3.13.

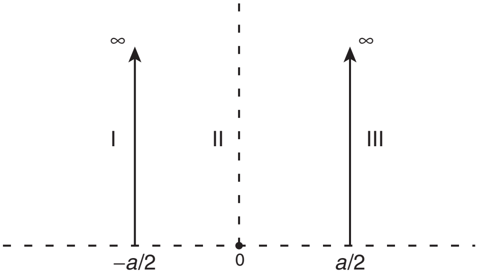


Fig. 3.13 Particle in a box (symmetric potential)

The potential energy $V(x)$ is given by

$$V(x)=\begin{cases}\infty & x<-\dfrac{a}{2}\\ 0 & -\dfrac{a}{2}<x<\dfrac{a}{2}\\ \infty & x>\dfrac{a}{2}\end{cases} \tag{3.81}$$

Since the potential $V(x)$ is infinite in regions I and III, the wave function $u(x) = 0$ in these regions. So we have

$$\begin{aligned} u(x)&=0 && x\le -\frac{a}{2}\\ \frac{d^2u}{dx^2}+\frac{2mE}{\hbar^2}u(x)&=0 && -\frac{a}{2}\le x\le \frac{a}{2}\\ u(x)&=0 && x\ge \frac{a}{2}\end{aligned} \tag{3.82}$$

Defining, $k^2=\dfrac{2mE}{\hbar^2}$, the Schrödinger equation becomes

$$\frac{d^2u(x)}{dx^2}+k^2u(x)=0 \qquad -\frac{a}{2}\le x\le\frac{a}{2} \tag{3.83}$$

It is obvious that $V(-x) = V(x)$ or $V(x)$ is symmetric about the origin. This implies that energy eigenfunctions $u(x)$ should be either odd or even.

i.e.,
$$u(-x)=\pm u(x) \tag{3.84}$$

So to get the energy eigenvalues and the eigenfunctions, we have to find the solutions of the Equation (3.83) subject to the boundary conditions

$$u(x)=0 \quad \text{at} \quad x=-\frac{a}{2}$$

$$u(x)=0 \quad \text{at} \quad x=\frac{a}{2}$$

and
$$u(-x)=\pm\, u(x)$$

The general solution of the Equation (3.83) is

$$u(x)=A\sin kx+B\cos kx$$

The boundary condition $u(-x) = u(x)$ demands that

$$-A\sin kx+B\cos kx=A\sin kx+B\cos kx$$

This is possible provided $A = 0$.

Therefore, the even-parity wave functions are given by $u(x) = B\cos kx$.

On the other hand, the boundary condition $u(-x) = -u(x)$ demands that

$$-A\sin kx+B\cos kx=-A\sin kx-B\cos kx.$$

$$\therefore B=0$$

Therefore, the odd-parity wave function is $u(x) = A\sin kx$

We have two sets of energy eigenfunctions, namely

$$u(x)=B\cos kx \quad \text{even parity}$$

$$u(x)=A\sin kx \quad \text{odd parity.}$$

Even-parity Wave Functions

The boundary condition $u(x) = 0$ at $x = \frac{a}{2}$ demands that

$$u\left(\frac{a}{2}\right) = B\cos\frac{ka}{2} = 0$$

$$\frac{ka}{2} = \left(n+\frac{1}{2}\right)\pi \quad \text{or} \quad k = (2n+1)\pi/a \quad n = 0, 1, 2, 3,..$$

The energy eigenvalues are

$$E = \frac{\hbar^2}{2m}\cdot\frac{(2n+1)^2\pi^2}{a^2} \qquad n = 0, 1, 2, 3,...$$

The list of energy eigenvalues are

$$\left\{\frac{\hbar^2}{2m}\cdot\left(\frac{\pi}{a}\right)^2, \frac{\hbar^2}{2m}\cdot\left(\frac{3\pi}{a}\right)^2, \frac{\hbar^2}{2m}\cdot\left(\frac{5\pi}{a}\right)^2, ...\right\}$$

Odd-parity Wave Functions

The boundary conditions $u(x) = 0$ at $x = \frac{a}{2}$ implies that

$$u\left(\frac{a}{2}\right) = A\sin\frac{ka}{2} = 0$$

$$\therefore \frac{ka}{2} = n\pi \quad \text{or} \quad k = \frac{2n\pi}{a} \qquad n = 1, 2, 3,...$$

The energy eigenvalues are

$$E = \frac{\hbar^2}{2m}\cdot\left(\frac{2n\pi}{a}\right)^2 \qquad n = 1, 2,...$$

The list of energy eigenvalues are given by

$$\left\{\frac{\hbar^2}{2m}\cdot\left(\frac{2\pi}{a}\right)^2, \frac{\hbar^2}{2m}\cdot\left(\frac{4\pi}{a}\right)^2, \frac{\hbar^2}{2m}\cdot\left(\frac{6\pi}{a}\right)^2, ...\right\}$$

Taking into account both sets of energy eigenvalues, we can arrange them into a single set with energy eigenvalues in ascending order and their corresponding energy eigenfunctions.

$$E_1 = \frac{\hbar^2}{2m}\left(\frac{\pi}{a}\right)^2 \quad u_1(x) = B\cos\frac{\pi x}{a} \tag{3.85}$$

$$E_2 = \frac{\hbar^2}{2m}\left(\frac{2\pi}{a}\right)^2 \quad u_2(x) = A\sin\frac{2\pi x}{a} \tag{3.86}$$

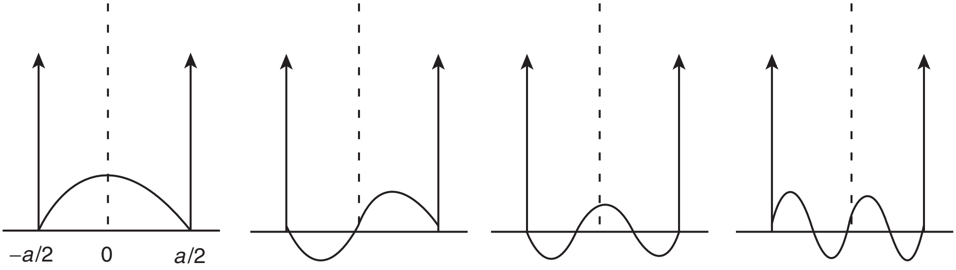
$$E_3 = \frac{\hbar^2}{2m}\left(\frac{3\pi}{a}\right)^2 \quad u_3(x) = B\cos\frac{3\pi x}{a} \tag{3.87}$$

$$E_4 = \frac{\hbar^2}{2m}\left(\frac{4\pi}{a}\right)^2 \quad u_4(x) = A\sin\frac{4\pi x}{a} \tag{3.88}$$

$$E_5 = \frac{\hbar^2}{2m}\left(\frac{6\pi}{a}\right)^2 \quad u_5(x) = B\cos\frac{5\pi x}{a} \tag{3.89}$$

$$E_6 = \frac{\hbar^2}{2m}\left(\frac{7\pi}{a}\right)^2 \quad u_6(x) = A\sin\frac{6\pi x}{a} \tag{3.90}$$

The sketches of some of the eigenfunctions are given in Fig. 3.14.



$u_1 = A\cos \pi x/a$ $u_2 = B\sin 2\pi x/a$ $u_3 = A\cos 3\pi x/a$ $u_4 = B\sin 4\pi x/a$

Fig. 3.14 Sketch of wave functions for a particle in a box (symmetric potential)

Example 3.5 Normalize the energy eigenfunctions of a particle in a box defined by $-\frac{a}{2} < x < \frac{a}{2}$.

Solution: Even-parity wave functions:
The even-parity wave functions $u_n(x)$ is given by

$$u_n(x) = \begin{cases} 0 & x < -\frac{a}{2} \\ B\cos\frac{(2n+1)\pi x}{a} & -\frac{a}{2} < x < \frac{a}{2} \\ 0 & x < \frac{a}{2} \end{cases}$$

The normalization constant B is determined by demanding

$$\int_{-\infty}^{\infty} |u(x)^2| dx = 1$$

$$\int_{-a/2}^{a/2} B^2\cos^2\frac{(2n+1)\pi x}{a} = B^2\int_{-a/2}^{a/2}\frac{1}{2}\left(1+\cos\frac{(2n+1)2\pi x}{a}\right)dx$$

$$= \frac{1}{2}B^2 a + \frac{1}{2}B^2\left[\sin\frac{(2n+1)2\pi x}{a}\cdot\frac{a}{(2n+1)2\pi}\right]_{-a/2}^{a/2} = \frac{B^2 a}{2} = 1$$

$$\therefore \quad B = \sqrt{\frac{2}{a}}.$$

The normalized even-parity wave function $u_n(x)$ is given by

$$u_n(x) = \sqrt{\frac{2}{a}}\cos\frac{(2n+1)\pi x}{a}$$

Odd-parity wave function:

The odd-parity wave function is given by

$$u_n(x) = \begin{cases} 0 & x < -\frac{a}{2} \\ A \sin\dfrac{2n\pi x}{a} & -\frac{a}{2} < x < \frac{a}{2} \\ 0 & x > \frac{a}{2} \end{cases}$$

The normalization constant A is determined as follows.

$$\int_{-a/2}^{a/2} A^2 \sin^2 \frac{2n\pi x}{a} dx = 1 \quad \text{or} \quad A^2 \int_{-a/2}^{a/2} \left[1 - \cos\frac{4n\pi x}{a}\right]\frac{1}{2} dx = \frac{A^2}{2} a = 1$$

$$\therefore A = \sqrt{\frac{a}{2}}$$

The normalized odd-parity wave function is

$$V_n(x) = \sqrt{\frac{2}{a}} \sin\frac{2n\pi x}{a}$$

Example 3.6 Show that the energy eigenfunctions for a particle confined in a box defined by $-\frac{a}{2} < x < \frac{a}{2}$ are orthogonal to each other.

Solution: Two functions $\phi_n(x)$ and $\phi_m(x)$ are said to be orthogonal to each other if they obey

$$\int_{-\infty}^{-\infty} \phi_m^*(x)\ \phi_n(x) dx = \delta_{mn}$$

The even-parity wave function $u_n(x)$ and the odd-parity wave function $V_n(x)$ are given by

$$u_n(x) = \sqrt{\frac{2}{a}} \cos\frac{(2n+1)\pi x}{a}$$

and
$$V_n(x) = \sqrt{\frac{2}{a}} \sin\frac{(2n+1)\pi x}{a}$$

In the context of our problem, by orthogonally we mean

$$\int_{-a/2}^{a/2} u_n^*(x) u_n(x) dx = \delta_{mn}, \quad \int_{-a/2}^{a/2} V_n^*(x) V_m(x) dx = \delta_{mn}$$

and
$$\int_{-a/2}^{a/2} u_n^*(x) V_m(x) dx = 0$$

$$\int_{-a/2}^{a/2} u_n^*(x) u_m(x) dx = \int_{-a/2}^{a/2} \frac{2}{a} \cos\frac{(2n+1)\pi x}{a} \cos\frac{(2m+1)\pi x}{a} dx$$

$$= \int_{-a/2}^{a/2} \frac{1}{a} \left\{\cos\frac{(2n+2m+2)\pi x}{a} + \cos\frac{2(n-m)\pi x}{a}\right\} dx$$

$$= \frac{1}{a\pi}\left\{\frac{a}{2(n+m+1)}\cdot\sin\frac{2(n+m+1)\pi x}{a}\Bigg|_{-a/2}^{a/2} + \frac{a}{2(n-m)}\sin\frac{2(n-m)\pi x}{a}\Bigg|_{-a/2}^{a/2}\right\}$$

$$= 0 \text{ provided } n \neq m$$

When $n = m$, we know then $\int_{-a/2}^{a/2} u_n^*(x)u_n(x)dx = 1$

$$\therefore \int_{-a/2}^{a/2} u_n^*(x)u_m(x)dx = \delta_{mn}$$

In the same way, for odd-parity wave functions, it is easy to check

$$\int_{-a/2}^{a/2} V_n^*(x)V_m(x)dx = \delta_{mn}$$

Now consider $\int_{-a/2}^{a/2} u_n^*(x)V_m(x)dx$

$$\int_{-a/2}^{a/2} u_n^*(x)V_m(x)dx = \int_{-a/2}^{a/2} \frac{2}{a}\cos\frac{(2n+1)\pi x}{a}\cdot\sin\frac{2m\pi x}{a}dx$$

$$= \frac{1}{a}\int_{-a/2}^{a/2}\left[\sin\frac{(2(m+n)+1)\pi x}{a} + \sin\frac{(2m-2n-1)\pi x}{a}\right]dx$$

$$= \frac{2}{a}\cdot\left[-\frac{a}{2(m+n)+1}\cos\frac{(2(m+n)+1)\pi x}{a} + \frac{a}{(2m-2n-1)}\cos\frac{(2m-2n-1)\pi x}{a}\right]_{-a/2}^{a/2}$$

$$= 0 \text{ provided } 2m \neq 2n+1$$

Consider the case when $2m = 2n + 1$

$$\int_{-a/2}^{a/2} V_n^*(x)V_m(x)dx = \frac{2}{a}\int_{-a/2}^{a/2} \cos\frac{2m\pi x}{a}\sin\frac{2m\pi x}{a}dx$$

$$= \frac{1}{a}\int_{-a/2}^{a/2} \sin\frac{4m\pi x}{a}dx$$

$$= \frac{1}{a}\cdot\frac{a}{4m\pi}\cos\frac{4m\pi x}{a}\Bigg|_{-a/2}^{a/2} \quad \text{provided } m \neq 0$$

$$= 0 \text{ provided } m \neq 0.$$

The minimum value of m is 1 for particle in a box. Therefore,

$$\int_{-a/2}^{a/2} u_n^*(x)V_m(x)dx = 0$$

3.2.3 Square-well Potential

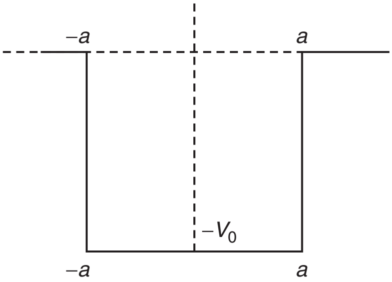


Fig. 3.15 Square-well potential for bound states

The potential is given by

$$V(x) = \begin{cases} 0 & x < -a \\ -V_0 & -a < x < a \\ 0 & x > a \end{cases} \tag{3.91}$$

We are looking for bound state solution for this potential. Therefore, the energy of the bound states should be such that $-V_0 < E < 0$. i.e., $E = -|E|$ and $|E| < V_0$.
The Schrödinger equation is

$$\frac{d^2u_{\text{I}}}{dx^2} + \frac{2mE}{\hbar^2}u_{\text{I}} = 0 \qquad x < -a$$

$$\frac{d^2u_{\text{II}}}{dx^2} + \frac{2m}{\hbar^2}(V_0 + E)u_{\text{II}} = 0 \qquad -a < x < a$$

$$\frac{d^2u_{\text{III}}}{dx^2} + \frac{2m}{\hbar^2}E\,u_{\text{III}} = 0 \qquad x > a$$

Let us define $\alpha^2 = -\frac{2mE}{\hbar^2}$ and $\beta^2 = -\frac{2m}{\hbar^2}(V_0 + E)$.

Note that α^2 is a positive number since E is less than zero. Since $V_0 > |E|$, $(V_0 - |E|)$ or $(V_0 + E)$is a positive number, β^2 is a positive number. The Schrödinger equation now becomes

$$\frac{d^2u_{\text{I}}}{dx^2} - \alpha^2 u_{\text{I}} = 0 \qquad x < -a$$

$$\frac{d^2u_{\text{II}}}{dx^2} + \beta^2 u_{\text{II}} = 0 \qquad -a < x < a$$

$$\frac{d^2u_{\text{III}}}{dx^2} - \alpha^2 u_{\text{III}} = 0 \qquad x > a$$

The solution to these equations are

$$\begin{aligned} u_{\text{I}} &= A\,e^{\alpha x} + B\,e^{-\alpha x} && x < -a \\ u_{\text{II}} &= C\cos\beta x + D\sin\beta x && -a < x < a \\ u_{\text{III}} &= F\,e^{\alpha x} + G\,e^{-\alpha x} && x > a \end{aligned}$$

Consider the solution u_{I} in region $-\infty < x < -a$. As $x \to -\infty$, the function $e^{\alpha x}$ goes to zero, but $e^{-\alpha x}$ becomes infinite. Similarly, consider the solution u_{III} in the region $a < x < \infty$. As $x \to \infty$, the function $e^{-\alpha x}$ goes to zero, but $e^{\alpha x}$ becomes infinite. So because of the boundary condition (3.6), the bound state wave function should go to zero as $x \to \pm\infty$. This can be met if only we set $B = 0$ and $F = 0$. Therefore, the solution becomes

$$u_{\text{I}} = A\, e^{\alpha x} \qquad x < -a \tag{3.92}$$

$$u_{\text{II}} = C \cos \beta x + D \sin \beta x \qquad -a < x < a \tag{3.93}$$

$$u_{\text{III}} = G\, e^{-\alpha x} \qquad x > a \tag{3.94}$$

The potential energy $V(x)$ is symmetric about $x = 0$; therefore, the solution $u(x)$ should have definite parity.

i.e., $u(-x) = u(x)$ even-parity function

$u(-x) = -u(x)$ odd-parity function

Even-parity Wave Function

From Equations (3.92), (3.93) and (3.94) we write $u(-x)$ as

$$u(-x) = \begin{cases} A\, e^{-\alpha x} & x > a \\ C \cos \beta x - D \sin \beta x & a > x > -a \\ G e^{\alpha x} & x < -a \end{cases}$$

The condition that $u(-x) = u(x)$ implies that

$$G = A \quad \text{and} \quad D = 0$$

The solution now becomes

$$u_{\text{I}} = A e^{\alpha x} \qquad x < -a \tag{3.95}$$

$$u_{\text{II}} = C \cos \beta x \qquad -a < x < a \tag{3.96}$$

$$u_{\text{III}} = A e^{-\alpha x} \qquad x > a \tag{3.97}$$

The boundary condition demanding the single valuedness of the wave function $u(x)$ and its derivatives $du(x)/dx$ at $x = -a$ leads to the following conditions.

$$u_{\text{I}}(x = -a) = u_{\text{II}}(x = -a) \quad \text{and} \quad \left.\frac{du_{\text{I}}(x)}{dx}\right|_{x=-a} = \left.\frac{du_{\text{II}}(x)}{dx}\right|_{x=-a}$$

These conditions imply that

$$A e^{-\alpha a} = C \cos \beta a \tag{3.98}$$

$$\alpha A e^{-\alpha a} = C\, \beta \sin \beta a \tag{3.99}$$

Using (3.98) in (3.99), we get

$$\alpha a = \beta a \tan \beta a \tag{3.100}$$

Note that.tan $\beta a = \alpha a/\beta a$. Since α and β are positive numbers, tan βa is also a positive number. This implies that βa has to be in either first quadrant or third quadrant.

i.e.,

$$2n\frac{\pi}{2} < \beta a < (2n+1)\frac{\pi}{2} \qquad n = 0, 1, 2, 3, \ldots \tag{3.101}$$

To get the energy eigenvalue corresponding to even-parity solution, we have to solve the Equation (3.100) subject to the range of βa given in the inequality (3.101).

We cannot get an algebraic expression for the solution of this equation. Only numerical solution is possible.

Before proceeding further, let us analyze the odd-parity solution.

Odd Parity Solution

The odd-parity solution is generated by the condition

$$u(-x) = -u(x) \tag{3.102}$$

From (3.95), (3.96) and (3.97), we have

$$u(-x) = \begin{cases} Ae^{-\alpha x} & x > a \\ C\cos\beta x - D\sin\beta x & a > x > -a \\ Ge^{\alpha x} & x < -a \end{cases}$$

The condition (3.102) implies that $G = -A$ and $C = 0$. Therefore, the odd-parity solution is given by

$$u(x) = \begin{cases} Ae^{\alpha x} & x < -a & (3.103) \\ D\sin\beta x & -a < x < a & (3.104) \\ -Ae^{-\alpha x} & x > a & (3.105) \end{cases}$$

The boundary condition demanding the wave function $u(x)$ and its derivatives $\frac{du(x)}{dx}$ be single valued at $x = -a$ implies the following conditions.

At $x = -a$ $\quad u_{\mathrm{I}}(x) = u_{\mathrm{II}}(x)$ and $\frac{du_{\mathrm{I}}(x)}{dx} = \frac{du_{\mathrm{II}}(x)}{dx}$

$$\therefore A\,e^{-\alpha a} = -D\sin\beta a \tag{3.106}$$

$$\alpha A e^{-\alpha a} = D\,\beta\cos\beta a \tag{3.107}$$

Making use of (3.106) in (3.107), we get

$$-\alpha a = \beta a\cot\beta a \tag{3.108}$$

or

$$\cot\beta a = -\frac{\alpha a}{\beta a} \tag{3.109}$$

Since α and β are positive number, cot βa is a negative number. This implies βa has to be either in second quadrant or fourth quadrant.

i.e.,

$$(2n+1)\frac{\pi}{2} < \beta a < 2(n+1)\frac{\pi}{2} \quad n = 0, 1, 2, 3, \ldots \tag{3.110}$$

To get the energy eigenvalue corresponding to the odd-parity solution, we have to solve the Equation (3.108) subject to the range of βa given in (3.110). Again, it is not possible to get an algebraic expression as the solution of this equation.

Graphical Solution

We can get a qualitative idea of the energy eigenvalues for different values of V_0 from graphical solution. Let us introduce three new variables X, Y and Z as follows:

$$X = \beta a \quad Y = \alpha a \quad Z^2 = X^2 + Y^2$$

$$Z^2 = X^2 + Y^2 = (\beta^2 + \alpha^2)a^2 = \frac{2m}{\hbar^2} V_0 a^2$$

To get the energy eigenvalues, we have to solve the following equations:

Even parity: $Y = X \tan X$: $\quad 2n\frac{\pi}{2} < X < (2n+1)\frac{\pi}{2}$

$$X^2 + Y^2 = Z^2$$

Odd parity: $Y = -X \cot X$: $\quad (2n+1)\frac{\pi}{2} < X < (2n+2)\frac{\pi}{2}$

$$X^2 + Y^2 = Z^2$$

In the graphical solution, the intersection of the two curves $Y = X \tan X$ and $X^2 + Y^2 = Z^2$ in the range $2n\,\pi/2 < X < (2n+1)\pi/2$ will determine the energy eigenvalue corresponding to even-parity solution. The intersection of the two curves $Y = -X \cot X$ and $X^2 + Y^2 = Z^2$ in the range $(2n+1)\pi/2 < X < (2n+2)\pi/2$ will determine the energy eigenvalues corresponding to odd-parity solution.

The sketches of the three curves namely $Y = X \tan X$, $X^2 + Y^2 = Z^2$ and $Y = -X \cot X$ are shown in Fig. 3.16.

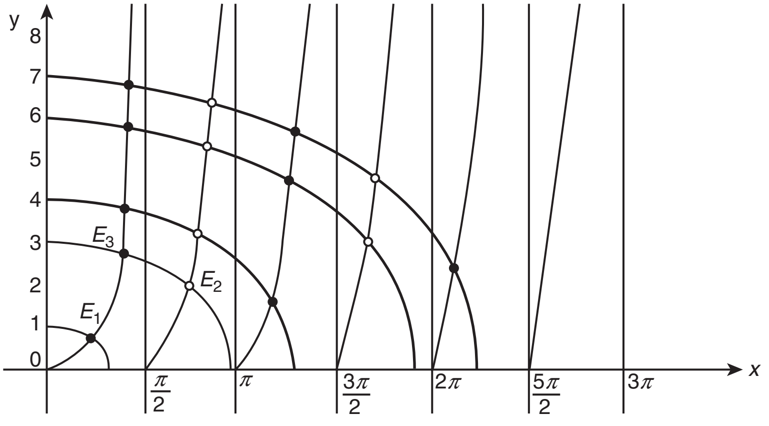


Fig. 3.16 Square-well potential – Graphical solution for energy eigenvalues

For five different values of Z, the intersection of the circle $X^2 + Y^2 = Z^2$ and the curves $Y = X \tan X$ and the curve $Y = -X \cot X$ is shown in Fig. 3.16. The small dots '•' represent the intersection between the circles $X^2 + Y^2 = Z^2$ and $Y = X \tan X$, and they correspond to even-parity solutions. The small circles 'o' represent the intersection between the circles $X^2 + Y^2 = Z^2$ and $Y = -X \cot X$, and they correspond to odd-parity solution.

Note that in the range $0 < \left(2m/\hbar^2 V_0 a^2\right)^{1/2} < \pi/2$, there is no odd-parity solution, and there is only one energy eigen state corresponding to even parity. In the range $\pi/2 < \left(2m V_0 a^2/\hbar^2\right)^{1/2} < \pi$, there are two energy eigen states: one solution corresponding to even parity and another solution corresponding to odd parity. The wave functions for both states are sketched in Fig. 3.16. In the range

$\pi < \left(2mV_0a^2/\hbar^2\right)^{1/2} < 3\pi/2$, there are three energy eigen states. Two energy eigenstates are of even parity and one energy eigenstate is of odd parity.

We can easily determine the number of possible energy eigenvalues for a given V_0 and a. It is important to recognize the number of possible energy eigenvalues are finite for a given V_0 and a. In fact, it is obvious from the above discussion; for a given potential V_0 which is such that $\left(2mV_0a^2/\hbar^2\right)^{1/2}$ lies in the range between $N\pi/2$ and $(N+1)\pi/2$, there are $(N+1)$ energy eigenvalues.

Even-parity Wave Function u(x)

In the case of even-parity wave function, the constant A in Equations (3.95),(3.96) and (3.97) can be determined in terms of C using (3.98), and it is given by

$$A = Ce^{\alpha a}\cos\beta a \tag{3.111}$$

Therefore, the even-parity wave function is given by

$$u_n(x) = \begin{cases} Ce^{\alpha a}\cos\beta_n a\ e^{\alpha x} & x < -a & (3.112)\\ C\cos\beta_n x & -a < x < a & (3.113)\\ Ce^{\alpha a}\cos\beta_n a\ e^{-\alpha x} & x > a & (3.114)\end{cases}$$

Odd-parity Wave Function

The constant A in equation can be determined in terms of D using (3.106), and it is given by

$$A = -De^{\alpha a}\sin\beta a \tag{3.115}$$

Therefore, the odd-parity wave function is given by

$$V_n(x) = \begin{cases} -D\ e^{\alpha a}\sin\beta_n a\ e^{\alpha x} & x < -a\\ D\ \sin\beta_n x & -a < x < a\\ D\ e^{\alpha a}\sin\beta_n a\ e^{-\alpha x} & x > a\end{cases} \tag{3.116}$$

These functions can be sketched easily. We have to bear in mind in which quadrant $\beta_n a$ lies. Correspondingly, the sign of coefficient of the exponential functions will become positive or negative. This is due to the fact $\cos\beta_n a$ and $\sin\beta_n a$ can become a positive or a negative number.

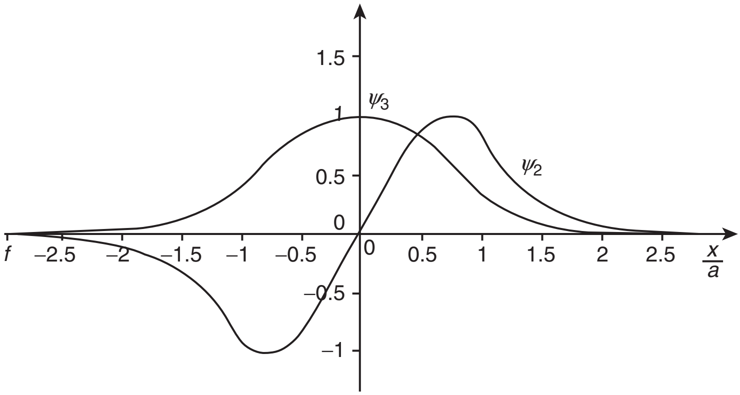


Fig. 3.17 Sketch of bound state wave functions for a particle in square-well potential

Particle in Classically Forbidden Region

We have already seen that the classically allowed region is when ($E - V$) is positive in a particular region. In the case of square-well potential, for bound states, we have

Region I $E - V = E = -|E|$; $E - V$ is negative; classically forbidden

Region II $E - V = V_0 - |E|$; $E - V$ is positive; classically allowed

Region III $E - V = E = -|E|$; $E - V$ is negative; classically forbidden

So, according to classical physics, the particle should not be found in region I and III. In quantum mechanics, the wave function is non-zero (though exponentially falls to zero) in regions I and III. Therefore, there is a definite probability of finding the particle in regions I and III. So the particles can be found in these regions also.

Example 3.7 Normalize the wave function for the bound state wave functions for a square-well potential.

Solution: The even-parity wave function is

$$u_n(x) = \begin{cases} -C\, e^{\alpha a} \cos \beta_n a\, e^{\alpha x} & x < -a \\ C \cos\beta_n x & -a < x < a \\ C\, e^{\alpha a} \cos\beta_n a\, e^{-\alpha x} & x > a \end{cases}$$

The odd-parity wave function is

$$V_n(x) = \begin{cases} -D\, e^{\alpha a} \sin \beta_n a\, e^{\alpha x} & x < -a \\ D \sin\beta_n x & -a < x < a \\ D\, e^{\alpha a} \sin\beta_n a\, e^{-\alpha x} & x > a \end{cases}$$

Condition for the normalization of even-parity wave function is $\int\limits_{-\infty}^{\infty} |u(x)|^2\, dx = 1$.

$$\int\limits_{-\infty}^{-a} |u_{\mathrm{I}}(x)|^2\, dx + \int\limits_{-a}^{a} |u_{\mathrm{II}}(x)|^2\, dx + \int\limits_{a}^{\infty} |u_{\mathrm{III}}(x)|^2\, dx = 1$$

$$\int\limits_{-\infty}^{-a} C^2 e^{2\alpha a} \cos^2 \beta a e^{2\alpha x} dx + \int\limits_{-a}^{a} C^2 \cos^2 \beta x dx + \int\limits_{a}^{\infty} C^2 e^{2\alpha a} \cos^2 \beta a e^{-2\alpha x} dx = 1$$

$$C^2 e^{2\alpha a} \cos^2 \beta a \left.\frac{e^{2\alpha x}}{2\alpha}\right|_{-\infty}^{-a} + C^2 \left.\left(\frac{1}{2} + \frac{\sin 2\beta x}{4\beta}\right)\right|_{-a}^{a} + C^2 e^{2\alpha a} \cos^2 \beta a \left.\frac{e^{-2\alpha x}}{-2\alpha}\right|_{a}^{\infty} = 1$$

$$C^2 \left[\frac{\cos^2 \beta a}{\alpha} + a + \frac{1}{\beta} \sin \beta a \cos \beta a\right] = 1$$

Since $\tan \beta a = \dfrac{\alpha}{\beta}$, we have

$$\sin \beta a = \frac{\alpha}{\sqrt{\alpha^2 + \beta^2}} \quad \text{and} \quad \cos \beta a = \frac{\beta}{\sqrt{\alpha^2 + \beta^2}}$$

Making use of these relations, we get

$$A^2\left(\frac{1}{\alpha}+a\right)=1$$

or

$$A=\left(\frac{\alpha}{1+\alpha a}\right)^{1/2}=\frac{1}{\sqrt{a}}\left(\frac{\alpha a}{1+\alpha a}\right)^{1/2}=\frac{1}{\sqrt{a}}\cdot\left(\frac{1}{\frac{1}{\alpha a}+1}\right)^{1/2}$$

In the same way, the normalization constant *B* for odd-parity wave function can be determined. *B* is found to be the same as *A*.

APPENDIX

Potentials and the Choice of the Coordinate Systems

To analyze a physical system using Schrödinger equation in non-relativistic equation, we employ a coordinate system to suit our convenience. This is possible since, in general, the physical results like energy eigenvalues or probability density should not depend on the choice of the coordinate system. We shall study this aspect in details in Chapter 9. Let us study the case of a particle in a box in this context.

In Section 3.7, we considered the particle in a box whose potential is

$$V(x)=\begin{cases}\infty & x<-\frac{a}{2}\\ 0 & -\frac{a}{2}<x<\frac{a}{2}\\ \infty & x<\frac{a}{2}\end{cases}\tag{3.117}$$

It is very common to use the potential given below to consider the particle in a box.

$$V(x)=\begin{cases}\infty & \text{for} \quad x<0\\ 0 & \text{for} \quad 0<x<a\\ \infty & \text{for} \quad x>a\end{cases}\tag{3.118}$$

Both potentials (3.117) and (3.118) describe the particle in a one-dimensional box of size *a*. The difference between them is that the potential (3.117) is an even function of *x* while the potential (3.118) is neither odd or even. This difference is due to the different choices in the coordinate systems. In the case of potential (3.117), the origin is at the centre of the box, while in the case of potential (3.118), the origin is at the left edge of the box. Let us verify that the potential (3.118) leads to the same physical conclusions.

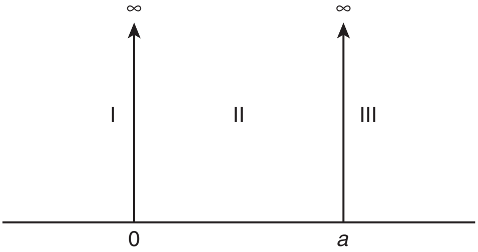


Fig. 3.18 Particle in a box – asymmetric potential

The Schrödinger equation is

$$\frac{d^2u}{dx^2}+\frac{2mE}{\hbar^2}u=0 \quad 0<x<a \tag{3.119}$$

Since the potential $V(x)$ is infinite in the regions I and III, we cannot find the particle in these regions. These conclusions imply that the wave function $u(x)$ is zero in regions I and III.

i.e.,

$$u(x)=0 \quad \text{for} \quad x<0$$
$$u(x)=0 \quad \text{for} \quad x>a$$

So we have Schrödinger's differential equation only in region II.

Define $$k^2=\frac{2mE}{\hbar^2}$$

The Equation (3.119) becomes

$$\frac{d^2u}{dx^2}+k^2u(x)=0$$

This differential equation has to be solved with the boundary conditions which arise from the requirement that the wave function $u(x)$ should be single valued at all points.

i.e., $$u(0)=0 \quad \text{and} \quad u(a)=0$$

The solution to the equation is given by

$$u(x)=A\sin kx+B\cos kx$$

The boundary condition $u(0)=0$ implies that $B=0$.

$$\therefore\ u(x)=A\sin kx.$$

The boundary condition $u(a)=0$ implies that

$$u(a)=A\sin ka=0 \quad \text{or} \quad ka\ n\pi \quad \text{or} \quad k=\frac{n\pi}{a}$$

It is more appropriate to write it as $k_n=\frac{n\pi}{a}$

Then, the wave function $u_n(x)$ is

$$u_n(x)=A\sin k_nx=A\sin\frac{n\pi x}{a}$$

The corresponding energy eigenvalues are given by

$$E_n=\frac{\hbar^2k_n^2}{2m}=\frac{\hbar^2n^2\pi^2}{2ma^2}$$

The Energy Eigenvalues and the Energy Eigenfunctions

The energy eigenvalues and the energy eigenfunctions are given by

$$E_n=\frac{\hbar^2n^2\pi^2}{2ma^2} \quad \text{and} \quad u_n(x)=\sqrt{\frac{2}{a}}\sin\frac{n\pi x}{a}$$

(See examples for the normalization constant.) Some of the energy eigenvalues and the corresponding energy eigenfunctions are listed below:

$$n=1\ \ E_1=\frac{\hbar^2\pi^2}{2ma^2} \qquad u_1(x)=\sqrt{\frac{2}{a}}\sin\frac{\pi x}{a} \tag{3.120}$$

$$n=2 \quad E_2 = \frac{\hbar^2\pi^2}{2ma^2} \qquad u_2(x) = \sqrt{\frac{2}{a}}\sin\frac{2\pi x}{a} \tag{3.121}$$

$$n=3 \quad E_3 = \frac{\hbar^2 9\pi^2}{2ma^2} \qquad u_3(x) = \sqrt{\frac{2}{a}}\sin\frac{3\pi x}{a} \tag{3.122}$$

$$n=4 \quad E_4 = \frac{\hbar^2 16\pi^2}{2ma^2} \qquad u_4 = \sqrt{\frac{2}{a}}\sin\frac{4\pi x}{a} \tag{3.123}$$

The sketches of these eigenfunctions are shown in Fig. 3.19. Compare the energy eigenvalues in Equations (3.120) – (3.123) with those in Equations (3.85) – (3.90). They are exactly the same. Though the functional form of the eigenfunctions given in Equations (3.85) – (3.90) is different from those given in (3.120) – (3.123), their sketches in Fig. 3.14 and Fig. 3.19 are the same, and so they will lead to the same physical consequences.

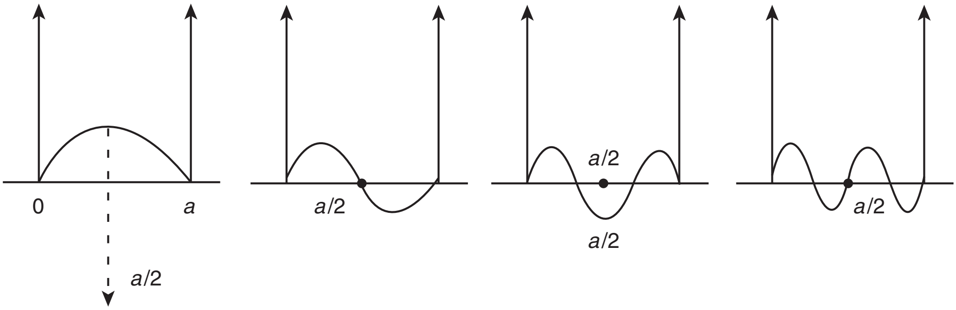


Fig. 3.19 Sketch of wave functions for a particle in a box (asymmetric potential)

Example 3.8 Normalize the eigenfunctions for a particle in a box.

Solution: The wave function for a particle in a box is $u_n(x) = A\sin\frac{n\pi x}{a}$

The normalization of $u_n(x)$ implies. $\int_{-\infty}^{\infty} |u_n(x)|^2\, dx = 1$

$$\int_{-\infty}^{0} |u_n(x)|^2\, dx + \int_{0}^{a} |u_n(x)|^2\, dx + \int_{a}^{\infty} |u_n(x)|^2\, dx = 1$$

Since $u_n(x)$ is zero in both regions I and II, the first and the third integrals are zero. Therefore, we have

$$\int_0^a |u_n(x)dx| = 1 \qquad \text{or} \qquad \int_0^a |A|^2 \sin^2\frac{n\pi x}{a}\, dx = 1$$

$$\int_0^a |A|^2 \frac{1}{2}\left(1-\cos\frac{2n\pi x}{a}\right) dx = \frac{a}{2}|A|^2 = 1 \qquad \text{or} \quad A = \sqrt{\frac{2}{a}}$$

The normalized wave function is

$$u_n(x) = \sqrt{\frac{2}{a}}\sin\frac{n\pi x}{a}$$

Example 3.9 Show that $\int_0^a u_n^*(x)u_m(x)dx = \delta_{nm}$ for a particle in a box where $u_n(x) = \sqrt{\frac{2}{a}}\sin\frac{n\pi x}{a}$ and

$$\delta_{nm} = \begin{cases} 1 & \text{if} \quad n = m \\ 0 & if \quad n \neq m \end{cases}$$

Solution:

$$\int_0^a u_n^*(x)u_m(x)dx = \int_0^a \frac{2}{a}\sin\frac{n\pi x}{a}\sin\frac{m\pi x}{a}dx = \frac{2}{a}\int_0^a \frac{1}{2}\left[\cos\frac{(n-m)\pi x}{a} - \cos\frac{(n+m)\pi x}{a}\right]$$

$$= \frac{1}{a}\cdot\left[\frac{a}{(n-m)}\sin\frac{(n-m)\pi x}{a} - \frac{a}{(n+m)}\sin\frac{(n+m)\pi x}{a}\right]_0^a \quad \text{provided } n \neq m = 0$$

Let us now consider when $n = m$. In that case, we have

$$\int_0^a u_n^*(x)u_m(x)dx = \int_0^a |u_n(x)|^2\, dx = 1$$

$$\int_0^a u_n^*(x)u_m(x)dx = \begin{cases} 0 & \text{if} \quad m \neq n \\ 1 & \text{if} \quad m = n \end{cases}$$

$$= \delta_{mn}$$

EXERCISES

1. A part of the incident beam is reflected by a potential. Is it a consequence of particle property or wave property?
2. What is meant by classically allowed and classically forbidden region?
3. What are the classically allowed and the classically forbidden regions for the following?
    (i) Step potential
    (ii) Square-well potential
    (iii) Rectangular barrier potential
4. Show that there is a classically forbidden region for α decay.
5. In a classically forbidden region, a particle cannot be present since $p = \sqrt{2m(E-V)}$ is imaginary. Yet quantum mechanics allows the particle in this region. How is this possible? Is there a contradiction?
6. For a particle in a box, why are we not implementing the boundary condition $\frac{\partial u}{\partial x}$ is single valued and continuous at the edges?
7. What is meant by translational invariance?
8. Two coordinate systems S and S' are employed to obtain the wave function of a system. S' is obtained from S by shifting the origin by ξ along x axis. How are the wave functions in the two coordinate systems related?
9. A particle is bound to the square-well potential, and its energy eigenfunctions are given by $u_n(x)$. Change the coordinate system so that the particle is bound in the region $0 < x < 2a$, and now the energy eigenfunctions are given by $V_n(x)$. Show that $V_n(x) = u_n(x - a)$.

10. A particle is bound to a box in a region defined by $0 < x < a$. Its energy eigenfunction is given by $u_n(x)$. Change the coordinate system where origin is at the midpoint of the box, and now its energy eigenfunctions are $V_n(x)$. Show that $V_n(x) = u_n\left(x + \frac{a}{2}\right)$.

11. What is meant by parity of the wave function?

12. Does the wave function of a particle has definite parity for all potentials?

13. Does the wave function of a particle depend on the coordinate system?

14. A particle is confined to a box in the region $0 < x < a$. Do the eigenfunctions have definite parity for such a particle?

15. For a particle bound in a square-well potential, the width $2a$ and the potential V_0 are such that $\left(\frac{2mV_0a^2}{\hbar^2}\right)^{1/2}$ is 32. How many energy eigenstates are there? How many of them are of even parity? What will be the sign of the coefficient of exponential function for these eigenstates?

REFERENCES

1. S. Gasiorowicz, 1974. Quantum Physics, New York: John Wiley & Sons.
2. P. M. Mathews and A. Venkatesh, *A Text Book of Quantum Mechanics*, New Delhi: Tata McGraw-Hill Publishing Company Limited.
3. B. H. Bransden and C. J. Joachain, 1989. *Introduction to Quantum Mechanics*. English Language Book Society. England: Longman.
4. J. Greensite, *Lecture Notes on Quantum Mechanics*. http://www.physics.sfsu.edu/~greensite/book.pdf
5. Jimwiss, http://online.physics.uiuc.edu/courses/phys386/fall03/Lectures/tunnelling.pdf

4

Mathematical Preliminaries

4.1 LINEAR VECTOR SPACE

The concept of linear vector space is the backbone of quantum mechanics. The important feature of quantum mechanics is the principle of superposition. If ψ_1 and ψ_2 are two states of a system, then there exists a state corresponding to $\psi = c_1\psi_1 + c_2\psi_2$, where c_1 and c_2 are complex numbers. It is the concept of linear vector space which expresses the idea of superposition. The simple concept of linear vector space is not sufficient in quantum mechanics. The structure of quantum mechanics requires imposition of additional properties on linear vector space. More particularly, quantum mechanics requires the linear vector space to be Hilbert space. The simplest way of expressing the Hilbert space is that it is a linear vector space with an inner product.

First, let us review the basic ideas of linear vector space. Let us start with a collection of geometric vectors-quantities with direction and magnitude. The best examples are position vector, velocity vector, force vector and so on. If $\mathbf{V}_1$ and $\mathbf{V}_2$ are two vectors representing velocity, then we can define a new vector $c_1\mathbf{V}_1 + c_2\mathbf{V}_2$ which also represents another velocity vector $\mathbf{V}$.

i.e.,

$$\mathbf{V} = c_1\mathbf{V}_1 + c_2\mathbf{V}_2 \tag{4.1}$$

Any vector $\mathbf{V}$ can be expressed in terms of unit vectors $\mathbf{i}, \mathbf{j}$ and $\mathbf{k}$ as follows:

$$\mathbf{V} = \mathrm{V}_1\,\mathbf{i} + \mathrm{V}_2\,\mathbf{j} + \mathrm{V}_3\,\mathbf{k} \tag{4.2}$$

For a given pair of vectors **A** and **B**, we can assign a number$(\mathbf{A}, \mathbf{B})$. In vector algebra, the most useful way of assigning such a number is the usual dot product.

$$(\mathbf{A}, \mathbf{B}) = \mathbf{A} \cdot \mathbf{B} = AB\cos\theta \tag{4.3}$$

Note that this expression is only a definition. One has to recognize the fact that there are many ways of assigning a number for any two vectors. However, within the vector algebra, the definition given in the Equation (4.3) is one of the most profitable definitions, especially in the context of classical physics and engineering. This definition satisfies the following properties:

$$\mathbf{A}\cdot\mathbf{B}=\mathbf{B}\cdot\mathbf{A}$$

$$\mathbf{A}\cdot(c_1\mathbf{B}+c_2\mathbf{C})=c_1\mathbf{A}\cdot\mathbf{B}+c_2\mathbf{A}\cdot\mathbf{C}$$

and $$\mathbf{A}\cdot\mathbf{A}\geq 0 \text{ with } \mathbf{A}\cdot\mathbf{A}=0 \text{ imply } \mathbf{A}=0 \tag{4.4}$$

We have just summarized a few important properties of geometrical vectors in order to develop the basic concepts of linear vector space.

The next thing to do is to leave the fact that the vector is a quantity with magnitude and direction and concentrate on some abstract properties satisfied by the geometrical vectors. In particular, we have to focus on the ideas of closure property, linear combinations, basis, dimension, scalar product and so on, and we will extend these ideas to other mathematical objects.

In what follows, we present a brief introduction to linear vector space.

A linear vector space L is a collection of (or a set of) objects $\{\phi_1,\phi_2,\phi_3,\ldots\}$ called vectors with an operation of addition, providing a rule to add two elements, and an operation of scalar multiplication, providing a rule to multiply a vector by scalar (real or complex number), which obey the following axioms:

1. For every ϕ_1 and ϕ_2, which are elements of L, we can define an addition operator + such that

   $$\phi_1+\phi_2=\phi=\phi_2+\phi_1$$

   where ϕ should also be an element of L. This is called closure property.
2. Addition is associative. $\phi_1+(\phi_2+\phi_3)=(\phi_1+\phi_2)+\phi_3$
3. There exists an element 0 of L called a null vector such that

   $$\phi+0=0+\phi=\phi$$
4. If c is a real or a complex number, then the multiplication of vectors by a scalar is distributive

   $$c(\phi_1+\phi_2)=c\phi_1+c\phi_2$$
5. Multiplication of vectors by scalars is distributive with respect to scalars.

   $$(c_1+c_2)\phi=c_1\phi+c_2\phi$$
6. Associative property with respect to scalar multiplication: $c_1(c_2\phi)=(c_1c_2)(\phi)$.
7. $$1\cdot\phi=\phi\cdot 1=\phi$$
8. $$0\cdot\phi=\phi\cdot 0=0 \tag{4.5}$$

It is very easy to see that all these axioms are satisfied by geometrical vectors $\mathbf{V}_1$, $\mathbf{V}_2$, $\mathbf{V}_3\ldots$. In fact, these axioms have been formulated from the study of geometrical vectors. However, they are also satisfied by a number of other mathematical objects. This can be seen by the list of examples given below for such linear vector space.

Linear vector space of column vectors with two elements:
The typical elements of this linear vector space are

$$\phi_1=\begin{pmatrix}a_1\\b_1\end{pmatrix},\phi_2=\begin{pmatrix}a_2\\b_2\end{pmatrix}\ldots.\ \phi=\begin{pmatrix}a\\b\end{pmatrix}$$

where a, a_1, a_2,… and b, b_1, b_2,… are either real or complex numbers.

We can see that $\phi_1+\phi_2=\begin{pmatrix}a_1\\b_1\end{pmatrix}+\begin{pmatrix}a_2\\b_2\end{pmatrix}=\begin{pmatrix}a_1+a_2\\b_1+b_2\end{pmatrix}$ which is again a column vector. So the closure property is satisfied.

The null vector of this linear vector space is $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$. It can be easily seen that

$$\begin{pmatrix} a \\ b \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} + \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} a \\ b \end{pmatrix}$$

It can be multiplied by a real number like, say, 5.

$$5\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 5a \\ 5b \end{pmatrix}$$

In fact, the verification of all these axioms by these sets of column vectors is trivial.

Linear vector space of continuous functions.

Let $$L = \{f_1(x), f_2(x), \ldots\}$$

It is obvious to see that the sum of two continuous functions $f_1(x)$ and $f_2(x)$ is again a continuous function.

$$f(x) = f_1(x) + f_2(x)$$

If $f(x)$ is the result of the addition, then $f(x)$ is also a continuous function.

Again, it is very easy to verify that all the axioms of a linear vector space are satisfied by the set of continuous functions.

Linear vector space of polynomials of degree n

For the sake of clarity, we choose $n = 2$, that is, polynomials of degree 2. Then L is given by

$$L = \{a_1 + a_2x + a_3x^2, b_1 + b_2x + b_3x^2, \ldots\}$$

It is obvious that the sum of two polynomial of degree 2 is again a polynomial of degree 2. Let $\phi_1 = a_1 + a_2x + a_3x^2$ and $\phi_2 = b_1 + b_2x + b_3x^2$. Then we have

$$\underbrace{(a_1 + a_2x + a_3x^2)}_{\phi_1} + \underbrace{(b_1 + b_2x + b_3x^2)}_{\phi_2} = \underbrace{(a_1 + b_1) + (a_2 + b_2)x + (a_3 + b_3)x^2}_{\phi_3}$$

The null element corresponds to $a = b = c = 0$. It is easy to verify that all the axioms of a linear vector space are satisfied by the set of polynomials of degree 2.

Note: The term *vector* no longer refers to an object with magnitude and direction.

Any object which satisfies all the eight axioms of a linear vector space is called a vector.

Thus, a vector may be $\phi = \begin{pmatrix} a \\ b \end{pmatrix}$ or $\phi = f(x)$ or $\phi = a + bx + cx^2$ etc

Let us briefly explain the other concepts of a linear vector space. For a detailed discussion, the reader can refer to a book on linear algebra. The important concepts associated with a linear vector space are (i) linear combination (ii) linear independence (iii) basis (iv) dimension of a linear vector space and (v) complete set of linearly independent vectors.

Linear Combination

If ϕ_1, ϕ_2,… ϕ_n are elements of the linear vector space, then we can form a new vector ϕ by

$$\phi = c_1\phi_1 + c_2\phi_2 ... \tag{4.6}$$

where c_1, c_2, c_3,… are complex numbers.

This is called a linear combination of the vectors $\phi_1, \phi_2, \ldots \phi_n$.

Linear Independence

The vectors ϕ_1, ϕ_2 … are said to be linearly independent if it is not possible to find non-zero scalars c_1, c_2,… c_n such that the linear combination

$$c_1\phi_1 + c_2\phi_2 + ... c_n\phi_n = 0 \tag{4.7}$$

The simplest example is the set $\{\mathbf{i}, \mathbf{j}, \mathbf{k}\}$. Let us write the linear combination

$$c_1\mathbf{i} + c_2\mathbf{j} + c_3\mathbf{k} = 0$$

This is possible only if $c_1 = c_2 = c_3 = 0$. Therefore, the vectors $\mathbf{i}, \mathbf{j}, \mathbf{k}$ are linearly independent.

Another example is $\{\phi_1 = 1, \phi_2 = x$ and $\phi_3 = x^2\}$. The linear combination of these vectors is $\phi = a + bx + cx^2$. If we demand $a + bx + cx^2 = 0$ for all x, the only possible solution is $a = b = c = 0$. Therefore, the set $\{a, x, x^2\}$ forms linearly independent vectors.

Basis

The linearly independent vectors e_1, e_2, e_3,… e_n form a basis of a linear vector space provided any vector ϕ of the linear vector space can be expressed as a linear combination of the vectors e_1, e_2,… e_n.

i.e.,
$$\phi = c_1e_1 + c_2e_2 + c_3e_3 + \ldots c_ne_n$$

For example $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ and $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ are the basis vectors for the linear vector space of 2 × 1 column vectors. Any vectors $\begin{pmatrix} a \\ b \end{pmatrix}$ can be expressed as

$$\begin{pmatrix} a \\ b \end{pmatrix} = a\begin{pmatrix} 1 \\ 0 \end{pmatrix} + b\begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Dimension

A linear vector space L is said to be of dimension n if the maximum number of linearly independent vectors in that space is n.

Continuing the example of linear vector space of 2 × 1 column vectors, let us choose three vectors, say, $\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ and $\begin{pmatrix} 2 \\ 3 \end{pmatrix}$ for which we can find a linear combination with non-zero values for c_1, c_2 and c_3 such that

$$c_1\begin{pmatrix} 1 \\ 0 \end{pmatrix} + c_2\begin{pmatrix} 0 \\ 1 \end{pmatrix} + c_3\begin{pmatrix} 2 \\ 3 \end{pmatrix} = 0$$

i.e., one solution is $c_1 = c_2 = c_3 = 0$; but there also exist solutions with $c_1 \neq 0, c_2 \neq 0, c_3 \neq 0$ (for instance, $(c_1 = 6, c_2 = 9, c_3 = -3)$.

Therefore, of the three vectors $\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}$ and $\begin{pmatrix}2\\3\end{pmatrix}$, only two are independent.

Note that $\begin{pmatrix}1\\0\end{pmatrix}$ and $\begin{pmatrix}0\\1\end{pmatrix}$ are not the only set of linearly independent vectors in L. We can form an infinite number of such pairs. For example, $\begin{pmatrix}1\\1\end{pmatrix}$ and $\begin{pmatrix}1\\-1\end{pmatrix}$ are a pair of two linearly independent vectors. Any vector $\begin{pmatrix}a\\b\end{pmatrix}$ of L can be expressed as a linear combination of $\begin{pmatrix}1\\1\end{pmatrix}$ and $\begin{pmatrix}1\\-1\end{pmatrix}$.

$$\begin{pmatrix}a\\b\end{pmatrix} = c_1\begin{pmatrix}1\\1\end{pmatrix} + c_2\begin{pmatrix}1\\-1\end{pmatrix} \text{ where } c_1 = \frac{a+b}{2} \text{ and } c_2 = (a-b)/2$$

However, of the four vectors $\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}, \begin{pmatrix}1\\1\end{pmatrix}$ and $\begin{pmatrix}1\\-1\end{pmatrix}$, only two are independent. So the linear vector space of 2×1 column vectors is a two-dimensional linear vector space.

Complete Set

The set of n vectors $\{\phi_1, \phi_2, \ldots \phi_n\}$ forms a complete set in a linear vector space L, provided they satisfy the following conditions:

1. They are linearly independent.
2. The dimension of the linear vector space is n.

This implies that any arbitrary vector ψ of the linear vector space L can be expressed as linear combination of $\phi_1, \phi_2, \ldots \phi_n$.

i.e.,
$$\psi = c_1\phi_1 + c_2\phi_2 + c_3\phi_3 + \ldots c_n\phi_n \tag{4.8}$$

Now let us introduce additional structures to a linear vector space. Depending on the kind of additional structure imposed, a linear vector space can become a metric space, Banach space, Hilbert space and so on. From the point of view of quantum mechanics, we are interested in a particular kind of linear vector space called Hilbert space. This requires the introduction of new concept called scalar product.

4.1.1 Scalar Product

In the cases of geometrical vectors, we have already seen that a scalar product assigns a number for a given pair of vectors **A** and **B** through the dot product $\mathbf{A}\cdot\mathbf{B} = AB\cos\theta$. In the same way, the concept of scalar product has to be extended to other linear vector spaces. We have to bear in mind that the term *vector* is no longer restricted to quantities with direction and magnitude only. Let us introduce the following notation to denote a scalar product:

Instead of writing $\psi\,\phi$ for the scalar product between two vectors ψ and ϕ, we write it as (ψ,ϕ). In general, the scalar product(ψ,ϕ) is a complex number. Mathematically, it is said that the scalar product (ψ,ϕ) maps the vectors ψ and ϕ into a complex number. That is, it assigns a complex number for a given pair of vectors ψ and ϕ. This is shown in Fig. 4.1.

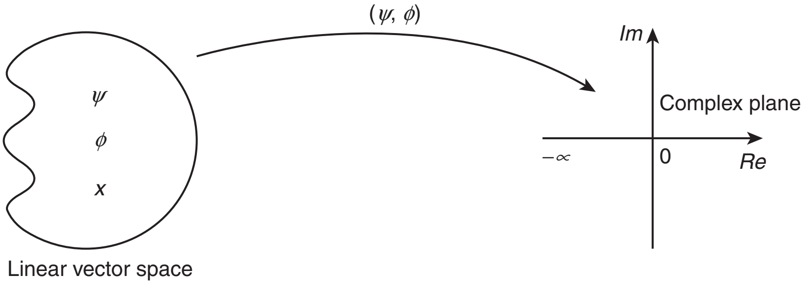


Fig. 4.1 Scalar product as a mapping from a linear vector space to complex number

The scalar product between two elements ψ and ϕ has to satisfy the following properties:

$$(\psi,\phi)=(\phi,\psi)^*$$
$$(\psi,c_1\,\phi_1+c_2\,\phi_2)=c_1(\psi_1,\phi_1)+c_2(\psi,\phi_2)$$
$$(\psi,\psi)\geq 0 \text{ with } (\psi,\psi)=0 \text{ implying } \psi=0 \tag{4.9}$$

There are no special reasons to demand these particular conditions on a scalar product, except for the fact that in quantum mechanics, only this kind of scalar product is useful.

Example 4.1 Prove that $(c\phi,\psi)=c^*(\phi,\psi)$.

Solution: From the first and second conditions in the Equation (4.9), we have

$$(c\phi,\psi)=(\psi,c\phi)^*=[c(\psi,\phi)]^*=c^*(\phi,\psi).$$

So we have

$$\begin{aligned}(\phi,c\psi)&=c(\phi,\psi)\\(c\phi,\psi)&=c^*(\phi,\psi)\end{aligned} \tag{4.10}$$

Norm of a Vector ψ

In analogy with geometrical vectors (where $\mathbf{A}\cdot\mathbf{A}=\|\mathbf{A}\|^2$), we define the norm of a vector ψ (we reiterate again that ψ is no longer an object with magnitude and direction) denoted by $\|\psi\|$ as

$$\|\psi\|^2=(\psi,\psi) \tag{4.11}$$

Orthogonality

Two vectors ψ and ϕ are said to be orthogonal to each other if their scalar product

$$(\psi,\phi)=0 \tag{4.12}$$

Scalar Product on Linear Vector Space of Continuous Functions

Let $\phi(x)$ and $\psi(x)$ be two functions which are continuous in the interval $[a, b]$. Then we can define the scalar product between $\phi(x)$ and $\psi(x)$ as

$$(\phi,\psi) = \int_a^b \phi^*(x)\psi(x)dx$$

Note that this definition satisfies all the conditions given in (4.9).

$$\begin{aligned}
(\phi,\psi) &= \int_a^b \phi^*(x)\psi(x)dx = \left[\int_a^b \psi^*(x)\phi(x)dx\right]^* \\
&= (\psi,\phi)^* \qquad \text{Condition 1 is satisfied} \\
(\psi, c_1\phi_1 + c_2\phi_2) &= \int_a^b \psi^*(x)[c_1\phi(x) + c_2\phi_2(x)]dx \\
&= c_1\int_a^b \psi^*(x)\phi_1(x)dx + c_2\int_a^b \psi^*(x)\phi_2(x)dx \\
&= c_1(\psi,\phi_1) + c_2(\psi,\phi_2) \qquad \text{Condition 2 is satisfied} \\
\text{Lastly,} (\psi,\psi) &= \int_a^b \psi^*(x)\psi(x)dx.
\end{aligned}$$

Since the integrand $\psi^*(x)\psi(x)$ is always positive, if we demand (ψ,ψ) to be zero, the only possibility is $\psi(x) = 0$ for all x in the interval $[a, b]$.

Scalar Product on Linear Vector Space of Column Vectors

Let ϕ and ψ be two column vectors given by $\phi = \begin{pmatrix} a \\ b \end{pmatrix}$, and $\psi = \begin{pmatrix} c \\ d \end{pmatrix}$.

We can define the scalar product (ϕ,ψ) as $(\phi,\psi) = \phi^\dagger\psi = (a^* b^*)\begin{pmatrix} c \\ d \end{pmatrix} = a^*c + b^*d$.

It is very easy to verify that this definition satisfies all the three axioms given (4.9). Take the first condition:

$$(\phi,\psi) = (a^*c + b^*d) = (ac^* + bd^*)^* = \left[(c^* d^*)\begin{pmatrix} a \\ b \end{pmatrix}\right]^* = (\psi\ \phi)^* = (\psi,\phi)^*.$$

Let us check the second condition.

Let $$\psi = \begin{pmatrix} a \\ b \end{pmatrix}, \phi_1 = \begin{pmatrix} e_1 \\ f_1 \end{pmatrix} \text{ and } \phi_2 = \begin{pmatrix} e_2 \\ f_2 \end{pmatrix}$$

Then
$$c_1\phi_1 + c_2\phi_2 = c_1\begin{pmatrix} e_1 \\ f_1 \end{pmatrix} + c_2\begin{pmatrix} e_2 \\ f_2 \end{pmatrix} = \begin{pmatrix} c_1e_1 + c_2e_2 \\ c_1f_1 + c_2f_2 \end{pmatrix}$$

$$\begin{aligned}
(\psi, c_1\phi_1 + c_2\psi_2) &= (a^* \; b^*)\begin{pmatrix} c_1e_1 + c_2e_2 \\ c_1f_1 + c_2f_2 \end{pmatrix} \\
&= a^*(c_1e_1 + c_2e_2) + b^*(c_1f_1 + c_2f_2) \\
c_1(\psi, \phi_1) + c_2(\psi, \phi_2) &= c_1\psi \ \phi_1 + c_2\psi \ \phi_2 \\
&= c_1(a^* \; b^*)\begin{pmatrix} e_1 \\ f_1 \end{pmatrix} + c_2(a^* \; b^*)\begin{pmatrix} e_2 \\ f_2 \end{pmatrix} \\
&= c_1a^*e_1 + c_1b^*f_1 + c_2a^*e_2 + c_2b^*f_2 \\
\therefore (\psi, c_1\phi_1 + c_2\phi_2) &= c_1(\psi, \phi_1) + c_2(\psi, \phi_2).
\end{aligned}$$

Consider (ψ, ψ)

$$\begin{aligned}
(\psi, \psi) = \psi^\dagger\psi &= (a^* \; b^*)\begin{pmatrix} a \\ b \end{pmatrix} = a^*a + b^*b \\
&= |a|^2 + |b|^2
\end{aligned}$$

(ψ, ψ) can be zero only if $a = 0$ and $b = 0$.

4.2 HILBERT SPACE

A linear vector space with an inner product is called Hilbert space. Note that this is not the way the mathematicians define a Hilbert space. In mathematics, the Hilbert space is introduced in a complex way. First, metric space and Cauchy sequences in metric space are considered. Then, the Hilbert space is seen as completion (another mathematical concept) of metric space. This way of introducing Hilbert space is more meaningful in the context of infinite dimensional linear vector space. Interested reader can refer to a simple exposition of these concepts by C. I. Tan.

In quantum mechanics, the stress is only on the inner product. Therefore, in quantum mechanics, both finite dimensional and infinite dimensional linear vector spaces with inner product are considered as Hilbert space.

4.2.1 Hilbert Spaces in Quantum Mechanics

Within non-relativistic quantum mechanics, broadly, we concentrate on the following three Hilbert spaces:

Hilbert Space of Bound State Wave Functions or Square Integrable Wave Functions

The Hilbert space $L_2(\mathrm{R}, dx)$ is a collection of square integrable wave functions. It is a collection of wave functions which go to zero as $x \to \pm\infty$. That is, if $\psi(x)$ is an element of $L_2(\mathrm{R}, dx)$, then, we have

$$\psi(x) \to 0 \quad \text{as} \quad x \to \pm\infty \quad \text{and} \quad \int |\psi(x)|^2 dx < \infty$$

The scalar product between two bound state wave functions $\phi(x)$ and $\psi(x)$ is defined as

$$(\phi,\psi) = \int_{-\infty}^{\infty} \phi^*(x)\psi(x)dx \tag{4.13}$$

This can be extended to other higher dimensional configuration spaces. In three-dimensional space, the bound state wave function $\psi(x, y, z)$ has to satisfy the boundary condition

$$\psi(x, y, z) \to 0 \quad \text{as} \quad x \to \pm\infty, \quad y \to \pm\infty, \quad z \to \pm\infty$$

The scalar product between $\phi(\mathbf{r})$ and $\psi(\mathbf{r})$ is defined as

$$(\phi,\psi) = \int \phi^*(\mathbf{r})\psi(\mathbf{r})d^3\mathbf{r} \tag{4.14}$$

In Cartesian coordinates,

$$(\phi,\psi) = \int_{-\infty}^{\infty} \cdot \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \phi^*(x, y, z)\psi(x, y, z)dxdydz \tag{4.15}$$

In spherical polar coordinates,

$$(\phi,\psi) = \int_0^{\infty} r^2 dr \int_0^{\pi} \sin\theta \, d\theta \int_0^{2\pi} d\theta \quad \phi^*(r,\theta,\phi) \quad \psi(r,\theta,\phi) \tag{4.16}$$

Hilbert Space of Square Summable Sequence l_2

The elements of this Hilbert spaces are $\phi = (c_1, c_2, c_3, \ldots c_n, \ldots)$ such that $|c_1|^2 + |c_2|^2 + \cdots |c_n|^2 + \cdots = 1$.

This sequences can be written as a column vector $\phi = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix}$.

The scalar product between $\phi = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix}$ and $\psi = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$ is defined as

$$(\phi,\psi) = \phi^\dagger \psi = (c_1^* c_2^* c_3^* \cdots c_n^*) \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$$

$$= c_1^* b_1 + c_2^* b_2 + c_3^* b_3 + \cdots c_n^* b_n \tag{4.17}$$

The simplest example is two-component spinor

$$\phi = \begin{pmatrix} a \\ b \end{pmatrix} \text{ and } \psi = \begin{pmatrix} c \\ d \end{pmatrix} \text{ or } \phi = \begin{pmatrix} f_1(\mathbf{r}) \\ f_2(\mathbf{r}) \end{pmatrix} \text{ and } \psi = \begin{pmatrix} g_1(\mathbf{r}) \\ g_2(\mathbf{r}) \end{pmatrix}.$$

In the last case,

$$\begin{aligned}(\phi, \psi) &= \int \phi^\dagger \psi \, d^3\mathbf{r} \\ &= \int [f_1^*(\mathbf{r}) \quad f_2^*(\mathbf{r})] \begin{bmatrix} g_1(\mathbf{r}) \\ g_2(\mathbf{r}) \end{bmatrix} d^3\mathbf{r} \\ &= \int [f_1^*(\mathbf{r}) g_1(\mathbf{r}) + f_2^*(\mathbf{r}) g_2(\mathbf{r})] d^3\mathbf{r}\end{aligned}$$

Another important Hilbert space in quantum mechanics is the Hilbert space of Dirac's bra and ket vectors. This is the most widely used Hilbert space in quantum mechanics. We will discuss this Hilbert space in a later chapter.

4.3 OPERATORS

It is an obvious fact that any operator acting on a mathematical object like wave function or a column vector produces a new function or column vector.

$$\mathrm{A}_{op} \psi = \phi$$

A_{op} is the operator acting on ψ whose result is ϕ.

In quantum mechanics, we require something more than the above equation for an operator. The Hilbert space operator A_{op} is defined by $\mathrm{A}_{op} \psi = \phi$, where both ϕ and ψ should be the elements of the same Hilbert space.

For example, if A_{op} acts on ψ, which is an element of $L_2(\mathrm{R}, dx)$ producing ϕ, then ϕ should also be an element of $L_2(\mathrm{R}, dx)$.

i.e.,

$$\mathrm{A}_{op} \psi = \phi$$

where $\int_{-\infty}^{\infty} |\phi(x)|^2 \, dx$ and $\int_{-\infty}^{\infty} |\psi(x)|^2 \, dx$ should be finite.

The second requirement leads to the definition of domain of an operator. We will discuss this aspect of an operator briefly here.

Domain of a Hilbert Space Operator

Consider x_{op} operating on an element of $L_2(\mathrm{R}, dx)$. x_{op} is defined as $x_{op} \psi = x \psi(x)$.

From the definition of Hilbert space operator, we require $\psi(x)$ as well as $x\ \psi(x)$ be square integrable.

i.e.,

$$\int_{-\infty}^{\infty} |\psi(x)|^2 \, dx < \infty \quad \text{and} \quad \int_{-\infty}^{\infty} x^2 |\psi(x)|^2 \, dx < \infty$$

This implies that within $L_2(\mathrm{R}, dx)$ only a subset or subspace will obey these two conditions.

For example, $\psi(x) = \dfrac{1}{(x^2 + a^2)^{1/2}}$ is a square integrable function.

$$\int_{-\infty}^{\infty} |\psi(x)|^2 \, dx = \int_{-\infty}^{\infty} \frac{dx}{(x^2 + a^2)} = \pi$$

On the other hand, $x\psi(x)$ is not a square integrable. This can be easily checked as follows:

$$\int_{-\infty}^{\infty} x^2 |\psi(x)|^2 \, dx = \int_{-\infty}^{\infty} \frac{x^2 dx}{x^2 + a^2}$$

$$\int_{-\infty}^{\infty} \left(1 - \frac{a^2}{x^2 + a^2}\right) dx = \int_{-\infty}^{\infty} dx - a^2 \int_{-\infty}^{\infty} \frac{dx}{x^2 + a^2}$$

$$= \infty - \pi a^2 = \infty$$

Obviously x_{op} is not defined for this function. The wave functions on which x_{op} can act form a subset of entire Hilbert space. The subspace for which x_{op} is defined is called the domain of x_{op}. The domain of x_{op} is smaller than the entire Hilbert space $L_2(\mathrm{R}, dx)$.

So, for each operator A_{op}, we have to specify the domain $D(A)$ of the operator also.

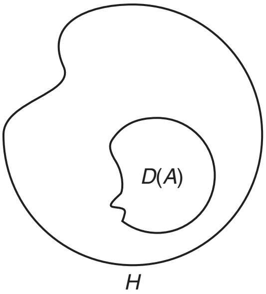


Fig. 4.2 *H* is Hilbert space and *D*(*A*) is domain of A_{op}. *D*(*A*) is a subspace of *H*

$D(A)$ may be the entire space H itself or a proper subspace of H. Strictly speaking, a Hilbert space operator is a pair of $(A_{op}, D(A))$. If B is another operator, the corresponding Hilbert space operator is $(B_{op}, D(B))$. Two Hilbert space operators are the same if

$$A_{op}\phi = B_{op}\phi \text{ for all } \phi \, \varepsilon \, D(A) = D(B)$$

In this case, one writes $A_{op} = B_{op}$.

For details, the reader can refer to a book on mathematical physics. (See Francois Gieres also.)

4.3.1 Linear Operator

An operator A_{op} is said to be a linear operator if it satisfies the following condition:

$$A_{op}(c_1\phi_1 + c_2\phi_2) = c_1 A_{op}\phi_1 + c_2 A_{op}\phi_2 \tag{4.18}$$

This implies that

$$A_{op}(c\phi) = cA_{op}\phi \tag{4.19}$$

Example 4.2 Is $A_{op} = \frac{d}{dx}$ a linear operator?

Solution: Let us check the condition (4.19).

$$\text{LHS}: A_{op}(c\phi) = \frac{d}{dx}(c\phi) = c\frac{d\phi}{dx} \qquad \text{RHS}: cA_{op}\phi = c\frac{d\phi}{dx}.$$

Let us check the condition (4.18).

$$\text{LHS}: A_{op}(c_1\phi_1 + c_2\phi_2) = \frac{d}{dx}(c_1\phi_1 + c_2\phi_2) = c_1\frac{d\phi_1}{dx} + c_2\frac{d\phi_2}{dx}$$

$$\text{RHS}: c_1A_{op}\phi_1 + c_2A_{op}\phi_2 = c_1\frac{d\phi_1}{dx} + c_2\frac{d\phi_2}{dx}$$

$$\therefore \text{LHS} = \text{RHS}$$

Therefore, $\frac{d}{dx}$ is a linear operator.

Example 4.3 Is the operator defined by $A_{op}\psi = \phi^2$ a linear operator?

Solution: Let us check the condition (4.19).

$$\text{LHS: } A_{op}(c\phi) = c^2\phi^2 \quad \text{RHS: } cA_{op}\phi = c\phi^2.$$

Obviously, LHS ≠ RHS.

Therefore, A_{op} is not a linear operator.

4.3.2 Adjoint of an Operator

For every linear operator, another operator called adjoint of A_{op}, denoted by $A_{op}^{\dagger}$, is defined by the relation

$$(\phi, A_{op}\psi) = (A_{op}^{\dagger}\phi, \psi) \quad \text{or} \quad (\phi, A_{op}\psi) = (\psi, A_{op}^{\dagger}\phi)^* \tag{4.20}$$

where ϕ and ψ are elements of the Hilbert space H.

Note $A_{op}^{\dagger}$ need not be the same as A_{op} itself. It is not even complex conjugate of the operator A_{op}. One has to do a detailed calculation to find the adjoint of a given operator.

Example 4.4 Find the adjoint of $A_{op} = \frac{d}{dx}$ where A_{op} acts on the elements of the Hilbert space of bound states.

Solution: Note that since ϕ and ψ are elements of the Hilbert space of bound states, the functions $\phi(x)$ and $\psi(x)$ are such that

$$\phi(x) \to 0 \quad \text{as} \quad x \to \pm\infty$$

$$\psi(x) \to 0 \quad \text{as} \quad x \to \pm\infty$$

$$
\begin{aligned}
(\phi,\ A_{op}\psi) &= \int_{-\infty}^{\infty} \phi^*(x)\frac{d}{dx}\psi(x)dx \\
&= [\phi^*(x)\psi(x)]_{-\infty}^{\infty} - \int_{-\infty}^{\infty} \frac{d\phi^*(x)}{dx}\psi(x)dx \\
&= \phi^*(\infty)\psi(\infty) - \phi^*(-\infty)\psi(-\infty) - \int_{-\infty}^{\infty} \frac{d\phi^*}{dx}\psi dx
\end{aligned}
$$

But since ϕ and ψ are bound state wave functions, the first two terms are zero.

$$
\begin{aligned}
\therefore (\phi, A_{op}\psi) &= -\int_{-\infty}^{\infty} \frac{d\phi^*}{dx}\psi\, dx = \int_{-\infty}^{\infty} \left[\left(\frac{-d}{dx}\right)\phi\right]^* \psi dx \\
&= \left(\left(-\frac{d}{dx}\right)\phi, \psi\right) = (A_{op}^{\dagger}\phi, \psi)
\end{aligned}
$$

$$
\therefore \qquad A_{op}^{\dagger} = -\frac{d}{dx}.
$$

In the case of Hilbert space of $m \times 1$ column vectors, the linear vector space is a finite-dimensional linear vector space. In such cases, if A is the matrix operator, $A^{\dagger}$ is the conventional Hermitian conjugate. This can be easily seen as follows:

Let ψ and ϕ be

$$
\psi = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{pmatrix} \qquad \phi = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}
$$

Let A be the $m \times m$ matrix operator. Then

$$
(\psi, A_{op}\phi) = \psi^{\dagger} A\phi
$$

where the dagger symbol indicates a Hermitian conjugate of the matrix. Let us recall that the scalar product between two column vectors ψ and ϕ is $\psi^{\dagger}\phi$

$$
\begin{aligned}
&(\psi, A\phi) = \psi^{\dagger}A\phi \\
&(A^{\dagger}\psi, \phi) = (A^{\dagger}\psi)^{\dagger}\phi = \psi^{\dagger}(A^{\dagger})^{\dagger}\phi = \psi^{\dagger}A\phi \\
\therefore \quad &(\psi, A\phi) = (A^{\dagger}\psi, \phi).
\end{aligned}
$$

4.3.3 Self-adjoint Operator

A linear operator A is said to be self-adjoint or Hermitian if $A^{\dagger}$ is the same as A itself.

Alternatively, a linear operator A is said to be self-adjoint if it satisfies the relation

$$
(\phi, A\psi) = (A\phi, \psi) \qquad \text{or} \qquad (\phi, A\psi) = (\psi, A\phi)^* \tag{4.21}
$$

Example 4.5 Prove that the momentum operator $p_{op} = -i\hbar\, d/dx$ is a self-adjoint operator when it acts on the elements of Hilbert space of bound state wave functions.

Solution:

$$
\begin{aligned}
(\phi, p_{op}\psi) &= \int_{-\infty}^{\infty} \phi^*(x)\left(-i\hbar\frac{d}{dx}\right)\psi(x)dx = -i\hbar\int_{-\infty}^{\infty} \phi^*(x)\frac{d}{dx}\psi(x)dx \\
&= [-i\hbar\phi^*(x)\psi(x)]_{-\infty}^{\infty} + i\hbar\int_{-\infty}^{\infty}\frac{d\phi^*}{dx}\psi dx \\
&= -i\hbar[\phi^*(\infty)\psi(\infty) - \phi^*(-\infty)\psi(-\infty)] + i\hbar\int_{-\infty}^{\infty}\frac{d\phi^*}{dx}\psi dx
\end{aligned}
$$

Since $\phi(x)$ and $\psi(x)$ are bound state wave functions, both $\phi(x)$ and $\psi(x)$ go to zero as $x \to \pm\infty$. Therefore, the first two terms are zero.

$$
\begin{aligned}
\therefore\quad (\phi, p_{op}\psi) &= i\hbar\int_{-\infty}^{\infty}\frac{d\phi^*(x)}{dx}\psi(x)dx = \int_{-\infty}^{\infty}\left(-i\hbar\frac{d}{dx}\phi(x)\right)^*\psi(x) \\
&= (p_{op}\phi, \psi)
\end{aligned}
$$

Therefore, $p_{op} = -i\hbar\dfrac{d}{dx}$ is a self-adjoint operator.

4.4 EIGENVALUE EQUATION

For every linear operator A_{op}, we can define an eigenvalue equation

$$A_{op}\phi_a = a\phi_a \tag{4.22}$$

Here a is called the eigenvalue and ϕ_a is called the eigenvector or eigenfucnction corresponding to the eigenvalue a. Note that eigenvalues and the eigenfunctions for a given operator are not arbitrary. We have to solve the eigenvalue Equation (4.22) to get the eigenvalues and the corresponding eigenfunctions for a given operator.

The list of all possible eigenvalues are called the eigenvalue spectrum of the operator A_{op}. The eigenvalues may be discrete or continuous. If they are discrete, they can be enumerated as $\{a_1, a_2, a_3,...\}$. For the same eigenvalue, if there exists more than one eigenfunction, such eigenvalue is called degenerate, and the corresponding eigenfunctions are called degenerate eigenfunctions.

Let us make the following observation from the point of view of quantum mechanics. Though the domain of the operator is restricted to a subspace of Hilbert space, its eigenfunctions need not be the elements of the Hilbert space. For instance, the eigenfunctions of the momentum operator are not the elements of Hilbert space of square integrable functions.

Eigenvalues and Eigenfunctions of a Self Adjoint Operator

The eigenvalues of a self-adjoint operator are real, and their eigenfunctions are orthogonal to each other. We shall prove this result by considering a restricted case, namely distinct eigenvalues and their eigenfunctions. However, this result is true, irrespective of the fact whether the eigenvalues are discrete or continuous and degenerate or non-degenerate.

Example 4.6 Prove that the eigenvalues of a self-adjoint operator are real and the eigenfunctions belonging to distinct eigenvalues are orthogonal to each other.

Solution: Let A_{op} be the self-adjoint operator. Therefore, A_{op} satisfies the relation

$$(\phi, A_{op}\psi) = (A_{op}\phi, \psi)$$

Let $a_1, a_2, a_3, \ldots$ be the set of eigenvalues of the operator A_{op}, and the corresponding eigenfunctions are $\phi_1, \phi_2, \ldots \phi_n$.

$$A_{op}\phi_i = a_i\phi_i \text{ and } A_{op}\phi_j = a_j\phi_j$$

The definition of self-adjoint operator leads to the relation

$$(\phi_i, A_{op}\phi_j) = (A_{op}\phi_i, \phi_j)$$

$$\text{LHS}: (\phi_i, A_{op}\phi_j) = (\phi_i, a_j\phi_j) = a_j(\phi_i, \phi_j)$$

$$\text{RHS}: (A_{op}\phi_i, \phi_j) = (a_i\phi_i, \phi_j) = a_i^*(\phi_i, \phi_j)$$

If the reader finds the above results difficult, he can work them in the Hilbert space $L_2(\text{R}, dx)$.

$$\{(\phi_i, a_j\phi_j) = \int \phi_i^* a_j \phi_j d_T = a_j \int \phi_i^* \phi_j d\tau$$

$$(a_i\phi_i, \phi_j) = \int (a_i\phi_i)^* \phi_j dT = a_i^* \int \phi_i^* \phi_j d\tau\}$$

$$\therefore a_j(\phi_i, \phi_j) = a_i^*(\phi_i, \phi_j)$$

$$(a_j - a_i^*)(\phi_i, \phi_j) = 0 \tag{4.23}$$

Case 1: $i = j$: The above equation becomes

$$(a_i - a_i^*)(\phi_i, \phi_i) = 0$$

Since (ϕ_i, ϕ_i) cannot be zero (see (4.9)),

$$a_i - a_i^* = 0 \quad \text{or} \quad a_i = a_i^*$$

Therefore, a_i is a real number.

Case 2: $a_i \neq a_j$

Then the Equation (4.23) becomes

$$(a_j - a_i)(\phi_i, \phi_j) = 0$$

Since $(a_j - a_i)$ cannot be zero, the only possibility is $(\phi_i, \phi_j) = 0$.

Orthonormal Eigenfunctions

In the previous example, note that $(\phi_i, \phi_i) \neq 0$. This can be made to be equal to 1, by suitably re-defining ϕ_i. Therefore, without losing any generality, we can write

$$(\phi_i, \phi_i) = 1 \text{ and } (\phi_i, \phi_j) = 0 \text{ if } i \neq j$$

Let us define Kronecker delta function δ_{ij} as

$$\delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

Using Kronecker delta function, we can write

$$(\phi_i, \phi_j) = \delta_{ij} \tag{4.24}$$

The set of linearly independent functions $\phi_i s$ is said to be orthonormal set if they satisfy (4.24). Therefore, the set of eigenfunctions of a self-adjoint operator forms an orthonormal set.

Complete Set

The eigenfunctions of a self-adjoint operator form a complete set. The meaning of this statement is that any arbitrary wave function in the Hilbert space can be expressed as a linear combination of this set of orthonormal functions. Let ψ be any arbitrary wave function in the Hilbert space. Then we have

$$\psi = \Sigma c_n \phi_n \tag{4.25}$$

This is very much similar to an arbitrary geometrical vector $\mathbf{F} = F_1\mathbf{i} + F_2\mathbf{j} + F_3\mathbf{k}$. The component or coefficient can be obtained by taking scalar product **F** with a basis vector. For instance, F_1 can be obtained by

$$F_1 = \mathbf{i} \cdot \mathbf{F}$$

Going back to the Equation (4.25), now the coefficient c_m can be found to be

$$c_m = (\phi_m, \psi) \tag{4.26}$$

This can be easily checked.

$$\begin{aligned} (\phi_m, \psi) &= (\phi_m, \sum_n c_n \phi_n) \\ &= \sum_n c_n (\phi_m, \phi_n) = \sum_n c_n \delta_{mn}. \end{aligned}$$

In the last summation $\sum_n c_n \delta_{mn}$, only one term, namely the term with $n = m$, alone will survive. i.e.,

$$\sum_n c_n \delta_{mn} = c_m$$

$$\therefore (\phi_m, \psi) = c_m$$

The formal mathematical treatment of Hilbert spaces concentrates considerably on the validity of the expansion of ψ as a linear combination of the basis function as given in (4.25). Though we have used the analogy of geometrical vectors to explain the meaning of the Equation (4.25), in the case of infinite dimensional linear vector space, things are not so simple, and the validity of the Equation (4.25) requires elaborate scrutiny.

4.5 DIRAC DELTA FUNCTION

There are many definitions of Dirac delta functions. Here we choose the simplest one, though this may not be mathematically satisfactory.

The Dirac delta function $\delta(x)$ is defined as

$$\delta(x) = \begin{cases} \infty & \text{at } x = 0 \\ 0 & \text{at } x \neq 0 \end{cases} \tag{4.27}$$

such that
$$\int_{-\infty}^{\infty} \delta(x)dx = 1$$

In the above integral, it is enough to have the upper and lower limits on either side of $x = 0$ so that $x = 0$ is included within these limits.

i.e.,
$$\int_{-\infty}^{\infty} \delta(x)dx = \int_{-\varepsilon}^{\varepsilon} \delta(x)dx = 1$$

This delta function $\delta(x)$ is said to be centred at $x = 0$. We can have a delta function centred at any point $x = a$. In such case, the Dirac delta function becomes

$$\delta(x-a) = \begin{cases} \infty & \text{at } x = a \\ 0 & \text{at } x \neq a \end{cases} \tag{4.28}$$

such that
$$\int_{-\infty}^{\infty} \delta(x-a)dx = \int_{a-\varepsilon}^{a+\varepsilon} \delta(x-a)dx = 1$$

We have a number of representations of Dirac delta function. One of the simplest representations of Dirac delta function is to consider $g_\varepsilon(x, a)$ in the limit $\varepsilon \to 0$ where $g_\varepsilon(x, a)$ is given by

$$g_\varepsilon(x,a) = \begin{cases} \dfrac{1}{2\varepsilon} & a-\varepsilon < x < +a+\varepsilon \\ 0 & \text{otherwise} \end{cases}$$

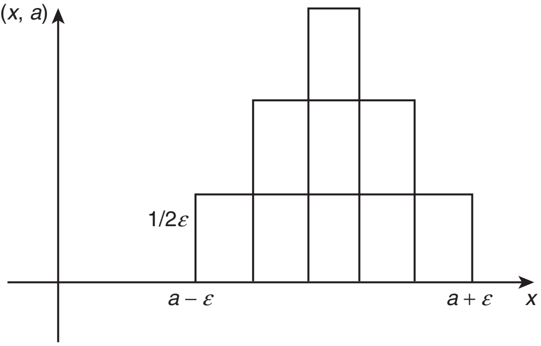


Fig. 4.3 g_ε (x, a) for different values of ε

In the Fig. 4.3, g_ε (x, a) is sketched for different values of ε. It is a rectangular box of width 2ε and height $1/2\varepsilon$ so that the area under the curve always is 1 for all values of ε. However, as ε goes to zero, the width becomes smaller, and the height becomes larger. This is essentially the Dirac delta function, so we can define Dirac delta function as

$$\delta(x-a) = \underset{\varepsilon\to 0}{\mathrm{Lt}}\, g_\varepsilon(x,a)$$

Example 4.7 Prove that $\int_{-\infty}^{\infty} f(x)\delta(x-a)dx = f(a)$.

Solution: This is one of the important properties of $\delta(x-a)$. Consider the integral $\int_{-\infty}^{\infty} f(x)\, g_\varepsilon(x,a)\, dx$.

$$\int_{-\infty}^{\infty} f(x)g_\varepsilon(x,a)dx = \int_{a-\varepsilon}^{a+\varepsilon} f(x)g_\varepsilon(x,a)\, dx = \int_{a-\varepsilon}^{a+\varepsilon} f(x)\cdot\frac{1}{2\varepsilon}dx$$

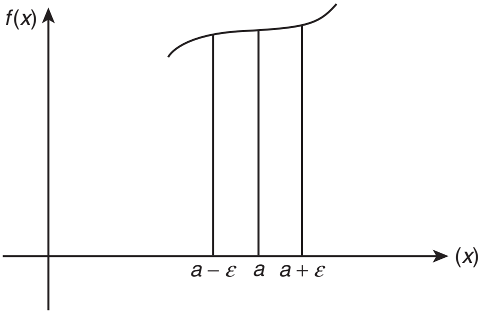


Fig. 4.4

In the limit $\varepsilon \to 0$, the function $f(x)$ does not change much from $f(a)$ in the integration range $a-\varepsilon$ to $a+\varepsilon$. Therefore, $f(x)$ can be replaced by $f(a)$ throughout this range of integration.

i.e.,
$$\int_{-\infty}^{\infty} f(x)\delta(x-a) = \underset{\varepsilon\to 0}{\mathrm{Lt}} \int_{a-\varepsilon}^{a+\varepsilon} f(x)g_\varepsilon(x,a)dx$$

$$= \underset{\varepsilon\to 0}{\mathrm{Lt}} \int_{a-\varepsilon}^{a+\varepsilon} f(a)\cdot\frac{1}{2\varepsilon}dx = f(a)\, \underset{\varepsilon\to 0}{\mathrm{Lt}} \int_{a-\varepsilon}^{a+\varepsilon} \frac{1}{2\varepsilon}dx = f(a)$$

$$\int_{-\infty}^{\infty} f(x)\delta(x-a)dx = f(a). \tag{4.29}$$

In fact, this result itself can be taken as the defining property of Dirac delta function.

Example 4.8 By considering Fourier transform and its inverse, show that $\int_{-\infty}^{\infty} e^{ik(x-a)}dk = 2\pi\,\delta(x-a)$. The Fourier transform and its inverse are given by

$$F(k)=\int_{-\infty}^{\infty} e^{ikx} f(x)dx \quad \text{and} \quad f(x)=\frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-ikx}F(k)dk$$

Solution: Let us start with the inverse Fourier transform.

$$f(a)=\frac{1}{2\pi}\int_{-\infty}^{\infty} dke^{-ika}F(k)=\frac{1}{2\pi}\int_{-\infty}^{\infty} dke^{-ika}\int_{-\infty}^{\infty} e^{ikx}f(x)dx=\frac{1}{2\pi}\int_{-\infty}^{\infty} dx\left[\int_{-\infty}^{\infty} dke^{ik(x-a)}\right]f(x)$$

$$=\int_{-\infty}^{\infty}\delta(x-a)f(x)dx$$

provided we make the following identification

$$\delta(x-a)=\frac{1}{2\pi}\int_{-\infty}^{\infty} dke^{ik(x-a)} \tag{4.30}$$

In three dimensions, we have $\delta(\mathbf{r}-\mathbf{r}')=\frac{1}{(2\pi)^3}\int e^{i\mathbf{k}\cdot(\mathbf{r}-\mathbf{r}')}d^3\mathbf{k}$ (4.31)

Example 4.9 Show that $\delta(-x)=\delta(x)$.

Solution:

$$\delta(-x)=\frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-ikx}dk$$

On changing the variable k to $k'=-k$, we get

$$\delta(-x)=\frac{1}{2\pi}\int_{\infty}^{-\infty} e^{ik'x}(-dk')=\frac{1}{2\pi}\int_{-\infty}^{\infty} e^{ik'x}dk'=\delta(x)$$

$$\delta(-x)=\delta(x) \tag{4.32}$$

Example 4.10 Prove that $\underset{\varepsilon\to 0+}{\text{Lt}}\ \frac{1}{\pi}\frac{\varepsilon}{x^2+\varepsilon^2}=\delta(x)$

Solution: One way of checking this result is as follows:

$$\text{As} \quad x\to 0, \quad \underset{\varepsilon\to 0+}{\text{Lt}}\ \frac{1}{\pi}\cdot\frac{\varepsilon}{\varepsilon^2}\to\infty$$

For other values of x, $\underset{\varepsilon\to 0+}{\text{Lt}}\ \frac{\varepsilon}{x^2+\varepsilon^2}=\underset{\varepsilon\to 0+}{\text{Lt}}\ \frac{\varepsilon}{x^2}=0$

$$\int_{-\infty}^{\infty}\underset{\varepsilon\to 0+}{\text{Lt}}\ \frac{1}{\pi}\cdot\frac{\varepsilon}{x^2+\varepsilon^2}dx=\underset{\varepsilon\to 0+}{\text{Lt}}\ \frac{1}{\pi}\cdot\varepsilon\cdot\frac{1}{\varepsilon}\tan^{-1}\frac{x}{\varepsilon}\bigg|_{-\infty}^{\infty}$$

$$=\frac{1}{\pi}\cdot\varepsilon\cdot\frac{1}{\varepsilon}\cdot\pi=1.$$

Alternative method: Consider the integral representation $\delta(x)=\frac{1}{2\pi}\int_{-\infty}^{\infty}e^{ikx}\,dk$

$$\delta(x)=\frac{1}{2\pi}\int_{-\infty}^{\infty}e^{ikx}dk=\frac{1}{2\pi}\int_{-\infty}^{0}e^{ikx}dk+\frac{1}{2\pi}\int_{0}^{\infty}e^{ikx}\,dk$$

The integral does not converge, so multiply each integrand by convergent factor. Multiply the integrand in the first integral by $e^{k\varepsilon}$(in this case, k varies from $-\infty$ to 0, so $e^{k\varepsilon}=e^{-|k|\varepsilon}$) and multiply the integrand in the second integral by$e^{-k\varepsilon}$. In the end, take $\varepsilon\to 0$ limit.

$$\delta(x)=\underset{\varepsilon\to 0}{\text{Lt}}\frac{1}{2\pi}\left[\int_{-\infty}^{0}e^{ikx}e^{k\varepsilon}dk+\int_{0}^{\infty}e^{ikx}e^{-k\varepsilon}dk\right]=\underset{\varepsilon\to 0}{\text{Lt}}\frac{1}{2\pi}\left[\int_{-\infty}^{0}e^{i(x-i\varepsilon)k}dk+\int_{0}^{\infty}e^{i(x+i\varepsilon)k}dk\right]$$

$$=\underset{\varepsilon\to 0}{\text{Lt}}\frac{1}{2\pi}\left\{\left[\frac{e^{i(x-i\varepsilon)k}}{i(x-i\varepsilon)}\right]_{-\infty}^{0}+\left[\frac{e^{i(x+i\varepsilon)k}}{i(x+i\varepsilon)}\right]_{0}^{\infty}\right\}$$

Note as $k\to-\infty$, $e^{ikx}e^{k\varepsilon}\to 0$ and as $k\to\infty$, $e^{ikx}e^{-k\varepsilon}\to 0$. Therefore, we have

$$\delta(x)=\underset{\varepsilon\to 0}{\text{Lt}}\frac{1}{2\pi}\left\{\frac{1}{i(x-i\varepsilon)}-\frac{1}{i(x+i\varepsilon)}\right\}$$

$$=\underset{\varepsilon\to 0}{\text{Lt}}\frac{1}{2\pi}\frac{2i\varepsilon}{i(x^2+\varepsilon^2)}=\underset{\varepsilon\to 0}{\text{Lt}}\frac{1}{\pi}\cdot\frac{\varepsilon}{x^2+\varepsilon^2}$$

Properties of δ function are listed below:

1. $\delta(ax)=\frac{1}{|a|}\delta(x)$
2. $\delta(x^2-a^2)=\frac{1}{2|a|}[\delta(x+a)+\delta(x-a)]$
3. $x\delta'(x)=-\delta(x)$
4. $\delta(g(x))=\sum_i\frac{1}{|g'(x_i)|}\delta(x-x_i)$

where x_i is the root of the equation $g(x)=0$.

These properties are to be understood in the context of integration; that is, if the integrand in an integral contains the quantity in the LHS as a factor, it can be replaced by the corresponding quantity in RHS (which will not alter the value of the integral). This is illustrated in the following examples:

Example 4.11 Prove that $\delta(ax)=\frac{1}{|a|}\delta(x)$.

Solution: Consider the integral $I=\int_{-\infty}^{\infty}f(x)\delta(ax)dx$.

Put $ax=y$. The upper and lower limits of the integration depend on the sign of a.

Case 1: $a > 0$; As $x \to -\infty$, $y \to -\infty$ and as $x \to \infty$, $y \to \infty$

Therefore, we have

$$I = \int_{-\infty}^{\infty} f\left(\frac{y}{a}\right) \delta(y) \frac{dy}{a} = \frac{1}{a} f(0)$$

But $\frac{1}{a} f(0)$ can be written as

$$\frac{1}{a} f(0) = \int_{-\infty}^{\infty} f(x) \frac{1}{a} \delta(x) dx$$

or

$$\int_{-\infty}^{\infty} f(x)\delta(ax)dx = \int_{-\infty}^{\infty} f(x)\frac{1}{a}\delta(x)dx \quad a > 0$$

Since this equation is true for any arbitrary *f*(*x*), we can write

$$\delta(ax) = \frac{1}{a}\delta(x) \quad a > 0$$

Case 2: $a < 0$; As $x \to \infty$, $y \to -\infty$ and as $x \to -\infty$, $y \to \infty$

$$I = \int_{-\infty}^{\infty} f(x)\delta(ax)dx = \int_{\infty}^{-\infty} f\left(\frac{y}{a}\right) \delta(y) \frac{1}{a} dy$$

$$= -\frac{1}{a} \int_{-\infty}^{\infty} f\left(\frac{y}{a}\right) \delta(y) dy = -\frac{1}{a} f(0).$$

$-\frac{1}{a} f(0)$ can be written as $\quad -\frac{1}{a} f(0) = \frac{1}{-a} \int_{-\infty}^{\infty} f(x)\delta(x)dx.$

$$\therefore \int_{-\infty}^{\infty} f(x)\delta(ax)dx = \int_{-\infty}^{\infty} f(x) \cdot \frac{1}{-a}\delta(x)dx \quad a < 0$$

Note $(-a)$ is a positive number. In fact, $-a = |a|$

$$\therefore \int_{-\infty}^{\infty} f(x)\delta(ax)dx = \int_{-\infty}^{\infty} f(x)\frac{1}{|a|}\delta(x)dx \quad a < 0.$$

Since this equation is true for any arbitrary *f*(*x*), we can write

$$\delta(ax) = \frac{1}{|a|}\delta(x) \quad a < 0.$$

Combining the result of case 1 and case 2, we can write

$$\delta(ax) = \frac{1}{|a|}\delta(ax) \quad \text{for all } a. \tag{4.33}$$

Example 4.12 Prove that $\delta(x^2 - a^2) = \frac{1}{2|a|}[\delta(x+a) + \delta(x-a)]$.

Solution: $\delta(x^2 - a^2)$ has to be zero for all values of x, except $x = -a$ and $x = a$. This is schematically shown in the Fig. 4.5.

Fig. 4.5

Since the integrand is zero in the ranges $-\infty < x < -(a + \varepsilon)$, $(-a + \varepsilon) < x < (a - \varepsilon)$ and $(a + \varepsilon < x < \infty)$, the integration in these ranges does not contribute to the integral.

$$\therefore \quad \int_{-\infty}^{\infty} f(x)\delta(x^2 - a^2)dx = \int_{-a-\varepsilon}^{-a+\varepsilon} f(x)\delta(x^2 - a^2)dx + \int_{a-\varepsilon}^{a+\varepsilon} f(x)\delta(x^2 - a^2)dx$$

In the range $(-a - \varepsilon < x < -a + \varepsilon)$, we have
as

$$x \to -a, \quad \delta[(x-a)(x+a)] \to \delta[-2a(x+a)]$$

But $\delta[-2a(x+a)] = \frac{1}{|2a|}\delta(x+a)$.

$$\therefore \quad \int_{-a-\varepsilon}^{-a+\varepsilon} f(x)\delta(x^2 - a^2)dx = \int_{-a-\varepsilon}^{-a+\varepsilon} f(x)\ \delta[(x-a)(x+a)]dx$$

$$= \int_{-a-\varepsilon}^{-a+\varepsilon} f(x)\delta[(-2a(x+a)]dx = \frac{1}{|2a|}\int_{-a-\varepsilon}^{-a+\varepsilon} f(x)\delta(x+a)dx = \frac{1}{|2a|}f(-a).$$

Similarly, in the range $(a - \varepsilon < x < a + \varepsilon)$, we have

as $$x \to a\,\delta[(x-a)(x+a)] \to \delta[2a(x-a)]$$

$$\therefore \quad \int_{a-\varepsilon}^{a+\varepsilon} f(x)\ \delta(x^2 - a^2)dx = \int_{a-\varepsilon}^{a+\varepsilon} f(x)\ [\delta(x-a)(x+a)]dx$$

$$= \int_{a-\varepsilon}^{a+\varepsilon} f(x)\delta[2a(x-a)]dx = \int_{a-\varepsilon}^{a+\varepsilon} f(x)\frac{1}{|2a|}\delta(x-a)dx = \frac{f(a)}{|2a|}$$

$$\therefore \quad \int_{-\infty}^{\infty} f(x)\delta(x^2 - a^2)dx = \frac{1}{|2a|}[f(-a) + f(a)]$$

$$= \int_{-\infty}^{\infty} f(x)\frac{1}{|2a|}[\delta(x+a) + \delta(x-a)]dx$$

Since this equation is true for any arbitrary function $f(x)$, we have

$$\delta(x^2 - a^2) = \frac{1}{|2a|}[\delta(x+a) + \delta(x-a)]$$

Example 4.13 Show that $f(x)\frac{d}{dx}\delta(x-x') = -\frac{df}{dx}\delta(x-x')$.

Solution: Consider the integral $\int f(x)\frac{d}{dx}\delta(x-x')$.

$$\int_{-\infty}^{\infty} f(x)\frac{d}{dx}\delta(x-x') = \underbrace{f(x)\delta(x-x')\Big|_{-\infty}^{\infty}}_{0} - \int_{-\infty}^{\infty}\frac{df(x)}{dx}\delta(x-x')dx = -\int_{-\infty}^{\infty}\frac{df(x)}{dx}\delta(x-x')dx$$

Since the above result is true for any arbitrary function $f(x)$ we have,

$$f(x)\frac{d}{dx}\delta(x-x') = -\frac{df}{dx}\delta(x-x') \tag{4.34}$$

EXERCISES

1. Explain what is meant by linear independence.
2. What is meant by scalar product in a linear vector space?
3. Give the conditions to be satisfied for a scalar product.
4. What are the differences between axioms for a scalar product defined in a linear vector space of geometrical vectors and the linear vector space in quantum mechanics?
5. Explain what is meant by Hilbert space.
6. Give the examples of Hilbert spaces that are used in non-relativistic quantum mechanics.
7. Explain what is meant by domain of an operator in quantum mechanics.
8. When are two operators said to be the same?
9. Define a linear operator.
10. Define the adjoint of an operator.
11. Define the self-adjoint operator.
12. Prove that $(AB)^\dagger = B^\dagger A^\dagger$.
13. Define the Dirac delta function.
14. Express $\begin{pmatrix} 3 \\ 5 \end{pmatrix}$ as a linear combination of basis vectors (i) $e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ (ii) $e_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and $e_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ (iii) $e_1 = \begin{pmatrix} 1 \\ i \end{pmatrix}$ and $e_2 = \begin{pmatrix} 1 \\ -i \end{pmatrix}$
15. Which of the following operators is linear operator?

    (i) $A_{op}\phi = \frac{d^2\phi}{dx^2}$ (ii) $A_{op}\phi = \left(\frac{d\phi}{dx}\right)^2$
16. Prove that $x\delta'(x) = -\delta(x)$, where the prime denotes differentiation with respect to x.
17. Prove that $\delta(x) = \underset{x\to\infty}{\mathrm{Lt}} \frac{\sin x}{\pi x}$.

REFERENCES

1. A. K. Ghatak, I. C. Goyal and S. J. Chua, 1995. *Mathematical Physics.* New Delhi: Macmillan India Limited.
2. K. F. Riley, M. D. Hobson and S. J. Bence, 1998. *Mathematical Methods for Physics and Engineering.* Cambridge University Press.
3. P. M. Mathews and K. Venkatesan, 1976. *A Textbook of Quantum Mechanics*, Tata McGraw-Hill Publishing Company Limited.
4. S. Gasiorowicz, 1974. *Quantum Physics.* New York: John Wiley & Sons, Inc.
5. J. L, Powell and B. Crasemann, 1961. *Quantum Mechanics,* Reading Massachusetts: Addison-Wesley Publishing Company, Inc.
6. Francois Gieres, 'Mathematical Surprises and Dirac's Formalism in Quantum Mechanics', Reports in Progress in Physics 63 (2008), 1893−1931.
7. C. I. Tan, 'Notes on Hilbert Space', http://jcbmac.chem.brown.edn/baird/quantumpdf/Tan_on_Hilbert-space.html.

5

General Formalism

Quantum mechanics is radically different from classical physics. It cannot be derived from classical physics. The description of the state, the rules to calculate a physical quantity and the interpretations in quantum mechanics are absolutely different from classical physics. One has to start the study of quantum mechanics with the study of set of postulates, based on which the whole quantum mechanics is developed. The validity of these postulates lies in the experimental verification of the results deduced from these postulates. In this respect, quantum mechanics is an extraordinarily successful theory. Postulatory approach to a subject is not new to physics. All the branches of physics have to start with a set of few basic postulates. Special theory of relativity is the best example, which is based on two simple postulates. However, the postulates of quantum mechanics are not that simple like those of special theory of relativity. The number of postulates itself varies from author to author. The statement of each postulate is fairly complex. Each postulate has to be followed by an elaborate commentary. Though the number of postulates varies from book to book, the sum total of all the postulates taken together should give the same amount of information. Here, we present seven postulates closely following Dicke and Wittki.

5.1 POSTULATES

Postulate P-1

The state of the system is described by a wave function $\psi(\mathbf{r},t)$.

We have already seen the interpretation and properties of wave functions in Chapter 2. Here we briefly recapitulate those ideas. The wave function $\psi(\mathbf{r},t)$ does not represent fluctuation in any physical property. The interpretation of the wave function $\psi(\mathbf{r},t)$ was provided by Max Born. $\psi(\mathbf{r},t)$ is interpreted as the probability amplitude, and $|\psi(\mathbf{r},t)|^2$ is interpreted as the probability density for the position of a particle. The probability density $\rho(\mathbf{r},t)$ is given by

$$\rho(\mathbf{r},t) = \psi^*(\mathbf{r},t)\psi(\mathbf{r},t) = \left|\psi(\mathbf{r},t)\right|^2 \tag{5.1}$$

The probability of finding the particle in a volume $d^3\mathbf{r}$ centred around $\mathbf{r}$ is given by

$$\rho(\mathbf{r},t)d^3\mathbf{r} = \left|\psi(\mathbf{r},t)\right|^2 d^3\mathbf{r}$$

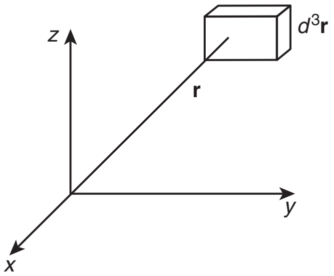


Fig. 5.1 Volume element d^3 **r** centred around **r**

We have already seen in Chapter 2 that the probability interpretation of $\psi(\mathbf{r},t)$ leads to the condition

$$\int \psi^*(\mathbf{r},t)\psi(\mathbf{r},t)d^3\mathbf{r} = 1 \tag{5.2}$$

Obviously, $\psi(\mathbf{r},t)$ is a square integrable function. In the case of N particles system, the state of the system is described by $\psi(\mathbf{r}_1,\mathbf{r}_2,\mathbf{r}_3,...,\mathbf{r}_N,t)$, and it is again interpreted as the probability amplitude.

i.e.,
$$\psi^*(\mathbf{r}_1,\mathbf{r}_2,...,\mathbf{r}_N,t)\psi(\mathbf{r}_1,\mathbf{r}_2,...,\mathbf{r}_N,t)\, d^3\mathbf{r}_1\, d^3\,\mathbf{r}_3 \ldots d^3\,\mathbf{r}_N \tag{5.3}$$

represents the probability of finding a particle in a volume $d^3\mathbf{r}_1$ (centred around $\mathbf{r}_1$), another particle in a volume $d^3\mathbf{r}_2$ (centred in $\mathbf{r}_2$) … and so on.

Postulate P-2

If $\psi_1(\mathbf{r},t)$ and $\psi_2(\mathbf{r},t)$ represent the states of a system, then there also exists a physical state of a system corresponding to $\psi(\mathbf{r},t) = c_1\psi_1(\mathbf{r},t) + c_2\psi_2(\mathbf{r},t)$.

This is called principle of superposition. Some authors combine these postulates to give a single equivalent statement saying that to each system, we can associate a Hilbert space H, whose elements $\psi(\mathbf{r}, t)$ describe the state of the system. Let us recall that the essential features of the Hilbert space are (i) the closure property, that is, if ψ_1 and ψ_2 are elements of H, then $c_1\psi_1 + c_2\psi_2$ is also an element of H (ii) and the existence of the inner product (ϕ, ψ). The significance of the principle of superposition will be discussed later.

Postulate P-3

To each dynamical variable A, we associate a linear self-adjoint operator A_{op}. The rule to assign a quantum mechanical operator is: If A is given by $A(\mathbf{r},\mathbf{p})$ in classical physics, then the corresponding quantum mechanical operator is $A_{op}(\mathbf{r}_{op},\mathbf{p}_{op})$,where $\mathbf{r}_{op} = \mathbf{r}$ and $\mathbf{p}_{op} = -i\hbar\nabla$.

This rule to assign an operator requires more careful scrutiny. Let us make the following observations.

1. This rule of assigning an operator can become ambiguous. For instance, in classical physics, $\mathbf{r}\cdot\mathbf{p}$ and $\mathbf{p}\cdot\mathbf{r}$ are the same. But in quantum mechanics, $\mathbf{r}\cdot(-i\hbar\nabla)$ and $(-i\hbar\nabla)$. $\mathbf{r}$ are different. In such circumstances, both combinations can be included. For instance, the operator corresponding to $\mathbf{r}\cdot\mathbf{p}$ is given by

$$\mathbf{r}\cdot\mathbf{p} \to \frac{1}{2}(\mathbf{r}_{op}\cdot\mathbf{p}_{op} + \mathbf{p}_{op}\cdot\mathbf{r}_{op}) \tag{5.4}$$

2. The operators A_{op} in quantum mechanics are Hilbert space operators; that is, in the operator statement $A_{op}\psi = \phi$, the wave functions ψ and ϕ are elements of Hilbert space H.

3. There are a number of dynamical variables which are unique to quantum mechanics. There is no classical analogue corresponding to them. For instance, spin angular momentum in quantum mechanics is different from that of classical physics. There is no classical analogue to quantum mechanical spin angular momentum. In such cases, the corresponding operators are obtained from other considerations.

In what follows, we give a list of some dynamical variables and their operators.

Table 5.1 Some dynamical variables and their operators

Classical	**Quantum mechanical**	
1. Position $\mathbf{r}$	$\mathbf{r}_{op} = \mathbf{r}$	$\mathbf{r}_{op}\psi = \mathbf{r}\psi$
2. Momentum $\mathbf{p}$	$\mathbf{p}_{op} = -i\hbar\nabla$	$\mathbf{p}_{op}\psi = -i\hbar\nabla\psi$
3. Orbital angular momentum $\mathbf{L} = \mathbf{r}\times\mathbf{p}$	$\mathbf{L}_{op} = \mathbf{r}_{op}\times\mathbf{p}_{op}$	$L_{op}\psi = -i\hbar\mathbf{r}\times\nabla\psi$
4. Kinetic energy $T = \dfrac{p^2}{2m}$	$T_{op} = \dfrac{p_{op}^2}{2m}$	$T_{op}\psi = \dfrac{1}{2m}\cdot(-\hbar^2\nabla^2\psi)$
5. Potential energy $V(\mathbf{r})$	$V_{op} = V(\mathbf{r})$	$V_{op}\psi = V(\mathbf{r})\psi$
6. Hamiltonian $H = \dfrac{p^2}{2m} + V(\mathbf{r})$	$H_{op} = \dfrac{p_{op}^2}{2m} + V(\mathbf{r})$	$H_{op}\psi = -\dfrac{\hbar^2}{2m}\nabla^2\psi + V\psi$

How do we choose these operators? Dirac has given a condition called Dirac's quantum condition, which provides a guidance to choose quantum mechanical operators. This condition necessitates to introduce the concept of commutation relation between two operators which play an important role in quantum mechanics.

The commutation relation between the operators A_{op} and B_{op} is defined as

$$[A_{op}, B_{op}] = A_{op}B_{op} - B_{op}A_{op}$$

Formally, we can equate the operator $C_{op}$to the commutator $[A_{op}, B_{op}]$ as follows. Consider the following equation:

$$A_{op}B_{op}\psi - B_{op}A_{op}\psi = C_{op}\psi$$

If this equation is true for an arbitrary function ψ, then, we can take

$$[A_{op}, B_{op}] = A_{op}B_{op} - B_{op}A_{op} = C_{op}$$

Dirac's Quantum Condition

The operators corresponding to the dynamical variables A and B are to be chosen such that the commutator between A_{op} and B_{op} is $i\hbar$ times the Poisson bracket between them in classical mechanics.

i.e., $$[A_{op}, B_{op}] = i\hbar[A, B]_{PB}.$$

where $$[A, B]_{PB} = \sum\left[\frac{\partial A}{\partial q_i}\frac{\partial B}{\partial p_i} - \frac{\partial A}{\partial p_i}\frac{\partial B}{\partial q_i}\right]$$

Dirac's quantum condition itself can be taken as the second postulate. However, there are a number of restrictions on this condition. This condition provides only a guideline to choose the operator corresponding to a dynamical variable. It does not tell what the operator is corresponding to the dynamical variable. In fact, there are many ways of choosing an operator for a dynamical variable. As an example, consider the dynamical valuables x and p_x. The Poisson bracket

$$[x, p_x]_{PB} = \frac{\partial x}{\partial x}\cdot\frac{\partial p_x}{\partial p_x} - \frac{\partial x}{\partial p_x}\cdot\frac{\partial p_x}{\partial x} + \frac{\partial x}{\partial y}\cdot\frac{\partial p_x}{\partial p_y} - \frac{\partial x}{\partial p_y}\frac{\partial p_x}{\partial y} + \frac{\partial x}{\partial z}\frac{\partial p_x}{\partial p_z} - \frac{\partial x}{\partial p_z}\cdot\frac{\partial p_x}{\partial z}$$

$$= 1$$

The corresponding quantum mechanical operators x_{op} and $(p_x)_{op}$ have to be chosen such that

$$[x_{op}, p_{xop}] = i\hbar$$

One choice is $$x_{op} = x \quad \text{and} \quad p_{xop} = -i\hbar\frac{\partial}{\partial x}$$

i.e., $$x_{op}\psi(x) = x\psi(x) \quad \text{and} \quad p_{op}\psi(x) = -i\hbar\frac{\partial\psi(x)}{\partial x}$$

Another choice is $$x_{op} = i\hbar\frac{\partial}{\partial p} \quad \text{and} \quad p_{op} = p_x$$

i.e., $$x_{op}\phi(p) = i\hbar\frac{\partial\phi(p)}{\partial p} \quad \text{and} \quad p_{op}\ \phi(p) = p_x\ \phi(p)$$

Dirac's quantum condition is true only when the dynamical variables are expressed in Cartesian coordinates.

Postulate P-4

Whenever a measurement of a dynamical variable A is made, the only possible experimental outcomes for any arbitrary state of the system are the eigenvalues of the corresponding quantum mechanical operator.

This postulate gives a prescription to determine the possible outcomes in an experiment. When an experiment is performed to determine the value of a dynamical variable, the possible outcomes are restricted to the list of eigenvalues of the corresponding operator. This result is true for any system. This implies that it is impossible to get any value other than a number in the list of the eigenvalues. Schematically, this postulate can be explained as follows:

Dynamical variable to be measured:	A
System:	any system
Quantum mechanical operator:	A_{op}
Eigenvalue equation:	$A_{op}\phi_i = a_i\phi_i$.
Eigenvalue spectrum:	$\{a_i\}$ or $\{a_1, a_2, a_3, \ldots\ldots\ldots\}$
List of possible experimental outcomes:	$\{a_i\}$ or $\{a_1, a_2, \ldots a_n\}$

Example 5.1 Z component of orbital angular momentum is measured for (i) a free electron and (ii) an electron bound to an atom. What are the possible experimental outcomes?

Solution: We have to use the postulates P-3 and P-4 to determine the possible experimental outcomes. We have to go through the following steps:

Step 1: Write down the expression for *z* component of angular momentum in classical physics.

$$L_z = (\mathbf{r}\times\mathbf{p})_z = xp_y - yp_x$$

Step 2: Replace L_z by $(L_z)_{op}$.

$$\begin{aligned}(L_z)_{op} &= x_{op}(p_y)_{op} - y_{op}(p_x)_{op}\\ &= -i\hbar x\frac{\partial}{\partial y} + i\hbar y\frac{\partial}{\partial x}\end{aligned}$$

Step 3: Find out the eigenvalues of this operator.

$$L_z\psi = m_1\hbar\psi$$

(Don't worry about the specific form in which the eigenvalue equation is written.)

The calculations are done in Chapter 7. Here we quote only the end results. The list of the possible eigenvalues are

$$\{\ldots -3\hbar, -2\hbar, 0\hbar, 1\hbar, 2\hbar, \ldots\}$$

Step 4: According to the postulate P-4, these are the only possible experimental outcomes. Therefore, whether it is free electron or an electron in an atom, the only possible values are from the list.

$$\{m_l\hbar\} \text{ or } \{\ldots -3\hbar, -2\hbar, -\hbar, 0\hbar, \hbar, 2\hbar, 3\hbar, \ldots\}$$

Postulate P-5

The eigenfunctions of a dynamical variable form a complete set.

i.e.,
$$A_{op}\phi_i = a_i\phi_i$$

$\{\phi_1, \phi_2, \phi_3, \ldots \phi_n, \ldots\}$ form a complete set.

It is trivial to check this postulate in the case of finite-dimensional Hilbert space. If the number of eigenfunctions forms an infinite set, then, in general, it is not obvious that they form a complete set. For quantum mechanical systems, this postulate ensures the completeness of the infinite set of eigenfunctions.

The significance of the complete set of eigenfunctions lies in the fact that any arbitrary wave function ψ can be written as

$$\psi = \sum c_n\phi_n$$

In general, the eigenvalues of an operator corresponding to a dynamical variable can be discrete and continuous. Let *A* be such dynamical variable, so the eigenvalue equation is

$$A\phi_i = a_i\phi_i \quad \text{and} \quad A\phi(a) = a\,\phi(a)$$

The eigenvalues are given by

$$\{a_1, a_2, a_3 \ldots\} \quad \text{and} \quad \{a : a \in \mathbb{R}\}$$

In such cases, we can write

$$\psi = \sum c_n \phi_n + \int c(a)\,\phi(a)\,da \tag{5.5}$$

The possibility that an arbitrary wave function ψ can be written as (5.5) makes us to call this postulate as the expansion postulate.

When the eigenfunctions are orthonormal, we can write

$$(\phi_n, \phi_m) = \delta_{mn} \quad \text{and} \quad (\phi(a), \phi(a')) = \delta(a - a')$$

and

$$c_n = \int \phi_n^* \psi d\tau \quad \text{and} \quad c(a) = \int \phi^*(a)\psi\, d\tau$$

Closure Relation for Complete Set of Functions

In the case of discrete eigenvalues, the complete set of eigenfunctions leads to the relation

$$\sum_n \phi_n^*(\mathbf{x}')\phi_n(\mathbf{x}) = \delta(\mathbf{x} - \mathbf{x}') \tag{5.6}$$

In the case of continuous eigenvalues, the complete set of the eigenfunctions leads to the relation

$$\int \phi^*(\mathbf{r}, a)\phi(\mathbf{r}', a)da = \delta(\mathbf{r} - \mathbf{r}') \tag{5.7}$$

These two relations are known as closure relation.

Example 5.2 Obtain the closure relation for a complete set of eigenfunctions when the eigenvalues of a dynamical variable A are discrete.

Solution: Then, the eigenvalue equation is

$$A_{op}\phi_n = a_n\phi_n \tag{5.8}$$

The set of eigenvalues are $\{a_1, a_2, \ldots a_n, \ldots\}$ and the corresponding eigenfunctions are $\{\phi_1, \phi_2, \phi_3, \ldots \phi_n, \ldots\}$.

Since, by definition, the set of $\{\phi_i\}$ forms a complete set (no other eigenfunction linearly independent of these ϕ_i's exists), we can write any arbitrary wave function ψ as

$$\psi = \sum c_n\phi_n = \sum (\phi_n, \psi)\,\phi_n, \quad \text{since} \quad c_n = (\phi_n, \psi)$$

$$\therefore \psi(\mathbf{r}) = \sum_n \int \phi_n^*(\mathbf{r}')\psi(\mathbf{r}')d^3\mathbf{r}'\phi_n(\mathbf{r}) = \int \underbrace{\sum_n \phi_n^*(\mathbf{r}')\phi_n(\mathbf{r})}_{\delta(\mathbf{r}-\mathbf{r}')}\ \psi(\mathbf{r}')d^3\mathbf{r}'$$

Note that the structure of the above equation is very much similar to

$$\psi(\mathbf{r}) = \int \delta(\mathbf{r} - \mathbf{r}')\ \psi(\mathbf{r}')d^3\mathbf{r}'$$

So, we get

$$\sum_n \phi_n^*(\mathbf{r}')\ \phi_n(\mathbf{r}) = \delta(\mathbf{r} - \mathbf{r}')$$

Example 5.3 Obtain the closure relation for the complete set of eigenfunctions of A when the eigenvalues are continuous.

The eigenvalue equation is

$$A\phi(a) = a\phi(a)$$

The orthogonality condition is

$$\int \phi^*(a;\mathbf{r})\phi(a';\mathbf{r})d^3\mathbf{r} = \delta(a-a').$$

From the definition of complete set of eigenfunctions (no other eigenfunction of A linearly independent of $\phi(a,\mathbf{r})$exists), we can write any arbitrary wave function as

$$\psi(\mathbf{r}) = \int c(a)\phi(a,\mathbf{r})\,da$$

But $c(a)$ is given by

$$c(a) = \int \phi^*(a,\mathbf{r}')\psi(\mathbf{r}')\,d^3\mathbf{r}'$$

$$\therefore\ \ \psi(\mathbf{r}) = \int \phi^*(a;\mathbf{r}')\psi(\mathbf{r}')\ \phi(a,\mathbf{r})\,da\,d^3\mathbf{r}'$$

$$= \int \underbrace{[\phi^*(a;\mathbf{r}')\phi(a;\mathbf{r})da]}_{\delta(\mathbf{r}-\mathbf{r}')}\ \psi(\mathbf{r})d^3\mathbf{r}'$$

Obviously, the above equation implies

$$\int \phi^*(a,\mathbf{r})\phi(a;\mathbf{r}')da = \delta(\mathbf{r}-\mathbf{r}')$$

Postulate P-6

The expectation value of a dynamical variable for an ensemble of identical systems, all of them being in the same state ψ, is given by

$$\langle A\rangle = \frac{(\psi, A_{op}\psi)}{(\psi,\psi)} \tag{5.9}$$

Generally, it is assumed that we are dealing with the normalized wave function, that is, we assume $(\psi, \psi) = 1$. Therefore, $\langle A\rangle$ is given by

$$\langle A\rangle = (\psi, A_{op}\psi) = \int \psi^* A_{op}\psi d^3\mathbf{r} \tag{5.10}$$

The postulates P-5 and P-6 together have far-reaching significance in quantum mechanics. So let us combine these two postulates. For the sake of clarity, we restrict ourselves to discrete eigenvalues.

Let A be the dynamical variable and the eigenvalue equation

$$A\phi_i = a_i\phi_i$$

Let us consider an arbitrary state of the system described by ψ. This need not be an eigenstate of A. Since ψ is arbitrary, from the postulate P-5, it can be written as

$$\psi = \sum c_n\phi_n$$

Let us substitute this expression in Equation (5.8) to calculate the expectation value $\langle A\rangle$.

$$\begin{aligned}\langle A\rangle &= \int \psi^* A_{op}\psi d\tau \\ &= \int\left(\sum_m c_m\phi_m\right)^* A_{op}\left(\sum c_n\phi_n\right)d\tau \\ &= \int\sum_m\sum_n c_m^* c_n\phi_m^* A_{op}\phi_n d\tau \\ &= \sum_m\sum_n c_m^* c_n\int\phi_m^* a_n\phi_n d\tau \\ &= \sum_m\sum_n c_m^* c_n a_n\int\phi_m^*\phi_n d\tau \\ &= \sum_m\sum_n c_m^* c_n a_n\delta_{mn} \\ &= \sum_m c_m^*\sum_n c_n a_n\delta_{mn}\end{aligned}$$

We have $\sum_n c_n a_n \delta_{mn} = c_m a_m$

$$\therefore \langle A\rangle = \sum_m c_m^* c_m a_m = \sum_m |c_m|^2 a_m$$

Let us recall that ψ should be a normalized wave function

i.e., $\int \psi^* \psi d\tau = 1$

$$\begin{aligned}\int\psi^*\psi d\tau &= \int\left(\sum_m c_m\phi_m\right)^*\left(\sum_n c_n\phi_n\right)d\tau \\ &= \sum_m\sum_n c_m^* c_n\int \phi_m^*\phi_n\, d\tau \\ &= \sum_m\sum_n c_m^* c_n\delta_{mn} \\ &= \sum_m c_m^*\sum_n c_n\delta_{mn} = \sum_m c_m^* c_m = \sum |c_m|^2 = 1\end{aligned}$$

Now, we have two results

$$\sum_m |c_m|^2 = 1$$

and
$$\langle A\rangle = \sum |c_m|^2 a_m \tag{5.11}$$

In statistics, if x is a random variable and $P(x_m)$ is the probability of getting a value x_m for random variable x, then we have

$$\sum_m P(x_m) = 1$$

and
$$\langle x\rangle = \sum_m P(x_m)x_m \tag{5.12}$$

As an example, consider the marks (the random variable x) obtained by 1000 students in a college. Let us say 60 students have 40 marks ($x_1 = 40$), 240 have 50 ($x_2 = 50$), 500 have 60 ($x_3 = 60$), 160 have 70 ($x_4 = 70$) and 40 have 80 ($x_5 = 80$). Probability of a student getting 40 marks is 60/1000 i.e.,$P(x_1) = \frac{60}{1000}$. Similarly, $P(x_2) = \frac{240}{1000}$, $P(x_3) = \frac{500}{1000}$, $P(x_4) = \frac{160}{1000}$ and $P(x_5) = \frac{40}{1000}$. The average mark is given by

$$\begin{aligned}\text{Average mark} &= \frac{\text{Total mark}}{\text{Number of students}} \\ &= \frac{60\times 40 + 240\times 50 + 500\times 60 + 160\times 70 + 70\times 80}{1000} \\ &= \frac{60}{1000}\times 40 + \frac{240}{1000}\times 50 + \frac{500}{1000}\times 60 + \frac{160}{1000}\times 70 + \frac{70}{1000}\times 80 \\ &= P(x_1)x_1 + P(x_2)x_2 + P(x_3)x_3 + P(x_4)x_4 + P(x_5)x_5\end{aligned}$$

Comparison of the set of the Equations (5.11) and (5.12) suggests that we can interpret $|c_m|^2$ as the probability, $P(a_m)$, of getting the result of measurement of A to be a_m when the system is in the state $\psi = \sum c_m \phi_m$.

i.e.,
$$P(a_m) = |c_m|^2$$

Note the statement of the postulate P-6. We have to first consider an ensemble or a collection of identical systems. All the systems are in the same state ψ. Still, when measurement of A is made on each system, we will get a statistical distribution of the experimental outcomes (the eigenvalues of A) a_1, $a_2, \ldots a_n$, instead of a single value. If one asks the question, what will be the result if the measurement of A be made on a single system, there is no definite answer. The answer provided by P-6 is that we can tell only the probability of getting the result of A to be a particular eigenvalue, say a_m, is $|c_m|^2$. The term *expectation value* itself indicates the statistical character of quantum mechanics. There is a statistical distribution of experimental outcomes when the value of a dynamical variable A is determined for all the systems in spite of the fact that all the systems are identical in identical state. This clearly brings out the indeterministic character of quantum mechanics. We will comment on this aspect later in detail.

Result of Measurement of a When the System is in An Eigenstate

Suppose we consider a system in a particular eigenstate ϕ_n. Then the state function ψ is given by

$$\psi = \phi_n$$

If we write ψ as $\psi = \sum c_m \phi_m$, then all the coefficient $c_m = 0$, excepting $c_n = 1$.

$$\begin{aligned}\text{i.e., } c_m &= \begin{cases} 1 & m = n \\ 0 & m \neq n \end{cases} \\ &= \delta_{mn}\end{aligned}$$

This implies that when a measurement of A is made on a system in state $\psi = \phi_n$, then the probability of getting the eigenvalue a_n is 1, whereas the probability of getting any other eigenvalue is zero.

$$P(a_n) = 1$$

and $$P(a_m) = 0 \text{ if } m \neq n$$

Note that in probability theory $P(a_n) = 1$ implies a definite and certain event. So, when we know the system to be in a particular eigenstate ϕ_n, we can boldly declare that the result of measurement of A will yield only a particular value, namely a_n. Therefore, we have two kinds of states for a system. One is the eigenstate and the other is the superposition of eigenstates. The distinction between these two states with respect to measurement is summarized below.

Ensemble of Identical Systems, All of them Being in the Same State

(a) State of each system: Superposition of eigenstates of A

$$\left.\begin{array}{l}\textit{Experimental outcomes of}\\ \textit{measurement of } A \textit{ for all systems}\end{array}\right\} \left\{\begin{array}{l}\textit{a statistical distribution of}\\ \textit{eigenvalues of } A\end{array}\right.$$

(b) State of each system: Eigenstate of A

$$\left.\begin{array}{l}\textit{Experimental outcomes of}\\ \textit{measurement of } A \textit{ for all systems}\end{array}\right\} \left\{\begin{array}{l}\textit{Identical outcomes for all the systems;}\\ \textit{No statistical distribution}\end{array}\right.$$

Single System

(a) State of the system: Superposition of eigenstates of A

$$\left.\begin{array}{l}\textit{Experimental outcomes of}\\ \textit{measurement of } A \textit{ for this system}\end{array}\right\} \left\{\begin{array}{l}\textit{Outcome is uncertain or unpredictable;}\\ \textit{any eigenvalue is possible}\end{array}\right.$$

(b) State of the system: Eigenstate of A

$$\left.\begin{array}{l}\textit{Experimental outcome of}\\ \textit{measurement of } A \textit{ for this system}\end{array}\right\} \left\{\begin{array}{l}\textit{definite outcome, without uncertainty;}\\ \textit{only one particular eigenvalue is possible}\end{array}\right.$$

Collapse of Wave Function

Superposition principle has more interesting physics than presented here. In fact, the superposition principle is something very special in quantum mechanics. Many of the philosophical problems in quantum physics centre around this principle.

A measurement of the dynamical variable A is made on a system which is in the state of superposition $\psi = \sum c_i\phi_i$, and suppose one gets the value of A to be a_n. What happens to the state of the system? The whole process of measurement is interpreted as follows in quantum mechanics: Before measurement, the system was in the state $\psi = \sum c_i\phi_i$. The measurement process has changed the state from $\psi = \sum c_i\phi_i$ to a state $\psi = \phi_n$. This is termed as the collapse of the wave function $\psi = \sum c_i \phi_i$ to the wave function $\psi = \phi_n$. The state of the system has now become $\psi = \phi_n$ immediately after measurement. So, the result of a measurement is the change of the state of the system from superposition to a particular eigenstate. If one measures A immediately, since the system is in $\psi = \phi_n$, the experimental value will be a_n again.

Non-classical Character of Superposition States – Illustration Using Spin

To illustrate the quantum mechanical consequence of superposition of states, we consider the states of spin angular momentum. The reader who is not familiar with spin angular momentum can come back to this discussion after learning spin angular momentum in Chapter 10. The spin angular momentum operators are given by

$$S_x = \frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad S_y = \frac{\hbar}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \qquad S_z = \frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

The eigenvalues of S_z are $\hbar/2$ and $-\hbar/2$, and the corresponding eigenvectors are given by $\chi(\uparrow) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\chi(\downarrow) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$.

$$S_z\chi(\uparrow) = \frac{\hbar}{2}\chi(\uparrow) \quad \text{and} \quad S_z\chi(\downarrow) = -\frac{\hbar}{2}\chi(\downarrow)$$

The states $\chi(\uparrow)$ and $\chi(\downarrow)$ are called Z spin-up and Z spin-down states for the following reasons. Though it is not proper in quantum mechanics to visualize them as spinning objects, we can consider these states as follows:

$\chi(\uparrow)$ represents anticlockwise spin, so the spin angular momentum vector points up along Z axis. Similarly, $\chi(\downarrow)$ represents clockwise spin, so the spin angular momentum points down along the Z axis.

Example 5.4 Let the state of a spin system be given by $\psi = 0.8\,\chi(\uparrow) + 0.6\,\chi(\downarrow)$. What is the probability of getting (i) spin up (ii) spin down along Z axis when Z component of the spin is measured?

Solution: Probability of getting spin up $= 0.8^2 = 0.64$

Probability of getting spin down $= 0.6^2 = 0.36$

Example 5.5 In the previous problem, on measuring Z component of spin, we get it to be $\hbar/2$. What is the state of the system immediately after the measurement process? If, again, the Z component of the spin is measured, what will be the experimental outcome?

Solution: The measurement process changes the state $\psi = 0.8\,\chi(\uparrow) + 0.6\,\chi(\downarrow)$ to the state $\chi(\uparrow)$. So, immediately after the measurement, the state of the system is $\psi = \chi(\uparrow)$.

When we measure the Z component of the spin again, the experimental outcome will be $\frac{\hbar}{2}$.

Example 5.6 Determine the average value of x component of spin when the system is in state (i) $\chi(\uparrow)$ (ii) $\chi(\downarrow)$

Solution: When the state of the system is in state $\chi(\uparrow)$

$$\langle S_x \rangle = \chi(\uparrow)^\dagger S_x \chi(\uparrow) = (1 \quad 0)\frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = 0$$

When the state of the system is in the state $\chi(\downarrow)$

$$\langle S_x \rangle = \chi(\downarrow)^\dagger S_x \chi(\downarrow) = (0 \quad 1)\frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \end{pmatrix} = 0$$

Example 5.7 Consider an ensemble of 1000 systems in which 640 systems are in state $\chi(\uparrow)$ and 360 systems are in state $\chi(\downarrow)$. Measure the X component of spin of each system. Determine the average value of X component of spin for the ensemble.

Solution:

$$\langle S_x \rangle = \frac{640 \times \langle S_x \rangle \text{ for state } \chi(\uparrow) + 360 \times \langle S_x \rangle \text{ for state } \chi(\downarrow)}{1000} = 0$$

Example 5.8 Determine the average value of X component of spin for an ensemble of 1000 systems in which all the systems are in the state $\psi = 0.8\chi(\uparrow) + 0.6\chi(\downarrow)$

Solution:

$$\psi = 0.8\,\chi(\uparrow) + 0.6\,\chi(\downarrow) = \begin{pmatrix} 0.8 \\ 0.6 \end{pmatrix}$$

$$\langle S_x \rangle = \psi^\dagger S_x \psi = (0.8 \quad 0.6)\frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0.8 \\ 0.6 \end{pmatrix} = 0.48\hbar$$

Now, let us ask the following question: What is the value of Z component of spin of a system whose state is $\psi = 0.8\,\chi(\uparrow) + 0.6\,\chi(\downarrow)$? Can we claim that the the system in this state has a definite value for Z spin which is either $\hbar/2$ or $-\hbar/2$?. Examples 5.7 and 5.8 together rule out such a possibility. If we assume a system in state $\psi = 0.8\,\chi(\uparrow) + 0.6\,\chi(\downarrow)$ has definite spin, either $\hbar/2$ or $-\hbar/2$, then in an ensemble of 1000 systems, roughly 640 systems will have Z spin up and roughly 360 systems will have Z spin down. The ensemble in 5.7 is such an ensemble for which $\langle S_x \rangle$ is zero. In Example 5.8, we do not make any assumption of the value of Z component of spin, and in this case, $\langle S_x \rangle = 0.48\ \hbar$. The only difference between these two ensembles is that in the ensemble in Example 5.7, each system has a well-defined Z spin value, while in the ensemble in Example 5.8, there is no value for the Z spin for a system. The fact that they lead to different results for $\langle S_x \rangle$ suggests these two ensembles are not equivalent. Therefore, it is not correct to claim that a system has definite eigenvalue either $\hbar/2$ or $-\hbar/2$, for Z component of spin when the system is in the superposition state $\psi = 0.8\,\chi(\uparrow) + 0.6\,\chi(\downarrow)$. However, after a measurement, the system has a definite value for Z spin.

So, we have three possible states in quantum mechanics. They are as follows:

1. State $\chi(\uparrow)$ representing a system with Z component of spin $\frac{\hbar}{2}$. The classical analogue is an object spinning anticlockwise.
2. State $\chi(\downarrow)$ representing system with Z component of spin $-\frac{\hbar}{2}$. The classical analogue is an object spinning clockwise.
3. Superposition of eigenstates like $\psi = 0.8\,\chi(\uparrow) + 0.6\,\chi(\downarrow)$. The Z component of the spin in this state is neither $\hbar/2$ nor $-\hbar/2$ (no other value is possible); it is undefined. There is no classical analogue to this state of an object, so this state is not comprehensible with our experience in macro world.

Extending this discussion to a general system, we conclude that the value of a dynamical variable A is undefined when the system is in the state $\psi = \sum C_n \phi_n$.

Postulate P-7

The time evolution of the state of the system is determined by the Schrödinger equation

$$i\hbar\frac{\partial\psi}{\partial t}(\mathbf{r},t)=H_{op}\psi(\mathbf{r},t)$$

or
$$i\hbar\frac{\partial\psi}{\partial t}(\mathbf{r},t)=\frac{-\hbar^2}{2m}\nabla^2\psi(\mathbf{r},t)+V(\mathbf{r},t)\psi(\mathbf{r},t) \tag{5.13}$$

We have already observed there is no way of deriving Schrodinger equation from any other fundamental principle. That is why this equation is given as a postulate of quantum mechanics.

All the postulates of quantum mechanics have been presented in full detail. In the rest of the chapter, we will concentrate on some general aspects of quantum mechanics without going into the specific systems. These topics will complete the general formalism.

5.2 COMMUTATION RELATION BETWEEN *x* AND p_x

The operators for x and p_x are given by

$$x_{op}=x \quad \text{and} \quad p_{xop}=-i\hbar\frac{\partial}{\partial x}$$

For the sake of simplicity, let us drop the subscripts 'op'.

For an arbitrary wave function ψ,

$$(xp_x-p_xx)\psi=x\,(p_x\psi)-p_x(x\psi)$$

$$=x\left(-i\hbar\frac{\partial\psi}{\partial x}\right)-\left(-i\hbar\frac{\partial}{\partial x}\right)(x\psi)=-i\hbar x\frac{\partial\psi}{\partial x}+i\hbar x\frac{\partial\psi}{\partial x}+i\hbar\psi=i\hbar\psi.$$

Thus, we have

$$(xp_x-p_xx)\psi=i\hbar\psi$$

This equation is true for any arbitrary ψ. Therefore, we conclude

$$xp_x-p_xx=i\hbar$$

We have not proved Dirac quantum condition. The operators x and p_x have been chosen in such a way that it is consistent with Dirac quantum condition.

Example 5.9 Evaluate $[x_i, p_j]$.

Solution: A more compact way of expressing the commutation relation among the three Cartesian coordinates *x, y* and *z* and the three corresponding momentum operators p_x, p_y and p_z are to use the following notation:

Let us define x_1, x_2 and x_3, p_1, p_2 and p_3, as

$$x_1=x,\quad x_2=y,\quad x_3=z$$

$$p_1=p_x,\quad p_2=p_y,\quad p_3=p_z$$

The operators p_x, p_y and p_z can be written as

$$p_j = -i\hbar \frac{\partial}{\partial x_j}$$

Now we have

$$[x_i, p_j]\psi = [x_i p_j - p_j x_i]\psi = \left[x_i\left(-i\hbar\frac{\partial \psi}{\partial x_j}\right) - \left(-i\hbar\frac{\partial}{\partial x_j}\right)(x_i\psi)\right]$$

$$= -i\hbar x_i \frac{\partial \psi}{\partial x_j} + i\hbar x_i \frac{\partial \psi}{\partial x_j} + i\hbar\psi\frac{\partial x_i}{\partial x_j}$$

$$= i\hbar\delta_{ij}\psi$$

i.e., $$[x_i, p_j]\psi = i\hbar\delta_{ij}\psi$$

Since this equation is true for any arbitrary ψ,

$$[x_i, p_j] = i\hbar\delta_{ij}$$

Example 5.10 Evaluate $[x_i, x_j]$.

Solution: $[x_i, x_j]\psi = [x_i x_j - x_j x_i]\ \psi = 0$

Since this equation is true for any arbitrary ψ, we write

$$[x_i, x_j] = 0$$

Similarly, it can be easily shown that $[p_i, p_j] = 0$.

We can summarize all the above commutation as follows:

$$[x_i, x_j] = [p_i, p_j] = 0 \text{ and } [x_i, p_j] = i\hbar\delta_{ij} \tag{5.14}$$

This means that in any expression involving x, y and z and p_x, p_y and p_z, the ordering between x and p_x, y and p_y and z and p_z have to be maintained as it is. Otherwise, other quantities can be flipped around. For instance, consider xp_yp_zy. Within this expression, the order p_yy alone has to be maintained.

i.e., $$xp_yp_zy = xp_zp_yy = p_yxp_zy = p_yp_xxy = p_zp_yxy = p_zxp_yy$$

Example 5.11 Prove that

1. $[A, BC] = [A, B]C + B[A, C]$
2. $[AB, C] = A[B, C] + [A, C]B$

Solution:

1. $$[A, BC] = ABC - BCA = ABC - BAC + BAC - BCA$$
$$= (AB - BA)C + B(AC - CA)$$
$$= [A, B]C + B\,[A, C] \tag{5.15}$$

2. $$[A, BC] = ABC - CAB$$
$$= ABC - ACB + ACB - CAB$$
$$= A\,(BC - CB) + (AC - CA)B$$
$$= A\,[B, C] + [A, C]B \tag{5.16}$$

Example 5.12 Prove that $[L_x, L_y] = i\hbar L_z$

Solution: The components L_x, L_y and L_z are given by

$$
\begin{aligned}
L_x &= yp_z - zp_y, \quad L_y = zp_x - xp_z, \quad L_z = xp_y - yp_x \\
[L_x, L_y] &= L_xL_y - L_yL_x = (yp_z - zp_y)\,(zp_x - xp_z) - (zp_x - xp_z)\,(yp_z - zp_y) \\
&= yp_zzp_x - yp_zxp_z - zp_yzp_{\mathrm{x}} + zp_yxp_z - zp_xyp_z + zp_xzp_y + xp_zyp_z - xp_zzp_y \\
&= xp_y\,(zp_z - p_zz) - yp_x\,(zp_z - p_zz) \\
&= i\hbar\,[xp_y - yp_x] = i\hbar L_z
\end{aligned}
\tag{5.17}
$$

We have used the fact that in the second, third, sixth and seventh terms, the order of four quantities is immaterial, and these terms cancel among themselves.

Example 5.13 Evaluate $[f(x), p_x]$.

Solution: Consider $[f(x), p_x]\ \psi$

$$
\begin{aligned}
[f(x), p_x]\psi &= [f(x)\,p_x - p_xf(x)]\,\psi \\
&= f(x)\left(-i\hbar\frac{\partial\psi}{\partial x}\right) - \left(-i\hbar\frac{\partial}{\partial x}\right)(f(x)\psi) \\
&= -i\hbar\,f(x)\frac{\partial\psi}{\partial x} + i\hbar\frac{\partial f}{\partial x}\psi + i\hbar\,f(x)\frac{\partial\psi}{\partial x} \\
&= i\hbar\frac{\partial f}{\partial x}\psi
\end{aligned}
$$

i.e.,
$$[f(x), p_x]\psi = i\hbar\frac{\partial f}{\partial x}\psi.$$

Since this equation is true for any arbitrary ψ,

$$[f(x), p_x] = i\hbar\frac{\partial f}{\partial x} \tag{5.18}$$

5.3 COMPATIBLE OBSERVABLES

Two dynamical observables A and B are said to be compatible if there exists a complete set of functions which are eigenfunctions of both A and B.

If $\{\phi_i\}$ are complete set of functions such that

$$A\phi_i = a_i\phi_i \text{ and } B\phi_i = b_i\phi_i$$

then A and B are said to be compatible.

Example 5.14 Prove that if A and B are compatible observables, then their operators commute.

Solution: We have to first note that in order to prove $[A, B] = 0$, we have to prove $(AB - BA)\ \psi = 0$, where ψ is an arbitrary function.

Since A and B are compatible observables,

$$A\phi_i = a_i\phi_i, \quad B\phi_i = b_i\phi_i$$

Let us consider an arbitrary wave function $\psi = \sum c_n \phi_n$

$$AB\psi = \sum c_n AB\phi_n = \sum_n c_n Ab_n\phi_n = \sum c_n a_n b_n \phi_n$$

$$BA\psi = \sum c_n BA\phi_n = \sum_n c_n Ba_n\phi_n = \sum c_n a_n b_n \phi_n$$

$$\therefore (AB - BA)\psi = \sum c_n a_n b_n \phi_n - \sum c_n a_n b_n \phi_n = 0$$

Since this equation is true for an arbitrary ψ,

$$AB - BA = 0 \quad \text{or} \quad AB = BA.$$

Example 5.15 Prove that if A and B are two commuting operators, then they have a set of common eigenfunctions.

Solution: This is proved by considering two cases, namely in the first case, the eigenfunctions of one of the operators, say A, is non-degenerate and in the second case, it is relaxing this condition.
Case 1: The eigenvalues of A are non-degenerate.

The eigenvalue equation for A is

$$A\phi_i = a_i\phi_i$$

$$BA\phi_i = Ba_i\phi_i = a_iB\phi_i$$

But since $BA = AB$, we have $BA\phi_i = AB\phi_i$

$$\therefore AB\phi_i = a_iB\phi_i$$

i.e., $B\phi_i$ is an eigenfunction of a_i. This implies that $B\phi_i$ and ϕ_i have to be the same or at most they can differ by a multiplicative constant.

i.e., $$B\phi_i = b_i\phi_i$$

$\therefore \phi_i$ is common eigenfunction of both A and B.
Case 2: The eigenvalues of A are degenerate.

Consider the eigenvalue a, which is α-fold degenerate, that is, there exists α linearly independent eigenfunctions of A, corresponding to the same eigenvalue a.

i.e., we have $$A\phi_j = a\phi_j \quad j = 1, 2, 3, \ldots \alpha$$

Note that any linear combination $\psi = \sum d_j\phi_j$ is also an eigenfunction of A corresponding to the eigenvalue a.

$$A\,\psi = A\sum d_j\phi_j = \sum_j d_jA\phi_j = \sum d_ja\phi_j$$

$$= a\sum d_j\phi_j = a\psi$$

This implies that to prove the existence of a set of common eigenfunctions of A and B, we have to prove that it is possible to find a linear combination of $\phi_i s$, which is also an eigenfunction of B; in other words we have to find a set of r_1, r_2, … such that

$$B\psi = b\psi$$

where
$$\psi = \sum r_n \phi_n = r_1 \phi_1 + r_2 \phi_2 + \ldots r_n \phi_n \tag{5.19}$$

Consider $BA\phi$

$$BA\phi_j = Ba\phi_j = aB\phi_j$$

But
$$BA\phi_j = AB\phi_j$$

$$\therefore\ A(B\phi_j) = a(B\phi_j)$$

i.e., $B\phi_j$ is also an eigenfunction of A with eigenvalue a. This implies that $B\phi_j$ should be a linear combination $\phi_j s$, since, as we have mentioned earlier, any linear combination of $\phi_j s$ is also an eigenfunction of A with eigenvalue a. Therefore, we can write

$$B\phi_j = \sum c_{jk}\phi_k \tag{5.20}$$

Let us write the eigenvalue equation for B as

$$B\psi = b\psi$$

To make this discussion more transparent, let us assume a is three-fold degenerate eigenvalue, that is, there are three linearly independent eigenfunctions ϕ_1, ϕ_2 and ϕ_3, which have the same eigenvalue

i.e.,
$$A\phi_1 = a\phi_1; \quad A\phi_2 = a\phi_2; \quad A\phi_3 = a\phi_3$$

The meaning of (5.20) is

$$\begin{aligned} B\phi_1 &= c_{11}\phi_1 + c_{12}\phi_2 + c_{13}\phi_3 \\ B\phi_2 &= c_{21}\phi_1 + c_{22}\phi_2 + c_{23}\phi_3 \\ B\phi_3 &= c_{31}\phi_1 + c_{32}\phi_2 + c_{33}\phi_3 \end{aligned} \tag{5.21}$$

The eigenvalue equation of B is (use (5.19))

$$B\psi = b\psi = b[r_1\phi_1 + r_2\phi_2 + r_3\phi_3] \tag{5.22}$$

But
$$\begin{aligned} B\psi &= B[r_1\phi_1 + r_2\phi_2 + r_3\phi_3] \\ &= r_1 B\phi_1 + r_2 B\phi_2 + r_3 B\phi_3 \\ &= r_1c_{11}\phi_1 + r_1c_{12}\phi_2 + r_1c_{13}\phi_3 + r_2c_{21}\phi_1 + r_2c_{22}\phi_2 + r_2c_{23}\phi_3 + r_3c_{31}\phi_1 + r_3c_{32}\phi_2 + r_3c_{33}\phi_3 \\ &= [r_1c_{11} + r_2c_{21} + r_3c_{31}]\phi_1 + [r_1c_{12} + r_2c_{22} + r_3c_{32}]\phi_2 + [r_1c_{13} + r_2c_{23} + r_3c_{33}]\phi_3 \end{aligned} \tag{5.23}$$

From Equations (5.22) and (5.23), we get

$$\begin{aligned} r_1 b &= r_1c_{11} + r_2c_{21} + r_3c_{31} \\ r_2 b &= r_2c_{21} + r_2c_{22} + r_3c_{32} \\ r_3 b &= r_3c_{13} + r_2c_{23} + r_3c_{33} \end{aligned}$$

$$b(r_1 \quad r_2 \quad r_3) = (r_1 \quad r_2 \quad r_3)\begin{pmatrix} c_{11} & c_{21} & c_{31} \\ c_{12} & c_{22} & c_{32} \\ c_{13} & c_{23} & c_{33} \end{pmatrix}$$

or taking transpose

$$\begin{pmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{pmatrix}\begin{pmatrix} r_1 \\ r_2 \\ r_3 \end{pmatrix} = b\begin{pmatrix} r_1 \\ r_2 \\ r_3 \end{pmatrix} \tag{5.24}$$

This is the standard eigenvalue equation. By solving these equations, we will get a set of eigenvalues b as well as the corresponding eigenvectors. In our case, we will get the eigenvalues b_1, b_2 and b_3. Thus, we can find a set of values for $r_1, r_2, \ldots$ satisfying the Equation (5.18) that is, we have

1. $A\psi_1 = a\psi_1$, $\psi_1 = r_{11}\phi_1 + r_{12}\phi_2 + r_{13}\phi_3$
   $B\psi_1 = b_1\psi_1$ (r_{11}, r_{12}, r_{13}) is the first eigenvector of Equation (5.24)
2. $A\psi_2 = a\psi_2$, $\psi_2 = r_{21}\phi_1 + r_{22}\phi_2 + r_{23}\phi_3$
   $B\psi_2 = b_2\psi_2$ (r_{21}, r_{22}, r_{23}) is the second eigenvector of Equation (5.24)
3. $A\psi_3 = a\psi_3$, $\psi_3 = r_{31}\phi_1 + r_{32}\phi_2 + r_{33}\phi_3$
   $B\psi_3 = b_3\psi_3$ (r_{31}, r_{32}, r_{33}) is the third eigenvector of Equation (5.24).

Though we have explicitly demonstrated the existence of common eigenfunctions of A and B in a particular case where the eigenvalue a is three-fold degenerate, it is true for any arbitrary m-fold degenerate eigenvalue a.

5.4 UNCERTAINTY PRINCIPLE

We have already discussed the Heisenberg uncertainty principle in Chapter 1. Here, we establish the uncertainty principle in a more rigorous way. Let us consider an ensemble of identical systems in which all the systems are in the same state ψ.

Let A and B be two dynamical variables such that their commutation relation is given by

$$[A_{op}, B_{op}] = i\hbar C_{op}$$

The uncertainty ΔA and the uncertainty ΔB are defined through the relations

$$(\Delta A)^2 = \langle (A_{op} - \langle A\rangle)^2\rangle \tag{5.25}$$

and

$$(\Delta B)^2 = \langle (B_{op} - \langle B\rangle)^2\rangle \tag{5.26}$$

For a system in the state ψ, $\langle A\rangle$ and $\langle B\rangle$ are given by

$$\langle A\rangle = \int \psi^* A_{op}\psi d\tau \quad \text{and} \quad \langle B\rangle = \int \psi^* B_{op}\psi d\tau.$$

Let us define U_{op} and V_{op} as

$$U_{op} = A_{op} - \langle A\rangle \text{ and } V_{op} = B_{op} - \langle B\rangle$$

Then, the uncertainty relations (5.25) and (5.26) imply

$$(\Delta A)^2 = \int \psi^* U_{op}^2 \psi d\tau$$

$$(\Delta B)^2 = \int \psi^* V_{op}^2 \psi d\tau$$

Note that since A and B are dynamical variables, A_{op} and B_{op} are Hermitian operators, which in turn imply that U_{op} and V_{op} are Hermitian operators. Let us recall that for any Hermitian operator P, we have

$$(\psi, P\phi) = (P\psi, \phi)$$

or

$$\int \psi^* P\phi d\tau = \int (P\psi)^* \phi d\tau \tag{5.27}$$

Let us define a new function ϕ as

$$\phi = (U_{op} + i\lambda V_{op})\psi$$

where λ is a real number. Then, the integral I is defined as

$$I = \int \phi^* \phi\, d\tau$$

Since I can only be a function of free parameters in ϕ, we can take I to be a function of λ. i.e.,

$$I = I(\lambda) = \int \phi^* \phi d\tau$$

Since the integrand $\phi^*\phi$ is always positive, $I(\lambda)$ has to be positive number.

i.e., $$I(\lambda) \geq 0$$

The equality $I(\lambda) = 0$ implies $\phi = 0$. i.e.,

$$(U_{op} + i\lambda V_{op})\psi = 0 \tag{5.28}$$

Let us evaluate $I(\lambda)$.

$$\begin{aligned} I(\lambda) &= \int \phi^*\phi\, d\tau \\ &= \int [(U_{op} + i\lambda V_{op})\psi]^* [(U_{op} + i\lambda V_{op})\psi] d\tau \\ &= \int [(U_{op}\psi)^* - i\lambda(V_{op}\psi)^*]\, [(U_{op}\psi) + i\lambda(V_{op}\psi)] d\tau \\ &= \int [U_{op}\psi]^* U_{op}\psi d\tau + \lambda^2 \int (V_{op}\psi)^* V_{op}\psi d\tau + i\lambda \\ &\quad \left[\int (U_{op}\psi)^* V_{op}\psi d\tau - (V_{op}\psi)^* (U_{op}\psi) d\tau\right] \end{aligned} \tag{5.29}$$

Since U_{op} is an Hermitian operator, making use of (5.27), we can write the first term in (5.28) as

$$\begin{aligned} \int [U_{op}\psi]^* (U_{op}\psi) d\tau &= \int \psi^* U_{op} [U_{op}\psi]\, d\tau \\ &= \int \psi^* U_{op}^2 \psi d\tau \end{aligned}$$

The other integrals in (5.28) can be rewritten in the same way.

$$\int (U_{op}\psi)^* V_{op}\psi d\tau = \int \psi^* U_{op} V_{op}\psi d\tau$$

$$\int (V_{op}\psi)^* U_{op}\psi d\tau = \int \psi^* V_{op} U_{op}\psi d\tau$$

and $$\int (V_{op}\psi)^*(V_{op}\psi)d\tau = \int \psi^* V_{op}^2 \psi d\tau$$

Writing all these terms together in (5.29), we get

$$\begin{aligned} I(\lambda) &= \int \psi^* U_{op}^2 \psi d\tau + \lambda^2 \int \psi^2 V_{op}^2 \psi d\tau + i\lambda \int \psi^*(U_{op}V_{op} - V_{op}U_{op})\psi d\tau \\ &= (\Delta A)^2 + \lambda^2 (\Delta B)^2 + i\lambda \langle [U_{op}, V_{op}] \rangle \end{aligned}$$

We have, $[U_{op}, V_{op}] = U_{op}V_{op} - V_{op}U_{op}$

$$\begin{aligned} &= (A_{op} - \langle A \rangle)(B_{op} - \langle B \rangle) - (B_{op} - \langle B \rangle)(A_{op} - \langle A \rangle) \\ &= A_{op}B_{op} - A_{op}\langle B \rangle - \langle A \rangle B_{op} + \langle A \rangle\langle B \rangle - B_{op}A_{op} + B_{op}\langle A \rangle + \langle B \rangle A_{op} - \langle A \rangle \langle B \rangle \end{aligned}$$

Since $\langle A \rangle$ and $\langle B \rangle$ are numbers, $A_{op}\langle B \rangle = \langle B \rangle A_{op}$, $\langle A \rangle B_{op} = B_{op}\langle A \rangle$ and $\langle A \rangle\langle B \rangle = \langle B \rangle \langle A \rangle$.

Therefore, $[U_{op}, V_{op}]$ is given by

$$[U_{op}, V_{op}] = [A_{op}B_{op} - B_{op}A_{op}] = [A_{op}, B_{op}]$$

So, $I(\lambda)$ is given by

$$I(\lambda) = (\Delta A)^2 + \lambda^2 (\Delta B)^2 + i\lambda \langle [A_{op}, B_{op}] \rangle$$

Since $I(\lambda)$ is always positive, the minimum value of $I(\lambda)$ has also to be always positive for any ψ. The general form may look like the curves as shown in Fig. 5.2.

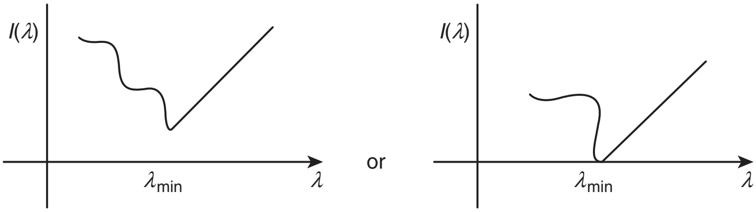


Fig. 5.2 Sketch of $i(\lambda)$

Let us, therefore, minimize $I(\lambda)$. The $\lambda_{\min}$ is obtained by solving $\frac{d}{d\lambda} I(\lambda) = 0$.

$$\frac{d}{d\lambda} I(\lambda) = 2\lambda (\Delta B)^2 + i\langle [A_{op}, B_{op}] \rangle = 0$$

or $$\lambda_{\min} = -\frac{i\langle [A_{op}, B_{op}] \rangle}{2(\Delta B)^2}$$

The minimum value of $I(\lambda)$ has to be greater than or equal to zero as shown in Fig. 5.2.

$$\min I(\lambda) \geq 0 \tag{5.30}$$

i.e.,
$$(\Delta A)^2 + \left\{\frac{-i\langle[A_{op}, B_{op}]\rangle}{2(\Delta B)^2}\right\}^2 (\Delta B)^2 + i\left\{\frac{-i\langle[A_{op}, B_{op}]\rangle}{2(\Delta B)^2}\right\}\langle[A_{op}, B_{op}]\rangle \geq 0$$

$$(\Delta A)^2 - \frac{\langle[A_{op}, B_{op}]\rangle^2}{4(\Delta B)^4}(\Delta B)^2 + \frac{\langle[A_{op}, B_{op}]\rangle^2}{2(\Delta B)^2} \geq 0$$

Multiplying throughout by $(\Delta B)^2$, we have

$$(\Delta A)^2(\Delta B)^2 - \frac{1}{4}\langle[A_{op}, B_{op}]\rangle^2 + \frac{1}{2}\langle[A_{op}, B_{op}]\rangle^2 \geq 0$$
$$(\Delta A)^2(\Delta B)^2 \geq -\frac{1}{4}\langle[A_{op}, B_{op}]\rangle^2$$

This is the inequality for the product of uncertainties ΔA and ΔB. In particular, the uncertainty relation for the position x and momentum p is given by

$$(\Delta x)^2(\Delta p)^2 \geq -\frac{1}{4}\langle[x_{op}, p_{op}]\rangle^2$$

Now
$$[x_{op}, p_{op}] = i\hbar$$

Therefore,

$$\langle[x_{op}, p_{op}]\rangle = \int \psi^*[x_{op}, p_{op}]\psi d\tau = \int \psi^* i\hbar \psi d\tau = i\hbar$$
$$\therefore \langle[x_{op}, p_{op}]\rangle^2 = (i\hbar)^2 = -\hbar^2$$

Therefore, we have

$$(\Delta x)^2(\Delta p_x)^2 \geq \frac{\hbar^2}{4}$$

or
$$\Delta x \Delta p \geq \frac{\hbar}{2} \tag{5.31}$$

Example 5.16 Show that $(\Delta A)^2 = \left\langle \left(A_{op} - \langle A\rangle\right)^2 \right\rangle = \langle A^2\rangle - \langle A\rangle^2$

Solution:

$$(\Delta A)^2 = \int \psi^*(A_{op} - \langle A\rangle)^2 \psi d\tau$$
$$= \int \psi^*[A_{op}^2 - \langle A\rangle A_{op} - A_{op}\langle A\rangle + \langle A\rangle^2]\psi d\tau$$
$$= \langle A^2\rangle - \langle A\rangle^2$$

Eigenstate and Uncertainty

Note that the value of the uncertainty of a dynamical variable depends on the state of the system. There are a group of states for which the uncertainty of a dynamical variable A is zero. These are the eigenstates of A. Next, consider the other type of states. Let the state of all the systems in an ensemble be $\psi = \sum C_n\phi_n$, where ϕ_n is an eigenstate of A with eigenvalue a_n. If one measures the dynamical variable A for each system in this ensemble, then the experimental outcomes will be a statistical distribution of eigenvalues of A, and hence the uncertainty ΔA for this state will be non-zero. In fact, it is very easy to calculate the uncertainty ΔA.

$$(\Delta A)^2 = \sum |C_n|^2 a_n^2 - \left(\sum |C_n|^2 a_n\right)^2$$

We can use this expression for systems in eigenstates as well since an eigenstate is a special case of superposition of states. If the states of all the systems in an ensemble are described by a particular eigenstate, say $\phi_m, (\Delta A)^2 = \langle A^2 \rangle - \langle A \rangle^2 = (a_m^2) - (a_m)^2 = 0$. So we have

$$\Delta A \neq 0 \quad \text{for superposition of eigenstates of A}$$

$$\Delta A = 0 \quad \text{for eigenstate of A}$$

Uncertainties and Experimental Errors

The uncertainties defined through the relations (5.25) or (5.26) are different from the experimental errors. The uncertainty defined through the relations (5.25) or (5.26) have nothing to do with details of instruments like least count. They arise due to intrinsic probabilistic interpretation of quantum mechanics. In principle, the uncertainty may even be greater than the least count of the instrument, so the uncertainty may be detected. (Least count may be 0.5 Å while the uncertainty may be 2 Å). There are many sources of experimental errors. Normally, the experimental errors will be large enough to wash out these uncertainties. The continuous improvements in technology will reduce the errors. However, the uncertainties can never be eliminated (see L. Ballentine).

Energy–Time Uncertainty Relation

The above derivation of the uncertainty cannot be used to obtain the energy–time uncertainty relation $\Delta E \Delta t \sim h$. Mathematically, we can have the commutation relation $[E,t] = [i\hbar\,\partial/\partial t, t] = i\hbar$ leading to the uncertainty relation $\Delta E \Delta t \geq \hbar/2$. However, this argument is physically invalid. The reason is that time t is not a dynamical variable like position or momentum. It is a parameter like mass or charge in the theory. So, no operator can be associated with time. In fact, there can be no statistical spread for time similar to position or momentum. If one measures the energy of a system, we will get a statistical spread ΔE in energy, depending on the time taken to measure the energy. The time taken to measure energy of the system is denoted by Δt. The energy–time relation $\Delta E \Delta t \sim h$ will be derived later in Chapter 14. Note that the uncertainty in energy E is similar to the uncertainties in position or momentum. This arises from the fact that we can have a state that is a superposition of the energy eigenstates.(See Chapter 2.)

5.4.1 Wave Function Corresponding to Minimum Uncertainty Product

It is obvious from the inequality given in (5.31) that the minimum value of the product $\Delta x \Delta p$ is $\frac{\hbar}{2}$. i.e.,

$$\Delta x\, \Delta p = \frac{\hbar}{2}$$

The inequality (5.31) has become an equality.

Tracking back the inequality from (5.31), we observe that for $\Delta x\,\Delta p = \frac{\hbar}{2}$, we should have min $I(\lambda) = 0$.

$$\min I(\lambda) = (U_{op} + i\lambda_{\min} V_{op})\,\psi = 0 \tag{5.32}$$

But

$$\lambda_{\min} = \frac{-i\langle[x_{op}, p_{op}]\rangle}{2(\Delta p)^2} = \frac{-ii\hbar}{2(\Delta p)^2} = \frac{\hbar}{2(\Delta p)^2}$$

Therefore, (5.32) implies

$$\left[x_{op} - \langle x\rangle + \frac{i\hbar}{2(\Delta p)^2}\left(p_{op} - \langle p\rangle\right)\right]\psi = 0$$

Substituting$-i\hbar\frac{d}{dx}$ for p_{op}, we get

$$\left[x - \langle x\rangle + \frac{i\hbar}{(2\Delta p)^2}\left[-i\hbar\frac{d}{dx} - \langle p\rangle\right]\right]\psi(x) = 0$$

$$\frac{\hbar^2}{2(\Delta p)^2}\frac{d\psi}{dx} = -x\psi + \langle x\rangle\psi + \frac{i\hbar\langle p\rangle}{2(\Delta p)^2}\psi$$

Rearranging this equation, we get

$$\frac{d\psi}{\psi} = \left[-2\frac{(\Delta p)^2}{\hbar^2}x + \frac{2(\Delta p)^2}{\hbar^2}\langle x\rangle + \frac{i\langle p\rangle}{\hbar}\right]dx$$

Integrating the above equation, we get

$$\begin{aligned}
\ln\frac{\psi}{\psi_0} &= \frac{-(\Delta p)^2}{\hbar^2}x^2 + \frac{2(\Delta p)^2}{\hbar^2}\langle x\rangle x + \frac{i}{\hbar}\langle p\rangle x \\
&= \frac{-(\Delta p)^2}{\hbar^2}\left[x^2 - 2\langle x\rangle x + 4\langle x\rangle^2 - 4\langle x\rangle^2\right] + \frac{i\langle p\rangle}{\hbar}x \\
&= \frac{-(\Delta p)^2}{\hbar^2}\left[x - \langle x\rangle\right]^2 + \frac{(\Delta p)^2}{\hbar^2}4\langle x\rangle^2 + \frac{i\langle p\rangle x}{\hbar} \\
\psi &= \psi_0 \exp\left[\frac{-(\Delta p)^2}{\hbar^2}(x - \langle x\rangle)^2\right]\cdot\exp\left[i\frac{\langle p\rangle}{\hbar}x\right]\exp\left[\frac{(\Delta p)^2}{\hbar^2}4\langle x\rangle^2\right]
\end{aligned}$$

Since exp $\left[(\Delta p)^2 4\langle x\rangle^2/\hbar^2\right]$ is a constant, it can be absorbed in the normalization constant N. i.e.,

$$\psi = Ne^{-} - \frac{(\Delta p)^2}{\hbar^2}\left(x - \langle x\rangle\right)^2 e^{i}\,\frac{\langle p\rangle x}{\hbar}$$

This is a Gaussian wave packet. So the wave function corresponding to $\Delta x\,\Delta p = \frac{\hbar}{2}$ is a Gaussian wave packet.

Example 5.17 Evaluate Δx and Δp for a system when its state is described by $\psi = Ne^{-\alpha x^2}$.

Solution:

Note: $$I(\beta) = \int_{-\infty}^{\infty} e^{-\beta x^2} dx = \sqrt{\frac{\pi}{\beta}}$$

$$\int_{-\infty}^{\infty} e^{-\beta x^2} x^2 dx = \frac{-d}{d\beta}\int_{-\infty}^{\infty} e^{-\beta x^2} = -\frac{dI(\beta)}{d\beta} = \frac{1}{2}\sqrt{\pi}\beta^{-3/2}$$

Let us first calculate the normalization constant N.

$$\int_{-\infty}^{\infty} \psi^* \psi dx = N^2 \int_{-\infty}^{\infty} e^{-2\alpha x^2} dx = N^2\sqrt{\frac{\pi}{2\alpha}} = 1$$

$$\therefore N = \left(\frac{2\alpha}{\pi}\right)^{1/4}$$

The expectation value $\langle x\rangle$ is given by

$$\langle x\rangle = \int_{-\infty}^{\infty} \psi^*(x) x \psi(x) dx$$

$$= \int_{-\infty}^{\infty} e^{-\alpha x^2} x e^{-\alpha x^2} dx = \int_{-\infty}^{\infty} e^{-2\alpha x^2} x dx = 0$$

The integral is zero since the integrand is an odd function of x in the interval $(-\infty, \infty)$.

$$\langle x^2\rangle = \int_{-\infty}^{\infty} \psi^* x^2 \psi dx$$

$$= \int_{-\infty}^{\infty} N^2 e^{-2\alpha x^2} x^2 dx = \frac{1}{2}\sqrt{\pi}(2\alpha)^{-3/2} N^2 = \frac{1}{2}\cdot N^2\sqrt{\frac{\pi}{2\alpha}}\cdot\frac{1}{2\alpha}$$

$$= \frac{1}{4\alpha} \quad \text{since} \quad N^2\sqrt{\frac{\pi}{2\alpha}} = 1$$

$$\langle p\rangle = \int_{-\infty}^{\infty} \psi^* p_{op} \psi dx = N^2 \int_{-\infty}^{\infty} e^{-\alpha x^2}\left(-i\hbar\frac{d}{dx}\right) e^{-\alpha x^2} dx$$

$$= N^2 \int_{-\infty}^{\infty} e^{-\alpha x^2}(-i\hbar)(-\alpha)2x e^{-\alpha x^2} dx$$

$$= 2i\hbar\alpha N^2 \int_{-\infty}^{\infty} e^{-2\alpha x^2} x dx = 0$$

$$\langle p^2\rangle = \int_{-\infty}^{\infty} \psi^* p_{op}^2 \psi dx = N^2 \int_{-\infty}^{\infty} e^{-\alpha x^2}\left(-\hbar^2\frac{d^2}{dx^2}\right) e^{-\alpha x^2} dx$$

$$= N^2 \int_{-\infty}^{\infty} e^{-\alpha x^2}(-\hbar^2)[-2\alpha e^{-\alpha x^2} + 4\alpha^2 x^2 e^{-\alpha x^2}] dx$$

$$\text{First term} = 2\alpha\hbar^2 N^2 \int_{-\infty}^{\infty} e^{-2\alpha x^2} dx = 2\alpha\hbar^2 N^2\cdot\sqrt{\frac{\pi}{2\alpha}} = 2\alpha\hbar^2$$

$$\text{Second term} = -4\alpha^2\hbar^2N^2\int_{-\infty}^{\infty} x^2 e^{-2\alpha x^2}\,dx = -4\alpha^2\hbar^2N^2\cdot\frac{1}{2}\sqrt{\pi}(2\alpha)^{-3/2} = -\alpha\hbar^2$$

$$\therefore\ \Delta p^2 = \alpha\hbar^2$$

$$\therefore (\Delta x)^2 = \langle x^2\rangle - \langle x\rangle^2 = \frac{1}{4\alpha}$$

$$\therefore (\Delta p)^2 = \langle p^2\rangle - \langle P\rangle^2 = \alpha\hbar^2$$

$$(\Delta x)^2(\Delta p)^2 = \frac{\hbar^2}{4}$$

$$\Delta x\,\Delta p = \hbar/2$$

Example 5.18 The state of a particle in a box of length L is described by

$$\psi = \sqrt{\frac{2}{L}}\sin\frac{n\pi x}{L} \qquad n = 1, 2, 3, \ldots$$

Calculate the uncertainty product $\Delta x\Delta p$.

Solution:

$$\begin{aligned}
\langle x\rangle &= \frac{2}{L}\int_0^L \sin\frac{n\pi x}{L}\cdot x\cdot\sin\frac{n\pi x}{L}\,dx\\
&= \frac{2}{L}\int_0^L \frac{x}{2}\left[1-\cos\frac{2n\pi x}{L}\right]dx\\
&= \frac{1}{L}\int_0^L x\,dx - \underbrace{\frac{1}{L}\int_0^L x\cos\frac{2n\pi x}{L}\,dx}_{0} = \frac{1}{L}\cdot\left.\frac{x^2}{2}\right|_0^L = \frac{L}{2}
\end{aligned}$$

$$\begin{aligned}
\langle x^2\rangle &= \frac{2}{L}\int_0^L \sin\frac{n\pi x}{L}\ x^2\sin\frac{n\pi x}{L}\,dx\\
&= \frac{2}{L}\int_0^L x^2\sin^2\frac{n\pi x}{L}\,dx = \frac{1}{L}\int_0^L x^2\left(1-\cos\frac{2n\pi x}{L}\right)dx\\
&= \frac{1}{L}\cdot\left.\frac{x^3}{3}\right|_0^L - \frac{1}{L}\int_0^L x^2\cos\frac{2n\pi x}{L}\,dx\\
&= \frac{L^2}{3} - \frac{1}{L}\left\{\underset{\Downarrow 0}{\left.\frac{L}{2n\pi}x^2\sin\frac{2n\pi x}{L}\right|_0^L} - \frac{L}{2n\pi}\int_0^L 2x\cdot\sin\frac{2n\pi x}{L}\,dx\right\}\\
&= \frac{L^2}{3} + \frac{1}{n\pi}\left\{\left[x\left(\frac{L}{2n\pi}\right)\left(-\cos\frac{2n\pi x}{L}\right)\right]_0^L + \frac{L}{2n\pi}\int_0^L\cos\frac{2n\pi x}{L}\,dx\right\}\\
&= \frac{L^2}{3} - \frac{L^2}{2n^2\pi^2}
\end{aligned}$$

$$\langle p\rangle = \frac{2}{L}\int_0^L \sin\frac{n\pi x}{L}\left(-i\hbar\frac{d}{dx}\right)\sin\frac{n\pi x}{L}\,dx = \left(-\frac{i2n\pi\hbar}{L^2}\right)\int_0^L \sin\frac{n\pi x}{L}\cos\frac{n\pi x}{L}\,dx = 0$$

$$\begin{aligned}\langle p^2\rangle &= \frac{2}{L}\int_0^L \sin\frac{n\pi x}{L}\left(-\hbar^2\frac{d^2}{dx^2}\right)\sin\frac{n\pi x}{L}\,dx \\ &= \frac{2}{L}\int_0^L \sin\frac{n\pi x}{L}(-\hbar^2)\left(-\frac{n^2\pi^2}{L^2}\sin\frac{n\pi x}{L}\right)dx \\ &= \frac{\hbar^2 n^2\pi^2}{L^3}\int_0^L 2\sin^2\frac{n\pi x}{L}\,dx = \frac{\hbar^2 n^2\pi^2}{L^3}\int_0^L\left(1-\cos\frac{2n\pi x}{L}\right)dx = \frac{\hbar^2 n^2\pi^2}{L^2}\end{aligned}$$

$$\therefore\ (\Delta x)^2 = \langle x^2\rangle - \langle x\rangle^2 = \frac{L^2}{3} - \frac{L^2}{2n^2\pi^2} - \frac{L^2}{4} = \frac{L^2}{12} - \frac{L^2}{2n^2\pi^2}$$

$$(\Delta p)^2 = \langle p^2\rangle - \langle p\rangle^2 = \frac{\hbar^2 n^2\pi^2}{L^2}$$

$$\begin{aligned}\therefore (\Delta x)^2(\Delta p)^2 &= \left[\frac{L^2}{12} - \frac{L^2}{2n^2\pi^2}\right]\frac{\hbar^2 n^2\pi^2}{L^2} \\ &= \frac{\hbar^2 n^2\pi^2}{12} - \frac{\hbar^2}{2} = \frac{\hbar^2}{4}\left(\frac{n^2\pi^2}{3} - 2\right)\end{aligned}$$

$$\therefore \Delta x\Delta p = \frac{\hbar}{2}\left(\frac{n^2\pi^2}{3} - 2\right)^{1/2}$$

Note this is always greater than $\frac{\hbar}{2}$ for all n.

5.5 COMPLETE SET OF COMPATIBLE OBSERVABLES

What is the significance of compatible observables in the context of uncertainty principle?

Let us first start with a dynamical variable A and measure it, and suppose we get the eigenvalue a. If the eigenvalue a is non-degenerate, we know that the system is in the eigenstate ϕ_a immediately after the measurement ($A\phi_a = a\phi_a$). On the other hand, if the eigenvalue a is m-fold degenerate, then the experimental result 'a' is not sufficient to determine the state of the system uniquely. Any linear combination of these m states would lead to the same experimental value a. We should have a method of determining other parameters to distinguish each of these m-fold degenerate states. In such a situation, the concept of compatible observables comes to our help.

The eigenvalue equation for A is given by

$$A_{op}\phi_a = a\phi_\alpha \quad \alpha = 1, 2, 3,\ldots m$$

Let us find a dynamic variable B such that $[A_{op}, B_{op}] = 0$. Since A and B are commuting operators, they have common eigenfunctions. Using the results of Example 15.5 we can construct *a* state ψ_{ab} such that

$$\psi_{ab} = \Sigma c_n\phi_n$$

$$A_{op}\psi_{ab} = a\psi_{ab}$$

$$B_{op}\psi_{ab} = b\psi_{ab}$$

We can construct m such linearly independent functions leading to m eigenfunctions of B with eigenvalues of b. Each of the m states is now specified by (a, b). If there are m non-degenerate eigenvalues for B_{op}, then there is one-to-one correspondence between (a, b) and the eigenstate ψ_{ab}; we have

$$\begin{array}{ccccc} (a,b_1) & (a,b_2) & (a,b_3) & \ldots & (a,b_m) \\ \psi_{ab_1} & \psi_{ab_2} & \psi_{ab_3} & \ldots & \psi_{ab_m} \end{array}$$

However, if B has also degenerate eigenvalues, the measurements of A and B are not sufficient to determine the state uniquely. We have to find another dynamical variable C such that A_{op}, B_{op} and C_{op} commute among themselves.

$$[A_{op}, B_{op}] = 0, \quad [A_{op}, C_{op}] = 0, \quad [B_{op}, C_{op}] = 0$$

Now we can construct a linear combination of $\phi_\alpha s$ such that

$$A_{op}\psi_{abc} = a\psi_{abc}$$
$$B_{op}\psi_{abc} = b\psi_{abc}$$
$$C_{op}\psi_{abc} = c\psi_{abc}$$

Now each state is specified by the set (a, b, c). We can continue this exercise by going to other operators D, E, F, M, …until we get a situation where there is a one-to-one correspondence between an eigenstate and the set of eigenvalues $(a, b, c, d, \ldots)$. That is, each set of eigenvalues $(a, b, c, d, \ldots)$ represents a non-degenerate state.

The set of all commuting dynamical variables is called complete set of compatible dynamical variables. Which plays an important role in specifying a quantum state. In fact, the study of a system begins with the identification of the complete set of compatible dynamical variables enabling us to describe the state of the system. For instance, in the case of a hydrogen atom (non-relativistic, without taking into account electron spin) H, L_2 and L_z form the complete set of commuting variables and so it is their common eigenfunctions that specify the state of the hydrogen atom.

Simultaneous Measurement and Complete Set of Compatible Dynamical Variables

Let A, B, C,… form a complete set of compatible dynamical variables for a given system. Then the operators A_{op}, B_{op}, C_{op}… commute among themselves. Let the eigenstate of the system be described by $\psi_{abc}\ldots$ such that

$$A_{op}\psi_{abcd} = a\psi_{abcd}\cdots$$
$$B_{op}\psi_{abcd} = b\psi_{abcd}\cdots$$
$$C_{op}\psi_{abcd} = c\psi_{abcd}\cdots$$

All the dynamical variables $A, B, C, \ldots$ can be measured simultaneously when the system is in this eigenstate in the following sense. First let us measure A. Since $\psi_{abc}\ldots$ is an eigenstate of A_{op} with eigenvalue a, the experimental outcome will be a, and the system will continue to be in the state $\psi_{abc}\ldots$. Immediately, let us measure B. Since ψ_{abc} is an eigenstate of B_{op} with eigenvalue b, the experimental outcome will be b. So A and B have been measured ‘simultaneously’. We can continue this procedure, and we can measure all the dynamical variables ‘simultaneously’ provided they are compatible.

Quantum Numbers and States

The eigenvalues of a dynamical variable are known as the quantum numbers. So let us concentrate on the quantum numbers corresponding to the complete set of compatible observables. The fact that $\psi_{abc}\ldots$ is a non-degenerate eigenstate of A, B, C, … implies that there is one-to-one correspondence between the set of quantum numbers (a, b, c,…) and the eigenstate $\psi_{abc}\ldots$. Therefore, specifying a state by the set of quantum numbers (a, b, c,…) is equivalent to specifying the state by explicitly giving the eigenfunction $\psi_{abc}\ldots$. However, specifying the state by the quantum numbers (a, b, c,…) has more objectivity since these numbers are experimentally measurable whereas the wave function $\psi_{abc}\ldots$ is not a measurable quantity. For instance, in the case of hydrogen atom, we have $[H, L^2] = 0$, $[H, L_z] = 0$ and $[L^2, L_z] = 0$. So for the hydrogen atom, the complete set of compatible observables are H, L^2 and L_z. The corresponding quantum numbers are n, l and m_l. So we have the following correspondence:

(nlm_l):	(2 0 0)	(2 1 0)	(2 1 1)	(21 − 1)
	ψ_{200}	ψ_{210}	ψ_{211}	ψ_{21-1}

The explicit forms of the functions ψ_{200}, ψ_{210}, ψ_{211} and ψ_{21-1} are given in Chapter 7.

Simultaneous Measurement of Position and Momentum

Since $[x, p_x] \neq 0$, they are not compatible set of dynamical variables, so we cannot find eigenstate common to both x and p_x. Therefore, x and p_x cannot be measured simultaneously. Depending on the way the state is prepared, we will get different kinds of statistical distribution (or the uncertainties).

Let us recall that (i) there is no uncertainty in A if the system is in the eigenstate of A and (ii) there is uncertainty in A if the state of the system is a superposition of eigenstates of A.

If the prepared state is a position eigenstate, then it will be a superposition of momentum eigenstates. There will be no uncertainty in position, but momentum is completely uncertain. If the prepared state is momentum eigenstate, then it will be superposition of position eigenstates. There will be no uncertainty in momentum, but the position is completely uncertain. If the prepared state is the eigenstate of neither position nor momentum, then it can be expressed as a superposition of position eigenstates or as a superposition of momentum eigenstates. So there will be uncertainty both in position and momentum.

5.6 CONSTANTS OF MOTION

In classical physics, A is a constant of motion if $\frac{dA}{dt} = 0$. This condition reduces to

$$\frac{dA}{dt} = [A, H]_{PB} + \frac{\partial A}{\partial t} = 0$$

where $[A, H]_{PB}$ is the Poisson bracket.

In quantum mechanics, A is said to be a constant of motion if

$$\frac{d}{dt}\langle A\rangle = 0 \tag{5.33}$$

A_{op} may or may not have explicit time dependence. If A_{op} has explicit time dependence, then, $\partial A_{op}/\partial t \neq 0$. Even if A_{op} has no explicit time dependence, in the expectation value $\langle A\rangle$, the time dependence can enter through the wave function $\psi(\mathbf{r}, t)$.

$$\begin{aligned}\therefore \frac{d}{dt}\langle A\rangle &= \frac{d}{dt}\int \psi^*(\mathbf{r},t)A_{op}\psi(\mathbf{r},t)d\tau \\ &= \int \frac{d\psi^*}{dt}A_{op}\psi d\tau + \int \psi^* A_{op}\frac{d\psi}{dt} \\ &\quad + \int \psi^* \frac{\partial A_{op}}{\partial t}\psi d\tau\end{aligned}$$

From Schrödinger equation,

$$\frac{d\psi}{dt} = \frac{1}{i\hbar}H_{op}\psi \quad \text{and} \quad \frac{d\psi^*}{dt} = -\frac{1}{i\hbar}(H_{op}\psi)^*$$

$$\begin{aligned}\frac{d}{dt}\langle A\rangle &= \int -\frac{1}{i\hbar}(H_{op}\psi)^* A_{op}\psi d\tau + \int \psi^* A_{op}\frac{1}{i\hbar}H_{op}\psi d\tau \\ &\quad + \int \psi^* \frac{\partial A_{op}}{\partial t}\psi d\tau \\ &= \int -\frac{1}{i\hbar}\psi^* H_{op}A_{op}\psi d\tau + \int \psi^* A_{op}\frac{1}{i\hbar}H_{op}\psi d\tau \\ &\quad + \int \psi^* \frac{\partial A_{op}}{\partial t}\psi d\tau \\ &= \int \psi^* \left\{\frac{1}{i\hbar}(A_{op}H_{op} - H_{op}A_{op}) + \frac{\partial A_{op}}{\partial t}\right\}\psi d\tau \\ &= \left\langle \frac{1}{i\hbar}[A_{op}, H_{op}] + \frac{\partial A_{op}}{\partial t}\right\rangle \qquad (5.34)\end{aligned}$$

For A to be a constant of motion,

$$\frac{1}{i\hbar}[A_{op}, H_{op}] + \frac{\partial A_{op}}{\partial t} = 0 \qquad (5.35)$$

In general, most of the basic operators in quantum mechanics have no explicit time dependence. In such cases, the condition for A to be constant of motion is

$$[A_{op}, H_{op}] = 0 \qquad (5.36)$$

For instance, the dynamical variables momentum $\mathbf{p}_{op}$ and angular momentum operator $\mathbf{L}_{op}$ do not have explicit time dependence. Therefore, we have

$$\begin{aligned}&[\mathbf{p}_{op}, H_{op}] = 0 \Rightarrow \quad \text{Conservation of momentum} \\ &[\mathbf{L}_{op}, H_{op}] = 0 \Rightarrow \quad \text{Conservation of angular momentum} \qquad (5.37)\end{aligned}$$

Example 5.19 Show that the angular momentum is conserved for a particle in central potential.

Solution: The Hamiltonian for a particle in central potential is

$$H = \frac{p^2}{2m} + V(r)$$

Consider the z component of angular momentum, L_z, is given by

$$L_z = xp_y - yp_x$$

(For the sake of clarity, the subscript *op* is omitted.)

$$[L_z, H] = \left[L_z, \frac{p^2}{2m} + V(r)\right] = \left[L_z, \frac{p^2}{2m}\right] + [L_z, V(r)]$$

To evaluate the commutation relations, we use $[A, BC] = B[A, C] + [A, B]C$ and $[AB, C] = A[B, C] + [A, C]B$

$$\left[L_z, \frac{p^2}{2m}\right] = \frac{1}{2m}[L_z, p_x^2 + p_y^2 + p_z^2] = \frac{1}{2m}\{[L_z, p_x^2] + [L_z, p_y^2] + [L_z, p_z^2]\}$$

$$[L_z, p_x^2] = p_x[L_z, p_x] + [L_z, p_x]p_x$$

$$\begin{aligned}[L_z, p_x] &= [xp_y - yp_x, p_x] = [xp_y, p_x] - [yp_x, p_x]\\ &= x[p_y, p_x] + [x, p_x]p_y - [y, p_x]p_x - y[p_x, p_x]\\ &= 0 + i\hbar p_y - 0 - 0 = -i\hbar p_y\end{aligned}$$

$$\therefore [L_z, p_x^2] = 2i\hbar p_x p_y$$

Similarly, we can prove $[L_z, p_y^2] = -2i\hbar p_x p_y$ and $[L_z, p_z^2] = 0$

$$\therefore \left[L_z, \frac{p^2}{2m}\right] = \frac{1}{2m}[2i\hbar p_x p_y - 2i\hbar p_x p_y] = 0$$

Let us now evaluate $[L_z, V(r)]$

$$\begin{aligned}[L_z, V(r)] &= [xp_y - yp_x, V(r)]\\ &= [xp_y, V(r)] - [yp_x, V(r)]\\ &= \underset{\Downarrow\, 0}{[x, V(r)]}p_y + x[p_y, V(r)] - y[p_x, V(r)] - \underset{\Downarrow\, 0}{[y, V(r)]}p_x\\ &= x\left[-i\hbar\frac{\partial V}{\partial y}\right] - y\left(-i\hbar\frac{\partial V}{\partial x}\right)\\ &= -i\hbar\left[x\frac{\partial V}{\partial y} - y\frac{\partial V}{\partial x}\right]\\ &= -i\hbar\left[x\frac{\partial V}{\partial r}\cdot\frac{\partial r}{\partial y} - y\frac{\partial V}{\partial r}\cdot\frac{\partial r}{\partial x}\right]\\ &= -i\hbar\left[x\cdot\frac{y}{r}\cdot\frac{\partial V}{\partial r} - y\cdot\frac{x}{r}\cdot\frac{\partial V}{\partial r}\right] = 0\end{aligned}$$

$$[L_z, H] = 0$$

We can extend the same result for L_x and L_y. i.e., $[L_x, H] = 0$ and $[L_y, H] = 0$.

Therefore, L_x, L_y and L_z are constants of motion; i.e., angular momentum is a conserved quantity for a central potential.

5.7 EHRENFEST THEOREM

Ehrenfest's theorem states that the mean values of a dynamical variable evolve according to classical equation of motion. This statement requires further qualifications, which will be discussed later. We closely follow Gaziorowicz here.

The classical equations of motion are

$$\frac{d\mathbf{r}}{dt} = \frac{\mathbf{p}}{m}$$

and

$$\frac{d\mathbf{p}}{dt} = F(\mathbf{r}) = -\nabla V(\mathbf{r}).$$

In one dimension, these equations become

$$\frac{dx}{dt} = \frac{p}{m} \quad \text{and} \quad \frac{dp}{dt} = F(x) = -\frac{\partial V}{dx}$$

The Ehrenfest theorem states that in quantum mechanics,

$$\frac{d}{dt}\langle x\rangle = \frac{\langle p\rangle}{m}$$

and

$$\frac{d}{dt}\langle p\rangle = -\left\langle \frac{\partial V}{\partial x}\right\rangle$$

These two results can be easily proved.

$$\frac{d}{dt}\langle x\rangle = \frac{1}{i\hbar}\langle [x,H]\rangle$$

$$[x,H] = \left[x, \frac{p^2}{2m} + V(x)\right] = \left[x, \frac{p^2}{2m}\right] + \underset{\Downarrow 0}{[x,V(x)]}$$

$$= \left[x, \frac{p^2}{2m}\right] = \frac{1}{2m}[x,p^2] = \frac{1}{2m}\{p[x,p] + [x,p]p\}$$

$$= \frac{1}{2m}\cdot 2i\hbar p = i\hbar\frac{p}{m}$$

$$\therefore \frac{d}{dt}\langle x\rangle = \frac{1}{i\hbar}\left\langle i\hbar\frac{p}{m}\right\rangle = \frac{\langle p\rangle}{m} \tag{5.38}$$

Now consider $\frac{d}{dt}\langle p\rangle$

$$\frac{d}{dt}\langle p\rangle = \frac{1}{i\hbar}\langle [p,H]\rangle$$

$$[p,H] = \underset{\Downarrow 0}{\left[p, \frac{p^2}{2m}\right]} + [p,V(x)] = -i\hbar\frac{\partial V}{\partial x}$$

$$\therefore \frac{d}{dt}\langle p\rangle = \frac{1}{i\hbar}\left\langle -i\hbar\frac{\partial V}{\partial x}\right\rangle = \left\langle -\frac{\partial V}{\partial x}\right\rangle \tag{5.39}$$

The Equations (5.38) and (5.39) are almost similar to Newton's equation of motion, and they compel us to conclude the position in classical physics as the mean value of position in the corresponding quantum version of the system. i.e., we would like to identify $x_{cl} = \langle x\rangle$. If we make such an identification, Newton's equation of motion should be

$$\frac{d}{dt}\langle p\rangle = -\frac{\partial V(x_{cl})}{\partial x_{cl}} = -\frac{\partial V(\langle x\rangle)}{\partial\langle x\rangle} \tag{5.40}$$

However, this equation is different from the Equation (5.39). Let us determine the conditions under which we can write

$$-\frac{\partial V(\langle x\rangle)}{\partial\langle x\rangle} = \left\langle -\frac{\partial V}{\partial x}\right\rangle \quad \text{or} \quad F(\langle x\rangle) = \langle F(x)\rangle$$

Let us expand $F(x)$ as

$$F(x) = F(\langle x\rangle) + (x-\langle x\rangle)F'(\langle x\rangle) + \frac{(x-\langle x\rangle)^2}{2!}F''(\langle x\rangle) + \cdots$$

Taking the expectation value, this equation becomes

$$\langle F(x)\rangle = \langle F(\langle x\rangle)\rangle + \langle (x-\langle x\rangle)F'(\langle x\rangle)\rangle + \frac{\langle (x-\langle x\rangle)^2 F''(\langle x\rangle)\rangle}{2!} + \cdots$$

Note
$$\langle F(\langle x\rangle)\rangle = \int \psi^*(x)F(\langle x\rangle)\psi(x)dx$$
$$= F(\langle x\rangle)\int \psi^*(x)\psi(x)dx = F(\langle x\rangle),$$

since $F(\langle x\rangle)$ is a number and not a function of x.

$$\begin{aligned}\langle (x-\langle x\rangle)F'(\langle x\rangle)\rangle &= \int \psi^*(x-\langle x\rangle)F'(\langle x\rangle)\psi dx\\ &= F'(\langle x\rangle)\int \psi^*(x-\langle x\rangle)\psi dx\\ &= F'(\langle x\rangle)\left[\int \psi^* x\psi dx - \langle x\rangle\int \psi^*\psi dx\right]\\ &= F'(\langle x\rangle)[\langle x\rangle - \langle x\rangle] = 0\end{aligned}$$

Similarly,

$$\begin{aligned}\langle (x-\langle x\rangle)^2 F''(\langle x\rangle)\rangle &= F''(\langle x\rangle)\int \psi^*(x-\langle x\rangle)^2\psi dx\\ &= F''(\langle x\rangle)(\Delta x)^2\end{aligned}$$
$$\therefore \langle F(x)\rangle = F(\langle x\rangle) + \frac{(\Delta x)^2}{2}F''(\langle x\rangle) + \cdots$$

If the uncertainty (Δx) is small, we can write

$$\langle F(x)\rangle = F(\langle x\rangle)$$

Under such conditions, we can identify $x_{cl} = \langle x\rangle$ and

$$\frac{d}{dt}\langle p\rangle = F(\langle x\rangle) = -\frac{\partial V(\langle x\rangle)}{\partial\langle x\rangle} \tag{5.41}$$

This is an enormously useful result. There are a number of circumstances where we can use classical equation of motion to study the motion of subatomic particles like electron, proton, etc. This is possible if the uncertainty in x is very small.

5.8 CLASSICAL PHYSICS AND QUANTUM PHYSICS

The key idea in postulates P-1 and P-6 is the concept of probability. The term *probability* implies that the experimental outcome of the measurement made on a single system is not unique or well defined. We may get any one of the possible value. There is no definiteness about the result.

In what follows, we use the relative frequencies of various outcomes to assign the number for the value for probability.

Do We Get Probability in Classical Physics?

The answer is yes. We do employ the concept of probability in classical physics also. For instance, toss of a coin is an random event. The weather is another example where we describe everything in terms of probability.

Toss of a Coin

In the toss of a coin, the possible results or outcomes are either head or tail. We will not be able to tell what exactly will be the outcome in a single trial. When we make a large number of trials using the same coin or when we toss a large number of identical coins, we will get head for nearly half the number of trials and tail for nearly half the number of trials. Here, we have assumed the coins to be fair coins. If the coin is a biased coin, the probability of getting the head in a single toss will not be 0.5, but some number between 0 to 1, say 0.3. In this case also, there is no way of predicting the result for a single trial. But if we consider a larger number of identical coins (for all which the probability of getting a head is 0.3 and tail is 0.7), we will get head nearly in 30% of the trials.

The important thing to recognize is that the moment we introduce the probability in our discussion, there is no way of predicting the results or outcome for a single system. Any possible outcome has to be accepted.

How do We Get Probability in a Deterministic Classical Physics?

We know a tossed coin obeys laws of classical physics. If so, it should obey Newton's equation of motion, and we should be able to predict the result of a single trial with 100 per cent assurance. Strictly speaking, there is no necessity to introduce the concept of probability in classical mechanics. If we know the forces acting on the coin and the initial conditions, we can make an assured prediction even in the case of a single trial. Still why do we introduce the concept of randomness or unpredictability in the discussions of the outcomes of a toss? The catch lies in the fact that if we know fully the laws obeyed by the coin and the initial conditions, we can make exact prediction. The randomness in the case of coin is due to the fact that we do not have all the details about the coin. It may be the lack of knowledge about the initial conditions - with what velocity the coin is flipped? – or the lack of knowledge of its interactions with the surrounding or the forces acting on it.. The same is true with the prediction of weather.

So in classical physics, randomness or unpredictability is introduced because full knowledge about the system is not available.

Probability in Quantum Mechanics

First, let us try to understand the implication of postulate P-1. For the sake of illustration, let us consider an ensemble of hydrogen atoms. That is, all the systems in the ensemble are identical. Let us further assume that all the hydrogen atoms are in the same state, say, the ground state, corresponding to quantum numbers $n = 1$, $l = 0$ and $m_l = 0$.

So we are considering a collection of identical system under identical conditions. Let us measure the distance of the electron from the nucleus for each of the hydrogen atom. Since we have a collection of identical systems under identical condition, we expect the outcome for all the hydrogen atoms to be identical. But postulate P-1 implies that we get a statistical distribution of values for the distance of the electrons from the nucleus.

An imaginary experimental outcome for the measurement of distance of electrons from the nucleons will look as follows:

Table 5.2 An imaginary experimental outcome for the measurement of distance of electrons from the nucleons

Distance in Å	0.1	0.3	0.5	0.8	1.0	1.2	1.5	2.0	2.2
Number of hydrogen atoms all in the state $n = 1$ $l = 0$ $m_l = 0$	1120	1290	1350	1300	900	820	730	590	80

The significant point is that we are not getting the same value for all the hydrogen atoms in spite of the fact that they are in the same state; i.e., we get a statistical distribution of values in spite of the fact that the experiment is performed for identical systems under identical conditions. (We have reached the same conclusion as a result of Postulate P-6) This is in total contrast with classical mechanics.

So we have the following

Classical systems: Identical systems under identical conditions ⇒ identical outcomes

Quantum systems: Identical systems under identical conditions ⇒ outcomes are not identical

Maximum Information About the System

What is the source of this randomness? In classical physics, the randomness is due to lack of knowledge about the system and the errors in the instruments. In quantum mechanics also, errors in the instruments and the observation process lead to randomness. But there is an additional source of the randomness which is not due to any fault in the instrument or some unknown factor. It is an intrinsic property of a quantum system at microscopic level.

In fact, it should be noted that within the structure of quantum mechanics, maximum information about a system is contained in the wave functions. We cannot get any information more than that is given by $\psi(\mathbf{r}, t)$. This implies that the probabilities associated with a quantum system are intrinsic property of that system within the structure of quantum mechanics. Lack of knowledge or errors in the instruments will worsen the situation and add further uncertainties or randomness.

Loss of Concept of Trajectory

One of the consequences of probabilistic interpretation is the loss of the concept of trajectory. The trajectory or a particle is the lifeline of classical physics. The ultimate purpose of solving Newton's equation of motion is to obtain the trajectory $\mathbf{r}(t)$ of the particle. The position $\mathbf{r}(t)$ of a particle at different instants gives the trajectory of a particle. So the determination of the trajectory of the particle is the determination of position **r** at different instants of time.

In quantum mechanics, all that we can tell is the probability of finding a particle at **r** at the instant *t*. Since the position of a particle at any instant is unpredictable or random, there is no possibility of defining a trajectory of a particle in quantum mechanics.

Therefore, the Bohr's model (the electrons moving in a circular orbit) or the Sommerfeld model (electrons moving in an elliptical orbit) are no longer valid. Once the concept of the trajectory is lost, the velocity and the momentum of a particle can no longer be defined as $d\mathbf{r}(t)/dt$ and $m\, d\mathbf{r}/dt$. Yet in quantum mechanics, a particle has a property called momentum. It is not the same as the classical concept $m\, d\mathbf{r}/dt$. However, it is a property that can be measured with a suitable experimental arrangement. The velocity of a particle is given by the same expression $\mathbf{p}/m$ both in classical and in quantum physics.

Do We Employ the Concept of Trajectory for Subatomic Particles?

There are a number of circumstances where electrons are treated like a classical particle which has definite trajectory. Description of electrons or a proton or an antiproton in accelerators is one such example. We accept the fact that photons are travelling from the sun to the earth. In the case of cathode ray oscilloscope, the idea of trajectory is used to explain the behavior of the beam.

In all these cases, the Ehrenfest theorem comes to our help. If the uncertainty is very small, the average value evolves in time like a classical system. In such cases, the classical equations are sufficient.

In terms of the wave function, the concept which comes closest to trajectory of classical particle is the motion of wave packet. The wave function describing the state of an electron is the wave packet travelling as shown in Fig. 5.3.

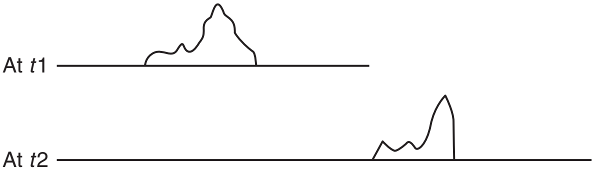


Fig. 5.3 Wave packets at different instants of time

Even here, the electron can be anywhere inside the region enclosed by the wave envelope. We can approximately take the velocity of the wave packet as the velocity of the particle. This resembles the classical concept of the velocity of a particle.

Uncertainty in Classical Physics

Do we employ the concept of uncertainty in classical physics? The answer is yes. Within the theories explaining the physical phenomena, there is no place for uncertainty in classical physics. However, in experimental physics, we employ the concept of uncertainties.

For instance, we determine the value of acceleration due to gravity in an experiment. The final result will be expressed in the format $g = 9.81 \pm 0.02$ m/sec^2 (Don't worry about the exact numerical value given here). If any dynamical variable Q is measured, the final result will be expressed in the format $q \pm \delta q$, where δq is the uncertainty. What is the source of this uncertainty in classical physics? A full theory of error analysis is devoted to estimate δq in experimental physics. Leaving aside many details, we can say δq as the standard deviation obtained from the statistical distribution of values obtained from observation.

Let q_1, q_2,…q_n be the observed values of Q in n trials. Then the average value and the standard deviation are given by

$$\bar{q} = (q_1 + q_2 + \cdots q_n)/n$$

$$\delta q = [(q_1 - \bar{q})^2 + (q_2 - \bar{q})^2 + \cdots]^{1/2}/n \tag{5.42}$$

The experimental value is quoted as

$$q = \bar{q} \pm \delta q$$

The question is why we get a statistical distribution of q_1, q_2,… in the experiment, though the theory suggests a definite value. When a measurement is made, there is a large number of sources of error: the instruments may be defective, may not have the required least count, may not have been calibrated properly, etc. All these factors contribute to the fact that we are not getting the exact unique value for Q. Can we reduce this uncertainty δq? With the advent of new technology, can we make the measurement of Q more accurate so that δq becomes as small as possible? Can we make δq to be zero? This is a hypothetical question. Within classical physics, there is no restriction on the possible range of δq, and hence in principle, we make $\delta q = 0$. Note that this is also a hypothetical answer.

Are the Uncertainties in Classical and Quantum Physics the Same?

The uncertainty Δq in quantum mechanics does not refer to the δq in the experimental physics. We have already seen that in any experimental measurement of a quantum system, we can only tell the probability of obtaining a particular eigenvalue. The statistical distribution of experimental outcomes is intrinsic to quantum mechanics.

Going back to the discussion on the imaginary experiment of determining the distance of the electron from the nucleus of a hydrogen atoms, the table of experimental outcomes is a statistical distribution, and we apply the expression (5.42).

The existence of average and standard deviation has nothing to do with the distinction between classical and quantum physics. As long as a statistical distribution is there, we can define average and standard deviation. However, the reasons for the existence of the statistical distribution of experimental outcomes are different in classical and quantum physics.

EXERCISES

1. What is the difference between probability and probability density?
2. Evaluate $[\phi(x, y, z), \mathbf{p}_{op}]$.
3. What is the significance of two commuting operators representing dynamical variables?

4. What is meant by a complete set of compatible observables? Why do we need them?
5. A dynamical variable A has two eigenvalue a and $-a$; i.e., $A\phi_1 = a\phi_1$ and $A\phi_2 = -a\phi_2$. Consider a system whose wave function is $\phi = 0.6\phi_1 + 0.8\phi_2$. When a measurement of A is made, what is the probability A to be (i) a (ii) $-a$ (iii) what is the average value of A and (iv) what is the uncertainty in A?
6. If A is dynamical variable such that $A^2 = I$, evaluate (i) $e^{\lambda A}$ and (ii) $e^{i\lambda A}$.
7. If ϕ_n is an eigenstate of the Hamiltonian H with energy eigenvalue E_n, then show that $\exp(iHt/\hbar)\phi_n = \exp(iE_n t/\hbar)\phi_n$.
8. Determine $\psi(x, t)$ when $\psi(x, 0) = [\Sigma c_n \phi_n(x)]$ and when $H\phi_n = E_n\phi_n$.
9. Show that momentum is conserved for a free particle with $H = p^2/2m$.
10. Show that angular momentum is conserved for a free particle with $H = \frac{p^2}{2m}$.
11. Show that $e^{\lambda A} B e^{\lambda A} = B + \lambda[A, B] + \frac{\lambda^2}{2!}[A,[A, B]] + \frac{\lambda^3}{3!}[A,[A, B]] + \cdots$ (take $f(\lambda) = e^{\lambda A} B e^{\lambda A}$ and use $f(\lambda) = f(0) + \lambda f'(0) + \frac{\lambda^2}{2!} f''(0) + \cdots$).
12. Prove that $\exp\left(ia\frac{p}{\hbar}\right) f(x) = f(x + a)$, where $p = -i\hbar\frac{d}{dx}$.
13. Obtain a general expression for $\frac{d}{dt}\langle x^2 \rangle$ in terms of $[x, H]$.

REFERENCES

1. R. H. Dicke and J. F. Wittki, 1960. *Introduction to Quantum Mechanics*. Reading Massachusetts: Addison-Wesley Publishing Company, Inc.
2. P. M. Mathews and K. Venkatesan, 1976. *A Textbook of Quantum Mechanics.* New Delhi: Tata McGraw-Hill Company.
3. S. Gasiorowicz, 1974. *Quantum Physics.* New York: John Wiley & Sons, Inc.
4. J. L. Powell and B. Crasemann, 1961. *Quantum Mechanics*. Reading Massachusetts: Addison-Wesley Publishing Company, Inc..
5. L. E. Ballentine, 1998. *Quantum Mechanics – A Modern Development.*, Singapore: World Scientific Co., Pvt, Ltd.

6

The Simple Harmonic Oscillator

The simple harmonic oscillator is one of the most important topics in quantum mechanics. The energy eigenvalues of a harmonic oscillator is $E_n = (n+1/2)\hbar\omega$. The energy levels are evenly spaced, and this result has enormous significance. A number of physical systems, in the first–order approximation, can be reduced to a collection of simple harmonic oscillators representing fluctuations in various physical quantities. For instance, in quantum field theory, fields become a collection of oscillators leading to the 'emergence' of 'particle property' of all quantum objects. In this chapter, we first study the harmonic oscillator in wave mechanics. Then we study the harmonic oscillators in terms of abstract operators called ladder operators. It is this formulation of simple harmonic oscillator which makes harmonic oscillators as the most widely used object in diverse areas like lattice vibration, electromagnetic fields, fluctuations in magnetic dipole moments in solids and so on.

6.1 WAVE MECHANICS

6.1.1 Schrödinger Equation and Harmonic Oscillator

The classical Hamiltonian for a harmonic oscillator is

$$H = \frac{p^2}{2m} + \frac{1}{2}m\omega^2 x^2 \tag{6.1}$$

To make a transition to quantum mechanics, we replace the dynamical variables by their operators. In the coordinate representation, the position x becomes $x_{op} = x$, the momentum p becomes the operator $p_{op} = -i\hbar\, d/dx$ and the Hamiltonian H becomes H_{op} given by

$$\begin{aligned} H_{op} &= \frac{p_{op}^2}{2m} + \frac{1}{2}m\omega^2 x_{op}^2 \\ &= -\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + \frac{1}{2}m\omega^2 x^2 \end{aligned}$$

The time-independent Schrödinger equation is given by

$$-\frac{\hbar^2}{2m}\frac{d^2 u}{dx^2} + \frac{1}{2}m\omega^2 x^2 u = Eu$$

This equation can be rewritten as

$$\frac{d^2u}{dx^2}+\frac{2m}{\hbar^2}\left(E-\frac{1}{2}m\omega^2x^2\right)u=0 \tag{6.2}$$

We are interested in determining the energy eigenvalues of the bound state solution of this equation. This implies that the above differential equation has to be solved subject to the boundary condition that $u(x)$ being single valued and finite and in the large x limit, its behaviour is given by

$$u\,(x)\rightarrow 0 \text{ as } x\rightarrow \pm\infty \tag{6.3}$$

The fact that $V=1/2\,m\omega^2x^2$ is symmetric about the origin implies that the energy eigenstate for the harmonic oscillator has definite parity. Therefore, the energy eigenfunctions have to be either even-parity or odd–parity functions.

To solve this equation, let us go to a new variable ρ instead of x. The new variable ρ is defined by

$$\rho=\alpha x$$

where α is an arbitrary constant that can be chosen in such a way that that differential equation in terms of ρ becomes simpler.

First, let us write the Equation (6.2) in terms of ρ. We have

$$\therefore \alpha^2\frac{d^2u}{d\rho^2}+\frac{2mE}{\hbar^2}u-\frac{m^2\omega^2}{\hbar^2}\cdot\frac{\rho^2}{\alpha^2}u=0$$

Dividing throughout by α^2, we get

$$\frac{d^2u}{d\rho^2}+\frac{2mE}{\hbar^2\alpha^2}u-\frac{m^2\omega^2}{\hbar^2}\cdot\frac{1}{\alpha^4}\cdot\rho^2u=0$$

Now let us choose $$\frac{m^2\omega^2}{\hbar^2\alpha^4}=1 \quad \text{or} \quad \alpha=\left(\frac{m\omega}{\hbar}\right)^{1/2} \tag{6.4}$$

Let us define λ as

$$\lambda=\frac{2mE}{\hbar^2\alpha^2}=\frac{2E}{\hbar\omega} \tag{6.5}$$

Therefore, the Schrödinger equation becomes

$$\frac{d^2u}{d\rho^2}+(\lambda-\rho^2)u=0 \tag{6.6}$$

Let us try the solution $u(\rho)=e^{-\rho^2/2}$. L.H.S. of the Equation (6.6) $=-e^{-\rho^2/2}+\rho^2e^{-\rho^2/2}+\lambda e^{-\rho^2/2}-\rho^2e^{-\rho^2/2}$. This can become zero provided $\lambda=1$.

Therefore, $e^{-\rho^2/2}$ is a solution to (6.6) corresponding to a special case of λ. This suggests that we can write the solution to equation (for all values of λ) as

$$u(\rho)=e^{-\frac{\rho^2}{2}}h(\rho) \tag{6.7}$$

This equation leads to

$$\frac{d^2u}{d\rho^2}=e^{-\frac{\rho^2}{2}}\left[\frac{d^2h(\rho)}{d\rho^2}-2\rho\frac{dh(\rho)}{d\rho}+(\rho^2-1)h(\rho)\right] \tag{6.8}$$

Substituting this equation in (6.6) and factoring out $e^{-\rho^2/2}$, we get

$$\frac{d^2h(\rho)}{d\rho^2}-2\rho\frac{dh(\rho)}{d\rho}+(\lambda-1)h(\rho)=0. \tag{6.9}$$

This equation becomes Hermite's differential equation, had it been $2n\, h(\rho)$ instead of $(\lambda-1)\, h(\rho)$ in the last term. It will be shown that $(\lambda-1)$ indeed is $2n$, as a consequence of the boundary condition.

One of the standard methods of solving a differential equation of this type is the series solution method. The solution $h(\rho)$ can be written as

$$h(\rho)=\sum_{s=0}^{\infty}C_s\rho^S \tag{6.10}$$

The determination of the constants C_s amounts to determining the function $h(\rho)$.

One of the boundary conditions for the harmonic oscillator wave functions is either $u(-x) = u(x)$ (even parity) or $u(-x) = -u(x)$ (odd parity). Implementing this boundary condition, we get two kinds of series solutions.

$$\text{Even parity}: u(\rho)=e^{-\frac{\rho^2}{2}}[C_0+C_2\rho^2+C_4\rho^4+\ldots]$$

$$\text{Odd parity}: u(\rho)=e^{-\frac{\rho^2}{2}}[C_1\rho+C_3\rho^3+C_5\rho^5+\ldots]$$

Let us determine the coefficients C_s.

Differentiating the expression (6.10), we get

$$\frac{dh}{d\rho}=\sum_{s=0}^{\infty}C_s s\rho^{s-1}\quad\text{and}\quad\frac{d^2h}{d\rho^2}=\sum_{s=0}^{\infty}C_s s(s-1)\rho^{s-2}=\sum_{s=2}^{\infty}C_s s(s-1)\rho^{s-2}$$

Substituting these expressions in (6.10), we get

$$\sum_{s=2}^{\infty}C_s s(s-1)\rho^{s-2}-\sum_{s=0}^{\infty}[2s-(\lambda-1)]\;C_s\rho^s=0 \tag{6.11}$$

In the first term, put $r = s - 2 \;\therefore\; s = r + 2$ and r varies from 0 to ∞.

$$\sum_{s=2}^{\infty}C_s s(s-1)\rho^{s-2}=\sum_{r=0}^{\infty}C_{r+2}(r+2)(r+1)\rho^r$$

Since r is a summation index, we can replace it by s in the RHS.

$$\sum_{r=0}^{\infty}C_{r+2}(r+2)(r+1)\rho^r=\sum_{s=0}^{\infty}C_{s+2}(s+2)(s+1)\rho^s$$

Substituting this expression in (6.11), we get

$$\sum_{s=0}^{\infty}[C_{s+2}(s+2)(s+1)-C_s(2s-(\lambda-1))]\rho^s=0$$

Equating the coefficient of each power to zero, we get

$$\begin{aligned}C_{s+2}(s+2)(s+1)&=[2s-(\lambda-1)]C_s\\ \text{or}\;\; C_{s+2}&=\frac{2s-(\lambda-1)}{(s+2)(s+1)}C_s\end{aligned} \tag{6.12}$$

Alternatively, we can rewrite the second term in (6.11) as follows:

$$\sum_{s=0}^{\infty}[2s-(\lambda-1)]C_s\rho^s = \sum_{r=2}^{\infty}[2(r-2)-(\lambda-1)]C_{r-2}\,\rho^{r-2}$$

$$= \sum_{s=2}^{\infty}[2(s-2)-(\lambda-1)]C_{s-2}\,\rho^{s-2}$$

Then Equation (6.11) becomes $\sum_{s=2}^{\infty}[C_s s(s-1)-(2(s-2)-(\lambda-1))C_{s-2}]\rho^{s-2}=0$

Equating the coefficient of each power separately to zero, we get

$$C_{s-2} = \frac{s(s-1)}{2(s-2)-(\lambda-1)}C_s \tag{6.13}$$

Both Equations (6.12) and (6.13) are the recursion relations, and they are equivalent to each other. Equation (6.12) helps us to calculate C_{s+2} from C_s and the Equation (6.13) helps us to calculate C_{s-2} from C_s. Let us use the recursion relation (6.12).

So we have

$$\text{even parity}: C_{s+2} = \frac{2s-(\lambda-1)}{(s+2)(s+1)}C_s \qquad s=0,2,4,\ldots$$

$$\text{odd parity}: C_{s+2} = \frac{2s-(\lambda-1)}{(s+2)(s+1)}C_s \qquad s=1,3,5,\ldots$$

The boundary condition for the harmonic oscillator wave function is that as $x \to \pm\infty$, $u(x) \to 0$. So we have to determine how $e^{-\rho^2/2}h(\rho)$ behaves as $\rho \to \pm\infty$. There is a standard method of studying the behaviour of an infinite series for large ρ. One has to first determine $\underset{s\to\infty}{\text{Lt}}\, C_{s+2}/C_s$ for the given series. Then, compare this result with the limit $\underset{s\to\infty}{\text{Lt}}\, C_{s+2}/C_s$ for a well–known series. From this knowledge, we can determine the asymptotic behaviour of a series. So let us consider the limit $\underset{s\to\infty}{\text{Lt}}\, C_{s+2}/C_s$ for the the above series.

From the Equation (6.12), we have

$$\underset{s\to\infty}{\text{Lt}}\frac{C_{s+2}}{C_s} \to \frac{2s}{s^2} = \frac{2}{s} \tag{6.14}$$

Let us compare this with e^{ρ^2}, which can be written as

$$\exp(\rho^2) = 1+\rho^2+\frac{\rho^4}{2!}+\ldots\ldots = \sum C_s\rho^s$$

provided $$C_s = \frac{1}{\left(\frac{s}{2}\right)!} \qquad s=0,2,4,\ldots$$

$$\text{Lt}\frac{C_{s+2}}{C_s} = \frac{\left(\frac{s}{2}!\right)}{\left(\frac{(s+2)}{2}!\right)} = \frac{\frac{s}{2}!}{\left(\frac{s}{2}+1\right)!} = \frac{1}{\left(\frac{s}{2}+1\right)}$$

$$\therefore \qquad \underset{s\to\infty}{\text{Lt}}\frac{C_{s+2}}{C_s} \to \frac{2}{s} \tag{6.15}$$

Comparing the results (6.14) and (6.15), we conclude that the series $h(\rho)$ behaves like e^{ρ^2} in the large ρ limit.

i.e.,
$$u(\rho) = e^{-\frac{\rho^2}{2}} h(\rho) \xrightarrow[\rho\to\infty]{} e^{\frac{\rho^2}{2}}$$

Obviously, this is not an acceptable solution in quantum mechanics. We would like to have a solution $u(\rho) \to 0$ as $\rho \to \infty$. The solution $u(\rho) = e^{-\rho^2/2} h(\rho)$ will go to zero for large ρ, if $h(\rho)$ is a polynomial in ρ (a series with finite number of terms) instead of an infinite series. So we have to terminate the series suitably so that the infinite series becomes a polynomial of degree n, where n is finite.

This requirement can be implemented for both even parity and odd parity solutions by suitably choosing $(\lambda - 1)$ and the coefficients C_n as follows:

Even parity: Choose $(\lambda - 1)$ as $(\lambda - 1) = 2n$ with $C_n \neq 0$ and $C_1 = 0$ for even n.
Odd parity: Choose $(\lambda - 1)$ as $(\lambda - 1) = 2n$ with $C_n \neq 0$ and $C_0 = 0$ for odd n.

From (6.12) for $s = n$, we have

$$C_{n+2} = \frac{2n-(\lambda-1)}{(n+2)(n+1)} = \frac{2n-2n}{(n+2)(n+1)} C_n = 0$$

This will make $C_{n+2} = C_{n+4} = C_{n+6} = \cdots = 0$.

For $\lambda = 2n - 1$, $h(\rho)$ will become

$$h(\rho) = C_0 + C_2\rho^2 + C_4\rho^4 + \ldots \qquad \text{even } n$$

and
$$h(\rho) = C_1\rho + C_3\rho^3 + C_5\rho^5 + \ldots \qquad \text{odd } n.$$

The choice of choosing $\lambda - 1 = 2n$ makes the Equation (6.9) as Hermite's differential equation, and the polynomial $h(\rho)$ is, in fact, Hermite's polynomial. Therefore, the energy eigenstates of harmonic oscillator are given by

$$u_n(\rho) = N_n e^{-\frac{\rho^2}{2}} H_n(\rho) \tag{6.16}$$

where $H_n(\rho)$ is Hermite's polynomial, and N_n is normalization constant. Though $H_n(\rho)$ tends to infinity for large ρ, the factor $e^{-\rho^2/2}$ ensures $u_n(\rho)$ goes to zero as $\rho \to \infty$.

Energy Eigenvalue

From the Equation (6.5), we know that the condition $\lambda - 1 = 2n$ implies

$$\frac{2E}{\hbar\omega} = (2n+1)$$
$$2E = \hbar\omega(2n+1)$$

or more appropriately,

$$E_n = \hbar\omega\left(n+\frac{1}{2}\right) \qquad n = 0,1,2,\ldots \tag{6.17}$$

The energy eigenvalues are discrete and evenly spaced. The corresponding eigenfunctions are

$$u_n(x) = N_n e^{-\frac{\alpha^2 x^2}{2}} H_n(\alpha x)$$

where
$$\alpha = \left(\frac{m\omega}{\hbar}\right)^{1/2} \tag{6.18}$$

Normalization and Orthoganality

The normalization constants can be determined from

$$(u_n(x), u_m(x)) = \delta_{nm}$$

$$\text{or } \int_{-\infty}^{\infty} u_n^*(x) u_m(x)\,dx = \delta_{nm}$$

or
$$\int_{-\infty}^{\infty} N_m N_n e^{-\alpha^2 x^2} H_n(\alpha x) H_m(\alpha x)\,dx = \delta_{nm}$$

In terms of $\rho = \alpha x$, this condition becomes

$$N_n N_m \int_{-\infty}^{\infty} e^{-\rho^2} H_n(\rho) H_m(\rho)\,d\rho = \alpha\,\delta_{nm} \tag{6.19}$$

We can write this integral as

$$I_{nm} = \int_{-\infty}^{\infty} e^{-\rho^2} H_n(\rho) H_m(\rho)\,d\rho = \frac{1}{N_n N_m}\alpha\,\delta_{nm} \tag{6.20}$$

The standard method of establishing the orthonormality of the eigenfunctions for the harmonic oscillator and the determination of the normalization constant N_n uses the generating function $G(\rho, t)$ for Hermite's polynomials, which is given by

$$G(\rho, t) = e^{-t^2+2t\rho} = \sum_{n=0}^{\infty} \frac{H_n(\rho)}{n!} t^n \tag{6.21}$$

The coefficients of $t^n/n!$ are the Hermite's polynomials. For the sake of manipulation, let us again write the generating function as $G(\rho, s)$ with s as a parameter.

$$G(\rho, s) = e^{-s^2+2s\rho} = \sum_{m=0}^{\infty} \frac{H_m(\rho) s^m}{m!} \tag{6.22}$$

Let us evaluate the integral $\int_{-\infty}^{\infty} e^{-\rho^2} G(\rho,t) G(\rho,s)\,d\rho$. Making use of (6.21) and (6.23),

$$\int_{-\infty}^{\infty} e^{-\rho^2} G(\rho,t)\; G(\rho,s)\,d\rho = \int_{-\infty}^{\infty} e^{-\rho^2} \left(\sum_{n=0}^{\infty} \frac{H_n(\rho) t^n}{n!}\right)\left(\sum_{m=0}^{\infty} \frac{H_m(\rho) s^m}{m!}\right) d\rho \tag{6.23}$$

$$= \sum_{n=0}^{\infty}\sum_{m=0}^{\infty} \int_{-\infty}^{\infty} e^{-\rho^2} H_n(\rho) H_m(\rho)\,d\rho \frac{1}{n!}\frac{1}{m!} t^n s^m = \sum_{n=0}^{\infty}\sum_{m=0}^{\infty} \frac{1}{n!}\frac{1}{m!} t^n s^m I_{nm}$$

But L.H.S. of (6.23) is also given by

$$\int_{-\infty}^{\infty} e^{-\rho^2} G(\rho,t)\; G(\rho,s)\,d\rho = \int_{-\infty}^{\infty} e^{-\rho^2} e^{-t^2+2t\rho} e^{-s^2+2s\rho}\,d\rho$$

$$= \int_{-\infty}^{\infty} e^{-(\rho^2+t^2+s^2-2t\rho-2s\rho+2st)} e^{2st}\,d\rho$$

$$= e^{2st} \int_{-\infty}^{\infty} e^{-(\rho-t-s)^2}\,d\rho$$

Change the integration variable from ρ to u where $u = \rho - t - s$. Then

$$\int_{-\infty}^{\infty} e^{-(\rho-t-s)^2} d\rho = \int_{-\infty}^{\infty} e^{-u^2} du = \sqrt{\pi}$$

$$\therefore \int_{-\infty}^{\infty} e^{-\rho^2} G(\rho,t)G(\rho,s)d\rho = \sqrt{\pi}e^{2st} \tag{6.24}$$

$$= \sqrt{\pi}\sum_{n=0}^{\infty}\frac{(2st)^n}{n!} = \sqrt{\pi}\sum_{n=0}^{\infty}\frac{2^n s^n t^n}{n!}$$

Comparing (6.23) and (6.24), we get

$$\begin{aligned}\sum_{n=0}^{\infty}\sum_{m=0}^{\infty}\frac{t^n}{n!}\frac{s^m}{m!} I_{mn} &= \sqrt{\pi}\sum_{n=0}^{\infty}\frac{2^n s^n t^n}{n!} \\ &= \sqrt{\pi}\sum_{n=0}^{\infty}\frac{2^n t^n}{n!}\sum_{m=0}^{\infty} s^m \frac{m!}{m!}\delta_{mn} \qquad \left(\text{use} \quad \sum_{m=0}^{\infty} s^m \delta_{mn} = s^n\right) \\ &= \sum_{n=0}^{\infty}\sum_{m=0}^{\infty}\sqrt{\pi}\,2^n \frac{t^n}{n!}\frac{s^m m!}{m!}\delta_{mn} \\ &= \sum_{n=0}^{\infty}\sum_{m=0}^{\infty}\frac{t^n}{n!}\frac{s^m}{m!}\underbrace{\sqrt{\pi}\,2^n m!\,\delta_{mn}}_{I_{mn}} \end{aligned} \tag{6.25}$$

This is possible, provided $\quad I_{mn} = n!\sqrt{\pi}\,2^n \delta_{mn}$

From (6.20), we get

$$N_m = \left(\frac{\alpha}{m!2^m\sqrt{\pi}}\right)^{1/2} \tag{6.26}$$

Therefore, the normalized wave function is

$$u_n(x) = \left(\frac{\alpha}{n!2^n\sqrt{\pi}}\right)^{1/2} e^{\frac{-\alpha^2x^2}{2}} H_n(\alpha x) \tag{6.27}$$

Energy Eigenvalues and Their Eigenfunctions

Let us list some of the energy eigenfunctions below.

Energy	Wave function
$E_0 = \frac{\hbar\omega}{2}$	$u_0 = \left(\frac{\alpha}{\sqrt{\pi}}\right)^{1/2} e^{-\frac{\alpha^2x^2}{2}}$
$E_1 = \frac{3}{2}\hbar\omega$	$u_1 = \left(\frac{\alpha}{2\pi}\right)^{1/2} 2\alpha\, xe^{-\frac{\alpha^2x^2}{2}}$

$$E_2 = \frac{5}{2}\hbar\omega \qquad u_2 = \left(\frac{\alpha}{2\sqrt{\pi}}\right)^{1/2}(2\alpha^2x^2 - 1)e^{-\frac{\alpha^2x^2}{2}}$$

$$E_3 = \frac{7}{2}\hbar\omega \qquad u_3 = \left(\frac{\alpha}{3\sqrt{\pi}}\right)^{1/2}(2\alpha^3x^3 - 3\alpha x)e^{-\frac{\alpha^2x^2}{2}}$$

$$E_4 = \frac{9}{2}\hbar\omega \qquad u_4 = \left(\frac{\alpha}{24\sqrt{\pi}}\right)^{1/2}(4\alpha^4x^4 - 12\alpha^2x^2 + 3)e^{-\frac{\alpha^2x^2}{2}}$$

The sketch of these functions is given in Fig. 6.1.

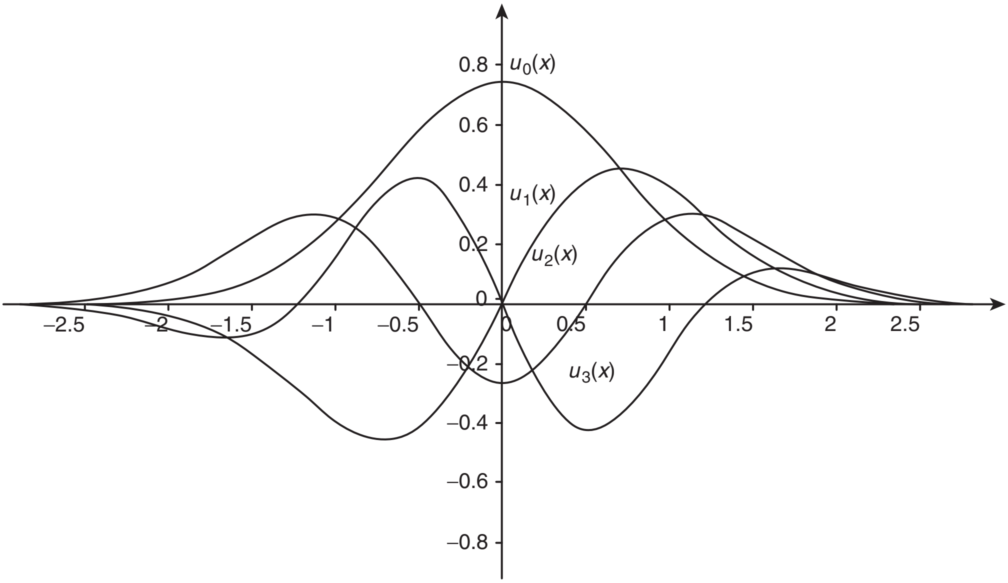


Fig. 6.1 Harmonic oscillator wave functions for $n = 0, 1, 2, 3.(\alpha = 1.0)$

6.1.2 Zero Point Energy

The energy of a simple harmonic oscillator is given by

$$E_n = \left(n + \frac{1}{2}\right)\hbar\omega \tag{6.28}$$

The lowest possible energy of a simple harmonic oscillator corresponds to

$$E_0 = \frac{\hbar\omega}{2}$$

This is in contrast to classical physics where the minimum energy of the harmonic oscillator is zero, which happens when the particle is at rest, i.e., $x = 0$ and $p = 0$. It amounts to specifying both the position and momentum exactly; neither of which is possible in quantum mechanics. This will violate Heisenberg's uncertainty principle. An exact location of a particle means $\Delta x = 0$, which will make Δp

very large, and hence the energy of the particle should become very large. A particle can be located only within a small range Δx, which implies that there will be a spread of Δp in momentum also. Consequently, the minimum energy ΔE cannot be zero.

It is an interesting exercise to estimate the zero point energy from the uncertainty principle.

Example 6.1 Estimate the lowest possible energy of a harmonic oscillator using the uncertainty relation $\Delta x \, \Delta p \geq \frac{\hbar}{2}$.

Solution: Classically, for the lowest energy, $x = 0$ and $p = 0$. Quantum mechanically, this is not possible. Let Δx and Δp be the spread in position and momentum around $x = 0$ and $p = 0$. The energy of the particle is

$$E = \frac{(\Delta p)^2}{2m} + \frac{1}{2} m\omega^2 (\Delta x)^2$$

From the uncertainty relation, we have $\Delta p \geq \frac{\hbar}{2\Delta x}$

$$\therefore E \geq \frac{\hbar^2}{8m(\Delta x)^2} + \frac{1}{2} m\omega^2 (\Delta x)^2$$

From this relation, the uncertainty Δx corresponding to the minimum energy E can be determined from

$$\frac{dE}{d(\Delta x)} = 0$$

$$-\frac{2\hbar^2}{8m(\Delta x)^3} + \frac{1}{2} m\omega^2 \Delta x = 0$$

or

$$(\Delta x)^2 = \frac{\hbar}{2m\omega}$$

$$\therefore \; E_{\min} = \frac{\hbar 2}{8m \cdot \frac{\hbar}{2m\omega}} + \frac{1}{2} m\omega^2 \cdot \frac{\hbar}{2m\omega} = \frac{\hbar\omega}{2}$$

It is a coincidence that a rough way of estimating minimum energy yields the same value which is obtained by a long rigourous calculation.

6.1.3 Comparison with Classical Physics

The standard example of a harmonic oscillator is a simple pendulum or an object attached to a spring. It executes a simple harmonic motion described by

$$x(t) = a \sin\omega t$$

$$v(t) = a\omega \cos\omega t$$

$$p(t) = ma\omega \cos\omega t$$

and

$$E = \frac{1}{2} ma^2\omega^2$$

The velocity is maximum at $x = 0$ and minimum at the extreme points. The position of a particle is in the range $-\sqrt{2E/m\omega^2} < x < \sqrt{2E/m\omega^2}$. The points $x = \sqrt{2E/m\omega^2}$ and $x = -\sqrt{2E/m\omega^2}$ are known as the classical turning points for a particle of energy E executing simple harmonic motion.

We have already seen the difference between classical physics and quantum physics with respect to the possible lowest energy of a harmonic oscillator. In classical mechanics, the lowest possible energy of a harmonic oscillator is zero, in which case, the particle is at rest. In quantum mechanics, this condition violates the uncertainty principle. The minimum energy of quantum harmonic oscillator is $\hbar\omega/2$, which is consistent with uncertainty principle.

There is another important difference pertaining to the position probability for a harmonic oscillator between classical and quantum physics. Though the position $x(t)$ is deterministic in classical physics, we can still introduce the concept of probability of finding the particle in an interval dx around x in the following way (this is artificial). Let T be the period of the simple harmonic motion of the particle. Let the particle be between x and $x + dx$ in the time interval between t and dt. Then the probability of finding the particle between x and $x + dx$ can be defined as

$$P(x)dx = \frac{dt}{T}$$

Since

$$x = a \sin \omega t$$

$$dx = a\omega \cos \omega t \, dt$$

$$dt = \frac{dx}{a\omega \cos \omega t}$$

$$\therefore \; P(x)dx = \frac{dx}{T.a\omega \cos \omega t} = \frac{dx}{2\pi a \cos \omega t}$$

$$= \frac{dx}{2\pi(a^2 - x^2)^{1/2}} = \frac{dx}{2\pi\left(\frac{2E}{m\omega^2} - x^2\right)^{1/2}}$$

The probability of finding the particle is minimum near the origin and maximum near the classical turning points. This can be compared with the quantum mechanical probability density $P(x)dx = |u_n(x)|^2 \, dx$. For the ground state, the quantum mechanical distribution differs completely from classical physics. For small values of n, they differ widely. But when n becomes very large, the probability distribution $|u_n(x)|^2$ fluctuates rapidly as shown in Fig. 6.3, but the average value tends to be closer to classical probability distribution.

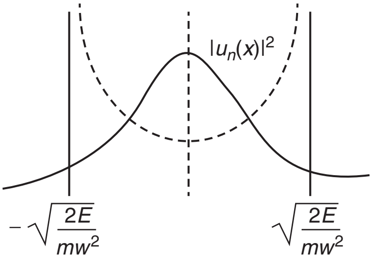


Fig. 6.2 $|u_0(x)|^2$ (solid curve) and classical $P(x)$ (dashed curve)

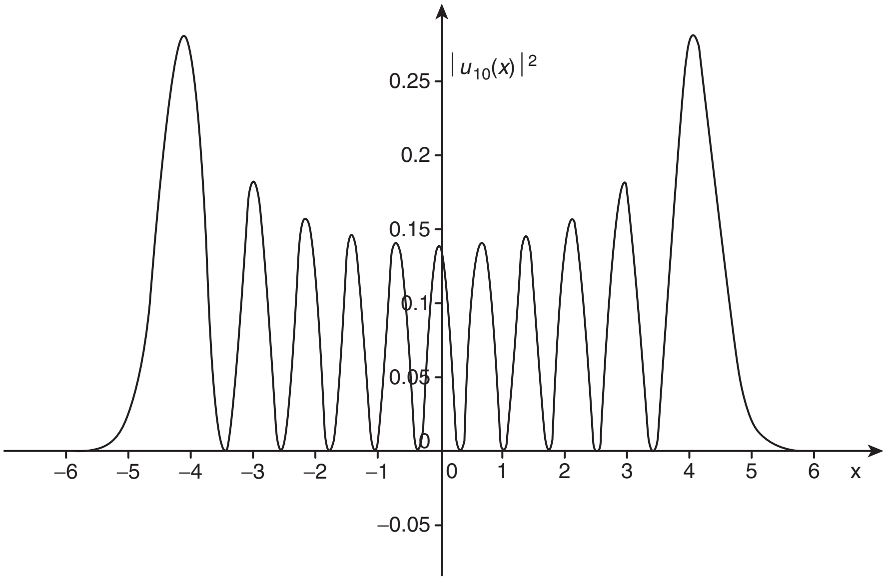


Fig. 6.3 Sketch of $\left|u_{10}(x)\right|^2$

Example 6.2 Evaluate the matrix element x_{mn} for harmonic oscillator where

$$x_{mn} = (u_m, xu_n)$$

Solution:

$$\begin{aligned}(u_m, xu_n) &= \int_{-\infty}^{\infty} N_n N_m e^{-\alpha^2 x^2} x\, H_m(\alpha x) H_n(\alpha x) dx \\ &= \frac{N_n N_m}{\alpha^2} \int_{-\infty}^{\infty} e^{-\rho^2} \rho H_m(\rho)\, H_n(\rho) d\rho \\ &= \frac{N_n N_m}{\alpha^2} I_{mn} \end{aligned} \tag{6.29}$$

where
$$I_{mn} = \int_{-\infty}^{\infty} e^{-\rho^2} \rho H_m(\rho) H_n(\rho) d\rho$$

To evaluate this integral, we again use the generating function for Hermite's polynomials.

$$\begin{aligned}\int_{-\infty}^{\infty} e^{-\rho^2} \rho\; G(\rho,t)\; G(\rho,s) d\rho &= \int_{-\infty}^{\infty} e^{-\rho^2} e^{-t^2+2t\rho} e^{-s^2+2s\rho} \rho d\rho \\ &= \sum_{n=0}^{\infty}\sum_{m=0}^{\infty} \int_{-\infty}^{\infty} e^{-\rho^2}\, \frac{H_n(\rho)}{n!} \frac{H_m(\rho)\rho}{m!} t^n s^m d\rho \\ &= \sum_{n=0}^{\infty}\sum_{m=0}^{\infty} \frac{t^n}{n!} \frac{s^m}{m!} I_{mn} \end{aligned} \tag{6.30}$$

$$= \int_{-\infty}^{\infty} e^{-\rho^2} \rho \, e^{-t^2+2t\rho} e^{-s^2+2s\rho} d\rho$$

$$= \int_{-\infty}^{\infty} e^{-(\rho^2+t^2+s^2-2t\rho-2s\rho+2st)} \rho e^{2st} d\rho$$

$$= e^{2st} \int_{-\infty}^{\infty} e^{-(\rho-t-s)^2} \rho d\rho$$

Put $u = \rho - t - s$

$$\int_{-\infty}^{\infty} e^{-\rho^2} \rho G(\rho,t) G(\rho,s) d\rho = e^{2st} \int_{-\infty}^{\infty} e^{-u^2} (u+t+s) du = e^{2st}(t+s)\sqrt{\pi}$$

(Note $\int_{-\infty}^{\infty} e^{-u^2} u du = 0$, since the integrand is an odd function of u).

$$\therefore \int_{-\infty}^{\infty} e^{-\rho^2} \rho G(\rho,t) G(\rho,s) d\rho = \sqrt{\pi}(t+s) \sum_{n=0}^{\infty} \frac{2^n t^n s^n}{n!}$$

$$= \sqrt{\pi} \sum_{n=0}^{\infty} 2^n \frac{t^{n+1} s^n}{n!} + \sqrt{\pi} \sum_{n=0}^{\infty} 2^n \frac{t}{n!} s^{n+1} \tag{6.31}$$

From (6.30) and (6.31), we get

$$= \sqrt{\pi} \sum_{m=0}^{\infty} 2^m \frac{s^m}{m!} \underbrace{\sum_{n=0}^{\infty} \frac{t^n}{n!} n! \delta_{n,m+1}}_{t^{m+1}} + \sqrt{\pi} \sum_{n=0}^{\infty} 2^n \frac{t^n}{n!} \underbrace{\sum_{m=0}^{\infty} \frac{s^m}{m!} m! \delta_{m,n+1}}_{s^{n+1}}$$

$$= \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} \frac{t^n}{n!} \frac{s^m}{m!} \underbrace{\left\{ \sqrt{\pi} 2^m n! \delta_{n,m+1} + \sqrt{\pi} 2^n m! \delta_{m,n+1} \right\}}_{I_{mn}}$$

$$\therefore I_{mn} = \begin{cases} \sqrt{\pi} 2^m (m+1)! & \text{if} \quad n = m+1 \\ \sqrt{\pi} 2^{m-1} m! & \text{if} \quad m = n+1 \\ 0 & \text{Otherwise} \end{cases}$$

$$\therefore x_{\text{mn}} = \frac{N_n N_m}{\alpha^2} I_{mn}$$

$$= \begin{cases} \dfrac{1}{\alpha} \cdot \left(\dfrac{m+1}{2} \right)^{1/2} & \text{if} \quad m = n-1 \\ \dfrac{1}{\alpha} \left(\dfrac{m}{2} \right) & \text{if} \quad m = n+1 \\ 0 & \text{Otherwise} \end{cases}$$

Example 6.3 Evaluate the matrix element x^2_{mn} for simple harmonic oscillator.

Solution:

$$(x^2)_{mn} = \int u_m^*(x) x^2 u_n(x) dx = N_n N_m \int e^{-\alpha^2 x^2} H_n(\alpha x) H_m(\alpha x) x^2 dx$$

$$= \frac{N_n N_m}{\alpha^3}\int_{-\infty}^{\infty} e^{-\rho^2}\rho^2 H_n(\rho)H_m(\rho) = \frac{N_n N_m}{\alpha^3} I_{mn}$$

where $$I_{mn} = \int_{-\infty}^{\infty} e^{-\rho^2}\rho^2 H_m(\rho)H_n(\rho)$$

Consider the integral $$\int_{-\infty}^{\infty} e^{-\rho^2}\rho^2 G(\rho,t)G(\rho,s)d\rho$$

$$\int_{-\infty}^{\infty} e^{-\rho^2}\rho^2 G(\rho,t)G(\rho,s)d\rho = \int_{-\infty}^{\infty} e^{-\rho^2}\rho^2 e^{-t^2+2\rho t-s^2+2\rho s}d\rho = e^{2st}\int_{-\infty}^{\infty} e^{-(\rho-t-s)^2}\rho^2 d\rho$$

Put $u = \rho - t - s$

$$\begin{aligned}\int_{-\infty}^{\infty} e^{-\rho^2}\rho^2 G(\rho,t)G(\rho,s)d\rho &= e^{2st}\int_{-\infty}^{\infty} e^{-u^2}(u+t+s)^2 du \\ &= e^{2st}\int_{-\infty}^{\infty} e^{-u^2}(u^2+t^2+s^2+2ut+2us+2ts)du \\ &= e^{2st}\left\{\int_{-\infty}^{\infty} e^{-u^2}u^2 du + \int_{-\infty}^{\infty} e^{-u^2}(t^2+s^2+2ts)du\right\} \\ &= e^{2st}\left\{\frac{\sqrt{\pi}}{2}s + \sqrt{\pi}(t^2+s^2+2ts)\right\}\end{aligned}$$

But $$\begin{aligned}\int_{-\infty}^{\infty} e^{-\rho^2}\rho^2 G(\rho,t)G(\rho,s)d\rho &= \sum_{n=0}^{\infty}\sum_{m=0}^{\infty}\int_{-\infty}^{\infty} e^{-\rho^2}\rho^2 H_n(\rho)H_m(\rho)\frac{t^n}{n!}\frac{s^m}{m!} \\ &= \sum_{n=0}^{\infty}\sum_{m=0}^{\infty}\frac{t^n}{n!}\frac{s^m}{m!}I_{mn}\end{aligned}$$

$$\begin{aligned}&= \frac{\sqrt{\pi}}{2}\sum_{n=0}^{\infty}\frac{2^n t^n}{n!}\sum_{m=0}^{\infty}\frac{s^m}{m!}m!\delta_{mn} + \sqrt{\pi}\sum_{m=0}^{\infty}\frac{2^m s^m}{m!}\sum_{n=0}^{\infty}\frac{t^n}{n!}n!\delta_{n,m+2} \\ &\quad + \sqrt{\pi}\sum_{n=0}^{\infty}2^n\frac{t^n}{n!}\sum_{m=0}^{\infty}\frac{s^m}{m!}m!\delta_{m,n+2} + \sqrt{\pi}\sum_{n=0}^{\infty}\frac{2^{n-1}t^n n}{n!}\sum_{m=0}^{\infty}\frac{s^m}{m!}m!\delta_{mn} \\ &= \sum_{n=0}^{\infty}\sum_{m=0}^{\infty}\frac{t^n}{n!}\frac{s^m}{m!}\underbrace{\left\{\sqrt{\pi}\,2^n n!\delta_{mn} + \sqrt{\pi}\,2^m n!\delta_{n,m+2} + \sqrt{\pi}\,2^n m!\delta_{m,n+2}\right\}}_{I_{mn}}\end{aligned}$$

$$\therefore I_{mn} = \begin{cases}\sqrt{\pi}n!2^n\left(\frac{1}{2}+n\right) & \text{if} \quad m = n \\ \sqrt{\pi}\,2^m n! & \text{if} \quad n = m+2 \\ \sqrt{\pi}\,2^n m! & \text{if} \quad m = n+2 \\ 0 & \text{Otherwise}\end{cases}$$

$$(x^2)_{mn} = \begin{cases} \dfrac{(2n+1)}{\alpha^2} & \text{if} \quad m = n \\ \dfrac{[(m+1)(m+2)]^{1/2}}{2\alpha^2} & \text{if} \quad m = n+2 \\ \dfrac{[(m+3)(m+4)]^{1/2}}{2\alpha^2} & \text{if} \quad m = n-2 \\ 0 & \text{Otherwise} \end{cases}$$

Example 6.4 Evaluate the matrix element $\langle p\rangle_n$ for a harmonic oscillator state u_n.

Solution:

$$\langle p\rangle_n = \int u_n^*(x)\left(-i\hbar\frac{d}{dx}\right)u_n(x)dx$$

$$pu_n(x) = -i\hbar\frac{d}{dx}\left[e^{-\frac{\alpha^2x^2}{2}}H_n(\alpha x)\right]$$

$$= -i\hbar\left[-\alpha^2xe^{-\frac{\alpha^2x^2}{2}}H_n(\alpha x) + e^{-\frac{\alpha^2x^2}{2}}\frac{dH_n}{dx}\right.$$

$$= -i\hbar e^{-\frac{\alpha^2x^2}{2}}[-\alpha^2xH_n + 2nH_{n-1}]$$

$$\therefore \langle p\rangle_n = (-i\hbar)N_n^2\int_{-\infty}^{\infty} e^{-\frac{\alpha^2x^2}{2}}H_n(\alpha x)\cdot(-\alpha^2x)e^{-\frac{\alpha^2x^2}{2}}H_n(\alpha x)dx$$

$$-2ni\hbar\frac{N_n}{N_{n-1}}\int_{-\infty}^{\infty} N_ne^{-\frac{\alpha^2x^2}{2}}H_n(\alpha x)N_{n-1}e^{-\frac{\alpha^2x^2}{2}}H_{n-1}(\alpha x)$$

$$= i\hbar\alpha^2x_{nn} - 2in\hbar\frac{N_n}{N_{n-1}}(u_n, u_{n-1})$$

Both the terms are zero.

$$\therefore \langle p\rangle_n = 0$$

Example 6.5 Evaluate the matrix element $\langle p^2\rangle_n$.

Solution:

$$\langle p^2\rangle_n = \int_{-\infty}^{\infty} u_n^*(x)\left(-\hbar^2\frac{d^2}{dx^2}\right)u_n(x)dx$$

$$p^2u_n(x) = -\hbar^2\frac{d^2u_n}{dx^2}$$

The Schrödinger equation is

$$-\frac{\hbar^2}{2m}\frac{d^2u}{dx^2} + \frac{1}{2}m\omega^2x^2u_n = E_nu_n$$

$$\therefore -\hbar^2\frac{d^2u}{dx^2} = 2mE_nu_n - m^2\omega^2x^2u_n$$

$$
\begin{aligned}
(p^2)_{nn} &= \int u_n^*(x)2mE_n u_n(x)dx - m^2\omega^2\int u_n^* x^2 u_n dx \\
&= 2mE_n - m^2\omega^2 x_{nn}^2 = 2mE_n - m^2\omega^2\frac{(2n+2)}{\alpha^2} \\
&= \frac{(2n+1)}{2}m\hbar\omega
\end{aligned}
$$

Example 6.6 Evaluate the uncertainty product $\Delta x\Delta p$ for an arbitrary energy eigenstate u_n of a simple harmonic oscillator.

Solution: The uncertainty $(\Delta x)^2$ is given by

$$
\begin{aligned}
(\Delta x)^2 &= \langle x^2\rangle_n - (\langle x\rangle_n)^2 = \frac{(2n+1)}{2\alpha^2} \\
(\Delta p)^2 &= \langle p^2\rangle_n - (\langle p\rangle_n)^2 = \frac{(2n+1)}{2}m\hbar\omega
\end{aligned}
$$

$$
\begin{aligned}
(\Delta x)^2(\Delta p)^2 &= \frac{(2n+1)^2\hbar^2}{4} \\
\therefore\ \Delta x\,\Delta p &= \frac{(2n+1)\hbar}{2}
\end{aligned}
$$

The ground state $n = 0$ corresponds to minimum uncertainty state.

Example 6.7 The wave function for a harmonic oscillator at $t = 0$ is given by

$$\psi(x, 0) = C_0u_0(x) + C_1u_1(x)$$

Assume C_0 and C_1 to be real numbers. Show that the probability density $|\psi(x,t)|^2$ oscillates with a frequency ω.

Solution: For a harmonic oscillator, $\psi(x, t)$ is given by

$$\psi(x,t) = e^{-\frac{iHt}{\hbar}}\psi(x,0) = C_0\ u_0(x)e^{-\frac{iE_0t}{\hbar}} + C_1\ u_1(x)e^{-\frac{iE_1t}{\hbar}}$$

Substituting $E_0 = \frac{\hbar\omega}{2}$ and $E_1 = \frac{3}{2}\hbar\omega$, we get

$$
\begin{aligned}
\psi(x,t) &= C_0\ u_0(x)e^{-\frac{i\omega t}{2}} + C_1\ u_1(x)e^{-\frac{i3\omega t}{2}} \\
&= e^{-i\omega t}\left[C_0\ u_0(x)e^{\frac{i\omega t}{2}} + C_1\ u_1(x)e^{-\frac{i\omega t}{2}}\right]
\end{aligned}
$$

$$
\begin{aligned}
|\psi(x,t)|^2 &= \left[C_0u_0(x)e^{-\frac{i\omega t}{2}} + C_1u_1e^{-\frac{i\omega t}{2}}\right]\cdot\left[C_0u_0e^{-\frac{i\omega t}{2}} + C_1u_1e^{-\frac{i\omega t}{2}}\right] \\
&= C_0^2u_0^2 + C_1^2u_1^2 + C_1C_0u_0u_1e^{i\omega t} + C_1C_0u_0u_1e^{-i\omega t} \\
&= C_0u_0^2 + C_1^2u_1^2 + 2C_0C_1u_0u_1\cos\omega t
\end{aligned}
$$

6.2 ABSTRACT OPERATOR METHOD

So far, the simple harmonic oscillator was analyzed in terms of Schrödinger wave equation. The energy eigenvalues and the eigenfunctions were obtained by solving the Schrödinger's differential equation explicitly. The matrix elements like x_{mn}, x^2_{mn}, the uncertainties Δx and Δp and so on were all evaluated using wave functions in the coordinate representation. There exists another method of studying harmonic oscillator, which does not require solving a differential equation or evaluation of an integral. This is called abstract operator or ladder operator method.

In using the Schrödinger equation, the operators x_{op} and p_{op} were taken as $x_{op} = x$ and $p_{op} = -i\hbar(d/dx)$. In the abstract operator method, explicit form of the operators x_{op} and p_{op} is not used, and only the commutation relation $[x_{op}, p_{op}] = i\hbar$ is used. In fact, it is because of abstract operator method, harmonic oscillator finds its applications extensively in other fields.

Let us recall that the Hamiltonian of the oscillator is $H = \frac{p_{op}^2}{2m} + \frac{1}{2}m\omega^2 x_{op}^2$.

This Hamiltonian along with the commutation relation $[x_{op}, p_{op}] = i\hbar$ are sufficient to determine the energy eigenvalues and other properties of a harmonic oscillator.

6.2.1 Ladder Operators *a* and *a*†

Let us introduce two new operators a and $a^\dagger$ which are a combination of x_{op} and p_{op}.

$$a = C_1 x_{op} + iC_2 p_{op}$$
$$a^\dagger = C_1 x_{op} - i\,C_2 p_{op}$$

The standard choice is $C_1 = \left(\frac{m\omega}{2\hbar}\right)^{1/2}$ and $C_2 = \frac{1}{(2\hbar m\omega)^{1/2}}$. so the operator a and $a^\dagger$ are given by

$$a = \left(\frac{m\omega}{2\hbar}\right)^{1/2} x_{op} + i\frac{1}{(2\hbar m\omega)^{1/2}} p_{op} \tag{6.32}$$

and

$$a^\dagger = \left(\frac{m\omega}{2\hbar}\right)^{1/2} x_{op} - i\frac{1}{(2\hbar m\omega)^{1/2}} p_{op} \tag{6.33}$$

One cannot give any immediate rationale behind the choice of these coefficients accompanying x_{op} and p_{op}. The only justification is that this choice leads to a very important commutation relation $[a, a^\dagger] = 1$. It is this commutation relation which forms the backbone of harmonic oscillator and its extensive applications.

Commutator [*a*, *a*†]

Let us evaluate the commutation relation $[a, a^\dagger]$.

$$[a, a^\dagger] = \left[\left(\frac{m\omega}{2\hbar}\right)^{1/2} x + \frac{i}{(2\hbar m\omega)^{1/2}} p, \left(\frac{m\omega}{2\hbar}\right)^{1/2} x - \frac{ip}{(2\hbar m\omega)^{1/2}}\right]$$

$$= \left(\frac{m\omega}{2\hbar}\right)^{1/2} \cdot \left(\frac{-i}{(2\hbar m\omega)^{1/2}}\right)[x, p] + \frac{i}{(2\hbar m\omega)^{1/2}} \cdot \left(\frac{m\omega}{2}\right)^{1/2}[p, x]$$

$$= \frac{1}{2} + \frac{1}{2} = 1$$

i.e., $$[a, a^\dagger] = 1 \tag{6.34}$$

Example 6.8 The operators a and $a^\dagger$ are given by $a = C_1x + iC_2p$ and $a^\dagger = C_1x - iC_2p$. Find the conditions on C_1 and C_2, which lead to $[a, a^\dagger] = 1$

Solution:

$$\begin{aligned}[a, a^\dagger] &= [C_1x + iC_2p, C_1x - iC_2p] \\ &= C_1 - iC_1C_2[x, p] + iC_2C_1[p, x] \\ &= 2C_1C_2\hbar \;= 1 \\ \therefore C_1C_2 &= \frac{1}{2\hbar}\end{aligned}$$

Note that $C_1 = m\omega/2\hbar^{1/2}$and $C_2 = 1/(2\hbar m\omega)^{1/2}$ satisfy this condition. Note both a and $a^\dagger$ are dimensionless quantities.

6.2.2 Consequences of [*a*, *a*†] = 1

Number Operator

Let us define the number operator N as $$N = a^\dagger a \tag{6.35}$$

Let us now determine the eigenvalues of the number operator N. The eigenvalue equation is

$$Nu = \lambda u$$

i.e., $$a^\dagger au = \lambda u$$

Consider the scalar product of au with itself.

$$(au, au) = (u, a^\dagger au) = (u, \lambda u) = \lambda(u, u) = \lambda$$

For any state vector, the scalar product of it with itself is a positive number.

$$(au, au) \geq 0 \;\therefore\; \lambda \geq 0$$

i.e., the eigenvalues of the number operator are always positive.

a†-Raising or Creation Operator

Let u be the eigenstate of the number operator with eigenvalue λ.

i.e., $$Nu = \lambda\, u$$

From this state, construct another state $a^\dagger u$. Now, consider $Na^\dagger u$.

$$Na^\dagger u = a^\dagger a a^\dagger u$$

From $a\,a^\dagger - a^\dagger a = 1$, we write

$$\begin{aligned}Na^\dagger u &= a^\dagger(1 + a^\dagger a)u \;= a^\dagger u + a^\dagger \underbrace{a^\dagger au}_{\lambda u} \\ &= a^\dagger u + \lambda a^\dagger u = (\lambda + 1)a^\dagger \mathrm{u}\end{aligned}$$

This is a very interesting result. If u is an eigenstate of N with eigenvalue λ, $a^\dagger u$ is also an eigenstate of the number operator N with eigenvalue $(\lambda + 1)$; i.e., $a^\dagger$ acting on u produces a new state for which eigenvalue of N increases by 1. So, $a^\dagger$ is called raising operator. In quantum field theory, it is called creation operator. So we have

$u \rightarrow$ eigenstate of N with eigenvalue λ
$a^\dagger u \rightarrow$ eigenstate of N with eigenvalue $(\lambda + 1)$

a – Lowering or Annihilation Operator

From the state u, now construct another state au. Now consider Nau.

$$Nau = a^\dagger aau$$

From $$a\,a^\dagger - a^\dagger a = 1, \text{ we get}$$

$$N\,au = (a\,a^\dagger - 1)au$$

$$= a\,\underbrace{a^{+}au}_{\lambda u} - au = \lambda au - au = (\lambda - 1)au$$

If u is an eigenstate of the number operator N with eigenvalue λ, then au is also an eigenstate of N with eigenvalue of $(\lambda - 1)$; i.e., a acting on the state u produces a new state for which eigenvalue of N has decreased by 1. So a is called lowering operator. In quantum field theory, it is called annihilation operator. So we have

u: eigenstate of N with eigenvalue λ
au: eigenstate of N with eigenvalue $(\lambda - 1)$

Lowest Eigenvalue of Number Operator

Let us start with the state u with the corresponding eigenvalue λ. The operator a acting on u produces a new state with an eigenvalue $(\lambda - 1)$. Let a act on this new state. It should produce another state with eigenvalue $(\lambda - 2)$. Repeating this procedure, we can go on producing new states with eigenvalues $(\lambda - 1)$, $(\lambda - 2)$, $(\lambda - 3)$ and so on.

State	Eigenvalue of N
u	λ
au	$\lambda - 1$
a^2u	$\lambda - 2$
a^3u	$\lambda - 3$
$\vdots$	$\vdots$

However, this cannot be repeated indefinitely since at some stage, we will get the eigenvalue of N to be a negative number. So by repeated application of lowering operator, we should reach a state from which it should not be possible for a to produce a new state. This is possible provided the sequence $\{\lambda, (\lambda - 1), (\lambda - 2),...\}$ ends in zero. i.e., the sequence should be $\{\lambda, (\lambda - 1), (\lambda - 2),...,0\}$. This implies λ should be a positive integer. (Otherwise, sequence will not end. For instance, suppose $\lambda = 1.2$. Then au will have eigenvalue 0.2, a^2u will have eigenvalue –0.8 and so on.)

So we should have a state u_0 such that

$$au_0 = 0$$

and so $$Nu_0 = a^\dagger au = 0$$

To emphasize the fact that λ is a positive integer, we replace it by letter n. Then we have

$$Nu_n = nu_n \tag{6.36}$$

$$Na^\dagger u_n = (n+1)\, a^\dagger u_n \tag{6.37}$$

$$Nau_n = (n-1)\, au_n \tag{6.38}$$

$$au_0 = 0 \tag{6.39}$$

and $$Nu_0 = 0 \tag{6.40}$$

6.2.3 Energy Eigenvalues of Harmonic Oscillator

From (6.32) and (6.33), we get

$$x_{op} = \frac{1}{2}\cdot\left(\frac{2\hbar}{m\omega}\right)^{1/2}[a+a^\dagger] \quad p_{op} = \frac{1}{2i}\cdot(2\hbar m\omega)^{1/2}[a-a^\dagger]$$

Now let us write the Hamiltonian H in terms of a and $a^\dagger$.

$$\begin{aligned} H &= \frac{p^2}{2m} + \frac{1}{2}m\omega^2 x^2 \\ &= -\frac{1}{2m}\cdot\frac{1}{4}(2\hbar m\omega)\cdot[a-a^\dagger][a-a^\dagger] + \frac{1}{4}\cdot\frac{1}{2}m\omega^2\,\frac{2\hbar}{m\omega}\cdot(a+a^\dagger)(a+a^\dagger) \\ &= \frac{\hbar\omega}{2}[a^\dagger a + aa^\dagger + 1] = \frac{\hbar\omega}{2}[2a^\dagger a + 1] = \hbar\omega\left(a^\dagger a + \frac{1}{2}\right) \\ &= \hbar\omega\left(N + \frac{1}{2}\right) \end{aligned} \tag{6.41}$$

Obviously, the Hamiltonian H commutes with the number operator N.

i.e., $$[H, N] = 0 \tag{6.42}$$

Therefore, H and N together have common eigenstate.

Since $$Nu_n = n\, u_n$$

$$Hu_n = \left(N + \frac{1}{2}\right)u_n = \left(n + \frac{1}{2}\right)\hbar\omega u_n \tag{6.43}$$

Therefore, the energy eigenvalues are

$$E_n = \left(n + \frac{1}{2}\right)\hbar\omega$$

6.2.4 [a, $a^\dagger$] = 1 – Some Observations

It is significant to realize that the consequences (from (6.36) to (6.40)) of $[a, a^\dagger] = 1$ as listed in Section 6.2.1. Have nothing to do with simple harmonic oscillators. They are the results of the commutation relation $[a, a^\dagger] = 1$. The existence of the number operator N and its eigenvalues, the properties of a and $a^\dagger$ as lowering and raising operators and the existence of eigenstate u_0 such that $au_0 = 0$ and $Nu_0 = 0$ are all results that have nothing to do with simple harmonic oscillator. In fact, in the derivation of these properties from $[a, a^\dagger] = 1$, the simple harmonic oscillator Hamiltonian $H = p^2/2m + 1/2m\omega^2x^2$ is nowhere used. All that matters is $[a, a^\dagger] = 1$. If, somehow, in a particular context, this algebra is obtained, then the rest of the results of section follow immediately. These results have to be interpreted in that particular context. For instance, in quantum field theory, N is interpreted as number operator corresponding to the number of particles; a and $a^\dagger$ are interpreted as annihilation and creation operators.

Example 6.9 Show that $[a,(a^\dagger)^n] = n(a^\dagger)^{n-1}$

Solution:

$$a(a^\dagger)^n = aa^\dagger(a^\dagger)^{n-1} = (a^\dagger a+1)(a^\dagger)^{n-1} = a^\dagger a(a^\dagger)^{n-1} + (a^\dagger)^{n-1}$$

$$\therefore a(a^\dagger)^n = a^\dagger a(a^\dagger)^{n-1} + (a^\dagger)^{n-1} \tag{6.44}$$

Using this equation, we can write

$$a(a^\dagger)^{n-1} = a^\dagger a(a^\dagger)^{n-2} + (a^\dagger)^{n-2}$$

Substituting this expression for $a(a^\dagger)^{n-1}$ in the RHS of the Equation (6.44), we get

$$a(a^\dagger)^n = a^\dagger[a^\dagger a(a^\dagger)^{n-2} + (a^\dagger)^{n-2}] + (a^\dagger)^{n-1} = (a^\dagger)^2 a(a^\dagger)^{n-2} + 2(a^\dagger)^{n-1}$$

Using the Equation (6.44), we can write $a(a^\dagger)^{n-2}$ as

$$a(a^\dagger)^{n-2} = a^\dagger a(a^\dagger)^{n-3} + (a^\dagger)^{n-3}$$

Making use of this expression for $a(a^\dagger)^{n-2}$ we get

$$a(a^\dagger)^n = (a^\dagger)^2[a^\dagger a(a^\dagger)^{n-3} + (a^\dagger)^{n-3}] + 2(a^\dagger)^{n-3} = (a^\dagger)^3 a(a^\dagger)^{n-3} + 3(a^\dagger)^{n-1}$$

Repeating this procedure, we get

$$a(a^\dagger)^n = (a^\dagger)^n a + n(a^\dagger)^{n-1}.$$

This gives

$$[a,(a^\dagger)^n] = n(a^\dagger)^{n-1} \tag{6.45}$$

6.2.5 Construction of Energy Eigenstates Using Ladder Operators

Let us start with the ground state u_0 which satisfies

$$au_0 = 0$$

We can get eigenstates corresponding to higher energies using the raising operator $a^\dagger$ as follows. $a^\dagger u_0$ is an eigenstate with energy $3/2\hbar\omega$; $(a^\dagger)^2u$ is an eigenstate with energy $5/2\hbar\omega$. Repeating this procedure, we can obtain eigenstates corresponding to any particular energy eigenvalue $E_n = (n+1/2)\hbar\omega$. Let u_n be the normalized eigenstate corresponding to energy $E_n = (n+1/2)\hbar\omega$.

$$\therefore u_n = C_n (a^\dagger)^n u_0$$

where C_n is a constant of normalization.

$$\begin{aligned}(u_n, u_n) &= (C_n (a^\dagger)^n u_0, C_n (a^\dagger)^n u_0) \\ &= |C_n|^2\, (u_0, a^n (a^\dagger)^n u_0)\end{aligned}$$

Let us now evaluate $a^n (a^\dagger)^n u_0$. Using (6.45), we get

$$a^n (a^\dagger)^n u_0 = a^{n-1} a (a^\dagger)^n u_0 = a^{n-1}[n(a^\dagger)^{n-1} + a^{\dagger n}] a u_0 = n a^{n-1}(a^\dagger)^{n-1} u_0 + a^{n-1}(a^\dagger)^n \underbrace{a\, u_0}_{\substack{\Downarrow \\ 0}}$$

$$= n a^{n-1}(a^\dagger)^{n-1} u_0 = n a^{n-2}\, \underbrace{a (a^\dagger)^{n-1}}\, u_0 = n a^{n-2}\, \underbrace{[(n-1)(a^\dagger)^{n-2} + (a^\dagger)^{n-2} a]}\, u_0$$

$$= n(n-1) a^{n-2} (a^\dagger)^{n-2} u_0$$

Proceeding in the same way, we get

$$a^n (a^\dagger)^n u_0 = n(n-1)\ldots 1\, u_0 = n!\, u_0 \tag{6.46}$$

$$(u_n, u_n) = C_n^2 n!\, (u_0, u_0) = C_n^2\, n!$$

$$\therefore C_n = \frac{1}{(n!)^{1/2}}$$

$\therefore$ The normalized wave function is given by

$$u_n = \frac{1}{(n!)^{1/2}} (a^\dagger)^n u_0 \tag{6.47}$$

Example 6.10 Show that $a^\dagger u_n = (n+1)^{1/2} u_{n+1}$

Solution:

$$\begin{aligned}a^\dagger u_n &= a^\dagger \cdot \frac{(a^\dagger)^n}{\sqrt{n!}} u_0 = \frac{1}{\sqrt{n!}} (a^\dagger)^{n+1} u_0 \\ &= \frac{1}{\sqrt{n!}} \sqrt{(n+1)!}\, u_{n+1} = \sqrt{n+1}\, u_{n+1}\end{aligned}$$

Example 6.11 Show that $a u_n = \sqrt{n}\, u_{n-1}$.

Solution:

$$au_n = a\frac{(a^\dagger)^n}{\sqrt{n!}}u_0 = \frac{1}{\sqrt{n!}}\left\{n(a^\dagger)^{n-1} + (a^\dagger)^n\, a\right\}u_0 = \frac{1}{\sqrt{n!}}n(a^\dagger)^{n-1}u_0$$

$$= \frac{n}{\sqrt{n!}}\sqrt{(n-1)!}\; u_{n-1} = \sqrt{n}\; u_{n-1}$$

Example 6.12 Evaluate $a^m(a^\dagger)^n u_0$.

Solution:

Case 1: $m = n$. From (6.47), we get $a^m(a^\dagger)^n u_0 = a^n(a^\dagger)^n u_0 = n!u_0$

Case 2: $m > n$.

$$a^m(a^\dagger)^n u_0 = a^{m-n}\ a^n(a^\dagger)^n u_0 = n!a^{m-n}u_0 = 0$$

Case 3: $m < n$ $\quad a^m(a^\dagger)^n u_0 = a^{m-1}a(a^\dagger)^n u_0 = a^{m-1}[n(a^\dagger)^{n-1}u_0 + (a^\dagger)^n a\ u_0] = na^{m-1}(a^\dagger)^{n-1}u_0$

Proceeding in the same way, we get

$$a^m(a^\dagger)^n u_0 = n(n-1)\cdots(n-m+1)(a^\dagger)^{n-m}u_0$$

Example 6.13 Evaluate x_{mn}.

Solution: Note $x = \frac{1}{\sqrt{2}\alpha}(a + a^\dagger)$

$$\therefore x_{mn} = (u_m, xu_n) = \left(\frac{(a^\dagger)^m}{\sqrt{m!}}u_0, \frac{(a+a^\dagger)}{\sqrt{2}\alpha}\frac{(a^\dagger)^n}{\sqrt{n!}}u_0\right)$$

$$= \frac{1}{\sqrt{m!}}\frac{1}{\sqrt{n!}}\frac{1}{\sqrt{2}\alpha}((a^\dagger)^m u_0, (a+a^\dagger)(a^\dagger)^n u_0)$$

$$= \frac{1}{\sqrt{m!}}\frac{1}{\sqrt{n!}}\frac{1}{\sqrt{2}\alpha}\{(u_0, a^{m+1}(a^\dagger)^n u_0) + (u_0, a^m(a^\dagger)^{n+1}u_0)\}$$

$$= \frac{1}{\sqrt{m!}}\frac{1}{\sqrt{n!}}\frac{1}{\sqrt{2}}\cdot\frac{1}{\alpha}\cdot\{n!\delta_{m+1,n} + m!\delta_{m,n+1}\}$$

$$= \frac{1}{\sqrt{2}}\cdot\frac{1}{\alpha}\begin{cases}(m+1)^{1/2} & \text{if} \quad m+1 = n\\ \sqrt{m} & \text{if} \quad m = n+1\\ 0 & \text{Otherwise}\end{cases}$$

Example 6.14 Evaluate x^2_{mn}.

Solution: The matrix element x^2_{mn} is given by

$$x^2_{mn} = (u_m, x^2 u_n)$$

The operator x is given by

$$x = \frac{1}{\sqrt{2}\alpha}(a + a^\dagger)$$

$$\therefore x^2 = \frac{1}{2\alpha^2}[aa^\dagger + a^2 + (a^\dagger)^2 + a^\dagger a] = \frac{1}{2\alpha^2}[a^2 + (a^\dagger)^2 + 2aa^\dagger - 1]$$

$$\therefore (u_m, x^2 u_n) = \left(\frac{1}{\sqrt{m!}}(a^\dagger)^m u_0, \frac{1}{2\alpha^2}(a^2 + (a^\dagger)^2 + 2aa^\dagger - 1)\frac{1}{\sqrt{n!}}(a^\dagger)^n u_0 \right)$$

$$= \frac{1}{\sqrt{m!}}\cdot\frac{1}{\sqrt{n!}}\cdot\frac{1}{2\alpha^2}\{(u_0, a^m (a^\dagger)^{n+2} u_0) + (u_0, a^{m+2}(a^\dagger)^n u_0) +$$
$$2(u_0, a^{m+1}(a^\dagger)^{n+1} u_0) - (u_0, a^m (a^\dagger)^n u_0)\}$$

$$= \begin{cases} \dfrac{1}{2\alpha^2}\dfrac{m!}{\sqrt{m!}}\dfrac{1}{\sqrt{(m-2)!}} & m = n+2 \\ \dfrac{1}{2\alpha^2}\dfrac{(m+2)!}{\sqrt{m}!(m+2)!} & m+2 = n \\ \dfrac{1}{2\alpha^2}\left[\dfrac{2(m+1)!}{\sqrt{m}!\sqrt{m!}} - \dfrac{m!}{\sqrt{m}!\sqrt{m!}}\right] & m = n \\ 0 & \text{Otherwise} \end{cases}$$

$$= \begin{cases} \dfrac{1}{2\alpha^2}[m(m-1)]^{1/2} & m = n+2 \\ \dfrac{1}{2\alpha^2}[(m+1)(m+2)]^{1/2} & m+2 = n \\ \dfrac{1}{2\alpha^2}(2m+1) & m = n \\ 0 & \text{Otherwise} \end{cases}$$

Example 6.15 Evaluate p_{mn}.

Solution: The matrix element p_{mn} is given by $(u_m, p\ u_n)$

The operator p is given by $$p = \frac{i\hbar\alpha(a^\dagger - a)}{\sqrt{2}}$$

$$\therefore p_{mn} = i\frac{\hbar\alpha}{\sqrt{2}}\cdot\frac{1}{\sqrt{m!}}\frac{1}{\sqrt{n!}}((a^\dagger)^m u_0, (a^\dagger - a)(a^\dagger)^n u_0)$$

$$= i\frac{\hbar\alpha}{\sqrt{2}}\frac{1}{\sqrt{m!}}\frac{1}{\sqrt{n!}}\{(u_0, a^m (a^\dagger)^{n+2} u_0) - (u_0, a^{m+1}(a^\dagger)^n u_0)\}$$

$$= i\frac{\hbar\alpha}{\sqrt{2}}\frac{1}{\sqrt{m!}}\frac{1}{\sqrt{n!}}\{[m!\delta_{m,n+1} - (m+1)!\delta_{m+1,n}]\}$$

$$= i\frac{\hbar\alpha}{\sqrt{2}}\begin{cases} \sqrt{m} & m = n+1 \\ -\sqrt{m+1} & m+1 = n \\ 0 & \text{otherwise} \end{cases}$$

Example 6.16 Evaluate p^2_{mn}.

Solution: The operator p^2 is given by

$$p^2 = -\frac{\hbar^2\alpha^2}{2}[(a^\dagger)^2 + a^2 - a^\dagger a - aa^\dagger] = \frac{\hbar^2\alpha^2}{2}[2aa^\dagger - 1 - (a^\dagger)^2 - a^2]$$

$$\therefore\ p^2_{mn} = \frac{\hbar^2\alpha^2}{2}\left(\frac{1}{\sqrt{m!}}(a^\dagger)^m u_0, (2aa^\dagger - 1 - (a^\dagger)^2 - a^2)\frac{1}{\sqrt{n!}}(a^\dagger)^n u_0\right)$$

$$= \frac{\hbar^2\alpha^2}{2\sqrt{m!}\sqrt{n!}}\{(u_0, 2a^{m+1}(a^\dagger)^{n+1}u_0) - (u_0, a^m(a^\dagger)^n u_0) - (u_0, a^m(a^\dagger)^{n+2}u_0) - (u_0, a^{m+2}(a^\dagger)^n u_0)\}$$

$$= \frac{\hbar^2\alpha^2}{2\sqrt{m!}\sqrt{n!}}\{2(m+1)!\delta_{mn} - m!\delta_{mn} - m!\delta_{m,n+2} - (m+2)!\delta_{m+2,n}\}$$

$$= \begin{cases} \dfrac{(2m+1)\hbar^2\alpha^2}{2} & m = n \\ -\left[m(m-1)\right]^{1/2}\dfrac{\hbar^2\alpha^2}{2} & m = n+2 \\ -\left[m(m+1)\right]^{1/2}\dfrac{\hbar^2\alpha^2}{2} & m+2 = n \end{cases}$$

Example 6.17 Evaluate the uncertainty product $\Delta x\Delta p$ for an eigenstate u_n for a harmonic oscillator.

Solution

$$(\Delta x)^2 = (u_n, x^2u_n) - (u_n, xu_n)^2$$

$$(\Delta p)^2 = (u_n, p^2u_n) - (u_n, pu_n)^2$$

From the examples 6.13 – 6.16, we have

$$(u_n, xu_n) = 0$$

$$(u_n, x^2u_n) = \frac{(2n+1)}{2\alpha^2}$$

$$(u_n, pu_n) = 0$$

$$(u_n, p^2u_n) = (2n+1)\frac{\hbar^2\alpha^2}{2}$$

$$\therefore\ (\Delta x)^2(\Delta p)^2 = (2n+1)^2\frac{\hbar^2}{4}$$

$$\therefore\ \Delta x\,\Delta p = (2n+1)\frac{\hbar}{2}$$

State Function in Coordinate Representation from Ladder Operator Method

Let us start with the ground state wave function.

$$a\,u_0(x) = 0$$

In the coordinate representation

$$x_{op} = x \quad \text{and} \quad p_{op} = -i\hbar \frac{d}{dx}$$

$$\therefore a = \left(\frac{m\omega}{2\hbar}\right)^{1/2} x + \frac{i}{(2\hbar m\omega)^{1/2}}\left(-i\hbar \frac{d}{dx}\right)$$

$$= \frac{1}{\sqrt{2}}\alpha x + \frac{1}{\sqrt{2}\alpha}\frac{d}{dx} = \frac{1}{\sqrt{2}}\left(\frac{d}{d\rho} + \rho\right)$$

where $\rho = \alpha x$

$$\therefore \left(\frac{d}{d\rho} + \rho\right) u_0(\rho) = 0$$

$$\frac{du_0}{d\rho} = -\rho u_0 \quad \text{or} \quad \frac{du_0}{u_0} = -\rho d\rho$$

On integrating, $\ln \frac{u_0}{N_0} = -\frac{\rho^2}{2}$

$$\therefore u_0 = N_0 e^{-\frac{\rho^2}{2}}$$

The operator $a^\dagger$ is given by

$$a^\dagger = \left(\frac{m\omega}{2\hbar}\right)^{1/2} x_{op} - \frac{i}{(2\hbar m\omega)^{1/2}} p_{op} = \frac{1}{\sqrt{2}}\left(\rho - \frac{d}{d\rho}\right)$$

The higher energy states are given by

$$u_n = \frac{1}{\sqrt{n!}}(a^\dagger)^n u_0 = \frac{1}{\sqrt{n!}}\frac{1}{2^{n/2}}\left(\rho - \frac{d}{d\rho}\right)^n u_0. \tag{6.48}$$

Consider $$\frac{d}{d\rho}\left(e^{-\frac{\rho^2}{2}}\phi\right)$$

$$\frac{d}{d\rho}\left(e^{-\frac{\rho^2}{2}}\phi\right) = e^{-\frac{\rho^2}{2}}\left(-\rho + \frac{d}{d\rho}\right)\phi = e^{-\frac{\rho^2}{2}} \underbrace{(-1)\left(\frac{d}{d\rho} - \rho\right)\phi}_{\psi} = e^{-\frac{\rho^2}{2}}\psi$$

Next, consider $$\frac{d^2}{d\rho^2}\left(e^{-\frac{\rho^2}{2}}\phi\right)$$

$$\frac{d^2}{d\rho^2}\left(e^{-\frac{\rho^2}{2}}\phi\right) = \frac{d}{d\rho}\cdot\frac{d}{d\rho}\left(e^{-\frac{\rho^2}{2}}\phi\right) = \frac{d}{d\rho}\left(e^{-\frac{\rho^2}{2}}\psi\right)$$

$$= e^{-\frac{\rho^2}{2}}.(-1)\left(\frac{d}{d\rho}-\rho\right)\psi = e^{-\frac{\rho^2}{2}}(-1)^2\left(\frac{d}{d\rho}-\rho\right)^2\phi$$

Extending this result further, we get

$$\frac{d^n}{d\rho^n}\left(e^{-\frac{\rho^2}{2}}\phi\right) = e^{-\frac{\rho^2}{2}}(-1)^n\left(\frac{d}{d\rho}-\rho\right)^n\phi.$$

$$\therefore \left(\frac{d}{d\rho}-\rho\right)^n\phi = e^{\frac{\rho^2}{2}}(-1)^n\frac{d^n}{d\rho^n}\left(e^{-\frac{\rho^2}{2}}\phi\right) = e^{-\frac{\rho^2}{2}}\cdot e^{\rho^2}(-1)^n\frac{d^n}{d\rho^n}\left(e^{-\frac{\rho^2}{2}}\phi\right)$$

Making use of this result in (6.48), we get

$$\begin{aligned}
u_n &= \frac{1}{\sqrt{n\,!}}\frac{1}{2^{n/2}}\cdot e^{-\frac{\rho^2}{2}}e^{\rho^2}(-1)^n\frac{d^n}{d\rho^n}\left(e^{-\frac{\rho^2}{2}}u_0\right)\\
&= \frac{N_0}{\sqrt{n\,!}}\frac{1}{2^{\frac{n}{2}}}e^{-\frac{\rho^2}{2}}(-1)^n e^{\rho^2}\frac{d^n}{d\rho^n}\left(e^{-\frac{\rho^2}{2}}\cdot e^{-\frac{\rho^2}{2}}\right)\\
&= \frac{N_0}{\sqrt{n\,!}}\cdot\frac{1}{2^{\frac{n}{2}}}e^{-\frac{\rho^2}{2}}\underbrace{(-1)^n e^{\rho^2}\frac{d^n}{d\rho^n}e^{-\rho^2}}_{H_n(\rho)}\\
&= \frac{N_0}{\sqrt{n\,!}}\frac{1}{2^{\frac{n}{2}}}e^{-\frac{\rho^2}{2}}H_n(\rho).
\end{aligned}\tag{6.49}$$

6.3 COHERENT STATE

Coherent state in the case of a harmonic oscillator is one of the important concepts in quantum mechanics. It was first worked by Schrödinger to illustrate that there exist states of harmonic oscillator which oscillates like classical harmonic oscillator.

The simplest definition of a coherent state of a simple harmonic oscillator is that it is an eigenstate of the annihilation operator.

i.e., $$a\phi = \alpha\phi$$

where a is the annihilation operator and α is the eigenvalue of a. So far, we have been using $(m\omega/\hbar)^{1/2}$ as α. In this section, we are not taking α as $(m\omega/\hbar)^{1/2}$. α is now the eigenvalue of the operator a and, in general, α is a complex number.

There is another definition of coherent states of a harmonic oscillator; it is a state with minimum uncertainty product, i.e., any wave packet for which $\Delta x\Delta p = \hbar/2$ represents a coherent state. It happens to be true that such a wave packet executes a simple harmonic motion.

In general, at introductory level, there is always an emphasis on the development of a wave packet whose time evolution resembles like the motion of a classical harmonic oscillator.

Change of Notation

Before proceeding further, let us mention the changes in the notation to be used in this section. Let us now define β as $\beta = (m\omega/\hbar)^{1/2}$. In the previous sections, the Greek letter α was used to represent this quantity. As mentioned earlier in this section, α is used as the eigenvalue of the annihilation operator a. So the harmonic oscillator wave function is given by

$$u_n(x) = \left(\frac{\beta}{\sqrt{\pi}\, 2^n\, n!}\right)^{1/2} e^{-\rho^2/2} H_n(\rho) \text{ where } \rho = \beta x \tag{6.50}$$

Classical Harmonic Oscillator

The classical harmonic oscillator is represented by

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

and

$$p(t) = -m\omega A \sin \omega t$$

So one will expect to find a wave packet for harmonic oscillator such that $\langle x(t)\rangle$ and $\langle p(t)\rangle$ are similar to the above functions. It is enough, even if the peak of the wave packet executes a simple harmonic motion.

The total energy of the classical harmonic oscillator is given by

$$E = \frac{p^2}{2m} + \frac{1}{2} m\omega^2 x^2$$

$$\text{or } (p^2 + m^2\omega^2 x^2) = 2mE$$

Dividing this equation by $2\hbar m\omega$,

$$\frac{(m^2\omega^2 x^2 + p^2)}{2\hbar m\omega} = \frac{2mE}{2\hbar m\omega} = \frac{E}{\hbar\omega}$$

This can be rewritten as

$$\frac{(m\omega x + ip)}{(2\hbar m\omega)^{1/2}} \cdot \frac{(m\omega x - ip)}{(2\hbar m\omega)^{1/2}} = \frac{E}{\hbar\omega} \tag{6.51}$$

It is immediately obvious, $\frac{m\omega x + ip}{(2\hbar m\omega)^{1/2}} = \left[\left(\frac{m\omega}{2\hbar}\right)^{1/2} x + \frac{i}{(2\hbar m\omega)^{1/2}} p\right]$ is the classical quantity corresponding to the quantum mechanical annihilation operator $a = \left[\left(\frac{m\omega}{2\hbar}\right)^{1/2} x_{op} + \frac{i}{(2\hbar m\omega)^{1/2}} p_{op}\right]$.

Let us now define α as

$$\alpha = \frac{m\omega x + ip}{(2\hbar m\omega)} \tag{6.52}$$

where x and p are numbers. Therefore, α is a number. We have the following correspondence between classical physics.

Classical physics	**Quantum mechanics**
x	x_{op}
p	p_{op}
$\frac{m\omega x + ip}{(2\hbar m\omega)^{1/2}}$	$a = \left[\left(\frac{m\omega}{2\hbar}\right)^{1/2} x_{op} + \frac{i}{(2\hbar m\omega)^{1/2}} p_{op}\right]$

It is interesting to see the simple harmonic motion in phase space. If one plots $p(t)$ against $m\omega x(t)$, the simple harmonic motion is represented by a circle in phase space. On the other hand, we can go to a complex plane where $m\omega x$ is along x axis and p is along imaginary axis. In this complex plane, $\alpha(t) = \alpha = (m\omega x(t) + ip(t))/(2\hbar m\omega)$ traces a circle as time progresses.

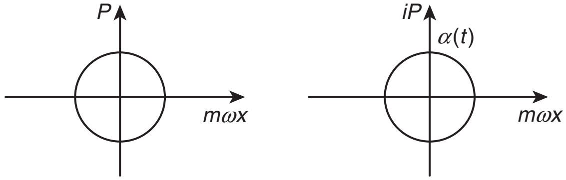


Fig. 6.4 Trajectory in phase space corresponding to simple harmonic motion

So, in quantum mechanics, for a system in the eigenstate of the annihilation operator a, we would expect the expectation value of x and p, to exhibit a classical harmonic motion. Therefore, the immediate task is to find the eigenstate of the quantum mechanical operator $a = ((m\omega/2\hbar)^{1/2} x_{op}) + (1/(2\hbar m\omega)^{1/2} ip_{op})$ and analyze its properties.

Before proceeding further, let us study wave packet corresponding to the uncertainty product $\Delta x\, \Delta p$ again.

Minimum Uncertainty Wave Packet and Coherent State

In the derivation of the uncertainty relation $\Delta x \Delta p \geq \hbar/2$, we define two operators $U = x_{op} - \langle x \rangle$ and $V = p_{op} - \langle p \rangle$, and the wave packet corresponding to minimum uncertainty product is given by

$$(U + i\lambda V)\,\phi(x) = 0 \tag{6.53}$$

where
$$\lambda = \frac{-i\langle [x, p] \rangle}{2(\Delta p)^2} = \frac{\hbar}{2(\Delta p)^2} = \frac{2(\Delta x)^2}{\hbar} \tag{6.54}$$

Let x_0 and p_0 be the expectation value of x and p, i.e., $\langle x \rangle = x_0$ and $\langle p \rangle = p_0$. Then, from (6.53), we get

$$(x - x_0)\phi(x) + i\frac{2(\Delta x)^2}{\hbar}(p - p_0)\phi(x) = 0$$

$$\left(x_{op} + i\frac{2(\Delta x)^2}{\hbar}p_{op}\right)\phi(x) = \left(x_0 + i\frac{2(\Delta x)^2}{\hbar}p_0\right)\phi(x) \tag{6.55}$$

Let us choose a particular width Δx, namely

$$(\Delta x)^2 = \frac{1}{2\beta^2} = \frac{\hbar}{2m\omega}$$

This is the width of the ground state of harmonic oscillator. Substituting this expression in the Equation (6.55), we get

$$\left(x_{op} + \frac{i}{m\omega} p_{op}\right)\phi = \left(x_0 + \frac{i}{m\omega} p_0\right)\phi$$

Multiplying both sides by $\left(m\omega/2\hbar\right)^{1/2}$ we get

$$\left[\left(\frac{m\omega}{2\hbar}\right)^{1/2} x_{op} + \frac{i}{(2\hbar m\omega)^{1/2}} p_{op}\right]\phi(x) = \left[\left(\frac{m\omega}{2\hbar}\right)^{1/2} x_0 + \frac{i}{(2\hbar m\omega)^{1/2}} p_0\right]\phi(x)$$

Or $$a\phi(x) = \alpha\phi(x) \tag{6.56}$$

So the wave packet corresponding to minimum uncertainty product and width $2\Delta x = 2(\hbar/m\omega)^{1/2}$ is the eigenstate of the annihilation operator and hence such a wave packet represents a coherent state. It is to be noted that all such wave packets (and hence all the eigenstates of the annihilation operator) have the same width $(\Delta x)^2 = 1/2\beta^2 = \hbar/2m\omega$ (as that of the ground state of harmonic oscillator).

Ground State as Coherent State

Let us recall that $\Delta x\Delta p = \hbar/2$ for ground state u_0. It is a state corresponding to minimum uncertainty product. Therefore, ground state $u_0 = (\beta/\sqrt{\pi})^{1/2} e^{-(\beta^2 x^2)/2}$ is a coherent state. What about other stationary energy eigenstates? We have already worked out $\Delta x\ \Delta p$ for an arbitrary energy eigenstate u_n, and it is found to be

$$\Delta x\ \Delta p = \left(n + \frac{1}{2}\right)\hbar$$

Obviously, none of the other energy eigenstates satisfies the minimum uncertainty product relation. Hence, it is not possible to use them as coherent states.

So if one constructs a coherent state which is distinct from the ground state, it will not be an energy eigenstate. Since all the energy eigenstates form a complete set, the coherent state can be expressed as a superposition of all the energy eigenstates.

Eigenstates of Annihilation Operator

The eigenvalue equation of the coherent state is

$$a\phi(x) = \alpha\,\phi(x)$$

Note that $a \neq a^\dagger$. Therefore, a is not a Hermitian operator, and hence its eigenvalue can be a real or a complex number. So it is convenient to write the eigenvalue equation as

$$a\phi(x) = \alpha e^{i\theta}\phi(x) \tag{6.57}$$

Our interest is two–fold now: one is the determination of $\phi(x)$ for a given $\alpha e^{i\theta}$ and the other is to determine how this state evolves in time for a harmonic oscillator. In other words, let us choose $\psi(x, 0)$ as

$$\psi(x, 0) = \phi(x) \tag{6.58}$$

The time evolution of this state is given by

$$\psi(x,t) = e^{-\frac{iHt}{\hbar}}\psi(x,0) = e^{-\frac{iHt}{\hbar}}\phi(x) \tag{6.59}$$

So we have to determine (i) $\phi(x)$ and (ii) $\psi(x, t)$.

Determination of $\phi(x)$

The eigenvalue equation for the annihilation operator is

$$a\phi(x) = \alpha e^{i\theta}\phi(x) \tag{6.60}$$

To make a correspondence with the classical harmonic oscillator, $\alpha e^{i\theta}$ may be written as

$$\alpha e^{i\theta} = \left(\frac{m\omega}{2\hbar}\right)^{1/2} x_0 + \frac{ip_0}{(2\hbar m\omega)^{1/2}} \tag{6.61}$$

At classical level, this represents a simple harmonic motion described by

$$x(t) = X\cos(\omega t - \theta) \tag{6.62}$$

where $$X = (x_0^2 + p_0^2/m^2\omega^2)^{1/2}$$

This represents a vibrating particle which at $t = 0$ is at position x_0 with momentum p_0.

i.e., at $$t = 0,\ x(0) = x_0 = X\cos\theta \tag{6.63}$$

and $$p(0) = p_0 = m\omega X\sin\theta \tag{6.64}$$

From (6.63) and (6.64) A simple algebra will give the relation $X = (x_0^2 + p_0^2/m^2\omega^2)^{1/2}$. For such a classical harmonic oscillator from (6.61), (6.62), (6.63) and (6.64), we get the following relations which will be useful later.

$$\sqrt{2}\alpha = \left(\frac{m\omega}{\hbar}\right)^{1/2} X = \beta X \tag{6.65}$$

$$\sqrt{2}\alpha\cos\theta = \beta X\cos\theta = \beta x_0 \tag{6.66}$$

$$\sqrt{2}\alpha\sin\theta = \beta X\sin\theta = \frac{p_0}{\hbar\beta} \tag{6.67}$$

$$\sqrt{2}\alpha\cos(\omega t - \theta) = \beta X\cos(\omega t - \theta) \tag{6.68}$$

$$\sqrt{2}\alpha\sin(\omega t - \theta) = \beta X\sin(\omega t - \theta) = -\frac{p(t)}{\hbar\beta} \tag{6.69}$$

Since the energy eigenstates $u_n(x)$ of a harmonic oscillator form a complete set, they can serve as a basis functions to expand $\phi(x)$. Therefore, let us write

$$\phi(x) = \sum_n C_n u_n(x), \text{with} \sum_n |C_n|^2 = 1$$

Hence $u_n(x)$ is given by

$$u_n(x) = N_n e^{-\beta^2 x^2/2} H_n(\beta x)$$

$$a\phi(x) = \sum_{n=0}^{\infty} C_n a u_n(x) = \sum_{n=0}^{\infty} C_n \sqrt{n}\; u_{n-1}(x) = \sum_{n=1}^{\infty} C_n \sqrt{n}\; u_{n-1}(x)$$

Feeding this expression in Equation (6.60), we get

$$\left[\sum_{n=1}^{\infty} C_n \sqrt{n}\; u_{n-1} - \alpha e^{i\theta} \sum_{n=0} C_n u_n \right] = 0 \tag{6.70}$$

The first term can be rewritten as follows:

$$\sum_{n=1}^{\infty} C_n \sqrt{n}\; u_{n-1} = \sum_{n=0}^{\infty} C_{n+1} \sqrt{n+1} u_n$$

Therefore, the Equation (6.70) becomes

$$\sum_{n=0}^{\infty} \left[C_{n+1} \sqrt{n+1} - C_n \alpha e^{i\theta} \right] u_n = 0$$

Since the functions $u_n s$ are linearly independent, we can equate the coefficient of each term to zero.

$$\therefore C_{n+1} = \frac{\alpha e^{i\theta}}{\sqrt{n+1}} C_n$$

Obviously C_n, corresponding to this recursion relation, is given by

$$C_n = \frac{\alpha^n e^{in\theta}}{\sqrt{n!}} C_0 \tag{6.71}$$

Since $C_n s$ are the coefficients of the basis functions in a linear combination, we have

$$\sum_n |C_n|^2 = 1$$

$$\therefore \sum \frac{\alpha^{2n}}{n!} |C_0|^2 = |C_0|^2 \sum \frac{\alpha^{2n}}{n!} = |C_0|^2 e^{\alpha^2} = 1$$

$$\therefore |C_0|^2 = e^{-\alpha^2} \text{or } C_0 = e^{-\frac{\alpha^2}{2}}$$

$$\therefore \phi(x) = e^{-\frac{\alpha^2}{2}} \sum_{n=0}^{\infty} \frac{(\alpha e^{i\theta})^n}{\sqrt{n!}} u_n(x) \tag{6.72}$$

This can be rewritten as follows:

$$\begin{aligned} \phi(x) &= e^{-\frac{\alpha^2}{2}} \sum_{n=0}^{\infty} \frac{\left(\alpha e^{i\theta}\right)^n}{\sqrt{n!}} \cdot \left(\frac{\beta}{2^n \sqrt{\pi}\; n!} \right)^{1/2} e^{-\frac{\rho^2}{2}} H_n(\rho) \\ &= e^{-\frac{\alpha^2}{2}} e^{-\frac{\rho^2}{2}} \left(\frac{\beta}{\sqrt{\pi}} \right)^{1/2} \sum_{n=0}^{\infty} \left(\frac{\alpha e^{i\theta}}{\sqrt{2}} \right)^n \cdot \frac{1}{n!} H_n(\rho) \end{aligned} \tag{6.73}$$

We know the generating function for Hermite's polynomials.

$$G(\rho,t) = e^{-t^2+2t\rho} = \sum_n \frac{t^n}{n!} H_n(\rho) \tag{6.74}$$

Comparing (6.73) and (6.74), $\left(\text{taking } \dfrac{\alpha e^{i\theta}}{\sqrt{2}} \text{ as } t\right)$, we get

$$\phi(x) = \exp\left(-\frac{\alpha^2}{2} - \frac{\rho^2}{2}\right)\left(\frac{\beta}{\sqrt{\pi}}\right)^2 \exp\left(-\left(\frac{\alpha}{\sqrt{2}} e^{i\theta}\right)^2 + 2\frac{\alpha}{\sqrt{2}} e^{i\theta}\rho\right)$$

$$= \left(\frac{\beta}{\sqrt{\pi}}\right)^{1/2} \exp\left[-\frac{\rho^2}{2} - \frac{\alpha^2}{2} - \frac{\alpha^2}{2} e^{2i\theta} + \sqrt{2}\alpha e^{i\theta}\rho\right]$$

$$= \left(\frac{\beta}{\sqrt{\pi}}\right)^{1/2} \exp\left[-\frac{\rho^2}{2} - \frac{\alpha^2}{2} - \frac{\alpha^2}{2}(\cos 2\theta + i\sin 2\theta) + \sqrt{2}\alpha(\cos\theta + i\sin\theta)\rho\right]$$

Making use of the relation $1 + \cos 2\theta = 2\cos^2\theta$, we get

$$\phi(x) = \left(\frac{\beta}{\pi}\right)^{1/2} \exp\left[-\frac{\rho^2}{2} - \alpha^2\cos^2\theta + \sqrt{2}\ \alpha\rho\cos\theta\right] \exp\left(-\frac{i}{2}\alpha^2\sin 2\theta\right) \exp(i\sqrt{2}\,\alpha\,\rho\,\sin\theta)$$

$$= \left(\frac{\beta}{\sqrt{\pi}}\right)^{1/2} \exp\left(-\frac{i}{2}\alpha^2\sin 2\theta\right).\exp(i\sqrt{2}\alpha\rho\sin\theta).\exp\left(\frac{-1}{2}(\rho - \sqrt{2}\alpha\cos\theta)^2\right)$$

Making use of the relations ((6.66) and (6.67)) $\sqrt{2}\alpha\cos\theta = \beta x_0$ and $\sqrt{2}\alpha\sin\theta = \dfrac{2p_0}{\hbar\beta}$,

we write $\phi(x)$ as

$$\phi(x) = \left(\frac{\beta}{\sqrt{\pi}}\right)^{1/2} \exp\left(-\frac{i}{2}\alpha^2\sin 2\theta\right)\exp\left(\frac{ip_0x}{\hbar}\right)\exp\left(-\frac{\beta^2}{2}(x - x_0)^2\right)$$

Leaving aside the phase factor, we can write $\phi(x)$ as

$$\phi(x) = \left(\frac{\beta}{\sqrt{\pi}}\right)^{1/2} \cdot \exp(ip_0x/\hbar)\exp\left(-\frac{\beta^2}{2}(x - x_0)^2\right) \tag{6.75}$$

This is Gaussian wave function centered around x_0; i.e., a Gaussian function shifted by a distance x_0 from the centre.

The probability density $|\phi(x)|^2$ is given by

$$|\phi(x)|^2 = \frac{\beta}{\sqrt{\pi}} e^{-\beta^2(x-x_0)^2} \tag{6.76}$$

This is the standard Gaussian distribution. It is almost the same as the probability density for ground state names $|u_0(x)|^2$ except for the fact that its peak is at x_0 and has the same width $2/\beta$ around x_0. So it is a Gaussian distribution displaced by x_0.

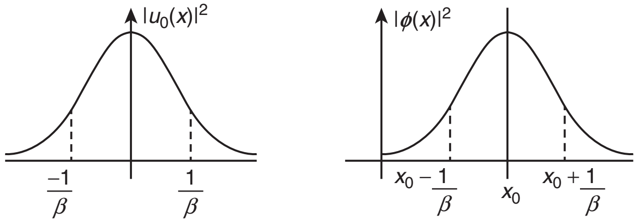


Fig. 6.5 Sketches of $\left|u_0(x)\right|^2$ and $\left|\phi(x)\right|^2$ for coherent state

Note that it is not energy eigenfunction for harmonic oscillator.

Time Evolution of the Coherent State

The time evolution of the coherent state is given by

$$\psi(x,t)=e^{-iHt/\hbar}\phi(x)$$

From Equation (6.69), we write $\psi(x, t)$ as

$$\psi(x,t)=e^{-iHt/\hbar}e^{-\alpha^{2}/2}\sum_{n=0}^{\infty}\frac{(\alpha e^{i\theta})^n}{\sqrt{n!}}u_n(x)$$

$$=e^{-\alpha^{2}/2}\sum_{n=0}^{\infty}\frac{(\alpha e^{i\theta})^n}{\sqrt{n!}}e^{-i\frac{Ht}{\hbar}}u_n(x)$$

But $$e^{-iHt/\hbar}u_n(x)=e^{-iE_nt/\hbar}u_n(x)=e^{-i\left(n+\frac{1}{2}\right)\omega t}u_n(x)=e^{-i\omega t/2}e^{-in\omega t}u_n(x)$$

$$\therefore\psi(x,t)=e^{-\alpha^{2}/2}\sum_{n=0}^{\infty}\frac{(\alpha e^{i\theta})^n}{\sqrt{n!}}e^{-i\omega t/2}e^{-in\omega t}u_n(x)$$

$$=\exp\left(-i\frac{\omega t}{2}\right)\left(\frac{\beta}{\sqrt{\pi}}\right)^{1/2}e^{-\alpha^{2}/2}\sum_{n=0}^{\infty}\left(\frac{\alpha e^{i\theta}e^{-i\omega t}}{\sqrt{2}}\right)^n\frac{1}{n!}e^{-\frac{\rho^2}{2}}H_n(\rho)$$

$$=\exp\left(-i\frac{\omega t}{2}\right)\left(\frac{\beta}{\sqrt{\pi}}\right)^{1/2}e^{-\frac{\alpha^2}{2}}e^{-\frac{\rho^2}{2}}G(\rho,t)\quad\text{with}\quad t=\frac{\alpha e^{i\theta}e^{-i\omega t}}{\sqrt{2}}$$

$$=\exp\left(-i\frac{\omega t}{2}\right)\left(\frac{\beta}{\sqrt{\pi}}\right)^{1/2}\exp\left[-\frac{\rho^2}{2}-\frac{\alpha^2}{2}-\frac{\alpha^2}{2}e^{-2i(\omega t-\theta)}+\sqrt{2}\alpha e^{-i(\omega t-\theta)}\rho\right]$$

$$=\left(\frac{\beta}{\sqrt{\pi}}\right)^{1/2}\exp\left(-i\frac{\omega t}{2}\right)\exp\left[i\frac{\alpha^2}{2}\sin 2(\omega t-\theta)-i\sqrt{2}\alpha\sin(\omega t-\theta)\rho\right]$$

$$\exp\left[-\frac{\rho^2}{2}-\frac{\alpha^2}{2}-\frac{\alpha^2}{2}\cos 2(\omega t-\theta)+\sqrt{2}\alpha\rho\cos(\omega t-\theta)\right]$$

$$= \left(\frac{\beta}{\sqrt{\pi}}\right)^{1/2} \exp\left[-i\left(\frac{\omega t}{2} - \frac{\alpha^2}{2}\sin 2(\omega t-\theta) + \sqrt{2}\alpha\rho\sin(\omega t-\theta)\right)\right].$$

$$\exp\left[-\tfrac{1}{2}(\rho-\sqrt{2}\alpha\cos(\omega t-\theta))^2\right]$$

But we know from (6.68) that $\sqrt{2}\alpha\cos(\omega t-\theta) = \beta X\cos(\omega t-\theta)$. It is β times the position $x(t)$ for a classical particle in simple harmonic motion. We also know that (from (6.69)) $-\rho\sqrt{\alpha}\sin(\omega t-\theta) = -\sqrt{\alpha}\beta x\sin(\omega t-\theta) = (p(t)x)/\hbar$ where $p(t)$ is the momentum of the classical particle which is executing simple harmonic motion.

$$\therefore \psi(x,t) = \left(\frac{\beta}{\sqrt{\pi}}\right)^{1/2} \exp\left[-i\left(\frac{\omega t}{2} - \frac{\alpha^2}{2}\sin 2(\omega t-\theta)\right)\right].\exp\left(i\frac{px}{\hbar}\right).\exp\left(-\tfrac{1}{2}(\rho-\beta X\cos(\omega t-\theta))^2\right)$$

Leaving aside the phase factors, we write $\psi(x, t)$ as

$$\therefore \psi(x,t) = \left(\frac{\beta}{\sqrt{\pi}}\right)^{1/2} \exp\left(i\frac{px}{\hbar}\right)\exp\left(-\tfrac{1}{2}(\rho-\beta X\cos(\omega t-\theta)\right)^2$$

$$\therefore \psi(x,t) = \left(\frac{\beta}{\sqrt{\pi}}\right)^{1/2} \exp\left(i\frac{px}{\hbar}\right)\exp\left[-\tfrac{1}{2}\beta^2(x-X\cos(\omega t-\theta))^2\right]$$

$$\therefore \psi(x,t) = \left(\frac{\beta}{\sqrt{\pi}}\right)^{1/2} \exp\left(i\frac{p_{cl}x}{\hbar}\right)\exp\left(-\tfrac{1}{2}\beta^2(x-x_{cl})\right)^2 \tag{6.77}$$

where x_{cl} is the classical position given by $x_{cl} = X\cos(\omega t - \theta)$ and p_{cl} is the classical momentum given by $p_{cl} = -m\omega x\sin(\omega t - \theta)$.

This wave packet literally behaves like a classical harmonic oscillator. It is a wave packet centred about $x_{cl}(t)$; i.e., the position of the peak executes a simple harmonic motion.

The probability density $|\psi(x,t)|^2$ is given by $|\psi(x,t)|^2 = \beta/\sqrt{\pi}\exp[-\beta^2(x-x_{cl}(t))^2]$. It is a Gaussian distribution with a mean at $x_{cl}(t) = x\cos(\omega t - \theta)$. i.e., the mean of the Gaussian distribution executes a simple harmonic motion. Thus, we have constructed a wave packet executing simple harmonic motion. The significance of this result can be realized only if we recognize the fact that the time evolution of this wave packet has been determined by the Hamiltonian of simple harmonic oscillator. Thus, we have succeeded in constructing a quantum state (which is the solution of time-dependent Schrödinger equation for harmonic oscillator) executing simple harmonic oscillations like a classical object. In addition, the product of uncertainties of position and momentum for this wave packet is $\hbar/2$. So we have succeeded in constructing a coherent state for simple harmonic oscillator.

Example 6.18 Show that $\Delta x\Delta p = \frac{\hbar}{2}$ for the coherent state ϕ defined by $a\phi = \alpha\phi$.

Solution: To calculate the uncertainty product $\Delta x\Delta p$, we make use of the following results:

$$(\psi, A\psi) = (A^\dagger\psi,\psi) \qquad (c\psi,\psi) = c^*(\psi,\psi) \qquad (\psi,c\psi) = c(\psi,\psi)$$

For harmonic oscillator, we have

$$x = \frac{1}{\sqrt{2}\beta}(a^\dagger + a) \qquad x^2 = \frac{1}{2\beta^2}(a^{\dagger 2} + a^2 + aa^\dagger + a^\dagger a) = \frac{1}{2\beta^2}(a^{\dagger 2} + a^2 + 2a^\dagger a + 1)$$

$$p = \frac{i\hbar\beta}{\sqrt{2}}(a^\dagger - a) \qquad p^2 = -\frac{\hbar^2\beta^2}{2}(a^{\dagger 2} + a^2 - a^\dagger a - aa^\dagger) = -\frac{\hbar^2\beta^2}{2}(a^{\dagger 2} + a^2 - 2a^\dagger a - 1)$$

$$\begin{aligned}\langle x\rangle &= \frac{1}{\sqrt{2}\beta}(\phi,(a^\dagger + a)\phi) = \frac{1}{\sqrt{2}\beta}[(\phi,a^\dagger\phi) + (\phi,a\phi)] = \frac{1}{\sqrt{2}\beta}[(a\phi,\phi) + (\phi,a\phi)]\\ &= \frac{1}{\sqrt{2}\beta}[\alpha^*(\phi,\phi) + \alpha(\phi,\phi)] = \frac{1}{\sqrt{2}\beta}[\alpha^* + \alpha]\end{aligned}$$

$$\begin{aligned}\langle x^2\rangle &= \frac{1}{2\beta^2}[(\phi,(a^{\dagger 2} + a^2 + 2a^\dagger a + 1)\phi) = \frac{1}{2\beta^2}[(\phi,a^{\dagger 2}\phi) + (\phi,a^2\phi) + (\phi,2a^\dagger a\phi) + (\phi,\phi)]\\ &= \frac{1}{2\beta^2}((\alpha^* + \alpha)^2 + 1)\end{aligned}$$

$$\langle p\rangle = \left(\phi, \frac{i\hbar\beta}{\sqrt{2}}(a^\dagger - a)\phi\right) = \frac{i\hbar\beta}{\sqrt{2}}(\alpha^* - \alpha)$$

$$\langle p^2\rangle = \left(\phi, -\frac{\hbar^2\beta^2}{2}(a^{\dagger 2} + a^2 - 2a^\dagger a - 1)\phi\right) = -\frac{\hbar^2\beta}{2}[(\alpha^* - \alpha)^2 - 1]$$

From the above calculations, we get

$$(\Delta x)^2 = \langle x^2\rangle - \langle x\rangle^2 = \frac{1}{2\beta^2} \quad \text{and} \quad (\Delta p)^2 = \langle p^2\rangle - \langle p\rangle^2 = \frac{\hbar^2\beta^2}{2}$$

$$\therefore (\Delta x)^2(\Delta p)^2 = \frac{\hbar^2}{4} \qquad \text{or} \qquad \Delta x\Delta p = \frac{\hbar}{2}$$

Example 6.19 Show that $e^{\lambda a^\dagger - \lambda^* a}u_0 = e^{-|\lambda|^2/2}e^{\lambda a^\dagger}u_0$

Solution: Making use of the result $e^{A+B} = e^A \cdot e^B \cdot e^{-[A,B]/2}$

$$e^{\lambda a^\dagger - \lambda^* a} = e^{\lambda a^\dagger} \cdot e^{-\lambda^* a} \cdot e^{-[\lambda a^\dagger, -\lambda^* a]/2}$$

$$[\lambda a^\dagger, -\lambda^* a] = -\lambda\lambda^*[a^\dagger, a] = \lambda\lambda^* = |\lambda|^2$$

$$\therefore e^{\lambda a^\dagger - \lambda^* a} = e^{-\frac{|\lambda|^2}{2}} e^{\lambda a^\dagger} e^{-\lambda^* a}$$

Note that

$$e^{-\lambda^* a}u_0 = \left(1 - \lambda^* a + \frac{(\lambda^*)^2 a^2}{2!} + \cdots\right)u_0 = u_0$$

$$\begin{aligned}\therefore e^{\lambda a^\dagger - \lambda^* a}u_0 &= e^{-\frac{|\lambda|^2}{2}} e^{\lambda a^\dagger} e^{-\lambda^* a}u_0\\ &= e^{-\frac{|\lambda|^2}{2}} e^{\lambda a^\dagger}u_0\end{aligned}$$

Example 6.20 Show that $\phi = e^{\lambda a^\dagger - \lambda^* a} u_0$ is an eigenstate the annihilation operator a with the eigenvalue λ.

Solution:

$$a\ \phi = a\ e^{\lambda a^\dagger - \lambda^* a} u_0$$

$$a\ e^{\lambda a^\dagger - \lambda^* a} u_0 = a\ e^{-\frac{|\lambda|^2}{2}} e^{\lambda a^\dagger} u_0$$

$$a\ e^{\lambda a^\dagger} u_0 = a\left[1 + \lambda a^\dagger + \lambda^2 \frac{(a^\dagger)^2}{2!} + \lambda^3 \frac{(a^\dagger)^3}{3!} + \cdots\right] u_0$$

$$= \lambda a a^\dagger u_0 + \lambda^2 \frac{a(a^\dagger)^2}{2!} u_0 + \lambda^3 \frac{a(a^\dagger)^3}{3!} u_0 + \cdots$$

But we know that $a(a^\dagger)^n = (a^\dagger)na + n(a^\dagger)^{n-1}$

$$\therefore a(a^\dagger)^n u_0 = n(a^\dagger)^{n-1} u_0$$

$$\therefore ae^{\lambda a^t} u_0 = \left[\lambda + \frac{2\lambda^2 a^\dagger}{2!} + \frac{3\lambda^3 (a^\dagger)^2}{3!} + \frac{4\lambda^4 (a^\dagger)^3}{4!} + \ldots\right] u_0$$

$$= \lambda\left[1 + \lambda a^\dagger + \frac{\lambda^2}{2!}(a^\dagger)^2 + \frac{\lambda^3}{3!}(a^\dagger)^3 + \ldots\right] u_0 = \lambda e^{\lambda a^\dagger} u_0$$

$$\therefore ae^{\lambda a^\dagger - \lambda^* a} u_0 = e^{-\frac{|\lambda^2|}{2}} \lambda e^{\lambda a^\dagger} u_0 = \lambda e^{\lambda a^\dagger - \lambda^* a} u_0 = \lambda \phi$$

or $a\phi = \lambda\phi$

EXERCISES

1. What are the boundary conditions on the wave functions for a simple harmonic oscillator?
2. Why should the energy eigenstate of a harmonic oscillator have a definite parity?
3. What is meant by zero point energy? How will you explain in terms of uncertainty principle?
4. What is the dimension of $\alpha = \left(\frac{m\omega}{\hbar}\right)^{1/2}$?
5. What is the dimension of $\rho = ax$?
6. How do you implement the boundary condition of finiteness of the wave function when it is expressed in terms of a series for a harmonic oscillator?
7. How does the polynomial $(C_0 + C_2\ \rho^2 + C_4\ \rho^4)$ behave as $\rho \to \infty$?
8. How does $e^{-\rho^2/2}(C_0 + C_2\rho^2 + C_4\rho^4)$ behave as $\rho \to \infty$?
9. Why do you think that the solution $\left(e^{-\frac{\rho^2}{2}} \sum_{s=0}^{\infty} C_s \rho^s\right)$ diverges to infinity while $\left(e^{-\frac{\rho^2}{2}} \sum_{s=0}^{n} C_s \rho^s\right)$ goes to zero as $\rho \to \infty$?
10. Obtain the series for $\lambda = 7$ using the recursion relation

    (i) $C_{s+2} = \dfrac{2s - (\lambda - 1)}{(s+2)(s+1)} C_s$

    (ii) $C_{s-2} = \dfrac{s(s-1)}{2(s-2) - (\lambda - 1)} C_s$

11. Normalize the energy eigenfunction $u(x)$ given by

$$u(x) = Ne^{-\alpha^2x^2/2}(4\alpha^2x^2 - 2)$$

12. Show that $u_0(x) = N_0e^{-\alpha^2x^2/2}$ and $u_1(x) = N_1e^{-\alpha^2x^2/2}$. $2x$ are orthogonal to each other.
13. Evaluate the matrix elements x_{12} and x_{21}.
14. Evaluate the matrix elements $(x^2)_{00}$, $(x^2)_{22}$, $(x^2)_{02}$ and $(x^2)_{20}$ of x_{op}. (Note $(x^2)_{\text{mn}} = \int u_m^* x^2 u_n dx$).
15. Evaluate the matrix elements p_{12} and p_{21} of the momentum operator p_{op}.
16. Obtain $u_2(x)$ using $(a^\dagger)^2/\sqrt{2}\, u_0((\alpha x))$, where $u_0(\alpha x) = \left(\alpha/\sqrt{\pi}\right)^{1/2} e^{(-\alpha^2x^2)/2}$.
17. Determine $[a, (a^\dagger)^3]$.
18. Determine

(i) $a^3 (a^\dagger)^3 u_0$
(ii) $a^5 (a^\dagger)^3 u_0$
(iii) $a^3 (a^\dagger)^5 u_0$

19. The wave function $\phi(x) = e^{-\alpha^2(x-x_0)^2/2}$ is written as $\phi(x) = \Sigma C_n u_n(x)$. Obtain the coefficients C_n.
20. What is meant by coherent state?
21. A ball in the simple pendulum executes a simple harmonic motion $x(t) = A \cos \omega t$. Determine its trajectory in phase space.
22. Explain how the classical variable a corresponding to the quantum mechanical operator $\dfrac{m\omega x + ip}{\sqrt{2m\hbar\omega}}$ (where x and p are numbers) represents the trajectory of a harmonic oscillator in phase space.
23. Obtain the coherent state wave packet corresponding a classical particle executing simple harmonic motion such that

(i) $x_{cl}(0) = 0,\ p_{cl}(0) = p_0$
(ii) $x_{cl}(0) = A$ and $p_{cl}(0) = 0$

REFERENCES

1. P. M. Mathew and A. Venkatesan, 1976. *A Text Book of Quantum Mechanics*. New Delhi: Tata McGraw–Hill Publishing Company.
2. B. H. Bransden and C. J. Joachain, 1990. *Introduction to Quantum Mechanics*. England: English Language Book Society/Longman.
3. Marcelo Alonso and Henry Valk, 1973. *Quantum Mechanics: Principles and Applications*. Reading, Massachusetts: Addison – Wesley Publishing Company.
4. Michael Fowler. http://galilo.phys.virginca.edu/classes/751.mf1.fa1102/coherentstates.htm.

7

Orbital Angular Momentum

The classical angular momentum of a particle at **r** of the momentum **p** is given by $\mathbf{L} = \mathbf{r} \times \mathbf{p}$. In classical physics, this definition takes care of all the situations like the orbital motion of earth around the sun, the spinning motion of the earth around its own axis and so on. In quantum mechanics, there exists a corresponding operator $\mathbf{L}_{op} = \mathbf{r}_{op} \times \mathbf{p}_{op}$. However, this definition of angular momentum is not sufficient to cover all the situations. Within quantum mechanics, there exists two kinds of angular momenta: orbital and spin angular momentum. Orbital angular momentum operator is defined as $\mathbf{L}_{op} = \mathbf{r}_{op} \times \mathbf{p}_{op}$. Spin angular momentum is defined later in Chapter 10. In this chapter let us concentrate on the properties of orbital angular momentum.

7.1 COMPATIBLE SET OF ANGULAR MOMENTUM OPERATORS

The classical orbital angular momentum **L** is given by

$$\mathbf{L} = \mathbf{r} \times \mathbf{p} = \hat{\mathbf{x}}(yp_z - zp_y) + \hat{\mathbf{y}}(zp_x - xp_z) + \hat{\mathbf{z}}(xp_y - yp_x) = \hat{\mathbf{x}}L_x + \hat{\mathbf{y}}L_y + \hat{\mathbf{z}}L_z \tag{7.1}$$

Quantum mechanically, the corresponding operators are given by

$$\begin{aligned}(L_x)_{op} &= y_{op}(p_z)_{op} - z_{op}(p_y)_{op}\\ (L_y)_{op} &= z_{op}(p_x)_{op} - x_{op}(p_z)_{op}\\ (L_z)_{op} &= x_{op}(p_y)_{op} - y_{op}(p_x)_{op}\end{aligned}$$

We will drop the subscript '*op*' hereafter. Therefore, the quantum mechanical operators are given by

$$L_x = i\hbar\left(z\frac{\partial}{\partial y} - y\frac{\partial}{\partial z}\right) \tag{7.2}$$

$$L_y = i\hbar\left(x\frac{\partial}{\partial z} - z\frac{\partial}{\partial x}\right) \tag{7.3}$$

$$L_z = i\hbar\left(y\frac{\partial}{\partial x} - x\frac{\partial}{\partial y}\right) \tag{7.4}$$

We have already shown in Chapter 5 that L_x, L_y and L_z satisfy the following commutation relations:

$$[L_x, L_y] = i\hbar\, L_z$$
$$[L_y, L_z] = i\hbar\, L_x$$
$$[L_z, L_x] = i\hbar\, L_y$$

Therefore, it can be easily checked that $[L^2, L_x] = 0, [L^2, L_y] = 0, [L^2, L_z] = 0$

Example 7.1 Show that $[L^2, L_z] = 0$

Solution:

$$\begin{aligned}[L^2, L_z] &= [L_x^2 + L_y^2 + L_z^2, L_z] = [L_x^2, L_z] + [L_y^2, L_z] \\ &= L_x[L_x, L_z] + [L_x, L_z]L_x + L_y[L_y, L_z] + [L_y, L_z]L_y \\ &= -i\hbar L_x L_y - i\hbar L_y L_x + i\hbar L_y L_x + i\hbar L_x L_y = 0\end{aligned}$$

So, of the four operators L^2, L_x, L_y and L_z, we have the following pair of commuting operators: $\{L^2, L_x\}, \{L^2, L_y\}$ and $\{L^2, L_z\}$. We can choose any particular pair from the above list. The standard choice of compatible observables are L^2 and L_z.

So our task is to find the eigenvalues and eigenfunctions common to both L^2 and L_z.

7.2 ORBITAL ANGULAR MOMENTUM OPERATORS IN SPHERICAL POLAR COORDINATES

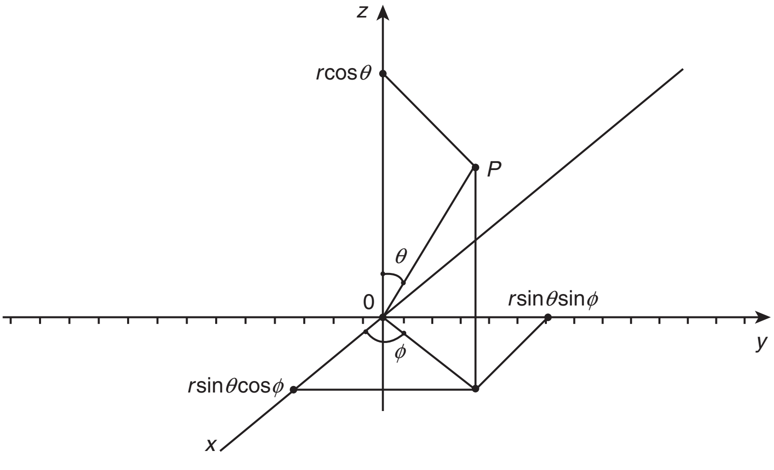


Fig. 7.1 Spherical polar coordinates

In spherical polar coordinate system, a point P is specified by three coordinates: r, θ, and ϕ as shown in Fig. 7.1. The coordinates r, θ and ϕ have the following ranges: $0 \le r \le \infty$, $0 \le \theta \le \pi$, $0 \le \phi \le 2\pi$.

The radius vector $\mathbf{r}$ is given by

$$\mathbf{r} = \mathbf{OP} = x\hat{\mathbf{x}} + y\hat{\mathbf{y}} + z\hat{\mathbf{z}}$$

We have $r^2 = x^2 + y^2 + z^2$.

From Fig. 7.1, it is obvious that

$$\mathbf{r} = \mathbf{OP} = r\sin\theta\cos\phi\hat{\mathbf{x}} + r\sin\theta\sin\phi\hat{\mathbf{y}} + r\cos\theta\hat{\mathbf{z}}$$

$$x = r\sin\theta\cos\phi \tag{7.5}$$

$$y = r\sin\theta\sin\phi \tag{7.6}$$

$$z = r\cos\theta \tag{7.7}$$

From the above equations, we get

$$r = (x^2 + y^2 + z^2)^{1/2} \tag{7.8}$$

$$\theta = \cos^{-1}\frac{z}{r} \tag{7.9}$$

and

$$\phi = \tan^{-1}\frac{y}{x} \tag{7.10}$$

The following relations are useful in obtaining the expressions for L_x, L_y and L_z in spherical polar coordinates.

From Equation (7.8) we get

$$\frac{\partial r}{\partial x} = \frac{x}{r}, \quad \frac{\partial r}{\partial y} = \frac{y}{r}, \quad \frac{\partial r}{\partial z} = \frac{z}{r} \tag{7.11}$$

From the Equation (7.9) we get

$$\frac{\partial\theta}{\partial x} = \frac{xz}{(x^2+y^2)^{1/2}r^2}, \quad \frac{\partial\theta}{\partial y} = \frac{yz}{(x^2+y^2)^{1/2}r^2}, \quad \frac{\partial\theta}{\partial z} = -\frac{(x^2+y^2)^{1/2}}{r^2} \tag{7.12}$$

From the Equation (7.10) we get

$$\frac{\partial\phi}{\partial x} = -\frac{y}{x^2+y^2}, \qquad \frac{\partial\phi}{\partial y} = \frac{x}{x^2+y^2} \tag{7.13}$$

L_x in spherical polar coordinates

$$\begin{aligned}
L_x &= i\hbar\left(z\frac{\partial}{\partial y} - y\frac{\partial}{\partial z}\right)\\
&= i\hbar\left\{z\left[\frac{\partial r}{\partial y}\cdot\frac{\partial}{\partial r} + \frac{\partial\theta}{\partial y}\cdot\frac{\partial}{\partial\theta} + \frac{\partial\phi}{\partial y}\cdot\frac{\partial}{\partial\phi}\right] - y\left[\frac{\partial r}{\partial z}\cdot\frac{\partial}{\partial r} + \frac{\partial\theta}{\partial z}\cdot\frac{\partial}{\partial\theta} + \frac{\partial\phi}{\partial z}\cdot\frac{\partial}{\partial\phi}\right]\right\}\\
&= i\hbar\left\{\left(z\frac{\partial r}{\partial y} - y\frac{\partial r}{\partial z}\right)\frac{\partial}{\partial r} + \left(z\frac{\partial\theta}{\partial y} - y\frac{\partial\theta}{\partial z}\right)\frac{\partial}{\partial\theta} + \left(z\frac{\partial\phi}{\partial y} - y\frac{\partial\phi}{\partial z}\right)\frac{\partial}{\partial\phi}\right\}\\
&= i\hbar\left[\frac{y}{(x^2+y^2)^{1/2}}\frac{\partial}{\partial\theta} + \frac{zx}{x^2+y^2}\frac{\partial}{\partial\phi}\right]\\
&= i\hbar\left[\sin\phi\frac{\partial}{\partial\theta} + \cot\theta\cos\phi\frac{\partial}{\partial\phi}\right]
\end{aligned} \tag{7.14}$$

L_y in spherical polar coordinates

$$
\begin{aligned}
L_y &= i\hbar\left(x\frac{\partial}{\partial z} - z\frac{\partial}{\partial x}\right)\\
&= i\hbar\left\{\left[x\frac{\partial r}{\partial z} - z\frac{\partial r}{\partial x}\right]\frac{\partial}{\partial r} + \left[x\frac{\partial \theta}{\partial z} - z\frac{\partial \theta}{\partial x}\right]\frac{\partial}{\partial \theta} + \left[x\frac{\partial \phi}{\partial z} - z\frac{\partial \phi}{\partial x}\right]\frac{\partial}{\partial \phi}\right\}\\
&= i\hbar\left[-\cos\phi\frac{\partial}{\partial\theta} + \cot\theta\sin\phi\frac{\partial}{\partial\phi}\right]
\end{aligned} \tag{7.15}
$$

L_z in spherical polar coordinates

$$
\begin{aligned}
L_z &= i\hbar\left(y\frac{\partial}{\partial x} - x\frac{\partial}{\partial y}\right)\\
&= i\hbar\left\{\left(y\frac{\partial r}{\partial x} - x\frac{\partial r}{\partial y}\right)\frac{\partial}{\partial r} + \left(y\frac{\partial \theta}{\partial x} - x\frac{\partial \theta}{\partial y}\right)\frac{\partial}{\partial \theta} + \left(y\frac{\partial \phi}{\partial x} - x\frac{\partial \phi}{\partial y}\right)\frac{\partial}{\partial \phi}\right\}\\
L_z &= -i\hbar\frac{\partial}{\partial\phi}
\end{aligned} \tag{7.16}
$$

L^2 in spherical polar coordinates

Let us first evaluate $L_x^2\psi$, where ψ is an arbitrary function of θ and ϕ. i.e., $\psi = \psi(\theta, \phi)$.

$$
\begin{aligned}
L_x^2\psi(\theta,\phi) &= L_x(L_x\psi)\\
&= -\hbar^2\left(\sin\phi\frac{\partial}{\partial\theta} + \cot\theta\cos\phi\frac{\partial}{\partial\phi}\right)\left(\sin\phi\frac{\partial\psi}{\partial\theta} + \cot\theta\cos\phi\frac{\partial\psi}{\partial\phi}\right)
\end{aligned}
$$

Similarly, $L_y^2\psi(\theta,\phi)$ is given by

$$
\begin{aligned}
L_y^2\psi(\theta,\phi) &= L_y(L_y\psi)\\
&= -\hbar^2\left(-\cos\phi\frac{\partial}{\partial\theta} + \cot\theta\sin\phi\frac{\partial}{\partial\phi}\right)\left(-\cos\phi\frac{\partial\psi}{\partial\theta} + \cot\theta\sin\phi\frac{\partial\psi}{\partial\phi}\right)\\
L_z^2\psi(\theta,\phi) &= -\hbar^2\frac{\partial^2\psi}{\partial\phi^2}
\end{aligned}
$$

$L^2\psi(\theta,\phi)$ is given by

$$
\begin{aligned}
L^2\psi(\theta,\phi) &= (L_x^2 + L_y^2 + L_z^2)\psi(\theta,\phi)\\
&= -\hbar\left(\frac{\partial^2\psi}{\partial\theta^2} + \frac{1}{\sin^2\theta}\frac{\partial^2\psi}{\partial\phi^2} + \cot\theta\frac{\partial\psi}{\partial\theta}\right)\\
&= -\hbar^2\left(\frac{\cos\theta}{\sin\theta}\frac{\partial\psi}{\partial\theta} + \frac{\partial^2\psi}{\partial\theta^2} + \frac{1}{\sin^2\theta}\frac{\partial^2\psi}{\partial\phi^2}\right)\\
&= -\hbar^2\left(\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\psi}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2\psi}{\partial\phi^2}\right)
\end{aligned}
$$

Since this equation is true for any arbitrary $\psi(\theta, \phi)$, the operator L^2 is given by

$$L^2 = -\hbar^2\left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right)+\frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2}\right] \tag{7.17}$$

7.3 PARITY

We have already seen what is meant by parity in one – dimensional problems. The same concept can be extended to wave functions in three dimensional space. The parity operator P is defined as

$$P\psi(\mathbf{r}) = \psi(-\mathbf{r})$$

or

$$P\psi(x,y,z) = \psi(-x,-y,-z) \tag{7.18}$$

The eigenvalue equation for the parity operator is

$$P\psi(\mathbf{r}) = \lambda\psi(\mathbf{r}) \tag{7.19}$$

It is easy to check that the eigenvalues λ are ±1.

Example 7.2 Determine the eigenvalues of the parity operator P which is defined as $P\psi(\mathbf{r}) = \psi(-\mathbf{r})$.

Solution: From the Equation (7.18), we get

$$P^2\psi(\mathbf{r}) = PP\psi(\mathbf{r}) = P\psi(-\mathbf{r}) = \psi(\mathbf{r}) \tag{7.20}$$

From the eigenvalue Equation (7.19), we get

$$P^2\psi(\mathbf{r}) = PP\psi(\mathbf{r}) = P\lambda\psi(\mathbf{r}) = \lambda P\psi(\mathbf{r}) = \lambda^2\psi(\mathbf{r})$$

Comparing the last two equations, we get

$$\lambda^2 = 1 \text{ or } \lambda = \pm 1 \tag{7.21}$$

The eigenfunction corresponding to $\lambda = 1$ is called even – parity function, and the eigenfunction corresponding to $\lambda = -1$ is called odd – parity function.

$$P\psi(\mathbf{r}) = \psi(-\mathbf{r}) = \psi(\mathbf{r}) \Rightarrow \text{even-parity wave function} \tag{7.22}$$

$$P\psi(\mathbf{r}) = \psi(-\mathbf{r}) = -\psi(\mathbf{r}) \Rightarrow \text{odd-parity wave function} \tag{7.23}$$

Example 7.3 Show that the parity operator P and the orbital angular momentum operator $\mathbf{L}$ commute.

Solution: Let us note that for an arbitrary wave function $\psi(\mathbf{r})$

$$\mathbf{L}\psi(\mathbf{r}) = -i\hbar\mathbf{r}\times\nabla\psi(\mathbf{r})$$

Consider the action of $(P\mathbf{L}-\mathbf{L}P)$ on an arbitrary wave function $\psi(\mathbf{r})$.

$$\begin{aligned}(P\mathbf{L}-\mathbf{L}P)\psi(\mathbf{r}) &= P(\mathbf{L}\psi(\mathbf{r})) - \mathbf{L}P\psi(\mathbf{r})\\ &= -i\hbar P[\mathbf{r}\times\nabla\psi(\mathbf{r})] + i\hbar\mathbf{r}\times\nabla\psi(-\mathbf{r})\\ &= -i\hbar(-\mathbf{r})\times(-\nabla\psi(-\mathbf{r})) + i\hbar\mathbf{r}\times\nabla\psi(-\mathbf{r})\\ &= -i\hbar\mathbf{r}\times\nabla\psi(-\mathbf{r}) + i\hbar\mathbf{r}\times\nabla\psi(-\mathbf{r}) = 0\end{aligned}$$

$$\therefore\ (P\,\mathbf{L}-\mathbf{L}P)\psi(\mathbf{r})=0$$

This equation is true for an arbitrary wave function $\psi(\mathbf{r})$.

Therefore, $P\,\mathbf{L}-\mathbf{L}P=0$

Or
$$[P,\mathbf{L}]=0 \tag{7.24}$$

Significance of [*P*, L] = 0

This commutation relation implies that the eigenfunction of angular momentum operator should also be the eigenfunction of the parity operator. In other words, the angular momentum eigenfunctions should have definite parity; i.e they should have either even parity or odd parity.

7.4 PARITY OPERATOR IN SPHERICAL POLAR COORDINATES

The parity transformation takes a point P to a new point P' where the coordinates of P are (x, y, z) and the coordinates of P' are $(x', y', z') = (-x, -y, -z)$. So, under parity transformation, we have

$$x\rightarrow x'=-x$$
$$y\rightarrow y'=-y$$
$$z\rightarrow z'=-z$$

In spherical polar coordinates, the coordinates of P are (r, θ, ϕ) and P' are (r', θ', ϕ'). What are the relationships between these two sets of coordinates?

i.e. , under parity transformation

$$r\rightarrow r'=?$$
$$\theta\rightarrow\theta'=?$$
$$\phi\rightarrow\phi'=?$$

We have to note the range of these coordinates $0\le r\le\infty$; $0\le\theta\le\pi$; $0\le\phi\le 2\pi$; $0\le r'\le\infty$; $0\le\theta'\le\pi$; $0\le\phi'\le 2\pi$.

It is obvious that $r'=((x')^2+(y')^2+(z')^2)^{1/2}=(x^2+y^2+z^2)^{1/2}=r$.

We have to work out the transformations for θ and ϕ.

$$x'=-x\Rightarrow\qquad r\sin\theta'\cos\phi'=-r\sin\theta\cos\phi \tag{7.25}$$

$$y'=-y\Rightarrow\qquad r\sin\theta'\sin\phi'=-r\sin\theta\sin\phi \tag{7.26}$$

$$z'=z\Rightarrow\qquad r\cos\theta'=-r\cos\theta \tag{7.27}$$

From (7.27), we get

$$\cos\theta'+\cos\theta=0\ \text{ or }\ 2\cos\frac{\theta'+\theta}{2}\cos\frac{\theta'-\theta}{2}=0$$

or $$\theta' = \pi - \theta \quad \text{or} \quad \theta' = \pi + \theta$$

Since, θ' is a number that has to be between 0 and π, we take θ' as

$$\theta' = \pi - \theta$$

Similarly the Equations (7.25) and (7.26) lead to

$$\cos\phi' + \cos\phi = 0$$

and $$\sin\phi' + \sin\phi = 0$$

$$2\cos\frac{\phi'+\phi}{2}\cos\frac{\phi'-\phi}{2} = 0 \Rightarrow \phi' = \pi + \phi \text{ or } \phi' = \pi - \phi$$

$$2\sin\frac{\phi'+\phi}{2}\cos\frac{\phi'-\phi}{2} = 0 \Rightarrow \phi' = \pi + \phi \text{ or } \phi' = -\phi$$

Obviously, only the solution $\phi' = \pi + \phi$ satisfies both equations.

Therefore, in spherical coordinate system the parity transformation

$$r \to r' = r \tag{7.28}$$

$$\theta \to \theta' = \pi - \theta \tag{7.29}$$

$$\phi \to \phi' = \pi + \phi \tag{7.30}$$

So, the parity operator in spherical polar coordinates is defined as

$$P\psi(r, \theta, \phi) = \psi(r, \pi - \theta, \pi + \phi) \tag{7.31}$$

Example 7.4 Prove that the angular momentum operator **L** and the parity operator **P** commute where the parity operator P is defined as $P\psi(r, \theta, \phi) = \psi(r, \pi - \theta, p + \phi)$.

Solution: Let us consider $(L_x P - PL_x)\ \psi(r, \theta, \phi)$

where $$\psi(r, \theta, \phi)$$

$$(L_x P - PL_x)\ \psi(r, \theta, \phi) = L_x(P\psi(r, \theta, \phi)) - P(L_x\psi(r, \theta, \phi))$$

$$= i\hbar\left[\sin\phi\frac{\partial}{\partial\theta} + \cot\theta\cos\phi\frac{\partial}{\partial\phi}\right]P\psi(r,\theta,\phi) - i\hbar P\left[\sin\phi\frac{\partial\psi(r,\theta,\phi)}{\partial\theta} + \cot\theta\cos\phi\frac{\partial\psi(r,\theta,\phi)}{\partial\phi}\right]$$

$$= i\hbar\left(\begin{array}{l} \sin\phi\dfrac{\partial}{\partial\theta}\psi(r,\pi-\theta,\pi+\phi) + \cot\theta\cos\phi\dfrac{\partial}{\partial\phi}\psi(r,\pi-\theta,\pi+\phi) - \sin(\pi+\phi)\dfrac{\partial}{\partial(\pi-\theta)} \\ \psi(r,\pi-\theta,\pi+\phi) - \cot(\pi-\theta)\cos(\pi+\phi)\dfrac{\partial}{\partial(\pi+\phi)}\psi(r,\pi-\theta,\pi+\phi) \end{array}\right)$$

$$= i\hbar\begin{pmatrix} \sin\phi\dfrac{\partial}{\partial\theta}\psi(r,\pi-\theta,\pi+\phi)+\cot\theta\cos\phi\dfrac{\partial}{\partial\phi}\psi(r,\pi-\theta,\pi+\phi)-\sin\phi\dfrac{\partial}{\partial\theta}\psi(r,\pi-\theta,\pi+\phi) \\ -\cot\theta\cos\phi\dfrac{\partial}{\partial\phi}\psi(r,\pi-\theta,\pi+\phi) \end{pmatrix}$$

$$= 0$$

i.e.,
$$(L_xP - PL_x)\ \psi(r,\ \theta,\ \phi) = 0$$

Since this equation is true for any arbitrary wave function $\psi(r,\ \theta,\ \phi)$, we have

$$[L_x, P] = 0$$

In the same way, we can prove that the parity operator P commutes with L_y and L_z.

7.5 EIGENVALUES OF L^2 AND L_z

Note that the dimension of $\mathbf{L}$ is $[L]\ [MLT^{-1}] = [M]\ [L^2]\ [T^{-2}]\ [T]$, which is the dimension of the product of energy and time, and so its unit is Joule.sec. Within quantum mechanics, we have one constant namely $\hbar$ whose unit is also Joule.sec. Therefore, we can write the eigenvalue of L_z as $m\hbar$ where m is a constant which has to be determined by solving the eigenvalue equation. Similarly, we can write the eigenvalue of L^2 as $\lambda\hbar^2$ where λ is a constant also which has to be also determined. So, we have

$$L^2u(\theta,\phi) = \lambda\hbar^2u(\theta,\phi) \tag{7.32}$$

$$L_zu(\theta,\phi) = m\hbar u(\theta,\phi) \tag{7.33}$$

or
$$-\hbar^2\left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right)+\frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2}\right]u(\theta,\phi) = \lambda\hbar^2u(\theta,\phi) \tag{7.34}$$

$$-i\hbar\frac{\partial u(\theta,\phi)}{\partial\phi} = m\hbar u(\theta,\phi) \tag{7.35}$$

The structure of these equations suggests that we can write

$$u(\theta,\phi) = \Theta(\theta)\ \Phi(\phi) \tag{7.36}$$

Substituting the Equation (7.36) in (7.35) , we get

$$-i\hbar\frac{d\Phi}{d\phi} = m\hbar\Phi \quad \text{or} \quad \Phi(\phi) = Ae^{im\phi} \tag{7.37}$$

Since ϕ is an angular variable, ϕ as well $\phi + 2\pi$ should represent the same angular position. Therefore the wave function at ϕ and at $\phi + 2\pi$ should be the same.

i.e.,
$$\Phi(\phi + 2\pi) = \Phi(\phi)$$

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

$$\therefore e^{im2\pi} = 1$$

or
$$m = 0, \pm 1, \pm 2, \pm 3\ldots \tag{7.38}$$

The eigenvalues of L_z are integer times $\hbar$. As it stands now, m can be any integer ranging from $-\infty$ to $+\infty$, i.e., there is no upper limit or lower limit for m. We get restrictions on the possible values of m only on solving the Equation (7.34).

Before proceeding further, let us determine the normalization constant A. The normalization constant A is determined by demanding

$$\int_0^{2\pi} \Phi^*\Phi d\phi = 1 \quad \text{or} \quad \int_0^{2\pi} |A|^2 d\phi = 1$$

Assuming A to be real number, we have

$$A^2 \cdot 2\pi = 1 \ \text{ or } \ A = \frac{1}{\sqrt{2\pi}}$$

$$\therefore \Phi(\phi) = \frac{1}{\sqrt{2\pi}} e^{im\phi}$$

or more appropriately

$$\Phi_m(\phi) = \frac{1}{\sqrt{2\pi}} e^{im\phi} \tag{7.39}$$

Example 7.5 Show that the eigenfunctions of L_z form an orthonormal set.

Solution: We have to show that $(\Phi_m, \Phi_n) = \delta_{mn}$.

$$\begin{aligned}(\Phi_m, \Phi_n) &= \int_0^{2\pi} \Phi_m^*(\phi)\Phi_n(\phi) d\phi \\ &= \frac{1}{2\pi}\int_0^{2\pi} e^{i(n-m)\phi} d\phi \\ &= \frac{1}{2\pi}\frac{e^{i(n-m)\phi}}{(n-m)}\Bigg|_0^{2\pi} \quad \text{provided } n \neq m \\ &= \frac{1}{2\pi}[e^{i(n-m)2\pi} - 1]\frac{1}{n-m} = 0 \quad \text{if } n \neq m\end{aligned}$$

If $n = m$, we know that

$$(\Phi_m, \Phi_m) = 1$$

$$(\Phi_m, \Phi_n) = \delta_{mn}$$

7.5.1 Solution $\Phi(\theta)$

Substituting the function (7.39) in (7.34), we get

$$-\hbar^2\left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\Theta}{\partial\theta}\right) - \frac{m^2}{\sin^2\theta}\Theta\right] = \lambda\hbar^2\Theta \tag{7.40}$$

Put $\cos\theta = \omega$

$$\therefore \frac{d}{d\omega}\left[(1-\omega^2)\frac{d\Theta}{d\omega}\right] + \left(\lambda - \frac{m^2}{1-\omega^2}\right)\Theta = 0 \quad -1 \le \omega \le 1 \tag{7.41}$$

Since $\cos\theta$ varies from -1 to $+1$, the range of ω is -1 to $+1$. Note, $\omega = \pm 1$ are singularities for the above equation.

Let us write $\Theta(\theta)$ as $(1-\omega^2)^a h(\omega)$ where a is a constant which can be chosen in a suitable way to suit our convenience. On substituting this expression in the Equation (7.41) we get a new equation for $h(\omega)$ containing a term with singularity and with proper choice of a we can remove this term.

Now with $\Theta(\omega) = (1-\omega^2)^a h(\omega)$,

$$\frac{d}{d\omega}\left[(1-\omega^2)\frac{d\Theta}{d\omega}\right] = (1-\omega^2)^{a+1}\frac{d^2h}{d\rho^2} - 2\omega(2a+1)\frac{dh}{d\rho}(1-\omega^2)^a$$
$$+4\omega^2a^2(1-\omega^2)^{a-1}h - 2a(1-\omega^2)^a h$$

Substituting this result in and making use of the result $4a^2\omega^2 = -4a^2(1-\omega^2) + 4a^2$ in (7.41) we get

$$(1-\omega^2)^{a+1}\frac{d^2h}{d\omega^2} - 2\omega(2a+1)(1-\omega^2)^a\frac{dh}{d\omega} - 4a^2(1-\omega^2)(1-\omega^2)^{a-1}h$$
$$+4a^2(1-\omega^2)^{a-1}h - 2a(1-\omega^2)^a h + \lambda(1-\omega^2)^a h - m^2(1-\omega^2)^{a-1}h = 0$$

Dividing throughout by $(1-\omega^2)^{a+1}$ we get

$$(1-\omega^2)\frac{d^2h}{d\omega^2} - 2\omega(2a+1)\frac{dh}{d\omega} - [4a^2+2a-\lambda]h + \frac{4a^2-m^2}{(1-\omega^2)}h = 0 \tag{7.42}$$

The last term is singular at $\omega = \pm 1$. Therefore, this term should be avoided. We can remove this term by making $4a^2 = m^2$

i.e.
$$a = \pm\frac{|m|}{2}$$

$$\therefore \Theta(\omega) = (1-\omega^2)^{|m|/2}h(\omega) \qquad \text{or} \qquad \Theta(\omega) = (1-\omega^2)^{-|m|/2}h(\omega)$$
$$= (1-\omega^2)^{-|m|/2}h(\omega)$$

The second function is singular at $\omega = \pm 1$.

Therefore, let us choose $a = \frac{|m|}{2}$

$$\therefore \Theta(\omega) = (1-\omega^2)^{|m|/2}h(\omega) \tag{7.43}$$

Therefore, the differential equation for $h(\omega)$ becomes

$$(1-\omega^2)\frac{d^2h}{d\omega^2} - 2\omega(|m|+1)\frac{dh}{d\omega} - [m^2 - |m| - \lambda]h = 0 \tag{7.44}$$

Let us try the series solution

$$h(\omega) = \sum_{s=0}^{\infty} C_s\omega^s \qquad -1 \le \omega \le 1$$

Substituting the series solution in Equation (7.44) we get

$$\sum_{s=0}^{\infty} C_s s(s-1)\,\omega^{s-2} - \sum_{s=0}^{\infty} C_s s(s-1)\,\omega^s$$
$$-2(|m|+1)\sum_{s=0}^{\infty} C_s\omega^s - [m^2 + 2|m| - \lambda]\sum_{s=0}^{\infty} C_s\omega^s = 0$$

Let us the write the first term of the series as follows:

$$\sum_{s=0}^{\infty} C_s s(s-1)\omega^{s-2} = \sum_{s=2}^{\infty} C_s s(s-1)\omega^{s-2} = \sum_{s=0}^{\infty} C_{s+2}(s+2)(s+1)\omega^{s}$$

Then, we have

$$\sum_{s=0}^{\infty} C_{s+2}(s+2)(s+1)\omega^{s} - \sum C_s\omega^{s}[-\lambda + m^2 + |m| + 2|m|s + 2s + s^2 - s] = 0$$

Equating the coefficient of of ω^s to zero, we get

$$C_{s+2} = \frac{(s+|m|)(s+|m|+1)-\lambda}{(s+2)(s+1)}C_s \tag{7.45}$$

Let us study the asymptotic behaviour of this series. Let us evaluate $\underset{s\to\infty}{\mathrm{Lt}}\frac{C_s+2}{C_s}$.

For the recurrence relation (7.45), we have

$$\begin{aligned}\underset{s\to\infty}{\mathrm{Lt}}\frac{C_{s+2}}{C_s} &= \frac{s\left(1+\frac{|m|}{s}\right)s\left(1+\frac{|m|+1}{s}\right)}{s\left(1+\frac{2}{s}\right)s\left(1+\frac{1}{s}\right)}\\ &= \left(1+\frac{|m|}{s}\right)\left(1+\frac{|m|+1}{s}\right)\left(1-\frac{2}{s}+....\right)\left(1-\frac{1}{s}+...\right)\\ &= 1+\frac{2|m|-2}{s}\end{aligned} \tag{7.46}$$

Let us compare this result with the asymptotic behaviour of the series expansion $(1-\omega^2)^{-|m|}$. For positive m, we have

$$(1-\omega^2)^{-m} = 1 + m\omega^2 + \frac{m(m+1)}{2!}\omega^4 + \ldots\ldots\ldots\ldots. = \sum C_r\omega^r$$

where

$$C_r = \frac{m(m+1)\cdots\cdots\left(m+\frac{r}{2}-1\right)}{\left(\frac{r}{2}\right)!}$$

Then

$$\frac{C_{r+2}}{C_r} = \frac{m(m+1)........\left(m+\frac{r}{2}\right)}{\left(\frac{r}{2}+1\right)!}\cdot\frac{\frac{r}{2}!}{m(m+1).........\left(m+\frac{r}{2}-1\right)}$$

$$= \frac{m+\frac{r}{2}}{\frac{r}{2}+1} = \frac{\frac{r}{2}+1-1+m}{\frac{r}{2}+1} = 1+\frac{m-1}{\frac{r}{2}+1}$$

In the large r limit, $\frac{C_{r+2}}{C_r} = 1+\frac{2(m-1)}{r}$

This result is the same as the Equation (7.46). So the asymptotic behaviour of $h(\omega)$ as determined by the recursion relation (7.45) has the asymptotic behaviour similar to that of $(1-\omega^2)^{-m}$. Therefore, the asymptotic behaviour of $\Theta(\omega) = (1-\omega^2)^{m/2}h(\omega)$ is similar to $(1-\omega^2)^{-m/2}$ (for $m > 0$), which becomes infinite for $\omega = \pm 1$.

This will make $\Theta = (1-\omega^2)^{|m|/2}h(\omega)$ as unacceptable wave function in quantum mechanics. The best way to make $\Theta(\omega)$ finite–valued wave function for the entire range $-1 \le \omega \le 1$ is to make the infinite series generated by the recursion relation (7.45) as a finite series.

i.e., Θ should become

$$\Theta = \begin{cases} (1-\omega^2)^{|m|/2}(C_0 + C_2\omega^2 + ...C_n\omega^n) & \text{even parity} \\ (1-\omega^2)^{|m|/2}(C_1\omega + C_3\omega^3 + ...C_n\omega^n) & \text{odd parity} \end{cases}$$

i.e. $h(\omega)$ should be a polynomial in ω instead of an infinite series.

How do you reduce the infinite series for $h(\omega)$ as polynomial in ω?

Even–parity Solution

We would like to have an even–parity solution as follows:

$$h(\omega) = C_0 + C_2\omega^2 + \ldots + C_n\omega^n$$

This can be achieved by demanding

$$\lambda = (n+|m|)(n+|m|+1), \quad C_n \neq 0, \quad C_1 = 0$$

Odd–parity Solution

The odd parity-solution should be as follows:

$$h(\omega) = C_1\omega + C_3\omega^3 + \ldots + C_n\omega^{\mathrm{n}}$$

This can be achieved by demanding

$$\lambda = (n+|m|)(n+|m|+1), \quad C_n \neq 0, \quad C_0 = 0$$

Note in both conditions, n and $|m|$ are integers, and so λ is an integer. If we define

$$l = n+|m|, \text{ then } \lambda = l(l+1) \tag{7.47}$$

This implies that λ is always a positive integer. The Equation (7.47) can be used to get the allowed values for m. Since $l = n + |m|$, then

$$|m| = l - n \qquad n = 0,1,2,\ldots..$$

For given l, what are the possible values m and n?

$$\begin{aligned} |m| &= l && \text{for} \quad n = 0 \\ |m| &= l - 1 && \text{for} \quad n = 1 \end{aligned}$$

Continuing further, we get

$$|m| = 0 \qquad \text{for} \quad n = l$$

We cannot increase the value of n further, for it will make $|m|$ negative.

$$\therefore \qquad |m| = l,\ l-1,\ l-2, \ldots, 0$$

or

$$m = \pm l, \pm(l-1), \pm(l-2), \ldots, 0 \tag{7.48}$$

Therefore, the possible values of m for a given l are

$$m = l, l-1, l-2, \ldots, 1, 0, -1, \ldots, -(l-1), -l \tag{7.49}$$

Eigenvalues of Angular Momentum

From (7.47) and (7.49), we find that the eigenvalue of the square of the angular momentum is $l(l+1)\hbar^2$ and the eigenvalue of Z component angular momentum is $m\hbar$

The eigenvalue of L^2: $l(l+1)\hbar^2 \quad l = 0,1,2,3\ldots\ldots$

The eigenvalue of L_z: $m_l\hbar$ with m_l taking the values $l, l-1, l-2, \ldots.. -(l-1), -l$

7.6 ANGULAR MOMENTUM EIGEN FUNCTIONS

For $\lambda = l(l+1)$, the differential equation becomes

$$(1-\omega^2)\frac{d^2 h_{lm}}{d\omega^2} - 2(|m|+1)\frac{dh_{lm}}{d\omega} + [l(l+1) - m^2 - |m|]h_{lm} = 0 \tag{7.50}$$

7.6.1 Solution h_{lm} for $m \geq 0$

Let us use the notation $D^m f = \dfrac{d^m f}{d\omega^m}$

Consider the Equation (7.50) for $m = 0$.

$$(1-\omega^2)^{1/2}\frac{d^2 h_{l0}}{d\omega^2} - 2\omega\frac{dh_{l0}}{d\omega} + l(l+1)h_{l0} = 0 \tag{7.51}$$

This is the well–known Legendre differential equation, and so

$$h_{l0}(\omega) = P_l(\omega)$$

where $P_l(\omega)$ is the Legendre polynomial.

Let us write Equation (7.51) on

$$(1-\omega^2)\ D^2P_l - 2\omega\ DP_l + l(l+1)P_l = 0$$

Let us differentiate this equation m times. We use Leibnitz formula in this regard.

$$\begin{aligned} D^m[uv] &= {}^mC_0 uD^m v + {}^mC_1 Du\cdot D^{m-1}v + {}^mC_2 D^2uD^{m-2}v + {}^mC_m D^m u\cdot v \\ &= uD^m\ v + mDuD^{m-1}v + \frac{m(m-1)}{2!}D^2vD^{m-2}v\ldots \end{aligned}$$

$$\begin{aligned} \therefore D^m[(l-\omega^2)D^2P_l] &= (l-\omega^2)D^mD^2P_l + mD(1-\omega^2)D^{m-1}D^2P_l \\ &\quad + \frac{m(m-1)}{2}D^2(1-\omega^2)D^{m-2}D^2P_l \\ &= (1-\omega^2)D^2D^mP_l - 2m\omega\, D\,D^mP_l - m(m-1)D^mP_l \end{aligned}$$

$$D^m 2\omega DP_l = 2\omega\, D^mDP_l + mD(2\omega)D^{m-1}DP_l = 2\omega\, DD^mP_l + 2mD^mP_l$$

Making use of these results in the Equation (7.51), we get

$$(1-\omega^2)D^2D^mP_l - 2\omega(m+1)DD^mP_l + [l(l+1) - m(m+1)]D^mP_l = 0$$

Comparing this equation with (7.50), we get

$$h_{lm} = c.D^mP_l, \text{ where c is a constant.} \tag{7.52}$$

Let us now introduce the associated Legendre functions P_{lm} which are given by

$$P_{lm} = (1-\omega^2)^{m/2}D^m P_l \quad m > 0 \tag{7.53}$$

More explicitly, we have

$$P_{lm} = (1-\omega^2)^{|m|/2}D^mP_l \tag{7.54}$$

The Legendre polynomial P_l is given by

$$P_l(\omega) = \frac{1}{2^l l!}D^l(\omega^2-1)^l \tag{7.55}$$

$$\therefore P_{lm} = \frac{1}{2^l l!}(1-\omega^2)^{m/2}D^{l+m}(\omega^2-1)^l \tag{7.56}$$

Obviously, $\Theta(\theta)$ is constant times the associated Legendre function P_{lm}.

7.6.2 Solution for Negative Values of *m*

For positive values of m, the solution to the Equation (7.41) is found to be P_{lm} (ω) and given by

$$P_{lm}(\omega) = \frac{1}{2^l}\frac{l}{l!}(1-\omega^2)^{|m|/2} D^{l+|m|}(\omega^2-1)^l \tag{7.57}$$

What is the solution corresponding to negative values of m? The Equation (7.57) helps us to extend the solution to negative values of m. $P_{l-|m|}$ is defined as

$$P_{l-|m|} = \frac{1}{2^l}\cdot\frac{1}{l!}(1-\omega^2)^{-\frac{|m|}{2}} D^{l-|m|}(\omega^2-1)^l \tag{7.58}$$

It is also written as

$$P_{l-m} = \frac{1}{2^l}\cdot\frac{1}{l!}(1-\omega^2)^{-m/2} D^{l-m}(\omega^2-1)^l \qquad m>0 \tag{7.59}$$

$P_{l-|m|}$ in (7.58) and P_{l-m} in (7.59) refer to the same quantity.

For example, for $m = 1$, P_{lm} is given by

$$P_{l1}(\omega) = \frac{1}{2^l}\cdot\frac{1}{l!}(1-\omega^2)^{1/2} D^{l+1}(\omega^2-1)^l$$

and for $m = -1$, P_{l-1} is given by

$$P_{l-1} = \frac{1}{2^l}\cdot\frac{1}{l!}(1-\omega^2)^{-1/2} D^{l-1}(\omega^2-1)^l$$

Is P_{l-m} as defined in (7.59) a solution to the Equation (7.41)?

It can be shown that P_{l-m} is constant times P_{lm}. Therefore, P_{l-m} satisfies the Equation (7.41). In fact, P_{l-m} is given by

$$P_{l-m} = (-1)^m \frac{(l-m)!}{(l+m)!} P_{lm}(x) \qquad m>0 \tag{7.60}$$

(See Example 7.15)

7.6.3 Normalization of Θ_{lm}

The normalized Θ_{lm} is defined as

$$\Theta_{lm} = N_{lm}P_{lm}(\omega)$$

The normalization constant N_{lm} is determined by the condition

$$\int_{-1}^{1} \Theta_{lm}^{*}\Theta_{lm}\,d\omega = 1$$

The normalization constant N_{lm} is found to be

$$N_{lm} = \left[\frac{(2l+1)}{2} \cdot \frac{(l-m)!}{(l+m)!}\right]^{1/2} \tag{7.61}$$

The orthoganality relation for Θ_{lm} is given by

$$\int_{-1}^{1} \Theta_{lm}^{*} \Theta_{l'm} d\Omega = \delta_{ll'} \delta_{mm'} \tag{7.62}$$

7.7 SPHERICAL HARMONICS AS EIGENFUNCTIONS OF ANGULAR MOMENTUM

The full eigenfunction of the angular momentum operator is the product of $\Theta_{lm}(\theta)$ and $\Phi(\phi)$. Such functions are called spherical harmonics Y_{lm}, which are defined as

$$\begin{aligned} Y_{lm}(\theta,\phi) &= (-1)^m \Theta_{lm}(\theta)\Phi(\phi) \\ &= \left[\frac{(2l+1)(l-m)!}{2(l+m)!}\right]^{1/2} (-1)^m P_{lm}(\cos\theta) \cdot \frac{1}{\sqrt{2\pi}} e^{im\phi} \\ &= \left[\frac{(2l+1)(l-m)!}{4\pi(l+m)!}\right]^{1/2} (-1)^m P_{lm}(\cos\theta) e^{im\phi} \end{aligned} \tag{7.63}$$

The spherical harmonics obey the orthoganality relation.

$$\begin{aligned} &\int Y_{lm}^{*}(\theta,\phi) Y_{l'm'}(\theta,\phi)\, d\Omega \\ &= \int_0^{\pi} \Theta_{lm}(\theta,\phi)\ \Theta_{l'm'}(\theta,\phi) \sin\theta\ d\theta \int_0^{2\pi} \frac{1}{2\pi} e^{im\phi} e^{-im'\phi} \\ &= \int_0^{\pi} \Theta_{lm}(\theta,\phi)\ \Theta_{l'm'}(\theta,\phi) \sin\theta\ d\theta\ \delta_{mm'} \\ &= \delta_{ll'} \delta_{mm'} \end{aligned} \tag{7.64}$$

One of the important properties of spherical harmonics is

$$Y_{l-m}(\theta,\phi) = (-1)^m Y_{lm}^{*}(\theta,\phi)$$

Let us list some useful recursion relations for P_{lm} and Y_{lm}

$$(1-\omega^2)^{1/2} P_{lm} = \frac{1}{(2l+1)} P_{l+1,m+1} - \frac{1}{2l+1} P_{l-1,m+1} \tag{7.65}$$

$$\omega P_{lm} = \frac{(l-m+1)}{2l+1} P_{l+1,m} + \frac{(l+m)}{2l+1} P_{l-1,m} \tag{7.66}$$

$$\sin\theta\, e^{i\phi} Y_{lm} = -\left[\frac{(l+m+1)(l+m+2)}{(2l+1)(2l+3)}\right]^{1/2} Y_{l+1,m+1} + \left[\frac{(l-m)(l-m-1)}{(2l+1)(2l-1)}\right]^{1/2} Y_{l-1,m+1} \tag{7.67}$$

$$\cos\theta\; Y_{lm} = \left[\frac{(l+m+1)(l-m+1)}{(2l+1)(2l+3)}\right]^{1/2} Y_{l+1,m} + \left[\frac{(l+m)(l-m)}{(2l+1)(2l-1)}\right]^{1/2} Y_{l-1,m} \tag{7.68}$$

$$\sin\theta\, e^{-i\phi} Y_{lm} = \left[\frac{(l-m+1)(l-m+2)}{(2l+1)(2l-1)}\right]^{1/2} Y_{l-1,m-1} - \left[\frac{(l-m)(l-m-1)}{(2l+1)(2l-1)}\right]^{1/2} Y_{l-1,m-1} \tag{7.69}$$

7.8 PARITY OF $Y_{lm}(\theta, \phi)$

Let us consider the behaviour of the spherical harmonics under parity transformation.

Under parity transformation $Y_{lm}(\theta,\phi)$ changes to $Y_{lm}(\pi-\theta, \pi+\phi)$

$$\begin{aligned} Y_{lm}(\pi-\theta,\pi+\phi) &= \Theta_{lm}(\pi-\theta)(-1)^m \cdot \frac{1}{\sqrt{2\pi}} e^{im\pi+im\phi} \\ &= (-1)^{2m}\Theta_{lm}(\pi-\theta)\cdot \frac{1}{\sqrt{2\pi}} e^{im\phi} \end{aligned}$$

The transformation property of Θ_{lm} can be obtained by considering Equation (7.56).

Under the transformation $\quad \theta \to \pi-\theta \qquad \cos\theta \to -\cos\theta \quad$ or $\quad \omega \to -\omega$

$$\begin{aligned} &\frac{d^{l+m}}{d\omega^{l+m}} \to (-1)^{l+m} \frac{d^{l+m}}{d\omega^{l+m}} \\ &(1-\omega^2)^{1/2} \frac{d^{l+m}}{d\omega^{l+m}} (\omega^2-1)^l \to (-1)^{l+m}(1-\omega^2)^{1/2} \frac{d^{l+m}}{d\omega^{l+m}} (\omega^2-1)^l \\ &\therefore \Theta_{lm}(\pi-\theta) = (-1)^{l+m}\Theta_{lm}(\theta) \end{aligned}$$

$$\begin{aligned} Y_{lm}(\pi-\theta,\pi+\theta) &= (-1)^l(-1)^m\Theta_{lm}(\theta)\frac{1}{\sqrt{2\pi}} e^{im\phi} \\ &= (-1)^l Y_{lm}(\theta,\phi) \end{aligned} \tag{7.70}$$

Therefore, the parity of the orbital angular momentum wave function is determined by $(-1)^l$

So, for spherical harmonics, we have

Even l even parity (7.71)

Odd l odd parity (7.72)

7.9 SKETCHING OF SPHERICAL HARMONICS

List of Spherical Harmonics

l	m	Y_{lm}
0	0	$Y_{00} = \left(\frac{1}{4\pi}\right)^{1/2}$
1	1	$Y_{11} = \left(\frac{3}{8\pi}\right)^{1/2} \sin\theta\, e^{i\phi}$
1	0	$Y_{10} = \left(\frac{3}{4\pi}\right)^{1/2} \cos\theta$
1	–1	$Y_{1-1} = \left(\frac{3}{8\pi}\right)^{1/2} \sin\theta\, e^{-i\phi}$
2	2	$Y_{22} = \left(\frac{15}{32\pi}\right)^{1/2} \sin^2\theta\, e^{i2\phi}$
2	1	$Y_{21} = -\left(\frac{15}{8\pi}\right)^{1/2} \sin^2\theta \cos\theta\, e^{i\phi}$
2	0	$Y_{20} = \left(\frac{5}{16\pi}\right)^{1/2} (3\cos^2\theta - 1)$
2	–1	$Y_{2-1} = \left(\frac{15}{8\pi}\right)^{1/2} \sin\theta \cos\theta\, e^{-i\phi}$
2	–2	$Y_{2-2} = \left(\frac{15}{32\pi}\right)^{1/2} \sin^2\theta\, e^{-i2\phi}$

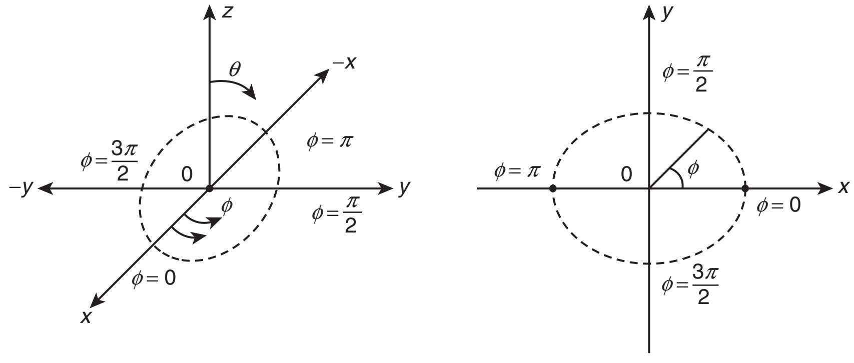


Fig. 7.2 Variation of θ and ϕ

$Y_{lm}(\theta,\phi)$ is a complex function in θ and ϕ. So, we can plot $|Y_{lm}(\theta,\phi)|^2$. Alternatively, real and imaginary part of $Y_{lm}(\theta,\phi)$ can be separately plotted.

θ is measured from z-axis, and ϕ varies in the horizontal *xy* plane as shown in Fig. 7.2. For a given θ and ϕ, along the radial line, mark a point whose length is $|Y_{lm}(\theta,\phi)|^2\,\theta$ as shown in Fig. 7.4.

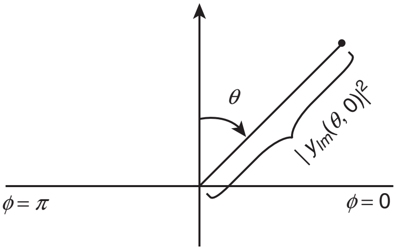


Fig. 7.3 $|Y_{lm}|^2$ in polar coordinates

Obviously for the entire range of θ and ϕ namely $0 \le \theta \le \pi$, and $0 \le \phi \le 2\pi$, all such points together form a two-dimensional surface in three-dimensional space. Normally, a section of this surface in the *zx* plane is shown. The entire surface can be generated by rotating this curve about *z*-axis.

Same procedure can be adopted for the sketch of real part and imaginary part of $Y_{lm}(\theta,\phi)$.

***l* = 0**

$$Y_{00}(\theta,\phi) = \frac{1}{\sqrt{4\pi}}$$

For all the values of θ and ϕ, length of $|Y_{00}(\theta,\phi)|^2$ is the same as $1/4\pi$. As θ varies from 0 to π, $|Y_{00}(\theta,\phi)|^2$ traces the right semicircle for $\phi = 0$ and $|Y_{00}(\theta,\phi)|^2$ traces left semicircle for $\phi = \pi$ as shown in Fig. 7.4.

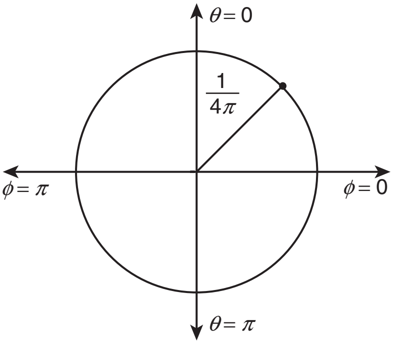


Fig. 7.4 Sketch of $|Y_{00}(\theta,\phi)|^2$

Obviously, as ϕ varies, this semicircle is rotated about *z* axis, and we get a spherical surface of radius $1/4\pi$.

***l* = 1, *m* = 0**

$$Y_{10} = \left(\frac{3}{4\pi}\right)^{1/2} \cos\theta \quad \text{or} \quad |Y_{10}|^2 = \frac{3}{4\pi}\cos^2\theta$$

Maximum value of $|Y_{10}|^2$ is $3/4\pi$ corresponding $\theta = 0$ and $\theta = \pi$. It slowly decreases as θ increases from 0 to $\pi/2$ And it becomes zero at $\theta = 0$ and subsequently increases to the maximum.

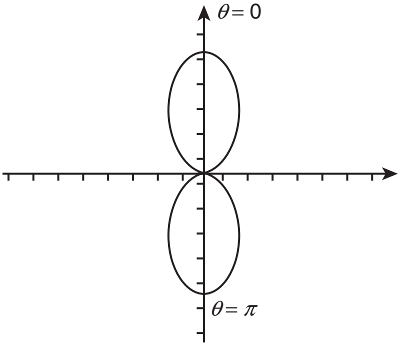


Fig. 7.5 Sketch of $|Y_{10}|^2$

***l* = 1, *m* = ±1**

$$|Y_{11}|^2 = |Y_{1-1}|^2 = \frac{3}{8\pi}\sin^2\theta$$

For $\theta = 0$, $|Y_{11}|^2 = 0$ and for $\theta = \frac{\pi}{2}, |Y_{11}|^2 = \frac{3}{8\pi}$. The sketch of $\left|Y_{11}\right|^2$ is as shown in Fig. 7.6

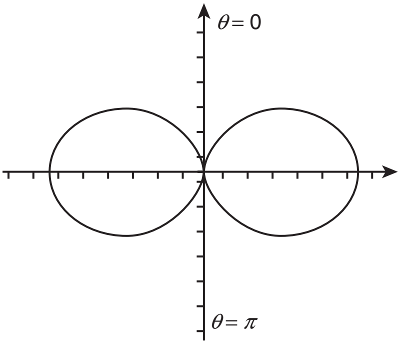


Fig. 7.6 Sketch of $|Y_{11}|^2$

***l* = 2, *m* = ±2**

$$|Y_{22}|^2 = |Y_{2-2}|^2 = \frac{15}{32\pi}\sin^4\theta$$

$|Y_{22}|^2$ is zero for $\theta = 0$ and $\theta = \pi$ and has a value of $\frac{15}{32\pi}$ at $\theta = \frac{\pi}{2}$. The sketch of $|Y_{22}|^2$ or $|Y_{2-2}|^2$ is shown in Fig. 7.7.

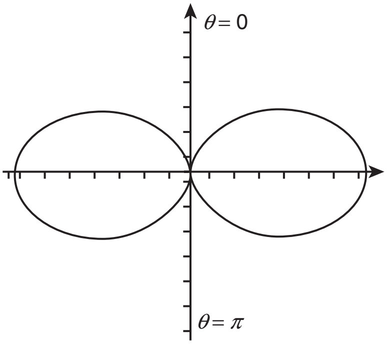


Fig. 7.7 Sketch of $|Y_{22}|^2$ or $|Y_{2-2}|^2$

***l* = 2, *m* = ±1**

$$|Y_{21}|^2 = |Y_{2-1}|^2 = 15/8\pi \sin^2\theta \cos^2\theta = 15/2\pi (\sin 2\theta)^2$$

$|Y_{21}|^2$ is zero for $\theta = 0$ and $\theta = \pi/2$. It has maximum value of $15/2\pi$ for $\theta = \pi/4$ and $\theta = 3\pi/4$.

So, at $\theta = 0$, $|Y_{21}|^2 = 0$. As θ increases, $|Y_{21}|^2$ increases and reaches a maximum at $\theta = \pi/4$ and subsequently decreases to become zero at $\theta = \pi/2$. Again, it increases to reach a maximum at $\theta = 3\pi/4$ and subsequently decreases to zero at $\theta = \pi$

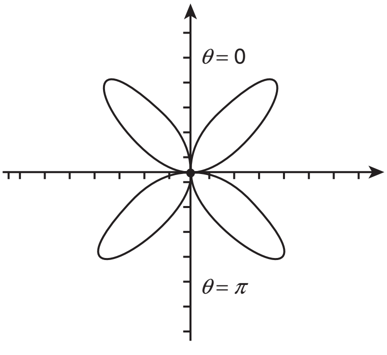


Fig. 7.8 Sketch of $|Y_{21}|^2$

***l* = 2, *m* = 0**

$$|Y_{20}|^2 = \frac{5}{16\pi}(3\cos^2\theta - 1)^2$$

The function has got maxima at $\theta = 0$, $\theta = \pi/2$ and $\theta = \pi$. The function is minimum for the value θ determined from the equation $3\cos^2\theta - 1 = 0$ or $\theta \simeq 55°$.

At $\theta = 0$, $|Y_{20}|^2 = 5/4\pi$. As θ increases, $|Y_{20}|^2$ decreases and becomes zero at $\theta = \tan^{-1} - 1/\sqrt{3} \simeq 55°$. Subsequently, it increases a maximum value of $5/16\pi$ at $\theta = 0$. It again decreases and becomes zero at $\theta = \tan^{-1} - 1/\sqrt{3}$ or $\theta \simeq 125°$. Subsequently it increases to $5/4\pi$ at $\theta = \pi$. This is sketched in Fig. 7.9.

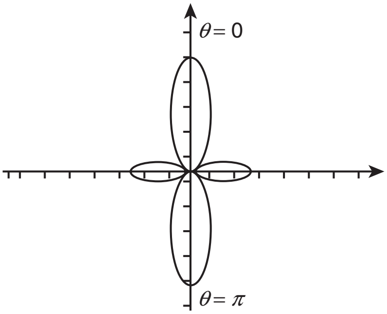


Fig. 7.9 Sketch of $|Y_{20}|^2$

7.10 RAISING AND LOWERING OPERATOR

Consider two operators L_+ and L_- given by

$$L_+ = L_x + iL_y \tag{7.73}$$

and

$$L_- = L_x - iL_y \tag{7.74}$$

These operators are known as raising and lowering operators for the following reasons. It can be shown that

$$L_+ Y_{lm}(\theta,\phi) = [(l-m)(l+m+1)]^{1/2} Y_{l,m+1} \tag{7.75}$$

and

$$L_- Y_{lm}(\theta,\phi) = [(l+m)(l-m+1)]^{1/2} Y_{l,m-1} \tag{7.76}$$

When L_+ acts on the state Y_{lm}, the resulting state is constant times $Y_{l\,m+1}$. i.e., it creates a new state whose m value has increased by 1. Therefore, L_+ is called the raising operator.

L_- acting on Y_{lm} produces a state proportional to $Y_{l\,m-1}$ (constant times $Y_{l\,m-1}$). i.e., it creates a state whose m value has decreased by 1. So this is called lowering operator.

Note that the maximum value of m is l. So, when L_+ acts on Y_{ll}, the result should be zero as can be seen from the Equation (7.75).

$$L_+ Y_{ll} = 0 \tag{7.77}$$

Similarly, the minimum value of m is $-l$. Therefore, when L_- acts on Y_{l-l} the result is zero as can be seen from the Equation (7.76).

$$L_- Y_{l\,-l} = 0 \tag{7.78}$$

7.10.1 L_+ and L_- in Spherical Polar Coordinates

$$\begin{aligned}
L_+ &= L_x + iL_y \\
&= i\hbar\left[\sin\phi\frac{\partial}{\partial\theta} + \cot\theta\cos\phi\frac{\partial}{\partial\phi} - i\cos\theta\sin\phi\frac{\partial}{\partial\theta} - i\cot\theta\sin\phi\frac{\partial}{\partial\phi}\right] \\
&= i\hbar e^{-i\phi}\left[-i\frac{\partial}{\partial\theta} + \cot\theta\frac{\partial}{\partial\phi}\right]
\end{aligned}$$

Put $\omega = \cos\theta$. Then, we have

$$L_+ = i\hbar e^{-i\phi}\left[i(1-\omega^2)^{1/2}\frac{\partial}{\partial\omega} + \frac{\omega}{(1-\omega^2)^{1/2}}\frac{\partial}{\partial\phi}\right] \tag{7.79}$$

$$\begin{aligned}
L_- &= L_x - iL_y \\
&= i\hbar\left[\sin\phi\frac{\partial}{\partial\theta} + \cot\theta\cos\phi\frac{\partial}{\partial\phi} + i\cos\phi\frac{\partial}{\partial\theta} - i\cot\theta\sin\phi\frac{\partial}{\partial\phi}\right] \\
&= i\hbar e^{-i\phi}\left[i\frac{\partial}{\partial\theta} + \cot\theta\frac{\partial}{\partial\phi}\right] \\
&= i\hbar e^{-i\phi}\left[-i\sin\theta\frac{\partial}{\partial\omega} + \cot\theta\frac{\partial}{\partial\phi}\right] \\
&= i\hbar e^{-i\phi}\left[-i(1-\omega^2)^{1/2}\frac{\partial}{\partial\omega} + \frac{\omega}{(1-\omega^2)^{1/2}}\frac{\partial}{\partial\phi}\right]
\end{aligned} \tag{7.80}$$

Example 7.6 Show that $L_+Y_{lm} = [(l-m)(l+m+1)]^{1/2}\hbar Y_{lm+1}$

Solution: Note that $Y_{lm}(\theta,\phi)$ is given by

$$Y_{lm}(\theta,\phi) = (-1)^m N_{lm}(1-\omega^2)^{m/2} D^{l+m}(\omega^2-1)^l \frac{1}{\sqrt{2\pi}} e^{im\phi}$$

$$\therefore L_+Y_{lm} = i\hbar\cdot(-1)^m \frac{N_{lm}e^{i\phi}}{\sqrt{2\pi}}\left[i(1-\omega^2)^{1/2}D + \frac{\omega}{(1-\omega^2)^{1/2}}\frac{\partial}{\partial\phi}\right](1-\omega^2)^{m/2}\cdot D^{l+m}(\omega^2-1)^l e^{im\phi}$$

$$\begin{aligned}
&= i\hbar\cdot(-1)^m \frac{N_{im}e^{i\phi}}{\sqrt{2\pi}}\Big\{i(1-\omega^2)^{1/2}\left[D(1-\omega^2)^{m/2}\right]\cdot D^{l+m}(\omega^2-1)^l e^{im\phi} \\
&\quad + i(1-\omega^2)^{1/2}(1-\omega^2)^{m/2}D^{l+m+1}(\omega^2-1)^l e^{im\phi} \\
&\quad + \frac{\omega}{(1-\omega^2)^{1/2}}im.(1-\omega^2)^{m/2}D^{l+m}(\omega^2-1)^l e^{im\phi}\Big\}
\end{aligned}$$

First and last terms cancel each other.

$$
\begin{aligned}
\therefore L_+Y_{lm} &= i\hbar(-1)^{m+1}\cdot(-1)\frac{N_{lm}}{\sqrt{2\pi}}i(1-\omega^2)^{(m+1)/2}D^{l+m+1}(\omega^2-1)^l e^{i(m+1)\phi} \\
&= \hbar\frac{N_{lm}}{N_{lm+1}}\cdot Y_{lm+!}(\theta,\phi) \\
&= \hbar\cdot\left[\frac{(2l+1)}{2}\cdot\frac{(l-m)!}{(l+m)!}\cdot\frac{2(l+m+1)!}{(l-m-1)!(2l+1)}\right]^{1/2} Y_{lm+1} \\
&= \hbar[(l-m)(l+m+1)]^{1/2}Y_{lm+1}
\end{aligned}
$$

Example 7.7 Show that $L\_Y_{lm} = [(l+m)(l-m+1)]^{1/2}Y_{lm}$

Solution:

$$
\begin{aligned}
L\_Y_{lm} &= i\hbar e^{-i\phi}\left[-i(1-\omega^2)^{1/2}D+\frac{\omega}{(1-\omega^2)^{1/2}}\frac{\partial}{\partial\phi}\right]Y_{lm} \\
&= i\hbar e^{-i\phi}(-1)^m\frac{N_{lm}}{\sqrt{2\pi}}\left[-i(1-\omega^2)^{1/2}D+\frac{\omega}{(1-\omega^2)^{1/2}}\frac{\partial}{\partial\phi}\right]\cdot(1-\omega^2)^{m/2}D^{l+m}(\omega^2-1)^l e^{im\phi} \\
&= i\hbar e^{-i\phi}(-1)^m\frac{N_{lm}}{\sqrt{2\pi}}\left\{\begin{array}{l}-i(1-\omega^2)^{1/2}(-m\omega)(1-\omega^2)^{\left(\frac{m}{2}-1\right)}D^{l+m}(\omega^2-1)^l e^{im\phi} \\ -i(1-\omega^2)^{1/2}(1-\omega^2)^{m/2}D^{l+m+1}(\omega^2-1)^l e^{im\phi} \\ +\dfrac{\omega}{(1-\omega^2)^{1/2}}im(1-\omega^2)^{m/2}D^{l+m}(\omega^2-1)^l e^{im\phi}\end{array}\right\} \\
&= i\hbar e^{i(m-1)\phi}(-1)^m\frac{N_{lm}\,i}{\sqrt{2\pi}}\left\{\frac{2m\omega}{(1-\omega^2)^{1/2}}P_{lm}(\omega)-P_{lm+1}(\omega)\right\} \qquad (7.81)
\end{aligned}
$$

The associated Legendre function satisfies the following recursion relation (Butkov).

$$
P_{lm+1}(\omega)-\frac{2m\omega}{(1-\omega^2)^{1/2}}P_{lm}+(l+m)(l+m+1)P_{lm-1}=0
$$

Making use of this relation in (7.81) we get

$$
\begin{aligned}
L\_Y_{lm} &= -\hbar e^{i(m-1)\phi}(-1)^m\cdot\frac{N_{lm}}{\sqrt{2\pi}}(l+m)(l-m+1)P_{lm-1} \\
&= \hbar\cdot\frac{N_{lm}}{N_{lm-1}}Y_{lm-1}(l+m)(l-m+1) \\
&= \hbar\cdot\left[\frac{(2l+1)(l-m)!}{2(l+m)!}\,\frac{2(l+m-1)!}{(2l+1)(l-m+1)!}\right]^{1/2}(l+m)(l-m+1)Y_{lm-1} \\
&= [(l+m)(l-m+1)]^{1/2}\hbar Y_{lm-1}
\end{aligned}
$$

7.11 EIGENSTATES OF L_x AND L_y

We have already seen that L_x, L_y and L_z do not commute with each other. Therefore, Y_{lm} cannot be an eigenstate of L_x or L_y. However, from the structure of the commutation relation, it is obvious that the eigenvalues of L_x and L_y are also $m\hbar$, where $m = l, l - l, \ldots, -l$. So we can construct the eigenstate of L_x or L_y as a linear combination of the spherical harmonics $Y_{lm}s$.

So we have

$$L^2\psi_{1\nu} = l(l+1)\hbar^2\psi_{1\nu} \tag{7.82}$$

$$L_x\psi_{1\nu} = \nu\hbar\psi_{1\nu} \qquad \nu = l, l-1, \ldots -l \tag{7.83}$$

where

$$\psi_{l\nu} = \sum_{m=l}^{-l} C_m Y_{lm} \tag{7.84}$$

Similar results hold for L_y also.

Example 7.8 Construct the eigenstates of L_x for $l = 1$ in terms of the spherical harmonics Y_{11}, Y_{10} and Y_{1-1}

Solution: Let $\psi_{1\nu}$ be the eigenfunction of L_x.

i.e.

$$L_x\psi_{l\nu} = \nu\hbar\,\psi_{l\nu}$$

Let us write $\psi_{l\nu}$ as

$$\psi_{l\nu} = c_1Y_{11} + c_0Y_{10} + c_{-1}Y_{1-1}$$

L_x is given by

$$L_x = \frac{1}{2}(L_+ + L_-).$$

$$L_x\psi_{l\nu} = \frac{1}{2}(L_+ + L_-)(c_1Y_{11} + c_0Y_{10} + c_{-1}Y_{1-1})$$

$$= \frac{\hbar}{\sqrt{2}}c_0Y_{11} + \frac{\hbar}{\sqrt{2}}(c_1 + c_{-1})Y_{10} + \frac{\hbar}{\sqrt{2}}c_0Y_{1-1} \tag{7.85}$$

But

$$L_x\psi_{l\nu} = \nu\hbar\psi_{l\nu} = \nu\hbar[c_1Y_{11} + c_0Y_{10} + c_{-1}Y_{1-1}] \tag{7.86}$$

Comparing (7.85) and (7.86) we get

$$\frac{1}{\sqrt{2}}c_0 = \nu c_1, \qquad \frac{1}{\sqrt{2}}(c_1 + c_{-1}) = \nu c_0, \qquad \frac{1}{\sqrt{2}}c_0 = \nu c_{-1}$$

In the matrix form, we can write the above equations as

$$\begin{bmatrix} \nu & -\frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & -\nu & \frac{1}{\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} & -\nu \end{bmatrix}\begin{bmatrix} c_1 \\ c_0 \\ c_{-1} \end{bmatrix} = 0$$

For the non-trivial solution of c_1, c_0, and c_{-1}, the determinant should be zero

$$\therefore \nu\left[\nu^2-\frac{1}{2}\right]-\frac{1}{\sqrt{2}}\frac{\nu}{\sqrt{2}}=0$$

$$\nu(\nu^2-1)=0 \qquad \therefore \nu=0,\pm 1.$$

$$\nu=1 \qquad c_1=\frac{c_0}{\sqrt{2}} \qquad c_{-1}=\frac{c_0}{\sqrt{2}}$$

$$\therefore \psi_{11}=c_0\left[\frac{1}{\sqrt{2}}Y_{11}+Y_{10}+\frac{1}{\sqrt{2}}Y_{1-1}\right]$$

$$\frac{c_0^2}{2}+c_0^2+\frac{c_0^2}{2}=1$$

$$\therefore c_0=\frac{1}{\sqrt{2}}$$

$$\psi_{11}=\left[\frac{Y_{11}}{2}+\frac{1}{\sqrt{2}}Y_{10}+Y_{\frac{1-1}{2}}\right] \tag{7.87}$$

$$\nu=\ 0: \qquad \psi_{10}=\ c_1(Y_{11}-Y_{1-1})$$

Normalization leads to $c_1=\frac{1}{\sqrt{2}}$

$$\therefore \psi_{10}=\frac{1}{\sqrt{2}}(Y_{11}-Y_{1-1}) \tag{7.88}$$

$$\nu=-1 \qquad c_1=-\frac{c_0}{\sqrt{2}} \qquad c_{-1}=-\frac{c_0}{\sqrt{2}}$$

$$\therefore \psi_{1-1}=c_0\left[-\frac{1}{\sqrt{2}}Y_{11}+Y_{10}-\frac{1}{\sqrt{2}}Y_{1-1}\right]$$

Normalization leads to

$$\frac{c_0^2}{2}+c_0^2+\frac{c_0^2}{2}=1$$

$$c_0=\frac{1}{\sqrt{2}}$$

$$\therefore \psi_{1-1}=\left[-\frac{1}{2}Y_{11}+\frac{1}{\sqrt{2}}Y_{10}-\frac{1}{2}Y_{1-1}\right] \tag{7.89}$$

Example 7.9 Calculate $\langle L_x\rangle$ for an angular momentum state $Y_{lm}(\theta,\phi)$

Solution:

$$\langle L_x\rangle=\int Y_{lm}^{*}L_xY_{lm}d\Omega$$

$$
\begin{aligned}
&= \int Y_{lm}^{*} \frac{1}{2}[L_{+} + L_{-}]Y_{lm}\, d\Omega \\
&= \frac{1}{2}\int Y_{lm}^{*}[(l-m)(l+m+1)]^{1/2}\, Y_{lm-1}\, d\Omega \\
&\quad + \frac{1}{2}\int Y_{lm}^{*}[(l+m)(l-m+1)]^{1/2}\, Y_{lm-1}\, d\Omega \\
&= 0
\end{aligned}
$$

These integrals are zero due to orthoganality of the spherical harmonics.

$$\langle L_x \rangle = 0$$

Similarly, $\langle L_y \rangle = 0$

Example 7.10 Calculate $\langle L_x^2 \rangle$ for a state $Y_{lm}(\theta,\phi)$

$$\langle L_x^2 \rangle = \int Y_{lm} L_x^2 Y_{lm}\, d\Omega$$

$$
\begin{aligned}
\langle L_x^2 \rangle Y_{lm} &= \frac{1}{4}(L_{+} + L_{-})(L_{+} + L_{-})Y_{lm} \\
&= \frac{1}{4}(L_{+})^2 Y_{lm} + \frac{1}{4}(L_{-})^2 Y_{lm} + \frac{1}{4}(L_{+}L_{-} + L_{-}L_{+})Y_{lm}
\end{aligned}
$$

Let us write $(L_{+})^2 Y_{lm} = aY_{lm+2}$ $\quad (L_{-})^2 Y_{lm} = bY_{lm-2}$

where *a* and *b* are constants which can be determined easily.

$$\int Y_{lm}^{*}(L_{+})^2 Y_{lm}\, d\Omega = a\int Y_{lm}^{*} Y_{lm+2} = 0$$

$$\int Y_{lm}^{*}(L_{-})^2 Y_{lm}\, d\Omega = b\int Y_{lm}^{*} Y_{lm-2} = 0$$

$$
\begin{aligned}
L_{+}L_{-} + L_{-}L_{+} &= (L_x + iL_y)(L_x - iL_y) + (L_x - iL_y)(L_x + iL_y) \\
&= L_x^2 - iL_xL_y + iL_yL_x + L_y^2 + L_x^2 - iL_yL_x + iL_xL_y + L_y^2 \\
&= 2\,(L_x^2 + L_y^2) = 2\,(L^2 - L_z^2)
\end{aligned}
$$

$$
\begin{aligned}
\langle L_x^2 \rangle &= \frac{1}{2}\int Y_{lm}^{*}(L^2 - L_z^2)Y_{lm}\, d\Omega \\
&= \frac{1}{2}\int Y_{lm}^{*}[l(l+1)\hbar^2 - m^2\hbar^2]Y_{lm}\, d\Omega = \frac{1}{2}[l(l+1) - m^2]\hbar^2.
\end{aligned}
$$

Example 7.11 Calculate $\langle L_y^2 \rangle$ for a given state $Y_{lm}(\theta,\phi)$.

Solution:

$$L_y = \frac{1}{2i}(L_{+} - L_{-})$$

$$\begin{aligned}L_y^2 &= \frac{1}{4}[L_+ - L_-][L_+ - L_-] = -\frac{1}{4}\{L_+^2 + L_-^2 - (L_+L_- + L_-L_+)\}\\ &= -\frac{1}{4}(L_+^2 + L_-^2) + \frac{1}{4}(L_+L_- + L_-L_+)\\ \langle L_y^2\rangle &= \frac{1}{2}[l(l+1) - m^2]\hbar^2\end{aligned}$$

Example 7.12 Determine the product of uncertainties $\Delta L_x\, \Delta L_y$ for the state $Y_{ml}(\theta,\phi)$.

Solution:

$$\begin{aligned}(\Delta L_x)^2 &= \langle L_x^2\rangle - \langle L_x\rangle^2\\ &= \frac{1}{2}[l(l+1) - m^2]\hbar^2.\\ (\Delta L_y)^2 &= \langle L_y^2\rangle - \langle L_y\rangle^2 = \frac{1}{2}[l(l+1) - m^2]\hbar^2\\ \Delta L_x &= \frac{1}{\sqrt{2}}[l(l+1) - m^2]^{1/2}\hbar\\ \Delta L_y &= \frac{1}{\sqrt{2}}[l(l+1) - m^2]^{1/2}\hbar\\ \therefore \Delta L_x \Delta L_y &= \frac{1}{2}[l(l+1) - m^2]\hbar^2\end{aligned}$$

7.12 ORBITAL ANGULAR MOMENTUM IN QUANTUM AND CLASSICAL PHYSICS

We note that L_x, L_y and L_z do not commute among themselves and so they do not have common eigenfunctions. This is summarized as follows:

$$L^2Y_{lm} = l(l+1)\hbar^2 Y_{lm} \qquad L^2\psi_{l\nu} = l(l+1)\hbar^2\psi_{l\nu} \qquad L^2\phi_{l\mu} = l(l+1)\hbar^2\phi_{l\mu}$$

$$L_zY_{lm} = m\hbar Y_{lm} \qquad L_x\psi_{l\nu} = \nu\hbar\psi_{l\nu} \qquad L_y\phi_{l\mu} = \mu\hbar\phi_{l\mu}$$

The fact that the operators corresponding to the three components do not commute among themselves makes the properties of angular momentum in quantum mechanics to be significantly different from that in classical physics. In classical physics, each of three components can be independently fixed. In quantum mechanics, depending on the choice of the basis functions, only one component is well defined. For instance, if we choose Y_{lm} as the basis functions to describe a system, they are eigenstates of L_z but they are not eigenstates of L_x and L_y. Therefore, only L^2 and L_z have well-defined values. L_x and L_y do not have any definite value. What does this mean? When the system is in state Y_{lm}, if one measures square of the total orbital angular momentum, the result will be $l(l+1)\hbar^2$; if one measures z component of the orbital angular momentum, the result will be $m\hbar$; but if we measure L_x or L_y, the result can be any of the $(2l+1)$ possible values: $l\hbar$, $(l-1)\hbar$, …, $-l\hbar$ since the functions Y_{lm} are superposition of either ψ_{lm} or $\phi_{lm}\left(Y_{lm} = \sum_\nu c_\nu\psi_{l\nu} \text{ or } Y_{lm} = \sum_\mu d_\mu\phi_{l\mu}\right)$. In fact , no value can be attributed for L_x or L_y, before measurement, when the system's state is a superposition of eigenstates of these observables.

If we try to construct a classical picture for **L** and L_z, then **L** is a vector of length $\sqrt{l(l+1)}\hbar$ and its z-component is $L_z = m\hbar = l\hbar, (l-1)\hbar, \ldots, -l\hbar$. This result is shown in Fig. 7.10.

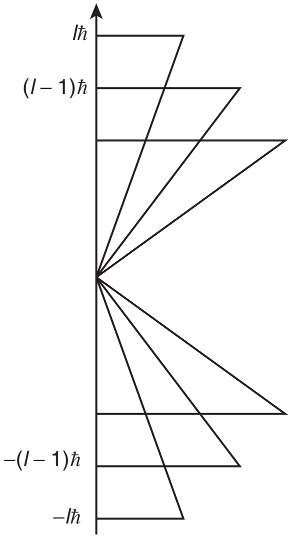


Fig. 7.10 Z component of angular momentum

Classically, the maximum value of L_z is L itself. Quantum mechanically, this is not possible. $L = \sqrt{l(l+1)}\hbar$ but the maximum value of L_z is $l\hbar$. Classically, there should not be any difference in the behaviour of L_x, L_y and L_z. However, quantum mechanically, L_z is well defined but L_x and L_y are not well defined. In fact, if we go by the uncertainly relation,

$$(\Delta L_x)^2(\Delta L_y)^2 \geq -\frac{1}{4}\langle [L_x, L_y] \rangle^2$$

$$(\Delta L_x)^2(\Delta L_y)^2 \geq -\frac{1}{4}\langle [i\hbar L_z] \rangle^2$$

or

$$(\Delta L_x)^2(\Delta L_y)^2 \geq \frac{1}{4} m^2\hbar^4$$

$\therefore$ For a system in state $Y_{lm}(\theta, \phi)$

$$\Delta L_x \Delta L_y \geq \frac{m\hbar^2}{2} \tag{7.90}$$

If we extend the classical picture further, one can picture this argument as the precession of angular momentum vector along z axis.

7.13 RIGID ROTOR

Rigid rotor is a system of two particles separated by a fixed distance r_0. It is called a rigid rotor because the distance between the particles is constant.

The Hamiltonian of a rigid rotor due to rotational motion is

$$H = \frac{L^2}{2I} = \frac{L^2}{2\mu r_0^2}$$

where r_0 is the fixed distance separating the two particles and μ is the reduced mass.
The quantum mechanical Hamiltonian operator for the rigid rotor is

$$H_{op} = \frac{L_{op}^2}{2\mu r_0^2} \tag{7.91}$$

The energy eigenvalue equation is

$$H_{op}\chi(\theta,\phi) = \frac{L_{op}^2}{2\mu r_0^2}\chi(\theta,\phi)$$

Obviously, $\chi(\theta, \phi)$ is the spherical harmonic $Y_{lm}(\theta, \phi)$

$$\therefore \frac{L_{op}^2}{2\mu r_0^2} Y_{lm}(\theta,\phi) = \frac{l(l+1)\hbar^2}{2\mu r_0^2} Y_{lm}(\theta,\phi)$$

Therefore, the energy eigenvalues of a rigid rotor are given by

$$E_l = \frac{l(l+1)\hbar^2}{2\mu r_0^2} \tag{7.92}$$

This is one of the important results in molecular spectroscopy. The significance of the rigid rotor model for a diatomic molecule can be seen as follows. The Hamiltonian for a system of two particles of masses m_1 and m_2, which interact through a central force, is given by

$$\begin{aligned} H &= \frac{p_1^2}{2m_1} + \frac{p_2^2}{2m_2} + V\left(|\mathbf{r}_1 - \mathbf{r}_2|\right) \\ &= -\frac{\hbar^2}{2m_1}\nabla_1^2 - \frac{\hbar^2}{2m_2}\nabla_2^2 + V\left(\left\|\mathbf{r}_1 - \mathbf{r}_2\right\|\right) \end{aligned}$$

In Chapter 8, it will be shown that this Hamiltonian can be split into two parts:

$$H = H_{cm} + H_{rel}$$

H_{cm} describes the motion of the centre of mass whose coordinate is given by $\mathbf{R} = (m_1\mathbf{r}_1 + m_2\mathbf{r}_2)/(m_1 + m_2)$. H_{rel} describes the relative motion between the two particles. It depends on $\mathbf{r} = \mathbf{r}_1 - \mathbf{r}_2$. H_{rel} is given by

$$\begin{aligned} H_{rel} &= -\frac{\hbar^2}{2\mu^2}\nabla^2 + V(r) \\ &= -\frac{\hbar^2}{2\mu^2}\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) - \frac{L^2}{2\mu r^2} + V(r) \end{aligned} \tag{7.93}$$

In the rigid body approximation, the term $L^2/2\mu r^2$ is replaced by $L^2/2\mu r_0^2$ where r_0 is the equilibrium distance between the particles.

$$\therefore H_{rel} = H(r) + H(\theta,\phi)$$

The structure of the Hamiltonian suggests that $u(r,\theta,\phi)$ can be written as $u(r,\theta,\phi) = R(r)\chi(\theta,\phi)$ so that $E = E_r + E_{rot}$

where

$$\left[-\frac{\hbar^2}{2\mu}\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + V(r)\right]R(r) = E_r R(r)$$

and

$$\frac{L^2}{2\mu r_0^2}\chi(\theta,\phi) = E_{rot}\chi(\theta,\phi)$$

So the rigid rotor approximation separates the rotational motion from the radial motion of the diatomic molecule. In molecular spectroscopy, the important radial motion is the vibrational motion of the atom.

Example 7.13 Obtain Y_{ll} from the equation $L_+ Y_{ll} = 0$.

Solution:

$$Y_{ll} = N_{ll}\Phi_{ll}(\theta)e^{il\phi}$$

$$L_+ Y_{ll} = i\hbar e^{i\phi}\left[-i\frac{\partial}{\partial\theta} + \cot\theta\frac{\partial}{\partial\phi}\right]N_{ll}\Phi_{ll}e^{il\phi} = 0$$

$$\left(\frac{\partial}{\partial\theta} - l\cot\theta\right)\Phi_{ll} = 0 \quad \text{or} \quad \frac{\partial\Phi_{ll}}{\partial\theta} = l\frac{\cos\theta}{\sin\theta}\Phi_{ll}$$

$$\frac{d\Phi_{ll}}{\Phi_{ll}} = l\frac{\cos\theta}{\sin\theta}d\theta$$

On integration, we get

$$d(\ln\Phi_{ll}) = l\,d(\ln\sin\theta)$$
$$\Phi_{ll} = A\sin^l\theta$$

The constant A can be determined from the condition

$$\int_0^\pi |\Phi_{ll}|^2 \sin\theta\, d\theta = 1$$

Assuming A to be real, we get

$$\int_0^\pi |\Phi_{ll}|^2 \sin\theta d\theta = \int_{-1}^{1} A^2(1-\omega^2)^l d\omega$$

$$= \int_{-1}^{1} A^2(1+\omega)^l(1-\omega)^l d\omega$$

Put $$1+\omega=t$$
$$-1-\omega=2-t$$

$$\therefore A^2\int_{-1}^{1}(1+\omega)^l(1-\omega)^l\,d\omega = A^2\int_0^2(2-t)^l t^l dt = A^2\cdot\frac{t^{l+1}}{l+1}(2-t)^l\,\Big|_0^2 + A^2 l\int_0^2\frac{t^{l+1}}{(l+1)}(2-t)^{l-1}$$

So we have,

$$\int_0^2(2-t)^l t^l dt = \frac{l}{l+1}\cdot\int_0^2 t^{l+1}\cdot(2-t)^{l-1}$$

Proceeding in the same way, we get

$$\int_0^2(2-t)^l t^l dt = \frac{l!}{(l+1)(1+2)..2l}\int_0^2 t^{2l}dt$$

$$= \frac{2^{2l+1}l!}{(l+1)(l+2)..(2l+1)} = \frac{2^{2l+1}(l!)^2}{(2l+1)!}$$

$$\therefore A = -\left[\frac{(2l+1)!}{2^{2l+1}(l!)^2}\right]^{1/2}$$

The negative sign is included to make the solution spherical harmonic as defined in the section ‘Spherical harmonics as eigenfunctions of angular momentum’.

$$\therefore \Phi_{ll} = -\left[\frac{(2l+1)!}{2^{2l+1}(l!)^2}\right]^{1/2}\sin^l\theta$$

$$Y_{ll} = -\left[\frac{(2l+1)!}{2^{2l+1}(l!)^2}\right]^{1/2}\sin^l\theta\cdot\frac{1}{\sqrt{2\pi}}e^{il\phi}$$

Example 7.14 Assuming $Y_{ll} = -\sqrt{\frac{3}{8\pi}}\sin\theta\, e^{i\phi}$ obtain Y_{10} and Y_{11} using the lowering operator L.

Solution:

$$L_- = i\hbar e^{-i\phi}\left[i\frac{\partial}{\partial\theta}+\cot\theta\frac{\partial}{\partial\phi}\right]$$

$$L_-Y_{11} = i\hbar e^{-i\phi}\left(-\sqrt{\frac{3}{8\pi}}\right)\cdot\left[i\frac{\partial}{\partial\theta}+\cot\theta\frac{\partial}{\partial\phi}\right]\sin\theta\, e^{i\phi}$$

$$= \hbar\cdot\sqrt{\frac{3}{8\pi}}2\cos\theta$$

$$L_-Y_{11} = \sqrt{2}\hbar Y_{10}$$

$$\therefore \sqrt{2}\hbar Y_{10} = \hbar \cdot \sqrt{\frac{3}{8\pi}} 2\cos\theta$$

$$\therefore Y_{10} = \sqrt{\frac{3}{4\pi}} \cos\theta$$

$$L\_Y_{10} = i\hbar e^{-i\phi}\left[i\frac{\partial}{\partial\theta} + \cot\theta\frac{\partial}{\partial\phi}\right] \cdot \sqrt{\frac{3}{4\pi}}\cos\theta$$

$$= i\hbar\sqrt{\frac{3}{4\pi}}\sin\theta\, e^{-i\phi}$$

But $$L\_Y_{10} = \sqrt{2}\hbar Y_{1-1}$$

$$\therefore Y_{1-1} = \sqrt{\frac{3}{8\pi}}\sin\theta e^{-i\phi}$$

Example 7.15 Prove that

$$P_{l-m}(x) = (-1)^m \frac{(l-m)!}{(l+m)!} P_{lm}(x)$$

Solution: $P_{lm}(x)$ is given by

$$P_{lm}(x) = \frac{1}{2^l} \cdot \frac{1}{l!}(1-x^2)^{\frac{m}{2}} D^{l+m}(x^2-1)^l$$

Using Leibnitz rule, we have

$$\begin{aligned} D^{l+m}(x^2-1)^l &= D^{l+m}(x+1)^l(x-1)^l \\ &= \sum_{r=0}^{l+m} {}^{l+m}C_r D^{l+m-r}(x+1)^l D^r(x-1)^l \end{aligned} \tag{7.94}$$

where the binomial coefficient is given by ${}^{l+m}C_r = \dfrac{(l+m)!}{r!(l+m-r)!}$

Note that $D^{l+m-r}(x+1)^l D^r(x-1)^l$ is not zero provided $l+m-r \le l$ and $r \le l$. Or equivalently $r \ge m$ and $r \le l$.

$$\therefore D^{l+m-r}(x+1)^l D^r(x-1)^l \begin{cases} \neq 0 & \text{if } m \le r \le l \\ 0 & \text{otherwise} \end{cases}$$

Further, $D^r(x \pm 1)^l = \dfrac{l!}{(l-r)!}(x \pm 1)^{l-r}$ provided $r \le l$

Making use of these results in the Equation (7.94), r takes the values from m to l in the summation, and so we get

$$\begin{aligned}{}^{l+m}(x^2-1)^l &= {}^{l+m}C_m D^l(x+1)^l D^m(x-1)^l + {}^{l+m}C_{m+1}D^{l-1}(x+1)^l D^{m+1}(x-1)^l + ..\\ &\quad +...+ {}^{l+m}C_l D^m(x+1)^l D^l(x-1)^l\\ &= \sum_{r=m}^{l} {}^{l+m}C_r D^{l+m-r}(x+1)^l D^r(x-1)^l\\ &= \sum_{r=m}^{l} {}^{l+m}C_r \frac{l!}{(r-m)!}(x+1)^{r-m}\frac{(x-1)^{l-r}\,l!}{(l-r)!}\end{aligned}$$

$$\therefore P_{lm} = \frac{1}{2^l}\cdot\frac{1}{l!}\sum_{r=m}^{l}\frac{(l+m)!}{r!(l+m-r)!}\cdot\frac{(l!)^2}{(r-m)!(l-r)!}(x+1)^{r-m}(x-1)^{l-r} \tag{7.95}$$

Now consider $D^{l-m}(x^2-1)^l$.

$$D^{l-m}(x^2-1)^l = \sum_{s}^{l-m} {}^{l-m}C_s D^{l-m-s}(x+1)^l D^s(x-1)^l$$

Put $s+m=r$

$\therefore$ r varies from m to l.

$$\begin{aligned}\therefore D^{l-m}(x^2-1)^l &= \sum_{r=m}^{l} {}^{l-m}C_{r-m} D^{l-r}(x+1)^l D^{r-m}(x-1)^l\\ &= \sum_{r=m}^{l}\frac{(l-m)!}{(r-m)!(l-r)!}\cdot\frac{l!}{r!}\cdot\frac{l!}{(l-r+m)!}(x+1)^r(x-1)^{l-r+m} \qquad (7.96)\end{aligned}$$

$$\begin{aligned}\therefore P_{l-m} &= \frac{1}{2^l}\cdot\frac{1}{l!}(1-x^2)^{-\frac{m}{2}}D^{l-m}(x^2-1)^l\\ &= \frac{1}{2^l}\cdot\frac{1}{l!}(1-x^2)^{\frac{m}{2}}\cdot\frac{D^{l-m}(x^2-1)^l}{(1-x^2)^m}\\ &= \frac{1}{2^l}\cdot\frac{1}{l!}(1-x^2)^{\frac{m}{2}}\cdot\frac{D^{l-m}(x^2-1)^l}{(-1)^m(x^2-1)^m}\\ &= (-1)^m\frac{1}{2^l}\cdot\frac{1}{l!}(1-x^2)^{\frac{m}{2}}\cdot\frac{D^{l-m}(x^2-1)^l}{(x+1)^m(x-1)^m}\end{aligned}$$

$$\begin{aligned}P_{l-m} &= \frac{1}{2^l}\cdot\frac{(-1)^m}{l!}(1-x^2)^{m/2}\cdot\sum_{r=m}^{l}\frac{(l-m)!}{(r-m)!(l-r)!}\frac{l!}{r!}\cdot\frac{l!}{(l-r+m)!}\frac{(x+1)^{l+r}(x-1)^{l-r+m}}{(x+1)^m(x-1)^m}\\ &= (-1)^m\frac{(l-m)!}{(l+m)!}\cdot\frac{1}{2^l}(1-x^2)^{m/2}\sum_{r=m}^{l}\frac{(l+m)!}{(r-m)!}\frac{(l!)^2}{(l-r)!(l-r+m)!}(x+1)^{r-m}(x-1)^{l-r}\\ &= (-1)^m\frac{(l-m)!}{(l+m)!}P_{lm}(x) \qquad \text{(from (7.95))}\end{aligned}$$

Example 7.16 For an atom interacting with electromagnetic field, the interaction Hamiltonian in the dipole approximation can be written as $-e\boldsymbol{\varepsilon}\cdot\mathbf{r}$ where $\boldsymbol{\varepsilon}$ is the electric field. Evaluate the matrix element $x_{lm,l'm'}$, $y_{lm,l'm'}$ and $z_{lm,l'm'}$. Hence obtain the selection rule due to dipole interaction.

Solution: The matrix elements are $(x)_{lm,l'm'}$, $y_{lm,l'm'}$ and $z_{lm,l'm'}$ given by

$$(x)_{lm,l'm'} = \int Y_{lm}^{*}(\theta,\phi)\,xY_{l'm'}(\theta,\phi)d\Omega$$

$$(y)_{lm,l'm'} = \int Y_{lm}^*(\theta,\phi)\, y\, Y_{l'm'}(\theta,\phi)\, d\Omega$$

$$(z)_{lm,l'm'} = \int Y_{lm}^*(\theta,\phi)\, z\, Y_{l'm'}(\theta,\phi)\, d\Omega$$

In the spherical coordinates

$$x = r\sin\theta\cos\phi = r\sin\theta(e^{i\phi} + e^{-i\phi})\frac{1}{2}$$

$$y = r\sin\theta\sin\phi = \frac{1}{2i} r\sin\theta(e^{i\phi} - e^{-i\phi})$$

$$z = r\cos\theta$$

Using the recursion relations (7.67) and (7.69), we get

$$\begin{aligned}
xY_{l'm'} &= \frac{r}{2}[\sin\theta\, e^{i\phi} Y_{l'm'} + \sin\theta\, e^{-i\phi} Y_{l'm'}] \\
&= \frac{r}{2}\left\{\begin{aligned} &-\left[\frac{(l'+m'+1)(l'+m'+2)}{(2l'+1)(2l'+3)}\right]^{1/2} Y_{l'+1,m'+1} + \left[\frac{(l'-m')(l'-m'+1)}{(2l'+1)(2l'-1)}\right]^{1/2} Y_{l'-1,m'+1} \\ &-\left[\frac{(l'-m'+1)(l'-m'+2)}{(2l'+1)(2l'+3)}\right]^{1/2} Y_{l'+1,m'-1} + \left[\frac{(l'+m')(l'+m'+1)}{(2l'+1)(2l'-1)}\right]^{1/2} Y_{l'-1,m'-1} \end{aligned}\right\} \\
&= \frac{r}{2}[a_1 Y_{l'+1,m'+1} + b_1 Y_{l'-1,m'+1} + c_1 Y_{l'+1,m'-1} + c_2 Y_{l'-1,m'-1}]
\end{aligned}$$

$$\begin{aligned}
\therefore (x)_{lm,l'm'} &= \int Y_{lm}^*(\theta,\phi)\, x Y_{l'm'}\, d\Omega \\
&= \frac{ra_1}{2}\int Y_{lm}^* Y_{l'+1,m'+1} d\Omega + \frac{rb_1}{2}\int Y_{lm}^* Y_{l'-1,m'+1} d\Omega + \frac{rc_1}{2}\int Y_{lm}^* Y_{l'+1,m'-1} d\Omega + \frac{rd_1}{2}\int Y_{lm}^* Y_{l'-1,m'-1} d\Omega \\
&= \frac{ra_1}{2}\delta_{l,l'+1}\delta_{m,m'+1} + \frac{rb_1}{2}\delta_{l,l'-1}\delta_{m,m'+1} + \frac{rc_1}{2}\delta_{l,l'+1}\delta_{m,m'-1} + \frac{rd_1}{2}\delta_{l,l'-1}\delta_{m,m'-1}
\end{aligned}$$

Similarly, we find

$$(y)_{lm,l'm'} = \frac{ra_2}{2i}\delta_{l,l'+1}\delta_{m,m'+1} + \frac{rb_2}{2i}\delta_{l,l'-1}\delta_{m,m'+1} + \frac{rc_2}{2i}\delta_{l,l'+1}\delta_{m,m'-1} + \frac{rd_2}{2i}\delta_{l,l'-1}\delta_{m,m'-1}$$

$$(z)_{lm,l'm'} = ra_3\delta_{l,l'+1}\delta_{mm'} + rb_3\delta_{l,l'-1}\delta_{m,m'}.$$

The matrix element for the dipole $e\boldsymbol{\varepsilon}\cdot\mathbf{r}$ is given by

$$\begin{aligned}
&\int Y_{lm}^* e\boldsymbol{\varepsilon}\cdot\mathbf{r}\, Y_{l'm'} d\Omega \\
&= \frac{er\varepsilon_x}{2}[a_1\delta_{l,l'+1}\delta_{m,m'+1} + b_1\delta_{l,l'-1}\delta_{m,m'+1} + c_1\delta_{l,l'+1}\delta_{m,m'-1} + d_1\delta_{l,l'-1}\delta_{m,m'-1}] \\
&\quad - \frac{ier\varepsilon_y}{2}[a_2\delta_{l,l'+1}\delta_{m,m'+1} + b_2\delta_{l,l'-1}\delta_{m,m'+1} + c_2\delta_{l,l'+1}\delta_{m,m'-1} + d_2\delta_{l,l'-1}\delta_{m,m'-1}] \\
&\quad + er\varepsilon_z[a_3\delta_{l,l'+1}\delta_{mm'} + b_3\delta_{l,l'-1}\delta_{mm'}]
\end{aligned}$$

Therefore, we have

$$(e\boldsymbol{\varepsilon}\cdot\mathbf{r})_{lm,l'm'}\begin{cases}\neq 0 & \begin{array}{l}\text{provided } l = l' \pm 1\\ \qquad\qquad m = m', m = m' \pm 1\end{array}\\ = 0 & \text{otherwise}\end{cases}$$

In atomic physics $l_f m_f$ refers to the final state of the atom, and $l_i m_i$ refers to the individual state of the atom. The dipole moment $e\boldsymbol{\varepsilon}\cdot\mathbf{r}$ induces an atomic transition from the state with $l_{if}m_i$ to $l_{ff}m_f$. So the matrix element of $e\boldsymbol{\varepsilon}\cdot\mathbf{r}$ is now written as

$$(e\boldsymbol{\varepsilon}\cdot\mathbf{r})_{l_f m_f, l_i m_i}\begin{cases}\neq 0 & \begin{array}{l}\text{provided } l_f = l_i \pm 1\\ \qquad\qquad m_f = m_i, m_f = m_i \pm 1\end{array}\\ = 0 & \text{otherwise}\end{cases}$$

Therefore the selection rules are

$$l_f = l_i \pm 1, m_f = m_i, m_i \pm 1.$$

EXERCISES

1. Evaluate (i) $[L_x, x]$ (ii) $[L_y, x]$ (iii) $[L_z, x]$ (iv) $[L_x, y]$ (v) $[L_y, y]$ (vi) $[L_z, z]$ (vii) $[L_x, z]$ (viii) $[L_y, z]$ (ix) $[L_z, z]$
2. Show that $[L^2, \mathbf{r}] = i\hbar[\mathbf{L}\times\mathbf{r} - \mathbf{r}\times\mathbf{L}]$
3. Evaluate (i) $[L_x, p_x]$ (ii) $[L_y, p_x]$ (iii) $[L_z, p_x]$
4. Show that $[L^2, \mathbf{p}] = i\hbar[-\mathbf{L}\times\mathbf{p} + \mathbf{p}\times\mathbf{L}]$
5. Starting from Y_{22}, obtain Y_{21}, Y_{20}, Y_{2-1} and Y_{2-2} using L_- operator.
6. Starting from Y_{2-2}, obtain Y_{2-1}, Y_{20}, Y_{21} and Y_{22} using L_+ operator.
7. Obtain Y_{l-l} by solving the equation $L_- Y_{ll} = 0$.
8. Distinguish the angular momentum in classical and quantum physics.
9. Construct the eigenfunctions of L_y in terms of the spherical harmonics.
10. A system is in a state Y_{lm}. What are the possible values of (i) z component of $\mathbf{L}$ (ii) x component of $\mathbf{L}$.
11. A system is in a state Y_{2m}. When L_x is measured, what is the probability of getting L_x to be $2\hbar, \hbar, 0, -\hbar$ and $-2\hbar$?
12. Obtain the matrix elements $(x^2)_{lm,l'm'}$, $(y^2)_{lm,l'm'}$, $(z^2)_{lm,l'm'}$, $(xy)_{lm,l'm'}$, $(yz)_{lm,l'm'}$, $(xz)_{lm,l'm'}$.

REFERENCES

1. M. Alonso and H. Valk, 1972. *Quantum Mechanics: Principles and Applications.* Massachusetts: Addison–Wesley Publishing Company.
2. P. M. Mathews and A. Venkatesan, 1976. *A Textbook of Quantum Mechanics.* New Delhi: Tata–McGraw Hill Publishing Company Ltd.
3. Eugene Butkov, 1968. *Mathematical Physics:* Massachusetts: Addison–Wesley Publishing Company.

8

Time-Independent Schrödinger Equation in Three Dimensions

In this chapter, we consider a few simple potentials for which we determine the bound state, as well as the scattering state, solutions of the time-independent Schrödinger equation given by

$$\left[-\frac{\hbar^2}{2m}\nabla^2+V(\mathbf{r})\right]u(\mathbf{r})=Eu(\mathbf{r})$$

Therefore, depending on the nature of the potential, we can employ either Cartesian coordinates or spherical polar coordinates or any other curvilinear coordinate systems.

8.1 PARTICLE IN A RECTANGULAR BOX

In one dimension, we considered a particle confined to a region $0 < x < L$.

We can extend this idea to three dimensions.

The potential $V(\mathbf{r})$ is now given by

$$V(\mathbf{r})=\begin{cases}0 & 0<x<a;\ 0<y<b;\ 0<z<c\\ \infty & \text{otherwise}\end{cases} \tag{8.1}$$

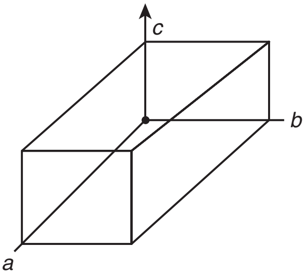


Fig. 8.1 Rectangular box potential

Therefore, the Schrödinger equation is

$$-\frac{\hbar^2}{2m}\nabla^2 u(\mathbf{r}) = Eu(\mathbf{r}) \qquad 0 < x < a; \quad 0 < y < b; \quad 0 < z < c \tag{8.2}$$

We have to solve this equation, subject to the boundary condition

$$\begin{aligned} \psi(0, 0, 0) &= \psi(a, y, z) = 0 \\ \psi(x, b, z) &= \psi(x, y, c) = 0 \end{aligned} \tag{8.3}$$

The Schrödinger equation now becomes

$$-\frac{\hbar^2}{2m}\frac{\partial^2 u}{\partial x^2} - \frac{\hbar^2}{2m}\frac{\partial^2 u}{\partial y^2} - \frac{\hbar^2}{2m}\frac{\partial^2 u}{\partial z^2} = Eu \tag{8.4}$$

The structure of this equation suggests that $u(x, y, z)$ can be written as

$$u(x, y, z) = X(x)\,Y(y)\,Z(z) \tag{8.5}$$

The Equation (8.4) now becomes

$$-\frac{\hbar^2}{2m}\frac{d^2X}{dx^2}YZ - \frac{\hbar^2}{2m}\frac{d^2Y}{dy^2}XZ - \frac{\hbar^2}{2m}\frac{d^2Z}{dz^2}XY = EXYZ$$

Dividing throughout by XYZ, we get

$$-\frac{\hbar^2}{2m}\frac{1}{X}\frac{d^2X}{dx^2} - \frac{\hbar^2}{2m}\frac{1}{Y}\frac{d^2Y}{dy^2} - \frac{\hbar^2}{2m}\frac{1}{Z}\frac{d^2Z}{dz^2} = E \tag{8.6}$$

Let us write $\frac{2mE}{\hbar^2}$ as

$$\frac{2mE}{\hbar^2} = k^2 = k_x^2 + k_y^2 + k_z^2 \tag{8.7}$$

The Equation (8.6) becomes

$$\left(\frac{1}{X}\frac{d^2X}{dx^2} + k_x^2\right) + \left(\frac{1}{Y}\frac{d^2Y}{dy^2} + k_y^2\right) + \left(\frac{1}{Z}\frac{d^2Z}{dz^2} + k_z^2\right) = 0$$

Each term has to be made separately to zero. Therefore, we have

$$\frac{d^2X}{dx^2} + k_x^2 X = 0 \tag{8.8}$$

$$\frac{d^2Y}{dy^2} + k_y^2 Y = 0 \tag{8.9}$$

$$\frac{d^2Z}{dz^2} + k_z^2 Z = 0 \tag{8.10}$$

The boundary condition $\psi(0, y, z) = \psi(a, y, z) = 0$ becomes

$$X(0) = X(a) = 0 \tag{8.11}$$

Similarly, the other boundary conditions lead to a new set of boundary conditions

$$Y(0) = Y(b) = 0 \tag{8.12}$$

$$Z(0) = Z(c) = 0 \tag{8.13}$$

Solutions to (8.8), (8.9) and (8.10), consistent with these boundary conditions, are

$$X(x) = \sin k_x x \qquad k_x = \frac{n_x \pi}{a} \qquad n_x = 1, 2, 3... \tag{8.14}$$

$$Y(y) = \sin k_y y \qquad k_y = \frac{n_y \pi}{b} \qquad n_y = 1, 2, 3... \tag{8.15}$$

$$Z(z) = \sin k_z z \qquad k_z = \frac{n_z \pi}{b} \qquad n_z = 1, 2, 3... \tag{8.16}$$

Therefore, the energy eigenvalues and the energy eigenfunctions are given by

$$u(x, y, z) = N \sin k_x x \sin k_y y \sin k_z z \tag{8.17}$$

$$E_{n_x, n_y, n_z} = \frac{\hbar^2}{2m} \pi^2 \left[\frac{n_x^2}{a^2} + \frac{n_y^2}{b^2} + \frac{n_z^2}{c^2} \right] \tag{8.18}$$

The wave function $u(x, y, z)$ can be normalized by demanding

$$\int_0^a dx \int_0^b dy \int_0^c dz \, |u(x, y, z)|^2 = 1$$

The normalization constant N is found to be

$$N = \left(\frac{8}{abc} \right)^{1/2}$$

The energy eigenfunction is

$$u(x, y, z) = \left(\frac{8}{abc} \right)^{1/2} \sin \frac{n_x \pi x}{a} \sin \frac{n_y \pi y}{b} \sin \frac{n_z \pi z}{c} \tag{8.19}$$

8.1.1 Particle in a Cubical Box

For a cubical box, $a = b = c = L$

Then, the energy eigenfunctions and eigenvalues become

$$E = \frac{\hbar^2}{2m} \frac{\pi^2}{L^2} (n_x^2 + n_y^2 + n_z^2) \tag{8.20}$$

and

$$u(x, y, z) = \left(\frac{2}{L} \right)^{3/2} \sin \frac{n_x \pi x}{L} \sin \frac{n_y \pi y}{L} \sin \frac{n_z \pi z}{L} \tag{8.21}$$

The energy levels are generally degenerate energy levels. Some of the energy levels and the corresponding degeneracies are listed below:

Table 8.1 Some of the energy levels and the corresponding degeneracies

Energy	**(n_x, n_y, n_z)**	**Degeneracy**
$\frac{\hbar^2}{2m}\cdot\frac{3\pi^2}{L^2}$	(1, 1, 1)	1
$\frac{\hbar^2}{2m}\cdot\frac{6\pi^2}{L^2}$	(1, 1, 2), (2, 1, 1), (1, 2, 1)	3
$\frac{\hbar^2}{2m}\cdot\frac{9\pi^2}{L^2}$	(1, 2, 2), (2, 1, 2), (2, 2, 1)	3
$\frac{\hbar^2}{2m}\cdot\frac{12\pi^2}{L^2}$	(2, 2, 2)	1
$\frac{\hbar^2}{2m}\cdot\frac{14\pi^2}{L^2}$	(1, 2, 3), (1, 3, 2), (3, 2, 1), (2, 1, 3),(3, 1, 2), (2, 3, 1)	6

8.2 HARMONIC OSCILLATOR IN THREE DIMENSIONS

The Schrödinger equation for a harmonic oscillator in three dimensions (most general case) is

$$\left[-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}+\frac{1}{2}m\omega_1 x^2-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial y^2}+\frac{1}{2}m\omega_2 y^2-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial z^2}+\frac{1}{2}m\omega_3{}^2 z^2\right]u(x,y,z)=Eu(x,y,z) \tag{8.22}$$

Let us write $u(x, y, z)$ as $X(x)\ Y(y)\ Z(z)$. The Schrödinger equation becomes

$$\left[-\frac{\hbar^2}{2m}\frac{d^2X}{dx^2}+\frac{1}{2}m\omega_1^2x^2X\right]YZ+\left[-\frac{\hbar^2}{2m}\frac{d^2Y}{dy^2}+\frac{1}{2}m\omega_2^2y^2Y\right]XZ+\left[-\frac{\hbar^2}{2m}\frac{d^2Z}{dz^2}+\frac{1}{2}m\omega_3^2z^2Z\right]XY=EXYZ$$

Divide throughout by XYZ and write E as $E=\varepsilon_1+\varepsilon_2+\varepsilon_3$. We now get

$$\begin{aligned}&\frac{1}{X}\left[-\frac{\hbar^2}{2m}\frac{d^2X}{dx^2}+\frac{1}{2}m\omega_1^2X-\varepsilon_1X\right]+\frac{1}{Y}\left[-\frac{\hbar^2}{2m}\frac{d^2Y}{dy^2}+\frac{1}{2}m\omega_2^2Y-\varepsilon_2Y\right]\\&\quad+\frac{1}{Z}\left[-\frac{\hbar^2}{2m}\frac{d^2Z}{dz^2}+\frac{1}{2}m\omega_3^2Z-\varepsilon_3Z\right]=0\end{aligned} \tag{8.23}$$

Setting each term separately to zero, we get

$$\left[-\frac{\hbar^2}{2m}\frac{d^2X}{dx^2}+\frac{1}{2}m\omega_1^2X\right]=\varepsilon_1X \tag{8.24}$$

$$\left[-\frac{\hbar^2}{2m}\frac{d^2Y}{dy^2}+\frac{1}{2}m\omega_2^2Y\right]=\varepsilon_2Y \tag{8.25}$$

$$\left[-\frac{\hbar^2}{2m}\frac{d^2Z}{dz^2}+\frac{1}{2}m\omega_3^2Z\right]=\varepsilon_3 Z \tag{8.26}$$

It is obvious that we have

$$\varepsilon_1=\left(n_1+\frac{1}{2}\right)\hbar\omega_1,\ \varepsilon_2=\left(n_2+\frac{1}{2}\right)\hbar\omega_2,\ \varepsilon_3=\left(n_3+\frac{1}{2}\right)\hbar\omega_3$$

$$\therefore E=\left(n_1+\frac{1}{2}\right)\hbar\omega_1+\left(n_2+\frac{1}{2}\right)\hbar\omega_2+\left(n_3+\frac{1}{2}\right)\hbar\omega_3 \tag{8.27}$$

The energy eigenfunction is

$$u(x,y,z)=N_1N_2N_3e^{-((\alpha_1x)^2+(\alpha_2y)^2+(\alpha_3z)^2)/2}H_{n_1}(\alpha_1x)H_{n_2}(\alpha_2x)H_{n_3}(\alpha_3x) \tag{8.28}$$

8.3 ISOTROPIC HARMONIC OSCILLATOR (CARTESIAN COORDINATES)

Let us consider a special case where the force constants k_1, k_2 and k_3 are the same. This is known as isotropic harmonic oscillator for which $\omega_1=\omega_2=\omega_3=\omega$. Therefore, the energy eigenvalues E is given by

$$E=\left(n_1+n_2+n_3+\frac{3}{2}\right)\hbar\omega \tag{8.29}$$

The states are characterized by a set of three integers (n_1, n_2, n_3). The sum $n_1+n_2+n_3=n$ can be written as n.

$$\therefore E=\left(n+\frac{3}{2}\right)\hbar\omega \qquad n=n_1+n_2+n_3 \tag{8.30}$$

The ground state corresponds to $n_1=n_2=n_3=0$. Therefore, the ground state energy is

$$E=\frac{3}{2}\hbar\omega \tag{8.31}$$

So the ground state is non-degenerate; i.e., there is only one ground state with minimum energy $3/2\,\hbar\omega$. States with higher energies are degenerate. For instance, for $n=1$, the possible values of (n_1, n_2, n_3) are (1, 0, 0), (0, 1, 0) and (0, 0, 1). For $n=2$, the possible values of (n_1, n_2, n_3) are (2, 0, 0), (1, 1, 0), (1, 0, 1), (0, 2, 0), (0, 1, 1) and (0, 0, 2). It can be shown that for a given n, there are $\frac{1}{2}(n+1)(n+2)$ states.

Example 8.1 Show that for an isotropic harmonic oscillator, the energy level $E=(n+3/2)\hbar\omega$ is $1/2(n+1)(n+2)$ fold degenerate.

Solution: For a given n, one has to find the total number of combinations of three integers n_1, n_2 and n_3 such that $n_1+n_2+n_3=n$.

For a given n, if one fixes n_1 and n_2, n_3 naturally gets fixed. So this problem reduces to find in how many ways n_1 and n_2 can be varied consistent with a given n. Let us enumerate the possible values of n_1 and n_2 consistent with a given n.

Table 8.2 Counting the degeneracy

Values of n_1	Possible values of n_2	n_1, n_2, n_3	Degeneracy
n	0	$(n, 0, 0)$	1
$n-1$	1, 0	$(n-1, 1, 0)$, $(n-1, 0, 1)$	2
$n-2$	2, 1, 0	$(n-2, 2, 0)$, $(n-2, 1, 1)$, $(n-2, 0, 2)$	3
$n=0$	$n, n-1, \ldots, 1, 0$		$n+1$

Total number of degenerate states D is given by

$$D = 1 + 2 + \ldots (n+1) = \frac{1}{2}(n+1)(n+2)$$

8.4 CENTRAL POTENTIALS: GENERAL PROPERTIES

If $V(r, \theta, \phi)$ is independent of θ and ϕ, then it is said to be spherically symmetric potential.

i.e., $$V(r, \theta, \phi) = V(r)$$

Such potentials are called central potentials. To describe a particle in central forces, the best coordinate system is spherically polar coordinates.

The Laplacian operator in spherical polar coordinates is given by

$$\nabla^2 = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + \frac{1}{r^2 \sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{r^2}\frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2} \tag{8.32}$$

The Schrödinger equation in spherical polar coordinates is given by

$$-\frac{\hbar^2}{2m}\left[\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial u}{\partial r}\right) + \frac{1}{r^2}\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial u}{\partial\theta}\right) + \frac{1}{r^2}\frac{1}{\sin^2\theta}\frac{\partial^2 u}{\partial\phi^2}\right] + V(r)u = Eu$$

Let us recall that the angular momentum operator L^2 is given by

$$L^2 = -\hbar^2\left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2}\right]$$

The Schrödinger equation now becomes

$$-\frac{\hbar^2}{2m}\left[\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial u}{\partial r}\right) - \frac{L^2}{\hbar^2 r^2}u\right] + V(r)u = Eu \tag{8.33}$$

It can be easily shown that for any central potential, the Hamiltonian commutes with L^2 and L_z.

i.e., $$[H, L^2] = 0,\ [H, L_z] = 0$$

Example 8.2 Show that the Hamiltonian for central potential $V(r)$ commutes with L^2 and L_z.

Solution:

$$H = -\frac{\hbar^2}{2m}\left[\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) - \frac{L^2}{\hbar^2 r^2}\right] + V(r)$$

$$= -\frac{\hbar^2}{2m}\left[\frac{\partial}{\partial r^2} + \frac{2}{r}\frac{\partial}{\partial r} - \frac{L^2}{\hbar^2 r^2}\right] + V(r)$$

$$\left[H, L^2\right] = \left[-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial r^2} - \frac{\hbar^2}{2m}\frac{2}{r}\frac{\partial}{\partial r} + \frac{L^2}{2mr^2} + V(r), L^2\right]$$

$$= \left[-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial r^2} - \frac{\hbar^2}{2m}.\frac{2}{r}\frac{\partial}{\partial r}, L^2\right] + \left[\frac{L^2}{2mr^2}, L^2\right] + [V(r), L^2]$$

Since L^2 depends only on θ and ϕ, each of the first two commutates with L^2.

$$\therefore [H, L^2] = 0$$

In the same way, we can prove that $[H, L_z] = 0$.

8.5 RADIAL EQUATION

The observables H, L^2 and L_z together form a set of compatible observables. Therefore, these operators should have common eigenfunctions.

Let us recall that

$$L^2 Y_{lm}(\theta,\phi) = l(l+1)\hbar^2 Y_{lm}(\theta,\phi)$$

$$L_z Y_{lm}(\theta,\phi) = m_l \hbar Y_{lm}(\theta,\phi)$$

Therefore, the eigenfunctions of H can be written as

$$u(r, \theta, \phi) = R(r)\, Y_{lm}(\theta, \phi)$$

Substituting this expression in (8.33), we get

$$-\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial r^2} + \frac{2}{r}\frac{\partial}{\partial r}\right)R(r)Y_{lm}(\theta,\phi) + \frac{L^2}{2mr^2}R(r)Y_{lm}(\theta,\phi) + V(r)R(r)Y_{lm}(\theta,\phi) = E\,R(r)Y_{lm}(\theta,\phi)$$

$$-\frac{\hbar^2}{2m}\left[\frac{\partial^2}{\partial r^2} + \frac{2}{r}\frac{\partial}{\partial r}\right]R(r)Y_{lm} + \frac{l(l+1)\hbar^2}{2mr^2}R(r)Y_{lm} + V(r)R(r)Y_{lm} = E\,R(r)Y_{lm}$$

$$-\frac{\hbar^2}{2m}\left[\frac{d^2R}{dr^2} + \frac{2}{r}\frac{dR}{dr}\right] + \frac{l(l+1)\hbar^2}{2mr^2}R(r) + V(r)R = ER(r)$$

Note that R depends on E and l. So we can rewrite this equation as

$$\frac{d^2R_{El}}{dr^2} + \frac{2}{r}\frac{dR_{El}}{dr} + \frac{2m}{\hbar^2}\left[E - V(r) - \frac{l(l+1)}{2mr^2}\hbar^2\right]R_{El}(r) = 0$$

Generally, quantum numbers n and l are used to specify the radial wave function instead of E and l. Therefore, the above equation becomes

$$\frac{d^2R_{nl}}{dr^2}+\frac{2}{r}\frac{dR_{nl}}{dr}+\frac{2m}{\hbar^2}\left[E-V(r)-\frac{l(l+1)}{2mr^2}\hbar^2\right]R_{nl}(r)=0 \tag{8.34}$$

This equation is known as radial equation.

8.6 CENTRIFUGAL POTENTIAL

The radial wave equation looks like one-dimensional Schrödinger equation in an effective potential V_{eff}, which is given by

$$V_{\text{eff}}(r)=V(r)+\frac{\hbar^2}{2mr^2}l(l+1) \tag{8.35}$$

If we try to understand $V_{eff}(r)$ in terms of force and potential at classical level, we can write it as

$$F=-\frac{d}{dr}\left[V(r)+\frac{L^2}{2mr^2}\right]$$

$$-\frac{d}{dr}\frac{L^2}{2mr^2}=\frac{L^2}{mr^3}=\frac{m^2v^2r^2}{mr^3}=\frac{mv^2}{r}$$

The term mv^2/r represents centrifugal force, and so $L^2/2mr^2=l(l+1)\hbar^2/2mr^2$ is interpreted as centrifugal potential. It represents a repulsive interaction. Normally, $V(r)$ is an attractive potential, and $l(l+1)\hbar^2/2mr^2$ is a repulsive potential. So the effective potential looks like the curve in Fig. 8.2, where r_1 and r_2 are the classical turning points.

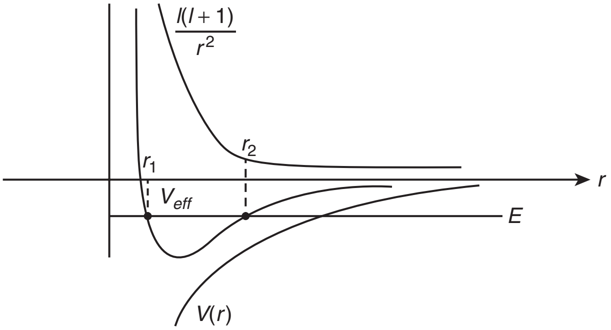


Fig. 8.2 Effective potential

8.7 RADIAL PROBABILITY DISTRIBUTION FUNCTION

We know that the probability of finding the particle in a volume $d^3\mathbf{r}$ centred around $\mathbf{r}$ is

$$\begin{aligned}P(\mathbf{r})d^3\mathbf{r}&=u*(\mathbf{r})u(\mathbf{r})d^3\mathbf{r}\\&=u_{nl}^*(\mathbf{r})u_{nl}(\mathbf{r})r^2\sin\theta d\theta d\phi\\&=|R_{nl}(r)|^2|Y_{lm}(\theta,\phi)|^2\ r^2\sin\theta d\theta d\phi\end{aligned}$$

Integrating over the solid angle $d\Omega$, we get

$$\begin{aligned} P_{nl}(r)dr &= \int_0^{\pi} \sin\theta d\theta \int_0^{2\pi} d\phi \, |Y_{lm}(\theta,\phi)|^2 |R_{nl}(r)|^2 \, r^2 dr \\ &= |R_{nl}(r)|^2 \, r^2 dr \end{aligned} \tag{8.36}$$

This term can be interpreted as probability of finding the particle between r and $r + dr$.

Since $\int P(\mathbf{r})d^3\mathbf{r} = 1$, we get

$$\int_0^{\infty} P_{nl}(r)dr = \int_0^{\infty} |R_{nl}(r)|^2 \, r^2 dr = 1 \tag{8.37}$$

This is the normalization condition for the radial wave function.

8.7.1 Boundary Conditions for $R_{nl}(r)$

Let us recall that in one-dimensional Schrödinger equation, the boundary conditions on the wave function $u(x)$ are determined by the requirement that the wave function should be single valued and finite at all points. This is necessary for probability interpretation and the normalization condition. For the radial wave functions also, we impose the same requirements. The normalization condition for $R_{nl}(r)$ is

$$\int_0^{\infty} r^2 \, |R_{nl}(r)|^2 \, dr = 1$$

The convergence of this integral demands that $rR_{nl}(r) \to 0$ as $r \to \infty$. It is to be noted that this is true for bound states.

There is one more boundary condition, namely the behaviour of the wave function near $r = 0$.

Let us write the solution $R_{nl}(r)$ as

$$R_{nl}(r) = r^s f(r) = r^s \sum_n a_n r^n, \qquad \text{with } a_0 \neq 0$$

Substituting this solution in the Equation (8.4) and then dividing by r^{s-2}, we get

$$r^2 \frac{d^2 f}{dr^2} + 2(s+1)r\frac{df}{dr} + \left[s(s+1) - l(l+1)\right] f(r) + \frac{2m}{\hbar^2} r^2 E_{nl} f(r) - \frac{2m}{\hbar^2} V(r) r^2 f(r) = 0$$

As $r \to 0$, this equation becomes $[s(s+1) - l(l+1]f(r) \simeq 0$

$$\begin{aligned} [s(s+1) - l(l+1)][a_0 + a_1 r + a_2 r^2 + ...] &\simeq 0 \\ \left[s(s+1) - l(l+1)\right] a_0 &\simeq 0 \end{aligned} \tag{8.38}$$

Here we have assumed that $r^2 V(r) \to 0$ as $r \to 0$.

The Equation (8.38) can be rewritten as

$$\begin{aligned} (s - l)\,(s + l + 1) &= 0 \\ \therefore s = l \text{ or } s &= -(l + 1) \end{aligned} \tag{8.39}$$

i.e., as $r \to 0$, the solution $R_{nl}(r)$ behaves like r^l or $r^{-(l+1)}$. Obviously, the acceptable behaviour is r^l as $r \to 0$.

So the boundary condition for the radial wave function $R_{nl}(r)$ is

1. $R_{nl}(r)$ is single valued, finite and continuous
2. $R_{nl}(r) \to r^l$ as $r \to 0$
3. For bound states, $rR_{nl}(r) \to 0$ as $r \to \infty$

8.8 RADIAL MOMENTUM IN QUANTUM MECHANICS

In classical mechanics, the radial momentum is the component of momentum along the radial direction. It is given by

$$p_r = \mathbf{r} \cdot \mathbf{p} = \frac{\mathbf{r}}{r} \cdot \mathbf{p} \qquad \text{(classical)}$$

In classical physics, $\frac{\mathbf{r}}{r} \cdot \mathbf{p}$ and $\mathbf{p} \cdot \frac{\mathbf{r}}{r}$ are the same. But in quantum mechanics, $\mathbf{r}/r$ and $\mathbf{p}$ are operators and they do not commute. Therefore, what is the correct expression for the operator p_r, whether $\mathbf{p} \cdot \frac{\mathbf{r}}{r}$ or $\frac{\mathbf{r}}{r} \cdot \mathbf{p}$?

The standard way of defining p_r is

$$p_r = \frac{1}{2}\left(\frac{\mathbf{r}}{r} \cdot \mathbf{p} + \mathbf{p} \cdot \frac{\mathbf{r}}{r}\right) \tag{8.40}$$

Example 8.3 Show that the radial momentum operator $\frac{1}{2}\left(\frac{\mathbf{r}}{r} \cdot \mathbf{p} + \mathbf{p} \cdot \frac{\mathbf{r}}{r}\right) = \frac{1}{r}[\mathbf{r} \cdot \mathbf{p} - i\hbar]$.

Solution: Consider $\frac{1}{2}\left(\frac{\mathbf{r}}{r} \cdot \mathbf{p} + \mathbf{p} \cdot \frac{\mathbf{r}}{r}\right)\psi$, where ψ is an arbitrary wave function.

$$\frac{1}{2}\left(\frac{\mathbf{r}}{r} \cdot \mathbf{p} + \mathbf{p} \cdot \frac{\mathbf{r}}{r}\right)\psi = \frac{1}{2r}[xp_x + yp_y + zp_z]\psi + \frac{1}{2}\left[p_x\left(\frac{x}{r}\psi\right) + p_y\left(\frac{y}{r}\psi\right) + p_z\left(\frac{z}{r}\psi\right)\right]$$

Note

$$\begin{aligned} p_x\left(\frac{x}{r}\psi\right) &= -i\hbar\frac{\partial}{\partial x}\left(\frac{x}{r}\psi\right) \\ &= -i\hbar\frac{1}{r}\psi + i\hbar\frac{x^2}{r^3}\psi + \frac{x}{r}p_x\psi \end{aligned}$$

$$p_y\left(\frac{x}{r}\psi\right) = -i\hbar\frac{1}{r}\psi + i\hbar\frac{y^2}{r^3}\psi + \frac{y}{r}p_y\psi$$

$$p_z\left(\frac{z}{r}\psi\right) = -i\hbar\frac{1}{r}\psi + i\hbar\frac{z^2}{r^3}\psi + \frac{z}{r}p_z\psi$$

$$\frac{1}{r}\mathbf{p} \cdot \mathbf{r}\psi = \frac{1}{2}\left[-3i\hbar\frac{1}{r} \cdot \psi + i\hbar\frac{(x^2+y^2+z^2)}{r^3}\psi + \frac{\mathbf{r}}{r} \cdot \mathbf{p}\psi\right]$$

$$\therefore \quad p_r\psi = \frac{1}{2}\left(\frac{\mathbf{r}}{r} \cdot \mathbf{p} + \mathbf{p} \cdot \frac{\mathbf{r}}{r}\right)\psi = \frac{1}{r}[\mathbf{r} \cdot \mathbf{p} - i\hbar]\psi$$

Since this equation is true for an arbitrary function ψ, the operator p_r is given by

$$p_r = \frac{1}{r}[\mathbf{r} \cdot \mathbf{p} - i\hbar] \tag{8.41}$$

Example 8.4 Show that $p_r^2 f(r) = -\hbar^2\left[\frac{d^2 f}{dr^2} + \frac{2}{r}\frac{df}{dr}\right]$,

where $f(r)$ is an arbitrary function of radial coordinate r.

Solution:

$$
\begin{aligned}
p_r f(r) &= \frac{1}{r}\left[\mathbf{r}\cdot\mathbf{p} - i\hbar\right] f(r)\\
\mathbf{r}\cdot\mathbf{p} f(r) &= x p_x f(r) + y p_y f(r) + z p_z f(r)\\
&= -i\hbar\left[x\frac{\partial f(r)}{\partial x} + y\frac{\partial f(r)}{\partial y} + z\frac{\partial f(r)}{\partial z}\right]\\
&= -i\hbar\left[x\cdot\frac{df}{dr}\cdot\frac{\partial r}{\partial x} + y\frac{df}{dr}\cdot\frac{\partial r}{dy} + z\frac{df}{dr}\cdot\frac{\partial r}{\partial z}\right]\\
&= -i\hbar\frac{df}{dr}\left[x\cdot\frac{x}{r} + y\cdot\frac{y}{r} + z\cdot\frac{z}{r}\right]\\
&= -i\hbar r\frac{df}{dr}\\
\therefore\ p_r f(r) &= -i\hbar\left[\frac{df}{dr} + \frac{f}{r}\right] = -i\hbar\left[\frac{d}{dr} + \frac{1}{r}\right] f(r)\\
\therefore\ p_r^2 f(r) &= -\hbar^2\left[\frac{d}{dr} + \frac{1}{r}\right]\left[\frac{df}{dr} + \frac{1}{r} f(r)\right]\\
&= -\hbar^2\left[\frac{d^2 f}{dr^2} + \frac{2}{r}\frac{df}{dr}\right]
\end{aligned}
$$

8.9 HAMILTONIAN IN TERMS OF RADIAL MOMENTUM

The Hamiltonian for a particle in spherical polar coordinates is given by

$$
\begin{aligned}
H &= -\frac{\hbar^2}{2m}\nabla^2 + V(r)\\
&= -\frac{\hbar^2}{2m}\left[\frac{d^2}{dr^2} + \frac{2}{r}\frac{d}{dr} - \frac{L^2}{\hbar^2 r^2}\right] + V(r)\\
&= \frac{p_r^2}{2m} + \frac{L^2}{2mr^2} + V(r) \qquad (8.42)
\end{aligned}
$$

Here, we have assumed the potential to be central potential.

8.10 FREE PARTICLE IN SPHERICAL POLAR COORDINATES

The Hamiltonian for a free particle is

$$
\begin{aligned}
H &= -\frac{\hbar^2}{2m}\nabla^2\\
&= \frac{p_r^2}{2m} + \frac{L^2}{2mr^2} \qquad (8.43)
\end{aligned}
$$

The energy eigenvalue equation is

$$Hu(r, \theta, \phi) = Eu(r, \theta, \phi)$$

The operators which commute with H are L^2 and L_z.

i.e., $$[H, L^2] = 0, [H, L_z] = 0, [L^2, L_z] = 0$$

Therefore, H, L^2 and L_z have got the same eigenfunctions. Since the spherical harmonics $Y_{lm}(\theta, \phi)$ are eigenfunctions L^2 and L_z, the wave function $u(r, \theta, \phi)$ can be written as

$$u(r, \theta, \phi) = R(r)\, Y_{lm}(\theta, \phi)$$

This leads to the radial equation

$$\begin{gathered} \frac{d^2R}{dr^2} + \frac{2}{r}\frac{dR}{dr} + \frac{2m}{\hbar^2}\left[E - \frac{\hbar^2}{2mr^2} l(l+1)\right]R = 0 \\ \frac{d^2R}{dr^2} + \frac{2}{r}\frac{dR}{dr} + \left[\frac{2mE}{\hbar^2} - \frac{1}{r^2} l(l+1)\right]R = 0 \end{gathered} \tag{8.44}$$

Since $E = \hbar^2 k^2/2m$, we write the above equation as

$$\frac{d^2R_{kl}}{dr^2} + \frac{2}{r}\frac{dR_{kl}}{dr} + \left[k^2 - \frac{1}{r^2} l(l+1)\right]R_{kl} = 0$$

Define $\rho = kr$

$$\frac{d^2R_{kl}}{d\rho^2} + \frac{2}{\rho}\frac{dR_{kl}}{d\rho} + \left[1 - \frac{l(l+1)}{\rho^2}\right]R_{kl} = 0 \tag{8.45}$$

This is a well-known equation whose solutions are spherical Bessel functions $j_l(kr)$ and Neumann functions $n_l(kr)$.

$$\therefore R_{kl}(r) = A_l j_l(kr) + B_l n_l(kr)$$

Some of the spherical Bessel functions and Neumann functions are listed below:

$$j_o(kr) = \frac{\sin kr}{kr} \qquad n_o(kr) = -\frac{\cos kr}{kr}$$

$$j_1(kr) = \frac{\sin kr}{k^2r^2} - \frac{\cos kr}{kr} \qquad n_1(kr) = -\frac{\cos kr}{k^2r^2} - \frac{\sin kr}{kr}$$

$$j_2(kr) = \left(\frac{3}{k^3r^3} - \frac{1}{kr}\right)\sin kr - \frac{3}{k^2r^2}\cos kr \qquad n_3(kr) = -\left(\frac{3}{k^2r^3} - \frac{1}{kr}\right)\cos kr - \frac{3}{k^2r^2}\sin kr$$

The functions $n_l(kr)$ are not finite at $r = 0$. So we choose $B_l = 0$.

Therefore, the free particle wave function $u_{klm}(r, \theta, \phi)$ is given by

$$u_{klm}(r, \theta, \phi) = A_l j_l(kr) Y_{lm}(\theta, \phi) \tag{8.46}$$

8.10.1 Difference Between $e^{i\frac{\mathbf{p}\cdot\mathbf{r}}{\hbar}}$ and $j_l(kr)Y_{lm}(\theta, \phi)$

Both $e^{i\mathbf{k}\cdot\mathbf{r}}$ and $j_l(kr)Y_{lm}(\theta, \phi)$ are eigenfunctions of the Hamiltonian $H = -\hbar^2/2m\nabla^2$. But they are eigenfunctions of two different sets of compatible dynamical variables. One set is $\{H, \mathbf{p}\}$ and another set is $\{H, L^2, L_z\}$. But $\mathbf{p}$ and L^2 do not commute with each other. So we have

$$\mathbf{p}_{op}e^{i\mathbf{k}\cdot\mathbf{r}} = \hbar\mathbf{k}e^{i\mathbf{k}\cdot\mathbf{r}}$$

$$He^{i\mathbf{k}\cdot\mathbf{r}} = \frac{\hbar^2k^2}{2m}e^{i\mathbf{k}\cdot\mathbf{r}}$$

$$L^2 j_l(kr)Y_{lm}(\theta,\phi) = l(l+1)\hbar^2 j_l(kr)Y_{lm}(\theta,\phi)$$

$$Hj_l(kr)Y_{lm}(\theta,\phi) = \frac{\hbar^2k^2}{2m} j_l(kr)Y_{lm}(\theta,\phi).$$

$e^{i\mathbf{k}\cdot\mathbf{r}}$ is an eigenfunction of momentum operator $\mathbf{p}_{op}$; but it is not an eigenfunction L^2. Therefore, the state represented by $e^{i\mathbf{k}\cdot\mathbf{r}}$ has a definite momentum, but it has not a well-defined angular momentum.

On the other hand, $j_l(kr)Y_{lm}(\theta, \phi)$ is an eigenfunction of L^2 and L_z; but it is not an eigenfunction of momentum operator $\mathbf{p}_{op}$. Therefore, the state represented by $j_l(kr)Y_{lm}(\theta, \phi)$ does not have a well-defined momentum.

It can be shown that the plane wave function is given by $e^{i\,\mathbf{k}\cdot\mathbf{r}} = 4\pi\sum_{l,m} i^l j_l(kr)Y_{lm}(\theta_k,\phi_m)Y^*_{lm}(\theta,\phi)$

where $\mathbf{k} = k\sin\theta_k\cos\phi_k\hat{\mathbf{x}} + k\sin\theta_k\sin\phi_k\hat{\mathbf{y}} + k\cos\theta_k\hat{\mathbf{z}}$. In the special coordinate system, where $\mathbf{k}$ is chosen along Z axis, we have

$$e^{ikr\cos\theta} = 4\pi\sum_{l,m} i^l\left[\frac{(2l+1)}{4\pi}\right]^{1/2} j_l(kr)Y_{l0}(\theta,\phi)$$

$$= \sum i^l(2l+1)j_l(kr)P_l(\cos\theta)$$

8.11 PARTICLE IN A SPHERICALLY SYMMETRIC BOX

The potential $V(r)$ is given by

$$V(r) = \begin{cases} 0 & 0 < r < a \\ \infty & \text{otherwise} \end{cases}$$

i.e., the potential is zero for a particle inside a sphere of radius a but infinite outside the sphere. So the particle has to be confined to this region. The radial equation is

$$\frac{d^2R}{dr^2} + \frac{2}{r}\frac{dR}{dr} + \left[\frac{2mE}{\hbar^2} - \frac{l(l+1)}{r^2}\right]R(r) = 0 \qquad 0 < r < a \tag{8.47}$$

$$R(r) = 0 \qquad\qquad r > a \tag{8.48}$$

Let us define k^2 as $2mE/\hbar^2 = k^2$

The Equation (8.47) now becomes

$$\frac{d^2R}{dr^2} + \frac{2}{R}\frac{dR}{dr} + \left[k^2 - \frac{l(l+1)}{r^2}\right]R(r) = 0 \tag{8.49}$$

This is the same as the Equation (8.45). Therefore, the solution is given by

$$R_{kl}(r) = \begin{cases} A_l j_l(kk) & 0 < r < a \\ 0 & r > a \end{cases} \tag{8.50}$$

k is now not a continuous variable; k is restricted to a set of discrete values determined by $j_l(ka) = 0$.

The wave function $u(r, \theta, \phi)$ is given

$$u(r,\theta,\phi) = \begin{cases} A_l j_l(kr)Y_{lm}(\theta,\phi) & 0 < r < a \\ 0 & r > a \end{cases} \tag{8.51}$$

Example 8.5 Determine the energy eigenvalue for a particle in a spherically symmetric box for $l = 0$.

Solution: The boundary condition for the radial wave function is $j_l(ka) = 0$.

$$j_1(ka) = \frac{\sin ka}{ka}$$

$$\therefore ka = n\pi \qquad\qquad n = 1,2,3...$$

$$\therefore E_{no} = \frac{\hbar^2k^2}{2m} = \frac{\hbar^2n^2\pi^2}{2ma^2} \qquad n = 1,2,3...$$

Example 8.6 Show that the energy eigenvalue of a particle in a spherically symmetric box for $l = 1$ is determined by tan $ka = ka$.

Solution: The boundary condition for the radial wave function is

$$j_l(ka) = 0$$

$$\therefore \frac{\sin ka}{k^2a^2} - \frac{\cos ka}{ka} = 0$$

$$\therefore \tan ka = ka$$

The solution is provided by the intersection of the two curves

$$y = ka \text{ and } y = \tan ka$$

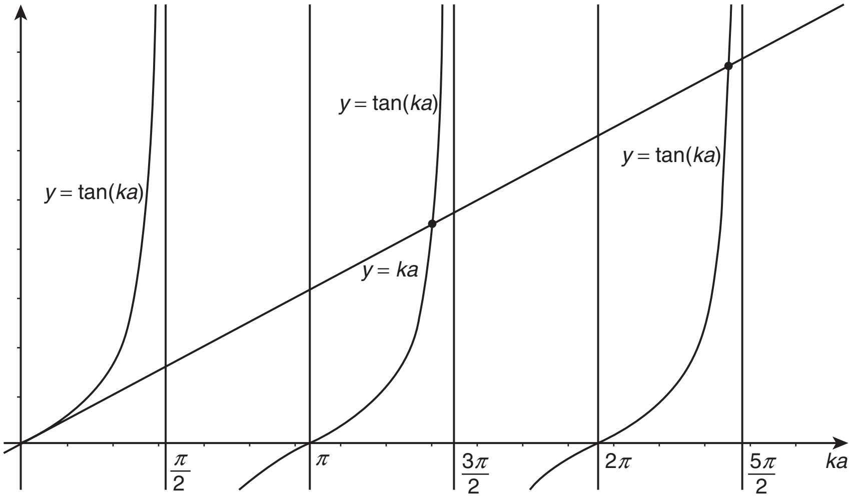


Fig. 8.3 Graphical solution for a particle in a spherically symmetric box

8.12 BOUND-STATE SOLUTION FOR SPHERICALLY SYMMETRIC POTENTIAL WELL

The spherically symmetric potential well is given by

$$V(r) = \begin{cases} -V_o & 0 < r < a \\ 0 & r > a \end{cases}$$

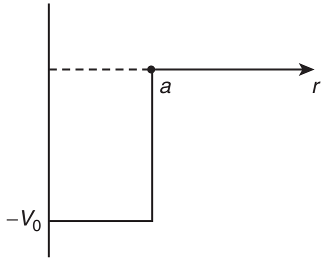


Fig. 8.4 Spherically symmetric square well

For bound states, E is negative and $|E| > V_0$.

The radial equation is given by

$$\frac{d^2R_I}{dr^2} + \frac{2}{r}\frac{dR_I}{dr} + \left[\frac{2m}{\hbar^2}(-|E| + V_0) - \frac{l(l+1)}{r^2}\right]R_I = 0 \qquad 0 < r < a \tag{8.52}$$

$$\frac{d^2R_{II}}{dr^2} + \frac{2}{r}\frac{dR_{II}}{dr} - \left[\frac{2m}{\hbar^2}|E| + \frac{l(l+1)}{r^2}\right]R_{II} = 0 \qquad r > a \tag{8.53}$$

Let us define α^2 and β^2 as

$$\alpha^2 = \frac{2m}{\hbar^2}(V_0 - |E|) \quad \beta^2 = \frac{2m}{\hbar^2}|E|$$

The radial equation becomes

$$\frac{d^2R_I}{dr^2} + \frac{2}{r}\frac{dR_I}{dr} + \left[\alpha^2 - \frac{l(l+1)}{r^2}\right]R_I = 0 \quad 0 \le r \le a \tag{8.54}$$

$$\frac{d^2R_{II}}{dr^2} + \frac{2}{r}\frac{dR_{II}}{dr} + \left[-\beta^2 - \frac{l(l+1)}{r^2}\right]R_{II} = 0 \quad r \ge a \tag{8.55}$$

Solution to Equations (8.54) and (8.55) is given by

$$R_I = A_l j_l(\alpha r) \quad 0 \le r \le a \tag{8.56}$$

$$R_{II} = B_l j_l(i\beta r) + C_l n_l(i\beta r) \quad r \ge a \tag{8.57}$$

For a bound-state solution, we would like to have the radial wave function to fall rapidly to zero for $r > a$. This behaviour is not reflected by $j_l(i\beta r)$ or $n_l(i\beta r)$ since they become infinite. A better solution reflecting this behaviour can be obtained by a linear combination $j_l(i\beta r)$ and $n_l(i\beta r)$. Such linear combinations are provided by Hankel functions $h_l^{(+)}$ and defined as

$$h_l^{(+)}(\rho) = j_l(\rho) + i n_l(\rho) \quad \text{and} \quad h_l^{(-)}(\rho) = j_l(\rho) - i n_l(\rho) \tag{8.58}$$

Using the spherical Bessel and Neumann functions given in the Section 8.10, we can work out some of the spherical Hankel functions.

$$\begin{aligned} h_0^{(\pm)}(\rho) &= \pm\left(\frac{i}{\rho}\right)e^{\pm i\rho} \\ h_1^{(\pm)}(\rho) &= \left(-\frac{1}{\rho} \pm \frac{i}{\rho^2}\right)e^{\pm i\rho} \end{aligned} \tag{8.59}$$

$$h_2^{(\pm)}(\rho) = \left(\pm\frac{i}{\rho} - \frac{3}{\rho^2} \mp \frac{3i}{\rho^3}\right)e^{\pm i\rho} \tag{8.60}$$

$h_l^{(\pm)}(\rho)$ have the general form $f(\rho)e^{\pm i\rho}$.

$$\therefore h_l^{(+)}(i\beta r) = f(i\beta r)e^{-\beta r} \tag{8.61}$$

and

$$h_l^{(-)}(i\beta r) = f(i\beta r)e^{\beta r} \tag{8.62}$$

Obviously, its $h_l^{(+)}(i\beta r)$ reflects the required behaviour of the radial wave function. So we have

$$R(r) = \begin{cases} A_l j_l(\alpha r) & 0 < r < a \quad (8.63) \\ B_l h_l^{(+)}(i\beta r) & r > a \quad (8.64) \end{cases}$$

The single value of the radial wave function $R(r)$ and $dR(r)/dr$ requires

$$A_l j_l(\alpha a) = B_l h_l^{(+)}(i\beta a) \tag{8.65}$$

$$\frac{d}{dr} A_l j_l(\alpha r)\Big|_{r=a} = \frac{d}{dr} B_l h_l^{(+)}(i\beta r)\Big|_{r=a}$$

These two equations can be combined to get the following single equation:

$$\frac{1}{j_l(\alpha a)}\frac{d}{dr} j_l(\alpha r)\bigg|_{r=a} = \frac{1}{h_l^{+}(i\beta a)}\frac{d}{dr} h_l^{+}(i\beta r)\bigg|_{r=a}$$

$$\frac{1}{j_l(\alpha a)}\,\alpha\frac{d}{d\rho} j_l(\rho)\bigg|_{\rho=\alpha a} = \frac{1}{h_l^{+}(i\beta a)}\, i\beta\frac{d}{d\rho} h_l^{+}(\rho)\bigg|_{\rho=i\beta a} \tag{8.66}$$

This equation is sufficient to determine the energy eigenvalues of bound states.

8.12.1 Special Case: *l* = 0 Bound-state Solution

This is the simplest case in determining the energy eigenvalue for a particle in spherically symmetric potential well.

For $l = 0$, the radial wave function $R(r)$ is given by

$$R_{\text{I}}(r) = A_0 j_0(\alpha r) = A_0 \frac{\sin \alpha r}{\alpha r} \qquad 0 \le r \le a$$

$$R_{\text{II}}(r) = B_0 h_0^{(+)}(i\beta r) = -\frac{B_0}{\beta r} e^{-\beta r} \qquad r \ge a$$

Making use of the Equation (8.66), we get

$$\frac{\alpha a \cos \alpha a - \sin \alpha a}{\sin \alpha a} = -\beta a - 1$$

$$\therefore\ \alpha a \cot \alpha a = -\beta a \tag{8.67}$$

$$\begin{aligned} \beta^2 a^2 + \alpha^2 a^2 &= \left[\frac{2m}{\hbar^2}(V_0 - |E|) + \frac{2m}{\hbar^2}|E|\right]a^2 \\ &= \frac{2ma^2}{\hbar^2} V_0 \end{aligned} \tag{8.68}$$

The Equations (8.67) and (8.68) are exactly the same as the equations determining energy eigenvalues of bound states with odd parity in the case of square well potential in one dimension.

Let us define X, Y and Z as $X = \alpha a$, $Y = \beta a$ and $Z^2 = 2ma^2/\hbar^2\, V_0$. Then the Equations (8.67) and (8.68) become

$$Y = -\cot X$$

and

$$X^2 + Y^2 = Z^2$$

The solution is determined by the intersection of the curves in the first quadrants as shown in Fig. 8.5.

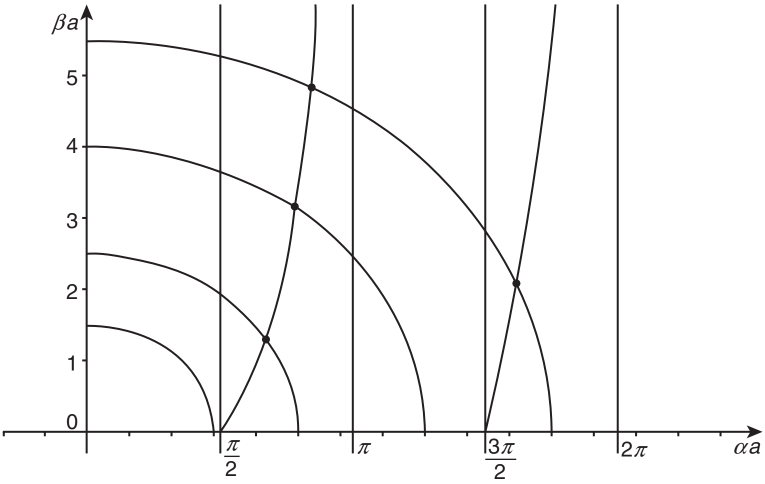


Fig. 8.5 Spherically symmetric square well potential – bound-state solution for $l = 0$

The existence of bound states depends on the value of $Z = (2ma^2V_0/\hbar^2)^{1/2}$. There is no bound state for Z value less than $\pi/2$. There is one bound state for $\pi/2 \le Z \le 3\pi/2$. Other bound states can be determined in the same way.

8.12.2 Condition for the Existence of Bound States for Square Well Potential

Let us now investigate the conditions for the existence of bound states for square well potential as a function of potential depth V_0. This result is important in the context of low energy scattering, which will be discussed in Chapter 15. For square well potential, the bound state energy is determined by the Equation (8.65). So, the existence of a solution to this equation implies the existence of the bound state. For instance, in the case of $l = 0$, this equation cannot be satisfied for values of V_0 less than $\hbar^2\pi^2/8ma^2$.We closely follow Gasiorowicz here.

First, let us consider the case of $l = 0$. From the graphical solution of the Equation (8.67) in Fig. 8.5, no bound state exists if $K_0a = (2mV_0/\hbar^2)^{1/2}a$ is less than $\pi/2$. As V_0 increases, we reach a threshold value of V_0, determined by $K_0a = \pi/2$, at which the formation of a bound state is possible. For the value of V_0 in the range $\pi/2 < K_0a < 3\pi/2$, there exists only one bound state. As V_0 is increased further, when it reaches $K_0a = 3\pi/2$, the existence of second bound state becomes possible. Continuing in the same way, we get the threshold value of V_0 for the nth bound state, which is determined by $K_0a = (2n-1)\pi/2$.

Let us next consider the case $l \ne 0$. When $l \ne 0$, it is difficult to get an analytic expression which can be used to get the condition for the existence of the bound state. However, when the potential V_0 is large, it is possible to get an approximate condition. This approximation depends on the use of asymptotic limit of $j_l(\rho)$, which is valid for $\rho \gg l$. The asymptotic limit of $j_l(\rho)$ is given by

$$j_l(\rho) \xrightarrow[\rho \gg l]{} \frac{\sin\left(\rho - \frac{l\pi}{2}\right)}{\rho}$$

$$\frac{d}{d\rho} j_l(\rho) \xrightarrow[\rho >> l]{} -\frac{\sin\left(\rho - \frac{l\pi}{2}\right)}{\rho^2} + \frac{\cos\left(\rho - \frac{l\pi}{2}\right)}{\rho}$$

or
$$\frac{d}{d\rho} j_l(\rho) \approx \frac{\cos\left(\rho - \frac{l\pi}{2}\right)}{\rho} \qquad \text{for } \rho >> l$$

Substituting these expressions in the Equation (8.66), we get

$$\alpha a \cot\left(\alpha a - \frac{l\pi}{2}\right) - \frac{1}{h_l^+(i\beta a)}\, i\beta a \frac{d}{d\rho} h_l^+(\rho)\Bigg|_{\rho = i\beta a} \approx 0$$

Note the second term does not depend on V_0. When $|E|$ is much less than V_0, the second term can be neglected.

$$\therefore \quad \alpha a \cot\left(\alpha a - \frac{l\pi}{2}\right) \approx 0$$

This is possible provided $\cot\left(\alpha a - \frac{l\pi}{2}\right) \approx 0$

or
$$\alpha a - \frac{l\pi}{2} = \left(n + \frac{1}{2}\right)\pi$$

This is the condition for the existence of bound states for square well potential when $\alpha a >> l$ and $V_0 >> |E|$.

8.13 SCATTERING STATE SOLUTION FOR SQUARE WELL POTENTIAL

In the previous section, we considered the cases where $E = -|E|$ in which $|E| < V_0$ was leading to the existence of bound states. Now let us consider $E > 0$. The solution to the Schrödinger equation no longer describes the bound state. In fact, it represents the scattering states.

The radial equation is

$$\frac{d^2 R_I}{dr^2} + \frac{2}{r}\frac{dR_I}{dr} + \left[\frac{2mE}{\hbar^2} + \frac{2m}{\hbar^2}V_0 - \frac{l(l+1)}{r^2}\right] R_I = 0 \quad 0 \le r \le a$$

$$\frac{d^2 R_{II}}{dr^2} + \frac{2}{r}\frac{dR_{II}}{dr} + \left[\frac{2mE}{\hbar^2} - \frac{l(l+1)}{r^2}\right] R_{II} = 0 \quad r \ge a \tag{8.69}$$

Let us define K^2 and k^2 as

$$\frac{2m(E+V_0)}{\hbar^2} = K^2 \text{ and } \frac{2mE}{\hbar^2} = k^2$$

The Schrödinger equation becomes

$$\frac{d^2R_I}{dr^2} + \frac{2}{r}\frac{dR_I}{dr} + \left[K^2 - \frac{l(l+1)}{r^2}\right]R_I = 0 \quad 0 \le r \le a \tag{8.70}$$

$$\frac{d^2R_{II}}{dr^2} + \frac{2}{r}\frac{dR_{II}}{dr} + \left[k^2 - \frac{l(l+1)}{r^2}\right]R_{II} = 0 \quad r \ge a \tag{8.71}$$

The solution to these equations are given by

$$R_I = A_l j_l(Kr) \quad 0 \le r \le a \tag{8.72}$$

$$R_{II} = C_l j_l(kr) + B_l n_l(kr) \quad r \ge a \tag{8.73}$$

In anticipation of phase-shift analysis in scattering theory, we can define $C_l = D_l \cos\delta_l$ and $B_l = D_l \sin\delta_l$.

$$\therefore R_{II} = D_l \cos\delta_l j_l(kr) + D_l \sin\delta_l n_l(kr) \tag{8.74}$$

The importance of the choice of the constants given in the Equation (8.74) can be understood by considering the large r behaviour of $j_l(kr)$ and $n_l(kr)$.

$$j_l(kr) \xrightarrow[r\to\infty]{} \frac{\sin\left(kr - \frac{l\pi}{2}\right)}{kr}$$

$$n_l(kr) \xrightarrow[r\to\infty]{} \frac{\cos\left(kr - \frac{l\pi}{2}\right)}{kr} \tag{8.75}$$

$$\therefore R_{II} \xrightarrow[r\to\infty]{} \frac{1}{kr}\left[D_l \cos\delta_l \sin\left(kr - \frac{l\pi}{2}\right) + D_l \sin\delta_l \cos\left(kr - \frac{l\pi}{2}\right)\right]$$

$$\text{or } R_{II} \xrightarrow[r\to\infty]{} \frac{1}{kr} D_l \sin\left(kr - \frac{l\pi}{2} + \delta_l\right)$$

The significance of these equation is discussed in Chapter 15.

8.14 TWO-PARTICLE SYSTEM IN A CENTRAL POTENTIAL

Let us consider a system of two particles of masses m_1 and m_2 interacting through a central potential. Our immediate interest in considering such a system is hydrogen or hydrogen-like atoms. In hydrogen-like atoms, m_1 and m_2 represent masses of nucleus and electron. Two-body problem under central potential can be reduced to two single body problems.

The Hamiltonian H for two particles in central potential is

$$\begin{aligned}
H &= \frac{p_1^2}{2m_1} + \frac{p_2^2}{2m_2} + V(|\mathbf{r}_1 - \mathbf{r}_2|) \\
&= -\frac{\hbar^2}{2m_1}\nabla_1^2 - \frac{\hbar^2}{2m_2}\nabla_2^2 + V(|\mathbf{r}_1 - \mathbf{r}_2|) \\
&= -\frac{\hbar^2}{2m_1}\left(\frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial y_1^2} + \frac{\partial^2}{\partial z_1^2}\right) - \frac{\hbar^2}{2m_2}\left(\frac{\partial^2}{\partial x_2^2} + \frac{\partial^2}{\partial y_2^2} + \frac{\partial^2}{\partial z_2^2}\right) + V(|\mathbf{r}_1 - \mathbf{r}_2|) \qquad (8.76)
\end{aligned}$$

Let us define $\mathbf{R}$ as the coordinate of the centre of mass. Let $\mathbf{r}$ be the relative coordinate of particle 2 with respect to particle 1.

$$\mathbf{R} = \frac{m_1\mathbf{r}_1 + m_2\mathbf{r}_2}{m_1 + m_2} = \frac{m_1\mathbf{r}_1 + m_2\mathbf{r}_2}{M}$$

$$\mathbf{r} = \mathbf{r}_2 - \mathbf{r}_1$$

Let us write $\mathbf{R}$ and $\mathbf{r}$ as

$$\mathbf{R} = X\,\mathbf{i} + Y\,\mathbf{j} + Z\,\mathbf{k} \text{ and } \mathbf{r} = x\,\mathbf{i} + y\,\mathbf{j} + z\,\mathbf{k}$$

So we have

$$X = \frac{m_1x_1 + m_2x_2}{M}, Y = \frac{m_1y_1 + m_2y_2}{M}, Z = \frac{m_1z_1 + m_2z_2}{M}$$

$$x = x_2 - x_1 \quad y = y_2 - y_1 \quad z = z_2 - z_1$$

$$\frac{\partial}{\partial x_1} = \frac{\partial X}{\partial x_1}\cdot\frac{\partial}{\partial X} + \frac{\partial x}{\partial x_1}\cdot\frac{\partial}{\partial x} = \frac{m_1}{M}\frac{\partial}{\partial X} - \frac{\partial}{\partial x}$$

$$\frac{\partial}{\partial x_2} = \frac{\partial X}{\partial x_2}\cdot\frac{\partial}{\partial X} + \frac{\partial x}{\partial x_2}\cdot\frac{\partial}{\partial x} = \frac{m_2}{M}\frac{\partial}{\partial X} + \frac{\partial}{\partial x}$$

$$\begin{aligned}
\frac{\partial^2}{\partial x_1^2} &= \left(\frac{m_1}{M}\frac{\partial}{\partial X} - \frac{\partial}{\partial x}\right)\left(\frac{m_1}{M}\frac{\partial}{\partial X} - \frac{\partial}{\partial x}\right) \\
&= \frac{m_1^2}{M^2}\frac{\partial^2}{\partial X^2} - \frac{2m_1}{M}\frac{\partial^2}{\partial x\partial X} + \frac{\partial^2}{\partial x^2}
\end{aligned}$$

$$\begin{aligned}
\frac{\partial^2}{\partial x_2^2} &= \left(\frac{m_2}{M}\frac{\partial}{\partial X} + \frac{\partial}{\partial x}\right)\left(\frac{m_2}{M}\frac{\partial}{\partial X} + \frac{\partial}{\partial x}\right) \\
&= \frac{m_2^2}{M^2}\frac{\partial^2}{\partial X^2} + \frac{2m_2}{M}\frac{\partial^2}{\partial x\partial X} + \frac{\partial^2}{dx^2}
\end{aligned}$$

$$-\frac{\hbar^2}{2m_1}\frac{\partial^2}{\partial x_1^2} - \frac{\hbar^2}{2m_2}\frac{\partial^2}{\partial x_2^2} = -\frac{\hbar^2}{2M}\frac{\partial^2}{\partial X^2} - \frac{\hbar^2}{2}\left(\frac{1}{m_1} + \frac{1}{m_2}\right)\frac{\partial^2}{\partial x^2}$$

Working out other terms also in the Equation (8.76), we get

$$H = -\frac{\hbar^2}{2M}\left[\frac{\partial^2}{\partial X^2}+\frac{\partial^2}{\partial Y^2}+\frac{\partial^2}{\partial Z^2}\right]-\frac{\hbar^2}{2}\left(\frac{1}{m_1}+\frac{1}{m_2}\right)\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2}\right)+V(r) \tag{8.77}$$

Let us define the reduced mass μ through the relation

$$\frac{1}{\mu}=\frac{1}{m_1}+\frac{1}{m_2} \tag{8.78}$$

$$H = -\frac{\hbar^2}{2M}\nabla_R^2-\frac{\hbar^2}{2\mu}\nabla^2+V(r) \tag{8.79}$$

Thus, the Hamiltonian splits into two parts: one part describing the motion of the centre of mass and the second part describing the motion of a single particle of reduced mass μ in terms of relative coordinate under central force.

$$Hu_T(\mathbf{r}_1,\mathbf{r}_2)=E_T u_T(\mathbf{r}_1,\mathbf{r}_2)$$

where the subscript T refers to the total system.

$$Hu_T(\mathbf{r}_1,\mathbf{r}_2)=\left[-\frac{\hbar^2}{2M}\nabla_R^2-\frac{\hbar^2}{2\mu}\nabla^2+V(r)\right]u_T(\mathbf{r}_1,\mathbf{r}_2)=E_T u_T(\mathbf{r}_1,\mathbf{r}_2)$$

The structure of the Hamiltonian operator suggests that the wave function $u_T(\mathbf{r}_1,\mathbf{r}_2)$ can be written as $u_T(\mathbf{r}_1,\mathbf{r}_2)=u_c(\mathbf{R})u_E(\mathbf{r})$ with $E_T=E_C+E$.

$$-\frac{\hbar^2}{2M}\nabla_R^2 u_C(\mathbf{R})\cdot u_E(\mathbf{r})+\left(-\frac{\hbar^2}{2\mu}\nabla^2 u_E(\mathbf{r})+V(r)u_E(\mathbf{r})\right)u_C(\mathbf{R})=E_T u_C(\mathbf{R})u_E(\mathbf{r})$$

Dividing throughout by $u_C(\mathbf{R})\ u_E(\mathbf{r})$, we get,

$$\left[-\frac{1}{u_C(\mathbf{R})}\frac{\hbar^2}{2M}\nabla_R^2 u_C(\mathbf{R})-E_C\right]+\left[\frac{1}{u_E(\mathbf{r})}\cdot\left(-\frac{\hbar^2}{2\mu}\nabla^2+V(r)\right)u_E(\mathbf{r})-E\right]=0$$

Setting each term to zero separately we get,

$$-\frac{\hbar^2}{2M}\nabla_R^2 u_E(\mathbf{R})=E_C u_E(\mathbf{R}) \tag{8.80}$$

$$\left[-\frac{\hbar^2}{2\mu}\nabla^2+V(r)\right]u_E(\mathbf{r})=Eu_E(\mathbf{r}) \tag{8.81}$$

The Equation (8.80) describes the motion of the centre of mass. This is a free particle equation whose solution is

$$u_C(\mathbf{R})=\frac{1}{L^{\frac{3}{2}}}e^{\frac{i\mathbf{P}\cdot\mathbf{R}}{\hbar}}$$

where **P** is the momentum of the centre of mass, and the energy E_C is given by $E_C = P^2/2M$.

The total wave function $u_T(\mathbf{r}_1, \mathbf{r}_2)$ is given by

$$u_T(\mathbf{r}_1, \mathbf{r}_2) = \frac{1}{L^{3/2}} e^{\frac{i\mathbf{P}\cdot\mathbf{R}}{\hbar}} u_E(\mathbf{r})$$

So solving time-independent Schrödinger equation for a two-particle system under central potential reduces to solving a single-particle equation for a particle of reduced mass μ.

i.e., we have to solve the equation

$$\left[-\frac{\hbar^2}{2\mu}\nabla^2 + V(r)\right] u_E(r,\theta,\phi) = E u_E(r,\theta,\phi)$$

8.15 RADIAL EQUATION FOR COULOMB POTENTIAL

The Coulomb potential between the changes *Ze* and –*e* is

$$V(r) = -\frac{Ze^2}{r}$$

The single-particle equation now becomes

$$\left[-\frac{\hbar^2}{2\mu}\nabla^2 - \frac{Ze^2}{r}\right] u_E(r,\theta,\phi) = E u_E(r,\theta,\phi)$$

where $u_E(r,\theta,\phi)$ is given by $u_E(r,\theta,\phi) = R(r)Y_{lm}(\theta,\phi)$.

The radial equation for Coulomb potential is

$$\frac{d^2R}{dr^2} + \frac{2}{r}\frac{dR}{dr} + \left[\frac{2\mu}{\hbar^2}\left(E + \frac{Ze^2}{r}\right) - \frac{l(l+1)}{r^2}\right] R = 0 \tag{8.82}$$

The energy eigenvalue E may be greater than or less than zero. The case $E > 0$ corresponds to the scattering of electron by the nucleus due to Coulomb interaction. The case $E < 0$ corresponds to the bound state.

8.16 HYDROGEN ATOM

Here we concentrate only on the bound-state solution of hydrogen or hydrogen-like atom, for which energy eigenvalue should be less than zero.

i.e., $$E = -|E|$$

Therefore, the radial equation becomes

$$\frac{d^2R}{dr^2} + \frac{2}{r}\frac{dR}{dr} + \left[\frac{2\mu}{\hbar^2}\left(\frac{Ze^2}{r} - |E|\right) - \frac{l(l+1)}{r^2}\right] R(r) = 0$$

Let us define $\rho = \alpha r$.

$$\alpha^2 \frac{d^2R}{d\rho^2} + \frac{2\alpha^2}{\rho}\frac{dR}{d\rho} + \left[\frac{2\mu}{\hbar^2}\cdot\frac{Ze^2\alpha}{\rho} - \frac{2m}{\hbar^2}|E| - \frac{\alpha^2(l+1)}{\rho^2}\right]R = 0$$

$$\frac{d^2R}{d\rho^2} + \frac{2}{\rho}\frac{dR}{d\rho} + \left[\frac{2\mu}{\hbar^2}\frac{Ze^2}{\alpha\rho} - \frac{2m|E|}{\hbar^2\alpha^2} - \frac{l(l+1)}{\rho^2}\right]R = 0$$

We can choose α in a convenient way. The standard choice of α^2 is such that

$$\frac{2\mu|E|}{\hbar^2\alpha^2} = \frac{1}{4} \quad \text{or} \quad \alpha^2 = \frac{8\mu|E|}{\hbar^2} \tag{8.83}$$

The reason for the factor 1/4 is that it will make the asymptotic behaviour of $R(\rho)$ as $e^{-\rho/2}$ for large ρ.

Define λ as

$$\lambda = \frac{2\mu Ze^2}{\hbar^2\alpha} \tag{8.84}$$

Therefore, the radial equation becomes

$$\frac{d^2R}{d\rho^2} + \frac{2}{\rho}\frac{dR}{d\rho} + \left[\frac{\lambda}{\rho} - \frac{1}{4} + \frac{l(l+1)}{\rho^2}\right]R = 0 \tag{8.85}$$

Let us study the asymptotic behaviour of $R(\rho)$. As $\rho \to \infty$, the Equation (8.85) becomes

$$\frac{d^2R}{d\rho^2} - \frac{1}{4}R \simeq 0$$

$$\therefore R(\rho) \xrightarrow[\rho\to\infty]{} Ae^{-\frac{1}{2}\rho} + Be^{\frac{1}{2}\rho}$$

The boundary condition for the radial wave function to be acceptable for a bound state is $R(\rho) \to 0$ as $\rho \to \infty$. Therefore, we can choose $B = 0$.

This suggests $R(\rho)$ as

$$R(\rho) = e^{-\frac{\rho}{2}}G(\rho)$$

Another boundary condition for the radial wave function is $R(\rho) \to \rho^l$ as $\rho \to 0$. Therefore, we can write

$$R(\rho) = e^{-\frac{\rho}{2}}\rho^l F(\rho)$$

$$\frac{dR(\rho)}{dP} = e^{-\frac{\rho}{2}}\left[-\frac{1}{2}\rho^l F + l\rho^{l-1}F + \rho^l\frac{dF}{d\rho}\right] \tag{8.86}$$

$$\frac{d^2R}{d\rho^2} = e^{-\frac{\rho}{2}}\left[\rho^l\frac{d^2F}{d\rho^2} + \frac{dF}{d\rho}(-\rho^l + 2l\rho^{l-1}) + F\left(\frac{1}{4}\rho^l - l\rho^{l-1} + l(l-1)\rho^{l-2}\right)\right]$$

Substituting these two equations in the Equation (8.85), we get

$$\rho^2 \frac{d^2F}{d\rho^2} + \rho[2(l+1)-\rho]\frac{dF}{d\rho} + [\lambda-(l+1)]\rho F = 0$$

or
$$\rho \frac{d^2F}{d\rho^2} + [2(l+1)-\rho]\frac{dF}{d\rho} + [\lambda-(l+1)]F = 0 \tag{8.87}$$

Let us try the solution $F(\rho) = \sum_{r=0}^{\infty} C_r \rho^r$

Substituting this solution in the Equation (8.87), we get

$$\sum_{r=0}^{\infty} C_r[r(r-1)+2(l+1)r]\rho^{r-1} - \sum_{r=0}^{\infty}(r-\lambda+l+1)C_r\rho^r = 0$$

Let us put $r - 1 = s$ in the first term. Then we get

$$\sum_{r=0}^{\infty} C_r[r(r-1)+2(l+1)r]\rho^{r-1} = \sum_{r=1}^{\infty} C_r[r(r-1)+2(l+1)r]\rho^{r-1}$$

$$= \sum_{s=0}^{\infty} C_{s+1}[(s+1)s+2(l+1)(s+1)]\rho^s = \sum_{r=0}^{\infty} C_{r+1}[(r+1)r+2(l+1)(r+1)]\rho^r$$

(We have made use of the fact that the summation index is a dummy variable.)

The recursion relation becomes

$$\sum_{r=0}^{\infty}\left\{C_{r+1}[(r+1)r+2(l+1)(r+1)] - C_r(r-\lambda+l+1)\right\}\rho^r = 0$$

Equating the coefficient of ρ^r to zero, we get

$$C_{r+1} = \frac{r-\lambda+l+1}{(r+1)(2l+2+r)}C_r \tag{8.88}$$

To check the convergence of the series, let us evaluate $\operatorname*{Lt}_{r\to\infty} C_{r+1}/C_r$ and compare it with such limit of a well-known series.

$$\operatorname*{Lt}_{r\to\infty}\frac{C_{r+1}}{C_r} = \frac{r}{r^2} = \frac{1}{r}$$

It is easy to check e^ρ has the same behaviour for large r.

$$e^\rho = 1 + \rho + \frac{\rho^2}{2!} + \frac{\rho^3}{3!} + \ldots = \sum_r C_r\rho^r \text{ with } C_r = \frac{1}{r!}$$

$$\therefore \frac{C_{r+1}}{C_r} = \frac{r!}{(r+1)!} = \frac{1}{r+1}$$

$$\operatorname*{Lt}_{r\to\infty}\frac{C_{r+1}}{C_r} = \frac{1}{r}$$

Therefore, the series $F(\rho)=\sum_{r=0}^{\infty}C_r\rho^r$ behaves like e^{ρ} for large ρ.

Let us recall $R_l(r)$ is given by $R_l(\rho)=\rho^l e^{-\frac{\rho}{2}}F$.

$$\therefore R_l(\rho)\xrightarrow[\rho\to\infty]{}\rho^l e^{-\frac{\rho}{2}}e^{\rho}\to\infty$$

i.e., $R_l(\rho)$ diverges in spite of the presence of the factor $e^{-\frac{\rho}{2}}$. This divergence can be avoided if we have a finite series (i.e. a polynomial) instead of an infinite series for $F(\rho)$. So we have to terminate the infinite series for $F(\rho)$ at some term so that it becomes a polynomial. This can be implemented by demanding

$$\lambda = n, \text{ with } n \text{ as an integer}$$

Then
$$C_{r+1}=\frac{r+l+1-n}{(r+1)(2l+2+r)}C_r$$

When r becomes n–l– 1, C_{r+1} and all the other subsequent coefficients become zero.

So the condition for terminating the infinite series is $r = n - l - 1$.

Since r has to be a positive integer, for a given n, l can take only some values; i.e., the only possible values consistent with n are 0, 1, 2, …, $(n - 1)$.

So for a given n, we have different radial functions: R_{n0}, R_{n1}, R_{n2}, …, $R_{n,n-1}$,

The energy eigenvalue E is is determined from the Equations (8.83) and (8.84).

$$\lambda = n=\frac{2\mu Ze^2}{\hbar^2\alpha}=\frac{2\mu Ze^2}{\hbar^2(-8\mu E)^{\frac{1}{2}}} \tag{8.89}$$

$$E=-\frac{\mu Z^2 e^4}{2n^2\hbar^2} \tag{8.90}$$

or
$$E_n=-\frac{\mu Z^2e^4}{2n^2\hbar^2}\qquad n=1,2,3\ldots \tag{8.91}$$

This is the same expression for energy values obtained from Bohr model. It is an unusual coincidence.

It is to be noted that quantum mechanical understanding of hydrogen atom is very much different from Bohr model. In Bohr model, the electron moves in a circular orbit of definite radius; i.e., the electron has a definite trajectory. We have already seen in Chapter 5 that quantum mechanics does not allow the concept of definite trajectory for a particle. So, in a sense, Bohr model is inconsistent with quantum mechanics.

8.16.1 ρ in terms Bohr radius a_0.

$$\rho=\alpha r=\left(-\frac{8\mu E}{\hbar^2}\right)^{1/2}r=\left[-8\mu\left(-\frac{\mu Z^2e^4}{2n^2\hbar^2}\right)\right]^{1/2}r$$

$$\therefore \rho = \frac{2\mu Z e^2}{\hbar n} r \tag{8.92}$$

The Bohr radius a_0 is defined as

$$a_0 = \frac{\hbar}{\mu e^2} \tag{8.93}$$

In terms of Bohr radius,

$$\rho = \frac{2Z}{na_0} r; \quad \alpha_n = \frac{2Z}{na_0} \tag{8.94}$$

8.16.2 Radial Wave Function

When $\lambda = n$, the radial equation becomes

$$\rho \frac{d^2 F}{d\rho^2} + [2(l+1) - \rho] \frac{dF}{d\rho} + [n - (l+1)]F = 0 \tag{8.95}$$

The structure of this differential equation is the same as the differential equation for the associated Laguerre polynomial L_q^k, which is given by

$$x \frac{d}{dx^2} L_q^k + (k+1-x) \frac{d}{dx} L_q^k - (k-q) L_q^k = 0 \tag{8.96}$$

The similarity between the two equations suggests that

$$k + 1 = 2(l + 1) \quad \text{and} \quad q - k = n - (l + 1)$$

Or

$$k = 2l + 1 \quad \text{and} \quad q = n + l \tag{8.97}$$

$$\therefore F(\rho) = L_{n+l}^{2l+1}(\rho) \tag{8.98}$$

The radial wave function $R_{rl}(r)$ is given by

$$R_{nl}(r) = N_{nl} e^{-\frac{Zr}{na_0}} \left(\frac{2Zr}{na_0} \right)^l L_{n+l}^{2l+1}(\rho) \tag{8.99}$$

The normalization constant N_{nl} can be determined by

$$N_{nl} = -\left(\frac{2Z}{na_0} \right)^{3/2} \left[\frac{(n-l-1)!}{2n[(n+l)!]^3} \right]^{1/2} \tag{8.100}$$

8.16.3 Degeneracy

The wave function for the hydrogen atom is given by

$$u_{ulm}(r,\theta,\phi) = R_{nl}(r)Y_{lm}(\theta,\phi)$$

$$= N_{nl}e^{-\frac{Zr}{na_o}} \cdot \left(\frac{2Zr}{na_0}\right)^l L_{n+l}^{2l+1}(\rho)Y_{lm}(\theta,\phi) \tag{8.101}$$

The energy eigenvalues are given by

$$E_n = -\frac{\mu Z^2 e^4}{2n^2\hbar^2} \tag{8.102}$$

Note that the energy eigenvalues depend only on n. But different sets of quantum numbers n and l represent different states. For a given n, we have a number of states having the same energy. These states are degenerate states corresponding to a given energy level E_n. We can easily calculate the degeneracy for a given energy level.

For a given n, the possible values of l are 0, 1, 2, …, $(n - 1)$. For a given l, there are $(2l + 1)$ values of m. Therefore, the degeneracy for a given n is

$$D = \sum_{l=0}^{n-1}(2l+1) = \sum_{l=0}^{n-1}2l + \sum_{l=0}^{n-1}1 = \frac{2(n-1)n}{2} + n = n^2$$

8.16.4 Radial Wave Function from Associated Laguerre Polynomial

It is a simple exercise to construct the radial wave function for hydrogen-like atoms, which is given by

$$R_{nl}(r) = N_{nl}\rho^l e^{-\frac{\rho}{2}} L_{n+l}^{2l+1}(\rho), \quad \text{with} \quad \rho = \frac{2Zr}{na_0} \tag{8.103}$$

So our task reduces to determining the associated Laguerre polynomial $L_{n+l}^{2l+1}(\rho)$.

The generating function for $L_q^k(\rho)$ is given by

$$(-1)^k s^k (1-s)^{-k-1} e^{-\frac{\rho s}{(1-s)}} = \sum \frac{s^q}{q!} L_q^k(\rho) \tag{8.104}$$

$L_q^k(\rho)$ can be obtained from the above equation. Alternatively, $L_q^k(\rho)$ can be obtained from

$$L_q^k(\rho) = \frac{q!}{(q-k)!} e^{\rho} \frac{d^q}{d\rho^q}[e^{-\rho}\rho^{q-k}] \tag{8.105}$$

Some of the radial wave functions are given in Table 8.3. The radial wave for atomic state is generally characterized by the quantum numbers n and l. In spectroscopy, states with different values of l are denoted by different letters like *s, p, d, f* … The choice of these letters are only by convention. The standard convention is as follows:

$l = 0 \quad s$

$l = 1 \quad p$

$$l = 2 \quad d$$
$$l = 3 \quad f$$
$$l = 4 \quad g$$

So the radial wave functions are labelled by (n, l). For instance, $3p$ means $n = 3$, $l = 1$.

Table 8.3 Some radial wave functions

n	l	Spectroscopic state	R_{nl}	ρ
1	0	1s	$R_{10} = 2\left(\frac{Z}{a_0}\right)^{3/2} e^{-\frac{\rho}{2}}$	$\rho = \frac{2Zr}{a_0}$
2	0	2s	$R_{20} = \frac{1}{2\sqrt{2}}\left(\frac{Z}{a_0}\right)^{3/2} (2-\rho)e^{-\frac{\rho}{2}}$	$\rho = \frac{Zr}{a_0}$
	1	2p	$R_{21} = \frac{1}{2\sqrt{6}}\left(\frac{Z}{a_0}\right)^{3/2} \rho e^{-\frac{\rho}{2}}$	
3	0	3s	$R_{30} = \frac{1}{9\sqrt{3}}\left(\frac{Z}{a_0}\right)^{3/2} (6-6\rho+\rho^2)e^{-\frac{\rho}{2}}$	$\rho = \frac{2Z}{3a_0}$
	1	3p	$R_{31} = \frac{1}{9\sqrt{6}}\left(\frac{Z}{a_0}\right)^{3/2} \rho(4-\rho)e^{-\frac{\rho}{2}}$	
	2	3d	$R_{32} = \frac{1}{9\sqrt{30}}\left(\frac{Z}{a_0}\right)^{3/2} \rho^2 e^{-\frac{\rho}{2}}$	

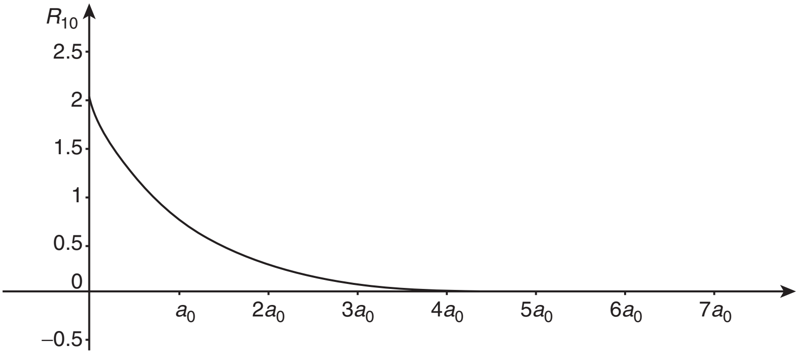


Fig. 8.6(i) Radial wave function R_{10} $(a_0 = 1)$

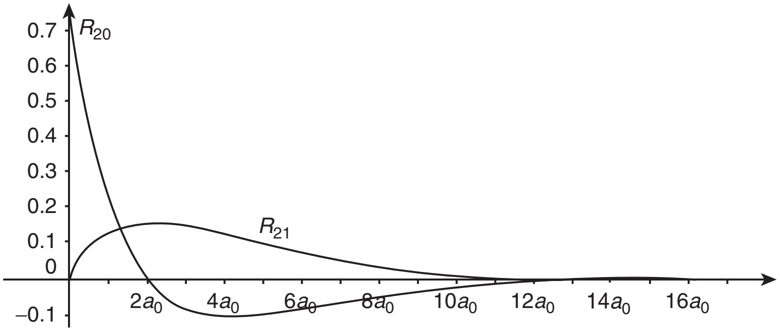


Fig. 8.6(ii) Radial wave functions R_{20} and R_{21} ($a_0 = 1$)

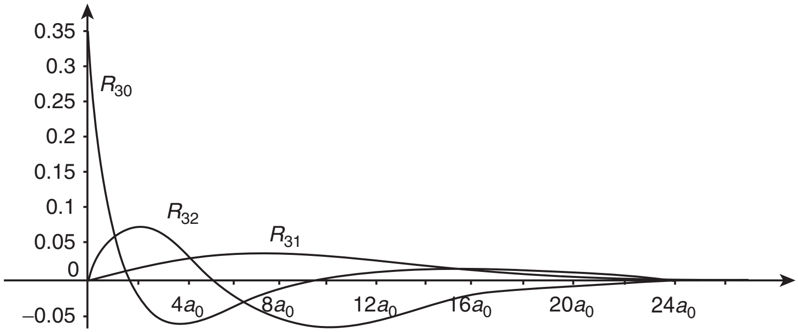


Fig. 8.6(iii) Radial wave functions R_{30}, R_{31} and R_{32} ($a_0 = 1$)

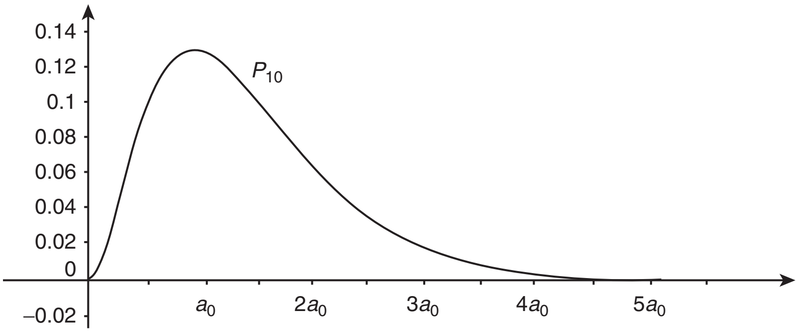


Fig. 8.6(iv) Radial probability density P_{10} ($a_0 = 1$)

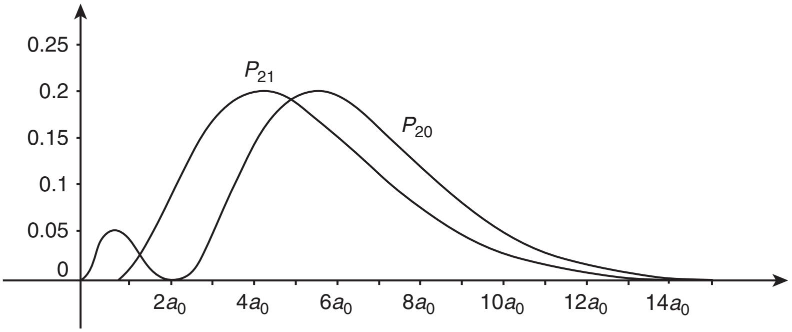


Fig. 8.6(v) Radial probability density P_{21} and P_{20} ($a_0 = 1$)

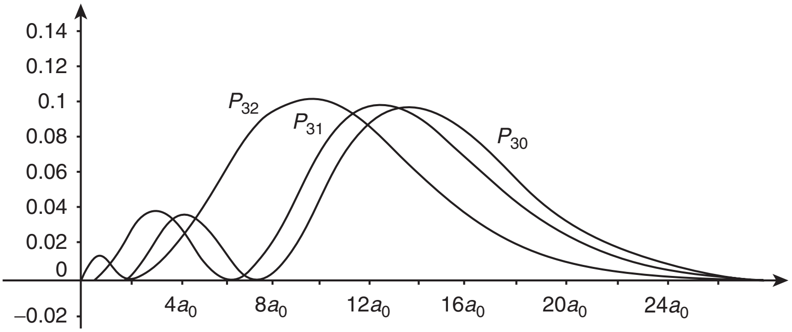


Fig. 8.6(vi) Radial probability density P_{32}, P_{31} and P_{30} ($a_0 = 1$)

These wave functions are sketched in Fig. 8.6. It can be generally seen that the wave function falls to zero after a few number of Bohr radii. The Bohr radius is $a_0 = \hbar^2/\mu e^2 = 0.53\text{Å}$. For instance, the wave function R_{10} becomes very small at $r = 3\text{Å}$ ($\sim 6a_0$). In region beyond this radius, it is almost impossible to find an electron. It can be seen that for larger n values, the wave function is significant for larger values of r.

8.16.5 Radial Probability Density

The radial probability density is given by

$$P_{nl}(r) = r^2 \left|R_{nl}(r)\right|^2 \tag{8.106}$$

It can be conveniently written as

$$P_{nl}(r) = \left(\frac{n^2 a_0^{\ 2}}{4Z^2}\right)\rho^2 \left|R_{nl}(\rho)\right|^2 \tag{8.107}$$

The radial probability density for some of the states are listed in Table 8.4:

Table 8.4 Probability density for some states of hydrogen

n	l	Spectroscopic state	P_{nl}	ρ
1	0	1*s*	$P_{10}(\rho) = \left(\dfrac{Z}{a_0}\right)\rho^2 e^{-\rho}$	$\rho = \dfrac{2Zr}{a_0}$
2	0	2*s*	$P_{20}(\rho) = \dfrac{1}{8}\left(\dfrac{\rho}{a_0}\right)\rho^2(2-\rho)^2 e^{-\rho}$	$\rho = \dfrac{Zr}{a_{10}}$
	1	2*p*	$P_{21}(\rho) = \dfrac{1}{24}\left(\dfrac{Z}{a_0}\right).\rho^4 e^{-\rho}$	
3	0	3*s*	$P_{30}(\rho) = \dfrac{1}{108}\left(\dfrac{Z}{a_0}\right)\rho^2(6-6\rho+\rho^2)^2 e^{-\rho}$	$\rho = \dfrac{2Zr}{3a_0}$
	1	3*p*	$P_{31}(\rho) = \dfrac{1}{216}\left(\dfrac{Z}{a_0}\right)\rho^2(4-\rho)^2 e^{-\rho}$	
	2	3*d*	$P_{32}(\rho) = \dfrac{1}{1080}\left(\dfrac{Z}{a_0}\right)\rho^6 e^{-\rho}$	

The radial probability density function $P_{nl}(r)$ for these functions is sketched in Fig. 8.6.

Example 8.7 Determine the maxima and minima of the radial probability distribution function $P_{30}(r)$.

Solution:

$$P_{30}(r) = \frac{1}{108}\left(\frac{Z}{a_0}\right)\rho^2(6-6\rho+\rho^2)e^{-\rho} \text{ with } \rho = \frac{2Zr}{3a_0}$$

The extreme values of $P_{30}(r)$ are determined by

$$\frac{dP_{30}}{dr} = 0 \quad \text{or} \quad \frac{dP_{30}}{d\rho} = 0$$

$$\therefore \frac{d}{d\rho}[\rho^2(6-6\rho+\rho^2)^2 e^{-\rho}] = 0$$

$$[2\rho(6-6\rho+\rho^2)^2 + \rho^2(6-6\rho+\rho^2)2(-6+2\rho) - \rho^2(6-6\rho+\rho^2)^2]e^{-\rho} = 0$$

$$\rho(6-6\rho+\rho^2)(12-30\rho+12\rho^2-\rho^3) = 0$$

The solution to quadratic equation are $\rho = 1.27$ and $\rho = 4.73$.
The solutions to cubic equation are $\rho = 0.49$, $\rho = 2.79$ and $\rho = 8.72$.
The corresponding values of r are determined from the relation $r = 3/2a_0 r$.

Substituting the values of these roots in P_{30}, we get the following values:

ρ	0	0.49	1.27	2.79	4.73	8.72
r	0	$0.74a_0$	$1.91a_0$	$4.19a_0$	$7.1a_0$	$13.08a_0$
P_{30}	0	$0.01/a_0$	0	$0.04/a_0$	0	$0.10/a_0$

Note the dimension of probability density. It has a dimension of inverse of length. $P_{nl}(r)dr$ is the probability which is a dimensionless number.

Example 8.8 Calculate the probability of a $3d$ electron for hydrogen atom in the ranges
1. $0 < r < a_0$ 2. $a_0 < r < 3a_0$ 3. $3a_0 < r < 9a_0$ 4. $9a_0 < r < 12a_0$

Solution: The probability of finding the particle in a specific range $a < r < b$ is given by

$$P_{nl}(a < r < b) = \int_a^b P_{nl}(r)dr$$

A useful result in such calculations is

$$\int_a^b \rho^n e^{-\rho} d\rho = \left[e^{-\rho}[\rho^n + n\rho^{n-1} + n(n-1)\rho^{n-2} + ...n!]\right]_b^a$$

1. $0 < r < a_0$

The corresponding range in ρ is $0 < \rho < \frac{2}{3}$.

$$\therefore P_{3d}(0 < r < a_0) = \int_0^{a_0} P_{nl}(r)dr = \int_0^{\frac{2}{3}} P_{nl}\frac{dr}{d\rho}.d\rho$$

$$= \frac{3a_0}{2}\cdot\frac{1}{1080a_0}\int_0^{\frac{2}{3}} \rho^6 e^{-\rho} d\rho$$

$$= \frac{1}{720}\{(\rho^6 + 6\rho^5 + 30\rho^4 + 120\rho^3 + 360\rho^2 + 720\rho + 720)e^{-\rho}\}_{\frac{2}{3}}^{0}$$

$$= 1.7\times 10^{-3}.$$

Other results can be worked out in a similar way

2. $P_{3d}(a_0 < r < 3a_0) = \frac{1}{720}\int_{\frac{2}{3}}^{2} \rho^6 e^{-\rho} d\rho = 3.39\times 10^{-3}$
3. $P_{3d}(3a_0 < r < 9a_0) = \frac{1}{720}\int_2^6 \rho^6 e^{-\rho} d\rho = 0.39$
4. $P_{3d}(9a_0 < r < 12a_0) = \frac{1}{720}\int_6^8 \rho^6 e^{-\rho} d\rho = 0.29$

8.16.6 $\langle r^s \rangle_{nl}$

Since we know the radial probability density function, we can determine $\langle r^s \rangle_{nl}$ given

$$\langle r^s \rangle_{nl} = \int_0^\infty r^s P_{nl}(r)dr = \int_0^\infty r^s |R_{nl}(r)|^2 r^2 dr \tag{8.108}$$

To calculate $\langle r^s \rangle_{nl}$, the following results are useful:

1. Let us define an integral $I^k_{qr,p}$ given by

$$I^k_{qr,p} = \int_0^\infty e^{-\rho}\rho^{k-1}L^k_q(\rho)L^k_r(\rho)\rho^p d\rho \tag{8.109}$$

   This integral has been evaluated in Appendix I of this chapter.
   A special case is $I^k_{qq,p}$, which is given by

$$I^k_{qq,p} = (q!)^2\begin{Bmatrix} {}^{p-1}C_0\,{}^{p-1}C_0\dfrac{(q+p-1)!}{(q-k)!} + {}^{p-1}C_1\,{}^{p-1}C_1\dfrac{(q+p-2)!}{(q-k-1)!} \\ +{}^{p-1}C_2\,{}^{p-1}C_2\dfrac{(q+p-3)!}{(q-k-2)!} + \ldots{}^{p-1}C_{p-1}\,{}^{p-1}C_{p-1}\dfrac{q!}{(q-k+p+1)!} \end{Bmatrix} \tag{8.110}$$

2. A second useful relation is Kramer's recurrence relation (see the appendix for proof).

$$\frac{s+1}{n^2}\langle r^s \rangle_{nl} - \frac{a_0}{Z}(2s+1)\langle r^{s-1} \rangle_{nl} + \frac{s}{4}[(2l+1)^2 - s^2]\langle r^{s-2} \rangle_{nl}\frac{a_0{}^2}{Z^2} = 0 \tag{8.111}$$

Example 8.9 Evaluate the normalization constant N_{lm} for the radial wave function for hydrogen atom given by

$$R_{nl}(r) = N_{nl}\rho^l e^{-\frac{\rho}{2}}L^{2l+1}_{n+l}(\rho)$$

Solution: The normalization condition for R_{nl} is given by $\int_0^\infty |R_{nl}(r)|^2 r^2 dr = 1$. In terms of the variable ρ this condition becomes

$$\int_0^\infty N^2_{nl}e^{-\rho}\rho^{2l}.\left[L^{2l+1}_{n+1}(\rho)\right]^2\frac{\rho^2 d\rho}{\alpha_n{}^3} = 1$$

This integral can be evaluated using the integral $I^k_{qr,p}$ given in (8.110).

Let us define q and k as

$$q = n + 1 \text{ and } k = 2l + 1$$

The normalization condition becomes

$$\int_0^\infty \frac{N^2_{nl}}{\alpha^3_n}e^{-\rho}\rho^{k-1}[(L^k_q(\rho)]^2\rho^2 d\rho = 1$$

or

$$\frac{N^2_{nl}}{\alpha^3_n}I^k_{qq,2} = 1$$

Using the Equation (8.110), we get

$$
\begin{aligned}
\therefore \frac{N_{nl}^2}{\alpha_n^3} I_{qq,2}^k &= \frac{N_{nl}^2}{\alpha_n^3}\left[\frac{(q+1)!}{(q-k)!}+\frac{q!}{(q-k-1)!}\right](q!)^2 \\
&= \frac{N_{nl}^2}{\alpha_n^3}.(q!)^2.\frac{q!}{(q-k)!}[q+1+q-k] \\
&= \frac{N_{nl}^2}{\alpha_n^3}.(q!)^3.\frac{1}{(q-k)!}[2q+1-k] \\
&= \frac{N_{nl}^2}{\alpha_n^3}[(n+l)!]^3\frac{1}{(n-l-1)!}2n = 1 \\
\therefore N_{nl}^2 &= \left(\frac{2Z}{na_0}\right)^3.\frac{(n-l-1)!}{2n.[(n+l)!]^3}
\end{aligned}
$$

$$
N_{nl} = -\left(\frac{2Z}{na_0}\right)^{3/2}\left[\frac{(n-l-1)!}{2n[(n+l)!]^3}\right]^{1/2}
$$

Example 8.10 Determine $\left\langle\frac{1}{r}\right\rangle_{nl}$ for hydrogen atom.

Solution:

$$
\begin{aligned}
\left\langle\frac{1}{r}\right\rangle_{nl} &= \int_0^\infty \frac{1}{r}P_{nl}(r)dr = \int_0^\infty \frac{1}{r}[R_{nl}(r)]^2 r^2 dr = \frac{1}{\alpha_n^2}\int_0^\infty [R_{nl}(\rho)]^2\rho d\rho \\
&= \frac{N_{nl}^2}{\alpha_n^2}\int_0^\infty e^{-\rho}\rho^{2l}[L_{n+1}^{2l+1}(\rho)]\rho d\rho = \frac{N_{nl}^2}{\alpha_n^2}\cdot I_{qq,1}^k
\end{aligned}
$$

Here $q = n+l$ and $k = 2l+1$. Using the Equation (8.110), we get

$$
\begin{aligned}
\langle r\rangle_{nl} &= \frac{N_{nl}^2}{\alpha_n^2}\cdot(q!)^2\left\{{}^0C_0\ {}^0C_0\frac{q!}{(q-k)!}\right\} \\
&= \frac{\alpha_n^3}{\alpha_n^2}\frac{(n-l-1)!}{2n\cdot[(n+l)!]^3}\ [(n+l)!]^2\ \frac{(n+l)!}{(n-l-1)!} = \frac{\alpha_n}{2n} = \frac{Z}{n^2a_0}
\end{aligned}
$$

Example 8.11 Determine $<r>_{nl}$ for hydrogen atom.

Solution:

$$
\begin{aligned}
\langle r_{nl}\rangle &= \int_0^\infty r\,P_{nl}(r)dr = \int_0^\infty r[R_{nl}(r)]^2r^2dr = \frac{1}{\alpha_n^4}\int_0^\infty [R_{nl}(\rho)]^2\rho^3 d\rho \\
&= \frac{N_{nl}^2}{\alpha_n^4}\int_0^\infty e^{-\rho}\rho^{2l}[L_{n+1}^{2l+1}(\rho)]\rho^3 d\rho = \frac{N_{nl}^2}{\alpha_n^4}\cdot I_{qq,3}^k
\end{aligned}
$$

Here $q = n+l$ and $k = 2l+1$. Using the Equation (8.110), we get

$$\begin{aligned}
\langle r\rangle_{nl} &= \frac{N_{nl}^2}{\alpha_n^4}\cdot(q!)^2\left\{{}^2C_0{}^2C_0\frac{(q+2)!}{(q-k)!}+{}^2C_1{}^2C_1\frac{(q+1)!}{(q-k)!}+{}^2C_2{}^2C_2\frac{q!}{(q-k-2)!}\right\}\\
&= \frac{N_{nl}^2}{\alpha_n^4}(q!)^2\left\{\frac{(q+2)!}{(q-k)!}+\frac{4(q+1)!}{(q-k-1)!}+\frac{q!}{(q-k-2)!}\right\}\\
&= \frac{N_{nl}^2}{\alpha_n^4}(q!)^2\cdot\frac{q!}{(q-k)!}\{(q+1)(q+2)+4(q+1)(q-k)+(q-k-1)(q-k)\}\\
&= \frac{\alpha_n^3}{\alpha_n^4(q!)^3}\frac{(q-k)!}{(2q+1-k)}\frac{(q!)^3}{(q-k)!}\{(n+l+1)(n+l+2)+4(n+l+1)(n-l-1)+(n-l-1)(n-l-2)\}\\
&= \frac{a_0}{2}[3n^2-l(l+1)]
\end{aligned}$$

Example 8.12 Evaluate $<r^2>_{nl}$ for hydrogen atom using Kramer's recursion relation.

$$\frac{s+1}{n^2}\langle r^s\rangle_{nl}-a_0(2s+1)\langle r^{s-1}\rangle_{nl}+\frac{s}{4}[(2l+1)^2-s^2]\langle r^{s-2}\rangle_{nl}\,a_0^{\;2}=0.\text{ Assume }\langle r\rangle=\frac{a_0}{2}[3n^2-l(l+1)].$$

Solution: Let us take $s = 2$.

Then we have

$$\frac{3}{n^2}\langle r^2\rangle_{nl}-a_0 5\langle r\rangle_{nl}+\frac{2}{4}a_0^{\;2}[(2l+1)^2-4]\langle r^0\rangle=0$$

$$\langle r^0\rangle=\int_0^\infty\left|R_{nl}(r)\right|^2 r^2dr=1$$

$$\langle r\rangle=\int_0^\infty\left|R_{nl}(r)\right|^2 r.r^2dr=\frac{a_0}{2}[3n^2-l(l+1)]$$

$$\therefore\frac{3}{n^2}\langle r^2\rangle_{nl}=\frac{3a_0^{\;2}}{2}[5n^2+1-3l(l+1)]$$

$$\therefore\langle r^2\rangle_{nl}=\frac{a_0^{\;2}n^2}{2}[5n^2+1-3l(l+1)].$$

Example 8.13 Evaluate $\left\langle\frac{1}{r}\right\rangle_{nl}$ for hydrogen atom using Kramer's recursion relation.

Solution: Choose $s = 0$. Then, Kramer's relation becomes

$$\frac{1}{n^2}\left\langle r^0\right\rangle_{nl}-a_0\left\langle\frac{1}{r}\right\rangle_{nl}=0.$$

Since $\left\langle r^0\right\rangle_{nl}=1,$ we get $\left\langle\frac{1}{r}\right\rangle_{nl}=\frac{1}{n^2a_0}$

Example 8.14 Evaluate $\left\langle\frac{1}{r^2}\right\rangle_{nl}$ for hydrogen like atom.

Solution: $\left\langle\frac{1}{r^2}\right\rangle_{nl}$ is evaluated using Hellmann–Feynman theorem. A special technique is used to obtain the expression for $\left\langle\frac{1}{r^2}\right\rangle_{nl}$.

The energy eigenvalue equation is

$$Hu_{nlm} = HR_{nl}Y_{lm} = E_n R_{nl} Y_{lm} = E_n u_{nlm}$$

$$\begin{aligned} Hu_{nlm} &= \left\{ -\frac{\hbar^2}{2\mu}\left[\frac{d^2}{dr^2} + \frac{2}{r}\frac{d}{dr}\right] + \frac{L^2}{2\mu r^2} - \frac{Ze^2}{r} \right\} R_{nl} Y_{lm} \\ &= \left\{ -\frac{\hbar^2}{2\mu}\left[\frac{d^2}{dr^2} + \frac{2}{r}\frac{d}{dr}\right] + \frac{l(l+1)\hbar^2}{2\mu r^2} - \frac{Ze^2}{r} \right\} R_{nl} Y_{lm} = H_{eff} u_{nlm} = E_n u_{nlm} \end{aligned}$$

where H_{eff} is

$$H_{eff} = -\frac{\hbar^2}{2\mu}\left[\frac{d^2}{dr^2} + \frac{2}{r}\frac{d}{dr}\right] + \frac{l(l+1)\hbar^2}{2\mu r^2} - \frac{Ze^2}{r}$$

The energy eigenvalue E_n is given by

$$E_n = \int u^*_{nlm} H u_{nlm} d^3\mathbf{r} = \int u^*_{nlm} H_{eff} u_{nlm} d^3\mathbf{r}$$

The trick is to now use l as a parameter and use the expression $r = n - l - 1$ of Sec. 8.16. Here, r is not the radial coordinate but it is an integer.

$$n = r + l + 1$$

$$\frac{\partial E_n}{\partial l} = \frac{\partial}{\partial l}\int u^*_{nlm} H u_{nlm} d^3\mathbf{r} = \frac{\partial}{\partial l}\int u^*_{nlm} H_{eff} u_{nlm} d^3\mathbf{r}$$

LHS:

$$\frac{\partial E_n}{\partial l} = \frac{\partial}{\partial l}\left(-\frac{\mu Z^2 e^2}{2\hbar^2 n^2}\right) = \frac{\partial}{\partial l}\left(-\frac{\mu Z^2 e^2}{2\hbar^2 (r+l+1)^2}\right) = \frac{\mu Z^2 e^2}{\hbar^2 (r+l+1)^3} = \frac{\mu Z^2 e^2}{\hbar^2 n^3}$$

RHS:

$$\frac{\partial}{\partial l}\int u^*_{nlm} H_{eff} u_{nlm} d^3\mathbf{r} = \int \frac{\partial u^*_{nlm}}{\partial l} H_{eff} u_{nlm} d^3\mathbf{r} + \int u^*_{nlm} H_{eff} \frac{\partial u_{nlm}}{\partial l} d^3\mathbf{r} + \int u^*_{nlm} \frac{\partial H_{eff}}{\partial l} u_{nlm} d^3\mathbf{r}$$

Consider the first two terms.

$$\int \frac{\partial u^*_{nlm}}{\partial l} H_{eff} u_{nlm} d^3\mathbf{r} = E_n \int \frac{\partial u^*_{nlm}}{\partial l} u_{nlm} d^3\mathbf{r}$$

$$\int u^*_{nlm} H_{eff} \frac{\partial u_{nlm}}{\partial l} d^3\mathbf{r} = \int (H_{eff} u_{nlm})^* \frac{\partial u_{nlm}}{\partial l} d^3\mathbf{r} = E_n \int u^*_{nlm} \frac{\partial u_{nlm}}{\partial l} d^3\mathbf{r}$$

Adding these two terms we get

$$E_n \int \frac{\partial u^*_{nlm}}{\partial l} u_{nlm} d^3\mathbf{r} + E_n \int u^*_{nlm} \frac{\partial u_{nlm}}{\partial l} d^3\mathbf{r} = E_n \frac{\partial}{\partial l} \underbrace{\int u^*_{nlm} u_{nlm} d^3\mathbf{r}}_{1} = 0$$

We have $\frac{\partial H_{eff}}{\partial l}$ given by

$$\frac{\partial H_{eff}}{\partial l} = \frac{\partial}{\partial l}\left\{ -\frac{\hbar^2}{2\mu}\left[\frac{d^2}{dr^2} + \frac{2}{r}\frac{d}{dr}\right] + \frac{l(l+1)\hbar^2}{2\mu r^2} - \frac{Ze^2}{r} \right\} = \frac{(2l+1)\hbar^2}{2\mu r^2}$$

$$\therefore \int u_{nlm}^{*} \frac{\partial H_{eff}}{\partial l} u_{nlm} d^3\mathbf{r} = \int u_{nlm}^{*} \frac{(2l+1)\hbar^2}{2\mu r^2} u_{nlm} d^3\mathbf{r} = \frac{(2l+1)\hbar^2}{2\mu}\left\langle \frac{1}{r^2}\right\rangle_{nl}$$

Equating LHS and RHS we get,

$$\frac{\mu Z^2 e^2}{\hbar^2 n^3} = \frac{(2l+1)\hbar^2}{2\mu}\left\langle \frac{1}{r^2}\right\rangle_{nl}$$

$$\left\langle \frac{1}{r^2}\right\rangle_{nl} = \frac{2\mu^2 Z^2 e^2}{\hbar^4 n^3 (2l+1)} = \frac{2Z^2}{n^3(2l+1)a_0^2}$$

8.17 ISOTROPIC HARMONIC OSCILLATOR IN SPHERICAL COORDINATES

We have already seen isotropic harmonic oscillator in the Section 8.3 in Cartesian coordinates. Let us work out the same problem in spherical polar coordinates. The Hamiltonian for isotropic harmonic oscillator is

$$\begin{aligned} H &= \frac{p^2}{2M} + \frac{1}{2}M\omega^2(x^2+y^2+z^2) \\ &= \frac{p^2}{2M} + \frac{1}{2}M\omega^2 r^2 \end{aligned}$$

So the isotropic harmonic oscillator has become a problem in central potential. Therefore, the Schrödinger equation is

$$\left[-\frac{\hbar^2}{2M}\nabla^2 + \frac{1}{2}M\omega^2 r^2\right]u(r,\theta,\phi) = Eu(r,\theta,\phi)$$

Since $u(r,\theta,\phi)$ can be written as $R(r)Y_{lm}(\theta,\phi)$, we get the radial equation as

$$\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{dR}{dr}\right) + \frac{2M}{\hbar^2}\left[E - \frac{1}{2}M\omega^2 r^2 - \frac{l(l+1)}{2Mr^2}\hbar^2\right]R(r) = 0$$

Let us define $\rho = \alpha r$ where α is constant that can be chosen to suit our convenience.

$$\frac{\alpha^2}{\rho^2}\frac{d}{d\rho}\left(\rho^2\frac{dR}{d\rho}\right) + \frac{2M}{\hbar^2}\left[E - \frac{1}{2}\frac{M\omega^2\rho^2}{\alpha^2} - \frac{l(l+1)}{2M\rho}\hbar^2\alpha^2\right]R(\rho) = 0$$

Dividing throughout by α^2, we get

$$\frac{d^2R}{d\rho^2} + \frac{2}{\rho}\frac{dR}{d\rho} + \left[\frac{2ME}{\hbar^2\alpha^2} - \frac{M^2\omega^2}{\hbar^2\alpha^4}\rho^2 - \frac{l(l+1)}{\rho^2}\right]R(\rho) = 0$$

Let us now define α and λ as follows. α is chosen such that

$$\frac{M^2\omega^2}{\hbar^2\alpha^4} = 1 \quad \text{or} \quad \alpha^2 = \frac{M\omega}{\hbar}$$

$$\lambda = \frac{2ME}{\hbar^2\alpha^2} = \frac{2E}{\hbar\omega}$$

The radial equation now becomes

$$\frac{d^2R}{d\rho^2} + \frac{2}{\rho}\frac{dR}{d\rho} + \left[\lambda - \rho^2 - \frac{l(l+1)}{\rho^2}\right]R = 0 \tag{8.112}$$

For large ρ, the above equation becomes

$$\frac{d^2R}{d\rho^2} - \rho^2 R \simeq 0$$

Let us try $e^{-\rho^2/2}$ as a solution to the equation

$$\frac{dR}{d\rho} = -\rho e^{-\frac{\rho^2}{2}} \quad \frac{d^2R}{d\rho^2} = -e^{-\frac{\rho^2}{2}} + \rho^2 e^{-\frac{\rho^2}{2}}$$

$$\therefore \frac{d^2R}{d\rho^2} - \rho^2 R = \rho^2 e^{-\frac{\rho^2}{2}} - e^{-\frac{\rho^2}{2}} - \rho^2 e^{-\frac{\rho^2}{2}} \simeq 0 \text{ for large } \rho.$$

Therefore, the solution R (ρ) behaves like $e^{-\rho^2/2}$ for large ρ. For small ρ, the boundary condition demands $R(\rho)$ behaves like ρ^l. With these information, we write $R(\rho)$ as

$$R(\rho) = h(\rho)\rho^l e^{-\frac{\rho^2}{2}} \tag{8.113}$$

Let us change the variables from ρ and ξ by the relation

$$\xi = \rho^2$$

Then $R(\xi)$ becomes

$$R(\xi) = \xi^l e^{-\xi/2} h(\xi) \tag{8.114}$$

Substituting (8.114) in (8.112), we get

$$\xi \frac{d^2h}{d\xi^2} + \left[l + \frac{3}{2} - \xi\right]\frac{dh}{d\xi} + \frac{1}{4}(\lambda - 3 - 2l)h = 0 \tag{8.115}$$

This is almost similar to the Equation (8.93) for $F(\rho)$ in the case of hydrogen atom.

Let us try the solution

$$h(\xi) = \sum C_s \xi^s.$$

Substituting this solution in the Equation (8.114), we get

$$\sum_{s=0}^{\infty} C_s\left[s(s-1) + \left(l + \frac{3}{2}\right)s\right]\xi^{s-1} - C_s\left[s - \frac{1}{4}(\lambda - 3 - 2l)\right]\xi^s = 0$$

Consider the first term

$$\begin{aligned}\sum_{s=0}^{\infty} C_s\left[s(s-1) + \left(l + \frac{3}{2}\right)s\right]\xi^{s-1} &= \sum_{s=1}^{\infty} C_s\left[s(s-1) + \left(l + \frac{3}{2}\right)s\right]\xi^{s-1} \\ &= \sum_{r=0}^{\infty} C_{r+1}[(r+1)r + \left(l + \frac{3}{2}\right)(r+1)]\xi^r \\ &= \sum_{s=0}^{\infty} C_{s+1}[(s+1)s + \left(l + \frac{3}{2}\right)(s+1)]\xi^s\end{aligned}$$

$$\therefore \sum\left\{C_{s+1}\left[(s+1)s - \left(l + \frac{3}{2}\right)(s+1)\right] - C_s\left[s - \frac{1}{4}(\lambda - 3 - 2l)\right]\right\}\xi^s = 0$$

Equating the coefficient of ξ^s separately equal to zero, we get

$$C_{s+1} = \frac{s - \frac{1}{4}(\lambda - 3 - 2l)}{(s+1)\left(s - l - \frac{3}{2}\right)} C_s \tag{8.116}$$

Proceeding in the same way as in the case of hydrogen atom, we can show that the $R(r)$ diverges if $h(\xi)$ is given by infinite series. The only way of getting a finite, well-behaving solution is if we make the series as a polynomial is in ξ. This can be done by choosing $\frac{1}{4}(\lambda - 3 - 2l) = n'$, where n' is an integer.

i.e.,
$$\frac{1}{4}(\lambda - 3 - 2l) = n',\ n' \text{ is an integer}$$

$$\lambda = 4n' + 3 + 2l \tag{8.117}$$

∴ The energy eigenvalue E is given by

$$E = (4n' + 3 + 2l)\frac{\hbar\omega}{2}$$

$$= \left(2n' + l + \frac{3}{2}\right)\hbar\omega$$

Comparing this equation with the Equation (8.30), we can identify n as

$$n = 2n' + l \tag{8.118}$$

This relation implies both n and l together, either odd or even. For a given energy level, l and n' have to be such that the possible values of l are $n, n-2, n-4\ldots.$

$$l = \begin{cases} n, n-2, n-4, \ldots, 1 & \text{odd } n \\ n, n-2, n-4, \ldots, 0 & \text{even } n \end{cases}$$

Using (8.117), we get

$$\lambda = 2n + 3$$

Feeding this expression in (8.115), we get

$$\xi \frac{d^2 h}{d\xi^2} + \left[l + \frac{3}{2} - \xi\right]\frac{dh}{d\xi} + \frac{1}{2}[n - l]h = 0$$

Let us recall that the equation for associated Laguerre polynomial is

$$x\frac{d^2 L_q^k}{dx^2} + (k + 1 - x)\frac{dL_q^k}{dx} + (q - k)L_q^k = 0 \tag{8.119}$$

Comparing these two equations, we can conclude that $h(\xi)$ is Laguerre polynomial L_q^k, provided

$$k + 1 = l + \frac{3}{2} \text{ and } q - k = \frac{1}{2}(n - l)$$

$$\therefore k = l + \frac{1}{2} \text{ and } q = \frac{1}{2}(n + l + 1)$$

$$\therefore h(\xi) = L_{\frac{n+l+1}{2}}^{l+\frac{1}{2}}(\xi) = L_{\frac{n+l+1}{2}}^{l+\frac{1}{2}}(\alpha^2 r^2)$$

$$R_{nl}(r) = N_{nl}(\alpha r)^l e^{\frac{-\alpha^2 r^2}{2}} L^{l+\frac{1}{2}}_{\frac{n+l+1}{2}}(\alpha^2 r^2) \tag{8.120}$$

where N_{nl} is the normalization constant given by

$$N_{nl} = \left[\frac{\left(\frac{n+l}{2}\right)!(n-2l)!2^{n+l+2}\alpha^{2l+3}}{(n+l+1)!\pi^{1/2}} \right]^{1/2} \tag{8.121}$$

(Refer P. M. Mathews and A.Venkatesan)
If one is interested in determining the eigefunctions for a particular energy eigenvalue, it can be evaluated directly using the recursion relation (8.116) and normalization. For a given n, the function h_{nl} can be found as follows. From (8.118), we have

$$n' = \frac{n-l}{2}$$

The recursion relation is

$$C_{s+1} = \frac{s-n'}{(s+1)\left(s-l-\frac{3}{2}\right)} C_s$$

So, for a given set of n and l, first calculate $n' = n - l/2$. Then the series solution for h_{nl} is given by

$$h_{nl} = \begin{cases} c_0\left(1+\frac{c_2}{c_0}\rho^2+\frac{c_4}{c_0}\rho^4+.....\frac{c_{n'}}{c_0}\rho^{n'}\right) & \text{for even } n \\ c_1\left(\rho+\frac{c_3}{c_1}\rho^3-\frac{c_5}{c_1}\rho^5+.....\frac{c_{n'}}{c_1}\rho^{n'}\right) & \text{for odd } n \end{cases}$$

Normalization condition is

$$\int_0^\infty (N_{nl})^2(\alpha r)^{2l} e^{-\alpha^2 r^2}\left(1+\frac{c_2}{c_0}\alpha^2 r^2+\frac{c_4}{c_0}\alpha^4 r^4+.....\frac{c_{n'}}{c_0}\alpha^{n'} r^{n'}\right)^2 r^2\,dr = 1 \qquad \text{for even } n$$

$$\int_0^\infty (N_{nl})^2(\alpha r)^{2l} e^{-\alpha^2 r^2}\left(\alpha r+\frac{c_3}{c_1}\alpha^3 r^3+\frac{c_5}{c_1}\alpha^5 r^5+.....\frac{c_{n'}}{c_1}\alpha^{n'} r^{n'}\right)^2 r^2\,dr = 1 \qquad \text{for odd } n$$

Here, the constants c_0 and c_1 have been absorbed in the normalization constant N_{nl}.

Example 8.15 Determine the wave function for an isotropic harmonic oscillator corresponding to the energy eigenvalue $3/2\hbar\omega$.

Solution: The wave function for the isotropic harmonic oscillator in the spherical coordinates is given by

$$\psi_{nlm} = R_{nl}(r)Y_{lm}(\theta,\phi) = N_{nl}(\alpha r)^l e^{-\alpha^2 r^2/2} h_{nl}(r) Y_{lm}(\theta,\phi)$$

The quantum number corresponding to $3/2\hbar\omega$ is $n = 0$. Therefore, we have $n' = 0, l = 0$.

$$\therefore R_{00} = N_{00} e^{-\frac{\alpha^2 r^2}{2}}$$

The constant N_{00} is determined from $(N_{00})^2 \int_0^{\infty} e^{-\frac{\alpha^2 r^2}{2}} r^2 \, dr = 1.$

$$(N_{00})^2 \frac{\sqrt{\pi}}{4\alpha^3} = 1 \qquad \text{or} \quad N_{00} = \frac{2}{\pi^{1/4}} \alpha^{3/2}$$

$$\psi_{000}(r,\theta,\phi) = \frac{2}{\pi^{1/4}} \left(\frac{M\omega}{\hbar}\right)^{3/4} \exp\left(-\frac{M\omega}{2\hbar} r^2\right) Y_{00}$$

8.17.1 Degeneracy

The energy eigenstates of an isotropic harmonic oscillator are characterized by the set of three quantum numbers: (n, l, m). So we have

$$E_{n,l,m} = \left(n + \frac{3}{2}\right)\hbar\omega \text{ Eigenstates:}(n,\, l,\, m)$$

We have already seen that for a given n, the possible values of l are

$$l = \begin{cases} 0,\, 2,\, 4,\, \ldots,\, n-2, n & \text{even } n \\ 1,\, 3,\, 5,\, \ldots,\, n-2, n & \text{odd } n \end{cases}$$

The ground state has $n = 0$. This is a non-degenerate energy level with energy eigenvalue $3/2\hbar\omega$. The next energy eigenvalue is $5/2\hbar\omega$, corresponding to $n = 1$. For $n = 1$, the possible values of m are $1, 0, -1$. So $E_1 = 5/2\hbar\omega$ is three-fold degenerate. In general, for a given n, there are $1/2(n+1)(n+2)$ degenerate states corresponding to different values of l and m. The degeneracy D can be calculated as follows:

Even *n*		**Odd *n***	
l	2*l* + 1	*l*	2*l* + 1
0	1	1	3
2	5	3	7
4	9	5	11
:	:	:	:
n	2n + 1	*n*	2*n* + 1

For even n, there are $\left(\frac{n+2}{2}\right)$ possible values of l. The degeneracy D is given by

$$D = \frac{1}{2} \cdot \frac{(n+2)}{2} \cdot [1 + 2n + 1] = \frac{1}{2}(n+2)(n+1)$$

For odd n, there are $\left(\frac{n+1}{2}\right)$ the possible values of l.

The degeneracy D is given by

$$D = \frac{1}{2} \cdot \frac{(n+1)}{2} \cdot [3 + 2n + 1] = \frac{1}{2}(n+1)(n+2)$$

APPENDIX I

THE INTEGRAL $I^k_{qr,p}$ FOR ASSOCIATED LAGUERE POLYNOMIALS

In calculating various matrix elements involving hydrogen or hydrogen-like atoms radial wave function $R_{nl}(r)$, we regularly come across the following type of integral $I^k_{qr,p}$, which is defined as

$$I^k_{qr,p} = \int_0^\infty e^{-\rho} \rho^{k-1} L^k_q(\rho) L^k_r(\rho) \rho^p \, d\rho$$

From (8.110), we have

$$(-1)^k s^k (1-s)^{-k-1} \exp\left(-\frac{\rho s}{1-s}\right) = \sum_q \frac{s^q}{q!} L^k_q$$

$$(-1)^k t^k (1-t)^{-k-} \exp\left(-\frac{\rho t}{1-t}\right) = \sum_q \frac{t^q}{q!} L^k_q$$

Making use of these two equations, we get

$$\sum_q \sum_r \frac{s^q t^r}{q!r!} L^k_q \, L^k_r = s^k t^k (1-s)^{-k-1} \; (1-t)^{-k-1} \exp\left[-\rho\left(\frac{s}{1-s} + \frac{t}{1-t}\right)\right]$$

Multiply the above equation by $e^{-\rho} \;\; \rho^{k-1+p}$ and integrate to get

$$\sum_q \sum_r \frac{s^q t^r}{q!r!} \int_0^\infty L^k_q \, L^k_r e^{-\rho} \rho^{k-1+p} \, d\rho$$

$$= s^k t^k (1-s)^{-k-1} \; (1-t)^{-k-1} \int_0^\infty \exp\left[-\rho\left[\frac{s}{1-s} + \frac{t}{1-t} + 1\right]\right] \rho^{k-1+p} \, d\rho$$

$$= s^k t^k (1-s)^{-k-1} \; (1-t)^{-k-1} \int_0^\infty \exp\left[-\frac{\rho(1-st)}{(1-s)(1-t)}\right] \rho^{k-1+p} \, d\rho$$

It is easy to show that

$$\int_0^\infty x^n e^{-\alpha x} \, dx = \frac{n!}{\alpha^{n+1}}$$

Using the above result, we have

$$\int_0^\infty \rho^{k-1+p} \exp\left[-\frac{\rho(1-st)}{(1-s)(1-t)}\right] d\rho = \frac{(1-s)^{k+p} (1-t)^{k+p}}{(1-st)^{k+p}} (k+p-1)!$$

$$\therefore \sum_q \sum_r \frac{s^q t^r}{q!r!}\int_0^\infty L_q^k\, L_r^k e^{-\rho}\rho^{k-1+p}d\rho$$

$$= s^k t^k (1-s)^{-k-1}\ (1-t)^{-k-1}\,\frac{(1-s)^{k+p}(1-t)^{k+p}}{(1-st)^{k+p}}(k+p-1)!$$

Now, let us state another useful result.

$$(1-x)^{-n} = 1+nx+\frac{n(n+1)}{2!}x^2+\cdots=\sum_{\lambda=0}^{\infty}\frac{(n+\lambda-1)!}{\lambda!(n-1)!}x^{\lambda}$$

$$\therefore \sum_q \sum_r \frac{s^q t^r}{q!r!}\int_0^\infty L_q^k\, L_r^k e^{-\rho}\rho^{k-1+p}d\rho$$

$$= s^k t^k (1-s)^{p-1}(1-t)^{p-1}(k+p-1)!\sum_{\lambda=0}^{\infty}\frac{(st)^{\lambda}(k+p+\lambda-1)!}{\lambda!(k+p-1)!}$$

$$\sum_q \sum_r \frac{s^q t^r}{q!r!} I^k_{qr,p} = \sum_{\lambda=0}^{\infty}\frac{(st)^{k+\lambda}(1-t)^{p-1}(1-s)^{p-1}(k+p+\lambda-1)!}{\lambda!}$$

$$\therefore I^k_{qr,p} = q!r!\times \text{co.eff.of}\, s^q t^r \text{in}\left\{\sum_{\lambda=0}^{\infty}\frac{(st)^{k+\lambda}(1-t)^{p-1}(1-s)^{p-1}}{\lambda!}(k+p+\lambda-1)!\right\}$$

Special Case $I^k_{qq,p}$

$$I^k_{qq,p} = (q!)^2\times \text{co.eff.of}(st)^q \text{in}\left\{\sum_{\lambda=0}^{\infty}\frac{(st)^{k+\lambda}(1-t)^{p-1}(1-s)^{p-1}(k+p+\lambda-1!)}{\lambda!}\right\}$$

Using binomial theorem, we have

$$(1-s)^{p-1}(1-t)^{p-1} = \sum_{l=0}^{\infty}{}^{p-1}C_l(-s)^l\sum_{m=0}^{\infty}{}^{p-1}C_m(-t)^m$$

$$= \sum_{l=0}^{\infty}\sum_{m=0}^{\infty}{}^{p-1}C_l\,{}^{p-1}C_m(-1)^{l+m}s^l t^m$$

$$\therefore \sum_{\lambda=0}^{\infty}\frac{(st)^{k+\lambda}(1-s)^{p-1}(1-t)^{p-1}(k+p+\lambda-1)!}{\lambda!}$$

$$= \sum_{\lambda=0}^{\infty}\sum_{l=0}^{p-1}\sum_{m=0}^{p-1}\frac{{}^{p-1}C_l\,{}^{p-1}C_m s^{k+\lambda+l}t^{k+\lambda+m}(-1)^{l+m}}{\lambda!}(k+p+\lambda-1)!$$

We have to choose λ, l and m such that

$$k+l+\lambda=q,\ k+\lambda+m=q \text{ or } \lambda=q-k-l,\ \lambda=q-k-m$$

This implies that l and m should be equal. Maximum value of l or m is $(p-1)$. Therefore, possible values of l and m are given

$$(l, m)\text{: } (0, 0), (1, 1), (2, 2),\ldots, (p-1, p-1)$$

$$\therefore I^k_{qq,p} = (q!)^2 \left\{ {}^{p-1}C_0 {}^{p-1}C_0 \frac{(q+p-1)!}{(q-k)!} + {}^{p-1}C_1 \frac{(q+p-2)!}{(q-k-1)!} + \right.$$

$$\left. {}^{p-1}C_2 {}^{p-1}C_2 \frac{(q+p+-3)!}{(q-k-2)!} + \ldots {}^{p-1}C_{p-1} {}^{p-1}C_{p-1} \frac{q!}{(q-k-p+1)!} \right\}$$

APPENDIX II

KRAMER'S RECURSION RELATION

Kramer's recursion relation for hydrogen like atoms is given by

$$\frac{(s+1)}{n^2}\langle r^s\rangle_{nl} - a_0(2s+1)\langle r^{s-1}\rangle_{nl} + \frac{s}{4}\{(2l+1)^2 - s^2\}a_0^2\langle r^{s-2}\rangle_{nl} = 0$$

where

$$\langle r^s\rangle_{nl} = \int_0^\infty r^s P_{nl}(r)dr = \int_0^\infty r^s R_{nl}(r) r^2 dr$$

We closely follow the proof of this recursion relation given in Wikipedia.

The radial equation for hydrogen atom is $\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{dR_{nl}}{dr}\right) + \frac{2\mu}{\hbar^2}\left[E_n + \frac{Ze^2}{r} - \frac{l(l+1)}{2\mu r^2}\hbar^2\right]R_{nl} = 0$

Let us define $\chi_{nl}(r)$ as

$$R_{nl}(r) = \frac{\chi_{nl}(r)}{r}$$

Substituting this expression in the above equation, we get

$$\frac{d^2\chi_{nl}}{dr^2} + \left[\frac{2\mu}{\hbar^2}E_n + \frac{2\mu}{\hbar^2}\cdot\frac{Ze^2}{r} - \frac{l(l+1)}{r^2}\right]\chi_{nl} = 0 \qquad (8.122)$$

For hydrogen-like atom,

$$\frac{2\mu E_n}{\hbar^2} = -\frac{\mu Z^2 e^4}{2\hbar^2 n^2}\cdot\frac{2\mu}{\hbar^2} = -\frac{Z^2}{a_0^2 n^2}$$

and
$$\frac{2\mu Z}{\hbar^2}\cdot\frac{e^2}{r} = -\frac{2Z}{a_0 r}$$

For hydrogen atom, $Z = 1$. The Equation (8.122) now becomes

$$\frac{d^2\chi_{nl}}{dr^2} + \left[-\frac{1}{a_0^2 n^2} + \frac{2}{a_0 r} - \frac{l(l+1)}{r^2}\right]\chi_{nl} = 0$$

$$\therefore \frac{d^2\chi_{nl}}{dr^2} = \left[\frac{l(l+1)}{r^2} - \frac{2}{a_0 r} + \frac{1}{a_0^2 n^2}\right]\chi_{nl} \qquad (8.123)$$

In terms of $\chi_{nl}(r)$, we get

$$\langle r^s \rangle = \int_0^\infty r^s R_{nl}^2(r) r^2 dr = \int_0^\infty r^s \chi_{nl}^2(r) dr$$

Let us consider the integral $\int_0^\infty \chi_{nl} r^s \frac{d^2\chi_{nl}}{dr^2} dr$.Using Equation (8.123), we get

$$\int_0^\infty \chi_{nl} r^s \frac{d^2\chi_{nl}}{dr^2} dr = \int_0^\infty r^s \chi_{nl}\left[\frac{l(l+1)}{r^2} - \frac{2}{a_0 r} + \frac{1}{a_0^2 n^2}\right]\chi_{nl} dr \tag{8.124}$$

$$= l(l+1)\langle r^{s-2}\rangle - \frac{2}{a_0}\langle r^{s-1}\rangle + \frac{1}{a_0^2 n^2}\langle r^s \rangle \tag{8.125}$$

But the LHS of (8.124) can be evaluated directly.

$$\int_0^\infty \chi_{nl} r^s \frac{d^2\chi_{nl}}{dr^2} dr = \int_0^\infty \chi_{nl} r^s \frac{d}{dr}\left(\frac{d\chi_{nl}}{dr}\right) dr$$

$$= \chi_{nl} r^s \frac{d\chi_{nl}}{dr}\Big|_0^\infty - \int_0^\infty \frac{d}{dr}(\chi_{nl} r^s)\frac{d\chi_{nl}}{dr} dr$$

The first term is zero. So we have

$$\begin{aligned}\int_0^\infty \chi_{nl} r^s \frac{d^2\chi_{nl}}{dr^2} dr &= -\int_0^\infty \frac{d}{dr}(\chi_{nl} r^s)\frac{d\chi_{nl}}{dr} \\ &= -\int_0^\infty \frac{d\chi_{nl}}{dr} r^s \frac{d\chi_{nl}}{dr} dr - s\int_0^\infty \chi_{nl} r^{s-1}\frac{d\chi_{nl}}{dr} dr\end{aligned} \tag{8.126}$$

Now consider the integral $\int_0^\infty \chi_{nl} r^s \frac{d\chi_{nl}}{dr} dr$.

$$\int_0^\infty \chi_{nl} r^s \frac{d\chi_{nl}}{dr} dr = \chi_{nl} r^s \chi_{nl}\Big|_0^\infty - \int_0^\infty \frac{d}{dr}(\chi_{nl} r^s)\chi_{nl} dr$$

The first term is zero.

$$\therefore \int_0^\infty \chi_{nl} r^s \frac{d\chi_{nl}}{dr} dr = -\int_0^\infty \frac{d}{dr}(\chi_{nl} r^s)\chi_{nl} dr$$

$$= -\int_0^\infty \frac{d\chi_{nl}}{dr} r^s \chi_{nl} - \int_0^\infty \chi_{nl} s r^{s-1}\chi_{nl} dr$$

Rearranging the terms in the above equation, we get

$$\therefore \int_0^\infty \chi_{nl} r^s \frac{d\chi_{nl}}{dr} dr = -\tfrac{s}{2}\langle r^{s-1}\rangle \tag{8.127}$$

Consider the first term in (8.126).

$$-\int_0^\infty \frac{d\chi_{nl}}{dr} r^s \frac{d\chi_{nl}}{dr}\,dr = -\frac{1}{s+1} r^{s+1} \frac{d\chi_{nl}}{dr}\frac{d\chi_{nl}}{dr}\bigg|_0^\infty + \frac{1}{s+1}\int_0^\infty r^{s+1}\frac{d}{dr}\left[\frac{d\chi_{nl}}{dr}\right]^2 dr$$

The first term is zero. So we have

$$-\int_0^\infty \frac{d\chi_{nl}}{dr} r^s \frac{d\chi_{nl}}{dr}\,dr = \frac{2}{s+1}\int_0^\infty r^{s+1}\frac{d\chi_{nl}}{dr}\frac{d^2\chi_{nl}}{dr^2}\,dr$$

Using the Equation (8.123), we get

$$-\int_0^\infty \frac{d\chi_{nl}}{dr} r^s \frac{d\chi_{nl}}{dr}\,dr = \frac{2}{s+1}\int_0^\infty r^{s+1}\frac{d\chi_{nl}}{dr}\left[\frac{l(l+1)}{r^2} - \frac{2}{a_0 r} + \frac{1}{a_0 n^2}\right]\chi_{nl}\,dr \tag{8.128}$$

$$= \frac{2}{s+1}\left[-\frac{l(l+1)}{2}(s-1)\langle r^{s-2}\rangle + \frac{2}{a_0}\frac{s}{2}\langle r^{s-1}\rangle - \frac{1}{a_0^2 n^2}\frac{s+1}{2}\langle r^s\rangle\right] \tag{8.129}$$

We have made use of the Equation (8.127) to each term in (8.128) separately in arriving at (8.129). Again making use of (8.127), we get the second term in (8.126) to be

$$-s\int_0^\infty \chi_{nl} r^{s-1}\frac{d\chi_{nl}}{dr}\,dr = \frac{s(s-1)}{2}\langle r^{s-2}\rangle \tag{8.130}$$

Using (8.129) and (8.130) in (8.126), we get

$$\int_0^\infty \chi_{nl} r^s \frac{d^2\chi_{nl}}{dr^2}\,dr = \frac{2}{s+1}\left[-\frac{l(l+1)}{2}(s-1)\langle r^{s-2}\rangle + \frac{s}{a_0}\langle r^{s-1}\rangle - \frac{1}{a_0^2 n^2}\frac{s+1}{2}\langle r^s\rangle\right] + \frac{s(s-1)}{2}\langle r^{s-2}\rangle \tag{8.131}$$

Equating (8.131) and (8.125), we get

$$\frac{s+1}{n^2}\langle r^s\rangle - a_0(2s+1)\langle r^{s-1}\rangle + \frac{s}{4}[(2l+1)^2 - s^2]\langle r^{s-2}\rangle a_0^2 = 0$$

For hydrogen-like atom, which has a nucleus of charge *Ze*, the above relation becomes

$$\frac{s+1}{n^2}\langle r^s\rangle - \frac{a_0}{Z}(2s+1)\langle r^{s-1}\rangle + \frac{s}{4}[(2l+1)^2 - s^2]\langle r^{s-2}\rangle\frac{a_0^2}{Z^2} = 0$$

EXERCISES

1. What are the boundary conditions for radial wave functions?
2. How do you define radial momentum?
3. Why do you need to define radial momentum p_r as $p_r = \frac{1}{2}\left(\frac{\mathbf{r}}{r}\cdot\mathbf{p} + \mathbf{p}\cdot\frac{\mathbf{r}}{r}\right)$?
4. Determine the maxima and minima of the radial probability distribution for $P_{20}(r)$.

5. Calculate the probability of a 1*s* and 2*p* electron for hydrogen atom in the range
   (i) $0 < r < a_0$ (ii) $a_0 < r < 3a_0$
6. Evaluate $\langle r^2 \rangle_{nl}$ using Kramer's relation.
7. Show that $\int\limits_0^\infty x^n e^{-\alpha x} dx = \dfrac{n!}{\alpha^{n+1}}$.
8. What is the normalization condition for the radial wave function?
9. What are the quantum numbers describing the energy eigenstate for a three-dimensional isotropic harmonic oscillator ?
10. Determine the wave function for isotropic harmonic oscillator corresponding to the energy eigenvalue $\frac{5}{2}\hbar\omega$ in spherical and Cartesian coordinate systems. Comment on the result.
11. For the same eigenvalue of the isotropic harmonic oscillator, the wave functions in the spherical and Cartesian coordinate systems are different. How do you explain this difference?
12. The differential Equation (8.95) for the radial wave function for hydrogen atom is not in the form of Sturm Liouville equation. Get it in the standard form of Sturm Liouville equation and hence get the orthoganality relation for the associated Laguerre Polynomials.

REFERENCES

1. M. Alonso and H. Valk, 1972. *Quantum Mechanics Principles and Applications*, Massachusetts: Addison-Wesley Publishing Company.
2. P. M. Mathews and A. Venkatesan. *A Text Book of Quantum Mechanics*, New Delhi: Tata McGraw Hill Publishing Company Ltd.
3. S. Gasiorowicz, 1974. *Quantum Physics*, New York: John Wiley & Sons.
4. http/www.physics.the tangentbundle.net/wiki/Quantum mechanics/hydrogenatom/Kramers 27% relation.

9

Bra and Ket Vector Formalism and Symmetries

In the previous chapters, quantum mechanics, based on Schrödinger's differential equations called wave mechanics, was discussed. Here the operators were differential operators like $-i\hbar\, d/dx$ for momentum or multiplicative factors like x^n. This is not the only formulation of quantum mechanics. There is indeed another formulation of quantum mechanics called matrix mechanics in which all the dynamical variables are represented by matrices. For instance, the operators X_{op} and P_{op} are represented by infinite dimensional matrices $[x]$ and $[p]$ such that $[x][p]-[p][x]=i\hbar[1]$ where $[1]$ is an infinite dimensional unit matrix. In fact, matrix mechanics was developed by Heisenberg and others earlier to Schrödinger's wave mechanics. The equivalence between these two formalisms was established long time back. Dirac (1939) developed a new mathematical language called abstract-state vector formalism, which is very general, and it is independent of specific representations like differential operators or matrix operators. Both these formalisms emerge as particular representations of Dirac's abstract-state vector formalism.

9.1 KET VECTORS

We know that quantum states form a linear vector space. Dirac suggested that a new quantity called ket vector can be associated with each quantum state. The ket vector is written as $|\,\rangle$ (part of bracket notation $\langle\ \rangle$). The state vector is characterized by writing all the necessary information about the state of the system inside the ket symbol. For instance, suppose we want to describe the ground state of a harmonic oscillator. We can write the ket vector as

$$|\,\textit{ground state of harmonic oscillator}\rangle$$

A better way of representing the ground state of harmonic oscillator is to write the ket vector as

$$\left|E=\frac{1}{2}\hbar\omega\right\rangle \text{ or } |\,n=0\rangle$$

We can adopt a still simpler way. It is simply written as 10 >.

In the same way, the ground state of hydrogen atom can be written as $|\,\textit{ground state of hydrogen atom}\rangle$ or $|\,E = -13.6 \text{ ev}\,\rangle$ or $|\,n = 1, l = 0, m_l = 0\rangle$ or simply $|\,100\rangle$.

What do these ket vectors represent in concrete form? Can we take $|\,0\rangle$ as a function like $e^{-\alpha^2 x^2}$ or a column vector with some numbers? In Dirac's ket vector formalism, the ket vectors $|\,0\rangle$ or $|\,100\rangle$ do not represent any particular coordinate function or a column vector or a function of any other variables. They are simply taken as it is. So these vectors are taken as abstract vectors without assigning any specific (functional) form.

Obviously, one should form rules, notations and interpretations to manipulate such ket vectors consistent with the requirements of quantum mechanics. Note that in some books $|\,\psi\rangle$ is identified as $\psi(x)$ itself. This way of writing helps the students in handling the ket vectors initially.

All the ket vectors $|\,\psi\rangle$, $|\,\phi\rangle$, $|\,\chi\rangle$ are elements of the linear vector space H. They satisfy all the axioms of a linear vector space. So if $|\,\psi\rangle$ is a state vector, $c\,|\,\psi\rangle$ (where c is a complex number) is also a state vector. If $|\,\psi\rangle$ and $|\,\phi\rangle$ are state vectors $c_1\,|\,\psi\rangle + c_2\,|\,\phi\rangle$ is also a state vector.

9.2 BRA VECTORS

Along with the linear vector space of ket vectors, let us postulate one more linear vector space whose elements are another kind of vectors called bra vectors denoted by $\langle\psi\,|$. The introduction of the concept of bra vectors requires the concept of the dual of a vector space. However, we will not present here the idea of the dual of a vector space. It is sufficient to say that under certain conditions, there exists a one-to-one correspondence between a ket vector $|\,\psi\rangle$ and a bra vector $\langle\,\psi|$.

i.e.,
$$|\,\psi\rangle \leftrightarrow \langle\psi\,| \tag{9.1}$$

Note that $|\,\psi\rangle$ and $\langle\psi\,|$ belong to two different linear vector spaces. Dirac calls them complex imaginary of each other. They are known as Hermitian conjugate of each other.

i.e.,
$$|\,\psi\rangle^{\dagger} = \langle\psi\,| \tag{9.2}$$

$$(c\,|\,\psi\rangle)^{\dagger} = \langle\psi\,|\,c^* \tag{9.3}$$

9.3 SCALAR PRODUCT

Having defined two linear vector spaces of ket and bra vectors, we now define a product of these vectors and choose appropriate symbol to denote such a product. Let us recall that in the case of geometrical vectors, we have two kinds of product between two vectors: scalar product giving a number and vector product yielding a vector. In the case of ket vectors and bra vectors, we have three kinds of products: scalar product, direct product and outer product.

The scalar product between a bra vector $\langle\phi\,|$ and a ket vector $|\,\psi\rangle$ is defined as

$$\langle\phi\,|\,\psi\rangle \rightarrow \text{a complex number} \tag{9.4}$$

The scalar product between $\langle\phi|$ and $|\psi\rangle$ could have been written as $(\langle\phi|)\cdot(|\psi\rangle)$ (similar to $A.B$) or $(\langle\phi|,|\psi\rangle)$ (a symbol used for scalar product in Chapter 4). The notation $\langle\phi|\psi\rangle$ is a much simpler notation. Whenever we see the notation $\langle\phi|\psi\rangle$ we have to recognize that it is a complex number.

Of course, the concept of a scalar product is more than assigning a complex number for a given bra vector and ket vector. We have already discussed what is meant by a scalar product in the context of a linear vector space in Chapter 4. It is a rule to assign a number for the pair $\langle\phi|$ and $|\psi\rangle$ subject to the following conditions:

1. $\langle\phi|\psi\rangle = \langle\psi|\phi\rangle^*$ (9.5)
2. $\langle\phi|(c_1|\psi_1\rangle + c_2|\psi_2\rangle) = c_1\langle\phi|\psi_1\rangle + c_2\langle\phi|\psi_2\rangle$ (9.6)
3. $\langle\phi|\phi\rangle \geq 0$ with $\langle\phi|\phi\rangle = 0$ implying $|\phi\rangle = 0$ (9.7)

Note that $\langle\phi|C|\psi\rangle = C\langle\phi|\psi\rangle$. This follows from (9.6).

Orthoganality

Once a scalar product is defined, we can define orthoganality between $\langle\phi|$ and $|\psi\rangle$ as $\langle\phi|\psi\rangle = 0$. i.e., $\langle\phi|$ and $|\psi\rangle$ are orthogonal if $\langle\phi|\psi\rangle = 0$.

9.4 ABSTRACT OPERATOR

The important feature of an operator is that when it acts on an element of vector space, it produces another element. Let us define an abstract operator $\hat{A}$, which, when it acts on a ket vector, produces another ket.

i.e.,

$$\hat{A}\,|\psi\rangle = |\phi\rangle \tag{9.8}$$

Again it is to be noted that $\hat{A}$ is neither a differential nor a matrix operator. The reader may wonder how one can use such an operator which does not have any concrete form familiar to us. As we progress in this chapter, one can recognize that a number of important results can be obtained from such operators even if we are unable to visualize them in the familiar form.

Differentiation of a Ket

Just as we differentiate a wave function, we can also differentiate (or integrate) a ket vector. For instance, $\frac{d}{dt}|\psi(t)\rangle$ is defined as

$$\frac{d}{dt}|\psi(t)\rangle = \underset{\Delta t\to 0}{\mathrm{Lt}}\frac{|\psi(t+\Delta t)\rangle - |\psi(t)\rangle}{\Delta t} \tag{9.9}$$

However, this is only a formal definition. We cannot get a result like $d/dt\,\sin\omega t = \omega\cos\omega t$. We may simply write $d/dt\,|\psi(t)\rangle = |\phi(t)\rangle$. Note that the differential operator d/dt is not an abstract operator $\hat{A}$.

The abstract operator $\hat{A}$ can act on both objects to the right and left of it.

Let us consider $\hat{A}$ acting on $|\psi\rangle$.

$$\hat{A}\,|\psi\rangle = |\phi\rangle \tag{9.10}$$

Consider the scalar product between a bra vector $\langle\chi\,|$ and the ket vector $|\,\phi\rangle$.

$$\langle\chi\,|\,\phi\rangle = \langle\chi\,|\,\hat{A}\,|\,\psi\rangle \tag{9.11}$$

This structure suggests the following interpretation also.

$$\langle\chi\,|\,\phi\rangle = \underbrace{\langle\chi|\,\hat{A}}_{\langle\chi'|}|\psi\rangle \tag{9.12}$$

$$\langle\chi\,|\,\phi\rangle = \langle\chi'\,|\,\psi\rangle \tag{9.13}$$

provided we identify $\langle\chi'\,| = \langle\chi\,|\,\hat{A}$. So we have

$$\hat{A}\,|\,\psi\rangle = |\,\phi\rangle$$
$$\langle\chi\,|\,\hat{A} = \langle\chi'\,|$$

Thus, $\hat{A}$ can act on the ket from the right, and it can act on the bra from left.

Linear Operator

We will restrict ourselves to linear abstract operators. An operator $\hat{A}$ is said to be linear operator if it obeys

$$\hat{A}(c_1\,|\,\psi_1\rangle + c_2\,|\,\psi_2\rangle) = c_1\hat{A}\,|\,\psi_1\rangle + c_2\hat{A}\,|\,\psi_2\rangle \tag{9.14}$$

It implies that

$$\hat{A}(c\,|\,\phi\rangle) = c\hat{A}\,|\,\phi\rangle \tag{9.15}$$

9.4.1 Outer Product of a Ket Vector and a Bra Vector

The outer product between a ket vector $|\,\phi\rangle$ and a bra vector $\langle\psi\,|$ is defined as $|\,\phi\rangle\,\langle\psi\,|$. This is not a scalar number. To understand the nature of this product, let us put another ket vector $|\,\chi\rangle$ to the right of $|\,\phi\rangle\langle\psi\,|$. We get $|\,\phi\rangle\langle\psi\,|\,\chi\rangle$.

Note $\langle\psi\,|\,\chi\rangle$ is a complex number. So we can write the above expression as $|\,\phi\rangle\underbrace{\langle\psi\,|\,\chi\rangle}$ or $\underbrace{\langle\psi\,|\,\chi\rangle}|\,\phi\rangle$

$$\therefore\quad |\,\phi\rangle\langle\psi\,|\,\chi\rangle = \langle\psi\,|\,\chi\rangle\,|\,\phi\rangle = c\,|\,\phi\rangle \tag{9.16}$$

where c is a complex number.

The structure of this equation suggests that $|\phi\rangle\langle\psi|$ is an operator. When it acts on a ket $|\chi\rangle$ it produces another ket $|\phi\rangle$. We can further test another property of an abstract operator in the sense it can act on a bra vector from the right. Let us put $\langle\chi|$ to the left of $|\phi\rangle\langle\psi|$.

$$\langle\chi|\phi\rangle\langle\psi| = d\langle\psi| \tag{9.17}$$

where d is a complex number.

It is easy to prove that $\hat{P} = |\phi\rangle\langle\psi|$ is a linear operator.

$$\begin{aligned} \hat{P}(|\chi\rangle + |\chi'\rangle) &= |\phi\rangle\langle\psi|(|\chi\rangle + |\chi'\rangle) \\ &= |\phi\rangle\langle\psi|\chi\rangle + |\phi\rangle\langle\psi|\chi'\rangle = \hat{P}|\chi\rangle + \hat{P}|\chi'\rangle \\ \hat{P}(c|\chi\rangle) &= |\phi\rangle\langle\psi|(\langle c|\chi\rangle) \\ &= |\phi\rangle\langle\psi|c|\chi\rangle = c|\phi\rangle\langle\psi|\psi\rangle = c\hat{P}|\chi\rangle \end{aligned} \tag{9.18}$$

$\therefore$ $|\phi\rangle\langle\psi|$ is a linear operator.

9.5 ADJOINT OF AN OPERATOR

Let $\hat{A}$ be a linear operator. Then, we have

$$\hat{A}|\phi\rangle = |\psi\rangle \tag{9.19}$$

What is the operator that will produce $\langle\psi|$ from $\langle\phi|$? i.e., we need an operator $\hat{B}$ such that $\langle\psi| = \langle\phi|\hat{B}$.

Normally, we do not write a separate operator $\hat{B}$ which may be the same as $\hat{A}$ or may be different from it. The standard notation is to write $A^\dagger$ in the place of $\hat{B}$. So we have

$$\begin{aligned} |\psi\rangle &= \hat{A}|\phi\rangle \\ \langle\psi| &= \langle\phi|\hat{A}^\dagger \end{aligned} \tag{9.20}$$

Take another arbitrary ket $|\chi\rangle$ and form a scalar product as given below:

$$\langle\chi|\psi\rangle = \langle\chi|\underbrace{\hat{A}|\phi\rangle}_{|\psi\rangle} = \underbrace{\langle\phi|\hat{A}^\dagger}_{\langle\psi|}|\chi\rangle^*$$

i.e.,
$$\langle\chi|\hat{A}|\phi\rangle = \langle\phi|\hat{A}^\dagger|\chi\rangle^* \tag{9.21}$$

In fact, this relation is taken as the definition of the adjoint of the operator $\hat{A}$.

Self-adjoint Operator

A linear operator $\hat{A}$ is self-adjoint if its adjoint is the same as the operator itself.

$$\hat{A}^\dagger = \hat{A} \tag{9.22}$$

$$\therefore |\psi\rangle = \hat{A}|\phi\rangle \quad \Rightarrow \quad \langle\phi|\hat{A} = \langle\psi|$$

$$\langle\chi|\hat{A}|\phi\rangle = \langle\phi|\hat{A}|\chi\rangle^* \tag{9.23}$$

This relation can itself be taken as the definition of a self-adjoint operator.

Example 9.1 Prove that $(\hat{A}\hat{B})^\dagger = \hat{B}^\dagger\hat{A}^\dagger$

Solution: Consider $\langle\phi|\hat{A}\hat{B}|\psi\rangle$. This can be rewritten in two ways.

First, let us write $|\chi\rangle = \hat{A}\hat{B}|\psi\rangle$. Then, using (9.20), we have $\langle\chi| = \langle\psi|(\hat{A}\hat{B})^\dagger$

$$\therefore \langle\phi|\underbrace{\hat{A}\hat{B}|\psi\rangle}_{|\chi\rangle} = \langle\chi|\phi\rangle^* = \langle\psi|(\hat{A}\hat{B})^\dagger|\phi\rangle^*$$

Let us rewrite $\langle\phi|\hat{A}\hat{B}|\psi\rangle$ in a different way. Let $|\phi'\rangle$ and $|\psi'\rangle$ be given by

$$|\phi'\rangle = \hat{A}^\dagger|\phi\rangle \text{ and } |\psi'\rangle = \hat{B}|\psi\rangle$$

Then we have

$$\langle\phi|\hat{A} = \langle\phi'| \text{ and } \langle\psi|\hat{B}^\dagger = \langle\psi'|$$

$$\therefore \langle\phi|\hat{A}\hat{B}|\psi\rangle = \underbrace{\langle\phi|\hat{A}}_{\langle\phi'|}\underbrace{\hat{B}|\psi\rangle}_{|\psi'\rangle} = \langle\phi'|\psi'\rangle = \langle\psi'|\phi'\rangle^* = \langle\psi|\hat{B}^\dagger\hat{A}^\dagger|\phi\rangle^*$$

$$\langle\phi|\hat{A}\hat{B}|\psi\rangle = \langle\psi|(\hat{A}\hat{B})^\dagger|\phi\rangle^*$$

$$\langle\phi|\hat{A}\hat{B}|\psi\rangle = \langle\psi|(\hat{A}\hat{B})^\dagger|\phi\rangle^* = \langle\psi|\hat{B}^\dagger\hat{A}^\dagger|\phi\rangle^*$$

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

9.6 BASIS VECTORS

The concept of basis vectors in a linear vector space has already been discussed in Chapter 4. The same ideas can be extended to Hilbert space of Dirac's ket vectors.

Let $\{|i\rangle\}$ be a complete set of linearly independent vectors. Let us choose them as an orthonormal set. i.e., $\langle i|j\rangle = \delta_{ij}$. These sets of vectors are said to form basis vectors of the linear vector space.

This implies that any arbitrary vector $|\psi\rangle$ belonging to the given Hilbert space can be expressed as a linear combination of these vectors.

$$|\psi\rangle = \sum_n C_n |n\rangle \tag{9.24}$$

Eigenvalue Equation

The eigenvalue equation is

$$\hat{A}|a\rangle = a|a\rangle$$

Here $|a\rangle$ is the eigenket corresponding to eigenvalue a. The eigenvalues may be discrete or continuous.

In the case of discrete eigenvalues, the eigenvalue equation can be written as

$$\hat{A}|i\rangle = a_i|i\rangle \qquad i = 1,2,3,..$$

Example 9.2 Prove that the eigenvalues of a self-adjoint operator are real and the eigenkets belonging to distinct eigenvalues are orthogonal to each other.

Solution: The eigenvalue equation is

$$\hat{A}|i\rangle = a_i|i\rangle$$
$$\hat{A}|j\rangle = a_j|j\rangle$$

The self-adjointness of $\hat{A}$ implies

$$\langle i|\hat{A}|j\rangle = \langle j|\hat{A}|i\rangle^*$$
$$\langle i|a_j|j\rangle = \langle j|a_i|i\rangle^* \text{ or } a_j\langle i|j\rangle = a_i^*\langle \boldsymbol{j}|i\rangle^* = a_i^*\langle i|j\rangle$$
$$\therefore (a_j - a_i^*)\langle i|j\rangle = 0$$

Case 1 $\quad i = j \quad (a_i - a_i^*)\langle i|i\rangle = 0$

Since $\langle i|i\rangle$ cannot be zero, $(a_i - a_i^*) = 0$ or $a_i = a_i^*$

$\therefore a_i$ is a real number.

Case 2 i and j are different.

$$(a_i - a_j)\langle i|j\rangle = 0$$

Since $\quad (a_i - a_j) \neq 0,\ \langle i|j\rangle = 0$

Without losing generality, we can have $\langle i|i\rangle = 1$.

$$\therefore \langle i|j\rangle = \delta_{ij}$$

9.6.1 Complete Set

The eigenkets of a self-adjoint operator form a complete set. Then they form a basis set for a linear vector space. This implies that any arbitrary state vector $|\psi\rangle$ belonging to that linear vector space can be expanded as a linear combination of the eigenket $|i\rangle$.

$$\hat{A}|n\rangle = a_n|n\rangle \quad n = 1, 2, 3\ldots$$

$$\langle\psi| = \sum C_n|n\rangle$$

The coefficient C_n is given by $C_n = \langle n|\psi\rangle$

$$\therefore\ |\psi\rangle = \sum_n \underbrace{\langle n|\psi\rangle}_{\text{a complex number}} |n\rangle$$

$$= \sum_n |n\rangle\langle n|\psi\rangle \tag{9.25}$$

The structure of this equation suggests

$$|\psi\rangle = \hat{I}|\psi\rangle$$

provided we identify the operator $\Sigma|n\rangle|n\rangle$ as $\hat{I}$.

$$\therefore\ \hat{I} = \sum|n\rangle\langle n| \tag{9.26}$$

9.6.2 Projection Operator

The operator $|n\rangle\langle n|$ is known as projection operator. Let us consider the effect of $|n\rangle\langle n|$ on $|\psi\rangle$.

$$P_n|\psi\rangle = |n\rangle\langle n|\psi\rangle$$

$$= |n\rangle\langle n|\left(\sum_m C_m|m\rangle\right)$$

$$= \sum_m |n\rangle\langle n|m\rangle C_m = \sum_m |n\rangle C_m\delta_{mn} = C_n|n\rangle$$

i.e., P_n projects out the 'components' corresponding to the basis vector $|n\rangle$.

9.6.3 Continuous Eigenvalues

Let us now consider the eigenvalue equation.

$$\hat{A}|a\rangle = a|a\rangle \quad a: -\infty < a < \infty \tag{9.27}$$

The completeness of eigenvectors $|a\rangle$ implies that any arbitrary state vector $|\psi\rangle$ can be expressed as

$$|\psi\rangle = \int da\, |a\rangle\, C(a)$$

In analogy with the Equation (9.25), we have

$$C(a) = \langle a|\psi\rangle \tag{9.28}$$

$$\therefore\ \ |\psi\rangle = \int da\, |a\rangle\langle a|\psi\rangle = \hat{I}\,|\psi\rangle$$

provided we identify the identity operator *I* as

$$\hat{I} = \int da\, |a\rangle\langle a| \tag{9.29}$$

The normalization condition for the eigenvector $|a\rangle$ can be obtained from

$$\begin{aligned} C(a) &= \langle a|\psi\rangle = \langle a|\hat{I}|\psi\rangle \\ &= \int da'\langle a|a'\rangle\langle a'|\psi\rangle \\ &= \int da'\langle a|a'\rangle C(a') \end{aligned}$$

This suggests

$$\langle a|a'\rangle = \delta(a - a') \tag{9.30}$$

The Equations (9.26) and (9.29) are known as resolution of the identity operator. It is to be noted that the above equations are true provided the set $\{|n\rangle\}$ or the set $\{|a\rangle\}$ forms a complete set.

For a dynamical variable $\hat{A}$ in general, the eigenvalue spectrum consists of the set of both discrete eigenvalues and continuous eigenvalues. In this case, all the eigenvectors, corresponding to discrete eigenvalues and continuous eigenvalues, together form a complete set.

Complete set of eigenvectors of $\hat{A}$ = eigenvectors corresponding to discrete eigenvalues + eigenvectors corresponding to continuous eigenvalues.

$$\therefore\ \ |\psi\rangle = \sum_n C_n\,|n\rangle + \int C(a)da\,|a\rangle \tag{9.31}$$

9.7 POSTULATES OF QUANTUM MECHANICS

Let us reformulate the postulates of quantum mechanics in the language of Dirac’s bra and ket vector formalism.

P-1

The state of a system is described by state vector. $|\psi\rangle$ and all the state vectors are the elements of Hilbert space.

This is consistent with principle of superposition. If $|\psi_1\rangle$ and $|\psi_2\rangle$ represent two distinct states, then Hilbert space ensures that there exists a state $|\psi\rangle = c_1|\psi_1\rangle + c_2|\psi_2\rangle$.

P-2

To every dynamical variable A, an abstract self-adjoint operator $\hat{A}$ is associated with it. The operators $\hat{x}_i$ and $\hat{p}_j$ are chosen in such a way that

$$[\hat{x}_i, \hat{p}_j] = i\hbar\delta_{ij} \quad i, j = 1, 2, 3 \ldots \tag{9.32}$$

This is known as quantum condition.

Dirac suggested a way to make a transition from classical mechanics to quantum mechanics. For classical mechanics, the Poisson bracket between the generalized coordinates and the canonically conjugate momenta is given as

$$[q_i, p_j]_{PB} = \delta_{ij}$$

Dirac's quantum condition is to elevate classical variables q_i and p_j to operators $\hat{q}_i$ and $\hat{p}_j$ and replace the Poison bracket by the commutator whose value is $i\hbar$.

$$[q_i, p_j]_{PB} = \delta_{ij} \rightarrow [\hat{q}_i, \hat{p}_j] = i\hbar\delta_{ij}$$

P-3

When a measurement of a dynamical variable A is made, the only possible experimental results are its eigenvalues.

P-4

The average value of a dynamical variable A for an ensemble in which all the systems are in the state $|\psi\rangle$ is given by

$$\langle A\rangle = \langle\psi|\hat{A}|\psi\rangle \tag{9.33}$$

Let us write $|\psi\rangle$ as

$$|\psi\rangle = \sum C_n|n\rangle$$

C_n is given by $C_n = \langle n|\psi\rangle$. Let us proceed in the same way as done in Chapter 5 to conclude $|C_n|^2$ is the probability obtaining the value of the dynamical variable A to be a_n. Let us write $P_r(n, \psi)$ for $|C_n|^2$. $P_r(n, \psi)$ is the probability of obtaining the eigenvalue a_n on measuring A when the system is in state $|\psi\rangle$ i.e., $P_r(a_n, \psi) = |\langle n|\psi\rangle|^2$. In the case of continuous eigenvalues, we write $|\psi\rangle$ as $|\psi\rangle = \int da\,|a\rangle\langle a|\psi\rangle$. $P_r(a, \psi)da = |\langle a|\psi\rangle|^2\,da$ is now interpreted as the probability of obtaining the value of A to be in the range a and $a + da$ (or in the interval da centred around a).

Interpretation of $\langle \phi \mid \psi \rangle$

We have already seen the effect of measurement on the state of a system in Chapter 5. In the language of abstract-state vector formalism, we say that a measurement of a dynamical variable A causes the system to change its state from $|\psi\rangle$ to the eigenstate $|n\rangle$ with a probability $|C_n|^2$. Therefore, C_n is interpreted as the probability amplitude for a system to make a transition from state $|\psi\rangle$ to state $|n\rangle$ in a measurement process. This interpretation can be generalized to a scalar product $\langle \phi|\psi\rangle$, where the ket $|\phi\rangle$ is an arbitrary state vector and it need not be a basis vector like $|n\rangle$. Note that, in general, $\langle \phi|\psi\rangle$ is a complex number. Now, $\langle \phi|\psi\rangle$ is interpreted as probability amplitude for a system to make a transition from state $|\psi\rangle$ to a state $|\phi\rangle$ in a measurement process.

P-5

The time evolution of the state vector $|\psi(t)\rangle$ can be determined by the Schrödinger equation.

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

where $\hat{H}$ is the Hamiltonian operator.

9.8 REPRESENTATIONS

In the case of geometrical vectors, $\mathbf{V}$ can be represented in many ways. In terms of the basis vectors $\mathbf{e}_1, \mathbf{e}_2$ and $\mathbf{e}_3$, we can write $\mathbf{V}$ as

$$\mathbf{V} = c_1\mathbf{e}_1 + c_2\mathbf{e}_2 + c_3\mathbf{e}_3$$

The numerical values c_1, c_2 and c_3 are the components of $\mathbf{V}$ along the basis vectors $\mathbf{e}_1, \mathbf{e}_2$ and $\mathbf{e}_3$ and specifying the numerical values c_1, c_2 and c_3 amounts to describing the vector $\mathbf{V}$ completely. It is enough to write $\mathbf{V}$ as (c_1, c_2, c_3) or $\begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix}$. The column vector is said to represent $\mathbf{V}$ with respect to the basis $\mathbf{e}_1, \mathbf{e}_2$ and $\mathbf{e}_3$.

It is to be noted that the numerical values of the components of a vector depend on the choice of basis vectors. Instead of $\mathbf{e}_1, \mathbf{e}_2$ and $\mathbf{e}_3$, let us choose some other set $\mathbf{E}_1, \mathbf{E}_2$ and $\mathbf{E}_3$.

The same vector $\mathbf{V}$ can be written as

$$\mathbf{V} = b_1\mathbf{E}_1 + b_2\mathbf{E}_2 + b_3\mathbf{E}_3$$

In this case, the components (b_1, b_2, b_3) or the column vector $\begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$ is said to represent the same vector $\mathbf{V}$ with respect to the basis $\mathbf{E}_1, \mathbf{E}_2$ and $\mathbf{E}_3$.

This idea can be extended to Dirac's bra and ket vectors.

A Representation and *B* Representation

The term *A* representation means here that our choice of basis kets is the eigenkets of the operator *A*. So if we choose the eigenkets of the dynamical variable *B* as the basis kets, then such representation is called *B* representation.

A Representation of State Vector

Let us choose the eigenkets of the dynamical variable *A* to represent the state $|\psi\rangle$.

Let
$$\hat{A}\,|\,i\rangle = a_i\,|\,i\rangle$$

In terms of the basis kets $\{|\,i\rangle\}$ let us express $|\psi\rangle$ as

$$|\psi\rangle = \sum_n c_n\,|\,n\rangle \quad \text{with} \quad c_n = \langle n\,|\,\psi\rangle$$

The state vector $|\psi\rangle$ in *A* representation is written as

$$\psi_A = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \end{bmatrix} = \begin{bmatrix} \langle 1\,|\,\psi\rangle \\ \langle 2\,|\,\psi\rangle \\ \vdots \end{bmatrix} \tag{9.34}$$

A Representation of the Abstract Operator $\hat{F}$

Having represented a state vector $|\psi\rangle$ in *A* basis, we can represent an abstract operator $\hat{F}$ also in *A* basis.

Consider the following equation

$$\hat{F}\,|\,\psi\rangle = |\,\phi\rangle$$

The state vectors $|\psi\rangle$ and $|\phi\rangle$ are given by

$$|\psi\rangle = \sum_n c_n\,|\,n\rangle$$
$$|\phi\rangle = \sum_i d_i\,|\,i\rangle$$

The *A* representation of $|\psi\rangle$ and $|\phi\rangle$ is given by

$$\psi_A = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \end{bmatrix} \quad \text{and} \quad \phi_A = \begin{bmatrix} d_1 \\ d_2 \\ \vdots \end{bmatrix} \tag{9.35}$$

$$\hat{F}\,|\,\psi\rangle = \hat{F}\hat{I}\,|\,\psi\rangle = |\,\phi\rangle$$

$$\sum_j \hat{F}\,|\,j\rangle\langle j\,|\,\psi\rangle = |\,\phi\rangle$$

Taking scalar product with $\langle i|$, we get

$$\sum_j \langle i|\hat{F}|j\rangle\langle j|\psi\rangle = \langle i|\phi\rangle$$

We can write this equation as

$$\sum_j F_{ij}\, c_j = d_i \tag{9.36}$$

provided we identify

$$F_{ij} = \langle i|\hat{F}|j\rangle, \quad c_j = \langle j|\psi\rangle \text{ and } d_i = \langle i|\phi\rangle \tag{9.37}$$

This equation is a matrix equation as given below.

$$\begin{bmatrix} F_{11} & F_{12} & F_{13} & .. & .. \\ F_{21} & F_{22} & F_{23} & .. & .. \\ \vdots & \vdots & \vdots & .. & .. \\ \vdots & \vdots & \vdots & .. & .. \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ \vdots \end{bmatrix} = \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ \vdots \end{bmatrix} \tag{9.38}$$

Therefore, the representation of the abstract $\hat{F}$ is a matrix operator.

9.8.1 Change in Basis – Transformation Matrix

We will now investigate the relationship between two representations corresponding to dynamical variables A and B.

Let us consider two sets of basis vectors $\{|a_i\rangle\}$ and $\{|b_i\rangle\}$. They are eigenkets of $\hat{A}$ and $\hat{B}$, respectively.

$$\hat{A}|a_i\rangle = a_i|a_i\rangle \quad \hat{B}|b_j\rangle = b_j|b_j\rangle \tag{9.39}$$

In A representation, the state vector $|\psi\rangle$ and the abstract operator $\hat{F}$ are represented as follows:

$$\psi_A = \begin{bmatrix} \langle a_1|\psi\rangle \\ \langle a_2|\psi\rangle \\ \vdots \end{bmatrix} \qquad F_A = \begin{bmatrix} \langle a_1|\hat{F}|a_1\rangle & \langle a_1|\hat{F}|a_2\rangle & \cdots & \cdots \\ \langle a_2|\hat{F}|a_2\rangle & \langle a_2|\hat{F}|a_2\rangle & \cdots & \cdots \\ \vdots & \vdots & \cdots & \cdots \end{bmatrix} \tag{9.40}$$

The same state vector $|\psi\rangle$ and the operator in the B representation are given as follows:

$$\psi_B = \begin{bmatrix} \langle b_1|\psi\rangle \\ \langle b_2|\psi\rangle \\ \vdots \end{bmatrix} \qquad F_B = \begin{bmatrix} \langle b_1|\hat{F}|b_1\rangle & \langle b_1|\hat{F}|b_2\rangle & \cdots & \cdots \\ \langle b_2|\hat{F}|b_1\rangle & \langle b_2|\hat{F}|b_2\rangle & \cdots & \cdots \\ \vdots & \vdots & \cdots & \cdots \end{bmatrix} \tag{9.41}$$

Consider $\langle b_i|\psi\rangle$

$$\langle b_i|\psi\rangle = \langle b_i|\hat{I}|\psi\rangle = \sum_j \langle b_i|a_j\rangle\langle a_j|\psi\rangle$$

$$\langle b_i|\psi\rangle = \sum_j U_{ij}\langle a_j|\psi\rangle \tag{9.42}$$

U_{ij} is matrix element of the matrix U, which is known as transformation matrix. It is easy to show that U is a unitary matrix.

$$\begin{aligned}(U^\dagger U)_{ik} &= \sum_j (U^\dagger)_{ij}\, U_{jk} = \sum_j U_{ji}^* U_{jk} = \sum_j \langle b_j|a_i\rangle^* \langle b_j|a_k\rangle \\ &= \sum_j \langle a_i|b_j\rangle\langle b_j|a_k\rangle = \langle a_i|\hat{I}|a_k\rangle \\ &= \langle a_i|a_k\rangle = \delta_{ik}\end{aligned}$$

Therefore, U is a unitary matrix.

Now consider the relationship between F_A and F_B.

$$\begin{aligned}[F_B]_{ij} &= \langle b_i|\hat{F}|b_j\rangle = \langle b_i|\hat{I}\hat{F}\hat{I}|b_j\rangle \\ &= \sum_j\sum_k \langle b_i|a_j\rangle\langle a_i|\hat{F}|a_k\rangle\langle a_k|b_j\rangle \\ &= \sum_j\sum_k U_{ij}[F_A]_{jk}U_{jk}^* = \sum_j\sum_k U_{ij}[F_A]_{jk}(U^\dagger)_{kj}\end{aligned}$$

In terms of matrix notation, the above equation becomes

$$[F]_B = U[F]_A U^\dagger$$

So we have

$$\psi_B = U\psi_A \tag{9.43}$$

$$F_B = UF_A U^\dagger \tag{9.44}$$

9.9 COORDINATE REPRESENTATION

In the coordinate representation, the eigenkets of $\hat{x}$ are chosen as the basis vectors.

The eigenvalue equation for the coordinate x is given by

$$\hat{x}|x\rangle = x|x\rangle \qquad -\infty < x < \infty \tag{9.45}$$

The eigenvalue x is a continuous variable and varies from $-\infty$ to ∞. Therefore, the normalization condition for the coordinate eigenkets is

$$\langle x|x'\rangle = \delta(x - x') \tag{9.46}$$

State Vector $|\alpha\rangle$ in Coordinate Representation

Let us consider a state vector $|\alpha\rangle$. To get the coordinate representation of $|\alpha\rangle$, we have to determine the numerical value of the 'components' $\langle x|\alpha\rangle$ for all values of x. The components corresponding to various values of x, say $x = x_1, x_2, x_3 \ldots$ may be written as column vector as given below:

$$\begin{bmatrix} \langle x_1|\alpha\rangle \\ \langle x_2|\alpha\rangle \\ \langle x_3|\alpha\rangle \\ \vdots \end{bmatrix}$$

Note this vector should be infinite dimensional vector. Since x varies continuously from $-\infty$ to ∞, numerical values of all the components of $|\alpha\rangle$ (corresponding to the basis vectors $|x\rangle$) can be written as $\psi_\alpha(x)$. The component $\psi_\alpha(x)$ is a number for each value of x. i.e., it is a function of x in the well-known sense of calculus.

$$\therefore \psi_\alpha(x) = \langle x|\alpha\rangle \qquad x: -\infty < x < \infty \tag{9.47}$$

$\hat{x}$ in Coordinate Representation

Let $\hat{x}$ act on a ket $|\alpha\rangle$ to produce another ket $|\beta\rangle$. The wave functions corresponding to the kets are given by $\psi_\alpha(x)$ and $\psi_\beta(x)$.

$$\hat{x}|\alpha\rangle = |\beta\rangle \tag{9.48}$$

i.e., $$\psi_\alpha(x) = \langle x|\alpha\rangle \text{ and } \psi_\beta(x) = \langle x|\beta\rangle$$

We know that in coordinate representation,

$$x_{op}\psi_\alpha(x) = \psi_\beta(x) \tag{9.49}$$

What is the operator x_{op}? Or how do we get x_{op} from (9.48)?

Taking inner product of the Equation (9.48) with $\langle x|$, we get

$$\langle x|\hat{x}|\alpha\rangle = \langle x|\beta\rangle \tag{9.50}$$

Since $\langle x|\hat{x} = \langle x|x = x\langle x|$ we write the Equation (9.50) as

$$x\langle x|\alpha\rangle = \langle x|\beta\rangle$$

$$x\psi_\alpha(x) = \psi_\beta(x) \tag{9.51}$$

Comparing the Equations (9.49) and (9.51), we get

$$x_{op}\psi_\alpha(x) = x\psi_\alpha(x) \tag{9.52}$$

Hence, x_{op} is x in the coordinate representation.

$f(\hat{x})$ in Coordinate Representation

Let us now consider $f(\hat{x}) = a_1\hat{x} + a_2\hat{x}^2 + \cdots$ and its coordinate representation $f(x_{op})$.

Now consider
$$\hat{x}^2 |\alpha\rangle = |\beta\rangle \tag{9.53}$$

The corresponding coordinate representation is

$$x_{op}^2 \psi_\alpha(x) = \psi_\beta(x) \tag{9.54}$$

From (9.53), we get
$$\langle x|\hat{x}^2|\alpha\rangle = \langle x|\beta\rangle$$

We know that
$$\langle x|\hat{x}^2 = \langle x|x^2 = x^2\langle x|$$

$$\therefore \langle x|\hat{x}^2|\alpha\rangle = x^2\langle x|\alpha\rangle$$

$$x^2\langle x|\alpha\rangle = \langle x|\beta\rangle \tag{9.55}$$

Comparing (9.54) and (9.55), we get

$$x_{op}^2 \psi_\alpha(x) = x^2\psi(x) \tag{9.56}$$

This can be extended a general function $f(\hat{x})$

$$f(\hat{x})|\alpha\rangle = |\beta\rangle \tag{9.57}$$

The corresponding equation in coordinate representation is

$$f(x_{op})\psi_\alpha(x) = f(x)\cdot\psi_\alpha(x) = \psi_\beta(x) \tag{9.58}$$

9.9.1 Momentum Operator in Coordinate Representation

Let us consider

$$\hat{p}|\alpha\rangle = |\beta\rangle \tag{9.59}$$

In coordinate representation, the corresponding equation is

$$p_{\text{op}}\psi_\alpha(x) = \psi_\beta(x) \tag{9.60}$$

Or
$$p_{\text{op}}\langle x|\alpha\rangle = \langle x|\beta\rangle \tag{9.61}$$

Consider
$$\hat{x}\hat{p} - \hat{p}\hat{x} = i\hbar$$

$$\langle x|\hat{x}\hat{p} - \hat{p}\hat{x}|x'\rangle = \langle x|i\hbar|x'\rangle \tag{9.62}$$

We have
$$\langle x|\hat{x} = \langle x|x = x\langle x|$$

and $\hat{x}\,|\,x'\rangle = x'\,|\,x'\rangle$

Using these results in the LHS of the Equation (9.62), we get

$$\langle x\,|\,\hat{x}\hat{p} - \hat{p}\hat{x}\,|\,x'\rangle = \langle x - x'\rangle\langle x\,|\,\hat{p}\,|\,x'\rangle$$

$$\therefore (x - x')\langle x\,|\,\hat{p}\,|\,x'\rangle = i\hbar\delta(x - x') \tag{9.63}$$

The structure of this equation is similar to the Equation (4.34).

$$f(x)\frac{d}{dx}\delta(x - x') = -\frac{df}{dx}(x)\delta(x - x')$$

Let us choose $f(x)$ as $f(x) = i\hbar(x - x')$

$$\therefore i\hbar(x - x')\frac{d}{dx}\delta(x - x') = -i\hbar\delta(x - x')$$

Changing the variable of differentiation from x to x', we get

$$i\hbar(x - x')\frac{d}{dx'}\delta(x - x') = i\hbar\delta(x - x') \tag{9.64}$$

Comparing this equation with (9.62), we get

$$\langle x\,|\,\hat{p}\,|\,x'\rangle = i\hbar\frac{d}{dx'}\delta(x - x') \tag{9.65}$$

Now consider $\langle x\,|\,\hat{p}\,|\,\alpha\rangle$

$$\begin{aligned}\langle x\,|\,\hat{p}\,|\,\alpha\rangle &= \int dx'\langle x\,|\,\hat{p}\,|\,x'\rangle\langle x'\,|\,\alpha\rangle \\ &= \int dx'\left[i\hbar\frac{d}{dx'}\delta(x - x')\right]\psi_\alpha(x') \\ &= i\hbar\delta(x - x')\psi_\alpha(x')\Big|_{-\infty}^{\infty} - \int\left(i\hbar\frac{d}{dx'}\psi_\alpha(x')\right)\delta(x - x')dx \\ &= -i\hbar\frac{d}{dx}\psi_\alpha(x)\end{aligned} \tag{9.66}$$

So we have

$$\langle x\,|\,\hat{p}\,|\,\alpha\rangle = \langle x\,|\,\beta\rangle \quad \text{and} \quad p_{op}\psi_\alpha(x) = \psi_\beta(x)$$

$$-i\hbar\frac{d}{dx}\psi_\alpha(x) = \psi_\beta\langle x\rangle$$

$$\therefore p_{op} = -i\hbar\frac{d}{dx} \tag{9.67}$$

9.9.2 $\hat{p}|x\rangle$ and $\langle x|\hat{p}$

From the Equation (9.65), we have

$$\langle x|\hat{p}|x'\rangle = -i\hbar\frac{d}{dx}\delta(x-x') \text{ and } \langle x|\hat{p}|x'\rangle = i\hbar\frac{d}{dx'}\delta(x-x')$$

Since $\langle x|x'\rangle = \delta(x-x')$ these equations become

$$\langle x|\hat{p}|x'\rangle = -i\hbar\frac{d}{dx}\langle x|x'\rangle \tag{9.68}$$

$$\langle x|\hat{p}|x'\rangle = i\hbar\frac{d}{dx'}\langle x|x'\rangle \tag{9.69}$$

The Equation (9.68) suggests $\langle x|\hat{p} = -i\hbar\frac{d}{dx}\langle x|$

Note that both $-i\hbar\, d/dx$ as well as $i\hbar\, d/dx'$ are not abstract operators.

$$\therefore i\hbar\frac{d}{dx'}\langle x|x'\rangle = \langle x|\left(i\hbar\frac{d}{dx'}|x'\rangle\right)$$

Making use of this result in (9.69), we get

$$\therefore \qquad \langle x|\hat{p}|x'\rangle = \langle x|\left(i\hbar\frac{d}{dx'}|x'\rangle\right)$$

This equation suggests $\hat{p}|x'\rangle = i\hbar\frac{d}{dx'}|x'\rangle.$

So, changing the variable from x to x' in the above equation, we have

$$\hat{p}|x\rangle = i\hbar\frac{d}{dx}|x\rangle \tag{9.70}$$

and

$$\langle x|\hat{p} = -i\hbar\frac{d}{dx}\langle x|. \tag{9.71}$$

The Equation (9.71) is the Hermitian conjugate of (9.70).

Let us reiterate again that $i\hbar\, d/dx$ is not an abstract operator.

9.9.3 $f(\hat{p})$ in Coordinate Representation

Consider $$\hat{p}^2 |\alpha\rangle = |\beta\rangle$$

$$\langle x|\hat{p}^2|\alpha\rangle = \langle x|\hat{p}\,\underbrace{\hat{p}|\alpha\rangle}_{|\chi\rangle} = \langle x|\hat{p}|\chi\rangle$$

$$= -i\hbar\frac{d}{dx}\langle x|\chi\rangle = \left(-i\hbar\frac{d}{dx}\right)\langle x|\hat{p}|\alpha\rangle \qquad \text{(Using 9.71)}$$

$$= \left(-i\hbar\frac{d}{dx}\right)\left(-i\hbar\frac{d}{dx}\right)\langle x|\alpha\rangle \qquad \text{(Using 9.71)}$$

$$= \left(-i\hbar\frac{d}{dx}\right)^2\langle x|\alpha\rangle.$$

$$\therefore \langle x|\hat{p}^2|\alpha\rangle = \left(-i\hbar\frac{d}{dx}\right)^2\langle x|\alpha\rangle = p_{op}^2\psi_\alpha(x) \qquad (9.72)$$

This reasoning can be extended to any function $f(\hat{p})$.

$$f(\hat{p})|\alpha\rangle = |\beta\rangle \qquad (9.73)$$

$$\langle x|f(\hat{p})|\alpha\rangle = \langle x|\beta\rangle \qquad (9.74)$$

In coordinate representation, this equation becomes

$$f\left(-i\hbar\frac{d}{dx}\right)\langle x|\alpha\rangle = \langle x|\beta\rangle \text{ or } f\left(-i\hbar\frac{d}{dx}\right)\psi_\alpha(x) = \psi_\beta(x)$$

$$f(\hat{p})|\alpha\rangle = |\beta\rangle \;\Rightarrow\; f\left(-i\hbar\frac{d}{dx}\right)\psi_\alpha(x) = \psi_\beta(x) \qquad (9.75)$$

$\hat{A}(\hat{x}, \hat{p})$

Let $A(x, p)$ be the classical dynamical variable. In quantum mechanics, it becomes $\hat{A}(\hat{x}, \hat{p})$.
Now, we have

$$\hat{A}(\hat{x}, \hat{p})|\alpha\rangle = |\beta\rangle$$

$$\langle x|\hat{A}(\hat{x}, \hat{p})|\alpha\rangle = \langle x|\beta\rangle$$

In coordinate representation, this equation becomes

$$A_{op}\left(x, -i\hbar\frac{d}{dx}\right)\psi_\alpha(x) = \psi_\beta(x)$$

In three dimensions, we have

$$\hat{A}(\hat{\mathbf{r}}, \hat{\mathbf{p}})\,|\,\alpha\rangle = |\,\beta\rangle \tag{9.76}$$

$$\langle\mathbf{r}\,|\,\hat{A}(\mathbf{r},\mathbf{p})\,|\,\alpha\rangle = \langle\mathbf{r}\,|\,\beta\rangle \tag{9.77}$$

or
$$\hat{A}(\hat{\mathbf{r}},\hat{\mathbf{p}})\,|\,\alpha\rangle = |\,\beta\rangle \quad \Rightarrow \quad A_{op}(\mathbf{r},-i\hbar\nabla)\psi_\alpha(\mathbf{r}) = \psi_\beta(\mathbf{r}) \tag{9.78}$$

9.10 MOMENTUM REPRESENTATION

In momentum representation, the eigenkets of momentum operator $\hat{p}$ are the basis vector.

$$\hat{p}\,|\,p\rangle = p\,|\,p\rangle \qquad -\infty < p < \infty \tag{9.79}$$

The normalization condition for the eigenkets is

$$\langle p\,|\,p'\rangle = \delta(p-p') \tag{9.80}$$

Let us now try to find x_{op} in momentum representation.

$$\begin{gathered}\hat{x}\,|\,\alpha\rangle = |\,\beta\rangle \\ \langle p\,|\,\hat{x}\,|\,\alpha\rangle = \langle p\,|\,\beta\rangle\end{gathered} \tag{9.81}$$

In momentum representation, we have

$$x_{op}\phi_\alpha(p) = \phi_\beta(p) \tag{9.82}$$

Let us consider $(\hat{x}\hat{p} - \hat{p}\hat{x}) = i\hbar$

$$\langle p\,|\,x\hat{p} - \hat{p}\hat{x}\,|\,p'\rangle = i\hbar\langle p|\,p'\rangle = i\hbar\delta(p-p')$$

We know that

$$\langle p\,|\,\hat{p} = \langle p\,|\,p \text{ and } \hat{p}\,|\,p'\rangle = p'\,|\,p\rangle$$

$$\begin{gathered}\therefore \langle p\,|\,\hat{x}\hat{p} - \hat{p}\hat{x}\,|\,p'\rangle = -(p-p')\langle p\,|\,\hat{x}\,|\,p'\rangle \\ \therefore (p-p')\langle p\,|\,\hat{x}\,|\,p'\rangle = i\hbar\delta(p-p')\end{gathered}$$

Based on the same reasoning as done in the case of determination of the operator p_{op} in coordinate representation, we get

$$\langle p\,|\,\hat{x}\,|\,p'\rangle = -i\hbar\frac{d}{dp'}\delta(p-p') \tag{9.83}$$

Let us now consider $\langle p|\hat{x}|\alpha\rangle$

$$\begin{aligned}\langle p|\hat{x}|\alpha\rangle &= \int dp'\langle p|\hat{x}|p'\rangle\langle p'|\alpha\rangle\\ &= \int dp'\left(-i\hbar\frac{d}{dp'}\delta(p-p')\right)\phi_\alpha(p')\\ &= -i\hbar\delta(p-p')\phi_\alpha(p')\big|_{-\infty}^{\infty} + \int dp'\delta(p-p')i\hbar\frac{d}{dp'}\phi_\alpha(p')\\ &= i\hbar\frac{d}{dp}\phi_\alpha(p)\end{aligned}$$

$$\langle p|\hat{x}|\alpha\rangle = \langle p|\beta\rangle, \qquad i\hbar\frac{d}{dp}\phi_\alpha(p) = \phi_\beta(p)$$

$$\therefore x_{op} = i\hbar\frac{d}{dp} \tag{9.84}$$

Just as we worked in the case of coordinate representation, we can work out A_{op} in the momentum representation.

$$\hat{A}(\hat{x},\hat{p})|\alpha\rangle = |\beta\rangle \tag{9.85}$$

$$A_{op}\left(x_{op} = i\hbar\frac{d}{dp}, p\right)\phi_\alpha(p) = \phi_\beta(p) \tag{9.86}$$

9.11 UNITARY TRANSFORMATIONS

Let us consider a general unitary transformation transforming all the kets of the Hilbert space as given below.

$$|\alpha'\rangle = \hat{U}|\alpha\rangle \text{ and } |\beta'\rangle = \hat{U}|\beta\rangle \tag{9.87}$$

Under the unitary transformation, $\hat{F}$ transforms to $\hat{F}'$. Let us assume $\hat{F}$ transforms $|\alpha\rangle$ into $|\beta\rangle$ and $\hat{F}'$ transforms $|\alpha'\rangle$ into $|\beta'\rangle$.

i.e.,

$$\hat{F}|\alpha\rangle = |\beta\rangle \quad\text{and}\quad \hat{F}'|\alpha'\rangle = |\beta'\rangle$$

$$\hat{F}'\hat{U}|\alpha\rangle = \hat{U}|\beta\rangle \qquad \text{(using(9.87))}$$

Multiplying both sides by $U^\dagger$, we get

$$\underbrace{\hat{U}^\dagger\hat{F}'\hat{U}}_{\hat{F}}|\alpha\rangle = |\beta\rangle = \hat{F}|\alpha\rangle$$

$$\therefore \hat{U}^\dagger\hat{F}'\hat{U} = \hat{F} \text{ or } \hat{F}' = \hat{U}\hat{F}\hat{U}^\dagger \tag{9.88}$$

The matrix element $\langle \beta | \hat{F} | \alpha \rangle$ is invariant under the transformation $|\alpha\rangle \rightarrow |\alpha'\rangle = \hat{U} |\alpha\rangle$ and $\hat{F} \rightarrow \hat{U}\hat{F}\hat{U}^{\dagger}$. This can be easily seen as follows:

$$\begin{aligned} \langle \beta' | \hat{F}' | \alpha' \rangle &= \langle \beta | \hat{U}^{\dagger} U \hat{F} U^{\dagger} U | \alpha \rangle \\ &= \langle \beta | \hat{F} | \alpha \rangle \end{aligned} \tag{9.89}$$

The norm of a vector $|\alpha\rangle$ is invariant under the unitary transformation $|\alpha\rangle \rightarrow |\alpha'\rangle = \hat{U} |\alpha\rangle$.

$$\langle \alpha' | \alpha' \rangle = \langle \alpha | \hat{U}^{\dagger} \hat{U} | \alpha \rangle = \langle \alpha | \alpha \rangle \tag{9.90}$$

The invariance of the norm of the state vectors $|\phi\rangle$ $|\psi\rangle$ and the invariance of the matrix element $\langle \phi | \hat{F} | \psi \rangle$ (The kets $|\phi\rangle$ and $|\psi\rangle$ are arbitrary state vectors) under unitary transformation gives immense significance to unitary transformation in quantum mechanics.

9.12 TIME EVOLUTION OPERATOR

The Schrödinger equation for the state vector $|\alpha(t)\rangle$ is

$$i\hbar \frac{d}{dt} |\alpha(t)\rangle = \hat{H} |\alpha(t)\rangle \tag{9.91}$$

Time solution operator $T(t, t_0)$ is defined as an operator which takes the state vector $|\alpha(t_0)\rangle$ to $|\alpha(t)\rangle$.

$$\therefore |\alpha(t)\rangle = \hat{T}(t,t_0) |\alpha(t_0)\rangle \tag{9.92}$$

$$i\hbar \frac{d}{dt} |\alpha(t)\rangle = i\hbar \frac{d}{dt} \hat{T}(t,t_0) |\alpha(t_0)\rangle$$

$$\begin{aligned} \therefore i\hbar \frac{d}{dt} \hat{T}(t,t_0) |\alpha(t_0)\rangle &= \hat{H} |\alpha(t)\rangle \\ &= \hat{H}\hat{T}(t,t_0) |\alpha(t_0)\rangle \end{aligned}$$

$$\therefore i\hbar \frac{d}{dt} \hat{T}(t,t_0) = \hat{H}\hat{T}(t,t_0) \tag{9.93}$$

Note that the solution to (9.91) can be formally written as

$$|\alpha(t)\rangle = \exp\left[\frac{i\hat{H}(t-t_0)}{\hbar}\right] |\alpha(t_0)\rangle \tag{9.94}$$

Obviously, the time evolution operator $\hat{T}(t,t_0)$ is given by

$$\hat{T}(t,t_0) = e^{-i\hat{H}(t-t_0)/\hbar} \tag{9.95}$$

Since $\hat{H}$ is a self-adjoint operator ($\hat{H} = \hat{H}^\dagger$), the time evolution operator $\hat{T}(t, t_0)$ is a unitary operator which transforms the state $|\alpha(t_0)\rangle$ into the state $|\alpha(t)\rangle$.

9.12.1 Pictures

The invariance of the matrix elements $\langle\beta|\hat{F}|\alpha\rangle$ under unitary transformations $|\alpha\rangle \to \hat{U}|\alpha\rangle$, $|\beta\rangle \to \hat{U}|\beta\rangle$ and $\hat{F} \to \hat{U}\hat{F}\hat{U}^\dagger$ gives us a tool to deal with the time dependence of the state vectors and dynamical variables in a convenient way. This idea leads to the concept of pictures. There are three different pictures: Schrödinger picture, Heisenberg picture and interaction picture. There exists unitary transformation which relates these pictures.

Schrödinger Picture

The state vectors and the dynamical variable are denoted by $|\alpha(t)\rangle_s$ and $\hat{F}_s$. The subscript S refers to Schrödinger picture. The operators $\hat{F}_s$ may or may not have time dependence. Generally, the operators are time independent. For instance, the position operator $\hat{\mathbf{r}}$, the momentum operator $\hat{\mathbf{p}}$ and the angular momentum operators $\hat{L} = \hat{\mathbf{r}} \times \hat{\mathbf{p}}$ are time-independent operators. Let us assume that the Hamiltonian $\hat{H}$ is time independent. The time evolution of the state $|\alpha(t)\rangle_s$ is determined by

$$i\hbar\frac{d}{dt}|\alpha_s(t)\rangle = \hat{H}_s\,|\alpha_s(t)\rangle \tag{9.96}$$

$$|\alpha_s(t)\rangle = e^{-\frac{i\hat{H}_s t}{\hbar}}\,|\alpha_s(0)\rangle \tag{9.97}$$

Heisenberg Picture

Let us consider the matrix element $\langle\beta_s(t)|\hat{F}_s\,|\alpha_s(t)\rangle$, where $|\alpha_s(t)\rangle$ and $|\beta_s(t)\rangle$ are arbitrary state vectors. (The subscripts refer to Schrödinger picture.)

$$\langle\beta_s(t)|\hat{F}_s\,|\alpha_s(t)\rangle = \langle\beta_s(0)|\,e^{i\hat{H}_s t/\hbar}\hat{F}_s\,e^{-i\hat{H}_s t/\hbar}\,|\alpha_s(0)\rangle$$

This suggests a unitary transformation $\hat{U} = e^{i\hat{H}_s t/\hbar}$ which transforms all the states $|\alpha_s(t)\rangle$ to $|\alpha_H(t)\rangle$ which is independent of time. Here the subscript H refers to Heisenberg picture. The unitary transformation taking the state vectors and the operators from Schrodinger picture to Heisenberg picture is defined by

$$|\alpha_s(t)\rangle \to |\alpha_H(t)\rangle = e^{i\hat{H}_s t/\hbar}\,|\alpha_s(t)\rangle \tag{9.98}$$

and
$$\hat{F}_s \to \hat{F}_H = e^{i\hat{H}_s t/\hbar}\hat{F}_s e^{-i\hat{H}_s t/\hbar} \tag{9.99}$$

Let us first consider the time evolution of the state vector $|\alpha_H(t)\rangle$. Note that from (9.98), we have $|\alpha_H(0)\rangle = |\alpha_s(0)\rangle$. Making use of (9.97), we get

$$\begin{aligned}|\alpha_H(t)\rangle &= e^{i\hat{H}_s t/\hbar} e^{-i\hat{H}_s t/\hbar} |\alpha_s(0)\rangle \\ &= |\alpha_s(0)\rangle \end{aligned} \tag{9.100}$$

So, in Heisenberg picture, all the state vectors $|\alpha_H(t)\rangle$ are independent of time.

Let us now consider the time evolution of the operator $\hat{F}_H$ corresponding to a dynamical variable.

$$\begin{aligned}\frac{d\hat{F}_H}{dt} &= \frac{d}{dt}\left[e^{i\hat{H}_s t/\hbar}\hat{F}_S e^{-i\hat{H}_s t/\hbar}\right] \\ &= \left(\frac{d}{dt}e^{i\hat{H}_s t/\hbar}\right)\hat{F}_s e^{-i\hat{H}_s t/\hbar} + e^{-i\hat{H}_s t/\hbar}\hat{F}_s \frac{d}{dt}(e^{i\hat{H}_s t/\hbar}) + e^{i\hat{H}_s t/\hbar}\frac{\partial \hat{F}_s}{\partial t}e^{-i\hat{H}_s t/\hbar} \\ &= \frac{1}{i\hbar}\left[e^{i\hat{H}_s t/\hbar}\hat{F}_s\hat{H}_s e^{-i\hat{H}_s t/\hbar} - e^{-i\hat{H}_s t/\hbar}\hat{H}_s\hat{F}_s e^{-i\hat{H}_s t/\hbar}\right] + e^{i\hat{H}_s t/\hbar}\frac{\partial \hat{F}_s}{\partial t}e^{-i\hat{H}_s t/\hbar} \\ &= \frac{1}{i\hbar}\left[e^{i\hat{H}_s t/\hbar}\hat{F}_s e^{-i\hat{H}_s t/\hbar}e^{i\hat{H}_s t/\hbar}\hat{H}_s e^{-i\hat{H}_s t/\hbar} - e^{i\hat{H}_s t/\hbar}\hat{H}_s e^{-i\hat{H}_s t/\hbar}e^{i\hat{H}_s t/\hbar}\hat{F}_s e^{-i\hat{H}_s t/\hbar}\right] \\ &\quad + e^{i\hat{H}_s t/\hbar}\frac{\partial \hat{F}_s}{\partial t}e^{-i\hat{H}_s t/\hbar} \\ &= \frac{1}{i\hbar}[\hat{F}_H(t)\hat{H}_H(t) - \hat{H}_H(t)\hat{F}_H(t)] + \left(\frac{\partial F_s}{\partial t}\right)_H \\ &= \frac{1}{i\hbar}[\hat{F}_H, \hat{H}_H] + \left(\frac{\partial \hat{F}_s}{\partial t}\right)_H \end{aligned} \tag{9.101}$$

It is very important to realize the significance of the time evolution of the dynamical variable $\hat{F}_H$. There is a close similarity between the equation of motion of a dynamical variable in classical mechanics and the Equation (9.101). In classical mechanics, the time evolution of a dynamical variable is determined by Poisson bracket. For the sake of clarity, let us list the equation of motion in classical mechanics and quantum mechanics.

$$\frac{dF}{dt} = [F, H]_{PB} + \left(\frac{\partial F}{\partial t}\right)_{PB} \qquad \text{Classical Mechanics}$$

$$\frac{dF_H}{dt} = \frac{1}{i\hbar}[F_H, H_H] + \left(\frac{\partial F_s}{\partial t}\right)_H \qquad \text{Quantum Mechanics}$$

So, if we do quantum mechanics in Heisenberg picture, the transition from classical mechanics to quantum mechanics is very simple and straightforward. Replace the Poisson bracket between the dynamical variable F and the classical Hamiltonian by $1/i\hbar$ times the commutation between the operator $\hat{F}_H$ and the Hamiltonian operator $\hat{H}_H$.

i.e.,
$$[F,H]_{PB} \rightarrow \frac{1}{i\hbar}[F_H, H_H].$$

Of course, this prescription requires more careful consideration (see Merzbacher).

It is important to recognize the difference between Schrödinger picture and Heisenberg picture.

In the Schrödinger picture, the state vector $|\alpha(t)\rangle$ evolves in time, whereas the dynamical variables, in general, have no time dependence. In Heisenberg picture, the state vector does not evolve in time, and the time dependence has been completely shifted to the operator corresponding to a dynamical variable. For instance, the position operator $\hat{x}_s$ and the momentum operator $\hat{p}_s$ do not have time dependence. In Heisenberg picture, the position operator $\hat{x}_H(t)$ and the momentum operator $\hat{p}_H(t)$ are time dependent.

Example 9.3 Show that $[\hat{x}_H, \hat{p}_H] = i\hbar$ where $\hat{x}_H$ and $\hat{p}_H$ are position and momentum operators in Heisenberg picture.

Solution:

$$\hat{x}_H(t) = \hat{U}\hat{x}_s\hat{U}^\dagger \qquad \hat{p}_H(t) = \hat{U}\hat{p}_s\hat{U}^\dagger$$

where $\hat{x}_s$ and $\hat{p}_s$ are position and momentum operators in Schrödinger picture.

$$\hat{x}_s\hat{p}_s - \hat{p}_s\,\hat{x}_s = i\hbar$$

Multiply this equation from the left by $\hat{U}$ and the right by $\hat{U}^\dagger$.

$$\hat{U}\hat{x}_s\hat{p}_s\hat{U}^\dagger - \hat{U}\hat{p}_s\hat{x}_s\hat{U}^\dagger = \hat{U}i\hbar\hat{U}^\dagger$$

$$\underbrace{\hat{U}\hat{x}_s\hat{U}^+}_{\hat{x}_H}\,\underbrace{\hat{U}\hat{p}_s\hat{U}^\dagger}_{\hat{p}_H} - \underbrace{\hat{U}\hat{p}_s\hat{U}^\dagger}_{\hat{x}_H}\,\underbrace{\hat{U}\hat{x}_s\hat{U}^\dagger}_{\hat{p}_H} = i\hbar$$

$$\therefore [\hat{x}_H, \hat{p}_H] = i\hbar$$

Example 9.4 Show that the angular momentum operator $\mathbf{L}_z$ in Heisenberg's picture is given by $\hat{L}_{z,H} = \hat{x}_H\hat{p}_{yH} - \hat{y}_H\hat{p}_{xH}$.

Solution: In Schrödinger picture, $L_{z,s}$ is given by

$$L_{zs} = \hat{x}_s\hat{p}_{ys} - \hat{y}_s\hat{p}_{xs}$$

$$\begin{aligned}\hat{U}\hat{L}_{zs}\hat{U}^\dagger = \hat{L}_{zH} &= \hat{U}\hat{x}_s\hat{p}_{ys}\hat{U}^\dagger - \hat{U}y_s\hat{p}_{xs}\hat{U}^\dagger \\ &= \hat{U}\hat{x}_s\hat{U}\hat{U}^\dagger\hat{p}_{ys}\hat{U}^\dagger - \hat{U}\hat{y}_s\hat{U}\hat{U}^\dagger\hat{p}_{xs}\hat{U}^\dagger \\ &= \hat{x}_H\hat{p}_{yH} - \hat{y}_H\hat{p}_{xH}\end{aligned}$$

Example 9.5 The Hamiltonian of a particle is given by $\hat{H} = \dfrac{\hat{p}^2}{2m} + V(\hat{x})$.

Show that
$$\hat{x}_H = \frac{\hat{p}_H}{m} \text{ and } \frac{d\hat{p}_H}{dt} = -\frac{\partial V(\hat{x}_H)}{\partial \hat{x}_H}.$$

Solution: The Hamiltonian in the Heisenberg picture is given by

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

$$\frac{d\hat{x}_H}{dt} = \frac{1}{i\hbar}\left[\hat{x}_H, \frac{\hat{p}_H^2}{2m} + V(\hat{x}_H)\right]$$

$$= \frac{\hat{p}_H}{m}$$

$$\frac{d\hat{p}_H}{dt} = \frac{1}{i\hbar}\left[\hat{p}_H, \frac{\hat{p}_H^2}{2m} + V(\hat{x}_H)\right]$$

$$= \frac{1}{i\hbar}[\hat{p}_H, V(\hat{x}_H)] = -\frac{\partial V(\hat{x}_H)}{\partial \hat{x}_H}$$

Example 9.6 Obtain $\hat{x}_H(t)$ and $\hat{p}_H(t)$ for a free particle.

Solution: The Hamiltonian for a free particle is

$$\hat{H}_H = \frac{\hat{p}_H^2}{2m}$$

$$\frac{d\hat{x}_H}{dt} = \frac{1}{i\hbar}\left[\hat{x}_H, \frac{\hat{p}_H^2}{2m}\right] = \frac{\hat{p}_H}{m}$$

$$\frac{d\hat{p}_H}{dt} = \frac{1}{i\hbar}\left[\hat{p}_H, \frac{\hat{p}_H^2}{2m}\right] = 0$$

$$\therefore \frac{d^2\hat{x}_H}{dt^2} = \frac{1}{m}\frac{d\hat{p}_H}{dt} = 0$$

$$\therefore \hat{x}_H(t) = \hat{A} + t\hat{B}$$

At $t = 0, \hat{x}_H(0) = \hat{x}_s = \hat{A} \quad \therefore \hat{x}_H(t) = \hat{x}_H(0) + t\hat{B}$

$$\frac{d\hat{x}_H}{dt} = \hat{B} = \frac{\hat{p}_H}{m} \qquad \therefore \hat{B} = \frac{\hat{p}_H(0)}{m}$$

$$\therefore \hat{x}_H(t) = \hat{x}_H(0) + \frac{t\hat{p}_H(0)}{m}$$

Example 9.7 Obtain $\hat{x}_H(t)$ and $\hat{p}_H(t)$ for a simple harmonic oscillator.

Solution: The Hamiltonian for a simple harmonic oscillator in Schrödinger picture is

$$\hat{H}_s = \frac{\hat{p}_s^2}{2m} + \frac{1}{2}m\omega^2\hat{x}_S^2$$

In Heisenberg picture, it is given by

$$\hat{H}_H = \frac{\hat{p}_H^2}{2m} + \frac{1}{2}m\omega^2\hat{x}_H^2$$

$$\frac{d\hat{x}_H}{dt} = \frac{1}{i\hbar}\left[\hat{x}_H, \frac{\hat{p}_H^2}{2m} + \frac{1}{2}m\omega^2\hat{x}_H^2\right] = \frac{\hat{p}_H}{m}$$

$$\frac{d\hat{p}_H}{dt} = \frac{1}{i\hbar}\left[\hat{p}_H, \frac{\hat{p}_H^2}{2m} + \frac{1}{2}m\omega^2\hat{x}_H^2\right] = -m\omega^2\hat{x}_H$$

$$\frac{d^2\hat{x}_H}{dt} = \frac{1}{m}\frac{d\hat{p}_H}{dt} = -\omega^2\hat{x}_H$$

$$\frac{d^2\hat{p}_H}{dt^2} = -m\omega^2\frac{d\hat{x}_H}{dt} = -\omega^2\hat{p}_H$$

The solution to these equations is

$$\hat{x}_H(t) = \hat{A}\sin\omega t + \hat{B}\cos\omega t$$

$$\hat{p}_H(t) = m\frac{d\hat{x}_H}{dt} = m\omega(\hat{A}\cos\omega t - \hat{B}\sin\omega t)$$

At $t = 0$

$$\hat{x}_H(0) = \hat{B}$$

$$\hat{p}_H(0) = m\omega\hat{A}$$

$$\therefore \hat{x}_H(t) = \frac{1}{m\omega}\hat{p}_H(0)\sin\omega t + \hat{x}_H(0)\cos\omega t$$

$$\hat{p}_H(t) = \hat{p}_H(0)\cos\omega t - m\omega\,\hat{x}_H(0)\sin\omega t$$

Note that in classical mechanics, the position $q(t)$ and the momentum $p(t)$ are given by

$$q(t) = \frac{1}{m\omega}p_0\sin\omega t + q_0\cos\omega t$$

$$p(t) = p_0\cos\omega t - m\omega\,x_0\sin\omega t$$

where x_0 and p_0 are the initial position and the momentum of the particle.

Interaction Picture

In the Schrödinger picture, the state vectors are time dependent and the dynamical variables and time independent. This is useful in doing non-relativistic quantum mechanics. In the Heisenberg picture, the time dependence is completely shifted to the dynamical variables. So the operator corresponding to a dynamical variable depends on both space and time coordinates. This fact is very useful in relativistic quantum field theory, where the space coordinates and the time coordinates have to be treated an equal footing. Therefore, Heisenberg picture is very useful in doing relativistic quantum field theory.

We have another picture called interaction picture which is useful to do perturbation theory. In this picture, both the dynamical variable and the state vectors are time dependent. But the time dependence is in such a way that the perturbation theory can be effectively done.

Let us write the Hamiltonian $\hat{H}$ as

$$\hat{H} = \hat{H}_0 + \hat{H}' \tag{9.102}$$

where $\hat{H}_0$ has no time dependence and $\hat{H}'$ is the perturbation Hamiltonian which may be or may not be time dependent.

Let us denote the state vector by $|\alpha_I(t)\rangle$ and the dynamical variable by $\hat{F}_I(t)$ in the interaction picture.

$|\alpha_I(t)\rangle$ is defined as

$$|\alpha(t)\rangle_I = \exp\left[i\frac{\hat{H}_{os}t}{\hbar}\right]|\alpha(t)\rangle_s \tag{9.103}$$

$\hat{F}_I(t)$ is defined as

$$\hat{F}_I(t) = e^{i\hat{H}_{os}t/\hbar}\hat{F}_s e^{-i\hat{H}_{os}t/\hbar} \tag{9.104}$$

The state vector $|\alpha_s(t)\rangle$ satisfies the Schrödinger equation

$$i\hbar\frac{d}{dt}|\alpha(t)\rangle_s = (\hat{H}_{os} + \hat{H}'_s)|\alpha(t)\rangle_s \tag{9.105}$$

From (9.103), we get

$$|\alpha_s(t)\rangle = e^{-i\hat{H}_{os}t/\hbar}\,|\alpha_I(t)\rangle \tag{9.106}$$

Making use of the Equations (9.106) and (9.105), we get

$$i\hbar\frac{d}{dt}\left[e^{-i\hat{H}_{os}t/\hbar}\,|\alpha_I(t)\rangle\right] = (\hat{H}_{os} + \hat{H}'_s)|\alpha_I(t)\rangle$$

$$\hat{H}_{os}\underbrace{e^{-i\hat{H}_{os}t/h}\,|\alpha(t)\rangle_I}_{|\alpha(t)\rangle_s} + e^{-i\hat{H}_{os}t/\hbar}i\hbar\frac{d}{dt}|\alpha(t)\rangle_I = \hat{H}_{os}\,|\alpha(t)\rangle_s + \hat{H}'_s\,|\alpha(t)\rangle_s$$

$$\hat{H}_{os}\,|\alpha(t)\rangle_s + e^{-i\hat{H}_{os}t/\hbar}i\hbar\frac{d}{dt}|\alpha(t)\rangle_I = \hat{H}_{os}\,|\alpha(t)\rangle_s + \hat{H}'_s\,|\alpha(t)\rangle_s$$

$$\therefore e^{-i\hat{H}_{os}t/h}i\hbar\frac{d}{dt}|\alpha_I(t)\rangle = \hat{H}'_s\,|\alpha_s(t)\rangle$$

$$i\hbar\frac{d}{dt}|\alpha_I(t)\rangle = e^{i\hat{H}_{os}t/\hbar}\hat{H}'_s e^{-i\hat{H}_{os}t/\hbar}\,|\alpha_I(t)\rangle$$

$$= \hat{H}'_I \,|\alpha(t)\rangle_I \tag{9.107}$$

where
$$\hat{H}'_I = e^{i\hat{H}_{os}t/\hbar}\hat{H}'_s e^{-i\hat{H}_{os}t/\hbar}$$

$$\frac{d}{dt}\hat{F}_I(t) = \frac{1}{i\hbar}\begin{bmatrix} e^{i\hat{H}_{os}t/\hbar}\hat{H}_{os}\hat{F}_s e^{-i\hat{H}_{os}t/\hbar} \\ +e^{i\hat{H}_{os}t/\hbar}\hat{F}_s\hat{H}_{os}e^{-i\hat{H}_{os}t/\hbar} \end{bmatrix} + e^{i\hat{H}_{os}t/\hbar}\frac{\partial \hat{H}'_s}{\partial t}e^{-i\hat{H}_{os}t/\hbar}$$

$$= \frac{1}{i\hbar}\begin{bmatrix} e^{i\hat{H}_{os}t/\hbar}\hat{F}_s e^{-i\hat{H}_{os}t/\hbar}e^{i\hat{H}_{os}t/\hbar}\hat{H}_{os}e^{-i\hat{H}_{os}t/\hbar} \\ -e^{i\hat{H}_{os}t/\hbar}\hat{H}_{os}e^{-i\hat{H}_{os}t/\hbar}e^{i\hat{H}_{os}t/\hbar}\hat{F}_s e^{-i\hat{H}_{os}t/\hbar} \end{bmatrix} + e^{i\hat{H}_{os}t/\hbar}\frac{\partial}{\partial t}\hat{F}_s e^{-i\hat{H}_{os}t/\hbar}$$

$$= \frac{1}{i\hbar}[\hat{F}_I\hat{H}_{OI} - \hat{H}_{OI}\hat{F}_I] + \left(\frac{\partial \hat{F}_s}{\partial t}\right)_I$$

So we have

$$i\hbar\frac{d}{dt}|\alpha_I(t)\rangle = H'_I\,|\alpha_I(t)\rangle \tag{9.108}$$

$$\frac{d}{dt}\hat{F}_I(t) = \frac{1}{i\hbar}[\hat{F}_I, \hat{H}_{OI}] + \left(\frac{\partial \hat{F}_s}{\partial t}\right)_I \tag{9.109}$$

The time evolution of the state is determined by the interaction Hamiltonian, whereas the time evolution of the dynamical variable $\hat{F}_I(t)$ is determined by 'free' Hamiltonian.

Example 9.8 Determine $\hat{x}_I(t), \hat{p}_I(t)$ and $H'_I(t)$ for a simple harmonic oscillator in interaction picture.

Solution: The Hamiltonian for harmonic oscillator is

$$\hat{H}_s = \frac{\hat{p}_s^2}{2m} + \frac{1}{2}m\omega^2\hat{x}_s^2$$

The free Hamiltonian $\hat{H}_{os}$ and the interaction Hamiltonian $\hat{H}'_s$ are

$$\hat{H}_{os} = \frac{\hat{p}_s^2}{2m} \qquad \hat{H}'_s = \frac{1}{2}m\omega^2\hat{x}_s^2$$

In interaction picture, we have

$$\hat{H}_{oI} = \frac{\hat{p}_I^2}{2m} \qquad \hat{H}'_I = \frac{1}{2}m\omega^2\hat{x}_I^2(t)$$

$$\frac{d}{dt}\hat{x}_I(t) = \frac{1}{i\hbar}\left[\hat{x}_I, \frac{\hat{p}_I^2}{2m}\right] = \frac{\hat{p}_I}{m}$$

$$\frac{d\hat{p}_I(t)}{dt} = \frac{1}{i\hbar}\left[\hat{p}_I, \frac{\hat{p}_I^2}{2m}\right] = 0$$

$$\therefore \frac{d^2\hat{x}_I}{dt^2} = 0 \quad \Rightarrow \quad \hat{x}_I = \hat{A} + t\hat{B}$$

$$\hat{p}_I = m\frac{d\hat{x}_I}{dt} = m\hat{B}$$

$$\hat{x}_I(0) = \hat{A}$$

$$\therefore \hat{x}_I(t) = \hat{x}_I(0) + t\frac{\hat{p}_I(0)}{m}$$

$$\hat{p}_I(t) = \hat{p}_I(0)$$

$$H'_I = \frac{1}{2}m\omega^2[x_I(t)]^2$$

$$= \frac{1}{2}m\omega^2\left[\hat{x}_I(0) + t\frac{\hat{p}_I(0)}{m}\right]\left[\hat{x}_I(0) + t\frac{\hat{p}_I(0)}{m}\right]$$

$$= \frac{1}{2}m\omega^2\left\{(\hat{x}_I(0))^2 + \frac{t}{m}[\hat{x}_I(0)\hat{p}_I(0) + \hat{p}_I(0)\hat{x}_I(0)] + \frac{t^2}{m^2}(\hat{p}_I(0))^2\right\}$$

9.13 SPACE–TIME SYMMETRIES

There exists another type of transformation that arises due to space–time symmetry. Here, the changes in the system are with respect to space–time background. These changes are like shifting the system (translating the system) from one place to another or rotating the system as a whole from one orientation to another orientation and so on. The experimental outcomes are not expected to vary due to such changes imposed on the system. We have assumed that the space–time background, in which the apparatus exists, to be isotropic and homogenous. It means that the environment of the system is the same. For instance, translational symmetry means shifting the apparatus from one workbench to another workbench.

Let us see how this idea is implemented in quantum mechanics. In an experiment, the system S makes a transition from state $|\alpha\rangle$ to a state $|\beta\rangle$. Let us now say, translate the system or rotate the system. Let us now label the system as S'. Let $|\alpha'\rangle$ and $|\beta'\rangle$ are the states of S' corresponding to states $|\alpha\rangle$and $|\beta\rangle$ of S. The space–time symmetry implies $|\langle\beta|\alpha\rangle|^2 = |\langle\beta'|\alpha'\rangle|^2$. Let us assume that there exists a transformation $\hat{T}$ which transforms $|\alpha\rangle$ into $|\alpha'\rangle$ and $|\beta\rangle$ into $|\beta'\rangle$.

i.e.,
$$|\alpha'\rangle = \hat{T}|\alpha\rangle \qquad |\beta'\rangle = \hat{T}|\beta\rangle \tag{9.110}$$

There is a general theorem which states that the result $|\langle\beta|\alpha\rangle|^2 = |\langle\beta'|\alpha'\rangle|^2$ implies $\hat{T}$ is either unitary or anti-unitary transformation.

$$\langle\beta'|\alpha'\rangle = \langle\beta|\alpha\rangle \quad \text{Unitary transformation} \tag{9.111}$$

$$\langle\beta'|\alpha'\rangle = \langle\alpha|\beta\rangle \quad \text{Anti-unitary transformation} \tag{9.112}$$

The space–time symmetries of our interest here are translation symmetry in space and time, rotation transformation, space reversal or parity transformation and time reversal transformation. All these transformations, except time reversal transformation, are unitary transformations. The time reversal transformation proves to be anti-unitary.

9.14 INVARIANCE OF HAMILTONIAN

Let us first determine the conditions under which the state vectors $|\alpha\rangle$ and $\hat{U}|\alpha\rangle$ satisfy the Schrödinger equation.

$$|\alpha'\rangle = \hat{U}|\alpha\rangle$$

$$i\hbar\frac{\partial}{\partial t}|\alpha\rangle = \hat{H}|\alpha\rangle$$

Multiply this equation from left by U.

$$\therefore i\hbar\frac{\partial}{\partial t}\hat{U}|\alpha\rangle = \hat{U}\hat{H}|\alpha\rangle$$

or

$$i\hbar\frac{\partial}{\partial t}\underbrace{\hat{U}|\alpha\rangle}_{|\alpha'\rangle} = \underbrace{\hat{U}\ \hat{H}\ \hat{U}^{\dagger}}_{\hat{H}'}\underbrace{\hat{U}|\alpha\rangle}_{|\alpha'\rangle}$$

$$\therefore i\hbar\frac{\partial}{\partial t}|\alpha'\rangle = \hat{H}'|\alpha'\rangle$$

This equation is the same as the Schrödinger equation provided

$$\hat{H}' = \hat{H}$$

Or

$$\hat{U}\hat{H}\hat{U}^{\dagger} = \hat{H}$$

The Hamiltonian $\hat{H}$ is said to be invariant under the symmetry operator $\hat{U}$ if

$$\hat{U}\hat{H}\hat{U}^{\dagger} = \hat{H} \tag{9.113}$$

or

$$\hat{U}\hat{H} = \hat{H}\hat{U} \tag{9.114}$$

What is the significance of invariance of the Hamiltonian under a symmetry transformation $\hat{U}$? First, let us consider the case in which the unitary operator $\hat{U}$ can be written as an exponential of Hamiltonian operators.

i.e.,

$$\hat{U} = e^{i(s_1\hat{X}_1+s_2\hat{X}_2+s_3\hat{X}_3+..)} \tag{9.115}$$

where $s_1, s_2 \ldots$ are some parameters, and $\hat{X}_1, \hat{X}_2 \ldots$ are Hermitian operators.

For the sake of simplicity, let us take $\hat{U}$ to be $\hat{U} = e^{is\hat{X}}$.

Then $\hat{U}\hat{H} = \hat{H}\hat{U}$ means

$$\left[1 + is\hat{X} + \frac{(is)^2\,\hat{X}^2}{2!} + \ldots\right]\hat{H} = \hat{H}\left[1 + is\hat{X} + \frac{(is)^2\,\hat{X}^2}{2!}\,\hat{H}^2 + \ldots\right]$$

This implies that $\hat{X}^n\hat{H}^n = \hat{H}^n\hat{X}^n$ for all n.

$$\therefore\ \hat{X}\hat{H} = \hat{H}\hat{X} \tag{9.116}$$

Let us recall from Sec. 5.6 (see Equation (5.34)) that A is a constant of motion if $\dfrac{d\langle A\rangle}{dt} = 0$.

$$\begin{aligned}\frac{d}{dt}\langle A\rangle &= \frac{d}{dt}\langle\alpha\,|\,\hat{A}\,|\,\alpha\rangle\\ &= \frac{1}{i\hbar}\langle\alpha\,|\,[\hat{A},\hat{H}]\,|\,\alpha\rangle + \left\langle\alpha\,\middle|\,\frac{\partial\hat{A}}{\partial t}\,\middle|\,\alpha\right\rangle\end{aligned}$$

If $\hat{A}$ does not have explicit time dependence, A is said to be a constant of motion or a conserved quantity if

$$[\hat{A},\hat{H}] = 0 \tag{9.117}$$

Comparison of (9.116) and (9.117) implies that there exists a dynamical variable X (note that $\hat{X}$ is a Hermitian operator) which is a constant of motion if there is a symmetry represented by a unitary transformation. i.e., the invariance of the Hamiltonian $\hat{H}$ under unitary transformations of the type $\hat{U} = e^{is\hat{X}}$ implies that there exists a conserved quantity.

In the subsequent sections, we will see that the unitary transformations corresponding to space–time translation symmetry are as follows:

Translation symmetry in space: $\hat{U} = e^{i\,\mathbf{a}.\hat{\mathbf{p}}/\hbar}$

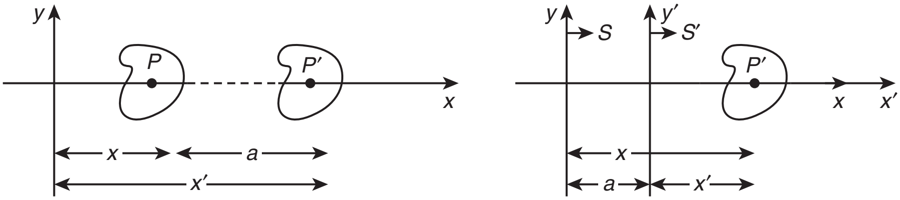
Translation symmetry in time: $\hat{U} = e^{it\hat{H}/\hbar}$

Rotational symmetry $\hat{U} = e^{i\boldsymbol{\theta}\cdot\hat{\mathbf{J}}/\hbar}$

There are symmetries for which $\hat{U}$ cannot be expressed in the Equation (9.115). But even in such case, the invariance of Hamiltonian has important significance. The fact that both $|\,\alpha\rangle$ and $|\,\alpha'\rangle = \hat{U}\,|\,\alpha\rangle$ obey Schrödinger equation has enormous significance. If there exists a physical state $|\,\alpha\rangle$ for a system, then there also exists a physical state corresponding to $|\,\alpha'\rangle = \hat{U}\,|\,\alpha\rangle$. This fact itself is taken as the exhibition of presence of symmetry of a physical property. Parity transformation is an excellent example in this respect. Parity transformation can be shown to be equivalent to mirror reflection. Experimental realization of a particular process and its mirror reflection are taken as conservation of parity. The reader can refer a good book on elementary particle physics for more details on parity.

9.15 ACTIVE AND PASSIVE TRANSFORMATIONS

There are two kinds of transformation related to space–time symmetries. They are active or passive transformations. In the active transformation, we work within the same coordinate systems, and we effect the changes on the system like translating the system or rotating the system and so on. In the passive transformation, the system remains as it is, but the coordinate system is changed like translating the coordinate system or rotating the coordinate system. These transformations are equivalent to each other. We illustrate these ideas by taking some specific cases.



(a) Active transformation (b) Passive transformation

Fig. 9.1 Translation symmetry

Fig. 9.1 refers to the translation symmetry. Fig. 9.1(a) represents the active transformation, and Fig. 9.1(b) represents the passive transformation.

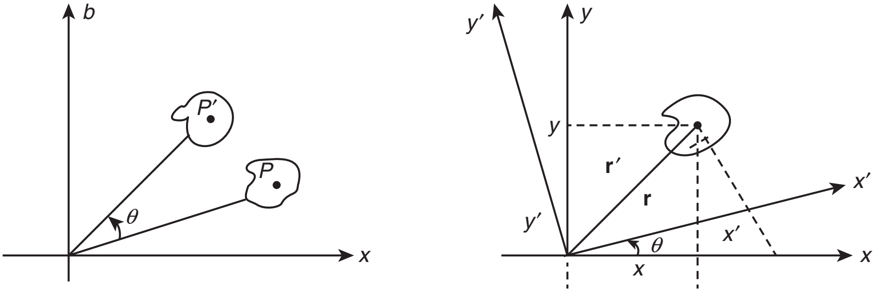
In the active transformation, the coordinate system remains the same. The system as a whole has been shifted along x axis by a distance. The point P on the object has become P'. So $\mathbf{r}$ and $\mathbf{r}'$ refer to two different points with respect to the same coordinate system.

$$x_{P'} = x_P + a \text{ or equivalently } x' = x + a \tag{9.118}$$

In the passive transformation, the coordinate system S has slided along x axis to become S'. The same point P on the object has two different coordinates with respect to S and S'. The coordinates x and x' refer to the same point.

$$\begin{aligned} x'_p &= x_p - a \\ x' &= x - a \end{aligned} \tag{9.119}$$

Now consider another transformation, namely the rotation about z axis.



(a) Active transformation (b) Passive transformation

Fig. 9.2 Rotation about z axis

Fig. 9.2(a) represents the active transformation, and Fig. 9.2(b) represents the passive transformation. In Fig. 9.2(a), the object is rotated by an angle θ about z axis. The point P becomes P'. So $\mathbf{r}$ and $\mathbf{r}'$ refer to points P and P'. The coordinate transformation is given by

$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} \tag{9.120}$$

In passive transformation (Fig. 9.2(b)), the object remains in the same position. The coordinate system has been rotated. The vectors $\mathbf{r}$ and $\mathbf{r}'$ refer to the same point P. The coordinate transformation relating $\mathbf{r}$ and $\mathbf{r}'$ is given by

$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} \tag{9.121}$$

We restrict ourselves only to active transformation in the following sections.

9.16 SPACE TRANSLATION

Let us assume the space, in which the object exists, is homogeneous. In such a space, we can place the system in any region, and the choice of the place of the system should not lead to changes in any experimental outcome.

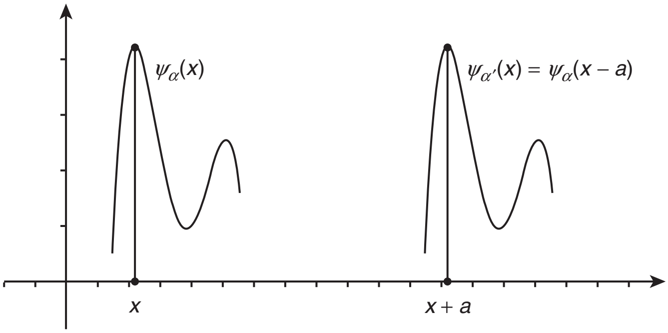


Fig. 9.3 Wave function of the system displaced by a

Let $|\alpha\rangle$ and $|\alpha'\rangle$ be the state vectors when the system is at $\mathbf{r}$ and $\mathbf{r}'$.

$$\mathbf{r}' = \mathbf{r} + \mathbf{a} \tag{9.122}$$

For the sake of simplicity, consider only one coordinate, namely x alone. Then $|x\rangle$ and $|x'\rangle$ are the basis vectors describing the system at x and $x + a$.

$$x' = x + a \tag{9.123}$$

When we displace the system from its original position by a, the wave function also gets displaced by a as shown in Fig. 9.3. Therefore, the space translation invariance implies that

$$\psi_{\alpha}(x)=\psi_{\alpha'}(x')$$

$$\therefore \psi_{\alpha'}(x')=\psi_{\alpha}(x'-a) \tag{9.124}$$

Note that both $\psi_{\alpha'}$ and ψ_{α} are functions of x'. So, we can relabel this variable as x in the Equation (9.124).

$$\therefore \psi_{\alpha'}(x)=\psi_{\alpha}(x-a) \tag{9.125}$$

In terms of Dirac's ket vector formalism, we have

$$|Tx\rangle=|x+a\rangle=\hat{U}|x\rangle \tag{9.126}$$

Multiplying by $\hat{U}^{\dagger}$ from left, we get

$$\hat{U}^{\dagger}|x+a\rangle=|x\rangle$$

Replacing x by $x - a$, we get

$$\hat{U}^{\dagger}|x\rangle=|x-a\rangle$$

$$\langle x|\hat{U}=\langle x-a| \tag{9.127}$$

$$\langle x|\hat{U}|\alpha\rangle=\langle x-a|\alpha\rangle \tag{9.128}$$

But

$$\hat{U}|\alpha\rangle=|\alpha'\rangle \tag{9.129}$$

$$\therefore\ \psi_{\alpha'}(x)=\psi_{\alpha}(x-a)=U\psi_{\alpha}(x) \tag{9.130}$$

Our task now is to determine U. The standard technique in finding such a transformation is to determine U for infinitesimal value of parameter and use them to construct U for the finite value of the parameter.

Let us first shift the system by an infinitesimal displacement ξ. From the Equation (9.130), we have

$$\begin{aligned}\psi_{\alpha'}(x)&=\psi_{\alpha}(x-\xi)=\psi_{\alpha}(x)-\xi\frac{\partial\psi_{\alpha}}{\partial x}\\&=\psi_{\alpha}(x)-i\frac{\xi}{\hbar}\left(-i\hbar\frac{\partial}{\partial x}\right)\psi_{\alpha}(x)=\psi_{\alpha}(x)-i\frac{\xi}{\hbar}p_x\psi_{\alpha}(x)\\&=\left(1-i\frac{\xi}{\hbar}p_x\right)\psi_{\alpha}(x)\end{aligned} \tag{9.131}$$

We have got the transformation for the infinitesimal displacement ξ. Let us displace the system N times in the steps of ξ so that $N\xi = a$. By making use of (9.131) for each step of ξ we get the transformation for finite displacement a as follows:

Translation by ξ: $\psi_{\alpha 1}(x)=\left(1-i\frac{\xi}{\hbar}p_x\right)\psi_{\alpha}(x)$

Translation by 2ξ:

$$\psi_{\alpha 2}(x) = \left(1 - i\frac{\xi}{\hbar}p_x\right)\psi_{\alpha 1}(x) = \left(1 - i\frac{\xi}{\hbar}p_x\right)\left(1 - i\frac{\xi}{\hbar}p_x\right)\psi_\alpha(x)$$

$$= \left(1 - i\frac{\xi}{\hbar}p_x\right)^2 \psi_\alpha(x)$$

Translation by 3ξ:

$$\psi_{\alpha 3}(x) = \left(1 - i\frac{\xi}{\hbar}p_x\right)\psi_{\alpha 2}(x) = \left(1 - i\frac{\xi}{\hbar}p_x\right)\left(1 - i\frac{\xi}{\hbar}p_x\right)^2 \psi_\alpha(x)$$

$$= \left(1 - i\frac{\xi}{\hbar}p_x\right)^3 \psi_\alpha(x)$$

Proceeding in the same way, after repeating N displacements of each of $\xi = (a/N)$, we get

$$\psi_{\alpha'}(x) = \lim_{N\to\infty}\left(1 - i\frac{\xi}{\hbar}p_x\right)^N \psi_\alpha(x) = \lim_{N\to\infty}\left(1 - i\frac{a}{N\hbar}p_x\right)^N \psi_\alpha(x) = e^{-iap_x/\hbar}\psi_\alpha(x)$$

$$\psi_{\alpha'}(x) = e^{-iap_x/\hbar}\psi_\alpha(x) \tag{9.132}$$

It is an interesting exercise to get the same result using ket vector formalism. In terms of ket vectors, the Equation (9.131) is written as

$$\langle x \mid \alpha' \rangle = \langle x \mid \alpha \rangle - i\frac{\xi}{\hbar}\left(-i\hbar\frac{\partial}{\partial x}\right)\langle x \mid \alpha \rangle$$

$$= \langle x \mid \alpha \rangle - i\frac{\xi}{\hbar}\langle x \mid \hat{p}_x \mid \alpha \rangle$$

$$\therefore \quad \mid \alpha' \rangle = \mid \alpha \rangle - i\frac{\xi}{\hbar}\hat{p}_x \mid \alpha \rangle = \left(1 - i\frac{\xi}{\hbar}\hat{p}_x\right) \mid \alpha \rangle$$

We can achieve the translation through a finite value a by repeatedly translating the system by ξ.

Translation by ξ: $\mid \alpha_1 \rangle = \left(1 - i\frac{\xi}{\hbar}\hat{p}_x\right) \mid \alpha \rangle$

Translation by 2ξ: $\mid \alpha_2 \rangle = \left(1 - i\frac{\xi}{\hbar}\hat{p}_x\right) \mid \alpha_1 \rangle$

$$= \left(1 - i\frac{\xi}{\hbar}\hat{p}_x\right)\left(1 - i\frac{\xi}{\hbar}\hat{p}_x\right) \mid \alpha \rangle$$

$$= \left(1 - i\frac{\xi}{\hbar}\hat{p}_x\right)^2 \mid \alpha \rangle$$

Translation by 3ξ: $|\alpha_3\rangle = \left(1 - i\frac{\xi}{\hbar}\hat{p}_x\right)|\alpha_2\rangle$

$$= \left(1 - i\frac{\xi}{\hbar}\hat{p}_x\right)\left(1 - i\frac{\xi}{\hbar}\hat{p}_x\right)^2 |\alpha\rangle$$

$$= \left(1 - i\frac{\xi}{\hbar}\hat{p}_x\right)^3 |\alpha\rangle$$

Proceedings in the same way, after N repeated translations, we get ($N\xi = a$).

$$|\alpha'\rangle = \underset{N\to\infty}{\text{Lt}}\left(1 - i\frac{\xi}{\hbar}\hat{p}_x\right)^N |\alpha\rangle = \underset{N\to\infty}{\text{Lt}}\left(1 - i\frac{a}{N\hbar}\hat{p}_x\right)^N |\alpha\rangle$$

$$= e^{-ia\,\hat{p}_x/\hbar}|\alpha\rangle \tag{9.133}$$

Now we can determine $\psi_{\alpha'}(x) = \langle x|\alpha'\rangle$. Making use of (9.71), we have

$$\langle x|e^{-ia\hat{p}_x/\hbar} = \langle x|\exp\left(-ia\left(-i\hbar\frac{d}{dx}\right)/\hbar\right) = \langle x|\exp(-iap_x/\hbar)$$

$$\psi_{\alpha'}(x) = \langle x|\alpha'\rangle = \langle x|\exp(-iap_x/\hbar)|\alpha\rangle = \langle x|\exp(-iap_x/\hbar)|\alpha\rangle$$

Though $\exp(-iap_x/\hbar)$ is a differential operator, it is not an abstract operator, and so it can be pulled out in the last term of the above equation.

$$\langle x|\exp(-iap_x/\hbar)|\alpha\rangle = \exp(-iap_x/\hbar)\langle x|\alpha\rangle$$

So we have

$$\psi_{\alpha'}(x) = e^{-ia\,p_x/\hbar}\psi_\alpha(x) \tag{9.134}$$

This can be extended to three dimensional displacement $\mathbf{a}$. Let us displace the system from $\mathbf{r}$ to $\mathbf{r}+\mathbf{a}$.

$$|T\mathbf{r}\rangle = |\mathbf{r}+\mathbf{a}\rangle = U|\mathbf{r}\rangle \tag{9.135}$$

$$|\alpha'\rangle = \exp(-i\mathbf{a}.\mathbf{p}/\hbar)|\alpha\rangle$$

$$\therefore \psi_\alpha(\mathbf{r}-\mathbf{a}) = e^{-\frac{i}{\hbar}\mathbf{a}\cdot\mathbf{p}}\psi_\alpha(\mathbf{r}) \tag{9.136}$$

where $\mathbf{p} = -i\hbar\nabla$

Example 9.9 Show that $e^{-iap_x/\hbar}u(x) = u(x-a)$.

Solution: We have $\frac{-iap_x}{\hbar} = \frac{(-ia)}{\hbar}\left(-i\hbar\frac{d}{dx}\right) = -a\frac{d}{dx}$

$$e^{-\frac{iap_x}{\hbar}} = \exp\left(-a\frac{d}{dx}\right) = 1 + \left(-a\frac{d}{dx}\right) + \frac{1}{2!}\left(-a\frac{d}{dx}\right)^2 + \frac{1}{3!}\left(-a\frac{d}{dx}\right)^3 + \cdots\cdots$$

$$e^{-\frac{iap_x}{\hbar}}u(x) = \left[1 + \left(-a\frac{d}{dx}\right) + \frac{1}{2!}\left(-a\frac{d}{dx}\right)^2 + \frac{1}{3!}\left(-a\frac{d}{dx}\right)^3 + \cdots\cdots\right]u(x)$$

$$= u(x) - a\frac{d}{dx}u(x) + \frac{a^2}{2!}\frac{d^2}{dx^2}u(x) - \frac{a^3}{3!}\frac{d^3}{dx^3}u(x) + \cdots\cdot$$

$$= u(x-a)$$

Example 9.10 A simple harmonic oscillator is centred around $x = 0$. It is then displaced by $a = 11/\alpha$, where $\alpha = [m\omega/\hbar]^{1/2}$. How are the ground state wave functions of these two harmonic oscillators related? Sketch their wave functions.

Solution: For the harmonic oscillator centered around $x = 0$, the ground state wave function is

$$u_0(x) = \frac{1}{\pi^{1/4}}e^{-\alpha^2x^2/2} = \frac{1}{\pi^{1/4}}e^{-\rho^2/2}$$

For the harmonic oscillator at $x = 11/\alpha$, the ground state wave function is

$$u_0'(x) = \frac{1}{\pi^{1/4}}e^{-\alpha^2\left(x-\frac{11}{\alpha}\right)^2/2} = \frac{1}{\pi^{1/4}}e^{-\frac{(\rho-11)^2}{2}}$$

The sketches of $u_0(x)$ and $u_0'(x)$ are given in Fig. 9.4.

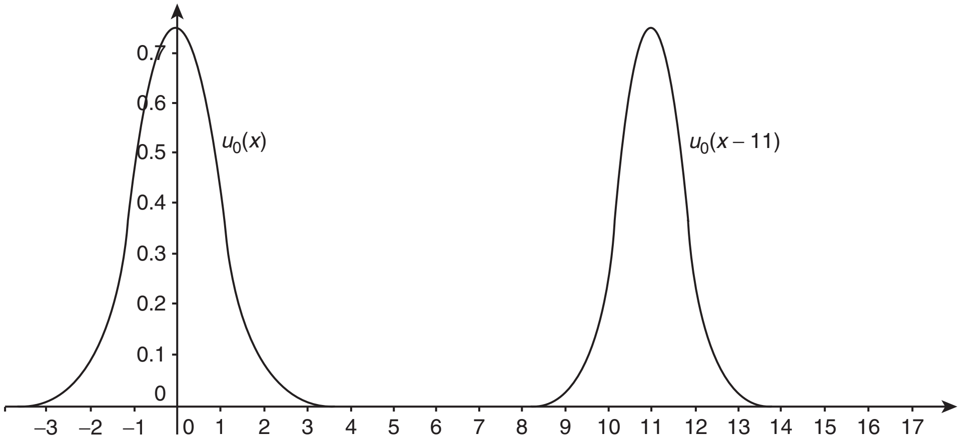


Fig. 9.4 Wave functions of harmonic oscillators centered around x = 0 and x = 11. ($\alpha = 1$)

9.16.1 Momentum as Generator for Space Translation

In group theory, any symmetry operation can be represented by an unitary operator.

$$U = e^{isX} \tag{9.137}$$

where s is some parameter and X is a Hermitian operator. X is said to be the generator of this transformation.

Comparing the Equations (9.136) and (9.137), we can interpret the momentum operator $\mathbf{p}_{op}$ **(or** $\hat{\mathbf{p}}$**)** as the generator for space translation symmetry.

9.17 TIME TRANSLATION SYMMETRY

The translation in time is very much similar to the translation in space. In space translation, we shift the system or the apparatus from one place to another. In time translation, the experiments are performed at different instants of time. Obviously, the environment should be the same at the instants t and $t+\tau$. This means that we are dealing with a system whose Hamiltonian does not change with time.

Our analysis of time translation symmetry is the same as the space translation symmetry.

Let $$t' = t + \tau$$

Time translation symmetry means

$$\psi_{\alpha'}(t') = \psi_\alpha(t)$$

$$\therefore \psi_{\alpha'}(t') = \psi_\alpha(t' - \tau)$$

ψ_α and $\psi_{\alpha'}$ are functions of t'. Let us now relabel this variable as t, and so both $\psi_{\alpha'}$ and ψ_α are functions of time.

$$\therefore \psi_{\alpha'}(t) = \psi_\alpha(t - \tau) \tag{9.138}$$

Proceeding in the same way as done in the case of space translation invariance, we get

$$|\alpha'(t)\rangle = e^{i\tau H/\hbar}\,|\alpha(t)\rangle \tag{9.139}$$

$$\psi_{\alpha'}(t) = \psi_\alpha(t - \tau) = e^{i\tau H/\hbar}\psi_\alpha(t) \tag{9.140}$$

The Hamiltonian $\hat{H}$ can now be interpreted as the generator of time translation symmetry.

9.18 ROTATIONAL INVARIANCE

Rotational invariance implies that the experimental outcomes of a system do not change when we change its orientation. We assume the external environment is the same in both the cases. For the sake of simplicity, let us consider the rotation about z axis. We have already seen (see Fig. 9.2(a)) that the point P of the system changes to P' on rotating the system. The coordinate vector of points P and P' is denoted by $\mathbf{r}$ and $\mathbf{r}'$

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

$$\mathbf{r}_{P'} = R_z(\theta)\mathbf{r}_P \tag{9.141}$$

It is understood that in Equation (9.141), $\mathbf{r}_p$ and $\mathbf{r}_{p'}$ refer to the column vectors $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$ and $\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix}$.

The behaviour of the wave functions under rotational symmetry operations is more complex than the corresponding changes under space-time translation. It is the spin angular momentum which makes the transformation property of the wave function of a system under rotation symmetry operation more complex.

We consider here two kinds of systems: Systems with spin 0 and system with spin 1.

9.18.1 Systems with Spin Zero

Let $|\alpha\rangle$ be the initial state of the system and $|\alpha'\rangle$ be the state vector of the rotated system. Let $\psi_\alpha(\mathbf{r})$ and $\psi_{\alpha'}(\mathbf{r}')$ be corresponding state vectors describing the system before and after rotation. Rotational invariance implies that the physical properties should be the same in both states. This can be achieved by demanding

$$\begin{aligned} \psi_{\alpha'}(\mathbf{r}') &= \psi_\alpha(\mathbf{r}) \\ \psi_{\alpha'}(\mathbf{r}') &= \psi_\alpha(R_z^{-1}\mathbf{r}') \end{aligned} \tag{9.142}$$

In the Equation (9.142) both, ψ_α and $\psi_{\alpha'}$ are functions of $\mathbf{r}'$. We can relabel this variable by $\mathbf{r}$ in the Equation (9.142)

$$\therefore \psi_{\alpha'}(\mathbf{r}) = \psi_\alpha(R_z^{-1}\mathbf{r}) \tag{9.143}$$

where $$R_z^{-1}\mathbf{r} = (x\cos\theta + y\sin\theta)\mathbf{e}_1 + (-x\sin\theta + y\cos\theta)\mathbf{e}_2 + z\mathbf{e}_3$$

$$\therefore \psi_{\alpha'}(x,y,z) = \psi_\alpha(x\cos\theta + y\sin\theta, -x\sin\theta + y\cos\theta, z) \tag{9.144}$$

For infinitesimal values of $\delta\theta$ ($\cos\delta\theta \sim 1, \sin\delta\theta \sim \delta\theta$) the Equation (9.144) becomes

$$\begin{aligned} \psi_{\alpha'}(x,y,z) &= \psi_\alpha(x + y\delta\theta, -x\delta\theta + y, z) \\ &= \psi_\alpha(x,y,z) + y\delta\theta\frac{\partial\psi_\alpha}{\partial x} - x\delta\theta\frac{\partial\psi_\alpha}{\partial y} \\ &= \psi_\alpha(x,y,z) - i\frac{\delta\theta}{\hbar}\left[x\left(-i\hbar\frac{\partial\psi_\alpha}{\partial y}\right) - y\left(-i\hbar\frac{\partial\psi_\alpha}{\partial x}\right)\right] \\ &= \psi_\alpha(x,y,z) - i\frac{\delta\theta}{\hbar}[xp_y - yp_x]\psi_\alpha \\ &= \psi_\alpha(x,y,z) - i\frac{\delta\theta}{\hbar}L_z\psi_\alpha(x,y,z) \\ &= \left(1 - i\frac{\delta\theta}{\hbar}L_z\right)\psi_\alpha(x,y,z) \end{aligned}$$

We have obtained the transformation for rotation through an infinitesimal angle $\delta\theta$ about z axis. For the rotation through a finite angle θ, following the same procedure as done in the case of translational invariance, we get

$$\psi_{\alpha'}(\mathbf{r}) = \psi_\alpha(R_z^{-1}\mathbf{r}) = e^{-i\theta L_z/\hbar}\psi_\alpha(\mathbf{r}) \tag{9.145}$$

In terms of ket vectors, we have

$$|R_z\mathbf{r}\rangle = \hat{U}\,|\mathbf{r}\rangle$$

Multiplying both sides by $\hat{U}^{\dagger}$ we get $\hat{U}^{\dagger}\,|\,R_z\mathbf{r}\rangle = |\,\mathbf{r}\rangle$

Or

$$\hat{U}^{\dagger}\,|\,\mathbf{r}\rangle = |\,R_z^{-1}\mathbf{r}\rangle$$

$$\therefore \langle \mathbf{r}|\hat{U} = \langle R_z^{-1}\mathbf{r}| \tag{9.146}$$

$$\therefore \langle \mathbf{r}|\hat{U}\,|\,\alpha\rangle = \langle R_z^{-1}\mathbf{r}\,|\,\alpha\rangle \tag{9.147}$$

Let us recall that $\quad |\,\alpha'\rangle = \hat{U}\,|\,\alpha\rangle, \quad \psi_{\alpha}(\mathbf{r}) = \langle \mathbf{r}\,|\,\alpha\rangle, \quad \text{and} \quad \psi_{\alpha'}(\mathbf{r}) = \langle \mathbf{r}\,|\,\alpha'\rangle$

$$\psi_{\alpha'}(\mathbf{r}) = \psi_{\alpha}(R_z^{-1}\mathbf{r}) = U\psi_{\alpha}(\mathbf{r}) \tag{9.148}$$

By comparing the Equations (9.148) with (9.145), U is given by $\hat{U} = e^{-i\theta\hat{L}_z/\hbar}$. Note this is a differential operator since L_z itself is a differential operator. The corresponding abstract operator is

$$\hat{U} = e^{-i\theta\hat{L}_z/\hbar} \tag{9.149}$$

This can be extended to rotation about an arbitrary axis **n** by an angle θ.

It is given by

$$|\,R\mathbf{r}\rangle = \hat{U}\,|\,\mathbf{r}\rangle = e^{-i\theta\mathbf{n}.\hat{\mathbf{L}}/\hbar}\,|\,\mathbf{r}\rangle \tag{9.150}$$

$$\psi_{\alpha'}(\mathbf{r}) = \psi_{\alpha}(R_n^{-1}(\theta)\mathbf{r}) = e^{-i\theta\mathbf{n}.\mathbf{L}/\hbar}\psi_{\alpha}(\mathbf{r})$$

Obviously, the orbital angular momentum operator $\mathbf{L}_{op}$ is the generator for the rotational symmetry transformation for spin zero particle.

9.18.2 Rotational Invariance and Systems with Spin

For non-zero spin angular momentum, the state vector has many components. For instance, a Dirac spinor describing spin 1/2 particles has four components. The spin 1 particle has three components. In general, the transformation of a multicomponent object is

$$\psi_{\alpha'i}(\mathbf{r}_{p'}) = \sum_j R_{ij}\psi_{\alpha j}(\mathbf{r}_p) \tag{9.151}$$

where R_{ij} depends on the spin of the particle

The simplest case is 'vector' particle (for which the spin is 1), for which the matrix R_{ij} is given by

$$R_{ij} = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{9.152}$$

The wave function of this particle ψ_α is given by

$$\psi_\alpha(\mathbf{r}) = \begin{bmatrix} \psi_{\alpha 1}(\mathbf{r}) \\ \psi_{\alpha 2}(\mathbf{r}) \\ \psi_{\alpha 3}(\mathbf{r}) \end{bmatrix} \tag{9.153}$$

We find that proceeding in the same way as in the earlier sections, under rotational symmetry operations, the transformation for $\psi_\alpha(\mathbf{r})$ is given by

$$\psi_{\alpha' i}(\mathbf{r}) = \sum_j R_{ij}\psi_{\alpha j}(R_z^{-1}\mathbf{r}) \tag{9.154}$$

For rotation by an infinitesimal angle $\delta\theta$,

$$\begin{bmatrix} \psi_{\alpha 1}(x,y,z) \\ \psi_{\alpha' 2}(x,y,z) \\ \psi_{\alpha' 3}(x,y,z) \end{bmatrix} = \begin{bmatrix} 1 & -\delta\theta & 0 \\ \delta\theta & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \psi_{\alpha 1}(R_z^{-1}\mathbf{r}) \\ \psi_{\alpha 2}(R_z^{-1}\mathbf{r}) \\ \psi_{\alpha 3}(R_z^{-1}\mathbf{r}) \end{bmatrix}$$

$$= \left\{ \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \frac{i\delta\theta}{\hbar}\hbar \begin{bmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \right\} \begin{bmatrix} \psi_{\alpha 1}(R_z^{-1}\mathbf{r}) \\ \psi_{\alpha 2}(R_z^{-1}\mathbf{r}) \\ \psi_{\alpha 3}(R_z^{-1}\mathbf{r}) \end{bmatrix}$$

$$= \left[1 - \frac{i\delta\theta}{\hbar}\hbar \sum\nolimits_z \right] \left[1 - \frac{i\delta\theta}{\hbar} L_z \right] \begin{bmatrix} \psi_{\alpha 1}(x,y,z) \\ \psi_{\alpha 2}(x,y,z) \\ \psi_{\alpha 3}(x,y,z) \end{bmatrix}$$

where Σ_z is given by

$$\sum\nolimits_z = \begin{bmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Spin angular momentum operator for spin 1 particle is given by $S_z = \hbar\Sigma_z$.

$$\begin{bmatrix} \psi_{\alpha' 1}(\mathbf{r}) \\ \psi_{\alpha' 2}(\mathbf{r}) \\ \psi_{\alpha' 3}(\mathbf{r}) \end{bmatrix} = \left(1 - \frac{i\delta\theta}{\hbar} S_z\right)\left(1 - \frac{i\delta\theta}{\hbar} L_z\right) \begin{bmatrix} \psi_{\alpha 1}(\mathbf{r}) \\ \psi_{\alpha 2}(\mathbf{r}) \\ \psi_{\alpha 3}(\mathbf{r}) \end{bmatrix}$$

$$= \left[1 - \frac{i\delta\theta}{\hbar}(S_z + L_z) \right] \begin{bmatrix} \psi_{\alpha 1}(\mathbf{r}) \\ \psi_{\alpha 2}(\mathbf{r}) \\ \psi_{\alpha 3}(\mathbf{r}) \end{bmatrix} \tag{9.155}$$

Or simply,

$$\psi_{\alpha'}(\mathbf{r}) = \left[1 - \frac{i\delta\theta}{\hbar}(L_z + S_z)\right]\psi_\alpha(\mathbf{r}) \tag{9.156}$$

For finite angle θ, this equation becomes

$$\psi_{\alpha'}(\mathbf{r}) = e^{-i\theta(L_z + S_z)/\hbar}\psi_\alpha(\mathbf{r}) \tag{9.157}$$

This can be extended to the rotation about x axis and y axis. Let us define S_x, S_y and S_z as $S_x = \hbar\Sigma_x, S_y = \hbar\Sigma_y$ and $S_z = \hbar\Sigma_z$.

$$S_x = \hbar\sum_x = \hbar\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{bmatrix} \tag{9.158}$$

$$S_y = \hbar\sum_y = \hbar\begin{bmatrix} 0 & 0 & i \\ 0 & 0 & 0 \\ -i & 0 & 0 \end{bmatrix} \tag{9.159}$$

$$S_z = \hbar\sum_z = \hbar\begin{bmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \tag{9.160}$$

It is easy to check whether S_x, S_y and S_z obey angular momentum algebra.

$$[S_x, S_y] = i\hbar S_z,\quad [S_y, S_z] = i\hbar S_x,\ \text{and}\ [S_z, S_x] = i\hbar S_y \tag{9.161}$$

Furthermore, the eigenvalue of S^2 is $2\hbar^2$ and the eigenvalues of S_z are +1, 0, −1. Therefore S_x, S_y and S_z are interpreted as angular momentum operators. In the next chapter, we can see that S_x, S_y and S_z are identified as spin angular momentum. There is no classical analogue to the spin angular momentum, and it is an intrinsic property of a particle in quantum mechanics in the sense that it is a property like charge or mass which cannot be taken away from the particle.

So the Equation (9.157) describes the effect of rotation about Z axis on the state vector of the system.The rotation of the system about an arbitrary vector **n** by an angle θ is given by

$$\begin{bmatrix} \psi_{\alpha'1}(\mathbf{r}) \\ \psi_{\alpha'2}(\mathbf{r}) \\ \psi_{\alpha'3}(\mathbf{r}) \end{bmatrix} = e^{-i\frac{\theta}{\hbar}\mathbf{n}.(\mathbf{L}+\mathbf{S})} \begin{bmatrix} \psi_{\alpha1}(\mathbf{r}) \\ \psi_{\alpha2}(\mathbf{r}) \\ \psi_{\alpha3}(\mathbf{r}) \end{bmatrix}$$

$$\begin{aligned} \psi_{\alpha'}(\mathbf{r}) &= e^{-i\frac{\theta}{\hbar}\mathbf{n}.(\mathbf{L}+\mathbf{S})}\psi_\alpha(\mathbf{r}) \\ &= e^{-i\theta\,\mathbf{n}.\mathbf{J}/\hbar}\psi_\alpha(\mathbf{r}) \end{aligned} \tag{9.162}$$

where **J** is the total angular momentum. Note that the total angular momentum operator **J** is the generator for rotational symmetry transformation for a particle with non-zero spin.

Example 9.11 Show that S_x, S_y and S_z given in (9.158)-(9.162) satisfy the angular momentum algebra.

Solution: Let as evaluate $[S_x, S_y]$.

$$[S_z,S_x]=\hbar^2\left\{\begin{bmatrix}0&0&0\\0&0&-i\\0&i&0\end{bmatrix}\begin{bmatrix}0&0&i\\0&0&0\\i&0&0\end{bmatrix}-\begin{bmatrix}0&0&i\\0&0&0\\i&0&0\end{bmatrix}\begin{bmatrix}0&0&0\\0&0&-i\\0&i&0\end{bmatrix}\right\}$$

$$=i\hbar\hbar\begin{bmatrix}0&-i&0\\-i&0&0\\0&0&0\end{bmatrix}=i\hbar\hbar\sum\nolimits_z=i\hbar S_z.$$

The commutation relation $[S_y, S_z]$ and $[S_z, S_x]$ can be checked in the same way.

Example 9.12 Show that $S^2 = S_x^2 + S_y^2 + S_z^2 = 2I$

Solution:

$$S_x^2=\hbar^2\sum\nolimits_x^2=\hbar^2\begin{bmatrix}0&0&0\\0&0&-i\\0&i&0\end{bmatrix}\begin{bmatrix}0&0&0\\0&0&-i\\0&i&0\end{bmatrix}$$

$$=\hbar^2\begin{bmatrix}0&0&0\\0&1&0\\0&0&1\end{bmatrix}$$

Similarly, we can find

$$S_y^2=\hbar^2\begin{bmatrix}1&0&0\\0&0&0\\0&0&1\end{bmatrix}\quad\text{and}\quad S_z^2=\hbar^2\begin{bmatrix}1&0&0\\0&1&0\\0&0&0\end{bmatrix}$$

$$\therefore S^2=2\hbar^2 I$$

$$S^2\chi=s(s+1)\hbar^2\chi$$

$$\therefore s=1$$

Example 9.13 Evaluate $e^{-i\theta S_y/\hbar}S_x e^{i\theta S_y/\hbar}$

Solution: Let us use the following formula to evaluate this expression:

$$e^{\alpha X}Ye^{-\alpha X}=Y+\alpha[X,Y]+\frac{\alpha^2}{2!}[X,[X,Y]]+\frac{\alpha^3}{3!}[X,[X,[X,Y]]]+\ldots$$

$$e^{-\frac{i\theta S_Y}{\hbar}}S_x e^{\frac{i\theta S_Y}{\hbar}}=S_x+\left(-\frac{i\theta}{\hbar}\right)[S_y,S_x]+\left(-\frac{i\theta}{\hbar}\right)^2[S_y,[S_y,S_x]+\ldots$$

$$[S_y,S_x]=-i\hbar S_Z;\ [S_y,-i\hbar S_z]=\hbar^2 S_x$$

$$[S_y,\hbar^2 S_x]=-i\hbar^3 S_Z;\ [S_y,-i\hbar^3 S_z]=\hbar^4 S_x\ldots$$

Making use of these results we get

$$
\begin{aligned}
e^{-i\theta S_Y/\hbar} S_x e^{i\theta S_Y/\hbar} &= S_x - \theta S_z - \frac{\theta^2}{2!} S_x + \frac{\theta^3}{3!} S_z + \frac{\theta^4}{4!} S_x + \dots \\
&= S_x \left(1 - \frac{\theta^2}{2!} + \frac{\theta^4}{4!} + \dots\right) - S_z \left(\theta - \frac{\theta^3}{3!} + \dots\right) \\
&= \cos\theta \ S_x - \sin\theta \ S_z
\end{aligned}
$$

In the same way, we can evaluate

$$
e^{-\frac{i\theta S_Y}{\hbar}} S_z e^{\frac{i\theta S_Y}{\hbar}} = \cos\theta \ S_z + \sin\theta \ S_x
$$

i.e., the operators S_x, S_y, and S_z, themselves transform like a vector.

$$
\begin{aligned}
e^{-i\theta S_Y/\hbar} \begin{bmatrix} S_x \\ S_y \\ S_z \end{bmatrix} e^{i\theta S_Y/\hbar} &= \begin{bmatrix} \cos\theta & 0 & -\sin\theta \\ 0 & 1 & 0 \\ \sin\theta & & \cos\theta \end{bmatrix} \begin{bmatrix} S_x \\ S_y \\ S_z \end{bmatrix} \\
\therefore e^{-i\theta S_y/\hbar} \mathbf{S}_{op} e^{i\theta S_y/\hbar} &= R_y(\theta) \ \mathbf{S}_{op}
\end{aligned}
$$

9.19 PARITY INVARIANCE – SPACE INVERSION

Space inversion corresponds to inversion of the system about the origin. Space inversion is also known as parity transformation. Parity is one of the important concepts in quantum mechanics. For a long time, it was believed that parity is also conserved like conservation of linear and angular momentum. Conservation of linear and angular momentum emerges as a consequence of space-time symmetries. Parity or space inversion is also a space-time symmetry. So it is reasonable to expect conservation of parity also. Now, we know that parity is conserved in strong and electromagnetic interactions, but it is not conserved in weak interactions.

The concept of parity is non-trivial in quantum mechanics. In the case of angular momentum, we have two kinds: orbital and spin angular momentum. There is a classical analogue to orbital angular momentum, but there is no classical analogue to spin angular momentum. In the same way, in the case of parity also, there are two kinds of parities: spatial and intrinsic parities. The spatial parity has a classical analogue, but the intrinsic parity is unique to quantum mechanics. Before going into these aspects, let us first study the transformation relating a system and the corresponding inverted system.

Under space inversion, the point P becomes P'. The position vectors P and P' are

$$
\begin{aligned}
\mathbf{r}_p &= \mathbf{r} = x\mathbf{e}_1 + y\mathbf{e}_2 + z\mathbf{e}_3 \\
\mathbf{r}_{p'} &= \mathbf{r}' = x'\mathbf{e}_1 + y'\mathbf{e}_2 + z'\mathbf{e}_3 = -x\mathbf{e}_1 - y\mathbf{e}_2 - z\mathbf{e}_3
\end{aligned}
$$

$$\therefore \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \underbrace{\begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}}_{\mathfrak{I}} \begin{bmatrix} x \\ y \\ z \end{bmatrix} \tag{9.163}$$

$$\mathbf{r}' = \mathfrak{I}\mathbf{r} \text{ and } \mathbf{r} = \mathfrak{I}\mathbf{r}' \tag{9.164}$$

where $\mathfrak{I}$ is the inversion matrix.

If we go by earlier experience in the consideration of space-time translation symmetry or rotation symmetry, we will write the transformation equation as follows:

$$\hat{U}\,|\,\mathbf{r}\rangle = |\,\mathfrak{I}\mathbf{r}\rangle \qquad \hat{U}\,|\,\alpha\rangle = |\,\alpha'\rangle$$

$$\psi_{\alpha'}(\mathbf{r}) = \psi_{\alpha}(-\mathbf{r}) = U\psi_{\alpha}(\mathbf{r}).$$

However, it turns out that these transformations are not sufficient to account for all phenomena involving elementary particles. So these transformations have to be modified by introducing a new factor $\eta_{i.}$

$$\psi_{\alpha'}(\mathbf{r}) = \eta_i \psi_{\alpha}(-\mathbf{r}) \tag{9.165}$$

But

$$\psi_{\alpha'}(\mathbf{r}) = U\psi_{\alpha}(\mathbf{r}) \tag{9.166}$$

$$\therefore U\psi_{\alpha}(\mathbf{r}) = \eta_i \psi_{\alpha}(-\mathbf{r}) \tag{9.167}$$

η_i is called intrinsic parity quantum number. The possible values of η_i can be determined easily.

Let us invert the system twice. Then we have

$$\begin{aligned} &P \xrightarrow{\text{inversion}} P' \xrightarrow{\text{inversion}} P \\ &\mathbf{r} \xrightarrow{\text{inversion}} \mathbf{r}' = -\mathbf{r} \xrightarrow{\text{inversion}} -\mathbf{r}' = \mathbf{r} \end{aligned} \tag{9.168}$$

The corresponding operations on the state vector will be as follows:

Under the first inversion (from 9.165), $\psi_{\alpha}(\mathbf{r}) \to \eta_i \psi_{\alpha}(-\mathbf{r})$.

Under the second inversion, we should get the wave function $\psi_{\alpha}(\mathbf{r})$.

i e.,

$$\eta_i \psi_{\alpha}(-\mathbf{r}) \to \eta_i \eta_i \psi_{\alpha}(\mathbf{r}) = \psi_{\alpha}(\mathbf{r})$$

$$\begin{aligned} &\psi_{\alpha}(\mathbf{r}) \xrightarrow{\text{inversion}} \eta_i \psi_{\alpha}(-\mathbf{r}) \xrightarrow{\text{inversion}} \eta_i(\eta_i \psi_{\alpha}(\mathbf{r})) = \psi_{\alpha}(\mathbf{r}) \\ &\therefore \eta_i^2 = 1 \qquad \therefore \eta_i = \pm 1. \end{aligned} \tag{9.169}$$

9.19.1 Intrinsic Parity

It is to be noted that the structure of the Equation (9.167) has nothing to do with odd or even function of $\mathbf{r}$. It relates two different states of a system. $\psi_{\alpha}(\mathbf{r})$ and $\psi_{\alpha'}(\mathbf{r})$ represent two different states. If $\psi_{\alpha}(\mathbf{r})$ represents the state of the system, $\psi_{\alpha'}(\mathbf{r})$ represents the state of the inverted system. If the wave function $\psi_{\alpha}(\mathbf{r})$ is either odd or even, $\psi_{\alpha}(-\mathbf{r}) = \pm\psi_{\alpha}(\mathbf{r})$. This equation is different from (9.169).

As we noted earlier, there are two kinds of parity quantum number: intrinsic and spatial parity quantum numbers. Different particles have different intrinsic parity quantum number. The value intrinsic quantum number (η_i is 1 or –1) is assigned for each particle in a way consistent with experimental observation. η_i for some particles are listed.

Particles	η_i
$e^-, \mu^-, \nu_e; \nu_\mu, \tau^-, \nu_\tau$	1
$e^+, \mu^+, \tau^+, \bar{\nu}_e, \bar{\nu}_\mu, \bar{\nu}_\tau$	–1
quarks : u, d, s, c, b, t	1
anti quarks : $\bar{u}, \bar{d}, \bar{s}, \bar{c}, \bar{b}, \bar{t}$	–1
p, n	1
$\bar{p}, \bar{n}$	–1
$\pi^+, \pi^0, \pi^-, K^+, K^-, K^0, \bar{K}^0$	–1
γ, W^+, W^-, Z, gluons	–1

In addition to intrinsic parity, we have spatial parity as given below:

$$\psi_\alpha(-\mathbf{r}) = \eta_s \psi_\alpha(\mathbf{r}) \qquad \begin{cases} \eta_s = 1 & \text{even function} \\ \eta_s = -1 & \text{odd function} \end{cases} \tag{9.170}$$

$$\therefore \psi_{\alpha'}(-\mathbf{r}) = \eta_i \eta_s \psi_\alpha(\mathbf{r}) \tag{9.171}$$

For instance, consider the hydrogen atom. The parity of the system is

$$\Pi = \eta_{ie^-} \eta_{ip} (-1)^l \tag{9.172}$$

where η_{ie-} and η_{ip} are the intrinsic parities of the electron and proton, and l is the orbital angular momentum of the particle.

Example 9.14 Assuming $U_I \psi_\alpha(\mathbf{r}) = \eta_i \psi_\alpha(-\mathbf{r})$ when U_1 is the parity operator, show that $U_I^\dagger \mathbf{r} U_I = -\mathbf{r}$.

Solution: We know that $\mathbf{r}_{op} \psi_\alpha(\mathbf{r}) = \mathbf{r}\, \psi_\alpha(\mathbf{r})$

Multiplying the above equation by U_I, we get

$$U_I \mathbf{r}_{op} U_I^\dagger \underbrace{U_I \psi_\alpha(\mathbf{r})}_{\eta_i \psi_\alpha(-\mathbf{r})} = \underbrace{U_I(\mathbf{r}\, \psi_\alpha(\mathbf{r}))}_{-\eta_i \mathbf{r} \psi_\alpha(-\mathbf{r})}$$

$$U_I \mathbf{r}_{op} U_I^\dagger \eta_i \psi_\alpha(-\mathbf{r}) = -\eta_i \mathbf{r}\, \psi_\alpha(-\mathbf{r})$$

$$\therefore U_I \mathbf{r}_{op} U_I^\dagger = -\mathbf{r}_{op}$$

Example 9.15 Show that $U_I \mathbf{p}_{op} U_I^{+} = -\mathbf{p}_{op}$.

Solution: We know that $\mathbf{p}_{op}\psi_\alpha(\mathbf{r}) = -i\hbar\nabla\psi_\alpha(\mathbf{r})$.

Multiplying this equation by U_I, we get

$$U_I \mathbf{p}_{op} U_I^\dagger \underbrace{U_I \psi_\alpha(\mathbf{r})}_{\eta_i \psi_\alpha(-\mathbf{r})} = \underbrace{-i\hbar U_I(\nabla\psi_\alpha(\mathbf{r}))}_{\eta_i i\hbar\nabla\psi_\alpha(-\mathbf{r})}$$

$$U_I \mathbf{p}_{op} U_I^\dagger \eta_i \psi_\alpha(-\mathbf{r}) = \eta_i(-\mathbf{p}_{op})\psi_\alpha(-\mathbf{r})$$

$$\therefore U_I \mathbf{p}_{op} U_I^\dagger = -\mathbf{p}_{op}$$

9.19.2 Conservation of Parity

Leaving aside gravitation, the basic interactions among the elementary particles are strong, electromagnetic and weak interactions. In the process or reactions governed by strong and electromagnetic interactions, parity is found to be conserved. i.e., initial parity and final parity will be the same. In weak interactions also, it was believed that parity is conserved. However, when $\theta - \tau$ puzzle was resolved, violation of parity in weak interaction was admitted. In fact, the famous experiment by Wu et al.in 1957 using the radioactive decay of Co-60 confirmed the violation of parity in weak interaction. The reader can consult a good book on nuclear physics for parity violation. The $\theta - \tau$ puzzle is presented here to illustrate the application of the concept of intrinsic parity.

$\theta - \tau$ Puzzle

In elementary particle physics, in the 1950s, new particles were continuously being discovered. There were many misunderstandings which got clarified later. One such puzzle is $\theta - \tau$ puzzle. Once, it was thought θ and τ were two different positively charged spin zero particles. Now, we know that they are the same particle K^+. It is the identification of θ^+ and τ^+ as K^+ is the resolution of the $\theta - \tau$ puzzle. To understand the $\theta - \tau$ puzzle, let us pretend that we do not know much about the θ^+ and τ^+ particles.

The parity of θ^+ and τ^+ was to be determined from their decay products. It is found that

$$\tau^+ \rightarrow \pi^+ + \pi^+ + \pi^-$$
$$\theta^+ \rightarrow \pi^0 + \pi^+$$

π^+ and π^0 and π^- are spin zero particles. Their intrinsic parity quantum number is −1. Assuming that the conservation of parity is valid in weak interaction (decay process), the intrinsic parity of τ^+ is determined to be −1 and the intrinsic parity of θ^+ is determined to be +1.

i.e.,
$$\eta_{\mathrm{i}}(\tau^+) = -1$$
$$\eta_i(\theta^+) = 1$$

What is the $\theta - \tau$ puzzle? The puzzle is that apart from the intrinsic parity η_i for all other parameters like charge, mass, spin and so on, θ^+ and τ^+ were identical suggesting θ^+ are τ^+ are the same particle. The only difference is the intrinsic parity quantum number. In fact, the resolution of this puzzle lies in the fact recognizing θ^+ and τ^+ as the same particle. (Now it is identified as K^+.). It has to be realized that the determination of intrinsic parity of θ^+ and τ^+ was based on the conservation of parity. Hence, this puzzle suggests the violation of the conservation of parity. If one accepts the violation of the conservation of parity in weak interaction, the determination of intrinsic parity through different decay modes is not acceptable. The intrinsic parity of K^+ is now found to be −1.

Example 9.16 The τ^+ and θ^+ particles decay as given below:

$$\tau^+ \to \pi^+ + \pi^+ + \pi^-$$
$$\theta^+ \to \pi^+ + \pi^0$$

Assuming the conservation of parity, determine the intrinsic parity of τ^+ and θ^+.

Solution: Let $\eta_i(\pi^+), \eta_i(\pi^-)$ and $\eta_i(\pi^0), \eta_i(\tau^+), \eta_i(\theta^+)$ be the intrinsic parties of π^+, π^-, π^0, τ^+ and θ^+, respectively. We know that

$$\eta_i(\pi^+) = \eta_i(\pi^-) = \eta_i(\pi^0) = -1$$

Let L_{12} be the relative orbital angular momentum of π^+ and π^0 in the θ^+ decay.

$\therefore$ Parity of the decay products in $\theta^+ = \eta_i(\pi^+)\eta_i(\pi^0)(-1)^{L_{12}}$

From the conservation of the angular momentum, we get L_{12} should be equal to the spin angular momentum of θ^+ which is zero.

$$\therefore L_{12} = 0$$
$$\therefore \eta_i(\theta^+) = \eta_i(\pi^+)\eta_i(\pi^0)(-1)^{L_{12}} = (-1)(-1)(-1)^0 = 1.$$

Next, let us consider τ^+ decay. Let l_{12} be the orbital angular momentum due to the relative motion between $\pi^+$and π^+. l_3 is the angular momentum of π^- due to its motion relative to π^- and π^+.

$$\begin{aligned}\therefore \eta_i(\tau^+) &= [\eta_i(\pi^+)\cdot\eta_i(\pi^+)\cdot(-1)^{l_{12}}]\cdot[\eta_i(\pi^-)(-1)^{l_3}] \\ &= (-1)\cdot(-1)(-1)^{l_{12}}(-1)(-1)^{l_3} \\ &= (-1)(-1)^{l_{12}+l_3}\end{aligned}$$

From the conservation of angular momentum, we know $l_{12} + l_3$ should be equal to the angular momentum of τ^+ which is zero.

i.e.,
$$l_{12} + l_3 = 0$$
$$\therefore \eta_i(\tau^+) = -1$$

9.20 TIME REVERSAL

Time reversal is similar to space inversion. In space inversion, the system is inverted about the origin; i.e., the point P of the system goes to P' where $\mathbf{r}_{P'} = -\mathbf{r}_P$. In the time reversal, t goes to $t' = -t$. We proceed in the same way as done in the case of space reversal or rotational transformation. In all the previous cases, our task is to determine the unitary transformation relating the states of the original system and the transformed system. In the same way, we would like to write the transformation equations as

$$|\alpha'(t)\rangle = \hat{\mathbf{T}}\,|\alpha(-t)\rangle \qquad (9.173)$$

Or

$$\psi_{\alpha'}(\mathbf{r},t) = T\psi_\alpha(\mathbf{r},-t) \qquad (9.174)$$

Can we get such transformation operator T? Even if we get such operator, will it be unitary operator? To answer these questions, let us start with time reversal transformation in classical mechanics.

In classical physics, the position vector $\mathbf{r}$, the momentum rector $\mathbf{p} = m\,d\mathbf{r}/dt$ and the angular momentum $\mathbf{L}$ transform as follows:

$$\begin{aligned}\mathbf{r} &\to \mathbf{r}\\ \mathbf{p} &\to -\mathbf{p}\\ \mathbf{L} &\to -\mathbf{L}\end{aligned}$$

This can be seen as follows:

$$\mathbf{r}(t) \to \mathbf{r}(-t)$$

$$\begin{aligned}\mathbf{p}(t) \to \mathbf{p}' = \frac{d\mathbf{r}'}{dt'} = -\frac{d\mathbf{r}(-t)}{dt} = -\mathbf{p}(-t)\\ \mathbf{L}(t) \to \mathbf{L}' = \mathbf{r}\times\mathbf{p} = -\mathbf{r}(-t)\times\mathbf{p}(-t) = -\mathbf{L}(-t)\end{aligned}$$

So we expect to retain this correspondence in quantum mechanics also.

$$\begin{aligned}\therefore \mathbf{r}_{op} &\to \mathbf{r}'_{op} = T\mathbf{r}_{op}T^{-1}\\ \mathbf{p}_{op} &\to \mathbf{p}'_{op} = T\mathbf{p}_{op}T^{-1} = -\mathbf{p}_{op}\\ \mathbf{L}_{op} &\to \mathbf{L}'_{op} = T\mathbf{L}_{op}T^{-1} = -\mathbf{L}_{op}\end{aligned}$$

Let us check whether these transformation equations are consistent with each other. Consider a commutation relation $[x, p_x] = i\hbar$.

$$xp_x - p_x x = i\hbar$$

Introduce T on the left and T^{-1} on the right.

$$\begin{aligned}Txp_xT^{-1} - Tp_xxT^{-1} &= Ti\hbar T^{-1}\\ TxT^{-1}Tp_xT^{-1} - Tp_xT^{-1}TxT^{-1} &= Ti\hbar T^{-1}\\ x'(-p'_x) - (-p'_x)(+x') &= Ti\hbar T^{-1}\\ -i\hbar &= Ti\hbar T^{-1}\end{aligned}$$

Obviously, a unitary operator cannot achieve this result. In fact, the structure of the equation is

$$TcT^{-1} = c* \tag{9.175}$$

where c is a constant. This, in turn, suggests that T should be an anti-linear operator.

9.20.1 Anti-linear Operator

Let us recall that a linear operator T satisfies the conditions

$$T(C_1\psi_\alpha + C_2\psi_\beta) = C_1T\psi_\alpha + C_2T\psi_\beta$$

An operator T is said to be anti-linear if it satisfies the condition

$$T(C_1\psi_\alpha + C_2\psi_\beta) = C_1^* T\psi_\alpha + C_2^* T\psi_\beta \tag{9.176}$$

This implies
$$TC = C^* T \tag{9.177}$$

This suggests that we can write T as

$$T = UK = KU \tag{9.178}$$

where U is an unitary operator and K is an operator which complex-conjugates the wave function on which it operates.

i.e.,
$$K\psi_\alpha = \psi_\alpha^* \tag{9.179}$$

$$\therefore \psi_{\alpha'}(\mathbf{r}, t) = T\psi_\alpha(\mathbf{r}, -t) = U\psi_\alpha^*(\mathbf{r}, -t) \tag{9.180}$$

Example 9.17 Show that the norm of a state vector is invariant under time reversal transformation.

Solution:

$$\begin{aligned}
\psi_{\alpha'} &= T\psi_\alpha = UK\psi_\alpha \\
\therefore (\psi_{\alpha'}, \psi_{\alpha'}) &= (UK\psi_\alpha, UK\psi_\alpha) \\
&= (K\psi_\alpha, U\ UK\psi_\alpha) = (K\psi_\alpha, K\psi_\alpha) \\
&= (\psi_\alpha^*, \psi_\alpha^*) = (\psi_\alpha, \psi_\alpha)^*
\end{aligned}$$

Since $(\psi_\alpha, \psi_\alpha)$ is a real number,

$$(\psi_{\alpha'}, \psi_{\alpha'}) = (\psi_\alpha, \psi_\alpha)$$

Note the scalar product between ψ_α and ψ_β is, in general, a complex number. Let $\psi_{\alpha'}$ and $\psi_{\beta'}$ be given by

$$\psi_{\alpha'} = T\psi_\alpha,\ \psi_{\beta'} = T\psi_\beta,$$

$$\begin{aligned}
(\psi_{\alpha'}, \psi_{\beta'}) &= (UK\psi_\alpha, UK\psi_\beta) \\
&= (K\psi_\alpha, U^\dagger UK\psi_\beta) = (K\psi_\alpha, K\psi_\beta) \\
&= (\psi_\alpha^*, \psi_\beta^*) = (\psi_\alpha, \psi_\beta)^* = (\psi_\beta, \psi_\alpha).
\end{aligned}$$

9.20.2 Condition for $\psi_{\alpha'}(t)$ to Satisfy Schrödinger Equation

$$\psi_{\alpha'}(\mathbf{r}, t) = T\psi_\alpha(\mathbf{r}, -t) = T\psi_\alpha(\mathbf{r}, t')$$

Note T is an anti-linear operator that operates on the state vectors which are the elements of Hilbert space. T should be independent of time.

$$T\frac{\partial}{\partial t} = \frac{\partial}{\partial t}T$$

Since T is an anti-linear operator, $TC = C^*T$. With their background information let us determine $i\hbar\frac{\partial}{\partial t}\psi_{\alpha'}$

$$\begin{aligned}
i\hbar\frac{\partial}{\partial t}\psi_{\alpha'}(\mathbf{r},t) &= i\hbar\frac{\partial}{\partial t}T\psi_\alpha(\mathbf{r},-t)\\
&= i\hbar T\frac{\partial}{\partial t}\psi_\alpha(\mathbf{r},-t)\\
&= T\left(-i\hbar\frac{\partial}{\partial t}\psi_\alpha(\mathbf{r},-t)\right) = T\left(i\hbar\frac{\partial}{\partial t'}\psi_\alpha(\mathbf{r},t')\right)\\
&= TH\psi_\alpha(\mathbf{r},t') = THT^{-1}\underbrace{T\psi_\alpha(\mathbf{r},t')}_{\psi_{\alpha'}(\mathbf{r},t)}\\
&= THT^{-1}\psi_{\alpha'}(\mathbf{r},t)
\end{aligned}$$

$\therefore \psi_{\alpha'}(\mathbf{r},t)$ satisfies the Schrödinger equation provided $THT^{-1} = H$. (9.181)

i.e., $\psi_{\alpha'}(\mathbf{r},t)$ satisfies the Schrödinger equation if the Hamiltonian is invariant under time reversal transformation.

Example 9.18 The plane wave function $\psi_\alpha(\mathbf{r},t) = Ne^{i(\mathbf{p}\cdot\mathbf{r}-Et)/\hbar}$ is a solution to the free-particle Schrodinger equation. Determine the corresponding solution $\psi_{\alpha'}(\mathbf{r},t)$ under time reversal symmetry.

Solution: Under time reversal symmetry, both $\psi_\alpha(\mathbf{r},t)$ and $\psi_{\alpha'}(\mathbf{r},t) = T\psi_\alpha(\mathbf{r},-t)$ satisfy the free-particle Schrodinger equation

$$i\hbar\frac{\partial}{\partial t}\psi(\mathbf{r},t) = -\frac{\hbar^2}{2m}\nabla^2\psi(\mathbf{r},t)$$

$\psi_\alpha(\mathbf{r},-t) = Ne^{i(\mathbf{p}\cdot\mathbf{r}+Et)/\hbar}$ is not obviously a solution to the free-particle Schrodinger equation. However, $\psi^*_\alpha(\mathbf{r},-t) = Ne^{-i(\mathbf{p}\cdot\mathbf{r}+Et)/\hbar}$ is a solution to the above equation. Therefore,

$$\psi_{\alpha'}(\mathbf{r},t) = Ne^{-i(\mathbf{p}\cdot\mathbf{r}+Et)/\hbar} = Ne^{i((-\mathbf{p})\cdot\mathbf{r}-Et)/\hbar}$$

$\psi_\alpha(\mathbf{r},t)$ describes the state of the particle with momentum $\mathbf{p}$. $\psi_{\alpha'}(\mathbf{r},t)$ describes the state of the time reversed particle which now has a momentum $-\mathbf{p}$.Time reversal invariance of the Schrodinger equation implies that if it allows $\psi_\alpha(\mathbf{r},t)$ as a possible physical state, then a physical state corresponding to $\psi_{\alpha'}(\mathbf{r},t)$ should also exist.

9.20.3 T for Spinless Particle

For spinless particle, the time-independent Hamiltonian H is

$$H = \frac{p^2}{2m} + V(\mathbf{r})$$

$$
\begin{aligned}
THT^{-1} &= T\frac{p^2}{2m}T^{-1} + TV(\mathbf{r})T^{-1} \\
&= \frac{1}{2m}T\mathbf{p}T^{-1}T\mathbf{p}T^{-1} + TV(\mathbf{r})T^{-1} \\
&= \frac{1}{2m}(-\mathbf{p})\cdot(-\mathbf{p}) + V(\mathbf{r}) = \frac{p^2}{2m} + V(\boldsymbol{r}) = H
\end{aligned}
$$

∴ The Hamiltonian $H = -\hbar^2/2m\nabla^2 + V(\mathbf{r})$ is invariant under time reversal transformation. This implies that both $\psi_\alpha(\mathbf{r},t)$ and $\psi_{\alpha'}(\mathbf{r},t)$ satisfy the Schrödinger equation.

$$i\hbar\frac{\partial}{\partial t}\psi_\alpha(\mathbf{r},t) = \left[-\frac{\hbar^2}{2m}\nabla^2 + V(\mathbf{r})\right]\psi_\alpha(\mathbf{r},t) \tag{9.182}$$

$$i\hbar\frac{\partial}{\partial t}\psi_{\alpha'}(\mathbf{r},t) = \left[-\frac{\hbar^2}{2m}\nabla^2 + V(\mathbf{r})\right]\psi_{\alpha'}(\mathbf{r},t) \tag{9.183}$$

Change the variable t to $-t$ in the Equation (9.182).

$$-i\hbar\frac{\partial}{\partial t}\psi_\alpha(\mathbf{r},-t) = \left[-\frac{\hbar^2}{2m}\nabla^2 + V(\mathbf{r})\right]\psi_\alpha(\mathbf{r},-t)$$

Take the complex conjugate of this equation.

$$\therefore i\hbar\frac{\partial}{\partial t}\psi_\alpha^*(\mathbf{r},-t) = \left[-\frac{\hbar^2}{2m}\nabla^2 + V(\mathbf{r})\right]\psi_\alpha^*(\mathbf{r},-t) \tag{9.184}$$

Comparing the Equations (9.183) and (9.184), it is obvious that

$$\psi_{\alpha'}(\mathbf{r},t) = \psi_\alpha^*(\mathbf{r},-t) \tag{9.185}$$

$$= T\psi_\alpha(\mathbf{r},-t) = UK\psi_\alpha(\mathbf{r},-t) \tag{9.186}$$

provided $U = 1$.

∴For spinless particle $T = 1\cdot K = K$

9.20.4 Particles with Spin 1/2

For spineless particles, we evaluated T to be K. For particles with non-zero spin, T is not so simple. Let us evaluate T for spin 1/2 particle. With respect to time reversal, the states and the operators transform as

$$\psi_{\alpha'}(\mathbf{r},t) = T\psi_\alpha(\mathbf{r},-t) \tag{9.187}$$

$$TFT^{-1} = F' \tag{9.188}$$

For spin angular momentum, we have

$$T\mathbf{S}T^{-1} = -\mathbf{S} \tag{9.189}$$

In the next chapter, we will see that for spin 1/2, the spin angular momentum operators S_x, S_y and S_z are given by

$$S_x = \frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},\ S_y = \frac{\hbar}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \text{ and } S_z = \frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

Consider the equation

$$\mathbf{S}\chi = \chi'$$

Taking complex conjugate, we get

$$\mathbf{S}^* \chi^* = \chi'^*$$

This equation can be obtained using K as follows:

$$KS\chi = K\chi'$$

$$\underbrace{KSK^{-1}}_{S^*}\ \underbrace{K\chi}_{\chi^*} = \underbrace{K\chi'}_{\chi'^*}$$

Since $K\chi = \chi^*$ and $K\chi' = \chi'^*$, we get

$$K\mathbf{S}K^{-1} = \mathbf{S}^* \tag{9.190}$$

Let us rewrite the Equation (9.189) as

$$TS_xT^{-1} = -S_x,\ TS_yT^{-1} = -S_y,\ TS_zT^{-1} = -S_z$$

$$UKS_xK^{-1}U^\dagger = -S_x;\ UKS_yK^{-1}U^\dagger = -S_y;\ UKS_zK^{-1}U^\dagger = -S_z$$

Making use of (9.189), we get

$$US_xU^\dagger = -S_x \quad US_yU^\dagger = S_y \quad US_zU^\dagger = -S_z \tag{9.191}$$

These equations suggest U as the operator rotating S_x and S_z into $-S_x$ and $-S_z$, respectively while leaving S_y unrotated. Therefore, we can take U as rotation about y axis by an angle π. In fact, we can check that $U = e^{-i\pi S_y/\hbar}$ achieve these equations (see Example 13).

Therefore, for spin 1/2 particles, T is given by

$$T = e^{-\frac{i\pi S_y}{\hbar}} K \tag{9.192}$$

EXERCISES

1. Explain what is meant by Dirac's ket vector and bra vector.
2. How do you describe the state of the system in terms of Dirac's ket vector?
3. Can we write $|0\rangle = e^{-\alpha^2x^2/2}$ for harmonic oscillator?

4. Why do you call Dirac's ket vector as abstract ket vector?
5. How will you define the adjoint of an abstract operator $\hat{A}$?
6. What is Hermitian conjugate of $|\psi\rangle$?
7. Given a bra vector $|\psi\rangle$ and a bra vector $|\phi\rangle$, define (i) scalar product (ii) outer product
8. Explain what is meant by decomposition of identity operator.
9. What is meant by coordinate representation?
10. In $\hat{p}\,|\,x\rangle = i\hbar\, d/dx\,|\,x\rangle$., can we write $\hat{p} = i\hbar\, d/dx$?
11. Show that the momentum operator in the coordinate representation is $-i\hbar\, d/dx$.
12. What is the significance of unitary operators in quantum mechanics?
13. Under unitary transformation, $|\alpha\rangle$ and $|\beta\rangle$ become $\hat{U}\,|\,\alpha\rangle$ and $\hat{U}\,|\,\beta\rangle$. If $\hat{F}\,|\,\alpha\rangle = |\,\beta\rangle$, show that $\hat{F}$ becomes $\hat{U}\hat{F}\hat{U}^{\dagger}$.
14. Show that the matrix element $\langle\phi\,|\,F\,|\,\psi\rangle$ is invariant under unitary transformation of the state vectors and the operators.
15. Show that the norm of a state vector is invariant under unitary transformation.
16. Show that the invariance of the Hamiltonian under unitary transformation leads to a conservation of a physical quantity.
17. Define unitary transformation relating the states under parity transformation.
18. Distinguish between intrinsic and spatial parity of a particle.
19. Explain what is meant by $\theta - \tau$ puzzle. How is this puzzle resolved?
20. What is meant by anti-linear operator?
21. Why do we need time reversal symmetry operator to be anti-linear?

REFERENCES

1. P. A. M. Dirac, 1977. *The Principles of Quantum Mechanics*. London: Oxford University Press.
2. M. Alonso and H. Valk, 1973. *Quantum Mechanics Principles and Applications*. Massachusetts: Addison-Wesley Publishing Company.
3. I. Schiff, 1968. *Quantum Mechanics*. New Delhi: McGraw-Hill International Book Company.
4. P. M. Mathews and A. Venkatesan, 1976. *A Text Book of Quantum Mechanics*. New Delhi: Tata McGraw-Hill Publishing Company Ltd.
5. E. Merzbacher, 1970. *Quantum Mechanics*. New York: John-Wiley of Sons, Inc.
6. Robert Littlejohn, http://bohr.physics.berkely.edu/classes/221/1011/notes.parity. pdf.
7. Robert Littlejohn, http://bohr.physics,berkely.edu/classes/221/1011/notes.timesrev.pdf.

10

Angular Momentum and Rotation Symmetry

In Chapter 7, we studied the orbital angular momentum $\mathbf{L} = \mathbf{r} \times \mathbf{p}$ by considering its differential operators in coordinate representation. Here, we discuss one more kind of angular momentum called spin angular momentum, which does not emerge from $\mathbf{r} \times \mathbf{p}$. It arises from the study of angular momentum, using the abstract operator formalism. However, there is another motivation to discuss angular momentum operators in terms of abstract operators, which come from the consideration of rotational symmetry.

We have already seen that under rotational symmetry

$$\begin{aligned} &\mathbf{r} \to R\,\mathbf{r} \\ &|\alpha\rangle \to \hat{U}|\alpha\rangle = e^{-i\boldsymbol{\theta}\boldsymbol{\cdot}\mathbf{J}/\hbar}|\alpha\rangle \end{aligned}$$

where $\hat{J} = \mathbf{x}\hat{J}_x + \mathbf{y}\hat{J}_y + \mathbf{z}\hat{J}_z$. The angular momentum operators $\hat{J}_x, \hat{J}_y$ and $\hat{J}_z$ obey angular momentum algebra

$$[\hat{J}_x, \hat{J}_y] = i\hbar\hat{J}_z, \quad [\hat{J}_y, \hat{J}_z] = i\hbar\hat{J}_x, \quad [\hat{J}_z, \hat{J}_x] = i\hbar\hat{J}_y \tag{10.1}$$

In Chapter 7, this algebra was proved for orbital angular momentum operators. In this chapter, we take these commutation relations as the definition of angular momentum algebra. The operators $\hat{J}_x, \hat{J}_y$, and $\hat{J}_z$ and hence $\hat{U}$ are abstract operators. We can have many representations depending on the choice of the basis vectors. The coordinate representation of these abstract operators becomes the differential operators for angular momentum operators L_x, L_y and L_z given in Chapter 7. The coordinate representation $\langle \mathbf{r}|\hat{U}|\alpha\rangle$ is given by

$$\langle \mathbf{r}|\hat{U}|\boldsymbol{\alpha}\rangle = e^{-\frac{i}{\hbar}\boldsymbol{\theta}\boldsymbol{\cdot}\mathbf{L}_{op}}\psi_\alpha(\mathbf{r})$$

where $$\mathbf{L}_{\mathrm{op}} = \mathbf{r}_{\mathrm{op}} \times \mathbf{p}_{\mathrm{op}} = -i\hbar\,\mathbf{r} \times \nabla$$

In this chapter, we are interested in studying the consequences of the angular momentum algebra of the abstract operators $\hat{J}_x, \hat{J}_y$, and $\hat{J}_z$ and this analysis is more general and brings a new and rich physics, which could not be obtained in coordinate representation. The set of Equations (10.1) is known as the angular momentum algebra. This algebra plus the fact that they are Hermitian operators are sufficient to get a lot of results. A better representation for the angular momentum happens to be the matrix representation compared to the coordinate representation.

10.1 EIGENVALUES OF $\hat{J}^2$ AND $\hat{J}_z$

The first thing to be noted is that the three operators $\hat{J}_x, \hat{J}_y$ and $\hat{J}_z$ do not commute among themselves. There exists another operator $\hat{J}^2 = \hat{J}_x^2 + \hat{J}_y^2 + \hat{J}_z^2$ which commutes with all the three operators.

$$[\hat{J}^2, \hat{J}_x] = 0 \quad [\hat{J}^2, \hat{J}_y] = 0 \quad [\hat{J}^2, \hat{J}_z] = 0 \tag{10.2}$$

Among the four operators $\hat{J}^2, \hat{J}_x, \hat{J}_y$ and $\hat{J}_z$, we have to select two operators which commute among themselves. The standard choice is $\hat{J}^2$ and $\hat{J}_z$. Since $\hat{J}^2$ and $\hat{J}_z$ commute, they have common eigenstates. So we have to determine the eigenvalues and eigenvectors of $\hat{J}^2$ and $\hat{J}_z$.

Since $\hbar$ has the dimension of angular momentum, we write the eigenvalue equation of $\hat{J}^2$ and $\hat{J}_z$ as

$$\hat{J}^2 |\lambda m\rangle = \lambda \hbar^2 |\lambda m\rangle \tag{10.3}$$

$$\hat{J}_z |\lambda m\rangle = m\hbar |\lambda m\rangle \tag{10.4}$$

At present, we do not know the possible values of λ and m. To evaluate λ and m, let us do some algebraic manipulations among the angular momentum operators. Let us define $\hat{J}_+$ and $\hat{J}_-$ as

$$\hat{J}_+ = \hat{J}_x + i\hat{J}_y \qquad \hat{J}_- = \hat{J}_x - i\hat{J}_y \tag{10.5}$$

Let us evaluate the following commutation relations:

$$[\hat{J}_z, \hat{J}_+] = [\hat{J}_z, \hat{J}_x + i\hat{J}_y] = \hbar \hat{J}_+ \tag{10.6}$$

$$[\hat{J}_z, \hat{J}_-] = [\hat{J}_z, \hat{J}_x - i\hat{J}_y] = -\hbar \hat{J}_- \tag{10.7}$$

$$\begin{aligned} \hat{J}_+\hat{J}_- &= (\hat{J}_x + i\hat{J}_y)(\hat{J}_x - i\hat{J}_y) \\ &= \hat{J}_x^2 + \hat{J}_y^2 - i(\hat{J}_x\hat{J}_y - \hat{J}_y\hat{J}_x) \\ &= \hat{J}_x^2 + \hat{J}_y^2 + \hbar\hat{J}_z = \hat{J}^2 - \hat{J}_z^2 + \hbar\hat{J}_z \end{aligned} \tag{10.8}$$

$$\begin{aligned} \hat{J}_-\hat{J}_+ &= (\hat{J}_x - i\hat{J}_y)(\hat{J}_x + i\hat{J}_y) \\ &= \hat{J}_x^2 + \hat{J}_y^2 + i(\hat{J}_x\hat{J}_y - \hat{J}_y\hat{J}_x) \\ &= \hat{J}_x^2 + \hat{J}_y^2 - \hbar\hat{J}_z = \hat{J}^2 - \hat{J}_z^2 - \hbar\hat{J}_z \end{aligned} \tag{10.9}$$

With these results, let us now determine the eigenvalues λ and m.

$m^2 \le \lambda$

Let us consider $\langle \lambda m | \hat{J}^2 | \lambda m \rangle$ and $\langle \lambda m | J_x^2 | \lambda m \rangle$

$$\begin{aligned}\langle \lambda m | \hat{J}^2 | \lambda m \rangle &= \langle \lambda m | \lambda \hbar^2 | \lambda m \rangle \\ &= \lambda \hbar^2 \langle \lambda m | \lambda m \rangle = \lambda \hbar^2\end{aligned}$$

$$\begin{aligned}\langle \lambda m | \hat{J}_x^2 | \lambda m \rangle &= \underbrace{\langle \lambda m | \hat{J}_x}_{\langle \phi |} \underbrace{\hat{J}_x | \lambda m \rangle}_{| \phi \rangle} \\ &= \langle \phi | \phi \rangle \ge 0\end{aligned}$$

where
$$| \phi \rangle = \hat{J}_x | \lambda m \rangle$$

Since $\langle \phi | \phi \rangle \ge 0$, let us say $\qquad \langle \lambda m | \hat{J}_x^2 | \lambda m \rangle = a^2 \hbar^2$

Similarly, we can write $\qquad \langle \lambda m | \hat{J}_y^2 | \lambda m \rangle = b^2 \hbar^2$

$$\langle \lambda m | \hat{J}_z^2 | \lambda m \rangle = \langle \lambda m | \hat{J}_z \hat{J}_z | \lambda m \rangle = m^2 \hbar^2 \langle \lambda m | \lambda m \rangle = m^2 \hbar^2$$

$$\begin{aligned}\therefore \langle \lambda m | \hat{J}_x^2 + \hat{J}_y^2 + \hat{J}_z^2 | \lambda m \rangle &= \langle \lambda m | \hat{J}^2 | \lambda m \rangle \\ \hbar^2 a^2 + \hbar^2 b^2 + m^2 \hbar^2 &= \lambda \hbar^2\end{aligned}$$

Since $a^2 + b^2$ is a positive quantity, we have

$$m^2 \le \lambda \tag{10.10}$$

Therefore, for a given value of λ, the set of allowed m values is bounded.

Let m' and m'' be the maximum and minimum values of m.

$$m'' \le m \le m' \tag{10.11}$$

$\hat{J}_+ | \lambda m \rangle$

Consider $\qquad \hat{J}_z \hat{J}_+ | \lambda m \rangle$

Since $\qquad \hat{J}_z \hat{J}_+ - \hat{J}_+ \hat{J}_z = \hbar \hat{J}_+$, we have

$$\begin{aligned}\hat{J}_z \hat{J}_+ &= \hat{J}_+ \hat{J}_z + \hbar \hat{J}_+ \\ \therefore \quad \hat{J}_z \hat{J}_+ | \lambda m \rangle &= (\hat{J}_+ \hat{J}_z + \hbar \hat{J}_+) | \lambda m \rangle = (m+1) \hbar \hat{J}_+ | \lambda m \rangle\end{aligned}$$

So we have

$$\left.\begin{aligned}\hat{J}_z | \lambda m \rangle &= m \hbar | \lambda m \rangle \\ \hat{J}_z \underbrace{\hat{J}_+ | \lambda m \rangle} &= (m+1) \hbar \underbrace{\hat{J}_+ | \lambda m \rangle}\end{aligned}\right\} \tag{10.12}$$

The Equation (10.12) suggests that while $|\lambda m\rangle$ is an eigenket of $\hat{J}_z$ with eigenvalue $m\hbar$, the ket $\hat{J}_+|\lambda m\rangle$ is an eigenket of $\hat{J}_z$ with an eigenvalue $(m+1)\hbar$. Thus $\hat{J}_+$ acting on the eigenket $|\lambda m\rangle$ produces an eigenket, for which the eigenvalue of $\hat{J}_z$ increases by 1. Therefore, $\hat{J}_+$ is called a raising operator. Let us consider the following equations:

$$\hat{J}_z|\lambda\overline{m+1}\rangle = (m+1)\hbar|\lambda\overline{m+1}\rangle \tag{10.13}$$

$$\hat{J}_z\hat{J}_+|\lambda m\rangle = (m+1)\hbar\hat{J}_+|\lambda m\rangle \tag{10.14}$$

By comparing the Equations (10.13) and (10.14), we conclude $\hat{J}_+|\lambda m\rangle$ is either $|\lambda\overline{m+1}\rangle$ or some constant times $|\lambda\overline{m+1}\rangle$. We drop the bar over (m + 1), with the understanding that $|\lambda\, m+1\rangle$ refers to $|\lambda\overline{m+1}\rangle$.

$$\therefore \hat{J}_+|\lambda m\rangle = C_+|\lambda m+1\rangle$$

where C_+ is a constant, which is yet to be determined.

For $m = m'$, $\hat{J}_+|\lambda m'\rangle$ cannot be $C_+|\lambda m'+1\rangle$ since m' is the maximum value of m.

$$\therefore \hat{J}_+|\lambda m'\rangle = 0 \tag{10.15}$$

$\hat{J}_-|\lambda m\rangle$

We now consider the relation

$$\hat{J}_z\hat{J}_- - \hat{J}_-\hat{J}_z = -\hbar\hat{J}_-$$

$$\hat{J}_z\hat{J}_-|\lambda m\rangle = (\hat{J}_-\hat{J}_z - \hbar\hat{J}_-)|\lambda m\rangle = (m-1)\hbar\hat{J}_-|\lambda m\rangle \tag{10.16}$$

$\hat{J}_-|\lambda m\rangle$ is an eigenket of $\hat{J}_z$ with an eigenvalue of $(m-1)\hbar$. i.e., $\hat{J}_-$ acting on the eigenket $|\lambda m\rangle$ produces an eigenket for which the eigenvalue of $\hat{J}_z$ has decreased by 1. Therefore, $\hat{J}_-$ is called a lowering operator. Consider the following equations:

$$\hat{J}_z|\lambda m\rangle = m\hbar|\lambda m\rangle; \quad \hat{J}_z\hat{J}_-|\lambda m\rangle = (m-1)\hbar\hat{J}_-|\lambda m\rangle$$

These equations suggest

$$\hat{J}_-|\lambda m\rangle = C_-|\lambda\, m-1\rangle \tag{10.17}$$

Since the minimum value of m is m'', $\hat{J}_-|jm''\rangle$ cannot be $C_-|jm''-1\rangle$. Therefore,

$$\hat{J}_-|jm''\rangle = 0 \tag{10.18}$$

$\lambda = m'(m'+1)$

Making use of the Equation (10.15), we get

$$\hat{J}_-\hat{J}_+ \,|\lambda m'\rangle = 0$$

But
$$\hat{J}_-\hat{J}_+ = \hat{J}^2 - \hat{J}_z^2 - \hbar\hat{J}_z$$

$$\therefore \hat{J}_-\hat{J}_+ \,|\lambda m'\rangle = (\hat{J}^2 - \hat{J}_z^2 - \hbar\hat{J}_z)\,|\lambda m'\rangle = 0$$

$$\therefore (\lambda - m'^2 - m')\hbar^2 \,|\lambda m'\rangle = 0$$

or
$$\lambda = m'(m'+1) \tag{10.19}$$

$\lambda = m''(m''-1)$

Making use of the Equation (10.18), we get

$$\hat{J}_+\hat{J}_- \,|\lambda m''\rangle = 0$$

$$\hat{J}_+\hat{J}_- \,|\lambda m''\rangle = (\hat{J}^2 - \hat{J}_z^2 + \hbar\hat{J}_z^2)\,|\lambda m''\rangle = 0$$

$$(\lambda - m''^2 + m'')\hbar^2 \,|\lambda m'\rangle = 0$$

$$\therefore \lambda = m''(m''-1) \tag{10.20}$$

From (10.19) and (10.20), we have

$$\therefore m'(m'+1) = m''(m''-1)$$

$$m'^2 + m' - m''^2 + m'' = 0$$

$$(m'+m'')(m'-m''+1) = 0$$

$$m'' = -m' \quad \text{or} \quad m'' = m'+1$$

The second solution is ruled out since the maximum value of m is m'. So the only possible solution is

$$m'' = -m'$$

Let us call m' as j

$$\therefore m' = j \qquad m'' = -j$$

Therefore, the possible value of m varies from $-j$ to j.

i.e.,
$$-j \le m \le j$$

From (10.19), we write
$$\lambda = j(j+1).$$

Therefore, we can write $|\lambda m\rangle$ as $|jm\rangle$. Then, we have

$$\hat{J}_z\,|jj\rangle = j\hbar\,|jj\rangle$$
$$\hat{J}_z\,|j-j\rangle = -j\hbar\,|j-j\rangle$$
$$\hat{J}_+\,|jj\rangle = \hat{J}_-\,|j-j\rangle = 0$$

Consider the effect of $\hat{J}_z\hat{J}_-$ on $|jj\rangle$.

$$\hat{J}_z\hat{J}_-\,|jj\rangle = (j-1)\,\hbar\hat{J}_-\,|jj\rangle$$

$\therefore \hat{J}_-\,|jj\rangle$ is an eigenket $\hat{J}_z$ with an eigenvalue $(j-1)\hbar$. By repeatedly using $\hat{J}_-$, we can get other eigenvalues which are given by

$$j\hbar, (j-1)\hbar, (j-2)\hbar, \cdots - j\hbar$$

So the possible values of m are as given below:

$$j,(j-1),(j-2)\cdots - j$$

There are $(2j+1)$ terms in the sequence. This can be seen in many ways. For instance, the above sequence is an arithmetic progression with initial term $a = j$ and common difference $d = -1$. If n is the number of terms in this sequence

$$j+(n-1)(-2) = -j$$

or
$$n = (2j+1) \tag{10.21}$$

Note that $(2j+1)$ is an integer. This is possible provided $j = 0, 1/2, 1, 3/2, \ldots$ i.e., j can be either integers or half-integers. This result has to be compared with the eigenvalues for orbital angular momentum. The allowed values of j are only set of integers; i.e., $j = 0, 1, 2, \ldots$

$$\hat{J}^2\,|jm\rangle = j(j+1)\hbar^2\,|jm\rangle \quad j = 0, \frac{1}{2}, 1, \frac{3}{2}, \ldots \tag{10.22}$$

$$\hat{J}_z\,|jm\rangle = m\hbar\,|jm\rangle \quad m = j, j-1, \ldots - j \tag{10.23}$$

C_+ and C_-

Let us determine the constants C_+ and C_- now.

$$\hat{J}_+\,|jm\rangle = C_+|jm+1\rangle \text{ and } \langle jm|\hat{J}_- = \langle jm+1|\,C_+^*$$

$$\therefore \langle jm|\,\hat{J}_-\hat{J}_+\,|jm\rangle = |C_+|^2\,\langle j\ m+1|j\ m+1\rangle = |C_+|^2$$

But
$$\begin{aligned}\langle jm|\hat{J}_-\hat{J}_+\,|jm\rangle &= \langle jm|\,(\hat{J}^2 - \hat{J}_z^2 - \hbar\hat{J}_z)\,|jm\rangle\\ &= (j-m)(j+m+1)\hbar^2\langle jm|jm\rangle = (j-m)(j+m+1)\hbar^2\\ \therefore |C_+|^2 &= (j-m)(j+m+1)\hbar^2\end{aligned}$$

By convention, both C_+ and C_- are chosen as a positive real number.

$$\therefore C_+ = [(j-m)(j+m+1)]^{1/2}\hbar \qquad (10.24)$$

In the same way, C_- can be evaluated as

$$C_- = [(j+m)(j-m+1)]^{1/2}\hbar \qquad (10.25)$$

So we have

$$\hat{J}^2\,|jm\rangle = j(j+1)\hbar^2\,|jm\rangle$$

$$\hat{J}_z\,|jm\rangle = m\hbar\,|jm\rangle$$

$$\hat{J}_+\,|jm\rangle = [(j-m)(j+m+1)]^{1/2}\hbar\,|\,jm+1\rangle$$

$$\hat{J}_-\,|jm\rangle = [(j+m)(j-m+1)]^{1/2}\hbar\,|\,jm-1\rangle$$

Using these relations, we get

$$\begin{aligned}\hat{J}_x\,|jm\rangle &= \frac{1}{2}(\hat{J}_+ + \hat{J}_-)\,|jm\rangle\\ &= \frac{1}{2}\begin{Bmatrix}[(j-m)(j+m+1)]^{1/2}\hbar\,|\,jm+1\rangle + \\ [(j+m)(j-m+1)]^{1/2}\hbar\,|\,jm-1\rangle\end{Bmatrix}\end{aligned}$$

$$\hat{J}_y\,|jm\rangle = \frac{1}{2i}\begin{Bmatrix}[(j-m)(j+m+1)]^{1/2}\hbar\,|\,jm+1\rangle \\ -[(j+m)(j-m+1)]^{1/2}\hbar\,|\,jm-1\rangle\end{Bmatrix}$$

10.2 MATRIX REPRESENTATION OF ANGULAR MOMENTUM OPERATORS

Let us recall that a general vector $|\psi\rangle$ can be expressed as a linear combination of the basis vectors. The eigenkets of the pair $\hat{J}^2$ and $\hat{J}_z$ form a complete set, and so they can be considered as basis vectors. So we have

$$|\psi\rangle = C_1\,|\,jj\rangle + C_2\,|\,jj-1\rangle + \ldots C_n\,|\,j-j\rangle$$

In matrix representation, ψ is given by

$$\psi = \begin{bmatrix} C_1 \\ C_2 \\ \vdots \end{bmatrix}$$

Now $|\,jj\rangle$ can be written as

$$|\,jj\rangle = |\,jj\rangle + 0.|\,jj-1\rangle + 0.|\,jj-2\rangle ...$$

$$\therefore \psi_{jj} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \end{bmatrix} \tag{10.26}$$

Similarly, we have

$$\therefore \psi_{j,j-1} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \end{bmatrix}, \psi_{j,j-2} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ \vdots \end{bmatrix} ... \tag{10.27}$$

There are $(2j + 1)$ column vectors.

The matrix representation of any operator $\hat{F}$ can be obtained as follows:

$$F = \begin{array}{cc} & \begin{array}{ccc} |jj\rangle & \quad |jj-1\rangle & \quad |jj-2\rangle \end{array} \\ \begin{array}{c} \langle jj| \\ \langle jj-1| \\ \end{array} & \begin{bmatrix} \langle jj|\hat{F}|jj\rangle & \langle jj|\hat{F}|jj-1\rangle & jj|\hat{F}|jj-2\rangle ... \\ \langle jj-1|\hat{F}|jj\rangle & \langle jj-1|\hat{F}|jj-1\rangle & \langle jj-1|\hat{F}|jj-2\rangle ... \\ \vdots & \vdots & \vdots \end{bmatrix} \end{array}$$

$\hat{F}$ may be $\hat{J}^2$ or $\hat{J}_x$ or $\hat{J}_y$ or $\hat{J}_z$ or any other combination of these operators.

In particular, $\hat{J}^2$ and $\hat{J}_z$ are diagonal matrices. The elements $\hat{J}^2$ and $\hat{J}_z$ are given by

$$\langle jm'|\,\hat{J}^2\,|jm\rangle = j(j+1)\hbar^2\delta_{mm'}$$

$$\langle jm'|\,\hat{J}_z\,|jm\rangle = m\hbar\delta_{mm'}$$

$$\begin{bmatrix} j(j+1)\hbar^2 & & \cdots & \cdots & \cdots & \cdots \\ & j(j+1)\hbar^2 & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & j(j+1)\hbar^2 & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & j(j+1)\hbar^2 \end{bmatrix} \tag{10.28}$$

$$= j(j+1)\hbar^2 \begin{bmatrix} 1 & 0 & 0 & \cdots & \cdots & \cdots \\ 0 & 1 & 0 & \cdots & \cdots & \cdots \\ 0 & 0 & 1 & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & 1 \end{bmatrix}$$

$$J_z = \begin{bmatrix} j\hbar & 0 & 0 & \cdots \\ 0 & (j-1)\hbar & 0 & \cdots \\ 0 & 0 & \left(j-2\right)\hbar & \cdots \\ \vdots & \vdots & \vdots & -j\hbar \end{bmatrix} \tag{10.29}$$

Example 10.1 Determine the matrices J^2, J_x, J_y and J_z for $j = 1$.

Solution: For $j = 1$, these matrices have got the dimension $2j + 1 = 3$. i.e., these matrices are 3×3 matrices.

$$J^2 = \begin{array}{c} \\ |11\rangle \\ |10\rangle \\ |1-1\rangle \end{array} \begin{array}{c} \begin{array}{ccc} |11\rangle & |10\rangle & |1-1\rangle \end{array} \\ \begin{bmatrix} |11\rangle \hat{J}^2 |11\rangle & |11\rangle \hat{J}^2 |10\rangle & |11\rangle \hat{J}^2 |1-1\rangle \\ |10\rangle \hat{J}^2 |11\rangle & |10\rangle \hat{J}^2 |10\rangle & |10\rangle \hat{J}^2 |1-1\rangle \\ |1-1\rangle \hat{J}^2 |11\rangle & |1-1\rangle \hat{J}^2 |10\rangle & |1-1\rangle \hat{J}^2 |1-1\rangle \end{bmatrix} \end{array}$$

$$= \begin{bmatrix} 2\hbar^2 & 0 & 0 \\ 0 & 2\hbar^2 & 0 \\ 0 & 0 & 2\hbar^2 \end{bmatrix}$$

$$J_z = \begin{bmatrix} |11\rangle \hat{J}_z |11\rangle & |11\rangle \hat{J}_z |10\rangle & |11\rangle \hat{J}_z |1-1\rangle \\ |10\rangle \hat{J}_z |11\rangle & |10\rangle \hat{J}_z |10\rangle & |10\rangle \hat{J}_z |1-1\rangle \\ |1-1\rangle \hat{J}_z |11\rangle & |1-1\rangle \hat{J}_z |10\rangle & |1-1\rangle \hat{J}_z |1-1\rangle \end{bmatrix}$$

$$= \begin{bmatrix} \hbar & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -\hbar \end{bmatrix}$$

$$J_+ = \begin{bmatrix} |11\rangle \hat{J}_+ |11\rangle & |11\rangle \hat{J}_+ |10\rangle & |11\rangle \hat{J}_+ |1-1\rangle \\ |10\rangle \hat{J}_+ |11\rangle & |10\rangle \hat{J}_+ |10\rangle & |10\rangle \hat{J}_+ |1-1\rangle \\ |1-1\rangle \hat{J}_+ |11\rangle & |1-1\rangle \hat{J}_+ |10\rangle & |1-1\rangle \hat{J}_+ |1-1\rangle \end{bmatrix}$$

$$= \begin{bmatrix} 0 & \sqrt{2}\hbar & 0 \\ 0 & 0 & \sqrt{2}\hbar \\ 0 & 0 & 0 \end{bmatrix}$$

Similarly, J_- is found to be

$$J_- = \begin{bmatrix} 0 & 0 & 0 \\ \sqrt{2}\hbar & 0 & 0 \\ 0 & \sqrt{2}\hbar & 0 \end{bmatrix}$$

$$J_x = \frac{1}{2}(J_+ + J_-) = \frac{\hbar}{\sqrt{2}} \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

$$J_y = \frac{1}{2i}(J_+ - J_-) = \frac{\hbar}{\sqrt{2}} \begin{bmatrix} 0 & -i & 0 \\ i & 0 & -i \\ 0 & i & 0 \end{bmatrix}$$

$$J_x = \frac{\hbar}{\sqrt{2}} \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, \quad J_y = \frac{\hbar}{\sqrt{2}} \begin{bmatrix} 0 & -i & 0 \\ i & 0 & -i \\ 0 & i & 0 \end{bmatrix} \quad J_z = \begin{bmatrix} \hbar & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -\hbar \end{bmatrix}$$

10.3 SPIN ANGULAR MOMENTUM

$\hat{J}_x, \hat{J}_y, \hat{J}_z$ and $\hat{J}^2$ are abstract operators, and the ket vector $|\, jm\rangle$ is an abstract state vector. We can choose different sets of basis vectors to get different representations. We have already seen that if one chooses the position vector $|\mathbf{r}\rangle$ as the basis vector, then $\langle \mathbf{r} \,|\, jm\rangle$ becomes $\langle \theta, \phi \,|\, lm\rangle$, which are the spherical harmonies.

$$\langle \mathbf{r} | jm\rangle = \langle \theta, \phi | lm\rangle = Y_{lm}(\theta, \phi)$$

On the other hand, we can choose the matrix representation. The operators J_x, J_y, J_z and J^2 are now independent of spatial coordinates.

$$J^2 \psi_{jm} = j(j+1)\hbar^2 \psi_{jm} \qquad j = 0, \frac{1}{2}, 1, \frac{3}{2}, \ldots$$
$$J_z \psi_{jm} = m\hbar \psi_{jm}$$

J^2 and J_z are now matrix operators, and ψ_{jm} is a column vector. They can be now taken to represent a new kind of angular momentum called spin angular momentum. Spin angular momentum does not emerge from $\mathbf{r} \times \mathbf{p}$. We cannot find a classical dynamical variable $A(\mathbf{r}, \mathbf{p})$ representing angular momentum such that the eigenvalues of the corresponding quantum mechanical operator are half integers. So there is no classical analogue to spin angular momentum. It is an intrinsic property of a particle since it cannot be removed or changed. For the sake of clarity, let us list all the possible quantum numbers of orbital and spin angular momenta.

$$l \,:\, 0, 1, 2, 3, \ldots$$
$$s \,:\, 0, \frac{1}{2}, 1, \frac{3}{2}, 2, \ldots$$

Since spin angular momentum operator is a matrix operator, it commutes with many dynamical variables, including Hamiltonian, provided it contains only differential operators.

10.4 SPIN 1/2

Let us now consider particles with spin quantum number $s = 1/2$.

For spin 1/2 particle, there are two states: $\left|\frac{1}{2}\frac{1}{2}\right\rangle$ and $\left|\frac{1}{2}-\frac{1}{2}\right\rangle$

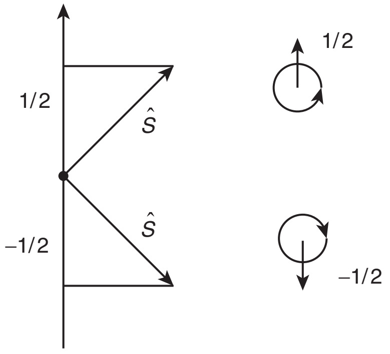
$$\hat{S}^2\left|\frac{1}{2}\frac{1}{2}\right\rangle = \frac{1}{2}\left(\frac{1}{2}+1\right)\hbar^2\left|\frac{1}{2}\frac{1}{2}\right\rangle$$

$$\hat{S}_z\left|\frac{1}{2}\frac{1}{2}\right\rangle = \frac{1}{2}\hbar\left|\frac{1}{2}\frac{1}{2}\right\rangle$$

$$\hat{S}^2\left|\frac{1}{2}-\frac{1}{2}\right\rangle = \frac{1}{2}\left(\frac{1}{2}+1\right)\hbar^2\left|\frac{1}{2}-\frac{1}{2}\right\rangle$$

$$\hat{S}_z\left|\frac{1}{2}-\frac{1}{2}\right\rangle = -\frac{1}{2}\hbar\left|\frac{1}{2}-\frac{1}{2}\right\rangle$$

Pictorially, these two states $\left|\frac{1}{2}\frac{1}{2}\right\rangle$ and $\left|\frac{1}{2}-\frac{1}{2}\right\rangle$ can be described as shown in Fig.10.1.



(a) (b)

Fig. 10.1 Representation of spin = 1/2 states

Many times, especially in popular level books, $m_s = 1/2$ and $m_s = -1/2$ states of an election are visualized as given in Fig.10.1(b). This picture is misleading since it presents electron as spinning object like spinning top or spinning earth. Quantum mechanics does not suggest any such picture. Quantum mechanics only says that an electron has two distinct states which are characterized as $\left|\frac{1}{2}\frac{1}{2}\right\rangle$ and $\left|\frac{1}{2}-\frac{1}{2}\right\rangle$.

However, visualization of $\text{spin} = 1/2$ states as in Fig.10.1(b) is very much helpful. In fact, many times, the states $\left|\frac{1}{2}\frac{1}{2}\right\rangle$ and $\left|\frac{1}{2}-\frac{1}{2}\right\rangle$ are simply written as $|\uparrow\rangle$ and $|\downarrow\rangle$.

i.e.,

$$|\uparrow\rangle = \left|\frac{1}{2}\frac{1}{2}\right\rangle$$

$$|\downarrow\rangle = \left|\frac{1}{2}-\frac{1}{2}\right\rangle$$

The matrix representation of the state vectors and the abstract operators can be easily determined. The state vector $\left|\frac{1}{2}\frac{1}{2}\right\rangle$ can be written as

$$\left|\frac{1}{2}\frac{1}{2}\right\rangle = 1\cdot\left|\frac{1}{2}\frac{1}{2}\right\rangle + 0\cdot\left|\frac{1}{2}-\frac{1}{2}\right\rangle$$

$$\therefore \psi_{\frac{1}{2}\frac{1}{2}} = \psi(\uparrow) = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \tag{10.30}$$

$$\left|\frac{1}{2}-\frac{1}{2}\right\rangle = 0\cdot\left|\frac{1}{2}\frac{1}{2}\right\rangle + 1\cdot\left|\frac{1}{2}-\frac{1}{2}\right\rangle$$

$$\therefore \psi_{\frac{1}{2}-\frac{1}{2}} = \psi(\downarrow) = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{10.31}$$

The matrix representation for various operators is given by

$$\begin{array}{c|cc|} & \left|\frac{1}{2}\frac{1}{2}\right\rangle & \left|\frac{1}{2}-\frac{1}{2}\right\rangle \\ \hline \left\langle\frac{1}{2}\frac{1}{2}\right| & & \\ \left\langle\frac{1}{2}-\frac{1}{2}\right| & & \\ \hline \end{array}$$

The matrix operator S^2 is given by

$$S^2 = \begin{bmatrix} \left\langle\frac{1}{2}\frac{1}{2}\right|\hat{S}^2\left|\frac{1}{2}\frac{1}{2}\right\rangle & \left\langle\frac{1}{2}\frac{1}{2}\right|\hat{S}^2\left|\frac{1}{2}-\frac{1}{2}\right\rangle \\ \left\langle\frac{1}{2}-\frac{1}{2}\right|\hat{S}^2\left|\frac{1}{2}\frac{1}{2}\right\rangle & \left\langle\frac{1}{2}-\frac{1}{2}\right|\hat{S}^2\left|\frac{1}{2}-\frac{1}{2}\right\rangle \end{bmatrix}$$

$$= \begin{bmatrix} s(s+1)\hbar^2 & 0 \\ 0 & s(s+1)\hbar^2 \end{bmatrix} = \frac{3}{4}\hbar^2\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

In the same way, we can work out S_+ and S_-, which are found to be

$$S_+ = \hbar\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad S_- = \hbar\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

S_x and S_y are determined from S_+ and S_-.

$$S_x = \frac{1}{2}(S_+ + S_-) \text{ and } S_y = \frac{1}{2i}(S_+ - S_-)$$

The matrices S_x, S_y and S_z are given by

$$S_x = \frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad S_y = \frac{\hbar}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad S_z = \frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{10.32}$$

Eigenvectors and Eigenvalues of S_x, S_y and S_z

It is easy to work out the eigenvalues and eigenvectors of these operators. The eigenvalues of each of these operators are ± $\hbar/2$.

Operators	Eigenvalue	Eigenvectors
S_x	$\hbar/2$	$\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}$
	$-\hbar/2$	$\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix}$
S_y	$\hbar/2$	$\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ i \end{pmatrix}$
	$-\hbar/2$	$\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -i \end{pmatrix}$
S_z	$\hbar/2$	$\begin{pmatrix} 1 \\ 0 \end{pmatrix}$
	$-\hbar/2$	$\begin{pmatrix} 0 \\ 1 \end{pmatrix}$

Example 10.2 Determine the eigenvalues and eigenvectors of S_x.

Solution:

$$\begin{aligned} S_x\chi &= \lambda\chi \\ \frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} &= \lambda\begin{pmatrix} a \\ b \end{pmatrix} \\ \begin{vmatrix} -\lambda & \hbar/2 \\ \hbar/2 & -\lambda \end{vmatrix} &= 0 \end{aligned}$$

$$\therefore \lambda^2 = \hbar^2/4 \;\text{ or }\; \lambda = \pm\hbar/2$$

$$\frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} = \pm\frac{\hbar}{2}\begin{pmatrix} a \\ b \end{pmatrix}$$

Solving this equation, we get the eigenvector to be $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$. The normalized eigenvectors corresponding to the eigenvalues $m_s = \pm\hbar/2$ are

$$m_s = \frac{\hbar}{2}: \quad \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

$$m_s = -\frac{\hbar}{2}: \quad \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

10.5 PAULI MATRICES

Let us define the three Pauli matrices σ_x, σ_y and σ_z as

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{10.33}$$

The spin matrices S_x, S_y and S_z are now given by

$$S_x = \frac{\hbar}{2}\sigma_x, \quad S_y = \frac{\hbar}{2}\sigma_y \quad \text{and} \quad S_z = \frac{\hbar}{2}\sigma_z \tag{10.34}$$

It is a customary practice to write the spin angular momentum operator as

$$\mathbf{S} = \frac{\hbar}{2}\boldsymbol{\sigma} \tag{10.35}$$

Though σ_x, σ_y, and σ_z are matrices, they are written with arrows like a geometrical vector. We can thus write $\boldsymbol{\sigma}$ as

$$\sigma = \hat{\boldsymbol{x}}\sigma_x + \hat{\boldsymbol{y}}\sigma_y + \hat{\boldsymbol{z}}\sigma_z \tag{10.36}$$

Properties of Pauli Matrices

1. σ_x, σ_y, and σ_z, are Hermitian matrices. It is easy to see that

$$\sigma_x^\dagger = \sigma_x, \quad \sigma_y^\dagger = \sigma_y \quad \text{and} \; \sigma_z^\dagger = \sigma_z$$

2. They are anti-commuting set of matrices. $\sigma_x\sigma_y + \sigma_y\sigma_x = 0$

$$\sigma_x\sigma_z + \sigma_z\sigma_x = 0$$

$$\sigma_y\sigma_z + \sigma_z\sigma_y = 0$$

3. $$\sigma_x^2 = I,\ \sigma_y^2 = I,\ \sigma_z^2 = I$$

i.e., $$\sigma_x^2 = \sigma_y^2 = \sigma_z^2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

4. $$\sigma_x\sigma_y = i\sigma_z;\quad \sigma_y\sigma_z = i\sigma_x;\quad \sigma_z\sigma_x = i\sigma_y$$

$\boldsymbol{\sigma}\cdot\mathbf{A}$:

We can have a scalar product $\boldsymbol{\sigma}$ with a geometrical vector $\mathbf{A}$.

$$\begin{aligned}
\boldsymbol{\sigma}\cdot\mathbf{A} &= (\hat{\boldsymbol{x}}\sigma_x + \hat{\boldsymbol{y}}\sigma_y + \hat{\boldsymbol{z}}\sigma_z)(\hat{\boldsymbol{x}}A_x + \hat{\boldsymbol{y}}A_y + \hat{\boldsymbol{z}}A_z) \\
&= \sigma_x A_x + \sigma_y A_y + \sigma_z A_z \\
&= \begin{pmatrix} 0 & A_x \\ A_x & 0 \end{pmatrix} + \begin{pmatrix} 0 & -iA_y \\ iA_y & 0 \end{pmatrix} + \begin{pmatrix} A_z & 0 \\ 0 & -A_z \end{pmatrix} \\
&= \begin{pmatrix} A_z & A_x - iA_y \\ A_y + iA_y & -A_z \end{pmatrix}
\end{aligned}$$

Example 10.3 Show that $(\boldsymbol{\sigma}\cdot\mathbf{A})(\boldsymbol{\sigma}\cdot\mathbf{B}) = \mathbf{A}\cdot\mathbf{B} + i\boldsymbol{\sigma}\cdot(\mathbf{A}\times\mathbf{B})$.
where $\mathbf{A}$ and $\mathbf{B}$ are geometrical vectors.

Solution:

$$\begin{aligned}
\mathbf{A} &= \hat{\boldsymbol{x}}A_x + \hat{\boldsymbol{y}}A_y + \hat{\boldsymbol{z}}A_z \\
\mathbf{B} &= \hat{\boldsymbol{x}}B_x + \hat{\boldsymbol{y}}B_y + \hat{\boldsymbol{z}}B_z
\end{aligned}$$

$$\begin{aligned}
\boldsymbol{\sigma}\cdot\mathbf{A}\,\boldsymbol{\sigma}\cdot\mathbf{B} &= (\sigma_x A_x + \sigma_y A_y + \sigma_z A_z)(\sigma_x B_x + \sigma_y B_y + \sigma_z B_z) \\
&= \sigma_x^2 A_x B_x + \sigma_x\sigma_y A_x B_y + \sigma_x\sigma_z A_x B_z + \sigma_y\sigma_x A_y B_x + \sigma_y^2 A_y B_y \\
&\quad + \sigma_y\sigma_z A_y B_z + \sigma_z\sigma_x A_z B_x + \sigma_z\sigma_y A_z B_y + \sigma_z^2 A_z B_z \\
&= (A_x B_x + A_y B_z + A_z B_z)I + i\sigma_x(A_y B_z - A_z B_y) \\
&\quad + i\sigma_y(A_z B_x - A_x B_z) + i\sigma_z(A_x B_y - A_y B_x)
\end{aligned}$$

Here, we have used the above property (4).

$$\therefore \boldsymbol{\sigma}\cdot\mathbf{A}\,\boldsymbol{\sigma}\cdot\mathbf{B} = \mathbf{A}\cdot\mathbf{B}I + i\boldsymbol{\sigma}\cdot(\mathbf{A}\times\mathbf{B})$$

Normally, the unit matrix *I* is not explicitly written.

$$\therefore (\boldsymbol{\sigma}\cdot\mathbf{A})(\boldsymbol{\sigma}\cdot\mathbf{B}) = \mathbf{A}\cdot\mathbf{B} + i\boldsymbol{\sigma}\cdot(\mathbf{A}\times\mathbf{B})$$

10.6 EIGENVALUES AND EIGENVECTOR OF S·n

Let **n** be an arbitrary unit vector given by $\mathbf{n} = \hat{\boldsymbol{x}}n_x + \hat{\boldsymbol{y}}n_y + \hat{\boldsymbol{z}}n_z$.

$\mathbf{S}\cdot\mathbf{n}$ gives the projection of the spin in the direction of **n**.

The eigenvalue equation for $\mathbf{S}\cdot\mathbf{n}$ is

$$\mathbf{S}\cdot\mathbf{n}\chi = \lambda\chi$$

Multiply the above equation by $\mathbf{S}\cdot\mathbf{n}$ from the left.

$$\mathbf{S}\cdot\mathbf{n}\cdot\mathbf{S}\cdot\mathbf{n}\chi = \lambda\underbrace{\mathbf{S}\cdot\mathbf{n}\chi}_{\lambda\chi} = \lambda^2\chi$$

But

$$\mathbf{S}\cdot\mathbf{n}\mathbf{S}\cdot\mathbf{n} = \frac{\hbar^2}{4}\boldsymbol{\sigma}\cdot\mathbf{n}\boldsymbol{\sigma}\cdot\mathbf{n}$$

$$= \frac{\hbar^2}{4}[\mathbf{n}\cdot\mathbf{n}I + i\boldsymbol{\sigma}\cdot\mathbf{n}\times\mathbf{n}] = \frac{\hbar^2}{4}I$$

$$\therefore \frac{\hbar^2}{4}I\chi = \lambda^2\chi$$

$$\lambda^2 = \frac{\hbar^2}{4} \quad \text{or} \quad \lambda = \pm\hbar/2$$

Obviously, the projection of **S** along any arbitrary vector **n** is either ℏ/2 or – ℏ/2.

To determine the eigenvector, let us write

$$\mathbf{S}\cdot\mathbf{n} = \frac{\hbar}{2}\boldsymbol{\sigma}\cdot\mathbf{n} = \frac{\hbar}{2}\begin{bmatrix} n_z & n_x - in_y \\ n_z + in_y & -n_z \end{bmatrix}$$

$$\frac{\hbar}{2}\begin{bmatrix} n_z & n_x - in_y \\ n_z + in_y & -n_z \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \pm\frac{\hbar}{2}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

It is a simple exercise to solve this equation. The eigenvectors are given below:

$$\frac{\hbar}{2}: \quad \psi(\uparrow) = \left(\frac{1+n_z}{2}\right)^{1/2}\begin{pmatrix} 1 \\ \dfrac{n_x + in_y}{(1+n_z)} \end{pmatrix}$$

$$-\frac{\hbar}{2}: \quad \psi(\downarrow) = \left(\frac{1+n_z}{2}\right)^{1/2}\begin{pmatrix} -\dfrac{n_x - in_y}{(1+n_z)} \\ 1 \end{pmatrix}$$

In spherical polar coordinates, we have

$$n_x = \sin\theta\cos\phi,\; n_y = \sin\theta\sin\phi \text{ and } n_z = \cos\theta.$$

$$\left(\frac{1+n_z}{2}\right)^{1/2} = \cos\frac{\theta}{2}; \quad \left(\frac{1-n_z}{2}\right)^{1/2} = \sin\frac{\theta}{2}$$

$$n_x + in_y = \sin\theta e^{i\phi}$$

Making use of these results, we get the following eigenvectors (leaving aside a phase factor):

$$\psi(\uparrow) = \begin{bmatrix} \cos\dfrac{\theta}{2} \\ \sin\dfrac{\theta}{2}e^{i\phi} \end{bmatrix}$$

$$\psi(\downarrow) = \begin{bmatrix} \sin\dfrac{\theta}{2} \\ -\cos\dfrac{\theta}{2}e^{i\phi} \end{bmatrix}$$

10.7 ROTATION MATRIX FOR SPIN 1/2 PARTICLE

We have already seen that if a system is rotated by an angle θ about **n**, then the transformation of the wave function ψ_α is given by

$$\psi_{\alpha'} = e^{i\boldsymbol{\theta}\cdot\mathbf{J}/\hbar}\psi_\alpha \tag{10.37}$$

In the case of spin $\frac{1}{2}$ particle, this equation becomes

$$\psi_{\alpha'} = e^{-i\theta\,\mathbf{n}\cdot\mathbf{S}/\hbar}\psi_\alpha \tag{10.38}$$

Note that both $\psi_{\alpha'}$ and ψ_α are column vectors.

i.e., $$\psi_{\alpha'} = \begin{pmatrix} a' \\ b' \end{pmatrix} \quad \text{and} \quad \psi_\alpha = \begin{pmatrix} a \\ b \end{pmatrix}$$

For the sake of simplicity, consider the rotation about Z axis by an angle θ.

$$\psi_{\alpha'} = e^{-i\theta\, S_z/\hbar}\psi_\alpha = e^{-i\theta\,\sigma_z/2}\psi_\alpha \tag{10.39}$$

It can be shown that

$$e^{-i\theta\sigma_z/2} = I\cos\frac{\theta}{2} + \sigma_z\sin\frac{\theta}{2} \tag{10.40}$$

$$\therefore \psi_{\alpha'} = \left(I\cos\frac{\theta}{2} + \sigma_z\sin\frac{\theta}{2}\right)\psi_\alpha \tag{10.41}$$

Let us now rotate the system by 360° about Z axis. Physically, we will get back the system to the original orientation. So we expect the wave function of the rotated system also to return to the original wave function. However, for spin 1/2 particle, this does not happen.

$$\psi_{\alpha'} = e^{-i2\pi S_z/\hbar}\psi_\alpha = (I\cos\pi + \sigma_z\sin\pi)\psi_\alpha = -\psi_\alpha \tag{10.42}$$

Only a rotation of the system by another 360° brings the function to be the same. Physically, this property has no observable consequence, since wave functions which differ only by a phase factor describe the same state.

Example 10.4 Show that $e^{-i\theta\mathbf{n}\cdot\mathbf{S}/\hbar} = I\cos\frac{\theta}{2} - \boldsymbol{\sigma}\cdot\mathbf{n}\sin\frac{\theta}{2}$, where S is the spin angular momentum operator for spin $\frac{1}{2}$.

Solution:

$$\text{Since } \mathbf{S} = \frac{\hbar}{2}\boldsymbol{\sigma}, \text{ we have } e^{-i\theta\mathbf{n}\cdot\mathbf{S}/\hbar} = e^{-i\theta\mathbf{n}\cdot\boldsymbol{\sigma}/2}$$

$$e^{-i\theta\mathbf{n}\cdot\boldsymbol{\sigma}/2} = 1 - i\frac{i\theta}{2}\boldsymbol{\sigma}\cdot\mathbf{n} + \frac{\left(-i\frac{\theta}{2}\right)^2}{2!}(\boldsymbol{\sigma}\cdot\mathbf{n})^2 + \frac{\left(-i\frac{\theta}{2}\right)^3}{3!}(\boldsymbol{\sigma}\cdot\mathbf{n})^3 + \frac{\left(-i\frac{\theta}{2}\right)^4}{4!}(\boldsymbol{\sigma}\cdot\mathbf{n})^4 + \ldots$$

$$\boldsymbol{\sigma}\cdot\mathbf{n}^2 = (\boldsymbol{\sigma}\cdot\mathbf{n})(\boldsymbol{\sigma}\cdot\mathbf{n}) = n^2 = 1$$

$$\therefore (\boldsymbol{\sigma}\cdot\mathbf{n})^4 = (\boldsymbol{\sigma}\cdot\mathbf{n})^6 = (\boldsymbol{\sigma}\cdot\mathbf{n})^8 = \cdots = 1$$

$$(\boldsymbol{\sigma}\cdot\mathbf{n})^3 = (\boldsymbol{\sigma}\cdot\mathbf{n})^5 = (\boldsymbol{\sigma}\cdot\mathbf{n})^7 = \cdots = \boldsymbol{\sigma}\cdot\mathbf{n}$$

$$e^{-i\theta\mathbf{n}\cdot\boldsymbol{\sigma}/2} = 1 - i\frac{\theta}{2}\boldsymbol{\sigma}\cdot\mathbf{n} - \frac{\left(\frac{\theta}{2}\right)^2}{2!} + i\left(\frac{\theta}{2}\right)^3\boldsymbol{\sigma}\cdot\mathbf{n} + \ldots$$

$$= \left[1 - \frac{\left(\frac{\theta}{2}\right)^2}{2!} + \frac{\left(\frac{\theta}{2}\right)^4}{4!} + \cdots\right] - i\boldsymbol{\sigma}\cdot\mathbf{n}\left[\frac{\theta}{2} - \frac{\left(\frac{\theta}{2}\right)^3}{3!} + \cdots\right]$$

$$= \cos\frac{\theta}{2}I - i\boldsymbol{\sigma}\cdot\mathbf{n}\sin\frac{\theta}{2}$$

10.8 ADDITION OF ANGULAR MOMENTA

We have already seen that each particle has spin and orbital angular momenta. How to add these two angular momenta? Consider another system, say an atom. An atom has Z electrons, and so we have to add angular momenta of all the electrons to get the effective angular momentum of the atom. In general, angular momenta of the sub-parts of a system have to be added. How to add two angular momenta in a way consistent with quantum mechanics?

Let $\hat{J}_1$ and $\hat{J}_2$ be two angular momenta to be added. First, we have to realize that the angular momenta $\hat{J}_1$ and $\hat{J}_2$ act on different linear vector spaces. This is similar to the operators d/dx and d/dy. They operate on the functions of different variables. d/dx operates on $f(x)$, and d/dy operates on $g(y)$.

Suppose $\hat{J}_1$ and $\hat{J}_2$ represent orbital and spin angular momentum, respectively; i.e., $\hat{J}_1 = \hat{L}$ and $\hat{J}_2 = \hat{S}$. It is obvious that their linear vector spaces are different. Both $\hat{J}_1$ and $\hat{J}_2$ may be orbital angular momenta; i.e., $\hat{J}_1 = \mathbf{r}_1 \times \mathbf{p}_1$ $\hat{J}_2 = \mathbf{r}_2 \times \mathbf{p}_2$, where $\mathbf{r}_1$ and $\mathbf{r}_2$ are position vectors of particles 1 and 2. So we see again that $\hat{J}_1$ and $\hat{J}_2$ act on two different linear vector spaces.

Let us write $\hat{J}$ as

$$\hat{J} = \hat{J}_1 + \hat{J}_2 \tag{10.43}$$

$\hat{J}_1$ acts on a linear vector space V_1; $\hat{J}_2$ acts on a different linear vector space V_2; what about $\hat{J}$? Its vector space V_3 is different from V_1 and V_2. In mathematics, there is a way of handling this type of problem. It needs tensor language in which the linear vector space of $\hat{J}$ is a direct product of V_1 and V_2. We will not go into such details. For our purpose, the fact that $\hat{J}_1$ and $\hat{J}_2$ act on linear vector spaces has an important consequence. Simply because they operate on different linear vector spaces, they commute with each other.

$$\left[\hat{\mathbf{J}}_1, \hat{\mathbf{J}}_2\right] = 0 \tag{10.44}$$

This is similar to the commutation between the operators $\frac{d}{dx}$ and $\frac{d}{dy}$.

$\hat{J}$-Angular Momentum

Let us first show that $\mathbf{J} = \mathbf{J}_1 + \mathbf{J}_2$ is also an angular momentum operator.

$$\begin{aligned}\left[\hat{J}_x, \hat{J}_y\right] &= \left[\hat{J}_{1x} + \hat{J}_{2x}, \hat{J}_{1y} + \hat{J}_{2y}\right] \\ &= \left[\hat{J}_{1x}, \hat{J}_{1y}\right] + \underbrace{\left[\hat{J}_{1x}, \hat{J}_{2y}\right]}_{0} + \underbrace{\left[\hat{J}_{2x}, \hat{J}_{1y}\right]}_{0} + \left[\hat{J}_{2x}, \hat{J}_{2y}\right] = i\hbar\left[\hat{J}_{1z} + \hat{J}_{2z}\right] = i\hbar\hat{J}_z\end{aligned}$$

Similarly, we can prove that

$$\left[\hat{J}_y, \hat{J}_z\right] = i\hbar\hat{J}_x \quad \text{and} \quad \left[\hat{J}_z, \hat{J}_x\right] = i\hbar\hat{J}_y.$$

$\therefore$ $\hat{J} = \hat{J}_1 + \hat{J}_2$ is also an angular momentum operator.

Uncoupled Representation $\hat{J}_1^2, \hat{J}_2^2, \hat{J}_{1z}, \hat{J}_{2z}$

All these four operators commute among themselves.

$$\begin{aligned}&\left[\hat{J}_1^2, \hat{J}_2^2\right] = 0 \; \left[\hat{J}_1^2, \hat{J}_{1z}\right] = 0 \; \left[\hat{J}_1^2, \hat{J}_{2z}\right] = 0 \; \left[\hat{J}_2^2, \hat{J}_{1z}\right] = 0 \\ &\left[\hat{J}_2^2, \hat{J}_{2z}\right] = 0 \; \left[\hat{J}_{1z}, \hat{J}_{2z}\right] = 0\end{aligned} \tag{10.45}$$

Therefore, these operators have common eigenstates. These eigenstates form a basis for a representation called uncoupled or direct representation. The quantum numbers characterizing these states are j_1, j_2, m_1 and m_2. So these four quantum numbers can be written inside the ket. The basis kets of the uncoupled representations are written as $|\, j_1 m_1; j_2 m_2\rangle$. In fact, $|\, j_1 m_1; j_2 m_2\rangle$ is the direct product of the kets $|\, j_1 m_1\rangle$ and $|\, j_2 m_2\rangle$.

$$|j_1 m_1; j_2 m_2\rangle = |j_1 m_1\rangle\, |j_2 m_2\rangle \tag{10.46}$$

$$\hat{J}_1^2\, |\, j_1 m_1; j_2 m_2\rangle = j_1(j_1 + 1)\hbar^2\, |\, j_1 m_1; j_2 m_2\rangle \tag{10.47}$$

$$\hat{J}_2^2 \,|\, j_1 m_1; j_2 m_2\rangle = j_2(j_2+1)\hbar^2 \,|\, j_1 m_1; j_2 m_2\rangle \tag{10.48}$$

$$\hat{J}_{1Z} \,|\, j_1 m_1; j_2 m_2\rangle = m_1\hbar \,|\, j_1 m_1; j_2 m_2\rangle \tag{10.49}$$

$$\hat{J}_{2Z} \,|\, j_1 m_1; j_2 m_2\rangle = m_2\hbar \,|\, j_1 m_1; j_2 m_2\rangle \tag{10.50}$$

Coupled Representation: $\hat{J}^2, \hat{J}_z, \hat{J}_1^2, \hat{J}_2^2$

$$\begin{aligned}
&[\hat{J}^2, \hat{J}_z] = 0, \quad [\hat{J}^2, \hat{J}_1^2] = 0, \quad [\hat{J}^2, \hat{J}_2^2] = 0 \\
&[\hat{J}_z, \hat{J}_1^2] = 0, \quad [\hat{J}_z, \hat{J}_2^2] = 0, \quad [\hat{J}_1^2, \hat{J}_2^2] = 0
\end{aligned} \tag{10.51}$$

These four operators commute among themselves. Therefore, $\hat{J}^2, \hat{J}_z, \hat{J}_1^2$ and $\hat{J}_2^2$ have common eigenstates, and these eigenstates form a basis for the representation called coupled representation. The quantum numbers of $\hat{J}_1^2, \hat{J}_2^2, \hat{J}^2$ and $\hat{J}_z$ should be used to describe these states. Therefore, the quantum numbers j_1, j_2, j and m can be written inside the ket vectors $|\, j_1 j_2; jm\rangle$.

$$\hat{J}^2 \,|\, j_1 j_2; jm\rangle = j(j+1)\hbar^2 \,|\, j_1 j_2; jm\rangle \tag{10.52}$$

$$\hat{J}_Z \,|\, j_1 j_2; jm\rangle = m\hbar \,|\, j_1 j_2; jm\rangle \tag{10.53}$$

$$\hat{J}_1^2 \,|\, j_1 j_2; jm\rangle = j_1(j_1+1)\hbar^2 \,|\, j_1 j_2; jm\rangle \tag{10.54}$$

$$\hat{J}_2^2 \,|\, j_1 j_2; jm\rangle = j_2(j_2+1)\hbar^2 \,|\, j_1 j_2; jm\rangle \tag{10.55}$$

Note that $[\hat{J}^2, \hat{J}_{1z}] \neq 0$ and $[\hat{J}^2, \hat{J}_{2z}] \neq 0$. Therefore, m_1 and m_2 cannot be included in the list of quantum numbers used to describe the basis vectors in the coupled representation.

Example 10.5 Evaluate $[\hat{J}^2, \hat{J}_{1z}]$ and $[\hat{J}^2, \hat{J}_{2z}]$.

Solution:

$$\begin{aligned}
[\hat{J}^2, \hat{J}_{1z}] &= [\hat{J}_1^2 + \hat{J}_2^2 + 2\hat{J}_1.\hat{J}_2, \hat{J}_{1z}] \\
&= [\hat{J}_1^2, \hat{J}_{1z}] + [\hat{J}_2^2, \hat{J}_{1z}] + [2\hat{J}_1.\hat{J}_2, \hat{J}_{1z}]
\end{aligned}$$

$$\begin{aligned}
[\hat{J}_1.\hat{J}_2, \hat{J}_{1z}] &= [\hat{J}_{1x}\hat{J}_{2x} + \hat{J}_{1y}\hat{J}_{2y} + \hat{J}_{1z}\hat{J}_{2z}, \hat{J}_{1z}] \\
&= [\hat{J}_{1x}, \hat{J}_{1z}]\hat{J}_{2x} + [\hat{J}_{1y}, \hat{J}_{1z}]\hat{J}_{2y} \\
&= -i\hbar J_{1y}J_{2x} + i\hbar J_{1x}J_{2y}
\end{aligned}$$

$$[\hat{J}^2, \hat{J}_{1z}] = 2(-i\hbar J_{1y}J_{2x} + i\hbar J_{1x}J_{2y})$$

Similarly, we have

$$[\hat{J}_1 \cdot \hat{J}_2, \hat{J}_{2z}] = -i\hbar J_{1x}J_{2y} + i\hbar \hat{J}_{1y}\hat{J}_{2x}$$

$$\therefore [\hat{J}_1 \cdot \hat{J}_2, \hat{J}_{2z}] \neq 0$$

So we have

$$[\hat{J}^2, \hat{J}_{1z}] \neq 0, \ [\hat{J}^2, \hat{J}_{2z}] \neq 0$$

But

$$[\hat{J}^2, \hat{J}_{1z} + \hat{J}_{2z}] = [\hat{J}, \hat{J}_z] = 0$$

10.9 NUMBER OF BASIS VECTORS IN COUPLED AND UNCOUPLED REPRESENTATION

Both coupled and uncoupled representations are the result of addition of angular momenta $\hat{J}_1$ and $\hat{J}_2$. Therefore, the number of basis vectors in the coupled and uncoupled representations is the same.

In the uncoupled representations, the basis vectors are given by

$$| j_1 m_1; j_2 m_2\rangle = | j_1 m_1\rangle \, | j_2 m_2\rangle$$

It is obvious that there are $(2j_1+1)(2j_2+2)$ basis states in the uncoupled representations.

In the coupled representation, the basis vectors are written as $| j_1 j_2; jm\rangle$. In a shorter notation, this can be written as $| jm\rangle$.

$$| j_1 j_2 jm\rangle \to | jm\rangle \quad \text{(shorter notation)} \tag{10.56}$$

For a given j, there are $(2j + 1)$ states. The number $(2j + 1)$ cannot be equal to $(2j_1+1)(2j_2+2)$. This implies that for a given j_1 and j_2, j cannot be a single value. j has to vary over a range from $j_{\max}$ to $j_{\min}$ such that the sum of all the states for all the values of j is the same as $(2j_1+1)(2j_2+1)$.

$$\sum_{j_{\min}}^{j_{\max}} (2j+1) = (2j_1+1)(2j_2+1) \tag{10.57}$$

These two basis vectors are related by unitary transformation.

Let us recall $|\psi\rangle$ can be written as

$$|\psi\rangle = \sum C_n |\phi_n\rangle \ \text{ with } \ C_n = \langle\phi_n|\psi\rangle$$

In the place of $|\psi\rangle$, let us write $| j_1 j_2; jm\rangle$, and in the place of $|\phi_n\rangle$, let us write $| j_1 m_1\rangle | j_2 m_2\rangle$. The coefficient C_n is replaced by a complicated symbol $C(j_1 m_1; j_2 m_2 jm)$ or $\langle j_1 m_1; j_2 m_2 | j_1 j_2; jm\rangle$to include as much information as possible.

$$\underbrace{|\psi\rangle}_{|j_1 j_2 jm\rangle} = \sum \underbrace{C_n}_{\langle j_1 m_1; j_2 m_2 | jm\rangle} \underset{|j_1 m_1\rangle|j_2 m_2\rangle}{|\phi_n\rangle}$$

$$|j_1 j_2; jm\rangle = \sum_{m_1, m_2} \underbrace{\langle j_1 m_1; j_2 m_2 | j_1 j_2; jm\rangle}_{\text{a number}} | j_1 m_1; j_2 m_2\rangle \tag{10.58}$$

In a more compact way, this can be written as

$$\therefore | jm\rangle = \sum_{m_1, m_2} \underbrace{\langle j_1 m_1; j_2 m_2 | jm\rangle}_{\text{a number}} | j_1 m_1; j_2 m_2\rangle \tag{10.59}$$

The coefficients $\langle j_1m_1; j_2m_2 \mid jm\rangle$ are called Clebsh–Gordan coefficients. There are many symbols used to write Clebsh–Gordan coefficients like $C(j_1m_1; j_2m_2|jm)$ or $C^{jm}_{j_1m_1;j_2m_2}$ or $\langle m_1; m_2|jm\rangle$. We choose to write the Clebsh–Gordan coefficients as $\langle j_1m_1; j_2m_2 \mid jm\rangle$.

In the same way, $|j_1m_1; j_2m_2\rangle$ can be expressed as a linear combination of kets $|jm\rangle$.

$$|j_1m_1; j_2m_2\rangle = \sum_{jm} \langle j_1j_2; jm|j_1m_1; j_2m_2\rangle \, |j_1j_2; jm\rangle$$

$$\text{or } |j_1m_1; j_2m_2\rangle = \sum_{jm} \underbrace{\langle jm|j_1m_1; j_2m_2\rangle}_{\text{a number}} |jm\rangle \tag{10.60}$$

Obviously, $\langle j_1m_1; j_2m_2|jm\rangle = \langle jm|j_1m_1; j_2m_2\rangle^*$. In principle, the Clebsh–Gordan coefficient can be complex numbers. However, by convention, all the Clebsh–Gordan coefficients are chosen as real numbers $\langle j_1m_1; j_2m_2|jm\rangle = \langle jm|j_1m_1; j_2m_2\rangle$.

10.10 POSSIBLE VALUES OF *J* IN COUPLED REPRESENTATION

The angular momenta $\hat{J}_1$ and $\hat{J}_2$ are added to give angular momentum $\hat{J}$. Let us recall that

$$\hat{J}_1^2 \,|j_1m_1\rangle = j_1(j_1+1)\hbar^2 \,|j_1m_1\rangle$$
$$\hat{J}_2^2 \,|j_2m_2\rangle = j_2(j_2+1)\hbar^2 \,|j_2m_2\rangle$$
$$\hat{J}^2 \,|jm\rangle = j(j+1)\hbar^2 \,|jm\rangle$$

We have already seen that the values of j vary from $j_{\max}$ to $j_{\min}$. We have not yet determined the possible values of j in terms of j_1 and j_2.

In terms of the basis of uncoupled representation, $|jm\rangle$ is given by

$$|jm\rangle = \sum_{m_1m_2} \langle j_1m_1; j_2m_2 \mid jm\rangle \,|j_1m_1; j_2m_2\rangle \tag{10.61}$$

Let us consider $\hat{J}_z \mid jm\rangle$.

$$\hat{J}_z \,|jm\rangle = (\hat{J}_{1z} + \hat{J}_{2z})\,|jm\rangle$$

$$\begin{aligned}
\text{LHS} &= \hat{J}_z \,|jm\rangle = m\hbar\,|jm\rangle \\
&= \sum_{m_1;m_2} m\hbar\langle j_1m_1; j_2m_2|jm\rangle \,|j_1m_1\rangle\,|j_2m_2\rangle \\
\text{RHS} &= (\hat{J}_{1z} + \hat{J}_{2z})\,|jm\rangle \\
&= \sum_{m_1m_2} (\hat{J}_{1z} + \hat{J}_{2z})\,\langle j_1m_1; j_2m_2|jm\rangle\,|j_1m_1\rangle\,|j_2m_2\rangle \\
&= \sum_{m_1m_2} \langle j_1m_1; j_2m_2|jm\rangle\,(m_1+m_2)\hbar\,|j_1m_1\rangle\,|j_2m_2\rangle
\end{aligned}$$

Equating LHS and RHS, we get

$$\sum_{m_1;m_2} (m - m_1 - m_2)\hbar\langle j_1 m_1; j_2 m_2 | jm\rangle \; |j_1 m_1; j_2 m_2\rangle = 0$$

Since each vector $|j_1 m_1; j_2 m_2\rangle$ is a linearly independent vector, the coefficient of each term in the above summation is zero.

$$\therefore \underbrace{\langle j_1 m_1; j_2 m_2 | jm\rangle}_{\text{a number}}(m - m_1 - m_2) = 0 \tag{10.62}$$

$\therefore$ Either $\langle j_1 m_1; j_2 m_2 | jm\rangle = 0$ or $m = m_1 + m_2$

If $m \neq m_1 + m_2$, the coefficient $\langle j_1 m_1; j_2 m_2 | jm\rangle$ has to be zero. So the only coefficients which survive in the summation are such that $m = m_1 + m_2$.

This condition fixes the value for a given m_1 and m_2. However, the j value is not fixed uniquely. For instance a state, say $|m_1 = 2, m_2 = 0\rangle$, is associated with two states $|j = 3, m = 2\rangle$ and $|j = 2, m = 2\rangle$. So, in general, any state $|j_1 m_1\rangle |j_2 m_2\rangle$ can be associated with a number of state $|j, m = m_1 + m_2\rangle$, corresponding to different values of j. The number of such possible states for a given m_1 and m_2 is known as multiplicity or degeneracy, corresponding to $|j_1 m_1\rangle |j_2 m_2\rangle$.

Returning back to the Equation (10.62), we write it as

$$|jm\rangle = \sum_{m_1} \langle j_1 m_1; j_2 m_2 | jm\rangle \quad |j_1 m_1; j_2 m_2\rangle \quad \text{with } m_1 + m_2 = m$$

$$|jm\rangle = \sum_{m_1} \langle j_1 m_1; j_2\, m - m_1 | jm\rangle \quad |j_1 m_1; j_2\, m - m_1\rangle \tag{10.63}$$

With these information, we can determine the possible values of j.

Let us start with the maximum values of m_1 and m_2. Now we have $m_1 = j_1$ and $m_2 = j_2$. The corresponding direct product state in the uncoupled representation is $|j_1 j_1\rangle\, |j_2 j_2\rangle$.

There can be only one term in the RHS of (10.63) with $|j_1 j_1\rangle\, |j_2 j_2\rangle$.

$$\therefore |j, j_1 + j_2\rangle = |j_1 j_1\rangle\, |j_2 j_2\rangle \tag{10.64}$$

The corresponding value of j has to be $(j_1 + j_2)$.

i.e., $$j = j_1 + j_2$$

This is the maximum possible value for j. It is easy to check this result (see Examples 10.6 and 10.7).

$$\therefore j_{\max} = j_1 + j_2 \tag{10.65}$$

$$\hat{J}^2\, |j_1 j_1\rangle\, |j_2 j_2\rangle = (j_1 + j_2)(j_1 + j_2 + 1)\hbar^2\, |j_1 j_1\rangle\, |j_2 j_2\rangle$$

This implies that the state $|j\, \overline{j_1 + j_2}\rangle$ in the coupled representation is given by

$$|j\, \overline{j_1 + j_2}\rangle = |j_1 j_1\rangle\, |j_2 j_2\rangle \quad \text{where } j = j_{\max} = j_1 + j_2$$

Next, let us decrease the value of m_1 or m_2 of 1.

So we have two possible direct product states with the same m value:

$$|\,j_1\overline{j_1-1}\rangle\,|\,j_2 j_2\rangle \quad \text{and} \quad |\,j_1 j_1\rangle\,|\,j_2\overline{j_2-1}\rangle$$

The m value of these states is $j_1 + j_2 - 1$. Then a state in the coupled representation is given by

$$\therefore |j\ \overline{j_1+j_2-1}\rangle = c_1\,|j_1\ \overline{j_1-1}\rangle\,|j_2 j_2\rangle + c_2\,|j_1 j_1\rangle\,|j_2\ \overline{j_2-1}\rangle$$

What are the possible values of j? Both $j = j_1 + j_2$ and $j = j_1 + j_2 - 1$ are consistent with RHS of the above equation.

So the possible values of j are $j_1 + j_2$ and $j_1 + j_2 - 1$.

Again decrease m_1 or m_2 by 1. So we have there possible direct product states with the same value for m.

$$|j_1 j_1\rangle\,|j_2\overline{j_2-2}\rangle, |\,j_1\overline{j_1-1}\rangle\,|j_2\overline{j_2-1}\rangle \quad \text{and} \quad |j_1\overline{j_1-2}\rangle\,|\,j_2 j_2\rangle$$

The m values for all these states are $j_1 + j_2 - 2$.

This means that we can have a state in the coupled representation given by

$$\therefore |\,j\ \overline{j_1+j_2-2}\rangle = c_1\,|\,j_1\ \overline{j_1-2}\rangle\,|j_1 j_2\rangle + c_2\,|j_1\ \overline{j_1-1}\rangle\,|j_2\ \overline{j_2-1}\rangle + c_3\,|j_1 j_1\rangle\,|j_2\ \overline{j_2-2}\rangle$$

The value of j consistent with $m = j_1 + j_2 - 2$ are

$$j = j_1 + j_2, \quad j = j_1 + j_2 - 1, \quad j = j_1 + j_2 - 2$$

Repeating this procedure, we can list all the possible values of j. They extend from $j_{\max}$ to $j_{\min}$. It can be shown that $j_{\min} = |\,j_1 - j_2|$.

The possible values of j are

$$j_1 + j_2, \quad j_1 + j_2 - 1, \quad j_1 + j_2 - 2, \ldots |j_1 - j_2| \tag{10.66}$$

It is easy to check $j_{\min}$ is $|j_1 - j_2|$.

Let us assume $j_1 > j_2$. Then, the possible values of j are claimed to be

$$j = j_1 + j_2, j_1 + j_2 - 1, \ldots, j_1 - j_2 \tag{10.67}$$

There are $(2j_2 + 1)$ terms in this sequence, or the number of possible values of j is $2j_2 + 1$. This can be seen in many ways. One way is to write the possible values of j as

$$j_1 + (j_2), \quad j_1 + (j_2 - 1), \quad j_1 + (j_2 - 2), \ldots j_1 + (-j_2)$$

The quantities in the brackets are $j_2, j_2 - 1, \ldots, -j_2$. Obviously, there are $(2j_2 + 1)$ terms in the sequence given in (10.67).

Each j has $(2j + 1)$ states in coupled representation. The total number of basis states should be $(2j_1 + 1)(2j_2 + 1)$. The total number of states N is given by

$$\begin{aligned} N &= \sum_{j=j_1-j_2}^{j_1+j_2} (2j+1) = \frac{\text{No. of terms}}{2}(\text{First term} + \text{Last term}) \\ &= \frac{(2j_2+1)}{2}[2j_1 + 2j_2 + 1 + 2j_1 - 2j_2 + 1] \\ &= (2j_1+1)(2j_2+1) \end{aligned}$$

Example 10.6 Show that $\hat{J}^2 = (\hat{J}_1 + \hat{J}_2)^2 = \hat{J}_1^2 + \hat{J}_2^2 + \hat{J}_{2-}\hat{J}_{1+} + \hat{J}_{1-}\hat{J}_{2+} + 2\hat{J}_{1z}\hat{J}_{2z}$

Solution:

$$\begin{aligned}\hat{J}^2 &= (\hat{J}_1 + \hat{J}_2)^2 = \hat{J}_1^2 + \hat{J}_2^2 + 2\hat{J}_1 \cdot \hat{J}_2 \\ &= \hat{J}_1^2 + \hat{J}_2^2 + 2\hat{J}_{1x}\hat{J}_{2x} + 2\hat{J}_{1y}\hat{J}_{2y} + 2\hat{J}_{1z}\hat{J}_{2z}\end{aligned}$$

We have

$$\hat{J}_x = \frac{1}{2}\left(\hat{J}_+ + \hat{J}_-\right) \text{ and } \hat{J}_y = \frac{1}{2i}\left(\hat{J}_+ - \hat{J}_-\right)$$

$$\begin{aligned}2\hat{J}_{1x}\hat{J}_{2x} &= \frac{1}{2}\left(\hat{J}_{1+} + \hat{J}_{1-}\right)\left(\hat{J}_{2+} + \hat{J}_{2-}\right) \\ &= \frac{1}{2}\left[\hat{J}_{1+}\hat{J}_{2+} + \hat{J}_{1+}\hat{J}_{2-} + \hat{J}_{1-}\hat{J}_{2+} + \hat{J}_{1-}\hat{J}_{2-}\right] \\ 2\hat{J}_{1y}\hat{J}_{2y} &= \frac{-1}{2}\left[\hat{J}_{1+}\hat{J}_{2+} - \hat{J}_{1+}\hat{J}_{2-} - \hat{J}_{1-}\hat{J}_{2+} + \hat{J}_{1-}\hat{J}_{2-}\right]\end{aligned}$$

Making use of these results, we get

$$\therefore \hat{J}^2 = \hat{J}_1^2 + \hat{J}_2^2 + \hat{J}_{1+}\hat{J}_{2-} + \hat{J}_{1-}\hat{J}_{2+} + 2\hat{J}_{1z}\hat{J}_{2z}$$

Example 10.7 Show that $\hat{J}^2\,|j_1 j_1\rangle\,|j_2 j_2\rangle = (j_1 + j_2)(j_1 + j_2 + 1)\hbar^2\,|j_1 j_1\rangle\,|j_2 j_2\rangle$

Solution:

$$\hat{J}^2|\,j_1 j_1\rangle|j_2 j_2\rangle = (\hat{J}_1^2 + \hat{J}_2^2 + \hat{J}_{1+}\hat{J}_{2-} + \hat{J}_{1-}\hat{J}_{2+} + 2\hat{J}_{1z}\hat{J}_{2z})\,|j_1 j_1\rangle\,|j_2 j_2\rangle$$

$$\hat{J}_{1+}\hat{J}_{2-}\,|\,j_1 j_1\rangle\,|\,j_2 j_2\rangle = (\hat{J}_{1+}\,|\,j_1 j_1\rangle)\cdot(\hat{J}_{2-}\,|\,j_2 j_2\rangle) = 0$$

$$\hat{J}_{1-}\hat{J}_{2+}\,|\,j_1 j_1\rangle\,|\,j_2 j_2\rangle = (\hat{J}_{1-}\,|\,j_1 j_1\rangle)\cdot(\hat{J}_{2+}\,|\,j_2 j_2\rangle) = 0$$

$$\hat{J}_{1z}\hat{J}_{2z}\,|\,j_1 j_2\rangle\,|\,j_2 j_2\rangle = j_1 j_2 \hbar^2\,|\,j_1 j_1\rangle\,|\,j_2 j_2\rangle$$

$$\begin{aligned}\therefore \hat{J}^2|j j_1\rangle\,|j_2 j_2\rangle &= [j_1(j_1+1)\hbar^2 + j_2(j_2+1)\hbar^2 + 2 j_1 j_2 \hbar^2]\,|j_1 j_1\rangle\,|\,j_2 j_2\rangle \\ &= (j_1 + j_2)(j_1 + j_2 + 1)\hbar^2|j_1 j_1\rangle\,|j_2 j_2\rangle\end{aligned}$$

Example 10.8 When two angular momenta $\hat{J}_1$ and $\hat{J}_2$ are added, the resulting angular momentum is $\hat{J}$. The possible values are

$$j = j_1 + j_2,\ j_1 + j_2 - 1,\ \ldots,\ j_{\min}$$

Show that

$$j_{\min} = |j_1 - j_2|.$$

Solution: Let N be the total number of states for all j values.

$$N = \sum_{j_{\max}}^{j_{\min}} (2j+1)$$

$$\begin{aligned}&= [2(j_1 + j_2) + 1] + [2(j_1 + j_2 - 1) + 1] + \ldots + (2j_{\min} + 1) \\ &= (2j_1 + 2j_2 + 1) + [2j_1 + 2j_2 - 2 + 1] + \ldots + (2j_{\min} + 1)\end{aligned}$$

This is a sum of terms in arithmetic progressions. The first term is $a = 2j_2 + 2j_2 + 1$, the last term is $l = 2j_{\min} + 1$ and the common difference is $d = -2$.

$$\therefore N = \frac{(l+a)(l-a+d)}{2d}$$

This should be equal to $(2j_1+1)(2j_2+1)$.

$$\therefore N = \frac{(2j_{\min} + 2j_1 + 2j_2 + 2)(2j_{\min} + 1 - 2j_1 - 2j_2 - 1 - 2)}{-4}$$

$$\begin{aligned}\therefore (j_{\min} + j_1 + j_2 + 1)(j_{\min} - j_1 - j_2 - 1) &= -(2j_1+1)(2j_2+1)\\ j_{\min}^2 &= (j_1 + j_2 + 1)^2 - (2j_1+1)(2j_2+1)\\ &= (j_1 - j_2)^2\\ \therefore j_{\min} &= |j_1 - j_2|\end{aligned}$$

10.11 CLEBSH–GORDAN COEFFICIENTS

The number of basis vectors in both coupled and uncoupled representations is $(2j_1+1)(2j_2+1)$. Therefore, the transformation matrix relating the two bases is $(2j_1+1)(2j_2+1)$ dimensional unitary matrix.

$$|jm\rangle = \sum_{m_1}\sum_{m_2} \langle j_1m_1; j_2m_2|jm\rangle\ |j_1m_1; j_2m_2\rangle \tag{10.68}$$

$$|jm_1; jm_2\rangle = \sum_{j}\sum_{m} \langle jm|j_1m_1; j_2m_2\rangle\, |jm\rangle \tag{10.69}$$

Using the Hermitian conjugate of (10.68), we have

$$\langle j'm'| = \sum_{m_1'}\sum_{m_2'} \langle j_1m_1'; j_2m_2'|j'm'\rangle\, \langle j_1m_1'; j_2m_2'|$$

Using the above equation and (10.68), we get

$$\begin{aligned}\langle j'm'|jm\rangle &= \sum_{m_1}\sum_{m_2}\sum_{m_1'}\sum_{m_2'} \langle j_1m_1; j_2m_2|jm\rangle\, \langle j_1m_1'; j_2m_2'|j'm'\rangle\, \langle j_1m_1'; j_2m_2'|j_1m_1; j_2m_2\rangle\\ &= \sum_{m_1}\sum_{m_2}\sum_{m_1'}\sum_{m_2'}\ \langle j_1m_1; j_2m_2|jm\rangle\langle j_1m_1'; j_2m_2'|j'm'\rangle\delta_{m_1m_1'}\ \delta_{m_2m_2'}\\ &= \sum_{m_1}\sum_{m_2} \langle j_1m_1; j_2m_2|jm\rangle\langle j_1m_1; j_2m_2|j'm'\rangle\end{aligned}$$

$$\text{LHS} = \delta_{jj'}\delta_{mm'}$$

$$\therefore \sum_{m_1}\sum_{m_2} \langle j_1m_1; j_2m_2|jm\rangle\langle j_1m_1; j_2m_2|j'm'\rangle = \delta_{jj'}\delta_{mm'} \tag{10.70}$$

Similarly, using the Hermitian conjugate of (10.69), we can write

$$\langle j_1m_1'; j_2m_2'| = \sum_{j'}\sum_{m'} \langle j'm'|j_1m_1'; j_2m_2'\rangle\langle j'm'|$$

$$\therefore \langle j_1m_1'; j_2m_2'|j_1m_1; j_2m_2\rangle = \sum_j \sum_m \sum_{j'} \sum_{m'} \langle jm|j_1m_1; j_2m_2\rangle\langle j'm'|j_1m_1'; j_2m_2'\rangle\langle j'm'|jm\rangle$$

$$= \sum_j \sum_m \sum_{j'} \sum_{m'} \langle jm|j_1m_1'; j_2m_2{}'\rangle\langle j'm'|j_1m_1'; j_2m_2'\rangle\delta_{jj'}\delta_{m'm}$$

$$= \sum_j \sum_m \langle jm|j_1m_1; j_2m_2\rangle\langle jm|j_1m_1'; j_2m_2'\rangle$$

But $$\text{LHS} = \delta_{m_1'm_1}\,\delta_{m_2'm_2}$$

$$\therefore \sum_j \sum_m \langle jm\,|\,j_1m_1; j_2m_2\rangle\langle jm\,|\,j_1m_1'; j_2m_2'\rangle = \delta_{m_1'm_1}\delta_{m_2m_2'} \tag{10.71}$$

The relations (10.70) and (10.71) are known as orthoganality relations for the Clebsh–Gordan coefficient. The Clebsh–Gordan coefficients can be presented in the form of a matrix as shown below:

Table 10.1 Matrix representation of Clebsh–Gordan coefficients

	m $\mathbf{j_1 + j_2 = j}$	**m** $\mathbf{j - 1}$		**m** $\mathbf{j - 2}$			• • •	**m** $\mathbf{-j}$
	$\vert jj\rangle$	$\vert j\,\overline{j-1}\rangle$	$\vert j-1,$ $j-1\rangle$	$\vert j\,\overline{j-2}\rangle$	$\vert j-1,$ $j-2\rangle$	$\vert j-2,$ $j-2\rangle$	• • •	$\vert -j,$ $-j\rangle$
$m = j_1 + j_2$ $\langle j_1j_1; j_2j_2\,\vert$	1 × 1							
$m = j_1 + j_2 - 1$ $\langle j_1j_1; j_2\overline{j_2-1}\,\vert$ $\langle j_1\overline{j_1-1}; j_2j_2\,\vert$		2 × 2						
$m = j_1 + j_2 - 2$ $\langle j_1; j_2 - 2\,\vert$ $\langle j_1 - 1; j_2 - 1\vert$ $\langle j_2 - 2; j_2\,\vert$				3 × 3			• • • • • •	
• ⋮• •								
$m = -j_1 - j_2 + 1$ $\langle j_1\overline{-j_1+1}; j_2\overline{-j_2}\,\vert$ $\langle j_1\overline{-j_1}; j_2\overline{-j_2+1}\vert$					••• ••• •••			
$m = -j_1 - j_2$ $\langle j_1\overline{-j_1}; j_2\overline{-j_2}\,\vert$								1 × 1

The orthoganality relation (10.70) indicates the orthoganality between the column vectors in the unitary matrix. Similarly, the orthoganality relation (10.71) is the orthoganality relation between the rows in the unitary matrix.

10.12 RECURSION RELATIONS FOR CLEBSH–GORDAN COEFFICIENT

Let us consider $|\, jm\rangle$.

$$|jm\rangle = \sum_{m_1}\sum_{m_2}\langle j_1 m_1; j_2 m_2|jm\rangle\,|j_1 m_1; j_2 m_2\rangle$$

Let $\hat{J}_+$ act on either side of this equation.

$$\hat{J}_+\,|jm\rangle = \sum_{m_1}\sum_{m_2}\left(\hat{J}_{1+} + \hat{J}_{2+}\right)\langle j_1 m_1; j_2 m_2|jm\rangle\,|j_1 m_1; j_2 m_2\rangle$$

$$\begin{aligned}\text{RHS} &= \sum_{m_1}\sum_{m_2}\langle j_1 m_1; j_2 m_2|jm\rangle\,(j_1 - m_1)^{1/2}(j_1 + m_1 + 1)^{1/2}\hbar\,|j_1 \overline{m_1 + 1}; j_2 m_2\rangle \\ &\quad + \sum_{m_1}\sum_{m_2}\langle j_1 m_1; j_2 m_2|jm\rangle\,(j_2 - m_2)^{1/2}(j_2 + m_2 + 1)^{1/2}\hbar\,|j_1 m_1; j_2\,\overline{m_2 + 1}\rangle \\ &= T_1 + T_2\end{aligned}$$

Consider the first term T_1.

Define $m_1' = m_1 + 1$ $\qquad\qquad\therefore m_1 = m_1' - 1$

$$\therefore T_1 = \sum_{m_1'}\sum_{m_2}\langle j_1\,\overline{m_1' - 1}; j_2 m_2|jm\rangle\,(j_1 - m_1' + 1)^{1/2}(j_1 + m_1')^{1/2}\hbar\,|j_1 m_1'; j_2 m_2\rangle$$

Since m_1' is a dummy variable, replace it by m_1.

$$\therefore T_1 = \sum_{m_1}\sum_{m_2}\langle j_1\,\overline{m_1 - 1}; j_2 m_2|jm\rangle\,(j_1 - m_1 + 1)^{1/2}(j_1 + m_1)^{1/2}\hbar\,|j_1 m_1; j_2 m_2\rangle$$

In the same way, T_2 can be written as

$$\therefore T_2 = \sum_{m_1}\sum_{m_2}\langle j_1 m_1; j_2\,\overline{m_2 - 1}|jm\rangle\,(j_2 - m_2 + 1)^{1/2}(j_2 + m_2)^{1/2}\hbar\,|j_1 m_1; j_2 m_2\rangle$$

Now LHS is given.

$$\begin{aligned}\text{LHS} &= (j - m)^{1/2}(j + m + 1)^{1/2}\hbar\,|jm + 1\rangle \\ &= \sum_{m_1}\sum_{m_2}(j - m)^{1/2}(j + m + 1)^{1/2}\hbar\,\langle j_1 m_1; j_2 m_2|j\,\overline{m + 1}\rangle\,|j_1 m_1; j_2 m_2\rangle\end{aligned}$$

Equating LHS and RHS, we get

$$\begin{aligned}
&(j-m)^{1/2}(j+m+1)^{1/2}\langle j_1 m_1; j_2 m_2 | jm+1\rangle \\
&= (j_1 - m_1 + 1)^{1/2}(j_1 + m_1)^{1/2}\langle j_1\, \overline{m_1 - 1}; j_2 m_2 | jm\rangle \\
&\quad + (j_2 - m_2 + 1)^{1/2}(j_2 + m_2)^{1/2}\langle j_1 m_1; j_2\, \overline{m_2 - 1} | jm\rangle
\end{aligned} \tag{10.72}$$

We can obtain another recurrence relation using $\hat{J}_-$.

$$\hat{J}_- \,|\, jm\rangle = (\hat{J}_{1-} + \hat{J}_{2-})\sum_{m_1}\sum_{m_2}\langle j_1 m_1; j_2 m_2 \,|\, jm\rangle \,|\, j_1 m_1; j_2 m_2\rangle$$

Working in the same way as done for $\hat{J}_+$, we get

$$\begin{aligned}
&(j+m)^{1/2}(j-m+1)^{1/2}\langle j_1 m_1; j_2 m_2 | j\, \overline{m-1}\rangle \\
&= (j_1 + m_1 + 1)^{1/2}(j_1 - m_1)^{1/2}\langle j_1\, \overline{m_1 + 1}; m_2 | jm\rangle \\
&\quad + (j_2 + m_2 + 1)^{1/2}(j_2 - m_2)^{1/2}\langle j_1 m_1; j_2\, \overline{m_2 + 1} | jm\rangle
\end{aligned} \tag{10.73}$$

Both relations (10.72) and (10.73) are recurrence relations for Clebsh–Gordan coefficient. If we know the Clebsh–Gordan coefficient for one state corresponding to a angular momentum quantum number *j*, then the Clebsh–Gordan coefficients for all the (2*j* + 1) states corresponding to *j*. The given *j* can be obtaining using either (10.72) or (10.73).

Evaluation of the Clebsh–Gordan Coefficient

There are many situations when we require the knowledge of Clebsh–Gordan coefficients. With the help of the lowering and raising operators and orthoganality relations, we can evaluate the Clebsh–Gordan coefficient. While using orthoganality relations, the signs of some Clebsh–Gordan coefficient will be arbitrarily chosen. So we have to follow certain conventions in fixing the sign of Clebsh–Gordan coefficients, In addition, the Clebsh–Gordan coefficients obey a lot of symmetry relations which are useful in evaluating the Clebsh–Gordan coefficients. A complete analytic expression for Clebsh–Gordan coefficients is also available.

Some of the symmetry relations are given here (Alonso and Valk).

$$\begin{aligned}
\langle j_1 m_1; j_2 m_2 | j_1 j_2; j_3 m_3\rangle &= (-1)^{j_1 + j_2 - j_3}\langle j_2 m_2; j_1 m_2 | j_2 j_1; j_3 m_3\rangle \\
&= (-1)^{j_2 + m_2}\left[\frac{2j_3 + 1}{2j_1 + 1}\right]^{1/2}\langle j_3 - m_3; j_2 m_2 | j_3 j_2; j_1 - m_2\rangle \\
&= (-1)^{j_1 + j_2 - j_3}\langle j_1 - m_1; j_2 - m_2 | j_1 j_2; j_3 - m_3\rangle
\end{aligned}$$

As an illustration, let us work out the Clebsh–Gordan coefficient for some simple cases. The following convention will be very much helpful to us:

For each angular momentum quantum number *j*, the Clebsh–Gordan coefficient $\langle j_1 j_1\; ; j_2\; \overline{j - j_1} | jj\rangle$ is a positive number.

Example 10.9 Determine the Clebsh–Gordan coefficient when we add two angular momenta with quantum numbers $j_1 = 1/2$ and $j_2 = 1/2$.

Solution: The possible values of j are $j = 1, 0$. According to the phase convention, $\left\langle \frac{1}{2}\frac{1}{2};\frac{1}{2}\frac{1}{2}\middle|11\right\rangle$ and $\left\langle \frac{1}{2}\frac{1}{2};\frac{1}{2}-\frac{1}{2}\middle|00\right\rangle$ are positive.

We have to bear in mind that $m = m_1 + m_2$.

We have four basis vectors in coupled representation. They are $|11\rangle, |10\rangle, |1-1\rangle$ and $|00\rangle$. They can be written as follows:

$$
\begin{aligned}
|11\rangle &= \left\langle \frac{1}{2}\frac{1}{2};\frac{1}{2}\frac{1}{2}\middle|11\right\rangle \left|\frac{1}{2}\frac{1}{2};\frac{1}{2}\frac{1}{2}\right\rangle \\
|10\rangle &= \left\langle \frac{1}{2}\frac{1}{2};\frac{1}{2}-\frac{1}{2}\middle|10\right\rangle \left|\frac{1}{2}\frac{1}{2};\frac{1}{2}-\frac{1}{2}\right\rangle + \left\langle \frac{1}{2}-\frac{1}{2};\frac{1}{2}\frac{1}{2}\middle|10\right\rangle \left|\frac{1}{2}-\frac{1}{2};\frac{1}{2}\frac{1}{2}\right\rangle \\
|1-1\rangle &= \left\langle \frac{1}{2}-\frac{1}{2};\frac{1}{2}-\frac{1}{2}\middle|1-1\right\rangle \left|\frac{1}{2}-\frac{1}{2};\frac{1}{2}-\frac{1}{2}\right\rangle \\
|00\rangle &= \left\langle \frac{1}{2}\frac{1}{2};\frac{1}{2}-\frac{1}{2}\middle|00\right\rangle \left|\frac{1}{2}\frac{1}{2};\frac{1}{2}-\frac{1}{2}\right\rangle + \left\langle \frac{1}{2}-\frac{1}{2};\frac{1}{2}\frac{1}{2}\middle|00\right\rangle \left|\frac{1}{2}-\frac{1}{2};\frac{1}{2}\frac{1}{2}\right\rangle
\end{aligned}
$$

Let us remember that $|j_1m_1; j_2m_2\rangle = |j_1m_1\rangle|j_2m_2\rangle$. It is obvious that $\left\langle \frac{1}{2}\frac{1}{2};\frac{1}{2}\frac{1}{2}\middle|11\right\rangle = 1$.

i.e.,

$$|11\rangle = \left|\frac{1}{2}\frac{1}{2};\frac{1}{2}\frac{1}{2}\right\rangle = \left|\frac{1}{2}\frac{1}{2}\right\rangle\left|\frac{1}{2}\frac{1}{2}\right\rangle.$$

$$\hat{J}_- |11\rangle = (\hat{J}_{1-} + \hat{J}_{2-})\left|\frac{1}{2}\frac{1}{2}\right\rangle\left|\frac{1}{2}\frac{1}{2}\right\rangle$$

$$
\begin{aligned}
&\text{LHS}: \hat{J}_- |11\rangle = \sqrt{2}\hbar|10\rangle \\
&\text{RHS}: (\hat{J}_{1-} + \hat{J}_{2-})\left|\frac{1}{2}\frac{1}{2}\right\rangle\left|\frac{1}{2}\frac{1}{2}\right\rangle = \left(\hat{J}_{1-}\left|\frac{1}{2}\frac{1}{2}\right\rangle\right)\left|\frac{1}{2}\frac{1}{2}\right\rangle + \left|\frac{1}{2}\frac{1}{2}\right\rangle\left(\hat{J}_{2-}\left|\frac{1}{2}\frac{1}{2}\right\rangle\right) \\
&\qquad = \hbar\left|\frac{1}{2}-\frac{1}{2}\right\rangle\left|\frac{1}{2}\frac{1}{2}\right\rangle + \hbar\left|\frac{1}{2}\frac{1}{2}\right\rangle\left|\frac{1}{2}-\frac{1}{2}\right\rangle \\
&\qquad = \hbar\left|\frac{1}{2}-\frac{1}{2};\frac{1}{2}\frac{1}{2}\right\rangle + \hbar\left|\frac{1}{2}\frac{1}{2};\frac{1}{2}-\frac{1}{2}\right\rangle
\end{aligned}
$$

$$
\begin{aligned}
\therefore \sqrt{2}\hbar|10\rangle &= \hbar\left|\frac{1}{2}-\frac{1}{2};\frac{1}{2}\frac{1}{2}\right\rangle + \hbar\left|\frac{1}{2}\frac{1}{2};\frac{1}{2}-\frac{1}{2}\right\rangle \\
|10\rangle &= \frac{1}{\sqrt{2}}\left|\frac{1}{2}-\frac{1}{2};\frac{1}{2}\frac{1}{2}\right\rangle + \frac{1}{\sqrt{2}}\left|\frac{1}{2}\frac{1}{2};\frac{1}{2}-\frac{1}{2}\right\rangle
\end{aligned}
$$

We can straightaway write

$$|1-1\rangle = \left|\frac{1}{2}-\frac{1}{2};\frac{1}{2}-\frac{1}{2}\right\rangle$$

Alternatively, we calculate $\hat{J}_- \,|10\rangle$.

$$\begin{aligned}\hat{J}_-\,|10\rangle &= \frac{1}{\sqrt{2}}(\hat{J}_{1-}+\hat{J}_{2-})\left|\frac{1}{2}\frac{1}{2};\frac{1}{2}-\frac{1}{2}\right\rangle + \frac{1}{\sqrt{2}}(\hat{J}_{1-}+\hat{J}_{2-})\left|\frac{1}{2}-\frac{1}{2};\frac{1}{2}\frac{1}{2}\right\rangle \\ &= \frac{1}{\sqrt{2}}(\hat{J}_{1-}+\hat{J}_{2-})\left|\frac{1}{2}\frac{1}{2}\right\rangle\left|\frac{1}{2}-\frac{1}{2}\right\rangle + \frac{1}{\sqrt{2}}(\hat{J}_{1-}+\hat{J}_{2-})\left|\frac{1}{2}-\frac{1}{2}\right\rangle\left|\frac{1}{2}\frac{1}{2}\right\rangle \\ &= \frac{1}{\sqrt{2}}\cdot 2\hbar\left|\frac{1}{2}-\frac{1}{2}\right\rangle\left|\frac{1}{2}-\frac{1}{2}\right\rangle = \sqrt{2}\hbar\left|\frac{1}{2}-\frac{1}{2};\frac{1}{2}-\frac{1}{2}\right\rangle.\end{aligned}$$

$$\therefore \sqrt{2}\hbar\,|1-1\rangle = \sqrt{2}\hbar\left|\frac{1}{2}-\frac{1}{2};\frac{1}{2}-\frac{1}{2}\right\rangle$$

$$|1-1\rangle = \left|\frac{1}{2}-\frac{1}{2};\frac{1}{2}-\frac{1}{2}\right\rangle$$

The states $|10\rangle$ and $|\,00\rangle$ have to be orthogonal. So let us write

$$|00\rangle = C_1\left|\frac{1}{2}\frac{1}{2};\frac{1}{2}-\frac{1}{2}\right\rangle + C_2\left|\frac{1}{2}-\frac{1}{2};\frac{1}{2}\frac{1}{2}\right\rangle$$

The orthoganality between $|10\rangle$ and $|\,00\rangle$ is $\langle 10|00\rangle = 0$.

$$\therefore \frac{C_1}{\sqrt{2}} + \frac{C_2}{\sqrt{2}} = 0 \quad \text{or} \quad C_2 = -C_1$$

The normalization condition is

$$C_1^{\,2} + C_2^{\,2} = 1$$

$$\therefore C_1 = \pm\frac{1}{\sqrt{2}}$$

However, by phase convention, C_1 is positive.

$$\therefore C_1 = \frac{1}{\sqrt{2}} \quad \text{and} \quad C_2 = -\frac{1}{\sqrt{2}}$$

$$\therefore |00\rangle = \frac{1}{\sqrt{2}}\left|\frac{1}{2}\frac{1}{2};\frac{1}{2}-\frac{1}{2}\right\rangle - \frac{1}{\sqrt{2}}\left|\frac{1}{2}-\frac{1}{2};\frac{1}{2}\frac{1}{2}\right\rangle$$

So we have

$$\begin{aligned}|11\rangle &= \left|\frac{1}{2}\frac{1}{2};\frac{1}{2}\frac{1}{2}\right\rangle \\ |10\rangle &= \frac{1}{\sqrt{2}}\left|\frac{1}{2}\frac{1}{2};\frac{1}{2}-\frac{1}{2}\right\rangle + \frac{1}{\sqrt{2}}\left|\frac{1}{2}-\frac{1}{2};\frac{1}{2}\frac{1}{2}\right\rangle \\ |00\rangle &= \frac{1}{\sqrt{2}}\left|\frac{1}{2}\frac{1}{2};\frac{1}{2}-\frac{1}{2}\right\rangle - \frac{1}{\sqrt{2}}\left|\frac{1}{2}-\frac{1}{2};\frac{1}{2}\frac{1}{2}\right\rangle \\ |1-1\rangle &= \left|\frac{1}{2}-\frac{1}{2};\frac{1}{2}-\frac{1}{2}\right\rangle\end{aligned}$$

The Clebsh–Gordan coefficients are

$$\left\langle \tfrac{1}{2}\tfrac{1}{2};\tfrac{1}{2}\tfrac{1}{2}\middle|11\right\rangle = 1$$

$$\left\langle \tfrac{1}{2}\tfrac{1}{2};\tfrac{1}{2}-\tfrac{1}{2}\middle|10\right\rangle = \frac{1}{\sqrt{2}} \; ; \; \left\langle \tfrac{1}{2}-\tfrac{1}{2};\tfrac{1}{2}\tfrac{1}{2}\middle|10\right\rangle = \frac{1}{\sqrt{2}}$$

$$\left\langle \tfrac{1}{2}\tfrac{1}{2};\tfrac{1}{2}-\tfrac{1}{2}\middle|00\right\rangle = \frac{1}{\sqrt{2}} \quad \left\langle \tfrac{1}{2}-\tfrac{1}{2};\tfrac{1}{2}\tfrac{1}{2}\middle|00\right\rangle = -\frac{1}{\sqrt{2}}$$

The transformation matrix is given below:

Table 10.2 Clebsh–Gordan coefficients for $j_1 = 1/2$ and $j_2 = 1/2$

	$\lvert 11\rangle$	$\lvert 10\rangle$	$\lvert 00\rangle$	$\lvert 1-1\rangle$
$\left\langle \tfrac{1}{2}\tfrac{1}{2};\tfrac{1}{2}\tfrac{1}{2}\right\rvert$	1			
$\left\langle \tfrac{1}{2}\tfrac{1}{2};\tfrac{1}{2}-\tfrac{1}{2}\right\rvert$		$1/\sqrt{2}$	$1/\sqrt{2}$	
$\left\langle \tfrac{1}{2}-\tfrac{1}{2};\tfrac{1}{2}\tfrac{1}{2}\right\rvert$		$1/\sqrt{2}$	$-1/\sqrt{2}$	
$\left\langle \tfrac{1}{2}-\tfrac{1}{2};\tfrac{1}{2}-\tfrac{1}{2}\right\rvert$				1

Example 10.10 Determine the Clebsh–Gordan coefficient when two angular momenta of quantum numbers $j_1 = 1$ and $j_2 = 1/2$ are added.

Solution: The possible values of $j = \frac{3}{2}$ and $j = \frac{1}{2}$.

So we have six states.

$$\left|\tfrac{3}{2}\tfrac{3}{2}\right\rangle, \left|\tfrac{3}{2}\tfrac{1}{2}\right\rangle, \left|\tfrac{3}{2}-\tfrac{1}{2}\right\rangle, \left|\tfrac{3}{2}-\tfrac{3}{2}\right\rangle \quad \text{and} \quad \left|\tfrac{1}{2}\tfrac{1}{2}\right\rangle, \left|\tfrac{1}{2}-\tfrac{1}{2}\right\rangle$$

By phase convention,

$$\left\langle 11;\tfrac{1}{2}\tfrac{1}{2}\middle|\tfrac{3}{2}\tfrac{3}{2}\right\rangle \text{ and } \left\langle 11;\tfrac{1}{2}-\tfrac{1}{2}\middle|\tfrac{1}{2}\tfrac{1}{2}\right\rangle \text{ are positive.}$$

$$\left|\tfrac{3}{2}\tfrac{3}{2}\right\rangle = \left\langle 11;\tfrac{1}{2}\tfrac{1}{2}\middle|\tfrac{3}{2}\tfrac{3}{2}\right\rangle \left|11;\tfrac{1}{2}\tfrac{1}{2}\right\rangle$$

$$\left|\tfrac{3}{2}\tfrac{1}{2}\right\rangle = \left\langle 11;\tfrac{1}{2}-\tfrac{1}{2}\middle|\tfrac{3}{2}\tfrac{1}{2}\right\rangle \left|11;\tfrac{1}{2}-\tfrac{1}{2}\right\rangle + \left\langle 10;\tfrac{1}{2}\tfrac{1}{2}\middle|\tfrac{3}{2}\tfrac{1}{2}\right\rangle \left|10;\tfrac{1}{2}\tfrac{1}{2}\right\rangle$$

$$\left|\tfrac{3}{2}-\tfrac{1}{2}\right\rangle = \left\langle 10;\tfrac{1}{2}-\tfrac{1}{2}\middle|\tfrac{3}{2}-\tfrac{1}{2}\right\rangle \left|10;\tfrac{1}{2}-\tfrac{1}{2}\right\rangle + \left\langle 1-1;\tfrac{1}{2}\tfrac{1}{2}\middle|\tfrac{3}{2}-\tfrac{1}{2}\right\rangle \left|1-1;\tfrac{1}{2}\tfrac{1}{2}\right\rangle$$

$$\left|\tfrac{3}{2}-\tfrac{3}{2}\right\rangle = \left\langle 1-1;\tfrac{1}{2}-\tfrac{1}{2}\middle|\tfrac{3}{2}-\tfrac{3}{2}\right\rangle \left|1-1;\tfrac{1}{2}-\tfrac{1}{2}\right\rangle$$

$$\left|\tfrac{1}{2}\tfrac{1}{2}\right\rangle = \left\langle 11;\tfrac{1}{2}-\tfrac{1}{2}\middle|\tfrac{1}{2}\tfrac{1}{2}\right\rangle \left|11;\tfrac{1}{2}-\tfrac{1}{2}\right\rangle + \left\langle 10;\tfrac{1}{2}\tfrac{1}{2}\middle|\tfrac{1}{2}\tfrac{1}{2}\right\rangle \left|10;\tfrac{1}{2}\tfrac{1}{2}\right\rangle$$

$$\left|\tfrac{1}{2}-\tfrac{1}{2}\right\rangle = \left\langle 10;\tfrac{1}{2}-\tfrac{1}{2}\middle|\tfrac{1}{2}-\tfrac{1}{2}\right\rangle \quad \left|10;\tfrac{1}{2}-\tfrac{1}{2}\right\rangle + \left\langle 1-1;\tfrac{1}{2}\tfrac{1}{2}\middle|\tfrac{1}{2}-\tfrac{1}{2}\right\rangle \quad \left|1-1;\tfrac{1}{2}\tfrac{1}{2}\right\rangle.$$

$$\left|\tfrac{3}{2}\tfrac{3}{2}\right\rangle = \left|11;\tfrac{1}{2}\tfrac{1}{2}\right\rangle.$$

$$\hat{J}_-\left|\tfrac{3}{2}\tfrac{3}{2}\right\rangle = (\hat{J}_{1-}+\hat{J}_{2-})\,|11\rangle\left|\tfrac{1}{2}\tfrac{1}{2}\right\rangle$$

$$\left|\tfrac{3}{2}\tfrac{1}{2}\right\rangle = \frac{1}{\sqrt{3}}\,|11\rangle\left|\tfrac{1}{2}-\tfrac{1}{2}\right\rangle + \sqrt{\frac{2}{3}}|10\rangle\left|\tfrac{1}{2}\tfrac{1}{2}\right\rangle$$

$$\hat{J}_-\left|\tfrac{3}{2}\tfrac{1}{2}\right\rangle = (\hat{J}_{1-}+\hat{J}_{2-})\left[\frac{1}{\sqrt{3}}\,|11\rangle\left|\tfrac{1}{2}-\tfrac{1}{2}\right\rangle + \sqrt{\frac{2}{3}}\,|10\rangle\left|\tfrac{1}{2}\tfrac{1}{2}\right\rangle\right]$$

$$\left|\tfrac{3}{2}-\tfrac{1}{2}\right\rangle = \sqrt{\frac{2}{3}}\,|10\rangle\left|\tfrac{1}{2}-\tfrac{1}{2}\right\rangle + \frac{1}{\sqrt{3}}\,|1-1\rangle\left|\tfrac{1}{2}\tfrac{1}{2}\right\rangle$$

$$\left|\tfrac{3}{2}-\tfrac{3}{2}\right\rangle = |1-1\rangle\left|\tfrac{1}{2}-\tfrac{1}{2}\right\rangle$$

Let us write $\left|\tfrac{1}{2}\tfrac{1}{2}\right\rangle$ as

$$\left|\tfrac{1}{2}\tfrac{1}{2}\right\rangle = C_1\,|11\rangle\left|\tfrac{1}{2}-\tfrac{1}{2}\right\rangle + C_2\,|10\rangle\left|\tfrac{1}{2}\tfrac{1}{2}\right\rangle$$

This state is orthogonal to $\left|\tfrac{3}{2}\tfrac{1}{2}\right\rangle$

$$\therefore \left\langle\tfrac{3}{2}\tfrac{1}{2}\middle|\tfrac{1}{2}\tfrac{1}{2}\right\rangle = 0 \Rightarrow \frac{C_1}{\sqrt{3}} + \sqrt{\frac{2}{3}}C_2 = 0 \quad C_2 = -\frac{C_1}{\sqrt{2}}$$

The normalization condition is

$$C_1^2 + C_2^2 = 1$$

$$\therefore C_1 = \sqrt{\frac{2}{3}}$$

By phase convention, $C_1 = \sqrt{\frac{2}{3}}$

$$\therefore \left|\tfrac{1}{2}\tfrac{1}{2}\right\rangle = \sqrt{\frac{2}{3}}\,|11\rangle\left|\tfrac{1}{2}-\tfrac{1}{2}\right\rangle - \sqrt{\frac{1}{3}}\,|10\rangle\left|\tfrac{1}{2}\tfrac{1}{2}\right\rangle$$

By operating $\hat{J}_-$ or $\left|\tfrac{1}{2}\tfrac{1}{2}\right\rangle$,we get

$$\left|\tfrac{1}{2}-\tfrac{1}{2}\right\rangle = \frac{1}{\sqrt{3}}|10\rangle\left|\tfrac{1}{2}-\tfrac{1}{2}\right\rangle - \sqrt{\frac{2}{3}}\,|1-1\rangle\left|\tfrac{1}{2}\tfrac{1}{2}\right\rangle$$

The Clebsh–Gordan coefficients are given below:

$$\left\langle 11;\tfrac{1}{2}\tfrac{1}{2}\middle|\tfrac{3}{2}\tfrac{3}{2}\right\rangle = 1$$

$$\left\langle 11;\frac{1}{2}-\frac{1}{2}\middle|\frac{3}{2}\frac{1}{2}\right\rangle=\frac{1}{\sqrt{3}}\qquad\left\langle 10;\frac{1}{2}\frac{1}{2}\middle|\frac{3}{2}\frac{1}{2}\right\rangle=\sqrt{\frac{2}{3}}$$

$$\left\langle 11;\frac{1}{2}-\frac{1}{2}\middle|\frac{1}{2}\frac{1}{2}\right\rangle=\sqrt{\frac{2}{3}}\qquad\left\langle 10;\frac{1}{2}\frac{1}{2}\middle|\frac{1}{2}\frac{1}{2}\right\rangle=-\frac{1}{\sqrt{3}}$$

$$\left\langle 10;\frac{1}{2}-\frac{1}{2}\middle|\frac{1}{2}-\frac{1}{2}\right\rangle=\frac{1}{\sqrt{3}}\qquad\left\langle 1-1;\frac{1}{2}\frac{1}{2}\middle|\frac{1}{2}-\frac{1}{2}\right\rangle=-\sqrt{\frac{2}{3}}$$

$$\left\langle 1-1;\frac{1}{2}-\frac{1}{2}\middle|\frac{3}{2}-\frac{3}{2}\right\rangle=1$$

The transformation matrix is given below:

Table 10.3 Clebsh–Gordan coefficients for $j_1 = 1$ and $j_2 = 1/2$

	$\left\|\frac{3}{2}\frac{3}{2}\right\rangle$	$\left\|\frac{3}{2}\frac{1}{2}\right\rangle$	$\left\|\frac{1}{2}\frac{1}{2}\right\rangle$	$\left\|\frac{3}{2}-\frac{1}{2}\right\rangle$	$\left\|\frac{1}{2}-\frac{1}{2}\right\rangle$	$\left\|\frac{3}{2}-\frac{3}{2}\right\rangle$
$\left\langle 11;\frac{1}{2}\frac{1}{2}\right\|$	1					
$\left\langle 11;\frac{1}{2}-\frac{1}{2}\right\|$		$\frac{1}{\sqrt{3}}$	$\sqrt{\frac{2}{3}}$			
$\left\langle 10;\frac{1}{2}\frac{1}{2}\right\|$		$\sqrt{\frac{2}{3}}$	$-\frac{1}{\sqrt{3}}$			
$\left\langle 10;\frac{1}{2}-\frac{1}{2}\right\|$				$\sqrt{\frac{2}{3}}$	$\frac{1}{\sqrt{3}}$	
$\left\langle 1-1;\frac{1}{2}\frac{1}{2}\right\|$				$\sqrt{\frac{1}{3}}$	$-\sqrt{\frac{2}{3}}$	
$\left\langle 1-1;\frac{1}{2}-\frac{1}{2}\right\|$						1

Example 10.11 Obtain the Clebsh–Gordan coefficient for all the states belonging $j = 2$, resulting from the addition of angular momenta with quantum number $j_1 = 1$ and $j_2 = 1$.

Solution:

$$\begin{aligned}
|22\rangle &= \langle 11;11|22\rangle\,|11;11\rangle\\
|21\rangle &= \langle 11;10|21\rangle\,|11;10\rangle+\langle 10;11|21\rangle\,|10;11\rangle\\
|20\rangle &= \langle 11;1-1|20\rangle\,|11;1-1\rangle+\langle 10;10|20\rangle\,|10;10\rangle+\langle 1-1;11|20\rangle\,|1-1;11\rangle\\
|2-1\rangle &= \langle 10;1-1|2-1\rangle\,|10;1-1\rangle+\langle 1-1;10|2-1\rangle\,|1-1;10\rangle\\
|2-2\rangle &= \langle 1-1;1-1|2-2\rangle\,|1-1;1-1\rangle
\end{aligned}$$

Let us use the following recursion relation.

$$\begin{aligned}
&(j+m)^{1/2}(j-m+1)^{1/2}\langle j_1m_1;\,j_2m_2\,|\,jm-1\rangle\\
&=(j_1-m_1)^{1/2}(j_1+m_1+1)^{1/2}\langle j_1\,\overline{m_1+1};\,j_2m_2\,|\,jm\rangle+(j_2-m_2)^{1/2}(j_2+m_2+1)^{1/2}\langle j_1m_1;\,j_2\,\overline{m_2+1}|\,jm\rangle
\end{aligned}$$

Let us start with $\langle 11;11|22\rangle$

$$\langle 11;11|22\rangle = 1$$

Substituting the values $m_1 = 1; m_2 = 0; m = 2$ in the recursion relation, we get

$$(2+2)^{1/2}.1.\langle 11;10|21\rangle = 0\times\sqrt{2}\langle 12;10|22\rangle + \sqrt{2}\langle 11;11|22\rangle$$

$$= \sqrt{2}$$

$$\therefore \langle 11;0\,|\,11\rangle = \frac{1}{\sqrt{2}}$$

$$m_1 = 0; m_2 = 1; m = 2$$

$$(2+2)^{1/2}.1.\langle 0;1|21\rangle = \sqrt{2}.\langle 11;11|22\rangle + 0\times\sqrt{2}\langle 12;11|22\rangle$$

$$\therefore \langle 10;11\,|\,21\rangle = \frac{1}{\sqrt{2}}$$

$$m_1 = 1; m_2 = -1; m = 1$$

$$\sqrt{3}.\sqrt{2}\langle 1;-1|20\rangle = 0.\sqrt{3}.\langle 12;1-1|21\rangle + 1.\sqrt{2}.\langle 11;10|21\rangle$$

$$\therefore \langle 11;-1\,|\,20\rangle = \frac{1}{\sqrt{6}}$$

$$m_1 = 0; m_2 = 0; m = 1$$

$$\sqrt{3}.\sqrt{2}\langle 10;10|20\rangle = \sqrt{2}.\langle 11;10|21\rangle + \sqrt{2}\langle 10;11|21\rangle$$

$$\therefore \langle 10;10\,|\,20\rangle = \sqrt{\frac{2}{3}}$$

$$m_1 = -1; m_2 = 1; m = 1$$

$$\sqrt{3}.\sqrt{2}\langle 1-1;11|20\rangle = \sqrt{2}.\langle 10;11|21\rangle + 0.\sqrt{3}\langle 1-1;12|21\rangle$$

$$\therefore \langle 1-1;11|20\rangle = \frac{1}{\sqrt{6}}$$

$$m_1 = 0; m_2 = -1; m = 0$$

$$\sqrt{2}\cdot\sqrt{3}\langle 10;1-1|2-1\rangle = \sqrt{2}\langle 11;1-1|20\rangle + \sqrt{2}\langle 10;10|20\rangle$$

$$\therefore \langle 10;1-1\,|\,2-1\rangle = \frac{1}{\sqrt{2}}$$

$$m_1 = -1; m_2 = 0; m = 0$$

$$\sqrt{2}\cdot\sqrt{3}\langle 1-1;10|2-1\rangle = \sqrt{2}\langle 10;10|20\rangle + \sqrt{2}\langle 1-1;11|20\rangle$$

$$\therefore \langle 1-1;10|2-1\rangle = \frac{1}{\sqrt{2}}$$

$$m_1 = -1; m_2 = -1; m = -1$$

$$1\cdot 2\cdot \langle 1-1;1-1|2-2\rangle = \sqrt{2}\langle 10;1-1|2-1\rangle + \sqrt{2}\langle 1-1;10|2-1\rangle$$

$$\therefore \langle 1-1;1-1|2-2\rangle = 1.$$

Example 10.12 Obtain the general expression for the Clebsh–Gordan coefficient when two angular momenta of quantum numbers j_1 and $j_2 = 1/2$ are added.

Solution: The possible j values are $j_1 + 1/2$ and $j_1 - 1/2$.

$$\left|j_1+\frac{1}{2}m\right\rangle = \left\langle j_1\ m-\frac{1}{2};\frac{1}{2}\frac{1}{2}\middle|j_1+\frac{1}{2}m\right\rangle \left|j_1 m-\frac{1}{2};\frac{1}{2}\frac{1}{2}\right\rangle + \left\langle j_1 m+\frac{1}{2};\frac{1}{2}-\frac{1}{2}\middle|j_1+\frac{1}{2}m\right\rangle \left|j_1 m+\frac{1}{2};\frac{1}{2}-\frac{1}{2}\right\rangle$$

$$\left|j_1-\frac{1}{2}m\right\rangle = \left\langle j_1\ m-\frac{1}{2};\frac{1}{2}\middle|j_1-\frac{1}{2}m\right\rangle \left|j_1 m-\frac{1}{2};\frac{1}{2}\frac{1}{2}\right\rangle + \left\langle j_1 m+\frac{1}{2};\frac{1}{2}-\frac{1}{2}\middle|j_1-\frac{1}{2}m\right\rangle \left|j_1 m+\frac{1}{2};\frac{1}{2}-\frac{1}{2}\right\rangle$$

Let us simplify these equations as follows:

$$\left|j_1+\frac{1}{2}m\right\rangle = C_1\left|j_1 m-\frac{1}{2};\frac{1}{2}\frac{1}{2}\right\rangle + C_2\left|j_1 m+\frac{1}{2};\frac{1}{2}-\frac{1}{2}\right\rangle$$

$$\left|j_1-\frac{1}{2}m\right\rangle = C_3\left|j_1 m-\frac{1}{2};\frac{1}{2}\frac{1}{2}\right\rangle + C_4\left|j_1 m+\frac{1}{2};\frac{1}{2}-\frac{1}{2}\right\rangle$$

Let us first consider the equation.

The normalization condition for this state is

$$C_1^2 + C_2^2 = 1.$$

Let us recall

$$\hat{J}^2 = \hat{J}_1^2 + \hat{J}_2^2 + \hat{J}_{1+}\hat{J}_{2-} + \hat{J}_{1-}\hat{J}_{2+} + 2\hat{J}_{1z}\hat{J}_{2z}$$

$$\hat{J}^2\left|j_1+\frac{1}{2}m\right\rangle = C_1\hat{J}^2\left|j_1\ m-\frac{1}{2};\frac{1}{2}\frac{1}{2}\right\rangle + C_2\hat{J}^2\left|j_1\ m+\frac{1}{2};\frac{1}{2}-\frac{1}{2}\right\rangle$$

$$\text{LHS} = \hat{J}^2\left|j_1+\frac{1}{2}m\right\rangle = \left(j_1+\frac{1}{2}\right)\left(j_1+\frac{3}{2}\right)\hbar^2\left|j_1+\frac{1}{2}m\right\rangle$$

Note that $\hat{J}_{2+}\left|\frac{1}{2}\frac{1}{2}\right\rangle = 0\ \ \hat{J}_{2-}\left|\frac{1}{2}-\frac{1}{2}\right\rangle = 0$

$$\text{RHS} = C_1[\hat{J}_1^2 + \hat{J}_2^2 + 2\hat{J}_{1z}\hat{J}_{2z}]\left|j_1\ m-\frac{1}{2};\frac{1}{2}\frac{1}{2}\right\rangle + C_1\hat{J}_{1+}\hat{J}_{2-}\left|j_1 m-\frac{1}{2};\frac{1}{2}\frac{1}{2}\right\rangle$$

$$+ C_2[\hat{J}_1^2 + \hat{J}_2^2 + 2\hat{J}_{1z}\hat{J}_{2z}]\left|j_1 m+\frac{1}{2};\frac{1}{2}-\frac{1}{2}\right\rangle + C_2\hat{J}_{1-}\hat{J}_{2+}\left|j_1 m+\frac{1}{2};\frac{1}{2}-\frac{1}{2}\right\rangle$$

$$= C_1\left[j_1(j_1+1)+m+\frac{1}{4}\right]\left|j_1\, m-\frac{1}{2};\frac{1}{2}\,\frac{1}{2}\right\rangle + C_1\left(j_1-m+\frac{1}{2}\right)^{1/2}\left(j_1+m+\frac{1}{2}\right)^{1/2}\left|j_1\, m+\frac{1}{2};\frac{1}{2}-\frac{1}{2}\right\rangle$$
$$+C_2\left[j_1(j_1+1)-m+\frac{1}{4}\right]\left|j_1\, m+\frac{1}{2};\frac{1}{2}-\frac{1}{2}\right\rangle + C_2\left[\left(j_1+m+\frac{1}{2}\right)\left(j_1-m+\frac{1}{2}\right)\right]^{1/2}\left|j_1\, m-\frac{1}{2};\frac{1}{2}\,\frac{1}{2}\right\rangle$$

Let us equate LHS and RHS.

From the terms involving $\left|j_1\, m+\frac{1}{2};\frac{1}{2}-\frac{1}{2}\right\rangle$ we get

$$\left(j_1-m+\frac{1}{2}\right)C_1 - C_2\left[\left(j_1+\frac{1}{2}\right)^2-m^2\right]^{1/2} = 0$$

From the terms involving $\left|j_1\, m-\frac{1}{2};\frac{1}{2}\,\frac{1}{2}\right\rangle$ we get

$$-\left[\left(j_1+\frac{1}{2}\right)^2-m^2\right]^{1/2} C_2 + C_1\left(j_1+m+\frac{1}{2}\right) = 0$$

$$C_2 = \frac{\left(j_1-m+\frac{1}{2}\right)^{1/2}}{\left(j_1+m+\frac{1}{2}\right)^{1/2}}C_1 \tag{10.74}$$

The normalization condition is

$$C_1^2 + C_2^2 = 1$$

Substituting the Equation (10.74) in the normalization condition, we get

$$C_1^2 = \left[\frac{j_1+m+\frac{1}{2}}{2j_1+1}\right]^{1/2}$$

$$C_1 = \pm\left[\frac{j_1+m+\frac{1}{2}}{2j_1+1}\right]^{1/2}$$

i.e.,
$$\left\langle j_1\, m-\frac{1}{2};\frac{1}{2}\,\frac{1}{2}\right|\left.j_1+\frac{1}{2}\; m\right\rangle = \pm\left[\frac{j_1+m+\frac{1}{2}}{2j_1+1}\right]^{1/2} \tag{10.75}$$

To choose the correct sign, let us consider the state $\left|j_1+\frac{1}{2}\, j_1+\frac{1}{2}\right\rangle$ which is given by

$$\left|j_1+\frac{1}{2}\; j_1+\frac{1}{2}\right\rangle = \left\langle j_1 j_1;\frac{1}{2}\,\frac{1}{2}\right|\left.j_1+\frac{1}{2}\; j_1+\frac{1}{2}\right\rangle\left|j_1 j_1;\frac{1}{2}\,\frac{1}{2}\right\rangle$$

$$\therefore \left\langle j_1 j_1; \frac{1}{2}\frac{1}{2}\middle| j_1+\frac{1}{2}\, j_1+\frac{1}{2}\right\rangle = 1$$

∴ This is consistent with the choice of positive sign in (10.75) as can be easily checked below. From (10.75), we have

$$\left[\frac{j_1+m+\frac{1}{2}}{2j_1+1}\right]^{1/2} = \left[\frac{j_1+j_1+\frac{1}{2}+\frac{1}{2}}{2j_1+1}\right]^{1/2} = 1$$

$$\therefore \left\langle j_1 m-\frac{1}{2}; \frac{1}{2}\frac{1}{2}\middle| j_1+\frac{1}{2}\, m\right\rangle = +\left[\frac{j_1+m+\frac{1}{2}}{2j_1+1}\right]^{1/2}$$

From (10.74), we have

$$\left\langle j_1\, m+\frac{1}{2}; \frac{1}{2}-\frac{1}{2}\middle| j_1+\frac{1}{2}\, m\right\rangle = \left[\frac{j_1-m+\frac{1}{2}}{2j_1+1}\right]^{1/2}$$

Similarly, the other Clebsh–Gordan coefficient can be calculated.

$$\left\langle j_1 m-\frac{1}{2}; \frac{1}{2}\frac{1}{2}\middle| j_1-\frac{1}{2}\, m\right\rangle = -\left[\frac{\left(j_1-m+\frac{1}{2}\right)}{2j_1+1}\right]^{1/2}$$

$$\left\langle j_1 m+\frac{1}{2}; \frac{1}{2}-\frac{1}{2}\middle| j_1-\frac{1}{2}\, m\right\rangle = \left[\frac{\left(j_1+m+\frac{1}{2}\right)}{(2j_1+1)}\right]^{1/2}$$

These results are tabulated in a table.

Table 10.4 Clebsh–Gordan coefficients $\langle j_1\, \overline{m-m_2};\ \frac{1}{2}\, m_2 \,|\, jm\rangle$ for $j_1 = j_1$ and $\hat{j}_2 = 1/2$

$m_{2\rightarrow}$ $j_{\downarrow}$	$m_2 = \frac{1}{2}$ $m_1 = m - \frac{1}{2}$	$m_2 = \frac{-1}{2}$ $m_1 = m + \frac{1}{2}$
$j_1 + \frac{1}{2}$	$\left[\frac{j_1+m+\frac{1}{2}}{2j_1+1}\right]^{1/2}$	$\left[\frac{j_1-m+\frac{1}{2}}{2j_1+1}\right]^{1/2}$
$j_1 - \frac{1}{2}$	$-\left[\frac{j_1-m+\frac{1}{2}}{2j_1+1}\right]^{1/2}$	$\left[\frac{j_1+m+\frac{1}{2}}{2j_1+1}\right]^{1/2}$

In the same way we can determine the expressions for C.G. coefficients in the additions of angular momenta corresponding to $j = j_1$ and $j = 1$

Table 10.5 Clebsh–Gordan coefficients $\langle j_1\ \overline{m - m_2};\ 1 m_2 | jm\rangle$ for $j_1 = j_1$ and $j_2 = 1$

j \ $m_{2\rightarrow}$	$m_2 = 1$ $m_1 = m - 1$	$m_2 = 0$ $m_1 = m$	$m_2 = -1$ $m_1 = m + 1$
$j_1 + 1$	$\left[\frac{(j_1+m)(j_1+m+1)}{[(2j_1+1)(2j_1+2)]}\right]^{1/2}$	$\left[\frac{(j_1-m+1)(j_1+m+1)}{[(2j_1+1)(j_1+1)]}\right]^{1/2}$	$\left[\frac{(j_1-m)(j_1-m+1)}{(2j_1+1)(2j_1+2)}\right]^{1/2}$
j_1	$-\left[\frac{(j_1+m)(j_1-m+1)}{2j_1(j_1+1)}\right]^{1/2}$	$\frac{m}{[j_1(j_1+1)]^{1/2}}$	$\left[\frac{(j_1-m)(j_1+m+1)}{2j_1(j_1+1)}\right]^{1/2}$
$j_2 - 1$	$\left[\frac{(j_1-m)(j_1-m+1)}{2j_1(2j_1+1)}\right]^{1/2}$	$-\left[\frac{(j_1-m)(j_1+m)}{j_1(2j_1+1)}\right]^{1/2}$	$\left[\frac{(j_1+m+1)(j_1+m)}{2j_1(2j_1+1)}\right]^{1/2}$

10.13 TRANSFORMATIONS OF OPERATORS UNDER ROTATIONS

We have already seen under rotation transformation,

$$|\psi\rangle \rightarrow |\psi'\rangle = \hat{U}(R)|\psi\rangle$$

and
$$\hat{A} \rightarrow \hat{A}' = \hat{U}(R)\hat{A}\,\hat{U}^{\dagger}(R) \tag{10.76}$$

The unitary transformation $\hat{U}(R)$ is given by $\hat{U}(R) = e^{-i\theta\mathbf{n}.\hat{\mathbf{J}}/\hbar}$.

$\hat{U}(R)\hat{A}\,\hat{U}^{\dagger}(R)$ may be $\hat{A}$ itself. This need not be true for all operators. In general, one studies the transformation properties of a group of operators $\hat{A}_1, \hat{A}_2, \ldots$ In such cases, $\hat{U}(R)\hat{A}\hat{U}^{\dagger}(R)$ may be a linear combination of the operators within this group.

The operators are now classified into different kinds based on the relationship between $\hat{A}_i$ and $\hat{U}(R)\hat{A}_i\hat{U}^{\dagger}(R)$.

Scalar Operator

An operator $\hat{S}$ is a scalar operator if it satisfies

$$\hat{U}(R)\hat{S}\hat{U}^{\dagger}(R) = \hat{S} \tag{10.77}$$

$$\hat{U}(R)\hat{S} = \hat{S}\,\hat{U}(R)$$
$$\therefore e^{-i\theta\mathbf{n}.\hat{\mathbf{J}}/\hbar}\hat{S} = \hat{S}e^{-i\theta\mathbf{n}.\hat{\mathbf{J}}/\hbar} \tag{10.78}$$

This implies that
$$[\hat{S}, \hat{J}_x] = 0, [\hat{S}, \hat{J}_y] = 0, [\hat{S}, \hat{J}_z] = 0 \tag{10.79}$$

These relations can be used to verify whether a particular operator is a scalar operator or not under rotation.

A simple property of a scalar operator is that its eigenvalues are independent of the quantum number m. This can be easily seen as follows:

Since $\hat{S}$ commutes with all the three angular momentum operators $\hat{J}_x, \hat{J}_y$ and $\hat{J}_z$, all the eigenfunctions of $\hat{S}$ should also be simultaneously the eigenfunction of $\hat{J}_x, \hat{J}_y$ and $\hat{J}_z$. This is possible only if the eigenvalue of $\hat{S}$ is independent of m.

Let us elaborate this aspect further with Example 10.13. Let us characterize a state by a set of quantum numbers α, j and m, where α is one or more number of quantum numbers which are different from j and m. Let the state $|\psi\rangle$ and $|\phi\rangle$ be given by

$$|\psi\rangle = |\alpha jm\rangle \text{ and } |\phi\rangle = |\alpha' j' m'\rangle \tag{10.80}$$

Example 10.13 Show that $\langle \alpha jm| S |\alpha jm\rangle$ is independent of m

Solution: One way of seeing this result is that the eigenvalue of S is independent of m since it commutes with all the angular momentum operators J_x, J_y and J_z. Another way is to show that the matrix element $\langle \alpha jm| S |\alpha jm\rangle$ is the same for different values of m. Let us evaluate $\langle \alpha j m+1| S |\alpha j m+1\rangle$.

Note that $\qquad J_+ |jm\rangle = C_+ |jm+1\rangle = [(j-m)(j+m+1)]^{1/2} |jm+1\rangle$

Therefore $|jm+1\rangle$ is given by

$$|jm+1\rangle = \frac{1}{C_+} J_+ |jm\rangle$$

Similarly, we have $\qquad \langle jm+1| = \frac{1}{C_+} \langle jm+1| J_-$

Making use of these results we get

$$\langle \alpha j\overline{m+1}|S|\alpha j\overline{m+1}\rangle = \frac{1}{C_+^2} \langle \alpha jm|J_- S J_+|\alpha jm\rangle$$

Since S commutes with J_+ and J_-, we have $J_- S J_+ = S J_- J_+$.

$$\langle \alpha j\overline{m+1}| S |\alpha j\overline{m+1}\rangle = \frac{1}{C_+^2} \langle \alpha jm| J_- S J_+ |\alpha jm\rangle = \frac{1}{C_+^2} \langle \alpha jm| S J_- J_+ |\alpha jm\rangle$$
$$= \frac{1}{C_+^2} C_+^2 \langle \alpha jm| S |\alpha jm\rangle = \langle \alpha jm| S |\alpha jm\rangle$$

So, $\langle \alpha j\overline{m+1}| S |\alpha j\overline{m+1}\rangle$ and $\langle \alpha jm| S |\alpha jm\rangle$ are the same. Since it is true for any m, it is independent of m.

Example 10.14 Show that $\langle \alpha' j' m'| S |\alpha jm\rangle = C(\alpha, j)\delta_{\alpha\alpha'}\delta_{jj'}\delta_{mm'}$.

Solution: Since $[S, J^2] = 0$ for a scalar operator S, we have

$$\langle \alpha' j' m'| [S, J^2] |\alpha jm\rangle = \langle \alpha' j' m'| (SJ^2 - J^2 S) |\alpha jm\rangle$$
$$= [j(j+1) - j'(j'+1)]\hbar^2 \langle \alpha' j' m'| S |\alpha jm\rangle = 0$$

The matrix element $\langle \alpha' j' m' | S | \alpha j m \rangle = 0 \text{ if } j \neq j'$

In the same way let us make use of $[S, J_Z] = 0$ for a scalar operator

$$\begin{aligned}\langle \alpha' j' m' | [S, J_z] | \alpha j m \rangle &= \langle \alpha' j' m' | (SJ_z - J_z S) | \alpha j m \rangle \\ &= (m - m')\hbar \langle \alpha' j' m' | S | \alpha j m \rangle = 0\end{aligned}$$

The matrix element $\langle \alpha' j' m' | S | \alpha j m \rangle = 0 \text{ if } m \neq m'$

With these results and the fact that $\langle \alpha j m | S | \alpha j m \rangle$ is independent of *m*, it is obvious to write

$$\langle \alpha' j' m' | S | \alpha j m \rangle = C(\alpha, j) \delta_{\alpha\alpha'} \delta_{jj'} \delta_{mm'}$$

where $C(\alpha, j)$ is a constant independent of *m*.

Example 10.15 Show that $\frac{\mathbf{S} \cdot \mathbf{r}}{r}$ is a scalar operator under rotation transformation. (Note $\mathbf{S}$ is the spin angular momentum operator.)

Solution:

$$\mathbf{J} = \mathbf{L} + \mathbf{S}$$

Let us evaluate the following commutation relations:

$$\left[J_x, \frac{\mathbf{S} \cdot \mathbf{r}}{r} \right] = ? \quad \left[J_y, \frac{\mathbf{S} \cdot \mathbf{r}}{r} \right] = ? \quad \left[J_z, \frac{\mathbf{S} \cdot \mathbf{r}}{r} \right] = ?$$

$$\frac{\mathbf{S} \cdot \mathbf{r}}{r} = \frac{S_x x}{r} + \frac{S_y y}{r} + \frac{S_z z}{r}$$

$$\left[J_x, S_x \frac{x}{r} \right] = [J_x, S_x] \frac{x}{r} + S_x \left[J_x, \frac{x}{r} \right]$$

$$[J_x, S_x] = [L_x + S_x, S_x] = 0$$

$$\left[J_x, \frac{x}{r} \right] = \left[L_x, \frac{x}{r} \right] + \left[S_x, \frac{x}{r} \right] = 0$$

$$\therefore \left[J_x, \frac{S_x}{r} x \right] = 0$$

$$\begin{aligned}\left[J_x, S_y \frac{y}{r} \right] &= \left[L_x, S_y \frac{y}{r} \right] + \left[S_x, S_y \frac{y}{r} \right] \\ &= S_y \left[L_x, \frac{y}{r} \right] + \frac{y}{r} [S_x, S_y] = i\hbar \left[\frac{S_z y}{r} - \frac{S_y z}{r} \right]\end{aligned}$$

Similarly, we have $$\left[J_x, \frac{S_z z}{r}\right] = i\hbar\left[\frac{S_y z}{r} - \frac{S_z y}{r}\right]$$

$$\therefore \left[J_x, \frac{\mathbf{S}\cdot\mathbf{r}}{r}\right] = 0$$

Similarly, we can show that $$\left[J_y, \frac{\mathbf{S}\cdot\mathbf{r}}{r}\right] = \left[J_z, \frac{\mathbf{S}\cdot\mathbf{r}}{r}\right] = 0$$

$\therefore \frac{\mathbf{S}\cdot\mathbf{r}}{r}$ is a scalar operator under rotation.

Vector Operator

Let us now consider a set of three operators V_1, V_2 and V_3. These three operators are said to be vector operators if they obey the following transformation law:

$$\begin{aligned} |\psi\rangle &\to \hat{U}|\psi\rangle \\ \langle\psi'|V_i|\psi'\rangle &= \sum_j R_{ij}\langle\psi|V_j|\psi\rangle \end{aligned} \tag{10.81}$$

where R is the rotation matrix relating the coordinates of the vector $\mathbf{r}$ and $\mathbf{r}'$.

i.e., $$\mathbf{r}' = R\mathbf{r}$$

It can be shown that this definition leads to the following results:

$$\begin{array}{lll} [\hat{V}_x, \hat{J}_y] = i\hbar\hat{V}_z & [\hat{V}_x, \hat{J}_z] = -i\hbar\hat{V}_y & [\hat{V}_x, \hat{J}_x] = 0 \\ [\hat{V}_y, \hat{J}_z] = i\hbar\hat{V}_x & [\hat{V}_y, \hat{J}_x] = -i\hbar\hat{V}_z & [\hat{V}_y, \hat{J}_y] = 0 \\ [\hat{V}_z, \hat{J}_x] = i\hbar\hat{V}_y & [\hat{V}_z, \hat{J}_y] = -i\hbar\hat{V}_x & [\hat{V}_z, \hat{J}_z] = 0 \end{array} \tag{10.82}$$

These relations themselves can be taken as a definition of vector operators.

Example 10.16 Show that the operators $\mathbf{r}_{op}$ and $\mathbf{p}_{op}$ are vector operators under rotation.

Solution:

$$\mathbf{r}_{op} = \hat{\boldsymbol{x}}\, x_{op} + \hat{\boldsymbol{y}}\, y_{op} + \hat{\boldsymbol{z}}\, z_{op}$$

We now have a set of three operators: x_{op}, y_{op} and z_{op}.

It is easy to check the following commutation relations:

$$\begin{array}{lll} [L_x, x] = 0 & [L_y, x] = -i\hbar z & [L_z, x] = i\hbar y \\ [L_x, y] = i\hbar z & [L_y, y] = 0 & [L_z, y] = -i\hbar x \\ [L_x, z] = -i\hbar y & [L_y, z] = i\hbar x & [L_z, z] = 0 \end{array}$$

For instance, $[L_x, y] = [yp_z - zp_y, y] = -z[p_y, y] = i\hbar z$

$\therefore$ x_{op}, y_{op} and z_{op} form a vector operator.

In the same way, the operators p_x, p_y and p_z together form a vector operator.

10.14 SPHERICAL TENSOR OPERATORS T_q^k

This classification emphasizes a different aspect of rotation transformation. The definition of spherical tensor requires the definition of rotation matrix $D^j_{m'm}$.

The unitary transformation corresponding to rotation is given by

$$\hat{U} = e^{-i\theta\,\mathbf{n}.\hat{\mathbf{J}}/\hbar}$$

We can build matrix representation of this operator. Let us choose the angular momentum state vectors $|jm\rangle$ as basis states to construct the matrices representing $\hat{U}$. There are $(2j+1)$ basis states, and so the matrices representing $\hat{U}$ are $(2j+1)$ dimensional. The rotation matrix $D^j_{m'm}$ is defined as

$$D^j_{m'm} = \langle jm'|\,\hat{U}\,|jm\rangle = \langle jm'|\,e^{-i\theta\,\mathbf{n}.\hat{\mathbf{J}}/\hbar}\,|jm\rangle \tag{10.83}$$

$D^j_{m'm}$ is $(m'm)^{\text{th}}$ the elements of the matrix $e^{-i\theta\,\mathbf{n}.\mathbf{J}/\hbar}$.

In the spherical tensors T_q^k, the indices k and q are very similar to j and m. The components of T_q^k are

$$\underbrace{T^k_K, T^k_{K-1}, T^k_{K-2} \ldots . T^k_{-K}}_{(2k+1)\text{ components}}$$

For instance, a second-rank spherical tensor T_q^2 is given by a set of five operators.

$$T_2^2, T_1^2, T_0^2, T_{-1}^2, T_{-2}^2$$

A set of $(2k+1)$ operators forms a spherical tensor of rank k if it transforms as given below:

$$\hat{U}\hat{T}_q^k\ \hat{U}^\dagger = \sum_{q'} \hat{T}^k_{q'} D^k_{q'q} \tag{10.84}$$

This definition leads to the following results:

$$[J_z, T_q^k] = q\hbar\ T_q^k \tag{10.85}$$

$$[J_+, T_q^k] = [(k-q)(k+q+1)]^{1/2}\, T^k_{q+1} \tag{10.86}$$

$$[J_-, T_q^k] = [(k+q)(k-q+1)]^{1/2}\, T^k_{q+1} \tag{10.87}$$

These results themselves can be taken as definition of spherical tensors of rank k.

Example 10.17 Show that $T_1^1 = -\frac{1}{\sqrt{2}}(x+iy)$, $T_0^1 = z$ and $T_{-1}^1 = \frac{1}{\sqrt{2}}(x-iy)$ form a spherical tensor of rank 1.

Solution: We have to verify whether the following results are true or not.

$$\begin{array}{lll}
[J_z, T_1^1] = \hbar T_1^1 & [J_z, T_0^1] = 0 & [J_z, T_{-1}^1] = -\hbar T_{-1}^1 \\
[J_+, T_1^1] = 0 & [J_+, T_0^1] = \sqrt{2}\hbar T_1^1 & [J_+, T_{-1}^1] = \sqrt{2}\hbar T_0^1 \\
[J_-, T_1^1] = \sqrt{2}\hbar T_0^1 & [J_-, T_0^1] = \sqrt{2}\hbar T_{-1}^1 & [J_-, T_{-1}^1] = 0
\end{array}$$

Let us evaluate $[J_z, T_1^1]$.

$$\begin{aligned}
[J_z, T_1^1] &= \left[L_z, -\frac{1}{\sqrt{2}}(x+iy)\right] = -\frac{1}{\sqrt{2}}[L_z, x] - \frac{i}{\sqrt{2}}[L_z, y] \\
&= -\frac{\hbar}{\sqrt{2}}[x+iy] = \hbar T_1^1
\end{aligned}$$

The other relations can be evaluated in the same way.

Example 10.18 Show that the components $T_1^1 = -\frac{1}{\sqrt{2}}(J_x + iJ_y), T_0^1 = J_z$ and $T_{-1}^1 = \frac{1}{\sqrt{2}}(J_x - iJ_y)$ form a spherical tensor of rank 1.

Solution: We have to verify whether the following relations are true or not.

$$\begin{array}{lll}
[J_z, T_1^1] = \hbar T_1^1 & [J_z, T_0^1] = 0 & [J_z, T_{-1}^1] = -\hbar T_{-1}^1 \\
[J_+, T_1^1] = 0 & [J_+, T_0^1] = \sqrt{2}\hbar T_1^1 & [J_+, T_{-1}^1] = \sqrt{2}\hbar T_0^1 \\
[J_-, T_1^1] = \sqrt{2}\hbar T_0^1 & [J_-, T_0^1] = \sqrt{2}\hbar T_{-1}^1 & [J_-, T_{-1}^1] = 0
\end{array}$$

Let us evaluate one relation

$$\begin{aligned}
[J_z, T_1^1] &= \left[J_z, -\frac{1}{\sqrt{2}}(J_x + iJ_y)\right] = -\frac{1}{\sqrt{2}}[J_z, J_x] - \frac{i}{\sqrt{2}}[J_z, J_y] \\
&= -\frac{\hbar}{\sqrt{2}}(J_x + iJ_y) = \hbar T_1^1
\end{aligned}$$

Similarly, other relations can be verified.

Example 10.19 Show that the spherical harmonics Y_{lm} form a spherical tensor of rank l.

Solution: The spherical harmonics $Y_{lm}(\theta, \phi)$ are eigenfunctions of L^2 and L_z. But they can be also viewed as operators. When they act on arbitrary function $f(\theta, \phi)$, it amounts to multiplying Y_{lm} and $f(\theta, \phi)$.

We have to verify

$$
\begin{aligned}
[L_z, Y_{lm}] &= m\hbar Y_{lm} \\
[L_+, Y_{lm}] &= [(l-m)(l+m+1)]^{1/2} Y_{lm+1} \\
[L_-, Y_{lm}] &= [(l+m)(l-m+1)]^{1/2} Y_{lm-1}
\end{aligned}
$$

Consider $[L_z, Y_{lm}]\psi$, where ψ is an arbitrary function ψ.

$$
\begin{aligned}
[L_z, Y_{lm}]\psi &= [L_z Y_{lm} - Y_{lm} L_z]\psi \\
&= L_z(Y_{lm}\psi) - Y_{lm} L_z \psi \\
&= (L_z Y_{lm})\psi + Y_{lm} L_z \psi - Y_{lm} L_z \psi \\
&= m\hbar Y_{lm}\psi \\
\therefore [L_z, Y_{lm}] &= m\hbar Y_{lm}
\end{aligned}
$$

The other relations can be verified in the same way.

10.15 WIGNER–ECKART THEOREM

Let $|\alpha jm\rangle$ and $|\alpha' j' m'\rangle$ be two-state vectors of a system where α and α' stand for quantum numbers other than the angular momentum quantum numbers j and m. Wigner–Eckart theorem states that the matrix elements of spherical tensor T_q^k between states $|\alpha jm\rangle$ and $|\alpha' j' m'\rangle$ can be factored into two components: one component which depends only geometrical quantum numbers q, m, m' and the other component which is independent of these quantum numbers.

$$
\langle \alpha' j' m' | T_q^k | \alpha jm\rangle = \langle jm; kq | j'm'\rangle \langle \alpha' j' \| T^k \| \alpha j'\rangle \tag{10.88}
$$

where $\langle jm; kq | j'm'\rangle$ is the Clebsh–Gordan coefficient and $\langle \alpha' j' \| T^k \| \alpha j\rangle$ is called reduced matrix element which is independent of angular momentum quantum numbers m, m' and q. $\langle \alpha' j' \| T^k \| \alpha j\rangle$ is a symbolic representation of a factor of the matrix element which is independent of the quantum number m, m' and q.

Before analyzing the matrix elements of spherical terms, let us consider the following relations:

$$
\mathbf{J} + \mathbf{K} = \mathbf{J}'
$$

where $\mathbf{J}$, $\mathbf{K}$ and $\mathbf{J}'$ are angular momentum operators.

$$
\begin{aligned}
\hat{J}^2 |jm\rangle &= j(j+1)\hbar^2 |jm\rangle & \hat{J}_z |jm\rangle &= m\hbar |jm\rangle \\
\hat{K}^2 |kq\rangle &= k(k+1)\hbar^2 |kq\rangle & \hat{K}_z |kq\rangle &= q\hbar |kq\rangle \\
\hat{J}'^2 |j'm'\rangle &= j'(j'+1)\hbar^2 |j'm'\rangle & \hat{J}'_z |j'm'\rangle &= m'\hbar |j'm'\rangle
\end{aligned}
$$

The recursion relations (10.72) and (10.73) between the Clebsh–Gordan coefficients now become

$$
\begin{aligned}
&(j'-m')^{1/2}(j'+m'+1)^{1/2}\langle jm; kq | j'm'+1\rangle \\
&\quad = (j+m)^{1/2}(j-m+1)^{1/2}\langle j\,\overline{m+1}; kq | j'm'\rangle + (k+q)^{1/2}(k-q+1)^{1/2}\langle jm; k\,\overline{q+1} | j'm'\rangle
\end{aligned}
\tag{10.89}
$$

$$
\begin{aligned}
&(j'+m')^{1/2}(j'-m'+1)^{1/2}\langle jm;kq|j'\,\overline{m'-1}\rangle \\
&\quad = (j-m)^{1/2}(j+m+1)^{1/2}\langle j\,\overline{m-1};kq|j'm'\rangle + (k-q)^{1/2}(k+q+1)^{1/2}\langle jm;k\,\overline{q-1}|j'm'\rangle
\end{aligned} \tag{10.90}
$$

Let us now take the commutation relation

$$
\begin{aligned}
J_zT_q^k - T_q^kJ_z &= q\hbar T_q^k \\
\langle j'm'|\,J_zT_q^k - T_q^kJ_z\,|jm\rangle &= q\hbar\langle j'm'|\,T_q^k\,|jm\rangle \\
(m-m-q)\langle j'm'\,|\,T_q^k\,|jm\rangle &= 0 \\
\therefore \langle j'm'\,|\,T_q^k\,|jm\rangle &= 0 \quad \text{unless} \quad m' = m+q
\end{aligned} \tag{10.91}
$$

This is similar to the results of the addition of angular momenta **J** and **K**.

Now, let us consider the other commutation relations.

$$
\begin{aligned}
J_+T_q^k - T_q^kJ_+ &= (k-q)^{1/2}(k+q+1)^{1/2}T_{q+1}^k \\
J_-T_q^k - T_q^kJ_- &= (k+q)^{1/2}(k-q+1)^{1/2}T_{q-1}^k
\end{aligned}
$$

These equations can be rewritten as follows:

$$
J_+T_q^k = T_q^kJ_+ + (k-q)^{1/2}(k+q+1)^{1/2}T_{q+1}^k \tag{10.92}
$$

$$
J_-T_q^k = T_q^kJ_- + (k+q)^{1/2}(k-q+1)^{1/2}T_{q-1}^k \tag{10.93}
$$

From (10.92), we have

$$
\langle j'm'\,|J_+T_q^k\,|\,jm\rangle = \langle j'm'\,|\,T_q^kJ_+\,|\,jm\rangle + (k-q)^{1/2}(k+q+1)^{1/2}\langle j'm'\,|\,T_{q+1}^k|\,jm\rangle
$$

The above equation becomes

$$
\begin{aligned}
&(j'+m')^{1/2}(j'-m'+1)^{1/2}\langle j'\,\overline{m'-1}|T_q^k\,|jm\rangle \\
&\quad = (j-m)^{1/2}(j+m+1)^{1/2}\langle j'm'|T_q^k\,|j\,\overline{m+1}\rangle + (k-q)^{1/2}(k+q+1)^{1/2}\langle j'm'|T_{q+1}^k\,|jm\rangle
\end{aligned} \tag{10.94}
$$

From (10.93), we have

$$
\langle j'm'\,|J_-T_q^k\,|\,jm\rangle = \langle j'm'\,|\,T_q^kJ_-\,|\,jm\rangle + (k+q)^{1/2}(k-q+1)^{1/2}\langle j'm' + T_{q-1}^k\,|\,jm\rangle
$$

The above equation becomes

$$
\begin{aligned}
&(j'-m')(j'+m'+1)^{1/2}\langle j'\overline{m'+1}\,|\,T_q^k\,|\,jm\rangle \\
&\quad = (j+m)^{1/2}(j'-m+1)^{1/2}\langle jm'\,|\,T_q^k\,|\,jm\rangle + (k+q)^{1/2}(k-q+1)^{1/2}\langle j'm'\,|\,T_{q-1}^k\,|\,jm\rangle
\end{aligned} \tag{10.95}
$$

The Equation (10.91) and the comparison of (10.94) with (10.89) suggests the following relations. There are three matrix elements in (10.94) and there are three Clebsh–Gordan coefficients in (10.89). The factors accompanying these matrix elements and the Clebsh–Gordan coefficients suggest to make the following one to one correspondence between them.

$$\langle j'm'|\, T_q^k \,|jm\rangle \Leftrightarrow \langle jm;kq|j'm'\rangle$$

$$\langle j'\,\overline{m'-1}|\, T_q^k \,|jm\rangle \Leftrightarrow \langle jm;kq|j'\,\overline{m'-1}\rangle$$

$$\langle j'm'|\, T_q^k \,|j\,\overline{m+1}\rangle \Leftrightarrow \langle j\,\overline{m+1};kq|j'm'\rangle$$

$$\langle j'm'|\, T_{q+1}^k \,|jm\rangle \Leftrightarrow \langle jm;k\,\overline{q+1}|j'm'\rangle$$

Note the first relation comes from Equation (10.91).In fact , the four matrix elements differ from the corresponding Clebsh–Gordan coefficients only by a constant factor C and so they can be written as follows.

$$\langle j'm'|\, T_q^k \,|jm\rangle = C\langle jm;kq|j'm'\rangle$$

$$\langle j'\,\overline{m'-1}|\, T_q^k \,|jm\rangle = C\langle jm;kq|j'\,\overline{m'-1}\rangle$$

$$\langle j'm'|\, T_q^k \,|j\,\overline{m+1}\rangle = C\langle j\,\overline{m+1};kq|j'm'\rangle$$

$$\langle j'm'|\, T_{q+1}^k \,|jm\rangle = C\langle jm;k\,\overline{q+1}|j'm'\rangle$$

Similarly, the comparision of (10.95) and (10.90) suggest

$$\langle j'\,\overline{m'+1}|\, T_q^k \,|jm\rangle = C\langle jm;kq|j'\,\overline{m'+1}\rangle$$

$$\langle j'm'|\, T_q^k \,|j\,\overline{m-1}\rangle = C\langle j\,\overline{m-1};kq|j'm'\rangle$$

$$\langle j'm'|\, T_{q-1}^k \,|jm\rangle = C\langle jm;k\,\overline{q-1}\,|j'm\rangle$$

All these relations are consistent with each other. Note that the constant C is independent of q, m and m'. Therefore, we can write this constant C as $\langle \alpha' j' \| T^k \| \alpha j\rangle$. This is a symbolic representation of the fact that $\langle j'm'|\, T_q^k \,|jm\rangle$ is a factor which is independent of q, m and m'.

$$\therefore \langle j'm'|\, T_q^k \,|jm\rangle = \langle jm;kq|j'm'\rangle\langle \alpha' j'\| T^k \|\alpha j\rangle$$

EXERCISES

1. How does spin angular momentum differ from orbital angular momentum?
2. Is the spin of electron similar to the spin of the earth?
3. Can we describe the spin of the earth using $\mathbf{r}\times\mathbf{p}$?
4. Can we describe the spin of the electron using $\mathbf{r}\times\mathbf{p}$?
5. Show that $\sigma_i\sigma_j + \sigma_j\sigma_i = 2\delta_{ij}$
6. Verify the relation $\sigma_i\sigma_j = \delta_{ij} + i\varepsilon_{ijk}\sigma_k$. Using this relation, prove that $\boldsymbol{\sigma}\cdot\mathbf{A}\boldsymbol{\sigma}\cdot\mathbf{B} = \mathbf{A}\cdot\mathbf{B} + i\boldsymbol{\sigma}\cdot\mathbf{A}\times\mathbf{B}$
7. In the matrix representation, we have $J_x = \hbar\Sigma_x, J_y = \hbar\Sigma_y, J_z = \hbar\Sigma_x$ for spin 1. Can we write $\Sigma_i\Sigma_j + \Sigma_j\Sigma_i = \boldsymbol{\delta}_{ij}$? The matrices Σ_x, Σ_y and Σ_z are given by

$$\Sigma_x = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{bmatrix}, \qquad \Sigma_y = \begin{bmatrix} 0 & 0 & i \\ 0 & 0 & 0 \\ -i & 0 & 0 \end{bmatrix}, \qquad \Sigma_z = \begin{bmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

8. Evaluate $e^{-i\theta\boldsymbol{\Sigma}\cdot\mathbf{n}}$ where **n** is a unit vector.
9. Why does spin angular momentum commute with all the differential operators?
10. What is meant by Clebsh–Gordan coefficient?
11. Construct Clebsh–Gordan coefficient for $j_1 = \frac{1}{2}$ and $j_2 = \frac{1}{2}$ using the recursion relations for Clebsh–Gordan coefficients (Equations 10.71 and 10.72).
12. Verify the orthoganality relation for Clebsh–Gordan coefficients for $j_1 = \frac{1}{2}$ and $j_2 = \frac{1}{2}$.
13. Using the Table 10.4, determine the Clebsh–Gordan coefficient for $j_1 = 1$ and $j_2 = \frac{1}{2}$.
14. Prove that $(\hat{J}_-)^m\,|jj\rangle = \left[\dfrac{(2j)!m!}{(2j-m)!}\right]^{1/2} |j\,\overline{j-m}\rangle$
15. Prove that $(\hat{J}_-)^{j-m}\,|jj\rangle = \left[\dfrac{(2j)!(j-m)!}{(j+m)!}\right]^{1/2} |jm\rangle$
16. What happens to the wave function of spin $\frac{1}{2}$ system when it is rotated through 360°?
17. What is the significance of Wigner–Eckart theorem?

REFERENCES

1. M. Alonso and H. Valk, 1973. *Quantum Mechanics: Principle and Application*. Massachusetts: Addison-Wesley Publishing Company.
2. I. Schiff, 1968. *Quantum Mechanics*. New Delhi: McGraw-Hill International Book Company.

11

Many-Particle Systems and Quantum Statistics

In this chapter, we discuss two important concepts related to many-particle systems. They are the principles of indistinguishability and the density matrix. The principle of indistinguishability is special to quantum mechanics. It is because of this principle that we have to divide all the elementary particles into two categories: bosons and fermions. There are drastic differences in physical properties between macroscopic objects that are made of bosons and fermions. We first discuss the concept of identical particles, principle of indistinguishability and their consequences. Then we introduce the concept of density matrix.

Density matrix is a general formalism suitable to study an ensemble of identical particles whose states may be the same or different. The density matrix is so general that the concept of state vector can be replaced by density matrix and the quantum mechanics can be formulated in terms of density matrices. However, only an outline of density matrix is presented here.

11.1 IDENTICAL PARTICLES AND INDISTINGUISHABILITY

Each particle has two kinds of properties: intrinsic and extrinsic. Extrinsic properties depend on the environment. Intrinsic properties are unique to the particles, and these properties cannot be removed from a particle. The intrinsic properties which are of immediate interest to us are mass, charge and spin. The concept of identical particles and the principle of indistinguishability also belong to the category of intrinsic properties.

Two particles are said to be identical if they have same charge, mass and spin. We can have identical protons, identical electrons, identical photons and so on. Obviously protons and electrons cannot be termed as identical particles since they differ in many respects including the mass.

In classical physics, though the particles are identical, they can be distinguished from each other. This is the most important aspect in the context of this chapter. How is this possible? One way of distinguishing them is to stick labels of #1 and #2 (or particle *a* and *b* or some names like Tom and Jerry). To say that these objects are distinguishable, their identities should be permanent; i.e., they should have same identities at all instants of time. So the stickers with #1 and #2 should survive long enough to recognize that these two objects are distinct, though they are identical. If we are unable to find such stickers, then there is no point in labelling the particular object as #1 and #2. The question is, can we find such stickers? The answer is ‘Yes’ in classical physics and ‘No’ in quantum mechanics.

To distinguish two particles in classical physics, there is no necessity to put stickers with labels #1 and #2. Newton’s laws help us to label the particles based on their trajectories. Consider the positions and momenta of the two particles at time t and t'. Let us first determine the position and the momenta of the two particles at time t. Our experimental result is: one particle with $\mathbf{r}_1\mathbf{p}_1$ and another particle with $\mathbf{r}_2\mathbf{p}_2$. Let us label the particle with $\mathbf{r}_1\mathbf{p}_1$ as #1 and the particle with $\mathbf{r}_2\mathbf{p}_2$ as #2. So we have labelled the particles as #1 and #2. These are their identities at time t. What about their identities at a later time t'?

At time t', measure the position and the momenta of the particles. The experimental result is: one particle with $\mathbf{r}_1'\ \mathbf{p}_1'$ and another particle with $\mathbf{r}_2'\mathbf{p}_2'$. What about their identities?

The answer depends on the theory. Using Newton's law, let us determine the trajectory of both particles #1 and #2. Suppose we find from the theory that #1 at $\mathbf{r}_1\ \mathbf{p}_1$ at time t would have evolved to $\mathbf{r}_1'\mathbf{p}_1'$ at later time t' and #2 at $\mathbf{r}_2\ \mathbf{p}_2$ at t would have evolved to $\mathbf{r}_2'\ \mathbf{p}_2'$ at t'. Comparing these theoretical predictions with our experimental results, we immediately conclude that the particle with $\mathbf{r}_1'\mathbf{p}_1'$ is #1 and the particle with $\mathbf{r}_1'\mathbf{p}_2'$ is #2. So the particles retain their identities between t and t' and we have distinguished them. This is shown in Fig. 11.1.

Can we do the same analysis to particles like electrons in quantum mechanics?

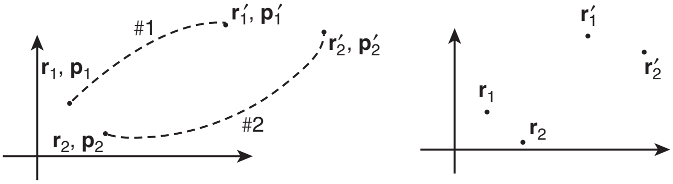


Fig. 11.1 Indistinguishability in classical and quantum physics

Within quantum mechanics, only the wave function ψ evolves; i.e., the wave function evolves from $\psi(\mathbf{r}_1, \mathbf{r}_2, t)$ to $\psi(\mathbf{r}_1', \mathbf{r}_2', t')$. The wave function ψ tells only the probability of finding a particle at position $\mathbf{r}$ at time t. Suppose we find a particle at $\mathbf{r}_1$ and a particle at $\mathbf{r}_2$ at time t. Let us label the particle at $\mathbf{r}_1$ as #1 and the particle at $\mathbf{r}_2$ as #2. So we have given particular identities to each particle at time t. At a later time t', again we find one particle at $\mathbf{r}_1'$ and another particle at $\mathbf{r}_2'$. In the absence of the trajectory, there is no way of telling whether the particle #1 has evolved from $\mathbf{r}_1$ to $\mathbf{r}_1'$ or whether the particle #2 has evolved from $\mathbf{r}_2$ to $\mathbf{r}_1'$. In the same way, we cannot tell which of the two particle has evolved to $\mathbf{r}_2'$. So there is no point in labelling or giving identity to any particular particle as #1 or #2. This implies that there is no way of distinguishing these two particles.

The principle of indistinguishability is brought out in a more forceful way in the following thought experiment (Feynman).

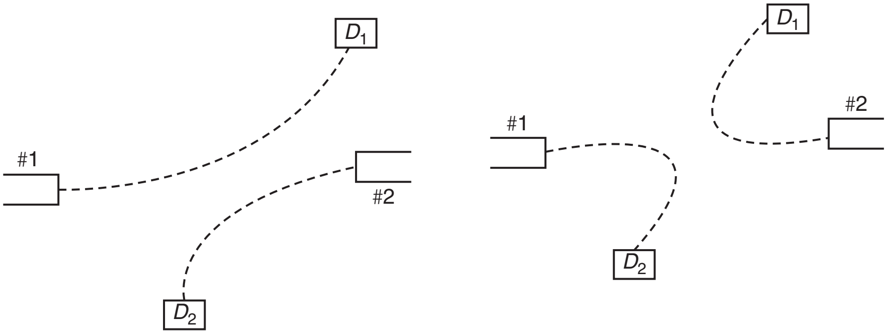


Fig. 11.2 Particles from sources #1 and #2; D_1 and D_2 are detectors of electrons

Electrons are from sources #1 and #2. They scatter each other, and in the end, they are detected by detectors D_1 and D_2. Suppose D_1 detects one electron and D_2 detects another electron. The question is, which electron was detected by D_1? i.e., whether the electron detected by D_1 is from source #1 or source #2. Fig. 11.2 gives both possibilities as seen by classical physics. By working backwards, we can determine the trajectory, and hence we can determine whether it is from source #1 or #2. But within quantum mechanics, there is no way of telling it since there is no concept of trajectory. So the two electrons do not have distinct identities, and here they are indistinguishable.

So far, our discussion highlighted the fact that there is no way of labelling a particular particle as #1, #2 … in a collection of identical particles. This idea has to be translated in a more technical language within quantum mechanics of a multiparticle system. To accomplish this, we have to introduce a new concept called interchange symmetry or permutation symmetry.

11.2 INTERCHANGE SYMMETRY OR PERMUTATION SYMMETRY

Let us consider a collection of two identical particles having the same mass, charge and spin. For the sake of simplicity, let us consider a system of two particles. Their wave function is $\psi(\mathbf{r}_1 s_1;\mathbf{r}_2 s_2)$. This can be simplified by writing ψ as $\psi(1, 2)$.

i.e.,

$$\psi(1,2) = \psi(\mathbf{r}_1 s_1;\mathbf{r}_2 s_2) \tag{11.1}$$

The interchange operator or permutation operator P_{12} is defined as

$$P_{12}\,\psi(1,2) = \psi(2,1) \tag{11.2}$$

Or

$$P_{12}\,\psi(\mathbf{r}_1 s_1;\mathbf{r}_2 s_2) = \psi(\mathbf{r}_2 s_2;\mathbf{r}_1 s_1)$$

i.e., all the variables describing the particle 1 particle 2 are interchanged.

At the outset, $\psi(\mathbf{r}_2 s_2;\mathbf{r}_1 s_1)$ need not be the same as $\psi(\mathbf{r}_1 s_1;\mathbf{r}_2 s_2)$. So let us denote it by ψ'.

$$\psi' = P_{12}\psi \tag{11.3}$$

The interchange operator or the permutation operator has the following properties:

1. P_{12} is a Hermitian operator.
2. P_{12} is a unitary operator.
3. $P_{12}^2 = 1$ and so its eigenvalues are ±1.

Example 11.1 Prove that P_{12} is a Hermitian operator.

Solution: Let us take $\psi(1,2) = \psi(\mathbf{r}_1, \mathbf{r}_2)$.

If P_{12} is a Hermitian operator, then it should satisfy the relation $(\psi, P_{12}\phi) = (P_{12}\psi, \phi)$.

$$\begin{aligned}(\psi, P_{12}\phi) &= \int \psi^*(\mathbf{r}_1;\mathbf{r}_2) P_{12}\phi(\mathbf{r}_1;\mathbf{r}_2)\, d^3\mathbf{r}_1 d^3\mathbf{r}_2 \\ &= \int \psi^*(\mathbf{r}_1;\mathbf{r}_2)\phi(\mathbf{r}_2;\mathbf{r}_1)\, d^3\mathbf{r}_1 d^3\mathbf{r}_2\end{aligned}$$

Since $\mathbf{r}_1$ and $\mathbf{r}_2$ are dummy variables, we can relabel them as $\mathbf{r}_2$ and $\mathbf{r}_1$. i.e., wherever $\mathbf{r}_1$ comes, rewrite it as $\mathbf{r}_2$, and similarly wherever $\mathbf{r}_2$ comes, rewrite it as $\mathbf{r}_1$.

$$\begin{aligned}\therefore (\psi, P_{12}\phi) &= \int \psi^*(\mathbf{r}_2, \mathbf{r}_1)\phi(\mathbf{r}_1, \mathbf{r}_2)d^3\mathbf{r}_2 d^3\mathbf{r}_1 \\ &= \int [P_{12}\psi(\mathbf{r}_1, \mathbf{r}_2)]^* \phi(\mathbf{r}_1, \mathbf{r}_2)d^3\mathbf{r}_1 d^3\mathbf{r}_2 \\ &= (P_{12}\psi, \phi)\end{aligned}$$

$\therefore P_{12}$ is a Hermitian operator.

Example 11.2 Show that $P_{12}^2 = I$.

Solution:

$$\begin{aligned}P_{12}\psi(1, 2) &= \psi(2, 1) \\ P_{12}\psi(2, 1) &= \psi(1, 2) \\ \therefore P_{12}^2\psi(1, 2) &= P_{12}(P_{12}\psi(1, 2)) = P_{12}\psi(2, 1) = \psi(1, 2) \\ \therefore P_{12}^2 &= I\end{aligned}$$

Example 11.3 Determine the eigenvalues of P_{12}.

Solution: The eigenvalue equation is

$$P_{12}\psi(1, 2) = \lambda\psi(1, 2)$$

Operate on both sides by P_{12}. Using the eigenvalue equation, we get

$$\begin{aligned}\underbrace{P_{12}P_{12}}_{I}\psi(1, 2) &= \lambda P_{12}\psi(1, 2) = \lambda^2\psi(1, 2) \\ \therefore \psi(1, 2) &= \lambda^2\psi(1, 2) \\ \text{or } \lambda &= \pm 1 \\ \therefore P_{12}\psi(1, 2) &= \psi(1, 2) \text{ or } P_{12}\psi(1, 2) = -\psi(1, 2)\end{aligned}$$

11.3 INDISTINGUISHABILITY AND THE INTERCHANGE OPERATOR

What is the relationship between indistinguishability and the interchange operator? Indistinguishability is a concept related to identical particles, and it states that it is not possible to distinguish one particle from other particles and it is equivalent to telling that the interchange of two particles cannot be detected in any observation. Since it is impossible to distinguish two identical particles, a physical situation with particle #1 at $\mathbf{r}_1 s_1$ and particle #2 at $\mathbf{r}_2 s_2$ is the same as a physical situation with particle #1 at $\mathbf{r}_2 s_2$ and particle #2 at $\mathbf{r}_1 s_1$.

$$\#1 \text{ at } \mathbf{r}_1 s_1 \text{ and } \#2 \text{ at } \mathbf{r}_2 s_2 \xleftrightarrow[\text{equivalent}]{} \#2 \text{ at } \mathbf{r}_1 s_1 \text{ and } \#1 \text{ at } \mathbf{r}_2 s_2$$

This is essentially the same as the invariance of identical particles under the interchange symmetry operation. So the principle of indistinguishability implies that

$$|\psi(1, 2)|^2 = |P\psi(1, 2)|^2 \tag{11.4}$$

and

$$P_{12}A_{op}(1, 2) = A_{op}(1, 2)P_{12} \tag{11.5}$$

where A is any possible dynamical variable. These requirements make it impossible to prepare any state of the system or make any measurement where it is possible to distinguish two particles.

The Equation (11.5) implies that the operators corresponding to a dynamical variable (any dynamical variable) should be symmetric with respect to interchange of particles. Making use of the Equation (11.5), we have

$$P_{12}A_{op}(1,2)\psi(1,2) = A_{op}(1,2)P_{12}\psi(1,2)$$

$$\text{LHS} = A_{op}(2,1)\psi(2,1) \text{ and } \text{RHS} = A_{op}(1,2)\psi(2,1)$$

$$\therefore A_{op}(2,1)\psi(2,1) = A_{op}(1,2)\psi(2,1)$$

Since this equation is there for any arbitrary wave function $\psi(2, 1)$

$$A_{op}(2, 1) = A_{op}(1, 2) \tag{11.6}$$

In particular, the Hamiltonian operator should be symmetric with respect to interchange of the particle coordinates.

i.e.,

$$H(2, 1) = H(1, 2) \tag{11.7}$$

Since $[P_{12}, H] = 0$, both H and P_{12} should have common eigenfunctions; i.e., the eigenfunction of the Hamiltonian should be either symmetric or antisymmetric with respect to interchange of particles.

$$H\psi_s(1, 2) = E\psi_s(1, 2) \text{ with } \psi_s(1, 2) = \psi_s(2, 1) \tag{11.8}$$

Or

$$H\psi_a(1, 2) = E\psi_a(1, 2) \text{ with } \psi_a(2, 1) = -\psi_a(1, 2) \tag{11.9}$$

These results can be extended to a system with more than two particles. For a system of N particles, the interchange operator P_{ij} is defined as

$$\psi' = P_{ij}\psi(1, 2,\ldots, i, \ldots, j\ldots) = \psi(1, 2, \ldots, j, \ldots, i\ldots) \tag{11.10}$$

The principle of indistinguishability implies

$$|\psi(1, 2, \ldots, i, \ldots, j\ldots)|^2 = |P_{ij}\psi(1, 2, \ldots, j, \ldots, i\ldots)|^2 \tag{11.11}$$

$$P_{ij}A_{op}(1, 2, \ldots, i, \ldots j\ldots) = A_{op}(1, 2, \ldots, i, \ldots, j\ldots)\, P_{ij} \tag{11.12}$$

The extension from two-particle case to three or more number of particles is more subtle than what is presented here (see Ballentine).

11.4 SYMMETRIZATION POSTULATE

For a multiparticle system, the eigenfunction of an interchange operator is either symmetric or antisymmetric with respect to the interchange of the particles *i* and *j*. The symmetrization postulates generalize this idea to any state vector.

Any allowed state vector $\psi(1, 2, \ldots, i, \ldots, j\ldots)$ has to be either symmetric (bosons) or antisymmetric (fermions) with respect to interchange of any pair of particles *i* and *j*.

$$\psi = \psi_s(1, 2, \ldots, i, \ldots, j\ldots) = \psi_s(1, 2, \ldots, j, \ldots, i, \ldots) \tag{11.13}$$

$$\psi = \psi_a(1, 2, \ldots, i, \ldots, j\ldots) = -\psi_a(1, 2, \ldots, j, \ldots, i\ldots) \tag{11.14}$$

The symmetric wave functions describe bosons, and the antisymmetric wave functions describe fermions.

From quantum field theory, it can be shown that the particles with integral spins ($s = 0, 1, 2 \ldots$) are described by symmetric wave functions and particles with half integer spins ($s = 1/2, 3/2\ldots$) are described by antisymmetric wave functions.

11.5 BOSONS – CONSTRUCTION OF SYMMETRIC WAVE FUNCTIONS

Let us first consider a two-particle system and let us start with a multiparticle wave function $\psi(1, 2)$, which is neither symmetric nor antisymmetric.

Now we can construct the symmetric wave function as follows:

$$\psi_s(1, 2) = \frac{1}{\sqrt{2}}[\psi(1, 2) + \psi(2, 1)] \tag{11.15}$$

Next, let us considers three-particle case. Again, let us start with an arbitrary function $\psi(1, 2, 3)$.Then, we can construct the symmetric wave functions as follows.

$$\psi_s(1,2,3) = \frac{1}{\sqrt{6}}[\psi(1,\ 2,\ 3) + \psi(2,\ 1,\ 3) + \psi(1,\ 3,\ 2) + \psi(3,\ 2,\ 1) + \psi(2,\ 3,\ 1) + \psi(3,\ 1,\ 2)] \tag{11.16}$$

This amounts to performing permutation of the particle coordinates. This can be generalized by to any number of particles.

$$\psi_s(1, 2\ldots) = \frac{1}{\sqrt{N!}}\sum_P P\psi(1, 2, \ldots, N) \tag{11.17}$$

where *P* in the summation sign denotes permutations.

Two Non-interacting Bosons

The Hamiltonian for such a system is

$$H(1, 2) = h(1) + h(2) \tag{11.18}$$

Obviously, the Hamiltonian $H(1, 2)$ is symmetric with respect to the coordinates of particles 1 and 2.

Let u_α and u_β be two distinct eigenstates.

$$hu_\alpha = E_\alpha u_\alpha; \quad hu_\beta = E_\beta u_\beta \tag{11.19}$$

If the particles are distinguishable, each particle can be identified or labelled by specifying its eigenstate. There are two such distinguishable states.

$$\psi_I(1, 2) = u_\alpha(1)u_\beta(2); \quad \psi_{II}(1, 2) = u_\beta(1)u_\alpha(2) \tag{11.20}$$

$u_\alpha(1)u_\beta(2)$ describes particle #1 in state u_α and particle #2 in state u_β. This is different from $u_\beta(1)$ $u_\alpha(2)$ which describes particle #1 in state u_β and particle #2 in u_α. They are neither symmetric nor antisymmetric.

$$Hu_\alpha(1)u_\beta(2) = (E_\alpha + E_\beta)u_\alpha(1)u_\beta(2) \tag{11.21}$$

$$Hu_\alpha(2)u_\beta(1) = (E_\alpha + E_\beta)u_\beta(1)u_\alpha(2) \tag{11.22}$$

The symmetrized wave function is

$$\psi_s(1, 2) = \frac{1}{\sqrt{2}}[u_\alpha(1)u_\beta(2) + u_\alpha(2)u_\beta(1)] \tag{11.23}$$

Note
$$H\psi_s(1, 2) = (E_\alpha + E_\beta)\psi_s(1, 2) \tag{11.24}$$

Three Non-interacting Bosons

The Hamiltonian for such a system is

$$H = h(1) + h(2) + h(3) \tag{11.25}$$

First, let us assume that one particle in u_α, one in u_β and one in u_γ.

If the particles are distinguishable, then there are six distinct multiparticle states. They are

$$\begin{aligned} \psi_I &= u_\alpha(1)u_\beta(2)u_\gamma(3); & \psi_{II} &= u_\alpha(2)u_\beta(1)u_\gamma(3) \\ \psi_{III} &= u_\alpha(1)u_\beta(3)u_\gamma(2); & \psi_{IV} &= u_\alpha(3)u_\beta(2)u_\gamma(2)u_\gamma(1) \\ \psi_V &= u_\alpha(2)u_\beta(3)u_\gamma(1); & \psi_{VI} &= u_\alpha(3)u_\beta(1)u_\gamma(2) \end{aligned} \tag{11.26}$$

All these six states are energy eigenstates with energy $E = E_\alpha + E_\beta + E_\gamma$. Principle of indistinguishability makes all these sates to describe the same physical state given by

$$\psi_s(1, 2, 3) = \frac{1}{\sqrt{6}}[\psi_I + \psi_{II} + \psi_{III} + \psi_{IV} + \psi_V + \psi_{VI}] \tag{11.27}$$

Note that
$$H\psi_s(1, 2, 3) = (E_\alpha + E_\beta + E_\gamma)\psi_s(1, 2, 3)$$

11.6 FERMION: CONSTRUCTION OF ANTISYMMETRIC STATE

Let us first consider a two-particle system. Let us start with a two-particle wave function $\psi(1,2)$, which is neither symmetric nor antisymmetric. We construct an antisymmetric wave function as follows:

$$\psi_a(1,2) = \frac{1}{\sqrt{2}}[\psi(1,2) - \psi(2,1)]$$

It is obvious that $\psi_a(2,1) = -\psi_a(1,2)$

Next, let us consider a system of three fermions. We start with a function $\psi(1,2,3)$, which is neither symmetric nor antisymmetric. The antisymmetric function can be constructed as follows:

$$\begin{aligned}\psi_a(1,2,3) = \frac{1}{\sqrt{6}}[&\psi(1,2,3) - \psi(2,1,3) - \psi(1,3,2)\\ &-\psi(3,2,1) + \psi(2,3,1) + \psi(3,1,2)]\end{aligned} \tag{11.28}$$

It is easy to check that $\psi_a(1,2,3)$ in (11.28) is antisymmetric with respect to interchange of the coordinate of any two particles. For instance, $\psi_a(2,1,3)$ is given by

$$\psi_a(2,1,3) = \frac{1}{\sqrt{6}}[\psi(2,1,3) - \psi(1,2,3) - \psi(2,3,1)$$
$$-\psi(3,1,2) + \psi(1,3,2) + \psi(3,2,1)] = -\psi_a(1,2,3)$$

The signs of the second, third, and fourth terms of (11.28) are different from the signs of the last two terms. These three terms are obtained from $\psi(1,2,3)$ by single permutation.

$$\psi(1,2,3) \xrightarrow{P_{12}} \psi(2,1,3)$$
$$\psi(1,2,3) \xrightarrow{P_{23}} \psi(1,3,2)$$
$$\psi(1,2,3) \xrightarrow{P_{13}} \psi(3,2,1)$$

Each of the last two expressions in (11.28) requires two permutations to get them from $\psi(1,2,3)$.

$$\psi(1,2,3) \xrightarrow{P_{12}} \psi(2,1,3) \xrightarrow{P_{13}} \psi(2,3,1)$$
$$\psi(1,2,3) \xrightarrow{P_{23}} \psi(1,3,2) \xrightarrow{P_{13}} \psi(3,1,2)$$

Terms obtained by single permutation have negative sign and terms with two permutations have positive sign. Therefore, each permutation accompanies a factor (−1). So, $\psi_a(1,2,3)$ can be written as

$$\psi_a(1, 2, 3) = \frac{1}{\sqrt{6}} \sum_P (-1)^P P\psi(1, 2, 3) \tag{11.29}$$

This can be generalized to *N* particle system

$$\psi_a(1, 2, 3, \ldots, N) = \frac{1}{\sqrt{N!}} \sum_P (-1)^P P\psi(1, 2, 3, \ldots, N) \tag{11.30}$$

Non-interacting Fermions

For a system of two fermions, let us assume that one particle is in state u_α and another particle is in state u_β. If the two particles are treated as distinguishable, then we have two distinct states.

$$\psi_I = u_\alpha(1)u_\beta(2), \qquad \psi_{II} = u_\alpha(2)u_\beta(1) \tag{11.31}$$

The principle of indistinguishability reduces these two states to one state given by

$$\psi_a(1, 2) = \frac{1}{\sqrt{2}}[u_\alpha(1)u_\beta(2) - u_\alpha(2)u_\beta(1)] \tag{11.32}$$

For a system of three fermions, let us assume that one particle in each of state u_α, u_β, and u_γ. If the three fermions are treated as distinguishable particles, then there are six distinct states given in (11.26).

The principle of indistinguishablity reduces these six distinct states to a single state given by

$$\begin{aligned}\psi_a(1, 2, 3) = \frac{1}{\sqrt{6}}[&u_\alpha(1)u_\beta(2)u_\gamma(3) - u_\alpha(1)u_\beta(3)u_\gamma(2)\\ &- u_\alpha(3)u_\beta(2)u_\gamma(1) - u_\alpha(2)u_\beta(1)u_\gamma(3)\\ &+ u_\alpha(2)u_\beta(3)u_\gamma(1) + u_\alpha(3)u_\beta(1)u_\gamma(2)\end{aligned} \tag{11.33}$$

Slater's Determinant

It is easy to recognize that the wave functions for fermions can be written in the form of determinants.

$$\psi_a(1, 2) = \frac{1}{\sqrt{2}}\begin{vmatrix} u_\alpha(1) & u_\alpha(2) \\ u_\beta(1) & u_\beta(2) \end{vmatrix} \tag{11.34}$$

and

$$\psi_a(1, 2, 3) = \frac{1}{\sqrt{6}}\begin{vmatrix} u_\alpha(1) & u_\alpha(2) & u_\alpha(3) \\ u_\beta(1) & u_\beta(2) & u_\beta(3) \\ u_\gamma(1) & u_\gamma(2) & u_\gamma(3) \end{vmatrix} \tag{11.35}$$

So, for a system of non-interacting fermions, the antisymmetric wave functions can be constructed using determinants. Such determinants are known as Slater's determinant. For a system of N non-interacting fermions, Slater's determinant is given by

$$\psi_{\alpha}(1,2,\ldots,N)=\frac{1}{\sqrt{N!}}\begin{vmatrix} u_{\alpha}(1) & u_{\alpha}(2) & \ldots & u_{\alpha}(N) \\ u_{\beta}(1) & u_{\beta}(2) & \ldots & u_{\beta}(N) \\ \vdots & & & \\ u_{\lambda}(1) & u_{\lambda}(2) & \ldots & u_{\lambda}(N) \end{vmatrix} \tag{11.36}$$

11.7 PAULI'S EXCLUSION PRINCIPLE

Pauli's exclusion principle says that at the most only one electron (fermions) can be found in a state; i.e., it is impossible for two electrons to be in the same state. This is a consequence of the fact that the wave functions for fermions should be antisymmetric.

Let us consider a system of two electrons. Let us start with two different states, assigning one state for each particle.

$$\psi_{a}(1,2)=\frac{1}{\sqrt{2}}[u_{\alpha}(1)u_{\beta}(2)-u_{\alpha}(2)u_{\beta}(1)] \tag{11.37}$$

Then, demanding two electrons in the same state makes u_{α} and u_{β} to the same.

$$\therefore \psi_{a}(1,2)=\frac{1}{\sqrt{2}}[u_{\alpha}(1)u_{\alpha}(2)-u_{\alpha}(2)u_{\alpha}(1)]=0 \tag{11.38}$$

This confirms Pauli's exclusion principle. The same argument can be extended in N particle system. Start with N different single particle states u_{α}, u_{β}, …, u_{λ}, assigning one state for each of the particles. For fermions, the antisymmetric state is the Slater's determinant. Demanding two electrons in the same state amounts to making two states, say, u_{α}, and u_{β}, to be the same; i.e., $u_{\beta}=u_{\alpha}$. Then the Slater's determinant becomes

$$\psi_{a}(1\ldots)=\frac{1}{\sqrt{N!}}\begin{vmatrix} u_{\alpha}(1) & u_{\alpha}(2) & \ldots & \ldots \\ u_{\alpha}(1) & u_{\alpha}(2) & \ldots & \ldots \\ u_{\gamma}(1) & u_{\gamma}(2) & \ldots & \ldots \\ u_{\lambda}(1) & u_{\lambda}(2) & \ldots & \ldots \end{vmatrix}=0 \tag{11.39}$$

11.8 DISTINGUISHABILITY AS AN APPROXIMATION

Under some circumstances, even within quantum mechanics, a multiparticle system can be approximated as a collection of distinguishable particles. This happens when the particles are so far away from each other that the wave functions $u_{\alpha}(\mathbf{r}_1), u_{\beta}(\mathbf{r}_2)\ldots$ are well separated from each other, and so there is no overlapping between them. This can be illustrated by considering a special case in which there is a system of two non-interacting particles that are separated as shown in Figs. 11.3 and 11.4.

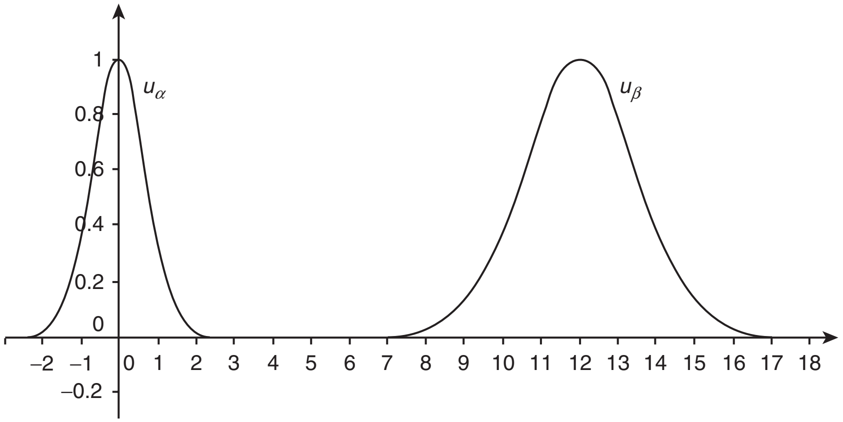


Fig. 11.3 Wave functions of two far away particles

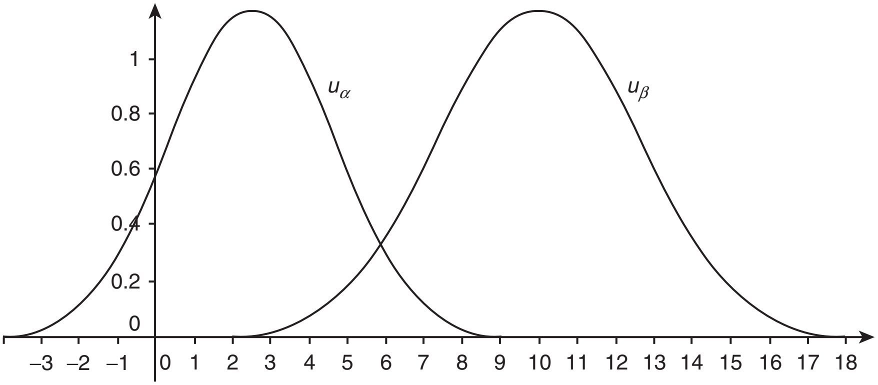


Fig. 11.4 Wave functions of two closer particles

$u_\alpha(x)$ is centered around x_{10}, and $u_\beta(x)$ is centered around x_{20}. For instance, for the wave functions shown in the Fig. 11.3, $x_{10} = 0$ and $x_{20} = 12$. The antisymmetric wave function is

$$\psi(1, 2) = \frac{1}{\sqrt{2}}[u_\alpha(x_1)u_\beta(x_2) - u_\alpha(x_2)u_\beta(x_1)]$$

For any set of values (x_1, x_2), only one term will be significantly different from zero, and the other term will be zero. A simple mumerical example will help the reader. For instance, consider two pairs of values: $(x_1 = 1, x_2 = 10)$ and $(x_1 = 10, x_2 = 1)$.

$$\psi(x_1 = 1, x_2 = 10) = \frac{1}{\sqrt{2}}\left[\underbrace{u_\alpha(1)u_\beta(10)}_{\text{Non-zero}} - \underbrace{u_\alpha(10)u_\beta(1)}_{\text{zero}}\right]$$

$$\psi(x_1 = 10, x_2 = 1) = \frac{1}{\sqrt{2}}\left[\underbrace{u_\alpha(10)u_\beta(1)}_{\text{zero}} - \underbrace{u_\alpha(1)u_\beta(10)}_{\text{Non-zero}}\right]$$

For any pair of values (x_1, x_2)

$$\psi(x_1, x_2) \text{ is either } u_\alpha(x_1)u_\beta(x_2) \text{ or } u_\alpha(x_2)u_\beta(x_1)$$

as shown above.

So, for any pair of values (x_1, x_2), it is only a product of two single functions very much like a distinguishable systems. When they come near, there is an overlap of the two single particle state function as shown in Fig. 11.4, and so both terms in the antisymmetric wave functions are significant. For instance, consider the coordinates $x_1 = 4,\ x_2 = 5$.

$$\psi(x_1 = 4, x_2 = 5) = \frac{1}{\sqrt{2}}\left[\underbrace{u_\alpha(4)u_\beta(5)}_{\text{Non-zero}} - \underbrace{u_\alpha(5)u_\beta(4)}_{\text{Non-zero}}\right]$$

So when the particles in a multiparticle system are so far away that there is no overlap between their single particle wave functions, they can be treated as distinguishable system. This is good approximation for such systems.

Example 11.4 The wave function for a system of two parties is given by

$$\psi(1, 2) = \frac{1}{\sqrt{2}}[u_\alpha(x_1)u_\beta(x_2) + u_\alpha(x_2)u_\beta(x_1)]$$

Here, $$u_\alpha(x) = \left(\frac{1}{\sqrt{\pi}}\right)^{1/2} \exp\left(-\frac{(x-a)^2}{2}\right)$$

and $$u_\beta(x) = \left(\frac{1}{8\sqrt{\pi}}\right)^{1/2} \exp\left(-\frac{(x-b)^2}{2}\right)[4(x-b)^2 - 2]$$

(These are harmonic oscillator wave functions centered around $x = a$ and $x = b$)

Can we approximate them as distinguishable particles for

1. $a = 0 \quad b = 12$ 2. $a = 4 \quad b = 8$

Solution:

$$\psi(1, 2) = \left(\frac{1}{8\pi}\right)^{1/2}\left\{e^{-\frac{x_1^2}{2}} \cdot e^{-\frac{(x_2-16)^{2/2}}{2}}(4(x_2-16)^2 - 2) + e^{-\frac{x_2^2}{2}} e^{-(x_1-b)^2}[4(x_1-16)^2 - 2]\right\}$$

For $a = 0$, $b = 12$, $\psi(1, 2)$ is given by

$$\psi(1, 2) = \left(\frac{1}{16\pi}\right)^{1/2}\left\{e^{-\frac{x_1{}^2}{2}} \cdot e^{-(x_2-16)^2/2}[4(x_2-16)^2 - 2]\right.$$
$$\left. + e^{-\frac{x_2{}^2}{2}} e^{-(x_2-16)^2/2}[4(x_1-16)^2 - 2]\right\}$$

The wave functions $u_\alpha(x)$ and $u_\beta(x)$ are sketched in Fig. 11.5 for $a = 0$ and $b = 12$, and obviously, do not overlap. Therefore, the two particles can be approximated as distinguishable particles.

For $a = 4$ $b = 8$, $\psi(1, 2)$ is given by

$$\psi(1,2) = \frac{1}{\sqrt{8\pi}}\left\{\exp\left(-\frac{(x_1-4)^2}{2}\right)\exp\left(-\frac{(x_2-8)^2}{2}\right)[4(x_2-8)^2-2]\right.$$
$$\left.+\exp\left(-\frac{(x_2-4)^2}{2}\right)\exp\left(-\frac{(x_1-8)^2}{2}\right)[4(x_1-8)^2-2]\right\}$$

For these values of the parameters a and b, the wave functions $u_\alpha(x)$ and $u_\beta(x)$ overlap as shown in Fig. 11.6. So they cannot be approximated as distinguishable particles.

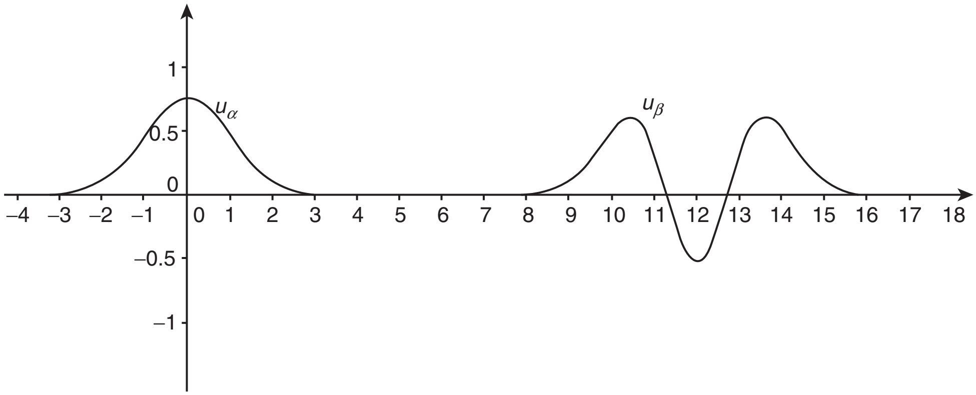


Fig. 11.5 Wave functions of two particles without overlapping

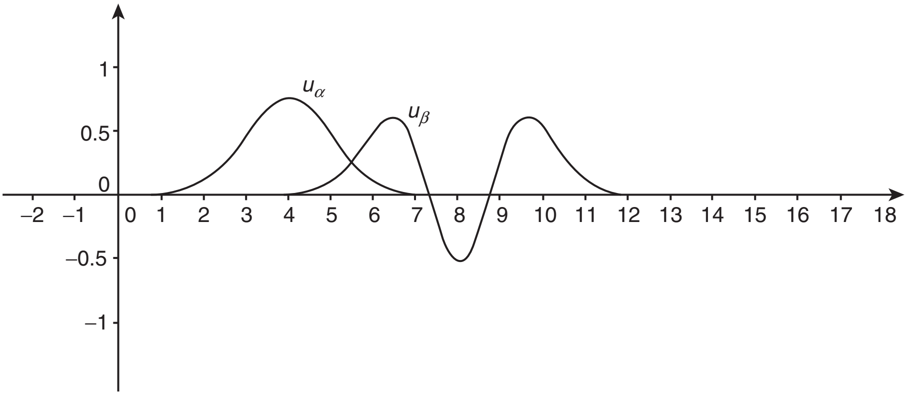


Fig. 11.6 Wave functions of two particles with overlapping

11.9 TWO-ELECTRON SYSTEM HELIUM

Let us consider a two-electron system as an illustration of non-trivial consequences of principle of indistinguishability. A simple example is helium atom.

The Hamiltonian for helium atom is

$$H = \frac{p_1^2}{2m} + \frac{p_2^2}{2m} - \frac{e^2}{r_1} - \frac{e^2}{r_2} + \frac{e^2}{|\mathbf{r}_1 - \mathbf{r}_2|} \tag{11.40}$$

$$= h(1) + h(2) + V_{12} \tag{11.41}$$

The interaction involving spins of the electrons has been neglected. Let us neglect the other interactions further. Then it becomes

$$H = h(1) + h(2) \tag{11.42}$$

Since the Hamiltonian H is independent of spin of the particles, the wave function $\psi(\mathbf{r}_1 s_1; \mathbf{r}_2 s_2)$ can be written as the product of space coordinate function and spin coordinate function.

$$\psi(\mathbf{r}_1 s_1; \mathbf{r}_2 s_2) = \Phi(\mathbf{r}_1, \mathbf{r}_2)\chi(s_1, s_2) \tag{11.43}$$

Since the electrons are fermions, its total wave function has to be antisymmetric with respect to the interchange $\mathbf{r}_1 s_1$ and $\mathbf{r}_2 s_2$.

$$\therefore \psi(\mathbf{r}_1 s_1; \mathbf{r}_2 s_2) = \psi_a(\mathbf{r}_1 s_1; \mathbf{r}_2 s_2) \tag{11.44}$$

i.e.,

$$\psi_a(\mathbf{r}_2 s_2; \mathbf{r}_1 s_1) = -\psi_a(\mathbf{r}_1 s_1; \mathbf{r}_2 s_2) \tag{11.45}$$

Since the wave function $\psi(\mathbf{r}_1 s_1; \mathbf{r}_2 s_2)$ can be written in the form given in the Equation (11.45), antisymmetry requirement can be achieved by two methods.

$$\psi_a(\mathbf{r}_1 s_1; \mathbf{r}_2 s_2) = \Phi_s(\mathbf{r}_1, \mathbf{r}_2)\chi_a(s_1, s_2) \tag{11.46}$$

$$\psi_a(\mathbf{r}_1 s_1; \mathbf{r}_2 s_2) = \Phi_a(\mathbf{r}_1, \mathbf{r}_2)\chi_s(s_1, s_2) \tag{11.47}$$

In the Equation (11.46), Φ_s is symmetric with respect to interchange of space coordinates $\mathbf{r}_1$ and $\mathbf{r}_2$, and χ_a is antisymmetric with respect to exchange of spin coordinates. In the Equation (11.47), space coordinate function Φ_a is antisymmetric, and spin coordinate function χ_s is symmetric.

In general, states with symmetric space coordinate function are known as para states and states with antisymmetric space coordinate function are called ortho states.

11.9.1 Spin Wave Function

If $\mathbf{s}_1$ and $\mathbf{s}_2$ are the spin angular momenta of the two electrons, the total spin angular momentum $\mathbf{s}$ is given by

$$\mathbf{s} = \mathbf{s}_1 + \mathbf{s}_2 \tag{11.48}$$

Instead of writing the spin states as $\left|\frac{1}{2}\frac{1}{2}\right\rangle$ and $\left|\frac{1}{2}-\frac{1}{2}\right\rangle$, we write them as α and β.

$$\therefore S_1^2\alpha(1) = \frac{3}{4}\hbar^2\alpha(1); \quad S_1^2\beta(1) = \frac{3}{4}\hbar^2\beta(1)$$

$$S_{1z}\alpha(1) = \frac{\hbar}{2}\alpha(1); \qquad S_{1z}\beta(1) = -\frac{\hbar}{2}\beta(1)$$

$$S_2^2\,\alpha(2) = \frac{3}{4}\hbar^2\alpha(2); \quad S_2^2\,\beta(2) = \frac{3}{4}\hbar^2\beta(2)$$

$$S_{1z}\alpha(2) = \frac{\hbar}{2}\alpha(2); \qquad S_{1z}\beta(2) = -\frac{\hbar}{2}\beta(2)$$

When S_1 and S_2 are combined, the quantum number s of the resultant angular momentum is $s = 1$ or $s = 0$. Let $\chi_{11}, \chi_{10}, \chi_{1-1}$ be the states with $s = 1$ and χ_{00} be the state with $s = 0$.

$$\chi_{11} = \alpha(1)\alpha(2) \tag{11.49}$$

$$\chi_{10} = \frac{1}{\sqrt{2}}[\alpha(1)\beta(2) + \alpha(2)\beta(1)] \tag{11.50}$$

$$\chi_{1\text{-}1} = \beta(1)\beta(2) \tag{11.51}$$

$$\chi_{00} = \frac{1}{\sqrt{2}}[\alpha(1)\beta(2) - \alpha(2)\beta(1)] \tag{11.52}$$

The states $\chi_{11}, \chi_{10}, \chi_{1-1}$, are known as spin triplet and χ_{00} is known as spin singlet.

Note that the pair $\alpha(1)$, $\beta(1)$ and the pair $\alpha(2)$, $\beta(2)$ belong to different Hilbert spaces. They can be written as follows:

$$\alpha(1) = \begin{pmatrix}1\\0\end{pmatrix}_1 ; \beta(1) = \begin{pmatrix}0\\1\end{pmatrix}_1 ; \alpha(2) = \begin{pmatrix}1\\0\end{pmatrix}_2 ; \beta(2) = \begin{pmatrix}0\\1\end{pmatrix}_2$$

The products in the Equations (11.49)–(11.52) are direct products of these spinors. For instance,

$$\alpha(1)\beta(2) = \begin{pmatrix}1\\0\end{pmatrix}_1 \otimes \begin{pmatrix}1\\0\end{pmatrix}_2 = \begin{pmatrix}1\begin{pmatrix}1\\0\end{pmatrix}\\ 0\begin{pmatrix}1\\0\end{pmatrix}\end{pmatrix} = \begin{pmatrix}1\\0\\0\\0\end{pmatrix}$$

The wave functions are now given by

$$\psi_a = \Phi_a(\mathbf{r}_1, \mathbf{r}_2)\begin{cases}\chi_{11}\\ \chi_{10} \quad \text{triplet; para state}\\ \chi_{1-1}\end{cases} \tag{11.53}$$

$$\psi_a = \Phi_s(\mathbf{r}_1, \mathbf{r}_2)\chi_{00} \quad \text{singlet; ortho state} \tag{11.54}$$

For helium, if the Hamiltonian is given by $H = h(1) + h(2)$, then both ortho state and para state have the same energy. The space symmetric Φ_s and antisymmetric Φ_a are given by in terms of hydrogen atom wave functions.

$$\Phi_{s,a}(\mathbf{r}_1,\mathbf{r}_2) = \frac{1}{\sqrt{2}}[u_{n_1 l_1 m_1}(\mathbf{r}_1)u_{n_2 l_2 m_2}(\mathbf{r}_2) \pm u_{n_2 l_2 m_2}(\mathbf{r}_2)u_{n_1 l_1 m_1}(\mathbf{r}_1)] \tag{11.55}$$

$$H\Phi_s(\mathbf{r}_1, \mathbf{r}_2)\chi_a(1, 2) = [E_{n1} + E_{n2}]\Phi_s(\mathbf{r}_1, \mathbf{r}_2)\chi_a(1, 2) \tag{11.56}$$

$$H\Phi_a(\mathbf{r}_1, \mathbf{r}_2)\chi_s(1, 2) = [E_{n1} + E_{n2}]\Phi_a(\mathbf{r}_1, \mathbf{r}_2)\chi_s(1, 2) \tag{11.57}$$

However, if we take into account of the interaction between the electrons by $H' = e^2/r_{12}$, then within first-order perturbation theory (see Chapter13),

$$E_{\text{ortho}} = E_{n1} + E_{n2} + K + J \tag{11.58}$$

and
$$E_{\text{para}} = E_{n1} + E_{n2} + K - \mathrm{J} \tag{11.59}$$

where
$$K = \int \mathrm{d}^3\mathbf{r}_1\mathrm{d}^3\mathbf{r}_2 \left|u_{n_1l_1m_1}(\mathbf{r}_1)\right|^2 \frac{\mathrm{e}^2}{\mathrm{r}_{12}} \left|u_{n_2l_2m_2}(\mathbf{r}_2)\right|^2 \tag{11.60}$$

and
$$J = \int \mathrm{d}^3\mathbf{r}_1\mathrm{d}^3\mathbf{r}_2 (u_{n_1l_1m_1}(\mathbf{r}_1))^*(u_{n_2l_2m_2}(\mathbf{r}_2))^* \frac{\mathrm{e}^2}{\mathrm{r}_{12}} u_{n_1l_1m_1}(\mathbf{r}_2)u_{n_2l_2m_2}(\mathbf{r}_1) \tag{11.61}$$

The first term represents the well-known Coulomb interaction between the charge distribution $\rho(\mathbf{r}_1)$ and $\rho(\mathbf{r}_2)$.

$$K = \int \mathrm{d}^3\mathbf{r}_1\mathrm{d}^3\mathbf{r}_2\rho(\mathbf{r}_1)\frac{\mathrm{e}^2}{\mathrm{r}_{12}}\rho(\mathbf{r}_2) \tag{11.62}$$

There is no classical analogue to the second term. It is known as exchange interaction. It is a consequence of antisymmetrization of electron wave function.

11.10 INDISTINGUISHABILITY AND COUNTING MICROSTATE IN STATITICAL MECHANICS

One of the methods used to study a system of particles in statistical mechanics is to use the concept of microcanonical formalism, in which the system is isolated from the environment. The total energy E, the total number of particles N and the volume V of a system are fixed in this formalism. These are the thermo dynamical variables which describe the thermodynamic state of a system. In statistical mechanics, we have two kinds of descriptions of the state of a thermodynamic system. They are known as macrostate and microstate. The set of the thermo dynamical variables (E, N, V) specifies the macrostate. On the other hand, a microstate corresponds to the specification of the multiparticle states of all the N particles. So the microstate can be written as $\psi(1, 2, 3, \ldots, N)$. For a given macrostate, we can have a large number of microstates. Let Ω be the number of microstates corresponding to a macrostate (E, N, V). The entropy S for a given macrostate is given by

$$S = k_B \ln \Omega$$

Our interest here is to understand the implications or the consequences of the principle of indistinguishability in counting the number possible microstates Ω. For a given macrostate (E, V, N), the values of Ω for the distinguishable particles and the indistinguishable particles are different. Within the class of indistinguishable particles, the values of Ω for bosons and fermions are different. So, for a given set (E, N, V),

Ω for distinguishable particles ≠ Ω for bosons ≠ Ω for fermions

These differences have deep implications, and they lead to three kinds of statistical distributions in statistical mechanics.

distinguishable particles:		Maxwell–Boltzman distribution
Indistinguishable:	Bosons	Bose–Einstein distribution
	Fermions	Fermi–Dirac distribution

These distributions will not be derived here. Our interest here is to demonstrate the implications of distinguishability and indistinguishability in counting the number of multiparticle states for a given number of particles. Let n_s be the number of particles and g_s be the degeneracy of the energy level. We are interested in determining the number of multiparticle states $\Omega(n_s, g_s)$.

11.10.1 Case 1: Distinguishable Particles

The simplest case is to consider $n_s = 2$ and $g_s = 2$. The particles can be labelled as a and b. The states can be taken as u_α and u_β. There are four multiparticle states, and they are given by

$$\psi_1(a,b) = u_\alpha(a)u_\alpha(b)$$

$$\psi_2(a,b) = u_\alpha(a)u_\beta(b)$$

$$\psi_3(a,b) = u_\alpha(b)u_\beta(a)$$

$$\psi_4(a,b) = u_\beta(a)u_\beta(b)$$

In statistical mechanics, there is a simple method of writing these states, using a box for single particle states. They are written as follows:

$$\psi_1(a,b) = u_\alpha(a)u_\alpha(b) \rightarrow \begin{array}{|c|c|}\hline ab & \\ \hline\end{array}$$

$$\psi_2(a,b) = u_\alpha(a)u_\beta(b) \rightarrow \begin{array}{|c|c|}\hline a & b \\ \hline\end{array}$$

$$\psi_3(a,b) = u_\alpha(b)u_\beta(a) \rightarrow \begin{array}{|c|c|}\hline b & a \\ \hline\end{array}$$

$$\psi_4(a,b) = u_\beta(a)u_\beta(b) \rightarrow \begin{array}{|c|c|}\hline & ab \\ \hline\end{array}$$

This method of writing the multiparticle states simplifies the calculations. Let us ask the reverse question. How do we generate these multiparticle states? Let us start with the empty boxes $\begin{array}{|c|c|}\hline & \\ \hline\end{array}$. The box for the particle a can be found in two ways. They are $\begin{array}{|c|c|}\hline a & \\ \hline\end{array}$ and $\begin{array}{|c|c|}\hline & a \\ \hline\end{array}$. In each of these arrangements, the particle b can be put in two ways. $\begin{array}{|c|c|}\hline a & \ \\ \hline\end{array}$ becomes $\begin{array}{|c|c|}\hline ab & \ \\ \hline\end{array}$ and $\begin{array}{|c|c|}\hline a & b \\ \hline\end{array}$. The arrangement $\begin{array}{|c|c|}\hline \ & a \\ \hline\end{array}$ becomes $\begin{array}{|c|c|}\hline b & a \\ \hline\end{array}$ and $\begin{array}{|c|c|}\hline \ & ab \\ \hline\end{array}$. So we have two ways of finding the box for each particle. The number of multiparticle states is given by

$$\begin{aligned}\Omega(n_s = 2, g_s = 2) = {} & \text{(2 ways of finding boxes (or states) for particle a)} \times \\ & \text{(2 ways of finding boxes for b)} = 2^2 = 4.\end{aligned}$$

This reasoning can be extended to general cases to get

$$\Omega(n_s, g_s) = (g_s)^{n_s}$$

Example 11.5 Write down the number of multiparticle states for three distinguishable particles a, b and c, assuming two states u_α and u_β are available to them.

Solution: Since two states are available, we should take two boxes $\begin{array}{|c|c|}\hline \ & \ \\ \hline\end{array}$. Let us put particle a in these boxes to get $\begin{array}{|c|c|}\hline a & \ \\ \hline\end{array}$ and $\begin{array}{|c|c|}\hline \ & a \\ \hline\end{array}$. In each of these arrangements, let us put particle b in two ways to get $\left\{\begin{array}{|c|c|}\hline ab & \ \\ \hline\end{array}, \begin{array}{|c|c|}\hline a & b \\ \hline\end{array}\right\}$ and $\left\{\begin{array}{|c|c|}\hline b & a \\ \hline\end{array}, \begin{array}{|c|c|}\hline \ & ab \\ \hline\end{array}\right\}$. Now let us put particle c in each of these four arrangements to get $\left\{\begin{array}{|c|c|}\hline abc & \ \\ \hline\end{array}, \begin{array}{|c|c|}\hline ab & c \\ \hline\end{array}\right\}, \left\{\begin{array}{|c|c|}\hline ac & b \\ \hline\end{array}, \begin{array}{|c|c|}\hline a & bc \\ \hline\end{array}\right\}, \left\{\begin{array}{|c|c|}\hline bc & a \\ \hline\end{array}, \begin{array}{|c|c|}\hline b & ac \\ \hline\end{array}\right\}$ and $\left\{\begin{array}{|c|c|}\hline c & ab \\ \hline\end{array}, \begin{array}{|c|c|}\hline \ & abc \\ \hline\end{array}\right\}$.

The corresponding multiparticle states are given below:

$$\begin{array}{|c|c|}\hline abc & \ \\ \hline\end{array} \rightarrow \psi_1(a,b,c) = u_\alpha(a)u_\alpha(b)u_\alpha(c)$$

$$\begin{array}{|c|c|}\hline ab & c \\ \hline\end{array} \rightarrow \psi_2(a,b,c) = u_\alpha(a)u_\alpha(b)u_\beta(c)$$

$$\begin{array}{|c|c|}\hline ac & b \\ \hline\end{array} \rightarrow \psi_3(a,b,c) = u_\alpha(a)u_\beta(b)u_\alpha(c)$$

$$\begin{array}{|c|c|}\hline a & bc \\ \hline\end{array} \rightarrow \psi_4(a,b,c) = u_\alpha(a)u_\beta(b)u_\beta(c)$$

$$\begin{array}{|c|c|}\hline bc & a \\ \hline\end{array} \rightarrow \psi_5(a,b,c) = u_\beta(a)u_\alpha(b)u_\alpha(c)$$

$$\begin{array}{|c|c|}\hline b & ac \\ \hline\end{array} \rightarrow \psi_6(a,b,c) = u_\beta(a)u_\alpha(b)u_\beta(c)$$

$$\begin{array}{|c|c|}\hline c & ab \\ \hline\end{array} \rightarrow \psi_7(a,b,c) = u_\beta(a)u_\beta(b)u_\alpha(c)$$

$$\begin{array}{|c|c|}\hline \ & abc \\ \hline\end{array} \rightarrow \psi_8(a,b,c) = u_\beta(a)u_\beta(b)u_\beta(c)$$

Thus, we have obtained all the eight possible multiparticle states.

11.10.2 Case 2: Indistinguishable Particles – Bosons

The principle of indistinguishability reduces the number of microstates enormously.

Let us consider the above case, namely $n_s = 3, g_s = 2$. Since there is no way of distinguishing the particles, there is no point in labelling the particles as *a* or *b* or *c*. When they were considered as distinguishable particles as in Example (11.5), there were eight distinct states. Now when they become indistinguishable, things begin to change. Consider, for instance, ψ_2, ψ_3 and ψ_5 in Example 11.5. All these wave functions imply that two particles are in state u_α and one particle in state u_β. For indistinguishable particles, these terms merge into one state, which can be written as

$$\Phi_{2,1} = \frac{1}{\sqrt{3}}[u_\alpha(a)u_\alpha(b)u_\beta(c) + u_\alpha(a)u_\beta(b)u_\alpha(c) + u_\beta(a)u_\alpha(b)u_\alpha(c)]$$

Note the change in notation. $\Phi_{2,1}$ means a multiparticle state in which two particles in single particle state u_α and one particle in single particle state u_β. In terms of the box notation, $\Phi_{2,1}$ can be written as $\begin{array}{|c|c|}\hline \bullet\bullet & \bullet \\ \hline \end{array}$

$$\Phi_{2,1} \to \begin{array}{|c|c|}\hline \bullet\bullet & \bullet \\ \hline \end{array}$$

$$\begin{array}{|c|c|}\hline \bullet\bullet & \bullet \\ \hline \end{array} = \frac{1}{\sqrt{3}}[u_\alpha(a)u_\alpha(b)u_\beta(c) + u_\alpha(a)u_\alpha(c)u_\beta(b) + u_\alpha(b)u_\alpha(c)u_\beta(a)]$$

So $\begin{array}{|c|c|}\hline \bullet\bullet & \bullet \\ \hline \end{array}$ represents a microstate where two particles in state u_α and one particle in state u_β. Proceeding in the same way, we write all the four possible microstates as given below:

3 particles in state u_α: $\begin{array}{|c|c|}\hline \bullet\bullet\bullet & \quad \\ \hline \end{array}$

2 in state u_α, 1 in state u_β: $\begin{array}{|c|c|}\hline \bullet\bullet & \bullet \\ \hline \end{array}$

1 in state u_α, 2 in state u_β: $\begin{array}{|c|c|}\hline \bullet & \bullet\bullet \\ \hline \end{array}$

3 particles in u_β: $\begin{array}{|c|c|}\hline \quad & \bullet\bullet\bullet \\ \hline \end{array}$

In terms of boxes and dots, the above four microstates can be interpreted as follows. Look at all the four multiparticle states. There are two boxes separated by a partition and three dots. Consider the multiparticle state $\begin{array}{|c|c|}\hline \bullet\bullet & \bullet \\ \hline \end{array}$. Interchange the order of the last dot and the partition to get a new multiparticle state $\begin{array}{|c|c|}\hline \bullet\bullet\bullet & \quad \\ \hline \end{array}$. i.e., by permuting a dot and a partition, we have obtained a new multiparticle state. The four multiparticle states can be seen as a result of permutation of the dots and partitions. This idea can be generalized immediately. Take the dots and the partitions together as a single set of objects. (In our example, there are three dots and one partition, and so we have four objects). The number of dots plus the number of partition makes $(g_s + n_s - 1)$ objects, where n_s objects are dots and $(g_s - 1)$ objects are partitions. The permutation of $(g_s + n_s - 1)$ objects will result in $(g_s + n_s - 1)!$ multiparticle states. Have we generated all the possible multiparticle states by permuting the $(g_s + n_s - 1)$ objects? We have to be a bit more careful. Consider a multiparticle state $\begin{array}{|c|c|c|}\hline \bullet\bullet & \bullet & \bullet \\ \hline \end{array}$. By permuting third and fourth dots, we will not get a new state. Similarly, by permuting the two partitions, we will not get a new state. Therefore, the permutation among n_s dots and the permutation among $(g_s - 1)$ partitions do not lead to any new state, and therefore, their permutations lead to overcounting, which has to be avoided. In our example, there are four objects. $(n_s = 3, g_s = 2)$ permutation of next four objects should lead to 24 multiparticle states, of which only four states are distinct. Each of these four distinct states is accompanied by six redundant states arising from permutation of three dots, and so we have 24/6 distinct multiparticle states. Arguing in the same way, we get the number of distinct multiparticle states $\Omega(n_s, g_s)$ as

$$\Omega(n_s, g_s) = \frac{(g_s + n_s - 1)!}{n_s!(g_s - 1)!} \tag{11.63}$$

11.10.3 Case 3: Indistinguishable Particles – Fermions

In the case of fermions, we have to take into account of Pauli's exclusion principle. Since the particles are indistinguishable, we can replace them with dots. Each box (state) can accommodate either zero or one particle. For instance if $n_s = g_s = 2$, we have

$$\begin{array}{|c|c|}\hline \bullet & \bullet \\ \hline \end{array} = u_\alpha(a)u_\beta(b) - u_\alpha(b)u_\beta(a)$$

If $g_s = 4$ and $n_s = 3$, then we have

One particle in states u_α, u_β, u_γ $\quad\begin{array}{|c|c|c|c|}\hline \bullet & \bullet & \bullet & \\ \hline \end{array}$

One in u_α, u_β, u_δ $\quad\begin{array}{|c|c|c|c|}\hline \bullet & \bullet & & \bullet \\ \hline \end{array}$

One in u_α, u_γ, u_δ $\quad\begin{array}{|c|c|c|c|}\hline \bullet & & \bullet & \bullet \\ \hline \end{array}$

One in u_β, u_γ, u_δ $\quad\begin{array}{|c|c|c|c|}\hline & \bullet & \bullet & \bullet \\ \hline \end{array}$

Obviously, this process amounts to selecting n_s boxes from available g_s boxes. Therefore, the number of possible multiparticle states is

$$\Omega(n_s, g_s) = \frac{g_s!}{(g_s - n_s)!n_s!} \tag{11.64}$$

11.10.4 Factorization of Canonical Partition Function for Distinguishable Particles

In evaluating the canonical partition function in statistical mechanics, the property distinguishability of particles plays an important role. Let us consider a system of N particles in equilibrium at temperature.

The state ψ_j of the system of N particles is given by

$$H\psi_j(a, b, \ldots\ldots) = E_j\psi_j(a, b, \ldots\ldots)$$

The canonical partition function Z for this system is defined as

$$Z = \sum_j e^{-\beta E_j}$$

where the summation j is overall the possible multiparticle states of the system. This is true for distinguishable as well as indistinguishable particles. However, the system of non-interacting distinguishable particles is a special case since Z becomes factorizable as explained below. Let the single particle state u_α be defined as

$$hu_\alpha = \varepsilon_\alpha u_\alpha$$

For non-interacting distinguishable particles, Z can be written as

$$Z = z^N$$

where z is given by
$$z = \sum_i e^{-\beta \varepsilon_i}$$

Here the summation i is over all the possible single particle states. This factorization is true only for distinguishable particles. Similarly, the grand canonical partition function is factorizable only for free indistinguishable particles.

Example 11.6 Two single-particle states u_1 and u_2 of energy ε_1 and ε_2 are available for three distinguishable particles a, b and c. Evaluate the canonical partition function Z and show that it can be written as

$$Z = z^3 = (e^{-\beta\varepsilon_1} + e^{-\beta\varepsilon_2})^3$$

Solution: There are eight multiparticle states, and they are listed below:

$$\psi_1 = \begin{array}{|c|c|}\hline abc & \\ \hline\end{array}, \text{with } E_1 = 3\varepsilon_1 \qquad \psi_2 = \begin{array}{|c|c|}\hline a & bc \\ \hline\end{array}, \text{with } E_2 = \varepsilon_1 + 2\varepsilon_2$$

$$\psi_3 = \begin{array}{|c|c|}\hline b & ac \\ \hline\end{array}, \text{with } E_3 = \varepsilon_1 + 2\varepsilon_2 \qquad \psi_4 = \begin{array}{|c|c|}\hline c & ab \\ \hline\end{array}, \text{with } E_4 = \varepsilon_1 + 2\varepsilon_2$$

$$\psi_5 = \begin{array}{|c|c|}\hline ab & c \\ \hline\end{array}, \text{with } E_5 = 2\varepsilon_1 + \varepsilon_2 \qquad \psi_6 = \begin{array}{|c|c|}\hline ac & b \\ \hline\end{array}, \text{with } E_6 = 2\varepsilon_1 + \varepsilon_2$$

$$\psi_7 = \begin{array}{|c|c|}\hline bc & a \\ \hline\end{array}, \text{with } E_7 = 2\varepsilon_1 + \varepsilon_2 \qquad \psi_8 = \begin{array}{|c|c|}\hline & abc \\ \hline\end{array}, \text{with } E_8 = 3\varepsilon_2$$

$$\begin{aligned} Z &= e^{-3\beta\varepsilon_1} + \underbrace{e^{-\beta\varepsilon_1 - 2\beta\varepsilon_2} + e^{-\beta\varepsilon_1 - 2\beta\varepsilon_2} + e^{-\beta\varepsilon_1 - 2\beta\varepsilon_2}}_{3e^{-\beta\varepsilon_1 - 2\beta\varepsilon_2}} + \underbrace{e^{-2\beta\varepsilon_1 - \beta\varepsilon_2} + e^{-2\beta\varepsilon_1 - \beta\varepsilon_2} + e^{-2\beta\varepsilon_1 - \beta\varepsilon_2}}_{3e^{-2\beta\varepsilon_1 - \beta\varepsilon_2}} + e^{-3\beta\varepsilon_2} \\ &= e^{-3\beta\varepsilon_1} + 3e^{-\beta\varepsilon_1 - 2\beta\varepsilon_2} + 3e^{-2\beta\varepsilon_1 - \beta\varepsilon_2} + e^{-3\beta\varepsilon_2} \\ &= (e^{-\beta\varepsilon_1} + e^{-\beta\varepsilon_2})^3 \end{aligned}$$

Let us compare this expression for the canonical partition function for bosons. Note for bosons, the multiparticle states are as listed below:

$$\phi_1 = \begin{array}{|c|c|}\hline \bullet\bullet\bullet & \\ \hline\end{array}, \text{with } E_1 = 3\varepsilon_1 \qquad \phi_2 = \begin{array}{|c|c|}\hline \bullet & \bullet\bullet \\ \hline\end{array}, \text{with } E_2 = \varepsilon_1 + 2\varepsilon_2$$

$$\phi_3 = \begin{array}{|c|c|}\hline \bullet\bullet & \bullet \\ \hline\end{array}, \text{with } E_3 = 2\varepsilon_1 + \varepsilon_2 \qquad \phi_4 = \begin{array}{|c|c|}\hline & \bullet\bullet\bullet \\ \hline\end{array}, \text{with } E_3 = 3\varepsilon_2$$

$$Z = e^{-3\beta\varepsilon_1} + e^{-\beta\varepsilon_1 - 2\beta\varepsilon_2} + e^{-2\beta\varepsilon_1 - \beta\varepsilon_2} + e^{-3\beta\varepsilon_2}$$

Obviously, this term is not factorizable.

11.11 DENSITY MATRIX

Let us now turn our attention to another concept called density matrix, which plays an important role in statistical mechanics.

Let us consider an ensemble of particles. One type of ensemble is a collection of particles in which all the particles are in the same state. This is like a collection of people with the same height or weight. This is a rare type of collection. Normally, ensembles will look like a collection of people with different heights or weights. So, in general, the particles in an ensemble have physical properties whose values vary over a wide range. For instance, in the case of kinetic theory of gases, there is a statistical distribution of particles with different velocities: n_1 particles with velocity $v_1; n_2$ particles with velocity v_2 and so on. This statistical distribution is a result of different particles in different states.

But, within quantum mechanics, the sources of statistical distribution of experimental outcomes are twofold. One source is the statistical distribution of state of particles; i.e., n_1 particles in state ψ_1, n_2 particles in state ψ_2 and so on. The second source is inherent to quantum mechanics. Even if all the particles are in the same state ψ, there will be a statistical distribution of eigenvalues as experimental outcomes.

Based on the above considerations, we divide the ensembles of quantum particles into two kinds: mixed ensemble and pure ensemble.

Pure State

The ensemble is pure if every system (or every particle) is in the same state. The state of all such systems may be described by $\psi = \Sigma C_n \phi_n$.

Mixed State

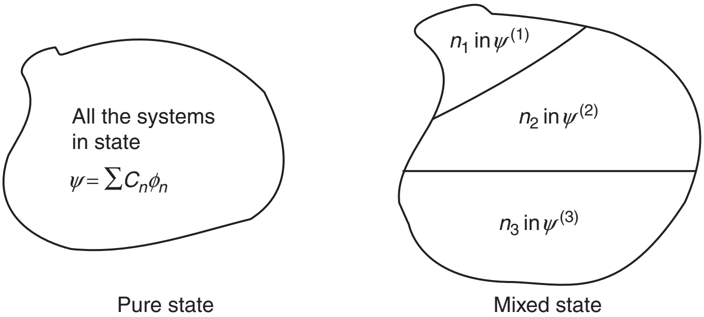
The ensemble is a mixed ensemble if the states of different systems (particles) are different.

$$n_1 \text{ particles in state } \psi^{(1)} = \sum C_n^{(1)} \phi_n$$

$$n_2 \text{ particles in state } \psi^{(2)} = \sum C_n^{(2)} \phi_n$$

$$\vdots$$

This is illustrated schematically in Fig. 11.7.



Pure state

Mixed state

Fig. 11.7 Ensembles of different kinds

Can we write a single wave function for all the particles in a mixed ensemble?

If we consider a particular particle in this ensemble, we cannot tell its state (by specifying $\psi^{(i)}$ or by specifying the set $\{c_n^{(i)}\}$). We can tell only the probability of finding the particle in states $\psi^{(i)}$ (or the set $\{c_n^{(i)}\}$). When we write a wave function for a system particle, it implies that we know all the possible information about the system. We cannot write a single wave function in this case since we lack some information for a particle in the mixed ensemble, and so we cannot write a definite set of coefficient $\{c_n\}$ representing all the particles in this ensemble.

11.12 PURE STATE DENSITY MATRIX

Let us first start with pure state ensemble. The state of each system in the ensemble is given by

$$|\psi\rangle = \sum C_n |n\rangle$$

where the set $\{|n\rangle\}$ is the eigenkets of a dynamical variable A.

In matrix representation, the state vector ψ is given by

$$\psi = \begin{pmatrix} C_1 \\ C_2 \\ \vdots \end{pmatrix}$$

The density matrix is defined by

$$\rho = \psi\psi^\dagger = \begin{pmatrix} C_1 \\ C_2 \\ \vdots \end{pmatrix} (C_1^* C_2^* ..) \tag{11.65}$$

$$\rho_{mn} = C_m C_n^* \tag{11.66}$$

In terms of Dirac's ket vectors, the density operation $\hat{\rho}$ is designed as

$$\hat{\rho} = |\psi\rangle\langle\psi| \tag{11.67}$$

$$\rho_{mn} = \langle m|\hat{\rho}|n\rangle = \langle m|\psi\rangle\langle\psi|n\rangle = C_m C_n^* \tag{11.68}$$

ρ Hermitian

The density matrix ρ is a Hermitian matrix

$$\rho^\dagger = (\psi\psi^\dagger)^\dagger = (\psi^\dagger)^\dagger\psi^\dagger = \psi\psi^\dagger \tag{11.69}$$

In terms of matrix elements,

$$(\rho_{nm})^* = (C_n C_m^*)^* = C_m C_n^* = \rho_{mn} \tag{11.70}$$

Tr (ρ) = 1

$$Tr(\rho) = \sum_n \rho_{nn} = \sum_n C_n C_n^* = \sum |C_n|^2 = 1 \tag{11.71}$$

$\rho^2 = \rho$

$$\hat{\rho}^2 = \hat{\rho}\hat{\rho} = |\psi\rangle\langle\psi|\psi\rangle\langle\psi| = |\psi\rangle\langle\psi| = \hat{\rho} \tag{11.72}$$

Therefore, the matrices corresponding to $\hat{\rho}^2$ and $\hat{\rho}$ should be the same.

ρ-Non-negative

$\hat{\rho}$ is non-negative definite, which means $\langle\phi|\hat{\rho}|\phi\rangle$ is always a positive number.

i.e., $\langle\phi|\hat{\rho}|\phi\rangle \geq 0$

Let $$|\phi\rangle = \sum a_n |n\rangle \quad \text{and} \quad \langle\phi| = \sum a_m^* \langle m|$$

$$\begin{aligned}\therefore \langle\phi|\hat{\rho}|\phi\rangle &= \sum_n \sum_m \langle m|a_m^*\hat{\rho}a_n|n\rangle \\ &= \sum_n \sum_m a_m^* a_n \langle m|\hat{\rho}|n\rangle = \sum_n \sum_m a_m^* a_n C_m C_n^* \\ &= \sum_n a_n C_n^* \sum_m a_m^* C_m = \left|\sum_x a_n C_n^*\right|^2 \geq 0\end{aligned} \tag{11.73}$$

Tr (ρ^2) = 1

Since $$\rho^2 = \rho, Tr(\rho^2) = Tr(\rho) = 1 \tag{11.74}$$

Expectation Values of F

$$\begin{aligned}\langle F\rangle &= \langle\psi|\hat{F}|\psi\rangle = \langle\psi|\hat{1}\hat{F}\hat{1}|\psi\rangle \\ &= \sum_n \sum_m \langle\psi|n\rangle\langle n|\hat{F}|m\rangle\langle m|\psi\rangle \\ &= \sum_n \sum_m C_n^* F_{nm} C_m = \sum_n \sum_m F_{nm} C_m C_n^* \\ &= \sum_n (F\rho)_{nn} = Tr(F\rho)\end{aligned} \tag{11.75}$$

where F is a matrix operator

Diagonal Elements of ρ

The diagonal element ρ_{nn} is given by

$$\rho_{nn} = C_n C_n^* = |C_n|^2$$

Therefore, ρ_{nn} is the probability of obtaining the value a_n in a measurement of A.

Expectation Value of the Projection Operator

The projection operator $\hat{P}_i$ is given by

$$
\begin{aligned}
\hat{P}_i &= |\, i\rangle\langle i\,| \\
\langle P_i\rangle &= Tr(P_i\rho) \\
\sum_n \langle n\,|\,\hat{P}_i\hat{\rho}\,|\,n\rangle &= \sum_n \langle n\,|\,\hat{P}\,|\,i\rangle\langle i\,|\,\hat{\rho}\,|\,n\rangle \\
&= \sum_n \delta_{ni} C_i C_n^{*} = C_i C_i^{*} = |\,C_i\,|^2
\end{aligned}
\tag{11.76}
$$

Time Evolution of Density Operator

The state vector $|\,\psi(t)\rangle$ is given by

$$
\begin{aligned}
|\,\psi(t)\rangle &= \sum C_n(t)\,|\,n\rangle \\
i\hbar\frac{\partial}{\partial t}|\,\psi(t)\rangle &= \hat{H}\,|\,\psi(t)\rangle \\
-i\hbar\frac{\partial}{\partial t}\langle\psi(t)\,| &= \langle\psi(t)\,|\,\hat{H}
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial}{\partial t}\hat{\rho} &= \frac{\partial}{\partial t}|\,\psi(t)\rangle(\langle\psi(t)\,|) + |\,\psi(t)\rangle\frac{\partial}{\partial t}\langle\psi(t)\,| \\
&= \frac{1}{i\hbar}[\hat{H}\,|\,\psi(t)\rangle\langle\psi(t)\,| - |\,\psi(t)\rangle\langle\psi(\mathrm{t})\,|\,\hat{H}] \\
&= \frac{1}{i\hbar}[\hat{H}\hat{\rho} - \hat{\rho}\hat{H}] = \frac{1}{i\hbar}[\hat{H},\hat{\rho}]
\end{aligned}
\tag{11.77}
$$

In the matrix representation, this result becomes

$$
\frac{\partial\rho}{\partial t} = \frac{1}{i\hbar}[\rho, H]
\tag{11.78}
$$

Diagonal Matrix ρ_d for a Pure State

Since ρ is a Hermitian matrix, it can be diagonalized by a unitary matrix U. The diagonal matrix is given by

$$
\rho_d = U\rho U^{\dagger} = U\psi\psi^{\dagger}U^{\dagger} = (U\psi)(U\psi)^{\dagger}
\tag{11.79}
$$

Obviously, the elements of a diagonal matrix ρ_d will be the eigenvalues of the density matrix ρ. For pure ensemble, only one eigenvalue will be 1 and the remaining eigenvalues are zero.

$$
\begin{pmatrix}
0 & & & & & & \\
 & \bullet & & & & & \\
 & & \bullet & & & & \\
 & & & 1 & & & \\
 & & & & \bullet & & \\
 & & & & & \bullet & \\
 & & & & & & 0
\end{pmatrix}
\tag{11.80}
$$

This can be seen as follows. The eigenvalue equation is

$$\rho\chi = \lambda\chi$$

Consider $\rho^2\chi$.

$$\rho^2\chi = \rho\underbrace{\rho\chi}_{\lambda\chi} = \lambda\underbrace{\rho\chi}_{\lambda\chi} = \lambda^2\chi$$

But since $\rho^2 = \rho$, we also have $\rho^2\chi = \rho\chi = \lambda\chi$.

$$\lambda\chi = \lambda^2\chi \Rightarrow \lambda = 0 \text{ or } 1$$

The fact that the diagonal elements of ρ_d are the eigenvalues of ρ and the fact that its trace is 1 together make all the elements of ρ_d to be zero except one element whose value is 1.

U transforms the basis states $\{|u_n\rangle\}$ to a new basis state $\{|v_n\rangle\}$. In the basis states $\{|u_n\rangle\}$, the state vector $|\psi\rangle$ is a superposition of basis vectors $|u_n\rangle$ given as

$$|\psi\rangle = \sum_n C_n\,|u_n\rangle \text{ leading to } \rho = \begin{bmatrix} C_1C_1^* & C_1C_2^* & \cdots & \cdots \\ C_2C_1^* & C_2C_2^* & \cdots & \cdots \\ \bullet & \bullet & & \\ \bullet & \bullet & & \\ \bullet & \bullet & & \\ & & & \end{bmatrix}$$

i.e., $|\psi\rangle$. Now in the new basis $\{|v_n\rangle\}$, the state vector becomes

$$|\psi\rangle = |v_n\rangle \text{ leading to } \rho_d = \begin{bmatrix} 0 & 0 & \cdots & \cdots \\ 0 & 1 & \cdots & \cdots \\ \bullet & \bullet & & \\ \bullet & \bullet & 0 & \\ \bullet & \bullet & & \\ & & & 0 \end{bmatrix}$$

If $\{|u_n\rangle\}$ are the eigenvectors of a dynamical variable A, we can always find an operator A' for which $|v_n\rangle = \Sigma C_n\,|u_n\rangle$ will be an eigenvector. Of course, A' is given by

$$UAU^\dagger = A'$$

Example 11.7 Obtain the eigenvalues of the density matrix for an ensemble of spin 1/2 particles whose state vector is given by

$$\psi = C_1\chi_{\frac{1}{2}} + C_2\chi_{-\frac{1}{2}}$$

Solution: The density matrix ρ is given by

$$\rho = \begin{pmatrix} C_1C_1^* & C_1C_2^* \\ C_2C_1^* & C_2C_2^* \end{pmatrix} \text{ with } |C_1|^2 + |C_2|^2 = 1$$

The eigenvalues equation is

$$\begin{vmatrix} |C_1|^2-\lambda & C_1C_2^* \\ C_2C_1^* & |C_2|^2-\lambda \end{vmatrix} = 0$$

$$(|C_1|^2-\lambda)(|C_2|^2-\lambda)-|C_1|^2|C_2|^2=0$$

$$\lambda^2-\lambda(|C_1|^2+|C_2|^2)=0$$

$$\lambda^2-\lambda=0 \qquad \lambda(\lambda-1)=0 \quad \therefore \lambda=0 \text{ or } \lambda=1$$

11.13 MIXED STATE DENSITY MATRIX

The above formulation of density matrix for pure ensemble can be extended to a mixed ensemble. Let n_i be the number of particles in state $|\psi^{(i)}\rangle$. If N is the total member of the particles in the mixed ensemble, then the density operator $\hat{\rho}$ is defined as

$$\hat{\rho}=\sum\left(\frac{n_i}{N}\right)\hat{\rho}^{(i)} \tag{11.81}$$

$$=\sum_i p_i\,|\psi^{(i)}\rangle\langle\psi^{(i)}| \tag{11.82}$$

where
$$|\psi^{(i)}\rangle=\sum_n C_n^{(i)}\,|n\rangle \tag{11.83}$$

Note that the different state vector $|\psi^{(i)}\rangle$ need not be orthogonal to each other.

i.e.,
$$\langle\psi^{(i)}|\psi^{(j)}\rangle\neq\delta_{ij} \text{ (in general)} \tag{11.84}$$

p_i can be taken to represent the fraction of the total number of particles in state $|\psi^{(i)}\rangle$. p_i can also be interpreted as the probability of finding the system in state $|\psi^{(i)}\rangle$.

The matrix ρ_{mn} is given by

$$\begin{aligned}\rho_{mn}&=\langle m|\hat{\rho}|n\rangle=\sum_i\langle m|p_i|\psi^{(i)}\rangle\langle\psi^{(i)}|n\rangle\\ \sum_i p_i\langle m|\psi^{(i)}\rangle\langle\psi^{(i)}|n\rangle&=\sum_i p_i\rho_{mn}^{(i)}=\sum_i p_iC_m^{(i)}C_n^{(i)}*\end{aligned} \tag{11.85}$$

$$\therefore \rho=\sum_i p_i\ \psi^{(i)}\psi^{(i)\dagger} \tag{11.86}$$

It shares almost all the properties of the density for pure ensemble:

It is a Hermitian matrix and non-negative definite matrix.

$$\langle\phi|\hat{\rho}|\phi\rangle=\langle\phi|\sum_i p_i\hat{\rho}_i|\phi\rangle=\sum_i p_i\langle\phi|\hat{\rho}_i|\phi\rangle$$

Making use of (11.73), we get

$$\langle\phi|\hat{\rho}|\phi\rangle=\sum_i p_i\left|\sum_n a_nC_n^{(i)}*\right|^2\geq 0$$

Trρ = 1

$$\mathrm{Tr}\rho = \sum_n \rho_{nn} = \sum_n \sum_i C_n^{(i)} C_n^{(i)*} = \sum_i p_i \sum_n |C_n^{(i)}|^2 = \sum p_i = 1$$

$\langle$F$\rangle$ = Tr(Fρ)

The expectation value of F in state $\langle \psi^{(\mathrm{i})} \rangle$ is given by

$$\langle F \rangle^{(i)} = \langle \psi^{(i)} \mid \hat{F} \mid \psi^{(i)} \rangle$$

This expectation value of $\langle F^{(i)} \rangle$ is the consequence of probabilistic interpretation intrinsic to quantum mechanics. The classical statistical average due to lack of information is given by

$$\begin{aligned}
\langle A \rangle &= \sum p_i \langle F \rangle^i = \sum_i p_i \langle \psi^{(i)} \mid \hat{F} \mid \psi^{(i)} \rangle \\
&= \sum_i p_i \sum_m \sum_n \langle \psi^{(i)} \mid m \rangle \langle m \mid \hat{F} \mid n \rangle \langle n \mid \psi^{(i)} \rangle \\
&= \sum_m \sum_n F_{mn} \sum_i p_i C_n^i C_m^{(i)*} \\
&= \sum_m \sum_n F_{mn} \rho_{nm} = Tr(F\rho) \qquad (11.87)
\end{aligned}$$

The Equation Motion d$\hat{\rho}$/dt

$$\begin{aligned}
\frac{d\hat{\rho}}{dt} &= \sum_i p_i \frac{d}{dt} (|\psi^{(i)}\rangle\langle\psi^{(i)}|) \\
&= \frac{1}{i\hbar}\left[\sum_i p_i \hat{H} \,|\psi^{(i)}\rangle\langle\psi^{(i)}| - \psi \sum_i p_i \,|\psi^{(i)}\rangle\langle\psi^{(i)}|\, \hat{H}\right] \\
&= \frac{1}{i\hbar}\left[\hat{H} \sum_i p_i \,|\psi^{(i)}\rangle\langle\psi^{(i)}| - \sum_i p_i \,|\psi^{(i)}\rangle\langle\psi^{(i)}|\, \hat{H}\right] \qquad (11.88) \\
&= \frac{1}{i\hbar}[\hat{H}, \hat{\rho}]
\end{aligned}$$

Trρ^2 for Mixed Ensemble

It is this property which is different from that of pure ensemble.

$$\begin{aligned}
Tr(\rho^2) = Tr(\rho\rho) &= \sum_n \sum_m \rho_{nm} \rho_{mn} \\
&= \sum_n \left(\sum_i p_i C_n^{(i)} C_m^{(i)*} \right) \left(\sum_j p_j C_m^{(j)} C_n^{(j)*} \right)
\end{aligned}$$

$$= \sum_i \sum_j p_i p_j \sum_n \sum_m \langle \psi^{(i)} | m \rangle \underbrace{\langle n | \psi^{(i)} \rangle} \overbrace{\langle m | \psi^{(j)} \rangle} \langle \psi^{(j)} | n \rangle$$

$$= \sum_i \sum_j p_i p_j \sum_n \sum_m \langle \psi^{(i)} | m \rangle \langle m | \psi^{(j)} \rangle \langle \psi^{(j)} | n \rangle \langle n | \psi^{(i)} \rangle$$

Let us use Schwartz inequality now.

$$\langle \psi^{(i)} | \psi^{(j)} \rangle^2 \le \langle \psi^{(i)} | \psi^{(i)} \rangle \langle \psi^{(j)} | \psi^{(j)} \rangle$$

(This is similar to $(\mathbf{A} \bullet \mathbf{B}) = A^2 B^2 \cos^2 \theta \le A^2 B^2$ in geometrical vectors)

$$\begin{aligned} Tr\rho^2 &\le \sum_i \sum_j p_i p_j \\ Tr\rho^2 &\le 1 \end{aligned} \tag{11.89}$$

Eigenvalues of Density Matrix for Mixed Ensemble

Since the matrix is non-negative definite, its eigenvalues are always positive.

$$\hat{\rho} | \chi \rangle = \lambda | \chi \rangle$$
$$\langle \chi | \hat{\rho} | \chi \rangle = \lambda \langle \chi | \chi \rangle = \lambda$$

Since $\langle \chi | \hat{\rho} | \chi \rangle \ge 0$, we conclude that $\lambda \ge 0$.

Let $\lambda_1, \lambda_2 \ldots$ be the eigenvalues of ρ. The two properties, namely $\lambda_i \ge 0$ and $\lambda_1 + \lambda_2 + \ldots = 1$, imply $0 \le \lambda_1 \le 1$, $0 \le \lambda_2 \le 1$, $0 \le \lambda_3 \le 1 \ldots$

Let us diagonalize the matrix ρ by a unitary transformation U.

$$\rho_d = U \rho U^\dagger = \begin{pmatrix} \lambda_1 & & \\ & \lambda_2 & \\ & & \ddots \end{pmatrix} S$$

Now $U \rho^2 U^\dagger = U \rho U^\dagger U \rho U^\dagger = \rho_d^2$

$$= \begin{pmatrix} \lambda_1^2 & & \\ & \lambda_2^2 & \\ & & \ddots \end{pmatrix}$$

Making use of the property $TrU\rho^2U^\dagger = Tr\rho_d^2$, we get

$$Tr\rho^2 = Tr\rho_d^2 = (\lambda_1^2 + \lambda_2^2 + \ldots) \leq 1$$

Note: It is possible for $Tr(\rho^2) = 1$ provided

$$\lambda_1^2 + \lambda_2^2 \ldots = 1$$

This is possible only when all the eigenvalues are zero except only one eigenvalue which is equal to 1.

i.e., $$\lambda_i = 1, \quad \text{and} \quad \lambda_j = 0, \quad \text{if } j \neq i$$

This corresponds to a pure state.

Example 11.8 An ensemble consists of spin 1/2 particles in the state

$$|\psi\rangle = 0.8\,|Z\uparrow\rangle + 0.6\,|Z\downarrow\rangle$$

Determine the density matrix ρ and hence compute $\langle S_x \rangle$.

Solution:

$$P = \begin{pmatrix} C_1C_1^* & C_1{C_2}^* \\ C_2C_1^* & C_2{C_2}^* \end{pmatrix} = \begin{pmatrix} 0.64 & 0.48 \\ 0.48 & 0.64 \end{pmatrix}$$

$$S_x = \frac{\hbar}{2}\sigma_x = \begin{pmatrix} 0 & \frac{\hbar}{2} \\ \frac{\hbar}{2} & 0 \end{pmatrix}$$

$$\therefore \langle S_x \rangle = Tr\left[\begin{pmatrix} 0 & \frac{\hbar}{2} \\ \frac{\hbar}{2} & 0 \end{pmatrix}\begin{pmatrix} 0.64 & 0.48 \\ 0.48 & 0.48 \end{pmatrix}\right] = 0.96\hbar$$

Example 11.9 Consider a mixed ensemble in which 64% particles are spin up and 36% particles with spin down along Z axis. Determine $\langle S_x \rangle$.

Solution: The density matrices and combination of the density matrices corresponding to pure states $|\psi^{(1)}\rangle = |Z\uparrow\rangle$ and $|\psi^{(2)}\rangle = |Z\downarrow\rangle$

$$\rho^{(1)} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \text{ and } \rho^{(2)} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

$$\rho = p_1\rho^{(1)} + p_2\rho^{(2)} = \begin{pmatrix} 0.64 & 0 \\ 0 & 0.36 \end{pmatrix}$$

$$\langle S_x \rangle = Tr\left[\begin{pmatrix} 0 & \frac{\hbar}{2} \\ \frac{\hbar}{2} & 0 \end{pmatrix}\begin{pmatrix} 0.64 & 0 \\ 0 & 0.36 \end{pmatrix}\right] = 0$$

Note that in the first ensemble, $|\psi\rangle = 0.6\,|Z\uparrow\rangle + 0.6\,|Z\downarrow\rangle$. In the measurement of Z component of the spin, probability getting $+\hbar/2$ and $-\hbar/2$ are 0.64 and 0.36. In the second ensemble also, the probability of getting of Z component of the spin as $\hbar/2$and $-\hbar/2$ are 0.64 and 0.36, respectively. Yet they describe different states. In one ensemble, $\langle S_x\rangle = 0.96\hbar$, and in the second ensemble, $\langle S_x\rangle = 0$. We have already seen the significance of these two different situations when we discussed the nonclassical nature of superposition state in quantum mechanics (see Chapter5).

Example 11.10 Obtain the density matrix for a pure ensemble of spin 1/2 particles in the state. $\psi = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\1\end{bmatrix}$. Hence compute $\langle S_x\rangle$ and $\langle S_z\rangle$.

Solution: The density matrix ρ is given by

$$\rho = \psi\psi^{\dagger} = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix}\frac{1}{\sqrt{2}}(1 \quad 1) = \frac{1}{2}\begin{pmatrix}1 & 1\\1 & 1\end{pmatrix}$$

$$\langle S_Z\rangle = Tr(S_Z\rho) = Tr\left\{\begin{bmatrix}\frac{\hbar}{2} & 0\\ 0 & -\frac{\hbar}{2}\end{bmatrix}\frac{1}{2}\begin{bmatrix}1 & 1\\1 & 1\end{bmatrix}\right\} = 0$$

$$\langle S_x\rangle = Tr\left\{\frac{1}{2}\begin{bmatrix}0 & \frac{\hbar}{2}\\ \frac{\hbar}{2} & 0\end{bmatrix}\begin{bmatrix}1 & 1\\1 & 1\end{bmatrix}\right\} = \frac{\hbar}{2}$$

Since $\langle S_x\rangle = \frac{\hbar}{2}$, obviously $\psi = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\1\end{bmatrix}$ is an eigenstate of S_x.

$$S_x\phi_x(\uparrow) = \frac{\hbar}{2}\phi_x(\uparrow)$$

In this basis, $$\psi = C_1\phi_x(\uparrow) + C_2\phi_x(\downarrow)$$

For $$\psi = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix}, C_1 = 1, C_2 = 0$$

$$\therefore \rho_d = \begin{pmatrix}1 & 0\\0 & 0\end{pmatrix}$$

11.14 SPIN DENSITY MATRIX

Consider a beam of spin 1/2 particles in a mixed state. Since there are two basis vectors, the density matrix should be 2×2 matrix. From matrix theory, we know that any 2×2 Hermitian matrix can be written as a linear combination of the identity matrix and the three Pauli matrices σ_x, σ_y and σ_z. Thus,

$$\rho = C_0 I + C_1\sigma_x + C_2\sigma_y + C_3\sigma_z$$

$$= \begin{pmatrix}C_0 + C_3 & C_1 - iC_2\\ C_1 + iC_2 & C_0 - C_3\end{pmatrix}$$

$$\text{Tr}\rho = 2C_0 = 1 \quad \therefore C_0 = \frac{1}{2}$$

$$\rho = \frac{1}{2}I + C_1\sigma_x + C_2\sigma_y + C_3\sigma_z$$

$$= \frac{1}{2}(I + 2C_1\sigma_x + 2C_2\sigma_y + 2C_3\sigma_3)$$

Instead if $2C_1$, $2C_2$ and $2C_3$, let us write them as P_x, P_y and P_z.

$$\therefore \rho = \frac{1}{2}(I + \boldsymbol{\sigma} \cdot \mathbf{P})$$

It is easy to show that $\langle S_x \rangle = P_x \cdot \frac{\hbar}{2}$, $\langle S_y \rangle = P_y \frac{\hbar}{2}$ and $\langle S_z \rangle = P_z \frac{\hbar}{2}$.

So P_x, P_y and P_z can be interpreted as the components of the polarization vector **P**.

EXERCISES

1. How do you define identical particles?
2. What is meant by indistinguishability?
3. Is indistinguishability the same as identical?
4. Is indistinguishability unique to quantum mechanics?
5. Is the concept of indistinguishability present in classical mechanics?
6. Can we call proton and neutron as identical particles?
7. Can we call electron and positron as identical particles?
8. What is meant by interchange operator?
9. Show that the interchange operator P_{12} is unitary operator.
10. Is $x_1p_2 - x_2p_1$ an allowed operator in quantum mechanics? (Subscripts 1 and 2 refer to particle 1 and particle 2.)
11. By considering three-particle wave function $\psi(1,2,3) = u_a(1)\,u_\gamma(2)\,u_\beta(3)$, show that $P_{12}P_{13} \neq P_{13}P_{12}$.
12. For the three particles, wave function ψ_s (1, 2, 3) given by (11.28). Show that $P_{12}\,P_{13} = P_{13}\,P_{12}$.
13. For a system of fermions, is it true that the operators corresponding to dynamical variable should be antisymmetric, i.e., is it true that $A_{op}(\mathbf{r}_1\text{s}_1;\mathbf{r}_2\text{s}_2) = -A_{op}(\mathbf{r}_2\text{s}_2;\mathbf{r}_1\text{s}_1)$?
14. What is the significance of symmetrization postulate?
15. The single particle energy level ε_s is two-fold degenerate. Determine the number of multiparticle states for four distinguishable particles. If the two particles are indistinguishable, how many multiparticle states are possible?
16. Consider a system of two particles for which four states of energy $\varepsilon_1, \varepsilon_2, \varepsilon_3$ and ε_4 are available to them. Determine the canonical partition function for the system, assuming the particles to be (i) distinguishable (ii) bosons and (iii) fermions.
17. Show that Pauli's exclusion principle is a consequence of antisymmetrization of a wave function of the multiparticle system.
18. In a pure ensemble, the wave function of each particle is given by

$$\psi = \frac{1}{\sqrt{2}}\phi_1 + \frac{1}{\sqrt{3}}\phi_2 + \frac{1}{\sqrt{6}}\phi_3$$

Obtain the density matrix ρ. Determine eigenvalues of the density matrix. Determine ρ^2. Diagonalize the matrix ρ.

19. In an ensemble 60% of the systems are in the state $\psi = 1/\sqrt{2}\phi_1 + 1/\sqrt{3}\phi_2 + 1/\sqrt{6}\phi_3$ and 40% of them are in the state $\psi = 1/\sqrt{3}\phi_1 + 1/\sqrt{3}\phi_2 + 1/\sqrt{3}\phi_3$. Determine (i) ρ (ii) ρ^2 and (iii) the eigenvalues of ρ and ρ^2.

REFERENCES

1. L. Ballentine, 1998. *Quantum Mechanics – A Modern Development*. Singapore: World Scientific Publishing Co. Pte. Ltd.
2. eg. bucknell.edu/~koutslts/Ph332/Text/Chapt8_JQPMaster.pdf.
3. J. D. McGervy, 1971. *Introduction to Modern Physics*. New York: Academic Press.
4. H. B. Callen, 1985. *Thermodynamics and an Introduction to Thermostatistics*, New York: John Wiley & Sons.
5. P. M. Mathews and K. Venkatesan, 1976. *A Text Book of Quantum Mechanics*. New Delhi: Tata Mcgraw-Hill Publishing Company Limited.
6. Frank Porter, http://www.cithep.caltech.edu/~fcp/physics/quantumMechanics/densityMatrix/densityMatrix.pdf5.

12

Electron in Magnetic Fields and Other Two-State Problems

We have already seen in the earlier chapters many problems for which exact solutions are available. In this chapter, we study some more exactly solvable problems. First, we consider the case of an electron in constant magnetic field. Problems involving electron in constant magnetic field are one of the simplest applications of quantum mechanics. For instance, if we consider orbital angular momentum alone, free electron in a constant magnetic field is exactly solvable, leading to Landau levels. On the other hand, if one concentrates on spin angular momentum alone, the electron in a magnetic field provides a simple model for the study of two-state systems.

Then, we consider gauge invariance in quantum mechanics. Though this principle is one of the major ideas in particle physics, here we restrict ourselves to Aharonov–Bohm effect, which brings out another important phenomenon where quantum mechanics and classical physics widely differ.

Next, we present the general analysis of two-state problems, which occur in wide variety of situations. In addition to the electron in a magnetic field, we consider ammonia molecule and neutrino oscillation. Though these problems are totally disconnected, they share important features of two-state problems.

12.1 ELECTRONS IN CLASSICAL ELECTROMAGNETIC FIELD

We present here the important features of classical electromagnetism relevant to the study of an electron in constant magnetic field. In classical electromagnetism, the electric field **E** and the magnetic field **B** are the basic quantities (SI units are used in this chapter). Maxwell's equations determine **E** and **B** for a given charge density and current density. On the other hand, Lorentz force describes the effect of **E** and **B** on the motion of a charged particle. In addition to **E** and **B**, two more concepts, namely the scalar potential ϕ and the vector potential **A**, are widely used in classical electromagnetism. They are related to **E** and **B** as follows:

$$\mathbf{E} = -\nabla\phi - \frac{\partial \mathbf{A}}{\partial t} \qquad \mathbf{B} = \nabla \times \mathbf{A} \tag{12.1}$$

In classical physics, we can study electromagnetism either using electric field and magnetic field or using the scalar potential and vector potential. But even in classical electromagnetism, Hamiltonian formalism can be done only using scalar potential and vector potential. So, in quantum mechanics, the interaction between a charged particle and electromagnetic field can be described using only scalar and vector potential.

Note that for a given **E** and **B**, **A** and ϕ cannot be chosen uniquely. There are many sets of ϕ and **A,** leading to the same fields **E** and **B**. This is expressed through a transformation called gauge transformation.

$$\phi \to \phi' = \phi - \frac{\partial \chi}{\partial t} \tag{12.2}$$

$$\mathbf{A} \to \mathbf{A}' = \mathbf{A} + \nabla \chi \tag{12.3}$$

Both sets $(\phi, \mathbf{A})$ and $(\phi', \mathbf{A}')$ lead to the same electric field **E** and magnetic field **B**.

$$\begin{aligned} \mathbf{E}' &= -\nabla \phi' - \frac{\partial \mathbf{A}'}{\partial t} = -\nabla \phi + \nabla \frac{\partial \chi}{\partial t} - \frac{\partial \mathbf{A}}{\partial t} - \frac{\partial}{\partial t} \nabla \chi \\ &= -\nabla \phi - \frac{\partial \mathbf{A}}{\partial t} = \mathbf{E} \end{aligned} \tag{12.4}$$

Similarly,

$$\mathbf{B}' = \nabla \times \mathbf{A}' = \nabla \times (\mathbf{A} + \nabla \chi) = \nabla \times \mathbf{A} + \nabla \times \nabla \chi = \nabla \times \mathbf{A} = \mathbf{B}$$

We will discuss the significance of gauge transformation in quantum mechanics a little later. We can take advantage of this arbitrariness and choose a potential for a given **B** to suit our convenience. Our interest in this chapter is the electron in a constant magnetic field. The following potentials lead to constant magnetic field **B**:

1. $\mathbf{A} = -\frac{1}{2} \mathbf{r} \times \mathbf{B} \qquad \nabla \times \mathbf{A} = \mathbf{B}$
2. $\mathbf{A} = (-yB, 0, 0) \qquad \nabla \times \mathbf{A} = B\mathbf{z}$
3. $\mathbf{A} = (0, xB, 0) \qquad \nabla \times \mathbf{A} = B\mathbf{z}$

Let us summarize other important results of classical electromagnetism which are also relevant to this chapter.

The Hamiltonian *H* for a charged particle in an electromagnetic field **E** and **B** is given by

$$\mathrm{H} = \frac{(\mathbf{p} - q\mathbf{A})^2}{2m} + V(\mathbf{r}) + q\phi \tag{12.5}$$

Magnetic Moment

A charged particle of charge *q* and angular momentum **L** has a magnetic moment μ given by

$$\mu = \frac{q}{2m} \mathbf{L} \tag{12.6}$$

For an electron of charge $-e$, the magnetic moment due to orbital angular momentum is

$$\mu = -\frac{e}{2m} \mathbf{L}$$

The potential energy of a dipole in the magnetic field **B** is given by

$$V = -\boldsymbol{\mu}\cdot\mathbf{B} \tag{12.7}$$

Larmor's Precession

A dipole of magnetic moment $\boldsymbol{\mu}$ in a magnetic field **B** experiences a torque $\boldsymbol{\tau} = \boldsymbol{\mu}\times\mathbf{B}$.

$$\boldsymbol{\tau} = \frac{d\mathbf{L}}{dt} = \boldsymbol{\mu}\times\mathbf{B}$$

A bar magnet in our laboratory is the simplest example of a dipole. If we place a bar magnet in a magnetic field, the torque on it tends to align it along the magnetic field.

In the case of a charged particle of angular momentum **L**, the torque induces a precessional motion of **L** or $\boldsymbol{\mu}$ around the field **B**.

This can be seen as follows:

$$\frac{d\mathbf{L}}{dt} = \frac{q}{2m}\mathbf{L}\times\mathbf{B}$$

Taking scalar product with **L**, we get

$$\mathbf{L}\cdot\frac{d\mathbf{L}}{dt} = \frac{1}{2}\frac{d}{dt}(L^2) = \mathbf{L}\cdot\left(\frac{q}{2m}\mathbf{L}\times\mathbf{B}\right) = 0$$

Therefore, L^2 does not change with time. Taking scalar product with magnetic field **B**, we get

$$\mathbf{B}\cdot\frac{d\mathbf{L}}{dt} = \frac{d}{dt}(\mathbf{B}\cdot\mathbf{L}) = \mathbf{B}\cdot\left(\frac{q}{2m}\mathbf{L}\times\mathbf{B}\right) = 0$$

Therefore, the angle between **L** and **B** does not change with time. These two results imply that **L** or $\boldsymbol{\mu}$ precesses around **L** with a frequency known as Larmor's frequency ω_L, which is given by

$$\omega_L = \frac{eB}{2m} \tag{12.8}$$

12.2 LANDAU ENERGY LEVELS

The concept of Landau levels serves as an important pedagogic tool in explaining quantum Hall effect. In solid state physics, Landau levels are used to explain de Haas-van Alphen effect in determining the Fermi surface. Landau levels are the energy levels of an electron in a constant magnetic field.

We consider a free electron subjected to constant magnetic field. We ignore the spin angular momentum here. The Hamiltonian for the electron of charge $-e$ is

$$H = \frac{1}{2m}(\mathbf{p}_{op} + e\mathbf{A})\cdot(\mathbf{p}_{op} + e\mathbf{A}) \tag{12.9}$$

The energy eigenvalue equation is

$$\frac{1}{2m}(\mathbf{p}_{op} + e\mathbf{A})\cdot(\mathbf{p}_{op} + e\mathbf{A})u(x,y,z) = Eu(x,y,z) \tag{12.10}$$

Let us choose the vector potential **A** as $\mathbf{A} = xB\hat{\mathbf{y}}$. It is easy to check that $\nabla\times(xB\hat{\mathbf{y}}) = B\hat{\mathbf{z}}$.

The Hamiltonian is

$$H = \frac{1}{2m}[\hat{\mathbf{x}}p_x + \hat{\mathbf{y}}(p_y + exB) + \hat{\mathbf{z}}p_z]^2 \tag{12.11}$$

This Hamiltonian commutes with p_y and p_z.

i.e., $[H, p_y] = [H, p_z] = 0$. Since the eigenfunctions of the operators p_y and p_z are $e^{ik_y y}$ and $e^{ik_z z}$, we write $u(x, y, z)$ as

$$u(x,y,z) = \phi(x)e^{ik_y y}e^{ik_z z} \tag{12.12}$$

The Schrödinger equation is

$$\left[-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + \frac{1}{2m}\left(-i\hbar\frac{\partial}{\partial y} + exB\right)^2 - \frac{\hbar^2}{2m}\frac{\partial^2}{\partial z^2}\right]u = Eu \tag{12.13}$$

$$\left[-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + \frac{1}{2m}\left(-i\hbar\frac{\partial}{\partial y} + exB\right)^2 - \frac{\hbar^2}{2m}\frac{\partial^2}{\partial z^2}\right]\phi(x)e^{ik_y y}e^{ik_z z} = E\phi(x)e^{ik_y y}e^{ik_z z} \tag{12.14}$$

Note
$$\left(-i\hbar\frac{\partial}{\partial y} + exB\right)^2 e^{ik_y y} = (\hbar k_y + exB)^2 e^{ik_y y} \tag{12.15}$$

and
$$-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial z^2}e^{ik_z z} = \frac{\hbar^2 k_z^2}{2m}e^{ik_z z} \tag{12.16}$$

Making use of these results in (12.14), we get

$$-\frac{\hbar^2}{2m}\frac{\partial^2\phi}{\partial x^2} + \frac{1}{2m}(\hbar k_y + exB)^2\phi + \frac{\hbar^2}{2m}k_z^2\phi = E\phi \tag{12.17}$$

This can be rewritten as

$$-\frac{\hbar^2}{2m}\frac{\partial^2\phi}{\partial x^2} + m\frac{e^2B^2}{2m^2}\left(x + \frac{\hbar k_y}{eB}\right)^2\phi = \left(E - \frac{\hbar^2 k_z^2}{2m}\right)\phi = \varepsilon\phi$$

$$-\frac{\hbar^2}{2m}\frac{\partial^2\phi}{\partial x^2} + \frac{1}{2}m\omega_0^2\,(x - x_0)^2\phi = \varepsilon\phi \tag{12.18}$$

where
$$\omega_0 = \frac{eB}{m} \text{ and } x_0 = -\frac{\hbar k_y}{eB} \text{ and } \varepsilon = E - \frac{\hbar^2 k_z^2}{2m}$$

The Equation (12.18) is the Schrödinger equation for harmonic oscillator of frequency ω_0 and centered around $x_0 = -\hbar k_y/eB$. Therefore, ε is given by

$$\varepsilon = \left(n + \frac{1}{2}\right)\hbar\omega_0$$

Or
$$E = \left(n + \frac{1}{2}\right)\hbar\omega_0 + \frac{\hbar^2 k_z^2}{2m} \tag{12.19}$$

The energy eigenfunction $u(x, y, z)$ is given by

$$u_{n,k_y,k_z} = N_n e^{-\frac{\alpha^2(x-x_0)^2}{2}} H_n(\alpha(x - x_0))e^{ik_y y}e^{ik_z z} \tag{12.20}$$

The energy levels (12.19) are known as Landau levels.

12.2.1 Degeneracy

It is to be noted that though the energy eigenfunctions u_n, k_y and k_z contain three quantum numbers n, k_y and k_z, the energy eigenvalues are independent of k_y. This implies that the Landau levels are degenerate. The degeneracy or the number of degenerate states corresponding to a given Landau level $(E_{n,k_z} = (n+1/2)\hbar\omega_0 + \hbar^2 k_z^2/2m)$ is determined by the boundary condition. It can be infinite or finite. For instance, in a solid of sides L_x, L_y and L_z, the degeneracy D is $D = \dfrac{L_x L_y eB}{2\pi\hbar}$.

Example 12.1 Determine the degeneracy corresponding to the Landau levels for an electron in a solid of dimensions L_x, L_y and L_z.

Solution: The wave function of an electron in a solid corresponding to Landau levels is

$$u_{n,k_y,k_z} = N_n e^{-\frac{\alpha^2(x-x_0)^2}{2}} H_n(\alpha(x-x_0)) e^{ik_y y} e^{ik_z z}$$

The boundary condition for an electron in a solid is given by

$$u_{n,k_y,k_z}(x, y+L_y, z+L_z) = u_{n,k_y,k_z}(x,y,z)$$

This implies $$e^{ik_y(y+L_y)} = e^{ik_y y} \quad \text{or} \quad e^{ik_y L_y} = 1$$

Therefore, the allowed values of k_y are $k_y = 2\pi n_y/L_y$ where $n_y = \pm1, \pm2, \pm3....$

In general, n_y can take any value. However, there is an upper bound for the n_y values for an electron inside a solid. This is due to the fact that the harmonic oscillator part of the wave function is centered around $x_0 = -\hbar k_y/eB$ and this has to be within the solid. This implies $0 \le \hbar k_y/eB \le L_x$. Therefore, the maximum value of $\hbar k_y/eB$is L_x. So the maximum value of n_y is $n_y = L_x L_y eB/2\pi\hbar$. So the degeneracy is $D = L_x L_y eB/2\pi\hbar$.

12.3 NORMAL ZEEMAN EFFECT

We now consider an atom in a uniform magnetic field B. The calculation is simplified if one ignores the spin angular momentum. In such cases, the effect of magnetic field is to lift $(2l+1)$-fold degeneracy (corresponding to $(2l+1)$ values of m_l), which causes a spectroscopic line to become three lines. This is known as normal Zeeman effect. The inclusion of spin angular momentum is more realistic, and it leads to anomalous Zeeman effect (see Chapter 13). Here we restrict ourselves with normal Zeeman effect, which occurs for atoms with total spin $S = 0$.

Since spin is not taken into account, the Hamiltonian is given by

$$H = \frac{(\mathbf{p}_{op} + e\mathbf{A})^2}{2m} + V(\mathbf{r}) \tag{12.21}$$

Let us choose the potential $\mathbf{A}$ as

$$\begin{aligned} \mathbf{A} &= -\frac{1}{2}\mathbf{r}\times\mathbf{B} \\ &= \frac{1}{2}\hat{\mathbf{x}}(zB_y - yB_z) + \frac{1}{2}\hat{\mathbf{y}}(xB_z - zB_x) + \frac{1}{2}\hat{\mathbf{z}}(yB_x - xB_y) \end{aligned} \tag{12.22}$$

It is easy to check $$\nabla \times (-1/2\mathbf{r}\times\mathbf{B}) = \mathbf{B} \tag{12.23}$$

The energy eigenvalue equation is

$$\left[\frac{1}{2m}(\mathbf{p}_{op}+e\mathbf{A})^2+V(\mathbf{r})\right]u=Eu \tag{12.24}$$

Let us evaluate $(\mathbf{p}_{op}+e\mathbf{A})^2u$

$$\begin{aligned}(\mathbf{p}_{op}+e\mathbf{A})^2u &= (\mathbf{p}_{op}+e\mathbf{A})\cdot(\mathbf{p}_{op}u+e\mathbf{A}u)\\ &= \mathbf{p}_{op}\cdot\mathbf{p}_{op}u+\mathbf{p}_{op}\cdot(e\mathbf{A}u)+e\mathbf{A}\cdot\mathbf{p}_{op}u+e^2A^2u\end{aligned}$$

$$\mathbf{p}_{op}\cdot(e\mathbf{A}u)=e(\mathbf{p}_{op}\cdot\mathbf{A})u+e\mathbf{A}\cdot\mathbf{p}_{op}u$$

But $$(\mathbf{p}_{op}\cdot\mathbf{A})=-i\hbar\,\nabla\cdot\mathbf{A}=0$$

We can check this result with $\mathbf{A}$ given in (12.22).

$$\therefore(\mathbf{p}_{op}+e\mathbf{A})^2u=p_{op}^2u+2e\mathbf{A}\cdot\mathbf{p}_{op}u+e^2A^2u$$

We can write the term $\mathbf{A}\cdot\mathbf{p}_{op}$ as

$$\mathbf{A}\cdot\mathbf{p}_{op}=-\frac{1}{2}\mathbf{r}\times\mathbf{B}\cdot\mathbf{p}_{op}=\frac{1}{2}\mathbf{r}\times\mathbf{p}_{op}\cdot\mathbf{B}=\frac{1}{2}\mathbf{L}\cdot\mathbf{B}$$

$$\therefore(\mathbf{p}_{op}+e\mathbf{A})^2u=p_{op}^2u+e\mathbf{L}\cdot\mathbf{B}\,\mathrm{u}+e^2A^2u$$

$$\begin{aligned}H&=\frac{p^2}{2m}+V(\mathbf{r})+\frac{e}{2m}\mathbf{L}\cdot\mathbf{B}+\frac{e^2}{8m}(B^2r^2-(\mathbf{r}\cdot\mathbf{B})^2)\\&=\frac{p^2}{2m}+V(\mathbf{r})-\mu\cdot\mathbf{B}+\frac{e^2}{8m}[B^2r^2-(\mathbf{r}\cdot\mathbf{B})^2]\end{aligned} \tag{12.25}$$

If we choose $\mathbf{B}=B_z\mathbf{z}$, then the Hamiltonian H is

$$H=\frac{p^2}{2m}+V(\mathbf{r})+\frac{e}{2m}BL_z+\frac{e^2}{8m}B^2(x^2+y^2) \tag{12.26}$$

$$H=\frac{p^2}{2m}+V(\mathbf{r})+\mu_B B\frac{L_z}{\hbar}+\frac{e^2B^2}{8m}(x^2+y^2) \tag{12.27}$$

where the μ_B is the Bohr magneton given by $\mu_B=e\hbar/2m$

Let us estimate the magnitude of the second and the third term. Take L_z as $m_l\hbar$. Substituting $\mu_B\sim 9.3\times10^{-24}J\cdot T^{-1}$, we get

$$\frac{\mu_B BL_z}{\hbar}\sim m_l\mu_B B\sim 9.3\times10^{-24}B\sim 10^{-23}B \tag{12.28}$$

where B is in Tesla.

For the third term, we can replace x and y by a_0, the Bohr radius.

$$\therefore \frac{e^2B^2}{8m}(x^2+y^2) \simeq \frac{e^2B^2}{4m}a_0{}^2 \simeq \frac{(1.6\times10^{-19}C)^2(0.5\times10^{-10}m)^2}{4\times9\times10^{-31}\,\text{Kg}}B^2$$
$$\simeq 10^{-31}B^2 \tag{12.29}$$

In terms of eV, the second term will be of the order 10^{-4} eV, which itself is very small, compared to the energy levels of an atom. (The ground state of hydrogen atom is −13.6 eV). The third term is much smaller, compared to the second term and so it can be safely ignored.

$$\therefore H = \frac{p^2}{2m} + V(\mathbf{r}) + \frac{\mu_B B}{\hbar}L_z \tag{12.30}$$

For central potential, L_z commutes with H. Therefore, the eigenfunctions of $H_0 = p^2/2m + \text{V(r)}$ are also the eigenfunctions of H. Let $\psi_{nlm}(r,\theta,\phi)$ are the eigenfunctions of H.

$$H_0\psi_{nlm} = E_{nl}\psi_{nlm} \quad \text{and} \quad L_z\psi_{nlm} = m_l\hbar\psi_{nlm}$$
$$\therefore H\psi_{nlm} = E_{nlm}\psi_{nlm} = (E_{nl} + m_l\mu_B B)\psi_{nlm} \tag{12.31}$$

In the absence of magnetic field, the energy E_{nl} is independent of the quantum number m_l and so each energy level is $(2l+1)$-fold degenerate. The presence of magnetic field splits each of this energy level into $(2l+1)$ closely spaced sublevels. The effect of magnetic field is to lift the $(2l+1)$-fold degeneracy. This is illustrated by considering $l = 2$ and $l = 1$ states with possible transitions. The spacing between each subenergy levels corresponding to both $l = 1$ and $l = 2$ is the same.

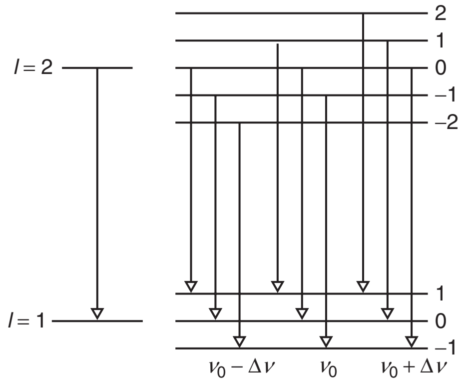


Fig. 12.1 Normal Zeeman effect

The numbers of transitions are restricted by the selection rules $\Delta l = 0, \Delta m_l = \pm1$.

$$\Delta\nu = \frac{\mu_B B}{\hbar}$$

12.4 GAUGE INVARIANCE OF SCHRÖDINGER EQUATION

Let us now introduce the concept of gauge invariance in quantum mechanics. Gauge invariances play a major role in elementary particle physics. However, our interest here is to study the gauge invariance of Schrödinger equation leading to an important phenomenon known as Aharonov–Bohm effect.

The Schrödinger equation for an electron in electromagnetic potential $(\phi, \mathbf{A})$ is

$$\left(i\hbar\frac{\partial}{\partial t}+e\phi\right)\psi=\frac{(\mathbf{p}_{op}+e\mathbf{A})^2}{2m}\psi+V\psi$$

We saw that the electric field **E** and the magnetic field **B** are invariant under the gauge transformation.

$$\phi\to\phi'=\phi-\frac{\partial\chi}{\partial t}$$
$$\mathbf{A}\to\mathbf{A}'=\mathbf{A}+\nabla\chi$$

What about the Schrödinger equation? Let ψ transforms to ψ' describing the particles in potentials $\mathbf{A}'$ and ϕ'. The invariance of Schrödinger equation under gauge transformation implies the following equations:

$$\left(i\hbar\frac{\partial}{\partial t}+e\phi\right)\psi=\frac{(\mathbf{p}_{op}+e\mathbf{A})^2}{2m}\psi+V\psi \tag{12.32}$$

$$\left(i\hbar\frac{\partial}{\partial t}+e\phi'\right)\psi'=\frac{(\mathbf{p}_{op}+e\mathbf{A}')^2}{2m}\psi'+V\psi' \tag{12.33}$$

If both sets $(\phi, \mathbf{A})$ and $(\phi', \mathbf{A}')$ describe the same physical situation, then ψ and ψ' should describe the same state. This is possible provided ψ and ψ' are related by a phase factor. Therefore, ψ' should be given by

$$\psi'=e^{i\xi(\mathbf{r},t)}\psi \tag{12.34}$$

Let us now verify whether ψ transforms to ψ' as given in (12.34). Let us consider $(i\hbar\partial/\partial t+e\phi')\psi'$ and $((\mathbf{p}_{op}+e\mathbf{A}'^2)/2m)\psi'$.

$$\begin{aligned}\left(i\hbar\frac{\partial}{\partial t}+e\phi'\right)\psi'&=\left(i\hbar\frac{\partial}{\partial t}+e\phi-e\frac{\partial\chi}{\partial t}\right)e^{i\xi}\psi\\&=e^{i\xi}\left[i\hbar\frac{\partial}{\partial t}+e\phi-e\frac{\partial\chi}{\partial t}-\hbar\frac{\partial\xi}{\partial t}\right]\psi\end{aligned}$$

$$\begin{aligned}(\mathbf{p}_{op}+e\mathbf{A}')\psi'&=(-i\hbar\nabla+e\mathbf{A}+e\nabla\chi)e^{i\xi}\psi\\&=e^{i\xi}(-i\hbar\nabla+e\mathbf{A}+e\nabla\chi+\hbar\nabla\xi)\psi\\&=e^{i\xi}(\mathbf{p}_{op}+e\mathbf{A}+e\nabla\chi+\hbar\nabla\xi)\psi\end{aligned}$$

$$
\begin{aligned}
&(\mathbf{p}_{op} + e\mathbf{A}')\cdot(\mathbf{p}_{op} + e\mathbf{A}')\psi' \\
&\qquad = (\mathbf{p}_{op} + e\mathbf{A} + e\nabla\chi)e^{i\xi}(\mathbf{p}_{op} + e\mathbf{A} + e\nabla\chi + \hbar\nabla\xi)\psi \\
&\qquad = e^{i\xi}(\mathbf{p}_{op} + e\mathbf{A} + e\nabla\chi + \hbar\nabla\xi)\cdot(\mathbf{p} + e\mathbf{A} + e\nabla\chi + \hbar\nabla\xi)\psi
\end{aligned}
$$

Therefore, the Equation(12.33) reduces to Equation (12.32) if

$$\hbar\frac{\partial\xi}{\partial t} = -e\frac{\partial\chi}{\partial t} \text{ and } \hbar\nabla\xi = -e\nabla\chi$$

Or

$$\xi = -\frac{e\chi}{\hbar}$$

Thus, the Schrödinger equation is invariant under gauge transformation provided $\psi' = e^{-ie\chi/\hbar}\psi$.

So the generalized gauge transformation is defined as

$$\phi \to \phi' = \phi - \frac{\partial\chi}{\partial t} \tag{12.35}$$

$$\mathbf{A} \to \mathbf{A}' = \mathbf{A} + \nabla\chi \tag{12.36}$$

$$\psi \to \psi' = e^{-ie\chi/\hbar}\psi \tag{12.37}$$

These transformations make sure that both sets $\{\phi, \mathbf{A}, \psi\}$ and $\{\phi' = \phi - \partial\chi/\partial t, \mathbf{A}' = \mathbf{A} + \nabla\chi, \psi' = e^{-ie\chi/\hbar}\psi\}$ describe the same physical states.

12.5 AHARONOV–BOHM EFFECT

Aharonov–Bohm effect raised serious questions about the nature of electric and magnetic fields and the vector potentials. Within classical physics, electric field and magnetic field are considered as real physical entities. They were deliberately introduced for the sake of avoiding action at a distance concept. A brief account of this concept is presented here.

In Newtonian mechanics, the general understanding is that physical contact between two objects is necessary for each object to act on the other. This is in agreement with everyday experience. The rail engine pulling the coaches is a best example. There was one exception to this scheme, namely gravitational force between two objects. The sun and the earth are separated by a vast distance. The sun does not pull the earth through any physical contact. Such actions are called actions at a distance. Can we formulate the forces describing the gravitation also in such a way to avoid the action at a distance?

Let us come back to electromagnetism, which also faced similar situations. For instance, consider the Coulomb force between charges q_1 at $\mathbf{r}_1$ and q_2 at $\mathbf{r}_2$. Though there is no physical contact between the charges q_1 and q_2, the charge q_2 experiences a force (action) q_1q_2/r^2. This is again an example of action at a distance. The concepts of electric and magnetic fields were introduced to avoid this possibility of action at a distance (or non-local interaction as generally known today). In terms of fields, the charge q_1 produces field $\mathbf{E}$ in the space around it, and $\mathbf{E}$ at $\mathbf{r}_2$ is given by $E = q_1/r^2$. There is a contact between the field $\mathbf{E}$ and the change q_2 at $\mathbf{r}_2$, and so it experiences a force $\mathbf{F} = q_2\mathbf{E} = (q_2q_1/r^2)\hat{\mathbf{r}}$. This is a simple account of the origin of fields in classical physics. Thus, in field theory, there is no action at a distance.

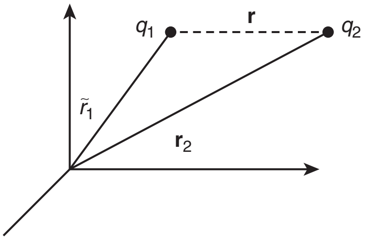
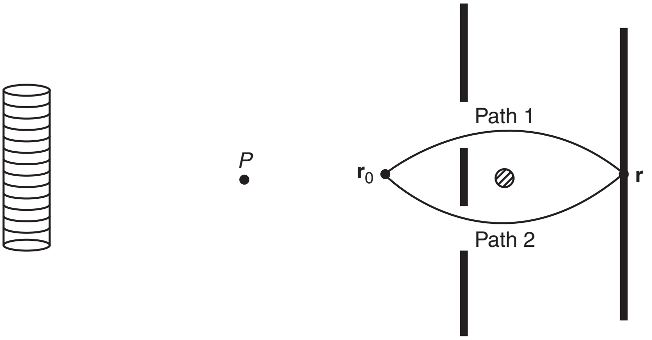


Fig. 12.2 Coulomb interaction between charges q_1 and q_2

We have already seen that in classical physics, the electric and magnetic fields are observable quantities, whereas the vector potentials are not measurable. Aharonov–Bohm effect challenges this simple picture. Consider a long, thin solenoid. When a current flows through the solenoid, magnetic field due to this current can be easily calculated. It is found that

$$\mathbf{B} = 0 \ \text{ Outside the solenoid}$$
$$\mathbf{B} \neq 0 \ \text{ Inside the solenoid}$$



(i) (ii)

Fig. 12.3 (i) Solenoid (ii) Aharonov–Bohm effect

Consider a point P well outside the solenoid. Since P is outside the solenoid, magnetic field $\mathbf{B}$ is zero at P. However, the vector potential may be or may not be zero at P, depending on whether $\mathbf{B}$ is zero or non-zero inside the solenoid. If the potential $\mathbf{A}$ is not zero, then it should be such that $\nabla \times \mathbf{A} = 0$. Obviously $\mathbf{A} = \nabla\chi$, where $\chi(\mathbf{r}, t)$ is an arbitrary function. Thus, we have the potential $\mathbf{A}$ at a point P outside the solenoid as given below:

$$\text{inside: } \mathbf{B} = 0; \qquad \text{outside: } \mathbf{B} = 0, \mathbf{A} = 0 \tag{12.38}$$

$$\text{inside: } \mathbf{B} \neq 0; \qquad \text{outside: } \mathbf{B} = 0, \mathbf{A} = \nabla\chi \tag{12.39}$$

Here, $\mathbf{A}$ has to be such that $\oint_C \mathbf{A} \cdot d\mathbf{r} = \Phi$, where Φ is the flux enclosed by the contour C. Note that the vector potential $\mathbf{A}$ can be arbitrary within gauge invariance. Let $\psi(\mathbf{r}, t)$ be the wave function of the electron at P corresponding to $\mathbf{A} = 0$. Then corresponding to $\mathbf{A} = \nabla\chi$, the wave function ψ' at P is given by

$$\psi' = e^{-ie\chi/\hbar}\psi \tag{12.40}$$

For a given **A**, χ can be calculated easily.

$$\mathbf{A}\cdot d\mathbf{r} = \nabla\chi\cdot d\mathbf{r} = \frac{\partial\chi}{\partial x}dx + \frac{\partial\chi}{\partial y}dy + \frac{\partial\chi}{\partial z}dz = d\chi$$

$$\therefore \chi = \int_{\mathbf{r}_0}^{\mathbf{r}} A(\mathbf{r}')\cdot d\mathbf{r}' \tag{12.41}$$

where $\mathbf{r}_0$ is some arbitrary reference point, and the line integral is along an arbitrary curve connecting $\mathbf{r}_0$ and $\mathbf{r}$.

$$\therefore \psi' = \exp\left(-\frac{ie}{\hbar}\int_{r_0}^{\mathbf{r}} \mathbf{A}\cdot d\mathbf{r}'\right)\psi(\mathbf{r},t) \tag{12.42}$$

So, a non-zero vector potential **A** produces a phase factor for the wave function. Since the wave function is arbitrary up to a phase factor, in general, such phase factors need not be taken seriously. However, Aharonov–Bohm considered a physical situation in which this phase can lead to observable consequences.

Consider two beams of electrons: one beam travelling path 1 and another beam travelling path 2, both beams being well outside the solenoid. Since two beams travelling through different paths are combined, an interference pattern will be formed independent of the fact whether solenoid is present or not or whether a current flows through the solenoid or not.

Let us recall that $\mathbf{B} = 0$ outside the solenoid whether a current flows or not through the coil. Only inside the coil, **B** can be zero or non-zero, depending on whether a current flows or not.

The non-zero potential **A** at **r** (see Fig. 12.3(ii)) along these two paths introduces phases for the two beams as can be seen in the following. The phases ϕ_1 and ϕ_2 for the two paths are given by

$$\phi_1 = -\frac{e}{\hbar}\underset{\text{path 1}}{\int_{\mathbf{r}_0}^{\mathbf{r}}} \mathbf{A}(\mathbf{r}')\cdot d\mathbf{r}' \quad \text{for path 1}$$

$$\phi_2 = -\frac{e}{\hbar}\underset{\text{path 2}}{\int_{\mathbf{r}_0}^{\mathbf{r}}} \mathbf{A}(\mathbf{r}')\cdot d\mathbf{r}' \quad \text{for path 2}$$

For electrons along path 1, $\psi_1' = e^{i\phi_1}\psi_1$ and for electrons along path 2, $\psi_2' = e^{i\phi_2}\psi_2$. At **r** on the screen, they superimpose.

$$\begin{aligned}\therefore \psi' &= e^{i\phi_1}\psi_1 + e^{i\phi_2}\psi_2\\ &= e^{i\phi_1}[\psi_1 + e^{i(\phi_2-\phi_1)}\psi_2]\end{aligned}$$

$$\phi_2 - \phi_1 = -\frac{e}{\hbar}\left(-\underset{\text{path 1}}{\int_{\mathbf{r}_0}^{\mathbf{r}}} \mathbf{A}(\mathbf{r}')\cdot d\mathbf{r}' + \underset{\text{path 2}}{\int_{\mathbf{r}_0}^{\mathbf{r}}} \mathbf{A}(\mathbf{r}')\cdot d\mathbf{r}\right)$$

$$= -\frac{e}{\hbar}\left(\underset{\text{path 1}}{\int_{\mathbf{r}}^{\mathbf{r}_0}} \mathbf{A}(\mathbf{r}')\cdot d\mathbf{r}' + \underset{\text{path 2}}{\int_{\mathbf{r}_0}^{\mathbf{r}}} \mathbf{A}(\mathbf{r}')\cdot d\mathbf{r}\right)$$

$$= -\frac{e}{\hbar}\oint_C \mathbf{A}(\mathbf{r})\cdot d\mathbf{r} = -\frac{e}{\hbar}\Phi \tag{12.43}$$

where Φ is the flux of magnetic field enclosed by the contour C.

$$\therefore \psi' = e^{i\phi_1}(\psi_1 + e^{-ie\Phi/\hbar}\psi_2) \tag{12.44}$$

The experiment is to start with zero **B** inside the solenoid. Since the electrons travel through two different paths, they will produce an interference pattern. The wave function on the screen is

$$\psi = \psi_1 + \psi_2$$

This superposition produces a set of interference fringes on the screen. Now make $\mathbf{B} \neq 0$ inside the solenoid. The wave function on the screen now becomes

$$\psi' = \psi_1' + \psi_2' = e^{i\phi_1}\left(\psi_1 + e^{\frac{-ie}{\hbar}\Phi}\psi_2\right) \tag{12.45}$$

The presence of the field inside solenoid produces a phase shift $e/\hbar\,\Phi$ between the beams, and so the interference fringe pattern on the screen will shift as a whole. The resulting shift for the intensity distribution (graphical representation of the Equation (12.45)) in the case of double slit experiment is shown in Fig. 12.4.

The shift is known as Aharonov–Bohm effect. This shift is non-zero for any closed path provided non-zero magnetic flux is enclosed. So it is a topological phase.

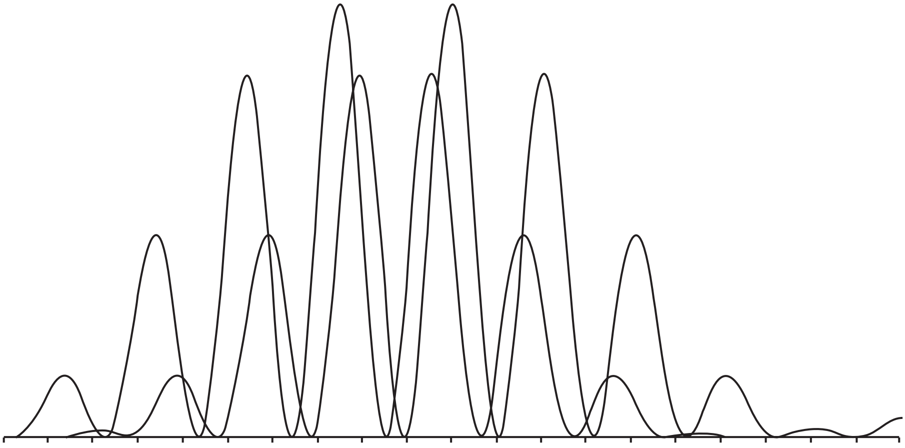


Fig. 12.4 Shift in interference fringe due to B in the solenoid

Experimentally, such fringe shift has been observed. It is not easy to verify Aharonov–Bohm effect since it is difficult to produce magnetic field only in one region without a leakage of magnetic field in the neighbouring region. In fact, this is cited as the reason for the skepticism towards the verification of Aharonov–Bohm effect. However, the experimental verification of Aharonov–Bohm effect is generally accepted today. (See R. G. Chambers (1960), A. Tonomura (1986)).

What do we conclude from such a result for the nature of electromagnetic field? Or how do you interpret Aharonov–Bohm effect?

If we analyze the Aharonov–Bohm effect in terms of the magnetic field, the presence of field in a region (within the solenoid) affects the behaviour of electrons in distant places (though there is no field in these regions). This is in contrast with our understanding on the physical nature of electromagnetic fields in classical physics.

There is a different interpretation of this effect. In classical physics, the magnetic field **B** is considered as real and observable quantity. The vector potential **A** is considered as a mathematical construct devised for the purpose of calculations. In quantum mechanics, it is the vector potential and not the magnetic field which enters into the calculations directly. In Aharonov–Bohm effect, it is the presence of **A** (though its value is determined by the value of **B** inside the solenoid) which determines the phase of the electron's wave function. So Aharonov–Bohm effect is interpreted to imply that the vector potential $\mathbf{A}$ is more fundamental than the magnetic field $\mathbf{B}$. Of course, the vector potential $\mathbf{A}$ is still arbitrary due to gauge invariance. It is the gauge invariant quantities which are important with respect to observation. So, in this interpretation, Aharonov–Bohm effect establishes the physical significance of the vector potential $\mathbf{A}$ within the constraints imposed by gauge invariance. This interpretation also contradicts the view in classical physics on the physical nature of vector potential.

The reader can consult an advanced book for the interpretations like topological phase. Aharonov–Bohm effect is one of the most extensively discussed topics in quantum mechanics.

12.6 TWO-STATE PROBLEM

Let us turn our attention to another set of phenomena involving systems with two-basis states. This class of phenomena is known as two-state problem. Let us consider a system for which the Hamiltonian is given by

$$h = h(\mathbf{r}) + h_{\text{spin}}$$

For such Hamiltonian, the spin dependent behaviour of the system is decoupled from the coordinate dependent part. So we can safely ignore the space coordinate part and consider the spin dependent part alone. The vector space of quantum mechanical states now becomes two-dimensional Hilbert space.

However, the two-state problem is not unique to spin systems alone. There is a large number of phenomena which can be described using the two-basis states. For instance, ammonia molecule's configuration (state of nitrogen) and neutrino oscillation have nothing to do with either spin or interaction with magnetic field, but their state vector spaces are also, under some circumstances, two-dimensional Hilbert space.

12.7 GENERAL ANALYSIS OF TWO-STATE PROBLEMS

12.7.1 Time-independent Hamiltonian

Let us consider a Hamiltonian whose eigenstates are $|E_1\rangle$ and $|E_2\rangle$.

$$\hat{H}_0 |E_1\rangle = E_1 |E_1\rangle; \quad \hat{H}_0 |E_2\rangle = E_2 |E_2\rangle \tag{12.46}$$

Let the initial state of the system $|\psi(0)\rangle$, be given by

$$|\psi(0)\rangle = C_1|E_1\rangle + C_2|E_2\rangle \tag{12.47}$$

This state evolves to $|\psi(t)\rangle$, which is determined by the Schrödinger equation

$$i\hbar\frac{\partial}{\partial t}|\psi(t)\rangle = \hat{H}_0\,|\psi(t)\rangle$$

$$|\psi(t)\rangle = e^{-i\frac{\hat{H}_0 t}{\hbar}}\,|\psi(0)\rangle \tag{12.48}$$

It is simple to see that

$$e^{-i\hat{H}_0 t/\hbar}\,|E_1\rangle = e^{-iE_1 t/\hbar}\,|E_1\rangle \tag{12.49}$$

and

$$e^{-i\hat{H}_0 t/\hbar}\,|E_2\rangle = e^{-iE_2 t/\hbar}\,|E_2\rangle \tag{12.50}$$

$$\begin{aligned}\therefore |\psi(t)\rangle &= e^{-iH_0 t/\hbar}[C_1|E_1\rangle + C_2|E_2\rangle] \\ &= C_1 e^{-iE_1 t/\hbar}|E_1\rangle + C_2 e^{-iE_2 t/\hbar}|E_2\rangle\end{aligned} \tag{12.51}$$

In terms of matrix representation, the Hamiltonian $\hat{H}_0$ becomes a 2 × 2 matrix H_0. Let us write the eigenvectors corresponding to $|\mathrm{E}_1\rangle$ and $|\mathrm{E}_2\rangle$ as $\psi(\uparrow)$ and $\psi(\downarrow)$, or $\psi(+)$ and $\psi(-)$.

$$H_0\psi(+) = E_1\psi(+); \quad H_0\psi(-) = E_2\psi(-) \tag{12.52}$$

$$\therefore \psi(0) = C_1\psi(+) + C_2\psi(-) \tag{12.53}$$

$$\psi(t) = C_1 e^{-iE_1 t/\hbar}\psi(+) + C_2 e^{-iE_2 t/\hbar}\psi(-) \tag{12.54}$$

Alternatively, the time evolution of the state $\psi(t)$ can be determined in an easier way as follows:

Note that any 2 × 2 Hamiltonian matrix H_0 can be written as a linear combination I, σ_x, σ_y and σ_z.

i.e.,

$$H_0 = aI + b\,\sigma_x + c\sigma_y + d\sigma_z. = aI + \boldsymbol{\sigma}\cdot\mathbf{n}$$

Therefore, $\psi(t)$ is given by

$$\psi(t) = e^{-(aI+\boldsymbol{\sigma}\cdot\mathbf{n})t/\hbar}\psi(0) = e^{-at/\hbar}e^{-\boldsymbol{\sigma}\cdot\mathbf{n}t/\hbar}\psi(0) \tag{12.55}$$

These results for spin angular momentum in magnetic field hold good for other physical systems ((nitrogen atom) in ammonia molecule, neutrino oscillation, etc.) for which the Hilbert space of states is a two-dimensional linear vector space.

12.7.2 Time-dependent Hamiltonian

Let us next consider time-dependent Hamiltonian. Let us write

$$\hat{H} = \hat{H}_0 + \hat{H}' \tag{12.56}$$

where $\hat{H}_0$ is time-independent and $\hat{H}'$ is time-dependent part. Let us choose again the eigenstates of $\hat{H}_0$ as the basis states. Then we can write any arbitrary state vector $|\psi(t)\rangle$ as

$$|\psi(t)\rangle = C_1(t)e^{-iE_1t/\hbar}|E_1\rangle + C_2(t)e^{-iE_2t/\hbar}|E_2\rangle \tag{12.57}$$

$|\psi(t)\rangle$ is determined by

$$i\hbar\frac{\partial}{\partial t}|\psi(t)\rangle = (\hat{H}_0 + \hat{H}')\,|\psi(t)\rangle \tag{12.58}$$

LHS:

$$i\hbar\frac{\partial}{\partial t}|\psi(t)\rangle = \sum_n i\hbar\dot{C}_n\, e^{-iE_nt/\hbar}|E_n\rangle + \sum_n C_n e^{-iE_nt/\hbar}E_n|E_n\rangle \tag{12.59}$$

RHS:

$$\begin{aligned}\hat{H}\,|\psi(t)\rangle &= \hat{H}_0|\psi(t)\rangle + \hat{H}'|\psi(t)\rangle \\ &= \sum_n C_n e^{-iE_nt/\hbar}E_n|E_n\rangle + \sum_n C_n e^{-iE_nt/\hbar}H'|E_n\rangle\end{aligned} \tag{12.60}$$

Equating these two equations, we get

$$\sum i\hbar\dot{C}_n(t)e^{-iE_nt/\hbar}\,|E_n\rangle = \sum_n C_n e^{-iE_nt/\hbar}\,|E_n\rangle$$

Consider the equation for C_1.

$$\begin{aligned}&i\hbar\dot{C}_1\, e^{-iE_1t/\hbar}\,|E_1\rangle + i\hbar\dot{C}_2\, e^{-iE_2t/\hbar}\,|E_2\rangle \\ &\qquad = C_1e^{-iE_1t/\hbar}\,|E_1\rangle + C_2e^{-iE_2t/\hbar}\,|E_2\rangle\end{aligned} \tag{12.61}$$

Taking scalar product with $\langle E_1|$, we get

$$\begin{aligned}i\hbar\dot{C}_1\, e^{-iE_1t/\hbar} &= C_1e^{-iE_1t/\hbar}H'_{11} + C_2e^{-iE_2t/\hbar}H'_{12} \\ \dot{C}_1 &= \frac{1}{i\hbar}[C_1H'_{11} + C_2e^{i(E_1-E_2)t/\hbar}H'_{12}]\end{aligned} \tag{12.62}$$

Or

$$\dot{C}_1 = \frac{1}{t\hbar}[C_1H'_{11} + C_2e^{i\omega_{12}t}H'_{12}]$$

Similarly, $\dot{C}_2$ is found to be

$$\dot{C}_2 = \frac{1}{i\hbar}[C_1e^{-i\omega_{12}t}H'_{21} + C_2H'_{22}] \tag{12.63}$$

Solution to these equations determines $|\psi(t)\rangle$.

The above analysis suggests the following steps to analyze the two-state problems which involve time-dependent Hamiltonian.

1. Determine the energy eigenvalues E_1 and E_2 for the given Hamiltonian H_0 and the corresponding eigenstates $\psi(+)$ and $\psi(-)$.
2. The wave function $\psi(t)$ is given by $\psi(t)=C_1(t)e^{-iE_1 \cdot t/\hbar}\psi(+)+C_2(t)e^{-iE_2 t/\hbar}\psi(-)$.
3. Choose an initial state $\psi(0)$ and write it as $\psi(0)=C_1(0)e^{-iE_1 \cdot t/\hbar}\psi(+)+C_2(0)e^{-iE_2 t/\hbar}\psi(-)$. Determine the coefficients $C_1(0)$ and $C_2(0)$.
4. Determine $C_1(t)$ and $C_2(t)$ using the Equations (12.62) and (12.63).

Example 12.2 Prove that any 2×2 Hermitian matrix A can be written as $A = aI + b\sigma_x + c\sigma_y + d\sigma_z$.

Solution: The arbitrary Hermitian matrix A is written as

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{12}^{*} & a_{22} \end{bmatrix}$$

Since a_{12} is a complex number, it can be written as $a_{12} = b - ic$. a_{11} and a_{22} can be written as $a_{11} = a + d$ and $a_{22} = a - d$.

$$\therefore A = \begin{bmatrix} a+d & b-ic \\ b+ic & a-d \end{bmatrix}$$
$$= aI + b\sigma_x + c\sigma_y + d\sigma_z$$

Example 12.3 Determine the eigenvalue of A given by

$$A = \begin{bmatrix} a+d & b-ic \\ b+ic & a-d \end{bmatrix}$$

Solution: The eigenvalue equation is

$$\begin{vmatrix} a+d-\lambda & b-ic \\ b+ic & a-d-\lambda \end{vmatrix} = 0$$
$$(a+d-\lambda)(a-d-\lambda)-(b^2+c^2)=0$$
$$a^2-(b^2+c^2+d^2)-2a\lambda+\lambda^2=0$$
$$\lambda = a+(b^2+c^2+d^2)^{1/2},\ a-(b^2+c^2+d^2)^{1/2}$$

Example 12.4 Evaluate $e^{ir\boldsymbol{\sigma}\cdot\mathbf{n}t}$.

Solution:

$$e^{ir\boldsymbol{\sigma}\cdot\mathbf{n}t} = 1 + ir\boldsymbol{\sigma}\cdot\mathbf{n}t - \frac{r^2t^2(\boldsymbol{\sigma}\cdot\mathbf{n})^2}{2!} - \frac{ir^3t^3(\boldsymbol{\sigma}\cdot\mathbf{n})^3}{3!} + \ldots\ldots$$

Since $(\boldsymbol{\sigma}\cdot\mathbf{n})^2 = 1$ we have

$$e^{ir\boldsymbol{\sigma}\cdot\mathbf{n}t} = \left(1 - \frac{r^2t^2}{2!} + \ldots\right) + i\boldsymbol{\sigma}\cdot\mathbf{n}\left(rt - \frac{r^3t^3}{3!} + \ldots\right)$$
$$= \cos rt + i\boldsymbol{\sigma}\cdot\mathbf{n}\sin rt$$

12.8 MAGNETIC MOMENT AND SPIN

We have already seen that in classical electromagnetism, the magnetic moment due to orbital angular momentum is

$$\boldsymbol{\mu} = \frac{q}{2m}\mathbf{L}$$

We would like to extend this idea to spin angular momentum also. Things are not so simple. In the case of electron, we would expect the corresponding magnetic moment to $\boldsymbol{\mu} = (-e/2m)\mathbf{S}$. However, experimentally, it is found to be different from this expression. The magnetic moment of an electron due to spin angular momentum is generally written as

$$\boldsymbol{\mu} = -g_s \frac{e}{2m}\mathbf{S} \tag{12.64}$$

where g_s is known as gyromagnetic ratio or simply g factor. We can take $g_s = 2$, though the experimental value g_s is little more than 2. Theoretically, one can calculate g_s from relativistic quantum field theory, which agrees well with the experimental value. For our purpose, we take $g_s = 2$.

For other particles like proton and neutron, the relationship between spin angular momentum and magnetic moment is more complex. As a general case, we write $\boldsymbol{\mu}$ as

$$\boldsymbol{\mu} = -\gamma\mathbf{S} \tag{12.65}$$

12.9 PRECESSION OF SPIN ANGULAR MOMENTUM

In classical physics, if one places a dipole (moment) in a magnetic field, there is a torque on it. If the magnetic moment is $\boldsymbol{\mu} = -\gamma\mathbf{S}$, the torque makes the spin angular momentum to precess around $\mathbf{B}$ with Larmor frequency $\omega = \gamma B$.

In quantum mechanics, the corresponding result holds good for mean value of $\mathbf{S}$. We will demonstrate here that $\langle\mathbf{S}\rangle$ processes around $\mathbf{B}$ with Larmor frequency $\omega = \gamma B$.

Consider an electron in a magnetic field $B = B\mathbf{z}$. The Hamiltonian H_0 is given by

$$\begin{aligned} H_0 &= -\boldsymbol{\mu}\cdot\mathbf{B} = g_S\frac{e}{2m}\mathbf{S}\cdot\mathbf{B} = \frac{g_S e}{2m}BS_z \\ &= \frac{\hbar\omega_0}{2}\sigma_z \end{aligned} \tag{12.66}$$

where $$\omega_0 = g_S eB/2m \tag{12.67}$$

The first step is to determine the energy eigenvalue and eigenvectors of $H_0 = (\hbar\omega_0/2)\sigma_z$. The eigenvalues and the corresponding eigenvectors are given by

$$E_1 = \frac{\hbar\omega_0}{2}; \qquad \psi(\uparrow) = \begin{pmatrix}1\\0\end{pmatrix} \tag{12.68}$$

$$E_2 = -\frac{\hbar\omega_0}{2}; \qquad \psi(\downarrow) = \begin{pmatrix}0\\1\end{pmatrix} \tag{12.69}$$

We are interested in determining the time evolution of an arbitrary state vector. So let us choose the eigenvector of $\mathbf{S}\cdot\mathbf{n}$ (see Section (10.6)) as the state vector $\psi(0)$ at $t = 0$, where

$$\mathbf{n} = (n_x, n_y, n_z) = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$$

$$\therefore \psi(0) = \begin{bmatrix} \cos\dfrac{\theta}{2} \\ \sin\dfrac{\theta}{2}e^{i\phi} \end{bmatrix}$$

Now, let us express $\psi(0)$ as a linear combination of $\begin{pmatrix}1\\0\end{pmatrix}$ and $\begin{pmatrix}0\\1\end{pmatrix}$.

$$\therefore \psi(0) = \cos\frac{\theta}{2}\begin{pmatrix}1\\0\end{pmatrix} + \sin\frac{\theta}{2}e^{i\phi}\begin{pmatrix}0\\1\end{pmatrix} \tag{12.70}$$

The time evolution of $\psi(t)$ is given by

$$\begin{aligned}\psi(t) &= e^{-iE_1t/\hbar}\cos\frac{\theta}{2}\begin{pmatrix}1\\0\end{pmatrix} + e^{-iE_2t/\hbar}\sin\frac{\theta}{2}e^{i\phi}\begin{pmatrix}0\\1\end{pmatrix} \\ &= e^{-iE_1t}\left[\cos\frac{\theta}{2}\begin{pmatrix}1\\0\end{pmatrix} + e^{i(E_1-E_2)t/\hbar}\sin\frac{\theta}{2}e^{i\phi}\begin{pmatrix}0\\1\end{pmatrix}\right] \\ &= e^{-iE_1t}\begin{bmatrix}\cos\dfrac{\theta}{2} \\ \sin\dfrac{\theta}{2}e^{i(\phi+\omega_0t)}\end{bmatrix}\end{aligned} \tag{12.71}$$

The expectation value of S_x is given by

$$\begin{aligned}\langle S_x\rangle &= \psi^\dagger(t)S_x\psi(t) = \frac{\hbar}{2}\psi^\dagger(t)\sigma_x\psi(t) \\ &= \frac{\hbar}{2}\begin{bmatrix}\cos\dfrac{\theta}{2} & \sin\dfrac{\theta}{2}e^{-i(\phi+\omega_0t)}\end{bmatrix}\begin{bmatrix}0 & 1\\1 & 0\end{bmatrix}\begin{bmatrix}\cos\dfrac{\theta}{2} \\ \sin\dfrac{\theta}{2}e^{i(\theta+\omega_0t)}\end{bmatrix} \\ &= \frac{\hbar}{2}\sin\theta\cos(\phi+\omega_0t)\end{aligned} \tag{12.72}$$

Similarly, $\langle S_y\rangle$ and $\langle S_z\rangle$ are found to be

$$\begin{aligned}\langle S_y\rangle &= \psi^+(t)S_y\psi(t) = \frac{\hbar}{2}\psi^+(t)\sigma_y\psi(t) \\ &= \frac{\hbar}{2}\sin\theta\sin(\phi+\omega_0t)\end{aligned} \tag{12.73}$$

and

$$\begin{aligned}\langle S_z\rangle &= \psi^\dagger(t)S_z\psi(t) = \frac{\hbar}{2}\psi^\dagger(t)\sigma_z\psi(t)\\ &= \frac{\hbar}{2}\cos\theta\end{aligned} \tag{12.74}$$

$$\langle \mathbf{S}\rangle = \frac{\hbar}{2}\cos\theta\cos(\phi_0+\omega t)\,\hat{\mathbf{x}} + \frac{\hbar}{2}\sin\theta\sin(\phi_0+\omega t)\,\hat{\mathbf{y}} + \frac{\hbar}{2}\cos\theta\,\hat{\mathbf{z}} \tag{12.75}$$

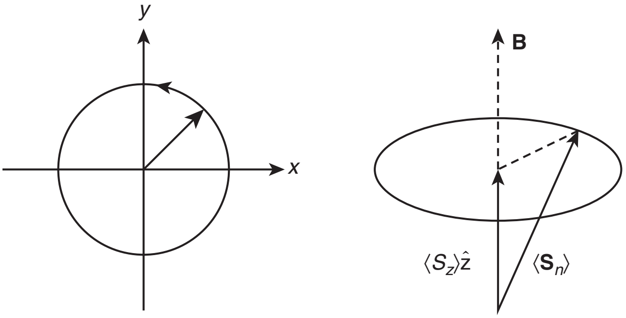


Fig. 12.5 Larmor precession

The average value $\langle S_x\rangle\hat{\mathbf{x}} + \langle S_y\rangle\hat{\mathbf{y}}$ traces a circle in $x-y$ plane and so $\langle\mathbf{S}\rangle$ executes a precessional motion around **B**.

Example 12.5 An electron in a state of spin-up along ***Z*** axis is subjected to a magnetic field $\mathbf{B} = B_0\hat{\mathbf{z}} + B_1\hat{\mathbf{x}}$. How does the state vector of the electron evolve in time *t* in the presence of the magnetic field? What is the probability of the electron to flip from spin-up state to spin-down state?

Solution: Let us write the magnetic field **B** as

$$\mathbf{B} = B\cos\theta\ \hat{\mathbf{z}} + B\sin\theta\ \hat{\mathbf{x}} \tag{12.76}$$

The Hamiltonian H_0 is given by

$$H_0 = g_S\frac{e}{2m}\mathbf{S}\cdot\mathbf{B} \tag{12.77}$$

We can write **B** as $\mathbf{B} = B\mathbf{n}$ where

$$\mathbf{n} = (n_x, n_y, n_z) = (\sin\theta, 0, \cos\theta)$$

The Hamiltonian H_0 becomes

$$H_0 = g_S\frac{eB}{2m}\mathbf{S}.\mathbf{n} = \frac{\hbar\omega}{2}\boldsymbol{\sigma}\bullet\mathbf{n} = \frac{\hbar\omega}{2}\begin{bmatrix}\cos\theta & \sin\theta\\ \sin\theta & -\cos\theta\end{bmatrix} \tag{12.78}$$

where

$$\omega = g_S\frac{e}{2m}B$$

From Schrödinger equation, we have

$$\psi(t) = e^{-iH_0 t/\hbar}\psi(0) = \exp\left(-i\frac{\omega t}{2}\boldsymbol{\sigma}\cdot\mathbf{n}\right)\psi(0)$$

We know that

$$\exp\left(-i\frac{\omega t}{2}\boldsymbol{\sigma}\cdot\mathbf{n}\right) = \cos\left(\frac{\omega t}{2}\right)I - i\sin\left(\frac{\omega t}{2}\right)\boldsymbol{\sigma}\cdot\mathbf{n} \tag{12.79}$$

In our problem, the electron begins in a state with spin-up along *Z* axis. Therefore, $\psi(0)$ is given by

$$\psi(0) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

$$\psi(t) = \left\{\begin{bmatrix} \cos(\omega t/2) & \\ & \cos(\omega t/2) \end{bmatrix} - i\sin\left(\frac{\omega t}{2}\right)\begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{bmatrix}\right\}\begin{pmatrix} 1 \\ 0 \end{pmatrix} \tag{12.80}$$

$$= \begin{pmatrix} \cos\left(\frac{\omega t}{2}\right) - i\sin\left(\frac{\omega t}{2}\right)\cos\theta \\ -i\sin\left(\frac{\omega t}{2}\right)\sin\theta \end{pmatrix} \tag{12.81}$$

The probability amplitude to get spin-down along Z axis is given by $\beta^\dagger\psi$, where $\beta = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$.

$$\beta^\dagger\psi(t) = \begin{pmatrix} 0 & 1 \end{pmatrix}\begin{pmatrix} \cos\left(\frac{\omega t}{2}\right) - i\sin\left(\frac{\omega t}{2}\right)\cos\theta \\ -i\sin\left(\frac{\omega t}{2}\right)\sin\theta \end{pmatrix} = -i\sin\left(\frac{\omega t}{2}\right)\sin\theta \tag{12.82}$$

Probability of getting spin-down is given by

$$P(\downarrow;t) = |\beta^\dagger\psi(t)|^2 \tag{12.83}$$

Since we start with a system with spin-up at $t = 0$, $P(\downarrow;\mathrm{t})$ represents the probability for the system to flip from spin-up to spin-down, and so we can denote it as $P(\uparrow\rightarrow\downarrow)$.

$$\begin{aligned} P(\uparrow\rightarrow\downarrow) = P(\downarrow;t) &= |\beta^\dagger\psi(t)|^2 \\ &= \sin^2\theta\sin^2\omega t \\ &= \sin^2\theta\sin^2\frac{(E_1 - E_2)}{\hbar}t \end{aligned} \tag{12.84}$$

The probability for the flip of spin-up (along *Z* axis) to spin-down depends on the difference in energy levels of the two states.

Let us write $\sin\theta$ as

$$\sin\theta = \frac{B_1}{B} = \frac{B_1}{\sqrt{B_0^2 + B_1^2}} = \frac{\frac{g_S e}{2m} B_1}{\frac{g_S e}{2m}\sqrt{B_0^2 + B_1^2}} = \frac{\omega_1}{\sqrt{\omega_0^2 + \omega_1^2}}$$

$$\therefore P(\uparrow\rightarrow\downarrow) = \frac{\omega_1}{[\omega_0^2 + \omega_1^2]^{\frac{1}{2}}} \sin^2 \frac{\sqrt{\omega_0^2 + \omega_1^2}}{2} t$$

If $\theta = \pi/2,\ \omega_0 = 0,\ \omega = \omega_1$

$$\therefore P(\uparrow\rightarrow\downarrow) = \sin^2 \frac{\omega_1 t}{2} \tag{12.85}$$

$$P(\uparrow\rightarrow\uparrow) = \cos^2 \frac{\omega_1 t}{2} \tag{12.86}$$

This is an interesting result. At $t = \pi/\omega_1, 3\pi/\omega_1 \ldots$ $P(\downarrow) = 1$ and at $t = 0, 2\pi/\omega_1, 4\pi/\omega_1 \ldots$ $P(\uparrow) = 1$. Let us recall that if the probability for an event to occur is 1, then it is a certain or definite event. This means that the electron has spin-up (definitely) at the instants of time $t = 0, 2\pi/\omega_1, 4\pi/\omega_1 \ldots$ and spin-down (definitely) at the instants of time $t = \pi/\omega_1, 3\pi/\omega_1 \ldots$ This result is true for any magnetic field applied in the ***XY*** plane. So if an electron starts with *Z* spin-up, a magnetic field applied perpendicular to *Z* axis makes the *Z* spin to flip flop up and down periodically with a frequency π/ω_1 in time.

12.10 MAGNETIC RESONANCE

Now consider an electron in a magnetic field

$$\mathbf{B} = B_0\hat{\mathbf{z}} + B_1 \cos\omega t\hat{\mathbf{x}} + B_1 \sin\omega t\hat{\mathbf{y}} \tag{12.87}$$

The field $B_0\hat{\mathbf{z}}$ creates two states with energies $E_1 = \hbar\omega_0/2, E_2 = -\hbar\omega_0/2$ corresponding to spin-up and -down along *Z* axis. There is a resonance phenomenon in this case similar to the case of forced oscillation. Rabi used this phenomenon to develop a method called molecular method to determine the magnetic moment of the proton, neutron and nuclei. Let the particle start in spin-up state at $t = 0$. The field $B_1 \cos\omega t\ \hat{\mathbf{x}} + B \sin\omega t\ \hat{\mathbf{y}}$ causes the spin to flip between spin-up and -down state. Let us determine the probability for the spin to become spin-down, which becomes maximum at a frequency called resonance frequency $\omega = (E_1 - E_2)/\hbar = \omega_0 = (g_S(e/2m))B_0$.

The Hamiltonian of the system is

$$\begin{aligned} H &= g_S \frac{e}{2m} \mathbf{S}\cdot\mathbf{B} \\ &= g_S \frac{eB_0}{2m}\cdot\frac{\hbar}{2}\sigma_z + g_S \frac{eB_1}{2m}\cdot\frac{\hbar}{2}\sigma_x \cos\omega t + g_S \frac{eB_1}{2m}\cdot\frac{\hbar}{2}\sigma_y \cdot \sin\omega t \\ &= \frac{\hbar}{2}\omega_0\sigma_z + \frac{\hbar}{2}\omega_1\sigma_x \cos\omega t + \frac{\hbar}{2}\omega_1\sigma_y \sin\omega t \end{aligned} \tag{12.88}$$

where $$\omega_0 = g_S eB_0/2m \text{ and } \omega_1 = g_S eB_1/2m. \tag{12.89}$$

This Hamiltonian H can be written as

$$H = H_0 + H'$$

where $$H_0 = \frac{\hbar\omega_0}{2}\sigma_Z = \frac{\hbar\omega_0}{2}\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \tag{12.90}$$

and $$H' = \frac{\hbar\omega_1}{2}\sigma_x \cos\omega t + \frac{\hbar\omega_1}{2}\sigma_y \sin\omega t$$

$$= \frac{\hbar\omega_1}{2}\begin{bmatrix} 0 & e^{-i\omega t} \\ e^{i\omega t} & 0 \end{bmatrix} \tag{12.91}$$

Let $\psi(\uparrow)$ and $\psi(\downarrow)$ are the eigenstates of H_0 with eigenvalues $\hbar\omega_0/2$ and $-\hbar\omega_0/2$.

$$H_0\psi(\uparrow) = \frac{\hbar\omega_0}{2}\psi(\uparrow); \quad \psi(\uparrow) = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \tag{12.92}$$

$$H_0\psi(\downarrow) = -\frac{\hbar\omega_0}{2}\psi(\downarrow); \quad \psi(\downarrow) = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{12.93}$$

The state $\psi(t)$ of the system is given by the Schrödinger equation

$$i\hbar\frac{\partial\psi}{\partial t}(t) = (H_0 + H')\psi(t)$$

In Section 12.8, we have already seen that the solution for such Hamiltonian is given by

$$\psi(t) = C_1(t)e^{-iE_1t/\hbar}\psi(\uparrow) + C_2(t)e^{-iE_2t/\hbar}\psi(\downarrow) \tag{12.94}$$

We are interested in determining that the time evolution of the system where initial state is $\psi(\uparrow)$.

i.e., $$\psi(0) = \psi(\uparrow)$$

$$\therefore C_1(0) = 1 \text{ and } C_2(0) = 0 \tag{12.95}$$

$C_1(t)$ and $C_2(t)$ are determined by solving the Equations (12.62) and (12.63) in Section 12.7.

To solve the equations, we require to evaluate $H'_{11}, H'_{22}, H'_{12}$ and H'_{21}.

Making use of the Equations (12.91), (12.92) and (12.93),

$$H'_{11} = \psi(\uparrow)^\dagger H'\psi(\uparrow) = 0$$

$$H'_{22} = \psi(\downarrow)^\dagger H'\psi(\downarrow) = 0$$

$$H'_{12} = \psi(\uparrow)^{\dagger} H'\psi(\downarrow) = \frac{\hbar\omega_1}{2} e^{-i\omega t}$$

$$H'_{21} = \psi(\downarrow)^{\dagger} H'\psi(\uparrow) = \frac{\hbar\omega_1}{2} e^{i\omega t}$$

Substituting these matrix elements in (12.62) and (12.63), we get

$$\dot{C}_1 = -i\frac{\omega_1}{2} e^{i(\omega_{12}-\omega)t} C_2 = -i\frac{\omega_1}{2} e^{i(\omega_0-\omega)t} C_2 \qquad (12.96)$$

$$\dot{C}_2 = -i\frac{\omega_1}{2} e^{-i(\omega_{12}-\omega)t} C_1 = -i\frac{\omega_1}{2} e^{-i(\omega_0-\omega)t} C_1 \qquad (12.97)$$

Differentiating the Equation (12.97) with respect to time, we get

$$\ddot{C}_2 = -i\frac{\omega_1}{2}\left[\frac{d}{dt} e^{-i(\omega_0-\omega)t}\right] C_1 - i\frac{\omega_1}{2} e^{-i(\omega_0-\omega)t} \frac{dC_1}{dt}$$

Making use of the Equation (12.97) in first term and the Equation (12.96) in second term, we get

$$\ddot{C}_2 + i(\omega_0 - \omega)\dot{C}_2 + \omega_1^2 C_2 = 0 \qquad (12.98)$$

Similarly, for C_1, we get

$$\ddot{C}_1 - i(\omega_0 - \omega)\dot{C}_1 + \omega_1^2 C_1 = 0 \qquad (12.99)$$

Since C_1 and C_2 are related through the Equation (12.96), it is enough to solve either the Equation (12.99) or the Equation (12.98).

The Equation (12.98) is a second-order differential equation with constant coefficient. The auxiliary equation is

$$\alpha^2 + i(\omega_0 - \omega)\alpha + \frac{\omega_1^2}{4} = 0$$

Solving this equation, we get

$$C_2(t) = e^{-i(\omega_0-\omega)t/2}\left[A_1 e^{i\sqrt{(\omega_0-\omega)^2+\omega_1^2}\,t/2} + A_2 e^{-i\sqrt{(\omega_0-\omega)^2+\omega_1^2}\,t/2}\right]$$

The boundary condition $C_2(0) = 0$ implies $A_2 = -A_1$. Therefore,

$$C_2(t) = e^{-i(\omega_0-\omega)t/2} A_1\left[e^{i\sqrt{(\omega_0-\omega)^2+\omega_1^2}\,t/2} - e^{-i\sqrt{(\omega_0-\omega)^2+\omega_1^2}\,t/2}\right]$$

$$= 2iA_1 e^{-i(\omega_0-\omega)t/2} \sin\frac{[(\omega_0-\omega)^2+\omega_1^2]^{\frac{1}{2}} t}{2}$$

From (12.97), we have

$$C_1(t) = \frac{2i}{\omega_1} e^{i(\omega_0-\omega)t} \dot{C}_2$$
$$= \frac{-4A_1}{\omega_1} e^{i(\omega_0-\omega)t/2} \left[-i\frac{(\omega_0-\omega)}{2}\sin\left(\frac{\sqrt{(\omega_0-\omega)^2+\omega_1^2 t}}{2}\right) + \frac{\sqrt{(\omega_0-\omega)+\omega_1^2}}{2}\cos\frac{\sqrt{(\omega_0-\omega)^2+\omega_1^2 t}}{2} \right]$$

Boundary condition C_1 (0) = 1 implies

$$A_1 = -\frac{\omega_1}{2[(\omega_0-\omega)^2+\omega_1^2]^{1/2}}$$

$$C_1(t) = \frac{1}{[(\omega_0-\omega)^2+\omega_1^2]^{1/2}} e^{i(\omega_0-\omega)t/2} \left\{ -i(\omega_0-\omega)\sin\frac{\sqrt{(\omega_0-\omega)^2+\omega_1^2}\, t}{2} + \sqrt{(\omega_0-\omega)^2+\omega_1^2}\cos\frac{\sqrt{(\omega_0-\omega)^2+\omega_1^2}\, t}{2} \right\} \tag{12.100}$$

$$C_2(t) = -\frac{i\omega_1}{\sqrt{(\omega_0-\omega)^2+\omega_1^2}} e^{-i(\omega_0-\omega)t/2} \sin\frac{\sqrt{(\omega_0-\omega)^2+\omega_1^2}\, t}{2} \tag{12.101}$$

The probability of finding the particle in states $\psi(\uparrow)$ and $\psi(\downarrow)$ is given by $|C_1(t)|^2$ and $|C_2(t)|^2$.

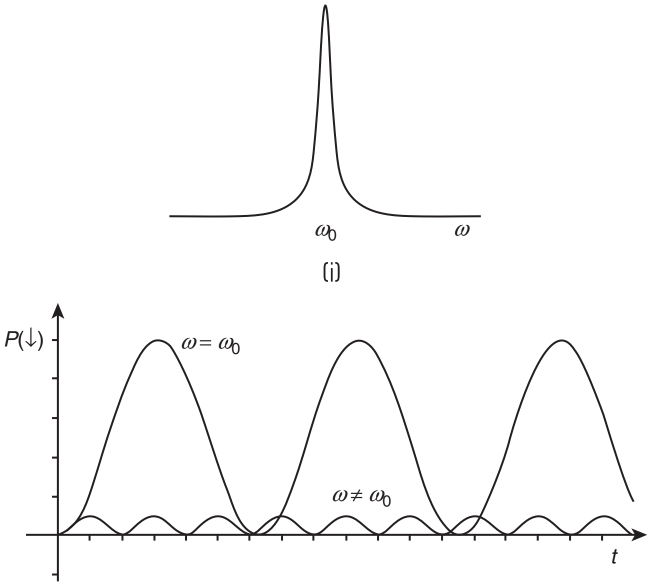
$$P(\downarrow) = |C_2(t)|^2 = \frac{\omega_1^2}{[(\omega_0-\omega)^2+\omega_1^2]}\sin^2\frac{\sqrt{(\omega_0-\omega)^2+\omega_1^2}\, t}{2} \tag{12.102}$$

and

$$P(\uparrow) = 1 - P(\downarrow)$$
$$= 1 - \frac{\omega_1^2}{[(\omega_0-\omega)^2+\omega_1^2]}\sin^2\frac{\sqrt{(\omega_0-\omega)^2+\omega_1^2}\, t}{2} \tag{12.103}$$

Remember that we have obtained this result by assuming $C_1(0) = 1$ and C_2 (0) = 0. i.e., we start with the system in state $\psi(\uparrow)$ at $t = 0$. Therefore, $|C_2(t)|^2$ can be interpreted as the transition probability to flip from spin-up to spin-down.

$$P(\uparrow\rightarrow\downarrow) = |C_2(t)|^2 = \frac{\omega_1^2}{(\omega_0-\omega)^2+\omega_1^2}\sin^2\frac{\sqrt{(\omega_0-\omega)^2+\omega_1^2}\, t}{2} \tag{12.104}$$



(i)

(ii)

Fig. 12.6 (i) Amplitude of the probability $P(\downarrow)$ (ii) $P(\downarrow)$ as a function of time

This expression reflects the behaviour of a resonance phenomenon with respect to variation in ω as can be seen in Fig. 12.6(i). The resonance occurs at $\omega = \omega_0$. The probability to flip from spin-up and spin-down varies periodically in time. The amplitude of periodic variation reaches a maximum at $\omega = \omega_{12} = \omega_0$, the resonance frequency.

i.e.,
$$\hbar\omega_{12} = \hbar\omega_0 = (E_1 - E_2)$$

At resonance frequency,

$$P(\downarrow) = \sin^2\left(\frac{\omega_1 t}{2}\right) \tag{12.105}$$

$$P(\uparrow) = \cos^2\left(\frac{\omega_1 t}{2}\right) \tag{12.106}$$

where
$$\omega_1 = g_s \frac{e}{2m} B_1.$$

If we choose the field strength and the dimensions of the apparatus such that $\omega_1 t = \pi$, then we can have maximum probability for the spin flip. Rabi (1939) employed three magnets and an radio frequency oscillator in their experiment and determined the nuclear magnetic moments for lithium isotopes and fluorine.

Note that this resonance phenomenon is true for any particle with magnetic moment $\boldsymbol{\mu} = \gamma \mathbf{S}$. For instance, in nuclear magnetic resonance, it is the proton which flips between spin-up and spin-down at resonance frequency.

12.11 AMMONIA MOLECULE

Ammonia is a molecule consisting of three hydrogen atoms and one nitrogen atom. The geometrical configuration of the molecule is a tetrahedron with nitrogen at one vertex and three hydrogen atoms forming the base triangle. But what is the position of nitrogen with respect to the base triangle? Is it to the left or right of the base triangle as shown in Fig. 12.7? Each configuration is a mirror reflection of the other. There is no reason to choose a particular configuration. Once the position of nitrogen is fixed, a particular configuration has been chosen. This configuration cannot change on its own to the second configuration. But, in quantum mechanics, things are different. Describing the entire molecule in terms of Schrödinger equation for all the electrons and nuclei is very complex. Leaving aside all the complexities, if one is interested only in nitrogen with respect to hydrogen base, one can argue for a simple model of a particle in a simple double potential well.

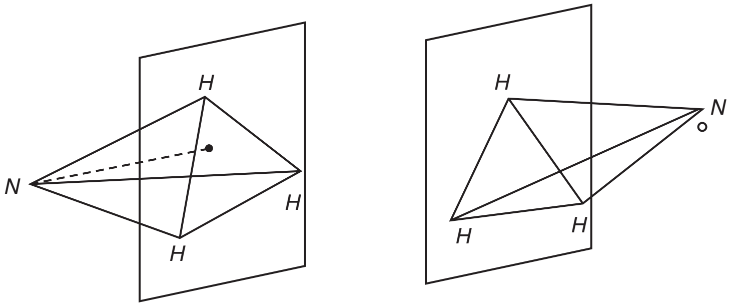


Fig. 12.7 Ammonia molecule

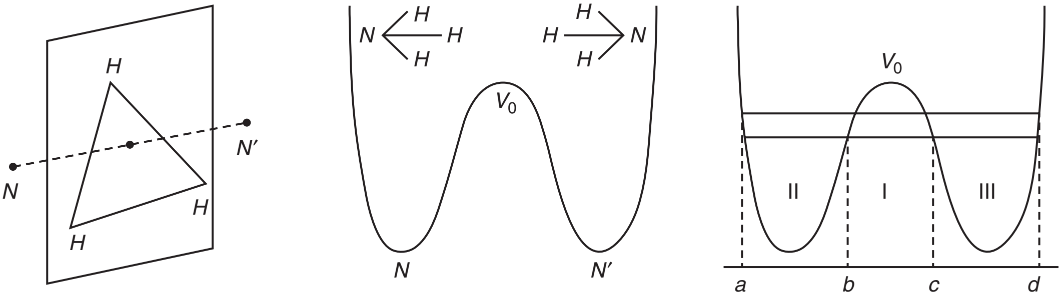


Fig. 12.8 Potential V and the ground-state energy levels for nitrogen

Let us first consider the ground state of the nitrogen from the point of view of classical physics. If N and N' are the equilibrium positions they should represent the minima of a potential. This potential should have a barrier of height V_0, preventing the change of one configuration to another. This can be represented by a double-well potential as shown in Fig. 12.8. There are three regions. For energy $E < V_0$, the regions II and III are classically allowed regions, and region I is a classically forbidden region. So, once a particle is in region II, it cannot move to region III.

The quantum mechanical description is based on solving Schrödinger equation for the particle in the double-well potential V. We can get the ground-state energy eigenvalues and eigenfunctions by numerical computation. Here, we quote only the end results (J.L. Basdevant and J. Dalibard (2002)). The ground state has two closely separated energy levels as shown in Fig. 12.9. The corresponding states ψ_s and ψ_a are symmetric and antis-symmetric wave function as shown in Fig. 12.7.

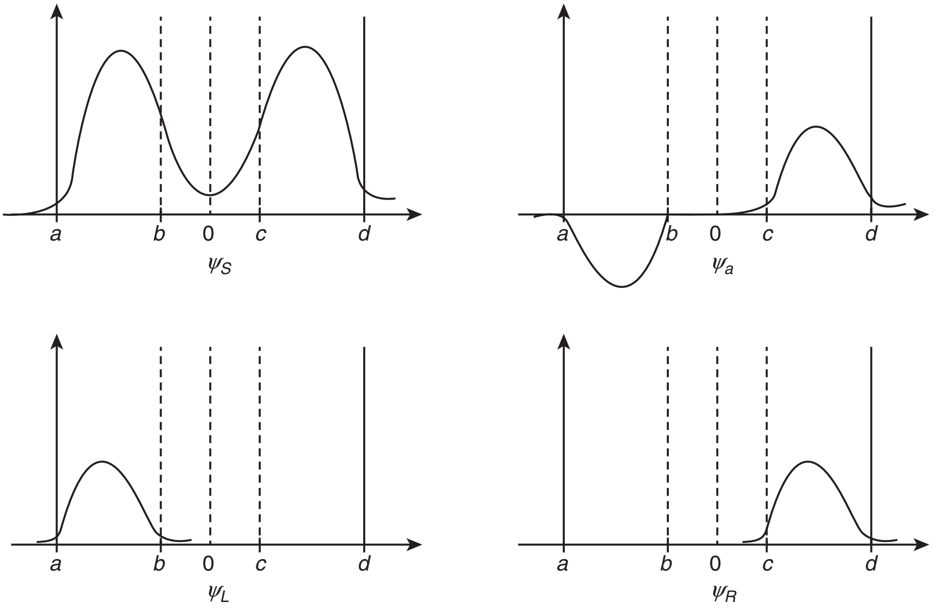


Fig. 12.9 Wave functions ψ_s, ψ_a, ψ_L and ψ_R

Note that both ψ_s and ψ_a are the non-zero region I which is a classically forbidden region. The presence of non-zero wave function in this region corresponds to tunnelling phenomenon.

The wave functions ψ_s and ψ_a are eigenstates with $E_s = E_0 - A$ and $E_a = E_0 + A$. We can form a superposition of these two states in two different ways as follows:

$$\psi_R = \frac{1}{\sqrt{2}}(\psi_s + \psi_a) \tag{12.107}$$

$$\psi_L = \frac{1}{\sqrt{2}}(\psi_s - \psi_a) \tag{12.108}$$

The sketches of ψ_s and ψ_a are shown in Fig. 12.7.

ψ_L is almost zero in region I and III, and so this represents the configuration of ammonia molecule with nitrogen at the left vertex. Similarly, ψ_R represents configuration with nitrogen at the right vertex. For a detailed numerical calculation, the reader can refer to Basdevant and Dalibard.

Here we follow a different route based on Feynman's discussion in his famous book *Feynman's Lectures in Physics Vol. III*. Feynman presents a set of plausibility arguments for the Hamiltonian matrix and establishes the existence of the states with different energy eigenvalues and finally demonstrates the fluctuation between the two configurations.

Let $|\mathrm{L}\rangle$ and $|\mathrm{R}\rangle$ be the state vectors describing the left and right configuration. Let us write the general state vector $|\psi\rangle$ as

$$|\psi\rangle = C_1(t)\,|\mathrm{R}\rangle + C_2(t)\,|\mathrm{L}\rangle \tag{12.109}$$

Further, let us assume that the system is in the state $|\mathrm{R}\rangle$ at $t = 0$.

i.e., $$C_1(0) = 1 \; C_2(0) = 0 \tag{12.110}$$

The time evolution of the state vector $|\psi(0)\rangle$ is determined by the Schrödinger equation

$$i\hbar\frac{\partial}{\partial t}|\psi(\mathrm{t})\rangle = \hat{H}\,|\psi(\mathrm{t})\rangle \tag{12.111}$$

In terms of matrices, we have

$$\psi(t) = \begin{pmatrix} C_1(t) \\ C_2(t) \end{pmatrix}$$

and the Hamiltonian H is

$$H = \begin{pmatrix} \langle R|\hat{H}|R\rangle & \langle R|\hat{H}|L\rangle \\ \langle L|\hat{H}|R\rangle & \langle L|\hat{H}|L\rangle \end{pmatrix}$$

Feynman does not solve the Schrödinger equation in any particular potential for this problem. He argues for the particular form of Hamiltonian from the general considerations. An outline of his argument is given here.

Let us start with the clues from classical physics. If we start with a particular configuration at $t = 0$, the molecule stays in the same configuration. In quantum mechanics, this corresponds to dealing with stationary eigenstates. So we have to assume $H_{RR} \neq 0$, $H_{LL} \neq 0$. But $H_{LR} = 0$ and $H_{RL} = 0$. For stationary eigenstates $C_1(t) = e^{-iH_{RR}t/\hbar}$ and $C_2(t) = e^{-iH_{LL}t/\hbar}$. Since each configuration is a mirror reflection, we expect $H_{RR} = H_{LL}$. Let us take

$$H_{RR} = H_{LL} = E_0$$

However, we know that this cannot be the complete story in quantum mechanics. Unlike classical physics, in quantum mechanics, there is a non-zero probability of finding the particle in region I (a classically forbidden region). We expect this probability amplitude to be very small. So, it is not correct to make H_{RL} and H_{LR} to be zero. We take H_{RL} and H_{LR} to be a small non-zero number. Again, from symmetry considerations, we take

$$H_{RL} = H_{LR}$$

$$\therefore H_{RL} = H_{LR} = -A$$

where A is a small number compare to E_0. So the Hamiltonian matrix is

$$H = \begin{pmatrix} E_0 & -A \\ -A & E_0 \end{pmatrix} \tag{12.112}$$

Since H is not diagonal, the state $|L\rangle$ and $|R\rangle$ are not energy eigenstates of the Hamiltonian.

The Hamiltonian can be diagonalized easily. The eigenvalue equation is

$$\begin{vmatrix} E_0 - \lambda & -A \\ -A & E_0 - \lambda \end{vmatrix} = 0$$

Or $$\lambda = E_0 - A, E_0 + A$$

The corresponding eigenvectors are given below:

$$E_S = E_0 - A; \qquad \frac{1}{\sqrt{2}}\begin{bmatrix}1\\1\end{bmatrix} \tag{12.113}$$

$$E_a = E_0 + A; \qquad \frac{1}{\sqrt{2}}\begin{bmatrix}1\\-1\end{bmatrix} \tag{12.114}$$

The diagonalizing matrix U is given by

$$U = \frac{1}{\sqrt{2}}\begin{bmatrix}1 & 1\\1 & -1\end{bmatrix}$$

So we have

$$\hat{H}\,|\psi_s\rangle = (E_0 - A)\,|\psi_s\rangle \tag{12.115}$$

Or

$$\hat{H}\,|\psi_a\rangle = (E_0 + A)\,|\psi_a\rangle \tag{12.116}$$

The state vectors $|R\rangle$ and $|L\rangle$ are given by

$$|R\rangle = \frac{1}{\sqrt{2}}[|\psi_s\rangle + |\psi_a\rangle] \tag{12.117}$$

and

$$|L\rangle = \frac{1}{\sqrt{2}}[|\psi_s\rangle - |\psi_a\rangle] \tag{12.118}$$

Let us write $|\psi(0)\rangle$ as

$$|\psi(0)\rangle = d_1\,|\psi_s\rangle + d_2\,|\psi_a\rangle \tag{12.119}$$

where d_1 and d_2 are constants.

$$\therefore\ |\psi(t)\rangle = d_1 e^{-i(E_0 - A)t/\hbar}\,|\psi_s\rangle + d_2 e^{-i(E_0 + A)t/\hbar}\,|\psi_a\rangle \tag{12.120}$$

Now $C_1(t)$ and $C_2(t)$ are given by

$$C_1(t) = \langle R\,|\,\psi\rangle \text{ and } C_2(t) = \langle L\,|\,\psi\rangle$$

$$C_1(t) = \left(\frac{1}{\sqrt{2}}\langle\psi_s| + \frac{1}{\sqrt{2}}\langle\psi_a|\right)|\psi(t)\rangle$$

$$= e^{-iE_0 t/\hbar}\,\frac{1}{\sqrt{2}}\left[e^{iAt/\hbar}d_1 + e^{-iAt/\hbar}d_2\right]$$

Similarly, we have

$$C_2(t) = e^{-iE_0 t/\hbar}\,\frac{1}{\sqrt{2}}[e^{iAt/\hbar}d_1 - e^{-iAt/\hbar}d_2]$$

From $C_2(0) = 0$, we get $d_1 = d_2$.

From $C_1(0) = 1$, we get $$d_1 = d_2 = \frac{1}{\sqrt{2}}$$

$$\therefore C_1(t) = e^{-iE_0 t/\hbar} \cos A\, t/\hbar \tag{12.121}$$

and $$C_2(t) = ie^{-iE_0 t/\hbar} \sin A\, t/\hbar \tag{12.122}$$

Let $P(L)$ and $P(R)$ be the probability of getting the configurations with nitrogen atoms in left vertex and right vertex respectively.

$$\therefore P(R) = |C_1(t)|^2 = \cos^2 \frac{At}{\hbar} \tag{12.123}$$

and $$P(L) = |C_2(t)|^2 = \sin^2 \frac{At}{\hbar} \tag{12.124}$$

These two equations are exactly the same as $P(\uparrow)$ and $P(\downarrow)$ for an electron in a constant magnetic field $B = B_1\hat{\mathbf{x}}$ in Example 12.5 (Equations (12.85) and (12.86)) and for an electron at resonant frequency in the case magnetic resonance (Equations (12.105) and (12.106)).

$$P(R) = 1 \text{ at } t = 0, \frac{\hbar\pi}{A}, \frac{2\hbar\pi}{A} \ldots$$

and

$$P(L) = 1 \text{ at } t = 0, \frac{\hbar\pi}{2A}, \frac{3\hbar\pi}{2A} \ldots$$

i.e., the nitrogen atom flips between the left vertex and the right vertex periodically in time. Assuming 2A ~ 10^{-4} eV, we can work out this inversion frequency which is found to be $\nu = 24$ GH_Z (or λ=1.25 cm).

12.12 NEUTRINO OSCILLATION

As one more example of two-state problem, we discuss the neutrino problem which brings out the salient features of the Hilbert space structure of quantum mechanics, especially choosing a basis to suit our convenience. Neutrino oscillation in details is too complex to be discussed in textbook at this level. We consider a simple model involving only two kinds of neutrinos with emphasis on the application of two-state problem. (See a simple and lucid exposition of neutrino oscillation by Chris Waltham (2003)). It takes too long to provide background information on neutrino oscillation.

To explain the continuous β-ray spectrum of radioactive nuclei, Pauli postulated a new kind of chargeless particle called neutrino. The nuclear radioactive decay process is now understood as

$$n \to p + e^- + \bar{\nu}_e$$

$$p \to n + e^+ + \nu_e$$

where ν_e and $\bar{\nu}_e$ stand for electron neutrino and electron antineutrino (the second reaction takes place only when it is a part of the nucleus). Neutrinos interact only through weak interaction. This makes it very difficult to detect them. We now know that there are six leptons. They are electron e^- and electron neutrino ν_e, muon μ^- and muon neutrino ν_μ, tau lepton τ^- and tau neutrino ν_τ. There exists corresponding six antiparticles e^+ and $\bar{\nu}_e$, μ^+ and $\bar{\nu}_\mu$, τ^+ and $\bar{\nu}_\tau$.

These particles are grouped into three generations:

First generation: $\begin{pmatrix} \nu_e \\ e^- \end{pmatrix}$

Second generation: $\begin{pmatrix} \nu_\mu \\ \mu^- \end{pmatrix}$

Third generation: $\begin{pmatrix} \nu_\tau \\ \tau^- \end{pmatrix}$

The other term for generation is flavour (this term has more broader meaning in elementary particle physics). Similarly, antiparticles can be grouped into three generations: $\begin{pmatrix} \bar{\nu}_e \\ e^+ \end{pmatrix}, \begin{pmatrix} \bar{\nu}_\mu \\ \mu^+ \end{pmatrix}, \begin{pmatrix} \bar{\nu}_\tau \\ \tau^+ \end{pmatrix}$. The mass of the electron is $\boldsymbol{m}_e = 0.5$ MeV; $m_\mu \simeq 200 m_e$; $m_\tau \simeq 3500 m_e$. Generally, it was believed for a longtime that the neutrinos are massless particles. However, neutrinos should have non-zero mass for neutrino oscillation. So the existence of neutrino oscillations implies that they have some non-zero mass though they may be very small. Each generation was assigned a lepton quantum member.

$\begin{pmatrix} \nu_e \\ e^- \end{pmatrix}$: electron lepton number $L_e = 1$

$\begin{pmatrix} \nu_\mu \\ \mu^- \end{pmatrix}$: muon lepton number $L_\mu = 1$

$\begin{pmatrix} \nu_\tau \\ \tau^- \end{pmatrix}$: τ lepton number $L_\tau = 1$

$\begin{pmatrix} \bar{\nu}_e \\ e^+ \end{pmatrix}$: electron lepton number $L_e = -1$

$\begin{pmatrix} \bar{\nu}_\mu \\ \mu^+ \end{pmatrix}$: muon lepton number $L_\mu = -1$

$\begin{pmatrix} \bar{\nu}_\tau \\ \tau^+ \end{pmatrix}$: tau lepton number $L_\tau = -1$

In elementary particle reactions, these three quantum numbers were believed to be conserved individually.

12.12.1 Solar Neutrino Problem

The concept of neutrino oscillation emerged as a solution to the solar neutrino problem. A brief account of solar neutrino problem is presented here. The interior of the sun is so hot that a number of nuclear reactions producing neutrinos are possible. Some of the reaction are given below:

$$\begin{aligned} p+p &\to {}^2H+e^+ +\nu_e \\ p+p+e^- &\to {}^2H+\nu_e \\ {}^7Be+e^- &\to {}^7Li+\nu_e \\ {}^8B &\to {}^8Be+e^+ +\nu_e \end{aligned}$$

By 1960s, a good astrophysical model for the sun has been developed. Using this model, Bahcall estimated theoretically the solar neutrino flux to be $6\times10^6\,/\mathrm{cm}^2/\mathrm{sec}$. In spite of this big number, it is very difficult to detect them since their interaction with the matter on the earth is very weak, and hence almost all of them simply pass through the earth without being detected. So it requires special effort to detect such neutrinos.

Raymond Davis constructed the first solar neutrino observatory deep inside a mine consisting of a huge tank filled with 100,000 gallons of cleaning fluid perchloroethylene, which contains 2×10^{30} chlorine atoms. The basic reaction is

$$\nu_e+n\to p+e^-$$

This in turn leads to the reaction $\nu_e + {}^{37}Cl \to {}^{38}Ar+e^-$. From the amount of argon atom collected, we can determine the solar neutrino flux. However, over a period of 30 years, the number of electron neutrinos detected is very less, and it is roughly only 30% of the neutrino flux computed theoretically. Same types of results have been obtained in other experiments also. So we always have

$$\nu_{\text{detected}}<\nu_{\text{theory}}\ \text{ for }\ \nu_e$$

- Why do we need neutrino oscillation?

The resolution of this anomaly has led to a new phenomenon called neutrino oscillation, which has been experimentally observed. Let us recall that in the case of an electron in a magnetic field, the Z spin sate is a superposition of spin-up and spin-down states and consequently the Z component flip-flops between spin-up and spin-down states. Neutrino oscillation is a similar phenomenon. The flavour takes the role of the Z component of the spin, and so we now have a state which is superposition of different flavour states. The effect of fluctuation in the flavour is more dramatic since changes in the flavour correspond to changes in identity as we know. We have to bear in mind that electron neutrino, muon neutrino and tau neutrino correspond to different flavours. In the neutrino oscillation, the neutrino fluctuates from one flavour to another flavour as it travels. For instance, the source may produce an electron neutrino. As it travels, it will change its identity (flavour) to become moun neutrino. The switching between electron neutrino and muon neutrino may take place many times before they reach the detector. This creates a new possibility. Though a source produces an electron neutrino beam, by the time they reach the detector, a fraction of these particles would have switched to become moun neutrinos. So the number of electron neutrinos detected will be less than the computed number from the source.

The explanation of neutrino oscillation requires masses of the neutrinos to be non-zero. It further requires two kinds of eigenstates, namely flavour eigenstates and mass eigenstates.

Flavour Eigenstates

There are three flavour eigenstates corresponding to the three kinds of neutrinos. They are $|\nu_e\rangle$ $|\nu_\mu\rangle$ and $|\nu_\tau\rangle$. These states have definite lepton numbers.

$$|\nu_e\rangle : L_e = 1;\ L_\mu = 0\ L_\tau = 0 \tag{12.125}$$

$$|\nu_\mu\rangle : L_e = 0\ L_\mu = 1\ L_\tau = 0 \tag{12.126}$$

$$|\nu_\tau\rangle : L_e = 0\ L_\mu = 0\ L_\tau = 1 \tag{12.127}$$

Mass Eigenstates

There are three distinct states $|\nu_1\rangle$, $|\nu_2\rangle$ and $|\nu_3\rangle$, each state representing particles with definite mass.

For a particle of mass m and momentum p, the energy E is given by

$$E = [p^2c^2 + m^2c^4]^{1/2}$$

Since neutrino mass is very small, we can write

$$E = pc\left[1 + \frac{m^2c^4}{p^2c^2}\right]^{1/2} \simeq pc + \frac{1}{2}\frac{m^2c^4}{pc} \tag{12.128}$$

For the sake of clarity, we consider only two neutrinos, namely ν_e and ν_μ.

The mass eigenstates $|\nu_1\rangle$ and $|\nu_2\rangle$ satisfy the following equations:

$$\hat{H}\,|\nu_1\rangle = \left(pc + \frac{1}{2}\frac{m_1^2c^4}{pc}\right)|\nu_1\rangle \tag{12.129}$$

and

$$\hat{H}\,|\nu_2\rangle = \left(pc + \frac{1}{2}\frac{m_2^2c^4}{pc}\right)|\nu_2\rangle \tag{12.130}$$

There are two sets of basis states, namely the mass eigenstates $\{|\nu_1\rangle, |\nu_2\rangle\}$ and the flavour eigenstates $\{|\nu_e\rangle, |\nu_\mu\rangle\}$. We can choose any particular basis states to suit our experimental situation. It is the flavour eigenstates which are physically produced and physically detected in our laboratories.

Let us write the flavour eigenstates $|\nu_e\rangle$ and $|\nu_\mu\rangle$ as a linear combination of mass eigenstates $|\nu_1\rangle$ and $|\nu_2\rangle$.

$$|\nu_e\rangle = C_1\,|\nu_1\rangle + C_2\,|\nu_2\rangle \tag{12.131}$$

$$|\nu_\mu\rangle = C_3\,|\nu_1\rangle + C_4\,|\nu_2\rangle \tag{12.132}$$

The normalization relations $\langle\nu_e|\nu_e\rangle = 1$ and $\langle\nu_\mu|\nu_\mu\rangle = 1$ lead to the equations

$$|C_1|^2 + |C_2|^2 = 1; \qquad |C_3|^2 + |C_4|^2 = 1$$

The orthoganality relation $\langle \nu_e | \nu_\mu \rangle = 0$ leads to

$$C_1^* C_3 + C_2^* C_4 = 0$$

These three relations can be satisfied if we write

$$C_1 = \cos\,\theta,\ C_2 = \sin\,\theta,\ C_3 = -\sin\,\theta,\ C_4 = \cos\,\theta$$

i.e.,
$$|\nu_e\rangle = \cos\theta\,|\nu_1\rangle + \sin\theta\,|\nu_2\rangle \tag{12.133}$$

$$|\nu_\mu\rangle = -\sin\theta\,|\nu_1\rangle + \cos\theta\,|\nu_2\rangle \tag{12.134}$$

θ is called mixing angle. The states $|\nu_1\rangle$ and $|\nu_2\rangle$ in turn can be written as

$$|\nu_1\rangle = \cos\theta\,|\nu_e\rangle - \sin\theta\,|\nu_\mu\rangle \tag{12.135}$$

$$|\nu_2\rangle = \sin\theta\,|\nu_e\rangle + \cos\theta\,|\nu_\mu\rangle \tag{12.136}$$

Let us assume that the source creates an electron neutrino at $t = 0$. Therefore, the state vector

$$|\nu(0)\rangle = |\nu_e\rangle = \cos\theta\,|\nu_1\rangle + \sin\theta\,|\nu_2\rangle \tag{12.137}$$

The time evolution of the state $|\nu(t)\rangle$ is given by

$$|\nu(t)\rangle = \cos\theta e^{-iE_1 t/\hbar}\,|\nu_1\rangle + \sin\theta\, e^{-iE_2 t/\hbar}\,|\nu_2\rangle \tag{12.138}$$

The probability of amplitude to get the state $|\nu_\mu\rangle$ is given by $\langle \nu_\mu | \nu(t)\rangle$.

$$\begin{aligned}\langle \nu_\mu | \nu(t)\rangle &= (-\sin\theta\langle \nu_1| + \cos\theta\langle \nu_2|)\,|\nu(t)\rangle \\ &= -\sin\theta\cos\theta e^{-iE_1 t/\hbar} + \sin\theta\cos\theta e^{-iE_2 t/\hbar} \\ &= e^{-iE_2 t/\hbar}\sin\theta\cos\theta[1 - e^{-i(E_2 - E_1)t/\hbar}]\end{aligned} \tag{12.139}$$

Similarly, $\langle \nu_e | \nu(t)\rangle$ is given by

$$\langle \nu_e | \nu(t)\rangle = e^{-iE_1 t/\hbar}[\cos^2\theta + \sin^2\theta e^{-i(E_2 - E_1)t/\hbar}] \tag{12.140}$$

$$\begin{aligned}\therefore |\nu(t)\rangle = e^{-iE_1 t/\hbar}[\cos^2\theta + \sin^2\theta e^{-i(E_2 - E_1)t/\hbar}]\,|\nu_e\rangle \\ + e^{-iE_1 t/\hbar}\sin\theta\cos\theta[1 - e^{-i(E_2 - E_1)t/\hbar}]\,|\nu_\mu\rangle\end{aligned} \tag{12.141}$$

Compare $|\nu(0)\rangle$ and $|\nu(t)\rangle$. The state $|\nu(0)\rangle$ is the flavour eigenstates corresponding to electron neutrino. So we start with an electron neutrino at $t = 0$. At a later time t, the state vector $|\nu(t)\rangle$ is a superposition of two flavour eigenstates $|\nu_e\rangle$ and $|\nu_\mu\rangle$. So if one tries to detect the neutrino, the result may be an electron neutrino or a muon neutrino. We can tell only the probability of detecting the electron neutrino or muon neutrino. The probability of getting the muon neutrino $P(\nu_\mu)$ is given by

$$\begin{aligned} P(\nu_\mu) &= |e^{-iE_1 t/\hbar}(1-e^{i(E_2-E_1)t/\hbar})|^2 \sin^2\theta\cos^2\theta \\ &= 4\sin^2\theta\cos^2\theta\sin^2\frac{\Delta Et}{2\hbar} \\ &= \sin^2 2\theta\sin^2\frac{\Delta Et}{2\hbar} \end{aligned} \tag{12.142}$$

$$\Delta E = E_2 - E_1 = \frac{m_2{}^2c^4 - m_1{}^2c^4}{2pc} \simeq \frac{(m_2{}^2 - m_1^2)c^4}{2E} \tag{12.143}$$

where we have used the fact $pc \simeq E$.

Let us write ΔE as

$$\Delta E = \frac{\Delta m^2c^4}{2E}$$

$$\therefore P(\nu_e \to \nu_\mu) = \sin^2 2\theta\sin^2\frac{\Delta m^2c^4t}{4E\hbar} \tag{12.144}$$

Since the mass of the neutrino is almost zero, we can assume that it travels with a velocity c. Let L be the distance travelled by the particle in time t. i.e., $L \simeq ct$.

$$\therefore P(\nu_e \to \nu_\mu) = \sin^2 2\theta\sin^2\frac{\Delta m^2c^3L}{4E\hbar} \tag{12.145}$$

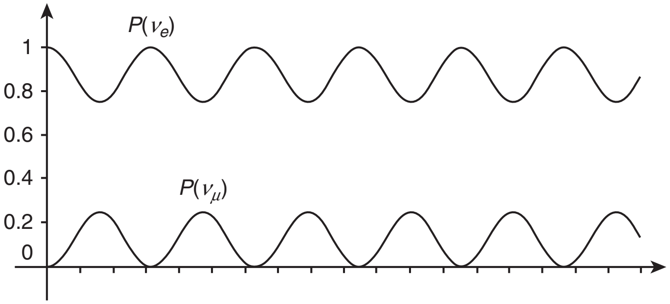


Fig. 12.10 The probabilities $P(\nu_e \to \nu_e)$ and $P(\nu_e \to \nu_\mu)$

For a neutrino ν, which starts as electron neutrino ν_e, the probability of detecting it as a moun neutrino ν_μ varies from 0 to a maximum value $\sin^2 2\theta$ periodically as a function of distance. Hence, this phenomenon is called neutrino oscillation.

The probability of finding the electron neutrino varies from 1 to a minimum value of $\cos^2 2\theta$. So if we are interested only in the number of electron neutrinos, only a fraction of the initial number will be detected as electron neutrino at a later time t.

We have presented the neutrino oscillation as an example for two-state problems. The real neutrino oscillation problem is more complex. The neutrino oscillations have been verified in a number of experiments, extensive literature or the experimental details and available. Though the concept of neutrino oscillation originated from the solar neutrino problem, it is true for neutrinos from any other source. For instance, neutrinos can be from the nuclear reactors or the cosmic ray reactions in the earth's atmosphere. Neutrino oscillations have been verified for such neutrinos also.

For the sake of completion, let us go back to solar neutrino problem. In the Sudbury Neutrino Observatory, three kinds of reactions were studied, of which one reaction involves the detection of neutrinos independent of flavour and another reaction involves the detection of electron neutrinos alone. The observational data from Sudbury Neutrino Observatory is in good agreement with the concept of neutrino oscillation. The detected electron neutrino flux was verified to be 34%, and the detected flux due to muon neutrinos and tau neutrinos were verified to be 66% of the solar neutrinos.

The total neutrino flux was found to be $(5.25\pm0.16(\text{stat.})^{+0.11}_{-0.13}(\text{syst.}))\times10^6/\text{cm}^2/\text{sec}$, which has to be compared with theoretical prediction $(5.88\pm0.65)\times10^6/\text{cm}^2/\text{sec}$ and $(4.85\pm0.58)\times10^6/\text{cm}^2/\text{sec}$ calculated using two theoretical models.

EXERCISES

1. The Hamiltonian for a charged particle is

$$H=\frac{p^2}{2m}+V(\mathbf{r})$$

What is the Hamiltonian for this particle in the presence of electric field **E** and magnetic field **B**?

2. Why do we need vector potential **A** and scalar potential ϕ when we can do electromagnetism using electric field **E** and magnetic field **B**?
3. What is meant by gauge invariance?
4. What is the gauge transformation for the quantum mechanical wave function of a particle?
5. Obtain the Landau level by choosing the vector potential $\mathbf{A}$ as $\mathbf{A}=(-yB,0,0)$.
6. Obtain the Landau level when the vector $\mathbf{A}$ is $\mathbf{A}=-1/2\,\mathbf{r}\times\mathbf{B}$.
7. What is meant by two-state problem?
8. Mention some of the basis states to describe the states of a spin 1/2 particle.
9. What is meant by Larmor's precession?
10. An electron in the Z spin-up state is subjected to a constant magnetic field $\mathbf{B}=B_0\hat{\mathbf{z}}+B_1\hat{\mathbf{x}}+B_2\hat{\mathbf{y}}$. Determine the time evolution of the state of the electron and the probability of finding the electron in Z spin-down state.
11. Give the justification for using double-well potential to describe ammonia molecule.
12. Distinguish between mass eigenstates and flavour eigenstates.
13. What is meant by neutrino oscillation?
14. Can we have neutrino oscillation if the neutrinos are massless particles?
15. Does neutrino oscillation conserve the flavour quantum members L_e and L_μ?.

REFERENCES

1. Gaziorowicz, 1974. *Quantum Physics.* New York: John Wiley & Sons.
2. R. G. Chambers, 1960. 'Shift of an Electron Interference Pattern Recognized by Enclosed Magnetic Flux', *Phys.Rev.Lett*. 5,3.
3. N. Osakabe, T. Matguda, T. Kawasaki, J. Endo, A. Tonomura, S. Yano, H. Yamada, 1986. 'Experimental Confirmation of Aharonov–Bohm Effect Using a Toraidal Magnetic Field Confined by a Superconductor', *Phys.Rev.A* 34, 815.
4. J. L. Basdevant and J. Dalibard, 2002. *Quantum Mechanics*. Berlin: Spinger-Verlag.

5. I. I. Rabi, S. Millman, P. Kusch and J. R. Zachario, 1939. 'The Molecular Beam Magnetic Resonance Method for Nuclear Magnetic Moments', *Phy.Rev.* 55,526.

6. M. Alonso and H. Valk, 1973. *Quantum Mechanics: Principles and Application.* Reading Massachusetts: Addison Wesley Publishing Company.

7. R. P. Feynman, R. B. Leighton and M. Sands, 1973. *The Feynman Lectures on Physics III, Quantum mechanics.* Reading Massachusetts: Addison Wesley Publishing Company.

8. Chris Waltham, Neutrino Oscillations for Dummies-arXiv:physics/0303116v1 [physics.ed-ph] 28, March 2003.

9. A.McDonald,SolarNeutrinos,SNO,SNOLAB-http://www.powershow.com/view1/5bda4-NDE0O/Solar_Neutrinos_SNO_and_SNOLAB_powerpoint_ppt_presentation

10. Aharmin et al. Combined Analysis of All the Three Phases of Solar Neurino Data from Sudbury Neutrino Observatory-arXiv-1109.0763v1 [nucl-ex]-4 Sep, 2011.

13

Time-Independent Schrödinger Equation – Approximations

In Chapters 3, 6, 7 and 8, the Schrödinger equation was solved for a number of simple potentials. However, only in a very few cases, such exact analytical solutions for Schrödinger equations are available. For other potentials, we have to resort various kinds of approximations. For time-independent Schrödinger equations, broadly we consider three kinds of approximations: perturbation theory, variation method and WKB approximation.

A. Time-Independent Perturbation Theory

In perturbation theory, we split the Hamiltonian into two parts, H_0 and $\lambda H'$.

$$H = H_0 + \lambda H' \tag{13.1}$$

The Hamiltonian H_0 is chosen in such a way that exact solutions are available for H_0.

$$H_0 \,|\, u_n\rangle = E_n^{(0)} \,|\, u_n\rangle \tag{13.2}$$

The eigenvalue equation for H is

$$H \,|\, \psi_n\rangle = E_n \,|\, \psi_n\rangle \tag{13.3}$$

Obviously,

$$E_n = E_n^{(0)} + \Delta E_n \tag{13.4}$$

and

$$|\, \psi_n\rangle = |\, u_n\rangle + |\, \Delta u_n\rangle \tag{13.5}$$

ΔE_n is the correction to energy eigenvalue $E_n^{(0)}$, and $|\, \Delta u_n\rangle$ is the correction to the eigenstate $|\, u_n\rangle$.

We consider the following two cases in perturbation theory:

1. Non-degenerate perturbation theory to determine the correction to a non-degenerate eigenvalue $E_n^{(0)}$.
2. Degenerate perturbation theory to determine the correction to a degenerate eigenvalue.

13.1 NON-DEGENERATE PERTURBATION THEORY

Let us write E_n and $|\psi_n\rangle$ as

$$E_n = E_n^{(0)} + \lambda E_n^{(1)} + \lambda^2 E_n^{(2)} + \cdots \tag{13.6}$$

$$|\psi_n\rangle = |u_n\rangle + \lambda|\psi_n^{(1)}\rangle + \lambda^2|\psi_n^{(2)}\rangle \cdots \tag{13.7}$$

From the Equation (13.3), we get

$$(\hat{H}_0 + \lambda\hat{H}' - E_n)|\psi_n\rangle = 0$$

$$(\hat{H}_0 + \lambda\hat{H}' - E_n^{(0)} - \lambda E_n^{(1)} - \lambda^2 E_n^{(2)} \ldots)\big(|u_n\rangle + \lambda\,|\psi_n^{(1)}\rangle + \cdots\big) = 0$$

$$\begin{aligned}&[\hat{H}_0 - E_n^{(0)}]\,|u_n\rangle + \lambda[(\hat{H}_0 - E_n^{(0)})\,|\psi_n^{(1)}\rangle - (E_n^{(1)} - \hat{H}')\,|u_n\rangle]\\ &+\lambda^2[(\hat{H}_0 - E_n^{(0)})\,|\psi_n^{(2)}\rangle - (E_n^{(1)} - \hat{H}')\,|\psi_n^{(1)}\rangle - E_n^{(2)}\,|\psi_n^{(2)}\rangle] + \cdots = 0\end{aligned} \tag{13.8}$$

This is an expansion in terms of λ. Equating the coefficients of different powers of λ separately to zero, we get

$$[\hat{H}_0 - E_n^{(0)}]\,|u_n\rangle = 0 \tag{13.9}$$

$$[\hat{H}_0 - E_n^{(0)}]\,|\psi_n^{(1)}\rangle = (E_n^{(1)} - \hat{H}')\,|u_n\rangle \tag{13.10}$$

$$[\hat{H}_0 - E_n^{(0)}]\,|\psi_n^{(2)}\rangle = (E_n^{(1)} - \hat{H}')\,|\psi_n^{(1)}\rangle + E_n^{(2)}\,|u_n\rangle \tag{13.11}$$

$$[\hat{H}_0 - E_n^{(0)}]\,|\psi_n^{(3)}\rangle = (E_n^{(1)} - \hat{H}')\,|\psi_n^{(2)}\rangle + E_n^{(2)}\,|\psi_n^{(1)}\rangle + E_n^{(1)}\,|u_n\rangle \tag{13.12}$$

Consider the Equation (13.10). The unknown quantities in this equation are $|\psi_n^{(1)}\rangle$ and $E_n^{(1)}$, which are to be determined. The LHS of the Equation (13.10) is invariant under $|\psi_n^{(1)}\rangle + C_n\,|u_n\rangle$, which can be verified easily.

$$\begin{aligned}(\hat{H}_0 - E_n^{(0)})(|\psi_n^{(1)}\rangle + C_n\,|u_n\rangle) &= (\hat{H}_0 - E_n^{(0)})\,|\psi_n^{(1)}\rangle + C_n\,\underbrace{(\hat{H}_0 - E_n^{(0)})\,|u_n\rangle}_{0}\\ &= (\hat{H}_0 - E_n^{(0)})\,|\psi_n^{(1)}\rangle\end{aligned}$$

Therefore, both the corrections $|\psi_n^{(1)}\rangle$ and $|\psi_n^{(1)}\rangle + C_n\,|u_n\rangle$ will lead to the same value $E_n^{(1)}$.

The same reasoning can be applied for other equations. So, for a given $E_n^{(r)}$, there is an element of arbitrariness in the correction terms $|\psi_n^{(r)}\rangle$. This arbitrariness can be removed by a suitable condition which is chosen as

$$\langle u_n\,|\psi_n^{(r)}\rangle = 0 \tag{13.13}$$

i.e., all the correction terms for the wave function are orthogonal to $|u_n\rangle$. This enables us to write

$$|\psi_n^{(r)}\rangle = \sum_{k\neq n} C_{nk}^{(r)}\,|u_k\rangle \tag{13.14}$$

The exclusion of the term $|u_n\rangle$ from the summation in the RHS ensures the orthoganality between $|u_n\rangle$and $|\psi_n^{(r)}\rangle$.

$$\therefore |\psi_n\rangle = |u_n\rangle + \lambda\sum_{k\neq n} C_{nk}^{(1)}\,|u_k\rangle + \lambda^2\sum_{k\neq n} C_{nk}^{(2)}\,|u_k\rangle \ldots \tag{13.15}$$

This state is not normalized to unity. We will discuss the normalization later.

13.1.1 First-order Perturbation Theory

In the first-order perturbation theory, we restrict E_n and $|\psi_n\rangle$ to the first order in λ.

$$E_n = E_n^{(0)} + \lambda E_n^{(1)} \tag{13.16}$$

$$|\psi_n\rangle = |u_n\rangle + \lambda\sum_{k\neq n} C_{nk}^{(1)}\,|u_k\rangle \tag{13.17}$$

Taking the scalar product of the Equation (13.10) with $\langle u_n|$, we get

$$\langle u_n|(\hat{H}_0 - E_n^{(0)})|\psi_n^{(1)}\rangle = E_n^{(1)}\langle u_n|u_n\rangle - \langle u_n|\hat{H}'|u_n\rangle$$

The left-hand side of the equation becomes

$$\langle u_n|(\hat{H}_0 - E_n^{(0)})|\psi_n^{(1)}\rangle = \langle u_n|(E_n^{(0)} - E_n^{(0)})|\psi_n^{(1)}\rangle = 0.$$

$$\therefore E_n^{(1)} = \langle u_n|H'|u_n\rangle = H'_{nn} \tag{13.18}$$

To determine the wave function $|\psi_n^{(1)}\rangle$, let us take the scalar product of the Equation (13.10) with $\langle u_k|$.

$$\therefore \langle u_k|(\hat{H}_0 - E_n^{(0)})|\psi_n^{(1)}\rangle = E_n^{(1)}\langle u_k|u_n\rangle - \langle u_k|\hat{H}'|u_n\rangle$$

$$\begin{aligned} \text{LHS} &= \langle u_k|(E_k^{(0)} - E_n^{(0)})|\psi_n^{(1)}\rangle \\ &= [E_k^{(0)} - E_n^{(0)}]\langle u_k|\psi_n^{(1)}\rangle \\ &= (E_K{}^{(0)} - E_n{}^{(0)})\, C_{nk}^{(1)} \end{aligned}$$

where $C_{nk}^{(1)}$ is given by

$$C_{nk}^{(1)} = \langle u_k|\psi_n^{(1)}\rangle$$

$$\text{RHS} = -\langle u_k|\hat{H}'|u_n\rangle = -H'_{kn}$$

$$\therefore C_{nk}^{(1)} = \frac{H'_{kn}}{E_n^{(0)} - E_k^{(0)}} \tag{13.19}$$

$$\therefore |\psi_n\rangle = |u_n\rangle + \lambda\sum_{k\neq n}\frac{H'_{kn}}{(E_n^{(0)} - E_k^{(0)})}|u_k\rangle \tag{13.20}$$

Let us write the normalized wave function as

$$|\psi_n\rangle = N\left[|u_n\rangle + \lambda\sum_{k\neq n} C_{nk}^{(1)}\,|u_k\rangle\right]$$

The normalization of this wave function leads to the following equation. Let us recall that the normalization of $\psi = \sum_n C_n\phi_n$ implies $\sum_n |C_n|^2 = 1$.

$$\langle\psi_n|\psi_n\rangle = N^2\left[1+\lambda^2\sum_{k\neq n}|C_{nk}^{(1)}|^2\right] = 1$$

However, consistency would require us to restrict $\langle\psi_n|\psi_n\rangle$ up to order λ since we have restricted E_n as well as $|\psi_n\rangle$ up to order λ.

$$\therefore \langle\psi_n|\psi_n\rangle = N^2 = 1$$

Therefore, the normalized wave function in the first order is

$$|\psi_n\rangle = |u_n\rangle + \lambda\sum_{k\neq n} C_{nk}^{(1)}\,|u_k\rangle \tag{13.21}$$

13.1.2 Second-order Perturbation Theory

To find the second-order correction, take the scalar product of the Equation (13.11) with $\langle u_n|$.

$$\langle u_n|(\hat{H}_0 - E_n^{(0)})|\psi_n^{(2)}\rangle = u_n|(E_n^{(1)} - \hat{H}')|\psi_n^{(1)}\rangle + E_n^{(2)}$$

The left-hand side of the above equation is given by

$$\langle u_n|(\hat{H}_0 - E_n^{(0)})|\psi_n^{(2)}\rangle = \langle u_n|(E_n^{(0)} - E_n^{(0)})|\psi_n^{(2)}\rangle = 0$$

$$\therefore E_n^{(2)} = \langle u_n|\hat{H}'|\psi_n^{(1)}\rangle - \underbrace{\langle u_n|E_n^{(1)}|\psi_n^{(1)}\rangle}_{0}$$

$$= \langle u_n|\hat{H}'\sum_{k\neq n} C_{nk}^{(1)}|u_k\rangle = \sum_{k\neq n} C_{nk}^{(1)}\langle u_n|\hat{H}'|u_k\rangle$$

$$= \sum_{k\neq n} C_{nk}^{(1)} H'_{nk} = \sum_k \frac{H'_{nk}H'_{nk}}{E_n^{(0)} - E_k^{(0)}} = \sum_k \frac{|H'_{nk}|^2}{E_n^{(0)} - E_k^{(0)}} \tag{13.22}$$

- Note that for the ground state $E_n^{(0)}$ is the lowest energy. Therefore $E_n^{(2)}$ is always negative for ground state.

The second-order correction to the wave function can be determined as follows:

Taking scalar product of the Equation (13.11) with $\langle u_l|$, we get

$$\langle u_l|(\hat{H}_0 - E_n^{(0)})|\psi_n^{(2)}\rangle = \langle u_l|E_n^{(1)}|\psi_n^{(1)}\rangle - \langle u_l|H'|\psi_n^{(1)}\rangle \quad \text{for} \quad l\neq n$$

The left-hand side of the above equation is given by

$$
\begin{aligned}
\langle u_l\,|(\hat{H}_0 - E_n^{(0)})\,|\psi_n^{(2)}\rangle &= \langle u_l\,|E_l^{(0)} - E_n^{(0)})\,|\psi_n^{(2)}\rangle \\
&= -(E_n^{(0)} - E_l^{(0)})\langle u_l\,|\psi_n^{(2)}\rangle \\
&= -(E_n^{(0)} - E_l^{(0)})C_{nl}^{(2)}
\end{aligned}
$$

$$
\begin{aligned}
\therefore C_{nl}^{(2)} &= \frac{\langle u_l\,|\,\hat{H}'\,|\psi_n^{(1)}\rangle}{E_n^{(0)} - E_l^{(0)}} - \frac{\langle u_l\,|\psi_n^{(1)}\rangle E_n^{(1)}}{E_n^{(0)} - E_l^{(0)}} \\
&= \frac{\langle u_l\,|\,\hat{H}'\,|\psi_n^{(1)}\rangle}{E_n^{(0)} - E_l^{(0)}} - \frac{C_{nl}^{(1)} E_n^{(1)}}{E_n^{(0)} - E_l^{(0)}} \\
&= \frac{\langle u_l\,|\,\hat{H}'\sum\limits_{k\neq n} C_{nk}^{(1)}\,|u_k\rangle}{E_n^{(0)} - E_l^{(0)}} - \frac{H'_{l\,n}}{[E_n^{(0)} - E_l^{(0)}]^2}H'_{nn}
\end{aligned}
$$

$$
C_{nl}^{(2)} = \sum_{k\neq n}\frac{H'_{lk}H'_{kn}}{[E_n^{(0)} - E_l^{(0)}][E_n^{(0)} - E_k^{(0)}]} - \frac{H'_{l\,n}}{[E_n^{(0)} - E_l^{(0)}]^2}H'_{nn} \qquad l \neq n \tag{13.23}
$$

This wave function is not a normalized function. The normalized wave function can be obtained easily. The normalized wave function can be written as

$$
|\psi_n\rangle = N\left[|u_n\rangle + \left(\lambda\sum_{k\neq n} C_{nk}^{(1)} + \lambda^2\sum_{k\neq n} C_{nk}^{(2)}\right)|u_k\rangle\right]
$$

The normalization implies

$$
N^2\left[1 + \lambda^2\sum_{k\neq n}\left|C_{nk}^{(1)}\right|^2 + \cdots\right] = 1
$$

Since we are discussing second-order perturbation theory, we can take up to λ^2 terms. Then we have

$$
N = \left[1 + \lambda^2\sum_{k\neq n}\left|C_{nk}^{(1)}\right|^2 \cdots\right]^{-1} = \left[1 - \frac{\lambda^2}{2}\sum_{k\neq n}\left|C_{nk}^{(1)}\right|^2\right]
$$

$$
\begin{aligned}
\therefore |\psi_n\rangle &= \left[1 - \frac{\lambda^2}{2}\sum_{k\neq n}\left|C_{nk}^{(1)}\right|^2\right]\left[|u_n\rangle + \left(\lambda\sum_{k\neq n} C_{nk}^{(1)} + \lambda^2\sum_{k\neq n} C_{nk}^{(2)}\right)|u_k\rangle\right] \\
&= |u_n\rangle + \lambda\sum_{k\neq n} C_{nk}^{(1)}\,|u_k\rangle + \left(\lambda^2\sum_{k\neq n} C_{nk}^{(2)}\right)|u_k\rangle - \frac{\lambda^2}{2}\sum_{k\neq n}\left|C_{nk}^{(1)}\right|^2\,|u_n\rangle
\end{aligned}
$$

Example 13.1 The Hamiltonian for an Harmonic oscillator is given by

$$H = \frac{p^2}{2m} + \frac{1}{2}m\omega^2 x^2 + \beta x^4$$

Taking βx^4 as the perturbation, determine the first-order correction to the energy $E_n^{(0)} = \left(n + \frac{1}{2}\right)\hbar\omega$.

Solution:

$$\Delta E_n = \langle n | H' | n \rangle$$

For Harmonic oscillator, $$\hat{x} = \left(\frac{\hbar}{2m\omega}\right)^{1/2} [a + a^\dagger]$$

$$\therefore \Delta E_n = \beta\left(\frac{\hbar}{2m\omega}\right)^2 \langle n | (a + a^\dagger)^4 | n \rangle$$

Note that every operator $a^\dagger$ acting on a ket $|n\rangle$ produces a new ket $|n+1\rangle$ and every operator a operating on $|n\rangle$ produces a new ket $|n-1\rangle$. Consider an operator $\hat{F}$ consisting of m annihilation operators and n creation operators. Irrespective of the order of the individual operators,

$$\hat{F} | 0 \rangle = \text{Const } |n - m\rangle$$

$$\therefore \langle 0 | \hat{F} | 0 \rangle = \text{Const } \langle 0 | n - m \rangle = 0 \text{ if } n \neq m$$

$\therefore \langle 0 | \hat{F} | 0 \rangle \neq 0$ only when $\hat{F}$ consists of equal number of creation and annihilation operators.

The same result holds good for $\langle n | \hat{F} | n \rangle$ also.

i.e., $\langle n | \hat{F} | n \rangle \neq 0$ only when $\hat{F}$ consists of equal number of creation and annihilation operators.

$$\therefore \langle n | (a + a^\dagger)^4 | n \rangle = \langle n | (a^2 + a^{\dagger 2} + aa^\dagger + a^\dagger a)^2 | n \rangle$$

By considering the terms only with equal number of creation and annihilation operators, we have

$$\langle n | (a + a^\dagger)^4 | n \rangle = \langle n | a^2 a^{\dagger 2} | n \rangle + \langle n | a^\dagger a\, a^\dagger a + a^\dagger a\, aa^\dagger + aa^\dagger a^\dagger a + aa^\dagger aa^\dagger | n \rangle + \langle n | a^{\dagger 2} a^2 | n \rangle$$

Let us recall that

$$a^\dagger |n\rangle = \sqrt{n+1}\, |n+1\rangle \quad ; \quad \langle n| a = \sqrt{n+1} \langle n+1|$$
$$a |n\rangle = \sqrt{n}\, |n-1\rangle \quad ; \quad \langle n| a^\dagger = \sqrt{n} \langle n-1|$$

We have

$$(a^\dagger)^2 |n\rangle = \sqrt{(n+1)(n+2)}\, |n+2\rangle$$
$$\langle n| a^2 = \sqrt{(n+1)(n+2)}\, \langle n+2|$$

$$\therefore \langle n| a^2 a^{\dagger 2} |n\rangle = (n+1)(n+2)$$
$$\langle n| a^\dagger a \underbrace{a^\dagger a |n\rangle}_{n|n\rangle} = n^2$$

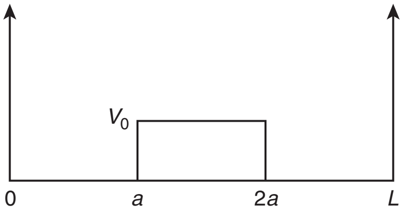
We have, $a\,a^{\dagger}\left|n\right\rangle = (n+1)\left|n\right\rangle$. Making use of the this equation, we get the following results:

$$\begin{aligned}\therefore\, a^{\dagger}aaa^{\dagger}\,|n\rangle &= (n+1)n\,|n\rangle\\ aa^{\dagger}a^{\dagger}a\,|n\rangle &= n(n+1)\,|n\rangle\\ aa^{\dagger}aa^{\dagger}\,|n\rangle &= (n+1)^2\,|n\rangle\end{aligned}$$

Using all these results, we get

$$\begin{aligned}\therefore\,\langle n|\,x^4\,|n\rangle &= \beta\left(\frac{\hbar}{2m\omega}\right)^2\{(n+1)(n+2)+n^2+2(n+1)n+(n+1)^2+n(n-1)\}\\ &= \frac{3}{2}\beta\left(\frac{\hbar}{m\omega}\right)^2\left(n^2+n+\frac{1}{2}\right).\end{aligned}$$

Example 13.2 Consider a particle in a box of length L. Take the rectangular potential shown in the box as a small perturbation. Determine the correction to the energy levels for the particle in the box using the first-order perturbation theory.



Solution: The unperturbed eigenvalues and the corresponding eigenfunctions are given by

$$E_n = \frac{\hbar^2 n^2 \pi^2}{2m}; \qquad u_n = \sqrt{\frac{2}{L}}\sin\frac{n\pi x}{L}$$

The perturbation H' is given by

$$H' = \begin{cases} V_0 & a \le x \le 2a \\ 0 & \text{otherwise}\end{cases}$$

The correction to the energy is given by

$$\begin{aligned}\Delta E_n &= \langle u_n|\,H'\,|u_n\rangle = \int_a^{2a} u_n^{*} V_0 u_n\,dx\\ &= \int_a^{2a}\frac{2}{L}V_0\sin^2\frac{n\pi x}{L}\,dx = \frac{2}{L}V_0\int_a^{2a}\frac{1}{2}\left(1-\cos\frac{2n\pi x}{L}\right)dx\\ &= V_0\left[\frac{a}{L}-\frac{1}{2n\pi}\left(\sin\frac{4\pi na}{L}-\sin\frac{2\pi na}{L}\right)\right]\end{aligned}$$

13.2 DEGENERATE PERTURBATION THEORY

In non-degenerate perturbation theory, the expression for the energy of the state $|\psi_n\rangle$ is given by

$$E_n = E_n^{(0)} + \langle u_n|\,\hat{H}'\,|u_n\rangle + \sum_{k\neq n}\frac{|H'_{kn}|^2}{E_n^{(0)} - E_k^{(0)}} \tag{13.24}$$

For a non-degenerate energy level of $\hat{H}_0$, the expression for the correction to the energy level $E_n^{(0)}$ is always finite. The second term is much less than the first, and the third term is much less than the second term. Only then, the perturbation theory is meaningful.

We will see soon that if we try to use this expression for degenerate energy levels, the third term becomes infinite. Obviously, we have to modify the expression for the correction to a degenerate energy level.

The expression was obtained by assuming the energy levels of the Hamiltonian H_0 are non-degenerate; i.e., for each energy eigenvalue $E_n^{(0)}$, there is only one eigenstate $|u_n\rangle$. These eigenstates are chosen as the basis vectors for Hilbert space for the given problem, and so $|\psi_n\rangle$ is written as

$$|\psi_n\rangle = \sum C_n\,|u_n\rangle$$

Let us now assume that some of the energy levels are degenerate. The eigenstates are now characterized by $|u_{n\alpha}\rangle$, where $\alpha = 1, 2, 3\ldots$

i.e.,
$$\hat{H}_0\,|u_{n\alpha}\rangle = E_n^{(0)}\,|u_{n\alpha}\rangle \quad \alpha = 1, 2\,\ldots. \tag{13.25}$$

The complete set of eigenstates is now given by

$$\left\{\underbrace{|u_{11}\rangle, |u_{12}\rangle, \ldots, |u_{1r}\rangle}_{E_1^{(0)}}, \ldots, \underbrace{|u_{n1}\rangle, |u_{n2}\rangle, \ldots, |u_{ns}\rangle}_{E_n^{(0)}}, \ldots\right\}$$

Choice of the Basis States

The existence of a set of degenerate states offers a new flexibility. Let the set of states $\{u_{n1}, u_{n2}, u_{n3}\ldots\}$ be a set of degenerate states corresponding to the energy level $E_n^{(0)}$. In principle, they themselves can be taken as the basis states, or a linear combination of them also can serve as basis states. We can choose any set of degenerate states as the basis states to suit our convenience.

A simple extension of non-degenerate perturbation theory suggests that the set of states $\left\{\underbrace{|u_{11}\rangle, |u_{12}\rangle, \ldots, |u_{1r}\rangle}_{E_1^{(0)}}, \ldots, \underbrace{|u_{n1}\rangle, |u_{n2}\rangle, \ldots, |u_{ns}\rangle}_{E_n^{(0)}}, \ldots\right\}$ themselves may be chosen as the basis vectors.

$$|\psi_n\rangle = \sum_n \sum_\alpha C_{n\alpha}\,|u_{n\alpha}\rangle \tag{13.26}$$

With this choice of basis vectors, the first- and the second-order energy $E_{n\alpha}$ is given by

$$E_{n\alpha} = E_n^{(0)} + \langle u_{n\alpha}| H' |u_{n\alpha}\rangle + \sum_{n'\alpha' \neq n\alpha} \frac{\left|\langle n\alpha| H' |n'\alpha'\rangle\right|^2}{E_n^{(0)} - E_{n'}^{(0)}} \tag{13.27}$$

Note that in the case of degenerate states, it is possible to have α and α' as different for the same n. In such cases, the denominator becomes zero, and so the third term becomes infinite. How do we avoid this problem?

Note that in the summation, the diagonal elements have been excluded. Let us now make use of this fact to avoid the third term in (13.27). For a diagonal matrix, all the non-diagonal elements are zero The standard method is to choose a new basis states (eigenstates of H_0) for which the matrix H' is diagonal.

i.e., if $|V_{n\alpha}\rangle$ is new basis state, it should be such that

$$\langle V_{n\alpha}| \hat{H}' |V_{n'\alpha'}\rangle = E_{n\alpha}^{(1)} \delta_{nn'} \delta_{\alpha\alpha'} \tag{13.28}$$

i.e.,
$$\hat{H}' |V_{n\alpha}\rangle = E_{n\alpha}^{(1)} |V_{n\alpha}\rangle \tag{13.29}$$

Perturbation Theory in New Basis

For the sake of clarity, let us consider two-fold degenerate eigenstate of H_0.

i.e.,
$$\hat{H}_0 |u_{n1}\rangle = E_n^{(0)} |u_{n1}\rangle \qquad \hat{H}_0 |u_{n2}\rangle = E_n^{(0)} |u_{n2}\rangle \tag{13.30}$$

Let us write $|V_{n1}\rangle$ and $|V_{n2}\rangle$ as a linear combination of $|u_{n1}\rangle$ and $|u_{n2}\rangle$.

i.e.,
$$|V_{n1}\rangle = a_{11} |u_{n1}\rangle + a_{21} |u_{n2}\rangle \tag{13.31}$$

$$|V_{n2}\rangle = a_{12} |u_{n1}\rangle + a_{22} |u_{n2}\rangle \tag{13.32}$$

Note that we have not introduced any change in the nature of the system for now also the energy level E_n is two-fold degenerate with two linearly independent eigenstates. At present, the coefficients a_{11}, a_{12}, a_{21} and a_{22} in the linear combinations in (13.31) and (13.32) are arbitrary. We can choose them at our convenience a little later.

$$\hat{H}' |V_{n1}\rangle = E_{n1}^{(1)} |V_{n1}\rangle \tag{13.33}$$

$$\hat{H}' |V_{n2}\rangle = E_{n2}^{(1)} |V_{n2}\rangle \tag{13.34}$$

Let us take the scalar product of the Equation (13.33) with $|u_{n1}\rangle$. Substituting (13.31) in (13.33), we get

$$\begin{aligned}
\langle u_{n1}| \hat{H}' |V_{n1}\rangle &= \langle u_{n1}| \hat{H}'(a_{11} |u_{n1}\rangle + a_{21} |u_{n2}\rangle) \\
&= a_{11}\langle u_{n1}| \hat{H}' |u_{n1}\rangle + a_{21}\langle u_{n1}| \hat{H}' |u_{n2}\rangle \\
\langle u_{n1}| E_{n1}^{(1)} |V_{n1}\rangle &= E_{n1}^{(1)} \left\{ \underbrace{\langle u_{n1}| a_{11} |u_{n1}\rangle}_{a_{11}} + \underbrace{\langle u_{n1}| a_{21} |u_{n2}\rangle}_{0} \right\}
\end{aligned}$$

$$\therefore \; a_{11}\langle u_{n1}| H' |u_{n1}\rangle + a_{21}\langle u_{n1}| H' |u_{n2}\rangle = E_{n1}^{(1)} a_{11} \tag{13.35}$$

Similarly, taking the scalar product of the Equation (13.34) with $|u_{n2}\rangle$, we get

$$a_{12}\langle u_{n2}|\,H'\,|u_{n1}\rangle + a_{22}\langle u_{n2}|\,H'\,|u_{n2}\rangle = E_{n1}^{(1)}a_{22} \tag{13.36}$$

The Equations (13.35) and (13.36) can be written as

$$a_{11}H'_{11} + a_{21}H'_{12} = E_{n1}^{(1)}a_{11}$$
$$a_{12}H'_{21} + a_{22}H'_{22} = E_{n1}^{(1)}a_{21}$$

Or
$$\begin{bmatrix} H'_{11} & H'_{12} \\ H'_{21} & H'_{22} \end{bmatrix}\begin{bmatrix} a_{11} \\ a_{21} \end{bmatrix} = E_{n1}^{(1)}\begin{bmatrix} a_{11} \\ a_{21} \end{bmatrix} \tag{13.37}$$

Similarly, from the Equation (13.34), we get

$$\begin{bmatrix} H'_{11} & H'_{12} \\ H'_{21} & H'_{22} \end{bmatrix}\begin{bmatrix} a_{12} \\ a_{22} \end{bmatrix} = E_{n2}^{(1)}\begin{bmatrix} a_{12} \\ a_{22} \end{bmatrix} \tag{13.38}$$

These equations are eigenvalue equations of the matrix $\begin{bmatrix} H'_{11} & H'_{12} \\ H'_{21} & H'_{22} \end{bmatrix}$. The eigenvalues $E_{n1}^{(1)}$ and $E_{n2}^{(1)}$ have to be determined. So both the Equations (13.37) and (13.38) can be put together by a single eigenvalue equation.

$$\begin{bmatrix} H'_{11} & H'_{12} \\ H'_{21} & H'_{22} \end{bmatrix}\begin{bmatrix} a_{1} \\ a_{2} \end{bmatrix} = E_{n}^{(1)}\begin{bmatrix} a_{1} \\ a_{2} \end{bmatrix} \tag{13.39}$$

Or
$$\begin{vmatrix} H'_{11} - E_n^{(1)} & H'_{12} \\ H'_{21} & H'_{22} - E_n^{(1)} \end{vmatrix} = 0 \tag{13.40}$$

This is the well-known secular equation. This can be extended to *m*-fold degenerate levels. For *m*-fold degenerate states,

$$\begin{vmatrix} H'_{11} - E_n^{(1)} & H'_{12} & H'_{13} & .. & .. \\ H'_{21} & H'_{22} - E_n^{(1)} & H'_{23} & .. & .. \\ H'_{31} & H'_{32} & H'_{33} - E_n^{(1)} & .. & .. \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{vmatrix} = 0 \tag{13.41}$$

13.3 STARK EFFECT FOR HYDROGEN ATOM

The Stark effect is the splitting of the spectroscopic lines or splitting the energy levels in the presence of a constant electric field. We consider here the hydrogen atom in the presence of electric field $\boldsymbol{\varepsilon}$ along *z* axis.

i.e.,
$$\boldsymbol{\varepsilon} = \varepsilon\,\hat{\mathbf{z}} \tag{13.42}$$

The dipole moment of an electron in a hydrogen atom is given by

$$\mathbf{d} = -e\,\mathbf{r}$$

The potential energy of a dipole due to its interaction with electric field ε is

$$V = -\mathbf{d}\cdot\varepsilon = e\mathbf{r}\cdot\varepsilon = e\,\mathbf{r}\cdot\varepsilon\,\hat{\mathbf{z}} = e\,\varepsilon\,z$$

The Hamiltonian for the hydrogen atom is now given by

$$H = \underbrace{\frac{p^2}{2m} + V(r)}_{H_0} + \underbrace{e\,\varepsilon\,z}_{H'} \tag{13.43}$$

First-order Correction to the Ground State

The ground state of hydrogen atom is a non-degenerate state. So we can apply first-order non-degenerate perturbation theory.

The ground-state energy E_1 is

$$E_1 = -\frac{\mu Z^2 e^4}{2\hbar^2}$$

i.e.,

$$H_0\phi_{100} = -\frac{\mu Z^2 e^4}{2\hbar^2}\phi_{100}$$

where u_{100} is the ground-state wave function corresponding to $n = 1$, $l = 0$, and $m_l = 0$.

The first-order correction is

$$\Delta E_1^{(1)} = \int \phi_{100}^{*} e\,\varepsilon\,z\,\phi_{100}\,d^3\mathbf{r}$$

The product $u_{100}^{*}\,u_{100}$ has even parity and the factor $e\varepsilon z$ has odd parity, and so the integral is zero.

$$\therefore \Delta E_1^{(1)} = 0 \tag{13.44}$$

Therefore, the first-order correction to the ground-state energy due to Stark effect is zero.

Second-order Correction to the Ground State

The second-order correction to the ground state is

$$\Delta E_1^{(2)} = \sum_{nlm\neq 100} \frac{\left|\langle \phi_{nlm}|\,H'\,|\phi_{100}\rangle\right|^2}{E_1^{(0)} - E_n^{(0)}}$$

This is an infinite series. There are special methods to evaluate this series. For instance, we can use Dalgarno and Lewis method (see Mathews and Venkatesan), and it is found to be

$$\Delta E_1^{(2)} = -\frac{9}{2}a^2\varepsilon_0^2 \tag{13.45}$$

Correction to *n* = 2 Energy Level

This is one of the simplest examples to illustrate degenerate perturbation theory. There are four eigenstates of H_0 corresponding $n = 2$. They are $|n=2, l=1, m_l=1\rangle$, $|210\rangle$, $|21-1\rangle$ and $|200\rangle$. All these four states have the same energy $E_2^{(0)} = -\mu Z^2e^4/8\hbar^2$. The perturbing Hamiltonian is $H' = e\varepsilon z$. The matrix elements are $\langle n'l'm'|\, H'\, |nlm\rangle$.

The wave function ψ_{nlm} is given by

$$\psi_{nlm} = a_1u_{211} + a_2u_{210} + a_3u_{21-1} + a_4u_{200}$$

The eigenvalues of the matrix H' are the first-order correction to the energy, and the corresponding eigenvectors give the coefficients (a_1, a_2, a_3, a_4). Consider the general matrix element $\langle n'l'm'|\, H'\, |nlm\rangle$. Since the wave function for the hydrogen atom is given by $u_{nlm}(r,\theta,\phi) = R_{nl}(r)Y_{lm}(\theta,\phi)$, we have

$$e\varepsilon\langle n'l'm' \,|\, z \,|\, nlm\rangle = e\varepsilon\int_0^\infty R^*_{n'l'}(r)R_{nl}(r)r^3dr\cdot\int d\Omega Y_{l'm'}\cos\theta Y_{lm} \tag{13.46}$$

The integrand in the angular part is given by

$$Y^*_{l'm'}\cos\theta\ \mathrm{Y}_{lm}$$

We know that

$$\cos\theta = \left(\frac{4\pi}{3}\right)^{1/2} Y_{10}$$

$$\therefore Y^*_{l'm'}\cos\theta\ Y_{lm} = Y^*_{l'm'}\left(\frac{4\pi}{3}\right)^{1/2} Y_{10}Y_{lm}$$

The parity of this function is $(-1)^{l'+l}\cdot(-1)$. If l' and l are the same, the integrand has odd parity and hence the integral is zero.

$$\therefore \int d\Omega Y^*_{l'm'}\cos\theta Y_{lm} = 0 \text{ if } l' = l$$

In order words
$$\int d\Omega Y^*_{1m'}\cos\theta Y_{1m} = 0 \quad \text{since} \quad l' = l = 1$$

and
$$\int d\Omega Y^*_{00}\cos\theta Y_{00} = 0 \quad \text{since} \quad l' = l = 0$$

This makes all the elements in the diagonal blocks H' (nine elements in the 3×3 block and one element in the 1×1 block) zero. So we have to evaluate two groups of elements (six elements):

$$\int d\Omega Y_{1m}^{*} \cos\theta Y_{00} = \left(\frac{1}{4\pi}\right)^{1/2} . \left(\frac{4\pi}{3}\right)^{1/2} \int d\Omega Y_{1m}^{*} Y_{10} = 0 \quad \text{if } m \neq 0$$

Therefore, there are only two non-zero elements, namely $e\varepsilon\langle 210 \,|\, z \,|\, 200\rangle$ and $e\varepsilon\langle 200 \,|\, z \,|\, 210\rangle$. We have

$$R_{21} = \frac{1}{\sqrt{24}}\left(\frac{Z}{a}\right)^{5/2} r\,\mathrm{e}^{-Zr/2a}$$

and

$$R_{20} = \frac{1}{\sqrt{2}}\left(\frac{Z}{a}\right)^{3/2}\left(1-\frac{Zr}{2a}\right)e^{-Zr/2a}$$

$$\begin{aligned}
\therefore e\varepsilon\langle 210|z|200\rangle &= e\varepsilon \cdot \frac{1}{\sqrt{2}} \cdot \frac{1}{\sqrt{24}} \cdot \left(\frac{3}{4\pi}\right)^{1/2} \left(\frac{Z}{a}\right)^{4} \\
&\quad \times \int_0^\infty \left(1-\frac{Zr}{2a}\right) r^2 e^{-Zr/2a} e^{-Zr/2a} dr \int_0^\pi \cos^2\theta \sin\theta d\theta \int_0^{2\pi} d\phi \\
&= \frac{e\varepsilon}{12\pi} \cdot \frac{1}{a^4} \int_0^\infty \left(1-\frac{r}{2a}\right) r^4 e^{-r/a} \\
&= -3e\varepsilon a
\end{aligned}$$

We have used the fact $\int_0^\infty r^n e^{-r/a} = n!\,a^{n+1}$. Note that we have taken Z as 1 for hydrogen atom. All the matrix elements are presented in Table 13.1.

Table 13.1 Matrix H': $\langle n'l'm'|\, H'|\, nlm\rangle$

	$\vert 211\rangle$	$\vert 210\rangle$	$\vert 21-1\rangle$	$\vert 200\rangle$
$\langle 211\vert$	0	0	0	0
$\langle 210\vert$	0	0	0	$-3ea\varepsilon$
$\langle 21-1\vert$	0	0	0	
$\langle 200\vert$	0	$-3ea\varepsilon$	0	0

The secular equation is

$$\begin{vmatrix} -\lambda & 0 & 0 & 0 \\ 0 & -\lambda & 0 & -3e\varepsilon a \\ 0 & 0 & -\lambda & 0 \\ 0 & -3ea\varepsilon & 0 & -\lambda \end{vmatrix} = 0$$

Or

$$\lambda = 0,\ \lambda = 0,\ \lambda = \pm\, 3e\varepsilon a.$$

The eigenvectors are determined by the equation

$$\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -3e\varepsilon a \\ 0 & 0 & 0 & 0 \\ 0 & -3e\varepsilon a & 0 & 0 \end{bmatrix}\begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{bmatrix} = \pm 3e\varepsilon a \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{bmatrix}$$

The eigenvectors are given below:

$$\lambda = 3ea\varepsilon : \begin{bmatrix} 0 \\ 1 \\ 0 \\ -1 \end{bmatrix}; \lambda = -3ea\varepsilon : \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix}$$

The correction to the energy eigenvalue and the corresponding wave functions is given below:

$$\Delta E_2^{(1)} = 0 \qquad \psi_{211} = u_{211}$$

$$\Delta E_2^{(1)} = 3ea\varepsilon \qquad \psi_{210} = \frac{1}{\sqrt{2}}[u_{210} - u_{200}]$$

$$\Delta E_2^{(1)} = 0 \qquad \psi_{21-1} = u_{21-1}$$

$$\Delta E_2^{(1)} = -3ea\varepsilon \qquad \psi_{200} = \frac{1}{\sqrt{2}}[u_{210} + u_{200}]$$

13.4 HYDROGEN ATOM – RELATIVISTIC EFFECTS AND *LS* COUPLING

Hydrogen atom is one of the best examples to illustrate the application of perturbation theory. We have already seen Stark effect in the case of hydrogen atom. In this section, we will see some corrections that are intrinsic to hydrogen atom. We will also describe hydrogen in an external magnetic field.

The simple Hamiltonian for hydrogen takes into account the electrostatic interaction between the electron and the nucleus. Number of corrections are required to describe hydrogen in a more realistic way. The major corrections are due to the H_r describing the relativistic effects and H_{LS} describing the spin-orbit coupling.

So we have

$$H_0 = \frac{p^2}{2m} - \frac{Ze^2}{r}$$

If we include the relativistic correction and the spin-orbit coupling,

$$H = H_0 + \underbrace{H_r + H_{LS}}_{\text{perturbations to } H_0} \tag{13.47}$$

$$E_n = E_n^{(0)} + \Delta E_n$$

where $$\Delta E_n = \Delta E_{nr} + \Delta E_{nLS}$$

In what follows, though we use the symbol m for the mass, it should be taken as representing reduced mass. Let us also recall the following expressions:

The fine structure constant is $\alpha = e^2/\hbar c$ and the Bohr radius is $a = \hbar^2/me^2$.

In the following problems, we use two groups of commuting operators (i) H_0, L^2, S^2, L_Z and S_Z and (ii) H_0, L^2, S^2, J^2 and J_Z. The corresponding eigenstates are $|nlm_l m_s\rangle$ and $|nljm_j\rangle$. So we have

$$\begin{aligned} H_0\,|nlm_l m_s\rangle &= E_n^{(0)}\,|nlm_l m_s\rangle \\ L^2\,|nlm_l m_s\rangle &= l(l+1)\hbar^2\,|nlm_l m_s\rangle \\ S^2\,|nlm_l m_s\rangle &= s(s+1)\hbar^2\,|nlm_l m_s\rangle \\ L_z\,|nlm_l m_s\rangle &= m_l\hbar\,|nlm_l m_s\rangle \\ S_z\,|nlm_l m_s\rangle &= m_s\hbar\,|nlm_l m_s\rangle \end{aligned}$$

Similarly, we have

$$\begin{aligned} H_0\,|nljm_j\rangle &= E_n^{(0)}\,|nljm_j\rangle \\ L^2\,|nljm_j\rangle &= l(l+1)\hbar^2\,|nljm_j\rangle \\ S^2\,|nljm_j\rangle &= s(s+1)\hbar^2\,|nljm_j\rangle \\ J^2\,|nljm_j\rangle &= j(j+1)\hbar^2\,|nljm_j\rangle \\ J_z\,|nljm_j\rangle &= m_j\hbar\,|nljm_j\rangle \end{aligned}$$

Depending on the choice of Hamiltonian, we can employ either of the two sets of eigenstates.

13.4.1 Relativistic Correction

The kinetic energy of an electron in the non-relativistic classical physics is

$$T = \frac{p^2}{2m}$$

In relativistic classical physics, we have

$$E = [m^2c^4 + p^2c^2]^{1/2}$$

The non-relativistic limit of this expression is

$$\begin{aligned} E &= mc^2\left[1 + \frac{p^2c^2}{m^2c^4}\right]^{1/2} \\ &= mc^2\left[1 + \frac{p^2}{2m^2c^2} - \frac{1}{8}\frac{p^4}{m^4c^4}\cdots\right] \end{aligned}$$

Therefore, the kinetic energy is

$$T = E - mc^2 \simeq \frac{p^2}{2m} - \frac{1}{2mc^2}\left(\frac{p^2}{2m}\right)^2 \qquad (13.48)$$

Making a transition to quantum mechanics, we have

$$T_{op} = \frac{p_{op}^2}{2m} - \frac{1}{2mc^2}\left(\frac{p_{op}^2}{2m}\right)^2$$

So the Hamilton becomes

$$H = \underbrace{\frac{p_{op}^2}{2m} - \frac{Ze^2}{r}}_{H_0} - \underbrace{\frac{1}{2mc^2}\left(\frac{p_{op}^2}{2m}\right)^2}_{H_r} \tag{13.49}$$

$$= H_0 + H_r$$

where

$$H_r = -\frac{1}{2mc^2}\left(\frac{p_{op}^2}{2m}\right)^2 \tag{13.50}$$

For hydrogen atom, let us write $\frac{p_{op}^2}{2m}$ as

$$\frac{p_{op}^2}{2m} = H_0 + \frac{Ze^2}{r}$$

$$\therefore H_r = -\frac{1}{2mc^2}\left(H_0 + \frac{Ze^2}{r}\right)^2$$

H_r is independent of spin or orbital angular momentum, and so H_r is diagonal in both representations. Let us choose the states $|nljm_j\rangle$ as the basis states since they can be used in the case of spin-orbit correction also. Therefore, ΔE_{nr} is given by

$$\begin{aligned}\Delta E_{nr} &= \langle nljm_j| -\frac{1}{2mc^2}\left(H_0 + \frac{Ze^2}{r}\right)\left(H_0 + \frac{Ze^2}{r}\right)|nljm_j\rangle \\ &= -\frac{1}{2mc^2}\langle nljm_j|\left(E_n^{(0)} + \frac{Ze^2}{r}\right)\left(E_n^{(0)} + \frac{Ze^2}{r}\right)|nljm_j\rangle \\ &= -\frac{1}{2mc^2}\langle nl|\,(E_n^{(0)})^2 + 2E_n^{(0)}\frac{Ze^2}{r} + \frac{Z^2e^4}{r^2}\,|nl\rangle \\ &= -\frac{1}{2mc^2}E_n^{(0)}\left[E_n^{(0)} + 2Ze^2\left\langle\frac{1}{r}\right\rangle_{nl}\right] - \frac{1}{2mc^2}Z^2e^4\left\langle\frac{1}{r^2}\right\rangle_{nl}\end{aligned}$$

Let us make use of the Equation (8.91) for E_n^0 and Examples 8.12 and 8.13 of Chapter 8 for $\left\langle\frac{1}{r}\right\rangle_{nl}$ and $\left\langle\frac{1}{r^2}\right\rangle_{nl}$

$$\left\langle\frac{1}{r}\right\rangle_{nl} = \frac{Z}{n^2a} \quad \text{and} \quad \left\langle\frac{1}{r^2}\right\rangle_{nl} = \frac{Z^2}{n^3a^2\left(l+\frac{1}{2}\right)}$$

Then, we have

$$= -\frac{1}{2mc^2}E_n^{(0)}\left[-\frac{1}{2}\frac{mc^2Z^2\alpha^2}{n^2} + 2Ze^2\frac{Z}{an^2}\right] - \frac{1}{2mc^2}\cdot Z^2e^4\cdot\frac{Z^2}{a^2n^3\left(l+\frac{1}{2}\right)}$$

$$= -\frac{E_n^{(0)}Z^2\alpha^2}{n^2}\left[-\frac{n}{l+\frac{1}{2}} + \frac{3}{4}\right] \tag{13.51}$$

13.4.2 Spin-orbit Interaction

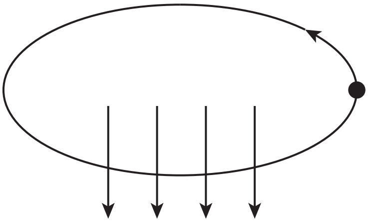


Fig. 13.1 Internal magnetic field due to orbital motion of electron

The orbital motion of an electron produces a magnetic field (Fig. 13.1). This is an internal magnetic field. There is an interaction between internal magnetic field due to orbital motion and magnetic moment due to spin angular momentum. This interaction is known as spin-orbit interaction.

The corresponding Hamiltonian is found to be

$$H_{LS} = \frac{1}{2m^2c^2}\frac{1}{r}\frac{dV}{dr}\mathbf{L}\cdot\mathbf{S} = \frac{Ze^2}{2m^2c^2}\frac{1}{r^3}\mathbf{L}\cdot\mathbf{S} \tag{13.52}$$

(The reader can find the derivation of this expression in any standard book on quantum mechanics.) The spin-orbit interaction is an intrinsic property of an electron in an atom.

Since J^2 and J_Z commute with $\mathbf{L}\cdot\mathbf{S}$, we will choose the basis states $|nljm_j\rangle$. In this basis, H_0 as well as H_{LS} are diagonal.

$$\Delta E_{LS} = \langle nljm_j|\frac{Ze^2}{2m^2c^2}\cdot\frac{1}{r^3}\cdot\mathbf{L}\cdot\mathbf{S}|nljm_j\rangle \tag{13.53}$$

$\mathbf{L}\cdot\mathbf{S}$ can now be written as

$$\mathbf{L}\cdot\mathbf{S} = (J^2 - L^2 - S^2)/2$$

$$\begin{aligned}\Delta E_{LS} &= \frac{Ze^2}{2m^2c^2}\langle nljm_j|\frac{1}{r^3}\mathbf{L}\cdot\mathbf{S}|nljm_j\rangle \\ &= \frac{Ze^2}{2m^2c^2}\langle nljm_j|\frac{1}{2}\frac{1}{r^3}(J^2-L^2-S^2)|nljm_j\rangle \\ &= \frac{Ze^2}{4m^2c^2}\left\langle\frac{1}{r^3}\right\rangle_{nl}[j(j+1)-l(l+1)-s(s+1)]\hbar^2 \\ &= \frac{Ze^2\hbar^2}{4m^2c^2}\left\langle\frac{1}{r^3}\right\rangle_{nl}\left[j(j+1)-l(l+1)-\frac{3}{4}\right]\end{aligned}$$

The possible values of j are $j_1 = l + \frac{1}{2}$ and $j_2 = l - \frac{1}{2}$.

$$
\begin{aligned}
\Delta E_{LS} &= \frac{Ze^2\hbar^2}{4m^2c^2}\left\langle \frac{1}{r^3}\right\rangle_{nl}\begin{cases} l & \text{for} \quad j = l + \frac{1}{2} \\ -(l+1) & \text{for} \quad j = l - \frac{1}{2} \end{cases} \\
&= -\frac{Ze^2\hbar^2}{4m^2c^2}\cdot\frac{Z^3}{a^3n^3l\left(l+\frac{1}{2}\right)(l+1)}\begin{cases} l \\ -(l+1) \end{cases} \\
&= -\frac{E_n^{(0)}Z^2\alpha^2}{2nl\left(l+\frac{1}{2}\right)(l+1)}\begin{cases} l \\ -(l+1) \end{cases} \\
&= -\frac{E_n^{(0)}Z^2\alpha^2}{n^2}\begin{cases} \dfrac{n}{2\left(l+\frac{1}{2}\right)(l+1)} & \text{for} \quad j = l + \frac{1}{2} \\ -\dfrac{n}{2l\left(l+\frac{1}{2}\right)} & \text{for} \quad j = l - \frac{1}{2} \end{cases}
\end{aligned}
\tag{13.54}
$$

$\Delta E_{nr} + \Delta E_{nLs}$

The total correction due to H_r and H_{LS} is given by

$$
\begin{aligned}
\Delta E_{nr} + \Delta E_{nLs} &= -\frac{E_n^{(0)}Z^2\alpha^2}{n^2}\begin{cases} \left(-\frac{n}{l+1}+\frac{3}{4}\right) & \text{for} \ \ j = l + \frac{1}{2} \\ \left(-\frac{n}{l}+\frac{3}{4}\right) & \text{for} \ \ j = l - \frac{1}{2} \end{cases} \\
&= -\frac{E_n^{(0)}Z^2\alpha^2}{n^2}\begin{cases} \left(-\dfrac{n}{j+\frac{1}{2}}+\dfrac{3}{4}\right) & \text{for} \ \ j = l + \frac{1}{2} \\ \left(-\dfrac{n}{j+\frac{1}{2}}+\dfrac{3}{4}\right) & \text{for} \ \ j = l - \frac{1}{2} \end{cases} \\
&= -\frac{E_n^{(0)}Z^2\alpha^2}{n^2}\cdot\left[-\frac{n}{\left(j+\frac{1}{2}\right)}+\frac{3}{4}\right]
\end{aligned}
\tag{13.55}
$$

This agrees well with non-relativistic approximation of the solution to the relativistic Dirac equation for hydrogen atom. Note that $E_n^{(0)}$ does not depend on l or m_s and so there is a degeneracy $2n^2$. H_r and H_{LS} together partially lift these degeneracy. However, the energy now depends on j.

Example 13.3 Determine the energy splitting of the hydrogen atom corresponding to $n = 3$ due to H_r and H_{LS}.

Solution: For $n = 3$, values of $l = 2, 1, 0$. For $l = 2$, $j = 5/2$ or $3/2$; for $l = 1$, $j = 3/2$ or $1/2$; for $l = 0$, $j = 1/2$. So any three possible values of j are $j = 5/2, 3/2, 1/s$. So the energy $E_3^{(0)}$ is split into three levels. They are given below in Fig(13.2).

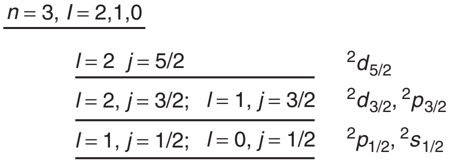


Fig. 13.2 Energy splitting for hydrogen due to H_r and H_{LS}

13.5 HYDROGEN ATOM IN MAGNETIC FIELD

Let us now place the hydrogen atom in a magnetic field B. The Hamiltonian H_B represents the interaction of the electron with the magnetic field, and it is given by

$$H_B = \frac{e}{2mc}(\mathbf{L} + 2\mathbf{S}) \cdot \mathbf{B} \tag{13.56}$$

Now we have three kinds of perturbation: H_r, H_{LS} and H_B. The Hamiltonian H_r and H_{LS} are intrinsic to hydrogen atom, and they are always comparable to each other. Therefore, they always occur as pair. On the other hand, H_B is under our control. So we can have $H_B >> H_r + H_{LS}$ or $H_r + H_{LS} >> H_B$ or intermediate strengths. The total Hamiltonian H is

$$H = H_0 + H_r + H_{LS} + H_B \tag{13.57}$$

Depending on the strength of the external magnetic field B, we have different kinds of approximations. Let us first list the various possible approximations before considering them in detail. We consider here the following extreme limits:

Strong Field Approximation

$$H_0 >> H_B >> H_r + H_{LS} \tag{13.58}$$

The effect due to H_B is so large compared to $H_r + H_{LS}$ that we neglect the effect of these Hamiltonians. As a first approximation, the Hamiltonian becomes

$$\therefore H = H_0 + \underbrace{H_B}_{\text{Perturbation}}$$

The corresponding energy eigenvalue is $E_n = E_n^{(0)} + \Delta E_{nB}$.

Paschen Back Effect

Field B is not large enough to completely ignore $H_r + H_{LS}$, and so their effect has to be included. Let us further assume that the correction due to magnetic field B is still larger compared to the correction due to $H_r + H_{LS}$. So we solve this problem in two stages:

First, we consider

$$H = H_0 + \underbrace{H_B}_{\text{Perturbation}} \tag{13.59}$$

$$E_n = E_n^{(0)} + \Delta E_{nB} \tag{13.60}$$

Next, we take the term $H_r + H_{LS}$. as perturbation to this Hamiltonian.

$$H = H_0 + H_B + \underbrace{H_r + H_{LS}}_{\text{Perturbation to } (H_0 + H_B)} \tag{13.61}$$

$$\therefore E_n = (E_n^{(0)} + \Delta E_{nB}) + (\Delta E_{nr} + \Delta E_{nLS}) \tag{13.62}$$

Weak Magnetic Field-anomalous Zeeman Effect

In this case, we consider a different situation where the effects of magnetic field B are small compared to $H_r + H_{LS}$. Again, we solve the problems in two stages:

$$H_0 >> H_r + H_{LS} >> H_B \tag{13.63}$$

First, we consider

$$H = H_0 + \underbrace{H_r + H_{LS}}_{\text{Perturbation to } H_0} \tag{13.64}$$

$$E_n = E_n^{(0)} + \Delta E_{nr} + \Delta E_{nLS} \tag{13.65}$$

Next, we take H_B as further perturbation.

$$H = (H_0 + H_r + H_s) + \underbrace{H_B}_{\text{Perturbation to} (H_0 + H_r + H_s)}$$

$$E_n = (E_n^{(0)} + \Delta E_{nr} + \Delta E_{nLS}) + \Delta E_{nB}$$

13.5.1 Strong Zeeman Effect

As mentioned above, the correction due to H_B is so large compared to the corrections raising from $(H_{LS} + H_r)$ that it can be neglected.

$$H = H_o + H_B \tag{13.66}$$

Since $H_B = \frac{eB}{2mc}(L_s + 2S_z)$, it commutes with L^2, S^2, L_Z, and S_Z.

i.e., $$[H_B, L_z] = 0 \; [H_B, S_z] = 0$$

Therefore, we have to choose the states $|nlm_l m_s\rangle$ as the basis states.

$$\frac{\mu_B B}{\hbar}(L_Z + 2S_Z)\,|nlm_l m_s\rangle = \frac{\mu_B B\hbar}{\hbar}[m_l + 2m_s]\,|nlm_l m_s\rangle \tag{13.67}$$

The Hamiltonian H' is a diagonal matrix.

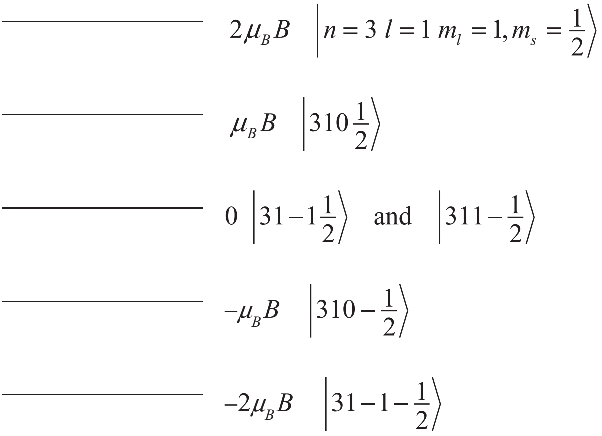
i.e., $$\langle n'l'm_l'm_s'|\,H'\,|nlm_l m_s\rangle = \mu_B B[m_l + 2m_s]\delta_{nn'}\delta_{ll'}\delta_{m_l m_l'}\delta_{m_s m_s'}$$

$$\therefore \Delta E_{nlm_l m_s} = (m_l + 2m_s)\mu_B B \tag{13.68}$$

The Hamiltonian H' does not lift the degeneracy completely. There are non-degenerate as well as degenerate states.

Example 13.4 Enumerate the eigenstates of $H_0 + H_B$ for a hydrogen-like atom for $n = 3$, $l = 1$.

Solution: The eigenstates of H_o are $(2l + 1)(2s + 1)$ or six-fold degenerate for $n = 3$, $l = 1$.



The matrix H' is

	$\left\|1\frac{1}{2}\right\rangle$	$\left\|0\frac{1}{2}\right\rangle$	$\left\|-1\frac{1}{2}\right\rangle$	$\left\|1-\frac{1}{2}\right\rangle$	$\left\|0-\frac{1}{2}\right\rangle$	$\left\|-1-\frac{1}{2}\right\rangle$
$\left\langle 1\frac{1}{2}\right\|$	$2\mu_B B$	0	0	0	0	0
$\left\langle 0\frac{1}{2}\right\|$	0	$\mu_B B$	0	0	0	0
$\left\langle -1\frac{1}{2}\right\|$	0	0	0	0	0	0
$\left\langle 1-\frac{1}{2}\right\|$	0	0	0	0	0	0
$\left\langle 0-\frac{1}{2}\right\|$	0	0	0	0	$-\mu_B B$	0
$\left\langle -1-\frac{1}{2}\right\|$	0	0	0	0	0	$-2\mu_B B$

13.5.2 Paschen-back Effect

The magnetic field B is such that the correction due to H_{LS} cannot be neglected and it is very small. So the Hamiltonian H_{LS} is considered as perturbation to $H_0 + H_B$.

$$\therefore H = \underbrace{H_0 + H_B}_{\text{Unperturbed}} + \underbrace{H_{LS} + H_r}_{\text{Perturbation}} \tag{13.69}$$

$$\begin{aligned}\Delta E_{LS} &= \langle nlm_l m_s |\, \frac{1}{2m^2c^2}\frac{1}{r^3}\cdot Ze^2 \mathbf{L}\cdot\mathbf{S}\, | nlm_l m_s\rangle \\ &= \frac{Ze^2}{2m^2c^2}\left\langle\frac{1}{r^3}\right\rangle_{nl}\cdot\langle m_l m_s |\, \mathbf{L}\cdot\mathbf{S}\, | m_l m_s\rangle\end{aligned}$$

$\mathbf{L}\cdot\mathbf{S}$ can be written as

$$\begin{aligned}\mathbf{L}\cdot\mathbf{S} &= L_z S_z + L_x S_x + L_y S_y \\ &= L_z S_z + \frac{1}{4}(L_+ + L_-)(S_+ + S_-) - \frac{1}{4}(L_+ - L_-)(S_+ - S_-) \\ &= L_z S_z + \frac{1}{2}(L_+ S_- + L_- S_+)\end{aligned}$$

$$\Delta E_{LS} = \frac{Ze^2}{2m^2c^2}\left\langle\frac{1}{r^3}\right\rangle_{nl}\left\{\langle m_l m_s |\, L_Z S_Z\, | m_l m_s\rangle + \langle m_l m_s |\, \frac{1}{2}(L_+ S_- + L_- S_+)\, | m_l m_s\rangle\right\}$$

The second term is zero. Using the Kramer's recursion relation (8.111), $\left\langle \frac{1}{r^3} \right\rangle_{nl}$ can be calculated, and it is found to be

$$\left\langle \frac{1}{r^3} \right\rangle_{nl} = \frac{Z^3}{n^3 a^3 l(l+1)\left(l+\frac{1}{2}\right)}$$

Making use of this result, we get

$$\therefore \Delta E_{LS} = \frac{Ze^2}{2m^2c^2} \cdot \frac{Z^3}{n^3 a^3 l(l+1)\left(l+\frac{1}{2}\right)} \cdot m_l m_s \hbar^2 \tag{13.70}$$

Let us recall $a = \frac{\hbar^2}{me^2}, \alpha = \frac{e^2}{\hbar c}$ and $E_n^{(0)} = -\frac{1}{2} mc^2 \cdot \frac{Z^2\alpha^2}{n^2}$

Then we have

$$\Delta E_{LS} = -\frac{\alpha^2 Z^2}{n} E_n^{(0)} \frac{m_l m_s}{l(l+1)\left(l+\frac{1}{2}\right)} \tag{13.71}$$

The energy E_n is given by

$$E_n = E_n^{(0)} + (m_l + 2m_s)\mu_B B + \Delta E_{LS} + \Delta E_r$$

ΔE_r has already been determined in Section 13.4.1. Note that whether we use $|nljm_j\rangle$ or $|nlm_l m_s\rangle$ as the basis states, ΔE_r is the same. So we can take the same expression as given in (13.51) for ΔE_r. Therefore, we have

$$E_n = E_n^{(0)} + (m_l + 2m_s)\mu_B B - \frac{\alpha^2 Z^2}{n} \cdot \frac{E_n^{(0)} m_l m_s}{l(l+1)\left(l+\frac{1}{2}\right)} - \frac{E_n^{(0)} Z^2 \alpha^2}{n^2}\left[-\frac{n}{l+\frac{1}{2}} + \frac{3}{4}\right] \tag{13.72}$$

Example 13.5 Determine the energy eigenvalues and the corresponding eigenfunctions in Paschen-Back effect for 3p levels of hydrogen.

Solution: There are six states. They are $\left|n=3, l=1, m_l=1, m_s=\frac{1}{2}\right\rangle, \left|310-\frac{1}{2}\right\rangle, \left|31-1\frac{1}{2}\right\rangle,$ $\left|311-\frac{1}{2}\right\rangle, \left|300-\frac{1}{2}\right\rangle$ and $\left|31-1-\frac{1}{2}\right\rangle$. All these states have same $E_3^{(0)}$. They get correction due to H_B and H_{LS} and H_r. They are shown in Fig.13.3 (Correction due to H_r is not shown.)

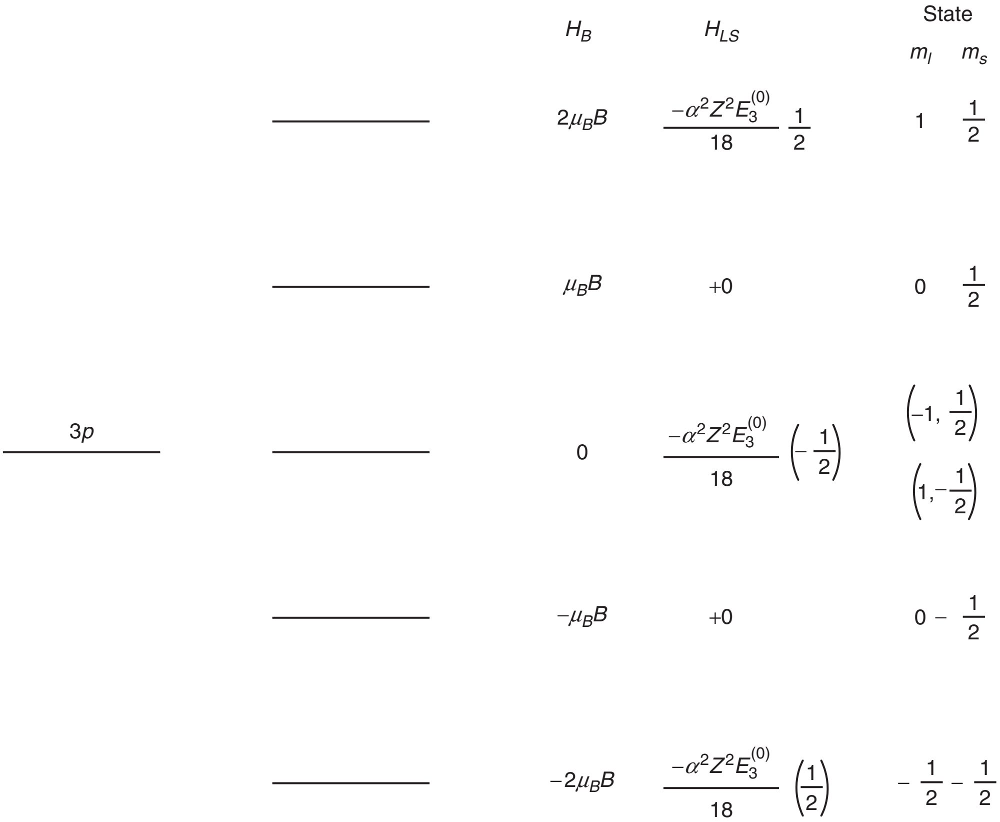


Fig. 13.3 Energy splitting of 3*p* level

13.5.3 Anomalous Zeeman Effect

Now we have
$$H_o >> H_r + H_{LS} >> H_B \tag{13.73}$$

In this case, the coupled representation is better choice; i.e., $|nljm_j\rangle$ are now more suitable basis. Both H_{LS} and H_B are diagonal in this representation.

The Hamiltonian H_B is

$$H_B = \frac{e}{2mc}(\mathbf{L} + 2\mathbf{S}) \cdot \mathbf{B}$$

The total Hamiltonian H can be now seen as

$$H = \underbrace{H_0 + H_{L \cdot S} + H_r +}_{\text{Unperturbed}} \overbrace{H_B}^{\text{Perturbation}} \tag{13.74}$$

Since J^2 and J_Z commute with H_B, $|\,nljmj\rangle$ serves as the basis states for H_B also. It can be shown that

$$\frac{\mu_B}{\hbar}(\mathbf{J}+\mathbf{S})\cdot\mathbf{B}\,|\,jm\rangle = g_J m_J \mu_B B\,|\,jm\rangle \tag{13.75}$$

where g_J is Lande's g factor (which is discussed below).

$$\therefore \Delta E_B = \langle nljm|\,\frac{\mu_B}{\hbar}(\mathbf{J}+\mathbf{S})\cdot\mathbf{B}\,|nljm\rangle$$

$$= g_J m_J \mu_B B \tag{13.76}$$

$$\therefore E_n = E_n^{(0)} + (\Delta E_{LS} + \Delta H_r) + \Delta E_B$$

The energy splitting is shown in Fig. 13.4.

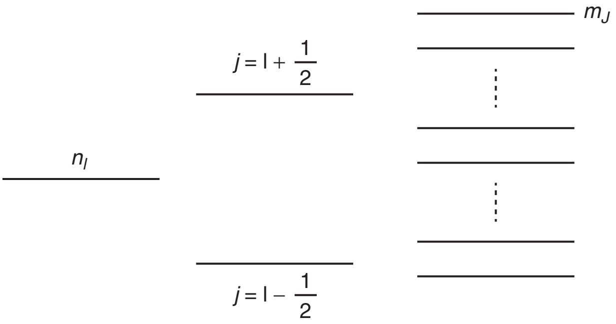


Fig. 13.4 Energy splitting due to anomalous zeeman effect

13.5.4 Lande's *g* Factor

The magnetic moments of an electron due to orbital and spin angular moment are given by

$$\mu_L = -\frac{e}{2mc}\mathbf{L} \quad \text{and} \quad \mu_s = -\frac{e}{2mc}g_s\mathbf{S} \tag{13.77}$$

The total magnetic moment is

$$\mu = \mu_L + \mu_s = -\frac{e}{2mc}(\mathbf{L}+2\mathbf{S}) = -\frac{e}{2mc}(\mathbf{J}+\mathbf{S}) \tag{13.78}$$

Note that $\boldsymbol{\mu}_L$ is in the direction of **L** and $\boldsymbol{\mu}_s$ is in the direction of **S** but $\boldsymbol{\mu}$ is not in the direction of **J**. So, in general, $\boldsymbol{\mu}$ cannot be written as $\boldsymbol{\mu} = -\frac{e}{2mc} g_J \mathbf{J}$. However, with respect to states $|jm\rangle$, we can write

$$\boldsymbol{\mu}|\,jm\rangle = -\frac{e}{2mc} g_J\,\mathbf{J} = -\frac{\mu_B}{\hbar} g_J \mathbf{J}\,|jm\rangle \tag{13.79}$$

This can be established as follows. We closely follow here Davydov:

Let us first write $\boldsymbol{\mu}\,|jm\rangle$ as

$$\boldsymbol{\mu}\,|\,jm\rangle = \hat{G}\mathbf{J}\,|\,jm\rangle \tag{13.80}$$

where $\hat{G}$ is an operator which has to be determined.

$$\begin{aligned}\hat{G}\mathbf{J}\cdot\mathbf{J} &= \boldsymbol{\mu}\cdot\mathbf{J} = -\frac{\mu_B}{\hbar}(\mathbf{L}+2\mathbf{S})\cdot\mathbf{J} = -\frac{\mu_B}{\hbar}(\mathbf{L}+\mathbf{S}+\mathbf{S})\cdot\mathbf{J}\\ &= -\frac{\mu_B}{\hbar}(J^2+\mathbf{S}\cdot\mathbf{J})\end{aligned}$$

Now, $\mathbf{J} = \mathbf{L}+\mathbf{S}$ or $\mathbf{J}-\mathbf{S} = \mathbf{L}$

$$J^2 + S^2 - 2\mathbf{J}\cdot\mathbf{S} = L^2$$

$$\therefore\ \mathbf{J}\cdot\mathbf{S} = \frac{J^2+S^2-L^2}{2}$$

$$\hat{G}\mathbf{J}\cdot\mathbf{J}\,|\,jm\rangle = --\frac{\mu_B}{\hbar}\left[J^2+\frac{J^2+S^2-L^2}{2}\right]|\,jm\rangle$$

$$j(j+1)\hbar^2\hat{G}\,|\,jm\rangle = -\frac{\mu_B}{\hbar}\left[\,j(j+1)+\frac{j(j+1)+s(s+1)-l(l+1)}{2}\right]\hbar^2\,|\,jm\rangle$$

$$\begin{aligned}\therefore\ \hat{G}\,|\,jm\rangle &= -\frac{\mu_B}{\hbar}\left[1+\frac{j(j+1)+s(s+1)-l(l+1)}{2j(j+1)}\right]|\,jm\rangle\\ &= -\frac{\mu_B}{\hbar} g_J\,|\,jm\rangle\end{aligned}$$

$$\begin{aligned}\therefore\ \boldsymbol{\mu}\cdot\mathbf{B}\,|\,jm\rangle &= \hat{G}\mathbf{J}\cdot\mathbf{B}\,|jm\rangle\\ &= \hat{G}BJ_Z\,|jm\rangle = \hat{G}Bm_J\hbar\,|jm\rangle = m_J\hbar B\hat{G}\,|jm\rangle\\ &= -m_J\hbar B\frac{\mu_B}{\hbar}g_J\,|jm\rangle\end{aligned} \tag{13.81}$$

$$\langle jm|-\boldsymbol{\mu}\cdot\mathbf{B}\,|\,jm\rangle = m_J\mu_B Bg_J \tag{13.82}$$

The same result can be obtained using Wigner–Eckart theorem, which again indicates $\boldsymbol{\mu}$ can be written as $\boldsymbol{\mu} = -\mu_B/\hbar\, g_J\mathbf{J}$ provided we consider the matrix element between states $|\,jm\rangle$.

13.6 GROUND STATE OF HELIUM—FIRST-ORDER PERTURBATION THEORY

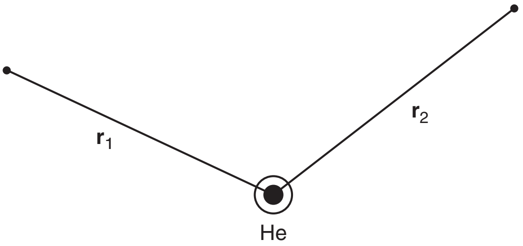


Fig. 13.5 Helium atoms

As an exercise in perturbation theory, let us now determine the ground-state energy of helium atom.

The Hamiltonian H for the helium atom is

$$H = \underbrace{\frac{p_1^2}{2m} - \frac{Ze^2}{r_1}}_{H_1} + \underbrace{\frac{p_2^2}{2m} - \frac{Ze^2}{r_2}}_{H_2} + \underbrace{\frac{e^2}{|\mathbf{r}_1 - \mathbf{r}_2|}}_{H'} \tag{13.83}$$

$$H = H_1 + H_2 + H'$$

Let us take $H_0 = H_1 + H_2$ as the unperturbed Hamiltonian and H' is the interaction Hamiltonian.

$$H_0 = H_1 + H_2 \tag{13.84}$$

The ground-state H_0 is $\psi^{(0)}(\mathbf{r}_1, \mathbf{r}_2)$ and is given by

$$\psi^{(0)}(\mathbf{r}_1, \mathbf{r}_2) = u_{100}(\mathbf{r}_1)u_{100}(\mathbf{r}_2) \tag{13.85}$$

where u_{100} is the ground state of hydrogen atom.

$$\left[-\frac{\hbar^2}{2m}\nabla^2 - \frac{Ze^2}{r}\right] = u_{100}(\mathbf{r}) = -\frac{1}{2}mc^2Z^2\alpha^2u_{100} \tag{13.86}$$

where $$u_{100} = (\mathbf{r}) = R_{10}(r)Y_{100}(\theta,\phi)$$

$$= 2\cdot\left(\frac{Z}{a}\right)^{3/2} e^{-Zr/a}\cdot\frac{1}{\sqrt{4\pi}} \tag{13.87}$$

Note that $\psi^{(0)}(\mathbf{r}_1, \mathbf{r}_2)$ is symmetric with respect to the exchange of coordinates. So if we include spin also, the spin wave function will be antisymmetric with respect to the exchange of particles 1 and 2. However, if H' does not have spin-dependent term, the inclusion of spin in the wave function does not have any significant effect. Since we restrict ourselves to electrostatic case alone, which has no spin dependence and so we are not including the, spin-dependent (singlet) wave function here.

$$\begin{aligned} H_0\psi^{(0)}(\mathbf{r}_1,\mathbf{r}_2) &= H_0 u_{100}(\mathbf{r}_1)u_{100}(\mathbf{r}_2) \\ &= [H_1 u_{100}(\mathbf{r}_1)]u_{100}(\mathbf{r}_2) + u_{100}(\mathbf{r}_1)H_2 u_{100}(\mathbf{r}_2) \\ &= \left[-\frac{1}{2}mc^2Z^2\alpha^2 - \frac{1}{2}mc^2Z^2\alpha^2\right]u_{100}(\mathbf{r}_1)u_{100}(\mathbf{r}_2) \end{aligned} \tag{13.88}$$

where m is the reduced mass and α is the fine structure constant $\alpha = e^2/\hbar c$. In the case of helium, $1/m = 1/M + 1/m_e$.

where M is the mass of the helium nucleus and m_e is the mass of the electron. Therefore, $m \simeq m_e$ and $Z = 2$ for helium nucleus.

$$\therefore\ H_0\psi^{(0)}(\mathbf{r}_1,\mathbf{r}_2) = E_0^{(0)}\psi^{(0)} \tag{13.89}$$

when $$E_0^{(0)} = -m_e c^2 Z^2\alpha^2 = -m_e c^2 4\alpha^2 \simeq 8\times(-13.6eV) = -108.8eV$$

This is too high compared to the actual ground-state energy of helium atom. Let us now calculate the correction due to perturbation H'. From time-independent perturbation theory,

$$\begin{aligned} \Delta E_0 &= \langle\psi^{(0)}|\ H'\ |\psi^{(0)}\rangle \\ &= \frac{1}{\pi}\int_0^\infty r_1^2dr_1\int_0^\infty r_2^2dr_2\int d\Omega_1\int d\Omega_2\left(\frac{Z}{a}\right)^6 \\ &\quad\times e^{-2Zr_1/a}e^{-2Zr_2/a}\frac{1}{[r_1^2 + r_2^2 + 2|\mathbf{r}_1 - \mathbf{r}_2|]} \end{aligned} \tag{13.90}$$

To carry out $d\Omega_1$ and $d\Omega_2$ integration, we have to choose two z axes. Note that the choice of z axis is always arbitrary. So we choose one z axis for $\mathbf{r}_1$ integration and choose $\mathbf{r}_1$ itself as z axis for $\mathbf{r}_2$ integration.

$$\begin{aligned} \therefore \int d\Omega_2\frac{1}{(r_1^2 + r_2^2 - 2\mathbf{r}_1\cdot\mathbf{r}_2)^{1/2}} &= \int d\Omega_2\frac{1}{(r_1^2 + r_2^2 - 2r_1r_2\cos\theta_2)^{1/2}} \\ &= \int_0^\pi \sin\theta_2 d\Omega_2\int_0^{2\pi} d\phi_2\frac{1}{(r_1^2 + r_2^2 - 2r_1r_2\cos\theta_2)} \end{aligned} \tag{13.91}$$

Define u as $u = \cos\ \theta_2$.

$$\begin{aligned} \therefore \int d\Omega_2\frac{1}{(r_1^2 + r_2^2 - 2\mathbf{r}_1\cdot\mathbf{r}_2)^{1/2}} &= 2\pi\int_{-1}^{1}\frac{du}{[r_1^2 + r_2^2 - 2r_1r_2u]} \\ &= \frac{2\pi}{r_1r_2}\left\{(r_1^2 + r_2^2 + 2r_1r_2)^{1/2} - (r_1^2 + r_2^2 - 2r_1r_2)^{1/2}\right\} \end{aligned} \tag{13.92}$$

Note that the square root of a quantity is always positive by convention. Therefore, $(r_1^2 + r_2^2 - 2r_1r_2)^{1/2}$ is either $(r_1 - r_2)$ or $(r_2 - r_1)$, depending on whether $r_1 > r_2$ or $r_1 < r_2$.

So

$$
\begin{aligned}
(r_1^2 + r_2^2 - 2r_1r_2)^{1/2} &= \begin{cases} r_1 - r_2 & \text{if } r_1 > r_2 \\ r_2 - r_1 & \text{if } r_1 < r_2 \end{cases} \\
&= |r_1 - r_2|.
\end{aligned}
$$

$$
\begin{aligned}
\Delta E = \left(\frac{Z}{a}\right)^6 & \int_0^\infty dr_1 r_1^2 dr_1 e^{\frac{-2Zr_1}{a}} \int_0^\infty dr_2 r_2^2 e^{-2Zr_2/a} \\
& \int d\Omega_1 . \frac{2\pi}{r_1r_2}\{(r_1 + r_2) - |r_1 - \mathbf{r}_2|\}
\end{aligned}
\tag{13.93}
$$

There is no θ_1 or ϕ_1 dependence in any of the factors in the integrand of (13.93).

$$\therefore \int d\Omega_1 = 4\pi$$

$$\Delta E = 8 \cdot \left(\frac{Z}{a}\right) \int_0^\infty dr_1 r_1 e^{\frac{-2Zr_1}{a}} \int_0^\infty dr_2 r_2 \{(r_2 + r_2) - |r_1 - r_2|\} e^{-2Zr_2/a} \tag{13.94}$$

Let us split r_2 integral into two parts.

$$\int_0^\infty dr_2() = \int_0^{r_1} dr_2() + \int_{r_1}^\infty dr_2()$$

In the range 0 to r_1, r_2 is less than r_1 and so $|r_1 - r_2| = r_1 - r_2 \therefore r_1 + r_2 - |r_1 - r_2| = 2r_2$
In the range r_1 to ∞, r_2 is greater than r_1 and so $|r_1 - r_2| = r_2 - r_1$. $\therefore$ $r_1 + r_2 - |r_1 - r_2| = 2r_1$

Therefore, r_2 integral becomes

$$\int_0^{r_1} dr_2 2r_2^2 e^{-2Zr_2/a} + \int_{r_1}^\infty 2r_1r_2 e^{-2Zr_2/a}$$

These integrals are easy to evaluate.

$$
\begin{aligned}
\int_{r_1}^{r_1} dr_2 r_2^2 e^{-2Zr_2/a} &= \int_0^{r_1} dr_2 r_2^2 e^{-\beta r_2} dr_2 \\
&= \left(-\frac{r_1^2}{\beta} - \frac{2r_1}{\beta^2} - \frac{2}{\beta^3}\right) e^{-\beta r_1} + \frac{2}{\beta^3}
\end{aligned}
$$

$$\int_{r_1}^\infty r_2 e^{-2Zr_2/a} = \frac{1}{\beta} r_1 e^{-\beta r_1} + \frac{1}{\beta^2} e^{-\beta r_1}$$

Here $$\beta = \frac{2Z}{a}$$

Substituting these results in (13.94), we get

$$\Delta E = 8e^2 \left(\frac{Z}{a}\right)^6 \left\{ \frac{-2}{\beta^2} \int_0^\infty r_1^2 e^{-2\beta r_1} dr_1 - \frac{4}{\beta^3} \int_0^\infty r_1 e^{-2\beta r_1} dr_1 + \frac{4}{\beta^3} \int_0^\infty r_1 e^{-\beta r_1} dr_1 \right\}$$

$$= 8e^2 \left(\frac{Z}{a}\right)^6 \cdot \frac{5}{2} \beta^5 = 8e^2 \left(\frac{Z}{a}\right)^6 \cdot \frac{5}{64} \cdot \left(\frac{a}{2Z}\right)^5$$

$$= \frac{5}{8} \frac{Ze^2}{a} = \frac{5}{8} \cdot \frac{2e^2}{a} = \frac{5}{2} \cdot \left(\frac{e^2}{2a}\right) = \frac{5}{2} \times 13.6eV = 34eV$$

Therefore, the energy of the ground state of helium in the first-order perturbation theory is

$$E = -108.8 + 34.0 = -74.8eV$$

The experimental value is $-78.975\ eV$. So the result of the first-order perturbation theory is closer to the experimental value, but still the difference is substantial.

B. Variational Method

13.7 LEAST UPPER BOUND FOR GROUND-STATE ENERGY

Another method of finding approximate solution to the time-independent Schrödinger equation is the variational method. This is a very useful method in quantum chemistry.

The time-independent Schrödinger equation is

$$H\phi_n = E_n \phi_n$$

We may not be able to get exact eigenvalue E_n and the corresponding eigenfunctions ϕ_n. Still, we know that these eigenfunctions ϕ_ns should form a complete set. Formally, this implies that any arbitrary wave function ψ can be expressed as a linear combination of ϕ_ns. Let us choose an arbitrary wave function ψ for which we can calculate $\langle \psi | H | \psi \rangle$. The wave function ψ is called the trial wave function.

As mentioned above, since ϕ_ns forms a complete set, let us write

$$\psi = \sum \mathrm{C}_n \phi_n$$

The normalized condition $\langle \psi | \psi \rangle$ leads to the condition

$$\sum_n |C_n|^2 = 1$$

$$\langle \psi | H | \psi \rangle = \sum_m \sum_n C_m^* C_n \langle \phi_m | H | \phi_n \rangle$$

$$= \sum_m \sum_n C_m^* C_n E_n \delta_{mn}$$

$$= \sum |C_n|^2 \ E_n$$

$$= E_0 \ |C_0|^2 + E_1 \ |C_1|^2 + E_2^2 \ |C_2|^2 + \ldots$$

Since E_0 is a ground-state energy, all the other energy eigenvalues are greater than E_0. i.e., $E_0 < E_1$, $E_0 < E_2 \ldots$

$$\therefore \langle \psi \,|\, H \,|\, \psi \rangle \geq E_0 \; |C_0|^2 + E_0 \; |C_1|^2 + E_0 \; |C_2|^2 \cdots$$

Or
$$\langle \psi \,|\, H \,|\, \psi \rangle \geq E_0 \tag{13.95}$$

since
$$|C_0|^2 + |C_1|^2 \cdots = 1$$

So we can choose any arbitrary function and evaluate $\langle \psi \,|\, H \,|\, \psi \rangle$. The ground-state energy E_0 will be equal to or less than this number. We can make use of this fact to estimate the least upper bound for E_0. The choice of the initial wave function is not arbitrary as it sounds. From physics considerations, we can choose an appropriate trial wave function, which is expected to give $\langle \psi \,|\, H \,|\, \psi \rangle$, a value very close to E_0.

In fact, a better way of doing this estimate is to choose an arbitrary trial wave function whose parameters can be varied. $\langle \psi \,|\, H \,|\, \psi \rangle = E(\alpha_1, \alpha_2, ..., \alpha_n)$, where α_1, $\alpha_2 \ldots$ are the parameters. Minimize this expression with respect to α_1, $\alpha_2 \ldots$

$$\frac{\partial E}{\partial \alpha_1} = \frac{\partial E}{\partial \alpha_2} = \ldots \frac{\partial E}{\partial \alpha_n} = 0 \tag{13.96}$$

Note that the variation method is not restricted to the estimation of ground state alone. This method can be extended to estimate the energy of the excited states also (see Mathews and Venkatesan).

We study the helium atom to illustrate the application of variational method.

13.8 VARIATIONAL METHOD FOR HELIUM ATOM

The Hamiltonian for helium atom is

$$H = \underbrace{\frac{p_1^2}{2m} - \frac{Ze^2}{r_1}}_{H_1} + \underbrace{\frac{p_2^2}{2m} - \frac{Ze^2}{r_2}}_{H_2} + \frac{e^2}{|\mathbf{r}_1 - \mathbf{r}_2|}$$

where
$$H_1 = \frac{p_1^2}{2m} - \frac{Ze^2}{r_1}$$

and
$$H_2 = \frac{p_2^2}{2m} - \frac{Ze^2}{r_2}$$

If the last term is not present, then H will be a sum of two Hamiltonians for two non-interacting hydrogen atoms. In that case, the wave function of ground state will be

$$\psi_0(\mathbf{r}_1, \mathbf{r}_2) = u_{100}(\mathbf{r}_1) u_{100}(\mathbf{r}_2)$$

where
$$u_{100}(\mathbf{r}) = \left(\frac{Z}{a}\right)^{3/2} e^{-Zr/2a} \frac{1}{(4\pi)^{1/2}}$$

This suggests a simple choice for the trial wave function. Let us choose $u_{100}(\mathbf{r}_1)u_{100}(\mathbf{r}_2)$ as the trial wave function with Z' in the place of Z so that Z' is the variation parameter. The trial wave function is

$$\psi(\mathbf{r}_1, \mathbf{r}_2) = 2\cdot\left(\frac{Z'}{a}\right)^{3/2} e^{-Z'r_1/2a}\left(\frac{1}{4\pi}\right)^{1/2}\cdot 2\left(\frac{Z'}{a}\right)^{3/2} e^{-Z'r_2/2a}\left(\frac{1}{4\pi}\right)^{1/2} \tag{13.97}$$

Evaluate $E(Z') = \langle\psi|\,H\,|\psi\rangle$ for this trial wave function and minimize $E\,(Z')$ to estimate the ground-state energy. Note that for actual hydrogen atom

$$\left(\frac{p^2}{2m} - \frac{Ze^2}{r}\right)u_{100}(Z, \mathbf{r}) = E_0(Z)u_{100}(Z, \mathbf{r})$$

with

$$E_0(Z) = -\frac{1}{2}mc^2(Z\alpha)^2$$

Therefore, the trial wave function $u_{100}(Z', \mathbf{r})$ satisfies

$$\left(\frac{p^2}{2m} - \frac{Z'e^2}{r}\right)u_{100}(Z', \mathbf{r}) = E_0(Z')u_{100}(Z', \mathbf{r})$$

with

$$E_0(Z') = -\frac{1}{2}mc^2(Z'\alpha)^2$$

$$\begin{aligned}\langle\psi\,|\,H\,|\,\psi\rangle &= \int\psi^*(\mathbf{r}_1, \mathbf{r}_2)H\psi(\mathbf{r}_1,\mathbf{r}_2)d\mathbf{r}_1 d^3\mathbf{r}_2\\ &= \int\psi^*(\mathbf{r}_1, \mathbf{r}_2)H_1\psi(\mathbf{r}_1, \mathbf{r}_2)d^3\mathbf{r}_1 d^3\mathbf{r}_2\\ &\quad+\int\psi^*(\mathbf{r}_1, \mathbf{r}_2)H_2\psi(\mathbf{r}_1, \mathbf{r}_2)d^3\mathbf{r}_1 d^3\mathbf{r}_2\\ &\quad+\int\psi^*(\mathbf{r}_1, \mathbf{r}_2)H'\psi(\mathbf{r}_1, \mathbf{r}_2)d^3\mathbf{r}_1 d^3\mathbf{r}_2\end{aligned}$$

Consider $H_1 u_{100}(Z', \mathbf{r}_1)$

$$\begin{aligned}H_1 u_{100}(Z', \mathbf{r}_1) &= \left(\frac{p_1^{\,2}}{2m} - \frac{Ze^2}{r_1}\right)u_{100}(Z'; \mathbf{r}_1)\\ &= \left[\frac{p_1^{\,2}}{2m} - \frac{Z'e^2}{r_1} + \frac{(Z'-Z)e^2}{r_1}\right]u_{100}(Z'; \mathbf{r}_1)\\ &= E_0(Z')u_{100}(Z', \mathbf{r}_1) + \frac{(Z'-Z)e^2}{r_1}u_{100}(Z'; \mathbf{r}_1)\end{aligned}$$

$$\begin{aligned}\int\psi^*(\mathbf{r}_1, \mathbf{r}_2)H_1\psi(\mathbf{r}_1, \mathbf{r}_2)d^3\mathbf{r}_1 d^3\mathbf{r}_2 &= \int u_{100}^*(Z'; \mathbf{r}_1)u_{100}^*(Z'; \mathbf{r}_2)H_1 u_{100}(Z'; \mathbf{r}_1)u_{100}(Z'; \mathbf{r}_2)d^3\mathbf{r}_1 d^3\mathbf{r}_2\\ &= \int u_{100}^*(Z'; \mathbf{r}_1)H_1 u_{100}(Z'; \mathbf{r}_1)d^3\mathbf{r}_1 \underbrace{\int u_{100}^*(Z'; \mathbf{r}_2)u_{100}(Z'; \mathbf{r}_2)d^3 r_2}_{1}\\ &= \int u_{100}^*(Z'; \mathbf{r}_1)H_1 u_{100}(Z'; \mathbf{r}_1)d^3\mathbf{r}_1\end{aligned}$$

$$= \int u^*_{100}(Z',\mathbf{r}_1)E_0(Z')u_{100}(Z';\mathbf{r}_1)d^3r_1 + \int u^*_{100}(Z';\mathbf{r}_1)\frac{(Z'-Z)e^2}{r_1}u_{100}(Z';\mathbf{r}_1)d^3\mathbf{r}_1$$

$$= E_0(Z') + \langle u_{100}(Z';\mathbf{r})|\frac{(Z'-Z)e^2}{r}|u_{100}(Z';\mathbf{r})\rangle$$

$$= E_0(Z') + (Z'-Z)e^2\left\langle\frac{1}{r}\right\rangle$$

From Chapter 8, we have

$$\left\langle\frac{1}{r}\right\rangle = \frac{Z'}{a}$$

$$\therefore \langle\psi|H_1|\psi\rangle = E_0(Z') + \frac{(Z'-Z)e^2Z'}{a} \tag{13.98}$$

Similarly, we have

$$\langle\psi|H_2|\psi\rangle = E_0(Z') + (Z'-Z)e^2\frac{Z'}{a} \tag{13.99}$$

$$\left\langle\psi\left|\frac{e^2}{|\mathbf{r}_1-\mathbf{r}_2|}\right|\psi\right\rangle = \int u^*_{100}(Z';\mathbf{r}_1)u^*_{100}(Z';\mathbf{r}_2)\frac{e^2}{|\mathbf{r}_1-\mathbf{r}_2|}u_{100}(Z';\mathbf{r}_1)u_{100}(Z';\mathbf{r}_2)d^3\mathbf{r}_1d^3\mathbf{r}_2$$

This integral has already been evaluated in Section 13.6.

$$\therefore \left\langle\psi\left|\frac{e^2}{|\mathbf{r}_1-\mathbf{r}_2|}\right|\psi\right\rangle = \frac{5}{8}\frac{Z'e^2}{a} = \frac{5}{8}mc^2\alpha^2Z'$$

$$\therefore E(Z') = -\frac{1}{2}mc^2\alpha\left[2Z'^2 + 4Z'(Z-Z') - \frac{5}{4}Z'\right] \tag{13.100}$$

To minimize this equation, let us set $\frac{dE}{dZ'} = 0$.

$$-4Z' + 4Z - \frac{5}{4} = 0 \quad \text{or} \quad Z' = \left(Z - \frac{5}{16}\right)$$

$$\min\ E(Z') = -\frac{1}{2}mc^2\alpha^2\cdot 2\left(Z-\frac{5}{16}\right)^2$$

$$\therefore\ E_0 \le -\frac{1}{2}mc^2\alpha^2\cdot 2\left(Z-\frac{5}{16}\right)^2 \tag{13.101}$$

$$-\frac{1}{2}mc^2\alpha^2 2\left(Z-\frac{5}{\pi}\right)^2 = -77.38\ eV$$

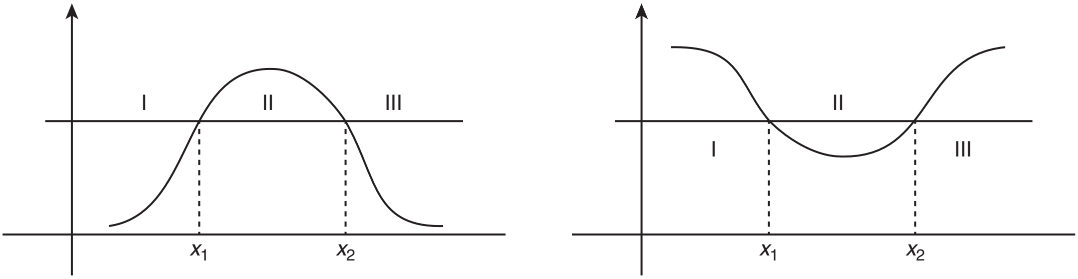
$$\therefore E_0 \le -77.38\ eV$$

This is to be compared with the experimental value −78.975 *eV*. Let us emphasize the meaning of inequality in (13.101). E_0 is estimated to be less than or equal to −77.38 *eV*. Note that the variational method does not say that E_0 is equal to −77.38 *eV*. This is the least upper bound as determined by this wave function.

C. WKB Approximation

13.9 SLOWLY VARYING POTENTIAL AND WKB APPROXIMATION

WKB method is yet another approximation to solve time-independent Schrödinger equation. The letters WKB stand for Wentzel, Kramers and Brillouin. This method is used to determine the bound-state energies and the tunnelling transmission probabilities. The general potentials $V(x)$ of interest for WKB approximation is as shown in Fig. 13.6.



(a) (b)

Fig. 13.6 Potentials with classical turning points

Both potentials share a common feature, namely that there are classical turning points x_1 and x_2. At these points, $E = V(x)$, and so the momentum $p(x) = [2m(E - V(x))]^{1/2} = 0$. These classical turning points separate the classically allowed and classically forbidden regions. For the potential in Fig. 13.6(a), regions I and III are classically allowed regions and region II is a classically forbidden region. For the potential in Fig. 13.6(b), regions I and III are classically forbidden regions, and region II is a classically allowed region.

It is for this kind of potentials the WKB approximations are widely used. We are interested in a particular kind of potentials which vary very slowly over a large distance. For a constant potential, the momentum p is a constant, and so the De Broglie wavelength $\lambda = h/p$ is constant. When the potential $V = V(x)$, a varying function, then the De Broglie wavelength is

$$\lambda(x) = \frac{h}{p} = \frac{h}{\sqrt{2m(E - V(x)}} \tag{13.102}$$

The De Broglie wavelength changes with distance. i.e., $\lambda = \lambda(x)$. The WKB approximation relies on the fact that $\lambda(x)$ changes very slowly.

Let us further elaborate the variation of $\lambda(x)$ over a large distance. Let $d\lambda(x)$ be change in the wavelength over a distance λ.

$$d\lambda = \frac{d\lambda}{dx}\lambda$$

$$\left|\frac{d\lambda}{\lambda}\right| = \left|\frac{d\lambda}{dx}\right|$$

Slowly varying De Broglie wavelength λ means

$$\left|\frac{d\lambda}{\lambda}\right| << 1 \text{ or } \left|\frac{d\lambda}{dx}\right| << 1$$

The De Broglie wavelength λ is given by

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

$$\frac{d\lambda}{dx} = -\frac{h}{p^2}\frac{dp}{dx} = -\frac{h}{p^2}\frac{d}{dx}[2m(E-V(x)]^{1/2}$$
$$= \frac{h}{p^3}m\frac{\partial V}{\partial x}$$

Therefore, the condition $\left|\frac{d\lambda}{\lambda}\right| << 1$ implies that

$$\left|\frac{m\hbar}{p^3}\frac{\partial V}{\partial x}\right| << 1 \tag{13.103}$$

This condition ensures that $\lambda(x)$ varies slowly.

13.10 WKB EXPANSION

For a free particle, the wave function is $u(x) = Ae^{ipx/\hbar}$, where p is the momentum. In the case of constant potential also, the free particle wave function is $Ae^{ipx/\hbar}$, where $p = [2m(E-V)]^{1/2}$. This suggests that for slowly varying potential, $u(x)$ can be written as

$$u(x) = Ae^{iS(x)/\hbar} \tag{13.104}$$

The Schrödinger equation is now written as

$$\frac{d^2u}{dx^2} + \frac{p^2(x)}{\hbar^2}u(x) = 0 \tag{13.105}$$

where $$p^2(x) = 2m(E-V(x)) \tag{13.106}$$

Substituting (13.104) in (13.105), we get

$$\left(\frac{dS}{dx}\right)^2 - p^2 - \hbar i\frac{d^2S}{dx^2} = 0 \tag{13.107}$$

Let us expand S as a power series in $\hbar$.

$$\therefore S = S_0 + \hbar S_1 + \frac{\hbar^2}{2}S_2 \cdots \tag{13.108}$$

$$\therefore \frac{dS}{dx} = \left(\frac{dS_0}{dx}\right) + \hbar\left(\frac{dS_1}{dx}\right) + \frac{\hbar^2}{2}\left(\frac{dS_2}{dx}\right)\cdots$$

Squaring the above equation, we get

$$\left(\frac{dS}{dx}\right)^2 = \left(\frac{dS_0}{dx}\right)^2 + 2\hbar\left(\frac{dS_0}{dx}\right)\left(\frac{dS_1}{dx}\right) + \hbar^2\left[\left(\frac{dS_1}{dx}\right)^2 + \frac{dS_0}{dx}\frac{dS_2}{dx}\cdots\right]$$

From (13.108), we get

$$\frac{d^2S}{dx^2} = \frac{d^2S_0}{dx^2} + \hbar\frac{d^2S_1}{dx^2} + \hbar^2\frac{d^2S_2}{dx^2}\cdots$$

Substituting this equation in (13.107), we get

$$\left[\left(\frac{dS_0}{dx}\right)^2 - p(x)\right] + \hbar\left[2\left(\frac{dS_0}{dx}\right)\left(\frac{dS_1}{dx}\right) - i\left(\frac{d^2S_0}{dx^2}\right)\right]$$
$$+\hbar^2\left[\left(\frac{dS_1}{dx}\right)^2 + \left(\frac{dS_0}{dx}\right)\left(\frac{dS_2}{dx}\right) - i\frac{d^2S_1}{dx^2}\right] + \cdots = 0$$

Equating the coefficients of different powers of $\hbar$ separately, we have

$$\left(\frac{dS_0}{dx}\right)^2 - p^2(x) = 0 \tag{13.109}$$

$$2\,\frac{dS_0}{dx}\frac{dS_1}{dx} - i\frac{d^2S_0}{dx^2} = 0 \tag{13.110}$$

$$\frac{dS_2}{dx} = \frac{i\dfrac{d^2S_1}{dx^2} - \left(\dfrac{dS_1}{dx}\right)^2}{\left(\dfrac{dS_0}{dx}\right)} \tag{13.111}$$

We have restricted up to $\hbar^2$ term. It will be shown soon that $S_2(x)$ can be neglected for slowly varying potential, which essentially makes $S(x) = S_0(x) + \hbar S_1(x)$.

Solving the Equation (13.109), we get

$$\frac{dS_0}{dx} = [p(x)] \tag{13.112}$$

or

$$S_0(x) = \int^x [p(x')]dx' \tag{13.113}$$

The constant of integration can be taken care by taking suitable lower limit.

From the Equation (13.110), we get

$$\frac{dS_1}{dx} = \frac{i}{2}\frac{\dfrac{d^2S_0}{dx^2}}{\left(\dfrac{dS_0}{dx}\right)} \tag{13.114}$$

We can rewrite the Equation (13.114) as

$$\frac{dS_1}{dx} = \frac{i}{2}\frac{d}{dx}\left[\ln\left(\frac{dS_0}{dx}\right)\right]$$

$$\therefore S_1 = \frac{i}{2}\ \ln\left(\frac{dS_0}{dx}\right) = i\ln[p(x)]^{1/2} \tag{13.115}$$

Let us now evaluate $S_2(x)$ from (13.111). From Equation (13.115) we get

$$\frac{dS_1}{dx} = \frac{i}{2}\frac{1}{p(x)}\frac{dp(x)}{dx}$$

From $p^2(x) = 2m\,(E - V(x))$, we get

$$\frac{dp}{dx} = -\frac{m}{p}\frac{dV}{dx} \tag{13.116}$$

$$\therefore \frac{dS_1}{dx} = -\frac{im}{2p^2}\frac{dV}{dx}$$

Differentiating this equation (and using (13.116)), we get

$$\begin{aligned}\frac{d^2S_1}{dx^2} &= \frac{im}{p^3}\frac{dp}{dx}\cdot\frac{dV}{dx} - \frac{im}{2p^2}\frac{d^2V}{dx^2}\\ &= -\frac{im^2}{p^4}\left(\frac{dV}{dx}\right)^2 - \frac{im}{2p^2}\frac{d^2V}{dx^2}\end{aligned}$$

Substituting this expression in (13.111), we get

$$\frac{dS_2}{dx} = \frac{5}{4}\frac{m^2}{p^5}\left(\frac{dV}{dx}\right)^2 + \frac{m}{2p^3}\frac{d^2V}{dx^2}$$

Integrating this equation, we get

$$S_2 = \int\frac{5}{4}\frac{m^2}{p^4}\left(\frac{dV}{dx}\right)^2 dx + \int\frac{m}{2p^3}\frac{d^2V}{dx^2}dx$$

Consider the second integral.

$$\begin{aligned}\int\frac{m}{2p^3}\frac{d^2V}{dx^2}dx &= \int\frac{m}{2p^3}\frac{d}{dx}\left(\frac{dV}{dx}\right)dx\\ &= \frac{m}{2p^3}\frac{dV}{dx} + \int\frac{3m}{2p^4}\cdot\frac{dp}{dx}\cdot\frac{dV}{dx}dx\end{aligned}$$

$$= \frac{m}{2p^3}\frac{dV}{dx} + \int \frac{3m}{2p^4}\cdot\left(-\frac{m}{p}\frac{dV}{dx}\right)\cdot\frac{dV}{dx} \qquad \text{(Use (13.116))}$$

$$= \frac{m}{2p^3}\frac{dV}{dx} - \int \frac{3m^2}{2p^5}\left(\frac{dV}{dx}\right)^2 dx$$

$$\therefore\ S_2 = \frac{m}{2p^3}\left(\frac{dV}{dx}\right) - \frac{1}{4}\int \frac{m^2}{p^5}\left(\frac{dV}{dx}\right)^2 dx.$$

$$\frac{\hbar^2}{2}S_2 = \frac{\hbar}{2}\cdot\frac{m}{p^3}\cdot\left(\hbar\frac{dV}{dx}\right) - \frac{1}{4}\cdot\int \frac{m}{p^2}\cdot\left(\frac{m\hbar}{p^3}\cdot\frac{dV}{dx}\right)dx$$

The quantity $\left|\frac{m\hbar}{p^3}\frac{dV}{dx}\right|$ is $\left|\frac{d\lambda}{\lambda}\right|$. For slowly varying potential, $\frac{m\hbar}{p^3}\frac{dV}{dx} << 1$. Therefore, $\frac{\hbar^2}{2}S_2$ can be neglected.

So, in the WKB approximation, we have

$$S = S_0 + \hbar S_1$$

$$\therefore u(x) = A e^{iS_0/\hbar}\cdot e^{iS_1}$$

$$= A e^{iS_0/\hbar}\cdot e^{-\frac{1}{2}\ln p}$$

$$= A e^{iS_0/\hbar}\cdot e^{\ln p^{-1/2}} = \frac{A}{[p(x)]^{1/2}} e^{iS_0/\hbar}$$

Note that

$$S_0(x) = \pm\int^x p(x')dx'$$

Therefore, $u(x)$ is given by

$$u(x) = \frac{A}{[p(x)]^{1/2}} \exp\left[\pm\frac{i}{\hbar}\int^x p(x')dx'\right]$$

The wave function depends on $p(x)$, which is given by $[2m(E - V(x))]^{1/2}$. Naturally, the wave function $u(x)$ in the classically allowed region differs from the wave function in the classically forbidden region. So we have to consider the following two cases.

E > *V* Classically Allowed Region

The wave function $u(x)$, in this region in WKB approximation, is given by

$$u(x) = \frac{A}{[p(x)]^{1/2}} \exp\left[\pm\frac{i}{\hbar}\int^x p(x')dx'\right]$$

$E < V$ Classically Forbidden Region

The factor $[2m(E-V(x))]^{1/2}$ is now imaginary.

i.e., $$p(x) = i\sqrt{2m(V-E)}$$

The $u(x)$ becomes

$$u(x) = \frac{A}{|p(x)|^{1/2}} \ \exp\left[\pm\frac{1}{\hbar}\int^{x}|p(x')|\,dx'\right]$$

In the denominator, the factor i will not make any difference, and so it is omitted.

Now we are ready to write the general expression for $u(x)$.

The general expression for $u(x)$ is

$$u(x) = \frac{C_1'}{p^{1/2}} \ \exp\left[\frac{i}{\hbar}\int_{x_0}^{x} p(x')dx'\right] + \frac{C_2'}{p^{1/2}} \ \exp\left[\frac{-i}{\hbar}\int_{x_0}^{x} p(x')dx'\right] \quad E > V$$

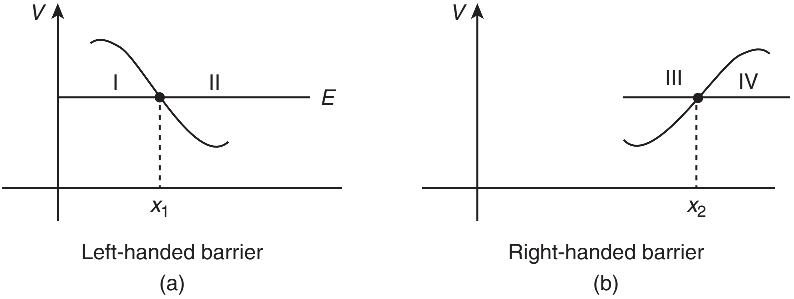
Or equivalently, this can be expressed as

$$u(x) = \frac{C_1}{p^{1/2}} \ \sin\left[\frac{1}{\hbar}\int_{x_0}^{x} p(x')dx' + \delta\right] + \frac{C_2}{p^{1/2}} \ \cos\left[\frac{1}{\hbar}\int_{x_0}^{x} p(x')dx' + \delta\right] \quad E > V$$

Similarly, we have

$$u(x) = \frac{d_1}{|p|^{1/2}} \ \exp\left[\frac{1}{\hbar}\int_{x_0}^{x}|p(x')|\,dx'\right] + \frac{d_2}{|p|^{1/2}} \ \exp\left[-\frac{1}{\hbar}\int_{x_0}^{x}|p(x')|\,dx'\right] \quad E < V$$

Let us extend these results to the regions near the classical turning points x_1 and x_2 as shown in Fig. 13.7, which will be useful later.



Left-handed barrier
(a)

Right-handed barrier
(b)

Fig. 13.7 Potentials near classical turning points

Left-handed Barrier

The barrier is to the left of the classical turning point x_1 as shown in Fig.13.7. The regions I and II are classically forbidden and allowed regions. The general form of the wave functions in these regions is as follows:

$$u_{\text{I}} = \frac{1}{|p|^{1/2}}\left\{d_1 \exp\left(\frac{1}{\hbar}\int_{x_1}^{x} |p(x')|\, dx'\right) + d_2 \exp\left(-\frac{1}{\hbar}\int_{x_1}^{x} |p(x')| dx'\right)\right\} \qquad x < x_1 \qquad (13.117)$$

$$u_{\text{II}} = \frac{1}{[p(x)]^{1/2}}\left[c_1 \sin\left(\int_{x_1}^{x} p(x')dx' + \delta\right) + c_2 \cos\left(\int_{x_1}^{x} p(x')dx' + \delta\right)\right] \qquad x > x_1 \qquad (13.118)$$

Or equivalently

$$u_{\text{II}} = \frac{1}{[p(x)]^{1/2}}\left[C_1' \exp\left(\frac{i}{\hbar}\int_{x_1}^{x} p(x')dx'\right) + C_2' \exp\left(\frac{i}{\hbar}\int_{x_1}^{x} p(x')dx'\right)\right] \qquad x > x_1 \qquad (13.119)$$

Right-handed Barrier

In Fig. 13.7(b), the barrier is at the right of x_2. Region III and region IV are classically allowed and forbidden regions. The general forms of the wave functions are as follows:

$$u_{\text{III}} = \frac{C_3}{[p(x)]^{1/2}} \sin\left(-\frac{1}{\hbar}\int_{x_2}^{x} p(x')dx' + \beta\right) + \frac{C_4}{[p(x)]^{1/2}} \cos\left(-\frac{1}{\hbar}\int_{x_2}^{x} p(x')dx' + \beta\right) \qquad x < x_2 \qquad (13.120)$$

Or equivalently

$$u_{\text{III}} = \frac{1}{[p(x)]^{1/2}}\left[C_3' \exp\left(-\frac{i}{\hbar}\int_{x_2}^{x} p(x')dx'\right) + C_4' \exp\left(-\frac{i}{\hbar}\int_{x_2}^{x} p(x')dx'\right)\right] \qquad x < x_2 \qquad (13.121)$$

$$u_{\text{IV}} = \frac{1}{|p(x)|^{1/2}}\left[d_3 \exp\left(\frac{1}{\hbar}\int_{x_2}^{x} |p(x')| dx'\right) + d_4\left(-\frac{1}{\hbar}\int_{x_2}^{x} |p(x')|\, dx'\right)\right] \qquad x > x_2 \qquad (13.122)$$

13.11 CONNECTION FORMULAE

Normally, we will have solutions in different regions. Of course, there should be only one wave function from $-\infty$ to ∞. This condition is achieved by matching the wave functions at the boundaries. However, this cannot be done in a straightforward way in the case of WKB approximation. Consider the regions I and II in the potential as shown in Fig. 13.7. Region I is a classically forbidden region, and region II is a classically allowed region. The solutions in these regions are different. The constants C_1, C_2 and d_1, d_2 are such that $u_{\text{I}}(x_1) = u_{\text{II}}(x_1)$. However, at x_1, $p(x_1) = \sqrt{2m(E-V)} = 0$ and therefore, the WKB wave functions u_{I} and u_{II} become infinite. Obviously, WKB solutions are valid only in a region far away from classical turning points (both for classically allowed and forbidden). Therefore, WKB solution is only an asymptotic solution in the place of an exact solution, so we have only incomplete information which make the matching of the wave functions difficult.

Still, we have to find a method of matching these asymptotic solutions in different regions.

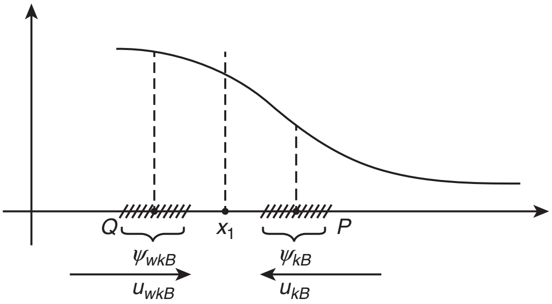


Fig. 13.8 Regions where ψ_{WKB} and u_{WKB} agree with each other (approximately)

Let us consider two potentials $V(x)$ and $V_1(x)$ for which the solutions are $u(x)$ and $\psi(x)$.

$$\frac{d^2u(x)}{dx^2} + \frac{2m}{\hbar^2}(E - V)u(x) = 0 \qquad \text{for whole region}$$

$$\frac{d^2\psi(x)}{dx^2} + \frac{2m}{\hbar^2}(E - V_1)\psi(x) = 0 \qquad \text{near } x_1$$

The reason for considering two different potentials will become clear soon. In fact, they are same near x_1. The functions u_{WKB} and ψ_{WKB} represent the solutions to the above equations in the WKB approximations. Our strategy is to determine the coefficients in u_{WKB} using the coefficients in ψ_{WKB}. In what follows, it emerges that both these WKB solutions have the same coefficients.

Let us consider two shaded regions centered around *P* and *Q* as shown in Fig.13.8. *P* and *Q* are far away from x_1. They are sufficiently far away from x_1 so that it is valid to use the WKB approximation for both $u(x)$ and $\psi(x)$.

Let us choose $V_1(x)$ as Taylor's expansion of $V(x)$ near x_1 in the left-handed barrier as shown in Fig. 13.8.

The Taylor's expansion of $V(x)$ near x_1 is

$$V(x) = V(x_1) + \left.\frac{\partial V}{\partial x}\right|_{x_1}(x - x_1) = E - A(x - x_1)$$

i.e.,

$$V_1(x) = E - A\,(x - x_1)$$

It is important to recognize the difference in the region of validity for $u(x)$ and $\psi(x)$. The term *validity* here refers to meaningfulness of the wave function in the context of our problem. It is the potential $V(x)$ for which we have to find the solution to the Schrödinger equation. For a given potential $V(x)$, the solution $u(x)$ is meaningful over the entire region. Since the Taylor's expansion is valid only for small region around x_1, the Schrödinger equation for $V_1(x)$ as well as its solution $\psi(x)$ are restricted to a smaller region around x_1. Correspondingly, the ranges of the WKB solutions for these two potentials also differ, though both of them are valid only when they are far away from x_1. The range of u_{WKB} is larger than that of ψ_{WKB}. Since $V(x)$ and $V_1(x)$ are the same near x_1, the WKB solutions u_{WKB} and ψ_{WKB} are also the same in a region called overlap region. This is shown as shaded region in Fig.13.8.

$$\therefore \quad u_{WKB} \sim \psi_{WKB} \quad \text{in the shaded region}$$

Therefore, it is enough to determine the coefficients that occur in ψ_{WKB}.

How to Determine the Coefficients in ψ_{WKB}

The determination of the coefficients in ψ_{WKB} depends on the following observations. The advantage of the potential $V_1(x) = E - A(x - x_1)$ is that the Schrödinger equation for this potential is exactly solvable. It is a superposition of Airy's function $A_i(x)$ and $B_i(x)$. The general solution corresponding to the linear potential is

$$\psi = aA_i(x) + bB_i(x)$$

Let us recall that P and Q are so far away from x_1 that the WKB approximation is valid in regions near these points. Normally, the terms *asymptotic limit of a function* mean its behaviour in the limit $x \to \pm\infty$ (or for large values of x). Though the points P and Q are not that far away, we take the function $\psi(x)$ in these regions by its asymptotic form. So we approximate $\psi(x)$ as follows:

$$\psi(x) \simeq \psi(x \to \infty) \qquad \text{near } P$$
$$\psi(x) \simeq \psi(x \to -\infty) \qquad \text{near } Q$$

These asymptotic limits can be easily determined from the exact analytical solutions, namely Airy's functions $A_i(x)$ and $B_i(x)$. By matching these solutions with the WKB solution ψ_{WKB}, we can determine the coefficients in the ψ_{WKB}.

Compare the asymptotic solution $\psi(x \to \infty)$ and ψ_{WKB} in the shaded region around P. Compare the asymptotic solution $\psi(x \to -\infty)$ and ψ_{WKB} in the shaded region around Q. From these comparisons, we determine the coefficients of C_1, C_2, d_1 and d_2.

The same coefficients hold good for u_{WKB}.

13.11.1 Exact Solution for Linear Potential and Its Asymptotic Limit

The Taylor's expansion of $V(x)$ near x_1 is

$$V(x) = V(x) + \left.\frac{\partial V}{\partial x}\right|_{x=x_1}(x - x_1) = E - A(x - x_1)$$

$p^2(x)$ is given by

$$p^2(x) = 2m(E - V_1) = 2mA(x - x_1) \tag{13.123}$$

The Schrödinger equation is

$$\frac{d^2\psi}{dx^2}+\frac{p^2(x)}{\hbar^2}\psi=0$$

$$\frac{d^2\psi}{dx^2}+\frac{2mA}{\hbar^2}(x-x_1)\psi=0 \tag{13.124}$$

Define $q(x)$ as

$$q(x)=\left(\frac{2mA}{\hbar^2}\right)^{1/3}(x-x_1) \tag{13.125}$$

$$dq=\left(\frac{2mA}{\hbar^2}\right)^{1/3}dx$$

Substituting these expressions in the Equation (13.124), we get

$$\frac{d^2\psi}{dq^2}+q\psi=0$$

The solutions to this equation are the Airy functions $A_i(-q)$ and $B_i(-q)$.

$$\therefore \psi(q)=a\,A_i(-q)+bB_i(-q) \tag{13.126}$$

The asymptotic behaviours of $A_i(-q)$ and $B_i(-q)$ are given by

$$A_i(-q)\rightarrow\begin{cases}\dfrac{1}{\pi^{1/2}}\cdot\dfrac{1}{q^{1/4}}\cos\left(\dfrac{2}{3}q^{1/3}-\dfrac{\pi}{4}\right) & \text{as} \quad q\rightarrow\infty \qquad (13.127)\\[2ex] \dfrac{1}{2\pi^{1/2}}\cdot\dfrac{1}{|q|^{1/4}}\ \exp\left(-\dfrac{2}{3}|q|^{3/2}\right) & \quad q\rightarrow-\infty \qquad (13.128)\end{cases}$$

$$B_i(-q)\rightarrow\begin{cases}-\dfrac{1}{(\pi)^{1/2}}\cdot\dfrac{1}{q^{1/4}}\ \sin\left(\dfrac{2}{3}q^{3/2}-\dfrac{\pi}{4}\right) & \text{as} \quad q\rightarrow\infty \qquad (13.129)\\[2ex] \dfrac{1}{\pi^{1/2}}\cdot\dfrac{1}{|q|^{1/4}}\ \exp\left(\dfrac{2}{3}|q^{3/2}|\right) & \text{as} \quad q\rightarrow-\infty \qquad (13.130)\end{cases}$$

The asymptotic solution $\psi(x)$ (13.126) is given by

$$\psi(x\rightarrow\infty)=\frac{1}{\pi^{1/2}}\cdot\frac{1}{q^{1/4}}\left[a\,\cos\left(\frac{2}{3}q^{3/2}-\frac{\pi}{4}\right)-b\,\sin\left(\frac{2}{3}q^{3/2}-\frac{\pi}{4}\right)\right]\ \text{as}\ q\rightarrow\infty \tag{13.131}$$

$$\psi(x\rightarrow-\infty)-\frac{1}{\pi^{1/2}}\frac{1}{|q|^{1/4}}\left[\frac{a}{2}\ \exp\left(\frac{-2}{3}|q|^{3/2}\right)+b\ \exp\left(\frac{2}{3}|q|^{3/2}\right)\right]\ \text{as}\ q\rightarrow-\infty \tag{13.132}$$

13.11.2 WKB Solution for Linear Potential

Case 1: $x > x_1$
From the Equation (13.123), we have

$$x - x_1 = \left(\frac{\hbar^2}{2mA}\right)^{1/3} q \qquad (13.133)$$

$$p^2(x) = (2mA)\left(\frac{\hbar^2}{2mA}\right)^{1/3} q = (2mA\hbar)^{2/3} q$$

$$\frac{p(x)}{\hbar} dx = \frac{(2mA\hbar)^{1/3}}{\hbar} \cdot \left(\frac{\hbar}{2mA}\right)^{1/3} q^{1/2} dq = q^{1/2} dq$$

$$\therefore \int_{x_1}^{x} \frac{p(x')}{\hbar} dx' = \int_{0}^{q} (q')^{1/2} dq' = \frac{2}{3} q^{3/2}$$

From the Equation (13.118), we have

$$\psi_{\text{WKB}} = \frac{C_1}{(2mA\hbar)^{1/6} q^{1/4}} \sin\left(\frac{2}{3} q^{3/2} + \delta\right) + \frac{C_2}{(2mA\hbar)^{1/6} q^{1/4}} \cos\left(\frac{2}{3} q^{3/2} + \delta\right) \quad x > x_1 \qquad (13.134)$$

Case 2: $x < x_1$
From (13.125), it can be seen that $q(x)$ is negative in this region. Let us define $z(x)$ as

$$z(x) = -q(x)$$

$\therefore z(x)$ is a positive quantity for $x < x_1$.

$$p^2 = (2mA\hbar)^{2/3} q = -(2mA\hbar)^{2/3} z$$

$$\therefore p = i(2mA\hbar)^{2/3} z^{1/2} \quad \therefore |p| = (2mA\hbar)^{1/3} z^{1/2}$$

$$\text{since } z(x) = \left(\frac{2mA}{\hbar^2}\right)^{1/3} (x_1 - x), \quad dx = -\left(\frac{\hbar^2}{2mA}\right)^{1/3} dz$$

$$\therefore \frac{|p|\, dx'}{\hbar} = -\frac{(2mA\hbar)^{1/3}}{\hbar} \left(\frac{\hbar^2}{2mA}\right)^{1/3} z^{1/2} dz = -z^{1/2} dz$$

Making use of the fact that $|q| = -q$, we get

$$\int_{x_1}^{x} \frac{|p|}{\hbar} dx' = \int_{0}^{z} -(z')^{1/2} dz' = -\frac{2}{3} z^{3/2} = -\frac{2}{3} (-q)^{3/2} = -\frac{2}{3} |q|^{3/2}$$

$$\therefore \psi_{\text{WKB}} = \frac{1}{(2mA\hbar)^{1/6} \, |q|^{1/6}} \left[d_1 e^{-\frac{2}{3}|q|^{3/2}} + d_2 e^{\frac{2}{3}|q|^{3/2}} \right] \qquad x < x_1 \qquad (13.135)$$

Suppose we have the boundary condition that the wave function $u(x)$ goes to zero to the left of the classical turning point. i.e., $u(x) \to 0$ as $x \to -\infty$. This implies that as $q \to -\infty$, $\psi \to 0 \therefore d_2 = 0$.

$$\therefore \psi_{WKB} = \frac{1}{(2mA\hbar)^{\frac{1}{6}}\left|q\right|^{4}} d_1 e^{-\frac{2}{3}|q|^{3/2}} \quad \text{as} \quad q \to -\infty \tag{13.136}$$

Comparison of this ψ_{WKB} with the asymptotic behaviour in (13.132) implies

$$b = 0 \ \text{ and } \ \frac{d_1}{(2mA\hbar)^{1/6}} = \frac{a}{2\sqrt{\pi}} \tag{13.137}$$

For the wave function to the right of x_1, $x - x_1$ is positive, and so q is positive. Therefore, we have to match (13.134) with (13.131) (Take $b = 0$ in (13.131)). This implies that $C_1 = 0$ and $\delta = -\pi/4$ and $\frac{C_2}{(2mA\hbar)^{1/6}} = \frac{a}{\sqrt{\pi}}$.

$$\therefore C_2 = 2d_1 \tag{13.138}$$

For left-handed barrier, we have $C_1 = 0$, $d_2 = 0$ and $C_2 = 2d_1$.

So, as we mentioned earlier, the values of the coefficients C_1, C_2, d_1, d_2, and δ hold good for the asymptotic solutions of $u(x)$, which is a solution corresponding to a general potential $V(x)$. So we conclude that for the boundary condition $u(x) \to 0$ as $x \to -\infty$,

$$\frac{1}{|p|^{1/2}} \exp\left(\int_{x_1}^{x} \frac{|p(x')|}{\hbar} dx'\right) \to \frac{2}{[p(x)]^{1/2}} \cos\left[\int_{x_1}^{x} \frac{p(x')}{\hbar} dx' - \frac{\pi}{4}\right] \tag{13.139}$$

This expression is known as connection formula.

There are four such connection formulae. For the sake of completeness, we give the list of all the formulae.

1. Barrier to the left of the turning point x_1.

$$\frac{1}{|p(x)|^{1/2}} \exp\left[\frac{1}{\hbar}\int_{x_1}^{x} |p(x')|\, dx'\right] \to \frac{2}{[p(x)]^{1/2}} \cos\left[\frac{1}{\hbar}\int_{x_1}^{x} p(x') dx' - \frac{\pi}{4}\right] \tag{13.140}$$

$$-\frac{1}{|p(x)|^{1/2}} \exp\left[\int_{x_1}^{x} |p(x')| dx'\right] - \leftarrow \frac{1}{[p(x)]^{1/2}} \frac{1}{\hbar} \sin\left[\frac{1}{\hbar}\int_{x_1}^{x} p(x') dx' - \frac{\pi}{4}\right] \tag{13.141}$$

In the LHS, $x < x_{1,}$ and in the RHS, $x > x_1$.

2. Barrier to the right of the turning point x_2.

$$\frac{2}{|p(x)|^{1/2}} \cos\left[-\frac{1}{\hbar}\int_{x_1}^{x} |p(x')|\, dx' - \frac{\pi}{4}\right] \leftarrow \frac{1}{[p(x)]^{1/2}} \exp\left[-\frac{1}{\hbar}\int_{x_1}^{x} p(x') dx'\right] \tag{13.142}$$

$$\frac{1}{|p(x)|^{1/2}} \sin\left[-\frac{1}{\hbar}\int_{x_2}^{x} p(x') dx' - \frac{\pi}{4}\right] \to \frac{-1}{|p(x)|^{\frac{1}{2}}} \exp\left[\frac{1}{\hbar}\int_{x_2}^{x} |p(x')|\, dx'\right] \tag{13.143}$$

In the LHS, $x < x_2$ and in the RHS $x > x_{2.}$

The presence of arrows indicates that the matching between WKB solution requires more care. We derived the connection formula (13.139) under a simple boundary condition, namely $u(x) \to 0$ as $x \to -\infty$, and so we assumed that only the exponentially decreasing function alone is present in the left of the turning point x_1. The boundary condition for the wave function need not be so simple. For instance, in the second formula (13.141), we have exponentially increasing function in the left of x_1. This term alone cannot be present in this region since the wave function will become infinite as $x \to -\infty$. There should be other terms also. It is such details that put restrictions indicated by the arrows for the applications of the specific correction formula.

13.12 BOHR-SOMMERFIELD QUANTIZATION RULE

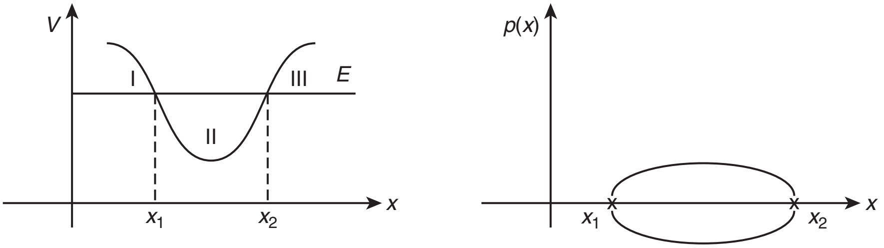


Fig. 13.9 Potential $V(x)$ and classical phase are for a particle of energy E

Consider a simple potential as shown in Fig.13.9, where x_1 and x_2 are the classical turning points. Regions I and III are classically forbidden regions. Classically, a particle of energy E executes a vibratory motion between x_1 and x_2 and the phase space corresponding to the classical vibrating particle is shown in Fig.13.9.

Quantum mechanically, the wave functions are u_{I}, u_{II}, and u_{III}. The wave functions in regions I and III should behave like exponentially decreasing function.

$$u_I = \frac{D_1}{|p(x)|^{1/2}} \exp\left[\frac{1}{\hbar}\int_{x_1}^{x} |p(x')|\,dx'\right] \tag{13.144}$$

and

$$u_{III} = \frac{D_3}{|p(x)|^{1/2}} \exp\left[-\frac{1}{\hbar}\int_{x_2}^{x} |p(x')|\,dx'\right] \tag{13.145}$$

Therefore, the connecting formulae for the wave functions inside region II are as follows:

$$D_1 \frac{1}{|p(x)|^{1/2}} \exp\left[\frac{1}{\hbar}\int_{x_1}^{x} |p(x')|dx'\right] \to \frac{2D_1}{[p(x)]^{1/2}} \cos\left(\frac{1}{\hbar}\int_{x_1}^{x} p(x')dx' - \frac{\pi}{4}\right) \tag{13.146}$$

$$D_2 \frac{1}{|p(x)|^{1/2}} \exp\left[-\frac{1}{\hbar}\int_{x_2}^{x} |p(x')|dx'\right] \to \frac{2D_2}{[p(x)]^{1/2}} \cos\left(-\frac{1}{\hbar}\int_{x_2}^{x} p(x')dx' - \frac{\pi}{4}\right) \tag{13.147}$$

The single valuedness of the wave function in region II implies

$$\frac{D_1}{|p(x)|^{1/2}}\cos\left[\frac{1}{\hbar}\int_{x_1}^{x}|p(x')|dx'-\frac{\pi}{4}\right]=\frac{D_2}{[p(x)]^{1/2}}\cos\left(-\int_{x_2}^{x}p(x')dx'-\frac{\pi}{4}\right)$$

Let us define θ as $\theta=-\frac{1}{\hbar}\int_{x_2}^{x}(\ )dx'-\frac{\pi}{4}$. Then the above equation can be rewritten as

$$\cos\left[\frac{1}{\hbar}\int_{x_1}^{x}p(x')dx'-\frac{\pi}{4}\right]=D\cos\left(-\frac{1}{\hbar}\int_{x_2}^{x}p(x')dx'-\frac{\pi}{4}\right)=D\cos\theta \qquad (13.148)$$

where $$D=\frac{D_2}{D_1}$$

The argument of the cosine function in the LHS can be rewritten as

$$\frac{1}{\hbar}\int_{x_1}^{x}(\ )dx'-\frac{\pi}{4}=\underbrace{\frac{1}{\hbar}\int_{x_1}^{x}(\ )dx'+\frac{1}{\hbar}\int_{x}^{x_2}(\ )dx'}_{\int_{x_1}^{x_2}(\ )dx'}-\underbrace{\frac{1}{\hbar}\int_{x}^{x_2}(\ )dx'}_{-\int_{x_2}^{x}(\ )dx'}-\frac{\pi}{4}$$

$$=\frac{1}{\hbar}\int_{x_1}^{x_2}(\ )dx'+\frac{1}{\hbar}\int_{x_2}^{x}(\ )dx'+\frac{\pi}{4}-\frac{\pi}{2}$$

$$=\frac{1}{\hbar}\int_{x_1}^{x_2}(\ )dx'-\left(-\frac{1}{\hbar}\int_{x_2}^{x}(\ )dx'-\frac{\pi}{4}\right)-\frac{\pi}{2}=\frac{1}{\hbar}\int_{x_1}^{x_2}(\ )dx'-\theta-\frac{\pi}{2} \qquad (13.149)$$

Using the Equation (13.149) for the LHS of the Equation (13.148), we get

$$\cos\left[\frac{1}{\hbar}\int_{x_1}^{x_2}p(x')dx'-\theta-\frac{\pi}{2}\right]=D\cos\theta$$

$$\cos\left[\frac{1}{\hbar}\int_{x_1}^{x_2}p(x')dx'-\frac{\pi}{2}\right]\cos\theta+\sin\left(-\frac{1}{\hbar}\int_{x_1}^{x_2}p(x')dx'-\frac{\pi}{2}\right)\sin\theta=D\cos\theta$$

The solution to this equation is

$$\frac{1}{\hbar}\int_{x_1}^{x}p(x')dx'-\frac{\pi}{2}=n\pi \quad n=0,1,2\ldots$$

$$\frac{1}{\hbar}\int_{x_1}^{x_2}p(x')dx'=\left(n+\frac{1}{2}\right)\pi \qquad (13.150)$$

From Fig.13.9, it can be seen that

$$\int_{x_1}^{x_2} p(x')dx' = \frac{1}{2}\oint p(x')dx'$$

$$\therefore \oint p(x')dx' = 2\left(n+\frac{1}{2}\right)\hbar\pi = \left(n+\frac{1}{2}\right)h \tag{13.151}$$

This condition is known as Bohr-Sommerfeld quantization rule.

13.13 ENERGY EIGENVALUES OF BOUND STATES USING WKB APPROXIMATION

Bohr-Sommerfeld quantization can be used to evaluate the energy eigenvalue of a bound state. The simplest case is harmonic oscillator. The energy of classical harmonic oscillator is given by

$$E = \frac{p^2}{2m} + \frac{1}{2}m\omega^2 x^2$$

$$p^2 = 2m\left(E - \frac{1}{2}m\omega^2 x^2\right)$$

$$\therefore p = (2mE)^{1/2}\left(1 - \frac{m\omega^2}{2E}x^2\right)^{1/2}$$

$$\text{Put } x = \left(\frac{2E}{m\omega^2}\right)^{1/2}\sin\theta \quad \therefore dx = \left(\frac{2E}{m\omega^2}\right)^{1/2}\cos\theta d\theta$$

The classical turning points are

$$x = \pm\left(\frac{2E}{m\omega^2}\right)^{1/2}$$

i.e.

$$x_1 = \left(\frac{2E}{m\omega^2}\right)^{1/2}\sin\theta_1 = \left(\frac{2E}{m\omega^2}\right)^{1/2} \qquad \therefore \theta_1 = \frac{\pi}{2}$$

$$x_2 = +\left(\frac{2E}{m\omega^2}\right)^{1/2}\sin\theta_2 = -\left(\frac{2E}{m\omega^2}\right)^{1/2} \qquad \therefore \theta_2 = \frac{-\pi}{2}$$

$$\frac{1}{\hbar}\int_{x_1}^{x_2} pdx = \int_{-\pi/2}^{\pi/2}(2mE)^{1/2}\left(\frac{2E}{m\omega^2}\right)^{1/2}\cos^2\theta d\theta = \frac{E\pi}{\omega} = \left(n+\frac{1}{2}\right)\pi \qquad \text{(Use (13.150))}$$

$$\therefore E = \left(n+\frac{1}{2}\right)\hbar\omega$$

13.14 BARRIER PENETRATION

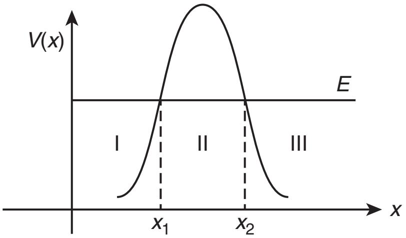


Fig. 13.10 Barrier penetration

We have already seen tunnelling phenomenon for rectangular potential for which exact solutions are available. For tunnelling involving other potentials, we have to resort to approximations. The most important application for WKB approximation is the tunnelling problem for an arbitrary potential. Let us consider the potential as shown in Fig. 13.10. Let *E* be the energy of the incident particle.

We have to solve the Schrödinger equation

$$\frac{d^2u}{dx^2}+\frac{2m}{\hbar^2}(E-V)u=0$$

Regions I and III are classically allowed. The boundary conditions for this problem are as follows:

1. There should be a wave function representing particles travelling to the right
2. There should be a wave function representing a linear combination of incident and reflected beams.

The barrier at x_1 is a right-handed barrier (the barrier is at the right of the classical turning point x_1). Therefore, in the WKB approximation, the solution u_I is given by

$$u_I=\frac{C_1}{p^{1/2}}\cos\left(-\frac{1}{\hbar}\int_{x_1}^{x}pdx'-\frac{\pi}{4}\right)+\frac{C_2}{p^{1/2}}\sin\left(-\frac{1}{\hbar}\int_{x_1}^{x}pdx'-\frac{\pi}{4}\right) \tag{13.152}$$

$$u_{II}=\begin{cases}\dfrac{C_3}{|p|^{1/2}}\exp\left(\dfrac{1}{\hbar}\displaystyle\int_{x_1}^{x}|p|\,dx'\right)+\dfrac{C_4}{|p|^{1/2}}\exp\left(-\dfrac{1}{\hbar}\displaystyle\int_{x_1}^{x}|p|\,dx'\right) & (13.153)\\ \text{OR} & \\ \dfrac{C_3'}{|p|^{1/2}}\exp\left(\dfrac{1}{\hbar}\displaystyle\int_{x_2}^{x}|p|\,dx'\right)+\dfrac{C_4'}{|p|^{1/2}}\exp\left(-\dfrac{1}{\hbar}\displaystyle\int_{x_2}^{x}|p|\,dx'\right) & (13.154)\end{cases}$$

The barrier at x_2 is a left-handed barrier.

$$\therefore u_{III}=\frac{C_5}{p^{1/2}}\cos\left(\frac{1}{\hbar}\int_{x_2}^{x}pdx'-\frac{\pi}{4}\right)+\frac{C_6}{p^{1/2}}\sin\left(\frac{1}{\hbar}\int_{x_2}^{x}pdx'-\frac{\pi}{4}\right) \tag{13.155}$$

Note that we have two terms (13.153) and (13.154) for the wave function in region II. The reason is that that we can obtain the WKB expansion using any classical turning point. We have used classical turning points $x_1$for (13.153) and x_2(13.154) as can be seen in the lower limits of the integrals in these two expressions. The boundary condition for region III requires a travelling wave to the right. If the potential is constant, u_{III} will have the form $e^{ipx/\hbar}$. This suggests $C_6 = iC_5$.

$$\therefore u_{III} = \frac{C_5}{p^{1/2}}\left[\cos\left(\frac{1}{\hbar}\int_{x_2}^{x} pdx' - \frac{\pi}{4}\right) + i\sin\left(\frac{1}{\hbar}\int_{x_2}^{x} pdx' - \frac{\pi}{4}\right)\right]$$

$$= \frac{C_5}{p^{1/2}}\exp\left(\frac{i}{\hbar}\int_{x_2}^{x} pdx' - \frac{\pi}{4}\right) \tag{13. 156}$$

From the connection formulae (13.140) and (13.141) for the left-handed barrier at x_2, we have

$$C_3' = \frac{C_5}{2} \quad \text{and} \quad C_4' = -C_6 = -iC_5$$

$$\therefore u_{II} = \frac{1}{2\,|p|^{1/2}}\left(C_5\exp\left(\frac{1}{\hbar}\int_{x_2}^{x}|p|\,dx'\right) - iC_5\exp\left(-\frac{1}{\hbar}\int_{x_2}^{x}|p|\,dx'\right)\right) \tag{13.157}$$

The wave function in the region should be the same whether it is written as (13.153) or (13.154). This fact can be used to get C_3 and C_4 in terms of C_3' and C_4', for which we need to get the integral $\int_{x_2}^{x}$ in terms of $\int_{x_1}^{x}$.

We have

$$\int_{x_1}^{x_2} = \int_{x_1}^{x} + \int_{x}^{x_2} = \int_{x_1}^{x} - \int_{x_2}^{x} \tag{13.158}$$

Let us define Λ as

$$\Lambda = \exp\left(\frac{1}{\hbar}\int_{x_1}^{x_2}|p|\,dx'\right)$$

$$\therefore \exp\left(\frac{1}{\hbar}\int_{x_2}^{x}|p|dx'\right) = \exp\left(\frac{1}{\hbar}\int_{x_1}^{x}|p|dx'\right)\exp\left(-\frac{1}{\hbar}\int_{x_1}^{x_2}|p|dx'\right) \qquad \text{(Use (13.158))}$$

$$= \frac{\exp\left(\frac{1}{\hbar}\int_{x_1}^{x}|p|\,dx'\right)}{\Lambda} \tag{13.159}$$

Similarly,

$$\exp\left(-\frac{1}{\hbar}\int_{x_2}^{x}|p|\,dx'\right) = \Lambda\exp\left(-\frac{1}{\hbar}\int_{x_1}^{x}|p|\,dx'\right) \tag{13.160}$$

Therefore, the Equation (13.154) becomes

$$u_{II} = \frac{C_3'}{2|p|^{1/2}}\cdot\frac{1}{\Lambda}\exp\left(\frac{1}{\hbar}\int_{x_1}^{x} pdx'\right) + \frac{C_4'\Lambda}{p^{1/2}}\exp\left(-\frac{1}{\hbar}\int_{x_1}^{x} pdx'\right)$$

$$\therefore \quad C_3 = C_3'\frac{1}{\Lambda} = C_5\frac{1}{\Lambda} \qquad \text{and} \qquad C_4 = C_4'\Lambda = -iC_5\frac{1}{\Lambda}$$

$$\therefore \quad u_{II} = \frac{C_5}{2|p|^{1/2}}\cdot\frac{1}{\Lambda}\exp\left(\frac{1}{\hbar}\int_{x_1}^{x} pdx'\right) - \frac{-iC_5\Lambda}{p^{1/2}}\exp\left(-\frac{1}{\hbar}\int_{x_1}^{x} pdx'\right) \tag{13.161}$$

The barrier at x_1 is a right handed barrier and so we use the connection formulae (13.142) and (13.143).

Therefore, $$C_1 \text{ is } (-2iC_5\Lambda) \text{ and } C_2 \text{ is} \left(-\frac{C_5}{2\Lambda}\right)$$

$$\therefore u_I = \frac{-2iC_5\Lambda}{p^{1/2}}\cos\left(-\frac{1}{\hbar}\int_{x_1}^{x} pdx' - \frac{\pi}{4}\right) - \frac{C_5}{2\Lambda p^{1/2}}\sin\left(-\frac{1}{\hbar}\int_{x_1}^{x} pdx' - \frac{\pi}{4}\right) \tag{13.162}$$

Let us define ϕ as $$\phi = -\frac{1}{\hbar}\int_{x_1}^{x} pdx' - \frac{\pi}{4}.$$

$$u_I = \frac{-2iC_5\Lambda}{p^{1/2}}\frac{(e^{i\phi}+e^{-i\phi})}{2} - \frac{C_5}{p^{1/2}2\Lambda}\cdot\frac{(e^{i\phi}-e^{-i\phi})}{2i}$$

$$= \frac{iC_5}{p^{1/2}}e^{i\phi}\left(-\Lambda + \frac{1}{4\Lambda}\right) - \frac{iC_5}{p^{1/2}}e^{-i\phi}\left(\frac{1}{4\Lambda} + \Lambda\right)$$

To identify the factor representing incident beam, let us consider $e^{i\phi}$ and $e^{-i\phi}$ terms under constant potential.

$$\phi = -\frac{1}{\hbar}\int_{x_1}^{x} pdx' - \frac{\pi}{4} = -\frac{p}{\hbar}(x - x_1) - \frac{\pi}{4}$$

Obviously, $e^{-i\phi}$ represents the wave travelling from the left to the barrier. Therefore, we can write

$$u_I = u_{I,\,\text{inc}} + u_{I,\,\text{ref}}$$

where $$u_{I,\,\text{inc}} = -iC_5\left(\frac{1}{4\Lambda} + \Lambda\right)\frac{1}{p^{1/2}}\exp\left(\frac{i}{\hbar}\int_{x_1}^{x} pdx' + i\frac{\pi}{4}\right)$$

and $$u_{I,\,\text{ref}} = -iC_5\left(\frac{1}{4\Lambda} - \Lambda\right)\frac{1}{p^{1/2}}\exp\left(-\frac{i}{\hbar}\int_{x_1}^{x} pdx' - i\pi/4\right)$$

The transmission coefficient for the potential barrier is found to be

$$T = \frac{|C_5|^2}{\left|-iC_5\left(\frac{1}{4\Lambda} + \Lambda\right)\right|^2} = \frac{1}{\Lambda^2 + \frac{1}{2} + \frac{1}{16\Lambda^2}}$$

The dominant term in the denominator is Λ^2. Therefore, we have

$$T \simeq \frac{1}{\Lambda^2} \sim e^{-\frac{2}{\hbar}\int_{x_1}^{x_2}|p|dx'} \tag{13.163}$$

13.15 α DECAY AND WKB APPROXIMATION

We have already seen the α decay process as tunnelling phenomenon. The parent nucleus before decaying is considered as a system consisting of the daughter nucleus, and the α particle, i.e., the alpha particle is confined to move within the volume of the daughter nucleus. The strong interaction between the daughter nucleus and the alpha particle is restricted to a distance R, the radius of the daughter nucleus. Beyond R, the alpha particle experiences only Coulomb interaction $Ze \cdot Z'e/r$, where Ze and $Z'e$ are the charges of the daughter nucleus and the alpha particle, respectively. For an alpha particle of energy E, the regions I and III are classically allowed regions, and region II is a classically forbidden region.

Let R and r_0 be the turning points, R is the nuclear radius and r_0 is determined from the equation $ZZ'e^2/r_0 = E$.

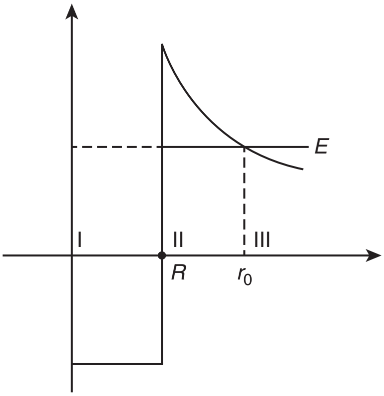


Fig. 13.11 Potential for α decay

The transmission coefficient T is given by

$$T = e^{-2\Lambda}$$

where

$$\Lambda = \frac{1}{\hbar}\int_R^{r_0} (2m)^{1/2}(V-E)^{1/2}\,dr$$

$$= \frac{(2m)^{1/2}}{\hbar}\cdot\int_R^{r_0}\left(\frac{ZZ'e^2}{r_0} - E\right)^{1/2} dr = \frac{(2mZZ'e^2)^{1/2}}{\hbar^2}\cdot\int_R^{r_0}\left(\frac{1}{r} - \frac{E}{ZZ'e^2}\right)^{1/2} dr$$

$$= \frac{(2mZZ'e^2)^{1/2}}{\hbar^2}\cdot\int_R^{r_0}\left(\frac{1}{r} - \frac{1}{r_0}\right)^{1/2} dr$$

To evaluate this integral, put $r = r_0 \cos^2\theta$

$$\therefore \int_R^{r_0}\left[\frac{1}{r_0}(\sec^2\theta - 1)\right]^{1/2}(-2r_0\cos\theta\sin\theta)d\theta = -2\sqrt{r_0}\int \sin^2\theta d\theta$$

$$= \sqrt{r_0}\int[-1+\cos^2\theta]d\theta = \sqrt{r_0}\left[-\theta + \sin\theta\cos\theta\right]_R^{r_0}$$

$$= \sqrt{r_0}\left[\cos^{-1}\frac{R}{r_0} - \cos^{-1}\frac{r_0}{r_0} - \left(\frac{R}{r_0}\right)^{1/2}\left(1-\left(\frac{R}{r_0}\right)\right)^{1/2}\right]$$

$$= \sqrt{r_0}\left[\cos^{-1}\left(\frac{R}{r_0}\right)^{1/2} - \left(\frac{R}{r_0}\right)^{1/2}\left(1-\left(\frac{R}{r_0}\right)\right)^{1/2}\right]$$

The Taylor's expansion for $\cos^{-1} x$ is

$$\cos^{-1} x \simeq \frac{\pi}{2} - x$$

Since $R/r_0 << 1$ for a nucleus, we have

$$\int_R^{r_0}\left(\frac{1}{r_0}-\frac{1}{r_0}\right)^{1/2} dr \simeq \sqrt{r_0}\left[\frac{\pi}{2}-\left(\frac{R}{r_0}\right)^{1/2}-\left(\frac{R}{r_0}\right)^{1/2}\right]$$

$$\therefore \quad \Lambda = \left[\frac{2mZZ'e^2}{\hbar^2}\cdot r_0\right]^{1/2}\left(\frac{\pi}{2}-2\left(\frac{R}{r_0}\right)^{1/2}\right)$$

Substituting the expression $ZZ'e^2/E$ for r_0 in the above equation, we get

$$\Lambda = \frac{\pi}{\sqrt{2}}\frac{ZZ'e^2}{E^{1/2}} - \left(\frac{8ZZ'e^2R}{\hbar^2}\right)^{1/2}$$

$$\therefore \quad T = e^{-2\Lambda} = \underbrace{\exp\left[-\frac{\sqrt{2}\pi ZZ'e^2}{\sqrt{E}}\right]}_{T_G}\cdot\exp\left[(32ZZ'e^2R/\hbar^2)^{1/2}\right]$$

where T_G is known as Gamow factor. T is the probability of a particle to be transmitted whenever it collides with the Coulomb barrier. This implies that if $1/T$ collisions take place, the probability of an α particle coming out of the barrier is $T \cdot 1/T \sim 1$. Therefore, $1/T$ collisions are required for the emission of an alpha particle. If v is the velocity of an α particle, then $v/2R$ collisions will take place in 1 sec. Then, the time required to make $1/T$ collision is

$$\tau = \frac{1}{T}\cdot\frac{2R}{v}$$

where τ is the lifetime of the α particle. It is given by

$$\tau = \frac{2R}{v}\cdot\exp\left[-(32ZZ'e^2R/\hbar^2)^{1/2}\right]\cdot\exp\left[\sqrt{2}\pi ZZ'e^2/E^{1/2}\right]$$

$$\therefore \ln\tau = \ln\frac{2R}{\upsilon} - \frac{(32ZZ'e^2R)^{1/2}}{\hbar^2} + \sqrt{2}\pi\frac{ZZ'e^2}{E^{1/2}}$$

$$= \frac{C_1}{E^{1/2}} - C_2$$

Experimentally, the functional form of this equation is found to hold good, though the numerical values of C_1 and C_2 obtained from the equations are not the same as the experimental value.

EXERCISES

1. $|u_n\rangle$ is an eigenstate of H_0 with energy eigenvalue $E_n^{(0)}$. Assume $E_n^{(0)}$ is non-degenerate. Do we need $|u_n\rangle$ to be eigenstates of $\lambda H'$ to do perturbation theory?
2. In the previous question, suppose $E_n^{(0)}$ is degenerate. Why do we need the eigenstates of H_0 to be eigenstates of $\lambda H'$ also?
3. Show that the correction due to second-order perturbation theory for a non-degenerate ground state is always negative.
4. Show that there is no shift in energy due to Stark effect for the ground state of hydrogen atom.
5. Determine energy levels of a hydrogen atom when H_r and H_{LS} are included.
6. What are the basis states suitable to describe the hydrogen atom when H_r and H_{LS} are included as perturbation?
7. What are the suitable basis states to describe hydrogen atom when $H_B = -\frac{e}{2mc}(\mathbf{L}+2\mathbf{S})\cdot\mathbf{S}$ is the perturbation?
8. Determine the energy splitting for hydrogen atom corresponding to $n = 1$, 2 and 3 H_r, H_{LS} and H_r, $+ H_{LS}$.

REFERENCES

1. P. M. Mathews and A. Venkatesan, 1976. *A Text Book of Quantum Mechanics.* New Delhi: Tata-McGraw Hill Publishing Company Limited.
2. S. Gaziorowicz, 1974. *Quantum Physics.* New York: John Wiley & Sons, Inc.
3. R. L. Liboff, 1980. *Introduction to Quantum Mechanics*. San Francisco: Holden-Day.
4. B. H. Bransden and C. J. Joachain, 1989. *Introduction to Quantum Mechanics*. London: English Language Book Society/Longman.
5. David Park, 1974. *Introduction to the Quantum Theory.* Tokyo: McGraw-Hill Kogakusha Ltd.
6. A. S. Davydov, 1965. *Quantum Mechanics.* Oxford: Pergamon Press.

14

Time-Dependent Perturbation Theory

When the Hamiltonian is independent of time, one obtains the stationary state solution. The solution to the time-dependent Schrödinger equation in such cases is restricted to the dynamical phase factor $e^{-iE_n t/\hbar}$. When the Hamiltonian becomes time dependent, the solution to the Schrödinger equation becomes more complex. In only a few cases, exact analytical solutions are available, and hence we have to resort to approximation methods as we did in the case of time-independent Schrödinger equation. Three kinds of approximations have been presented here. They are time-dependent perturbation theory, adiabatic approximation and sudden approximation.

14.1 TIME-DEPENDENT PERTURBATION THEORY–BASIS

Let us write the Hamiltonian H as

$$H = H_0 + \lambda H'(t) \tag{14.1}$$

where H_0 is time-independent part and H' is time-dependent part. Let us assume the exact analytical solutions are available for H_0. Let $\{u_n(\mathbf{r})\}$ be the eigenfunctions of H_0.

i.e.,

$$H_0 u_n = E_n u_n \tag{14.2}$$

The corresponding time-dependent part is given by

$$\phi_n = e^{-iE_n t/\hbar} u_n(\mathbf{r}) \tag{14.3}$$

The wave function ϕ_n satisfies the Schrödinger equation

$$i\hbar \frac{d\phi_n}{dt} = H_0 \phi_n \tag{14.4}$$

Both the sets $\{u_n\}$ and $\{\phi_n\}$ form a complete set. Let ψ be the solution to the time-dependent Schrödinger equation.

$$i\hbar \frac{d\psi}{dt} = H\psi = (H_0 + \lambda H')\psi \tag{14.5}$$

Then $\psi(\mathrm{t})$ can be written as

$$\psi = \sum C_n(t)\phi_n(t) \tag{14.6}$$

The complete description of the state of the system is now given by $C_n(t)$ or more precisely by $\begin{pmatrix} C_1(t) \\ C_2(\mathrm{t}) \\ \vdots \end{pmatrix}$.

The study of time evolution of the state, therefore, amounts to determine the function $C_n(t)$. So our task is to find the set $\{C_n(t)\}$ for a set of given initial conditions $\{C_n(0)\}$.

i.e., for a given set $\begin{bmatrix} C_1(0) \\ C_2(0) \\ \vdots \end{bmatrix}$, what is the set $\begin{bmatrix} C_1(t) \\ C_2(t) \\ \vdots \end{bmatrix}$?

Let us start by substituting the Equation (14.6) in (14.5).

LHS:

$$i\hbar\frac{d\psi}{dt} = \sum_k i\hbar\dot{C}_k\phi_k + \sum_k C_k i\hbar\dot{\phi}_k \tag{14.7}$$

RHS:

$$(H_0 + \lambda H')\psi = \sum_k C_k H_0\phi_k + \sum_k \lambda C_k H'\phi_k \tag{14.8}$$

Equating these equations, we get

$$\sum_k \dot{C}_k\phi_k = \frac{1}{i\hbar}\sum_k C_k \lambda H' \lambda\phi_k \tag{14.9}$$

Taking scalar product with ϕ_n, we get

$$\sum_k \dot{C}_k\langle\phi_n|\phi_k\rangle = \frac{1}{i\hbar}\sum_k \lambda C_k\langle\phi_n|H'|\phi_k\rangle$$

Since $\langle\phi_n|\phi_k\rangle = \delta_{nk}$, we get

$$\dot{C}_n = \frac{1}{i\hbar}\sum_k \lambda C_k\langle\phi_n|H'|\phi_k\rangle \tag{14.10}$$

These are the sets of coupled equations. All that we have done is to convert the differential equation for ψ into a set of coupled differential equations for $C_n(t)$.

To suit the perturbation theory, let us expand $C_n(t)$ as a power series in λ.

$$C_n(t) = C_n^{(0)}(t) + \lambda C_n^{(1)}(t) + \lambda^2 C_n^{(2)}(t)\cdots \tag{14.11}$$

$C_n^{(k)}$ is given by

$$C_n^{(k)} = \frac{1}{k!}\frac{d^k C_n(t)}{dt^k}\bigg|_{t=0}$$

Substituting the Equation (14.11) in (14.10), we get

$$\dot{C}_n^{(0)} + \lambda \dot{C}_n^{(1)} + \lambda^2 \dot{C}_n^{(2)} + \cdots = \frac{1}{i\hbar} \sum_k \lambda H'_{nk} \left[C_k^{(0)} + \lambda C_k^{(1)} + \lambda^2 C_k^{(2)} \cdots \right]$$

Equating the coefficients of different power of λ, we get

$$\dot{C}_n^{(0)} = 0 \tag{14.12}$$

$$i\hbar \dot{C}_n^{(1)} = \sum C_k^{(0)} \langle \phi_n | H' | \phi_k \rangle \tag{14.13}$$

$$i\hbar \dot{C}_n^{(2)} = \sum C_k^{(1)} \langle \phi_n | H' | \phi_k \rangle \tag{14.14}$$

$$\vdots \qquad \vdots$$

Given $C_k^{(0)}$, the coefficient $C_k^{(1)}$ is determined from the Equation (14.13). Knowing from $C_k^{(1)}$, $C_n^{(2)}$ can be determined and so on. So, in principle, we can solve all the coupled equations. In practice, we will restrict ourselves to the first order.

The solution to the Equation (14.12) is very simple.

$$C_n^{(0)}(t) = \text{const}$$

The initial time may be chosen as $t = -\infty$ or $t = 0$ or simply t_0. Let us choose $t = 0$ as the initial time. Since $C_n^{(0)}(t) = \text{const}$, we have

$$C_n^{(0)}(t) = C_n^{(0)}(0) \tag{14.15}$$

Let us now consider $C_n(t)$ at $t = 0$.

$$C_n(0) = C_n^{(0)}(0) + \underbrace{\lambda C_n^{(1)}(0) + \lambda^2 C_n^{(2)}(0) \cdots}_{\text{terms due to interaction}} \tag{14.16}$$

Since $t = 0$ is chosen as the initial time, the perturbation itself starts at $t = 0$. Therefore, the terms due to interaction in (14.16) are zero.

$$C_n^{(1)}(0) = C_n^{(2)}(0) = C_n^{(3)}(0) = \cdots = 0 \tag{14.17}$$

$$\therefore C_n(0) = C_n^{(0)}(0) \tag{14.18}$$

Let us assume that the system starts in the initial state u_i.

i.e., $$\psi(0) = \phi_i(0) = u_i \tag{14.19}$$

$$C_n^{(0)} = \begin{cases} 1 \text{ for } n = i \\ 0 \text{ for } n \neq i \end{cases} \tag{14.20}$$
$$= \delta_{ni}$$

or
$$C_n^{(0)}(0) = \delta_{ni} \quad \text{or } C_n^{(0)}(t) = \delta_{ni} \tag{14.21}$$

14.2 TRANSITION PROBABILITY

Before going into the first-order perturbation theory, let us explain how the concept of transition probability arises in time-dependent perturbation theory. This is the most significant consequence of time-dependent perturbation theory.

Let us recall that $\psi(\mathbf{r},t)$ is given by

$$\psi(\mathbf{r},t) = \sum_n C_n(t)\phi_n(\mathbf{r},t) = \sum_n C_n(t)e^{-i\,E_n t/\hbar}u_n(\mathbf{r})$$

Now, by choice, we have assumed that the system starts in state $u_i(\mathbf{r})$ at $t = 0$.

$$\therefore \psi(\mathbf{r},0) = \sum_n C_n(0)u_n(\mathbf{r}) = u_i(\mathbf{r})$$
$$\therefore C_i(0) = 1$$

and
$$C_n(0) = 0 \text{ if } n \neq i$$

We know that $|C_n(t)|^2$ or $(|C_n(t)e^{-iE_n t/\hbar}|^2)$ gives the probability of finding the system (which starts in state u_i at $t = 0$) in state u_n at time t. Therefore, $|C_n(t)|^2$ is interpreted as the transition probability to make a transition from the initial state u_i (at $t = 0$) to a final state u_n (at t); i.e., the transition probability $P_{i\to n}$ is given by $P_{i\to n} = \left|C_n(t)\right|^2$.

Note the difference between the time-independent and time-dependent perturbation theories. In time-independent perturbation, the eigenstates of both H_0 and $H = H_0 + \lambda H'$ are stationary eigenstates. The effect of perturbation $\lambda H'$ is to produce the correction to the eigenvalue of stationary eigenstate of H_0. In time-dependent perturbation theory, we get a transition from the eigenstate of H_0 to another eigenstate of H_0.

14.3 FIRST-ORDER PERTURBATION

From the Equation (14.13), we have

$$\frac{dC_n^{(1)}}{dt} = \frac{1}{i\hbar}\sum_k C_k^{(0)}(t)\langle\phi_n|H'|\phi_k\rangle$$
$$= \frac{1}{i\hbar}\sum_k \delta_{ki}\langle\phi_n|H'|\phi_k\rangle = \frac{1}{i\hbar}\langle\phi_n|H'|\phi_i\rangle$$

Integrating both sides, we get

$$\int_0^t \frac{dC_n^{(1)}(t')}{dt'}dt' = \frac{1}{i\hbar}\int_0^t \langle\phi_n|H'|\phi_i\rangle dt'$$

LHS $= C_n^{(1)}(t) - C_n^{(1)}(0) = C_n^{(1)}(t)$

$$\begin{aligned} C_n^{(1)}(t) &= \frac{1}{i\hbar}\int_0^t \langle\phi_n|H'|\phi_i\rangle dt' \\ &= \frac{1}{i\hbar}\int_0^t \langle u_n|\, e^{i\omega_n t'} H' e^{-i\omega_i t'}\, |u_i\rangle\, dt' \\ &= \frac{1}{i\hbar}\int_0^t e^{i\omega_{ni}t}\langle u_n|H'|u_i\rangle dt' \end{aligned} \tag{14.22}$$

where $\omega_{ni} = [E_n - E_i]/\hbar$

Let us emphasize again that the transition probability $P_{i\to n}$ is given by

$$P_{i\to n} = \left|C_n^{(1)}(t)\right|^2$$

14.3.1 Second-order Perturbation

From the Equation (14.14), we have

$$\dot{C}_n^{(2)} = \sum_k \langle\phi_n|H'|\phi_k\rangle C_k^{(1)}$$

$$\begin{aligned} C_n^{(2)} &= \sum_k \frac{1}{i\hbar}\int_0^t dt_1 \langle\phi_n(t_1)|H'(t_1)|\phi_k(t_1)\rangle C_k^{(1)}(t_1) \\ &= \sum_k \frac{1}{(i\hbar)^2}\int_0^t dt_1 \langle\phi_n(t_1)\,|H'(t_1)|\phi_k(t_1)\rangle \times \int_0^{t_1} \langle\phi_k(t_2)\,|H'(t_2)|\phi_i(t_2)\rangle dt_2 \\ &= \sum_k \frac{1}{(i\hbar)^2}\int_0^t e^{i\omega_{nk}t_1}\langle u_n|H'|u_k\rangle dt_1 \int_0^{t_1} e^{i\omega_{ki}t_2}\langle u_k|H'|u_i\rangle dt_2 \end{aligned} \tag{14.23}$$

14.4 PRINCIPLE OF DETAILED BALANCING

Let us recall that the system starts in state $u_i(\mathbf{r})$ at $t = 0$ and $C_n^{(1)}(t)$ is the probability amplitude for the system to make a transition to the state $u_n(\mathbf{r})$ at time t. So, let us write the Equation (14.22) as

$$C_{i\to n}^{(1)}(t) = \frac{1}{i\hbar}\int_0^t e^{i\omega_{ni}t'}\langle u_n|H'(t)|u_i\rangle dt' \tag{14.24}$$

The transition probability from state $u_i(\mathbf{r})$ to final state $u_n(\mathbf{r})$ is given by

$$P_{i\to n}(t) = |\lambda C^{(1)}_{i\to n}(t)|^2 \tag{14.25}$$

Now let us consider the reverse situation. The system starts in state $u_n(\mathbf{r})$at $t = 0$. The transition probability amplitude to make the transition to state $u_i(\mathbf{r})$ is given by

$$C^{(1)}_{n\to i}(t) = \frac{1}{i\hbar}\int_0^t e^{i\omega_{in}t'}\langle u_i|H'(t)|u_n\rangle dt \tag{14.26}$$

The transition probability $P_{n\to i}$ is given by

$$P_{n\to i} = \left|\lambda C^{(1)}_{n\to i}(t)\right|^2 \tag{14.27}$$

Since H' is a Hermitian operator,

$$\langle u_n|H'(t)|u_i\rangle = \langle u_i|H'(t)|u_n\rangle^*$$

$$\begin{aligned}
(C^{(1)}_{i\to n}(t))^* &= -\frac{1}{i\hbar}\int_0^t e^{-i\omega_{ni}t'}\langle u_n|H'(t)|u_i\rangle^* dt' \\
&= -\frac{1}{i\hbar}\int_0^t e^{i\omega_{in}t'}\langle u_i|H'(t)|u_n\rangle dt' \\
&= -C^{(1)}_{n\to i}(t)
\end{aligned} \tag{14.28}$$

Therefore, we have

$$P_{i\to n} = P_{n\to i} \tag{14.29}$$

Therefore, for a given Hamiltonian $H'(t)$, the transition probability from the initial state $u_i(\mathbf{r})$ to the final state $u_n(\mathbf{r})$ is the same as the transition probability from the initial state $u_n(\mathbf{r})$ to the final state $u_i(\mathbf{r})$.

This result is known as the principle of detailed balance.

14.5 CONSTANT PERTURBATION

Consider a time-dependent perturbation H' given by

$$H' = \begin{cases} 0 & t \le 0 \\ V(\mathbf{r}) & t \ge 0 \end{cases} \tag{14.30}$$

The only time dependence is that it is switched on at time $t = 0$. Once it is switched on, H' has no time dependence subsequently. H' may be written as

$$H' = V(\mathbf{r})\theta(t)$$

where $\theta(t)$ is a step function.

$$\theta(t) = \begin{cases} 0 & t \le 0 \\ 1 & t \ge 0 \end{cases}$$

In the first-order perturbation theory,

$$C_n^{(1)}(t) = \frac{1}{i\hbar}\int_0^t \langle \phi_n | H' | \phi_i \rangle dt' = \frac{1}{i\hbar}\int_0^t e^{i\omega_{ni}t'} \langle u_n | V(\mathbf{r}) | u_i \rangle dt' \tag{14.31}$$

Since $V(\mathbf{r})$ is independent of time, the matrix element $V_{ni} = \langle u_n | V(\mathbf{r}) | u_i \rangle$ is independent of time.

$$\begin{aligned} C_n^{(1)}(t) &= \frac{1}{i\hbar} V_{ni} \int_0^t e^{i\omega_{ni}t'} dt' = \frac{1}{i\hbar}\frac{1}{i\omega_{ni}} V_{ni}(e^{i\omega_{ni}t} - 1) \\ &= \frac{1}{i\hbar}\cdot\frac{1}{i\omega_{ni}} V_{ni} e^{i\omega_{ni}t/2}[e^{i\omega_{ni}t/2} - e^{-i\omega_{ni}t/2}] \\ &= \frac{1}{i\hbar\omega_{ni}} V_{ni} e^{i\omega_{ni}t/2} 2\sin\frac{\omega_{ni}t}{2} \end{aligned}$$

$$\begin{aligned} \therefore \left|C_n^{(1)}(t)\right|^2 &= \frac{4\cdot\left|V_{ni}\right|^2}{\hbar^2\omega_{ni}^2}\cdot\sin^2(\omega_{ni}t/2) \\ &= \frac{\left|V_{ni}\right|^2}{\hbar^2}\cdot\frac{\sin^2 \omega_{ni}t/2}{(\omega_{ni}/2)^2} \end{aligned}$$

Note that $C_n(t) = \lambda C_n^{(1)}(t)$ for $n \ne i$. We have already seen that the transition probability from state u_i to state u_n is $\left|C_n(t)\right|^2$. Therefore, the transition probability $P_{i\to n}$ is given by

$$P_{i\to n} = \frac{\left|\lambda V_{ni}\right|^2}{\hbar^2}\cdot\frac{\sin^2 \omega_{ni}t/2}{(\omega_{ni}/2)^2} \tag{14.32}$$

Let us sketch the function $\frac{\sin^2 \omega_{ni}t/2}{\omega_{ni}^2}$. In the limit $t \to 0$, $\frac{\sin^2 \omega_{ni}t/2}{\omega_{ni}^2} \sim \frac{\omega_{ni}^2 t^2}{4\omega_{ni}^2} = t^2/4$. The function is zero for

$$\frac{\omega_{ni}t}{2} = \pi, 2\pi, 3\pi \ldots \tag{14.33}$$

A sketch of this function is shown in Fig.14.1.

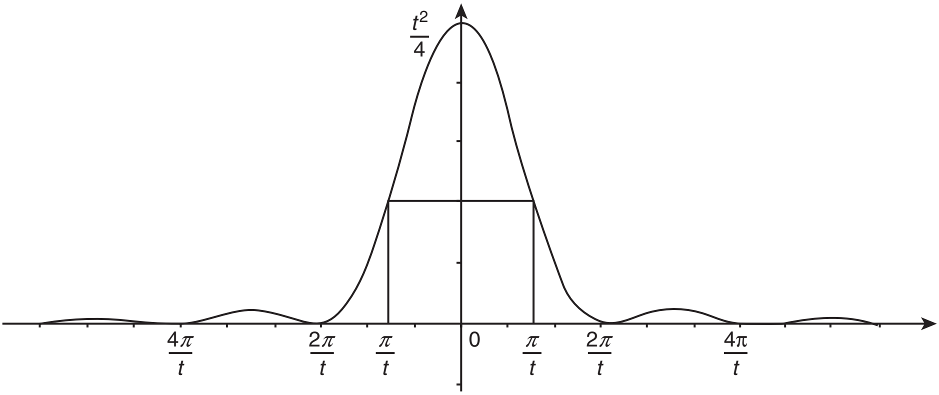


Fig. 14.1 Sketch of $\frac{\sin^2 \omega_{ni} t/2}{\omega_{ni}^2}$

This function has got a peak at $\omega_{ni} = 0$, and the central peak has a half-width $2\pi/t$ (width at half the maximum). The transition probability is significant only in the range $|\omega_{ni}| \sim 2\pi/t$ or $|E_n - E_i| \sim 2\pi\hbar/t$. This has got an interesting interpretation. Let us define $|E_n - E_i|$ as ΔE and introduce the notation Δt for t. Then the expression $|E_n - E_i| \sim 2\pi\hbar/t$ becomes

$$\Delta E \sim \frac{2\pi\hbar}{\Delta t} \tag{14.34}$$

$$\Delta E \Delta t \sim 2\pi\hbar \tag{14.35}$$

ΔE is interpreted as the uncertainty in energy ΔE, and Δt is the time for which the perturbation acts on the system, and the relation (14.35) is the well-known time energy uncertainty relation.

We have already seen the difference between this uncertainty relation and the relation $\Delta x \Delta p_x \sim h$ in Chapter 4. Since x and p_x are dynamical variables, we can associate Hermitian operators with them, and it is the non-commutation between these operators which play an important role in establishing the concept of uncertainties for Δx and Δp_x. But time t is not a dynamical variable. So though mathematically we can have the relation $[t, i\hbar\, \partial/\partial t] = i\hbar$, this relation has no meaning in quantum mechanics. So Δt cannot be interpreted as uncertainty in time.

Let us try to understand ΔE. The system starts with a definite energy E_i. So the initial state of the system is the eigenstate of H_0. Due to the presence of time-dependent perturbation H', the state of the system becomes a superposition of eigenstates of H_0 at a later instant time t. So, with respect to H_0, the energy of the system is uncertain. There is a definite probability that the energy of the system will be different from E_i. So ΔE can be interpreted as the statistical spread in the energy of the system resulting from the perturbation H'. However, Δt is not such a spread. Δt or, more appropriately, t is the time interval for which the perturbation has acted. As t increases, the width $\Delta E \sim 2\pi\hbar/\Delta t$ decreases.

Finding the system in states with energies other than initial energy E_i is also sometimes interpreted as violation of the conservation of energy. Perturbation H' is interpreted as the measurement process, and getting a value other than E_i looks like violation of conservation of energy. However, it is generally argued that if we include energy of the perturbation source also, there will be no violation of conservation of energy.

14.6 FERMI'S GOLDEN RULE

The expression (14.32) for $P_{i\to n}$ gives the transition probability for a transition from the initial state u_i to a specific final state u_n. Now, let us extend this discussion to consider the transition to group of closely spaced final states, for which the energy eigenvalues are centered around E_n. The energy eigenvalues of this group of states are between $E_n - \Delta$ and $E_n + \Delta$. Let us calculate the transition probability for the transition from state u_i to this group of states.

$$
\begin{aligned}
P_{i\to[n]} &= \sum_{E_n-\Delta}^{E_n+\Delta} P_{i\to m} = \sum \left|\lambda C_{mi}^{(1)}\right|^2 \\
&= \sum \frac{\left|\lambda V_{mi}\right|^2 \sin^2\left(\dfrac{E_m - E_i}{2\hbar}\right)t}{\hbar^2\left(\dfrac{E_m - E_i}{2\hbar}\right)^2}
\end{aligned}
$$

This is a sum over discrete energy eigenvalues. If the energy eigenvalues are very closely spaced, we can replace the summation by integration. The standard technique is

$$\sum P_{i\to m} \to \int P_{i\to m}\rho(E_m)dE_m \tag{14.36}$$

The Equation (14.36) can be explained as follows. Let us divide the energy range between $E_n - \Delta$ and $E_n + \Delta$ into many sub-intervals dE_m. All the states within the energy range E_m and $E_m + dE_m$ have almost the same values for $P_{i\to m}$. Therefore, the summation of $P_{i\to m}$ for these states can be replaced by $\rho(E_m)dE_m P_{i\to m}$. The integral represents the sum over dE_m, covering all the states within the range $E_n - \Delta$ to $E_n + \Delta$. The quantity $\rho(E_m)$ is called density of states; i.e., $\rho(E_m)$ is the number of states per unit energy range between E_m and $E_m + dE_m$. So $\rho(E_m)dE_m$ gives the number of states between E_m and $E_m + dE_m$.

$$
\begin{aligned}
\therefore P_{i\to[n]} &= \int_{En-\Delta}^{En+\Delta} P_{i\to m}\rho(E_m)dE_m \\
&= \int_{E_n-\Delta}^{E_n+\Delta} \frac{|\lambda V_{mi}|^2}{\hbar^2} \frac{\sin^2\left(\dfrac{E_m - E_i}{2\hbar}\right)t}{[(E_m - E_i)/2\hbar]^2}\rho(E_m)dE_m
\end{aligned}
\tag{14.37}
$$

Let us further assume $|V_{mi}|^2 \simeq |V_{ni}|^2$ and $\rho(E_m) \simeq \rho(E_n)$ for all the states in the energy range $(E_n - \Delta) < E_m < (E_n + \Delta)$. So we can pull $\rho(E_n)$ and $|V_{ni}|^2$ outside the integral

$$\therefore P_{i\to[n]} = \frac{|V_{ni}|^2}{\hbar^2}\rho(E_n)\int_{E_n-\Delta}^{E_n+\Delta} \frac{\sin^2\left(\dfrac{E_m - E_i}{2\hbar}\right)t}{[(E_m - E_i)/2\hbar]^2}dE_m \tag{14.38}$$

The integrand $(\sin^2(E_m - E_i)t/2\hbar)/[(E_m - E_i)/2\hbar]^2$ is significant only a small range and is zero practically outside the range $E_n - \Delta < E_n < E_n + \Delta$. Therefore, we can make the limits of the integral from $-\infty$ to $+\infty$. i.e., $\int_{E_n-\Delta}^{E_n+\Delta} dE_m \simeq \int_{-\infty}^{\infty} dE_m$

$$\therefore P_{i\to[n]} = \frac{|\lambda V_{ni}|^2\ t^2}{\hbar^2} \int_{-\infty}^{\infty} \frac{\sin^2(E_m - E_i)t/2\hbar}{[(E_m - E_i)t/2\hbar]^2}\, dE_m$$

Define x as

$$\frac{(E_m - E_i)t}{2\hbar} = x$$

$$P_{i\to[n]} = \frac{|V_{ni}|^2}{\hbar}\rho(E_n)2t \int_{-\infty}^{\infty} \frac{\sin^2 x}{x} dx$$

$$= \frac{2\pi}{\hbar}|V_{ni}|^2\ t\rho(E_n)$$

We have to bear in mind that the integral (14.38) is significantly different from zero only for $E_n \sim E_i$, especially for large t. Therefore, write $P_{i\to[n]}$ as

$$P_{i\to[n]} = \frac{2\pi}{\hbar}|V_{ni}|^2\ t\ \rho(E_n)\Big|_{E_n \approx E_i} \tag{14.39}$$

The transition probability to a group of final states is now proportional to time t, and so the transition probability per unit time is given by

$$W_{i\to[n]} = \frac{2\pi}{\hbar}|V_{ni}|^2\ \rho(E_n)\Big|_{E_n \sim E_i}$$

This is known as Fermi's golden rule.

14.6.1 Aliter

We can get the same results in a different way. Let us start with the expression

$$P_{i\to n} = \frac{|\lambda V_{ni}|^2}{\hbar^2} \frac{\sin^2 \omega_{ni}t/2}{(\omega_{ni}/2)^2} \tag{14.40}$$

Let us write this expression as

$$P_{i\to n} = \frac{|\lambda V_{ni}|^2}{\hbar^2} t \frac{4\sin^2 \omega_{ni}t/2}{\omega_{ni}^2 t}$$

For very large values of t,

$$P_{i\to n} = \frac{|\lambda V_{ni}|^2}{\hbar^2} t\, 2\pi\delta(\omega_{ni})$$

where we have used the fact

$$\underset{t\to\infty}{\mathrm{Lt}} \frac{\sin^2(\omega_{ni}t/2)}{\left(\frac{\omega_{ni}}{2}\right)^2 t} = 2\pi\delta(\omega_{ni})$$

$$\begin{aligned}\therefore P_{i\to n} &= \frac{|\lambda V_{ni}|^2}{\hbar^2}\cdot t\cdot 2\pi\delta\left(\frac{E_n - E_i}{\hbar}\right)\\ &= \frac{|\lambda V_{ni}|^2\, 2\pi t}{\hbar}\delta(E_n - E_i)\end{aligned}$$

We have used the fact $\delta(ax) = 1/|a|\,\delta(x)$. We have to note that this expression is zero when $E_n \neq E_i$ and infinite at $E_n = E_i$. This reflects the fact that we are using the Equation (14.40) beyond the range of its validity. Still this is a good approximation in view of the presence of Dirac delta function, which is meaningful only in the context of some integrals.

Indeed, such a situation arises when we consider the transition probability from initial state u_i to a group of closed, spaced final states.

$$\begin{aligned}P_{i\to[n]} &= \sum P_{i\to m} = \int P_{i\to m}\rho(E_m)dE_m\\ &= \frac{2\pi t}{\hbar}\int |\lambda V_{ni}|^2 \rho(E_n)\delta(E_n - E_i)dE_n\\ &= \frac{2\pi t}{\hbar}|V_{ni}|^2\ \rho(E_n)\,|_{E_n \simeq E_i}\end{aligned}$$

This is the same result as the Equation (14.40). We summarize all these relations below as they are very useful in the subsequent calculations:

$$P_{i\to n} = \frac{|\lambda V_{ni}|^2}{\hbar} 2\pi t\delta(E_n - E_i) \tag{14.41}$$

$$P_{i\to[n]} = \frac{2\pi t}{\hbar}|\lambda V_{ni}|^2\rho(E_n)|_{(E_n \sim E_i)} \tag{14.42}$$

$$W_{i\to n} = \frac{2\pi}{\hbar}|\lambda V_{ni}|^2\delta(E_n - E_i) \tag{14.43}$$

$$W_{i\to[n]} = \frac{2\pi}{\hbar}|\lambda V_{ni}|^2\rho(E_n)|_{E_n \simeq E_i} \tag{14.44}$$

14.7 HARMONIC PERTURBATION

The general form of the Hamiltonian for harmonic perturbation is given by

$$H' = V(\mathbf{r})e^{-i\omega t} + V^{\dagger}(\mathbf{r})e^{i\omega t} \tag{14.45}$$

Let us assume again that the system starts in the initial state u_i at $t = 0$.

$$C_i^{(0)} = 1 \quad C_n^{(0)} = \delta_{ni}$$

In the first-order perturbation theory,

$$\begin{aligned}
C_n^{(1)}(t) &= \frac{1}{i\hbar}\int_0^t \langle \phi_n | H' | \phi_i \rangle dt' \\
&= \frac{1}{i\hbar}\left[\int_0^t \langle \phi_n | V(\mathbf{r})e^{-i\omega t} | \phi_i \rangle + \int_0^t \langle \phi_n | V^{\dagger}(\mathbf{r})e^{i\omega t} | \phi_i \rangle\right] \\
&= \frac{1}{i\hbar}\left[\int_0^t \langle u_n | e^{i\omega_n t} V(\mathbf{r})e^{-i\omega t} e^{-i\omega_i t} | u_i \rangle + \int_0^t \langle u_n | e^{i\omega_n t} V^{\dagger}(\mathbf{r})e^{i\omega t} e^{-i\omega_i t} | u_i \rangle\right] \\
&= \frac{1}{i\hbar}\left[V_{ni}\int_0^t e^{i(\omega_{ni}-\omega)t'} dt' + (V^{\dagger})_{ni}\int_0^t e^{i(\omega_{ni}+\omega)t'} dt'\right] \\
&= \frac{1}{i\hbar}\left[V_{ni}\frac{(e^{i(\omega_{ni}-\omega)t} - 1)}{i(\omega_{ni}-\omega)} + (V^{\dagger})_{ni}\frac{(e^{i(\omega_{ni}+\omega)t} - 1)}{i(\omega_{ni}+\omega)}\right] \\
&= \frac{1}{i\hbar}\left\{\begin{aligned} &\frac{V_{ni}e^{i(\omega_{ni}-\omega)t/2}}{i(\omega_{ni}-\omega)}[e^{i(\omega_{ni}-\omega)t/2} - e^{-i(\omega_{ni}-\omega)t/2}] + \\ &\frac{V^{\dagger}{}_{ni}e^{i(\omega_{ni}+\omega)t/2}}{i(\omega_{ni}+\omega)}[e^{i(\omega_{ni}+\omega)t/2} - e^{-i(\omega_{ni}+\omega)t/2}] \end{aligned}\right\} \\
&= \frac{1}{i\hbar}\left\{\begin{aligned} &2\frac{V_{ni}e^{i(\omega_{ni}-\omega)t/2}}{(\omega_{ni}-\omega)}\sin(\omega_{ni}-\omega)t/2 + \\ &\frac{2(V^{\dagger})_{ni}}{(\omega_{ni}+\omega)}e^{i(\omega_{ni}+\omega)t/2}\sin(\omega_{ni}+\omega)t/2 \end{aligned}\right\} \qquad (14.46) \\
&= A + B
\end{aligned}$$

Let us now calculate $|C_n^{(1)}(t)|^2$.

$$\therefore |C_n^{(1)}(t)|^2 = (A^* + B^*)(A + B) = |A|^2 + |B|^2 + (A^*B + AB^*)$$

The term A is very small (almost zero), excepting in a small region around $\omega_{ni} \simeq \omega$; i.e., A is significant only in the range $\omega - 2\pi/t < \omega_{ni} < \omega + 2\pi/t$. In this region, the second term B is almost zero.

Similarly, the term B is very small for all values of ω_{ni}, excepting when $\omega_{ni} \sim -\omega$ or more specifically $-\omega - 2\pi/t < \omega_{ni} < \omega_{ni} < -\omega + 2\pi/t$ and in this range, A is almost zero. Altogether, we have

$$C_n^{(1)}(t) = \begin{cases} A & \omega - \dfrac{2\pi}{t} < \omega_{ni} < \omega + \dfrac{2\pi}{t} \\ B & -\omega - \dfrac{2\pi}{t} < \omega_{ni} < -\omega + \dfrac{2\pi}{t} \\ \sim 0 & \text{for other values of } \omega_{ni} \end{cases}$$

(We assume that there is no overlap between their two ranges.) Therefore $|C_n^{(1)}|^2$ is given by

$$|C_n^{(1)}(t)|^2 = \begin{cases} |A|^2 & \omega - \dfrac{2\pi}{t} < \omega_{ni} < \omega + \dfrac{2\pi}{t} \\ |B|^2 & -\omega - \dfrac{2\pi}{t} < \omega_{ni} < -\omega + \dfrac{2\pi}{t} \\ \sim 0 & \text{for other values of } \omega_{ni} \end{cases}$$

Let us now consider these two extreme cases.

Case 1: $\omega_{ni} \simeq \omega$

In this case,

$$(E_n - E_i) \simeq \hbar\omega \quad \text{or} \quad E_n \simeq E_i + \hbar\omega$$

Obviously, the system in state u_i absorbs energy $\hbar\omega$ from the source of perturbation and goes to state u_n. This corresponds to absorption process. In the case of an atom interacting with electromagnetic radiation frequency ω, this represents absorption of photons. The transition probability for absorption process is

$$P_{i\to n} = \frac{|\lambda V_{ni}|^2}{\hbar^2} \cdot \frac{\sin^2((\omega_{ni} - \omega)t/2)}{(\omega_{ni} - \omega)/2} \tag{14.47}$$

Case 2: $\omega_{ni'} \simeq -\omega$

In this case,

$$E_n - E_i \simeq -\hbar\omega \quad E_n \simeq E_i - \hbar\omega$$

The system in state u_i gives energy $\hbar\omega$ to the source of perturbation and goes to state u_n. This corresponds to emission process. In the case of an atom interacting with the electromagnetic radiation, this represents the emission of photon of energy $\hbar\omega$.

$$P_{i\to n} = \frac{|(\lambda V^{\dagger})_{ni}|}{\hbar^2}\frac{\sin^2((\omega_{ni}+\omega)t/2)}{[(\omega_{ni}+\omega)t/2]} \tag{14.48}$$

Note that this explanation of absorption or emission in terms of photons does not emerge from non-relativistic quantum mechanics. It requires quantum field theory, though these explanations are obvious.

Working out the calculations in the same way as done in constant perturbation theory, we get the following results:

For absorption,

$$P_{i\to n} = \frac{2\pi}{\hbar}t|\lambda V_{ni}|^2\delta(E_n - E_i - \hbar\omega) \tag{14.49}$$

$$P_{i\to[n]} = \frac{2\pi}{\hbar}t|\lambda V_{ni}|^2\rho(E_n)\big|_{E_n\simeq E_i+\hbar\omega} \tag{14.50}$$

$$W_{i\to n} = \frac{2\pi}{\hbar}|\lambda V_{ni}|^2\delta(E_n - E_i - \hbar\omega) \tag{14.51}$$

$$W_{i\to[n]} = \frac{2\pi}{\hbar}t|\lambda V_{ni}|^2\rho(E_n)\big|_{E_n\simeq E_i+\hbar\omega} \tag{14.52}$$

For emission,

$$P_{i\to n} = \frac{2\pi t}{\hbar}|\lambda V^{\dagger}{}_{ni}|^2\delta(E_n - E_i + \hbar\omega) \tag{14.53}$$

$$P_{i\to[n]} = \frac{2\pi t}{\hbar}|\lambda V^{\dagger}{}_{ni}|^2\rho(E_n)\big|_{E_n\simeq E_i-\hbar\omega} \tag{14.54}$$

$$W_{i\to n} = \frac{2\pi}{\hbar}|\lambda V^{\dagger}{}_{ni}|^2\delta(E_n - E_i + \hbar\omega) \tag{14.55}$$

$$W_{i\to[n]} = \frac{2\pi}{\hbar^2}|\lambda V^{\dagger}{}_{ni}|^2\rho(E_n)\big|_{E_n\simeq E_i-\hbar\omega} \tag{14.56}$$

14.8 ATOMS AND RADIATION – THREE KINDS OF TRANSITION

It is interesting to analyze the possible kinds of atomic transitions due to electromagnetic interactions. From Bohr model, we have the following naive picture. To make a transition from state u_i to u_m (assume $E_m > E_i$), the presence of a photon energy $\hbar\omega = E_m - E_i$ is necessary. This process is called absorption. On the other hand, if an atom in state u_m is left alone, it will make a transition to E_i by emitting a photon of energy $\hbar\omega$. The absorption of the photon is induced absorption, and the emission process is spontaneous process.

Does non-relativistic quantum mechanics support this understanding? The non-relativistic quantum mechanics rules out the possibility of spontaneous emission, but it introduces a new process called induced emission.

Why is spontaneous emission not possible in non-relativistic quantum mechanics? If there is no electromagnetic field, there is no time-dependent perturbation. If there is no time-dependent perturbation, the state of an atom is a stationary eigenstate, and so if an atom is left in state u_m, it will continue to remain in the same state. On the other hand, the harmonic perturbation leads to both emission as well as absorption. This emission is induced emission due to harmonic perturbation of electromagnetic field.

What about spontaneous emission? There are induced emission as well as absorption as a result of electromagnetic radiation interacting with the atom. In addition to these two processes, spontaneous emissions do take place, but the non-relativistic quantum mechanics is insufficient to describe the spontaneous emission process. The explanation of spontaneous emission comes from quantum field theory. The atom lives for ever in a fluctuating background electromagnetic field. This field is not directly observable. However, its effect can be seen indirectly through processes like spontaneous emission.

14.9 ATOMS AND RADIATION – SEMI-CLASSICAL THEORY

The important application of harmonic perturbation theory is the study of interaction between electromagnetic waves and atom. We know that an atom absorbs and emits electromagnetic radiation. In fact, in Bohr model, the atom absorbs a photon of energy $\hbar\omega$ and goes from the states of energies E_i to E_m with $E_m = E_i + \hbar\omega$. In the same way, the atom makes a transition from u_i to u_n by emitting a radiation of energy $\hbar w$ ($E_i = E_n + \hbar\omega$). The description of electromagnetic waves in terms of photons requires quantum field theory, which makes electromagnetic field themselves as operators. Here, we treat electromagnetic waves as a classical vector function. So, in the description of the interaction between the atoms and electromagnetic waves, we use quantum mechanical formulation like operators, eigenvalues and wave functions for atoms while electromagnetic waves are treated as a classical physics variable. This treatment is called semi-classical theory.

14.9.1 Classical Electromagnetic Waves

Let us recollect some basic features of the electromagnetic radiation. We use four quantities namely $\mathbf{A}(x, y, z, t)$ and $\phi(x, y, z, t)$ to describe the electromagnetic radiation. The four quantities $A_x(x, y, z, t)$, $A_y(x, y, z, t)$, $A_z(x, y, z, t)$ and $\phi(x, y, z, t)$ are very much similar to the atomic state wave function $\psi(x, y, z, t)$. Do we need four wave functions to describe electromagnetic waves? From classical electromagnetism, we know that only two quantities are required to describe electromagnetic waves. Use of four functions, therefore, creates a redundancy in the theory. How do we get rid of this redundancy?

The standard trick is to impose additional conditions on $\mathbf{A}$ and ϕ such that we have got only two independent quantities. This process is known as gauge fixing. Without going into details, let us take the following conditions for our purpose.

$$\nabla \cdot \mathbf{A} = 0 \quad \phi = 0$$

This is called Coulomb gauge (or radiation gauge).

The vector potential $\mathbf{A}$ describing the electromagnetic wave of wave vector $\mathbf{k}$ and frequency ω is given by

$$\mathbf{A} = 2A_0\boldsymbol{\varepsilon}\cos(\mathbf{k}\cdot\mathbf{r} - \omega t) \tag{14.57}$$

where ε is polarization vector. The electric field **E** and the magnetic fields **B** are given below:

$$\begin{aligned}\mathbf{E} &= -\frac{1}{c}\frac{\partial \mathbf{A}}{\partial t} = 2\frac{\omega}{c}A_0\boldsymbol{\varepsilon}\sin(\mathbf{k}\cdot\mathbf{r}-\omega t)\\ &= 2A_0k\boldsymbol{\varepsilon}\sin(\mathbf{k}\cdot\mathbf{r}-\omega t)\\ \mathbf{B} &= \nabla\times\mathbf{A} = -2A_0(\mathbf{k}\times\boldsymbol{\varepsilon})\sin(\mathbf{k}\cdot\mathbf{r}-\omega t)\end{aligned}$$

The radiation gauge condition is $\nabla\cdot\mathbf{A}=0$.

$$\nabla\cdot\mathbf{A} = 2A_0\mathbf{k}\cdot\boldsymbol{\varepsilon}\sin(\mathbf{k}\cdot\mathbf{r}-\omega t) = 0$$

Therefore, the polarization vector $\boldsymbol{\varepsilon}$ is perpendicular to $\mathbf{k}$; the electric field **E** is along the direction of ε and perpendicular to $\mathbf{k}$. The magnetic field **B** is perpendicular to both $\boldsymbol{\varepsilon}$ and $\mathbf{k}$.

The energy density of the electromagnetic wave is given by

$$\begin{aligned}u &= \frac{1}{8\pi}(\mathbf{E}^2+\mathbf{B}^2)\\ &= \frac{1}{\pi}k^2A_0^2\sin^2(\mathbf{k}\cdot\mathbf{r}-\omega t)\end{aligned}$$

This is a fluctuating function, and experimentally, more relevant quantity is the average energy density, which is given by

$$\bar{u} = \frac{1}{2\pi}k^2A_0^2$$

The corresponding energy flux is given by

$$I = \bar{u}c = \frac{c}{2\pi}k^2A_0^2 = \frac{\omega^2}{2\pi c}A_0^2$$

So far, we were concerned with monochromatic radiation. For our purpose, we consider an incoherent beam of radiation with a spread in frequencies. Let $\rho(\omega)d\omega$ be the number of modes between ω and $\omega + d\omega$. The flux density corresponding to each mode is $\omega^2A_0^2/2\pi c$. Therefore, the flux density of the incoherent beam of radiation with frequencies between ω and $\omega+d\omega$ is

$$I(\omega)d\omega = \frac{\omega^2}{2\pi c}A_0^2\rho(\omega)d\omega \tag{14.58}$$

14.9.2 Transition Probability for the Atom

The Schrödinger equation for an electron in the atom is given by

$$i\hbar\frac{\partial\psi}{\partial t} = \left[\frac{\left(\mathbf{p}+\frac{e\mathbf{A}}{c}\right)^2}{2m} + V(\mathbf{r}) - e\phi\right]\psi$$

Let us first evaluate $\frac{1}{2m}\left(\mathbf{p}_{op}+\frac{e}{c}\mathbf{A}\right)^2\psi$

$$\frac{1}{2m}\left(\mathbf{p}_{op}+\frac{e}{c}\mathbf{A}\right)^2\psi=\frac{1}{2m}\left(\mathbf{p}_{op}+\frac{e}{c}\mathbf{A}\right)\cdot\left(\mathbf{p}_{op}\psi+\frac{e}{c}\mathbf{A}\psi\right)$$

$$=\frac{1}{2m}\left(\mathbf{p}_{op}\cdot\mathbf{p}_{op}\psi+\mathbf{p}_{op}\cdot\left(\frac{e}{c}\mathbf{A}\psi\right)+\frac{e}{c}\mathbf{A}\cdot\mathbf{p}_{op}\psi+\frac{e^2}{c^2}A^2\psi\right)$$

$$\mathbf{p}_{op}\cdot(\mathbf{A}\psi)=e(\mathbf{p}_{op}\cdot\mathbf{A})\psi+e\mathbf{A}\cdot\mathbf{p}_{op}\psi$$

But $$(\mathbf{p}_{op}\cdot\mathbf{A})=-i\hbar\,\nabla\cdot\mathbf{A}=0$$

$$\therefore\quad\frac{1}{2m}\left(\mathbf{p}_{op}+\frac{e}{c}\mathbf{A}\right)^2\psi=\frac{1}{2m}\left(p_{op}^2+2\frac{e}{c}\mathbf{A}\cdot\mathbf{p}_{op}+\frac{e^2}{c^2}A^2\right)\psi$$

Since we are using radiation gauge, $\phi=0$. Then, the Schrodinger equation becomes

$$i\hbar\frac{\partial\psi}{\partial t}=\left[\frac{p^2}{2m}+V(\mathbf{r})+\frac{e}{mc}\mathbf{A}\cdot\mathbf{p}+\frac{e^2A^2}{m^2c^2}\right]\psi$$

The last term is estimated to be very small. Therefore, we have

$$i\hbar\frac{\partial\psi}{\partial t}=\left[\frac{p^2}{2m}+V(\mathbf{r})+\frac{e}{mc}\mathbf{A}\cdot\mathbf{p}\right]\psi$$

$$=(H_0+H')\psi$$

where, $$H_0=\frac{p^2}{2m}+V(\mathbf{r})\tag{14.59}$$

and $$H'=\frac{e}{mc}\mathbf{A}\cdot\mathbf{p}\tag{14.60}$$

First, let us consider monochromatic radiation of frequency ω. Then H' is

$$H'=\frac{2e}{mc}A_0(\omega)\cos(\mathbf{k}\cdot\mathbf{r}-\omega t)\boldsymbol{\varepsilon}\cdot\mathbf{p}$$

$$=\frac{A_0(\omega)}{mc}\left[e^{(i\mathbf{k}\cdot\mathbf{r}-\omega t)}+e^{-i(\mathbf{k}\cdot\mathbf{r}-\omega t)}\right]\boldsymbol{\varepsilon}\cdot\mathbf{p}$$

$$=V(t)e^{-i\omega t}+V^\dagger(t)e^{i\omega t}$$

where $$V=\frac{A_0(\omega)}{mc}e^{i\mathbf{k}\cdot\mathbf{r}}\boldsymbol{\varepsilon}\cdot\mathbf{p}\tag{14.61}$$

and $$V^\dagger=\frac{A_0(\omega)}{mc}e^{-i\mathbf{k}\cdot\mathbf{r}}\boldsymbol{\varepsilon}\cdot\mathbf{p}\tag{14.62}$$

From harmonic perturbation theory, we know that the transition probability rates $W_{i\to n}$ and $W_{n\to i}$ are given by

$$W_{i\to n} = \frac{2\pi}{\hbar}|V_{ni}|^2\delta(E_n - E_i - \hbar\omega) \qquad \text{Absorption}$$

$$W_{n\to i} = \frac{2\pi}{\hbar}|V_{ni}^\dagger|^2\delta(E_n - E_i + \hbar\omega) \qquad \text{Emission}$$

From the principle of detailed balance, we have

$$W_{i\to n} = W_{n\to i}$$

In the discussion of harmonic perturbation theory, we considered the transition probability from state u_i to a group of final states whose energy eigenvalues centered around E_n. We assumed that harmonic perturbation is monochromatic. Now we are interested in a different scenario. We now consider two specific states of the atom, namely u_i and u_n. However, the incident radiation is not monochromatic. The source now produces an incoherent beam with a spread of frequencies. To find the transition probability $P_{i\to n}$, we add the probabilities due to all the frequencies in the incoherent beam.

i.e.,
$$P_{i\to \mathrm{n}} = \sum_{\omega} P_{i\to n} \quad \text{for each mode}$$

Normally, the frequencies of different modes are very close to each other, and so we can take ω as continuous variable. And in that case, the summation over ω becomes an integral. If $\rho(\omega)$ is the number of modes between ω and $\omega + d\omega$, then the transition probability per unit time, $W_{i\to n}$, is given by

$$\begin{aligned}W_{i\to n} = W_{n\to i} &= \int \frac{2\pi}{\hbar}|V_{ni}|^2\delta(E_n - E_i - \hbar\omega)\rho(\omega)d\omega \\ &= \int \frac{2\pi}{\hbar}\cdot\frac{e^2}{m^2c^2}A_0^2(\omega)|\langle u_n|e^{i\mathbf{k}\cdot\mathbf{r}}\boldsymbol{\varepsilon}\cdot\mathbf{p}\,|u_i\rangle|^2\ \delta(E_n - E_i - \hbar\omega)\rho(\omega)d\omega\end{aligned}$$

$$\begin{aligned}\delta(E_n - E_i - \hbar\omega) &= \delta(\hbar(\omega_n - \omega_i - \omega)) = \frac{1}{\hbar}\delta(\omega_n - \omega_i - \omega) \\ &= \frac{1}{\hbar}\delta(\omega_{ni} - \omega)\end{aligned}$$

$$\begin{aligned}\therefore W_{i\to n} = W_{n\to i} &= \int \frac{2\pi}{\hbar^2}\cdot\frac{e^2}{m^2c^2}A_0^2(\omega)|\langle u_n|e^{i\mathbf{k}\cdot\mathbf{r}}\boldsymbol{\varepsilon}\cdot\mathbf{p}|u_i\rangle|^2\rho(\omega)\delta(\omega_{ni} - \omega)d\omega \\ &= \frac{2\pi}{\hbar^2}\cdot\frac{e^2}{m^2c^2}A_0^2(\omega_{ni})\rho(\omega_{ni})|\langle u_n|e^{i\mathbf{k}\cdot\mathbf{r}}\boldsymbol{\varepsilon}\cdot\mathbf{p}|u_i\rangle|^2\end{aligned}$$

From (14.58), we get

$$\begin{aligned}A_0^2(\omega_{ni})\rho(\omega_{ni}) &= \frac{I(\omega_{ni})2\pi c}{\omega_{ni}^2} \\ \therefore W_{i\to n} &= \frac{4\pi e^2}{\hbar^2 m^2 c}\cdot\frac{I(\omega_{ni})}{\omega_{ni}^2}|\langle u_n|e^{i\mathbf{k}\cdot\mathbf{r}}\boldsymbol{\varepsilon}\cdot\mathbf{p}|u_i\rangle|^2\end{aligned} \tag{14.63}$$

14.9.3 Dipole Approximation

It will be interesting to estimate $e^{i\mathbf{k}\cdot\mathbf{r}}$ for an electron in an atom. The average scale of the distance of an electron from the nucleus is few Bohr radius a. Therefore, we can replace $e^{i\mathbf{k}\cdot\mathbf{r}}$ by e^{ika}. For optical region of electromagnetic spectrum, let us take $\lambda \sim 5000\text{Å}$.

$$\begin{aligned} ka &= \frac{2\pi}{\lambda}a \simeq \frac{2\pi\times 0.5\text{Å}}{5000\text{Å}} \simeq 10^{-3} \\ e^{ika} &= 1 + ika + \frac{(ika)^2}{2}\cdots \\ \therefore e^{ika} &\simeq 1 \end{aligned}$$

This is known as dipole approximation. The reason for this terminology will become clear soon. Now the matrix element in (14.63) becomes

$$\langle u_n|e^{i\mathbf{k}\cdot\mathbf{r}}\boldsymbol{\varepsilon}\cdot\mathbf{p}|u_i\rangle \simeq \langle u_n|\boldsymbol{\varepsilon}\cdot\mathbf{p}|u_i\rangle \tag{14.64}$$

The operator $\boldsymbol{\varepsilon}\cdot\mathbf{p}$ can be expressed in terms of the Hamiltonian H and the position vector $\mathbf{r}$.

Let us first evaluate $[H,\mathbf{r}]$.

$$[H,\mathbf{r}] = [H,x]\mathbf{x} + [H,y]\mathbf{y} + [H,z]\mathbf{z}$$

$[H,x]$ is given by

$$\begin{aligned} [H,x] &= \left[\frac{p^2}{2m} + V(\mathbf{r}), x\right] = \left[\frac{p^2}{2m}, x\right] \\ &= \left[\frac{p_x^2}{2m}, x\right] + \underbrace{\left[\frac{p_y^2}{2m}, x\right]}_{0} + \underbrace{\left[\frac{p_z^2}{2m}, x\right]}_{0} = \frac{-i\hbar}{m}p_x \end{aligned}$$

Evaluating other terms also in the same way, we get

$$\begin{aligned} \therefore [H,\mathbf{r}] &= \frac{-i\hbar}{m}\mathbf{x}p_x - \frac{i\hbar}{m}\mathbf{y}\,p_y \frac{-i\hbar}{m}\mathbf{z}p_z = -\frac{i\hbar}{m}\mathbf{p} \\ \therefore \mathbf{p} &= \frac{im}{\hbar}[H,\mathbf{r}] \\ \therefore \boldsymbol{\varepsilon}\cdot\mathbf{p} &= \frac{im}{\hbar}[H,\boldsymbol{\varepsilon}\cdot\mathbf{r}] \end{aligned}$$

$$\begin{aligned} \therefore \langle u_n|\boldsymbol{\varepsilon}\cdot\mathbf{p}\,|u_i\rangle &= \frac{im}{\hbar}\langle u_n|[H,\boldsymbol{\varepsilon}\cdot\mathbf{r}]|u_i\rangle = \frac{im}{\hbar}\langle u_n|H\boldsymbol{\varepsilon}\cdot\mathbf{r} - \boldsymbol{\varepsilon}\cdot\mathbf{r}\,H|u_i\rangle \\ &= \frac{im(E_n - E_i)}{\hbar}\langle u_n|\boldsymbol{\varepsilon}\cdot\mathbf{r}|u_i\rangle = im\,\omega_{ni}\langle u_n|\boldsymbol{\varepsilon}\cdot\mathbf{r}|u_i\rangle \end{aligned} \tag{14.65}$$

If $-e$ is the charge of the electron and if we assume the nucleus is at the origin, then the dipole moment $\mathbf{d} = -e\mathbf{r}$.

$$\begin{aligned}\therefore \langle u_n|\boldsymbol{\varepsilon}\cdot\mathbf{r}|u_i\rangle &= \frac{-im\omega_{ni}}{e}\langle u_n|\,\boldsymbol{\varepsilon}\cdot\mathbf{d}|u_i\rangle \\ &= \frac{-im\omega_{ni}}{e}\boldsymbol{\varepsilon}\cdot\mathbf{d}_{ni}\end{aligned} \tag{14.66}$$

where $\mathbf{d}_{ni}$ is the matrix element of the dipole moment $\mathbf{d}$.

The transition probability rate from state u_i to state u_n is

$$W_{i\to n} = \frac{4\pi^2 I(\omega_{ni})}{\hbar^2 c}|\boldsymbol{\varepsilon}\cdot\mathbf{d}_{ni}|^2 \tag{14.67}$$

14.10 SELECTION RULES

The matrix elements $\langle u_f|\boldsymbol{\varepsilon}\cdot\mathbf{d}|u_i\rangle$ is non-zero for a few pairs of states u_f and u_i. Therefore, the transitions between only such pair of states are allowed. Other transitions are not allowed due to the matrix element $\langle u_f|\boldsymbol{\varepsilon}\cdot\mathbf{d}|u_i\rangle$ becoming zero. These are called forbidden transitions. Therefore, the dipole matrix element imposes certain conditions on the orbital angular momentum quantum numbers l_i and l_f, and the azimuthal quantum numbers m_i and m_f for a transition to be allowed. These are known as selection rules.

These selection rules can be obtained in many ways. Let us write $\boldsymbol{\varepsilon}\cdot\mathbf{d}$ as

$$\begin{aligned}\boldsymbol{\varepsilon}\cdot\mathbf{d} &= -e[\boldsymbol{\varepsilon}_x x + \boldsymbol{\varepsilon}_y y + \boldsymbol{\varepsilon}_z z] \\ &= -e[\boldsymbol{\varepsilon}_x r\sin\theta\cos\phi + \boldsymbol{\varepsilon}_y r\sin\theta\sin\phi + \boldsymbol{\varepsilon}_z r\cos\theta]\end{aligned}$$

Let us recall

$$\begin{aligned} &= \left(\frac{3}{4\pi}\right)^{1/2}\cos\theta \\ Y_{11} &= \left(\frac{3}{8\pi}\right)^{1/2}\sin\theta(\cos\phi + i\sin\phi) \\ Y_{1-1} &= \left(\frac{3}{8\pi}\right)^{1/2}\sin\theta(\cos\phi - i\sin\phi)\end{aligned}$$

Making use of these relations, we get

$$\begin{aligned}\boldsymbol{\varepsilon}\cdot\mathbf{d} &= er\left(\frac{2\pi}{3}\right)^{1/2}\left[-(\boldsymbol{\varepsilon}_x + i\boldsymbol{\varepsilon}_y)Y_{11} + (\varepsilon_x - i\boldsymbol{\varepsilon}_y)Y_{1-1} - \left(\frac{4\pi}{3}\right)^{1/2}\boldsymbol{\varepsilon}_z Y_{10}\right] \\ &= er[d_1 Y_{11} + d_0 Y_{10} + d_{-1}Y_{1-1}] \\ &= \sum_m erd_m Y_{1m}\end{aligned} \tag{14.68}$$

The wave functions of the atom can be written as

$$u_i = R_{n_i l_i}(r)Y_{l_i m_i}(\theta,\phi) \quad \text{and}\; u_f = R_{n_f l_f}(r)Y_{l_f m_f}(\theta,\phi)$$

$$\therefore (\boldsymbol{\varepsilon}\cdot\mathbf{d})_{fi} = e\int_0^\infty R^*_{n_f l_f}(r)R_{n_i l_i}(r)r^3dr.\left[\sum_m d_m\int Y^*_{l_f m_f}Y_{1m}Y_{l_i m_i}d\Omega\right] \tag{14.69}$$

Let us consider $Y_{1m}Y_{l_i m_i}$. This is a basis vector in uncoupled representation. Let us recall that when two angular momenta of quantum number j_1 and j_2 are added, we can write

$$|j_1 m_1\rangle \cdot |j_2 m_2\rangle = \sum_j \langle j_1 m_1; j_2 m_2 | j\,\overline{m_1 + m_2}\rangle\, |j\,\overline{m_1 + m_2}\rangle$$

where $\langle j_1 m_1; j_2 m_2 | j m_1 + m_2\rangle$ are the Clebsh–Gordan coefficients. Here, the two orbital angular momentum quantum numbers are $j_l = l_i$ and $j_2 = 1$. The resulting angular momentum has quantum numbers $l_i + 1$, l_i and $l_i - 1$ provided $l_i \neq 0$. So we have to consider two cases, namely $l_i \neq 0$ and $l_i \neq 0$

Case 1:

$$Y_{1m}Y_{l_i m_i} = C_1 Y_{l_i+1, m+m_i} + C_2 Y_{l_i, m+m_i} + C_3 Y_{l_i-1, m+m_i} \tag{14.70}$$

where C_1, C_2 and C_3 are Clebsh–Gordan coefficients. The parity of $Y_{1m}Y_{l_i m_i}$ is $(-1)^{l_i+1}$ and the parity of $Y_{l_i, m+m_i}$ is $(-1)^{l_i}$. Since the functions in the LHS and RHS should have the same parity, C_2 has to be zero.

$$C_2 = 0$$

$$\begin{aligned}\therefore \int Y^*_{l_f m_f} Y_{1m} Y_{l_i m_i}\, d\Omega &= C_1 \int Y^*_{l_f m_f} Y_{l_i+1, m+m_i}\, d\Omega + C_3 \int Y^*_{l_f m_f} Y_{l_i-1, m+m_i}\, d\Omega \\ &= C_1 \delta_{l_f, l_i+1} \delta_{m_f, m+m_i} + C_3 \delta_{l_f, l_i-1} \delta_{m_f, m+mi}\end{aligned} \tag{14.71}$$

The integral is non-zero provided

$$l_f = l_i + 1, \quad m_f = m_i + m$$

or

$$l_f = l_i - 1, \quad m_f = m_i + m$$

Note that the possible values of m are ±1, 0.

Therefore, the selection rules are

$$l_f - l_i = \pm 1, \quad m_f - m_i = m = \pm 1, 0$$

Case 2: $l_i = 0$

In this case,

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

$$\begin{aligned}\therefore \int Y^*_{l_f m_f} Y_{1m} Y_{00}\, d\Omega &= \int Y^*_{l_f m_f} Y_{1m} \frac{1}{4\pi}\, d\Omega \\ &= \frac{1}{4\pi} \delta_{l_f 1} \delta_{m_f m}\end{aligned}$$

$$\text{i.e., } l_f = 1, m_f = m = 0, \pm 1.$$

$$\therefore (\boldsymbol{\varepsilon} \cdot \mathbf{d})_{fi} \neq 0 \ \text{ only when } \ l_f - l_i = \pm 1,\ m_f - m_i = 0, \pm 1 \ \text{ for } l_i \neq 0 \tag{14.72}$$

and

$$l_f = 1, \quad m_f = 0, \pm 1 \quad \text{for} \quad l_i = 0 \tag{14.73}$$

These are the selection rules for a transition to be allowed in the dipole approximation. The transitions for which the selection rules are not obeyed are termed as forbidden transitions. However, we have to bear in mind that these selection rules are consequences of the dipole approximation. If we include other terms in $e^{i\mathbf{k}\cdot\mathbf{r}}$, other transitions will become allowed. However, their strength will be very small compared to those allowed by dipole approximation.

14.11 EINSTEIN COEFFICIENTS

Much before the advent of quantum mechanics, as we know today, Einstein in 1917 arrived at the idea of induced emission by considering the thermodynamical equilibrium of black body radiation inside a cavity with the atoms constituting the cavity walls.

What do we mean by thermodynamic equilibrium between the atoms and radiation? Both the radiation and the atoms are at the same temperature T; the number of atom in each state remains the same; the number of photons at a particular frequency remains the same. It has to be understood that all these conditions are dynamical. For instance, if we consider two specific states 1 and 2, the number of atoms which change from state 1 to state 2 should be the same as the number of atom which change from state 2 to state 1.

Let us consider two specific states 1 and 2. Since the atoms are in thermodynamic equilibrium, the numbers N_1 and N_2 are fixed. It is given by

$$\frac{N_2}{N_1} = \frac{e^{-E_2/kT}}{e^{-E_1/kT}} \tag{14.74}$$

The energy density of the radiation inside the cavity is

$$u(\omega) = \frac{\hbar\omega^3}{\pi^2 c^2 e^{\hbar\omega/kT} - 1} \tag{14.75}$$

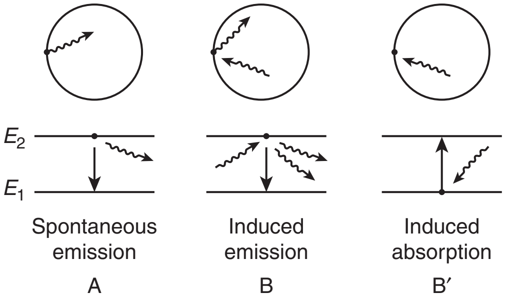


Fig. 14.2 Atomic transitions in a black body

An atom in state 1 can absorb a photon from the radiation inside the cavity and make a transition to state 2. This is induced absorption. An atom in state 2 can make a spontaneous transition to state 1 by emitting radiation of energy $\hbar\omega$. The atom does not need any photon from the radiation for this process. The atom in state 2 can also make transition to a state 1 by induced emission. This requires a photon from the radiation.

Now, we would like to calculate ΔN_1 and ΔN_2. Here ΔN_1 is the number of atoms making a transition from state 1 to state 2 in one second. Similarly ΔN_2 is the number of atoms making transitions from state 2 to state 1 in one second. For thermodynamic equilibrium, $\Delta N_1 = \Delta N_2$.

Let A be the transition probability rate for spontaneous transition from state 2 to state 1.

Similarly B and B' are the probability rates for induced emission and absorption, respectively. These coefficients are known as Einstein's coefficients.

Let us now determine ΔN_1, the number of atoms per second which change from state 1 to state 2 by induced absorption. This number should be proportional to the number of atoms in state 1, the energy density of the radiation and the Einstein's coefficient B'.

i.e.,
$$\therefore \Delta N_1 = N_1 u(\omega) B'$$

Similarly, we have to determine ΔN_2. It is given by

$$\Delta N_2 = N_2 uB + N_2 A$$

$N_2 uB$ and $N_2 A$ correspond to induced emission and spontaneous emission, respectively. Note the absence of u in the second term. This is due to the fact that spontaneous emission independent of the cavity radiation.

Since $\Delta N_1 = \Delta N_2$, we have

$$N_2 uB + N_2 A = N_1 uB'$$

$$u(N_1 B' - N_2 B) = N_2 A$$

$$\therefore u = \frac{N_2 A}{N_1 B' - N_2 B} = \frac{A/B'}{\frac{N_1}{N_2} - \frac{B}{B'}}$$

$$= \frac{A/B'}{e^{\hbar\omega/kT} - \frac{B}{B'}}$$

Comparing this expression with the Planck's law given in (14.75), we get

$$B = B' \quad \text{and} \quad \frac{A}{B'} = \frac{\hbar\omega^3}{\pi^2 c^2}$$

14.12 ADIABATIC APPROXIMATION

In time-dependent perturbation theory, the Hamiltonian $H(t)$ is split into two parts, namely H_0 and $H'(t)$ where $H'(t)$ is taken as the perturbation to the system, and the effect of $H'(t)$ is to make a transition from one stationary eigenstate of H_0 to another stationary eigenstate. Now, we discuss the adiabatic approximation to describe the system under different condition. In the adiabatic approximation, the Hamiltonian varies with time very slowly. In fact, we can consider systems where Hamiltonian depends on external parameters like magnetic field, which can be made to vary very slowly to make the adiabatic approximations valid.

The Schrödinger equation is

$$i\hbar \frac{\partial \psi(t)}{\partial t} = H(t)\psi(t) \tag{14.76}$$

More explicitly, we can write the Hamilton as $H(R_i(t))$, where $R_1(t), R_2(t)\ldots$ are time-dependent parameters that occur in the Hamiltonian.

i.e.,
$$i\hbar\frac{\partial\psi(t)}{\partial t} = H(R_i(t))\psi(t) \tag{14.77}$$

If the Hamiltonian is independent of time, the energy eigenstates of the Hamiltonian are stationary states, and the energy eigenvalues are independent of time as shown in Fig. 14.3.

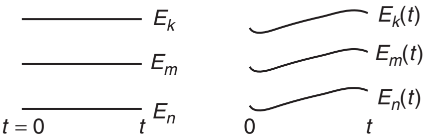


Fig. 14.3 Stationary energy eigenstates and instantaneous energy eigenstates

Under some conditions, even now it is possible to define instantaneous eigenvalues $E_n(t)$ and instantaneous eigenstates $\phi_n(t)$ given by

$$H(t)\ \phi_n(t) = E_n(t)\ \phi_n(t) \tag{14.78}$$

We restrict ourselves to systems whose instantaneous eigenvalues $E_n(t)$ are as shown in Fig. 14.3, where they do not cross each other.

14.12.1 Adiabatic Theorem

Adiabatic theorem states that when the Hamiltonian varies slowly with time for a system whose initial state $\psi(0)$ is $\phi_n(0)$, the state $\psi(t)$ of the system evolves to the instantaneous eigenstate ϕ_n(t).

Let us elaborate this theorem. Let us assume that the instantaneous eigenstates $\{\phi_n(t)\}$ form a complete set. Then, we can write $\psi(t)$ as

$$\psi(t) = \sum_n C_m(t)e^{-i\theta_m(t)}\phi_m(t)$$

where $\theta_m(t)$ is called dynamic phase factor given by

$$\theta_m(t) = \frac{1}{\hbar}\int_0^t E_m(t')dt'$$

We have assumed the initial state as $\psi(0) = \phi_n(0)$. Let us write $\psi(t)$ as

$$\psi(t) = C_n(t)e^{-i\theta_n(t)}\phi_n(t) + \sum_{m\neq n} C_m(t)e^{-i\theta_m(t)}\phi_m(t)$$

Adiabatic theorem says that for a slowly varying time, we can approximate $\psi(t)$ as

$$\psi(t) \simeq C_n(t)e^{-i\theta_n(t)}\phi_n(t) \tag{14.79}$$

Further, it can be shown that $C_n(t) \simeq e^{i\gamma_n(t)}$.

$$\therefore \psi(t) \simeq e^{i\gamma_n(t)} e^{-i\theta_n(t)} \phi_n(t) \tag{14.80}$$

What is the significance of adiabatic theorem?

First, note that the instantaneous eigenstates behave like stationary eigenstates. The expression $\psi(t)$ in (14.79) almost looks like the time evolution of a stationary state, namely $\psi(t) = e^{-iE_n t/\hbar} u_n$. This suggests that for a slowly varying time-dependent Hamiltonian, the solution to the Schrödinger equation is obtained by solving the corresponding time-independent Schrödinger equation, taking time *t* as one more parameter in the theory. Example 14.1 illustrates this aspect.

There is another important significant aspect which is of recent origin, namely Berry phase. There are two phase factors: $\gamma_n(t)$ and $\theta_n(t)$. The phase $\gamma_n(t)$ is called Berry phase and the phase $\theta_n(t)$ is called dynamic phase. Generally, such phase factors of a wave function are ignored as inconsequential. The phase factor $\gamma_n(t)$ was also ignored for a long time. However, Berry showed that this phase has observable consequences under some circumstances.

Here we shall establish the conditions under which the adiabatic theorem is valid and subsequently develop the concept of Berry phase.

Example 14.1 A particle is confined to a box of length *L*(*t*) by the potential

$$V(t) = \begin{Bmatrix} 0 & 0 < x < L(t) \\ \infty & \text{otherwise} \end{Bmatrix}$$

Determine the energy of the particle assuming the adiabatic theorem is valid.

Solution: Since adiabatic theorem is valid, the problem is treated like solving time-independent Schrödinger equation, taking *t* as parameter. For stationary energy eigenstates, we have

$$\frac{d^2u}{dx^2} + \frac{2m}{\hbar^2}u = 0 \quad 0 < x < L$$

$$u = \sqrt{\frac{2}{L}} \sin\frac{n\pi x}{L} \quad \text{and} \quad E_n = \frac{\hbar^2}{2m} \cdot \frac{\pi^2}{L^2}$$

Therefore, due to adiabatic theorem, we write the instantaneous eigenstate and the instantaneous eigenvalues as

$$\phi_n(t) = \sin\frac{n\pi x}{L(t)}$$

and

$$E_n(t) = \frac{\hbar^2}{2m} \frac{\pi^2}{L^2(t)}$$

14.13 VALIDITY OF ADIABATIC THEOREM

Let the system start in state $\phi_n(t)$. i.e., $\psi(0) = \phi_n$. The general expression for $\psi(t)$ is

$$\psi(t) = \sum_k C_k(t) e^{-i\theta_k} \phi_k(t) \tag{14.81}$$

θ_k is given by
$$\theta_k = \frac{1}{\hbar}\int_0^t E_k(t')dt'$$

$$\frac{d\theta_k}{dt} = \frac{1}{\hbar}E_k(t)$$

$|C_k(t)|^2$ gives the probability of finding the system in state $\phi_k(t)$. Therefore, $|C_k(t)|^2$ represents the transition probability, $P_{n\to k}$, from state ϕ_n to ϕ_k. Let us now determine the conditions under which $\left|P_{n\to k}\right|^2 << 1$.

$$i\hbar\frac{d\psi}{dt} = i\hbar\sum_k [\dot{C}_k e^{-i\theta_k}\phi_k + C_k(-i\dot{\theta}_k)e^{-i\theta_k}\ \phi_k + C_k e^{-i\theta_k}\dot{\phi}_k]$$

$$i\hbar\frac{d\psi}{dt} = i\hbar\sum_k [\dot{C}_k e^{-i\theta_k}\phi_k + C_k\left(-i\frac{1}{\hbar}E_k(t)\right)e^{-i\theta_k}\ \phi_k + C_k e^{-i\theta_k}\dot{\phi}_k] \tag{14.82}$$

$$H(t)\psi(t) = \sum_k C_k(t)H(t)\phi_k(t) = \sum_k C_k(t)E_k(t)\phi_k(t) \tag{14.83}$$

Equating (14.82) and (14.83), we get

$$\sum_k \dot{C}_k e^{-i\theta_k}\phi_k = -\sum_k C_k e^{-i\theta_k}\dot{\phi}_k$$

Taking scalar product with ϕ_m, we get

$$\dot{C}_m e^{-i\theta_m} = -\sum_k C_k e^{-i\theta_k}\langle\phi_m\ |\dot{\phi}_k\rangle$$

$$\therefore \dot{C}_m = -\sum_k C_k e^{i(\theta_k-\theta_m)}\langle\phi_m\ |\dot{\phi}_k\rangle \tag{14.84}$$

Let us rewrite this equation as

$$\dot{C}_m = -\langle\phi_m\ |\dot{\phi}_m\rangle C_m - \sum_{k\neq m} C_m e^{i(\theta_m-\theta_k)}\langle\phi_m\ |\dot{\phi}_k\rangle \tag{14.85}$$

Let us first evaluate $\langle\phi_m\ |\dot{\phi}_k\rangle$. We have

$$H(t)\phi_k(t) = E_k(t)\phi_k(t)$$

Differentiating this equation, we get

$$\dot{H}\phi_k + H\dot{\phi}_k = \dot{E}_k\phi_k + E_k\dot{\phi}_k$$

Taking scalar product with ϕ_m, we get

$$\langle\phi_m|\dot{H}|\phi_k\rangle + \langle\phi_m|H|\dot{\phi}_k\rangle = \dot{E}_k\langle\phi_m|\phi_k\rangle + E_k\langle\phi_m\ |\dot{\phi}_k\rangle$$

Since $k \neq m$, $\langle \phi_m | \phi_k \rangle = 0$.

$$\therefore (E_k - E_m)\langle \phi_m \,|\, \dot{\phi}_k \rangle = \langle \phi_m \,|\dot{H}|\, \phi_k \rangle$$

(We have used the fact $\langle \phi_m | H = \langle \phi_m | E_m$).

$$\therefore \langle \phi_m \,|\, \dot{\phi}_k \rangle = \frac{1}{E_k - E_m} \langle \phi_m \,|\, \dot{H} \,|\, \phi_k \rangle \text{ for } m \neq k \tag{14.86}$$

$$\dot{C}_m = -\langle \phi_m \,|\, \dot{\phi}_m \rangle C_m - \sum_{k \neq m} C_k e^{i(\theta_m - \theta_k)} \frac{1}{(E_k - E_m)} \langle \phi_m \,|\dot{H}|\, \phi_k \rangle \tag{14.87}$$

Let us now evaluate $\langle \phi_m \,|\, \dot{\phi}_m \rangle$. The normalization condition for ϕ_m is

$$\langle \phi_m \,|\, \phi_m \rangle = 1$$

Differentiating this equation, we get

$$\langle \phi_m \,|\, \dot{\phi}_m \rangle + \langle \dot{\phi}_m \,|\, \phi_m \rangle = 0 \tag{14.88}$$

Since $\langle \phi_m \,|\, \dot{\phi}_m \rangle$ and $\langle \dot{\phi}_m \,|\, \phi_m \rangle$ are complex conjugate of each other, we conclude that $\langle \phi_m \,|\, \dot{\phi}_m \rangle$ is an imaginary number. Let us denote $\langle \phi_m \,|\, \dot{\phi}_m \rangle$ by $i\alpha(t)$.

$$\therefore \langle \phi_m \,|\, \dot{\phi}_m \rangle = i\alpha(t) \tag{14.89}$$

Let us define ϕ'_m and $\phi'_m = e^{i\gamma_m(t)} \phi_m$ (14.90)

$$\therefore \langle \dot{\phi}'_m | \phi'_m \rangle = \left\langle \phi_m \left| e^{-i\gamma} \frac{d}{dt} e^{i\gamma} \right| \phi_m \right\rangle$$

$$= \langle \phi_m | \dot{\phi}_m \rangle + i\dot{\gamma} \langle \phi_m \,| \phi_m \rangle$$

$$= i\alpha(t) + i\dot{\gamma}(t) \tag{14.91}$$

Choose γ in such a way that $\dot{\gamma} = -\alpha$ which makes $\langle \phi'_m \,|\, \dot{\phi}'_m \rangle = 0$. Let us now change the basis set

$$\{\phi_k\} \text{ to } \{\phi'_k = e^{i\gamma_m} \phi_k\} \tag{14.92}$$

Note that all the basis states have been multiplied by the same phase factor $e^{i\gamma_m}$.

We have $\phi_k = e^{-i\gamma_m} \phi'_k$. (14.93)

Substituting the Equation (14.93) in (14.81), we get

$$\psi(t) = \sum_k C_k(t)e^{-i\theta_k}e^{-i\gamma_m}\phi_k'(t) = \sum_k C_k(t)e^{-i(\theta_k+\gamma_m)}\phi_k'(t)$$

$$= \sum_k C_k(t)e^{-i\theta_k'}\phi_k'(t) \tag{14.94}$$

where $\theta_k' = \theta_k + \gamma_m$

The summation over n in the above equation includes the state ϕ_m' also. Therefore, in terms of the new set of basis states $\{\phi_k'(t)\}$, the Equation (14.84) becomes (note that we have $\langle\phi_m' \mid \dot{\phi}_m'\rangle = 0$).

$$\dot{C}_m = -\sum_k C_k e^{i(\theta_m'-\theta_k')}\langle\phi_m' \mid \dot{\phi}_k'\rangle$$

$$= -C_m \underbrace{\langle\phi_m'|\dot{\phi}_m'\rangle}_{0} - \sum_{k\neq m} C_k e^{i(\theta_m'-\theta_k')}\langle\phi_m'|\dot{\phi}_k'\rangle$$

$$= -\sum_{k\neq m} C_k e^{i(\theta_m'-\theta_k')}\langle\phi_m' \mid \dot{\phi}_k'\rangle$$

Let us drop the primes from this expression with the understanding we are working in the new basis.

$$\dot{C}_m = -\sum_{k\neq m} C_k e^{i(\theta_m-\theta_k)}\langle\phi_m \mid \dot{\phi}_k\rangle \tag{14.95}$$

$$= \sum_{k\neq m} C_k(t)\frac{e^{i(\theta_m-\theta_k)}}{(E_m-E_k)}\langle\phi_m|\dot{H}|\phi_k\rangle \quad \text{(Use (14.86))}$$

$$\therefore C_m(t) = \sum_{k\neq m}\int_0^t C_k(t')\frac{e^{i(\theta_m-\theta_k)}}{(E_m-E_k)}\langle\phi_m|\dot{H}|\phi_k\rangle dt' \tag{14.96}$$

We are interested in estimating $C_m(t)$ for a system whose initial state is ϕ_n.

i.e., $C_m(0) = \delta_{mn}$ or $C_k(0) = \delta_{kn}$

Let us make the following approximation in the RHS of the Equation (14.96): $C_m(t) \sim C_m(0)$ and $\theta_m = \frac{1}{\hbar}\int_0^t E_m(t')dt' \sim \frac{E_m t}{\hbar} = \omega_m t$ and let us further assume the matrix element $\langle\phi_m|\dot{H}|\phi_k\rangle$ does not vary much in this time interval. Under these conditions, the Equation (14.96) becomes

$$C_m(t) = \sum_{k\neq m}\delta_{kn}\frac{1}{\hbar\omega_{mk}}\langle\phi_m|\dot{H}|\phi_k\rangle\int_0^t e^{i\omega_{mk}t'}dt'$$

$$= \frac{1}{\hbar\omega_{mn}}\cdot\langle\phi_m|\dot{H}|\phi_n\rangle\int_0^t e^{i\omega_{mn}t'}dt'$$

$$= \frac{1}{\hbar\omega_{mn}} \cdot \langle\phi_m\, |\dot{H}|\, \phi_n\rangle \cdot \frac{(e^{i\omega_{mn}t'} - 1)}{i\omega_{mn}}$$

$$= \frac{1}{\hbar\omega_{mn}^2} \langle\phi_m\, |\dot{H}|\, \phi_n\rangle 2i \sin \omega_{mn} t/2$$

$$\therefore\ P_{n\to m} \simeq \frac{4}{\hbar^2\omega_{mn}^4} \left| \left\langle \phi_m \left| \frac{\partial H}{\partial t} \right| \phi_n \right\rangle \right|^2 \sin^2 \omega_{mn} t/2 \qquad (14.97)$$

The maximum value of $\sin^2\omega_{mt}t/2$ is 1.
Therefore,

$$P_{n\to m} \le \frac{4}{\hbar^2\omega_{mn}^4} \left| \left\langle \phi_m \left| \frac{\partial H}{\partial t} \right| \phi_n \right\rangle \right|^2 \quad n \ne m$$

If $\partial H/\partial t$ is very small, the $P_{n\to m}$ is very small. So the condition for the adiabatic approximation is the time variation of the Hamiltonian should be very small. In other words, the Hamiltonian should be slowly varying with respect to time for adiabatic approximation. Let us make it more precise with the time scale available within the theory. Let us define T as $T = 2\pi/\omega_{mn}$.

Then we have

$$\frac{1}{\omega_{mn}} \cdot \left\langle \phi_m \left| \frac{\partial H}{\partial t} \right| \phi_n \right\rangle = \frac{T}{2\pi} \left\langle \phi_m \left| \frac{\partial H}{\partial t} \right| \phi_n \right\rangle$$

$$\therefore\ P_{n\to m} \le \left| \frac{2}{\hbar\omega_{mn}} \cdot \frac{T}{2\pi} \langle\phi_m|\, \frac{\partial H}{\partial t}\, |\phi_n\rangle \right|^2$$

or

$$P_{n\to m} \le \left| \frac{2}{(E_m - E_n)} \cdot \frac{T}{2\pi} \cdot \left\langle \phi_m \left| \frac{\partial H}{\partial t} \right| \phi_n \right\rangle \right|^2 \qquad (14.98)$$

If the RHS is much less than 1, then $P_{n\to m}$ is much less than 1.

So if

$$\left| \frac{2}{(E_m - E_n)} \cdot \frac{T}{2\pi} \cdot \left\langle \phi_n \left| \frac{\partial H}{\partial t} \right| \phi_n \right\rangle \right|^2 << 1, \text{ then } P_{n\to m} \simeq 0$$

This is the condition for validity of adiabatic approximation.

14.14 ADIABATIC THEOREM FROM TIME-DEPENDENT PERTURBATION

It is very informative to obtain the adiabatic theorem from time-dependent perturbation theory.

Let us recall that H can be written as

$$H = H_0 + \lambda H'(t)$$

The wave function $\psi(\mathbf{r}, t)$ can be written as

$$\psi(\mathbf{r}, t) = \sum C_m(t) e^{-iE_m^{(0)}t/\hbar} u_n(\mathbf{r}) = \sum C_m(t) e^{-i\omega_m t} u_m(\mathbf{r})$$

where $$H_0 u_m(\mathbf{r}) = E_m^{(0)} u_m(\mathbf{r})$$

To the first order,

$$C_m(t) = C_m^{(0)}(t) + \lambda C_m^{(1)}(t)$$

Let us assume the initial state of the system to be $\psi(0) = u_n(\mathbf{r})$. Then $C_{mn}^{(0)} = \delta_{nm}$.

We know that $C_{mn}^{(0)}(t)$ is given by

$$C_{mn}^{(1)}(t) = \frac{1}{i\hbar}\int_0^t e^{i\omega_{mn}t'} \langle u_m | H' | u_n \rangle dt' \qquad m \neq n$$

Integrating by parts, we get

$$C_m^{(1)}(t) = \frac{1}{i\hbar}\frac{e^{i\omega_{mn}t'}}{i\omega_{mn}} \cdot H'_{mn}\Bigg|_0^t + \frac{1}{\hbar\omega_{mn}}\int_0^t dt' e^{i\omega_{mn}t'} \left\langle u_m \left| \frac{\partial H'}{\partial t} \right| u_n \right\rangle \tag{14.99}$$

Assuming $\left\langle u_m \left| \frac{\partial H'}{\partial t} \right| u_n \right\rangle$ is very small, we can neglect the second term. Then

$$C_m^{(1)}(t) = \frac{1}{i\hbar} \cdot \frac{e^{i\omega_{mn}t}}{i\omega_{mn}} \cdot \langle u_m | H' | u_n \rangle_{m\neq n}$$

$$\therefore \psi(\mathbf{r}, t) = \sum_m [C_m^{(0)}(t) + \lambda C_m^{(1)}(t)] e^{-i\omega_m t} u_m(\mathbf{r}) \tag{14.100}$$

$$= \left[\sum_m \delta_{mn} e^{-i\omega_m t} + \sum_{m\neq n} \lambda C_m^{(1)}(t) e^{-i\omega_m t}\right] u_m(\mathbf{r})$$

Note $$e^{i\omega_{mn}t} e^{-iE_m^{(0)}t/\hbar} = e^{i(\omega_m - \omega_n)t} e^{-i\omega_m t} = e^{-i\omega_n t}$$

Then

$$\psi(\mathbf{r},t) = e^{-i\omega_n t}\left[u_n(\mathbf{r}) + \sum_{m\neq n} \frac{\langle u_m | \lambda H' | u_n \rangle}{E_n^{(0)} - E_m^{(0)}} u_m(\mathbf{r})\right] \tag{14.101}$$

The quantity inside the bracket is very much like the wave function including the correction in the first-order time-independent perturbation theory. (The second term is the correction.) The only difference is that $\lambda H'$ is time independent in the time-independent perturbation theory, whereas here $\lambda H'$ depends on time. If we treat t as a parameter like mass m, then this function can be interpreted as a wave function in the first-order time-independent perturbation theory. Then, we have

$$H_0 u_n = E_n^{(0)} u_n \tag{14.102}$$

$$[H_0 + \lambda H'] w_n = E_n w_n = (E_n^{(0)} + \Delta E_n) w_n \tag{14.103}$$

More appropriately, the Equation (14.103) can be written as (remembering t as a parameter)

$$[H_0 + \lambda H'(t)] w_n(t) = E_n(t) w_n(t)$$

or $$H(t)w_n(t) = E_n(t)w_n(t)$$

$$\Delta E_n = \langle u_n |\lambda H'(t)| u_n\rangle$$

These equations suggest that the state of the system goes from u_n at $t = 0$ to state $w_n(t)$ at time t. The energy eigenvalue goes from $E_n^{(0)}$ at $t = 0$ to $E_n(t)$ at time t.

The fact that $E_n = E_n^{(0)} + \Delta E_n$ suggests that the behaviour of $E_n(t)$ is very much similar to Fig. 14.3. These equations prove the adiabatic theorem. Further, it is instructive to get the condition for validity of adiabatic theorem. This amounts to investigating the conditions under which the second term in (14.99) can be neglected. Let us write $\psi(t)$ as the Equation (14.100), without neglecting the second term in the Equation (14.99).

$$\psi(t) = \underbrace{\left[u_n + \sum_{m\neq n}\frac{\langle u_m|\lambda H'|u_n\rangle}{E_n^{(0)} - E_m^{(0)}}u_m\right]e^{-i\omega_n t}}_{w(t)}$$

$$+\left[\sum_{m\neq n}\frac{1}{\hbar\omega_{mn}}\int_0^t dt'\left\langle u_m\left|\lambda\frac{\partial H'}{\partial t'}\right|u_n\right\rangle e^{i\omega_{mn}t'}\right]e^{-i\omega_m t}u_m$$

$$= w_i(t) + \left[\sum_{m\neq n}\frac{\lambda}{\hbar\omega_{mn}}\int_0^t dt' e^{i\omega_{mn}t'}\left\langle u_m\left|\frac{\partial H'}{\partial t'}\right|u_n\right\rangle\right]e^{-i\omega_m t}u_m \tag{14.104}$$

The second term represents the transition probability amplitude from u_n to u_m. Obviously,

$$P_{n\to m} = \frac{1}{\hbar^2\omega_{mn}^4}\sin^2\frac{\omega_{mn}t}{2}\cdot\left|\lambda\left(\frac{\partial H'}{\partial t}\right)_{mn}\right|^2 \tag{14.105}$$

This is the same as the condition (14.97).

Example 14.2 An electric field is switched on at $t = -\infty$ to a harmonic oscillator, which is in ground state at $t = -\infty$. The interaction Hamiltonian is given by

$$H' = e\varepsilon_0 x e^{-t^2/\tau^2}$$

1. Determine the transition probability to other excited states at $t = \infty$ using the first-order time-dependent perturbation theory.
2. Determine the instantaneous energy eigenvalues.

Solution:

1. The transition probability from the ground state to an excite state is given by

$$P_{0\to n} = \left|C_n^{(1)}(t)\right|^2$$

$$C_n^{(1)}(t) = \frac{1}{i\hbar}\int_{-\infty}^{\infty} e^{-i\omega_{no}t'}\langle u_n|H'|u_0\rangle dt'$$

where $$\omega_{no} = \frac{(E_n - E_0)}{\hbar} = \frac{\left(n + \frac{1}{2}\right)\hbar\omega - \frac{\hbar\omega}{2}}{\hbar} = n\omega$$

$$\therefore C_n^{(1)}(t) = \frac{1}{i\hbar}\int_{-\infty}^{\infty} e^{-in\omega t'}\langle u_n|H'|u_0\rangle dt'$$

We can evaluate $\langle u_n \mid H' \mid u_0\rangle$ using the ladder operators. In terms of the ladder operators,

$$x = \left(\frac{\hbar}{2m\omega}\right)^{1/2}[a + a^{\dagger}]$$

$$\therefore \langle u_n|H'|u_0\rangle = \sqrt{\frac{\hbar}{2m\omega}}\langle u_n|(a + a^{\dagger})|0\rangle e^{-t^2/\tau^2}$$

Since $a \mid 0\rangle = 0$, we have

$$\langle u_n|H'|u_0\rangle = e\varepsilon_0\sqrt{\frac{\hbar}{2m\omega}}\cdot\langle u_n|a^{\dagger}|0\rangle\cdot e^{-t^2/\tau^2}$$

$$= e\varepsilon_0\sqrt{\frac{\hbar}{2m\omega}}\delta_{n1}e^{-t^2/\tau^2}$$

$\therefore$ The matrix element $\langle u_n|H'|u_0\rangle$ makes only the transition to the first excited state as allowed transition.

$$\therefore C_1^{(1)} = \frac{e\varepsilon_0}{i\hbar}\sqrt{\frac{\hbar}{2m\omega}}\int_{-\infty}^{\infty} e^{-t'^2/\tau^2} - e^{i\omega t'}dt'$$

Consider the term $-\frac{t'^2}{\tau^2} - i\omega t'$ in the exponential function.

$$-\frac{t'^2}{\tau^2} - i\omega t' = -(t'^2 + i\omega t'\tau^2)/\tau^2 = \frac{-1}{\tau^2}\left(t'^2 + \frac{2i\omega t'\tau^2}{2} - \frac{\omega^2\tau^4}{4}\right) - \frac{\omega^2\tau^2}{4}$$

$$= -\left[\left(t' + \frac{i\omega\tau^2}{2}\right)^2 + \frac{\omega^2\tau^4}{4}\right]\Big/\tau^2$$

$$\therefore C_1^{(1)} = \frac{e\varepsilon_0}{i\hbar}\sqrt{\frac{\hbar}{2m\omega}}\cdot\int_{-\infty}^{\infty} e^{-\left(t' + \frac{i\omega\tau^2}{2}\right)^2} e^{-\omega^2\tau^2/4}dt'$$

$$= \frac{e\varepsilon_0}{i\hbar}\cdot\sqrt{\frac{\hbar}{2m\omega}}\cdot\sqrt{\pi}e^{-\omega^2\tau^2/4}$$

$$P_{0\to 1} = \frac{\pi e^2\varepsilon_0^{\,2}}{2m\omega\hbar}\tau^2 e^{-\omega^2\tau^2/2}$$

2. The instantaneous eigenstate can be determined as follows:

$$H(t)=\frac{p^2}{2m}+\frac{1}{2}m\omega^2x^2+e\varepsilon_0xe^{-t^2/\tau^2}$$

The corresponding time-independent Hamiltonian is

$$H=\frac{p^2}{2m}+\frac{1}{2}m\omega^2x^2+kx$$

where $k=e\varepsilon_0e^{-t^2/\tau^2}$. Let us rewrite the Hamiltonian H as

$$\therefore H=\frac{p^2}{2m}+\frac{1}{2}m\omega^2\left[x^2+\frac{2kx}{x\omega^2}+\frac{k^2}{m^2\omega^4}\right]-\frac{1}{2}m\omega^2\cdot\frac{k^2}{m^2\omega^4}$$

$$=\frac{p^2}{2m}+\frac{1}{2}m\omega^2\left(x+\frac{k}{m\omega^2}\right)^2-\frac{1}{2}\frac{k^2}{m\omega^2}$$

The energy eigenvalue of this Hamiltonian is

$$E_n=\left(n+\frac{1}{2}\right)\hbar\omega-\frac{1}{2}\frac{k^2}{m\omega^2}$$

The instantaneous eigenvalue $E_n(t)$ is given by

$$E_n(t)=\left(n+\frac{1}{2}\right)\hbar\omega-\frac{1}{2}\cdot\frac{e^2\varepsilon_0{}^2}{m\omega^2}\cdot e^{-2t^2/\tau^2}$$

And the corresponding eigenfunction is given by

$$\phi_n(t)=N_n\exp\left(-\frac{\alpha^2}{2}\left(x-\frac{k}{m\omega^2}\right)^2\right)H_n\left(\alpha\left(x-\frac{k}{m\omega^2}\right)\right)$$

14.15 BERRY PHASE

Let us go back to the Equation (14.81) or (14.94).

$$\psi(t)=\sum_m C_m(t)e^{-i\theta_m}\phi_m(t)$$

From adiabatic theorem, we know that for a slowly varying Hamiltonian, the states of a system continue to evolve in the instantaneous eigenstate $\phi_n(t)$ if the system starts in the state $\phi_n(0)$. This amounts to dropping the second term in (14.85) for $\dot{C}_m(t)$.

So, in the adiabatic approximation, the equation for $\dot{C}_m(t)$ is given by

$$\dot{C}_m(t)=-C_m(t)\langle u_m\,|\,\dot{u}_m\rangle \tag{14.106}$$

$$=-C_m(t)(i\alpha(t))$$

The solution to this equation is

$$C_m(t) = C_m(0)e^{-i\int_0^t \alpha_m(t')dt'} = C_m(0)e^{i\gamma_m(t)} \tag{14.107}$$

Since $C_m(0) = \delta_{mn}$

$$C_m(t) = \delta_{mn}e^{ir_m(t)}$$

or $C_n(t) = e^{i\gamma_n(t)}$

$$\therefore \psi(t) = e^{i\gamma_n(t)}e^{-i\theta_n(t)}\phi_n(t) \tag{14.108}$$

Let us rephrase these results in such a way that it brings out the Berry phase clearly.

The slowly varying Hamiltonian $H(t)$ is written as

$$H(t) = H(R_i(t)) \quad i = 1, 2, 3 \ldots \tag{14.109}$$

$R_i(t)$ are the parameters which can be varied very slowly. For the sake of clarity, let us restrict ourselves to three parameters. Therefore,

$$\mathbf{R}(t) = R_1(t)\mathbf{e}_1 + R_2(t)\mathbf{e}_2 + R_3(t)\mathbf{e}_3 \tag{14.110}$$

i.e., $\mathbf{R}(t)$ is a vector in the parameter space.

The eigenvalue equation is

$$H(R_i(t))\phi_n(R_i(t)) = E_n(R_i(t))\phi_n(R_i(t)) \tag{14.111}$$

The adiabatic theorem means

$$\psi(t) = \begin{cases} \phi_n(R_i(0)) & t = 0 \\ e^{-i\theta_n}e^{i\gamma_n}\phi_n(R_i(t)) & t \neq 0 \end{cases} \tag{14.112}$$

where $\theta_n = \frac{1}{\hbar}\int_0^t E_n(R_i(t'))\,dt'$

and

$$\gamma_n = i\int_0^t \langle \phi_n(R_i(t') \,|\, \dot{\phi}_n(R_i(t'))\rangle dt' \tag{14.113}$$

$\theta_n(t)$ is called dynamic phase and $\gamma_n(t)$ is called geometric phase. In the case of stationary eigenstate, we have $\theta_n = E_n t/\hbar$ (or $e^{-i\theta_n} = e^{-iE_n t/\hbar}$).

$$\begin{aligned}\frac{d}{dt}\phi_n(R_i(t)) &= \frac{\partial \phi_n}{\partial R_1}\frac{\partial R_1}{\partial t} + \frac{\partial \phi_n}{\partial R_2}\cdot\frac{\partial R_2}{\partial t} + \frac{\partial \phi_n}{\partial R_3}\cdot\frac{\partial R_3}{\partial t} \\ &= \left(\hat{\mathbf{e}}_1\frac{\partial \phi_n}{\partial R_1} + \hat{\mathbf{e}}_2\frac{\partial \phi_n}{\partial R_2} + \hat{\mathbf{e}}_3\frac{\partial \phi_n}{\partial R_3}\right)\cdot\left(\hat{\mathbf{e}}_1\frac{dR_1}{dt} + \hat{\mathbf{e}}_2\frac{dR_2}{dt} + \hat{\mathbf{e}}_3\frac{dR_3}{dt}\right) \\ &= \nabla_R\phi_n\cdot\frac{d\mathbf{R}}{dt}\end{aligned}$$

It is important to realize that the derivative as well as the basis vectors $\hat{\mathbf{e}}_1$, $\hat{\mathbf{e}}_2$, $\hat{\mathbf{e}}_3$ are with respect to the parameter space; i.e., we associate a separate space for the variation in the parameters (Fig. 14.4).

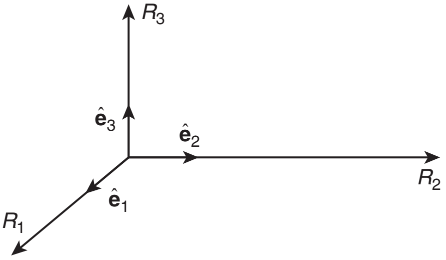


Fig. 14.4 Parameter space

$$\therefore \gamma_n = i\int_0^t \langle \phi_n|\nabla_R \phi_n\rangle \cdot \frac{d\mathbf{R}}{dt}\,dt$$

$$= i\int_{R(0)}^{\mathbf{R}(t)} \langle \phi_n\, |\nabla_R \phi_n\rangle \cdot d\mathbf{R} \tag{14.114}$$

Consider a special case when the parameters return to the same values after a time interval T; i.e., $\mathbf{R}(T) = \mathbf{R}(0)$. In that case, the integral becomes

$$\gamma_n(\mathrm{C}) = \gamma_n(T) = i\oint_C \langle \phi_m\, |\nabla_R \phi_n\rangle \cdot d\mathbf{R} \tag{14.115}$$

where C is a closed contour in the parameter space representing the time evolution of $\mathbf{R}(t)$. The phase as given by (14.115) is a line integral over a closed contour in the parameter space, and it is called Berry phase.

This has a very interesting structure. This is very much similar to $\oint \mathbf{A}_{em} \cdot d\mathbf{r}$, where $\mathbf{A}_{em}$ is the vector potential in classical electromagnetism.

So γ_m can be written as

$$\gamma_n = \oint \mathbf{A}_n(t)\cdot d\mathbf{R}(t)$$

where
$$\mathbf{A}_n = i\langle \phi_n|\nabla_R \phi_n\rangle$$

Note that though we use the symbol $\mathbf{A}_n$ here, it is not electromagnetic vector potential $\mathbf{A}$. But it behaves likes electromagnetic vector potential.

As t changes, the value of the parameter changes, and so $\mathbf{R}(t)$ traces a curve in the parameter space. At $t = T$, the parameter $\mathbf{R}(T)$ comes back to initial print $\mathbf{R}(0)$. So $\mathbf{R}(t)$ traces a closed loop C in the parameter space. Since we have restricted the number of parameters to three, the parameter space is three. All the possible values of $\mathbf{R}(t)$ together from a geometrical surface in the parameter space. This geometrical surface may be a plane or curved surface like sphere or a surface of any other shape. So the closed loop C lies on this surface. The value of $\oint \mathbf{A}_n(t)\cdot d\mathbf{R}$ depends on the geometrical nature of surface in the parameter space. So Berry's phase is a geometrical phase.

The Berry phase can be shown to be gauge invariant. We have obtained the expression for Berry phase using the eigenstates $\{\phi_n(t)\}$. Had we chosen a different basis like $\{\phi_n'\}$, will we still get the same Berry phase? This can be checked easily as follows. Let us consider the following transformation:

$$\phi_n \to \phi_n'(t) = e^{i\xi_n(\mathbf{R})}\phi_n(t) \tag{14.116}$$

This is obviously the well-known gauge transformation. Of course, we have to recognize that this gauge transformation also happens in parameter space.

The gauges transformation induces a transformation of $\mathbf{A}_n(\mathbf{R})$ as follows:

$$\mathbf{A}_n(\mathbf{R}) \rightarrow \mathbf{A}'_n(\mathbf{R}) = i\langle \phi'_n | \dot{\phi}'_n \rangle$$

$$\nabla_R \phi'_n = \nabla_R e^{i\xi_n(\mathbf{R})} \phi_n(\mathbf{R})$$

$$= i\nabla_R \xi_n(\mathbf{R}) e^{i\xi_n(\mathbf{R})} \phi_n(\mathbf{R}) + e^{i\xi_n(\mathbf{R})} \nabla_R \phi_n$$

$$\therefore i\langle \phi' | \nabla_R \phi'_n \rangle = i\langle \phi_n | \nabla_R \phi_n \rangle + \nabla_R \xi_n(\mathbf{R})$$

$$\mathbf{A}'_n = \mathbf{A}_n + \nabla_R \xi_n(\mathbf{R}) \tag{14.117}$$

$$\therefore \gamma'_n(c) = \oint \mathbf{A}'_n \cdot d\mathbf{R} = \oint \mathbf{A}_n \cdot d\mathbf{R} + \underbrace{\oint \nabla_R \xi(\mathbf{R}) \cdot d\mathbf{R}}_{0} = \gamma_n(c)$$

The transformations (14.116) and (14.117) are the well-known pairs of gauge transformations. So the fact that $\gamma'_n(C) = \gamma_n(C)$ means that $\gamma_n(C)$ is a gauge invariant quantity. i.e., Berry phase is a gauge invariant quantity. Therefore, the existence of Berry phase is independent of the choice of the basis. The consequence of this result is that Berry phase cannot be avoided by changing the basis states.

14.16 BERRY PHASE FOR ROTATING MAGNETIC FIELDS

This is the simplest example illustrating the non-trivial character of Berry phase. Consider a particle in the magnetic field $\mathbf{B}$ given by

$$\mathbf{B} = B_0 \sin \upsilon \cos \omega t \, \hat{\mathbf{x}} + B_0 \sin \upsilon \sin \omega t \, \hat{\mathbf{y}} + B_0 \cos \upsilon \, \hat{\mathbf{z}}.$$

$$B_x^2 + B_y^2 + B_z^2 = B_0^2$$

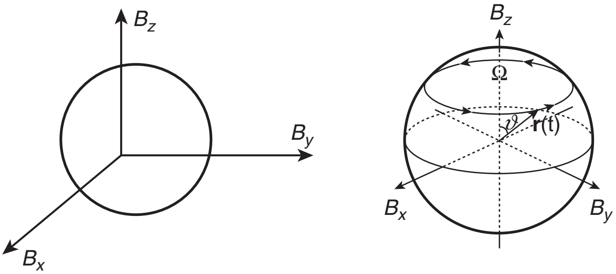


Fig. 14.5 Parameter space for rotating magnetic field

The above equation is the equation of a sphere in the parameter space. (The parameters are B_x, B_y and B_z.) Therefore, as time progresses, $\mathbf{B}(t)$ traces a circle as shown in Fig. 14.5. Let us choose the spherical polar coordinate r, υ and ϕ. It is to be emphasized that these are coordinates in parameter space:

$$R_1 = r = B; \; R_2 = \upsilon; \; R_3 = \varphi = \omega t$$

The Hamiltonian $H(\mathbf{R},t)$ is given by

$$H(\mathbf{R},t) = \boldsymbol{\mu}\cdot\mathbf{B} = \mu_B B_0 \begin{bmatrix} \cos\upsilon & e^{-i\omega t}\sin\upsilon \\ e^{i\omega t}\sin\upsilon & -\cos\upsilon \end{bmatrix}$$

The eigenvalues of this Hamiltonian are $E(\uparrow) = \mu_B B_0$ and $E(\downarrow) = -\mu_B B_0$.

$$H(\mathbf{R},t)\chi(\uparrow;t) = \mu_B B_0 \chi(\uparrow;t)$$

$$H(\mathbf{R},t)\chi(\downarrow;t) = -\mu_B B_0 \chi(\downarrow;t)$$

$$\chi(\uparrow;t) = \begin{pmatrix} \cos\upsilon/2 \\ e^{i\omega t}\sin\upsilon/2 \end{pmatrix} = \begin{pmatrix} \cos\upsilon/2 \\ e^{i\varphi}\sin\upsilon/2 \end{pmatrix}$$

$$\chi(\downarrow;t) = \begin{pmatrix} -\sin\upsilon/2 \\ e^{i\omega t}\cos\upsilon/2 \end{pmatrix} = \begin{pmatrix} -\sin\upsilon/2 \\ e^{i\varphi}\cos\upsilon/2 \end{pmatrix}$$

The operator ∇_R in spherical coordinate (in parameter space) is given by

$$\nabla_R = \hat{\mathbf{r}}\frac{\partial}{\partial r} + \hat{\upsilon}\frac{1}{r}\frac{\partial}{\partial \upsilon} + \hat{\phi}\frac{1}{r\sin\upsilon}\frac{\partial}{\partial\phi}$$

$$\nabla_R\chi(\uparrow;t) = \hat{\upsilon}\frac{1}{r}\frac{\partial}{\partial\upsilon}\chi(\uparrow;t) + \frac{\hat{\varphi}}{r\sin\upsilon}\frac{\partial}{\partial\varphi}\chi(\uparrow;t)$$

$$= \hat{\upsilon}\frac{1}{r}\begin{bmatrix} -\frac{1}{2}\sin\frac{\upsilon}{2} \\ \frac{1}{2}e^{i\varphi}\cos\frac{\upsilon}{2} \end{bmatrix} + \hat{\varphi}\frac{1}{r\sin\upsilon}\begin{bmatrix} 0 \\ i\varphi\sin\frac{\upsilon}{2} \end{bmatrix}$$

$$\therefore \langle \chi(\uparrow,t) \,|\, \nabla_R \chi(\uparrow,t)\rangle = \chi(\uparrow;t)^\dagger \nabla_R \chi(\uparrow;t)$$

$$= \frac{i}{r\sin\upsilon}\sin^2\frac{\upsilon}{2}\hat{\varphi}$$

For our choice of magnetic field,

$$d\mathbf{R} = r\sin\upsilon\, d\varphi\hat{\varphi}$$

$$\therefore \oint_c i\langle\chi(\uparrow,t)\,|\,\nabla_R\chi(\uparrow,t)\rangle\cdot d\mathbf{R} = -\int_0^{2\pi}\frac{\sin^2\upsilon/2}{r\sin\upsilon}\cdot r\sin\upsilon\, d\varphi$$

$$= -2\pi\sin^2(\upsilon/2) = -\pi(1-\cos\upsilon)$$

$$\therefore \gamma_C = -\pi(1-\cos\upsilon)$$

Note that γ_C is independent T. It depends only on the geometrical property of parameter space. In fact, we can show that $\pi(1-\cos\upsilon)\pi = \frac{1}{2}\Delta\Omega$ where $\Delta\Omega$ is the solid angle subtended at the centre by the part of the spherical surface for which C is the boundary.

The dynamical phase $\theta_n(t)$ is given by

$$\theta_n = -\frac{1}{\hbar}\int_0^T E(\uparrow,t)\,dt' = -\frac{1}{\hbar}\mu_B B_0 T$$

$$\therefore \psi(B,T) = e^{-i\mu_B B_0 T/\hbar} e^{-i(1-\cos\theta)\pi} \chi(\uparrow,t)$$

Berry's phase is not the only geometrical phase. There are a number of geometrical phases similar to Berry's phase, and these phases have been experimentally observed.

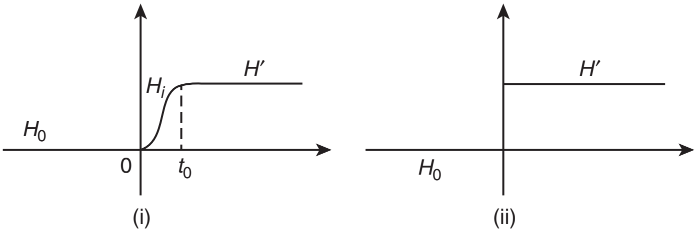
14.17 SUDDEN APPROXIMATION

In the previous sections, we discussed the adiabatic approximation which is suitable for the systems for which the Hamiltonian is varying very slowly with respect to time. The time dependence in the sudden approximation refers to situation which represents the other extreme.

The Hamiltonian is H_0 for $t < 0$, and the interaction is switched on at $t = 0$. The Hamiltonian changes very fast in a small interval of time t_0, and it becomes H' at t_0 and remains the same after t_0. So we have

$$H = \left\{\begin{array}{ll} H_0 & t \le 0 \\ H_i(t) & 0 \le t \le t_0 \\ H' & t \ge t_0 \end{array}\right| \tag{14.118}$$

H_0 and H' are of our interest here.



(i) (ii)

Fig. 14.6 Actual Hamiltonian and sudden approximation

The sudden approximation makes the following approximation for the Hamiltonian H.

$$H \simeq \begin{cases} H_0 & t \le 0 \\ H' & t \ge 0 \end{cases} \tag{14.119}$$

This looks valid approximation provided t_0 is very small. Our task now is to make this statement more precise. Under what conditions the Hamiltonian given in (14.118) can be replaced by (14.119)?

First, let us analyze the evolution of the state under the Hamiltonian (14.119).

Let $\{u_k\}$ and $\{v_n\}$ be the sets of eigenstates of the Hamiltonians H_0 and H'.

$$\therefore H_0 u_k = E_k^{(0)} u_k$$

and $$H' v_n = E'_n v_n$$

The general solution of the Schrödinger equation can be written as

$$\psi = \begin{cases} \psi_0 = \sum_k a_k e^{-iE_k^{(0)} t/\hbar} u_k & t \le 0 \\ \psi' = \sum_k c_n e^{-iE'_n t/\hbar} v_n & t \ge 0 \end{cases}$$

Obviously, the coefficient $\{a_k\}$ describes the initial state and the coefficient $\{c_n\}$ describes the final state. Our interest always is to determine the transition probability from initial state to final state. Therefore, we are interested in determining c_n in terms of a_n.

The wave function $\psi(t)$ should be continuous at all times. Therefore, at $t = 0$, the wave function should be continuous.

i.e., $$\psi_0(t=0) = \psi'(t=0)$$

$$\therefore \sum_k a_k u_k = \sum_n c_n v_n$$

$$\therefore c_n = \langle v_n | \sum_k a_k u_k \rangle$$

$$= \sum_k a_k \langle v_n | u_k \rangle \tag{14.120}$$

Suppose the system is in state u_i at $t = 0$.

$$\therefore a_k = \delta_{ki}$$

$$c_n = \sum_k \delta_{ki} \langle v_n | u_k \rangle = \langle v_n | u_i \rangle$$

This is the essence of sudden approximation. The transition probability $P_{i\to n}$ from the eigenstate u_i before $t = 0$ to the eigenstate v_n after $t = 0$ is given by

$$P_{i\to n} = |c_n|^2 = |\langle v_n | u_i \rangle|^2 \tag{14.121}$$

To determine the validity of sudden approximation, let us write the Hamiltonian as

$$H = \begin{cases} H_0 & t \le 0 \\ H_i(t) & 0 \le t \le t_0 \\ H' & t \ge t_0 \end{cases} \tag{14.122}$$

Since we are interested only in estimating the validity of sudden approximation, let us take H_i also as constant in time. So we now have three sets of eigenstates: $\{u_k\}$, $\{w_s\}$, $\{v_n\}$.

$$\begin{aligned} H_0 u_k &= E_k^{(0)} u_k \\ H_i w_s &= E_s^{(i)} w_s \\ H' v_n &= E'_n v_n \end{aligned}$$

The general solution ψ is given by

$$\psi = \begin{cases} \psi_0 = \sum_k a_k u_k e^{-iE_k^{(0)}t/\hbar} & t \le 0 \\ \psi_i = \sum_s b_s w_s e^{-iE_s^{(i)}t/\hbar} & 0 \le t \le t_0 \\ \psi' = \sum_n c_n v_n e^{-iE'_n\,t/\hbar} & t \ge t_0 \end{cases}$$

The continuity of the wave function at $t = 0$ implies

$$\psi_0(t=0) = \psi_i(t=0)$$

$$\therefore \sum a_k u_k = \sum b_s w_s$$

$$\therefore b_s = \sum_k a_k \langle w_s \mid u_k \rangle \tag{14.123}$$

The continuity of the wave function at t_0 implies that

$$\psi_i(t=t_0) = \psi'\,(t=t_0)$$

$$\therefore \sum b_s w_s e^{-iE_s^{(0)}t_0/\hbar} = \sum c_m v_m e^{-iE'_m\,t_0/\hbar}$$

Taking scalar product with v_n, we get

$$c_n e^{-iE'_n t_0/\hbar} = \sum b_s \langle v_n \mid w_s \rangle e^{-iE_s^{(i)}\,t_0/\hbar}$$

$$\therefore c_n = \sum b_s \langle v_n \mid w_s \rangle e^{i(E'_n - E^{(i)}{}_s)t_0/\hbar}$$

Making use of (14.123),

$$\begin{aligned} c_n &= \sum_s \sum_k a_k \langle w_s | u_k \rangle \langle v_n | w_s \rangle e^{i(E_n{}' - E_s^{(i)})t_0/\hbar} \\ &= \sum_k \sum_s a_k \langle v_n | w_s \rangle \langle w_s\ | u_k \rangle e^{i(E'_n - E_s^{(i)})t_0/\hbar} \end{aligned} \tag{14.124}$$

If the exponential term has been 1 (or $t_0 \simeq 0$), then the summation over s becomes $\Sigma\, |w_s\rangle\langle w_s|$, which is an identity operator making c_n to be the same as (14.120). For very small t_0, we can write

$$e^{i(E'_n - E_s^{(i)})t_0/\hbar} \simeq 1 + \frac{i\left(E'_n - E_s^{(i)}\right)t_0}{\hbar}$$

$$\begin{aligned} c_n &= \sum_k \sum_s \langle v_n | w_s \rangle \langle w_s | u_k \rangle a_k + \sum_k \sum_s \langle v_n | w_s \rangle \langle w_s | u_k \rangle \left(i\,\frac{(E_n{}' - E_s^{(i)})t_0}{\hbar} \right) \\ &= \sum_k \langle v_n | u_k \rangle a_k + \sum_k \sum_s \langle v_n | w_s \rangle \langle w_s\ | u_k \rangle \left(i\,\frac{(E_n{}' - E_s^{(i)})t_0}{\hbar} \right) \end{aligned}$$

The first term represents sudden approximation, and the second term is the correction to the sudden approximation.

The correction term can be written as

$$\Delta c_n = \sum_k \sum_s \langle v_n | i \frac{(E_n' - E_s^{(i)})t_0}{\hbar} |w_s\rangle\langle w_s|u_k\rangle$$

$$= \sum_k \sum_s \langle v_n | i \frac{(H_n' - H^{(i)})t_0}{\hbar} |w_s\rangle\langle w_s|u_k\rangle$$

Since H' and H are operators, we can take $|w_s\rangle\langle w_s| = 1$.

$$\therefore \Delta c_n = \langle v_n | \frac{it_0}{\hbar}(H' - H_i)|u_k\rangle$$

The correction Δc_n can be estimated, and if $\Delta c_n << 1$, then the sudden approximation is valid. This is possible when t_0 is very small.

Examples 14.3 Initially, the particle is confined to a box of size a in the ground state. The size of the box is suddenly extended to $2a$. Determine the probability of finding the particle in the ground state and the first excited state.

Solution: The Hamiltonian H is given below:

For $t < 0$, $H_0 = \begin{cases} 0 & 0 \le x \le a \\ \infty & \text{Otherwise} \end{cases}$

For $t > 0$, $H' = \begin{cases} 0 & 0 \le x \le 2a \\ \infty & \text{Otherwise} \end{cases}$

$$H_0 u_n = E_n^{(0)} u_n$$

The eigenvalues and the eigenfunctions of the H_0 are

$$E_n^{(0)} = \frac{\hbar^2\pi^2}{2ma^2} \qquad u_n = \begin{cases} \sqrt{\frac{2}{a}} \sin\frac{n\pi x}{a} & 0 < x < a \\ 0 & \text{Otherwise} \end{cases}$$

The eigenvalues and the eigenfunctions of H' are

$$E_n' = \frac{\hbar^2}{2m} \cdot \frac{\pi^2}{4a^2} \qquad v_n = \begin{cases} \sqrt{\frac{2}{2a}} \sin\frac{n\pi x}{2a} & 0 < x < 2a \\ 0 & \text{Otherwise} \end{cases}$$

The transition probability to make a transition from u_1 to v_n is given by $\langle v_n | u_1\rangle$.

$$
\begin{aligned}
\langle v_n \mid u_1 \rangle &= \int_0^{2a} v_n^* u_1 dx \\
&= \int_0^{a} v_n^* u_1 dx + \underbrace{\int_a^{2a} v_n^* u_1 dx}_{0} \\
&= \sqrt{\frac{2}{2a}} \cdot \sqrt{\frac{2}{a}} \cdot \int_0^{a} \sin\frac{n\pi x}{2a} \cdot \sin\frac{\pi x}{a} dx \\
&= \frac{1}{\sqrt{2a}} \int_0^{a} \frac{1}{2}\left[\cos\left(\frac{n\pi x}{2a} - \frac{\pi x}{a}\right) - \cos\left(\frac{n\pi x}{2a} + \frac{\pi x}{2a}\right)\right] dx \\
&= \frac{1}{2\sqrt{2}} \left\{ \frac{\sin\left(\frac{n}{2}-1\right)\pi}{\left(\frac{n}{2}-1\right)\pi} - \frac{\sin\left(\frac{n}{2}+1\right)\pi}{\left(\frac{n}{2}+1\right)\pi} \right\}
\end{aligned}
$$

$$
\begin{aligned}
\langle v_1 \mid u_1 \rangle &= \frac{4\sqrt{2}}{3\pi} \\
\langle v_2 \mid u_1 \rangle &= 0
\end{aligned}
$$

The probability of finding the particle in ground state after $t = 0$ is

$$P_{1\to 1} = |\langle v_1 \mid u_1 \rangle|^2 = \frac{32}{9\pi^2}$$

Example 14.4 A simple harmonic oscillator was subjected to an electric field suddenly at $t = 0$. Assuming the oscillator was in ground state when $t < 0$, determine the probability of finding the system in the ground state after $t > 0$.

Solution: The Hamiltonian is given by

$$
H = \begin{cases} H_0 = \dfrac{p^2}{2m} + \dfrac{1}{2} n\omega^2 x^2 & t \le 0 \\[2ex] H' = \dfrac{p^2}{2m} + \dfrac{1}{2} m\omega^2 x^2 + e\varepsilon x & t \ge 0 \end{cases}
$$

The Hamiltonian H' can be rewritten as

$$H' = \frac{p^2}{2m} + \frac{1}{2} m\omega^2 \left(x + \frac{e\varepsilon}{m\omega^2}\right)^2 - \frac{e^2\varepsilon^2}{m\omega^2}$$

The ground state of H_0 is

$$u_0 = \left(\frac{\alpha}{\sqrt{\pi}}\right)^{1/2} \exp\left(-\frac{\alpha^2 x^2}{2}\right)$$

The ground state of H' is

$$v_0 = \left(\frac{\alpha}{\sqrt{\pi}}\right)^{1/2} \exp\left(-\frac{\alpha^2}{2}\left(x+\frac{e\varepsilon}{m\omega^2}\right)^2\right)$$

The probability of finding the system in ground state V_0 is given by

$$P_{0\to 0} = |\langle v_0|u_0\rangle|^2$$

$$\langle v_0\,|u_0\rangle = \int_{-\infty}^{\infty} \frac{\alpha}{\sqrt{\pi}}\cdot \exp\left(-\frac{\alpha^2}{2}\left(x+\frac{e\varepsilon}{m\omega^2}\right)^2\right)\cdot \exp\left(-\frac{\alpha^2}{2}x^2\right)dx$$

$$= \int_{-\infty}^{\infty} \frac{\alpha}{\sqrt{\pi}}\cdot \exp\left(-\frac{\alpha^2}{2}\left[2x^2+\frac{e^2\varepsilon^2}{m^2\omega^4}+\frac{2e\varepsilon}{m\omega^2}x\right]\right)dx$$

$$= \int_{-\infty}^{\infty} \frac{\alpha}{\sqrt{\pi}}\cdot \exp\left(-\alpha^2\left[x+\frac{1}{2}\frac{e\varepsilon}{m\omega^2}\right]^2\right)\exp\left(-\frac{1}{4}\alpha^2\frac{e^2\varepsilon^2}{m^2\omega^4}\right)dx$$

$$= \frac{\alpha}{\sqrt{\pi}}\cdot\frac{\sqrt{\pi}}{\alpha}\cdot \exp\left(-\frac{1}{4}\alpha^2\frac{e^2\varepsilon^2}{m^2\omega^4}\right)$$

$$\therefore\quad \left|\langle v_0\,|u_0\rangle\right|^2 = \exp\left(-\frac{1}{2}\alpha^2\frac{e^2\varepsilon^2}{m^2\omega^4}\right)$$

Probability of finding the oscillator in an excited state is

$$1-\left|\langle v_0\,|u_0\rangle\right|^2 = 1-\exp\left(-\frac{1}{2}\frac{\alpha^2e^2\varepsilon^2}{m^2\omega^4}\right)$$

EXERCISES

1. What is the distinction between the time-independent and time-dependent perturbation theory?
2. Why do you need gauge-fixing term for electromagnetic fields?
3. What is meant by principle of detailed balance?
4. In the inequality $\Delta E\,\Delta t \sim \hbar$, can we define Δt as uncertainty in time?
5. In what way the inequalities $\Delta x\Delta p \sim \hbar$ and $\Delta E\,\Delta t \sim \hbar$ differ?
6. In deriving the uncertainly relation, we used the commutator $[x, p] = i\hbar$. Why cannot we use $[E,t]=i\hbar$ to get the uncertainty relation $\Delta E\,\Delta t \sim \hbar$?
7. The transition probability in the time-dependent perturbation theory under harmonic perturbation is given by

$$P_{i\to n} = \frac{2\pi t}{\hbar}|\lambda V_{ni}|^2\delta(E_n - E_i - \hbar\omega)$$

   When $E_n = E_i + \hbar\omega$, this becomes infinite. How do we justify this expression?
8. What is meant by adiabatic theorem?

9. In the case of time-independent Hamiltonian, $Hu_n = E_n u_n$ and $\psi(\mathbf{r},t) = e^{-iE_n t/\hbar} u_n(\mathbf{r})$. For time-dependent Hamiltonian, can we write $\psi(\mathbf{r},t) = e^{-iE_n(t)t} u_n(\mathbf{r},t)$?
10. Under what conditions we can have instantaneous eigenstate?
11. Can we define Berry's phase for stationary states?
12. Why do you call Berry's phase as geometric phase?
13. The frequency of a harmonic oscillator is changed from ω to ω' suddenly at $t = 0$. The initial state of the harmonic oscillator is the ground state corresponding to $\hbar\omega/2$. Determine the probability of finding the oscillator in (i) ground state (ii) excited state after $t > 0$.

REFERENCES

1. S. Gasirowicz, 1974. *Quantum Physics.* New York: John Wiley & Sons.
2. I. Schiff, 1968. *Quantum Mechanics*. New Delhi: McGraw-Hill International Book Company.
3. R. L. Liboff, 1980. *Introductory Quantum Mechanics*. San Francisco: Holder-Day Inc.
4. D. J. Griffiths, 2006. *Introduction to Quantum Mechanics*. Pearson Education.
5. B. H. Bransden and C. J. Joachain, 1989. *Introduction to Quantum Mechanics*, English Language Book Society, Longman.

15

The Scattering Theory

The major source of knowledge about sub-nuclear world is through the scattering experiments. For instance, it is from Rutherford's scattering experiments we got the planetary model for atoms. Scattering experiments play an important role in the development of elementary particle physics. Of course, they are in relativistic region, and it requires relativistic quantum field theory. In this chapter, we present the scattering theory in non-relativistic quantum mechanics.

Diverse ideas are presented here, and hence, an outline of this chapter is given here. First, the scattering cross section is introduced, the boundary condition for the Schrödinger wave function is discussed and the relation between scattering cross section and scattering amplitude is developed. Then, the partial wave analysis is considered. The concept of phase shift δ_l is introduced, and the relation between scattering cross section and phase shift δ_l has been developed. Two expressions for δ_l have been obtained: one in the form of integral and the second using the continuity of logarithmic derivative of the wave function.

Then, low-energy scattering is presented. We introduce the basic ideas of *s*-wave scattering, and these are illustrated using square well potential. Square well potential serves as the best pedagogic tool to present the major ideas of low-energy scattering, where it is enough to consider the scattering of partial waves corresponding to $l = 0$ case alone. The behaviour of phase shift δ_0 and the total cross section σ_0 for various depths of the square well potential V_0 are considered in detail. The concept of resonant and non-resonant scattering is presented. Scattering length for different potential strength V_0 is considered. Finally, Breit–Wigner formula is established.

We next develop Born approximation for scattering amplitude, using Green's function. It is applied to Coulomb scattering. The validity of Born approximation is considered for various potentials. We also obtain the first-order Born approximation using time-dependent perturbation theory.

Last, we consider two distinct topics. The first is the relationship for scattering cross sections in centre of mass frame and laboratory frame. The second is the scattering of identical particles, which takes into account the consequence of principle of indistinguishability.

15.1 SCATTERING CROSS SECTION

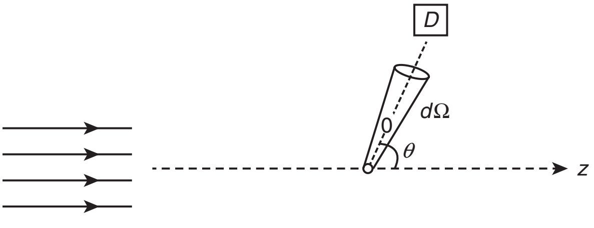


Fig.15.1 A standard scattering experiment

The standard scenario of a scattering experiment is shown in Fig.15.1. There is an incident beam of particles: there is a target and there is a scattered beam. We assume the incident beam to be a collection of non-interacting particles, i.e., there is no interaction among these particles. Initially, they are very far away from the target, and hence they behave like free particles. (Therefore, we can represent the incident particles by a plane wave function.) Once they are in the vicinity of the target, there is an interaction between the target and the incident particles, which results in the incident particles being scattered in all directions. After being scattered, the particles move very far away from the target to reach the detectors, where the influence of the target will not be felt by the scattered particles. (We assume that the interaction potential between the target and the incident particle is a short-range potential.)

Let ΔN be the number of particles scattered per second into the solid angle $\Delta\Omega$. We expect this number to be proportional to the solid angle $\Delta\Omega$ and the flux of incident particles F (number of particles crossing per unit area per unit sec).

The standard way of writing ΔN is

$$\Delta N = \left(\frac{d\sigma}{d\Omega}\right) \cdot F \cdot \Delta\Omega \tag{15.1}$$

Let us first establish that the quantity $d\sigma$ has the dimension of area. A simple dimensional analysis will indicate this. Since ΔN is the number of particles scattered in the solid angle ΔN per second, $[\Delta N] = T^{-1}$. The dimension of the flux is $[L]^{-2}T^{-1}$.

$$\begin{aligned} \therefore [\Delta N] &= [d\sigma] \cdot [F] \\ T^{-1} &= [d\sigma] \cdot L^{-2}T^{-1} \end{aligned}$$

Therefore, $d\sigma$ has the dimension of area, and hence it is called scattering cross section. The quantity $\frac{d\sigma}{d\Omega}$ is called differential cross section. The total cross section σ is given by

$$\sigma = \int \frac{d\sigma}{d\Omega} \cdot d\Omega \tag{15.2}$$

The unit used for σ is barn. It is defined as

$$1 \text{ barn} = 10^{-28} m^2$$

Significance of $d\sigma$

The Equation (15.1) can be rewritten as

$$\left.\begin{array}{l} \Delta N \text{ particles per second} \\ \text{in the solid angle } \Delta\Omega \end{array}\right\} = F d\sigma$$

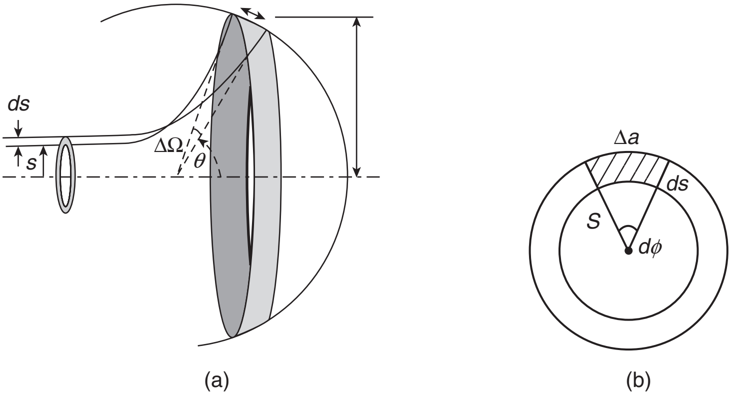
$$= \text{Number of particles crossing the area } d\sigma \text{ in 1 second}$$

So ΔN is the number of particles which cross the area $d\sigma$ that are scattered into the solid angle $\Delta\Omega$. In other words, ΔN particles crossing the area $d\sigma$ have been removed from the incident beam. This picture is used both in classical and quantum physics. We have to bear in mind that the above equation is true for both classical and quantum physics. However, there is a subtle difference in the interpretation of $d\sigma$ in classical physics and quantum physics.

There are two kinds of quantities which have the dimensions of area. One is the geometrically measured area Δa, located at a specific geometrical point at $\mathbf{r}$. In spherical coordinates, it is given by $r^2 \sin\theta d\theta d\phi$, and in cylindrical coordinates system, it is given by $s\, ds\, d\phi$. The other quantity is $d\sigma$ (which is also a function of θ and ϕ (or s and ϕ in cylindrical coordinates)), but it is calculated from the expression $\Delta N/F$. Are these two quantities one and the same? Are they related to each other? In classical physics, they are the same, and in quantum mechanics, they are different.

Let us take Rutherford scattering as an example to bring the subtle difference between classical and quantum physics in scattering process.

The Rutherford scattering in classical physics is shown in Fig.15.2.



(a) (b)

Fig.15.2 Rutherford scattering

S in Fig. 15.2(a) is the impact parameter. A circular ring of width ds is shown in Fig.15.2(b). Δa is a small area in this ring. The location of this area is specified by (z, s, ϕ), and it is given by $\Delta a = s d\phi ds$. Using Newtonian mechanics, we can show that for all the particles which cross the area Δa, their trajectories go through the region enclosed by the solid angle $\Delta\Omega$. i.e., all the particles which cross the area Δa are scattered into the solid angle $\Delta\Omega$. Since all the particles crossing the area Δa are scattered in the solid angle $\Delta\Omega$, we have

$$d\sigma = \Delta a$$

$$\therefore\ \Delta a = d\sigma \left(\text{which is calculated from } \frac{\Delta N}{F}\right)$$

So, in classical physics, the scattering cross section is identified as the geometrical cross section at the location (z, s, ϕ) or (r, θ, ϕ).

Now let us turn our attention to quantum mechanics. Let us recall that in quantum mechanics, there is no concept of trajectory. Let us compare and contrast the answer to the following question from classical and quantum physics. In the case of particles crossing the geometrical area Δa (located at (z, s, ϕ)), where do they end up? In classical physics, the existence of trajectories leads them into the solid angle $\Delta\Omega$. In quantum mechanics, we cannot tell whether these particles will be found in solid angle $\Delta\Omega$. We can tell only the probability of finding these particles in the solid angle $\Delta\Omega$. The particles crossing Δa may or may not be found in the solid angle $\Delta\Omega$. Therefore, there is no relationship between the geometrical area Δa and $d\sigma$ computed from $\Delta N/F$. They will not be the same even numerically. In fact, conceptually, they are different quantities.

Still, we associate a classical picture with the $d\sigma(= \Delta N/F)$ in quantum mechanics also. It is interpreted as though all the particles crossing the area $d\sigma(\theta\,\phi)$ are scattered into the solid angle $\Delta\Omega$. The quantity $d\sigma$ is interpreted as an effective area presented by the target to the incident particles to be scattered out or removed from the incident beam.

Obviously $d\sigma$, in quantum mechanics, is related to the probability of an incident particle being scattered into the solid angle $\Delta\Omega$. This can be seen as follows:

Let N particles cross an area A per second. Then the flux F is

$$F = \frac{N}{A} \text{ per second per unit area}$$

$$\therefore\ \Delta N \text{ per sec in } \Delta\Omega = \frac{N}{A}\cdot d\sigma \text{ per sec}$$

$$= N \text{ per sec}\left(\frac{d\sigma}{A}\right)$$

The quantity $d\sigma/A$ is the fraction of an area which can be taken to represent the probability of an incident particle to be scattered in the solid angle $\Delta\Omega$. This implies that though the quantity $d\sigma$ is not probability by itself, it is a measure of probability.

In the case of a single incident particle, ΔN represents the probability rate of a particle being scattered into solid angle $\Delta\Omega$. The probability rate dW is

$$dW = F d\sigma$$

We can formulate this equation in terms of the probability rate $dW_{i\to[f]}$ in time-dependent perturbation theory as follows:

Let u_i represent the initial state of an incident particle and the state u_f represents the particle in the scattered direction. Scattering into solid angle $\Delta\Omega$ means selecting directions that are within $\Delta\Omega$. So, from the point of view of time-dependent perturbation theory, the scattering process corresponds to transitions from state u_i to a group of final states $\{u_{[\Delta\Omega]}\}$. Then, $dW_{i\to[\Delta\Omega]}$ is the transition probability rate given by

$$dW_{i\to[\Delta\Omega]} = F\,d\sigma \tag{15.3}$$

15.2 SCATTERING AMPLITUDE AND DIFFERENTIAL SCATTERING CROSS SECTION

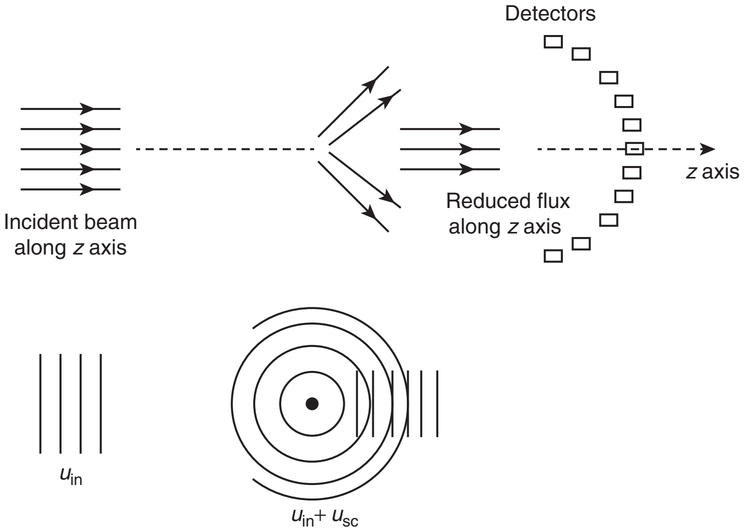


Fig.15.3 Scattering process in terms of wave mechanics

A general scattering process is shown in Fig.15.3. We have to describe this process in terms of wave mechanics as shown in Fig.15.3. We assume that the scattering process of our interest is a steady process. There are a steady incident beam and a steady scattered beam. The steady state situation means that we can use time-independent Schrödinger equation. We have already seen such processes in one-dimensional potentials. For instance, in the case of barrier penetration by a rectangular potential barrier, a single wave function, which is a solution to the Schrödinger wave equation, describes the incident beam, the reflected beam and the transmitted beam. In the same way, we would like to have a single solution for Schrödinger equation which describes the process shown in Fig.15.3.

The time-independent Schrödinger equation is

$$\left[-\frac{\hbar^2}{2m}\nabla^2 + V(\mathbf{r})\right]u(\mathbf{r}) = Eu(\mathbf{r})$$

We are looking for solutions reflecting the scattering process shown in Fig.15.3. In the vicinity of the target, we do not have enough details. But in the far-off regions, we should have an incident beam and a scattered beam. So we expect the wave function $u(\mathbf{r})$ to be composed of two parts: u_{in} and u_{sc} i.e.,

$$u(\mathbf{r}) \xrightarrow[r\to\infty]{} u_{in} + u_{sc} \tag{15.4}$$

where u_{in} describes the incident particles and u_{sc} describes the scattered particles. Let us assume that the incident particles are moving along the direction of z. When they are far away from the target, we can consider them to behave like free particles. To describe the free particles, we choose a plane wave function ψ_{in} given by

$$\psi_{\text{in}} = Ne^{i(kz-Et)/\hbar}$$

Since we are dealing with the time-independent Schrödinger equation, we omit the time-dependent part $e^{-iEt/\hbar}$ in the subsequent discussion.

$$\therefore u_{in} = Ne^{ipz/\hbar} = Ne^{ikz} \tag{15.5}$$

where N is the normalization constant. In scattering theory, it is convenient to choose N such that the flux F_{in} corresponding to u_{in} is $\frac{\hbar k}{m}$.

The flux F_{in} is given by

$$\begin{aligned} F_{in} &= \frac{\hbar}{2im}[u_{in}^{*}\nabla u_{in} - c.c.] \\ &= \frac{\hbar}{2im}|N|^2\, 2ik = |N|^2\frac{\hbar k}{m} \end{aligned}$$

$$\therefore |N|^2 = 1 \quad \text{or} \quad N = 1$$

This normalized wave function also corresponds to a number density of one particle per unit volume (the number density $n = |u_{in}|^2 = 1$). We can employ other normalization constants also (in the previous chapters, we used $N = 1/\sqrt{V}$ or $1/(2\pi\hbar)^{3/2}$) since the final result will be independent of the normalization constant.

As we said earlier, the wave function near the target depends on $V(\mathbf{r})$, and it will be known only after solving the Schrödinger equation. However, we would like to have a general form or structure for the wave function at large r for all potentials. This is possible only if we consider short-range potentials. In the vicinity of the target, the wave function can have any arbitrary shape (determined by the nature of $V(\mathbf{r})$), but at large distances, they become spherically divergent wavefronts. Therefore, the square of the modulus of wave function (or intensity) representing the scattered particles at large values of r (irrespective of the potential) should fall off like $1/r^2$. Therefore, the scattered wave should be proportional to e^{ikr}/r. So we write $u(\mathbf{r})$ as

$$u(\mathbf{r}) \longrightarrow e^{ikz} + f(\theta,\phi)\frac{e^{ikr}}{r} \tag{15.6}$$

$f(\theta,\phi)$ is known as scattering amplitude.

Obviously, the RHS is the asymptotic behaviour of the solution to the Schrödinger equation.

Where does the detail about $V(\mathbf{r})$ enter here? It is the nature of $V(\mathbf{r})$ that determines functional form of the scattering amplitude $f(\theta, \phi)$.

Let us now determine the relationship between the scattering amplitude and the differential cross section. To find such a relation, let us calculate ΔN from Schrödinger equation. For large r, the wave function u_{sc} represents the scattered particles. The flux corresponding to the scattered wave is $|u_{sc}|^2\, V$ (note that this is good approximation in the scattering processes as shown in Fig.15.3. See Example 15.1). The flux of the scattered beam $= |u_{sc}|^2\, V = \frac{|f(\theta,\phi)|^2}{r^2}\cdot\frac{\hbar k}{m}$. This expression is radial flux. ΔN is the number of particles scattered into solid angle $\Delta\Omega$ per second. $r^2\Delta\Omega$ is the area dS of a small part of the spherical surface of the sphere of radius r. This surface dS subtends an angle $\Delta\Omega$ at the centre of the sphere.

$$\begin{aligned}\therefore \Delta N &= \text{Radial flux of the scattered particles} \times \text{Area } dS\\ &= \left(\frac{|f(\theta,\phi)|^2}{r^2}\frac{\hbar k}{m}\text{ particles per second per unit area} \times dS\right)\\ &= \frac{|f(\theta,\phi)^2|}{r^2}\cdot\frac{\hbar k}{m}\cdot r^2\Delta\Omega \text{ per second}\end{aligned}$$

But $\frac{\hbar k}{m}$ is the flux of the incident particles.

$$\begin{aligned}\therefore \Delta N &= |f(\theta,\phi)|^2\, F_{in}\\ \therefore \frac{d\sigma}{d\Omega} &= |f(\theta,\phi)|^2\end{aligned} \tag{15.7}$$

So our task has become clear now. To describe a scattering process, we have to solve the Schrödinger equation

$$\left[-\frac{\hbar^2}{2m}\nabla^2 + V(\mathbf{r})\right]u(\mathbf{r}) = Eu(\mathbf{r}) \tag{15.8}$$

subject to the following boundary condition

$$u(\mathbf{r}) \xrightarrow[r\to\infty]{} e^{ikz} + f(\theta,\phi)\frac{e^{ikr}}{r} \tag{15.9}$$

The scattering problem reduces to the calculation of the scattering amplitude $f(\theta, \phi)$ for a scattering process.

The general practice is to obtain the nature of the potential from the experimental results on cross section. So one has to develop different kinds of approximations so that they are suitable to extract as much information as possible about the system from the analysis of the experimental data on the cross section.

Example 15.1 Determine the probability current density for the scattered particles in spherical coordinates represented by

$$\psi_{sc} = f(\theta,\phi)\frac{e^{ikr}}{r}$$

Solution: The flux density j_{sc} is given by

$$j_{sc} = \frac{\hbar}{2im}\left[\hat{r}\psi_{sc}^{*}\frac{\partial\psi_{sc}}{\partial r} + \hat{\theta}\psi_{sc}^{*}.\frac{1}{r}\frac{\partial}{\partial\theta}\psi_{sc} + \hat{\phi}\psi_{sc}^{*}\frac{1}{r\sin\theta}\frac{\partial}{\partial\phi}\psi_{sc} - c.c\right]$$

In spherical coordinates, we have

$$\nabla \psi_{sc} = \hat{\boldsymbol{r}}\frac{\partial}{\partial r}\left[f(\theta,\phi)\frac{e^{ikr}}{r}\right] + \hat{\boldsymbol{\theta}}\frac{1}{r}\cdot\frac{\partial}{\partial\theta}\left[f(\theta,\phi)\frac{e^{ikr}}{r}\right] + \hat{\boldsymbol{\phi}}\frac{1}{r\sin\theta}\frac{\partial}{\partial\phi}\left[f(\theta,\phi)\frac{e^{ikr}}{r}\right]$$

$$= \hat{\boldsymbol{r}}\, f(\theta,\phi) ik\,\frac{e^{ikr}}{r} - \hat{\boldsymbol{r}}\, f(\theta,\phi)\frac{e^{ikr}}{r^2} + \hat{\boldsymbol{\theta}}\frac{1}{r}\frac{\partial f}{\partial\theta}\frac{e^{ikr}}{r} + \hat{\boldsymbol{\phi}}\frac{1}{r\sin\theta}\frac{\partial f}{\partial\phi}\frac{e^{ikr}}{r}$$

$$j_{sc} = \frac{\hbar}{2im}\left\{\hat{\boldsymbol{r}}\,|f(\theta,\phi)|^2\,\frac{ik}{r^2} - \hat{\boldsymbol{r}}\frac{|f(\theta,\phi)|^2}{r^3} + \hat{\boldsymbol{\theta}}\, f^*\,\frac{\partial f}{\partial\theta}\cdot\frac{1}{r^3} + \hat{\boldsymbol{\phi}}\frac{1}{r^3\sin\theta}\, f^*\,\frac{\partial f}{\partial\phi} - c\cdot c\right\}$$

$$= \frac{\hbar}{2im}\left\{\hat{\boldsymbol{r}}\frac{|f(\theta,\phi)|^2\; 2ik}{r^2} + \hat{\boldsymbol{\theta}}\frac{1}{r^3}2Im\, f^*\,\frac{\partial f}{\partial\theta} + \hat{\boldsymbol{\phi}}\frac{1}{r^3\sin\theta}2Im\, f^*\,\frac{\partial f}{\partial\phi}\right\}$$

$$= \frac{\hbar}{2im}\left(\frac{\hat{\boldsymbol{r}}\,|\,f(\theta,\phi)\,|^2\; 2ik}{r^2}\right) + O\left(\frac{1}{r^3}\right)$$

For large r, the terms of the order $O(1/r^3)$ are smaller than the radial component.

$$\therefore\; j_{sc} \simeq \frac{|\,f(\theta,\phi)\,|^2}{r^2}\frac{\hbar k}{m}\cdot\hat{r}$$

A. Partial Waves – Basics

15.3 PHASE SHIFT ANALYSIS

Let us consider the potentials which are spherically symmetric; i.e., $V(\mathbf{r}) = V(r)$. The Hamiltonian for such potentials commutes with L^2; i.e., $[H, L^2] = 0$. Therefore, we can write the general solution as

$$u(\mathbf{r}) = \sum_{lm} R_l(r) Y_{lm}(\theta,\phi) \tag{15.10}$$

For a particle moving in the direction of z, the z component of the angular momentum is always zero; i.e., $m = 0$ in $Y_{lm}(\theta,\phi)$. This implies that Y_{lm} becomes Y_{l0}. Note that $Y_{l0} = \sqrt{2l+1/4\pi}\; P_l(\cos\theta)$. Therefore, the Equation (15.10) becomes

$$u(\mathbf{r}) = \sum_{l=0}^{\infty} a_l R_l(r) P_l(\cos\theta) \tag{15.11}$$

The potentials of our interest are short range. Let r_0 be the range of the potential. Then $V(r)$ rapidly falls to zero beyond r_0.

$$V(r) \sim 0 \;\text{ for }\; r > r_0$$

Let us solve the radial equation in the absence as well as in the presence of the potential and compare their behaviour.

Radial Equation for Free Particle

We have already studied the radial wave equation in Chapter 8. We briefly review them in the context of scattering theory. The radial equation for free particle is given by

$$\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{dR_l^{(0)}}{dr}\right)+\left[\frac{2\mu E}{\hbar^2}-\frac{l(l+1)}{r^2}\right]R_l^{(0)}(r)=0 \tag{15.12}$$

Define k^2 as $k^2=\frac{2\mu E}{\hbar^2}$, and let us write $R_l^{(0)}(r)=\frac{\chi_l^{(0)}}{r}$. Then, we get

$$\frac{d^2\chi_l^{(0)}}{dr^2}+\left[k^2-\frac{l(l+1)}{r^2}\right]\chi_l^{(0)}=0 \tag{15.13}$$

The solution to the Equation (15.12) is

$$R_l^{(0)}=b_l j_l(kr)+b_l' n_l(kr) \tag{15.14}$$

Since $n_l(kr)$ is singular at $r = 0$, $R_l^{(0)}$ is

$$R_l^{(0)}=b_l j_l(kr) \tag{15.15}$$

Let us recall the asymptotic behaviour of $j_l\,(kr)$ and $n_l(kr)$.

$$j_l(kr)\xrightarrow[r\to\infty]{}\sin\left(kr-\frac{l\pi}{2}\right)\Big/kr;$$

$$n_l(kr)\xrightarrow[r\to\infty]{}-\cos\left(kr-\frac{l\pi}{2}\right)\Big/kr$$

$$\therefore\ R_l^{(0)}\xrightarrow[r\to\infty]{}b_l\sin\left(kr-\frac{l\pi}{2}\right)\Big/kr;$$

and

$$\chi_l^{(0)}(r)\xrightarrow[r\to\infty]{}\frac{b_l}{k}\sin\left(kr-\frac{l\pi}{2}\right)$$

Radial Equation in the Presence of *V*(*r*)

The radial equation in the presence of $V(r)$ is given by

$$\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{dR_l}{dr}\right)+\left[k^2-\frac{2\mu}{\hbar^2}V(r)-\frac{l(l+1)}{r^2}\right]R_l(r)=0 \tag{15.16}$$

Or

$$\frac{d^2\chi_l}{dr^2}+\left[k^2-\frac{2\mu}{\hbar^2}V(r)-\frac{l(l+1)}{r^2}\right]\chi_l=0 \tag{15.17}$$

The exact solutions can be obtained only when $V(r)$ is known. However, the asymptotic behaviour can be studied. Since we are mainly concerned with the short-range potential, we have $V(r) \to 0$ as $r \to \infty$. In the large r limit,

$$\frac{d^2\chi_l}{dr^2}+k^2\chi_l\simeq 0 \tag{15.18}$$

or
$$\chi_l \xrightarrow[r\to\infty]{} \frac{A_l}{k}\sin\left(kr-\frac{l\pi}{2}\right)+\frac{B_l}{k}\cos\left(kr-\frac{l\pi}{2}\right) \tag{15.19}$$

The phase angles and the constants have been chosen to make the comparison with the free particle solution $\chi_l^{(0)}$. The Equation (15.19) can be rewritten as

$$\chi_l \xrightarrow[r\to\infty]{} \frac{C_l}{k}\sin\left(kr-\frac{l\pi}{2}+\delta_l\right)$$

$$=\frac{C_l}{k}\cos\delta_l\sin\left(kr-\frac{l\pi}{2}\right)+\frac{C_l}{k}\sin\delta_l\cos\left(kr-\frac{l\pi}{2}\right) \tag{15.20}$$

We also have

$$R_l \xrightarrow[r\to\infty]{} \frac{C_l}{kr}\sin\left(kr-\frac{l\pi}{2}+\delta_l\right) \tag{15.21}$$

The same results could have been obtained from the Equation (15.16) by recognizing the fact $V(r)\to 0$ for $r>r_0$ leading to the equation (see the Equation (8.75) in Section 8.13).

$$\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{dR_l}{dr}\right)+\left[k^2-\frac{l(l+1)}{r^2}\right]R_l(r)\simeq 0 \quad r>r_0 \tag{15.22}$$

or
$$R_l(r)=C_l\cos\delta_l j_l(kr)-C_l\sin\delta_l n_l(kr) \quad r>r_0 \tag{15.23}$$

The Equation (15.23) becomes the Equation (15.21) in the large r limit.

Phase Shift

So we have

$$\chi_l^{(0)} \xrightarrow[r\to\infty]{} \frac{a_l}{k}\sin\left(kr-\frac{l\pi}{2}\right) \qquad V(r)=0$$

$$\chi_l \xrightarrow[r\to\infty]{} \frac{C_l}{k}\sin\left(kr-\frac{l\pi}{2}+\delta_l\right) \qquad V(r)\neq 0$$

In the asymptotic limit, the phase of the function $\chi_l^{(0)}$ is $(kr-l\pi/2)$ and the phase of χ_l for a given potential $V(r)$ is $(kr-l\pi/2+\delta_l)$. So the effect of potential $V(r)$ is to change the phase of the asymptotic solution from $(kr-l\pi/2)$ to $(kr-l\pi/2+\delta_l)$. Therefore, δ_l is called phase shift.

15.4 SCATTERING AMPLITUDE AND PHASE SHIFT

Let us now determine the scattering amplitude in terms of the phase shift. Let us recall the boundary condition for the wave function in the scattering theory is

$$u(r,\theta,\phi)\xrightarrow[r\to\infty]{} e^{ikz}+f(\theta,\phi)\frac{e^{ikr}}{r} \tag{15.24}$$

We also know that from the Equation (15.11) and the asymptotic form (15.21) of the radial wave function,

$$u(r,\theta,\phi)\xrightarrow[r\to\infty]{}\sum_{l=0}^{\infty}\frac{C_l}{kr}\sin\left(kr-\frac{l\pi}{2}+\delta_l\right)P_l\cos\theta \tag{15.25}$$

Obviously the Equation (15.25) should be brought to the form of the Equation (15.24) to determine the scattering amplitude $f(\theta,\phi)$.

Let us first concentrate on Equation (15.24). The plane wave function e^{ikz} can now be expressed as follows:

$$e^{ikz}=e^{ikr\cos\theta}=\sum_{l=0}^{\infty}(2l+1)\,i^l j_l(kr)P_l(\cos\theta) \tag{15.26}$$

This is true for all values of r.

In the asymptotic limit, this equation becomes

$$\begin{aligned} e^{ikz} &\longrightarrow \sum_{l=0}^{\infty}(2l+1)i^l\sin\left(kr-\frac{l\pi}{2}\right)\cdot P_l(\cos\theta)\\ &=-\sum_{l=0}^{\infty}\frac{(2l+1)i^l}{2ikr}\left[e^{-i\left(kr-\frac{l\pi}{2}\right)}-e^{i\left(kr-\frac{l\pi}{2}\right)}\right]P_l(\cos\theta)\end{aligned} \tag{15.27}$$

This is a sum of infinite number of incoming spherical waves (e^{-ikr}/r term) and outgoing spherical waves (e^{ikr}/r term). Making use of the Equation (15.27) in (15.24), we get

$$u(\mathbf{r})\xrightarrow[r\to\infty]{}\sum_{l=0}^{\infty}\frac{(-1)(2l+1)i^l}{2ikr}\left[e^{-i\left(kr-\frac{l\pi}{2}\right)}-e^{i\left(kr-\frac{l\pi}{2}\right)}\right]+\frac{e^{ikr}}{r}f(\theta,\phi) \tag{15.28}$$

It is obvious that the coefficients of e^{-ikr}/r are the same in both (15.27) and (15.28). The coefficients of e^{-ikr}/r are different for Equations (15.27) and (15.28). The simplest way to take into account these facts is to write $u(\mathbf{r})$ as follows:

$$u(\mathbf{r})\to\sum_{l=0}^{\infty}\frac{(-1)(2l+1)}{2ikr}i^l\left[e^{-i\left(kr-\frac{l\pi}{2}\right)}-S_l(k)e^{i\left(kr-\frac{l\pi}{2}\right)}\right]P_l(\cos\theta) \tag{15.29}$$

This equation can be rewritten as

$$\begin{aligned} u(\mathbf{r}) &=\sum_{l=0}^{\infty}\frac{(-1)(2l+1)i^l}{2ikr}\left[e^{-i\left(kr-\frac{l\pi}{2}\right)}-e^{i\left(kr-\frac{l\pi}{2}\right)}-(S_l(k)-1)e^{i\left(kr-\frac{l\pi}{2}\right)}\right]P_l(\cos\theta)\\ &=e^{ikz}+\sum_{l=r}^{\infty}\frac{(2l+1)i^l}{2ikr}(S_l(k)-1)e^{i\left(kr-\frac{l\pi}{2}\right)}P_l(\cos\theta)\end{aligned} \tag{15.30}$$

Since, $i^l = e^{il\pi/2}$, we get

$$\begin{aligned} u(\mathbf{r}) &= e^{ikz} + \sum_{l=0}^{\infty} \frac{(2l+1)}{2ikr}(S_l(k)-1)P_l(\cos\theta)e^{ikr} \\ &= e^{ikz} + f(\theta,\phi)\frac{e^{ikr}}{r} \end{aligned} \tag{15.31}$$

Provided, we have

$$f(\theta,\phi) = \sum_{l=0}^{\infty}(2l+1)f_l(k)P_l(\cos\theta) \tag{15.32}$$

where
$$f_l(k) = \frac{S_l(k)-1}{2ik} \tag{15.33}$$

What have we done so far? We have transformed the equation for $f(\theta,\phi)$ (Equation (15.24)) into an equation for $f_l(k)$ or $S_l(k)$. So, the determination of $S_l(k)$ amounts to the determination of $f(\theta,\phi)$. We can determine $S_l(k)$ by comparing the Equations (15.25) and (15.29).

From the Equation (15.25), we have

$$\begin{aligned} u(\mathbf{r}) &\xrightarrow[r\to\infty]{} \sum_{l=0}^{\infty} \frac{C_l}{kr}\sin\left(kr - \frac{l\pi}{2} + \delta_l\right)P_l(\cos\theta) \\ &= \sum_{l=0}^{\infty} -\frac{C_l}{2ikr}\left[e^{-i\left(kr-\frac{l\pi}{2}+\delta_l\right)} - e^{i\left(kr-\frac{l\pi}{2}+\delta_l\right)}\right]P_l(\cos\theta) \\ &= \sum_{l=r}^{\infty} \frac{-C_l e^{-i\delta_l}}{2ikr}\left[e^{-i\left(kr-\frac{l\pi}{2}\right)} - e^{2i\delta_l}\, e^{i\left(kr-\frac{l\pi}{2}\right)}\right]P_l(\cos\theta) \end{aligned} \tag{15.34}$$

Comparing the Equations (15.34) and (15.29), we get

$$C_l e^{-i\delta_l} = (2l+1)i^l$$

and

$$S_l(k) = e^{2i\delta_l} \tag{15.35}$$

Therefore, we have (from the Equations (15.32) and (15.33))

$$f(\theta,\phi) = \sum_{l=0}^{\infty}(2l+1)\frac{(e^{2i\delta_l}-1)}{2ik}P_l(\cos\theta) \tag{15.36}$$

$$= \sum_{l=0}^{\infty}(2l+1)\, f_l(k)P_l(\cos\theta) \tag{15.37}$$

where
$$f_l(k) = \frac{e^{2i\delta_l}-1}{2ik}$$

The scattering amplitude $f(\theta, \phi)$ can be written as

$$\begin{aligned} f(\theta,\phi) &= \sum_{l=1}^{\infty} \frac{(2l+1)e^{i\delta_l}(e^{i\delta_l} - e^{-i\delta_l})}{2ik}P_l(\cos\theta) \\ &= \sum_{l=0}^{\infty} \frac{(2l+1)\sin\delta_l e^{i\delta_l}P_l(\cos\theta)}{k} \end{aligned}$$

Before determining the total cross section, let us complete this discussion by writing asymptotic form of the wave function $R_l(r)$ using the Equations (15.23) and (15.35).

$$R_l \to (2l+1)i^l\, e^{i\delta_l}[\cos\delta_l j_l(kr) - \sin\delta_l n_l(kr)] \text{ for } r >> r_0 \tag{15.38}$$

This is a very useful result for later calculations. Note this is true for any short-range potential.

Total Cross Section

The total cross section σ is given by

$$\begin{aligned}\sigma &= \int |f(\theta,\phi)|^2 d\Omega \\ &= \frac{1}{k^2}\sum_{l=0}\sum_{l'=0}(2l+1)(2l'+1)\sin\delta_l \sin\delta_{l'} e^{i\delta_l}\, e^{-i\delta_{l'}} \int P_l(\cos\theta)P_{l'}(\cos\theta)d\Omega \\ &= \frac{1}{k^2}\sum_{l=0}^{\infty}\sum_{l'=0}^{\infty}(2l+1)(2l'+1)\sin\delta_l \sin\delta_{l'} e^{i\delta_l}\, e^{-i\delta_{l'}} \cdot \frac{4\pi}{(2l+1)}\delta_{ll'} \\ &= \frac{4\pi}{k^2}\sum_{l=0}^{\infty}(2l+1)\sin^2\delta_l\end{aligned} \tag{15.39}$$

In terms of $f_l(k)$, we can write σ as

$$\sigma = \sum_{l=0}^{\infty}(2l+1)|f_l(k)|^2 \tag{15.40}$$

The structure of the scattering amplitude as well as the total cross section σ is $\sum_{l=0}^{\infty}$. Obviously, they are sum of the contribution from each of l. This is the reason why this analysis is called partial wave analysis.

Optical Theorem

It is interesting to determine the forward scattering amplitude corresponding to $\theta = 0$.

Note $\qquad P_l(\cos\theta) = P_l(1) = 1$

$$\therefore \quad f(0) = \sum_{l=0}^{\infty}\frac{1}{k}(2l+1)\sin\delta_l(\cos\delta_l + i\sin\delta_l)$$

$$\frac{4\pi}{k}\operatorname{Im} f(0) = \sum_{l=0}^{\infty}\frac{4\pi}{k}\cdot\frac{1}{k}(2l+1)\sin^2\delta_l$$

But the RHS is the total cross section σ

$$\therefore \sigma = \frac{4\pi}{k}\operatorname{Im} f(0).$$

The form of the Schrödinger wave function in scattering theory $u_{in} + u_{sc} = e^{ikz} + u_{sc}$. There can be interference between incident wave and scattered wave only in the forward direction behind the target. This interference corresponds to the reduction in flux along z axis behind the target. This reduction should be the same as the number of particles crossing σ ($F\sigma$ is the total number of particles removed from the incident beam by the scattering process). So there should be a relationship between σ and forward scattering amplitude. Optical theorem reflects this fact.

In Appendix I, this result is brought but more clearly.

Example 15.2 For elastic scattering, show that

$$f_l(k) = \frac{e^{2i\delta_l} - 1}{2ik} = \frac{1}{k\cot\delta_l - ik}$$

Solution:

$$\begin{aligned} f_l(k) &= \frac{e^{2i\delta_l} - 1}{2ik} = \frac{\cos 2\delta_l + i\sin 2\delta_l - 1}{2ik} \\ &= \frac{1 - 2\sin^2\delta_l + 2i\sin\delta_l\cos\delta_l - 1}{2ik} \\ &= -\frac{2\sin^2\delta_l(1 - i\cot\delta_l)}{2ik} \\ &= \frac{1}{-ik}\,\frac{1}{\operatorname{cosec}^2\delta_l}(1 - i\cot\delta_l) \\ &= \frac{1}{-ik}\cdot\frac{(1 - i\cot\delta_l)}{(1 + i\cot\delta_l)(1 - i\cot\delta_l)} \\ &= \frac{1}{k\cot\delta_l - ik} \end{aligned} \tag{15.41}$$

This expression will be useful when we discuss the concept of scattering length.

15.5 NUMBER OF PARTIAL WAVES NEEDED

The expression for the scattering amplitude and the cross section as given in Section 15.4 is summation over infinite number of partial waves. If one has to evaluate all these terms, then the partial wave analysis will not be useful. One would like to restrict to few partial waves. In particular, we would like to describe situations where $l = 0$ partial wave is sufficient. A partial wave with $l = 0$ is called *s* wave. We are interested in determining the conditions under which the *s* wave scattering is enough.

We use a semi-classical argument to determine $l_{\max}$, the maximum value of l for which the partial waves are necessarily to be considered. Equivalently, we determine the conditions for the partial waves with $l \geq l_{\max}$ to be ignored.

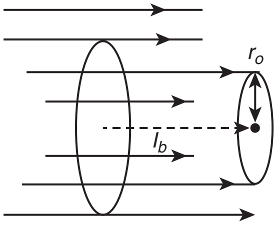


Fig.15.4 $L > pr_0$ are not scattered by a hard disk

Consider the classical scattering where the target is a hard disk radius r_0 as shown in Fig. 15.4. Let A be the area of the incident beam of particles. A particle with an impact parameter b has an angular momentum $L = pb$. There is a maximum value of b equal to r_0 beyond which the incident particles will not be scattered. In terms of angular momentum, particles with $L > L_{\max}$ will not be scattered where $L_{\max} = pr_0$. Particles with $L > pr_0$ can be ignored as far as the scattering by the target is concerned.

This result can be taken over to quantum mechanics where the quantity r_0 is now the range of the potential *V*(*r*). Taking the maximum angular momentum to be $l_{\max}\hbar$ and the momentum to be $p = \hbar k$, we have

$$l_{\max}\hbar = \hbar k r_0$$

or $l_{\max} = kr_0 \therefore l \geq kr_0 \Rightarrow$ Partial waves with such *l* to be ignored.

Obviously for considering *s* wave alone, partial waves with even $l = 1$ have to be ignored.

$$\therefore\ 1 >> kr_0$$

Therefore, the condition for low-energy scattering (or for restricting our consideration to *s* wave scattering only) is

or
$$kr_0 << 1 \tag{15.42}$$

If the above condition is satisfied, it is enough to consider *s* waves alone. Such scattering is called low-energy scattering.

Example 15.3 Show that the condition for low-energy scattering is $kr_0 << 1$ by considering the centrifugal potential barrier.

Solution: The effective potential is

$$V_{eff}(r) = V(r) + \frac{l(l+1)}{r^2} \cdot \frac{\hbar^2}{2\mu}$$

The general shape of V_{eff} is shown in Fig.15.5, which is a potential barrier. The barrier height increases as the orbital angular momentum quantum number *l* increases. For a given energy *E*, depending on the value of *l*, the barrier height of V_{eff} may be less than or greater than *E*. Let $l_{\max}$ be the maximum value for which the barrier height of V_{eff} is less than *E*. This means that for $l > l_{\max}$, the effective potential V_{eff} is greater than *E* for some range of *r* as shown in Fig.15.5.

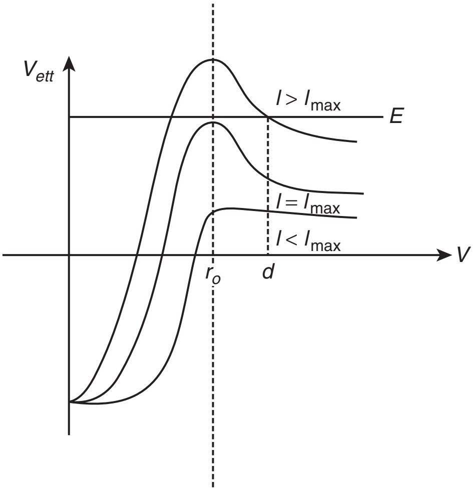


Fig. 15.5 Effective potential. Here r_0 is effective range and *d* is the shortest distance

A particle with energy E has to penetrate the potential barrier $\frac{l(l+1)}{r^2}\cdot\frac{\hbar^2}{2\mu}$ to reach r_0. Then only it can experience the potential $V(r)$. For $l > l_{\max}$, the shortest distance, which the incident particle can reach, is d, which is determined by

$$\frac{l(l+1)\hbar^2}{2\mu\, d^2} = \frac{\hbar^2 k^2}{2\mu}$$

or

$$\frac{l^2}{d^2} \simeq k^2 \text{ or } l \simeq kd$$

For the particle to reach r_0, the shortest distance d should be the same as r_0. Therefore, for $l_{\max}$, we have $l_{\max} \simeq kr_0$.

Thus, if we have $l >> kr_0$, such partial waves need not be considered. Therefore, if $kr_0 << 1$, only s wave need to be considered.

Example 15.4 Assuming the range of interaction between neutron and proton to be $r_0 = 2\,fm$, estimate the energy range in the centre of mass frame for low-energy scattering.

Solution: For low-energy scattering $kr_0 << 1$

$$\left(\frac{2\mu E}{\hbar^2}\right)^{1/2} r_0 << 1 \quad \text{or} \quad E << \frac{\hbar^2}{2\mu r_0^{\,2}}$$

The reduced mass μ is

$$\frac{1}{\mu} = \frac{1}{m_p} + \frac{1}{m_n} = \left(\frac{1}{1.67265} + \frac{1}{1.67492}\right)\times 10^{27}\,kg^{-1}$$

$$\therefore E << \frac{(1.05459\times10^{-34})^2 \times 1.1949\times10^{27}}{2\times(2\times10^{-15})^2\times(1.60219)\times10^{-13}}\,MeV$$

or

$$E << 10 \text{ MeV}$$

15.6 INTEGRAL FORMULA FOR PHASE SHIFT δ_l

Let us get back to the radial equation for particle in the presence of $V(r)$ and in the absence of $V(r)$.

$$\frac{d^2\chi_l}{dr^2} + \left[k^2 - \frac{2\mu}{\hbar^2}V(r) - \frac{l(l+1)}{r^2}\right]\chi_l = 0 \qquad V(r)\neq 0 \tag{15.43}$$

$$\frac{d^2\chi_l^{(0)}}{dr^2} + \left[k^2 - \frac{l(l+1)}{r^2}\right]\chi_l^{(0)} = 0 \qquad V(r) = 0 \tag{15.44}$$

$$\chi_l^{(0)} \times (15.43) - \chi_l \times (15.44)$$

$$\chi_l^{(0)}\frac{d^2\chi_l}{dr^2} - \chi_l\frac{d^2\chi_l^{(0)}}{dr^2} = \frac{2\mu}{\hbar^2}V(r)\chi_l(r)\chi_l^{(0)}(r)$$

$$\frac{d}{dr}\left[\chi_l^{(0)}\frac{d\chi_l}{dr} - \frac{d\chi_l^{(0)}}{dr}\chi_l\right] = \frac{2\mu}{\hbar^2}\chi_l^{(0)}(r)\chi_l(r)V(r)$$

Integrating from 0 to r,

$$\chi_l^{(0)}\frac{d\chi_l}{dr}-\frac{d\chi_l^{(0)}}{dr}\chi_l(r)=\frac{2\mu}{\hbar^2}\int_0^r\chi_l^{(0)}(r')\chi_l(r')V(r')dr' \tag{15.45}$$

Let us make r very large.

For large r,

$$\chi_l^{(0)}=\frac{b_l}{k}\sin\left(kr-\frac{l\pi}{2}\right);\quad \frac{d\chi_l^{(0)}}{dr}=b_l\cos\left(kr-\frac{l\pi}{2}\right)$$

$$\chi_l=\frac{C_l}{k}\sin\left(kr-\frac{l\pi}{2}+\delta_l\right);\quad \frac{d\chi_l}{dr}=C_l\cos\left(kr-\frac{l\pi}{2}+\delta_l\right)$$

Since $\chi_l^{(0)}$ and χ_l appear in both LHS and RHS, the normalization constants will cancel, and so we can drop both b_1 and C_l.

In the large r limit, the LHS of the Equation (15.45) becomes

$$\text{LHS}=\frac{1}{k}\sin\left(kr-\frac{l\pi}{2}\right)\cos\left(kr-\frac{l\pi}{2}+\delta_l\right)-\sin\left(kr-\frac{l\pi}{2}+\delta_l\right)\cos\left(kr-\frac{l\pi}{2}\right)$$

$$=\frac{-\sin\delta_l}{k}$$

$$\text{RHS}=\frac{2\mu}{\hbar^2}\int_0^\infty r'j_l(kr')\chi_l(kr')V(r')dr'$$

We have made use of the fact $\chi_l^{(0)}=rR_l^{(0)}(r)$.

$$\therefore\frac{-\sin\delta_l}{k}=\frac{2\mu}{\hbar^2}\int_0^\infty r'j_l(kr')\chi_l(kr')V(r')dr'$$

$$\sin\delta_l=\frac{-2\mu}{\hbar^2}\int_o^\infty kr'j_l(kr')\chi_l(kr')V(r')dr' \tag{15.46}$$

Let as make a further approximation in which $\chi_l(kr)$ is replaced by $\chi_l^{(0)}(kr)=rj_l(kr)$ in the integral in (15.46).

$$\therefore\sin\delta_l=\frac{-2\mu}{\hbar^2}\int_o^\infty kr^2[j_l(kr)]^2V(r)dr \tag{15.47}$$

Sign of Phase Shift

The sign of the phase shift depends on the nature of interaction namely, whether it is attractive or repulsive. The above equation can be used to determine the sign of the phase shift. In the integrand of (15.47), the factor $[rj_l(kr)]^2$ is always positive. For an attractive interaction, $V(r)$ is negative, which makes the RHS positive. Therefore, δ_l is a positive number. In the same way, we can argue that for a repulsive interaction, δ_l is negative.

Positive δ_l → attractive interaction
Negative δ_l → repulsive interaction

15.7 EXPRESSION FOR δ_l USING LOGARITHMIC DERIVATIVE

We can get another general expression for phase shift δ_l for short-range potential. This time, we use the continuity of the function and its derivative. The potential $V(r)$ is written as

$$V(r) \simeq \begin{cases} V(r) & r < a \\ 0 & r > a \end{cases}$$

$V(r)$ may be a square well potential or any other general short-range potential.

Then we have two solutions: one solution inside the range a and the other outside the range a.

$$\frac{d^2\chi_l}{dr^2} + \left[k^2 - \frac{2\mu}{\hbar^2}V(r) - \frac{l(l+1)}{r^2}\right]\chi_l = 0 \quad r < a \tag{15.48}$$

$$\frac{d^2\chi_l^{(0)}}{dr^2} + \left[k^2 - \frac{l(l+1)}{r^2}\right]\chi_l^{(0)} = 0 \quad r > a \tag{15.49}$$

Inside the region $r < a$, the wave function can be determined only if $V(r)$ is known. In the region $r > a$, we know that the asymptotic solution $\chi_l^{(0)}$ is given by (see (15.23))

$$\begin{aligned} \chi_l^{(o)} &= r\,R_l^{(o)} \\ &= C_l r[\cos\delta_l\, j_l(kr) - \sin\delta_l n_l(kr)] \end{aligned} \tag{15.50}$$

$$\frac{d\chi_l^{(0)}}{dr} = C_l[\cos\delta_l\ j_l - \sin\delta_l\ n_l] + C_l rk[\cos\delta_l\ j_l' - \sin\delta_l\ n_l'] \tag{15.51}$$

where $$j_l' = \frac{dj_l(\rho)}{d\rho} \text{ and } \chi_l' = \frac{d\chi_l(\rho)}{d\rho}$$

Since we have two solutions, we can match their logarithmic derivatives. First, let us evaluate the logarithmic derivative for the solution corresponding to the region outside the range a.

$$\begin{aligned} \left.\frac{1}{\chi_l^{(0)}}\frac{d\chi_l^{(0)}}{dr}\right|_{r=a} &= \frac{[j_l(ka) + ka\ j_l'(ka)]\cos\delta_l - [n_l(ka) + ka\ n_l'(ka)]\sin\delta_l}{j_l(ka)\cos\delta_l - n_l(ka)\sin\delta_l} \\ &= \frac{j_l(ka) + ka\ j_l'(ka) - [n_l(ka) + ka\ n_l'(ka)]\tan\delta_l}{j_l(ka) - n_l(ka)\tan\delta_l} \end{aligned} \tag{15.52}$$

Since we do not know $V(r)$, we cannot evaluate the logarithmic derivative for the wave function inside $r < a$. However, let us define $\lambda_l a$ as

$$\frac{1}{\chi_l}\cdot\frac{d\chi_l}{dr} \xrightarrow{r\to a} \frac{1}{\lambda_l a} \tag{15.53}$$

Then we have,

$$\frac{1}{\chi_l^{(0)}}\cdot\frac{d\chi_l^{(0)}}{dr}\bigg|_{r=a}=\frac{1}{\lambda_l a} \tag{15.54}$$

$$\therefore \frac{1}{\lambda_l a}=\frac{(j_l+ka\ j_l')-(n_l+ka\ n_l')\tan\delta_l}{j_l-n_l\tan\delta_l}$$

Rearranging the terms, we get

$$\begin{aligned}\tan\delta_l&=\frac{\lambda_l ka\ j_l'(ka)-(1-\lambda_l)j_l(ka)}{\lambda_l ka\ n_l'(ka)-(1-\lambda_l)n_l(ka)}\\&=\frac{\dfrac{\lambda_l}{(1-\lambda_l)}ka\ j_l'(ka)-j_l(ka)}{\dfrac{\lambda_l}{(1-\lambda_l)}ka\ n_l'(ka)-n_l(ka)}\\&=\frac{\gamma_l\ ka\ j_l'(ka)-j_l(ka)}{\gamma_l\ ka\ n_l'(ka)-n_l(ka)}\end{aligned} \tag{15.55}$$

Where $\gamma_l=\dfrac{\lambda_l}{1-\lambda_l}$

Note that in the absence of knowledge of potential inside the range a, the quantity λ_l or γ_l is just an unknown factor in the Equations (15.53)–(15.55).

Example 15.5 Using the expression (15.55) for the phase shift δ_0, show that

$$\lim_{k\to 0} k\cot\delta_0=-\frac{1}{\alpha}+\frac{1}{2}r_{eff}k^2$$

Solution: For small values of k, assume that γ_0 can be written as

$$\gamma_0(k)=\gamma_0(0)+\varepsilon k^2$$

From (15.55), we have

$$\cot\delta_0=\frac{\gamma_0 ka\,n_0'(ka)-n_0(ka)}{\gamma_0 kaj_0'(ka)-j_0(ka)}$$

The functions j_0 and n_0 and their derivatives are given by

$$j_0(\rho)=\frac{\sin\rho}{\rho}\qquad j_0'(\rho)=\frac{\cos\rho}{\rho}-\frac{\sin\rho}{\rho^2}$$

$$n_0(\rho)=-\frac{\cos\rho}{\rho}\qquad n_0'(\rho)=\frac{\sin\rho}{\rho}+\frac{\cos\rho}{\rho^2}$$

where $\rho=ka$

Substituting these expressions in the Equation (15.55), we get

$$\cot\delta_0 = \frac{\gamma_0\rho\left(\dfrac{\sin\rho}{\rho}+\dfrac{\cos\rho}{\rho^2}\right)+\dfrac{\cos\rho}{\rho}}{\gamma_0\rho\left(\dfrac{\cos\rho}{\rho}-\dfrac{\sin\rho}{\rho^2}\right)-\dfrac{\sin\rho}{\rho}}$$

$$\rho\cot\delta_0 = \frac{\gamma_0(\rho\sin\rho+\cos\rho)+\cos\rho}{\gamma_0\left(\cos\rho-\dfrac{\sin\rho}{\rho}\right)-\dfrac{\sin\rho}{\rho}}$$

In the limit $\rho \to 0$, restricting ourselves to ρ^2 terms

$$\lim_{\rho\to 0}\rho\cot\delta_0 = \frac{\gamma_0\left(\rho^2+1-\dfrac{\rho^2}{2}\right)+\left(1-\dfrac{\rho^2}{2}\right)}{\gamma_0\left(1-\dfrac{\rho^2}{2}-1+\dfrac{\rho^2}{6}\right)-\left(1-\dfrac{\rho^2}{6}\right)}$$

$$= -(\gamma_0+1)+\frac{\rho^2}{3}(1-\gamma_0+\gamma_0{}^2)$$

Since we have assumed $\gamma_0(\rho)=\gamma_0(0)+\varepsilon\rho^2$, we get

$$\lim_{k\to 0} ka\cot\delta_0 = -(\gamma_0(0)+1)+\frac{k^2a^2}{3}(1-\gamma_0(0)+\gamma_0(0)^2+3\varepsilon)$$

$$\lim_{k\to 0} k\cot\delta_0 = -\underbrace{\frac{(\gamma_0(0)+1)}{a}}_{\frac{1}{\alpha}}+k^2\underbrace{\frac{a}{3}(1-\gamma_0(0)+\gamma_0(0)^2+3\varepsilon)}_{\frac{1}{2}r_{eff}}$$

We can write this equation as

$$\lim_{k\to 0} k\cot\delta_0 = -\frac{1}{\alpha}+\frac{1}{2}r_{eff}k^2$$

The above expression was computed for an arbitrary short-range potential of range r_0. The quantity γ_0 depends on the potential $V(r)$. The reason for the term 'the effective range' for the coefficient of k^2 will become clear if we compare the expression with the corresponding expression for square well potential of range a. See Example 15.8.

15.8 SCATTERING BY HARD SPHERE

This is one of the few problems, which is exactly solvable. The target is a hard sphere of radius R. This means the particle cannot be found inside the sphere. So the potential of hard sphere is

$$V(r)=\begin{cases}\infty & r\le R\\ 0 & r\ge R\end{cases}$$

Since the potential inside R is ∞, the incident particles cannot penetrate. Therefore, the wave function $R_l(r)$ is zero inside R, and it is a free particles solution outside R.

Therefore, making use of (15.23), we can write

$$R_l(r) = \begin{cases} 0 & r \le R \\ C_l[\cos\delta_l j_l(kr) - \sin\delta_l n_l(kr)] & r \ge R \end{cases}$$

From (15.35), we know that the constant C_l is given by

$$C_l = (2l+1)i^l e^{i\delta_l}$$

Let us restrict ourselves to $l = 0$ case. The solution R_0 is given by

$$R_0 = \begin{cases} 0 & r \le R \\ C_0[\cos\delta_0 j_0(kr) - \sin\delta_0 n_0(kr)] & r \ge R \end{cases}$$

So for $r \ge R$, we have

$$\begin{aligned} R_0 &= e^{i\delta_0}\left[\cos\delta_0 \frac{\sin kr}{kr} + \sin\delta_0 \frac{\cos kr}{kr}\right] \\ &= \frac{e^{i\delta_0}}{kr}\sin(kr+\delta_0) \end{aligned}$$

Matching the $R_0(r)$ at $r = R$, we get

$$R_0(r) = 0 \text{ at } r = R$$

$$kR + \delta_0 = 0 \quad \text{or} \quad \delta_0 = -kR$$

$$\therefore R_0 = e^{-ikR}\sin k\,(r - R)$$

In the low-energy limit, we assume $kR << 1$, and we restrict ourselves to $l = 0$ partial wave. Then the total cross section

$$\begin{aligned} \sigma &= \frac{4\pi}{R^2}\sin^2\delta_0 \simeq \frac{4\pi}{k^2}\cdot k^2R^2 \\ &= 4\pi R^2 \end{aligned} \tag{15.56}$$

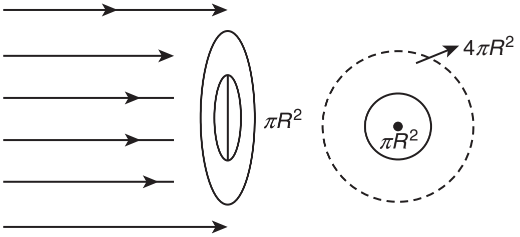


Fig.15.6 Scattering by hard disc

This is much greater than the geometrical cross section πR^2 as shown in Fig. 15.6. So the target presents a large area $4\pi R^2$ to remove all the incident particles which cross this area. It is to be noted that the phase shift δ_0 is negative, confirming the fact that a repulsive interaction produces a negative phase shift.

The phase shift generally depends on the energy of the incident particles. For very low energies, the scattering cross section is independent of the energy.

B. Low-energy Scattering – General Short-range Potential

15.9 PHASE SHIFT IN THE LOW-ENERGY APPROXIMATION

For a potential of range r_0, low energy means the energy range is such that $kr_0 << 1$. Consider the equation for δ_l given in (15.47).

For small values of kr, the function $j_l(kr)$ obeys the following relation

$$j_l(kr) \simeq \frac{(kr)^l}{1.3.5(2l+1)} \qquad kr << kr_0 << 1 \tag{15.57}$$

Since the range of the potential is r_0, we can restrict the upper limit of the integral to r_0.

i.e.,

$$\int_0^\infty \to \int_0^{r_0}$$

$$\therefore \sin\delta_l = -\frac{2\mu}{\hbar^2}\int_0^{r_o} \frac{k^{2l} r^{2l}}{[1.3.5...(2l+1)^2]} k \cdot r^2 V(r)dr$$

$$= -\frac{2\mu}{\hbar^2}\cdot\frac{(kr_0)^{2l+1}}{[1.3.5...(2l+1)]^2}\int_0^{r_0} V(r)\left(\frac{r}{r_o}\right)^{2l+1} rdr \tag{15.58}$$

A similar result can be obtained from the Equation (15.55). Let us note

$$n_l(kr) = -[1.3.5\ldots(21-1)]\frac{1}{(kr)^{l+1}} \tag{15.59}$$

Making use of (15.57) and (15.59) in (15.55), we get

$$\tan\delta_l \sim \frac{(kr_0)^{2l+1}\cdot(2l+1)}{[1.3.5....(2l+1)]^2}\cdot\left\{\frac{(l+1)\lambda_l - 1}{l\lambda_l + 1}\right\} \tag{15.60}$$

15.10 SCATTERING LENGTH IN LOW-ENERGY SCATTERING

Two new concepts called scattering length and effective range are introduced in low-energy scattering. We concentrate here only on the concept of scattering length.

Let us recall that

$$\sigma = \sum_{l=0}^{\infty} 4\pi\,|f_l(k)|^2$$

Since we are dealing with low-energy scattering, it is enough to restrict $l = 0$ partial wave alone. In that case,

$$\sigma \simeq \sigma_0 = 4\pi|f_0(k)|^2 \tag{15.61}$$

From the Equation (15.41) (see Example 15.2), We have

$$f_0(k) = \frac{1}{k\cot\delta_0 - ik} \tag{15.62}$$

Scattering length α is defined as

$$\underset{k\to 0}{\text{Lt}}\, k\cot\delta_0 = -\frac{1}{\alpha} \tag{15.63}$$

This can be improved by next approximation

$$k\cot\delta_0 = -\frac{1}{\alpha} + \frac{1}{2} r_{\text{eff}} k^2$$

This relation has already been proved in Example 15.5. Let us concentrate on the Equations (15.62) and (15.63),

$$\begin{aligned} \underset{k\to 0}{\text{Lt}}\, f_0(k) &= \frac{1}{-1/\alpha} = -\alpha \\ \therefore \sigma_0 &= 4\pi\alpha^2 \end{aligned} \tag{15.64}$$

To understand the significance of the scattering length, we have to compare the above expression with the cross section for the scattering by hard sphere of radius R.

The scattering cross section for the hard sphere is

$$\sigma = 4\pi R^2$$

It is at $r = R$, the wave function $R_0(r)$ or $\chi_0(r)$ is zero. Can we interpret α as the distance at which the wave function $R_0(r)$ or $\chi_0(r)$ vanishes for the general potential $V(r)$ also?

To answer this question, let us consider the asymptotic behaviour of $\chi_0(r)$ in the context of low-energy scattering.

The radial equation for $l = 0$ wave is

$$\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{dR_0}{dr}\right) + \left(k^2 - \frac{2\mu}{\hbar^2}V(r)\right)R_0 = 0 \tag{15.65}$$

or

$$\frac{d^2\chi_0}{dr^2} + \left(k^2 - \frac{2\mu}{\hbar^2}V(r)\right)\chi_0 = 0 \tag{15.66}$$

Here,

$$R_0(r) = \frac{\chi_0(r)}{r}$$

For large r, making use of (15.20), we have

$$\chi_0(r) \xrightarrow{\text{large } r} \frac{C_l}{k}[\cos\delta_0 \sin kr + \sin\delta_0 \cos kr] \qquad kr_0 << kr << 1$$

Since $kr << 1$, we have

$$\chi_0(r) \xrightarrow[\text{large } r]{} \frac{C_l}{k}[kr\cos\delta_0 + \sin\delta_0]$$

$$= \frac{C_l}{k}\cos\delta_0[kr + \tan\delta_0] \tag{15.67}$$

$\chi_0(r)$ can become zero at $r = \alpha$ if

$$k\alpha + \tan\,\delta_0 = 0$$

or
$$k\alpha = -\tan\delta_0$$

or
$$k\cot\,\delta_0 = -1/\alpha$$

So $\chi_0(r)$ becomes zero at $r = \alpha$ if $k\cot\delta_0 = -\dfrac{1}{\alpha}$

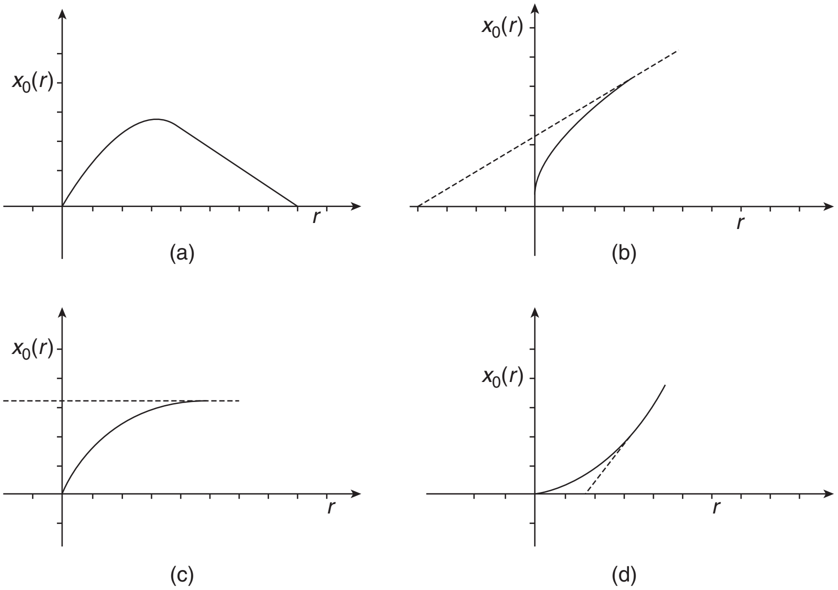
However, one has to be careful in interpreting the Equation (15.67). It is an equation of limiting behaviour of $\chi_0(r)$ in the large r limit for very small values of k. Note that we did not determine the exact solution of (15.66) to get (15.67) as the limiting case. In fact, it is a solution to the Equation (15.66) (after dropping the terms k^2 and $2\mu V(r)/\hbar^2$)

$$\frac{d^2\chi_0}{dr^2} \simeq 0 \tag{15.68}$$

So it is an equation for the asymptote. (15.67) is the equation in r for a straight line. The r intercept of this straight line is

$$r = \alpha = -\frac{\tan\delta_0}{k}$$

The various possibilities for the wave function and their asymptotes are shown in Fig.15.7.



(a) (b)

(c) (d)

Fig.15.7 Scattering amplitudes-different possibilities

The Equation (15.66) can be used to interpret the conditions under which these possibilities arise. The behaviour of χ_0 in Fig.15.7 (a) is possible if the interaction is attractive enough to form a bound state. We can substantiate this conclusion as follows. We have to bear in mind that a bound-state solution goes like e^{-kr} with negative slope for large r. Let us come back to the Equation (15.66). In the case of an attractive potential, $V(r)$ is negative, which makes the solution χ_0 of the Equation (15.66) to concave downwards. The slope of $\chi_0(r)$ is negative, and so it can match a bound-state solution like e^{-kr}. So if the attractive potential is strong enough, a bound state like Fig.15.7 (a) exists.

In Fig.15.7 (b) also, $\chi_0(r)$ concaves down, and so this case also corresponds to attractive potential. However, the attractive potential is not strong enough to form a bound state. Fig.15.7(c) represents a resonant scattering in which the scattering length becomes very large. In Fig.15.7 (d), $\chi_0\,(r)$ concaves upwards, and so the potential $V(r)$ has to be positive, which represents a repulsive potential. These ideas will be illustrated when we consider scattering by square well potential.

C. Square Well Potential – Illustrative Example for Partial Waves

15.11 PHASE SHIFT FOR SQUARE WELL POTENTIAL

Square well potential is the best example to illustrate many aspects of low-energy scattering theory. It is a simple model to illustrate resonant and non-resonant scattering, the concept of scattering length, resonance, and bound-state formation. The real advantage is that exact solutions are available.

The square well potential is

$$V(r) = \begin{cases} -V_o & \text{for} \quad r \le a \\ 0 & \text{for} \quad r \ge a \end{cases}$$

The Schrödinger equation for $l = 0$ partial wave is

$$\frac{d^2\chi_{in}}{dr^2} + \frac{2\mu}{\hbar^2}(E+V_o)\chi_{\text{in}} = 0 \quad r \le a \tag{15.69}$$

$$\frac{d^2\chi_{\text{out}}}{dr^2} + \frac{2\mu}{\hbar^2}E\,\chi_{\text{out}} = 0 \quad r \ge a \tag{15.70}$$

We use the following symbols:

For $E > 0$

$$k^2 = \frac{2\mu E}{\hbar^2}; \quad K^2 = \frac{2\mu(E+V_0)}{\hbar^2}; \quad k_0^2 = \frac{2\mu V_0}{\hbar^2};$$

α : scattering length

$$\beta^2 = -\frac{2\mu E}{\hbar^2} \qquad \text{For bound state} \quad (E<0)$$

$$K'^2 = \frac{2\mu(E+V_0)}{\hbar^2} \qquad \text{For bound state} \quad (E<0)$$

The solutions χ_{in} and χ_{out} are given by

$$\chi_{\text{in}} = A \sin \text{Kr} \quad r \le a \tag{15.71}$$

$$\chi_{\text{out}} = \frac{C_0}{k}[\cos\delta_0 \sin kr + \sin\delta_0 \cos kr] \qquad r \ge a$$

$$= \frac{e^{i\delta_o}}{k}\sin(kr+\delta_0) \tag{15.72}$$

These solutions are exact solutions. The constants A, C_0/k or $e^{i\delta_0}/k$ do not play any role in our discussion. The continuity of the wave function and its derivatives at $r = a$ implies

$$\frac{1}{\chi_{\text{in}}}\frac{d\chi_{\text{in}}}{dr}\bigg|_{r=a} = \frac{1}{\chi_{\text{out}}}\frac{d\chi_{\text{out}}}{dr}\bigg|_{r=a}$$

$$K\cot Ka = k\cot(ka+\delta_0)$$

$$\tan\,(ka+\delta_0) = \frac{k}{K}\tan Ka \tag{15.73}$$

This equation can be rewritten as follows:

$$\frac{\tan ka + \tan\delta_0}{1-\tan ka\tan\delta_0} = \frac{k}{K}\tan Ka$$

Rearranging this equation, we get

$$\tan\delta_0 = \frac{k\tan Ka - K\tan ka}{K + k\tan ka\tan Ka} \tag{15.74}$$

$$= \frac{\frac{k}{K}\tan Ka - \tan ka}{1+\frac{k}{K}\tan ka\tan Ka} \tag{15.75}$$

The structure of this equation is very similar to $\tan\,(\mathrm{A}-\mathrm{B}) = \dfrac{\tan A - \tan B}{1+\tan A\,\tan B}$ provided we identify $\tan A$ as

$$\tan A = \frac{k}{K}\tan Ka$$

or
$$A = \tan^{-1}\left(\frac{k}{K}\tan Ka\right)$$

Then,
$$\tan\,\delta_0 = \tan\,(A-B)$$

or
$$\delta_0 = n\pi + \tan^{-1}\left(\frac{k}{K}\tan ka\right) - ka \tag{15.76}$$

This is a useful result. Let us go back to the Equation (15.74), which can be rewritten as

$$\tan\delta_0 = \frac{\frac{ka}{Ka}\tan Ka - \tan ka}{1+\frac{ka}{Ka}\tan Ka.\tan ka} \tag{15.77}$$

So far, no approximation has been done. Since we are interested in the low-energy scattering we have $ka \ll 1$. Therefore, the Equation (15.77) becomes

$$\tan\delta_0 = \frac{ka}{Ka}\tan Ka - ka$$

$$= ka\left(\frac{\tan Ka}{Ka} - 1\right) \tag{15.78}$$

We are interested in studying how δ_0 and, hence, σ_0 vary as the potential depth V_0 is varied. However, before considering such variation in V_0, let us explain how the two kinds of scattering processes namely, resonant and non-resonant scattering emerge from (15.78).

Resonant and Non-resonant Scattering

The value of tan δ_0 and, hence, δ_0 depends on the right-hand side of the Equation (15.78). This, in turn, depends on $\frac{ka}{Ka}$ and tan Ka.

Note that
$$K^2 = \frac{2\mu V_0}{\hbar^2} + \frac{2\mu E}{\hbar^2} = k_0^2 + k^2 \tag{15.79}$$

In the low-energy scattering, k_0^2 is the dominant term. Therefore, $\frac{ka}{Ka}$ is very small.

If tan Ka is not very large, $\left[\frac{ka}{Ka}\tan Ka - 1\right]$ will be very small so that we can make the approximation

$$\tan\delta_0 \simeq a \,\text{small number}$$

or
$$\tan\delta_0 \simeq \delta_0$$

We may also have a situation where in spite of the fact that $ka/Ka << 1$, the quantity $[ka/Ka \tan Ka - 1]$ may not be very small.

If the potential V_0 is sufficiently strong, the approximation tan $\delta_0 \sim \delta_0$ will break down. In fact, when $Ka = \pi/2$, tan Ka and, hence, tan δ_0 become infinite, and δ_0 is now $n\pi/2$. The scattering under this condition is called resonant scattering. On the other hand, for non-resonant scattering, the conditions should be such that δ_0 should be far away from $n\pi/2$.

Having seen what is meant by resonant and non-resonant scattering, let us see how these scattering processes occur as the potential depth V_0 is varied.

Weak Potential – No Bound State – Non-resonant Scattering

The potential V_0 is not strong enough to form a bound state consisting of the incident and the target particles, and so the final states of both incident particles and the target particles are scattering states. So it is now a non-resonant scattering.

Since tan Ka is greater than ka (see (15.79)), δ_0 is now positive, and it is in the first quadrant. For very weak potential, the RHS of (15.78) is very small, and so we can make the approximation $\tan\delta_0 \sim \delta_0$.

$$\therefore \delta_0 = ka\left(\frac{\tan ka}{Ka} - 1\right) \tag{15.80}$$

We can make a further approximation for low-energy scattering by replacing Ka by $k_0 a$ $\left(Ka = \sqrt{k_0^2 + k^2 a} \simeq k_0 a\right)$

$$\therefore \delta_0 = ka\left(\frac{\tan k_0 a}{k_0 a} - 1\right) \tag{15.81}$$

∴ The total cross section is

$$\sigma_0 = \frac{4\pi}{k^2}\sin^2\delta_0 \simeq \frac{4\pi}{k^2}\delta_0$$

$$= 4\pi a^2\left(\frac{\tan k_0 a}{k_0 a} - 1\right)^2 \tag{15.82}$$

Since $k_0 a$ is independent of incident particle's energy E, the total cross section σ_0 is also independent of the incident particle's energy E. But it depends on V_0. As V_0 increases, the approximation tan $\delta_0 \sim \delta_0$ breaks down.

Threshold Potential V_0 for Bound-state-Resonant Scattering

As V_0 increases further, $\tan ka/Ka$ increases. When V_0 reaches a value such that $Ka = \pi/2$, $\tan\delta_0$ becomes infinite and $\delta_0 = \frac{\pi}{2}$. Since $\sin^2\delta_0$ is maximum for $\delta_0 = \frac{\pi}{2}$, the scattering cross section σ_0 becomes maximum.

$$\sigma_0 = \frac{4\pi}{k^2} = 4\pi a^2 \frac{1}{k^2 a^2} \tag{15.83}$$

Since $ka << 1$, σ_0 becomes very large now. The cross section now depends on the incident particles energy. What does it mean by $Ka = \pi/2$? For low-energy scattering, $Ka \simeq k_0 a = \pi/2$. We know that in the case of square well potential in three dimensions, it is not possible to have a bound state if V_0 is less than a threshold value (see Section 8.12 in Chapter 8).

The threshold condition for the formation of the bound state $k_0 a = \pi/2$. For $k_0 a$ close to $\pi/2$, we can show that a bound-state solution for the Schrödinger equation exists.

This is called zero energy bound state since its energy is $E = 0^- = -\Delta$, where Δ is very small. So in resonant scattering, the incident particle and the target particle together form a bound state.

Further Increase of V_0: Ramsauer–Townsend Effect

As the depth of the scattering potential V_0 increases further, the phase shift δ_0 becomes more than $\pi/2$. The scattering has now become non-resonant scattering. As we increase V_0 further, δ_0 also increases further, and a new phenomenon takes place when δ_0 becomes π; i.e., $\delta_0 = \pi$. For this potential strength, $\tan\delta_0 = \tan\pi = 0$, and so the scattering cross section becomes zero. This means that there is a target present on the path of the incident particle and yet there is no scattering. This phenomenon is called Ramsauer–Townsend effect. What is the condition for Ramsauer–Townsend effect to takes place?

$$\tan\delta_0 = ka\left(\frac{\tan Ka}{Ka} - 1\right) = 0$$

or

$$\tan ka = Ka \tag{15.84}$$

This happens when $Ka = 4.493$. This analysis of Ramsauer–Townsend effect requires a bit more careful consideration. The zero scattering has been obtained by considering *s*-wave scattering alone. The other partial waves are also present, though their contribution is very small compared to *s* wave for total cross section in low-energy scattering. There will be some scattering from higher-order partial waves when $\delta_0 = \pi$, and so, some scattering will take place even now. This will be reflected by the fact that the scattering cross section takes a dip and reaches a minimum value.

The actual experiment demonstrating Ramsauer–Townsend effect is the scattering of electron by heavy inert gas atoms. In terms of theoretical model, square well potential is the simplest model exhibiting zero scattering.

Resonance Again – Second Bound State

As the depth V_0 increases further when Ka reaches $3\pi/2$, the phase shift becomes $\delta_0 = 3\pi/2$, and the scattering cross section becomes maximum. The condition $Ka \simeq k_0 a = 3\pi/2$ corresponds to threshold potential V_0 to form a second bound state.

All these results can be generalized to the following results:

$$\delta_0 = n\pi + ka\left(\frac{\tan ka}{Ka} - 1\right) \text{ (n bound states)} \tag{15.85}$$

15.11.1 Behaviour of Scattering Length for Various Potential Strengths V_0

So far, square well potential was used to explain resonant and non-resonant scattering in terms of the phase shift δ_0 and the corresponding scattering cross section σ_0. Now let us illustrate how the scattering length α changes as Ka changes.

Let us recall (from (15.71) and (15.72)) that the s-wave solution to Schrödinger equation for square well potential is given by

$$\chi_{\text{in}} = \text{A} \sin Kr \quad r < a$$

$$\chi_{\text{out}} = \frac{C_0}{k}[\cos\delta_0 \sin kr + \cos kr \sin\delta_0] \quad r > a \tag{15.86}$$

χ_{out} is written in a functional form which is true only in the asymptotic limit for a general potential $V(r)$. But in the case of square well potential, this function is true for all $r \geq a$; i.e., the exact solution and the asymptotic form are the same.

For low-energy scattering,

$$ka << kr << 1.$$

$$\therefore \cos kr \simeq 1 \text{ and } \sin kr \simeq kr$$

Then, χ_{out} is now written as

$$\begin{aligned} \chi_{out} &= \frac{C_0}{k}[\sin\delta_0 + kr\cos\delta_0] \\ &= \frac{C_0}{k}\sin\delta_0[1 + (k\cot\delta_0)r] \end{aligned} \tag{15.87}$$

In the low-energy limit, we have

$$k\cot\delta_0 = -\frac{1}{\alpha}$$

where α is the scattering length.

$$\therefore \chi_{out} = \frac{C_0 \sin\delta_0}{k}\left(1 - \frac{r}{\alpha}\right) \quad \text{for } r > a$$

Let us define

$$\phi(r) = 1 - \frac{r}{\alpha} \tag{15.88}$$

This is an equation for the asymptote for χ_{out}. We have obtained this equation by considering $r > a$. However, let us extrapolate this equation till one gets $\phi(r) = 0$. It is an equation of a straight line with r-intercept α. At $r = \alpha$, $\phi(r) = 0$. Since we consider extrapolated function for $\phi(r)$, α may be negative or positive or even infinite.

Let us consider the scattering length for the following three cases namely, $Ka < \frac{\pi}{2}, Ka = \frac{\pi}{2}$ and $Ka > \frac{\pi}{2}$

Case 1:
$$Ka < \frac{\pi}{2}$$

The scattering process is now non-resonant scattering. δ_0 is positive, and hence, it is in first quadrant. Therefore, $\cot\delta_0$ is positive, and hence, α is negative.

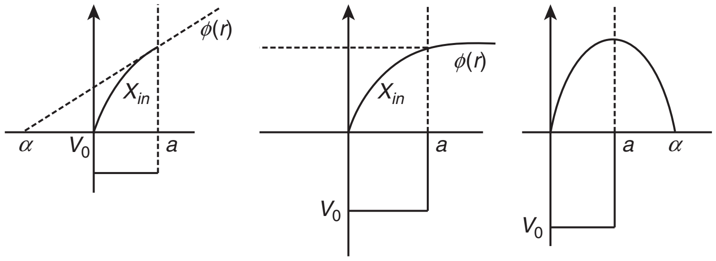
Case 2:
$$Ka = \frac{\pi}{2}$$

Scattering process is now resonant scattering. Now $\delta_0 = \frac{\pi}{2}$, and so $\cot\delta_0 = 0$. Therefore, α is infinite.

Case 3:
$$Ka < \frac{\pi}{2}$$

The scattering has become non-resonant again. δ_0 is in the second quadrant. So $\cot\delta_0$ is negative, which means α is positive.

These three cases are shown in Fig.15.8, in which χ_{in} and $\phi(r)$ are matched at $r = a$.



$Ka \simeq k_0 a < \frac{\pi}{2}$; $\delta_0 < \frac{\pi}{2}$

$Ka \simeq k_0 a = \frac{\pi}{2}$ $\delta_0 = \frac{\pi}{2}$

$Ka \simeq k_0 a > \frac{\pi}{2}$; $\delta_0 > \frac{\pi}{2}$

$$\chi(r) = \chi_{\text{in}} = A \sin kr \qquad \phi = \text{const}\left(1 - \frac{r}{\alpha}\right)$$

Fig.15.8 Scattering length for square well potential χ_{in} and ϕ is matched at $r = a$

Example 15.6 Show the condition for the existence of zero energy bound state and the condition for resonant scattering ($\delta_0 = \pi/2$) are the same in the case of square well potential of depth V_0.

Solution: By zero energy bound state, we mean that it is a bound state with energy E close to zero. We denote it by $E = 0^-$.

i.e.,
$$E = 0^- = -\Delta'$$

The Schrödinger equation for zero energy bound state is

$$\frac{d^2\chi'_{in}}{dr^2}+\frac{2m}{\hbar^2}(E+V_0)\chi'_{in}=0 \qquad r\le a \tag{15.89}$$

$$\frac{d^2\chi'_{out}}{dr^2}+\frac{2m}{\hbar^2}E\,\chi'_{out}=0 \quad r\ge a$$

Define $$K'^2=\frac{2m}{\hbar^2}(E+V_0)=\frac{2m}{\hbar^2}(-\Delta'+V_0)$$

$$\beta^2=-\frac{2mE}{\hbar^2}=\frac{2m|E|}{\hbar^2}=\frac{2m\Delta}{\hbar^2} \tag{15.90}$$

$$\frac{d^2\chi'_{in}}{dr^2}+K'^2\chi'_{in}=0 \quad r\le a$$

$$\frac{d^2\chi'_{out}}{dr^2}-\beta^2\chi'_{out}=0 \quad r\ge a$$

The solution to these equations is

$$\chi'_{\text{in}}=\text{A}\sin K'r \quad r\le a$$

$$\chi'_{\text{out}}=Be^{-\beta r} \quad r\ge a$$

Matching the logarithmic derivatives at $r=a$, we have

$$\frac{1}{\chi'_{in}}\frac{d\chi'_{in}}{dr}\bigg|_{r=a}=\frac{1}{\chi'_{out}}\cdot\frac{d\chi'_{out}}{dr}\bigg|_{r=a}$$

or $$K'\cot K'a=-\beta \tag{15.91}$$

Since the binding energy E is very small,

$$K'=\left[\frac{2\mu V_0}{\hbar^2}\right]^{1/2}=k_0$$

$$k_0\cot k_0a=-\beta \tag{15.92}$$

$$k_0\tan\left(\frac{\pi}{2}-k_0a\right)=-\beta$$

This represents the condition for the formation of a zero energy bound state.

For resonant scattering, k_0a is very close to $\pi/2$. i.e., the factor $(\pi/2-k_0a)$ is a very small number, and so we have $\tan(\pi/2-k_0a)\simeq\pi/2-k_0a$.

$$\therefore k_0\left(\frac{\pi}{2}-k_0a\right)=-\beta$$

or $$k_0a\simeq\frac{\pi}{2}+\frac{\beta}{k_0}$$

Therefore, the resonance condition $k_0a=\frac{\pi}{2}$ corresponds to zero energy bound-state formation.

Example 15.7 Determine the binding energy of a zero energy bound state in terms of the scattering length α of zero energy scattering state for a square well potential.

Solution: There are two terms here: zero energy bound state and zero energy scattering state. Their energies are denoted as $E = 0^-$ and $E = 0^+$.

Zero energy bound state: $E = 0^- = -\Delta'$

Zero energy scattering state: $E = 0^+ = \Delta$

Here Δ and Δ' are very small positive numbers. For bound states, we use the results of the previous example.

The Schrödinger equation for a scattering state is

$$\frac{d^2\chi_{\text{in}}}{dr^2} + \frac{2\mu}{\hbar^2}(E+V_0)\chi_{\text{in}} = 0 \qquad r < a$$

$$\frac{d^2\chi_{\text{out}}}{dr^2} + \frac{2\mu}{\hbar^2}E\,\chi_{\text{out}} = 0 \qquad r > a$$

Here $$K^2 = \frac{2\mu}{\hbar^2}(E+V_0) = \frac{2\mu}{\hbar^2}(\Delta+V_0) \approx \frac{2\mu}{\hbar^2}V_0$$

$$\therefore K^2 \simeq k_0^2 \text{ and } k^2 = \frac{2\mu}{\hbar^2}\Delta$$

The Schrdinger equation becomes

$$\frac{d^2\chi_{\text{in}}}{dr^2} + k_0^2\chi_{\text{in}} = 0 \qquad\qquad r < a$$

$$\frac{d^2\chi_{\text{out}}}{dr^2} + k^2\chi_{\text{out}} = 0 \text{ or } \frac{d^2\chi_{\text{out}}}{dr^2} \simeq 0 \qquad r > a$$

Since Δ is very small, we have neglected the term $k^2\chi_{\text{out}}$ in the second equation. We have already seen the solution to these equations, which is given by

$$\chi_{in} = A\sin Kr \qquad r < a$$

$$\chi_{\text{out}} = B(r-\alpha) \qquad r > a$$

Matching the logarithmic derivatives at $r = a$,

$$K\cot Kr = \frac{1}{r-\alpha} \tag{15.93}$$

Let us compare the wave functions inside r for the zero energy bound state and the zero energy scattering state.

$$\chi'_{in} = A\sin K'r \qquad\qquad E = 0^-$$

$$\chi_{in} = A\sin Kr \qquad\qquad E = 0^+$$

$$K'^2 = \frac{2m}{\hbar^2}(V_0 - \Delta') \text{ and } K^2 = \frac{2m}{\hbar}(V_0 + \Delta)$$

Since Δ and Δ' are very small, $K' \simeq K$.

$$\therefore K' \cot K'a \simeq K \cot Ka$$

or $$-\beta \simeq \frac{1}{a-\alpha} \quad \text{(See the Equations (15.91) and (15.93))}$$

For large α, we write

$$\frac{1}{\alpha} \sim \beta \tag{15.94}$$

Let us recall that α is the scattering length of the zero energy scattering state and β is the width of the zero energy bound state which corresponds to resonant scattering.

Therefore, the zero energy scattering with large scattering length represents the resonant scattering corresponding to the zero energy bound state.

The bound-state energy is

$$E = 0^- = -\Delta' = -\frac{\hbar^2\beta^2}{2\mu} = -\frac{\hbar^2}{2\mu\alpha^2} \tag{15.95}$$

This is an interesting result since it relates bound-state energy of the system consisting of incident and target particles with the scattering length of a scattered state near resonance phenomenon. A number of approximations have been made in deriving this relationship. So the validity of this relationship has to be addressed with care. For instance, consider the neutron-proton scattering. Deuteron is the bound state of the neutron and the proton. The experimental value of binding energy of deuteron is 2.22 Mev. The binding energy from the expression (15.95) is found to be 1.4 Mev (see Sakurai).

Example 15.8 Show that $k \cot\delta_0 = -\frac{1}{\alpha} + \frac{1}{2}ak^2$ for a square well potential, which has a zero energy bound state.

Solution: Let us go back to the Equation (15.73).

$$K \cot Ka = k \cot (ka + \delta_0) \tag{15.96}$$

Let us summarize the results for zero energy bound states ((15.90) and 15.92)).

$$E = 0^- = -\frac{\hbar^2\beta^2}{2\mu}$$

and $$k_0 \cot k_0 a = -\beta \tag{15.97}$$

As a first approximation, neglect ka in both sides of the Equation (15.96). Then, we have

$$k_0 \cot k_0 a = k \cot\delta_0 \tag{15.98}$$

$$\therefore k \cot\delta_0 = -\beta \quad \text{(use (15.92))}$$

Let us improve this approximation by including ka also. The following result is useful in further manipulation.

$$\begin{aligned}\cot(\theta+\Delta) &= \cot\theta - \Delta.\ \mathrm{cosec}^2\theta \\ &= \cot\theta - \Delta.\ (1+\cot^2\theta)\end{aligned}$$

LHS of (15.96):

$$\cot(ka+\delta_0) = \cot\delta_0 - ka(1+\cot^2\delta_0)$$

$$\therefore k\cot(ka+\delta_0) = k\cot\delta_0 - k^2a - \underbrace{k^2a\cot^2\delta_0}_{\beta^2 a} = k\cot\delta_0 - k^2a - \beta^2 a$$

$$= k\cot\delta_0 - (k^2+\beta^2)a \tag{15.99}$$

RHS of (15.96):

First let us define K^2 and K_1^2 as

$$K^2 = k_0^2 + k^2$$

$$K_1^2 = k_0^2 - \beta^2$$

$$K^2 = K_1^2 + k^2 + \beta^2 = K_1^2\left(1+\frac{k^2+\beta^2}{K_1^2}\right)$$

$$\therefore K = K_1\left(1+\frac{1}{2}\cdot\frac{k^2+\beta^2}{K_1^2}\right) = K_1 + \underbrace{\frac{1}{2}\cdot\frac{k^2+\beta^2}{K_1}}_{\varepsilon} = K_1+\varepsilon \tag{15.100}$$

$$\therefore Ka = K_1a + a\varepsilon$$

Let us now evaluate $K\cot Ka$.

$$\therefore K\cot Ka = (K_1+\varepsilon)\cot(K_1+\varepsilon)a = (K_1+\varepsilon)[\cot K_1a - a\varepsilon\,\text{cosec}^2 K_1a]$$

$$= (K_1+\varepsilon)[\cot K_1a - a\varepsilon\cdot(1+\cot^2 K_1a)]$$

$$= K_1\cot K_1a + \varepsilon[\cot K_1a - K_1a(1+\cot^2 K_1a)]$$

We have neglected the ε^2 term here.

Let us take $K_1 \simeq k_0$.

$$K\cot Ka \simeq k_0\cot k_0a + \varepsilon[\cot k_0a - k_0a(1+\cot^2 k_0a)]$$

$$\simeq -\beta + \varepsilon\left[-\frac{\beta}{k_0} - k_0a - k_0a\frac{\beta^2}{k_0^2}\right]$$

Neglecting the terms $\frac{\beta}{k_0}$ and $k_0a\frac{\beta^2}{k_0^2}$ we have

$$K\cot Ka \simeq -\beta - \varepsilon k_0a$$

$$\simeq -\beta - \frac{1}{2}(k^2+\beta^2)a \tag{15.101}$$

Here we have used Equation (15.97) and $\varepsilon = \frac{1}{2}\frac{k^2+\beta^2}{k_0}$

We have used the Equation (15.97) and $\varepsilon \sim \frac{1}{2}\frac{k^2+\beta^2}{k_0}$.

Equating (15.100) and (15.101) (LHS and RHS of the Equation (15.98)), we get

$$k\cot\delta_0 - a(k^2+\beta^2) = -\beta - \frac{1}{2}(k^2+\beta^2)a$$

$$\therefore k\cot\delta_0 = -\beta + \frac{1}{2}\beta^2 a + \frac{1}{2}ak^2$$

$$= -\frac{1}{\alpha} + \frac{1}{2}ak^2$$

where $$\frac{1}{\alpha} = -\beta + \frac{1}{2}\beta^2 a$$

This equation represents the general form of effective range r_{eff} for a short-range potential.

$$k\cot\delta_0 = -\frac{1}{\alpha} + \frac{1}{2}r_{\text{eff}}k^2$$

where r_{eff} is the effective range of the potential.

15.12 BREIT–WIGNER FORMULA

In this section, we get an expression for total scattering cross section for resonant scattering. This expression is known as Breit–Wigner formula. In resonant scattering, a quasi-bound state of the incident and the target particles is formed. In nuclear reactions, they are known as compound nucleus. They are basically unstable states, and they decay exponentially. Naturally, such states are accompanied by a statistical spread in energy Γ, known as width. Breit–Wigner formula for the total cross section reflects these ideas well.

The effective potential $V_{\text{eff}}(r)$ is

$$V_{eff}(r) = V(r) + \frac{l(l+1)2\mu}{\hbar^2 r^2}$$

This is shown in Fig.15.9 in which the bound states and the quasi bound states are shown. The bound states have energy $E < 0$ (like E_1 and E_2 as shown in Fig.15.9), and quasi bound states have energy $E > 0$ (like E_3 and E_4 as shown in Fig.15.9). Quasi bound states are not really bound states since the energy of the system is positive. Particles trapped in state of energy, say E_3, are classically prevented from escaping the region I. Quantum mechanically, it can reach the region III by tunnelling phenomenon. A trapped particle inside the centrifugal barriers will remain there for a long time (within the atomic scale) but leaks to region III ultimately. So, in effect, the system behaves as a bound state, which has a definite lifetime τ. This is known as resonance.

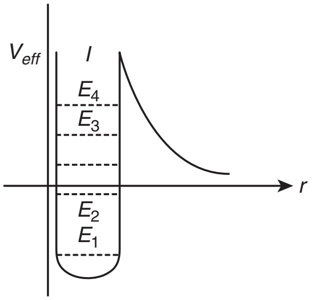


Fig.15.9 Bound states and quasi bound states

Our interest on quasi bound states in scattering theory is resonant scattering. It is found that when the energy of the incident particle is close to the quasi bound state energy, the scattering becomes resonant scattering.

Let E_r be the energy at which the resonance occurs for *l*th partial wave. The corresponding phase shift $\delta_l(E_r)$ is given by

$$\delta_l(E_r) = (2n+1)\frac{\pi}{2} \tag{15.102}$$

$$\therefore \cot\delta_l(E_r) = 0 \tag{15.103}$$

For energies E closer to E_r, let us make Taylor's expansion around E_r.

$$\begin{aligned}\cot\delta_l(E) &= \cot\delta_l(E_r) + (E-E_r)\frac{d}{dE}\cot\delta_l\,\Big|_{E=Er} + \frac{1}{2}(E-E_r)^2\frac{d^2}{dE^2}\cot\delta_l\,\Big|_{E=E_r}\\ &= -(E-E_r)\text{cosec}^2\delta_l\cdot\left.\frac{d\delta_l(E)}{dE}\right|_{E=E_r}\\ &= -(E-E_r)\cdot\left.\frac{1}{\sin^2\delta_l}\right|_{E=E_r}\cdot\left.\frac{d\delta_l}{dE}\right|_{E=E_r}\end{aligned}$$

Note that $\sin^2\delta_l = 1$ at resonance. We can rewrite the above equation by defining a new quantity Γ_l as

$$\left.\frac{d}{dE}\delta_l(E)\right|_{E=E_r} = \frac{2}{\Gamma_l}$$

$$\therefore \cot\delta_l(E) = -\frac{2}{\Gamma_l}(E-E_r)$$

The total cross section for *l*th partial wave is

$$\begin{aligned}\sigma_l &= \frac{4\pi}{k^2}\cdot(2l+1)\frac{1}{1+\cot^2\delta_l}\\ &= \frac{4\pi}{k^2}\cdot(2l+1)\frac{\Gamma_l^2/4}{(E-E_r)^2+\Gamma_l^2/4}\end{aligned}$$

Let us recall that the cross section is related to transition probability from an initial state to a group of final states. In the resonant scattering, it can be shown that Breit–Wigner formula represents the probability rate for the transition from an initial scattering state to an unstable state, which decays exponentially (see Cottingham and Greenwood).

Example 15.9 For a square well of depth V_0, show that the condition for resonance scattering of *l*th partial wave and the condition formation of the bound state corresponding to the given *l* are the same. Assume $Ka >> l >> ka$.

Solution: For a square well potential, exact solutions are available both inside and outside the range *a*. From Equation (8.72) we have.

$$\chi_{l,in} = A_l r j_l(Kr) \quad r < a$$

Let us recall from Equation (15.53)

$$\frac{1}{\lambda_l a} = \frac{1}{\chi_{l,in}} \cdot \left.\frac{d\chi_{l,in}}{dr}\right|_{r=a}$$

Substituting the expression for $\chi_{l,in}$ we get

$$\frac{1}{\lambda_l} = \frac{j_l(Ka) + Ka\, j_l'(Ka)}{j_l(Ka)}$$

In the low-energy scattering ($ka << 1$), from Equation (15.60) we have the expression for shift δ_l given by

$$\tan\delta_l = \frac{(ka)^{2l+1}(2l+1)}{[1\cdot3\cdot5\ldots(2l+1)]^2} \cdot \left\{\frac{(l+1)\lambda_l - 1}{l\lambda_l + 1}\right\}$$

Substituting the expression for λ_l in the above equation, we get

$$\tan\delta_l = \frac{(ka)^{2l+1}\cdot(2l+1)}{[1\cdot3\cdot5\cdots(2l+1)]^2} \cdot \left\{\frac{lj_l(Ka) - Kaj_l'(Ka)}{(l+1)\, j_l(Ka) + Ka\, j_l'(Ka)}\right\}$$

For resonant scattering, $\delta_l = \frac{\pi}{2}$ or $\tan\delta_l = \infty$. So the condition for resonant scattering

$$(l+1)\, j_l(Ka) + Kaj_l'(Ka) = 0 \qquad (15.104)$$

For a deep well, we can take Ka large enough to consider the asymptotic form of $j_l(Ka)$.

Then we have

$$j_l(\rho) \simeq \frac{\sin\left(\rho - \frac{l\pi}{2}\right)}{\rho}$$

and

$$j_l'(\rho) = \frac{\cos\left(\rho - \frac{l\pi}{2}\right)}{\rho} - \frac{\sin\left(\rho - \frac{l\pi}{2}\right)}{\rho^2} \simeq \frac{\cos\left(\rho - \frac{l\pi}{2}\right)}{\rho}$$

Using these results in (15.104), we get

$$\frac{(l+1)}{Ka}\sin\left(Ka - \frac{l\pi}{2}\right) + \cos\left(Ka - \frac{l\pi}{2}\right) = 0$$

$$\frac{(l+1)}{Ka}\sin\left(Ka - \frac{(l+1)\pi}{2} + \frac{\pi}{2}\right) + \cos\left(Ka - \frac{(l+1)\pi}{2} + \frac{\pi}{2}\right) = 0$$

or

$$\frac{(l+1)}{Ka}\cos\left(Ka - \frac{(l+1)\pi}{2}\right) - \sin\left(Ka - \frac{(l+1)\pi}{2}\right) = 0$$

$$\tan\left(Ka - \frac{(l+1)\pi}{2}\right) \simeq \frac{(l+1)}{Ka}.$$

$$\therefore Ka - \frac{(l+1)\pi}{2} = \tan^{-1}\left(\frac{l+1}{Ka}\right)$$

Since $Ka >> l$, $\frac{l+1}{Ka}$ is a very small number and

$$\tan^{-1}\left(\frac{l+1}{Ka}\right) = n\pi + \frac{(l+1)}{Ka} \simeq n\pi$$

or
$$Ka \simeq \left(n + \frac{1}{2}\right)\frac{\pi}{2} + \frac{l\pi}{2}$$

But this is the condition for the formation of bound states as we have already seen it in Section 8.12 in Chapter 8.

D. Born Approximation

15.13 GREEN'S FUNCTION TECHNIQUE IN NON-RELATIVISTIC SCATTERING THEORY

We have seen that the scattering problem in non-relativistic quantum mechanics reduces to solving the Schrödinger equation.

$$\left(\nabla^2 + \frac{2mE}{\hbar^2}\right)u(\mathbf{r}) = \frac{2m}{\hbar^2}V(r)u(\mathbf{r}) \tag{15.105}$$

subject to the boundary condition

$$u(\mathbf{r}) \to e^{ikz} + \frac{f(\theta,\phi)}{r}e^{ikr}$$

One of the methods of solving this problem is to convert this differential equation into an integral equation, which takes into account the boundary condition. The standard method is to use Green's function technique.

The Schrödinger Equation (15.105) can be rewritten as

$$(\nabla^2 + k^2)u(\mathbf{r}) = U(\mathbf{r})u(\mathbf{r}) \tag{15.106}$$

where
$$k^2 = \frac{2\mu E}{\hbar^2} \text{ and } U(r) = \frac{2\mu}{\hbar^2}V(r)$$

The Green's function of the linear operator $(\nabla^2 + k^2)$ is given by

$$(\nabla^2 + k^2)G(\mathbf{r},\mathbf{r}') = \delta(\mathbf{r} - \mathbf{r}') \tag{15.107}$$

It is easy to check that the solution $u(\mathbf{r})$ in terms of the Green's function $G(\mathbf{r},\mathbf{r}')$ is

$$\begin{aligned}
u(\mathbf{r}) &= \int U(\mathbf{r}')u(\mathbf{r}')G(\mathbf{r},\mathbf{r}')d^3\mathbf{r}' \\
(\nabla^2 + k^2)u(\mathbf{r}) &= (\nabla^2 + k^2)\int U(\mathbf{r}')u(\mathbf{r}')G(\mathbf{r},\mathbf{r}')d^3\mathbf{r}' \\
&= \int U(\mathbf{r}')u(\mathbf{r}')\underbrace{(\nabla^2 + k^2)G(\mathbf{r},\mathbf{r}')}_{\delta(\mathbf{r}-\mathbf{r}')}d^3\mathbf{r}' \\
&= U(\mathbf{r})u(\mathbf{r})
\end{aligned}$$

We have used the fact that ∇^2 can operate only on functions of $\mathbf{r}$.

Let $u_0(\mathbf{r})$ be a solution to the homogeneous equation

$$(\nabla^2 + k^2)u_0(\mathbf{r}) = 0$$

The most general solution is

$$u(\mathbf{r}) = u_0(\mathbf{r}) + \int G(\mathbf{r},\mathbf{r}')\frac{2m}{\hbar^2}V(\mathbf{r}')u(\mathbf{r}')d^3\mathbf{r}'$$

In scattering theory, we choose $u_0(\mathbf{r})$ to represent the incident beam. So $u_0(\mathbf{r})$ is

$$u_0(\mathbf{r}) = e^{ikz}$$

$$\therefore u(\mathbf{r}) = e^{ikz} + \int G(\mathbf{r},\mathbf{r}')\frac{2m}{\hbar^2}V(\mathbf{r}')u(\mathbf{r}')d^3\mathbf{r}' \tag{15.108}$$

Green's Function for Scattering Theory

For a differential equation, different sets of boundary conditions lead to different solutions. In Green's function technique, different sets of boundary conditions lead to different Green's function. So, for a particular linear operator, we can get many Green's functions, but we have to choose one among them which takes into account our boundary condition. In the scattering theory, we have to find the Green's function $G(\mathbf{r},\mathbf{r}')$ of $(\nabla^2 + k^2)$, which will lead to the solution

$$u(\mathbf{r}) \to e^{ikz} + f(\theta,\phi)\frac{e^{ikr}}{r}$$

It is not difficult to find such Green's function. It is given by

$$G(\mathbf{r},\mathbf{r}') = -\frac{1}{4\pi}\frac{e^{ik|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|} \tag{15.109}$$

For proof, see Appendix II.

Scattering Amplitude

The solution to Schrödinger Equation (15.108) is

$$u(\mathbf{r}) = e^{ikz} - \frac{1}{4\pi}\int \frac{e^{ik|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|}\cdot\frac{2m}{\hbar^2}V(\mathbf{r}')u(\mathbf{r}')d^3\mathbf{r} \tag{15.110}$$

Note that we have not really solved the Schrödinger equation. We have just converted a differential equation $u(\mathbf{r})$ into an integral equation for $u(\mathbf{r})$.

Let us write the second term as

$$I = -\frac{1}{4\pi}\int_{\Omega}\frac{e^{ik|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|}\cdot\frac{2m}{\hbar^2}V(\mathbf{r}')u(\mathbf{r}')d^3\mathbf{r}' \tag{15.111}$$

where Ω is a large region of space around the target as shown in Fig.15.10.

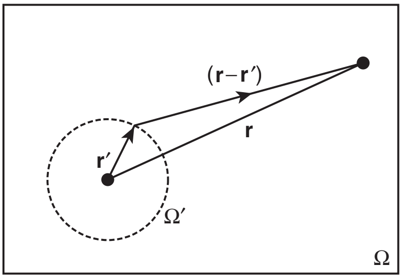


Fig.15.10 Integration region

We assume the potential to be short range. Though the integral extends over the entire volume Ω, $V(r)$ is significant only in a small region Ω' around the target.

$$\therefore I = -\frac{1}{4\pi}\int_{\Omega}\frac{e^{ik|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|}U(\mathbf{r}')u(\mathbf{r}')d^3\mathbf{r}'$$

$$\simeq -\frac{1}{4\pi}\int_{\Omega'}\frac{e^{ik|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|}U(\mathbf{r}')u(\mathbf{r}')d^3\mathbf{r}'$$

From Fig.15.10, it is obvious

$$|\mathbf{r}-\mathbf{r}'| = [(\mathbf{r}-\mathbf{r}')\cdot(\mathbf{r}-\mathbf{r}')]^{1/2}$$

$$= [r^2 - 2\mathbf{r}\cdot\mathbf{r}' + r'^2]^{1/2}$$

$$= r\left[1 - \frac{2\mathbf{r}\cdot\mathbf{r}'}{r^2} + \frac{r'^2}{r^2}\right]^{1/2}$$

$$\simeq r\left[1 - \frac{\mathbf{r}\cdot\mathbf{r}'}{r^2}\right] = r - \frac{\mathbf{r}}{r}\cdot\mathbf{r}' = r - \hat{\mathbf{r}}\cdot\mathbf{r}'$$

where $\hat{\mathbf{r}}$ is a unit vector. Since $\hat{\mathbf{r}}\cdot\mathbf{r}'$ is small compared to r, it can be neglected in the denominator.

$$I = -\frac{1}{4\pi}\int_{\Omega'}\frac{e^{ik(r-\hat{\mathbf{r}}\cdot\mathbf{r}')}}{r}U(\mathbf{r}')u(\mathbf{r}')d^3\mathbf{r}'$$

$$= -\frac{1}{4\pi}\frac{e^{ikr}}{r}\int_{\Omega'}e^{-ik\hat{\mathbf{r}}\cdot\mathbf{r}'}U(\mathbf{r}')u(\mathbf{r}')d^3\mathbf{r}'$$

Since $\mathbf{k} = k\hat{\mathbf{r}}$, we can write I as

$$I = -\frac{1}{4\pi}\frac{e^{ikr}}{r}\int_{\Omega'} e^{-i\mathbf{k}\cdot\mathbf{r}'}U(\mathbf{r}')u(\mathbf{r}')d^3\mathbf{r}'$$

Since the integrand is zero outside the volume Ω', we can replace the volume of the integration by Ω.

$$\text{ie} \int_{\Omega'} \simeq \int_{\Omega}$$

$$\therefore I = -\frac{1}{4\pi}\frac{e^{ikr}}{r}\int_{\Omega} e^{-i\mathbf{k}\cdot\mathbf{r}'}U(\mathbf{r}')u(\mathbf{r}')d^3\mathbf{r}'$$

or

$$I = -\frac{1}{4\pi}\frac{e^{ikr}}{r}\int_{\Omega} e^{-i\mathbf{k}\cdot\mathbf{r}}U(\mathbf{r})u(\mathbf{r})d^3\mathbf{r} \tag{15.112}$$

(We have replaced $\mathbf{r}'$ by $\mathbf{r}$, since $\mathbf{r}'$ is only an integration variable, a dummy variable.)

The wave function $u(\mathbf{r})$ is given by

$$u(\mathbf{r}) = u_0(\mathbf{r}) - \frac{1}{4\pi}\frac{e^{ikr}}{r}\int_{\Omega} e^{-i\mathbf{k}\cdot\mathbf{r}}U(\mathbf{r})u(\mathbf{r})d^3\mathbf{r} \tag{15.113}$$

Obviously, the scattering amplitude $f(\theta, \phi)$ is given by

$$f(\theta,\phi) = -\frac{1}{4\pi}\int e^{-i\mathbf{k}\cdot\mathbf{r}}\cdot\frac{2m}{\hbar^2}V(r)\cdot u(\mathbf{r})d^3\mathbf{r} \tag{15.114}$$

Born Series

The Equation (15.113) is an integral equation for $u(\mathbf{r})$. The advantage of the integral equation is that we can employ an iteration produce to obtain the solution, which lends itself to an approximation.

Let us go back to the Equation (15.108).

$$u(\mathbf{r}) = u_0(\mathbf{r}) + \int G(\mathbf{r},\mathbf{r}')U(\mathbf{r}')u(\mathbf{r}')d^3\mathbf{r} \tag{15.115}$$

The zeroth-order approximation is

$$u(\mathbf{r}) = u_0(\mathbf{r})$$

We can improve this wave function by replacing $u(\mathbf{r}')$ by $u_0(\mathbf{r})$ in the Equation (15.115).

i.e.,

$$u_{\mathrm{I}}(\mathbf{r}) = u_0(\mathbf{r}) + \int G(\mathbf{r},\mathbf{r}')\,U(\mathbf{r}')u_0(\mathbf{r}')d^3\mathbf{r}' \tag{15.116}$$

This wave function $u_{\mathrm{I}}(\mathbf{r})$ can be improved further as follows:

$$\begin{aligned} u_{\mathrm{II}}(\mathbf{r}) &= u_0(\mathbf{r}) + \int G(\mathbf{r},\mathbf{r}')\,U(\mathbf{r}')u_{\mathrm{I}}(\mathbf{r}')d^3\mathbf{r}' \\ &= u_0(\mathbf{r}) + \int G(\mathbf{r},\mathbf{r}')U(\mathbf{r}')u_0(\mathbf{r}')d^3\mathbf{r}' + \\ &\qquad \iint G(\mathbf{r},\mathbf{r}')U(\mathbf{r}')G(\mathbf{r}',\mathbf{r}'')U(\mathbf{r}'')u_0(\mathbf{r}'')d^3\mathbf{r}'d^3\mathbf{r}'' \end{aligned} \tag{15.117}$$

The wave function $u_{\mathrm{II}}(\mathbf{r})$ can be improved further as follows:

$$u_{\mathrm{III}}(\mathbf{r}) = u_0(\mathbf{r}) + \int G(\mathbf{r},\mathbf{r}')U(\mathbf{r}')u_{\mathrm{II}}(\mathbf{r}')d^3\mathbf{r}'$$

We can go on doing this procedure. In the end, we have

$$\begin{aligned} u(\mathbf{r}) = {} & \underbrace{u_0(\mathbf{r})}_{\text{zeroth order}} + \underbrace{\int G(\mathbf{r},\mathbf{r}')U(\mathbf{r}')u_0(\mathbf{r}')d^3\mathbf{r}'}_{\text{1st order}} + \\ & \underbrace{\iint G(\mathbf{r},\mathbf{r}')U(\mathbf{r}')G(\mathbf{r}',\mathbf{r}'')U(\mathbf{r}'')u_0(\mathbf{r}'')d^3\mathbf{r}'\cdot d^3\mathbf{r}''}_{\text{2nd order}} \\ & + \underbrace{\iiint G(\mathbf{r},\mathbf{r}')U(\mathbf{r}')G(\mathbf{r}',\mathbf{r}'')U(\mathbf{r}'')G(\mathbf{r}'',\mathbf{r}''')u_0(\mathbf{r}''')d^3\mathbf{r}'\cdot d^3\mathbf{r}''d^3\mathbf{r}''' + \dots}_{\text{3rd order}} \end{aligned} \tag{15.118}$$

This is known as Born series. Born approximation consists of restricting the Born series only to few terms. Then we have

$$\text{zeroth order}: \quad u_0(\mathbf{r}) \tag{15.119}$$

$$\text{first order}: \quad u_0(\mathbf{r}) + \int G(\mathbf{r},\mathbf{r}')U(\mathbf{r}')u_0(\mathbf{r}')d^3\mathbf{r}' \tag{15.120}$$

$$\begin{aligned} \text{second order}: \quad & u_0(\mathbf{r}) + \int G(\mathbf{r},\mathbf{r}')U(\mathbf{r}')u_0(\mathbf{r}')d^3\mathbf{r}' + \\ & \iint G(\mathbf{r},\mathbf{r}')U(\mathbf{r}')G(\mathbf{r},\mathbf{r}'')U(\mathbf{r}'')u_0(\mathbf{r}'')d^3\mathbf{r}'\cdot d^3\mathbf{r}'' \end{aligned} \tag{15.121}$$

15.14 FIRST-ORDER BORN APPROXIMATION

Let us recall that the wave function $u(\mathbf{r})$ is (Equation 15.110)

$$u(\mathbf{r}) = u_0(\mathbf{r}) - \frac{1}{4\pi}\int \frac{e^{ik|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|}U(\mathbf{r}')u(\mathbf{r}')d^3\mathbf{r}'$$

The second term can now be interpreted as the deviation of the wave function $u(\mathbf{r})$ from the incident wave function, $u_0(\mathbf{r})$. The deviation $g(\mathbf{r})$ is defined as

$$g(\mathbf{r}) = u(\mathbf{r}) - u_0(\mathbf{r}) \tag{15.122}$$

$$|g(\mathbf{r})| = \frac{1}{4\pi}\left|\int \frac{e^{ik|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|}U(\mathbf{r}')u(\mathbf{r}')d^3\mathbf{r}'\right| \tag{15.123}$$

The deviation $|g(\mathbf{r})|$ is very small, if

$$|g(\mathbf{r})| << |u_0(\mathbf{r})|$$

But $\qquad |u_0(\mathbf{r})| = |e^{ikz}| = 1$

Therefore, the deviation of $u(\mathbf{r})$ from $u_0(\mathbf{r})$ is small provided

$$|g(\mathbf{r})| << 1$$

If $|g(\mathbf{r})| << 1$, then we can restrict ourselves to first-order Born approximation. In the first order, we have

$$u(\mathbf{r}) = e^{ikz} - \frac{1}{4\pi}\int \frac{e^{ik|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|} U(\mathbf{r}')u_0(\mathbf{r}')d^3\mathbf{r}' \qquad \text{(from(15.110))}$$

$$\simeq e^{ikz} - \frac{e^{ikr}}{r}\frac{1}{4\pi}\int e^{-i\mathbf{k}\cdot\mathbf{r}}U(\mathbf{r})u_0(\mathbf{r})d^3\mathbf{r} \qquad \text{(from(15.113))}$$

The scattering amplitude $f(\theta, \phi)$ is

$$f(\theta,\phi) = -\frac{1}{4\pi}\int e^{-i\mathbf{k}\cdot\mathbf{r}}\frac{2m}{\hbar^2}V(\mathbf{r})u_0(\mathbf{r})d^3\mathbf{r} \tag{15.124}$$

Let us write $u_0(\mathbf{r})$ as $u_0(\mathbf{r}) = e^{ikz} = e^{i\mathbf{k}_0\cdot\mathbf{r}}$

where $$\mathbf{k}_0 = k\hat{z}$$

$$\therefore f(\theta,\phi) = -\frac{1}{4\pi}\int e^{-i(\mathbf{k}-\mathbf{k}_0)\cdot\mathbf{r}}\frac{2m}{\hbar^2}V(\mathbf{r})d^3\mathbf{r} \tag{15.125}$$

This is the expression for scattering amplitude in the first-order Born approximation.

Central Potential

Let us take the potential V to a central potential.

i.e., $$V = V(r)$$

Let us define $$\mathbf{K} = \mathbf{k} - \mathbf{k}_0$$

$$\hbar\mathbf{K} = \hbar\mathbf{k} - \hbar\mathbf{k}_0$$

$\hbar\mathbf{K}$ is called momentum transfer See Fig. 15.11.

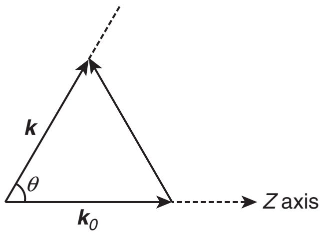


Fig.15.11 $\mathbf{K} = \mathbf{k} - \mathbf{k}_0$ for elastic scattering

$$\begin{aligned} K^2 &= [(\mathbf{k}-\mathbf{k}_0)\cdot(\mathbf{k}-\mathbf{k}_0)]^2 = k^2 + k_0^2 - 2\mathbf{k}_0\cdot\mathbf{k} \\ &= k^2 + k^2 - 2\mathbf{k}_0\cdot\mathbf{k} = 2k^2(1-\cos\theta) = 4k^2\sin\frac{\theta}{2} \end{aligned} \tag{15.126}$$

$$\begin{aligned}\therefore\quad f(\theta) &= -\frac{1}{4\pi}\cdot\frac{2\mu}{\hbar^2}\int_0^\infty r^2drV(r)\int_0^\pi \sin\theta d\theta\int_0^{2\pi} d\phi e^{i\mathbf{K}\cdot\mathbf{r}}\\ &= -\frac{\mu}{\hbar^2}\int_0^\infty r^2drV(r)\int_0^\pi e^{iKr\cos\theta}\sin\theta d\theta\\ &= -\frac{\mu}{\hbar^2}\int_0^\infty r^2drV(r)\int_{-1}^1 e^{-iKru}du\\ &= -\frac{2\mu}{k\hbar^2}\int_0^\infty rV(r)\sin Kr\ dr \qquad (15.127)\end{aligned}$$

Scattering Amplitude for Yukawa and Coulomb Potential

The Coulomb potential is

$$V(r) = -\frac{ZZ'e^2}{r}$$

The Yukawa potential is $V(r) = \frac{e^{-\lambda r}}{r}$. For our purpose, we can write Yukawa potential as

$$V(r) = -\frac{ZZ'e^2}{r}\cdot e^{-\lambda r}$$

Let us first try to evaluate the scattering amplitude for Coulomb potential.

$$\begin{aligned}f(\theta,\phi) &= -\frac{2\mu}{k\hbar^2}\cdot\int_0^\infty r\left(-\frac{ZZ'e^2}{r}\right)\sin Krdr\\ &= \frac{2\mu ZZ'e^2}{\hbar^2 k}\int_0^\infty \sin Krdr\end{aligned}$$

The integrand is an oscillating function, and so the integral does not converge. The usual trick is to multiply the integrand by a convergent factor $e^{-\lambda r}$ and in the end make $\lambda \to 0$. But this is literally replacing Coulomb potential by Yukawa potential. So we write

$$\begin{aligned}f(\theta,\phi) &= \frac{2\mu ZZ'e^2}{\hbar^2 K}\int_0^\infty \sin Kre^{-\lambda r}dr\\ &= \frac{2\mu ZZ'e^2}{\hbar^2}\cdot\frac{1}{K^2+\lambda^2}\\ &= \frac{2\mu ZZ'e^2}{\left(4p^2\sin^2\frac{\theta}{2}+\hbar^2\lambda^2\right)} \qquad \text{(use 15.126)}\end{aligned}$$

$\therefore$ The scattering cross section is

$$\frac{d\sigma}{d\Omega}=\begin{cases}\dfrac{\mu^2 Z^2 Z'^2 e^4}{4p^4\sin^4\dfrac{\theta}{2}} & \text{Coulomb potential}\\[2ex] \dfrac{4\mu^2 Z^2 Z'^2 e^4}{\left(4p^2\sin^2\dfrac{\theta}{2}+\hbar^2\lambda^2\right)^2} & \text{Yukawa potential}\end{cases}$$

15.15 VALIDITY OF BORN APPROXIMATION

The Born approximation is valid if the deviation from the incident wave function is very small. We have already seen that the deviation $g(\mathbf{r})$ is given by

$$g(\mathbf{r})=-\frac{1}{4\pi}\int_{\Omega}\frac{e^{ik|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|}U(\mathbf{r}')u(\mathbf{r}')d^3\mathbf{r}'$$

For Born approximation to be valid,

$$|g(\mathbf{r})|<<1$$

$$\therefore |g(\mathbf{r})|=\left|\frac{2\mu}{\hbar^2}\cdot\frac{1}{4\pi}\int_{\Omega}\frac{e^{ik|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|}V(\mathbf{r}')e^{ikz'}d^3\mathbf{r}'\right|<<1 \tag{15.128}$$

The maximum deviation is expected to occur near the origin where the target particle is situated. Therefore, the condition for the validity of Born approximation is

$$|\,g(0)\,|<<1$$

$$|g(0)|=\left|\frac{2\mu}{\hbar^2}\cdot\frac{1}{4\pi}\int\frac{e^{ikr'}}{r'}V(r')u_0(\mathbf{r}')e^{ikz'}d^3\mathbf{r}'\right|<<1 \tag{15.129}$$

For a spherically symmetric potential, we have

$$|g(0)|=\left|\frac{2\mu}{\hbar^2}\cdot\frac{1}{4\pi}\int_0^{\infty}r^2\frac{e^{ikr}}{r}V(r)\,dr\int_0^{\pi}e^{ikr\cos\theta}\sin\theta d\theta\int_0^{2\pi}d\phi\right|$$

$$=\left|\frac{\mu}{k\hbar^2}\int_0^{\infty}(e^{2ikr}-1)\cdot V(r)dr\right| \tag{15.130}$$

Yukawa Potential

The validity of Born approximation for Yukawa potential $V=\dfrac{V_0e^{-\lambda r}}{r}$ is determined by computing $|\,g(0)\,|$.

$$|g(0)|=\left|\frac{\mu V_0}{k\hbar^2}\int_0^{\infty}(e^{2ikr}-1)\cdot\frac{e^{-\lambda r}}{r}\right|$$

From integral calculus, we have

$$\int_0^\infty \frac{e^{-ar} - e^{-br}}{r} dr = \int_0^\infty dr \int_a^b e^{-\alpha r} d\alpha$$

$$= \int_a^b d\alpha \int_0^\infty e^{-\alpha r} dr = \ln \frac{b}{a}$$

$$\therefore \int_0^\infty (e^{2ikr} - 1) e^{-\lambda r} dr = \ln \frac{\lambda}{\lambda - 2ik} = -\ln \frac{\lambda - 2ik}{\lambda}$$

$$= -\ln\left(1 - \frac{2ik}{\lambda}\right)$$

$$= -\ln\left\{\left(1 + \frac{4k^2}{\lambda^2}\right)^{1/2} e^{-i\theta}\right\}$$

$$= -\ln\left(1 + \frac{4k^2}{\lambda^2}\right)^{1/2} + i\theta$$

$$= -\ln\left(1 + \frac{4k^2}{\lambda^2}\right)^{1/2} + i\tan^{-1}\frac{2k}{\lambda}$$

$$\therefore |g(0)|^2 = \left(\frac{\mu V_0}{k\hbar^2}\right)^2 \left\{\left[\ln\left(1 + \frac{4k^2}{\lambda^2}\right)^{1/2}\right]^2 + \left[\tan^{-1}\frac{2k}{\lambda}\right]^2\right\} << 1 \tag{15.131}$$

Let us first consider low-energy limit. For small values of k, $\tan^{-1}\frac{2k}{\lambda} \sim \frac{2k}{\lambda}$ and $\ln\left(1 + \frac{4k^2}{\lambda^2}\right) \sim \frac{4k^2}{\lambda^2}$. Retaining terms up to $\frac{k}{\lambda}$, we have

$$|g(0)|^2 = \frac{\mu_0 |V_0|}{k\hbar^2} \cdot \frac{2k}{\lambda} = \frac{2\mu |V_0|}{\hbar^2 \lambda} << 1 \tag{15.132}$$

For low-energy scattering, the deviation from the incident wave function is independent of velocity or energy of the incident particle. V_0 is supposed to be very small. V_0 should be sufficiently weak enough not to form a bound state. If bound states are formed, Born approximation fails.

Let us now consider high-energy limit. When k is very large, $\ln\sqrt{\left(1 + \frac{4k^2}{\lambda^2}\right)} \simeq \ln\frac{2k}{\lambda}$. Compared to in $\frac{2k}{\lambda}$, arc $\tan 2k/\lambda$ is small, and it can be neglected.

$$|g(0)| = \frac{2\mu |V_0|}{k\hbar^2} \ln\frac{2k}{\lambda}$$

$$= \frac{2|V_0|}{v\hbar} \ln\frac{2k}{\lambda} \tag{15.133}$$

Here, v is the velocity of the particle $\hbar k/\mu$. The larger the velocity, the smaller is the deviation from the incident wave. So Born approximation is a high-energy approximation.

Square Well Potential

The validity of Born approximation can be determined by computing $|g(0)|$ for square well potential.

$$|g(0)| = \left|\frac{\mu V_0}{k\hbar^2}\int_0^a (e^{2ikr} - 1)dr\right|$$
$$= \left|\frac{\mu V_0}{k\hbar^2}\cdot\frac{(e^{2ika} - 2ika - 1)}{2ik}\right|$$
$$= \frac{\mu V_0}{k^2\hbar^2}|e^{2ika} - 2ika - 1| \tag{15.134}$$

Let us consider low-energy limit first.

$$e^{2ika} - 2ika - 1 \simeq \left[1 + 2ika - \frac{4k^2a^2}{2} + \ldots\right] - 2ika - 1$$
$$\simeq -2k^2a^2$$
$$\therefore |g(0)| = \frac{\mu V_0}{2\hbar}a^2 \tag{15.135}$$

Again, we find that the deviation $|g(0)|$ is independent of velocity or energy of the incident particle. Let us now consider high-energy limit.

$$|(e^{2ika} - 2ika - 1)| = |\cos 2ka + i\sin 2ka - 2ika - 1|$$
$$= [(\cos 2ka - 1)^2 + (2ka - \sin 2ka)^2]^{1/2}$$
$$\simeq 2ka$$

All the other terms can be neglected compared to $2ka$ for large value of ka.

$$\therefore |\mathrm{g}(0)| = \frac{\mu|V_0|}{k^2\hbar^2}\cdot 2ka = \frac{2\mu|V_0|\,a}{\hbar v} \tag{15.136}$$

This is essentially the same as the condition for Yukawa potential.

15.16 BORN APPROXIMATION FROM TIME-DEPENDENT PERTURBATION THEORY

Let us now obtain the scattering amplitude of first-order Born approximation from time-dependent perturbation theory. The derivation of scattering cross section from time-dependent perturbation theory is more general. It is time-dependent perturbation theory which is more useful in scattering theory in quantum field theory.

Let us recall from the Equation (15.13) that the cross section $d\sigma$ is

$$d\sigma = \frac{1}{F}\cdot W_{i\to[\Delta\Omega]} \tag{15.137}$$

where F is the flux of the incident particle and $W_{i\to[\Delta\Omega]}$ is the transition probability rate for transition from an initial state i to a group of final states.

From Fermi's golden rule, $W_{i\to[\Delta\Omega]}$ is

$$W_{i\to[\Delta\Omega]} = \frac{2\pi}{\hbar}\cdot|\langle u_f|\,H'\,|u_i\rangle|^2\ \rho(E_f)$$

u_i and u_f are the eigenstates of the Hamiltonian H_0, which represents the free particle in the scattering theory. The wave functions representing the initial and the final states are given by

$$u_i(\mathbf{r}) = \frac{1}{\sqrt{V}}e^{i\mathbf{p}_i\cdot\mathbf{r}/\hbar}$$

and
$$u_f(\mathbf{r}) = \frac{1}{\sqrt{V}}e^{i\mathbf{p}_f\cdot\mathbf{r}/\hbar}$$

We have chosen the box normalized wave functions to describe the incident and scattered particle. The values of p_x, p_y and p_z are now discrete and are given by

$$0,\pm\frac{2\pi\hbar}{L},\pm\frac{4\pi\hbar}{L},\pm\frac{6\pi\hbar}{L},\ldots$$

Our first task is to determine the density of states $\rho(E_f)$; i.e., $\rho(E_f)dE_f$ is the number of states between E_f and $E_f + dE_f$. This amounts to counting the number of states in this range. Each state is now specified by a set of three numbers: (p_x, p_y, p_z). They form a set of grid points in three-dimensional p-space.

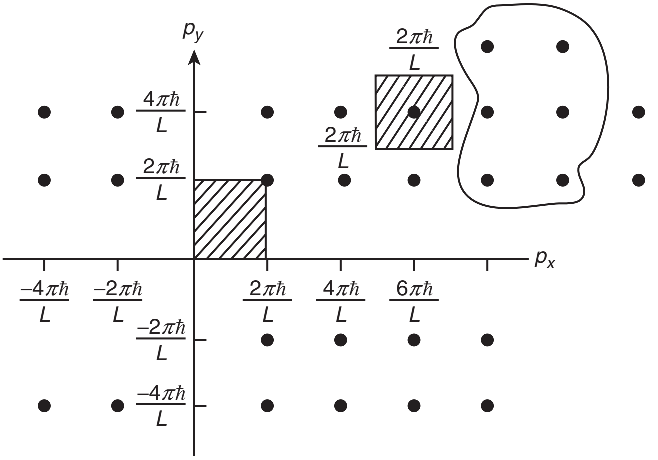


Fig.15.12 Representation of state in p-space

For the sake of clarity, we just show the two-dimensional p-space in Fig.15.12. The shaded region is a square of area $(2\pi\hbar/L)^2$. Each square of this area contains only one grid point or state. Therefore, the number of states within the area A is $A/(2\pi\hbar/L)^2$. Extending this idea to three-dimensional space, each cube of volume $(2\pi\hbar/L)^3$ contains one state. The number of states in volume ν_p is $\nu_p/(2\pi\hbar/L)^3$ or $\frac{L^3}{8\pi^3\hbar^3}\nu_p$. Let us determine the volume ν_p of our interest in scattering theory. The scattering states corresponding to the particles travelling in solid angle $\Delta\Omega$ are represented by the grid points within the corresponding solid angle $\Delta\Omega$ in p-space as shown in Fig.15.13. The number of states between p and $p + dp$ in this cone is the number of grid points within a volume slice of length dp and cross section area $p^2 d\Omega$.

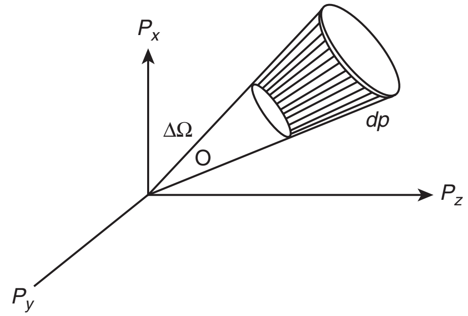


Fig.15.13 Volume $v_b = p^2 d\Omega dp$ for state with momenta between P and $p + dp$

The number of states between p and $p + dp$ is

$$g(p)dp = \frac{p^2 d\Omega dp}{(2\pi\hbar/L)^3} = \frac{L^3}{8\pi\hbar^3} \cdot p^2 dp d\Omega \tag{15.138}$$

Number of states between E and $E + dE$ is

$$\begin{aligned} \rho(E_f)dE_f &= g(p_f)\frac{dp_f}{dE_f} \cdot dE_f \\ &= \frac{(2\mu)^{3/2} E_f^{\ 1/2}}{2} \cdot \frac{L^3}{8\pi^3\hbar^3} \cdot d\Omega \end{aligned} \tag{15.139}$$

(We have used the relation $p = (2\mu E)^{1/2}$.)

$$\rho(E_f)dE_f = \mu \cdot (2\mu E_f)^{1/2} \frac{V}{8\pi^3\hbar^3}$$

$$\therefore W_{i\to[d\Omega]} = \frac{2\pi}{\hbar} \cdot |\langle u_f|H'|u_i\rangle|^2 \ \mu \frac{(2\mu E_f)^{1/2} V}{8\pi^3\hbar^3} \cdot d\Omega$$

Corresponding to the wave function $u_i = \frac{1}{\sqrt{V}} e^{i\mathbf{p}_i\cdot\mathbf{r}/\hbar}$, the incident flux F is

$$F = \frac{\hbar}{2i\mu}[u_i^* \nabla u_i - c.c] = \frac{\hbar k}{\mu V}$$

Therefore,

$$\frac{d\sigma}{d\Omega} = \frac{W_{i\to[d\Omega]}}{d\Omega \cdot F} = \frac{\mu^2 V^2}{4\pi^2\hbar^4} |\langle u_f|H'|u_i\rangle|^2 \tag{15.140}$$

The scattering amplitude $f(\theta, \phi)$ is given by

$$f(\theta,\phi) = -\frac{\mu V}{2\pi\hbar^2}\cdot |\langle u_f|H'|u_i\rangle| \tag{15.141}$$

The negative sign is chosen to match with the expression for $f(\theta, \phi)$ in (15.125).

$$\begin{aligned}\therefore f(\theta,\phi) &= -\frac{\mu V}{2\pi\,\hbar^2}\int u_f^* V(r) u_i d^3\mathbf{r} \\ &= -\frac{\mu V}{2\pi\,\hbar^2}\cdot\frac{1}{V}\int e^{i(\mathbf{p}_i-\mathbf{p}_f)\cdot\mathbf{r}/\hbar} \\ &= -\frac{\mu}{2\pi\,\hbar^2}\cdot\int e^{-i\mathbf{K}\cdot\mathbf{r}} V(\mathbf{r}) d^3\mathbf{r} \end{aligned}\tag{15.142}$$

which is the same as (15.125).

E. Lab Frame and Centre of Mass Frame in Scattering Theory

15.17 TRANSFORMATION FOR SCATTERING CROSS SECTIONS IN TWO FRAMES

The scattering theory process involves a two-particle system: an incident particle and a target particle. So the Schrödinger equation is

$$\left[-\frac{\hbar^2}{2m_1}\nabla_1^2 - \frac{\hbar^2}{2m_2}\nabla_2^2 + V(\mathbf{r}_1,\mathbf{r}_2)\right]\psi(\mathbf{r}_1,\mathbf{r}_2) = E_T\psi(\mathbf{r}_1,\mathbf{r}_2) \tag{15.143}$$

Instead of $\mathbf{r}_1$ and $\mathbf{r}_2$, we employ two other coordinates namely, the centre of mass coordinate $\mathbf{R}$ and relative coordinate $\mathbf{r}$, which are given by

$$\mathbf{R} = \frac{m_1\mathbf{r}_1 + m_2\mathbf{r}_2}{m_1+m_2} \quad \text{and } \mathbf{r} = \mathbf{r}_1 - \mathbf{r}_2 \tag{15.144}$$

It has been shown that in Chapter 8 the Equation (15.143) becomes

$$\left[-\frac{\hbar^2}{2M}\nabla_R^2 - \frac{\hbar^2}{2\mu}\nabla_r^2 + V(\mathbf{r})\right]\Phi(\mathbf{R})u(\mathbf{r}) = (E_{cm}+E)\Phi(\mathbf{R})u(\mathbf{r})$$

where

$$-\frac{\hbar^2}{2M}\nabla_R^2\Phi(\mathbf{R}) = E_{cm}\Phi(\mathbf{R}) \tag{15.145}$$

and

$$\left[\frac{-\hbar^2}{2\mu}\nabla^2 + V(\mathbf{r})\right]u(\mathbf{r}) = Eu(\mathbf{r}) \tag{15.146}$$

In scattering theory, the coordinate system based on $\mathbf{r}_1$ and $\mathbf{r}_2$ is called laboratory coordinate system. The coordinate system based on $\mathbf{R}$ and $\mathbf{r}$ are is centre of mass coordinate system. All the analyses of scattering theory in the previous sections were using the Schrödinger equation of the type (15.146).

Therefore, all the theoretical calculations in the previous sections have been done in the centre of mass coordinate systems where the centre of mass is at rest.

Consider the experimental situations, in which, the target is at rest. The centre of mass is moving, and so $\mathbf{r}_1$ and $\mathbf{r}_2$ are better coordinates in such situations. This is, of course, laboratory coordinate system.

So we have two types of situations:

Theoretical calculations – Centre of mass coordinate system

Experiments – Laboratory coordinate system

Therefore, there is a necessity to find the relationship relating the values of scattering cross sections in two different coordinate systems.

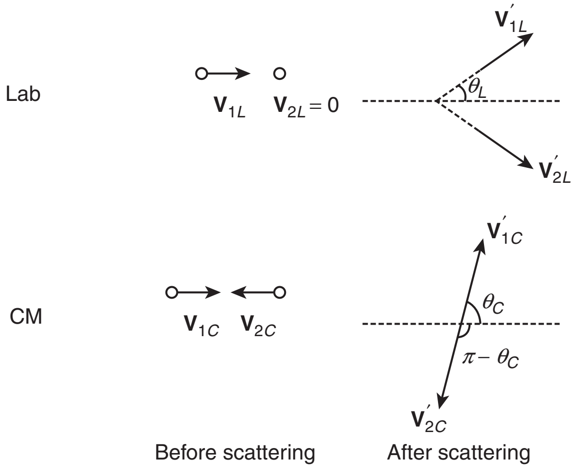


Fig.15.14 Scattering process in lab and centre of mass coordinate systems

Let us first obtain some kinematical relations, which will be useful in relating the solid angle in the two coordinate systems. Let $\mathbf{V}_{1L}$ and $\mathbf{V}_{2L}$ be the velocities of the incident particle and the target particle in the lab coordinate system before scattering. They become $\mathbf{V}'_{1L}$ and $\mathbf{V}'_{2L}$ after scattering.

The corresponding quantities in the centre of mass frame are $\mathbf{V}_{1C}$ and $\mathbf{V}_{2C}$ before scattering and $\mathbf{V}'_{1C}$ and $\mathbf{V}'_{2C}$ after scattering as shown in Fig.15.14.

The velocity of the centre of mass (as measured in the lab frame) is

$$\mathbf{V}_C = \frac{m_1\mathbf{V}_{1L} + m_2\mathbf{V}_{2L}}{m_1 + m_2} = \frac{m_1\mathbf{V}_{1L} + m_2\mathbf{V}_{2L}}{M}$$

$\mathbf{V}_{1C}$ and $\mathbf{V}_{2C}$ are the velocities of the particles 1 and 2 before scattering as measured in the centre of mass coordinate system. i.e., they are the velocities with respect to the centre of mass.

$$\therefore \mathbf{V}_{1C} = \mathbf{V}_{1L} - \mathbf{V}_{CM} = \mathbf{V}_{1L} - \frac{m_1\mathbf{V}_{1L} + m_2\mathbf{V}_{2L}}{m_1 + m_2}$$

$$= \frac{m_2}{m_1 + m_2}(\mathbf{V}_{1L} - \mathbf{V}_{2L}) = \frac{\mu}{m_1}\mathbf{V}_r \tag{15.147}$$

$$\mathbf{V}_{2C} = \mathbf{V}_{2L} - \mathbf{V}_{CM} = \mathbf{V}_{2L} - \frac{m_1\mathbf{V}_{1L} + m_2\mathbf{V}_{2L}}{m_1 + m_2} = -\frac{\mu}{m_2}\mathbf{V}_r$$

The momenta of the particles 1 and 2 in the centre of mass coordinate system are

$$\mathbf{p}_{1C} = m_1\mathbf{V}_{1C} = \mu\mathbf{V}_r$$

and
$$\mathbf{p}_{2C} = m_2\mathbf{V}_{2C} = -\mu\mathbf{V}_r$$

Therefore, the total momentum is zero in the centre of mass frame. Though we have derived these equations for the particles before scattering, these relations hold good after scattering also. Therefore, the total momentum $\mathbf{p} = \mathbf{p}_{1C} + \mathbf{p}_{2C} = 0$ always.

The total kinetic energy in the centre of mass frame is a conserved quantity.

$$\frac{\mathbf{p}_{1C}{}^2}{2m_1} + \frac{p_{2C}{}^2}{2m_2} = \frac{\mathbf{p}'_{1C}{}^2}{2m_1} + \frac{\mathbf{p}'_{2C}{}^2}{2m_2}$$

or
$$\frac{p_{1C}{}^2}{2}\left(\frac{1}{m_1} + \frac{1}{m_2}\right) = \frac{p'_{1C}{}^2}{2}\cdot\left(\frac{1}{m_1} + \frac{1}{m_2}\right)$$

or
$$\mathbf{p}_{1C}^2 = \mathbf{p}'^2_{1C}$$

$$\therefore |\mathbf{p}_{1C}| = |\mathbf{p}'_{1C}|$$

or
$$|\mathbf{V}_{1C}| = |\mathbf{V}'_{1C}|$$

So, in the centre of mass coordinate system, only the direction of the momentum changes after scattering, i.e., the magnitude of the momentum before scattering and after scattering is the same. We will use this result to determine the geometrical relationship between θ_L and θ_C.

Since the target is at rest initially, $\mathbf{V}_{2L} = 0$

$$\therefore \mathbf{V}_{CM} = \frac{m_1\mathbf{V}_{1L}}{m_1 + m_2} \tag{15.148}$$

Therefore, the direction of the velocity of the centre of mass is always along the incident direction of the beam.

Let us consider the velocities of the particles after scattering.

$$\mathbf{V}'_{1L} = \mathbf{V}'_{1C} + \mathbf{V}_{CM}$$
$$\mathbf{V}'_{2L} = \mathbf{V}'_{2C} + \mathbf{V}_{CM}$$

Since $\mathbf{V}_{CM}$ is parallel to the direction of the incident beam, only the parallel components of $\mathbf{V}'_{1L}$ and $\mathbf{V}'_{1C}$ differ, while the vertical components of $\mathbf{V}'_{1L}$ and $\mathbf{V}'_{1C}$ are the same.

$$\therefore V'_{1L}\cos\theta_L = V'_{1C}\cos\theta_C + V_{CM}$$
$$V'_{1L}\sin\theta_L = V'_{1C}\sin\theta_C$$

$$\therefore \tan\theta_L = \frac{V'_{1C}\sin\theta_C}{V'_{1C}\cos\theta_C + V_{CM}} = \frac{\sin\theta_C}{\cos\theta_C + \dfrac{V_{CM}}{V'_{1C}}}$$

Let us now determine $\dfrac{V_{CM}}{V'_{1C}}$

From (15.147) and (15.148), we have

$$\mathbf{V}_{1C} = \frac{m_2}{m_1 + m_2}\mathbf{V}_{1L} = \frac{m_2}{m_1}\mathbf{V}_{CM} \tag{15.149}$$

Since $|\mathbf{V}'_{1C}| = |\mathbf{V}_{1C}|$, we have

$$\mathbf{V}'_{1C} = \frac{m_2}{m_1}V_{CM}$$

$$\therefore \tan\theta_L = \frac{\sin\theta_C}{\cos\theta_C + \dfrac{m_1}{m_2}} \tag{15.150}$$

The observer in the lab frame finds ΔN particles per second to be scattered into the solid angle $\Delta\Omega$. This number should be the same with respect to an observer in the centre of the mass frame also. But the value of the solid angle for these ΔN particles as measured by the observer in the centre of mass frame is $\Delta\Omega_{CM}$, which is different from $\Delta\Omega_{Lab}$.

$$\therefore \Delta N = \left(\frac{d\sigma}{d\Omega}\right)_{Lab}\cdot(\Delta\Omega)_{Lab} = \left(\frac{d\sigma}{d\Omega}\right)_{CM}\cdot(\Delta\Omega)_{CM}$$

$$\left(\frac{d\sigma}{d\Omega}\right)_L\cdot\sin\theta_L d\theta_L d\phi_L = \left(\frac{d\sigma}{d\Omega}\right)_C \sin\theta_C d\theta_C\cdot d\phi_C$$

It is easy to see that $d\phi_L = d\phi_C$.

$$\therefore \left(\frac{d\sigma}{d\Omega}\right)_L = \left(\frac{d\sigma}{d\Omega}\right)_C \left|\frac{\sin\theta_C}{\sin\theta_L}\right|\cdot\left|\frac{d\theta_C}{d\theta_L}\right| \tag{15.151}$$

From (15.150), we can get

$$\sin\theta_L = \frac{\sin\theta_C}{\left[1 + \left(\dfrac{m_1}{m_2}\right)^2 + \dfrac{2m_1}{m_2}\cos\theta_C\right]^{1/2}}$$

and

$$\cos\theta_L = \frac{\cos\theta_C + \dfrac{m_1}{m_2}}{\left[1 + \left(\dfrac{m_1}{m_2}\right)^2 + \dfrac{2m_1}{m_2}\cos\theta_C\right]^{1/2}}$$

$$\left(\text{If } \tan\theta = \frac{a}{b}, \sin\theta = \frac{a}{\sqrt{a^2+b^2}}, \cos\theta = \frac{b}{\sqrt{a^2+b^2}}\right).$$

Differentiating Equation (15.150) with respect to θ_L, we get

$$\sec^2\theta_L = \frac{\left(\cos\theta_C\left(\cos\theta_C + \frac{m_1}{m_2}\right) + \sin^2\theta_C\right)\frac{d\theta_C}{d\theta_L}}{\left(\cos\theta_C + \frac{m_1}{m_2}\right)^2} = \frac{\left(1 + \frac{m_1}{m_2}\cos\theta_C\right)\frac{d\theta_C}{d\theta_L}}{\left(\cos\theta_C + \frac{m_1}{m_2}\right)^2}$$

Making use of the expression for $\cos\theta_L$ we get

$$\frac{d\theta_C}{d\theta_L} = \frac{\left[1 + \left(\frac{m_1}{m_2}\right)^2 + 2\frac{m_1}{m_2}\cos\theta_C\right]}{\left(1 + \frac{m_1}{m_2}\cos\theta_C\right)}$$

So, from (15.151), we get

$$\therefore \left(\frac{d\sigma}{d\Omega}\right)_L = \left(\frac{d\sigma}{d\Omega}\right)_C \cdot \frac{\left[1 + \left(\frac{m_1}{m_2}\right)^2 + \frac{2m_1}{m_2}\cos\theta_C\right]^{3/2}}{\left[1 + \frac{m_1}{m_2}\cos\theta_C\right]} \tag{15.152}$$

F. Identical Particles and Scattering Amplitude

15.18 SCATTERING CROSS SECTION AND INDISTINGUISHABILITY

If the incident and the target particles happen to be the same kind, the whole system is a system of identical particles. For instance, it may be a proton-proton or neutron-neutron scattering. The wave function describing such systems should take into account the symmetric or antisymmetric nature with respect to exchange of particle coordinates.

Consider the collision of two identical particles *A* and *A*.

In the case of indistinguishable particles, the wave function for the system is either symmetric or antisymmetric with respect to the exchange of particles.

$$\psi(\mathbf{r}_1, \mathbf{r}_2) : \begin{cases} \psi_s(\mathbf{r}_2, \mathbf{r}_1) = \psi_s(\mathbf{r}_1, \mathbf{r}_2) \\ \psi_a(\mathbf{r}_2, \mathbf{r}_1) = -\psi_a(\mathbf{r}_1, \mathbf{r}_2) \end{cases}$$

Instead of $\mathbf{r}_1$ and $\mathbf{r}_2$, we have been working in terms $\mathbf{R}$ and $\mathbf{r}$, where $\mathbf{R} = \frac{m_1\mathbf{r}_1 + m_2\mathbf{r}_2}{m_1 + m_2} = \frac{\mathbf{r}_1 + \mathbf{r}_2}{2}$ and $\mathbf{r} = \mathbf{r}_1 - \mathbf{r}_2$. Here $\mathbf{R}$ is the coordinate of the centre of mass and $\mathbf{r}$ is the relative coordinate.

Under $$\mathbf{r}_1 \leftrightarrow \mathbf{r}_2, \mathbf{R} \to \mathbf{R}, \mathbf{r} \to -\mathbf{r} \tag{15.153}$$

Therefore, for a symmetric wave function, we have

$$\begin{aligned} &\psi_s(\mathbf{r}_2,\mathbf{r}_1) = \psi_s(\mathbf{r}_1,\mathbf{r}_2) \\ \text{or} \quad &\psi_s(\mathbf{R},-\mathbf{r}) = \psi_s(\mathbf{R},\mathbf{r}) \end{aligned} \tag{15.154}$$

Similarly, for an antisymmetric wave function, we have

$$\begin{aligned} &\psi_a(\mathbf{r}_2,\mathbf{r}_1) = -\psi_a(\mathbf{r}_1,\mathbf{r}_2) \\ \text{or} \quad &\psi_a(\mathbf{R},-\mathbf{r}) = -\psi_a(\mathbf{R},\mathbf{r}) \end{aligned} \tag{15.155}$$

Let us write $\psi(\mathbf{r}_1,\mathbf{r}_2)$ as $\psi(\mathbf{r}_1,\mathbf{r}_2) = \chi(\mathbf{R})\phi(\mathbf{r})$

Since **R** remains the same with respect to the exchange of particles 1 and 2, it is enough to consider $\phi(\mathbf{r})$.

Let us recall that in spherical polar coordinates,

$$\mathbf{r} \to -\mathbf{r} \quad \Rightarrow \quad r \to r, \theta \to \pi - \theta, \phi \to \pi + \phi \tag{15.156}$$

The wave functions for bosons and fermions are

$$\psi_s = (e^{ikz} + e^{-ikz}) + \left[\frac{f(\theta) + f(\pi - \theta)}{r}\right] e^{ikr} - \text{Bosons} \tag{15.157}$$

$$\psi_a = (e^{ikz} - e^{-ikz}) + \left[\frac{f(\theta) - f(\pi - \theta)}{r}\right] e^{ikr} - \text{Fermions} \tag{15.158}$$

The corresponding processes are shown in Fig.15.15.

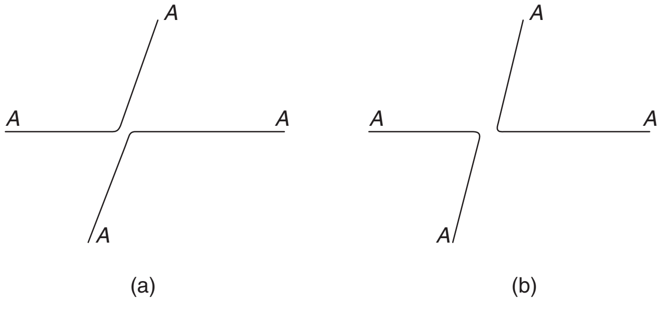


Fig.15.15 Scattering of identical particles

Suppose the upper detector detects *A*. Which *A*? Is it from right (Fig. 15. 15(a)) or is it from left (Fig. 15. 15(b))? In classical physics, two identical particles are distinguishable with the help of trajectories. By tracing back through their trajectory, we can determine whether the particle *A* has come from left or right. These two processes are mutually exclusive. Therefore, the probability of *A* being from left or right is

$$\begin{aligned}\frac{d\sigma}{d\Omega} &= \left|\frac{d\sigma}{d\Omega}\right|_{\text{left}} + \left|\frac{d\sigma}{d\Omega}\right|_{\text{right}} \\ &= |f(\theta,\phi)|^2 + |f(\pi-\theta,\pi+\phi)|^2\end{aligned}$$

In the case of quantum mechanics, the particles are indistinguishable. From the final outcome, it is not possible to tell whether the process in Fig.15.15(a) or Fig.15.19(b) has occurred. We have to go back to the wave function which takes both processes into account.

The wave function is now given by

$$\psi = (e^{ikz} \pm e^{-ikz}) + [f(\theta,\phi) \pm f(\pi-\theta,\pi+\phi)]\frac{e^{ikr}}{r}$$

The scattering cross section $\frac{d\sigma}{d\Omega}$ is

$$\begin{aligned}\frac{d\sigma}{d\Omega} &= |f(\theta,\phi) \pm f(\pi-\theta,\pi+\phi)|^2 \\ &= [f^*(\theta,\phi) \pm f^*(\pi-\theta,\pi+\phi)][f(\theta,\phi) \pm f(\pi-\theta,\pi+\phi)] \\ &= |f(\theta,\phi)|^2 + |f(\pi-\theta,\pi+\phi)|^2 \\ &\quad \pm 2\,\mathrm{Re}\ f(\theta,\phi)\ f(\pi-\theta,\pi+\phi)\end{aligned}$$

The first two terms show the same result as it is based on classical physics in which the two particles are distinguishable. The second term is unique to quantum mechanics. It is due to interference between the two processes shown in Fig.15.15.

In partial wave analysis, the scattering amplitude for distinguishable particles is

$$f(\theta,\phi) = \frac{1}{k}\sum_{l=0}^{\infty}(2l+1)\,e^{i\delta_l}\,\sin\delta_l P_l(\cos\theta)$$

Note that $$P_l(\cos(\pi-\theta)) = P_l(-\cos\theta) = (-1)^l P_l(\cos\theta)$$

So we have $$f(\pi-\theta) = \frac{1}{k}\sum_{l=0}^{\infty}(2l+1)\,e^{i\delta_l}\,\sin\delta_l P_l(\cos(\pi-\theta))$$

$$= \frac{1}{k}\sum_{l=0}^{\infty}(2l+1)\,e^{i\delta_l}\,\sin\delta_l (-1)^l P_l(\cos\theta)$$

The scattering amplitude for indistinguishable particles is

$$f(\theta,\phi) \pm f(\theta,\phi) = \frac{1}{k}\sum_{l=0}^{\infty}(2l+1)\,e^{i\delta_l}\,\sin\delta_l P_l(\cos\theta)[1 \pm (-1)^l]$$

For bosons, the partial waves of odd *l* do not contribute to the scattering amplitude. Similarly, even *l* partial waves do not contribute for fermions.

Identical Fermions – Including Spin

So far, we have not taken into account the spin degree of freedom. When spin degree of freedom is included, the wave functions for the system itself become more involved. Let us consider the scattering of two spin 1/2 fermions. This may be the case of electron-electron, proton-proton, or neutron-neutron scattering. When angular momenta of quantum number $s_1 = 1/2$ and $s_2 = 1/2$ are combined, the resulting angular momentum is either $s = 1$ (triplet states) or $s = 0$ (singlet states). The wave function describing the system of the two functions is $\psi_a(\mathbf{r}_1 s_1; \mathbf{r}_2 s_2)$.

$$\psi_a(\mathbf{r}_2 s_2; \mathbf{r}_1 s_1) = -\psi_a(\mathbf{r}_1 s_1; \mathbf{r}_2 s_2)$$

Let us write the wave function ψ_a as

$$\psi_a(\mathbf{r}_1 s_1; \mathbf{r}_2 s_2) = \chi(s_1, s_2)\Phi(\mathbf{r}_1, \mathbf{r}_2)$$

Now we have two possibilities.

$$\psi_a(\mathbf{r}_1 s_1; \mathbf{r}_2 s_2) = \chi_t(s_1, s_2)\Phi_a(\mathbf{r}_1, \mathbf{r}_2)$$

and

$$\psi_a(\mathbf{r}_1 s_1; \mathbf{r}_2 s_2) = \chi_s(s_1, s_2)\Phi_s(\mathbf{r}_1, \mathbf{r}_2)$$

The spin dependent wave functions χ_t and χ_s are symmetric and antisymmetric with respect to exchange of spin coordinates. The corresponding spatial wave functions $\Phi_a(\mathbf{r}_1, \mathbf{r}_2)$ and $\Phi_s(\mathbf{r}_1, \mathbf{r}_2)$ have to be antisymmetric and symmetric with respect to exchange of position coordinates.

Therefore, we have the following results:

Triplet:

$$\psi_a = (e^{ikz} - e^{-ikz}) + (f_t(\theta,\phi) - f_t(\pi-\theta,\phi+\pi))\frac{e^{ikr}}{r}$$

$$\left(\frac{d\sigma}{d\Omega}\right)_t = |f_t(\theta,\phi)|^2 + |f_t(\pi-\theta,\pi+\phi)|^2 - 2\,\mathrm{Re}\, f_t^*(\theta,\phi) f_t(\pi-\theta,\pi+\phi)$$

Singlet:

$$\psi_s = (e^{ikz} + e^{-ikz}) + (f_s(\theta,\phi) + f_s(\pi-\theta,\phi+\pi))\frac{e^{ikr}}{r}$$

$$\left(\frac{d\sigma}{d\Omega}\right)_s = |f_s(\theta,\phi)|^2 + |f_s(\pi-\theta,\pi+\phi)|^2 + 2\,\mathrm{Re}\, f_s^*(\theta,\phi) f_s(\pi-\theta,\pi+\phi)$$

It is interesting to consider a special case, namely $\theta = \frac{\pi}{2}$ for the case of scattering amplitudes which are independent of ϕ.

For

$$\theta = \frac{\pi}{2},$$

$$\left(\frac{d\sigma}{d\Omega}\right)_t = 0$$

$$\left(\frac{d\sigma}{d\Omega}\right)_s = 4\left|f_s\left(\frac{\pi}{2}\right)\right|^2$$

For a system of two particles, there are three possible values of m_s (± 1, 0), and there is one singlet states. So, in general, there are four possible spin polarization states. There are experimental situations where details of polarization of the incident and target particles are not involved. In such cases, we can take the statistical average of the cross sections given by

$$\frac{d\sigma}{d\Omega}=\frac{3}{4}\left(\frac{d\sigma}{d\Omega}\right)_t+\frac{1}{4}\left(\frac{d\sigma}{d\Omega}\right)_s$$

The above analysis does not take into account other details like spin-dependent interaction.

APPENDIX I

Optical Theorem

There is a relationship between the total cross section and the imaginary part of scattering amplitude. This is based on the fact that the number of particles scattered out of the incident beam is the same as the number of particles missing in the forward direction (in the direction of the incident beam) behind the target.

The wave function for the incident beam is

$$\psi_{\text{in}}=e^{ikz}$$

The incident flux is $F = \hbar k/m$, and so the total number of particles scattered out of incident beam is $F\sigma=\frac{\hbar k}{m}\sigma$. Therefore, behind the scattering target, there should be a depletion of these particles in the forward direction. To check this conclusion, let us consider the number of particles in the forward direction through an area A See Fig. 15.6. If the scattering target is not present, the flux in the forward direction will remain unchanged; hence, the number particles crossing A should be $\frac{\hbar k}{m}A$. The presence of scattering target will change this number, and our task is to prove that the new number is $\frac{\hbar k}{m}A-\frac{\hbar k}{m}\sigma_T$ explicitly using the Schrdinger wave function.

The outgoing scattered wave is

$$\psi_{\text{out}}=e^{ikz}+f(\theta,\phi)\frac{e^{ikr}}{r}$$

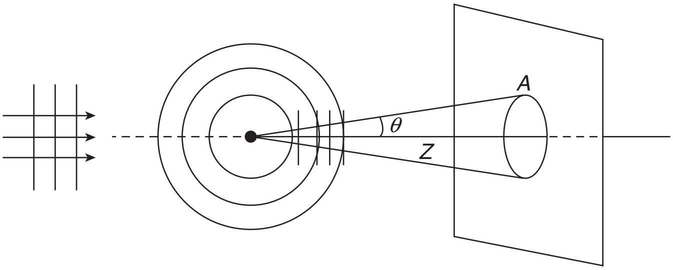


Fig.15.16 Optical theorem

The outgoing wave function in the forward direction is

$$\psi_{\text{out}}=e^{ikz}+\frac{f(0)}{r}e^{ikr}$$

Consider a region A in the plane at z axis. This area A subtends a small angle θ on the target.

$$r = [z^2 + x^2 + y^2]^{1/2} = z\left(1 + \frac{x^2 + y^2}{z^2}\right)^{1/2}$$

$$= z + \frac{1}{2}\left(\frac{x^2 + y^2}{z}\right) \qquad x, y << z$$

$$\therefore \psi_{\text{out}} = e^{ikz} + \frac{f(0)}{z} e^{ikz} \cdot \exp\left(ik\,\frac{(x^2 + y^2)}{2z}\right)$$

$$= e^{ikz}\left[1 + \frac{f(0)}{z}\exp\left(ik\,\frac{(x^2 + y^2)}{2z}\right)\right]$$

The outgoing flux is

$$\mathbf{j}_{\text{out}} = \frac{\hbar}{2im}[\psi^*_{\text{out}} \nabla \psi_{\text{out}} - c \cdot c]$$

$$\nabla \psi_{\text{out}} = (\nabla e^{ikz})\left[1 + \frac{f(0)}{z} e^{ik(x^2+y^2)/2z}\right]$$

$$+ e^{ikz} \nabla\left[1 + \frac{f(0)}{z} e^{ik(x^2+y^2)/2z}\right]$$

Since the term $\frac{f(0)}{z} e^{ik(x^2+y^2)/2z}$ itself is small, we can neglect the derivative of this term.

$$\therefore \nabla \psi_{\text{out}} = (ik) e^{ikz}\left[1 + \frac{f(0)}{z} e^{ik(x^2+y^2)/2z}\right]$$

$$= ik\psi$$

$$\therefore j_{\text{out}} = \frac{\hbar}{2im}[ik|\psi|^2 - c \cdot c] = \frac{\hbar k}{m}|\psi|^2$$

$$|\psi|^2 = \left[1 + \frac{f^*(0)}{z} e^{-ik(x^2+y^2)/2z}\right] \cdot \left[1 + \frac{f(0)}{z} e^{ik(x^2+y^2)/2z}\right]$$

$$= 1 + \frac{f(0)}{z} e^{ik(x^2+y^2)/2z} + \frac{f^*(0)}{z} e^{-ik(x^2+y^2)/2z} + \ldots$$

$$= 1 + \frac{2}{z}\text{Re}[f(0) e^{ik(x^2+y^2)/2z}]$$

Let N_{out} be the number of outgoing particle through area A.

$$N_{\text{out}} = \int_A j_{\text{out}}\, dxdy = \int_A |\psi|^2\, \frac{\hbar k}{m} dxdy$$

$$= \int_A \left[1 + \frac{2}{z}\text{Re}(f(0) e^{ik(x^2+y^2)/2z})\right] \frac{\hbar k}{m}\, dxdy$$

$$= \frac{\hbar k}{m} \cdot A + \frac{2}{z} \cdot \frac{\hbar k}{m} \cdot \text{Re}\left[f(0) \int_A e^{ik(x^2+y^2)/2z} dxdy\right]$$

For large z, A will be sufficiently large enough to make the integral limits to extend from $-\infty$ to ∞.

$$\therefore \int_A e^{ik(x^2+y^2)/2z^2} dx dy = \int_{-\infty}^{\infty} e^{ikx^2/2z} dx \cdot \int_{-\infty}^{\infty} e^{iky^2/2z} dy$$

$$= \frac{\pi}{-ik/2z} = \frac{i2z\pi}{k}$$

$$\begin{aligned}\therefore N_{\text{out}} &= \frac{\hbar k}{m} A + \frac{2}{z} \cdot \frac{2z\pi}{k} \operatorname{Re}[(\operatorname{Re} f(0) + i \operatorname{Im} f(0))(i)] \\ &= \frac{\hbar k}{m} A - \frac{4\pi}{k} \operatorname{Im} f(0) \cdot \frac{\hbar k}{m} \\ &= \frac{\hbar k}{m} A - \sigma_T \cdot \frac{\hbar k}{m}\end{aligned}$$

where
$$\sigma_T = \frac{4\pi}{k} \operatorname{Im} f(0) \tag{15.159}$$

Let us reiterate that $\frac{\hbar k}{m} \cdot A$ is the number of particles that would have crossed the area A in the absence of scattering and $\frac{\hbar k}{m} \sigma_T$ is the number of particles that have been removed from the incident beam. Obviously, this is the number of particles that have been scattered out of the incident beam.

In terms of probability, the probability current density for the particle along z axis is $\frac{\hbar k}{m}$ Behind the target, there is a reduction in the probability current along z axis by a factor $-2/z \operatorname{Re}[f(0)e^{ik(x^2+y^2)/2z}]$. So the optical theorem is an expression of conservation of probability. This reduction can be linked to wave property of the matter. In wave mechanics, everything is discussed only in terms of waves. The reduction in intensity ($|\psi|^2 = 1$ for incident beam) in the forward direction behind the target can be achieved by a destructive interference due to the superposition of the incident wave and the scattered wave. (Note that this is not total destructive interference to make the intensity zero.)

APPENDIX II

Green's Function $-\frac{1}{4\pi}\frac{e^{ik|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|}$

There are many ways for obtaining Green's function from the definition

$$(\nabla^2 + k^2)G(\mathbf{r}, \mathbf{r}) = \delta(\mathbf{r} - \mathbf{r}') \tag{15.160}$$

We choose a particular method which uses the techniques of complex variable. This method illustrates how different boundary conditions lead to different Green's function. This method will help us to find similar Green's function in relativistic quantum field theory where it plays a major role.

To evaluate $G(\mathbf{r}, \mathbf{r}')$, let us consider its Fourier transform $g(\mathbf{s})$.

$$G(\mathbf{r}, \mathbf{r}') = \frac{1}{(2\pi)^3} \int g(\mathbf{s}) e^{i\mathbf{s}\cdot(\mathbf{r}-\mathbf{r}')} d^3\mathbf{s} \tag{15.161}$$

The Dirac delta function $\delta(\mathbf{r} - \mathbf{r}')$ can be written as

$$\delta(\mathbf{r}-\mathbf{r}') = \frac{1}{(2\pi)^3}\int e^{i\mathbf{s}\cdot(\mathbf{r}-\mathbf{r}')} d^3\mathbf{s} \tag{15.162}$$

Substituting (15.161) and (15.162) in the Equation (15.160), we get

$$(\nabla^2+k^2)\cdot\frac{1}{(2\pi)^3}\int g(\mathbf{s})e^{i\mathbf{s}\cdot(\mathbf{r}-\mathbf{r}')}d^3\mathbf{s} = \frac{1}{(2\pi)^3}\int e^{i\mathbf{s}\cdot(\mathbf{r}-\mathbf{r}')}d^3\mathbf{s}$$

$$\int g(s)(-s^2+k^2)e^{i\mathbf{s}\cdot(\mathbf{r}-r')}d^3\mathbf{s} = \int e^{i\mathbf{s}\cdot(\mathbf{r}-\mathbf{r}')}d^3\mathbf{s}$$

$$\therefore g(s) = -\frac{1}{s^2-k^2}$$

$$\therefore G(\mathbf{r},\mathbf{r}') = -\frac{1}{(2\pi)^3}\int\frac{1}{s^2-k^2}e^{i\mathbf{s}\cdot(\mathbf{r}-\mathbf{r}')}d^3\mathbf{r} \tag{15.163}$$

Let us evaluate this integral in spherical polar coordinates.

$$\therefore G(\mathbf{r},\mathbf{r}') = -\frac{1}{(2\pi)^3}\int_0^\infty \frac{s^2}{s^2-k^2}\,ds\int_0^\pi \sin\theta\, d\theta\int_0^{2\pi} d\phi\; e^{i\mathbf{s}\cdot(\mathbf{r}-\mathbf{r}')}$$

$$= -\frac{1}{(4\pi^2)}\cdot\int_0^\infty \frac{s^2}{s^2-k^2}\,ds\int_0^\pi e^{is|\mathbf{r}-\mathbf{r}'|\cos\theta}\sin\theta\, d\theta$$

$$= -\frac{1}{(4\pi^2)}\int_0^\infty \frac{s^2}{s^2-k^2}\,ds\cdot\frac{1}{is\,|\mathbf{r}-\mathbf{r}'|}\left[e^{is|\mathbf{r}-\mathbf{r}'|}-e^{-is|\mathbf{r}-\mathbf{r}'|}\right]$$

Consider the second term $\displaystyle\int_0^\infty \frac{se^{-is|\mathbf{r}-\mathbf{r}'|}}{s^2-k^2}\,ds$

Change the integration variable s to $s'=-s$.

$$\int_0^\infty \frac{s}{s^2-k^2}\,e^{-is|\mathbf{r}-\mathbf{r}'|}ds \to \int_0^{-\infty} e^{is'|\mathbf{r}-\mathbf{r}'|}\,\frac{(-s')}{s'^2-k^2}(-ds')$$

$$= \int_{-\infty}^{0} e^{is|\mathbf{r}-\mathbf{r}'|}\,\frac{s}{s^2-k^2}\,ds$$

(We have used the fact that s' is a dummy variable and so it is replaced by s in the last step.)

$$\therefore G(\mathbf{r},\mathbf{r}') = \frac{i}{4\pi^2}\left\{\int_0^\infty \frac{s}{s^2-k^2}\,e^{is|\mathbf{r}-\mathbf{r}'|}ds + \int_{-\infty}^{0} \frac{s}{s^2-k^2}\,e^{is|\mathbf{r}-\mathbf{r}'|}ds\right\}\frac{1}{|\mathbf{r}-\mathbf{r}'|}$$

$$= \frac{i}{4\pi^2}\cdot\frac{1}{|\mathbf{r}-\mathbf{r}'|}\int_{-\infty}^{\infty} \frac{s}{s^2-k^2}\,e^{is|\mathbf{r}-\mathbf{r}'|}ds \tag{15.164}$$

This has a singularity at $s=\pm k$. Let us evaluate this integral by making s a complex variable. s now becomes $s = s_R + is_i$, where s_R and s_i are real numbers. Then, the above integral becomes a contour integral.

$$\int_{-\infty}^{\infty} \frac{s}{s^2-k^2}\, e^{is|\mathbf{r}-\mathbf{r}'|} ds \rightarrow \int_{-\infty}^{\infty} \frac{s}{s^2-k^2}\, e^{is|\mathbf{r}-\mathbf{r}'|} ds + \int_{\Gamma} \frac{s e^{is|\mathbf{r}-\mathbf{r}'|}}{s^2-k^2}\, ds \tag{15.165}$$

$$= \oint \frac{s}{s^2-k^2}\, e^{is|\mathbf{r}-\mathbf{r}'|} ds \tag{15.166}$$

We have to choose the contour in such a way that (i) the correct pole for our purpose is included and (ii) the line integral $\int_{\Gamma} \rightarrow 0$ for large s. The convergence of the integral decides the curve Γ to be a semicircle in the upper half plane. This can be easily seen as follows:

$$\begin{aligned} e^{is|\mathbf{r}-\mathbf{r}'|} &= e^{i(s_R+is_i)|\mathbf{r}-\mathbf{r}'|} \\ &= e^{is_R|\mathbf{r}-\mathbf{r}'|} e^{-s_i|\mathbf{r}-\mathbf{r}'|} \end{aligned}$$

For large s on the contour Γ, the factor

$$e^{-s_i|\mathbf{r}-\mathbf{r}'|} \rightarrow 0 \text{ and so } \int_{\Gamma} \rightarrow 0.$$

As far as the inclusion of the poles is concerned, we can include the pole at $s = k$ or $s = -k$ or both. Each of these choice leads to different Green's function and hence different solutions $u(\mathbf{r})$.

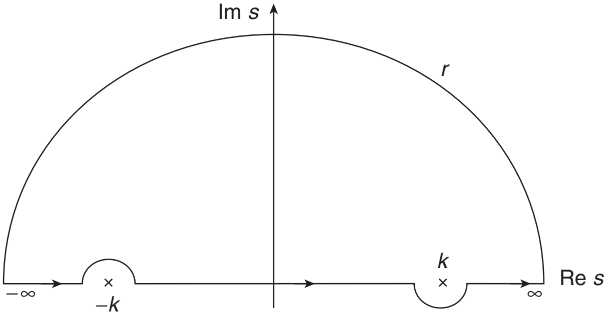


Fig.15.17 Contour for Green's function

It is true that there are many possible solutions for Schrödinger equation. But in scattering theory, we look for solutions which will lead to the asymptotic behaviour given in Equation (15.9). This criterion leads to the choice of contour shown in Fig.15.17, which includes the pole at $s = k$.

$$\begin{aligned} G(\mathbf{r},\mathbf{r}') &= \frac{i}{4\pi^2} \oint \frac{s}{(s+k)(s-k)} \frac{e^{is|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|} ds \\ &= -\frac{1}{4\pi} \cdot \frac{e^{ik|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|} \end{aligned} \tag{15.167}$$

EXERCISES

1. What is meant by scattering cross section?
2. What is the difference between cross sections in classical and quantum mechanics?
3. Are the scattering cross section and the geometrical cross section the same?

4. The geometrical cross section of a nucleus is $\pi r_0^2 A^{2/3}$. Assuming $r_0 = 1.4$ F and $A = 197$ for gold atom, estimate the geometrical cross section. Can scattering cross section be greater than this number?
5. What is meant by optical theorem?
6. How will you interpret optical theorem in terms of wave properties?
7. What is meant by phase shift?
8. What is meant by scattering amplitude?
9. Is there a difference between scattering amplitude and scattering length?
10. Does the radial wave function for *s* wave in the low-energy scattering become zero at $r = \alpha$ where α is the scattering length?
11. What is meant by resonant scattering?
12. What is meant by zero energy bound state and zero energy scattering state?
13. The interaction potential between the incident and the target particles is $V(r) = V_0 e^{-r^2/2a^2}$. Show that $\frac{d\sigma}{d\Omega}$ is the first-order approximation in

$$\frac{d\sigma}{d\Omega} = \frac{2\pi\mu^2 a^6 V_0^{\,2}}{\hbar^4} e^{-k^2a^2}$$

14. The interaction potential between the incident particle and the target particles is

$$V(r) = \begin{cases} V_0 & r \le a \\ 0 & r \ge a \end{cases}$$

Show that
$$\frac{d\sigma}{d\Omega} = \frac{4\mu^2 V_0^{\,2}}{\hbar^4 k^4}\left[a\cos Ka - \frac{\sin ka}{K}\right]^2$$

REFERENCES

1. S. Gasiorowicz, 1974. *Quantum Physics*. New York: John Wiley & Sons.
2. D. Park, 1974. *Introduction to Quantum Theory*, Tokyo: McGraw-Hill Kogakusha Ltd.
3. A. S. Davydov, 1988. *Quantum Mechanics*, Oxford: Pergamon Press.
4. En. Wikipedia org/wiki/optical theorem.
5. http://www.phys.ufl.edu/~pjh/teaching/phys 4605/notes/scatter1.pdf.
6. R. Fitzpatrick-Farside.ph.utexas.edu/teaching/qmech/qmech.pdf.
7. J. J. Sakurai, 1994. *Modern Quantum Mechanics*, Reading Massachusetts: Addison Wesley Publishing Company Inc.
8. W. N. Cottingham and D. A. Greenwood, 2001. *An Introduction to Nuclear Physics*, Cambridge: Cambridge University Press.

16

Relativistic Wave Equation

The Schrödinger equation is non-relativistic wave equation. Non-relativistic and relativistic domains are generally determined by the velocity of the particle. If v/c is much less than 1, we call it non-relativistic domain, and if v/c is closer to 1, then it is called relativistic domain. Within quantum mechanics, whether a system is non-relativistic or relativistic is determined by the fact that whether the non-relativistic Hamiltonian or relativistic Hamiltonian is quantized; i.e., whether the transition from classical mechanics to quantum mechanics is done through non-relativistic or relativistic Hamilton.

Irrespective of the velocity of the particle, there is a compelling necessity to combine quantum mechanics and special theory of relativity. This comes from the fact that all the laws of physics should be Lorentz covariant. This is a very strong requirement.

The non-relativistic quantum mechanics is based on a set of postulates. The special theory of relativity and the quantum mechanics have to be combined in a way consistent with those postulates. An attempt in this direction ended up with a set of relativistic wave equations and quantum field theory.

In non-relativistic quantum mechanics, we have only one equation namely, the Schrödinger equation to determine the time evolution of the state $\psi(\mathbf{r},t)$ of the system. In relativistic quantum mechanics, we have many equations for $\psi(\mathbf{r},t)$, each describing a particular category of particles depending on spin angular momentum. Let us mention here some of these equations.

1. Klien Gordan (KG) equation for spin zero particles
2. Dirac equation for spin $1/2$ particles
3. Maxwell's equation for photon

Here, we concentrate only on the Klien Gordan and Dirac equations. The Klien Gordan equation and the Dirac equation have been developed in this chapter in the same way normally presented in a textbook at this level. First, the arguments leading to the formulation of the Klien Gordan equation and its 'shortcomings' are presented. Then, the Dirac equation is developed as an alternative to the Klien Gordan equation. This sounds as if the Klien Gordan equation is wrong. We have to mention that both the equations are valid equations, but they describe different kinds of systems, and both of them have to be reformulated in terms of quantum fields.

16.1 KLIEN GORDAN EQUATION

In non-relativistic quantum mechanics, the classical Hamiltonian $H = p^2/2m + V$ is the starting point. To make a transition to quantum mechanics, we write the wave equation for ψ as

$$E\psi = H\psi \tag{16.1}$$

where E is replaced by $E = i\hbar\, \partial/\partial t$ and H is replaced by H_{op}, which, in turn, is obtained by replacing $\mathbf{r}$ by $\mathbf{r}$ and $\mathbf{p}$ by $-i\hbar\nabla$. The Hamiltonian operator H_{op} is given by $H_{op} = -\hbar^2/2m\nabla^2 + V(\mathbf{r},t)$. So the equation becomes

$$i\hbar\frac{\partial\psi}{\partial t} = \left[-\frac{\hbar^2}{2m}\nabla^2 + V(\mathbf{r},t)\right]\psi \tag{16.2}$$

which is the Schrödinger equation.

Let us extend the same method to relativistic case also to obtain the relativistic wave equation. This is not a derivation but a set of arguments leading to the formulation of Klien Gordan equation. The classical Hamiltonian for the relativistic free particle is $H = \sqrt{p^2c^2 + m^2c^4}$. To make the transition to quantum mechanics, the wave equation is now written as

$$E\psi = H\psi$$

or

$$i\hbar\frac{\partial}{\partial t}\psi = [-c^2\hbar^2\nabla^2 + m^2c^4]^{1/2}\psi \tag{16.3}$$

What is meant by the operator in the RHS of the Equation (16.3)? What is the effect of this operator on a wave function (for instance, the effect of d/dx on sin kx is to produce k cos kx)?. It is very difficult to interpret the partial derivatives inside a square root. To circumvent this problem, let us go to a different equation where partial derivatives inside a square root is not involved.

We can achieve this result as follows. In classical physics, we have

$$E^2 = p^2c^2 + m^2c^4$$

So the quantum mechanical wave equation is written as

$$E^2\psi = (p^2c^2 + m^2c^4)\psi$$

Replacing E by $i\hbar\dfrac{\partial}{\partial t}$ and $\mathbf{p}$ by $-i\hbar\nabla$, we get

$$\left(i\hbar\frac{\partial}{\partial t}\right)\left(i\hbar\frac{\partial}{\partial t}\right)\psi = [c^2(-i\hbar\nabla)\cdot(-i\hbar\nabla) + m^2c^4]\psi$$

or

$$-\hbar^2\frac{\partial^2\psi}{\partial t^2} = [-c^2\hbar^2\nabla^2 + m^2c^4]\psi \tag{16.4}$$

This is the well-known Klien Gordan equation. This equation can be rewritten in many ways. Dividing throughout by $-c^2\hbar^2$, we get

$$\frac{1}{c^2}\frac{\partial^2\psi}{\partial t^2} = \left[\nabla^2 - \frac{m^2c^2}{h^2}\right]\psi$$

or

$$\frac{1}{c^2}\frac{\partial^2\psi}{\partial t^2} - \nabla^2\psi + \frac{m^2c^2}{\hbar^2}\psi = 0$$

i.e.,

$$\left(\frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \nabla^2\right)\psi + \frac{m^2c^2}{\hbar^2}\psi = 0 \tag{16.5}$$

16.2 FREE PARTICLE SOLUTION TO KG EQUATION

Having constructed a wave equation, we have to find the solutions describing different systems. The simplest system is the free particle, and so let us analyze the case of free particle using Klien Gordan equation. For free particles, the Klien Gordan equation is

$$-\hbar^2 \frac{\partial^2 \psi}{\partial t^2} = (-\hbar^2 c^2 \nabla^2 + m^2 c^4)\psi$$

We have to construct a solution which describes the free particle in a way consistent with the known physics. From physics considerations, we expect the free particle to have definite momentum **p** and energy E. Therefore, it should be an eigenfunction of momentum operator $-i\hbar\nabla$. Then, $\psi = Ne^{i(\mathbf{p}.\mathbf{r}-Et)/\hbar}$ will be the appropriate wave function to describe a free particle. This is also consistent with De Broglie wave property of a free particle. So let us try $\psi = Ne^{i(\mathbf{p}.\mathbf{r}-Et)/\hbar}$ as a solution to the Klien Gordan equation. We have to check whether this function satisfies the KG equation.

LHS of KG equation

$$-\hbar^2 \frac{\partial^2 \psi}{\partial t^2} = -\hbar^2 N\, e^{i(\mathbf{p}.\mathbf{r}-Et)/\hbar} \left(-\frac{iE}{\hbar}\right)^2 = E^2 N e^{i(\mathbf{p}.\mathbf{r}-Et)/\hbar}$$

RHS of KG equation

$$\nabla^2 \psi = N\nabla^2 e^{i(\mathbf{p}.\mathbf{r}-Et)/\hbar} = N\left(\frac{i\mathbf{p}}{\hbar}\right)^2 e^{i(\mathbf{p}.\mathbf{r}-Et)/\hbar} = -N\frac{p^2}{\hbar^2} e^{i(\mathbf{p}.\mathbf{r}-Et)/\hbar}$$

$$\therefore -\hbar^2 c^2 \nabla^2 \psi + m^2 c^4 \psi = N(p^2c^2 + m^2c^4)e^{i(\mathbf{p}.\mathbf{r}-Et)/\hbar}$$

The wave function $\psi = Ne^{i(\mathbf{p}\cdot\mathbf{r}-Et)/\hbar}$ is solution to the Equation (16.5) if LHS and RHS are equal.

$$\therefore E^2 N e^{i(\mathbf{p}.\mathbf{r}-Et)/\hbar} = N(p^2c^2 + m^2c^4)e^{i(\mathbf{p}.\mathbf{r}-Et)/\hbar}$$

This is possible provided $E^2 = p^2c^2 + m^2c^4$. Taking the square root of E^2, we get E to be $E = \pm\sqrt{p^2c^2 + m^2c^4}$. So, for a free particle obeying KG equation, we have two energy eigenvalues $E = +\sqrt{p^2c^2 + m^2c^4}$ and $E = -\sqrt{p^2c^2 + m^2c^4}$.

Let us define E_p as $E_p = \sqrt{p^2c^2 + m^2c^2}$ where E_p is a positive quantity by definition. In terms of E_p, the possible energy eigenvalues are $E = E_p$ and $E = -E_p$. So, for a free particle KG equation, we have two possible solutions:

Solution I: $\psi_{\mathrm{I}} = Ne^{i(\mathbf{p}.\mathbf{r}-E_p t)/\hbar}$. This represents a particle with momentum **p** and energy $E = E_p$.

Solution II: $\psi_{\mathrm{II}} = Ne^{i(\mathbf{p}.\mathbf{r}+E_p t)/\hbar}$. This represents a particle with momentum **p** and energy $E = -E_p$. The second solution represents a free particle with negative energy $-E_p$.

16.3 PROBLEM WITH NEGATIVE ENERGY SOLUTION

The negative value for energy of a free particle causes considerable difficulties. From the structure of quantum mechanics, we know that whenever we make a measurement of a dynamical variable A, the possible experimental outcome has to be one among the eigenvalues of the operation A_{op}. If the Klien Gordan is true for a relativistic particle, experimentally, we should observe a free particle with negative energy. But, experimentally, no such particle has been observed. Since no such particle is found experimentally, the existence of negative energy solution poses a serious problem. We cannot simply dismiss the negative energy solution as unphysical. The structure of quantum mechanics does not have a provision to take some subset of the eigenvalues as physical and disallow remaining subset of eigenvalues as unphysical.

In fact, one cannot leave out any particular eigenstate since all the eigenstates are required to form a complete set. This is very much required, since any general solution to Klien Gordan equation can be expressed as a linear combination of eigenstates only if all the eigenstates of a complete set are available.

16.4 PROBABILITY INTERPRETATION AND KG EQUATION

In non-relativistic quantum mechanics, from the Schrödinger equation, we can derive the equation of continuity given by

$$\frac{\partial}{\partial t}(\psi^*\psi)+\nabla\cdot\left[\frac{\hbar}{2mi}(\psi^*\nabla\psi-\psi\nabla\psi^*)\right]=0 \tag{16.6}$$

This equation is a consequence of the fact that ψ satisfies the Schrödinger equation.

The expression $\psi^*\psi$ is interpreted as probability density $\rho(\mathbf{r},t)$. This is consistent with (i) the mathematical requirement that the expression for the probability should always be positive and (ii) the requirement from the physics consideration that the total probability is a conserved quantity

$$\left(\int\psi^*\psi d\tau=\int P(\mathbf{r},t)d\tau \text{ is independent time}\right)$$

In relativistic quantum mechanics, ψ satisfies a different differential equation namely, KG equation. We would like to derive the equation of continuity as a consequence of KG equation and subsequently identify suitable mathematical expressions as the probability density $\rho(\mathbf{r},t)$ and probability current density $\mathbf{S}(\mathbf{r},t)$.

The KG equation is

$$-\hbar^2\frac{\partial^2\psi}{\partial t^2}=-c^2\hbar^2\nabla^2\psi+m^2c^4\psi \tag{16.7}$$

Its complex conjugate is

$$-\hbar^2\frac{\partial^2\psi^*}{\partial t^2}=-c^2\hbar^2\nabla^2\psi^*+m^2c^4\psi^* \tag{16.8}$$

Multiply the Equation (16.7) by ψ^* and the Equation (16.8) by ψ to get

$$-\hbar^2\psi^*\frac{\partial^2\psi}{\partial t^2}=-c^2h^2\psi^*\nabla^2\psi+m^2c^4\psi^*\psi \tag{16.9}$$

and

$$-\hbar^2 \frac{\partial^2 \psi^*}{\partial t^2}\psi = -c^2h^2\nabla^2\psi^*\psi + m^2c^4\psi^*\psi \tag{16.10}$$

Subtracting the Equation (16.9) from the Equation (16.10), we get

$$\hbar^2\left[\psi^*\frac{\partial^2\psi}{\partial t^2} - \frac{\partial^2\psi^*}{\partial t^2}\psi\right] = c^2\hbar^2[\psi^*\nabla^2\psi - \nabla^2\psi^*\psi]$$

Dividing throughout by $-2imc^2\hbar$, we get

$$\frac{\hbar}{2imc^2}\left[\frac{\partial^2\psi^*}{\partial t^2}\psi - \psi^*\frac{\partial^2\psi}{\partial t^2}\right] = \frac{\hbar}{2im}[\nabla^2\psi^*\psi - \psi^*\nabla^2\psi]$$

$$\frac{\hbar}{2imc^2}\left[\frac{\partial^2\psi^*}{\partial t^2}\psi - \psi^*\frac{\partial^2\psi}{\partial t^2}\right] + \frac{\hbar}{2im}[\psi^*\nabla^2\psi - \nabla^2\psi^*\psi] = 0$$

$$\frac{\hbar}{2imc^2}\frac{\partial}{\partial t}\left[\frac{\partial\psi^*}{\partial t}\psi - \psi^*\frac{\partial\psi}{\partial t}\right] + \frac{\hbar}{2im}\nabla\cdot[\psi^*\nabla\psi - \nabla\psi^*\psi] = 0$$

This equation has the structure

$$\frac{\partial\rho}{\partial t} + \nabla\cdot\mathbf{S} = 0$$

provided we identify ρ and $\mathbf{S}$ as

$$\rho(\mathbf{r},t) = \frac{\hbar}{2imc^2}\left[\frac{\partial\psi^*}{\partial t}\psi - \psi^*\frac{\partial\psi}{\partial t}\right] \tag{16.11}$$

and

$$\mathbf{S}(\mathbf{r},t) = \frac{h}{2im}[\psi^*\nabla\psi - \nabla\psi^*\psi] \tag{16.12}$$

The second equation is exactly the same as the expression for probability current in non-relativistic quantum mechanics. This fact immediately suggests that the RHS of the Equation (16.11) to be interpreted as probability density.

However, this expression has two terms, of which one term is subtracted from the other. Depending on the values of the terms $\psi^*\partial\psi/\partial t$ and $\psi\,\partial\psi^*/\partial t$, the expression for ρ may be positive or negative. Hence, ρ is not a positive definite quantity. Since probability density has to be always a positive quantity, this expression cannot be interpreted as probability density.

In fact, for free particles, substitution of the wave function $\psi = Ne^{i(\mathbf{p}.\mathbf{r}-Et)/\hbar}$ in the expression for ρ leads to

$$\rho = \frac{\hbar}{2imc^2}\left[Ne^{i(\mathbf{p}.\mathbf{r}-Et)/\hbar}N^*\left(\frac{iE}{\hbar}\right)e^{-i(\mathbf{p}.\mathbf{r}-Et)/\hbar} - N^*e^{-i(\mathbf{p}.\mathbf{r}-Et)/\hbar}N\left(\frac{-iE}{\hbar}\right)e^{i(\mathbf{p}.\mathbf{r}-Et)/\hbar}\right]$$

$$= \frac{\hbar}{2imc^2}|N|^2\frac{2iE}{\hbar} = \frac{|N|^2 E}{mc^2}$$

For positive energy particle $E = E_P = \sqrt{p^2c^2 + m^2c^4}$, and $\rho = |N|^2 E_p/mc^2$ is a positive quantity.

For negative energy particle $E = -E_P = -\sqrt{p^2c^2 + m^2c^4}$, and so $\rho = -|N|^2 E_p/mc^2$ is a negative quantity.

This brings the second problem of Klien Gordan equation namely, the expression for probability density is not consistent with the requirement that the probability density should be positive definite.

The problems regarding the existence of negative energy eigenstates for a free particle and the interpretation of the expression for ρ are solved only in quantum field theory in which the negative energy particle is shown to be equivalent to positive energy antiparticle and ρ is not interpreted as the probability density to find the position of a particle, but it is interpreted as the number density operator in Fock space for particle and antiparticles.

16.5 KLIEN GORDAN EQUATION AND HYDROGEN ATOM

Having analyzed the consequences of Klien Gordan equation for free particle, let us now get the bound-state energy levels for hydrogen atom. Let us recall that one of the successes of Schrödinger equation is the derivation of Bohr energy levels correctly. However, the fine structure of the hydrogen atom could not be accounted well. If the Klien Gordan equation works well for hydrogen atom, then we should get the Bohr energy level as well as the fine structure splitting of these levels correctly.

In the hydrogen atom, the electron is in the Coulomb potential due to charge *Ze* of the nucleus. The potential energy of the electron is given by $V = (-e)(Ze)/r = -Ze^2/r$.

In non-relativistic quantum mechanics, the Schrödinger equation for hydrogen atom is

$$\left(-\frac{\hbar^2}{2m}\nabla^2 - \frac{Ze^2}{r}\right)u(r,\theta,\Phi) = Eu(r,\theta,\Phi)$$

This equation has be solved subject to the boundary condition $u(r, \theta, \phi). \to 0$ as $r \to \infty$

Let us define the fine structure constant α as $\alpha = e^2/\hbar c$. From the Schrödinger equation, the energy eigenvalues of the hydrogen atom are found to be $E_n = -mc^2Z^2\alpha^2/2n^2$.

Let us now solve Klien Gordan equation for hydrogen atom. The Klien Gordan equation for an electron of charge $-e$ in the electromagnetic field is obtained by replacing E by $E + e\Phi$ and $\mathbf{p}$ by $\mathbf{p} + e\mathbf{A}/c$ in the free particle equation $E^2\psi = (p^2c^2 + m^2c^4)\psi$.

In the operator form, this amounts to replace $i\hbar\, \partial/\partial t$ by $(i\hbar\, \partial/\partial t + e\Phi)$ and $-i\hbar\nabla$ by $\left(-i\hbar\nabla + \frac{e\mathbf{A}}{c}\right)$. So the Klien Gordan equation for hydrogen atom is

$$\left(i\hbar\frac{\partial}{\partial t} + e\Phi\right)^2 \psi = c^2\left(-i\hbar\nabla + \frac{e\mathbf{A}}{c}\right)^2 \psi + m^2c^4\psi$$

For hydrogen or hydrogen-like atoms, $\mathbf{A} = 0$ and $\Phi = Ze/r$. Therefore, we have

$$\left(i\hbar\frac{\partial}{\partial t} + \frac{Ze^2}{r}\right)\psi = -c^2\hbar^2\nabla^2\psi + m^2c^4\psi \tag{16.13}$$

Let us write ψ as

$$\psi = e^{-iEt/\hbar}u(\mathbf{r}) \tag{16.14}$$

Then, LHS of the Equation (16.13) is

$$\left(i\hbar\frac{\partial}{\partial t}+\frac{Ze^2}{r}\right)\left(i\hbar\frac{\partial}{\partial t}+\frac{Ze^2}{r}\right)e^{-iEt/\hbar}u(\mathbf{r})=\left(i\hbar\frac{\partial}{\partial t}+\frac{Ze^2}{r}\right)\left(E+\frac{Ze^2}{r}\right)u(\mathbf{r})e^{-iEt/\hbar}$$

$$=\left(E+\frac{Ze^2}{r}\right)^2 u(\mathbf{r})e^{-iEt/\hbar}$$

RHS of the Equation (16.13) is

$$(-c^2\hbar^2\nabla^2u+m^2c^4u)e^{-iEt/\hbar}$$

Equating LHS and RHS, we get

$$\left(E+\frac{Ze^2}{r}\right)^2 u=-c^2\hbar^2\nabla^2u+m^2c^4u$$

Let us rewrite the above equation as

$$\nabla^2u-\frac{m^2c^2}{\hbar^2}u+\frac{1}{c^2\hbar^2}\left(E+\frac{Ze^2}{r}\right)^2 u=0 \tag{16.15}$$

Let us write the solution u as

$$u(\mathbf{r})=u(r,\theta,\Phi)=R(r)Y_{lm}(\theta,\varphi)$$

Let us recall that

$$L^2Y_{lm}(\theta,\varphi)=l(l+1)\hbar^2Y_{lm}(\theta,\varphi)$$

and

$$\nabla^2=\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{d}{dr}\right)-\frac{L^2}{r^2\hbar^2}$$

$$\therefore \nabla^2u=\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{d}{dr}u\right)-\frac{L^2u}{r^2\hbar^2}=\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{dR}{dr}\right)Y_{lm}-\frac{RL^2Y_{lm}}{r^2\hbar^2}$$

$$=\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{dR}{dr}\right)Y_{lm}-\frac{l(l+1)}{r^2}RY_{lm} \tag{16.16}$$

Making use of the Equation (16.16) in the Equation (16.15), we have

$$\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{dR}{dr}\right)Y_{lm}-\frac{l(l+1)}{r^2}RY_{lm}-\frac{m^2c^2}{\hbar^2}RY_{lm}+\frac{1}{c^2\hbar^2}\left(E+\frac{Ze^2}{r}\right)^2 RY_{lm}=0$$

Factoring out Y_{lm} and rearranging different terms, we get

$$\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{dR}{dr}\right)+\frac{1}{\hbar^2c^2}[E^2-m^2c^4]R+\frac{2EZe^2}{c^2\hbar^2 r}R-\frac{l(l+1)-Z^2e^4/c^2\hbar^2}{r^2}R=0 \tag{16.17}$$

Let us define s as

$$s(s+1)=l(l+1)-\frac{Z^2e^4}{\hbar^2c^2} \tag{16.18}$$

The fine structure constant is $\alpha = e^2/\hbar c$, and its value is 1/137. Therefore, $s(s+1)/r^2$ is almost same as the term $l(l+1)/r^2$ in the radial equation for hydrogen atom in the non-relativistic quantum mechanics. Therefore, s should be a positive quantity.

Let us rewrite the Equation (16.18) as

$$s^2+s-l^2-l+Z^2\,\alpha^2=0$$

This is a quadratic equation in s. Solving this equation, we get

$$s=\frac{-1\pm\sqrt{1+4l^2+4l-4Z^2\alpha^2}}{2}\ \frac{-1\pm\sqrt{(2l+1)^2-4Z^2\alpha^2}}{2} \tag{16.19}$$

Since s has to be positive, we take only one root, and so we have

$$s=-\frac{1}{2}+\frac{1}{2}\left[(2l+1)^2-4Z^2\alpha^2\right]^{1/2}=-\frac{1}{2}+\frac{1}{2}2\left[\left(l+\frac{1}{2}\right)^2-Z^2\alpha^2\right]^{1/2}$$

$$=-\frac{1}{2}+\left[\left(l+\frac{1}{2}\right)^2-Z^2\alpha^2\right]^{1/2} \tag{16.20}$$

It terms of s, the Equation (16.17) is written as

$$\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{dR}{dr}\right)+\left[\frac{E^2-m^2c^4}{\hbar^2c^2}+\frac{2EZe^2}{\hbar^2c^2 r}-\frac{s(s+1)}{r^2}\right]R=0 \tag{16.21}$$

Note that for hydrogen atom, the bound state energy E is less than the free particle energy mc^2. Therefore, $(E^2-m^2c^4)$ is negative quantity.

Define $\rho=\chi r$ where χ is an arbitrary constant, which can be chosen to suit our convenience. Then, the Equation (16.21) becomes

$$\frac{\chi^2}{\rho^2}\frac{d}{d\rho}\left[\rho^2\frac{dR}{d\rho}\right]+\left[\frac{E^2-m^2c^4}{\hbar^2c^2}+\frac{2EZe^2}{\hbar^2c^2\rho}\chi-\frac{s(s+1)\chi^2}{\rho^2}\right]R=0$$

Dividing throughout by χ^2, we get

$$\frac{1}{\rho^2}\frac{d}{d\rho}\left[\rho^2\frac{dR}{d\rho}\right]+\left[\frac{E^2-m^2c^4}{\hbar^2c^2\chi^2}+\frac{2EZe^2}{\hbar^2c^2\rho\chi}-\frac{s(s+1)}{\rho^2}\right]R=0 \tag{16.22}$$

Let us recall that the non-relativistic radial equation for hydrogen atom is

$$\frac{1}{\rho^2}\frac{d}{d\rho}\left[\rho^2\frac{dR}{d\rho}\right]+\left[\frac{\lambda}{\rho}-\frac{1}{4}-\frac{l(l+1)}{\rho^2}\right]R=0 \tag{16.23}$$

The Equation (16.22) will have exactly the same structure as that of the Equation (16.23), provided we have

$$\frac{E^2-m^2c^4}{\hbar^2c^2\chi^2}=-\frac{1}{4} \text{ and } \frac{2EZe^2}{\hbar^2c^2\chi}=\lambda$$

So we have

$$\chi=\frac{2}{\hbar c}(m^2c^4-E^2)^{1/2} \text{ and } \lambda=\frac{2EZe^2}{\hbar^2c^2\chi}=\frac{2EZ\alpha}{\hbar c\chi} \tag{16.24}$$

The Equation (16.22) now becomes

$$\frac{1}{\rho^2}\frac{d}{d\rho}\left[\rho^2\frac{dR}{d\rho}\right]+\left[\frac{\lambda}{\rho}-\frac{1}{4}-\frac{s(s+1)}{p^2}\right]R=0 \tag{16.25}$$

This equation has to be solved, subject to the boundary condition $R\,(r)\to 0$ as $r\to\infty$.

Working with a series of solution as done in the non-relativistic case, it can be shown that this boundary condition is satisfied, provided

$$\lambda=n'+s+1 \text{ with } n'=0, 1, 2\ldots$$

$$=n'-\frac{1}{2}+\left[\left(l+\frac{1}{2}\right)^2-Z^2\alpha^2\right]^{1/2}+1$$

$$=n'+\frac{1}{2}+\left[\left(l+\frac{1}{2}\right)^2-Z^2\alpha^2\right]^{1/2} \tag{16.26}$$

Expanding in powers of α, we can write

$$\lambda=n'+\frac{1}{2}+\left(l+\frac{1}{2}\right)\left[1-\frac{Z^2\alpha^2}{\left(l+\frac{1}{2}\right)^2}\right]^{1/2}$$

$$=n'+\frac{1}{2}+\left(l+\frac{1}{2}\right)\left[1-\frac{Z^2\alpha^2}{2\left(l+\frac{1}{2}\right)^2}-\frac{Z^4\alpha^4}{8\left(l+\frac{1}{2}\right)^4}+\ldots.\right]$$

$$=n'+\frac{1}{2}+l+\frac{1}{2}-\frac{Z^2\alpha^2}{2\left(l+\frac{1}{2}\right)}-\frac{Z^4\alpha^4}{8\left(l+\frac{1}{2}\right)^3}$$

$$= n' + l + 1 - \frac{Z^2\alpha^2}{2\left(l+\frac{1}{2}\right)} - \frac{Z^4\alpha^4}{8\left(l+\frac{1}{2}\right)^3} \tag{16.27}$$

Put $n' + l + 1 = n$. The possible values of n are $n = 1, 2, 3\ldots$

$$\therefore \lambda = n - \frac{Z^2\alpha^2}{2\left(l+\frac{1}{2}\right)} - \frac{Z^4\alpha^4}{8\left(l+\frac{1}{2}\right)^3} \tag{16.28}$$

Let us rewrite this equation as follows:

$$\lambda = n\left(1 - \frac{Z^2\alpha^2}{2n\left(l+\frac{1}{2}\right)} - \frac{Z^4\alpha^4}{8n\left(l+\frac{1}{2}\right)^3}\right) = n(1 - a\alpha^2 - b\alpha^4) \tag{16.29}$$

where $a = \dfrac{Z^2}{2n\left(l+\frac{1}{2}\right)}$ and $b = \dfrac{Z^4}{8n\left(l+\frac{1}{2}\right)^3}$

From the Equation (16.24), we have

$$\lambda^2 = \frac{4E^2Z^2\alpha^2}{\hbar^2c^2\chi^2} = \frac{E^2Z^2\alpha^2}{m^2c^4 - E^2}$$

$$\therefore E^2(Z^2\alpha^2 + \lambda^2) = m^2c^4\lambda^2 \tag{16.30}$$

$$E^2 = \frac{m^2c^4\lambda^2}{\lambda^2 + Z^2\alpha^2} = \frac{m^2c^4}{\left(1 + \frac{Z^2\alpha^2}{\lambda^2}\right)}$$

or

$$E = mc^2\left(1 + \frac{Z^2\alpha^2}{\lambda^2}\right)^{-1/2} \tag{16.31}$$

Let us now evaluate $Z^2\alpha^2/\lambda^2$ using binomial expansion up to the order α^4. Using the Equation (16.29), we get

$$\begin{aligned}
\frac{Z^2\alpha^2}{\lambda^2} &= \frac{Z^2\alpha^2}{n^2}(1 - a\alpha^2 - b\alpha^4)^{-2} \\
&= \frac{Z^2\alpha^2}{n^2}(1 + 2(a\alpha^2 + b\alpha^4) + 3(a\alpha^2 + b\alpha^4)^2 + \cdots) \\
&= \frac{Z^2\alpha^2}{n^2}(1 + 2(a\alpha^2 + b\alpha^4) + 3a^2\alpha^4 + \cdots) \\
&= \frac{Z^2\alpha^2}{n^2} + \frac{2aZ^2\alpha^4}{n^2} + \cdots
\end{aligned} \tag{16.32}$$

Using the Equation (16.32) in the Equation (16.31), let us write E as

$$
\begin{aligned}
E &= mc^2\left(1+\frac{Z^2\alpha^2}{n^2}+\frac{2aZ^2\alpha^4}{n^2}\right)^{-1/2} \\
&= mc^2\left(1-\frac{1}{2}\left(\frac{Z^2\alpha^2}{n^2}+\frac{2aZ^2\alpha^4}{n^2}\right)+\frac{3}{8}\left(\frac{Z^2\alpha^2}{n^2}+\frac{2aZ^2\alpha^4}{n^2}\right)^2+\cdots\right) \qquad (16.33) \\
&= mc^2\left(1-\frac{Z^2\alpha^2}{2n^2}-\frac{aZ^2\alpha^4}{n^2}+\frac{3Z^2\alpha^4}{8n^4}+\cdots\right)
\end{aligned}
$$

$$
= mc^2\left(1-\frac{Z^2\alpha^2}{2n^2}-\frac{Z^2\alpha^4}{n^2}\cdot\frac{Z^2}{2n\left(l+\frac{1}{2}\right)}+\frac{3Z^2\alpha^4}{8n^4}+\cdots\right)
$$

$$
\therefore E = mc^2\left[1-\frac{Z^2\alpha^2}{2n^2}-\frac{Z^4\alpha^4}{2n^4}\left[\frac{n}{\left(l+\frac{1}{2}\right)}-\frac{3}{4}\right]\right]
$$

$$
= mc^2 - mc^2\frac{Z^2\alpha^2}{2n^2}-\frac{Z^4\alpha^4mc^2}{2n^4}\left[\frac{n}{l+\frac{1}{2}}-\frac{3}{4}\right] \qquad (16.34)
$$

The first term is the rest mass energy of the free electron. This is consistent with special theory of relativity. The next two terms are the consequences of interaction between the electron and the nucleus. The overall decrease in energy from mc^2 indicates that a bound state has been formed.

The second term is exactly the same as the expression obtained in the non-relativistic case. The third term is proportioned to $\alpha^4 (\alpha \sim 1/137)$, and so it is very small compared to the second term. This is correction to the energy levels obtained in non-relativistic quantum mechanics. For a given n, this correction depends on l. The non-relativistic energy levels are spilt into fine levels as shown in Fig.16.1.

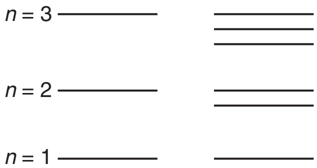


Fig.16.1 Energy levels for hydrogen atom

The sub-levels lead to fine structure to the emission line spectrum of hydrogen atom, and this is experimentally observed. However, the agreement with experimental results is not exact. Though the existence of the fine structure is predicated by KG equation, there is a discrepancy in the energy level splitting between theory and experiment.

16.6 MOTIVATIONS LEADING TO THE FORMULATION OF DIRAC EQUATION

When Klien Gordan equation was developed as an extension of non-relativistic to relativistic case, certain shortcomings were noted. So Klien Gordan equation was dropped and alternative relativistic equations were attempted. Today, we know that Klien Gordan equation is the correct equation to describe spin zero particles, and with the development of quantum field theory, there is no problem with KG equation. However, historically, Dirac formulated a new relativistic equation as an attempt to circumvent the problems of KG equation.

Let us, first, list the shortcomings of KG equation:

1. The KG equation leads to a state of free particle with negative energy.
2. The expression for probability density was not positive definite.
3. When it was applied to hydrogen atom, we get the existence of fine structure. However, it was not an agreement in details, and there are discrepancies between theory and experiment for the splitting of energy levels.

The interesting thing is the comparison of the expressions for probability density arising from the Schrödinger and KG equations.

- The Schrödinger equation is

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

$$\rho = \psi^*\psi \qquad S = \frac{\hbar}{2im}[\psi^*\nabla\psi - \nabla\psi^*\psi^*]$$

- The KG equation is

$$-\hbar^2\frac{\partial^2\psi}{\partial t^2} = c^2\hbar^2\nabla^2\psi + m^2c^4\psi$$

$$\rho = \frac{\hbar}{2imc^2}\left[\frac{\partial\psi^*}{\partial t}\psi - \frac{\partial\psi}{\partial t}\psi^*\right] \qquad \mathbf{S} = \frac{\hbar}{2im}[\psi^*\nabla\psi - \nabla\psi^*\psi^*]$$

In obtaining the expression for ρ and **S**, same procedure was used in the Schrödinger and KG equations. By comparing the Schrödinger and Klien Gordan equations, one can immediately recognize a significant difference between them. The time derivative in the Schrödinger equation is first-order derivative, leading to positive definite expression $\psi^*\psi$. The time derivative in the Klien Gordan equation is second-order, leading to positive indefinite expression for ρ. Dirac realized that a suitable relativistic wave equation should be first order in time derivative. But special theory of relativity demands that time and space coordinates occur in a symmetric manner. By this, it is meant that if the time derivative is first order, then the derivatives with respect to spatial coordinates x, y, and z should also be first order (note that in the Schrödinger equation, the spatial coordinate derivatives are second order). It should be mentioned that Dirac argued from different considerations that the wave equation for a quantum mechanical system be linear in $\partial/\partial t$; i.e., the general form of equation of motion for a quantum mechanical system should be $i\hbar\,\partial\psi/\partial t = H\psi$ (see Dirac). So Dirac constructed an equation in the following form:

$$E\psi = c\boldsymbol{\alpha}.\mathbf{p}\psi + \beta mc^2\psi$$

or

$$E\psi = H\psi \tag{16.35}$$

Dimensionally, each term has the dimension of energy. But α and β are totally unknown quantities whose nature has to be determined from other physics considerations. By replacing E and **p** by $i\hbar\, \partial/\partial t =$ and $-i\hbar\nabla$, respectively, we get the well-known Dirac equation.

$$i\hbar\frac{\partial\psi}{\partial t} = -i\hbar c\boldsymbol{\alpha}\cdot\nabla\,\psi + \beta mc^2\psi \tag{16.36}$$

16.7 MEANING OR NATURE OF α AND β

The Dirac equation is

$$i\hbar\frac{\partial\psi}{\partial t} = (c\boldsymbol{\alpha}.\mathbf{p} + \beta mc^2)\psi$$

In terms of E and H, this equation is

$$E\psi = H\psi \tag{16.37}$$

There are four quantities namely, α_x, α_y, α_z and β, which are to be determined. These four quantities cannot be the functions of coordinate x, y, z or t, and they cannot be functions of E and **p** also. If they depend on such variables, the Equation (16.36) will no longer be the equation for free particles, for the only potentials will be the functions of x, y, z, t or E and **p**.

This implies that these quantities are either numbers or matrices with constant numbers. In any case, they will commute with $E = i\hbar\,\partial/\partial t$ and $\mathbf{p} = -i\hbar\nabla$. We also have

$$EE\psi = EH\psi = HE\psi = HH\psi$$

$$\therefore E^2\psi = H^2\psi$$

Consistency with special theory of relativity demands H^2 should be $p^2c^2 + m^2c^4$; i.e., the above equation should be the same as the Klien Gordan equation.

$$-\hbar^2\frac{\partial\psi}{\partial t^2} = -\hbar^2c^2\nabla^2\psi + m^2c^4\psi$$

i.e.,

$$H^2\psi = (p^2c^2 + m^2c^4)\psi \tag{16.38}$$

This equation can be used to determine the exact nature of the four quantities $\alpha_x, \alpha_y, \alpha_z$ and β.

$$\begin{aligned}
\text{LHS} &= H^2\psi \\
&= [c\alpha_x p_x + c\alpha_y p_y + c\alpha_z p_z + \beta mc^2][c\alpha_x p_x + c\alpha_y p_y + c\alpha_z p_z + \beta mc^2]\psi \\
&= \big[c^2\alpha_x^2 p_x^2 + c^2 p_x p_y \alpha_x \alpha_y + c^2 p_x p_z \alpha_x \alpha_z + c^3 m p_x \alpha_x \beta \\
&\quad + c^2 p_x p_y \alpha_y \alpha_x + c^2 p_y^2 \alpha_y^2 + c^2 p_y p_z \alpha_y \alpha_z + mc^3 p_y \alpha_y \beta \\
&\quad + c^2 p_x p_z \alpha_z \alpha_x + c^2 p_y p_z \alpha_z \alpha_y + c^2 p_z^2 \alpha_z^2 + mc^3 p_y \alpha_z \beta \\
&\quad + mc^3 p_x \beta\alpha_x + mc^3 p_y \beta\alpha_y + mc^3 p_z \beta\alpha_z + m^2c^4\beta^2\big]\psi
\end{aligned}$$

$$= \Big[c^2 \alpha_x^2 p_x^2 + c^2 p_y^2 \alpha_y^2 xy^2 + c^2 p_z^2 \alpha_z^2 + m^2 c^4 \beta^2$$

$$+ c^2 p_x p_y (\alpha_x \alpha_y + \alpha_y \alpha_x) + c^2 p_x p_z (\alpha_x \alpha_z + \alpha_z \alpha_x)$$

$$+ c^2 p_y p_z (\alpha_y \alpha_z + \alpha_z \alpha_y) + mc^3 p_x (\alpha_x \beta + \beta \alpha_x)$$

$$+ mc^3 p_y (\alpha_y \beta + \beta \alpha_y) + mc^3 p_z [\alpha_z \beta + \beta \alpha_z] \psi$$

RHS $= [p_x c^2 + p_y^2 c^2 + p_z^2 c^2 + m^2 c^4] \psi$

If LHS and RHS are to be equal, then it implies that

$$a_x^2 = \alpha_y^2 = \alpha_z^2 = \beta^2 = 1$$

$$\alpha_x \alpha_y + \alpha_y \alpha_x = 0; \ \alpha_x \alpha_z + \alpha_z \alpha_x = 0; \ \alpha_y \alpha_z + \alpha_z \alpha_y = 0$$

$$\alpha_x \beta + \beta \alpha_x = 0; \ \alpha_y \beta + \beta \alpha_y = 0; \ \alpha_z \beta + \beta \alpha_z = 0 \qquad (16.39)$$

Obviously, α_x, α_y, α_z, and β are either numbers or set of anti-commuting matrices. α_x, α_y, α_z and β cannot be pure numbers. This can be easily seen as follows. The Equation (16.39) suggests $\alpha_x = +1$ or -1, $\alpha_y =$ +1 or -1,$\alpha_z = +1$ or -1 and $\beta = +1$ or -1. Suppose we choose $\alpha_x = +1$. Neither $\alpha_y = 1$ nor $\alpha_y = -1$ will satisfy the relations between them in the Equation (16.39). In the same way, one can argue for $\alpha_x = -1$. So no combination of these numbers will satisfy the relation (16.39).

Next choice is α_x, α_y, α_z and β are the set of anti-commuting matrices. If so, what is the order of these matrices and what are their elements? A number of general conclusions about these matrices can be reached using the anti-commutation relation (16.39).

Properties of α and β

1. Eigenvalues
   Consider the eigenvalue equation $\alpha_x \chi = \lambda \chi$.
   Multiplying both sides of the above equation by α_x, we get

$$\alpha_x \alpha_x \chi = \alpha_x \lambda \chi = \lambda \alpha_x \chi = \lambda^2 \chi$$

   Since $\alpha_x^2 = 1$, $\alpha_x^2 \chi$ becomes χ in the LHS.

$$\chi = \lambda^2 \chi$$

   i.e., $\lambda^2 = 1$

   or $\lambda = \pm 1$

   The same results hold good for other matrices. The eigenvalues of all the four matrices are +1 or −1.

2. Traceless matrices
   Let us evaluate the trace of particular matrix, say α_x.

$$tr\ \alpha_x = tr\ (\alpha_x \beta^2) \quad (\text{since } \beta^2 = I)$$

$$= tr\,(\alpha_x \beta\beta)$$

$$= tr\,(\beta\alpha_x\beta) \quad (tr\,(AB) = tr\,(BA))$$

$$= tr(-\alpha_x\beta\beta) \quad (\beta\alpha_x = -\alpha_x\beta)$$

$$= tr(-\alpha_x) \quad (\beta^2 = I)$$

$$\therefore tr\alpha_x = 0 \tag{16.40}$$

We can extend this argument to all the matrices.

3. Even-dimensional matrices

Take any particular matrix, say α_x. Diagonalize this matrix with the help of a suitable matrix S. The diagonal elements of the diagonal matrix will be the eigenvalues $\lambda_1, \lambda_2 \ldots$

$$S\alpha_x S^{-1} = \alpha_x^{(d)} = \begin{pmatrix} \lambda_1 & 0 & 0 & \ldots \\ \vdots & \lambda_2 & \ldots & \ldots \\ \vdots & \vdots & \lambda_3 & \ldots \\ & & & \end{pmatrix}$$

From matrix theory, we know that $tr\alpha_x$ = Sum of the eigenvalues of α_x.

i.e. $$\lambda_1 + \lambda_2 + \cdots = 0$$

But $$\lambda_1 = \pm 1,\ \lambda_2 = \pm 1,\ \lambda_3 = \pm 1 \ldots$$

This is possible provided half the diagonal elements of $\alpha_x^{(d)}$ should be +1 and other half should be −1. This implies that $\alpha_x^{(d)}$ or α_x is even dimensional.

The dimensions of α_x, α_y, α_z and β and can be 2, 4, 6…With these three properties, let us now try to construct them.

Are they 2 × 2 matrices?

From angular momentum theory, we know that Pauli matrices σ_x, σ_y and σ_z are anti-commuting 2 × 2 matrices. Suppose we choose $\alpha_x = \sigma_x, \alpha_y = \sigma_y$, and $\alpha_z = \sigma_z$. Now we have to find one more 2 × 2 matrix β, which has to anti-commute with σ_x, σ_y, and σ_z. We can show that it is not possible to find such a 2 × 2 matrix. This can be easily seen as follows:

Let us choose

$$\beta = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

Then

$$\sigma_x\beta + \beta\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix} + \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

$$= \begin{pmatrix} c+b & a+d \\ a+d & b+c \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

$$\sigma_y\beta+\beta\sigma_y=\begin{pmatrix}0 & -i\\ i & 0\end{pmatrix}\begin{pmatrix}a & b\\ c & d\end{pmatrix}+\begin{pmatrix}a & b\\ c & d\end{pmatrix}\begin{pmatrix}0 & -i\\ i & 0\end{pmatrix}$$

$$=\begin{pmatrix}i(b-c) & i(a-d)\\ i(a-d) & i(b-c)\end{pmatrix}=\begin{pmatrix}0 & 0\\ 0 & 0\end{pmatrix}$$

This is possible provided $a = b = c = d = 0$. Therefore, we will not be able to find a set four anti-commuting 2×2 matrices.

4 × 4 *matrices*

Next, we have to construct a set of 4×4 matrices, satisfying the anti-commutation relation (16.39)

In fact, we do find α_x, α_y, α_z and β as

$$\alpha_x=\begin{pmatrix}0 & \sigma_x\\ \sigma_x & 0\end{pmatrix},\ \alpha_y=\begin{pmatrix}0 & \sigma_y\\ \sigma_y & 0\end{pmatrix},\ \alpha_z=\begin{pmatrix}0 & \sigma_z\\ \sigma_z & 0\end{pmatrix}$$

and
$$\beta=\begin{pmatrix}I & 0\\ 0 & -I\end{pmatrix}$$

These are 4×4 Hermitian matrices. Note that there is no unique way of choosing elements α_x, α_y, α_z and β. The following is another choice of these matrices:

$$\alpha_x=\begin{pmatrix}\sigma_x & 0\\ 0 & -\sigma_x\end{pmatrix},\alpha_y=\begin{pmatrix}0 & I\\ I & 0\end{pmatrix},\alpha_z=\begin{pmatrix}\sigma_z & 0\\ 0 & -\sigma_z\end{pmatrix},\beta=\begin{pmatrix}0 & iI\\ -iI & 0\end{pmatrix}.$$

16.8 EQUATION OF CONTINUITY

Having constructed a new equation, we should check whether at least some problems of Klien Gordan equations have been solved. The simplest test is whether the expression for probability density is positive definite.

So let us derive the equation of continuity for a particle obeying the Dirac equation.

The Dirac equation is

$$\begin{aligned}i\hbar\frac{\partial\psi}{\partial t}&=c\alpha_x p_x\psi+c\alpha_y p_y\psi+c\alpha_z p_z\psi+\beta mc^2\psi\\ i\hbar\frac{\partial\psi}{\partial t}&=-i\hbar c\alpha_x\frac{\partial\psi}{\partial x}-i\hbar c\alpha_y\frac{\partial\psi}{\partial y}-i\hbar c\alpha_z\frac{\partial\psi}{dz}+\beta mc^2\psi\end{aligned}\tag{16.41}$$

Taking Hermitian conjugate, we have

$$-i\hbar\frac{\partial\psi^\dagger}{\partial t}=i\hbar c\frac{\partial\psi^\dagger}{\partial x}\alpha_x^\dagger+i\hbar c\frac{\partial\psi^\dagger}{\partial y}\alpha_y^\dagger+i\hbar c\frac{\partial\psi^\dagger}{\partial z}\alpha_z^\dagger+mc^2\psi^\dagger\beta^\dagger$$

$$-i\hbar\frac{\partial\psi^\dagger}{\partial t}=i\hbar c\frac{\partial\psi^\dagger}{\partial x}\alpha_x+i\hbar c\frac{\partial\psi^\dagger}{\partial y}\alpha_y+i\hbar c\frac{\partial\psi^\dagger}{\partial z}\alpha_z+mc^2\psi^\dagger\beta\tag{16.42}$$

Multiplying the Equation (16.41) by $\psi^\dagger$ from left and the Equation (16.42) by ψ from right, we get

$$i\hbar\psi^\dagger\frac{\partial\psi}{\partial t} = -i\hbar c\psi^\dagger\alpha_x\frac{\partial\psi}{\partial x} - i\hbar c\psi^\dagger\alpha_y\frac{\partial\psi}{\partial y} - i\hbar c\psi^\dagger\alpha_z\frac{\partial\psi}{\partial z} + mc^3\psi^\dagger\beta\psi \tag{16.43}$$

and

$$-i\hbar\frac{\partial\psi^\dagger}{\partial t}\psi = i\hbar c\frac{\partial\psi^\dagger}{\partial x}\alpha_x\psi + i\hbar c\frac{\partial\psi^\dagger}{\partial y}\alpha_y\psi + i\hbar\frac{\partial\psi^\dagger}{\partial z}\alpha_z\psi + mc^3\psi^\dagger\beta\psi \tag{16.44}$$

Subtracting the Equation (16.44) from the Equation (16.43), we get

$$i\hbar\left(\frac{\partial\psi^\dagger}{\partial t}\psi + \partial\psi^\dagger\frac{\partial\psi}{\partial t}\right) = -i\hbar c\left(\psi^\dagger\alpha_x\frac{\partial\psi}{\partial x} + \frac{\partial\psi^\dagger}{\partial x}\alpha_x\psi\right) - i\hbar c\left(\psi^\dagger\alpha_y\frac{\partial\psi}{\partial y} + \frac{\partial\psi^\dagger}{\partial y}\alpha_y\psi\right)$$

$$-i\hbar c\left(\psi^\dagger\alpha_z\frac{\partial\psi}{\partial z} + \frac{\partial\psi^\dagger}{\partial z}\alpha_z\psi\right)$$

This can be rewritten as

$$i\hbar\frac{\partial}{\partial t}(\psi^\dagger\psi) = -i\hbar c\frac{\partial}{\partial x}(\psi^\dagger\alpha_x\psi) - i\hbar c\frac{\partial}{\partial y}(\psi^\dagger\alpha_y\psi) - i\hbar c\frac{\partial}{\partial z}(\psi^\dagger\alpha_z\psi)$$

i.e.,

$$\frac{\partial}{\partial t}(\psi^\dagger\psi) = -c\nabla.(\psi^\dagger\boldsymbol{\alpha}\psi)$$

$$\frac{\partial}{\partial t}(\psi^\dagger\psi) = \nabla.(c\psi^\dagger\boldsymbol{\alpha}\psi) = 0$$

$$S = c\psi^\dagger\alpha\psi$$

This equation has the form

$$\frac{\partial\rho}{\partial t} + \nabla\cdot\mathbf{S} = 0$$

provided we identify ρ and **S** as $\rho = \psi^\dagger\psi$ and $\mathbf{S} = c\psi^\dagger\boldsymbol{\alpha}\psi.$

Note that $\rho = \psi^\dagger\psi$ is always a positive quantity. Therefore, it can be interpreted as probability density.

So the Dirac equation gets over one of the problems of the KG equation. We should mention that this interpretation of wave function in the Dirac equation also runs into difficulty as will be shown later.

16.9 SOLUTION TO FREE PARTICLE DIRAC EQUATION

The free particle Dirac equation is

$$i\hbar\frac{\partial\psi}{\partial t} = c\boldsymbol{\alpha}.(-i\hbar\nabla)\psi + \beta mc^2\psi \tag{16.45}$$

Since we are interested in the free particle solution, we expect the particle to have definite momentum and energy. The natural solution is $\psi = Ae^{i(\mathbf{p}.\mathbf{r}-Et)/\hbar}$. As in the case of the Klien Gordan equation, this is consistent with De Broglie wave concept.

Since α and β are 4×4 matrices, the wave function should also be a four-component column vector. So the plane wave function has to be modified. This suggests that the solution to the free particle equation is

$$\psi = Au(\mathbf{p})e^{i(\mathbf{p}.\mathbf{r}-Et)/\hbar} \tag{16.46}$$

Note that $u(\mathbf{p})$ has to depend only on momentum; otherwise, ψ will not be an eigenfunction of momentum operator. $u(\mathbf{p})$ is a column vector with four rows; i.e., it is a 4×1 matrix. It is called spinor, and the reason for such term will become clear a little later.

We have to check whether the Equation (16.46) is solution to the Equation (16.45), and in the process, we have to find $u(\mathbf{p})$ as well as energy eigenvalue E.

Substituting the Equation (16.46) in the Equation (16.45), we have the LHS given by

$$i\hbar\frac{\partial\psi}{\partial t} = i\hbar\frac{\partial}{\partial t}[Au(\mathbf{p})e^{i(\mathbf{p}.\mathbf{r}-Et)/\hbar}] = EAu(\mathbf{p})e^{i(\mathbf{p}.\mathbf{r}-Et)/\hbar}$$

And RHS as

$$c\boldsymbol{\alpha}\cdot(-i\hbar\nabla\psi) + \beta mc^2\psi = c\boldsymbol{\alpha}\cdot(-i\hbar\nabla(Au(\mathbf{p})e^{i(\mathbf{p}.\mathbf{r}-Et)/\hbar}) + \beta mc^2 Au(\mathbf{p})e^{i(\mathbf{p}.\mathbf{r}-Et)/\hbar}$$

$$= c\boldsymbol{\alpha}\cdot\mathbf{p}Au(\mathbf{p})e^{i(\mathbf{p}.\mathbf{r}-Et)/\hbar} + \beta mc^2 Au(\mathbf{p})e^{i(\mathbf{p}.\mathbf{r}-Et)/\hbar}$$

Equating LHS and RHS, we get

$$EAu(\mathbf{p})e^{i(\mathbf{p}.\mathbf{r}-Et/\hbar)} = c\boldsymbol{\alpha}.(\mathbf{p})Au(\mathbf{p})e^{i(\mathbf{p}.\mathbf{r}-Et)/\hbar} + \beta mc^2 Au(\mathbf{p})e^{i(\mathbf{p}.\mathbf{r}-Et)/\hbar}$$

$$\therefore Eu(\mathbf{p}) = c\boldsymbol{\alpha}\cdot\mathbf{p}u(\mathbf{p}) + \beta mc^2 u(\mathbf{p}) \tag{16.47}$$

This is energy eigenvalue equation. This can be written as

$$(EI - c\boldsymbol{\alpha}\cdot\mathbf{p} - \beta mc^2)u(\mathbf{p}) = 0 \tag{16.48}$$

Here I is 4×4 identity matrix. Note that the same letter I is used to denote both 2×2 or 4×4 unit matrix. The order of the matrix can be understood from the context.

Let us write I, α_x, α_y, α_z and β (all of them 4×4 matrix) in the block form as follows.

$$I = \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix}, \alpha_x = \begin{pmatrix} 0 & \sigma_x \\ \sigma_x & 0 \end{pmatrix}, \alpha_y = \begin{pmatrix} 0 & \sigma_y \\ \sigma_y & 0 \end{pmatrix}, \alpha_z = \begin{pmatrix} 0 & \sigma_z \\ \sigma_z & 0 \end{pmatrix} \text{ and } \beta = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} \tag{16.49}$$

Each element here is 2×2 matrix. In particular, we have

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \text{ and } \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{16.50}$$

Let us write $u(\mathbf{p})$ also in the block form. i.e.,

$$u(\mathbf{p}) = \begin{pmatrix} V \\ W \end{pmatrix} \tag{16.51}$$

where V and W are column vectors (2×1 matrix) with two rows. They are called two-component spinors.

The Equation (16.48) now becomes

$$\left[\begin{pmatrix} EI & 0 \\ 0 & EI \end{pmatrix} - \begin{pmatrix} 0 & c\boldsymbol{\sigma}.\mathbf{p} \\ c\boldsymbol{\sigma}.\mathbf{p} & 0 \end{pmatrix} - \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} mc^2 \right] \begin{bmatrix} V \\ W \end{bmatrix} = 0$$

$$\begin{bmatrix} (E - mc^2)I & -c\boldsymbol{\sigma}.\mathbf{p} \\ -c\boldsymbol{\sigma}.\mathbf{p} & (E + mc^2)I \end{bmatrix} \begin{bmatrix} V \\ W \end{bmatrix} = 0$$

i.e.,

$$(E - mc^2)V = c\boldsymbol{\sigma}.\mathbf{p}W \tag{16.52}$$

and

$$(E + mc^2)W = c\boldsymbol{\sigma}.\mathbf{p}V \tag{16.53}$$

Multiplying the Equation (16.52) by $E + mc^2$, we get

$$(E + mc^2)(E - mc^2)V = c\boldsymbol{\sigma}.\mathbf{p}(E + mc^2)W$$

$$= c\boldsymbol{\sigma} \cdot \mathbf{p} c\boldsymbol{\sigma}.\mathbf{p}V \quad \text{(Use the Equation (16.53))}$$

Let us make use of the identity

$$(\boldsymbol{\sigma} \cdot \mathbf{A})(\boldsymbol{\sigma} \cdot \mathbf{B}) = \mathbf{A} \cdot \mathbf{B} + i\boldsymbol{\sigma} \cdot (\mathbf{A} \times \mathbf{B})$$

We have

$$(\boldsymbol{\sigma}.\mathbf{p})(\boldsymbol{\sigma}.\mathbf{p}) = p^2$$

Then, we get

$$\therefore (E^2 - m^2c^4 - p^2c^2)V = 0 \tag{16.54}$$

The non-trivial solution is possible, provided

$$E^2 = p^2c^2 + m^2c^4$$

or

$$E = \pm\sqrt{p^2c^2 + m^2c^4} \tag{16.55}$$

So, in the case of the Dirac equation also (like in the KG equation), we get two energy eigenvalues namely, $E = +\sqrt{p^2c^2 + m^2c^4}$ and $E = -\sqrt{p^2c^2 + m^2c^4}$. Let us define $E_p = +\sqrt{p^2c^2 + m^2c^4}$. Therefore, the energy eigenvalues are $E = E_p$ and $E = -E_p$. Note that one of the energy eigenvalues is negative value. We will see later Dirac's interpretation of the negative energy solution.

Let us continue to find the eigenfunctions corresponding to these two eigenvalues.

Let us write the Equations (16.52) and (16.53).

$$W = \frac{c\boldsymbol{\sigma}.\mathbf{p}}{E + mc^2} V \tag{16.56}$$

and

$$V = \frac{c\boldsymbol{\sigma}.\mathbf{p}}{E - mc^2} W \tag{16.57}$$

Eigensolutions for $E = E_p$

Consider the solution given in the Equation (16.57).

$$V = \frac{c\boldsymbol{\sigma}.\mathbf{p}}{E_p - mc^2} W$$

As $p \to 0$, $E_p = \sqrt{p^2c^2 + m^2c^4}$ tends to mc^2, and so the solution blows up. So this solution is not an admissible solution.

Next, consider the solution given in the Equation (16.56).

$$W = \frac{c\boldsymbol{\sigma}.\mathbf{p}}{E_p + mc^2} V$$

As $p \to 0$, $E_p \to mc^2$, and the solution is well behaved in this limit. So this is an admissible solution. Since V is a two-component spinor, we can write V in terms of two linearly independent column vectors χ_1, and χ_2. One choice is $\chi_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\chi_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$. But these are not the only choices. Let us choose χ_1 and χ_2 such that $\chi_1^\dagger \chi_1 = \chi_2^\dagger \chi_2 = 1$ and $\chi_1^\dagger \chi_2 = 0$.

Then, the eigensolutions corresponding to $E = E_p$ are

$$\psi_{\mathrm{I}} = N \begin{pmatrix} \chi_1 \\ \dfrac{c\boldsymbol{\sigma}.\mathbf{p}\chi_1}{E_p + mc^2} \end{pmatrix} e^{i(\mathbf{p}.\mathbf{r} - E_p t)/\hbar} \tag{16.58}$$

and

$$\psi_{\mathrm{II}} = N \begin{pmatrix} \chi_2 \\ \dfrac{c\boldsymbol{\sigma}.\mathbf{p}\chi_2}{E_p + mc^2} \end{pmatrix} e^{i(\mathbf{p}.\mathbf{r} - E_p t)/\hbar} \tag{16.59}$$

Eigensolutions for $E = -E_p$

Note that as $p \to 0$, $E = -E_p = -\sqrt{p^2c^2 + m^2c^4}$ tends to $-mc^2$. Now consider the solution given in the Equation (16.56).

$$W = \frac{c\boldsymbol{\sigma}.\mathbf{p}}{-E_p + mc^2} V$$

As $p \rightarrow 0$, $E_{\text{p}} \rightarrow mc^2$, and so this solution blows up in this limit. Hence, this solution is not an admissible solution. Next, consider the solution given in the Equation (16.57). Substituting $E = -E_p$ in this equation, we get

$$V = \frac{c\boldsymbol{\sigma}.\mathbf{p}}{-E_p - mc^2} W$$

This function is a well-behaved function in the limit $p \rightarrow 0$ or $E_p \rightarrow mc^2$, and so this is an accepted solution. As in the earlier case, W is an arbitrary two-component spinor. So we can write W in terms of two linearly independent two-component spinors χ_1 and χ_2. Therefore, the eigensolution corresponding to $E = - E_p$ is

$$\psi_{\text{III}} = N \begin{bmatrix} -\dfrac{c\boldsymbol{\sigma}.\mathbf{p}}{E_p + mc^2} \chi_1 \\ \chi_1 \end{bmatrix} e^{i(\mathbf{p}.\mathbf{r}+E_p t)/\hbar} \tag{16.60}$$

and

$$\psi_{\text{IV}} = N \begin{bmatrix} -\dfrac{c\boldsymbol{\sigma}.\mathbf{p}}{E_p + mc^2} \chi_2 \\ \chi_2 \end{bmatrix} e^{i(\mathbf{p}.\mathbf{r}+E_p t)/\hbar} \tag{16.61}$$

There are four linearly independent solutions for the free particle Dirac equation. How, are these four states distinguished?

Apart from Hamiltonian and momentum, we need to introduce one more operator namely, helicity operator to characterize these solutions. There are three commuting operators here. They are Hamiltonian $\hat{H}$, momentum $\hat{\mathbf{p}}$, and helicity $\hat{\Lambda} = \frac{\hbar}{2}\boldsymbol{\Sigma}\cdot\mathbf{p} = \frac{\hbar}{2}\begin{pmatrix} \boldsymbol{\sigma}\cdot\mathbf{p} & 0 \\ 0 & \boldsymbol{\sigma}\cdot\mathbf{p} \end{pmatrix}$. Therefore, the eigenfunction should be eigenfunction of $\hat{H}$, $\hat{p}$ and $\hat{\Lambda}$. In the next section, we will show that $\hbar/2\, \boldsymbol{\Sigma}\cdot\mathbf{p}$ is the projection of spin along the direction $\mathbf{p}$. Therefore, all the four solutions can be specified as follows:

$$\psi_{\text{I}} = N \begin{pmatrix} \chi_1 \\ \dfrac{c\boldsymbol{\sigma}\cdot\mathbf{p}}{E_p + mc^2} \chi_1 \end{pmatrix} e^{i(\mathbf{p}.\mathbf{r}-E_p t)/\hbar} \quad \text{momentum } \mathbf{p}\text{, energy } E_{\text{p}} \text{ and spin} \uparrow$$

$$\psi_{\text{II}} = N \begin{pmatrix} \chi_2 \\ \dfrac{c\boldsymbol{\sigma}\cdot\mathbf{p}}{E_p + mc^2} \chi_2 \end{pmatrix} e^{i(\mathbf{p}.\mathbf{r}-E_p t)/\hbar} \quad \text{momentum } \mathbf{p}\text{, energy } E_{\text{p}} \text{ and spin} \downarrow$$

$$\psi_{\text{III}} = N \begin{pmatrix} -\dfrac{c\boldsymbol{\sigma}\cdot\mathbf{p}}{E_p + mc^2} \chi_1 \\ \chi_1 \end{pmatrix} e^{i(\mathbf{p}.\mathbf{r}-E_p t)/\hbar} \quad \text{momentum } \mathbf{p}\text{, energy } -E_p\text{, and spin} \uparrow$$

$$\psi_{IV} = N\begin{pmatrix} -\dfrac{c\boldsymbol{\alpha}\cdot\mathbf{p}}{E_p + mc^2}\chi_1 \\ \chi_2 \end{pmatrix} e^{i(\mathbf{p}.\mathbf{r}-E_p t)/\hbar} \quad \text{momentum } \mathbf{p}\text{; energy } -E_p\text{, and spin } \downarrow$$

Normalization of Eigenfunction

The normalization of the Dirac spinor is done by demanding the condition that

$$\int \psi^\dagger \psi d^3\mathbf{r} = 1$$

Since Dirac spinor involves a plane wave function, the normalization can be done either by δ function normalization $\left(\int \psi^\dagger \psi d\mathbf{r} = \delta(\mathbf{p}-\mathbf{p})\right)$ or by box normalization $\left(\int \psi^\dagger \psi d^3\mathbf{r} = 1\right.$, where the volume of the integration is restricted to a box of side L).

For the sake of simplicity, let us adopt box normalization. Let us take any particular solution, say ψ_I. Then, we have

$$\int \psi^\dagger \psi d^3\mathbf{r} = \int_0^L\int_0^L\int_0^L dx\,dy\,dz\,\psi^\dagger\psi$$

$$= \int_0^L\int_0^L\int_0^L dx\,dy\,dz\,N^2\left[\chi_1^\dagger \quad \frac{\chi_1^\dagger c\boldsymbol{\sigma}.\mathbf{p}}{E_p + mc^2}\right] \times \begin{bmatrix} \chi_1 \\ \dfrac{c\boldsymbol{\sigma}.\mathbf{p}\chi_1}{E_p + mc^2} \end{bmatrix} e^{-i(\mathbf{p}.\mathbf{r}-E_p t)/\hbar} e^{i(\mathbf{p}.\mathbf{r}-E_p t)/\hbar}$$

$$= \int_0^L\int_0^L\int_0^L dx\,dy\,dz\,N^2\left[\chi_1^\dagger\chi_1 + \frac{\chi_1^\dagger c\cdot\boldsymbol{\sigma}.\mathbf{p}c\boldsymbol{\sigma}.\mathbf{p}\chi_1}{(E_p + mc^2)^2}\right]$$

$$= \int_0^L\int_0^L\int_0^L dx\,dy\,dz\,N^2\left[\chi_1^\dagger\chi_1 + \frac{\chi_1^\dagger c^2 p^2 \chi_1}{(E_p + mc^2)^2}\right]$$

where we have used the fact $(\boldsymbol{\sigma}\cdot\mathbf{p})(\boldsymbol{\sigma}\cdot\mathbf{p}) = p^2$. Note $\chi_1^\dagger\chi_1 = 1$.

$$\therefore \left[\chi_1^\dagger\chi_1 + \frac{\chi_1^\dagger c^2 p^2 \chi_1}{(E_p + mc^2)^2}\right] = \left[1 + \frac{p^2c^2}{(E_p + mc^2)^2}\right]$$

$$= \frac{E_p^2 + 2mc^2E_p + m^2c^4 + p^2c^2}{(E_p + mc^2)^2}$$

$$= \frac{2E_p(E_p + mc^2)}{(E_p + mc^2)^2} = \frac{2E_p}{(E_p + mc^2)}$$

$$\therefore \int \psi^{\dagger}\psi^{3}\mathbf{r} = N^2 \frac{2E_p}{(E_p + mc^2)} \int_0^L\int_0^L\int_0^L dxdydz$$

$$= N^2 \frac{2E_p}{(E_p + mc^2)} L^3 = \frac{N^2 2E_p V}{(E_p + mc^2)} = 1$$

$$\therefore N = \sqrt{\frac{(E_p + mc^2)}{2E_p V}}$$

It can be shown that the normalization factor N is the same for all the four solutions.

16.10 DIRAC'S INTERPRETATION OF NEGATIVE ENERGY STATES

Existence of Negative Energy States

There are four linearly independent solutions to Dirac's free particle equation. Of these four solutions, two solutions represent states with positive energy $E = E_p = \sqrt{p^2c^2 + m^2c^4}$, and other two solutions represent states with negative energy $E = -E_p = -\sqrt{p^2c^2 + m^2c^4}$. Like the Klien Gordan equation, the Dirac equation also has negative energy solutions. In view of the fact that no free electron with negative energy has been experimentally observed, what does one do with negative energy states? It is not possible to discard negative energy states as unphysical and ignore them.

The Dirac spinors are 4 × 1 column vectors. Therefore, the linear vector space of the general solutions of the Dirac equation requires four linearly independent solutions to form a complete set. If the positive energy spinors are accepted as an allowed solution, one has to necessarily accept the negative energy solutions also as allowed ones.

Dirac accepted the negative energy states as allowed states and came with an ingenious interpretation, which is presented below.

Instability of an Electron

If the negative energy states are allowed states, what are their consequences? Let us first start with the energy spectrum of an electron. There are two groups of states: one group of states with $E = E_p = -\sqrt{p^2c^2 + m^2c^4}$. and another group of states with $E = -E_p = -\sqrt{p^2c^2 + m^2c^4}$. Positive energy states start from $E = mc^2$ and extend up to $+\infty$. Similarly, negative energy states start from $E = -mc^2$ and extend blow to $-\infty$. This is shown in Fig.16.2.

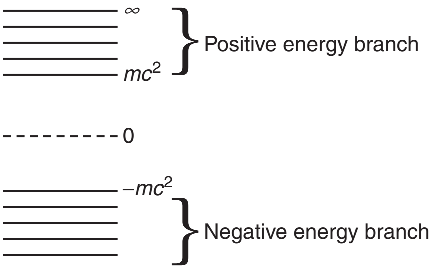


Fig.16.2 Energy spectrum for dirac electron

There are infinite number of states both in positive and negative energy branches. The two branches are separated by a gap $2mc^2$. To go from one negative energy state to a positive energy state, it requires energy more than $2mc^2$.

Let us start with an electron in a positive energy state. Since negative energy states are lower in energy, compared with this positive energy state, it will make a transition to negative energy state. The transitions will not stop, since infinite number of negative energy states is available for the electron. It can go on making transitions to lower and lower energy states. Thus, we can never find a stable positive energy electron even if it has a minimum energy $E = mc^2$. So the existence of negative energy states means that no stable positive energy electron can be found. We do not find any such instability, and electron is a very stable particle. How do we resolve this contradiction?

Negative Energy Sea and Pauli's Exclusion Principle

To avoid this instability, Dirac suggested that all the negative energy states are occupied by electrons. The infinite negative energy electrons together constitute negative energy sea. How is the instability of the electron prevented by the negative energy sea? Since the electrons are fermions, they obey Pauli's exclusion principle, and so there can be only one electron in any state. Since all the negative energy states are occupied, a positive energy electron cannot make a transition to any one of the negative energy states, and hence its stability is assured.

Ground State – Unobservable Negative Energy Sea

In the Dirac's interpretation, each positive energy electron is accompanied by a negative energy sea. The negative energy sea should have an infinite energy, infinite mass and infinite charge. Then, how can a single electron have a finite (though very small) mass, charge and energy? The answer to this question is based on defining a new ground state. The entire negative energy states constitute the ground state for an electron (see Fig.16.3). We have to conclude that the positive energy states are excitations over the ground state. The ground state itself may not be observed directly, but excited states and other deviations from the ground state should be observable. As an analogy, this concept is similar to the electrostatic case, where only the potential difference is measured, while the absolute potential is not measurable.

It is true that by postulating a new type of ground state for an electron, we are able to explain why electrons with infinite energy, infinite mass and infinite charge have not been observed. It is further claimed that this ground state as such is not observable, but only the excitations are observed. If we stop here, this 'explanation' does not explain much since it cannot be observed. However, if it leads to some consequences, which can be verified experimentally, then this particular type of ground state can be taken more seriously.

Though this ground state itself may not be directly observable, the deviations and the modifications of the ground state (if such a ground state exists) should lead to observable consequences, which can be verified experimentally. Indeed, we present some interesting consequences due to the changes introduced in the ground state. It should be mentioned that though we discuss this concept of ground state in detail here, it has been superseded by the quantum field theory, which provides a different kind of ground state.

16.11 OBSERVABLE CONSEQUENCES OF NEGATIVE ENERGY SEA

Antiparticle

One of the important observable consequences of Dirac's interpretation of negative energy state is the emergence of a new concept called antiparticle. The concept of antiparticle is purely a quantum mechanical idea.

The antiparticle of electron is positron, and let us explain how the concept of positron naturally emerges in the Dirac equation. Let E_0 and Q_0 be the total energy and charge of the ground state. Let us modify this ground state by removing an electron (of charge $-e$) from a negative energy state $-|E|$. The new state has energy $E_0-(-|E|)$ and $Q_0-(-e)$; i.e., we have the following situation

Ground state	New State
	One vacancy in ground state with energy $-E$
E_0, Q_0	$(E_0-(-E)),\ (Q_0-(-e))$

(16.62)

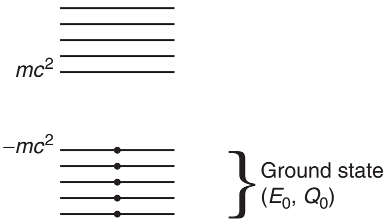


Fig.16.3 Ground state of dirac electron. Dots are electrons

The following analogy will be useful. Suppose the total energy and the charge of the negative energy sea are –100J and –100e (don't worry about the numbers). By removing one particle from the negative energy sea, we have modified the ground state. The total energy of the new state is –99J and –99e. The new state is equivalent to adding a particle with energy 1J and charge $+e$ to the ground state. The second column in the Equation (16.62) has ground state plus a particle with positive energy $+E$ and positive charge $+e$. This leads to the following interpretation.

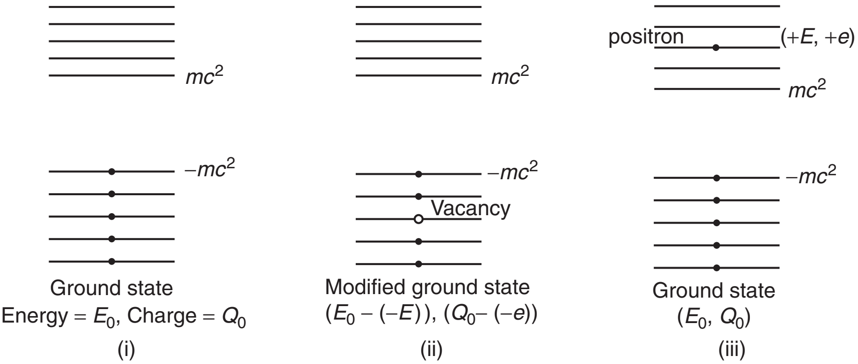


Fig.16.4 Positron state. Representations (ii) and (iii) are equivalents

A vacancy in the state of $-|E|$ in the negative energy sea is equivalent to the presence of a particle with positive energy $|E|$ and charge $+e$. If a particle has positive energy, it should be observable. Dirac originally tried to identify it as proton. Later on, it was identified as positron. It is called the antiparticle of the electron.

So now the Dirac equation describes two kinds of particles: electron of mass m_e and charge $-e$ and positron of mass m_e and charge $+e$. Within quantum mechanics, the description of positron of positive energy $|E|$ is equivalent to the vacancy of state with $-|E|$ in the negative energy sea of electron. This is shown in Fig.16.4. We can describe the system in terms of an electron alone by taking the account of both positive and negative energy states or equivalently describe the system in terms of two positive energy particles namely, electron and positron. Quantum field theory clarifies this equivalence more precisely. A negative energy electron propagating backwards in time is equivalent to a positron propagating forward in time.

Pair Creation

One of the important consequences of negative energy sea is the existence of a process called pair creation. In this process, a photon creates a pair of positron e^+ and an electron e^-.

$$\gamma \rightarrow e^- + e^+$$

(Actually this process is not allowed since conservation of energy and momentum together cannot be satisfied. A related process is γ + nucleus $X \rightarrow e^- + e^+$ + nucleus X'.) This process has a very simple explanation in terms of negative energy sea. Fig.16.5(a) is the ground state of negative energy, which itself is not directly observable. A photon of energy greater than $2mc^2$ interacts with one negative energy electron, moving it to a positive energy state, leaving a vacancy in the negative energy sea of negative energy electrons.

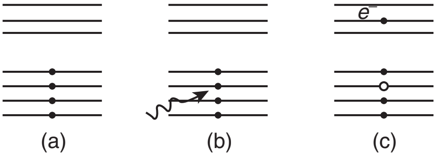


Fig.16.5 (a) Unobservable ground state (b) Interaction of a photon with a negative energy electron (c) Positive energy electron and a vacancy in a negative energy state

This electron absorbs the photon, and its energy becomes positive, and hence it is observable. At the same time, the ground state is altered to a new state (Fig.16.5(c)), in which one negative energy electron is missing; i.e., there is a vacancy in the negative energy sea, which is equivalent to the presence of positive energy positron (see Fig.16.6). This is schematically represented by

$$\gamma \rightarrow e^- + e^+$$

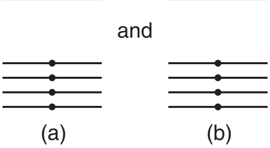


Fig.16.6 An electron, a positron and the ground state in the pair creation process. Equivalent to Fig.16.5(c)

This is consistent with charge conservation. The only necessity is that new particle called positron should exist. The charge before and after the pair creation is zero. $2mc^2$ is the minimum energy of the electron and the positron together, and so the photon should have the minimum energy $2mc^2$ for this process to take place. The same conclusion is reached from the energy-level diagram. Such a process has indeed been experimentally observed.

Pair Annihilation

The reverse process of pair creation is pair annihilation. In this process, an electron e^- and a positron e^+ annihilate themselves to produce a photon.

i.e.,
$$e^- + e^+ \to \gamma$$

Consider a pair of electron and positron. In the energy-level diagram, an electron can be represented by an occupied state (a dot in positive energy state), and a positron can be represented by an empty state (a circle) in the negative energy sea as in Fig.16.7.

Since one of the negative energy states is unoccupied, a positive energy electron can make transition to such state, bringing back the unobservable ground state. In effect, a positive energy electron as well as the altered negative energy states with a vacancy are lost. So we do not observe any particle: neither the electron nor the positron. This is precisely implied by the equation

$$e^- + e^+ \to \gamma$$

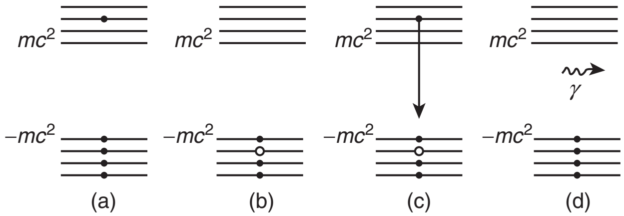


Fig.16.7 (a) An electron (b) A positron (c) Pair annihilation as transition from positive energy state to a negative energy state (d) Photon and the electron ground state

It should be mentioned that the concept of assuming the Dirac's negative energy sea as electron's ground state is successful, in explaining the existence of antiparticle for an electron. This scheme will not work in the case of bosons since Pauli's exclusion principle is not applicable to bosons, and hence it cannot be the mechanism to prohibit the positive energy boson from making transition to a negative energy state. However, this concept of ground state has been dropped now. The existence of the antiparticle has been formulated in a more general way for all the particles, including bosons in quantum field theory.

Negative Energy Electron in Em Field – Justification for Positron Picture

Let us try to justify the equivalence between the positive energy positron state and the vacancy in the negative energy state by considering the Dirac equation for an electron in electromagnetic field $(\Phi, \mathbf{A})$.

The Dirac equation for an electron in the presence of an electromagnetic filed $(\Phi, \mathbf{A})$ is

$$\left[i\hbar\frac{\partial}{\partial t} + e\Phi\right]\psi = c\boldsymbol{\alpha}.\left(-i\hbar\nabla + e\frac{\mathbf{A}}{c}\right)\psi + \beta mc^2\psi \tag{16.63}$$

Consider a stationary state solution $\psi(\mathbf{r}) = e^{-iEt/\hbar}u(\mathbf{r})$. Then, the above equation becomes

$$\left(i\hbar\frac{\partial}{\partial t} + e\phi\right)e^{-iEt/\hbar}u(\mathbf{r})$$

$$= c\boldsymbol{\alpha}.\left(-i\hbar\nabla + e\frac{\mathbf{A}}{c}\right)e^{-iEt/\hbar}u(\mathbf{r}) + \beta mc^2 e^{-iEt/\hbar}u(\mathbf{r}) \tag{16.64}$$

This equation becomes

$$(E + e\Phi)\, u(\mathbf{r}) = (-i\hbar c\boldsymbol{\alpha}.\nabla u(\mathbf{r})) + e\boldsymbol{\alpha}.\mathbf{A}u + \beta mc^2 u \tag{16.65}$$

Let us assume that effect of electromagnetic interaction is so small that the energy spectrum resembles very much like that of a free electron; a series of positive energy states separated by a gap exists; i.e., a positive energy branch and a negative energy branch exist.

The four matrices α_x, α_y, α_z and β are anti-commuting matrices. Let us choose α_x, α_y, α_z and β such that $\alpha_x^* = \alpha_x, \alpha_y^* = \alpha_y, \alpha_z^* = \alpha_z$ and $\beta^* = -\beta$. One such choice is

$$\alpha_x = \begin{pmatrix} \sigma_x & 0 \\ 0 & -\sigma_x \end{pmatrix}, \alpha_y = \begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix}, \alpha_z = \begin{pmatrix} \alpha_z & 0 \\ 0 & -\alpha_z \end{pmatrix} \; \beta = \begin{pmatrix} 0 & iI \\ -iI & 0 \end{pmatrix} \tag{16.66}$$

Note that these four matrices are also anti-commuting matrices. Now let us consider a negative energy state with energy $E = -|E|$. Therefore, the Equation (16.65) becomes

$$-|E|\, u(\mathbf{r}) = c\boldsymbol{\alpha}.(-i\hbar\nabla u) + e\boldsymbol{\alpha}\cdot\mathbf{A} + \beta mc^2 u - e\Phi u \tag{16.67}$$

Taking complex conjugate of this equation, we get

$$-|E|\, u^*(\mathbf{r}) = c\boldsymbol{\alpha}^*\cdot(i\hbar\nabla u^*) + e\boldsymbol{\alpha}^*\cdot\mathbf{A}^* + \beta^* mc^2 u - e\Phi^* u^* \tag{16.68}$$

Assume the potentials $\mathbf{\Phi}$ and $\mathbf{A}$ are real, i.e., $\mathbf{\Phi}^* = \mathbf{\Phi}$ and $\mathbf{A}^* = \mathbf{A}$.

Therefore, Equation (16.67) becomes

$$-|E|\, u^*(\mathbf{r}) = \boldsymbol{\alpha}\cdot(i\hbar\nabla u^*) + e\boldsymbol{\alpha}\cdot\mathbf{A}u^* - \beta mc^2 u^* - e\Phi u^*$$

or

$$|E|\, u^*(\mathbf{r}) = \boldsymbol{\alpha}\cdot(-i\hbar\nabla u^*) - e\boldsymbol{\alpha}\cdot\mathbf{A}u^* + \beta mc^2 u^* + e\Phi u^* \tag{16.69}$$

We can rewrite this equation as

$$(|E| - e\Phi)u^*(\mathbf{r}) = c\boldsymbol{\alpha}\cdot\left(-i\hbar\nabla - \frac{e\mathbf{A}}{c}\right)u^* + \beta mc^2 u^* \tag{16.70}$$

A comparison of this equation with the Equation (16.65) suggests that u^* is a solution of the Dirac equation for a positively charged particle of mass with positive energy $|E|$ in the presence of electromagnetic field.

So we have established that a state $u(\mathbf{r})$ of energy $-|E|$ for a particle of charge $-e$ and mass m is equivalent to a state $u^*(\mathbf{r})$ of energy $|E|$ of its antiparticle of charge e and mass m.

16.12 UNTENABILITY OF INTERPRETATION OF ψ IN THE NON-RELATIVISTIC SENSE

In the non-relativistic quantum mechanics, we can interpret $\psi(\mathbf{r}, t)$ as the wave function describing the state of a single particle and $\psi^*(\mathbf{r}, t)\psi(\mathbf{r}, t)d^3\mathbf{r}$ as the probability of finding the particle in volume $d^3\mathbf{r}$. Dirac's interpretation of negative energy state with the help of the negative energy sea rules out the possibility of existence of a single isolated electron. Strictly speaking, to each positive energy electron, we should associate a wave function $\psi = \psi(\mathbf{r}, \mathbf{r}_1, \mathbf{r}_2 ...)$, where $\mathbf{r}_1, \mathbf{r}_2 ...$ are position vectors of the electrons in the accompanying negative energy sea. This makes the interpretation of $\psi^\dagger \psi$ as the probability density questionable. This problem is resolved in quantum field theory, where $\psi^\dagger \psi$ is interpreted as an operator in Fock space for the charge density, taking into account the particle and antiparticles.

16.13 SPIN ANGULAR MOMENTUM OF AN ELECTRON

In non-relativistic quantum mechanics, the Schrödinger equation describes particles only with orbital angular momentum. Within non-relativistic quantum mechanics, spin angular momentum has to be accommodated in an ad hoc way. The Schrödinger equation has to be modified to get Pauli equation, which describes a non-relativistic particle with spin.

Spin angular momentum naturally emerges from the Dirac equation. We know that angular momentum is conserved quantity. Any theory has to accommodate conservation of angular momentum. One of the important requirements of a theory is that all the basic conservation laws like energy, momentum and angular momentum have to emerge as a consequence of the theory. So if the Dirac equation is claimed as the equation describing electron, then one should get the conservation of angular momentum as a consequence of that equation.

So our immediate task is to check whether the angular momentum is conserved for a particle described by the Dirac equation. Within the structure of quantum mechanics, a dynamical variable A is conserved if $[A, H] = 0$, provided A has no explicit time dependence.

Is Orbital Angular Momentum Conserved?

The operator for orbital angular momentum is

$$\mathbf{L}_{op} = \mathbf{r}_{op} \times \mathbf{p}_{op}$$

Let us check the z component of the orbital angular momentum, which is given by

$$L_z = xp_y - yp_x$$

Note that here x, p_y, y and p_x are operators.

The Dirac Hamiltonian for a free particle is

$$\begin{aligned} H_D &= c\boldsymbol{\alpha}\cdot\mathbf{p} + \beta mc^2 \\ &= c\alpha_x p_x + c\alpha_y p_y + c\alpha_z p_z + \beta mc^2 \end{aligned}$$

Let us evaluate $[L_z, H_D]$.

$$\begin{aligned} [L_z, H_D] &= [L_z, c\alpha_x p_x + c\alpha_y p_y + c\alpha_z p_z + \beta mc^2] \\ &= [L_z, c\alpha_x p_x] + [L_z, c\alpha_y p_y] + [L_z, c\alpha_z p_z] + [L_z, \beta mc^2] \end{aligned} \tag{16.71}$$

$$[L_z, c\alpha_x p_x] = [xp_y - yp_x, c\alpha_x p_x] = [xp_y, c\alpha_x p_x] - [yp_x, c\alpha_x p_x] \tag{16.72}$$

$$[xp_y, -c\alpha_x p_x] = xp_y c\alpha_x p_x - c\alpha_x p_x xp_y \tag{16.73}$$

In the RHS, only x and p_x are non-commuting; i.e., only their order in each term has to be preserved as it is.

$$\therefore [xp_y, c\alpha_x p_x] = cp_y \alpha_x [xp_x - p_x x] = i\hbar c p_y \alpha_x$$

This second term in the Equation (16.72) is

$$[yp_y, c\alpha_x p_x] = yp_x c\alpha_z p_z - c\alpha_x p_x yp_x$$

All the quantities in the RHS commute among themselves, and so it is zero.

i.e.,
$$[yp_x, c\alpha_x p_x] = 0$$

$$\therefore [L_z, c\alpha_x p_x] = i\hbar c\alpha_x p_y \tag{16.74}$$

Next, let us evaluate $[L_z, c\alpha_y p_y]$.

$$\begin{aligned}[L_z, c\alpha_y p_y] &= [xp_y - yp_x, c\alpha_y p_y] \\ &= [xp_y, c\alpha_y p_y] - [yp_x, c\alpha_y p_y] = -yp_x c\alpha_y p_y + c\alpha_y p_y yp_x \\ &= -c\alpha_y p_x [yp_y - p_y y] = i\hbar c\alpha_y p_x \end{aligned} \tag{16.75}$$

Working out in the same way, the third and the fourth terms in the Equation (16.71) can be shown to be zero.

$$\therefore [L_z, H_D] = i\hbar c[\alpha_x p_y - \alpha_y p_x]$$

$$\therefore [L_z, H_D] \neq 0 \tag{16.76}$$

So if the Dirac equation is taken as the correct equation for electron, the orbital angular momentum is not conserved.

New Operator – Spin Angular Momentum

Let us now see how the spin angular momentum emerges naturally from the Dirac equation and restore the conservation of angular momentum in the theory. Classically, the angular momentum of any system has to be described by terms $\mathbf{r} \times \mathbf{p}$. This is true whether the object is revolving about an axis or spinning about its own axis. If we restrict ourselves to the angular momentum $\mathbf{r} \times \mathbf{p}$ alone, then the Dirac equation predicts violation of angular momentum. The surprising thing is that we can construct a new operator which behaves like the angular momentum. This new angular momentum is the spin angular momentum.

The spin angular momenta S_x, S_y and S_z are defined as

$$S_x = -i\frac{\hbar}{2}\alpha_y \alpha_z, \qquad S_y = -i\frac{\hbar}{2}\alpha_z \alpha_x \quad \text{and} \quad S_z - i\frac{\hbar}{2}\alpha_x \alpha_y \tag{16.77}$$

Note that these operators have neither $\mathbf{r}$ dependence nor $\mathbf{p}$ dependence. So there is no quantity in classical mechanics from which we can construct these operators. Spin angular momentum is unique to quantum mechanics. Before investigating their properties, let us check whether the spin angular momentum is conserved.

Is S_z Conserved

Let us evaluate $[S_z, H_D]$.

$$[S_z, H_D] = [S_z, c\alpha_x p_x + c\alpha_y p_y + c\alpha_z p_z + \beta mc^2]$$

$$= [S_z, c\alpha_x p_x] + [S_z, c\alpha_y p_y] + [S_z, c\alpha_z p_z] + [S_z, \beta mc^2] \tag{16.78}$$

In evaluating these commutation relations, we will make use of the following property: if *A* and *B* are anti-commuting matrices, flipping them will introduce a negative sign; i.e.,

$$AB = -BA$$

$$[S_z, c\alpha_x p_x] = \left[-\frac{i\hbar}{2}\alpha_x\alpha_y, c\alpha_x p_x\right]$$

$$= -\frac{i\hbar}{2} cp_x[\alpha_x\alpha_y\alpha_x - \alpha_x\alpha_x\alpha_y]$$

Note that since α_x and α_y are anti-commuting matrices,

$$\alpha_x\alpha_y = -\alpha_y\alpha_x$$

$$\therefore \alpha_x\alpha_y\alpha_x = -\alpha_x\alpha_x\alpha_y = -\alpha_y (\because \alpha_x^2 = 1)$$

$$\therefore [S_z, c\alpha_x p_x] = -\frac{i\hbar}{2} cp_x(-2\alpha_y) = \hbar cp_x\alpha_y \tag{16.79}$$

Next, let us evaluate $[S_z, c\alpha_y p_y]$.

$$[S_z, c\alpha_y p_y] = \left[-\frac{i\hbar}{2}\alpha_x\alpha_y, c\alpha_y p_y\right]$$

$$= -\frac{i\hbar cp_y}{2}[\alpha_x\alpha_y\alpha_y - \alpha_y\alpha_x\alpha_y]$$

Now we have $\alpha_y\alpha_x\alpha_y = -\alpha_y\alpha_y\alpha_x = -\alpha_x$

$$\therefore [S_z, c\alpha_y p_y] = -i\hbar cp_y\alpha_x. \tag{16.80}$$

Working out in the same way, we can show that the third and the last term in the Equation (16.78) are zero.

$$[S_z, H_D] = i\hbar c[\alpha_y p_x - \alpha_x p_y] \tag{16.81}$$

$$\therefore [S_z, H_D] \neq 0 \tag{16.82}$$

So if we consider the spin angular momentum alone, it is also not conserved for a particle described by the Dirac equation.

Conservation of Total Angular Momentum

Now we have two kinds of angular momenta, neither of which is individually conserved.

Let us define the total angular momentum **J** by

$$\mathbf{J} = \mathbf{L} + \mathbf{S}$$

or

$$J_z = L_z + S_z \tag{16.83}$$

Let us check whether **J** is conserved. As mentioned earlier, it is enough to check whether J_z is conserved.

$$\begin{aligned}[J_z, H_D] &= [L_z + S_z, H_D] \\ &= [L_z, H_D] + [S_z, H_D] \\ &= i\hbar c[\alpha_x p_y - \alpha_y p_x] + i\hbar c[\alpha_y p_x - \alpha_x p_y] \\ &= 0\end{aligned} \tag{16.84}$$

Therefore, the total angular momentum is conserved. So the conservation of angular momentum for a Dirac particle requires the existence of spin angular momentum.

Properties of Spin Operators

1. First, let us check the operators S_x, S_y and S_z satisfy the angular momentum algebra.
   Let us evaluate $[S_x, S_y]$.

$$\begin{aligned}[S_x, S_y] &= \left[-\frac{i\hbar}{2}\alpha_y\alpha_z, -\frac{i\hbar}{2}\alpha_z\alpha_x\right] \\ &= \left(-\frac{i\hbar}{2}\right)^2 [\alpha_y\alpha_z\alpha_z\alpha_x - \alpha_z\alpha_x\alpha_y\alpha_z] \\ &= \left(-\frac{i\hbar}{2}\right)^2 [\alpha_y\alpha_x - \alpha_z\alpha_z\alpha_x\alpha_y] \\ &= \left(-\frac{i\hbar}{2}\right)^2 [\alpha_y\alpha_x - \alpha_x\alpha_y] = \left(-\frac{i\hbar}{2}\right)\left(-\frac{i\hbar}{2}\right)(-2\alpha_x\alpha_y) \\ &= i\hbar S_z\end{aligned} \tag{16.85}$$

   So S_x, S_y, and S_z satisfy the angular momentum algebra.

2. $S_x^2 = S_y^2 = S_z^2 = \dfrac{\hbar^2}{4} I$ (16.86)

   This can be easily checked.

$$\begin{aligned}S_z^2 &= \left(-i\frac{\hbar}{2}\alpha_x\alpha_y\right)\left(-i\frac{\hbar}{2}\alpha_x\alpha_y\right) \\ &= -\frac{\hbar^2}{4}\alpha_x\alpha_y\alpha_x\alpha_y = \frac{\hbar^2}{4}\alpha_x\alpha_x\alpha_y\alpha_y = \frac{\hbar^2}{4}\end{aligned}$$

3. Eigenvalues

   The eigenvalue equations for S^2 and S_z are

$$S^2\chi = s(s+1)\hbar^2\chi$$

and

$$S_z\chi = m_s\hbar\chi \tag{16.87}$$

Since

$$S_x^2 = S_y^2 = S_z^2 = \frac{1}{4}\hbar^2 I,$$

we have

$$S^2\chi = (S_x^2 + S_y^2 + S_z^2)\chi = \frac{3}{4}\hbar^2\chi = \frac{1}{2}\left(\frac{1}{2}+1\right)\hbar^2\chi \tag{16.88}$$

Immediately, we identity that $s = \frac{1}{2}$.

Starting from the Equation (16.87), we have

$$S_z(S_z\chi) = S_z m_s\hbar\chi = m_s\hbar S_z\chi = m_s^2\hbar^2\chi$$

But

$$S_z^2 = \frac{1}{4}\hbar^2 I \;\therefore \frac{1}{4}\hbar^2\chi = m_s^2\hbar^2\chi$$

or

$$m_s^2 = \frac{1}{4} \;\therefore m_s = \pm\frac{1}{2}$$

i.e.,

$$m_s = \frac{1}{2} \text{ or } \frac{1}{2} \tag{16.89}$$

The spin quantum s is $s = \frac{1}{2}$ and z component quantum number $m_s = \pm\frac{1}{2}$.

4. Matrix representation of S_x, S_y and S_z

   Consider S_x

$$S_x = -\frac{i\hbar}{2}\alpha_y\alpha_z = -\frac{i\hbar}{2}\begin{pmatrix} 0 & \sigma_y \\ \sigma_y & 0 \end{pmatrix}\begin{pmatrix} 0 & \sigma_z \\ \sigma_z & 0 \end{pmatrix}$$

$$= -\frac{i\hbar}{2}\begin{pmatrix} \sigma_y\sigma_z & 0 \\ 0 & \sigma_y\sigma_z \end{pmatrix}$$

$$= -\frac{i\hbar}{2}\begin{pmatrix} i\sigma_x & 0 \\ 0 & i\sigma_x \end{pmatrix} = \frac{\hbar}{2}\begin{pmatrix} \sigma_x & 0 \\ 0 & \sigma_x \end{pmatrix} = \frac{\hbar}{2}\Sigma_x$$

where

$$\Sigma_x = \begin{pmatrix} \sigma_x & 0 \\ 0 & \sigma_x \end{pmatrix}$$

So we have

$$S_x = \frac{\hbar}{2}\Sigma_x = \frac{\hbar}{2}\begin{pmatrix} \sigma_x & 0 \\ 0 & \sigma_x \end{pmatrix} \tag{16.90}$$

$$S_y = \frac{\hbar}{2}\Sigma_y = \frac{\hbar}{2}\begin{pmatrix}\sigma_y & 0 \\ 0 & \sigma_y\end{pmatrix} \tag{16.91}$$

and
$$S_z = \frac{\hbar}{2}\Sigma_z = \frac{\hbar}{2}\begin{pmatrix}\sigma_z & 0 \\ 0 & \sigma_z\end{pmatrix} \tag{16.92}$$

Example 16.1 Show that

$$(\mathbf{\Sigma}\cdot\mathbf{A})(\mathbf{\Sigma}\cdot\mathbf{B}) = \mathbf{A}\cdot\mathbf{B} + i\mathbf{\Sigma}\cdot(\mathbf{A}\times\mathbf{B})$$

Solution: We know that

$$(\boldsymbol{\sigma}\cdot\mathbf{A})(\boldsymbol{\sigma}\cdot\mathbf{B}) = \mathbf{A}\cdot\mathbf{B} + i\boldsymbol{\sigma}\cdot(\mathbf{A}\times\mathbf{B})$$

$$(\mathbf{\Sigma}\cdot\mathbf{A})(\mathbf{\Sigma}\cdot\mathbf{B}) = \begin{pmatrix}\boldsymbol{\sigma}\cdot\mathbf{A} & 0 \\ 0 & \boldsymbol{\sigma}\cdot\mathbf{A}\end{pmatrix}\begin{pmatrix}\boldsymbol{\sigma}\cdot\mathbf{B} & 0 \\ 0 & \boldsymbol{\sigma}\cdot\mathbf{B}\end{pmatrix} = \begin{pmatrix}\boldsymbol{\sigma}\cdot\mathbf{A}\,\boldsymbol{\sigma}\cdot\mathbf{B} & 0 \\ 0 & \boldsymbol{\sigma}\cdot\mathbf{A}\,\boldsymbol{\sigma}\cdot\mathbf{B}\end{pmatrix}$$

$$= \begin{pmatrix}\mathbf{A}\cdot\mathbf{B}I_2 + i\boldsymbol{\sigma}\cdot(\mathbf{A}\times\mathbf{B}) & 0 \\ 0 & \mathbf{A}\cdot\mathbf{B}I_2 + i\boldsymbol{\sigma}\cdot(\mathbf{A}\cdot\mathbf{B})\end{pmatrix}$$

$$= \mathbf{A}\cdot\mathbf{B}I + i\mathbf{\Sigma}\cdot\mathbf{A}\times\mathbf{B}.$$

Here I_2 is 2×2 unit matrix and I is 4×4 matrix. Generally, I is not written in the final expression.

16.14 HELICITY OPERATOR

Let us introduce the concept of helicity, which is the projection of the spin angular momentum along the direction of the momentum. Classically, if **S** is a vector representing the spin angular momentum, then

$$\Lambda = \frac{\mathbf{S}\cdot\mathbf{p}}{|\mathbf{p}|}$$

Geometrically, this is shown in Fig.16.8.

In quantum mechanics, the helicity operator of an electron is defined as

$$\hat{\Lambda} = \frac{\mathbf{S}\cdot\mathbf{p}_{op}}{|\mathbf{p}|} = \frac{\hbar}{2}\frac{\mathbf{\Sigma}\cdot\mathbf{p}_{op}}{|\mathbf{p}|} \tag{16.93}$$

It is interesting to note that the spin angular momentum $\hbar/2\Sigma$ does not commute with the Dirac Hamiltonian H_D, but $\hat{\Lambda} = \frac{\hbar}{2}\frac{\mathbf{\Sigma}\cdot\mathbf{p}_{op}}{|\mathbf{p}|}$ commutes with the Hamiltonian H_D.

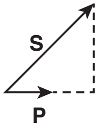


Fig.16.8 Projection of **S** along **P**

i.e., $$\left[\frac{\hbar}{2}\boldsymbol{\Sigma}, H_D\right] \neq 0 \text{ but } \left[\frac{\hbar}{2}\frac{\boldsymbol{\Sigma}\cdot\mathbf{p}_{op}}{|\mathbf{p}|}, H_D\right] = 0 \qquad (16.94)$$

The commutation relation can be checked easily.

Let us first write H_D and Λ in the matrix form.

$$\Lambda = \frac{\hbar}{2}\frac{\boldsymbol{\Sigma}\cdot\mathbf{p}_{op}}{|\mathbf{p}|} = \frac{\hbar}{2\,|\mathbf{p}|}\begin{bmatrix}\boldsymbol{\sigma}\cdot\mathbf{p}_{op} & 0\\ 0 & \boldsymbol{\sigma}\cdot\mathbf{p}_{op}\end{bmatrix}$$

$$\therefore \Lambda H_D - H_D\Lambda = \frac{\hbar}{2\,|\mathbf{p}|}\left\{\begin{bmatrix} mc^2 & c\boldsymbol{\sigma}\cdot\mathbf{p}_{op}\\ c\boldsymbol{\sigma}\cdot\mathbf{p}_{op} & -mc^2\end{bmatrix}\begin{bmatrix}\boldsymbol{\sigma}\cdot\mathbf{p}_{op} & 0\\ 0 & \boldsymbol{\sigma}\cdot\mathbf{p}_{op}\end{bmatrix}\right.$$

$$\left. - \begin{bmatrix}\boldsymbol{\sigma}\cdot\mathbf{p}_{op} & 0\\ 0 & \boldsymbol{\sigma}\cdot\mathbf{p}_{op}\end{bmatrix}\cdot\begin{bmatrix} mc^2 & c\boldsymbol{\sigma}\cdot\mathbf{p}_{op}\\ c\boldsymbol{\sigma}\cdot\mathbf{p}_{op} & -mc^2\end{bmatrix}\right\}$$

$$= 0$$

Eigenvalues of Helicity Operator

$$\Lambda\psi = \frac{\hbar}{2\,|\mathbf{p}|}\boldsymbol{\Sigma}\cdot\mathbf{p}_{op}\psi = \lambda\psi$$

Since ψ is also an eigenstate $\mathbf{p}_{op}$, we can write $\mathbf{p}_{op}\psi = \mathbf{p}\,\psi$

$$\therefore \Lambda\psi = \frac{\hbar}{2\,|\mathbf{p}|}\boldsymbol{\Sigma}\cdot\mathbf{p}\,\psi = \lambda\psi$$

$$\Lambda\Lambda\psi = \left[\frac{\hbar}{2\,|\mathbf{p}|}\right]^2 (\boldsymbol{\Sigma}\cdot\mathbf{p})(\boldsymbol{\Sigma}\cdot\mathbf{p})\,\psi$$

$$= \frac{\hbar^2}{4\,|\mathbf{p}|^2}|\mathbf{p}|^2\,\psi = \lambda^2\psi$$

$$\therefore \lambda = \pm\frac{\hbar}{2} \qquad (16.95)$$

Sometimes $\hat{\Lambda}$ is defined as $\hat{\Lambda} = \dfrac{\boldsymbol{\Sigma}\cdot\mathbf{p}}{|\mathbf{p}|}$. The advantage of this definition is that the eigenvalues are ±1.

Compatible Operators for the Dirac Equation

The complete set of compatible observables for the Dirac equation is momentum $\mathbf{p}_{op}$, Hamiltonian H_D and the helicity operator Λ.

i.e., $$[\mathbf{p}, H_D] = 0 \quad [\mathbf{p}, \Lambda] = 0 \quad [\Lambda, H_D] = 0 \qquad (16.96)$$

Obviously, the quantum numbers which characterize the solution for the Dirac equation for free particles are the eigenvalues of momentum, Hamiltonian and the helicity.

Example 16.2 Show that for a Dirac electron, $\mathbf{S}\cdot\mathbf{n}$ does not commute with free particle Hamiltonian H_D unless $\mathbf{n}$ is a unit vector in the direction of momentum.

Solution:

$$\begin{aligned}
[\mathbf{S}.\mathbf{n}, H_D] &= \left[\frac{\hbar}{2}\boldsymbol{\Sigma}\cdot\mathbf{n}, H_D\right] \\
&= \frac{\hbar}{2}\left\{\begin{pmatrix} \boldsymbol{\sigma}\cdot\mathbf{n} & 0 \\ 0 & \boldsymbol{\sigma}\cdot\mathbf{n}\end{pmatrix}\begin{pmatrix} mc^2 & c\boldsymbol{\sigma}\cdot\mathbf{p} \\ c\boldsymbol{\sigma}\cdot\mathbf{p} & -mc^2\end{pmatrix} - \begin{pmatrix} mc^2 & c\boldsymbol{\sigma}\cdot\mathbf{p} \\ c\boldsymbol{\sigma}\cdot\mathbf{p} & -mc^2\end{pmatrix}\begin{pmatrix} \boldsymbol{\sigma}\cdot\mathbf{n} & 0 \\ 0 & \boldsymbol{\sigma}\cdot\mathbf{n}\end{pmatrix}\right\} \\
&= \frac{\hbar}{2}\begin{pmatrix} 0 & c(\boldsymbol{\sigma}\cdot\mathbf{n})(\boldsymbol{\sigma}\cdot\mathbf{p}) - c(\boldsymbol{\sigma}\cdot\mathbf{p})(\boldsymbol{\sigma}\cdot\mathbf{n}) \\ c(\boldsymbol{\sigma}\cdot\mathbf{n})(\boldsymbol{\sigma}\cdot\mathbf{p}) - c(\boldsymbol{\sigma}\cdot\mathbf{p})(\boldsymbol{\sigma}\cdot\mathbf{n}) & 0\end{pmatrix} \\
&= \frac{\hbar}{2}\begin{pmatrix} 0 & ci\boldsymbol{\sigma}\cdot(\mathbf{n}\times\mathbf{p} - \mathbf{p}\times\mathbf{n}) \\ ci\boldsymbol{\sigma}\cdot(\mathbf{n}\times\mathbf{p} - \mathbf{p}\times\mathbf{n}) & 0\end{pmatrix}
\end{aligned}$$

This is not zero unless $\mathbf{n}$ and $\mathbf{p}$ are parallel or antiparallel to each other.

Example 16.3 Determine the free particle Dirac spinors which are also the eigenstates of helicity operator.

Solution: Consider the positive energy eigenspinor given by

$$u = N\begin{pmatrix} \chi \\ \dfrac{\boldsymbol{\sigma}.\mathbf{p}\chi}{E_p + mc^2}\end{pmatrix}$$

The eigenvalue equation for the helicity operator is

$$\Lambda u = \lambda u$$

$$\frac{\hbar}{2}\begin{pmatrix} \dfrac{\boldsymbol{\sigma}\cdot\mathbf{p}}{p} & 0 \\ 0 & \dfrac{\boldsymbol{\sigma}\cdot\mathbf{p}}{p}\end{pmatrix} N\begin{pmatrix} \chi \\ \dfrac{\boldsymbol{\sigma}\cdot\mathbf{p}\chi}{E_p + mc^2}\end{pmatrix} = N\frac{\hbar}{2}\begin{pmatrix} \dfrac{\boldsymbol{\sigma}\cdot\mathbf{p}}{p}\chi \\ \dfrac{\dfrac{\boldsymbol{\sigma}\cdot\mathbf{p}}{p}\boldsymbol{\sigma}\cdot\mathbf{p}\chi}{E_p + mc^2}\end{pmatrix} = \pm\frac{\hbar}{2}N\begin{pmatrix} \chi \\ \dfrac{\boldsymbol{\sigma}\cdot\mathbf{p}\chi}{E_p + mc^2}\end{pmatrix}$$

This is possible, provided we have $\dfrac{\boldsymbol{\sigma}\cdot\mathbf{p}}{p}\chi = \pm\chi$.

It is easy to solve this equation. After some algebra, we find

$$\chi_1 = \left(\frac{p+p_z}{2p}\right)^{1/2}\begin{pmatrix} 1 \\ \dfrac{p_x+ip_y}{p+p_z}\end{pmatrix} \text{and } \chi_2 = \left(\frac{p+p_z}{2p}\right)^{1/2}\begin{pmatrix} -\dfrac{p_x-ip_y}{p+p_z} \\ 1\end{pmatrix}$$

16.15 MAGNETIC MOMENT OF AN ELECTRON DUE TO SPIN ANGULAR MOMENTUM

One of the remarkable successes of the Dirac equation is that it leads to correct value for the magnetic moment of an electron due to the spin angular momentum apart from a small correction. Let us recapture some of our earlier discussions of the magnetic moment and the angular momentum in Chapter 12. In classical electrodynamics, a charge $-e$ with an angular momentum **L** has a magnetic moment μ given by

$$\mu = -\frac{e}{2mc}\mathbf{L}$$

The same ideas can be extended to quantum mechanics. For an electron, we define the magnetic moment due to orbital angular momentum as

$$\boldsymbol{\mu}_{op} = -\frac{e}{2mc}\mathbf{L}_{op} = -\frac{e\hbar}{2mc}\frac{\mathbf{L}_{op}}{\hbar} = -\mu_B\frac{\mathbf{L}_{op}}{\hbar}$$

and $$\mu_z = -\frac{e}{2mc}L_{z,op} = -\frac{e\hbar}{2mc}\cdot\frac{L_{z,op}}{\hbar} = -\frac{\mu_B}{\hbar}L_{z,op}$$

Extending the same considerations for magnetic moment, due to the spin angular momentum, we expect

$$\boldsymbol{\mu}_{op} = -\frac{e}{2mc}\mathbf{S}_{op} = -\frac{e\hbar}{2mc}\frac{\mathbf{S}_{op}}{\hbar} = -\mu_B\frac{\mathbf{S}_{op}}{\hbar}$$

and μ_z to be given as

$$\mu_z = -m_s\mu_B$$

where $$m_s = \pm\frac{1}{2}$$

However, this is not in agreement with experimental values, which are almost twice this value. So we have to come with new definition for magnetic moment due to the spin angular momentum. The magnetic moments $\boldsymbol{\mu}_{op}$ and $\boldsymbol{\mu}_{z,op}$ for the spin angular momentum are defined as

$$\boldsymbol{\mu}_{op} = g_s\mu_B\frac{1}{\hbar}\mathbf{S}_{op} \tag{16.97}$$

and $$\boldsymbol{\mu}_{z,op} = -g_s\mu_B\frac{1}{\hbar}S_z \tag{16.98}$$

The experimental value for g_s is found to be $g_s = 2.002$. The Dirac equation leads to $g_s = 2$, and the remaining factors are accounted in quantum electrodynamics. But to obtain $g_s = 2$ is itself a remarkable achievement for any theory.

Let us now obtain the values of g_s from the Dirac equation.

For an electron in a magnetic field **B**, the Dirac equation is

$$i\hbar\frac{\partial\psi}{\partial t} = c\boldsymbol{\alpha}\cdot\left(\mathbf{p}+\frac{e\mathbf{A}}{c}\right)\psi + \beta mc^2\psi$$

Again, let us consider the stationary state solution $\psi(\mathbf{r},t) = u(\mathbf{r})e^{-iEt/\hbar}$

Substituting this function in the Dirac equation, we get

$$Eu(\mathbf{r})e^{-iEt/\hbar} = c\boldsymbol{\alpha}\cdot\left(\mathbf{p}+\frac{e\mathbf{A}}{c}\right)u(\mathbf{r})e^{-iEt/\hbar} + \beta mc^2u(\mathbf{r})e^{-iEt/\hbar}$$

Factoring out $e^{-iEt/\hbar}$ we get

$$Eu(\mathbf{r}) = c\boldsymbol{\alpha}\cdot\left(\mathbf{p}+\frac{e\mathbf{A}}{c}\right)u(\mathbf{r}) + \beta mc^2u(\mathbf{r})$$

Let us write $$u = \begin{pmatrix} V \\ W \end{pmatrix}$$

Then, we have

$$E\begin{pmatrix} V \\ W \end{pmatrix} = \left[\begin{pmatrix} 0 & c\boldsymbol{\sigma}\cdot\left(\mathbf{p}+\frac{c\mathbf{A}}{c}\right) \\ c\boldsymbol{\sigma}\cdot\left(\mathbf{p}+\frac{c\mathbf{A}}{c}\right) & 0 \end{pmatrix} + \begin{pmatrix} mc^2 & 0 \\ 0 & -mc^2 \end{pmatrix}\right]\begin{bmatrix} V \\ W \end{bmatrix}$$

$$EV = mc^2V + c\boldsymbol{\sigma}\cdot\left(\mathbf{p}+\frac{e\mathbf{A}}{c}\right)W \tag{16.99}$$

and $$EW = -mc^2W + c\boldsymbol{\sigma}\cdot\left(\mathbf{p}+\frac{e\mathbf{A}}{c}\right)V \tag{16.100}$$

Rewriting these two equations, we get

$$(E - mc^2)V = c\boldsymbol{\sigma}\cdot\left(\mathbf{p}+\frac{e\mathbf{A}}{c}\right)W \tag{16.101}$$

and $$(E + mc^2)W = c\boldsymbol{\sigma}\cdot\left(\mathbf{p}+\frac{e\mathbf{A}}{c}\right)V \tag{16.102}$$

Now let us make the non-relativistic approximation. In the non-relativistic approximation,

$$E \sim E' + mc^2$$

Then, $$(E - mc^2)\,V \sim E'V$$

and $$(E + mc^2)\,W \sim 2mc^2W$$

Now the Equations (16.101) and (16.102) become

$$E'V = c\boldsymbol{\sigma}\cdot\left(\mathbf{p}+\frac{e\mathbf{A}}{c}\right)W \tag{16.103}$$

and

$$2mc^2W = c\boldsymbol{\sigma}\cdot\left(\mathbf{p}+\frac{e\mathbf{A}}{c}\right)V \tag{16.104}$$

Substituting the Equation (16.104) in the Equation (16.103) for W, we get

$$E'V = \frac{1}{2m}\boldsymbol{\sigma}\cdot\left(\mathbf{p}+\frac{e\mathbf{A}}{c}\right)\boldsymbol{\sigma}\cdot\left(\mathbf{p}+\frac{e\mathbf{A}}{c}\right)V \tag{16.105}$$

$$= \frac{1}{2m}\left(\mathbf{p}+\frac{e\mathbf{A}}{c}\right)\cdot\left(\mathbf{p}+\frac{e\mathbf{A}}{c}\right)V + \frac{1}{2m}i\boldsymbol{\sigma}\cdot\left(\mathbf{p}+\frac{e\mathbf{A}}{c}\right)\times\left(\mathbf{p}+\frac{e\mathbf{A}}{c}\right)V \tag{16.106}$$

Consider the second term in the above equation.

$$\left(\mathbf{p}+\frac{e\mathbf{A}}{c}\right)\times\left(\mathbf{p}+\frac{e\mathbf{A}}{c}\right)V$$

$$= \left(-i\hbar\nabla+\frac{e\mathbf{A}}{c}\right)\times\left(-i\hbar\nabla+\frac{e\mathbf{A}}{c}\right)V$$

$$= \left(-i\hbar\nabla+\frac{e\mathbf{A}}{c}\right)\times\left(-i\hbar\nabla V+\frac{e\mathbf{A}V}{c}\right)$$

$$= -i\hbar\nabla\times(-i\hbar\nabla V)+\frac{e^2}{c^2}\mathbf{A}\times\mathbf{A}V - i\hbar\frac{e}{c}\nabla\times(\mathbf{A}V) - i\hbar\frac{e}{c}\mathbf{A}\times\nabla V \tag{16.107}$$

The first two terms are zero. In the third term, we use $\nabla\times(\mathbf{A}V) = (\nabla\times\mathbf{A})V - \mathbf{A}\times\nabla V$.

Therefore, we have

$$\left(\mathbf{p}+\frac{e\mathbf{A}}{c}\right)\times\left(\mathbf{p}+\frac{e\mathbf{A}}{c}\right)V$$

$$= -\frac{ie\hbar}{c}(\nabla\times\mathbf{A})V + \frac{ie\hbar}{c}\mathbf{A}\times\nabla V - \frac{ie\hbar}{c}\mathbf{A}\times\nabla V$$

$$= -\frac{ie\hbar}{c}(\nabla\times\mathbf{A})V = -\frac{ie\hbar}{c}\mathbf{B}V \ (\because \nabla\times\mathbf{A} = \mathbf{B}) \tag{16.108}$$

Therefore, the Equation (16.105) becomes

$$E'V = \frac{1}{2m}\left(\mathbf{p}-\frac{e\mathbf{A}}{c}\right)^2 + \frac{1}{2m}\cdot\frac{e\hbar}{c}\boldsymbol{\sigma}\cdot\mathbf{B}V \tag{16.109}$$

Making of the fact that the spin angular momentum $\mathbf{S} = \frac{\hbar}{2}\boldsymbol{\sigma}$, we get

$$E'V = \frac{1}{2m}\cdot\left(\mathbf{p}-\frac{e\mathbf{A}}{c}\right)^2 + \frac{e\hbar}{2mc}.2.\frac{\mathbf{S}}{\hbar}\cdot\mathbf{B}$$

$$= \frac{1}{2m}\cdot\left(\mathbf{p}-\frac{e\mathbf{A}}{c}\right)^2 - \boldsymbol{\mu}_{op}\cdot\mathbf{B}$$

provided we identify the magnetic moment μ_{op} as

$$\boldsymbol{\mu} = -2 \cdot \mu_B \cdot \frac{\mathbf{S}}{\hbar} \tag{16.110}$$

This is the magnetic moment due to the spin angular momentum. Comparing the Equations (16.97) and (16.110), we get

$$g_s = 2 \tag{16.111}$$

16.16 HYDROGEN ATOM

In non-relativistic quantum mechanics, it is very simple to get the radial wave equation for hydrogen atom. In the case of the Dirac equation, we have to do a long background work to arrive at the radial wave equation. We, first, define a new operator *K* which commutes with the Dirac Hamiltonian and study its property in details. Next, we introduce parity operator. Then, we define radial momentum p_r and radial matrix α_r, and we get the Hamiltonian in terms of these quantities. The Dirac Hamiltonian, the *K* operator, the parity operator and the operators J^2 and J_z together form a set of compatible observables. This helps us to write general form of Dirac spinor for hydrogen atom and finally arrive at the radial wave equations.

16.16.1 *K* Operator

To solve the Dirac equation for hydrogen atom, we have to define a new operator *K* operator. It is defined as

$$K = (\boldsymbol{\Sigma} \cdot \mathbf{L} + \hbar)\beta \tag{16.112}$$

Why do we need do define a new operator? Let us recapitulate the salient features of solution to the Schrödinger equation for hydrogen atom in non-relativistic quantum mechanics. The Hamiltonian *H* is

$$H = -\frac{\hbar^2}{2m}\nabla^2 - \frac{Ze^2}{r}$$

This Hamiltonian commutes with L^2 and L_z

i.e., $$[H, L^2] = 0 \quad \text{and} \quad [H, L_z] = 0$$

Therefore, the eigenfunction of *H* should also be the eigenfunctions of L^2 and L_z. It is important to note that ∇^2 is given by

$$\nabla^2 = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2 \frac{\partial}{\partial r}\right) - \frac{1}{\hbar^2 r^2} L^2 \tag{16.113}$$

So the Hamiltonian can be written in terms of L^2.

i.e., $$H = \frac{\hbar^2}{2m}\left[\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2 \frac{\partial}{\partial r}\right) - \frac{1}{\hbar^2 r^2} L^2\right] + V(r) \tag{16.114}$$

It becomes obvious now that the eigenfunctions of H are also the eigenfunctions of L^2 and L_z. The fact that H can be written in terms of L^2 enables us to write $u(\mathbf{r}) = R_{nl}(r)Y_{lm}(\theta,\varphi)$, where $L^2u = l(l+1)\hbar^2 u$, $L_z\, u = m\hbar u$ and $H_u = E_n u$.

The Dirac Hamiltonian for hydrogen atom is

$$H_D = c\boldsymbol{\alpha}\cdot\mathbf{p} + \beta mc^2 + V(r) \tag{16.115}$$

where
$$V(r) = -\frac{Ze^2}{r}$$

It is easy to check that $[H_D, J^2] = 0$ and $[H_D, J_z] = 0$. Therefore, the eigenfunctions of H_D should be the eigenfunctions of J^2 and J_z. This demand is very much similar to non-relativistic quantum mechanics. However, unlike non-relativistic case, we cannot write H_D in terms of J^2 (similar to the Equations (16.114) and (16.113)). So it becomes very difficult to construct an eigenfunction common to both H_D and J^2. It is this difficulty which suggests a new operator K related to J^2 to be defined. The K operator should be as follows:

1. $[H_D, K] = 0$
2. $[K, J_z] = 0$
3. H_D should be expressed in terms of K. (16.116)

Let us now try to 'obtain' the operator K. Let us recall that the spin angular momentum $\mathbf{S}$ is given by

$$\mathbf{S} = \frac{\hbar}{2}\boldsymbol{\Sigma} \text{ where } \boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\sigma} & 0 \\ 0 & \boldsymbol{\sigma} \end{pmatrix}$$

$\mathbf{S}^2$ is given by

$$S^2 = S_x^2 + S_y^2 + S_z^2 = \frac{\hbar^2}{4}\boldsymbol{\Sigma}_x^2 + \frac{\hbar^2}{4}\boldsymbol{\Sigma}_y^2 + \frac{\hbar^2}{4}\boldsymbol{\Sigma}_x^2$$

$$= \frac{3}{4}\hbar^2 I \text{ (since } \boldsymbol{\Sigma}_x^2 = \boldsymbol{\Sigma}_y^2 = \boldsymbol{\Sigma}_z^2 = I\text{)}$$

We know that

$$\boldsymbol{\Sigma}\cdot\mathbf{A}\,\boldsymbol{\Sigma}\cdot\mathbf{B} = \mathbf{A}\cdot\mathbf{B} + i\boldsymbol{\Sigma}\cdot(\mathbf{A}\times\mathbf{B})$$

Making use of this result, we get

$$\boldsymbol{\Sigma}\cdot\mathbf{L}\,\boldsymbol{\Sigma}\cdot\mathbf{L} = L^2 + i\Sigma\cdot(\mathbf{L}\times\mathbf{L})$$

$$= \mathbf{L}^2 - \hbar\Sigma\cdot\mathbf{L} \qquad \text{(Since } \mathbf{L}\times\mathbf{L} = i\hbar\mathbf{L}\text{)}$$

or
$$\mathbf{L}^2 = (\boldsymbol{\Sigma}\cdot\mathbf{L})^2 + \hbar\boldsymbol{\Sigma}\cdot\mathbf{L} \tag{16.117}$$

Now J^2 is given by

$$\mathbf{J}^2 = (\mathbf{L}+\mathbf{S})^2 = \mathbf{L}^2 + \mathbf{S}^2 + 2\mathbf{L}\cdot\mathbf{S}$$

$$= \mathbf{L}^2 + \frac{3}{4}\hbar^2 I + 2\left(\frac{\hbar\Sigma}{2}\cdot\mathbf{L}\right)$$

$$= \mathbf{L}^2 + \frac{3}{4}\hbar^2 + \hbar\boldsymbol{\Sigma}\cdot\mathbf{L} \tag{16.118}$$

$$= (\mathbf{\Sigma}\cdot\mathbf{L})^2 + 2\hbar\mathbf{\Sigma}\cdot\mathbf{L} + \frac{3}{4}\hbar^2 \quad \text{(from(16.117))} \tag{16.119}$$

Adding $\frac{1}{4}\cdot\hbar^2$ to both sides, we get

$$J^2 + \frac{1}{4}\hbar^2 = (\mathbf{\Sigma}\cdot\mathbf{L})^2 + 2\hbar(\mathbf{\Sigma}\cdot\mathbf{L})^2 + \hbar^2$$

$$= (\mathbf{\Sigma}\cdot\mathbf{L} + \hbar)^2$$

$$= [(\mathbf{\Sigma}\cdot\mathbf{L} + \hbar)\beta]^2 \text{ since } \beta^2 = I$$

K is defined as $K = (\mathbf{\Sigma}\cdot\mathbf{L} + \hbar)\beta$ and so we have

$$K^2 = J^2 + \frac{1}{4}\hbar^2 \tag{16.120}$$

It is obvious that K^2 commutates with J^2.

It can be shown that H_D can be expressed in terms of K.

Example 16.4 Show that the operator $K = \beta(\Sigma\cdot\mathbf{L} + \hbar)$ commutes with the Dirac Hamiltonian for a particle in central potential.

Solution:

$$K = \beta(\Sigma\cdot\mathbf{L} + \hbar)$$

$$\therefore K = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix}\begin{pmatrix} \boldsymbol{\sigma}.\mathbf{L} + \hbar & 0 \\ 0 & \boldsymbol{\sigma}.\mathbf{L} + \hbar \end{pmatrix} = \begin{pmatrix} \boldsymbol{\sigma}.\mathbf{L} + \hbar I & 0 \\ 0 & -\boldsymbol{\sigma}.\mathbf{L} - \hbar I \end{pmatrix}$$

$$[H_D, K] = [\mathbf{c}\boldsymbol{\alpha}\cdot\mathbf{p} + \beta mc^2 + V(r), K]$$

$$= [\mathbf{c}\boldsymbol{\alpha}\cdot\mathbf{p}, K] + [\beta mc^2, K] + [V(r), K]$$

$$[c\boldsymbol{\alpha}\cdot\mathbf{p}, K] = c\boldsymbol{\alpha}\cdot\mathbf{p}K - cK\boldsymbol{\alpha}\cdot\mathbf{p}$$

$$= \begin{bmatrix} 0 & c\boldsymbol{\sigma}\cdot\mathbf{p} \\ c\boldsymbol{\sigma}\cdot\mathbf{p} & 0 \end{bmatrix}\begin{bmatrix} \boldsymbol{\sigma}\cdot\mathbf{L} + \hbar & 0 \\ 0 & -\boldsymbol{\sigma}\cdot\mathbf{L} - \hbar \end{bmatrix}$$

$$- \begin{bmatrix} \boldsymbol{\sigma}\cdot\mathbf{L} + \hbar & 0 \\ 0 & -\boldsymbol{\sigma}\cdot\mathbf{L} - \hbar \end{bmatrix}\begin{bmatrix} 0 & c\boldsymbol{\sigma}\cdot\mathbf{p} \\ c\boldsymbol{\sigma}\cdot\mathbf{p} & 0 \end{bmatrix}$$

$$= \begin{bmatrix} 0 & -c\boldsymbol{\sigma}\cdot\mathbf{p}\boldsymbol{\sigma}\cdot\mathbf{L} - c\hbar\boldsymbol{\sigma}\cdot\mathbf{p} \\ c\boldsymbol{\sigma}\cdot\mathbf{p}\boldsymbol{\sigma}\cdot\mathbf{L} + c\hbar\boldsymbol{\sigma}\cdot\mathbf{p} & 0 \end{bmatrix}$$

$$- \begin{bmatrix} 0 & c\boldsymbol{\sigma}\cdot\mathbf{L}\boldsymbol{\sigma}\cdot\mathbf{p} + c\hbar\boldsymbol{\sigma}\cdot\mathbf{p} \\ -c\boldsymbol{\sigma}\cdot\mathbf{L}\boldsymbol{\sigma}\cdot\mathbf{p} - c\hbar\boldsymbol{\sigma}\cdot\mathbf{p} & 0 \end{bmatrix}$$

$$= \begin{bmatrix} 0 & -c(\boldsymbol{\sigma}\cdot\mathbf{p}\boldsymbol{\sigma}\cdot\mathbf{L}+\boldsymbol{\sigma}\cdot\mathbf{L}\boldsymbol{\sigma}\cdot\mathbf{p})-2c\hbar\boldsymbol{\sigma}\cdot\mathbf{p} \\ c(\boldsymbol{\sigma}\cdot\mathbf{L}\boldsymbol{\sigma}\cdot\mathbf{L}+\boldsymbol{\sigma}\cdot\mathbf{L}\boldsymbol{\sigma}\cdot\mathbf{L})+2c\hbar\boldsymbol{\sigma}\cdot\mathbf{p} & 0 \end{bmatrix}$$

$$\boldsymbol{\sigma}\cdot\mathbf{p}\boldsymbol{\sigma}\cdot\mathbf{L}+\boldsymbol{\sigma}\cdot\mathbf{L}\boldsymbol{\sigma}\cdot\mathbf{p} = \mathbf{p}\cdot\mathbf{L}+i\boldsymbol{\sigma}\cdot(\mathbf{p}\times\mathbf{L})+\mathbf{L}\cdot\mathbf{p}+i\boldsymbol{\sigma}\cdot(\mathbf{L}\times\mathbf{p})$$

$$= \mathbf{p}\cdot\mathbf{L}+\mathbf{p}\cdot\mathbf{L}+i\boldsymbol{\sigma}\cdot(\mathbf{p}\times\mathbf{L}+\mathbf{L}\times\mathbf{p})$$

It is easy to verify that the first two terms are zero.

Let us now evaluate the x component of $(\mathbf{p}\times\mathbf{L}+\mathbf{L}\times\mathbf{p})_x$.

$$(\mathbf{p}\times\mathbf{L}+\mathbf{L}\times\mathbf{p})_x = p_y L_z - p_z L_y + L_y p_z - L_z p_y$$

$$= p_y(xp_y - yp_x) - p_z(zp_x - xp_z) + (zp_x - xp_z)p_z - (xp_y - yp_x)p_y$$

$$= (-p_y y + yp_y)p_x + (-p_z z + zp_z)p_x = 2i\hbar p_x$$

Similarly, evaluations of other components give

$$\boldsymbol{\sigma}\cdot(\mathbf{p}\times\mathbf{L}+\mathbf{L}\times\mathbf{p}) = 2i\hbar\sigma_x p_x + 2i\hbar\sigma_y p_y + 2i\hbar\sigma_z p_z = 2i\hbar\boldsymbol{\sigma}.\mathbf{p}$$

$$[c\boldsymbol{\alpha}\cdot\mathbf{p}, K] = \begin{bmatrix} 0 & -i2i\hbar c\boldsymbol{\sigma}\cdot\mathbf{p}-2c\hbar\boldsymbol{\sigma}\cdot\mathbf{p} \\ i2i\hbar c\boldsymbol{\sigma}\cdot\mathbf{p}+2c\hbar\boldsymbol{\sigma}\cdot\mathbf{p} & 0 \end{bmatrix} = 0$$

It is easy to evaluate the other two terms.

$$[\beta mc^2, K] = \begin{bmatrix} mc^2 & 0 \\ 0 & -mc^2 \end{bmatrix}\begin{bmatrix} \boldsymbol{\sigma}\cdot\mathbf{L}+\hbar & 0 \\ 0 & -\boldsymbol{\sigma}\cdot\mathbf{L}-\hbar \end{bmatrix}$$

$$-\begin{bmatrix} \boldsymbol{\sigma}\cdot\mathbf{L}+\hbar & 0 \\ 0 & -\boldsymbol{\sigma}\cdot\mathbf{L}-\hbar \end{bmatrix}\begin{bmatrix} mc^2 & 0 \\ 0 & -mc^2 \end{bmatrix} = 0$$

$$[V(r), K] = 0$$

$$\therefore [H_D, K] = 0 \tag{16.121}$$

Example 16.5 Determine the eigenvalues operator K.

Solution: The eigenvalue equation is

$$K\psi = k\hbar\psi$$

where ψ is a four-component spinor.

$$K^2\psi = KK\psi = K(k\hbar\psi)$$

$$K^2\psi = KK\psi = K(k\hbar\psi) = k^2\hbar^2\psi$$

But

$$K^2 = J^2 + \frac{1}{4}\hbar^2$$

$$\therefore K^2\psi = \left(J^2 + \frac{1}{4}\hbar^2\right)\psi$$

$$= J^2\psi + \frac{1}{4}\hbar\psi = j(j+1)\hbar^2\psi + \frac{1}{4}\hbar^2\psi$$

$$= \left(j^2 + j + \frac{1}{4}\right)\hbar^2\psi = \left(j + \frac{1}{2}\right)^2 \hbar^2\psi$$

$$\therefore k^2 = \left(j + \frac{1}{2}\right)^2 \tag{16.122}$$

$$\therefore k = \pm\left(j + \frac{1}{2}\right) = \pm 1, \pm 2, \pm 3 \cdots \tag{16.123}$$

Example 16.6 Determine the orbital angular momentum quantum numbers of upper and lower components *V* and *W* of eigenvectors of *K*.

Solution: Consider the eigenvalue equation

$$K\begin{pmatrix} V \\ W \end{pmatrix} = k\hbar\begin{pmatrix} V \\ W \end{pmatrix}$$

The *K* operator is given by

$$K = (\boldsymbol{\sigma}\cdot\mathbf{L} + \hbar)\beta = \begin{pmatrix} \boldsymbol{\sigma}\cdot\mathbf{L}+\hbar & 0 \\ 0 & \boldsymbol{\sigma}\cdot\mathbf{L}+\hbar \end{pmatrix}\begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} = \begin{pmatrix} \boldsymbol{\sigma}\cdot\mathbf{L}+\hbar & 0 \\ 0 & -\boldsymbol{\sigma}\cdot\mathbf{L}-\hbar \end{pmatrix}$$

$$\therefore K\begin{pmatrix} V \\ W \end{pmatrix} = \begin{bmatrix} \boldsymbol{\sigma}\cdot\mathbf{L}+\hbar & 0 \\ 0 & -\boldsymbol{\sigma}\cdot\mathbf{L}-\hbar \end{bmatrix}\begin{bmatrix} V \\ W \end{bmatrix} = k\hbar\begin{bmatrix} V \\ W \end{bmatrix}$$

$$\boldsymbol{\sigma}\cdot\mathbf{L}V + \hbar V = k\hbar V \text{ or } \boldsymbol{\sigma}\cdot\mathbf{L}V = (k-1)\hbar V \tag{16.124}$$

$$-\boldsymbol{\sigma}\cdot\mathbf{L}W - \hbar W = k\hbar W \text{ or } \boldsymbol{\sigma}.\mathbf{L}W = -(k+1)\hbar W \tag{16.125}$$

The eigenvector of *K* should also be the eigenvector of J^2.

$$\therefore J^2\begin{pmatrix} V \\ W \end{pmatrix} = j(j+1)\hbar^2\begin{pmatrix} V \\ W \end{pmatrix}$$

But

$$J^2 = \mathbf{L}^2 + S^2 + 2\mathbf{S}\cdot\mathbf{L} = \mathbf{L}^2 + \frac{3}{4}\hbar^2 + \hbar\boldsymbol{\Sigma}\cdot\mathbf{L}$$

$$= \begin{pmatrix} \mathbf{L}^2 + \frac{3}{4}\hbar^2 + \boldsymbol{\sigma}.\mathbf{L} & 0 \\ 0 & \mathbf{L}^2 + \frac{3}{4}\hbar^2 + \boldsymbol{\sigma}.\mathbf{L} \end{pmatrix}$$

$$\therefore J^2\begin{pmatrix} V \\ W \end{pmatrix} = \begin{bmatrix} \mathbf{L}^2 + \frac{3}{4}\hbar^2 + \boldsymbol{\sigma}.\mathbf{L}\hbar & 0 \\ 0 & \mathbf{L}^2 + \frac{3}{4}\hbar^2 + \hbar\boldsymbol{\sigma}.\mathbf{L} \end{bmatrix}\begin{bmatrix} V \\ W \end{bmatrix} = j(j+1)\hbar^2\begin{bmatrix} V \\ W \end{bmatrix}$$

$$\therefore \left(\mathbf{L}^2 + \frac{3}{4}\hbar^2 + \hbar\boldsymbol{\sigma}.\mathbf{L}\right)V = j(j+1)\hbar^2 V \tag{16.126}$$

and

$$\left(\mathbf{L}^2 + \frac{3}{4}\hbar^2 + \hbar\boldsymbol{\sigma}.\mathbf{L}\right)W = j(j+1)\hbar W \tag{16.127}$$

Eigenvalue l for V

Making use of the Equation (16.124) in the Equation (16.126), we get

$$L^2 V = \left[\, j(j+1)\hbar^2 - \frac{3}{4}\hbar^2 - (k-1)\hbar^2 \,\right]V$$

Using $\mathbf{L}^2\, V = l(l+1)\hbar^2 V$ in the above equation, we get

$$l(l+1) = j(j+1) - \frac{3}{4} - (k-1)$$

$$= \left(j^2 + j + \frac{1}{4}\right)\hbar^2\psi = \left(j + \frac{1}{2}\right)^2 \hbar^2\psi$$

Adding $\frac{1}{4}$ to both, we get

$$l^2 + l + \frac{1}{4} = j^2 + j + \frac{1}{2} - k$$

or

$$\left(l + \frac{1}{2}\right)^2 = j^2 + j + \frac{1}{2} - k$$

For $k = j + \frac{1}{2}$

$$\left(l + \frac{1}{2}\right)^2 = j^2 + j + \frac{1}{2} - \left(j + \frac{1}{2}\right) = j^2$$

$$\therefore l + \frac{1}{2} = j \text{ or } l = j - \frac{1}{2}$$

For $k = -\left(j + \frac{1}{2}\right)$

$$\left(l + \frac{1}{2}\right)^2 = j^2 + j + \frac{1}{2} + \left(j + \frac{1}{2}\right) = (j+1)^2$$

$$\left(l + \frac{1}{2}\right) = j + 1 \text{ or } l = j + \frac{1}{2}$$

Eigenvalue l for W

Working in the same way, we get

for $$k = j+\frac{1}{2}, \quad l = j+\frac{1}{2}$$

and for $$k = -1\left(j+\frac{1}{2}\right), \quad l = j-\frac{1}{2}.$$

Summarizing, we have

$$K\begin{pmatrix} V \\ W \end{pmatrix} = k\hbar\begin{pmatrix} V \\ W \end{pmatrix}$$

$$L^2V = l(l+1)\hbar^2V: \quad \begin{cases} l = j-\frac{1}{2} & \text{for} \quad k = j+\frac{1}{2} \\ l = j+\frac{1}{2} & \text{for} \quad k = -\left(j+\frac{1}{2}\right) \end{cases}$$

$$L^2W = l(l+1)\hbar W: \quad \begin{cases} l = j+\frac{1}{2} & \text{for} \quad k = j+\frac{1}{2} \\ l = j-\frac{1}{2} & \text{for} \quad k = -\left(j+\frac{1}{2}\right) \end{cases} \tag{16.128}$$

16.6.2 Parity Operator

We have already seen the concept parity invariance in Chapter 9. Let us recall the space reversal of the coordinates is defined by

$$\mathbf{r} \to \mathbf{r}' = -\mathbf{r}$$

The invariance of parity transformation of the Dirac equation means that for every state $\psi(\mathbf{r},t)$ satisfying the Dirac equation, there exists a corresponding state $\psi'(-\mathbf{r},t)$, which also satisfies the Dirac equation. This implies that there should exist a transformation P such that

$$\psi'(-\mathbf{r},t) = P\psi(\mathbf{r},t)$$

Finding such transformation amounts to proving the parity invariance of the Dirac equation. Here, we quote only the end result without giving the proof. The parity operator P is given by

$$P\psi(\mathbf{r},t) = \beta\, P_0\psi(\mathbf{r},t) = \beta\psi(-\mathbf{r},t) \tag{16.129}$$

Normally, the determination of this operator is a part of the proof of covariance of the Dirac equation. Note that the effect of the operator P_0 on $\psi(\mathbf{r},t)$ is to replace $\mathbf{r}$ by $-\mathbf{r}$ in the function $\psi(\mathbf{r},t)$.

$$P_0\psi(\mathbf{r},t) = \psi(-\mathbf{r},t)$$

Our interest on this operator in this chapter lies with the fact that P commutes with the Dirac Hamiltonian for a particle in a central potential, and therefore, we have to choose that eigenstate of H_D which is also an eigenstate of P.

Example 16.7 Prove that $P\psi(\mathbf{r},t) = \beta P_0\psi(\mathbf{r},t) = \beta\psi(-\mathbf{r},t)$ satisfies the Dirac equation when $\psi(\mathbf{r},t)$ satisfies the Dirac equation.

Solution: The Dirac equation for $\psi(\mathbf{r},t)$ is

$$i\hbar\frac{\partial}{\partial t}\psi(\mathbf{r},t) = -ic\hbar\boldsymbol{\alpha}.\nabla\,\psi(\mathbf{r},t) + mc^2\beta\psi(\mathbf{r},t)$$

Now operate P on either side of the equation.

$$\begin{aligned}
i\hbar\frac{\partial}{\partial t}P\psi(\mathbf{r},t) &= -ic\hbar P(\boldsymbol{\alpha}.\nabla\,\psi(\mathbf{r},t)) + mc^2 P\beta\psi(\mathbf{r},t)\\
i\hbar\frac{\partial}{\partial t}\beta P_0\psi(\mathbf{r},t) &= -ic\hbar\beta P_0(\boldsymbol{\alpha}.\nabla\,\psi(\mathbf{r},t)) + mc^2\beta P_0\beta\psi(\mathbf{r},t)\\
i\hbar\frac{\partial}{\partial t}\beta\psi(-\mathbf{r},t) &= -ic\hbar\beta\boldsymbol{\alpha}.(-\nabla\,\psi(-\mathbf{r},t)) + mc^2\beta\beta\psi(-\mathbf{r},t)
\end{aligned}$$

Note that the effect of P_0 is to change $\mathbf{r}$ *by* $-\mathbf{r}$ in all the functions on which it acts, and so ∇ becomes $-\nabla$. Using the fact $\beta\boldsymbol{\alpha} = -\boldsymbol{\alpha}\beta$, we get

$$i\hbar\frac{\partial}{\partial t}\underbrace{\beta\psi(-\mathbf{r},t)} = -ic\hbar\boldsymbol{\alpha}.\nabla\,\underbrace{\beta\psi(-\mathbf{r},t)} + mc^2\beta\,\underbrace{\beta\psi(-\mathbf{r},t)}$$

Therefore, $\beta\psi(-\mathbf{r},t)$ satisfies the Dirac equation.

Example 16.8 Show that the parity operator defined by $P\psi(\mathbf{r},t) = \beta P_o\psi(\mathbf{r},t) = \beta\psi(-\mathbf{r},t)$ commutes with the Dirac Hamiltonian $H_D = c\boldsymbol{\alpha}\cdot\mathbf{p}_{op} + \beta mc^2 + V(r)$.

Solution: The effect of P_o is to replace $\mathbf{r}$ by $-\mathbf{r}$ in all expressions on which P_o acts.

i.e., $$P_o\psi(\mathbf{r},t) = \psi(-\mathbf{r},t)$$

$$P_o(\mathbf{p}_{op}\psi(\mathbf{r},t)) = P_o(-i\hbar\nabla\,\psi(\mathbf{r},t)) = i\hbar\nabla\,\psi(-\mathbf{r},t) = -\mathbf{p}_{op}\psi(-\mathbf{r},t)$$

Let us consider an arbitrary Dirac spinor $\psi(\mathbf{r},t)$.

$$\begin{aligned}
[H_D, P]\psi(\mathbf{r},t) &= [H_D P - PH_D]\psi(\mathbf{r},t)\\
&= [c\boldsymbol{\alpha}.\mathbf{p}_{op} + \beta mc^2 + V(r)]\beta P_o\psi(\mathbf{r},t)\\
&\quad -\beta P_o[c\boldsymbol{\alpha}.\mathbf{p}_{op} + \beta mc^2 + V(r)]\psi(\mathbf{r},t)\\
&= c\boldsymbol{\alpha}.\mathbf{p}_{op}\beta\psi(-\mathbf{r},t) + mc^2\psi(-\mathbf{r},t) + V(r)\beta\psi(-\mathbf{r},t)\\
&\quad -\beta(-c\boldsymbol{\alpha}.\mathbf{p}_{op}\psi(-\mathbf{r},t)) - mc^2\psi(-\mathbf{r},t) - V(r)\beta\psi(-\mathbf{r},t)\\
&= c\boldsymbol{\alpha}\cdot\mathbf{p}_{op}\beta\psi(-\mathbf{r},t) + c\beta\boldsymbol{\alpha}\cdot\mathbf{p}_{op}\psi(-\mathbf{r},t)
\end{aligned}$$

Since $\boldsymbol{\alpha}$ and β are anti-commuting matrices,

$$\boldsymbol{\alpha}.\mathbf{p}_{op}\beta = -\beta\boldsymbol{\alpha}.\mathbf{p}_{op}$$

$$[H_D, P]\psi(\mathbf{r},t) = c\boldsymbol{\alpha}.\mathbf{p}_{op}\beta\psi(-\mathbf{r},t) - c\boldsymbol{\alpha}.\mathbf{p}_{op}\beta\psi(-\mathbf{r},t) = 0$$

Since this equation is true for any arbitrary Dirac spinor $\psi(\mathbf{r},t)$, we conclude

$$[H_D, P] = 0$$

16.6.3 Radial Momentum $\mathbf{P_r}$ and Radial Matrix α_r and Hamiltonian

Let us first define α_r and p_r.

Since $\boldsymbol{\alpha}$ and $\mathbf{r}$ are commuting, the definition of the radial component of $\boldsymbol{\alpha}_r$ is straightforward.

$$\boldsymbol{\alpha}.\mathbf{r} = \boldsymbol{\alpha}.\frac{\mathbf{r}}{r} \tag{16.130}$$

In the case of p_r, in classical physics, $\mathbf{p}.\mathbf{r}$ and $\mathbf{r}.\mathbf{p}$ are the same. However, since they are non-commuting operators in quantum mechanics, we define p_r as

$$p_r = \frac{1}{2}\left(\mathbf{p}.\frac{\mathbf{r}}{r} + \frac{\mathbf{r}}{r}.\mathbf{p}\right) \tag{16.131}$$

From the Example 8.3 in Chapter 8, we have

$$p_r = \frac{1}{r}(\mathbf{r}\cdot\mathbf{p} - \hbar)$$

Let us consider $(\boldsymbol{\alpha}\cdot\mathbf{r})^{-1}$. It is obvious $(\boldsymbol{\alpha}\cdot\mathbf{r})^{-1} = \frac{1}{r^2}(\boldsymbol{\alpha}\cdot\mathbf{r})$. We can easily check this result.

$$(\boldsymbol{\alpha}\cdot\mathbf{r})^{-1}(\boldsymbol{\alpha}\cdot\mathbf{r}) = \frac{1}{r^2}(\boldsymbol{\alpha}\cdot\mathbf{r})(\boldsymbol{\alpha}\cdot\mathbf{r}) = \frac{1}{r^2}r^2 = 1$$

Next consider $\boldsymbol{\alpha}\cdot\mathbf{r}\boldsymbol{\alpha}\cdot\mathbf{p}$.

$$\begin{aligned}\boldsymbol{\alpha}\cdot\mathbf{r}\boldsymbol{\alpha}\cdot\mathbf{p} &= \mathbf{r}.\mathbf{p} + i\boldsymbol{\Sigma}\cdot(\mathbf{r}\times\mathbf{p})\\ &= \mathbf{r}\cdot\mathbf{p} + i\boldsymbol{\Sigma}\cdot\mathbf{L}\end{aligned} \tag{16.132}$$

From the definition of K_{op}, we get

$$\beta(\boldsymbol{\Sigma}\cdot\mathbf{L} + \hbar) = K$$

or

$$\boldsymbol{\Sigma}\cdot\mathbf{L} + \hbar = \beta K \text{ or } \boldsymbol{\Sigma}\cdot\mathbf{L} = \beta K - \hbar$$

$$\therefore \boldsymbol{\alpha}\cdot\mathbf{r}\boldsymbol{\alpha}\cdot\mathbf{p} = \mathbf{r}\cdot\mathbf{p} + i(\beta K - \hbar) \tag{16.133}$$

Multiply this equation by $\frac{1}{r^2}\boldsymbol{\alpha}\cdot\mathbf{r}$ on either side.

$$\frac{1}{r^2}\boldsymbol{\alpha}\cdot\mathbf{r}\boldsymbol{\alpha}\cdot\mathbf{r}\boldsymbol{\alpha}\cdot\mathbf{p} = \frac{\boldsymbol{\alpha}\cdot\mathbf{r}}{r^2}[\mathbf{r}\cdot\mathbf{p} - i\hbar + i\beta K]$$

$$\therefore\ \boldsymbol{\alpha}\cdot\mathbf{p} = \frac{\boldsymbol{\alpha}\cdot\mathbf{r}}{r}\frac{1}{r}[\mathbf{r}\cdot\mathbf{p} - i\hbar] + \frac{\alpha_r}{r}i\beta K \qquad \left(\text{Since } \frac{1}{r^2}\boldsymbol{\alpha}.\mathbf{r}\boldsymbol{\alpha}.\mathbf{r} = 1\right)$$

$$= \alpha_r p_r + i\frac{1}{r}\alpha_r\beta K$$

The Dirac Hamiltonian H_D is given by

$$H_D = c\alpha_r p_r + ic\frac{1}{r}\alpha_r\beta K + \beta mc^2 - \frac{Ze^2}{r} \tag{16.134}$$

It can be shown that K commutes with H_D.

Example 16.9 Show that $[H_D, K] = 0$ where H_D is given by

$$H_D = c\alpha_r p_r + ic\frac{1}{r}\alpha_r\beta K + \beta mc^2 - \frac{Ze^2}{r}$$

Solution:

$$[H_D, K] = [c\alpha_r p_r, K] + ic\left[\frac{1}{r}\alpha_r\beta K, K\right] + [\beta mc^2, K] - \left[\frac{Ze^2}{r}, K\right]$$

$$[c\alpha_r p_r, K] = c[\alpha_r, K]p_r + c\alpha_r[p_r, K]$$

$$[\alpha_r, K] = \begin{bmatrix} 0 & \frac{1}{r}\boldsymbol{\sigma}\cdot\mathbf{r} \\ \frac{1}{r}\boldsymbol{\sigma}\cdot\mathbf{r} & 0 \end{bmatrix}\begin{bmatrix} \boldsymbol{\sigma}\cdot\mathbf{L} + \hbar & 0 \\ 0 & -\boldsymbol{\sigma}\cdot\mathbf{L} - \hbar \end{bmatrix}$$

$$- \begin{bmatrix} \boldsymbol{\sigma}\cdot\mathbf{L} + \hbar & 0 \\ 0 & -\boldsymbol{\sigma}\cdot\mathbf{L} - \hbar \end{bmatrix}\begin{bmatrix} 0 & \frac{1}{r}\boldsymbol{\sigma}\cdot\mathbf{r} \\ \frac{1}{r}\boldsymbol{\sigma}\cdot\mathbf{r} & 0 \end{bmatrix}$$

$$= \begin{bmatrix} 0 & -\frac{1}{r}\boldsymbol{\sigma}\cdot\mathbf{r}\boldsymbol{\sigma}\cdot\mathbf{L} - \boldsymbol{\sigma}\cdot\mathbf{L}\frac{1}{r}\boldsymbol{\sigma}\cdot\mathbf{r} - 2\hbar\boldsymbol{\sigma}\cdot\mathbf{r} \\ \frac{1}{r}\boldsymbol{\sigma}\cdot\mathbf{r}\boldsymbol{\sigma}\cdot\mathbf{L} + \boldsymbol{\sigma}\cdot\mathbf{L}\frac{1}{r}\boldsymbol{\sigma}\cdot\mathbf{r} + 2\hbar\boldsymbol{\sigma}\cdot\mathbf{r} & \end{bmatrix} \tag{16.135}$$

Consider $$\frac{1}{r}\boldsymbol{\sigma}\cdot\mathbf{r}\boldsymbol{\sigma}\cdot\mathbf{L} + \boldsymbol{\sigma}\cdot\mathbf{L}\frac{1}{r}\boldsymbol{\sigma}\cdot\mathbf{r}$$

Since $\mathbf{L}$ is an operator containing θ and ϕ only, $\boldsymbol{\sigma}.\mathbf{L}$ and $\frac{1}{r}$ commute.

Therefore, we have

$$\frac{1}{r}[\boldsymbol{\sigma}\cdot\mathbf{r}\boldsymbol{\sigma}\cdot\mathbf{L} + \boldsymbol{\sigma}\cdot\mathbf{L}\boldsymbol{\sigma}\cdot\mathbf{r}] = \frac{1}{r}[\underbrace{\mathbf{r}\cdot\mathbf{L}}_{0} + \underbrace{\mathbf{L}\cdot\mathbf{r}}_{0} + i\boldsymbol{\sigma}\cdot(\mathbf{r}\times\mathbf{L} + \mathbf{L}\times\mathbf{r})]$$

Consider the x component of $\mathbf{r}\times\mathbf{L}+\mathbf{L}\times\mathbf{r}$.

$$\begin{aligned}[\mathbf{r}\times\mathbf{L}+\mathbf{L}\times\mathbf{r}]_x &= [yL_z - zL_y + L_y z - L_z y]\\ &= [y, L_z] + [z, L_y] = 2i\hbar x\end{aligned}$$

Taking into account the other components also, we get

$$\begin{aligned}\therefore \frac{1}{r}[\boldsymbol{\sigma}\cdot\mathbf{r}\boldsymbol{\sigma}\cdot\mathbf{L}+\boldsymbol{\sigma}\cdot\mathbf{L}\boldsymbol{\sigma}\cdot\mathbf{r}] &= i\sigma_x 2i\hbar x + i\sigma_y 2i\hbar y + i\sigma_z i\hbar y\\ &= -2\hbar\boldsymbol{\sigma}\cdot\mathbf{r}\end{aligned}$$

Using these results in the Equation (16.135), we get

$$[\alpha_r, K] = 0$$

Let us evaluate $[p_r, K]$ next.

$$\begin{aligned}[p_r, K] &= \left[\frac{1}{r}(\mathbf{r}\cdot\mathbf{p}-i\hbar), K\right] = \left[\frac{1}{r}\mathbf{r}\cdot\mathbf{p}, K\right]\\ &= \frac{1}{r}\mathbf{r}\cdot\mathbf{p}\begin{bmatrix}\boldsymbol{\sigma}\cdot\mathbf{L}+\hbar & 0\\ 0 & -\boldsymbol{\sigma}\cdot\mathbf{L}-\hbar\end{bmatrix} - \begin{bmatrix}\boldsymbol{\sigma}\cdot\mathbf{L}+\hbar & 0\\ 0 & -\boldsymbol{\sigma}\cdot\mathbf{L}-\hbar\end{bmatrix}\frac{1}{r}\mathbf{r}\cdot\mathbf{p}\\ &= \begin{bmatrix}\frac{1}{r}\mathbf{r}\cdot\mathbf{p}\boldsymbol{\sigma}\cdot\mathbf{L}-\boldsymbol{\sigma}\cdot\mathbf{L}\frac{1}{r}\mathbf{r}\cdot\mathbf{p} & 0\\ 0 & -\frac{1}{r}\mathbf{r}\cdot\mathbf{p}\boldsymbol{\sigma}\cdot\mathbf{L}+\boldsymbol{\sigma}\cdot\mathbf{L}\frac{1}{r}\mathbf{r}\cdot\mathbf{p}\end{bmatrix}\end{aligned}$$

$$\begin{aligned}\frac{1}{r}\mathbf{r}\cdot\mathbf{p}\boldsymbol{\sigma}\cdot\mathbf{L}-\boldsymbol{\sigma}\cdot\mathbf{L}\frac{1}{r}\mathbf{r}\cdot\mathbf{p} &= \frac{1}{r}(\mathbf{r}\cdot\mathbf{p}\boldsymbol{\sigma}\cdot\mathbf{L}-\boldsymbol{\sigma}\cdot\mathbf{L}\mathbf{r}\cdot\mathbf{p})\\ &= \frac{1}{r}\left\{[\mathbf{r}\cdot\mathbf{p}, \sigma_x L_x] + [\mathbf{r}\cdot\mathbf{p}, \sigma_y L_y] + [\mathbf{r}\cdot\mathbf{p}, \sigma_z L_z]\right\}\end{aligned}$$

$$\begin{aligned}[\mathbf{r}\cdot\mathbf{p}, \sigma_x L_x] &= \sigma_x[\mathbf{r}\cdot\mathbf{p}, L_x]\\ &= \sigma_x[xp_x + yp_y + zp_z, L_x]\\ &= \sigma_x[xp_x, L_x] + \sigma_x[yp_y, L_x] + \sigma_x[zp_z, L_x]\\ &= \sigma_x\left\{y[p_y, L_x] + [y, L_x]p_y + z[p_z, L_x] + [z, L_z]p_z\right\}\\ &= \sigma_x[-i\hbar yp_z - i\hbar zp_y + i\hbar zp_y + i\hbar xp_z]\\ &= 0\end{aligned}$$

Similarly, it can be shown $[\mathbf{r}\cdot\mathbf{p}, \sigma_y L_y] = [\mathbf{r}\cdot\mathbf{p}, \sigma_z L_z] = 0$.

Consider $\left[\frac{1}{r}\alpha_r\beta K, K\right]$

This commutator is zero since each factor separately commutes with K. i.e., $\frac{1}{r}$, α_r, and β commute with K.

$$\therefore \left[\frac{1}{r}\alpha_r \beta K, K\right] = 0$$

Similarly, it can be shown $[\beta mc^2, K] = 0$ and $\left[-\frac{Ze^2}{r}, K\right] = 0$

$$\therefore\ [H_D, K] = 0$$

16.6.4 General Form of Spinor for Hydrogen Atom

For the Dirac Hamiltonian H_D, the set of operators which commute with H_D are J^2, J_z, K and P.

$$\begin{aligned} &[H_D, J^2] = 0 \quad [H_D, J_Z] = 0 \quad [H_D, P] \quad [H_D, K] = 0 \\ &[J^2, J_Z] = 0 \quad [J^2, P] = 0 \quad [K, P] = 0 \\ &[J_Z, K] = 0 \quad [J_Z, P] = 0 \end{aligned} \tag{16.136}$$

Since the eigenvalue k is $\left(j+\frac{1}{2}\right)$ or $-\left(j+\frac{1}{2}\right)$, only the quantum numbers n, j and m need to be considered.

i.e.,

$$\begin{aligned} H_D\psi_{njm} &= E_n\psi_{njm} \\ J^2\psi_{njm} &= j(j+1)\hbar^2\psi_{njm} \\ J_z\psi_{njm} &= m\hbar\psi_{njm} \end{aligned}$$

For the sake of clarity, we will omit the subscripts n. Therefore, we have

$$\begin{aligned} J^2\psi_{jm} &= j(j+1)\hbar^2\psi_{jm} \\ J_Z\psi_{jm} &= m\hbar\ \psi_{jm} \\ K\psi_{jm} &= \pm\left(j+\frac{1}{2}\right)\hbar\psi_{jm} \\ P\psi_{jm} &= \lambda\psi_{jm}, \text{ with } \lambda = \pm 1 \end{aligned} \tag{16.137}$$

We have already seen that for a given k, the upper component has one l value and the lower component has another l value. Therefore, we can write

$$\psi_{jm} = \begin{pmatrix} \psi_{jl_1m}(\mathbf{r}) \\ \psi_{jl_2m}(\mathbf{r}) \end{pmatrix} \tag{16.138}$$

Since the potential is spherically symmetric, we can write the spinor as a product of radial and angular momentum functions. i.e., ψ_{jm} is written as

$$\psi_{jm} = \begin{pmatrix} \psi_{jl_1m}(\mathbf{r}) \\ \psi_{jl_2m}(\mathbf{r}) \end{pmatrix} = \psi_{jm} = \begin{pmatrix} v(r)\phi_{jl_1m} \\ iw(r)\phi_{jl_2m} \end{pmatrix} \tag{16.139}$$

The value of l_1 and l_2 depends on the value of k. If we choose $k = (j+1/2)$, then $l_1 = j-1/2$ and $l_2 = j+1/2$. If we choose $k = -(j+1/2)$, then $l_1 = j+1/2$ and $l_2 = j-1/2$. This is also consistent with the fact $\mathbf{J} = \mathbf{L}+\mathbf{S}$ or $\mathbf{L} = \mathbf{J}-\mathbf{S}$. For a given value of j, the only possible values of l are $j+1/2$ and $j-1/2$.

Let us consider the effect of P on ψ_{jm}.

$$P\psi_{jm} = \beta P_o \psi_{jm} = \beta \begin{bmatrix} \psi_{jl_1m}(-\mathbf{r}) \\ \psi_{jl_2m}(-\mathbf{r}) \end{bmatrix}$$

$$= \begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix} \begin{bmatrix} \psi_{jl_1m}(-\mathbf{r}) \\ \psi_{jl_2m}(-\mathbf{r}) \end{bmatrix} = \begin{bmatrix} \psi_{jl_1m}(-\mathbf{r}) \\ -\psi_{jl_2m}(-\mathbf{r}) \end{bmatrix}$$

The parity eigenvalue has to be either +1 or −1.

i.e., $$P\psi_{jm} = \psi_{jm} \quad \text{or} \quad P\psi_{jm} = -\psi_{jm}$$

In the first case,

$$P\psi_{jm} = \begin{bmatrix} \psi_{jl_1m}(-\mathbf{r}) \\ -\psi_{jl_2m}(-\mathbf{r}) \end{bmatrix} = \begin{bmatrix} \psi_{jl_1m}(\mathbf{r}) \\ \psi_{jl_2m}(\mathbf{r}) \end{bmatrix}$$

$$\psi_{jl_1m}(-\mathbf{r}) = \psi_{jl_1m}(\mathbf{r}), \psi_{jl_2m}(-\mathbf{r}), = -\psi_{jl_2m}(\mathbf{r}),$$

In the second case,

$$P\psi_{jm} = \begin{bmatrix} \psi_{jl_1m}(-\mathbf{r}) \\ -\psi_{jl_2m}(-\mathbf{r}) \end{bmatrix} = \begin{bmatrix} -\psi_{jl_1m}(\mathbf{r}) \\ -\psi_{jl_2m}(\mathbf{r}) \end{bmatrix}$$

$$\psi_{jl_1m}(-\mathbf{r}) = -\psi_{jl_1m}(\mathbf{r}), \quad \psi_{jl_2m}(-\mathbf{r}) = \psi_{jl_2m}(\mathbf{r})$$

So if

$$P_o\psi_{jl_1m} = \lambda\psi_{jl_1m}, \text{ then } P_o\psi_{jl_2m} = -\lambda\psi_{jl_2m} \tag{16.140}$$

In either case, they have opposite parity.

16.6.5 Radial Wave Equation

Let us use the Hamiltonian H_D given in the Equation (16.134) to develop the radial equations. Since $[H_D, K] = 0$, the operators H_D and K have common eigenstates.

i.e., $$H_D\psi = E\psi \text{ and } K\psi = k\hbar\psi$$

$$\therefore H_D\psi = \left(c\alpha_r p_r + icr^{-1}\alpha_r\beta K + \beta mc^2 - \frac{Ze^2}{r}\right)\psi$$

$$= \left(c\alpha_r p_r + icr^{-1}\alpha_r\beta k\hbar + \beta mc^2 - \frac{Ze^2}{r}\right)\psi$$

The 'new' Hamiltonian is

$$H = c\alpha_r p_r + icr^{-1}\alpha_r\beta k\hbar + \beta mc^2 - \frac{Ze^2}{r}$$

Note σ_r and p_r commute with each other. This is due to the fact that σ_r has no r dependence, whereas p_r contains derivative of r only. Therefore, we can write

$$H = cp_r\alpha_r + icr^{-1}\alpha_r\beta k\hbar + \beta mc^2 - \frac{Ze^2}{r}$$

To proceed further, let us write down the energy eigenvalue equation

$$H\begin{pmatrix}\psi_{jl_1m}\\ \psi_{jl_2m}\end{pmatrix} = E\begin{pmatrix}\psi_{jl_1m}\\ \psi_{jl_2m}\end{pmatrix} \tag{16.141}$$

From the Equation (16.139), we have

$$\psi_{jl_1m} = v(r)\ \phi_{jl_1m} \tag{16.142}$$

and
$$\psi_{jl_2m} = \mathrm{i}\omega(r)\ \phi_{jl_2m} \tag{16.143}$$

where ϕ_{jl_1m} and ϕ_{jl_2m} are the angular momentum parts.

$$\therefore H\begin{pmatrix}\psi_{jl_1m}\\ \psi_{jl_2m}\end{pmatrix} = E\begin{bmatrix}v(r)\phi_{jl_1m}\\ i\omega(r)\phi_{jl_2m}\end{bmatrix}$$

$$K\begin{pmatrix}\psi_{jl_1m}\\ \psi_{jl_2m}\end{pmatrix} = k\hbar\begin{pmatrix}\psi_{jl_1m}\\ \psi_{jl_2m}\end{pmatrix} = k\hbar\begin{bmatrix}v(r)\phi_{jl_1m}\\ i\omega(r)\phi_{jl_2m}\end{bmatrix}$$

$$cp_r\alpha_r\begin{pmatrix}\psi_{jl_1m}\\ \psi_{jl_2m}\end{pmatrix} = cp_r\begin{pmatrix}0 & \sigma_r\\ \sigma_r & 0\end{pmatrix}\begin{pmatrix}v\phi_{jl_1m}\\ i\omega\phi_{jl_2m}\end{pmatrix} = cp_r\begin{pmatrix}\sigma_r i\omega(r)\phi_{jl_2m}\\ \sigma_r v(r)\phi_{jl_1m}\end{pmatrix}$$

It can be shown that (see Examples 16.9 and 16.10)

$$\sigma_r\psi_{jl_2m} = \sigma_r i\omega(r)\phi_{jl_2m} = i\omega(r)\sigma_r\phi_{jl_2m} = -i\omega(r)\phi_{jl_1m}$$

$$\text{and } \sigma_r\psi_{jl_1m} = \sigma_r v(r)\phi_{jl_2m} = v(r)\sigma_r\phi_{jl_1m} = -v(r)\phi_{jl_2m}$$

$$\therefore cp_r\alpha_r\begin{pmatrix}\psi_{jl_1m}\\ \psi_{jl_2m}\end{pmatrix} = cp_r\begin{pmatrix}-i\omega(r)\phi_{jl_1m}\\ -v(r)\phi_{jl_2m}\end{pmatrix} \tag{16.144}$$

Similarly, we have

$$icr^{-1}\alpha_r\beta K\begin{bmatrix}\psi_{jl_1m}\\ \psi_{jl_2m}\end{bmatrix} = icr^{-1}k\hbar\begin{pmatrix}0 & \sigma_r\\ \sigma_r & 0\end{pmatrix}\begin{pmatrix}I & 0\\ 0 & -I\end{pmatrix}\begin{pmatrix}\psi_{jl_1m}\\ \psi_{jl_2m}\end{pmatrix}$$

$$= icr^{-1}k\hbar\begin{pmatrix}0 & \sigma_r\\ \sigma_r & 0\end{pmatrix}\begin{pmatrix}\psi_{jl_1m}\\ -\psi_{jl_2m}\end{pmatrix}$$

$$= icr^{-1}k\hbar\begin{pmatrix}-\sigma_r\psi_{jl_2m}\\ \sigma_r\psi_{jl_1m}\end{pmatrix} = icr^{-1}k\hbar\begin{pmatrix}i\omega(r)\phi_{jl_1m}\\ -v(r)\phi_{jl_2m}\end{pmatrix}$$

Therefore, the Equation (16.141) becomes

$$cp_r\begin{pmatrix}-i\omega\phi_{jl_1m}\\ -v\phi_{jl_2m}\end{pmatrix} + \frac{ic\hbar k}{r}\begin{pmatrix}i\omega\phi_{jl_1m}\\ -v\phi_{jl_2m}\end{pmatrix} + \begin{bmatrix}\left(mc^2 - E - \frac{Ze^2}{r}\right) & 0\\ 0 & -\left(mc^2 + E + \frac{Ze^2}{r}\right)\end{bmatrix}\begin{bmatrix}v\phi_{jl_1m}\\ i\omega\phi_{jl_2m}\end{bmatrix} = 0$$

We can factor out ϕ_{jl_1m} and ϕ_{jl_2m} since they occur as multiplicative factor in all terms. Therefore, we have

$$cp_r\begin{pmatrix}-i\omega\\ -v\end{pmatrix} + \frac{ic\hbar k}{r}\begin{pmatrix}i\omega\\ -v\end{pmatrix} + \begin{pmatrix}\left(mc^2 - E - \frac{Ze^2}{r}\right)v\\ -\left(mc^2 + E + \frac{Ze^2}{r}\right)i\omega\end{pmatrix} = 0 \tag{16.145}$$

In spherical coordinates,

$$\nabla = \mathbf{e}_r\frac{\partial}{\partial r} + \mathbf{e}_\theta\frac{1}{r}\frac{\partial}{\partial\theta} + \mathbf{e}_\phi\frac{1}{r\sin\theta}\frac{\partial}{\partial\phi}$$

$$\mathbf{r}\cdot\mathbf{p} = -i\hbar\mathbf{r}\cdot\nabla = -i\hbar r\frac{\partial}{\partial r}$$

$$\therefore p_r = \frac{1}{r}\mathbf{r}\cdot\mathbf{p} - \frac{i\hbar}{r} = -i\hbar\frac{\partial}{\partial r} = -i\hbar\left(\frac{\partial}{\partial r} + \frac{1}{r}\right)$$

Therefore, from the Equation (16.145), we have

$$-i\hbar c\left(\frac{\partial}{\partial r} + \frac{1}{r}\right)(-i\omega) + \frac{ic\hbar k}{r}i\omega + \left(mc^2 - E - \frac{Ze^2}{r}\right)v = 0 \tag{16.146}$$

and

$$-i\hbar c\left(\frac{\partial}{\partial r} + \frac{1}{r}\right)(-v) + \frac{ic\hbar k}{r}(-v) - \left(mc^2 + E + \frac{Ze^2}{r}\right)i\omega = 0 \tag{16.147}$$

These equations can be written as

$$\left(\frac{\partial}{\partial r} + \frac{1}{r}\right)\omega + \frac{k}{r}\omega - \frac{1}{\hbar c}\left(mc^2 - E - \frac{Ze^2}{r}\right)v = 0 \tag{16.148}$$

$$\left(\frac{\partial}{\partial r} + \frac{1}{r}\right)v - \frac{k}{r}v - \frac{1}{\hbar c}\left(mc^2 + E + \frac{Ze^2}{r}\right)\omega = 0 \tag{16.149}$$

Let us write $v(r) = \dfrac{F(r)}{r}$ and $\omega(r) = \dfrac{G(r)}{r}$.

$$\left(\frac{\partial}{\partial r} + \frac{1}{r}\right)\frac{G}{r} = \frac{1}{r}\frac{dG}{dr} - \frac{1}{r}\frac{G}{r} + \frac{G}{r^2} = \frac{1}{r}\frac{dG}{dr}$$

Similarly we have
$$\left(\frac{\partial}{\partial r}+\frac{1}{r}\right)\frac{F}{r}=\frac{1}{r}\frac{dF}{dr}.$$

Then, the Equations (16.148) and (16.149) become

$$\frac{1}{r}\frac{dG}{dr}+\frac{k}{r}.\frac{G}{r}-\frac{1}{\hbar c}\left(mc^2-E-\frac{Ze^2}{r}\right)\frac{F}{r}=0$$

and

$$\frac{1}{r}\frac{dF}{dr}-\frac{k}{r}\frac{F}{r}-\frac{1}{\hbar c}\left(mc^2+E+\frac{Ze^2}{r}\right)\frac{G}{r}=0$$

or

$$\frac{dF}{dr}-\frac{k}{r}F-\frac{1}{\hbar c}\left(mc^2+E+\frac{Ze^2}{r}\right)G=0 \tag{16.150}$$

$$\frac{dG}{dr}+\frac{k}{r}G-\frac{1}{\hbar c}\left(mc^2-E-\frac{Ze^2}{r}\right)F=0 \tag{16.151}$$

These are the two radial Dirac equations for hydrogen atom.

16.6.6 Solution to Radial Equations

Let us first study the asymptotic behaviour of solutions to the Equations (16.150) and (16.151). As $r\to\infty$, we can drop the terms $\frac{k}{r}F$ and Ze^2/r. Therefore, we have

$$\frac{dF}{dr}-\frac{1}{\hbar c}(mc^2+E)G\simeq 0 \text{ and } \frac{dG}{dr}-\frac{1}{\hbar c}(mc^2-E)\simeq 0 \tag{16.152}$$

Differentiating the above equations, we get

$$\frac{d^2F}{dr^2}-\frac{1}{\hbar c}(mc^2+E)\frac{dG}{dr}=\frac{d^2F}{dr^2}-\frac{1}{(\hbar c)^2}(mc^2+E)(mc^2-E)F\simeq 0 \tag{16.153}$$

Let us define χ, χ_1, and χ_2 as

$$\chi_1=\frac{mc^2+E}{\hbar c},\chi_2=\frac{mc^2-E}{\hbar c} \text{ and } \chi^2=\chi_1\chi_2=\frac{m^2c^4-E^2}{\hbar^2c^2}$$

Then, the Equation (16.153) becomes

$$\frac{d^2F}{dr^2}\simeq\chi^2F$$

The solution of this equation is

$$F=\text{const } e^{\chi r}+\text{const } e^{-\chi r}$$

We would like to have the solution which goes to zero as $r\to\infty$. Therefore, we have

As $r\to\infty$ $\quad F(r)\to\text{const } e^{-\chi r}$

and $G(r)\to\text{const } e^{-\chi r}$

This suggests $F(r)$ and $G(r)$ can be written as

$$F=e^{-\chi r}f \text{ and } G=e^{-\chi r}g$$

Let us define $\rho = \chi r$. Then, the Equations (16.150) and (16.151) become

$$\frac{dF}{d\rho} - k\frac{F}{\rho} - \left(\frac{\chi_1}{\chi} + \frac{Z\alpha}{\rho}\right)G = 0 \tag{16.154}$$

$$\frac{dG}{d\rho} + \frac{kG}{\rho} - \left(\frac{\chi_2}{\chi} - \frac{Z\alpha}{\rho}\right)F = 0 \tag{16.155}$$

Let us try the series solution for these equations. $F(\rho)$ and $G(\rho)$ can be written as

$$F(\rho) = e^{-\rho}\sum_{n=0}^{\infty} a_n \rho^{s+n} \tag{16.156}$$

and

$$G(\rho) = e^{-\rho}\sum_{n=0}^{\infty} b_n \rho^{s+n} \tag{16.157}$$

Substituting the Equations (16.156) and (16.157) in the Equation (16.154) and cancelling the factor $e^{-\rho}$ after effecting the differentiation, we get

$$\sum_{n=0}^{\infty}\left(-a_n\rho^{s+n} + a_n(s+n-k)\rho^{s+n-1} - \frac{\chi_1}{\chi}b_n\rho^{s+n} - Z\alpha b_n\rho^{s+n-1}\right) = 0 \tag{16.158}$$

The Equation (16.158) becomes

$$\sum_{n=1}^{\infty}\left[-a_{n-1} - \frac{\chi_1}{\chi}b_{n-1}\right]\rho^{s+n-1} + \sum_{n=0}^{\infty}(a_n(s+n-k) - Z\alpha b_n)\rho^{s+n-1} = 0$$

Consider the terms involving ρ^{s+n} and replace n by $(n - 1)$ in the summation (the series now starts from $n = 1$).

$$\sum_{n=0}^{\infty}\left[-a_n\rho^{s+n} - \frac{\chi_1}{\chi}b_n\rho^{s+n}\right] = \sum_{n=1}^{\infty}\left[-a_{n-1} - \frac{\chi_1}{\chi}b_{n-1}\right]\rho^{s+n-1}$$

Consider the second summation in the above equation and expand it by separating $n = 0$ term from the rest of the series.

$$\sum_{n=0}^{\infty}[a_n(s+n-k) - Z\alpha\, b_n]\rho^{s+n-1} = [a_0(s-k) - Z\alpha\, b_0]\rho^{s-1} + \sum_{n=1}^{\infty}[a_n(s+n-k) - Z\alpha\, b_n]\rho^{s+n-1}$$

Making use of these results, we can write the Equation (16.158) as

$$[a_0(s-k) - Z\alpha b_0]\,\rho^{s-1} + \sum_{n=1}^{\infty}\left\{-a_{n-1} + (s+n-k)a_n - \frac{\chi_1}{\chi}b_{n-1} - Z\alpha\, b_n\right\}\rho^{s+n-1} = 0$$

Equating the coefficient of each power of ρ to zero, we get

$$a_0(s-k) - Z\alpha b_0 = 0 \quad \text{for } n = 0 \tag{16.159}$$

and

$$-a_{n-1} + (s+n-k)a_n - \frac{\chi_1}{\chi} b_{n-1} - Z\alpha\, b_n = 0 \quad \text{for } n > 0 \tag{16.160}$$

In the same way, from the Equation (16.155), we get

$$(s+n+k)a_n - b_{n-1} + Z\alpha\, a_n - \frac{\chi_2}{\chi} a_{n-1} = 0 \quad \text{for } n > 0 \tag{16.161}$$

and

$$(s+k)b_0 + Z\alpha\, a_0 = 0 \quad \text{for } n = 0 \tag{16.162}$$

First, let us solve the equation for a_0 and b_0.

The Equations (16.159) and (16.162) form a pair of equations for a_0 and b_0, and the non-zero solution exists, provided

$$\begin{vmatrix} s-k & -Z\alpha \\ Z\alpha & s+k \end{vmatrix} = 0$$

or

$$(s-k)(s+k) + Z^2\alpha^2 = 0$$

or

$$s^2 = (k^2 - Z^2\alpha^2)$$

and

$$s = \pm[k^2 - Z^2\alpha^2]^{1/2}$$

The first term ($n = 0$) in the series solutions for $F(\rho)$ and $G(\rho)$ is $e^{-\rho}\rho^{(k^2-z^2\alpha^2)^{1/2}}$ or $e^{-\rho}/\rho^{(k^2-z^2\alpha^2)^{1/2}}$.

As $\rho \to 0$, the second solution diverges. Therefore, this solution is not considered. The first solution is a well-behaved solution. Let us note that the minimum value of k^2 is 1 for which both the solutions $v(r) = F(r)/r$ and $\omega(r) = G(r)/r$ will diverge as $r \to 0$. However, in the calculation of matrix elements, this divergence goes away. Therefore, we choose

$$s = +[k^2 - Z^2\alpha^2]^{1/2} \tag{16.163}$$

Multiply the Equations (16.160) by χ and (16.161) by χ_1 and subtract the second from the first:

(16.160) χ – (16.161) χ_1:

$$[\chi(s+n-k) - Z\alpha\chi_1]a_n = [\chi_1(s+n+k) + Z\alpha\chi]b_n \quad \text{for } n \neq 0 \tag{16.164}$$

This recursion relation decides the series in the Equations (16.156) and (16.157). Let us study the convergence of these series.

We have

$$a_n = \frac{(s+n+k)\chi_1 + Z\alpha\chi}{(s+n-k)\chi_1 - Z\alpha\chi_1} b_n$$

As $n \to \infty$

$$a_n \simeq \frac{n\chi_1}{n\chi} b_n \quad \text{or} \quad a_n \simeq \frac{\chi_1}{\chi} b_n \tag{16.165}$$

From the Equation (16.160), we have

As $n \rightarrow \infty \quad -a_{n-1} + na_n - \frac{\chi_1}{\chi} b_{n-1} \simeq 0$

Using the Equation (16.165), we get

$$-a_{n-1} + na_n - \frac{\chi_1}{\chi} . \frac{\chi}{\chi_1} a_{n-1} \simeq 0$$

$$a_n \simeq \frac{2}{n} a_{n-1}$$

Similarly, we have $\quad b_n \simeq \frac{2}{n} b_{n-1}$

If we make a series expansion for $e^{2\rho}$ we get the same recurrence relationship. Therefore, both the series $\sum_{n=1}^{\infty} a_n \rho^{s+n}$ and $\sum_{n=1}^{\infty} b_n \rho^{s+n}$ behave like $e^{2\rho}$. Obviously, $F = e^{-\rho} \sum_{n=0}^{\infty} a_n \rho^{s+n}$ and $G = e^{-\rho} \sum_{n=0}^{\infty} b_n \rho^{n+s}$ diverge for larger ρ. This divergence can be avoided, provided we terminate the series for some value of n so that the infinite series becomes a polynomial of finite degree. Then, the multiplicative factor $e^{-\rho}$ will make F and G go to zero for large ρ.

Let us terminate the series by implementing the following condition.

$$a_{n'} \neq 0, \ b_{n'} \neq 0 \ \text{ and } \ a_{n'+1} = b_{n'+1} = 0 \ \text{ for } \ n' \neq 0$$

Let us put $n = n' + 1$ in the Equation (16.160).

$$-a_{n'} + (s + n' + 1 - k) a_{n'+1} - \frac{\chi_1}{\chi} b_{n'} - Z\alpha_1 b_{n'+1} = 0$$

Since we choose the boundary conditions $a_{n'+1} = b_{n'+1} = 0$, the above equation leads to the following results.

$$-a_{n'} - \frac{\chi_1}{\chi} b_{n'} = 0 \ \text{ or } \ a_{n'} = -\frac{\chi_1}{\chi} b_{n'} \tag{16.166}$$

Now, put $n = n'$ in the Equation (16.164).

$$[\chi(s + n' - k) - Z\alpha\chi_1] a_{n'} = [\chi_1(s + n' + k) + Z\alpha\chi] b_{n'}$$

$$[\chi(s + n' - k) - Z\alpha\chi_1] \left(-\frac{\chi_1}{\chi} b_{n'} \right) = [\chi_1(s + n' + k) + Z\alpha\chi] b_{n'}$$

$$-[\chi(s + n' - k) - Z\alpha\chi_1]\chi_1 = \chi_1\chi(s + n' + k) + Z\alpha\chi^2$$

or $$-[\chi(s + n' - k) - Z\alpha\chi_1]\chi_1 = \chi_1\chi(s + n' + k) + Z\alpha\chi_1\chi_2$$

$$2\chi(s + n') = Z\alpha(\chi_1 - \chi_2)$$

Squaring this equation, we get

$$4\chi^2(s + n')^2 = Z^2\alpha^2(\chi_1 - \chi_2)^2$$

or
$$\frac{4(m^2c^4-E^2)}{\hbar^2c^2}(s+n')^2=\frac{Z^2\alpha^2 4E^2}{\hbar^2c^2}$$
$$E^2[(s+n')^2+Z^2\alpha^2]=m^2c^4(s+n')^2$$
$$E^2=\frac{m^2c^4(s+n')^2}{\left[(s+n')^2+Z^2\alpha^2\right]}=\frac{m^2c^4}{\left[1+\dfrac{Z^2\alpha^2}{(s+n')^2}\right]}$$
$$E=mc^2\left[1+\frac{Z^2\alpha^2}{(s+n')^2}\right]^{-1/2} \tag{16.167}$$

We have to recognize the fact that this result is derived using the Equation (16.164), which is true only for $n'=1, 2, 3\ldots.$

From the Equations (16.156) and (16.157), we have for $n'=0$, $F(\rho)=e^{-\rho}a_0\rho^s$ and $G(\rho)=e^{-\rho}b_0\rho^s$. Substituting these expressions in the Equations (16.154) and 16.155), we get

$$\left[-a_0-\frac{\chi_1}{\chi}b_0\right]\rho^s+[a_0s-a_0k-Z\alpha b_0]\rho^{s-1}=0$$
$$\left[-b_0-\frac{\chi_2}{\chi}a_0\right]\rho^s+[b_0s+b_0k-Z\alpha a_0]\rho^{s-1}=0$$

Equating the coefficients of ρ^s and ρ^{s-1} to zero, we get

$$\left(-a_0-\frac{\chi_1}{\chi}b_0\right)=0 \quad\text{or}\quad \frac{a_0}{b_0}=-\frac{\chi_1}{\chi}$$
$$a_0(s-k)-Z\alpha b_0=0 \quad\text{or}\quad \frac{a_0}{b_0}=\frac{Z\alpha}{s-k}$$
$$b_0(s+k)+Z\alpha a_0=0 \quad\text{or}\quad \frac{a_0}{b_0}=-\frac{s+k}{Z\alpha} \tag{16.168}$$

From these equations, we get

$$\chi_1(s-k)=-\chi Z\alpha$$
$$\chi(s+k)=\chi_1 Z\alpha$$
$$\chi_1\underbrace{\chi_1(s-k)}_{-Z\alpha\chi}=-\chi\underbrace{\chi_1 Z\alpha}_{(s+k)\chi}=-\chi^2(s+k)=-\chi_1\chi_2(s+k)$$
$$\therefore\quad \chi_1^2(s-k)=-\chi_1\chi_2(s+k)$$

Rearranging this equation, we get

$$s(\chi_1+\chi_2)=k(\chi_1-\chi_2)$$
$$s\frac{2mc^2}{\hbar c}=k\frac{2E}{\hbar c} \qquad \text{or}\quad E=\frac{mc^2s}{k} \tag{16.169}$$

Note that since E is positive, k has to be positive. So, on the other hand, in the expression for E in the Equation (16.167), k can be positive or negative. Therefore, from the Equations (16.167) and (16.169), we have

$$E = \left\{\begin{array}{ll} \dfrac{mc^2 s}{k} & k = 1, 2, 3\cdots \text{ for } n' = 0 \\ mc^2\left[1+\dfrac{Z^2\alpha^2}{(s+n')^2}\right]^{-\frac{1}{2}} & k = \pm 1, \pm 2 \cdots n' \neq 0 \end{array}\right\} \tag{16.170}$$

In fact, if we put $n' = 0$ in the Equation (16.167), we get the Equation (16.169) (use the Equation (16.163)).

$$E = mc^2\left[1+\frac{Z^2\alpha^2}{s^2}\right]^{-1/2} = \frac{mc^2 s}{(s^2+Z^2\alpha^2)^{1/2}} = \frac{mc^2 s}{k}$$

Let us expand the energy E in terms of α^2. In what follows, we make binomial expansion consistently retaining only up to order α^4.

$$\begin{aligned} s = (k^2 - Z^2\alpha^2)^{1/2} &= |k|\left(1-\frac{Z^2\alpha^2}{k^2}\right)^{1/2} \\ &= |k|\left(1-\frac{1}{2}\frac{Z^2\alpha^2}{k^2}-\frac{1}{8}\frac{Z^4\alpha^4}{k^4}\right) \\ &= |k| - \frac{1}{2}\frac{Z^2\alpha^2}{|k|} - \frac{1}{8}\frac{Z^4\alpha^4}{|k|^3} \end{aligned} \tag{16.171}$$

Note that we have used the result $(k^2)^{1/2} = |k|$.

$$s + n' = n' + |k| - \frac{1}{2}\frac{Z^2\alpha^2}{|k|} - \frac{1}{8}\frac{Z^4\alpha^4}{|k|^3}$$

Let us define n as $\qquad n = n' + |k|$

The possible values of n' are 0, 1, 2, …, and the possible values of n are 1, 2, 3 … (since $|k|$ has to be non-zero).

$$s + n' = n - \frac{1}{2}\frac{Z^2\alpha^2}{|k|} - \frac{1}{8}\frac{Z^4\alpha^4}{|k|^3} = n\left(1-\frac{Z^2\alpha^2}{2n\,|k|}-\frac{Z^4\alpha^4}{8n\,|k|^3}\right)$$

$$\begin{aligned} (s+n')^{-2} &= \frac{1}{n^2}\left(1-\frac{Z^2\alpha^2}{2n\,|k|}-\frac{Z^4\alpha^4}{8n\,|k|^3}\right)^{-2} \\ &= \frac{1}{n^2}\left(1+\frac{Z^2\alpha^2}{n\,|k|}+\frac{Z^4\alpha^4}{4n\,|k|^3}+\frac{3Z^4\alpha^4}{4n^2\,|k|^2}\cdots\right) \end{aligned}$$

$$\begin{aligned} [1+Z^2\alpha^2(s+n')^{-2}]^{-1/2} &= \left[1+\frac{Z^2\alpha^2}{n^2}\left(1+\frac{Z^2\alpha^2}{n\,|k|}\cdots\right)\right]^{-1/2} \\ &= \left[1+\frac{Z^2\alpha^2}{n^2}+\frac{Z^4\alpha^4}{n^3\,|k|}+\cdots\right]^{-1/2} \end{aligned}$$

$$= \left[1 - \frac{1}{2}\frac{Z^2\alpha^2}{n^2} - \frac{1}{2}\frac{Z^4\alpha^4}{n^3|k|} + \frac{3}{8}\frac{Z^4\alpha^4}{n^4} \cdots \right]$$

$$= \left[1 - \frac{1}{2}\frac{Z^2\alpha^2}{n^2} - \frac{Z^4\alpha^4}{2n^4}\left(\frac{n}{|k|} - \frac{3}{4} \right) \cdots \right]$$

Substituting this expression in Equation (16.170) we get

$$E = mc^2 \left[1 - \frac{1}{2}\frac{Z^2\alpha^2}{n^2} - \frac{Z^4\alpha^4}{2n^4}\left(\frac{n}{|k|} - \frac{3}{4} \right) + \cdots \right]$$

$$= mc^2 \left[1 - \frac{1}{2}\frac{Z^2\alpha^2}{n^2} - \frac{Z^4\alpha^4}{2n^4}\left(\frac{n}{j+\frac{1}{2}} - \frac{3}{4} \right) \right]$$

The first term is the rest mass energy of the electron. The second and third terms are bound-state energy of the hydrogen atom. The second term is the same as the expression for Bohr energy level of hydrogen obtained in the non-relativistic quantum mechanics. The third term is the correction due to the relativistic quantum mechanics and is responsible for the fine structure. The fine structure as determined by this equation agrees well with the experimental values. This result is yet another success of the Dirac equation. Note that the correction to the Bohr energy level depends directly only on the quantum number *j*, and it is the same as the expression obtained in the time-independent perturbation theory due to relativistic correction and LS coupling.

16.6.7 Upper Component and Lower Component of the Dirac Spinor

For a given value *k*, the upper component and lower angular momentum eigenvalues are l_1 and l_2. So, strictly speaking, the orbital angular momentum quantum number cannot be used to characterize the hydrogen atom. However, it is customary to characterize the energy eigenstates using the angular momentum quantum number of the upper component. One way of justifying this idea is that the upper component is numerically much larger than the lower component. From the Equations (16.168) and (16.171), we have

$$\frac{b_0}{a_0} = -\frac{Z\alpha}{s+k} \simeq -\frac{Z\alpha}{2k} << 1$$

Similarly,

$$\frac{b_n}{a_n} = -\frac{\chi}{\chi_1} = -\frac{(\chi_1\chi_2)^{1/2}}{\chi_2}$$

$$= -\left(\frac{\chi_2}{\chi_1} \right)^{1/2} = -\left[\frac{mc^2 - E}{mc^2 + E} \right] << 1$$

Characterizing Eigenstates of Hydrogen Atom

We use the three quantum numbers n, l_1 and j to describe the energy eigenstate of the hydrogen atom. The quantum numbers which directly enter the expression for energy are *n* and *k*. They are related by the expression

$$n = n' + |k|$$

or
$$n' = n - |k|$$

Since n' and n are integers, for a given n, the values of n' and $|k|$ are restricted to select values. Correspondingly, possible values of j and l_1 are also restricted. We illustrate these ideas by considering the following cases. We use the following notation $n\,l_j$ to represent a state. (Here, l refers to l_1 of the upper component of the Dirac spinor.)

Example 16.10 Determine the possible states of hydrogen for $n = 1$, 2 and 3.

Solution: It will be useful to recall the following results in determining the various possible states of the hydrogen atom.

Possible values of n': $n' = 0, 1, 2, \ldots$

Possible values of n: $n = 1, 2, \ldots$

$$\text{Possible values of } k: \begin{cases} n' = 0 : k = 1, 2, 3\cdots \\ n' \neq 0 : k \pm 1, \pm 2, \pm 3, \cdots \end{cases}$$

$$\text{Values of } k \text{ and } l: \begin{cases} \text{Positive } k\text{: } k = j + \frac{1}{2}; \quad l_1 = j - \frac{1}{2}; l_2 = j + \frac{1}{2} \\ \text{Negative } k\text{: } k = -\left(j + \frac{1}{2}\right); l_1 = j + \frac{1}{2}; l_2 = j - \frac{1}{2} \end{cases}$$

With these information, we can determine the possible states of the hydrogen atom.

Case 1:

$$n = 1$$

$$n' = 1 - |k|$$

The only possible value is
$$|k| = 1$$

$$\therefore\ n' = 0 \ \text{ when } \ |k| = 1; k = 1 \quad \therefore\ j + \frac{1}{2} = 1 \text{ or } j = \frac{1}{2}$$

Therefore, we have

$$n = 1,\ j = \frac{1}{2} \text{ and } l_1 = 0$$

The corresponding state is $1s_{\frac{1}{2}}$.

Case 2:

$$n = 2$$

$$n' = 2 - |k|$$

The possible values of k are $|k| = 1, 2$

a) $n' = 1$ when $|k| = 1$. The values of k are $k = \pm 1$
   For $k = 1$, $j + 1/2 = 1 \Rightarrow j = 1/2$. This in turn implies $l_1 = j - 1/2 = 0$
   For $k = -1$, $-(j + 1/2) = -1$ or $j = 1/2$; This implies $l_1 = j + 1/2 = 1$

b) When $|k| = 2, n' = 0 \therefore k = 2$; $j + 1/2 = 2$ or $j = 3/2$. This, in turn, implies $l_1 = j - 1/2 = 3/2 - 1/2 = 1$. Therefore, for $n = 2$, there are three possible states:
They are

$$\left.\begin{array}{l} n = 2,\ j = \frac{1}{2},\ l_1 = 0 \Rightarrow 2s_{1/2}. \\ n = 2,\ j = \frac{1}{2},\ l_1 = 1 \Rightarrow 2p_{1/2}. \end{array}\right\}\text{Same energy}$$

$$n = 2,\ j = \frac{3}{2},\ l_1 = 1 \Rightarrow 2p_{3/2}$$

Case 3:

$$n = 3$$

$$n' = 3 - |k|$$

The possible values $\quad |k| = 1, 2, 3$ or $k = \pm 1, \pm 2, 3$

a) $|k| = 1;\ k = +1 \Rightarrow j + \frac{1}{2} = 1 \ \therefore j = \frac{1}{2};\ l_1 = j - \frac{1}{2} = 0$

b) $|k| = 1;\ k = -1 \Rightarrow -\left(j + \frac{1}{2}\right) = -1 \ \therefore j = \frac{1}{2};\ l_1 = j + \frac{1}{2} = 1$

c) $|k| = 2;\ k = 2 \Rightarrow j + \frac{1}{2} = 2 \ \therefore j = \frac{3}{2};\ l_1 = j - \frac{1}{2} = \frac{3}{2} - \frac{1}{2} = 1$

d) $|k| = 2;\ k = -2 \Rightarrow -\left(j + \frac{1}{2}\right) = -2 \ \therefore j = \frac{3}{2};\ l_1 = j + \frac{1}{2} = 2$

e) $|k| = 3;\ k = 3 \Rightarrow j + \frac{1}{2} = 3 \ \therefore j = \frac{5}{2};\ l_1 = j - \frac{1}{2} = \frac{5}{2} - \frac{1}{2} = 2$

Therefore, for $n = 3$, there are five states. They are

$$\left.\begin{array}{l} n = 3, j = \frac{1}{2}, l_1 = 0 \Rightarrow 3s_{1/2} \\ n = 3, j = \frac{1}{2}, l_1 = 1 \Rightarrow 3p_{1/2}. \end{array}\right\} \text{ Same energy}$$

$$\left.\begin{array}{l} n = 3, j = \frac{3}{2}, l_1 = 0 \Rightarrow 3p_{3/2} \\ n = 3, j = \frac{3}{2}, l_1 = 2 \Rightarrow 3d_{3/2}. \end{array}\right\} \text{ Same energy}$$

$$n = 3, j = \frac{5}{2}, l_1 = 2 \Rightarrow 3d_{5/2}$$

The energy-level diagram can be drawn as follows:

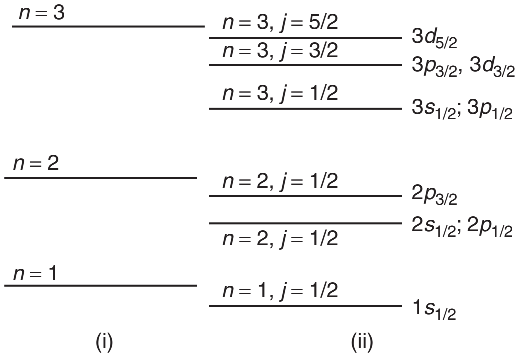


Fig.16.9 Hydrogen energy levels (not drawn to scales): (i) Bohr level (ii) Relativistic dirac equation

The spectral lines representing a transition from one Bohr level to another are actually closely spaced lines which require high-resolution instrument to resolve. The fine structure as calculated from the Dirac equation agrees well with the experiment.

Example 16.11 Show that $\frac{\boldsymbol{\sigma}.\mathbf{r}}{r}\phi_{jl_1m} = a\phi_{jl_2m}$ using the transformation property of these functions under parity transformation.

Solution: From the Equation (16.134), we have $P_0\phi_{jl_1m} = \lambda\phi_{jl_1m}$ and $P_0\phi_{jl_2m} = -\lambda\phi_{jl_1m}$. Now we have

$$P_0\frac{\boldsymbol{\sigma}\cdot\mathbf{r}}{r}\phi_{jl_1m} = -\frac{\boldsymbol{\sigma}\cdot\mathbf{r}}{r}\phi_{jl_1m}(-\mathbf{r}) = -\frac{\boldsymbol{\sigma}\cdot\mathbf{r}}{r}P_0\phi_{jl_1m} = -\frac{\boldsymbol{\sigma}\cdot\mathbf{r}}{r}\lambda\phi_{jl_1m}$$

$$\therefore P_0\frac{\boldsymbol{\sigma}\cdot\mathbf{r}}{r}\phi_{jl_1m} = -\lambda\frac{\boldsymbol{\sigma}\cdot\mathbf{r}}{r}\phi_{jl_1m}$$

$\therefore \frac{\boldsymbol{\sigma}\cdot\mathbf{r}}{r}\phi_{jl_1m}$ is an eigenstate of P_0 with eigenvalue $-\lambda$.

$$\therefore \frac{\boldsymbol{\sigma}\cdot\mathbf{r}}{r}\phi_{jl_1m} = \text{Constant}\times\phi_{jl_2m} = a\phi_{jl_2m}$$

Example 16.12 Prove that $\frac{\boldsymbol{\sigma}\cdot\mathbf{r}}{r}\phi_{jl_1m} = -\phi_{jl_2m}$ and $\frac{\boldsymbol{\sigma}\cdot\mathbf{r}}{r}\phi_{jl_2m} = -\phi_{jl_1m}$.

Solution: Here, we closely follow Alonso and Valk. We have

$$L^2\phi_{jl_1m} = l_1(l_1+1)\hbar^2\phi_{jl_1m}; \qquad l_1 = j+\frac{1}{2}$$

$$L^2\phi_{jl_2m} = l_2(l_2+1)\hbar^2\phi_{jl_2m}; \qquad l_2 = j-\frac{1}{2}$$

ϕ_{jl_1m} can be constructed from spherical harmonics representing the orbital angular momentum and spinors representing the spin angular momentum.

$$\phi_{jl_1m} = c_1 Y_{l_1 m-\frac{1}{2}}\chi_{\frac{1}{2}\frac{1}{2}} + c_2 Y_{l_1 m+\frac{1}{2}}\chi_{\frac{1}{2}-\frac{1}{2}}$$

where c_1 and c_2 are the Clebsch–Gordan coefficients. In this case, we have $l_1 = j+\frac{1}{2}$or $j = l_1 - \frac{1}{2}$, and from Table 10.4 in Chapter 10, we get

$$c_1 = \left\langle l_1\,\overline{m-\frac{1}{2}};\frac{1}{2}\,\frac{1}{2}\middle|\overline{l_1-\frac{1}{2}}\,m\right\rangle = -\left[\frac{l_1-m+\frac{1}{2}}{2l_1+1}\right]^{1/2}$$

$$= -\frac{\left(j+\frac{1}{2}-m+\frac{1}{2}\right)^{1/2}}{\left(2\left(j+\frac{1}{2}\right)+1\right)^{1/2}} = -\frac{(j-m+1)^{1/2}}{(2j+2)^{1/2}}$$

Similarly,

$$c_2 = \left\langle l_1\,\overline{m+\frac{1}{2}};\frac{1}{2}-\frac{1}{2}\middle|\overline{l_1-\frac{1}{2}}\,m\right\rangle = \frac{\left(l_1+m+\frac{1}{2}\right)^{1/2}}{2l_1+1} = \left[\frac{j+m+1}{2j+2}\right]^{1/2}$$

Let us choose $$\chi_{\frac{1}{2}\frac{1}{2}} = \begin{pmatrix}1\\0\end{pmatrix} \text{ and } \chi_{\frac{1}{2}-\frac{1}{2}} = \begin{pmatrix}0\\1\end{pmatrix}$$

$$\therefore \phi_{jl_1m} = \begin{pmatrix} -\dfrac{(j-m+1)^{1/2}}{(2j+2)^{1/2}}Y_{l_1m-1/2} \\ \dfrac{(j+m+1)^{1/2}}{(2j+2)^{1/2}}Y_{l_1m+1/2}\end{pmatrix}$$

Similarly, ϕ_{jl_2m} can be written as

$$\phi_{jl_2m} = \begin{pmatrix} \left(\dfrac{j+m}{2j}\right)^{1/2} Y_{l_2m-1/2} \\ \left(\dfrac{j-m}{2j}\right)^{1/2} Y_{l_2m+1/2}\end{pmatrix}$$

$$\frac{\boldsymbol{\sigma}.\mathbf{r}}{r} = \begin{pmatrix}\cos\theta & \sin\theta\, e^{-i\phi} \\ \sin\theta\, e^{i\phi} & -\cos\theta\end{pmatrix}$$

$$\frac{\boldsymbol{\sigma}.\mathbf{r}}{r}\phi_{jl_1m} = \begin{bmatrix}\cos\theta & \sin\theta\, e^{-i\phi} \\ \sin\theta\, e^{i\phi} & -\theta\end{bmatrix}\begin{bmatrix} -\dfrac{(j-m+1)^{1/2}}{(2j+2)^{1/2}}Y_{l_1m-1/2} \\ \dfrac{(j+m+1)^{1/2}}{(2j+2)^{1/2}}Y_{l_1m+1/2}\end{bmatrix}$$

$$= \begin{pmatrix} -\dfrac{(j-m+1)^{1/2}}{(2j+2)^{1/2}}\cos\theta\, Y_{l_1 m-\frac{1}{2}} + \sin\theta\, e^{-i\phi}\dfrac{(j+m+1)^{1/2}}{(2j+2)^{1/2}} Y_{l_1 m+1/2} \\ -\dfrac{(j-m+1)^{1/2}}{(2j+2)^{1/2}}\sin\theta\, e^{i\phi} Y_{l_1 m-\frac{1}{2}} - \dfrac{(j+m+1)^{1/2}}{(2j+2)^{1/2}}\cos\theta\, Y_{l_1 m+1/2} \end{pmatrix}$$

Using the Equations (7.68) and (7.69) from Chapter 7 (note that $l_1 = j + \frac{1}{2}$), we get

$$-\frac{(j-m+1)^{1/2}}{(2j+2)^{1/2}}\cos\theta\, Y_{l_1 m-\frac{1}{2}} + \frac{(j+m+1)^{1/2}}{(2j+2)^{1/2}}\sin\theta\, e^{-i\phi} Y_{l_1 m-\frac{1}{2}}$$

$$= -\frac{(j-m+1)^{1/2}}{(2j+2)^{1/2}}\left\{\frac{\left(j+\frac{1}{2}+m-\frac{1}{2}+1\right)^{1/2}\left(j+\frac{1}{2}-m+\frac{1}{2}+1\right)^{1/2}}{\left[2\left(j+\frac{1}{2}\right)+1\right]^{1/2}\left[2\left(j+\frac{1}{2}\right)+3\right]^{1/2}} Y_{j+\frac{3}{2}m-\frac{1}{2}}\right.$$

$$\left. + \frac{\left(j+\frac{1}{2}+m-\frac{1}{2}\right)^{1/2}\left(j+\frac{1}{2}-m+\frac{1}{2}\right)^{1/2}}{\left[2\left(j+\frac{1}{2}\right)+1\right]^{1/2}\left[2\left(j+\frac{1}{2}\right)-1\right]^{1/2}} Y_{j-\frac{1}{2},m-\frac{1}{2}}\right\}$$

$$+ \frac{(j+m+1)^{1/2}}{(2j+2)^{1/2}}\left\{\frac{\left(j+\frac{1}{2}-m-\frac{1}{2}+1\right)^{1/2}\left(j+\frac{1}{2}-m-\frac{1}{2}+2\right)^{1/2}}{\left[2\left(j+\frac{1}{2}\right)+1\right]^{1/2}\left[2\left(j+\frac{1}{2}\right)+3\right]^{1/2}} Y_{j+\frac{3}{2}m-\frac{1}{2}}\right.$$

$$\left. - \frac{\left(j+\frac{1}{2}+m+\frac{1}{2}\right)^{1/2}\left(j+\frac{1}{2}+m+\frac{1}{2}-1\right)^{1/2}}{\left[2\left(j+\frac{1}{2}\right)+1\right]^{1/2}\left[2\left(j+\frac{1}{2}\right)-1\right]^{1/2}} Y_{j-\frac{1}{2},m-\frac{1}{2}}\right\}$$

$$= -\left(\frac{j+m}{2j}\right)^{1/2} Y_{j-\frac{1}{2}m-\frac{1}{2}}$$

Similarly, we can show that

$$-\frac{(j-m+1)^{1/2}}{(2j+2)^{1/2}}\sin\theta\, e^{i\phi} Y_{l_1 m-\frac{1}{2}} - \frac{(j+m+1)^{1/2}}{(2j+2)^{1/2}} Y_{l_1 m+\frac{1}{2}}$$

$$= -\left(\frac{j-m}{2j}\right)^{1/2} Y_{j-\frac{1}{2}m+\frac{1}{2}}$$

$$\therefore \frac{\boldsymbol{\sigma}.\mathbf{r}}{r}\phi_{jl_1m} = \begin{pmatrix} -\left(\dfrac{j+m}{2j}\right)^{1/2} & Y_{j-\frac{1}{2}\,m-\frac{1}{2}} \\ -\left(\dfrac{j-m}{2j}\right)^{1/2} & Y_{j-\frac{1}{2}\,m+\frac{1}{2}} \end{pmatrix} = -\phi_{jl_2m}$$

Since $\left(\frac{\boldsymbol{\sigma}.\mathbf{r}}{r}\right)\left(\frac{\boldsymbol{\sigma}.\mathbf{r}}{r}\right) = 1,$ we have

$$\frac{\boldsymbol{\sigma}.\mathbf{r}}{r}\left[\frac{\boldsymbol{\sigma}.\mathbf{r}}{r}\phi_{jl_1m}\right] = \frac{\boldsymbol{\sigma}.\mathbf{r}}{r}(-\phi_{jl_2m})$$

But

$$\frac{\boldsymbol{\sigma}.\mathbf{r}}{r}\frac{\boldsymbol{\sigma}.\mathbf{r}}{r}\phi_{jl_1m} = \phi_{jl_1m}$$

$$\therefore \frac{\boldsymbol{\sigma}.\mathbf{r}}{r}(-\phi_{jl_2m}) = \phi_{jl_1m}$$

EXERCISES

1. Give the arguments leading to Klien Gordan equation.
2. Why is it not possible to ignore or discard the negative energy eigenstates?
3. Obtain the equation of continuity $\partial\rho/\partial t + \nabla.\mathbf{S} = 0$ for a free particle, obeying Klien Gordan equation. Explain the difficulty involved in interpreting ρ as the probability density.
4. What are the shortcomings of Klien Gordan equation?
5. Give the arguments leading to the formation of the Dirac equation.
6. Why do we expect the time derivative in a relativistic wave equation to be first order?
7. In the Dirac equation $i\hbar\,\partial\psi/\partial t = (c\boldsymbol{\alpha}\cdot\mathbf{p}_{op} + \beta mc^2)\psi$, give the arguments to determine α and β.
8. Show that β cannot be a 2 × 2 matrix which anti-commutes with the Pauli matrices $\sigma_{x,}\sigma_y$ and σ_z.
9. What are the complete sets of compatible dynamical variable for the Dirac equation describing (a) free particle and (b) particle in Coulomb field.
10. Explain Dirac's interpretation of negative energy states for electrons.
11. Explain the nature of the ground state in Dirac's theory of an electron.
12. Is it possible to extend Dirac's concept of negative energy sea to bosons also?
13. Explain the concept of helicity for a particle.
14. Which of the following operators commutes with the Dirac Hamiltonian for a free particle: (a) $\hbar/2\Sigma_z$ and (b) $\boldsymbol{\Sigma}\cdot\mathbf{p}_{op}/|\mathbf{p}|$?
15. Explain the nature of g_s in the expression for the magnetic moment of an electron due to the spin angular momentum.
16. Give the arguments leading to the operator K to describe the Dirac equation for a particle in central potential.
17. Why do you need to define the radial component of momentum operator as $p_r = \frac{1}{2}\left(\frac{\mathbf{r}}{r}\mathbf{p} + \mathbf{p}\cdot\frac{\mathbf{r}}{r}\right)$?
18. What are the quantum numbers used to specify the energy eigenstates resulting from the Dirac equation for hydrogen atom?
19. How do you explain the use of the orbital angular momentum quantum number l in specifying the hydrogen atom in spite of the fact the energy eigenstates for the Dirac equation do not have well-defined l values?

REFERENCES

1. P. M. Mathews and K. Venkatesan, 1976. *A Textbook of Quantum Mechanics.* New Delhi: Tata McGraw-Hill Company.
2. M. Alonso and H. Valk, 1960. *Quantum Mechanics: Principles and Applications*. Reading, Massachusetts: Addison Wesley Publishing Company, Inc.
3. W. Greiner, 1990. *Relativistic Quantum Mechanics – Wave Equations*. Berlin: Springer Verlag.
4. M. E. Rose, 1957. *Elementary Theory of Angular Momentum*. New York: John Wiley & Sons, Inc.
5. L. I. Schiff, 1968. *Quantum Mechanics*. New Delhi: McGraw-Hill International Book Company.
6. P. A. M. Dirac, 1958. *The Principles of Quantum Mechanics*. London: Oxford University Press.

Index