Quantum Mechanics
Lecture Notes

J. W. Van Orden
Department of Physics
Old Dominion University

August 21, 2007

Contents

Chapter 1

Introduction: The Old Quantum Theory

Quantum Mechanics is the physics of matter at scales much smaller than we are able to observe of feel. As a result, we have no direct experience of this domain of physics and therefore no intuition of how such microscopic systems behave. The behavior of quantum systems is very different for the macroscopic systems of Classical Mechanics. For this reason, we will begin by considering the many historical motivations for quantum mechanics before we proceed to develop the mathematical formalism in which quantum mechanics is expressed. The starting point is to review some aspects of classical physics and then to show how evidence accumulated for atomic systems that could not be explained in the context of classical physics.

1.1   Classical Physics

What we now call "Classical Physics" is the result of the scientific revolution of the sixteenth and seventeenth centuries that culminated in Newtonian mechanics. The core of this physics is Newton's laws describing the motion of particles of matter. The particles are subject to forces and Newton's Second Law $\boldsymbol{F} = m\boldsymbol{a}$ can then be used to describe the motion of the particle in terms of a second-order differential equation. By specifying the position and velocity of the particle at some initial time, the motion of the particle is determined at all subsequent time. There is nothing in principle in classical physics that prevents the initial conditions from being determined to arbitrary accuracy. This property that allows all subsequent motion to be predicted from Newton's Laws, the force laws and the initial conditions is called "classical causality" and classical physics is said to be deterministic since the motion is determined by the initial conditions.

The major direction of physics after Newton was to incorporate as much physical phenomena as possible into the framework of Newtonian physics. Thus Newton's laws were applied to the motions of extended objects, to the motion of fluids and elastic

bodies and to link mechanics to thermodynamics by means of the global conservation of energy.

One application, which will be of particular interest in this course, was the description of wave motion in fluids and elastic materials. The waves are the manifestation of the collective motion of a macroscopically continuous medium. Since the waves are seen as deviations of some quantity such as the height of the surface of water from some average value, the wave is characterized by the amplitude of this deviation and the sign of the deviation at any point and time may be either positive or negative. Since the waves are solutions to linear differential equations, waves can be added by simply adding the deviation of the wave from equilibrium at every point at any given time. The fact that this deviation may be either positive or negative leads to the wave motion being either cancelled or enhanced at different points which produces the typical wave phenomena of interference and diffraction. The average energy density carried by a mechanical wave is proportional to the square of the amplitude of the wave, and independent of the frequency.

At the beginning of the nineteenth century, mechanics, thermodynamics, electromagnetic phenomena and optics were not yet united in any meaningful way. Perhaps the most glaring problem was the question of the nature of light. The long and heated argument as to whether light was corpuscular of a wave was finally settled in the nineteenth century by demonstrating in such experiments as Young's two slit experiment that light could interfere and diffract and was therefore a wave.

During the course of this century, the understanding electromagnetic phenomena developed rapidly culminating in Maxwell's equations for electromagnetic fields. The great triumph of Maxwell's equations was the prediction of wave solutions to Maxwell's equations that led to the unification of electrodynamics and optics. The Maxwell's equations were also verified by the discovery of radio waves by Hertz. There were still obstacles to the unification of electromagnetism with mechanics. The nature of electromagnetic currents was not understood until the very last part of the century and there appeared to be no supporting medium for electromagnetic waves as was the case with all mechanical waves. An attempt to deal with the latter problem was to propose the existence of a all-pervading medium called the *ether* that was the medium that supported electromagnetic waves. It was the unsatisfactory nature of this hypothesis that led Einstein to develop the special theory of relativity in 1905.

The situation with thermodynamics was much more satisfactory. The discovery that heat was a form of energy that could be created from mechanical energy or work and that could be in turn used to produce mechanical energy led to the concept of global conservation of energy. That is, energy can be transformed from one type to another but cannot be created or destroyed. During the nineteenth century the laws and logical structure of thermodynamics were codified and applied to a variety of phenomena. In addition, through the kinetic theory of gasses and the development of statistical mechanics by Boltzman and Gibbs it was shown that thermodynamics could be described by the average motions of complicated systems of very large num-

bers of particles composing either fluids or solids. Since the exact motion of such large collections of particles could not be determined, statistical methods are used to describe average properties of macroscopic systems. As a result thermodynamics and statistical mechanics are not deterministic. It should be emphasized that this is because the exact knowledge of the microscopic state of the system is impractical but not impossible in principle.

It is also important for an understanding of the motivation behind the development of quantum mechanics to note that the nineteenth century is also the time when chemistry became a quantitative science. It was noted that the specific gravities of various elements were approximately integer multiples of that of hydrogen. Experiments by Faraday on electrolysis indicated that the change in electric charges of various ions in this process indicated that charge appeared as multiples of some fixed elementary charge. It was discovered that the light spectrum given of by various materials when heated by a flame or an electric arc showed discrete lines rather than a continuous distribution of wavelengths. It was shown that these lines were characteristic of each element and that, therefore, the spectra of materials could be used to identify the presence of known elements and to find new ones. This was used Kirchhoff in 1859 to show that the absorbtion spectra of the sun indicated the existence of sodium in the stellar atmosphere. Finally the periodic table of elements (Mendeleev, 1869) showed that the chemical properties of various elements display a regular pattern as a function of the atomic number. As the century progressed, more elements were discovered (some of which did not actually exist) and an ever increasing collection of of improved spectroscopic information about these elements was amassed. All of this led to the belief on the part of chemists and physicists that the atomic description of matter was at least a useful tool if not a reality. The reality of atoms and molecules as chemically fundamental constituents of matter was not demonstrated until Einstein's paper of 1905 on Brownian motion, where the erratic motion of small particles suspended in fluids was described as the result of the collective result of large numbers of collisions between the molecules of the fluid and the small particles.

One result of this period of great progress in the unification of classical physics and in the development of chemistry was that a great amount of new measurements were accumulated that were not fully incorporated into the structure of classical physics and which ultimately proved this structure to be inadequate. We will now examine some of these problems and see how they led to the development of what is now called the old quantum theory.

1.2 Thermodynamics and Statistical Physics

1.2.1 Black Body Radiation

One of the phenomena that had long been known to man was that when an object is heated to a sufficiently high temperature that it begins to glow and that the color of the glowing object is related to the temperature of the object. This was used for millennia by metal smiths to determine when metal was sufficiently hot to be easily worked with hammers. In 1859 Kirchhoff showed on the basis of thermodynamics that the energy per unit area per unit time (the energy current density) of light given off by an a completely absorbing body is a function only of the temperature of the object and the frequency of the light emitted by the object and not to any particular physical properties of the emitter. A completely absorbing object is referred to as a *black body*. In practice a black body was constructed as a closed box which was sealed to light and heated to a uniform temperature. A small hole was placed in the box to allow the measurement of the intensity of the light in the cavity as a function of frequency. The black body distribution problem was of interest not only for its intrinsic scientific value, but also because black bodies could be used as means of calibrating various kinds of lamps used for scientific and commercial purposes . The verification of Kirchhoff's prediction and the actual measurement of the spectral density for black body radiation posed a serious technical problem at the time and it was not until the last part of the nineteenth century that sufficiently reliable experiments were available.

In the mean time several theoretical contributions to this problem were obtained. In 1879, Stefan proposed that the total electromagnetic energy in the cavity is proportional to the fourth power of the absolute temperature. In 1884 Boltzmann provided a proof of this using thermodynamics and electrodynamics. He showed that

$$J(\nu, T) = \frac{c}{8\pi} \rho(\nu, T) \tag{1.1}$$

where $J(\nu, T)$ is the energy current density of emitted radiation and $\rho(\nu, T)$ is the spectral density or the energy per unit volume per unit time of radiation in the cavity and c is the speed of light. From this Boltzmann derived the Stefan-Boltzmann law

$$E(T) = V \int d\nu \, \rho(\nu, T) = aVT^4 \tag{1.2}$$

where $E(T)$ is the energy of the radiation in the cavity for temperature T and V is the volume enclosed by the cavity.

In 1893, Wien derived the Wien displacement law

$$\rho(\nu, T) = \nu^3 f\left(\frac{\nu}{T}\right) \tag{1.3}$$

which states that the spectral distribution is proportional the ν^3 times some function of the ratio of ν to T. However on the basis of just thermodynamics and electrodynamics it is not possible to determine this function. Wien conjectured that the spectral distribution was of the form

$$\rho(\nu, T) = \alpha \nu^3 e^{-\beta \frac{\nu}{T}} \tag{1.4}$$

where α and β were unknown constants to be determined by data. In 1893 Paschen presented data from the near infrared that was in excellent agreement with Wien's formula.

In 1900, Rayleigh derived a new formula using statistical mechanics that is given by

$$\rho(\nu, T) = \frac{8\pi\nu^2}{c^3} kT \,. \tag{1.5}$$

where k is the Boltzmann constant. In fact Rayleigh did not actually determine the constants in this expression. These were correctly determined by Jeans in 1905 and the equation is now called the Rayleigh-Jeans Law. Note that this law will not satisfy the Stefan-Boltzmann condition since increases and ν^2 and, therefore, has an infinite integral over ν. This was referred to as the *ultraviolet catastrophe.*

Also in 1900, two groups in Berlin consisting of Lummer and Pringsheim, and Rubens and Kurlbaum obtained data at lower frequencies further into the infrared. This data, however, was not in good agreement with Wien's Law. This data was immediately shown to Planck and he was told that the data was linear in T at small temperatures. It is not clear whether he knew of Rayleigh's work at that time, however. Planck knew that Wien's Law worked well at high frequencies and that the spectral density had to be linear in T as low temperatures. Using this he quickly guessed at a formula that would interpolate between the two regions. This is Planck's Law

$$\rho(\nu, T) = \frac{8\pi h\nu^3}{c^3} \frac{1}{e^{\frac{h\nu}{kT}} - 1} \tag{1.6}$$

which contained a new constant h (*Planck's constant*) that could be determined from the data. Indeed this law provides an excellent representation of the black body spectral distribution. Figure 1.1 shows the three laws for the distribution functions at $T = 1000K$.

Planck was now in a situation which is not uncommon for theoretical physicists, he had a formula that fit the data, but did not have a proof. He then proceeded to try to derive his Law. He modeled the black body as a set of charges that were attached to harmonic oscillators. The acceleration of these particles then produced radiation. The assumption was that these oscillators were in thermal equilibrium with the radiation in the cavity. In obtaining his proof, he did two things which were not consistent with classical physics, he used a counting law for determining the probability of various configurations that was not consistent with classical statistical
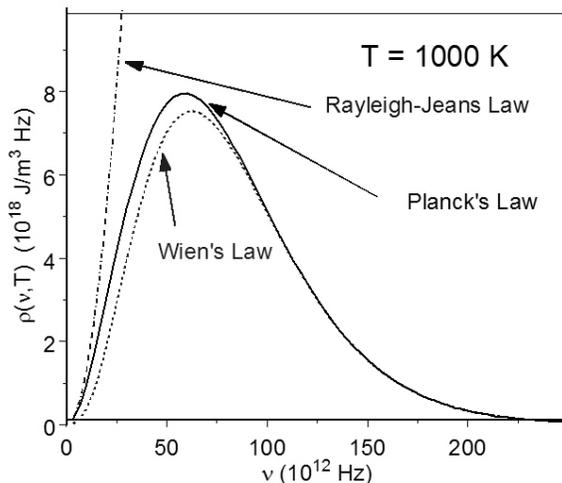
Figure 1.1: Spectral distributions for the Wien, Rayleigh-Jeans and Planck laws at $T = 1000K$.

mechanics and he was required to assume that the oscillators in the walls of the cavity could only radiate at a specific energy

$$E = h\nu \,. \tag{1.7}$$

He referred to these bundles of energy as "quanta." This is a very radical departure from classical mechanics because according to electrodynamics the energy of the radiation should be determined by the magnitude of the oscillations and be independent of the frequency. Planck assumed that there must be some unknown physics associated with the production of radiation of the oscillators, but that the description of the radiation in the cavity should still be consistent with Maxwell's equations which had recently been verified by an number of experiments. This is historically the beginning of quantum mechanics.

1.2.2   The Photoelectric Effect

One of Einstein's papers of 1905 considered Planck's derivation of his radiation law. Einstein was well aware of the errors and conjectures that were necessary to Planck's derivation. As a result Einstein believed that Plank's Law was consistent with experiment but not with existing theory while the Rayleigh-Jeans Law was consistent with existing theory but not with experiment. He then proceeded to use Boltzmann statistics to examine the radiation in the regime where Wien's Law is consistent with data

and derived this result. This proof was also deficient[1], but in the process he made the hypothesis that the light in the cavity was quantized, in contrast to Planck's assumption that it was the material oscillators that were quantized. The light-quantum hypothesis was of course not consistent with the classical physics of Maxwell's equations where the energy of the wave is related to the square of the electric field and not to the frequency. This led to the work that was to lead to Einstein's 1922 Nobel prize.

The first application of the light-quantum hypothesis was to the photoelectric effect. This effect was first seen by Hertz in 1887 in connection with his experiments with electromagnetic radiation. He noticed that the light from one electric arc could effect the magnitude of the current in a second arc. In 1888 Hallwachs showed that ultraviolet light falling on a conductor could give it a positive charge. Since this was before the discovery of the electron, the nature of this effect was a mystery.

In 1899, J. J. Thomson showed that the charges emitted in the photolectric effect were electrons which he had identified in cathode rays in 1897.

In 1902 Lenard examined the photoelectric effect using a carbon arc light source which could be varied in intensity by a factor of 1000. He made the surprising discovery that the maximum energy of electrons given off in the photoelectric effect was independent of the intensity of the light. This is in contradiction to classical electrodynamics where the energy provided by the light source depends only on the intensity. In addition, he determined that the energy of the electrons increased with the frequency of the light, again in contradiction to classical theory.

In 1905 Einstein proposed that the photoelectric effect could be understood in terms of the light-quantum hypothesis. If the light quanta have an energy of $h\nu$ then the maximum energy of the emitted electrons should follow the formula

$$E_{\max} = h\nu - P \tag{1.8}$$

where P is the amount of energy required to remove an electron from the conductor and is called the work function. Experimental confirmation of this formula was provided in 1916 by Millikan who showed that this formula was consistent with his data to within 0.5%.

In spite of this stunning confirmation, the light-quantum hypothesis was viewed with considerable skepticism by the majority of physicists at the time.

1.3   Atomic Physics

The other path that led to the establishment of quantum mechanics was through atomic physics. As we have already seen a considerable amount of information had been collected during the nineteenth century associated with the regularities seen in

---

[1]Einstein would return to this problem several times during the career, but a completely satisfactory derivation of Planck's Law was not achieved until Dirac did so in 1927.

the chemical properties of various elements and with the very large amount of data that had been obtained on the spectra of the elements. Any acceptable theory of the atom would necessarily need to account for these phenomena. It had been noted by Maxwell in 1875 that atoms must have many degrees of freedom in order to produce the complicated spectra that were being observed. This implied that the atoms must have some complicated structure since a rigid body with only six degrees of freedom would not be sufficient to describe the data. The problem here is that until the last few years of the nineteenth century any clues as to the physical structure of the atom was missing.

1.3.1 Cathode Rays

One of the first advances in this area involved the study of cathode rays. Cathode rays are seen as luminous discharges when current flows through partially evacuated tubes. This phenomenon had been known since the eighteenth century and demonstrations of it had been a popular entertainment. However, since pressures in these tubes could only be lowered by a small amount compared to atmospheric pressure, there was a sufficient amount of gas in the tube that a great many secondary effects were present, so it was difficult to study the cathode rays themselves. About a third of the way through the nineteenth century it became possible to produce tubes with much higher vacuums and to start to consider the primary effect. The nature of these rays was a topic of some dispute. Some physicists (mostly English) believed that the rays were due to the motions of charged particles while others (mostly German) believed that the rays were actually due to flows or disturbances of the ether.

Hertz showed in 1891 that cathode rays could pass through thin metal foils. His student Lenard then produced tubes with thin metal windows which would allow the cathode rays to be extracted from the tube. Using these he showed in 1894 that the cathode rays could not be molecular and that they could be bent in an external electric field. Also in 1894 Thomson showed that the cathode rays moved with a velocity substantially smaller than the speed of light. In 1895 Perrin placed a small metal cup in a cathode ray tube to collect the rays and showed that the cathode rays carried a negative charge as had been indicated by the direction of deviation of the rays external fields.

In 1897 Wiechert, Kaufmann and Thomson each performed experiments with the deflection of cathode rays in magnetic (or in the case of Thomson electric and magnetic) fields. By measuring the deflection of the cathode rays it was possible to determine the ratio of the charge to the mass, e/m. In all cases it was shown that the this ratio was on the order of 2000 times that of singly ionized hydrogen. This could of course be either due to a large charge or a small mass. Both Wiechert and Thomson speculated that the cause was the small mass of the particles constituting the cathode rays. In 1899 Thomson was able to measure the charge of the constituents of the cathode rays separately using the newly invented Wilson cloud chamber. He

showed that this was roughly the same as the charge of ionized hydrogen determined from electrolysis. This then proved that the mass of the cathode ray particle was indeed much smaller than the hydrogen mass. Thus, the electron was born as the first subatomic particle.

Also in 1899, as was previously noted, Thomson determined that the particles emitted by the photoelectric effect were also electrons and thus that ionization was the result of removing electrons from atoms. That is the atom was no longer immutable and could be broken down into constituent parts. This led to Thomson's model of the atom. This model assumed that the atom was composed of electrons moving in the electric field of some positive background charge which was assumed to uniformly distributed over the volume of the atom. Initially Thomson proposed that there were as many as a thousand electrons in the atom, but later came to believe that the number of electrons was on the order of the atomic number of the atom. This model is often referred to as the "plum pudding" model of the atom.

1.3.2 Radioactive Decay

In 1895 Roentgen discovered X-rays while experimenting with cathode ray tubes. While this has no direct impact on the development of quantum mechanics, it stimulated a great deal of experimental activity. One of those who was stimulated to look into the problem was Becquerel. Since the source of the X-rays appeared to come from the luminous spot on the wall of the tube struck by the cathode ray, Becquerel hypothesized that X-rays were associated with florescence. To test this hypothesis, in 1896 he began studying whether a phosphorescent uranium salt could expose photographic plates wrapped in thick black paper. He would expose the salt to sun light to cause them to fluoresce and then would place it on top of the plate. He found that it did indeed expose the photographic plate. At one point during the experiment the weather turned cloudy and he was unable to subject the salt to sunlight so he placed it along with the plate in a closed cupboard. After several days, he developed the plate and discovered that it had been exposed. Therefore the exposure of the plate was not the result of the florescence after all. He also discovered that any uranium salt, even those that were not phosphorescent, also exposed the plates. The radiation was, therefore, a property of uranium. He found that the uranium continued to radiate energy continuously with no apparent diminution over a considerable period of time.

The hunt was now on for more sources of Becquerel rays. In 1897 both the Curies and Rutherford became engaged in the problem. The Curies soon began to discover a variety of known elements such as thorium which gave off the rays and also new radioactive elements such as radium and radon. They also proceeded to give themselves radiation poisoning. To this day their laboratory books are sufficiently radioactive that they are kept in a lead lined vault.

Rutherford showed in 1898 that uranium gave off two different types of rays which

he called α- and β-rays. The β-rays were much more penetrating than the α-rays. The β-rays were soon identified as electrons. It was suspected that the α-rays were related to helium since this element seemed to appear in the gases given off by radioactive materials, but it was not until 1908 that Rutherford made a completely convincing case that the α-rays were doubly-ionized helium atoms or helium nuclei.

For the purposes of our introduction to quantum mechanics this has two important consequences, one which adds another puzzle to the list of problems with classical physics and another which led to a greater understanding of the structure of the atom. The first of these resulted from the publication by Rutherford and Soddy of the transformation theory. In this they proposed that radioactive decay causes one type of atom to be transformed into another kind. This was consistent with the pattern of radioactive elements that was being established. This is clearly a problem for those who thought that atoms were immutable, but for our purposes the real problem lies in the other part of this theory. It was proposed that the number of atoms that decayed in a given time period was proportional to the number of atoms present. Mathematically this is expressed as

$$\frac{dN}{dt} = -\lambda N \tag{1.9}$$

where $N(t)$ is the number of atoms of a given type at time t. This leads to the exponential decay law

$$N(t) = N(0)e^{-\lambda t} . \tag{1.10}$$

The question that arises from this is: Why does one uranium atom decay now while another seemingly identical atom decays in 10,000 years? Clearly, from classical physics it should be expected that once the atom is created it should be possible to determine exactly when it will decay. You could argue that the atoms were created 10,000 years apart and were indeed decaying in the same way, but it can be shown that radioactive elements created by the decay of another radioactive element within a short period of time will also satisfy the same decay law. The apparent statistical nature of radioactive decays is a direct challenge to classical physics.

The second aspect of radioactive decays that is important to the story of quantum mechanics is also associated with Rutherford. Because energetic α-particles were given off by radioactive decays and these α-particles could be columnated into a beam, the α-particles could be used as a probe of the structure of the atom. In 1908 Rutherford and Geiger demonstrated that they could indeed be scattered from atoms. These experiments were very tedious and required that observers sit in the dark looking for light flashes given off by the scattered α-particles when the hit a florescent screen. Each of these events had to be carefully counted and recorded. In 1909 Rutherford suggested to Geiger and a young undergraduate named Marsden, who was assisting him with the counting, that they should look for α-particles that were scattered through more than 90 degrees. They found that 1 in 8000 of the α-particles were indeed deflected by more than 90 degrees. This was extremely surprising since

if the positive charge in the atom was distributed over the volume of the atom, many fewer α-particles should be scattered at such large angles. This immediately led Rutherford to see that the positive charge in the atom should be concentrated in a very small part of the atom. This, in turn, resulted in the Rutherford model of the atom where the electrons orbit around a very small positive nucleus.

This is clearly an improvement over the Thomson model since it explains the new scattering data, but as a classical model it does little to satisfy the criteria that the model deal with all of the previous data collected for atoms. There is nothing in either of these models that explains the regularities discovered in the atoms nor does it explain why the spectra should show discrete lines rather than a continuous spectrum which should be expected from a classical orbital model. In addition, neither of these two models is stable. Since any electron moving in a confined space must accelerate, classical electrodynamics predicts that the electrons in these atomic models should be continually emitting light until they spiral into the center of the atom and come to a stop.

1.3.3   The Bohr-Sommerfeld Model

Before proceeding to the Bohr model it is necessary to make a small digression. In 1885 Balmer considered four spectral lines in the spectrum of hydrogen measured by Ångstrum in 1868. These were referred to as H_α, H_β, H_γ and H_δ. He noted that the ratios of the frequencies of the these states could be written as simple fractions and that these could be summarized by the expression

$$\nu = R \left(\frac{1}{4} - \frac{1}{n^2} \right) . \tag{1.11}$$

After reporting this work he was informed that there were additional data available and these fit his formula with very good accuracy.

Apparently, Bohr did not know of the Balmer formula until he was informed of it by a Danish colleague in 1913. Once he knew of the Balmer formula he saw a way to obtain an expression that would reproduce the formula for hydrogen-like atoms. To do this he assumed that there had to be stable solutions which he called stationary states the describe the ground and excited states of atoms. He then assumed that the spectra were due to light emitted when an electron moves from a higher energy state to a lower energy state. That is,

$$h\nu = E_m - E_n \tag{1.12}$$

where $E_m > E_n$.

The model that he constructed assumed that the atom was described by an electron in a circular orbit around a nucleus with positive charge Ze. Classically, these

circular orbits will be stable if the Coulomb force on the electron produces the appropriate centripetal acceleration. This is given by[2]

$$\frac{Ze^2}{r^2} = m_e \frac{v^2}{r} \tag{1.13}$$

where m_e is the mass of the electron and v is its orbital speed. This implies that

$$\frac{Ze^2}{r} = m_e v^2 . \tag{1.14}$$

This is, of course, true for any circular orbit and classically there are a continuous set of such orbits corresponding to any choice the radius r. To select the set of allowable stationary states, Bohr imposed the quantum condition that the kinetic energy for the stationary states is fixed such that

$$\frac{1}{2} m_e v^2 = \frac{n}{2} h f \tag{1.15}$$

where h is Planck's constant and f is the orbital frequency of the electron. The orbital frequency is in turn given by

$$f = \frac{v}{2\pi r} . \tag{1.16}$$

Substituting (1.16) into (1.15) leads to the expression

$$v = n \frac{h}{2\pi} \frac{1}{m_e r} = \frac{n\hbar}{m_e r} \tag{1.17}$$

where we have defined the constant $\hbar = \frac{h}{2\pi} = 1.05457 \times 10^{-34}$ J·s. We can now substitute (1.17) into the stability condition (1.14) which yields

$$\frac{Ze^2}{r} = \frac{n^2 \hbar^2}{m_e r^2} . \tag{1.18}$$

This can now be solved for the radius of the stationary state corresponding to the integer n to give

$$r_n = \frac{n^2 \hbar^2}{m_e Z e^2} . \tag{1.19}$$

We can now now calculate the energy. First using (1.14) we can write

$$E = \frac{1}{2} m_e v^2 - \frac{Ze^2}{r} = \frac{1}{2} \frac{Ze^2}{r} - \frac{Ze^2}{r} = -\frac{1}{2} \frac{Ze^2}{r} . \tag{1.20}$$

---

[2]Here I am choosing to express the Coulomb force with constants appropriate for the esu system of electromagnetic units. This is somewhat simpler and the expressions can always be rewritten in terms of dimensionless quantities with values independent of the system of units.

The energy of the nth stationary state can now be calculated using this and (1.19) giving

$$E_n = -\frac{1}{2}\frac{Ze^2}{\frac{n^2\hbar^2}{m_e Ze^2}} = -\frac{Z^2e^4m_e}{2\hbar^2n^2}. \tag{1.21}$$

We can now define a new dimensionless constant

$$\alpha = \frac{e^2}{\hbar c} \cong \frac{1}{137}. \tag{1.22}$$

The energy can now be rewritten as

$$E_n = -\frac{Z^2\alpha^2m_ec^2}{2n^2}. \tag{1.23}$$

The frequency of light that is emitted from the transition from a state n to a state m, where $n > m$ can now be written as

$$\nu_{nm} = \frac{E_n - E_m}{h} = \frac{Z^2\alpha^2m_ec^2}{4\pi\hbar}\left(\frac{1}{m^2} - \frac{1}{n^2}\right). \tag{1.24}$$

If we identify

$$R = \frac{Z^2\alpha^2m_ec^2}{4\pi\hbar} \tag{1.25}$$

it is clear that the Balmer series corresponds to the special case where $m = 2$.

It also useful to consider two alternate forms of (1.17). First we can rewrite this equation as

$$m_e rv = n\hbar. \tag{1.26}$$

The left-hand side of this is just the angular momentum for a particle moving with uniform speed in a circle, so

$$L = n\hbar \tag{1.27}$$

means that we can also state the quantization condition as the quantization of angular momentum. The second form is

$$m_e v 2\pi r = nh. \tag{1.28}$$

The left-hand side of this is just the momentum times the circumference of the circular orbit. This can be generalized as

$$\int d\boldsymbol{l} \cdot \boldsymbol{p} = nh. \tag{1.29}$$

The left-hand side is called the *action* so this form of the condition means that the action is quantized. When Sommerfeld extended the Bohr model to allow for elliptical orbits and relativistic corrections it was the action form of the quantization condition that was used.

13

Now lets return to the expression for the radius of the Bohr orbitals (1.19) for the case of hydrogen ($Z = 1$). This can be rewritten as

$$r_n = \frac{\hbar}{m_e c \alpha} n^2 = a_0 n^2 \tag{1.30}$$

$a_0 = 5.292 \times 10^{-11}$m is called the Bohr radius and is the radius of the ground state of the Bohr hydrogen atom. Although a_0 is a small number, n^2 grows very rapidly. For the the radius to be 1mm,

$$n = \sqrt{\frac{1.0 \times 10^{-3}\text{m}}{5.292 \times 10^{-11}\text{m}}} \cong 4350 \,. \tag{1.31}$$

Bohr had now introduced a new theory for atoms, but all of the accumulated theory and observations show that classical physics works a macroscopic scales. It is, therefore, necessary for the quantum theory to reproduce classical physics when the size of the object becomes on the macroscopic scale. This called the *classical correspondence principle*. Bohr stated this by observing that for large values of n his theory should reproduce the classical result. We can see how this occurs for the Bohr atom by noting that a classical electron moving in a circle with a positive charge at the center will radiate at the a frequency equal to the orbital frequency of the electron. That is

$$\nu_{\text{cl}} = f = \frac{v}{2\pi r} = \frac{\frac{n\hbar}{m_e r}}{2\pi r} = \frac{n\hbar}{2\pi m_e r^2} = \frac{n\hbar}{2\pi m_e} \frac{m_e^2 c^2 \alpha^2}{\hbar^2 n^4} = \frac{\alpha^2 m_e c^2}{2\pi \hbar n^2} \,. \tag{1.32}$$

Now consider the frequency for the Bohr model when the electron moves from a state n to a state $n - 1$ using (1.24) for the case of hydrogen. This gives

$$\nu_{n,n-1} = \frac{\alpha^2 m_e c^2}{4\pi \hbar} \left(\frac{1}{(n-1)^2} - \frac{1}{n^2} \right) = \frac{\alpha^2 m_e c^2}{4\pi \hbar} \frac{2n-1}{n^2(n-1)^2} \,. \tag{1.33}$$

In the limit where n becomes large this yields

$$\nu_{n,n-1} \cong \frac{\alpha^2 m_e c^2}{2\pi \hbar n^3} \tag{1.34}$$

which agrees with the classical result. So the Bohr atom obeys the classical correspondence principle.

As we will see, quantum mechanics in its current form is constructed such that it satisfies the classical correspondence principle.

The Bohr atom was revolutionary. For the first time it was possible to reproduce spectroscopic data and a great number of advances were made in physics under its influence. It did, however, have substantial problems. While it predicts spectra well for hydrogen and singly ionized helium, it does a poor job of reproducing the spectra of the neutral helium atom. Even when spectra a predicted, the model cannot account for the intensity of the spectral lines or for their widths. Bohr was able to make remarkable number of predictions with the model in conjunction with the correspondence principle, but in the end it did not provide a sufficient basis to move forward with the study of quantum systems.

1.4   Wave Particle Duality

At this point let's return to Einstein and the light-quantum hypothesis. In 1905 when Einstein first put forward this hypothesis, he simply stated that light consisted of quanta with energy $E = h\nu$, he did not propose that light was actually composed of particles. Over time, as Einstein applied the light-quantum hypothesis to more problems, his thinking on this problem evolved and by 1917 he was willing to state that the light quantum carried a momentum with magnitude given by

$$p = \frac{h\nu}{c} \, . \tag{1.35}$$

Note that for a massless particle special relativity requires that

$$p = \frac{E}{c} \, . \tag{1.36}$$

Using the light quantum value for E then gives the momentum result proposed by Einstein. At this point we now have light quanta with the properties of massless particle which would later be given the name of *photon*.

   We are now left with a sizable dilemma: light behaves like waves since we know that light can produce interference and diffraction but we also know that light behaves like particles in black body radiation and the photoelectric effect. How can something be both a particle and a wave? That light has both wave and particle properties is called *wave-particle duality*.

   The immediate problem for Einstein was that there were few people who believed that light really had particle-like properties. The skepticism was eliminated by the Compton scattering experiment of 1923.

1.4.1   Compton Scattering

Consider a massless photon scattering form an electron at rest. The photon scatters at an angle θ from the incident photon. Let $\boldsymbol{k}$ be the three-momentum of the initial photon, $\boldsymbol{k}'$ be the momentum of the scattered photon and $\boldsymbol{p}'$ be the momentum of the scattered electron. Since the electron is initially at rest, momentum conservation requires that

$$\boldsymbol{k} = \boldsymbol{k}' + \boldsymbol{p}' \tag{1.37}$$

and energy conservation requires that

$$h\nu + m_e c^2 = h\nu' + \sqrt{c^2 \boldsymbol{p}'^2 + m_e^2 c^4} \, . \tag{1.38}$$

Equation (1.37) can be rewritten as

$$\boldsymbol{k} - \boldsymbol{k}' = \boldsymbol{p}' \, . \tag{1.39}$$

Squaring this we get

$$\boldsymbol{p}'^2 = \boldsymbol{k}^2 + \boldsymbol{k}'^2 - 2\boldsymbol{k} \cdot \boldsymbol{k}' = \boldsymbol{k}^2 + \boldsymbol{k}'^2 - 2|\boldsymbol{k}||\boldsymbol{k}'|\cos\theta\,. \tag{1.40}$$

Now we can rewrite (1.38) as

$$\sqrt{c^2\boldsymbol{p}'^2 + m_e^2 c^4} = h(\nu - \nu') + m_e c^2\,. \tag{1.41}$$

Squaring this gives

$$c^2\boldsymbol{p}'^2 + m_e^2 c^4 = h^2(\nu - \nu')^2 + 2m_e c^2 h(\nu - \nu') + m_e^2 c^4 \tag{1.42}$$

or

$$c^2\boldsymbol{p}'^2 = h^2(\nu' - \nu)^2 - 2m_e c^2 h(\nu' - \nu)\,. \tag{1.43}$$

Substituting (1.40) into (1.43) gives

$$c^2\boldsymbol{k}^2 + c^2\boldsymbol{k}'^2 - 2c^2|\boldsymbol{k}||\boldsymbol{k}'|\cos\theta = h^2(\nu - \nu')^2 - 2m_e c^2 h(\nu - \nu')\,. \tag{1.44}$$

Since

$$|\boldsymbol{k}| = \frac{h\nu}{c} \tag{1.45}$$

and

$$|\boldsymbol{k}'| = \frac{h\nu'}{c}\,, \tag{1.46}$$

this can be rewritten as

$$h^2\nu^2 + h^2\nu'^2 - 2h^2\nu\nu'\cos\theta = h^2(\nu - \nu')^2 + 2m_e c^2 h(\nu - \nu')\,. \tag{1.47}$$

Adding and subtracting $2h^2\nu\nu'$ from the left-hand side to complete the square gives

$$h^2(\nu - \nu')^2 + 2h^2\nu\nu' - 2h^2\nu\nu'\cos\theta = h^2(\nu - \nu')^2 + 2m_e c^2 h(\nu - \nu')\,. \tag{1.48}$$

Canceling the first terms on each side and dividing the equation by $2h^2\nu\nu'$ yields

$$1 - \cos\theta = \frac{m_e c^2}{h}\left(\frac{1}{\nu'} - \frac{1}{\nu}\right)\,. \tag{1.49}$$

Finally, using $\lambda = \frac{c}{\nu}$ we can rewrite this as

$$\lambda' - \lambda = \frac{h}{m_e c}(1 - \cos\theta)\,. \tag{1.50}$$

The constant $\frac{h}{m_e c}$ must have the dimension of length and is called the *Compton wavelength*.

Both Debye and Compton produced papers with proofs of this result. In addition Compton and his collaborators provided excellent experimental verification of the result. This directly demonstrated that light could behave as a particle.

16
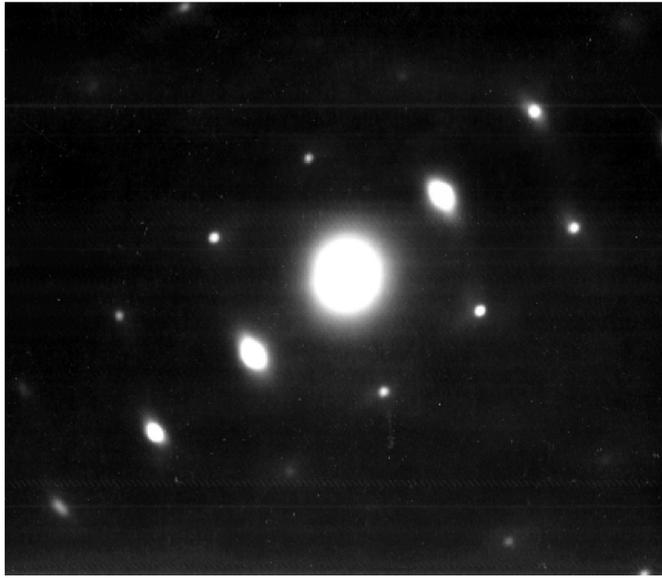
Figure 1.2: Electron diffraction pattern for a silver-gold alloy.

1.4.2 Electron Diffraction

In 1923 Louis Victor, Duc de Broglie was working in the physics laboratory of his older brother Maurice, Prince de Broglie. He had been giving considerable attention to Einstein's light-quantum hypothesis. It occurred to him that if light which is classically a wave could behave as a particle, then classical particles could also behave as quantum waves. That is, he extended the wave-particle duality from light to particles. He proposed that the wavelength of a massive particle could be given by

$$\lambda = \frac{h}{p}\,. \tag{1.51}$$

This is now called the *de Broglie wavelength.*

The implications of this for the Bohr model can be seen by considering the quantization condition (1.28). This can be rewritten as

$$2\pi r = n\frac{h}{p} = n\lambda\,. \tag{1.52}$$

That is, the circumference of the Bohr orbit is equal to an integral number of de Broglie wavelengths. de Broglie called these waves *pilot waves.*

If particles can act as waves, it should be possible to see the diffraction and interference effects characteristic of waves with particles as well. Davisson and Germer demonstrated the diffraction of electrons from crystals in 1927. An example of an electron diffraction pattern for an silver-gold alloy is shown in Fig. 1.2. This phenomenon is now routinely used as an analytic tool in science and engineering.

We are now left with the problem of how we can interpret a world where waves are particle and particles are waves. The resolution of this problem is associated with the development of quantum matrix mechanics by Heisenberg and of quantum wave mechanics of Schrödinger. Although the matrix mechanics appeared first in 1925 with the wave mechanics appearing about half of a year later in 1926, we will begin our treatment of quantum dynamics by considering Schrödinger's formulation of the theory.

Chapter 2

Mathematical Background to Quantum Mechanics

Before beginning our development of quantum mechanics, it is useful to first review linear algebra and analysis which forms the mathematical structure that we will use to describe the theory.

2.1   Vector Spaces

Consider a set of vectors for which the operations of vector addition and multiplication by a scalar are defined. The set forms a linear vector space if any operation of addition or scalar multiplication yields a vector in the original set. This means that the vector space is *closed* under these operations. It is also required that for arbitrary vectors $\boldsymbol{A}$, $\boldsymbol{B}$ and $\boldsymbol{C}$ the operation of addition has the properties:

$$\boldsymbol{A} + \boldsymbol{B} = \boldsymbol{B} + \boldsymbol{A} \tag{2.1}$$

and

$$\boldsymbol{A} + (\boldsymbol{B} + \boldsymbol{C}) = (\boldsymbol{A} + \boldsymbol{B}) + \boldsymbol{C} \,. \tag{2.2}$$

There must also be a *null vector* $\boldsymbol{0}$ which acts as the identity under addition. That is

$$\boldsymbol{A} + \boldsymbol{0} = \boldsymbol{A} \,. \tag{2.3}$$

Each vector $\boldsymbol{A}$ must also have an inverse $-\boldsymbol{A}$ such that

$$\boldsymbol{A} + (-\boldsymbol{A}) = \boldsymbol{0} \,. \tag{2.4}$$

The operation of scalar multiplication has the properties

$$\begin{aligned}
\alpha(\boldsymbol{A} + \boldsymbol{B}) &= \alpha\boldsymbol{A} + \alpha\boldsymbol{B} & (2.5) \\
(\alpha + \beta)\boldsymbol{A} &= \alpha\boldsymbol{A} + \beta\boldsymbol{A} & (2.6) \\
\alpha(\beta\boldsymbol{A}) &= (\alpha\beta)\boldsymbol{A} \,, & (2.7)
\end{aligned}$$

where α and β are scalars. If the scalars are restricted to real numbers, then the the vector space is real. If the scalars are complex numbers, then the vector space is complex. For the moment we will restrict ourselves to real vector spaces. Clearly, the usual three-dimensional vectors that we use to describe particle motion satisfy these properties.

A set of vectors $\{\boldsymbol{V}_1, \boldsymbol{V}_2, \ldots, \boldsymbol{V}_n\}$ is said to be *linearly independent* if

$$\sum_{i=1}^{n} \alpha_i \boldsymbol{V}_i \neq \boldsymbol{0} \tag{2.8}$$

except when all of the α_i are zero. Any set linearly independent vectors form a *basis* and for an n-dimensional vector space the will be n vectors in any basis set. A basis set is said to *span* the vector space. The definition of implies that if we add some arbitrary vector $\boldsymbol{V}$ to the basis set we could now write

$$\boldsymbol{V} - \sum_{i=1}^{n} \alpha_i \boldsymbol{V}_i = \boldsymbol{0} \tag{2.9}$$

or

$$\boldsymbol{V} = \sum_{i=1}^{n} \alpha_i \boldsymbol{V}_i \, . \tag{2.10}$$

This means that any vector in the vector space can be written as a linear combination of the basis vectors. This is the principle of linear superposition. The coefficients α_i in the expansion are the *components* of the vector relative to the chosen basis set. For a three-dimensional vector space, any three non-coplanar vector will form a basis.

One of the reasons for decomposing vectors into components relative to a set of basis vectors is that if

$$\boldsymbol{A} = \sum_{i=1}^{n} \alpha_i \boldsymbol{V}_i \tag{2.11}$$

and

$$\boldsymbol{B} = \sum_{i=1}^{n} \beta_i \boldsymbol{V}_i \, , \tag{2.12}$$

then

$$\boldsymbol{A} + \boldsymbol{B} = \sum_{i=1}^{n} \alpha_i \boldsymbol{V}_i + \sum_{i=1}^{n} \beta_i \boldsymbol{V}_i = \sum_{i=1}^{n} (\alpha_i + \beta_i) \boldsymbol{V}_i \, . \tag{2.13}$$

The addition of vectors then becomes equivalent to the addition of components.

2.1.1 The Scalar Product

An additional property that we associate with vectors is definition of a *scalar or inner product* which combines two vectors to give a scalar. A vector space on which a scalar

product is defined is called an *inner product space.* The scalar product which we are used to is defined as

$$\boldsymbol{A} \cdot \boldsymbol{B} = |\boldsymbol{A}||\boldsymbol{B}| \cos\gamma \,, \tag{2.14}$$

where $\boldsymbol{A}$ and $\boldsymbol{B}$ are real vectors, $|\boldsymbol{A}|$ and $|\boldsymbol{B}|$ are the lengths of the two vectors, and γ is the angle between the vectors. This immediately implies that the inner product of a vector with itself is

$$\boldsymbol{A} \cdot \boldsymbol{A} = |\boldsymbol{A}|^2 \tag{2.15}$$

and that if $\boldsymbol{A}$ and $\boldsymbol{B}$ are perpendicular to one another ($\gamma = 90°$) then

$$\boldsymbol{A} \cdot \boldsymbol{B} = 0 \,. \tag{2.16}$$

The inner product is distributive. That is

$$\boldsymbol{A} \cdot (\boldsymbol{B} + \boldsymbol{C}) = \boldsymbol{A} \cdot \boldsymbol{B} + \boldsymbol{A} \cdot \boldsymbol{C} \,. \tag{2.17}$$

The definition of the inner product in terms of components relative to a set of basis vectors can be simplified by choosing a set of basis vectors which have unit length and are all mutually perpendicular. Such a basis is called an *orthonormal* basis. If we label the orthonoral basis vectors as $\boldsymbol{e}_i$ for $i = 1, \ldots, n$ this means that

$$\boldsymbol{e}_i \cdot \boldsymbol{e}_j = \delta_{ij} \qquad \text{for } i, j = 1, \ldots, n \,. \tag{2.18}$$

If

$$\boldsymbol{A} = \sum_{i=1}^{n} \alpha_i \boldsymbol{e}_i \tag{2.19}$$

and

$$\boldsymbol{B} = \sum_{j=1}^{n} \beta_j \boldsymbol{e}_j \,, \tag{2.20}$$

Then

$$\boldsymbol{A} \cdot \boldsymbol{B} = \sum_{i=1}^{n} \alpha_i \beta_i \,. \tag{2.21}$$

This implies that we can now identify the components relative to an orthonormal basis as

$$\boldsymbol{A} \cdot \boldsymbol{e}_i = \boldsymbol{A} = \sum_{j=1}^{n} \alpha_j \boldsymbol{e}_j \cdot \boldsymbol{e}_i = \sum_{j=1}^{n} \alpha_j \delta_{ij} = \alpha_i \,. \tag{2.22}$$

It should be noted that it is quite possible to deal with vector algebra in nonorthogonal bases, but the mathematics is much more complicated.

21

2.1.2 Operators

Now consider the possibility that there are operators that can act on a vector in the vector space and transform it into another vector in the space. That is

$$\boldsymbol{\Omega x} = \boldsymbol{y} \,, \tag{2.23}$$

where $\boldsymbol{x}$ and $\boldsymbol{y}$ are vectors in the space and $\boldsymbol{\Omega}$ is some operator on the vectors in the space. An example of such an operator would be a rotation of a vector about some axis. If the operator has the properties

$$\boldsymbol{\Omega}(c\boldsymbol{x}) = c(\boldsymbol{\Omega x}) \tag{2.24}$$

and

$$\boldsymbol{\Omega}(\boldsymbol{x} + \boldsymbol{y}) = \boldsymbol{\Omega x} + \boldsymbol{\Omega y} \tag{2.25}$$

the operator is a *linear operator*.

2.1.3 Matrix Notation

It is convenient to represent a vector in terms of a $1 \times n$ column matrix which has the components of the vector relative to an orthonormal basis as elements. For example the three-dimensional vector $\underline{x}$ can be represented given by the column vector

$$\underline{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \,, \tag{2.26}$$

where $x_i = \boldsymbol{x} \cdot \boldsymbol{e}_i$. The transpose of the matrix is

$$\underline{x}^T = \begin{pmatrix} x_1 & x_2 & x_3 \end{pmatrix} \,. \tag{2.27}$$

For two real vectors $\boldsymbol{x}$ and $\boldsymbol{y}$, the inner product can then be written as

$$\boldsymbol{x} \cdot \boldsymbol{y} = \underline{x}^T \underline{y} = \sum_{i=1}^{3} x_i y_i \,. \tag{2.28}$$

If the vectors are complex, the inner product must be defined as

$$\boldsymbol{x} \cdot \boldsymbol{y} = \sum_{i=1}^{3} x_i^* y_i \tag{2.29}$$

since the inner product of a vector with itself must give the length of the vector which is a real number. If we define the *hermitian conjugate* of a vector as

$$\underline{x}^\dagger = (\underline{x}^*)^T \,, \tag{2.30}$$

then the inner product can be defined generally as

$$\boldsymbol{x} \cdot \boldsymbol{y} = \underline{x}^\dagger \underline{y} \,. \tag{2.31}$$

A linear operator acting on a vector produces another vector in the vector space. So, if we represent the vectors as column matrices, the operator acting on vector must also result in a column matrix. This means that an operator can be represented by an $n \times n$ matrix and the action of the operator on a vector is simply described as matrix multiplication. That is

$$\underline{y} = \underline{\underline{\Omega}} \, \underline{x} \tag{2.32}$$

where the components of $\underline{\underline{\Omega}}$ are Ω_{ij} and the components of $\underline{y}$ can be written as

$$y_i = \sum_{j=1}^{n} \Omega_{ij} x_j. \tag{2.33}$$

In order to simplify the notation for this kind of matrix transformation, it is convenient to use the *Einstein summation convention*. This convention assumes that any two repeated indices are to be summed over the appropriate range unless it is specifically stated otherwise. Using this convention we can write

$$y_i = \Omega_{ij} x_j \,, \tag{2.34}$$

where it is assumed that the repeated index j is summed from 1 to n.

As an example the operator that produces a rotation of a three-dimensional vector through an angle ϕ about $\boldsymbol{e}_3$ is represented by the matrix

$$\underline{\underline{R}}_3(\phi) = \begin{pmatrix} \cos\phi & -\sin\phi & 0 \\ \sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix} \,. \tag{2.35}$$

Since the operators are represented as matrices, they have the algebraic properties of matrices under multiplication. These properties are:

1. Using the distributive property of matrix multiplication,

$$\underline{\underline{A}} \left(\underline{x}_1 + \underline{x}_2 \right) = \underline{\underline{A}} \, \underline{x}_1 + \underline{\underline{A}} \, \underline{x}_2 \,. \tag{2.36}$$

2. Since matrix multiplication is associative,

$$\left(\underline{\underline{A}} \, \underline{\underline{B}} \right) \underline{x} = \underline{\underline{A}} \left(\underline{\underline{B}} \, \underline{x} \right) \,. \tag{2.37}$$

3. Multiplying a matrix by a constant is defined such that

$$(c\underline{\underline{A}})_{ij} = c A_{ij} \,. \tag{2.38}$$

4. In general,

$$\underline{\underline{A}}\,\underline{\underline{B}} \neq \underline{\underline{B}}\,\underline{\underline{A}}\,, \tag{2.39}$$

or

$$\left[\underline{\underline{A}}, \underline{\underline{B}}\right] \neq 0\,, \tag{2.40}$$

where

$$\left[\underline{\underline{A}}, \underline{\underline{B}}\right] \equiv \underline{\underline{A}}\,\underline{\underline{B}} - \underline{\underline{B}}\,\underline{\underline{A}} \tag{2.41}$$

defines the commutator of the matrices $\underline{\underline{A}}$ and $\underline{\underline{B}}$.

5. The identity matrix is the matrix with elements

$$(\underline{\underline{1}})_{ij} = \delta_{ij}\,. \tag{2.42}$$

Then,

$$\underline{\underline{1}}\,\underline{\underline{A}} = \underline{\underline{A}}\,\underline{\underline{1}} = \underline{\underline{A}}\,. \tag{2.43}$$

6. If $\det(\underline{\underline{A}}) \neq 0$, then $\underline{\underline{A}}$ has an inverse $\underline{\underline{A}}^{-1}$ such that

$$\underline{\underline{A}}^{-1}\underline{\underline{A}} = \underline{\underline{A}}\,\underline{\underline{A}}^{-1} = \underline{\underline{1}}\,. \tag{2.44}$$

7. The *transpose* of a matrix is defined such that

$$\left(\underline{\underline{A}}^T\right)_{ij} = A_{ji}\,. \tag{2.45}$$

8. The *complex conjugate* of a matrix is defined such that

$$\left(\underline{\underline{A}}^*\right)_{ij} = (A_{ij})^*\,. \tag{2.46}$$

9. The *hermitian conjugate* of a matrix is defined such that

$$\underline{\underline{A}}^\dagger = \left(\underline{\underline{A}}^*\right)^T\,. \tag{2.47}$$

So,

$$\left(\underline{\underline{A}}^\dagger\right)_{ij} = (A_{ji})^*\,. \tag{2.48}$$

10. $\underline{\underline{A}}$ is *hermitian* if

$$\underline{\underline{A}}^\dagger = \underline{\underline{A}}\,. \tag{2.49}$$

11. $\underline{\underline{A}}$ is *antihermitian* if

$$\underline{\underline{A}}^\dagger = -\underline{\underline{A}}\,. \tag{2.50}$$

24

12. $\underline{\underline{A}}$ is *unitary* if

$$\underline{\underline{A}}^\dagger = \underline{\underline{A}}^{-1} \,. \tag{2.51}$$

Unitary transformations of the form

$$\underline{y} = \underline{\underline{A}}\,\underline{x} \tag{2.52}$$

have the hermitian conjugate

$$\underline{y}^\dagger = \underline{x}^\dagger \underline{\underline{A}}^\dagger = \underline{x}^\dagger \underline{\underline{A}}^{-1} \,. \tag{2.53}$$

So,

$$\underline{y}^\dagger \underline{y} = \underline{x}^\dagger \underline{\underline{A}}^{-1} \underline{\underline{A}}\,\underline{x} = \underline{x}^\dagger \underline{x} \,. \tag{2.54}$$

This means that unitary transformations preserve the *norm* of a vector.

2.1.4 The Eigenvalue Problem

A problem of particular interest to us is the *eigenvalue problem*

$$\underline{\underline{A}}\,\underline{x} = \lambda \underline{x} \tag{2.55}$$

where λ is a constant. That is, we want to find all vectors $\underline{x}$ which when multiplied by $\underline{\underline{A}}$ give the original vector times some constant λ. This can be rewritten as

$$\left(\underline{\underline{A}} - \lambda \underline{\underline{1}}\right)\underline{x} = 0. \tag{2.56}$$

If $\underline{\underline{A}} - \lambda \underline{\underline{1}}$ has an inverse, then the only solution would be the trivial solution where $\underline{x} = 0$. Therefore, this must not have an inverse which implies that

$$\det\left(\underline{\underline{A}} - \lambda \underline{\underline{1}}\right) = 0 \,. \tag{2.57}$$

In three dimensions this will produce a cubic polynomial with three roots. So there will be three *eigenvalues* λ_i and three corresponding *eigenvectors* $\underline{x}_i$. In n dimensions this will be a polynomial of order n and there will be n eigenvalues with n corresponding eigenvectors.

Consider the case where $\underline{\underline{A}}$ is hermitian. The eigenvalue equation for the eigenvalue λ_i is

$$\underline{\underline{A}}\,\underline{x}_i = \lambda_i \underline{x}_i \tag{2.58}$$

where in this case the repeated index is not summed. If $\underline{\underline{A}}$ is hermitian, the hermitian conjugate of the eigenvalue equation for the eigenvalue equation for eigenvalue λ_j is

$$\underline{x}_j^\dagger \underline{\underline{A}}^\dagger = \underline{x}_j^\dagger \underline{\underline{A}} = \lambda_j^* \underline{x}_j^\dagger \,. \tag{2.59}$$

We can no multiply the first expression on the left by $\underline{x}_j^\dagger$ and the second on the right by $\underline{x}_i$ and then subtract the two expressions to give

$$\underline{x}_j^\dagger \left(\underline{\underline{A}} - \underline{\underline{A}}\right)\underline{x}_i = \left(\lambda_i - \lambda_j^*\right)\underline{x}_j^\dagger \underline{x}_i \,, \tag{2.60}$$

25

or
$$\left(\lambda_i - \lambda_j^*\right) \underline{x}_j^\dagger \underline{x}_i = 0 \,. \tag{2.61}$$

For the case where $i = j$, $\underline{x}_i^\dagger \underline{x}_i > 0$ if the eigenvector is to be nontrivial. This then requires that

$$\lambda_i - \lambda_i^* = 0 \,. \tag{2.62}$$

Therefore, the λ_i must be real for all i. Now, for $i \neq j$, if the eigenvalues are not degenerate this requires that

$$\underline{x}_j^\dagger \underline{x}_i = 0 \,. \tag{2.63}$$

So the eigenvectors are orthogonal. In the case where two or more eigenvalues have the same value which means that they are *degenerate*, the eigenvectors associated with these eigenvalues will be orthogonal to all of the remaining eigenvectors, but may not be mutually orthogonal. In this case it is necessary to construct linear combinations of the eigenvectors that are mutually orthogonal. Now multiply (2.58) by some arbitrary constant c. This gives

$$c\underline{\underline{A}}\,\underline{x}_i = c\lambda_i \underline{x}_i \,, \tag{2.64}$$

or

$$\underline{\underline{A}}\,(c\underline{x}_i) = \lambda_i(c\underline{x}_i) \,. \tag{2.65}$$

This means that the normalization of the eigenvectors are not determined by the eigenvalue equation. As a result we can always choose the the eigenvectors to be of unit length. We have now shown that for a hermitian matrix we will always obtain real eigenvalues and the eigenvectors can be normalized to for an orthonormal set of basis vectors.

Now consider the case where a set of vectors $\underline{x}_i$ are eigenvector for two different operators $\underline{\underline{\Lambda}}$ and $\underline{\underline{\Omega}}$ such that

$$\underline{\underline{\Lambda}}\,\underline{x}_i = \lambda_i \underline{x}_i \tag{2.66}$$

and

$$\underline{\underline{\Omega}}\,\underline{x}_i = \omega_i \underline{x}_i \,. \tag{2.67}$$

Using the eigenvalue equation we can write

$$\underline{\underline{\Lambda}}\,\underline{\underline{\Omega}}\,\underline{x}_i = \underline{\underline{\Lambda}}\omega_i \underline{x}_i = \omega_i \underline{\underline{\Lambda}}\,\underline{x}_i = \omega_i \lambda_i \underline{x}_i = \underline{\underline{\Omega}}(\lambda_i \underline{x}_i) = \underline{\underline{\Omega}}\,\underline{\underline{\Lambda}}\,\underline{x}_i \,, \tag{2.68}$$

where it has been assumed that the operators are linear in the second and fourth steps. This implies that

$$(\underline{\underline{\Lambda}}\,\underline{\underline{\Omega}} - \underline{\underline{\Omega}}\,\underline{\underline{\Lambda}})\underline{x}_i = [\underline{\underline{\Lambda}}, \underline{\underline{\Omega}}]\underline{x}_i = \underline{0} \,. \tag{2.69}$$

Since this must be true for all of the eigenvectors in the set, it follows that

$$[\underline{\underline{\Lambda}}, \underline{\underline{\Omega}}] = \underline{\underline{0}} \,. \tag{2.70}$$

26

That is, operators with common eigenvectors commute.

Next, let the vectors $\underline{x}_i$ represent a normalized set of eigenvectors of the operator $\underline{\underline{A}}$ such that

$$\underline{\underline{A}}\,\underline{x}_i = \lambda_i \underline{x}_i\,. \tag{2.71}$$

We can now define the *outer product* of these eigenvectors as

$$\underline{\underline{\Xi}}_i = \underline{x}_i \underline{x}_i^\dagger\,. \tag{2.72}$$

By the rules of matrix multiplication, the $\underline{\underline{\Xi}}_i$ must be an $n \times n$ matrix. Now, define

$$\underline{\underline{\Xi}} = \sum_{i=1}^n \underline{\underline{\Xi}}_i = \sum_{i=1}^n \underline{x}_i \underline{x}_i^\dagger\,. \tag{2.73}$$

Consider the case where this operates on an arbitrary eigenvector in this set. That is

$$\underline{\underline{\Xi}}\,\underline{x}_j = \sum_{i=1}^n \underline{\underline{\Xi}}_i \underline{x}_j = \sum_{i=1}^n \underline{x}_i \underline{x}_i^\dagger \underline{x}_j = \sum_{i=1}^n \underline{x}_i \delta_{ij} = \underline{x}_j\,. \tag{2.74}$$

Therefore,

$$\sum_{i=1}^n \underline{x}_i \underline{x}_i^\dagger = \underline{\underline{1}}\,. \tag{2.75}$$

This is called the *completeness relation* for the set of eigenvectors.

Again, consider an operator $\underline{\underline{A}}$ which has eigenvectors $\underline{x}_i$ with eigenvalues λ_i. That is,

$$\underline{\underline{A}}\,\underline{x}_i = \lambda_i \underline{x}_i\,. \tag{2.76}$$

With the solution of the eigenvalue problem, we now have two bases that have been defined, the original basis set given by the vectors $\underline{e}_i$ and the basis set composed of the normalized eigenvectors $\underline{x}_i$. That is the eigenvectors are defined in terms of the original basis vectors as

$$\underline{x}_i = \sum_{j=1}^n \underline{e}_j \underline{e}_j^\dagger \underline{x}_i \tag{2.77}$$

and the components of the operator matrix are given by

$$(\underline{\underline{A}})_{ij} = \underline{e}_i^\dagger \underline{\underline{A}}\,\underline{e}_j\,. \tag{2.78}$$

We can also expand the original basis vectors in terms of the eigenvectors by using the completeness of the eigenvectors to give

$$\underline{e}_i = \underline{\underline{1}}\,\underline{e}_i = \sum_{j=1}^n \underline{x}_j \underline{x}_j^\dagger \underline{e}_i\,. \tag{2.79}$$

Similarly, we can now define a new matrix operator $\widetilde{\underline{\underline{A}}}$ which is defined in the basis of eigenvectors and has components given by

$$\widetilde{\underline{\underline{A}}}_{kl} = \underline{x}_k^\dagger \underline{\underline{A}}\, \underline{x}_l = \underline{x}_k^\dagger \lambda_l \underline{x}_l = \lambda_l \underline{x}_k^\dagger \underline{x}_l = \lambda_l \delta_{kl} \,. \tag{2.80}$$

This means that the operator represented in the basis of eigenvectors is diagonal. These components can be rewritten in terms of the original matrix representation of the operator by using the completeness of the basis vectors $\underline{e}_i$ as

$$
\begin{aligned}
(\widetilde{\underline{\underline{A}}})_{kl} &= \underline{x}_k^\dagger \sum_{i=1}^n \underline{e}_i \underline{e}_i^\dagger \underline{\underline{A}} \sum_{j=1}^n \underline{e}_j \underline{e}_j^\dagger \underline{x}_l \\
&= \sum_{i=1}^n \sum_{j=1}^n \underline{x}_k^\dagger \underline{e}_i \underline{e}_i^\dagger \underline{\underline{A}}\, \underline{e}_j \underline{e}_j^\dagger \underline{x}_l \\
&= \sum_{i=1}^n \sum_{j=1}^n \underline{x}_k^\dagger \underline{e}_i (\underline{\underline{A}})_{ij} \underline{e}_j^\dagger \underline{x}_l \,.
\end{aligned}
\tag{2.81}
$$

Now, define the matrix $\underline{\underline{U}}$ with components given by

$$(\underline{\underline{U}})_{jl} = \underline{e}_j^\dagger \underline{x}_l = (\underline{x}_l)_j \tag{2.82}$$

and note that

$$\underline{x}_k^\dagger \underline{e}_i = \sum_{m=1}^n (\underline{x}_k)_m^* (\underline{e}_i)_m = \sum_{m=1}^n \left((\underline{e}_i)_m^* (\underline{x}_k)_m\right)^* = \left(\underline{e}_i^\dagger \underline{x}_k\right)^* = (\underline{\underline{U}})_{ik}^* = (\underline{\underline{U}}^\dagger)_{ki} \,. \tag{2.83}$$

Then,

$$(\widetilde{\underline{\underline{A}}})_{kl} = \sum_{i=1}^n \sum_{j=1}^n (\underline{\underline{U}}^\dagger)_{ki} (\underline{\underline{A}})_{ij} (\underline{\underline{U}})_{jl} \tag{2.84}$$

or

$$\widetilde{\underline{\underline{A}}} = \underline{\underline{U}}^\dagger \, \underline{\underline{A}}\, \underline{\underline{U}} \,. \tag{2.85}$$

This transformation then converts from one representation of the operator to the other. From the definition of $\underline{\underline{U}}$ we can write that

$$\sum_{k=1}^n (\underline{\underline{U}}^\dagger)_{ik} (\underline{\underline{U}})_{kj} = \sum_{k=1}^n \underline{x}_i^\dagger \underline{e}_k \underline{e}_k^\dagger \underline{x}_j = \underline{x}_i^\dagger \underline{x}_j = \delta_{ij} \tag{2.86}$$

or

$$\underline{\underline{U}}^\dagger \underline{\underline{U}} = \underline{\underline{1}} \,. \tag{2.87}$$

This shows that

$$\underline{\underline{U}}^\dagger = \underline{\underline{U}}^{-1} \,, \tag{2.88}$$

28

so this matrix is unitary and the transformation

$$\widetilde{\underline{\underline{A}}} = \underline{\underline{U}}^{-1}\underline{\underline{A}}\,\underline{\underline{U}} \tag{2.89}$$

is called a *unitary transformation*. This means that the operator $\underline{\underline{A}}$ is diagonalized by a unitary transformation. This transformation can be inverted to give

$$\underline{\underline{A}} = \underline{\underline{U}}\,\widetilde{\underline{\underline{A}}}\,\underline{\underline{U}}^{-1}. \tag{2.90}$$

Eventually we will need to define functions of operators. In this case, the function is always defined as a power series

$$f(\underline{\underline{A}}) = \sum_{n=0}^{\infty} c_n \underline{\underline{A}}^n \tag{2.91}$$

where the expansion coefficients are the same as the Taylor series for the same function of a scalar. That is

$$f(z) = \sum_{n=0}^{\infty} c_n z^n. \tag{2.92}$$

Unless the matrix $\underline{\underline{A}}$ has special algebraic properties, the evaluation of the function as an infinite series can be very difficult. This can be simplified by using the unitary transformation that diagonalizes the operator. To see how this works consider the third power of the operator. Using the unitary of the transformation matrix we can write

$$\underline{\underline{A}}^3 = \underline{\underline{A}}\,\underline{\underline{A}}\,\underline{\underline{A}} = \underline{\underline{U}}\,\underline{\underline{U}}^{-1}\underline{\underline{A}}\,\underline{\underline{U}}\,\underline{\underline{U}}^{-1}\underline{\underline{A}}\,\underline{\underline{U}}\,\underline{\underline{U}}^{-1}\underline{\underline{A}}\,\underline{\underline{U}}\,\underline{\underline{U}}^{-1} = \underline{\underline{U}}\,\widetilde{\underline{\underline{A}}}\,\widetilde{\underline{\underline{A}}}\,\widetilde{\underline{\underline{A}}}\,\underline{\underline{U}}^{-1} = \underline{\underline{U}}\,\widetilde{\underline{\underline{A}}}^3\underline{\underline{U}}^{-1}. \tag{2.93}$$

This helpful because any power of a diagonal matrix is given by simply raising the diagonal elements to the same power. That is,

$$(\widetilde{\underline{\underline{A}}}^n)_{ij} = \lambda_i^n \delta_{ij}. \tag{2.94}$$

It is easy to see how this method can be extended to any power of the operator. Then,

$$f(\underline{\underline{A}}) = \sum_{n=0}^{\infty} c_n \underline{\underline{U}}\,\widetilde{\underline{\underline{A}}}^n\underline{\underline{U}}^{-1} = \underline{\underline{U}}\left(\sum_{n=0}^{\infty} c_n \widetilde{\underline{\underline{A}}}^n\right)\underline{\underline{U}}^{-1} = \underline{\underline{U}}f(\widetilde{\underline{\underline{A}}})\,\underline{\underline{U}}^{-1}, \tag{2.95}$$

where

$$\left(f(\widetilde{\underline{\underline{A}}})\right)_{ij} = f(\lambda_i)\delta_{ij}. \tag{2.96}$$

Example: Find the eigenvalues and eigenvectors for the matrix

$$\underline{\underline{A}} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}. \tag{2.97}$$

The eigenvalues are obtained by requiring that

$$\det \left(\underline{\underline{A}} - \lambda \underline{\underline{1}} \right) = 0. \tag{2.98}$$

Or

$$0 = \det \begin{pmatrix} -\lambda & 1 & 0 \\ 1 & -\lambda & 1 \\ 0 & 1 & -\lambda \end{pmatrix} = -\lambda \left(\lambda^2 - 1 \right) - 1 \left(-\lambda \right) = -\lambda \left(\lambda^2 - 2 \right) \tag{2.99}$$

So

$$\lambda = 0, \pm \sqrt{2}. \tag{2.100}$$

Now the eigenvectors must satisfy the equation

$$\begin{pmatrix} -\lambda & 1 & 0 \\ 1 & -\lambda & 1 \\ 0 & 1 & -\lambda \end{pmatrix} \begin{pmatrix} 1 \\ a \\ b \end{pmatrix} = 0. \tag{2.101}$$

Only two of the linear equations produced by this matrix expression can be independent since the matrix is required to have a determinant of zero. The eigenvalue problem determines the eigenvectors only up to a multiplicative constant. For this reason, the top component of the column vector has been arbitrarily chosen to be 1. We will choose the normalization of the eigenvectors such that they are unit vectors. This leads to three linear equations

$$-\lambda + a = 0 \tag{2.102}$$
$$1 - \lambda a + b = 0 \tag{2.103}$$
$$a - \lambda b = 0. \tag{2.104}$$

The first of these can be solved to give

$$a = \lambda. \tag{2.105}$$

The second can be solved to give

$$b = \lambda a - 1 = \lambda^2 - 1. \tag{2.106}$$

So for $\lambda = \sqrt{2}$,

$$\underline{x}_1 = N_1 \begin{pmatrix} 1 \\ \sqrt{2} \\ 1 \end{pmatrix}. \tag{2.107}$$

30

The normalization constant is chosen such that $\underline{x}_i^\dagger \underline{x}_i = 1$. So

$$1 = N_i^2 \begin{pmatrix} 1 & a^* & b^* \end{pmatrix} \begin{pmatrix} 1 \\ a \\ b \end{pmatrix} = N_i^2 (1 + |a|^2 + |b|^2) . \tag{2.108}$$

Solving this for N_i gives

$$N_i = \frac{1}{\sqrt{1 + |a|^2 + |b|^2}} . \tag{2.109}$$

Therefore, for this eigenvector

$$N_1 = \frac{1}{\sqrt{1 + 2 + 1}} = \frac{1}{2} . \tag{2.110}$$

For $\lambda = 0$,

$$\underline{x}_2 = N_2 \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} , \tag{2.111}$$

where

$$N_2 = \frac{1}{\sqrt{1 + 0 + 1}} = \frac{1}{\sqrt{2}} . \tag{2.112}$$

For $\lambda = -\sqrt{2}$,

$$\underline{x}_3 = N_2 \begin{pmatrix} 1 \\ -\sqrt{2} \\ 1 \end{pmatrix} , \tag{2.113}$$

where

$$N_3 = \frac{1}{\sqrt{1 + 2 + 1}} = \frac{1}{2} . \tag{2.114}$$

Note that for these eigenvectors $\underline{x}_i^\dagger \underline{x}_j = \delta_{ij}$. So the eigenvectors for a set of orthonormal basis vectors.

2.2 The Continuum Limit

Consider a system of n equal masses connected by equivalent springs between the masses and connecting the end masses to fixed walls. The system is allowed to move in one dimension. The case of $n = 5$ is shown in Fig. 2.1. The equilibrium position of mass i is given by x_i. The equations of motion of the masses can be written in
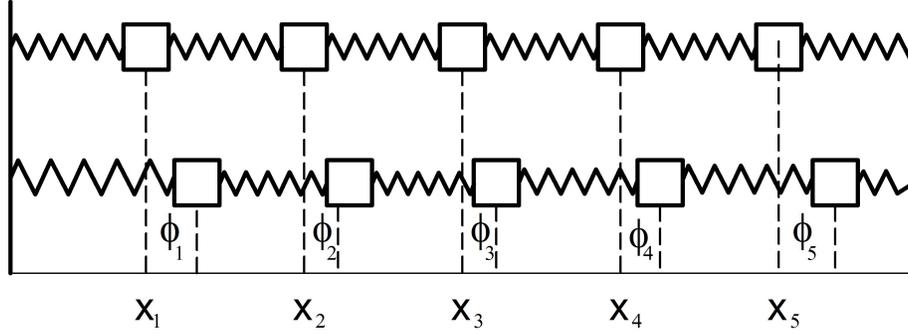
Figure 2.1: Five equal masses coupled by identical springs. The top part of the figure shows the equilibrium position while the bottom part shows the system with displaced masses. The equilibrium positions are given by x_i and the displacement of each mass from equilibrium is given by ϕ_i.

terms of the displacement of the masses ϕ_i from their equilibrium positions by

$$
\begin{aligned}
m\ddot{\phi}_1(t) &= k\left(\phi_2(t) - \phi_1(t)\right) - k\phi_1(t) \\
m\ddot{\phi}_2(t) &= k\left(\phi_3(t) - \phi_2(t)\right) - k\left(\phi_2(t) - \phi_1(t)\right) \\
m\ddot{\phi}_3(t) &= k\left(\phi_4(t) - \phi_3(t)\right) - k\left(\phi_3(t) - \phi_2(t)\right) \\
&\vdots \\
m\ddot{\phi}_{n-1}(t) &= k\left(\phi_n(t) - \phi_{n-1}(t)\right) - k\left(\phi_{n-1}(t) - \phi_{n-2}(t)\right) \\
m\ddot{\phi}_n(t) &= -k\phi_n(t) - k\left(\phi_n(t) - \phi_{n-1}(t)\right)
\end{aligned}
\tag{2.115}
$$

This set of coupled second-order differential equation can be written in matrix form as

$$
\begin{pmatrix}
\ddot{\phi}_1(t) \\
\ddot{\phi}_2(t) \\
\ddot{\phi}_3(t) \\
\ddot{\phi}_4(t) \\
\vdots \\
\ddot{\phi}_{n-1}(t) \\
\ddot{\phi}_n(t)
\end{pmatrix}
= \frac{k}{m}
\begin{pmatrix}
-2 & 1 & 0 & 0 & \cdots & 0 & 0 & 0 \\
1 & -2 & 1 & 0 & \cdots & 0 & 0 & 0 \\
0 & 1 & -2 & 1 & \cdots & 0 & 0 & 0 \\
0 & 0 & 1 & -2 & \cdots & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & 0 & \cdots & 1 & -2 & 1 \\
0 & 0 & 0 & 0 & \cdots & 0 & 1 & -2
\end{pmatrix}
\begin{pmatrix}
\phi_1(t) \\
\phi_2(t) \\
\phi_3(t) \\
\phi_4(t) \\
\vdots \\
\phi_{n-1}(t) \\
\phi_n(t)
\end{pmatrix} .
\tag{2.116}
$$

Since we would expect that the masses will oscillate when set in motion, it is logical to look for solutions of the form

$$
\phi_i(t) = \eta_i e^{i\omega t} .
\tag{2.117}
$$

Substituting this into the matrix form of the equations of motion gives

$$
\frac{k}{m}
\begin{pmatrix}
2 & -1 & 0 & 0 & \cdots & 0 & 0 & 0 \\
-1 & 2 & -1 & 0 & \cdots & 0 & 0 & 0 \\
0 & -1 & 2 & -1 & \cdots & 0 & 0 & 0 \\
0 & 0 & -1 & -2 & \cdots & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & 0 & \cdots & -1 & 2 & -1 \\
0 & 0 & 0 & 0 & \cdots & 0 & -1 & 2
\end{pmatrix}
\begin{pmatrix}
\eta_1 \\ \eta_2 \\ \eta_3 \\ \eta_4 \\ \vdots \\ \eta_{n-1} \\ \eta_n
\end{pmatrix}
= \omega^2
\begin{pmatrix}
\eta_1 \\ \eta_2 \\ \eta_3 \\ \eta_4 \\ \vdots \\ \eta_{n-1} \\ \eta_n
\end{pmatrix} .
\tag{2.118}
$$

This is now an eigenvalue problem with n eigenvalues ω_i^2 and corresponding n eigenvectors η_i. This is an example of a common problem in classical mechanics involving small oscillations of coupled oscillators. The eigenvectors are called the normal modes of the system and the corresponding frequencies are the normal mode frequencies. This can be solved for any value of n.

It is interesting to consider the case where $n \to \infty$ with the total mass fixed. We can do this be examining one of the equations of motion

$$
m\ddot{\phi}_i(t) = k\left(\phi_{i+1}(t) - \phi_i(t)\right) - k\left(\phi_i(t) - \phi_{i-1}(t)\right) .
\tag{2.119}
$$

It is convenient at this point to change the way that we label the masses. Instead of using the subscript i, we can label each mass by its equilibrium position $x_i = i\Delta x$ where

$$
\Delta x = \frac{L}{n+1} ,
\tag{2.120}
$$

with L being the separation between the anchoring walls. The differential equation can then be rewritten as

$$
m\frac{\partial^2}{\partial t^2}\phi(x_i, t) = k\left(\phi(x_{i+1}, t) - \phi(x_i, t)\right) - k\left(\phi(x_i, t) - \phi(x_{i-1}, t)\right) .
\tag{2.121}
$$

If the total mass of the system is M then the mass of each particle is

$$
m = \frac{M}{n} .
\tag{2.122}
$$

We can now rewrite the differential equation as

$$
m\frac{\partial^2}{\partial t^2}\phi(x_i, t) = k\left(\phi(x_i + \Delta x, t) - \phi(x_i, t)\right) - k\left(\phi(x_{i-1} + \Delta x, t) - \phi(x_{i-1}, t)\right) .
\tag{2.123}
$$

Since as n becomes large Δx will be come small for fixed L, we can expand the displacement functions to first order in Δx to give

$$
\begin{aligned}
m\frac{\partial^2}{\partial t^2}\phi(x_i, t) &= k\frac{\partial \phi}{\partial x}(x_i, t)\Delta x - k\frac{\partial \phi}{\partial x}(x_{i-1}, t)\Delta x \\
&= k\Delta x\left(\frac{\partial \phi}{\partial x}(x_i, t) - \frac{\partial \phi}{\partial x}(x_i - \Delta x, t)\right) \\
&= k(\Delta x)^2\frac{\partial^2 \phi}{\partial x^2}(x_i, t) .
\end{aligned}
\tag{2.124}
$$

Note that

$$m = \frac{M}{n} = \frac{M}{L}\frac{L}{n} = \frac{M}{L}\frac{n+1}{n}\Delta x \cong \frac{M}{L}\Delta x \qquad (2.125)$$

for large n. If we define the linear mass density as

$$\mu = \frac{M}{L} \qquad (2.126)$$

and Young's Modulus as

$$Y = k\Delta x\,, \qquad (2.127)$$

we obtain the wave equation

$$\frac{\partial^2}{\partial x^2}\phi(x,t) - \frac{1}{v^2}\frac{\partial^2}{\partial t^2}\phi(x,t) = 0\,, \qquad (2.128)$$

where

$$v^2 = \frac{Y}{M}\,. \qquad (2.129)$$

This is a second-order partial differential equation in x and t. Since the ends of the system are fixed, the amplitude function must satisfy the spatial boundary conditions

$$\phi(0,t) = \phi(L,t) = 0\,. \qquad (2.130)$$

Any particular solution must also involve the imposition of initial conditions for ϕ and $\frac{\partial \phi}{\partial t}$.

Recall that for any finite n, there will be n eigenvalues (eigenfrequencies) and eigenvectors. Since the wave equation is the result of the limiting behavior of the system when the $n \to \infty$, we should expect that we should be able to express the solution of the wave equation in terms of an eigenvalue problem with an infinite number of eigenfrequencies and eigenvectors. The standard method for obtaining such a solution involves the procedure of separation of variables. To do this we assume that the amplitude function can be factored such that

$$\phi(x,t) = X(x)T(t)\,. \qquad (2.131)$$

Substituting this into (2.128) and rearranging gives

$$\frac{1}{X(x)}\frac{\partial^2 X(x)}{\partial x^2} - \frac{1}{v^2 T(t)}\frac{\partial^2 T(t)}{\partial t^2} = 0\,. \qquad (2.132)$$

Since each of the two terms depends only on a single variable, the only way this can be satisfied for all values of x and t is for each of the two terms to be constant, with the constants adding to 0. If we choose the constant such that

$$\frac{1}{X(x)}\frac{\partial^2 X(x)}{\partial x^2} = -k^2\,, \qquad (2.133)$$

34

then we obtain the differential equation

$$\frac{\partial^2 X(x)}{\partial x^2} = -k^2 X(x) \,, \tag{2.134}$$

or

$$\frac{\partial^2 X(x)}{\partial x^2} + k^2 X(x) = 0 \,. \tag{2.135}$$

This is the familiar harmonic equation which has solutions of the form

$$X(x) = c_1 \cos(kx) + c_2 \sin(kx) \,. \tag{2.136}$$

The first of the spatial boundary conditions requires that $X(0) = 0$ which requires that $c_1 = 0$. The second boundary condition requires that $X(L) = c_2 \sin(kL) = 0$. This can only be satisfied if

$$kL = n\pi \,, \tag{2.137}$$

where $n > 0$. Then for each value of n we have

$$k_n = \frac{\pi n}{L} \tag{2.138}$$

Note that (2.134) is of a form similar to a matrix eigenequation, where the matrix is replaced by the *differential operator* $\frac{\partial^2}{\partial x^2}$ and the eigenvalue is given by $-k^2$. The eigenvectors are replaced by the functions

$$X_n(x) = N_n \sin\left(\frac{n\pi x}{L}\right) \,. \tag{2.139}$$

Since there are an infinite number of values for n, there are an infinite number of eigenvalues and eigenvectors or *eigenfunctions*.

The differential operator is *linear differential operator* since

$$\frac{\partial^2}{\partial x^2} a f(x) = a \frac{\partial^2}{\partial x^2} f(x) \,, \tag{2.140}$$

where a is a complex constant, and

$$\frac{\partial^2}{\partial x^2} \left(f(x) + g(x)\right) = \frac{\partial^2 f(x)}{\partial x^2} + \frac{\partial^2 g(x)}{\partial x^2} \,. \tag{2.141}$$

This means that we can expand any nonsingular function defined on the interval $0 \leq x \leq L$ as a linear combination of the eigenfunctions. That is

$$f(x) = \sum_{n=1}^{\infty} a_n X_n(x) \,, \tag{2.142}$$

where the a_n are complex constants.

We can now define an *inner product* as

$$f \cdot g = \int_0^L dx \, f^*(x) g(x) \,. \tag{2.143}$$

The inner product of any two eigenfunctions is then

$$\int_0^L dx \, X_m^*(x) X_n(x) = N_m N_n \int_0^L dx \, \sin\left(\frac{m\pi x}{2L}\right) \sin\left(\frac{n\pi x}{2L}\right) = N_n^2 \frac{L}{2} \delta_{mn} \,. \tag{2.144}$$

This means that the eigenfunctions are orthogonal and can be normalized by choosing

$$N_n = \sqrt{\frac{2}{L}} \,. \tag{2.145}$$

The completeness relation for the normalized eigenfunctions is

$$\sum_{n=1}^\infty X_n(x) X_n^*(x') = \delta(x - x') \,, \tag{2.146}$$

where $\delta(x - x')$ is the Dirac delta function and is defined such that:

1. $\delta(x - a) = 0$ for all $x \neq a$.

2.
$$\int_{x_0}^{x_1} dx \, \delta(x - a) = \begin{cases} 1 & \text{if } x_0 < a < x_1 \\ 0 & \text{if } a < x_0 \text{ or } a > x_1 \end{cases} \tag{2.147}$$

3.
$$\int_{x_0}^{x_1} dx \, f(x) \delta(x - a) = \begin{cases} f(a) & \text{if } x_0 < a < x_1 \\ 0 & \text{if } a < x_0 \text{ or } a > x_1 \end{cases} \tag{2.148}$$

4.
$$\int_{x_0}^{x_1} dx \, f(x) \delta'(x - a) = \begin{cases} -f'(a) & \text{if } x_0 < a < x_1 \\ 0 & \text{if } a < x_0 \text{ or } a > x_1 \end{cases} \tag{2.149}$$

5.
$$\delta(f(x)) = \sum_i \frac{1}{\left|\frac{df}{dx}\right|_{x=x_i}} \delta(x - x_i) \tag{2.150}$$

where the x_i are all of the simple zeros of $f(x)$. A common particular case of this identity is

$$\delta(cx) = \frac{1}{|c|} \delta(x) \,, \tag{2.151}$$

where c is a constant.

This kind of infinite-dimensional inner-product vector space with the eigenvectors composed of continuous functions defined on some interval is called a *Hilbert space.*

If we now substitute (2.133) into (2.132) and rearrange the result, we obtain

$$\frac{\partial^2 T(t)}{\partial t^2} + v^2 k^2 T(t) = 0 \,. \tag{2.152}$$

Again this is the harmonic equation. Since we have already fixed the eigenvalue using the spatial equation, if we define

$$\omega_n = v k_n \,, \tag{2.153}$$

the solution to this differential equation is of the form

$$T_n(t) = A_n \cos(\omega_n t) + B_n \sin(\omega_n t) \,. \tag{2.154}$$

The complete wave function for a given n is then

$$\phi_n(x, t) = X_n(t) T_n(t) = \sqrt{\frac{2}{L}} \sin(k_n x) \left(A_n \cos(\omega_n t) + B_n \sin(\omega_n t) \right) \tag{2.155}$$

and a general solution for the wave function is

$$\phi(x, t) = \sum_{n=1}^{\infty} \phi_n(x, t) = \sum_{n=1}^{\infty} \sqrt{\frac{2}{L}} \sin(k_n x) \left(A_n \cos(\omega_n t) + B_n \sin(\omega_n t) \right) \,. \tag{2.156}$$

The constants A_n and B_n are determined by imposing the initial conditions in time,

$$\phi(x, 0) = f(x) \tag{2.157}$$

and

$$\left. \frac{\partial \phi(x, t)}{\partial t} \right|_{t=0} = g(x) \,. \tag{2.158}$$

We have now outlined the mathematical foundations that we will use to describe quantum mechanics. The mathematical solution of most quantum mechanical problems begins with the *time-dependent Schrödinger equation* which is a partial differential equation that is second order in the spatial variables and first order in time. A particular solution to this equation requires imposing appropriate boundary conditions in space and time. The basic procedure for solution follows that presented above and can be summarized as:

1. A separable solution in space and time is assumed.

2. Substituting this solution in the Schrödinger equation and separating variables results in an eigenvalue equation in the spatial variables called the *time-independent Schrödinger equation.*

3. Imposing the spatial boundary conditions determines the eigenvalues and eigenfunctions.

4. We will show that the differential operator in the spatial equation is hermitian. This implies that the eigenvalues will be real and the eigenfunctions are orthonormal.

5. The differential equation in time can be solved for each eigenvalue.

6. The complete wave function for each eigenvalue can then be constructed.

7. The general solution of the time-dependent Schrödinger equation can be constructed as a linear combination of these solutions where the coefficients in the expansion are determined by the initial condition in time.

We will now examine the physical origins and interpretation of the Schrödinger equation.

Chapter 3

The Schrödinger Equation

3.1 Wave Particle Duality

Now we are left with the problem of having to make sense out the reality that waves can also behave like particles and particles can behave like waves. To motivate how we are going to find a formulation of quantum mechanics that allows this to happen, it is useful to consider the classical situation. In particular consider the case of light waves and the two-slit diffraction problem.

Assume that we have a plane wave illuminating the slits from the left. When the waves strike the screen with the slits, we can describe the part of the wave which passes through the slits using Huygen's principle. As the wave front strikes the slit, each point in the slit radiates a spherical wave and the electromagnetic wave emanating from the slit is the total of the spherical waves from all of the points. The wave from the upper slit can be represented by the electric field $\boldsymbol{E}_1(\boldsymbol{r}, t)$ and that from the lower slit as $\boldsymbol{E}_2(\boldsymbol{r}, t)$. The intensity of the wave is proportional to the absolute square of the electric field. If the bottom slit is covered and light can pass only through the upper slit we see an image of the slit which is proportional to $|\boldsymbol{E}_1(\boldsymbol{r}, t)|^2$ on the second screen and similarly covering the upper slit while light passes through the lower slit gives an image of slit two proportional to $|\boldsymbol{E}_2(\boldsymbol{r}, t)|^2$. This situation is illustrated in Fig. 3.1 where the intensity of the light along the second screen is illustrated by the curve to the right of the screen.

Now in the case when both slits are open, the total field is given by the sum of the fields due to each slit $\boldsymbol{E}(\boldsymbol{r}, t) = \boldsymbol{E}_1(\boldsymbol{r}, t) + \boldsymbol{E}_2(\boldsymbol{r}, t)$ since Maxwell's equations are linear. The intensity at the screen will then be proportional to

$$|\boldsymbol{E}_1(\boldsymbol{r}, t) + \boldsymbol{E}_2(\boldsymbol{r}, t)|^2 = |\boldsymbol{E}_1(\boldsymbol{r}, t)|^2 + |\boldsymbol{E}_2(\boldsymbol{r}, t)|^2 + 2\boldsymbol{E}_1(\boldsymbol{r}, t) \cdot \boldsymbol{E}_2(\boldsymbol{r}, t) \,. \qquad (3.1)$$

The result is illustrated by Fig. 3.2. The interference pattern is due to the cross term (the interference term).

Now consider the situation where the the two slits are now being illuminated by a beam of electrons which is larger than the distance between the slits and where the
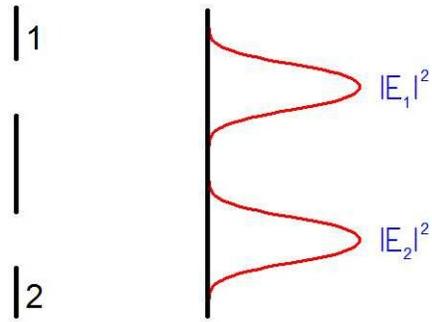
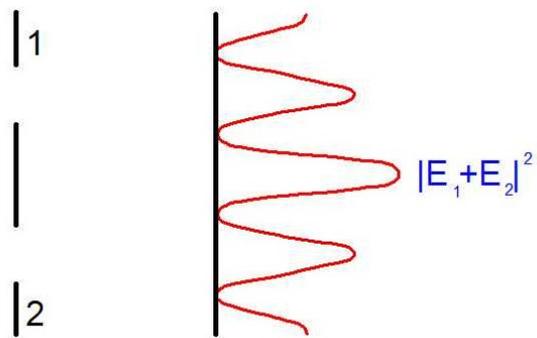Figure 3.1: Images of the slits is illuminated individually

Figure 3.2:

directions of the velocities of the electrons vary by a small amount. If the electrons behaved only like particles the result would appear as in the case of Fig. 3.1 with the total count of electrons hitting the screen represented by the sum of the two curves on the right. However, since we know from our previous discussions that the electrons have wave-like properties, the result is an interference pattern as illustrated in Fig. 3.2. Furthermore, the interference pattern is made up of a large number of individual electrons striking the screen so that the electrons can indeed still be seen as particles. One way of looking at this phenomenon is to introduce de Broglie's pilot waves. These waves determine in some fashion where the electrons will go. This means that we would replace the electric field from slit one in our classical example by a *wave function* $\Psi_1(\boldsymbol{r}, t)$ and that from slit 2 by $\Psi_2(\boldsymbol{r}, t)$. In order for us to get the interference pattern, these two wave functions must add to give the total wave function. This requires that the wave equation, of which these wave functions Ψ_1 and Ψ_2 are solutions, must be a linear differential equation. The number of electrons that strike the second screen at a given time must then be represented by the absolute square of the total wave function. That is, $|\Psi_1(\boldsymbol{r}, t) + \Psi_2(\boldsymbol{r}, t)|^2$.

It is reasonable to ask the question: "Which slit does each electron pass through?" Imagine that we place a laser beam on the right side of the screen containing the two slits with the beam parallel to this screen. We can then detect whether an electron passes through slit 1 or slit 2 by looking for a flash of light due to Compton scattering of a photon in the laser beam from each electron as it passes through a slit. Now we perform the experiment again and record not only the position of the electron as it strikes the second screen, but also which slit it passes through. All of the events can then be classified by which slit the electron passed through. If we now assemble all of these events together we will see that the distribution is given by $|\Psi_1|^2 + |\Psi_2|^2$. The interference pattern has disappeared and we get the result that we would get if the electrons only behaved as particles.

The source of this problem is that in order to measure the positions of the electrons to determine through which slit they passed, we have scattered photon's from the electrons. This scattering affects the motion of the electrons and eliminates the interference pattern. One way we can imagine that we could reduce this effect is to decrease the intensity of the laser beam. This means that there are fewer photons to scatter from the electrons so we will decrease the chance that such a scattering takes place. The result is that we will now see events where the electrons strike the second screen but for which we see no flash indicating which hole they went through. We will now have three classes of events: those which are seen to pass through slit 1, those that are seen to pass through slit 2 and those where we do not know through which slit they passed. If we look at electrons in the first two categories we again see no interference, but those in the third category where there is no information about which slot they passed through will show an interference pattern. So we have done nothing to resolve our problem.

The problem might be that we have chosen the wrong approach. After all, we

know from Einstein that the photons will each have energy $E = h\nu$ so those photons that scatter will still change the motion of the electrons in the same way regardless of the intensity of the beam. For this reason we can consider another approach. We will keep the high intensity but decrease the frequency (increase the wavelength). If we now redo the experiment under these conditions, we will still see that for electrons that are measured to go through either of the slits we will have no interference pattern. However, as the wavelength increases as we reduce the energy of the photons, we will reach the point that the wavelength becomes greater than the separation between the two slits and we will no longer be able to resolve the positions of the electrons to determine which slit they passed through. We will then start to see the interference pattern again. This shows that there is no way that we can determine the path of the electron without affecting the motion of the electrons. Although we have framed this discussion in terms of photons scattering from electrons, any means of determining the slit through which an electron passes will have the same problems.

We then have a fundamental limitation on our ability to determine the trajectories of a given electron. We can no longer follow the trajectory of the electron through the system to the point where it strikes the second screen without changing that motion. One of the basic requirements that Heisenberg used in constructing his matrix mechanics was that only the aspects of the theory that described observable quantities are essential to the theory. Since we have now demonstrated that we cannot measure the trajectories of the electrons, the concept that quantum particles must follow definite trajectories, as in classical mechanics, need not be contained in quantum mechanical theories.

3.2   Wave Packets

To get a better understanding of how this comes about, lets return to the basic problem of the motion of classical waves. If we want to try to localize the energy carried by such a wave, we need to construct a *wave packet.* This is done by representing the wave function as a linear superposition of plane waves. In one dimension this can be written as

$$\Psi(x,t) = \int_{-\infty}^{\infty} dk \phi(k) e^{i(kx-\omega t)} , \tag{3.2}$$

where ω is some function of k. At $t = 0$ this gives

$$\Psi(x,0) = \int_{-\infty}^{\infty} dk \phi(k) e^{ikx} \tag{3.3}$$

so the initial wave packet is just the Fourier transform of $\phi(k)$.

For the purposes of our discussion, consider the case where $\phi(k)$ is a gaussian

$$\phi(k) = e^{-\frac{1}{2}\alpha^2(k-k_0)^2} \tag{3.4}$$

where α and k_0 are constants. The wave packet at $t = 0$ is then given by

$$\Psi(x, 0) = \int_{-\infty}^{\infty} dk e^{-\frac{1}{2}\alpha^2(k-k_0)^2} e^{ikx} . \tag{3.5}$$

We can now change the integration variable to $q = k - k_0$ so that

$$\Psi(x, 0) = \int_{-\infty}^{\infty} dq e^{-\frac{1}{2}\alpha^2 q^2} e^{i(q+k_0)x} = e^{ik_0 x} \int_{-\infty}^{\infty} dq e^{-\frac{1}{2}\alpha^2 q^2 + ixq} . \tag{3.6}$$

This integral can be found in any good set of integral tables. The result is then

$$\Psi(x, 0) = \frac{\sqrt{2\pi}}{\alpha} e^{ik_0 x} e^{-\frac{x^2}{2\alpha^2}} . \tag{3.7}$$

The intensity of the wave will be related to

$$|\Psi(x, 0)|^2 = \frac{2\pi}{\alpha^2} e^{-\frac{x^2}{\alpha^2}} \tag{3.8}$$

which is a gaussian.

Now we can define the average value of some power of x over this distribution as

$$\langle x^n \rangle = \frac{\int_{-\infty}^{\infty} dx \, x^n |\Psi(x, 0)|^2}{\int_{-\infty}^{\infty} dx |\Psi(x, 0)|^2} . \tag{3.9}$$

For the special case we are considering

$$\int_{-\infty}^{\infty} dx |\Psi(x, 0)|^2 = \frac{2\pi}{\alpha^2} \int_{-\infty}^{\infty} dx e^{-\frac{x^2}{\alpha^2}} = \frac{2\pi^{\frac{3}{2}}}{\alpha} , \tag{3.10}$$

$$\int_{-\infty}^{\infty} dx \, x |\Psi(x, 0)|^2 = \frac{2\pi}{\alpha^2} \int_{-\infty}^{\infty} dx \, x e^{-\frac{x^2}{\alpha^2}} = 0 , \tag{3.11}$$

since the integrand is odd and

$$\int_{-\infty}^{\infty} dx \, x^2 |\Psi(x, 0)|^2 = \frac{2\pi}{\alpha^2} \int_{-\infty}^{\infty} dx \, x^2 e^{-\frac{x^2}{\alpha^2}} = \pi^{\frac{3}{2}}\alpha . \tag{3.12}$$

Using these expressions yields

$$\langle x \rangle = 0 \tag{3.13}$$

and

$$\langle x^2 \rangle = \frac{\pi^{\frac{3}{2}}\alpha}{\frac{2\pi^{\frac{3}{2}}}{\alpha}} = \frac{\alpha^2}{2} . \tag{3.14}$$

If we now define the root mean square deviation as

$$\Delta x \equiv \sqrt{\langle |x - \langle x \rangle|^2 \rangle} = \sqrt{\langle x^2 \rangle - \langle x \rangle^2} , \tag{3.15}$$

43

for the special case we are considering here we get

$$\Delta x = \sqrt{\frac{\alpha^2}{2} - 0^2} = \frac{\alpha}{\sqrt{2}} \, . \tag{3.16}$$

We can now do something similar for the wave number k by defining

$$\langle k^n \rangle = \frac{\int_{-\infty}^{\infty} dk \, k^n |\phi(k)|^2}{\int_{-\infty}^{\infty} dk \, |\phi(k)|^2} \, . \tag{3.17}$$

For the special case being considered here

$$\int_{-\infty}^{\infty} dk \, |\phi(k)|^2 = \int_{-\infty}^{\infty} dk \, e^{-\alpha^2(k-k_0)^2} = \int_{-\infty}^{\infty} dq \, e^{-\alpha^2 q^2} = \frac{\sqrt{\pi}}{\alpha} \, , \tag{3.18}$$

$$\begin{aligned}
\int_{-\infty}^{\infty} dk \, k |\phi(k)|^2 &= \int_{-\infty}^{\infty} dk \, k e^{-\alpha^2(k-k_0)^2} = \int_{-\infty}^{\infty} dq \, (q+k_0) e^{-\alpha^2 q^2} \\
&= \int_{-\infty}^{\infty} dq \, q e^{-\alpha^2 q^2} + k_0 \int_{-\infty}^{\infty} dq \, e^{-\alpha^2 q^2} = k_0 \frac{\sqrt{\pi}}{\alpha}
\end{aligned} \tag{3.19}$$

and

$$\begin{aligned}
\int_{-\infty}^{\infty} dk \, k^2 |\phi(k)|^2 &= \int_{-\infty}^{\infty} dk \, k^2 e^{-\alpha^2(k-k_0)^2} = \int_{-\infty}^{\infty} dq \, (q+k_0)^2 e^{-\alpha^2 q^2} \\
&= \int_{-\infty}^{\infty} dq \, q^2 e^{-\alpha^2 q^2} + 2k_0 \int_{-\infty}^{\infty} dq \, q e^{-\alpha^2 q^2} + k_0^2 \int_{-\infty}^{\infty} dq \, e^{-\alpha^2 q^2} \\
&= \frac{\sqrt{\pi}}{2\alpha^3} + k_0^2 \frac{\sqrt{\pi}}{\alpha} \, .
\end{aligned} \tag{3.20}$$

So,

$$\langle k \rangle = \frac{k_0 \frac{\sqrt{\pi}}{\alpha}}{\frac{\sqrt{\pi}}{\alpha}} = k_0 \tag{3.21}$$

and

$$\langle k^2 \rangle = \frac{\frac{\sqrt{\pi}}{2\alpha^3} + k_0^2 \frac{\sqrt{\pi}}{\alpha}}{\frac{\sqrt{\pi}}{\alpha}} = \frac{1}{2\alpha^2} + k_0^2 \, . \tag{3.22}$$

The rms deviation Δk is then given by

$$\Delta k = \sqrt{\langle k^2 \rangle - \langle k \rangle^2} = \sqrt{\frac{1}{2\alpha^2} + k_0^2 - k_0^2} = \frac{1}{\sqrt{2}\alpha} \, . \tag{3.23}$$

We can now note that the product of the rms deviations in wave number and position is given by

$$\Delta k \Delta x = \frac{1}{\sqrt{2}\alpha} \frac{\alpha}{\sqrt{2}} = \frac{1}{2} \, . \tag{3.24}$$

Since this product has a fixed value, if we try to make the energy of the wave packet more localized by decreasing Δx, we must necessarily increase Δk, the width of the wave-number distribution, to compensate. If we try to narrow the range of values of k that contribute to the wave packet we will necessarily increase the width of the wave packet.

3.3 The Schrödinger Equation

To this point we have argued that we must have a wave function $\Psi(\boldsymbol{r}, t)$ that will describe a quantum mechanical particle. We can establish some requirements for the equation that produces the wave function from our previous discussion. The properties that are required are:

1. The wave function must be the solution of a linear differential equation. This is required so that the a general solution can be composed by a linear superposition of particular solutions of the equation, as is required if we are to obtain wave-like interference phenomena.

2. The wave equation must be consistent with the classical correspondence principle. That is, the wave equation should be constructed such that we can recover classical physics for systems of macroscopic size.

Now consider the expression for the wave packet in one dimension (3.2). We can introduce quantum mechanics into this expression by noting that the de Broglie wave length is

$$\lambda = \frac{h}{p} \tag{3.25}$$

and that the wave number is defined as

$$k = \frac{2\pi}{\lambda} = \frac{2\pi}{\frac{h}{p}} = \frac{2\pi p}{h} = \frac{p}{\hbar} \ . \tag{3.26}$$

The energy is given by

$$E = h\nu = h\frac{\omega}{2\pi} = \hbar\omega \ . \tag{3.27}$$

So,

$$\omega = \frac{E}{\hbar} \ . \tag{3.28}$$

Using the expressions for the wave number and angular frequency in (3.2) yields

$$\Psi(x, t) = \int_{-\infty}^{\infty} dp\, \phi(p) e^{\frac{i}{\hbar}(px - Et)} \ . \tag{3.29}$$

Now consider the spatial derivative of this wave packet

$$\frac{\partial}{\partial x}\Psi(x,t) = \frac{\partial}{\partial x}\int_{-\infty}^{\infty}dp\phi(p)e^{\frac{i}{\hbar}(px-Et)} = \frac{i}{\hbar}\int_{-\infty}^{\infty}dpp\phi(p)e^{\frac{i}{\hbar}(px-Et)}\,. \tag{3.30}$$

This can be rewritten as

$$\frac{\hbar}{i}\frac{\partial}{\partial x}\Psi(x,t) = \int_{-\infty}^{\infty}dpp\phi(p)e^{\frac{i}{\hbar}(px-Et)}\,. \tag{3.31}$$

Similarly,

$$\left(\frac{\hbar}{i}\frac{\partial}{\partial x}\right)^2\Psi(x,t) = \int_{-\infty}^{\infty}dpp^2\phi(p)e^{\frac{i}{\hbar}(px-Et)}\,. \tag{3.32}$$

The time derivative of the wave packet is

$$\frac{\partial}{\partial t}\Psi(x,t) = \frac{\partial}{\partial t}\int_{-\infty}^{\infty}dp\phi(p)e^{\frac{i}{\hbar}(px-Et)} = -\frac{i}{\hbar}\int_{-\infty}^{\infty}dpE\phi(p)e^{\frac{i}{\hbar}(px-Et)} \tag{3.33}$$

which can be rewritten as

$$i\hbar\frac{\partial}{\partial t}\Psi(x,t) = \int_{-\infty}^{\infty}dpE\phi(p)e^{\frac{i}{\hbar}(px-Et)}\,. \tag{3.34}$$

We now want to introduce the classical correspondence principle by considering the classical expression for the energy of a free particle

$$\frac{p^2}{2m} = E\,. \tag{3.35}$$

Multiplying both sides by $\phi(p)e^{\frac{i}{\hbar}(px-Et)}$ and integrating over all p yields

$$\frac{1}{2m}\int_{-\infty}^{\infty}dpp^2\phi(p)e^{\frac{i}{\hbar}(px-Et)} = \int_{-\infty}^{\infty}dpE\phi(p)e^{\frac{i}{\hbar}(px-Et)}\,, \tag{3.36}$$

or using (3.32) and (3.34)

$$\frac{1}{2m}\left(\frac{\hbar}{i}\frac{\partial}{\partial x}\right)^2\Psi(x,t) = i\hbar\frac{\partial}{\partial t}\Psi(x,t) \tag{3.37}$$

which can be simplified as

$$-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}\Psi(x,t) = i\hbar\frac{\partial}{\partial t}\Psi(x,t)\,. \tag{3.38}$$

This is the Schrödinger equation for a free particle in one dimension.

3.4 The Interpretation of the Wave Function

Our discussion of the two slit problem clearly indicates that it is meaningless to talk about a classical linear trajectory for quantum mechanical particles. In the two slit problem we introduced the wave function Ψ in order to describe the interference seen in the distribution of electrons at the second screen but noted that the electrons were still emitted and detected as particles. So what does the wave function mean? For the two slit diffraction problem cannot know which path each electron follows in reaching the second screen, but by measuring the distributions of electrons along the screen for a given number of electrons, we can determine the probability that an electron will strike a particular point on the screen. The wave function is therefore related to probability. Since we need to find $|\Psi|^2$ to give the interference pattern, we can now define the *probability density distribution* for a quantum mechanical particle in one dimension as

$$\rho(x,t) = |\Psi(x,t)|^2 \ . \tag{3.39}$$

If we are not to have particles appearing and disappearing at will, the particle must be somewhere in the interval $-\infty < x < \infty$ at any given time. That is, the probability for the particle to be in this interval must be unity at any given time. So,

$$\int_{-\infty}^{\infty} dx \rho(x,t) = \int_{-\infty}^{\infty} dx \, |\Psi(x,t)|^2 = 1 \, . \tag{3.40}$$

Now consider the one-dimensional Schrödinger equation

$$i\hbar \frac{\partial}{\partial t} \Psi(x,t) = -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} \Psi(x,t) \, . \tag{3.41}$$

The complex conjugate of this expression is

$$-i\hbar \frac{\partial}{\partial t} \Psi^*(x,t) = -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} \Psi^*(x,t) \tag{3.42}$$

where we have assumed that the potential is a real function. If we now multiply (3.41) on the left by $\Psi^*(x,t)$ and (3.42) on the right by $\Psi(x,t)$ and then subtract the two resulting equations, we get

$$i\hbar \left[\Psi^*(x,t)\frac{\partial}{\partial t}\Psi(x,t) + \left(\frac{\partial}{\partial t}\Psi^*(x,t)\right)\Psi(x,t) \right]$$
$$= -\frac{\hbar^2}{2m}\left[\Psi^*(x,t)\frac{\partial^2}{\partial x^2}\Psi(x,t) - \left(\frac{\partial^2}{\partial x^2}\Psi^*(x,t)\right)\Psi(x,t) \right] \, . \tag{3.43}$$

The left-hand side of this can be rewritten using

$$\Psi^*(x,t)\frac{\partial}{\partial t}\Psi(x,t) + \left(\frac{\partial}{\partial t}\Psi^*(x,t)\right)\Psi(x,t) = \frac{\partial}{\partial t}\left(\Psi^*(x,t)\Psi(x,t)\right) = \frac{\partial}{\partial t}\rho(x,t). \tag{3.44}$$

The right-hand side can be rewritten using

$$\Psi^*(x,t)\frac{\partial^2}{\partial x^2}\Psi(x,t) - \left(\frac{\partial^2}{\partial x^2}\Psi^*(x,t)\right)\Psi(x,t)$$

$$= \Psi^*(x,t)\frac{\partial^2}{\partial x^2}\Psi(x,t) + \left(\frac{\partial}{\partial x}\Psi^*(x,t)\right)\left(\frac{\partial}{\partial x}\Psi(x,t)\right)$$

$$- \left(\frac{\partial}{\partial x}\Psi^*(x,t)\right)\left(\frac{\partial}{\partial x}\Psi(x,t)\right) - \left(\frac{\partial^2}{\partial x^2}\Psi^*(x,t)\right)\Psi(x,t)$$

$$= \frac{\partial}{\partial x}\left[\Psi^*(x,t)\frac{\partial}{\partial x}\Psi(x,t) - \left(\frac{\partial}{\partial x}\Psi^*(x,t)\right)\Psi(x,t)\right]. \tag{3.45}$$

If we now define the probability current density distribution as

$$j(x,t) = \frac{\hbar}{2im}\left[\Psi^*(x,t)\frac{\partial}{\partial x}\Psi(x,t) - \left(\frac{\partial}{\partial x}\Psi^*(x,t)\right)\Psi(x,t)\right], \tag{3.46}$$

we obtain the continuity equation

$$\frac{\partial}{\partial t}\rho(x,t) + \frac{\partial}{\partial x}j(x,t) = 0. \tag{3.47}$$

This is of the same form as the continuity equation for charge conservation in electromagnetism. Here, this equation represents the conservation of probability. To see this, consider the integration of this equation over the interval $a \leq x \leq b$

$$\int_a^b dx\frac{\partial}{\partial t}\rho(x,t) + \int_a^b dx\frac{\partial}{\partial x}j(x,t) = 0. \tag{3.48}$$

The first term can be rewritten as

$$\int_a^b dx\frac{\partial}{\partial t}\rho(x,t) = \frac{d}{dt}\int_a^b dx\rho(x,t) \tag{3.49}$$

and is the time rate of change of the probability that the particle is in the interval. The second term is

$$\int_a^b dx\frac{\partial}{\partial x}j(x,t) = j(b,t) - j(a,t). \tag{3.50}$$

So,

$$\frac{\partial}{\partial t}\int_a^b dx\rho(x,t) = -j(b,t) + j(a,t). \tag{3.51}$$

The time rate of change of the probability for the particle to be in the interval $a \leq x \leq b$ is determined by the flow of probability out of the interval at b plus the flow of probability into the interval at a. This means that the probability is neither created or destroyed.

Note that in the limit $a \rightarrow -\infty$ and $b \rightarrow \infty$, the probability becomes unity and

$$\frac{\partial}{\partial t}1 = 0 = -j(\infty,t) + j(-\infty,t) \tag{3.52}$$

which is satisfied if the probability current vanishes at $\pm\infty$.

48

3.5 Coordinate Space and Momentum Space Representations

We now want to examine some aspects of the interpretation of the physical meaning of the wave functions that will help us to understand how we can generalize the Schrödinger equation to more three dimensions and to allow for the presence of more than one particle. To do this, we will first define the fourier transform of the wave function to *momentum space* for *coordinate space* as

$$\tilde{\Psi}(p,t) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} dx \, e^{-\frac{i}{\hbar}px} \Psi(x,t) \,. \tag{3.53}$$

The inverse of this fourier transform is

$$\Psi(x,t) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} dp \, e^{\frac{i}{\hbar}px} \tilde{\Psi}(p,t) \,. \tag{3.54}$$

The normalization of the momentum-space wave function can be calculated form this definition giving

$$
\begin{aligned}
\int_{-\infty}^{\infty} dp \, \tilde{\Psi}^*(p,t)\tilde{\Psi}(p,t) &= \frac{1}{2\pi\hbar} \int_{-\infty}^{\infty} dp \int_{-\infty}^{\infty} dx \, e^{\frac{i}{\hbar}px} \Psi^*(x,t) \int_{-\infty}^{\infty} dx' \, e^{-\frac{i}{\hbar}px'} \Psi(x',t) \\
&= \frac{1}{2\pi\hbar} \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dx' \, \Psi^*(x,t)\Psi(x',t) \int_{-\infty}^{\infty} dp \, e^{\frac{i}{\hbar}px} e^{-\frac{i}{\hbar}px'} \\
&= \frac{1}{2\pi\hbar} \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dx' \, \Psi^*(x,t)\Psi(x',t) \, 2\pi\delta\left(\frac{1}{\hbar}(x-x')\right) \\
&= \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dx' \, \Psi^*(x,t)\Psi(x',t) \, \delta(x-x') \\
&= \frac{1}{2\pi\hbar} \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dx' \, \Psi^*(x,t)\Psi(x',t) \, 2\pi\delta\left(\frac{1}{\hbar}(x-x')\right) \\
&= \int_{-\infty}^{\infty} dx \, \Psi^*(x,t)\Psi(x,t) = 1 \tag{3.55}
\end{aligned}
$$

In fact we have chosen the constants in the definition of the momentum-space or p-space wave function such that is normalized to unity. We can then define a *momentum density distribution* as

$$\tilde{\rho}(p,t) \equiv \tilde{\Psi}^*(p,t)\tilde{\Psi}(p,t) \,, \tag{3.56}$$

which represents the probability of finding a particle with momentum between p and $p+dp$ at time t.

The average position of the particle can be defined

$$\langle x \rangle = \int_{-\infty}^{\infty} dx \, x\rho(x,t) = \int_{-\infty}^{\infty} dx \, x\Psi^*(x,t)\Psi(x,t) = \int_{-\infty}^{\infty} dx \, \Psi^*(x,t)x\Psi(x,t) \tag{3.57}$$

49

and, similarly, the average momentum of the particle can be defined as

$$\langle p \rangle = \int_{-\infty}^{\infty} dp\, p\tilde{\rho}(p,t) = \int_{-\infty}^{\infty} dp\, p\tilde{\Psi}^*(p,t)\tilde{\Psi}(p,t) = \int_{-\infty}^{\infty} dp\, \tilde{\Psi}^*(p,t)p\tilde{\Psi}(p,t)\,. \quad (3.58)$$

Note that while the average position is given using the coordinate-space wave function, the average momentum is given by the momentum-space wave function. It is convenient to reexpress the average momentum in terms of the coordinate-space wave function. First note that from the definition of the momentum-space wave function we can write

$$
\begin{aligned}
p\tilde{\Psi}(p,t) &= p\frac{1}{\sqrt{2\pi\hbar}}\int_{-\infty}^{\infty} dx\, e^{-\frac{i}{\hbar}px}\Psi(x,t) = \frac{1}{\sqrt{2\pi\hbar}}\int_{-\infty}^{\infty} dx\, p e^{-\frac{i}{\hbar}px}\Psi(x,t)\\
&= \frac{1}{\sqrt{2\pi\hbar}}\int_{-\infty}^{\infty} dx\, \left(-\frac{\hbar}{i}\frac{\partial}{\partial x}e^{-\frac{i}{\hbar}px}\right)\Psi(x,t)\\
&= -\frac{\hbar}{i}e^{-\frac{i}{\hbar}px}\Psi(x,t)\Big|_{-\infty}^{\infty} - \frac{1}{\sqrt{2\pi\hbar}}\int_{-\infty}^{\infty} dx\, e^{-\frac{i}{\hbar}px}\left(-\frac{\hbar}{i}\frac{\partial}{\partial x}\Psi(x,t)\right)\\
&= \frac{1}{\sqrt{2\pi\hbar}}\int_{-\infty}^{\infty} dx\, e^{-\frac{i}{\hbar}px}\left(\frac{\hbar}{i}\frac{\partial}{\partial x}\Psi(x,t)\right)\,, \quad (3.59)
\end{aligned}
$$

where we have used the fact that the coordinate-space wave function must vanish at $\pm\infty$ to eliminate the surface term arising from the integration by parts. This can now be used to evaluate the average momentum as

$$
\begin{aligned}
\langle p \rangle &= \frac{1}{2\pi\hbar}\int_{-\infty}^{\infty} dp \int_{-\infty}^{\infty} dx\, e^{\frac{i}{\hbar}px}\Psi^*(x,t)\int_{-\infty}^{\infty} dx'\, e^{-\frac{i}{\hbar}px'}\left(\frac{\hbar}{i}\frac{\partial}{\partial x'}\Psi(x',t)\right)\\
&= \frac{1}{2\pi\hbar}\int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dx'\, \Psi^*(x,t)\left(\frac{\hbar}{i}\frac{\partial}{\partial x'}\Psi(x',t)\right)\int_{-\infty}^{\infty} dp\, e^{\frac{i}{\hbar}px}e^{-\frac{i}{\hbar}px'}\\
&= \int_{-\infty}^{\infty} dx\, \Psi^*(x,t)\frac{\hbar}{i}\frac{\partial}{\partial x}\Psi(x,t) = \int_{-\infty}^{\infty} dx\, \Psi^*(x,t)\hat{p}\Psi(x,t)\,, \quad (3.60)
\end{aligned}
$$

where

$$\hat{p} \equiv \frac{\hbar}{i}\frac{\partial}{\partial x} \quad (3.61)$$

is the *coordinate-space momentum operator*. From (3.57) we can also identify the *coordinate-space position operator* as

$$\hat{x} = x\,. \quad (3.62)$$

We could also decide that we want to work with the momentum-space wave function. From (3.58) we can identify the *momentum-space momentum operator* as

$$\hat{\tilde{p}} = p \quad (3.63)$$

50

and, using the inverse fourier transform we can identify the *momentum-space coordinate operator* as

$$\hat{\hat{x}} = i\hbar \frac{\partial}{\partial p} \, . \tag{3.64}$$

Note that the coordinate-space and momentum-space representations are completely equivalent and the choice of representation is usually based on calculational convenience. We will revisit this problem after we have considered solutions to the one-dimensional Schrödinger equation. In either representation, the operators consist of complex numbers or *differential operators*.

3.6 Differential Operators

We can generalize this be writing that some operator $\hat{A}$ acts on the wave function to give

$$\hat{A} f(x) = g(x) \, , \tag{3.65}$$

where g is some other function in the vector space. We will use the "hat" over a symbol to indicate that it represents a differential operator. Differential operators have the properties:

1. The operator $\hat{A}$ is *linear* if

$$\hat{A} \left(c f(x) \right) = c \left(\hat{A} f(x) \right) \tag{3.66}$$

   and

$$\hat{A} \left(f(x) + g(x) \right) = \hat{A} f(x) + \hat{A} g(x) \, , \tag{3.67}$$

   where c is a complex constant.

2.
$$\left(\hat{A} + \hat{B} \right) f(x) = \hat{A} f(x) + \hat{B} f(x) \, . \tag{3.68}$$

3.
$$\left(\hat{A} \hat{B} \right) f(x) = \hat{A} \left(\hat{B} f(x) \right) \tag{3.69}$$

4. In general differential operators do not commute. That is,

$$\left[\hat{A}, \hat{B} \right] \neq 0 \, . \tag{3.70}$$

   Note that the commutator may yield either a complex constant or another differential operator.

5. There is an identity operator such that

$$\hat{1}f(x) = f(x) \tag{3.71}$$

and

$$\hat{1}\hat{A} = \hat{A}\hat{1} = \hat{A}. \tag{3.72}$$

6. The inverse of an operator $\hat{A}$ is denoted by $\hat{A}^{-1}$ and has the property that

$$\hat{A}^{-1}\hat{A} = \hat{1}. \tag{3.73}$$

Not all differential operators have an inverse.

7. The hermitian conjugate of an operator is defined such that

$$\int dx\, f^*(x)\hat{A}^\dagger g(x) = \int dx\, \left(\hat{A}f(x)\right)^* g(x). \tag{3.74}$$

8. An operator is hermitian if

$$\int dx\, f^*(x)\hat{A}^\dagger g(x) = \int dx\, f^*(x)\hat{A}g(x). \tag{3.75}$$

9. An operator is antihermitian if

$$\int dx\, f^*(x)\hat{A}^\dagger g(x) = -\int dx\, f^*(x)\hat{A}g(x). \tag{3.76}$$

10. An operator is unitary if

$$\hat{A}^\dagger = \hat{A}^{-1}. \tag{3.77}$$

It is always necessary to remember that the differential operators will always appear in conjunction with some function. To see what this means consider the product of the coordinate-space moment and position operators $\hat{p}\hat{x}$. To determine the properties of this composite operator we must consider

$$
\begin{aligned}
(\hat{p}\hat{x})\,\Psi(x,t) &= \hat{p}\,(\hat{x}\Psi(x,t)) = \frac{\hbar}{i}\frac{\partial}{\partial x}\left(x\Psi(x,t)\right) = \frac{\hbar}{i}\Psi(x,t) + x\frac{\hbar}{i}\frac{\partial}{\partial x}\Psi(x,t) \\
&= \frac{\hbar}{i}\Psi(x,t) + \hat{x}\hat{p}\Psi(x,t) \tag{3.78}
\end{aligned}
$$

or

$$\hat{p}\hat{x}\Psi(x,t) - \hat{x}\hat{p}\Psi(x,t) = -i\hbar\Psi(x,t). \tag{3.79}$$

From this we can determine that the commutator of the momentum and position operators is given by

$$[\hat{p}, \hat{x}] = -i\hbar. \tag{3.80}$$

52

This and the commutation relations

$$[\hat{x}, \hat{x}] = [\hat{p}, \hat{p}] = 0 \tag{3.81}$$

are called the *canonical commutation relations.*

A function of an operator is defined in terms of a power series. That is

$$f(\hat{A}) = \sum_{n=0}^{\infty} c_n \hat{A}^n \tag{3.82}$$

where the expansion coefficients are the same as the Taylor series for the same function of a scalar. That is

$$f(z) = \sum_{n=0}^{\infty} c_n z^n \,. \tag{3.83}$$

3.7   The Heisenberg Uncertainty Relation

In Section 3.2 we obtained a relationship between the rms width of the packet and the rms deviation in momentum of the packet for a gaussian wave packet. We can now obtain a similar, but general, result for the Schrödinger equation. This is the famous Heisenberg uncertainty relation.

To do this, we must first derive a general mathematical inequality called the Schwartz inequality. Since the integral of any positive quantity must be positive, we start with

$$\int dx \, |f(x) - \lambda g(x)|^2 \geq 0 \,, \tag{3.84}$$

where λ is any complex number. Expanding the integrand gives

$$\int dx \, |f(x)|^2 + |\lambda|^2 \int dx \, |g(x)|^2 - \lambda^* \int dx g^*(x) f(x) - \lambda \int dx f^*(x) g(x) \geq 0 \,. \tag{3.85}$$

If we now choose

$$\lambda = \frac{\int dx g^*(x) f(x)}{\int dx \, |g(x)|^2} \,, \tag{3.86}$$

this reduces to

$$\int dx \, |f(x)|^2 - \frac{\left| \int dx g^*(x) f(x) \right|^2}{\int dx \, |g(x)|^2} \geq 0 \,. \tag{3.87}$$

This can be rearranged to give

$$\int dx \, |f(x)|^2 \int dx \, |g(x)|^2 \geq \left| \int dx g^*(x) f(x) \right|^2 \,. \tag{3.88}$$

This is the *Schwartz inequality.*

We can now define the average position of a wave function as

$$\langle x \rangle = \int dx \Psi^*(x,t) \hat{x} \Psi(x,t) \tag{3.89}$$

and the average momentum as

$$\langle p \rangle = \int dx \Psi^*(x,t) \hat{p} \Psi(x,t) . \tag{3.90}$$

The mean square deviation in position is defined as

$$
\begin{aligned}
(\Delta x)^2 &= \int dx \Psi^*(x,t) \left| \hat{x} - \langle x \rangle \right|^2 \Psi(x,t) \\
&= \int dx \left[(\hat{x} - \langle x \rangle) \Psi(x,t) \right]^* \left[(\hat{x} - \langle x \rangle) \Psi(x,t) \right] \\
&\equiv \int dx \left| f(x) \right|^2 ,
\end{aligned} \tag{3.91}
$$

where to obtain the second line we have used the fact that $\hat{x}$ is hermitian. Similarly the mean square deviation in the momentum of the wave function is

$$
\begin{aligned}
(\Delta p)^2 &= \int dx \Psi^*(x,t) \left| \hat{p} - \langle p \rangle \right|^2 \Psi(x,t) \\
&= \int dx \left[(\hat{p} - \langle p \rangle) \Psi(x,t) \right]^* \left[(\hat{p} - \langle p \rangle) \Psi(x,t) \right] \\
&\equiv \int dx \left| g(x) \right|^2 .
\end{aligned} \tag{3.92}
$$

We can now use the Schwartz inequality (3.88) to write

$$
\begin{aligned}
(\Delta x)^2 (\Delta p)^2 &\geq \left| \int dx \left[(\hat{p} - \langle p \rangle) \Psi(x,t) \right]^* \left[(\hat{x} - \langle x \rangle) \Psi(x,t) \right] \right|^2 \\
&= \frac{1}{2} \left[\left| \int dx \left[(\hat{p} - \langle p \rangle) \Psi(x,t) \right]^* \left[(\hat{x} - \langle x \rangle) \Psi(x,t) \right] \right|^2 \right. \\
&\quad + \left. \left| \int dx \left[(\hat{x} - \langle x \rangle) \Psi(x,t) \right]^* \left[(\hat{p} - \langle p \rangle) \Psi(x,t) \right] \right|^2 \right] \\
&= \frac{1}{2} \left[\left| \int dx \Psi^*(x,t) (\hat{p} - \langle p \rangle)(\hat{x} - \langle x \rangle) \Psi(x,t) \right|^2 \right. \\
&\quad + \left. \left| \int dx \Psi^*(x,t) (\hat{x} - \langle x \rangle)(\hat{p} - \langle p \rangle) \Psi(x,t) \right|^2 \right]
\end{aligned} \tag{3.93}
$$

54

We can now use the inequality

$$|a|^2 + |b|^2 = \frac{1}{2}\left[|a-b|^2 + |a+b|^2\right] \geq \frac{1}{2}|a-b|^2 \tag{3.94}$$

to rewrite this as

$$(\Delta x)^2 (\Delta p)^2 \geq \frac{1}{4}\left|\int dx\, \Psi^*(x,t)\left[(\hat{p}-\langle p\rangle)(\hat{x}-\langle x\rangle) - (\hat{x}-\langle x\rangle)(\hat{p}-\langle p\rangle)\right]\Psi(x,t)\right|^2$$

$$= \frac{1}{4}\left|\int dx\, \Psi^*(x,t)\left[\hat{p},\hat{x}\right]\Psi(x,t)\right|^2. \tag{3.95}$$

We can now use the canonical commutation relation (3.80) to write this as

$$(\Delta x)^2 (\Delta p)^2 \geq \frac{1}{4}\left|\int dx\, \Psi^*(x,t)(-i\hbar)\Psi(x,t)\right|^2 = \frac{\hbar^2}{4}\left|\int dx\, \Psi^*(x,t)\Psi(x,t)\right|^2 = \frac{\hbar^2}{4}. \tag{3.96}$$

This implies that

$$\Delta x \Delta p \geq \frac{\hbar}{2} \tag{3.97}$$

which is the *Heisenberg uncertainty relation.*

We can see the consequences of this for the Bohr atom. We can examine what will happen if we confine a electron to within a distance comparable to that of the Bohr radius a_0. That is, let

$$\Delta x = a_0. \tag{3.98}$$

The uncertainty in the momentum of the electron is then

$$\Delta p \sim \frac{\hbar}{a_0}. \tag{3.99}$$

We can also find the momentum of the electron in the first Bohr orbit of hydrogen using (1.17) with $n = 1$ and $r = a_0$. This gives

$$p = \frac{\hbar}{a_0}. \tag{3.100}$$

The uncertainty in the energy of the electron due to the uncertainty principle is

$$\Delta E = \Delta\left(\frac{p^2}{2m_e}\right) = \frac{p\Delta p}{m_e} \sim \frac{\hbar^2}{m_e a_0^2}. \tag{3.101}$$

This gives

$$\Delta E \sim \frac{(1.05 \times 10^{-34}\,\mathrm{J\,s})^2}{9.11 \times 10^{-31}\mathrm{kg}(5.29 \times 10^{-11}\mathrm{kg})^2}\frac{1\mathrm{eV}}{1.60 \times 10^{-19}\mathrm{J}} \sim 27\mathrm{eV}. \tag{3.102}$$

55

The binding energy of an electron in the first Bohr orbit is only 13.5 eV. So the uncertainty in the energy for the Bohr atom would be enough to completely free the electron from the hydrogen atom. So the Bohr atom cannot be a satisfactory description of the atom.

Now consider a particle with a mass $m = 1\text{mg} = 1 \times 10^{-6}\text{kg}$.

$$\Delta p \Delta x \sim \hbar \,, \tag{3.103}$$

then

$$\Delta v \Delta x \sim \frac{\hbar}{m} = \frac{1.05 \times 10^{-34}\,\text{J s}}{1 \times 10^{-6}\text{kg}} \cong 1 \times 10^{-28}\text{m}^2/\text{s} \,. \tag{3.104}$$

So even if we could measure the position of the particle to $\Delta x \sim 10^{-10}$m, the uncertainty in speed would be $\Delta v \sim 1)^{-10}$m/s $\sim 3 \times 10^{-9}$m/century! This clearly implies that for macroscopic objects the uncertainty principle has no practical effect.

3.8   Review of Classical Mechanics

To see how we can generalize the Schrödinger equation, it is useful to review the lagrangian and hamiltonian formulations of classical mechanics. These formulations of classical mechanics are equivalent to Newton's laws, by are based on Hamilton's principle, or the principle of least action. For simplicity, we will first consider the case of motion in one dimension. The action is defined as

$$I(t_1, t_2) = \int_{t_1}^{t_2} dt L(x(t), \dot{x}(t), t) \,, \tag{3.105}$$

where $L(x, \dot{x}, t)$ is the lagrangian. For a classical particle with mass m moving in a potential the lagrangian is

$$L(x, \dot{x}, t) = T - V = \frac{1}{2}m\dot{x}^2 - V(x) \,. \tag{3.106}$$

Hamilton's principle states that the classical trajectory between to fixed points $x(t_1)$ and $x(t_2)$ will be the one that minimizes the action. Using the calculus of variations, it can be shown that the equation of motion of the particle is given by

$$\frac{d}{dt}\frac{\partial}{\partial \dot{x}}L(x, \dot{x}, t) - \frac{\partial}{\partial x}L(x, \dot{x}, t) = 0 \,. \tag{3.107}$$

The generalized momentum of the particle is defined as

$$p \equiv \frac{\partial}{\partial \dot{x}}L(x, \dot{x}, t) \,. \tag{3.108}$$

The lagrangian is defined as a function of the position $x(t)$ and the velocity $\dot{x}(t)$. Alternately, since the generalized momentum is a function of these variables, we can

obtain a formulation of Hamilton's principle in terms of the position $x(t)$ and the generalized momentum $p(t)$. This is done by using a Legendre transform to define the hamiltonian

$$H(p, x, t) = p\dot{x} - L(x, \dot{x}, t). \tag{3.109}$$

For a particle moving in one dimension in the presence of a potential the hamiltonian is

$$H(p, x) = \frac{p^2}{2m} + V(x). \tag{3.110}$$

Hamilton's principle then gives the equations of motion

$$\dot{x} = \frac{\partial H}{\partial p} \tag{3.111}$$

and

$$\dot{p} = -\frac{\partial H}{\partial x}. \tag{3.112}$$

The generalization of hamiltonian to describe a particle moving in three dimensions in the presence of a potential is

$$H(\boldsymbol{p}, \boldsymbol{r}) = \frac{\boldsymbol{p}^2}{2m} + V(\boldsymbol{r}) \tag{3.113}$$

and the equations of motion are

$$\dot{r}_i = \frac{\partial H}{\partial p_i} \tag{3.114}$$

and

$$\dot{p}_i = -\frac{\partial H}{\partial r_i}. \tag{3.115}$$

Similarly, the hamiltonian can be extended to systems of more than one particle by simply adding kinetic energy terms for each of the additional degree of freedom and by constructing a potential term describing all of the conservative forces acting on the system.

3.9 Generalization of the Schrödinger Equation

We can now reexpress the one-dimensional Schrödinger equation for a free particle (3.38) in terms of differential operators as

$$\frac{\hat{p}^2}{2m} \Psi(x, t) = \hat{E} \Psi(x, t), \tag{3.116}$$

where we define the energy operator as

$$\hat{E} = i\hbar \frac{\partial}{\partial t}. \tag{3.117}$$

Note that from (3.110), the hamiltonian for a free particle moving in one dimension is

$$H(p, x) = \frac{p^2}{2m}. \tag{3.118}$$

Comparing this with (3.116) suggests a method of extending the Schrödinger equation to more complicated situations. The method is to start with the hamiltonian $H(p, x)$ for a classical system analogous to the quantum system that we wish to describe. The *hamiltonian operator* is then obtained by replacing the classical momentum and position variables with the quantum mechanical momentum and position operators. That is

$$\hat{H} = H(\hat{p}, \hat{x}), \tag{3.119}$$

where the momentum and position operators satisfy the *canonical commutation relation*

$$[\hat{p}, \hat{x}] = i\hbar. \tag{3.120}$$

The Schrödinger equation is then written as

$$\hat{H}\Psi(x, t) = \hat{E}\Psi(x, t). \tag{3.121}$$

This method is called the canonical quantization procedure.

It is now simple to generalize the Schrödinger equation for a one-dimensional particle in a potential by defining the hamiltonian operator as

$$\hat{H} = \frac{\hat{p}^2}{2m} + V(\hat{x}). \tag{3.122}$$

In coordinate space this becomes

$$\hat{H} = -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + V(x). \tag{3.123}$$

The Schrödinger equation then becomes

$$-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}\Psi(x, t) + V(x)\Psi(x, t) = i\hbar\frac{\partial}{\partial t}\Psi(x, t). \tag{3.124}$$

In three dimensions we can define the coordinate-space momentum operator as

$$\hat{\boldsymbol{p}} = \frac{\hbar}{i}\boldsymbol{\nabla}, \tag{3.125}$$

where

$$\boldsymbol{\nabla} = \boldsymbol{e}_x\frac{\partial}{\partial x} + \boldsymbol{e}_y\frac{\partial}{\partial y} + \boldsymbol{e}_z\frac{\partial}{\partial z} \tag{3.126}$$

and $\boldsymbol{e}_x$, $\boldsymbol{e}_y$ and $\boldsymbol{e}_z$ are the usual cartesian basis vectors. The position operator is

$$\hat{\boldsymbol{r}} = x\boldsymbol{e}_x + y\boldsymbol{e}_y + z\boldsymbol{e}_z. \tag{3.127}$$

The hamiltonian operator for a particle moving in three-dimensions in a potential is then

$$\hat{H} = \frac{\hat{\boldsymbol{p}}^2}{2m} + V(\boldsymbol{r}) = -\frac{\hbar^2}{2m}\boldsymbol{\nabla}^2 + V(\boldsymbol{r}). \tag{3.128}$$

In this case the canonical commutation relations are

$$[\hat{p}_i, \hat{x}_j] = -i\hbar\delta_{ij} \tag{3.129}$$
$$[\hat{x}_i, \hat{x}_j] = 0 \tag{3.130}$$
$$[\hat{p}_i, \hat{p}_j] = 0 \tag{3.131}$$

for $i = 1, 2, 3$.

Chapter 4

The Time-Independent Wave Function

Using (3.128), the three-dimensional Schrödinger equation is given by

$$-\frac{\hbar^2}{2m}\nabla^2\Psi(\boldsymbol{r},t) + V(\boldsymbol{r})\Psi(\boldsymbol{r},t) = i\hbar\frac{\partial}{\partial t}\Psi(\boldsymbol{r},t)\,. \tag{4.1}$$

The first step in finding general solutions to partial differential equations is to apply the method of separation of variables. In the case of the (4.1) the wave functions are functions of the variables $\boldsymbol{r}$ and t. Assume that the wave function can be written as a product of spatial and time functions as

$$\Psi(\boldsymbol{r},t) = \psi(\boldsymbol{r})T(t)\,. \tag{4.2}$$

Substituting this wave function into (4.1) yields

$$-\frac{\hbar^2}{2m}T(t)\nabla^2\psi(\boldsymbol{r}) + V(\boldsymbol{r})\psi(\boldsymbol{r})T(t) = i\psi(\boldsymbol{r})\hbar\frac{\partial}{\partial t}T(t)\,. \tag{4.3}$$

Dividing this equation by $\psi(\boldsymbol{r})T(t)$ gives

$$\frac{1}{\psi(\boldsymbol{r})}\left[-\frac{\hbar^2}{2m}\nabla^2 + V(x)\psi(\boldsymbol{r})\right] = i\hbar\frac{1}{T(t)}\frac{\partial T(t)}{\partial t}\,. \tag{4.4}$$

Since this must be true for all values of $\boldsymbol{r}$ and t, each side must be equal to the same constant. The right-hand side of the equation is then

$$i\hbar\frac{1}{T(t)}\frac{dT(t)}{dt} = E\,, \tag{4.5}$$

or

$$i\hbar\frac{dT(t)}{dt} = ET(t)\,. \tag{4.6}$$

This differential equation is easily integrated to give

$$T(t) = T(0)e^{-\frac{i}{\hbar}Et}. \tag{4.7}$$

Since the wave functions must ultimately be normalized, the complex constant $T(0)$ can be subsumed into the normalization constant. We can then assume that

$$T(0) \to 1. \tag{4.8}$$

The left-hand side of the separated equation is

$$-\frac{\hbar^2}{2m}\nabla^2\psi(\boldsymbol{r}) + V(x)\psi(\boldsymbol{r}) = E\psi(\boldsymbol{r}), \tag{4.9}$$

or

$$\hat{H}\psi(x) = E\psi(x). \tag{4.10}$$

This is the time-independent Schrödinger equation. The solutions to this equation will be *stationary states.*

The normalization of the probability requires that

$$1 = \int d^3r\Psi^*(\boldsymbol{r},t)\Psi(\boldsymbol{r},t). \tag{4.11}$$

That is, the wave function must be *square integrable.* For this to be true, of course, the integral of the absolute square of the wave function must be finite. For this to occur, the wave function must be localized, which means that the wave function must vanish at infinity. That is

$$\lim_{|\boldsymbol{r}|\to\infty}\psi(\boldsymbol{r}) = 0. \tag{4.12}$$

This requirement is a boundary condition on the solution of the time-independent Schrödinger equation. In addition, the normalization condition implies that the wave function must be bounded everywhere in space, or that it must have only integrable singularities anywhere in space.

The time-independent Schrödinger equation along with the boundary conditions constitutes an eigenvalue problem and solutions to the equation will exist only for special values of E. These are the the eigenvalues or *eigenenergies* for the Schrödinger equation. If we label all of the eigenenergies by some integer n, then the wave equation will be for each of the eigenenergies will be

$$\hat{H}\psi_n(\boldsymbol{r}) = E_n\psi_n(\boldsymbol{r}), \tag{4.13}$$

where the $\psi_n(\boldsymbol{r})$ are the eigenfunctions or *eigenstates* of the equation.

Note that as long as the potential $V(\boldsymbol{r})$ is a real function, the Hamiltonian operator is hermitian. Consider the *matrix element* of the hamiltonian operator

$$\int d^3r\psi_m^*(\boldsymbol{r})\hat{H}\psi_n(\boldsymbol{r}) = E_n\int d^3r\psi_m^*(\boldsymbol{r})\psi_n(\boldsymbol{r}), \tag{4.14}$$

where we have used the fact that ψ_n is an eigenstate of the hamiltonian. Since $\hat{H}$ is hermitian, we can also write that

$$\int d^3 r \psi_m^*(\boldsymbol{r}) \hat{H} \psi_n(\boldsymbol{r}) = \int d^3 r \left(\hat{H} \psi_m(\boldsymbol{r}) \right)^* \psi_n(\boldsymbol{r}) = E_m^* \int d^3 r \psi_m^*(\boldsymbol{r}) \psi_n(\boldsymbol{r}) . \quad (4.15)$$

Subtracting these two expression gives

$$0 = (E_m^* - E_n) \int d^3 r \psi_m^*(\boldsymbol{r}) \psi_n(\boldsymbol{r}) . \quad (4.16)$$

In the case where $m = n$ we have

$$\int d^3 r \, |\psi_n(\boldsymbol{r})|^2 > 0 . \quad (4.17)$$

So,

$$E_n^* - E_n = 0 \quad (4.18)$$

which means that the eigenenergies are real. For states where $E_m \neq E_n$, (4.16) requires that

$$\int d^3 r \psi_m^*(\boldsymbol{r}) \psi_n(\boldsymbol{r}) = 0 . \quad (4.19)$$

So the states are orthogonal. The final possibility is that there may be more than one eigenstate having a given eigenvalue. Then (4.16) does not require that the states are orthogonal. However, these states will always be orthogonal to the other states with different eigenvalues and can be made mutually orthogonal by using an orthogonalization procedure. For example, suppose that we have two different eigenstates $\psi_1(\boldsymbol{r})$ and $\psi_2(\boldsymbol{r})$ where the corresponding eigenstates are the same. That is, $E_1 = E_2$. We can find a new set of states that are mutually orthogonal be choosing

$$\psi_1'(\boldsymbol{r}) = \psi_1(\boldsymbol{r}) \quad (4.20)$$

and

$$\psi_2'(\boldsymbol{r}) = \psi_2(\boldsymbol{r}) - \frac{\int d^3 r' \psi_1^*(\boldsymbol{r}') \psi_2(\boldsymbol{r}')}{\int d^3 r' \, |\psi_1(\boldsymbol{r}')|^2} \psi_1(\boldsymbol{r}) . \quad (4.21)$$

By construction, we then have that

$$\int d^3 r \psi_1'^*(\boldsymbol{r}) \psi_2'(\boldsymbol{r}) = 0 . \quad (4.22)$$

We can choose to normalized all of the orthogonal states so we can then write that

$$\int d^3 r \psi_m^*(\boldsymbol{r}) \psi_n(\boldsymbol{r}) = \delta_{mn} . \quad (4.23)$$

So all of the states are orthonormal.

Since we now have a set of eigenstates that span three-dimensional space, we can expand any complex function defined on that space as

$$f(\boldsymbol{r}) = \sum_n c_n \psi_n(\boldsymbol{r}) \,. \tag{4.24}$$

This in combination with the orthonormality of the states gives the completeness relation

$$\sum_n \psi_n(\boldsymbol{r}) \psi_n^*(\boldsymbol{r}') = \delta(\boldsymbol{r} - \boldsymbol{r}') = \delta(x - x')\delta(y - y')\delta(z - z') \,. \tag{4.25}$$

Using (4.7), we can now obtain an infinite set of solutions to the time-dependent Schrödinger equation as

$$\Psi_n(\boldsymbol{r}, t) = \psi_n(\boldsymbol{r}) e^{-\frac{i}{\hbar} E_n t} \,. \tag{4.26}$$

These are the *stationary state* solutions of the Schrödinger equation. These states are also orthonormal since

$$
\begin{aligned}
\int d^3 r\, \Psi_m^*(\boldsymbol{r}, t) \Psi_n(\boldsymbol{r}, t) &= \int d^3 r\, \psi_m^*(\boldsymbol{r}) e^{\frac{i}{\hbar} E_m t} \psi_n(\boldsymbol{r}) e^{-\frac{i}{\hbar} E_n t} \\
&= e^{\frac{i}{\hbar}(E_m - E_n) t} \int d^3 r\, \psi_m^*(\boldsymbol{r}) \psi_n(\boldsymbol{r}) \\
&= e^{\frac{i}{\hbar}(E_m - E_n) t} \delta_{mn} = \delta_{mn} \,.
\end{aligned}
\tag{4.27}
$$

Similarly, the *expectation value* of the hamiltonian for any of the stationary states is

$$\langle H \rangle_n \equiv \int d^3 r\, \Psi_n^*(\boldsymbol{r}, t) \hat{H} \Psi_n(\boldsymbol{r}, t) = E_n \,. \tag{4.28}$$

The mean-square deviation of the energy is given by

$$\left\langle \Delta H^2 \right\rangle_n = \left\langle (H - \langle H \rangle_n)^2 \right\rangle_n = \left\langle H^2 \right\rangle_n - 2 \langle H \rangle_n^2 + \langle H \rangle_n^2 = E_n^2 - 2E_n^2 + E_n^2 = 0 \,. \tag{4.29}$$

This means that the energy has a definite value that is independent of time. Note that the same situation will occur for any operator for which the $\psi_n(\boldsymbol{r})$ are eigenvectors.

Any general solution of the time-dependent Schrödinger equation can be written as a linear combination of the stationary state wave functions. That is,

$$\Psi(\boldsymbol{r}, t) = \sum_n a_n \Psi_n(\boldsymbol{r}, t) \,. \tag{4.30}$$

The requirement that the probability be normalized gives

$$
\begin{aligned}
1 = \int d^3 r\, \Psi^*(\boldsymbol{r}, t) \Psi(\boldsymbol{r}, t) &= \sum_m \sum_n a_m^* a_n \int d^3 r\, \Psi_m^*(\boldsymbol{r}, t) \Psi_n(\boldsymbol{r}, t) \\
&= \sum_m \sum_n a_m^* a_n \delta_{mn} = \sum_n |a_n|^2 \,.
\end{aligned}
\tag{4.31}
$$

Since the Schrödinger equation is first order in time, a particular solution requires that the wave function has to be specified as at some time, which is usually $t = 0$. This condition then can be written as

$$\Psi(\boldsymbol{r}, 0) = f(\boldsymbol{r}) = \sum_n a_n \Psi_n(\boldsymbol{r}, 0) = \sum_n a_n \psi_n(\boldsymbol{r}) \,. \tag{4.32}$$

Multiplying this expression from the left by $\psi_m^*(\boldsymbol{r})$ and then integrating over all space gives.

$$\int d^3 r \psi_m^*(\boldsymbol{r}) f(\boldsymbol{r}) = \sum_n a_n \int d^3 r \psi_m^*(\boldsymbol{r}) \psi_n(\boldsymbol{r}) = \sum_n a_n \delta_{mn} = a_n \tag{4.33}$$

The coefficients in the expansion of the time-dependent wave function can therefore be determined by knowing the wave function at $t = 0$.

The general approach to obtaining solutions to the Schrödinger equation is to:

1. Solve the time-independent Schrödinger equation (4.13) for the eigenenergies E_n and the corresponding eigenstates ψ_n.

2. Construct the stationary state solutions using (4.26).

3. Impose the initial condition (4.32) and use (4.33) to find the coefficients of the expansion of the general solution (4.30).

Chapter 5

Solutions to the One-Dimensional Schrödinger Equation

Using (3.123) we can write the one-dimensional time-independent Schrodinger equation as

$$-\frac{\hbar^2}{2m}\frac{d^2}{dx^2}\psi(x) + V(x)\psi(x) = E\psi(x) \tag{5.1}$$

subject to the boundary conditions

$$\lim_{x\to\pm\infty}\psi(x) = 0 \,. \tag{5.2}$$

Before studying solutions for particular choices of $V(x)$, it is useful to study the general properties of the of the one-dimensional time-independent wave functions. To do this, recall that the Heisenberg uncertainty relation (3.97) is

$$\Delta x \Delta p \geq \frac{\hbar}{2} \,. \tag{5.3}$$

This means that $\hbar$ sets the scale for the applicability of quantum mechanics. It is the fact that $\hbar$ is very small that tells us that quantum mechanics is important only for very small objects at very short distances. In fact, if $\hbar$ were zero we wouldn't have to worry about quantum mechanics.

In general, we can rewrite the time-independent wave function for the one-dimensional Schrödinger equation as

$$\psi(x) = e^{\frac{i}{\hbar}\phi(x)} \,, \tag{5.4}$$

where $\phi(x)$ is a position-dependent phase function. Substituting this into (5.1) gives

$$-\frac{i\hbar}{2m}\phi''(x) + \frac{1}{2m}\phi'^2(x) + V(x) = E \,. \tag{5.5}$$

We want to now consider an approximation. We will assume that the potential $V(x)$ changes slowly compared to the local wavelength of the wave function. Recall that the de Broglie wavelength is defined as

$$\lambda = \frac{2\pi\hbar}{p} \,, \tag{5.6}$$

where p is the momentum of the particle. The wavelength will be small if either p is large or $\hbar$ is small. We can then obtain the small wavelength approximation by expanding in powers of $\hbar$ around the point $\hbar = 0$. That is, we can expand the phase function as

$$\phi(x) = \phi_0(x) + \hbar\phi_1(x) + \ldots \tag{5.7}$$

Substituting this into (5.5) and keeping only terms up to order $\hbar$ gives

$$-\frac{i\hbar}{2m}\phi_0''(x) + \frac{1}{2m}\left(\phi_0'^2(x) + 2\hbar\phi_0'(x)\phi_1'(x)\right) + V(x) = E \,. \tag{5.8}$$

Equating powers of $\hbar$ yields

$$\frac{1}{2m}\phi_0'^2(x) + V(x) = E \tag{5.9}$$

and

$$-\frac{i}{2m}\phi_0''(x) + \frac{1}{2m}2\phi_0'(x)\phi_1'(x) = 0 \,. \tag{5.10}$$

The first of these can be solved as

$$\phi_0'(x) = \pm\sqrt{2m(E - V(x))} \equiv \pm k_0(x) \,, \tag{5.11}$$

which can be integrated to give

$$\phi_0(x) = \pm \int dx\, k_0(x) \,. \tag{5.12}$$

The first order correction to the phase function is given by the second equation and is

$$\phi_1'(x) = \frac{i}{2}\frac{\phi_0''(x)}{\phi_0'(x)} = \frac{i}{2}\frac{d}{dx}\ln\phi_0'(x) = i\frac{d}{dx}\ln\sqrt{\phi_0'(x)} = i\frac{d}{dx}\ln\sqrt{k_0(x)} \,. \tag{5.13}$$

The wave function is then approximated by

$$\psi(x) \cong \frac{N}{\sqrt{k_0(x)}}e^{\pm\frac{i}{\hbar}\int dx\, k_0(x)} \,, \tag{5.14}$$

where N is a normalization constant. This is called the semiclassical or WKB approximation. A detailed description of how this can be used to obtain the eigenenergies

and eigenfunctions is given in most graduate quantum mechanics texts. Here we are interested primarily in using this approximation to understand some of the general features of the solutions to the one-dimensional Schrödinger equation.

Consider the local wave number

$$k_0(x) = \sqrt{2m(E - V(x))}. \tag{5.15}$$

When $E > V(x)$, $k_0(x)$ is real and the wave function will be oscillatory. This is the condition for the motion of classical particle in the same potential so this is called the *classically allowed* region and points where $E = V(x)$ are called the *classical turning points*. For $E < V(x)$, $k_0(x)$ will be imaginary and the wave function will either rise or fall exponentially. This is shown schematically in Fig. 5.1. In Fig. 5.1a, for the energy E_1, there are two classical turning points. In the region between the two points the potential is less than the energy and the solution is oscillatory. Outside of this region the solution is chosen to be exponentially damped so that the wave function is normalizable.

In Fig. 5.1b, for energy E_2, there is one classical turning point and to the right of this point the energy is always greater than the potential so the wave function is oscillatory. To the left of the classical turning point, the energy is less than the potential and the wave function is exponentially damped. Note that since the wave function is oscillatory and undamped from the turning point to ∞, the wave function is not square integrable. This is in contradiction to our assumptions about the Hilbert space and we will deal with these complications in more detail later.

In Fig. 5.1c, for energy E_3, there are no classical turning points and the energy is always greater than the potential so the wave function is oscillatory for $-\infty < x < \infty$. Again, the wave function is not normalizable.

The WKB wave function in the classically allowed region for bound states is given by

$$\psi(x) = N \cos\left(\int_{x_-}^{x} dx' k_0(x') - \frac{\pi}{4}\right) \tag{5.16}$$

and the quantization condition for the WKB approximation for bound states is

$$\frac{1}{\hbar} \int_{x_-}^{x_+} dx k_0(x) = \left(n + \frac{1}{2}\right)\pi, \tag{5.17}$$

where x_- and x_+ are the lower and upper turning points given by $k_0(x_\pm) = 0$, and $n = 0, 1, 2, 3, \ldots$. The quantization condition can only be satisfied for certain discrete values of the energies. The expression for the wave function and the quantization condition imply that the wave function will have n zeros or nodes between the classical turning points. In addition, since the classical turning points and k_0 depend on the energy, there can only be one eigenstate with a given energy. Furthermore, the energy of the eigenstates must increase with n. In other words, the greater the number of nodes in the eigenfunction, the greater the eigenenergy.
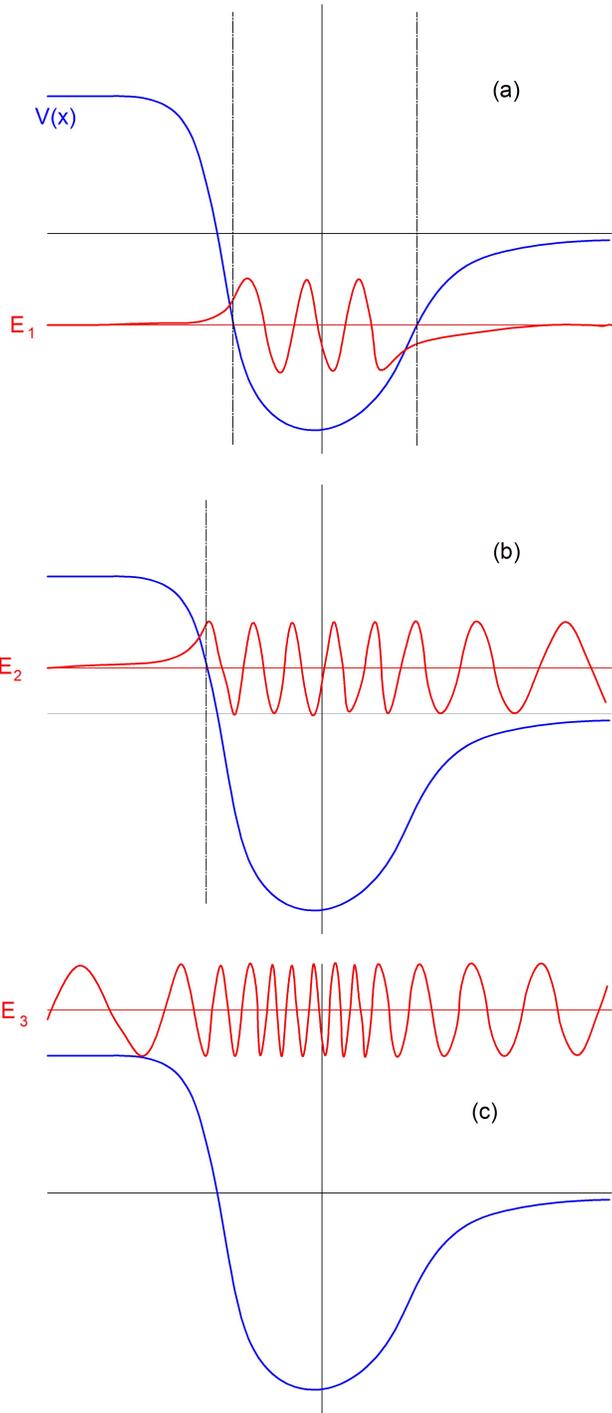
Figure 5.1:

5.1 The One-dimensional Infinite Square Well Potential

We will now solve the one-dimensional time-independent Schrödinger equation where the potential is of the form

$$V(x) = \begin{cases} 0 & \text{for } -\frac{a}{2} < x < \frac{a}{2} \\ \infty & \text{otherwise} \end{cases} \tag{5.18}$$

From (5.15), we can see that as $V \to \infty$, $k_0 \to i\infty$ and, therefore, $\psi \to 0$. This means that the wave function must vanish outside of the interval $-a/2 < x < a/2$. Since the wave function will be oscillatory within this interval for $E > 0$, the wave functions must satisfy the boundary conditions

$$\psi(-a/2) = \psi(a/2) = 0. \tag{5.19}$$

For $-a/2 < x < a/2$, the Schrödinger equation is

$$-\frac{\hbar^2}{2m}\frac{d^2}{dx^2}\psi(x) = E\psi(x). \tag{5.20}$$

Dividing both sides of the equation by $-\frac{\hbar^2}{2m}$ gives

$$\frac{d^2}{dx^2}\psi(x) = -\frac{2mE}{\hbar^2}\psi(x), \tag{5.21}$$

or

$$\frac{d^2}{dx^2}\psi(x) + k^2\psi(x) = 0, \tag{5.22}$$

where

$$k^2 = \frac{2mE}{\hbar^2}. \tag{5.23}$$

This a harmonic equation which has solutions of the form

$$\psi(x) = A\cos kx + B\sin kx. \tag{5.24}$$

We now need to impose the boundary conditions. First

$$0 = \psi(-a/2) = A\cos\left(-\frac{ka}{2}\right) + B\sin\left(-\frac{ka}{2}\right) = A\cos\left(\frac{ka}{2}\right) - B\sin\left(\frac{ka}{2}\right) \tag{5.25}$$

and

$$0 = \psi(a/2) = A\cos\left(\frac{ka}{2}\right) + B\sin\left(\frac{ka}{2}\right). \tag{5.26}$$

The two equations can only be satisfied simultaneously is either $A = 0$ or $B = 0$. If $B = 0$, then the boundary conditions require that

$$\cos\left(\frac{ka}{2}\right) = 0\,,\tag{5.27}$$

which in turn requires that

$$\frac{ka}{2} = \frac{n\pi}{2}\tag{5.28}$$

with n being an odd integer. Then

$$k_n = \frac{n\pi}{a}\,.\tag{5.29}$$

For $A = 0$,

$$\sin\left(\frac{ka}{2}\right) = 0\,,\tag{5.30}$$

which in turn requires that

$$\frac{ka}{2} = n\pi\,.\tag{5.31}$$

with n being an integer. Then

$$k_n = \frac{2n\pi}{a}\,.\tag{5.32}$$

These two solutions can be combined by writing the wave function as

$$\psi_n(x) = N_n \begin{cases} \cos\frac{n\pi x}{a} & \text{for } n \text{ odd} \\ \sin\frac{n\pi x}{a} & \text{for } n \text{ even} \end{cases}\tag{5.33}$$

where $n = 1, 2, 3, \ldots$. This implies that

$$k_n = \frac{n\pi}{a}\tag{5.34}$$

and

$$E_n = \frac{\hbar^2\pi^2 n^2}{2ma^2}\,.\tag{5.35}$$

Now consider

$$\int_{-a/2}^{a/2} dx\,\psi_m(x)\psi_n(x) = N_m N_n \int_{-a/2}^{a/2} dx\,\cos\frac{m\pi x}{a}\cos\frac{n\pi x}{a}\,.\tag{5.36}$$

If we change variables to $y = \frac{\pi x}{a}$ this becomes

$$N_m N_n \frac{a}{\pi} \int_{-\pi/2}^{\pi/2} dx\,\cos my \cos ny = N_m N_n \frac{a}{2\pi} \int_{-\pi/2}^{\pi/2} dx\,[\cos(m-n)y + \cos(m+n)y]\,.\tag{5.37}$$

For $m = n$ we have

$$N_n^2 \frac{a}{2\pi} \int_{-\pi/2}^{\pi/2} dy \left[1 + \cos 2ny\right] = N_n^2 \frac{a}{2\pi} \left[y + \frac{\sin 2ny}{2n}\right]_{-\pi/2}^{\pi/2} = N_n^2 \frac{a}{2} \,. \tag{5.38}$$

For $m \neq n$,

$$N_m N_n \frac{a}{2\pi} \int_{-\pi/2}^{\pi/2} dy \left[\cos(m-n)y + \cos(m+n)y\right]$$

$$= B_m B_n \frac{a}{2\pi} \left[\frac{\sin(m-n)y}{m-n} + \frac{\sin(m+n)y}{m+n}\right]\Big|_{-\pi/2}^{\pi/2} = 0 \tag{5.39}$$

since $m - n$ and $m + n$ are even. Similarly, it can be shown that

$$N_m N_n \frac{a}{\pi} \int_{-\pi/2}^{\pi/2} dy \cos my \sin ny = 0 \tag{5.40}$$

and

$$N_m N_n \frac{a}{\pi} \int_{-\pi/2}^{\pi/2} dy \sin my \sin ny = N_n^2 \frac{a}{2} \delta_{mn} \,. \tag{5.41}$$

So if we choose $N_n = \sqrt{\frac{2}{a}}$,

$$\int_0^a dx \psi_m(x)\psi_n(x) = \delta_{mn} \,. \tag{5.42}$$

So the wave functions are orthonormal. The wave functions for the first four values of n are shown in Fig. 5.2.

The general form of the time-dependent wave function can now be written as

$$\Psi(x,t) = \sqrt{\frac{2}{a}} \sum_{n=1}^{\infty} \left(A_n \cos\frac{(2n-1)\pi x}{a} e^{-\frac{i}{\hbar}E_{2n-1}t} + B_n \sin\frac{2n\pi x}{a} e^{-\frac{i}{\hbar}E_{2n}t}\right), \tag{5.43}$$

where the coefficients A_n and B_n are determined from the form of the wave function at $t = 0$. For example, consider the case where

$$\Psi(x,0) = \sqrt{\frac{2}{a}}\theta(x)\theta(a/2 - x) \,. \tag{5.44}$$

The θ-function is defined such that

$$\theta(x) = \begin{cases} 0 & \text{for } x < 0 \\ 1 & \text{for } x > 0 \end{cases} \,. \tag{5.45}$$

73
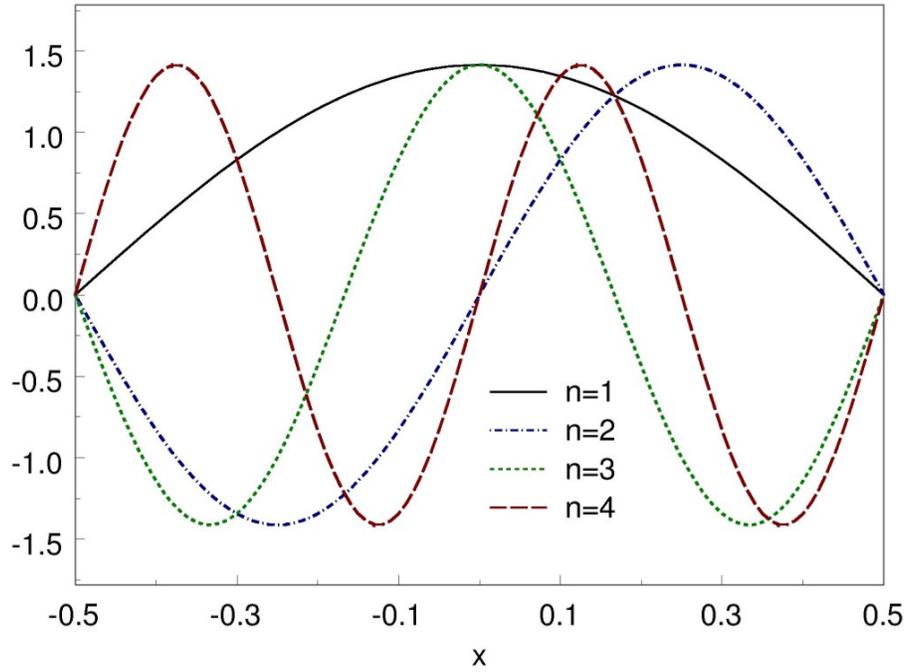
Figure 5.2: The infinite square well wave functions for $n = 1, 2, 3, 4$.

So this means that the wave function at $t = 0$ vanishes for $-a/2 < x < 0$ and is a constant for $0 < x < a/2$. The expansion coefficients are given by

$$
\begin{aligned}
A_n &= \sqrt{\frac{2}{a}} \int_{-a/2}^{a/2} dx \cos \frac{(2n-1)\pi x}{a} \Psi(x, 0) = \frac{2}{a} \int_0^{a/2} dx \cos \frac{(2n-1)\pi x}{a} \\
&= \frac{2}{a} \left[\frac{a}{(2n-1)\pi} \sin \frac{(2n-1)\pi x}{a} \right]_0^{a/2} = \frac{2}{(2n-1)\pi} \sin \frac{(2n-1)\pi}{2} \\
&= \frac{2}{(2n-1)\pi} (-1)^{n+1}
\end{aligned}
\tag{5.46}
$$

and

$$
\begin{aligned}
B_n &= \sqrt{\frac{2}{a}} \int_{-a/2}^{a/2} dx \sin \frac{2n\pi x}{a} \Psi(x, 0) = \frac{2}{a} \int_0^{a/2} dx \sin \frac{2n\pi x}{a} \\
&= -\frac{2}{a} \left[\frac{a}{2n\pi} \cos \frac{2n\pi x}{a} \right]_0^{a/2} = -\frac{1}{n\pi} [\cos n\pi - 1] \\
&= \frac{1}{n\pi} [1 - (-1)^n] .
\end{aligned}
\tag{5.47}
$$

74

Now consider the expectation value of the position

$$
\begin{aligned}
\langle x(t)\rangle &= \int_{-a/2}^{a/2} dx\, x \Psi^*(x,t)\Psi(x,t) \\
&= \frac{2}{a}\int_{-a/2}^{a/2} dx\, x \sum_{n=1}^{\infty}\left(A_n\cos\frac{(2n-1)\pi x}{a}e^{\frac{i}{\hbar}E_{2n-1}t}+B_n\sin\frac{2n\pi x}{a}e^{\frac{i}{\hbar}E_{2n}t}\right) \\
&\quad \times \sum_{n'=1}^{\infty}\left(A_{n'}\cos\frac{(2n'-1)\pi x}{a}e^{-\frac{i}{\hbar}E_{2n'-1}t}+B_{n'}\sin\frac{2n'\pi x}{a}e^{-\frac{i}{\hbar}E_{2n'}t}\right).\ (5.48)
\end{aligned}
$$

The integrals necessary to evaluate this are

$$
\int_{-a/2}^{a/2} dx\, x \cos\frac{(2n-1)\pi x}{a}\cos\frac{(2n'-1)\pi x}{a} = 0\,, \tag{5.49}
$$

$$
\int_{-a/2}^{a/2} dx\, x \sin\frac{2n\pi x}{a}\sin\frac{2n'\pi x}{a} = 0 \tag{5.50}
$$

and

$$
\int_{-a/2}^{a/2} dx\, x \cos\frac{(2n-1)\pi x}{a}\sin\frac{2n'\pi x}{a} = (-1)^{n+n'}\frac{8a^2 n'(2n-1)}{\pi^2[(2n-1)^2-4n'^2]^2}\,. \tag{5.51}
$$

Using these the expectation value of the position operator can be evaluated as

$$
\begin{aligned}
\langle x(t)\rangle &= \frac{16a}{\pi^2}\sum_{n=1}^{\infty}\sum_{n'=1}^{\infty}\Bigg\{A_n B_{n'}(-1)^{n+n'}\frac{n'(2n-1)}{[(2n-1)^2-4n'^2]^2}e^{\frac{i}{\hbar}(E_{2n-1}-E_{2n'})t} \\
&\quad +A_{n'}B_n(-1)^{n+n'}\frac{n(2n'-1)}{[(2n'-1)^2-4n^2]^2}e^{\frac{i}{\hbar}(E_{2n}-E_{2n'-1})t}\Bigg\} \\
&= \frac{16a}{\pi^2}\sum_{n=1}^{\infty}\sum_{n'=1}^{\infty}A_n B_{n'}(-1)^{n+n'}\frac{n'(2n-1)}{[(2n-1)^2-4n'^2]^2} \\
&\quad \times \left(e^{\frac{i}{\hbar}(E_{2n-1}-E_{2n'})t}+e^{-\frac{i}{\hbar}(E_{2n-1}-E_{2n'})t}\right) \\
&= \frac{32a}{\pi^2}\sum_{n=1}^{\infty}\sum_{n'=1}^{\infty}A_n B_{n'}(-1)^{n+n'}\frac{n'(2n-1)}{[(2n-1)^2-4n'^2]^2} \\
&\quad \times \cos\left(\frac{1}{\hbar}(E_{2n-1}-E_{2n'})t\right), \tag{5.52}
\end{aligned}
$$

where in the second step the labels of the summation indices have been interchanged.

For the example we are considering here

$$A_n B_{n'} (-1)^{n+n'} \frac{n'(2n-1)}{[(2n-1)^2 - 4n'^2]^2} = \frac{2}{(2n-1)\pi} (-1)^{n+1} \frac{1}{n'\pi} \left[1 - (-1)^{n'} \right]$$

$$\times (-1)^{n+n'} \frac{n'(2n-1)}{[(2n-1)^2 - 4n'^2]^2}$$

$$= \frac{2}{\pi^2} \frac{1-(-1)^{n'}}{[(2n-1)^2 - 4n'^2]^2} \,. \tag{5.53}$$

The argument of the cosine can be reexpressed using

$$\frac{1}{\hbar} (E_{2n-1} - E_{2n'}) = \frac{1}{\hbar} \left(\frac{\hbar^2 \pi^2 (2n-1)^2}{2ma^2} - \frac{\hbar^2 \pi^2 4n'^2}{2ma^2} \right) = \frac{\hbar \pi^2}{2ma^2} \left[(2n-1)^2 - 4n'^2 \right] \,. \tag{5.54}$$

The expectation value of the position operator is then given by

$$\langle x(t) \rangle = \frac{64a}{\pi^4} \sum_{n=1}^{\infty} \sum_{n'=1}^{\infty} \frac{1-(-1)^{n'}}{[(2n-1)^2 - 4n'^2]^2} \cos \left(\frac{\hbar \pi^2}{2ma^2} \left[(2n-1)^2 - 4n'^2 \right] t \right) \,. \tag{5.55}$$

We can also obtain the expectation value of the momentum operator as

$$\langle p(t) \rangle = \int_{-a/2}^{a/2} dx \Psi^*(x,t) \frac{\hbar}{i} \frac{\partial}{\partial x} \Psi(x,t)$$

$$= \frac{2\hbar}{ia} \int_{-a/2}^{a/2} dx \sum_{n=1}^{\infty} \left(A_n \cos \frac{(2n-1)\pi x}{a} e^{\frac{i}{\hbar} E_{2n-1} t} + B_n \sin \frac{2n\pi x}{a} e^{\frac{i}{\hbar} E_{2n} t} \right)$$

$$\times \frac{\partial}{\partial x} \sum_{n'=1}^{\infty} \left(A_{n'} \cos \frac{(2n'-1)\pi x}{a} e^{-\frac{i}{\hbar} E_{2n'-1} t} + B_{n'} \sin \frac{2n'\pi x}{a} e^{-\frac{i}{\hbar} E_{2n'} t} \right)$$

$$= \frac{2\hbar}{ia} \int_{-a/2}^{a/2} dx \sum_{n=1}^{\infty} \left(A_n \cos \frac{(2n-1)\pi x}{a} e^{\frac{i}{\hbar} E_{2n-1} t} + B_n \sin \frac{2n\pi x}{a} e^{\frac{i}{\hbar} E_{2n} t} \right)$$

$$\times \sum_{n'=1}^{\infty} \left(-A_{n'} \frac{(2n'-1)\pi}{a} \sin \frac{(2n'-1)\pi x}{a} e^{-\frac{i}{\hbar} E_{2n'-1} t} \right.$$

$$\left. + B_{n'} \frac{2n'\pi}{a} \cos \frac{2n'\pi x}{a} e^{-\frac{i}{\hbar} E_{2n'} t} \right) \,. \tag{5.56}$$

Evaluation of this requires the integrals

$$\int_{-a/2}^{a/2} dx \cos \frac{(2n-1)\pi x}{a} \sin \frac{(2n'-1)\pi x}{a} = 0 \,, \tag{5.57}$$

$$\int_{-a/2}^{a/2} dx \sin \frac{2n\pi x}{a} \cos \frac{2n'\pi x}{a} = 0 \,, \tag{5.58}$$

76

$$\int_{-a/2}^{a/2} dx \cos \frac{(2n-1)\pi x}{a} \cos \frac{2n'\pi x}{a} = (-1)^{n'+n} \frac{2a}{\pi} \frac{2n-1}{4n'^2 - (2n-1)^2} \qquad (5.59)$$

and

$$\int_{-a/2}^{a/2} dx \sin \frac{2n\pi x}{a} \sin \frac{(2n'-1)\pi x}{a} = (-1)^{n'+n} \frac{2a}{\pi} \frac{2n}{4n^2 - (2n'-1)^2} . \qquad (5.60)$$

The expectation value of the momentum operator is then

$$
\begin{aligned}
\langle p(t) \rangle &= \frac{2\hbar}{ia} \sum_{n=1}^{\infty} \sum_{n'=1}^{\infty} \left\{ A_n B_{n'} \frac{2n'\pi}{a} (-1)^{n'+n} \frac{2a}{\pi} \frac{2n-1}{4n'^2 - (2n-1)^2} e^{\frac{i}{\hbar}(E_{2n-1} - E_{2n'})t} \right. \\
&\qquad \left. - A_{n'} B_n \frac{(2n'-1)\pi}{a} (-1)^{n'+n} \frac{2a}{\pi} \frac{2n}{4n^2 - (2n'-1)^2} e^{-\frac{i}{\hbar}(E_{2n'-1} - E_{2n})t} \right\} \\
&= \frac{8\hbar}{ia} \sum_{n=1}^{\infty} \sum_{n'=1}^{\infty} A_n B_{n'} (-1)^{n'+n} \frac{n'(2n-1)}{4n'^2 - (2n-1)^2} \\
&\qquad \times \left(e^{\frac{i}{\hbar}(E_{2n-1} - E_{2n'})t} - e^{\frac{i}{\hbar}(E_{2n-1} - E_{2n'})t} \right) \\
&= \frac{16\hbar}{a} \sum_{n=1}^{\infty} \sum_{n'=1}^{\infty} A_n B_{n'} (-1)^{n'+n} \frac{n'(2n-1)}{4n'^2 - (2n-1)^2} \sin \left(\frac{1}{\hbar}(E_{2n-1} - E_{2n'})t \right)
\end{aligned}
$$
$$\qquad (5.61)$$

For the example considered here

$$A_n B_{n'} (-1)^{n+n'} \frac{n'(2n-1)}{4n'^2 - (2n-1)^2} = \frac{2}{\pi^2} \frac{1 - (-1)^{n'}}{4n'^2 - (2n-1)^2} . \qquad (5.62)$$

Thus,

$$\langle p(t) \rangle = \frac{32\hbar}{\pi^2 a} \sum_{n=1}^{\infty} \sum_{n'=1}^{\infty} \frac{1 - (-1)^{n'}}{4n'^2 - (2n-1)^2} \sin \left(\frac{\hbar \pi^2}{2ma^2} \left[(2n-1)^2 - 4n'^2 \right] t \right) . \qquad (5.63)$$

The average position and momentum as a function of time are shown in Fig. 5.3

Using our result for $\langle x(t) \rangle$, the time rate of change of the average position is given by

$$
\begin{aligned}
\frac{d}{dt} \langle x(t) \rangle &= -\frac{64a}{\pi^4} \sum_{n=1}^{\infty} \sum_{n'=1}^{\infty} \frac{1 - (-1)^{n'}}{[(2n-1)^2 - 4n'^2]^2} \\
&\qquad \times \frac{\hbar \pi^2}{2ma^2} \left[(2n-1)^2 - 4n'^2 \right] \sin \left(\frac{\hbar \pi^2}{2ma^2} \left[(2n-1)^2 - 4n'^2 \right] t \right) \\
&= \frac{32\hbar}{\pi^2 ma} \sum_{n=1}^{\infty} \sum_{n'=1}^{\infty} \frac{1 - (-1)^{n'}}{4n'^2 - (2n-1)^2} \sin \left(\frac{\hbar \pi^2}{2ma^2} \left[(2n-1)^2 - 4n'^2 \right] t \right) \\
&= \frac{\langle p(t) \rangle}{m} .
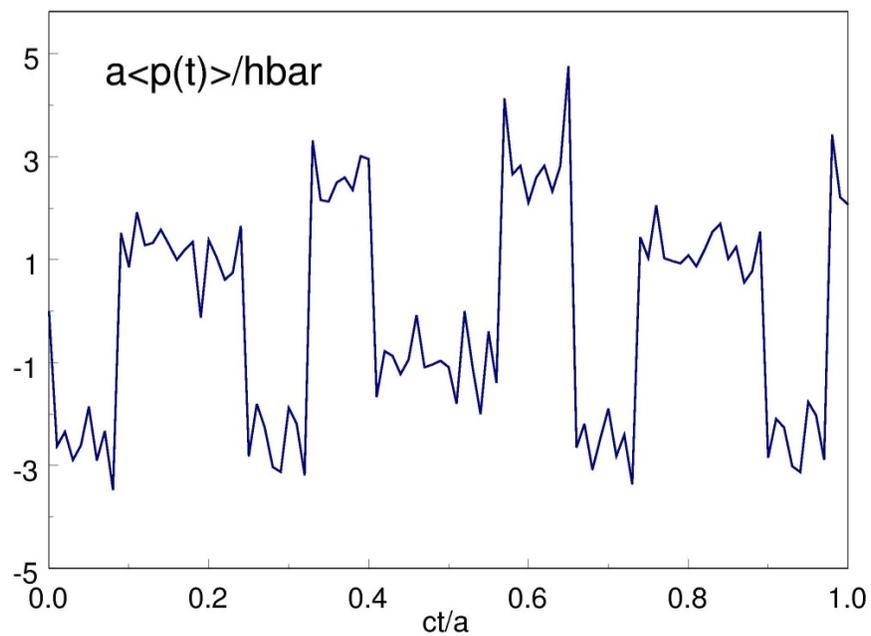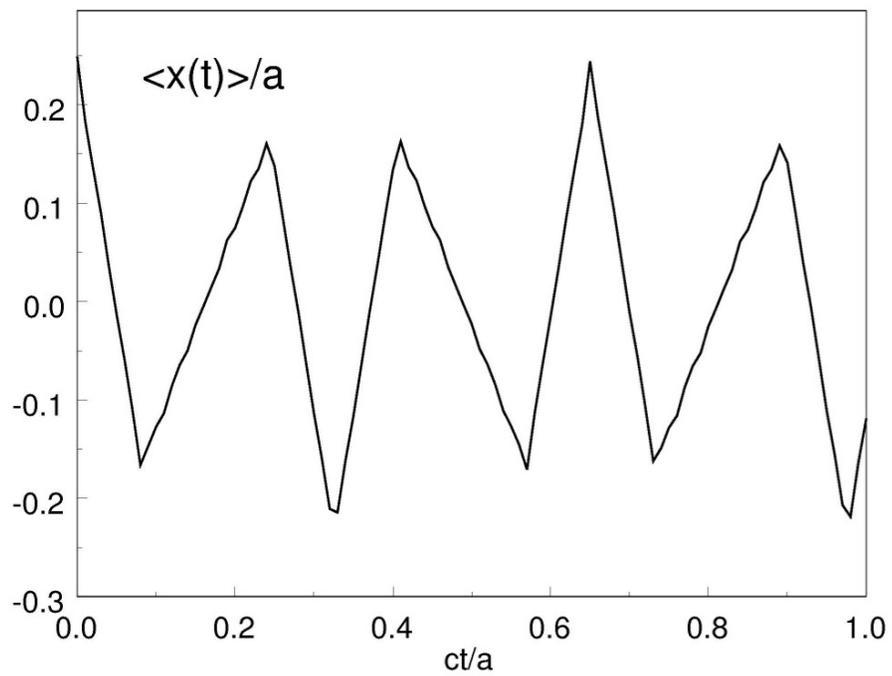\end{aligned}
$$
$$\qquad (5.64)$$

Figure 5.3: Average position and momentum as a function of time.

This shows that the rate of change of the average position is related to the average momentum in exactly the wave that the velocity and momentum of a classical particle are related. This is in fact a general result.

To see this, consider the expectation value of some operator $\hat{A}$. The time rate of change of this operator is

$$
\begin{aligned}
\frac{d}{dt}\langle A(t)\rangle &= \frac{d}{dt}\int_{-\infty}^{\infty} dx\, \Psi^*(x,t)\hat{A}\Psi(x,t) \\
&= \int_{-\infty}^{\infty} dx\left[\left(\frac{\partial}{\partial t}\Psi^*(x,t)\right)\hat{A}\Psi(x,t) + \Psi^*(x,t)\hat{A}\frac{\partial}{\partial t}\Psi(x,t)\right. \\
&\quad \left. +\Psi^*(x,t)\frac{\partial\hat{A}}{\partial t}\Psi(x,t)\right].
\end{aligned}
\tag{5.65}
$$

The time derivatives of the wave function can be replaced using the Schrödinger equation to give

$$
\begin{aligned}
\frac{d}{dt}\langle A(t)\rangle &= \int_{-\infty}^{\infty} dx\left[\left(\frac{1}{i\hbar}\hat{H}\Psi(x,t)\right)^* \hat{A}\Psi(x,t) + \Psi^*(x,t)\hat{A}\frac{1}{i\hbar}\hat{H}\Psi(x,t)\right. \\
&\quad \left. +\Psi^*(x,t)\frac{\partial\hat{A}}{\partial t}\Psi(x,t)\right] \\
&= \int_{-\infty}^{\infty} dx\,\Psi^*(x,t)\left\{\frac{1}{i\hbar}[\hat{A},\hat{H}] + \frac{\partial\hat{A}}{\partial t}\right\}\Psi(x,t) \\
&= \left\langle\frac{1}{i\hbar}[\hat{A},\hat{H}]\right\rangle + \left\langle\frac{\partial\hat{A}}{\partial t}\right\rangle
\end{aligned}
\tag{5.66}
$$

In the case of the position operator this becomes

$$
\frac{d}{dt}\langle x(t)\rangle = \left\langle\frac{1}{i\hbar}[\hat{x},\hat{H}]\right\rangle.
\tag{5.67}
$$

The commutator of the position operator and the hamiltonian operator is

$$
[\hat{x},\hat{H}(\hat{p},\hat{x})] = -\frac{\hbar}{i}\frac{\partial}{\partial\hat{p}}\hat{H}(\hat{p},\hat{x}) = i\hbar\frac{\hat{p}}{m}.
\tag{5.68}
$$

The time rate of change of the expectation value of the position operator is then

$$
\frac{d}{dt}\langle x(t)\rangle = \frac{\langle p\rangle}{m},
\tag{5.69}
$$

which we have obtained for a special case. Similarly, for the momentum operator we obtain

$$
\frac{d}{dt}\langle p(t)\rangle = \left\langle\frac{1}{i\hbar}[\hat{p},\hat{H}]\right\rangle
\tag{5.70}
$$

79

Here the commutation relation is

$$[\hat{p}, \hat{H}(\hat{p}, \hat{x})] = \frac{\hbar}{i}\frac{\partial}{\partial \hat{x}}\hat{H}(\hat{p}, \hat{x}) = -i\hbar\frac{\partial}{\partial \hat{x}}V(\hat{x}) \tag{5.71}$$

giving the result

$$\frac{d}{dt}\langle p(t)\rangle = -\left\langle \frac{\partial}{\partial \hat{x}}V(\hat{x})\right\rangle . \tag{5.72}$$

What we have now shown is that in general the average position and momentum operators obey classical hamilton's equations. This result is known as *Ehrenfest's theorem*.

5.2 Bound States in a Finite Square Well Potential

We will now look at solutions for the finite square well potential defined as

$$V(x) = \begin{cases} 0 & \text{for } x < -\frac{a}{2} \\ -V_0 & \text{for } -\frac{a}{2} < x < \frac{a}{2} \\ 0 & \text{for } \frac{a}{2} < x \end{cases} . \tag{5.73}$$

The finite square-well potential is illustrated in Fig. 5.4. This naturally implies that there are for regions to be considered: Region I for $x < -a/2$, Region II for $-a/2 < x < a/2$ and Region III for $a/2 < x$. The Schrödinger equation with this potential will have localized bound states for $E < 0$. In regions I and III, the time-independent Schrödinger equation will be

$$-\frac{\hbar^2}{2m}\frac{d^2}{dx^2}\psi(x) = E\psi(x) \tag{5.74}$$

or

$$\frac{d^2}{dx^2}\psi(x) - \frac{2m|E|}{\hbar^2}\psi(x) = 0 . \tag{5.75}$$

This has solutions of the form

$$\psi(x) = c_1 e^{\kappa x} + c_2 e^{-\kappa x} , \tag{5.76}$$

where

$$\kappa = \sqrt{\frac{2m|E|}{\hbar^2}} . \tag{5.77}$$

In Region II the time-independent Schrödinger equation is

$$-\frac{\hbar^2}{2m}\frac{d^2}{dx^2}\psi(x) - V_0\psi(x) = E\psi(x) \tag{5.78}$$

or

$$\frac{d^2}{dx^2}\psi(x) + \frac{2m(V_0 - |E|)}{\hbar^2}\psi(x) = 0 . \tag{5.79}$$

80
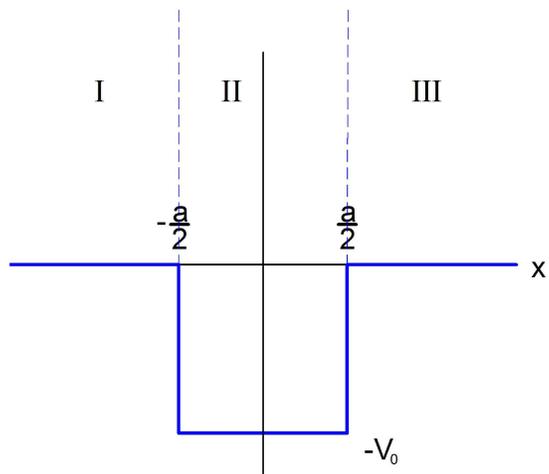
Figure 5.4: Finite square well potential.

As long as $V_0 > |E|$, this will have oscillatory solutions of the form

$$\psi(x) = c_3 \cos qx + c_4 \sin qx,\qquad(5.80)$$

where

$$q = \sqrt{\frac{2m(V_0 - |E|)}{\hbar^2}}.\qquad(5.81)$$

In order for the wave function to be normalizable, we require that $\psi(-\infty) = \psi(\infty) = 0$. Imposing the condition on the wave function, we can define the wave function piecewise on the three intervals as

$$\psi^{(I)}(x) = Ae^{\kappa x}\qquad(5.82)$$
$$\psi^{(II)}(x) = B\cos qx + C\sin qx\qquad(5.83)$$
$$\psi^{(III)}(x) = De^{-\kappa x}.\qquad(5.84)$$

Since the potential is discontinuous at the boundaries between the regions, we need to define boundary conditions at these boundaries that join the the wave function in the various regions and determine three of the four constants. The physical requirement that determines the boundary conditions is that the probability current at the boundaries must be continuous. For example at $x = a/2$ we require that the current leaving Region II must be the equal to the current entering Region III. This can be stated mathematically as

$$\frac{\hbar}{2im}\left[\psi^{(II)*}(a/2)\psi^{(II)\prime}(a/2) - \psi^{(II)\prime*}(a/2)\psi^{(II)}(a/2)\right]$$
$$= \frac{\hbar}{2im}\left[\psi^{(III)*}(a/2)\psi^{(III)\prime}(a/2) - \psi^{(III)\prime*}(a/2)\psi^{(III)}(a/2)\right]\qquad(5.85)$$

This is satisfied if

$$\psi^{(II)}(a/2) = \psi^{(III)}(a/2)\qquad(5.86)$$

and

$$\psi^{(II)\prime}(a/2) = \psi^{(III)\prime}(a/2).\qquad(5.87)$$

Similarly, the other boundary

$$\psi^{(I)}(a/2) = \psi^{(II)}(a/2)\qquad(5.88)$$

and

$$\psi^{(I)\prime}(a/2) = \psi^{(II)\prime}(a/2).\qquad(5.89)$$

Applying these boundary conditions to our general solution gives

$$Ae^{-\kappa a/2} = B\cos\frac{qa}{2} - C\sin\frac{qa}{2}\qquad(5.90)$$
$$\kappa Ae^{-\kappa a/2} = qB\sin\frac{qa}{2} + qC\cos\frac{qa}{2}\qquad(5.91)$$
$$B\cos\frac{qa}{2} + C\sin\frac{qa}{2} = De^{-\kappa a/2}\qquad(5.92)$$
$$-qB\sin\frac{qa}{2} + qC\sin\frac{qa}{2} = -\kappa De^{-\kappa a/2}\qquad(5.93)$$

If we now divide (5.91) by (5.90) we obtain

$$\kappa = q \frac{B \sin \frac{qa}{2} + C \cos \frac{qa}{2}}{B \cos \frac{qa}{2} - C \sin \frac{qa}{2}}. \tag{5.94}$$

Similarly, if we divide (5.93) by (5.92) we obtain

$$-q \frac{B \cos \frac{qa}{2} - C \sin \frac{qa}{2}}{B \cos \frac{qa}{2} + C \sin \frac{qa}{2}} = -\kappa. \tag{5.95}$$

We now solve (5.94) to give

$$B \left(\frac{\kappa}{q} \cos \frac{qa}{2} - \sin \frac{qa}{2} \right) = C \left(\frac{\kappa}{q} \sin \frac{qa}{2} + \cos \frac{qa}{2} \right) \tag{5.96}$$

and (5.95) to give

$$-B \left(\frac{\kappa}{q} \cos \frac{qa}{2} - \sin \frac{qa}{2} \right) = C \left(\frac{\kappa}{q} \sin \frac{qa}{2} + \cos \frac{qa}{2} \right). \tag{5.97}$$

It is easy to see that these equations are incompatible since the right sides differ by an overall sign while the right sides are identical. The only way they can be made compatible is if either $B = 0$ or $C = 0$. If we choose $C = 0$, then (5.96) becomes

$$B \left(\frac{\kappa}{q} \cos \frac{qa}{2} - \sin \frac{qa}{2} \right) = 0 \tag{5.98}$$

which will have a nontrivial solution only if

$$\frac{\kappa}{q} \cos \frac{qa}{2} - \sin \frac{qa}{2} = 0 \tag{5.99}$$

or

$$\frac{\kappa}{q} = \tan \frac{qa}{2}. \tag{5.100}$$

If we choose $B = 0$, then (5.96) becomes

$$0 = C \left(\frac{\kappa}{q} \sin \frac{qa}{2} + \cos \frac{qa}{2} \right) \tag{5.101}$$

which requires that

$$\frac{\kappa}{q} \sin \frac{qa}{2} + \cos \frac{qa}{2} = 0 \tag{5.102}$$

or

$$\frac{\kappa}{q} = -\cot \frac{qa}{2}. \tag{5.103}$$

The solutions to (5.100) and (5.103) then give the eigenvalues for the bound states.

83

These equations can be easily solved numerically and can be solved approximately using graphical methods. To do this, note that

$$q^2 = \frac{2mV_0}{\hbar^2} - \frac{2m|E|}{\hbar^2} = \frac{2mV_0}{\hbar^2} - \kappa^2 \tag{5.104}$$

or

$$\kappa = \sqrt{\frac{2mV_0}{\hbar^2} - q^2} \, . \tag{5.105}$$

Therefore,

$$\frac{\kappa}{q} = \sqrt{\frac{2mV_0}{\hbar^2 q^2} - 1} = \sqrt{\frac{ma^2V_0}{2\hbar^2(\frac{qa}{2})^2} - 1} = \sqrt{\frac{b^2}{(\frac{qa}{2})^2} - 1} \, . \tag{5.106}$$

Then (5.100) and (5.103) can be written as

$$\sqrt{\frac{b^2}{(\frac{qa}{2})^2} - 1} = \tan\frac{qa}{2} \tag{5.107}$$

and

$$\sqrt{\frac{b^2}{(\frac{qa}{2})^2} - 1} = -\cot\frac{qa}{2} \, . \tag{5.108}$$

First consider (5.107). The top panel of panel of Fig. 5.5 shows a plot of $\tan\frac{qa}{2}$ and plots of $\sqrt{\frac{b^2}{(\frac{qa}{2})^2} - 1}$ for three values of b. Equation (5.107) implies that eigenvalues will occur when a curve for $\sqrt{\frac{b^2}{(\frac{qa}{2})^2} - 1}$ crosses the curve for $\tan\frac{qa}{2}$. Since the square-root term is nonzero only for $0 < \frac{qa}{2} < b$, the number of eigenvalues is limited by the the value of b. The lower panel of this figure shows a similar set of curves for (5.108). Here the curve for $-\cot\frac{qa}{2}$ is similar to that for $\tan\frac{qa}{2}$ but is shifted to the right by $\frac{\pi}{2}$. As a result the eigenvalues for this case are always larger than for the corresponding eigenvalue of the first case.

We can now determine the coefficients in the wave functions. For the case where $B \neq 0$ and $C = 0$, we can solve (5.90) to give

$$A = e^{\kappa a/2} \cos\frac{qa}{2} \, B \tag{5.109}$$

and (5.92) to give

$$D = e^{\kappa a/2} \cos\frac{qa}{2} \, B \, . \tag{5.110}$$

The remaining constant B is then fixed by normalizing the wave functions. For the case where $C \neq 0$ and $B = 0$, we can solve (5.90) to give

$$A = -e^{\kappa a/2} \sin\frac{qa}{2} \, C \tag{5.111}$$

84
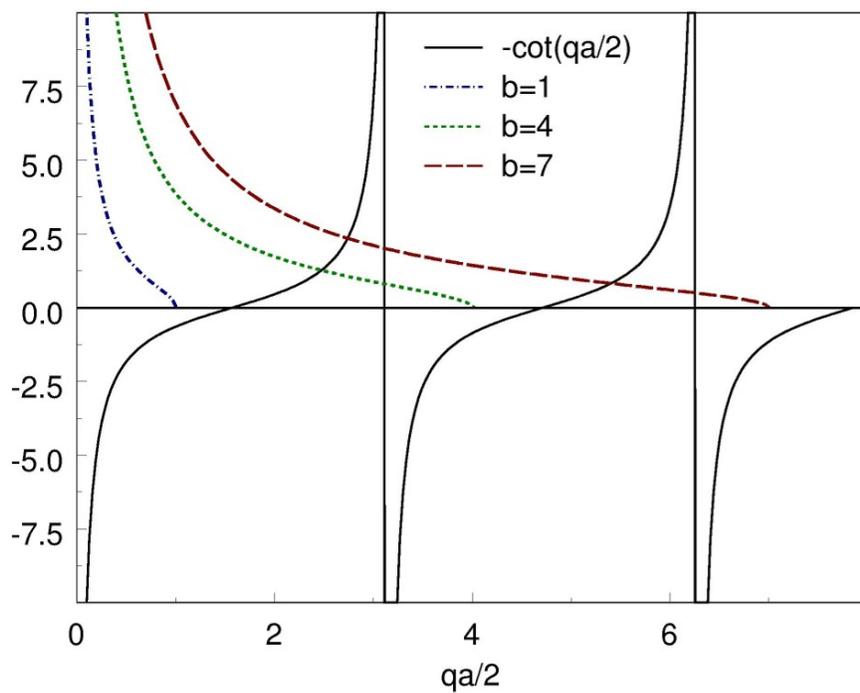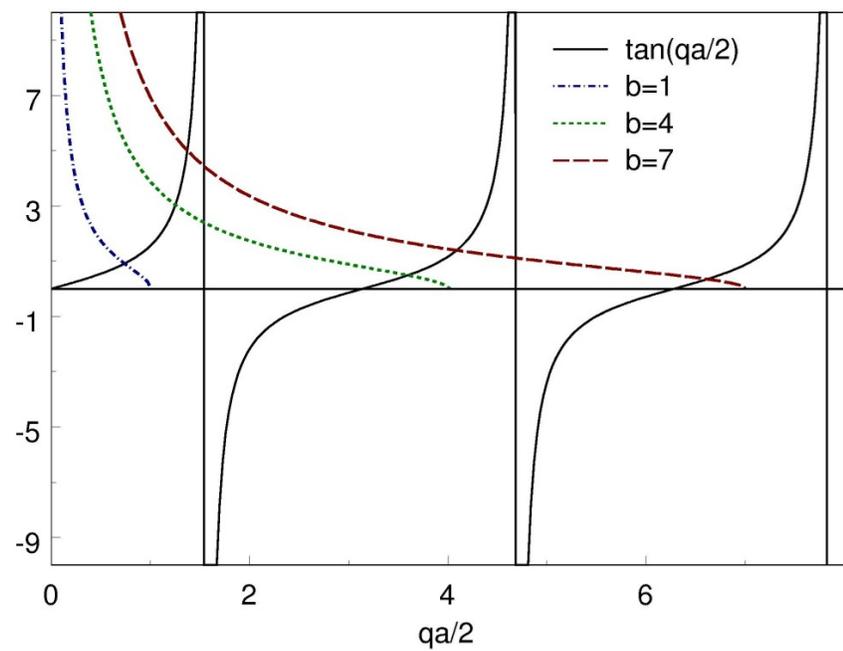
Figure 5.5: Graphical representations of the solutions to the eigenvalue equations.
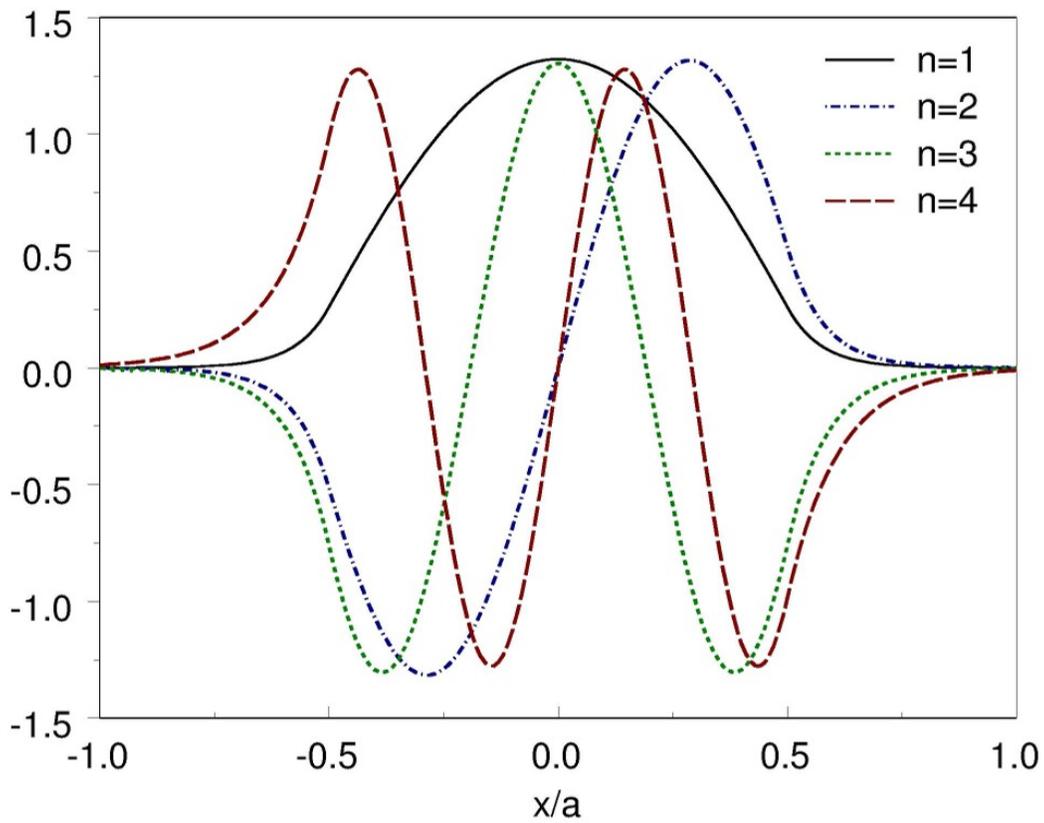
Figure 5.6: Wave functions for the first four eigenstates in the finite square potential with $b = 7$ and $a = 1$.

and (5.92) to give

$$D = e^{\kappa a/2} \sin \frac{qa}{2} C \,. \qquad (5.112)$$

C is then fixed by normalization.

The five wave functions corresponding to $b = 7$ are shown in (5.6) for $a = 1$. Since the wave functions for $n = 1, 3, 5$ correspond to the case where $C = 0$ and the wave function for $-a/2 < x < a/2$ are cosines, these wave functions are even about 0. The wave functions for $n = 2, 4$ correspond to $B = 0$ where the wave function is a sine and are odd about zero. This is in fact the result of a property of the hamiltonian in this case that can be shown to have more general consequences.

5.2.1   Parity

The symmetry about the origin is related to the parity operator $\hat{\mathcal{P}}$ which is defined such that

$$\hat{\mathcal{P}}\psi(x) = \eta\psi(-x) \qquad (5.113)$$

where η is an arbitrary phase. For simplicity, we will choose $\eta = 1$. Using this definition,

$$\hat{\mathcal{P}}^2\psi(x) = \hat{\mathcal{P}}\left(\hat{\mathcal{P}}\psi(x)\right) = \hat{\mathcal{P}}\psi(-x) = \psi(x) \,. \qquad (5.114)$$

So,

$$\hat{\mathcal{P}}^2 = 1 \,. \qquad (5.115)$$

Now consider the application of the parity operator to both sides of the eigenequation

$$\hat{\mathcal{P}}\hat{H}\psi(x) = E\hat{\mathcal{P}}\psi(x) \,. \qquad (5.116)$$

Inserting 1 in the form of (5.115) between the hamiltonian and the wave function yields

$$\hat{\mathcal{P}}\hat{H}\hat{\mathcal{P}}\hat{\mathcal{P}}\psi(x) = E\hat{\mathcal{P}}\psi(x) \qquad (5.117)$$

or

$$\hat{\mathcal{P}}\hat{H}\hat{\mathcal{P}}\psi(-x) = E\psi(-x) \,. \qquad (5.118)$$

So $\psi(-x)$ is an eigenstate of $\hat{\mathcal{P}}\hat{H}\hat{\mathcal{P}}$. This is only interesting if we know the effect of the parity operator on the hamiltonian. To determine this we need to define the action of the parity operator on the canonical operators. These are defined such that

$$\hat{\mathcal{P}}x\hat{\mathcal{P}} = -x \qquad (5.119)$$

and

$$\hat{\mathcal{P}}\hat{p}\hat{\mathcal{P}} = -\hat{p} \,. \qquad (5.120)$$

If we have a hamiltonian operator of the form

$$\hat{H} = \frac{\hat{p}^2}{2m} + V(x) \,, \qquad (5.121)$$

then

$$\hat{\mathcal{P}}\hat{H}\hat{\mathcal{P}} = \frac{1}{2m}\hat{\mathcal{P}}\hat{p}\hat{\mathcal{P}}\hat{\mathcal{P}}\hat{p}\hat{\mathcal{P}} + \hat{\mathcal{P}}V(x)\hat{\mathcal{P}} = \frac{\hat{p}^2}{2m} + V(-x).$$ (5.122)

Therefore, if

$$V(-x) = V(x),$$ (5.123)

then

$$\hat{\mathcal{P}}\hat{H}\hat{\mathcal{P}} = \hat{H}$$ (5.124)

or

$$\left[\hat{\mathcal{P}}, \hat{H}\right] = 0.$$ (5.125)

To see what this means consider the matrix element of some operator $\hat{\mathcal{O}}$ between to time-dependent wave functions $\Psi_f(x, t)$ and $\Psi_i(x, t)$. Then

$$\mathcal{O}_{fi} = \int_{-\infty}^{\infty} dx \Psi_f^*(x, t)\hat{\mathcal{O}}\Psi_i(x, t).$$ (5.126)

The time rate of change of this matrix element is then

$$\frac{d}{dt}\mathcal{O}_{fi} = \int_{-\infty}^{\infty} dx \left\{ \left(\frac{\partial}{\partial t}\Psi_f^*(x, t)\right)\hat{\mathcal{O}}\Psi_i(x, t) + \Psi_f^*(x, t)\frac{\partial \hat{\mathcal{O}}}{\partial t}\Psi_i(x, t) \right.$$

$$\left. + \Psi_f^*(x, t)\hat{\mathcal{O}}\frac{\partial}{\partial t}\Psi_i(x, t) \right\}.$$ (5.127)

We can now use the Schrödinger equation to give

$$\frac{\partial}{\partial t}\Psi_f^*(x, t) = -\frac{1}{i\hbar}\Psi_f^*(x, t)\hat{H}$$ (5.128)

and

$$\frac{\partial}{\partial t}\Psi_i(x, t) = \frac{1}{i\hbar}\hat{H}\Psi_i(x, t).$$ (5.129)

The time rate of change of the matrix element can then be written as

$$\frac{d}{dt}\mathcal{O}_{fi} = \int_{-\infty}^{\infty} dx \Psi_f^*(x, t)\left\{\frac{\partial \hat{\mathcal{O}}}{\partial t} + \frac{1}{i\hbar}\left[\hat{\mathcal{O}}, \hat{H}\right]\right\}\Psi_i(x, t).$$ (5.130)

This means that if the operator has no explicit time dependence and commutes with the hamiltonian, then

$$\frac{d}{dt}\mathcal{O}_{fi} = 0.$$ (5.131)

That is all of the matrix elements of the operator are constants of motion.

Now consider the matrix elements of $\left[\hat{\mathcal{O}}, \hat{H}\right] = 0$ between to eigenstates of the energy

$$\int_{-\infty}^{\infty} dx \psi_m^*(x)\left(\hat{\mathcal{O}}\hat{H} - \hat{H}\hat{\mathcal{O}}\right)\psi_n(x) = 0.$$ (5.132)

Since $\hat{H}$ is hermitian, this can be rewritten as

$$0 = \int_{-\infty}^{\infty} dx \psi_m^*(x) \left(\hat{\mathcal{O}} E_n - E_m \hat{\mathcal{O}} \right) \psi_n(x) = (E_n - E_m) \int_{-\infty}^{\infty} dx \psi_m^*(x) \hat{\mathcal{O}} \psi_n(x). \quad (5.133)$$

Therefore for any to nondegenerate states we must have

$$\int_{-\infty}^{\infty} dx \psi_m^*(x) \hat{\mathcal{O}} \psi_n(x) = 0. \quad (5.134)$$

Since this must be true for any m where $E_m \neq E_n$, this requires that

$$\hat{\mathcal{O}} \psi_n(x) = \alpha \psi_n(x) \quad (5.135)$$

where α is a constant.

In summary, any operator $\hat{\mathcal{O}}$ that is time independent and commutes with the hamiltonian $\hat{H}$ can have eigenstates which are also eigenstates of the hamiltonian. If the eigenvalue is to correspond to an observable, we must also require that $\hat{\mathcal{O}}$ be hermitian.

Since we have shown that for potentials which are even in x, the parity operator $\hat{\mathcal{P}}$ commutes with the hamiltonian, we know that we can construct eigenstates of $\hat{\mathcal{P}}$ which will also be energy eigenstates. The eigenstates of $\hat{\mathcal{P}}$ are easily constructed as

$$\psi^+(x) = \frac{1}{2} \left(\psi(x) + \psi(-x) \right) \quad (5.136)$$

and

$$\psi^-(x) = \frac{1}{2} \left(\psi(x) - \psi(-x) \right) \quad (5.137)$$

where

$$\hat{\mathcal{P}} \psi^{\pm}(x) = \pm \psi^{\pm}(x). \quad (5.138)$$

Note that $\psi^+(x)$ is even in x while $\psi^-(x)$ is odd in x.

This means that the eigenstates of a hamiltonian which is even in x will have eigenstates which are either even or odd in x. This is indeed what we have found for the case of the square well.

5.3   The One-Dimensional Harmonic Oscillator

We will now consider the case of the one-dimensional harmonic oscillator. The classical hamiltonian for this is

$$H = \frac{p^2}{2m} + \frac{1}{2} k x^2 \quad (5.139)$$

where the spring constant can be written as $k = m\omega^2$ and ω is the angular frequency of the oscillator.

Using this, the time-independent Schrödinger equation is

$$-\frac{\hbar^2}{2m}\frac{d^2}{dx^2}\psi(x) + \frac{1}{2}m\omega^2 x^2\psi(x) = E\psi(x)\,. \tag{5.140}$$

This can be rearranged to give

$$\frac{d^2}{dx^2}\psi(x) + \frac{2mE}{\hbar^2}\psi(x) - \frac{m^2\omega^2 x^2}{\hbar^2}\psi(x) = 0\,. \tag{5.141}$$

In order to solve this differential equation, we will follow a procedure that is a routine approach to the solution of such equations. Since this potential will only have bound states, we know that the solutions must vanish at $\pm\infty$ so that the wave functions will be square integrable. The first step is to determine what this fall off at large $|x|$ looks like. Now examine (5.141). The potential term goes like $x^2\psi(x)$ while the energy term goes like $\psi(x)$. Therefore, for large values of x the potential term should dominate and the Schrödinger equation in this region is well approximated by

$$\frac{d^2}{dx^2}\psi(x) - \frac{m^2\omega^2 x^2}{\hbar^2}\psi(x) = 0\,. \tag{5.142}$$

Note that the second derivative of the wave function must result in some constant times x^2 times the original function. This suggests that the wave function should be an exponential of some power of x. We will then choose the trial function

$$\psi(x) = e^{-\alpha x^\beta}\,. \tag{5.143}$$

The derivatives of this function are

$$\frac{d\psi}{dx} = -\alpha\beta x^{\beta-1}e^{-\alpha x^\beta} \tag{5.144}$$

and

$$\frac{d^2\psi}{dx^2} = -\alpha\beta(\beta-1)x^{\beta-2}e^{-\alpha x^\beta} + \alpha^2\beta^2 x^{2\beta-2}e^{-\alpha x^\beta}\,. \tag{5.145}$$

We can drop the first term of the second derivative since we are only interested in large x and it grows more slowly than the second. So we will use

$$\frac{d^2\psi}{dx^2} \cong \alpha^2\beta^2 x^{2\beta-2}e^{-\alpha x^\beta}\,. \tag{5.146}$$

Substituting the trail function and its second derivative into (5.142) yields

$$\alpha^2\beta^2 x^{2\beta-2}e^{-\alpha x^\beta} - \frac{m^2\omega^2 x^2}{\hbar^2}e^{-\alpha x^\beta}\,. \tag{5.147}$$

The exponentials can be divided out and in order for the equation to be true for all values of x, the powers of x must be the same in both terms. This implies that

$$2\beta - 2 = 2 \tag{5.148}$$

or
$$\beta = 2 \,. \tag{5.149}$$

We can now cancel the powers of x and substitute to value of β to give

$$4\alpha^2 - \frac{m^2\omega^2}{\hbar^2} \tag{5.150}$$

or

$$\alpha = \frac{1}{2}\frac{m\omega}{\hbar} \,. \tag{5.151}$$

We have now determined that for large values of x the wave function behaves as

$$\psi(x) \sim e^{-\frac{1}{2}\frac{m\omega}{\hbar}x^2} \,. \tag{5.152}$$

It is convenient to express the differential equation in terms of dimensionless quantities. Since the argument of the exponential in (5.152) must be dimensionless, we can define the dimensionless variable

$$y = \sqrt{\frac{m\omega}{\hbar}}x \,. \tag{5.153}$$

Changing variables in (5.141) from x to y yields

$$\frac{m\omega}{\hbar}\frac{d^2}{dy^2}\psi(y) + \frac{2mE}{\hbar^2}\psi(y) - \frac{m^2\omega^2}{\hbar^2}\frac{\hbar}{m\omega}y^2\psi(y) = 0 \tag{5.154}$$

or

$$\frac{d^2}{dy^2}\psi(y) + \frac{2E}{\hbar\omega}\psi(y) - y^2\psi(y) = 0 \,. \tag{5.155}$$

If we now define the dimensionless variable

$$\varepsilon = \frac{2E}{\hbar\omega} \,, \tag{5.156}$$

this becomes

$$\frac{d^2}{dy^2}\psi(y) + \varepsilon\psi(y) - y^2\psi(y) = 0 \,. \tag{5.157}$$

The next step is to remove the asymptotic behavior of the wave function from the differential equation. To do this we write

$$\psi(y) = f(y)e^{-\frac{1}{2}y^2} \,. \tag{5.158}$$

The derivatives of (5.158) are

$$\frac{d\psi}{dy} = f'(y)e^{-\frac{1}{2}y^2} - yf(y)e^{-\frac{1}{2}y^2} \tag{5.159}$$

and

$$\frac{d^2\psi}{dy^2} = f''(y)e^{-\frac{1}{2}y^2} - 2f'(y)ye^{-\frac{1}{2}y^2} - f(y)e^{-\frac{1}{2}y^2} + y^2 f(y)e^{-\frac{1}{2}y^2}. \tag{5.160}$$

Now, substituting (5.158) and (5.160) into (5.157) to give

$$f''(y)e^{-\frac{1}{2}y^2} - 2f'(y)ye^{-\frac{1}{2}y^2} - f(y)e^{-\frac{1}{2}y^2} + y^2 f(y)e^{-\frac{1}{2}y^2} + \varepsilon f(y)e^{-\frac{1}{2}y^2} - y^2 f(y)e^{-\frac{1}{2}y^2} = 0. \tag{5.161}$$

Dividing out the exponentials and simplifying give the differential equation for $f(y)$

$$f''(y) - 2yf'(y) + (\varepsilon - 1)f(y) = 0. \tag{5.162}$$

The final step is to see if we can find a solution of (5.162) in terms of the power series

$$f(y) = \sum_{i=0}^{\infty} a_i y^i. \tag{5.163}$$

The derivatives of this series are

$$f'(y) = \sum_{i=0}^{\infty} i a_i y^{i-1} = \sum_{i=1}^{\infty} i a_i y^{i-1} \tag{5.164}$$

and

$$f''(y) = \sum_{i=1}^{\infty} i(i-1)a_i y^{i-2} = \sum_{i=2}^{\infty} i(i-1)a_i y^{i-2}. \tag{5.165}$$

Substituting (5.164) and (5.165) into (5.162) gives

$$\sum_{i=2}^{\infty} i(i-1)a_i y^{i-2} - 2\sum_{i=1}^{\infty} i a_i y^i + (\varepsilon - 1)\sum_{i=0}^{\infty} a_i y^i = 0. \tag{5.166}$$

The a_i can be found by requiring the coefficients of like powers of y vanish. We can make this job easier if we change the sum in the first term by making the substitution $i \to i+2$ and changing the lower limit of the sum accordingly. Then

$$\sum_{i=0}^{\infty}(i+2)(i+1)a_{i+2}y^i - 2\sum_{i=0}^{\infty} i a_i y^i + (\varepsilon - 1)\sum_{i=0}^{\infty} a_i y^i = 0. \tag{5.167}$$

This can be rewritten as

$$\sum_{i=0}^{\infty} \left[(i+2)(i+1)a_{i+2} - [2i - \varepsilon + 1]\, a_i\right] y^i = 0. \tag{5.168}$$

This requires that

$$a_{i+2} = \frac{2i - \varepsilon + 1}{(i+2)(i+1)}a_i. \tag{5.169}$$

92

This is a recursion relation for generating the coefficients of the power series. Since this recursion relations connects $i + 2$ to i, a_0 and a_1 can be chosen independently. If we choose $a_1 = 0$ and a_0 to be finite, then the series will only have even powers of y. If $a_0 = 0$ and a_1 is finite, then the series contains only odd powers of y. Since the oscillator potential is even in x we expect the solutions to be either even or odd so the two situations above will generate solutions of definite parity, as required.

Note that

$$\frac{a_{i+2}}{a_i} = \frac{2i - \varepsilon + 1}{(i+2)(i+1)} \sim \frac{2}{i} \tag{5.170}$$

for large values of i. This indicates the infinite series cannot converge over the entire region $-\infty < x < \infty$. This means that the only way the series solution can be valid only if a finite number of terms have nonzero coefficients. This will happen if the recursion relation for the coefficients will cause one of the coefficients to vanish for some value of i. Then all higher coefficients will vanish and the series will truncate to a finite number of terms. Let n be the value of i for which the coefficient $a_{n+2} = 0$. This then requires that

$$2n - \varepsilon + 1 = 0 \tag{5.171}$$

or

$$\varepsilon = 2n + 1 \,. \tag{5.172}$$

Using (5.156), this gives

$$\frac{2E}{\hbar\omega} = 2n + 1 \tag{5.173}$$

or

$$E_n = \hbar\omega \left(n + \frac{1}{2} \right) \tag{5.174}$$

where $n = 0, 1, 2, \ldots$. This differs from the spectrum for the Bohr quantized oscillator by having $n + \frac{1}{2}$ rather than n.

The solutions for f(y) are now polynomials of order n. If we make the choice that

$$a_0 = (-1)^{\frac{n}{2}} \frac{n!}{\frac{n}{2}!} \tag{5.175}$$

for even n and

$$a_1 = 2n(-1)^{\frac{n-1}{2}} \frac{(n-1)!}{\frac{n-1}{2}!} \tag{5.176}$$

for odd n, then $f_n(y) = H_n(y)$ where $H_n(y)$ are the *Hermite polynomials*. The first

few of these polynomials are

$$
\begin{align}
H_0(x) &= 1 \tag{5.177}\\
H_1(x) &= 2x \tag{5.178}\\
H_2(x) &= 4x^2 - 2 \tag{5.179}\\
H_3(x) &= 8x^3 - 12x \tag{5.180}\\
H_4(x) &= 16x^4 - 48x^2 + 12 \tag{5.181}\\
H_5(x) &= 32x^5 - 160x^3 + 120x \tag{5.182}\\
H_6(x) &= 64x^6 - 480x^4 + 720x^2 - 120 \tag{5.183}
\end{align}
$$

Hermite polynomials for larger values of n can be constructed using the recurrence relation

$$
H_{n+1}(x) = 2xH_n(x) - 2nH_{n-1}(x) \tag{5.184}
$$

and the first derivatives of the Hermite polynomials are given by

$$
H'_n(x) = 2nH_{n-1}(x) . \tag{5.185}
$$

The orthogonality relation for the Hermite polynomials is

$$
\int_{-\infty}^{\infty} dx\, e^{-x^2} H_m(x) H_n(x) = 2^n \pi^{\frac{1}{2}} n! \delta_{mn} . \tag{5.186}
$$

We can now write the eigenfunctions for the one-dimensional harmonic oscillator as

$$
\psi_n(x) = N_n H_n \left(\sqrt{\frac{m\omega}{\hbar}} x \right) e^{-\frac{1}{2}\frac{m\omega}{\hbar}x^2} . \tag{5.187}
$$

The normalization constant N_n can be determined from

$$
\begin{align}
1 &= \int_{-\infty}^{\infty} dx\, |\psi_n(x)|^2 = N_n^2 \int_{-\infty}^{\infty} dx\, H_n^2 \left(\sqrt{\frac{m\omega}{\hbar}} x \right) e^{-\frac{m\omega}{\hbar}x^2} \\
&= N_n^2 \sqrt{\frac{\hbar}{m\omega}} \int_{-\infty}^{\infty} dy\, H_n^2(y)\, e^{-y^2} = N_n^2 \sqrt{\frac{\hbar}{m\omega}} 2^n \pi^{\frac{1}{2}} n! . \tag{5.188}
\end{align}
$$

So,

$$
N_n = \left(\frac{m\omega}{\hbar\pi} \right)^{\frac{1}{4}} \frac{1}{\sqrt{2^n n!}} . \tag{5.189}
$$

The eigenfunctions are then given by

$$
\psi_n(x) = \left(\frac{m\omega}{\hbar\pi} \right)^{\frac{1}{4}} \frac{1}{\sqrt{2^n n!}} H_n \left(\sqrt{\frac{m\omega}{\hbar}} x \right) e^{-\frac{1}{2}\frac{m\omega}{\hbar}x^2} . \tag{5.190}
$$

The first four wave functions for the one-dimensional harmonic oscillator are shown in Fig. 5.7
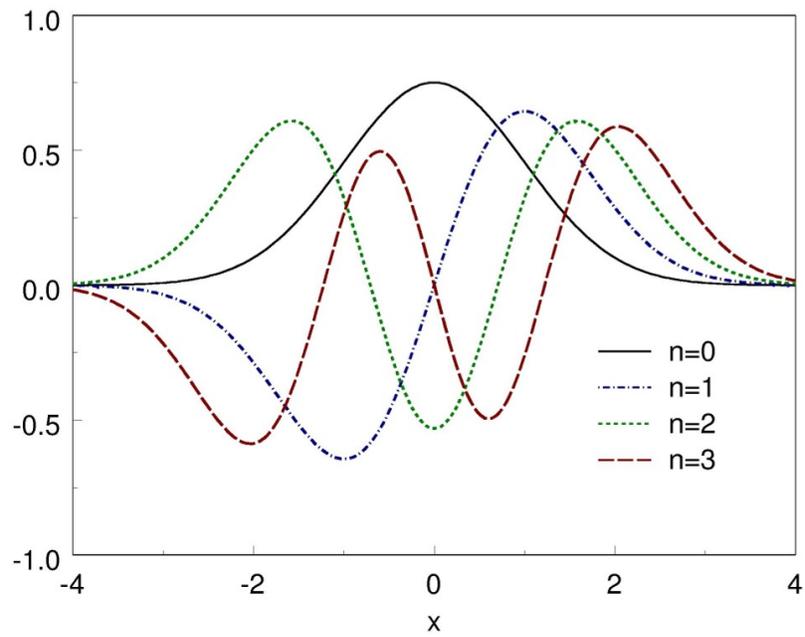
Figure 5.7: One-dimensional harmonic oscillator wave functions for $n = 0, 1, 2, 3$ with the choice of parameters such that $\frac{m\omega}{\hbar} = 1$.

Chapter 6

Scattering in One Dimension

6.1 The Free-Particle Schrödinger Equation

For a free particle the time-independent Schrödinger equation is

$$-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}\psi(x) = E\psi(x)\,. \tag{6.1}$$

Proceeding as we did for the infinite square well, this can be rewritten as

$$\frac{\partial^2}{\partial x^2}\psi(x) + k^2\psi(x) = 0 \tag{6.2}$$

where

$$k^2 = \frac{2mE}{\hbar^2}\,, \tag{6.3}$$

or

$$E = \frac{\hbar^2 k^2}{2m}\,. \tag{6.4}$$

The solutions to this equation are

$$\psi_k(x) = N e^{ikx}\,. \tag{6.5}$$

This solutions is defined on the interval $-\infty < x < \infty$ and does not satisfy the boundary conditions $\psi(\pm\infty) = 0$. The eigenvalue k is therefore continuous and can take on any value in the range $-\infty < k < \infty$. This can be seen from the square well solution. Previously, we showed that for the infinite square well,

$$k_n = \frac{n\pi}{a}\,. \tag{6.6}$$

The spacing between two adjacent states is

$$\Delta k = k_{n+1} - k_n = \frac{\pi}{a}\,. \tag{6.7}$$

Therefore,

$$\lim_{a \to \infty} \Delta k = 0 \,. \tag{6.8}$$

So, this shows that k becomes continuous for infinite intervals.

If we operator on this wave function with the momentum operator, then

$$\hat{p}\psi_k(x) = \frac{\hbar}{i}\frac{\partial}{\partial x}Ne^{ikx} = \frac{\hbar}{i}ikNe^{ikx} = \hbar k\psi_k(x) \,. \tag{6.9}$$

Therefore, this wave function is an eigenfunction of of the momentum operator. If we define the momentum eigenvalue as p, then $p = \hbar k$ and

$$\hat{p}\psi_k(x) = p\psi_k(x) \,. \tag{6.10}$$

This means that we can label the plane wave state by the momentum eigenvalue. Therefore

$$\psi_p(x) = N_p e^{\frac{i}{\hbar}px} \tag{6.11}$$

and

$$E_p = \frac{p^2}{2m} \,. \tag{6.12}$$

The problem that arises here is that this plane-wave wave function is not localized with the result that the wave function is not square integrable. To understand the consequences of this, it is useful to examine the probability density

$$\rho(x) = \psi_p^*(x)\psi_p(x) = N_p^* e^{-\frac{i}{\hbar}px}N_p e^{\frac{i}{\hbar}px} = |N_p|^2 \tag{6.13}$$

and the probability current density

$$
\begin{aligned}
j(x) &= \frac{\hbar}{2im}\left[\psi_p^*(x)\frac{\partial}{\partial x}\psi_p(x) - \left(\frac{\partial}{\partial x}\psi_p^*(x)\right)\psi_p(x)\right] \\
&= |N_p|^2\frac{\hbar}{2im}\left[e^{-\frac{i}{\hbar}px}\frac{\partial}{\partial x}e^{\frac{i}{\hbar}px} - \left(\frac{\partial}{\partial x}e^{-\frac{i}{\hbar}px}\right)e^{\frac{i}{\hbar}px}\right] \\
&= |N_p|^2\frac{\hbar}{2im}\left[e^{-\frac{i}{\hbar}px}\frac{i}{\hbar}pe^{\frac{i}{\hbar}px} - \left(-\frac{i}{\hbar}pe^{-\frac{i}{\hbar}px}\right)e^{\frac{i}{\hbar}px}\right] \\
&= |N_p|^2\frac{p}{m} \,. \tag{6.14}
\end{aligned}
$$

Note that both $\rho(x)$ and $j(x)$ are independent of x, so integration of either of these quantities over $-\infty < x < \infty$ will be infinite. This indicates that the probability can not be normalized. The current density indicates that there is a constant flow of probability from $x = -\infty$ to $x = \infty$ if $p > 0$. Therefore probability must be created at $-\infty$ and destroyed at ∞. This type of solution, which is not normalizable and has a continuous spectrum, does not constitute a hilbert space in the sense that we have defined previously. Great care then has to be taken in using these continuum solutions for calculating quantum mechanical processes.

One approach to addressing this problem is the *box normalization* procedure where all such processes treated as if they are contained inside a large box defined by $-\frac{a}{2} \leq x \leq \frac{a}{2}$ where a will be taken to be large compared to the size of the system that we are considering. This is a reasonable approach since quantum mechanical processes typically take place over small distances and the imposition of boundary conditions at distances much larger than the size of the interaction region should not have a large effect on the observables. All observables are the calculated in the box and then the limit $a \to \infty$ is taken. To implement this we will assume that the boundary conditions at the walls of the box are periodic. That is

$$\psi\left(-\frac{a}{2}\right) = \psi\left(\frac{a}{2}\right) . \tag{6.15}$$

This boundary condition allows for the propagation of waves across the interval without generating reflections at the boundaries as occurs with the boundary conditions for the infinite square well. Imposing this condition on the plane wave gives

$$N_k e^{-ik\frac{a}{2}} = N_k e^{ik\frac{a}{2}} . \tag{6.16}$$

This will be satisfied if

$$k\frac{a}{2} = -k\frac{a}{2} + 2\pi n \tag{6.17}$$

where n is any positive or negative integer. This can be rewritten as

$$ka = 2\pi n \tag{6.18}$$

or

$$k_n = \frac{2\pi n}{a} . \tag{6.19}$$

The corresponding momentum is then

$$p_n = \hbar k_n = \frac{2\pi \hbar n}{a} , \tag{6.20}$$

and the eigenenergy is

$$E_n = \frac{p_n^2}{2m} = \frac{4\pi^2 \hbar^2 n^2}{2ma^2} . \tag{6.21}$$

The eigenstates can now be written as

$$\psi_n(x) = N_n e^{\frac{i}{\hbar} p_n x} . \tag{6.22}$$

The normalization condition for these wave functions is

$$1 = \int_{-\frac{a}{2}}^{\frac{a}{2}} \psi_n^*(x)\psi_n(x) = |N_n|^2 \int_{-\frac{a}{2}}^{\frac{a}{2}} e^{-i\frac{2\pi nx}{a}} e^{i\frac{2\pi nx}{a}} = |N_n|^2 a . \tag{6.23}$$

The wave functions are then normalized by choosing

$$N_n = \frac{1}{\sqrt{a}} \,.$$

(6.24)

The normalized wave function is then

$$\psi_n(x) = \frac{1}{\sqrt{a}} e^{\frac{i}{\hbar} p_n x} \,.$$

(6.25)

The completeness relation requires that

$$\delta(x - x') = \sum_{n=-\infty}^{\infty} \psi_n(x) \psi_n^*(x') = \frac{1}{a} \sum_{n=-\infty}^{\infty} e^{\frac{i}{\hbar} p_n x} e^{-\frac{i}{\hbar} p_n x} = \frac{1}{a} \sum_{n=-\infty}^{\infty} e^{\frac{i}{\hbar} p_n (x - x')} \,.$$

(6.26)

The generalized solution to the time-dependent Schrödinger equation is

$$\Psi(x, t) = \sum_{n=-\infty}^{\infty} c_n \psi_n(x) e^{-\frac{i}{\hbar} E_n t} = \frac{1}{\sqrt{a}} \sum_{n=-\infty}^{\infty} c_n e^{\frac{i}{\hbar}(p_n x - E_n t)} \,,$$

(6.27)

where

$$c_n = \int_{-a/2}^{a/2} dx \, \psi_n^*(x) \Psi(x, 0)$$

(6.28)

and these coefficients must satisfy the normalization condition

$$\sum_{n=-\infty}^{\infty} |c_n|^2 = 1 \,.$$

(6.29)

We are interested in the case where $a \to \infty$. In this limit we have shown that the spectrum becomes continuous. This limit can be performed by carefully rewriting the infinite sums of the discreet eigenstates. The first step is to note that the difference between any to adjacent eigenmomenta is

$$\Delta p_n = p_n - p_{n-1} = \frac{2\pi\hbar}{a}$$

(6.30)

or

$$\frac{a}{2\pi\hbar} \Delta p_n = 1 \,.$$

(6.31)

Now consider the completeness relation (6.26). Using (6.31), this can be rewritten as

$$\delta(x - x') = \frac{1}{2\pi\hbar} \sum_{n=-\infty}^{\infty} \Delta p_n e^{\frac{i}{\hbar} p_n (x - x')} \,.$$

(6.32)

This looks like a Riemann sum and in the limit $a \to \infty$ where $\Delta p_n \to 0$ this becomes an integral. So,

$$\delta(x - x') = \frac{1}{2\pi\hbar} \int_{-\infty}^{\infty} dp \, e^{\frac{i}{\hbar} p(x-x')} = \frac{1}{2\pi\hbar} 2\pi\delta\left(\frac{i}{\hbar}(x - x')\right) = \delta(x - x') . \tag{6.33}$$

So the completeness relation is satisfied in this limit.

Next consider the normalization constraint on the coefficents (6.29). This can be rewritten as

$$1 = \frac{a}{2\pi\hbar} \sum_{n=-\infty}^{\infty} \Delta p_n |c_n|^2 = \frac{a}{2\pi\hbar} \sum_{n=-\infty}^{\infty} \Delta p_n |c(p_n)|^2 , \tag{6.34}$$

where we have redefined the notation of the coefficient by noting that it can be written as a function of the eigenmomentum. Again, as a becomes large the sum can be approximated by an integral as

$$1 = \frac{a}{2\pi\hbar} \int_{-\infty}^{\infty} dp |c(p)|^2 . \tag{6.35}$$

This condition can be retained in the limit $a \to \infty$ only if $c(p) \sim 1/\sqrt{a}$. Assuming that this is indeed the case, we can define a new function

$$\phi(p) \equiv \sqrt{\frac{a}{2\pi\hbar}} \, c(p) . \tag{6.36}$$

The normalization condition then becomes

$$1 = \int_{-\infty}^{\infty} dp |\phi(p)|^2 . \tag{6.37}$$

Finally, consider the general solution to the time-dependent Schrödinger equation (6.27). This can be rewritten as

$$\Psi(x, t) = \frac{1}{\sqrt{a}} \frac{a}{2\pi\hbar} \sum_{n=-\infty}^{\infty} \Delta p_n \, c_n e^{\frac{i}{\hbar}(p_n x - E_n t)} = \frac{\sqrt{a}}{2\pi\hbar} \sum_{n=-\infty}^{\infty} \Delta p_n \, c(p_n) e^{\frac{i}{\hbar}(p_n x - E(p_n)t)} . \tag{6.38}$$

For large a, this can be written as

$$\Psi(x, t) = \frac{\sqrt{a}}{2\pi\hbar} \int_{-\infty}^{\infty} dp \, c(p) e^{\frac{i}{\hbar}(px - E(p)t)} . \tag{6.39}$$

Using (6.36), this becomes

$$\Psi(x, t) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} dp \, \phi(p) e^{\frac{i}{\hbar}(px - E(p)t)} . \tag{6.40}$$

This is a wave packet. This means that one way of imposing the box normalization condition is to calculate all scattering processes using wave packets. This can be done, but is often rather tedious. However, it is instructive to consider the behavior of these wave packets in one-dimension.

6.2 The Free Wave Packet

Before looking at scattering of wave packets, it is useful to first consider the motion of a free Schrödinger wave packet

$$\Psi(x,t) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} dp\, \phi(p) e^{\frac{i}{\hbar}\left(px - \frac{p^2}{2m}t\right)}. \tag{6.41}$$

The corresponding momentum-space wave function is given by

$$
\begin{aligned}
\Psi(p,t) &= \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} dx\, e^{-\frac{i}{\hbar}px} \Psi(x,t) = \frac{1}{2\pi\hbar} \int_{-\infty}^{\infty} dx\, e^{-\frac{i}{\hbar}px} \int_{-\infty}^{\infty} dp'\, \phi(p') e^{\frac{i}{\hbar}\left(p'x - \frac{p'^2}{2m}t\right)} \\
&= \frac{1}{2\pi\hbar} \int_{-\infty}^{\infty} dp'\, \phi(p') e^{-\frac{i}{\hbar}\frac{p'^2}{2m}t} \int_{-\infty}^{\infty} dx\, e^{\frac{i}{\hbar}(p'-p)x} = \int_{-\infty}^{\infty} dp'\, \phi(p') e^{-\frac{i}{\hbar}\frac{p'^2}{2m}t} \delta(p'-p) \\
&= \phi(p) e^{-\frac{i}{\hbar}\frac{p^2}{2m}t}.
\end{aligned}
\tag{6.42}
$$

The average momentum can now be calculated as

$$\langle p(t) \rangle = \int_{-\infty}^{\infty} dp\, \Psi^*(p,t) p \Psi(p,t) = \int_{-\infty}^{\infty} dp\, p\, |\phi(p)|^2 \tag{6.43}$$

and, similarly, the mean-square momentum is given by

$$\langle p^2(t) \rangle = \int_{-\infty}^{\infty} dp\, p^2\, |\phi(p)|^2. \tag{6.44}$$

The average position can be calculated in momentum space as

$$\langle x(t) \rangle = \int_{-\infty}^{\infty} dp\, \Psi^*(p,t) \left(-\frac{\hbar}{i}\frac{\partial}{\partial p}\right) \Psi(p,t) = i\hbar \int_{-\infty}^{\infty} dp\, \Psi^*(p,t) \frac{\partial}{\partial p} \Psi(p,t) \tag{6.45}$$

and the mean-square position as

$$\langle x^2(t) \rangle = -\hbar^2 \int_{-\infty}^{\infty} dp\, \Psi^*(p,t) \frac{\partial^2}{\partial p^2} \Psi(p,t). \tag{6.46}$$

Now consider the special case where

$$\phi(p) = \frac{1}{\eta^{\frac{1}{2}}\pi^{\frac{1}{4}}} e^{-\frac{(p-p_0)^2}{2\eta^2}} e^{-\frac{i}{\hbar}px_0}, \tag{6.47}$$

where η, p_0 and x_0 are constants. In this case, the average momentum is

$$\langle p(t) \rangle = \frac{1}{\eta\pi^{\frac{1}{2}}} \int_{-\infty}^{\infty} dp\, p\, e^{-\frac{(p-p_0)^2}{\eta^2}} = \frac{1}{\pi^{\frac{1}{2}}} \int_{-\infty}^{\infty} dy\, (\eta y + p_0)\, e^{-y^2} = p_0. \tag{6.48}$$

This wave packet is, therefore, moving with a constant average momentum p_0. The mean-square momentum is

$$\langle p^2(t)\rangle = \frac{1}{\eta\pi^{\frac{1}{2}}}\int_{-\infty}^{\infty}dp\,p^2\,e^{-\frac{(p-p_0)^2}{\eta^2}} = \frac{1}{\pi^{\frac{1}{2}}}\int_{-\infty}^{\infty}dy(\eta y+p_0)^2\,e^{-y^2} = \frac{1}{2}\eta^2-p_0^2\,. \quad (6.49)$$

The root-mean-square deviation in momentum is then

$$\Delta p = \sqrt{\langle p^2(t)\rangle - \langle p(t)\rangle^2} = \frac{\eta}{\sqrt{2}}\,. \quad (6.50)$$

The width of the wave packet in momentum space is, therefore, constant and proportional to the parameter η.

To calculate the average position and mean-square position, we need the derivatives of the momentum space wave function

$$\Psi(p,t) = \frac{1}{\eta^{\frac{1}{2}}\pi^{\frac{1}{4}}}e^{-\frac{(p-p_0)^2}{2\eta^2}}e^{-\frac{i}{\hbar}px_0}e^{\frac{i}{\hbar}\frac{p^2}{2m}t}\,. \quad (6.51)$$

The first derivative of this wave function is

$$\frac{\partial}{\partial p}\Psi(p,t) = -\frac{1}{\eta^{\frac{1}{2}}\pi^{\frac{1}{4}}}\left[\frac{p-p_0}{\eta^2}+\frac{i}{\hbar}\left(x_0+\frac{p}{m}t\right)\right]e^{-\frac{(p-p_0)^2}{2\eta^2}}e^{-\frac{i}{\hbar}px_0}e^{\frac{i}{\hbar}\frac{p^2}{2m}t} \quad (6.52)$$

and the second derivative is

$$\frac{\partial^2}{\partial p^2}\Psi(p,t) = \frac{1}{\eta^{\frac{1}{2}}\pi^{\frac{1}{4}}}\left\{\left[\frac{p-p_0}{\eta^2}+\frac{i}{\hbar}\left(x_0+\frac{p}{m}t\right)\right]^2-\frac{1}{\eta^2}-\frac{it}{\hbar m}\right\}e^{-\frac{(p-p_0)^2}{2\eta^2}}e^{-\frac{i}{\hbar}px_0}e^{\frac{i}{\hbar}\frac{p^2}{2m}t}\,. \quad (6.53)$$

Using these we can evaluate the average position to be

$$\begin{aligned}
\langle x(t)\rangle &= -\frac{i\hbar}{\eta\pi^{\frac{1}{2}}}\int_{-\infty}^{\infty}dp\,e^{-\frac{(p-p_0)^2}{\eta^2}}\left[\frac{p-p_0}{\eta^2}+\frac{i}{\hbar}\left(x_0+\frac{p}{m}t\right)\right]\\
&= -\frac{i\hbar}{\pi^{\frac{1}{2}}}\int_{-\infty}^{\infty}dy\,e^{-y^2}\left[\frac{y}{\eta}+\frac{i}{\hbar}\left(x_0+\frac{p_0}{m}t+\frac{\eta y}{m}t\right)\right]\\
&= x_0+\frac{p_0}{m}t\,. \quad (6.54)
\end{aligned}$$

This means that the average position of the wave packet starts at x_0 at $t=0$ and then moves with the constant velocity $v_0 = p_0/m$ for subsequent times.

The mean-square position is

$$\begin{aligned}
\langle x^2(t)\rangle &= -\frac{\hbar^2}{\eta\pi^{\frac{1}{2}}}\int_{-\infty}^{\infty}dp\,e^{-\frac{(p-p_0)^2}{\eta^2}}\left\{\left[\frac{p-p_0}{\eta^2}+\frac{i}{\hbar}\left(x_0+\frac{p}{m}t\right)\right]^2-\frac{1}{\eta^2}-\frac{it}{\hbar m}\right\}\\
&= -\frac{\hbar^2}{\pi^{\frac{1}{2}}}\int_{-\infty}^{\infty}dy\,e^{-y^2}\left\{\left[\frac{y}{\eta^2}+\frac{i}{\hbar}\left(x_0+\frac{p_0}{m}t+\frac{\eta y}{m}t\right)\right]^2-\frac{1}{\eta^2}-\frac{it}{\hbar m}\right\}\\
&= \frac{\eta^2 t^2}{2m^2}+\frac{\hbar^2}{2\eta^2}+\left(x_0+\frac{p_0}{m}t\right)^2\,. \quad (6.55)
\end{aligned}$$

The root-mean-square deviation in the position is then

$$\Delta x = \sqrt{\langle x^2(t)\rangle - \langle x(t)\rangle^2} = \sqrt{\frac{\eta^2 t^2}{2m^2} + \frac{\hbar^2}{2\eta^2}} = \Delta x_0 \sqrt{1 + \frac{t^2}{\tau^2}}, \qquad (6.56)$$

where

$$\Delta x_0 = \frac{\hbar}{\sqrt{2}\eta} \qquad (6.57)$$

and

$$\tau = \frac{\hbar m}{\eta^2}. \qquad (6.58)$$

This means that the wave packet starts with an initial width of Δx_0 and then increases with time at a rate determined by τ. This spreading of the wave packet with time is characteristic of the Schrödinger equation.

The origin of the spreading is straightforward. Consider a plane wave of the form

$$e^{\frac{i}{\hbar}(px - E(p)t)}. \qquad (6.59)$$

We can find the position of a surface of constant phase by requiring that

$$px - E(p)t = \phi_0. \qquad (6.60)$$

This can be solved to give

$$x = \frac{\phi_0}{p} + \frac{E(p)}{p}t. \qquad (6.61)$$

A surface of constant phase then propagates with the velocity

$$v_p = \frac{E(p)}{p}, \qquad (6.62)$$

which is called the *phase velocity*. In the case of the Schrödinger equation the phase velocity is

$$v_p = \frac{p}{2m}. \qquad (6.63)$$

Now consider the wave packet

$$\Psi(x,t) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} dp\, \phi(p) e^{\frac{i}{\hbar}(px - E(p)t)}. \qquad (6.64)$$

If $\phi(p)$ is sharply peaked about some momentum p_0, then the wave packet will receive appreciable contributions from momenta close to this value and we can approximate the expression for the wave packet by expanding the phase about p_0. That is

$$px - E(p)t \cong px - E(p_0)t - E'(p_0)t(p - p_0) = p(x - E'(p_0)t) - E(p_0)t - E'(p_0)p_0 t \quad (6.65)$$

104

The wave packet is then approximately given by

$$\Psi(x,t) = \frac{e^{-\frac{i}{\hbar}(E(p_0)+p_0 E'(p_0))t}}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} dp\, \phi(p) e^{\frac{i}{\hbar}p(x-E'(p_0)t)} . \tag{6.66}$$

This implies that the motion of the wave packet as a whole is given by the *group velocity*

$$v_g = E'(p_0) . \tag{6.67}$$

For the Schrödinger equation the group velocity is

$$v_g = \frac{p_0}{m} , \tag{6.68}$$

which is the same as the velocity of the average position that we found for our particular example of $\phi(p)$. The spreading of the Schrödinger wave packet occurs because each plane-wave component of the wave packet has a different phase velocity that depends on the momentum of that component. Since some components are slower than the group velocity of the packet and some are greater, the packet will tend to spread as a result of the accumulated differences in phase among the various components.

The rate of spreading of the packet can be decreased by making τ larger. Since

$$\tau = \frac{\hbar m}{\eta^2} , \tag{6.69}$$

this means making η to be smaller. But since,

$$\Delta x_0 = \frac{\hbar}{\sqrt{2}\eta} , \tag{6.70}$$

this increases the width of the wave packet. Conversely, decreasing the initial width of the wave packet will increase the rate with which the packet spreads.

6.3   The Step Potential

Consider the situation where a wave packet initially centered at some $x_0 < 0$ and moving to the right with average momentum p_0 moves in a step-function potential

$$V(x) = \begin{cases} 0 & \text{for } x < 0 \\ V_0 & \text{for } x > 0 \end{cases} . \tag{6.71}$$

If $|x_0|$ is sufficiently large compared to the width of the packet, the wave function will initially be localized in a region where there is no potential. As the packet moves to the right, it will start to overlap with the region of finite potential and we will expect to see some part of the wave function to be reflected back into the negative

x region (Region I) and some part of the wave function may be transmitted through the barrier and propagate to the right in the positive x region (Region II). The wave function in Region I will, therefore, have two contributions: the initial wave packet moving to the right and a reflected wave packet moving to the left. We can then write the wave function in this region as

$$\Psi_I(x,t) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} dp\, \phi(p) e^{\frac{i}{\hbar}\left(px - \frac{p^2}{2m}t\right)} + \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} dp\, \phi_-(p) e^{-\frac{i}{\hbar}\left(px + \frac{p^2}{2m}t\right)}. \quad (6.72)$$

Note that the momentum distribution $\phi_-(p)$ is assumed to be different from that of the initial wave packet.

The time-independent wave function in Region II must be a solution to the Schrödinger equation

$$\frac{d^2}{dx^2}\psi(x) + \frac{2m}{\hbar^2}\left[E - V_0\right]\psi(x) = 0. \quad (6.73)$$

This has plane wave solutions of the form

$$\psi(x) = N e^{\frac{i}{\hbar}qx}, \quad (6.74)$$

where

$$q = \pm\sqrt{2m(E - V_0)}. \quad (6.75)$$

Since the energy eigenvalue of the plane-wave solutions must be independent of the region,

$$E = \frac{p^2}{2m}. \quad (6.76)$$

We can, therefore, write

$$q(p) = \pm\sqrt{p^2 - 2mV_0}. \quad (6.77)$$

The time-dependent wave function in Region II can then be expressed as

$$\Psi_{II}(x,t) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} dp\, \phi_+(p) e^{\frac{i}{\hbar}\left(q(p)x - \frac{p^2}{2m}t\right)}. \quad (6.78)$$

The boundary conditions at the discontinuity in the potential are

$$\Psi_I(0,t) = \Psi_{II}(0,t) \quad (6.79)$$

and

$$\left.\frac{\partial}{\partial x}\Psi_I(x,t)\right|_{x=0} = \left.\frac{\partial}{\partial x}\Psi_{II}(0,t)\right|_{x=0}. \quad (6.80)$$

The first boundary condition (6.79) requires that

$$\frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} dp\, (\phi(p) + \phi_-(p)) e^{-\frac{i}{\hbar}\frac{p^2}{2m}t} = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} dp\, \phi_+(p) e^{-\frac{i}{\hbar}\frac{p^2}{2m}t}, \quad (6.81)$$

106

or
$$\phi(p) + \phi_-(p) = \phi_+(p) \tag{6.82}$$

The second boundary condition (6.80) requires that

$$\frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} dp \, \frac{i}{\hbar} p(\phi(p) - \phi_-(p)) e^{-\frac{i}{\hbar}\frac{p^2}{2m}t} = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} dp \, \frac{i}{\hbar} q(p)\phi_+(p) e^{-\frac{i}{\hbar}\frac{p^2}{2m}t}, \tag{6.83}$$

or

$$p(\phi(p) - \phi_-(p)) = q(p)\phi_+(p). \tag{6.84}$$

This can be rewritten as

$$\phi(p) - \phi_-(p) = \frac{q(p)}{p}\phi_+(p). \tag{6.85}$$

Adding (6.82) and (6.85) yields

$$2\phi(p) = \left(1 + \frac{q(p)}{p}\right)\phi_+(p). \tag{6.86}$$

which can be solved for $\phi_+(p)$ to give

$$\phi_+(p) = \frac{2}{1 + \frac{q(p)}{p}}\phi(p) = \frac{2p}{p + q(p)}\phi(p). \tag{6.87}$$

Using (6.82) and (6.87) we can solve for

$$\phi_-(p) = \phi_+(p) - \phi(p) = \frac{2p}{p + q(p)}\phi(p) - \phi(p) = \frac{p - q(p)}{p + q(p)}\phi(p). \tag{6.88}$$

The wave functions in the two regions can now be written as

$$\Psi_I(x,t) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} dp \, \phi(p) \left[e^{\frac{i}{\hbar}\left(px - \frac{p^2}{2m}t\right)} + \frac{p - q(p)}{p + q(p)} e^{-\frac{i}{\hbar}\left(px + \frac{p^2}{2m}t\right)} \right]. \tag{6.89}$$

and

$$\Psi_{II}(x,t) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} dp \, \phi(p)\frac{2p}{p + q(p)} e^{\frac{i}{\hbar}\left(q(p)x - \frac{p^2}{2m}t\right)}. \tag{6.90}$$

The actual evaluation of these expressions can be complicated. To simplify our discussion, let us assume that the width in p of $\phi(p)$ is small compared to the average momentum $p_0 > 0$. This means that the contributions to the wave packet will come almost exclusively from momenta $p > 0$. So we don't have to worry about the sign ambiguity in (6.77) and can take only the positive solution. For wave packet components where $p \geq \sqrt{2mV_0}$, $q(p)$ will be real and these components will give oscillating contributions to the wave packet in Region II. For components with momenta
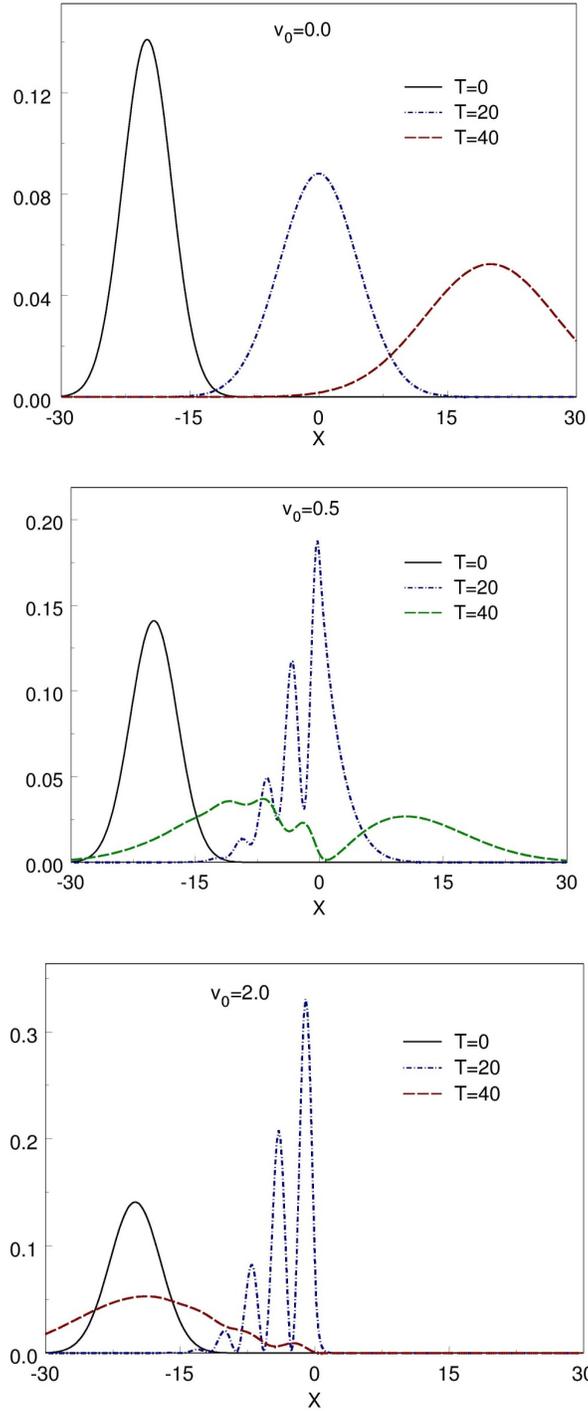
Figure 6.1: Plots of the probability density $\frac{\hbar}{p_0}\rho(x,t)$ as a function of $X = \frac{p_0 x}{\hbar}$ at dimensionless times $T = \frac{p_0^2 t}{2\hbar m} = 0$, 20 and 40 for $v_0 = \frac{mV_0}{p_0^2} = 0.0$, 0.5 and 2.0. The paramaters of the calculations are chosen such that $\frac{\eta}{p_0} = 0.25$ and $X_0 = -20$.
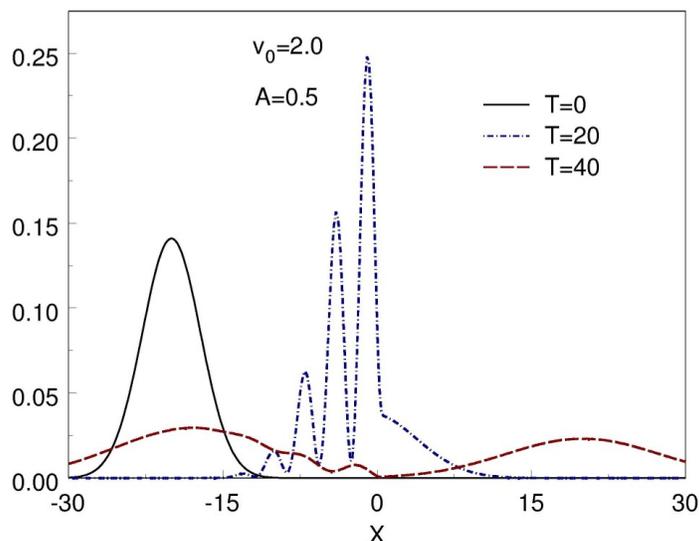
Figure 6.2:

$p < \sqrt{2mV_0}$, $q(p) = i\sqrt{2mV_0 - p^2}$ and the corresponding contributions to the wave function in Region II will be exponentially damped.

Figure 6.1 shows the probability density $\rho(x,t)$ calculated using the momentum distribution (6.47) for three choices of the potential. In the top figure the potential is chosen to be $V_0 = 0$. Since there is no barrier in this case, the wave packet is free and there is no reflected wave. As discussed above, we see the spreading of the free wave packet as a function of time. The second panel shows the probability density for $V_0 = 0.5\frac{p_0^2}{m}$. Here a substantial fraction of the momentum distribution is form momenta above the damping threshold $p = \sqrt{2mV_0}$. As a result, a substantial fraction of the probability is transmitted through the barrier. The final panel is for $V_0 = 2\frac{p_0^2}{m}$. Here virtually none of the momentum distribution is for momenta above the threshold value and, as a result, no appreciable probability is transmitted through the barrier and the wave is completely reflected. Note, however, that at $t = 20\frac{2m\hbar}{p_0^2}$ where the wave packet is interacting with the barrier, the probability density penetrates into the region just above $x = 0$. This gives rise to one of the more interesting implications of quantum mechanics. If the potential is chosen to be

$$
V(x) = \begin{cases} 0 & \text{for } x < 0 \\ V_0 & \text{for } 0 \leq x \leq a \\ 0 & \text{for } a < x \end{cases} .
$$

(6.91)

The width of the barrier can be chosen such that the wave can tunnel through the barrier. This is shown in Fig. 6.2 for $a = 0.5\frac{\hbar}{p_0}$.

109

Chapter 7

The Solution of the Schrödinger Equation in Three Dimensions

Consider the three-dimensional time-idependent Schrödinger equation

$$\left(-\frac{\hbar^2}{2m}\nabla^2 + V(\boldsymbol{r}) \right) \psi(\boldsymbol{r}) = E\psi(\boldsymbol{r}) \,. \tag{7.1}$$

This is a three-dimensional partial differential equation and the approach to the solution of this equation depends very much on the character of the potential.

A simple example of this is to consider the three-dimensional harmonic oscillator equation with the potential

$$V_{\text{HO}}(\boldsymbol{r}) = \frac{1}{2}m\omega^2\boldsymbol{r}^2 \,. \tag{7.2}$$

The Schrödinger equation for this potential can be written in Cartesian coordinates as

$$\left[-\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) + \frac{1}{2}m\omega^2\left(x^2 + y^2 + z^2 \right) \right]\psi(x,y,z) = E\psi(x,y,z) \,. \tag{7.3}$$

This equation can be solved using separation of variables by defining

$$\psi(x,y,z) = X(x)Y(y)Z(z) \,. \tag{7.4}$$

Substituting this into the Schrödinger equation and dividing by the wave function gives

$$-\frac{\hbar^2}{2mX(x)}\frac{\partial^2 X(x)}{\partial x^2} + \frac{1}{2}m\omega^2 x^2 - \frac{\hbar^2}{2mY(y)}\frac{\partial^2 Y(y)}{\partial y^2} + \frac{1}{2}m\omega^2 y^2$$
$$-\frac{\hbar^2}{2mZ(z)}\frac{\partial^2 Z(z)}{\partial z^2} + \frac{1}{2}m\omega^2 z^2 = E \,. \tag{7.5}$$

This clearly results in a separation of the three coordinates. If this equation is to be true for all possible values of x, y and z, it is necessary that each separated term be constant. This allows us, for example, to write

$$-\frac{\hbar^2}{2mX_{n_x}(x)}\frac{\partial^2 X_{n_x}(x)}{\partial x^2} + \frac{1}{2}m\omega^2 x^2 = E_{n_x}, \tag{7.6}$$

or

$$-\frac{\hbar^2}{2m}\frac{\partial^2 X_{n_x}(x)}{\partial x^2} + \frac{1}{2}m\omega^2 x^2 X_{n_x}(x) = E_{n_x} X_{n_x}(x). \tag{7.7}$$

This is the one-dimensional harmonic oscillator equation which we have already solved. The solution is the wave function described by (5.190),

$$X_{n_x}(x) = \psi_{n_x}(x) \tag{7.8}$$

where $n_x = 0, 1, 2, \ldots$ and

$$E_{n_x} = \hbar\omega\left(n_x + \frac{1}{2}\right). \tag{7.9}$$

Similarly,

$$Y_{n_y}(y) = \psi_{n_y}(y) \tag{7.10}$$

and

$$Z_{n_z}(z) = \psi_{n_z}(z). \tag{7.11}$$

The eigenfunctions of the three-dimensional wave equation are given by

$$\psi_{n_x n_y n_z}(x, y, z) = \psi_{n_x}(x)\psi_{n_y}(y)\psi_{n_z}(z) \tag{7.12}$$

and the eigenenergies are given by

$$E_{n_x n_y n_z} = E_{n_x} + E_{n_y} + E_{n_z} = \hbar\omega\left(n_x + n_y + n_z + \frac{3}{2}\right). \tag{7.13}$$

Note that the states are now labeled by three quantum numbers n_x, n_y and n_z. This comes about because the hamiltonian can be written as a sum of three one-dimensional hamiltonians as

$$\hat{H} = \hat{H}_x + \hat{H}_y + \hat{H}_z \tag{7.14}$$

and these three one-dimensional hamiltonians commute with the total hamiltonian

$$[\hat{H}, \hat{H}_i] = 0 \tag{7.15}$$

and the one-dimensional hamiltonians are mutually communting

$$[\hat{H}_i, \hat{H}_j] = 0. \tag{7.16}$$
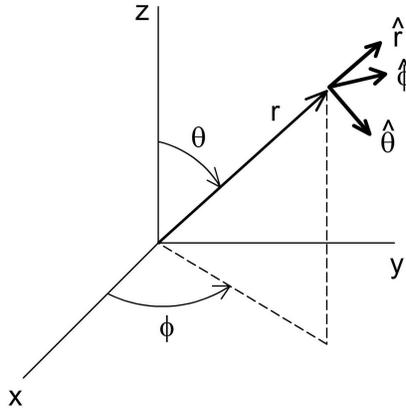
Figure 7.1: The spherical coordinate system.

As we will see, eigenstates are labeled by a set eigenvalues of operators which commute with the hamiltonian and are mutually commuting.

This problem can be solved in another way by noting that the harmonic oscillator potential is a function of the magnitude of the distance from the origin. That is

$$V_{HO}(\boldsymbol{r}) = V(|\boldsymbol{r}|) \,. \tag{7.17}$$

This is among a large class of common potential that are functions only of the distance from the origin. These are called *central potentials*. We will now consider the solution of the Schrödinger equation for this class of potentials.

Note that the potential is a function of the length of the coordinate vector $|\boldsymbol{r}|$ and is, therefore, spherically symmetrical. As a result it is also possible to separate the equation in spherical coordinates as shown in Fig. 7.1. Potentials that depend only on the magnitude of the coordinate vector are called central potentials and we will now consider solutions of the Schrödinger equation for potentials of this class.

7.1   The Schrödinger Equation with a Central Potential

We now consider the solution of the three-dimensional time-independent Schrödinger equation for a central potential

$$V(\boldsymbol{r}) = V(|\boldsymbol{r}|) = V(r) \,. \tag{7.18}$$

Since the potential depends only on the distance from the origin, the hamiltonian is spherically symmetric. It is, therefore, convenient to represent the Schrödinger equation in terms of the standard spherical coordinate system shown in Fig. 7.1. The laplacian in this coordinate system is given by

$$\nabla^2 = \frac{1}{r} \frac{\partial^2}{\partial r^2} r + \frac{1}{r^2 \sin\theta} \frac{\partial}{\partial \theta} \left(\sin\theta \frac{\partial}{\partial \theta} \right) + \frac{1}{r^2 \sin^2\theta} \frac{\partial^2}{\partial \phi^2} \, . \tag{7.19}$$

Therefore, the time-independent Schrödinger equation in spherical coordinates is

$$\left\{ -\frac{\hbar^2}{2m} \left[\frac{1}{r} \frac{\partial^2}{\partial r^2} r + \frac{1}{r^2 \sin\theta} \frac{\partial}{\partial \theta} \left(\sin\theta \frac{\partial}{\partial \theta} \right) + \frac{1}{r^2 \sin^2\theta} \frac{\partial^2}{\partial \phi^2} \right] + V(r) \right\} \psi(\boldsymbol{r}) = E\psi(\boldsymbol{r}) \, . \tag{7.20}$$

This can be simplified by treating it as a mathematical problem. We will proceed to do this and then will explore the physical consequences afterward.

The standard approach to solving partial differential equations is to use the method of separation of variables, which we have already used several times. We assume that the wave function is a product of functions each depending upon a single variable. Let

$$\psi(\boldsymbol{r}) = R(r)P(\theta)Q(\phi) \, . \tag{7.21}$$

Substituting this into (7.20), dividing both sides by $R(r)P(\theta)Q(\phi)$ and multiplying both sides by $\frac{2m}{\hbar^2} r^2 \sin^2\theta$ gives

$$\frac{r\sin^2\theta}{R(r)} \frac{\partial^2}{\partial r^2} rR(r) + \frac{\sin\theta}{P(\theta)} \frac{\partial}{\partial \theta} \left(\sin\theta \frac{\partial}{\partial \theta} P(\theta) \right) + \frac{2m}{\hbar^2} r^2 \sin^2\theta \, (E - V(r)) + \frac{1}{Q(\phi)} \frac{\partial^2}{\partial \phi^2} Q(\phi) = 0 \, . \tag{7.22}$$

Note that only the last term depends upon ϕ and that no other variable appears in this term. This equation can be true for all values of r, θ and ϕ only if the last term is equal to a constant. Let

$$\frac{1}{Q(\phi)} \frac{\partial^2}{\partial \phi^2} Q(\phi) = -m^2 \, . \tag{7.23}$$

This is can be rewritten as

$$\frac{\partial^2}{\partial \phi^2} Q(\phi) + m^2 Q(\phi) = 0 \, . \tag{7.24}$$

This is the familiar harmonic equation that has solutions of the form

$$Q(\phi) = e^{im\phi} \, . \tag{7.25}$$

Since the solution must satisfy $Q(\phi) = Q(\phi + 2\pi)$, the constant m must be an integer.

Using (7.23) in (7.22), and dividing both sides by $\sin^2 \theta$ gives

$$\frac{r}{R(r)} \frac{\partial^2}{\partial r^2} r R(r) + \frac{2m}{\hbar^2} r^2 \left(E - V(r) \right) + \frac{1}{\sin \theta P(\theta)} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} P(\theta) \right) - \frac{m^2}{\sin^2 \theta} = 0 \,.$$
(7.26)

Here the first two terms depend only upon r while the third and fourth terms depend only upon θ. For this equation to be satisfied for all possible values of r and θ, the third and forth terms must add to a constant. Let

$$\frac{1}{\sin \theta P(\theta)} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} P(\theta) \right) - \frac{m^2}{\sin^2 \theta} = -l(l+1) \,.$$
(7.27)

This can be rewritten as

$$\frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} P(\theta) \right) - \frac{m^2}{\sin^2 \theta} P(\theta) + l(l+1) P(\theta) = 0 \,.$$
(7.28)

The solution to this equation is simplified by changing variables to $x = \cos \theta$. In terms of the new variable (7.28) becomes

$$\frac{\partial}{\partial x} \left(\left(1 - x^2 \right) \frac{\partial}{\partial x} P(x) \right) + \left(l(l+1) - \frac{m^2}{1 - x^2} \right) P(x) = 0 \,.$$
(7.29)

This is Legendre's equation and is known to have finite solutions on the interval $-1 \le x \le 1$ only if l is an integer with $l \ge 0$.

First consider the solutions for m=0. These solutions are the Legendre polynomials $P_l(x)$. Traditionally these polynomials are normalized such that

$$P_l(1) = 1$$
(7.30)

and the first five solutions are

$$P_0(x) = 1$$
(7.31)
$$P_1(x) = x$$
(7.32)
$$P_2(x) = \frac{1}{2} \left(3x^2 - 1 \right)$$
(7.33)
$$P_3(x) = \frac{1}{2} \left(5x^3 - 3x \right)$$
(7.34)
$$P_4(x) = \frac{1}{8} \left(35x^4 - 30x^2 + 3 \right) \,.$$
(7.35)

Note that these are polynomials of order l and that they are even when l is even and odd when l is odd. A more compact representation of the Legendre polynomials is given by Rodrigues's formula

$$P_l(x) = \frac{1}{2^l l!} \frac{d^l}{dx^l} \left(x^2 - 1 \right)^l \,.$$
(7.36)

115

For $m \neq 0$, the solutions to (7.29) are the associated Legendre functions. For $m > 0$ the associated Legendre functions are given by

$$P_l^m(x) = \left(1 - x^2\right)^{\frac{m}{2}} \frac{d^m}{dx^m} P_l(x) \,. \tag{7.37}$$

Since $P_l(x)$ is a polynomial of degree l this implies that the only nonvanishing associated Legendre functions are for $m \leq l$. Note that (7.29) depends on m only through m^2. This implies that there are also solutions for negative m and that these are related to the solutions for positive m. The $m < 0$ solutions are defined as

$$P_l^{-m}(x) \equiv (-1)^m \frac{(l-m)!}{(l+m)!} P_l^m(x) \,. \tag{7.38}$$

The permissable values of m are then given by $-l \leq m \leq l$.

These associated Legendre functions satisfy the orthogonality relation

$$\int_{-1}^{1} dx \, P_l^m(x) P_{l'}^m(x) = \frac{2(l+m)!}{(2l+1)(l-m)!} \delta_{ll'} \,. \tag{7.39}$$

7.1.1 Spherical Harmonics

We now have solutions for angular functions $Q(\phi)$ and $P(\theta)$. It is convenient to combine these into a single normalized function of both angles. These functions are called the *spherical harmonics* and are defined as

$$Y_{lm}(\theta, \phi) = (-1)^m \left[\frac{(2l+1)(l-m)!}{4\pi(l+m)!} \right]^{\frac{1}{2}} P_l^m(\cos\theta) e^{im\phi} \,. \tag{7.40}$$

In particular,

$$Y_{l0}(\theta, \phi) = \sqrt{\frac{2l+1}{4\pi}} P_l(\cos\theta) \tag{7.41}$$

The definition of the associated Legendre functions for negative m can be used to show that

$$Y_{lm}^*(\theta, \phi) = (-1)^m Y_{l,-m}(\theta, \phi) \,. \tag{7.42}$$

The spherical Harmonics satisfy the orthogonality condition

$$\int_0^{2\pi} d\phi \int_0^{\pi} d\theta \, \sin\theta \, Y_{l'm'}^*(\theta, \phi) Y_{lm}(\theta, \phi) = \delta_{l'l} \, \delta_{m'm} \,. \tag{7.43}$$

The spherical harmonics are also complete on the space $0 \leq \phi \leq 2\pi$ and $0 \leq \theta \leq \pi$ and the completeness relation is

$$\sum_{l=0}^{\infty} \sum_{m=-l}^{l} Y_{lm}(\theta, \phi) Y_{lm}^*(\theta', \phi') = \delta(\phi - \phi') \, \delta(\cos\theta - \cos\theta') \,. \tag{7.44}$$

116

This means that any function of the two angles defined on the complete 4π solid angle can be expanded in the form

$$F(\theta, \phi) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} c_{lm} Y_{lm}(\theta, \phi),$$ (7.45)

where

$$c_{lm} = \int_0^{2\pi} d\phi \int_0^{\pi} d\theta \, \sin\theta \, Y_{lm}^*(\theta, \phi) F(\theta, \phi).$$ (7.46)

7.2   The Radial Equation

Substituting (7.27) into (7.26) gives

$$\frac{r}{R(r)} \frac{\partial^2}{\partial r^2} r R(r) + \frac{2m}{\hbar^2} r^2 \left(E - V(r) \right) - l(l+1) = 0,$$ (7.47)

which can be rewritten as

$$-\frac{\hbar^2}{2m} \left(\frac{1}{r} \frac{d^2}{dr^2} r - \frac{l(l+1)}{r^2} \right) R(r) + V(r) R(r) = E R(r).$$ (7.48)

This is the eigenvalue equation for the radial coordinate and since it explicitly depends upon the integer l, the eigenfunctions will in general be represented by $R_{nl}(r)$ where n is the energy quantum number.

The complete solutions of the Schrödinger equation can then written as

$$\psi_{nlm}(\boldsymbol{r}) = R_{nl}(r) Y_{lm}(\theta, \phi)$$ (7.49)

where the particular form of $R_{nl}(r)$ will depend upon the choice of potential $V(r)$. Note that the wave function not only depends on the energy quantum n but also on l and m.

The wave functions for bound states must be normalizable. That is

$$\int d^3 r \psi_{nlm}^*(\boldsymbol{r}) \psi_{nlm}(\boldsymbol{r}) = 1.$$ (7.50)

This implies that

$$1 = \int_0^{\infty} dr \, r^2 R_{nl}^2(r) \int_0^{2\pi} d\phi \int_0^{\pi} d\theta \sin(\theta) Y_{lm}^*(\theta, \phi) Y_{lm}(\theta, \phi) = \int_0^{\infty} dr \, r^2 R_{nl}^2(r).$$ (7.51)

For (7.51) to be satisfied the integrand must be constrained at both small and large r and it must be finite over the interval $0 < r < \infty$. Assume that $R(r) \sim r^\alpha$ for $r << 1$. Then the normalization integral in this region goes like

$$\int dr r^2 r^{2\alpha} = r^{2\alpha+3}.$$ (7.52)

117

In order for this to be finite at $r = 0$, $2\alpha + 3 \geq 0$ which implies that $\alpha \geq -\frac{3}{2}$.

Assume that $R(r) \sim r^\beta$ for $r \to \infty$. The normalization integral for large r then goes like

$$\int dr\, r^2 r^{2\beta} = r^{2\beta+3} \,. \tag{7.53}$$

In order for this vanish as $r \to \infty$, $2\beta + 3 < 0$ which implies that $\beta < -\frac{3}{2}$.

In many cases it is convenient to define a reduced radial wave function $u(r)$ that is defined such that

$$R(r) = \frac{u_{(r)}}{r} \,. \tag{7.54}$$

The differential equation for the reduced wave function is then

$$-\frac{\hbar^2}{2m}\left(\frac{d^2}{dr^2} - \frac{l(l+1)}{r^2}\right)u(r) + V(r)u(r) = Eu(r)\,. \tag{7.55}$$

7.3 Solution of the Radial Equation for a Free Particle

For a free particle $V(r) = 0$ leading to the radial wave equation

$$-\frac{\hbar^2}{2m}\left(\frac{1}{r}\frac{d^2}{dr^2}r - \frac{l(l+1)}{r^2}\right)R(r) = ER(r)\,. \tag{7.56}$$

Multiplying both sides by $-\frac{2m}{\hbar^2}$ and then moving all terms to the left-hand side gives

$$\left[\frac{1}{r}\frac{d^2}{dr^2}r + \frac{2m}{\hbar^2}E - \frac{l(l+1)}{r^2}\right]R(r) = 0\,. \tag{7.57}$$

This equation is related to Bessel's equation of half-odd-integer order. The solutions are the spherical Bessel functions. Since this is a second order differential equation, there are two sets of solutions denoted as $j_l(kr)$ and $n_l(kr)$ where $k = \sqrt{\frac{2mE}{\hbar^2}}$. These functions are given by

$$j_l(x) = (-x)^l\left(\frac{1}{x}\frac{d}{dx}\right)^l\frac{\sin x}{x} \tag{7.58}$$

and

$$n_l(x) = -(-x)^l\left(\frac{1}{x}\frac{d}{dx}\right)^l\frac{\cos x}{x} \,. \tag{7.59}$$

The first three of each of these functions are given by

$$j_0(x) = \frac{\sin x}{x}$$

$$j_1(x) = \frac{\sin x}{x^2} - \frac{\cos x}{x}$$

$$j_2(x) = \left(\frac{3}{x^3} - \frac{1}{x}\right)\sin x - \frac{3}{x^2}\cos x$$

$$n_0(x) = -\frac{\cos x}{x}$$

$$n_1(x) = -\frac{\cos x}{x^2} - \frac{\sin x}{x}$$

$$n_2(x) = -\left(\frac{3}{x^3} - \frac{1}{x}\right)\cos x - \frac{3}{x^2}\sin x \tag{7.60}$$

Note that the spherical Bessel functions $j_l(x)$ are regular at $x = 0$ while the spherical Neumann functions $n_l(x)$ are singular at $x = 0$.

In some applications it is useful to replace the singular Neumann functions by linear combinations of the Bessel and Neumann functions called the spherical Hankel functions defined as

$$h_l^{(1)}(x) = j_l(x) + in_l(x) \tag{7.61}$$

and

$$h_l^{(2)}(x) = \left[h^{(1)}(x)\right]^* . \tag{7.62}$$

The first three spherical Hankel functions are

$$h_0^{(1)}(x) = \frac{e^{ix}}{ix}$$

$$h_1^{(1)}(x) = -\frac{e^{ix}}{x}\left(1 + \frac{i}{x}\right)$$

$$h_2^{(1)}(x) = \frac{ie^{ix}}{x}\left(1 + \frac{3i}{x} - \frac{3}{x^2}\right) . \tag{7.63}$$

It is often useful to know the asymptotic forms of these functions for very small or large values of x. For $x << l$ the asymptotic forms are

$$j_l(x) \sim \frac{x^l}{(2l+1)!!} \tag{7.64}$$

and

$$n_l(x) \sim \frac{(2l+1)!!}{x^{l+1}} . \tag{7.65}$$

For $x >> l$,

$$j_l(x) \sim \frac{1}{x}\sin\left(x - \frac{l\pi}{2}\right) \tag{7.66}$$

119
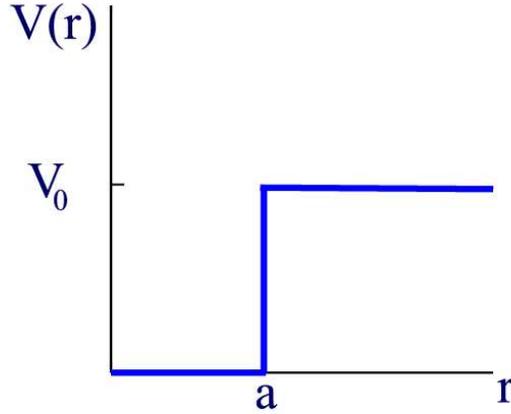
Figure 7.2: The finite spherical well potential.

$$n_l(x) \sim -\frac{1}{x} \cos\left(x - \frac{l\pi}{2}\right) \tag{7.67}$$

and

$$h_l^{(1)}(x) \sim -\frac{i}{x} e^{i\left(x - \frac{l\pi}{2}\right)} \,. \tag{7.68}$$

The requirement that the solutions to this continuum solution be normalizable in a box implies that we must choose the regular functions $j_l(x)$ as the appropriate solutions. So,

$$R_{kl}(r) \propto j_l(kr) \,. \tag{7.69}$$

7.4 The Finite Spherical Well

Next we consider the case where the potential is defined as

$$V(r) = V_0 \, \theta(r - a) \,. \tag{7.70}$$

This potential is shown graphically in Fig. 7.2.

For $r < a$ the radial equation is

$$\left[\frac{1}{r}\frac{d^2}{dr^2}r + \frac{2m}{\hbar^2}E - \frac{l(l+1)}{r^2}\right] R(r) = 0 \,, \tag{7.71}$$

which is the same as for the case of a free particle. The solutions are then the same as in the previous section

$$R(r) = A j_l(kr) \,, \tag{7.72}$$

120

where

$$k = \sqrt{\frac{2mE}{\hbar^2}} \tag{7.73}$$

and the constant A will be determined by the normalization condition.

For $r > a$ the radial equation is

$$\left[\frac{1}{r}\frac{d^2}{dr^2}r + \frac{2m}{\hbar^2}\left(E - V_0\right) - \frac{l(l+1)}{r^2}\right] R(r) = 0 \,, \tag{7.74}$$

Since this region does not contain the origin, the solution can have contributions from both the regular and irregular solutions so

$$R(r) = B j_l(k'r) + C n_l(k'r) \,, \tag{7.75}$$

where

$$k' = \sqrt{\frac{2m}{\hbar^2}\left(E - V_0\right)} \tag{7.76}$$

and B and C are determined by the boundary conditions at $r = a$ and $r = \infty$. The boundary conditions require that the radial wave function be continuous and have continuous derivative at $r = a$. That is

$$A j_l(ka) = B j_l(k'a) + C n_l(k'a) \tag{7.77}$$

and

$$\frac{\partial}{\partial r} A j_l(kr)\bigg|_{r=a} = \frac{\partial}{\partial r}\left(B j_l(k'r) + C n_l(k'r)\right)\bigg|_{r=a}. \tag{7.78}$$

The boundary condition for $r \to \infty$ depends upon whether the wave function represents a bound or continuum state. For bound states, the boundary condition is that the wave function fall rapidly enough that the radial wave function is normalizable as discussed above. For continuum solutions the boundary condition requires that the solution look like incoming or outgoing spherical waves, or some combination of the two.

7.4.1 Bound States

Bound states occur for $E < V_0$. In this case, the argument of the square root in (7.76) becomes negative and therefore k' becomes imaginary. We can then write that

$$k' = i\beta \tag{7.79}$$

where

$$\beta = \sqrt{\frac{2m}{\hbar^2}\left(V_0 - E\right)} \,. \tag{7.80}$$

121

The solution for $r > a$ can then be written as

$$A j_l(ka) = B j_l(i\beta a) + C n_l(i\beta a) \,. \tag{7.81}$$

Examination of (7.60) shows that for imaginary arguments the sines and cosines become hyperbolic functions which are singular at $r = \infty$. This violates the normalization condition for bound states. However, by choosing $C = iB$ and using (7.61) the solution becomes a spherical hankel function of the first kind which can be seen from (7.63) vanishes exponentially for large r. The first three solutions for $r > a$ are

$$
\begin{aligned}
h_0^{(1)}(i\beta r) &= -\frac{e^{-\beta r}}{\beta r} \\
h_1^{(1)}(i\beta r) &= i\frac{e^{-\beta r}}{\beta r}\left(1 + \frac{1}{\beta r}\right) \\
h_2^{(1)}(i\beta r) &= \frac{e^{-\beta r}}{\beta r}\left(1 + \frac{3}{\beta r} + \frac{3}{\beta^2 r^2}\right) \,.
\end{aligned}
\tag{7.82}
$$

The boundary conditions then become

$$A j_l(ka) = B h_l^{(1)}(i\beta a) \tag{7.83}$$

and

$$\left.\frac{\partial}{\partial r} A j_l(kr)\right|_{r=a} = \left.\frac{\partial}{\partial r} B h_l^{(1)}(i\beta r)\right|_{r=a} \,. \tag{7.84}$$

From (7.83) we can write

$$B = A\frac{j_l(ka)}{h_l^{(1)}(i\beta a)} \,. \tag{7.85}$$

Dividing (7.84) by (7.83) yields

$$\left.\frac{1}{j_l(kr)}\frac{\partial}{\partial r} j_l(kr)\right|_{r=a} = \left.\frac{1}{h_l^{(1)}(i\beta r)}\frac{\partial}{\partial r} h_l^{(1)}(i\beta r)\right|_{r=a} \,. \tag{7.86}$$

Recall that both k and β depend on the energy E. Equation (7.86) will only be satisfied for certain values of E and will thus determine the energy eigenvalues. This eigencondition is a fairly complicated transcendental equation for the energy. For example for the $l = 0$ case it gives

$$\sqrt{\frac{2mE}{\hbar^2}} \cot\left(\sqrt{\frac{2mE}{\hbar^2}}a\right) + \sqrt{\frac{2m(V_0 - E)}{\hbar^2}} = 0 \,. \tag{7.87}$$

These equations can be easily and rapidly solved numerically.

7.4.2 The Infinite Spherical Well

An interesting limiting case of the spherical well is when $V_0 \to \infty$. This implies that $\beta \to \infty$ also. For $r > a$ we can now use

$$\lim_{\beta \to \infty} h_l^{(1)}(i\beta r) = 0 \tag{7.88}$$

to show that the boundary condition at $r = a$ is simply

$$A j_l(ka) = 0 \tag{7.89}$$

or

$$j_l(ka) = 0. \tag{7.90}$$

If we define the nth root of the lth spherical Bessel function as α_{nl}, then then

$$k_{nl} = \frac{\alpha_{nl}}{a} \tag{7.91}$$

and the energy eigenvalues are

$$E_{nl} = \frac{\hbar^2}{2m} \left(\frac{\alpha_{nl}}{a} \right)^2. \tag{7.92}$$

The values of the α_{nl} can be found in some mathematical tables or determined numerically. The search for these roots is simplified for $x >> l$ by using the asymptotic form

$$j_l(x) \sim \frac{1}{x} \sin \left(x - \frac{l\pi}{2} \right) \tag{7.93}$$

which requires that

$$\alpha_{nl} - \frac{l\pi}{2} \cong n\pi \qquad \text{for} \qquad n > 0. \tag{7.94}$$

This implies that

$$\alpha_{nl} \cong \pi \left(n + \frac{l}{2} \right). \tag{7.95}$$

The wave functions for the infinite spherical well are thus given by

$$\psi_{nlm}(\boldsymbol{r}) = A_{nl} j_l \left(\frac{\alpha_{nl}}{a} r \right) \theta(a - r) Y_{lm}(\vartheta, \varphi). \tag{7.96}$$

The normalization constant can be determined from the normalization condition

$$1 = \int_0^\infty dr \, r^2 R_{nl}^2(r) = A_{nl}^2 \int_0^\infty dr \, r^2 j_l^2 \left(\frac{\alpha_{nl}}{a} r \right) \theta(a - r) = A_{nl}^2 \int_0^a dr \, r^2 j_l^2 \left(\frac{\alpha_{nl}}{a} r \right). \tag{7.97}$$

This integral can be performed using the identity

$$\int_0^a dr \, r^2 j_l \left(\frac{\alpha_{nl}}{a} r \right) j_l \left(\frac{\alpha_{n'l}}{a} r \right) = \frac{a^3}{2} \left[j_{l+1}(\alpha_{nl}) \right]^2 \delta_{nn'}. \tag{7.98}$$

7.5 The Coulomb Potential

A simple version of hydrogen-like atoms can be obtained by solving the Schrödinger equation with a Coulomb potential

$$V(r) = -\frac{Ze^2}{r}\,. \tag{7.99}$$

Here we will focus on the bound state solutions of the Schrödinger equation which will occur for negative energies. That is where $E = -|E|$. In this case we can write the radial equation as

$$-\frac{\hbar^2}{2m}\left(\frac{1}{r}\frac{d^2}{dr^2}r - \frac{l(l+1)}{r^2}\right)R(r) - \frac{Ze^2}{r}R(r) = -|E|R(r)\,. \tag{7.100}$$

Multiplying both sides by $-\frac{2m}{\hbar^2}$, moving all terms to the left-hand side and expanding the second derivative term gives

$$\left(\frac{d^2}{dr^2} + \frac{2}{r}\frac{d}{dr} - \frac{l(l+1)}{r^2} + \frac{2mZe^2}{\hbar^2 r} - \frac{2m|E|}{\hbar^2}\right)R(r) = 0\,. \tag{7.101}$$

The first step in finding a solution to this equation is to extract the asymptotic behavior of the radial wave function for $r \to \infty$. Assuming that the wave function and its derivatives are all of roughly similar size at large r, then the terms containing negative powers of r in their coefficients will tend to be small as r becomes large. Then the wave equation for large r is approximately

$$\left(\frac{d^2}{dr^2} - \frac{2m|E|}{\hbar^2}\right)R(r) = 0\,. \tag{7.102}$$

The solutions to this equation are exponentials of real argument. For the state to be normalizable, the wave function must correspond to a damped exponential so

$$R(r) \sim e^{-\frac{\beta r}{2}}\,, \tag{7.103}$$

where

$$\beta = \sqrt{\frac{8m|E|}{\hbar^2}}\,. \tag{7.104}$$

It is convenient to define a dimensionless radial variable

$$\rho = \beta r\,. \tag{7.105}$$

The radial wave equation in terms of ρ is then

$$\left(\beta^2\frac{d^2}{d\rho^2} + \beta^2\frac{2}{\rho}\frac{d}{d\rho} - \beta^2\frac{l(l+1)}{\rho^2} + \beta\frac{2mZe^2}{\hbar^2\rho} - \frac{\beta^2}{4}\right)R(\rho) = 0\,, \tag{7.106}$$

or dividing the equation by β^2

$$\left(\frac{d^2}{d\rho^2} + \frac{2}{\rho}\frac{d}{d\rho} - \frac{l(l+1)}{\rho^2} + \frac{\lambda}{\rho} - \frac{1}{4}\right)R(\rho) = 0\,, \tag{7.107}$$

where

$$\lambda = \frac{2mZe^2}{\beta\hbar^2} = \frac{Ze^2}{\hbar}\sqrt{\frac{m}{2|E|}} = Z\alpha\sqrt{\frac{mc^2}{2|E|}}\,, \tag{7.108}$$

with

$$\alpha = \frac{e^2}{\hbar c} \tag{7.109}$$

being the fine structure constant.

The radial wave function can now be written in a form that explicitly shows the its asymptotic behavior as

$$R(\rho) = F(\rho)e^{-\frac{\rho}{2}}\,. \tag{7.110}$$

Using

$$\frac{d}{d\rho}R(\rho) = \frac{dF(\rho)}{d\rho}e^{-\frac{\rho}{2}} - \frac{1}{2}F(\rho)e^{-\frac{\rho}{2}} \tag{7.111}$$

and

$$\frac{d^2}{d\rho^2}R(\rho) = \frac{d^2F(\rho)}{d\rho^2}e^{-\frac{\rho}{2}} - \frac{dF(\rho)}{d\rho}e^{-\frac{\rho}{2}} + \frac{1}{4}F(\rho)e^{-\frac{\rho}{2}}\,, \tag{7.112}$$

the differential equation becomes

$$\left[\frac{d^2F(\rho)}{d\rho^2} - \frac{dF(\rho)}{d\rho} + \frac{1}{4}F(\rho) + \frac{2}{\rho}\left(\frac{dF(\rho)}{d\rho} - \frac{1}{2}F(\rho)\right) - \frac{l(l+1)}{\rho^2}F(\rho)\right.$$
$$\left. + \frac{\lambda}{\rho}F(\rho) - \frac{1}{4}F(\rho)\right]e^{-\frac{\rho}{2}} = 0\,, \tag{7.113}$$

This can be simplified to give

$$\frac{d^2F(\rho)}{d\rho^2} + \left(\frac{2}{\rho} - 1\right)\frac{dF(\rho)}{d\rho} - \frac{l(l+1)}{\rho^2}F(\rho) + \frac{\lambda-1}{\rho}F(\rho) = 0\,. \tag{7.114}$$

The next step is to see if it is possible to construct a solution in the form of a power series

$$F(\rho) = \rho^s\sum_{k=0}^{\infty}a_k\rho^k = \sum_{k=0}^{\infty}a_k\rho^{k+s}\,. \tag{7.115}$$

Using

$$\frac{d}{d\rho}F(\rho) = \sum_{k=0}^{\infty}(k+s)a_k\rho^{k+s-1} \tag{7.116}$$

125

and

$$\frac{d^2}{d\rho^2}F(\rho) = \sum_{k=0}^{\infty}(k+s)(k+s-1)a_k\rho^{k+s-2}\,,\qquad(7.117)$$

and substituting into (7.114) gives

$$
\begin{aligned}
0 &= \sum_{k=0}^{\infty}(k+s)(k+s-1)a_k\rho^{k+s-2} + \left(\frac{2}{\rho}-1\right)\sum_{k=0}^{\infty}(k+s)a_k\rho^{k+s-1}\\
&\quad -\frac{l(l+1)}{\rho^2}\sum_{k=0}^{\infty}a_k + \frac{\lambda-1}{\rho}\sum_{k=0}^{\infty}a_k\\
&= \sum_{k=0}^{\infty}(k+s)(k+s-1)a_k\rho^{k+s-2} + 2\sum_{k=0}^{\infty}(k+s)a_k\rho^{k+s-2} - \sum_{k=0}^{\infty}(k+s)a_k\rho^{k+s-1}\\
&\quad -l(l+1)\sum_{k=0}^{\infty}a_k\rho^{k+s-2} + (\lambda-1)\sum_{k=0}^{\infty}a_k\rho^{k+s-1}\\
&= \sum_{k=0}^{\infty}\left[(k+s)(k+s+1)-l(l+1)\right]a_k\rho^{k+s-2} + \sum_{k=0}^{\infty}\left[\lambda-1-(k+s)\right]a_k\rho^{k+s-1}\quad(7.118)
\end{aligned}
$$

In order for the power series to sum to zero, the coefficients of each power have to vanish. It is useful, therefore, to rearrange the above expression to clearly show similar powers. We can do this be changing the dummy index in the second term such that $k \to k-1$. We then obtain

$$
\begin{aligned}
0 &= \sum_{k=0}^{\infty}\left[(k+s)(k+s+1)-l(l+1)\right]a_k\rho^{k+s-2} + \sum_{k=1}^{\infty}\left[\lambda-1-(k-1+s)\right]a_{k-1}\rho^{k+s-2}\\
&= \left[s(s+1)-l(l+1)\right]a_0\rho^{s-2}\\
&\quad + \sum_{k=1}^{\infty}\left\{\left[(k+s)(k+s+1)-l(l+1)\right]a_k + \left[\lambda-k-s\right]a_{k-1}\right\}\rho^{k+s-2}\,.\qquad(7.119)
\end{aligned}
$$

For the coefficient of the first term to vanish requires that

$$s(s+1) - l(l+1) = 0\,.\qquad(7.120)$$

So, $s = l$. Using this, the vanishing of the rest of the coefficients requires that

$$\left[(k+l)(k+l+1) - l(l+1)\right]a_k + \left[\lambda - k - l\right]a_{k-1} = 0\,,\qquad(7.121)$$

or

$$a_k = \frac{k+l-\lambda}{k(k+2l+1)}a_{k-1}\,.\qquad(7.122)$$

This is equivalent to

$$a_{k+1} = \frac{k+l+1-\lambda}{(k+1)(k+2l+2)}a_k\,.\qquad(7.123)$$

126

Now note that for $k \gg l$, this implies that

$$\frac{a_{k+1}}{a_k} \cong \frac{1}{k} . \tag{7.124}$$

An infinite power series with coefficients in this ratio cannot converge. This implies that for this solution to be finite, only a finite number of terms in the power series can contribute. This means that there must be some maximum value of k for which the coefficient of a_k vanishes in the recursion relation for the coefficients. This will guarantee that all higher coefficients vanish and that the series is then truncated to a finite number of terms. If this maximum value of k is called n', then

$$n' + l + 1 - \lambda = 0 \tag{7.125}$$

which implies that

$$\lambda = n = n' + l + 1 \tag{7.126}$$

where n' is called the radial quantum number and n is called the total quantum number. Using our previous definition of λ,

$$n = Z\alpha \sqrt{\frac{mc^2}{2|E|}} . \tag{7.127}$$

So,

$$E_n = -|E_n| = -\frac{Z^2\alpha^2 mc^2}{2n^2} , \tag{7.128}$$

which is in agreement with the energy spectrum of the Bohr atom.

7.5.1 Degeneracy

Since the energy depends on on n and this in turn depends on n' and l, there will be states that are degenerate in energy. With $n' \geq 0$ and $l \geq 0$, then $n \geq 1$. This means that $0 \leq l \leq n - 1$. Now for each value of l, $-l \leq m \leq l$, so there are $2l + 1$ degenerate states for each l. The total degeneracy will then be given by

$$\sum_{l=0}^{n-1}(2l + 1) = 2\frac{n(n - 1)}{2} + n = n^2 . \tag{7.129}$$

While the degeneracy in m occurs of any central force, the degeneracy in l is characteristic of the Coulomb force.

127

7.5.2 Radial Wave Functions

Now using (7.122) we can write

$$a_1 = \frac{l - n + 1}{2l + 2} a_0 \tag{7.130}$$

and

$$a_2 = \frac{l - n + 2}{2(2l + 2)} a_1 = \frac{l - n + 2}{2(2l + 3)} \frac{l - n + 1}{2l + 3} a_0 . \tag{7.131}$$

By induction we can write

$$a_k = \frac{(l - n + k)(l - n + k - 1) \cdots (l - n + 1)}{k!(2l + 1 + k)(2l + k) \cdots (2l + 2)} a_0 . \tag{7.132}$$

This can be written in a more compact form by noting that

$$(2l + 1 + k)(2l + k) \cdots (2l + 2) = \frac{(2l + 1 + k)!}{(2l + 1)!} . \tag{7.133}$$

Since $k \leq n - l - 1$, we have

$$l - n + k \leq l - n + n - l - 1 = -1 . \tag{7.134}$$

So, all of the factors in the numerator of (7.132) are negative. We can then rewrite these factors as

$$(l - n + k)(l - n + k - 1) \cdots (l - n + 1)$$
$$= (-1)^k (n - l - k)(n - l - k + 1) \cdots (n - l - 1)$$
$$= (-1)^k \frac{(n - l - 1)!}{(n - l - k - 1)!} . \tag{7.135}$$

We can put everything back together to give

$$a_k = (-1)^k \frac{(n - l - 1)!(2l + 1)!}{k!(2l + 1 + k)!(n - l - k - 1)!} a_0 . \tag{7.136}$$

We are free to choose a_0 to be any function of n and l. If we choose

$$a_0 = (-1)^{2l+1} \frac{[(n + l)!]^2}{(n - l - 1)!(2l + 1)!} , \tag{7.137}$$

then

$$a_k = (-1)^{k+2l+1} \frac{[(n + l)!]^2}{k!(2l + 1 + k)!(n - l - k - 1)!} . \tag{7.138}$$

128

We can now write

$$F(\rho) = \rho^l \sum_{k=0}^{n-l-1} \frac{(-1)^{k+2l+1}[(n+l)!]^2 \rho^k}{k!(k+2l+1)!(n-l-k-1)!} = \rho^l L_{n+l}^{2l+1}(\rho),$$
(7.139)

where $L_p^q(\rho)$ is the associated Laguerre polynomial.

Using the identity

$$\int_0^\infty d\rho\, \rho^2\, \rho^{2l} e^{-\rho} \left[L_{n+l}^{2l+1}(\rho) \right]^2 = \frac{2n[(l+n)!]^3}{(n-l-1)!},$$
(7.140)

the normalized radial wave function is

$$R_{nl}(r) = - \left\{ \left(\frac{2Z}{na_0} \right)^3 \frac{(n-l-1)!}{2n[(n+l)!]^3} \right\}^{\frac{1}{2}} \rho^l e^{-\frac{\rho}{2}} L_{n+l}^{2l+1}(\rho),$$
(7.141)

where

$$a_0 = \frac{\hbar^2}{me^2} = \frac{\hbar}{\alpha mc}$$
(7.142)

is the Bohr radius and should not be confused with the similarly named coefficient used above.

Chapter 8

Formal Foundations for Quantum Mechanics

8.1 Review

In this chapter we will describe some of the formal mathematical aspects of quantum mechanics. It is useful to first summarize what we have learned to this point. The expressions below will be for the case of a particle in three-dimensional space. Our starting point for quantum mechanics is has been the Schrödinger equation

$$\hat{H}\Psi(\boldsymbol{r},t) = i\hbar\frac{\partial}{\partial t}\Psi(\boldsymbol{r},t)\,. \tag{8.1}$$

The hamiltonian operator is obtained by starting with the hamiltonian of a corresponding classical system and replacing all of the momenta and coordinates with momentum and coordinate operators. That is

$$\hat{H} = H(\hat{\boldsymbol{p}},\hat{\boldsymbol{r}})\,. \tag{8.2}$$

The components of the coordinate and momentum operators must satisfy the canonical commutation relations

$$[\hat{r}_i,\hat{p}_j] = i\hbar\delta_{ij} \tag{8.3}$$

and

$$[\hat{r}_i,\hat{r}_j] = [\hat{p}_i,\hat{p}_j] = 0\,. \tag{8.4}$$

We have proven for the case of the one-dimensional Schrödinger equation that the canonical commutation relations lead to the Heisenberg uncertainty relations. In three-dimensions this is

$$\Delta p_i\Delta r_i \geq \frac{\hbar}{2}\,. \tag{8.5}$$

In fact, the proof that we considered applies to any to operators that satisfy the commutation relation

$$[\hat{A},\hat{B}] = i\hbar \tag{8.6}$$

resulting in the uncertainty relation

$$\Delta A \Delta B \geq \frac{\hbar}{2} \,. \tag{8.7}$$

The physical interpretation of the Schrödinger wave function $\Psi(\boldsymbol{r}, t)$ is that it represents a probability amplitude. The probability is given by

$$\rho(\boldsymbol{r}, t) = \Psi^*(\boldsymbol{r}, t)\Psi(\boldsymbol{r}, t) \,. \tag{8.8}$$

Since the total probability must be unity,

$$\int d^3 r \rho(\boldsymbol{r}, t) = 1 \,. \tag{8.9}$$

The probability current density is defined by

$$\boldsymbol{j}(\boldsymbol{r}, t) = \frac{\hbar}{2im} \left[\Psi^*(\boldsymbol{r}, t)\boldsymbol{\nabla}\Psi(\boldsymbol{r}, t) - \left(\boldsymbol{\nabla}\Psi^*(\boldsymbol{r}, t)\right)\Psi(\boldsymbol{r}, t) \right] \,, \tag{8.10}$$

and the probability density and current density must satisfy the continuity equation

$$\boldsymbol{\nabla} \cdot \boldsymbol{j}(\boldsymbol{r}, t) + \frac{\partial}{\partial t}\rho(\boldsymbol{r}, t) = 0 \,, \tag{8.11}$$

which requires that probability is conserved.

The time-dependent Schrödinger wave function is obtained by first satisfy the time-independent equation

$$\hat{H}\psi_n(\boldsymbol{r}) = E_n\psi_n(\boldsymbol{r}) \tag{8.12}$$

with eigenenergies E_n and eigenfunctions $\psi_n(\boldsymbol{r})$. In order for the eigenfunctions to form the basis of a hilbert space, the wave functions must be orthonormal. That is,

$$\int d^3 r \psi_n^*(\boldsymbol{r})\psi_m(\boldsymbol{r}) = \delta_{nm} \,. \tag{8.13}$$

The wave functions must be square integrable, which implies that they are localized. The eigenfunctions then form a complete, orthonormal basis set that spans the hilbert space. These eigenfunctions satisfy the completeness relation

$$\sum_{n=0}^{\infty} \psi_n(\boldsymbol{r})\psi_n^*(\boldsymbol{r}') = \delta(\boldsymbol{r} - \boldsymbol{r}') \,. \tag{8.14}$$

The time-dependent wave functions which correspond to the eigen functions are

$$\Psi_n(\boldsymbol{r}, t) = \psi_n(\boldsymbol{r})e^{-\frac{i}{\hbar}E_n t} \,. \tag{8.15}$$

which are also orthonormal. Any solution to the time-independent Schrödinger equation can be expanded in these solutions in the form

$$\Psi(\boldsymbol{r}, t) = \sum_{n=0}^{\infty} c_n \Psi_n(\boldsymbol{r}, t), \tag{8.16}$$

where the expansion coefficients are obtained from the initial condition using

$$c_n = \int d^3 r \psi_n^*(\mathrm{r}) \Psi(\mathrm{r}, 0). \tag{8.17}$$

Normalization of probability for wave functions of this form requires that the expansion coefficients satisfy the constraint

$$\sum_{n=0}^{\infty} |c_n|^2 = 1. \tag{8.18}$$

Consider the action of the hamiltonian operator on the time-dependent wave function $\Psi_n(\boldsymbol{r}, t)$. This gives

$$\hat{H}\Psi_n(\boldsymbol{r}, t) = \hat{H}\psi_n(\boldsymbol{r})e^{-\frac{i}{\hbar}E_n t} = E_n \psi_n(\boldsymbol{r})e^{-\frac{i}{\hbar}E_n t} = E_n \Psi_n(\boldsymbol{r}, t). \tag{8.19}$$

This is the result of the fact that $\psi_n(\mathrm{r})$ is an eigenstate of the hamiltonian operator. This result shows that the application of the hamiltonian on this state returns the same state multiplied by the energy of the state. This means that the hamiltonian can be used to measure the energy of the state with out disturbing the system. The energy is, therefore, an *observable* of the system since it can be measured repeatedly without changing the state of the system.

Now consider the action of the hamiltonian operator on an arbitrary solution to the time-dependent Schrödinger equation. This gives

$$\hat{H}\Psi(\boldsymbol{r}, t) = \hat{H}\sum_{n=0}^{\infty} c_n \Psi_n(\boldsymbol{r}, t) = \sum_{n=0}^{\infty} c_n \hat{H}\Psi_n(\boldsymbol{r}, t) = \sum_{n=0}^{\infty} c_n E_n \Psi_n(\boldsymbol{r}, t) \neq E\Psi_n(\boldsymbol{r}, t). \tag{8.20}$$

This means that for such a general solution, an attempt to measure the energy of the system with the hamiltonian operator changes the state of the system. The special property of the eigenstates of the hamiltonian is that they can be measured without changing the state of the system.

Although we have made the above argument based on the hamiltonian operator, it is possible to find eigenstates of other operators in the hilbert space. Consider to operators that have no explicit time dependence. This means that

$$\frac{\partial}{\partial t}\hat{A} = \frac{\partial}{\partial t}\hat{B} = 0. \tag{8.21}$$

In the context of our discussion of the parity operator we have shown that if the two operators satisfy the commutation relation

$$[\hat{A}, \hat{B}] = 0 \qquad (8.22)$$

that if $\phi(\boldsymbol{r})$ is an eigenfunction of operator $\hat{A}$ satifying

$$\hat{A}\phi(\boldsymbol{r}) = \alpha\phi(\boldsymbol{r}), \qquad (8.23)$$

then $\phi(\boldsymbol{r})$ is also an eigenfunction of the operator $hatB$ satifying

$$\hat{B}\phi(\boldsymbol{r}) = \beta\phi(\boldsymbol{r}). \qquad (8.24)$$

We can, therefore, label the eigenfunctions by the *quantum numbers* labelling the eigenvalues of both operators. That is

$$\phi(\boldsymbol{r}) = \phi_{\alpha\beta}(\boldsymbol{r}). \qquad (8.25)$$

The Schrödinger wave functions that we use as a basis for the hilbert space are eigenfunctions of the hamiltonian. However, they may also be eigenfunctions of other operators. According to the argument given above, provided that the operators satisfy the commutation relations

$$[\hat{H}, \hat{\mathcal{O}}_i] = 0. \qquad (8.26)$$

and

$$[\hat{\mathcal{O}}_i, \hat{\mathcal{O}}_j] = 0 \qquad (8.27)$$

for $i, j = 1, \ldots, N$, with the operators satisfying the eigenequations

$$\hat{\mathcal{O}}_i\psi_n(\boldsymbol{r}) = \alpha_i\psi_n(\boldsymbol{r}), \qquad (8.28)$$

the wave functions can be labelled by all of the quantum numbers associated with these operators as

$$\psi_n(\boldsymbol{r}) = \psi_{n,\alpha_1,\ldots\alpha_N}(\boldsymbol{r}). \qquad (8.29)$$

The implication of this is that each of these operators will be associated with some dynamical variable of the hamiltonian and that since the states are eigenfunctions of all of the operators, all of the values of all of all of these dynamical variables can be measured without changing the state of the system. That is, all of these operators represent observables of the system. We also require that the eigenvalues associated with each operator be real numbers, which requires that the operators be hermitian. Therefore, we require that the states of a system be uniquely specified by the eigenvalues of a complete set of hermitian operators that commute with the hamiltonian and are mutually commuting.

8.2 Dirac Notation

We have shown previously that, starting with the solution to time-dependent Schrödinger wave functions, we can define a momentum-space wavefunction

$$\Psi(\boldsymbol{p}, t) = \frac{1}{(2\pi\hbar)^{\frac{3}{2}}} \int d^3 r\, e^{-\frac{i}{\hbar}\boldsymbol{p}\cdot\boldsymbol{r}} \Psi(\boldsymbol{r}, t) \tag{8.30}$$

as the Fourier transform of the coordinate-space wave function. The inverse of the this transform is

$$\Psi(\boldsymbol{r}, t) = \frac{1}{(2\pi\hbar)^{\frac{3}{2}}} \int d^3 p\, e^{\frac{i}{\hbar}\boldsymbol{p}\cdot\boldsymbol{r}} \Psi(\boldsymbol{p}, t)\,. \tag{8.31}$$

By defining the Fourier transforms as given by these two equations, we have shown that the momentum-space wave functions also satisfy a normalization condintion

$$\int d^3 p\, |\Psi(\boldsymbol{p}, t)|^2 = 1\,. \tag{8.32}$$

This means that we can define a momentum-space probability density as

$$\rho(\boldsymbol{p}, t) = |\Psi(\boldsymbol{p}, t)|^2 \tag{8.33}$$

and this can be used for obtaining averages of operators. We have shown that if the coordinate-space operators are,

$$\hat{\boldsymbol{r}} = \boldsymbol{r} \tag{8.34}$$

$$\hat{\boldsymbol{p}} = \frac{\hbar}{i}\boldsymbol{\nabla} \tag{8.35}$$

The corresponding momentum space operators are

$$\hat{r} = -\frac{\hbar}{i}\boldsymbol{\nabla}_p \tag{8.36}$$

$$\hat{\boldsymbol{p}} = \boldsymbol{p}\,, \tag{8.37}$$

where

$$\boldsymbol{\nabla}_p \equiv \sum_{i=1}^{3} \boldsymbol{e}_i \frac{\partial}{\partial p_i}\,. \tag{8.38}$$

Both sets of operators satisfy the canonical commutation relations and, therefore, both coordinate-space and momentum-space representations are physically equivalent.

Although it most common to see quantum mechanical calculation performed in coordinate space, we have already seen that is sometimes convenient to calculate in momentum space. In a great number of cases calculation of quantum mechanical matrix elements is very complicated and simplification of the results becomes is very

tedious when all of the necessary integrals over coordinates or momenta are taken into account. In many of these situations the same steps must be followed in simplifying the expressions whether the calculation is done in coordinate or momentum space. With the object of simplifying the notation, P. A. M. Dirac introduced a simplified abstract notation for quantum mechanics that greatly aids in performing many complicated quantum mechanical calculations.

The basic simplification of the Dirac notation is to introduce an abstract state that is an eigenvector of the position operator. That is

$$\hat{\boldsymbol{r}} \ket{\boldsymbol{r}} = \boldsymbol{r} \ket{\boldsymbol{r}} , \tag{8.39}$$

where the state $\ket{\boldsymbol{r}}$ is called a *ket* and its conjugate

$$\bra{\boldsymbol{r}} = \ket{\boldsymbol{r}}^* \tag{8.40}$$

is called a *bra*. We require that

$$\braket{\boldsymbol{r} | \boldsymbol{r}'} = \delta(\boldsymbol{r} - \boldsymbol{r}') \tag{8.41}$$

and

$$\int d^3 \ket{\boldsymbol{r}} \bra{\boldsymbol{r}} = \hat{1} . \tag{8.42}$$

Similarly, a momentum eigenket can be defined such that

$$\hat{\boldsymbol{p}} \ket{\boldsymbol{p}} = \boldsymbol{p} \ket{\mathrm{p}} \tag{8.43}$$

with

$$\bra{\boldsymbol{p}} = \ket{\boldsymbol{p}}^* , \tag{8.44}$$

$$\braket{\boldsymbol{p} | \boldsymbol{p}'} = \delta(\boldsymbol{p} - \boldsymbol{p}') \tag{8.45}$$

and

$$\int d^3 p \ket{\boldsymbol{p}} \bra{\boldsymbol{p}} = \hat{1} . \tag{8.46}$$

Next, we define an abstract state vector $\ket{\Psi(t)}$ such that the time-dependent coordinate-space Schrödinger wave function is given by

$$\Psi(\boldsymbol{r}, t) \equiv \braket{\boldsymbol{r} | \Psi(t)} . \tag{8.47}$$

The conjugate of the wave function is then

$$\Psi^*(\boldsymbol{r}, t) \equiv \braket{\Psi(t) | \boldsymbol{r}} . \tag{8.48}$$

We can now use the completeness of the momentum eigenstates (8.46) to introduce the unit operator between the coordinate eigenstate and the Schrödinger eigenstate to give

$$\Psi(\boldsymbol{r}, t) = \bra{\boldsymbol{r}} \left(\int d^3 p \ket{\boldsymbol{p}} \bra{\boldsymbol{p}} \right) \ket{\Psi(t)} = \int d^3 p \braket{\boldsymbol{r} | \boldsymbol{p}} \braket{\boldsymbol{p} | \Psi(t)} . \tag{8.49}$$

136

If we now define the momentum space wave function as

$$\Psi(\boldsymbol{p}, t) \equiv \langle \boldsymbol{p} | \Psi(t) \rangle , \tag{8.50}$$

the coordinate-space wave function can be written as

$$\Psi(\boldsymbol{r}, t) = \int d^3 p \, \langle \boldsymbol{r} | \boldsymbol{p} \rangle \, \Psi(\boldsymbol{p}, t) . \tag{8.51}$$

Comparing this with (8.31), we can see that

$$\langle \boldsymbol{r} | \boldsymbol{p} \rangle \equiv \frac{1}{(2\pi\hbar)^{\frac{3}{2}}} e^{\frac{i}{\hbar} \boldsymbol{p} \cdot \boldsymbol{r}} . \tag{8.52}$$

The conjugate of this is then

$$\langle \boldsymbol{p} | \boldsymbol{r} \rangle = \langle \boldsymbol{r} | \boldsymbol{p} \rangle^* = \frac{1}{(2\pi\hbar)^{\frac{3}{2}}} e^{-\frac{i}{\hbar} \boldsymbol{p} \cdot \boldsymbol{r}} . \tag{8.53}$$

We can now consider the expectation value of abstract operator $\hat{\mathcal{O}}$ between to abstract state for some abstract state $|\Psi_i(t)\rangle$. That is,

$$\langle \mathcal{O} \rangle = \langle \Psi_i(t) | \, \mathcal{O} \, | \Psi_i(t) \rangle . \tag{8.54}$$

This can be realized in coordinate space by inserting unity as given by (8.42) on either side of the operator giving

$$
\begin{aligned}
\langle \Psi(t) | \, \hat{\mathcal{O}} \, | \Psi(t) \rangle &= \langle \Psi(t) | \left(\int d^3 r' \, | \boldsymbol{r}' \rangle \langle \boldsymbol{r}' | \right) \hat{\mathcal{O}} \left(\int d^3 r \, | \boldsymbol{r} \rangle \langle \boldsymbol{r} | \right) | \Psi(t) \rangle \\
&= \int d^3 r' \int d^3 r \, \langle \Psi(t) | \, \boldsymbol{r}' \rangle \langle \boldsymbol{r}' | \hat{\mathcal{O}} \, | \boldsymbol{r} \rangle \langle \boldsymbol{r} | \, \Psi(t) \rangle .
\end{aligned}
\tag{8.55}
$$

Similarly, in momentum space

$$
\begin{aligned}
\langle \Psi(t) | \, \hat{\mathcal{O}} \, | \Psi(t) \rangle &= \langle \Psi(t) | \left(\int d^3 p' \, | \boldsymbol{p}' \rangle \langle \boldsymbol{p}' | \right) \hat{\mathcal{O}} \left(\int d^3 p \, | \boldsymbol{p} \rangle \langle \boldsymbol{p} | \right) | \Psi(t) \rangle \\
&= \int d^3 p' \int_d^3 p \, \langle \Psi(t) | \, \boldsymbol{p}' \rangle \langle \boldsymbol{p}' | \hat{\mathcal{O}} \, | \boldsymbol{p} \rangle \langle \boldsymbol{p} | \, \Psi(t) \rangle .
\end{aligned}
\tag{8.56}
$$

Therefore, the coordinate-space representation of the operator is $\langle \boldsymbol{r}' | \hat{\mathcal{O}} \, | \boldsymbol{r} \rangle$ and the momentum-space representation of the operator is $\langle \boldsymbol{p}' | \hat{\mathcal{O}} \, | \boldsymbol{p} \rangle$.

The coordinate space representation of the position operator is

$$\langle \boldsymbol{r}' | \, \hat{\boldsymbol{r}} \, | \boldsymbol{r} \rangle = \langle \boldsymbol{r}' | \, \boldsymbol{r} \, | \boldsymbol{r} \rangle = \boldsymbol{r} \, \langle \boldsymbol{r}' | \, \boldsymbol{r} \rangle = \boldsymbol{r} \delta(\boldsymbol{r}' - \boldsymbol{r}) . \tag{8.57}$$

The delta function indicates that this operator depends only on one position variable. Such and operator is said to be *local* or *diagonal*.

The coordinate space representation of the momentum operator is

$$
\begin{aligned}
\langle \boldsymbol{r}'| \hat{\boldsymbol{p}} |\boldsymbol{r}\rangle &= \langle \boldsymbol{r}'| \hat{\boldsymbol{p}} \left(\int d^3p \, |\boldsymbol{p}\rangle \langle \boldsymbol{p}| \right) |\boldsymbol{r}\rangle = \int d^3p \, \langle \boldsymbol{r}'| \hat{\boldsymbol{p}} |\boldsymbol{p}\rangle \langle \boldsymbol{p}| \, \boldsymbol{r}\rangle \\
&= \int d^3p \, \boldsymbol{p} \, \langle \boldsymbol{r}'| \, \boldsymbol{p}\rangle \langle \boldsymbol{p}| \, \boldsymbol{r}\rangle = \frac{1}{(2\pi\hbar)^3} \int d^3p \, \boldsymbol{p} e^{\frac{i}{\hbar}\boldsymbol{p}\cdot\boldsymbol{r}'} e^{-\frac{i}{\hbar}\boldsymbol{p}\cdot\boldsymbol{r}} \\
&= \frac{1}{(2\pi\hbar)^3} \int d^3p \, e^{\frac{i}{\hbar}\boldsymbol{p}\cdot\boldsymbol{r}'} \left(-\frac{\hbar}{i}\boldsymbol{\nabla} e^{-\frac{i}{\hbar}\boldsymbol{p}\cdot\boldsymbol{r}} \right) = -\frac{\hbar}{i}\boldsymbol{\nabla} \frac{1}{(2\pi\hbar)^3} \int d^3p \, e^{\frac{i}{\hbar}\boldsymbol{p}\boldsymbol{r}'} e^{-\frac{i}{\hbar}\boldsymbol{p}\boldsymbol{r}} \\
&= -\frac{\hbar}{i}\boldsymbol{\nabla}\delta(\boldsymbol{r}' - \boldsymbol{r}) \doteq \delta(\boldsymbol{r}' - \boldsymbol{r})\frac{\hbar}{i}\boldsymbol{\nabla} \, .
\end{aligned}
\tag{8.58}
$$

The coordinate space representations of the position and momentum operators mean that the expectation values of these operators will be

$$
\langle \Psi(t)| \hat{\boldsymbol{r}} |\Psi(t)\rangle = \int d^3r' \int d^3r \, \Psi^*(\boldsymbol{r}',t)\boldsymbol{r}\delta(\boldsymbol{r}' - \boldsymbol{r})\Psi(\boldsymbol{r},t) = \int d^3r \, \Psi^*(\boldsymbol{r},t)\boldsymbol{r}\Psi(\boldsymbol{r},t)
\tag{8.59}
$$

and

$$
\langle \Psi(t)| \hat{\boldsymbol{p}} |\Psi(t)\rangle = \int d^3r' \int d^3r \, \Psi^*(\boldsymbol{r}',t)\delta(\boldsymbol{r}' - \boldsymbol{r})\frac{\hbar}{i}\boldsymbol{\nabla}\Psi(\boldsymbol{r},t) = \int d^3 \, \Psi^*(\boldsymbol{r},t)\frac{\hbar}{i}\boldsymbol{\nabla}\Psi(\boldsymbol{r},t)
\tag{8.60}
$$

which reproduces our definition of the coordinate-space matrix expectation values of these operators which we have used so far.

8.3 Heisenberg Representation

So far we have introduced the dynamics of quantum mechanics through the Schrödinger equation, which in abstract form is

$$
\hat{H} |\Psi(t)\rangle = i\hbar\frac{\partial}{\partial t} |\Psi(t)\rangle \, .
\tag{8.61}
$$

Historically, Heisenberg introduced described quantum mechanics in terms of a matrix formulation which we will soon describe. The relationship between these two approaches or representations of quantum mechanics can be easily understood by introducing an alternate representation of the wave equation called the Heisenberg representation.

To see how this works consider a general Schrödinger state which can be represented in Dirac notation as

$$
|\Psi(t)\rangle = \sum_{n=1}^{\infty} c_n |\Psi_n(t)\rangle = \sum_{n=1}^{\infty} c_n e^{-\frac{i}{\hbar}E_n t} |\psi_n\rangle \, ,
\tag{8.62}
$$

where
$$\hat{H} \left| \psi_n \right\rangle = E_n \left| \psi_n \right\rangle . \tag{8.63}$$

Using
$$e^{-\frac{i}{\hbar} E_n t} \left| \psi_n \right\rangle = e^{-\frac{i}{\hbar} \hat{H} t} \left| \psi_n \right\rangle , \tag{8.64}$$

then
$$\left| \Psi(t) \right\rangle = \sum_{n=1}^{\infty} c_n e^{-\frac{i}{\hbar} \hat{H} t} \left| \psi_n \right\rangle = e^{-\frac{i}{\hbar} \hat{H} t} \sum_{n=1}^{\infty} c_n \left| \psi_n \right\rangle = e^{-\frac{i}{\hbar} \hat{H} t} \left| \Psi(0) \right\rangle . \tag{8.65}$$

So,
$$\left| \Psi(t) \right\rangle = e^{-\frac{i}{\hbar} \hat{H} t} \left| \Psi(0) \right\rangle . \tag{8.66}$$

This can be inverted to give
$$\left| \Psi(0) \right\rangle = e^{\frac{i}{\hbar} \hat{H} t} \left| \Psi(t) \right\rangle . \tag{8.67}$$

Note that since this left-hand side of this is independent of time, we can write
$$\left| \Psi(0) \right\rangle = e^{\frac{i}{\hbar} \hat{H} t'} \left| \Psi(t') \right\rangle . \tag{8.68}$$

Substituting this into (8.66) yields
$$\left| \Psi(t) \right\rangle = e^{-\frac{i}{\hbar} \hat{H} t} e^{\frac{i}{\hbar} \hat{H} t'} \left| \Psi(t') \right\rangle = U(t, t') \left| \Psi(t') \right\rangle , \tag{8.69}$$

where
$$\hat{U}(t, t') = e^{-\frac{i}{\hbar} \hat{H} (t - t')} \tag{8.70}$$

is called the *time evolution operator*. The time evolution operator gives the state at time t by acting on the state at any other time t'. Constructing the time evolution operator is equivalent to solving the Schrödinger equation. Note that since the hamiltonian operator is hermitian, the time evolution operator is unitary.

The time-independent state $\left| \Psi(0) \right\rangle$ is called the *Heisenberg wave function* where
$$\left| \Psi_H \right\rangle = \left| \Psi(0) \right\rangle = e^{\frac{i}{\hbar} \hat{H} t} \left| \Psi(t) \right\rangle = \hat{U}^\dagger(t, 0) \left| \Psi(t) \right\rangle . \tag{8.71}$$

This can be inverted to give
$$\left| \Psi(t) \right\rangle = e^{-\frac{i}{\hbar} \hat{H} t} \left| \Psi_H \right\rangle . \tag{8.72}$$

Now consider the expectation value of some operator $\hat{\mathcal{O}}$. Using (8.66) and (8.71) we can write this as
$$\left\langle \mathcal{O}(t) \right\rangle = \left\langle \Psi(t) \right| \hat{\mathcal{O}} \left| \Psi(t) \right\rangle = \left\langle \Psi_H \right| e^{\frac{i}{\hbar} \hat{H} t} \hat{\mathcal{O}} e^{-\frac{i}{\hbar} \hat{H} t} \left| \Psi_H \right\rangle = \left\langle \Psi_H \right| \hat{\mathcal{O}}_H(t) \left| \Psi_H \right\rangle , \tag{8.73}$$

where
$$\hat{\mathcal{O}}_H(t) = e^{\frac{i}{\hbar} \hat{H} t} \hat{\mathcal{O}} e^{-\frac{i}{\hbar} \hat{H} t} \tag{8.74}$$

is the *Heisenberg representation* of the operator $\hat{\mathcal{O}}$. Note that even if the Schrödinger operator is time independent, the Heisenberg operator will be time dependent. From the definition (8.74) we can calculate

$$
\begin{aligned}
\frac{d}{dt}\hat{\mathcal{O}}_H(t) &= \frac{d}{dt}e^{\frac{i}{\hbar}\hat{H}t}\hat{\mathcal{O}}e^{-\frac{i}{\hbar}\hat{H}t} \\
&= \frac{i}{\hbar}\hat{H}e^{\frac{i}{\hbar}\hat{H}t}\hat{\mathcal{O}}e^{-\frac{i}{\hbar}\hat{H}t} + e^{\frac{i}{\hbar}\hat{H}t}\frac{\partial\hat{\mathcal{O}}}{\partial t}e^{-\frac{i}{\hbar}\hat{H}t} - e^{\frac{i}{\hbar}\hat{H}t}\hat{\mathcal{O}}e^{-\frac{i}{\hbar}\hat{H}t}\frac{i}{\hbar}\hat{H} \\
&= \frac{i}{\hbar}[\hat{H},\hat{\mathcal{O}}_H(t)] + e^{\frac{i}{\hbar}\hat{H}t}\frac{\partial\hat{\mathcal{O}}}{\partial t}e^{-\frac{i}{\hbar}\hat{H}t} \tag{8.75}
\end{aligned}
$$

Or,

$$
\frac{d}{dt}\hat{\mathcal{O}}_H(t) = \frac{i}{\hbar}[\hat{H},\hat{\mathcal{O}}_H(t)] + e^{\frac{i}{\hbar}\hat{H}t}\frac{\partial\hat{\mathcal{O}}}{\partial t}e^{-\frac{i}{\hbar}\hat{H}t}, \tag{8.76}
$$

which is called the *Heisenberg equation of motion* for the operator $\hat{\mathcal{O}}$. Note that if the corresponding Schrödinger representation of the operator is time independent and

$$
[\hat{H},\hat{\mathcal{O}}_H(t)] = 0, \tag{8.77}
$$

Then

$$
\frac{d}{dt}\hat{\mathcal{O}}_H(t) = 0. \tag{8.78}
$$

The Heisenberg operator is then a constant of motion.

8.4   Matrix Representation

Another common representation of quantum mechanics is the matrix representation. Consider some general Heisenberg state

$$
|\Psi_H\rangle = |\Psi(0)\rangle = \sum_n c_n |\psi_n\rangle. \tag{8.79}
$$

The expectation value of an operator $\hat{\mathcal{O}}$ in the Heisenberg representation can be written as

$$
\begin{aligned}
\langle\mathcal{O}(t)\rangle &= \langle\Psi_H|\hat{\mathcal{O}}_H(t)|\Psi_H\rangle = \sum_{n'}c_{n'}^*\langle\psi_{n'}|\hat{\mathcal{O}}_H(t)\sum_n c_n|\psi_n\rangle \\
&= \sum_{n'}\sum_n c_{n'}^*\langle\psi_{n'}|\hat{\mathcal{O}}_H(t)|\psi_n\rangle c_n \\
&= \sum_{n'}\sum_n c_{n'}^*\langle\Psi_{n'}(t)|\hat{\mathcal{O}}|\Psi_n(t)\rangle c_n. \tag{8.80}
\end{aligned}
$$

The quantity

$$
\langle\psi_{n'}|\hat{\mathcal{O}}_H(t)|\psi_n\rangle = \langle\Psi_{n'}(t)|\hat{\mathcal{O}}|\Psi_n(t)\rangle = \langle\psi_{n'}|\hat{\mathcal{O}}|\psi_n\rangle e^{\frac{i}{\hbar}(E_{n'}-E_n)t} \tag{8.81}
$$

is the probability amplitude for a transition from state n to state n' at time t. *Transition probability* for transitions from n to n' is then

$$P_{nn'} = \left| \langle \psi_{n'} | \hat{\mathcal{O}} | \psi_n \rangle \right|^2 . \tag{8.82}$$

Note that the expansion coefficients c_n have the role as vector components do in the case of regular vectors. That is, we have an infinite dimensional basis of eigenstates $|\psi_n\rangle$ and the coefficients $c_n = \langle \psi_n | \Psi(0) \rangle$ represent the components of the vector $|\Psi(0)\rangle$ projected onto this basis. This means that one way to represent this vector is as an infinite-dimensional vector of the coefficients. That is

$$\underline{\psi} = \begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ \vdots \end{pmatrix} . \tag{8.83}$$

Any operator $\hat{\mathcal{O}}$ can then be represented as an infinite dimensional square matrix in this basis, where the components of this matrix are defined as

$$\left(\underline{\mathcal{O}} \right)_{n'n} = \langle \psi_{n'} | \hat{\mathcal{O}} | \psi_n \rangle . \tag{8.84}$$

For this reason transition amplitudes of the form $\langle \psi_{n'} | \hat{\mathcal{O}} | \psi_n \rangle$ are usually called *matrix elements* of the operator $\hat{\mathcal{O}}$. For example, the matrix elements of the hamiltonian operator are given by

$$\left(\underline{H} \right)_{n'n} = \langle \psi_{n'} | \hat{H} | \psi_n \rangle = \langle \psi_{n'} | E_n | \psi_n \rangle = E_n \langle \psi_{n'} | E_n | \psi_n \rangle = E_n \delta_{n'n} . \tag{8.85}$$

The hamiltonian matrix is, therefore, diagonal. This will also be true for all of the mutually commuting operators that commute with the hamiltonian and make up set of observables. The matrix representations of hermitian operators with be hermitian matrices and the eigenvalues of any operator can be found in the matrix representation as solutions of the infinite dimensional matrix equation

$$\underline{\mathcal{O}} \, \underline{\psi} = \lambda \underline{\psi} . \tag{8.86}$$

In fact all of the properties and techniques associated with linear matrix algebra will also apply to the matrix representation of quantum mechanics.

One major application of the matrix representation of operators is to operators which have no explicit representation in any other form such as operators for internal degrees of freedom of particles such as spin and isospin.

An additional practical application is to situation where the hamiltonian operator can be written in the form

$$\hat{H} = \hat{H}_0 + \hat{H}_1 , \tag{8.87}$$

141

where H_0 is a hamiltonian for which eigenstates can be obtained easily and H_1 is some additional interaction which complicates the construction of eigenstates for the full hamiltonian. If we define the eigenstates of $\hat{H}_0$ such that

$$\hat{H}_0 \ket{\phi_n} = E^{(0)} \ket{\phi_n} , \tag{8.88}$$

the states $\ket{\phi_n}$ form a complete orthonormal basis forming a hilbert space. In particular these states satisfy the completeness relation

$$\sum_n \ket{\phi_n} \bra{\phi_n} = \hat{1} . \tag{8.89}$$

Now consider the time-independent Schrödinger equation for the full hamiltonian

$$\hat{H} \ket{\psi} = E \ket{\psi} . \tag{8.90}$$

We can now expand the expression by using the completeness relation (8.89) to write

$$\hat{H} \left(\sum_n \ket{\phi_n} \bra{\phi_n} \right) \ket{\psi} = E \ket{\psi} . \tag{8.91}$$

Multiplying this from the left by $\bra{\phi_{n'}}$ gives

$$\sum_n \bra{\phi_{n'}} \hat{H} \ket{\phi_n} \braket{\phi_n | \psi} = E \braket{\phi_{n'} | \psi} . \tag{8.92}$$

We can now represent this in matrix form relative this basis as

$$\underline{\underline{H_\phi}} \, \underline{\psi_\phi} = E \underline{\psi_\phi} \tag{8.93}$$

This is a matrix eigenvalue problem, although one of infinite dimension. For most problems, the contributions to the wave functions of lower energy states will be little affected by basis states with large energies $E_n^{(0)}$. This means that a good numerical approximation can be obtained be retaining only a finite number of basis states leading to a finite-dimensional matrix equation that can be solved numerically. We will devote more attention to various applications of this kind of problem in the future.

Chapter 9

Symmetries, Constants of Motion and Angular Momentum in Quantum Mechanics

9.1 Translations

To translate a coordinate vector is to add a constant vector. We define a translation operator $\mathcal{D}$ such that

$$\mathcal{D}(\boldsymbol{a})\,\boldsymbol{r} = \boldsymbol{r}' = \boldsymbol{r} + \boldsymbol{a}\,. \tag{9.1}$$

This operator obeys the group multiplication property

$$\mathcal{D}(\boldsymbol{a}')\mathcal{D}(\boldsymbol{a})\,\boldsymbol{r} = \mathcal{D}(\boldsymbol{a})\mathcal{D}(\boldsymbol{a}')\,\boldsymbol{r} = \boldsymbol{r} + \boldsymbol{a} + \boldsymbol{a}'\,. \tag{9.2}$$

Clearly, the inverse of this operation is

$$\mathcal{D}^{-1}(\boldsymbol{a}) = \mathcal{D}(-\boldsymbol{a})\,. \tag{9.3}$$

We now want consider a quantum translation operator that when acting on a wave function gives

$$\hat{\mathcal{D}}(\boldsymbol{a})\Psi(\boldsymbol{r}, t) = \Psi'(\boldsymbol{r}, t) \tag{9.4}$$

Figure 9.1 shows a schematic representation of the translation of a wave function. Clearly, since the wave function is displaced by a constant amount for all values of the position, the shape of the wave function is the same and for each r there is an r' with same value of the wave function. Therefore,

$$\Psi'(\boldsymbol{r}', t) = \Psi(\boldsymbol{r}, t)\,, \tag{9.5}$$

where $\boldsymbol{r}'$ is given by (9.1). This means that we can write

$$\Psi'(\boldsymbol{r}, t) = \Psi(\mathcal{D}^{-1}(\boldsymbol{a})\,\boldsymbol{r}, t) = \Psi(\boldsymbol{r} - \boldsymbol{a}, t)\,. \tag{9.6}$$
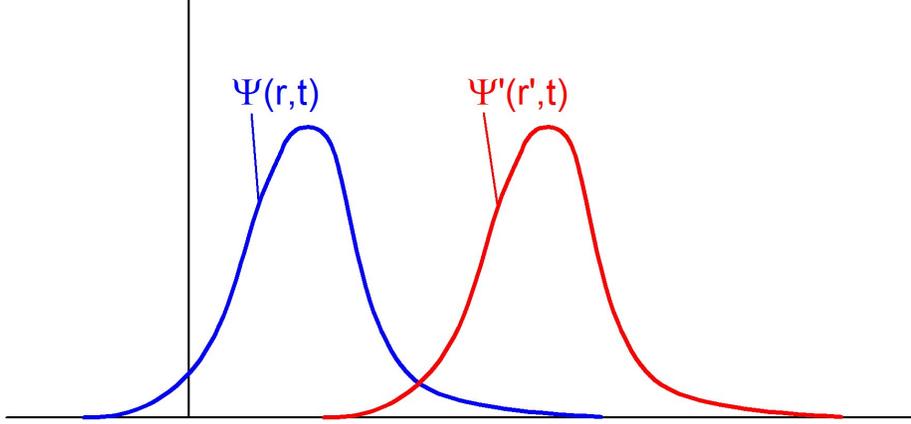
Figure 9.1: Schematic representation of the translation of a wave function.

Now consider the case where the translation is infinitesimal. That is, $\boldsymbol{a} \to \delta\boldsymbol{a}$, where $|\delta\boldsymbol{a}| << 1$. We can then expand the wave function in a three-dimensional Taylor series about $\delta\boldsymbol{a} = \boldsymbol{0}$. This gives

$$\Psi'(\boldsymbol{r}, t) = \Psi(\boldsymbol{r} - \boldsymbol{a}, t) \cong \Psi(\boldsymbol{r}, t) - \delta\boldsymbol{a} \cdot \boldsymbol{\nabla}\Psi(\boldsymbol{r}, t) = \Psi(\boldsymbol{r}, t) - \frac{i}{\hbar}\delta\boldsymbol{a} \cdot \hat{\boldsymbol{p}}\Psi(\boldsymbol{r}, t). \quad (9.7)$$

So to first order,

$$\hat{\mathcal{D}}(\delta\boldsymbol{a}) \cong \hat{1} - \frac{i}{\hbar}\delta\boldsymbol{a} \cdot \hat{\boldsymbol{p}}. \quad (9.8)$$

A finite translation can be compounded of a large number of infinitesimal translations using the group multiplication property (9.2). If we define

$$\delta\boldsymbol{a} = \frac{\boldsymbol{a}}{N}, \quad (9.9)$$

then

$$\hat{\mathcal{D}}(\boldsymbol{a}) \cong \left(\hat{1} - \frac{i}{\hbar}\frac{\boldsymbol{a}}{N} \cdot \hat{\boldsymbol{p}}\right)^{N}. \quad (9.10)$$

This result will be exact in the limit $N \to \infty$. We can use the identity

$$\lim_{N\to\infty} \left(1 + \frac{a}{N}\right)^{N} = e^{a} \quad (9.11)$$

To right the translation operator as

$$\hat{\mathcal{D}}(\boldsymbol{a}) = e^{-\frac{i}{\hbar}\boldsymbol{a}\cdot\hat{\boldsymbol{p}}}. \quad (9.12)$$

Note that since $\boldsymbol{a}$ is real and $\hat{\boldsymbol{p}}$ is hermitian, the translation operator is unitary. The momentum operator is called the *generator* of translations while the components of the vector $\boldsymbol{a}$ are parameters that describe a particular translation.

Note that if $[\hat{\boldsymbol{p}}, \hat{H}] = \mathbf{0}$, then

$$\hat{H}' = \hat{\mathcal{D}}^\dagger(\boldsymbol{a})\hat{H}\hat{\mathcal{D}}(\boldsymbol{a}) = \hat{H}\,. \tag{9.13}$$

The hamiltonian is then said to be *invariant under translations*. From the previous chapter, we also recognize that this is the condition for $\hat{\boldsymbol{p}}$ to be a constant of motion and for the hamiltonian and momentum to have the same eigenfunctions. An example of hamiltonian is invariant under translations is the free hamiltonian

$$\hat{H} = \frac{\hat{\boldsymbol{p}}^2}{2m}\,, \tag{9.14}$$

where the eigenstates of the hamiltonian are plane wave which are also eigenstates of the momentum operator.

9.2 Time Translation

Time translation is a displacement in time. We can define a time translation operator such that

$$\mathcal{T}(\tau)\,t = t' = t + \tau\,. \tag{9.15}$$

We now want consider a quantum time-translation operator that when acting on a wave function gives

$$\hat{\mathcal{T}}(\tau)\,|\Psi(t)\rangle = |\Psi'(t)\rangle \tag{9.16}$$

Since we are simply translating all times by a constant value, we expect that

$$|\Psi'(t')\rangle = |\Psi(t)\rangle\,, \tag{9.17}$$

where t' is given by (9.15). This means that we can write

$$|\Psi'(t)\rangle = \left|\Psi(\mathcal{T}^{-1}(\tau)\,t)\right\rangle = |\Psi(t - \tau)\rangle\,. \tag{9.18}$$

We can use the time evolution operator (8.70) to rewrite this as

$$|\Psi'(t)\rangle = \hat{U}(t - \tau, t)\,|\Psi(t)\rangle\,. \tag{9.19}$$

So the time translation operator is

$$\hat{\mathcal{T}}(\tau) = \hat{U}(t - \tau, t) = e^{-\frac{i}{\hbar}\hat{H}(t-\tau-t)} = e^{\frac{i}{\hbar}\hat{H}\tau}\,. \tag{9.20}$$

Since $\hat{H}$ is hermitian, $\hat{\mathcal{T}}(\tau)$ is unitary. The generator of time translations is the hamiltonian operator.

9.3 Rotations

9.3.1 Rotations of a Vector

A rotation of a position vector can be written in matrix form as

$$\boldsymbol{r}' = \boldsymbol{R}\boldsymbol{r}\,, \tag{9.21}$$

where in this case $\boldsymbol{R}$ is a three-dimensional rotation matrix. We can also write this in component form as

$$r_i' = R_{ij}r_j \tag{9.22}$$

where the Einstein summation convention is assumed to give a sum of the index j. In this discussion we will assume that the rotation is an *active* rotation where the rotation actually rotates a vector from one direction to another rather than a passive rotation where the vector remains fixed and the coordinate system is rotated. Furthermore, parameters will be chosen such that the rotation is right handed for positive angles. Taking partial derivatives of both sides of (9.22) gives

$$R_{ij} = \frac{\partial r_i'}{\partial r_j}\,. \tag{9.23}$$

The length of a vector remains unchanged under rotation or is said to be *invariant* under rotation. We can use this to determine one of the properties of the rotation matrix. This means that

$$r_j r_j = r_i' r_i' = R_{ij} r_j R_{ik} r_k = R_{ij} R_{ik} r_j r_k\,. \tag{9.24}$$

This requires that

$$\delta_{jk} = R_{ij} R_{ik} = R_{ji}^T R_{ik}\,, \tag{9.25}$$

or in other words

$$\boldsymbol{R}^T \boldsymbol{R} = \boldsymbol{1}\,. \tag{9.26}$$

This implies that

$$\boldsymbol{R}^T = \boldsymbol{R}^{-1}\,. \tag{9.27}$$

So the rotation matrix is an orthogonal matrix.

To determine the explicit form of $\boldsymbol{R}$, it is useful to consider the effects of infinitesimal rotations. Since rotation about some axis through a zero angle transforms any vector into itself, such a rotation is represented by the unit matrix $\boldsymbol{1}$. Any infinitesimal rotation is, therefore, different from the identity by addition of a matrix where all of its components are infinitesimal. That is,

$$\boldsymbol{R} = \boldsymbol{1} + \boldsymbol{\mathcal{R}}\,, \tag{9.28}$$

where alll of the components of $\boldsymbol{\mathcal{R}}$ are infinitesimally small. We can now substitute this into (9.26) to give

$$1 = (1 + \boldsymbol{\mathcal{R}}^T)(1 + \boldsymbol{\mathcal{R}}) \cong 1 + \boldsymbol{\mathcal{R}} + \boldsymbol{\mathcal{R}}^T. \tag{9.29}$$

This means that

$$\boldsymbol{\mathcal{R}}^T = -\boldsymbol{\mathcal{R}}. \tag{9.30}$$

So $\boldsymbol{\mathcal{R}}$ is an antisymmetrix matrix which will have the form

$$\boldsymbol{\mathcal{R}} = \begin{pmatrix} 0 & \mathcal{R}_{12} & \mathcal{R}_{13} \\ -\mathcal{R}_{12} & 0 & \mathcal{R}_{23} \\ -\mathcal{R}_{13} & -\mathcal{R}_{23} & 0 \end{pmatrix}. \tag{9.31}$$

This matrix is, therefore, determined by three independent real numbers and can be written as a linear combination of three linearly independent antisymmetric matrices. A convenient choice for the set of matrices is the set of antisymmetric, hermitian matrices

$$J_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \tag{9.32}$$

$$J_2 = \begin{pmatrix} 0 & 0 & i \\ 0 & 0 & 0 \\ -i & 0 & 0 \end{pmatrix} \tag{9.33}$$

and

$$J_3 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{9.34}$$

It is easily verified that these matrices satisfy the commutation relations

$$[J_i, J_j] = i\epsilon_{ijk} J_k, \tag{9.35}$$

where ϵ_{ijk} the Levi-Civita tensor density or the completely antisymmetric tensor of rank 3. It is defined such that $\epsilon_{123} = \epsilon_{231} = \epsilon_{312} = 1$, $\epsilon_{213} = \epsilon_{321} = \epsilon_{132} = -1$ and any combination of indices where any two have the same value is 0. This is a completely antisymmetric tensor since the interchange of any two indices changes the sign. This tensor has some very useful properties. First, any cross product can be written as

$$(\boldsymbol{A} \times \boldsymbol{B})_i = \epsilon_{ijk} A_j B_k. \tag{9.36}$$

The matrix for infinitesimal rotations can then be parameterized as

$$\boldsymbol{R} = 1 - i\delta\boldsymbol{\theta} \cdot \boldsymbol{J}, \tag{9.37}$$

where a factor of i is required to make the rotation matrix real, and the minus sign will be needed to make the rotation right handed. A convenient notation is to write

$$\delta\boldsymbol{\theta} = \boldsymbol{n}\delta\theta\,, \tag{9.38}$$

where $\boldsymbol{n}$ is a unit vector which gives the direction about which the rotation takes place, and $\delta\theta$ is an infinitesimal angle of rotation.

To construct the rotation matrix for finite rotations, we need to determine the group properties of the rotations. First, note that the composition of two rotations about different directions do not commute and are equivalent to a single rotation about a third direction. However, two rotations about the same direction will commute and will satisfy the composition rule

$$\boldsymbol{R}(\boldsymbol{n},\theta_1)\boldsymbol{R}(\boldsymbol{n},\theta_2) = \boldsymbol{R}(\boldsymbol{n},\theta_1+\theta_2)\,. \tag{9.39}$$

Using this rule we can construct a rotation through a finite angle as

$$\boldsymbol{R}(\boldsymbol{n},\theta) = \lim_{N\to\infty}\left(1-i\frac{\theta}{N}\boldsymbol{n}\cdot\boldsymbol{J}\right)^N = e^{-i\theta\boldsymbol{n}\cdot\boldsymbol{J}} = e^{-i\boldsymbol{\theta}\cdot\boldsymbol{J}}\,. \tag{9.40}$$

To check that this does indeed give the rotation matrix, consider the case where $\boldsymbol{n} = \boldsymbol{e}_3$ and $\theta = \phi$. That is, a rotation through the angle ϕ about the z-axis. This calculation can be simplified by noting that

$$J_3^2 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}^2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}\,, \tag{9.41}$$

which implies that

$$J_3^{2k-1} = J_3 \tag{9.42}$$

and

$$J_3^{2k} = J_3^2\,. \tag{9.43}$$

Now, consider the rotation

$$\boldsymbol{R}(\boldsymbol{z},\phi) = e^{-i\phi J_3} = \sum_{j=0}^{\infty}\frac{(-i\phi J_3)^j}{j!} = 1 + \sum_{j=1}^{\infty}\frac{(-i\phi J_3)^j}{j!}\,. \tag{9.44}$$

Separating the series into even and odd powers gives

$$
\begin{aligned}
\boldsymbol{R}(\boldsymbol{z},\phi) &= 1 + \sum_{k=1}^{\infty}\frac{(-i\phi J_3)^{2k}}{(2k)!} + \sum_{k=1}^{\infty}\frac{(-i\phi J_3)^{2k-1}}{(2k-1)!} \\
&= 1 + \sum_{k=1}^{\infty}\frac{(-1)^k\phi^{2k}}{(2k)!}J_3^{2k} + i\sum_{k=1}^{\infty}\frac{(-1)^k\phi^{2k-1}}{(2k-1)!}J_3^{2k-1} \\
&= 1 + \left(\sum_{k=1}^{\infty}\frac{(-1)^k\phi^{2k}}{(2k)!}\right)J_3^2 + i\left(\sum_{k=1}^{\infty}\frac{(-1)^k\phi^{2k-1}}{(2k-1)!}\right)J_3 \\
&= 1 + (\cos\phi - 1)J_3^2 - i\sin\phi\,J_3\,.
\end{aligned} \tag{9.45}
$$

In matrix form this becomes

$$\boldsymbol{R}(\boldsymbol{z}, \phi) = \begin{pmatrix} \cos\phi & -\sin\phi & 0 \\ \sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{9.46}$$

which is the correct matrix for a right-handed active rotation about the z-axis. Rotations about the x- and y-axes can be found in a similar fashion.

9.3.2 Rotations of Wave Functions

We can now proceed to examine the rotation operator for wave functions in a manner similar to that used in describing translations. As before, if we define the operation of a rotation operator on a wave function

$$\hat{R}(\boldsymbol{n}, \theta)\Psi(\boldsymbol{r}, t) = \Psi'(\boldsymbol{r}, t), \tag{9.47}$$

we expect that

$$\Psi'(\boldsymbol{r}', t) = \Psi(\boldsymbol{r}, t), \tag{9.48}$$

where $\boldsymbol{r}' = \boldsymbol{R}(\boldsymbol{n}, \theta)\,\boldsymbol{r}$. So,

$$\Psi'(\boldsymbol{r}, t) = \Psi(\boldsymbol{R}^{-1}(\boldsymbol{n}, \theta)\,\boldsymbol{r}, t), . \tag{9.49}$$

From the group composition property (9.39), we can see that

$$\boldsymbol{R}^{-1}(\boldsymbol{n}, \theta) = \boldsymbol{R}(\boldsymbol{n}, -\theta). \tag{9.50}$$

So, for a rotation about $\boldsymbol{n}$ through and infinitesimal angle $\delta\theta$,

$$\boldsymbol{R}^{-1}(\boldsymbol{n}, \delta\theta)\,\boldsymbol{r} = \boldsymbol{R}(\boldsymbol{n}, -\delta\theta)\,\boldsymbol{r} = (1 + i\delta\theta\boldsymbol{n} \cdot \boldsymbol{J})\boldsymbol{r} = \boldsymbol{r} + \delta\boldsymbol{r}, \tag{9.51}$$

where

$$\delta\boldsymbol{r} = i\delta\theta\boldsymbol{n} \cdot \boldsymbol{J}\,\boldsymbol{r} = i\delta\boldsymbol{\theta} \cdot \boldsymbol{J}\,\boldsymbol{r}. \tag{9.52}$$

We can now write

$$\Psi'(\boldsymbol{r}, t) = \Psi(\boldsymbol{r} + \delta\boldsymbol{r}, t) \cong \Psi(\boldsymbol{r}, t) + \delta r_k \frac{\partial}{\partial r_k}\Psi(\boldsymbol{r}, t). \tag{9.53}$$

Therefore,

$$\hat{R}(\boldsymbol{n}, \delta\theta) \cong 1 + \delta r_k \frac{\partial}{\partial r_k}. \tag{9.54}$$

It is useful to rewrite the second term of this expression. Consider

$$\begin{aligned} \delta r_k \frac{\partial}{\partial x_k} &= i(\delta\boldsymbol{\theta} \cdot \boldsymbol{J}\,\boldsymbol{r})_k \frac{\partial}{\partial r_k} = i(\delta\boldsymbol{\theta} \cdot \boldsymbol{J})_{kl}\, r_l \frac{\partial}{\partial r_k} = i(\delta\theta_j J_j)_{kl}\, r_l \frac{\partial}{\partial r_k} \\ &= i\delta\theta_j (J_j)_{kl}\, r_l \frac{\partial}{\partial r_k} \end{aligned} \tag{9.55}$$

149

Careful examination of the matrices J_j shows that their components can be written as

$$(J_j)_{kl} = -i\epsilon_{jkl} \,. \tag{9.56}$$

Therefore,

$$
\begin{aligned}
\delta r_k \frac{\partial}{\partial r_k} &= \delta\theta_j \epsilon_{jkl}\, r_l \frac{\partial}{\partial r_k} = -\delta\theta_j \epsilon_{jlk}\, r_l \frac{\partial}{\partial r_k} = -\frac{i}{\hbar}\delta\theta_j \epsilon_{jlk}\, r_l \frac{\hbar}{i}\frac{\partial}{\partial r_k} \\
&= -\frac{i}{\hbar}\delta\theta_j \epsilon_{jlk}\, r_l \hat{p}_k = -\frac{i}{\hbar}\delta\theta_j (\boldsymbol{r} \times \hat{\boldsymbol{p}})_j = -\frac{i}{\hbar}\delta\boldsymbol{\theta} \cdot (\boldsymbol{r} \times \hat{\boldsymbol{p}}) \,. \tag{9.57}
\end{aligned}
$$

We can now define the orbital angular momentum operator as

$$\hat{\boldsymbol{L}} = \hat{\boldsymbol{r}} \times \hat{\boldsymbol{p}} \,. \tag{9.58}$$

This follows the canonical quantization procedure by replace the coordinates and momenta in the definition of classical orbital angular momentum with the corresponding quantum operators. We know have

$$\hat{R}(\boldsymbol{n}, \delta\theta) \cong \boldsymbol{1} - \frac{i}{\hbar}\delta\theta\boldsymbol{n} \cdot \hat{\boldsymbol{L}} \,. \tag{9.59}$$

The rotation operator for a finite rotation is then given by the composition of an infinite number of infinitesimal rotations as

$$\hat{R}(\boldsymbol{n}, \theta) = \lim_{N\to\infty} \left(\boldsymbol{1} - \frac{i}{\hbar}\frac{\theta}{N}\boldsymbol{n} \cdot \hat{\boldsymbol{L}}\right)^N = e^{-\frac{i}{\hbar}\theta\boldsymbol{n}\cdot\hat{\boldsymbol{L}}} = e^{-\frac{i}{\hbar}\boldsymbol{\theta}\cdot\hat{\boldsymbol{L}}} \,. \tag{9.60}$$

The angular momentum operator is, therefore, the generator of rotations. Since both $\hat{\boldsymbol{r}}$ and $\hat{\boldsymbol{p}}$ are hermitian, $\hat{\boldsymbol{L}}$ is also hermitian. Therefore, the rotation operator is unitary.

If $[\hat{L}_i, \hat{H}] = 0$ for $i = 1, 2, 3$ then

$$\hat{H}' = \hat{R}(\boldsymbol{n}, \theta)\hat{H}\hat{R}^{-1}(\boldsymbol{n}, \theta) = \hat{H} \tag{9.61}$$

for all $\boldsymbol{\theta}$ and, therefore, the hamiltonian is invariant under rotations. This commutation relation also implies that the components of angular momentum operator are constants of motion. It is easy to show that the hamiltonian for a particle in a central potential is rotationally invariant.

Now consider the definition of the angular momentum operator in component form

$$\hat{L}_i = \epsilon_{ijk}\hat{r}_j\hat{p}_k \,. \tag{9.62}$$

Using the identity

$$\epsilon_{ijk}\epsilon_{ilm} = \delta_{jl}\delta_{km} - \delta_{jm}\delta_{kl} \,, \tag{9.63}$$

we can write

$$
\begin{aligned}
\epsilon_{ijk}\hat{L}_i &= \epsilon_{ijk}\epsilon_{ilm}\hat{r}_l\hat{p}_m = \epsilon_{ijk}\epsilon_{ilm}\hat{r}_l\hat{p}_m \\
&= (\delta_{jl}\delta_{km} - \delta_{jm}\delta_{kl})\,\hat{r}_l\hat{p}_m = \hat{r}_j\hat{p}_k - \hat{r}_k\hat{p}_j \,. \tag{9.64}
\end{aligned}
$$

150

The commutation relations for the components of the angular momentum are given by

$$
\begin{aligned}
[\hat{L}_i, \hat{L}_j] &= [\epsilon_{ikl}\hat{r}_k\hat{p}_l, \epsilon_{jmn}\hat{r}_m\hat{p}_n] = \epsilon_{ikl}\epsilon_{jmn}[\hat{r}_k\hat{p}_l, \hat{r}_m\hat{p}_n] \\
&= \epsilon_{ikl}\epsilon_{jmn}(\hat{r}_k\hat{p}_l\hat{r}_m\hat{p}_n - \hat{r}_m\hat{p}_n\hat{r}_k\hat{p}_l) \, .
\end{aligned}
\tag{9.65}
$$

The canonical commutation relations can be used to write

$$
\begin{aligned}
\hat{r}_k\hat{p}_l\hat{r}_m\hat{p}_n &= \hat{r}_k(\hat{r}_m\hat{p}_l - i\hbar\delta_{lm})\hat{p}_n = \hat{r}_k\hat{r}_m\hat{p}_l\hat{p}_n - i\hbar\delta_{lm}\hat{r}_k\hat{p}_n \\
&= \hat{r}_m\hat{r}_k\hat{p}_n\hat{p}_l - i\hbar\delta_{lm}\hat{r}_k\hat{p}_n = \hat{r}_m(\hat{p}_n\hat{r}_k + i\hbar\delta_{kn})\hat{p}_l - i\hbar\delta_{lm}\hat{r}_k\hat{p}_n \\
&= \hat{r}_m\hat{p}_n\hat{r}_k\hat{p}_l + i\hbar\delta_{kn}\hat{r}_m\hat{p}_l - i\hbar\delta_{lm}\hat{r}_k\hat{p}_n \, .
\end{aligned}
\tag{9.66}
$$

Substituting this back into (9.65) gives

$$
\begin{aligned}
[\hat{L}_i, \hat{L}_j] &= \epsilon_{ikl}\epsilon_{jmn}(i\hbar\delta_{kn}\hat{r}_m\hat{p}_l - i\hbar\delta_{lm}\hat{r}_k\hat{p}_n) \\
&= i\hbar\epsilon_{ikl}\epsilon_{jmn}(\delta_{kn}\hat{r}_m\hat{p}_l - \delta_{lm}\hat{r}_k\hat{p}_n) \, .
\end{aligned}
\tag{9.67}
$$

Since k, l, m and n are dummy indices, we can rename them with out changing the result. In the second term we will make the exchanges in labels $k \leftrightarrow l$ and $m \leftrightarrow n$ giving

$$
\begin{aligned}
[\hat{L}_i, \hat{L}_j] &= i\hbar(\epsilon_{ikl}\epsilon_{jmn}\delta_{kn}\hat{r}_m\hat{p}_l - \epsilon_{ilk}\epsilon_{jnm}\delta_{kn}\hat{r}_l\hat{p}_m) \\
&= i\hbar(\epsilon_{ikl}\epsilon_{jmn}\delta_{kn}\hat{r}_m\hat{p}_l - (-\epsilon_{ikl})(-\epsilon_{jmn})\delta_{kn}\hat{r}_l\hat{p}_m) \\
&= i\hbar\epsilon_{ikl}\epsilon_{jmn}\delta_{kn}(\hat{r}_m\hat{p}_l - \hat{r}_l\hat{p}_m) \\
&= i\hbar\epsilon_{ikl}\epsilon_{jmk}(\hat{r}_m\hat{p}_l - \hat{r}_l\hat{p}_m) \, .
\end{aligned}
\tag{9.68}
$$

Using (9.64) to write

$$
\hat{r}_m\hat{p}_l - \hat{r}_l\hat{p}_m = \epsilon_{mlq}\hat{L}_q \, ,
\tag{9.69}
$$

then (9.68) can be rewritten as

$$
\begin{aligned}
[\hat{L}_i, \hat{L}_j] &= i\hbar\epsilon_{ikl}\epsilon_{jmk}\epsilon_{mlq}\hat{L}_q \\
&= i\hbar\epsilon_{ikl}\epsilon_{jmk}\epsilon_{mlq}\hat{L}_q \\
&= i\hbar\epsilon_{ikl}\epsilon_{mkj}\epsilon_{mlq}\hat{L}_q \\
&= i\hbar\epsilon_{ikl}(\delta_{kl}\delta_{jq} - \delta_{kq}\delta_{jl})\hat{L}_q \\
&= i\hbar\epsilon_{ikl}(-\delta_{kq}\delta_{jl})\hat{L}_q \\
&= -i\hbar\epsilon_{ikj}\hat{L}_k = i\hbar\epsilon_{ijk}\hat{L}_k \, .
\end{aligned}
\tag{9.70}
$$

Since for any i and j there can be at most one value of k that will contribute to the sum, we can summarize this as

$$
[\hat{L}_i, \hat{L}_j] = i\hbar\epsilon_{ijk}\hat{L}_k \, .
\tag{9.71}
$$

Note that, aside from the factor of $\hbar$, this is of the same form as the commutation relation for the matrices J_i.

Recall that if $\hat{H}$ is rotationally invariant $[\hat{L}_i, \hat{H}] = 0$. This means that the components of the angular momentum operator can have eigenstates in common with the hamiltonian. However, since the components of the angular momentum are not mutually commuting, the will not all have common eigenstates. This means that eigenvalue of only one of the components of the angular momentum operator can be used to label states which are also energy eigenstates. However, we have seen that the eigenstates of a particle in a central potential, which is rotationally invariant, that the states must be labelled with three quantum numbers. This means that there must be another operator which will also commute with the hamiltonian and the chosen component of the angular momentum operator. However, using (9.71) it can be easily proved that

$$[\hat{L}_i, \hat{\boldsymbol{L}}^2] = 0 \,. \tag{9.72}$$

This implies that the three operators whose eigenvalues can be used to label the eigenstates are $\hat{H}$, $\hat{\boldsymbol{L}}^2$ and one component of the angular momentum operator which is conventionally chosen to be $\hat{L}_3 = \hat{L}_z$. To see that is indeed the case we need to construct the angular momentum operators in coordinate space using spherical coordinates.

9.4 The Angular Momentum Operators in Spherical Coordinates

In order to represent the angular momentum operators in spherical coordinates, we need to do a little mathematics. From Fig. 7.1, we can represent the cartesian coordinates of the vector r in terms of the spherical coordinates as

$$x = r \cos\phi \sin\theta \tag{9.73}$$
$$y = r \sin\phi \sin\theta \tag{9.74}$$
$$z = r \cos\theta \,. \tag{9.75}$$

This set of equations can be inverted to give the spherical coordinates in terms of the cartesian coordinates as

$$r = \sqrt{x^2 + y^2 + z^2} \tag{9.76}$$
$$\theta = \cos^{-1}\frac{z}{r} \tag{9.77}$$
$$\phi = \tan^{-1}\frac{y}{x} \,. \tag{9.78}$$

It is also convenient to define a new set of orthogonal unit vectors appropriate to the spherical coordinate system. These are

$$\boldsymbol{e}_r = \frac{\boldsymbol{r}}{r} = \cos\phi\sin\theta\,\boldsymbol{e}_x + \sin\phi\sin\theta\,\boldsymbol{e}_y + \cos\theta\,\boldsymbol{e}_z \tag{9.79}$$

$$\boldsymbol{e}_\theta = \frac{\partial\boldsymbol{e}_r}{\partial\theta} = \cos\phi\cos\theta\,\boldsymbol{e}_x + \sin\phi\cos\theta\,\boldsymbol{e}_y - \sin\theta\,\boldsymbol{e}_z \tag{9.80}$$

$$\boldsymbol{e}_\phi = \frac{1}{\sin\theta}\frac{\partial\boldsymbol{e}_r}{\partial\phi} = -\sin\phi\,\boldsymbol{e}_x + \cos\phi\,\boldsymbol{e}_y \tag{9.81}$$

and are shown in Fig. 7.1.

We can now rewrite the gradient operator by using the chain rule for partial differentiation as

$$
\begin{aligned}
\boldsymbol{\nabla} &= \boldsymbol{e}_x\frac{\partial}{\partial x} + \boldsymbol{e}_y\frac{\partial}{\partial y} + \boldsymbol{e}_z\frac{\partial}{\partial z} \\
&= \boldsymbol{e}_x\left(\frac{\partial r}{\partial x}\frac{\partial}{\partial r} + \frac{\partial\theta}{\partial x}\frac{\partial}{\partial\theta} + \frac{\partial\phi}{\partial x}\frac{\partial}{\partial\phi}\right) + \boldsymbol{e}_y\left(\frac{\partial r}{\partial y}\frac{\partial}{\partial r} + \frac{\partial\theta}{\partial y}\frac{\partial}{\partial\theta} + \frac{\partial\phi}{\partial y}\frac{\partial}{\partial\phi}\right) \\
&\quad + \boldsymbol{e}_z\left(\frac{\partial r}{\partial z}\frac{\partial}{\partial r} + \frac{\partial\theta}{\partial z}\frac{\partial}{\partial\theta} + \frac{\partial\phi}{\partial z}\frac{\partial}{\partial\phi}\right).
\end{aligned}
\tag{9.82}
$$

The various partial derivatives can be evaluated using (9.76), (9.77) and (9.78) to give

$$\frac{\partial r}{\partial x} = \frac{x}{r} = \cos\phi\sin\theta \tag{9.83}$$

$$\frac{\partial r}{\partial y} = \sin\phi\sin\theta \tag{9.84}$$

$$\frac{\partial r}{\partial z} = \cos\theta \tag{9.85}$$

$$\frac{\partial\theta}{\partial x} = \frac{1}{r}\cos\phi\cos\theta \tag{9.86}$$

$$\frac{\partial\theta}{\partial y} = \frac{1}{r}\sin\phi\cos\theta \tag{9.87}$$

$$\frac{\partial\theta}{\partial z} = -\frac{1}{r}\sin\theta \tag{9.88}$$

$$\frac{\partial\phi}{\partial x} = -\frac{\sin\phi}{r\sin\theta} \tag{9.89}$$

$$\frac{\partial\phi}{\partial y} = \frac{\cos\phi}{r\sin\theta} \tag{9.90}$$

$$\frac{\partial\phi}{\partial z} = 0. \tag{9.91}$$

These along with the definitions of the spherical unit vectors can be used to write the gradient as

$$\boldsymbol{\nabla} = \boldsymbol{e}_r\frac{\partial}{\partial r} + \frac{\boldsymbol{e}_\theta}{r}\frac{\partial}{\partial\theta} + \frac{\boldsymbol{e}_\phi}{r\sin\theta}\frac{\partial}{\partial\phi}. \tag{9.92}$$

We can now write the angular momentum operator as

$$
\begin{aligned}
\hat{\boldsymbol{L}} &= \boldsymbol{r} \times \frac{\hbar}{i}\boldsymbol{\nabla} = \frac{\hbar}{i}r\,\boldsymbol{e}_r \times \left(\boldsymbol{e}_r\frac{\partial}{\partial r} + \frac{\boldsymbol{e}_\theta}{r}\frac{\partial}{\partial \theta} + \frac{\boldsymbol{e}_\phi}{r\sin\theta}\frac{\partial}{\partial \phi} \right) \\
&= \frac{\hbar}{i}r\left(\frac{\boldsymbol{e}_\phi}{r}\frac{\partial}{\partial \theta} - \frac{\boldsymbol{e}_\theta}{r\sin\theta}\frac{\partial}{\partial \phi} \right) = \frac{\hbar}{i}\left(\boldsymbol{e}_\phi\frac{\partial}{\partial \theta} - \frac{\boldsymbol{e}_\theta}{\sin\theta}\frac{\partial}{\partial \phi} \right) \\
&= \frac{\hbar}{i}\left[\left(-\sin\phi\,\boldsymbol{e}_x + \cos\phi\,\boldsymbol{e}_y\right)\frac{\partial}{\partial \theta} - \frac{\cos\phi\cos\theta\,\boldsymbol{e}_x + \sin\phi\cos\theta\,\boldsymbol{e}_y - \sin\theta\,\boldsymbol{e}_z}{\sin\theta}\frac{\partial}{\partial \phi} \right] \\
&= \frac{\hbar}{i}\left[\boldsymbol{e}_x\left(-\sin\phi\frac{\partial}{\partial \theta} - \frac{\cos\phi\cos\theta}{\sin\theta}\frac{\partial}{\partial \phi}\right) + \boldsymbol{e}_y\left(\cos\phi\frac{\partial}{\partial \theta} - \frac{\sin\phi\cos\theta}{\sin\theta}\frac{\partial}{\partial \phi}\right) + \boldsymbol{e}_z\frac{\partial}{\partial \phi} \right].
\end{aligned}
$$
$$(9.93)$$

From this we can extract the cartesian components of the angular momentum vector in terms of spherical coordinates yielding

$$
\hat{L}_x = \frac{\hbar}{i}\left(-\sin\phi\frac{\partial}{\partial \theta} - \frac{\cos\phi\cos\theta}{\sin\theta}\frac{\partial}{\partial \phi} \right) \tag{9.94}
$$

$$
\hat{L}_y = \frac{\hbar}{i}\left(\cos\phi\frac{\partial}{\partial \theta} - \frac{\sin\phi\cos\theta}{\sin\theta}\frac{\partial}{\partial \phi} \right) \tag{9.95}
$$

$$
\hat{L}_z = \frac{\hbar}{i}\frac{\partial}{\partial \phi}. \tag{9.96}
$$

From the definition of the spherical harmonics (7.40), we find that

$$
\hat{L}_z Y_{lm}(\theta,\phi) = \frac{\hbar}{i}\frac{\partial}{\partial \phi}Y_{lm}(\theta,\phi) = \frac{\hbar}{i}im Y_{lm}(\theta,\phi) = \hbar m Y_{lm}(\theta,\phi). \tag{9.97}
$$

So m is related to the eigenvalue of $\hat{L}_z$. By choosing our coordinate system such that the azimuthal angle ϕ is defined in terms of rotation about the z axis we have implicitly chosen a *quantization axis* such that L_z is chosen as the one allowed component of the angular momentum to used to label the eigenstates of the Schrödinger equation.

The spherical representation of the components of the angular momentum opera-

tor can be used to express the square of the angular momentum operator as

$$
\begin{aligned}
\hat{\boldsymbol{L}}^2 &= -\hbar^2 \left[\sin^2\phi \frac{\partial^2}{\partial\theta^2} + \sin\phi\cos\phi \frac{\partial\cot\theta}{\partial\theta}\frac{\partial}{\partial\theta} + \frac{\sin\phi\cos\phi\cos\theta}{\sin\theta}\frac{\partial^2}{\partial\theta\partial\phi} \right. \\
&\quad + \frac{\cos^2\phi\cos\theta}{\sin\theta}\frac{\partial}{\partial\theta} + \frac{\sin\phi\cos\phi\cos\theta}{\sin\theta}\frac{\partial^2}{\partial\phi\partial\theta} - \frac{\sin\phi\cos\phi\cos^2\theta}{\sin^2\theta}\frac{\partial}{\partial\phi} \\
&\quad + \frac{\cos^2\phi\cos^2\theta}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2} \\
&\quad + \cos^2\phi\frac{\partial^2}{\partial\theta^2} - \sin\phi\cos\phi\frac{\partial\cot\theta}{\partial\theta}\frac{\partial}{\partial\theta} - \frac{\sin\phi\cos\phi\cos\theta}{\sin\theta}\frac{\partial^2}{\partial\theta\partial\phi} \\
&\quad + \frac{\sin^2\phi\cos\theta}{\sin\theta}\frac{\partial}{\partial\theta} - \frac{\sin\phi\cos\phi\cos\theta}{\sin\theta}\frac{\partial^2}{\partial\phi\partial\theta} + \frac{\sin\phi\cos\phi\cos^2\theta}{\sin^2\theta}\frac{\partial}{\partial\phi} \\
&\quad \left. + \frac{\sin^2\phi\cos^2\theta}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2} + \frac{\partial^2}{\partial\phi^2} \right].
\end{aligned}
\tag{9.98}
$$

The colored terms cancel in pairs and the remaining terms can be simplified to give

$$
\hat{\boldsymbol{L}}^2 = -\hbar^2 \left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2} \right].
\tag{9.99}
$$

This can be used to rewrite (7.20) as

$$
\left\{ -\frac{\hbar^2}{2m}\left[\frac{1}{r}\frac{\partial^2}{\partial r^2}r + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2}{\partial\phi^2} \right] + V(r) \right\} \psi(\boldsymbol{r})
$$

$$
= \left\{ -\frac{\hbar^2}{2m}\frac{1}{r}\frac{\partial^2}{\partial r^2}r + \frac{\hat{\boldsymbol{L}}^2}{2mr^2} + V(r) \right\} \psi(\boldsymbol{r}) = E\psi(\boldsymbol{r}).
\tag{9.100}
$$

Comparison of this to the radial wave equation (7.48) shows that

$$
\hat{\boldsymbol{L}}^2 Y_{lm}(\theta,\phi) = \hbar^2 l(l+1) Y_{lm}(\theta,\phi)
\tag{9.101}
$$

We have now demonstrated that the spherical harmonics are eigenstates of the operators $\hat{\boldsymbol{L}}^2$ and $\hat{L}_z$.

We have, of course, performed all of the derivations here in coordinate space. We can represent the eigenstates of the of the angular momentum operator in Dirac notation by defining

$$
Y_{lm}(\theta,\phi) = \langle \boldsymbol{r} | \, lm \rangle,
\tag{9.102}
$$

where

$$
\hat{\boldsymbol{L}}^2 |lm\rangle = \hbar^2 l(l+1) |lm\rangle
\tag{9.103}
$$

and

$$
\hat{L}_z |lm\rangle = \hbar m |lm\rangle.
\tag{9.104}
$$

9.5 Matrix Representations of the Angular Momentum Operators

We now want to generalize the concept of angular momentum operators by introducing a new set of three operators $\hat{J}_i$ that obey the same commutations relations as the orbital angular momentum operators. That is,

$$\left[\hat{J}_i, \hat{J}_j\right] = i\hbar\epsilon_{ijk}\hat{J}_k \tag{9.105}$$

and the operators are assumed to be hermitian. These commutation relations imply that

$$\left[\hat{\boldsymbol{J}}^2, \hat{J}_i\right] = 0 \,. \tag{9.106}$$

We can therefore characterize the eigenstates of the operators as $|jm\rangle$ where

$$\hat{J}_z \left|jm\right\rangle = \hbar m \left|jm\right\rangle \tag{9.107}$$

and

$$\hat{\boldsymbol{J}}^2 \left|jm\right\rangle = \hbar^2 \lambda_j \left|jm\right\rangle \,, \tag{9.108}$$

where λ_j is some function of j.

It is useful to define two new operators

$$\hat{J}_\pm = \hat{J}_x \pm i\hat{J}_y \,. \tag{9.109}$$

These satisfy the commutation relations

$$\begin{align}
\left[\hat{\boldsymbol{J}}^2, \hat{J}_\pm\right] &= 0 \tag{9.110} \\
\left[\hat{J}_z, \hat{J}_+\right] &= \hbar\hat{J}_+ \tag{9.111} \\
\left[\hat{J}_z, \hat{J}_-\right] &= -\hbar\hat{J}_- \tag{9.112} \\
\left[\hat{J}_+, \hat{J}_-\right] &= 2\hbar\hat{J}_z \,. \tag{9.113}
\end{align}$$

We can now use these commutators to determine the properties of the operators and their eigenstates. First, (9.106) implies that

$$\hat{\boldsymbol{J}}^2 \hat{J}_i = \hat{J}_i \hat{\boldsymbol{J}}^2 \,. \tag{9.114}$$

Then,

$$\hat{\boldsymbol{J}}^2 \left(\hat{J}_i \left|jm\right\rangle\right) = \hat{J}_i \hat{\boldsymbol{J}}^2 \left|jm\right\rangle = \hbar^2 \hat{J}_i \lambda_j \left|jm\right\rangle = \hbar^2 \lambda_j \left(\hat{J}_i \left|jm\right\rangle\right) \,. \tag{9.115}$$

This means that all of the states contributing to $\hat{J}_i \left|jm\right\rangle$ must have the same value of j, but may have differing values of m.

Next, (9.111) means that

$$\hat{J}_z \hat{J}_+ - \hat{J}_+ \hat{J}_z = \hbar \hat{J}_+ \,. \tag{9.116}$$

Taking the matrix element of this expression gives

$$\langle jm' | \, \hat{J}_z \hat{J}_+ \, | jm \rangle - \langle jm' | \, \hat{J}_+ \hat{J}_z \, | jm \rangle = \hbar \, \langle jm' | \, \hat{J}_+ \, | jm \rangle \,. \tag{9.117}$$

Since $\hat{J}_z$ is hermitian it can act on eigenstates to the left or right to give

$$\hbar m' \, \langle jm' | \, \hat{J}_+ \, | jm \rangle - \hbar m \, \langle jm' | \, \hat{J}_+ \, | jm \rangle = \hbar \, \langle jm' | \, \hat{J}_+ \, | jm \rangle \,. \tag{9.118}$$

Simplifying this yields

$$(m' - m - 1) \, \langle jm' | \, \hat{J}_+ \, | jm \rangle = 0 \,. \tag{9.119}$$

In order for the matrix element of $\hat{J}_+$ to be nonzero, we must have $m' = m + 1$. So $\hat{J}_+$ is a raising operator for m. This implies that

$$\hat{J}_+ \, | jm \rangle = \hbar \kappa_m \, | j, m + 1 \rangle \,, \tag{9.120}$$

where κ_m is some function of m which has yet to be determined. A similar argument using (9.112) gives

$$\hat{J}_- \, | jm \rangle = \hbar \kappa'_m \, | j, m - 1 \rangle \,. \tag{9.121}$$

So $\hat{J}_-$ is a lowering operator for m.

From (9.109) it can be seen that

$$J_+^\dagger = J_- \,. \tag{9.122}$$

Taking a matrix element of this expression yields

$$\langle jm' | \, J_+^\dagger \, | jm \rangle = \langle jm' | \, J_- \, | jm \rangle \,. \tag{9.123}$$

We can now use (9.120) and (9.121) to rewrite this as

$$\hbar \kappa_{m'}^* \, \langle j, m' + 1 | \, jm \rangle = \hbar \kappa'_m \, \langle jm' | \, j, m - 1 \rangle \,, \tag{9.124}$$

which implies that

$$\hbar \kappa_{m'}^* \delta_{m'+1,m} = \hbar \kappa'_m \delta_{m',m-1} \,, \tag{9.125}$$

or

$$\kappa_{m'}^* = \kappa'_{m'+1} \,. \tag{9.126}$$

From (9.113),

$$\hat{J}_+ \hat{J}_- - \hat{J}_- \hat{J}_+ = 2\hbar \hat{J}_z \,. \tag{9.127}$$

157

Using (9.122), this can be rewritten as

$$\hat{J}_-^\dagger \hat{J}_- - \hat{J}_+^\dagger \hat{J}_+ = 2\hbar \hat{J}_z \,. \tag{9.128}$$

Taking a matrix element of this gives

$$\langle jm' | \hat{J}_-^\dagger \hat{J}_- | jm \rangle - \langle jm' | \hat{J}_+^\dagger \hat{J}_+ | jm \rangle = 2\hbar \langle jm' | \hat{J}_z | jm \rangle \,. \tag{9.129}$$

Using (9.107), (9.120), (9.121) and the orhogonality of the states, this reduces to

$$\hbar^2 \kappa_m'^* \kappa_m' \delta_{m'm} - \hbar^2 \kappa_m^* \kappa_m \delta_{m'm} = 2\hbar^2 m \delta_{m'm} \,. \tag{9.130}$$

Finally, using (9.126) this can be simplified to the difference equaiton

$$|\kappa_{m-1}|^2 - |\kappa_m|^2 = 2m \,. \tag{9.131}$$

The general solution to this difference equation is

$$|\kappa_m|^2 = C - m(m+1) \,, \tag{9.132}$$

where C is some constant. Note that the left-hand side of this equation must be nonnegative while for sufficiently large or small m the right-hand side can be become negative. In order for this expression to be valid it is necessary that the right-hand side must vanish for some maximum value of m, which we will label as $\overline{m}$ and some minimum value of m which we will label as $\underline{m}$. These can be found from the two solutions of

$$C - m(m+1) = 0 \,. \tag{9.133}$$

Therefore,

$$\overline{m} = -\frac{1}{2} + \frac{1}{2}\sqrt{1 + 4C} \tag{9.134}$$

and

$$\underline{m} = -\frac{1}{2} - \frac{1}{2}\sqrt{1 + 4C} \,. \tag{9.135}$$

These two equations can be solved to give

$$\underline{m} = -\overline{m} - 1 \tag{9.136}$$

and

$$C = \overline{m}(\overline{m} + 1) \,. \tag{9.137}$$

So,

$$|\kappa_m|^2 = \overline{m}(\overline{m} + 1) - m(m+1) \,, \tag{9.138}$$

Now consider

$$\hat{\boldsymbol{J}}^2 = \frac{1}{2}\left(\hat{J}_+ \hat{J}_- + \hat{J}_- \hat{J}_+\right) + \hat{J}_z^2 \,. \tag{9.139}$$

158

Using (9.122), this can be rewritten as

$$\boldsymbol{\hat{J}}^2 = \frac{1}{2}\left(\hat{J}_-^\dagger \hat{J}_- + \hat{J}_+^\dagger \hat{J}_+\right) + \hat{J}_z^2 \, . \tag{9.140}$$

Taking the diagonal matrix element of this yields

$$\langle jm | \boldsymbol{\hat{J}}^2 | jm\rangle = \frac{1}{2}\left(\langle jm | \hat{J}_-^\dagger \hat{J}_- | jm\rangle + \langle jm | \hat{J}_+^\dagger \hat{J}_+ | jm\rangle\right) + \langle jm | \hat{J}_z^2 | jm\rangle \, . \tag{9.141}$$

Using (9.107), (9.108), (9.120), (9.126) (9.121) and the orthogonality of the states gives

$$\hbar^2 \lambda_j = \frac{\hbar^2}{2}\left(|\kappa_{m-1}|^2 + |\kappa_m|^2\right) + m^2 \hbar^2 \, . \tag{9.142}$$

We can now use (9.138) to obtain

$$\lambda_j = \overline{m}(\overline{m}+1) \, . \tag{9.143}$$

This implies that $\overline{m}$ can only depend upon j, so we will identify $\overline{m} = j$ based on our experience with orbital angular momentum. Then we have

$$\boldsymbol{\hat{J}}^2 | jm\rangle = \hbar^2 j(j+1) | jm\rangle \tag{9.144}$$

and the requirement that $-j \le m \le j$.

Equation (9.138) is now

$$|\kappa_m|^2 = j(j+1) - m(m+1) \, , \tag{9.145}$$

which implies that

$$\kappa_m = e^{i\phi}\sqrt{j(j+1) - m(m+1)} \, , \tag{9.146}$$

where ϕ is some arbitrary phase angle which is chosen by convention to be zero. So

$$\kappa_m = \sqrt{j(j+1) - m(m+1)} \, . \tag{9.147}$$

Using (9.126),

$$\kappa_m' = \kappa_{m-1}^* = \sqrt{j(j+1) - m(m-1)} \, . \tag{9.148}$$

We can now rewrite (9.120) and (9.121) as

$$\hat{J}_+ | jm\rangle = \hbar\sqrt{j(j+1) - m(m+1)} \, | j, m+1\rangle \, , \tag{9.149}$$

and

$$\hat{J}_- | jm\rangle = \hbar\sqrt{j(j+1) - m(m-1)} \, | j, m-1\rangle \, . \tag{9.150}$$

From (9.149) we can see that the coefficient vanishes when $m = j$ so this is the largest value of m for which a state exists. Similarly (9.150) shows that the smallest value of is $m = -j$. Since the various states with fixed j can be reached by

repeated application of the raising or lowering operators which change m by ± 1, the difference between the largest and smallest values of m must be an integer. Therefore, $m_{max} - m_{min} = j - (-j) = 2j$ is an integer. Then j can take on the following values: $j = 0, \frac{1}{2}, 1, \frac{3}{2}, 2, \ldots$.

The matrix elements of (9.107), (9.144), (9.149) and (9.150) are then

$$\langle jm' | \hat{J}_z | jm \rangle = \hbar m \delta_{m'm}, \tag{9.151}$$

$$\langle jm' | \hat{\boldsymbol{J}}^2 | jm \rangle = \hbar^2 j(j+1)\delta_{m'm}, \tag{9.152}$$

$$\langle jm' | \hat{J}_+ | jm \rangle = \hbar \sqrt{j(j+1) - m(m+1)}\, \delta_{m',m+1}, \tag{9.153}$$

and

$$\langle jm' | \hat{J}_- | jm \rangle = \hbar \sqrt{j(j+1) - m(m-1)}\, \delta_{m',m-1}. \tag{9.154}$$

Using (9.109), we can solve for

$$\hat{J}_x = \frac{1}{2}\left(\hat{J}_- + \hat{J}_+\right) \tag{9.155}$$

and

$$\hat{J}_y = \frac{i}{2}\left(\hat{J}_- - \hat{J}_+\right). \tag{9.156}$$

These then have the matrix elements

$$\langle jm' | \hat{J}_x | jm \rangle = \frac{\hbar}{2}\left(\sqrt{j(j+1) - m(m-1)}\, \delta_{m',m-1} + \sqrt{j(j+1) - m(m+1)}\, \delta_{m',m+1}\right) \tag{9.157}$$

and

$$\langle jm' | \hat{J}_y | jm \rangle = \frac{i\hbar}{2}\left(\sqrt{j(j+1) - m(m-1)}\, \delta_{m',m-1} - \sqrt{j(j+1) - m(m+1)}\, \delta_{m',m+1}\right). \tag{9.158}$$

Note that everything which we have obtained in this section ultimately depends only on the commutation relations (9.105) and is true for any operators that satisfy these commutation relations, not just for the orbital angular momentum operator that we have studied previously. This will allow us to generalize the concept of angular momentum to include the concept of spin angular momentum.

We can now use (9.157), (9.158) and (9.151) to construct matrix representations of the angular momentum operators. There is a matrix representation for each value of j with matrices of dimension $2j+1$. For example, for $j = 1$ the matrix representations have a dimension of 3. These matrices are:

$$\underline{\underline{J_x}} = \frac{\hbar}{\sqrt{2}}\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \tag{9.159}$$

$$\underline{\underline{J_y}} = \frac{\hbar}{\sqrt{2}} \begin{pmatrix} 0 & -i & 0 \\ i & 0 & -i \\ 0 & i & 0 \end{pmatrix} \qquad (9.160)$$

and

$$\underline{\underline{J_z}} = \hbar \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}. \qquad (9.161)$$

where the matrix indices 1, 2 and 3 correspond to the values $m = 1, 0, -1$. Note that these matrices are hermitian as required for observables. Also, for any two matrices $\boldsymbol{A}$ and $\boldsymbol{B}$,

$$\mathrm{Tr}(\underline{\underline{A}}\,\underline{\underline{B}}) = \mathrm{Tr}(\underline{\underline{B}}\,\underline{\underline{A}}). \qquad (9.162)$$

As a result,

$$\mathrm{Tr}([\underline{\underline{A}}, \underline{\underline{B}}]) = 0. \qquad (9.163)$$

Taking the trace of

$$[\underline{\underline{J_i}}, \underline{\underline{J_j}}] = i\hbar\epsilon_{ijk}\underline{\underline{J_k}} \qquad (9.164)$$

then requires that

$$\mathrm{Tr}\underline{\underline{J_k}} = 0. \qquad (9.165)$$

Note that the matrices that we have derived for $j = 1$ are traceless as this requires.

9.6   Spin

In solving the time-independent Schrödinger equation for a central potential we obtained solutions that are eigenstates of orbital angular momentum. These solutions are for integer values of l and m. However, in the previous section we showed that operators satisfying the commutation relations for angular momenta can also have solutions corresponding to half-odd-integer angular momenta j. These solutions have no classical analog and can only be represented in matrix form.

Our previous solution of the Schrödinger equation in a coulomb potential describes a simple model of a hydrogen-like atom where the electron and proton interacted through a simple Coulomb potential. For this model we obtained a spectrum that depends on the quantum numbers n and l. Careful examination of the hydrogen spectrum reveals that each of these states is in fact split into two states. This can be understood by attributing a new property to the electron called *spin angular momentum*. Although this suggests that this quantity is the result of the electron spinning about some axis, the electron is an elementary point-like particle and the concept of the electron spinning about some axis has no physical meaning. Instead this is an intrinsic property of the electron for which there is no differential operator such as for the case of the orbital angular momentum. The representation of spin

is therefore accomplished entirely in the form of matrix representations of a set of operators that satisfy the commutation relations

$$\left[\hat{S}_i, \hat{S}_j\right] = i\hbar\epsilon_{ijk}\hat{S}_k. \tag{9.166}$$

9.6.1 Spin 1/2

The electron has spin $\frac{1}{2}$. We can then construct the matrix representations of this operator by using (9.157), (9.158) and (9.151) with $j = \frac{1}{2}$ to construct a two-dimensional matrix representation of the spin-$\frac{1}{2}$ angular momentum operators. These are

$$\underline{\underline{S_x}} = \frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \tag{9.167}$$

$$\underline{\underline{S_y}} = \frac{\hbar}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \tag{9.168}$$

and

$$\underline{\underline{S_z}} = \frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{9.169}$$

It is conventional to write these matrices as

$$\underline{\underline{S_i}} = \frac{\hbar}{2}\underline{\underline{\sigma_i}}, \tag{9.170}$$

where the $\underline{\underline{\sigma_i}}$ are the *Pauli spin matrices*

$$\underline{\underline{\sigma_x}} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \tag{9.171}$$

$$\underline{\underline{\sigma_y}} = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \tag{9.172}$$

and

$$\underline{\underline{\sigma_z}} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{9.173}$$

The Pauli spin matrices satisfy the commutation relations

$$\left[\underline{\underline{\sigma_i}}, \underline{\underline{\sigma_j}}\right] = 2i\epsilon_{ijk}\underline{\underline{\sigma_k}} \tag{9.174}$$

and

$$\underline{\underline{\sigma_x}}^2 = \underline{\underline{\sigma_y}}^2 = \underline{\underline{\sigma_z}}^2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{9.175}$$

The Pauli matrices satisfy the anti-commutation relation

$$\{\underline{\underline{\sigma_i}}, \underline{\underline{\sigma_j}}\} \equiv \underline{\underline{\sigma_i}}\,\underline{\underline{\sigma_j}} + \underline{\underline{\sigma_j}}\,\underline{\underline{\sigma_i}} = 2\delta_{ij}\underline{\underline{1}}. \tag{9.176}$$

Adding (9.174) and (9.176) gives the identity

$$\underline{\underline{\sigma_i}}\,\underline{\underline{\sigma_j}} = \underline{\underline{1}}\delta_{ij} + i\epsilon_{ijk}\underline{\underline{\sigma_k}}\,. \tag{9.177}$$

The eigenstates of σ_z are the *Pauli spinors*

$$\underline{\chi}_{\frac{1}{2}} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \tag{9.178}$$

and

$$\underline{\chi}_{-\frac{1}{2}} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}\,. \tag{9.179}$$

These spinors form a complete orthonormal basis in two dimensions. The spinors satisfy the eigenvalue equation

$$\underline{\underline{S_z}}\underline{\chi}_{\pm\frac{1}{2}} = \pm\frac{1}{2}\hbar\underline{\chi}_{\pm\frac{1}{2}}\,. \tag{9.180}$$

The rotation operator in the two-dimensional spinor space can be written using (9.60) as

$$\underline{\underline{R}}(\boldsymbol{n},\theta) = e^{-\frac{i}{\hbar}\theta\boldsymbol{n}\cdot\underline{\underline{\boldsymbol{S}}}} = e^{-\frac{i}{2}\theta\boldsymbol{n}\cdot\underline{\underline{\boldsymbol{\sigma}}}}\,. \tag{9.181}$$

It is easy to show from (9.177) that

$$\left(\boldsymbol{n}\cdot\underline{\underline{\boldsymbol{\sigma}}}\right)^2 = \underline{\underline{1}}\,. \tag{9.182}$$

Expanding (9.181) in a power series and using this identity the rotation operator can be shown to be

$$\underline{\underline{R}}(\boldsymbol{n},\theta) = \underline{\underline{1}}\cos\frac{\theta}{2} + i\boldsymbol{n}\cdot\underline{\underline{\boldsymbol{\sigma}}}\sin\frac{\theta}{2} \tag{9.183}$$

Note that

$$\underline{\underline{R}}(\boldsymbol{n},2\pi) = -\underline{\underline{1}}\,. \tag{9.184}$$

That is, the rotation of a spinor through 2π about any axis yields the negative of the spinor rather than the spinor itself. The identity operator is obtained by rotation through integral multiples of 4π. This is characteristic of the group SU(2).

It is often convenient to use an abstract Dirac notation to describe the spin angular momentum in a fashion similar to that which we have used for describing orbital angular momentum. In this case we will represent the spin-1/2 eigenstates as

$$\underline{\chi}_{m_s} \rightarrow \left|\tfrac{1}{2}, m_s\right\rangle\,, \tag{9.185}$$

where

$$\hat{S}_z \left|\tfrac{1}{2}, m_s\right\rangle = \hbar m_s \left|\tfrac{1}{2}, m_s\right\rangle \tag{9.186}$$

and

$$\hat{\boldsymbol{S}}^2 \left|\tfrac{1}{2}, m_s\right\rangle = \hbar^2\tfrac{1}{2}(\tfrac{1}{2}+1)\left|\tfrac{1}{2}, m_s\right\rangle = \tfrac{3}{4}\hbar^2\left|\tfrac{1}{2}, m_s\right\rangle\,. \tag{9.187}$$

9.6.2 The Intrinsic Magnetic Moment of Spin-1/2 Particles

For a charge $-e$ moving in a circle with angular momentum L the classical magnetic moment is

$$\boldsymbol{\mu}_c = -\frac{e\boldsymbol{L}}{2mc} \, . \tag{9.188}$$

Since the spin of a particle is not associated with such a simple flow of charge, it is not surprising that the intrinsic magnetic moment of the electron does not satisfy this relation. The magnetic moment of an electron is given by

$$\boldsymbol{\mu} = -\frac{eg\boldsymbol{S}}{2m_e} = -\frac{eg\hbar\boldsymbol{\sigma}}{4m_e c} \, , \tag{9.189}$$

where $g \cong 2$ is called the *Landé g factor*. Note that for simplicity of notation we have dropped the explicit designation of matrices and column vectors denoted by underlines. We will now assume that any expression involving the Pauli matrices is necessarily a matrix equation.

For an electron at rest in an external magnetic field $\boldsymbol{B}$ the hamiltonian is given by

$$\hat{H} = -\boldsymbol{\mu} \cdot \boldsymbol{B} = \frac{eg\hbar}{4m_e c}\boldsymbol{\sigma} \cdot \boldsymbol{B} \, . \tag{9.190}$$

The time-dependent Schrödinger equation is then

$$i\hbar\frac{\partial}{\partial t}\Psi(t) = \frac{eg\hbar}{4m_e c}\boldsymbol{\sigma} \cdot \boldsymbol{B}\Psi(t) \, . \tag{9.191}$$

where the time dependent wave function $\Psi(t)$ is two-dimensional Pauli spinor that can be written as

$$\Psi(t) = \Psi_+(t)\chi_{\frac{1}{2}} + \Psi_-(t)\chi_{-\frac{1}{2}} = \begin{pmatrix} \Psi_+(t) \\ \Psi_-(t) \end{pmatrix} \, . \tag{9.192}$$

If the magnetic field is aligned along the z-axis, then

$$\hat{H} = \frac{eg\hbar B}{4m_e c}\sigma_z \tag{9.193}$$

and

$$\hat{H}\chi_{\pm\frac{1}{2}} = \pm\frac{eg\hbar B}{4m_e c}\chi_{\pm\frac{1}{2}} = \pm\hbar\omega_0\chi_{\pm\frac{1}{2}} \tag{9.194}$$

where

$$\omega_0 = \frac{egB}{4m_e c} \, . \tag{9.195}$$

The operator form of the solution to the time-dependent Schrödinger equation is

$$\Psi(t) = e^{-\frac{i}{\hbar}\hat{H}t}\Psi(0) \, . \tag{9.196}$$

For $\boldsymbol{B}$ along the z-axis this becomes

$$
\begin{aligned}
\Psi(t) &= e^{-\frac{i}{\hbar}\hat{H}t}\left(\Psi_+(0)\chi_{\frac{1}{2}} + \Psi_-(0)\chi_{-\frac{1}{2}}\right) = e^{-i\omega_0 t}\Psi_+(0)\chi_{\frac{1}{2}} + e^{i\omega_0 t}\Psi_-(0)\chi_{-\frac{1}{2}} \\
&= \left(\begin{array}{c} e^{-i\omega_0 t}\Psi_+(0) \\ e^{i\omega_0 t}\Psi_-(0) \end{array}\right).
\end{aligned}
\tag{9.197}
$$

As an example, consider the case where the spin is initially polarized along the x-axis. That is

$$
\sigma_x \Psi(0) = \Psi(0).
\tag{9.198}
$$

This state can be found by solving the eigenvalue problem

$$
\left(\begin{array}{cc} 0 & 1 \\ 1 & 0 \end{array}\right)\left(\begin{array}{c} \Psi_+(0) \\ \Psi_-(0) \end{array}\right) = \lambda\left(\begin{array}{c} \Psi_+(0) \\ \Psi_-(0) \end{array}\right).
\tag{9.199}
$$

For this equation to have a solution, we must have

$$
\det\left(\begin{array}{cc} -\lambda & 1 \\ 1 & -\lambda \end{array}\right) = \lambda^2 - 1 = 0.
\tag{9.200}
$$

This implies that $\lambda = \pm 1$. The solution we are looking for corresponds to the solution $\lambda = 1$ and has the normalized eigenvector

$$
\left(\begin{array}{c} \Psi_+(0) \\ \Psi_-(0) \end{array}\right) = \frac{1}{\sqrt{2}}\left(\begin{array}{c} 1 \\ 1 \end{array}\right).
\tag{9.201}
$$

The time-dependent wave function for this case is then

$$
\Psi(t) = \frac{1}{\sqrt{2}}\left(\begin{array}{c} e^{-i\omega_0 t} \\ e^{i\omega_0 t} \end{array}\right).
\tag{9.202}
$$

We can now use this solution to calculate the average values of the components of the spin operator

$$
\langle S_i(t)\rangle = \Psi^\dagger(t)\frac{\hbar}{2}\sigma_i\Psi(t).
\tag{9.203}
$$

So,

$$
\langle S_x\rangle = \frac{\hbar}{4}\left(\begin{array}{cc} e^{i\omega_0 t} & e^{-i\omega_0 t} \end{array}\right)\left(\begin{array}{cc} 0 & 1 \\ 1 & 0 \end{array}\right)\left(\begin{array}{c} e^{-i\omega_0 t} \\ e^{i\omega_0 t} \end{array}\right) = \frac{\hbar}{4}\left(e^{i2\omega_0 t} + e^{-i2\omega_0 t}\right) = \frac{\hbar}{2}\cos 2\omega_0 t,
\tag{9.204}
$$

$$
\langle S_y\rangle = \frac{\hbar}{4}\left(\begin{array}{cc} e^{i\omega_0 t} & e^{-i\omega_0 t} \end{array}\right)\left(\begin{array}{cc} 0 & -i \\ i & 0 \end{array}\right)\left(\begin{array}{c} e^{-i\omega_0 t} \\ e^{i\omega_0 t} \end{array}\right) = \frac{\hbar}{4}\left(-ie^{i2\omega_0 t} + ie^{-i2\omega_0 t}\right) = \frac{\hbar}{2}\sin 2\omega_0 t
\tag{9.205}
$$

and

$$
\langle S_z\rangle = \frac{\hbar}{4}\left(\begin{array}{cc} e^{i\omega_0 t} & e^{-i\omega_0 t} \end{array}\right)\left(\begin{array}{cc} 1 & 0 \\ 0 & -1 \end{array}\right)\left(\begin{array}{c} e^{-i\omega_0 t} \\ e^{i\omega_0 t} \end{array}\right) = 0.
\tag{9.206}
$$

165

This implies that the average spin vector is precessing about the z-axis in the xy-plane with a frequency of $2\omega_0$.

With this in mind it is interesting to rewrite (9.196) as

$$\Psi(t) = e^{-\frac{i}{\hbar}\varphi(t)\hat{S}_z}\Psi(0),\tag{9.207}$$

where $\varphi(t) = 2\omega_0 t$ is the azimuthal angle of the average spin vector at time t. We can recognize this operator as $\hat{R}(\boldsymbol{z},\varphi(t))$. So, this represents a rotation of the spinor about the z-axis at a constant angular velocity.

9.6.3 Paramagnetic Resonance

Now suppose that we add a small time dependent component to the magnetic field in the x direction. That is $\boldsymbol{B}(t) = B_0\hat{z} + B_1\cos\omega t\hat{x}$. The hamiltonian in this case is then

$$\begin{aligned}\hat{H} &= \frac{eg\hbar}{4m_e}[B_0\sigma_z + B_1\cos\omega t\sigma_x] = \frac{eg\hbar}{4m_e}\begin{pmatrix} B_0 & B_1\cos\omega t \\ B_1\cos\omega t & -B_0 \end{pmatrix} \\ &= \hbar\begin{pmatrix} \omega_0 & \omega_1\cos\omega t \\ \omega_1\cos\omega t & -\omega_0 \end{pmatrix},\end{aligned}\tag{9.208}$$

where $\omega_0 = \frac{egB_0}{4m_e c}$ and $\omega_1 = \frac{egB_1}{4m_e c}$.

The time-dependent Schrödinger equation can then be written as

$$\hbar\begin{pmatrix} \omega_0 & \omega_1\cos\omega t \\ \omega_1\cos\omega t & -\omega_0 \end{pmatrix}\begin{pmatrix} \Psi_+(t) \\ \Psi_-(t) \end{pmatrix} = i\hbar\frac{d}{dt}\begin{pmatrix} \Psi_+(t) \\ \Psi_-(t) \end{pmatrix}.\tag{9.209}$$

Since $B_0 \gg B_1$, $\omega_0 \gg \omega_1$. As a result the greatest part of the time dependence is determined by ω_0. For this reason it is convenient to write the time-dependent wave function as

$$\begin{pmatrix} \Psi_+(t) \\ \Psi_-(t) \end{pmatrix} = \begin{pmatrix} \phi_+(t)e^{-i\omega_0 t} \\ \phi_-(t)e^{i\omega_0 t} \end{pmatrix}.\tag{9.210}$$

Substituting this into the time-dependent Schrödinger equation gives

$$\begin{pmatrix} \omega_0\phi_+(t)e^{-i\omega_0 t} + \omega_1\cos\omega t\phi_-(t)e^{i\omega_0 t} \\ \omega_1\cos\omega t\phi_+(t)e^{-i\omega_0 t} - \omega_0\phi_-(t)e^{i\omega_0 t} \end{pmatrix} = \begin{pmatrix} i\left(\frac{d}{dt}\phi_+(t)\right)e^{-i\omega_0 t} + \omega_0\phi_+(t)e^{-i\omega_0 t} \\ i\left(\frac{d}{dt}\phi_-(t)\right)e^{i\omega_0 t} - \omega_0\phi_-(t)e^{i\omega_0 t} \end{pmatrix}.\tag{9.211}$$

Simplifying this expression yields

$$\begin{pmatrix} \omega_1\cos\omega t\phi_-(t)e^{i2\omega_0 t} \\ \omega_1\cos\omega t\phi_+(t)e^{-i2\omega_0 t} \end{pmatrix} = \begin{pmatrix} i\frac{d}{dt}\phi_+(t) \\ i\frac{d}{dt}\phi_-(t) \end{pmatrix}.\tag{9.212}$$

Now consider

$$\cos\omega t\, e^{i2\omega_0 t} = \frac{1}{2}\left(e^{i\omega t} + e^{-i\omega t}\right)e^{i2\omega_0 t} = \frac{1}{2}\left(e^{i(2\omega_0+\omega)t} + e^{i(2\omega_0-\omega)t}\right).\tag{9.213}$$

We will be primarily interested in considering the solution to the time-dependent Schrödinger equation for driving frequencies ω where $\omega \cong 2\omega_0$. In this case $e^{i(2\omega_0-\omega)t}$ will vary slowly with time while $e^{i(2\omega_0+\omega)t}$ will oscillate rapidly with time. If the solution is to be averaged over some interval of time where the second term is varying rapidly, this contribution will tend to average to zero while the other term will not be noticeably changed by the averaging. As a result we can make the "rotating field approximation"

$$\cos \omega t \, e^{i2\omega_0 t} \cong \frac{1}{2} e^{i(2\omega_0-\omega)t} . \tag{9.214}$$

Similarly

$$\cos \omega t \, e^{-i2\omega_0 t} \cong \frac{1}{2} e^{-i(2\omega_0-\omega)t} . \tag{9.215}$$

With these approximations, the time-dependent Schrödinger equation reduces to the pair of coupled first-order differential equations

$$i\frac{d}{dt}\phi_+(t) = \frac{1}{2}\omega_1 e^{i(2\omega_0-\omega)t}\phi_-(t) \tag{9.216}$$

and

$$i\frac{d}{dt}\phi_-(t) = \frac{1}{2}\omega_1 e^{-i(2\omega_0-\omega)t}\phi_+(t) . \tag{9.217}$$

These equations can be decoupled by taking the time derivative of (9.216) to give

$$\frac{d^2}{dt^2}\phi_+(t) = (2\omega_0-\omega)\frac{1}{2}\omega_1 e^{i(2\omega_0-\omega)t}\phi_-(t) - \frac{1}{2}\omega_1 e^{i(2\omega_0-\omega)t}i\frac{d}{dt}\phi_-(t) . \tag{9.218}$$

All terms containing $\phi_-(t)$ can be reexpressed in terms of $\phi_+(t)$ by using (9.216) and (9.217). This yields

$$\begin{aligned}\frac{d^2}{dt^2}\phi_+(t) &= (2\omega_0-\omega)i\frac{d}{dt}\phi_+(t) - \frac{1}{2}\omega_1 e^{i(2\omega_0-\omega)t}\frac{1}{2}\omega_1 e^{-i(2\omega_0-\omega)t}\phi_+(t) \\ &= i(2\omega_0-\omega)\frac{d}{dt}\phi_+(t) - \frac{1}{4}\omega_1^2\phi_+(t) . \end{aligned} \tag{9.219}$$

If we assume that $\phi_+(t) = e^{i\Omega t}$ then the differential equation requires that

$$-\Omega^2 = -(2\omega_0-\omega)\Omega - \frac{1}{4}\omega_1^2 . \tag{9.220}$$

This can be solved to give

$$\Omega_\pm = \omega_0 - \frac{\omega}{2} \pm \sqrt{\left(\omega_0 - \frac{\omega}{2}\right)^2 + \frac{1}{4}\omega_1^2} . \tag{9.221}$$

$\phi_+(t)$ can then be written as a linear combination of these two solutions

$$\phi_+(t) = A_+ e^{i\Omega_+ t} + A_- e^{i\Omega_- t} . \tag{9.222}$$

167

From (9.216) we can write

$$
\begin{aligned}
\phi_-(t) &= \frac{2}{\omega_1} e^{-i(2\omega_0-\omega)t} i \frac{d}{dt} \phi_+(t) = \frac{2}{\omega_1} e^{-i(2\omega_0-\omega)t} \left(-A_+\Omega_+ e^{i\Omega_+ t} - A_-\Omega_- e^{i\Omega_- t} \right) \\
&= -\frac{2}{\omega_1} \left(A_+\Omega_+ e^{i(\Omega_+ - 2\omega_0+\omega)t} + A_-\Omega_- e^{i(\Omega_- - 2\omega_0+\omega)t} \right) .
\end{aligned}
\tag{9.223}
$$

Using

$$
\Omega_\pm - 2\omega_0 + \omega = -\Omega_\mp ,
\tag{9.224}
$$

Then

$$
\phi_-(t) = -\frac{2}{\omega_1} \left(A_+\Omega_+ e^{-i\Omega_- t} + A_-\Omega_- e^{-i\Omega_+ t} \right) .
\tag{9.225}
$$

The components of the time dependent wave function can now be written as

$$
\Psi_+(t) = \phi_+(t) e^{-i\omega_0 t} = A_+ e^{i(\Omega_+ -\omega_0)t} + A_- e^{i(\Omega_- -\omega_0)t}
\tag{9.226}
$$

and

$$
\Psi_-(t) = \phi_-(t) e^{i\omega_0 t} = -\frac{2}{\omega_1} \left(A_+\Omega_+ e^{-i(\Omega_- -\omega_0)t} + A_-\Omega_- e^{-i(\Omega_+ -\omega_0)t} \right) .
\tag{9.227}
$$

Now consider the special case where the spin is initially aligned along the z-axis. That is

$$
\Psi(0) = \begin{pmatrix} 1 \\ 0 \end{pmatrix} .
\tag{9.228}
$$

Then

$$
\Psi_+(0) = A_+ + A_- = 1
\tag{9.229}
$$

and

$$
\Psi_-(0) = -\frac{2}{\omega_1} \left(A_+\Omega_+ + A_-\Omega_- \right) = 0 .
\tag{9.230}
$$

These two equations can be solved to give

$$
A_+ = \frac{\Omega_-}{\Omega_- - \Omega_+}
\tag{9.231}
$$

and

$$
A_- = -\frac{\Omega_+}{\Omega_- - \Omega_+} .
\tag{9.232}
$$

Using these coefficients, the wave function components become

$$
\Psi_+(t) = \frac{\Omega_- e^{i(\Omega_+ -\omega_0)t} - \Omega_+ e^{i(\Omega_- -\omega_0)t}}{\Omega_- - \Omega_+}
\tag{9.233}
$$

and

$$
\Psi_-(t) = -\frac{2}{\omega_1} \frac{\Omega_+\Omega_-}{\Omega_- - \Omega_+} \left(e^{-i(\Omega_- -\omega_0)t} - e^{-i(\Omega_+ -\omega_0)t} \right) .
\tag{9.234}
$$

168

Since the wave function is normalized with for all times

$$1 = \Psi^\dagger(t)\Psi(t) = \Psi^\dagger(t)\left(\chi_{\frac{1}{2}}\chi_{\frac{1}{2}}^\dagger + \chi_{-\frac{1}{2}}\chi_{-\frac{1}{2}}^\dagger\right)\Psi(t) = \left|\chi_{\frac{1}{2}}^\dagger\Psi(t)\right|^2 + \left|\chi_{-\frac{1}{2}}^\dagger\Psi(t)\right|^2 \quad (9.235)$$

The first term at the right is the probability that the spin will be $1/2$ at time t, while the second term is the probability that the spin will be $-1/2$. Consider the second term. That is,

$$
\begin{aligned}
P_{-\frac{1}{2}}(t) &= \left|\chi_{-\frac{1}{2}}^\dagger\Psi(t)\right|^2 = |\Psi_-(t)|^2 = \frac{4}{\omega_1^2}\left(\frac{\Omega_+\Omega_-}{\Omega_- - \Omega_+}\right)^2\left|e^{-i\Omega_- t} - e^{-i\Omega_+ t}\right|^2 \\
&= \frac{8}{\omega_1^2}\left(\frac{\Omega_+\Omega_-}{\Omega_- - \Omega_+}\right)^2(1 - \cos[(\Omega_- - \Omega_+)t]) \\
&= \frac{1}{8}\frac{\omega_1^2}{\left(\omega_0 - \frac{\omega}{2}\right)^2 + \frac{\omega_1^2}{4}}\left[1 - \cos\left(2\sqrt{\left(\omega_0 - \frac{\omega}{2}\right)^2 + \frac{\omega_1^2}{4}}\, t\right)\right]. \quad (9.236)
\end{aligned}
$$

Consider the factor

$$f(\omega) = \frac{1}{4}\frac{\omega_1^2}{\left(\omega_0 - \frac{\omega}{2}\right)^2 + \frac{\omega_1^2}{4}} = \frac{\frac{\omega_1^2}{\omega_0^2}}{\left(2 - \frac{\omega}{\omega_0}\right)^2 + \frac{\omega_1^2}{\omega_0^2}}. \quad (9.237)$$

Since the remaining contribution to (9.236) is oscillatory and bounded between 0 and 1, $f(\omega)$ determines the amplitude of the probability function. The maximum of $f(\omega)$ occurs when the denominator is at its minimal value. This clearly occurs when $\frac{\omega}{\omega_0} = 2$ or $\omega = 2\omega_0$. At this point $f(2\omega_0) = 1$. The function $f(\omega)$ falls of rapidly for values of ω above or below this point. This can be seen in Fig. 9.2 which plots $f(\omega)$ as a function of ω/ω_0. This is a classic example of a resonating system. When the system is on resonance where $\omega = 2\omega_0$,

$$P_{-\frac{1}{2}} = \frac{1}{2}\left(1 - \cos\omega_1 t\right). \quad (9.238)$$
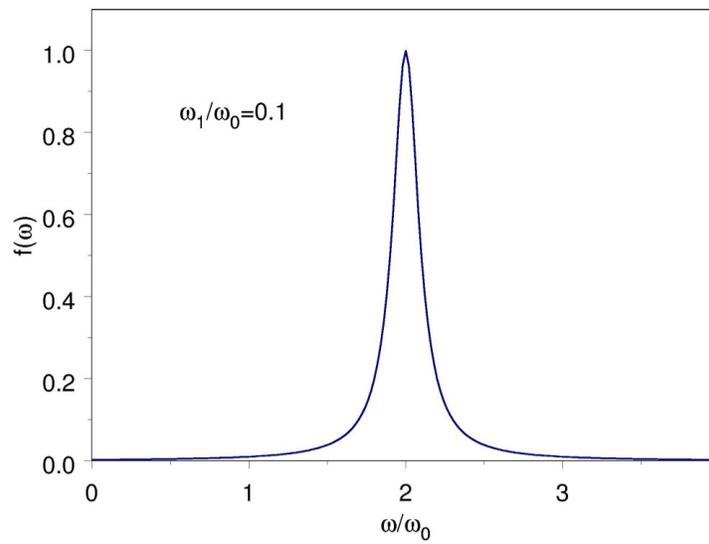
Figure 9.2: The function $f(\omega)$ as a function of ω/ω_0.

Chapter 10

Addition of Angular Momenta

10.1 Introduction

We can now think of the hydrogen atom as a spin-1/2 electron moving in the Coulomb potential of a proton. As we have already shown, the electron has an orbital angular momentum associated with the operator $\hat{\boldsymbol{L}}$, where

$$[\hat{L}_i, \hat{L}_j] = i\hbar\epsilon_{ijk}\hat{L}_k \tag{10.1}$$

and

$$[\hat{\boldsymbol{L}}^2, \hat{L}_i] = 0 \,. \tag{10.2}$$

The eigenstates of the orbital angular momentum are $|lm_l\rangle$ such that

$$\hat{L}_z |lm_l\rangle = \hbar m_l |lm_l\rangle \tag{10.3}$$

and

$$\hat{\boldsymbol{L}}^2 |lm_l\rangle = \hbar^2 l(l+1) |lm_l\rangle \tag{10.4}$$

with $-l \leq m_l \leq l$. These orbital angular momentum states are part of the infinite-dimensional basis of the Hilbert space spanning three-dimensional space.

In addition to this we now have the spin angular momentum of the electron associated with the operator $\hat{\boldsymbol{S}}$, where

$$[\hat{S}_i, \hat{S}_j] = i\hbar\epsilon_{ijk}\hat{S}_k \tag{10.5}$$

and

$$[\hat{\boldsymbol{S}}^2, \hat{S}_i] = 0 \,. \tag{10.6}$$

The eigenstates of the orbital angular momentum are $\left|\tfrac{1}{2}m_s\right\rangle$ such that

$$\hat{S}_z |lm_l\rangle = \hbar m_s \left|\tfrac{1}{2}m_s\right\rangle \tag{10.7}$$

and

$$\hat{\boldsymbol{S}}^2 \left|\tfrac{1}{2}m_s\right\rangle = \tfrac{3}{4}\hbar^2 \left|\tfrac{1}{2}m_s\right\rangle \tag{10.8}$$

with $m_s = \pm\frac{1}{2}$. These spin angular momentum states form the basis of a two-dimensional Hilbert space.

These two Hilbert spaces are independent of each other which means that

$$[\hat{L}_i, \hat{S}_j] = 0 \,. \tag{10.9}$$

The total Hilbert space describing the electron in a Coulomb potential is the sum of these two Hilbert spaces and the eigenstates that span the complete Hilbert space are direct productions of the eigenstaes of the subspaces. That is, the states

$$\left|nlm_l; \tfrac{1}{2}m_s\right\rangle = \left|nlm_l\right\rangle \left|\tfrac{1}{2}m_s\right\rangle \tag{10.10}$$

form the basis states for the complete Hilbert space of the electron.

Classically, we know that we can add two different angular momenta to give a total angular momentum. This suggest that it should be a possible to represent state of the electron in terms of eigenstates of a total angular momentum operator

$$\hat{\boldsymbol{J}} = \hat{\boldsymbol{L}} + \hat{\boldsymbol{S}} \,. \tag{10.11}$$

To see how this can be done, we will consider the general problem of the addition of two angular momenta.

10.2 Addition of Two Angular Momenta

Consider two angular momentum operators $\hat{\boldsymbol{J}}^{(1)}$ and $\hat{\boldsymbol{J}}^{(2)}$ such that

$$[\hat{J}_i^{(1)}, \hat{J}_j^{(1)}] = i\hbar\epsilon_{ijk}\hat{J}_k^{(1)} \,, \tag{10.12}$$

$$[\hat{J}_i^{(2)}, \hat{J}_j^{(2)}] = i\hbar\epsilon_{ijk}\hat{J}_k^{(2)} \tag{10.13}$$

and

$$[\hat{J}_i^{(1)}, \hat{J}_j^{(2)}] = 0 \,. \tag{10.14}$$

The eigenstates of the two angular momentum operators are defined such that

$$\hat{J}_z^{(1)} \left|j_1 m_1\right\rangle = \hbar m_1 \left|j_1 m_1\right\rangle \tag{10.15}$$
$$\hat{\boldsymbol{J}}^{(1)2} \left|j_1 m_1\right\rangle = \hbar^2 j_1(j_1+1) \left|j_1 m_1\right\rangle \tag{10.16}$$

and

$$\hat{J}_z^{(2)} \left|j_2 m_2\right\rangle = \hbar m_2 \left|j_2 m_2\right\rangle \tag{10.17}$$
$$\hat{\boldsymbol{J}}^{(2)2} \left|j_2 m_2\right\rangle = \hbar^2 j_2(j_2+1) \left|j_2 m_2\right\rangle \,. \tag{10.18}$$

Since the Hilbert spaces of the two operators are independent, the total basis states of the total Hilbert space are given by

$$\left|j_1 m_1; j_2 m_2\right\rangle = \left|j_1 m_1\right\rangle \left|j_2 m_2\right\rangle \,. \tag{10.19}$$

Now consider the total angular momentum operator

$$\hat{\boldsymbol{J}} = \hat{\boldsymbol{J}}^{(1)} + \hat{\boldsymbol{J}}^{(2)} . \tag{10.20}$$

The commutation relations for $\hat{\boldsymbol{J}}^{(1)}$ and $\hat{\boldsymbol{J}}^{(2)}$ can be used to show that

$$
\begin{align}
[\hat{J}_i, \hat{J}_j] &= i\hbar \epsilon_{ijk} \hat{J}_k \tag{10.21} \\
[\hat{\boldsymbol{J}}^2, \hat{J}_i] &= 0 \tag{10.22} \\
[\hat{\boldsymbol{J}}^2, \hat{\boldsymbol{J}}^{(1)2}] &= 0 \tag{10.23} \\
[\hat{\boldsymbol{J}}^2, \hat{\boldsymbol{J}}^{(2)2}] &= 0 \tag{10.24} \\
[\hat{\boldsymbol{J}}^2, \hat{J}_i^{(1)}] &= 2i\hbar \left(\hat{\boldsymbol{J}}^{(1)} \times \hat{\boldsymbol{J}}^{(2)} \right)_i \tag{10.25} \\
[\hat{\boldsymbol{J}}^2, \hat{J}_i^{(2)}] &= -2i\hbar \left(\hat{\boldsymbol{J}}^{(1)} \times \hat{\boldsymbol{J}}^{(2)} \right)_i \tag{10.26}
\end{align}
$$

Since the direct product states are labeled by the quantum numbers of four operators, we expect that the new total angular momentum basis will also be labeled by the quantum numbers of for operators. Equations (10.21), (10.22), (10.23) and (10.24) indicate that we can choose these to be $\hat{\boldsymbol{J}}^2$, $\hat{J}_z$, $\hat{\boldsymbol{J}}^{(1)2}$ and $\hat{\boldsymbol{J}}^{(2)2}$. Equations (10.21) and (10.22) imply that the eigenstates of the total angular momentum operator can be defined such that

$$
\begin{align}
\hat{J}_z |j_1 j_2 J M\rangle &= \hbar M |j_1 j_2 J M\rangle \tag{10.27} \\
\hat{\boldsymbol{J}}^2 |j_1 j_2 J M\rangle &= \hbar^2 J(J+1) |j_1 j_2 J M\rangle \tag{10.28} \\
\hat{\boldsymbol{J}}^{(1)2} |j_1 j_2 J M\rangle &= \hbar^2 j_1(j_1+1) |j_1 j_2 J M\rangle \tag{10.29} \\
\hat{\boldsymbol{J}}^{(2)2} |j_1 j_2 J M\rangle &= \hbar^2 j_2(j_2+1) |j_1 j_2 J M\rangle , \tag{10.30}
\end{align}
$$

where $-J \leq M \leq J$.

We have now defined two different basis sets that span the same space, the direct product set and the total angular momentum set. We now want to construct and expansion of the total angular momentum states in terms of the direct product states. As a first step, note that

$$
\begin{align}
\hat{J}_z |j_1 m_1; j_2 m_2\rangle &= \left(\hat{J}_z^{(1)} + \hat{J}_z^{(2)} \right) |j_1 m_1\rangle |j_2 m_2\rangle \\
&= \left(\hat{J}_z^{(1)} |j_1 m_1\rangle \right) |j_2 m_2\rangle + |j_1 m_1\rangle \hat{J}_z^{(2)} |j_2 m_2\rangle \\
&= \hbar m_1 |j_1 m_1\rangle |j_2 m_2\rangle + \hbar m_2 |j_1 m_1\rangle |j_2 m_2\rangle \\
&= \hbar(m_1 + m_2) |j_1 m_1; j_2 m_2\rangle . \tag{10.31}
\end{align}
$$

So, the product states are eigenstates of z-component of the total angular momentum. Next, note that using

$$\hat{\boldsymbol{J}}^2 = \hat{\boldsymbol{J}}^{(1)2} + \hat{\boldsymbol{J}}^{(2)2} + 2\hat{\boldsymbol{J}}^{(1)} \cdot \hat{\boldsymbol{J}}^{(2)} = \hat{\boldsymbol{J}}^{(1)2} + \hat{\boldsymbol{J}}^{(2)2} + \hat{J}_+^{(1)} \hat{J}_-^{(2)} + \hat{J}_-^{(1)} \hat{J}_+^{(2)} + 2\hat{J}_z^{(1)} \hat{J}_z^{(2)} , \tag{10.32}$$

173

$$\hat{J}_+ |jm\rangle = \hbar\sqrt{j(j+1) - m(m+1)}\,|j, m+1\rangle = \hbar\sqrt{(j+m+1)(j-m)}\,|j, m+1\rangle \tag{10.33}$$

and

$$\hat{J}_- |jm\rangle = \hbar\sqrt{j(j+1) - m(m-1)}\,|j, m-1\rangle = \hbar\sqrt{(j+m)(j-m+1)}\,|j, m-1\rangle \; ; \tag{10.34}$$

we obtain

$$\hat{\boldsymbol{J}}^2 |j_1 m_1; j_2 m_2\rangle = \hbar^2 \left[j_1(j_1+1) + j_2(j_2+1) + 2m_1 m_2 \right] |j_1 m_1; j_2 m_2\rangle$$
$$+\hbar^2 \sqrt{(j_1+m_1+1)(j_1-m_1)}\sqrt{(j_2+m_2)(j_2-m_2+1)}\,|j_1, m_1+1; j_2, m_2-1\rangle$$
$$+\hbar^2 \sqrt{(j_1+m_1)(j_1-m_1+1)}\sqrt{(j_2+m_2+1)(j_2-m_2)}\,|j_1, m_1-1; j_2, m_2+1\rangle$$
$$. \tag{10.35}$$

The direct product states are not eigenstates of the square of the total angular momentum. However, since both basis sets are complete on the same space, we can expand the total angular momentum states in terms of the direct product states as

$$|j_1 j_2 JM\rangle = \sum_{m_1, m_2} A^{JM}_{j_1 m_1; j_2 m_2} |j_1 m_1; j_2 m_2\rangle \; , \tag{10.36}$$

where the orthogonality of the direct product states can be used to extract the expansion coefficients as

$$A^{JM}_{j_1 m_1; j_2 m_2} = \langle j_1 m_1; j_2 m_2 | j_1 j_2 JM\rangle \; . \tag{10.37}$$

Applying the z-component of the total angular momentum to this expression gives

$$\hat{J}_z |j_1 j_2 JM\rangle = \sum_{m_1, m_2} \langle j_1 m_1; j_2 m_2 | j_1 j_2 JM\rangle\, \hat{J}_z |j_1 m_1; j_2 m_2\rangle \; . \tag{10.38}$$

Using the fact that both the total angular momentum states and the direct product states are eigenstates of this operator yields

$$\hbar M |j_1 j_2 JM\rangle = \hbar \sum_{m_1, m_2} \langle j_1 m_1; j_2 m_2 | j_1 j_2 JM\rangle\, (m_1 + m_2) |j_1 m_1; j_2 m_2\rangle \; . \tag{10.39}$$

The expression can only be true if the the factor $m_1 + m_2$ can be moved outside of the sum to reproduce the total angular momentum eigenstate. This means that the only terms that contribute to the sum are those for which $M = m_1 + m_2$. The total angular momentum eigenstate can now be written as

$$|j_1 j_2 JM\rangle = \sum_{m_1, m_2} \langle j_1 m_1; j_2 m_2 | j_1 j_2 JM\rangle\, |j_1 m_1; j_2 m_2\rangle \; , \tag{10.40}$$

where

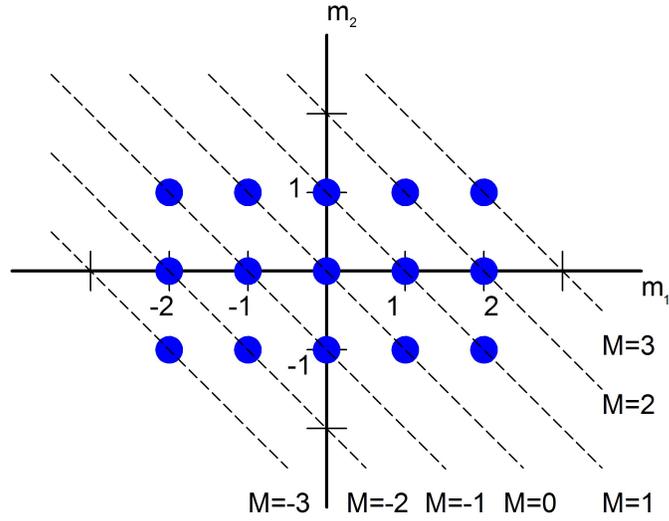$$\langle j_1 m_1; j_2 m_2 | j_1 j_2 JM\rangle = 0 \text{ if } M \neq m_1 + m_2 \; . \tag{10.41}$$

174

Figure 10.1: Direct product states for $j_1 = 2$ and $j_2 = 1$.

The expansion coefficients $\langle j_1 m_1; j_2 m_2 | j_1 j_2 JM \rangle$ are called *Clebsch-Gordan coefficients*.

The problem now is that in order to obtain the expansion we need to evaluate the Clebsch-Gordan coefficients. In fact, what will do now is to construct the direct product states and use this to obtain the values for the various Clebsch-Gordan coefficients. To see how this is done it is useful to consider Fig. 10.1. Here, all of the states corresponding to the direct product states for $j_1 = 2$ and $j_2 = 1$ are designated by the quantum numbers m_1 and m_2. The states with common values of $M = m_1 + m_2$ lie along the diagonal dashed lines. These are the states that can contribute to the expansion for a total angular momentum state for a given M. Note that for the maximum value $M = j_1 + j_2 = 3$ there is only one state, so there will only be one term in the expansion of the corresponding total angular momentum state. This state can be identified by evaluating

$$\begin{aligned}
\hat{\boldsymbol{J}}^2 |j_1 j_1; j_2 j_2\rangle &= \hbar^2 \left[j_1(j_1 + 1) + j_2(j_2 + 1) + 2 j_1 j_2 \right] |j_1 j_1; j_2 j_2\rangle \\
&= \hbar^2 (j_1 + j_2)(j_1 + j_2 + 1) |j_1 j_1; j_2 j_2\rangle \ . \qquad (10.42)
\end{aligned}$$

Therefore, this state has a total angular momentum $J = j_1 + j_2$. We can then write that

$$|j_1 j_2, j_1 + j_2, j_1 + j_2\rangle = |j_1 j_1; j_2 j_2\rangle \ . \qquad (10.43)$$

We can now obtain all of the states with value of J by noting that the lowering

operator for the total angular momentum states can be written as

$$\hat{\boldsymbol{J}}_- = \hat{\boldsymbol{J}}_-^{(1)} + \hat{\boldsymbol{J}}_-^{(2)} . \tag{10.44}$$

Applying this to (10.43) gives

$$\hat{\boldsymbol{J}}_- |j_1 j_2, j_1 + j_2, j_1 + j_2\rangle = \left(\hat{\boldsymbol{J}}_-^{(1)} + \hat{\boldsymbol{J}}_-^{(2)}\right) |j_1 j_1; j_2 j_2\rangle . \tag{10.45}$$

Using (10.34), the left-hand side of this expression can be evaluated as

$$\hat{\boldsymbol{J}}_- |j_1 j_2, j_1 + j_2, j_1 + j_2\rangle = \hbar\sqrt{2(j_1 + j_2)} |j_1 j_2, j_1 + j_2, j_1 + j_2 - 1\rangle \tag{10.46}$$

and the right-hand side as

$$\left(\hat{\boldsymbol{J}}_-^{(1)} + \hat{\boldsymbol{J}}_-^{(2)}\right) |j_1 j_1; j_2 j_2\rangle = \hbar\sqrt{2j_1} |j_1, j_1 - 1; j_2 j_2\rangle + \hbar\sqrt{2j_2} |j_1 j_1; j_2, j_2 - 1\rangle . \tag{10.47}$$

Equating these to expressions and solving for the total angular momentum state gives

$$|j_1 j_2, j_1 + j_2, j_1 + j_2 - 1\rangle = \sqrt{\frac{j_1}{j_1 + j_2}} |j_1, j_1 - 1; j_2 j_2\rangle + \sqrt{\frac{j_2}{j_1 + j_2}} |j_1 j_1; j_2, j_2 - 1\rangle . \tag{10.48}$$

We can then identify the two Clebsch-Gordan coefficients as

$$\langle j_1, j_1 - 1; j_2 j_2 | j_1 j_2, j_1 + j_2, j_1 + j_2 - 1\rangle = \sqrt{\frac{j_1}{j_1 + j_2}} \tag{10.49}$$

and

$$\langle j_1 j_1; j_2, j_2 - 1 | j_1 j_2, j_1 + j_2, j_1 + j_2 - 1\rangle = \sqrt{\frac{j_2}{j_1 + j_2}} . \tag{10.50}$$

Following this procedure and applying the total angular momentum lowering operator to (10.48), we can obtain the next state

$$\begin{aligned}
|j_1 j_2, j_1 + j_2, j_1 + j_2 - 2\rangle &= \sqrt{\frac{j_1}{j_1 + j_2}}\sqrt{\frac{2j_1 - 1}{2(j_1 + j_2) - 1}} |j_1, j_1 - 2; j_2 j_2\rangle \\
&\quad + 2\sqrt{\frac{j_1 j_2}{(j_1 + j_2)[2(j_1 + j_2) - 1]}} |j_1, j_1 - 1; j_2, j_2 - 1\rangle \\
&\quad + \sqrt{\frac{j_2}{j_1 + j_2}}\sqrt{\frac{2j_2 - 1}{2(j_1 + j_2) - 1}} |j_1 j_1; j_2 j_2 - 2\rangle . \tag{10.51}
\end{aligned}$$

We can continue in this manner until we obtain the last state for this value of J where $M = -j_1 - j_2$.

Next consider the case of $M = j_1 + j_2 - 1$. There are two direct product state which contribute to the expansion of the total angular momentum states with this value of M. We have obtained the state for $J = j_1 + j_2$, but since there are two direct product states, there must be two orthogonal combinations of these states. This means that there must be another total angular momentum state such that

$$\langle j_1 j_2 J, j_1 + j_2 - 1 | j_1 j_2, j_1 + j_2, j_1 + j_2 - 1 \rangle = 0 \tag{10.52}$$

and

$$\langle j_1 j_2 J, j_1 + j_2 - 1 | j_1 j_2 J, j_1 + j_2 - 1 \rangle = 1 \,. \tag{10.53}$$

These orthonormality conditions are satisfied by the state

$$|j_1 j_2 J, j_1 + j_2 - 1\rangle = \sqrt{\frac{j_1}{j_1 + j_2}} \, |j_1 j_1; j_2, j_2 - 1\rangle - \sqrt{\frac{j_2}{j_1 + j_2}} \, |j_1, j_1 - 1; j_2 j_2\rangle \,. \tag{10.54}$$

Applying the square of the total angular momentum operator to this state and using (10.35) gives

$$\hat{\boldsymbol{J}}^2 |j_1 j_2 J, j_1 + j_2 - 1\rangle = \hbar^2 (j_1 + j_2 - 1)(j_1 + j_2 - 1 + 1) |j_1 j_2 J, j_1 + j_2 - 1\rangle \,. \tag{10.55}$$

So, for this state, $J = j_1 + j_2 - 1$. That is,

$$|j_1 j_2, j_1 + j_2 - 1, j_1 + j_2 - 1\rangle = \sqrt{\frac{j_1}{j_1 + j_2}} \, |j_1 j_1; j_2, j_2 - 1\rangle - \sqrt{\frac{j_2}{j_1 + j_2}} \, |j_1, j_1 - 1; j_2 j_2\rangle \,. \tag{10.56}$$

We can now obtain all of the states for this value of J by applying the lowering operator to this state and each successive state.

Next, we need to consider the states with $M = j_1 + j_2 - 2$. In general, there are three direct product states that contribute to the total angular momentum states with this value of M, so there are three possible orthogonal states that can be constructed from them. Two of them are for $J = j_1 + j_2$ and $J = j_1 + j_2 - 1$. The third is obtained by writing a general linear combination of the three direct product states. The coefficients of this combination are obtained by orthogonalizing this combination to the other two states and then normalizing. The resulting state has $J = j_1 + j_2 - 2$. The lowering operator can then be used to obtain all of the states with this J. This procedure of orthogonalizing and lowering can be used until all possible states have been obtained for the given values of j_1 and j_2.

Note that as we have worked through this procedure we have obtained states with successively lower values of J. The fact that there are a fixed number of direct product states for a given j_1 and j_2 implies that there must by some minimum value of $J_{min} \geq 0$ which can be constructed. This can be determined by counting states. Since the direct product states and the total angular momentum states span the

same hilbert space, the number of states in each set must be the same. For the direct product states the number of states is given by

$$N_{\text{states}} = (2j_1 + 1)(2j_2 + 1).$$ (10.57)

For the total angular momentum states, the maximum value of J is $j_1 + j_2$ and for each value of J there are $2J + 1$ states corresponding to the possible values of M. We can calculate the total number of total angular momentum states as

$$
\begin{aligned}
N_{states} &= (2j_1 + 1)(2j_2 + 1) = \sum_{J=J_{min}}^{j_1+j_2} (2J + 1) = \sum_{J=0}^{j_1+j_2} (2J + 1) - \sum_{J=0}^{J_{min}-1} (2J + 1) \\
&= (j_1 + j_2 + 1)(j_1 + j_2) + j_1 + j_2 + 1 - J_{min}(J_{min} - 1) - J_{min} \\
&= (j_1 + j_2)^2 + 2(j_1 + j_2) + 1 - J_{min}^2.
\end{aligned}
$$ (10.58)

Solving this for J_{min}^2 gives

$$
\begin{aligned}
J_{min}^2 &= (j_1 + j_2)^2 + 2(j_1 + j_2) + 1 - (2j_1 + 1)(2j_2 + 1) \\
&= j_1^2 + 2j_1j_2 + j_2^2 + 2j_1 + 2j_2 + 1 - 4j_1j_2 - 2j_1 - 2j_2 - 1 \\
&= j_1^2 - 2j_1j_2 + j_2^2 = (j_1 - j_2)^2
\end{aligned}
$$ (10.59)

or

$$J_{min} = |j_1 - j_2|.$$ (10.60)

Therefore for given values of j_1 and j_2,

$$|j_1 - j_2| \le J \le j_1 + j_2.,$$ (10.61)

10.2.1   General Properties of Clebsch-Gordan Coefficients

Clearly, the procedure outlined in the previous section can become very tedious, particularly when the values of either of the two angular momenta in the direct product states become large. However, since this precess is very important in atomic and nuclear physics, the problem of coupling of angular momenta has been studied exhaustively and results have been tabulated in a number of books.[1] In addition the computer program *Mathematica* can generate values for Clebsch-Gordan coefficients.

It should also be noted that in describing the procedure of the previous section, we have implicitly made choices concerning the phase of each set of coefficients for a given J. The phases chosen are those of Edmonds which are now largely standard. With this choice all of the Clebsch-Gordan coefficients are real.

---

[1]Books in my own collection are: A. R. Edmonds, *Angular Momentum in Quantum Mechanics*, Princeton University Press (1996); M. E. Rose, *Elementary Theory of Angular Momentum*, Dover (1995); D.A. Varshalovich, A. N. Moskalev and V. K. Khersonskii, *Quantum Theory of Angular Momentum*, World Scientific (1988).

Calculations involving Clebsch-Gordan coefficients can be simplified by using a few general properties of these coefficients which are derived in most of the standard texts. First note that the notation that we have been using to this point for the Clebsch-Gordan coefficients is redundant in that the values of j_1 and j_2 appear twice in the designation of any coefficient. We will therefore introduce the simplification of notation

$$\langle j_1 m_1; j_2 m_2| \, j_3 m_3 \rangle \equiv \langle j_1 m_1; j_2 m_2| \, j_1 j_2 j_3 m_3 \rangle \, . \tag{10.62}$$

The unitarity properties of the Clebsch-Gordan coefficients are the simple result of the completeness and orthogonality of the states in the the two basis sets. Therefore,

$$\sum_{j_3=|j_1-j_2|}^{j_1+j_2} \sum_{m_3=-j_3}^{j_3} \langle j_1 m_1; j_2 m_2| \, j_3 m_3 \rangle \langle j_3 m_3| \, j_1 m_1'; j_2 m_2' \rangle = \delta_{m_1 m_1'} \delta_{m_2 m_2'} \tag{10.63}$$

and

$$\sum_{m_1=-j_1}^{j_1} \sum_{m_2=-j_2}^{j_2} \langle j_3 m_3| \, j_1 m_1; j_2 m_2 \rangle \langle j_1 m_1; j_2 m_2| \, j_3' m_3' \rangle = \delta_{j_3 j_3'} \delta_{m_3 m_3'} \delta(j_1 j_2 j_3) \, , \tag{10.64}$$

where $\delta(j_1 j_2 j_3) = 1$ if $|j_1 - j_2| \leq j_3 \leq j_1 + j_2$ and is zero otherwise.

The Clebsch-Gordan coefficients have a number of symmetry properties. Coefficients which differ only in the order of the values for the two product states are related by

$$\langle j_1 m_1; j_2 m_2| \, j_3 m_3 \rangle = (-1)^{j_1+j_2-j_3} \langle j_2, m_2; j_1, m_1| \, j_3, m_3 \rangle \, . \tag{10.65}$$

States that differ by the signs of all of the magnetic quantum numbers are related by

$$\langle j_1 m_1; j_2 m_2| \, j_3 m_3 \rangle = (-1)^{j_1+j_2-j_3} \langle j_1, -m_1; j_2, -m_2| \, j_3, -m_3 \rangle \, . \tag{10.66}$$

Note that this relation means that if we are forced to construct the states using the procedure of the previous section, it is only necessary to calculate the states for which $M \geq 0$ since the remaining coefficients can be obtained using this relation.

Cyclic permutations of the arguments of the Clebsch-Gordan coefficients are given by

$$\langle j_1 m_1; j_2 m_2| \, j_3 m_3 \rangle = (-1)^{j_2+m_2} \left(\frac{2j_3+1}{2j_1+1} \right)^{\frac{1}{2}} \langle j_2, -m_2; j_3 m_3| \, j_1 m_1 \rangle \tag{10.67}$$

and

$$\langle j_1 m_1; j_2 m_2| \, j_3 m_3 \rangle = (-1)^{j_1-m_1} \left(\frac{2j_3+1}{2j_2+1} \right)^{\frac{1}{2}} \langle j_3 m_3; j_1, -m_1| \, j_2, m_2 \rangle \, . \tag{10.68}$$

In actual calculations, may people use a more symmetric form of the vector coupling coefficients called *Wigner 3-j symbols* which are related to the Clebsch-Gordan coefficients. This is beyond the scope of this presentation.

10.2.2 Example: Addition of $j_1 = 1$ and $j_2 = 1$

As an example of the procedure that we have outlined for the construction of the total angular momentum states, consider the case where $j_1 = 1$ and $j_2 = 1$. Coupling of these two angular momenta results in states with $0 \geq J \geq 2$. Using (10.43) we obtain

$$|1122\rangle = |11; 11\rangle . \tag{10.69}$$

Equation (10.48) gives

$$|1121\rangle = \frac{1}{\sqrt{2}} |11; 10\rangle + \frac{1}{\sqrt{2}} |10; 11\rangle \tag{10.70}$$

and (10.51) gives

$$|1120\rangle = \frac{1}{\sqrt{6}} |11; 1, -1\rangle + \sqrt{\frac{2}{3}} |10; 10\rangle + \frac{1}{\sqrt{6}} |1, -1; 11\rangle . \tag{10.71}$$

The remaining $J = 2$ states can be obtained using (10.66).

Using (10.56) we can obtain the first state for $J = 1$

$$|1111\rangle = \frac{1}{\sqrt{2}} |11; 10\rangle - \frac{1}{\sqrt{2}} |10; 11\rangle . \tag{10.72}$$

To obtain the remaining states for $J = 1$ we need to apply the lowering operator to both sides of the state above. For the left-hand side

$$\hat{J}_- |1111\rangle = \hbar \sqrt{2} |1110\rangle \tag{10.73}$$

and for the right-hand side

$$(\hat{J}_-^{(1)} + \hat{J}_-^{(2)}) \left(\frac{1}{\sqrt{2}} |11; 10\rangle - \frac{1}{\sqrt{2}} |10; 11\rangle \right)$$
$$= \frac{\hbar}{\sqrt{2}} \left(\sqrt{2} |10; 10\rangle + \sqrt{2} |11; 1, -1\rangle \right) - \frac{\hbar}{\sqrt{2}} \left(\sqrt{2} |1, -1; 11\rangle + \sqrt{2} |10; 10\rangle \right)$$
$$= \hbar |11; 1, -1\rangle - \hbar |1, -1; 11\rangle . \tag{10.74}$$

Equating these two expression and solving gives

$$|1110\rangle = \frac{1}{\sqrt{2}} |11; 1, -1\rangle - \frac{1}{\sqrt{2}} |1, -1; 11\rangle . \tag{10.75}$$

The remaining $J = 1$ state can be found using (10.66).

The only remaining state is for $J = M = 0$. Note that there are three direct product states that have $M = 0$ which implies that there are three possible orthogonal states composed of linear combinations of these three states. We have already found

two of these, one for $J = 2$, $M = 0$ and the other for $J = 1$, $M = 0$. To obtain the $J = 0$ state we assume that the state is given by a general linear combination of the direct product states

$$|1100\rangle = a\,|11; 1, -1\rangle + b\,|10; 10\rangle + c\,|1, -1; 11\rangle \tag{10.76}$$

and then require that this state is orthogonal to the other two $M = 0$ states and is properly normalized. The orthogonality of the $M = 0$ states for $J = 1$ and $J = 0$ gives

$$0 = \langle 1110\,|1100\rangle = \frac{1}{\sqrt{2}}a - \frac{1}{\sqrt{2}}c\,, \tag{10.77}$$

which requires that

$$c = a\,. \tag{10.78}$$

The orthogonality of the $M = 0$ states for $J = 2$ and $J = 0$ gives

$$0 = \langle 1120\,|1100\rangle = \frac{1}{\sqrt{6}}a + \sqrt{\frac{2}{3}}b + \frac{1}{\sqrt{6}}c \tag{10.79}$$

which requires that

$$b = -a\,. \tag{10.80}$$

Finally, the normalization of the total angular momentum states for $J = 0$ gives

$$1 = \langle 1100\,|1100\rangle = 3a^2\,. \tag{10.81}$$

The yields

$$a = \frac{1}{\sqrt{3}}\,. \tag{10.82}$$

The $J = 0$ state is then given by

$$|1100\rangle = \frac{1}{\sqrt{3}}\,|11; 1, -1\rangle - \frac{1}{\sqrt{3}}\,|10; 10\rangle + \frac{1}{\sqrt{3}}\,|1, -1; 11\rangle\,. \tag{10.83}$$

By comparing the expressions for the total angular momentum states to the general expansion in terms of Clebsch-Gordan coefficients and direct product states, we can obtain the coefficients for this case. These are summarized in Table 10.1

Table 10.1: Clebsch-Gordan coefficients for $j_1 = j_2 = 1$.

j_3	m_3	m_1	m_2	$\langle 1m_1; 1m_2 \vert\, j_3 m_3 \rangle$
2	2	1	1	1
2	1	1	0	$\frac{1}{\sqrt{2}}$
2	1	0	1	$\frac{1}{\sqrt{2}}$
2	0	1	-1	$\frac{1}{\sqrt{6}}$
2	0	0	0	$\sqrt{\frac{2}{3}}$
2	0	-1	1	$\frac{1}{\sqrt{6}}$
1	1	1	0	$\frac{1}{\sqrt{2}}$
1	1	0	1	$-\frac{1}{\sqrt{2}}$
1	0	1	-1	$\frac{1}{\sqrt{2}}$
1	0	0	0	0
1	0	-1	1	$-\frac{1}{\sqrt{2}}$
0	0	1	-1	$\frac{1}{\sqrt{3}}$
0	0	0	0	$-\frac{1}{\sqrt{3}}$
0	0	-1	1	$\frac{1}{\sqrt{3}}$

Chapter 11

Rotation Matices and Spherical Tensor Operators

11.1 Matrix Representation of Rotations

In a previous chapter we derived the operators for rotations of the wave functions. We parameterized the rotation by specifying an axis about which the rotation occurs and the angle for rotation about this axis. In classical mechanics it is convenient to describe rotation of rigid bodies by parameterizing the rotation in terms of a set of three angles known as the Euler angles. This is done by constructing the final rotation as the result of three rotations about specified axes as is illustrated in Fig. 11.1. In the convention that we will use, the first rotation is through the angle α about the z axis. The second is a rotation through the angle β about the rotated y axis. The third is a rotation about the resulting z axis through the angle γ. The rotation can then be written as

$$\boldsymbol{R}(\boldsymbol{n},\theta) = \boldsymbol{R}(\boldsymbol{e}_{z''},\gamma)\boldsymbol{R}(\boldsymbol{e}_{y'},\beta)\boldsymbol{R}(\boldsymbol{e}_z,\alpha) = e^{-i\gamma J_{z''}}e^{-i\beta J_{y'}}e^{-i\alpha J_z} . \tag{11.1}$$

Similarly, we can parameterize the quantum mechanical rotation operator in terms of the Euler angles as

$$\hat{R}(\boldsymbol{n},\theta) = \hat{R}(\boldsymbol{e}_{z''},\gamma)\hat{R}(\boldsymbol{e}_{y'},\beta)\hat{R}(\boldsymbol{e}_z,\alpha) \equiv \hat{D}(\alpha,\beta,\gamma) = e^{-\frac{i}{\hbar}\gamma\hat{J}_{z''}}e^{-\frac{i}{\hbar}\beta\hat{J}_{y'}}e^{-\frac{i}{\hbar}\alpha\hat{J}_z} . \tag{11.2}$$

This is inconvenient since it refers two rotations about to axis which vary in direction according to the parameters of the rotation. We can reexpress this as rotations relative to a fixed set of coordinates by noting that operators for these operators can also be constructed by rotations defined in the initial fixed reference frame. For example

$$e^{-\frac{i}{\hbar}\beta\hat{J}_{y'}} = e^{-\frac{i}{\hbar}\alpha\hat{J}_z}e^{-\frac{i}{\hbar}\beta\hat{J}_y}e^{\frac{i}{\hbar}\alpha\hat{J}_z} \tag{11.3}$$

Similarly,

$$e^{-\frac{i}{\hbar}\gamma\hat{J}_{z''}} = e^{-\frac{i}{\hbar}\beta\hat{J}_{y'}}e^{-\frac{i}{\hbar}\gamma\hat{J}_{z'}}e^{\frac{i}{\hbar}\beta\hat{J}_{y'}} \tag{11.4}$$
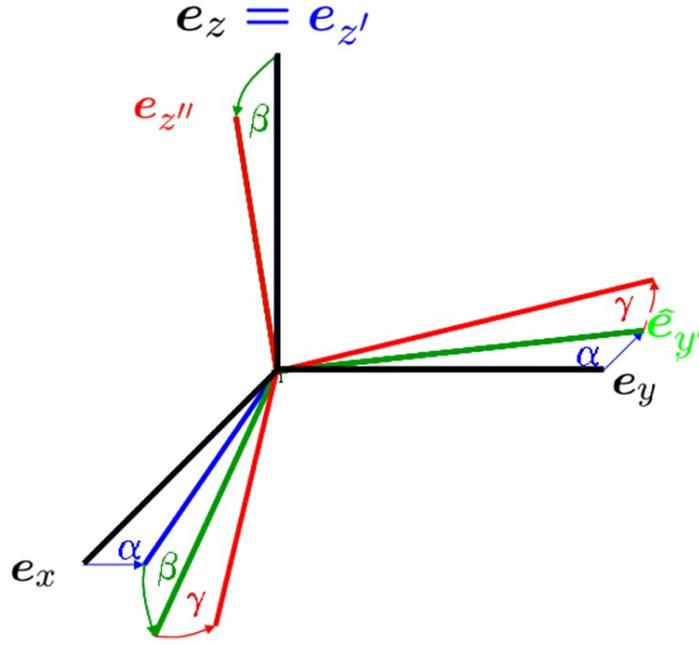
Figure 11.1: Rotations parameterized with the Euler angles.

and

$$e^{-\frac{i}{\hbar}\gamma\hat{J}_{z'}} = e^{-\frac{i}{\hbar}\alpha\hat{J}_z}e^{-\frac{i}{\hbar}\gamma\hat{J}_z}e^{\frac{i}{\hbar}\alpha\hat{J}_z} \,. \tag{11.5}$$

We can now rewrite the rotation operator as

$$\begin{aligned}
\hat{D}(\alpha,\beta,\gamma) &= e^{-\frac{i}{\hbar}\beta\hat{J}_{y'}}e^{-\frac{i}{\hbar}\gamma\hat{J}_{z'}}e^{\frac{i}{\hbar}\beta\hat{J}_{y'}}e^{-\frac{i}{\hbar}\beta\hat{J}_{y'}}e^{-\frac{i}{\hbar}\alpha\hat{J}_z} = e^{-\frac{i}{\hbar}\beta\hat{J}_{y'}}e^{-\frac{i}{\hbar}\gamma\hat{J}_{z'}}e^{-\frac{i}{\hbar}\alpha\hat{J}_z} \\
&= e^{-\frac{i}{\hbar}\alpha\hat{J}_z}e^{-\frac{i}{\hbar}\beta\hat{J}_y}e^{\frac{i}{\hbar}\alpha\hat{J}_z}e^{-\frac{i}{\hbar}\alpha\hat{J}_z}e^{-\frac{i}{\hbar}\gamma\hat{J}_z}e^{\frac{i}{\hbar}\alpha\hat{J}_z}e^{-\frac{i}{\hbar}\alpha\hat{J}_z} \\
&= e^{-\frac{i}{\hbar}\alpha\hat{J}_z}e^{-\frac{i}{\hbar}\beta\hat{J}_y}e^{-\frac{i}{\hbar}\gamma\hat{J}_z} \,,
\end{aligned} \tag{11.6}$$

where the rotations are now all defined in a fixed frame.

Our objective is to obtain the matrix representation of this operator in terms of the eigenstates of $\hat{\boldsymbol{J}}^2$ and J_z. Note that since

$$[\hat{\boldsymbol{J}}^2, \hat{J}_i] = 0 \,, \tag{11.7}$$

the commutator of $\hat{\boldsymbol{J}}^2$ with the rotation operator is

$$[\hat{\boldsymbol{J}}^2, \hat{D}(\alpha,\beta,\gamma)] = 0 \,. \tag{11.8}$$

Using this,

$$\hat{\boldsymbol{J}}^2\hat{D}(\alpha,\beta,\gamma)\ket{jm} = \hat{D}(\alpha,\beta,\gamma)\hat{\boldsymbol{J}}^2\ket{jm} = \hbar^2 j(j+1)\hat{D}(\alpha,\beta,\gamma)\ket{jm} \,. \tag{11.9}$$

184

This means that the rotation mixes states of the same j but different values of m. The matrix element of the rotation operator is then

$$\langle j'm'|\,\hat{D}(\alpha,\beta,\gamma)\,|jm\rangle = \delta_{j'j}\,\langle jm'|\,\hat{D}(\alpha,\beta,\gamma)\,|jm\rangle \,. \qquad (11.10)$$

We will define the matrix element of the rotation operator as

$$D^j_{m'm}(\alpha,\beta,\gamma) \equiv \langle jm'|\,\hat{D}(\alpha,\beta,\gamma)\,|jm\rangle \,. \qquad (11.11)$$

This is called the *Wigner rotation functions*. Using the definition of the rotation operator, we can write the rotation matrix element as

$$
\begin{aligned}
D^j_{m'm}(\alpha,\beta,\gamma) &= \langle jm'|\,e^{-\frac{i}{\hbar}\alpha\hat{J}_z}e^{-\frac{i}{\hbar}\beta\hat{J}_y}e^{-\frac{i}{\hbar}\gamma\hat{J}_z}\,|jm\rangle \\
&= e^{-i\alpha m'}\,\langle jm'|\,e^{-\frac{i}{\hbar}\beta\hat{J}_y}\,|jm\rangle\,e^{-i\gamma m} \\
&\equiv e^{-i\alpha m'}d^j_{m'm}(\beta)e^{-i\gamma m}\,. \qquad (11.12)
\end{aligned}
$$

By parameterizing the rotation in terms of the Euler angles, we can simplify the matrix elements of the rotation operator to consist of simple factors times the matrix element of a rotation about a single fixed axis.

Using the completeness of the eigenstates, the rotation of a state can be written as

$$\hat{D}(\alpha,\beta,\gamma)\,|jm\rangle = \sum_{j'm'}|j'm'\rangle\,\langle j'm'|\,\hat{D}(\alpha,\beta,\gamma)\,|jm\rangle = \sum_{m'}|jm'\rangle\,D^j_{m'm}(\alpha,\beta,\gamma)\,. \qquad (11.13)$$

For the special case where $j = l$ is an integer, we can use this to describe the effect of the rotation operator on a spherical harmonic as

$$\hat{D}(\alpha,\beta,\gamma)Y_{lm}(\boldsymbol{e}_r) = \sum_{m'}D^l_{m'm}(\alpha,\beta,\gamma)Y_{lm'}(\boldsymbol{e}_r)\,, \qquad (11.14)$$

where we have indicated the angles in the argument of the spherical harmonic in terms of the radial unit vector. We will return to this expression later.

We can now derive a number of identities involving the rotation functions. First, since $\hat{D}(\alpha,\beta,\gamma)$ is unitary,

$$\hat{D}^\dagger(\alpha,\beta,\gamma)\hat{D}(\alpha,\beta,\gamma) = 1\,. \qquad (11.15)$$

The matrix element of the inverse rotation can then be determined to be

$$\langle jm|\,\hat{D}^\dagger(\alpha,\beta,\gamma)\,|jm'\rangle = \langle jm'|\,\hat{D}(\alpha,\beta,\gamma)\,|jm\rangle^* = D^{j*}_{m'm}(\alpha,\beta,\gamma)\,. \qquad (11.16)$$

Taking the matrix element of the unitarity condition, gives

$$
\begin{aligned}
\langle j'm' \, | jm \rangle &= \delta_{j'j}\delta_{m'm} = \langle j'm'| \, \hat{D}^\dagger(\alpha,\beta,\gamma)\hat{D}(\alpha,\beta,\gamma)\, |jm \rangle \\
&= \sum_{JM} \langle j'm'| \, \hat{D}^\dagger(\alpha,\beta,\gamma) \, |JM \rangle \, \langle JM| \, \hat{D}(\alpha,\beta,\gamma) \, |jm \rangle \\
&= \sum_{JM} \delta_{j'J} D^{j'*}_{Mm'}(\alpha,\beta,\gamma)\delta_{Jj} D^{j}_{Mm}(\alpha,\beta,\gamma) \\
&= \delta_{j'j} \sum_{M} D^{j*}_{Mm'}(\alpha,\beta,\gamma) D^{j}_{Mm}(\alpha,\beta,\gamma) \,. \qquad (11.17)
\end{aligned}
$$

So,

$$
\delta_{m'm} = \sum_{M} D^{j*}_{Mm'}(\alpha,\beta,\gamma) D^{j}_{Mm}(\alpha,\beta,\gamma) \,. \qquad (11.18)
$$

Next, consider the Clebsch-Gordan series

$$
|jm \rangle = \sum_{m_1 m_2} \langle j_1 m_1; j_2 m_2 \, | jm \rangle \, |j_1 m_1; j_2 m_2 \rangle \,. \qquad (11.19)
$$

Using the orthogonality properties of the Clebsch-Gordan coefficients the can be inverted to give

$$
|j_1 m_1; j_2 m_2 \rangle = \sum_{jm} \langle j_1 m_1; j_2 m_2 \, | jm \rangle \, |jm \rangle \,. \qquad (11.20)
$$

Acting on (11.19) with the rotation operator,

$$
\begin{aligned}
\hat{D}(\alpha,\beta,\gamma) \, |jm \rangle &= \sum_{m_1 m_2} \langle j_1 m_1; j_2 m_2 \, | jm \rangle \, \hat{D}(\alpha,\beta,\gamma) \, |j_1 m_1; j_2 m_2 \rangle \\
&= \sum_{m_1 m_2} \langle j_1 m_1; j_2 m_2 \, | jm \rangle \, \hat{D}(\alpha,\beta,\gamma) \, |j_1 m_1 \rangle \, \hat{D}(\alpha,\beta,\gamma) \, |j_2 m_2 \rangle \\
&= \sum_{m_1 m_2} \sum_{m'_1 m'_2} \langle j_1 m_1; j_2 m_2 \, | jm \rangle \, |j_1 m'_1 \rangle \, \langle j_1 m'_1| \, \hat{D}(\alpha,\beta,\gamma) \, |j_1 m_1 \rangle \\
&\qquad \times |j_2 m'_2 \rangle \, \langle j_2 m'_2| \, \hat{D}(\alpha,\beta,\gamma) \, |j_2 m_2 \rangle \\
&= \sum_{m_1 m_2} \sum_{m'_1 m'_2} \langle j_1 m_1; j_2 m_2 \, | jm \rangle \, D^{j_1}_{m'_1 m_1}(\alpha,\beta,\gamma) D^{j_2}_{m'_2 m_2}(\alpha,\beta,\gamma) \\
&\qquad \times |j_1 m'_1; j_2 m'_2 \rangle \,. \qquad (11.21)
\end{aligned}
$$

Substituting for the direct product state in the last line using (11.20) yields

$$
\begin{aligned}
\hat{D}(\alpha,\beta,\gamma) \, |jm \rangle &= \sum_{m_1 m_2} \sum_{m'_1 m'_2} \langle j_1 m_1; j_2 m_2 \, | jm \rangle \, D^{j_1}_{m'_1 m_1}(\alpha,\beta,\gamma) D^{j_2}_{m'_2 m_2}(\alpha,\beta,\gamma) \\
&\qquad \sum_{j''m''} \langle j_1 m'_1; j_2 m'_2 \, | j''m'' \rangle \, |j''m'' \rangle \,. \qquad (11.22)
\end{aligned}
$$

Multiplying this from the left with $\langle jm'|$ and using the orthonormality of the eigenstates gives the identity

$$D^j_{m'm}(\alpha, \beta, \gamma) = \sum_{m_1 m_2} \sum_{m'_1 m'_2} \langle j_1 m_1; j_2 m_2 \,|jm\rangle \langle j_1 m'_1; j_2 m'_2 \,|jm'\rangle$$
$$\times D^{j_1}_{m'_1 m_1}(\alpha, \beta, \gamma) D^{j_2}_{m'_2 m_2}(\alpha, \beta, \gamma). \tag{11.23}$$

Starting with (11.20) and following a similar procedure,

$$\hat{D}(\alpha, \beta, \gamma) \,|j_1 m_1; j_2 m_2\rangle = \hat{D}(\alpha, \beta, \gamma) \,|j_1 m_1\rangle \hat{D}(\alpha, \beta, \gamma) \,|j_2 m_2\rangle$$
$$= \sum_{jm} \langle j_1 m_1; j_2 m_2 \,|jm\rangle \sum_{j'm'} |j'm'\rangle \langle j'm'| \,\hat{D}(\alpha, \beta, \gamma) \,|jm\rangle$$
$$= \sum_{jmm'} \langle j_1 m_1; j_2 m_2 \,|jm\rangle D^j_{m'm}(\alpha, \beta, \gamma) \,|jm'\rangle$$
$$= \sum_{jmm'} \langle j_1 m_1; j_2 m_2 \,|jm\rangle D^j_{m'm}(\alpha, \beta, \gamma)$$
$$\times \sum_{m''_1 m''_2} \langle j_1 m''_1; j_2 m''_2 \,|jm'\rangle \,|j_1 m''_1; j_2 m''_2\rangle. \tag{11.24}$$

Multiplying from the left by $\langle j_1 m'_1; j_2 m'_2|$ and using the orthogonality of the direct product states yields the identity

$$D^{j_1}_{m'_1 m_1}(\alpha, \beta, \gamma) D^{j_2}_{m'_2 m_2}(\alpha, \beta, \gamma) = \sum_{jmm'} \langle j_1 m_1; j_2 m_2 \,|jm\rangle \langle j_1 m'_1; j_2 m'_2 \,|jm'\rangle D^j_{m'm}(\alpha, \beta, \gamma).$$
$$\tag{11.25}$$

Returning to the problem of the transformation properties of the spherical harmonics given by (11.14), we first want to define a column vector containing the components of the radial unit vector

$$\underline{e}_r = \begin{pmatrix} \sin\theta\cos\phi \\ \sin\theta\sin\phi \\ \cos\theta \end{pmatrix}. \tag{11.26}$$

Now define a column vector containing the components of the spherical harmonic with $l = 1$

$$\underline{Y}_1(\boldsymbol{e}_r) = \begin{pmatrix} Y_{11}(\theta, \phi) \\ Y_{10}(\theta, \phi) \\ Y_{1-1}(\theta, \phi) \end{pmatrix} = \sqrt{\frac{3}{4\pi}} \begin{pmatrix} -\frac{1}{\sqrt{2}}\sin\theta e^{i\phi} \\ \cos\theta \\ \frac{1}{\sqrt{2}}\sin\theta e^{-i\phi} \end{pmatrix} = \sqrt{\frac{3}{4\pi}} \begin{pmatrix} -\frac{1}{\sqrt{2}}(e_{r1} + ie_{r2}) \\ e_{r3} \\ \frac{1}{\sqrt{2}}(e_{r1} - ie_{r2}) \end{pmatrix}.$$
$$\tag{11.27}$$

This can be rewritten as

$$\underline{Y}_1(\boldsymbol{e}_r) = \sqrt{\frac{3}{4\pi}} \begin{pmatrix} -\frac{1}{\sqrt{2}} & -\frac{i}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \\ \frac{1}{\sqrt{2}} & -\frac{i}{\sqrt{2}} & 0 \end{pmatrix} \underline{e}_r = \sqrt{\frac{3}{4\pi}} \underline{\underline{U}}\, \underline{e}_r, \tag{11.28}$$

where the matrix $\underline{\underline{U}}$ defines a transformation of a vector from a cartesian basis to a spherical basis. The inverse of this expression is then

$$e_r = \sqrt{\frac{4\pi}{3}} \underline{\underline{U}}^{-1} \underline{Y}_1(e_r) \,. \tag{11.29}$$

We can now use the transformation property of wave functions under rotation to give

$$\begin{aligned}
\hat{D}(\alpha, \beta, \gamma) Y_{1m}(e_r) &= Y_{1m}(R^{-1}e_r) = \sqrt{\frac{3}{4\pi}} U_{mi} R_{ij}^{-1} e_{rj} \\
&= \sqrt{\frac{3}{4\pi}} U_{mi} R_{ij}^{-1} \sqrt{\frac{4\pi}{3}} (U^{-1})_{jm} Y_{1m'}(e_r) \\
&= (UR^{-1}U^{-1})_{mm'} Y_{1m'}(e_r) \,.
\end{aligned} \tag{11.30}$$

Substituting this into (11.14),

$$\sum_{m'} (UR^{-1}U^{-1})_{mm'} Y_{1m'}(e_r) = \sum_{m'} D^1_{m'm}(\alpha, \beta, \gamma) Y_{1m'}(e_r) \,. \tag{11.31}$$

Comparison of the two side of this expression allows us to identify

$$(UR^{-1}U^{-1})_{mm'} = D^1_{m'm}(\alpha, \beta, \gamma) \,. \tag{11.32}$$

This provides a direct method for calculating the rotation functions of rank 1.

11.2   Irreducible Spherical Tensor Operators

We can now obtain the transformation properties of the coordinate operator. The coordinate operator can be expressed in the spherical basis using

$$\hat{r}_m = U_{mi} \hat{r}_i \,. \tag{11.33}$$

The rotation of this operator is defined as

$$\hat{D}(\alpha, \beta, \gamma) \hat{r}_m \hat{D}^{-1}(\alpha, \beta, \gamma) = U_{mi} \hat{D}(\alpha, \beta, \gamma) \hat{r}_i \hat{D}^{-1}(\alpha, \beta, \gamma) \,. \tag{11.34}$$

Since this expression is true for any rotation we can consider the case of an infinitesimal rotation where

$$\begin{aligned}
\hat{D}(\alpha, \beta, \gamma) \hat{r}_i \hat{D}^{-1}(\alpha, \beta, \gamma) &\cong \left(1 - \frac{i}{\hbar} \theta \boldsymbol{n} \cdot \hat{\boldsymbol{L}}\right) \hat{r}_i \left(1 + \frac{i}{\hbar} \theta \boldsymbol{n} \cdot \hat{\boldsymbol{L}}\right) \\
&\cong \hat{r}_i + \frac{i}{\hbar} \theta n_j [\hat{r}_i, \hat{L}_j] \,.
\end{aligned} \tag{11.35}$$

The commutator of the coordinates of position operator and those of the orbital angular momentum operator is

$$[\hat{r}_i, \hat{L}_j] = i\hbar\epsilon_{ijk}\hat{r}_k .\tag{11.36}$$

Recalling that the antisymmetric tensor is related to the coordinates of the generators of rotation of three-dimension vectors by $-i\epsilon_{ijk} = (J_i)_{jk}$, the rotation of the coordinate operator can be written as

$$\begin{aligned}
\hat{D}(\alpha,\beta,\gamma)\hat{r}_i\hat{D}^{-1}(\alpha,\beta,\gamma) &\cong \hat{r}_i + \frac{i}{\hbar}\theta n_j i\hbar\epsilon_{ijk}\hat{r}_k = \hat{r}_i + i\theta n_j(-i\epsilon_{jik})\hat{r}_k \\
&= \hat{r}_i + i\theta n_j(J_j)_{ik}\hat{r}_k = (\delta_{ik} + i\theta(\boldsymbol{n}\cdot\boldsymbol{J})_{ik})\,\hat{r}_k \\
&\cong (R^{-1})_{ik}\hat{r}_k .
\end{aligned}\tag{11.37}$$

The rotation of the position operator in the spherical basis then becomes

$$\begin{aligned}
\hat{D}(\alpha,\beta,\gamma)\hat{r}_m\hat{D}^{-1}(\alpha,\beta,\gamma) &= U_{mi}(R^{-1})_{ik}\hat{r}_k = U_{mi}(R^{-1})_{ik}(U^{-1})_{km'}\hat{r}_{m'} \\
&= (UR^{-1}U^{-1})_{mm'}\hat{r}_{m'} .
\end{aligned}\tag{11.38}$$

From our examination of the rotation properties of Y_{1m} we can now write the transformation of the position operator in the spherical basis as

$$\hat{D}(\alpha,\beta,\gamma)\hat{r}_m\hat{D}^{-1}(\alpha,\beta,\gamma) = \sum_{m'} D^1_{m'm}(\alpha,\beta,\gamma)\hat{r}'_m .\tag{11.39}$$

This means that the position vector operator in the spherical basis transforms like an object with $J = 1$. We can generalize this by defining a class of operators called *irreducible spherical tensor operators* which are defined such that the transform under rotations as

$$\hat{D}(\alpha,\beta,\gamma)\hat{T}^J_M\hat{D}^{-1}(\alpha,\beta,\gamma) = \sum_{M'} D^J_{M'M}(\alpha,\beta,\gamma)\hat{T}^J_{M'}\tag{11.40}$$

A spherical tensor operator $\hat{T}^J_M$ is to said to be of *rank J* and *projection M*.

It can be shown that if $\hat{T}^{(1)J_1}_{M_1}$ and $\hat{T}^{(2)J_2}_{M_2}$ are spherical tensor operators, then

$$\hat{T}^J_M = \sum_{M_1 M_2} \langle J_1 M_1; J_2 M_2 \,|JM\rangle\, \hat{T}^{(1)J_1}_{M_1}\hat{T}^{(2)J_2}_{M_2}\tag{11.41}$$

is also a spherical tensor operator. Using the orthogonality condition for the Clebsch-Gordan coefficients, this can be inverted to give

$$\hat{T}^{(1)J_1}_{M_1}\hat{T}^{(2)J_2}_{M_2} = \sum_{JM} \langle J_1 M_1; J_2 M_2 \,|JM\rangle\, \hat{T}^J_M .\tag{11.42}$$

This product of spherical tensor operators is *reducible* since it can be written as a sum over irreducible tensor operators.

Consider the case of (11.41) where $J = M = 0$. This can only happen when $J_1 = J_2$. Then,

$$
\begin{aligned}
\hat{T}_0^0 &= \sum_{M_1 M_2} \langle J_1 M_1; J_1 M_2 \,|00\rangle \, \hat{T}_{M_1}^{(1)J_1} \hat{T}_{M_2}^{(2)J_1} = \sum_{M_1} \langle J_1 M_1; J_1, -M_1 \,|00\rangle \, \hat{T}_{M_1}^{(1)J_1} \hat{T}_{-M_1}^{(2)J_1} \\
&= \sum_{M_1} \frac{1}{\sqrt{2J_1 + 1}} (-1)^{J_1 - M_1} \hat{T}_{M_1}^{(1)J_1} \hat{T}_{-M_1}^{(2)J_1} \, .
\end{aligned}
\tag{11.43}
$$

If we apply this to two spherical vector operators $\hat{T}_{M_1}^{(1)1} = \hat{A}_{M_1}$ and $\hat{T}_{M_2}^{(2)1} = \hat{B}_{M_1}$ then

$$
\begin{aligned}
\hat{T}_0^0 &= \frac{1}{\sqrt{3}} \sum_{M_1} (-1)^{1 - M_1} \hat{A}_{M_1} \hat{B}_{-M_1} = \frac{1}{\sqrt{3}} \left(\hat{A}_1 \hat{B}_{-1} - \hat{A}_0 \hat{B}_0 + \hat{A}_{-1} \hat{B}_1 \right) \\
&= -\frac{1}{\sqrt{3}} \hat{\boldsymbol{A}} \cdot \hat{\boldsymbol{B}} \, .
\end{aligned}
\tag{11.44}
$$

For this reason, it is conventional to define the inner product of two irreducible tensor operators of equal, integral rank as

$$
\hat{\boldsymbol{T}}^{(1)J} \cdot \hat{\boldsymbol{T}}^{(2)J} = \sum_M (-1)^M \hat{T}_M^{(1)J} \hat{T}_{-M}^{(2)J} \, .
\tag{11.45}
$$

Now consider the components of the $J = 1$ spherical tensor operator obtained from combining the vector operators. This is

$$
\hat{T}_M^1 = \sum_{M_1 M_2} \langle 1 M_1; 1 M_2 \,|1M\rangle \, \hat{A}_{M_1} \hat{B}_{M_2} = \sum_{M_1} \langle 1 M_1; 1, M - M_1 \,|1M\rangle \, \hat{A}_{M_1} \hat{B}_{M - M_1} \, .
\tag{11.46}
$$

From this,

$$
\hat{T}_1^1 = \langle 11; 10 \,|11\rangle \, \hat{A}_1 \hat{B}_0 + \langle 10; 11 \,|11\rangle \, \hat{A}_0 \hat{B}_1 = \frac{1}{\sqrt{2}} \left(\hat{A}_1 \hat{B}_0 - \hat{A}_0 \hat{B}_1 \right) ,
\tag{11.47}
$$

$$
\begin{aligned}
\hat{T}_0^1 &= \langle 11; 1, -1 \,|10\rangle \, \hat{A}_1 \hat{B}_{-1} + \langle 10; 10 \,|10\rangle \, \hat{A}_0 \hat{B}_0 + \langle 1, -1; 11 \,|10\rangle \, \hat{A}_{-1} \hat{B}_1 \\
&= \frac{1}{\sqrt{2}} \left(\hat{A}_1 \hat{B}_{-1} - \hat{A}_{-1} \hat{B}_1 \right)
\end{aligned}
\tag{11.48}
$$

and

$$
\hat{T}_{-1}^1 = \langle 10; 1, -1 \,|1, -1\rangle \, \hat{A}_0 \hat{B}_{-1} + \langle 1, -1; 10 \,|1, -1\rangle \, \hat{A}_{-1} \hat{B}_0 = \frac{1}{\sqrt{2}} \left(\hat{A}_0 \hat{B}_{-1} - \hat{A}_{-1} \hat{B}_0 \right) .
\tag{11.49}
$$

These expressions can be summarized as

$$
\hat{T}_M^1 = \frac{i}{\sqrt{2}} (\hat{\boldsymbol{A}} \times \hat{\boldsymbol{B}})_M \, .
\tag{11.50}
$$

190

11.2.1 The Tensor Force

An application of spherical tensor operators to the simplification of an operator is the case of the tensor force between nucleons. The contribution to the total interaction potential for two nucleons is of the form

$$V_T(\hat{\boldsymbol{r}}) = F(\hat{r}) \left(\boldsymbol{\sigma}^{(1)} \cdot \hat{\boldsymbol{r}} \boldsymbol{\sigma}^{(2)} \cdot \hat{\boldsymbol{r}} - \frac{1}{3} \hat{r}^2 \boldsymbol{\sigma}^{(1)} \cdot \boldsymbol{\sigma}^{(2)} \right) , \qquad (11.51)$$

where $\boldsymbol{\sigma}^{(1)}$ and $\boldsymbol{\sigma}^{(1)}$ are the sets of Pauli spin matrices for nucleon 1 and nucleon 2, and $\hat{\boldsymbol{r}}$ is the operator for the relative position of the two nucleons.

If we define the spherical tensors,

$$\hat{R}_M^J = \sum_{M_1 M_2} \langle 1M_1; 1M_2 \, | JM \rangle \, \hat{r}_{M_1} \hat{r}_{M_2} \qquad (11.52)$$

and

$$\hat{S}_M^J = \sum_{M_1 M_2} \langle 1M_1; 1M_2 \, | JM \rangle \, \hat{\sigma}_{M_1}^{(1)} \hat{\sigma}_{M_2}^{(2)} , \qquad (11.53)$$

then

$$\hat{r}^2 = -\sqrt{3} \hat{R}_0^0 \qquad (11.54)$$

and

$$\boldsymbol{\sigma}^{(1)} \cdot \boldsymbol{\sigma}^{(2)} = -\sqrt{3} \hat{S}_0^0 . \qquad (11.55)$$

Next, consider

$$
\begin{aligned}
\boldsymbol{\sigma}^{(1)} \cdot \hat{\boldsymbol{r}} \boldsymbol{\sigma}^{(2)} \cdot \hat{\boldsymbol{r}} &= \sum_\lambda (-1)^\lambda \sigma_\lambda^{(1)} \hat{r}_{-\lambda} \sum_{\lambda'} (-1)^{\lambda'} \sigma_{\lambda'}^{(2)} \hat{r}_{-\lambda'} = \sum_{\lambda\lambda'} (-1)^{\lambda+\lambda'} \sigma_\lambda^{(1)} \sigma_{\lambda'}^{(2)} \hat{r}_{-\lambda} \hat{r}_{-\lambda'} \\
&= \sum_{\lambda\lambda'} (-1)^{\lambda+\lambda'} \sum_{JM} \langle 1\lambda; 1\lambda' \, | JM \rangle \hat{S}_M^J \sum_{J'M'} \langle 1, -\lambda; 1, -\lambda' \, | J'M' \rangle \hat{R}_{M'}^{J'} \\
&= \sum_{\lambda\lambda'} (-1)^{\lambda+\lambda'} \sum_{JM} \sum_{J'M'} \langle 1\lambda; 1\lambda' \, | JM \rangle \langle 1, -\lambda; 1, -\lambda' \, | J'M' \rangle \hat{S}_M^J \hat{R}_{M'}^{J'} \\
&= \sum_{\lambda\lambda'} \sum_{JM} \sum_{J'M'} (-1)^M \langle 1\lambda; 1\lambda' \, | JM \rangle (-1)^{1+1-J'} \langle 1, \lambda; 1, \lambda' \, | J', -M' \rangle \\
&\qquad \times \hat{S}_M^J \hat{R}_{M'}^{J'} \\
&= \sum_{JM} \sum_{J'M'} (-1)^M (-1)^{J'} \sum_{\lambda\lambda'} \langle 1\lambda; 1\lambda' \, | JM \rangle \langle 1, \lambda; 1, \lambda' \, | J', -M' \rangle \\
&\qquad \times \hat{S}_M^J \hat{R}_{M'}^{J'} \\
&= \sum_{JM} \sum_{J'M'} (-1)^M (-1)^{J'} \delta_{JJ'} \delta_{M,-M'} \hat{S}_M^J \hat{R}_{M'}^{J'} = \sum_{JM} (-1)^M (-1)^J \hat{S}_M^J \hat{R}_{-M}^J
\end{aligned}
$$
$$(11.56)$$

Note that since $\hat{R}^1_M$ is related to the cross product of $\hat{r}$ with itself, the $J = 1$ term in this sum will be zero. The tensor force can now be written as

$$V_T(\hat{r}) = F(\hat{r}) \left(\sum_{JM}(-1)^M(-1)^J \hat{S}^J_M \hat{R}^J_{-M} - \hat{S}^0_0 \hat{R}^0_0 \right) = F(\hat{r}) \sum_M (-1)^M \hat{S}^2_M \hat{R}^2_{-M} .$$

(11.57)

11.3   The Wigner-Eckart Theorem

From a practical standpoint, the purpose of introducing all of this technology is that the matrix elements of spherical tensors are relatively easy to calculate due to the Wigner-Eckart theorem. To see how this works consider the matrix element of (11.40). This gives

$$
\begin{aligned}
\sum_{M'} D^J_{M'M} \langle j'm'| \hat{T}^J_{M'} |jm\rangle &= \langle j'm'| \hat{D}\hat{T}^J_M \hat{D}^{-1} |jm\rangle \\
&= \sum_{\mu'\mu} \langle j'm'| \hat{D} |j'\mu'\rangle \langle j'\mu'| \hat{T}^J_M |j\mu\rangle \langle j\mu| \hat{D}^{-1} |jm\rangle \\
&= \sum_{\mu'\mu} D^{j'}_{m'\mu'} \langle j'\mu'| \hat{T}^J_M |j\mu\rangle D^{j*}_{m\mu} \\
&= \sum_{\mu'\mu} D^{j'}_{m'\mu'}(-1)^{m-\mu} D^j_{-m,-\mu} \langle j'\mu'| \hat{T}^J_M |j\mu\rangle \quad (11.58)
\end{aligned}
$$

where we have used the identity

$$D^{j*}_{mm'} = (-1)^{m-m'} D^j_{-m,-m'} .$$

(11.59)

We can now use (11.25) to rewrite this as

$$
\begin{aligned}
\sum_{M'} D^J_{M'M} \langle j'm'| \hat{T}^J_{M'} |jm\rangle &= \sum_{\mu'\mu} \sum_{\mathcal{JMM'}} (-1)^{m-\mu} \langle j'\mu'; j, -\mu \,|\mathcal{JM}\rangle \langle j'm'; j, -m \,|\mathcal{JM'}\rangle \\
&\qquad \times D^{\mathcal{J}}_{\mathcal{M'M}} \langle j'\mu'| \hat{T}^J_M |j\mu\rangle .
\end{aligned}
$$

(11.60)

For the two sides of this expression to be equivalent, the matrix elements of the spherical tensor operator must be of the form

$$\langle j'm'| \hat{T}^J_M |jm\rangle = (-1)^{j-m} \langle j'm'; j, -m \,|JM\rangle (2J+1)^{-\frac{1}{2}} \langle j'\| \hat{T}^J \|j\rangle ,$$

(11.61)

where the factor $(2J+1)^{-\frac{1}{2}}$ is chosen for convenience and the factor $\langle j'\| \hat{T}^J \|j\rangle$ the *reduced matrix element*. This is the *Wigner-Eckart Theorem*. Note that all of the information about the projections quantum numbers m, m' and M is contained in the Clebsch-Gordan coefficient and that this coefficient requires that $|j' - j| \le J \le j' + j$

and $m' = m + M$. This is equivalent to requiring that angular momentum be conserved. The only factor that contains information about the particular operator involved in the matrix element is the reduced matrix element. The rest of the expression is generic to all operators of a given rank and is geometric rather than dynamic.

We now have the problem of how to obtain the reduced matrix element. The solution to this is contained in the Wigner-Eckart theorem itself. In order to obtain the reduced matrix element, we need to evaluate the full matrix element with one particular choice of values for the projection quantum numbers. Obviously, this should be done for the case where the calculation of the matrix element will be the simplest. The Wigner-Eckart theorem can then be inverted to obtain the reduced matrix element which can then be used in the theorem to produce all other matrix elements.

For example, consider the case of the Pauli matrices in the spherical basis,

$$\sigma_{\pm 1} = \mp \frac{1}{\sqrt{2}} (\sigma_1 \pm i\sigma_2) \tag{11.62}$$

$$\sigma_0 = \sigma_3. \tag{11.63}$$

Clearly, it is easy to calculate

$$\left\langle \tfrac{1}{2}\tfrac{1}{2} \right| \sigma_0 \left| \tfrac{1}{2}\tfrac{1}{2} \right\rangle = 1. \tag{11.64}$$

Inverting Wigner-Eckart theorem (11.61), the reduced matrix element is given by

$$\langle j' \| \hat{T}^J \| j \rangle = (-1)^{j-m} \sqrt{2J+1} \frac{\langle j'm' | \hat{T}^J_{M'} | jm \rangle}{\langle j'm'; j, -m | JM \rangle}. \tag{11.65}$$

For our example, this gives

$$\langle \tfrac{1}{2} \| \sigma \| \tfrac{1}{2} \rangle = (-1)^{\frac{1}{2}-\frac{1}{2}} \sqrt{3} \frac{\left\langle \tfrac{1}{2}\tfrac{1}{2} \right| \sigma_0 \left| \tfrac{1}{2}\tfrac{1}{2} \right\rangle}{\left\langle \tfrac{1}{2}\tfrac{1}{2}; \tfrac{1}{2}, -\tfrac{1}{2} | 10 \right\rangle} = \sqrt{3} \frac{1}{\frac{1}{\sqrt{2}}} = \sqrt{6}. \tag{11.66}$$

We can then use the Wigner-Eckart Theorem to write

$$\left\langle \tfrac{1}{2}m' \right| \sigma_\lambda \left| \tfrac{1}{2}m \right\rangle = (-1)^{\frac{1}{2}-m} \left\langle \tfrac{1}{2}m'; \tfrac{1}{2}, -m | 1\lambda \right\rangle \frac{1}{\sqrt{3}} \sqrt{6} = (-1)^{\frac{1}{2}-m} \sqrt{2} \left\langle \tfrac{1}{2}m'; \tfrac{1}{2}, -m | 1\lambda \right\rangle \tag{11.67}$$

193

Chapter 12

Time-Independent Perturbation Theory

In this chapter we will be developing the formalism of time-independent perturbation theory. Since the general treatment of this problem will become abstract and complicated, it is useful to first focus on a simple problem that will help motivate this discussion and will help to clarify some of the technical issues that arise. The problem that we will consider is that of determining the spectrum and wave functions for the hydrogen atom.

12.1 The Hydrogen Atom

We have already constructed a simple model of the hydrogen atom in Section 7.5 where we solved the three-dimensional Schrödinger equation with a simple Coulomb potential. The Hamiltonian for this model is

$$\hat{H}_0 = \frac{\hat{\boldsymbol{p}}^2}{2m_e} - \frac{e^2}{r} \,. \tag{12.1}$$

where we have assumed that the proton in the hydrogen atom is sufficiently more massive that the electron that we can consider only the motion of the electron in the Coulomb field of the proton. The the wave functions for this hamiltonian have a well-known analytic form and the energy eigenvalues of $\hat{H}_0$ are given by

$$E_n^{(0)} = -\frac{\alpha^2 m_e c^2}{2n^2} \,, \tag{12.2}$$

where the principle quantum number n can be related to the highest power appearing in the polynomial part of the wave function n' and the angular momentum quantum number l by

$$n = n' + l + 1 \,. \tag{12.3}$$

For a given value of n, all states with $0 \le l \le n-1$ are degenerate with $2l+1$ values of m for each l. There are two types of degeneracy here. The first is common to any solution of the Schrödinger equation with a central potential. That is for every value of l, there are states with $2l+1$ values of m that are degenerate. This degeneracy is related to the spherical symmetry of the hamiltonian and is manifest in the set of mutually commuting operators $\hat{H}_0$, $\hat{\boldsymbol{L}}^2$ and $\hat{L}_z$. The other type of degeneracy is the degeneracy of the states with $0 \le l \le n-1$ which is the dynamical result of the particular form of the potential.

The hamiltonian $\hat{H}_0$ would provide a correct description of the hydrogen-like atoms if electrons had no spin. Electron spin was proposed by Uhlenbeck and Goudsmit in 1926 to account for the anomalous Zeeman effect. Due to its spin, the electron has a magnetic moment which can interact with the electric field of the nucleus. The classical picture of this situation is that in the rest frame of an electron moving in an electric field, the electron sees a magnetic field given by

$$\boldsymbol{B} = -\frac{\boldsymbol{v} \times \boldsymbol{E}}{c} , \tag{12.4}$$

which results in a potential energy for the interaction of the electron interacting with this magnetic field of

$$V = -\boldsymbol{\mu} \cdot \boldsymbol{B} = \boldsymbol{\mu} \cdot \frac{\boldsymbol{v} \times \boldsymbol{E}}{c} . \tag{12.5}$$

The electric field can be written in terms of the electric scalar potential of the nucleus as

$$\boldsymbol{E} = -\boldsymbol{\nabla}\phi(\boldsymbol{r}) = -\boldsymbol{\nabla}\frac{e}{r} = \frac{e}{r^3}\boldsymbol{r} . \tag{12.6}$$

Using this along with the expression for the electron magnetic moment in terms of its spin yields

$$V = -\frac{eg}{2m_e c}\boldsymbol{S} \cdot \frac{1}{c}\boldsymbol{v} \times \frac{e}{r^3}\boldsymbol{r} = -\frac{e^2 g}{2m_e c^2 r^3}\boldsymbol{S} \cdot \frac{\boldsymbol{p}}{m_e} \times \boldsymbol{r} = \frac{e^2 g}{2m_e^2 c^2 r^3}\boldsymbol{S} \cdot \boldsymbol{r} \times \boldsymbol{p} = \frac{e^2 g}{2m_e^2 c^2}\frac{\boldsymbol{L} \cdot \boldsymbol{S}}{r^3} . \tag{12.7}$$

Because this potential involves the interaction of the electron spin and its orbital angular momentum, it is referred to as the *spin-orbit potential*. This interaction was used in perturbative calculations of the hydrogen spectrum and was found to produce multiplet splittings which were too large by a factor of 2. It was then shown by Thomas that this is related to the fact that in this derivation the electron was treated as if it were in uniform motion. If the the electron is allowed to accelerate, as it will in the presence of the nucleus, a relativistic kinematical effect results in a precession of the electron spin that requires that the above expression be multiplied by a factor of $1/2$. This effect is called the Thomas precession and is a symptom of the relativistic nature of electromagnetism. The correct form of the spin-orbit interaction is then given by

$$\hat{H}_{SO} = \frac{ge^2}{4m_e^2 c^2}\frac{\hat{\boldsymbol{L}} \cdot \hat{\boldsymbol{S}}}{r^3} . \tag{12.8}$$

It should be noted here that this problem can be eliminated by using the "correct" relativistic wave equation for the spin-1/2 electron which is the Dirac equation. Consideration of the equation and its solution is beyond the scope of this course, but we will compare the results of our calculations of the hydrogen spectrum to that obtained from the Dirac equation in the discussion below.

Because of the relativistic nature of the spin-orbit potential, it is necessary to also consider other possible contributions from relativity of similar order. The motivation for a kinematic correction of this order is often motivated by considering the relativistic energy of a particle as

$$E = \sqrt{\boldsymbol{p}^2 c^2 + m^2 c^4} \,. \tag{12.9}$$

The nonrelativistic limit of this expression occurs when the the first term under the radical is small compared to the second. We can therefore construct a Taylor series expansion of the energy as

$$E = mc^2 \sqrt{1 + \frac{\boldsymbol{p}^2}{m^2 c^2}} = mc^2 + \frac{\boldsymbol{p}^2}{2m} - \frac{\boldsymbol{p}^4}{8m^3 c^2} + \ldots \tag{12.10}$$

This then suggests that the leading relativistic correction to the kinetic energy for the electron in hydrogen should be

$$\hat{H}_{rel} = -\frac{\hat{\boldsymbol{p}}^4}{8m_e^3 c^2} \,. \tag{12.11}$$

This is in fact the correct result although the derivation is misleading. The actual origin of this correction is in a complicated nonrelativistic reduction of the Dirac equation.

With these corrections, the hamiltonian of the hydrogen atom can be written as

$$\hat{H} = \hat{H}_0 + \hat{H}_1 \,, \tag{12.12}$$

where

$$\hat{H}_1 = \hat{H}_{SO} + \hat{H}_{rel} \,. \tag{12.13}$$

The situation can be summarized by noting that energy eigenvalues and wave functions for part of the hamiltonian $\hat{H}_0$ can be obtained analytically and that the remaining part of the hamiltonian $\hat{H}_1$ is small.

12.2 Perturbation Theory for Non-Degenerate States

The hydrogen atom is an example of a relatively common situation in quantum mechanics where the hamiltonian for the system we wish to solve is very close to that of

a simple problem that is readily solvable analytically. In such a case we can separate the hamiltonian into two pieces

$$\hat{H} = \hat{H}_0 + \hat{H}_1 \,, \tag{12.14}$$

where $\hat{H}_0$ is the hamiltonian for the solvable problem and $\hat{H}_1$ contains the small corrections.

The eignestates of $\hat{H}_0$ are given by the Schrödinger equation

$$\hat{H}_0 |\phi_\alpha\rangle = E_\alpha^{(0)} |\phi_\alpha\rangle \,, \tag{12.15}$$

while the Schrödinger equation for the full hamiltonian is

$$\left(\hat{H}_0 + \hat{H}_1 \right) |\psi_\alpha\rangle = E_\alpha |\psi_\alpha\rangle \,. \tag{12.16}$$

Here we will use Greek letters to represent sets of eigenvalues corresponding to a complete set of mutually commuting operators necessary to specify the eigenstates of the hamiltonian $\hat{H}_0$ and Roman letters to represent other summation indices. Since the eigenstates of $\hat{H}_0$ form a complete set of states, one solution to the problem would be to reduce this problem to matrix form by expanding in theses states. This gives

$$\langle\phi_\beta| \left(\hat{H}_0 + \hat{H}_1 \right) \left(\sum_\gamma |\phi_\gamma\rangle \langle\phi_\gamma| \right) |\psi_\alpha\rangle = E_\alpha \langle\phi_\beta| \psi_\alpha\rangle \tag{12.17}$$

or

$$\sum_\gamma \langle\phi_\beta| \left(\hat{H}_0 + \hat{H}_1 \right) |\phi_\gamma\rangle \langle\phi_\gamma| \psi_\alpha\rangle = E_\alpha \langle\phi_\beta| \psi_\alpha\rangle \,. \tag{12.18}$$

Using (12.15) this can be reduced to

$$\sum_\gamma \left(E_\beta^{(0)} \delta_{\beta\gamma} + \langle\phi_\beta| \hat{H}_1 |\phi_\gamma\rangle \right) \langle\phi_\gamma| \psi_\alpha\rangle = E_\alpha \langle\phi_\beta| \psi_\alpha\rangle \,. \tag{12.19}$$

This can be viewed as representing the components of a matrix equation which can be solved for the eigenvalues E_n and the coefficients $\langle\phi_k| \psi_n\rangle$ which determine the expansion of the state

$$|\psi_\alpha\rangle = \sum_\gamma |\phi_\gamma\rangle \langle\phi_\gamma| \psi_\alpha\rangle \,. \tag{12.20}$$

In general, this represents an infinite set of equations. However, as we have discussed previously, this can be approximated in most cases by simply keeping a large but finite number of states of $\hat{H}_0$. This can be done regardless of the size of $\hat{H}_1$.

If all of the matrix elements $\hat{H}_1$ are small in comparison to the energies $E_\alpha^{(0)}$, then it is usually easier to approximate the result by using perturbation theory. To derive the results of time-independent perturbation theory it is convenient to make the

replacement $\hat{H}_1 \to \lambda \hat{H}_1$ where λ will be used to characterized the size of corrections and will ultimately be replaced by $\lambda = 1$ in any actual calculations.

Since we expect that the actual eigenvalues will be close to $E_n^{(0)}$, we can assume that the eigenvalues can be represented as

$$E_\alpha = E_\alpha^{(0)} + \lambda E_\alpha^{(1)} + \lambda^2 E_\alpha^{(2)} + \ldots = \sum_{i=0}^{\infty} \lambda^i E_\alpha^{(i)} \,. \tag{12.21}$$

This guaranties that in the limit $\lambda \to 0$, the unperturbed result is recovered. Similarly, we expect that the eigenfunction will be close to that of $|\phi_\alpha\rangle$ with all corrections involving the admixture of the other eigenstates of $\hat{H}_0$ so that we can express it as

$$|\psi_\alpha\rangle = N \left(|\phi_\alpha\rangle + \sum_{\gamma \neq \alpha} \sum_{j=1}^{\infty} \lambda^j C_{\alpha\gamma}^{(j)} |\phi_\gamma\rangle \right) , \tag{12.22}$$

where N is a normalization constant that must be recalculated for each state at any given level of approximation in the perturbation series. Note that the sum over γ is restricted such that $\gamma \neq \alpha$. Including the term for $\gamma = \alpha$ produces a contribution proportional to the zeroth order contribution and can be absorbed into it by simply redefining the normalization constant. It is, therefore, redundant. It is convenient to rewrite (12.22) as

$$|\psi_\alpha\rangle = N \left(|\psi_\alpha^{(0)}\rangle + \sum_{j=1}^{\infty} \lambda^j |\psi_\alpha^{(j)}\rangle \right) , \tag{12.23}$$

where

$$|\psi_\alpha^{(0)}\rangle = |\phi_\alpha\rangle \tag{12.24}$$

and

$$|\psi_\alpha^{(j)}\rangle = \sum_{\gamma \neq \alpha} C_{\alpha\gamma}^{(j)} |\phi_\gamma\rangle \,. \tag{12.25}$$

Equation (12.22) implies that

$$\langle\phi_\beta| \psi_\alpha\rangle = N \left(\delta_{\beta\alpha} + \sum_{\gamma \neq \alpha} \sum_{j=1}^{\infty} \lambda^j C_{\alpha\gamma}^{(j)} \delta_{\beta\gamma} \right) = N \left(\delta_{\beta\alpha} + \sum_{j=1}^{\infty} \lambda^j C_{\alpha\beta}^{(j)} (1 - \delta_{\beta\alpha}) \right) \,. \tag{12.26}$$

Now rewriting (12.19) as

$$\sum_\gamma \langle\phi_\beta| \lambda \hat{H}_1 |\phi_\gamma\rangle \langle\phi_\gamma| \psi_\alpha\rangle = \left(E_\alpha - E_\beta^{(0)} \right) \langle\phi_\beta| \psi_\alpha\rangle \,, \tag{12.27}$$

199

and using (12.21) and (12.26), we can write

$$\sum_{\gamma} \langle \phi_{\beta} | \, \lambda \hat{H}_1 \, | \phi_{\gamma} \rangle \, N \left(\delta_{\gamma\alpha} + \sum_{j=1}^{\infty} \lambda^j C_{\alpha\gamma}^{(j)} \left(1 - \delta_{\gamma\alpha} \right) \right)$$

$$= \left(\sum_{i=0}^{\infty} \lambda^i E_{\alpha}^{(i)} - E_{\beta}^{(0)} \right) N \left(\delta_{\beta\alpha} + \sum_{j=1}^{\infty} \lambda^j C_{\alpha\beta}^{(j)} \left(1 - \delta_{\beta\alpha} \right) \right) ,$$

$$= \left(E_{\alpha}^{(0)} - E_{\beta}^{(0)} + \sum_{i=1}^{\infty} \lambda^i E_{\alpha}^{(i)} \right) N \left(\delta_{\beta\alpha} + \sum_{j=1}^{\infty} \lambda^j C_{\alpha\beta}^{(j)} \left(1 - \delta_{\beta\alpha} \right) \right) ,$$

$$\tag{12.28}$$

Canceling the normalization constant from both sides and expanding yields

$$\lambda \langle \phi_{\beta} | \, \hat{H}_1 \, | \phi_{\alpha} \rangle + \sum_{j=1}^{\infty} \lambda^{j+1} \sum_{\gamma} \langle \phi_{\beta} | \, \hat{H}_1 \, | \phi_{\gamma} \rangle \, C_{\alpha\gamma}^{(j)} \left(1 - \delta_{\gamma\alpha} \right)$$

$$= \left(E_{\alpha}^{(0)} - E_{\beta}^{(0)} \right) \sum_{j=1}^{\infty} \lambda^j C_{\alpha\beta}^{(j)} \left(1 - \delta_{\beta\alpha} \right) + \sum_{i=1}^{\infty} \lambda^i E_{\alpha}^{(i)} \delta_{\beta\alpha}$$

$$+ \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \lambda^{i+j} E_{\alpha}^{(i)} C_{\alpha\beta}^{(j)} \left(1 - \delta_{\beta\alpha} \right) . \tag{12.29}$$

This can be rewritten as

$$\lambda \langle \phi_{\beta} | \, \hat{H}_1 \, | \phi_{\alpha} \rangle + \sum_{k=2}^{\infty} \lambda^k \sum_{\gamma} \langle \phi_{\beta} | \, \hat{H}_1 \, | \phi_{\gamma} \rangle \, C_{\alpha\gamma}^{(k-1)} \left(1 - \delta_{\gamma\alpha} \right)$$

$$= \left(E_{\alpha}^{(0)} - E_{\beta}^{(0)} \right) \sum_{k=1}^{\infty} \lambda^k C_{\alpha\beta}^{(k)} \left(1 - \delta_{\beta\alpha} \right) + \sum_{k=1}^{\infty} \lambda^k E_{\alpha}^{(k)} \delta_{\beta\alpha}$$

$$+ \sum_{k=2}^{\infty} \sum_{i=1}^{k-1} \lambda^k E_{\alpha}^{(i)} C_{\alpha\beta}^{(k-i)} \left(1 - \delta_{\beta\alpha} \right) . \tag{12.30}$$

Since the two sides of this expression are polynomials in λ, we can equate the coefficients of each power separately. Note that the coefficient of the term of order λ^0 is zero. Equating the coefficients of λ yields

$$\langle \phi_{\beta} | \, \hat{H}_1 \, | \phi_{\alpha} \rangle = \left(E_{\alpha}^{(0)} - E_{\beta}^{(0)} \right) C_{\alpha\beta}^{(1)} \left(1 - \delta_{\beta\alpha} \right) + E_{\alpha}^{(1)} \delta_{\beta\alpha} . \tag{12.31}$$

For the case where $\beta = \alpha$ we get

$$E_{\alpha}^{(1)} = \langle \phi_{\alpha} | \, \hat{H}_1 \, | \phi_{\alpha} \rangle , \tag{12.32}$$

which gives the first order energy shift due to the presence of $\hat{H}_1$. For the case where $\beta \neq \alpha$, we get

$$\langle\phi_\beta| \hat{H}_1 |\phi_\alpha\rangle = \left(E_\alpha^{(0)} - E_\beta^{(0)}\right) C_{\alpha\beta}^{(1)}, \tag{12.33}$$

which can be solved for the first order correction to the wave function expansion coefficients

$$C_{\alpha\beta}^{(1)} = \frac{\langle\phi_\beta| \hat{H}_1 |\phi_\alpha\rangle}{E_\alpha^{(0)} - E_\beta^{(0)}}. \tag{12.34}$$

Using this,

$$\left|\psi_\alpha^{(1)}\right\rangle = \sum_{\gamma\neq\alpha} C_{\alpha\gamma}^{(1)} |\phi_\gamma\rangle = \sum_{\gamma\neq\alpha} \frac{\langle\phi_\gamma| \hat{H}_1 |\phi_\alpha\rangle}{E_\alpha^{(0)} - E_\gamma^{(0)}} |\phi_\gamma\rangle \tag{12.35}$$

For $k \geq 2$,

$$\sum_\gamma \langle\phi_\beta| \hat{H}_1 |\phi_\gamma\rangle C_{\alpha\gamma}^{(k-1)} (1 - \delta_{\gamma\alpha}) = \left(E_\alpha^{(0)} - E_\beta^{(0)}\right) C_{\alpha\beta}^{(k)} (1 - \delta_{\beta\alpha}) + E_\alpha^{(k)}\delta_{\beta\alpha}$$

$$+ \sum_{i=1}^{k-1} E_\alpha^{(i)} C_{\alpha\beta}^{(k-i)} (1 - \delta_{\beta\alpha}). \tag{12.36}$$

For $\alpha = \beta$, this gives

$$E_\alpha^{(k)} = \sum_\gamma \langle\phi_\alpha| \hat{H}_1 |\phi_\gamma\rangle C_{\alpha\gamma}^{(k-1)} (1 - \delta_{\gamma\alpha}). \tag{12.37}$$

For $\alpha \neq \beta$,

$$\sum_\gamma \langle\phi_\beta| \hat{H}_1 |\phi_\gamma\rangle C_{\alpha\gamma}^{(k-1)} (1 - \delta_{\gamma\alpha}) = \left(E_\alpha^{(0)} - E_\beta^{(0)}\right) C_{\alpha\beta}^{(k)} + \sum_{i=1}^{k-1} E_\alpha^{(i)} C_{\alpha\beta}^{(k-i)}. \tag{12.38}$$

This can be solved to give

$$C_{\alpha\beta}^{(k)} = \frac{1}{E_\alpha^{(0)} - E_\beta^{(0)}} \left(\sum_\gamma \langle\phi_\beta| \hat{H}_1 |\phi_\gamma\rangle C_{\alpha\gamma}^{(k-1)} (1 - \delta_{\gamma\alpha}) - \sum_{i=1}^{k-1} E_\alpha^{(i)} C_{\alpha\beta}^{(k-i)}\right). \tag{12.39}$$

We now have a recursive method for obtaining perturbative corrections to the energies and wave functions to any order.

Note, however that the expression for $C_{\alpha\beta}^{(1)}$ contains the factor $(E_\alpha^{(0)} - E_\beta^{(0)})^{-1}$. This will be singular if $E_\alpha^{(0)} = E_\beta^{(0)}$ for $\alpha \neq \beta$ and will cause this coefficient to become infinite unless the matrix element $\langle\phi_\beta| \hat{H}_1 |\phi_\alpha\rangle$ also vanishes. This means that, in general, this approach to obtaining perturbative corrections fails if there are degenerate states. As a result, perturbation theory for systems with degenerate states has to be treated very carefully by using *degenerate-state perturbation theory* which we

will consider shortly. Unfortunately, the states of the simple Coulomb hamiltonian are degenerate and we must develop the non-degenerate state perturbation theory before tackling this problem in general. However, the first order correction to the energy can always be calculated since it does not contain the problematic energy denominator.

For non-degenerate states, the appearance of this denominator has the consequence that if the matrix elements of $\hat{H}_1$ are of roughly comparable size, the perturbative corrections will be dominated by contributions from states close in energy to $|\phi_n\rangle$.

Before considering considering degenerate-state perturbation theory, we will consider the case of a non-degenerate system and a few examples of first order energy corrections for the hydrogen atom.

12.2.1 One-Dimensional Harmonic Oscillator in an Electric Field

As a simple example of the application of perturbation theory consider a simple harmonic oscillator with the hamiltonian

$$\hat{H}_0 = \frac{\hat{p}^2}{2m} + \frac{1}{2}m\omega^2\hat{x}^2 \,. \tag{12.40}$$

If this is now placed in a constant electric field $\mathcal{E}$ lying along the x-axis, an additional term

$$\hat{H}_1 = q\mathcal{E}x \tag{12.41}$$

is added to the hamiltonian.

This calculation can be simplified by introducing the raising and lowering operators

$$\hat{a}^\dagger = \sqrt{\frac{m\omega}{2\hbar}}\hat{x} + i\frac{\hat{p}}{\sqrt{2m\hbar\omega}} \tag{12.42}$$

and

$$\hat{a} = \sqrt{\frac{m\omega}{2\hbar}}\hat{x} - i\frac{\hat{p}}{\sqrt{2m\hbar\omega}} \,. \tag{12.43}$$

We can calculate the commutation relation for these operators using the canonical commutation relations as

$$\begin{aligned}
[\hat{a}, \hat{a}^\dagger] &= \left[\sqrt{\frac{m\omega}{2\hbar}}\hat{x}, i\frac{\hat{p}}{\sqrt{2m\hbar\omega}}\right] + \left[-i\frac{\hat{p}}{\sqrt{2m\hbar\omega}}, \sqrt{\frac{m\omega}{2\hbar}}\hat{x}\right] \\
&= \frac{i}{2\hbar}[\hat{x}, \hat{p}] - \frac{i}{2\hbar}[\hat{p}, \hat{x}] = \frac{i}{\hbar}[\hat{x}, \hat{p}] = 1 \,.
\end{aligned} \tag{12.44}$$

Similarly,

$$[\hat{a}, \hat{a}] = [\hat{a}, \hat{a}] = 0 \,. \tag{12.45}$$

In order to represent the harmonic oscillator hamiltonian in terms of the raising and lowering operators, it is useful to invert (12.43) and (12.42) to give

$$\hat{x} = \sqrt{\frac{\hbar}{2m\omega}}(\hat{a} + \hat{a}^\dagger) \tag{12.46}$$

and

$$\hat{p} = i\sqrt{\frac{m\hbar\omega}{2}}(\hat{a} - \hat{a}^\dagger). \tag{12.47}$$

The harmonic oscillator hamiltonian can now be reexpressed as

$$
\begin{aligned}
\hat{H}_0 &= \frac{\hat{p}^2}{2m} + \frac{1}{2}m\omega^2\hat{x}^2 = -\frac{1}{2m}\frac{m\hbar\omega}{2}(\hat{a} - \hat{a}^\dagger)^2 + \frac{1}{2}m\omega^2\frac{\hbar}{2m\omega}(\hat{a} + \hat{a}^\dagger)^2 \\
&= \frac{\hbar\omega}{4}\left[(\hat{a} + \hat{a}^\dagger)^2 - (\hat{a} - \hat{a}^\dagger)^2\right] \\
&= \frac{\hbar\omega}{4}\left(\hat{a}\hat{a} + \hat{a}\hat{a}^\dagger + \hat{a}^\dagger\hat{a} + \hat{a}^\dagger\hat{a}^\dagger - \hat{a}\hat{a} + \hat{a}\hat{a}^\dagger + \hat{a}^\dagger\hat{a} - \hat{a}^\dagger\hat{a}^\dagger\right) \\
&= \frac{\hbar\omega}{2}\left(\hat{a}\hat{a}^\dagger + \hat{a}^\dagger\hat{a}\right) = \frac{\hbar\omega}{2}\left(\hat{a}^\dagger\hat{a} + 1 + \hat{a}^\dagger\hat{a}\right) = \hbar\omega\left(\hat{a}^\dagger\hat{a} + \tfrac{1}{2}\right),
\end{aligned}
\tag{12.48}
$$

where we have used the commutation relation (12.44) in the next to last step. Applying this to the eigenstate for the the nth excitation of the oscillator gives

$$\hat{H}_0 \ket{n} = \hbar\omega\left(\hat{a}^\dagger\hat{a} + \tfrac{1}{2}\right)\ket{n}. \tag{12.49}$$

From our previous study of the one-dimensional harmonic oscillator, we know that

$$\hat{H}_0 \ket{n} = \hbar\omega\left(n + \tfrac{1}{2}\right)\ket{n}. \tag{12.50}$$

Comparing these two expressions we can identify the operator

$$\hat{n} = \hat{a}^\dagger\hat{a} \tag{12.51}$$

such that

$$\hat{n}\ket{n} = n\ket{n}. \tag{12.52}$$

This is called the number operator. We can then rewrite the hamiltonian as

$$\hat{H} = \hbar\omega\left(\hat{n} + \frac{1}{2}\right). \tag{12.53}$$

The commutators

$$[\hat{a}^\dagger, \hat{H}_0] = -\hbar\omega\hat{a}^\dagger \tag{12.54}$$

and

$$[\hat{a}, \hat{H}_0] = \hbar\omega\hat{a} \tag{12.55}$$

can be calculated using the commutation relations for the creation and annihilation operators. Applying the first of these to an eigenstate of $\hat{H}_0$ gives

$$[\hat{a}^\dagger, \hat{H}_0]\,|n\rangle = \hat{a}^\dagger \hat{H}_0\,|n\rangle - \hat{H}_0 \hat{a}^\dagger\,|n\rangle = \hbar\omega\left(n + \tfrac{1}{2}\right)\hat{a}^\dagger\,|n\rangle - \hat{H}_0 \hat{a}^\dagger\,|n\rangle = -\hbar\omega\hat{a}^\dagger\,. \quad (12.56)$$

This can be arranged as

$$\hat{H}_0(\hat{a}^\dagger\,|n\rangle) = \hbar\omega\left(n + 1 + \tfrac{1}{2}\right)(\hat{a}^\dagger\,|n\rangle)\,. \quad (12.57)$$

So the state $\hat{a}^\dagger\,|n\rangle$ has the eigenvalue of the state $|n+1\rangle$. Therefore,

$$\hat{a}^\dagger\,|n\rangle = c_n\,|n+1\rangle \quad (12.58)$$

where c_n is a constant. Using the second commutator, a similar argument leads to

$$\hat{a}\,|n\rangle = d_n\,|n-1\rangle\,. \quad (12.59)$$

To obtain the values of the constants c_n and d_n consider

$$n = \langle n|\,\hat{n}\,|n\rangle = \langle n|\,\hat{a}^\dagger \hat{a}\,|n\rangle = (\hat{a}\,|n\rangle)^\dagger \hat{a}\,|n\rangle = (d_n\,|n\rangle)^\dagger d_n\,|n\rangle = |d_n|^2\,. \quad (12.60)$$

This requires that

$$|d_n|^2 = n\,. \quad (12.61)$$

So,

$$d_n = e^{i\varphi}\sqrt{n} \quad (12.62)$$

where φ is an arbitrary phase which we will choose to be $\varphi = 0$. Next, consider

$$n\,|n\rangle = \hat{a}^\dagger \hat{a}\,|n\rangle = \hat{a}^\dagger d_n\,|n-1\rangle = c_{n-1}d_n\,|n\rangle = \quad (12.63)$$

which requires that

$$c_{n-1}d_n = n\,. \quad (12.64)$$

Solving this for

$$c_{n-1} = \frac{n}{d_n} = \frac{n}{\sqrt{n}} = \sqrt{n} \quad (12.65)$$

gives

$$c_n = \sqrt{n+1}\,. \quad (12.66)$$

Therefore,

$$\hat{a}\,|n\rangle = \sqrt{n}\,|n-1\rangle \quad (12.67)$$

and

$$\hat{a}^\dagger\,|n\rangle = \sqrt{n+1}\,|n+1\rangle\,,\,. \quad (12.68)$$

We can now consider perturbative corrections to the energy and wave functions of the unperturbed hamiltonian to first and second order. Using (12.46), we can write

$$
\begin{aligned}
\langle m | \hat{H}_1 | n \rangle &= q\mathcal{E}\sqrt{\frac{\hbar}{2m\omega}} \langle m | \left(\hat{a} + \hat{a}^\dagger \right) | n \rangle \\
&= q\mathcal{E}\sqrt{\frac{\hbar}{2m\omega}} \left(\sqrt{n} \langle m | n-1 \rangle + \sqrt{n+1} \langle m | n+1 \rangle \right) \\
&= q\mathcal{E}\sqrt{\frac{\hbar}{2m\omega}} \left(\sqrt{n}\, \delta_{m,n-1} + \sqrt{n+1}\, \delta_{m,n+1} \right) .
\end{aligned}
\tag{12.69}
$$

The first order energy shift is

$$
E_n^{(1)} = \langle n | \hat{H}_1 | n \rangle = 0
\tag{12.70}
$$

since the matrix elements of $\hat{H}_1$ have no diagonal terms.

The first order expansion coefficients are given by

$$
\begin{aligned}
C_{nm}^{(1)} &= \frac{\langle m | \hat{H}_1 | n \rangle}{E_n^{(0)} - E_m^{(0)}} \\
&= \frac{1}{\hbar\omega \left(n + \frac{1}{2} \right) - \hbar\omega \left(m + \frac{1}{2} \right)} q\mathcal{E}\sqrt{\frac{\hbar}{2m\omega}} \left(\sqrt{n}\, \delta_{m,n-1} + \sqrt{n+1}\, \delta_{m,n+1} \right) \\
&= \frac{1}{\hbar\omega} q\mathcal{E}\sqrt{\frac{\hbar}{2m\omega}} \left(\frac{\sqrt{n}}{1}\, \delta_{m,n-1} + \frac{\sqrt{n+1}}{-1}\, \delta_{m,n+1} \right) \\
&= \frac{1}{\hbar\omega} q\mathcal{E}\sqrt{\frac{\hbar}{2m\omega}} \left(\sqrt{n}\, \delta_{m,n-1} - \sqrt{n+1}\, \delta_{m,n+1} \right) .
\end{aligned}
\tag{12.71}
$$

The second order energy shift is given by

$$
\begin{aligned}
E_n^{(2)} &= \sum_{k \neq n} \frac{\langle n | \hat{H}_1 | k \rangle \langle k | \hat{H}_1 | n \rangle}{E_n^{(0)} - E_k^{(0)}} = \sum_{k \neq n} \frac{\left| \langle k | \hat{H}_1 | n \rangle \right|^2}{E_n^{(0)} - E_k^{(0)}} \\
&= \sum_{k \neq n} \frac{\left| q\mathcal{E}\sqrt{\frac{\hbar}{2m\omega}} \left(\sqrt{n}\, \delta_{k,n-1} + \sqrt{n+1}\, \delta_{k,n+1} \right) \right|^2}{\hbar\omega \left(n + \frac{1}{2} \right) - \hbar\omega \left(k + \frac{1}{2} \right)} \\
&= \frac{\mathcal{E}^2 q^2}{\hbar\omega} \frac{\hbar}{2m\omega} \sum_{k \neq n} \frac{n\, \delta_{k,n-1} + (n+1)\, \delta_{k,n+1}}{n - k} \\
&= \frac{\mathcal{E}^2 q^2}{\hbar\omega} \frac{\hbar}{2m\omega} \left(n - (n+1) \right) = -\frac{\mathcal{E}^2 q^2}{2m\omega^2} .
\end{aligned}
\tag{12.72}
$$

The second order correction to the expansion coefficients is

$$
\begin{aligned}
C_{nm}^{(2)} &= \frac{1}{E_n^{(0)} - E_m^{(0)}} \left[\sum_{k \neq n} \langle m| \hat{H}_1 |k\rangle C_{nk}^{(1)} - E_n^{(1)} C_{nm}^{(1)} \right] \\
&= \frac{1}{E_n^{(0)} - E_m^{(0)}} \sum_{k \neq n} \frac{\langle m| \hat{H}_1 |k\rangle \langle k| \hat{H}_1 |n\rangle}{E_n^{(0)} - E_k^{(0)}} \\
&= \frac{q^2 \mathcal{E}^2}{\hbar^2 \omega^2} \frac{\hbar}{2m\omega} \frac{1}{n-m} \sum_{k \neq n} \frac{1}{n-k} \left(\sqrt{k}\, \delta_{m,k-1} + \sqrt{k+1}\, \delta_{m,k+1} \right) \\
&\qquad \times \left(\sqrt{n}\, \delta_{k,n-1} + \sqrt{n+1}\, \delta_{k,n+1} \right) \\
&= \frac{q^2 \mathcal{E}^2}{\hbar 2m\omega^3} \frac{1}{n-m} \Big[\left(\sqrt{n-1}\sqrt{n}\, \delta_{m,n-2} + n\delta_{mn} \right) \\
&\qquad - \left((n+1)\, \delta_{m,n} + \sqrt{n+2}\sqrt{n+1}\delta_{m,n+2} \right) \Big] . \qquad (12.73)
\end{aligned}
$$

At this point it might appear that there is some trouble since the terms where $m = n$ would be singular. However it should be remembered that the sums over the expansion coefficients explicitly exclude these terms so we can write

$$
C_{nm}^{(2)} = \frac{q^2 \mathcal{E}^2}{\hbar 4 m\omega^3} \left(\sqrt{n-1}\sqrt{n}\, \delta_{m,n-2} - \sqrt{n+2}\sqrt{n+1}\delta_{m,n+2} \right) \qquad (12.74)
$$

Because of the characteristics of the oscillator potential it is possible to obtain an exact solution to this problem. The potential for this problem is

$$
V(\hat{x}) = \frac{1}{2}m\omega^2 \hat{x}^2 + q\mathcal{E}\hat{x} = \frac{1}{2}m\omega^2 \left(\hat{x}^2 + \frac{2q\mathcal{E}}{m\omega^2}\hat{x} \right) . \qquad (12.75)
$$

The square can be completed by adding and subtracting a term. That is,

$$
V(\hat{x}) = \frac{1}{2}m\omega^2 \left(\hat{x}^2 + \frac{2q\mathcal{E}}{m\omega^2}\hat{x} + \frac{q^2\mathcal{E}^2}{m^2\omega^4} \right) - \frac{q^2\mathcal{E}^2}{2m\omega^2} = \frac{1}{2}m\omega^2 \left(\hat{x} + \frac{q\mathcal{E}}{m\omega^2} \right)^2 - \frac{q^2\mathcal{E}^2}{2m\omega^2}
$$

$$
(12.76)
$$

This corresponds to an oscillator potential with a shifted origin plus a shift in magnitude. This results in an energy shift equal to the second order result.

12.3   Degenerate State Perturbation Theory

We will now deal with the problem of perturbation theory for systems with degenerate states. The problem that we saw in our original derivation of the perturbation series was that the denominators for terms mixing degenerate states vanish resulting in infinite contributions. These terms are associate with the mixing of the degenerate

states by the perturbative contribution to the hamiltonian. In addition we have assumed that the eigenstates of the unperturbed hamiltonian differ from those of the perturbed system by only a small amount that would vanish as the size of the perturbation goes to zero. That this assumption is false can be seen by considering the solution of a simple two level system with degenerate states for the uncoupled system. The matrix representation of the hamiltonian of this system is given by

$$H = \begin{pmatrix} E_0 & V_{12} \\ V_{12} & E_0 \end{pmatrix}, \tag{12.77}$$

where the matrix element V_{12} is assumed to be real. The eigenvalues for this hamiltonian are

$$E_\pm = E_0 \pm V_{12} \tag{12.78}$$

and the corresponding eigenvalues are

$$\psi_\pm = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ \pm 1 \end{pmatrix}. \tag{12.79}$$

Note that the eigenvectors are independent of the size of the interaction potential V_{12} and do not approach the "unperturbed" eigenvectors.

Before tackling the construction of the perturbation expansion for degenerate states, it is necessary to introduce the concept of the projection operator. For a given state $|\phi_\alpha\rangle$ we define the projection operator

$$\hat{P}_\alpha = |\phi_\alpha\rangle \langle \phi_\alpha|. \tag{12.80}$$

The product of the projection operator with itself is

$$\hat{P}_\alpha \hat{P}_\alpha = |\phi_\alpha\rangle \langle \phi_\alpha | \phi_\alpha\rangle \langle \phi_\alpha| = |\phi_\alpha\rangle \langle \phi_\alpha| = \hat{P}_\alpha. \tag{12.81}$$

The property that the square of an operator is equal to the operator itself is called *idempotency*.

Now consider some arbitrary state in the Hilbert space

$$|\psi\rangle = \sum_\beta a_\beta |\phi_\beta\rangle. \tag{12.82}$$

Acting on this state with the projection operator for state α yields

$$\hat{P}_\alpha |\psi\rangle = \sum_\beta a_\beta \hat{P}_\alpha |\phi_\beta\rangle = \sum_\beta a_\beta |\phi_\alpha\rangle \langle \phi_\alpha | \phi_\beta\rangle = \sum_\beta a_\beta \delta_{\alpha\beta} |\phi_\alpha\rangle = a_\alpha |\phi_\alpha\rangle. \tag{12.83}$$

That is $\hat{P}_\alpha$ projects out the contribution of the state α to the complete state. Note that the completeness relation for the Hilbert space can be rewritten as.

$$\sum_\alpha \hat{P}_\alpha = \hat{1} \tag{12.84}$$

Now consider the case of a system with degenerate states. We can characterize the energy of each state with the energy quantum number n. If there are degenerate states for a given energy, then we can label this set of states as $\{n\}$ and can define a projection operator onto the set of degenerate states as

$$\hat{P}_n = \sum_{\alpha \in \{n\}} \hat{P}_\alpha \,. \tag{12.85}$$

The complement to this subspace is given by

$$\hat{Q}_n = \sum_{\alpha \ni \{n\}} \hat{P}_\alpha = \hat{1} - \hat{P}_n \,. \tag{12.86}$$

From this definition, it is clear that

$$\hat{P}_n \hat{Q}_n = \hat{Q}_n \hat{P}_n = 0 \,. \tag{12.87}$$

We can now use these operators to project the Schrödinger equation onto these two subspaces. First, using $\hat{P}_n + \hat{Q}_n = \hat{1}$,

$$\left(\hat{H}_0 + \hat{H}_1 \right) \left(\hat{P}_n + \hat{Q}_n \right) |\psi_\alpha\rangle = E_\alpha |\psi_\alpha\rangle \,. \tag{12.88}$$

Multiplying this from the left by $\hat{P}_n$ gives

$$\left[\hat{P}_n \left(\hat{H}_0 + \hat{H}_1 \right) \hat{P}_n + \hat{P}_n \left(\hat{H}_0 + \hat{H}_1 \right) \hat{Q}_n \right] |\psi_\alpha\rangle = E_\alpha \hat{P}_n |\psi_\alpha\rangle \,. \tag{12.89}$$

Since $\hat{H}_0$ is diagonal in basis, $\hat{P}_n \hat{H}_0 \hat{Q}_n = \hat{Q}_n \hat{H}_0 \hat{P}_n = 0$. So,

$$\left[\hat{P}_n \hat{H}_0 \hat{P}_n + \hat{P}_n \hat{H}_1 \hat{P}_n + \hat{P}_n \hat{H}_1 \hat{Q}_n \right] |\psi_\alpha\rangle = E_\alpha \hat{P}_n |\psi_\alpha\rangle \,. \tag{12.90}$$

Similarly, by multiplying the Schrödinger equation from the left by $\hat{Q}_n$,

$$\left[\hat{Q}_n \hat{H}_0 \hat{Q}_n + \hat{Q}_n \hat{H}_1 \hat{Q}_n + \hat{Q}_n \hat{H}_1 \hat{P}_n \right] |\psi_\alpha\rangle = E_\alpha \hat{Q}_n |\psi_\alpha\rangle \,. \tag{12.91}$$

Assume that we are interested in finding the perturbative corrections to some state α where $\alpha \in \{n\}$. The problem that we had with our previous derivation was associated with the situation where degenerate states were mixed by the perturbation $\hat{H}_1$. This is associated in (12.90) with the term $\hat{P}_n \hat{H}_1 \hat{P}_n$. We can eliminate this term by redefining the hamiltonian such that

$$\hat{\bar{H}}_0 = \hat{H}_0 + \hat{P}_n \hat{H}_1 \hat{P}_n \tag{12.92}$$

and

$$\hat{\bar{H}}_1 = \hat{H}_1 - \hat{P}_n \hat{H}_1 \hat{P}_n \,. \tag{12.93}$$

We now construct an new expansion basis by requiring that

$$\hat{\bar{H}}_0 \left| \bar{\phi}_\alpha \right\rangle = \bar{E}_\alpha^{(0)} \left| \bar{\phi}_\alpha \right\rangle . \tag{12.94}$$

Note that since $\hat{\bar{H}}_0$ differs from $\hat{H}_0$ only in the subspace defined by $\{n\}$, this amounts to simply diagonalizing the full hamiltonian in this finite dimensional subspace. The states outside of this subspace are unchanged by the redefinition of the parts of the hamiltonian. Since the two sets of states $\left| \bar{\phi}_\alpha \right\rangle$ and $\left| \phi_\alpha \right\rangle$ with $\alpha \in \{n\}$ span the same subspace,

$$\hat{\bar{P}}_n = \sum_{\alpha \in \{n\}} \left| \bar{\phi}_\alpha \right\rangle \left\langle \bar{\phi}_\alpha \right| = \sum_{\alpha \in \{n\}} \left| \phi_\alpha \right\rangle \left\langle \phi_\alpha \right| = \hat{P}_n \tag{12.95}$$

and

$$\hat{\bar{Q}}_n = \hat{Q}_n . \tag{12.96}$$

Note that

$$\hat{P}_n \hat{\bar{H}}_1 \hat{Q}_n = \hat{P}_n \hat{H}_1 \hat{Q}_n \tag{12.97}$$

$$\hat{Q}_n \hat{\bar{H}}_1 \hat{P}_n = \hat{Q}_n \hat{H}_1 \hat{P}_n \tag{12.98}$$

$$\hat{Q}_n \hat{\bar{H}}_1 \hat{Q}_n = \hat{Q}_n \hat{H}_1 \hat{Q}_n \tag{12.99}$$

$$\hat{Q}_n \hat{\bar{H}}_0 \hat{Q}_n = \hat{Q}_n \hat{H}_0 \hat{Q}_n . \tag{12.100}$$

We can now rewrite (12.90) and (12.91) as

$$\left[\hat{P}_n \hat{\bar{H}}_0 \hat{P}_n + \lambda \hat{P}_n \hat{H}_1 \hat{Q}_n \right] \left| \psi_\alpha \right\rangle = E_\alpha \hat{P}_n \left| \psi_\alpha \right\rangle . \tag{12.101}$$

and

$$\left[\hat{Q}_n \hat{H}_0 \hat{Q}_n + \lambda \hat{Q}_n \hat{H}_1 \hat{Q}_n + \lambda \hat{Q}_n \hat{H}_1 \hat{P}_n \right] \left| \psi_\alpha \right\rangle = E_\alpha \hat{Q}_n \left| \psi_\alpha \right\rangle , \tag{12.102}$$

where we have made the substitutions $\hat{H}_1 \to \lambda \hat{H}_1$ for the purpose of constructing a perturbative expansion.

To construct the perturbation series, we once again expand the energy and eigenstate in powers of λ as

$$E_\alpha = \bar{E}_\alpha^{(0)} + \sum_{i=1}^{\infty} \lambda^i \bar{E}_\alpha^{(i)} \tag{12.103}$$

and

$$\left| \psi_\alpha \right\rangle = N \left(\left| \bar{\phi}_\alpha \right\rangle + \sum_\gamma \sum_{j=1}^{\infty} \lambda^j C_{\alpha\gamma}^{(j)} \left| \bar{\phi}_\gamma \right\rangle (1 - \delta_{\alpha\gamma}) \right) . \tag{12.104}$$

209

The projections of the eigenstate onto the subspaces give

$$
\begin{aligned}
\hat{P}_n \left| \psi_\alpha \right\rangle &= \sum_{\kappa \in \{n\}} \left| \bar{\phi}_\kappa \right\rangle \left\langle \bar{\phi}_\kappa \left| \psi_\alpha \right\rangle \right. \\
&= \sum_{\kappa \in \{n\}} \left| \bar{\phi}_\kappa \right\rangle N \left(\left\langle \bar{\phi}_\kappa \left| \bar{\phi}_\alpha \right\rangle \right. + \sum_\gamma \sum_{j=1}^\infty \lambda^j C_{\alpha\gamma}^{(j)} \left\langle \bar{\phi}_\kappa \left| \bar{\phi}_\gamma \right\rangle \right. (1 - \delta_{\alpha\gamma}) \right) \\
&= N \sum_{\kappa \in \{n\}} \left| \bar{\phi}_\kappa \right\rangle \left(\delta_{\kappa\alpha} + \sum_\gamma \sum_{j=1}^\infty \lambda^j C_{\alpha\gamma}^{(j)} \delta_{\kappa\gamma} (1 - \delta_{\alpha\gamma}) \right) \\
&= N \sum_{\kappa \in \{n\}} \left| \bar{\phi}_\kappa \right\rangle \left(\delta_{\kappa\alpha} + \sum_{j=1}^\infty \lambda^j C_{\alpha\kappa}^{(j)} (1 - \delta_{\alpha\kappa}) \right) \quad (12.105)
\end{aligned}
$$

and

$$
\begin{aligned}
\hat{Q}_n \left| \psi_\alpha \right\rangle &= N \sum_{\kappa \ni \{n\}} \left| \bar{\phi}_\kappa \right\rangle \left(\delta_{\kappa\alpha} + \sum_{j=1}^\infty \lambda^j C_{\alpha\kappa}^{(j)} (1 - \delta_{\alpha\kappa}) \right) \\
&= N \sum_{\kappa \ni \{n\}} \left| \bar{\phi}_\kappa \right\rangle \sum_{j=1}^\infty \lambda^j C_{\alpha\kappa}^{(j)} \quad (12.106)
\end{aligned}
$$

since $\alpha \in \{n\}$.

For $\beta \in \{n\}$,

$$
\left\langle \bar{\phi}_\beta \right| \hat{P}_n = \left\langle \bar{\phi}_\beta \right| . \quad (12.107)
$$

Multiplying (12.101) on the left with $\left\langle \bar{\phi}_\beta \right|$, using (12.107) and the fact that this state is an eigenstate of $\hat{\bar{H}}_0$ gives

$$
\bar{E}_\beta^{(0)} \left\langle \bar{\phi}_\beta \right| \hat{P}_n \left| \psi_\alpha \right\rangle + \lambda \left\langle \bar{\phi}_\beta \right| \hat{H}_1 \hat{Q}_n \left| \psi_\alpha \right\rangle = E_\alpha \left\langle \bar{\phi}_\beta \right| \hat{P}_n \left| \psi_\alpha \right\rangle . \quad (12.108)
$$

Using (12.105), (12.106) and (12.103), this becomes

$$
\left(\bar{E}_\alpha^{(0)} + \sum_{i=1}^\infty \lambda^i \bar{E}_\alpha^{(i)} - \bar{E}_\beta^{(0)} \right) N \left(\delta_{\alpha\beta} + \sum_{j=1}^\infty \lambda^j C_{\alpha\beta}^{(j)} (1 - \delta_{\alpha\beta}) \right)
$$
$$
= \lambda N \sum_{j=1}^\infty \lambda^j \sum_{\kappa \ni \{n\}} \left\langle \bar{\phi}_\beta \right| \hat{H}_1 \left| \bar{\phi}_\kappa \right\rangle C_{\alpha\kappa}^{(j)} \quad (12.109)
$$

Expanding the left-hand side,

$$
\delta_{\alpha\beta} \sum_{i=1}^\infty \lambda^i \bar{E}_\alpha^{(i)} + \left(\bar{E}_\alpha^{(0)} - \bar{E}_\beta^{(0)} \right) \sum_{j=1}^\infty \lambda^j C_{\alpha\beta}^{(j)} (1 - \delta_{\alpha\beta})
$$
$$
+ \sum_{i=1}^\infty \lambda^i \bar{E}_\alpha^{(i)} \sum_{j=1}^\infty \lambda^j C_{\alpha\beta}^{(j)} (1 - \delta_{\alpha\beta}) = \sum_{j=1}^\infty \lambda^{j+1} \sum_{\kappa \ni \{n\}} \left\langle \bar{\phi}_\beta \right| \hat{H}_1 \left| \bar{\phi}_\kappa \right\rangle C_{\alpha\kappa}^{(j)} .
$$
$$
(12.110)
$$

Rearranging the sums to simplify the identification of the various powers of λ as we did for the non-degenerate case, yields

$$
\delta_{\alpha\beta} \sum_{k=1}^{\infty} \lambda^k \bar{E}_{\alpha}^{(k)} + \left(\bar{E}_{\alpha}^{(0)} - \bar{E}_{\beta}^{(0)} \right) \sum_{k=1}^{\infty} \lambda^k C_{\alpha\beta}^{(k)} (1 - \delta_{\alpha\beta})
$$

$$
+ \sum_{k=2}^{\infty} \sum_{i=1}^{k-1} \lambda^k \bar{E}_{\alpha}^{(i)} C_{\alpha\beta}^{(k-i)} (1 - \delta_{\alpha\beta}) = \sum_{k=2}^{\infty} \lambda^k \sum_{\kappa \ni \{n\}} \left\langle \bar{\phi}_{\beta} \right| \hat{H}_1 \left| \bar{\phi}_{\kappa} \right\rangle C_{\alpha\kappa}^{(k-1)}
$$

$$(12.111)$$

The terms linear in λ give

$$
\delta_{\alpha\beta} \bar{E}_{\alpha}^{(1)} + \left(\bar{E}_{\alpha}^{(0)} - \bar{E}_{\beta}^{(0)} \right) C_{\alpha\beta}^{(1)} (1 - \delta_{\alpha\beta}) = 0 \,. \tag{12.112}
$$

For $\alpha = \beta$, this requires that

$$
\bar{E}_{\alpha}^{(1)} = 0 \,. \tag{12.113}
$$

For $\alpha \neq \beta$, we get

$$
\left(\bar{E}_{\alpha}^{(0)} - \bar{E}_{\beta}^{(0)} \right) C_{\alpha\beta}^{(1)} = 0 \,. \tag{12.114}
$$

If $\bar{E}_{\alpha}^{(0)} \neq \bar{E}_{\beta}^{(0)}$, this requires that

$$
C_{\alpha\beta}^{(1)} = 0 \,. \tag{12.115}
$$

If, on the other hand, there are still degenerate states after the diagonalizing of $\hat{\bar{H}}_0$, it is not possible to determine the value of $C_{\alpha\beta}^{(1)}$ and we will have to modify the problem again by moving additional parts of the reaction terms into a new unperturbed hamiltonian to break the degeneracy. We will not consider that situation in detail here.

For $k \geq 2$, (12.111) gives

$$
\delta_{\alpha\beta} \bar{E}_{\alpha}^{(k)} + \left(\bar{E}_{\alpha}^{(0)} - \bar{E}_{\beta}^{(0)} \right) C_{\alpha\beta}^{(k)} (1 - \delta_{\alpha\beta})
$$

$$
+ \sum_{i=1}^{k-1} \bar{E}_{\alpha}^{(i)} C_{\alpha\beta}^{(k-i)} (1 - \delta_{\alpha\beta}) = \sum_{\kappa \ni \{n\}} \left\langle \bar{\phi}_{\beta} \right| \hat{H}_1 \left| \bar{\phi}_{\kappa} \right\rangle C_{\alpha\kappa}^{(k-1)} \,.
$$

$$(12.116)$$

For $\alpha = \beta$, this yields

$$
\bar{E}_{\alpha}^{(k)} = \sum_{\kappa \ni \{n\}} \left\langle \bar{\phi}_{\alpha} \right| \hat{H}_1 \left| \bar{\phi}_{\kappa} \right\rangle C_{\alpha\kappa}^{(k-1)} \,. \tag{12.117}
$$

For $\alpha \neq \beta$, (12.116) yields

$$
\left(\bar{E}_{\alpha}^{(0)} - \bar{E}_{\beta}^{(0)} \right) C_{\alpha\beta}^{(k)} + \sum_{i=1}^{k-1} \bar{E}_{\alpha}^{(i)} C_{\alpha\beta}^{(k-i)} = \sum_{\kappa \ni \{n\}} \left\langle \bar{\phi}_{\beta} \right| \hat{H}_1 \left| \bar{\phi}_{\kappa} \right\rangle C_{\alpha\kappa}^{(k-1)} \,. \tag{12.118}
$$

Assuming that there is no degeneracy, this can be solved to give

$$C_{\alpha\beta}^{(k)} = \frac{1}{\bar{E}_\alpha^{(0)} - \bar{E}_\beta^{(0)}} \left[\sum_{\kappa \ni \{n\}} \langle \bar{\phi}_\beta | \, \hat{H}_1 \, | \bar{\phi}_\kappa \rangle C_{\alpha\kappa}^{(k-1)} - \sum_{i=1}^{k-1} \bar{E}_\alpha^{(i)} C_{\alpha\beta}^{(k-i)} \right] . \tag{12.119}$$

Note that (12.117) and (12.119) contain coefficients $C_{\alpha\kappa}^{(k-1)}$ where $\kappa \ni \{n\}$. We have expressions for the coefficients only for the case where both indices are in $\{n\}$. Consider the case where $\beta \ni \{n\}$. Then,

$$\langle \bar{\phi}_\beta | \, \hat{Q}_n = \langle \bar{\phi}_\beta | \, . \tag{12.120}$$

Multiplying (12.102) from the left with $\langle \bar{\phi}_\beta |$, and using (12.120) and the fact that the state is an eigenstate of $\hat{H}_0$ yields

$$\bar{E}_\beta^{(0)} \langle \bar{\phi}_\beta | \, \hat{Q}_n \, | \psi_\alpha \rangle + \lambda \langle \bar{\phi}_\beta | \, \hat{H}_1 \hat{Q}_n \, | \psi_\alpha \rangle + \lambda \langle \bar{\phi}_\beta | \, \hat{H}_1 \hat{P}_n \, | \psi_\alpha \rangle = E_\alpha \langle \bar{\phi}_\beta | \, \hat{Q}_n \, | \psi_\alpha \rangle . \tag{12.121}$$

Using (12.105), (12.106) and (12.103), this becomes

$$\left(\bar{E}_\alpha^{(0)} - \bar{E}_\beta^{(0)} + \sum_{i=1}^{\infty} \lambda^i \bar{E}_\alpha^{(i)} \right) N \sum_{j=1}^{\infty} \lambda^j C_{\alpha\beta}^{(j)} = \lambda \langle \bar{\phi}_\beta | \, \hat{H}_1 N \sum_{\kappa \ni \{n\}} \sum_{j=1}^{\infty} \lambda^j \, | \bar{\phi}_\kappa \rangle C_{\alpha\kappa}^{(j)}$$

$$+ \lambda \langle \bar{\phi}_\beta | \, \hat{H}_1 N \sum_{\kappa \in \{n\}} | \bar{\phi}_\kappa \rangle \left(\delta_{\kappa\alpha} + \sum_{j=1}^{\infty} \lambda^j C_{\alpha\kappa}^{(j)} (1 - \delta_{\kappa\alpha}) \right) . \tag{12.122}$$

Canceling the normalization factor and expanding the products on both sides yields

$$\left(\bar{E}_\alpha^{(0)} - \bar{E}_\beta^{(0)} \right) \sum_{j=1}^{\infty} \lambda^j C_{\alpha\beta}^{(j)} + \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \lambda^{i+j} \bar{E}_\alpha^{(i)} C_{\alpha\beta}^{(j)} = \sum_{\kappa \ni \{n\}} \sum_{j=1}^{\infty} \lambda^{j+1} \langle \bar{\phi}_\beta | \, \hat{H}_1 \, | \bar{\phi}_\kappa \rangle C_{\alpha\kappa}^{(j)}$$

$$+ \lambda \langle \bar{\phi}_\beta | \, \hat{H}_1 \, | \bar{\phi}_\alpha \rangle + \sum_{j=1}^{\infty} \lambda^{j+1} \sum_{\kappa \in \{n\}} \langle \bar{\phi}_\beta | \, \hat{H}_1 \, | \bar{\phi}_\kappa \rangle C_{\alpha\kappa}^{(j)} (1 - \delta_{\kappa\alpha}) . \tag{12.123}$$

Changing the sums to expose common powers of λ gives

$$\left(\bar{E}_\alpha^{(0)} - \bar{E}_\beta^{(0)} \right) \sum_{k=1}^{\infty} \lambda^k C_{\alpha\beta}^{(k)} + \sum_{k=2}^{\infty} \lambda^k \sum_{i=1}^{k-1} \bar{E}_\alpha^{(i)} C_{\alpha\beta}^{(k-i)} = \sum_{k=2}^{\infty} \lambda^k \sum_{\kappa \ni \{n\}} \langle \bar{\phi}_\beta | \, \hat{H}_1 \, | \bar{\phi}_\kappa \rangle C_{\alpha\kappa}^{(k-1)}$$

$$+ \lambda \langle \bar{\phi}_\beta | \, \hat{H}_1 \, | \bar{\phi}_\alpha \rangle + \sum_{k=2}^{\infty} \lambda^k \sum_{\kappa \in \{n\}} \langle \bar{\phi}_\beta | \, \hat{H}_1 \, | \bar{\phi}_\kappa \rangle C_{\alpha\kappa}^{(k-1)} (1 - \delta_{\kappa\alpha}) . \tag{12.124}$$

Equating the coefficients of the terms linear in λ yields

$$\left(\bar{E}_\alpha^{(0)} - \bar{E}_\beta^{(0)} \right) C_{\alpha\beta}^{(1)} = \langle \bar{\phi}_\beta | \, \hat{H}_1 \, | \bar{\phi}_\alpha \rangle . \tag{12.125}$$

Since α and β are in different subspaces in this case there is no possibility that $\bar{E}_\alpha^{(0)} - \bar{E}_\beta^{(0)}$ will vanish. So we can always write

$$C_{\alpha\beta}^{(1)} = \frac{\langle \bar{\phi}_\beta | \hat{H}_1 | \bar{\phi}_\alpha \rangle}{\bar{E}_\alpha^{(0)} - \bar{E}_\beta^{(0)}} \, . \tag{12.126}$$

For powers $k \geq 2$, (12.124) gives

$$\left(\bar{E}_\alpha^{(0)} - \bar{E}_\beta^{(0)} \right) C_{\alpha\beta}^{(k)} + \sum_{i=1}^{k-1} \bar{E}_\alpha^{(i)} C_{\alpha\beta}^{(k-i)} = \sum_{\kappa \ni \{n\}} \langle \bar{\phi}_\beta | \hat{H}_1 | \bar{\phi}_\kappa \rangle C_{\alpha\kappa}^{(k-1)}$$

$$+ \sum_{\kappa \in \{n\}} \langle \bar{\phi}_\beta | \hat{H}_1 | \bar{\phi}_\kappa \rangle C_{\alpha\kappa}^{(k-1)} (1 - \delta_{\kappa\alpha}) \, . \tag{12.127}$$

This can be solved to give

$$C_{\alpha\beta}^{(k)} = \frac{1}{\bar{E}_\alpha^{(0)} - \bar{E}_\beta^{(0)}} \left[\sum_{\kappa \ni \{n\}} \langle \bar{\phi}_\beta | \hat{H}_1 | \bar{\phi}_\kappa \rangle C_{\alpha\kappa}^{(k-1)} \right.$$

$$\left. + \sum_{\kappa \in \{n\}} \langle \bar{\phi}_\beta | \hat{H}_1 | \bar{\phi}_\kappa \rangle C_{\alpha\kappa}^{(k-1)} (1 - \delta_{\kappa\alpha}) - \sum_{i=1}^{k-1} \bar{E}_\alpha^{(i)} C_{\alpha\beta}^{(k-i)} \right] \, . \tag{12.128}$$

In the case that all of the degeneracy in the subspace $\{n\}$ is removed by diagonalizaing $\hat{\bar{H}}_0$, the perturbation series is completely described by (12.113), (12.115), (12.117), (12.119), (12.126) and (12.128). To see how this works in practice, we will now consider several calculations involving the hydrogen atom.

12.4 Leading Order Corrections to the Hydrogen Spectrum

The hamitonian for the hydrogen atom contains the spin-dependent spin-orbit potential, so we need to include the spin in to unperturbed wave function. The simplest form of this is the direct-product wave function

$$\left| \phi_{nlm_l; \frac{1}{2}m_s} \right\rangle = | \phi_{nlm} \rangle \left| \tfrac{1}{2}m_s \right\rangle \, . \tag{12.129}$$

To understand the perturbative corrections to the energy eigenvalues and wave functions of hydrogen it is useful to consider the constants of motion of the full hamiltonian. For $\hat{H}_0$ we have set of mutually commuting operators $\hat{H}_0$, $\hat{\boldsymbol{L}}^2$, $\hat{L}_z$, $\hat{\boldsymbol{S}}^2$ and

$\hat{S}_z$. The relativistic correction hamiltonian can be shown to satisfy the commutation relations

$$[\hat{H}_{rel}, \hat{\boldsymbol{L}}^2] = 0 \tag{12.130}$$

$$[\hat{H}_{rel}, \hat{L}_z] = 0 \tag{12.131}$$

$$[\hat{H}_{rel}, \hat{\boldsymbol{S}}^2] = 0 \tag{12.132}$$

$$[\hat{H}_{rel}, \hat{S}_z] = 0 \tag{12.133}$$

so it satisfies the same commutation relations as $\hat{H}_0$. In the absence of the spin-orbit potential, the quantum numbers labeling the eigenstates of $\hat{H}_0 + \hat{H}_{rel}$ would be n, $l,m_l,s = \frac{1}{2}$ and m_s. In this case, (12.19) could be used to obtain the coefficients

$$\left\langle \phi_{n'l'm'_l;\frac{1}{2}m'_s} \middle| \psi_{nlm_l;\frac{1}{2}m_s} \right\rangle = \delta_{l'l}\delta_{m'_l,m_l}\delta_{m'_s m_s} \left\langle \phi_{n'lm_l;\frac{1}{2}m_s} \middle| \psi_{nlm_l;\frac{1}{2}m_s} \right\rangle \tag{12.134}$$

and, due to the commutation relations, the matrix element of $\hat{H}_{rel}$ will be

$$\left\langle \phi_{n'l'm'_l;\frac{1}{2}m'_s} \middle| \hat{H}_{rel} \middle| \phi_{nlm_l;\frac{1}{2}m_s} \right\rangle = \delta_{l'l}\delta_{m'_l,m_l}\delta_{m'_s m_s} \left\langle \phi_{n'lm_l;\frac{1}{2}m_s} \middle| \hat{H}_{rel} \middle| \phi_{nlm_l;\frac{1}{2}m_s} \right\rangle . \tag{12.135}$$

This means that the linear equations described by (12.19) for each set of values for l, m_l and m_s will decouple. That is, only states with a given l, m_l and m_s, and different values of n will be mixed. This means that the degenerate states for any given n will not be mixed and the perturbative corrections can be calculated without recourse to extra complexities of degenerate-state perturbation theory.

If we now include the spin-orbit interaction, things get a little more complicated. This happens because

$$[\hat{\boldsymbol{L}} \cdot \hat{\boldsymbol{S}}, \hat{L}_z] = i\hbar(\hat{\boldsymbol{L}} \times \hat{\boldsymbol{S}})_z \tag{12.136}$$

and

$$[\hat{\boldsymbol{L}} \cdot \hat{\boldsymbol{S}}, \hat{S}_z] = -i\hbar(\hat{\boldsymbol{L}} \times \hat{\boldsymbol{S}})_z . \tag{12.137}$$

This means that states for a given n and l but different values of m_l and m_s would mix. As a result, the spin-orbit potential would mix degenerate state and recourse to degenerate-state perturbation theory would be necessary.

It is, however, possible to partially diagonalize the hamiltonian and avoid degenerate-state perturbation theory by noting that we can use

$$\hat{\boldsymbol{J}}^2 = (\hat{\boldsymbol{L}} + \hat{\boldsymbol{S}})^2 = \hat{\boldsymbol{L}}^2 + 2\hat{\boldsymbol{L}} \cdot \hat{\boldsymbol{S}} + \hat{\boldsymbol{S}}^2 \tag{12.138}$$

to write

$$\hat{\boldsymbol{L}} \cdot \hat{\boldsymbol{S}} = \frac{1}{2}(\hat{\boldsymbol{J}}^2 - \hat{\boldsymbol{L}}^2 - \hat{\boldsymbol{S}}^2) . \tag{12.139}$$

This immediately implies that

$$[\hat{H}_{SO}, \hat{\boldsymbol{J}}^2] = 0 \tag{12.140}$$

$$[\hat{H}_{SO}, \boldsymbol{J}_z] = 0 \tag{12.141}$$

$$[\hat{H}_{SO}, \hat{\boldsymbol{L}}^2] = 0 \tag{12.142}$$

$$[\hat{H}_{SO}, \hat{\boldsymbol{S}}^2] = 0 , \tag{12.143}$$

and $\hat{H}_0$ and $\hat{H}_{rel}$ will also satisfy these commutation relations as we have seen in our discussion of angular momentum coupling. This means that the eigenstates of the full hamiltonian will be labeled by the set of quantum numbers n, l, $s = \frac{1}{2}$, j and m_j. That is

$$|\psi_\alpha\rangle = \left|\psi_{nl\frac{1}{2}jm_j}\right\rangle \tag{12.144}$$

We can, therefore, avoid using degenerate-state perturbation theory by using an un-perturbed basis set that has the same angular momentum quantum numbers as $|\psi_\alpha\rangle$. This can be done by simply coupling the orbital and spin angular momentum of the product states to give comparable states in the total angular momentum basis by defining

$$\left|\phi_{nl\frac{1}{2}jm_j}\right\rangle = \sum_{m_l m_s} \langle l, m_l; \tfrac{1}{2}m_s \left| jm_j\rangle \left|\phi_{nlm_l}\right\rangle \left|\tfrac{1}{2}m_s\right\rangle . \tag{12.145}$$

This basis then decouples the degenerate states and we can proceed without recourse to degenerate-perturbation theory.

We can now calculate the first-order corrections to the hydrogen spectrum. We will start with the calculation of the relativistic kinematical correction. This can be simplified by rewriting the interaction as

$$\hat{H}_{rel} = -\frac{1}{2m_e c^2}\left(\frac{\hat{\boldsymbol{p}}^2}{2m_e}\right)^2 . \tag{12.146}$$

Note that a straightforward approach to the calculation of the energy shift from this term would involve applying the square of the laplacian to the unperturbed wave functions which would be possible but tedious. We eliminate this problem by using the hamiltonian $\hat{H}_0$ to obtain

$$\frac{\hat{\boldsymbol{p}}^2}{2m_e} = \hat{H}_0 + \frac{e^2}{r} = \hat{H}_0 + \frac{\alpha\hbar c}{r} . \tag{12.147}$$

We can use this to rewrite the kinematical correction to the hamiltonian as

$$\hat{H}_{rel} = -\frac{1}{2m_e c^2}\left(\hat{H}_0 + \frac{\alpha\hbar c}{r}\right)^2 . \tag{12.148}$$

We can now calculate the first order energy shift due to the relativistic kinematical

correction

$$
\begin{aligned}
E_{nl}^{rel(1)} &= \left\langle \phi_{nl\frac{1}{2}jm_j} \middle| \hat{H}_{rel} \middle| \phi_{nl\frac{1}{2}jm_j} \right\rangle = -\frac{1}{2m_e c^2} \left\langle \phi_{nl\frac{1}{2}jm_j} \middle| \left(\frac{\hat{\boldsymbol{p}}^2}{2m_e} \right)^2 \middle| \phi_{nl\frac{1}{2}jm_j} \right\rangle \\
&= -\frac{1}{2m_e c^2} \left\langle \phi_{nl\frac{1}{2}jm_j} \middle| \left(\hat{H}_0 + \frac{\alpha \hbar c}{r} \right)^2 \middle| \phi_{nl\frac{1}{2}jm_j} \right\rangle \\
&= -\frac{1}{2m_e c^2} \left\langle \phi_{nl\frac{1}{2}jm_j} \middle| \left(\hat{H}_0^2 + \hat{H}_0 \frac{\alpha \hbar c}{r} + \frac{\alpha \hbar c}{r} \hat{H}_0 + \frac{\alpha^2 \hbar^2 c^2}{r^2} \right) \middle| \phi_{nl\frac{1}{2}jm_j} \right\rangle \\
&= -\frac{1}{2m_e c^2} \left(\hat{E}_n^{(0)\,2} + 2\hat{E}_n^{(0)} \alpha \hbar c \left\langle \phi_{nl\frac{1}{2}jm_j} \middle| \frac{1}{r} \middle| \phi_{nl\frac{1}{2}jm_j} \right\rangle \right. \\
&\qquad \left. + \alpha^2 \hbar^2 c^2 \left\langle \phi_{nl\frac{1}{2}jm_j} \middle| \frac{1}{r^2} \middle| \phi_{nl\frac{1}{2}jm_j} \right\rangle \right)
\end{aligned}
\tag{12.149}
$$

To simplify this further, consider the matrix element

$$
\begin{aligned}
\left\langle \phi_{nl\frac{1}{2}jm_j} \middle| \hat{r}^k \middle| \phi_{nl\frac{1}{2}jm_j} \right\rangle &= \int_0^\infty dr\, r^2 \int_0^{2\pi} d\varphi \int_0^\pi d\theta \sin\theta \sum_{m_l' m_s'} \left\langle l, m_l'; \tfrac{1}{2} m_s' \middle| jm_j \right\rangle \\
&\quad \times R_{nl}(r) Y_{lm_l'}^*(\theta, \varphi) \left\langle \tfrac{1}{2} m_s' \middle| r^k \right. \\
&\quad \times \sum_{m_l m_s} \left\langle l, m_l; \tfrac{1}{2} m_s \middle| jm_j \right\rangle R_{nl}(r) Y_{lm_l}(\theta, \varphi) \left| \tfrac{1}{2} m_s \right\rangle \\
&= \int_0^\infty dr\, r^{2+k} R_{nl}^2(r) \sum_{m_l' m_s'} \sum_{m_l m_s} \left\langle \tfrac{1}{2} m_s' \middle| \tfrac{1}{2} m_s \right\rangle \\
&\quad \times \int_0^{2\pi} d\varphi \int_0^\pi d\theta \sin\theta\, Y_{lm_l'}^*(\theta, \varphi) Y_{lm_l}(\theta, \varphi) \\
&\quad \times \left\langle l, m_l'; \tfrac{1}{2} m_s' \middle| jm_j \right\rangle \left\langle l, m_l; \tfrac{1}{2} m_s \middle| jm_j \right\rangle \\
&= \int_0^\infty dr\, r^{2+k} R_{nl}^2(r) \\
&\quad \times \sum_{m_l m_s} \left\langle l, m_l; \tfrac{1}{2} m_s \middle| jm_j \right\rangle \left\langle l, m_l; \tfrac{1}{2} m_s \middle| jm_j \right\rangle \\
&= \int_0^\infty dr\, r^{2+k} R_{nl}^2(r)\,.
\end{aligned}
\tag{12.150}
$$

Here, we need the integrals

$$
\int_0^\infty dr\, r R_{nl}^2(r) = \frac{\alpha m_e c}{\hbar n^2}
\tag{12.151}
$$

and

$$
\int_0^\infty dr\, R_{nl}^2(r) = \frac{(\alpha m_e c)^2}{\hbar^2 n^3 \left(l + \frac{1}{2} \right)}\,.
\tag{12.152}
$$

216

We can now write the first-order energy shift due to the kinematical relativistic correction as

$$
\begin{aligned}
E_{nl}^{rel(1)} &= -\frac{1}{2m_e c^2}\left(\frac{\alpha^4 m_e^2 c^4}{4n^4} - 2\frac{\alpha^2 m_e c^2}{2n^2}\alpha\hbar c\frac{\alpha m_e c}{\hbar n^2} + \alpha^2\hbar^2 c^2\frac{(\alpha m_e c)^2}{\hbar^2 n^3\left(l+\frac{1}{2}\right)}\right) \\
&= \frac{\alpha^4 m_e c^2}{2}\left[\frac{3}{4n^4} - \frac{1}{n^3\left(l+\frac{1}{2}\right)}\right].
\end{aligned} \tag{12.153}
$$

Now we will consider the spin-orbit interaction. This can be calculated to give

$$
\begin{aligned}
E_{nlj}^{SO(1)} &= \left\langle \phi_{nl\frac{1}{2}jm_j}\left|\hat{H}_{SO}\right|\phi_{nl\frac{1}{2}jm_j}\right\rangle = \frac{\alpha\hbar c}{2m_e^2 c^2}\left\langle \phi_{nl\frac{1}{2}jm_j}\left|\frac{\hat{\boldsymbol{L}}\cdot\hat{\boldsymbol{S}}}{r^3}\right|\phi_{nl\frac{1}{2}jm_j}\right\rangle \\
&= \frac{\alpha\hbar c}{4m_e^2 c^2}\left\langle \phi_{nl\frac{1}{2}jm_j}\left|\frac{\hat{\boldsymbol{J}}^2 - \hat{\boldsymbol{L}}^2 - \hat{\boldsymbol{S}}^2}{r^3}\right|\phi_{nl\frac{1}{2}jm_j}\right\rangle \\
&= \frac{\alpha\hbar c}{4m_e^2 c^2}\hbar^2\left[j(j+1) - l(l+1) - \frac{3}{4}\right]\left\langle \phi_{nl\frac{1}{2}jm_j}\left|\frac{1}{r^3}\right|\phi_{nl\frac{1}{2}jm_j}\right\rangle \\
&= \frac{\alpha\hbar c}{4m_e^2 c^2}\hbar^2\left[j(j+1) - l(l+1) - \frac{3}{4}\right]\frac{(\alpha m_e c)^3}{\hbar^3 n^3 l\left(l+\frac{1}{2}\right)(l+1)} \\
&= \frac{\alpha^4 m_e c^2}{4}\frac{j(j+1) - l(l+1) - \frac{3}{4}}{n^3 l\left(l+\frac{1}{2}\right)(l+1)},
\end{aligned} \tag{12.154}
$$

where we have used the integral

$$
\int_0^\infty dr\, r^{-1} R_{nl}^2(r) = \frac{(\alpha m_e c)^3}{\hbar^3 n^3 l\left(l+\frac{1}{2}\right)(l+1)} \tag{12.155}
$$

and taken $g = 2$.

We can now compare the results of these calculations with the spectrum given by the Dirac equation with the Coulomb potential which has eigenenergies

$$
E_{nj}^{Dirac} = m_e c^2\left[\frac{1}{\sqrt{1 + \frac{\alpha^2}{\left[n-\left(j+\frac{1}{2}\right)+\sqrt{\left(j+\frac{1}{2}\right)^2-\alpha^2}\right]^2}}} - 1\right]. \tag{12.156}
$$

The results of these corrections are shown in Table 12.1 for the case of hydrogen. All states for $1 \le n \le 3$ are shown with the principle quantum number n, the angular momentum quantum number l and the total angular momentum quantum number j. The unperturbed spectrum $E_n^{(0)}$ depends only on the principal quantum number n and all states of a fixed n are therefore degenerate. The relativistic correction depends upon both n and l and therefore breaks the degeneracy in l but still produces states

217

Table 12.1: First order corrections to the hydrogen atom spectrum.

n	l	j	$E^{(0)}$	$E^{rel(1)}$	$E^{SO(1)}$	$E^{total(1)}$	E^{Dirac}
1	0	1/2	-13.6056923	-0.0009057	0.0000000	-13.6065980	-13.6044245
2	0	1/2	-3.4014231	-0.0001472	0.0000000	-3.4015702	-3.4012986
2	1	1/2	-3.4014231	-0.0000264	-0.0000302	-3.4014797	-3.4012986
2	1	3/2	-3.4014231	-0.0000264	0.0000151	-3.4014344	-3.4013438
3	0	1/2	-1.5117436	-0.0000470	0.0000000	-1.5117906	-1.5117100
3	1	1/2	-1.5117436	-0.0000112	-0.0000089	-1.5117637	-1.5117100
3	1	3/2	-1.5117436	-0.0000112	0.0000045	-1.5117503	-1.5117235
3	2	3/2	-1.5117436	-0.0000040	-0.0000027	-1.5117503	-1.5117235
3	2	5/2	-1.5117436	-0.0000040	0.0000018	-1.5117458	-1.5117279

of differing j for a given n and l which are still degenerate. This correction always degreases the energy. The spin-orbit interaction depends upon the values of n, l and j. This could be expected to further lift the degeneracy of the states. However, note that in the case of the $n = 3$ states, the sum of the the zeroth and first order corrections results in the two states with $j = 3/2$ but differing values of l being degenerate. Comparison shows that this perturbatively corrected spectrum with the spectrum from the solution of the Dirac equation with a Coulomb potential shows remarkably good agreement. Since the Dirac spectrum depends only upon n and j, the degeneracy in the $n = 3$, $j = 3/2$ states is explained and is not the result of an accident.

12.4.1 The Stark Effect

Consider the case of a hydrogen atom in a uniform electric field aligned along the z-axis. The contribution to the hamiltonian from the interaction of the charged electron with the external field is

$$\hat{H}_1 = -e\mathcal{E}\hat{z}\,, \tag{12.157}$$

where $\mathcal{E}$ is the magnitude of the electric field. Since this interaction is independent of the electron spin, we will neglect the spin to simplify the discussion. The commutators of $\hat{H}_1$ with the angular momentum operators are

$$[\hat{L}_z, \hat{H}_1] = 0 \tag{12.158}$$

and

$$[\hat{\boldsymbol{L}}^2, \hat{H}_1] = e\mathcal{E}[i\hbar(\hat{\boldsymbol{r}} \times \hat{\boldsymbol{L}})_3 + \hbar^2\hat{z}]\,. \tag{12.159}$$

This means that the interaction will not mix states with different m_l, but will mix states with different n and l. The degeneracy in m_l is not, therefore, a problem.

However, the interaction will mix degenerate states for a fixed n but different values of l. As a result we a compelled to use degenerate-state perturbation theory.

The matrix elements of the interaction hamiltonian are given by

$$\langle \phi_{n'l'm'} | \hat{H}_1 | \phi_{nlm} \rangle = -e\mathcal{E} \langle \phi_{n'l'm'} | \hat{z} | \phi_{nlm} \rangle = -e\mathcal{E} \int_0^\infty dr r^2 R_{n'l'}(r) R_{nl}(r) \langle l'm' | \hat{r}_0 | lm \rangle .$$
(12.160)

We can use the Wigner-Eckart theorem to calculate the matrix element of $\hat{r}_0$ between angular momentum states giving

$$\langle l'm' | \hat{r}_0 | lm \rangle = \frac{(-1)^{l-m}}{\sqrt{3}} \langle l'm'; l, -m | 10 \rangle \langle l' \| \hat{r} \| l \rangle .$$
(12.161)

The reduced matrix element is obtained by calculating the matrix element for $m' = m = 0$ directly. Then,

$$\langle l' \| \hat{r} \| l \rangle = \sqrt{3}(-1)^l \frac{\langle l'0 | \hat{r}_0 | l0 \rangle}{\langle l'0; l, 0 | 10 \rangle} .$$
(12.162)

The general matrix element can then be rewritten as

$$\langle l'm' | \hat{r}_0 | lm \rangle = (-1)^m \frac{\langle l'm'; l, -m | 10 \rangle}{\langle l'0; l, 0 | 10 \rangle} \langle l'0 | \hat{r}_0 | l0 \rangle .$$
(12.163)

The $m' = m = 0$ matrix element is then

$$
\begin{aligned}
\langle l'0 | \hat{r}_0 | l0 \rangle &= \int_0^{2\pi} d\varphi \int_0^\pi d\theta \sin\theta Y_{l'0}^*(\theta, \varphi) r \cos\theta Y_{l0}(\theta, \varphi) \\
&= r \int_0^{2\pi} d\varphi \int_0^\pi d\theta \sin\theta \sqrt{\frac{2l'+1}{4\pi}} P_{l'}(\cos\theta) \cos\theta \sqrt{\frac{2l+1}{4\pi}} P_l(\cos\theta) \\
&= \frac{r}{2} \sqrt{(2l'+1)(2l+1)} \int_0^\pi d\theta \sin\theta P_{l'}(\cos\theta) \cos\theta P_l(\cos\theta) \\
&= \frac{r}{2} \sqrt{(2l'+1)(2l+1)} \int_{-1}^1 dx P_{l'}(x) x P_l(x)
\end{aligned}
$$
(12.164)

We can use the recursion relation for the Legendre polynomials (Abramowitz and Stegun 8.5.3)

$$x P_l(x) = \frac{1}{2l+1} [(l+1) P_{l+1}(x) + l P_{l-1}(x)]$$
(12.165)

along with the orthogonality condition for these polynomials to evaluate the integral

$$
\begin{aligned}
\int_{-1}^1 dx P_{l'}(x) x P_l(x) &= \int_{-1}^1 dx \frac{1}{2l+1} [(l+1) P_{l'}(x) P_{l+1}(x) + l P_{l'}(x) P_{l-1}(x)] \\
&= \frac{1}{2l+1} \left[(l+1) \frac{2}{2l'+1} \delta_{l',l+1} + l \frac{2}{2l'+1} \delta_{l',l-1} \right] .
\end{aligned}
$$
(12.166)

Therefore,

$$\langle l'0| \hat{r}_0 |l0\rangle = \frac{r}{\sqrt{(2l'+1)(2l+1)}} \left[(l+1)\delta_{l',l+1} + l\delta_{l',l-1}\right] . \tag{12.167}$$

The Clebsch-Gordan coefficients that we will need are of the form of

$$\langle l+1, m; l, -m |10\rangle = (-1)^{l-m}\sqrt{\frac{3(l-m+1)(l+m+1)}{(l+1)(2l+1)(2l+3)}} \tag{12.168}$$

or

$$\langle l-1, m; l, -m |10\rangle = -(-1)^{l-m}\sqrt{\frac{3(l-m)(l+m)}{l(2l-1)(2l+1)}} . \tag{12.169}$$

Using these,

$$\frac{\langle l+1, m; l, -m |10\rangle}{\langle l+1, 0; l, 0 |10\rangle} = (-1)^m \frac{\sqrt{(l-m+1)(l+m+1)}}{l+1} \tag{12.170}$$

and

$$\frac{\langle l-1, m; l, -m |10\rangle}{\langle l-1, 0; l, 0 |10\rangle} = (-1)^m \frac{\sqrt{(l-m)(l+m)}}{l} . \tag{12.171}$$

The angular momentum matrix element is then

$$\langle l'm'| \hat{r}_0 |lm\rangle = \delta_{m'm} r \left(\sqrt{\frac{(l-m+1)(l+m+1)}{(2l+1)(2l+3)}}\delta_{l',l+1} + \sqrt{\frac{(l-m)(l+m)}{(2l-1)(2l+1)}}\delta_{l',l-1} \right) . \tag{12.172}$$

Notice that only states with $l' = l \pm 1$ are coupled by this matrix element.

For simplicity, we will confine ourselves to obtaining calculating the corrections to the states for $n = 2$. This contains degenerate states for $l = 0$ and $l = 1$ which will be mixed by the interaction. The only states with the same values of m that can be mixed are then the $l = 0$, $m = 0$ and $l = 1$, $m = 0$ states. The $l = 1$, $m = \pm 1$ states are not mixed and can be corrected perturbatively without recourse the degenerate-state perturbation theory.

The remaining problem is to calculate the radial integrals. The radial wave functions for $n = 1$ and $n = 2$ are

$$R_{10}(r) = \left(\frac{1}{a_0}\right)^{\frac{3}{2}} 2e^{-\frac{1}{a_0}r} \tag{12.173}$$

$$R_{20}(r) = \left(\frac{1}{2a_0}\right)^{\frac{3}{2}} \left(2 - \frac{1}{a_0}r\right) e^{-\frac{1}{2a_0}r} \tag{12.174}$$

$$R_{21}(r) = \left(\frac{1}{2a_0}\right)^{\frac{3}{2}} \frac{1}{\sqrt{3}a_0} re^{-\frac{1}{2a_0}r} . \tag{12.175}$$

We require the integral

$$
\begin{aligned}
\int_0^\infty dr\, r^3 R_{20}(r) R_{21}(r) &= \int_0^\infty dr\, r^3 \left(\frac{1}{2a_0}\right)^{\frac{3}{2}} \left(2 - \frac{1}{a_0} r\right) e^{-\frac{1}{2a_0} r} \left(\frac{1}{2a_0}\right)^{\frac{3}{2}} \frac{1}{\sqrt{3} a_0} r e^{-\frac{1}{2a_0} r} \\
&= \frac{1}{\sqrt{3}} \left(\frac{1}{2a_0}\right)^3 \int_0^\infty dr\, r^3 \frac{1}{a_0} r \left(2 - \frac{1}{a_0} r\right) e^{-\frac{1}{a_0} r} \\
&= \frac{1}{\sqrt{3}} \left(\frac{1}{2a_0}\right)^3 a_0^4 \int_0^\infty dx\, x^4 \left(2 - x\right) e^{-x} \\
&= \frac{1}{8\sqrt{3}} a_0(-72) = -\frac{9}{\sqrt{3}} a_0\,.
\end{aligned} \tag{12.176}
$$

We can now obtain the two non-zero matrix elements of $\hat{H}_1$ for $n = 2$,

$$
\langle \phi_{210}| \hat{H}_1 |\phi_{200}\rangle = \langle \phi_{200}| \hat{H}_1 |\phi_{210}\rangle = 3e\mathcal{E} a_0\,. \tag{12.177}
$$

The first step in obtaining the perturbative correction to these states is to diagonalize the hamiltonian in the subspace of the two mixing states. This is done by solving the eigenvalue problem for the matrix

$$
\begin{pmatrix} E_2^{(0)} & \langle \phi_{210}| \hat{H}_1 |\phi_{200}\rangle \\ \langle \phi_{200}| \hat{H}_1 |\phi_{210}\rangle & E_2^{(0)} \end{pmatrix} = \begin{pmatrix} E_2^{(0)} & 3e\mathcal{E} a_0 \\ 3e\mathcal{E} a_0 & E_2^{(0)} \end{pmatrix} \tag{12.178}
$$

This has eigenenergies

$$
\bar{E}_{2,\pm,0}^{(0)} = E_2^{(0)} \pm 3e\mathcal{E} \tag{12.179}
$$

with eigenvectors

$$
\left|\bar{\phi}_{2,\pm,0}\right\rangle = \frac{1}{\sqrt{2}} \left(|\phi_{200}\rangle \pm |\phi_{210}\rangle\right)\,. \tag{12.180}
$$

This diagonalization clearly breaks the degeneracy and the calculations using perturbation series for degenerate-state perturbation theory as given by (12.113), (12.115), (12.117), (12.119), (12.126) and (12.128) can now be carried out.

Chapter 13

Feynman Path Integrals

In this chapter we will introduce the Feynman path integral as means of describing the time evolution of a Schrödinger wave function. This provides an alternate way of understanding the physical content of quantum mechanics and forms the basis which is most commonly used to construct quantum field theories.

The starting point for our derivation is (8.69) which we will rewrite as

$$|\Psi(t)\rangle = U(t, t') |\Psi(t')\rangle , \tag{13.1}$$

where

$$\hat{U}(t, t') = e^{-\frac{i}{\hbar} \hat{H}(t - t')} . \tag{13.2}$$

In coordinate space, we can express this as

$$\Psi(\boldsymbol{r}, t) = \langle \boldsymbol{r} | \Psi(t) \rangle = \langle \boldsymbol{r} | U(t, t') | \Psi(t') \rangle = \int d^3 r' \langle \boldsymbol{r} | U(t, t') | \boldsymbol{r}' \rangle \langle \boldsymbol{r}' | \Psi(t') \rangle \tag{13.3}$$

If we now define

$$K(\boldsymbol{r}, t; \boldsymbol{r}', t') \equiv \langle \boldsymbol{r} | \hat{U}(t, t') | \boldsymbol{r}' \rangle , \tag{13.4}$$

This can be rewritten as

$$\Psi(\boldsymbol{r}, t) = \int d^3 r' K(\boldsymbol{r}, t; \boldsymbol{r}', t') \Psi(\boldsymbol{r}', t') . \tag{13.5}$$

The function $K(\boldsymbol{r}, t; \boldsymbol{r}', t')$ is called the *propagator*. Knowledge of the propagator is equivalent to that of the complete solution to the Schrödinger equation since given the wave function at some initial time t' it gives the wave function at any other time t. Since the original solution of the Schrödinger equation requires the imposition of boundary conditions, the propagator must also contain boundary conditions. For example if $t > t'$, then $\Psi(\boldsymbol{r}', t')$ is the initial condition and the propagator must require that $t > t'$. That is $K(\boldsymbol{r}, t; \boldsymbol{r}', t') = 0$ for $t' > t$. This describes the *retarded propagator*.

The propagator can be rewritten in terms of a Feynman path integral. In performing this derivation, we will consider a simple system in one dimension that is describe by the hamiltonian operator

$$\hat{H} = H(\hat{p}, \hat{x}) \,. \tag{13.6}$$

For convenience, we will also assume that $\hat{H}$ is written in *normal form* with all occurrences of the the momentum operator to the left of any occurrences of the coordinate operator in any term in the hamiltonian. We can now use the eigenequations for the momentum and coordinate space operators

$$\hat{p} \ket{p} = p \ket{p} \tag{13.7}$$

and

$$\hat{x} \ket{x} = x \ket{x} \tag{13.8}$$

along with

$$\braket{x | p} = \frac{1}{\sqrt{2\pi\hbar}} e^{\frac{i}{\hbar} p x} \tag{13.9}$$

to write a mixed matrix element of the hamiltonian in normal form as

$$\bra{p} H(\hat{p}, \hat{x}) \ket{x} = H(p, x) \braket{p | x} = H(p, x) \frac{1}{\sqrt{2\pi\hbar}} e^{-\frac{i}{\hbar} p x} \,. \tag{13.10}$$

We can express the time-evolution operator in terms of a limit as

$$\hat{U}(t, t') = \lim_{N \to \infty} \left(1 - i \frac{\hat{H}(t - t')}{\hbar N} \right)^N = \lim_{N \to \infty} \left(1 - \frac{i}{\hbar} \hat{H} \delta t \right)^N , \tag{13.11}$$

where

$$\delta t = \frac{t - t'}{N} \,. \tag{13.12}$$

We have subdivided the time interval $t - t'$ into N subintervals of length δt. for convenience, we can label the end points of the subintervals as t_i, where $t_0 = t'$, $t_N = t$ and $t_n = n\delta t + t' = n\delta t + t_0$. The propagator can now be written as

$$K(x, t; x', t') = \lim_{N \to \infty} \bra{x} \left(1 - \frac{i}{\hbar} \hat{H} \delta t \right)^N \ket{x'} \,. \tag{13.13}$$

If we now write this by explicitly the power by a product an insert complete sets of states between every pair of factors in the product we obtain

$$K(x, t; x', t') = \lim_{N \to \infty} \int_{-\infty}^{\infty} \prod_{n=1}^{N-1} dx_n \bra{x_N} \left(1 - \frac{i}{\hbar} \hat{H} \delta t \right) \ket{x_{N-1}}$$

$$\times \bra{x_{N-1}} \left(1 - \frac{i}{\hbar} \hat{H} \delta t \right) \ket{x_{N-2}} \dots \bra{x_1} \left(1 - \frac{i}{\hbar} \hat{H} \delta t \right) \ket{x_0} \,, \tag{13.14}$$

224

where $x_N = x$ and $x_0 = x'$. Each of the factors

$$\langle x_n| \left(1 - \frac{i}{\hbar}\hat{H}\delta t\right) |x_{n-1}\rangle \tag{13.15}$$

is a propagator over the infinitesimal time interval $t_n - t_{n-1} = \delta t$ and connecting the point x_{n-1} to x_n. This propagator can be rewritten by inserting a complete set of momentum states to give

$$\langle x_n| \left(1 - \frac{i}{\hbar}\hat{H}\delta t\right) |x_{n-1}\rangle = \int_{-\infty}^{\infty} dp_{n-1} \langle x_n |p_{n-1}\rangle \langle p_{n-1}| \left(1 - \frac{i}{\hbar}\hat{H}\delta t\right) |x_{n-1}\rangle . \tag{13.16}$$

The mixed matrix element is

$$\langle p_{n-1}| \left(1 - \frac{i}{\hbar}\hat{H}\delta t\right) |x_{n-1}\rangle = \langle p_{n-1} |x_{n-1}\rangle - \frac{i}{\hbar}\delta t \langle p_{n-1}| \hat{H} |x_{n-1}\rangle$$

$$= \frac{1}{\sqrt{2\pi\hbar}} e^{-\frac{i}{\hbar}p_{n-1}x_{n-1}} - \frac{i}{\hbar}\delta t H(p_{n-1}, x_{n-1}) \frac{1}{\sqrt{2\pi\hbar}} e^{-\frac{i}{\hbar}p_{n-1}x_{n-1}}$$

$$= \frac{1}{\sqrt{2\pi\hbar}} e^{-\frac{i}{\hbar}p_{n-1}x_{n-1}} \left(1 - \frac{i}{\hbar}\delta t H(p_{n-1}, x_{n-1})\right)$$

$$= \frac{1}{\sqrt{2\pi\hbar}} e^{-\frac{i}{\hbar}p_{n-1}x_{n-1}} \left(e^{-\frac{i}{\hbar}\delta t H(p_{n-1}, x_{n-1})} + \mathcal{O}\left(\frac{1}{N^2}\right)\right)$$

$$= \frac{1}{\sqrt{2\pi\hbar}} e^{-\frac{i}{\hbar}(p_{n-1}x_{n-1}+\delta t H(p_{n-1},x_{n-1}))} + \mathcal{O}\left(\frac{1}{N^2}\right) \tag{13.17}$$

Using

$$\langle x_n |p_{n-1}\rangle = \frac{1}{\sqrt{2\pi\hbar}} e^{\frac{i}{\hbar}p_{n-1}x_n} , \tag{13.18}$$

$$\langle x_n| \left(1 - \frac{i}{\hbar}\hat{H}\delta t\right) |x_{n-1}\rangle = \frac{1}{2\pi\hbar} \int_{-\infty}^{\infty} dp_{n-1} e^{\frac{i}{\hbar}p_{n-1}x_n} e^{-\frac{i}{\hbar}(p_{n-1}x_{n-1}+\delta t H(p_{n-1},x_{n-1}))}$$

$$= \frac{1}{2\pi\hbar} \int_{-\infty}^{\infty} dp_{n-1} e^{\frac{i}{\hbar}[p_{n-1}(x_n-x_{n-1})-\delta t H(p_{n-1},x_{n-1})]} \tag{13.19}$$

So,

$$K(x,t;x',t') = \lim_{N\to\infty} \int_{-\infty}^{\infty} \prod_{n=1}^{N-1} dx_n \prod_{n=0}^{N-1} \frac{dp_n}{2\pi\hbar} e^{\frac{i}{\hbar}[p_{N-1}(x_N-x_{N-1})-\delta t H(p_{N-1},x_{N-1})]}$$

$$\times e^{\frac{i}{\hbar}[p_{N-2}(x_{N-1}-x_{N-2})-\delta t H(p_{N-2},x_{N-2})]} \dots e^{\frac{i}{\hbar}[p_0(x_1-x_0)-\delta t H(p_0,x_0)]}$$

$$= \lim_{N\to\infty} \int_{-\infty}^{\infty} \prod_{n=1}^{N-1} dx_n \prod_{n=0}^{N-1} \frac{dp_n}{2\pi\hbar} e^{\frac{i}{\hbar}\sum_{n=0}^{N-1}[p_n(x_{n+1}-x_n)-\delta t H(p_n,x_n)]}$$

$$= \lim_{N\to\infty} \int_{-\infty}^{\infty} \prod_{n=1}^{N-1} dx_n \prod_{n=0}^{N-1} \frac{dp_n}{2\pi\hbar} e^{\frac{i}{\hbar}\sum_{n=0}^{N-1}\delta t\left[p_n\frac{x_{n+1}-x_n}{\delta t} - H(p_n,x_n)\right]} . \tag{13.20}$$

Now parameterize $p(t)$ and $x(t)$ such that $p_n = p(t_n)$ and $x_n = x(t_n)$. In the limit $N \to \infty$ and $\delta t \to 0$, the argument of the exponential has the appearance of an integral over time. That is,

$$\lim_{N \to \infty} \sum_{n=0}^{N-1} \delta t \left[p_n \frac{x_{n+1} - x_n}{\delta t} - H(p_n, x_n) \right] = \int_{t_0 = t'}^{t_N = t} dt \left[p(t) \dot{x}(t) - H(p(t), x(t)) \right].$$

(13.21)

The propagator can then be written formally as

$$K(x, t; x', t') = \int \mathcal{D}x(t) \mathcal{D}p(t) e^{\frac{i}{\hbar} \int_{t'}^{t} dt [p(t)\dot{x}(t) - H(p(t), x(t))]},$$

(13.22)

where the measures $\mathcal{D}x$ and $\mathcal{D}p$ are defined only in terms of the limiting procedure given above.

The propagator in this form is an integral over all possible paths in phase space which connect the initial and final space-time points. This functional integral is weighted by an ex exponential of imaginary argument. This argument is determined by integrating the quantity

$$p\dot{x} - H(p, x)$$

(13.23)

over each path in phase space. This quantity looks like the canonical transformation from the hamiltonian to the lagrangian. We can see how this identification occurs by "integrating out" the momentum for a simple case.

Consider the hamiltonian operator for a particle moving in a potential

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

(13.24)

Since no term in this hamiltonian contains both the momentum and coordinate operator, it is clearly in normal form. Using this hamiltonian in (refFPdef),

$$K(x, t; x', t') = \lim_{N \to \infty} \int_{-\infty}^{\infty} \prod_{n=1}^{N-1} dx_n \prod_{n=0}^{N-1} \frac{dp_n}{2\pi\hbar} e^{\frac{i}{\hbar} \sum_{n=0}^{N-1} \delta t \left[p_n \frac{\delta x_n}{\delta t} - \frac{p_n^2}{2m} - V(x_n) \right]},$$

(13.25)

where $\delta x_n = x_{n+1} - x_n$. Consider the integral

$$\int_{-\infty}^{\infty} \frac{dp_n}{2\pi\hbar} e^{\frac{i}{\hbar} \delta t \left[p_n \frac{\delta x_n}{\delta t} - \frac{p_n^2}{2m} - V(x_n) \right]} = \int_{-\infty}^{\infty} \frac{dp_n}{2\pi\hbar} e^{\frac{i}{\hbar} \delta t \left[\frac{1}{2} m \left(\frac{\delta x_n}{\delta t} \right)^2 - \frac{1}{2} m \left(\frac{\delta x_n}{\delta t} \right)^2 + p_n \frac{\delta x_n}{\delta t} - \frac{p_n^2}{2m} - V(x_n) \right]}$$

$$= \int_{-\infty}^{\infty} \frac{dp_n}{2\pi\hbar} e^{\frac{i}{\hbar} \delta t \left[\frac{1}{2} m \left(\frac{\delta x_n}{\delta t} \right)^2 - V(x_n) - \frac{1}{2m} \left(p_n - m \frac{\delta x_n}{\delta t} \right)^2 \right]}$$

$$= e^{\frac{i}{\hbar} \delta t \left[\frac{1}{2} m \left(\frac{\delta x_n}{\delta t} \right)^2 - V(x_n) \right]} \int_{-\infty}^{\infty} \frac{dp_n}{2\pi\hbar} e^{\frac{i}{\hbar} \frac{\delta t}{2m} \left(p_n - m \frac{\delta x_n}{\delta t} \right)^2}.$$

(13.26)

The integral here is oscillatory and, therefore, not convergent. In order to make it well defined we must either introduce a convergence factor or analytically continue the integral to make it convergent. We will use the second approach. Let

$$\frac{\delta t}{2m\hbar} \equiv -i\tau \tag{13.27}$$

and

$$a \equiv m\frac{\delta x_n}{\delta t}. \tag{13.28}$$

Then

$$I \equiv \int_{-\infty}^{\infty} \frac{dp_n}{2\pi\hbar} e^{-\frac{i}{\hbar}\frac{\delta t}{2m}\left(p_n - m\frac{\delta x_n}{\delta t}\right)^2} = \int_{-\infty}^{\infty} \frac{dp_n}{2\pi\hbar} e^{-\tau(p_n - a)^2}. \tag{13.29}$$

Shifting the origin of the integral with $\xi = p_n - a$ yields

$$I = \int_{-\infty}^{\infty} \frac{d\xi}{2\pi\hbar} e^{-\xi^2} = \frac{1}{2\pi\hbar}\sqrt{\frac{\pi}{\tau}} = \sqrt{\frac{1}{4\pi\hbar^2\tau}} = \sqrt{\frac{2m\hbar}{4\pi\hbar^2 i\delta t}} = \sqrt{\frac{m}{i2\pi\hbar\delta t}}. \tag{13.30}$$

The propagator can therefore be written as

$$K(x,t;x',t') = \lim_{N\to\infty}\left(\frac{m}{i2\pi\hbar\delta t}\right)^{\frac{N}{2}} \int_{-\infty}^{\infty} \prod_{n=1}^{N-1} dx_n e^{\frac{i}{\hbar}\delta t \sum_{n=0}^{N-1}\left[\frac{1}{2}m\left(\frac{\delta x_n}{\delta t}\right)^2 - V(x_n)\right]}. \tag{13.31}$$

After integrating over the momenta, we are left with an integral depending only on positions and time. Note that the constant resulting from the momentum integrals is badly divergent as $N \to \infty$ (it goes like $N^{N/2}$). However, practical applications of the path integral formalism will generally involve normalized rations where the coefficient of the path integral is canceled. The value of the constant is unimportant so it is conventional to simply denote it with some symbol such as $\mathcal{N}$.

Taking the continuum limit, $N \to \infty$ and $\delta t \to 0$, the argument of the exponential again becomes an integral over time. We can formally express this limit as

$$K(x,t;x',t') = \mathcal{N} \int \mathcal{D}x \, e^{\frac{i}{\hbar}\int_{t'}^{t} dt\left[\frac{1}{2}m\dot{x}^2 - V(x)\right]}. \tag{13.32}$$

We can now identify the lagrangian as

$$L(\dot{x}, x) = \frac{1}{2}m\dot{x}^2 - V(x) \tag{13.33}$$

and the action as

$$S(x,t;x',t') = \int_{t'}^{t} dt L(\dot{x}, x). \tag{13.34}$$

The propagator is given by an integral overall possible paths connecting q' and t' to q and t, weighted by $e^{\frac{i}{\hbar}S}$ for each path. In situations where the action changes

rapidly as the path is varied, the exponential oscillates rapidly, contributing little to the path integral. However, major contributions to the path integral occur in regions where the action is slowly varying. Such regions occur near stationary points which are found by locating the extrema of the action. Indeed, the requirement that the action be a minimum leads directly to the classical equations of motion as derived from the lagrangian. In the path integral formulation of quantum mechanics, quantum mechanical effects arise from paths other than the classical trajectory. Note that the path integral form of the propagator contains no quantum operators. All of the integrals are over c-numbers.

13.1   The Propagator for a Free Particle

Unfortunately, path integral for the propagator can only be evaluated analytically in only a limited number of situations. Because the integrals are actually defined as the limit of integrations performed on discrete grid, it is generally possible to evaluate the propagator numerically.

Here we will consider the simplest possible situation of a free, one-dimensional particle where $V(x) = 0$. In this case, (13.31) can be written as

$$K(x,t;x',t') = \lim_{N \to \infty} \left(\frac{m}{i 2\pi\hbar\delta t} \right)^{\frac{N}{2}} \int_{-\infty}^{\infty} \prod_{n=1}^{N-1} dx_n e^{\frac{i}{\hbar} \frac{m}{2\delta t} \sum_{n=0}^{N-1} (x_{n+1} - x_n)^2}. \tag{13.35}$$

We can now perform the integrals over the coordinates one at a time. First note that if we write the exponential of the sum of terms as a product of exponentials, there will be two factors that contain x_1. The integral over x_1 can then be written as

$$\frac{m}{i 2\pi\hbar\delta t} \int_{-\infty}^{\infty} dx_1 e^{\frac{i}{\hbar} \frac{m}{2\delta t} (x_2 - x_1)^2} e^{\frac{i}{\hbar} \frac{m}{2\delta t} (x_1 - x_0)^2}$$

$$= \frac{m}{i 2\pi\hbar\delta t} \int_{-\infty}^{\infty} dx_1 e^{\frac{im}{2\hbar\delta t} \left(x_2^2 - 2x_2 x_1 + x_1^2 + x_1^2 - 2x_1 x_0 + x_0^2 \right)}$$

$$= \frac{m}{i 2\pi\hbar\delta t} \int_{-\infty}^{\infty} dx_1 e^{\frac{im}{\hbar\delta t} \left(x_1^2 - (x_2 + x_0)x_1 + \frac{1}{2}(x_2^2 + x_0^2) \right)}$$

$$= \frac{m}{i 2\pi\hbar\delta t} \int_{-\infty}^{\infty} dx_1 e^{\frac{im}{\hbar\delta t} \left(x_1^2 - (x_2 + x_0)x_1 + \frac{1}{4}(x_2 + x_0)^2 - \frac{1}{4}(x_2 + x_0)^2 + \frac{1}{2}(x_2^2 + x_0^2) \right)}$$

$$= \frac{m}{i 2\pi\hbar\delta t} \int_{-\infty}^{\infty} dx_1 e^{\frac{im}{\hbar\delta t} \left[\left(x_1 - \frac{1}{2}(x_2 + x_0) \right)^2 + \frac{1}{4}(x_2 - x_0)^2 \right]}$$

$$= \frac{m}{i 2\pi\hbar\delta t} e^{\frac{im}{4\hbar\delta t} (x_2 - x_0)^2} \int_{-\infty}^{\infty} dx_1 e^{\frac{im}{\hbar\delta t} \left(x_1 - \frac{1}{2}(x_2 + x_0) \right)^2}$$

$$= \frac{m}{i 2\pi\hbar\delta t} e^{\frac{im}{4\hbar\delta t} (x_2 - x_0)^2} \sqrt{\frac{i\pi\hbar\delta t}{m}} = \sqrt{\frac{m}{i 4\pi\hbar\delta t}} e^{\frac{im}{4\hbar\delta t} (x_2 - x_0)^2}, \tag{13.36}$$

We can now perform the integral over x_2 by multiplying this result by the remaining exponential factor containing x_2 and integrating to give

$$\sqrt{\frac{m}{i2\pi\hbar\delta t}} \int_{-\infty}^{\infty} dx_2 e^{\frac{im}{2\hbar\delta t}(x_3-x_2)^2} \sqrt{\frac{m}{i4\pi\hbar\delta t}} e^{\frac{im}{4\hbar\delta t}(x_2-x_0)^2}$$

$$= \sqrt{\frac{m}{i2\pi\hbar\delta t}} \sqrt{\frac{m}{i4\pi\hbar\delta t}} \int_{-\infty}^{\infty} dx_2 e^{\frac{im}{2\hbar\delta t}\left[(x_3-x_2)^2+\frac{1}{2}(x_2-x_0)^2\right]}$$

$$= \sqrt{\frac{m}{i2\pi\hbar\delta t}} \sqrt{\frac{m}{i4\pi\hbar\delta t}} \int_{-\infty}^{\infty} dx_2 e^{\frac{im}{2\hbar\delta t}\left[\frac{3}{2}x_2^2-(2x_3+x_0)x_2+x_3^2+\frac{1}{2}x_0^2\right]}$$

$$= \sqrt{\frac{m}{i2\pi\hbar\delta t}} \sqrt{\frac{m}{i4\pi\hbar\delta t}} \int_{-\infty}^{\infty} dx_2 e^{\frac{im}{2\hbar\delta t}\left[\frac{3}{2}\left(x_2^2-\frac{1}{3}(2x_3+x_0)\right)^2+\frac{1}{3}(x_3-x_0)^2\right]}$$

$$= \sqrt{\frac{m}{i2\pi\hbar\delta t}} \sqrt{\frac{m}{i4\pi\hbar\delta t}} e^{\frac{im}{6\hbar\delta t}(x_3-x_0)^2} \int_{-\infty}^{\infty} dx_2 e^{\frac{i3m}{4\hbar\delta t}\left(x_2-\frac{1}{3}(2x_3+x_0)\right)^2}$$

$$= \sqrt{\frac{m}{i2\pi\hbar\delta t}} \sqrt{\frac{m}{i4\pi\hbar\delta t}} e^{\frac{im}{6\hbar\delta t}(x_3-x_0)^2} \sqrt{\frac{i4\pi\hbar\delta t}{3m}}$$

$$= \sqrt{\frac{m}{i6\pi\hbar\delta t}} e^{\frac{im}{6\hbar\delta t}(x_3-x_0)^2} . \tag{13.37}$$

Comparing the results of the first two integrals, we can determine by induction that

$$K(x,t;x',t') = \lim_{N\to\infty} \sqrt{\frac{m}{i2N\pi\hbar\delta t}} e^{\frac{im}{2N\hbar\delta t}(x_N-x_0)^2}$$

$$= \lim_{N\to\infty} \sqrt{\frac{m}{i2\pi\hbar(t-t')}} e^{\frac{im}{2\hbar}\frac{(x-x')^2}{t-t'}}$$

$$= \sqrt{\frac{m}{i2\pi\hbar(t-t')}} e^{\frac{im}{2\hbar}\frac{(x-x')^2}{t-t'}} . \tag{13.38}$$

This solution is appropriate for stationary state solutions. However, if we want to obtain the propagator for the retarded propagator, it is necessary to require that $K(x,t;x',t') = 0$ for $t < t'$. This can be done by multiplying the above expression by $\theta(t-t')$. The retarded propagator is then

$$K_R(x,t;x',t') = \sqrt{\frac{m}{i2\pi\hbar(t-t')}} e^{\frac{im}{2\hbar}\frac{(x-x')^2}{t-t'}} \theta(t-t') . \tag{13.39}$$

Chapter 14

Time-Dependent Perturbation Theory

We now want to consider solutions to the Schrödinger equation in the case where the hamiltonian is time dependent. With the ultimate objective of obtaining time-dependent perturbation theory, we will choose the hamiltonian to be of the form

$$\hat{H}(t) = \hat{H}_0 + \hat{V}(t) \,, \tag{14.1}$$

where $\hat{H}_0$ is time independent and $V(t)$ is some time-dependent interaction. The Schrödinger equation with this interaction is then

$$\hat{H}(t) \left| \Psi(t) \right\rangle = i\hbar \frac{\partial}{\partial t} \left| \Psi(t) \right\rangle \,. \tag{14.2}$$

The solution of this problem is generally more difficult that the solution for a time-independent hamiltonian and requires a more sophisticated approach. Our first step is the obtain an expression for the time-evolution operator defined such that

$$\left| \Psi(t) \right\rangle = \hat{U}(t, t') \left| \Psi(t') \right\rangle \,. \tag{14.3}$$

Using this to replace the state vectors in (14.2) gives

$$\hat{H}(t)\hat{U}(t, t') \left| \Psi(t') \right\rangle = i\hbar \frac{\partial}{\partial t} \hat{U}(t, t') \left| \Psi(t') \right\rangle \,. \tag{14.4}$$

Since this must be true for any state vector which is a solution to the Schrödinger equation, we obtain a differential equation for the time-evolution operator

$$\hat{H}(t)\hat{U}(t, t') = i\hbar \frac{\partial}{\partial t} \hat{U}(t, t') \,. \tag{14.5}$$

Solution of this equation is equivalent to solution of the Schrödinger equation. Now, integrate (14.5) from t' to t

$$\int_{t'}^{t} d\tau \hat{H}(\tau)\hat{U}(\tau, t') = i\hbar \int_{t'}^{t} d\tau \frac{\partial}{\partial \tau} \hat{U}(\tau, t') = i\hbar \left(\hat{U}(t, t') - \hat{U}(t', t') \right) \tag{14.6}$$

If we now impose the boundary condition

$$\hat{U}(t', t') = \hat{1},$$ (14.7)

then we can solve (14.6) to give

$$\hat{U}(t, t') = \hat{1} - \frac{i}{\hbar} \int_{t'}^{t} d\tau \hat{H}(\tau) \hat{U}(\tau, t').$$ (14.8)

We have succeeded in turning the differential equation (14.5) into an integral equation. This may not seem like a positive step, but we will soon see that this can facilitate to construction of a perturbation expansion. Note that if the hamiltonian is time-independent, the evolution equation (8.70) satisfies this equation.

Since the probability must be normalized for all times

$$1 = \langle \Psi(t) | \Psi(t) \rangle = \langle \Psi(t') | \hat{U}^\dagger(t, t') \hat{U}(t, t') | \Psi(t) \rangle.$$ (14.9)

This means that

$$\hat{U}^\dagger(t, t') = \hat{U}^{-1}(t, t').$$ (14.10)

So the time-evolution operator is unitary.

For the case of the time-independent hamitlonian the composition property

$$\hat{U}(t, t'') \hat{U}(t'', t') = \hat{U}(t, t')$$ (14.11)

is trivially satisfied. To show that this is still the case for a time-dependent hamiltonian, note that

$$i\hbar \frac{\partial}{\partial t} \hat{U}(t, t'') \hat{U}(t'', t') = \left(i\hbar \frac{\partial}{\partial t} \hat{U}(t, t'') \right) \hat{U}(t'', t') = \hat{H}(t) \hat{U}(t, t'') \hat{U}(t'', t').$$ (14.12)

So $\hat{U}(t, t'') \hat{U}(t'', t')$ satisfies the same differential equation as $\hat{U}(t, t')$. The composition property will then be true for any t if it is true some particular t. If we choose $t = t''$, then

$$\hat{U}(t'', t'') \hat{U}(t'', t') = \hat{U}(t'', t').$$ (14.13)

This finishes the proof of the composition property.

Using the composition property,

$$\hat{U}(t, t') \hat{U}(t', t) = \hat{U}(t, t) = \hat{1}.$$ (14.14)

This implies that

$$\hat{U}(t, t') = \hat{U}^\dagger(t', t).$$ (14.15)

Now, if we take the hermitian conjugate of (14.8) and exchanging the variables t and t', we obtain

$$\hat{U}^\dagger(t', t) = \hat{1} + \frac{i}{\hbar} \int_{t}^{t'} d\tau \hat{U}^\dagger(\tau, t) \hat{H}(\tau).$$ (14.16)

Applying (14.15) to this gives

$$\hat{U}(t, t') = \hat{1} + \frac{i}{\hbar} \int_{t}^{t'} d\tau \hat{U}(t, \tau) \hat{H}(\tau) = \hat{1} - \frac{i}{\hbar} \int_{t'}^{t} d\tau \hat{U}(t, \tau) \hat{H}(\tau).$$ (14.17)

232

14.1 The Interaction Representation

In Section 8.3 we introduced the Heisenberg representation of of quantum mechanics where the Heisenberg states are defined as

$$|\Psi_H\rangle = |\Psi(0)\rangle = e^{\frac{i}{\hbar}\hat{H}t}|\Psi(t)\rangle = \hat{U}^\dagger(t,0)|\Psi(t)\rangle \ . \tag{14.18}$$

In this representation, the states are time independent and the operators

$$\hat{\mathcal{O}}_H(t) = e^{\frac{i}{\hbar}\hat{H}t}\hat{\mathcal{O}}e^{-\frac{i}{\hbar}\hat{H}t} \tag{14.19}$$

are time dependent even if the original Schrödinger operators are time independent. This transformation between the Schrödinger and Heisenberg representations is unitary which guarantees the matrix elements of will be the same in either representation.

It is possible to produce other representations that are defined as unitary transformations of the states and operators. One representation that is particularly useful for construction time-dependent perturbation theory is the *Interaction Representation* which is intermediate between the Schrödinger and Heisenberg represetations. The interaction representation state is defined in terms of the Schrödinger state as

$$|\Psi_I(t)\rangle = e^{\frac{i}{\hbar}\hat{H}_0 t}|\Psi(t)\rangle \ . \tag{14.20}$$

This transformation can be inverted to give

$$|\Psi(t)\rangle = e^{-\frac{i}{\hbar}\hat{H}_0 t}|\Psi_I(t)\rangle \ . \tag{14.21}$$

The operators in this representation are defined as

$$\hat{\mathcal{O}}_I(t) = e^{\frac{i}{\hbar}\hat{H}_0 t}\hat{\mathcal{O}}e^{-\frac{i}{\hbar}\hat{H}_0 t} \ . \tag{14.22}$$

Note that in this representation both the states and operators are time dependent.

We can find the time evolution of the interaction state by considering

$$\begin{aligned}
i\hbar\frac{\partial}{\partial t}|\Psi_I(t)\rangle &= i\hbar\frac{\partial}{\partial t}e^{\frac{i}{\hbar}\hat{H}_0 t}|\Psi(t)\rangle = -\hat{H}_0 e^{\frac{i}{\hbar}\hat{H}_0 t}|\Psi(t)\rangle + e^{\frac{i}{\hbar}\hat{H}_0 t}i\hbar\frac{\partial}{\partial t}|\Psi(t)\rangle \\
&= -\hat{H}_0 e^{\frac{i}{\hbar}\hat{H}_0 t}|\Psi(t)\rangle + e^{\frac{i}{\hbar}\hat{H}_0 t}\hat{H}|\Psi(t)\rangle = e^{\frac{i}{\hbar}\hat{H}_0 t}(\hat{H} - \hat{H}_0)|\Psi(t)\rangle \\
&= e^{\frac{i}{\hbar}\hat{H}_0 t}\hat{V}(t)e^{-\frac{i}{\hbar}\hat{H}_0 t}e^{\frac{i}{\hbar}\hat{H}_0 t}|\Psi(t)\rangle \ . \tag{14.23}
\end{aligned}$$

Using the definitions of the interaction states and operators, this becomes

$$i\hbar\frac{\partial}{\partial t}|\Psi_I(t)\rangle = H_1(t)|\Psi_I(t)\rangle \ , \tag{14.24}$$

where

$$H_1(t) = e^{\frac{i}{\hbar}\hat{H}_0 t}\hat{V}(t)e^{-\frac{i}{\hbar}\hat{H}_0 t} \ . \tag{14.25}$$

The time dependence of the operators can be found by considering

$$
\begin{aligned}
i\hbar \frac{\partial}{\partial t}\hat{\mathcal{O}}_I(t) &= i\hbar \frac{\partial}{\partial t} e^{\frac{i}{\hbar}\hat{H}_0 t}\hat{\mathcal{O}} e^{-\frac{i}{\hbar}\hat{H}_0 t} \\
&= -\hat{H}_0 e^{\frac{i}{\hbar}\hat{H}_0 t}\hat{\mathcal{O}} e^{-\frac{i}{\hbar}\hat{H}_0 t} + e^{\frac{i}{\hbar}\hat{H}_0 t}\hat{\mathcal{O}} e^{-\frac{i}{\hbar}\hat{H}_0 t}\hat{H}_0 + e^{\frac{i}{\hbar}\hat{H}_0 t}\frac{\partial}{\partial t}\hat{\mathcal{O}} e^{-\frac{i}{\hbar}\hat{H}_0 t} \quad (14.26)
\end{aligned}
$$

Using the definition of the interaction representation operators, this becomes

$$
i\hbar \frac{\partial}{\partial t}\hat{\mathcal{O}}_I(t) = [\hat{\mathcal{O}}_I(t), \hat{H}_0] + e^{\frac{i}{\hbar}\hat{H}_0 t}\frac{\partial}{\partial t}\hat{\mathcal{O}} e^{-\frac{i}{\hbar}\hat{H}_0 t} . \tag{14.27}
$$

We now want to find the interaction representation of the time evolution operator such that

$$
|\Psi_I(t)\rangle = \hat{U}_I(t, t') |\Psi_I(t')\rangle . \tag{14.28}
$$

Using the definition of the interaction representation states and the time evolution of the Schrödinger state gives

$$
|\Psi_I(t)\rangle = e^{\frac{i}{\hbar}\hat{H}_0 t}|\Psi(t)\rangle = e^{\frac{i}{\hbar}\hat{H}_0 t}\hat{U}(t, t')|\Psi(t')\rangle = e^{\frac{i}{\hbar}\hat{H}_0 t}\hat{U}(t, t')e^{-\frac{i}{\hbar}\hat{H}_0 t'}|\Psi_I(t')\rangle . \tag{14.29}
$$

This then yields

$$
\hat{U}_I(t, t') = e^{\frac{i}{\hbar}\hat{H}_0 t}\hat{U}(t, t')e^{-\frac{i}{\hbar}\hat{H}_0 t'} . \tag{14.30}
$$

A differential equation for the time evolution operator in the interaction representation can be obtained by considering

$$
\begin{aligned}
i\hbar \frac{\partial}{\partial t}\hat{U}_I(t, t') &= i\hbar \frac{\partial}{\partial t} e^{\frac{i}{\hbar}\hat{H}_0 t}\hat{U}(t, t')e^{-\frac{i}{\hbar}\hat{H}_0 t'} \\
&= -\hat{H}_0 e^{\frac{i}{\hbar}\hat{H}_0 t}\hat{U}(t, t')e^{-\frac{i}{\hbar}\hat{H}_0 t'} + e^{\frac{i}{\hbar}\hat{H}_0 t}i\hbar \frac{\partial}{\partial t}\hat{U}(t, t')e^{-\frac{i}{\hbar}\hat{H}_0 t'} \\
&= -\hat{H}_0 e^{\frac{i}{\hbar}\hat{H}_0 t}\hat{U}(t, t')e^{-\frac{i}{\hbar}\hat{H}_0 t'} + e^{\frac{i}{\hbar}\hat{H}_0 t}\hat{H}(t)\hat{U}(t, t')e^{-\frac{i}{\hbar}\hat{H}_0 t'} \\
&= e^{\frac{i}{\hbar}\hat{H}_0 t}\left(\hat{H}(t) - \hat{H}_0\right)\hat{U}(t, t')e^{-\frac{i}{\hbar}\hat{H}_0 t'} \\
&= e^{\frac{i}{\hbar}\hat{H}_0 t}\hat{V}(t)\hat{U}(t, t')e^{-\frac{i}{\hbar}\hat{H}_0 t'} = e^{\frac{i}{\hbar}\hat{H}_0 t}\hat{V}(t)e^{-\frac{i}{\hbar}\hat{H}_0 t'}e^{\frac{i}{\hbar}\hat{H}_0 t}\hat{U}(t, t')e^{-\frac{i}{\hbar}\hat{H}_0 t'} .
\end{aligned}
$$
$$\tag{14.31}$$

Using the definitions of the interaction representation operators this yields

$$
i\hbar \frac{\partial}{\partial t}\hat{U}_I(t, t') = \hat{H}_1(t)\hat{U}_I(t, t') . \tag{14.32}
$$

This can be integrated in the same way as we used for the Schrödinger time evolution operator to give

$$
\hat{U}_I(t, t') = \hat{1} - \frac{i}{\hbar}\int_{t'}^{t} d\tau \hat{H}_1(\tau)\hat{U}_I(\tau, t') . \tag{14.33}
$$

Solving this integral equation is equivalent to solving the Schrödinger equation and is at least as difficult. However, we can construct a perturbation expansion using this equation. To see how this works, consider the first iteration of (14.33) where the the time-evolution operator on the right-hand side is replaced by the using (14.33) itself. This gives

$$
\begin{aligned}
\hat{U}_I(t,t') &= \hat{1} - \frac{i}{\hbar} \int_{t'}^{t} d\tau_1 \hat{H}_1(\tau_1) \left(\hat{1} - \frac{i}{\hbar} \int_{t'}^{\tau_1} d\tau_2 \hat{H}_1(\tau_2) \hat{U}_I(\tau_2,t') \right) \\
&= \hat{1} - \frac{i}{\hbar} \int_{t'}^{t} d\tau_1 \hat{H}_1(\tau_1) + \left(\frac{i}{\hbar} \right)^2 \int_{t'}^{t} d\tau_1 \int_{t'}^{\tau_1} d\tau_2 \hat{H}_1(\tau_1) \hat{H}_1(\tau_2) \hat{U}_I(\tau_2,t') .
\end{aligned}
$$

(14.34)

This can be repeated an arbitrary number of times by simply replacing the time-evolution operator on the right-hand side of each iteration by (14.33). This leads to a series solution for the time-evolution operator in the interaction representation given by

$$
\hat{U}_I(t,t') = \hat{1} - \frac{i}{\hbar} \int_{t'}^{t} d\tau_1 \hat{H}_1(\tau_1) + \left(\frac{i}{\hbar} \right)^2 \int_{t'}^{t} d\tau_1 \int_{t'}^{\tau_1} d\tau_2 \hat{H}_1(\tau_1) \hat{H}_1(\tau_2) + \cdots . \quad (14.35)
$$

Note that each successive term includes one more factor of the perturbing hamiltonian $\hat{H}_1(t)$. We now have a perturbation series for the time-evolution operator in the interaction representation.

We can use this to construct a series in the Schrödinger representation by using the inverse of (14.30)

$$
\hat{U}(t,t') = e^{-\frac{i}{\hbar}\hat{H}_0 t} \hat{U}_I(t,t') e^{\frac{i}{\hbar}\hat{H}_0 t'} \tag{14.36}
$$

with (14.25) and (14.35) to give

$$
\hat{U}(t,t') = e^{-\frac{i}{\hbar}\hat{H}_0(t-t')} - \frac{i}{\hbar} e^{-\frac{i}{\hbar}\hat{H}_0 t} \int_{t'}^{t} d\tau_1 e^{\frac{i}{\hbar}\hat{H}_0\tau_1} \hat{V}(\tau_1) e^{-\frac{i}{\hbar}\hat{H}_0\tau_1} e^{\frac{i}{\hbar}\hat{H}_0 t'}
$$
$$
+ \left(\frac{i}{\hbar} \right)^2 e^{-\frac{i}{\hbar}\hat{H}_0 t} \int_{t'}^{t} d\tau_1 \int_{t'}^{\tau_1} d\tau_2 e^{\frac{i}{\hbar}\hat{H}_0\tau_1} \hat{V}(\tau_1) e^{-\frac{i}{\hbar}\hat{H}_0\tau_1} e^{\frac{i}{\hbar}\hat{H}_0\tau_2} \hat{V}(\tau_2) e^{-\frac{i}{\hbar}\hat{H}_0\tau_2} e^{\frac{i}{\hbar}\hat{H}_0 t'} + \cdots
$$

(14.37)

If we define

$$
\hat{U}_0(t,t') \equiv e^{-\frac{i}{\hbar}\hat{H}_0(t-t')} , \tag{14.38}
$$

then

$$
\begin{aligned}
\hat{U}(t,t') &= \hat{U}_0(t,t') - \frac{i}{\hbar} \int_{t'}^{t} d\tau_1 \hat{U}_0(t,\tau_1) \hat{V}(\tau_1) \hat{U}_0(\tau_1,t') \\
&\quad + \left(\frac{i}{\hbar} \right)^2 \int_{t'}^{t} d\tau_1 \int_{t'}^{\tau_1} d\tau_2 \hat{U}_0(t,\tau_1) \hat{V}(\tau_1) \hat{U}_0(\tau_1,\tau_2) \hat{V}(\tau_2) \hat{U}_0(\tau_2,t') + \cdots
\end{aligned}
$$

(14.39)

14.2 Transition Matrix Elements

Our objective is to find time-dependent transition matrix elements that describe the changes in a quantum mechanical system resulting from the introduction of the interaction. The basic assumption is that the interaction is turned on at some time t'. Before this time, the wave functions are solutions to the unperturbed hamiltonian and the wave function at time t' provides the initial conditions for the solution of the time-dependent equations. It is sufficient to assume that the system is in some eigenstate of $\hat{H}_0$ at t', since any more general solution can be written as a linear combination of such states. The wave function at some later time t is then given by

$$|\Psi_\alpha(t)\rangle = \hat{U}(t,t') |\Phi_\alpha(t')\rangle \,, \tag{14.40}$$

where

$$|\Phi_\alpha(t')\rangle = e^{-\frac{i}{\hbar}E_\alpha^{(0)}t'} |\phi_\alpha\rangle \tag{14.41}$$

Since we want the evolution to occur in the positive time direction, we need to build in the boundary conditions by requiring that

$$\hat{U}(t,t') \to \hat{U}(t,t')\theta(t-t') \,. \tag{14.42}$$

If we assume that the interaction is turned off at time t, the probability amplitude for the system to be in the eigenstate β after this time is then given by

$$S_{\beta\alpha}(t,t') = \langle \Phi_\beta(t) | \Psi_\alpha(t)\rangle = \langle \Phi_\beta(t)| \hat{U}(t,t') |\Phi_\alpha(t')\rangle \,. \tag{14.43}$$

The corresponding transition probability is then

$$P_{\beta\alpha}(t,t') = \left| \langle \Phi_\beta(t)| \hat{U}(t,t') |\Phi_\alpha(t)\rangle \right|^2 \,. \tag{14.44}$$

The transition amplitude can then be expanded in a series using (14.39) such that

$$S_{\beta\alpha}(t,t') = \sum_{n=0}^\infty S_{\beta\alpha}^{(n)}(t,t') \,. \tag{14.45}$$

The boundary condition is imposed in this series by requiring that

$$\hat{U}_0(t,t') = e^{-\frac{i}{\hbar}\hat{H}_0(t-t')}\theta(t-t') \,. \tag{14.46}$$

The leading order contribution to the series is

$$
\begin{aligned}
S_{\beta\alpha}^{(0)}(t,t') &= \langle \Phi_\beta(t)| \hat{U}_0(t,t') |\Phi_\alpha(t')\rangle = \langle \phi_\beta| e^{\frac{i}{\hbar}E_\beta^{(0)}t} e^{-\frac{i}{\hbar}\hat{H}_0(t-t')}\theta(t-t') e^{-\frac{i}{\hbar}E_\alpha^{(0)}t'} |\phi_\alpha\rangle \\
&= \delta_{\alpha\beta}\theta(t-t') \,,
\end{aligned}
\tag{14.47}
$$

which simply corresponds to the system remaining in its original state. The first order contribution is given by

$$
\begin{aligned}
S_{\beta\alpha}^{(1)}(t,t') &= -\frac{i}{\hbar} \int_{t'}^{t} d\tau_1 \, \langle \Phi_\beta(t)| \, \hat{U}_0(t,\tau_1)\hat{V}(\tau_1)\hat{U}_0(\tau_1,t') \, |\Phi_\alpha(t')\rangle \\
&= -\frac{i}{\hbar} \sum_\kappa \sum_\lambda \int_{t'}^{t} d\tau_1 \, \langle \phi_\beta| \, e^{\frac{i}{\hbar}E_\beta^{(0)}t}\hat{U}_0(t,\tau_1) \, |\phi_\kappa\rangle \, \langle \phi_\kappa| \, \hat{V}(\tau_1)) \, |\phi_\lambda\rangle \\
&\quad \langle \phi_\lambda| \, \hat{U}_0(\tau_1,t')e^{-\frac{i}{\hbar}E_\alpha^{(0)}t'} \, |\phi_\alpha\rangle \\
&= -\frac{i}{\hbar} \sum_\kappa \sum_\lambda \int_{t'}^{t} d\tau_1 e^{\frac{i}{\hbar}E_\alpha^{(0)}\tau_1}\delta_{\beta\kappa}\theta(t-\tau_1) \, \langle \phi_\kappa| \, \hat{V}(\tau_1)) \, |\phi_\lambda\rangle \\
&\quad \times e^{-\frac{i}{\hbar}E_\alpha^{(0)}t'}\delta_{\lambda\alpha}\theta(\tau_1-t') \\
&= -\frac{i}{\hbar} \int_{t'}^{t} d\tau_1 e^{\frac{i}{\hbar}E_\beta^{(0)}\tau_1}V_{\beta\alpha}(\tau_1)e^{-\frac{i}{\hbar}E_\alpha^{(0)}\tau_1}\theta(t-\tau_1)\theta(\tau_1-t')\,, \qquad (14.48)
\end{aligned}
$$

where

$$
V_{\beta\alpha}(\tau_1) = \langle \phi_\beta| \, \hat{V}(\tau_1) \, |\phi_\alpha\rangle \,. \qquad (14.49)
$$

The second-order contribution is

$$
\begin{aligned}
S_{\beta\alpha}^{(2)}(t,t') &= \left(-\frac{i}{\hbar}\right)^2 \sum_\gamma \int_{t'}^{t} d\tau_1 \int_{t'}^{t} d\tau_2 e^{\frac{i}{\hbar}E_\beta^{(0)}\tau_1}V_{\beta\gamma}(\tau_1)e^{-\frac{i}{\hbar}E_\gamma^{(0)}(\tau_1-\tau_2)} \\
&\quad \times V_{\gamma\alpha}(\tau_2)e^{-\frac{i}{\hbar}E_\alpha^{(0)}\tau_2}\theta(t-\tau_1)\theta(\tau_1-\tau_2)\theta(\tau_2-t')\,, \qquad (14.50)
\end{aligned}
$$

and higher-order contributions will follow this pattern.

This series then has a simple physical interpretation that is illustrated by the diagrams in Fig. 14.1. In these diagrams the time increases from bottom to top. The is zeroth order contribution is represented by diagram (a) and the system propagates in the state *alpha* from time t' to time t. The propagation is represented by a line. Diagram (b) illustrates the first-order correction. Here the system starts in state α at t' and propagates to time τ_1 where it experiences the interaction, represented by the cross, and is changed to state *beta* and then propagates in this state to time t. Diagram (c) represents the second-order contribution. Starting in state *alpha* at t', the system propagates to τ_2 where an interaction changes it to state *gamma*. This propagates to τ_1 where an interaction changes it to state β where it propagates to t. These are called *time-ordered diagrams*.

14.3 Time-Independent Interactions

Consider the case where $\hat{V}(t) = \hat{V}_0$. We want to calculate the transition probability amplitude from state α to state $\beta \neq \alpha$ to first order. Assume that the interaction is
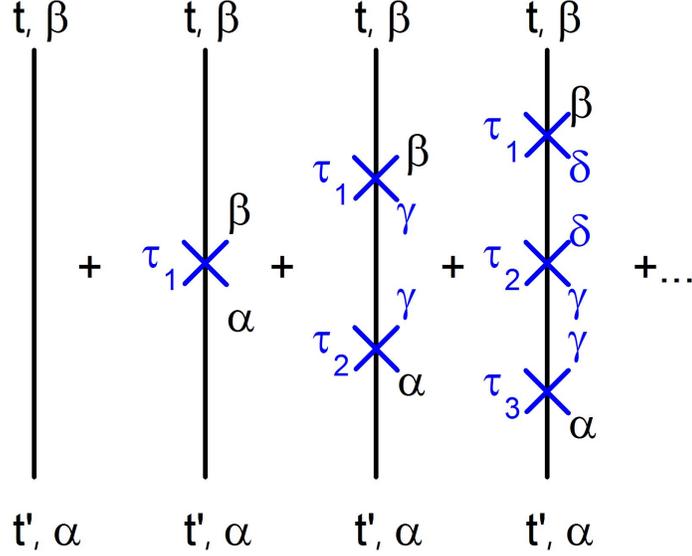
Figure 14.1: Time-ordered diagrams representing contributions to the transition matrix elements.

turned on at $t' = -T/2$ and turned off at $t = T/2$. The probability amplitude to first order is then given by

$$
\begin{aligned}
S_{\beta\alpha}^{(1)}(T/2, -T/2) &= -\frac{i}{\hbar} \int_{-T/2}^{T/2} d\tau_1 e^{\frac{i}{\hbar} E_\beta^{(0)} \tau_1} \langle \phi_\beta | \hat{V}_0 | \phi_\alpha \rangle e^{-\frac{i}{\hbar} E_\alpha^{(0)} \tau_1} \\
&= -\frac{i}{\hbar} \langle \phi_\beta | \hat{V}_0 | \phi_\alpha \rangle \int_{-T/2}^{T/2} d\tau_1 e^{\frac{i}{\hbar} (E_\beta^{(0)} - E_\alpha^{(0)}) \tau_1} \\
&= -\frac{2i}{\hbar \omega_{\beta\alpha}} \langle \phi_\beta | \hat{V}_0 | \phi_\alpha \rangle \sin \frac{\omega_{\beta\alpha} T}{2} .
\end{aligned}
$$

$$(14.51)$$

The transition probability to this order is

$$
P_{\beta\alpha}(T/2, -T/2) = \frac{4}{\hbar^2 \omega_{\beta\alpha}^2} \left| \langle \phi_\beta | \hat{V}_0 | \phi_\alpha \rangle \right|^2 \sin^2 \frac{\omega_{\beta\alpha} T}{2} . \tag{14.52}
$$

The total transition probability for all states β is given by

$$
P_\alpha \equiv \sum_\beta P_{\beta\alpha}(T/2, -T/2) = \sum_\beta \frac{4}{\hbar^2 \omega_{\beta\alpha}^2} \left| \langle \phi_\beta | \hat{V}_0 | \phi_\alpha \rangle \right|^2 \sin^2 \frac{\omega_{\beta\alpha} T}{2} . \tag{14.53}
$$

We can estimate the effect of the length of the time interval T on this by performing a integral over all frequencies of

$$
\int_{-\infty}^{\infty} d\omega \sin^2 \frac{\omega T}{2} = \frac{\pi}{2} T . \tag{14.54}
$$

238

The implies that the total transition probability increases linearly with T. For this reason, it is convenient to define the transition rate as

$$\Gamma_{\beta\alpha}(T) = \frac{1}{T}P_{\beta\alpha}(T/2, -T/2) = \frac{4}{\hbar^2\omega_{\beta\alpha}^2 T}\left|\langle\phi_\beta|\hat{V}_0|\phi_\alpha\rangle\right|^2 \sin^2\frac{\omega_{\beta\alpha}T}{2}. \qquad (14.55)$$

14.4   Gaussian Time Dependence

Now consider a potential with Gaussian time dependence given by

$$\hat{V}(t) = \left(\frac{2}{\pi}\right)^{\frac{1}{4}}\hat{V}_0 e^{-\frac{t^2}{T^2}}. \qquad (14.56)$$

Here, T gives the width of the potential in time. We can, therefore, calculate the transition matrix element from (14.48) by integrating over all times to give

$$
\begin{aligned}
S_{\beta\alpha}^{(1)}(\infty, -\infty) &= -\frac{i}{\hbar}\int_{-\infty}^{\infty} d\tau_1 e^{\frac{i}{\hbar}E_\beta^{(0)}\tau_1}\left(\frac{2}{\pi}\right)^{\frac{1}{4}}\langle\phi_\beta|\hat{V}_0|\phi_\alpha\rangle\, e^{-\frac{\tau_1^2}{T^2}}e^{-\frac{i}{\hbar}E_\alpha^{(0)}\tau_1} \\
&= -\frac{i}{\hbar}\left(\frac{2}{\pi}\right)^{\frac{1}{4}}\langle\phi_\beta|\hat{V}_0|\phi_\alpha\rangle\int_{-\infty}^{\infty} d\tau_1 e^{\frac{i}{\hbar}(E_\beta^{(0)}-E_\alpha^{(0)})\tau_1}e^{-\frac{\tau_1^2}{T^2}} \\
&= -\frac{i}{\hbar}\left(\frac{2}{\pi}\right)^{\frac{1}{4}}\langle\phi_\beta|\hat{V}_0|\phi_\alpha\rangle\int_{-\infty}^{\infty} d\tau_1 e^{-\frac{\tau_1^2}{T^2}+i\omega_{\beta\alpha}\tau_1} \\
&= -\frac{i\sqrt{\pi}}{\hbar}\left(\frac{2}{\pi}\right)^{\frac{1}{4}}\langle\phi_\beta|\hat{V}_0|\phi_\alpha\rangle\, Te^{-\frac{1}{4}\omega_{\beta\alpha}^2 T^2}. \qquad (14.57)
\end{aligned}
$$

The corresponding transition probability is then

$$P_{\beta\alpha}(\infty, -\infty) = \frac{\sqrt{2\pi}}{\hbar^2}\left|\langle\phi_\beta|\hat{V}_0|\phi_\alpha\rangle\right|^2 T^2 e^{-\frac{1}{2}\omega_{\beta\alpha}^2 T^2}. \qquad (14.58)$$

The transition rate can now be written as

$$\Gamma_{\beta\alpha}(T) = \frac{P_{\beta\alpha}(\infty, -\infty)}{T} = \frac{\sqrt{2\pi}}{\hbar^2}\left|\langle\phi_\beta|\hat{V}_0|\phi_\alpha\rangle\right|^2 T e^{-\frac{1}{2}\omega_{\beta\alpha}^2 T^2}. \qquad (14.59)$$

For any T, the transition rate is a gaussian distribution as a function of frequency for transitions to various states β. Now consider the case where $T \to \infty$. The peak of the distribution is maximum at $\omega_{\beta\alpha} = 0$. the transition rate increases with T and becomes increasingly narrow as a function of frequency. So as T in increases the peak increases is size and decreases in width. Furthermore, since

$$\int_{-\infty}^{\infty} d\omega Te^{-\frac{1}{2}\omega^2 T^2} = \sqrt{2\pi}, \qquad (14.60)$$

239

The area under the peak is fixed. This implies that in the limit at $T \to \infty$, The distribution becomes a delta function in the frequency. Using

$$\lim_{T \to \infty} T e^{-\frac{1}{2}\omega_{\beta\alpha}^2 T^2} = \sqrt{2\pi}\delta(\omega_{\beta\alpha}) , \tag{14.61}$$

We can define

$$\Gamma_{\beta\alpha} \equiv \lim_{T \to \infty} \Gamma_{\beta\alpha}(T) = \lim_{T \to \infty} \frac{\sqrt{2\pi}}{\hbar^2} \left| \langle \phi_\beta | \hat{V}_0 | \phi_\alpha \rangle \right|^2 T e^{-\frac{1}{2}\omega_{\beta\alpha}^2 T^2} = \frac{2\pi}{\hbar^2} \left| \langle \phi_\beta | \hat{V}_0 | \phi_\alpha \rangle \right|^2 \delta(\omega_{\beta\alpha}) . \tag{14.62}$$

14.5   Harmonic Time Dependence

Now assume that the the potential has a harmonic time dependence. Since the interaction must be limited to a finite region in time for the transition amplitude to be finite, we will also assume that there is an overall gaussian factor to localized the interaction in time. The potential is then

$$\hat{V}(t) = \left(\frac{2}{\pi}\right)^{\frac{1}{4}} \hat{V}_0 e^{\mp i\omega t} e^{-\frac{t^2}{T^2}} . \tag{14.63}$$

We can now proceed as in the previous case to calculate the transition amplitude as

$$
\begin{aligned}
S_{\beta\alpha}^{(1)}(\infty, -\infty) &= -\frac{i}{\hbar} \left(\frac{2}{\pi}\right)^{\frac{1}{4}} \int_{-\infty}^{\infty} d\tau_1 e^{\frac{i}{\hbar}E_\beta^{(0)}\tau_1} \langle \phi_\beta | \hat{V}_0 | \phi_\alpha \rangle e^{\mp i\omega\tau_1} e^{-\frac{\tau_1^2}{T^2}} e^{-\frac{i}{\hbar}E_\alpha^{(0)}\tau_1} \\
&= -\frac{i}{\hbar} \left(\frac{2}{\pi}\right)^{\frac{1}{4}} \langle \phi_\beta | \hat{V}_0 | \phi_\alpha \rangle \int_{-\infty}^{\infty} d\tau_1 e^{\frac{i}{\hbar}(E_\beta^{(0)} - E_\alpha^{(0)})\tau_1} e^{\mp i\omega\tau_1} e^{-\frac{\tau_1^2}{T^2}} \\
&= -\frac{i}{\hbar} \left(\frac{2}{\pi}\right)^{\frac{1}{4}} \langle \phi_\beta | \hat{V}_0 | \phi_\alpha \rangle \int_{-\infty}^{\infty} d\tau_1 e^{-\frac{\tau_1^2}{T^2} + i(\omega_{\beta\alpha} \mp \omega)\tau_1} \\
&= -\frac{i\sqrt{\pi}}{\hbar} \left(\frac{2}{\pi}\right)^{\frac{1}{4}} \langle \phi_\beta | \hat{V}_0 | \phi_\alpha \rangle T e^{-\frac{1}{4}(\omega_{\beta\alpha} \mp \omega)^2 T^2} . \tag{14.64}
\end{aligned}
$$

The corresponding transition probability is then

$$P_{\beta\alpha}(\infty, -\infty) = \frac{\sqrt{2\pi}}{\hbar^2} \left| \langle \phi_\beta | \hat{V}_0 | \phi_\alpha \rangle \right|^2 T^2 e^{-\frac{1}{2}(\omega_{\beta\alpha} \mp \omega)^2 T^2} \tag{14.65}$$

which results in the transition rate

$$\Gamma_{\beta\alpha}(T) = \frac{\sqrt{2\pi}}{\hbar^2} \left| \langle \phi_\beta | \hat{V}_0 | \phi_\alpha \rangle \right|^2 T e^{-\frac{1}{2}(\omega_{\beta\alpha} \mp \omega)^2 T^2} . \tag{14.66}$$

The is the same as in the previous case with the exception that the driving frequency $\pm\omega$ appears in the gaussian. Using the same argument as in the previous case, in the limit $T \to \infty$, we have

$$\Gamma_{\beta\alpha} = \frac{2\pi}{\hbar^2} \left| \langle \phi_\beta | \hat{V}_0 | \phi_\alpha \rangle \right|^2 \delta(\omega_{\beta\alpha} \mp \omega) \,. \tag{14.67}$$

This means that, in this limit, only transitions where $\omega_{\beta\alpha} = \pm\omega$ will contribute to the decay rate. Since $\omega \geq 0$, $\omega_{\beta\alpha} = \omega$ requires that $E_\beta \geq E_\alpha$ and $\omega_{\beta\alpha} = -\omega$ requires that $E_\beta \leq E_\alpha$. Therefore, the time dependence $e^{-i\omega t}$ cause the system to make a transition to a higher energy state, while $e^{i\omega t}$ cause the system to make a transition to a lower energy state.

14.6 Electromagnetic Transitions

We can now apply time-dependent perturbation theory to study the emission and absorption of photons by atoms and nuclei. As a first step, we need to review classical electrodynamics.

14.6.1 Classical Electrodynamics

The equations of motion for electrodynamics are Maxwell's equations (in the esu system)

$$\boldsymbol{\nabla} \cdot \boldsymbol{E} = 4\pi\rho \tag{14.68}$$

$$\boldsymbol{\nabla} \cdot \boldsymbol{B} = 0 \tag{14.69}$$

$$\boldsymbol{\nabla} \times \boldsymbol{E} + \frac{1}{c}\frac{\partial \boldsymbol{B}}{\partial t} = 0 \tag{14.70}$$

$$\boldsymbol{\nabla} \times \boldsymbol{B} - \frac{1}{c}\frac{\partial \boldsymbol{E}}{\partial t} = \frac{4\pi}{c}\boldsymbol{j} \,, \tag{14.71}$$

where we have assumed that the fields are in the vacuum. Here $\boldsymbol{E}$ and $\boldsymbol{B}$ and the electric and magnetic fields, and ρ and are source charge and current densities. The charge and current densities satisfy the continuity equation

$$\boldsymbol{\nabla} \cdot \boldsymbol{j} + \frac{\partial \rho}{\partial t} = 0 \tag{14.72}$$

which means that electric charge is conserved.

The usual strategy for solving Maxwell's equations is to redefine the electric and magnetic fields in terms of a (rotational) scalar potential field ϕ and a (rotational) vector potential field $\boldsymbol{A}$. Choosing

$$\boldsymbol{E} = -\boldsymbol{\nabla}\phi - \frac{1}{c}\frac{\partial \boldsymbol{A}}{\partial t} \tag{14.73}$$

and

$$\boldsymbol{B} = \boldsymbol{\nabla} \times \boldsymbol{A} \tag{14.74}$$

the two homogeneous equations (14.69) and (14.70) are satisfied identically. Then, (14.68) becomes

$$\nabla^2 \phi + \frac{1}{c} \frac{\partial}{\partial t} \left(\boldsymbol{\nabla} \cdot \boldsymbol{A} \right) = -4\pi\rho \tag{14.75}$$

and (14.71) becomes

$$\nabla^2 \boldsymbol{A} - \frac{1}{c^2} \frac{\partial^2 \boldsymbol{A}}{\partial t} - \boldsymbol{\nabla} \left(\boldsymbol{\nabla} \cdot \boldsymbol{A} + \frac{1}{c} \frac{\partial \phi}{\partial t} \right) = -\frac{4\pi}{c} \boldsymbol{j} \,. \tag{14.76}$$

This is a set of four second order coupled equations for ϕ and the three components of $\boldsymbol{A}$.

Since the electric and magnetic fields are defined in terms of derivatives of the scalar and vector potentials and the fields are directly related to the classical forces, the scalar and vector potentials are not uniquely determined. These potentials can be transformed using *gauge transformtions* that leave the electric and magnetic fields unchanged. A careful choice of gauge transformations can be chosen to obtain the Lorentz condition

$$\boldsymbol{\nabla} \cdot \boldsymbol{A} + \frac{1}{c} \frac{\partial \phi}{\partial t} = 0 \,. \tag{14.77}$$

It is still possible to find gauge transformations that preserve the fields and the Lorentz condition. This allows the condition

$$\boldsymbol{\nabla} \cdot \boldsymbol{A} = 0 \tag{14.78}$$

to also be imposed. This is called the *Coulomb gauge*. Then (14.75) and (14.76) become

$$\nabla^2 \phi = -4\pi\rho \tag{14.79}$$

and

$$\nabla^2 \boldsymbol{A} - \frac{1}{c^2} \frac{\partial^2 \boldsymbol{A}}{\partial t} = -\frac{4\pi}{c} \boldsymbol{j} \,. \tag{14.80}$$

Now consider the solution to (14.80) in a region of space where $j = 0$. The vector potential is then a solution to a homogeneous wave equation and has the form

$$\boldsymbol{A} = A_0 \boldsymbol{\varepsilon} e^{i(\boldsymbol{k} \cdot \boldsymbol{r} - \omega t)} \tag{14.81}$$

where $\boldsymbol{k}$ is the wave vector, ω is the angular frequency and $\boldsymbol{\varepsilon}$ is a unit vector in the direction of the vector potential that is called the polarization vector. Therefore, the polarization vector satisfies

$$\varepsilon^2 = 1 \,. \tag{14.82}$$

For the plane wave to be a solution to the wave equation the wave vector and angular frequency must satisfy the dispersion relation

$$k^2 = \frac{\omega^2}{c^2} \, . \tag{14.83}$$

If we require that (14.81) must satisfy the Coulomb condition (14.78) then

$$0 = \boldsymbol{\nabla} \cdot \boldsymbol{A} = iA_0 \boldsymbol{\varepsilon} \cdot \boldsymbol{k} e^{i(\boldsymbol{k} \cdot \boldsymbol{r} - \omega t)} \, , \tag{14.84}$$

which requires that

$$\boldsymbol{\varepsilon} \cdot \boldsymbol{k} = 0 \, , \tag{14.85}$$

This implies that the vector potential is perpendicular to the direction of motion of the wave, or is transverse.

14.6.2   Quantization

In order to consider the absorption and emission of photons from atoms and nuclei, it is necessary to quantize a system of charged particles interacting with electromagnetic fields. This is in general a difficult and mathematically complicated problem that was the subject of considerable attention beginning in the late 1920s and was not solved satisfactorily until around 1950. The problem of field quantization is treated in a more advanced course and will not be considered here. However, it can be proved from field theory that at leading order in perturbation theory, the problem of photon emission and absorption can be solved by assuming that the charged particles are subject to external classical electromagnetic fields. This leads to the semiclassical treatment of radiation.

To see how this works consider an electron in an external electromagnetic field. The classical hamiltonian for this is

$$H(\boldsymbol{p}, \boldsymbol{r}) = \frac{1}{2m_e} \left(\boldsymbol{p} + \frac{e}{c} \boldsymbol{A}(\boldsymbol{r}, t) \right)^2 - e\phi(\boldsymbol{r}, t) \, . \tag{14.86}$$

Note that the hamiltonian is expressed in terms of the scalar and vector potentials rather than the electric and magnetic fields. To obtain the semiclassical hamiltonian we replace the canonical coordinates and momenta by the corresponding quantum operators to give

$$
\begin{aligned}
\hat{H} &= \frac{1}{2m_e} \left(\hat{\boldsymbol{p}} + \frac{e}{c} \boldsymbol{A}(\hat{\boldsymbol{r}}, t) \right)^2 - e\phi(\hat{\boldsymbol{r}}, t) \\
&= \frac{\hat{\boldsymbol{p}}^2}{2m_e} - e\phi(\hat{\boldsymbol{r}}, t) + \frac{e}{2m_e c} \hat{\boldsymbol{p}} \cdot \boldsymbol{A}(\hat{\boldsymbol{r}}, t) + \frac{e}{2m_e c} \boldsymbol{A}(\hat{\boldsymbol{r}}, t) \cdot \hat{\boldsymbol{p}} + \frac{e^2}{2m_e c^2} \boldsymbol{A}^2(\hat{\boldsymbol{r}}, t) \\
&= \frac{\hat{\boldsymbol{p}}^2}{2m_e} - e\phi(\boldsymbol{r}, t) + \frac{\hbar e}{2im_e c} (\boldsymbol{\nabla} \cdot \boldsymbol{A}(\hat{\boldsymbol{r}}, t)) + \frac{e}{m_e c} \boldsymbol{A}(\hat{\boldsymbol{r}}, t) \cdot \hat{\boldsymbol{p}} + \frac{e^2}{2m_e c^2} \boldsymbol{A}^2(\hat{\boldsymbol{r}}, t) \\
&= \frac{\hat{\boldsymbol{p}}^2}{2m_e} - e\phi(\hat{\boldsymbol{r}}, t) + \frac{e}{m_e c} \boldsymbol{A}(\hat{\boldsymbol{r}}, t) \cdot \hat{\boldsymbol{p}} + \frac{e^2}{2m_e c^2} \boldsymbol{A}^2(\hat{\boldsymbol{r}}, t) \, , \tag{14.87}
\end{aligned}
$$

where the Coulomb condition (14.78) has been used to simplify the hamiltonian. Since we are working to first order in the perturbation, we can neglect the last term since is goes like e^2 rather than e. For calculation of photon absorption or emission, the we are only concerned with the vector potential so we can identify the time dependent potential as

$$
\begin{aligned}
V_\pm(t) &= \left(\frac{2}{\pi}\right)^{\frac{1}{4}} \frac{e}{m_e c} \boldsymbol{A}_\pm(\hat{\boldsymbol{r}}, t) \cdot \hat{\boldsymbol{p}}\, e^{-\frac{t^2}{T^2}} = \left(\frac{2}{\pi}\right)^{\frac{1}{4}} \frac{e}{m_e c} A_0 \boldsymbol{\varepsilon}\, e^{\pm i(\boldsymbol{k}\cdot\hat{\boldsymbol{r}}-\omega t)} \cdot \hat{\boldsymbol{p}}\, e^{-\frac{t^2}{T^2}} \\
&= \left(\frac{2}{\pi}\right)^{\frac{1}{4}} \frac{e A_0}{m_e c} e^{\pm i\boldsymbol{k}\cdot\hat{\boldsymbol{r}}} \boldsymbol{\varepsilon} \cdot \hat{\boldsymbol{p}}\, e^{-\frac{t^2}{T^2}} e^{\mp i\omega t} ,
\end{aligned}
\tag{14.88}
$$

where we have included the Gaussian factor so that the potential is in effect only over a finite time period. The positive (negative) subscript indicates that the photon is to be absorbed (emitted). The time structure of this potential is the same as in (14.63). We can, therefore, use (14.63) to obtain the first order contribution to the transition amplitude as

$$
S^{(1)\pm}_{\beta\alpha}(\infty, -\infty) = -\frac{i\sqrt{\pi}}{\hbar} \left(\frac{2}{\pi}\right)^{\frac{1}{4}} \frac{e A_0}{m_e c} \langle \phi_\beta | e^{\pm i\boldsymbol{k}\cdot\hat{\boldsymbol{r}}} \boldsymbol{\varepsilon} \cdot \hat{\boldsymbol{p}} | \phi_\alpha \rangle T e^{-\frac{1}{4}(\omega_{\beta\alpha}\mp\omega)^2 T^2} .
\tag{14.89}
$$

The corresponding transition probability is then

$$
P^\pm_{\beta\alpha}(\infty, -\infty) = \frac{\sqrt{2\pi}}{\hbar^2} \left|\frac{e A_0}{m_e c}\right|^2 \left|\langle \phi_\beta | e^{\pm i\boldsymbol{k}\cdot\hat{\boldsymbol{r}}} \boldsymbol{\varepsilon} \cdot \hat{\boldsymbol{p}} | \phi_\alpha \rangle\right|^2 T^2 e^{-\frac{1}{2}(\omega_{\beta\alpha}\mp\omega)^2 T^2}
\tag{14.90}
$$

which results in the transition rate

$$
\Gamma^\pm_{\beta\alpha}(T) = \frac{\sqrt{2\pi}}{\hbar^2} \frac{e^2 |A_0|^2}{m_e^2 c^2} \left|\langle \phi_\beta | e^{\pm i\boldsymbol{k}\cdot\hat{\boldsymbol{r}}} \boldsymbol{\varepsilon} \cdot \hat{\boldsymbol{p}} | \phi_\alpha \rangle\right|^2 T e^{-\frac{1}{2}(\omega_{\beta\alpha}\mp\omega)^2 T^2} .
\tag{14.91}
$$

We can now apply this to the calculation of the cross section for photo-absorption by the hydrogen atom. The first thing we need to do is to obtain an expression for the $|A_0|^2$. To do this in the semiclassical approximation it is necessary to make a particular choice for the form of the vector potential. Since classically, the vector potential must be real, we choose

$$
\boldsymbol{A}(\boldsymbol{r}, t) = A_0 \boldsymbol{\varepsilon} \left(e^{i(\boldsymbol{k}\cdot\boldsymbol{r}-\omega t)} + e^{-i(\boldsymbol{k}\cdot\boldsymbol{r}-\omega t)}\right) ,
\tag{14.92}
$$

where A_0 and $\boldsymbol{\varepsilon}$ are chosen to be real. This can be used to calculate the magnetic field from (14.74) as

$$
\boldsymbol{B} = \boldsymbol{\nabla} \times \boldsymbol{A} = i\boldsymbol{k} \times A_0 \boldsymbol{\varepsilon} \left(e^{i(\boldsymbol{k}\cdot\boldsymbol{r}-\omega t)} - e^{-i(\boldsymbol{k}\cdot\boldsymbol{r}-\omega t)}\right) .
\tag{14.93}
$$

Similarly, the contribution of the wave to the electric field is obtained from (14.73) as

$$
\boldsymbol{E} = -\frac{1}{c}\frac{\partial \boldsymbol{A}}{\partial t} = \frac{i\omega}{c} A_0 \boldsymbol{\varepsilon} \left(e^{i(\boldsymbol{k}\cdot\boldsymbol{r}-\omega t)} - e^{-i(\boldsymbol{k}\cdot\boldsymbol{r}-\omega t)}\right) .
\tag{14.94}
$$

The energy flux (the energy per unit area per unit time passing through a surface perpendicular to $\boldsymbol{k}$) is given by the Poynting vector

$$\boldsymbol{S} = \frac{c}{4\pi} \boldsymbol{E} \times \boldsymbol{B} \,. \tag{14.95}$$

Then,

$$
\begin{aligned}
\boldsymbol{E} \times \boldsymbol{B} &= \frac{i\omega}{c} A_0 \boldsymbol{\varepsilon} \left(e^{i(\boldsymbol{k}\cdot\boldsymbol{r}-\omega t)} - e^{-i(\boldsymbol{k}\cdot\boldsymbol{r}-\omega t)} \right) \times i\boldsymbol{k} \times A_0 \boldsymbol{\varepsilon} \left(e^{i(\boldsymbol{k}\cdot\boldsymbol{r}-\omega t)} - e^{-i(\boldsymbol{k}\cdot\boldsymbol{r}-\omega t)} \right) \\
&= -\frac{A_0^2 \omega}{c} \boldsymbol{\varepsilon} \times (\boldsymbol{k} \times \boldsymbol{\varepsilon}) \left(e^{i(\boldsymbol{k}\cdot\boldsymbol{r}-\omega t)} - e^{-i(\boldsymbol{k}\cdot\boldsymbol{r}-\omega t)} \right) \left(e^{i(\boldsymbol{k}\cdot\boldsymbol{r}-\omega t)} - e^{-i(\boldsymbol{k}\cdot\boldsymbol{r}-\omega t)} \right) \\
&= -\frac{A_0^2 \omega}{c} \left(\boldsymbol{k}\varepsilon^2 - \boldsymbol{\varepsilon}(\boldsymbol{k}\cdot\boldsymbol{\varepsilon}) \right) \left(e^{i2(\boldsymbol{k}\cdot\boldsymbol{r}-\omega t)} + e^{-i2(\boldsymbol{k}\cdot\boldsymbol{r}-\omega t)} - 2 \right) \\
&= \frac{2A_0^2 \omega}{c} \boldsymbol{k} \left(1 - \cos(2(\boldsymbol{k}\cdot\boldsymbol{r} - \omega t)) \right) \,. \tag{14.96}
\end{aligned}
$$

So,

$$\boldsymbol{S} = \frac{c}{4\pi} \frac{2A_0^2 \omega}{c} \boldsymbol{k} \left\{ 1 - \cos[2(\boldsymbol{k}\cdot\boldsymbol{r} - \omega t)] \right\} = \frac{A_0^2 \omega}{2\pi} \boldsymbol{k} \left\{ 1 - \cos[2(\boldsymbol{k}\cdot\boldsymbol{r} - \omega t)] \right\} \,. \tag{14.97}$$

If we average the Poynting vector over one period, the oscillatory piece will vanish and we can write the average Poynting vector as

$$\bar{\boldsymbol{S}} = \frac{A_0^2 \omega}{2\pi} \boldsymbol{k} \,. \tag{14.98}$$

The intensity of the wave is the magnitude of the time-averaged pointing vector so

$$I(\omega) = \frac{\omega^2}{2\pi c} |A_0|^2 \,, \tag{14.99}$$

or

$$|A_0|^2 = \frac{2\pi c I(\omega)}{\omega^2} \,. \tag{14.100}$$

The transition rate for photo-absorption can now be obtained from (14.91) using (14.100) to give

$$\Gamma_{\beta\alpha}^+(T) = \frac{\sqrt{2\pi}}{\hbar^2} \frac{e^2}{m_e^2 c^2} \frac{2\pi c I(\omega)}{\omega^2} \left| \langle \phi_\beta | e^{\pm i\boldsymbol{k}\cdot\hat{\boldsymbol{r}}} \boldsymbol{\varepsilon} \cdot \hat{\boldsymbol{p}} | \phi_\alpha \rangle \right|^2 T e^{-\frac{1}{2}(\omega_{\beta\alpha}-\omega)^2 T^2} \,. \tag{14.101}$$

In the limit $T \to \infty$, we have

$$\Gamma_{\beta\alpha}^+ = \frac{2\pi}{\hbar^2} \frac{e^2}{m_e^2 c^2} \frac{2\pi c I(\omega)}{\omega^2} \left| \langle \phi_\beta | e^{\pm i\boldsymbol{k}\cdot\hat{\boldsymbol{r}}} \boldsymbol{\varepsilon} \cdot \hat{\boldsymbol{p}} | \phi_\alpha \rangle \right|^2 \delta(\omega_{\beta\alpha} - \omega) \,. \tag{14.102}$$

245

The cross section per unit energy is just the transition rate divided by the intensity of the incoming radiation. Therefore,

$$
\begin{aligned}
\frac{d\sigma}{dE} &= \frac{\Gamma^+_{\beta\alpha}}{I(\omega)} = \frac{2\pi}{\hbar^2}\frac{e^2}{m_e^2 c^2}\frac{2\pi c}{\omega^2}\left|\langle\phi_\beta| e^{i\boldsymbol{k}\cdot\hat{\boldsymbol{r}}}\boldsymbol{\varepsilon}\cdot\hat{\boldsymbol{p}}\,|\phi_\alpha\rangle\right|^2 \delta(\omega_{\beta\alpha}-\omega) \\
&= \frac{2\pi}{\hbar^2}\frac{\alpha\hbar c}{m_e^2 c^2}\frac{2\pi\hbar^2 c}{(E_\beta - E_\alpha)^2}\left|\langle\phi_\beta| e^{i\boldsymbol{k}\cdot\hat{\boldsymbol{r}}}\boldsymbol{\varepsilon}\cdot\hat{\boldsymbol{p}}\,|\phi_\alpha\rangle\right|^2 \hbar\delta(E_\beta - E_\alpha - E) \\
&= \frac{4\pi^2\alpha\hbar^2}{m_e^2(E_\beta - E_\alpha)^2}\left|\langle\phi_\beta| e^{i\boldsymbol{k}\cdot\hat{\boldsymbol{r}}}\boldsymbol{\varepsilon}\cdot\hat{\boldsymbol{p}}\,|\phi_\alpha\rangle\right|^2 \delta(E_\beta - E_\alpha - E)\,. \qquad (14.103)
\end{aligned}
$$

Thus, the cross section can be determined by calculating the matrix element $\langle\phi_\beta| e^{i\boldsymbol{k}\cdot\hat{\boldsymbol{r}}}\boldsymbol{\varepsilon}\cdot\hat{\boldsymbol{p}}\,|\phi_\alpha\rangle$. General techniques have been developed for this calculation but are to complicated to describe here. We will consider a simple approximation that is more readily calculated.

Consider the case where the wave length of the photons is much larger than the size of the system that will absorb the proton. In this case the exponential $e^{i\boldsymbol{k}\cdot\hat{\boldsymbol{r}}}$ varies only slightly over the volume of the target so that we can make the approximation that

$$
e^{i\boldsymbol{k}\cdot\hat{\boldsymbol{r}}} \cong 1\,. \qquad (14.104)
$$

Then,

$$
\langle\phi_\beta| e^{i\boldsymbol{k}\cdot\hat{\boldsymbol{r}}}\boldsymbol{\varepsilon}\cdot\hat{\boldsymbol{p}}\,|\phi_\alpha\rangle \cong \langle\phi_\beta|\boldsymbol{\varepsilon}\cdot\hat{\boldsymbol{p}}\,|\phi_\alpha\rangle = \boldsymbol{\varepsilon}\cdot\langle\phi_\beta|\hat{\boldsymbol{p}}\,|\phi_\alpha\rangle\,. \qquad (14.105)
$$

This can be rewritten in a form which is more easily calculated by considering the expression for the transition matrix given by (14.48). For this case, this gives

$$
\begin{aligned}
S^{(1)}_{\beta\alpha}(t,t') &= -\frac{i}{\hbar}\int_{t'}^{t} d\tau_1 e^{\frac{i}{\hbar}E^{(0)}_\beta \tau_1}\langle\phi_\beta| V(\tau_1)\,|\phi_\alpha\rangle e^{-\frac{i}{\hbar}E^{(0)}_\alpha \tau_1} \\
&= -\frac{i}{\hbar}\int_{t'}^{t} d\tau_1 e^{\frac{i}{\hbar}E^{(0)}_\beta \tau_1}\langle\phi_\beta|\left(\frac{2}{\pi}\right)^{\frac{1}{4}}\frac{eA_0}{m_e c}e^{i\boldsymbol{k}\cdot\hat{\boldsymbol{r}}}\boldsymbol{\varepsilon}\cdot\hat{\boldsymbol{p}}\,e^{-\frac{\tau_1^2}{T^2}}e^{-i\omega\tau_1}\,|\phi_\alpha\rangle e^{-\frac{i}{\hbar}E^{(0)}_\alpha \tau_1} \\
&\cong -\frac{i}{\hbar}\left(\frac{2}{\pi}\right)^{\frac{1}{4}}\frac{eA_0}{m_e c}\int_{t'}^{t} d\tau_1 e^{\frac{i}{\hbar}E^{(0)}_\beta \tau_1}\langle\phi_\beta|\boldsymbol{\varepsilon}\cdot\hat{\boldsymbol{p}}\,|\phi_\alpha\rangle e^{-\frac{i}{\hbar}E^{(0)}_\alpha \tau_1}e^{-\frac{\tau_1^2}{T^2}}e^{-i\omega\tau_1} \\
&= -\frac{i}{\hbar}\left(\frac{2}{\pi}\right)^{\frac{1}{4}}\frac{eA_0}{m_e c}\int_{t'}^{t} d\tau_1 \boldsymbol{\varepsilon}\cdot\langle\phi_\beta| e^{\frac{i}{\hbar}\hat{H}_0\tau_1}\hat{\boldsymbol{p}}\,e^{-\frac{i}{\hbar}\hat{H}_0\tau_1}\,|\phi_\alpha\rangle e^{-\frac{\tau_1^2}{T^2}}e^{-i\omega\tau_1}
\end{aligned}
$$

$$(14.106)$$

Now consider

$$
\frac{d}{dt}e^{\frac{i}{\hbar}\hat{H}_0 t}\hat{r}_i\, e^{-\frac{i}{\hbar}\hat{H}_0 t} = e^{\frac{i}{\hbar}\hat{H}_0 t}\frac{i}{\hbar}[\hat{H}_0,\hat{r}_i]\,e^{-\frac{i}{\hbar}\hat{H}_0 t}\,. \qquad (14.107)
$$

For a central potential

$$
\hat{H}_0 = \frac{\hat{\boldsymbol{p}}^2}{2m_e} + V(\hat{r}) \qquad (14.108)
$$

246

and
$$\frac{i}{\hbar}[\hat{H}_0, \hat{r}_i] = \frac{1}{m_e}\hat{p}_i \, . \tag{14.109}$$

So,
$$e^{\frac{i}{\hbar}\hat{H}_0 t}\hat{p}_i\, e^{-\frac{i}{\hbar}\hat{H}_0 t} = m_e\frac{d}{dt}e^{\frac{i}{\hbar}\hat{H}_0 t}\hat{r}_i\, e^{-\frac{i}{\hbar}\hat{H}_0 t} \tag{14.110}$$

This can be used to write
$$\begin{aligned}
\langle\phi_\beta|\, e^{\frac{i}{\hbar}\hat{H}_0 t}\hat{p}_i\, e^{-\frac{i}{\hbar}\hat{H}_0 t}\,|\phi_\alpha\rangle &= m_e\langle\phi_\beta|\,\frac{d}{dt}e^{\frac{i}{\hbar}\hat{H}_0 t}\hat{r}_i\, e^{-\frac{i}{\hbar}\hat{H}_0 t}\,|\phi_\alpha\rangle \\
&= m_e\frac{d}{dt}\langle\phi_\beta|\, e^{\frac{i}{\hbar}\hat{H}_0 t}\hat{r}_i\, e^{-\frac{i}{\hbar}\hat{H}_0 t}\,|\phi_\alpha\rangle \\
&= m_e\frac{d}{dt}e^{\frac{i}{\hbar}(E_\beta^{(0)}-E_\alpha^{(0)})t}\langle\phi_\beta|\,\hat{r}_i\,|\phi_\alpha\rangle \\
&= m_e\frac{d}{dt}e^{i\omega_{\beta\alpha}t}\langle\phi_\beta|\,\hat{r}_i\,|\phi_\alpha\rangle \\
&= im_e\omega_{\beta\alpha}e^{i\omega_{\beta\alpha}t}\langle\phi_\beta|\,\hat{r}_i\,|\phi_\alpha\rangle \, .
\end{aligned} \tag{14.111}$$

We could then proceed to follow the derivation through again to obtain an approximate expression for the cross section. This would be equivalent to making the substitution
$$\langle\phi_\beta|\, e^{i\boldsymbol{k}\cdot\hat{\boldsymbol{r}}}\hat{\boldsymbol{p}}\,|\phi_\alpha\rangle \to m_e\omega_{\beta\alpha}\langle\phi_\beta|\,\hat{\boldsymbol{r}}\,|\phi_\alpha\rangle = \frac{m_e}{\hbar}(E_\beta^{(0)}-E_\alpha^{(0)})\langle\phi_\beta|\,\hat{\boldsymbol{r}}\,|\phi_\alpha\rangle \tag{14.112}$$

in (14.103). As a result
$$\frac{d\sigma}{dE} = 4\pi^2\alpha\,|\langle\phi_\beta|\,\boldsymbol{\varepsilon}\cdot\hat{\boldsymbol{r}}\,|\phi_\alpha\rangle|^2\,\delta(E_\beta - E_\alpha - E)\, . \tag{14.113}$$

Since the electric dipole moment of a proton and electron would be given by er, this is called the dipole approximation.

The cross section (14.113) assumes that the initial and final states a completely specified and that the photons are of fixed polarization. This is often not the case. If we assume that the incident light is unpolarized it is necessary to average (14.113) over the possible directions of polarization. For this purpose we will assume that the photons are incident along the z axis. That is, $\boldsymbol{k} = k\boldsymbol{e}_z$. Since the polarization vector must be normal to $\boldsymbol{k}$, it will lie in the xy plane. The direction of the polarization can then be parameterized in terms of the angle φ of the polarization vector to the x axis. Then,
$$\langle\phi_\beta|\,\boldsymbol{\varepsilon}\cdot\hat{\boldsymbol{r}}\,|\phi_\alpha\rangle = \langle\phi_\beta|\,(\hat{x}\cos\varphi + \hat{y}\sin\varphi)\,|\phi_\alpha\rangle \, . \tag{14.114}$$

The absolute square of this is
$$\begin{aligned}
|\langle\phi_\beta|\,\boldsymbol{\varepsilon}\cdot\hat{\boldsymbol{r}}\,|\phi_\alpha\rangle|^2 &= \cos^2\varphi\,|\langle\phi_\beta|\,\hat{x}\,|\phi_\alpha\rangle|^2 + \sin^2\varphi\,|\langle\phi_\beta|\,\hat{y}\,|\phi_\alpha\rangle|^2 \\
&\quad + \cos\varphi\sin\varphi\left(\langle\phi_\alpha|\,\hat{x}\,|\phi_\beta\rangle\langle\phi_\beta|\,\hat{y}\,|\phi_\alpha\rangle + \langle\phi_\alpha|\,\hat{y}\,|\phi_\beta\rangle\langle\phi_\beta|\,\hat{x}\,|\phi_\alpha\rangle\right) \, .
\end{aligned} \tag{14.115}$$

Averaging over ϕ yields

$$\frac{1}{2\pi}\int_0^{2\pi} d\varphi\,|\langle\phi_\beta|\,\boldsymbol{\varepsilon}\cdot\hat{\boldsymbol{r}}\,|\phi_\alpha\rangle|^2 \;=\; \frac{1}{2}\left(|\langle\phi_\beta|\,\hat{x}\,|\phi_\alpha\rangle|^2 + |\langle\phi_\beta|\,\hat{y}\,|\phi_\alpha\rangle|^2\right)$$

$$=\; \frac{1}{2}\left(|\langle\phi_\beta|\,\hat{r}_1\,|\phi_\alpha\rangle|^2 + |\langle\phi_\beta|\,\hat{r}_{-1}\,|\phi_\alpha\rangle|^2\right)\;.\text{(14.116)}$$

For the hydrogen atom, we need to calculate the matrix element

$$\langle\phi_{n'l'm'}|\,\hat{r}_\lambda\,|\phi_{nlm}\rangle \;=\; \int_0^\infty dr\,r^2 R_{n'l'}(r)\,\langle l'm'|\,\hat{r}_\lambda R_{nl}(r)\,|lm\rangle$$

$$=\; \int_0^\infty dr\,r^2 R_{n'l'}(r)R_{nl}(r)\,\langle l'm'|\,\hat{r}_\lambda\,|lm\rangle \qquad \text{(14.117)}$$

If the atom is not prepared in a with a fixed value of m and m' is not determined for the final state, then we must average of m and sum over m'. Doing this we can write the average cross section as

$$\overline{\frac{d\sigma}{dE}} = 2\pi^2\alpha\frac{1}{2l+1}\sum_m\sum_{m'}\left(|\langle\phi_{n'l'm'}|\,\hat{r}_1\,|\phi_{nlm}\rangle|^2 + |\langle\phi_{n'l'm'}|\,\hat{r}_{-1}\,|\phi_{nlm}\rangle|^2\right)\delta(E_{n'}^{(0)}-E_n^{(0)}-E)\,.$$

$$\text{(14.118)}$$

Chapter 15

Many-body Systems and Spin Statistics

To this point we have considered only the quantum mechanics of a single particle in a potential. We now want to generalize our treatment of the Schrödinger equation to allow for a system of n particles interacting by means of a potential. The hamiltonian for such a system is

$$\hat{H} = \sum_{i=1}^{N} \frac{\hat{\boldsymbol{p}}_i^2}{2m_i} + V(\hat{\boldsymbol{r}}_1, \hat{\boldsymbol{r}}_2, \ldots, \hat{\boldsymbol{r}}_{n-1}, \hat{\boldsymbol{r}}_n) \,. \tag{15.1}$$

The position and momentum operators satisfy the commutations relations

$$[(\hat{\boldsymbol{r}}_i)_m, (\hat{\boldsymbol{p}}_j)_n] = i\hbar\delta_{ij}\delta_{mn} \tag{15.2}$$

$$[(\hat{\boldsymbol{r}}_i)_m, (\hat{\boldsymbol{r}}_j)_n] = 0 \tag{15.3}$$

$$[(\hat{\boldsymbol{p}}_i)_m, (\hat{\boldsymbol{p}}_j)_n] = 0 \,. \tag{15.4}$$

Note that the potential may depend upon intrinsic properties of the particles such as spin in addition to the coordinates. The Schrödinger equation in coordinate space is then

$$\hat{H}\psi(\hat{\boldsymbol{r}}_1, \hat{\boldsymbol{r}}_2, \ldots, \hat{\boldsymbol{r}}_{n-1}, \hat{\boldsymbol{r}}_n; t) = i\hbar \frac{\partial \psi(\hat{\boldsymbol{r}}_1, \hat{\boldsymbol{r}}_2, \ldots, \hat{\boldsymbol{r}}_{n-1}, \hat{\boldsymbol{r}}_n; t)}{\partial t} \,. \tag{15.5}$$

The wave functions are normalized such that

$$\int d^3r_1 d^3r_2 \ldots d^3r_{n-1} d^3r_n \psi^\dagger(\hat{\boldsymbol{r}}_1, \hat{\boldsymbol{r}}_2, \ldots, \hat{\boldsymbol{r}}_{n-1}, \hat{\boldsymbol{r}}_n; t)\psi(\hat{\boldsymbol{r}}_1, \hat{\boldsymbol{r}}_2, \ldots, \hat{\boldsymbol{r}}_{n-1}, \hat{\boldsymbol{r}}_n; t) = 1$$

$$\tag{15.6}$$

The most common situation is for the particles to interact through two-body potentials $V_{ij}(\hat{\boldsymbol{r}}_i, \hat{\boldsymbol{r}}_j)$ where Newton's third law requires that

$$V_{ij}(\hat{\boldsymbol{r}}_i, \hat{\boldsymbol{r}}_j) = V_{ji}(\hat{\boldsymbol{r}}_j, \hat{\boldsymbol{r}}_i) \,. \tag{15.7}$$

In this case the hamitonian operator can be written as

$$\hat{H} = \sum_{i=1}^{N} \frac{\hat{\boldsymbol{p}}_i^2}{2m_i} + \sum_{i=1}^{N} \sum_{j>i}^{N} V_{ij}(\hat{\boldsymbol{r}}_i, \hat{\boldsymbol{r}}_j) \,. \tag{15.8}$$

The solution of the many-body Schrödinger equation is, in general, very difficult and an enormous amount of effort has been expended over the last 80 years in developing approximation methods and numerical techniques for describing many-body systems.

15.1 The Two-Body Problem

The simplest example of such a system is two particles interacting by means of a central potential. The classical Hamiltonian for such a system is given by

$$H = \frac{\boldsymbol{p}_1^2}{2m_1} + \frac{\boldsymbol{p}_2^2}{2m_2} + V(|\boldsymbol{r}_1 - \boldsymbol{r}_2|) \,. \tag{15.9}$$

This Hamiltonian can be separated in to relative and center of mass terms by defining the center-of-mass and relative coordinates as

$$\boldsymbol{R} = \frac{m_1 \boldsymbol{r}_1 + m_2 \boldsymbol{r}_2}{M} \tag{15.10}$$

and

$$\boldsymbol{r} = \boldsymbol{r}_1 - \boldsymbol{r}_2 \,, \tag{15.11}$$

where the total mass is given by

$$M = m_1 + m_2 \,. \tag{15.12}$$

The momentum conjugate to the center-of-mass vector is the total momentum

$$\boldsymbol{P} = \boldsymbol{p}_1 + \boldsymbol{p}_2 \tag{15.13}$$

and the momentum conjugate to the relative position vector is the relative momentum

$$\boldsymbol{p} = \frac{m_2 \boldsymbol{p}_1 - m_1 \boldsymbol{p}_2}{M} \,. \tag{15.14}$$

The Hamiltonian can be rewritten in terms of these new coordinates as

$$H = \frac{\boldsymbol{P}^2}{2M} + \frac{\boldsymbol{p}^2}{2\mu} + V(|\boldsymbol{r}|) \,, \tag{15.15}$$

where

$$\mu = \frac{m_1 m_2}{M} \tag{15.16}$$

is the reduced mass. This is the well known classical result that the motion of an interaction system can be separated into the motion of the center of mass and the relative motion of the two particles. The center of mass moves like a point particle of mass M while the relative motion is described by a one-body problem for a particle of mass μ.

The corresponding quantum mechanical problem can now be constructed using the canonical quantization and the classical correspondence principle. The classical coordinates and momenta are replaced by operators such that

$$[\hat{R}_i, \hat{P}_j] = i\hbar\delta_{ij} \tag{15.17}$$

and

$$[\hat{r}_i, \hat{p}_j] = i\hbar\delta_{ij} \tag{15.18}$$

with all other commutators being zero.

In coordinate space representation, the Schrödinger equation can then be written as

$$\left(-\frac{\hbar^2\boldsymbol{\nabla}_R^2}{2M} - \frac{\hbar^2\boldsymbol{\nabla}_r^2}{2\mu} + V(|\boldsymbol{r}|)\right)\Psi(\boldsymbol{r},\boldsymbol{R}) = E_{tot}\Psi(\boldsymbol{r},\boldsymbol{R}). \tag{15.19}$$

This is a six-dimensional differential equation and can be simplified by using the technique of separation of variables. Assume that the wave function can be written as a product of a center-of-mass and relative wave functions. That is

$$\Psi(\boldsymbol{r},\boldsymbol{R}) = \psi(\boldsymbol{r})\phi(\boldsymbol{R}). \tag{15.20}$$

Substituting this into the Schrödinger equation gives

$$-\frac{\psi(\boldsymbol{r})\hbar^2\boldsymbol{\nabla}_R^2\phi(\boldsymbol{R})}{2M} - \frac{\phi(\boldsymbol{R})\hbar^2\boldsymbol{\nabla}_r^2\psi(\boldsymbol{r})}{2\mu} + V(|\boldsymbol{r}|)\psi(\boldsymbol{r})\phi(\boldsymbol{R}) = E_{tot}\psi(\boldsymbol{r})\phi(\boldsymbol{R}). \tag{15.21}$$

Now dividing both sides by $\psi(\boldsymbol{r})\phi(\boldsymbol{R})$ yields

$$-\frac{\hbar^2\boldsymbol{\nabla}_R^2\phi(\boldsymbol{R})}{2M\phi(\boldsymbol{R})} - \frac{\hbar^2\boldsymbol{\nabla}_r^2\psi(\boldsymbol{r})}{2\mu\psi(\boldsymbol{r})} + V(|\boldsymbol{r}|) = E_{tot}. \tag{15.22}$$

The only way that we can satisfy this equation for all values of the coordinates and all solutions for the two wave functions is if the first term is equal to a constant and consequently the second and third terms add to a constant. Let

$$-\frac{\hbar^2\boldsymbol{\nabla}_R^2\phi(\boldsymbol{R})}{2M\phi(\boldsymbol{R})} = E_{CM}. \tag{15.23}$$

Then

$$-\frac{\hbar^2\boldsymbol{\nabla}_R^2\phi(\boldsymbol{R})}{2M} = E_{CM}\phi(\boldsymbol{R}), \tag{15.24}$$

251

This is the Schrödinger equation for a free particle of mass M and has the plane wave solution

$$\phi(\boldsymbol{R}) = \frac{1}{(2\pi\hbar)^{\frac{3}{2}}} \, e^{\frac{i}{\hbar}\boldsymbol{P}\cdot\boldsymbol{R}} . \qquad (15.25)$$

The energy eigenvalue is

$$E_{CM}(\boldsymbol{P}) = \frac{\boldsymbol{P}^2}{2M} . \qquad (15.26)$$

The Schrödinger equation for the relative wave function can now be written as

$$\left(-\frac{\hbar^2 \boldsymbol{\nabla}^2}{2\mu} + V(|\boldsymbol{r}|) \right) \psi(\boldsymbol{r}) = E\psi(\boldsymbol{r}) , \qquad (15.27)$$

where

$$E = E_{tot} - E_{CM} \qquad (15.28)$$

and we have dropped the subscript r from the laplacian for convenience.

The result is that we have reduced the two-body problem to an effective one-body problem for a particle with mass μ. For this reason the two-body problem is a very special situation.

Solution of problems with three or more particles become much more complicated. It is now possible to obtain exact numerical solutions to three-body problems with short-range interactions and exact results for selected states of systems with more particles. In general, however, approximations methods must be used to solve most many-body problem. Over the years a great variety of approximation techniques have been developed for studying the properties of many-body systems.

15.2 The Pauli Exclusion Principle

The nineteenth century was the time that chemistry became firmly established as a science. During this time many elements were discovered and empirical rules for combining these elements into various compounds were discovered. By mid century attempts were being made to establish some order for these new discoveries. In 1869, Dmitri Mendeleev, a Russian chemistry professor, proposed the first successful organization of the elements in his periodic table. Mendeleev showed that he could arrange the elements in order of atomic weight A in a table such that a regular pattern appeared for the valence of the elements. At its introduction, the table contained only 63 elements and there were apparent gaps where Mendeleev predicted that new elements were required to complete the pattern. Three of these elements, gallium, scandium, and germanium, were found during Mendeleev's lifetime and had properties remarkably close to those predicted by him. In some cases the order in atomic weight did not correctly correspond to the chemical properties of the elements in which case Mendeleev allowed the order to of the elements in the table to be reordered to correctly group the elements by chemical properties. The transition

metals also could not be properly arranged in Mendeleev's periodic table and were simply arranged according to atomic weight. In addition, none of the noble gases had yet been discovered and so one whole column was missing from the table.

The problem now was that although it was now shown that the elements can be arranged in a regular pattern reflective of their chemical properties, there was no way to understand why the chemical properties of elements should be determined by their atomic weight. In 1913, the British chemist Henry Mosely showed by means of x-ray spectroscopy that the the number obtained from the serial ordering of elements in the periodic table corresponded to the number of elementary positive charges in the nucleus of each element. Since this number Z also corresponds to the number of electrons in an atom of each element. This means that the chemical properties of an element are determined by the number of electrons it contains. This gave rise to the periodic table in its modern form as shown in Fig. 15.1.

This table clearly shows that there is some underlying structure which determines the chemical properties of elements. The elements in the rightmost column are the noble gases which are the least chemically active of the elements. The elements which differ in atomic number from the noble gases by ± 1 are the most reactive, The alkalies in the far left-hand column and the halogens in the column next to the nobel gases.

A rough picture of why this occurs can be obtained from considering a simple model of the atom in which Z electrons are placed in the coulomb potential of a nucleus with Z protons. Since the nucleus is much heavier than the electrons, it is a reasonable approximation to treat it as though it is fixed in space. For simplicity we will ignore the coulomb interactions among the electrons. The hamiltonian for this system is then given by

$$\hat{H} = \sum_{i=1}^{Z} \left(\frac{\hat{\boldsymbol{p}}_i^2}{2m_e} - \frac{Ze^2}{r_i} \right) = \sum_{i=1}^{Z} \hat{H}_i \,. \tag{15.29}$$

Since this is simply a sum of hamiltonians for each of the individual electrons, the eigenenergies of the system are

$$E_{n_1,n_2,\ldots,n_{Z-1},n_Z} = \sum_{i=1}^{Z} E_{n_i} \tag{15.30}$$

and the corresponding eigenstates are

$$\Psi_{n_1,l_1,m_1;\ldots;n_Z,l_Z,m_Z}(\boldsymbol{r}_1,\ldots,\boldsymbol{r}_Z) = \prod_{i=1}^{Z} \psi_{n_i,l_i,m_i}(\boldsymbol{r}_i) \,. \tag{15.31}$$

Previously we showed that the individual eigenenergies of the electrons is given by

$$E_n = -\frac{Z^2 \alpha^2 m_e c^2}{2n^2} \tag{15.32}$$

Figure 15.1: The periodic table.

Table 15.1: Table of states in a coulomb potential for $n \leq 3$.

n	l	Spectroscopic notation	# of states $(2 \times (2l+1))$
1	0	1s	2
2	0	2s	2
	1	2p	6
3	0	3s	2
	1	3p	6
	2	3d	10

where for each $n \geq 1$, $0 \leq l \leq n - 1$. We can now develop a table of the states that are available for $n \leq 3$. If we note that the electrons of spin-$\frac{1}{2}$, the number of states for each n and l is $2 \times (2l+1)$. We also introduce the spectroscopic notation where $l = 0, 1, 2, 3, 4, 5, ...$ are designated by the letters s, p, d, f, g, etc. Thus the spectrascopic notation $3d$ refers to the state with $n = 3$ and $l = 3$.

Now consider the periodic table, Table 15.1. Note that there are 2 elements in the first row, 8 in the second row and 8 in the third row. From Table 15.1, we can see that there are 2 states in the for $n = 1$ and 8 states for $n = 2$ and 8 states in the $3s$ and $3p$ states. We can then imagine that we can build the atoms by successively placing 1 electron in each possible states. This called the atomic shell model. In this scheme, the electron configurations for each atom are shown in Table 15.2. The exponent in the configuration indicates how many electrons are in each state. Therefore, $(2p)^5$ means that there are 5 electrons in the $2p$ state. This table also shows the ionization energy for each atom. This is the amount of energy that is required to remove an electron from the neutral atom. The larger the ionization energy, the harder it is to remove an electron from the atom, and the less chemically reactive the atom is. Note that the ionization energy reaches a local maximum whenever the number of electrons equals the number of states for each shell nl. The largest jumps in ionization energy occur when adding 1 electron to a nobel gas to obtain an alkali.

Clearly, this is a very simplified model of the atom since we have ignored the interactions among the electrons. The more electrons there are, the worse this approximation becomes. In fact, this scheme is clearly breaking down since we have not included the 10 states associated with the $3d$ shell which must in someway be associated with the next row of the periodic table. In addition, we have made a major assumption, that we can only place one electron in each state designated by the quantum numbers n, l, m_l, m_s. The lowest energy of each atom would be obtained by placing all of the electrons in the $1s$ shell, but this doesn't seem to fit the pattern of the periodic table. Pauli solved the problem by fiat by stating that only one electron can exist in each state and that all others are excluded from a state by the presence

Table 15.2: Shell model configurations of simple atoms.

Z	element	configuration	ionization energy (eV)
1	H	$1s$	13.5
2	He	$(1s)^2$	24.6
3	Li	$(1s)^2(2s)$	5.4
4	Be	$(1s)^2(2s)^2$	9.3
5	B	$(1s)^2(2s)^2(2p)$	8.3
6	C	$(1s)^2(2s)^2(2p)^2$	11.3
7	N	$(1s)^2(2s)^2(2p)^3$	14.5
8	O	$(1s)^2(2s)^2(2p)^4$	13.6
9	F	$(1s)^2(2s)^2(2p)^5$	17.4
10	Ne	$(1s)^2(2s)^2(2p)^6$	21.6
11	Na	$(1s)^2(2s)^2(2p)^6(3s)$	5.1
12	Mg	$(1s)^2(2s)^2(2p)^6(3s)^2$	7.6
13	Al	$(1s)^2(2s)^2(2p)^6(3s)^2(3p)$	6.0
14	Si	$(1s)^2(2s)^2(2p)^6(3s)^2(3p)^2$	8.1
15	P	$(1s)^2(2s)^2(2p)^6(3s)^2(3p)^3$	11.0
16	S	$(1s)^2(2s)^2(2p)^6(3s)^2(3p)^4$	10.4
17	Cl	$(1s)^2(2s)^2(2p)^6(3s)^2(3p)^5$	13.0
18	Ar	$(1s)^2(2s)^2(2p)^6(3s)^2(3p)^6$	15.8

of an electron in the state. This called the *Pauli exclusion principle.*

15.2.1 The Permutation Operator

The problem now is to find a way that this arbitrary assertion can be accommodated within the formalism of quantum mechanics. Our discussion of this solution requires that we first discuss an important distinction between classical and quantum mechanical particles. Classical, macroscopic objects are distinguishable. This means for example that if we have two basket balls it is possible to tell them apart by marking one of the balls. By examining the balls for the mark we can distinguish which ball is which. The mark will not transport itself from one ball to the other and we can always be sure which ball we have at any time. On the other hand, quantum mechanical particles are indistinguishable. The only way to characterize a hydrogen atom is by the values of its quantum numbers. If we have two hydrogen atoms in their ground state which are confined in a small box, there is no way of distinguishing between between the atoms since they are measurably exactly the same. This a characteristic that must be built into our quantum mechanical descriptions of systems of two or more identical particles.

To illustrate how this is done, consider the case of 2 non-interacting, identical, spin-$\frac{1}{2}$ particles in a potential well where

$$\hat{H} = \frac{\hat{\boldsymbol{p}}_1^2}{2m} + V(r_1) + \frac{\hat{\boldsymbol{p}}_2^2}{2m} + V(r_2) \,. \tag{15.33}$$

The eigenstates are of the form

$$\psi_{\alpha_1, \alpha_2}(\boldsymbol{r}_1, \sigma_1; \boldsymbol{r}_2, \sigma_2) = \psi_{\alpha_1}(\boldsymbol{r}_1, \sigma_1)\psi_{\alpha_2}(\boldsymbol{r}_2, \sigma_2) \tag{15.34}$$

where the α_i represents the complete set of eigenvalues necessary to designate the one-body state and the σ_i are the indexes for the corresponding spinor component. For example each particle could correspond to a particle with spin-$\frac{1}{2}$ in a central potential. Then

$$\psi_\alpha(\boldsymbol{r}, \sigma) = \psi_{nlm}(\boldsymbol{r}) \left(\chi_s\right)_\sigma \,. \tag{15.35}$$

Then, $\alpha = \{n, l, m, s\}$ and σ refers to the component of the spinor (i.e. the row of the spinor). For convenience, we will use the shorthand notation $1 = \boldsymbol{r}_1, \sigma_1$, $2 = \boldsymbol{r}_2, \sigma_2$, etc. Then the two-particle state can be written as

$$\psi_{\alpha_1, \alpha_2}(\boldsymbol{r}_1, \sigma_1; \boldsymbol{r}_2, \sigma_2) = \psi_{\alpha_1, \alpha_2}(1, 2) = \psi_{\alpha_1}(1)\psi_{\alpha_2}(2) \,. \tag{15.36}$$

We can now introduce the permutation operator $\hat{\mathcal{P}}_{12}$ which when acting on the two-particle state exchanges the positions and the spinor component labels. That is

$$\hat{\mathcal{P}}_{12}\psi_{\alpha_1, \alpha_2}(1, 2) = \psi_{\alpha_1, \alpha_2}(2, 1) = \psi_{\alpha_1}(2)\psi_{\alpha_2}(1) \,. \tag{15.37}$$

Since a second application of the permutation operator will return the state to its original form, the permutation operator must have the property that

$$\hat{\mathcal{P}}_{12}\hat{\mathcal{P}}_{12} = 1\,. \tag{15.38}$$

Now consider the Schrödinger equation

$$\hat{H}\psi_{\alpha_1,\alpha_2}(1,2) = E\Psi_{\alpha_1,\alpha_2}(1,2)\,. \tag{15.39}$$

Introducing unity between the hamiltonian and the wave function on the left-hand side of the equation in the form of (15.38) and then multiplying the equation on the left by the permutation operator gives

$$\hat{\mathcal{P}}_{12}\hat{H}\hat{\mathcal{P}}_{12}^{-1}\left(\hat{\mathcal{P}}_{12}\psi_{\alpha_1,\alpha_2}(1,2)\right) = E\left(\hat{\mathcal{P}}_{12}\psi_{\alpha_1,\alpha_2}(1,2)\right)\,. \tag{15.40}$$

So the permuted states are eigenstates of $\hat{\mathcal{P}}_{12}\hat{H}\hat{\mathcal{P}}_{12}^{-1}$. This combination defines the operation of the permutation operators on the hamiltonian. This operation simply exchanges the labels of the coordinates and any spin operators in the equation. In the case of the simple hamiltonian that we are considering here which contains simple central potentials with no spin operators,

$$\hat{\mathcal{P}}_{12}\hat{H}\hat{\mathcal{P}}_{12}^{-1} = \frac{\hat{\boldsymbol{p}}_2^2}{2m} + V(2) + \frac{\hat{\boldsymbol{p}}_1^2}{2m} + V(1) = \hat{H}\,. \tag{15.41}$$

This can be restated as

$$\hat{\mathcal{P}}_{12}\hat{H} = \hat{H}\hat{\mathcal{P}}_{12} \tag{15.42}$$

or

$$\left[\hat{\mathcal{P}}_{12}, \hat{H}\right] = 0\,. \tag{15.43}$$

Since the permutation operator and the hamiltonian commute, the eigenstates of the permutation operator will also be eigenstates of the hamiltonian. We can now introduce two wave functions

$$\psi_{S\alpha_1,\alpha_2}(1,2) = \frac{1}{\sqrt{2}}\left(\psi_{\alpha_1}(1)\psi_{\alpha_2}(2) + \psi_{\alpha_1}(2)\psi_{\alpha_2}(1)\right) \tag{15.44}$$

and

$$\psi_{A\alpha_1,\alpha_2}(1,2) = \frac{1}{\sqrt{2}}\left(\psi_{\alpha_1}(1)\psi_{\alpha_2}(2) - \psi_{\alpha_1}(2)\psi_{\alpha_2}(1)\right) \tag{15.45}$$

From the definition of the permutation operator

$$\hat{\mathcal{P}}_{12}\psi_{S\alpha_1,\alpha_2}(1,2) = \psi_{S\alpha_1,\alpha_2}(1,2) \tag{15.46}$$

and

$$\hat{\mathcal{P}}_{12}\psi_{A\alpha_1,\alpha_2}(1,2) = -\psi_{A\alpha_1,\alpha_2}(1,2)\,. \tag{15.47}$$

Thus these states are eigenstates of the permutation operator, and ψ_S is symmetric or even under permutation and ψ_A is antisymmetric or odd under permutation.

Note that

$$
\begin{aligned}
\psi_{S\alpha_1,\alpha_1}(1,2) &= \frac{1}{\sqrt{2}} \left(\psi_{\alpha_1}(1)\psi_{\alpha_1}(2) + \psi_{\alpha_1}(2)\psi_{\alpha_1}(1) \right) \\
&= \sqrt{2}\psi_{\alpha_1}(1)\psi_{\alpha_1}(2) \tag{15.48}
\end{aligned}
$$

and

$$
\psi_{A\alpha_1,\alpha_1}(1,2) = \frac{1}{\sqrt{2}} \left(\psi_{\alpha_1}(1)\psi_{\alpha_1}(2) - \psi_{\alpha_1}(2)\psi_{\alpha_1}(1) \right) = 0 \,. \tag{15.49}
$$

This implies that for symmetric wave functions we can have two particles in the same state, while for antisymmetric wave functions it is impossible to have two particles in the same state. This now leads to a general classification of quantum mechanical particles.

1. Particles with angular momentum of $\frac{1}{2}$, $\frac{3}{2}$, $\frac{5}{2}$, etc., are called *fermions* and only one particle can occupy each quantum state. This requires that the wave functions of fermions be antisymmetric under the interchange of any two sets of coordinates and spin indices. These particles are said to satisfy Fermi-Dirac statistics.

2. All particles with angular momenta 0, 1, 2, 3, etc., are called *bosons* and any number of particles can occupy a given quantum state. The wave functions for bosons must be symmetric under the interchange of any two sets of coordinates and spin indices. These particles are said to satisfy Bose-Einstein statistics.

15.2.2 n Particles in a Potential Well

For the case of n identical particles in a potential well, it is possible to construct completely antisymmetric wave functions in a straightforward way. The hamiltonian for such a system would be

$$
\hat{H} = \sum_{i=1}^{n} \left[\frac{\hat{\boldsymbol{p}}_i^2}{2m} + V(r_i) \right] \,. \tag{15.50}
$$

Since this is the sum of n one-body hamiltonians, the energy will be the sum of the individual one-particle energies and the wave functions will be a product of the one-body wave functions. The completely antisymmetric wave function can be written as a *Slater determinant*

$$
\Psi_{A\alpha_1,\ldots,\alpha_n}(1,\ldots,n) = \frac{1}{\sqrt{n!}}
\begin{vmatrix}
\psi_{\alpha_1}(1) & \psi_{\alpha_1}(2) & \cdots & \psi_{\alpha_1}(n) \\
\psi_{\alpha_2}(1) & \psi_{\alpha_2}(2) & \cdots & \psi_{\alpha_2}(n) \\
\vdots & & & \\
\psi_{\alpha_n}(1) & \psi_{\alpha_n}(2) & \cdots & \psi_{\alpha_n}(n)
\end{vmatrix}
\tag{15.51}
$$

259

Note that if any to particles have the same quantum numbers, for example if $\alpha_1 = \alpha_2$ the matrix will have two identical rows and the determinant will be 0.

15.2.3 When is symmetrization necessary?

The requirement that the total wave functions of particles be either symmetric or antisymmetric raises the question of whether it is really necessary to symmetrization the wave functions of all of the electrons in the universe at the same time that we are trying to solve for the wave functions of the helium atom. Common sense suggests that particles at large distances from one another should be independent and that it should only be necessary to consider the symmetry of particles which are all located within small distances. We can see obtain a reasonable criterion for choosing whether it is necessary to symmetrize from consideration of a simple example.

Consider a system with two one-dimensional potential wells with their centers located a distance a apart. Assume initially that a is large. Place one particle in the left potential well described by the wave function $\psi_l(x_1)$ and a second particle in the right well with wave function $\psi_r(x_2)$. The unsymmetrized wave function for this pair of particles is

$$\Psi(x_1, x_2) = \psi_l(x_1)\psi_r(x_2) \,. \tag{15.52}$$

The probability that these will be a particle at the position x is given by the one-body number density

$$
\begin{aligned}
n(x) &= \int_{-\infty}^{\infty} dx_1 \int_{-\infty}^{\infty} dx_2 \, \Psi^*(x_1, x_2) \left[\delta(x_1 - x) + \delta(x_2 - x)\right] \Psi(x_1, x_2) \\
&= \int_{-\infty}^{\infty} dx_1 \int_{-\infty}^{\infty} dx_2 \, |\psi_l(x_1)|^2 \, |\psi_r(x_2)|^2 \left[\delta(x_1 - x) + \delta(x_2 - x)\right] \\
&= |\psi_l(x)|^2 + |\psi_r(x)|^2 \,.
\end{aligned}
\tag{15.53}
$$

Note that this density is normalized to the total number of particles

$$\int_{-\infty}^{\infty} dx \, n(x) = 2 \,. \tag{15.54}$$

Now consider the situation where the wave function has been antisymmetrized to give

$$\Psi(x_1, x_2) = \frac{1}{\sqrt{2}} \left(\psi_l(x_1)\psi_r(x_2) - \psi_l(x_2)\psi_r(x_1)\right) \,. \tag{15.55}$$

The corresponding particle number density is

$$
\begin{aligned}
n_A(x) &= \int_{-\infty}^{\infty} dx_1 \int_{-\infty}^{\infty} dx_2\, \Psi_A^*(x_1,x_2)\left[\delta(x_1-x)+\delta(x_2-x)\right]\Psi_A(x_1,x_2) \\
&= \frac{1}{2}\int_{-\infty}^{\infty} dx_1 \int_{-\infty}^{\infty} dx_2\, \Big[|\psi_l(x_1)|^2\,|\psi_r(x_2)|^2 + |\psi_l^*(x_2)|^2\,|\psi_r(x_1)|^2 \\
&\quad -\psi_l^*(x_1)\psi_r(x_1)\psi_r^*(x_2)\psi_l(x_2) - \psi_l^*(x_2)\psi_r(x_2)\psi_r^*(x_1)\psi_l(x_1)\Big] \\
&\quad \times [\delta(x_1-x)+\delta(x_2-x)] \\
&= |\psi_l(x)|^2 + |\psi_r(x)|^2 - 2\Re\psi_l^*(x)\psi_r(x)\int_{-\infty}^{\infty} dx_2\,\psi_r^*(x_2)\psi_l(x_2)\,. \quad (15.56)
\end{aligned}
$$

Notice that

$$
n_A(x) - n(x) = -2\Re\psi_l^*(x)\psi_r(x)\int_{-\infty}^{\infty} dx_2\,\psi_r^*(x_2)\psi_l(x_2)\,, \quad (15.57)
$$

where it should be noted that ψ_l and ψ_r are not orthogonal since they are determined from different hamiltonians. This means that antisymmetrization will only be important when this term is appreciable. This term depends on the overlap integral between the two wave functions ψ_l and ψ_r. Since the wave functions will then to fall off exponentially, this overlap integral will be small if the two wells are separated by a distance which is large compared to the sizes of the two wave functions.

As a specific example assume that the potentials are harmonic oscillator potentials and that they are located at $x = \pm\frac{a}{2}$. If we assume that each particle is in the ground state of its respective well, then

$$
\psi_l(x) = \left(\frac{b}{\pi}\right)^{\frac{1}{4}} e^{-\frac{b}{2}\left(x+\frac{a}{2}\right)^2} \quad (15.58)
$$

and

$$
\psi_r(x) = \left(\frac{b}{\pi}\right)^{\frac{1}{4}} e^{-\frac{b}{2}\left(x-\frac{a}{2}\right)^2}\,, \quad (15.59)
$$

where $b = \frac{m\omega}{\hbar}$. The overlap integral is then

$$
\begin{aligned}
\int_{-\infty}^{\infty} dx_2\,\psi_r^*(x_2)\psi_l(x_2) &= \sqrt{\frac{b}{\pi}} \int_{-\infty}^{\infty} dx_2\, e^{-\frac{b}{2}\left(x_2+\frac{a}{2}\right)^2} e^{-\frac{b}{2}\left(x_2-\frac{a}{2}\right)^2} \\
&= \sqrt{\frac{b}{\pi}} \int_{-\infty}^{\infty} dx_2\, e^{-bx^2} e^{-b\frac{a^2}{4}} = e^{-b\frac{a^2}{4}}\,. \quad (15.60)
\end{aligned}
$$

Clearly, for large a this will be very small and the effect of antisymmetrization will also be very small. In other words, we only need to worry about antisymmetrization in those cases where the overlap between wave functions is nontrivial.

Antisymmetrization of the wave functions has another consequence. Suppose that the particles in the potential wells of the above example have spin-$\frac{1}{2}$. For the purposes of discussion, assume that the potentials are infinite square wells. Suppose that for large separations the particles are both in the ground state of their respective potentials with $m_s = \frac{1}{2}$. If the two systems are then brought together such that the respective potentials correspond, the Pauli principle states that we cannot have two particles with identical quantum numbers. So if the spin is fixed, one of the particles must be raised to the first excited state. This has a larger energy than the original separated systems. This means that there is an effective repulsive force do to the antisymmetrization of the wave functions.

15.3 The Fermi Gas

A useful example of a many-body system is a Fermi gas. This model consists of N fermions placed in a cubical infinite well. The potential is then

$$V(\boldsymbol{r}) = \begin{cases} 0 & \text{for } 0 \leq x, y, z \leq L \\ \infty & \text{otherwise} \end{cases} . \qquad (15.61)$$

Since the hamiltonian can be written as a sum of three one-dimensional hamiltonians, the wave function with be a product of one-dimensional wave functions

$$\psi_{n_x,n_y,n_z}(\boldsymbol{r}) = \sqrt{\frac{8}{L^3}} \sin\left(\frac{n_x \pi x}{L}\right) \sin\left(\frac{n_y \pi y}{L}\right) \sin\left(\frac{n_z \pi z}{L}\right) \qquad (15.62)$$

and the energy is the sum of three energies

$$E_{n_x,n_y,n_z} = E_{n_x} + E_{n_y} + E_{n_z} = \frac{\hbar^2 \pi^2}{2mL^2} \left(n_x^2 + n_y^2 + n_z^2\right) , \qquad (15.63)$$

where $1 \leq n_x, n_y, n_z$. The state of a particle is then determined by the triplet of numbers (n_x, n_y, n_z) and all states where $n_x^2 + n_y^2 + n_z^2$ have the same values will be degenerate. It is convenient to define the wave number

$$k_n = \frac{\pi}{L} n . \qquad (15.64)$$

Note that

$$\Delta k_n = k_n - k_{n-1} = \frac{\pi}{L} . \qquad (15.65)$$

As a result, as the size of the box L increases the values of the wave number become closer and closer together such that at some large values of L we can approximate the spectrum by continuous values of k.

Now suppose that we have a large number of spin-$\frac{1}{2}$ particles and use then to fill all of the states up to some fixed energy E_f called the *Fermi energy*. The total energy will then be given by

$$E = 2 \sum_{n_x,n_y,n_z} \frac{\hbar^2 \pi^2}{2mL^2} \left(n_x^2 + n_y^2 + n_z^2 \right) \theta(n_f^2 - n_x^2 - n_y^2 - n_z^2),$$ (15.66)

where

$$E_f = \frac{\hbar^2 \pi^2}{2mL^2} n_f^2.$$ (15.67)

Using $1 = \frac{L}{\pi} \Delta k_n$ and the definition of the wave number, we can rewrite the energy as

$$E = 2 \sum_{n_x,n_y,n_z} \frac{L^3}{\pi^3} \Delta k_{n_x} \Delta k_{n_y} \Delta k_{n_z} \frac{\hbar^2}{2m} \left(k_{n_x}^2 + k_{n_y}^2 + k_{n_z}^2 \right) \theta(k_f^2 - k_{n_x}^2 - k_{n_y}^2 - k_{n_z}^2),$$ (15.68)

where k_f is the magnitude of the wave vector of the highest filled state

$$E_f = \frac{\hbar^2 k_f^2}{2m}.$$ (15.69)

Now as the box size L becomes large along with the total number of particles N, the triple sum for the total energy resembles a Reimann sum and can be approximated by the integral

$$E = 2 \frac{L^3}{\pi^3} \frac{\hbar^2}{2m} \int_0^\infty dk_x \int_0^\infty dk_y \int_0^\infty dk_z \left(k_x^2 + k_y^2 + k_z^2 \right) \theta(k_f^2 - k_x^2 - k_y^2 - k_z^2).$$ (15.70)

The lower limits of the integrals and the θ function limit the integral to one octant of the wave number space. Since the integrand depends only on the magnitude of the wave vector $k = \sqrt{k_x^2 + k_y^2 + k_z^2}$, we can change the integral to include all octants and divide by 8 to compensate giving

$$E = 2 \frac{1}{8} \frac{L^3}{\pi^3} \frac{\hbar^2}{2m} \int d^3 k \, k^2 \theta(k_f - k) = 2 \frac{V}{(2\pi)^3} \int d^3 k \frac{\hbar^2 k^2}{2m} \theta(k_f - k),$$ (15.71)

where

$$V = L^3$$ (15.72)

is the volume of the box. This integral can now be performed easily in spherical coordinates to give

$$E = 2 \frac{V}{(2\pi)^3} \frac{\hbar^2}{2m} 4\pi \int_0^{k_f} dk \, k^2 \, k^2 = \frac{\hbar^2 V}{2\pi^2 m} \frac{k_f^5}{5}$$ (15.73)

263

At this point we do not know exactly what value the fermi wave vector k_f has but we know that it is related to the total number of fermions N. We can count the total number of fermions by adding them up level by level as

$$N = 2 \sum_{n_x, n_y, n_z} \theta(n_f^2 - n_x^2 - n_y^2 - n_z^2).$$ (15.74)

Following the same steps as used in calculating the energy, we can approximate this sum by an integral in the case where the box size is large and the are a large number of fermions as

$$N = 2\frac{V}{(2\pi)^3} \int d^3k \, \theta(k_f - k) = 2\frac{V}{(2\pi)^3} 4\pi \int_0^{k_f} dk \, k^2 = \frac{V k_f^3}{3\pi^2}.$$ (15.75)

This can be solved for k_f in terms of N to give

$$k_f = \left(\frac{3\pi^2 N}{V}\right)^{\frac{1}{3}}.$$ (15.76)

The total energy of the fermi gas can now be rewritten as

$$E = \frac{\hbar^2 (3\pi^2 N)^{\frac{5}{3}}}{10\pi^2 m V^{\frac{2}{3}}}.$$ (15.77)

Note that for fixed N, the energy will increase as the volume decreases. This means that the fermi gas is exerting a pressure on the walls of the box. This called the degeneracy pressure and is given by

$$p_{\text{deg}} = -\frac{\partial E}{\partial V} = \frac{\hbar^2 (3\pi^2 N)^{\frac{5}{3}}}{15\pi^2 m V^{\frac{5}{3}}}.$$ (15.78)

15.3.1   The Bulk Modulus of a Conductor

A conductor is characterized by having a large number of electrons that are free to move about the conductor. As a first approximation, the conductor can be viewed as a degenerate Fermi gas. The bulk modulus is defined as

$$B = -V\frac{\partial p}{\partial V},$$ (15.79)

where p is the pressure applied to the material. In this case, the applied pressure is equal to the degeneracy pressure and

$$B = \frac{\hbar^2 (3\pi^2 N)^{\frac{5}{3}}}{9\pi^2 m V^{\frac{5}{3}}}$$ (15.80)

For copper, the electron density $n_e = \frac{N}{V} = 8.47 \times 10^{28}$ electrons/m³. The expression for the bulk modulus of the degenerate Fermi gas is then $B = 6.5 \times 10^{10}$ N/m² whereas the experimental result is $B = 14 \times 10^{10}$ N/m². So the degenerate Fermi gas provides a rough estimate of the bulk modulus for a conductor.

15.4 The Deuteron

15.4.1 Isospin

The simplest nontrivial nucleus is the deuteron which consists of a bound state of a proton and a neutron. The deuteron has a mass of M_d=1875.612762(75) MeV. The mass of the proton is m_p=938.27200(4) MeV and the mass of the neutron is m_n=939.56533(4) MeV. The binding energy of the deuteron is then $E = -2.225$ MeV. The deuteron has no bound excited states.

Both the neutron and the proton are spin-$\frac{1}{2}$ particles and differ in mass by only about a 0.1 percent. It is convenient to treat the proton and neutron of two states of the same particle, the nucleon. The quantum number that determines the charge state of the nucleon is called the *isospin*. Since there are two charge states there must be two values of the isospin quantum numbers. This is reminiscent of the two magnetic quantum numbers associated with a spin-$\frac{1}{2}$ particle. The isospin is an intrinsic property of the nucleon that is treated in the same way as spin. The nucleons are assigned an isospin of $\frac{1}{2}$ which is represented by two-dimensional matrix operators $\hat{T}_i$, $i = 1, 2, 3$, which satisfy the commutation relations

$$\left[\hat{T}_i, \hat{T}_j\right] = i\epsilon_{ijk}\hat{T}_k \,. \tag{15.81}$$

These operators can be written as

$$\hat{T}_i = \tfrac{1}{2}\tau_i \tag{15.82}$$

where the τ_i are the same as the Pauli matrices σ_i and have been renamed to avoid confusion since the nucleons have both spin and isospin. The two isospin states of the nucleon are then represented by the two-dimensional spinors

$$\zeta_{\frac{1}{2}} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \tag{15.83}$$

and

$$\zeta_{-\frac{1}{2}} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{15.84}$$

where $\zeta_{\frac{1}{2}}$ represents the proton and $\zeta_{-\frac{1}{2}}$ represents the neutron. The charge of the nucleon (in units of e) is determined by the operator

$$\hat{e} = \frac{1}{2}\left(1 + \tau_3\right) \tag{15.85}$$

since

$$\hat{e}\zeta_{\frac{1}{2}} = 1 \tag{15.86}$$

and

$$\hat{e}\zeta_{-\frac{1}{2}} = 0 \,. \tag{15.87}$$

Since the isospin operators have the same algebra as angular momentum operators (with $\hbar = 1$), isospin can be added in exactly the same way as angular momentum using all of the same rules and Clebsch-Gordan coefficients.

15.4.2   A Simple Model

The deuteron can now be viewed as a bound state of two nucleons with a total charge of e. Since these nucleons are identical particles with spin-$\frac{1}{2}$ and isospin-$\frac{1}{2}$ it is necessary that the total wave function be antisymmetric. To see how this works, consider the simple case where the nucleon-nucleon interaction is described by a central potential so that the hamiltonian is

$$\hat{H} = \frac{\hat{\boldsymbol{p}}_1^2}{2m_N} + \frac{\hat{\boldsymbol{p}}_2^2}{2m_N} + V(|\boldsymbol{r}_1 - \boldsymbol{r}_2|), \tag{15.88}$$

where the nucleon mass is taken to be $m_N = \frac{1}{2}(m_p + m_n)$. Since this is a two-body problem the hamiltonian can be separated into relative and center-of-mass coordinates and reduced to an effective one-body equation in the relative coordinate with the effective hamiltonian

$$\hat{H}_{rel} = \frac{\hat{\boldsymbol{p}}^2}{2\mu} + V(r), \tag{15.89}$$

where $\mu = \frac{m_N}{2}$. This can now be solved in spherical coordinates as usual to yield the product of a radial wave function and a spherical harmonic.

The complete wave function must now consist of the spatial parts, and the spinors and isospinors for the two nucleons. For reasons that will soon become apparent, it is convenient to construct wave functions with the individual spins coupled to a total spin, and the isospins coupled to a total isospin. Doing this, the relative wave function is

$$\begin{aligned}
\Psi_{nlmSM_STM_T}(\boldsymbol{r}) &= R_{nl}(r)Y_{lm}(\theta,\phi)\sum_{m_{s_1}}\sum_{m_{s_2}}\left\langle \tfrac{1}{2}, m_{s_1}; \tfrac{1}{2}, m_{s_2} \,|\, S, M_S\right\rangle \chi^{(1)}_{m_{s_1}}\chi^{(2)}_{m_{s_2}} \\
&\times \sum_{m_{t_1}}\sum_{m_{t_2}}\left\langle \tfrac{1}{2}, m_{t_1}; \tfrac{1}{2}, m_{t_2} \,|\, T, M_T\right\rangle \zeta^{(1)}_{m_{t_1}}\zeta^{(2)}_{m_{t_2}}. \tag{15.90}
\end{aligned}$$

Now, consider what happens when we apply the permutation operator $\mathcal{P}_{12}$ to the wave function. First, this requires that we make the transformation $\boldsymbol{r}_1 \leftrightarrow \boldsymbol{r}_2$ which means that $\boldsymbol{r} \to -\boldsymbol{r}$. This, in turn, requires that $\theta \to \pi - \theta$ and $\phi \to \phi + \pi$ while the magnitude r remains unchanged. Using the basic properties of the spherical harmonics

$$Y_{lm}(\pi - \theta, \phi + \pi) = (-1)^l Y_{lm}(\theta, \phi). \tag{15.91}$$

Next, the permutation operator requires that the spinor indices be exchanged. This is equivalent to interchanging the superscripts of the spinors. That is $(1) \leftrightarrow (2)$. Since both particles are spin $\frac{1}{2}$, this means that the total spin can be either 0 or 1,

where the spin-0 state is called a singlet state and the spin-1 states are called spin triplet states. We have in fact already constructed these states using the the raising and lowering operators. These give

$$\sum_{m_{s_1}}\sum_{m_{s_2}} \langle \tfrac{1}{2}, m_{s_1}; \tfrac{1}{2}, m_{s_2} | 0, 0 \rangle \chi^{(1)}_{m_{s_1}} \chi^{(2)}_{m_{s_2}} = \frac{1}{\sqrt{2}} \left(\chi^{(1)}_{\frac{1}{2}} \chi^{(2)}_{-\frac{1}{2}} - \chi^{(1)}_{-\frac{1}{2}} \chi^{(2)}_{\frac{1}{2}} \right), \tag{15.92}$$

$$\sum_{m_{s_1}}\sum_{m_{s_2}} \langle \tfrac{1}{2}, m_{s_1}; \tfrac{1}{2}, m_{s_2} | 1, 1 \rangle \chi^{(1)}_{m_{s_1}} \chi^{(2)}_{m_{s_2}} = \chi^{(1)}_{\frac{1}{2}} \chi^{(2)}_{\frac{1}{2}}, \tag{15.93}$$

$$\sum_{m_{s_1}}\sum_{m_{s_2}} \langle \tfrac{1}{2}, m_{s_1}; \tfrac{1}{2}, m_{s_2} | 1, 0 \rangle \chi^{(1)}_{m_{s_1}} \chi^{(2)}_{m_{s_2}} = \frac{1}{\sqrt{2}} \left(\chi^{(1)}_{\frac{1}{2}} \chi^{(2)}_{-\frac{1}{2}} + \chi^{(1)}_{-\frac{1}{2}} \chi^{(2)}_{\frac{1}{2}} \right) \tag{15.94}$$

and

$$\sum_{m_{s_1}}\sum_{m_{s_2}} \langle \tfrac{1}{2}, m_{s_1}; \tfrac{1}{2}, m_{s_2} | 1, -1 \rangle \chi^{(1)}_{m_{s_1}} \chi^{(2)}_{m_{s_2}} = \chi^{(1)}_{-\frac{1}{2}} \chi^{(2)}_{-\frac{1}{2}}. \tag{15.95}$$

Note that the $S = 0$, or singlet state, is odd under the interchange of the superscripts, while the $S = 1$ triplet states are even under the interchange.

Finally, the permutation operator will also require that the isospinor indices be interchanged which is equivalent to interchanging the superscripts of the isospinors. Since the isospins are combined exactly like the spins, there will be an isoscalar $T = 0$ state which will be odd under the exchange of superscripts and isotriplet $T = 1$ states which will be even under the interchange. From the preceding arguments, we can write that

$$\begin{aligned} \mathcal{P}_{12} \Psi_{nlmSM_S TM_T}(\boldsymbol{r}) &= (-1)^l (-1)^{S+1} (-1)^{T+1} \Psi_{nlmSM_S TM_T}(\boldsymbol{r}) \\ &= (-1)^{l+S+T} \Psi_{nlmSM_S TM_T}(\boldsymbol{r}). \end{aligned} \tag{15.96}$$

Therefore, the Pauli principle requires that $l + S + T$ must be odd. Since the lowest energy state will occur for the s wave $l = 0$, this requires that $S + T$ must be odd. It has been shown experimentally that the total angular momentum of the deuteron is 1, so $S = 1$. This then requires that $T = 0$. So the deuteron is an spin triplet, isosinglet, s wave in this model.

A simple, relatively easily solvable model would be to use the potential for a finite spherical well where

$$V(r) = \begin{cases} -V_0 & \text{for } r \le a \\ 0 & \text{for } r > a \end{cases}. \tag{15.97}$$

Using our previous solution of the spherical well with $E \to E + V_0$, the radial wave function for $l = 0$ is

$$R_{10}(r) = N \left(\frac{\sin(kr)}{kr} \theta(a - r) + \frac{\beta \sin(ka)}{ke^{-\beta a}} \frac{e^{-\beta r}}{\beta r} \theta(r - a) \right), \tag{15.98}$$

where

$$k = \sqrt{\frac{2\mu}{\hbar^2}(E + V_0)},$$ (15.99)

$$\beta = \sqrt{-\frac{2\mu}{\hbar^2}E}$$ (15.100)

and

$$N = \sqrt{\frac{2k^3}{ka - \cos ka \sin ka + \frac{k}{\beta}\sin^2 ka}}.$$ (15.101)

The eigencondition for obtaining the energy is

$$k \cot ka + \beta = 0.$$ (15.102)

There are two free parameters, V_0 and a that have to be determined. We, therefore, need to have two measurable quantities to determine the values of these parameters. One obvious choice is to require that we obtain the experimental binding energy $E = -2.225$ MeV. For the second observable we can choose the root-mean-square radius of the deuteron which has been determined to be $r_{\text{rms}} = 1.98$ fm from a combination of theory and experiment. Using the wave function we can calculate the mean-square separation of the nucleons in the deuteron as

$$\begin{aligned} \langle r^2 \rangle &= \int_0^\infty dr\, r^4 R_{10}^2(r) \\ &= N^2 \left\{ \frac{1}{12k^5} \left[2k^2 a^2 \left(ka - 3\cos ka \sin ka \right) + 3ka \left(1 - 2\cos^2 ka \right) \right. \right. \\ &\quad \left. \left. + \cos ka \sin ka \right] + \frac{\sin^2 ka}{2k^2 \beta^3} \left(\beta^2 a^2 + \beta a + \frac{1}{2} \right) \right\}. \end{aligned}$$ (15.103)

This is related to r_{rms} by

$$r_{\text{rms}} = \frac{1}{2}\sqrt{\langle r^2 \rangle}.$$ (15.104)

Adjusting V_0 and a to obtain the experimental values of E and r_{rms} gives $V_0 = 30.0$ MeV and $a = 2.25$ fm.

We can now calculate the charge density of the deuteron (in units of e) which is given by

$$\rho(\boldsymbol{r}) = \int d^3 r' \Psi_{1001M_S00}^\dagger(\boldsymbol{r}') \left[\hat{e}^{(1)} \delta\left(\boldsymbol{r} - \frac{\boldsymbol{r}'}{2} \right) + \hat{e}^{(2)} \delta\left(\boldsymbol{r} + \frac{\boldsymbol{r}'}{2} \right) \right] \Psi_{1001M_S00}(\boldsymbol{r}).$$ (15.105)

It is easily verified that

$$\int d^3 r \rho(\boldsymbol{r}) = 1$$ (15.106)

as should be expected since the deuteron has one positive charge.

268

To evaluate this, first consider the isospin matrix element

$$\sum_{m'_{t_1}}\sum_{m'_{t_2}}\langle \tfrac{1}{2},m'_{t_1};\tfrac{1}{2},m'_{t_2}\left|0,0\right\rangle \zeta^{(1)\dagger}_{m'_{t_1}}\zeta^{(2)\dagger}_{m'_{t_2}}\hat{e}^{(1)}\sum_{m_{t_1}}\sum_{m_{t_2}}\langle \tfrac{1}{2},m_{t_1};\tfrac{1}{2},m_{t_2}\left|0,0\right\rangle \zeta^{(1)}_{m_{t_1}}\zeta^{(2)}_{m_{t_2}}$$

$$= \frac{1}{\sqrt{2}}\left(\zeta^{(1)\dagger}_{\frac{1}{2}}\zeta^{(2)\dagger}_{-\frac{1}{2}} - \zeta^{(1)\dagger}_{-\frac{1}{2}}\zeta^{(2)\dagger}_{\frac{1}{2}}\right)\hat{e}^{(1)}\frac{1}{\sqrt{2}}\left(\zeta^{(1)}_{\frac{1}{2}}\zeta^{(2)}_{-\frac{1}{2}} - \zeta^{(1)}_{-\frac{1}{2}}\zeta^{(2)}_{\frac{1}{2}}\right) = \frac{1}{2}. \qquad (15.107)$$

The same result is obtained for the matrix element of $\hat{e}^{(1)}$. Since the charge operator is independent of the nucleon spins, the spin matrix elements will be

$$\sum_{m'_{s_1}}\sum_{m'_{s_2}}\langle \tfrac{1}{2},m'_{s_1};\tfrac{1}{2},m'_{s_2}\left|1,M_S\right\rangle \chi^{(1)\dagger}_{m'_{s_1}}\chi^{(2)\dagger}_{m'_{s_2}}\sum_{m_{s_1}}\sum_{m_{s_2}}\langle \tfrac{1}{2},m_{s_1};\tfrac{1}{2},m_{s_2}\left|1,M_S\right\rangle \chi^{(1)}_{m_{s_1}}\chi^{(2)}_{m_{s_2}}$$

$$= 1. \qquad (15.108)$$

This can be obtained either by using the explicit expression for the three triplet spin wave functions, or by using the general property of the Clebsch-Gordan coefficients

$$\sum_{m_1 m_2}\langle j_1,m_1;j_2,m_2|\,J'M'\rangle\,\langle j_1,m_1;j_2,m_2|\,JM\rangle = \delta_{JJ'}\delta_{MM'}. \qquad (15.109)$$

Now, having evaluate the isospin and spin matrix elements, the charge density becomes

$$\begin{aligned}
\rho(\boldsymbol{r}) &= \int d^3r'\frac{1}{4\pi}R^2_{10}(r')\frac{1}{2}\left[\delta\left(\boldsymbol{r}-\frac{\boldsymbol{r'}}{2}\right)+\delta\left(\boldsymbol{r}+\frac{\boldsymbol{r'}}{2}\right)\right] \\
&= \frac{1}{8\pi}\int_0^\infty dr'r'^2\int_{-1}^1 d\cos\theta'\int_0^{2\pi}d\phi'R^2_{10}(r') \\
&\quad \times \left[\frac{1}{r^2}\delta\left(r-\frac{r'}{2}\right)\delta(cos\theta'-cos\theta)\delta(\phi'-\phi)\right. \\
&\quad \left.+\frac{1}{r^2}\delta\left(r-\frac{r'}{2}\right)\delta(cos(\pi-\theta')-cos\theta)\delta(\phi'+\pi-\phi)\right]. \qquad (15.110)
\end{aligned}$$

Since the wave function is independent of the angles, the angular integrals can be performed trivially using the δ-functions. Then using

$$\delta\left(r-\frac{r'}{2}\right) = 2\delta(r'-2r). \qquad (15.111)$$

the radial integral can also be performed using the δ-function to give

$$\rho(r) = \frac{2}{\pi}R^2_{10}(2r). \qquad (15.112)$$

Figure 15.2 shows the charge density distribution for the simple spherical well potential as compared to that obtained using a modern Argonne V18 (AV18) nucleon-nucleon potential. Although the charge density for the simple model is approximately
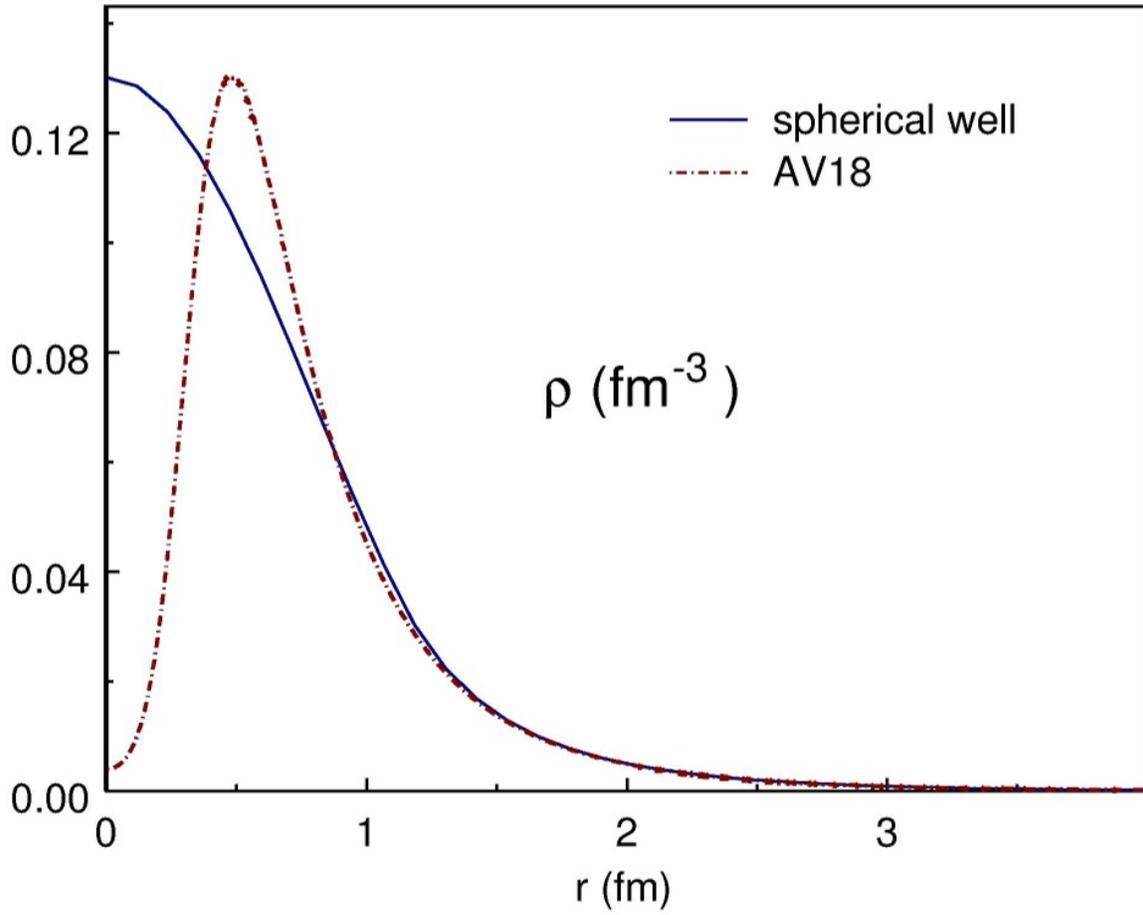
Figure 15.2: The deuteron charge density for the spherical well potential and for the modern AV18 nucleon-nucleon potential.

the same size due to the fact that we have fixed the radius to the physical value, it clearly does not have the correct behavior at small r. The fact that the charge density becomes small for small r is do to the strong repulsion of the nucleon-nucleon force for small separations. This is in part due to the fact that the nucleons are actually composite objects consisting of quarks and gluons. Since the quarks are spin-$\frac{1}{2}$ fermions, the Pauli principle results in a repulsion at small r.

15.4.3   The Nucleon-Nucleon Potential

The nucleon-nucleon potential is determined from a combination of theory and experiment. The long-range part of the interaction is believed to be the result of the exchange of π mesons. This is used to fix the long-range part of the force. The rest of the force is parameterized in terms of potentials multiplying various operators for angular momentum, spin and isospin. The parameters are then adjusted to fit certain properties of the deuteron and of nucleon-nucleon data. These potentials are now very sophisticated and are capable of producing a remarkable representation of the data. A minimal form of such a potential is given by

$$V(\boldsymbol{r}) = V_0(r) + V_\sigma(r)\boldsymbol{\sigma}^{(1)} \cdot \boldsymbol{\sigma}^{(2)} + V_{\mathrm{LS}}(r)\hat{\boldsymbol{L}} \cdot \hat{\boldsymbol{S}} + V_{\mathrm{T}}(r)S_{\mathrm{T}}(\boldsymbol{r}) \tag{15.113}$$

where the tensor operator is given by

$$S_{\mathrm{T}}(\boldsymbol{r}) = \frac{3}{r^2}\boldsymbol{\sigma}^{(1)} \cdot \boldsymbol{r}\,\boldsymbol{\sigma}^{(2)} \cdot \boldsymbol{r} - \boldsymbol{\sigma}^{(1)} \cdot \boldsymbol{\sigma}^{(2)} \,. \tag{15.114}$$

Each of the radial potentials $V_0(r)$, $V_\sigma(r)$, $V_{\mathrm{LS}}(r)$, and $V_{\mathrm{T}}(r)$ can also be separated into an isospin-independent part and an isospin-dependent part such that

$$V_i(r) = V_{i0}(r) + V_{i\tau}(r)\boldsymbol{\tau}^{(1)} \cdot \boldsymbol{\tau}^{(2)} \,. \tag{15.115}$$

Now, let's consider the various operators with the objectives of understanding their physical meaning and determining the constants of motion that can be used to label the wave function. First consider the operator $\boldsymbol{\sigma}^{(1)} \cdot \boldsymbol{\sigma}^{(2)} = 2(\sigma_-^{(1)}\sigma_+^{(2)} + \sigma_+^{(1)}\sigma_-^{(2)}) + \sigma_3^{(1)}\sigma_3^{(2)}$. The first two terms will be nonzero only when acting on states with opposite spins and will cause the spins of the two particles to be interchanged. The last term will be 1 for like spins and -1 for opposite spins, but will not change the spins. We have already demonstrated in a homework problem that $[\hat{\boldsymbol{S}}^2, \boldsymbol{\sigma}^{(1)} \cdot \boldsymbol{\sigma}^{(2)}] = 0$ and $[\hat{S}_z, \boldsymbol{\sigma}^{(1)} \cdot \boldsymbol{\sigma}^{(2)}] = 0$. So this term will not change the total spin or its projection in the z direction. The isospin operator $\boldsymbol{\tau}^{(1)} \cdot \boldsymbol{\tau}^{(2)}$ has similar properties.

The spin-orbit operator $\hat{\boldsymbol{L}} \cdot \hat{\boldsymbol{S}}$ does not commute with $\hat{L}_z$ or $\hat{S}_z$, but commutes with $\hat{\boldsymbol{L}}^2$, $\hat{\boldsymbol{S}}^2$, $\hat{\boldsymbol{J}}^2$ and $\hat{J}_z$ where $\hat{\boldsymbol{J}} = \hat{\boldsymbol{L}} + \hat{\boldsymbol{S}}$.

The tensor operator S_{T} can be shown to commute with $\hat{\boldsymbol{J}}^2$, $\hat{J}_z$ and $\hat{\boldsymbol{S}}^2$, but $[\hat{\boldsymbol{L}}^2, S_{\mathrm{T}}] \neq 0$. This means that the orbital angular momentum quantum number is not a good quantum number for deuteron. The wave function will then be labeled

by the quantum numbers associated with $\hat{H}$, $\hat{\boldsymbol{J}}^2$, $\hat{J}_z$ and $\hat{\boldsymbol{S}}^2$, $\hat{\boldsymbol{T}}^2$ and $\hat{T}_z$. This means that the deuteron wave function will look like

$$
\begin{aligned}
\Psi_{nSJM_JTM_T}(\boldsymbol{r}) \;=\; & \sum_l \sum_m \sum_{M_s} \langle l,m; S,M_s|\, J,M_J\rangle\, R_{nl}(r) Y_{lm}(\theta,\phi) \\
& \times \sum_{m_{s_1}} \sum_{m_{s_2}} \langle \tfrac{1}{2}, m_{s_1}; \tfrac{1}{2}, m_{s_2} \,|S,M_S\rangle\, \chi^{(1)}_{m_{s_1}} \chi^{(2)}_{m_{s_2}} \\
& \times \sum_{m_{t_1}} \sum_{m_{t_2}} \langle \tfrac{1}{2}, m_{t_1}; \tfrac{1}{2}, m_{t_2} \,|T,M_T\rangle\, \zeta^{(1)}_{m_{t_1}} \zeta^{(2)}_{m_{t_2}}. \qquad (15.116)
\end{aligned}
$$

Applying the interchange operator to this wave function gives

$$
\begin{aligned}
\mathcal{P}_{12}\Psi_{nSJM_JTM_T}(\boldsymbol{r}) \;=\; & \sum_l \sum_m \sum_{M_s} \langle l,m; S,M_s|\, J,M_J\rangle\, (-1)^{l+S+T} R_{nl}(r) Y_{lm}(\theta,\phi) \\
& \times \sum_{m_{s_1}} \sum_{m_{s_2}} \langle \tfrac{1}{2}, m_{s_1}; \tfrac{1}{2}, m_{s_2} \,|S,M_S\rangle\, \chi^{(1)}_{m_{s_1}} \chi^{(2)}_{m_{s_2}} \\
& \times \sum_{m_{t_1}} \sum_{m_{t_2}} \langle \tfrac{1}{2}, m_{t_1}; \tfrac{1}{2}, m_{t_2} \,|T,M_T\rangle\, \zeta^{(1)}_{m_{t_1}} \zeta^{(2)}_{m_{t_2}}. \qquad (15.117)
\end{aligned}
$$

The wave function will be antisymmetric only if $l + S + T$ is odd. The total angular momentum of the deuteron (its spin) is $J = 1$. The first term in the sum over l is $l = 0$. In this case $S = 1$ is the only allowed value and since S is a good quantum number, all other terms must also have $S = 1$. We have also shown that for this case $T = 0$ and since it is also a good quantum number, this must be true overall. Now with $J = 1$ and $S = 1$ the allowed values of l are $l = 0, 1, 2$. However, the wave function will only be antisymmetric if l is even, so the allowed values are $l = 0, 2$. Thus the deuteron will have both s- and d-wave contributions. The evidence that the deuteron has a d-wave contribution is that the deuteron has an charge quadrapole moment which can only occur if the deuteron is not spherically symmetric.

Chapter 16

Scattering Theory

One of the primary tools for studying quantum mechanical systems is the scattering of photons, electrons, nucleons, and atomic ions from these systems. We will now develop the basic framework of scattering theory and apply it to a simple model as a way of learning how useful information can be extracted from scattering data.

For the purposes of introducing the formalism, we will assume that the scattering will be from some fixed object that can be represented in terms of a potential $V(\boldsymbol{r})$. This can be easily generalized to allow for scattering from a free particle and we will discuss this later. We will assume that this potential is localized. That is, we assume that

$$\lim_{r \to \infty} rV(\boldsymbol{r}) = 0 \,. \tag{16.1}$$

The case of a long range potential such as the coulomb potential must be treated separately. Now assume that a beam of particles is prepared at some large distance (compared to the size of the scatterer) to the left of the scattering center. This beam will generally be large in cross section compared to the size of the scattering center. Since the potential is of short range, the initial beam at large distance can be treated as a plane wave moving from left to right. As the beam approaches the scattering center and begins to experience the potential, it will be distorted by the potential and part of the beam will be scattered away from the initial direction. This scattered beam continues outward from the scattering center and eventually moves beyond the range of the potential. At large distances from the scattering center, the scattered wave must again be a solution to the Schrödinger equation without an interaction.

In Section 7.3 we obtained the solutions to the Schrödinger equation for a free particle in spherical coordinates. There we found that the radial wave functions are given by spherical Bessel functions. The particular choice of Bessel function depends on the boundary condition. For the scattering case, the particle is free only at large distances, so we do not need to consider solutions that are regular at the origin. However, we want to have solutions that are incoming or outgoing waves. This means that the correct choice of spherical Bessel function for this case is the spherical Hankel function of the first kind $h^{(1)}(x)$ and the wave function at large distances has

the form

$$\psi_{\boldsymbol{k},lm}(\boldsymbol{r}) = A_l h^{(1)}(kr) Y_{lm}(\theta, \phi) \,. \tag{16.2}$$

For $x >> l$, the Hankel function behaves as

$$h_l^{(1)}(x) \sim -\frac{i}{x} e^{i\left(x - \frac{l\pi}{2}\right)} \,. \tag{16.3}$$

So the wave function for all values of l behaves as an outgoing spherical wave at large r.

With this in mind, we can specify a general form for the solution of the Schrödinger equation for our scattering problem. We expect the wave function to be of the form

$$\psi(\boldsymbol{r}) = \psi_{\text{inc}}(\boldsymbol{r}) + \psi_{\text{scat}}(\boldsymbol{r}) \,, \tag{16.4}$$

where

$$\psi_{\text{inc}}(\boldsymbol{r}) = \frac{1}{\sqrt{V}} e^{i\boldsymbol{k}\cdot\boldsymbol{r}} \tag{16.5}$$

is the incident wave and $\psi_{\text{scat}}(\boldsymbol{r})$ is the outgoing scattered wave. Furthermore, at large distances we expect that

$$\psi(\boldsymbol{r}) \rightarrow \frac{1}{\sqrt{V}} \left(e^{i\boldsymbol{k}\cdot\boldsymbol{r}} + f_k(\theta, \phi) \frac{e^{ikr}}{r} \right) \,, \tag{16.6}$$

where the box normalization described in Chapter 6 has been used. The function $f_k(\theta, \phi)$ is called the *scattering amplitude* and is in general a complex function.

16.1 The Scattering Cross Section

The quantity that is usually given to describe scattering is the *cross section*. For a beam of incident classical particles striking a classical target, only those particles that actually came into mechanical contact with the target would scatter from it. The number of scattered particles would then be proportional to the number of particles per unit area in the beam and the cross sectional area of the target in a plane perpendicular to the beam. By knowing the number of particles per unit area per unit time in the beam and counting the number of particles per unit time that are scattered by the target, the cross sectional area of the target can be determined. This can be generalized from a target which is some solid object to scattering of particles from a potential such as a gravitational field, or to the scattering of electromagnetic waves from some object.

To generalize this quantity to describe quantum mechanical scattering consider the situation where with a given incident beam scattered particles are detected in some device at a large distance form the scattering center at an angular position given by θ and ϕ that has a cross sectional area $d\boldsymbol{S}$, where the vector denotes a vector

along the radial vector locating the detector. The area of the detector then subtends a solid angle $d\Omega$. The incident beam of particles has a flux I_{inc} equal to the number of particles in the beam crossing a unit area perpendicular to the beam per unit time. The number of particles scattered into the solid angle $d\Omega$ per unit time is $I_{\text{inc}}d\sigma$ where $d\sigma$ is the differential cross section. The number of particles scattered into $d\Omega$ per unit time is the scattered flux I_{scat} times the area dS. The differential cross section is then given by

$$d\sigma = \frac{I_{\text{scat}}dS}{I_{\text{inc}}} . \tag{16.7}$$

This has units of area as is to be expected.

The incident flux is found by examining the beam at a large distance to the left of the target where the effect of the potential is negligible. The probability density of this wave is

$$\rho_{\text{inc}}(\boldsymbol{r}) = \psi^*_{\text{inc}}(\boldsymbol{r})\psi_{\text{inc}}(\boldsymbol{r}) = \frac{1}{V}e^{-i\boldsymbol{k}\cdot\boldsymbol{r}}e^{i\boldsymbol{k}\cdot\boldsymbol{r}} = \frac{1}{V} . \tag{16.8}$$

The probability current current density is then given by

$$
\begin{aligned}
\boldsymbol{j}_{\text{inc}}(\boldsymbol{r}) &= \frac{\hbar}{2im}\left[\psi^*_{\text{inc}}(\boldsymbol{r})\boldsymbol{\nabla}\psi_{\text{inc}}(\boldsymbol{r}) - \left(\boldsymbol{\nabla}\psi^*_{\text{inc}}(\boldsymbol{r})\right)\psi_{\text{inc}}(\boldsymbol{r})\right] \\
&= \frac{\hbar}{2imV}\left[e^{-i\boldsymbol{k}\cdot\boldsymbol{r}}\boldsymbol{\nabla}e^{i\boldsymbol{k}\cdot\boldsymbol{r}} - \left(\boldsymbol{\nabla}e^{-i\boldsymbol{k}\cdot\boldsymbol{r}}\right)e^{-i\boldsymbol{k}\cdot\boldsymbol{r}}\right] \\
&= \frac{\hbar}{2imV}\left[i\boldsymbol{k} - (-i\boldsymbol{k})\right] = \frac{\hbar\boldsymbol{k}}{mV} .
\end{aligned}
\tag{16.9}
$$

If the beam is along the z axis, the incident flux will then be

$$I_{\text{inc}} = \boldsymbol{e}_z \cdot \boldsymbol{j}_{\text{inc}} = \frac{\hbar k}{mV} = \rho_{\text{inc}}(\boldsymbol{r})v . \tag{16.10}$$

where v is the speed of the particles in the beam.

Using $d\boldsymbol{S} = r^2 d\Omega \boldsymbol{e}_r$, the flux scattered into $d\boldsymbol{S}$ will be

$$I_{\text{scat}}dS = d\boldsymbol{S} \cdot \boldsymbol{j}_{\text{scat}} = r^2\,d\Omega\,\boldsymbol{e}_r \cdot \boldsymbol{j}_{\text{scat}} . \tag{16.11}$$

We can now calculate

$$
\begin{aligned}
\boldsymbol{e}_r \cdot \boldsymbol{j}_{\text{scat}}(\boldsymbol{r}) &= \hat{r} \cdot \frac{\hbar}{2im}\left[\psi^*_{\text{scat}}(\boldsymbol{r})\boldsymbol{\nabla}\psi_{\text{scat}}(\boldsymbol{r}) - \left(\boldsymbol{\nabla}\psi^*_{\text{scat}}(\boldsymbol{r})\right)\psi_{\text{scat}}(\boldsymbol{r})\right] \\
&= \frac{\hbar}{2im}\left[\psi^*_{\text{scat}}(\boldsymbol{r})\frac{\partial}{\partial r}\psi_{\text{scat}}(\boldsymbol{r}) - \left(\frac{\partial}{\partial r}\psi^*_{\text{scat}}(\boldsymbol{r})\right)\psi_{\text{scat}}(\boldsymbol{r})\right] \\
&= \frac{\hbar}{2imV}|f_k(\theta,\phi)|^2\left[\frac{e^{-ikr}}{r}\frac{\partial}{\partial r}\frac{e^{ikr}}{r} - \left(\frac{\partial}{\partial r}\frac{e^{-ikr}}{r}\right)\frac{e^{ikr}}{r}\right] \\
&= \frac{\hbar}{2imV}|f_k(\theta,\phi)|^2\left[\frac{1}{r}\left(\frac{ik}{r} - \frac{1}{r^2}\right) - \left(-\frac{ik}{r} - \frac{1}{r^2}\right)\frac{1}{r}\right] \\
&= \frac{\hbar k}{mr^2V}|f_k(\theta,\phi)|^2 = \rho_{\text{inc}}(\boldsymbol{r})\frac{\hbar k}{mr^2}|f_k(\theta,\phi)|^2 .
\end{aligned}
\tag{16.12}
$$

So,

$$I_{\text{scat}} dS = \rho_{\text{inc}}(\boldsymbol{r}) \frac{\hbar k}{m} |f_k(\theta, \phi)|^2 \, d\Omega \,. \tag{16.13}$$

We can now determine the differential cross section to be

$$d\sigma = \frac{\rho_{\text{inc}}(\boldsymbol{r}) \frac{\hbar k}{m} |f_k(\theta, \phi)|^2 \, d\Omega}{\rho_{\text{inc}}(\boldsymbol{r}) \frac{\hbar k}{m}} = |f_k(\theta, \phi)|^2 \, d\Omega \,. \tag{16.14}$$

The differential cross section per unit solid angle is then

$$\frac{d\sigma}{d\Omega} = |f_k(\theta, \phi)|^2 \,. \tag{16.15}$$

The total cross section is

$$\sigma = \int d\Omega \frac{d\sigma}{d\Omega} = \int d\Omega \, |f_k(\theta, \phi)|^2 \,. \tag{16.16}$$

Clearly, the problem is to calculate the scattering amplitude. Before we do this, we will first explore a general property of the scattering amplitude.

16.2   The Optical Theorem

A general property of the scattering amplitude, called the *Optical Theorem*, can be obtained by starting with the continuity equation for probability

$$\boldsymbol{\nabla} \cdot \boldsymbol{j}(\boldsymbol{r}, t) + \frac{\partial}{\partial t} \rho(\boldsymbol{r}, t) = 0 \,. \tag{16.17}$$

For a stationary state,

$$\rho(\boldsymbol{r}, t) = \psi^*(\boldsymbol{r}, t)\psi(\boldsymbol{r}, t) = e^{\frac{i}{\hbar}Et}\psi^*(\boldsymbol{r})\psi(\boldsymbol{r})e^{\frac{i}{\hbar}Et} = \psi^*(\boldsymbol{r})\psi(\boldsymbol{r}) = \rho(\boldsymbol{r}) \,. \tag{16.18}$$

Since the probability density is independent of time, the continuity equation becomes

$$\boldsymbol{\nabla} \cdot \boldsymbol{j}(\boldsymbol{r}, t) = 0 \,. \tag{16.19}$$

We can now integrate this expression over some large sphere centered at the scattering center. Then

$$0 = \int_V d^3 r \boldsymbol{\nabla} \cdot \boldsymbol{j}(\boldsymbol{r}, t) = \int_S d\boldsymbol{S} \cdot \boldsymbol{j}(\boldsymbol{r}, t) = \lim_{r \to \infty} \int d\Omega r^2 \boldsymbol{e}_r \cdot \boldsymbol{j}(\boldsymbol{r}, t) \,, \tag{16.20}$$

where we have used the divergence theorem to write the volume integral of the divergence as a surface integral. Since the radius will be very large, we can use the asymptotic form of the wave function to calculate the probability current density. So,

$$\begin{aligned}
0 = & \lim_{r \to \infty} \frac{1}{V} \int d\Omega r^2 \left[\left(e^{-i\boldsymbol{k} \cdot \boldsymbol{r}} + f_k^*(\theta, \phi) \frac{e^{-ikr}}{r} \right) \frac{\partial}{\partial r} \left(e^{i\boldsymbol{k} \cdot \boldsymbol{r}} + f_k(\theta, \phi) \frac{e^{ikr}}{r} \right) \right. \\
& \left. - \left(\frac{\partial}{\partial r} \left(e^{-i\boldsymbol{k} \cdot \boldsymbol{r}} + f_k^*(\theta, \phi) \frac{e^{-ikr}}{r} \right) \right) \left(e^{i\boldsymbol{k} \cdot \boldsymbol{r}} + f_k(\theta, \phi) \frac{e^{ikr}}{r} \right) \right]
\end{aligned} \tag{16.21}$$

Using

$$\frac{\partial}{\partial r}\left(e^{i\boldsymbol{k}\cdot\boldsymbol{r}} + f_k(\theta,\phi)\frac{e^{ikr}}{r}\right) = ik\cos\theta e^{i\boldsymbol{k}\cdot\boldsymbol{r}} + f_k(\theta,\phi)\left(\frac{ik}{r} - \frac{1}{r^2}\right)e^{ikr},\qquad(16.22)$$

we have

$$
\begin{aligned}
0 &= \lim_{r\to\infty}\int d\Omega r^2\Bigg[ik\cos\theta + f_k(\theta,\phi)\left(\frac{ik}{r} - \frac{1}{r^2}\right)e^{ikr}e^{-i\boldsymbol{k}\cdot\boldsymbol{r}} + f_k^*(\theta,\phi)\frac{ik}{r}\cos\theta e^{-ikr}e^{i\boldsymbol{k}\cdot\boldsymbol{r}}\\
&\qquad + |f_k(\theta,\phi)|^2\frac{1}{r}\left(\frac{ik}{r} - \frac{1}{r^2}\right) + ik\cos\theta - f_k^*(\theta,\phi)\left(-\frac{ik}{r} - \frac{1}{r^2}\right)e^{-ikr}e^{i\boldsymbol{k}\cdot\boldsymbol{r}}\\
&\qquad + f_k(\theta,\phi)\frac{ik}{r}\cos\theta e^{ikr}e^{-i\boldsymbol{k}\cdot\boldsymbol{r}} - |f_k(\theta,\phi)|^2\frac{1}{r}\left(-\frac{ik}{r} - \frac{1}{r^2}\right)\Bigg]\\
&= \lim_{r\to\infty}\int d\Omega r^2\Bigg[2ik\cos\theta + f_k^*(\theta,\phi)\frac{ik}{r}\cos\theta e^{-ikr}e^{i\boldsymbol{k}\cdot\boldsymbol{r}} + f_k(\theta,\phi)\frac{ik}{r}\cos\theta e^{ikr}e^{-i\boldsymbol{k}\cdot\boldsymbol{r}}\\
&\qquad + f_k(\theta,\phi)\left(\frac{ik}{r} - \frac{1}{r^2}\right)e^{ikr}e^{-i\boldsymbol{k}\cdot\boldsymbol{r}} - f_k^*(\theta,\phi)\left(-\frac{ik}{r} - \frac{1}{r^2}\right)e^{-ikr}e^{i\boldsymbol{k}\cdot\boldsymbol{r}}\\
&\qquad + |f_k(\theta,\phi)|^2\frac{2ik}{r^2}\Bigg]\\
&= \lim_{r\to\infty}\int d\Omega r^2\Bigg[2ik\cos\theta + f_k^*(\theta,\phi)\left(\frac{ik}{r}\cos\theta + \frac{ik}{r} + \frac{1}{r^2}\right)e^{-ikr}e^{i\boldsymbol{k}\cdot\boldsymbol{r}}\\
&\qquad + f_k(\theta,\phi)\left(\frac{ik}{r}\cos\theta + \frac{ik}{r} - \frac{1}{r^2}\right)e^{ikr}e^{-i\boldsymbol{k}\cdot\boldsymbol{r}} + |f_k(\theta,\phi)|^2\frac{2ik}{r^2}\Bigg].\qquad(16.23)
\end{aligned}
$$

This can be simplified by noting that

$$\int d\Omega\cos\theta = 2\pi\int_{-1}^{1} d(\cos\theta)\cos\theta = 0.\qquad(16.24)$$

Next, consider an integral of the form

$$\int_{-1}^{1} d(\cos\theta)F(\cos\theta)e^{ikr\cos\theta} = \int_{-1}^{1} d\xi\, F(\xi)e^{ikr\xi}.\qquad(16.25)$$

Integrating this by parts gives

$$\int_{-1}^{1} d\xi\, F(\xi)e^{ikr\xi} = F(\xi)\frac{e^{ikr\xi}}{ikr}\bigg|_{-1}^{1} - \int_{-1}^{1} d\xi\, F'(\xi)\frac{e^{ikr\xi}}{ikr}.\qquad(16.26)$$

Integrating the righthand side again by parts gives

$$\int_{-1}^{1} d\xi\, F(\xi)e^{ikr\xi} = F(\xi)\frac{e^{ikr\xi}}{ikr}\bigg|_{-1}^{1} - F'(\xi)\frac{e^{ikr\xi}}{(ikr)^2}\bigg|_{-1}^{1} + \int_{-1}^{1} d\xi\, F''(\xi)\frac{e^{ikr\xi}}{(ikr)^2}.\qquad(16.27)$$

This shows that the first surface term falls of like $1/r$ and that the remaining contributions fall off at least as fast as $1/r^2$. Because of the limit of large radius for our sphere, we can now identify those terms which survive the limit. Therefore,

$$
\begin{aligned}
0 &= \int_0^{2\pi} d\phi\, f_k^*(\theta,\phi)\,(1+\cos\theta)\,e^{-ikr}e^{ikr\cos\theta}\big|_{\cos\theta=-1}^{+1} \\
&\quad - \int_0^{2\pi} d\phi\, f_k(\theta,\phi)\,(1+\cos\theta)\,e^{ikr}e^{-ikr\cos\theta}\big|_{\cos\theta=-1}^{+1} + \int d\Omega\,|f_k(\theta,\phi)|^2\,2ik \\
&= 2\int_0^{2\pi} d\phi f_k^*(0,\phi) - 2\int_0^{2\pi} d\phi f_k(\theta,\phi) + 2ik\int d\Omega \frac{d\sigma}{d\Omega}\,. \tag{16.28}
\end{aligned}
$$

For scattering from a spherically symmetric potential or scattering of unpolarized particles, the scattering amplitude is independent of the azimuthal angle ϕ. In this case the integrals over this angle simply reduce to a factor of 2π and the last term can be integrated trivially. So,

$$
0 = 4\pi f_k^*(0) - 4\pi f_k(0) + 2ik\sigma\,, \tag{16.29}
$$

or

$$
\Im(f_k(0)) = \frac{k}{4\pi}\sigma\,. \tag{16.30}
$$

This is called the *optical theorem*. This simply means that the imaginary part of the forward scattering amplitude is related to the total cross section.

16.3 Partial Waves and Phase Shifts

The objective now is to find a way to describe the scattering amplitude in terms of the wave functions that we can obtain from solving the time-independent Schrödinger equation. As a first step, recall that we can describe the solution to the three dimensional Schrödinger equation in free space in terms of spherical bessel functions. Since we are assuming that a some large distance the scattering potential is essentially zero, we can imagine that we enclose the potential in a large sphere that is concentric with the origin of the potential. The radius of the sphere is chosen to be sufficiently large that the potential is zero outside of the sphere. We can then find the solution of the Schrödinger equation, we solve the Schrödinger equation inside the sphere with the potential and then use the usual boundary conditions at the surface of the sphere to match this solution to the free-space solution outside of the sphere. Since the solution external to the sphere does not contain the origin, the solution will be constructed of spherical bessel functions that are both regular and irregular at the origin. For convenience we will choose to construct this solution in terms of the spherical hankel functions of the first and second kinds. If we assume that the potential is spherically

symmetric, the wave function should be independent of the azimuthal angle ϕ and can be written, in general, as

$$\psi_{\text{ext}}(\boldsymbol{r}) = \frac{1}{\sqrt{V}} \sum_{l=0}^{\infty} N_l \, i^l (2l+1) \left(h_l^{(2)}(kr) + S_l h_l^{(1)}(kr) \right) P_l(\cos\theta), \qquad (16.31)$$

where

$$k = \sqrt{\frac{2mE}{\hbar^2}} \qquad (16.32)$$

and N_l and $N_l S_l$ are the two expansion coefficients.

We know that for very large r, ψ_{ext} must match (16.6). To make this comparison we need to construct the partial wave expansion of the plane wave. Since the plane wave must be regular at the origin its expansion is

$$e^{i\boldsymbol{k}\cdot\boldsymbol{r}} = \sum_{l=0}^{\infty} i^l (2l+1) j_l(kr) P_l(\cos\theta) = \sum_{l=0}^{\infty} i^l \frac{2l+1}{2} \left(h_l^{(2)}(kr) + h_l^{(1)}(kr) \right) P_l(\cos\theta).$$
$$(16.33)$$

For $kr \gg l$, the Hankel functions behave as

$$h_l^{(1)}(kr) \sim -\frac{i}{kr} e^{i\left(kr - \frac{l\pi}{2}\right)} \qquad (16.34)$$

and

$$h_l^{(2)}(kr) \sim \frac{i}{kr} e^{-i\left(kr - \frac{l\pi}{2}\right)}. \qquad (16.35)$$

If we expand the scattering amplitude as

$$f_k(\theta) = \sum_{l=0}^{\infty} i^l (2l+1) f_l(k) P_l(\cos\theta), \qquad (16.36)$$

then the asymptotic form of the wave function becomes

$$\psi(\boldsymbol{r}) \;\; \to \;\; \frac{1}{\sqrt{V}} \left[\sum_{l=0}^{\infty} i^l \frac{2l+1}{2} \left(\frac{i}{kr} e^{-i\left(kr - \frac{l\pi}{2}\right)} - \frac{i}{kr} e^{i\left(kr - \frac{l\pi}{2}\right)} \right) P_l(\cos\theta) \right.$$
$$\left. + \sum_{l=0}^{\infty} i^l (2l+1) f_l(k) P_l(\cos\theta) \frac{e^{ikr}}{r} \right]. \qquad (16.37)$$

The asymptotic form of the external wave function is

$$\psi_{\text{ext}}(\boldsymbol{r}) \to \frac{1}{\sqrt{V}} \sum_{l=0}^{\infty} N_l \, i^l (2l+1) \left(\frac{i}{kr} e^{-i\left(kr - \frac{l\pi}{2}\right)} - S_l \frac{i}{kr} e^{i\left(kr - \frac{l\pi}{2}\right)} \right) P_l(\cos\theta). \quad (16.38)$$

279

Comparing these two expression we can immediately fix on constant for the external wave function as

$$N_l = \frac{1}{2} , \tag{16.39}$$

and the other constant S_l is related to the scattering amplitude by

$$-\frac{i}{2kr}e^{i\left(kr-\frac{l\pi}{2}\right)} + f_l(k)\frac{e^{ikr}}{r} = -S_l\frac{i}{2kr}e^{i\left(kr-\frac{l\pi}{2}\right)} . \tag{16.40}$$

Solving this for the partial wave scattering amplitude yields

$$f_l(k) = (-i)^l \frac{S_l - 1}{2ik} . \tag{16.41}$$

The total scattering amplitude then can be written as

$$f_k(\theta) = \sum_{l=0}^{\infty}(2l+1)\frac{S_l - 1}{2ik}P_l(\cos\theta) . \tag{16.42}$$

The scattering amplitude must satisfy the optical theorem (16.30) which will place a constraint on the constants S_l. The left-hand side of (16.30) can be written as

$$
\begin{aligned}
\Im f_k(0) &= \frac{1}{2i}\left(f_k(0) - f_k^*(0)\right) \\
&= \frac{1}{2i}\left(\sum_{l=0}^{\infty}(2l+1)\frac{S_l-1}{2ik}P_l(1) - \sum_{l=0}^{\infty}(2l+1)\frac{S_l^*-1}{-2ik}P_l(1)\right) \\
&= \frac{1}{2i}\sum_{l=0}^{\infty}\frac{2l+1}{2ik}\left(S_l + S_l^* - 2\right) \\
&= \frac{1}{4k}\sum_{l=0}^{\infty}(2l+1)\left(2 - S_l - S_l^*\right) .
\end{aligned} \tag{16.43}
$$

We can calculate the total cross section to be

$$
\begin{aligned}
\sigma &= \int d\Omega \, |f_k(\theta)|^2 \\
&= 2\pi\int_{-1}^{1} d(\cos\theta)\sum_{l=0}^{\infty}(2l+1)\frac{S_l^*-1}{-2ik}P_l(\cos\theta)\sum_{l'=0}^{\infty}(2l'+1)\frac{S_{l'}-1}{2ik}P_{l'}(\cos\theta) \\
&= \frac{4\pi}{4k^2}\sum_{l=0}^{\infty}(2l+1)\left(|S_l|^2 - S_l - S_l^* + 1\right) .
\end{aligned} \tag{16.44}
$$

Using these to expressions, (16.30) becomes

$$\frac{1}{4k}\sum_{l=0}^{\infty}(2l+1)\left(2 - S_l - S_l^*\right) = \frac{1}{4k}\sum_{l=0}^{\infty}(2l+1)\left(|S_l|^2 - S_l - S_l^* + 1\right) . \tag{16.45}$$

Cancelling identical terms results in

$$\sum_{l=0}^{\infty}(2l+1) = \sum_{l=0}^{\infty}(2l+1)|S_l|^2 \,. \tag{16.46}$$

This expression does not determine $|S_l|^2$. This can be done by generalizing derivation above to the case where the initial wave function is composed of the linear superposition of two plane waves with the same energy, but different directions. Doing this, one obtains

$$|S_l|^2 = 1 \,. \tag{16.47}$$

Since the S_l must have unit modulus, we can parameterize these constants as

$$S_l = e^{2i\delta_l} \,, \tag{16.48}$$

where the new constants δ_l are called the partial wave phase shifts. Using this parametrization, the total scattering amplitude becomes

$$f_k(\theta) = \sum_{l=0}^{\infty}(2l+1)\frac{e^{2i\delta_l}-1}{2ik}P_l(\cos\theta) = \sum_{l=0}^{\infty}(2l+1)\frac{e^{i\delta_l}\sin\delta_l}{k}P_l(\cos\theta) \tag{16.49}$$

and the corresponding total cross section is then

$$\sigma = \frac{4\pi}{k^2}\sum_{l=0}^{\infty}(2l+1)\sin^2\delta_l \,. \tag{16.50}$$

The solution can now be completed by determining the constants S_l. These can be obtained by matching the solution of the Schrödinger equation inside of our sphere with the exterior solution. We can write the interior wave function as

$$\psi_{\text{int}}(\boldsymbol{r}) = \sum_{l=0}^{\infty}i^l(2l+1)A_l g_l(r)P_l(\cos\theta) \,, \tag{16.51}$$

where $g_l(r)$ is the solution to the radial Schrödinger equation inside of the sphere. Requiring that the radial wave functions and their derivatives match at the boundary sphere gives

$$\frac{1}{2}\left(h_l^{(2)}(kr_0) + S_l h_l^{(1)}(kr_0)\right) = A_l g_l(r_0) \tag{16.52}$$

and

$$\frac{1}{2}\frac{\partial}{\partial r}\left(h_l^{(2)}(kr) + S_l h_l^{(1)}(kr)\right)\Big|_{r=r_0} = A_l \frac{\partial}{\partial r}g_l(r)\Big|_{r=r_0} \,. \tag{16.53}$$

Dividing the second expression by the first yields

$$\frac{\frac{\partial}{\partial r}\left(h_l^{(2)}(kr) + S_l h_l^{(1)}(kr)\right)\Big|_{r=r_0}}{h_l^{(2)}(kr_0) + S_l h_l^{(1)}(kr_0)} = \frac{\frac{\partial}{\partial r}g_l(r)\Big|_{r=r_0}}{g_l(r_0)} \,. \tag{16.54}$$

Using (16.48), the left side of this equation can be rewritten as

$$
\frac{\frac{\partial}{\partial r}\left(h_l^{(2)}(kr) + S_l h_l^{(1)}(kr)\right)\Big|_{r=r_0}}{h_l^{(2)}(kr_0) + S_l h_l^{(1)}(kr_0)} = \frac{k h_l^{(2)\prime}(kr_0) + e^{2i\delta_l} k h_l^{(1)\prime}(kr_0)}{h_l^{(2)}(kr_0) + e^{2i\delta_l} h_l^{(1)}(kr_0)}
$$

$$
= \frac{e^{-i\delta_l} k h_l^{(2)\prime}(kr_0) + e^{i\delta_l} k h_l^{(1)\prime}(kr_0)}{e^{-i\delta_l} h_l^{(2)}(kr_0) + e^{i\delta_l} h_l^{(1)}(kr_0)}
$$

$$
= k\frac{e^{-i\delta_l}\left(j_l'(kr_0) - i n_l'(kr_0)\right) + e^{i\delta_l}\left(j_l'(kr_0) + i n_l'(kr_0)\right)}{e^{-i\delta_l}\left(j_l(kr_0) - i n_l(kr_0)\right) + e^{i\delta_l}\left(j_l(kr_0) + i n_l(kr_0)\right)}
$$

$$
= k\frac{\cos\delta_l\, j_l'(kr_0) - \sin\delta_l\, n_l'(kr_0)}{\cos\delta_l\, j_l(kr_0) - \sin\delta_l\, n_l(kr_0)} = k\frac{j_l'(kr_0) - \tan\delta_l\, n_l'(kr_0)}{j_l(kr_0) - \tan\delta_l\, n_l(kr_0)}. \qquad (16.55)
$$

This can now be used in (16.54) and the equation solved to give

$$
\tan\delta_l = \frac{k j_l'(kr_0) g_l(r_0) - j_l(kr_0) g_l'(r_0)}{k n_l'(kr_0) g_l(r_0) - n_l(kr_0) g_l'(r_0)}. \qquad (16.56)
$$

If the internal radial wave function can be determined, it is now possible to obtain the phase shifts and to calculate the scattering amplitude and cross sections.

16.4 The Low-Energy Limit

It is useful to consider the behavior of the phase shifts in the limit $E \to 0$ which corresponds to $k \to 0$. This means that in this limit the wavelength of the scattering particle is approaching infinity. We can use the fact that for $x << l$,

$$
j_l(x) = \frac{x^l}{(2l+1)!!} + \cdots \qquad (16.57)
$$

and

$$
n_l(x) = -\frac{(2l+1)!!}{(2l+1)x^{l+1}} \qquad (16.58)
$$

to write

$$
\tan\delta_l = \frac{k\frac{l(kr_0)^{l-1}}{(2l+1)!!} g_l(r_0) - \frac{(kr_0)^l}{(2l+1)!!} g_l'(r)}{k\frac{(l+1)(2l+1)!!}{(2l+1)(kr_0)^{l+2}} g_l(r_0) + \frac{(2l+1)!!}{(2l+1)(kr_0)^{l+1}} g_l'(r)}
$$

$$
= \frac{\frac{(kr_0)^l}{(2l+1)!!}\left[\frac{l}{r_0} g_l(r_0) - g_l'(r)\right]}{\frac{(2l+1)!!}{(2l+1)(kr_0)^{l+1}}\left[\frac{1}{r_0}(l+1)g_l(r_0) + g_l'(r_0)\right]}
$$

$$
= \frac{(2l+1)(kr_0)^{2l+1}}{[(2l+1)!!]^2}\frac{l g_l(r_0) - r_0 g_l'(r_0)}{r_0 g_l'(r_0) + (l+1)g_l(r_0)}, \qquad (16.59)
$$

where $kr_0 << l$. This implies that for very low scattering energies, the phase shifts, and therefore the cross section, will be dominated by $l = 0$ or s-wave scattering. We can also express this limiting behavior as

$$k^{2l+1} \cot \delta_l = \frac{[(2l+1)!!]^2}{(2l+1)r_0^{2l+1}} \frac{r_0 g_l'(r_0) + (l+1)g_l(r_0)}{l g_l(r_0) - r_0 g_l'(r_0)} . \tag{16.60}$$

Then as $k \to 0$,

$$\lim_{k \to 0} k^{2l+1} \cot \delta_l = \frac{[(2l+1)!!]^2}{(2l+1)r_0^{2l+1}} \frac{g_l'(r_0) + (l+1)g_l(r_0)}{l g_l(r_0) - r_0 g_l'(r_0)} , \tag{16.61}$$

which is a constant with dimensions of $[L]^{-(2l+1)}$. In the particular case of $l = 0$, we have

$$\lim_{k \to 0} k \cot \delta_0 = -\frac{1}{r_0} \frac{r_0 g_0'(r_0) + g_0(r_0)}{r_0 g_0'(r_0)} = -\frac{1}{a} , \tag{16.62}$$

where a is called the scattering length. In fact, it is possible to write an expansion of $k \cot \delta_0$ about $k = 0$ which is called the *effective range expansion*.

16.5   Scattering from a Spherical Well

As a simple example, consider scattering from a spherical well defined by the potential

$$V(r) = V_0 \theta(r_0 - r) . \tag{16.63}$$

Since the potential vanishes at r_0 and will have free solutions outside of this radius, it is natural to match the solutions at this radius. Inside the radius the radial solutions are give by

$$g_l(r) = j_l(\kappa r) , \tag{16.64}$$

where

$$\kappa = \sqrt{\frac{2m}{\hbar^2}(E - V_0)} = \sqrt{k^2 - \frac{2mV_0}{\hbar^2}} . \tag{16.65}$$

For $V_0 < 0$, κ will be real.

This can be used to give

$$\tan \delta_l = \frac{k j_l'(kr_0) j_l(\kappa r_0) - \kappa j_l(kr_0) j_l'(\kappa r_0)}{k n_l'(kr_0) j_l(\kappa r_0) - \kappa n_l(kr_0) j_l'(\kappa r_0)} . \tag{16.66}$$

For s-wave scattering, this reduces to

$$\tan \delta_0 = \frac{\frac{k}{\kappa} \cos kr_0 \sin \kappa r_0 - \sin kr_0 \cos \kappa r_0}{\frac{k}{\kappa} \sin kr_0 \sin \kappa r_0 - \cos kr_0 \cos \kappa r_0} . \tag{16.67}$$

283

The low energy behavior of the phase shifts can be used to understand some characteristics of the phase shift. For $kr_0 << l$,

$$\tan \delta_l \cong \frac{(2l+1)(kr_0)^{2l+1}}{[(2l+1)!!]^2} \frac{lj_l(\kappa_0 r_0) - \kappa_0 r_0 j_l'(\kappa_0 r_0)}{\kappa_0 r_0 j_l'(\kappa_0 r_0) + (l+1)j_l(\kappa_0 r_0)} , \tag{16.68}$$

For s-wave scattering this becomes.

$$
\begin{aligned}
\tan \delta_0 &\cong -kr_0 \frac{\kappa_0 r_0 j_0'(\kappa_0 r_0)}{\kappa_0 r_0 j_0'(\kappa_0 r_0) + j_0(\kappa_0 r_0)} = -kr_0 \frac{\kappa_0 r_0 \left(\frac{\cos \kappa_0 r_0}{\kappa_0 r_0} - \frac{\sin \kappa_0 r_0}{\kappa_0^2 r_0^2} \right)}{\kappa_0 r_0 \left(\frac{\cos \kappa_0 r_0}{\kappa_0 r_0} - \frac{\sin \kappa_0 r_0}{\kappa_0^2 r_0^2} \right) + \frac{\sin \kappa_0 r_0}{\kappa_0 r_0}} \\
&= kr_0 \frac{\frac{\sin \kappa_0 r_0}{\kappa_0 r_0} - \cos \kappa_0 r_0}{\cos \kappa_0 r_0} = kr_0 \frac{\frac{1}{\kappa_0 r_0} - \cot \kappa_0 r_0}{\cot \kappa_0 r_0} , \tag{16.69}
\end{aligned}
$$

where

$$\kappa_0 = \sqrt{\frac{2m|V_0|}{\hbar^2}} \tag{16.70}$$

The scattering length is then given by

$$\frac{1}{a} = - \lim_{k \to 0} k \cot \delta_0 = \frac{1}{r_0} \frac{\cot \kappa_0 r_0}{\cot \kappa_0 r_0 - \frac{1}{\kappa_0 r_0}} . \tag{16.71}$$

So,

$$a = r_0 \frac{\cot \kappa_0 r_0 - \frac{1}{\kappa_0 r_0}}{\cot \kappa_0 r_0} . \tag{16.72}$$

Note that although k is small, $\tan \delta_0$ will become infinite when $\cot \kappa_0 r_0 = 0$. So if we start the potential such that $\kappa_0 r_0 < \pi$, the phase shift will be zero at $k = 0$. As the depth of the well increases, the phase shift will be infinite for small but finite k when $\kappa_0 r_0 = \pi$. This means that the phase shift becomes $\delta_0 = \pi/2$. As the well depth increases such that $\kappa_0 r_0 > \pi$, $\tan \delta_0$ becomes negative which implies that $\delta_0 = \pi$ at $k = 0$. Note also that the scattering length also becomes zero when $\cot \kappa_0 r_0 = 0$.

The physical interpretation of this phenomenon can be obtained by considering bound states in the same well. The eigencondition for bound states in the we is

$$\tilde{\kappa} \frac{j_l'(\tilde{\kappa} r_0)}{j_l(\tilde{\kappa} r_0)} = i\beta \frac{h^{(1)\prime}(i\beta r_0)}{h^{(1)\prime}(i\beta r_0)} , \tag{16.73}$$

where

$$\beta = \sqrt{\frac{2m|E|}{\hbar^2}} \tag{16.74}$$

and

$$\tilde{\kappa} = \sqrt{\frac{2m}{\hbar^2}(E - V_0)} = \sqrt{\kappa_0^2 - \beta^2} . \tag{16.75}$$
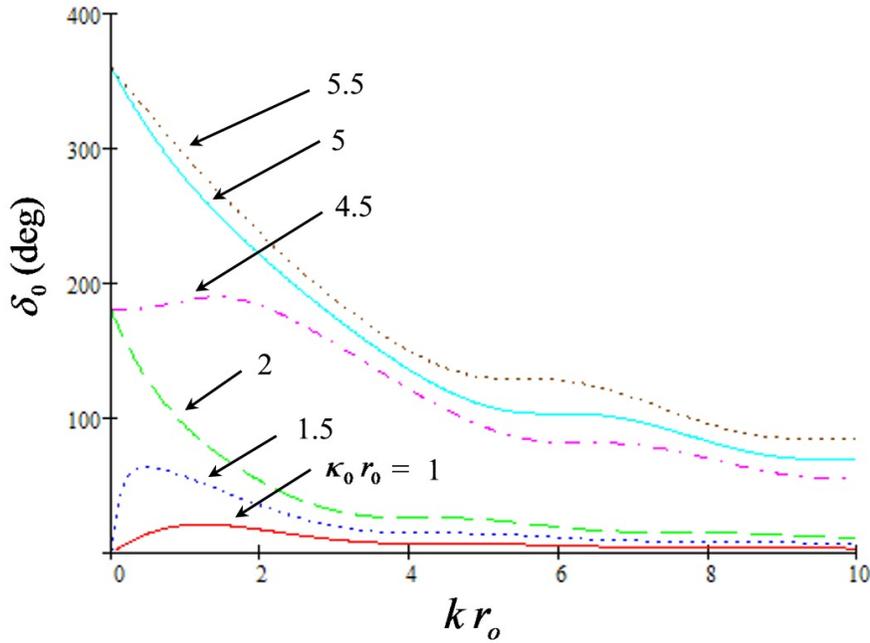
Figure 16.1: S-wave phase shifts in degrees as a function of kr_0 for $\kappa_0 r_0 = 1$, 1.5, 2, 4.5, 5 and 5.5.

For $l = 0$, this gives

$$\tilde{\kappa} \cot(\tilde{\kappa} r_0) + \beta = 0 \tag{16.76}$$

Now for a bound state at the top of the well, $E = 0$, $\beta = 0$ and $\kappa = \kappa_0$. The eigencondition for this case is then

$$\kappa_0 \cot(\kappa_0 r_0) = 0 \,, \tag{16.77}$$

or

$$\cot(\kappa_0 r_0) = 0 \,. \tag{16.78}$$

This is, of course, the condition for the phase shift at $k = 0$ to change form $\delta_0 = 0$ to $\delta_0 = \pi$. So this phenomenon occurs when the depth of the well is such that a bound state exists at $E = 0$. Once the phase shift at $k = 0$ becomes $\delta_0 = \pi$, The phase shift will then decrease with increasing k such that it passes through $\pi/2$. Recall that

$$f_l(k) \sim e^{i\delta_l} \sin \delta_l \,. \tag{16.79}$$

Whenever $\delta_l = \pi/2$, the amplitude becomes imaginary and $\sin^2 \delta_l = 1$ so that total cross section will have a local maximum at this point. This describes a resonance in the scattering partial wave.

Figure 16.1 shows the s-wave phase shifts in degrees as a function of kr_0 for $\kappa_0 r_0 = 1$, 1.5, 2, 4.5, 5 and 5.5. Zero energy bound states exist for

$$\kappa_0 r_0 = \left(n + \frac{1}{2}\right)\pi\,. \tag{16.80}$$

The first to of these occur at $\kappa_0 r_0 = 1.571$ and 4.712. This figure clearly shows the jump in phase shifts at $k = 0$ when these thresholds are passed.

We can also provide an interpretation of the scattering length a by examining the $k \to 0$ limit of the positive energy scattering wave function. For $l = 0$ we know that the internal part of the radial scattering wave function is $A_0 j_0(\kappa r)$ while the radial part of the external wave function is

$$
\frac{1}{2}\left(h^{(2)}(kr) + S_0 h^{(1)}(kr)\right) = \frac{1}{2}\left(\frac{e^{-ikr}}{-ikr} + e^{2i\delta_0(k)}\frac{e^{ikr}}{ikr}\right)
$$
$$
= \frac{e^{i\delta_0(k)}}{2ikr}\left(e^{ikr}e^{i\delta_0(k)} - e^{-ikr}e^{-i\delta_0(k)}\right) = \frac{e^{i\delta_0(k)}}{kr}\sin(kr + \delta_0(k))\,. \tag{16.81}
$$

Requiring that the wave functions match at r_0 gives

$$
A_0\frac{\sin \kappa r_0}{\kappa r_0} = \frac{e^{i\delta_0(k)}}{kr_0}\sin(kr_0 + \delta_0(k))\,. \tag{16.82}
$$

Then

$$
A_0 = \frac{\kappa}{k}e^{i\delta_0(k)}\frac{\sin(kr_0 + \delta_0(k))}{\sin \kappa r_0}\,. \tag{16.83}
$$

If we now define a reduced radial wave function $u_0(r)$ such that

$$
R(r) = \frac{e^{i\delta_0(k)}}{kr}u_0(r)\,, \tag{16.84}
$$

then

$$
u_0(r) = \begin{cases} \frac{\sin(kr_0 + \delta_0(k))}{\sin \kappa r_0}\sin \kappa r & r \leq r_0 \\ \sin(kr + \delta_0(k)) & r > r_0 \end{cases}\,. \tag{16.85}
$$

Figure 16.2 shows the reduced wave function $u_0(r)$ as a function of r/r_0 for $\kappa_0 r_0 = 1.7$ and $kr_0 = 0.001$. Note that outside of the well where $r/r_0 > 1$, the reduced radial wave function is essentially a straight line. Now consider the external reduced radial wave function for $k \ll 1$. Expanding to first order in k,

$$
\sin(kr + \delta_0(k)) \cong \sin(\delta_0(0)) + \cos(\delta_0(0))\left(r + \delta_0'(0)\right)k \tag{16.86}
$$

Expanding (16.69) gives

$$
\tan \delta_0(0) + \left(1 + \tan^2 \delta_0(0)\right)\delta_0'(0)k \cong kr_0\frac{\frac{1}{\kappa_0 r_0} - \cot \kappa_0 r_0}{\cot \kappa_0 r_0}\,. \tag{16.87}
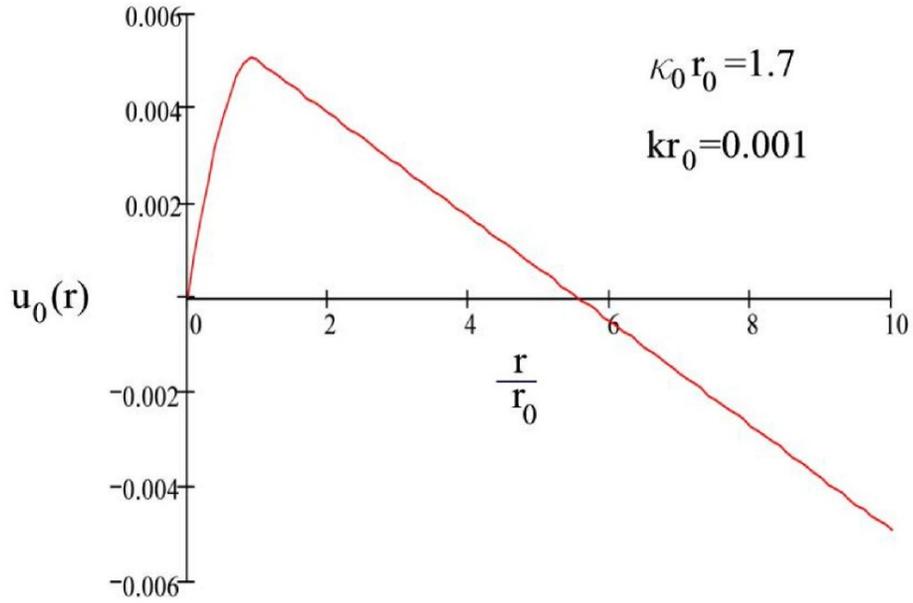$$

Figure 16.2: This figure shows the reduced wave function $u_0(r)$ as a function of r/r_0 for $\kappa_0 r_0 = 1.7$ and $kr_0 = 0.001$.

Since $\tan \delta_0(0) = 0$,

$$\delta_0'(0) = r_0 \frac{\frac{1}{\kappa_0 r_0} - \cot \kappa_0 r_0}{\cot \kappa_0 r_0} . \tag{16.88}$$

Therefore, since $\sin(\delta_0(0)) = 0$

$$\sin(kr + \delta_0(k)) \cong \cos(\delta_0(0)) (r - a) k . \tag{16.89}$$

This means that the scattering wave function for $k = 0$ vanishes at $r = a$. For the conditions of Fig. 16.2, $a/r_0 = 5.527$ which is consistent with the zero in the curve in this figure.