

Contents

1	Electron Spin	1
1.1	Intrinsic spin of an electron	1
1.2	A new description of spin	8
2	Spin Dynamics	19
2.1	Energy	19
2.2	Time dependent interaction	21
3	Two state systems	27
3.1	The two state approximation	27
4	N-State systems: General formulation	35
4.1	Quantum Mechanics of N-state systems	35
4.2	Expectation Value and Uncertainty	44
4.3	Compatible and Incompatible Observables: The Uncertainty Principle	49
4.4	Ehrenfest's Theorem	53
4.5	Quantum Mechanics of Angular Momentum: Higher Spin spaces	56
5	Quantum Mechanics of Point Particles	63
5.1	Propagation of an electron in a lattice	63
5.2	The continuum limit	71
5.3	Stationary states	88
5.4	Bound states	90
5.5	The quantum Harmonic Oscillator	101
6	Quantum Mechanics in three dimensions	111
6.1	Generalization to three dimensions	111
6.2	Orbital Angular Momentum	120
6.3	Spherically symmetric interaction: Hydrogen atom	128

Chapter 1

Electron Spin

1.1 Intrinsic spin of an electron

To ‘discover’ the laws of Quantum mechanics, we analyze a simple system. We completely ignore the propagation of the electron through space and instead focus on its ‘internal’ dynamics. We start with a classical description, analyze its implications, test these against experiments and see how this description falls short. Then, we will try to replace this description by its quantum version to account for the experimental results.

The electron is a charged ‘particle’, so we visualize it classically as either a truly point-like entity, or one with structure. Since no one has ever been able to split an electron into possible constituents, we will assume that it has a rigid structure (in the classical picture). If it does have a structure, it is easy to argue that since it is charged, it must possess a magnetic moment by virtue of its spin angular momentum. For simplicity, we visualize the electron as a rigid, charged sphere. If this sphere spins with angular momentum \mathbf{S} , it must possess a magnetic moment μ related to \mathbf{S} as

$$\mu = \frac{g(-e)}{2m_e} \mathbf{S} \quad (1.1)$$

where $-e$ is the electron charge, m_e is mass and g is a dimensionless constant, called a *structure constant*. To see this, consider a tiny section of the sphere which will move along a circle of some radius r with the same angular velocity as that of the sphere. The speed of this section will be $v = r\omega$ where ω is the angular velocity.

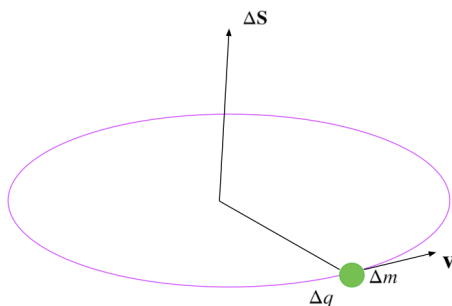


Figure 1.1: An element of mass Δm and charge Δq with angular momentum ΔS

The angular momentum of this section will be perpendicular to the plane of motion, with magnitude

$$\Delta S = \Delta m v r \quad (1.2)$$

At the same time, the motion of this section constitutes an effective current

$$\begin{aligned}\Delta I &= \frac{\Delta q}{T} \\ &= \frac{\Delta q v}{2\pi r}\end{aligned}\tag{1.3}$$

where T is the time it takes the section to once round the circle. Then, the magnetic moment associated with it has magnitude

$$\begin{aligned}\Delta\mu &= \Delta I(\pi r^2) \\ &= \frac{\Delta q v r}{2} \\ &= \left(\frac{\Delta q}{2\Delta m}\right)\Delta S\end{aligned}\tag{1.4}$$

which is proportional to the magnitude of the angular momentum. When we add up such magnetic moments, we will end up with the relation between \mathbf{S} and μ . Note that since the electron is negatively charged, \mathbf{S} will be in direction opposite to μ . The structure factor g will depend on the detailed charge and mass distributions.

Experimentally, it is known that electrons possess magnetic moment, which accounts for magnetism in matter. Therefore, there must be a spin angular momentum associated with an electron. Then, we have a classical picture of an electron as a spinning object with some shape (perhaps spherical), spinning with some angular momentum \mathbf{S}

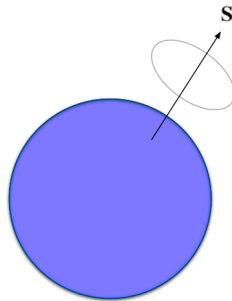


Figure 1.2: A mental ‘picture’ of an spinning electron.

We will soon see that such ‘pictures’ are inconsistent with experiments and have to be given up. It is this lack of visual intuition that makes quantum physics appear so counter-intuitive.

If the electron possesses a spin and an associated magnetic moment, we should be able to conduct experiments which can give us some information about the internal structure of the electron (encoded in the structure factor g in (1.1)). Since we are at the moment unfamiliar with the laws governing the ‘propagation’ of the electron in space, it will be helpful if we can conduct experiments on electron spin without worrying about its propagation in space. Isolating this internal degree of freedom from propagation in space is not simple, since the electron does not follow a trajectory in space (the double-slit experiment!). In such a situation, it is not clear what the ‘center of mass’ frame is, in which one need focus only on the spin motion.

A clever workaround is to perform experiments not on isolated electrons, but a complex *atom* within which the electron resides. For an atom, a classical description of motion is approximately acceptable (since quantum effects are relatively small). If the electron within this atom is responsible for the magnetic

moment of the atom, then magnetic experiments conducted on the atom can indirectly give us information about the spinning electron inside. The Potassium atom is a good candidate, since its magnetic moment arises solely due to a single electron's spin angular momentum¹. If a Potassium atom is made to interact with an external magnetic field, the interaction of this field with the magnetic moment of the atom will effect the dynamics of the atom. A study of these dynamics will indirectly give us information about the spin of the electron within the atom.

Assuming classical physics correctly describes electron spin, the 'spin-state' of the electron will be a 'three-vector' \mathbf{S} described by a set of three numbers (its components) S_x , S_y and S_z in some Cartesian coordinate system. There will be an associated magnetic moment, given by (1.1). In presence of an external magnetic field \mathbf{B} , this magnetic moment will interact with the field with an interaction energy

$$\begin{aligned} U_M &= -\boldsymbol{\mu} \cdot \mathbf{B} \\ &= \frac{ge}{2m_e} \mathbf{S} \cdot \mathbf{B} \end{aligned} \quad (1.5)$$

If a Potassium atom is in presence of this magnetic field, the interaction energy of the electron within it with the magnetic field is given by (1.5). This is also the interaction energy of the atom with the field. Therefore, the atom (visualised as a classical object) will experience a force given by

$$\begin{aligned} \mathbf{F} &= -\nabla(\boldsymbol{\mu} \cdot \mathbf{B}) \\ &= \frac{ge}{2m_e} \nabla(\mathbf{S} \cdot \mathbf{B}) \end{aligned} \quad (1.6)$$

This force will make the atom accelerate with acceleration

$$\begin{aligned} \mathbf{a} &= \frac{1}{M} \mathbf{F} \\ &= \frac{ge}{2m_e M} \nabla(\mathbf{S} \cdot \mathbf{B}) \end{aligned} \quad (1.7)$$

where M is the mass of the Potassium atom. This acceleration will be non-zero only if the magnetic field is inhomogeneous (since it is proportional to the gradient of the field). In addition, the electron spin will experience a torque given by

$$\boldsymbol{\tau} = \boldsymbol{\mu} \times \mathbf{B} \quad (1.8)$$

which will lead to a change in the direction of the spin

$$\begin{aligned} \frac{d\mathbf{S}}{dt} &= \boldsymbol{\tau} \\ &= -\frac{ge}{2m_e} \mathbf{S} \times \mathbf{B} \end{aligned} \quad (1.9)$$

If the magnetic field has magnitude B and direction along unit vector $\hat{\mathbf{n}}$, this equation can be written as

$$\frac{d\mathbf{S}}{dt} = -\omega \mathbf{S} \times \hat{\mathbf{n}} \quad (1.10)$$

where $\omega = geB/2m_e$ has dimensions of angular velocity. It is easy to solve this equation (Exercise 1.1.2). The solution shows that the spin vector \mathbf{S} will precess about the direction of the magnetic field with angular frequency ω , called *Larmor frequency*

¹The Potassium atom has a single electron in its outermost 'shell', whose quantum 'propagation' is such that it has zero angular momentum due to propagation in space. Technically, we say that this electron is in an 's-state'. Therefore, the magnetic moment of this electron arises solely due to its spin angular momentum. There are, of course, other electrons as well in the atom (in the inner 'shells'), but they conspire in a way such that the net magnetic moment due to them is zero.

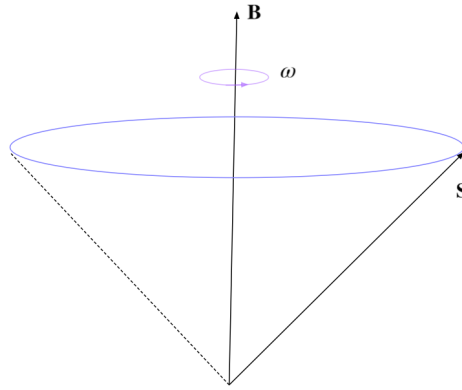


Figure 1.3: Classical spin vector \mathbf{S} precesses about the direction of the magnetic field with angular frequency $\omega = geB/2m_e$.

Then, classical physics predicts that if a Potassium atom is placed in an inhomogeneous magnetic field, it will experience a force proportional to the component of the electron spin along the magnetic field, and this spin will precess about the direction of the magnetic field. We will now discuss modern experimental techniques to verify this ‘picture’ that we have constructed, based on our intuition about classical physics.

We shall discuss a modern version of the classic Stern-Gerlach experiment, first carried out in 1922. The experiment involves a collimated beam of Potassium atoms, which is made to pass through a region of an inhomogeneous magnetic field². Due to the interaction of the magnetic moment of the electron in the atom and the inhomogeneous magnetic field, the atom experiences a force, which should deflect its trajectory. The atoms, after they emerge from the magnets, are collected on a detector screen

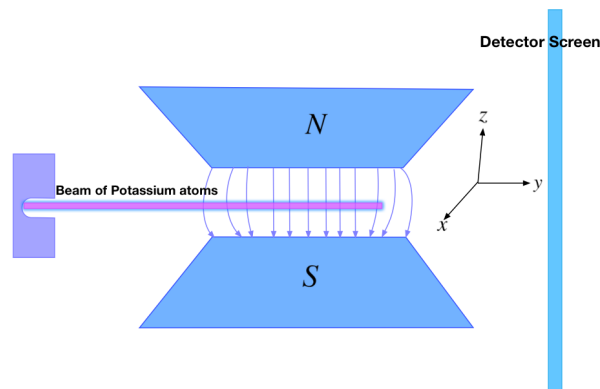


Figure 1.4: A modern Stern-Gerlach experiment setup.

The magnets in the setup have a shape such that only the z -component of the magnetic field depends on the z -coordinate³ (refer to the figure for the coordinate system used). Due to the inhomogeneity in the field, the potassium atom will experience a force (1.6) whose z -component will be

$$F_z = \frac{geS_z}{2m_e} \left(\frac{dB_z}{dz} \right) \quad (1.11)$$

This force will depend on the orientation of the electron spin within the atom, relative to the direction of

²The original Stern-Gerlach experiment used a beam of Silver atoms.

³A mathematical example of such a field could be $\mathbf{B} = -bx\hat{\mathbf{i}} + (B_0 + bz)\hat{\mathbf{k}}$, which is consistent with $\nabla \cdot \mathbf{B} = 0$ and $\nabla \times \mathbf{B} = 0$.

B. If the angle between \mathbf{S} and \mathbf{B} is θ , then the force will be

$$F_z = \frac{geS}{2m_e} \left(\frac{dB_z}{dz} \right) \cos \theta \quad (1.12)$$

where S is the magnitude of the spin. The Potassium atoms in the beam are expected to be emitted from the oven with the electron having a random spin direction. Therefore, θ in (1.12) will vary randomly between 0 and π . As a result, the Potassium atoms will experience force F_z between $F_z = +(geS/2m_e)dB_z/dz$ and $F_z = -(geS/2m_e)dB_z/dz$ with all values in this range equally probable. This force will displace a given atom in the beam along the z -direction by

$$\Delta z \sim \frac{1}{2} \frac{F_z}{M} t^2 \quad (1.13)$$

where M is the mass of the Potassium atom and $t \sim d/v$ where d is the distance between the magnets and the point where the atom will strike the detecting screen and v is the speed which which the atom is moving. Since the atoms are emitted from an oven, this speed will satisfy the Maxwell-Boltzmann distribution.

The deflected Potassium atoms are detected as follows. The detecting screen is a grid of hot Platinum wire maintained at some fixed potential. When an atom strikes the wire, it gets ionized. The bias voltage pulls the charged atom to a nearby plate, and the resulting current is measured. A current measured at some point on the grid signals the detection of an atom at that point.

Because of the random orientation of the spins, one should expect a smooth distribution of Potassium atoms on the screen, with the effect of the Maxwell-Boltzmann distribution of speed v superposed

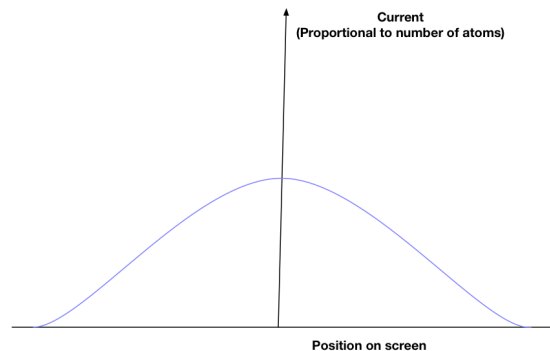


Figure 1.5: The classically expected Potassium atom distribution on the screen.

The experimentally measured distribution is as follows

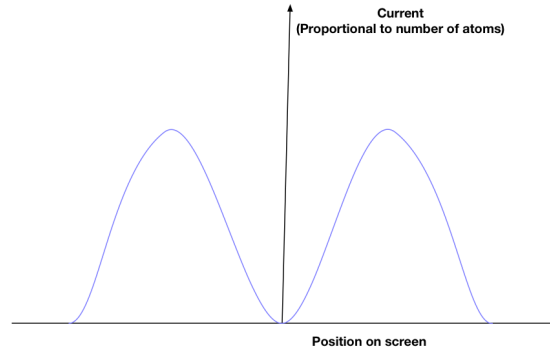


Figure 1.6: The experimentally observed Potassium atom distribution on the screen.

If we account for the effect of the Maxwell-Boltzmann distribution on the speed of the atoms, we can figure out the distribution on the screen if all the atoms had the same velocity. This distribution is striking

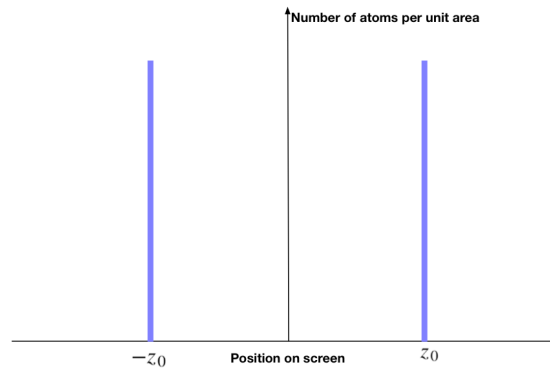


Figure 1.7: The Potassium atom distribution for a fixed velocity.

The distribution implies that electron spin in each Potassium atom had just two possible components along the direction of the magnetic field, say $S_z = \pm S_0$. The atoms with $S_z = +S_0$ are deflected along one direction and those with $S_z = -S_0$ are deflected in the other. It turns out that these two values are $S_z = \pm \hbar/2$ where $\hbar = h/2\pi$, h being Planck's constant⁴.

⁴The Stern-Gerlach experiment by itself would determine only the product gS_z , where g is the structure constant. Another experiment would be required to determine g and S_z individually.

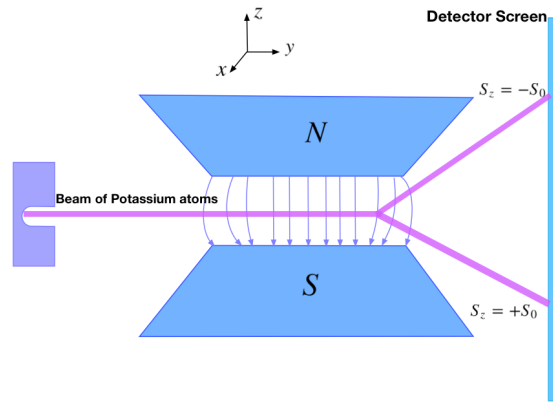


Figure 1.8: The electron beam splits into two, implying two possible components of spin along the magnetic field.

This experiment forces us to create the following ‘picture’ in our head

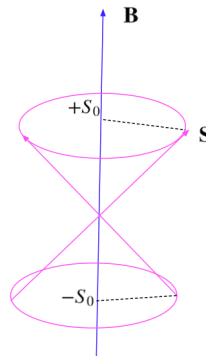


Figure 1.9: Classical interpretation of the Stern Gerlach experiment.

This interpretation is clearly problematic. Since there is nothing inherently special about the direction of the magnetic field, it is bizarre why the Potassium atom spins should be oriented in such a special way. To appreciate this more, let us say we repeat the experiment with the magnetic field along the x -direction. Once again, the Potassium atom beam will split into two, and we will conclude that the electron spin had two possible components along the magnetic field, this time $S_x = +S_0$ and $S_x = -S_0$. Similar experiment with the magnetic field along *any* direction (say \hat{n} will have us conclude that $S_n = +S_0$ and $S_n = -S_0$. Then, we will conclude that the Potassium atoms emitted from the oven have spins *with the same two components along any direction in space*. Clearly, there cannot exist a vector which has the same components along *all* directions in space. Therefore, it is incorrect to visualize the electron as a spinning object in a classical sense, with a spin vector pointing along some direction in space. Yet, it *is* true that it possesses a spin angular momentum, since it possesses a magnetic moment. How do we reconcile these conflicting observations? The way out is to accept that there exists a spin angular momentum, but that its classical description as a vector with three components is inadequate.

Exercises

Exercise 1.1.1. Visualize an electron as a homogeneous solid sphere of mass m_e with charge $-e$ uniformly distributed on its surface. Show that the spin angular momentum and the magnetic moment are related

as

$$\mu = -\frac{5e}{6m_e}\mathbf{S}$$

Exercise 1.1.2. Verify that equation (1.10) implies that the electron spin will precess about the direction of the magnetic field.

Hint: Choose a Cartesian coordinate system with unit vector $\hat{\mathbf{n}}$ along the z -direction and write the three component equations. The equations for S_x and S_y will be first order coupled differential equations. Decouple them by differentiating them, to obtain equations for simple harmonic motion.

1.2 A new description of spin

To develop a new formalism to describe electron spin, we perform a series of Stern-Gerlach experiments in succession. At this stage, it is useful to introduce a symbolic way of representing the Stern-Gerlach apparatus. The apparatus will be represented as a ‘black-box’ with the direction of the magnetic field specified. Often, it will be useful to block one of the two beams (into which the original beam splits), say by putting an obstruction in its path

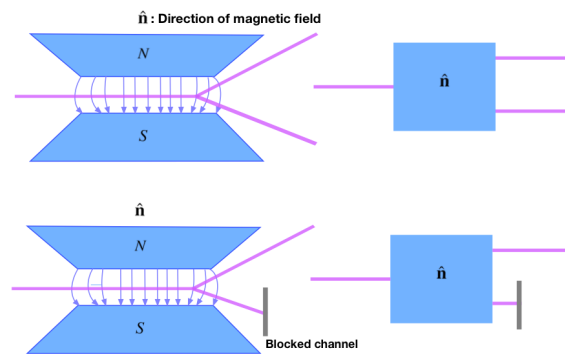


Figure 1.10: Symbolic representation of Stern-Gerlach ‘filters’.

We now analyze a series of experiments which will eventually allow us to determine a mathematical formalism to describe electron spin. In the first experiment, we take a beam of Potassium atoms emerging from an oven and pass it through a Stern Gerlach (SG) filter. It is observed that the beam splits into two beams of equal intensity, irrespective of the orientation of the filter. Then, if there are N_0 atoms in the initial beam, each beam that emerges will have *statistically* $N_0/2$ atoms. If we were to interpret this classically, we would infer that the two beams consist of atoms in which the electron spins have components $\pm\hbar/2$

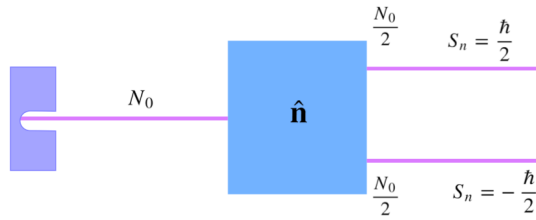


Figure 1.11: Experiment 1.

In the second experiment, we take two such filters, each with the magnetic field along the z -direction. The first filter splits a beam emerging from an oven into two beams, which we label ‘up’ and ‘down’ (corresponding to the deflection relative to the direction of the magnetic field). We block one of the two beams, say the ‘down’ beam, and place the second filter in front of the first. It is observed that if the ‘up’ channel of the second filter is blocked, no atoms pass through it. However, if its ‘down’ channel is blocked, all the atoms pass through the second filter. From this, we infer that the atoms in the two beams have some physical property, which distinguishes the two beams. An ‘up’ beam will only pass through a filter whose ‘up’ channel is open, and similarly a ‘down’ beam will only pass through a filter whose ‘down’ channel is open. Classically, this property is the z -component of the electron spin. In the classical interpretation, the ‘up’ beam consists of atoms with electrons having z -component of spin $S_z = +\hbar/2$ and the ‘down’ beam has atoms with electrons having $S_z = -\hbar/2$. The original beam emerging from the oven must have been a mixture of atoms with these two components of spin

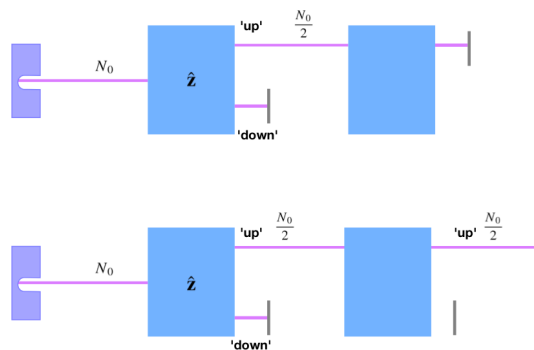


Figure 1.12: Experiment 2.

Even though classical visualisation of spin as a three-dimensional vector needs to be given up, the SG experiment *does* in a way measure the component of spin along the direction of the magnetic field, though perhaps not in the sense we are used to visualising it. Indeed, there is something different about the two beams that emerge, which classically would be the component of spin. We will progressively realise that physical quantities that are defined in classical physics are still defined in quantum physics, though they need to be re-interpreted. Then, we will *still* say that electrons (in Potassium atoms) that emerge out of a SG filter, oriented along the z -direction, have components of spin along this direction equal to $S_z = \pm\hbar/2$. However, we will *not* visualize S_z as component of a three dimensional vector representing the angular

momentum of a rigid, spinning object in the classical sense. It is useful to introduce a new notation here, originally due to Dirac, and widely accepted. We say that electrons in the two beams are in different *spin states*, symbolically denoted as $|+\rangle_z$ and $|-\rangle_z$. The $|+\rangle_z$ state is one in which $S_z = +\hbar/2$ and the $|-\rangle_z$ state is one in which $S_z = -\hbar/2$. The symbol $| \rangle$ is called a *ket*. It is a very powerful notation, as will progressively become clear. For now, this is just a symbolic way of saying that there is a component of spin, but it does not have a traditional interpretation. The classical interpretation would be that one state is represented by a three-dimensional vector $\mathbf{S}_+ = S_x \hat{\mathbf{i}} + S_y \hat{\mathbf{j}} + (\hbar/2) \hat{\mathbf{k}}$ and the other by $\mathbf{S}_- = S'_x \hat{\mathbf{i}} + S'_y \hat{\mathbf{j}} - (\hbar/2) \hat{\mathbf{k}}$ where S_x , S_y , S'_x and S'_y have not been measured yet, but which subsequent measurements of y and z components of spin would reveal.

In the third experiment, we take a beam of atoms, all in state $|+\rangle_z$, and make it incident on a SG filter oriented along the x -direction⁵. It is observed that the beam splits into two beams of (statistically) equal intensity. Similar result is observed if we take a beam of atoms in the $|-\rangle_z$ state and pass it through an x -filter. Once again, the beam splits into two beams of equal intensity

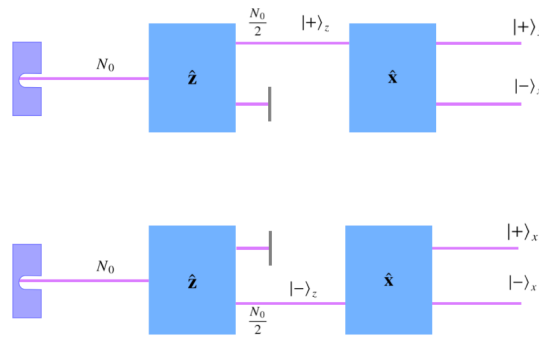


Figure 1.13: Experiment 3.

How do we interpret this result? Classically, we would interpret that this experiment has measured the x -component of spin. Then, if say the beam with $S_z = +\hbar/2$ emerges out of the x -filter such that the S_x component is measured to be $+\hbar/2$, the classical state would be described by the vector $\mathbf{S}_+ = \hbar/2 \hat{\mathbf{i}} + S_y \hat{\mathbf{j}} + (\hbar/2) \hat{\mathbf{k}}$, the S_y component still not determined. Note the ‘common sense’ notion that this component exists, even though we have not yet measured it. However, sticking to our notation, we will say that since the S_x component of spin is measured to be $+\hbar/2$, the state will be represented as $|+\rangle_x$. It might seem puzzling why this notation suppresses the fact that the z -component of spin is also $+\hbar/2$ (which has already been measured). The reason will become clear once we analyze the next experiment. In the fourth experiment, we take the $|+\rangle_x$ beam, and pass it through a z -filter. The arrangement is illustrated below. The beam splits into two beams of equal intensity, one corresponding to state $|+\rangle_z$ and the other corresponding to state $|+\rangle_z$

⁵In a coordinate invariant language, the second filter is oriented perpendicular to the first one.

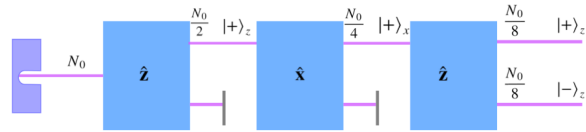


Figure 1.14: Experiment 4.

This result clearly implies that the atoms in the beam characterised by the state $|+\rangle_x$ did *not* all have z -component of spin equal to $\hbar/2$. In fact, half of them must have had $S_z = -\hbar/2$. This is odd, since before entering the x -filter, they were all in the state $|+\rangle_z$. The implications of this experiment are even more striking if we compare it against the second experiment. Following are two setups, next to each other

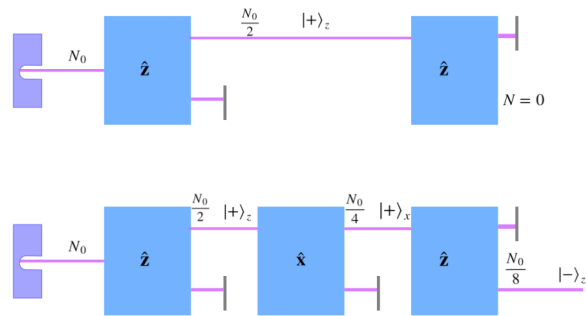


Figure 1.15: Experiment 5.

In the first setup, a beam emerging from an oven is made to pass through a z -filter with the $S_z = -\hbar/2$ channel blocked. The atoms emerging from this filter, all in state $|+\rangle_z$, are passed through another z -filter with the $S_z = +\hbar/2$ channel blocked. No atoms emerge from this filter, since the incident atoms did not have $S_z = -\hbar/2$. However, as strangely implied by the third experiment, if we introduce an x -filter between the two z -filters, some atoms *do* emerge out (a quarter, actually). It is as if measurement of the x -component somehow made component $S_z = -\hbar/2$ reappear (which had been blocked by the first z -filter). We could similarly introduce y -filters, or filters which measure spin along an arbitrary direction. In general, we will observe the following: if the measurement of a component of spin is followed by a measurement of some other component of spin, the result of the first measurement will in general be ‘erased’. For instance, if the z -component of spin is measured to be say $\hbar/2$, followed by a measurement of the x -component, a subsequent measurement of the z -component *need not* give $\hbar/2$. There is a possibility that the second measurement of S_z might give $-\hbar/2$.

There is another important, and more ‘disturbing’ aspect of electron spin, that seems to appear. Say, we take a Potassium atom in state $|+\rangle_z$ and pass it through an x -filter. We know that a beam of such atoms splits into two with atoms in one beam in state $|+\rangle_x$ and the atoms in the other in state $|-\rangle_x$. How does any one given atom decide whether to be a part of one beam, or the other? In other words, what

decides that a given atom will change its spin state from $|+\rangle_z$ to $|+\rangle_x$ or $|-\rangle_x$?

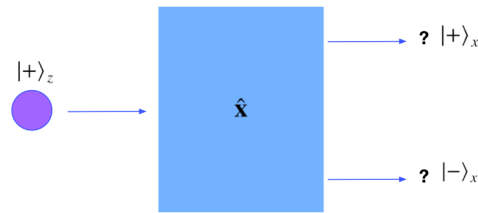


Figure 1.16: How does an atom decide?

We could generalize this to a situation in which a beam in state $|+\rangle_z$ is made to pass through a filter oriented along some arbitrary direction, the direction denoted by a unit vector $\hat{\mathbf{n}}$. This direction can be characterised by a pair of ‘spherical polar’ angles θ and ϕ , with $0 \leq \theta \leq \pi$ and $0 \leq \phi \leq 2\pi$. Once again, a beam of atoms in state $|+\rangle_z$ will split into two, but the two beams will in general not have the same intensity

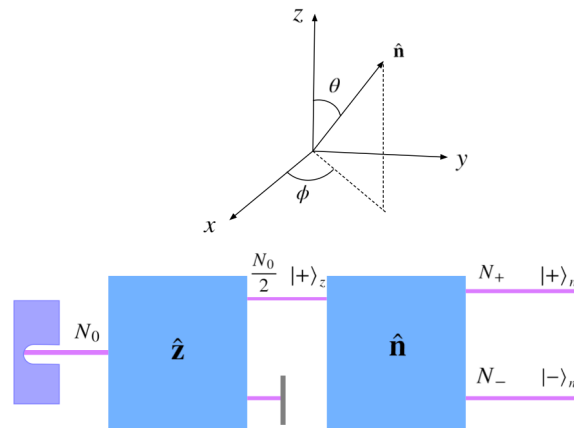


Figure 1.17: Stern Gerlach filter along an arbitrary direction. The beam still splits, but the two emerging beams in general do not have the same intensity.

Again, a certain fraction of atoms will change their spin state to $|+\rangle_n$, and others to $|-\rangle_n$. What decides what a given atom will do? All this is reminiscent of the double-slit experiment, where one faces a similar paradoxical situations: how does a given electron decide, which slit to pass through? The answer, as we saw, was that the electron does *not* take either alternative. There is a certain probability that it reaches some point on the screen, and this probability arises out of a complex amplitude, the *probability amplitude*. This amplitude satisfies a principle of superposition: the total probability amplitude is the sum of two amplitudes, one associated with the electron passing through one slit, and the other with the electron passing through the second slit. In some sense, the electron is in a ‘state’ of *superposition* of two states, each corresponding to a slit. We should expect the same principle to hold, even for the ‘internal’ dynamics of the electron. We need to recognize that we are encountering here (for electron spin) another situation in which we need to invoke some kind of principle of superposition, as we were forced to in the double slit experiment. We need to embrace the idea that an electron in state $|+\rangle_z$ has an inherent probability

of ending up in state $|+\rangle_n$, and a probability of ending up in state $|-\rangle_n$, when the component of spin along some direction $\hat{\mathbf{n}}$ is measured. Further, this probability is *fundamental*, in the sense that there is no underlying deterministic mechanism at work, which we do not know of. Further, just as in the double slit experiment, this probability must arise out of a *probability amplitude*, whose modulus squared should result in the probability. Without apologies, then, we will postulate the following idea: An electron in spin state $|+\rangle_z$ is neither in state $|+\rangle_n$ nor in state $|-\rangle_n$, but some kind of a linear superposition of these states. Conversely, an electron in state $|+\rangle_n$ is neither in state $|+\rangle_z$ nor in state $|-\rangle_z$, but a linear superposition of these states. If a beam of atoms in state $|+\rangle_n$ is passed through a z -filter, it will split into two, one corresponding to state $|+\rangle_z$ and the other to state $|-\rangle_z$. To make this idea mathematical, we will introduce the operation of adding these *ket* states, and multiplying them with complex numbers (we will see soon why these numbers need to be complex). Then, the state $|+\rangle_n$ can be expressed as

$$|+\rangle_n = c_+ |+\rangle_z + c_- |-\rangle_z \quad (1.14)$$

where c_+ and c_- are complex numbers, depending on the orientation of $\hat{\mathbf{n}}$ relative to the z -direction. This equation is still empty of content, until we interpret the numbers c_+ and c_- . These numbers should be related to the probability that a measurement of S_z on state $|+\rangle_n$ results in state $|+\rangle_z$ or $|-\rangle_z$. These numbers are just the probability amplitudes that we are looking for. We then postulate that the modulus squared of these numbers gives the probability that the corresponding spin state is produced. That is, $|c_+|^2$ is the probability that the state will end up in $|+\rangle_z$ and $|c_-|^2$ is the probability that the state will end up in $|-\rangle_z$. Since the total probability of obtaining these states should be unity, these complex amplitudes should satisfy a ‘normalization’ condition

$$|c_+|^2 + |c_-|^2 = 1 \quad (1.15)$$

Now, there is nothing special about states $|\pm\rangle_z$, and we should be able to express the state $|+\rangle_n$ as a linear superposition of, say, $|\pm\rangle_x$ states, since a beam of atoms in state $|+\rangle_n$, when passed through a x -filter, will again split. Then, the *same* spin state $|+\rangle_n$ can be written as

$$|+\rangle_n = c'_+ |+\rangle_x + c'_- |-\rangle_x \quad (1.16)$$

where c'_+ and c'_- are the probability amplitudes of measuring $S_x = \hbar/2$ and $S_x = -\hbar/2$ when an atom in state $|+\rangle_n$ is passed through a x -filter. Let us denote a general spin state as $|\psi\rangle$. We are assuming that this state is produced by passing a beam of atoms through a SG filter oriented along some direction, with one channel blocked. Now, we pass a beam of atoms in this state through a filter with magnetic field along some arbitrary direction, say $\hat{\mathbf{n}}$. Then, the formalism we are developing suggests that if any one atom in state $|\psi\rangle$ is made to pass through this filter, there is a probability amplitude $c_+(\hat{\mathbf{n}})$ that the spin state of the atom changes to $|+\rangle_n$, and an amplitude $c_-(\hat{\mathbf{n}})$ that it changes to $|-\rangle_n$. Mathematically, this is expressed as

$$|\psi\rangle = c_+(\hat{\mathbf{n}}) |+\rangle_n + c_-(\hat{\mathbf{n}}) |-\rangle_n \quad (1.17)$$

The point to notice here is that there is nothing special about the states $|\pm\rangle_n$, since there is nothing special about the direction $\hat{\mathbf{n}}$. We could have expressed $|\psi\rangle$ as a linear superposition of spin states along *any* direction in space. Of course, the associated probability amplitudes will depend on the direction. This is mathematically very similar to how a *vector* behaves in two dimensions. A simple example of a vector in two dimensions is that of a *directed arrow* in a plane. This arrow is characterized by its length (*magnitude*) and direction. Consider then a set of all such arrows in this plane. Geometrically, we can define two operations on these arrows: multiplication of an arrow by a number (*scalar*) and addition of two arrows. Given a number $c > 0$, the arrow $c\vec{a}$ is obtained by stretching (or shrinking, if $c < 1$) the arrow \vec{a} by a factor of c . If $c < 0$, we carry out the same operation, but flip the direction of the arrow. The operation of addition of two such arrows is carried out by the so-called *parallelogram law*. Given two such arrows \vec{a} and \vec{b} , we define an arrow $\vec{c} = \vec{a} + \vec{b}$ as follows: we move one of the arrow parallel to itself till its ‘tail’ touches the tail of the other arrow. We then complete a parallelogram with the two arrows as adjacent sides, to deduce the arrow representing $\vec{a} + \vec{b}$

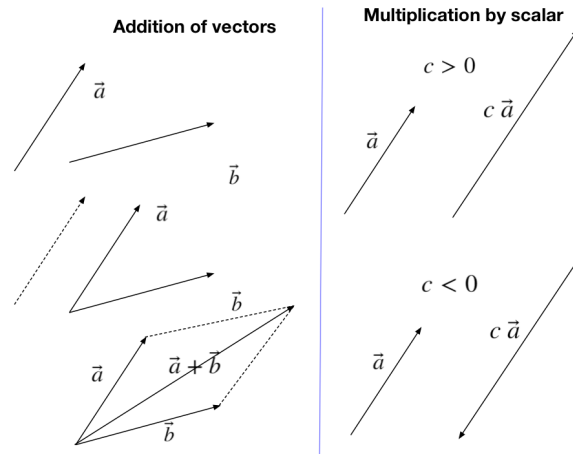


Figure 1.18: Vector algebra for arrows in a plane.

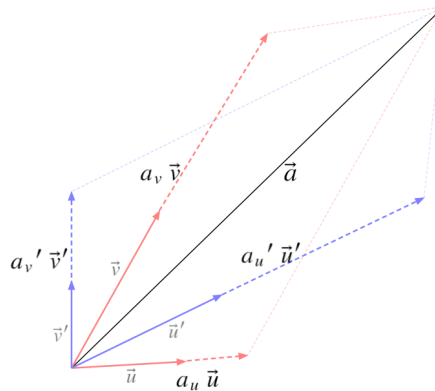
With these operations defined, we can express any arrow as a linear superposition of *two* arrows that are not in the same direction⁶. Then, given two arrows \vec{u} and \vec{v} that are not in the same direction, an arbitrary arrow \vec{a} can be expressed as

$$\vec{a} = a_u \vec{u} + a_v \vec{v} \quad (1.18)$$

where a_u and a_v are two real numbers, called *components* of \vec{a} along \vec{u} and \vec{v} respectively. Again, there is nothing special about \vec{u} and \vec{v} . Given any other pair of non-parallel arrows \vec{u}' and \vec{v}' , we can express the same arrow \vec{a} as

$$\vec{a} = a'_u \vec{u}' + a'_v \vec{v}' \quad (1.19)$$

where a'_u and a'_v components of \vec{a} along \vec{u}' and \vec{v}'

Figure 1.19: ‘Resolving’ a directed arrow in a plane. The sets \vec{u}, \vec{v} and \vec{u}', \vec{v}' are called *bases*.

The sets \vec{u}, \vec{v} and \vec{u}', \vec{v}' are called *bases*. Then, any directed arrow in a plane can be expressed as a superposition of a pair of basis vectors. Expressed as such a linear combination, the arrow is said to be *resolved* along the basis. Given two such arrows \vec{a} and \vec{b} , we can expand them in a suitable basis, say \vec{u}, \vec{v}

$$\begin{aligned} \vec{a} &= a_u \vec{u} + a_v \vec{v} \\ \vec{b} &= b_u \vec{u} + b_v \vec{v} \end{aligned} \quad (1.20)$$

Then, it is easy to figure out that the sum of \vec{a} and \vec{b} , as defined by the parallelogram law, is a vector whose expansion in this basis is

$$\vec{a} + \vec{b} = (a_u + b_u) \vec{u} + (a_v + b_v) \vec{v} \quad (1.21)$$

⁶This is why we say that these arrows are *two dimensional* vectors

The components of the vector $\vec{a} + \vec{b}$ are just the sum of components of \vec{a} and \vec{b} in the given basis. Similarly, if a vector \vec{a} is multiplied by a number c , the resulting vector $c\vec{a}$ is given by

$$c\vec{a} = (ca_u)\vec{u} + (ca_v)\vec{v} \quad (1.22)$$

Returning back to spin states, equation (1.17) can be interpreted in a similar way. The spin state $|\psi\rangle$ can be viewed as a ‘vector’ in an abstract two-dimensional vector space, in which spin states $|\pm\rangle_n$ along any direction $\hat{\mathbf{n}}$ form a ‘basis’. The vector $|\psi\rangle$ can be resolved along any one such basis, with the ‘components’ of the vector along the basis vectors having the physical interpretation of probability amplitudes. Then, the numbers $c_+(\hat{\mathbf{n}})$ and $c_-(\hat{\mathbf{n}})$ in (1.17) are components of the *ket* vector $|\psi\rangle$ in the $|\pm\rangle_n$ basis. Since vectors can be added, what does it mean to add two spin states? Unlike directed arrows in a plane, we cannot visualize the sum of two ket vectors $|\psi\rangle$ and $|\phi\rangle$ in a simple geometrical way (such as the parallelogram law for directed arrows). The reason is that unlike directed arrows, the components of these vectors are *complex* numbers. However, we can mathematically define addition of two *ket* vectors and multiplication of a vector with a (complex) number by referring to the basis. Say, we are given two ket vectors, $|\psi\rangle$ and $|\phi\rangle$. We expand them in a suitable basis, which consists of vectors $|\alpha\rangle_+$ and $|\alpha\rangle_-$, these being spin states along some direction

$$\begin{aligned} |\psi\rangle &= a_+ |\alpha\rangle_+ + a_- |\alpha\rangle_- \\ |\phi\rangle &= b_+ |\alpha\rangle_+ + b_- |\alpha\rangle_- \end{aligned} \quad (1.23)$$

where (a_+, a_-) and (b_+, b_-) are components of $|\psi\rangle$ and $|\phi\rangle$ in this basis. Next, analogous to two directed arrows, we *define* the sum of $|\psi\rangle$ and $|\phi\rangle$ as

$$|\psi\rangle + |\phi\rangle = (a_+ + b_+) |\alpha\rangle_+ + (a_- + b_-) |\alpha\rangle_- \quad (1.24)$$

and multiplication of $|\psi\rangle$ with a complex number as

$$c|\psi\rangle = (ca_+) |\alpha\rangle_+ + (ca_-) |\alpha\rangle_- \quad (1.25)$$

Since we are identifying spin states as abstract vectors, we would like to interpret the sum of two such vectors (which should be another vector) as representing some state. At this point, we encounter a problem. Given a state $|\psi\rangle$ expanded in a basis $\{|\alpha\rangle_+, |\alpha\rangle_-\}$ as

$$|\psi\rangle = a_+ |\alpha\rangle_+ + a_- |\alpha\rangle_- \quad (1.26)$$

we would like to interpret the modulus squared of complex coefficients a_\pm as the respective probabilities of state $|\psi\rangle$ changing to states $|\alpha\rangle_\pm$ during a measurement process⁷

$$\begin{aligned} P_\psi(\alpha_+) &= |a_+|^2 \\ P_\psi(\alpha_-) &= |a_-|^2 \end{aligned} \quad (1.27)$$

with $P_\psi(\alpha_+) + P_\psi(\alpha_-) = |a_+|^2 + |a_-|^2 = 1$. Similarly, given another state $|\phi\rangle$

$$|\phi\rangle = b_+ |\alpha\rangle_+ + b_- |\alpha\rangle_- \quad (1.28)$$

we will calculate the probability of the state changing to $|\alpha\rangle_\pm$ as

$$\begin{aligned} P_\phi(\alpha_+) &= |b_+|^2 \\ P_\phi(\alpha_-) &= |b_-|^2 \end{aligned} \quad (1.29)$$

⁷A measurement here implies passing the atom in state $|\psi\rangle$ through a filter which produces states $|\alpha\rangle_\pm$, corresponding to spin measurements $\pm\hbar/2$ along the direction of the magnetic field in the filter.

such that $P_\phi(\alpha_+) + P_\phi(\alpha_-) = |b_+|^2 + |b_-|^2 = 1$. However, now consider a linear superposition of these states

$$\begin{aligned} |\chi\rangle &= c|\psi\rangle + d|\phi\rangle \\ &= (ca_+ + db_+)|\alpha\rangle_+ + (ca_- + db_-)|\alpha\rangle_- \\ &= e_+|\alpha\rangle_+ + e_-|\alpha\rangle_- \end{aligned} \quad (1.30)$$

where

$$\begin{aligned} e_+ &= ca_+ + db_+ \\ e_- &= ca_- + db_- \end{aligned} \quad (1.31)$$

We *cannot* interpret $|e_\pm|^2$ as the probability of state $|\chi\rangle$ changing to states $|\alpha\rangle_\pm$. This is because given $|a_+|^2 + |a_-|^2 = 1$ and $|b_+|^2 + |b_-|^2 = 1$, it does *not* follow that $|e_+|^2 + |e_-|^2 = 1$ for arbitrary complex numbers c and d . Therefore, if we wish to identify spin states as vectors in some abstract *vector space* with spin states along some direction in space as a basis, we cannot identify the modulus squared of expansion coefficients as the *absolute* probabilities of measurement of the spin values. However, since we *do* wish to identify these coefficients with probabilities, we can get around this problem by identifying their moduli (squared) with *relative* probabilities. Then, given (1.26), we interpret that

$$\frac{P_\psi(\alpha_+)}{P_\psi(\alpha_-)} = \frac{|a_+|^2}{|a_-|^2} \quad (1.32)$$

There is another point of view that we can take (and one we will stick to). This point of view states that the possible spin states of an electron are vectors in an abstract vector space. As in any vector space, there are an infinite number of vectors in this space, since given a set of vectors, we can construct arbitrary linear combinations of these vectors to produce other vectors. We will assume that any such vector can be expressed in the form

$$|\psi\rangle = c_+|\alpha_+\rangle + c_-|\alpha_-\rangle \quad (1.33)$$

without any constraint on the expansion coefficients c_+ and c_- . However, the vectors which represent *physical* states of electron spin are such which satisfy the constraint

$$|c_+|^2 + |c_-|^2 = 1 \quad (1.34)$$

With this constraint, the moduli squared of the expansion coefficients will be equal to the associated probabilities. Given an arbitrary vector in this vector space, we can construct a vector representing a physical state as follows. Say, the vector is given by (1.33), but without the constraint (1.34). Consider the vector $|\psi\rangle_P$ proportional to $|\psi\rangle$

$$|\psi\rangle_P = \frac{1}{\sqrt{|c_+|^2 + |c_-|^2}} (c_+|\alpha_+\rangle + c_-|\alpha_-\rangle) \quad (1.35)$$

It is easy to see that the coefficients of this vector satisfy the constraint for a physical state. We will soon see that the physical states can be interpreted as *unit* vectors (vectors of unit ‘length’) in the vector space.

At this point, we can investigate why we need to define our vectors over *complex* numbers. After all, the vector space of directed arrows (with which we have drawn analogies) is defined over real numbers. The reason is that if we were to just stick to real numbers, we will not be able to represent all physical spin states as vectors in our vector space. To see this, say we choose spin states $|+\rangle_z$ and $|-\rangle_z$ along the z -direction as a basis in our vector space. Consider the spin state $|+\rangle_x$, which we express in this basis as

$$|+\rangle_x = c_+|+\rangle_z + c_-|-\rangle_z \quad (1.36)$$

Since we wish to interpret it as a physical state, it should satisfy (1.34). Now, we know that if a beam of atoms with spin component $+\hbar/2$ along some direction is made to pass through a SG filter with magnetic

field along a perpendicular direction, the beam splits into two beams of equal intensity, corresponding to spin components $\pm\hbar/2$ along the direction of the magnetic field. Then, a beam in state $|+\rangle_x$ should split into two beams of equal intensity when passed through a filter along the z -direction. The atoms in the two beams will be in states $|+\rangle_z$ and $|-\rangle_z$ respectively. Then, a given atom in state $|+\rangle_x$ has probability $1/2$ each of going to state $|+\rangle_z$ and $|-\rangle_z$. Then, given our interpretation of (1.36), it follows that $|c_+|^2 = 1/2$ and $|c_-|^2 = 1/2$. If the expansion coefficients were real, this would result in two physically distinct possibilities: $c_+ = 1/\sqrt{2}, c_- = 1/\sqrt{2}$ and $c_+ = 1/\sqrt{2}, c_- = -1/\sqrt{2}$ ⁸. One of these will correspond to the state $|+\rangle_x$. The other possibility will presumably correspond to state $|-\rangle_x$ (as we will eventually demonstrate). Then, the spin states along the x -direction will be

$$\begin{aligned} |+\rangle_x &= \frac{1}{\sqrt{2}} |+\rangle_z + \frac{1}{\sqrt{2}} |-\rangle_z \\ |-\rangle_x &= \frac{1}{\sqrt{2}} |+\rangle_z - \frac{1}{\sqrt{2}} |-\rangle_z \end{aligned} \tag{1.37}$$

We now encounter the following problem. Consider atoms with spin along the y -direction, say $S_y = +\hbar/2$. A beam of such atoms will again split into two beams of equal intensity when passed through a z -filter (since these are again perpendicular directions). If we try to mathematically construct vectors representing physical spin states $|+\rangle_y$ and $|-\rangle_y$, if we insist on using real coefficients, we will once again end up with the same linear combinations of $|+\rangle_z$ and $|-\rangle_z$ as in (1.37). However, this will imply that the y -states are identical to the x -states, which is not true. Indeed, a beam of atoms in an x spin state will once again split into two beams of equal intensity when passed through a y -filter, since these directions are once again perpendicular. Therefore, if we wish to represent the y -spin states as linear superpositions of the z -states, we will need to enlarge the set of numbers to complex numbers. Then, for instance, the y -spin states could be the following linear combinations, consistent with the constraint (1.34) and the fact that an atom in state $|+\rangle_y$ (or $|-\rangle_y$) has equal probability of ending up in states $|+\rangle_z$ and $|-\rangle_z$

$$\begin{aligned} |+\rangle_y &= \frac{1}{\sqrt{2}} |+\rangle_z + \frac{i}{\sqrt{2}} |-\rangle_z \\ |-\rangle_y &= \frac{1}{\sqrt{2}} |+\rangle_z - \frac{i}{\sqrt{2}} |-\rangle_z \end{aligned} \tag{1.38}$$

These linear combinations are clearly distinct from the combinations in (1.37).

⁸The other two possibilities, $c_+ = -1/\sqrt{2}, c_- = -1/\sqrt{2}$ and $c_+ = -1/\sqrt{2}, c_- = 1/\sqrt{2}$ will result in vectors proportional to those generated by the other two possibilities. Vectors proportional to each other represent the same physical state.

Chapter 2

Spin Dynamics

2.1 Energy

We have seen that a spin state $|\psi(t)\rangle = c_+ |+\rangle_z(t) + c_- |-\rangle_z(t)$, written as a column vector

$$\psi(t) = \begin{pmatrix} c_+(t) \\ c_-(t) \end{pmatrix}$$

satisfied the differential equation

$$i\hbar \frac{d\psi(t)}{dt} = H\psi(t) \quad (2.1)$$

with solution

$$\psi(t) = e^{-iH(t-t_0)/\hbar} \psi(t_0) \quad (2.2)$$

Matrix H is given by

$$\begin{aligned} H &= -\boldsymbol{\mu} \cdot \mathbf{B} \\ &\text{with} \\ \boldsymbol{\mu} &= -\frac{ge\mathbf{S}}{2m_e} \end{aligned} \quad (2.3)$$

g is the 'structure factor' for the electron, and is experimentally determined to be 2. The 'vector' $\mathbf{S} = \frac{\hbar}{2}\boldsymbol{\sigma}$ is a set of three 'spin-matrices' S_x, S_y, S_z , the quantum analogue of classical spin. Here, $\boldsymbol{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$ are the Pauli matrices

$$\begin{aligned} \sigma_x &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \\ \sigma_y &= \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \\ \sigma_z &= \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \end{aligned} \quad (2.4)$$

In this form, the matrix H has the same form as the classical energy of a (spin-induced) magnetic moment interacting with a magnetic field, except that the classical spin is replaced by the set of Hermitian matrices $\mathbf{S} = (S_x, S_y, S_z)$. Given this, it appears that H is the quantum analogue of energy, in the same spirit as \mathbf{S} is the quantum analogue of spin angular momentum. For this reason, the matrix H is known as the 'Hamiltonian' matrix (energy in classical mechanics is often referred to as Hamiltonian).

We have seen that given a direction in space represented by a unit vector $\hat{\mathbf{n}}$, the matrix $\mathbf{S} \cdot \hat{\mathbf{n}}$ represents the component of spin along $\hat{\mathbf{n}}$. Further, an eigenvector of this matrix physically represents a state in which spin along this direction is well-defined, with value equal to the corresponding eigenvalue. Let us

take this principle further. Given that the Hamiltonian matrix H is the quantum analogue of energy, it is conceivable that eigenvectors of H should represent states in which energy is well defined, with energy in such a state equal to the corresponding eigenvalue. This also implies that since an arbitrary state will not be an eigenvector of H , in such a state, energy will in general not be well-defined (just as a given spin component need not be well-defined in an arbitrary state). Let $|E\rangle = c_+(E)|+\rangle_z + c_-(E)|-\rangle_z$ be a state of definite energy, with energy equal to E . Then, the column vector $\psi_E = \begin{pmatrix} c_+ \\ c_- \end{pmatrix}$ must satisfy the equation

$$H\psi_E = E\psi_E \quad (2.5)$$

Let the magnetic field be along some direction $\hat{\mathbf{n}}$. Then the Hamiltonian matrix is given by $H = \left(\frac{eB}{m_e}\right)\mathbf{S}\cdot\hat{\mathbf{n}}$. Since H is proportional to $\mathbf{S}\cdot\hat{\mathbf{n}}$, it follows that eigenvectors of $\mathbf{S}\cdot\hat{\mathbf{n}}$ are eigenvectors of H . We have seen that the eigenvectors of $\mathbf{S}\cdot\hat{\mathbf{n}}$ are

$$\begin{aligned} \psi_{+n} &= \begin{pmatrix} \cos(\theta/2) \\ \sin(\theta/2)e^{i\phi} \end{pmatrix} \\ \psi_{-n} &= \begin{pmatrix} \sin(\theta/2) \\ -\cos(\theta/2)e^{i\phi} \end{pmatrix} \end{aligned} \quad (2.6)$$

Since $(\mathbf{S}\cdot\hat{\mathbf{n}})\psi_{+n} = \frac{\hbar}{2}\psi_{+n}$ and $(\mathbf{S}\cdot\hat{\mathbf{n}})\psi_{-n} = -\frac{\hbar}{2}\psi_{-n}$, it follows that

$$\begin{aligned} H\psi_{+n} &= E_+\psi_{+n} \\ H\psi_{-n} &= E_-\psi_{-n} \end{aligned} \quad (2.7)$$

where $E_+ = \frac{eB\hbar}{2m_e}$ and $E_- = -\frac{eB\hbar}{2m_e}$ are the energy eigenvalues. Then, in ket notation, the states of definite energy are $|E_+\rangle = |+\rangle_n$ and $|E_-\rangle = |-\rangle_n$ with $|+\rangle_n$ and $|-\rangle_n$ being spin states along direction $\hat{\mathbf{n}}$. Since these spin states form a basis in this vector space, so do these energy states. Then, an arbitrary state can be expressed as

$$|\psi\rangle = c_+|E_+\rangle + c_-|E_-\rangle \quad (2.8)$$

If a measurement of energy is made on this state, one of two values- E_+ and E_- will be measured, with probabilities $|c_+|^2$ and $|c_-|^2$ respectively. Further, if E_+ is measured, the state will collapse to $|E_+\rangle$ and if E_- is measured, it will collapse to $|E_-\rangle$.

States of definite energy are very special, and have interesting properties. Given an arbitrary state at some instant, due to interactions, such a state will in general change with time in a complicated way. However, if this happens to be a state of definite energy, it will remain 'frozen', not changing with time. To see this, let the state at instant $t = 0$ be $|\psi(0)\rangle = |E\rangle$, represented by the column vector ψ_E . Then, the state at instant $t > 0$ will be given by (2.2)

$$\psi(t) = e^{-iHt/\hbar}\psi_E$$

Using (2.5) and the result that $f(H)\psi_E = f(E)\psi_E$ where $f(H)$ is a function of matrix H , we get

$$\psi(t) = e^{-iEt/\hbar}\psi_E \quad (2.9)$$

which is physically the same state, since it is just the state $|E\rangle$ multiplied by a phase $e^{-iEt/\hbar}$. In ket notation, this tells us that if $|\psi(0)\rangle = |E\rangle$ then $|\psi(t)\rangle = e^{-iEt/\hbar}|E\rangle$. Therefore, the time-evolution of states of definite energy is very simple-they do not change with time. This can be thought of as the analogue of energy conservation in quantum mechanics. If we start with a state in which energy is well-defined, then this energy does not change with time. It is natural to wonder what happens to energy conservation if we let a state, which does not have a well-defined energy, evolve with time. Since energy eigenvectors form a basis in the vector space, we can expand the state at instant $t = 0$ as a linear superposition of these states. In case of our spin system, such an expansion will be $|\psi(0)\rangle = c_+(0)|E_+\rangle + c_-(0)|E_-\rangle$. In the column

vector representation, this is just $\psi(0) = c_+(0)\psi_{E_+} + c_-(0)\psi_{E_-}$ (where ψ_{E_+} and ψ_{E_-} are actually the same as ψ_{+n} and ψ_{-n}). This state will evolve with time according to (2.2)

$$\begin{aligned}\psi(t) &= e^{-iHt/\hbar}\psi(0) \\ &= e^{-iHt/\hbar}(c_+(0)\psi_{E_+} + c_-(0)\psi_{E_-}) \\ &= c_+(0)e^{-iHt/\hbar}\psi_{E_+} + c_-(0)e^{-iHt/\hbar}\psi_{E_-}\end{aligned}$$

where we have used the linearity of matrix multiplication. Using the fact that ψ_{E_+} and ψ_{E_-} are eigenvectors of H with eigenvalues E_+ and E_- respectively, we get

$$\begin{aligned}\psi(t) &= c_+(0)e^{-iE_+t/\hbar}\psi_{E_+} + c_-(0)e^{-iE_-t/\hbar}\psi_{E_-} \\ &= c_+(t)\psi_{E_+} + c_-(t)\psi_{E_-}\end{aligned}\tag{2.10}$$

where $c_+(t) = c_+(0)e^{-iE_+t/\hbar}$ and $c_-(t) = c_-(0)e^{-iE_-t/\hbar}$. If we make a measurement of energy on this state, the probability that we measure $E = E_+$ at instant t will be $P_{E_+}(t) = |c_+(t)|^2 = |c_+(0)e^{-iE_+t/\hbar}|^2 = |c_+(0)|^2$ which is the same as the probability of measuring this energy at $t = 0$. Similarly, the probability that we measure $E = E_-$ at instant t will be the same as the probability of measuring this energy at $t = 0$. Then, even though energy may not be well-defined in a general quantum state, the probability (and therefore statistical measurements) of measuring a certain value of energy is independent of time. This is the analogue of energy conservation in Quantum Mechanics.

2.2 Time dependent interaction

We now consider the dynamics of quantum spin in presence of a time-varying magnetic field. When the interaction depends on time, equation (2.2) does not correctly describe the time-evolution of the spin state. Instead, we need to start from first principles. Consider once again the state at instant t , expanded in the $|+\rangle_z, |-\rangle_z$ basis

$$|\psi(t)\rangle = c_+(t)|+\rangle_z + c_-(t)|-\rangle_z\tag{2.11}$$

As before, this will be represented by a column vector

$$\psi(t) = c_+(t)\psi_{+z} + c_-(t)\psi_{-z}\tag{2.12}$$

where $\psi_{+z} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\psi_{-z} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$. The state at instant $t' > t$ will be given by unitary evolution of $\psi(t)$

$$\psi(t') = U(t', t)\psi(t)\tag{2.13}$$

where $U(t', t)$ is the time evolution matrix, with properties

$$\begin{aligned}U(t', t)^\dagger U(t', t) &= I \\ U(t, t) &= I \\ U(t_2, t_1)U(t_1, t_0) &= U(t_2, t_0); t_2 > t_1 > t_0\end{aligned}\tag{2.14}$$

Given the state at instant t , the state at instant $t + dt$ will be

$$\begin{aligned}\psi(t + dt) &= U(t + dt, t)\psi(t) \\ &\simeq \left(I - \frac{idt}{\hbar}H(t) \right) \psi(t)\end{aligned}$$

where as before, to ensure that U is unitary, the matrix H will be Hermitian. The important difference now is that since the interaction depends on time, this matrix will depend on time. We reorganize terms to get

$$i\hbar \frac{1}{dt} (\psi(t+dt) - \psi(t)) = H(t)\psi(t)$$

Taking the limit $dt \rightarrow 0$, we finally get

$$i\hbar \frac{d\psi}{dt} = H(t)\psi(t) \quad (2.15)$$

which has the same form as (2.1). It is understood here that

$$\frac{d\psi}{dt} = \begin{pmatrix} \dot{c}_+(t) \\ \dot{c}_-(t) \end{pmatrix} \quad (2.16)$$

with $\dot{c}_+ = dc_+/dt$, etc. For logical consistency, matrix $H(t)$ must still be the quantum analogue of energy, except that it will be time-dependent. For instance, if the electron spin interacts with an external time-dependent magnetic field $\mathbf{B}(t)$, the form of $H(t)$ must be

$$H(t) = \frac{e}{m_e} \mathbf{B}(t) \cdot \mathbf{S} \quad (2.17)$$

Unlike (2.1), equation (2.15) cannot be formally integrated into a simple closed form. However, we can still solve it by reducing it to coupled differential equations involving the column vector entries of $\psi(t)$. As a concrete example, let us analyze the phenomenon of Magnetic Resonance. In presence of a static external magnetic field $\mathbf{B}_0 = B_0 \hat{\mathbf{z}}$, an electron has energy states $|E_{\pm}\rangle$ with energy $E_{\pm} = \pm \frac{eB_0\hbar}{2m_e}$. If we start with the electron in any of these two states, its state will not change with time. In particular, if it has energy E_- (corresponding to spin 'parallel' to \mathbf{B}_0), this energy will be conserved. What if we wish to 'excite' this electron to the higher energy $|E_+\rangle$? To achieve this, we add an additional, harmonic time-varying magnetic field $\mathbf{B}_1 = B_1 \cos \omega t \hat{\mathbf{x}}$. With this, the total magnetic field will be $\mathbf{B}(t) = B_0 \hat{\mathbf{z}} + B_1 \cos \omega t \hat{\mathbf{x}}$, and the Hamiltonian matrix will be

$$\begin{aligned} H(t) &= \frac{e}{m_e} \mathbf{B}(t) \cdot \mathbf{S} \\ &= \frac{e}{m_e} (B_0 \hat{\mathbf{z}} + B_1 \cos \omega t \hat{\mathbf{x}}) \cdot \mathbf{S} \\ &= \frac{e}{m_e} (B_0 S_z + B_1 \cos \omega t S_x) \end{aligned} \quad (2.18)$$

We write this compactly as

$$H = H_0 + H_1(t) \quad (2.19)$$

where $H_0 = \frac{eB_0\hbar}{2m_e} \sigma_z$ is the original time-independent Hamiltonian and $H_1 = \frac{eB_1\hbar}{2m_e} \cos \omega t \sigma_x$ is the additional time-dependent piece. If we write these matrices explicitly, we get

$$\begin{aligned} H_0 &= \begin{pmatrix} E_+ & 0 \\ 0 & E_- \end{pmatrix} \\ H_1(t) &= \begin{pmatrix} 0 & E_1 \cos \omega t \\ E_1 \cos \omega t & 0 \end{pmatrix} \end{aligned} \quad (2.20)$$

where $E_{\pm} = \pm \frac{eB_0\hbar}{2m_e}$ are the energy eigenvalues in absence of the time-dependent interaction, and $E_1 = \frac{eB_1\hbar}{2m_e}$ has dimensions of energy. Let us define $\omega_0 = \frac{E_+ - E_-}{\hbar}$ and $\omega_1 = \frac{2E_1}{\hbar}$ which have dimensions of angular frequency. With this, the Hamiltonian matrix takes the form

$$H(t) = \frac{\hbar}{2} \begin{pmatrix} \omega_0 & \omega_1 \cos \omega t \\ \omega_1 \cos \omega t & -\omega_0 \end{pmatrix} \quad (2.21)$$

Substituting this form in (2.15), along with $\psi(t) = \begin{pmatrix} c_+(t) \\ c_-(t) \end{pmatrix}$ gives

$$i\hbar \begin{pmatrix} \dot{c}_+(t) \\ \dot{c}_-(t) \end{pmatrix} = \frac{\hbar}{2} \begin{pmatrix} \omega_0 & \omega_1 \cos \omega t \\ \omega_1 \cos \omega t & -\omega_0 \end{pmatrix} \begin{pmatrix} c_+(t) \\ c_-(t) \end{pmatrix} \quad (2.22)$$

These are coupled, first order differential equations in expansion coefficients c_+ and c_- . Writing them out explicitly give

$$\begin{aligned} i \dot{c}_+(t) &= \frac{\omega_0}{2} c_+(t) + \frac{\omega_1}{2} \cos \omega t c_-(t) \\ i \dot{c}_-(t) &= \frac{\omega_1}{2} \cos \omega t c_+(t) - \frac{\omega_0}{2} c_-(t) \end{aligned} \quad (2.23)$$

If the time-varying magnetic field is not there, then $\omega_1 = 0$ and these equations reduce to

$$\begin{aligned} i \dot{c}_+(t) &= \frac{\omega_0}{2} c_+(t) \\ i \dot{c}_-(t) &= -\frac{\omega_0}{2} c_-(t) \end{aligned} \quad (2.24)$$

These are decoupled equations and are easily solved to get

$$\begin{aligned} c_+(t) &= c_+(0) e^{-i\omega_0 t/2} \\ c_-(t) &= c_-(0) e^{i\omega_0 t/2} \end{aligned} \quad (2.25)$$

It is easy to see that these solutions are what we would expect if the time-dependent interaction is not there. Since the static magnetic field is in the $\hat{\mathbf{z}}$ direction, states $|+\rangle_z$ and $|-\rangle_z$ are just states $|E_+\rangle$ and $|E_-\rangle$, the states of definite energy. Then, $c_{\pm}(t)$ also happen to be expansion coefficients of the state $|\psi\rangle(t)$ in the energy basis. That is, equation (2.12) is equivalent to

$$|\psi(t)\rangle = c_+(t) |E_+\rangle + c_-(t) |E_-\rangle \quad (2.26)$$

We have already encountered this in the column vector form in (2.11)

$$\psi(t) = c_+(0) e^{-iE_+ t/\hbar} \psi_{E_+} + c_-(0) e^{-iE_- t/\hbar} \psi_{E_-}$$

Then, $c_+(t) = c_+(0) e^{-iE_+ t/\hbar}$ and $c_-(t) = c_-(0) e^{-iE_- t/\hbar}$. It is easy to check that (2.25) is equivalent to this, given that $E_{\pm} = \pm (eB_0\hbar/2m_e)$ and $\omega_0 = (E_+ - E_-)/\hbar$.

Now, let us assume that at $t = 0$, the spin is in state $|E_-\rangle$. Then, $c_+(0) = 0$ and $c_-(0) = 1$. In the absence of the time-varying magnetic field \mathbf{B}_1 , equations (2.25) give $c_+(t) = 0$ and $c_-(t) = e^{i\omega_0 t/2} = e^{-iE_- t/\hbar}$. Then, $\psi(t) = e^{-iE_- t/\hbar} |E_-\rangle$. The state therefore does not change with time, and the probability that a measurement of energy on the state will give $E = E_+$ is zero for all instant t . We now demonstrate that with the time-dependent magnetic field present, this probability is not equal to zero, and by suitably tuning the frequency of this time-varying field, this probability can in fact even be unity. We now return to equations (2.23), with the initial condition that at $t = 0$, the spin state is $|E_-\rangle$. That is, $c_+(0) = 0$ and $c_-(0) = 1$. First, we will factor out the $e^{\mp i\omega_0 t/2}$ dependence in $c_{\pm}(t)$, and define a new pair of coefficients $a(t) = e^{i\omega_0 t/2} c_+(t)$ and $b(t) = e^{-i\omega_0 t/2} c_-(t)$. From (2.23), it follows that these satisfy

$$\begin{aligned} i \dot{a}(t) &= \frac{\omega_1}{2} \cos \omega t e^{i\omega_0 t} b(t) \\ i \dot{b}(t) &= \frac{\omega_1}{2} \cos \omega t e^{-i\omega_0 t} a(t) \end{aligned} \quad (2.27)$$

Using $\cos \omega t = (e^{i\omega t} + e^{-i\omega t})/2$, we get

$$\begin{aligned} i \dot{a}(t) &= \frac{\omega_1}{4} \left(e^{i(\omega_0+\omega)t} + e^{i(\omega_0-\omega)t} \right) b(t) \\ i \dot{b}(t) &= \frac{\omega_1}{4} \left(e^{-i(\omega_0-\omega)t} + e^{-i(\omega_0+\omega)t} \right) a(t) \end{aligned} \quad (2.28)$$

We now argue that relative to the terms $e^{\pm i(\omega_0-\omega)t}$, the terms $e^{\pm i(\omega_0+\omega)t}$ are not important for the dynamics of coefficients $a(t)$ and $b(t)$, and can be discarded. The idea is that the product of a rapidly varying and a smoothly varying function of time averages to zero. We will be eventually interested in the range of ω such that it is close to ω_0 . Then, terms such as $e^{\pm i(\omega_0+\omega)t}$ will give rise to rapidly oscillating functions $\cos(\omega_0+\omega)t$ and $\sin(\omega_0+\omega)t$, and $e^{\pm i(\omega_0-\omega)t}$ will give rise to slowly oscillating functions $\cos(\omega_0-\omega)t$ and $\sin(\omega_0-\omega)t$ (since $\omega - \omega_0$ is small). When multiplied with smoothly varying functions $a(t)$ and $b(t)$, terms of the first kind will average out to values much less than the averages involving terms of the second kind. Then, we can ignore the $e^{\pm i(\omega_0+\omega)t}$ terms to get simpler coupled differential equations in $a(t)$ and $b(t)$

$$\begin{aligned} i \dot{a}(t) &= \frac{\omega_1}{4} e^{i(\omega_0-\omega)t} b(t) \\ i \dot{b}(t) &= \frac{\omega_1}{4} e^{-i(\omega_0-\omega)t} a(t) \end{aligned} \quad (2.29)$$

These can be decoupled to get

$$\begin{aligned} \ddot{a}(t) + i(\omega - \omega_0)\dot{a}(t) + \left(\frac{\omega_1}{4}\right)^2 a(t) &= 0 \\ \ddot{b}(t) - i(\omega - \omega_0)\dot{b}(t) + \left(\frac{\omega_1}{4}\right)^2 b(t) &= 0 \end{aligned} \quad (2.30)$$

These are simple second order differential equations with constant coefficients and will have two linearly independent solutions of the form $e^{-i\lambda t}$ where the two roots $\lambda = \lambda_+$ and $\lambda = \lambda_-$ are determined (through substitution) to be

$$\lambda_{\pm} = \frac{\omega - \omega_0}{2} \left(1 \mp \sqrt{1 + \left(\frac{\omega_1}{2(\omega - \omega_0)}\right)^2} \right) \quad (2.31)$$

With this, the solution for $a(t)$ will have the general form

$$a(t) = Ae^{-i\lambda_+t} + Be^{-i\lambda_-t}$$

where A and B are constants of integration. These are fixed by the initial conditions $a(0) = 0$ and $b(0) = 1$. We are interested in the probability that at instant t if we measure the energy of the system, we get $E = E_+$. This probability is given by $|c_+(t)|^2 = |e^{-i\omega_0 t/2} a(t)|^2 = |a(t)|^2$. The explicit form is as follows

$$|c_+(t)|^2 = \frac{\omega_1^2/4}{(\omega - \omega_0)^2 + \omega_1^2/4} \sin^2 \left(\frac{\sqrt{(\omega - \omega_0)^2 + \omega_1^2/4}}{2} t \right) \quad (2.32)$$

This is known as Rabi's formula. This formula tells us the following. A spin interacting with a static magnetic field can have two states of definite energy, $|E_+\rangle$ and $|E_-\rangle$, with $E_- < E_+$. If it is prepared in the state $|E_-\rangle$, it cannot make a 'transition' to the state $|E_+\rangle$ by itself, since energy is conserved (states of definite energy do not evolve with time). To induce such a transition, we need to introduce an additional time-dependent interaction (in the form of a time-varying magnetic field), since time-dependent interactions do not conserve energy. If we introduce such an interaction which is a harmonic function of time (with frequency ω), there is a non-zero probability (given by equation (2.32)) that the system makes a transition to the state $|E_+\rangle$. This probability varies periodically with time, with period $T = 2\pi/\sqrt{(\omega - \omega_0)^2 + \omega_1^2/4}$.

The maximum probability is $P_{\max} = \omega_1^2 / (4(\omega - \omega_0)^2 + \omega_1^2)$. Following is a plot of the maximum probability vs angular frequency ω

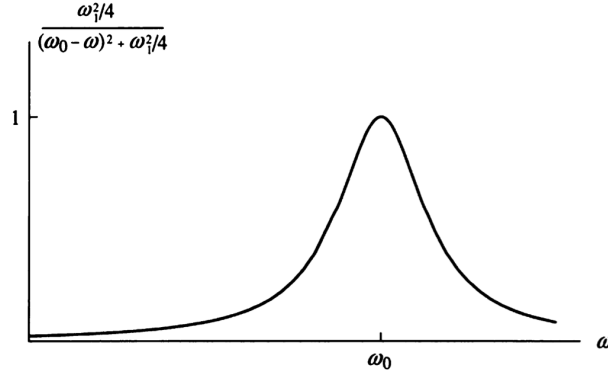


Figure 2.1: Magnetic-resonance transition probability as a function of frequency ω

The plot shows that if $\omega = \omega_0 = (E_+ - E_-)/\hbar$, the maximum probability is unity. This is the phenomenon of 'magnetic resonance'. With $\omega = \omega_0$, (2.32) reduces to

$$|c_+(t)|^2 = \sin^2\left(\frac{\omega_1 t}{4}\right) \quad (2.33)$$

which shows that the probability of detecting the electron in the higher energy state $|E_+\rangle$ oscillates between zero and one with period $T = 4\pi/\omega_1 = (4\pi m_e)/eB_1$. The resonance condition $\omega = (E_+ - E_-)/\hbar$ has a simple intuitive interpretation. An electromagnetic field oscillating with angular frequency ω can be visualised as a collection of photons, each with energy $\hbar\omega$. Then, at resonance, the transition of the system from a state of energy E_- to E_+ can be interpreted as the system having absorbed a photon of energy $\hbar\omega = E_+ - E_-$. This is consistent with the principle of conservation of energy. Then, with a time-dependent (harmonic) interaction, the spin system can get excited from a lower to higher energy state by absorbing the corresponding energy difference from the environment in the form of photons of the right frequency.

It is interesting to see what happens if the state at instant $t = 0$ is $|E_+\rangle$, the higher energy state. Then, the differential equations (2.30) need to be solved with initial conditions $a(0) = 1$ and $b(0) = 0$. It is easy to see from symmetry between $a(t)$ and $b(t)$ that the behavior of $b(t)$ will be similar to that of $a(t)$ in the previous situation. It can be checked easily that the probability that a measurement of energy of the system gives $E = E_-$ will be the same as in equation (2.32). In particular, at resonance ($\omega = \omega_0$), the system can make a transition to the lower energy state. This can be interpreted as the system having *emitted* a photon of energy $\hbar\omega = E_+ - E_-$.

Chapter 3

Two state systems

3.1 The two state approximation

A study of the quantum mechanics of electron spin provides us with tools and intuition to analyze more complex quantum systems. These systems are described by the same vector space formalism, except that the dimension of the vector space in which such quantum systems ‘live’ will generally be more than two (in fact, usually infinite!). However, often such complex systems can be approximately treated as if they live in a two dimensional vector space. Such an approximation, when applicable, can give important insight into certain aspects of these systems. The reason why such an approximate description often explains the physics of these systems quite well can be understood in the following way. Let us say that at some instant the quantum state of the system is a vector in a (complex) N-dimensional space. This vector can be expanded as a linear superposition over N linearly independent basis vectors, with the expansion coefficients containing numerical information about this state (relative probabilities, etc.) at this instant. For some special choice of basis, it may be possible that only two of the expansion coefficients have appreciable magnitude (compared with the other coefficients). Then, at this instant, we can approximately write

$$\psi = c_1 |\alpha_1\rangle + c_2 |\alpha_2\rangle \quad (3.1)$$

where $|\alpha_1\rangle$ and $|\alpha_2\rangle$ are those basis states (out of N) whose expansion coefficients (c_1 and c_2) are appreciable. It might seem that in such a situation we could perhaps pretend that the state lives in a two dimensional vector space spanned by these two states. However, it could be possible that as the state changes with time, the other coefficients (which we have ignored) could become appreciable at a later instant. In such a situation, this approximation will be a poor one. However, if the dynamics of the system is such that at all instants of physical interest the other expansion coefficients are negligible, then we may use this approximation.

Another situation in which we may use this approximation is one in which the system has approximately disjointed ‘degrees of freedom’, and one of these can be described by a state in a two dimensional vector space. For instance, when we talk about electron spin, we completely ignore its propagation in space. The situations in which we do this are usually such that the propagation of the electron is irrelevant (at least approximately) to the dynamics of its spin. For instance, for an electron in Potassium, its ‘propagation’ consists of it being bound to the Potassium nucleus in an orbital state, and as long as it is bound this way, we can ignore this aspect of its dynamics, and focus only on its spin. In such situations, it might even be possible to treat its spatial propagation classically (the Potassium atom’s ‘trajectory’), and spin dynamics quantum mechanically. In what follows, we often will not explicitly argue or demonstrate the validity of this approximation, but it will be assumed that the system it is being applied to satisfies such requirements.

Let us say that we have two states $|\alpha_1\rangle$ and $|\alpha_2\rangle$ in an N-dimensional vector space of quantum states, which happen to be of physical interest in the sense that they correspond to some physically identifiable configurations of the system. Further, we assume that these states are (at least approximately) orthogonal vectors in this vector space. We can always normalize each of these, so that $\langle\alpha_1|\alpha_1\rangle = \langle\alpha_2|\alpha_2\rangle = 1$ and $\langle\alpha_1|\alpha_2\rangle = \langle\alpha_2|\alpha_1\rangle = 0$. Assuming that our approximation of treating the vector space as two-dimensional

with these states forming a basis, a state of interest $|\psi\rangle$ can be expanded at an instant of time t (assuming that our approximation is valid at this instant) as

$$\psi = c_1(t) |\alpha_1\rangle + c_2(t) |\alpha_2\rangle \quad (3.2)$$

The time evolution of the state will be given by the analogue of (2.2)

$$\psi(t) = e^{-iH(t-t_0)/\hbar} \psi(t_0) \quad (3.3)$$

where H is a Hermitian matrix, the Hamiltonian matrix for this system. Based on our experience with electron spin, this will have the same form as the classical expression for energy, but with relevant dynamical variables (such as spin components in case of electron spin) represented by Hermitian matrices. This matrix will have two eigenvectors (since it is a 2x2 matrix) with corresponding eigenvalues. Let these eigenvectors be ψ_{E_1} and ψ_{E_2} with eigenvalues E_1 and E_2 respectively

$$\begin{aligned} H\psi_{E_1} &= E_1\psi_{E_1} \\ H\psi_{E_2} &= E_2\psi_{E_2} \end{aligned} \quad (3.4)$$

As before, we interpret ψ_{E_1} and ψ_{E_2} as column vector representations of states of definite energy $|E_1\rangle$ and $|E_2\rangle$ respectively, with eigenvalues E_1 and E_2 being the well-defined value of energy in these states. In general, states $|E_1\rangle$ and $|E_2\rangle$ will be *different* from states $|\alpha_1\rangle$ and $|\alpha_2\rangle$. It is a general mathematical result (discussed in 4) that eigenvalues of a Hermitian matrix are real and the corresponding eigenvectors (at least for distinct eigenvalues) are orthogonal. Then, states $|E_1\rangle$ and $|E_2\rangle$ are orthonormal (assuming that they have been normalized) and will also form a basis in this (approximately two-dimensional) vector space. In particular, states $|\alpha_1\rangle$ and $|\alpha_2\rangle$ will be linear superpositions of these states. Then, if we start with the initial state as either $|\alpha_1\rangle$ or $|\alpha_2\rangle$, such a state will change with time. In particular, it is possible that starting with one of these states, we can end up with the other one after a certain instant of time.

A special case of interest is as follows: the system has two states of well-defined energy, which are *degenerate*. What this implies is that there are two distinct states $|E_1\rangle$ and $|E_2\rangle$, both of which have the same energy, say $E_1 = E_2 = E_0$. Note that even though the energy of these states is the same, they are represented by *distinct* vectors, which are in fact orthogonal. It is easy to see that a system can have the same energy for different physical states. For instance, in classical physics, a free (non-interacting) particle at rest at two different points will have the same energy. Now, let us say we change the interaction of this system. Then, the matrix H will change, since it depends on the interaction of the system (the expression for energy of the system changes). As a result, states $|E_1\rangle$ and $|E_2\rangle$ (represented as column vectors) will no longer be eigenvectors of the new Hamiltonian matrix. This matrix will have new eigenvectors, representing the new states of definite energy. In general, these eigenvectors will have *different* eigenvalues. In such a situation, we say that *degeneracy has been lifted*, which is just a technical way of saying that with the introduction of the new interaction, the new energy states have different energies. At the same time, if we start the system in one of the old states of definite energy, this state will change with time, and it is possible to have a situation in which these states perform a kind of ‘flip-flop’ from one state to the other. Let us analyze the mathematical aspects of this situation. The idea is that even with the new interaction present, states $|E_1\rangle$ and $|E_2\rangle$ will still form a basis in this vector space (since they are a set of two orthonormal vectors). With these as a basis, their own column vector representations will trivially be

$$\begin{aligned} \psi_1 &= \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ \psi_2 &= \begin{pmatrix} 0 \\ 1 \end{pmatrix} \end{aligned} \quad (3.5)$$

Let us first consider the situation in which the external interaction is not there. Let the Hamiltonian matrix for this situation be H_0 . If the energy of the ‘degenerate’ states is E_0 , then in absence of the external interaction, the matrix form of H_0 will be

$$H_0 = \begin{pmatrix} E_0 & 0 \\ 0 & E_0 \end{pmatrix} \quad (3.6)$$

It is easy to check that ψ_1 and ψ_2 are eigenvectors of this matrix with eigenvalue(s) E_0 . With the new interaction present, the Hamiltonian matrix will change. Let the new Hamiltonian be H . We focus on a special case in which H has the form

$$H = \begin{pmatrix} E_0 & A \\ A & E_0 \end{pmatrix} \quad (3.7)$$

where A is real. It is clear that H is Hermitian, as it should be. Let us calculate the eigenvalues and eigenvectors of the new Hamiltonian. The eigenvalue equation for H (for eigenvalue E) is

$$\begin{vmatrix} E_0 - E & A \\ A & E_0 - E \end{vmatrix} = 0 \quad (3.8)$$

This gives eigenvalues

$$\begin{aligned} E_+ &= E_0 + A \\ E_- &= E_0 - A \end{aligned} \quad (3.9)$$

$$\begin{aligned} \psi_+ &= \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \\ \psi_- &= \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \end{aligned} \quad (3.10)$$

These eigenvectors can be expressed as a linear superposition of vectors ψ_1 and ψ_2

$$\begin{aligned} \psi_+ &= \frac{1}{\sqrt{2}}\psi_1 + \frac{1}{\sqrt{2}}\psi_2 \\ \psi_- &= \frac{1}{\sqrt{2}}\psi_1 - \frac{1}{\sqrt{2}}\psi_2 \end{aligned} \quad (3.11)$$

This tells us that the new states of definite energy are

$$\begin{aligned} |E_+\rangle &= \frac{1}{\sqrt{2}}|E_1\rangle + \frac{1}{\sqrt{2}}|E_2\rangle \\ |E_-\rangle &= \frac{1}{\sqrt{2}}|E_1\rangle - \frac{1}{\sqrt{2}}|E_2\rangle \end{aligned} \quad (3.12)$$

In turn, we can invert these equations to express states $|E_1\rangle$ and $|E_2\rangle$ in terms of the energy eigenstates

$$\begin{aligned} |E_1\rangle &= \frac{1}{\sqrt{2}}|E_+\rangle + \frac{1}{\sqrt{2}}|E_-\rangle \\ |E_2\rangle &= \frac{1}{\sqrt{2}}|E_+\rangle - \frac{1}{\sqrt{2}}|E_-\rangle \end{aligned} \quad (3.13)$$

Now, let us start with the system in state $|E_1\rangle$ at $t = 0$. Then, the state is described by the column vector

$$\psi(0) = \psi_1 = \frac{1}{\sqrt{2}}\psi_+ + \frac{1}{\sqrt{2}}\psi_- \quad (3.14)$$

which is the column vector representation of the first of equations (3.13). The state at instant t will then be given by (3.3)

$$\begin{aligned} \psi(t) &= e^{-iHt/\hbar}\psi(0) \\ &= \frac{1}{\sqrt{2}}e^{-iHt/\hbar}(\psi_+ + \psi_-) \\ &= \frac{1}{\sqrt{2}}\left(e^{-iHt/\hbar}\psi_+ + e^{-iHt/\hbar}\psi_-\right) \end{aligned} \quad (3.15)$$

since ψ_+ and ψ_- are eigenvectors of H with eigenvalues E_+ and E_- respectively, it follows that $e^{-iHt/\hbar}\psi_+ = e^{-iE_+t/\hbar}\psi_+$ and $e^{-iHt/\hbar}\psi_- = e^{-iE_-t/\hbar}\psi_-$. Then,

$$\psi(t) = \frac{1}{\sqrt{2}} \left(e^{-iE_+t/\hbar}\psi_+ + e^{-iE_-t/\hbar}\psi_- \right) \quad (3.16)$$

whose ket version is

$$|\psi(t)\rangle = \frac{1}{\sqrt{2}} \left(e^{-iE_+t/\hbar} |E_+\rangle + e^{-iE_-t/\hbar} |E_-\rangle \right) \quad (3.17)$$

Let us calculate the probability that the system will be detected in state $|E_2\rangle$ at some instant t . The probability amplitude for this to happen is $\langle E_2 | \psi(t)\rangle$, which can be calculated using the expansions for $|E_2\rangle$ and $|\psi(t)\rangle$ in the orthonormal energy basis (3.13)

$$\langle E_2 | \psi(t)\rangle = \frac{1}{2} \left(e^{-iE_+t/\hbar} - e^{-iE_-t/\hbar} \right) \quad (3.18)$$

The probability is $P_{E_1 \rightarrow E_2} = |\langle E_2 | \psi(t)\rangle|^2$, and is calculated to be

$$P_{E_1 \rightarrow E_2} = \sin^2 \left(\frac{\omega_0 t}{2} \right) \quad (3.19)$$

where $\omega_0 = (E_+ - E_-)/\hbar$. Similarly, the probability of detecting the system in state $|E_1\rangle$ will be $P_{E_1 \rightarrow E_1} = |\langle E_1 | \psi(t)\rangle|^2$, and is calculated to be

$$P_{E_1 \rightarrow E_1} = \cos^2 \left(\frac{\omega_0 t}{2} \right) \quad (3.20)$$

It is clear that the system performs a ‘flip-flop’ from $|E_1\rangle$ to $|E_2\rangle$ and vice-versa with a time period $T = 2\pi/\omega_0$. Note that the frequency of this flip-flop is the same as the frequency of a photon emitted by the system if it makes a transition from state $|E_+\rangle$ to $|E_-\rangle$.

Let us consider an application of the two-state approximation. We will try to get a feel for covalent-bond formation in molecules, and what it means to say that electrons are ‘shared’ by atoms in a covalent bond. To keep things simple, we will look at the covalent bond formation in a Hydrogen Molecular Ion, H_2^+ . This is a system of two protons and one electron, the electron being ‘shared’ by the two protons. Assume that the protons can be treated as classical spheres, located with their centres separated by distance d . The electron will interact with both protons. Classically, the potential energy experienced by the electron will have two minima, at the locations of the two protons. The actual problem is three-dimensional and the potential energy will be complicated, but we look at a one-dimensional ‘toy’ potential energy, which will broadly capture the essence of this situation. This is the so-called ‘double-well’ potential energy function

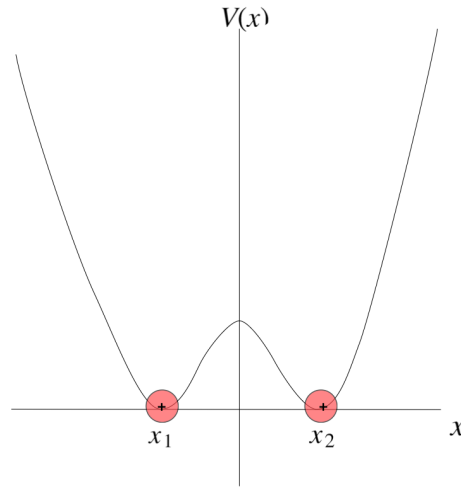


Figure 3.1: Double-Well potential with two local minima representing positions of the protons. Classically the electron has two minimum energy configurations in which it is at rest at the positions of the two minima

There are two classical minimum energy configurations of the electron, in which it is at rest at the location of either proton (points x_1 and x_2 in Figure 3.1). Now, let us analyze the quantum situation, and look at the analogue of the classical minimum energy configurations. Note that classically, as long as the electron is in one of the two configurations, it does not care if the other proton is there. If we remove the other proton, the position of the electron will still be one of minimum energy, though the overall potential energy function will change. The ‘toy’ potential energy function would be a ‘single-well’ potential energy function

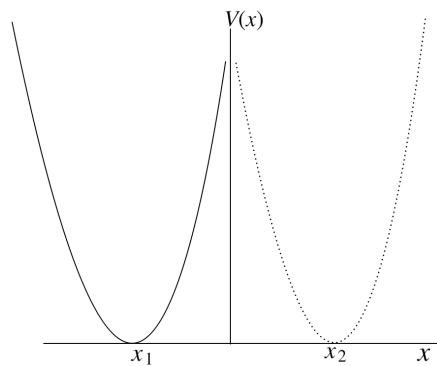


Figure 3.2: The ‘toy’ potential energy of the electron when only one proton is there

The quantum situation is quite different. The quantum analogue of the electron ‘not seeing the other proton’ would be the electron forming a Hydrogen atom with one proton, ignoring the existence of the other proton. Then, naively there will be two possible lowest energy states (‘ground-states’) $|E_1\rangle$ and $|E_2\rangle$ with the same energy E_0 , equal to the ground-state energy of an electron in a Hydrogen atom. Each state has a wavefunction which is localized around the proton to which the electron is bound

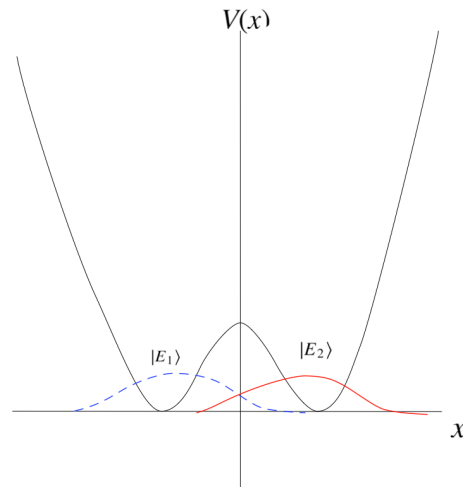


Figure 3.3: Cartoon representing the two hydrogen atom states. Each is represented by a real wavefunction

One can visualize the three-dimensional situation by plotting the corresponding Hydrogen-like probability distributions for the electron in each case

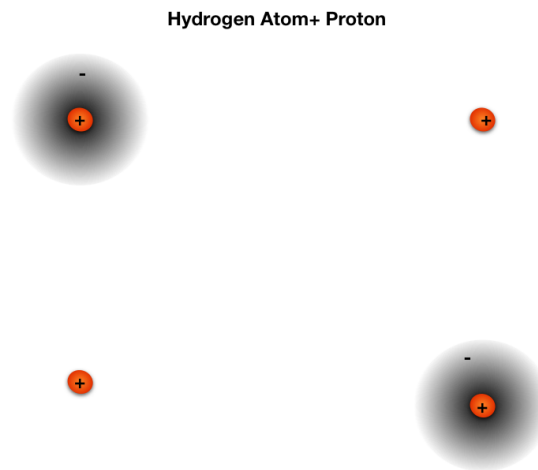


Figure 3.4: Electron probability distribution for ground state of Hydrogen represented by a continuous distribution of points in space around the protons. The density of points at a point is proportional to the modulus squared of the wavefunction at that point

Unlike the classical situation, since the electron is not localized, it ‘knows’ that the other proton is there. Figure 3.1 shows that each wavefunction has a ‘tail’ in the other well, and so the electron ‘knows’ that the other proton isn’t there. As a result, states $|E_1\rangle$ and $|E_2\rangle$ are not really states of well-defined energy, since they ignored the existence of the other proton. However, if the ‘tail’ of each electron wavefunction in the second well is tiny, $|E_1\rangle$ and $|E_2\rangle$ are a good zeroth order approximation to the energy states. To calculate the true energy states to the first order, we keep states $|E_1\rangle$ and $|E_2\rangle$ as a basis in our (approximately) two-dimensional vector space, represented by column vectors (3.5). To zeroth order, the energy matrix would be given by (3.6), with degenerate eigenvalue(s) E_0 , equal to the ground-state energy of the Hydrogen

atom. Then, a possible candidate for the corrected Hamiltonian could be the matrix (3.7)

$$H = \begin{pmatrix} E_0 & A \\ A & E_0 \end{pmatrix}$$

with the magnitude of A being a decreasing function of proton separation (closer the protons, larger the correction to the energy matrix). This form of the Hamiltonian is justified, since assuming that the change in the interaction (due to the presence of the second proton) is small, the most important change will be in the off-diagonal elements of H , since initially these were zero. There will in principle be a shift in the diagonal entries as well, but we will ignore it relative to the dominant term E_0 . We will assume that A is *negative*, this choice being consistent with physical bond-formation. Then, $A = -|A|$. The new energy eigenstates will be (3.12)

$$\begin{aligned} |E_+\rangle &= \frac{1}{\sqrt{2}} |E_1\rangle + \frac{1}{\sqrt{2}} |E_2\rangle \\ |E_-\rangle &= \frac{1}{\sqrt{2}} |E_1\rangle - \frac{1}{\sqrt{2}} |E_2\rangle \end{aligned}$$

with energy eigenvalues given by (3.9)

$$\begin{aligned} E_+ &= E_0 - |A| \\ E_- &= E_0 + |A| \end{aligned} \quad (3.21)$$

Since $E_+ < E_-$, the state $|E_+\rangle$ will be the equilibrium state of the molecule. Let us analyze properties of this state. If ψ_1 and ψ_2 represent the (Hydrogen atom) wavefunctions for states $|E_1\rangle$ and $|E_2\rangle$ respectively, it follows from (3.12) that the wavefunction corresponding to the ‘bonding’ state will be

$$\psi_{E_+} = \frac{1}{\sqrt{2}} \psi_{E_1} + \frac{1}{\sqrt{2}} \psi_{E_2} \quad (3.22)$$

Similarly, the wavefunction corresponding to the ‘anti-bonding’ state will be

$$\psi_{E_-} = \frac{1}{\sqrt{2}} \psi_{E_1} - \frac{1}{\sqrt{2}} \psi_{E_2} \quad (3.23)$$

Given the ‘toy’ wavefunctions ψ_1 and ψ_2 we had constructed, we can visualize the ‘toy’ form of the bonding and anti-bonding wavefunctions

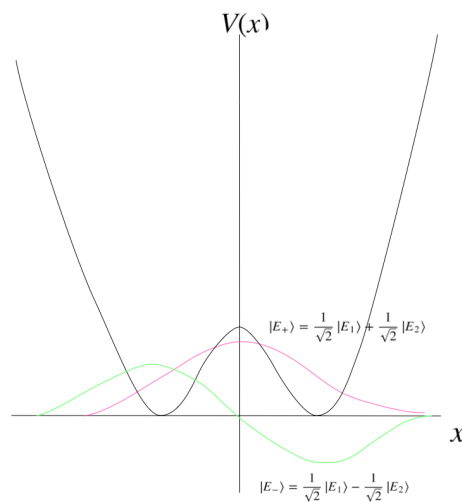


Figure 3.5: A cartoon of the ‘Bonding’ and ‘Anti-Bonding’ wavefunctions. The electron probability distribution for the bonding state is mostly localized between the protons. The anti-bonding distribution is localized away from the central region.

Let us analyze what determines the bond-length of the molecule. The electron energy $E_+ = E_0 - |A|$ decreases with decreasing proton separation (since $|A|$ increases with decreasing proton separation). The system will try to reduce this energy by drawing the protons closer. However, a competing influence is the proton-proton repulsion, which results in the energy of the system increasing with decreasing proton separation. At some separation, equilibrium is reached and this is the equilibrium bond length. Following is a possible three-dimensional visualization of the probability distributions in the Hydrogen, anti-bonding and bonding states

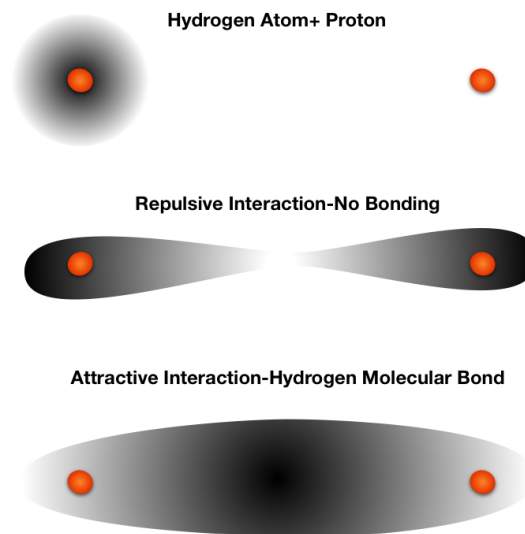


Figure 3.6: Visualization of three dimensional electron probability distributions corresponding to bonding and anti-bonding states

As analyzed before, if we start with the system in either state $|E_1\rangle$ or $|E_2\rangle$, the state will ‘flip-flop’ from one state to the other, since these are not energy eigenstates. Therefore, if we bring a proton close enough to a Hydrogen atom, such a flip-flop will be initiated and the only equilibrium, stationary configuration will be one which is a linear superposition of these states, with the correct phase (the ‘bonding’ state).

Chapter 4

N-State systems: General formulation

4.1 Quantum Mechanics of N-state systems

Having developed quantum mechanics of systems with two states, we now move on to more complex systems. Such systems ‘live’ in vector spaces with N dimensions, where N could even be infinite. Our experience with the two dimensional vector space of electron spin allows us to generalize the formalism to describe arbitrary quantum systems. We will also try to formalise the concepts we have developed in a more abstract notation, which is very versatile and powerful. There are several interesting examples of quantum systems whose states are described by vectors in N dimensional complex vector spaces. For instance, if we performed a Stern Gerlach experiment not on a beam of Potassium atoms but some other atomic species and observed the beam split into three (instead of two), we would need to describe the spin of such an atom as a vector in a three-dimensional vector space, instead of two. There are well-known examples of atoms whose beams, in a Stern Gerlach experiment, split into not two or three, but several beams, in general N . The spin of such an atom would be described by a vector in an N -dimensional space. As another example, consider the propagation of an electron in a regular, one dimensional lattice of N atoms. If an electron is introduced in this lattice, it can form a weakly bound negative ion with any one lattice atom. In such a quantum state, it will have a wavefunction localized around that atom (similar to what happens in a H_2^+ molecule, where an electron can form a Hydrogen atom state with one of the protons). Let us denote the quantum state in which the electron is localized to the n^{th} atom as $|x_n\rangle$

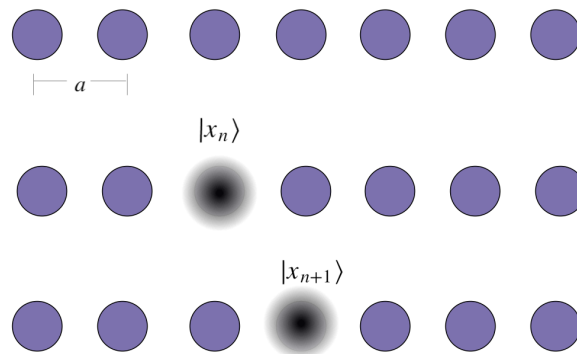


Figure 4.1: Electron states localized around lattice atoms

If the electron could not ‘see’ the other lattice atoms, these states would be states of well-defined energy, each with the same energy, say E_0 . However, as we saw in case of the H_2^+ molecule, since the electron’s wavefunction ‘leaks’ to neighboring sites, it *can* in a sense perceive the existence of the neighboring atoms. If the separation between the neighbours is large enough, it is a good approximation to assume that such an

electron can only ‘see’ the two nearest neighbors (if it is in state $|x_n\rangle$, it will interact weakly with the atoms at sites $n - 1$ and $n + 1$). The states $|x_n\rangle$ will not be states in which energy is well-defined. Instead, the states of definite energy will be linear superpositions of these states, which (approximately) form a basis in this vector space. Since the number of such states is N , we will be dealing with an N dimensional vector space. In this space, if we were to start with the electron in state $|x_n\rangle$, the state will change with time, with the electron ‘hopping’ from one lattice site to the next (the analogue of the flip-flop from one proton to the other in the H_2^+ molecule). This is then how an electron propagates in a solid. To better understand the quantum mechanics of such N dimensional systems, it is useful to introduce a more abstract formalism. This requires a study of Linear Algebra, and we briefly review the core ideas, applied to the vector space of quantum states.

Given a quantum system, we have seen that the state of the system is a vector in an N -dimensional complex vector space, the dimension N depending on the system. We denote such a vector by a *ket* symbol $|\psi\rangle$, introduced by Dirac. We shall call this vector space a *Hilbert Space*, denoted as \mathcal{H} . On these vectors, operations of multiplication of a vector by a complex number and addition of vectors are defined, such that

1. $a|\psi\rangle \in \mathcal{H} \quad \forall a \in \mathbb{C} \text{ and } |\psi\rangle \in \mathcal{H}$
2. $|\psi\rangle + |\phi\rangle = |\phi\rangle + |\psi\rangle \in \mathcal{H} \quad \forall |\psi\rangle, |\phi\rangle \in \mathcal{H}$

On this vector space, an inner product (analogue of the ‘dot product’) is defined. This is a rule that takes two vectors and gives a complex number. The inner product of vectors ψ and ϕ (denoted as $\langle\phi|\psi\rangle$) satisfies the following properties:

1. $\langle\phi|\psi\rangle \in \mathbb{C}$
2. $\langle\phi|\psi\rangle^* = \langle\psi|\phi\rangle$
3. $\langle\phi|a\psi_1 + b\psi_2\rangle = a\langle\phi|\psi_1\rangle + b\langle\phi|\psi_2\rangle$ where $|a\psi_1 + b\psi_2\rangle = a|\psi_1\rangle + b|\psi_2\rangle$ (Linearity)
4. $\langle\psi|\psi\rangle \geq 0 \quad \forall |\psi\rangle \in \mathcal{H}$

Using these properties, it is easy to see that the inner product also satisfies

$$\langle a\psi_1 + b\psi_2|\phi\rangle = a^* \langle\psi_1|\phi\rangle + b^* \langle\psi_2|\phi\rangle$$

In this Hilbert space, not every vector represents a physical quantum state. Physical states are those whose ‘norm’ $\langle\psi|\psi\rangle$ is unity. That is, they are ‘unit vectors’ in this vector space

$$\langle\psi|\psi\rangle = 1 \tag{4.1}$$

Given that the vector space has dimension N , it follows that any state vector can be expressed as a linear superposition of N linearly independent vectors, which form a ‘basis’ in this vector space. There will be an infinite number of such bases. Given a set of n linearly independent vectors, it is always possible to construct another set of n vectors all of which have unit norm and are orthogonal to each other¹. Such a set of vectors is called an *orthonormal set*. Then, it is always possible to construct orthonormal bases in the vector space. Unless stated otherwise, it will be assumed that a basis is orthonormal. Given an orthonormal basis $\{|\alpha_i\rangle\} = |\alpha_1\rangle, |\alpha_2\rangle, \dots, |\alpha_N\rangle$, a state $|\psi\rangle$ can be expanded as a linear superposition of these basis vectors

$$|\psi\rangle = \sum_{i=1}^N c_i |\alpha_i\rangle \tag{4.2}$$

where c_i are complex numbers. Since the basis is orthonormal, it follows that

$$\langle\alpha_i|\alpha_j\rangle = \delta_{ij} \tag{4.3}$$

¹Such a construction is called Gram Schmidt orthogonalisation, and is discussed in most textbooks on Linear Algebra

where δ_{ij} is the Kronecker delta symbol with the following property

$$\delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases} \quad (4.4)$$

Taking the inner product of (4.2) with $|\alpha_i\rangle$ and using (4.3) we get

$$c_i = \langle \alpha_i | \psi \rangle \quad (4.5)$$

This allows us to write (4.2) in a very useful form

$$|\psi\rangle = \sum_{i=1}^N |\alpha_i\rangle \langle \alpha_i | \psi \rangle \quad (4.6)$$

In this basis, the state vector $|\psi\rangle$ can be represented as a column vector

$$\psi = \begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ \vdots \\ \vdots \\ c_N \end{pmatrix} \quad (4.7)$$

where the entries are expansion coefficients of the state in the $\{|\alpha_i\rangle\}$ basis. The column vector representation of basis vector $|\alpha_i\rangle$ is

$$\alpha_i = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad (4.8)$$

where all entries are zeros, except the i^{th} entry which is 1. The column vector analogue of (4.2) is

$$\psi = \sum_{i=1}^N c_i \alpha_i \quad (4.9)$$

Let $|\phi\rangle$ be another vector in this vector space, with expansion

$$|\phi\rangle = \sum_{i=1}^N b_i |\alpha_i\rangle \quad (4.10)$$

Then, the inner product of $|\phi\rangle$ with $|\psi\rangle$ will be

$$\begin{aligned} \langle \phi | \psi \rangle &= \sum_i \sum_j b_i^* c_j \langle \alpha_i | \alpha_j \rangle \\ &= \sum_i \sum_j b_i^* c_j \delta_{ij} \\ &= \sum_i \sum_j b_i^* c_i \end{aligned} \quad (4.11)$$

where properties of the inner product and orthonormality of the basis vectors has been used. In the column vector representation, the inner product is given by

$$\langle \phi | \psi \rangle = \phi^\dagger \psi \quad (4.12)$$

where $\phi^\dagger = (\phi^T)^*$ is the Hermitian adjoint of ϕ .

In spin space of the electron, we saw that physically measurable quantities ('Observables') are represented by 2x2 matrices. For instance, we encountered matrix analogues of the three components of spin and the energy of the system when this spin interacts with an external magnetic field. We know that column vectors in this vector space represent abstract ket vectors in some basis. What abstract quantities do these matrices represent? Given a choice of basis, it turns out that *Linear Transformations* on vectors in a vector space are represented by matrices. A Linear Transformation is a rule that maps vectors in a vectors space to vectors in the same or different vector space, in a *linear* way (we will shortly define what we mean by 'linear'). For the purpose of quantum mechanics, of interest are linear transformations that map vectors in a vector space to vectors in the *same* vector space. Such linear transformations are called *Linear Operators*. Let \hat{T} be a linear operator. Then, by definition, it acts on a vector in \mathcal{H} to give another vector in \mathcal{H} . That is, if $|\psi\rangle \in \mathcal{H}$, then $\hat{T}|\psi\rangle \in \mathcal{H}$. The fact that it is a linear operator tells us about its action on a linear combination of vectors

$$\hat{T}(a|\psi\rangle + b|\phi\rangle) = a\hat{T}|\psi\rangle + b\hat{T}|\phi\rangle \quad (4.13)$$

It is easy to see that if a ket vector is represented as a column vector, a linear operator must be represented by an $N \times N$ matrix, given a choice of basis in \mathcal{H} . This is because an $N \times N$ matrix acts on a column vector to give another column vector in a linear manner, just as a linear operator acts on a ket vector. All that is left for us to do is to figure out what the matrix elements of a given linear operator must be in a given basis.

Let us consider the action of operator \hat{T} on a basis vectors $|\alpha_i\rangle$. Since the resulting object is another vector, it can be expanded in the same basis. Then,

$$\begin{aligned} \hat{T}|\alpha_i\rangle &= \sum_{j=1}^N T_{ji} |\alpha_j\rangle \\ &= T_{1i} |\alpha_1\rangle + T_{2i} |\alpha_2\rangle + \dots + T_{Ni} |\alpha_N\rangle \end{aligned} \quad (4.14)$$

This defines a set of N^2 (complex) numbers T_{ij} , which we can visualize as elements of a matrix T . Let \hat{T} act on a vector $|\psi\rangle$ to give a vector $|\phi\rangle$

$$\hat{T}|\psi\rangle = |\phi\rangle \quad (4.15)$$

Using the basis expansion of $|\psi\rangle$ and the linearity of \hat{T} , we can write (4.15) as

$$|\phi\rangle = \sum_{i=1}^N c_i \hat{T}|\alpha_i\rangle \quad (4.16)$$

Substituting for $\hat{T}|\alpha_i\rangle$ from (4.15) and reordering summations

$$|\phi\rangle = \sum_{j=1}^N \left(\sum_{i=1}^N T_{ji} c_i \right) |\alpha_j\rangle \quad (4.17)$$

We can see that if we expand $|\phi\rangle$ in the $|\alpha_j\rangle$ basis (an in (4.10)), the expansion coefficient b_j will be

$$b_j = \sum_{i=1}^N T_{ji} c_i \quad (4.18)$$

which is the same as the matrix equation

$$\phi = T \psi \quad (4.19)$$

where ϕ and ψ are column vector representations of $|\phi\rangle$ and $|\psi\rangle$ respectively. Then, the action of a linear operator \hat{T} on a ket vector $|\psi\rangle$ is the same as the action of the corresponding matrix T (with matrix elements defined in (4.15)) on the column vector representation of the vector. It clear that the matrix representation depends on the choice of basis. Since the basis we are using is orthonormal, there is a very simple way to relate the matrix elements to the linear operator. Taking the inner product of (4.15) with a basis vector $|\alpha_i\rangle$, and using orthonormality of the basis, we get

$$\langle \alpha_i | \hat{T} | \alpha_j \rangle = T_{ij} \quad (4.20)$$

where $\langle \alpha_i | \hat{T} | \alpha_j \rangle = \langle \alpha_i | (\hat{T} | \alpha_j \rangle)$. That is, the matrix element T_{ij} is the inner product of $|\alpha_i\rangle$ with vector $\hat{T} | \alpha_j \rangle$.

The product of two linear operators is defined as successive action of these operators on a vector. If \hat{A} and \hat{B} are linear operators, their product $\hat{A}\hat{B}$ is a linear operator such that given any vector $|\psi\rangle \in \mathcal{H}$

$$\hat{A}\hat{B} |\psi\rangle = \hat{A}(\hat{B} |\psi\rangle) \quad (4.21)$$

It is easy to check that the product $\hat{A}\hat{B}$ is represented by a matrix which is just the product of matrix representations of \hat{A} and \hat{B} .

The order in which operators are multiplied is important. In general, $\hat{A}\hat{B} \neq \hat{B}\hat{A}$. An easy way of seeing this is to realize that given matrices A and B , in general $AB \neq BA$. The *Commutator* of operators \hat{A} and \hat{B} is defined as

$$[\hat{A}, \hat{B}] = \hat{A}\hat{B} - \hat{B}\hat{A} \quad (4.22)$$

If the commutator of two operators is not zero ², it implies that the order in which they are multiplied matters.

We have already defined functions of (square) matrices, and we can similarly define functions of operators. First, we define $\hat{A}^n = \hat{A}\hat{A}..\hat{A}$ as successive operation of \hat{A} on a vector n times. Let $f(x)$ be a function of a real variable x which can be expressed as a series

$$f(x) = a_0 + a_1x + a_2x^2 + \dots + a_nx^n + \dots \quad (4.23)$$

Then, $f(\hat{A})$ is defined to be

$$f(\hat{A}) = a_0\hat{I} + a_1\hat{A} + a_2\hat{A}^2 + \dots + a_n\hat{A}^n + \dots \quad (4.24)$$

Since $f(\hat{A})$ is itself a linear operator, it must have a matrix representation. Since the matrix representation of a product of two operators is the product of their matrix representations, it is easy to see that the matrix representation of \hat{A}^n is just $A^n = AA...A$, matrix A multiplied n times. It easily follows that the matrix representation of $f(\hat{A})$ is just $f(A)$, where $f(A)$ is the matrix analogue of (4.24).

The Hermitian adjoint of an operator \hat{A} , denoted as \hat{A}^\dagger , is defined as follows: let $|\psi\rangle$ and $|\phi\rangle$ be any two vectors in \mathcal{H} . Then

$$\langle \psi | \hat{A}^\dagger | \phi \rangle = \langle \phi | \hat{A} | \psi \rangle^* \quad (4.25)$$

For the special case when $|\psi\rangle$ and $|\phi\rangle$ are basis vectors, using (4.20), it follows that the matrix elements of the Hermitian adjoint are given by

$$\hat{A}_{ij}^\dagger = A_{ji}^* \quad (4.26)$$

which is how we defined the Hermitian adjoint of a matrix before, as the transpose of the matrix followed by complex conjugation. Let us prove a useful result

$$\begin{aligned} \text{if } |\alpha\rangle &= \hat{A} |\phi\rangle \\ \text{then } \langle \alpha | \psi \rangle &= \langle \phi | \hat{A}^\dagger | \psi \rangle \end{aligned} \quad (4.27)$$

²A zero operator acts on any vector to give a null vector

To prove this, we start with $\langle \psi | \alpha \rangle = \langle \psi | \hat{A} | \phi \rangle$. Taking the complex conjugate, we get

$$\begin{aligned} \langle \psi | \alpha \rangle^* &= \langle \psi | \hat{A} | \phi \rangle^* \\ \Rightarrow \langle \alpha | \psi \rangle &= \langle \phi | \hat{A}^\dagger | \psi \rangle \end{aligned}$$

where we have used (4.25). It is easy to prove that the Hermitian adjoint of a product of two operators is the product of their Hermitian adjoint, in the reverse order

$$\left(\hat{A} \hat{B} \right)^\dagger = \hat{B}^\dagger \hat{A}^\dagger \quad (4.28)$$

A *Hermitian operator* is one which is equal to its Hermitian adjoint

$$\hat{A}^\dagger = \hat{A} \quad \text{Hermitian Operator} \quad (4.29)$$

In our study of electron spin, we saw that classical spin components in Cartesian coordinates are represented by Hermitian matrices S_x , S_y and S_z . Similarly, energy is represented by the Hermitian matrix H . Further, we saw that eigenvalues of these matrices correspond to possible results of measurements of the physical quantities they represent. Also, if a certain eigenvalue is measured, the state collapses to the corresponding eigenvector. Given the correspondence between matrices and linear operators, it must be that physical ‘observables’ are represented by Hermitian linear operators. Let us analyze what is so special about Hermitian operators. A vector $|\psi\rangle$ is said to be an eigenvector of a linear operator \hat{A} with eigenvalue a , if

$$\hat{A} |\psi\rangle = a |\psi\rangle \quad (4.30)$$

Given a linear operator, the number of eigenvectors it has depends on the properties of the operator. For instance, *all* vectors are trivially eigenvectors of the identity operator \hat{I} with eigenvalue 1. An eigenvector of an operator is also an eigenvector of any function of the operator with eigenvalue equal to the function of the corresponding eigenvalue. Let $\hat{B} |b\rangle = b |b\rangle$. Then, by successive action of \hat{B} , it is easy to see that $\hat{B}^n |b\rangle = b^n |b\rangle$. Then, given a function $f(\hat{B})$ in the form of a series (4.24),

$$\begin{aligned} f(\hat{B}) |b\rangle &= \left(a_0 \hat{I} + a_1 \hat{B} + a_2 \hat{B}^2 + \dots + a_n \hat{B}^n + \dots \right) |b\rangle \\ &= a_0 \hat{I} |b\rangle + a_1 \hat{B} |b\rangle + a_2 \hat{B}^2 |b\rangle + \dots + a_n \hat{B}^n |b\rangle + \dots \\ &= (a_0 + a_1 b + a_2 b^2 + \dots + a_n b^n + \dots) |b\rangle \\ &= f(b) |b\rangle \end{aligned} \quad (4.31)$$

The eigenvectors and eigenvalues of Hermitian operators have special properties, because of which physical observables are represented by such operators. We state without proof these properties³

1. The eigenvalues of Hermitian operators are real.
2. The eigenvectors of Hermitian operators corresponding to distinct eigenvalues are orthogonal

If we wish to associate results of measurements of physical quantities by eigenvalues of operators they represent, it is essential these eigenvalues be real, since measurements give real numbers. The postulate that these observables are represented by Hermitian operators guarantees this. Property (2) ensures the if a measurement of a physical observable yields an experimental number, if this measurement is immediately followed by a second measurement (with no time-delay), we get the same result. The way it works out is as follows: Let us say we make an observation of an observable \hat{A} , and get the experimental number a_1 . Then, the state of the system on this measurement will collapse to $|a\rangle$, where $\hat{A} |a\rangle = a |a\rangle$. Now, if we immediately measure \hat{A} again, we cannot get a different number, say a_2 , since the probability amplitude that this happens will be $\langle a_2 | a_1 \rangle$, which is zero, if $a_2 \neq a_1$ (property 2). Together, properties (1) and (2) justify representation of physical quantities by Hermitian operators. There is, however, a subtlety. What

³The proofs can be found in standard texts on Linear Algebra

happens if there are two distinct eigenvectors of \hat{A} with the *same* eigenvalue? What can we say about the inner product of these states? For instance, in the Hydrogen molecular ion, the Hydrogen atom states $|E_1\rangle$ and $|E_2\rangle$ had the same eigenvalue E_0 , but were still assumed to be orthogonal. In general, eigenvectors of a Hermitian operator with the same eigenvalue need not be orthogonal. However, given a set (say m) of such eigenvectors $|a_1\rangle, |a_2\rangle, \dots, |a_m\rangle$ all with the same eigenvalue a , it is always possible to find a *new* set $|b_1\rangle, |b_2\rangle, \dots, |b_m\rangle$ where the vectors $|b_i\rangle$ are linear combinations of the vectors $|a_i\rangle$ ($i = 1, 2, \dots, m$), such that vectors $|b_i\rangle$ are eigenvectors of \hat{A} with eigenvalue a , and are orthogonal to each other. That is,

$$\begin{aligned}\hat{A}|b_i\rangle &= a|b_i\rangle \\ \langle b_i|b_j\rangle &= 0 \text{ if } i \neq j\end{aligned}\quad (4.32)$$

where

$$|b_i\rangle = \sum_{j=1}^m c_{ij} |a_j\rangle \quad (4.33)$$

Given these properties and the fact that a Hermitian operator \hat{A} is represented by an $N \times N$ Hermitian matrix, we can always find a set of N orthogonal eigenvectors of this operator, with perhaps some of them corresponding to the same eigenvalue. Normalizing each of them, we will end up with N orthonormal eigenvectors of \hat{A} , which will form a basis of \mathcal{H} . Then, in an N dimensional Hilbert space, we can always choose the eigenvectors of a Hermitian operator as an orthonormal basis.

We next define a *Unitary* operator. A Unitary operator U is such that its Hermitian adjoint is its inverse

$$\hat{U}^\dagger \hat{U} = \hat{U} \hat{U}^\dagger = \hat{I} \quad \text{Unitary Operator} \quad (4.34)$$

Unitary operators have a very special property. A Unitary transformation on a pair of vectors leaves their inner product unchanged

$$\begin{aligned}\text{if } |\psi'\rangle &= \hat{U}|\psi\rangle \\ \text{and } |\phi'\rangle &= \hat{U}|\phi\rangle \\ \text{then } \langle\phi'|\psi'\rangle &= \langle\phi|\psi\rangle\end{aligned}\quad (4.35)$$

The proof goes as follows:

$$\begin{aligned}\langle\phi'|\psi'\rangle &= \langle\phi|\hat{U}^\dagger|\psi'\rangle \quad \text{Using (4.27)} \\ &= \langle\phi|\hat{U}^\dagger\hat{U}|\psi\rangle \\ &= \langle\phi|\hat{I}|\psi\rangle \quad \text{Using (4.34)} \\ &= \langle\phi|\psi\rangle\end{aligned}$$

We saw that the time evolution of spin state of an electron is represented by the action of a Unitary matrix on the column vector representing the spin state. Let us generalise this idea to an N dimensional Hilbert space, using the abstract notation. Let the state at instant t_0 be expanded in an orthonormal basis $\{|\alpha_i\rangle\}$ with expansion coefficients $\{c_i(t_0)\}$

$$|\psi(t_0)\rangle = \sum_{i=1}^N c_i(t_0) |\alpha_i\rangle \quad (4.36)$$

Since it is a physical state, it should satisfy $\langle\psi(t_0)|\psi(t_0)\rangle = 1$. In terms of expansion coefficients, this translated into

$$\sum_{i=1}^N |c_i(t_0)|^2 = 1 \quad (4.37)$$

As the state changes with time, the requirement that it remain a physical state (with unit norm) implies that $\langle\psi(t)|\psi(t)\rangle = 1$ at all instants t . We have seen that Unitary transformations preserve inner products,

and therefore, preserve the norm. Then, time-evolution of a state should be represented by a Unitary transformation

$$|\psi(t)\rangle = \hat{U}(t, t_0) |\psi(t_0)\rangle \quad (4.38)$$

where $\hat{U}(t, t_0)$ is a unitary operator. The form of \hat{U} is determined by the same line of reasoning as for the Unitary matrix representing time-evolution. In fact, it is easy to convince oneself that given that this operator generates time-evolution, it should satisfy

$$\begin{aligned} \hat{U}^\dagger(t, t_0)\hat{U}(t, t_0) &= \hat{I} \\ \hat{U}(t_2, t_1)\hat{U}(t_1, t_0) &= \hat{U}(t_2, t_0) \quad t_2 > t_1 > t_0 \\ \hat{U}(t_0, t_0) &= \hat{I} \end{aligned} \quad (4.39)$$

Then, if we consider time-evolution from instant t to $t + dt$ where dt is infinitesimal, we can write (to order dt)

$$\hat{U}(t + dt, t) = \hat{I} - \frac{idt}{\hbar} \hat{H} \quad (4.40)$$

Using $\hat{U}^\dagger\hat{U} = \hat{I}$ (and retaining terms upto order dt), we can deduce that $\hat{H}^\dagger = \hat{H}$. That is, H should be a Hermitian operator. This operator will in general depend explicitly on time, if interactions are time-dependent. Then, the state at instant $t + dt$ will be

$$\begin{aligned} |\psi(t + dt)\rangle &= \hat{U}(t + dt, t) |\psi(t)\rangle \\ &\sim \left(\hat{I} - \frac{idt}{\hbar} \hat{H} \right) |\psi(t)\rangle \\ &= |\psi(t)\rangle - \frac{idt}{\hbar} \hat{H} |\psi(t)\rangle \end{aligned}$$

Rearranging, we get

$$i\hbar \frac{1}{dt} (|\psi(t + dt)\rangle - |\psi(t)\rangle) = \hat{H} |\psi(t)\rangle$$

We now take the limit $dt \rightarrow 0$. Note that the expression $(1/dt) (|\psi(t + dt)\rangle - |\psi(t)\rangle)$ is a vector, since it involves the difference of two vectors ($|\psi(t + dt)\rangle$ and $|\psi(t)\rangle$), followed by multiplication by a number (dt). We define the derivative of $|\psi(t)\rangle$ w.r.t t as a vector which is the limit of this expression, as $dt \rightarrow 0$. Finally, we get

$$i\hbar \frac{d}{dt} |\psi(t)\rangle = \hat{H} |\psi(t)\rangle \quad (4.41)$$

which is Schrodinger equation for the N dimensional quantum system, written in the ket notation. The Hermitian operator H represents energy, and is called the Hamiltonian operator for this system. As mentioned, \hat{H} could in general depend on time explicitly, as it (its matrix analogue) did in the case of electron spin interacting with a time-varying magnetic field. In the special case when the interactions of the system are time-independent, this operator will not explicitly depend on time. Then, we can formally integrate (4.41) to get

$$|\psi(t)\rangle = e^{-i\hat{H}(t-t_0)/\hbar} |\psi(t_0)\rangle \quad (4.42)$$

Comparing with (4.38), we see that

$$\hat{U}(t, t_0) = e^{-i\hat{H}(t-t_0)/\hbar} \quad (4.43)$$

To verify that (4.42) is the solution to (4.41), we must understand that the exponential function of \hat{H} is defined in the spirit of (4.24). Then, we expand it in a series in (4.42), to get

$$|\psi(t)\rangle = \left[\hat{I} + \left(-i \frac{t-t_0}{\hbar} \right) \hat{H} + \frac{1}{2!} \left(-i \frac{t-t_0}{\hbar} \right)^2 \hat{H}^2 + \dots + \frac{1}{n!} \left(-i \frac{t-t_0}{\hbar} \right)^n \hat{H}^n + \dots \right] |\psi(t_0)\rangle$$

When we differentiate this expression with respect to time (after multiplication with $i\hbar$, since operator \hat{H} is time-independent, the derivatives will only act on terms of the form $(t - t_0)^n$. Factoring out one power

of \hat{H} , we will again get the series for $e^{-i\hat{H}(t-t_0)/\hbar}$, acting on $|\psi(t_0)\rangle$, which is just $|\psi(t)\rangle$, eventually leading to (4.41).

Since the Hamiltonian operator is Hermitian, it is always possible to find a set of N orthonormal eigenvectors of \hat{H} , corresponding to distinct eigenvalues, or some eigenvalues being the same. Let this set be $\{|E_i\rangle\}$, $i = 1, 2, \dots, N$. Since these are eigenvectors of \hat{H} , they satisfy

$$\hat{H} |E_i\rangle = E_i |E_i\rangle \quad (4.44)$$

Eigenvectors of \hat{H} are states in which the system's energy is well-defined, equal to the corresponding eigenvalue. Let us expand a state $|\psi\rangle$ in the energy basis

$$|\psi\rangle = \sum_{i=1}^N c_i |E_i\rangle \quad (4.45)$$

If a measurement of energy is made on this state, one of the eigenvalues will be obtained. The probability that eigenvalue E_j is obtained is computed from the probability amplitude $\langle E_j | \psi \rangle$

$$\begin{aligned} \langle E_j | \psi \rangle &= \sum_{i=1}^N c_i \langle E_j | E_i \rangle \\ &= \sum_{i=1}^N c_i \delta_{ij} \\ &= c_j \end{aligned} \quad (4.46)$$

where orthonormality of the eigenstates has been used. The probability of measuring energy E_j is then $|c_j|^2$. Then, given state $|\psi\rangle$ expanded in the energy basis, the probability of measuring a certain value of energy is equal to the modulus squared of the corresponding expansion coefficient. If energy E_i is measured, the state will collapse to $|E_i\rangle$. An interesting situation arises if there is a *degeneracy*. We say there is a 'degeneracy' in energy if corresponding to a given eigenvalue, there is more than one eigenstate. Let us say states $|E_1\rangle$, $|E_2\rangle$ and $|E_3\rangle$ are degenerate, with the same energy, say E . If we measure the energy of the system in the state $|\psi\rangle$ in (4.45) and obtain E , what will be the state after this measurement? Will it be $|E_1\rangle$, $|E_2\rangle$ or $|E_3\rangle$? The answer is as follows, to be taken as a *postulate*. If E is measured, *project out* the part of the state that is not a linear superposition of $|E_1\rangle$, $|E_2\rangle$ and $|E_3\rangle$. What remains is the linear combination $c_1 |E_1\rangle + c_2 |E_2\rangle + c_3 |E_3\rangle$. The state will collapse to *this* linear combination, after normalization has been carried out. The normalized vector proportional to this linear combination is

$$|\phi\rangle = \frac{1}{\sqrt{|c_1|^2 + |c_2|^2 + |c_3|^2}} (c_1 |E_1\rangle + c_2 |E_2\rangle + c_3 |E_3\rangle)$$

States of definite energy are special. If we start with a state of definite energy, it stays that way. To see this, let us start with state $|E_i\rangle$ at $t = 0$. Then, the state at instant t will be given by (4.42)

$$|\psi(t)\rangle = e^{-i\hat{H}t/\hbar} |E_i\rangle \quad (4.47)$$

Since $\hat{H} |E_i\rangle$, it follows (from 4.31) that

$$|\psi(t)\rangle = e^{-iE_i t/\hbar} |E_i\rangle \quad (4.48)$$

which is the same physical state, since multiplying a state vector by a phase does not change the result of any measurement on that vector. To see this, let $|\phi\rangle = e^{i\delta} |\psi\rangle$. Then, the probability that state $|\phi\rangle$ collapses to some state $|\alpha\rangle$ as a result of measurement is $|\langle \alpha | \phi \rangle|^2$. But, since $\langle \alpha | \phi \rangle = e^{i\delta} \langle \alpha | \psi \rangle$, therefore $|\langle \alpha | \phi \rangle|^2 = |\langle \alpha | \psi \rangle|^2$. Since this is true for any state $|\alpha\rangle$, it follows that statistical outcome of any measurement on states $|\psi\rangle$ and $e^{i\delta} |\psi\rangle$ will be the same. They are, therefore, physically the same state.

The fact that states of definite energy do not change with time is the quantum analogue of conservation of energy. Another consequence of this is that given an arbitrary state changing with time, the probability of measuring some energy eigenvalue is independent of time. To see this, let us calculate $\langle E_i | \psi(t) \rangle$. Using (4.38), it follows that

$$\langle E_i | \psi(t) \rangle = \langle E_i | e^{-i\hat{H}t/\hbar} | \psi(0) \rangle$$

It is easier to calculate the complex conjugate of this quantity. Using property of Hermitian adjoint (4.25), it follows that

$$\begin{aligned} \langle E_i | \psi(t) \rangle^* &= \langle E_i | e^{-i\hat{H}t/\hbar} | \psi(0) \rangle^* \\ &= \langle \psi(0) | e^{+i\hat{H}t/\hbar} | E_i \rangle \\ &= e^{iE_i t/\hbar} \langle \psi(0) | E_i \rangle \end{aligned}$$

where we have used the fact that $(e^{-i\hat{H}t/\hbar})^\dagger = e^{+i\hat{H}t/\hbar}$, since $\hat{H}^\dagger = \hat{H}$. Then, the probability that we measure energy E_i at instant t is

$$\begin{aligned} P(E_i, t) &= |\langle E_i | \psi(t) \rangle|^2 \\ &= |\langle E_i | \psi(t) \rangle^*|^2 \\ &= |\langle E_i | \psi(0) \rangle|^2 \\ &= P(E_i, 0) \end{aligned} \tag{4.49}$$

which is the same as the probability at $t = 0$. Then, this probability is independent of time. It must be understood that this is true only for energy measurements. If some other observable is measured, the probabilities of different outcomes (eigenvalues of the observable) will in general change with time.

The special properties of energy states makes it useful to choose them as basis states. Expanded in the energy basis, the time-evolution of a state (reflected as change in expansion coefficients with time) is simple to analyze. Say, we expand a state at $t = 0$ as a linear superposition of energy eigenstates

$$|\psi(0)\rangle = \sum_{i=0}^N c_i(0) |E_i\rangle \tag{4.50}$$

The state at instant t will be (from 4.42)

$$\begin{aligned} |\psi(t)\rangle &= \sum_{i=0}^N c_i(0) e^{-i\hat{H}t/\hbar} |E_i\rangle \\ &= \sum_{i=0}^N c_i(0) e^{-iE_i t/\hbar} |E_i\rangle \end{aligned} \tag{4.51}$$

since $e^{-i\hat{H}t/\hbar} |E_i\rangle = e^{-iE_i t/\hbar} |E_i\rangle$. This tells us that in the energy basis, the expansion coefficients will vary with time as

$$c_i(t) = c_i(0) e^{-iE_i t/\hbar} \tag{4.52}$$

From this, we can once again deduce the fact that if we measure the energy of the system, the probability of measuring energy E_i (equal to $|c_i(t)|^2$) is independent of time.

4.2 Expectation Value and Uncertainty

We have seen that quantum mechanics gives probabilistic predictions. To test quantum theory, therefore, we need to repeatedly perform identical experiments, and analyze statistical outcomes. For instance, say we are given that the system is in some quantum state $|\psi\rangle$, and we wish to predict the result of the

measurement of some physical observable, say \hat{A} , on this state. To do this, we expand the state as a linear superposition of orthonormal eigenvectors of this observable

$$|\psi\rangle = \sum_{i=1}^N c_i |\alpha_i\rangle \quad (4.53)$$

where $\hat{A}|\alpha_i\rangle = a_i|\alpha_i\rangle$. Then, the probability that the result of the measurement will be $A = a_i$ will be

$$\begin{aligned} P(a_i) &= |\langle\alpha_i|\psi\rangle|^2 \\ &= |c_i|^2 \end{aligned} \quad (4.54)$$

How do we test this hypothesis? A single measurement performed on state $|\psi\rangle$ will not suffice, since the result will be one of the eigenvalues, and the state of the system after the measurement will collapse to the corresponding eigenvector. Further measurements on this collapsed state will not give us any information about the initial state $|\psi\rangle$. Then, we need to prepare identical copies of the initial state, and perform repeated measurements of \hat{A} . Eventually, we will end up with a table telling us the number of times a given eigenvalue was measured, and this data can be analyzed to check whether the corresponding numbers agree with the predictions of the different probabilities. Since this analysis is statistical, it is useful to talk about standard statistical tools such as average and standard deviation, and connect it to the predictions of quantum mechanics. For instance, say we wish to predict what the average of the result of successive measurements of \hat{A} will be, the measurements performed on (identical copies of) state $|\psi\rangle$. Quantum mechanics tells us that we will measure only the eigenvalues of \hat{A} , with probabilities given by (4.54). Then, the average of the results of these measurements will be predicted to be

$$\langle\hat{A}\rangle = \sum_{i=1}^N P(a_i) \times a_i \quad (4.55)$$

where average of A is denoted as $\langle\hat{A}\rangle$. Given that $P(a_i)$ is given by (4.54) and the fact that $\hat{A}|\alpha_i\rangle = a_i|\alpha_i\rangle$, we can manipulate (4.55) to get a compact form

$$\begin{aligned} \langle\hat{A}\rangle &= \sum_{i=1}^N |\langle\alpha_i|\psi\rangle|^2 a_i \\ &= \sum_{i=1}^N \langle\alpha_i|\psi\rangle \langle\alpha_i|\psi\rangle^* a_i \\ &= \sum_{i=1}^N \langle\alpha_i|\psi\rangle \langle\psi|\alpha_i\rangle a_i \\ &= \sum_{i=1}^N \langle\alpha_i|\psi\rangle \langle\psi|\hat{A}|\alpha_i\rangle \end{aligned} \quad (4.56)$$

where we have used the fact that $\langle\psi|\hat{A}|\alpha_i\rangle = \langle\psi|\alpha_i\rangle a_i$. A little more manipulation gives

$$\begin{aligned} \langle\hat{A}\rangle &= \sum_{i=1}^N \langle\psi|\hat{A}|\alpha_i\rangle \langle\alpha_i|\psi\rangle \\ &= \langle\psi|\hat{A}\left(\sum_{i=1}^N |\alpha_i\rangle \langle\alpha_i|\psi\rangle\right) \end{aligned} \quad (4.57)$$

where linearity of \hat{A} has been used. The expression in parenthesis is just the expansion for $|\psi\rangle$. Finally, we get

$$\langle \hat{A} \rangle = \langle \psi | \hat{A} | \psi \rangle \quad (4.58)$$

The average of a physical quantity in quantum mechanics is referred to as its *Expectation Value*.

Let us find a similar, compact expression for the standard deviation in the measurement of A . The standard deviation is defined as usual, and denoted as $\Delta\hat{A}$. The standard deviation squared is, as usual, the average of deviation (squared) from the mean (expectation value)

$$\Delta\hat{A}^2 = \sum_{i=1}^N P(a_i) \times (a_i - \langle \hat{A} \rangle)^2 \quad (4.59)$$

This can be simplified as follows

$$\begin{aligned} \Delta\hat{A}^2 &= \sum_{i=1}^N P(a_i) \left(a_i^2 + \langle \hat{A} \rangle^2 - 2a_i \langle \hat{A} \rangle \right) \\ &= \sum_{i=1}^N P(a_i) a_i^2 + \langle \hat{A} \rangle^2 \sum_{i=1}^N P(a_i) - 2 \langle \hat{A} \rangle \sum_{i=1}^N P(a_i) a_i \\ &= \sum_{i=1}^N P(a_i) a_i^2 - \langle \hat{A} \rangle^2 \end{aligned} \quad (4.60)$$

where we have used $\sum_{i=1}^N P(a_i) = 1$ and $\sum_{i=1}^N P(a_i) a_i = \langle \hat{A} \rangle$. Since a_i^2 is the eigenvalue of operator \hat{A}^2 (corresponding to eigenstate $|\alpha_i\rangle$), it follows that $\sum_{i=1}^N P(a_i) a_i^2 = \langle \hat{A}^2 \rangle$. Substituting this in (4.60), we finally get a compact expression for standard deviation squared

$$\begin{aligned} \Delta\hat{A}^2 &= \langle \hat{A}^2 \rangle - \langle \hat{A} \rangle^2 \\ &= \langle \psi | \hat{A}^2 | \psi \rangle - \langle \psi | \hat{A} | \psi \rangle^2 \end{aligned} \quad (4.61)$$

The standard deviation $\Delta\hat{A} = \sqrt{\Delta\hat{A}^2}$ in quantum mechanics is referred to as *Uncertainty*.

Let us apply the formalism that we have developed to quantum mechanics of electron spin. It is a two-dimensional vector space, with physical observables being the three components of spin, represented by operators \hat{S}_x , \hat{S}_y and \hat{S}_z . Experiments show that these observables have two eigenvectors each, with eigenvalues $\pm\hbar/2$. We denote orthonormal eigenvectors of \hat{S}_z as $|+z\rangle$ and $|-z\rangle$, such that

$$\hat{S}_z |\pm z\rangle = \pm \frac{\hbar}{2} |\pm z\rangle \quad (4.62)$$

From (4.20), the matrix elements of \hat{S}_z will be

$$\begin{aligned} \langle +z | \hat{S}_z | +z \rangle &= \frac{\hbar}{2} \\ \langle +z | \hat{S}_z | -z \rangle &= 0 \\ \langle -z | \hat{S}_z | +z \rangle &= 0 \\ \langle -z | \hat{S}_z | -z \rangle &= -\frac{\hbar}{2} \end{aligned} \quad (4.63)$$

Then, the matrix representation of \hat{S}_z in this basis will be, as we have seen before,

$$S_z = \begin{pmatrix} \hbar/2 & 0 \\ 0 & -\hbar/2 \end{pmatrix} \quad (4.64)$$

What about the matrix elements of the other two spin operators? How does one find them from first principles? We will answer this question in a while. For now, let us assume that the matrix representations of \hat{S}_x and \hat{S}_y in the $|\pm_z\rangle$ basis are given to us, the same as when we first ‘discovered’ the spin matrices

$$\begin{aligned} S_x &= \begin{pmatrix} 0 & \hbar/2 \\ \hbar/2 & 0 \end{pmatrix} \\ S_y &= \begin{pmatrix} 0 & -i\hbar/2 \\ i\hbar/2 & 0 \end{pmatrix} \end{aligned} \quad (4.65)$$

Let us use this information to determine the action of \hat{S}_x and \hat{S}_y on $|\pm_z\rangle$. Let us operate on $|+z\rangle$ and $|-z\rangle$ with \hat{S}_x . Since $\hat{S}_x|+z\rangle$ and $\hat{S}_x|-z\rangle$ are vectors, we expand them in the $|\pm_z\rangle$ basis

$$\begin{aligned} \hat{S}_x|+z\rangle &= |+z\rangle\langle +z|\hat{S}_x|+z\rangle + |-z\rangle\langle -z|\hat{S}_x|+z\rangle \\ \hat{S}_x|-z\rangle &= |+z\rangle\langle +z|\hat{S}_x|-z\rangle + |-z\rangle\langle -z|\hat{S}_x|-z\rangle \end{aligned} \quad (4.66)$$

where $\langle +z|\hat{S}_x|+z\rangle = \langle +z|(\hat{S}_x|+z\rangle)$, etc. Since we are given the matrix representation of \hat{S}_x , we can read off the elements

$$\begin{aligned} \langle +z|\hat{S}_x|+z\rangle &= 0 \\ \langle +z|\hat{S}_x|-z\rangle &= \frac{\hbar}{2} \\ \langle -z|\hat{S}_x|+z\rangle &= \frac{\hbar}{2} \\ \langle -z|\hat{S}_x|-z\rangle &= 0 \end{aligned} \quad (4.67)$$

Then, the action of \hat{S}_x on the $|\pm_z\rangle$ basis is

$$\begin{aligned} \hat{S}_x|+z\rangle &= \frac{\hbar}{2}|-z\rangle \\ \hat{S}_x|-z\rangle &= \frac{\hbar}{2}|+z\rangle \end{aligned} \quad (4.68)$$

Similarly, the action of \hat{S}_y on the $|\pm_z\rangle$ basis is found to be

$$\begin{aligned} \hat{S}_y|+z\rangle &= \frac{i\hbar}{2}|-z\rangle \\ \hat{S}_y|-z\rangle &= \frac{-i\hbar}{2}|+z\rangle \end{aligned} \quad (4.69)$$

We have seen before that matrices S_x , S_y and S_z satisfy the following algebra

$$[S_i, S_j] = i\hbar \sum_{k=1}^3 \epsilon_{ijk} S_k \quad (4.70)$$

where ϵ_{ijk} is the *Levi Cevita* symbol, which is completely antisymmetric in its indices (it changes sign on permutation of any two indices) and $\epsilon_{123} = 1$ (which automatically fixes the other ‘components’). In particular, it is zero when any two (or more) indices are the same. Since every operator equation has a matrix analogue and vice-versa, it follows that the spin operators themselves should satisfy these relations

$$[\hat{S}_i, \hat{S}_j] = i\hbar \sum_{k=1}^3 \epsilon_{ijk} \hat{S}_k \quad (4.71)$$

Explicitly, these reduce to

$$\begin{aligned} [\hat{S}_x, \hat{S}_y] &= i\hbar \hat{S}_z \\ [\hat{S}_y, \hat{S}_z] &= i\hbar \hat{S}_x \\ [\hat{S}_z, \hat{S}_x] &= i\hbar \hat{S}_y \end{aligned} \quad (4.72)$$

Let us now calculate the expectation value of \hat{S}_x in the $|+_z\rangle$ state. This can be computed using (4.58) and (4.67) and is seen to be zero. Similarly, the expectation value of \hat{S}_x in the $|-_z\rangle$ state is zero. The same is true for the expectation values of \hat{S}_y in the $|+_z\rangle$ and $|-_z\rangle$ states (as can be verified). Is this what we expect? In the $|\pm_z\rangle$ states, if \hat{S}_x or \hat{S}_y are measured, we will get $+\hbar/2$ and $-\hbar/2$ with equal probability. Then, the average must be zero, which agrees with our results for expectation values. Now, let us calculate the uncertainty in say \hat{S}_x in state $|+_z\rangle$, using (4.61). To do this, we need to calculate $\langle+_z|\hat{S}_x^2|+_z\rangle$. But this is just the norm of the state $\hat{S}_x|+_z\rangle$. If \hat{A} is a Hermitian operator, then the expectation value of its square, calculated in some state will be

$$\begin{aligned} \langle \hat{A}^2 \rangle &= \langle \psi | \hat{A}^2 | \psi \rangle \\ &= \langle \psi | \hat{A} \hat{A} | \psi \rangle \\ &= \left(\langle \psi | \hat{A}^\dagger \right) \left(\hat{A} | \psi \rangle \right) \\ &= \langle \hat{A} \psi | \hat{A} \psi \rangle \end{aligned} \quad (4.73)$$

where $|\hat{A}\psi\rangle$ is shorthand for $\hat{A}|\psi\rangle$. Since $\hat{S}_x|+_z\rangle = \frac{\hbar}{2}|-_z\rangle$, therefore

$$\begin{aligned} \langle+_z|\hat{S}_x^2|+_z\rangle &= \left(\frac{\hbar}{2}\right)^2 \langle-_z| -_z\rangle \\ &= \left(\frac{\hbar}{2}\right)^2 \end{aligned}$$

Then, in state $|+_z\rangle$, the uncertainty (squared) in \hat{S}_x is

$$\begin{aligned} \Delta \hat{S}_x^2 &= \langle+_z|\hat{S}_x^2|+_z\rangle - \langle+_z|\hat{S}_x|+_z\rangle^2 \\ &= \left(\frac{\hbar}{2}\right)^2 \end{aligned} \quad (4.74)$$

which gives $\Delta \hat{S}_x = \hbar/2$.

4.3 Compatible and Incompatible Observables: The Uncertainty Principle

We have seen that if measurement of an observable is made on a state, we obtain one of the eigenvalues of the associated operator and the state collapses to the corresponding eigenvector. If the same observable is measured in quick succession (with in principle zero time delay), the same eigenvalue will be measured and the state will remain unchanged. However, if some other observable is measured, an eigenvalue of that observable will be obtained, and the state will collapse to the corresponding eigenvector. What happens if after measuring the second observable, we measure the first one again?

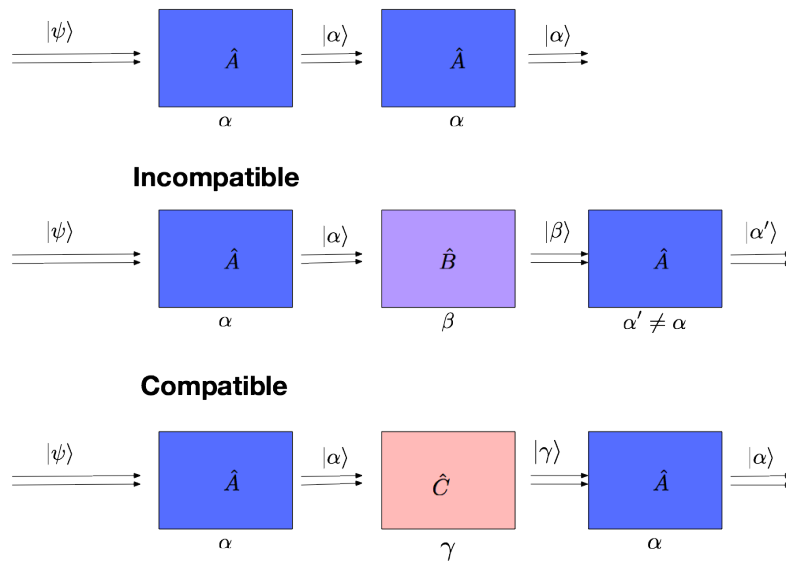


Figure 4.2: Observables \hat{A} and \hat{B} are incompatible, whereas \hat{A} and \hat{C} are compatible

Let us say we make a measurement of observable \hat{A} on state $|\psi\rangle$, and obtain eigenvalue α . Then, the state will collapse to $|\alpha\rangle$. In immediate succession, we measure another observable \hat{B} , measure eigenvalue β , and conclude that the state is now $|\beta\rangle$. What if we measure \hat{A} now (again, without delay)? Our experience with electron spin tells us that in general, we will *not* measure α again, but a different eigenvalue, say α' . What does this imply? Does it imply that a measurement of \hat{B} (with result of measurement equal to β) has converted the state from $|\alpha\rangle$ to $|\alpha'\rangle$, so that we obtain α' on measurement of \hat{A} again? No, since if we were to take a beam of particles, all in state $|\beta\rangle$ and passed it through an \hat{A} measuring apparatus, the beam will in general split into several components, each corresponding to a different eigenvalue of \hat{A} . This implies that a measurement of \hat{B} on the state $|\alpha\rangle$, which resulted in an eigenvector $|\beta\rangle$ of \hat{B} , made observable \hat{A} *uncertain*. That is, state $|\beta\rangle$ is not an eigenstate of \hat{A} , but a linear superposition of eigenstates of \hat{A} . In other words, a measurement of \hat{B} does not preserve the result of a measurement of \hat{A} . Such operators are said to be *Incompatible*. Note that in classical physics, *all* physical quantities are compatible, in the sense that they are all simultaneously well-defined. However, there also exist observables which are such that a measurement of one preserves the result of measurement of the other. Such observables are said to be *Compatible*. We shall now determine the necessary and sufficient condition for two observables to be compatible. Say observables \hat{A} and \hat{C} are compatible. Say we measure \hat{A} and get α , the state collapsing to the eigenvector $|\alpha\rangle$. Next, we measure \hat{C} and get γ , the resulting state $|\gamma\rangle$ being an eigenvector of \hat{C} . For compatibility, we require that if we now measure \hat{A} again, we should once again obtain α . One way this is possible is that the states $|\alpha\rangle$ and $|\gamma\rangle$ are the *same* state. That is, the state $|\alpha\rangle$ obtained after a measurement of \hat{A} is an eigenvector of \hat{C} as well, with perhaps a different eigenvalue γ . Such a *simultaneous*

eigenstate of \hat{A} and \hat{C} (with eigenvalues α and γ) is written as $|\alpha, \gamma\rangle$ and satisfies

$$\begin{aligned}\hat{A}|\alpha, \gamma\rangle &= \alpha|\alpha, \gamma\rangle \\ \hat{C}|\alpha, \gamma\rangle &= \gamma|\alpha, \gamma\rangle\end{aligned}\tag{4.75}$$

Then, if we measure \hat{A} and the result is α , the state collapses to $|\alpha, \gamma\rangle$, on which a measurement of \hat{C} is guaranteed to give γ , without changing the state. However, if the result of measurement of \hat{A} was some other eigenvalue, the incompatibility might still be there. The only way the two observables would be completely compatible would be for *all* eigenvectors of one to be eigenvectors of the other as well. Then, since a measurement of one of them is guaranteed to result in a state which is also an eigenstate of the other one, successive measurements will leave the state and the eigenvalues intact. We now need to find a mathematical condition to be satisfied by two observables to be compatible. The following theorem gives this condition: “*The necessary and sufficient condition for two Hermitian operators to have a complete set⁴ of simultaneous eigenvectors is that their commutator be equal to zero*”. That is, operators \hat{A} and \hat{B} will have all their eigenvectors common if and only if $[\hat{A}, \hat{B}] = 0$. The statement and its implications are a little non-trivial, so we will analyze it in a little detail. Let us assume first that all eigenvectors of \hat{A} are eigenvectors of \hat{B} as well. Then, given any (simultaneous) eigenvector $|\alpha_i, \beta_i\rangle$, such that $\hat{A}|\alpha_i, \beta_i\rangle = \alpha_i|\alpha_i, \beta_i\rangle$ and $\hat{B}|\alpha_i, \beta_i\rangle = \beta_i|\alpha_i, \beta_i\rangle$,

$$\begin{aligned}[\hat{A}, \hat{B}]|\alpha_i, \beta_i\rangle &= (\hat{A}\hat{B} - \hat{B}\hat{A})|\alpha_i, \beta_i\rangle \\ &= \hat{A}\hat{B}|\alpha_i, \beta_i\rangle - \hat{B}\hat{A}|\alpha_i, \beta_i\rangle \\ &= \beta_i\hat{A}|\alpha_i, \beta_i\rangle - \alpha_i\hat{B}|\alpha_i, \beta_i\rangle \\ &= \beta_i\alpha_i|\alpha_i, \beta_i\rangle - \alpha_i\beta_i|\alpha_i, \beta_i\rangle \\ &= |0\rangle\end{aligned}\tag{4.76}$$

where $|0\rangle$ is the null (or zero) vector. Then, the action of $[\hat{A}, \hat{B}]$ on any eigenvector gives a null vector. But since the set of all these eigenvectors forms a basis, it follows that the action of $[\hat{A}, \hat{B}]$ on *any* vector in \mathcal{H} will give a null vector. This implies that $[\hat{A}, \hat{B}] = 0$ (null operator). Next, we prove the converse. Assume that $[\hat{A}, \hat{B}] = 0$, so that $\hat{A}\hat{B} = \hat{B}\hat{A}$. For now, assume that all the eigenvalues of one of the operators, say \hat{A} , are distinct. Let $|\alpha\rangle$ be an eigenvector of \hat{A} with eigenvalue α . Then,

$$\begin{aligned}\hat{A}(\hat{B}|\alpha\rangle) &= \hat{A}\hat{B}|\alpha\rangle \\ &= \hat{B}\hat{A}|\alpha\rangle \\ &= \alpha\hat{B}|\alpha\rangle \\ &= \alpha(\hat{B}|\alpha\rangle)\end{aligned}\tag{4.77}$$

which shows that if $|\alpha\rangle$ is an eigenvector of \hat{A} with eigenvalue α , so is $\hat{B}|\alpha\rangle$, with the *same* eigenvalue. But, we assumed that all eigenvalues of \hat{A} were distinct. Then, the state $\hat{B}|\alpha\rangle$ must be proportional to $|\alpha\rangle$

$$\hat{B}|\alpha\rangle = \beta|\alpha\rangle\tag{4.78}$$

where β is some proportionality constant. But, this is the same as saying that $|\alpha\rangle$ is an eigenvector of \hat{B} with eigenvalue β . Therefore, if the commutator of two operators is zero and one of them has a ‘non-degenerate spectrum’⁵, all the eigenvectors of the operator with non-degenerate eigenvalues are eigenvectors of the other one, so the two operators have a complete set of simultaneous eigenstates. What if one or both

⁴By ‘complete set’, we mean the set of all N eigenvectors

⁵The spectrum of an operator is the set of its eigenvalues

have a degenerate spectrum? This is when the theorem gets a little tricky. The key idea is as follows: Let some eigenvectors of \hat{A} have the same eigenvalue. Then, in general, they need not be eigenvectors of \hat{B} . However, these degenerate eigenvectors form a subspace of \mathcal{H} , that is, any linear combination of these eigenvectors is an eigenvector of \hat{A} with the same, degenerate eigenvalue (as can be checked easily). These eigenvectors are linearly independent⁶, and so form a basis in this subspace. From this basis, it is always possible to form new linear combinations which *are* eigenvectors of \hat{B} ⁷. Then, even if one of the operators has a degenerate spectrum, it is always possible to find *at least one* complete set of eigenvectors which are also eigenvectors of the other operator. There is another subtlety that is worth talking about. For illustration, let two simultaneous eigenvectors of commuting operators \hat{A} and \hat{B} have the same eigenvalue α for \hat{A} , but distinct eigenvalues β_1 and β_2 for \hat{B} . We denote these as $|\alpha, \beta_1\rangle$ and $|\alpha, \beta_2\rangle$ respectively. Let us take a state $|\psi\rangle = c_1 |\alpha, \beta_1\rangle + c_2 |\alpha, \beta_2\rangle$ and make successive measurements of \hat{A} and \hat{B} in different orders. First, say we measure \hat{B} . The measurement will give one of the eigenvalues β_1 or β_2 . Say, we measure β_1 . Then, the state will collapse to $|\alpha, \beta_1\rangle$. Next, if we measure \hat{A} , we are bound to get α as a result of the measurement, and the state will remain unchanged. If we now measure \hat{B} once again, we will again get β_1 with unit probability. Then, as expected, the measurements of \hat{A} and \hat{B} are compatible. However, it should be noticed that after the first measurement, the state changed from $|\psi\rangle$ to $|\alpha, \beta_1\rangle$. Now, we again start with state $|\psi\rangle$, and measure \hat{A} first. We note that $|\psi\rangle = c_1 |\alpha, \beta_1\rangle + c_2 |\alpha, \beta_2\rangle$ is an eigenvector of \hat{A} with eigenvalue α (since both $|\alpha, \beta_1\rangle$ and $|\alpha, \beta_2\rangle$ are eigenvectors of \hat{A} with the same eigenvalue, and $|\psi\rangle$ lies in the degenerate subspace). Then, we will measure α , and *the state will remain unchanged*. This is unlike the first case, when we first measured \hat{B} , and since $|\psi\rangle$ is not an eigenstate of \hat{B} , it collapsed to one of the states $|\alpha, \beta_1\rangle$ or $|\alpha, \beta_2\rangle$ (we assumed it was $|\alpha, \beta_1\rangle$). Now, if we measure \hat{B} , we will either obtain β_1 or β_2 , and the state will collapse to one of the corresponding eigenvectors. Subsequently, if \hat{A} is measured, we will still get α , so the measurements are still compatible. Then, the two observables are compatible, in whatever order they are measured. However, the behavior of the state in the two cases is quite different.

We now deduce an extremely important inequality, which is central to quantum physics. Before deducing it, we need a few mathematical results. The first is that the expectation value of a Hermitian operator is always real. This is easy to see, since the eigenvalues of a Hermitian operator are real, and so it follows from (4.55) that the expectation value must also be real. Next, we define an operator to be *anti-Hermitian*, if it is minus of its Hermitian adjoint

$$\hat{A}^\dagger = -\hat{A} \quad \text{Anti-Hermitian Operator} \quad (4.79)$$

Let us prove that the eigenvalues of an anti-Hermitian operator are purely imaginary. Let $|\alpha\rangle$ be an eigenvector of \hat{A} with eigenvalue α

$$\hat{A}|\alpha\rangle = \alpha|\alpha\rangle \quad (4.80)$$

Taking the inner product of this equation with $|\alpha\rangle$, we get

$$\langle\alpha|\hat{A}|\alpha\rangle = \alpha\langle\alpha|\alpha\rangle \quad (4.81)$$

Taking the complex conjugate of this equation, using the definition of the Hermitian adjoint (4.25) and using the fact that \hat{A} is anti-Hermitian, we get

$$\begin{aligned} \langle\alpha|\hat{A}|\alpha\rangle^* &= \alpha^*\langle\alpha|\alpha\rangle^* \\ \Rightarrow \langle\alpha|\hat{A}^\dagger|\alpha\rangle &= \alpha^*\langle\alpha|\alpha\rangle \\ \Rightarrow \langle\alpha|\hat{A}|\alpha\rangle &= -\alpha^*\langle\alpha|\alpha\rangle \end{aligned} \quad (4.82)$$

From (4.81), it is clear that $\alpha^* = -\alpha$. That is, α is purely imaginary. We will also need the so-called *Schwarz Inequality*

$$|\langle\phi|\psi\rangle|^2 \leq \langle\phi|\phi\rangle\langle\psi|\psi\rangle \quad (4.83)$$

⁶Eigenvectors of a Hermitian operator are linearly independent, even corresponding to the same eigenvalue: see any textbook on Linear Algebra

⁷See any text on Linear Algebra for details of the proof

This is easily proved as follows: Define a vector $|\alpha\rangle = |\psi\rangle - |\phi\rangle \langle\phi|\psi\rangle / \langle\phi|\phi\rangle$. Using the fact that $\langle\alpha|\alpha\rangle \geq 0$ gives the inequality. Now, let \hat{A} and \hat{B} be two Hermitian operators. Let their expectation values in some state $|\psi\rangle$ be $\langle\hat{A}\rangle$ and $\langle\hat{B}\rangle$ respectively. Define operators $\hat{A}' = \hat{A} - \langle\hat{A}\rangle \hat{I}$ and $\hat{B}' = \hat{B} - \langle\hat{B}\rangle \hat{I}$. It is easy to see that \hat{A}' and \hat{B}' are also Hermitian. Next, define vectors $|\alpha\rangle$ and $|\beta\rangle$ as

$$\begin{aligned} |\alpha\rangle &= \hat{A}' |\psi\rangle \\ |\beta\rangle &= \hat{B}' |\psi\rangle \end{aligned} \quad (4.84)$$

where \hat{I} is the identity operator. We intend to apply the Schwarz Inequality to these vectors. It can be checked easily that $\langle\alpha|\alpha\rangle = \Delta\hat{A}^2$ and $\langle\beta|\beta\rangle = \Delta\hat{B}^2$. Now, let us calculate $\langle\alpha|\beta\rangle$. This will be

$$\begin{aligned} \langle\alpha|\beta\rangle &= \langle\psi|\hat{A}'\hat{B}'|\psi\rangle \\ &= \frac{1}{2} \langle\psi|(\hat{A}'\hat{B}' - \hat{B}'\hat{A}')|\psi\rangle + \frac{1}{2} \langle\psi|(\hat{A}'\hat{B}' + \hat{B}'\hat{A}')|\psi\rangle \\ &= \frac{1}{2} \langle\psi|[\hat{A}', \hat{B}']|\psi\rangle + \frac{1}{2} \langle\psi|[\hat{A}', \hat{B}']_+|\psi\rangle \\ &= \frac{1}{2} \langle\psi|[\hat{A}, \hat{B}]|\psi\rangle + \frac{1}{2} \langle\psi|[\hat{A}', \hat{B}']_+|\psi\rangle \end{aligned} \quad (4.85)$$

where $[\hat{A}', \hat{B}']_+ = \hat{A}'\hat{B}' + \hat{B}'\hat{A}'$ is called the *Anti-Commutator* of \hat{A}' and \hat{B}' . Also, the fact that $[\hat{A}', \hat{B}'] = [\hat{A}, \hat{B}]$ has been used (check it!). It should be observed that the commutator of two Hermitian operators is anti-Hermitian, whereas the anti-commutator is Hermitian (easily checked). Further, the expectation value of a Hermitian operator is real, and that of an anti-Hermitian operator is imaginary (since its eigenvalues are imaginary). Then, it follows that $\langle\psi|[\hat{A}, \hat{B}]|\psi\rangle$ is imaginary and $\langle\psi|[\hat{A}', \hat{B}']_+|\psi\rangle$ is real. Then,

$$\begin{aligned} |\langle\alpha|\beta\rangle|^2 &= \frac{1}{4} \left| \langle\psi|[\hat{A}, \hat{B}]|\psi\rangle \right|^2 + \frac{1}{4} \left| \langle\psi|[\hat{A}, \hat{B}]_+|\psi\rangle \right|^2 \\ &\geq \frac{1}{4} \left| \langle\psi|[\hat{A}, \hat{B}]|\psi\rangle \right|^2 \end{aligned} \quad (4.86)$$

Finally, using Schwarz Inequality (4.83), we get

$$\Delta\hat{A}^2 \Delta\hat{B}^2 \geq \frac{1}{4} \left| \langle\psi|[\hat{A}, \hat{B}]|\psi\rangle \right|^2 \quad (4.87)$$

Equation (4.87) is the mathematical statement of the *Generalized Uncertainty Principle*. The well-known position-momentum (Heisenberg) uncertainty principle is a consequence of this general principle. Given two Hermitian, non-commuting operators \hat{A} and \hat{B} , this principle puts a lower bound on the product of uncertainties in \hat{A} and \hat{B} . Let us assume that these operators are such that there is no state in which the expectation value of $[\hat{A}, \hat{B}]$ is zero. Then, it follows that the product $\Delta\hat{A}\Delta\hat{B}$ can never be zero, for any state. This implies that if we try to reduce, say $\Delta\hat{A}$ (by suitably choosing a state), it will increase $\Delta\hat{B}$ and vice-versa. It is not possible for uncertainties in the two observables to be arbitrarily small, because of this constraint. In particular, $\Delta\hat{A}$ and $\Delta\hat{B}$ will never be both zero. This is understandable, since if it were true for some state $|\psi\rangle$, it would have to be an eigenstate of both \hat{A} and \hat{B} , which is not possible, since their commutator is non-zero. Let us apply this principle to spin operators \hat{S}_x , \hat{S}_y and \hat{S}_z . The commutators of these operators with each other are non-zero. In particular, $[\hat{S}_x, \hat{S}_y] = i\hbar\hat{S}_z$. Then, given a state $|\psi\rangle$, it follows that

$$\Delta\hat{S}_x \Delta\hat{S}_y \geq \frac{\hbar}{2} \langle\psi|\hat{S}_z|\psi\rangle \quad (4.88)$$

Is it possible to arbitrarily minimize $\Delta\hat{S}_x$ or $\Delta\hat{S}_y$ in some state? It is, provided the expectation value of \hat{S}_z in the state is zero. We know that $\langle\hat{S}_z\rangle$ is zero in eigenstates of \hat{S}_x or \hat{S}_y . In any one of these eigenstates, either $\Delta\hat{S}_x$ or $\Delta\hat{S}_y$ will be zero (since the state is an eigenstate of one of the two operators). However, given a state in which $\langle\hat{S}_z\rangle$ is not zero, it is *not* possible to have either $\Delta\hat{S}_x$ or $\Delta\hat{S}_y$ zero. Then, if a state is such that the average of the component of spin along some direction is non-zero in the state, the spin component in any perpendicular direction cannot be precise. There will always be some uncertainty.

4.4 Ehrenfest's Theorem

We now derive a very important theorem, which in a sense establishes a connection between classical and quantum physics. It also gives us an operational way of quantizing an arbitrary physical system, whose classical behavior is known. The basic idea is as follows: unlike classical physics, quantum physics is fundamentally probabilistic. Given a set of identically prepared states $|\psi\rangle$, if a measurement of a physical observable \hat{A} is made, we will get statistically fluctuating results (eigenvalues of \hat{A} with different probabilities). In statistical terms, we will observe certain average behaviour (expectation value) with fluctuations about this average (uncertainty). These statistical parameters will depend on the state on which these observations are carried out. Then, if the state is changing with time, the average and fluctuations about the average will also change with time. Ehrenfest's Theorem is an equation describing how the average of a physical quantity varies with time, as a result of the state changing with time.

The state of a quantum system changes with time according to Schrodinger equation (4.41)

$$i\hbar \frac{d}{dt} |\psi(t)\rangle = \hat{H} |\psi(t)\rangle \quad (4.89)$$

As the state evolves with time, the expectation value of a physical observable \hat{A} will change with time

$$\langle \hat{A}(t) \rangle = \langle \psi(t) | \hat{A} | \psi(t) \rangle \quad (4.90)$$

We wish to deduce the rate at which this changes. That is, we wish to compute $\frac{d}{dt} \langle \hat{A}(t) \rangle$

$$\begin{aligned} \frac{d}{dt} \langle \hat{A}(t) \rangle &= \frac{d}{dt} \langle \psi(t) | \hat{A} | \psi(t) \rangle \\ &= \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \left\{ \langle \psi(t + \Delta t) | \hat{A} | \psi(t + \Delta t) \rangle - \langle \psi(t) | \hat{A} | \psi(t) \rangle \right\} \end{aligned} \quad (4.91)$$

It follows from Schrodinger equation that for infinitesimal Δt ,

$$\begin{aligned} i\hbar \frac{1}{\Delta t} (|\psi(t + \Delta t)\rangle - |\psi(t)\rangle) &= \hat{H} |\psi(t)\rangle \\ \Rightarrow |\psi(t + \Delta t)\rangle &= \left(\hat{I} - \frac{i\Delta t}{\hbar} \hat{H} \right) |\psi(t)\rangle \end{aligned} \quad (4.92)$$

Using this, let us calculate $\langle \psi(t + \Delta t) | \hat{A} | \psi(t + \Delta t) \rangle$ in the limit of infinitesimal Δt

$$\begin{aligned} \langle \psi(t + \Delta t) | \hat{A} | \psi(t + \Delta t) \rangle &= \langle \psi(t) | \left(\hat{I} + \frac{i\Delta t}{\hbar} \hat{H} \right) \hat{A} \left(\hat{I} - \frac{i\Delta t}{\hbar} \hat{H} \right) | \psi(t) \rangle \\ &= \langle \psi(t) | \left(\hat{A} + \frac{i\Delta t}{\hbar} (\hat{H} \hat{A} - \hat{A} \hat{H}) \right) | \psi(t) \rangle \\ &= \langle \psi(t) | \hat{A} | \psi(t) \rangle + \frac{i\Delta t}{\hbar} \langle \psi(t) | [\hat{H}, \hat{A}] | \psi(t) \rangle \end{aligned} \quad (4.93)$$

where $[\hat{H}, \hat{A}]$ is the commutator of \hat{H} and \hat{A} . Substituting this in (4.91), we finally get

$$\frac{d}{dt} \langle \hat{A}(t) \rangle = \frac{i}{\hbar} \langle \psi(t) | [\hat{H}, \hat{A}] | \psi(t) \rangle \quad (4.94)$$

Equation (4.94) is known as *Ehrenfest's Theorem*, and has deep implications. To explore these, let us apply this to spin space. We assume that there the spin is interacting with a static magnetic field along the z-direction, $\mathbf{B} = B\hat{\mathbf{z}}$. The interaction Hamiltonian will be the operator analogue of (2.4)

$$\begin{aligned} \hat{H} &= \frac{e}{m_e} \hat{\mathbf{S}} \cdot \mathbf{B} \\ &= \frac{e}{m_e} B \hat{S}_z \end{aligned} \quad (4.95)$$

where $\hat{\mathbf{S}} = (\hat{S}_x, \hat{S}_y, \hat{S}_z)$. We now apply Ehrenfest's Theorem to see how expectation values of \hat{S}_x , \hat{S}_y and \hat{S}_z , in some (time-varying) state $|\psi(t)\rangle$, change with time. From (4.94), it is clear that we need to calculate the commutators of these operators with the Hamiltonian operator, $\hat{H} = (e/m_e)B\hat{S}_z$. We will need the expressions for the commutators of the spin operators with each other, which we can get from (4.72). These commutators are as follows

$$\begin{aligned}
 [\hat{H}, \hat{S}_z] &= \frac{eB}{m_e} [\hat{S}_z, \hat{S}_z] \\
 &= 0 \\
 [\hat{H}, \hat{S}_x] &= \frac{eB}{m_e} [\hat{S}_z, \hat{S}_x] \\
 &= i\frac{eB\hbar}{m_e} \hat{S}_y \\
 [\hat{H}, \hat{S}_y] &= \frac{eB}{m_e} [\hat{S}_z, \hat{S}_y] \\
 &= -i\frac{eB\hbar}{m_e} \hat{S}_x
 \end{aligned} \tag{4.96}$$

Substituting for each commutator in (4.94), we get

$$\begin{aligned}
 \frac{d}{dt} \langle \hat{S}_z(t) \rangle &= 0 \\
 \frac{d}{dt} \langle \hat{S}_x(t) \rangle &= -\left(\frac{eB}{m_e}\right) \langle \hat{S}_y(t) \rangle \\
 \frac{d}{dt} \langle \hat{S}_y(t) \rangle &= \left(\frac{eB}{m_e}\right) \langle \hat{S}_x(t) \rangle
 \end{aligned}$$

These are the *same* as the classical equations for spin-precession, except that instead of classical spin components S_x, S_y, S_z , we have expectation values of their quantum counterparts. This seems to suggest the following connection between classical and quantum physics: The expectation values of Hermitian operators representing classical observables change with time the same way as the classical observables themselves.

If we take this as a general principle, does it have any predictive power? To see this, let us consider the following experiment. A beam of certain particle species is made to pass through a Stern-Gerlach filter, and is seen to split into n beams. It then follows that the particles are charged and have some spin angular momentum, which is such that along any direction, its components will have n possible values, corresponding to spin states $|\alpha_1\rangle, |\alpha_2\rangle \dots |\alpha_n\rangle$

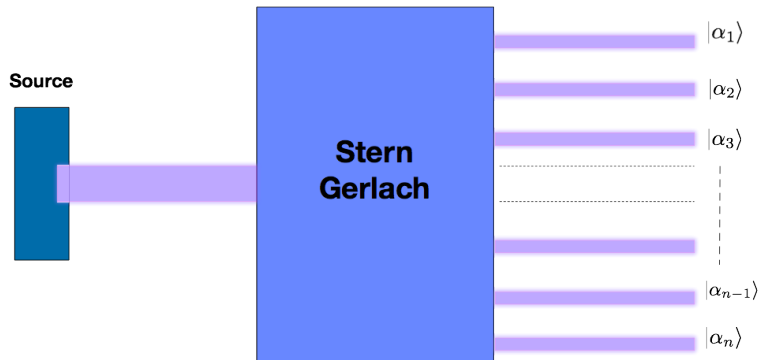


Figure 4.3: Stern Gerlach experiment with higher spin particles

We wish to describe the quantum mechanics of this (higher) spin system. Just as in the case of electron spin, there will be some magnetic moment associated with this spin, which interacts with the magnetic field. If we treat this spin angular momentum (denoted as \mathbf{J}) as classical, and let it interact with a magnetic field along the z-direction, its equations of motion will be

$$\begin{aligned}\frac{d}{dt}J_z(t) &= 0 \\ \frac{d}{dt}J_x(t) &= -\omega J_y(t) \\ \frac{d}{dt}J_y(t) &= \omega J_x(t)\end{aligned}$$

which are just the precession equations, with ω being the precession (angular) frequency. It should be noted that we are *not* saying that this is the spin of some fundamental particle (such as electron), but could be the total spin angular momentum of a complex object (an atom, say) originating out some combination of constituent particle angular momenta. Whatever the origin of this angular momentum, as long as it produces a magnetic moment, it *will* satisfy these equations. The spin components will be represented by Hermitian operators \hat{J}_x , \hat{J}_y and \hat{J}_z . The interaction of this system with the magnetic field will be described by the Hamiltonian operator \hat{H} , which we believe will have the *same* form as the classical energy, with physical quantities represented by their operator counterparts. This Hamiltonian will then have the form

$$\hat{H} = \lambda \mathbf{B} \cdot \hat{\mathbf{J}} \quad (4.97)$$

where $\hat{\mathbf{J}} = (\hat{J}_x, \hat{J}_y, \hat{J}_z)$ and λ will depend on some parameters of the system. The state of the system will be a vector in an n dimensional vector space, spanned by states $|\alpha_1\rangle, |\alpha_2\rangle \dots |\alpha_n\rangle$. If the system is prepared in some state $|\psi\rangle$, the state will change with time according to Schrodinger Equation (4.41). The expectation values of the three spin operators will change with time according to (4.94). Given that the magnetic field is along the z-direction, this will give us the following equations

$$\begin{aligned}\frac{d}{dt}\langle \hat{J}_z(t) \rangle &= i\frac{\lambda B}{\hbar} \langle \psi(t) | [\hat{J}_z, \hat{J}_z] | \psi(t) \rangle \\ \frac{d}{dt}\langle \hat{J}_x(t) \rangle &= i\frac{\lambda B}{\hbar} \langle \psi(t) | [\hat{J}_z, \hat{J}_x] | \psi(t) \rangle \\ \frac{d}{dt}\langle \hat{J}_y(t) \rangle &= i\frac{\lambda B}{\hbar} \langle \psi(t) | [\hat{J}_z, \hat{J}_y] | \psi(t) \rangle\end{aligned} \quad (4.98)$$

Now, we assume the correspondence between classical and quantum physics, implied by Ehrenfest Theorem, as a general principle. Then, comparing the above equations with the corresponding classical equations, and demanding that $\langle \hat{J}_x(t) \rangle$, $\langle \hat{J}_y(t) \rangle$ and $\langle \hat{J}_z(t) \rangle$ satisfy the same equations as the classical variables $J_x(t)$, $J_y(t)$ and $J_z(t)$, it is easy to see that operators \hat{J}_x , \hat{J}_y and \hat{J}_z must satisfy the *same* commutator algebra as \hat{S}_x , \hat{S}_y and \hat{S}_z (electron spin operators)

$$\begin{aligned} [\hat{J}_x, \hat{J}_y] &= i\hbar \hat{J}_z \\ [\hat{J}_y, \hat{J}_z] &= i\hbar \hat{J}_x \\ [\hat{J}_z, \hat{J}_x] &= i\hbar \hat{J}_y \end{aligned} \quad (4.99)$$

which can be written in the compact form

$$[\hat{J}_i, \hat{J}_j] = i\hbar \sum_{k=1}^3 \epsilon_{ijk} \hat{J}_k \quad (4.100)$$

This algebra should be satisfied by *any* kind of angular momentum, whatever its origin.

4.5 Quantum Mechanics of Angular Momentum: Higher Spin spaces

We now analyze the consequences of the ‘angular momentum algebra’ implied by (4.100). First, the fact that the commutators of angular momentum operators \hat{J}_x , \hat{J}_y and \hat{J}_z are non-zero tells us that there cannot exist simultaneous eigenstates of these operators. That is, there do not exist states in which more than one component of angular momentum is well-defined, whatever the origin of angular momentum. Now, let us construct the quantum analogue of the magnitude (squared) of angular momentum, just as we did in case of electron spin. This will be represented by an operator \hat{J}^2 , given by

$$\hat{J}^2 = \hat{J}_x^2 + \hat{J}_y^2 + \hat{J}_z^2 \quad (4.101)$$

Let us calculate the commutator of \hat{J}^2 and \hat{J}_z . We use the following property of commutators (easy to verify)

$$[\hat{A}, \hat{B}\hat{C}] = \hat{B}[\hat{A}, \hat{C}] + [\hat{A}, \hat{B}]\hat{C} \quad (4.102)$$

Then,

$$\begin{aligned} [\hat{J}_z, \hat{J}^2] &= [\hat{J}_z, \hat{J}_x^2] + [\hat{J}_z, \hat{J}_y^2] \\ &= \hat{J}_x[\hat{J}_z, \hat{J}_x] + [\hat{J}_z, \hat{J}_x]\hat{J}_x + \hat{J}_y[\hat{J}_z, \hat{J}_y] + [\hat{J}_z, \hat{J}_y]\hat{J}_y \\ &= i\hbar \hat{J}_x\hat{J}_y + i\hbar \hat{J}_y\hat{J}_x - i\hbar \hat{J}_y\hat{J}_x - i\hbar \hat{J}_x\hat{J}_y \\ &= 0 \end{aligned} \quad (4.103)$$

Since the commutator of \hat{J}^2 and \hat{J}_z is zero, we can construct a complete set of simultaneous eigenstates of these operators. It is easily checked that the commutators $[\hat{J}_x, \hat{J}^2]$ and $[\hat{J}_y, \hat{J}^2]$ are also zero. Then, we can also construct simultaneous eigenstates of \hat{J}^2 and \hat{J}_x or \hat{J}^2 and \hat{J}_y . However, since the commutator of any one component of $\hat{\mathbf{J}}$ with another component is non-zero, we can have simultaneous eigenstates of \hat{J}^2 and any *one* of the three components of $\hat{\mathbf{J}}$. In general, we can say that it is possible to construct simultaneous eigenstates of \hat{J}^2 and components of $\hat{\mathbf{J}}$ along any one direction in space. Traditionally, this direction is labelled z. Let us denote the simultaneous eigenstates of \hat{J}^2 and \hat{J}_z as $|a, b\rangle$ where a is the eigenvalue of \hat{J}^2 and b is the eigenvalue of \hat{J}_z such that

$$\begin{aligned} \hat{J}^2 |a, b\rangle &= a |a, b\rangle \\ \hat{J}_z |a, b\rangle &= b |a, b\rangle \end{aligned} \quad (4.104)$$

We will now determine the possible values of a and b , and the properties of the corresponding eigenvectors. First, we define two operators \hat{J}_+ and \hat{J}_- , called *Ladder Operators*

$$\hat{J}_\pm = \hat{J}_x \pm i\hat{J}_y \quad (4.105)$$

Because of the presence of i in (4.105), the operators \hat{J}_+ and \hat{J}_- are not Hermitian. In fact,

$$\begin{aligned} \hat{J}_+^\dagger &= \hat{J}_- \\ \hat{J}_-^\dagger &= \hat{J}_+ \end{aligned} \quad (4.106)$$

It can be checked that the ladder operators have the following properties

$$\begin{aligned} [\hat{J}_+, \hat{J}_-] &= 2\hbar\hat{J}_z \\ [\hat{J}_z, \hat{J}_\pm] &= \pm\hbar\hat{J}_\pm \end{aligned} \quad (4.107)$$

Also, since the commutator of \hat{J}^2 with \hat{J}_i is zero, it follows trivially that

$$[\hat{J}^2, \hat{J}_\pm] = 0 \quad (4.108)$$

Let us analyze the result of the operation of \hat{J}_+ and \hat{J}_- on states $|a, b\rangle$. First, let us act by \hat{J}_+ on $|a, b\rangle$ to get the vector $\hat{J}_+|a, b\rangle$. Next, we observe the action of \hat{J}_z on this state

$$\begin{aligned} \hat{J}_z(\hat{J}_+|a, b\rangle) &= (\hat{J}_z\hat{J}_+)|a, b\rangle \\ &= (\hat{J}_z\hat{J}_+ - \hat{J}_+\hat{J}_z + \hat{J}_+\hat{J}_z)|a, b\rangle \\ &= ([\hat{J}_z, \hat{J}_+] + \hat{J}_+\hat{J}_z)|a, b\rangle \\ &= (\hbar\hat{J}_+ + \hat{J}_+\hat{J}_z)|a, b\rangle \\ &= (b + \hbar)\hat{J}_+|a, b\rangle \end{aligned} \quad (4.109)$$

where we have used (4.104) and (4.107). It is evident from (4.109) that while $|a, b\rangle$ is an eigenvector of \hat{J}_z with eigenvalue b , $\hat{J}_+|a, b\rangle$ is an eigenvector of \hat{J}_z with eigenvalue $(b + \hbar)$. Then, the action of \hat{J}_+ on a state of well-defined \hat{J}_z raises the value of \hat{J}_z by \hbar . For this reason, \hat{J}_+ is called a *raising* operator. A similar calculation shows that the action of \hat{J}_- on $|a, b\rangle$ gives an eigenstate of \hat{J}_z with a lowered eigenvalue $b - \hbar$. That is,

$$\hat{J}_z(\hat{J}_-|a, b\rangle) = (b - \hbar)\hat{J}_-|a, b\rangle \quad (4.110)$$

For this reason, \hat{J}_- is called a *lowering* operator. We now check what the action of these ladder operators is on the eigenvalue of \hat{J}^2

$$\begin{aligned} \hat{J}^2(\hat{J}_\pm|a, b\rangle) &= (\hat{J}^2\hat{J}_\pm)|a, b\rangle \\ &= ([\hat{J}^2, \hat{J}_\pm] + \hat{J}_\pm\hat{J}^2)|a, b\rangle \\ &= (0 + \hat{J}_\pm\hat{J}^2)|a, b\rangle \\ &= a\hat{J}_\pm|a, b\rangle \end{aligned} \quad (4.111)$$

where we have used (4.108). Then, $\hat{J}_\pm|a, b\rangle$ is an eigenvector of \hat{J}^2 , with the same eigenvalue, a . Then, the action of \hat{J}_\pm leaves the value of \hat{J}^2 unchanged. Starting with a state $|a, b\rangle$, if we successively act by \hat{J}_+ , we will produce states which are all simultaneous eigenvectors of \hat{J}^2 and \hat{J}_z , with the same eigenvalue of \hat{J}^2 , but increasing eigenvalues of \hat{J}_z . Every action of \hat{J}_+ will increase this eigenvalue by \hbar . Similarly, successive action of \hat{J}_- will produce states which are all simultaneous eigenvectors of \hat{J}^2 and \hat{J}_z , again with the same

eigenvalue of \hat{J}^2 , but decreasing eigenvalues of \hat{J}_z . This process can be visualized as a *ladder* (hence the term *ladder operators*)

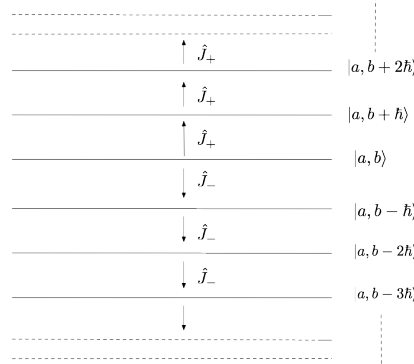


Figure 4.4: ‘Ladder’ of Angular Momentum states

Intuition tells us that starting with a state $|a, b\rangle$, we cannot indefinitely ‘climb’ such a ladder, either ‘up’ or ‘down’. Any step along this ladder leaves the magnitude of angular momentum unchanged (in the sense of unchanged eigenvalue of \hat{J}^2 while increasing or decreasing the z-component of angular momentum. Classically, the magnitude of a component of a vector can never be greater than the magnitude of the vector itself. One would expect something similar in quantum mechanics. That is, it should not be possible to construct states with a given magnitude of angular momentum, but arbitrarily large values of \hat{J}_z . The quantum analogue of this classical constraint can be derived as follows. We first rewrite (4.5) as

$$\hat{J}^2 - \hat{J}_z^2 = \hat{J}_x^2 + \hat{J}_y^2$$

Next, we take the expectation value of both sides with respect to the state $|a, b\rangle$

$$\langle a, b | \hat{J}^2 | a, b \rangle - \langle a, b | \hat{J}_z^2 | a, b \rangle = \langle a, b | \hat{J}_x^2 | a, b \rangle + \langle a, b | \hat{J}_y^2 | a, b \rangle \quad (4.112)$$

Next, let us write $\hat{J}_x |a, b\rangle = |\psi\rangle$ and $\hat{J}_y |a, b\rangle = |\phi\rangle$. Then, since \hat{J}_x and \hat{J}_y are Hermitian,

$$\begin{aligned} \langle a, b | \hat{J}_x^2 | a, b \rangle &= \langle \psi | \psi \rangle \\ &\geq 0 \\ \langle a, b | \hat{J}_y^2 | a, b \rangle &= \langle \phi | \phi \rangle \\ &\geq 0 \end{aligned} \quad (4.113)$$

Since the norm of a vector is greater than or equal to zero. Then, it follows that

$$\langle a, b | \hat{J}^2 | a, b \rangle - \langle a, b | \hat{J}_z^2 | a, b \rangle \geq 0 \quad (4.114)$$

But, $\hat{J}^2 |a, b\rangle = a |a, b\rangle$ and $\hat{J}_z^2 |a, b\rangle = b^2 |a, b\rangle$. Therefore, we get the following constraint on the eigenvalues a (of \hat{J}^2) and b (of \hat{J}_z)

$$a \geq b^2 \quad (4.115)$$

which just says that the z-component of angular momentum has to be less than or equal to the magnitude of angular momentum. Since this constraint needs to be satisfied, there must be an upper limit to which we can raise the eigenvalue b by successive action of \hat{J}_+ . Let this eigenvalue be b_{\max} and the corresponding state $|a, b_{\max}\rangle$. The only way action of \hat{J}_+ on this state does not give a state of higher b is that

$$\hat{J}_+ |a, b_{\max}\rangle = 0 \quad (4.116)$$

(where ‘0’ henceforth stands for the null vector $|\rangle_0$). Operating with \hat{J}_- , we get on this state does not give a state of higher b is that

$$\hat{J}_- \hat{J}_+ |a, b_{\max}\rangle = 0 \quad (4.117)$$

Now, we use the following result (easy to verify)

$$\begin{aligned} \hat{J}_- \hat{J}_+ &= \hat{J}_x^2 + \hat{J}_y^2 - i(\hat{J}_y \hat{J}_x - \hat{J}_x \hat{J}_y) \\ &= \hat{J}^2 - \hat{J}_z^2 - \hbar \hat{J}_z \end{aligned} \quad (4.118)$$

where we have used $[\hat{J}_x, \hat{J}_y] = i\hbar \hat{J}_z$. Substituting this in (4.117), we get

$$\begin{aligned} (\hat{J}^2 - \hat{J}_z^2 - \hbar \hat{J}_z) |a, b\rangle &= 0 \\ \Rightarrow (a - b_{\max}^2 - b_{\max} \hbar) |a, b\rangle &= 0 \end{aligned} \quad (4.119)$$

Since $|a, b\rangle \neq 0$, we get

$$a - b_{\max}^2 - b_{\max} \hbar = 0 \quad (4.120)$$

which gives

$$a = b_{\max} (b_{\max} + \hbar) \quad (4.121)$$

A similar argument tells us that there should exist a state $|a, b_{\min}\rangle$, such that

$$\hat{J}_- |a, b_{\min}\rangle = 0 \quad (4.122)$$

Operating with \hat{J}_+ , we get

$$\hat{J}_+ \hat{J}_- |a, b_{\min}\rangle = 0 \quad (4.123)$$

We can write $\hat{J}_+ \hat{J}_-$ as (verify!)

$$\hat{J}_+ \hat{J}_- = \hat{J}^2 - \hat{J}_z^2 + \hbar \hat{J}_z \quad (4.124)$$

Proceeding as before, we are eventually led to

$$a = b_{\min} (b_{\min} - \hbar) \quad (4.125)$$

Comparing (4.121) and (4.125), we conclude that

$$b_{\max} = -b_{\min} \quad (4.126)$$

Then, b lies in the range

$$-b_{\max} \leq b \leq b_{\max} \quad (4.127)$$

It is clear that starting with $|a, b_{\min}\rangle$, we should be able to reach $|a, b_{\max}\rangle$ after an integral number of operations with \hat{J}_+ . Since every action increases b by \hbar , it follows that

$$b_{\max} = b_{\min} + n\hbar \quad (4.128)$$

for some integer n . We then get

$$b_{\max} = \frac{n\hbar}{2} \quad (4.129)$$

so that from (4.121), it follows that a has to be of the form

$$a = \frac{n\hbar}{2} \left(\frac{n\hbar}{2} + \hbar \right) \quad (4.130)$$

Let us rewrite (4.127) in terms of n

$$-\frac{n\hbar}{2} \leq b \leq \frac{n\hbar}{2} \quad (4.131)$$

Equations (4.130) and (4.131) can be interpreted as follows: In a state in which the magnitude and component of angular momentum along some direction (here, z) are well-defined simultaneously, these values are characterized by an integer $n = 0, 1, 2, \dots$. For a given n , the magnitude can take a set of *discrete* values $|\mathbf{J}| = \hbar\sqrt{(n/2)(n/2 + 1)}$, and the component can take a *range* of discrete values between $-(n/2)\hbar$ and $(n/2)\hbar$ in steps of \hbar . It is convenient to define $j = (n/2)$, such that $j = 0, 1/2, 1, 3/2, \dots$. In terms of j , $|\mathbf{J}| = \hbar\sqrt{j(j+1)}$ and $-j\hbar \leq J_z \leq j\hbar$, in steps of \hbar . Since there is a one to one correspondence between eigenvalue of \hat{J}^2 and j , we rewrite the eigenvectors $|a, b\rangle$ as $|j, m_j\rangle$, such that

$$\begin{aligned}\hat{J}^2 |j, m_j\rangle &= j(j+1)\hbar^2 |j, m_j\rangle \\ \hat{J}_z |j, m_j\rangle &= m_j\hbar |j, m_j\rangle\end{aligned}\quad (4.132)$$

where

$$\begin{aligned}j &= 0, \frac{1}{2}, 1, \frac{3}{2}, 2, \frac{5}{2}, \dots \\ m_j &= -j, -j+1, -j+2, \dots, j-1, j\end{aligned}\quad (4.133)$$

For a given magnitude of angular momentum characterized by j , its component can take $2j+1$ values, give by (4.133). These are then the possible states of definite angular momentum (magnitude and one component), and they form an infinite set. We have just encountered an infinite dimensional Hilbert Space. The states $|j, m_j\rangle$ are assumed to be normalized. Further, since they are eigenvectors of (two) Hermitian operators, two such states $|j, m_j\rangle$ and $|j', m'_j\rangle$ will be orthogonal if *either* of the two eigenvalues are different. This can be expressed as

$$\langle j', m'_j | j, m_j \rangle = \delta_{j'j} \delta_{m'_j m_j} \quad (4.134)$$

We now determine the action of operators \hat{J}_x and \hat{J}_y on these states. First, we determine the action of the ladder operators \hat{J}_+ and \hat{J}_- . We know that $\hat{J}_+ |a, b\rangle$ is an eigenstate of \hat{J}^2 and \hat{J}_z with eigenvalue a unchanged and b changed to $b + \hbar$. In the $|j, m_j\rangle$ notation, this is equivalent to saying that

$$\hat{J}_+ |j, m_j\rangle = c_+ |j, m_j + 1\rangle \quad (4.135)$$

where c_+ is a complex number, to be determined. We take the norm of this equation, and remembering that $\hat{J}_+^\dagger = \hat{J}_-$, we get

$$\langle j, m_j | \hat{J}_- \hat{J}_+ |j, m_j\rangle = |c_+|^2 \quad (4.136)$$

Using (4.118), this gives

$$\begin{aligned}\langle j, m_j | \left(\hat{J}^2 - \hat{J}_z^2 - \hbar \hat{J}_z \right) |j, m_j\rangle &= |c_+|^2 \\ \Rightarrow (j(j+1) - m_j^2 - m_j) \hbar^2 &= |c_+|^2\end{aligned}\quad (4.137)$$

This determines c_+ (we assume it is real, dropping an overall phase)

$$c_+ = \sqrt{(j - m_j)(j + m_j + 1)} \hbar \quad (4.138)$$

Similar analysis shows that the action of \hat{J}_- on $|j, m_j\rangle$ should be

$$\hat{J}_- |j, m_j\rangle = c_- |j, m_j - 1\rangle \quad (4.139)$$

where c_- is determined to be

$$c_- = \sqrt{(j + m_j)(j - m_j + 1)} \hbar \quad (4.140)$$

We can summarize the action of \hat{J}_\pm as follows

$$\hat{J}_\pm |j, m_j\rangle = \sqrt{(j \mp m_j)(j \pm m_j + 1)} \hbar |j, m_j \pm 1\rangle \quad (4.141)$$

The action of \hat{J}_x and \hat{J}_y can now be determined, using (4.105)

$$\begin{aligned}\hat{J}_x |j, m_j\rangle &= \frac{1}{2} (\hat{J}_+ + \hat{J}_-) |j, m_j\rangle \\ &= \frac{\hbar}{2} \sqrt{(j - m_j)(j + m_j + 1)} |j, m_j + 1\rangle + \frac{\hbar}{2} \sqrt{(j + m_j)(j - m_j + 1)} |j, m_j - 1\rangle\end{aligned}\quad (4.142)$$

Similarly, the action of \hat{J}_y is

$$\begin{aligned}\hat{J}_y |j, m_j\rangle &= \frac{1}{2i} (\hat{J}_+ - \hat{J}_-) |j, m_j\rangle \\ &= \frac{\hbar}{2i} \sqrt{(j - m_j)(j + m_j + 1)} |j, m_j + 1\rangle - \frac{\hbar}{2i} \sqrt{(j + m_j)(j - m_j + 1)} |j, m_j - 1\rangle\end{aligned}\quad (4.143)$$

Chapter 5

Quantum Mechanics of Point Particles

5.1 Propagation of an electron in a lattice

An interesting example of quantum mechanics in an N dimensional vector space involves the propagation of an electron in a lattice of atoms. It is an experimental fact that low energy electrons propagate ‘freely’ in a regular crystal lattice. If an electron had a classical trajectory, it would have a mean free path of the order of just a few angstroms in the lattice (it would encounter a ‘rigid’ lattice atom every few angstroms) and would encounter a large resistance due to scattering off lattice atoms. The fact that an electron propagates according to the laws of quantum mechanics allows it to ‘move’ in such a lattice with little resistance. As an example, we will consider a one-dimensional lattice. We will in fact (for mathematical reasons) start with a regular chain of atoms that is not linear, but in the form of a circle. We will assume that there are $N \gg 1$ atoms, separated by distance b , so that this atomic chain has length $L = Nb$. The lattice sites have coordinates x_0, x_1, \dots, x_{N-1}

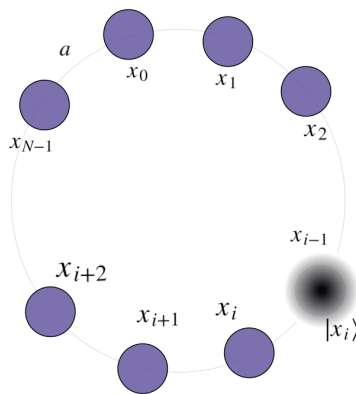


Figure 5.1: A circular lattice of atoms. An electron ‘injected’ into the lattice can form a weakly bound state $|x_i\rangle$ around a lattice atom at x_i

Say, an external electron is introduced into this lattice. What will be the lowest energy configuration of such an electron? If there was only one lattice atom, it could form a weakly bound ionic state with the atom, and this would be its lowest energy state. Now, imagine that the lattice atoms are separated by a large distance. Then, an electron bound this way to one of the atoms would not be able to ‘see’ the other atoms due to the large separation, and its state will not change with time. Say, the electron is bound this way to the i^{th} atom, located at x_i . We call this state $|x_i\rangle$. Since the other atoms are equivalent, it could form a similar weakly bound state with any of the other atoms, and all such states will have the same energy, say E_0 . We label these states as $|x_0\rangle, |x_1\rangle, \dots, |x_i\rangle, \dots, |x_{N-1}\rangle$. The situation is very similar to the one in a H_2^+ molecule, with these states being the analogues of the states in which the electron forms a

Hydrogen atom around one of the protons. Now, say we bring the lattice atoms closer. The electron, bound to atom located at x_i , will ‘sense’ the presence of the atoms located at x_{i-1} and x_{i+1} (nearest neighbors of x_i). Then, states $|x_i\rangle$ will no longer be states in which energy is well-defined, since the interaction of the electron with the neighboring atoms will change its state. We will now call states $|x_i\rangle$ as ‘states of definite position’, since the electron has an approximately well-defined position (upto an atom) in such a state, it being the position coordinate x_i of the lattice site. We will assume that states $|x_i\rangle$, with $i = 0, 1, \dots, N - 1$ form a basis in the vector space of electron states. Since the position of the electron in these states is well-defined, we define a Hermitian operator, the ‘position’ operator \hat{x} , which is such that these states are eigenstates of this operator, with eigenvalues equal to the position coordinate of the electron in the state

$$\hat{x} |x_i\rangle = x_i |x_i\rangle \quad (5.1)$$

We assume that these states are orthonormal

$$\langle x_i | x_j \rangle = \delta_{ij} \quad (5.2)$$

Let \hat{H} be the Hamiltonian operator, which is the Hermitian operator corresponding to energy of the electron. eigenvectors of this operator will be states of well-defined energy. We expect there to be N such states, since the vector space is N -dimensional. Any such state $|E\rangle$ satisfies

$$\hat{H} |E\rangle = E |E\rangle \quad (5.3)$$

If the lattice atoms were very far apart, states $|x_i\rangle$ would have well-defined energy, and would be eigenstates of \hat{H} . Then, in the $|x_i\rangle$ basis the matrix representation for \hat{H} would look like

$$H = \begin{pmatrix} E_0 & 0 & \cdot & \cdot & 0 \\ 0 & E_0 & 0 & \cdot & 0 \\ 0 & 0 & E_0 & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot & \cdot & E_0 \end{pmatrix} \quad (5.4)$$

However, since the nearest neighbors do interact with the electron, analogous to (3.7), the Hamiltonian matrix in the $|x_i\rangle$ basis will be

$$H = \begin{pmatrix} E_0 & -A & 0 & \cdot & \cdot & 0 & 0 \\ -A & E_0 & -A & 0 & \cdot & 0 & 0 \\ 0 & -A & E_0 & -A & 0 & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \cdot & 0 & -A & E_0 & -A & 0 \\ 0 & \cdot & \cdot & 0 & -A & E_0 & -A \\ 0 & 0 & \cdot & \cdot & 0 & -A & E_0 \end{pmatrix} \quad (5.5)$$

Because only the nearest neighbors are assumed to be relevant (apart from the site the electron is located on), only the diagonal and the next to diagonal entries are non-zero. The number A is assumed to be positive, so that the $(-A)$ appearing in the matrix corresponds to an attractive interaction. An energy eigenstate $|E\rangle$ can be written as linear superpositions of the basis states $|x_j\rangle$

$$|E\rangle = \sum_{j=0}^{N-1} c(x_j) |x_j\rangle \quad (5.6)$$

with a column vector representation

$$\psi_E = \begin{pmatrix} c(x_0) \\ c(x_1) \\ c(x_2) \\ \cdot \\ \cdot \\ \cdot \\ c(x_{N-1}) \end{pmatrix} \quad (5.7)$$

Note that we are writing the expansion coefficients as $c(x_i)$, instead of c_i . It is deliberate, and will become clear later why we are using such notation. Since the basis $\{|x_i\rangle\}$ is orthonormal, therefore the expansion coefficient $c(x_i)$ is

$$c(x_i) = \langle x_i | E \rangle \quad (5.8)$$

which can be interpreted as the probability amplitude of detecting the electron at lattice site x_i , if a position measurement is made on the state $|E\rangle$. In matrix form, equation (5.3) will be

$$H\psi_E = E\psi_E \quad (5.9)$$

The i^{th} element of this equation will be

$$\sum_{j=0}^{N-1} H_{ij}c(x_j) = Ec(x_i) \quad (5.10)$$

It is easy to see from (5.5) that the elements of matrix H are of the form

$$H_{ij} = E_0\delta_{i,j} - A\delta_{i,j+1} - A\delta_{i,j-1} \quad (5.11)$$

Substituting this in (5.10), we get

$$\sum_{j=0}^{N-1} (E_0\delta_{i,j} - A\delta_{i,j+1} - A\delta_{i,j-1})c(x_j) = Ec(x_i) \quad (5.12)$$

The summation over the kronecker deltas is easy, and we get

$$E_0 c(x_i) - A c(x_{i-1}) - A c(x_{i+1}) = E c(x_i) \quad (5.13)$$

To solve for $c(x_i)$, we use a trial solution

$$c(x_i) = e^{ikx_i} \quad (5.14)$$

Substituting it in (5.13), we get

$$E_0 e^{ikx_i} - A e^{ikx_{i-1}} - A e^{ikx_{i+1}} = E e^{ikx_i} \quad (5.15)$$

Since the spacing between the lattice points is b , $x_{i+1} = x_i + b$ and $x_{i-1} = x_i - b$. Substituting this in (5.15), we get

$$E_0 - A e^{-ikb} - A e^{ikb} = E \quad (5.16)$$

which can be solved for E to give

$$E = E_0 - 2A \cos(kb) \quad (5.17)$$

It might appear from (5.3) that k can take arbitrary values. However, there is a constraint that the expansion coefficients need to satisfy because of the fact that the lattice chain is a circle. This constraint can be imposed very easily in the following way. We let the position coordinates x_i run from $-\infty$ to ∞ (in steps of b), but make the identification

$$x_{i+N} \sim x_i \quad (5.18)$$

which says that points x_i and x_{i+N} are *identified*, in the sense that these coordinates actually represent the *same* lattice site (since it is a circle of N sites). Mathematically, this imposes the following constraint on the coefficients $c(x_i)$

$$c(x_i + Nb) = c(x_i) \quad (5.19)$$

Substituting this in (5.14), we get

$$e^{ik(x_i+Nb)} = e^{ikx_i} \quad (5.20)$$

which gives

$$e^{ikNb} = 1 \quad (5.21)$$

This constrains k to the following possible values

$$k_n = \frac{2n\pi}{Nb} \quad (5.22)$$

where n is an integer. It might appear that there are an infinite number of possible values of k , given that n can take any integer value. However, it is easy to see that only N such values are of physical significance. To see this, we first label the energy states by $|k_n\rangle$, corresponding to the allowed values of k_n . These states, when expanded in the position basis (5.6), will be

$$|k_n\rangle = \sum_{j=0}^{N-1} c_n(x_j) |x_j\rangle \quad (5.23)$$

where $c_n(x_j) = e^{ik_n x_j}$. Given a state $|k_n\rangle$ with this expansion, let us calculate the expansion for state $|k_{n+N}\rangle$ (assuming n takes arbitrary integer values). The expansion coefficients for this state will be

$$c_{n+N}(x_j) = e^{ik_{n+N} x_j} \quad (5.24)$$

From the form of k_n , it is easy to see that $k_{n+N} = k_n + 2\pi/b$. Using this, we see that

$$c_{n+N}(x_j) = e^{i\frac{2\pi x_j}{b}} e^{ik_n x_j} \quad (5.25)$$

But, $x_j = x_0 + j \times b$. Since $e^{i2\pi j} = 1$ (j is an integer), therefore

$$\begin{aligned} c_{n+N}(x_j) &= e^{i\frac{2\pi x_0}{b}} e^{ik_n x_j} \\ &= e^{i\frac{2\pi x_0}{b}} c_n(x_j) \end{aligned} \quad (5.26)$$

which implies that $|k_{n+N}\rangle = e^{i2\pi x_0/b} |k_n\rangle$. That is, $|k_{n+N}\rangle$ is physically the same state as $|k_n\rangle$. Then, there are just N independent states, corresponding to N possible integer values of n . We assume $N = 2M + 1$ (so that N is odd), and choose these values to be $n = 0, \pm 1, \pm 2, \dots, \pm M$. At the same time, we relabel the position coordinates as $x_0 = 0$, $x_j = jb$ with $j = \pm 1, \pm 2, \dots, \pm M$. Then, the energy states are given by

$$|k_n\rangle = \frac{1}{\sqrt{N}} \sum_{j=-M}^M e^{ik_n x_j} |x_j\rangle \quad (5.27)$$

with corresponding energy eigenvalues

$$E_n = E_0 - 2A \cos(k_n b) \quad (5.28)$$

with

$$k_n = \frac{2n\pi}{Nb}; \quad n = 0, \pm 1, \pm 2, \dots, \pm M \quad (5.29)$$

where $N = 2M + 1$. In (5.27), $1/\sqrt{N}$ is a normalization factor. Then, there are N states of definite energy (as it should be, since it is an N dimensional vector space). Let us calculate the maximum and minimum energy. First, we note that the maximum and minimum values of k_n are $k = \pm 2M\pi/Nb$. Since $N \gg 1$ and $N = 2M + 1$, we have

$$-\frac{\pi}{b} \leq k_n \leq \frac{\pi}{b} \quad (5.30)$$

in steps of $\Delta k = 2\pi/Nb$. Next, we observe that there is a *degeneracy* in energy. Since k_n appears as argument of a cos function, since states $|k_n\rangle$ and $|k_{-n}\rangle$ have the same energy, as $\cos(k_{-n}b) = \cos(-k_n b) = \cos(k_n b)$. However, this does not imply that states $|k_n\rangle$ and $|k_{-n}\rangle$ are physically the same. In fact, the N states form an orthonormal set. To see this, recalling that $x_j = jb$, we rewrite (5.27) in the following way

$$|k_n\rangle = \frac{1}{\sqrt{N}} \sum_{j=-M}^M \omega^{nj} |x_j\rangle \quad (5.31)$$

where $\omega = e^{i2\pi/N}$ is the N^{th} root of unity. This satisfies $\omega^N = 1$, $\omega^{N+j} = \omega^j$. Then,

$$\begin{aligned}
\langle k_n | k_m \rangle &= \frac{1}{N} \sum_{j=-M}^M \sum_{l=-M}^M (\omega^*)^{nj} (\omega)^{ml} \langle x_j | x_l \rangle \\
&= \frac{1}{N} \sum_{j=-M}^M \sum_{l=-M}^M (\omega^*)^{nj} (\omega)^{ml} \delta_{jl} \\
&= \frac{1}{N} \sum_{j=-M}^M (\omega^*)^{nj} (\omega)^{mj} \\
&= \frac{1}{N} \sum_{j=-M}^M \omega^{j(n-m)}
\end{aligned} \tag{5.32}$$

where we have used $(\omega^*)^j = \omega^{-j}$. The sum is a geometric series of the form $S = \sum_{j=-M}^M a^j$, where $a = \omega^{n-m}$. Let us first assume that $n \neq m$. Then, the series sums to $S = (a^{-M}/a - 1)(a^{2M+1} - 1)$. But, $a^{2M+1} = a^N = \omega^{N(n-m)} = 1$, since $\omega^N = 1$. Therefore, $S = 0$, and $\langle k_n | k_m \rangle = 0$. Next, if $n = m$, the series just sums to N (since there are N terms, all of them 1). Then, $\langle k_n | k_n \rangle = 1$. This can be summarised as

$$\langle k_n | k_m \rangle = \delta_{nm} \tag{5.33}$$

which just says that $\{|k_n\rangle\}$ is an orthonormal set of vectors. Clearly, there is more to states $|k_n\rangle$ than energy, since only half of them have distinct energies. Let us define a Hermitian operator \hat{p} , which we call the *Lattice Momentum* operator. It will be clear soon why we choose to call it ‘momentum’. The action of any operator is completely defined, if we define its action on a basis. Since $\{|k_n\rangle\}$ is a set of N orthonormal vectors, they form a basis in \mathcal{H} . Then, we define the operator \hat{p} by its action on the $\{|k_n\rangle\}$ basis

$$\hat{p} |k_n\rangle = \hbar k_n |k_n\rangle \tag{5.34}$$

that is, we define states $|k_n\rangle$ to be eigenvectors of \hat{p} , with eigenvalue $\hbar k_n$. At this point, it is useful to change notation. We define $p_n = \hbar k_n$ and relabel states $|k_n\rangle$ as $|p_n\rangle$. Then, (5.27) can be written as

$$|p_n\rangle = \frac{1}{\sqrt{N}} \sum_{j=-M}^M e^{ip_n x_j / \hbar} |x_j\rangle \tag{5.35}$$

with

$$p_n = \frac{2n\pi\hbar}{Nb}; \quad n = 0, \pm 1, \pm 2, \dots, \pm M \tag{5.36}$$

This equation can be easily inverted to express states $|x_j\rangle$ in terms of states $|p_n\rangle$. Since states $|p_n\rangle$ are orthonormal, we can write (see (4.6))

$$|x_j\rangle = \sum_{n=-M}^M |p_n\rangle \langle p_n | x_j \rangle \tag{5.37}$$

From (5.35), we can calculate $\langle p_n | x_j \rangle$ by calculating $\langle x_j | p_n \rangle$, followed by complex conjugation

$$\begin{aligned}
\langle x_j | p_n \rangle &= \frac{1}{\sqrt{N}} \sum_{i=-M}^M e^{ip_n x_i / \hbar} \langle x_j | x_i \rangle \\
&= \frac{1}{\sqrt{N}} \sum_{i=-M}^M e^{ip_n x_i / \hbar} \delta_{ji} \\
&= \frac{1}{\sqrt{N}} e^{ip_n x_j / \hbar}
\end{aligned} \tag{5.38}$$

Taking the complex conjugate of this equation, we get

$$\langle p_n | x_j \rangle = \frac{1}{\sqrt{N}} e^{-ip_n x_j / \hbar} \quad (5.39)$$

Then, the position states expanded in terms of the lattice momentum states will be

$$|x_j\rangle = \frac{1}{\sqrt{N}} \sum_{n=-M}^M e^{-ip_n x_j / \hbar} |p_n\rangle \quad (5.40)$$

In vector space \mathcal{H} , we can choose either $\{|x_i\rangle\}$ or $\{|p_n\rangle\}$ as a basis. Then, a given quantum state $|\psi\rangle$ can be expanded in either basis. If we expand it in the position basis, the expansion will be of the form

$$|\psi\rangle = \sum_{j=-M}^M \psi(x_j) |x_j\rangle \quad (5.41)$$

The expansion coefficients $\psi(x_i)$ can be visualized as a function ψ defined over a discrete set of points x_i . Since the basis is orthonormal, $\psi(x_i) = \langle x_i | \psi \rangle$, which can be thought of as the probability amplitude that the electron in this state is detected in vicinity of the lattice point x_i . In a real lattice, the lattice spacing would be of the order of a few angstroms, so if we are not interested in what happens at this scale but are interested in the broad properties of propagation of the electron, we could visualize $\psi(x)$ as an approximately continuous function of ‘position coordinate’ x . This ‘continuous’ function (which is strictly speaking a set of N probability amplitudes) describes the state $|\psi\rangle$ completely in the position basis. This function is called a *Wavefunction*. We could equivalently expand the same state in the ‘momentum’ basis $|p_n\rangle$, and the expansion would be

$$|\psi\rangle = \sum_{n=-M}^M \tilde{\psi}(p_n) |p_n\rangle \quad (5.42)$$

The expansion coefficient $\tilde{\psi}(p_n) = \langle p_n | \psi \rangle$ can be thought of as the probability amplitude that the electron in state $|\psi\rangle$ is found to have ‘momentum’ p_n (if a measurement of momentum is made). Again, for a real lattice (with $N \gg \gg 1$ and given the ‘smallness’ of Planck’s constant), the spacing in the momentum eigenvalues $\Delta p = 2\pi\hbar/Nb$ is too small to be of significance, and we could visualize $\tilde{\psi}(p)$ as an almost continuous function of p . This is called a *Momentum Space Wavefunction*. Since the two bases are related through (5.35) and (5.40), the wavefunctions $\psi(x_i)$ and $\tilde{\psi}(p_n)$ are related. Taking the inner product of (5.41) with $|p_n\rangle$, we get

$$\langle p_n | \psi \rangle = \sum_{j=-M}^M \psi(x_j) \langle p_n | x_j \rangle \quad (5.43)$$

Using (5.39), we get the desired relationship

$$\tilde{\psi}(p_n) = \frac{1}{\sqrt{N}} \sum_{j=-M}^M \psi(x_j) e^{-ip_n x_j / \hbar} \quad (5.44)$$

Taking the inner product of (5.58) with $|x_j\rangle$, we get the inverse relationship

$$\psi(x_j) = \frac{1}{\sqrt{N}} \sum_{n=-M}^M \tilde{\psi}(p_n) e^{ip_n x_j / \hbar} \quad (5.45)$$

Mathematically, these equations represent a *Discrete Fourier Transform*. The two kinds of wavefunctions are discrete fourier transforms of each other. These wavefunctions have an interesting property. Let us

say that state $|\psi\rangle$ is such that the electron is localized around a single lattice site, say x_0 . Then, its position space wavefunction will be

$$\begin{aligned}\psi(x_i) &= \langle x_i | \psi \rangle \\ &= \langle x_i | x_0 \rangle \\ &= \delta_{i0}\end{aligned}\tag{5.46}$$

which is zero unless $x_i = x_0$, something that we would expect. The momentum wavefunction will be

$$\begin{aligned}\tilde{\psi}(p_n) &= \frac{1}{\sqrt{N}} \sum_{j=-M}^M \psi(x_j) e^{-ip_n x_j / \hbar} \\ &= \frac{1}{\sqrt{N}} \sum_{j=-M}^M \delta_{j0} e^{-ip_n x_j / \hbar} \\ &= \frac{1}{\sqrt{N}} e^{-ip_n x_0 / \hbar}\end{aligned}\tag{5.47}$$

What if we measured the momentum of the particle in such a state? The probability that we measure it to be p_n will be

$$\begin{aligned}P(p_n) &= |\tilde{\psi}(p_n)|^2 \\ &= \frac{1}{N}\end{aligned}\tag{5.48}$$

which is independent of p_n ! Then, if the electron is localized around a lattice site, its momentum is *completely indefinite*. The probability of measuring any momentum is the same. Now, say that the electron is in a state of well-defined momentum, so that say $|\psi\rangle = |p_m\rangle$. Then, its momentum is well-defined and $\psi(p_n) = \langle p_n | p_m \rangle = \delta_{nm}$, but let us calculate its position space wavefunction. Using (5.45), it will be

$$\begin{aligned}\psi(x_j) &= \frac{1}{\sqrt{N}} \sum_{n=-M}^M \psi(p_n) e^{ip_n x_j / \hbar} \\ &= \frac{1}{\sqrt{N}} \sum_{n=-M}^M \delta_{nm} e^{ip_n x_j / \hbar} \\ &= \frac{1}{\sqrt{N}} e^{ip_m x_j / \hbar}\end{aligned}\tag{5.49}$$

Therefore, the probability of detecting the electron in the vicinity of the lattice site at x_j is

$$\begin{aligned}P(x_j) &= |\psi(x_j)|^2 \\ &= \frac{1}{N}\end{aligned}\tag{5.50}$$

which is independent of x_j ! The position of the electron is completely indefinite. Let us now analyze the energy of the electron. We have seen that states $|p_n\rangle$ are states of well-defined energy, such that

$$\hat{H} |p_n\rangle = E_n |p_n\rangle\tag{5.51}$$

where

$$E_n = E_0 - 2A \cos\left(\frac{p_n b}{\hbar}\right)\tag{5.52}$$

with p_n given in (5.36). Since E_n is an even function of p_n , it follows that states $|p_n\rangle$ and $|p_{-n}\rangle = |-p_n\rangle$ have the same energy. They are *degenerate*. Then, an interesting situation arises. Say, the electron state is $|\psi\rangle = (1/\sqrt{2})(|p_n\rangle + |-p_n\rangle)$. This is clearly an eigenstate of \hat{H}

$$\begin{aligned}\hat{H}|\psi\rangle &= \frac{1}{\sqrt{2}}(\hat{H}|p_n\rangle + \hat{H}|-p_n\rangle) \\ &= \frac{1}{\sqrt{2}}(E_n|p_n\rangle + E_n|-p_n\rangle) \\ &= E_n|\psi\rangle\end{aligned}\tag{5.53}$$

Then, a measurement of energy on this state will give $E = E_n$ for certain. However, if the momentum of the electron is measured, it is equally likely that it will be measured to be $+p_n$, as $-p_n$. The electron, with equal probability, will be detected moving to the ‘left’, or ‘right’! In an actual lattice, $N \gg \gg 1$. Then, the difference in successive values of momentum $\Delta p = 2\pi\hbar/Nb$ will be practically zero, and the momentum can take continuous values between $-\hbar\pi/b$ and $+\hbar\pi/b$

$$-\frac{\hbar\pi}{b} \leq p \leq \frac{\hbar\pi}{b}\tag{5.54}$$

The momentum then lies in a *band*. The corresponding energy also lies in a band, called the *Conduction Band*.

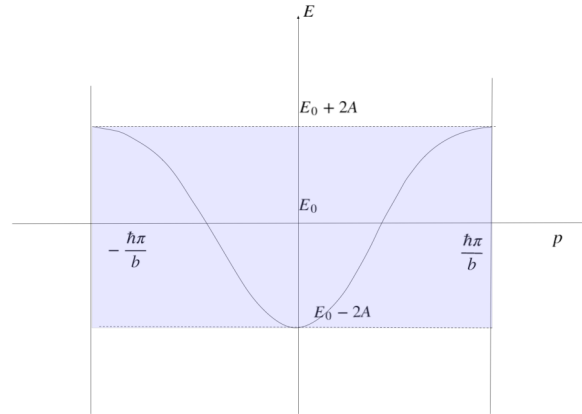


Figure 5.2: Band of allowed electron energies in a one-dimensional crystal lattice

Now, we justify identifying the states $|p_n\rangle$ as ‘momentum’ states, and the operator \hat{p} as a ‘momentum’ operator. In (5.52), let us expand the cos function about $p = 0$, and retain terms upto $O(p^2)$. This gives

$$E(p) = E_0 - 2A + \frac{Ab^2}{\hbar^2}p^2\tag{5.55}$$

We rewrite this as follows

$$E(p) = \frac{p^2}{2m_{\text{eff}}} + E_0 - 2A\tag{5.56}$$

where

$$m_{\text{eff}} = \frac{\hbar^2}{2Ab^2}\tag{5.57}$$

Equation (5.56) is very similar to the expression for the energy of a particle of mass m_{eff} and momentum p , and potential energy $E_0 - 2A$. We interpret this as follows: For ‘small’ values of p_n , the electron propagates in the lattice as if it has an *effective mass* m_{eff} , which is different from its inertial mass, and is a function of the interaction between the electron and the lattice atoms. The state $|p_n\rangle$ for small p_n is

a state of well-defined momentum p_n , and the energy of the electron in this state is given by (5.56). The details of the interaction of the electron with the lattice are ‘hidden’ in the effective mass, and the constant background potential energy $E_0 - 2A$.

We now analyze the time-evolution of an arbitrary quantum state on the lattice. Let the state at $t = 0$ be $|\psi(0)\rangle$, which we expand in the $|p_n\rangle$ basis, since this is also an energy eigenbasis

$$|\psi(0)\rangle = \sum_{n=-M}^M \tilde{\psi}(p_n, 0) |p_n\rangle \quad (5.58)$$

where $\tilde{\psi}(p_n, 0)$ are the expansion coefficients. The state at instant t will be given by (4.42)

$$\begin{aligned} |\psi(t)\rangle &= e^{-i\hat{H}t/\hbar} |\psi(0)\rangle \\ &= \sum_{n=-M}^M \tilde{\psi}(p_n, 0) e^{-i\hat{H}t/\hbar} |p_n\rangle \\ &= \sum_{n=-M}^M \tilde{\psi}(p_n, 0) e^{-iE_n t/\hbar} |p_n\rangle \end{aligned} \quad (5.59)$$

where E_n is given by (5.52). If we are given the position space expansion at $t = 0$, we can determine $\tilde{\psi}(p_n, 0)$ using (5.44)

$$\tilde{\psi}(p_n, 0) = \frac{1}{\sqrt{N}} \sum_{j=-M}^M \psi(x_j, 0) e^{-ip_n x_j/\hbar} \quad (5.60)$$

The position space expansion at instant t can be found by taking the inner product of (5.59) with respect to $|x_i\rangle$

$$\begin{aligned} \psi(x_i, t) &= \sum_{n=-M}^M \tilde{\psi}(p_n, 0) e^{-iE_n t/\hbar} \langle x_i | p_n \rangle \\ &= \frac{1}{\sqrt{N}} \sum_{n=-M}^M \tilde{\psi}(p_n, 0) e^{-iE_n t/\hbar} e^{ip_n x_i/\hbar} \\ &= \frac{1}{N} \sum_{n=-M}^M \sum_{j=-M}^M \psi(x_j, 0) e^{-iE_n t/\hbar} e^{ip_n(x_i - x_j)/\hbar} \end{aligned} \quad (5.61)$$

where we have used (5.38) and (5.44).

5.2 The continuum limit

We now analyze the quantum propagation of an electron in free space. To begin with, we will assume that the electron is confined to a circle, and eventually we will take the limit of the radius of the circle approaching infinity. To analyze free propagation on a circle, we start with our circular lattice, and take the limit $N \rightarrow \infty$ and $b \rightarrow 0$, such that $Nb = L$. At the same time, we take the limit that the interaction of the electron with the neighboring atoms approaches zero. We need to be careful in taking the limit $b \rightarrow 0$ and this interaction going to zero. The number A which enters the expression for energy is sensitive to b , and we expect it to increase with decreasing b (the ease with which the electron can ‘leak’ to a neighboring lattice atom should increase with decreasing b). With this in mind, we first analyze what happens to the momentum eigenvalues (5.36) in this limit. Since $Nb = L$, these will be

$$p_n = \frac{2n\pi\hbar}{L}; \quad n = 0, \pm 1, \pm 2, \dots \quad (5.62)$$

These appear to be the same as before, *except* that now since $N \rightarrow \infty$, n can take any integer value. Next, we observe that in the expression for energy (5.52), for any finite n , since $p_n b / \hbar = 2n\pi/N$ approaches zero, we can expand the cos term to retain terms upto $O(p^2)$. Further, we discard the p independent part $E_0 - 2A$ to get

$$E_n = \frac{Ab^2}{\hbar^2} p_n^2 \quad (5.63)$$

At this point, we wish to make the lattice disappear in a way that does not introduce any singular behavior. So, we assume that the limit $b \rightarrow 0$ is taken in such a way, that together with the neighbor interaction being tuned down, the quantity $\frac{Ab^2}{\hbar^2}$ approaches a finite value. We introduce a ‘mass’ parameter $m = \frac{\hbar^2}{2Ab^2}$, to get

$$E_n = \frac{p_n^2}{2m} \quad (5.64)$$

which is just the classical relation between energy and momentum, except that now momentum can take discrete values. We might as well call the parameter m as the inertial mass of the particle, since with the lattice gone, that is what it would experimentally be determined to be. We now shift attention to the states $|x_i\rangle$ and see how they behave in the limits we have taken. Since $b \rightarrow 0$, x_i will take continuous values. Then, the states of well-defined position will be labelled $|x\rangle$, with

$$-L/2 < x \leq L/2 \quad (5.65)$$

Notice that we have excluded the point $x = -L/2$. This is because on a circle, this coordinate refers to the same point as coordinate $x = +L/2$. Since x is continuous, we immediately encounter a problem with the normalization of $|x\rangle$ states. On the lattice, the states $|x_i\rangle$ were normalized such that $\langle x_i | x_j \rangle = \delta_{ij}$. On the continuous circle, this inner product will be replaced by $\langle x' | x \rangle$, where x' and x are continuous. Clearly, in the absence of a discrete index, there seems to be a problem in defining this inner product. We then need to look at the continuum analogue of the kronecker delta. This is just the Dirac delta ‘function’, defined as follows

$$\begin{aligned} \delta(x' - x) &= \begin{cases} \infty & x' = x \\ 0 & x' \neq x \end{cases} \\ \int_{-L/2}^{L/2} dx' \delta(x' - x) &= 1 \end{aligned} \quad (5.66)$$

These properties also lead to the following

$$\int_{-L/2}^{L/2} dx' f(x') \delta(x' - x) = f(x) \quad (5.67)$$

In all these properties, it is assumed that both x and x' lie in the domain (5.65). To make a smooth transition from the discrete to the continuum situation, it is useful to scale the position and momentum states as follows

$$\begin{aligned} |x_i\rangle &\rightarrow \frac{1}{\sqrt{b}} |x_i\rangle \\ |p_n\rangle &\rightarrow \sqrt{\frac{Nb}{2\pi\hbar}} |p_n\rangle \end{aligned} \quad (5.68)$$

With this, the inner product of states $|x_i\rangle$ will be

$$\langle x_i | x_j \rangle = \frac{1}{b} \delta_{ij} \quad (5.69)$$

It is easy to see that this is equivalent to a Dirac delta function in the limit $b \rightarrow 0$, with $x_i = x'$, $x_j = x$. First, if $x' \neq x$, this is zero because of the kronecker delta. If $x' = x$, this is equal to $1/b$ which diverges as

$b \rightarrow 0$. To complete the identification with the delta function, we take the integral of $\langle x' | x \rangle$ over x' . We first note that in the limit $b \rightarrow 0$, the summation over the lattice index can be replaced by an integral over x , with the identification

$$\begin{aligned} \sum_{i=-M/2}^{M/2} b &= \sum_{x=-L/2}^{L/2} \Delta x \\ &\rightarrow \int_{-L/2}^{L/2} dx \end{aligned} \quad (5.70)$$

where $\Delta x = b$ is the increment in x , which approaches dx as b approaches zero. Then,

$$\begin{aligned} \int_{-L/2}^{L/2} dx' \langle x' | x \rangle &= \lim_{b \rightarrow 0} \sum_i \left(\frac{1}{b} \delta_{ij} \right) b \\ &= 1 \end{aligned} \quad (5.71)$$

which completes the identification

$$\langle x' | x \rangle = \delta(x' - x) \quad (5.72)$$

With the scaling (5.68), the momentum normalization is

$$\langle p_m | p_n \rangle = \frac{L}{2\pi\hbar} \delta_{nm} \quad (5.73)$$

The expansions (5.35) and (5.40) also change, and with the identification (5.70), these are now

$$\begin{aligned} |p_n\rangle &= \frac{1}{\sqrt{2\pi\hbar}} \int_{-L/2}^{L/2} dx e^{ip_n x/\hbar} |x\rangle \\ |x\rangle &= \frac{\sqrt{2\pi\hbar}}{L} \sum_{n=-\infty}^{\infty} e^{-ip_n x/\hbar} |p_n\rangle \end{aligned} \quad (5.74)$$

The eigenvalue equations for position and momentum are

$$\begin{aligned} \hat{x} |x\rangle &= x |x\rangle \\ \hat{p} |p_n\rangle &= p_n |p_n\rangle \end{aligned} \quad (5.75)$$

with the eigenvalues given by (5.65) and (5.62). Given (5.64), the Hamiltonian operator will be

$$\hat{H} = \frac{1}{2m} \hat{p}^2 \quad (5.76)$$

with

$$\hat{H} |p_n\rangle = \frac{p_n^2}{2m} |p_n\rangle \quad (5.77)$$

Then, the energy of a free particle propagating on a circle forms a discrete set. As before, we can expand an arbitrary quantum state $|\psi\rangle$ in either the position or momentum basis. In the position basis, the expansion will be

$$|\psi\rangle = \int_{-L/2}^{L/2} dx \psi(x) |x\rangle \quad (5.78)$$

Where the sum over $|x_i\rangle$ is replaced by an integral. The expansion coefficients $\psi(x)$ constitute the wavefunction of the particle. It is important to remember that this function must be *single valued* at all points x on the circle. Therefore, it should satisfy the constraint $\psi(x + L) = \psi(x)$. In particular, $\psi(L/2) = \psi(-L/2)$

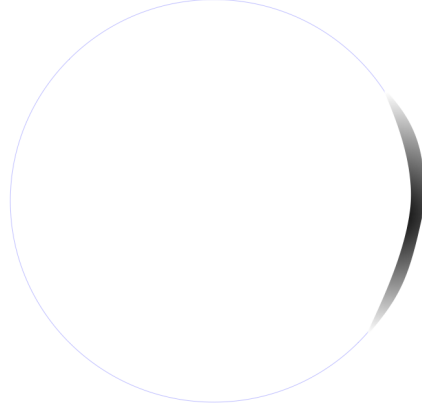


Figure 5.3: Quantum particle propagating on a circle. The wavefunction of the particle is shown spread out on the circle.

Taking the inner product of (5.78) with state $|x'\rangle$, we get

$$\begin{aligned}
 \langle x' | \psi \rangle &= \int_{-L/2}^{L/2} dx \psi(x) \langle x' | x \rangle \\
 &= \int_{-L/2}^{L/2} dx \psi(x) \delta(x' - x) \\
 &= \psi(x')
 \end{aligned} \tag{5.79}$$

Then, $\psi(x) = \langle x | \psi \rangle$, and we can write

$$|\psi\rangle = \int_{-L/2}^{L/2} dx |x\rangle \langle x | \psi \rangle \tag{5.80}$$

Because of the rescaling (5.68), the modulus squared of the wavefunction does not give the probability of detecting the particle at point x . Instead, it has the interpretation of probability *density*. To see this, let us start with the discrete formalism, and calculate the total probability of detecting the particle between points x_a corresponding to lattice index n_a and point x_b corresponding to lattice index n_b . From (5.41), this will be

$$\begin{aligned}
 P &= \sum_{i=n_a}^{n_b} |\psi(x_i)|^2 \\
 &= \sum_{i=n_a}^{n_b} |\langle x_i | \psi \rangle|^2
 \end{aligned} \tag{5.81}$$

where $|x_i\rangle$ is the unscaled state of well-defined position. In terms of the new, scaled states, this probability is

$$P = \sum_{i=n_a}^{n_b} b |\langle x_i | \psi \rangle|^2 \tag{5.82}$$

Taking the continuum limit, the probability of detecting the particle between points x_a and x_b will be

$$P = \int_{x_a}^{x_b} dx |\psi(x)|^2 \tag{5.83}$$

Clearly, $|\psi(x)|^2$ is a probability per unit ‘volume’ (here, length). We will talk about the momentum expansion later, when we take the limit $L \rightarrow \infty$. Let us now calculate the matrix elements of the momentum

operator in the $|x\rangle$ basis. We first expand the state $|x\rangle$ in the $|p_n\rangle$ basis (5.74) and operate on it by \hat{p} to get

$$\begin{aligned}\hat{p}|x\rangle &= \frac{\sqrt{2\pi\hbar}}{L} \sum_{n=-\infty}^{\infty} e^{-ip_n x/\hbar} \hat{p}|p_n\rangle \\ &= \frac{\sqrt{2\pi\hbar}}{L} \sum_{n=-\infty}^{\infty} e^{-ip_n x/\hbar} p_n |p_n\rangle\end{aligned}\quad (5.84)$$

where we have used the fact the $|p_n\rangle$ is an eigenvector of \hat{p} . Taking the inner product with $|x'\rangle$, we get

$$\begin{aligned}\langle x'|\hat{p}|x\rangle &= \frac{\sqrt{2\pi\hbar}}{L} \sum_{n=-\infty}^{\infty} e^{-ip_n x/\hbar} p_n \langle x'|p_n\rangle \\ &= \frac{\sqrt{2\pi\hbar}}{L} \sum_{n=-\infty}^{\infty} e^{-ip_n x/\hbar} p_n \frac{1}{\sqrt{2\pi\hbar}} e^{ip_n x'/\hbar} \\ &= \frac{1}{L} \sum_{n=-\infty}^{\infty} p_n e^{ip_n(x'-x)/\hbar}\end{aligned}\quad (5.85)$$

To evaluate this sum, we define $S(x', x) = (1/L) \sum_{n=-\infty}^{\infty} e^{ip_n(x'-x)/\hbar}$. We can then express $\langle x'|\hat{p}|x\rangle$ as a derivative of $S(x', x)$ with respect to x'

$$\langle x'|\hat{p}|x\rangle = \frac{\hbar}{i} \frac{d}{dx'} S(x', x) \quad (5.86)$$

It is easy to check that a term by term differentiation of $S(x', x)$ with respect to x' will generate the series. Next, we evaluate $S(x', x)$ as follows

$$\begin{aligned}S(x', x) &= \frac{1}{L} \lim_{N \rightarrow \infty} \sum_{n=-N}^N e^{ip_n(x'-x)/\hbar} \\ &= \frac{1}{L} \lim_{N \rightarrow \infty} \sum_{n=-N}^N a^n \\ &= \frac{1}{L} \lim_{N \rightarrow \infty} a^{-N} \left(\frac{a^{2N} - 1}{a - 1} \right)\end{aligned}\quad (5.87)$$

where $a = e^{2\pi i(x'-x)/L}$. Substituting for a and a little simplification gives

$$S(x', x) = \frac{1}{L} \lim_{N \rightarrow \infty} e^{-\pi i(x'-x)/L} \frac{\sin\left(\frac{2\pi(x'-x)N}{L}\right)}{\sin\left(\frac{\pi(x'-x)}{L}\right)} \quad (5.88)$$

We first note that if $x' \neq x$, in the limit $N \rightarrow \infty$, the numerator of $S(x', x)$ is a sin function that oscillates infinitely rapidly as a function of x' . Then, if we were to perform an integral of some smooth function $f(x')$ multiplied with $S(x', x)$ over a region in which $x' \neq x$, the rapid oscillations in $S(x', x)$ will average out the integral to zero. Then, unless $x' = x$, the function $S(x', x)$ is effectively zero. Let us then put $x' = x$ in the exponential function, which makes it equal to 1. Then,

$$S(x', x) = \frac{1}{L} \lim_{N \rightarrow \infty} \frac{\sin\left(\frac{2\pi(x'-x)N}{L}\right)}{\sin\left(\frac{\pi(x'-x)}{L}\right)} \quad (5.89)$$

Next, let us evaluate the integral of $S(x', x)$ over x' from $-L/2$ to $L/2$. We denote the integral as I

$$\begin{aligned} I &= \int_{-L/2}^{L/2} dx' S(x', x) \\ &= \frac{1}{L} \lim_{N \rightarrow \infty} \int_{-L/2}^{L/2} dx' \frac{\sin\left(\frac{2\pi(x'-x)N}{L}\right)}{\sin\left(\frac{\pi(x'-x)}{L}\right)} \end{aligned} \quad (5.90)$$

We make the substitution $y = 2\pi(x' - x)N/L$. In the limit $N \rightarrow \infty$, the limits for y in the integral will be $-\infty$ and ∞ . Then, we get

$$I = \frac{1}{\pi N} \lim_{N \rightarrow \infty} \int_{-\infty}^{\infty} dy \frac{\sin y}{\sin(y/2N)} \quad (5.91)$$

We can write $\sin(y/2N) = y/2N$ in this limit, and the explicit N dependence goes away

$$\begin{aligned} I &= \frac{1}{\pi} \int_{-\infty}^{\infty} dy \frac{\sin y}{y} \\ &= 1 \end{aligned} \quad (5.92)$$

This tells us that the integral of $S(x', x)$ is unity. Clearly, $S(x', x)$ is a Dirac delta function

$$\begin{aligned} S(x', x) &= \frac{1}{L} \lim_{N \rightarrow \infty} \sum_{n=-N}^N e^{ip_n(x'-x)/\hbar} \\ &= \delta(x' - x) \end{aligned} \quad (5.93)$$

with this and (5.86), we get the matrix elements of \hat{p} in the $|x\rangle$ basis

$$\langle x' | \hat{p} | x \rangle = \frac{\hbar}{i} \frac{d}{dx'} \delta(x' - x) \quad (5.94)$$

We observed that states in which the lattice position of the particle is well-defined, its momentum is completely uncertain, and vice-versa. This tells us that the commutator of the \hat{x} and \hat{p} operators must be non-zero. We compute the action of this commutator on an arbitrary state $|\psi\rangle$. We first prove the following important result

$$\langle x | \hat{p} | \psi \rangle = \frac{\hbar}{i} \frac{d\psi(x)}{dx} \quad (5.95)$$

To prove this, we expand $|\psi\rangle$ in the $|x'\rangle$ basis (5.78) with x' as a dummy coordinate, operate with \hat{p} and take the inner product with $|x\rangle$

$$\langle x | \hat{p} | \psi \rangle = \int_{-L/2}^{L/2} dx' \psi(x') \langle x | \hat{p} | x' \rangle \quad (5.96)$$

We use (5.94) and extract the derivative with respect to x outside the integral (the integration variable is x' , not x) to get

$$\begin{aligned} \langle x | \hat{p} | \psi \rangle &= \int_{-L/2}^{L/2} dx' \psi(x') \frac{\hbar}{i} \frac{d}{dx} \delta(x - x') \\ &= \frac{\hbar}{i} \frac{d}{dx} \int_{-L/2}^{L/2} dx' \psi(x') \delta(x - x') \\ &= \frac{\hbar}{i} \frac{d\psi(x)}{dx} \end{aligned} \quad (5.97)$$

which proves the result. Now, we determine the action of the commutator of \hat{x} and \hat{p} on a state $|\psi\rangle$. First, we evaluate $\hat{x}\hat{p}|\psi\rangle$. To do this, we visualize this as $\hat{x}(\hat{p}|\psi\rangle)$ and expand the vector $\hat{p}|\psi\rangle$ in the $|x\rangle$ basis

$$\begin{aligned}\hat{p}|\psi\rangle &= \int_{-L/2}^{L/2} dx |x\rangle \langle x|\hat{p}|\psi\rangle \\ &= \int_{-L/2}^{L/2} dx \frac{\hbar}{i} \frac{d\psi(x)}{dx} |x\rangle\end{aligned}\quad (5.98)$$

where we have used (5.95). Operating by \hat{x} , we get

$$\begin{aligned}\hat{x}\hat{p}|\psi\rangle &= \frac{\hbar}{i} \int_{-L/2}^{L/2} dx \frac{d\psi(x)}{dx} \hat{x}|x\rangle \\ &= \frac{\hbar}{i} \int_{-L/2}^{L/2} dx x \frac{d\psi(x)}{dx} |x\rangle\end{aligned}\quad (5.99)$$

where we have used $\hat{x}|x\rangle = x|x\rangle$. Next, we evaluate $\hat{p}\hat{x}|\psi\rangle$. Let $|\phi\rangle = \hat{x}|\psi\rangle$. Then,

$$\begin{aligned}\hat{p}\hat{x}|\psi\rangle &= \hat{p}|\phi\rangle \\ &= \frac{\hbar}{i} \int_{-L/2}^{L/2} dx \frac{d\phi(x)}{dx} |x\rangle\end{aligned}\quad (5.100)$$

where we have used (5.98), with $|\psi\rangle$ replaced by $|\phi\rangle$. But,

$$\begin{aligned}\phi(x) &= \langle x|\phi\rangle \\ &= \langle x|\hat{x}|\psi\rangle \\ &= x \langle x|\psi\rangle \\ &= x \psi(x)\end{aligned}\quad (5.101)$$

Then,

$$\begin{aligned}\hat{p}\hat{x}|\psi\rangle &= \frac{\hbar}{i} \int_{-L/2}^{L/2} dx \frac{d}{dx}(x\psi(x)) |x\rangle \\ &= \frac{\hbar}{i} \int_{-L/2}^{L/2} dx \left(x \frac{d\psi(x)}{dx} + \psi(x) \right) |x\rangle\end{aligned}\quad (5.102)$$

The action of $[\hat{x}, \hat{p}]$ on $|\psi\rangle$ can now be determined

$$\begin{aligned}[\hat{x}, \hat{p}]|\psi\rangle &= \hat{x}\hat{p}|\psi\rangle - \hat{p}\hat{x}|\psi\rangle \\ &= i\hbar \int_{-L/2}^{L/2} dx \psi(x) |x\rangle \\ &= i\hbar |\psi\rangle\end{aligned}\quad (5.103)$$

Since this is true for any state $|\psi\rangle$, it follows that

$$[\hat{x}, \hat{p}] = i\hbar \hat{I} \quad (5.104)$$

where \hat{I} is the identity operator. This is the fundamental commutation relation for position and momentum, and the origin of the celebrated Uncertainty Principle. In the generalized uncertainty principle (4.87), if we substitute $\hat{A} = \hat{x}$ and $\hat{B} = \hat{p}$, given (5.104), we get

$$\Delta\hat{x} \Delta\hat{p} \geq \frac{\hbar}{2} \quad (5.105)$$

for any state $|\psi\rangle$. This is Heisenberg's Uncertainty Principle.

We are now ready to take the limit $L \rightarrow \infty$. In this limit, we have the particle propagating along a line described by coordinate $-\infty < x < \infty$. Just a few things change when we take this limit. Firstly, the coordinate x in the position basis $|x\rangle$ can take any real value. The states $|x\rangle$ are eigenstates of Hermitian operator \hat{x} , such that

$$\hat{x}|x\rangle = x|x\rangle \quad (5.106)$$

These are states in which the particle has a well-defined position, mathematically a point in space. These states have normalization given by

$$\langle x'|x\rangle = \delta(x' - x) \quad (5.107)$$

A quantum state $|\psi\rangle$ can be expanded in this basis as follows

$$|\psi\rangle = \int_{-\infty}^{\infty} dx \psi(x) |x\rangle \quad (5.108)$$

where

$$\psi(x) = \langle x|\psi\rangle \quad (5.109)$$

The probability of detecting the particle between points x_a and x_b in this state is, as before,

$$P = \int_{x_a}^{x_b} dx |\psi(x)|^2 \quad (5.110)$$

In particular, the total probability of detecting the particle somewhere must be unity. Then, the wavefunction $\psi(x)$ must satisfy the constraint

$$\int_{-\infty}^{\infty} dx |\psi(x)|^2 = 1 \quad (5.111)$$

Using the expansion (5.108) and (5.107), it is easy to check that this is equivalent to the condition that $\langle\psi|\psi\rangle = 1$. Condition (5.111) imposes an important constraint on physical wavefunctions. For the integral to converge, the wavefunction has to vanish asymptotically

$$\lim_{x \rightarrow \pm\infty} \psi(x) = 0 \quad (5.112)$$

The observable momentum is represented by Hermitian operator \hat{p} , with eigenstates $|p\rangle$ that represent states of well-defined momentum p . In the $L \rightarrow \infty$ limit, it follows from (5.62) that p will take continuous values. Further, the summation over discrete p_n will be replaced by an integral over p . To see the explicit conversion, we observe that the difference in successive momentum eigenvalues (5.62) is $\Delta p = 2\pi\hbar/L$. Then, given a function f of p_n

$$\begin{aligned} \frac{2\pi\hbar}{L} \sum_n f(p_n) &= \sum_n f(p_n) \Delta p f(p_n) \\ &\rightarrow \int_{-\infty}^{\infty} dp f(p) \end{aligned} \quad (5.113)$$

The normalization of momentum states $|p\rangle$ will be given by (5.73), with the limit $L \rightarrow \infty$. In this limit, it is easy to check that the normalization is

$$\langle p'|p\rangle = \delta(p' - p) \quad (5.114)$$

To see this, we need to evaluate the integral of $\langle p'|p\rangle$ over p' and check that it is equal to one. To see this, we use the conversion between summation and integration

$$\begin{aligned} \int_{-\infty}^{\infty} dp' \langle p'|p\rangle &= \lim_{L \rightarrow \infty} \frac{2\pi\hbar}{L} \sum_{n'} \langle p_{n'}|p\rangle \\ &= \lim_{L \rightarrow \infty} \frac{2\pi\hbar}{L} \sum_{n'} \frac{L}{2\pi\hbar} \delta_{n'n} \\ &= 1 \end{aligned} \quad (5.115)$$

where we have used (5.73). In the $L \rightarrow \infty$ limit, the relations between the position and momentum bases given by (5.74) become

$$|p\rangle = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} dx e^{ipx/\hbar} |x\rangle \quad (5.116)$$

and

$$|x\rangle = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} dp e^{-ipx/\hbar} |p\rangle \quad (5.117)$$

In this limit, the expansion of a state $|\psi\rangle$ in the momentum basis will be

$$|\psi\rangle = \int_{-\infty}^{\infty} dp \tilde{\psi}(p) |p\rangle \quad (5.118)$$

where the ‘momentum space wavefunction’ $\tilde{\psi}(p)$ is related to the position space wavefunction $\psi(x)$ as

$$\tilde{\psi}(p) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} dx \psi(x) e^{-ipx/\hbar} \quad (5.119)$$

which can be inverted to get

$$\psi(x) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} dp \tilde{\psi}(p) e^{ipx/\hbar} \quad (5.120)$$

These relations follow from (5.44) and (5.45), along with the scaling (5.68), followed by the limit $L \rightarrow \infty$. It should be noted that states $|x\rangle$ and $|p\rangle$ are *not* physical states, but convenient mathematical constructions, necessitated by the need to introduce a normalization consistent with the continuum limit. Given two states $|\psi\rangle$ and $|\phi\rangle$ with wavefunctions $\psi(x)$ and $\phi(x)$, their inner product can be evaluated in terms of these wavefunctions as follows

$$\begin{aligned} \langle\phi|\psi\rangle &= \langle\phi|\int_{-\infty}^{\infty} dx \psi(x)|x\rangle \\ &= \int_{-\infty}^{\infty} dx \psi(x) \langle\phi|x\rangle \\ &= \int_{-\infty}^{\infty} dx \phi^*(x) \psi(x) \end{aligned} \quad (5.121)$$

where $\langle\phi|x\rangle = \langle x|\phi\rangle^* = \phi^*(x)$.

The Hamiltonian operator is the same function of the momentum operator as the classical energy is of momentum

$$\hat{H} = \frac{1}{2m} \hat{p}^2 \quad (5.122)$$

The states of definite momentum are eigenstates of \hat{H}

$$\hat{H}|p\rangle = \frac{p^2}{2m}|p\rangle \quad (5.123)$$

As before, it is possible to construct states that have well-defined energy, but not momentum. Such a state will be of the form $|\psi\rangle = c_+|p\rangle + c_-|-p\rangle$

$$\hat{H}(c_+|p\rangle + c_-|-p\rangle) = \frac{p^2}{2m}(c_+|p\rangle + c_-|-p\rangle) \quad (5.124)$$

We now evaluate the matrix element $\langle x|\hat{H}|\psi\rangle$, which is useful. Let $|\phi\rangle = \hat{p}|\psi\rangle$. Then,

$$\begin{aligned} \langle x|\hat{H}|\psi\rangle &= \frac{1}{2m} \langle x|\hat{p}^2|\psi\rangle \\ &= \frac{1}{2m} \langle x|\hat{p}|\phi\rangle \\ &= \frac{1}{2m} \frac{\hbar}{i} \frac{d\phi(x)}{dx} \\ &= -\frac{\hbar^2}{2m} \frac{d^2\psi(x)}{dx^2} \end{aligned} \quad (5.125)$$

where we have used (5.97) and $\phi(x) = \langle x | \hat{p} | \psi \rangle$. We now write Schrodinger Equation (4.41) for a free particle in the form of a differential equation. We take the inner product of both sides of the equation with $|x\rangle$. Then,

$$i\hbar \langle x | \frac{d}{dt} |\psi(t)\rangle = \langle x | \hat{H} |\psi(t)\rangle \quad (5.126)$$

Remembering that

$$\frac{d}{dt} |\psi(t)\rangle = \lim_{dt \rightarrow 0} \frac{1}{dt} (|\psi(t+dt)\rangle - |\psi(t)\rangle) \quad (5.127)$$

it is easy to see that

$$\langle x | \frac{d}{dt} |\psi(t)\rangle = \frac{\partial}{\partial t} \langle x | \psi(t)\rangle \quad (5.128)$$

Given (5.125), we finally get

$$i\hbar \frac{\partial \psi(x, t)}{\partial t} = -\frac{\hbar^2}{2m} \frac{\partial^2 \psi(x, t)}{\partial x^2} \quad (5.129)$$

This equation allows us to determine the wavefunction at all instants of time, if we know the wavefunction at any one instant of time. Let us determine an integral solution to (5.129). Say, we are given the wavefunction at instant $t = t_0$ to be $\psi(x, t_0)$. We expand the state at $t = t_0$ in the momentum basis (which also happens to be the energy basis)

$$|\psi(t_0)\rangle = \int_{-\infty}^{\infty} dp \tilde{\psi}(p, t_0) |p\rangle \quad (5.130)$$

where $\tilde{\psi}(p, t_0)$ is related to $\psi(x, t_0)$ by (5.119). The state at instant t is obtained by operationg on $|\psi(t_0)\rangle$ by the time-evolution operator

$$\begin{aligned} |\psi(t)\rangle &= e^{-i\hat{H}(t-t_0)/\hbar} |\psi(t_0)\rangle \\ &= \int_{-\infty}^{\infty} dp \tilde{\psi}(p, t_0) e^{-i\hat{H}(t-t_0)/\hbar} |p\rangle \\ &= \int_{-\infty}^{\infty} dp \tilde{\psi}(p, t_0) e^{-ip^2(t-t_0)/2m\hbar} |p\rangle \end{aligned} \quad (5.131)$$

The wavefunction at t is found by taking the inner product of this equation with $|x\rangle$

$$\begin{aligned} \psi(x, t) &= \int_{-\infty}^{\infty} dp \tilde{\psi}(p, t_0) e^{-ip^2(t-t_0)/2m\hbar} \langle x | p \rangle \\ &= \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} dp \tilde{\psi}(p, t_0) e^{-ip^2(t-t_0)/2m\hbar} e^{ipx/\hbar} \end{aligned} \quad (5.132)$$

Reorganizing terms gives us the following form for the integral solution

$$\psi(x, t) = \int_{-\infty}^{\infty} dx' U(x, t; x', t_0) \psi(x', t_0) \quad (5.133)$$

where the function $U(x, t; x', t_0)$, called a *propagator*, is given by

$$U(x, t; x', t_0) = \frac{1}{2\pi\hbar} \int_{-\infty}^{\infty} dp e^{-ip^2(t-t_0)/2m\hbar} e^{ip(x-x')/\hbar} \quad (5.134)$$

The propagator tells us how the quantum particle propagates. To see this, we just need to recognize that

$$U(x, t; x', t_0) = \langle x | e^{-i\hat{H}(t-t_0)/\hbar} |x'\rangle \quad (5.135)$$

which essentially is the amplitude that the particle, located at point x' at instant t_0 , propagates to point x at instant t . To check this, we just need to expand $|x'\rangle$ in the momentum basis, operate by the time-evolution operator, and take the inner product with $|x\rangle$. The complex integral in (5.134) is evaluated as follows. First, we complete the square in p in the exponent to reduce the integral to the form

$$U(x, t; x', t_0) = \frac{1}{2\pi\hbar} e^{im(x-x')^2/2\hbar(t-t_0)} \int_{-\infty}^{\infty} dp e^{(-it/2m\hbar)(p-\alpha)^2} \quad (5.136)$$

where $\alpha = m(x - x')/(t - t_0)$. Shifting the variable of integration p by α and using the result

$$\int_{-\infty}^{\infty} dx e^{-ax^2} = \sqrt{\frac{\pi}{a}} \quad (5.137)$$

for complex a , we finally get

$$U(x, t; x', t_0) = \sqrt{\frac{m}{2\pi i\hbar(t-t_0)}} e^{im(x-x')^2/2\hbar(t-t_0)} \quad (5.138)$$

Let us now determine the expressions for expectation value and uncertainty of a linear operator in this infinite dimensional vector space. Given a linear operator \hat{A} , its expectation value and uncertainty (squared) in a state $|\psi\rangle$ are given by (4.58) and (4.61) respectively. Let us evaluate in particular expressions for expectation value and uncertainty in operators \hat{x} and \hat{p} in state $|\psi\rangle$. To calculate the expectation value of \hat{x} , we expand the state in the position basis $|x\rangle$

$$\begin{aligned} \langle \hat{x} \rangle &= \langle \psi | \hat{x} | \psi \rangle \\ &= \langle \psi | \hat{x} \int_{-\infty}^{\infty} dx \psi(x) |x\rangle \\ &= \int_{-\infty}^{\infty} dx \psi(x) \langle \psi | \hat{x} |x\rangle \\ &= \int_{-\infty}^{\infty} dx \psi^*(x) x \psi(x) \end{aligned} \quad (5.139)$$

It is easy to calculate the expectation value of any function $f(\hat{x})$ of the position operator. Proceeding as in (5.139)

$$\begin{aligned} \langle f(\hat{x}) \rangle &= \langle \psi | f(\hat{x}) | \psi \rangle \\ &= \langle \psi | f(\hat{x}) \int_{-\infty}^{\infty} dx \psi(x) |x\rangle \\ &= \int_{-\infty}^{\infty} dx \psi(x) \langle \psi | f(\hat{x}) |x\rangle \\ &= \int_{-\infty}^{\infty} dx \psi^*(x) f(x) \psi(x) \end{aligned} \quad (5.140)$$

where we have used $f(\hat{x})|x\rangle = f(x)|x\rangle$.

To calculate the expectation value of momentum, we expand $\hat{p}|\psi\rangle$ in the position basis, and use (5.97)

$$\begin{aligned} \langle \hat{p} \rangle &= \langle \psi | \hat{p} | \psi \rangle \\ &= \langle \psi | \int_{-\infty}^{\infty} dx |x\rangle \langle x | \hat{p} | \psi \rangle \\ &= \langle \psi | \int_{-\infty}^{\infty} dx \frac{\hbar}{i} \frac{d\psi(x)}{dx} |x\rangle \\ &= \int_{-\infty}^{\infty} dx \frac{\hbar}{i} \frac{d\psi(x)}{dx} \langle \psi | x \rangle \\ &= \int_{-\infty}^{\infty} dx \psi^*(x) \frac{\hbar}{i} \frac{d\psi(x)}{dx} \end{aligned} \quad (5.141)$$

Next, we calculate the expectation value of a function $g(\hat{p})$ of the momentum operator. It is useful to first calculate the expectation value of \hat{p}^n , where n is an integer. Using (5.97) recursively, we first get

$$\langle x | \hat{p}^n | \psi \rangle = \left(\frac{\hbar}{i} \frac{d}{dx} \right)^n \psi(x) \quad (5.142)$$

Given that a function $g(\hat{p})$ can be expanded as a series in \hat{p} , it is easy to generalize (5.142) to

$$\langle x | g(\hat{p}) | \psi \rangle = g \left(\frac{\hbar}{i} \frac{d}{dx} \right) \psi(x) \quad (5.143)$$

where $g \left(\frac{\hbar}{i} \frac{d}{dx} \right)$ is to be interpreted as a series in the differential operator $\left(\frac{\hbar}{i} \frac{d}{dx} \right)$. To evaluate the expectation value of $g(\hat{p})$ in state $|\psi\rangle$, we expand $g(\hat{p})|\psi\rangle$ in the position basis

$$\begin{aligned} \langle g(\hat{p}) \rangle &= \langle \psi | g(\hat{p}) | \psi \rangle \\ &= \langle \psi | \int_{-\infty}^{\infty} dx |x\rangle \langle x | g(\hat{p}) | \psi \rangle \\ &= \langle \psi | \int_{-\infty}^{\infty} dx g \left(\frac{\hbar}{i} \frac{d}{dx} \right) \psi(x) |x\rangle \\ &= \int_{-\infty}^{\infty} dx g \left(\frac{\hbar}{i} \frac{d}{dx} \right) \psi(x) \langle \psi | x \rangle \\ &= \int_{-\infty}^{\infty} dx \psi^*(x) g \left(\frac{\hbar}{i} \frac{d}{dx} \right) \psi(x) dx \end{aligned} \quad (5.144)$$

To evaluate the uncertainty in position and momentum, we just need to calculate the expectation values of \hat{x}^2 and \hat{p}^2 . These are easily found using (5.140) and (5.144). Then, the uncertainties (squared) are

$$\begin{aligned} \Delta \hat{x}^2 &= \langle \hat{x}^2 \rangle - \langle \hat{x} \rangle^2 \\ &= \int_{-\infty}^{\infty} dx \psi^*(x) x^2 \psi(x) - \left(\int_{-\infty}^{\infty} dx \psi^*(x) x \psi(x) \right)^2 \end{aligned} \quad (5.145)$$

and

$$\begin{aligned} \Delta \hat{p}^2 &= \langle \hat{p}^2 \rangle - \langle \hat{p} \rangle^2 \\ &= \int_{-\infty}^{\infty} dx \psi^*(x) \left(-\hbar^2 \frac{d^2}{dx^2} \right) \psi(x) - \left(\int_{-\infty}^{\infty} dx \psi^*(x) \frac{\hbar}{i} \frac{d}{dx} \psi(x) \right)^2 \end{aligned} \quad (5.146)$$

We now implement the ideas developed to a simple problem. Let the state of a particle at $t = 0$ be described by the wavefunction

$$\psi_0(x) = N e^{-(x-a)^2/2b^2} \quad (5.147)$$

where N is a normalization constant, determined by the constraint (5.111). The constraint integral for this wavefunction is a simple Gaussian integral of the form (5.137). Evaluation of the integral gives $N = 1/(\pi b^2)^{1/4}$. Then, the normalized wavefunction at $t = 0$ is

$$\psi_0(x) = \frac{1}{(\pi b^2)^{1/4}} e^{-(x-a)^2/2b^2} e^{ip_0 x/\hbar} \quad (5.148)$$

The associated position probability distribution is

$$|\psi_0(x)|^2 = \frac{1}{(\pi b^2)^{1/2}} e^{-(x-a)^2/b^2} \quad (5.149)$$

This is a Gaussian distribution, with characteristic width $\Delta x \sim b$ and peak at $x = a$

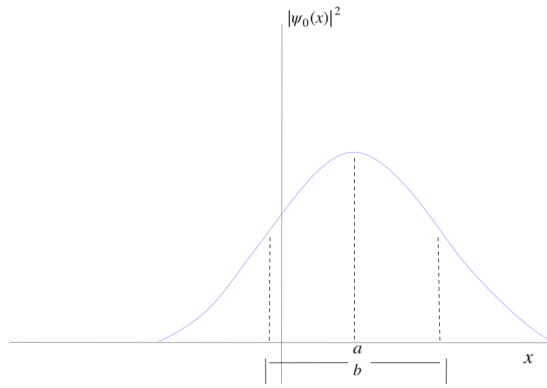


Figure 5.4: A Gaussian wavefunction with peak at $x = a$ and width $\Delta x = b$.

In this state, if the position of the particle is measured, the probability of measuring it in the neighborhood of $x = a$ will be the highest. Let us calculate the expectation value of \hat{x} and the uncertainty $\Delta\hat{x}$ in this state. From (5.139), it follows that the expectation value of \hat{x} is

$$\begin{aligned}\langle\hat{x}\rangle &= \int_{-\infty}^{\infty} dx |\psi_0(x)|^2 x \\ &= \frac{1}{(\pi b^2)^{1/2}} \int_{-\infty}^{\infty} dx e^{-(x-a)^2/b^2} x\end{aligned}\quad (5.150)$$

Changing the variable of integration from x to $y = (x - a)/b$ gives

$$\begin{aligned}\langle\hat{x}\rangle &= \frac{1}{(\pi b^2)^{1/2}} b \int_{-\infty}^{\infty} dy e^{-y^2} (by + a) \\ &= a\end{aligned}\quad (5.151)$$

where the integral over ye^{-y^2} is zero (this being an odd function), and we have used (5.137) for the other term. Then, the average value of position in this state is $x = a$, as we would expect from the probability distribution. To calculate the uncertainty in position, we need to evaluate the expectation value of \hat{x}^2

$$\begin{aligned}\langle\hat{x}^2\rangle &= \int_{-\infty}^{\infty} dx |\psi_0(x)|^2 x^2 \\ &= \frac{1}{(\pi b^2)^{1/2}} \int_{-\infty}^{\infty} dx e^{-(x-a)^2/b^2} x^2 \\ &= \frac{1}{(\pi b^2)^{1/2}} b \int_{-\infty}^{\infty} dy e^{-y^2} (a + by)^2 \\ &= a^2 + \frac{b^2}{2}\end{aligned}\quad (5.152)$$

where we have used

$$\begin{aligned}\int_{-\infty}^{\infty} dy e^{-y^2} y^2 &= \Gamma(3/2) \\ &= \frac{\sqrt{\pi}}{2}\end{aligned}\quad (5.153)$$

where

$$\Gamma(z) = 2 \int_0^{\infty} dy e^{-y^2} y^{2z-1}\quad (5.154)$$

is the ‘Gamma function’. The uncertainty in position is given by (5.145)

$$\begin{aligned}\Delta\hat{x}^2 &= \langle\hat{x}^2\rangle - \langle\hat{x}\rangle^2 \\ &= a^2 + \frac{b^2}{2} - a^2 \\ &= \frac{b^2}{2}\end{aligned}\tag{5.155}$$

Therefore

$$\begin{aligned}\Delta\hat{x} &= \sqrt{\Delta\hat{x}^2} \\ &= \frac{b}{\sqrt{2}}\end{aligned}\tag{5.156}$$

This means that if the position of the particle is measured, it will likely be detected in a region of size $\Delta x \sim b$ around the point $x = a$. Next, we measure the expectation value and uncertainty in momentum in this state. Notice that the phase $e^{ip_0x/\hbar}$ in the wavefunction (5.148) does not affect the probability of position measurement. However, it encodes crucial information regarding momentum measurement. The expectation value of momentum in this state is given by (5.141)

$$\begin{aligned}\langle\hat{p}\rangle &= \int_{-\infty}^{\infty} dx \psi_0^*(x) \frac{\hbar}{i} \frac{d\psi_0(x)}{dx} \\ &= \frac{1}{(\pi b^2)^{1/2}} \frac{\hbar}{i} \int_{-\infty}^{\infty} dx e^{-(x-a)^2/b^2} \left(-\frac{x-a}{b^2} + \frac{ip_0}{\hbar} \right) \\ &= p_0\end{aligned}\tag{5.157}$$

Then, the phase contains the information about the average momentum of the particle. The uncertainty in momentum is calculated using (5.146)

$$\begin{aligned}\Delta\hat{p}^2 &= \langle\hat{p}^2\rangle - \langle\hat{p}\rangle^2 \\ &= \frac{1}{(\pi b^2)^{1/2}} (-\hbar^2) \int_{-\infty}^{\infty} dx e^{-(x-a)^2/b^2} \left(\left(\frac{ip_0}{\hbar} - \frac{(x-a)}{b^2} \right)^2 - \frac{1}{b^2} \right) \\ &= \frac{\hbar^2}{2b^2}\end{aligned}\tag{5.158}$$

Therefore

$$\begin{aligned}\Delta\hat{p} &= \sqrt{\Delta\hat{p}^2} \\ &= \frac{\hbar}{\sqrt{2}b}\end{aligned}\tag{5.159}$$

The product of the position and momentum uncertainties is

$$\Delta\hat{x} \Delta\hat{p} = \frac{\hbar}{2}\tag{5.160}$$

which is consistent with (5.105). In fact, it turns out that the Gaussian wavefunction is the only wavefunction for which the inequality in (5.105) is replaced by an equality.

It is instructive to visualize the state (5.148) in momentum space. That is, we analyze the expansion of the state in the momentum basis. The momentum space wavefunction can be computed using (5.119)

$$\begin{aligned}\tilde{\psi}_0(p) &= \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} dx \psi_0(x) e^{-ipx/\hbar} \\ &= \frac{1}{\sqrt{2\pi\hbar}} \frac{1}{(\pi b^2)^{1/4}} \int_{-\infty}^{\infty} dx e^{-(x-a)^2/2b^2} e^{ip_0x/\hbar} e^{-ipx/\hbar} \\ &= \frac{1}{\sqrt{2\pi\hbar}} \frac{1}{(\pi b^2)^{1/4}} \int_{-\infty}^{\infty} dx e^{-(x-a)^2/2b^2} e^{-i(p-p_0)x/\hbar}\end{aligned}\tag{5.161}$$

The integral over x is evaluated by completing the square in the exponent, which will result in a Gaussian integral. The result is

$$\tilde{\psi}_0(p) = \left(\frac{b^2}{\pi \hbar^2} \right)^{1/4} e^{-b^2(p-p_0)^2/2\hbar^2} e^{-i(p-p_0)a/\hbar} \quad (5.162)$$

The form of the momentum-space wavefunction is strikingly similar to the position space wave-function (5.148). In fact, the momentum probability distribution is a Gaussian too

$$|\tilde{\psi}_0(p)|^2 = \left(\frac{b^2}{\pi \hbar^2} \right)^{1/2} e^{-b^2(p-p_0)^2/\hbar^2} \quad (5.163)$$

This Gaussian distribution is peaked around $p = p_0$, with characteristic width $\Delta p \sim \hbar/b$

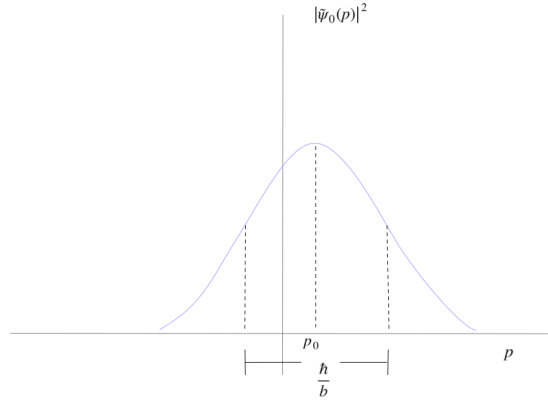


Figure 5.5: Momentum Space Gaussian wavefunction with peak at $p = p_0$ and width $\Delta p = \hbar/b$.

It can be seen that the analysis of average momentum and uncertainty in momentum is consistent with this distribution. Next, we consider the time-evolution of this state. For mathematical simplicity, we choose $a = 0$, so that the wavefunction at $t = 0$ is

$$\psi_0(x) = \frac{1}{(\pi b^2)^{1/4}} e^{-x^2/2b^2} e^{ip_0x/\hbar} \quad (5.164)$$

The wavefunction at instant t is given by (5.133)

$$\begin{aligned} \psi(x, t) &= \int_{-\infty}^{\infty} dx' U(x, t; x', 0) \psi_0(x') \\ &= \sqrt{\frac{m}{2\pi i \hbar t}} \frac{1}{(\pi b^2)^{1/4}} \int_{-\infty}^{\infty} dx' e^{im(x-x')^2/2\hbar t} e^{-x'^2/2b^2} e^{ip_0x'/\hbar} \end{aligned} \quad (5.165)$$

Once again, the integral can be solved by completing the square of x' in the exponent and evaluating a Gaussian integral. This gives

$$\psi(x, t) = \frac{1}{\sqrt{\pi^{1/2} (b + i\hbar t/m)}} e^{-(x-p_0t/m)^2/2b^2(1+i\hbar t/m)} e^{ip_0(x-p_0t/2m)/\hbar} \quad (5.166)$$

The probability distribution for this wavefunction is

$$|\psi(x, t)|^2 = \frac{1}{\pi^{1/2} (b^2 + \hbar^2 t^2/m^2 b^2)^{1/2}} e^{-(x-p_0t/m)^2/(b^2 + \hbar^2 t^2/m^2 b^2)} \quad (5.167)$$

which is Gaussian, peaked at $x = p_0t/m$ and with width $\Delta x \sim b\sqrt{1 + \hbar^2 t^2/m^2 b^4}$

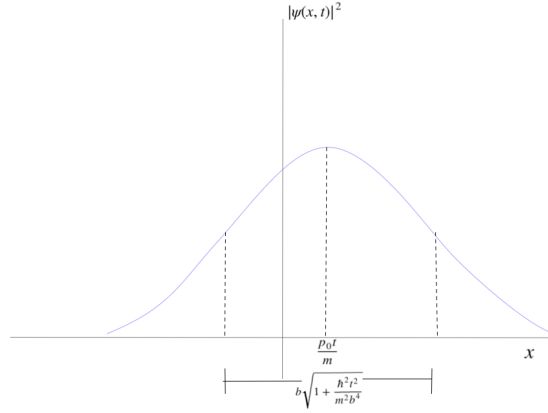


Figure 5.6: Evolution of a Gaussian Wavefunction. The probability distribution is Gaussian at every instant.

We now discuss propagation of a particle when it interacts with its environment, the interaction with the environment described by a potential energy function $V(x)$. We observe that the Hamiltonian for a free particle, given by (5.123), has the same form as the classical energy of the particle. We also saw that the Hamiltonian for spin interacting with an external magnetic field has the same form as the classical energy, with spin replaced by its Hermitian counterpart. This motivates us to postulate the following Hamiltonian for an interacting particle

$$\hat{H} = \frac{1}{2m}\hat{p}^2 + V(\hat{x}) \quad (5.168)$$

where $V(\hat{x})$ has the same form as the classical potential energy function, with coordinate x replaced by the Hermitian operator \hat{x} . Let us see if this Hamiltonian, through Ehrenfest's Theorem (4.94), reproduces the classical equations of motion. For this, we need to calculate the commutators of \hat{H} with \hat{x} and \hat{p} . We first calculate $[\hat{H}, \hat{x}]$. Observing that $[\hat{x}, V(\hat{x})] = 0$ and using property (4.102) of commutators and the position-momentum commutator (5.104), we get

$$\begin{aligned} [\hat{H}, \hat{x}] &= \frac{1}{2m} [\hat{p}^2, \hat{x}] \\ &= -\frac{1}{2m} (\hat{p} [\hat{x}, \hat{p}] + [\hat{x}, \hat{p}] \hat{p}) \\ &= \frac{-i\hbar}{m} \hat{p} \end{aligned} \quad (5.169)$$

To calculate $[\hat{H}, \hat{p}]$, we will need to calculate $[\hat{p}, V(\hat{x})]$. We first prove that

$$[\hat{p}, \hat{x}^n] = (-i\hbar)n\hat{x}^{n-1} \quad (5.170)$$

This is easily proved by induction. Let this be true for some integer n . Then,

$$\begin{aligned} [\hat{p}, \hat{x}^{n+1}] &= [\hat{p}, \hat{x} \hat{x}^n] \\ &= \hat{x} [\hat{p}, \hat{x}^n] + [\hat{p}, \hat{x}] \hat{x}^n \\ &= \hat{x}(-i\hbar)n\hat{x}^{n-1} - i\hbar \hat{I} \hat{x}^n \\ &= (-i\hbar)(n+1) \hat{x}^n \end{aligned}$$

which shows that it is true for $n+1$. Next, we visualize a series expansion for $V(\hat{x})$

$$V(\hat{x}) = \sum_n c_n \hat{x}^n \quad (5.171)$$

Then,

$$\begin{aligned}
 [\hat{p}, V(\hat{x})] &= \sum_n c_n [\hat{p}, \hat{x}^n] \\
 &= (-i\hbar) \sum_n n c_n \hat{x}^{n-1} \\
 &= (-i\hbar) V'(\hat{x})
 \end{aligned} \tag{5.172}$$

where $V'(\hat{x})$ is the derivative of the potential energy function with respect to x , with x replaced by \hat{x} . then, the commutator of \hat{H} and \hat{p} will be

$$\begin{aligned}
 [\hat{H}, \hat{p}] &= [V(\hat{x}), \hat{p}] \\
 &= (i\hbar) V'(\hat{x})
 \end{aligned} \tag{5.173}$$

Now, we use Ehrenfest's Theorem to calculate the rate of change of the expectation values of position and momentum operators. Let the instantaneous state of the particle be $|\psi(t)\rangle$. Then,

$$\begin{aligned}
 \frac{d}{dt} \langle \hat{x}(t) \rangle &= \frac{i}{\hbar} \langle \psi(t) | [\hat{H}, \hat{x}] | \psi(t) \rangle \\
 &= \frac{i}{\hbar} \frac{-i\hbar}{m} \langle \psi(t) | \hat{p} | \psi(t) \rangle \\
 &= \frac{1}{m} \langle \hat{p}(t) \rangle
 \end{aligned} \tag{5.174}$$

which is the same as the classical relation between velocity and momentum. Similarly,

$$\begin{aligned}
 \frac{d}{dt} \langle \hat{p}(t) \rangle &= \frac{i}{\hbar} \langle \psi(t) | [\hat{H}, \hat{p}] | \psi(t) \rangle \\
 &= \frac{i}{\hbar} (i\hbar) \langle \psi(t) | V'(\hat{x}) | \psi(t) \rangle \\
 &= -\langle V'(\hat{x}) \rangle
 \end{aligned} \tag{5.175}$$

which has the same form as Newton's Second Law.

Given the Hamiltonian (5.168), we write Schrodinger equation (4.41) as a differential equation for the wavefunction. As before, we take the inner product of both sides of (4.41) with state $|x\rangle$ to get

$$i\hbar \langle x | \frac{d}{dt} |\psi(t)\rangle = \langle x | \hat{H} |\psi(t)\rangle \tag{5.176}$$

where H is now the Hamiltonian (5.168). As before,

$$\langle x | \frac{d}{dt} |\psi(t)\rangle = \frac{\partial}{\partial t} \psi(x, t) \tag{5.177}$$

and

$$\frac{1}{2m} \langle x | \hat{p}^2 |\psi(t)\rangle = -\frac{\hbar^2}{2m} \frac{d^2 \psi(x, t)}{dx^2} \tag{5.178}$$

Further, since $V(\hat{x}) |x\rangle = V(x) |x\rangle$, therefore

$$\langle x | \hat{H} |\psi(t)\rangle = -\frac{\hbar^2}{2m} \frac{d^2 \psi(x, t)}{dx^2} + V(x) \psi(x, t) \tag{5.179}$$

Then, Schrodinger equation becomes

$$i\hbar \frac{\partial \psi(x, t)}{\partial t} = -\frac{\hbar^2}{2m} \frac{\partial^2 \psi(x, t)}{\partial x^2} + V(x) \psi(x, t) \tag{5.180}$$

Given the wavefunction at some instant t_0 , this equation determines it for all instants t uniquely.

5.3 Stationary states

Given the form of the Hamiltonian (5.168) for an interacting particle, we analyze states of definite energy, the so-called ‘stationary states’ (since they do not evolve with time). These states not only give information about possible values of the energy of the system, but also form a convenient basis for \mathcal{H} . Expanded in the energy basis, the time evolution of a given state is easy to determine. To recapitulate, let $\{|E_n\rangle\}$ be the complete set of eigenvectors of the Hamiltonian

$$\hat{H} |E_n\rangle = E_n |E_n\rangle \quad (5.181)$$

where though we have labelled these states with a discrete subscript n , some or all of these states could form a continuous set. For example, for a free particle, states of definite momentum are also states of definite energy, and the possible values of momentum form a continuous set. If there is an attractive interaction which would classically confine the particle, we will see that in general states of definite energy will partially form a discrete set (corresponding to energies for which the particle would be classically bound) and partially continuous set (corresponding to energies for which the particle’s motion is classically unbounded). When we write a summation over these states, it will be understood that part of the summation is in fact an integral over continuous states. We will assume that these states are orthonormal

$$\langle E_n | E_m \rangle = \delta_{nm} \quad (5.182)$$

Let the state of the particle at instant $t = t_0$ be $|\psi(t_0)\rangle$, which when expanded in the energy basis, has the expansion

$$|\psi(t_0)\rangle = \sum_n c_n(t_0) |E_n\rangle \quad (5.183)$$

where

$$c_n(t_0) = \langle E_n | \psi(t_0) \rangle \quad (5.184)$$

Then, the state of the particle at instant t will be given by (4.42)

$$\begin{aligned} |\psi(t)\rangle &= e^{-i\hat{H}(t-t_0)/\hbar} |\psi(t_0)\rangle \\ &= \sum_n c_n(t_0) e^{-i\hat{H}(t-t_0)/\hbar} |E_n\rangle \\ &= \sum_n c_n(t_0) e^{-iE_n(t-t_0)/\hbar} |E_n\rangle \end{aligned} \quad (5.185)$$

Say, we are given the wavefunction of the particle at $t = t_0$, and we wish to determine its wavefunction at instant t . We take the inner product of (5.185) with $|x\rangle$ to get

$$\begin{aligned} \psi(x, t) &= \sum_n c_n(t_0) e^{-iE_n(t-t_0)/\hbar} \langle x | E_n \rangle \\ &= \sum_n c_n(t_0) e^{-iE_n(t-t_0)/\hbar} \psi_n(x) \end{aligned} \quad (5.186)$$

where $\psi_n(x) = \langle x | E_n \rangle$ is the wavefunction for state $|E_n\rangle$. The information about the wavefunction at $t = t_0$ is encoded in the coefficients $c_n(t_0)$

$$\begin{aligned} c_n(t_0) &= \langle E_n | \psi(t_0) \rangle \\ &= \langle E_n | \int_{-\infty}^{\infty} dx \psi(x, t_0) |x\rangle \\ &= \int_{-\infty}^{\infty} dx \psi(x, t_0) \langle E_n | x \rangle \\ &= \int_{-\infty}^{\infty} dx \psi_n^*(x) \psi(x, t_0) \end{aligned} \quad (5.187)$$

where we have used $\langle E_n | x \rangle = \langle x | E_n \rangle^* = \psi_n^*(x)$. Then, to solve for the time-evolution of the wavefunction (5.186), we need the complete set of wavefunctions for the states of definite energy. To determine these wavefunctions (and associated energy eigenvalues), we need to express the eigenvalue equation (5.181) for \hat{H} in the form of a differential equation. Taking the inner product of (5.181) with $|x\rangle$ and using (5.179), we get

$$-\frac{\hbar^2}{2m} \frac{d^2\psi_n(x)}{dx^2} + V(x)\psi_n(x) = E_n\psi_n(x) \quad (5.188)$$

With appropriate boundary conditions, we can solve (5.188) to determine both E_n and $\psi_n(x)$.

Let us analyze the general behavior of energy eigenstates in presence of an interaction. A generic attractive interaction can be modelled by a potential energy function of the following form

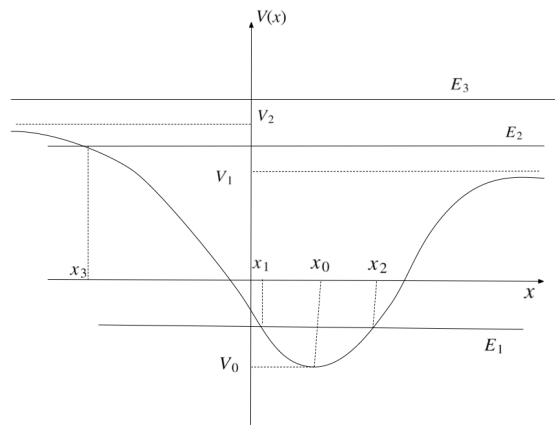


Figure 5.7: Generic potential energy function for an attractive interaction.

The potential energy function is attractive with minimum V_0 at $x = x_0$. It asymptotes to V_1 as $x \rightarrow \infty$ and V_2 for $x \rightarrow -\infty$ with $V_2 > V_1 > V_0$. The classical motion is bounded for energy in the range $V_0 \leq E < V_1$ and unbounded for $E > V_1$. For the bounded motion, there are two turning points x_1 and x_2 which are solutions to $V(x) = E$. For $V_1 < E < V_2$, the motion is unbounded for $x > x_3$, where x_3 is the sole turning point for this energy. If the particle has positive velocity, it will keep moving to ∞ , and if its velocity is negative, it will keep moving along the negative x direction, stop momentarily at $x = x_3$, and then proceed in the positive direction to ∞ . In other words, the particle will be scattered by the potential at $x = x_3$. For $E > V_2$, the motion is unbounded along either direction, and the particle will continue to move in the same direction without scattering. To solve for states of definite energy, we need to solve (5.188) with appropriate boundary conditions. In general, we will find that if we solve (5.188) with the boundary condition $\lim_{x \rightarrow \pm\infty} \psi_n(x) = 0$, we will get non-trivial solutions for E in the range $V_0 \leq E < V_1$. On the other hand, if we do not impose this condition, we will obtain non-trivial solutions for $E > V_1$. The boundary condition $\lim_{x \rightarrow \pm\infty} \psi_n(x) = 0$ physically implies that the probability of detecting the particle is non-zero in an effectively finite region of space. That is, the particle is localized. This is similar in some sense to classically bounded motion. It is therefore no coincidence that such solutions correspond to energy in the range $V_0 < E < V_1$, which is the range for classically bound motion. However, we will see that there are important differences. Firstly, the lowest possible energy will be found to be greater than, *not* equal to V_0 , the minimum value of the potential energy function. Secondly, the energy for such ‘bound states’ will form a discrete set. However, it will be seen that for $E > V_1$, the energy eigenvalues will form a continuous set, similar to the classical result. Such states are called ‘scattering states’

$$\begin{aligned} \lim_{x \rightarrow \pm\infty} \psi_n(x) &= 0 && \text{Bound States with discrete } E \\ \lim_{x \rightarrow \pm\infty} \psi_n(x) &\neq 0 && \text{Scattering States with continuous } E \end{aligned} \quad (5.189)$$

5.4 Bound states

We now analyse properties of bound states, and study a few interesting physical examples of quantum confinement by an attractive interaction. Given an attractive interaction of the generic form discussed above, let us first demonstrate that for bound states the lowest possible energy eigenvalue is greater than the least value of the potential energy. Multiplying (5.188) with $\psi_n^*(x)$ and integrating over all x , we get

$$\begin{aligned} -\frac{\hbar^2}{2m} \int_{-\infty}^{\infty} dx \psi_n^* \frac{d^2 \psi_n}{dx^2} + \int_{-\infty}^{\infty} dx V(x) |\psi_n|^2 &= E_n \int_{-\infty}^{\infty} dx |\psi_n|^2 \\ &= E_n \end{aligned} \quad (5.190)$$

where we have used the fact that ψ_n is normalized. Integrating the first term by parts, we get

$$\begin{aligned} -\frac{\hbar^2}{2m} \int_{-\infty}^{\infty} dx \psi_n^* \frac{d^2 \psi_n}{dx^2} &= -\frac{\hbar^2}{2m} \left[\psi_n^* \frac{d\psi_n}{dx} \right]_{-\infty}^{\infty} + \frac{\hbar^2}{2m} \int_{-\infty}^{\infty} dx \frac{d\psi_n^*}{dx} \frac{d\psi_n}{dx} \\ &= \frac{\hbar^2}{2m} \int_{-\infty}^{\infty} dx \left| \frac{d\psi_n}{dx} \right|^2 \\ &> 0 \end{aligned} \quad (5.191)$$

where we have used the boundary condition that ψ_n vanishes as $x \rightarrow \pm\infty$, and the fact that the integrand is positive definite (it cannot be zero, since it would imply a constant wavefunction, which would not be normalizable). If V_0 is the lowest value of the potential energy, it follows that

$$\begin{aligned} \int_{-\infty}^{\infty} dx V(x) |\psi_n|^2 &> V_0 \int_{-\infty}^{\infty} dx |\psi_n|^2 \\ &= V_0 \end{aligned} \quad (5.192)$$

Together with (5.190), these inequalities imply that $E > V_0$. This minimum energy, greater than V_0 , is often called ‘zero-point energy’. Let us now analyze the generic behavior of the energy eigenfunctions ψ_n for bound states. The following figure illustrates the wavefunction for a bound state with energy corresponding to classically bound motion between turning points x_1 and x_2

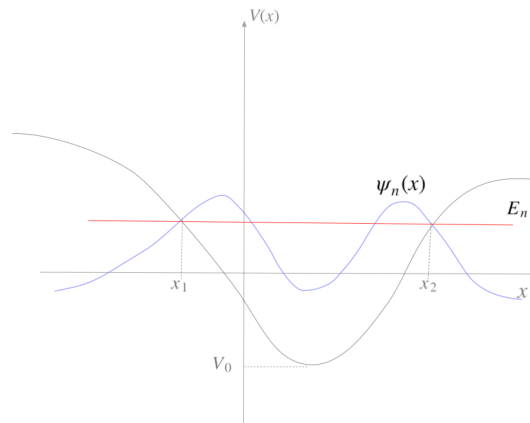


Figure 5.8: Wavefunction for a bound state. The wavefunction oscillates in the classically accessible region $x_1 < x < x_2$ and asymptotically falls to zero (approximately exponentially) in the classically forbidden region.

We can divide space into broadly two regions of interest: (a) Classically accessible region $x_1 \leq x \leq x_2$ in which the particle executes periodic, bound motion and (b) Classically ‘forbidden’ region $x < x_1$ and

$x > x_2$. In the classically accessible region, $E > V(x)$ for all x , and in the forbidden region, $V(x) > E$. This region is forbidden classically since the kinetic energy of the particle would be negative, which is not allowed. In the accessible region, we can write the eigenvalue equation as

$$\frac{d^2\psi_n}{dx^2} + k_n^2(x)\psi_n = 0 \quad (5.193)$$

where $k_n(x) = \sqrt{2m(E_n - V(x))}/\hbar$. If $V(x)$ is slowly varying (or the region of interest is small enough so that $V(x)$ varies very little), this equation has *locally* oscillatory solution of the form

$$\psi_n(x) \sim A \cos(k_n(x)x) + B \sin(k_n(x)x) \quad (5.194)$$

which oscillates with local ‘wavelength’ $\lambda_n(x) = 2\pi/k_n(x)$. Note that this local wavelength is shorter if the difference between E_n and $V(x)$ is larger. In the classically forbidden region, the eigenvalue equation can be cast in the form

$$\frac{d^2\psi_n}{dx^2} - \kappa_n^2(x)\psi_n = 0 \quad (5.195)$$

where $\kappa_n(x) = \sqrt{2m(V(x) - E_n)}/\hbar$. This has locally exponential solution of the form

$$\psi_n(x) \sim A' e^{\kappa_n(x)x} + B' e^{-\kappa_n(x)x} \quad (5.196)$$

since the wavefunction has to asymptotically vanish, for $x > x_2$, $A' = 0$ and the solution will have a local form

$$\psi_n(x) \sim B' e^{-\kappa_n(x)x} \quad x > x_2 \quad (5.197)$$

Similarly, for $x < x_1$, $B' = 0$ and the solution will have a local form

$$\psi_n(x) \sim A' e^{\kappa_n(x)x} \quad x < x_1 \quad (5.198)$$

As a result, the wavefunction oscillates with some local wavelength in the classically allowed region, and falls (locally) exponentially in the classically forbidden region. It should be noted that the wavefunction *does not* vanish in the classically forbidden region, but has an exponential ‘tail’ which ‘leaks’ into the region. Therefore, the probability of detecting the particle in the ‘forbidden’ region is non-zero. Classically, this would imply a negative kinetic energy. However, there is no paradox here, since a measurement of position of the particle collapses the state to a state of well-defined position, in which it does not have a well-defined energy. This behavior of the wavefunctions for bound states also allows us to conclude that the energy eigenvalues E_n will form a discrete set. The key here is that the exponential and oscillatory solutions have to *smoothly* match at the classical turning points. For an arbitrary value of E_n , this will *not* happen. Smooth solutions will be generated only for special values of energy

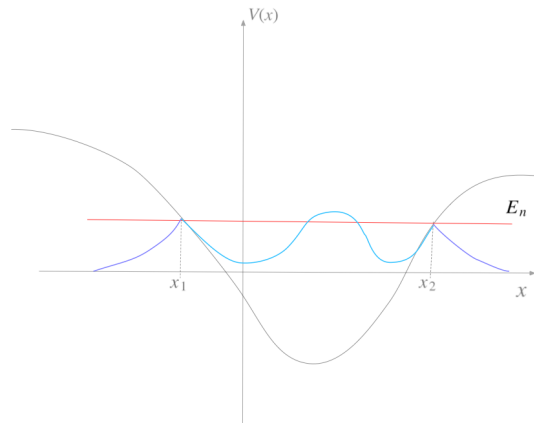


Figure 5.9: For arbitrary energy, the oscillating and exponential solutions do not match smoothly at the classical turning points.

We now prove a few important theorems concerning properties of bound state wavefunctions in one dimension. It should be noted that these theorems are valid only in one dimension.

Theorem 1: Wavefunctions for bound states in one dimension can always be chosen to be real.

Proof: The complex conjugate of the energy eigenvalue equation (5.188) is

$$-\frac{\hbar^2}{2m} \frac{d^2\psi_n^*(x)}{dx^2} + V(x)\psi_n^*(x) = E_n\psi_n^*(x) \quad (5.199)$$

where there is no complex conjugation of E_n , since it is real (eigenvalue of Hermitian operator \hat{H}). Multiplying (5.188) by ψ_n^* and (5.199) by ψ_n and subtracting, we get

$$\begin{aligned} \psi_n^* \frac{d^2\psi_n(x)}{dx^2} - \psi_n \frac{d^2\psi_n^*(x)}{dx^2} &= 0 \\ \Rightarrow \frac{d}{dx} \left(\psi_n^* \frac{d\psi_n}{dx} - \psi_n \frac{d\psi_n^*}{dx} \right) &= 0 \\ \Rightarrow \psi_n^* \frac{d\psi_n}{dx} - \psi_n \frac{d\psi_n^*}{dx} &= c_0 \end{aligned}$$

where c_0 is a constant. Since ψ_n and ψ_n^* vanish as $x \rightarrow \pm\infty$, c_0 must be zero. This gives

$$\begin{aligned} \psi_n^* \frac{d\psi_n}{dx} &= \psi_n \frac{d\psi_n^*}{dx} \\ \Rightarrow \int \frac{d\psi_n}{\psi_n} &= \int \frac{d\psi_n^*}{\psi_n^*} \\ \Rightarrow \psi_n^*(x) &= a \psi_n(x) \end{aligned}$$

where a is a constant of integration. Writing $\psi_n(x) = e^{i\delta(x)}\phi(x)$ where $\phi(x)$ is real, this gives $e^{2i\delta(x)} = a$, implying that δ is constant. Therefore, the wavefunction is $\psi_n(x) = e^{i\delta}\phi(x)$, which is physically the same as $\phi(x)$, which is real.

Theorem 2: In one dimension, there is no degeneracy in bound states. That is, there is a unique wavefunction corresponding to a given energy eigenvalue.

Proof: Let ψ_n and ϕ_n be two wavefunctions satisfying (5.188) for a given E_n . Then,

$$\begin{aligned} -\frac{\hbar^2}{2m} \frac{d^2\psi_n(x)}{dx^2} + V(x)\psi_n(x) &= E_n\psi_n(x) \\ &\text{and} \\ -\frac{\hbar^2}{2m} \frac{d^2\phi_n(x)}{dx^2} + V(x)\phi_n(x) &= E_n\phi_n(x) \end{aligned}$$

Multiplying the first equation with ϕ_n , the second with ψ_n and subtracting, we get

$$\begin{aligned} \phi_n \frac{d^2\psi_n(x)}{dx^2} - \psi_n \frac{d^2\phi_n(x)}{dx^2} &= 0 \\ \Rightarrow \frac{d}{dx} \left(\phi_n \frac{d\psi_n}{dx} - \psi_n \frac{d\phi_n}{dx} \right) &= 0 \\ \Rightarrow \phi_n \frac{d\psi_n}{dx} - \psi_n \frac{d\phi_n}{dx} &= c_0 \end{aligned}$$

where c_0 is a constant. Since both ψ_n and ϕ_n vanish asymptotically, this constant must be zero. Then, following the same steps as in Theorem 1, we get $\psi_n(x) = a \phi_n(x)$ where a is a constant. Then, the two wavefunctions are proportional to each other, and therefore physically the same wavefunction.

Theorem 3: In one dimension, wavefunctions for bound states corresponding to a potential energy function which is an even function of x are either even or odd functions of x .

Proof: Replacing x with $-x$ in (5.188) gives

$$\begin{aligned} -\frac{\hbar^2}{2m} \frac{d^2\psi_n(-x)}{dx^2} + V(-x)\psi_n(-x) &= E_n\psi_n(-x) \\ \Rightarrow -\frac{\hbar^2}{2m} \frac{d^2\psi_n(-x)}{dx^2} + V(x)\psi_n(-x) &= E_n\psi_n(-x) \end{aligned} \quad (5.200)$$

since $V(-x) = V(x)$. Multiplying (5.188) with $\psi_n(-x)$, (5.200) with $\psi_n(x)$ and subtracting, steps similar to ones in theorems 1 and 2 lead us to

$$\psi_n(-x) = a \psi_n(x) \quad (5.201)$$

Replacing x with $-x$ in this equation gives

$$\begin{aligned} \psi_n(x) &= a \psi_n(-x) \\ &= a^2 \psi_n(x) \end{aligned}$$

which implies that $a^2 = 1$ or $a = \pm 1$. Then,

$$\psi_n(-x) = \pm \psi_n(x) \quad (5.202)$$

which proves the theorem.

Let us illustrate these ideas using a simple ‘toy’ model for an attractive potential energy function. This is the so-called ‘Box’ potential energy function, given by

$$V(x) = \begin{cases} 0 & 0 < x < a \\ \infty & x < 0, x > a \end{cases} \quad (5.203)$$

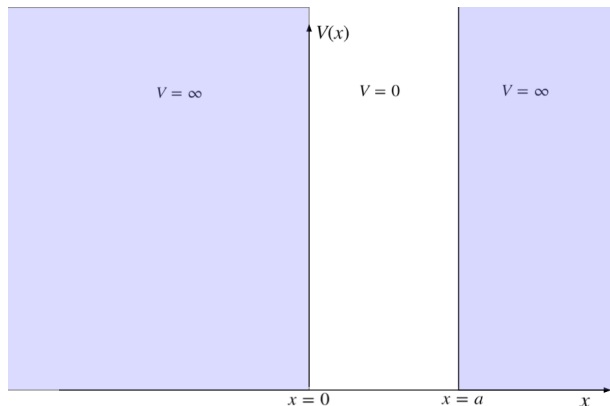


Figure 5.10: Particle experiencing a ‘Box’ interaction.

Let us solve (5.188) for bound states. First since the lowest value of the potential energy is zero, it follows that the energy eigenvalues will be positive (since they are always greater than the minimum potential energy for bound states). Given our intuition for the behavior of the energy eigenfunctions $\psi_E(x)$, for any finite energy E , the potential energy is infinitely greater than E in the regions $x < 0$ and $x > a$ (where $V = \infty$). Then, given that this is the classically forbidden region, the wavefunction will fall exponentially in these regions. In fact, this exponential fall will be infinitely steep, since V is infinite. Then, it follows that $\psi_E(x) = 0$ for $x < 0$ and $x > a$. The wavefunction will be non-zero only within the ‘box’ ($0 < x < a$). In this region, the wavefunction should satisfy the boundary conditions $\psi_E(0) = \psi_E(a) = 0$, since it should continuously match with the wavefunction in regions $x < 0$ and $x > a$ (where it is zero). Then, we need to solve (5.188) in the region $0 \leq x \leq a$ with $V(x) = 0$ and boundary conditions $\psi_E(0) = 0$ and $\psi_E(a) = 0$. The differential equation in this region is then

$$-\frac{\hbar^2}{2m} \frac{d^2\psi_E(x)}{dx^2} = E\psi_E(x) \quad 0 \leq x \leq a \quad (5.204)$$

which can be written as

$$\frac{d^2\psi_E(x)}{dx^2} = -k^2\psi_E(x) \quad (5.205)$$

where $k = \sqrt{2mE}/\hbar$. This has the general solution

$$\psi_E(x) = A \cos(kx) + B \sin(kx) \quad (5.206)$$

The boundary condition $\psi_E(0) = 0$ gives $A = 0$, and the condition $\psi_E(a) = 0$ gives

$$\begin{aligned} B \sin(ka) &= 0 \\ \Rightarrow k &= \frac{n\pi}{a}, \quad n = 0, \pm 1, \pm 2, \dots \end{aligned} \quad (5.207)$$

Then, k takes discrete values, resulting in discrete values of energy. It is easy to see that negative values of n are redundant, and need to be discarded. To see this, we note that for a given n , the energy eigenfunction has the form

$$\psi_n(x) = B_n \sin\left(\frac{n\pi x}{a}\right) \quad (5.208)$$

where B_n is a constant fixed by normalization. Under $n \rightarrow -n$, the wavefunction just changes sign, and therefore corresponds to the same physical state. Further, $n = 0$ needs to be discarded as well, since it gives a trivial solution $\psi_0(x) = 0$ for all x . The normalization constant B_n is fixed by condition (5.111), which since $\psi_n(x) = 0$ for $x < 0$ and $x > a$, reduces to

$$\int_0^a dx |\psi_n|^2 = 1 \quad (5.209)$$

Substituting the form of $\psi_n(x)$, this gives $B_n = \sqrt{2/a}$. Then, finally, the energy eigenfunctions are

$$\psi_n(x) = \begin{cases} \sqrt{\frac{2}{a}} \sin\left(\frac{n\pi x}{a}\right) & 0 \leq x \leq a \\ 0 & x < 0, x > a \end{cases} \quad n = 1, 2, 3, \dots \quad (5.210)$$

with energy eigenvalues

$$\begin{aligned} E_n &= \frac{\hbar^2 k_n^2}{2m} \\ &= \frac{n^2 \pi^2 \hbar^2}{2ma^2} \end{aligned} \quad (5.211)$$

The energy eigenvalues and eigenfunctions have all the properties that we proved in the theorems on bound states. The energy takes discrete values, the wavefunctions are real and there is no degeneracy (a given eigenvalue corresponds to a unique wavefunction). Further, the ground state energy is greater than zero, which agrees with the fact that the ground state energy for a bound state is greater than the lowest value of the potential energy. Also, the wavefunction has oscillatory behaviour in the classically allowed region ($0 < x < a$). Even though the ‘particle in a box’ model is a toy model, it reveals an interesting feature of quantum mechanics. The expression for energy eigenvalues (5.211) is inversely proportional to the square of the size of the region of confinement, a . If we decrease the size of the ‘box’ (confining the particle further), the ground state energy increases. This is in contrast with the classical situation in which there is no correlation between the size of the region of interaction and the energy of the particle. Whatever the value of a , the lowest value of energy would be zero. Following are the first three energy eigenfunctions. Notice that the number of oscillations (counted as the number of ‘nodes’ of the wavefunction) increases by one from one eigenstate to the next. This is a generic feature of bound states. If the position of the particle is measured in an energy eigenstate, then in the ground state, the probability of detecting the particle is maximum near the centre of the box. This behavior changes from one eigenfunction to the next. In the second excited state for instance, this probability is maximum at $x = a/4, a/2$ and $3a/4$. In the limit of very large n , it can be shown that the probability of detecting the particle is approximately the same at all points. This is in fact what would happen classically. Classically, if we were to randomly check the position of the particle, it would be equally likely to be found at any point. The important difference, though, is that classically the particle would be *there*, whereas quantum mechanics tells us that it will end up at a point as a result of a wavefunction collapse.

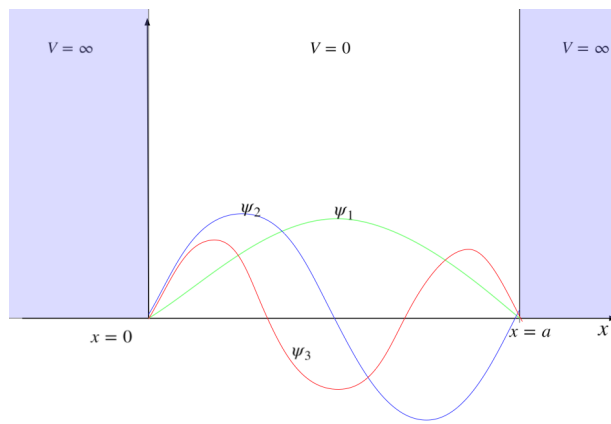


Figure 5.11: First three energy eigenstates for the box interaction.

We next look at a more physical example. Consider an attractive potential energy function in the form of a ‘square-well’

$$V(x) = \begin{cases} -V_0 & |x| < a/2 \\ 0 & |x| > a/2 \end{cases} \quad (5.212)$$

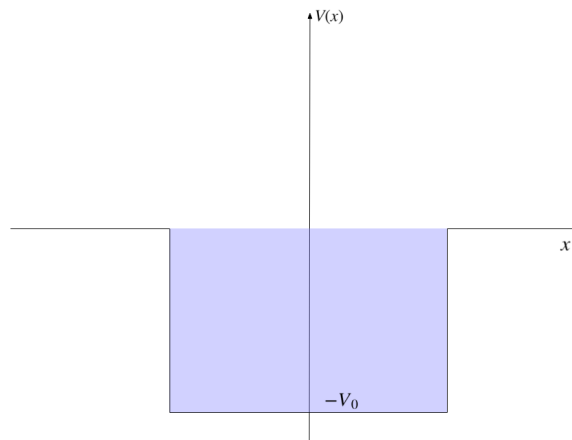


Figure 5.12: ‘Square-Well’ Potential energy function.

Bound states will correspond to energy $-V_0 < E < 0$ and the corresponding energy eigenfunctions will have oscillatory behaviour in the region $-a/2 < x < a/2$ (classically allowed region) and will be exponentially ‘damped’ in the region $|x| > a/2$ (classically forbidden). The choice of coordinate x is such that $V(-x) = V(x)$, so that the energy eigenfunctions will be even or odd functions of x . We divide space into three regions:

Region I: $x < -a/2$

In this region, $V(x) = 0$ and the eigenvalue equation reduces to

$$-\frac{\hbar^2}{2m} \frac{d^2\psi_I(x)}{dx^2} = E\psi_I(x) \quad (5.213)$$

where $\psi_I(x)$ is the eigenfunction $\psi_E(x)$ in the region $x < -a/2$. For bound states, $E = -|E|$, and this equation can be written as

$$\frac{d^2\psi_I(x)}{dx^2} = \kappa^2\psi_I(x) \quad (5.214)$$

where $\kappa = \sqrt{2m|E|}/\hbar$. This has the following solution, consistent with the boundary condition $\lim_{x \rightarrow -\infty} \psi_I = 0$

$$\psi_I(x) = Ae^{\kappa x} \quad (5.215)$$

The other independent solution $\psi = Be^{-\kappa x}$ is ruled out by the boundary condition.

Region II: $-a/2 < x < a/2$

In this region, $V(x) = -V_0$. Since $-V_0 < E < 0$, therefore $V_0 > |E|$. Then, the eigenvalue equation in this region has the form

$$\frac{d^2\psi_{II}(x)}{dx^2} = -k^2\psi_{II}(x) \quad (5.216)$$

where $k = \sqrt{2m(V_0 - |E|)}/\hbar$. This has the following solution

$$\psi_{II}(x) = C \cos(kx) + D \sin(kx) \quad (5.217)$$

Region III: $x > a/2$

In this region, $V(x) = 0$ and the eigenvalue equation reduces to

$$-\frac{\hbar^2}{2m} \frac{d^2\psi_{III}(x)}{dx^2} = E\psi_{III}(x) \quad (5.218)$$

where $\psi_{III}(x)$ is the eigenfunction $\psi_E(x)$ in the region $x > a/2$. As in Region I, this equation can be written as

$$\frac{d^2\psi_{III}(x)}{dx^2} = \kappa^2\psi_{III}(x) \quad (5.219)$$

with $\kappa = \sqrt{2m|E|}/\hbar$. In this region, the solution, consistent with the boundary condition $\lim_{x \rightarrow \infty} \psi_{III} = 0$ will be

$$\psi_I(x) = Be^{-\kappa x} \quad (5.220)$$

with other independent solution $\psi \sim e^{\kappa x}$ ruled out by the boundary condition. Since ψ_I , ψ_{II} and ψ_{III} are profiles of the same wavefunction $\psi_E(x)$ in the three regions the solutions should match at the boundaries of the regions in which they are defined. In particular, the wavefunction and its derivative should be continuous at the boundary. This imposes the following conditions

Boundary of Regions I and II:

$$\begin{aligned} \psi_I(-a/2) &= \psi_{II}(-a/2) \\ \psi'_I(-a/2) &= \psi'_{II}(-a/2) \end{aligned} \quad (5.221)$$

Boundary of Regions II and III:

$$\begin{aligned} \psi_{II}(a/2) &= \psi_I(a/2) \\ \psi'_{II}(a/2) &= \psi'_I(a/2) \end{aligned} \quad (5.222)$$

where prime ' refers to derivative with respect to x . Since the potential energy function is even, we can separately solve for even and odd eigenfunctions.

Even Functions:

These satisfy $\psi_E(-x) = +\psi_E(x)$. Then, in Region II, the coefficient of the sin term should be zero, since the sin function is odd. Further, the coefficients A (Region I) and B (Region III) should be equal, since these regions are reflections of each other about $x = 0$. Then, the even solutions $\psi_+(x)$ have the form

$$\psi_+(x) = \begin{cases} Ae^{\kappa x} & x < -a/2 \\ C \cos(kx) & |x| < a/2 \\ Ae^{-\kappa x} & x > a/2 \end{cases} \quad (5.223)$$

Boundary conditions (5.222) give the following equations

$$\begin{aligned} C \cos(ka/2) &= Ae^{-\kappa a/2} \\ Ck \sin(ka/2) &= A\kappa e^{-\kappa a/2} \end{aligned} \quad (5.224)$$

Dividing these equations gives the constraint on energy eigenvalues

$$k \tan(ka/2) = \kappa \quad (5.225)$$

Let us define dimensionless parameters $\alpha = ka/2$ and $\beta = \kappa a/2$. Given the form of k and κ , it can be easily checked that these parameters satisfy

$$\alpha^2 + \beta^2 = R^2; \quad R = \sqrt{(mV_0 a^2)/2\hbar^2} \quad (5.226)$$

It should be noted that by definition, $\alpha \geq 0$. Condition (5.231) can be expressed in terms of these parameters as

$$\alpha \tan \alpha = \beta \quad (5.227)$$

Simultaneous equations (5.226) and (5.227) will yield a discrete set of solution (α_n, β_n) . Given β_n , we can determine the associated energy eigenvalue through the relation between $\kappa = 2\beta/a$ and E

$$E_n = -\frac{2\beta_n^2 \hbar^2}{ma^2} \quad (5.228)$$

Points (α_n, β_n) can be found graphically as points of interections of the curves defined by equations (5.226) and (5.227). The curve representing equation (5.226) is a quarter of a circle with radius R , where R is a dimensionless parameter fixed by the nature of the potential (its width and depth). A ‘deeper’ and ‘wider’ potential well will lead to a larger R . Equation (5.227) is a transcendental equation

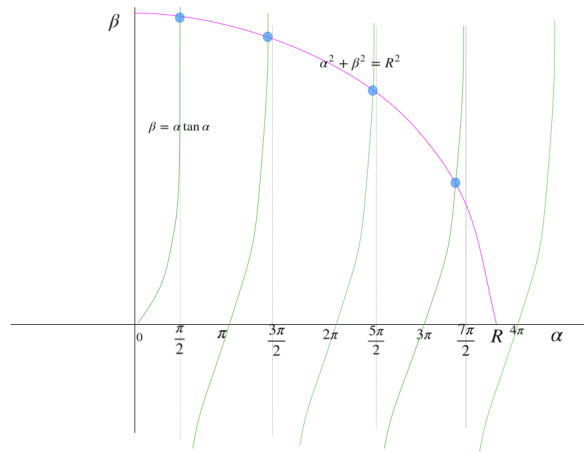


Figure 5.13: Even energy eigenvalues for a square-well potential.

Odd Functions:

These satisfy $\psi_E(-x) = -\psi_E(x)$. In Region II, the coefficient of the cos term should be zero, since the cos function is even. Further, since under $x \rightarrow -x$ the wavefunction should change sign, it implies that $\psi_{III}(x) = -\psi_I(-x)$ which gives $B = -A$. Then, the odd solutions $\psi_-(x)$ have the form

$$\psi_-(x) = \begin{cases} Ae^{\kappa x} & x < -a/2 \\ D \sin(kx) & |x| < a/2 \\ -Ae^{-\kappa x} & x > a/2 \end{cases} \quad (5.229)$$

Boundary conditions (5.222) give the following equations

$$\begin{aligned} D \sin(ka/2) &= -Ae^{-\kappa a/2} \\ Dk \cos(ka/2) &= A\kappa e^{-\kappa a/2} \end{aligned} \quad (5.230)$$

Dividing these equations gives the following constraint on energy eigenvalues

$$k \cot(ka/2) = -\kappa \quad (5.231)$$

Again, defining dimensionless parameters $\alpha = ka/2$ and $\beta = \kappa a/2$ gives, along with (5.226), the constraint equation

$$\alpha \cot \alpha = -\beta \quad (5.232)$$

The simultaneous equations (5.226) and (5.232) will once again yield a discrete set of solutions (α_n, β_n) corresponding to allowed energy states (with odd wavefunctions). As before, given β_n , we can determine the associated energy eigenvalue through (5.228). These can be found as points of intersections of the curves defined by equations (5.226) and (5.232). Equation (5.232) is again a transcendental equation

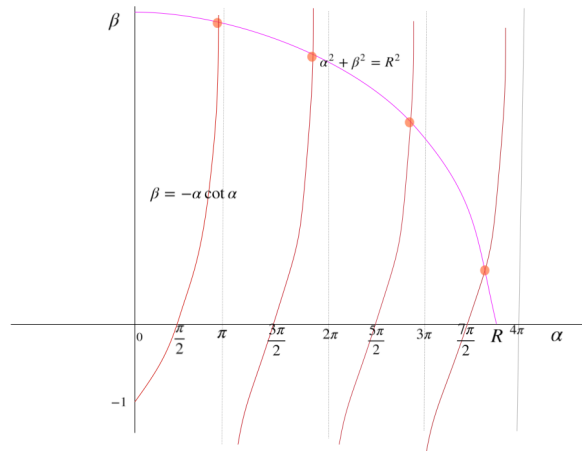


Figure 5.14: Odd energy eigenvalues for a square-well potential.

Let us now illustrate both sets of bound states on the same graph

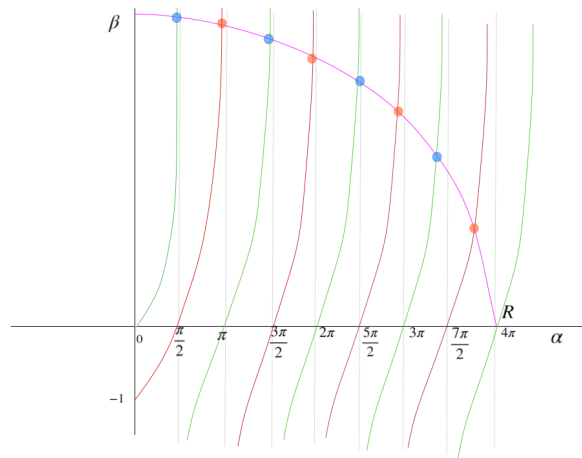


Figure 5.15: Complete set of bound state energy eigenvalues for a give potential width and depth. The even and odd states alternate with increasing energy. The ground state is an even function.

Since $\alpha = ka/2 = \left(\sqrt{2m(V_0 - |E|)} a\right) / 2\hbar$, increasing energy implies an increasing α (note that since $E < 0$, increasing energy implies decreasing $|E|$). Then, in order of increasing energy, we alternately encounter even and odd states. The ground state is an even state, the first excited state is an odd state, the second excited state an even state, and so on. It is important to note that there will always be a *finite* number of bound states, the number depending on the points of interection in the $\alpha - \beta$ plane, which in turn will depend on the dimensionless parameter R (the ‘radius’). Say, the width and depth of the

potential well are such that R is in the range $N\pi < R < (N + 1/2)\pi$. Then, there will be exactly $N + 1$ even and N odd bound states. It should be noted that no matter how small the value of R , there will always be at least one bound state, which will be an even state.

5.5 The quantum Harmonic Oscillator

We now analyze one of the most important quantum systems, a quantum harmonic oscillator. Mathematically, it can be visualised as a particle under the influence of a potential energy of the form $V(x) = (1/2)kx^2$ where k is some parameter. The classical motion of the particle under this influence is simple harmonic, with angular frequency $\omega = \sqrt{k/m}$ where m is the mass of the particle. Then, the harmonic oscillator potential energy function can be written as

$$V(x) = \frac{1}{2}m\omega^2x^2 \quad (5.233)$$

Even if potential energy function has a more complex form, as long as it has a minimum, the behavior of the system in the neighborhood of the minimum can be approximated by such a potential energy function. Then, if the potential energy $V(x)$ has a minimum at $x = x_0$, a Taylor expansion about this point gives

$$V(x) \simeq V(x_0) + \frac{1}{2}V''(x_0)(x - x_0)^2 + \dots \quad (5.234)$$

If we ignore the constant term $V(x_0)$ (which is not physically observable) and retain only the quadratic term in x , we are approximating the potential energy function by the form (5.233). The classical energy of the particle is then of the form

$$E = \frac{p^2}{2m} + \frac{1}{2}m\omega^2x^2 \quad (5.235)$$

where the origin of coordinates is chosen so that the equilibrium point has coordinate $x_0 = 0$. The quantum Hamiltonian for this system will then be

$$\hat{H} = \frac{1}{2m}\hat{p}^2 + \frac{1}{2}m\omega^2\hat{x}^2 \quad (5.236)$$

At this point, we could solve for states of definite energy, and analyze their properties. However, it turns out that the scope of this problem is much larger, and the ‘quantum harmonic oscillator’ is encountered in Physics in several different contexts, be it high energy particle physics, low energy condensed matter physics or quantum computers. Therefore, it is important to analyze the quantum mechanics of this system from a more general, abstract point of view. To see this problem emerge in a different context (of potential use in quantum computers), let us consider an electrical circuit consisting of an inductor and a capacitor

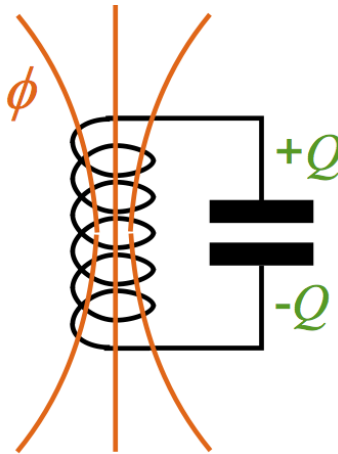


Figure 5.16: An LC circuit.

Should we apply laws of quantum mechanics to this circuit? Such circuits are macroscopic systems, consisting of an enormous number of atoms, so even if we were to somehow ‘quantize’ such a system, quantum effects would likely be wiped out by *decoherence* (environmental ‘noise’). However, if we make

the circuit sufficiently microscopic and also eliminate sources of decoherence (such as resistance) as much as possible, it the *do* show quantum behavior, as recent experiments suggest. In fact, they have the potential of being used as ‘qubits’ in quantum computers of the future. What does it mean to ‘quantize’ such a system? We start from first principles and identify classical physical observables, and represent them as Hermitian operators. These operators will satisfy some algebra, which will allow us to construct quantum states (quantum ‘statics’). The dynamics of the states will be generated by the quantum Hamiltonian, which is the Hermitian operator corresponding to the classical energy of the system. Clearly, the classical ‘observables’ for this circuit are the charge Q stored on the capacitor and the magnetic flux Φ stored in the inductor. The classical energy of the LC circuit in terms of these observables is

$$E = \frac{Q^2}{2C} + \frac{\Phi^2}{2L} \quad (5.237)$$

where C is the capacitance, and L the inductance. If an initial charge is applied to the capacitor, the instantaneous charge $Q(t)$ and current $I(t)$ will satisfy the following equations

$$\begin{aligned} \frac{dQ}{dt} &= I \\ L \frac{dI}{dt} &= -\frac{Q}{C} \end{aligned} \quad (5.238)$$

The current I can be expressed in terms of the magnetic flux Φ through the relation $\Phi = LI$. Then, the classical observables Q and Φ satisfy the equations

$$\begin{aligned} \frac{dQ}{dt} &= \frac{\Phi}{L} \\ \frac{d\Phi}{dt} &= -\frac{Q}{C} \end{aligned} \quad (5.239)$$

When we quantize the system, observables Q and Φ are represented by Hermitian operators \hat{Q} and $\hat{\Phi}$. These operators will satisfy some commutator algebra which will lead to a description of the quantum states of the system. The Hamiltonian operator for the system is

$$\hat{H} = \frac{\hat{Q}^2}{2C} + \frac{\hat{\Phi}^2}{2L} \quad (5.240)$$

To determine the commutator algebra of \hat{Q} and $\hat{\Phi}$, we invoke Ehrenfest’s Theorem (4.94) to calculate the rate of change of expectation values of \hat{Q} and $\hat{\Phi}$, and demand that the form of the equations should be the same as the classical equations (5.239). Ehrenfest’s equations for this system will be

$$\begin{aligned} \frac{d}{dt} \langle \hat{Q}(t) \rangle &= \frac{i}{\hbar} \langle \psi(t) | [\hat{H}, \hat{Q}] | \psi(t) \rangle \\ &= \frac{i}{\hbar} \frac{1}{2L} \langle \psi(t) | [\hat{\Phi}^2, \hat{Q}] | \psi(t) \rangle \\ \frac{d}{dt} \langle \hat{\Phi}(t) \rangle &= \frac{i}{\hbar} \langle \psi(t) | [\hat{H}, \hat{\Phi}] | \psi(t) \rangle \\ &= \frac{i}{\hbar} \frac{1}{2C} \langle \psi(t) | [\hat{Q}^2, \hat{\Phi}] | \psi(t) \rangle \end{aligned} \quad (5.241)$$

If these expectation values are to change with time the same way as the classical observables, these equations should reduce to

$$\begin{aligned} \frac{d}{dt} \langle \hat{Q}(t) \rangle &= \frac{1}{L} \langle \psi(t) | \hat{\Phi} | \psi(t) \rangle \\ \frac{d}{dt} \langle \hat{\Phi}(t) \rangle &= -\frac{1}{C} \langle \psi(t) | \hat{Q} | \psi(t) \rangle \end{aligned} \quad (5.242)$$

It is easy to check that the following prescription will ensure this

$$[\hat{\Phi}, \hat{Q}] = i\hbar\hat{I} \quad (5.243)$$

which had the same form as the position-momentum commutator for a particle! In fact, the Hamiltonian (5.240) has the same form as the harmonic oscillator Hamiltonian (5.236) with the following identifications

$$\begin{aligned} \hat{x} &\leftrightarrow \hat{\Phi} \\ \hat{p} &\leftrightarrow \hat{Q} \\ m &\leftrightarrow C \\ k &\leftrightarrow 1/L \end{aligned} \quad (5.244)$$

where k is the ‘stiffness constant’ for an imaginary spring visualised as a source of the potential energy of the harmonic oscillator.

We now proceed to construct energy eigenstates for the harmonic oscillator Hamiltonian (5.236). Instead of solving the energy eigenvalue equation in the form of a differential equation, we will use a powerful operator technique. First, we identify natural length and energy scales in the problem. The classical frequency ω and Planck’s constant induce a natural energy scale $\epsilon_0 = \hbar\omega$. Dividing the classical expression for energy (5.235) by ϵ_0 gives

$$\frac{E}{\epsilon_0} = \frac{p^2}{2m\omega\hbar} + \frac{m\omega x^2}{2\hbar} \quad (5.245)$$

Since each term in this equation is dimensionless, we see that there is a natural length scale $l = \sqrt{2\hbar/m\omega}$ (which is not surprising, since Planck’s constant naturally relates length L and energy E through $E \sim \hbar^2/mL^2$). We can factorise equation (5.5) as follows

$$\frac{E}{\epsilon_0} = \left(\sqrt{\frac{m\omega}{2\hbar}} x + i \frac{1}{\sqrt{2m\omega\hbar}} p \right) \left(\sqrt{\frac{m\omega}{2\hbar}} x - i \frac{1}{\sqrt{2m\omega\hbar}} p \right) \quad (5.246)$$

If we define a new complex variable

$$a = \sqrt{\frac{m\omega}{2\hbar}} x + i \frac{1}{\sqrt{2m\omega\hbar}} p \quad (5.247)$$

then we can write the expression for energy as

$$E = \epsilon_0 a^* a \quad (5.248)$$

The energy then exclusively depends on the variable a (and its complex conjugate). Exploiting this, we define an operator \hat{a} which is the same function of the position and momentum operators as the complex number a

$$\hat{a} = \sqrt{\frac{m\omega}{2\hbar}} \hat{x} + i \frac{1}{\sqrt{2m\omega\hbar}} \hat{p} \quad (5.249)$$

This operator is *not* Hermitian. Since \hat{x} and \hat{p} are Hermitian, the Hermitian adjoint of \hat{a} will be

$$\hat{a}^\dagger = \sqrt{\frac{m\omega}{2\hbar}} \hat{x} - i \frac{1}{\sqrt{2m\omega\hbar}} \hat{p} \quad (5.250)$$

Given the commutator of \hat{x} and \hat{p} , the commutator of \hat{a} and \hat{a}^\dagger is easily computed to be

$$[\hat{a}, \hat{a}^\dagger] = \hat{I} \quad (5.251)$$

The operators \hat{a} and \hat{a}^\dagger are called *annihilation* and *creation* operators respectively, for reasons which will become clear. The expression for classical energy (5.248) suggests that the Hamiltonian must be expressible in terms of \hat{a} and \hat{a}^\dagger . Given (5.249) and (5.250), let us calculate the product $\hat{a}^\dagger \hat{a}$. We get

$$\begin{aligned} \hat{a}^\dagger \hat{a} &= \frac{m\omega}{2\hbar} \hat{x}^2 + \frac{1}{2m\omega\hbar} \hat{p}^2 + \frac{i}{2\hbar} [\hat{x}, \hat{p}] \\ &= \frac{1}{\hbar\omega} \hat{H} - \frac{1}{2} \hat{I} \end{aligned} \quad (5.252)$$

where we have used the form of \hat{H} (5.236) and the position-momentum commutator. Then, the Hamiltonian can be expressed in terms of the creation and annihilation operators as follows

$$\hat{H} = \left(\hat{a}^\dagger \hat{a} + \frac{1}{2} \hat{I} \right) \hbar\omega \quad (5.253)$$

We now evaluate the commutators of \hat{a} and \hat{a}^\dagger with \hat{H} . Give (5.251), we get

$$\begin{aligned} [\hat{H}, \hat{a}] &= \hbar\omega [\hat{a}^\dagger \hat{a}, \hat{a}] \\ &= \hbar\omega [\hat{a}^\dagger, \hat{a}] \hat{a} \\ &= -\hbar\omega \hat{a} \end{aligned} \quad (5.254)$$

and

$$\begin{aligned} [\hat{H}, \hat{a}^\dagger] &= \hbar\omega [\hat{a}^\dagger \hat{a}, \hat{a}^\dagger] \\ &= \hbar\omega \hat{a}^\dagger [\hat{a}, \hat{a}^\dagger] \\ &= \hbar\omega \hat{a}^\dagger \end{aligned} \quad (5.255)$$

Given these, we can interpret the action of these operators on an energy eigenstate. Let $|E\rangle$ be an energy eigenstate

$$\hat{H} |E\rangle = E |E\rangle \quad (5.256)$$

Let us consider the state $\hat{a} |E\rangle$. Operating on this state with \hat{H} gives

$$\begin{aligned} \hat{H}(\hat{a} |E\rangle) &= ([\hat{H}, \hat{a}] + \hat{a}\hat{H}) |E\rangle \\ &= -\hbar\omega \hat{a} |E\rangle + E \hat{a} |E\rangle \\ &= (E - \hbar\omega) \hat{a} |E\rangle \end{aligned} \quad (5.257)$$

Therefore, the state $\hat{a} |E\rangle$ is an eigenstate of \hat{H} with eigenvalue $E - \hbar\omega$

$$\hat{a} |E\rangle = c_- |E - \hbar\omega\rangle \quad (5.258)$$

where c_- is a proportionality constant. Similarly,

$$\begin{aligned} \hat{H}(\hat{a}^\dagger |E\rangle) &= ([\hat{H}, \hat{a}^\dagger] + \hat{a}^\dagger \hat{H}) |E\rangle \\ &= \hbar\omega \hat{a}^\dagger |E\rangle + E \hat{a}^\dagger |E\rangle \\ &= (E + \hbar\omega) \hat{a}^\dagger |E\rangle \end{aligned} \quad (5.259)$$

Therefore, the state $\hat{a}^\dagger |E\rangle$ is an eigenstate of \hat{H} with eigenvalue $E + \hbar\omega$

$$\hat{a}^\dagger |E\rangle = c_+ |E + \hbar\omega\rangle \quad (5.260)$$

where c_+ is another proportionality constant. Clearly, starting with state $|E\rangle$, we can create a ladder of energy eigenstates by successively acting on this state by \hat{a} or \hat{a}^\dagger . Every action of \hat{a} decreases the energy by $\hbar\omega$, and action of \hat{a}^\dagger increases the energy by $\hbar\omega$. Since the lowest energy eigenvalue has to be greater than zero (the lowest value of the potential energy function is zero), therefore starting with a state $|E\rangle$, we cannot move arbitrarily low down this ‘ladder’. We have to eventually reach a state $|E_0\rangle$ such that the action of \hat{a} on this state will not produce any new state of lower energy. This will happen only if the action of \hat{a} on this state gives a null vector

$$\hat{a} |E_0\rangle = | \rangle_0 \quad (5.261)$$

This uniquely determines the properties of this state. To see this, let us operate by \hat{a}^\dagger on this state. Using (5.253), we get

$$\begin{aligned} \hat{a}^\dagger \hat{a} |E_0\rangle &= 0 \\ \Rightarrow \left(\frac{1}{\hbar\omega} \hat{H} - \frac{1}{2} \hat{I} \right) |E_0\rangle &= 0 \\ \Rightarrow \left(\frac{E_0}{\hbar\omega} - \frac{1}{2} \right) |E_0\rangle &= 0 \end{aligned} \quad (5.262)$$

Since $|E_0\rangle$ is not a null vector, it follows that

$$E_0 = \frac{\hbar\omega}{2} \quad (5.263)$$

This is then the ground state energy, and the state $|E_0\rangle$ is the ground state. This state is unique, since there is no degeneracy in energy in one dimension (for bound states). The other states can be generated by successive action of \hat{a}^\dagger on this state. Every action of \hat{a}^\dagger gives an energy eigenstate with eigenvalue increased by $\hbar\omega$. Then, acting n times with \hat{a}^\dagger will give a state with eigenvalue $E_n = E_0 + n\hbar\omega = (n + 1/2)\hbar\omega$. Then, the possible energy eigenvalues are of the form

$$E_n = \left(n + \frac{1}{2}\right) \hbar\omega \quad n = 0, 1, 2, 3, \dots \quad (5.264)$$

Note that the energy eigenvalues have a constant spacing, equal to $\hbar\omega$

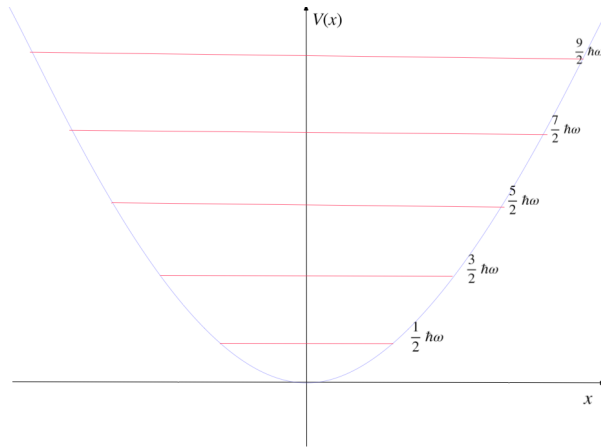


Figure 5.17: Energy eigenvalues for a harmonic oscillator. Note that the eigenvalue spacing is constant.

To see that these are the only energy eigenstates (and that we have not missed any), let us assume there is a state $|E\rangle$ such that $E_n < E < E_{n+1}$ for some n . Then, if we operate successively with \hat{a} on this state, we will again come to a state such that the action of \hat{a} on it will give a null vector. But, we have seen that such a state has the unique eigenvalue $E_0 = \hbar\omega/2$. Therefore, since every action of \hat{a} decreases the energy by $\hbar\omega$, the state that we started with must have had an eigenvalue which would be some odd half-integer times $\hbar\omega$. But then, it must have been equal to $(n + 1/2)\hbar\omega$ for some n , and so one of the set (5.264). We assume that this ground state is normalized, so that $\langle E_0 | E_0 \rangle = 1$. Let us generate the excited, normalized states, by acting successively on $|E_0\rangle$ by \hat{a}^\dagger . First, we observe that the action of \hat{a}^\dagger on state $|E_n\rangle$ gives a state proportional to $|E_{n+1}\rangle$, which is the same as $|E_{n+1}\rangle$. Similarly, the action of \hat{a} on $|E_n\rangle$ gives a state proportional to $|E_{n-1}\rangle$. Then,

$$\hat{a}^\dagger |E_n\rangle = c_+(n) |E_{n+1}\rangle \quad (5.265)$$

and

$$\hat{a} |E_n\rangle = c_-(n) |E_{n-1}\rangle \quad (5.266)$$

where $c_+(n)$ and $c_-(n)$ are complex constants. To determine these constants, we assume that states $|E_n\rangle$ are normalized such that $\langle E_n | E_n \rangle = 1$. Then, taking the inner product of the second equation with itself gives

$$\begin{aligned} |c_-(n)|^2 &= \langle E_n | \hat{a}^\dagger \hat{a} | E_n \rangle \\ &= \langle E_n | \left(\frac{1}{\hbar\omega} \hat{H} - \frac{1}{2} \hat{I} \right) | E_n \rangle \\ &= \langle E_n | \left(n + \frac{1}{2} - \frac{1}{2} \right) | E_n \rangle \\ &= n \end{aligned} \quad (5.267)$$

which gives $c_-(n) = \sqrt{n}$, upto an overall phase. Similarly, taking the inner product of the first equation with itself gives

$$\begin{aligned}
|c_+(n)|^2 &= \langle E_n | \hat{a} \hat{a}^\dagger | E_n \rangle \\
&= \langle E_n | \left([\hat{a}, \hat{a}^\dagger] + \hat{a}^\dagger \hat{a} \right) | E_n \rangle \\
&= \langle E_n | \left(\hat{I} + \frac{1}{\hbar\omega} \hat{H} - \frac{1}{2} \hat{I} \right) | E_n \rangle \\
&= \langle E_n | (n+1) | E_n \rangle \\
&= n+1
\end{aligned} \tag{5.268}$$

which gives $c_+(n) = \sqrt{n+1}$, upto an overall phase. Summarising, we get

$$\begin{aligned}
\hat{a}^\dagger | E_n \rangle &= \sqrt{n+1} | E_{n+1} \rangle \\
\hat{a} | E_n \rangle &= \sqrt{n} | E_{n-1} \rangle
\end{aligned} \tag{5.269}$$

To check consistency, let us verify that the eigenvalue of \hat{H} in state $| E_n \rangle$ is given by (5.264). Acting on $| E_n \rangle$ by \hat{H} gives

$$\begin{aligned}
\hat{H} | E_n \rangle &= \hbar\omega \left(\hat{a}^\dagger \hat{a} + \frac{1}{2} \hat{I} \right) | E_n \rangle \\
&= \hbar\omega \left(\hat{a}^\dagger \hat{a} | E_n \rangle + \frac{1}{2} | E_n \rangle \right) \\
&= \hbar\omega \left(\sqrt{n} \hat{a}^\dagger | E_{n-1} \rangle + \frac{1}{2} | E_n \rangle \right) \\
&= \hbar\omega \left(\sqrt{n} \sqrt{n} | E_n \rangle + \frac{1}{2} | E_n \rangle \right) \\
&= \left(n + \frac{1}{2} \right) \hbar\omega | E_n \rangle
\end{aligned}$$

which agrees with (5.264).

We now calculate the wavefunctions for these states. We start by calculating the wavefunction for the ground state $| E_0 \rangle$, given by the condition (5.261). Given the form of the annihilation operator in (5.249), condition (5.261) gives

$$\left(\sqrt{\frac{m\omega}{2\hbar}} \hat{x} + i \frac{1}{\sqrt{2m\omega\hbar}} \hat{p} \right) | E_0 \rangle = | \rangle \tag{5.270}$$

Taking the inner product with $| x \rangle$ and using (5.97), we get a differential equation for the ground state wavefunction $\psi_0(x) = \langle x | E_0 \rangle$

$$\begin{aligned}
\sqrt{\frac{m\omega}{2\hbar}} x \psi_0(x) + i \frac{1}{\sqrt{2m\omega\hbar}} \frac{\hbar}{i} \frac{d\psi_0(x)}{dx} &= 0 \\
\Rightarrow \frac{d\psi_0(x)}{dx} + \frac{m\omega x}{\hbar} \psi_0(x) &= 0
\end{aligned} \tag{5.271}$$

This equation is easily solved to give

$$\psi_0(x) = N e^{-m\omega x^2/2\hbar} \tag{5.272}$$

where N is a constant to be fixed by normalisation. Using the normalization condition (5.111) gives

$$\psi_0(x) = \left(\frac{m\omega}{\pi\hbar} \right)^{1/4} e^{-m\omega x^2/2\hbar} \tag{5.273}$$

To calculate the excited state wavefunctions, we use (5.269). Recursively using the first relation gives

$$\begin{aligned} |E_n\rangle &= \frac{1}{\sqrt{n!}} (\hat{a}^\dagger)^n |E_0\rangle \\ &= \frac{1}{\sqrt{n!}} \left(\sqrt{\frac{m\omega}{2\hbar}} \hat{x} - i \frac{1}{\sqrt{2m\omega\hbar}} \hat{p} \right)^n |E_0\rangle \end{aligned} \quad (5.274)$$

Taking the inner product with $|x\rangle$ and replacing \hat{x} and \hat{p} by the corresponding differential operators, we get

$$\psi_n(x) = \frac{1}{\sqrt{n!}} \left(\sqrt{\frac{m\omega}{2\hbar}} x - \sqrt{\frac{\hbar}{2m\omega}} \frac{d}{dx} \right)^n \psi_0(x) \quad (5.275)$$

where the differential operator

$$\hat{D} = \sqrt{\frac{m\omega}{2\hbar}} x - \sqrt{\frac{\hbar}{2m\omega}} \frac{d}{dx} \quad (5.276)$$

is understood to act n times in succession on the ground state wavefunction. At this point, it is useful to express coordinate x in terms of the fundamental length $l = \sqrt{\hbar/m\omega}$, as $x = ly$. Then, in terms of coordinate y , the excited state wavefunctions have the simple form

$$\psi_n(y) = \frac{1}{2^{n/2}\sqrt{n!}} \left(y - \frac{d}{dy} \right)^n \psi_0(y) \quad (5.277)$$

with

$$\psi_0(y) = \left(\frac{m\omega}{\pi\hbar} \right)^{1/4} e^{-y^2/2} \quad (5.278)$$

It is easy to see that

$$\left(y - \frac{d}{dy} \right)^n e^{-y^2/2} = H_n(y) e^{-y^2/2} \quad (5.279)$$

where $H_n(x)$ is a polynomial of degree n . This equation can be rewritten as

$$H_n(y) = e^{y^2/2} \left(y - \frac{d}{dy} \right)^n e^{-y^2/2} \quad (5.280)$$

These polynomials are known as *Hermite Polynomials*. The first few, generated using (5.280), are

$$\begin{aligned} H_0(x) &= 1 \\ H_1(y) &= 2y \\ H_2(y) &= -2(1 - 2y^2) \\ H_3(y) &= -12(1 - 2y^3/3) \end{aligned} \quad (5.281)$$

Then, the energy eigenfunctions are

$$\psi_n(y) = \left(\frac{m\omega}{\pi\hbar} \right)^{1/4} \frac{1}{2^{n/2}\sqrt{n!}} H_n(y) e^{-y^2/2} \quad (5.282)$$

In terms of the original position coordinate x , these will be

$$\psi_n(x) = \left(\frac{m\omega}{\pi\hbar 2^{2n}(n!)^2} \right)^{1/4} H_n \left(\sqrt{\frac{m\omega}{\hbar}} x \right) e^{-m\omega x^2/2\hbar} \quad (5.283)$$

These then are the wavefunctions for states of definite energy. It is important to note that the overall Gaussian envelope makes the wavefunction for any n fall exponentially beyond a region of size of the order of l . It is easy to prove using (5.280) that the Hermite polynomials satisfy

$$H_n(-x) = (-1)^n H_n(x) \quad (5.284)$$

which in turn leads to

$$\psi_n(-x) = (-1)^n \psi_n(x) \quad (5.285)$$

This behaviour is understandable, since the potential energy function (5.233) is an even function of x . The first few wavefunctions are plotted

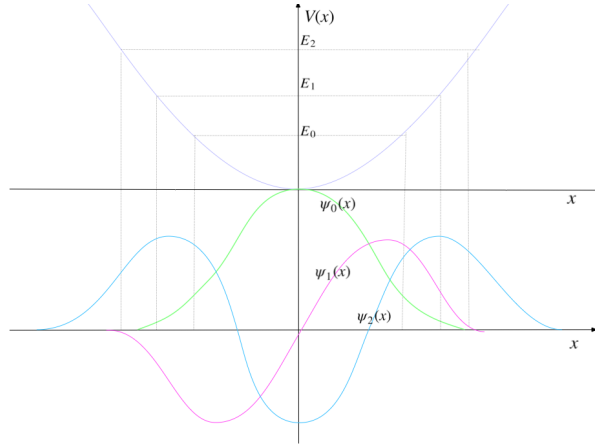


Figure 5.18: First three energy eigenfunctions for a Harmonic Oscillator.

From (5.283), it is clear that for any n , the wavefunctions are localised to a region of the order of $l = \sqrt{\hbar/m\omega}$ because of the Gaussian envelope.

Let us now calculate the matrix elements of position and momentum operators, their expectation values and uncertainties in the energy eigenstates. For a harmonic oscillator, this is a simple exercise because of the algebra of creation and annihilation operators. First, from (5.249) and (5.250), we express \hat{x} and \hat{p} in terms of \hat{a} and \hat{a}^\dagger

$$\begin{aligned} \hat{x} &= \left(\frac{\hbar}{2m\omega} \right)^{1/2} (\hat{a} + \hat{a}^\dagger) \\ \hat{p} &= i \left(\frac{m\omega\hbar}{2} \right)^{1/2} (\hat{a}^\dagger - \hat{a}) \end{aligned} \quad (5.286)$$

Let us calculate the expectation value of \hat{x} in state $|E_n\rangle$

$$\begin{aligned} \langle E_n | \hat{x} | E_n \rangle &= \left(\frac{\hbar}{2m\omega} \right)^{1/2} \left(\langle E_n | \hat{a} | E_n \rangle + \langle E_n | \hat{a}^\dagger | E_n \rangle \right) \\ &= \left(\frac{\hbar}{2m\omega} \right)^{1/2} \left(\sqrt{n} \langle E_n | E_{n-1} \rangle + \sqrt{n+1} \langle E_n | E_{n+1} \rangle \right) \\ &= 0 \end{aligned} \quad (5.287)$$

where we have used orthonormality of states $|E_n\rangle$. Similarly, it is easy to check that the expectation value of momentum in an energy eigenstate is zero. To calculate uncertainties, we need to calculate the

expectation value of \hat{x}^2 and \hat{p}^2 .

$$\begin{aligned}
\langle E_n | \hat{x}^2 | E_n \rangle &= \left(\frac{\hbar}{2m\omega} \right) \langle E_n | (\hat{a} + \hat{a}^\dagger)^2 | E_n \rangle \\
&= \left(\frac{\hbar}{2m\omega} \right) \langle E_n | (\hat{a}^2 + \hat{a}^{\dagger 2} + \hat{a}\hat{a}^\dagger + \hat{a}^\dagger\hat{a}) | E_n \rangle \\
&= \left(\frac{\hbar}{2m\omega} \right) \left(\langle E_n | \hat{a}\hat{a}^\dagger | E_n \rangle + \langle E_n | \hat{a}^\dagger\hat{a} | E_n \rangle \right) \\
&= \left(\frac{\hbar}{2m\omega} \right) (\sqrt{n+1}\sqrt{n+1} + \sqrt{n}\sqrt{n}) \\
&= (2n+1) \left(\frac{\hbar}{2m\omega} \right)
\end{aligned} \tag{5.288}$$

then, the uncertainty in position in state $|E_n\rangle$ is

$$\Delta x_n = \sqrt{(2n+1) \left(\frac{\hbar}{2m\omega} \right)} \tag{5.289}$$

similarly, it is straightforward to calculate the uncertainty in momentum

$$\Delta p_n = \sqrt{(2n+1) \left(\frac{m\omega\hbar}{2} \right)} \tag{5.290}$$

Then, the product of uncertainties will give

$$\Delta x_n \Delta p_n = (2n+1) \frac{\hbar}{2} \tag{5.291}$$

Clearly, the product is least for the ground state, which has a Gaussian wavefunction.

We had discussed that the harmonic oscillator potential energy function can be regarded as an approximation to a generic potential energy function, if we are interested only in the lowest few energy eigenvalues and wavefunctions. However, with increasing energy, the anharmonicity in the potential energy becomes important. This can be introduced as a ‘small’ correction to the harmonic oscillator potential energy function

$$V(x) = \frac{1}{2}m\omega^2x^2 - \alpha x^3 - \beta x^4 \tag{5.292}$$

where it is assumed that α and β are ‘small’ and positive. What are the new energy eigenvalues? The answer turns out to be the following: say, we have a Hamiltonian \hat{H} with eigenstates $|E_n\rangle$, such that

$$\hat{H} |E_n\rangle = E_n |E_n\rangle \tag{5.293}$$

Let the Hamiltonian be changed due to a small correction, so that it is of the form

$$\hat{H}' = \hat{H} + \lambda \hat{H}_1 \tag{5.294}$$

where λ is ‘small’. The eigenstates $|E'_n\rangle$ will now be different, corresponding to new eigenvalues

$$\hat{H}' |E'_n\rangle = E'_n |E'_n\rangle \tag{5.295}$$

Then, it can be shown that to lowest order, the new eigenvalues are

$$E'_n = E_n + \lambda \langle E_n | \hat{H}_1 | E_n \rangle \tag{5.296}$$

Then, the change in energy eigenvalue in the n^{th} state is equal to the expectation value of the correction to the Hamiltonian, calculated in the original energy eigenstate. Let us apply this result to the anharmonic oscillator Hamiltonian. From (5.292), the correction to the Hamiltonian will be

$$\lambda \hat{H}_1 = -\alpha \hat{x}^2 - \beta \hat{x}^4 \tag{5.297}$$

The n^{th} energy eigenvalue will now be

$$E'_n = \left(n + \frac{1}{2}\right) \hbar\omega - \alpha \langle E_n | \hat{x}^3 | E_n \rangle - \beta \langle E_n | \hat{x}^4 | E_n \rangle \quad (5.298)$$

Expressing \hat{x} in terms of \hat{a} and \hat{a}^\dagger , this can be easily calculated to be

$$E'_n = \left(n + \frac{1}{2}\right) \hbar\omega - \beta \left(\frac{\hbar}{2m\omega}\right)^2 [3 + 6n(n + 1)] \quad (5.299)$$

Given this, the energy spacing between the n^{th} and the $(n + 1)^{\text{th}}$ eigenvalues will be

$$\Delta E_n = \hbar\omega - 12\beta \left(\frac{\hbar}{2m\omega}\right)^2 (n + 1) \quad (5.300)$$

which decreases with increasing n . Note that this expression is valid for only those n for which the second term is small compared with the first one. For large enough n , this approximation will break down.

Chapter 6

Quantum Mechanics in three dimensions

6.1 Generalization to three dimensions

We now generalize the formalism developed for one-dimensional quantum propagation of a particle to three dimensions. The generalization is straightforward, but the addition of two dimensions introduces more physical observables (angular momentum). Also, some theorems that we proved for one-dimensional propagation will no longer be valid (theorems on properties of bound states). Of course, the real world is three-dimensional, so we will get to apply quantum mechanics to problems of real physical interest.

We begin by identifying the fundamental observables in three dimensions. These will be three position coordinates x, y, z and momentum components p_x, p_y, p_z of the particle, measured with respect to a suitable coordinate system. These will now be represented by Hermitian operators. Then, there will be three Hermitian operators for position ($\hat{x}, \hat{y}, \hat{z}$) and three for momentum ($\hat{p}_x, \hat{p}_y, \hat{p}_z$). We now wish to identify their commutator algebra. We will denote the three position operators generically as \hat{x}_i , where $i = 1, 2, 3$ and $\hat{x}_1 = \hat{x}$, $\hat{x}_2 = \hat{y}$ and $\hat{x}_3 = \hat{z}$. Similarly, for momentum as well. To determine the commutators of the three position operators with each other, we need to realize that in principle, it must be possible to create a state in which the position of the particle is well-defined. In such a state, the particle will occupy some point in space with coordinates x, y and z , measured with respect to the coordinate system. We denote such a state as $|x, y, z\rangle$ or $|\mathbf{x}\rangle$ where \mathbf{x} is the position vector of the point. Clearly, in this state, since the three coordinates are well-defined, it must be a simultaneous eigenstate of the operators \hat{x}_i , such that

$$\hat{x}_i |\mathbf{x}\rangle = x_i |\mathbf{x}\rangle \quad (6.1)$$

where $x_1 = x$, $x_2 = y$ and $x_3 = z$. Then, the commutators of the three operators \hat{x}_i with each other must be zero

$$[\hat{x}_i, \hat{x}_j] = 0 \quad (6.2)$$

Similarly, there must exist states of well-defined momentum, in which the three components of momentum are well-defined. We denote such states as $|p_x, p_y, p_z\rangle$ or simply $|\mathbf{p}\rangle$. This must be a simultaneous eigenstate of operators \hat{p}_i

$$\hat{p}_i |\mathbf{p}\rangle = p_i |\mathbf{p}\rangle \quad (6.3)$$

Then, it follows that

$$[\hat{p}_i, \hat{p}_j] = 0 \quad (6.4)$$

We now need to determine commutators of the form $[\hat{x}_i, \hat{p}_j]$. We know that if the motion of the particle is constrained to a straight line (say, the x axis), the position and momentum operators corresponding to such motion satisfy (5.104). Then, it must be true that

$$\begin{aligned} [\hat{x}, \hat{p}_x] &= i\hbar \hat{I} \\ [\hat{y}, \hat{p}_y] &= i\hbar \hat{I} \\ [\hat{z}, \hat{p}_z] &= i\hbar \hat{I} \end{aligned} \quad (6.5)$$

We now postulate that the rest of the commutators (cross-commutators) must be zero. For instance, $[\hat{x}, \hat{p}_y] = 0$, etc. The justification for this will come from Ehrenfest's Theorem, which will yield the correct classical equations, provided these cross-commutators are zero. We can express the position-momentum commutators concisely as follows

$$[\hat{x}_i, \hat{p}_j] = i\hbar \hat{I} \delta_{ij} \quad (6.6)$$

Given (6.6), it is possible to have quantum states in which different components of position and momentum are simultaneously well-defined. For instance, it is possible to have simultaneous eigenstates of \hat{x} , \hat{y} and \hat{p}_z . We can denote such a state by the simultaneous eigenvalues $|x, y, p_z\rangle$. In this state, the particle is confined to a line parallel to the z direction. Along this line, its position is not well-defined, but its momentum component (along this direction) is well-defined.

We now generalize the entire set of results for quantum propagation in one dimension to three dimensions. In three dimensions, the quantum state of the particle is a vector $|\psi\rangle$ in an infinite dimensional vector space \mathcal{H} with inner product defined (a *Hilbert Space*). In this vector space, states of definite position ($|x\rangle$) or definite momentum ($|p\rangle$) can be chosen as a continuous basis. These bases have a delta function normalization

$$\langle \mathbf{x}' | \mathbf{x} \rangle = \delta^3(\mathbf{x}' - \mathbf{x}) \quad (6.7)$$

and

$$\langle \mathbf{p}' | \mathbf{p} \rangle = \delta^3(\mathbf{p}' - \mathbf{p}) \quad (6.8)$$

The expansion of state $|\psi\rangle$ in the position basis has the form

$$|\psi\rangle = \int d^3\mathbf{x} \psi(\mathbf{x}) |\mathbf{x}\rangle \quad (6.9)$$

where $d^3\mathbf{x} = dx dy dz$ and

$$\psi(\mathbf{x}) = \langle \mathbf{x} | \psi \rangle \quad (6.10)$$

is the position space wavefunction, with $\psi(\mathbf{x}) = \psi(x, y, z)$ being the probability amplitude of detecting the particle at point (x, y, z) . The modulus squared of the wavefunction gives the probability density at this point. The probability of detecting the particle in a region of space \mathcal{V} is

$$P(\mathcal{V}) = \int_{\mathcal{V}} d^3\mathbf{x} |\psi(\mathbf{x})|^2 \quad (6.11)$$

Physical states are represented by normalized wavefunctions

$$\int_{\text{all space}} d^3\mathbf{x} |\psi(\mathbf{x})|^2 = 1 \quad (6.12)$$

The expansion of the state in the momentum basis is

$$|\psi\rangle = \int d^3\mathbf{p} \tilde{\psi}(\mathbf{p}) |\mathbf{p}\rangle \quad (6.13)$$

where $d^3\mathbf{p} = dp_x dp_y dp_z$ and

$$\tilde{\psi}(\mathbf{p}) = \langle \mathbf{p} | \psi \rangle \quad (6.14)$$

is the momentum space wavefunction, with $\tilde{\psi}(\mathbf{p}) = \tilde{\psi}(p_x, p_y, p_z)$ being the probability amplitude of detecting the particle with momentum \mathbf{p} . The modulus squared of this wavefunction gives the probability density of detecting this momentum. The two sets of bases are related through a Fourier integral

$$|\mathbf{p}\rangle = \frac{1}{(2\pi\hbar)^{3/2}} \int d^3\mathbf{x} e^{i\mathbf{p}\cdot\mathbf{x}/\hbar} |\mathbf{x}\rangle \quad (6.15)$$

and

$$|\mathbf{x}\rangle = \frac{1}{(2\pi\hbar)^{3/2}} \int d^3\mathbf{p} e^{-i\mathbf{p}\cdot\mathbf{x}/\hbar} |\mathbf{p}\rangle \quad (6.16)$$

The two sets of wavefunctions are similarly related

$$\tilde{\psi}(\mathbf{p}) = \frac{1}{(2\pi\hbar)^{3/2}} \int d^3\mathbf{x} e^{-i\mathbf{p}\cdot\mathbf{x}/\hbar} \psi(\mathbf{x}) \quad (6.17)$$

and

$$\psi(\mathbf{x}) = \frac{1}{(2\pi\hbar)^{3/2}} \int d^3\mathbf{p} e^{i\mathbf{p}\cdot\mathbf{x}/\hbar} \tilde{\psi}(\mathbf{p}) \quad (6.18)$$

The matrix elements of position and momentum operators in the position basis are

$$\langle \mathbf{x}' | \hat{x}_i | \mathbf{x} \rangle = x_i \delta^3(\mathbf{x}' - \mathbf{x}) \quad (6.19)$$

and

$$\langle \mathbf{x}' | \hat{p}_i | \mathbf{x} \rangle = \frac{\hbar}{i} \frac{\partial}{\partial x'_i} \delta^3(\mathbf{x}' - \mathbf{x}) \quad (6.20)$$

These further lead to

$$\langle \mathbf{x} | \hat{x}_i | \psi \rangle = x_i \psi(\mathbf{x}) \quad (6.21)$$

and

$$\langle \mathbf{x} | \hat{p}_i | \psi \rangle = \frac{\hbar}{i} \frac{\partial \psi(\mathbf{x})}{\partial x_i} \quad (6.22)$$

To gain familiarity with the formalism, let us determine the wavefunction for a particle in a state of well-defined momentum $|\mathbf{p}\rangle$. This state satisfies the eigenvalue equations (6.3). Taking the inner product with $|\mathbf{x}\rangle$ and using the matrix element (6.22), we get the following set of (three) simultaneous partial differential equations

$$\frac{\hbar}{i} \frac{\partial \psi_{\mathbf{p}}(\mathbf{x})}{\partial x_i} = p_i \psi_{\mathbf{p}}(\mathbf{x}) \quad (6.23)$$

where $\psi_{\mathbf{p}}(\mathbf{x}) = \langle \mathbf{x} | \mathbf{p} \rangle$ is the wavefunction for the state of definite momentum. Let us explicitly write this set of equations

$$\begin{aligned} \frac{\partial \psi_{\mathbf{p}}(\mathbf{x})}{\partial x} &= \frac{ip_x}{\hbar} \psi_{\mathbf{p}}(\mathbf{x}) \\ \frac{\partial \psi_{\mathbf{p}}(\mathbf{x})}{\partial y} &= \frac{ip_y}{\hbar} \psi_{\mathbf{p}}(\mathbf{x}) \\ \frac{\partial \psi_{\mathbf{p}}(\mathbf{x})}{\partial z} &= \frac{ip_z}{\hbar} \psi_{\mathbf{p}}(\mathbf{x}) \end{aligned} \quad (6.24)$$

The equations are separable, and can be easily solved to give

$$\begin{aligned} \psi_{\mathbf{p}}(\mathbf{x}) &= N e^{ip_x x/\hbar} e^{ip_y y/\hbar} e^{ip_z z/\hbar} \\ &= N e^{i(p_x x + p_y y + p_z z)/\hbar} \\ &= N e^{i\mathbf{p}\cdot\mathbf{x}/\hbar} \end{aligned} \quad (6.25)$$

where N is a normalization constant. This is determined from the condition (6.8) to be $1/(2\pi\hbar)^{3/2}$. Then, the wavefunction is

$$\psi_{\mathbf{p}}(\mathbf{x}) = \frac{1}{(2\pi\hbar)^{3/2}} e^{i\mathbf{p}\cdot\mathbf{x}/\hbar} \quad (6.26)$$

It is clear that in this state, the probability of detecting the particle at any point is the same, that is, uncertainty in position is infinite.

The Hamiltonian operator in three dimensions is the same function of position and momentum operators as the classical energy is of position and momentum

$$\hat{H} = \frac{1}{2m} \hat{\mathbf{p}}^2 + V(\hat{\mathbf{x}}) \quad (6.27)$$

where $\hat{\mathbf{p}}^2 = \hat{p}_x^2 + \hat{p}_y^2 + \hat{p}_z^2$ and $V(\hat{\mathbf{x}}) = V(\hat{x}, \hat{y}, \hat{z})$. Energy eigenstates satisfy (5.181), which can be reduced to a differential equation satisfied by the energy eigenfunctions

$$-\frac{\hbar^2}{2m}\nabla^2\psi_E(\mathbf{x}) + V(\mathbf{x})\psi_E(\mathbf{x}) = E\psi_E(\mathbf{x}) \quad (6.28)$$

where $\psi_E(\mathbf{x}) = \langle \mathbf{x} | E \rangle$. This can be solved for bound states by imposing the condition $\lim_{|\mathbf{x}| \rightarrow \infty} \psi_E(\mathbf{x}) = 0$. Scattering states do not satisfy this condition.

The time-evolution of the state is given by Schrodinger equation (4.41). Expressed as a differential equation, it gives the time-evolution of the wavefunction. The generalization of (5.180) to three dimensions is

$$i\hbar\frac{\partial\psi(\mathbf{x}, t)}{\partial t} = -\frac{\hbar^2}{2m}\nabla^2\psi(\mathbf{x}, t) + V(\mathbf{x})\psi(\mathbf{x}, t) \quad (6.29)$$

As before, this can be solved by expanding the initial state in terms of energy eigenstates. The wavefunction at instant t is given by the three dimensional analog of (5.186)

$$\psi(\mathbf{x}, t) = \sum_n c_n(t_0)e^{-iE_n(t-t_0)/\hbar}\psi_n(\mathbf{x}) \quad (6.30)$$

where $\psi_n(\mathbf{x}) = \langle \mathbf{x} | E_n \rangle$ is the wavefunction for state $|E_n\rangle$. The coefficients $c_n(t_0)$ are given by

$$c_n(t_0) = \int d^3\mathbf{x} \psi_n^*(\mathbf{x})\psi(\mathbf{x}, t_0) \quad (6.31)$$

Let us now determine the quantum dynamics of a free particle. The Hamiltonian for a free particle is

$$\hat{H} = \frac{1}{2m}(\hat{p}_x^2 + \hat{p}_y^2 + \hat{p}_z^2) \quad (6.32)$$

Just as in one dimension, states of definite momentum are states of definite energy

$$\begin{aligned} \hat{H}|\mathbf{p}\rangle &= \frac{1}{2m}(\hat{p}_x^2 + \hat{p}_y^2 + \hat{p}_z^2)|\mathbf{p}\rangle \\ &= \frac{1}{2m}(p_x^2 + p_y^2 + p_z^2)|\mathbf{p}\rangle \end{aligned} \quad (6.33)$$

We saw that in one dimension, given energy E , there were two degenerate states $|p\rangle$ and $|-p\rangle$ with $p = \sqrt{2mE}$ with the same energy, and any linear superposition of these states is a state of well-defined energy. In three dimensions, it is easy to see that there is an infinite degeneracy. Momentum eigenstates $|\mathbf{p}\rangle$ with a fixed *magnitude* of momentum $|\mathbf{p}|$ will have the same energy, whatever the direction of $|\mathbf{p}\rangle$. To construct a general superposition of these states, let us set up a coordinate system, with a given vector \mathbf{p} described by its magnitude p and two spherical polar angles θ and ϕ

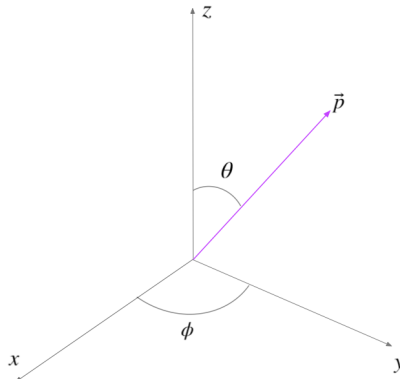


Figure 6.1: Direction of momentum denoted by spherical polar angles.

Then, a linear superposition of states corresponding to the same magnitude of momentum but different directions (measured through θ and ϕ) can be written as follows

$$|E\rangle = \int_0^\pi d\theta \sin\theta \int_0^{2\pi} d\phi f(\theta, \phi) |\mathbf{p}\rangle \quad (6.34)$$

This will be a state of definite energy for any function $f(\theta, \phi)$. The wavefunction for this state can be found by taking the inner product of this equation with $|\mathbf{x}\rangle$. It is convenient to choose the z direction along \mathbf{x} in this calculation

$$\begin{aligned} \psi_E(\mathbf{x}) &= \langle \mathbf{x} | E \rangle \\ &= \int_0^\pi d\theta \sin\theta \int_0^{2\pi} d\phi f(\theta, \phi) \langle \mathbf{x} | \mathbf{p} \rangle \\ &= \frac{1}{(2\pi\hbar)^{3/2}} \int_0^\pi d\theta \sin\theta \int_0^{2\pi} d\phi f(\theta, \phi) e^{i\mathbf{p}\cdot\mathbf{x}/\hbar} \\ &= \frac{1}{(2\pi\hbar)^{3/2}} \int_0^\pi d\theta \sin\theta \int_0^{2\pi} d\phi f(\theta, \phi) e^{ip|\mathbf{x}|\cos\theta/\hbar} \end{aligned} \quad (6.35)$$

A special choice of the function $f(\theta, \phi)$ could be $f = c$ where c is a constant. For this choice, we get a spherically symmetric superposition of momentum states. The integral is easy to perform and gives the following energy eigenfunction

$$\psi_E(\mathbf{x}) = \frac{4\pi c}{(2\pi\hbar)^{3/2}} \frac{\sin(p|\mathbf{x}|/\hbar)}{(p|\mathbf{x}|/\hbar)} \quad (6.36)$$

As expected, this wavefunction is spherically symmetric.

From (6.29), Schrodinger equation for a free particle is obtained by setting $V = 0$

$$i\hbar \frac{\partial\psi(\mathbf{x}, t)}{\partial t} = -\frac{\hbar^2}{2m} \nabla^2 \psi(\mathbf{x}, t) \quad (6.37)$$

Given the wavefunction at some instant t_0 , this allows us to compute the wavefunction at any instant of time for a free particle. As in one dimension, we can find an integral solution to this equation in the form of a ‘propagator’. To get it, we expand the state at instant t_0 in terms of states of definite momentum (which are also states of definite energy for a free particle)

$$|\psi(t_0)\rangle = \int d^3\mathbf{p} \tilde{\psi}(\mathbf{p}, t_0) |\mathbf{p}\rangle \quad (6.38)$$

The state at instant t is obtained by operating on $|\psi(t_0)\rangle$ by the time-evolution operator

$$\begin{aligned} |\psi(t)\rangle &= e^{-i\hat{H}(t-t_0)/\hbar} |\psi(t_0)\rangle \\ &= \int d^3\mathbf{p} \tilde{\psi}(\mathbf{p}, t_0) e^{-i\hat{H}(t-t_0)/\hbar} |\mathbf{p}\rangle \\ &= \int d^3\mathbf{p} \tilde{\psi}(\mathbf{p}, t_0) e^{-i\mathbf{p}^2(t-t_0)/2m\hbar} |\mathbf{p}\rangle \end{aligned} \quad (6.39)$$

where $\mathbf{p}^2 = p_x^2 + p_y^2 + p_z^2$. The wavefunction at t is found by taking the inner product of this equation with $|\mathbf{x}\rangle$

$$\begin{aligned} \psi(\mathbf{x}, t) &= \int d^3\mathbf{p} \tilde{\psi}(\mathbf{p}, t_0) e^{-i\mathbf{p}^2(t-t_0)/2m\hbar} \langle \mathbf{x} | \mathbf{p} \rangle \\ &= \frac{1}{(2\pi\hbar)^{3/2}} \int d^3\mathbf{p} \tilde{\psi}(\mathbf{p}, t_0) e^{-i\mathbf{p}^2(t-t_0)/2m\hbar} e^{i\mathbf{p}\cdot\mathbf{x}/\hbar} \end{aligned} \quad (6.40)$$

As in one dimension, this can be written in terms of a ‘propagator’

$$\psi(\mathbf{x}, t) = \int d^3\mathbf{x}' U(\mathbf{x}, t; \mathbf{x}', t_0) \psi(\mathbf{x}', t_0) \quad (6.41)$$

where the propagator $U(\mathbf{x}, t; \mathbf{x}', t_0)$ is given by

$$U(\mathbf{x}, t; \mathbf{x}', t_0) = \frac{1}{(2\pi\hbar)^3} \int d^3\mathbf{p} e^{-i\mathbf{p}^2(t-t_0)/2m\hbar} e^{i\mathbf{p}\cdot(\mathbf{x}-\mathbf{x}')/\hbar} \quad (6.42)$$

The propagator tells us how the quantum particle propagates in three dimensions, since it can be easily verified that

$$U(\mathbf{x}, t; \mathbf{x}', t_0) = \langle \mathbf{x} | e^{-i\hat{H}(t-t_0)/\hbar} | \mathbf{x}' \rangle \quad (6.43)$$

which is the amplitude that the particle, located at point \mathbf{x}' at instant t_0 , propagates to point \mathbf{x} at instant t . The integral in (6.42) can be evaluated easily, by expanding the exponents in the integrand in terms of components. This results in the integrals over x , y and z decoupling from each other, each reducing to the exact same Gaussian form, encountered in the one-dimensional case. After performing the identical integrals, we finally get

$$U(\mathbf{x}, t; \mathbf{x}', t_0) = \left(\frac{m}{2\pi i\hbar(t-t_0)} \right)^{3/2} e^{im(\mathbf{x}-\mathbf{x}')^2/2\hbar(t-t_0)} \quad (6.44)$$

We now study particle confinement in three dimensions. Given an attractive potential $V(\mathbf{x})$, bound states are obtained by imposing the boundary condition $\lim_{|\mathbf{x}|\rightarrow\infty} \psi_E(\mathbf{x}) = 0$ in the energy eigenvalue equation (6.28). We will see that in general, such states will correspond to discrete values of energy. If we do not impose this boundary condition, we get scattering states, and these will correspond to a continuum of energy eigenvalues. As in one dimension, the simplest example of confinement is provided by a ‘box’ potential energy function

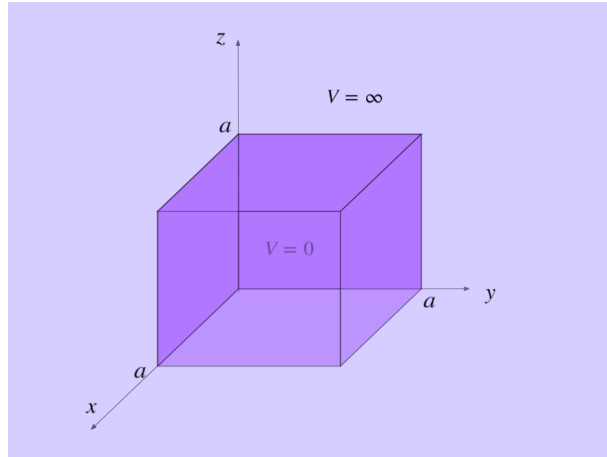


Figure 6.2: A three dimensional ‘box’ potential. The potential energy function is zero inside the cubic region and infinite outside.

The potential energy function is zero within a cubic region of volume $V = a^3$, bound by the planes $x = 0, x = a; y = 0, y = a; z = 0, z = a$. We write the eigenvalue equation (6.28) in Cartesian coordinates x, y, z . Within the cubic region, $V = 0$. Therefore, the equation has the form

$$-\frac{\hbar^2}{2m} \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) \psi_E(x, y, z) = E \psi_E(x, y, z) \quad (6.45)$$

which can be rewritten as

$$\frac{\partial^2 \psi_E}{\partial x^2} + \frac{\partial^2 \psi_E}{\partial y^2} + \frac{\partial^2 \psi_E}{\partial z^2} = -k^2 \psi_E \quad (6.46)$$

where $k = \sqrt{2mE}/\hbar$. Note that $E > 0$, since for bound states the lowest value of energy is greater than the least value of the potential energy function. The wavefunction outside the cubic region must

vanish, since the potential energy function is infinite everywhere in that region. From continuity, equation (6.45) must be solved with the boundary condition that the wavefunction vanishes on all faces of the cube. The rectangular symmetry of the boundary conditions allows us to use separation of variables. Writing $\psi_E(x, y, z) = X(x)Y(y)Z(z)$ and simplifying, we get

$$\frac{1}{X(x)} \frac{d^2 X}{dx^2} + \frac{1}{Y(y)} \frac{d^2 Y}{dy^2} + \frac{1}{Z(z)} \frac{d^2 Z}{dz^2} = -k^2 \quad (6.47)$$

Rearranging, we get

$$\frac{1}{X(x)} \frac{d^2 X}{dx^2} = - \left(k^2 + \frac{1}{Y(y)} \frac{d^2 Y}{dy^2} + \frac{1}{Z(z)} \frac{d^2 Z}{dz^2} \right) \quad (6.48)$$

The left side is a function of x only and the right side is a function of both y and z . The only way that this equation can be true is for both sides of the equation to be constant. We assume that this constant is negative (assume it is positive and carry out the analysis. You will get a trivial solution), and write it as $-\alpha^2$. Then, we get two sets of equations

$$\frac{1}{X(x)} \frac{d^2 X}{dx^2} = -\alpha^2 \quad (6.49)$$

and

$$\frac{1}{Y(y)} \frac{d^2 Y}{dy^2} + \frac{1}{Z(z)} \frac{d^2 Z}{dz^2} + k^2 = \alpha^2 \quad (6.50)$$

The solution to (6.49) is of the form

$$X(x) = A \cos(\alpha x) + B \sin(\alpha x) \quad (6.51)$$

We now impose the boundary condition that the wavefunction must vanish on the faces of the cube. Consider the faces $x = 0, 0 \leq y \leq a, 0 \leq z \leq a$ and $x = a, 0 \leq y \leq a, 0 \leq z \leq a$. The boundary condition applied to these faces gives $\psi(0, y, z) = \psi(a, y, z) = 0$. This should be satisfied for all y and z in the given range. The only way this is possible is for $X(x)$ to vanish on these faces. Then, $X(0) = 0$ and $X(a) = 0$. Applied to solution (6.51), the first condition gives $A = 0$. The second condition gives

$$\begin{aligned} B \sin(\alpha a) &= 0 \\ \Rightarrow \alpha_{n_x} &= \frac{n_x \pi}{a}; \quad n_x = 1, 2, 3, \dots \end{aligned} \quad (6.52)$$

so that the solution (characterized by positive integer n_x) is

$$X_{n_x}(x) = B \sin\left(\frac{n_x \pi x}{a}\right) \quad (6.53)$$

Negative values of n_x will give the same solutions (upto a sign) and $n_x = 0$ will make the wavefunction disappear everywhere, which is unphysical. We now rewrite (6.50) as

$$\frac{1}{Y(y)} \frac{d^2 Y}{dy^2} = - \left(\frac{1}{Z(z)} \frac{d^2 Z}{dz^2} + k^2 - \alpha^2 \right) \quad (6.54)$$

As before, each side of this equation must be a negative constant, say $-\beta^2$. The Y equation can be solved to give

$$Y(y) = C \cos(\beta y) + D \sin(\beta y) \quad (6.55)$$

We now impose the boundary condition that the wavefunction must vanish on the faces $y = 0, 0 \leq x \leq a, 0 \leq z \leq a$ and $y = a, 0 \leq x \leq a, 0 \leq z \leq a$. As before, this gives $C = 0$ and

$$\begin{aligned} D \sin(\beta a) &= 0 \\ \Rightarrow \beta_{n_y} &= \frac{n_y \pi}{a}; \quad n_y = 1, 2, 3, \dots \end{aligned} \quad (6.56)$$

which gives the solution

$$Y_{n_y}(y) = D \sin\left(\frac{n_y \pi y}{a}\right) \quad (6.57)$$

The Z equation now reduces to

$$\frac{1}{Z(z)} \frac{d^2 Z}{dz^2} + k^2 - \alpha^2 = \beta^2 \quad (6.58)$$

which can be rewritten as

$$\frac{1}{Z(z)} \frac{d^2 Z}{dz^2} = \alpha^2 + \beta^2 - k^2 \quad (6.59)$$

This should equal a negative constant, say $-\gamma^2$ (a positive constant will give a trivial solution, which can be verified). Then, the solution to (6.59) will be

$$Z(z) = E \cos(\beta z) + F \sin(\beta z) \quad (6.60)$$

Imposing the boundary condition that the wavefunction must vanish on the faces $z = 0, 0 \leq x \leq a, 0 \leq y \leq a$ and $z = a, 0 \leq x \leq a, 0 \leq y \leq a$, we get

$$Z_{n_z}(z) = F \sin\left(\frac{n_z \pi z}{a}\right) \quad (6.61)$$

with

$$\gamma_{n_z} = \frac{n_z \pi}{a} \quad (6.62)$$

where n_z is a positive integer. At the end of the calculation, we are also left with the constraint

$$k^2 = \alpha^2 + \beta^2 + \gamma^2 \quad (6.63)$$

which, given that $k = \sqrt{2mE}/\hbar$, gives the allowed energy eigenvalues

$$\begin{aligned} E_{n_x, n_y, n_z} &= \frac{\pi^2 \hbar^2}{2ma^2} (n_x^2 + n_y^2 + n_z^2) \\ &= \frac{\pi^2 \hbar^2}{2mV^{2/3}} (n_x^2 + n_y^2 + n_z^2) \end{aligned} \quad (6.64)$$

where $V = a^3$ is the volume of the cubic region. Then, states of definite energy are characterized by three ‘quantum numbers’ n_x , n_y and n_z . We shall denote these states as $|n_x, n_y, n_z\rangle$. Then,

$$\hat{H} |n_x, n_y, n_z\rangle = E_{n_x, n_y, n_z} |n_x, n_y, n_z\rangle \quad (6.65)$$

The wavefunctions for these states are of the form

$$\psi_{n_x n_y n_z}(x, y, z) = A \sin\left(\frac{n_x \pi x}{a}\right) \sin\left(\frac{n_y \pi y}{a}\right) \sin\left(\frac{n_z \pi z}{a}\right) \quad (6.66)$$

where A is a normalization constant, to be determined by (6.12). Since the wavefunction vanishes outside the cubic region, the normalization condition is

$$|A|^2 \int_0^a dx \sin^2\left(\frac{n_x \pi x}{a}\right) \int_0^a dy \sin^2\left(\frac{n_y \pi y}{a}\right) \int_0^a dz \sin^2\left(\frac{n_z \pi z}{a}\right) = 1 \quad (6.67)$$

Each integral is independent, and equal to $a/2$. Then, the normalization constant is $A = (2/a)^{3/2} = 2^{3/2}/\sqrt{V}$. The normalized energy eigenfunctions (within the ‘box’) are

$$\psi_{n_x n_y n_z}(x, y, z) = \frac{2^{3/2}}{\sqrt{V}} \sin\left(\frac{n_x \pi x}{a}\right) \sin\left(\frac{n_y \pi y}{a}\right) \sin\left(\frac{n_z \pi z}{a}\right) \quad (6.68)$$

The ground state corresponds to $n_x = n_y = n_z = 1$. This state has energy

$$E_{111} = \frac{3\pi^2\hbar^2}{2mV^{3/2}} \quad (6.69)$$

This state has a unique wavefunction

$$\psi_{111}(x, y, z) = \frac{2^{3/2}}{\sqrt{V}} \sin\left(\frac{\pi x}{a}\right) \sin\left(\frac{\pi y}{a}\right) \sin\left(\frac{\pi z}{a}\right) \quad (6.70)$$

There are *three* first excited states. These are $|112\rangle$, $|121\rangle$ and $|211\rangle$. The wavefunctions for these states are clearly different, but they have the same (degenerate) energy

$$E_{112} = \frac{3\pi^2\hbar^2}{mV^{3/2}} \quad (6.71)$$

Clearly, the no-degeneracy theorem for bound states that we had proved in one dimension is not applicable to three dimensions. There *can* be degenerate energy states in more than one dimension.

6.2 Orbital Angular Momentum

In three dimensions, apart from position, momentum and energy, there is another observable of fundamental physical interest. This is angular momentum, which is not defined in one dimension. We carried out a general analysis of quantum mechanics of angular momentum in Section (4.5). The analysis was very general, and can be applied to angular momentum of any system. We found that there can exist states in which the magnitude of angular momentum and any *one* component can be simultaneously well-defined. This component is traditionally taken to be ‘z’. We found that the states are characterized by two quantum numbers j and m_j , where j can take integer or half-integer values, and for any one such value, m_j can take $2j + 1$ values ranging from $-j$ to j in steps of 1. The magnitude (squared) and z component of angular momentum in these states are given by (4.132)

$$\begin{aligned}\hat{J}^2 |j, m_j\rangle &= j(j+1)\hbar^2 |j, m_j\rangle \\ \hat{J}_z |j, m_j\rangle &= m_j \hbar |j, m_j\rangle\end{aligned}\quad (6.72)$$

However, the general formalism does not tell us what possible values of j a given system can take. For instance, $j = 1/2$ for intrinsic angular momentum of an electron, and so takes only one value. The angular momentum arising out of the motion of a particle in three dimensional space is often called *Orbital* angular momentum, and is classically given by

$$\mathbf{L} = \mathbf{x} \times \mathbf{p} \quad (6.73)$$

where \mathbf{x} is the position vector of the particle (with respect to some origin) and \mathbf{p} its momentum. Its components are

$$\begin{aligned}L_x &= y p_z - z p_y \\ L_y &= z p_x - x p_z \\ L_z &= x p_y - y p_x\end{aligned}\quad (6.74)$$

To ‘quantize’ this orbital angular momentum, we represent these components by Hermitian operators, which should be same functions of position and momentum operators, as the classical components are of position and momentum. Then, we get the following three operators

$$\begin{aligned}\hat{L}_x &= \hat{y} \hat{p}_z - \hat{z} \hat{p}_y \\ \hat{L}_y &= \hat{z} \hat{p}_x - \hat{x} \hat{p}_z \\ \hat{L}_z &= \hat{x} \hat{p}_y - \hat{y} \hat{p}_x\end{aligned}\quad (6.75)$$

We have seen that in general, the order in which we multiply two operators is important. In (6.75), though, it does not matter whether the position operators premultiply or postmultiply the momentum operators. For instance, the expression for \hat{L}_x can equivalently be written as

$$\hat{L}_x = \hat{p}_z \hat{y} - \hat{p}_y \hat{z} \quad (6.76)$$

This is because the commutator of \hat{y} with \hat{p}_z is zero, and so is the commutator of \hat{z} with \hat{p}_y , which follows from (6.6). Since the commutators of position operators with momentum operators are known to us, we can use them to evaluate the commutators of the orbital angular momentum operators to check if they satisfy the algebra (4.100). Let us explicitly calculate the commutator of \hat{L}_x with \hat{L}_y

$$\begin{aligned}\left[\hat{L}_x, \hat{L}_y \right] &= \left[\hat{y} \hat{p}_z - \hat{z} \hat{p}_y, \hat{z} \hat{p}_x - \hat{x} \hat{p}_z \right] \\ &= \left[\hat{y} \hat{p}_z, \hat{z} \hat{p}_x \right] - \left[\hat{y} \hat{p}_z, \hat{x} \hat{p}_z \right] - \left[\hat{z} \hat{p}_y, \hat{z} \hat{p}_x \right] + \left[\hat{z} \hat{p}_y, \hat{x} \hat{p}_z \right]\end{aligned}\quad (6.77)$$

Using commutation relations (6.6) and property (4.102) of commutators, we can isolate the non-zero terms

$$\begin{aligned}
[\hat{L}_x, \hat{L}_y] &= [\hat{y} \hat{p}_z, \hat{z} \hat{p}_x] + [\hat{z} \hat{p}_y, \hat{x} \hat{p}_z] \\
&= \hat{y} [\hat{p}_z, \hat{z}] \hat{p}_x + \hat{x} [\hat{z}, \hat{p}_z] \hat{p}_y \\
&= \hat{y} (-i\hbar \hat{I}) \hat{p}_x + \hat{x} (i\hbar \hat{I}) \hat{p}_y \\
&= i\hbar (\hat{x} \hat{p}_y - \hat{y} \hat{p}_x) \\
&= i\hbar \hat{L}_z
\end{aligned} \tag{6.78}$$

which is consistent with (4.100). Similarly, the other two commutators can be checked. Orbital angular momentum, then, indeed satisfies the general angular momentum algebra. Then, there must exist simultaneous eigenstates of \hat{L}^2 and \hat{L}_z , characterized by two numbers l and m_l , such that

$$\begin{aligned}
\hat{L}^2 |l, m_l\rangle &= l(l+1)\hbar^2 |l, m_l\rangle \\
\hat{L}_z |l, m_l\rangle &= m_l \hbar |l, m_l\rangle \\
m_l &= -l, -l+1, \dots, l-1, l
\end{aligned} \tag{6.79}$$

where the general algebra allows l to in principle take both integer and half-integer values. We now need to figure out what subset of these values are physically realized for orbital angular momentum. It is physically reasonable to expect that the subset will not be a finite set, unlike electron spin (in which case only one value of j is realized). In some limit, results should agree with classical physics, which allows for arbitrarily large angular momentum. Therefore, there cannot be an upper bound on the allowed values of l . To determine the allowed values of l , we construct wavefunctions for states $|l, m_l\rangle$. To do this, we will need matrix elements of the angular momentum operators in the position basis. We could deduce them rigorously, but it is sufficient (based on our experience) to use the ansatz of representing operators with differential operators. For instance, the matrix element $\langle \mathbf{x} | \hat{L}_z | \psi \rangle$ will be

$$\begin{aligned}
\langle \mathbf{x} | \hat{L}_z | \psi \rangle &= \langle \mathbf{x} | (\hat{x} \hat{p}_y - \hat{y} \hat{p}_x) | \psi \rangle \\
&= \langle \mathbf{x} | \hat{x} \hat{p}_y | \psi \rangle - \langle \mathbf{x} | \hat{y} \hat{p}_x | \psi \rangle \\
&= x \frac{\hbar}{i} \frac{\partial \psi}{\partial y} - y \frac{\hbar}{i} \frac{\partial \psi}{\partial x} \\
&= \frac{\hbar}{i} \left(x \frac{\partial}{\partial y} - y \frac{\partial}{\partial x} \right) \psi
\end{aligned} \tag{6.80}$$

Similarly,

$$\begin{aligned}
\langle \mathbf{x} | \hat{L}_x | \psi \rangle &= y \frac{\hbar}{i} \frac{\partial \psi}{\partial z} - z \frac{\hbar}{i} \frac{\partial \psi}{\partial y} \\
&= \frac{\hbar}{i} \left(y \frac{\partial}{\partial z} - z \frac{\partial}{\partial y} \right) \psi
\end{aligned} \tag{6.81}$$

and

$$\begin{aligned}
\langle \mathbf{x} | \hat{L}_y | \psi \rangle &= z \frac{\hbar}{i} \frac{\partial \psi}{\partial x} - x \frac{\hbar}{i} \frac{\partial \psi}{\partial z} \\
&= \frac{\hbar}{i} \left(z \frac{\partial}{\partial x} - x \frac{\partial}{\partial z} \right) \psi
\end{aligned} \tag{6.82}$$

Also, since $\hat{L}^2 = \hat{L}_x^2 + \hat{L}_y^2 + \hat{L}_z^2$, the matrix element $\langle \mathbf{x} | \hat{L}^2 | \psi \rangle$ will be

$$\begin{aligned}
\langle \mathbf{x} | \hat{L}^2 | \psi \rangle &= \langle \mathbf{x} | \hat{L}_x^2 | \psi \rangle + \langle \mathbf{x} | \hat{L}_y^2 | \psi \rangle + \langle \mathbf{x} | \hat{L}_z^2 | \psi \rangle \\
&= -\hbar^2 \left[\left(y \frac{\partial}{\partial z} - z \frac{\partial}{\partial y} \right)^2 + \left(z \frac{\partial}{\partial x} - x \frac{\partial}{\partial z} \right)^2 + \left(x \frac{\partial}{\partial y} - y \frac{\partial}{\partial x} \right)^2 \right] \psi
\end{aligned} \tag{6.83}$$

These expressions have a simpler form in spherical polar coordinates. Let us denote points in space using coordinates (r, θ, ϕ) , where

$$\begin{aligned}x &= r \sin \theta \cos \phi \\y &= r \sin \theta \sin \phi \\z &= r \cos \theta\end{aligned}\tag{6.84}$$

Let us express the wavefunction as a function of position through these coordinates. Using chain rule for partial derivatives, we can write the matrix elements of angular momentum operators in terms of partial derivatives with respect to the coordinates r, θ, ϕ . The relations between the two sets of partial derivatives are

$$\begin{aligned}\frac{\partial}{\partial r} &= \frac{\partial x}{\partial r} \frac{\partial}{\partial x} + \frac{\partial y}{\partial r} \frac{\partial}{\partial y} + \frac{\partial z}{\partial r} \frac{\partial}{\partial z} \\&= \sin \theta \cos \phi \frac{\partial}{\partial x} + \sin \theta \sin \phi \frac{\partial}{\partial y} + \cos \theta \frac{\partial}{\partial z} \\ \frac{\partial}{\partial \theta} &= \frac{\partial x}{\partial \theta} \frac{\partial}{\partial x} + \frac{\partial y}{\partial \theta} \frac{\partial}{\partial y} + \frac{\partial z}{\partial \theta} \frac{\partial}{\partial z} \\&= r \cos \theta \cos \phi \frac{\partial}{\partial x} + r \cos \theta \sin \phi \frac{\partial}{\partial y} - r \sin \theta \frac{\partial}{\partial z} \\ \frac{\partial}{\partial \phi} &= \frac{\partial x}{\partial \phi} \frac{\partial}{\partial x} + \frac{\partial y}{\partial \phi} \frac{\partial}{\partial y} + \frac{\partial z}{\partial \phi} \frac{\partial}{\partial z} \\&= -r \sin \theta \sin \phi \frac{\partial}{\partial x} + r \sin \theta \cos \phi \frac{\partial}{\partial y}\end{aligned}\tag{6.85}$$

We can invert these relations to express partial derivatives with respect to x, y, z in terms of those with respect to r, θ, ϕ . Subsequently, we can substitute those expressions into those for the matrix elements of the angular momentum operators. We then get the following simple expressions

$$\begin{aligned}\langle \mathbf{x} | \hat{L}_z | \psi \rangle &= -i\hbar \frac{\partial \psi}{\partial \phi} \\ \langle \mathbf{x} | \hat{L}_x | \psi \rangle &= -i\hbar \left(-\sin \phi \frac{\partial}{\partial \theta} - \cot \theta \cos \phi \frac{\partial}{\partial \phi} \right) \psi \\ \langle \mathbf{x} | \hat{L}_y | \psi \rangle &= -i\hbar \left(\cos \phi \frac{\partial}{\partial \theta} - \cot \theta \sin \phi \frac{\partial}{\partial \phi} \right) \psi\end{aligned}\tag{6.86}$$

The expression for the matrix element of \hat{L}^2 in spherical polar coordinates is similarly obtained

$$\langle \mathbf{x} | \hat{L}^2 | \psi \rangle = -\hbar^2 \left[\frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \phi^2} + \frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) \right] \psi\tag{6.87}$$

It should be noted that the matrix elements of the orbital angular momentum operators have *no radial coordinate dependence*. This is a big simplification, and comes about because of the relationship between angular momentum and rotations (which keep radial coordinate unchanged). We are now ready to construct wavefunctions of simultaneous eigenstates of \hat{L}^2 and \hat{L}_z , which satisfy (6.79). Taking the inner product of the \hat{L}_z eigenvalue equation with $|\mathbf{x}\rangle$ and using the matrix element for \hat{L}_z in spherical polar coordinates, we get

$$-i\hbar \frac{\partial \psi_{lm_l}}{\partial \phi} = m_l \hbar \psi_{lm_l}\tag{6.88}$$

where $\psi_{lm_l}(r, \theta, \phi) = \langle \mathbf{x} | l, m_l \rangle$. Since r and θ do not appear explicitly in this equation, it can be integrated to give the following

$$\psi_{lm_l}(r, \theta, \phi) = F_{lm_l}(r, \theta) e^{im_l \phi}\tag{6.89}$$

where F is a function of r and θ , but not ϕ . Now, under $\phi \rightarrow \phi + 2\pi$, the wavefunction should not change. This is because (r, θ, ϕ) and $(r, \theta, \phi + 2\pi)$ are coordinates of the same point in space and the probability amplitude of detecting the particle (in state $|l, m_l\rangle$) at this point must be unique. This is referred to as *single valuedness* of the wavefunction. Imposing this condition give

$$e^{2\pi i m_l} = 1 \quad (6.90)$$

which should be satisfied for all values of m_l . This will be true *only if* m_l is an integer. Since m_l depends on l through (6.79), it constrains l to have integer values. *Half-integer values are not realized in case of orbital angular momentum.* Therefore, the allowed values of l and m_l satisfy

$$\begin{aligned} l &= 0, 1, 2, 3, \dots \\ m_l &= -l, -l+1, \dots, l-1, l \end{aligned} \quad (6.91)$$

Taking the inner product of the \hat{L}^2 eigenvalue equation with $|\mathbf{x}\rangle$ gives the following differential equation

$$-\hbar^2 \left[\frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \phi^2} + \frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) \right] \psi_{lm_l} = l(l+1)\hbar^2 \psi_{lm_l} \quad (6.92)$$

Substituting (6.89) in (6.92), we get

$$-\hbar^2 \left[\frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) - \frac{m_l^2}{\sin^2 \theta} \right] F_{lm_l}(r, \theta) = l(l+1)\hbar^2 F_{lm_l}(r, \theta) \quad (6.93)$$

Since there are no derivatives with respect to r , we can write

$$F_{lm_l}(r, \theta) = R(r)P_l^{m_l}(\cos \theta) \quad (6.94)$$

where the function $P_l^{m_l}$ depends only on θ . The reason we have written it as a function of $\cos \theta$ will become clear soon. With this, the wavefunction ψ_{lm_l} can be written as

$$\psi_{lm_l} = R(r)Y_l^{m_l}(\theta, \phi) \quad (6.95)$$

where the angular functions

$$Y_l^{m_l}(\theta, \phi) = P_l^{m_l}(\cos \theta)e^{im_l \phi} \quad (6.96)$$

are referred to as *spherical Harmonics*. Substituting this in (6.93) gives the following *ordinary* differential equation for $P_l^{m_l}$

$$\left[\frac{1}{\sin \theta} \frac{d}{d\theta} \left(\sin \theta \frac{d}{d\theta} \right) - \frac{m_l^2}{\sin^2 \theta} \right] P_l^{m_l} + l(l+1)P_l^{m_l} = 0 \quad (6.97)$$

Recognizing that $\sin \theta(d/d\cos \theta) = -d/d\theta$, we can rewrite (6.97) as

$$\left[\frac{d}{d\cos \theta} \left((1 - \cos^2 \theta) \frac{d}{d\cos \theta} \right) - \frac{m_l^2}{1 - \cos^2 \theta} \right] P_l^{m_l}(\cos \theta) + l(l+1)P_l^{m_l}(\cos \theta) = 0 \quad (6.98)$$

where it is now explicitly clear that $P_l^{m_l}$ is a function of $\cos \theta$. Define $u = \cos \theta$, so that $-1 \leq u \leq 1$. In terms of u , the equation for $P_l^{m_l}$ becomes

$$\left[\frac{d}{du} \left((1 - u^2) \frac{d}{du} \right) - \frac{m_l^2}{1 - u^2} \right] P_l^{m_l}(u) + l(l+1)P_l^{m_l}(u) = 0 \quad (6.99)$$

If $m_l = 0$ for a given l , this reduces to

$$\frac{d}{du} (1 - u^2) \frac{d}{du} P_l^0 + l(l+1)P_l^0(u) = 0 \quad (6.100)$$

This is *Legendre equation*, and solutions to this equation are polynomials of degree l , known as Legendre polynomials. Then, wavefunctions for states $|l, 0\rangle$ are of the form

$$\psi_{l0}(r, \theta, \phi) = R(r)P_l(\cos \theta) \quad (6.101)$$

when $m_l \neq 0$, solutions to (6.99) are again polynomial functions, known as *associated Legendre polynomials*. These polynomials are more complex. Using ladder operators introduced in Section (4.5), we can construct these polynomials, just as we constructed harmonic oscillator energy eigenstates using creation-annihilation operators. First, we determine the matrix elements of ladder operators $\hat{L}_+ = \hat{L}_x \pm i\hat{L}_y$ in spherical polar coordinates, using (6.86)

$$\begin{aligned} \langle \mathbf{x} | \hat{L}_\pm | \psi \rangle &= \langle \mathbf{x} | \hat{L}_x | \psi \rangle \pm i \langle \mathbf{x} | \hat{L}_y | \psi \rangle \\ &= -i\hbar e^{\pm i\phi} \left(\pm i \frac{\partial}{\partial \theta} - \cot \theta \frac{\partial}{\partial \phi} \right) \psi \end{aligned} \quad (6.102)$$

We start with the state $|l, l\rangle$ and use the property

$$\hat{L}_+ |l, l\rangle = 0 \quad (6.103)$$

Taking the inner product of (6.103) with $|\mathbf{x}\rangle$, we get

$$-i\hbar e^{i\phi} \left(i \frac{\partial}{\partial \theta} - \cot \theta \frac{\partial}{\partial \phi} \right) \psi_{ll} = 0 \quad (6.104)$$

where $\psi_{ll} = \langle \mathbf{x} | l, l \rangle$. From (6.89) and (6.94) and with $m_l = l$, $\psi_{ll} = R(r)P_l^l e^{il\phi}$. Substituting this in (6.104), we get

$$\frac{dP_l^l}{d\theta} - l \cot \theta P_l^l = 0 \quad (6.105)$$

which is easily solved to get

$$P_l^l = c_l (\sin^l \theta) \quad (6.106)$$

where c_l is a constant of proportionality, determined by normalization. At this point, we need to fix the normalization of the angular functions $P_l^{m_l}(\cos \theta)$. The full wavefunction $\psi_{lm_l} = R(r)P_l^{m_l}(\cos \theta)e^{im_l\phi}$ needs to satisfy the normalization condition (6.12). Expressing the volume element in the integral in spherical polar coordinates, the normalization condition is

$$\int_0^\infty dr r^2 \int_0^\pi d\theta \sin \theta \int_0^{2\pi} d\phi |\psi_{lm_l}|^2 = 1 \quad (6.107)$$

This reduces to

$$2\pi \int_0^\infty dr r^2 |R(r)|^2 \int_0^\pi d\theta \sin \theta |P_l^{m_l}|^2 = 1 \quad (6.108)$$

We are at liberty to choose the normalizations of the radial function $R(r)$ and the angular function $P_l^{m_l}(\cos \theta)$ any way we choose, as long as the overall normalization (6.108) is respected. We choose the following natural conditions

$$\int_0^\infty dr r^2 |R(r)|^2 = 1 \quad (6.109)$$

and

$$\int_0^\pi d\theta \sin \theta |P_l^{m_l}|^2 = 1 \quad (6.110)$$

Condition (6.110) fixes the normalization constant in (6.106)

$$2\pi |c_l|^2 \int_0^\pi d\theta \sin^{2l+1} \theta = 1 \quad (6.111)$$

The integral can be expressed in terms of the Beta function. The constant c_l is then determined to be

$$c_l = \frac{(-1)^l}{2^l l!} \sqrt{\frac{(2l+1)!}{4\pi}} \quad (6.112)$$

The angular part of the wavefunction for the state $|l, l\rangle$ is then

$$Y_l^l(\theta, \phi) = \frac{(-1)^l}{2^l l!} \sqrt{\frac{(2l+1)!}{4\pi}} (\sin^l \theta) e^{il\phi} \quad (6.113)$$

Let us now calculate the wavefunction for the state $|l, l-1\rangle$ by operating on state $|l, l\rangle$ with \hat{L}_-

$$\hat{L}_- |l, l\rangle = \hbar\sqrt{2l} |l, l-1\rangle \quad (6.114)$$

which, in terms of differential operators and wavefunctions, reduces to

$$-i\hbar e^{-i\phi} \left(-i \frac{\partial}{\partial \theta} - \cot \theta \frac{\partial}{\partial \phi} \right) Y_l^l(\theta, \phi) = \hbar\sqrt{2l} Y_l^{l-1}(\theta, \phi) \quad (6.115)$$

This easily allows us to calculate the spherical harmonic Y_l^{l-1} by differentiating the harmonic Y_l^l . This process can be repeated till the spherical harmonics for states $|l, m_l\rangle$ are determined for all $2l+1$ values of m_l . The result for $m_l \geq 0$ is

$$Y_l^{m_l}(\theta, \phi) = \frac{(-1)^l}{2^l l!} \sqrt{\frac{(2l+1)(l+m_l)!}{4\pi(l-m_l)!}} e^{im_l\phi} \frac{1}{\sin^{m_l} \theta} \frac{d^{l-m_l}}{d(\cos \theta)^{l-m_l}} \sin^{2l} \theta \quad (6.116)$$

For $m_l < 0$, the spherical harmonics are given by

$$Y_l^{-m_l}(\theta, \phi) = (-1)^{m_l} Y_l^{m_l*}(\theta, \phi) \quad (6.117)$$

Spherical harmonics for $l = 0, 1, 2$ are as follows

$$\begin{aligned} Y_0^0(\theta, \phi) &= \sqrt{\frac{1}{4\pi}} \\ Y_1^{\pm 1}(\theta, \phi) &= \mp \sqrt{\frac{3}{8\pi}} e^{\pm i\phi} \sin \theta \\ Y_1^0(\theta, \phi) &= \sqrt{\frac{3}{4\pi}} \cos \theta \\ Y_2^{\pm 2}(\theta, \phi) &= \sqrt{\frac{15}{32\pi}} e^{\pm 2i\phi} \sin^2 \theta \\ Y_2^{\pm 1}(\theta, \phi) &= \mp \sqrt{\frac{15}{8\pi}} e^{\pm i\phi} \sin \theta \cos \theta \\ Y_2^0(\theta, \phi) &= \sqrt{\frac{5}{16\pi}} (3 \cos^2 \theta - 1) \end{aligned} \quad (6.118)$$

We have found that states of well-defined orbital angular momentum have wavefunctions of the form (6.95)

$$\psi_{lm_l} = R(r) Y_l^{m_l}(\theta, \phi) \quad (6.119)$$

We observe that the radial wavefunction $R(r)$ is not uniquely determined. This implies that a measurement of angular momentum does not uniquely determine the state of the particle. To uniquely pin down the state (while preserving definiteness of angular momentum), we need to observe an additional physical quantity. This observable should be such that its commutator with \hat{L}^2 and \hat{L}_z is zero. Clearly, the observable must be some function of position and momentum operators. To construct this observable, we first evaluate the

commutators of the position and momentum operators with the angular momentum operators. These can be verified to be

$$\begin{aligned}\left[\hat{L}_i, \hat{x}_j\right] &= i\hbar \epsilon_{ijk} \hat{x}_k \\ \left[\hat{L}_i, \hat{p}_j\right] &= i\hbar \epsilon_{ijk} \hat{p}_k\end{aligned}\quad (6.120)$$

For instance, for \hat{L}_x , \hat{y} and \hat{p}_y , these reduce to

$$\begin{aligned}\left[\hat{L}_x, \hat{y}\right] &= i\hbar \hat{z} \\ \left[\hat{L}_x, \hat{p}_y\right] &= i\hbar \hat{p}_z\end{aligned}\quad (6.121)$$

Using these commutators, it is easy to verify that the commutator of all three angular momentum operators \hat{L}_i with operators $\hat{\mathbf{p}}^2 = \hat{p}_x^2 + \hat{p}_y^2 + \hat{p}_z^2$ and $\hat{\mathbf{r}}^2 = \hat{x}^2 + \hat{y}^2 + \hat{z}^2$ are zero. Let us check this for \hat{L}_z and $\hat{\mathbf{p}}^2$

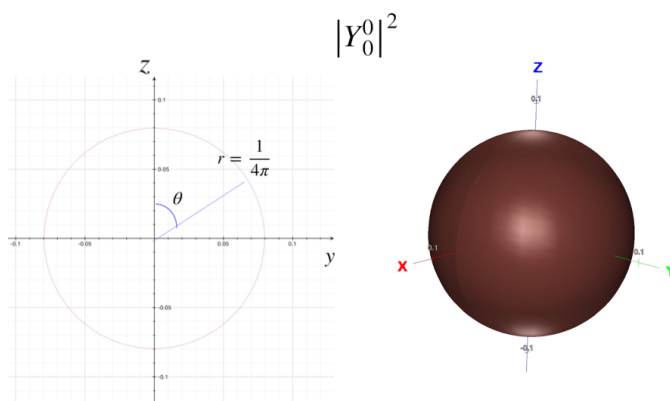
$$\begin{aligned}\left[\hat{L}_z, \hat{\mathbf{p}}^2\right] &= \left[\hat{L}_z, \hat{p}_x^2\right] + \left[\hat{L}_z, \hat{p}_y^2\right] + \left[\hat{L}_z, \hat{p}_z^2\right] \\ &= \hat{p}_x \left[\hat{L}_z, \hat{p}_x\right] + \left[\hat{L}_z, \hat{p}_x\right] \hat{p}_x + \hat{p}_y \left[\hat{L}_z, \hat{p}_y\right] + \left[\hat{L}_z, \hat{p}_y\right] \hat{p}_y \\ &= i\hbar \hat{p}_x \hat{p}_y + i\hbar \hat{p}_y \hat{p}_x - i\hbar \hat{p}_y \hat{p}_x - i\hbar \hat{p}_x \hat{p}_y \\ &= 0\end{aligned}\quad (6.122)$$

Similarly, it is easily checked that $[\hat{L}_z, \hat{\mathbf{r}}^2] = 0$. Since there is nothing special about the z direction, the commutators of \hat{L}_x and \hat{L}_y with $\hat{\mathbf{p}}^2$ and $\hat{\mathbf{r}}^2$ should also be zero. In particular, the commutator of \hat{L}_i with any function $f(\hat{\mathbf{r}}^2)$ of $\hat{\mathbf{r}}^2$ will be zero. Since the commutator of \hat{L}_i with $\hat{\mathbf{p}}^2$ is zero, there exist simultaneous eigenstates of the magnitude of angular momentum, one of its components and the *magnitude of momentum*. Therefore, if a measurement of angular momentum is followed by a measurement of the magnitude of momentum, we should end up with a well-defined state. Experimentally, we could instead measure the kinetic energy of the particle (which is proportional to the square of the magnitude of momentum). This will once again lead to a well-defined state.

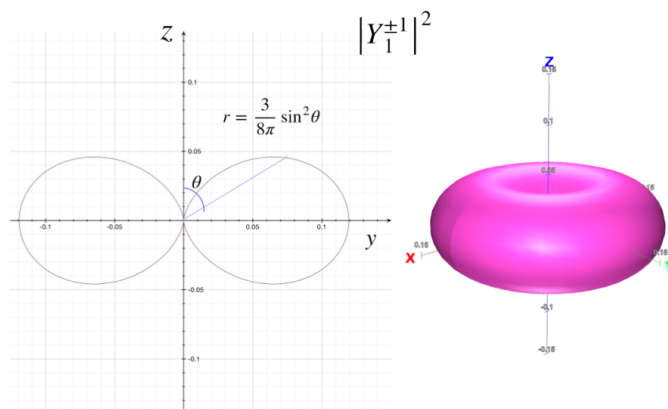
In states of well-defined angular momentum, it is useful to have an intuition for the angular part of the wavefunction, $Y_l^{m_l}(\theta, \phi)$. One way to develop this intuition is as follows. Given a state of definite angular momentum, we can ask for the probability of detecting the particle in a certain direction, *irrespective of the radial distance of the point*. For instance, we could visualize an array of detectors positioned within a narrow cone with a certain solid angle. If these detectors are switched on, what is the probability that the particle is detected within this cone? The answer is clearly proportional to $|Y_l^{m_l}(\theta, \phi)|^2$, since the radial dependence will be integrated out in computing this probability. Then, we need to be able to visualize the modulus squared of the spherical harmonics. This is achieved using *polar plots*. Since the modulus squared is independent of the angle ϕ (which appears as a phase in the spherical harmonics), the distribution will be symmetric under a rotation about the z -axis. So, it is sufficient to visualize the distribution in say the $z-y$ plane, and then rotate it about the z -axis to obtain the three dimensional distribution. In this plane, let $|Y_l^{m_l}(\theta, \phi)|^2 = f(\theta)$. To plot this distribution in this plane, for every angle θ , place a dot in this plane which is at a distance $f(\theta)$ from the origin. When such dots are placed for all values of θ , a curve will be obtained in this plane. Mathematically, this curve will be described by the equation

$$r = f(\theta)\quad (6.123)$$

For instance, let us plot $|Y_0^0|^2 = 1/4\pi$. The equation for this curve in the $z-y$ plane is $r = 1/4\pi$, which is a circle of radius $1/4\pi$. Rotating about the z -axis will generate a sphere of the same radius

Figure 6.3: The $l = 0, m_l = 0$ orbital.

Similarly, consider $|Y_1^{\pm 1}|^2 = (3/8\pi) \sin^2 \theta$. The two and three dimensional figures are as follows

Figure 6.4: The $l = 1, m_l = \pm 1$ orbital.

A few other such ‘orbitals’ are shown below

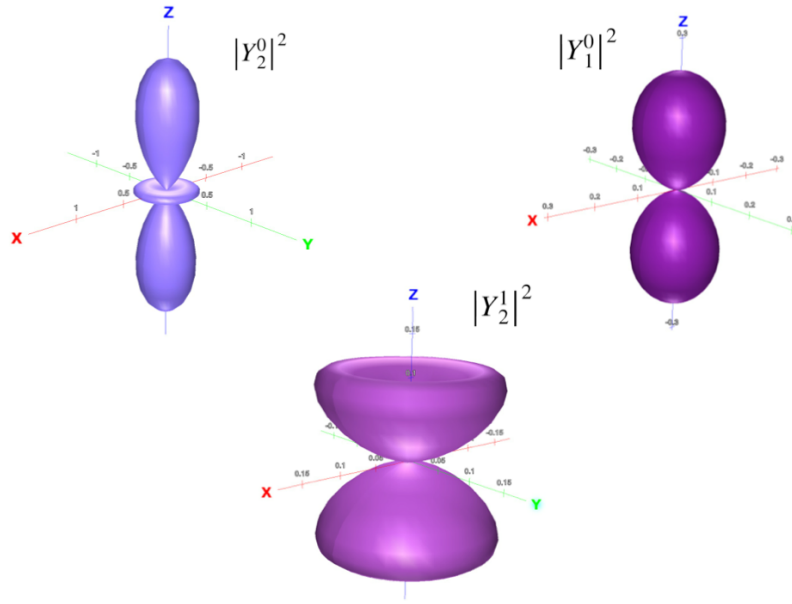


Figure 6.5: The $l = 1, m_l = 0$; $l = 2, m_l = 1$ and $l = 1, m_l = 0$ orbitals.

6.3 Spherically symmetric interaction: Hydrogen atom

A large class of interactions between physical systems are spherically symmetric. Such an interaction has a potential energy function associated with it which depends only on the radial separation between the systems. For instance, the interaction of an electron with the proton in a Hydrogen atom is spherically symmetric, described by the potential energy function

$$V(r) = \frac{-e^2}{4\pi\epsilon_0 r} \quad (6.124)$$

where e is the magnitude of the electron/proton charge and r is the electron-proton distance. Consider the quantum Hamiltonian for a system with a spherically symmetric potential energy function $V(r)$

$$\hat{H} = \frac{1}{2m}\hat{\mathbf{p}}^2 + V(\hat{r}) \quad (6.125)$$

Here the operator $V(\hat{r})$ is understood to be a function of $\hat{\mathbf{r}}^2 = \hat{x}^2 + \hat{y}^2 + \hat{z}^2$ (the same function as the classical potential energy function). Since the commutator of angular momentum operators with $\hat{\mathbf{r}}^2$ is zero, their commutator with $V(\hat{r})$ will also be zero. Then, given that $[\hat{L}_i, \hat{\mathbf{p}}^2] = 0$, it follows that $[\hat{H}, \hat{L}_i] = 0$. Then, there exist simultaneous eigenstates of \hat{H} , \hat{L}^2 and \hat{L}_z

$$\begin{aligned} \hat{H} |E, l, m_l\rangle &= E |E, l, m_l\rangle \\ \hat{L}^2 |E, l, m_l\rangle &= l(l+1)\hbar^2 |E, l, m_l\rangle \\ \hat{L}_z |E, l, m_l\rangle &= m_l\hbar |E, l, m_l\rangle \end{aligned} \quad (6.126)$$

These states will have wavefunctions $\psi_{Elm_l} = \langle \mathbf{x} | E, l, m_l \rangle$, such that

$$\psi_{Elm_l} = R_E(r)Y_l^{m_l}(\theta, \phi) \quad (6.127)$$

The energy eigenvalue equation in (6.126) has the standard form (6.28)

$$-\frac{\hbar^2}{2m}\nabla^2\psi_{Elm_l} + V(r)\psi_{Elm_l} = E\psi_{Elm_l} \quad (6.128)$$

where $V(r)$ is the spherically symmetric classical potential energy function. Because of spherical symmetry, it is useful to cast (6.129) in terms of spherical polar coordinates, by expressing the Laplacian operator ∇^2 in these coordinates

$$\frac{-\hbar^2}{2m} \frac{1}{r^2} \left(\frac{\partial}{\partial r} r^2 \frac{\partial}{\partial r} + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \phi^2} + \frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) \right) \psi_{Elm_l} + V(r) \psi_{Elm_l} = E \psi_{Elm_l} \quad (6.129)$$

Let us isolate the angular part of the left hand side of the equation. It is

$$\left[\frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \phi^2} + \frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) \right] \psi_{Elm_l} = \frac{-1}{\hbar^2} \langle \mathbf{x} | \hat{L}^2 | E, l, m_l \rangle \quad (6.130)$$

where we have identified the angular part with the matrix element of the operator \hat{L}^2 in (6.87). Since the state is an eigenstate of \hat{L}^2 (6.126), therefore this gives

$$\left[\frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \phi^2} + \frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) \right] \psi_{Elm_l} = -l(l+1) \psi_{Elm_l} \quad (6.131)$$

Substituting this in (6.129) and using the decomposition (6.127), we get the *Radial equation*

$$\frac{-\hbar^2}{2m} \left(\frac{1}{r^2} \frac{d}{dr} r^2 \frac{d}{dr} - \frac{l(l+1)}{r^2} \right) R_E(r) + V(r) R_E(r) = E R_E(r) \quad (6.132)$$

This equation, with appropriate boundary conditions, gives the energy eigenvalues and eigenfunctions (radial part). If the potential is attractive, then for bound states, we need to impose the boundary condition $\lim_{r \rightarrow \infty} R_E(r) = 0$. This will naturally give a discrete set of energy eigenvalues. This equation can be put in a simpler form by introducing an auxiliary function $U_E(r)$

$$R_E(r) = \frac{U_E(r)}{r} \quad (6.133)$$

In terms of this function, the radial eigenvalue equation is remarkably similar to the one-dimensional energy eigenvalue equation (5.188)

$$\frac{-\hbar^2}{2m} \frac{d^2 U_E(r)}{dr^2} + \left(V(r) + \frac{l(l+1)\hbar^2}{2mr^2} \right) U_E(r) = E U_E(r) \quad (6.134)$$

There are important differences, however. In addition to the potential energy term, there is a ‘centrifugal potential’ term, which reminds us that the problem is three dimensional. This term arises in classical physics also, when one writes the expression for energy of a particle under the influence of a central force. Also, the radial coordinate is restricted to the range $0 \leq r < \infty$, unlike the one-dimensional coordinate x which had the range $-\infty < x < \infty$. Then, for bound states, the appropriate boundary condition on $U_E(r)$ is that $U_E(0) = 0$ and $\lim_{r \rightarrow \infty} U_E(r) = 0$. The normalization for $U_E(r)$ is obtained from that of the radial wavefunction $R_E(r)$. Given (6.109) and (6.133), it follows that

$$\int_0^\infty dr |U_E(r)|^2 = 1 \quad (6.135)$$

which is similar to the normalization for a one-dimensional wavefunction.

A trivial case of a spherically symmetric interaction is a free particle, which does not experience any interaction. It follows that for a free particle, we can construct states of well-defined energy and angular momentum. For such states, the radial wavefunction can now be obtained from (6.134) with $V(r) = 0$

$$\frac{-\hbar^2}{2m} \frac{d^2 U_E(r)}{dr^2} + \frac{l(l+1)\hbar^2}{2mr^2} U_E(r) = E U_E(r) \quad (6.136)$$

Equation (6.136) can be solved in terms of spherical Bessel functions (we will not discuss the solution here). The important point to note is that for a free particle, we can either states of well-defined energy and momentum, *or* well-defined energy and angular momentum. The wavefunctions for the two sets of states will be completely different. Either of the two sets can be used as a basis in the Hilbert space of states.

We now determine the energy eigenvalues and wavfunctions for a particle under the influence of a Coulomb potential energy function (6.124). This has immediate application to the Hydrogen atom, though

it can be in principle applied to other systems as well, interacting via a Coulomb interaction (such as a hypothetical positronium atom, consisting of an electron bound to a positron. Such a system, though, is unstable and will annihilate to produce a pair of photons). In case of a Hydrogen atom, the mass parameter m in the potential (6.124) will be the reduced mass of the electron-proton system. For the Coulomb potential, the eigenvalue equation (6.134) reduces to

$$\frac{d^2 U_E(r)}{dr^2} + \frac{2m}{\hbar^2} \left(E + \frac{e^2}{4\pi\epsilon_0 r} - \frac{l(l+1)\hbar^2}{2mr^2} \right) U_E(r) = 0 \quad (6.137)$$

For bound states, the energy eigenvalues will be negative. We identify a suitable length scale in the problem, in terms of which we can measure the radial coordinate r . The energy eigenvalue E induces a length scale $a = \sqrt{\hbar^2/2m(-E)}$. Let us analyze the asymptotic behavior of this equation, for $r \gg a$. In this limit, the differential equation reduces to

$$\frac{d^2 U_E(r)}{dr^2} - \frac{2m}{\hbar^2} |E| U_E(r) = 0 \quad (6.138)$$

where $|E| = -E$, since we are solving for bound states, for which $E < 0$. The general solution to this equation has the form

$$U_E(r) = Ae^{-r/a} + Be^{r/a} \quad (6.139)$$

Since the wavefunction has to asymptotically vanish, $B = 0$. Then, the solution to (6.138) for large r must reduce to

$$U_E(r) \rightarrow e^{-r/a} \quad r \gg a \quad (6.140)$$

To solve (6.138), we introduce a dimensionless coordinate $\rho = r/a$. Then, for $\rho \gg 1$, the solution has the form $U_E \sim e^{-\rho}$. We factor out this dependence by defining a new function v_E such that $U_E = e^{-\rho} v_E$. Substituting this in (6.138), we get the following differential equation for v_E in terms of the variable ρ

$$\frac{d^2 v_E}{d\rho^2} - 2 \frac{dv_E}{d\rho} + \left[\left(\frac{e^2 \lambda}{4\pi\epsilon_0} \right) \frac{1}{\rho} - \frac{l(l+1)}{\rho^2} \right] v_E = 0 \quad (6.141)$$

where

$$\lambda = \left(\frac{2m}{\hbar^2(-E)} \right)^{1/2} \quad (6.142)$$

We solve this differential equation using the *Frobenius method*. Taking into account the behavior of the coefficient of v_E near $\rho = 0$, we assume a series solution of the form

$$v_E = \rho^{l+1} \sum_{k=0}^{\infty} C_k \rho^k \quad (6.143)$$

Substituting in (6.141), we get the following recursion relation

$$\frac{C_{k+1}}{C_k} = \frac{(-e^2 \lambda)/(4\pi\epsilon_0) + 2(k+l+1)}{(k+l+2)(k+l+1) - l(l+1)} \quad (6.144)$$

We analyze the behavior of the series in (6.143) for large k . As $k \rightarrow \infty$, the recursion relation tends to

$$\frac{C_{k+1}}{C_k} \rightarrow \frac{2}{k} \quad (6.145)$$

We need to figure out what kind of a function is represented by an infinite series $\sum_k C_k \rho^k$ with recursion (6.145). Inspection shows that a function of the form $f(\rho) = \rho^m e^{2\rho}$ will satisfy this relation. Then, the asymptotic form of v_E will be

$$v_E \rightarrow \rho^m e^{2\rho} \quad (6.146)$$

which will give the following asymptotic form for U_E

$$U_E \rightarrow \rho^m e^\rho \quad (6.147)$$

which diverges as $\rho \rightarrow \infty$. Since U_E should vanish at infinity, the recursion relation (6.144) must terminate for some finite k . Let us denote this by k_M . This is the value of k for which the numerator of (6.144) vanishes. That is

$$\frac{e^2 \lambda}{4\pi\epsilon_0} = 2(k_M + l + 1) \quad (6.148)$$

The energy eigenvalues are obtained from the relation (6.142)

$$E = - \left(\frac{e^2}{4\pi\epsilon_0} \right)^2 \frac{m}{2\hbar^2(k_M + l + 1)^2} \quad (6.149)$$

Clearly, the combination $k_M + l + 1$ is an integer. Let us denote this by n . Then, the energy eigenvalues are given by

$$E_n = - \left(\frac{me^4}{32\pi^2\epsilon_0^2\hbar^2} \right) \frac{1}{n^2} \quad (6.150)$$

Now, k_M can take values $0, 1, 2, \dots$ and l can take values $1, 2, \dots$. Then, the quantum number n can take values $n = 1, 2, 3, \dots$. n is called the *Principal quantum number*. The energy eigenvalue depends only on n . Given n , l is given by

$$l = n - k_M - 1 \quad (6.151)$$

For a given n , since l cannot be negative, k_M can take values $0, 1, 2, \dots, n - 1$. Correspondingly, l can take values $n - 1, n - 2, \dots, 1, 0$. Equivalently, we can view this as follows. Given an energy eigenvalue in terms of the Principal quantum number n , there are n degenerate states corresponding to $l = 0, 1, 2, \dots, n - 1$. We denote these states as $|n, l, m_l\rangle$. These are just the states $|E, l, m_l\rangle$, the simultaneous eigenstates of energy and angular momentum. The wavefunctions for these states are given by

$$\psi_{nlm_l}(r, \theta, \phi) = R_{nl}(r)Y_l^{m_l}(\theta, \phi) \quad n = 1, 2, 3, \dots; \quad l = 0, 1, 2, \dots, n - 1 \quad (6.152)$$

where

$$R_{nl}(r) = \frac{U_{nl}(r)}{r} \quad (6.153)$$

with the auxiliary function U_{nl} given by

$$U_{nl}(r) = v_{nl}(r/a_n)e^{-r/a_n} \quad (6.154)$$

where $a_n = \sqrt{\hbar^2/2m(-E_n)}$ is the length scale. The function $v_{nl}(r/a_n)$ is given by the (terminating) series (6.143)

$$v_{nl} = (r/a_n)^{l+1} \sum_{k=0}^{k_M} C_k (r/a_n)^k \quad (6.155)$$

where

$$k_M = n - 1 - l \quad (6.156)$$

From (6.148), the recursion relations (6.144) can be written as

$$\frac{C_{k+1}}{C_k} = \frac{-2n + 2(k + l + 1)}{(k + l + 2)(k + l + 1) - l(l + 1)} \quad (6.157)$$

Let us analyze the first few energy eigenstates. From (6.150), the ground state energy is obtained for $n = 1$

$$E_1 = - \frac{me^4}{32\pi^2\epsilon_0^2\hbar^2} \quad (6.158)$$

Applied to an electron in hydrogen (with m being the reduced mass of the electron and the proton), the ground state energy is calculated to be (approximately)

$$E_1 = -13.6 \text{ eV} \quad (6.159)$$

The higher energy eigenvalues are then

$$E_n = -\frac{13.6}{n^2} \text{ eV} \quad (6.160)$$

Coming back to the ground state, it follows from (6.152) that l can take only one value, $l = 0$. Then, in the ground state, an electron has zero angular momentum. To determine the wavefunction, it is convenient to express the length scale $a_n = \sqrt{\hbar^2/2m(-E_n)}$ appearing in (6.154) as

$$a_n = na_0 \quad (6.161)$$

where

$$a_0 = \frac{4\pi\hbar^2}{me^2} \quad (6.162)$$

For an electron, the length scale a_0 is called the *Bohr Radius* (for historical reasons), and is approximately $a_0 = 0.55$ Angstrom. Then, the ground state wavefunction is determined from (6.152) with $n = 0$ and $l = 0$. For these values, it follows from (6.156) that the series (6.155) will terminate for $k_M = 0$. Then, from (6.153) and (6.154), the radial wavefunction will be (upto a constant)

$$R_{10} = c_0 e^{-r/a_0} \quad (6.163)$$

After normalization imposed by (6.109), the complete wavefunction for the ground state is easily found to be

$$\psi_{100} = \left(\frac{1}{\pi a_0^3}\right)^{1/2} e^{-r/a_0} \quad (6.164)$$

We now determine the wavefunctions for the first excited state(s). For $n = 2$, the first excited state energy is

$$\begin{aligned} E_2 &= -\frac{13.6}{2^2} \text{ eV} \\ &= -3.4 \text{ eV} \end{aligned} \quad (6.165)$$

For $n = 2$, l can take values 0 and 1. For $l = 0$, the series (6.155) will truncate for $k_M = 2 - 1 - 0 = 1$ (from (6.156))

$$v_{20} = \frac{r}{2a_0} \left(C_0 + C_1 \frac{r}{2a_0} \right) \quad (6.166)$$

From recursion relations (6.157), it follows that $C - 1 = -C_0$. This will give

$$v_{20} = \frac{r}{2a_0} C_0 \left(1 - \frac{r}{2a_0} \right) \quad (6.167)$$

After normalization, the first excited state $|2, 0, 0\rangle$ is described by the wavefunction

$$\psi_{200} = \left(\frac{1}{32\pi a_0^3}\right)^{1/2} \left(2 - \frac{r}{a_0} \right) e^{-r/2a_0} \quad (6.168)$$

For the states $|2, 1, m_l\rangle$, since $l = 1$, the series (6.155) will truncate for $k_M = 2 - 1 - 1 = 0$. Then,

$$v_{21} = C_0 (r/2a_0)^2 \quad (6.169)$$

After normalization, the wavefunctions are obtained as

$$\psi_{210} = \left(\frac{1}{32\pi a_0^3}\right)^{1/2} \frac{r}{a_0} e^{-r/2a_0} \cos \theta \quad (6.170)$$

and

$$\psi_{21\pm 1} = \mp \left(\frac{1}{64\pi a_0^3} \right)^{1/2} \frac{r}{a_0} e^{-r/2a_0} \sin \theta e^{\pm i\phi} \quad (6.171)$$

Let us take a closer look at the series (6.155). The factor $(r/a_n)^{l+1}$ multiplies a polynomial in $(r/a_n)^{l+1}$ of degree $k_M = n - l - 1$. This polynomial is called the *associated Laguerre polynomial*, $L_{n-l-1}^{2l+1}(2r/a_n)$. In terms of this polynomial, the radial wavefunction R_{nl} will have the form (upto an overall normalization factor)

$$R_{nl} \sim \left(\frac{r}{na_0} \right)^l L_{n-l-1}^{2l+1} \left(\frac{2r}{na_0} \right) e^{-r/na_0} \quad (6.172)$$

The associated Laguerre polynomials are generated using the recursion relations (6.157). To visualize the probability density distributions corresponding to these wavefunctions, one uses the following visualization. Imagine a distribution of points in space whose density at any point is proportional to $|\psi_{nlm_l}(\mathbf{x})|^2$ at that point. These points will then form a ‘dust’-like distribution, the density of the dust at any point proportional to the probability of detecting the electron at that point. Alternatively, one can associate an ‘intensity’ distribution in which every point in space is given an intensity (brightness) proportional to the probability density at that point. such an intensity distribution for the first few energy eigenstates is illustrated below

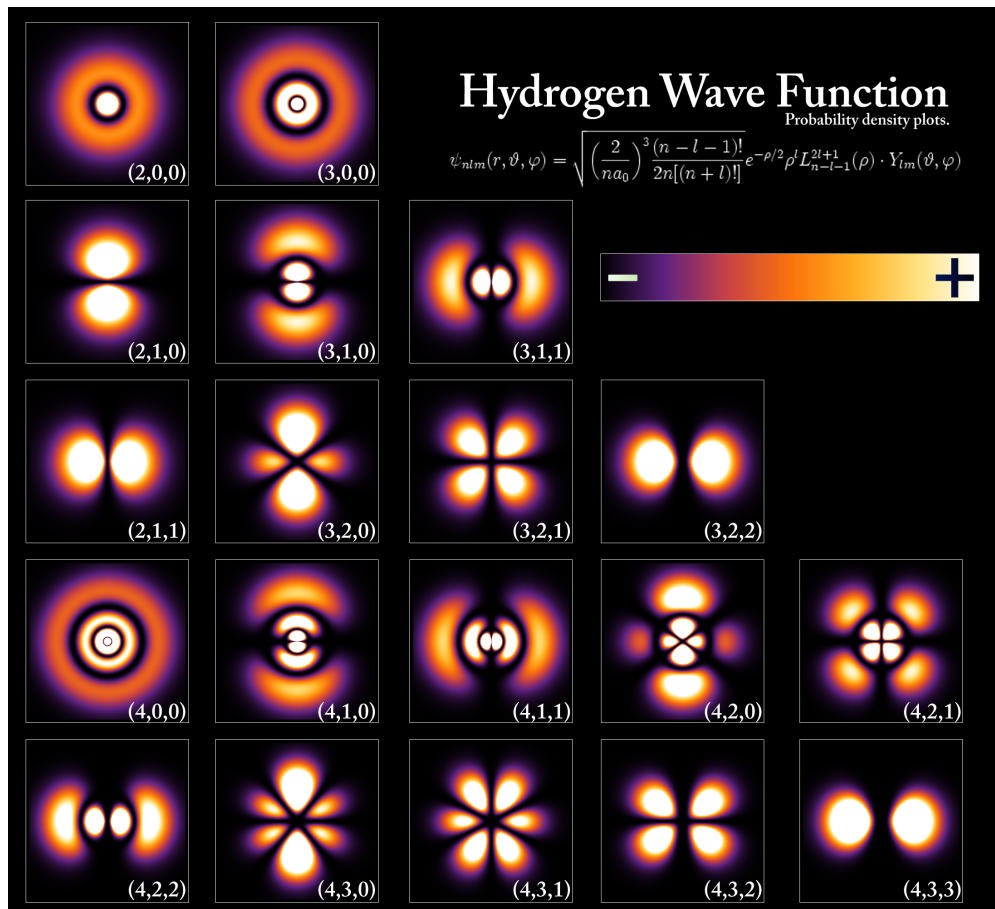


Figure 6.6: Electron probability distributions for the first few energy eigenstates of Hydrogen. The ‘intensity’ at any point is proportional to $|\psi_{nlm_l}|^2$.