Quantum Mechanics Made Simple: Lecture Notes

Weng Cho CHEW [1]

December 6, 2016

[1]The author is with U of Illinois, Urbana-Champaign.

Contents

Preface

This course ECE 487 was taught again during the Fall of 2016. Hence, the lecture notes are slightly modified and updated. I have chosen to reduce the number of chapters taught, so that longer time can be spent on eschewing some of the physical concepts. The chapters skipped are marked with asterisks. They are chapters 10, 13, and 14. Hence, the reader can skip these chapters on first reading. Also, for readers who are interested in studying quantum electromagnetics and quantum information, they can read the first five chapters, chapters 7, 9, 11, 12, and 15.

Weng Cho CHEW
December 2016

    This set of supplementary lecture notes is the outgrowth of a course I taught, ECE 487, Quantum Electronics, at ECE Department, University of Illinois at Urbana-Champaign. It was intended to teach quantum mechanics to undergraduate students as well as graduate students. The primary text book for this course is Quantum Mechanics for Scientists and Engineers by D.A.B. Miller that I have learnt a great deal from. But where I feel the book to be incomplete, I supplement them with my lecture notes. I try to reach into first principles as much as I could with these lecture notes. The only background needed for reading these notes is undergraduate wave physics, and linear algebra.

    I would still recommend using Miller's book as the primary text book for such a course, and use these notes as supplementary to teach this topic to undergraduates.

    Also, I am indebted to Aiyin Liu who has helped modified part of these notes, especially Chapter 2.

Weng Cho CHEW
June 2013

Acknowledgements

Chapter 1

Introduction

1.1 Introduction

Quantum mechanics (or quantum physics) is an important intellectual achievement of the 20th century. It is one of the more sophisticated fields in physics that has affected our understanding of nano-meter length scale systems important for chemistry, materials, optics, electronics, and quantum information. The existence of orbitals and energy levels in atoms can only be explained by quantum mechanics. Quantum mechanics can explain the behaviors of insulators, conductors, semi-conductors, and giant magneto-resistance. It can explain the quantization of light and its particle nature in addition to its wave nature (known as particle-wave duality). Quantum mechanics can also explain the radiation of hot body or black body, and its change of color with respect to temperature. It explains the presence of holes and the transport of holes and electrons in electronic devices.

The first observation of something quantum was probably by Michael Faraday in 1838 on cathode rays: the fact that when electrons bombard a cathode, light is given out. The next quantum phenomenon was the black body radiation of Gustav Kirchhoff in 1859-1860. Ludwig Boltzmann conjectured that atomic energy levels should be discrete around 1877. Heinrich Hertz first observed the photo-electric effect in 1887 which was later explained by Einstein in 1905 using the particle nature of light. Examples of famous physicists that contribute to quantum mechanics were Max Planck, Albert Einstein, Niels Bohr, Louis de Broglie, Max Born, Paul Dirac, Werner Heisenberg, Wolfgang Pauli, Erwin Schrdinger, Richard Feynman.

Marx Planck proposed the Planck radiation law in 1900 that

$$I(\nu) = \frac{2h\nu^3}{c^2} \frac{1}{e^{h\nu/k_b T} - 1} \qquad (1.1.1)$$

where $h \approx 6.626 \times 10^{-34}$J $\cdot$ s is Planck constant, c is velocity of light, ν the frequency of light, $k_b = 1.38064852(79) \times 10^{-23}$J/K is Boltzmann constant, and T the temperature in Kelvin. The above radiation law is obtained by assuming that light forms quanta of energy each of which is given by $h\nu$ when the quantum theory of light was not well understood. When the Planck constant becomes negligibly small, implying that the quantization of light

1

Chapter 1

Introduction

1.1 Introduction

Quantum mechanics (or quantum physics) is an important intellectual achievement of the 20th century. It is one of the more sophisticated fields in physics that has affected our understanding of nano-meter length scale systems important for chemistry, materials, optics, electronics, and quantum information. The existence of orbitals and energy levels in atoms can only be explained by quantum mechanics. Quantum mechanics can explain the behaviors of insulators, conductors, semi-conductors, and giant magneto-resistance. It can explain the quantization of light and its particle nature in addition to its wave nature (known as particle-wave duality). Quantum mechanics can also explain the radiation of hot body or black body, and its change of color with respect to temperature. It explains the presence of holes and the transport of holes and electrons in electronic devices.

The first observation of something quantum was probably by Michael Faraday in 1838 on cathode rays: the fact that when electrons bombard a cathode, light is given out. The next quantum phenomenon was the black body radiation of Gustav Kirchhoff in 1859-1860. Ludwig Boltzmann conjectured that atomic energy levels should be discrete around 1877. Heinrich Hertz first observed the photo-electric effect in 1887 which was later explained by Einstein in 1905 using the particle nature of light. Examples of famous physicists that contribute to quantum mechanics were Max Planck, Albert Einstein, Niels Bohr, Louis de Broglie, Max Born, Paul Dirac, Werner Heisenberg, Wolfgang Pauli, Erwin Schrdinger, Richard Feynman.

Marx Planck proposed the Planck radiation law in 1900 that

$$I(\nu) = \frac{2h\nu^3}{c^2} \frac{1}{e^{h\nu/k_b T} - 1} \qquad (1.1.1)$$

where $h \approx 6.626 \times 10^{-34}$J $\cdot$ s is Planck constant, c is velocity of light, ν the frequency of light, $k_b = 1.38064852(79) \times 10^{-23}$J/K is Boltzmann constant, and T the temperature in Kelvin. The above radiation law is obtained by assuming that light forms quanta of energy each of which is given by $h\nu$ when the quantum theory of light was not well understood. When the Planck constant becomes negligibly small, implying that the quantization of light

1

is unimportant, the above formula becomes

$$I(\nu) \approx \frac{2\nu^2}{c^2} k_b T \tag{1.1.2}$$

which becomes Rayleigh-Jeans law of radiation. This law fails to predict radiation of a black body correctly when the frequency is high.
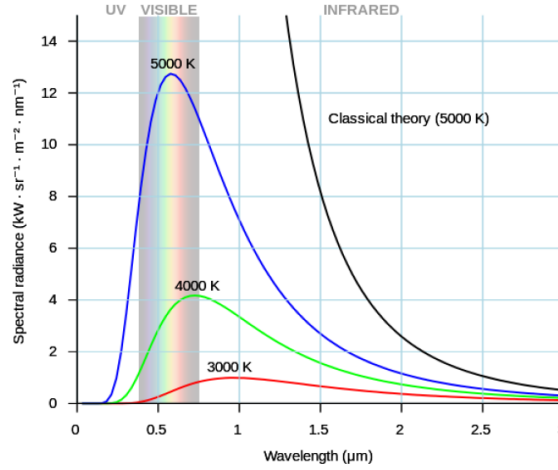
Figure 1.1: Planck black body radiation law: With decreasing temperature, the peak of the black body radiation curve shifts to longer wavelengths and also has lower intensities. The black body radiation curves (1862) at left are also compared with the early, classical limit model of Rayleigh and Jeans (1900) shown at right. The short wavelength side of the curves was already approximated in 1896 by the Wien distribution law. Courtesy of Darth Kule and Wikipedia.

Quantum mechanics has played an important role in photonics, quantum electronics, nano- and micro-electronics, nano- and quantum optics, quantum computing, quantum communication and crytography, solar and thermo-electricity, nano-electromechanical systems, etc. Many emerging technologies require the understanding of quantum mechanics; and hence, it is important that scientists and engineers understand quantum mechanics better.

In nano-technologies due to the recent advent of nano-fabrication techniques, nano-meter size systems are more common place. In electronics, as transistor devices become smaller, how the electrons behave in the device is quite different from when the devices are bigger: nano-electronic transport is quite different from micro-electronic transport. For small devices, the electron can travel for a distance without colliding with a lattice, when the electron travels through it, hence, giving rise to ballistic transport. For large devices, the electron collides many times with the lattice when traversing it, giving rise to collision dominated transport.

The quantization of electromagnetic field is important in the area of nano-optics and quantum optics. It explains how photons interact with atomic systems or materials. It also allows the use of electromagnetic or optical field to carry quantum information. Quantum

mechanics is certainly giving rise to interest in quantum information, quantum communication, quantum cryptography, and quantum computing. Moreover, quantum mechanics is also needed to understand the interaction of photons with materials in solar cells, as well as many topics in material science.

When two objects are placed close together, they experience a force called the Casimir force that can only be explained by quantum mechanics. (This force is there even when no electromagnetic field is present in the classical sense. Even at zero temperature, and "zero" field, a vacuum electromagnetic fluctuation exists, giving rise to attractive force between two neutral objects.) This is important for the understanding of micro/nano-electromechanical systems (M/NEMS). Moreover, the understanding of spins is important in spintronics, another emerging technology where giant magneto-resistance, tunneling magneto-resistance, and spin transfer torque are being used. It is obvious that the richness of quantum physics will greatly affect the future generation technologies in many aspects.

1.2   Quantum Mechanics is Bizarre

The development of quantum mechanics is a great intellectual achievement, but at the same time, it is bizarre. The reason is that quantum mechanics is quite different from classical physics. The development of quantum mechanics is likened to watching two players having a game of chess, but the observers have not a clue as to what the rules of the game are. By observations, and conjectures, finally the rules of the game are outlined. Often, equations are conjectured like conjurors pulling tricks out of a hat to match experimental observations. It is the interpretations of these equations that can be quite bizarre.

Quantum mechanics equations were postulated to explain experimental observations, but the deeper meanings of the equations often confused even the most gifted. Even though Einstein received the Nobel prize for his work on the photo-electric effect that confirmed that light energy is quantized, he himself was not totally at ease with the development of quantum mechanics as charted by the younger physicists. He was never comfortable with the probabilistic interpretation of quantum mechanics by Born and the Heisenberg uncertainty principle: "God doesn't play dice," was his statement assailing the probabilistic interpretation. He proposed "hidden variables" to explain the random nature of many experimental observations. He was thought of as the "old fool" by the younger physicists during his time.

Schrödinger came up with the bizarre "Schrödinger cat paradox" that showed the struggle that physicists had with quantum mechanics' interpretation. But with today's understanding of quantum mechanics, the paradox is a thing of yesteryear.

The latest twist to the interpretation in quantum mechanics is the parallel universe view that explains the multitude of outcomes of the prediction of quantum mechanics. All outcomes are possible, but with each outcome occurring in different universes that exist in parallel with respect to each other ...

1.3   The Wave Nature of a Particle—Wave-Particle Duality

The quantized nature of the energy of light was first proposed by Planck in 1900 to successfully explain the black body radiation. Einstein's explanation of the photoelectric effect (1905) further asserts the quantized nature of light, or light as a photon with a packet of energy given by[1]

$$E = \hbar\omega \tag{1.3.1}$$

where $\hbar = h/(2\pi)$ is the reduced Planck constant or Dirac constant, and h is the Planck's constant, given by

$$h \approx 6.626 \times 10^{-34} \text{Joule} \cdot \text{second} \tag{1.3.2}$$

However, it is well known that light is a wave since it can be shown to interfere as waves in the Newton ring experiment as far back as 1717. Hence, light exhibits wave-particle duality.

Electron was first identified as a particle by Thomson (1897). From the fact that a photon has energy $E = \hbar\omega$ and that it also has an energy related to its momentum p by $E = pc$ where c is the velocity of light,[2] De Broglie (1924) hypothesized that the wavelength of an electron, when it behaves like a wave, is

$$\lambda = \frac{h}{p} \tag{1.3.3}$$

where p is the electron momentum.[3] The wave nature of an electron is revealed by the fact that when electrons pass through a crystal, they produce a diffraction pattern. That can only be explained by the wave nature of an electron. This experiment was done by Davisson and Germer in 1927.[4] When an electron manifests as a wave, it is described by

$$\psi(z) \propto \exp(ikz) \tag{1.3.4}$$

where $k = 2\pi/\lambda$. Such a wave is a solution to[5]

$$\frac{\partial^2}{\partial z^2}\psi = -k^2\psi \tag{1.3.5}$$

---

[1]In the photoelectric effect, it was observed that electrons can be knocked off a piece of metal only if the light exceeded a certain frequency. Above that frequency, the electron gained some kinetic energy proportional to the excess frequency. Einstein then concluded that a packet of energy was associated with a photon that is proportional to its frequency.

[2]This follows from Einstein theory of relativity that says the $E^2 - (pc)^2 = (mc^2)^2$ where m is the rest mass of the particle. Photon has zero rest mass. When $p = 0$, this is the famous $E = mc^2$ formula.

[3]Typical electron wavelengths are of the order of nanometers. Compared to 400 nm of wavelength of blue light, they are much smaller. Energetic electrons can have even smaller wavelengths. Hence, electron waves can be used to make electron microscope whose resolution is much higher than optical microscope.

[4]Young's double slit experiment was conducted in early 1800s to demonstrate the wave nature of photons. Due to the short wavelengths of electrons, it was not demonstrated it until 2002 by Jonsson. But it has been used as a thought experiment by Feynman in his lectures.

[5]The wavefunction can be thought of as a "halo" that an electron carries that determine its underlying physical properties and how it interact with other systems.

A generalization of this to three dimensions yields

$$\nabla^2 \psi(\mathbf{r}) = -k^2 \psi(\mathbf{r}) \tag{1.3.6}$$

We can define, from de Broglie's formula that

$$p = \hbar k \tag{1.3.7}$$

where $\hbar = h/(2\pi)$.[6] Consequently, we arrive at an equation

$$-\frac{\hbar^2}{2m_0} \nabla^2 \psi(\mathbf{r}) = \frac{p^2}{2m_0} \psi(\mathbf{r}) \tag{1.3.8}$$

where

$$m_0 \approx 9.11 \times 10^{-31} \text{kg} \tag{1.3.9}$$

The expression $p^2/(2m_0)$ is the kinetic energy of an electron. Hence, the above can be considered an energy conservation equation.

The Schrödinger equation (1925) is motivated by further energy balance that total energy is equal to the sum of potential energy and kinetic energy. Defining the potential energy to be $V(\mathbf{r})$, the energy balance equation becomes

$$\left[-\frac{\hbar^2}{2m_0} \nabla^2 + V(\mathbf{r}) \right] \psi(\mathbf{r}) = E\psi(\mathbf{r}) \tag{1.3.10}$$

where E is the total energy of the system. The above is the time-independent Schrödinger equation. The *ad hoc* manner at which the above equation is arrived at usually bothers a beginner in the field. However, it predicts many experimental outcomes. It particular, it predicts the energy levels and orbitals of a trapped electron in a hydrogen atom with resounding success.

One can further modify the above equation in an *ad hoc* manner by noticing that other experimental finding shows that the energy of a photon is $E = \hbar\omega$. Hence, if we let

$$i\hbar \frac{\partial}{\partial t} \Psi(\mathbf{r}, t) = E\Psi(\mathbf{r}, t) \tag{1.3.11}$$

then

$$\Psi(\mathbf{r}, t) = e^{-i\omega t} \psi(\mathbf{r}) \tag{1.3.12}$$

Then we arrive at the time-dependent Schrödinger equation:

$$\left[-\frac{\hbar^2}{2m_0} \nabla^2 + V(\mathbf{r}) \right] \Psi(\mathbf{r}, t) = i\hbar \frac{\partial}{\partial t} \Psi(\mathbf{r}, t) \tag{1.3.13}$$

Another disquieting fact about the above equation is that it is written in terms of complex functions and numbers. In our prior experience with classical laws, they can all be written in real functions and numbers. We will later learn the reason for this.

---

[6]This is also called Dirac constant sometimes.

Mind you, in the above, the frequency is not unique. We know that in classical physics, the potential V is not unique, and we can add a constant to it, and yet, the physics of the problem does not change. So, we can add a constant to both sides of the time-independent Schrödinger equation (1.3.11), and yet, the physics should not change. Then the total E on the right-hand side would change, and that would change the frequency we have arrived at in the time-dependent Schrödinger equation. We will explain how to resolve this dilemma later on. Just like potentials, in quantum mechanics, it is the difference of frequencies that matters in the final comparison with experiments, not the absolute frequencies.

The setting during which Schrödinger equation was postulated was replete with knowledge of classical mechanics. It will be prudent to review some classical mechanics knowledge next.

Chapter 2

Classical Mechanics and Some Mathematical Preliminaries

2.1   Introduction

Quantum mechanics cannot be derived from classical mechanics, but classical mechanics can inspire quantum mechanics. Quantum mechanics is richer and more sophisticated than classical mechanics. Quantum mechanics was developed during the period when physicists had rich knowledge of classical mechanics. In order to better understand how quantum mechanics was developed in this environment, it is better to understand some fundamental concepts in classical mechanics. Classical mechanics can be considered as a special case of quantum mechanics. We will review some classical mechanics concepts here.

In classical mechanics, a particle moving in the presence of potential[1] $V(q)$ will experience a force given by

$$F(q) = -\frac{dV(q)}{dq} \tag{2.1.1}$$

where q represents the coordinate or the position of the particle. Hence, the particle can be described by the equations of motion

$$\frac{\mathrm{d}p}{\mathrm{d}t} = F(q) = -\frac{dV(q)}{dq}, \qquad \frac{\mathrm{d}q}{\mathrm{d}t} = p/m \tag{2.1.2}$$

For example, when a particle is attached to a spring and moves along a frictionless surface, the force the particle experiences is $F(q) = -kq$ where k is the spring constant. Then the equations of motion of this particle are

$$\frac{\mathrm{d}p}{\mathrm{d}t} = \dot{p} = -kq, \qquad \frac{\mathrm{d}q}{\mathrm{d}t} = \dot{q} = p/m \tag{2.1.3}$$

---

[1]The potential here refers to potential energy.
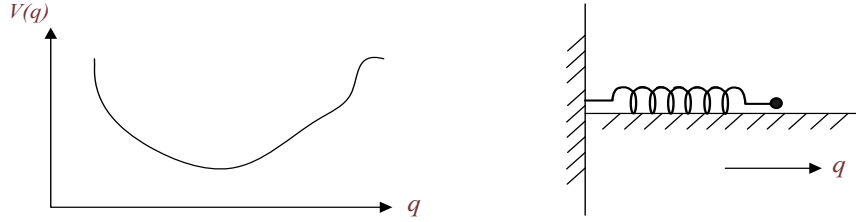
Figure 2.1: The left side shows a potential well in which a particle can be trapped. The right side shows a particle attached to a spring. The particle is subject to the force due to the spring, but it can also be described by the force due to a potential well.

Given p and q at some initial time t_0, one can integrate (2.1.2) or (2.1.3) to obtain p and q for all later time. A numerical analyst can think of that (2.1.2) or (2.1.3) can be solved by the finite difference method, where time-stepping can be used to find p and q for all later times. For instance, we can write the equations of motion more compactly as

$$\frac{d\mathbf{u}}{dt} = \mathbf{f}(\mathbf{u}) \tag{2.1.4}$$

where $\mathbf{u} = [p, q]^t$, and $\mathbf{f}$ is a general vector function of $\mathbf{u}$. It can be nonlinear or linear; in the event if it is linear, then $\mathbf{f}(\mathbf{u}) = \overline{\mathbf{A}} \cdot \mathbf{u}$.

Using finite difference approximation, we can rewrite the above as

$$\mathbf{u}(t + \Delta t) - \mathbf{u}(t) \doteq \Delta t \mathbf{f}(\mathbf{u}(t)),$$
$$\mathbf{u}(t + \Delta t) \doteq \Delta t \mathbf{f}(\mathbf{u}(t)) + \mathbf{u}(t) \tag{2.1.5}$$

The above can be used for time marching to derive the future values of $\mathbf{u}$ from past values.

The above equations of motion are essentially derived using Newton's law. However, there exist other methods of deriving these equations of motion. Notice that only two variables p and q are sufficient to describe the state of a particle.

2.2 Lagrangian Formulation

Another way to derive the equations of motion for classical mechanics is via the use of the Lagrangian and the principle of least action. A Lagrangian is usually defined as the difference between the kinetic energy and the potential energy, i.e.,

$$L(\dot{q}, q) = T - V \tag{2.2.1}$$

where $\dot{q}$ is the velocity. For a fixed t, q and $\dot{q}$ are independent variables, since $\dot{q}$ cannot be derived from q if it is only known at one given t. The equations of motion are derived from the principle of least action which says that $q(t)$ that satisfies the equations of motion between two times t_1 and t_2 should minimize the action integral

$$S = \int_{t_1}^{t_2} L(\dot{q}(t), q(t)) dt \tag{2.2.2}$$

Assuming that $q(t_1)$ and $q(t_2)$ are fixed, then the function $q(t)$ and $\dot{q}(t)$ that varies between t_1 and t_2 should minimize S, the action. In other words, a first order perturbation in q from the optimal answer that minimizes S should give rise to second order error in S. Hence, taking the first variation of (2.2.2), we have

$$\delta S = \delta \int_{t_1}^{t_2} L(\dot{q}, q)dt = \int_{t_1}^{t_2} L(\dot{q} + \delta\dot{q}, q + \delta q)dt - \int_{t_1}^{t_2} L(\dot{q}, q)dt$$

$$= \int_{t_1}^{t_2} \delta L(\dot{q}, q)dt = \int_{t_1}^{t_2} \left(\delta\dot{q}\frac{\partial L}{\partial \dot{q}} + \delta q\frac{\partial L}{\partial q} \right) dt = 0 \qquad (2.2.3)$$

In order to take the variation into the integrand, we have to assume that $\delta L(\dot{q}, q)$ is taken with constant time. At constant time, $\dot{q}$ and q are independent variables; hence, the partial derivatives in the next equality above follow. Using integration by parts on the first term, we have

$$\delta S = \delta q \left. \frac{\partial L}{\partial \dot{q}} \right|_{t_1}^{t_2} - \int_{t_1}^{t_2} \delta q\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}} \right) dt + \int_{t_1}^{t_2} \delta q\frac{\partial L}{\partial q}dt$$

$$= \int_{t_1}^{t_2} \delta q \left[-\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}} \right) + \frac{\partial L}{\partial q} \right] dt = 0 \qquad (2.2.4)$$

The first term vanishes because $\delta q(t_1) = \delta q(t_2) = 0$ because $q(t_1)$ and $q(t_2)$ are fixed. Since $\delta q(t)$ is arbitrary between t_1 and t_2, we must have

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}} \right) - \frac{\partial L}{\partial q} = 0 \qquad (2.2.5)$$

The above is called the Lagrange equation, from which the equation of motion of a particle can be derived. The derivative of the Lagrangian with respect to the velocity $\dot{q}$ is the momentum

$$p = \frac{\partial L}{\partial \dot{q}} \qquad (2.2.6)$$

The derivative of the Lagrangian with respect to the coordinate q is the force. Hence

$$F = \frac{\partial L}{\partial q} \qquad (2.2.7)$$

The above equation of motion is then

$$\dot{p} = F \qquad (2.2.8)$$

Equation (2.2.6) can be inverted to express $\dot{q}$ as a function of p and q, namely

$$\dot{q} = f(p, q) \qquad (2.2.9)$$

The above two equations can be solved in tandem to find the time evolution of p and q.

For example, the kinetic energy T of a particle is given by

$$T = \frac{1}{2}m\dot{q}^2 \tag{2.2.10}$$

Then from (2.2.1), and the fact that V is independent of $\dot{q}$,

$$p = \frac{\partial L}{\partial \dot{q}} = \frac{\partial T}{\partial \dot{q}} = m\dot{q} \tag{2.2.11}$$

or

$$\dot{q} = \frac{p}{m} \tag{2.2.12}$$

Also, from (2.2.1), (2.2.7), and (2.2.8), we have

$$\dot{p} = -\frac{\partial V}{\partial q} \tag{2.2.13}$$

The above pair, (2.2.12) and (2.2.13), form the equations of motion for this problem.

The above can be generalized to multidimensional problems. For example, for a one particle system in three dimensions, q_i has three degrees of freedom, and $i = 1, 2, 3$. (The q_i can represent x, y, z in Cartesian coordinates, but r, θ, ϕ in spherical coordinates.) But for N particles in three dimensions, there are $3N$ degrees of freedom, and $i = 1, \ldots, 3N$. The formulation can also be applied to particles constraint in motion. For instance, for N particles in three dimensions, q_i may run from $i = 1, \ldots, 3N - k$, representing k constraints on the motion of the particles. This can happen, for example, if the particles are constraint to move in a manifold (surface), or a line (ring) embedded in a three dimensional space.

Going through similar derivation, we arrive at the equation of motion

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}_i}\right) - \frac{\partial L}{\partial q_i} = 0 \tag{2.2.14}$$

In general, q_i may not have a dimension of length, and it is called the generalized coordinate (also called conjugate coordinate). Also, $\dot{q}_i$ may not have a dimension of velocity, and it is called the generalized velocity.

The derivative of the Lagrangian with respect to the generalized velocity is the generalized momentum (also called conjugate momentum), namely,

$$p_i = \frac{\partial L}{\partial \dot{q}_i} \tag{2.2.15}$$

The generalized momentum may not have a dimension of momentum. Hence, the equation of motion (2.2.14) can be written as

$$\dot{p}_i = \frac{\partial L}{\partial q_i} \tag{2.2.16}$$

Equation (2.2.15) can be inverted to yield an equation for $\dot{q}_i$ as a function of the other variables. This equation can be used in tandem (2.2.16) as time-marching equations of motion.

2.2.1 The Classical Harmonic Oscillator

As a simple example for the use of the Lagrangian formulation we investigate the harmonic oscillator. We take our model to be that sketched in Figure 2.1 in which a particle of mass m is attached to a spring of spring constant k and is constrained to move in the horizontal direction.

The very first step in composing the equations of motion consist of identifying the suitable coordinates for the system. In this case the horizontal distance is best referenced to the rest position of the spring. We shall measure the length that the spring is compressed or stretched by q. With this coordinate set up, we proceed to write down the Lagrangian.

$$L(q, \dot{q}) = T - V = \frac{1}{2}m\dot{q}^2 - \frac{1}{2}kq^2 \tag{2.2.17}$$

From the Lagrangian we write down the equation of motion for the particle using the Euler-Lagrange equation (2.2.5).

$$\frac{\partial L}{\partial q} = -kq$$

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}} = m\ddot{q}$$

Equating the above we get the familiar equation describing a mass attached to a spring.

$$m\ddot{q} = -kq \tag{2.2.18}$$

In spite of its simplicity, the harmonic oscillator model has many applications in physics. Most smooth potentials will look almost quadratic in the neighborhood of a minimum. The inter-atomic potential has this character. For this reason, a chain of atoms may be modeled using a chain of identical particles connected by identical springs. For small longitudinal motion, the Lagrangian for such a system can be written as a sum of harmonic oscillators.

$$L = \sum_n \frac{1}{2}m\,\dot{y}_n^2 - \sum_n \frac{1}{2}k\,[y_{n+1} - y_n]^2 \tag{2.2.19}$$

Here we denote the longitudinal displacement from equilibrium position of the n-th atom by y_n. The variable along the chain direction can be denoted by x and the inter-atomic distance at equilibrium by Δx. In other words, the displacement of the n-th atom, y_n is a function of x. Though they do not show up in the above Lagrangian, we shall use for them in the next example.

2.2.2 Continuum Mechanics of a Cylinder

We next turn to the slightly more complicated problem of the longitudinal motion of an elastic cylinder. The continuous cylinder can be seen as a limit of the atomic chain. To take the continuum limit of equation (2.2.19) we make the inter-atomic distance Δx explicit.

$$L = \sum_n \frac{1}{2}\frac{m}{\Delta x}\,\dot{y}_n^2 \Delta x - \sum_n \frac{1}{2}\,k\Delta x\,\left[\frac{y_{n+1} - y_n}{\Delta x}\right]^2 \Delta x \tag{2.2.20}$$
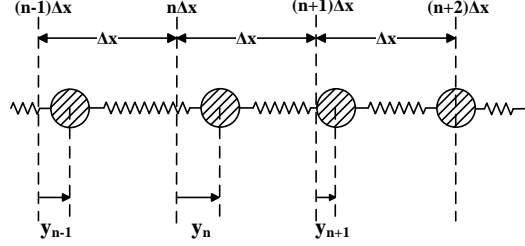
Figure 2.2: A chain of linear atoms can be used to approximate a longitudinal mode on a solid cylinder. The mass density per unit length $\rho = m/\Delta x$, and the spring constant k becomes large when Δx becomes small such that $k\Delta x = T$, the tension.

Nothing is changed by the insertion of the Δx factors. However, we can associated the discrete quantities $m/\Delta x$ and $k\Delta x$ with their continuum counterparts ρ and T, which are the linear mass density and string tension, respectively, because m becomes smaller, and k becomes larger when $\Delta x \to 0$. These parameters characterize the continuous string. Dimensional analysis justifies these associations.

Taking the continuum limit $\Delta x \to dx$ in equation (2.2.20), the discrete index n can be replaced by the continuous x since $n\Delta x \to x$; and the difference term in the second summation turns into a derivative. Inserting the continuum parameters we have

$$
\begin{aligned}
L &= \sum_n \frac{1}{2}\rho\, \dot{y}^2(x,t)\Delta x - \sum_n \frac{1}{2}T \left[\frac{\partial y(x,t)}{\partial x}\right]^2 \Delta x \\
L(\dot{y}, y') &= \int_0^\ell \left[\frac{1}{2}\rho\, \dot{y}^2(x,t) - \frac{1}{2}T\, y'^2(x,t)\right] dx
\end{aligned}
\tag{2.2.21}
$$

where $y'(x,t) = \partial y(x,t)/\partial x$, and ℓ is the total length of the cylinder. In equation (2.2.21), dot denotes time derivative and prime denotes spatial derivative. The limits of integration sets the horizontal length of the string. It is important to note that x has entered the Lagrangian simply as a replacement for the discrete index n and holds little physical significance. Despite seemingly having two spatial coordinates y and x, the mechanics of the string remains a one dimensional problem. The *dynamical variable* is y while x is only a *parameter*. The functional form of $y(x,t)$ describes the shape of the string at a certain time.

The quantity inside the integral over x is called the Lagrangian density $\mathcal{L}$.

$$
\mathcal{L}(\dot{y}, y') = \frac{1}{2}\rho\, \dot{y}^2 - \frac{1}{2}T\, y'^2
\tag{2.2.22}
$$

Continuum mechanics are characterized by Lagrangian densities. The action is still given by the time integral of the Lagrangian.

$$
S = \int_{t_1}^{t_2} dt \int_0^\ell dx\, \mathcal{L}(\dot{y}, y')
\tag{2.2.23}
$$

To write the equation of motion, we use the principle of least action. Make a variation to the shape of the string $y(x,t) \to y(x,t) + \delta y(x,t)$, the corresponding variations to the spatial and temporal derivatives of y are

$$
\begin{aligned}
\dot{y} &\to \dot{y} + \delta\dot{y} \\
y' &\to y' + \delta y'
\end{aligned}
$$

To compute the change in the action we identify the explicit dependence of the Lagrangian density on $\dot{y}$ and y' and do a Taylor expansion. Keeping only first order terms we have:

$$
\begin{aligned}
\delta S &= \int_{t_1}^{t_2} dt \int_0^\ell dx \left[\frac{\partial \mathcal{L}}{\partial \dot{y}} \delta\dot{y} + \frac{\partial \mathcal{L}}{\partial y'} \delta y' \right] \\
\delta S &= \int_{t_1}^{t_2} dt \int_0^\ell dx \left[\rho\dot{y}\,\delta\dot{y} - Ty'\,\delta y' \right]
\end{aligned}
\tag{2.2.24}
$$

The strategy then is to factor out the variation δy via integration by parts. The order of integration can be interchanged freely. We proceed as

$$
\begin{aligned}
\delta S &= \int_0^\ell dx \int_{t_1}^{t_2} dt\, \rho\dot{y}\,\delta\dot{y} - \int_{t_1}^{t_2} dt \int_0^\ell dx\, Ty'\delta y' \\
&= \int_0^\ell dx \left[\rho\dot{y}\,\delta y \Big|_{t_1}^{t_2} - \int_{t_1}^{t_2} dt\, \rho\ddot{y}\,\delta y \right] - \int_{t_1}^{t_2} dt \left[Ty'\delta y \Big|_0^\ell - \int_0^\ell dx\, Ty''\delta y \right] \\
\delta S &= \int_{t_1}^{t_2} dt \int_0^\ell dx \left[-\rho\ddot{y} + Ty'' \right] \delta y
\end{aligned}
\tag{2.2.25}
$$

Using the arbitrariness of δy we conclude that the stationarity of the action amounts to setting $-\rho\ddot{y} + Ty'' = 0$. We define a velocity $v = \sqrt{T/\rho}$ and recognize this as the familiar wave equation.

$$
\frac{\partial^2 y}{\partial x^2} = \frac{1}{v^2} \frac{\partial^2 y}{\partial t^2}
\tag{2.2.26}
$$

We must elaborate on a point about fixing boundary conditions here. In the line leading to equation (2.2.25), we have explicitly written the integrated out pieces. They are then both set to zero but for quite different reasons.

We set the boundary pieces in the dx integral (first integral) to zero because as is *always* done in deriving the equations of motion in Lagrangian mechanics, we require that the shape of the displacement, $y(x,t_1)$ and $y(x,t_2)$ to be fixed at these two times.

The boundary pieces in the dt integral are set to zero because of boundary conditions. If $y(x,t)$ is fixed at $x = 0$ and $x = \ell$ (Dirichlet boundary condition), or if $y'(x,t) = 0$ at $x = 0$ and $x = \ell$ (Neumann boundary condition), then these pieces are zero. Also, if one is solving a periodic problem, then $y'\delta y$ are equal to each other at these two end points $x = 0$ and $x = \ell$, and these two pieces cancel each other due to the periodic boundary condition.

With the example of the string we conclude our introduction of the Lagrangian formulation of classical mechanics.

2.3   Hamiltonian Formulation

For a multi-dimensional system, or a many particle system in multi-dimensions, the total time derivative of L is

$$\frac{\mathrm{d}L}{\mathrm{d}t} = \sum_i \left(\frac{\partial L}{\partial q_i} \dot{q}_i + \frac{\partial L}{\partial \dot{q}_i} \ddot{q}_i \right) \tag{2.3.1}$$

Since $\partial L/\partial q_i = \frac{\mathrm{d}}{\mathrm{d}t}(\partial L/\partial \dot{q}_i)$ from the Lagrange equation, we have

$$\frac{\mathrm{d}L}{\mathrm{d}t} = \sum_i \left[\frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{\partial L}{\partial \dot{q}_i} \right) \dot{q}_i + \frac{\partial L}{\partial \dot{q}_i} \ddot{q}_i \right] = \frac{\mathrm{d}}{\mathrm{d}t} \sum_i \left(\frac{\partial L}{\partial \dot{q}_i} \dot{q}_i \right) \tag{2.3.2}$$

or

$$\frac{\mathrm{d}}{\mathrm{d}t} \left(\sum_i \frac{\partial L}{\partial \dot{q}_i} \dot{q}_i - L \right) = 0 \tag{2.3.3}$$

The quantity

$$H = \sum_i \frac{\partial L}{\partial \dot{q}_i} \dot{q}_i - L \tag{2.3.4}$$

is known as the Hamiltonian of the system, and is a constant of motion, namely, $dH/dt = 0$. As shall be shown, the Hamiltonian represents the total energy of a system. It is a constant of motion because of the conservation of energy.

    The Hamiltonian of the system, (2.3.4), can also be written, after using (2.2.15), as

$$H = \sum_i \dot{q}_i p_i - L \tag{2.3.5}$$

where $p_i = \partial L/\partial \dot{q}_i$ is the generalized momentum. The first term has a dimension of energy, and in Cartesian coordinates, for a simple particle motion, it is easily seen that it is twice the kinetic energy. Hence, the above indicates that the Hamiltonian

$$H = T + V \tag{2.3.6}$$

The total variation of the Hamiltonian is

$$\delta H = \delta \left(\sum_i p_i \dot{q}_i \right) - \delta L$$

$$= \sum_i (\dot{q}_i \delta p_i + p_i \delta \dot{q}_i) - \sum_i \left(\frac{\partial L}{\partial q_i} \delta q_i + \frac{\partial L}{\partial \dot{q}_i} \delta \dot{q}_i \right) \tag{2.3.7}$$

Using (2.2.15) and (2.2.16), we have

$$\delta H = \sum_i (\dot{q}_i \delta p_i + p_i \delta \dot{q}_i) - \sum_i (\dot{p}_i \delta q_i + p_i \delta \dot{q}_i)$$

$$= \sum_i (\dot{q}_i \delta p_i - \dot{p}_i \delta q_i) \tag{2.3.8}$$

From the above, since the first variation of the Hamiltonian depends only on δp_i and δq_i, we gather that the Hamiltonian is a function of p_i and q_i. Taking the first variation of the Hamiltonian with respect to these variables, we arrive at another expression for its first variation, namely,

$$\delta H = \sum_i \left(\frac{\partial H}{\partial p_i} \delta p_i + \frac{\partial H}{\partial q_i} \delta q_i \right) \tag{2.3.9}$$

Comparing the above with (2.3.8), we gather that

$$\dot{q}_i = \frac{\partial H}{\partial p_i} \tag{2.3.10}$$

$$\dot{p}_i = -\frac{\partial H}{\partial q_i} \tag{2.3.11}$$

These are the equations of motion known as the Hamilton's equations.

The (2.3.4) is also known as the Legendre transformation. The original function L is a function of $\dot{q}_i$, q_i. Hence, δL depends on both $\delta \dot{q}_i$ and δq_i. After the Legendre transformation, δH depends on the differential δp_i and δq_i as indicated by (2.3.8). This implies that H is a function of p_i and q_i. The equations of motion then can be written as in (2.3.10) and (2.3.11).

Alternatively, one can argue that since the Hamiltonian represents the total energy of the system, it must be a constant of motion, or constant with respect to time due to energy conservation. Therefore, taking the time derivative of the Hamiltonian, one ends up with

$$\frac{\partial H}{\partial t} = \sum_i \left(\frac{\partial H}{\partial p_i} \frac{\partial p_i}{\partial t} + \frac{\partial H}{\partial q_i} \frac{\partial q_i}{\partial t} \right) = 0 \tag{2.3.12}$$

One can see that when the Hamilton's equations are satisfied, it implies that energy is conserved.

2.4   More on Hamiltonian

The Hamiltonian of a particle in classical mechanics is given by (2.3.6), and it is a function of p_i and q_i. For a non-relativistic particle in three dimensions, the kinetic energy

$$T = \frac{\mathbf{p} \cdot \mathbf{p}}{2m} \tag{2.4.1}$$

and the potential energy V is a function of $\mathbf{q}$. Hence, the Hamiltonian can be expressed as

$$H = \frac{\mathbf{p} \cdot \mathbf{p}}{2m} + V(\mathbf{q}) \tag{2.4.2}$$

in three dimensions. When an electromagnetic field is present, the Hamiltonian for an electron can be derived by letting the generalized momentum

$$p_i = m\dot{q}_i + eA_i \tag{2.4.3}$$

where $e = -|e|$ is the electron charge and A_i is component of the vector potential $\mathbf{A}$. Consequently, the Hamiltonian of an electron in the presence of an electromagnetic field is

$$H = \frac{(\mathbf{p} - e\mathbf{A}) \cdot (\mathbf{p} - e\mathbf{A})}{2m} + e\phi(\mathbf{q}) \tag{2.4.4}$$

The equation of motion of an electron in an electromagnetic field is governed by the Lorentz force law, which can be derived from the above Hamiltonian using the equations of motion provided by (2.3.10) and (2.3.11).

2.5   Poisson Bracket

Yet another way of expressing equations of motion in classical mechanics is via the use of Poisson bracket. This is interesting because Poisson bracket has a close quantum mechanics analogue. A Poisson bracket of two scalar variables $u(q, p)$ and $v(q, p)$ that are functions of q and p is defined as

$$\{u, v\} = \frac{\partial u}{\partial q}\frac{\partial v}{\partial p} - \frac{\partial v}{\partial q}\frac{\partial u}{\partial p} \tag{2.5.1}$$

In this notation, using (2.3.10) and (2.3.11),

$$\begin{aligned}
\frac{\mathrm{d}u}{\mathrm{d}t} &= \frac{\partial u}{\partial q}\frac{\mathrm{d}q}{\mathrm{d}t} + \frac{\partial u}{\partial p}\frac{\mathrm{d}p}{\mathrm{d}t} = \frac{\partial u}{\partial q}\frac{\partial H}{\partial p} - \frac{\partial u}{\partial p}\frac{\partial H}{\partial q} \\
&= \{u, H\}
\end{aligned} \tag{2.5.2}$$

which is valid for any variable u that is a function of p and q. Hence, from the Hamilton's equations of motion,

$$\frac{\partial q}{\partial t} = \frac{\partial H}{\partial p}, \qquad \frac{\partial p}{\partial t} = -\frac{\partial H}{\partial q} \tag{2.5.3}$$

we have the equations of motion as

$$\dot{q} = \{q, H\}, \quad \dot{p} = \{p, H\} \tag{2.5.4}$$

in the Poisson bracket notation. As we shall see later, similar equations will appear in quantum mechanics.

The algebraic properties of Poisson bracket are

$$\{u, v\} = -\{v, u\} \tag{2.5.5}$$
$$\{u + v, w\} = \{u, w\} + \{v, w\} \tag{2.5.6}$$
$$\{uv, w\} = \{u, w\}v + u\{v, w\} \tag{2.5.7}$$
$$\{u, vw\} = \{u, v\}w + v\{u, w\} \tag{2.5.8}$$
$$\{\{u, v\}, w\} + \{\{v, w\}, u\} + \{\{w, u\}, v\} = 0 \tag{2.5.9}$$

These properties are antisymmetry, distributivity, associativity and Jacobi's identity.

If we define a commutator operation between two noncommuting operator $\hat{u}$ and $\hat{v}$ as

$$[\hat{u}, \hat{v}] = \hat{u}\hat{v} - \hat{v}\hat{u}, \tag{2.5.10}$$

where $\hat{u}$ and $\hat{v}$ are analogous to matrix operators, it can be shown that the above commutator have the same algebraic properties as the Poisson bracket. An operator in quantum mechanics can be a matrix operator or a differential operator. In general, operators do not commute unless under very special circumstances.

2.6 Some Useful Knowledge of Matrix Algebra

Matrix algebra (or linear algebra) forms the backbone of many quantum mechanical concepts. Hence, it is prudent to review some useful knowledge of matrix algebra. Many of the mathematical manipulations in quantum mechanics can be better understood if we understand matrix algebra.

A matrix is a mathematical linear operator that when operate (also called "act") on a vector produces another vector, or

$$\mathbf{b} = \overline{\mathbf{A}} \cdot \mathbf{a} \tag{2.6.1}$$

where $\mathbf{a}$ and $\mathbf{b}$ are distinct vectors, and $\overline{\mathbf{A}}$ is a matrix operator other than the identity operator. The inner product between two vectors can be of the form of reaction inner product

$$\mathbf{v}^t \cdot \mathbf{w} \tag{2.6.2}$$

or energy inner product

$$\mathbf{v}^\dagger \cdot \mathbf{w} \tag{2.6.3}$$

where the † implies conjugation transpose or that $\mathbf{v}^\dagger = (\mathbf{v}^*)^t$. For a finite dimensional matrix and vectors, the above can be written as

$$b_j = \sum_{i=1}^N A_{ji} a_i \tag{2.6.4}$$

$$\mathbf{v}^t \cdot \mathbf{w} = \sum_{i=1}^N v_i w_i \tag{2.6.5}$$

$$\mathbf{v}^\dagger \cdot \mathbf{w} = \sum_{i=1}^N v_i^* w_i \tag{2.6.6}$$

The above are sometimes written with the summation sign removed, namely, as

$$b_j = A_{ji} a_i \tag{2.6.7}$$

$$\mathbf{v}^t \cdot \mathbf{w} = v_i w_i \tag{2.6.8}$$

$$\mathbf{v}^\dagger \cdot \mathbf{w} = v_i^* w_i \tag{2.6.9}$$

The summation is implied whenever repeated indices occur. This is known varyingly as the index notation, indicial notation, or Einstein notation.

The above concepts can be extended to infinite dimensional system by letting $N \to \infty$, but the convergence of the summations must be established in order for the inner product to be meaningful. It is quite clear that

$$\mathbf{v}^\dagger \cdot \mathbf{v} = \sum_{i=1}^{N} v_i^* v_i = \sum_{i=1}^{N} |v_i|^2 > 0 \qquad (2.6.10)$$

or $\mathbf{v}^\dagger \cdot \mathbf{v}$ is positive definite.

Furthermore, matrix operators satisfies associativity but not commutativity, namely,

$$\left(\overline{\mathbf{A}} \cdot \overline{\mathbf{B}}\right) \cdot \overline{\mathbf{C}} = \overline{\mathbf{A}} \cdot \left(\overline{\mathbf{B}} \cdot \overline{\mathbf{C}}\right) \qquad (2.6.11)$$

$$\overline{\mathbf{A}} \cdot \overline{\mathbf{B}} \neq \overline{\mathbf{B}} \cdot \overline{\mathbf{A}} \qquad (2.6.12)$$

2.6.1   Identity, Hermitian, Symmetric, Inverse and Unitary Matrices

For discrete, countable systems, the definition of the above is quite straightforward. The identity operator $\overline{\mathbf{I}}$ is defined such that

$$\overline{\mathbf{I}} \cdot \mathbf{a} = \mathbf{a} \qquad (2.6.13)$$

or the ij element of the matrix is

$$\left[\overline{\mathbf{I}}\right]_{ij} = \delta_{ij} \qquad (2.6.14)$$

or it is diagonal matrix with one on the diagonal. A Hermitian matrix A_{ij} has the property that

$$A_{ij} = A_{ji}^* \qquad (2.6.15)$$

or

$$\overline{\mathbf{A}}^\dagger = \overline{\mathbf{A}} \qquad (2.6.16)$$

A symmetric matrix A_{ij} is such that

$$A_{ij} = A_{ji} \qquad (2.6.17)$$

or

$$\overline{\mathbf{A}}^t = \overline{\mathbf{A}} \qquad (2.6.18)$$

The inverse of the matrix $\overline{\mathbf{A}}$, denoted as $\overline{\mathbf{A}}^{-1}$, has the property that

$$\overline{\mathbf{A}}^{-1} \cdot \overline{\mathbf{A}} = \overline{\mathbf{I}} \qquad (2.6.19)$$

So given the equation

$$\overline{\mathbf{A}} \cdot \mathbf{x} = \mathbf{b} \tag{2.6.20}$$

$\mathbf{x}$ can be found once $\overline{\mathbf{A}}^{-1}$ is known. Multiplying the above by $\overline{\mathbf{A}}^{-1}$, we have

$$\mathbf{x} = \overline{\mathbf{A}}^{-1} \cdot \mathbf{b} \tag{2.6.21}$$

A unitary matrix $\overline{\mathbf{U}}$ has the property that

$$\overline{\mathbf{U}}^{\dagger} \cdot \overline{\mathbf{U}} = \overline{\mathbf{I}} \tag{2.6.22}$$

In other words

$$\overline{\mathbf{U}}^{\dagger} = \overline{\mathbf{U}}^{-1} \tag{2.6.23}$$

2.6.2 Determinant

An $N \times N$ matrix $\overline{\mathbf{A}}$ can be written as a collection of column vectors

$$\overline{\mathbf{A}} = [\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3, \ldots, \mathbf{a}_N] \tag{2.6.24}$$

where the column vectors are of length N. Determinant of $\overline{\mathbf{A}}$, the $\det(\overline{\mathbf{A}})$, or $|\overline{\mathbf{A}}|$ can be thought of as the "volume" subtended by the vectors $\mathbf{a}_n \ldots \mathbf{a}_N$. Also, $\det(\overline{\mathbf{A}})$ changes sign when any two column vectors are swapped. It can be thought of as a generalized "cross-product". Some useful properties of determinants are

1. $\det\left(\overline{\mathbf{I}}\right) = 1$

2. $\det\left(\overline{\mathbf{A}}^t\right) = \det\left(\overline{\mathbf{A}}\right)$

3. $\det\left(\overline{\mathbf{A}} \cdot \overline{\mathbf{B}}\right) = \det\left(\overline{\mathbf{A}}\right) \cdot \det(\overline{\mathbf{B}})$

4. $\det\left(\overline{\mathbf{A}}^{-1}\right) = 1/\det\left(\overline{\mathbf{A}}\right)$

5. If $\overline{\mathbf{\Lambda}}$ is diagonal, then $\det\left(\overline{\mathbf{\Lambda}}\right) = \prod_{i=1}^{N} \lambda_{ii}$

6. $\det\left(\overline{\mathbf{A}}\right) = \prod_{i=1}^{N} \lambda_i$ where λ_i are the eigenvalues of $\overline{\mathbf{A}}$.

7. $\det\left(\overline{\mathbf{A}}\right) = 0$ implies that $\overline{\mathbf{A}}$ is singular, or there exists a vector $\mathbf{u}$ such that $\overline{\mathbf{A}} \cdot \mathbf{u} = 0$

8. $\det\left(c\overline{\mathbf{A}}\right) = c^N \det\left(\overline{\mathbf{A}}\right)$

2.6.3 Eigenvectors and Eigenvalues

An eigenvector $\mathbf{v}$ of a matrix $\overline{\mathbf{A}}$ satisfies the property that

$$\overline{\mathbf{A}} \cdot \mathbf{v} = \lambda \mathbf{v} \tag{2.6.25}$$

where λ is the eigenvalue. The above can be rewritten as

$$\left(\overline{\mathbf{A}} - \lambda \overline{\mathbf{I}}\right) \cdot \mathbf{v} = 0 \tag{2.6.26}$$

The above implies that

$$\det\left(\overline{\mathbf{A}} - \lambda \overline{\mathbf{I}}\right) = 0 \tag{2.6.27}$$

In general, an $N \times N$ matrix has N eigenvalues with corresponding N eigenvectors. When two or more eigenvalues are the same, they are known as degenerate. Some useful properties are:

1. If $\overline{\mathbf{A}}$ is Hermitian, then λ is real.

2. If $\lambda_i \neq \lambda_j$, then $\mathbf{v}_i^\dagger \cdot \mathbf{v}_j = 0$. In general, $\mathbf{v}_i^\dagger \cdot \mathbf{v}_j = C_n \delta_{ij}$ where C_n is real.

3. If $\overline{\mathbf{A}}$ is positive definite, then $\lambda_i > 0$ for all i and vice versa for negative definiteness.

4. $\overline{\mathbf{P}} \cdot \overline{\mathbf{A}} \cdot \overline{\mathbf{P}}^{-1}$ has the same eigenvalues as $\overline{\mathbf{A}}$.

2.6.4 Trace of a Matrix

The trace of a matrix $\overline{\mathbf{A}}$ is the sum of its diagonal elements, or

$$\text{tr}\left(\overline{\mathbf{A}}\right) = \sum_{i=1}^{N} A_{ii} \tag{2.6.28}$$

Some properties are:

1. $\text{tr}\left(\overline{\mathbf{A}} + \overline{\mathbf{B}}\right) = \text{tr}\left(\overline{\mathbf{A}}\right) + \text{tr}\left(\overline{\mathbf{B}}\right)$

2. $\text{tr}\left(c\overline{\mathbf{A}}\right) = c\, \text{tr}\left(\overline{\mathbf{A}}\right)$

3. $\text{tr}\left(\overline{\mathbf{A}}\right) = \text{tr}\left(\overline{\mathbf{A}}^t\right)$

4. $\text{tr}\left(\overline{\mathbf{A}} \cdot \overline{\mathbf{B}}\right) = \text{tr}\left(\overline{\mathbf{B}} \cdot \overline{\mathbf{A}}\right)$

5. $\text{tr}\left(\overline{\mathbf{P}}^{-1} \cdot \overline{\mathbf{A}} \cdot \overline{\mathbf{P}}\right) = \text{tr}\left(\left(\overline{\mathbf{A}} \cdot \overline{\mathbf{P}}\right) \cdot \overline{\mathbf{P}}^{-1}\right) = \text{tr}\left(\overline{\mathbf{A}}\right)$

6. $\text{tr}\left(\overline{\mathbf{A}}\right) = \sum_{i=1}^{N} \lambda_i$ where λ_i's are the eigenvalues of $\overline{\mathbf{A}}$.

2.6.5   Function of a Matrix

An example of a function of a matrix is

$$e^{\overline{\mathbf{A}}} \tag{2.6.29}$$

The above has no meaning unless it operates on a vector; namely

$$e^{\overline{\mathbf{A}}} \cdot \mathbf{x} = \mathbf{b} \tag{2.6.30}$$

We can Taylor series expand to get

$$e^{\overline{\mathbf{A}}} = \overline{\mathbf{I}} + \overline{\mathbf{A}} + \frac{\overline{\mathbf{A}}^2}{2!} + \frac{\overline{\mathbf{A}}^3}{3!} + \cdots + \frac{\overline{\mathbf{A}}^n}{n!} + \cdots \tag{2.6.31}$$

Then if $\mathbf{v}$ is an eigenvector of $\overline{\mathbf{A}}$ such that $\overline{\mathbf{A}} \cdot \mathbf{v} = \lambda \mathbf{v}$, then

$$e^{\overline{\mathbf{A}}} \cdot \mathbf{v} = \left(\overline{\mathbf{I}} + \overline{\mathbf{A}} + \frac{\overline{\mathbf{A}}^2}{2!} + \cdots + \frac{\overline{\mathbf{A}}^n}{n!} + \cdots \right) \cdot \mathbf{v} \tag{2.6.32}$$

$$= \left(1 + \lambda + \frac{\lambda^2}{2!} + \cdots + \frac{\lambda^n}{n!} + \cdots \right) \cdot \mathbf{v} \tag{2.6.33}$$

$$= e^{\lambda} \mathbf{v} \tag{2.6.34}$$

In general, we can expand

$$\mathbf{x} = \sum_{i=1}^{N} a_i \mathbf{v}_i \tag{2.6.35}$$

where $\mathbf{v}_i, i = 1, \cdots, N$ are the N eigenvectors of $\overline{\mathbf{A}}$ with the property that $\overline{\mathbf{A}} \cdot \mathbf{v}_i = \lambda_i \mathbf{v}_i$. Then

$$e^{\overline{\mathbf{A}}} \cdot \mathbf{x} = \sum_{i=1}^{N} a_i e^{\overline{\mathbf{A}}} \mathbf{v}_i = \sum_{i=1}^{N} a_i e^{\lambda_i} \mathbf{v}_i \tag{2.6.36}$$

We can use the above to easily prove that the solution to

$$\frac{d\mathbf{v}(t)}{dt} = \overline{\mathbf{A}} \cdot \mathbf{v}(t) \tag{2.6.37}$$

is

$$\mathbf{v(t)} = e^{\overline{\mathbf{A}}\mathbf{t}} \cdot \mathbf{v}(0) \tag{2.6.38}$$

Exercise 1

For a particle with charge q in the presence of an electric field, the classical Hamiltonian is given by

$$H = \frac{1}{2m} (\mathbf{p} - q\mathbf{A})^2 + q\phi \tag{2.6.39}$$

The equation of motion from the Hamiltonian equations are

$$\dot{x}_j = \partial H/\partial p_j, \quad \dot{p}_j = -\partial H/\partial x_j \tag{2.6.40}$$

Show that

$$\dot{x}_j = \frac{1}{m}(p_j - qA_j) \tag{2.6.41}$$

$$\dot{p}_j = \frac{q}{m}(\mathbf{p} - q\mathbf{A}) \cdot \frac{\partial \mathbf{A}}{\partial x_j} - q\frac{\partial \phi}{\partial x_j} \tag{2.6.42}$$

We can express

$$\mathbf{r} = \hat{x}x_1 + \hat{y}x_2 + \hat{z}x_3 = \hat{x}x + \hat{y}y + \hat{z}z, \quad \mathbf{p} = \hat{x}p_1 + \hat{y}p_2 + \hat{z}p_3 = \hat{x}p_x + \hat{y}p_y + \hat{z}p_z \tag{2.6.43}$$

where the hat quantities are unit vectors. With the help of indicial notation, show that the above can be written as

$$\dot{\mathbf{r}} = \frac{1}{m}(\mathbf{p} - q\mathbf{A}) \tag{2.6.44}$$

$$\dot{\mathbf{p}} = \frac{q}{m}\nabla\mathbf{A} \cdot (\mathbf{p} - q\mathbf{A}) - q\nabla\phi = q\nabla\mathbf{A} \cdot \dot{\mathbf{r}} - q\nabla\phi \tag{2.6.45}$$

Since the vector potential is associate with the electron charge location, show that

$$\dot{A}_j = \sum_i \frac{\partial A_j}{\partial x_i}\dot{x}_i + \frac{\partial A_j}{\partial t} \tag{2.6.46}$$

or that

$$\dot{\mathbf{A}} = \dot{\mathbf{r}} \cdot \nabla\mathbf{A} + \frac{\partial \mathbf{A}}{\partial t} \tag{2.6.47}$$

Derive that

$$m\ddot{\mathbf{r}} = q\nabla\mathbf{A} \cdot \dot{\mathbf{r}} - q\dot{\mathbf{r}} \cdot \nabla\mathbf{A} - q\nabla\phi - q\frac{\partial \mathbf{A}}{\partial t} \tag{2.6.48}$$

Show that the above is the same as the Lorentz force law that

$$\mathbf{F} = m\mathbf{a} = q\mathbf{v} \times \mathbf{B} + q\mathbf{E} \tag{2.6.49}$$

where $\mathbf{F}$ is the force on the electron, $\mathbf{a}$ is the acceleration, and $\mathbf{v}$ is the velocity, $\mathbf{B}$ is the magnetic field, and $\mathbf{E}$ is the electric field.

Chapter 3

Quantum Mechanics—Some Preliminaries

3.1 Introduction

With some background in classical mechanics, we may motivate the Schrödinger equation in a more sanguine fashion. Experimental evidence indicated that small particles such as electrons behave quite strangely and cannot be described by classical mechanics alone. In classical mechanics, once we know p and q and their time derivatives (or $\dot{p}$, $\dot{q}$) of a particle at time t_0, one can integrate the equations of motion

$$\dot{p} = F, \quad \dot{q} = p/m \tag{3.1.1}$$

or use the finite difference method to find p and q at $t_0 + \Delta t$, and at all subsequent times.

In quantum mechanics, the use of two variables p and q and their derivatives is insufficient to describe the state of a particle and derive its future states. The state of a particle has to be more richly endowed and described by a wave function or state function $\psi(q, t)$ where q is a position indicator. The state function (also known as a state vector) is a vector in the infinite dimensional space.[1]

At this juncture, the state function or vector is analogous to when we study the control theory of a highly complex system. In the state variable approach, the state of a control system is described by the state vector, whose elements are variables that we think are important to capture the state of the system. For example, the state vector $\mathbf{v}$, describing the state of the factory, can contain variables that represent the number of people in a factory, the number of machines, the temperature of the rooms, the inventory in each room, etc. The state equation of this factory can then be written as

$$\frac{d}{dt}\mathbf{v}(t) = \overline{\mathbf{A}} \cdot \mathbf{v}(t) \tag{3.1.2}$$

---

[1]This concept will be further elaborated in Chapter 5.

It describes the time evolution of the factory. The matrix $\overline{\mathbf{A}}$ describes the coupling between state variables as they evolve. It bears strong similarity to the time-dependent Schrödinger equation, which is used to describe the time-evolution of the state function or the wave function of an electron. In the wave function, a complex number is assigned to each location in space.
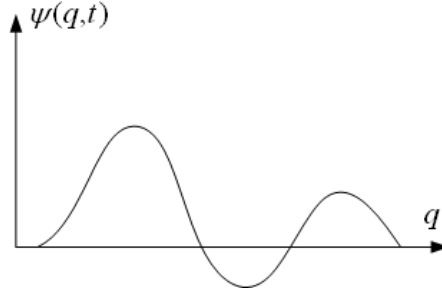
Figure 3.1: The state of a particle in quantum mechanics is described by a state function, which has infinitely many degrees of freedom.

In the Schrödinger equation, the wave function $\psi(q,t)$ is a continuous function of of the position variable q at any time instant t; hence, it is described by infinitely many numbers, and has infinite degrees of freedom. The time evolution of the wave function $\psi(q,t)$ is governed by the Schrödinger equation. It was motivated by experimental evidence and the works of many others such as Planck, Einstein, and de Broglie, who were aware of the wave nature of a particle and the dual wave-particle nature of light.

3.2   Probabilistic Interpretation of the Wave Function

The wave function of the Schrödinger equation has defied an acceptable interpretation for many years even though the Schrödinger equation was known to predict experimental outcomes. Some thought that it represented an electron cloud, and that perhaps, an electron, at the atomistic level, behaved like a charge cloud, and hence not a particle. The final, most accepted interpretation of this wave function (one that also agrees with experiments) is that its magnitude squared corresponds to the probability density function.[2] In other words, the probability of finding an electron in an interval $[x, x + \Delta x]$ is equal to

$$|\psi(x,t)|^2 \Delta x \tag{3.2.1}$$

For the 3D case, the probability of finding an electron in a small volume ΔV in the vicinity of the point $\mathbf{r}$ is given by

$$|\psi(\mathbf{r},t)|^2 \Delta V \tag{3.2.2}$$

---

[2]This interpretation is due to Born.

Since the magnitude squared of the wave function represents a probability density function, it must satisfy the normalization condition of a probability density function, viz.,

$$\int dV |\psi(\mathbf{r}, t)|^2 = 1 \tag{3.2.3}$$

with its counterparts in 1D and 2D. The magnitude squared of this wave function is like some kind of "energy" that cannot be destroyed. Electrons cannot be destroyed, and hence, charge conservation is upheld by the Schrödinger equation.

3.3   Time Evolution of the Hamiltonian Operator

Motivated by the conservation of the "energy" of the wave function, we shall consider an "energy" conserving system where the classical Hamiltonian will be a constant of motion. In this case, there is no "energy" loss from the system. The Schrödinger equation that governs the time evolution of the wave function $\psi(\mathbf{r}, t)$ is[3]

$$\left[-\frac{\hbar^2}{2m_0} \nabla^2 + V(\mathbf{r}) \right] \psi(\mathbf{r}, t) = i\hbar \frac{\partial}{\partial t} \psi(\mathbf{r}, t) \tag{3.3.1}$$

The above can be written as

$$\hat{H}\psi(\mathbf{r}, t) = i\hbar \frac{\partial \psi(\mathbf{r}, t)}{\partial t} \tag{3.3.2}$$

where $\hat{H}$ is the Hamiltonian operator.[4] One can solve (3.3.2) formally to obtain

$$\psi(\mathbf{r}, t) = e^{-i\frac{\hat{H}}{\hbar}t}\psi(\mathbf{r}, t = 0) \tag{3.3.3}$$

Since the above is a function of an operator, it has meaning only if this function acts on the eigenfunctions of the operator $\hat{H}$. As has been discussed in Chapter 2, Subsection 2.6.5, it can be shown easily that if $\overline{\mathbf{A}} \cdot \mathbf{v}_i = \lambda_i \mathbf{v}_i$,

$$\exp(\overline{\mathbf{A}}) \cdot \mathbf{v}_i = \exp(\lambda_i)\mathbf{v}_i \tag{3.3.4}$$

To simplify the expression (3.3.3), it is best to express $\psi(\mathbf{r}, t = 0)$ as an eigenfunction (eigenstate or eigenvector) of $\hat{H}$, or linear superposition of its eigenfunctions. If $\hat{H}$ is a Hermitian operator, then there exist eigenfunctions, or special wave functions, ψ_n, such that

$$\hat{H}\psi_n(\mathbf{r}) = E_n\psi_n(\mathbf{r}) \tag{3.3.5}$$

where E_n is real and $\psi_n(\mathbf{r})$ are orthogonal to each other for different n. Analogous to the Hermitian matrix operator, it can be shown that E_n is purely real and that the ψ_n are orthogonal to each other. In this case, the time evolution of ψ_n from (3.3.3) is

$$\psi(\mathbf{r}, t) = e^{-i\frac{E_n}{\hbar}t}\psi_n(\mathbf{r}) = e^{-i\omega_n t}\psi_n(\mathbf{r}) \tag{3.3.6}$$

---

[3] Please be reminded that this equation was postulated and was not derived.

[4] Rightfully, one should use the bra and ket notation to write this equation as $\hat{H}|\psi\rangle = i\hbar \frac{d}{dt}|\psi\rangle$. In the less rigorous notation in (3.3.2), we will assume that $\hat{H}$ is in the representation in which the state vector ψ is in. That is if ψ is in coordinate space representation, $\hat{H}$ is also in coordinates space representation.

In the above, $E_n = \hbar\omega_n$, or the energy E_n is related to frequency ω_n via the reduced Planck constant $\hbar$. The reduced Planck constant is related to the Planck constant by $\hbar = h/(2\pi)$ where $h = 6.626068 \times 10^{-34}$ J s. The fact that E_n is real means that ω_n is real, or that the magnitude squared of these functions are time independent or conserved, as is required by their probabilistic interpretation.

We can write

$$\psi(\mathbf{r}, t = 0) = \sum_n c_n \psi_n(\mathbf{r}) \tag{3.3.7}$$

where c_n can be found using the orthonormality relation of ψ_n. Then

$$\psi(\mathbf{r}, t) = e^{-\frac{i\hat{H}}{\hbar}t}\psi(\mathbf{r}, t = 0) = \sum_n c_n e^{-\frac{i\hat{H}}{\hbar}t}\psi_n(\mathbf{r}) = \sum_n c_n e^{-i\omega_n t}\psi_n(\mathbf{r}) \tag{3.3.8}$$

since $e^{-\frac{i\hat{H}}{\hbar}t}\psi_n(\mathbf{r}) = e^{-i\omega_n t}\psi_n(\mathbf{r})$. It can be easily shown that the above, which is derived from (3.3.3) is a solution to (3.3.2). Hence, (3.3.3) is the formal solution to (3.3.2).

Scalar variables that are measurable in classical mechanics, such as p and q, are known as observables in quantum mechanics. They are elevated from scalar variables to operators in quantum mechanics, denoted by a "ˆ" symbol here. In classical mechanics, for a one particle system, the Hamiltonian is given by

$$H = T + V = \frac{p^2}{2m} + V \tag{3.3.9}$$

The Hamiltonian contains the information from which the equations of motion for the particle can be derived. But in quantum mechanics, this is not sufficient, and H becomes an operator

$$\hat{H} = \frac{\hat{p}^2}{2m} + \hat{V} \tag{3.3.10}$$

This operator works in tandem with a wave function ψ to describe the state of the particle. The operator acts on a wave function $\psi(t)$, where in the one dimensional case and in the coordinate q representation, is $\psi(q, t)$. When $\psi(q, t)$ is an eigenfunction with energy E_n, it can be expressed as

$$\psi(q, t) = \psi_n(q)e^{-i\omega_n t} \tag{3.3.11}$$

where $E_n = \hbar\omega_n$. The Schrödinger equation for $\psi_n(q)$ then becomes[5]

$$\hat{H}\psi_n(q) = \left(\frac{\hat{p}^2}{2m} + \hat{V}\right)\psi_n(q) = E_n\psi_n(q) \tag{3.3.12}$$

For simplicity, we consider an electron moving in free space where it has only a constant kinetic energy but not influenced by any potential energy. In other words, there is no force acting on the electron. In this case, $\hat{V} = 0$, and this equation becomes

$$\frac{\hat{p}^2}{2m}\psi_n(q) = E_n\psi_n(q) \tag{3.3.13}$$

---

[5]For the Schrödinger equation in coordinate space, $\hat{V}$ turns out to be a scalar operator or a diagonal operator.

It has been observed by de Broglie that the momentum of a particle, such as an electron which behaves like a wave, has a momentum

$$p = \hbar k \tag{3.3.14}$$

where $k = 2\pi/\lambda$ is the wavenumber of the wave function. This motivates that the operator $\hat{p}$ can be expressed by

$$\hat{p} = -i\hbar \frac{d}{dq} \tag{3.3.15}$$

in the coordinate space representation. This is chosen so that if an electron is described by a state function $\psi(q, t) = c_1 e^{ikq - i\omega t}$, then $\hat{p}\psi(q, t) = \hbar k\psi(q, t)$. The above motivation for the form of the operator $\hat{p}$ is highly heuristic. We will see other reasons for the form of $\hat{p}$ when we study the correspondence principle and the Heisenberg picture.

Equation (3.3.13) for a free particle is then

$$-\frac{\hbar^2}{2m} \frac{d^2}{dq^2} \psi_n(q) = E_n \psi_n(q) \tag{3.3.16}$$

Since this is a constant coefficient ordinary differential equation, the solution is of the form

$$\psi_n(q) = e^{\pm i k_n q} \tag{3.3.17}$$

which when used in (3.3.16), yields

$$\frac{\hbar^2 k_n^2}{2m} = E_n \tag{3.3.18}$$

Namely, the kinetic energy T of the particle is given by

$$T = \frac{\hbar^2 k_n^2}{2m} \tag{3.3.19}$$

where $p = \hbar k_n$ is in agreement with de Broglie's finding.

In many problems, the operator $\hat{V}$ is a scalar operator in coordinate space representation which is a scalar function of position $V(q)$. This potential traps the particle within it acting as a potential well. In general, the Schrödinger equation for a particle becomes

$$\left[-\frac{\hbar^2}{2m} \frac{\partial^2}{\partial q^2} + V(q) \right] \psi(q, t) = i\hbar \frac{\partial}{\partial t} \psi(q, t) \tag{3.3.20}$$

For a particular eigenstate with energy E_n as indicated by (3.3.11), it becomes

$$\left[-\frac{\hbar^2}{2m} \frac{d^2}{dq^2} + V(q) \right] \psi_n(q) = E_n \psi_n(q) \tag{3.3.21}$$

The above is an eigenvalue problem with eigenvalue E_n and eigenfunction $\psi_n(q)$. These eigenstates are also known as stationary states, because they have a time dependence indicated by (3.3.11). Hence, their probability density functions $|\psi_n(q, t)|^2$ are time independent.

These eigenfunctions correspond to trapped modes (or bound states) in the potential well defined by $V(q)$ very much like trapped guided modes in a dielectric waveguide. These modes are usually countable and they can be indexed by the index n.

In the special case of a particle in free space, or the absence of the potential well, the particle or electron is not trapped and it is free to assume any energy or momentum indexed by the continuous variable k. In (3.3.18), the index for the energy should rightfully be k and the eigenfunctions are uncountably infinite. Moreover, the above can be generalized to two and three dimensional cases.

3.4   Simple Examples of Time Independent Schrödinger Equation

At this juncture, we have enough knowledge to study some simple solutions of time-independent Schrödinger equation such as a particle in a box, a particle impinging on a potential barrier, and a particle in a finite potential well.

3.4.1   Particle in a 1D Box

Consider the Schrödinger equation for the 1D case where the potential $V(x)$ is defined to be a function with zero value for $0 < x < a$ (inside the box) and infinite value outside this range. The Schrödinger equation is given by

$$\left[-\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + V(x) \right] \psi(x) = E\psi(x) \tag{3.4.1}$$

where we have replaced q with x. Since $V(x)$ is infinite outside the box, $\psi(x)$ has to be zero. Inside the well, $V(x) = 0$ and the above equation has a general solution of the form

$$\psi(x) = A\sin(kx) + B\cos(kx) \tag{3.4.2}$$

The boundary conditions are that $\psi(x = 0) = 0$ and $\psi(x = a) = 0$. For this reason, a viable solution for $\psi(x)$ is

$$\psi(x) = A\sin(kx) \tag{3.4.3}$$

where $k = n\pi/a$, $n = 1, \ldots, \infty$. There are infinitely many eigensolutions for this problem. For each chosen n, the corresponding energy of the solution is

$$E_n = \frac{(\hbar n\pi/a)^2}{2m} \tag{3.4.4}$$

These energy values are the eigenvalues of the problem, with the corresponding eigenfunctions given by (3.4.3) with the appropriate k. It is seen that the more energetic the electron is (high E_n values), the larger the number of oscillations the wave function has inside the box. The solutions that are highly oscillatory have higher k values, and hence, higher momentum or higher kinetic energy. The solutions which are even about the center of the box are said to have even parity, while those that are odd have odd parity.

One other thing to be noted is that the magnitude squared of the wave function above represents the probability density function. Hence, it has to be normalized. The normalized version of the wave function is

$$\psi(x) = \sqrt{2/a}\sin(n\pi x/a) \tag{3.4.5}$$

Moreover, these eigenfunctions are orthonormal to each other, viz.,

$$\int_0^a dx\,\psi_n^*(x)\psi_m(x) = \delta_{nm} \tag{3.4.6}$$

The orthogonality is the generalization of the fact that for a Hermitian matrix system, where the eigenvectors are given by

$$\overline{\mathbf{H}} \cdot \mathbf{v}_i = \lambda_i \mathbf{v}_i \tag{3.4.7}$$

then it can be proven easily that

$$\mathbf{v}_j^\dagger \cdot \mathbf{v}_i = C_j \delta_{ij} \tag{3.4.8}$$

Moreover, the eigenvalues are real.
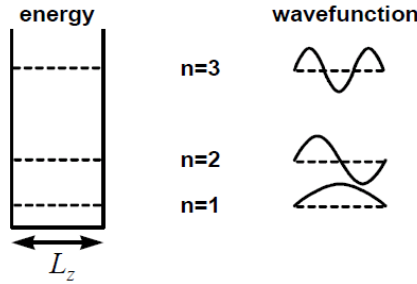
Figure 3.2: The wave functions of an electron trapped in a 1D box (from DAB Miller).

3.4.2 Particle Scattering by a Barrier

In the previous example, it is manifestly an eigenvalue problem since the solution can be found only at discrete values of E_n. The electron is trapped inside the box. However, in an open region problem where the electron is free to roam, the energy of the electron E can be arbitrary. We can assume that the potential profile is such that $V(x) = 0$ for $x < 0$ while $V(x) = V_0$ for $x > 0$. The energy of the electron is such that $0 < E < V_0$. On the left side, we assume an electron coming in from $-\infty$ with the wave function described by $A\exp(ik_1x)$. When this wave function hits the potential barrier, a reflected wave ensues, and the general solution on the left side of the barrier is given by

$$\psi_1(x) = A_1 e^{ik_1 x} + B_1 e^{-ik_1 x} \tag{3.4.9}$$

where $(\hbar k_1)^2/(2m) = E$ is the kinetic energy of the incident electron. On the right side, however, the Schrödinger equation to be satisfied is

$$\left[-\frac{\hbar^2}{2m}\frac{d^2}{dx^2} \right] \psi_2(x) = (E - V_0)\psi_2(x) \qquad (3.4.10)$$

The solution of the transmitted wave on the right is

$$\psi_2(x) = A_2 e^{ik_2 x} \qquad (3.4.11)$$

where

$$k_2 = \sqrt{2m(E - V_0)}/\hbar \qquad (3.4.12)$$

Given the known incident wave amplitude A_1, we can match the boundary conditions at $x = 0$ to find the reflected wave amplitude B_1 and the transmitted wave amplitude A_2. By eye-balling the Schrödinger equation (3.4.1), we can arrive at the requisite boundary conditions are that ψ and $\frac{d}{dx}\psi(x)$ are continuous at $x = 0$. This is because $\frac{d^2}{dx^2}\psi(x)$ in (3.4.1) has to be a finite quantity. Hence, $\frac{d}{dx}\psi(x)$ and $\psi(x)$ cannot have jump discontinuities.

Since $E < V_0$, k_2 is pure imaginary, and the wave is evanescent and decays when $x \to \infty$. This effect is known as tunneling. The electron as a nonzero probability of being found inside the barrier, albeit with decreasing probability into the barrier. The larger V_0 is compared to E, the more rapidly decaying is the wave function into the barrier.

However, if the electron is energetic enough so that $E > V_0$, k_2 becomes real, and then the wave function is no more evanescent. It penetrates into the barrier; it can be found even a long way from the boundary.
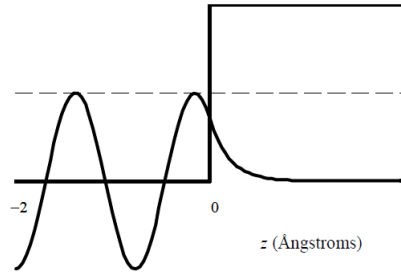
Figure 3.3: Scattering of the electron wave function by a 1D barrier (from DAB Miller).

It is to be noted that the wave function in this case cannot be normalized as the above represents a fictitious situation of an electron roaming over infinite space. However, the above example illustrates the wave physics at the barrier.

3.4.3   Particle in a Potential Well

If the potential profile is such that

$$V(x) = \begin{cases} V_1 & \text{if } x < -a/2, \text{ region 1} \\ V_2 = 0 & \text{if } |x| < a/2, \text{ region 2} \\ V_3 & \text{if } x > a/2, \text{ region 3} \end{cases} \tag{3.4.13}$$

then there can be trapped modes (or bound states) inside the well represented by standing waves, whereas outside the well, the waves are evanescent for eigenmodes for which $E < V_1$ and $E < V_3$.

The wave function for $|x| < a/2$ can be expressed as

$$\psi_2(x) = A_2 \sin(k_2 x) + B_2 \cos(k_2 x) \tag{3.4.14}$$

where $k_2 = \sqrt{2mE}/\hbar$. In region 1 to the left, the wave function is

$$\psi_1(x) = A_1 e^{\alpha_1 x} \tag{3.4.15}$$

where $\alpha_1 = \sqrt{2m(V_1 - E)}/\hbar$. The wave has to decay in the left direction. Similar, in region 3 to the right, the wave function is

$$\psi_3(x) = B_3 e^{-\alpha_3 x} \tag{3.4.16}$$

where $\alpha_3 = \sqrt{2m(V_3 - E)}/\hbar$. It has to decay in the right direction. Four boundary conditions can be imposed at $x = \pm a/2$ to eliminate the four unknowns A_1, A_2, B_2, and B_3. These four boundary conditions are continuity of the wave function $\psi(x)$ and its derivative $d\psi(x)/dx$ at the two boundaries. However, non-trivial eigensolutions can only exist at selected values of E which are the eigenvalues of the Schrödinger equation. The eigenequation from which the eigenvalues can be derived is a transcendental equation.

To illustrate this point, we impose that ψ is continuous at $x = \pm a/2$ to arrive at the following two equations:

$$A_1 e^{-\alpha_1 a/2} = -A_2 \sin(k_2 a/2) + B_2 \cos(k_2 a/2) \tag{3.4.17}$$
$$A_3 e^{-\alpha_3 a/2} = A_2 \sin(k_2 a/2) + B_2 \cos(k_2 a/2) \tag{3.4.18}$$

We further impose that $\partial\psi/\partial x$ is continuous at $x = \pm a$ to arrive at the following two equations:

$$\alpha_1 A_1 e^{-\alpha_1 a/2} = k_2 A_2 \cos(k_2 a/2) + k_2 B_2 \sin(k_2 a/2) \tag{3.4.19}$$
$$-\alpha_3 A_3 e^{-\alpha_3 a/2} = k_2 A_2 \cos(k_2 a/2) - k_2 B_2 \sin(k_2 a/2) \tag{3.4.20}$$

The above four equations form a matrix equation

$$\overline{\mathbf{M}} \cdot \mathbf{v} = 0 \tag{3.4.21}$$

where $\mathbf{v} = [A_1, A_2, B_2, B_3]^t$, and the elements of $\overline{\mathbf{M}}$, which depend on E (because $k_2 = \sqrt{2mE}/\hbar$, $\alpha_1 = \sqrt{2m(V_1 - E)}/\hbar$, and $\alpha_3 = \sqrt{2m(V_3 - E)}/\hbar$), can be gleaned off the above equations. Non-trivial solutions exists for $\mathbf{v}$ only if

$$\det\left(\overline{\mathbf{M}}(E)\right) = |\overline{\mathbf{M}}(E)| = 0 \tag{3.4.22}$$

The above is the transcendental eigenequation from which the eigenvalues E can be derived. The nontrivial solutions for $\mathbf{v}$ are in the null space of $\overline{\mathbf{M}}$. Having known $\mathbf{v} = [A_1, A_2, B_2, B_3]^t$, the eigenfunctions of the Schrödinger equation can be constructed. Notice that $\mathbf{v}$ is known to an arbitrary multiplicative constant. The normalization of the eigenfunction will pin down the value of this constant.

When the potential well is symmetric such that $V_1 = V_3 = V_0$, then the solutions can be decomposed into odd and even solutions about $x = 0$. In this case, either $A_2 = 0$ for even modes, or $B_2 = 0$ for odd modes. Furthermore, $A_1 = \pm B_3$ for these modes. The problem then has two unknowns, and two boundary conditions at one of the interfaces suffice to deduce the eigenequation.

The above problem is analogous to the 1D dielectric waveguide problem in classical electromagnetics. In most textbooks, the transcendental eigenequation is solved using a graphical method, which can be done likewise here. The plot of the eigenmodes and their energy levels are shown in Figure 3.4. It is an interesting example showing that a trapped electron exists with different energy levels. The more energetic (more kinetic energy) the electron is, the higher the energy level. For the case when $E > V_0$, the wave function is not evanescent outside the well, and the electron is free to roam outside the well.

Modern technology has allowed the engineering of nano-structures so small that a quantum well can be fabricated. Quantum well technology is one of the emerging nano-technologies. It will be shown later in the course that an electron wave propagating in a lattice is like an electron wave propagating in vacuum but with a different effective mass and seeing a potential that is the potential of the conduction band. The quantum wells are fabricated with III-V compound, for example, with alloys of the form $Al_xGa_{1-x}As$ forming a ternary compound. They have different values of valence and conduction band depending on the value of x. By growing heterostructure layers with different compounds, multiple quantum wells can be made. Hence, the energy levels of a quantum well can also be engineered so that laser technology of different wavelengths can be fabricated.[6]

In the case of a hydrogen atom, the Coulomb potential around the proton at the nucleus yields a potential well described by $-q^2/(\epsilon 4\pi r)$. This well can trap an electron into various eigenstates, yielding different electronic orbitals. The Schrödinger equation has predicted the energy levels of a hydrogen atom with astounding success.

3.5   The Quantum Harmonic Oscillator–A Preview

The pendulum, a particle attached to a spring, or many vibrations in atoms and molecules can be described as a harmonic oscillator. Hence, the harmonic oscillator is one of the most

---

[6]J. H. Davis, *The Physics of Low Dimensional Semiconductors: An Introduction*, Cambridge U Press, 1998.
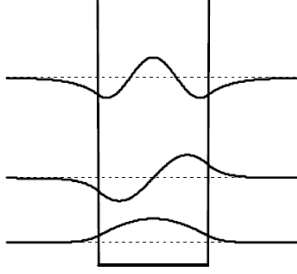
Figure 3.4: Trapped modes representing the electron wave functions inside a 1D potential well (from DAB Miller).

important examples in quantum mechanics. Its quantum mechanical version can be described by the 1D Schrödinger equation.

The classical equation for a harmonic oscillator is given by

$$m\frac{\mathrm{d}^2 z}{\mathrm{d}t^2} = -Kz \tag{3.5.1}$$

The above is Newton's law, and K is the spring constant, and the force provided by the spring is Kz. We can rewrite the above as

$$\frac{\mathrm{d}^2 z}{\mathrm{d}t^2} = -\omega_0^2 z \tag{3.5.2}$$

where $\omega_0 = \sqrt{K/m}$, and m is the mass of the particle. The above has a time harmonic solution of the form $\exp(\pm i\omega_0 t)$ where ω_0 is the oscillation frequency. Since the force $F = -\partial V/\partial z$, the potential energy of a particle attached to a spring is easily shown to be given by

$$V(z) = \frac{1}{2}m\omega_0^2 z^2 \tag{3.5.3}$$

Consequently, the above potential energy can be used in the Schrödinger equation to describe the trapping of wave modes (or bound states). The kinetic energy of the particle is described by a term proportional to the square of the momentum operator. Hence, the corresponding 1D Schrödinger equation is

$$\left[-\frac{\hbar^2}{2m}\frac{d^2}{dz^2} + \frac{1}{2}m\omega_0^2 z^2\right]\psi_n(z) = E_n\psi_n(z) \tag{3.5.4}$$

with a parabolic potential well. It turns out that this equation has closed-form solutions, yielding the wave function for an eigenstate as given by

$$\psi_n(z) = \sqrt{\frac{1}{2^n n!}\sqrt{\frac{m\omega_0}{\pi\hbar}}}\, e^{-\frac{m\omega_0}{2\hbar}z^2} H_n\left(\sqrt{\frac{m\omega_0}{\hbar}}z\right) \tag{3.5.5}$$

where $H_n(x)$ is a Hermite polynomial, and the wave function is Gaussian tapered by $e^{-\frac{m\omega_0}{2\hbar}z^2}$. The energy of the eigenstate is given by

$$E_n = \left(n + \frac{1}{2}\right)\hbar\omega_0 \tag{3.5.6}$$

The energy levels are equally spaced $\hbar\omega_0$ apart. Even the lowest energy state, the ground state, has a nonzero energy of $\hbar\omega_0/2$ known as the zero-point energy. The higher energy states correspond to larger amplitudes of oscillation, and vice versa for the lower energy states. In order to kick the quantum harmonic oscillator from the low energy state to a level above, it needs a packet of energy of $\hbar\omega_0$, the quantization energy of a photon. The physics of quantized electromagnetic oscillations (photons) and quantized mechanical oscillations (phonons) is intimately related to the quantum harmonic oscillator.

It is to be noted that if $|\psi_n(z)|^2$ has the meaning of probability density function, then the average location, or the expectation value of the location, of the oscillating particle for these eigenstates is given by

$$\langle z \rangle = \int_{-\infty}^{\infty} dz\, z |\psi_n(z)|^2 \tag{3.5.7}$$

Since z is an odd function while $|\psi_n(z)|^2$ is an even function of z, the above integral is always zero. Since the potential energy of the particle is proportional to z^2, it is seen that

$$\langle z^2 \rangle = \int_{-\infty}^{\infty} dz\, z^2 |\psi_n(z)|^2 \tag{3.5.8}$$

is not zero. Therefore, the potential energy of the particle for each of these eigenmodes is not zero, while the average location is zero. Moreover, their associated probability density functions are time-independent. These stationary states hence do not have a classical counterpart, and are termed non-classical states. They are also called the photon-number states since the n-th state corresponds to having the energy of n photons each with energy of $\hbar\omega_0$.
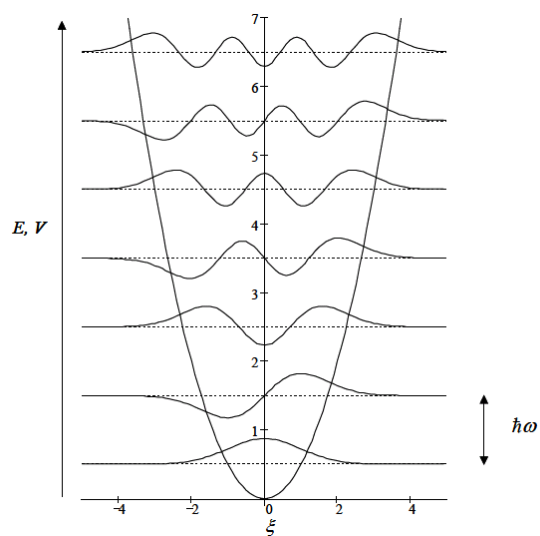
Figure 3.5: Sketch of the eigenstates, energy levels, and the potential well of a quantum harmonic oscillator (picture from DAB Miller).

Chapter 4

Time-Dependent Schrödinger Equation

4.1 Introduction

Each eigenstate of Schrödinger equation has its own time dependence of $\exp(-i\omega_n t)$. When we consider one eigenstate alone, its time dependence is unimportant, as the time dependence disappears when we convert the wavefunction into a probability density function by taking its magnitude squared. Moreover, the absolute frequency of an eigenstate is arbitrary. Hence, the probability density function is quite uninteresting. This is an antithesis to the classical harmonic oscillator where the position of the particle moves with respect to time.

However, when a quantum state is described by a wavefunction which is a linear superposition of two eigenstates, it is important that we take into account their individual frequency value and time dependence. The two eigenstates will "beat" with each other to produce a difference in frequency when we take the magnitude squared of the wavefunction.

4.2 Quantum States in the Time Domain

Consider a quantum state which is a linear superposition of two eigenstates

$$\psi(\mathbf{r}, t) = c_a e^{-i\omega_a t}\psi_a(\mathbf{r}) + c_b e^{-i\omega_b t}\psi_b(\mathbf{r}) \tag{4.2.1}$$

where c_a and c_b are properly chosen to normalize the corresponding probability density function. Then the probability function is

$$|\psi(\mathbf{r}, t)|^2 = |c_a|^2|\psi_a(\mathbf{r})|^2 + |c_b|^2|\psi_b(\mathbf{r})|^2 + 2\Re e\left[c_a c_b^*\psi_a(\mathbf{r})\psi_b^*(\mathbf{r})e^{-i(\omega_a-\omega_b)t}\right] \tag{4.2.2}$$

It is clearly time varying. We are seeing the motion of the particle through the potential well.

37
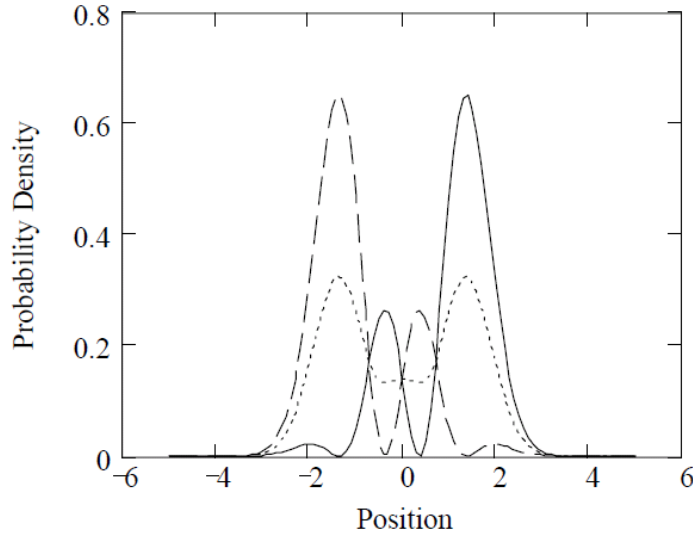
Figure 4.1: The time evolution of the linear superposition of the first two eigenstates of the simple quantum harmonic oscillator. (i) Beginning of a period (solid line). (ii) One quarter and three quarters way through the period (dotted line). (iii) Half way through the period (picture from DAB Miller).

4.3 Coherent State

The eigenstates of a quantum harmonic oscillator do not resemble its classical state. First, its magnitude squared is time independent, whereas a classical harmonic oscillator is time varying. In order to see the classical state emerging from the quantum harmonic oscillator eigenstates, we need to take a judicious linear superposition of them. Such a state is called the coherent state.[1] The coherent state is given by

$$\psi_N(\xi, t) = \sum_{n=0}^{\infty} c_{Nn} e^{-i(n+1/2)\omega_0 t} \psi_n(\xi) \tag{4.3.1}$$

where $\xi = \sqrt{m\omega_0/\hbar} z$, and $c_{Nn} = \sqrt{N^n e^{-N}/n!}$.[2] It forms a time-varying wave packet that emulates the motion of the classical harmonic oscillators such as a pendulum. The total energy of this wave packet is given by $(N + 1/2)\hbar\omega_0$ as shall be shown later. The larger N is, the more energy the particle has, and the closer is the coherent state to a classical state. A plot of the coefficient $|c_{Nn}|^2 = P_n$ as a function of n is shown in Fig. 4.2. Here, P_n is the probability of finding the state in the stationary state ψ_n. Notice the the probability of being in the eigenstates peak around $n = N$, the dominant eigenstate of the packet.

---

[1] The state was derived by R. Glauber who won the Nobel Prize in 2005 for his many contributions to quantum optics.

[2] The coefficient is also written as $e^{-|\alpha|^2/2} \frac{\alpha^n}{\sqrt{n!}}$ where $N = |\alpha|^2$.
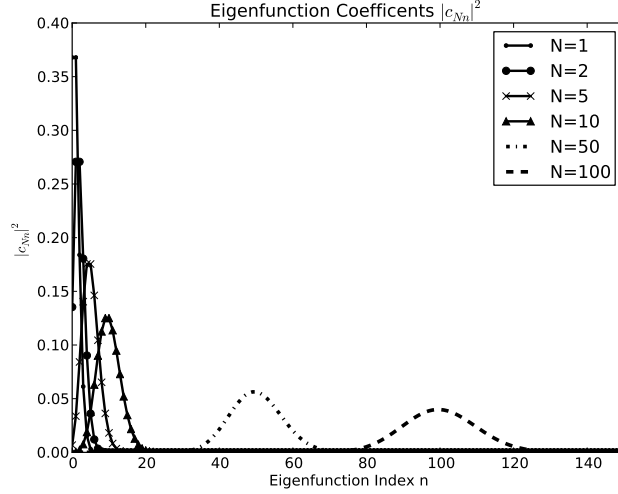
Figure 4.2: The plot of $|c_{Nn}|^2 = P_n$ as a function of n for different values of N.

As shall be seen, as N increases, the coherent state looks more like a localized wave packet describing a particle oscillating back and forth time harmonically. This is the case for a particle traveling in a parabolic potential well. Even when a particle is traveling in vacuum, it should be described by a localized wavefunction, such as a Gaussian wave packet. The study of the Gaussian wave packet is given in the Appendix. It can be seen that the classical limit emerges when the momentum of the particle becomes very large.

4.4   Measurement Hypothesis and Expectation Value

The wavefunction of a quantum system can be written as a linear superposition of the stationary states

$$\psi(\mathbf{r}, t) = \sum_n c_n(t)\psi_n(r) \tag{4.4.1}$$

where $c_n(t) = c_n(0)e^{-i\omega_n t}$ with $E_n = \hbar\omega_n$. The magnitude squared of this wavefunction should integrate to one due to its probabilistic interpretation. In other words,

$$\int_V |\psi(\mathbf{r}, t)|^2 d\mathbf{r} = \int_V \sum_n c_n(t)\psi_n(\mathbf{r}) \sum_{n'} c_{n'}^*(t)\psi_{n'}^*(\mathbf{r}) d\mathbf{r} = 1 \tag{4.4.2}$$

Exchanging the order of integration and summations, we arrive at

$$\int_V |\psi(\mathbf{r}, t)|^2 d\mathbf{r} = \sum_n \sum_{n'} c_n(t)c_{n'}^*(t) \int_V \psi_n(\mathbf{r})\psi_{n'}^*(\mathbf{r}) d\mathbf{r} \tag{4.4.3}$$
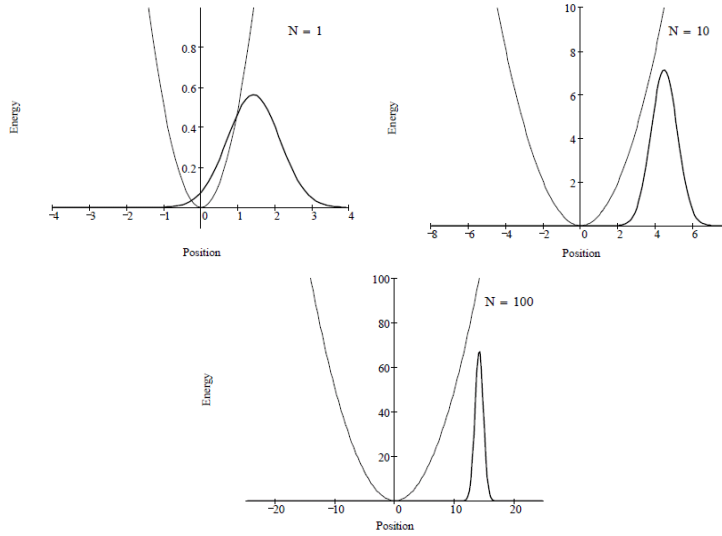
Figure 4.3: Coherent state for different values of N (picture from DAB Miller).

Using the orthonormality of the eigenstates, viz., $\int_V \psi_n(\mathbf{r})\psi_{n'}^*(\mathbf{r})d\mathbf{r} = \delta_{nn'}$, we arrive at the fact that

$$\sum_n |c_n(t)|^2 = 1 \tag{4.4.4}$$

Since the squares of the magnitudes add up to one, they can be assigned probabilistic interpretation as well. Hence, $P_n = |c_n(t)|^2$ represents the probability of finding the electron in eigenstate n.

The quantum measurement hypothesis states that before the measurement, the electron lives as a linear superposition of different eigenstates. After the measurement, the electron collapses into one of the eigenstates, say the n eigenstate, with probability proportional to $|c_n(t)|^2$.

The above can go further by saying that an electron is described by a wavefunction where its position is indeterminate: it can be found any place where the wavefunction is nonzero. Its probability of being found is proportional to the magnitude squared of the wavefunction. However, once the measurement is made to determine the location of the electron, the electron's position collapses to one location as is discovered by the experiment. Its position is determinate after the measurement.

One can think of the electron to be like "ghost" or "angel" which could exist as a linear superposition of many states before a measurement. After the measurement, it collapses to the state "discovered" by the measurement.

Due to this probabilistic interpretation, the expected energy of the quantum system is

given by

$$\langle E \rangle = \sum_n E_n P_n = \sum_n E_n |c_n(t)|^2 \tag{4.4.5}$$

The above is the expectation value of E. Notice that for the eigenstate $c_n(t) \propto e^{-i\omega_n t}$ and hence, $|c_n(t)|^2$ is time independent. The above expectation of the energy is a constant. This means that the quantum system is energy conserving, namely, it cannot lose energy, for instance, due to radiation. This is because this is an isolated quantum system. In order for this system to lose energy, it has to be coupled with other quantum systems, which is a subject of later study.

For the coherent state, the expectation value of E is

$$\langle E \rangle = \sum_{n=0}^{\infty} E_n \frac{N^n e^{-N}}{n!} \tag{4.4.6}$$

$$= \hbar\omega_0 \left[\sum_{n=0}^{\infty} n \frac{N^n e^{-N}}{n!} \right] + \frac{\hbar\omega_0}{2} \tag{4.4.7}$$

$$= \left(N + \frac{1}{2} \right) \hbar\omega_0 \tag{4.4.8}$$

where we have made use of the fact that $E_n = (n+1/2)\hbar\omega_0$, and that the above summations sum to N and 1.[3] Hence, the energy of the coherent state is that of N photons. The larger N is, the closer it is to a classical state.

The above expected energy value can also be found by taking the expectation of the Hamiltonian operator, viz.,

$$\langle E \rangle = \int d\mathbf{r} \, \psi^*(\mathbf{r}, t) \hat{H} \psi(\mathbf{r}, t) \tag{4.4.9}$$

The above equality can be easily shown by substituting (4.4.1) into the right-hand side and making use of the fact that $\psi_n(\mathbf{r})$ is the eigenfunction of the Hamiltonian operator

$$\hat{H} = -\frac{\hbar^2}{2m} \nabla^2 + V(\mathbf{r}) \tag{4.4.10}$$

In general, a quantity that is measurable, like the energy E, is replaced by an operator $\hat{H}$ in quantum mechanics. Such operators are called observables, such as momentum, position, etc. The relation between the classical value of an observable and its quantum mechanical counterpart, the operator, is via the expectation relation, viz.,

$$\left\langle \hat{A} \right\rangle = \int d\mathbf{r} \, \psi^*(\mathbf{r}, t) \hat{A} \psi(\mathbf{r}, t) \tag{4.4.11}$$

---

[3]It can be shown easily that $\sum_0^{\infty} \frac{N^n}{n!} = e^N$. Using the fact that $nN^n = N\frac{\partial}{\partial N} N^n$, it can be shown that $\sum_0^{\infty} nN^n = Ne^N$.

The above reduces to a scalar number, but it could still be time dependent. It is also the bridge between classical quantities and quantum mechanical entities.

It is seen that

$$\left\langle \hat{H} \right\rangle = \int d\mathbf{r}\psi^*(\mathbf{r},t)\hat{H}\psi(\mathbf{r},t) = \langle E \rangle = \sum_n E_n P_n \tag{4.4.12}$$

in accordance with (4.4.5) and (4.4.9). One can also do a direct substitution of (4.4.1) into the above to arrive at

$$\left\langle \hat{H} \right\rangle = \int d\mathbf{r} \sum_{n'} c_{n'}^*(t)\psi_{n'}^*(\mathbf{r})\hat{H} \sum_n c_n(t)\psi_n(\mathbf{r}) = \sum_{n'}\sum_n c_{n'}^*(t)c_n(t)E_n \int_V \psi_{n'}^*(\mathbf{r})\psi_n(\mathbf{r})d\mathbf{r}$$
$$= \sum_n |c_n(t)|^2 E_n = \sum_n E_n P_n = \langle E \rangle \tag{4.4.13}$$

where $\hat{H}\psi_n(\mathbf{r}) = E_n\psi_n(\mathbf{r})$ and the orthogonality of the eigenfunctions have been used. Hence, the expectation of the Hamiltonian operator with respect to a quantum state is the average of its eigenvalues with respect to the quantum state.

4.4.1 Uncertainty Principle–A Simple Version

The uncertainty principle is asserted by the fact that Fourier transform of a Gaussian function is another Gaussian function. For instance, if we have a wave packet that is formed by superposing waves with different k or momentum, we can express it as

$$\psi(z) = \int_{-\infty}^{\infty} \tilde{\psi}(k)e^{ikz}dk \tag{4.4.14}$$

where

$$\tilde{\psi}(k) = Ce^{-\left(\frac{k-k_0}{2\Delta k}\right)^2} \tag{4.4.15}$$

The above represents a Gaussian-tapered function with a spread of wavenumbers centered around $k = k_0$. It can be Fourier transformed in closed form. First, we find the Fourier transform of a Gaussian, namely,

$$\psi_g(z) = \int_{-\infty}^{\infty} e^{-k^2}e^{ikz}dk = \int_{-\infty}^{\infty} e^{-(k-iz/2)^2}e^{-z^2/4}dk = e^{-z^2/4}\int_{-\infty}^{\infty} e^{-\eta^2}d\eta \tag{4.4.16}$$

The second equality follows by completing the square of the exponent. The last equality follows by a change of variable by letting $k - iz/2 = \eta$. The integration of a Gaussian function can be obtained in closed form.[4] Subsequently,

$$\psi_g(z) = e^{-z^2/4}\sqrt{\pi} \tag{4.4.17}$$

---

[4]The integral can be evaluated in closed form by using the identity that $I = \int_{-\infty}^{\infty} dx e^{-x^2} = \sqrt{\pi}$. This can be proved by noticing that $I^2 = \int_{-\infty}^{\infty} dx e^{-x^2} \int_{-\infty}^{\infty} dy e^{-y^2} = \iint_{-\infty}^{\infty} dx dy e^{-(x^2+y^2)} = 2\pi\int_0^{\infty} \rho d\rho e^{-\rho^2} = \pi$. The last integral can be integrated in closed form.

proving the fact that the Fourier transform of a Gaussian is a Gaussian. The scaling property of Fourier transform says that

$$\psi_s(z) = \int_{-\infty}^{\infty} \tilde{\psi}(k/a)e^{ikz}dk = a\int_{-\infty}^{\infty} \tilde{\psi}(k/a)e^{i(k/a)az}d(k/a) = a\psi(az) \tag{4.4.18}$$

The above implies that if $\psi(z)$ is the Fourier transform of $\tilde{\psi}(k)$, then $a\psi(az)$ is the Fourier transform of $\tilde{\psi}(k/a)$. This together with the shifting property of Fourier transform allows us to deduce that the Fourier transform of (4.4.15) is

$$\psi(z) = 2\Delta k C e^{-(\Delta k z)^2 + ik_0 z} \tag{4.4.19}$$

The probability density functions are proportional to the magnitude squared of the above. The first pulse, whose probability density function is given by $|\tilde{\psi}(k)|^2 = C^2 e^{-(k-k_0)^2/[2(\Delta k)^2]}$, has standard deviation of Δk.[5] The standard deviation of $|\psi(z)|^2 \propto e^{-(2\Delta k z)^2/2}$ is $\Delta z = 1/(2\Delta k)$. The product of these two standard deviations yields

$$\Delta k \Delta z = \frac{1}{2} \tag{4.4.20}$$

or

$$\Delta p \Delta z = \frac{\hbar}{2} \tag{4.4.21}$$

4.4.2   Particle Current

It turns out that even though the state of an electron is defined by a wavefunction, other equations do not see this wavefunction. For instance, Maxwell's equations, will produce electric field from sources, but they will only produce the electric field from the charge cloud and the current produced by the charge cloud. We see that for stationary states of a trapped electron in a potential well, the charge cloud is static. Hence, it will not radiate according to electromagnetic theory. This resolves the conflict in early days as to why the electron, supposedly "orbiting" around the nucleus of an atom, does not radiate.

However, when an electron is in more than one stationary state, the charge cloud is time varying, and can potentially couple to an external electric field and radiate.[6] For conservation of charge, we have the relation that

$$\frac{\partial \rho_p}{\partial t} = -\nabla \cdot \mathbf{J}_p \tag{4.4.22}$$

where ρ_p is the particle density, and $\mathbf{J}_p$ is the particle current density. The particle density $\rho_p(\mathbf{r}, t) = |\psi(\mathbf{r}, t)|^2$. We can take the time derivative of ρ_p to yield

$$\frac{\partial}{\partial t}[\psi^*(\mathbf{r}, t)\psi(\mathbf{r}, t)] = \frac{\partial \psi^*(\mathbf{r}, t)}{\partial t}\psi(\mathbf{r}, t) + \psi^*(\mathbf{r}, t)\frac{\partial \psi(\mathbf{r}, t)}{\partial t} \tag{4.4.23}$$

---

[5] A Gaussian PDF of the form $Ce^{-x^2/(2\sigma^2)}$ has a standard deviation of σ.

[6] We will learn how to treat such coupling later on.

We can use Schrödinger equation to replace the time derivatives on the right-hand side to get

$$\frac{\partial}{\partial t}[\psi^*(\mathbf{r}, t)\psi(\mathbf{r}, t)] = -\frac{i}{\hbar}\left(\psi^*\hat{H}\psi - \psi\hat{H}^*\psi^*\right) \tag{4.4.24}$$

Substituting in for the definition of the Hamiltonian operator, we have further that

$$\frac{\partial}{\partial t}[\psi^*(\mathbf{r}, t)\psi(\mathbf{r}, t)] = \frac{i\hbar}{2m}\left(\psi^*\nabla^2\psi - \psi\nabla^2\psi^*\right) \tag{4.4.25}$$

$$\frac{\partial}{\partial t}[\psi^*(\mathbf{r}, t)\psi(\mathbf{r}, t)] = -\frac{i\hbar}{2m}\nabla \cdot (\psi\nabla\psi^* - \psi^*\nabla\psi) \tag{4.4.26}$$

Hence, we can define the particle current as

$$\mathbf{J}_p = \frac{i\hbar}{2m}\left(\psi\nabla\psi^* - \psi^*\nabla\psi\right) \tag{4.4.27}$$

When a stationary eigenstate is substituted into the above, due to the product of the function and its complex conjugate above, the particle current becomes time independent. Hence, stationary states can only give rise to non-time-varying current (static or DC current), and such current does not radiate according to electromagnetic theory.

Appendix

A   Introduction to Gaussian Wave Packet

In order to relate the wave nature to the particle nature of an object, it is necessary to obtain a wave packet picture of the object. Experimental observation has indicated that electrons are highly localized objects, so are photons. A wavefunction with a pure or single wavenumber k has equal amplitude everywhere. According to the probabilistic interpretation of quantum mechanics, it is untenable physically to think of a particle as all pervasive and can be found equally likely everywhere. A photon wave can have many k wavenumbers, and they can be linearly superposed to form a localized photon wave packet. Since the E-k relationship of the photon wave is linear (or almost linear in the material-media case), the formation of such a wave packet is quite straightforward. It is less clear for an electron wave, since the E-k relationship is nonlinear.

   In this appendix, we will explain the particle nature of an electron in the classical limit when the momentum of the electron becomes large. One expects that the wavefunction resembles that of a localized particle in this limit so that the electron is found with high probability only in a confined location. This understanding can be achieved by studying the Gaussian wave packet solution of Schrödinger equation.

   When the Schrödinger equation is solved in vacuum for electron, the solution is

$$e^{i\mathbf{k}\cdot\mathbf{r}}$$

This unlocalized solution means that the electron can be everywhere. A more realistic wavefunction for an electron is a wave packet which is localized. This is more akin to the motion that an electron is a localized particle. A wave packet can be constructed by linear superposing waves of different momenta $\hbar\mathbf{k}$ or different energies $\hbar\omega$. We will derive the Gaussian wave packet solution to Schrödinger equation. This can be constructed by studying the solution to the wave equation.

B   Derivation from the Wave Equation

It is well known that the wave equation

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} + k^2\right)\phi(x,y,z) = 0 \tag{B1}$$

admits solution of the form

$$\phi(x,y,z) = C\frac{e^{\pm ik\sqrt{x^2+y^2+(z-ib)^2}}}{\sqrt{x^2+y^2+(z-ib)^2}} \tag{B2}$$

for

$$\sqrt{x^2+y^2+(z-ib)^2} \neq 0 \tag{B3}$$

The above corresponds to a spherical wave where the source point is located in the complex coordinates $x = 0$, $y = 0$, and $z = ib$. It was first suggested by G. Deschamps. From the above, it is quite clear that, after letting $z = vt$,

$$\phi(x, y, vt) = A \frac{e^{\pm ik\sqrt{x^2 + y^2 + (vt - ib)^2}}}{\sqrt{x^2 + y^2 + (vt - ib)^2}} \tag{B4}$$

is a solution to

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{1}{v^2} \frac{\partial^2}{\partial t^2} + k^2 \right) \phi(x, y, vt) = 0 \tag{B5}$$

The above equation can be factored

$$\left(v^{-1} \partial_t - i\sqrt{k^2 + \partial_x^2 + \partial_y^2} \right) \left(v^{-1} \partial_t + i\sqrt{k^2 + \partial_x^2 + \partial_y^2} \right) \phi(x, y, vt) = 0 \tag{B6}$$

where $\partial_t = \partial/\partial t$, $\partial_x^2 = \partial^2/\partial x^2$ and so on. In the above, function of an operator has meaning only when the function is Taylor expanded into an algebraic series. One can assume that $k^2 \to \infty$, while ∂_x^2 and ∂_y^2 are small.[7] Taylor expanding the above and keeping leading order terms only, we have

$$\left(v^{-1} \partial_t - ik - \frac{i}{2k} \left(\partial_x^2 + \partial_y^2 \right) \right) \left(v^{-1} \partial_t + ik + \frac{i}{2k} \left(\partial_x^2 + \partial_y^2 \right) \right) \phi(x, y, vt) \cong 0 \tag{B7}$$

In (B4), we can let[8] $|vt - ib|^2 \gg x^2 + y^2$ to arrive at the approximation

$$\phi(x, y, vt) \approx A \frac{e^{\pm \left(ik(vt - ib) + ik \frac{x^2 + y^2}{vt - ib} \right)}}{vt - ib} \tag{B8}$$

It can be shown that when we pick the plus sign above in the $\pm$ sign, the above is the exact solution to

$$\left(v^{-1} \partial_t - ik - \frac{i}{2k} \left(\partial_x^2 + \partial_y^2 \right) \right) \phi_+(x, y, vt) = 0 \tag{B9}$$

where

$$\phi_+(x, y, vt) = A e^{ik(vt - ib)} \frac{e^{ik \frac{x^2 + y^2}{2(vt - ib)}}}{2(vt - ib)} = e^{ik(vt - ib)} \psi(x, y, t) \tag{B10}$$

and

$$\psi(x, y, t) = -A_0 \frac{ib e^{ik \frac{x^2 + y^2}{2(vt - ib)}}}{(vt - ib)} \tag{B11}$$

Furthermore, $\psi(x, y, t)$ is an exact solution to

$$\left(v^{-1} \partial_t - \frac{i}{2k} \left(\partial_x^2 + \partial_y^2 \right) \right) \psi(x, y, t) = 0 \tag{B12}$$

---

[7]This is known as the paraxial wave approximation. It implies that the variation of the solution is mainly in the vt direction with slow variation in the x and y directions.

[8]It can be shown that when this is valid, the paraxial wave approximation used in (B7) is good.

If we multiply the above by $i\hbar v$, the above becomes

$$\left(i\hbar\partial_t + \frac{\hbar v}{2k} \left(\partial_x^2 + \partial_y^2 \right) \right) \psi(x, y, t) = 0 \tag{B13}$$

By letting $v = \hbar k/m$, the above becomes

$$\left(i\hbar\partial_t + \frac{\hbar^2}{2m} \left(\partial_x^2 + \partial_y^2 \right) \right) \psi(x, y, t) = 0 \tag{B14}$$

which is just Schrödinger equation. Equation (B11) represents the Gaussian wave packet solution of Schrödinger equation. It can be studied further to elucidate its physical contents.

C   Physical Interpretation

In (B11), one can write

$$\frac{k\left(x^2 + y^2\right)}{2\left(vt - ib\right)} = \frac{k\left(x^2 + y^2\right)\left(vt + ib\right)}{2\left(v^2t^2 + b^2\right)} = \frac{k\left(x^2 + y^2\right)}{2R} + i\frac{x^2 + y^2}{W^2} \tag{C1}$$

where

$$R = \frac{v^2t^2 + b^2}{vt}, \qquad W^2 = \frac{2b}{k}\left(1 + \frac{v^2t^2}{b^2}\right) \tag{C2}$$

Then Gaussian wave packet can be more suggestively written as

$$\psi(x, y, t) = \frac{A_0}{\sqrt{1 + v^2t^2/b^2}} e^{-\frac{x^2+y^2}{W^2}} e^{ik\frac{x^2+y^2}{2R}} e^{-i\varphi(t)} \tag{C3}$$

where

$$\varphi(t) = \tan^{-1}(vt/b) \tag{C4}$$

The above reveals a wave packet which is Gaussian tapered with width W and modulated by oscillatory function of space and time. However, the width of this packet is a function of time as indicated by (C2).

At $vt = 0$, the width of the packet is given by

$$W_0 = \sqrt{\frac{2b}{k}}, \quad \text{or} \quad b = \frac{1}{2}kW_0^2 \tag{C5}$$

This width can be made independent of k if b is made proportional to k. Nevertheless, as time progresses with $vt > 0$, the width of the packet grows according to (C2). However, to maintain a fixed-width W_0, it is necessary that $b \to \infty$ as $k \to \infty$. Subsequently, the effect of v^2t^2/b^2 becomes small in (C2) as $b \to \infty$. This means that the width of the Gaussian wave packet remains almost constant for the time when $vt \leq b$, but b is a large number proportional to k. The duration over which the Gaussian wave packet's width remains unchange becomes increasingly longer as k becomes larger.

In the above, $\hbar k$ represents the momentum of a particle. When the particle carries high momentum, it can be represented by a Gaussian wave packet that hardly changes in shape with respect to time. This is what is expected of a classical picture of a moving particle: the Gaussian wave packet does reproduce the classical picture of a high momentum particle.

It is quite easy to design a wave packet for a photon that does not spread with time by linear superposing waves with different frequencies. This is because the ω-k diagram (or E-k diagram, since $E = \hbar\omega$) for photons is a straight line. It implies that all waves with different k's travel with the same phase velocity. These Fourier modes are locked in phase, and the pulse shape does not change as the wave packet travels.
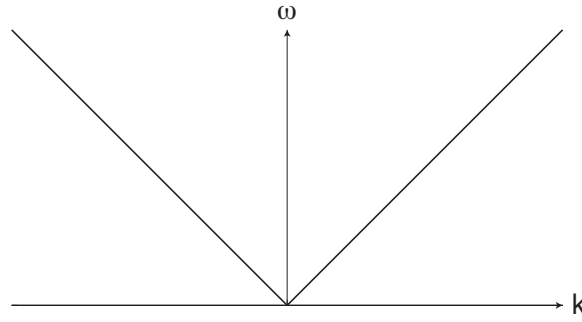
Figure C1: $\omega - k$ diagram for a photon wave which is a straight line.

But for electron wave in a vacuum, the ω-k diagram is quadratic. It implies that waves with different k numbers travel with different phase velocity, giving rise to distortion of the pulse shape. But if we have a narrow band pulse operating in the high k regime, the ω-k diagram is quasi-linear locally, and there should be little pulse spreading. Hence, one can construct a quasi-distortionless pulse in the high k regime.
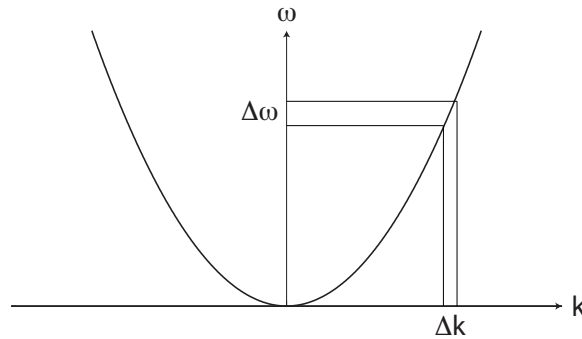
Figure C2: $\omega - k$ diagram of an electron wave in vacuum.

D Stability of the Plane Wave Solution

It is quite obvious from our understanding of quantum mechanics, wave packets, and coherence, that the plane wave solution e^{ikx} is not stable as $k \to \infty$. One can always express

$$e^{ikx} = \int_{-\infty}^{\infty} dx' e^{ikx'} \delta(x - x') \tag{D1}$$

If one can think of $\delta(x - x')$ as the limiting case of a wave packet, the above implies that a plane wave can be expanded as a linear superposition of wave packets at each location x, but bearing a phase $\exp(ikx)$. As $k \to \infty$, this phase is rapidly varying among the different wave packets. Hence, their coherence is extremely difficult to maintain, and upset easily by coupling to other quantum systems that exists in the environment. In other words, the particle cannot be stably maintained in the plane-wave state classically: it has to collapse to a wave-packet state. Consequently, classically, particles are localized when its momentum k becomes large.

Chapter 5

More Mathematical Preliminaries

5.1 A Function is a Vector

A function is actually an equivalence of a vector. In linear algebra, we denote a vector as $\mathbf{v} = [v_1, v_2, v_3, \ldots v_N]^t$. A short-hand notation for this vector is v_j where $j = 1, \ldots, N$. But for a function $f(x)$, the countably finite index j in v_j is now replaced by an uncountably infinite (nondenumerable) set of indices denoted by x. Hence, we can think of $f(x)$ as a vector in an infinite dimensional vector space. An inner product in linear algebra is written as

$$\mathbf{v}^\dagger \cdot \mathbf{u} = \sum_{j=1}^{N} v_j^* u_j \tag{5.1.1}$$

The analogue of the above for functions is

$$\langle f|g \rangle = \int_{-\infty}^{\infty} dx f(x)^* g(x) \tag{5.1.2}$$

The above is the inner product between two functions which are actually vectors.[1] The left-hand side is the compact notation for the inner product between two state vectors known as Dirac's bra-ket notation. The $\langle f|$ is the "bra" while $|g\rangle$ is the "ket". The "bra" can be thought of as the conjugate transpose of a vector, while the "ket" is analogous to an ordinary vector.

The set of functions, just as a set of vectors, can then span a vector space.[2] A vector space for which an inner product is defined as above is known as an inner product space.

---

[1] In the mathematics literature, the inner product is often denoted as $\langle f, g \rangle$ or (f, g).

[2] When any vector in the space can be represented as a linear superposition of the set of independent vectors, the set of vectors are said to span the vector space. They also form the basis set for spanning the space.

Moreover, an inner product space which is complete is a Hilbert space.[3] Hilbert spaces can be infinite dimensional. The above inner product facilitates the definition of a norm since $\langle f|f \rangle$ is a positive definite number. Hence, the norm of a vector can be defined to be

$$\|f\| = (\langle f|f \rangle)^{1/2} \tag{5.1.3}$$

It is the measure of the length of the vector. We can use this norm to define the distance between two vectors to be

$$d(f, g) = \|f - g\| \tag{5.1.4}$$

In general, if we have a set of orthonormal eigenfunctions, $\{\psi_n(x), n = 1, \dots \infty\}$, that spans a linear vector space, we can expand an arbitrary function in the same space as

$$g(x) = \sum_n d_n \psi_n(x) \tag{5.1.5}$$

The set $\{\psi_n(x), n = 1, \dots \infty\}$ also forms the orthonormal basis or the orthonormal basis set for spanning the vector space. A member of the set is known as a basis function or a basis vector.[4] Eigenfunctions of an operator can be used as basis functions.

The above can be written using Dirac's notation as

$$|g\rangle = \sum_n d_n |\psi_n\rangle \tag{5.1.6}$$

Inner product the above with $\langle \psi_m|$ from the left, we arrive at that

$$\langle \psi_m|g\rangle = \sum_n d_n \langle \psi_m|\psi_n\rangle \tag{5.1.7}$$

Using the orthonormality of the eigenfunction such that $\langle \psi_m|\psi_n\rangle = \delta_{nm}$, the above yields that

$$d_m = \langle \psi_m|g\rangle \tag{5.1.8}$$

Consequently, we have

$$|g\rangle = \sum_n |\psi_n\rangle \langle \psi_n|g\rangle \tag{5.1.9}$$

We can identify the operator

$$\hat{I} = \sum_n |\psi_n\rangle \langle \psi_n| \tag{5.1.10}$$

---

[3]It was studied by David Hilbert (1862-1943). Later Stefan Banach (1892-1945), and Sergei Sobolev (1908-1989) studied the generalization of Hilbert spaces, which are known respectively as Banach spaces and Sobolev spaces.

[4]We will use "function" and "vector" interchangeably since they are the same.

as an identity operator since when it operates on a vector, it returns the same vector. We can often construct an identity operator of a vector space once we have identify the set of orthonormal vectors that spans the space. The product $|\psi_n\rangle\langle\phi_n|$ in Dirac's notation, is analogous to outer product between two vectors $\mathbf{u}\cdot\mathbf{v}^\dagger$ in linear algebraic notation. This outer product produces a matrix or an operator.

For instance, if we are in a 3-space (the short for 3D vector space), the unit vectors that span the 3-space are $\{\hat{a}_1, \hat{a}_2, \hat{a}_3\}$. They are complete and orthonormal. Hence, the identity operator in a 3-space is[5]

$$\bar{\mathbf{I}} = \sum_{i=1}^{3} \hat{a}_i \hat{a}_i \tag{5.1.11}$$

To be strictly correct, the above is an outer product between two vectors, and a transpose sign should be attached to one of the unit vectors. But it is customary in the physics literature to ignore this transpose sign. Written in terms of x, y, z notation, the identity operator becomes

$$\bar{\mathbf{I}} = \hat{x}\hat{x} + \hat{y}\hat{y} + \hat{z}\hat{z} \tag{5.1.12}$$

One can easily be convinced that $\bar{\mathbf{I}} \cdot \mathbf{a} = \mathbf{a}$, confirming the above.[6]

In an N dimensional vector space spanned by a set of orthonormal vectors $\{\mathbf{U}_n, n = 1, \ldots, N\}$, the identity operator is formed by their outer products; namely,

$$\bar{\mathbf{I}} = \sum_{n=1}^{N} \mathbf{u}_n \cdot \mathbf{u}_n^\dagger \tag{5.1.13}$$

One can easily show that when $\bar{\mathbf{I}}$ operates on a member of the vector space, it returns the same member. Namely,

$$\bar{\mathbf{I}} \cdot \mathbf{v} = \sum_{n=1}^{N} \mathbf{u}_n \cdot \mathbf{u}_n^\dagger \cdot \mathbf{v} = \sum_{n=1}^{N} \mathbf{u}_n \left(\mathbf{u}_n^\dagger \cdot \mathbf{v}\right) = \mathbf{v} \tag{5.1.14}$$

The last equality follows because the second last term is just an orthonormal eigenvector expansion of the vector $\mathbf{v}$.

A vector space is also defined by a set of axioms. For $\mathbf{u}, \mathbf{v}, \mathbf{w}$ that belong to a vector

---

[5]Throughout these lecture notes, the hat ("ˆ") symbol is used to denote an operator, but here, it denotes a unit vector.

[6]It is to be noted that an outer product in matrix notation is $\mathbf{u} \cdot \mathbf{v}^\dagger$, while in physics notation for 3 space, it is often just written as $\mathbf{EB}$. In Dirac notation, an outer product is $|\psi\rangle\langle\phi|$.

space, the following axioms hold:

Associativity of addition	$\mathbf{u} + (\mathbf{v} + \mathbf{w}) = (\mathbf{u} + \mathbf{v}) + \mathbf{w}$	(5.1.15)
Commutativity of addition	$\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$	(5.1.16)
Identity element of addition	$\mathbf{v} + \mathbf{0} = \mathbf{v}$	(5.1.17)
Inverse elements of addition	$\mathbf{v} + (-\mathbf{v}) = \mathbf{0}$	(5.1.18)
Distributivity of scalar multiplication over vectors	$a(\mathbf{u} + \mathbf{v}) = a\mathbf{u} + a\mathbf{v}$	(5.1.19)
Distributivity of scalar multiplication by a vector	$(a + b)\mathbf{v} = a\mathbf{v} + b\mathbf{v}$	(5.1.20)
Compatibility of scalar multiplication	$a(b\mathbf{v}) = (ab)\mathbf{v}$	(5.1.21)
Identity element of scalar multiplication	$1\mathbf{v} = \mathbf{v}$	(5.1.22)

5.2 Operators

An operator maps vectors from one space to vectors in another space. It is denoted mathematically as $\hat{A} : V \to W$ where $\hat{A}$ is the operator, while V and W are two different vector spaces: V is the domain space, while W is the range space. In linear algebra, the operator is a matrix operator. In Hilbert spaces, it can be a differential operator such as

$$g(x) = \frac{d}{dx} f(x) \tag{5.2.1}$$

It can be an integral operator such as a Fourier transform operator

$$g(k) = \int_{-\infty}^{\infty} dx e^{ikx} f(x) \tag{5.2.2}$$

In Dirac notation,

$$|g\rangle = \hat{A}|f\rangle \tag{5.2.3}$$

Linear operators are defined such that

$$\hat{L}(c_1|g_1\rangle + c_2|g_2\rangle) = c_1\hat{L}|g_1\rangle + c_2\hat{L}|g_2\rangle \tag{5.2.4}$$

It is quite clear that matrix operators satisfy the above, and hence, they are linear. They are also known as linear maps. In general, like matrix operators, linear operators are not commutative; namely

$$\hat{A}\hat{B} \neq \hat{B}\hat{A} \tag{5.2.5}$$

5.2.1 Matrix Representation of an Operator

An operator equation can be written as[7]

$$|g\rangle = \hat{A}|f\rangle \tag{5.2.6}$$

---

[7]In the mathematics literature, this is often just denoted as $g = \hat{A}f$.

We can convert the above into a matrix equation by inserting an identity operator on the right-hand side to give

$$|g\rangle = \sum_n \hat{A}|\psi_n\rangle\langle\psi_n|f\rangle \tag{5.2.7}$$

Furthermore, we can multiply the above from the left by the basis vector $\langle\psi_m|, m = 1, \ldots, \infty$ to yield[8]

$$\langle\psi_m|g\rangle = \sum_n \langle\psi_m|\hat{A}|\psi_n\rangle\langle\psi_n|f\rangle, \qquad m = 1, \ldots, \infty \tag{5.2.8}$$

The above is an infinite dimensional matrix equation which can be written as

$$\mathbf{g} = \overline{\mathbf{A}} \cdot \mathbf{f} \tag{5.2.9}$$

where

$$\left[\overline{\mathbf{A}}\right]_{mn} = \langle\psi_m|\hat{A}|\psi_n\rangle \tag{5.2.10}$$

$$[\mathbf{g}]_m = \langle\psi_m|g\rangle \tag{5.2.11}$$

$$[\mathbf{f}]_n = \langle\psi_n|f\rangle \tag{5.2.12}$$

The matrix equation can be solved approximately by truncating its size to $N \times N$, or without truncation, it can be solved iteratively.

As a simple example of an iterative solution, we write $\overline{\mathbf{A}} = \overline{\mathbf{D}} + \overline{\mathbf{T}}$ where $\overline{\mathbf{D}}$ is the diagonal part of $\overline{\mathbf{A}}$ while $\overline{\mathbf{T}}$ is its off-diagonal part. Then (5.2.9) can be rewritten as

$$\overline{\mathbf{D}} \cdot \mathbf{f} = \mathbf{g} - \overline{\mathbf{T}} \cdot \mathbf{f} \tag{5.2.13}$$

The above can be solved iteratively as

$$\overline{\mathbf{D}} \cdot \mathbf{f}_n = \mathbf{g} - \overline{\mathbf{T}} \cdot \mathbf{f}_{n-1}, \qquad n = 0, 1, 2, \ldots, \tag{5.2.14}$$

until convergence is reached, or that $\mathbf{f}_n$ does not change with n. More sophisticated iterative methods are available in the literature.[9]

The matrix denoted by

$$A_{mn} = \langle\psi_m|\hat{A}|\psi_n\rangle \tag{5.2.15}$$

is the matrix representation of the operator $\hat{A}$. By the same token, $\langle\psi_m|g\rangle$ and $\langle\psi_n|f\rangle$ are the vector representations of the functions g and f respectively.

In the above, we have assumed that the range space and the domain space of the operator are the same, and hence, they can be spanned by the same basis set. For a Hermitian operator, this is usually the case. However, for some operators where the range space and the domain space are different, we may choose to test (5.2.7) with a different set of basis functions.

---

[8]This process is called testing or weighting, and ψ_m is called the testing or weighting function.

[9]An example is the Krylov subspace method by Aleksey Krylov (1863-1945).

5.2.2   Bilinear Expansion of an Operator

We have seen how the use of the identity operator allows us to expand a function in terms of a set of basis functions as in (5.1.9). The same can be done with an operator. Pre- and post-multiply an operator with the identity operator as given by (5.1.10), we have

$$\hat{A} = \sum_n \sum_m |\psi_n\rangle\langle\psi_n|\hat{A}|\psi_m\rangle\langle\psi_m| \tag{5.2.16}$$

The above can be rewritten as

$$\hat{A} = \sum_n \sum_m |\psi_n\rangle A_{nm}\langle\psi_m| = \sum_n \sum_m A_{nm}|\psi_n\rangle\langle\psi_m| \tag{5.2.17}$$

where A_{nm} is the matrix representation of the operator $\hat{A}$. The above is the bilinear expansion of the operator in terms of orthonormal functions. Notice that the expansion of an identity operator given by (5.1.10) is a bilinear expansion.

5.2.3   Trace of an Operator

The trace of a matrix operator is defined to be the sum of its diagonal elements; namely,

$$\text{tr}\left(\overline{\mathbf{M}}\right) = \sum_{i=1}^{N} M_{ii} \tag{5.2.18}$$

If an operator $\hat{A}$ has matrix representation given by $\langle\psi_i|\hat{A}|\psi_j\rangle$, the trace of the operator $\hat{A}$ is defined to be

$$\text{tr}\left(\hat{A}\right) = \sum_i \langle\psi_i|\hat{A}|\psi_i\rangle \tag{5.2.19}$$

It can be shown that the trace of an operator is independent of the basis used for its representation. To this end, we insert the identity operator

$$\hat{I} = \sum_m |\phi_m\rangle\langle\phi_m| \tag{5.2.20}$$

into (5.2.19) to get

$$\text{tr}\left(\hat{A}\right) = \sum_i \sum_m \langle\psi_i|\phi_m\rangle\langle\phi_m|\hat{A}|\psi_i\rangle \tag{5.2.21}$$

Exchanging the order of summation above, and the order of the two scalar numbers in the summand, we have

$$\text{tr}\left(\hat{A}\right) = \sum_m \sum_i \langle\phi_m|\hat{A}|\psi_i\rangle\langle\psi_i|\phi_m\rangle \tag{5.2.22}$$

The inner summation reduces to an identity operator which can be removed, and the above becomes

$$\text{tr}\left(\hat{A}\right) = \sum_m \langle \phi_m | \hat{A} | \phi_m \rangle \tag{5.2.23}$$

This is the trace of $\hat{A}$ using another basis set that is complete and orthonormal. Hence, the trace of an operator is invariant with respect to the choice of basis.

If we choose a basis function that is the eigenfunction of $\hat{A}$ such that $\hat{A}|\psi_i\rangle = \lambda_i |\psi_i\rangle$, then (5.2.19) becomes

$$\text{tr}\left(\hat{A}\right) = \sum_i \lambda_i \langle \psi_i | \psi_i \rangle = \sum_i \lambda_i \tag{5.2.24}$$

Hence, the trace of an operator is also the sum of its eigenvalues.

It can also be shown that

$$\text{tr}\left(\hat{A}\hat{B}\right) = \text{tr}(\hat{B}\hat{A}) \tag{5.2.25}$$

This is quite easy to show for matrix operators since

$$\text{tr}\left(\overline{\mathbf{A}} \cdot \overline{\mathbf{B}}\right) = \sum_i \sum_j A_{ij} B_{ji} = \sum_j \sum_i B_{ji} A_{ij} \tag{5.2.26}$$

Trace is usually used in quantum mechanics as an alternative way to write the expectation value of an operator. As mentioned before, for a quantum system in a state defined by the state function $|\psi\rangle$, the expectation value of a quantum operator in such a state is

$$\langle \psi | \hat{A} | \psi \rangle \tag{5.2.27}$$

For denumerable indices, the above is analogous to

$$\mathbf{u}^\dagger \cdot \overline{\mathbf{A}} \cdot \mathbf{u} = \sum_n \sum_m u_n^* A_{nm} u_m = \sum_n \sum_m A_{nm} u_m u_n^* = \text{tr}\left(\overline{\mathbf{A}} \cdot \mathbf{u} \cdot \mathbf{u}^\dagger\right) \tag{5.2.28}$$

where $u_m u_n^*$ is the outer product of two vectors u_m and u_n. Converting the above to Dirac notation, we have

$$\langle \psi | \hat{A} | \psi \rangle = \text{tr}\{\hat{A} | \psi \rangle \langle \psi |\} \tag{5.2.29}$$

The operator

$$\hat{\rho} = |\psi\rangle\langle\psi| \tag{5.2.30}$$

is known as the density operator. It is used in quantum mechanics to denote the state of a quantum system as an alternative to the state vector.

5.2.4   Unitary Operators

Unitary operators satisfy the property that

$$\overline{\mathbf{U}}^{\dagger} \cdot \overline{\mathbf{U}} = \overline{\mathbf{I}} \tag{5.2.31}$$

Unitary operators are operators that when acted on a vector does not change its length. In matrix notation, a unitary transform is

$$\overline{\mathbf{U}} \cdot \mathbf{v} = \mathbf{v}' \tag{5.2.32}$$

The length squared of the new vector $\mathbf{v}'$ is defined to be

$$\mathbf{v}^{\dagger} \cdot \overline{\mathbf{U}}^{\dagger} \cdot \overline{\mathbf{U}} \cdot \mathbf{v} = \mathbf{v}'^{\dagger} \cdot \mathbf{v}' \tag{5.2.33}$$

Making use of (5.2.31), the above implies that

$$\mathbf{v}^{\dagger} \cdot \mathbf{v} = \mathbf{v}'^{\dagger} \cdot \mathbf{v}' \tag{5.2.34}$$

or that the length of the vector has not changed. Furthermore, it can be easily shown that unitary transformations preserve the inner product between two vectors, namely,

$$\mathbf{v}^{\dagger} \cdot \mathbf{u} = \mathbf{v}'^{\dagger} \cdot \mathbf{u}' \tag{5.2.35}$$

For unitary matrix, it is clear that $\overline{\mathbf{U}}^{\dagger} = \overline{\mathbf{U}}^{-1}$. The above can be rewritten using Dirac notation. Since in quantum mechanics, $\langle \psi | \psi \rangle = 1$, the length of the vector is 1. The time evolution operator by integrating Schrödinger equation given below

$$\hat{\tau} = e^{-\frac{i}{\hbar}\hat{H}t} \tag{5.2.36}$$

where $\hat{H}$ is the Hamiltonian operator, is a unitary operator, since the state vector it acts on cannot change its length.

Since many basis set comprises orthonormal basis vectors, a unitary matrix is needed for the change of basis from one set to another. For example, we can define a Fourier transform to be

$$f(x) = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{\infty} dk\, e^{ikx} \tilde{f}(k).$$

It is used to transform the momentum representation of a wavefunction to the coordinate representation. Using the above, we can show that

$$\int dx\, f^{*}(x)\, g(x) = \int dk\, \tilde{f}^{*}(k)\, \tilde{g}(k).$$

The above can be proved using the fact that

$$\frac{1}{2\pi} \int\limits_{-\infty}^{\infty} e^{i(k'-k)x}\, dx = \delta(k'-k)$$

Hence, a Fourier transform operator is a unitary operator that preserves inner products between two vectors and their lengths.

5.2.5   Hermitian Operators

Hermitian operators appear frequently in quantum mechanics since they have real eigenvalues. As such, the expectation value of such operators are real valued, so that it can be connected to measurable or observable quantities in the real world.

A Hermitian matrix is one where its conjugate transpose (also called Hermitian transpose or Hermitian conjugate) is itself. For operators, the term "adjoint" is often used. A Hermitian operator is also called a self-adjoint operator. The notation for this is

$$\hat{M}^{\dagger} = \hat{M} \tag{5.2.37}$$

For matrix operator with denumerable indices, the above is the same as

$$M_{ij}^{*} = M_{ji} \tag{5.2.38}$$

For operators with nondenumerable indices, it is better to define the adjoint of the operator via inner products. Using the rule for conjugate transpose, we have

$$\left(\mathbf{a}^{\dagger} \cdot \overline{\mathbf{M}} \cdot \mathbf{b}\right)^{*} = \mathbf{b}^{\dagger} \cdot \overline{\mathbf{M}}^{\dagger} \cdot \mathbf{a} \tag{5.2.39}$$

Generalizing the above to infinite dimensional space, we have

$$\langle f|\hat{M}|g\rangle^{*} = \langle f|\hat{M}|g\rangle^{\dagger} = \langle g|\hat{M}^{\dagger}|f\rangle \tag{5.2.40}$$

The first equality follows because the above are scalar quantities: hence, the conjugate transpose of a scalar is the same as its conjugation. To obtain the second equality, we have used the rule that

$$\left(\overline{\mathbf{A}} \cdot \overline{\mathbf{B}} \cdot \overline{\mathbf{C}}\right)^{t} = \overline{\mathbf{C}}^{t} \cdot \overline{\mathbf{B}}^{t} \cdot \overline{\mathbf{A}}^{t} \tag{5.2.41}$$

in linear algebra, where the matrices need not be square. Hence, $\overline{\mathbf{A}}$ and $\overline{\mathbf{C}}$ can be $1 \times N$ and $N \times 1$ respectively representing vectors. This identity immediately implies that

$$\left(\overline{\mathbf{A}} \cdot \overline{\mathbf{B}} \cdot \overline{\mathbf{C}}\right)^{\dagger} = \overline{\mathbf{C}}^{\dagger} \cdot \overline{\mathbf{B}}^{\dagger} \cdot \overline{\mathbf{A}}^{\dagger} \tag{5.2.42}$$

These rules can easily be shown to apply to inner product of operators and vectors under the Dirac notation. Hence, (5.2.40) defines the adjoint of the operator $\hat{M}^{\dagger}$ when f and g are arbitrary. It does so by using just inner products.

A Hermitian operator has real eigenvalues and their eigenvectors are orthogonal. The proof is analogous to the proof for matrix operators, for which we will provide next. Given

$$\overline{\mathbf{M}} \cdot \mathbf{v}_i = \lambda_i \mathbf{v}_i \tag{5.2.43}$$

$$\overline{\mathbf{M}} \cdot \mathbf{v}_j = \lambda_j \mathbf{v}_j \tag{5.2.44}$$

Dot-multiply the first equation by $\mathbf{v}_j^{\dagger}$ from the left, and likewise for the second equation with $\mathbf{v}_i^{\dagger}$, we have

$$\mathbf{v}_j^{\dagger} \cdot \overline{\mathbf{M}} \cdot \mathbf{v}_i = \lambda_i \mathbf{v}_j^{\dagger} \cdot \mathbf{v}_i \tag{5.2.45}$$

$$\mathbf{v}_i^{\dagger} \cdot \overline{\mathbf{M}} \cdot \mathbf{v}_j = \lambda_j \mathbf{v}_i^{\dagger} \cdot \mathbf{v}_j \tag{5.2.46}$$

We can take the conjugate transpose of the first equation, using the rule in (5.2.42), the left-hand side becomes the same as that for the second equation after making use of the Hermitian property of the matrix $\overline{\mathbf{M}}$. On subtracting the two equations, the following equation ensues:

$$0 = (\lambda_i^* - \lambda_j)\mathbf{v}_i^\dagger \cdot \mathbf{v}_j \tag{5.2.47}$$

If $i = j$, on the right-hand side, we have $\mathbf{v}_i^\dagger \cdot \mathbf{v}_i = |\mathbf{v}_i|^2$ which is a positive definite number. The above can be zero only if $\lambda_i^* = \lambda_i$ or that λ_i is real. On the other hand, if $\lambda_i \neq \lambda_j$, it is necessary that $\mathbf{v}_i^\dagger \cdot \mathbf{v}_j = 0$. In other words,

$$\mathbf{v}_i^\dagger \cdot \mathbf{v}_j = C_n \delta_{ij} \tag{5.2.48}$$

The above proof can be repeated using Dirac notation, and the conclusion will be the same. The eigenvectors of a Hermitian operator are also complete in the space that the operator acts on. It is obvious in the finite dimensional case, but not so obvious in the infinite dimensional space.

We can use the rule expressed in (5.2.40) to see if an operator is Hermitian. For instance, the momentum operator is $\hat{p}$. Using (5.2.40), we have

$$\langle f|\hat{p}|g\rangle^* = \langle g|\hat{p}^\dagger|f\rangle \tag{5.2.49}$$

The above defines the adjoint of the operator $\hat{p} = -i\hbar d/dx$. Writing the above explicitly in 1D space using coordinate space representation,[10] we have on the left-hand side

$$\int_{-\infty}^{\infty} dx f(x) i\hbar \frac{d}{dx} g^*(x) = \int_{-\infty}^{\infty} dx g^*(x) \left(-i\hbar \frac{d}{dx}\right) f(x) \tag{5.2.50}$$

We have arrived at the form of the right-hand side by using integration by parts, and assuming that the functions are vanishing at infinity. By comparing the above, we identify that

$$\hat{p}^\dagger = -i\hbar \frac{d}{dx} \tag{5.2.51}$$

Therefore, we note that $(\hat{p})^\dagger = \hat{p}$ implying that it is Hermitian or self-adjoint. The eigenfunction of the momentum operator is $f_k(x) \propto e^{ikx}$. It is quite clear that $\hat{p}f_k(x) = \hbar k f_k(x)$, and hence, its eigenvalues are also real, as is required of a Hermitian operator.

The above can be generalized to 3D. It can also be shown that the kinetic energy operator

$$\hat{T} = \hat{p}^2/2m = -\frac{\hbar \hat{\nabla}^2}{2m} \tag{5.2.52}$$

is Hermitian. In 3D, $\hat{p}^2$ is proportional to the Laplacian operator ∇^2 in coordinate representation. We can use integration by parts in 3D to show that the above operator is also Hermitian. Using the fact that

$$\nabla \cdot [f(\mathbf{r})\nabla g^*(\mathbf{r})] = f(\mathbf{r})\nabla^2 g^*(\mathbf{r}) + \nabla f(\mathbf{r})\nabla g^*(\mathbf{r}) \tag{5.2.53}$$

---

[10]A note is in order here regarding the term "coordinate space", since there is only one Hilbert space. It is understood that when the term "coordinate space representation" is used, it means that the Hilbert space is represented in the space where the basis is the coordinate space basis. It is often just called "coordinate representation".

we have in coordinate representation

$$\int_V d\mathbf{r} f(\mathbf{r}) \nabla^2 g^*(\mathbf{r}) = \int_V d\mathbf{r} \left[\nabla \cdot (f(\mathbf{r}) \nabla g^*(\mathbf{r})) - \nabla f(\mathbf{r}) \nabla g^*(\mathbf{r}) \right] = - \int_V d\mathbf{r} \nabla f(\mathbf{r}) \nabla g^*(\mathbf{r})$$

(5.2.54)

The last form is symmetric between f and g, and hence, we can easily show that

$$\int_V d\mathbf{r} g^*(\mathbf{r}) \nabla^2 f(\mathbf{r}) = - \int_V d\mathbf{r} \nabla g^*(\mathbf{r}) \nabla f(\mathbf{r})$$

(5.2.55)

Consequently, we can show that

$$\begin{aligned}
\langle f | \hat{\nabla}^2 | g \rangle^* &= \left[\int_V d\mathbf{r} f^*(\mathbf{r}) \nabla^2 g(\mathbf{r}) \right]^* \\
&= \int_V d\mathbf{r} g^*(\mathbf{r}) \nabla^2 f(\mathbf{r}) \\
&= \langle g | \hat{\nabla}^2 | f \rangle
\end{aligned}$$

(5.2.56)

indicating the Hermitian property of the Laplacian operator. Consequently, the $\hat{T}$ operator is also Hermitian or self-adjoint. If we have chosen $f = g$ in the above, (5.2.55) is always negative, implying that the Laplacian operator is a negative definite operator.

The definition of (5.2.40) can be extended to defining the transpose of an operator. For matrix algebra,

$$\left(\mathbf{a}^\dagger \cdot \overline{\mathbf{M}} \cdot \mathbf{b} \right)^t = \mathbf{b}^t \cdot \overline{\mathbf{M}}^t \cdot \mathbf{a}^*$$

(5.2.57)

Generalizing the above to infinite dimensional space, we have

$$\langle f | \hat{M} | g \rangle^t = \langle g^* | \hat{M}^t | f^* \rangle$$

(5.2.58)

which is the defining equation for the transpose of an operator.

5.3 Commutation of Operators

Operators, like matrices, are non-commuting for most parts. In other words,

$$\hat{A}\hat{B} \neq \hat{B}\hat{A}$$

(5.3.1)

The commutator is defined to be

$$\left[\hat{A}, \hat{B} \right] = \hat{A}\hat{B} - \hat{B}\hat{A}$$

(5.3.2)

If $\hat{A}$ and $\hat{B}$ commute, then

$$\left[\hat{A}, \hat{B} \right] = 0$$

(5.3.3)

But for most parts,

$$\left[\hat{A}, \hat{B}\right] = i\hat{C} \tag{5.3.4}$$

It can be shown easily that $\hat{C}$ is a Hermitian operator if $\hat{A}$ and $\hat{B}$ are Hermitian.

If $\hat{A}$ and $\hat{B}$ are commuting, they share the same set of eigenfunctions, but not the same eigenvalues. Say if $|\psi_n\rangle$ is an eigenfunction of $\hat{B}$, with eigenvalue B_n, then

$$\hat{B}\hat{A}|\psi_n\rangle = \hat{A}\hat{B}|\psi_n\rangle = \hat{A}B_n|\psi_n\rangle = B_n\hat{A}|\psi_n\rangle \tag{5.3.5}$$

From the above, it is clear that $\hat{A}|\psi_n\rangle$ is also eigenfunction of $\hat{B}$ with the same eigenvalue B_n. Therefore, $\hat{A}|\psi_n\rangle$ is proportional to $|\psi_n\rangle$, or that $\hat{A}|\psi_n\rangle = A_n|\psi_n\rangle$. In other words, $|\psi_n\rangle$ is also an eigenfunction of $\hat{A}$, but with a different eigenvalue, since $A_n \neq B_n$.

Alternatively, one can argue that if $|\psi_n\rangle$ is an eigenfunction of both $\hat{A}$ and $\hat{B}$, the actions of these operators on their eigenfunctions just produce a scalar number or a c-number (classical number). In other words,

$$\hat{A}\hat{B}|\psi_n\rangle = A_n B_n|\psi_n\rangle = B_n A_n|\psi_n\rangle = \hat{B}\hat{A}|\psi_n\rangle \tag{5.3.6}$$

The above applies also to an arbitrary vector expanded in terms of the eigenfunctions shared by both operators.

As an example, the position operator $\hat{x} = x$ and the momentum operator $\hat{p} = -i\hbar d/dx$ do not commute with each other. It can be easily shown that

$$[x, \hat{p}] = i\hbar\hat{I} = i\hbar$$

Therefore, they cannot share the same set of eigenfunctions.

5.4   Expectation Value and Eigenvalue of Operators

The expectation value of a random variable is a concept from probability. In quantum mechanics, a measurable quantity or an observable in the real world is a random variable. For each observable in the real world, there is a corresponding operator in the quantum world.

The expectation value of an operator for a quantum system in state $|f\rangle$ is defined to be

$$\left\langle \hat{A} \right\rangle = \left\langle f \left| \hat{A} \right| f \right\rangle \tag{5.4.1}$$

It is analogous to

$$\mathbf{f}^\dagger \cdot \overline{\mathbf{A}} \cdot \mathbf{f} \tag{5.4.2}$$

The above are real value quantities if $\hat{A}$ and $\overline{\mathbf{A}}$ are Hermitian operators. It is prudent to understand this expectation value in terms of the eigenstates and eigenvalues of $\hat{A}$. Letting

$$|f\rangle = \sum_n |\psi_n\rangle f_n \tag{5.4.3}$$

then (5.4.1) can be rewritten as

$$\left\langle \hat{A} \right\rangle = \sum_n \sum_m \left\langle \psi_n \left| \hat{A} \right| \psi_m \right\rangle f_n^* f_m \tag{5.4.4}$$

When ψ_m is the eigenstate of $\hat{A}$ with eigenvalue A_m, the above can be written as

$$\left\langle \hat{A} \right\rangle = \sum_n \sum_m A_m \langle \psi_n | \psi_m \rangle f_n^* f_m = \sum_n A_n |f_n|^2 = \sum_n A_n P_n \tag{5.4.5}$$

where $|f_n|^2 = P_n$ is the probability of finding the state in eigenstate n.[11] Hence, the above is the statistical average or the expectation value of the eigenvalue A_n. The eigenvalue A_n can be thought of as a random variable. Hence, the expectation value of an operator is the average value of its eigenvalue when the quantum state is in state f. Therefore, it is customary is denote

$$\left\langle \hat{A} \right\rangle = \left\langle f \left| \hat{A} \right| f \right\rangle = \langle A \rangle = \bar{A} \tag{5.4.6}$$

where the scalar variable A denotes the eigenvalue of $\hat{A}$, which is a random variable, while the angular brackets over A, or $\langle A \rangle$, indicate statistical average. An overbar in the last equality above is often used as a short-hand for angular brackets to denote an average of a random variable.

The above indicates that if we prepare a quantum state that is exactly the eigenstate of the quantum operator $\hat{A}$, then its expectation value in this quantum eigenstate is just the eigenvalue of the eigenstate. In a sense, we can "measure" or "observe" the eigenvalue of the quantum operator through such an experiment. Namely,

$$\left\langle \hat{A} \right\rangle = \left\langle \psi_n \left| \hat{A} \right| \psi_n \right\rangle = A_n \tag{5.4.7}$$

where A_n is the eigenvalue of the eigenstate.

We see in the previous section that when two operators commute, they share the same eigenfunctions but with different eigenvalues. Therefore, if we prepare an eigenstate shared by these two operators, their respective eigenvalues can be "measured" exactly. In other words,

$$\left\langle \hat{A} \right\rangle = \left\langle \psi_n \left| \hat{A} \right| \psi_n \right\rangle = A_n \tag{5.4.8}$$

$$\left\langle \hat{B} \right\rangle = \left\langle \psi_n \left| \hat{B} \right| \psi_n \right\rangle = B_n \tag{5.4.9}$$

On the other hand, if the two operators do not commute, they do not share the same set of eigenstates. If we prepare an eigenstate ψ_n that is the eigenstate of $\hat{A}$, but it is not the eigenstate of $\hat{B}$. However, we can expand the eigenstate ψ_n in terms of the eigenstates of the operator $\hat{B}$; namely,

$$|\psi_n\rangle = \sum_i a_i |\phi_i\rangle \tag{5.4.10}$$

---

[11]Because in (5.4.3), $\langle f|f \rangle = 1$, it implies that $\sum_n |f_n|^2 = 1$. Therefore, $|f_n|^2 = P_n$, the probability of finding the particle in state $|\psi_n\rangle$.

where $|\phi_i\rangle$ are the eigenstates of $\hat{B}$. We now see that the expectation value of $\hat{B}$ due to the eigenstate $|\psi_n\rangle$ is

$$\left\langle \hat{B} \right\rangle = \left\langle \psi_n \left| \hat{B} \right| \psi_n \right\rangle = \sum_{ij} a_i^* a_j \langle \phi_i | \hat{B} | \phi_j \rangle = \sum_i |a_i|^2 B_i \qquad (5.4.11)$$

The expectation value of $\hat{B}$ is due to a range of eigenvalues of $\hat{B}$, and not just due to one pure, single eigenvalue like before. This is the gist of the uncertainty principle. When two operators do not commute, the precise measurement of the eigenvalue of one operator implies that the measurement of the second operator will involve a range of eigenvalues.

For example, the eigenfunction of the momentum operator $\hat{p} = -i\hbar d/dx$ is proportional to e^{ikx} with eigenvalue $\hbar k$, since $\hat{p}e^{ikx} = \hbar k e^{ikx}$. The eigenfunction of the position operator is $\delta(x - x_0)$ since $x\delta(x - x_0) = x_0\delta(x - x_0)$ where x_0 is a constant. Since

$$e^{ikx} = \int\limits_{-\infty}^{\infty} e^{ikx'} \delta(x - x') dx'$$

It is seen that the pure eigenstate of the momentum operator has to be expanded in terms of infinitely many eigenstates of the position operator. Conversely,

$$\delta(x - x') = \frac{1}{2\pi} \int\limits_{-\infty}^{\infty} e^{i(x-x')k} \, dk$$

indicates that the eigenstate of the position operator can be expanded in terms of infinitely many eigenstates of the momentum operator.[12]

5.5 Generalized Uncertainty Principle

We have seen the Heisenberg uncertainty principle expressed for uncertainty in momentum Δp and position Δx as $\Delta p \Delta x \geq \hbar/2$. In this case, p and x are both observables in the real world. In quantum mechanics, all observables are replaced by operators. To connect the operators to real world observables, we take the expectation values of the operators. That is

$$\bar{A} = \langle A \rangle = \langle f | \hat{A} | f \rangle \qquad (5.5.1)$$

where A is a scalar variable representing the eigenvalue of $\hat{A}$ and f is a state vector that defines that state the quantum system is in. The expectation value of an operator also gives the statistical mean of the observable expressed as the mean of the eigenvalue of the corresponding operator. In general,

$$\langle f | (\hat{A})^n | f \rangle = \langle (A)^n \rangle \qquad (5.5.2)$$

$$\langle f | F(\hat{A}) | f \rangle = \langle F(A) \rangle \qquad (5.5.3)$$

---

[12]It is to be noted that the eigenstates e^{ikx} and $\delta(x - x')$ are both not square integrable, and hence, should be thought of as limiting cases of some square integrable functions.

where the left-hand sides are the expectation values of the operators, while the right-hand sides are the expectation values of their eigenvalues.

We can define the deviation from its mean by the operator

$$\Delta\hat{A} = \hat{A} - \bar{A} \tag{5.5.4}$$

On the right-hand side, the first term is an operator while the second term is the mean.[13] The above operator has zero expectation value or zero mean.

We can study the variance of the operator $\left(\Delta\hat{A}\right)^2$ which has nonzero expectation value. Namely,

$$\begin{aligned}
\left\langle f|(\Delta\hat{A})^2|f\right\rangle &= \left\langle f|(\hat{A}-\bar{A})^2|f\right\rangle \\
&= \left\langle (A-\bar{A})^2\right\rangle \\
&= \left\langle A^2 - 2\bar{A}A + \bar{A}^2\right\rangle \\
&= \left\langle A^2\right\rangle - 2\bar{A}\langle A\rangle + \bar{A}^2 \\
&= \left\langle A^2\right\rangle - \bar{A}^2 = \left\langle (\Delta A)^2\right\rangle = \overline{(\Delta A)^2} = \sigma_A^2
\end{aligned} \tag{5.5.5}$$

The above is just the definition of variance of random variable A as in statistics. The standard deviation is obtained by taking the square root of the variance to get[14]

$$\sigma_A = \sqrt{\overline{(\Delta A)^2}} \tag{5.5.6}$$

We can derive a similar expression for σ_B, the standard deviation for B.

The generalized uncertainty principle is obtained by using the Schwartz inequality:

$$\left|\int f^*g\,dx\right|^2 \leq \left(\int |f|^2 dx\right)\left(\int |g|^2 dx\right) \tag{5.5.7}$$

We can rewrite the above using Dirac notation to get

$$|\langle f|g\rangle|^2 \leq \langle f|f\rangle\langle g|g\rangle \tag{5.5.8}$$

The above is the generalization of the cosine inequality we have for 3-vectors[15]

$$|\mathbf{A}\cdot\mathbf{B}|^2 = |\mathbf{A}|^2|\mathbf{B}|^2|\cos\theta|^2 \leq |\mathbf{A}|^2|\mathbf{B}|^2 \tag{5.5.9}$$

It can be generalized to N-vectors or vectors in N dimensional space, and then to vectors in infinite dimensional space if the integrals converge. If we define, for a quantum system in state ψ,

$$|f\rangle = \left(\hat{A} - \bar{A}\right)|\psi\rangle = \hat{a}|\psi\rangle \tag{5.5.10}$$

$$|g\rangle = \left(\hat{B} - \bar{B}\right)|\psi\rangle = \hat{b}|\psi\rangle \tag{5.5.11}$$

---

[13] To be strictly correct, we should multiply the second term by the identity operator, but this is usually understood.

[14] It is tempting to denote the standard deviation as ΔA but this could be confusing in view of (5.5.4).

[15] The integral for the inner products above can be generalized to 3D space.

where $\hat{a} = \hat{A} - \bar{A}$ and $\hat{b} = \hat{B} - \bar{B}$. Then

$$\langle f|f \rangle = \left\langle \psi \left| \hat{a}^2 \right| \psi \right\rangle = \left\langle \psi \left| \left(\hat{A} - \bar{A} \right)^2 \right| \psi \right\rangle = \overline{(\Delta A)^2} \tag{5.5.12}$$

$$\langle g|g \rangle = \left\langle \psi \left| \hat{b}^2 \right| \psi \right\rangle = \left\langle \psi \left| \left(\hat{B} - \bar{B} \right)^2 \right| \psi \right\rangle = \overline{(\Delta B)^2} \tag{5.5.13}$$

Using the inequality in (5.5.8), we have

$$\left(\overline{(\Delta A)^2} \right) \left(\overline{(\Delta B)^2} \right) \geq |\langle \psi|\hat{a}\hat{b}|\psi \rangle|^2 \tag{5.5.14}$$

It can be shown easily that if $[\hat{A}, \hat{B}] = i\hat{C}$

$$[\hat{a}, \hat{b}] = i\hat{C} \tag{5.5.15}$$

where $\hat{C}$ is Hermitian. Furthermore,

$$\hat{a}\hat{b} = \frac{\hat{a}\hat{b} + \hat{b}\hat{a}}{2} + \frac{\hat{a}\hat{b} - \hat{b}\hat{a}}{2} = \frac{\hat{a}\hat{b} + \hat{b}\hat{a}}{2} + i\frac{\hat{C}}{2} \tag{5.5.16}$$

Taking the expectation value of the above, we have

$$\langle \hat{a}\hat{b} \rangle = \left\langle \frac{\hat{a}\hat{b} + \hat{b}\hat{a}}{2} \right\rangle + \frac{i}{2}\langle \hat{C} \rangle \tag{5.5.17}$$

Since $\hat{a}$ and $\hat{b}$ are Hermitian, the operator in first term on the right-hand side is also Hermitian. Hence, its expectation value is purely real, while the second term is purely imaginary. Therefore, the amplitude squared of the above is

$$|\langle \hat{a}\hat{b} \rangle|^2 = \left| \left\langle \frac{\hat{a}\hat{b} + \hat{b}\hat{a}}{2} \right\rangle \right|^2 + \frac{1}{4}|\langle \hat{C} \rangle|^2 \geq \frac{1}{4}|\langle \hat{C} \rangle|^2 \tag{5.5.18}$$

Using the above in (5.5.14), we have

$$\left(\overline{(\Delta A)^2} \right) \left(\overline{(\Delta B)^2} \right) \geq \frac{1}{4}|\langle \hat{C} \rangle|^2 \tag{5.5.19}$$

The above is the generalized uncertainty principle for two observables A and B. We can take the square root of the above to get

$$\sigma_A \sigma_B \geq \frac{1}{2}|\langle \hat{C} \rangle| \tag{5.5.20}$$

Appendix

A   Identity Operator in a Continuum Space

We have seen the derivation of the identity operator when the basis functions are indexed by integers, such as the eigenfunctions of an infinite potential well, or those of a harmonic

oscillator in a parabolic potential well. In many situations, we are required to work with indices that are a continuum (nondenumerable) like the coordinate space indices x, y, z. When an electron is freely roaming, its energy values are also a continuum.

We have seen how we have used denumerable eigenfunctions, such as ψ_n to project a function into its vector representation. Namely,

$$f_n = \langle \psi_n | f \rangle \tag{A1}$$

Very often, a denumerable basis function is just written as $\langle n |$, and the above becomes

$$f_n = \langle n | f \rangle \tag{A2}$$

Similar, we can think of a coordinate (or position) basis function $\langle p_x |$ whose property is that its inner product with $|f\rangle$ yields the value of the function f at position x; namely,

$$f(x) = \langle p_x | f \rangle \tag{A3}$$

The above is often abbreviated as

$$f(x) = \langle x | f \rangle \tag{A4}$$

Assuming that this basis $|x\rangle$ is complete and orthogonal in $0 < x < a$, then we can define an identity operator such that

$$\bar{\mathbf{I}} = \int_0^a dx' |x'\rangle \langle x'| \tag{A5}$$

so that

$$|f\rangle = \int_0^a dx' |x'\rangle \langle x'|f\rangle = \int_0^a dx' |x'\rangle f(x') \tag{A6}$$

Notice that the above is quite different from the identity operator when the basis functions are denumerable, which is

$$\bar{\mathbf{I}} = \sum_{n=1}^{\infty} |f_n\rangle \langle f_n| \tag{A7}$$

Taking the product of (A6) with $\langle x|$, we have

$$\langle x | f \rangle = f(x) = \int_0^a dx' \langle x | x' \rangle f(x') \tag{A8}$$

Hence, in order for (A5) to be an identity operator, the basis function must satisfy

$$\langle x | x' \rangle = \delta(x - x') \tag{A9}$$

Notice that in the denumerable case, the orthonormal relation of the basis function is expressed in terms of Kronecker delta function δ_{ij}, but in the continuum case, the equivalent relationship is expressed in terms of the Dirac delta function $\delta(x - x')$.

We note that $|x\rangle$ is not orthonormal in the strict sense that (A9) is infinite when $x = x'$. The identity operator (A5) is different from (A7) because of the extra weight dx in the integral summation. One may think that $|x\rangle\sqrt{dx'}$ as analogous to the orthonormal vector of the countable case. So (A5) can be thought of as the limiting case of a countable summation

$$\bar{\mathbf{I}} = \sum_{i=1}^{N} \sqrt{\Delta x}|x_i\rangle\langle x_i|\sqrt{\Delta x} \tag{A10}$$

as $\Delta x \to 0$, so the orthonormal basis is actually $\sqrt{\Delta x}|x_i\rangle$. Similar, (A9) as the limiting case of

$$\langle x_i|x_j\rangle = \delta_{ij}/\Delta x \tag{A11}$$

The inner product between two vectors is written as

$$\langle f|g\rangle = \int_0^a dx\langle f|x\rangle\langle x|g\rangle = \int_0^a dx f^*(x)g(x) \tag{A12}$$

where we have inserted the identity operator (A5) in the first expression above to get the second expression, and making use of (A4) to get the third expression. Furthermore, we note that $\langle f|x\rangle$ is the complex conjugate of $\langle x|f\rangle$, since $\langle f|$ and $|x\rangle$ are conjugate transpose of $|f\rangle$ and $\langle x|$, respectively.

A vector $|f\rangle$ can have other representations. For example, we can define a set of "orthonormal" vectors $\langle k|$ such that

$$\langle k|f\rangle = f(k), \tag{A13}$$

where $f(k)$ is the Fourier transform of $f(x)$ via the relationship

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dk e^{ikx} f(k) \tag{A14}$$

The above can be written as

$$f(x) = \langle x|f\rangle = \int_{-\infty}^{\infty} dk\langle x|k\rangle\langle k|f\rangle \tag{A15}$$

where we have defined the identity operator

$$I = \int_{-\infty}^{\infty} dk|k\rangle\langle k| \tag{A16}$$

Comparing (A13), (A14) and (A15), we deduce that

$$\langle x|k\rangle = \frac{1}{\sqrt{2\pi}}e^{ikx} \tag{A17}$$

In other words, the coordinate representation of the vector $|k\rangle$ is given by (A17). Since $\hbar k$ in quantum mechanics is related to the momentum of a particle, $f(k) = \langle k|f\rangle$ is also called the momentum representation of the vector $|f\rangle$, while $f(x) = \langle x|f\rangle$ is the coordinate representation of the vector $|f\rangle$.

From the above, we can derive Parseval's theorem

$$
\begin{aligned}
\langle f|g\rangle &= \int_0^a dx\langle f|x\rangle\langle x|g\rangle = \int_0^a dx f^*(x)g(x) \\
&= \int_{-\infty}^\infty dk\langle f|k\rangle\langle k|g\rangle = \int_{-\infty}^\infty dk f^*(k)g(k)
\end{aligned}
\tag{A18}
$$

Parseval's theorem is the statement that the inner product between two vectors is invariant with respect to the basis that represent it.

The above can be generalized to 3-space where a position basis function is $\langle \mathbf{r}|$, with the property that

$$
f(\mathbf{r}) = \langle \mathbf{r}|f\rangle
\tag{A19}
$$

The identity operator can then be expressed as

$$
\overline{\mathbf{I}} = \int_V d\mathbf{r}'|\mathbf{r}'\rangle\langle\mathbf{r}'|
\tag{A20}
$$

And the "orthonormality" relationship is

$$
\langle \mathbf{r}|\mathbf{r}'\rangle = \delta(\mathbf{r} - \mathbf{r}')
\tag{A21}
$$

Similar basis can be defined for the 3D $\mathbf{k}$ space.

B   Changing Between Representations

If we have an operator equation such as

$$
\hat{p}|\psi\rangle = |g\rangle
\tag{B1}
$$

The above equation is analogous to the equation $\overline{\mathbf{P}} \cdot \mathbf{f} = \mathbf{g}$ in matrix algebra. The explicit forms and the representations of the matrix operator $\overline{\mathbf{P}}$ and the vector $\mathbf{f}$ are not given. In order to express the above explicitly, we first insert the identity operator to transform (B1) to

$$
\int dx\, \hat{p}|x\rangle\langle x|\psi\rangle = |g\rangle
\tag{B2}
$$

Then multiply the above from the left by $\langle x'|$, we have

$$
\int dx\langle x'|\hat{p}|x\rangle\langle x|\psi\rangle = \langle x'|g\rangle
\tag{B3}
$$

$$
\int dx p(x', x)\psi(x) = g(x')
\tag{B4}
$$

The above is the coordinate representation of (B1). It is valid for any operator in a 1D space.

B1 Momentum Operator

But if $\hat{p}$ is the momentum operator, we know that the above equation in coordinate space representation should be

$$-i\hbar\frac{d}{dx}\psi(x) = g(x')$$ (B5)

Hence we conclude that the coordinate representation of $\hat{p}$, which is $\langle x'|\hat{p}|x\rangle$ is

$$p(x', x) = -i\hbar\delta(x' - x)\frac{d}{dx}$$ (B6)

Therefore, a differential operator is a highly localized operator, or a quasi-diagonal operator.

Next, we will study the action of the momentum operator $\hat{p}$ on a momentum eigenstate $|k\rangle$; namely,

$$\hat{p}|k\rangle$$ (B7)

Transforming the above to coordinate representation similar to what have been done above, we have

$$\langle x'|\hat{p}|k\rangle = \int dx \langle x'|\hat{p}|x\rangle\langle x|k\rangle = \int dx\,\delta(x' - x)\left(-i\hbar\frac{d}{dx}\right)\langle x|k\rangle = \left(-i\hbar\frac{d}{dx'}\right)\langle x'|k\rangle$$ (B8)

The expression for $\langle x|k\rangle$ is given in equation (A17). Hence, the above becomes

$$\langle x'|\hat{p}|k\rangle = \hbar k\langle x'|k\rangle$$ (B9)

Therefore, we see that $|k\rangle$ is an eigenvector of the momentum operator $\hat{p}$. In other words,

$$\hat{p}|k\rangle = \hbar k|k\rangle$$ (B10)

In fact, any function of $\hat{p}$, when operating on a momentum eigenstate gives

$$f(\hat{p})|k\rangle = f(\hbar k)|k\rangle$$ (B11)

We can repeat (B2) and (B3) using the basis $|k\rangle$ to arrive at

$$\int dk \langle k'|\hat{p}|k\rangle\langle k|\psi\rangle = \langle k'|g\rangle$$ (B12)

Since $|k\rangle$ is an eigenvector of $\hat{p}$, the above becomes

$$\int dk\,\hbar k\langle k'|k\rangle\langle k|\psi\rangle = \int dk\,\hbar k\,\delta(k' - k)\langle k|\psi\rangle = \hbar k'\langle k'|\psi\rangle = \langle k'|g\rangle$$ (B13)

which is just a scalar equation. In other words, the momentum representation of (B1) is

$$\langle k'|\hat{p}|\psi\rangle = \hbar k'\langle k'|\psi\rangle = \hbar k'\psi(k') = g(k')$$ (B14)

B2 Position Operator

Similarly, we can show that the position operator $\hat{x}$, when operating on a vector, can be expressed as

$$\hat{x}|\psi\rangle = \int dk\,\hat{x}|k\rangle\langle k|\psi\rangle = \int dk \int dx\,\hat{x}|x\rangle\langle x|k\rangle\langle k|\psi\rangle = \int dk \int dx\,x|x\rangle\langle x|k\rangle\langle k|\psi\rangle \qquad \text{(B15)}$$

where we have used $\hat{x}|x\rangle = x|x\rangle$ to arrive at the last equality. Furthermore, we notice that $x|x\rangle\langle x|k\rangle = -\frac{i}{k}\frac{d}{dk}\left(|x\rangle\langle x|k\rangle\right)$ after making use of (A17). Consequently,

$$\hat{x}|\psi\rangle = \int dk \left(-\frac{i}{k}\frac{d}{dk}\int dx\,|x\rangle\langle x|k\rangle\right)\langle k|\psi\rangle = \int dk \left(-\frac{i}{k}\frac{d}{dk}|k\rangle\right)\langle k|\psi\rangle \qquad \text{(B16)}$$

Taking the inner product of the above with $\langle k'|$, we have

$$\langle k'|\hat{x}|\psi\rangle = \int dk \left(-\frac{i}{k}\frac{d}{dk}\langle k'|k\rangle\right)\langle k|\psi\rangle = \int dk \left(-\frac{i}{k}\frac{d}{dk}\delta(k'-k)\right)\psi(k)$$

$$= \int dk\,\delta(k'-k)\left(\frac{i}{k}\frac{d}{dk}\psi(k)\right) \qquad \text{(B17)}$$

The last equality follows from integration by parts. Therefore, the momentum representation of $\hat{x}|\psi\rangle$, which is just $\langle k'|\hat{x}|\psi\rangle$, is

$$\langle k'|\hat{x}|\psi\rangle = \frac{i}{k}\frac{d}{dk'}\psi(k') \qquad \text{(B18)}$$

B3 The Coordinate Basis Function

Using the coordinate basis function $|\mathbf{r}\rangle$, we can write many of our previously derived identities in coordinate representation. Say if the functions defined in (5.1.9) and (5.1.10) are functions of 3-space, indexed by $\mathbf{r}$, then multiplying (5.1.9) from the left by $\langle \mathbf{r}|$,

$$\langle \mathbf{r}|g\rangle = g(\mathbf{r}) = \sum_n \langle \mathbf{r}|\psi_n\rangle\langle\psi_n|g\rangle = \sum_n \psi_n(\mathbf{r})\langle\psi_n|g\rangle \qquad \text{(B19)}$$

We can pre- and post-multiply the identity operator defined in (5.1.10) by $\langle \mathbf{r}|$ and $|\mathbf{r}'\rangle$ respectively, we can identify the operator

$$\left\langle \mathbf{r}\left|\hat{I}\right|\mathbf{r}'\right\rangle = \langle \mathbf{r}|\mathbf{r}'\rangle = \delta(\mathbf{r}-\mathbf{r}') = \sum_n \langle \mathbf{r}|\psi_n\rangle\langle\psi_n|\mathbf{r}'\rangle = \sum_n \psi_n(\mathbf{r})\psi_n^*(\mathbf{r}') \qquad \text{(B20)}$$

where we have noted that $\langle\psi_n|\mathbf{r}'\rangle = \langle\mathbf{r}'|\psi_n\rangle^*$ since they are conjugate transpose of each other. The above is the bilinear eigenfunction expansion of the Dirac delta function. We can also apply the above to the bilinear expansion of an operator given by (5.2.17). Going through similar operations, we have

$$\left\langle \mathbf{r}\left|\hat{A}\right|\mathbf{r}'\right\rangle = A(\mathbf{r},\mathbf{r}') = \sum_n \sum_m A_{nm}\psi_n(\mathbf{r})\psi_m^*(\mathbf{r}') \qquad \text{(B21)}$$

We have let $A(\mathbf{r}, \mathbf{r}')$ be the coordinate representation of the operator $\hat{A}$. The above is just the bilinear expansion in coordinate representation.

We can also choose the momentum basis set $|\mathbf{k}\rangle$ (or any other basis sets) and project the previously obtained identities in momentum representation (or any other representations).

C Time Evolution of the Expectation Value of an Operator

The expectation value of an operator is given by

$$\langle \hat{A} \rangle = \langle \psi(t)|\hat{A}|\psi(t)\rangle \tag{C1}$$

Taking the time derivative of the above yields

$$\partial_t \langle \hat{A} \rangle = \langle \partial_t \psi(t)|\hat{A}|\psi(t)\rangle + \langle \psi(t)|\hat{A}|\partial_t \psi(t)\rangle \tag{C2}$$

Assuming that $\hat{A}$ is time independent, and using the fact that

$$i\hbar \partial_t |\psi(t)\rangle = i\hbar |\partial_t \psi(t)\rangle = \hat{H}|\psi(t)\rangle \tag{C3}$$

we have

$$\begin{aligned}
\partial_t \langle \hat{A} \rangle &= \frac{i}{\hbar} \left\langle \left[\hat{H}\psi(t)\right]|\hat{A}|\psi(t)\right\rangle - \frac{i}{\hbar}\langle \psi(t)|\hat{A}|\hat{H}\psi(t)\rangle \\
&= \frac{i}{\hbar}\langle \psi(t)|\hat{H}\hat{A}|\psi(t)\rangle - \frac{i}{\hbar}\langle \psi(t)|\hat{A}\hat{H}|\psi(t)\rangle \\
&= \frac{i}{\hbar}\langle \psi(t)|\hat{H}\hat{A} - \hat{A}\hat{H}|\psi(t)\rangle
\end{aligned} \tag{C4}$$

or

$$i\hbar \partial_t \langle \hat{A} \rangle = \langle \hat{A}\hat{H} - \hat{H}\hat{A}\rangle = \left\langle \left[\hat{A}, \hat{H}\right]\right\rangle \tag{C5}$$

In other words, if $\hat{A}$ commutes with $\hat{H}$, its expectation value is time independent, or it is a constant of motion. The operator $\hat{A}$ in this case represents an observable that is conserved, such as linear momentum or angular momentum. If such an operator commutes with the Hamiltonian of the quantum system, the observable that corresponds to the operator is conserved.

Furthermore, it can be shown that if $\hat{A}$ is time dependent, we need only to augment the above equation as

$$i\hbar \partial_t \langle \hat{A} \rangle = \left\langle \left[\hat{A}, \hat{H}\right]\right\rangle + i\hbar \langle \partial_t \hat{A}\rangle \tag{C6}$$

C1 Comparison to classical equations of motion

As a simple but illustrative example of the use of equation (C5) we consider the time evolution of the expectation values of the position and momentum operators, with a 1-D Hamiltonian of the form $\hat{H} = \hat{T} + \hat{V} = \frac{\hat{p}^2}{2m} + V(\hat{x})$. Although we do not use the conventional "hat" label for the potential energy operator in the latter form, it is understood as an operator function of the position operator $\hat{x}$. First we consider the time evolution of the expectation of $\hat{x}$.

An operator will commute with any function of itself, therefore the commutator of $\hat{x}$ with $\hat{H}$ reduces to $[\hat{x}, \hat{T}]$.

$$[\hat{x}, \hat{T}] = \frac{1}{2m}[\hat{x}, \hat{p}^2] = \frac{1}{2m}[\hat{x}\,\hat{p}^2 - \hat{p}^2\,\hat{x}] \tag{C7}$$

To evaluate the last commutator between $\hat{x}$ and $\hat{p}^2$ one can imagine an arbitrary state that the commutator acts on and work out everything using chain rule in the position basis. However, this method is cumbersome and unnecessary since the commutator between $\hat{x}$ and $\hat{p}$ is known. The straightforward way to compute is to take the operator $\hat{x}$ across $\hat{p}^2$ using $[\hat{x}, \hat{p}] = i\hbar$

$$\hat{x}\hat{p} = i\hbar + \hat{p}\hat{x} \tag{C8}$$

$$
\begin{aligned}
\hat{x}\hat{p}\hat{p} &= [\hat{p}\hat{x} + i\hbar]\hat{p} \\
&= i\hbar\hat{p} + \hat{p}[\hat{p}\hat{x} + i\hbar] \\
&= 2i\hbar\hat{p} + \hat{p}\hat{p}\hat{x} \\
\hat{x}\hat{p}^2 - \hat{p}^2\hat{x} &= 2i\hbar\hat{p}
\end{aligned}
\tag{C9}
$$

We therefore conclude that the time evolution of the expectation value of $\hat{x}$ is governed by:

$$\partial_t \langle \hat{x} \rangle = \frac{1}{m}\langle \hat{p} \rangle \tag{C10}$$

In the sense of expectation values this equation agrees with its classical counterpart, which is Newton's first law. A more interesting case is the time evolution of the expectation value of the momentum. We would expect to get a quantum mechanical analog of Newton's second law. But the calculation shows otherwise.

Now the commutator of interest is that between the momentum operator $\hat{p}$ and the potential energy operator $V(\hat{x})$. Since no explicit expression is given for the functional form of $V(\hat{x})$ it is not immediately clear how this can be evaluated. However, recall our result in equation (C9). Let us recast the formula into a different form. An interesting observation is that

$$[\hat{x}, \hat{p}^2] = 2i\hbar\hat{p} = [\hat{x}, \hat{p}]\,\partial_{\hat{p}}\hat{p}^2 \tag{C11}$$

In the right most part of the above equation we take the derivative of an operator function. This may be understood by imagining the whole structure operating on an eigenfunction of $\hat{p}$ and thus reduces to taking the derivative of an ordinary function. This suggest us to consider

the following commutator

$$
\begin{aligned}
[\hat{p}, \hat{x}^n] &= \hat{p}\hat{x}^n - \hat{x}^n\hat{p} \\
&= [\hat{x}\hat{p} - i\hbar]\hat{x}^{n-1} - \hat{x}^n\hat{p} \\
&= \ldots \\
&= -ni\hbar\hat{x}^{n-1} + \hat{x}^n\hat{p} - \hat{x}^n\hat{p} \\
[\hat{p}, \hat{x}^n] &= [\hat{p}, \hat{x}]\,\partial_{\hat{x}}\hat{x}^n
\end{aligned}
\tag{C12}
$$

Equation (C12) is true for any polynomial function of $\hat{x}$. Given a nice and smooth potential function $V(\hat{x})$ we can use Taylor expansion to turn it into a polynomial function. Then the commutator between $\hat{p}$ and $V(\hat{x})$ can be computed as

$$
[\hat{p}, V(\hat{x})] = -i\hbar\,\partial_{\hat{x}}V(\hat{x})
\tag{C13}
$$

The time evolution of the expectation value of the momentum operator is thus governed by

$$
\partial_t\langle\hat{p}\rangle = -\langle\partial_{\hat{x}}V(\hat{x})\rangle
\tag{C14}
$$

We can now take the obvious choice of defining the "force operator" in analogy with the classical force $F(x) = -\partial_x V(x)$. The "force operator" is again understood to be an operator function of $\hat{x}$. Thus we can write the quantum mechanical equation of motion

$$
\partial_t\langle\hat{p}\rangle = \langle F(\hat{x})\rangle
\tag{C15}
$$

At first sight one might say this is just Newton's second law. However, due to the expectation value being taken on the operator function $F(\hat{x})$ the time evolution in quantum mechanics is in general different from that for a classical particle following the classical path defined by $\langle\hat{x}\rangle$, i.e. $\langle F(\hat{x})\rangle \neq F(\langle\hat{x}\rangle)$.

It is curious to consider when quantum mechanics agrees with classical mechanics in the expectation value sense. For this purpose we can expand the "force operator" around the expectation value of $\hat{x}$. We denote this expectation value by $\bar{x}$

$$
F(\hat{x}) = F(\bar{x}) + \partial_{\hat{x}}F(\hat{x})\Big|_{\bar{x}}(\hat{x} - \bar{x}) + \frac{1}{2!}\partial_{\hat{x}}^2 F(\hat{x})\Big|_{\bar{x}}(\hat{x} - \bar{x})^2 + H.O.T
\tag{C16}
$$

When the expectation value is taken, it is clear that the 0-th order and 1-st order terms do not cause any difference between $\langle F(\hat{x})\rangle$ and $F(\bar{x})$. The higher order terms are responsible for the deviation. If the derivatives are vanishing above the 1-st order then the quantum mechanical equations of motion for the expectation values of position and momentum are in complete agreement with their classical counterparts. This corresponds to the case of no force, a constant force or a linear force.

In terms of potential problems, which is the more natural language for quantum mechanics, these are the free particle, the constant slope and the quadratic potential. This last potential will be studied in great detail in later chapters. We will make extensive connections with classical physics using the quadratic potential, the celebrated potential of the harmonic oscillator. These connections between quantum mechanics and classical mechanics are possible due to the linear nature of the force.

D   Periodic Boundary Condition

One technique often used to create countably infinite modes is to use periodic boundary condition. In this method, we require that the wavefunction satisfies the boundary condition that $\psi(x = 0) = \psi(x = L)$. This can be satisfied easily by a traveling wave

$$\psi(x) = Ce^{ikx} \tag{D1}$$

The periodic boundary condition implies that

$$e^{ikL} = 1 \tag{D2}$$

or that

$$k = k_n = \frac{2n\pi}{L} \tag{D3}$$

where we have indexed k with subscript n, and n is all integer values on the real line from $-\infty$ to ∞. Furthermore, we can pick $C = 1/\sqrt{L}$ to normalize the above function. It is quite easy to show that with

$$\psi_n(x) = \sqrt{\frac{1}{L}} e^{i2n\pi x/L} \tag{D4}$$

then

$$\langle \psi_n | \psi_m \rangle = \delta_{nm} \tag{D5}$$

Hence, given any function $f(x)$ defined between $0 < x < L$, we can expand

$$f(x) = \sum_{n=-\infty}^{\infty} f_n \psi_n(x) = \sum_{n=-\infty}^{\infty} f_n \sqrt{\frac{1}{L}} e^{i2n\pi x/L} \tag{D6}$$

The above is just a Fourier series expansion of the function $f(x)$. The Fourier coefficient f_n can be found by using the orthonormality of ψ_n, to get

$$f_n = \langle \psi_n | f \rangle = \sqrt{\frac{1}{L}} \int_0^L dx e^{-i2n\pi x/L} f(x) \tag{D7}$$

In the above, since $\psi_n(x)$ is a periodic function with period L, $f(x)$ is also a periodic function with period L.

Fourier transform can be derived from Fourier series expansion. To this end, we look at (D3), and note that the k_n values are equally spaced by $2\pi/L$ on the real k line. The spacing becomes increasingly small as $L \to \infty$. We can define the spacing to be $\Delta k = 2\pi/L$. Consequently, in view of changing a summation to an integral, we can rewrite (D6) more suggestively as

$$f(x) = \sum_{n=-\infty}^{\infty} \Delta k f_n \frac{L}{2\pi} \sqrt{\frac{1}{L}} e^{i2n\pi x/L} = \frac{1}{2\pi} \sum_{n=-\infty}^{\infty} \Delta k \tilde{f}(k_n) e^{ik_n x} \tag{D8}$$

where we have defined $\tilde{f}(k_n) = f_n\sqrt{L}$. From (D7), we have

$$\tilde{f}(k_n) = \sqrt{L}f_n = \int_0^L dx e^{-ik_n x} f(x) \tag{D9}$$

In the limit when L tends to infinity, $\Delta k \to 0$, the sum in (D8) becomes an integral; namely

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dk \tilde{f}(k) e^{ikx} \tag{D10}$$

The above is just a Fourier inverse transform integral with $\tilde{f}(k)$ from (D9)

$$\tilde{f}(k) = \int_0^{\infty} dx e^{-ikx} f(x) \tag{D11}$$

which is a Fourier transform. Here, $f(x)$ is originally a periodic function, but now the period L is infinite. Also, if $f(x) \to 0$ when $x \to \infty$, then by its infinite periodicity, $f(x) \to 0$ when $x \to -\infty$. We can replace the semi-infinite integral above with an infinite integral:

$$\tilde{f}(k) = \int_{-\infty}^{\infty} dx e^{-ikx} f(x) \tag{D12}$$

The above, (D10) and (D12), form a Fourier transform pair.

Chapter 6

Approximate Methods in Quantum Mechanics

6.1   Introduction

There are many problems in quantum mechanics where closed form or simple solutions cannot be found. What we have shown as examples are usually problems that have simple solutions, such a infinite potential well, harmonic oscillator, or a finite potential well. Even the case of finite potential well requires the solution of a transcendental equation. Nevertheless, these are called textbook problems, because they are wonderful examples in textbooks to teach beginning students on the subject matter. But in general, most quantum mechanics problems do not have closed form solution. We have to resort to approximate or numerical methods to solve such problems.

6.2   Use of an Approximate Subspace

We have seen that in general, a Schrödinger equation problem can be cast into a matrix equation by projecting it into a space spanned by countably infinite orthonormal basis functions. With such a basis, we define an identity operator

$$\hat{I} = \sum_{n=1}^{\infty} |\psi_n\rangle\langle\psi_n| \tag{6.2.1}$$

Given a general time independent Schrödinger equation, which is

$$\hat{H}|\psi\rangle = E|\psi\rangle \tag{6.2.2}$$

we can first insert an identity operator between $\hat{H}$ and $|\psi\rangle$. Then the above becomes

$$\sum_{n=1}^{\infty} \hat{H}|\psi_n\rangle\langle\psi_n|\psi\rangle = E|\psi\rangle \tag{6.2.3}$$

Testing the above by multiplying it from the left with $\langle\psi_m|$, $m = 1, \ldots, \infty$, we have[1]

$$\sum_{n=1}^{\infty} \langle\psi_m|\hat{H}|\psi_n\rangle\langle\psi_n|\psi\rangle = E\langle\psi_m|\psi\rangle, \; m = 1, \ldots, \infty \qquad (6.2.4)$$

The above is a matrix equation of the form

$$\sum_{n=1}^{\infty} H_{mn}f_n = Ef_m, \; m = 1, \ldots, \infty \qquad (6.2.5)$$

It corresponds to an infinite dimension matrix, but it can be truncated to form a finite matrix system

$$\overline{\mathbf{H}} \cdot \mathbf{f} = E\mathbf{f} \qquad (6.2.6)$$

The above is a matrix eigenvalue problem that can be solved numerically. If $\overline{\mathbf{H}}$ is $N \times N$, it will have N eigenvalues and N eigenvectors. If we choose our basis functions appropriately, and are only interested in certain eigenvalues, which represent the energy levels of the stationary states, we need only to pick a small set of basis functions that approximate these stationary states well. Therefore, N can be a small number. This is important because the number of computer operations needed to find the eigenvectors and eigenvalues of the above problem is proportional to N^3.

In solving the above, we will have to form the matrix elements H_{mn}. For an infinite potential well, whose well bottom has been distorted, we can use the eigenfunctions of the flat-bottom well as our basis functions. Then, the matrix elements H_{mn} are explicitly given as

$$H_{mn} = \int_0^L dx\,\psi_m^*(x)\left(-\frac{\hbar^2}{2m_o}\frac{d^2}{dx^2} + V(x)\right)\psi_n(x) \qquad (6.2.7)$$

We can use integration by parts to convert the term that involves second derivative to a form that involves only first derivatives. That is we let

$$\psi_m^*(x)\frac{d^2}{dx^2}\psi_n(x) = \frac{d}{dx}\left[\psi_m^*(x)\frac{d}{dx}\psi_n(x)\right] - \frac{d}{dx}\psi_m^*(x)\frac{d}{dx}\psi_n(x)$$

in the above to arrive at

$$H_{mn} = -\frac{\hbar^2}{2m_o}\psi_m^*(x)\frac{d}{dx}\psi_n(x)\bigg|_0^L + \int_0^L dx\left(\frac{\hbar^2}{2m_o}\frac{d}{dx}\psi_m^*(x)\frac{d}{dx}\psi_n(x) + \psi_m^*(x)V(x)\psi_n(x)\right)$$
$$(6.2.8)$$

The first term on the right-hand side can be made to vanish by the virtue of the boundary condition, leaving the second term, which is a more symmetric form.

---

[1]This procedure is called testing or weighting in the mathematics literature, $\langle\psi_m|$ is called the testing or weighting function.

Modern numerical methods have allowed us to use flexible basis set to expand our unknown functions. For example, we can let

$$|\psi\rangle = \sum_{n=1}^{N} a_n |f_n\rangle \tag{6.2.9}$$

where $f_n(x)$ need not be orthogonal or complete in Hilbert space, but they can approximate the unknown function $\psi(x)$ well. Also, $f_n(x), n = 1, \ldots, N$ spans an N dimensional vector space that is the subspace of the Hilbert space. We assume that the solution in the subspace is a good approximation of the solution in the original infinite dimensional Hilbert space. Examples of such finite basis set are triangle functions shown in Figure 6.1.

Using (6.2.9) in (6.2.2), one gets

$$\sum_{n=1}^{N} \hat{H}|f_n\rangle a_n = E \sum_{n=1}^{N} a_n |f_n\rangle \tag{6.2.10}$$

Multiplying the above from $\langle f_m|$ from the left, we have

$$\sum_{n=1}^{N} \langle f_m|\hat{H}|f_n\rangle a_n = E \sum_{n=1}^{N} \langle f_m|f_n\rangle a_n \tag{6.2.11}$$

The above is a matrix system of the form

$$\tilde{\overline{\mathbf{H}}} \cdot \mathbf{a} = E\overline{\mathbf{B}} \cdot \mathbf{a} \tag{6.2.12}$$

In the above,

$$\left[\tilde{\overline{\mathbf{H}}}\right]_{mn} = \langle f_m|\hat{H}|f_n\rangle$$
$$\left[\overline{\mathbf{B}}\right]_{mn} = \langle f_m|f_n\rangle$$
$$[\mathbf{a}]_n = a_n \tag{6.2.13}$$

The equation (6.2.12) above is a generalized eigenvalue problem with eigenvalue E and eigenvector $\mathbf{a}$. The difference of it from (6.2.6) is that a matrix $\overline{\mathbf{B}}$ is on the right-hand side.[2] But numerical software to seek the eigenvalues and eigenvectors for such generalized eigenvalue problems are widely available, e.g., in MATLAB.

The approximate subspace method is varyingly known as the subspace projection method, Galerkin's method (named after Boris Galerkin, 1871-1945), Petrov-Galerkin method, weighted-residual method, and the method of moments. The finite element method falls under the same category.

6.3 Tight Binding Model

The tight binding model is a way to find the solution to two weakly coupled quantum systems when the solution to each isolated quantum system is known (see Figure 6.1).

---

[2]The matrix $\overline{\mathbf{B}}$ is also known as the Gram matrix.
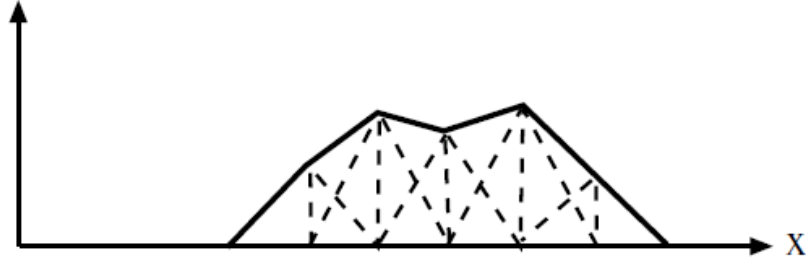
Figure 6.1: The triangle functions for a piecewise linear approximation of a function. This is a basis that is not orthogonal but yet can be used to seek approximate solutions to (6.2.12).

To solve this system, we can use the subspace projection method of nonorthogonal basis to derive the relevant matrix equation. Given the Schrödinger equation

$$\hat{H} \left| \psi \right\rangle = E \left| \psi \right\rangle \tag{6.3.1}$$

we let

$$\left| \psi \right\rangle = \sum_{i=1}^{2} a_i \left| \psi_i \right\rangle \tag{6.3.2}$$

where $\left| \psi_1 \right\rangle$ is the eigenfunction of the ground state of well 1 in isolation and $\left| \psi_2 \right\rangle$ is the eigenfunction of the ground state of well 2. Going through the same procedure we have had before, we arrive at

$$\sum_{i=1}^{2} a_i \left\langle \psi_j \right| \hat{H} \left| \psi_i \right\rangle = E \sum_{i=1}^{2} a_i \left\langle \psi_j \right| \psi_i \right\rangle \tag{6.3.3}$$

The above is the same as the matrix equation

$$\overline{\mathbf{H}} \cdot \mathbf{a} = E \overline{\mathbf{B}} \cdot \mathbf{a} \tag{6.3.4}$$

where

$$\left[\overline{\mathbf{H}} \right]_{ij} = \left\langle \psi_i \right| \hat{H} \left| \psi_j \right\rangle \tag{6.3.5}$$

$$\left[\overline{\mathbf{B}} \right]_{ij} = \left\langle \psi_i \right| \psi_j \right\rangle \tag{6.3.6}$$

It is easy to show that $\overline{\mathbf{H}}$ and $\overline{\mathbf{B}}$ are Hermitian matrices.[3] In the above, we can assume that $\left| \psi_i \right\rangle$ is quasi-orthonormal, so that

$$\left\langle \psi_i \right| \psi_j \right\rangle \simeq \delta_{ij} \tag{6.3.7}$$

---

[3]It can be easily shown that the matrix representation of a Hermitian operator remains Hermitian.
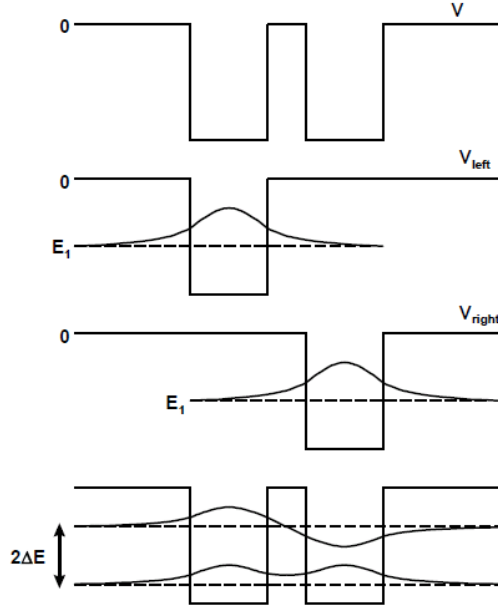
Figure 6.1: The tight binding model can be used to find the approximate eigenstates of two quantum wells that are weakly coupled to each other. The eigenstates of the isolated quantum wells are known (Figure is from DAB Miller).

In other words, the $\overline{\mathbf{B}}$ matrix is unity on the diagonal, and small on the off diagonal terms. Then we can write the system (6.3.4) as

$$\begin{bmatrix} E_1 & \Delta E \\ \Delta E^* & E_1 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = E \begin{bmatrix} 1 & \Delta B \\ \Delta B^* & 1 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \tag{6.3.8}$$

The off diagonal elements are conjugate of each other because the matrix is Hermitian. Moreover, $\Delta E \ll E_1$ and $\Delta B \ll 1$. These terms are small because the overlap between the eigenfunctions ψ_1 and ψ_2 is small. The further apart the wells are, the smaller these terms would be, as the overlap is smaller. When the wells are infinitely far apart, the off diagonal terms vanish and the two quantum systems are uncoupled from each other. The two eigenvalues are then degenerate. To begin, we can assume weak coupling between the modes in the wells.

The eigenvalue of (6.3.8) above is obtained by solving

$$\det \begin{bmatrix} E_1 - E & \Delta E - E\Delta B \\ \Delta E^* - E\Delta B^* & E_1 - E \end{bmatrix} = 0 \tag{6.3.9}$$

or

$$(E_1 - E)^2 - |\Delta E - E\Delta B|^2 = 0 \tag{6.3.10}$$

Upon solving the above, we have

$$E = E_1 \pm |\Delta E_1 - E\Delta B| \cong E_1 \pm |\Delta E_1 - E_1\Delta B| \qquad (6.3.11)$$

Since the $E\Delta B$ on the right-hand side of the first equality above is small, we do not incur much error by replacing it with $E_1\Delta B$. Another way of saying it is that $E \approx E_1$ is correct to leading order (or zeroth order), and $E\Delta B \approx E_1\Delta B$ is correct to second order.

Also, it can be shown that stationary states can be made completely real valued (or equal phase). Hence, ΔE and ΔB can be made to be real valued. Therefore, without the lost of generality, the above can be replaced with

$$E \cong E_1 \pm (\Delta E_1 - E_1\Delta B) \qquad (6.3.12)$$

Note that the degenerate eigenvalues are now split into two non-degenerate eigenvalues.

The equation to be satisfied by the eigenvectors is

$$\begin{bmatrix} \pm(\Delta E - E_1\Delta B) & \Delta E - E_1\Delta B \\ \Delta E - E_1\Delta B & \pm(\Delta E - E_1\Delta B) \end{bmatrix} \cdot \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = 0 \qquad (6.3.13)$$

The above equation is correct to first order, as the errors are second order. On solving, and normalizing,[4] we obtain that

$$\begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \qquad (6.3.14)$$

or

$$\begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix} \qquad (6.3.15)$$

The corresponding eigenvectors can be found, and the eigenfunctions are of the form

$$\psi_+(x) = \frac{1}{\sqrt{2}} \left(\psi_1(x) + \psi_2(x)\right), \qquad \psi_-(x) = \frac{1}{\sqrt{2}} \left(\psi_1(x) - \psi_2(x)\right) \qquad (6.3.16)$$

The eigenfunctions are due to even and odd coupling between the isolated quantum states of the two wells as they are brought closer together.

6.3.1   Variational Method

Given an eigenvalue problem denoted by

$$\hat{H} |\phi\rangle = E |\phi\rangle \qquad (6.3.17)$$

we can convert the above into a matrix generalized eigenvalue problem

$$\overline{\mathbf{H}} \cdot \mathbf{a} = E \, \overline{\mathbf{B}} \cdot \mathbf{a} \qquad (6.3.18)$$

---

[4]The normalization of the eigenfunction in (6.3.2) implies that $|a_1|^2 + |a_2|^2 = 1$.

We next convert the above into a functional by multiplying it with $\bar{\mathbf{a}}^\dagger$ to obtain

$$\mathbf{a}^\dagger \cdot \overline{\mathbf{H}} \cdot \mathbf{a} = E\, \mathbf{a}^\dagger \cdot \overline{\mathbf{B}} \cdot \mathbf{a} \qquad (6.3.19)$$

By setting $\mathbf{a} = \mathbf{a}_0 + \delta\mathbf{a}$, $E = E_0 + \delta E$, where $\mathbf{a}_0$ and E_0 are the exact solution to (6.3.18), and then taking the first variation of (6.3.19) we have

$$\delta\mathbf{a}^\dagger \cdot \overline{\mathbf{H}} \cdot \mathbf{a}_0 + \mathbf{a}_0^\dagger \cdot \overline{\mathbf{H}} \cdot \delta\mathbf{a} = \delta E\, \mathbf{a}_0^\dagger \cdot \overline{\mathbf{B}} \cdot \mathbf{a}_0 + E_0 \cdot \delta\mathbf{a}^\dagger \cdot \overline{\mathbf{B}} \cdot \mathbf{a}_0 \\ + E_0 \mathbf{a}_0^\dagger \cdot \overline{\mathbf{B}} \cdot \delta\mathbf{a} \qquad (6.3.20)$$

If $\mathbf{a}_0$ and E_0 are exact solution to (6.3.20), they satisfy

$$\overline{\mathbf{H}} \cdot \mathbf{a}_0 = E_0\, \overline{\mathbf{B}} \cdot \mathbf{a}_0 \qquad (6.3.21)$$

$$\mathbf{a}_0^\dagger \cdot \overline{\mathbf{H}} = E_0 \mathbf{a}_0^\dagger \cdot \overline{\mathbf{B}}_0 \qquad (6.3.22)$$

Then the term multiplied by $\delta\mathbf{a}^\dagger$ and $\delta\mathbf{a}$ cancel each other on both sides and we have, to first order,

$$0 = \delta E\, \mathbf{a}_0^\dagger \cdot \overline{\mathbf{B}} \cdot \mathbf{a}_0 \qquad (6.3.23)$$

The above implies that δE is zero. It means that if we were to use (6.3.19) to find E by writing it as

$$E = \frac{\mathbf{a}^\dagger \cdot \overline{\mathbf{H}} \cdot \mathbf{a}}{\mathbf{a}^\dagger \cdot \overline{\mathbf{B}} \cdot \mathbf{a}} \qquad (6.3.24)$$

and if we substitute inexact or approximate value for $\mathbf{a} = \mathbf{a}_0 + \delta\mathbf{a}$, the value we obtain for E is second order accurate since $\delta E = 0$. The above is known as the Rayleigh quotient.[5] To begin with, we need an estimate of the solution that is quite close to the exact solution so that $\delta\mathbf{a}$ is small. In other words, it allows one to substitute in approximate value of $\mathbf{a}$ that is just first-order accurate, and yet get an estimate of the eigenvalue or energy E that is second-order accurate.

A more general form of the Rayleigh quotient is to test (6.3.17) with $\langle\phi|$ and convert it into a functional. The corresponding Rayleigh quotient is

$$E = \frac{\langle\phi|\hat{H}|\phi\rangle}{\langle\phi|\phi\rangle} \qquad (6.3.25)$$

It assumes that $|\phi\rangle$ is not normalized. In the above, it can be proved by the same token that an approximate, first-order accurate value of $|\phi\rangle$ can yield a second-order accurate estimate of E. Rayleigh quotient is great for estimating the eigenvalues when only approximate eigenfunctions are known. These eigenfunctions are also called trial functions. A systematic way to estimate these eigenfunctions is known as the Rayleigh-Ritz method.

---

[5] This method is also known as Rayleigh-Ritz method, due to Lord Rayleigh (John W. Strutt, 1842-1919) and Walther Ritz (1878-1909).

6.4   Time Dependent Perturbation Theory

In time dependent perturbation theory, the Hamiltonian is perturbed with a perturbing Hamiltonian that is a function of time. This is contradictory to the requirement that Hamiltonian is a independent of time, since it represents total energy. However, we can think of the Hamiltonian of a closed quantum system that is energy conserving as consisting of three parts;

$$\hat{H} = \hat{H}_A + \hat{H}_I + \hat{H}_B \tag{6.4.1}$$

The above consists of quantum systems described by $\hat{H}_A$ and $\hat{H}_B$, and $\hat{H}_I$ accounts for the interaction between the two systems. Then each of the constituent Hamiltonian above can be time varying even though the total Hamiltonian is not. Assuming the interaction is weak, then we can focus on system A, and see what the effect of the interaction Hamiltonian H_I is on system A. To achieve energy conservation, one needs for ascertain the back-action of A on B, but this is assumed to be small.

Time varying Hamiltonian due to a small perturbation from the interaction Hamiltonian will be the spirit of the time dependent perturbation theory. The time dependent perturbation theory is important for studying field-atom interaction. When light impinges on an atom, it can cause the electron in a stationary state in the atom to be excited to a higher level by absorbing a photon, or that the electron drops to a lower level by emitting a photon (radiative transition).

To this end, we assume that the perturbing Hamiltonian is a function of time. In other words, we have a Schrödinger equation where

$$i\hbar \frac{\partial}{\partial t} |\Psi\rangle = \hat{H} |\Psi\rangle \tag{6.4.2}$$

and

$$\hat{H} = \hat{H}_0 + \gamma \hat{H}_p(t) \tag{6.4.3}$$

where $\hat{H}_p(t)$ is time varying.[6] It can come from an external electric field that is time-harmonic, for instance. We also assume that we can solve the unperturbed system in terms of stationary states, namely

$$\hat{H}_0 |\Psi_n\rangle = E_n |\Psi_n\rangle \tag{6.4.4}$$

The total time dependent wavefunction can be expanded in terms of the stationary states, namely

$$|\Psi\rangle = \sum_n a_n(t) e^{-i\omega_n t} |\Psi_n\rangle \tag{6.4.5}$$

where $\hbar\omega_n = E_n$, and $a_n(t)$ is a function of time.

---

[6]To be precise, the Hamiltonian of a system has to be a constant of motion, since it represents the conservation of energy. But when a quantum system is coupled to another quantum system to which energy is transferred, the first Hamiltonian may not be a constant anymore. But we can assume that the perturbing Hamiltonian is small.

Substituting the above into (6.4.2), we have

$$\frac{\partial}{\partial t}\ket{\Psi} = \sum_n \left(\dot{a}_n(t) - i\omega_n\right) e^{-i\omega_n t}\ket{\Psi_n} = \frac{1}{i\hbar}\hat{H}\sum_n a_n(t) e^{-i\omega_n t}\ket{\Psi_n} \tag{6.4.6}$$

$$= \frac{1}{i\hbar}\sum_n a_n(t) e^{-i\omega_n t}\left(E_n\ket{\Psi_n} + \gamma\hat{H}_p(t)\ket{\Psi_n}\right) \tag{6.4.7}$$

where $\dot{a}_n(t) = da_n(t)/dt$.

Since $E_n = \hbar\omega_n$, the E_n term on the right-hand side cancels the ω_n term on the left-hand side. Finally we have

$$\sum_n \dot{a}_n(t) e^{-i\omega_n t}\ket{\Psi_n} = \frac{\gamma}{i\hbar}\sum_n a_n(t) e^{-i\omega_n t}\hat{H}_p(t)\ket{\Psi_n} \tag{6.4.8}$$

To simplify the equation further, we test it with $\bra{\Psi_q}$ to get

$$\dot{a}_q(t) e^{-i\omega_q t} = \frac{\gamma}{i\hbar}\sum_n a_n(t) e^{-i\omega_n t}\bra{\Psi_q}\hat{H}_p(t)\ket{\Psi_n} \tag{6.4.9}$$

In the above, $\bra{\Psi_q}\hat{H}_p(t)\ket{\Psi_n}$ is the q, n element of the matrix representation of the Hamiltonian $\hat{H}_p$ in the basis formed by Ψ_n.

The above equation is exact as of this point. To solve it further, we expand

$$a_n = a_n^{(0)} + \gamma a_n^{(1)} + \gamma^2 a_n^{(2)}\cdots \tag{6.4.10}$$

Using the above in (6.4.9), and matching terms of like orders, we have

$$\dot{a}_q^{(0)}(t) = 0, \quad \forall q \tag{6.4.11}$$

or that $a_q^{(0)}$ is a constant. This means that if $\gamma = 0$, or there is no perturbation, then the only terms that remain are the leading order terms. Since these are stationary states, the quantum system remains in these stationary states if perturbation is turned off.

By matching terms of the first order, we have

$$\dot{a}_q^{(1)}(t) = \frac{1}{i\hbar}\sum_n a_n^{(0)} e^{i\omega_{qn} t}\bra{\Psi_q}\hat{H}_p(t)\ket{\Psi_n} \tag{6.4.12}$$

where $\omega_{qn} = \omega_q - \omega_n$. To make things simpler, we can assume that the electron is in the m-th stationary state before we turn on the perturbation $\hat{H}_p(t)$. Namely,

$$\ket{\Psi} = a_m^{(0)} e^{-i\omega_m t}\ket{\Psi_m} = e^{-i\omega_m t}\ket{\Psi_m} \tag{6.4.13}$$

In the above, since if only one state exists, $a_m^{(0)} = 1$ by normalization. Then, (6.4.12) becomes

$$\dot{a}_q^{(1)}(t) = \frac{1}{i\hbar}a_m^{(0)} e^{i\omega_{qm} t}\bra{\Psi_q}\hat{H}_p(t)\ket{\Psi_m} \tag{6.4.14}$$

We assume the perturbing Hamiltonian to be of the form:

$$\hat{H}_p(t) = \begin{cases} 0, & t < 0 \\ \hat{H}_{p0}(e^{-i\omega t} + e^{i\omega t}), & 0 < t < t_0 \\ 0, & t > t_0 \end{cases} \tag{6.4.15}$$

so that it is turned on only for a time window $0 < t < t_0$. Then for $t > t_0$, by integrating (6.4.14), we have

$$a_q^{(1)} = \frac{1}{i\hbar} \int_0^{t_0} \langle \Psi_q | \hat{H}_p(t') | \Psi_m \rangle e^{i\omega_{qm}t'} dt' \tag{6.4.16}$$

$$= \frac{1}{i\hbar} \langle \Psi_q | \hat{H}_{p0} | \Psi_m \rangle \int_0^{t_0} \left(e^{i(\omega_{qm}-\omega)t'} + e^{i(\omega_{qm}+\omega)t'} \right) dt' \tag{6.4.17}$$

$$= -\frac{1}{i\hbar} \langle \Psi_q | \hat{H}_{p0} | \Psi_m \rangle \left(\frac{e^{i(\omega_{qm}-\omega)t_0} - 1}{\omega_{qm} - \omega} + \frac{e^{i(\omega_{qm}+\omega)t_0} - 1}{\omega_{qm} + \omega} \right) \tag{6.4.18}$$

$$= -\frac{t_0}{i\hbar} \langle \Psi_q | \hat{H}_{p0} | \Psi_m \rangle \left[e^{\frac{i(\omega_{qm}-\omega)t_0}{2}} \operatorname{sinc}\left(\frac{(\omega_{qm}-\omega)t_0}{2} \right) + e^{\frac{i(\omega_{qm}+\omega)t_0}{2}} \operatorname{sinc}\left(\frac{(\omega_{qm}+\omega)t_0}{2} \right) \right] \tag{6.4.19}$$

The probability of finding the state in q-state is

$$P(q) = |a_q^{(1)}|^2 \tag{6.4.20}$$

or

$$P(q) \approx \frac{t_0^2}{\hbar^2} |\langle \Psi_q | \hat{H}_{p0} | \Psi_m \rangle|^2 \left[\operatorname{sinc}^2\left(\frac{(\omega_{qm}-\omega)t_0}{2} \right) + \operatorname{sinc}^2\left(\frac{(\omega_{qm}+\omega)t_0}{2} \right) \right] \tag{6.4.21}$$

where we have ignored the cross terms since the sinc functions in (6.4.19) are highly peaked at $\omega = \pm\omega_{qm}$ when $t_0 \to \infty$; and hence their cross terms are small.

The above equation means that the probability of transition from the m to the q eigenstate is high only if

$$\omega = \pm\omega_{qm} = \pm(\omega_q - \omega_m) \tag{6.4.22}$$

or that

$$\hbar\omega = \pm(\hbar\omega_q - \hbar\omega_m) = \pm(E_q - E_m) \tag{6.4.23}$$

The "+" sign corresponds to when the electron jumps from a low energy E_m to high energy E_q requiring the absorption of a photon with energy $\hbar\omega = (E_q - E_m)$. The "−" sign in (6.4.23) corresponds to the electron dropping from a high energy state E_m to a low energy state E_q emitting a photon with energy $\hbar\omega = (E_m - E_q)$. This is called stimulated emission.[7] The rule of energy level matching is known as the Fermi's golden rule (named after Enrico Fermi, 1901-1954). Moreover, the transition is possible only if the value of the matrix element $|\langle \Psi_q | \hat{H}_{p0} | \Psi_m \rangle|$ is larger than zero. Due to certain symmetry condition, this matrix element may evaluate to zero, forbidding transition between p and m stationary states.

---

[7]Spontaneous emission happens even when there is no field incident on the photon, and this can be understood by the presence of vacuum field fluctuation to be discussed later.

Appendix

A   Time Independent Perturbation Theory
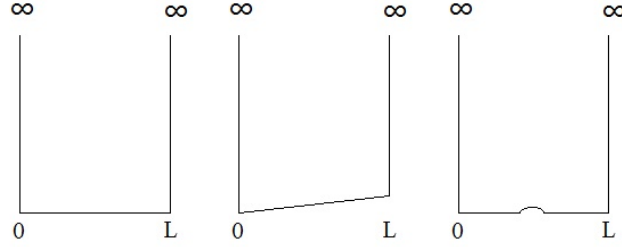
Figure A1: The infinite potential well on the left represents the unperturbed problem. The middle figure represents a perturbation due to a tiny electric field. The right figure represents a perturbation due to imperfection in fabrication or impurities.

Let us assume that we can solve a problem when it is unperturbed, such as the infinite potential well problem. We like to solve the problem when, say there is a small electric field in the well indicated by a potential with a gradient, or that the well is imperfect due to some impurities indicated by a small bump (see Figure A1). The Schrödinger equation for the perturbed system can be written as

$$(\hat{H}_0 + \gamma \hat{H}_p)|\phi\rangle = E|\phi\rangle \tag{A1}$$

where $\hat{H}_0$ is the Hamiltonian of the unperturbed system whose solution is known, and $\gamma \hat{H}_p$ is due to the small perturbation where γ is a small parameter. Here, $\hat{H}_0$ can be the Hamiltonian of the infinite potential well, for instance. In the above equation, $|\phi\rangle$ and E are both unknowns, but we can write them in a perturbation series or expansion, namely

$$|\phi\rangle = |\phi^{(0)}\rangle + \gamma|\phi^{(1)}\rangle + \gamma^2|\phi^{(2)}\rangle + \dots \tag{A2}$$

$$E = E^{(0)} + \gamma E^{(1)} + \gamma^2 E^{(2)} + \dots \tag{A3}$$

Upon substituting the above series into (A1), we obtain

$$\left(\hat{H}_0 + \gamma \hat{H}_p\right)\left(|\phi^{(0)}\rangle + \gamma|\phi^{(1)}\rangle + \gamma^2|\phi^{(2)}\rangle + \dots\right)$$
$$= \left(E^{(0)} + \gamma E^{(1)} + \gamma^2 E^{(2)} + \dots\right)\left(|\phi^{(0)}\rangle + \gamma|\phi^{(1)}\rangle + \gamma^2|\phi^{(2)}\rangle + \dots\right) \tag{A4}$$

The left-hand side of (A4) can be expanded and rewritten on a power series in γ

$$a_0 + a_1\gamma + a_2\gamma^2 + \dots \tag{A5}$$

while the right-hand side is similarly written as

$$b_0 + b_1\gamma + b_2\gamma^2 + ... \tag{A6}$$

These two power series in γ are equal only if $a_i = b_i$, $i = 0, 1, ..., \infty$.[8]

Equating the coefficients of the power series on both sides of ($A4$) we have the following equations:

Zeroth Order:

$$\hat{H}_0|\phi^{(0)}\rangle = E^{(0)}|\phi^{(0)}\rangle \tag{A7}$$

First Order:

$$\hat{H}_0|\phi^{(1)}\rangle + \hat{H}_p|\phi^{(0)}\rangle = E^{(0)}|\phi^{(1)}\rangle + E^{(1)}|\phi^{(0)}\rangle \tag{A8}$$

Second Order:

$$\hat{H}_0|\phi^{(2)}\rangle + \hat{H}_p|\phi^{(1)}\rangle = E^{(0)}|\phi^{(2)}\rangle + E^{(1)}|\phi^{(1)}\rangle + E^{(2)}|\phi^{(0)}\rangle \tag{A9}$$

We assume that the zeroth order equation is known in terms of an eigenstate $|\psi_m\rangle$ with energy E_m. In other words

$$|\phi^{(0)}\rangle = |\psi_m\rangle, \quad E^{(0)} = E_m \tag{A10}$$

We will use this knowledge to solve the first order equation ($A8$) above.

Before we proceed further, a note is in order regarding the uniqueness of the eigenvalue problem (A1). An eigenvector is known only within a multiplicative factor. Hence, its length is indeterminate. This non-uniqueness in its length manifests in the non-uniqueness of the value of the perturbation series (A2) as we shall see later. To achieve uniqueness, it is best to pin down the length of the total eigenvector given by (A2). We fix the length of the eigenvector $|\phi\rangle$ by requiring that

$$\langle\psi_m|\phi\rangle = 1 \tag{A11}$$

With this requirement, we substitute (A2) into the above. Since $\langle\psi_m|\phi^{(0)}\rangle = 1$, because $|\phi^{(0)}\rangle = |\psi_m\rangle$, it is easy to show that $\langle\psi_m|\phi^{(i)}\rangle = 0$, $i > 0$. As a consequence, $|\phi^{(i)}\rangle$ is orthogonal to $|\psi_m\rangle$. The perturbation series is not necessarily normalized, but it can be normalized later after the series has been calculated.

A1   First Order Perturbation

Next, to find the first order corrections to the eigenvalue and the eigenvector, we move the unknowns $|\phi^{(1)}\rangle$ to the left of ($A8$). We then have

$$\left(\hat{H}_0 - E_m\right)|\phi^{(1)}\rangle = E^{(1)}|\psi_m\rangle - \hat{H}_p|\psi_m\rangle \tag{A12}$$

---

[8]This can be easily proved by setting $\gamma = 0$, which immediately implies that $a_0 = b_0$. By differentiating the power series again and let $\gamma = 0$, we prove that $a_i = b_i$ etc.

where we have made use of (A10). Notice that the above equation is non-unique since the operator $\left(\hat{H}_0 - E_m\right)$ has a null space with a null space vector $|\psi_m\rangle$.

Testing the above equation with $\langle\psi_m|$, we have

$$\langle\psi_m|\hat{H}_0 - E_m|\phi^{(1)}\rangle = E^{(1)} - \langle\psi_m|\hat{H}_p|\psi_m\rangle \tag{A13}$$

But $\langle\psi_m|\left(\hat{H}_0 - E_m\right) = 0$ because $\hat{H}_0|\psi_m\rangle = E_m|\psi_m\rangle$. Hence, the above gives

$$E^{(1)} = \langle\psi_m|\hat{H}_p|\psi_m\rangle \tag{A14}$$

the first order correction to the energy of the perturbed system.

Therefore, one of the two unknowns in (A12) is found. The remaining unknown $|\phi^{(1)}\rangle$ can be expanded in terms of the eigenfunctions of the unperturbed system, that is

$$|\phi^{(1)}\rangle = \sum_n a_n^{(1)}|\psi_n\rangle \tag{A15}$$

First, testing the equation (A12) with $\langle\psi_i|$, we have

$$\langle\psi_i|\hat{H}_0 - E_m|\phi^{(1)}\rangle = E^{(1)}\langle\psi_i|\psi_m\rangle - \langle\psi_i|\hat{H}_p|\psi_m\rangle \tag{A16}$$

Upon substituting (A15) into the above, the left-hand side evaluates to

$$\langle\psi_i|E_i - E_m|\phi^{(1)}\rangle = (E_i - E_m)\sum_n \langle\psi_i|a_n^{(1)}|\psi_n\rangle = (E_i - E_m)a_i^{(1)} \tag{A17}$$

The right-hand side of (A16), for $i \neq m$, is just $-\langle\psi_i|\hat{H}_p|\psi_m\rangle$. Hence

$$a_i^{(1)} = \frac{\langle\psi_i|\hat{H}_p|\psi_m\rangle}{E_m - E_i}, i \neq m \tag{A18}$$

When $i = m$, (A17) evaluates to zero, while the right hand side of (A16) also evaluates to zero because of (A14). Hence $a_m^{(1)}$ is undefined. We choose $a_m^{(1)} = 0$ for a number of reasons: It makes the correction term unique since $|\psi^{(1)}\rangle$ is orthogonal to $|\psi^{(0)}\rangle$. It makes the normalization of the eigenvector $|\phi\rangle$ accurate to second order even though the correction is first order. It will also make the second order corrections much simpler to find.

A2   Second Order Perturbation

To find the second order corrections, we rewrite (A9) with the unknown $|\phi^{(2)}\rangle$ on the left hand side. Then (A9) becomes

$$\left(\hat{H}_0 - E_m\right)|\phi^{(2)}\rangle = E^{(1)}|\phi^{(1)}\rangle + E^{(2)}|\psi_m\rangle - \hat{H}_p|\phi^{(1)}\rangle \tag{A19}$$

Testing the above with $\langle\psi_m|$, the left hand side becomes zero as before.[9] Since we have made $|\phi^{(1)}\rangle$ orthogonal to $|\psi_m\rangle$, on the right-hand side, only the last two terms remain. Consequently,

$$0 = E^{(2)} - \langle\psi_m|\hat{H}_p|\phi^{(1)}\rangle \tag{A20}$$

or

$$E^{(2)} = \langle\psi_m|\hat{H}_p|\phi^{(1)}\rangle \tag{A21}$$

Letting

$$|\phi^{(2)}\rangle = \sum_n a_n^{(2)}|\psi_n\rangle \tag{A22}$$

and substituting it into the left side of (A19), testing with $\langle\psi_i|$, we have

$$(E_i - E_m)a_i^{(2)} = E^{(1)}\langle\psi_i|\phi^{(1)}\rangle + E^{(2)}\delta_{im} - \langle\psi_i|\hat{H}_p|\phi^{(1)}\rangle \tag{A23}$$

Therefore, for $i \neq m$,

$$a_i^{(2)} = \frac{\langle\psi_i|\hat{H}_p|\phi^{(1)}\rangle}{E_m - E_i} - \frac{E^{(1)}a_i^{(1)}}{E_m - E_i} \tag{A24}$$

When $i = m$, both sides of (A23) vanish, and $a_m^{(2)}$ is undefined. Again, we pick $a_m^{(2)} = 0$ to obtain a unique solution.

A3   Higher Order Corrections

The above procedure can be generalized to arbitrary order. By induction, we notice that the equivalence of (A9) to p-th order is

$$\hat{H}_0|\phi^{(p)}\rangle + \hat{H}_p|\phi^{(p-1)}\rangle = E^{(0)}|\phi^{(p)}\rangle + E^{(1)}|\phi^{(p-1)}\rangle + \cdots + E^{(p)}|\phi^{(0)}\rangle \tag{A25}$$

The above can be rewritten as

$$\left(\hat{H}_0 - E^{(0)}\right)|\phi^{(p)}\rangle = E^{(1)}|\phi^{(p-1)}\rangle + \cdots + E^{(p)}|\phi^{(0)}\rangle - \hat{H}_p|\phi^{(p-1)}\rangle \tag{A26}$$

Testing the above with $\langle\psi_m|$ gives

$$E^{(p)} = \langle\psi_m|\hat{H}_p|\phi^{(p-1)}\rangle \tag{A27}$$

Letting

$$|\phi^{(p)}\rangle = \sum_n a_n^{(p)}|\psi_n\rangle \tag{A28}$$

---

[9]Again, the above is non-unique for the same reason cited for (A12).

in (A25), and testing with $\langle \psi_i |$, one can easily show, for $i \neq m$, and after noting that $E^{(0)} = E_m$, $|\phi^{(0)}\rangle = |\psi_m\rangle$, that

$$a_i^{(p)} = \frac{1}{E_m - E_i} \left(\langle \psi_i | \hat{H}_p | \phi^{(p-1)} \rangle - E^{(1)} a_i^{(p-1)} - E^{(2)} a_i^{(p-2)} + \cdots - E^{(p-1)} a_i^{(1)} \right) \qquad \text{(A29)}$$

It is to be noted that with modern advent of computer technology, and given the availability of numerical methods, the calculation of perturbation theory to very high order is laborious and not necessary. However, a perturbation correction can give us insight on how a small change in the Hamiltonian can change the solution.

Chapter 7

Quantum Mechanics in Crystals

7.1 Introduction

The quantum mechanics of an electron in a crystalline lattice is becoming increasingly important because of the advent of nano-technology. If a crystalline material is perfect, in principle, an electron can move about in it with no resistance. But this is not true. The electron wave function interacts and scatters off the crystalline lattice due to non-ideal crystalline structure and its thermal vibration.

Also, the electron wave function will lose phase coherence as it travels due to its interaction with its environment such as lattice vibration and impurity scattering. The time for which a wave function loses phase coherence is the phase relaxation time. The distance for which the electron travels before it loses phase coherence is the coherence length. If the coherence length is much larger than the de Broglie wavelength of the electron, the electron can be approximately described by a particle.

Moreover, there are impurities or defects that would scatter the electron, or electron wave causing it to lose momentum. The distance that an electron would travel without experiencing scattering is called the mean-free path of the electron. If the bulk material is much larger that the mean-free path, then an electron will experience much scattering as it travels from one end of the bulk material to the other end. However, if the bulk material is small, so that the electron can travel from one end to the other without experiencing scattering, we call this ballistic transport. In this regime, the wave nature of the electron is important, and it is expedient to study an electron as a wave that propagates through crystalline materials.

When the bulk material is much larger than the mean-free path and the coherence length, we are in the regime of macroscopic transport where the electron behaves like a particle. Scattering events are important, and the electron motion is hindered by them. These scattering events develop into resistance for the electrons. The transport of the electron at the macroscopic level can be described by the drift-diffusion equations, which follow from Boltzmann's transport equation. The currents due to electron and hole movements are:

$$\mathbf{J}_n(\mathbf{r}) = -q\mu_n n(\mathbf{r})\nabla V(\mathbf{r}) + qD_n\nabla n(\mathbf{r}) \qquad (7.1.1)$$

$$\mathbf{J}_p(\mathbf{r}) = -q\mu_p p(\mathbf{r})\nabla V(\mathbf{r}) - qD_p\nabla p(\mathbf{r}) \qquad (7.1.2)$$

where q is the charge of the particle, μ_n and μ_p are the mobilities of the electrons and holes respectively, n and p are their densities respectively, and D_n and D_p are their diffusion coefficients. Since $-\nabla V$ gives the electric field, the first term represents the current due to the electric field called the drift current, while the second term expresses that the particle will move from region of high concentration to low concentration, a diffusion effect.

In this chapter, we will discuss mainly on mesoscopic transport regime where the bulk material is much smaller than the mean-free path of the electron. The movement of electron on the lengthscale of an atom is called microscopic transport regime as when an electron moves through a crystalline lattice. The wave nature of electron is important in these regimes.

In a crystal, the atoms are located in a periodic lattice. Hence, when an electron wave propagates on a lattice, it is propagating in a periodic structure. However, the real world is not that simple. Usually, there are more than one electron traveling in a lattice. The electrons see each other's electric field or potentials. There will be electron-electron interaction that has to be accounted for. Moreover, electrons are fermions meaning that they obey Pauli exclusion principle. Two electrons cannot be in the same state simultaneously. There are many-body effects, but we will ignore them here. The many-body effect can be lumped approximately into an effective potential. We will assume that the effective potential is still periodic. This is true when the system is in equilibrium or quiescent state, and may not be true when the electrons are not in equilibrium with respect to each other.

Another effect is that as an electron moves through a lattice, the attractive force between the electron and a nucleus of the atoms distorts the lattice locations. This gives rise to lattice vibrations called phonons. We will ignore electron-phonon coupling here.

In a periodic lattice, the atoms are located at the lattice points given by

$$\mathbf{R}_L = n_1 \mathbf{a}_1 + n_2 \mathbf{a}_2 + n_3 \mathbf{a}_3 \tag{7.1.3}$$

where $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$ are lattice vectors, and n_1, n_2, n_3 are integers. Every lattice point can be related to every other lattice point by a proper choice of n_1, n_2, and n_3.

Assuming that an electron sees a periodic potential in a lattice, then

$$V_P(\mathbf{r} + \mathbf{R}_L) = V_P(\mathbf{r}) \tag{7.1.4}$$

The single electron sees this periodic potential and its wavefunction satisfies

$$\left[-\frac{\hbar^2}{2m_e} \nabla^2 + V_P(\mathbf{r}) \right] \psi(\mathbf{r}) = E \psi(\mathbf{r}) \tag{7.1.5}$$

We will study next the kind of wave that can propagate in this periodic potential, known as Bloch-Floquet waves.[1] The Bloch-Floquet wave analysis will also explain the formation of band structures, such as valence and conduction bands, in crystalline solids.

7.2   Bloch-Floquet Waves

[2]

---

[1] In the subsequent discussions, we will replace m_e with m, with implicit understanding that this is the mass of the electron.

[2] Named after Andre Bloch (1893-1948) and Gaston Floquet (1847-1920).

Let us study the case of a wave propagating in a 1-D periodic structure, where $V(x)$ may look like that of Figure 7.1.
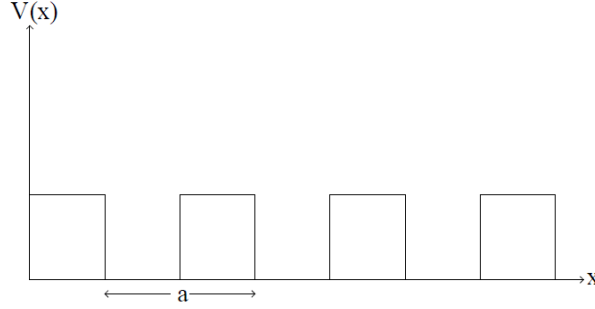
Figure 7.1: A 1D periodic structure of a potential profile for $V(x)$ where a Bloch-Floquet wave can travel on it.

The periodicity of the potential is a. Let us assume that the wave that can exist on such a structure is of the form

$$\psi(k, x) = e^{ikx} u_p(x) \tag{7.2.1}$$

This wavefunction must satisfy the 1-D Schrödinger equation that

$$\left[-\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + V_P(x) \right] \psi(k, x) = E\psi(k, x) \tag{7.2.2}$$

In order for $\psi(k, x)$ to satisfy the above, $u_p(x)$ has to satisfy a certain condition. Since the above is a periodic structure, if one translate the coordinate such that $x \Rightarrow x + na$, then $V_P(x) \Rightarrow V_P(x + na) = V_P(x)$ which remains unchanged. Letting $x \Rightarrow x + na$ in (7.2.1), we have

$$\psi(k, x) \Rightarrow \psi(k, x + na) = e^{ikx} e^{ikna} u_p(x + na) \tag{7.2.3}$$

This new wavefunction must be a solution of Schrödinger equation as well. In order for it to satisfy (7.2.2), we require that

$$u_p(x + na) = u_p(x), \qquad \text{for all } n. \tag{7.2.4}$$

In other words, $u_p(x)$ is a periodic function of x. Since e^{ikna} is just a constant, it is clear that $\psi(x + na)$ satisfies the same equation as (7.2.2). Hence, the form of (7.2.1) that satisfies (7.2.2) has to be such that $u_p(x)$ is a periodic function. Equation (7.2.1) represents a traveling wave modulated by a periodic function $u_p(x)$. However, $u_p(x)$ is more rapidly varying than e^{ikx}. Such a wave is a Bloch wave or a Bloch-Floquet wave. It is to be noted that $u_p(x)$ is a function of k as well, and we should have written $u_p(k, x)$, but we assume that the k dependence is implied.

Since $u_p(x)$ is periodic function with period a, it can be expanded as a Fourier series, namely

$$u_p(x) = \sum_{n=-\infty}^{\infty} a_n e^{i2n\pi x/a} \tag{7.2.5}$$

Noting that the second derivative of (7.2.1) is given by

$$\frac{\mathrm{d}^2}{\mathrm{d}x^2}\left[e^{ikx}u_p(x)\right] = -k^2 e^{ikx}u_p(x) + 2ik e^{ikx}u_p'(x) + e^{ikx}u_p''(x), \tag{7.2.6}$$

substituting (7.2.1) into (7.2.2), and using the above, we have

$$-\frac{\hbar^2}{2m}\left[-k^2 u_p(x) + 2ik u_p'(x) + u_p''(x)\right] + V_p(x)u_p(x) = Eu_p(x) \tag{7.2.7}$$

In the above, if $V_p(x) = 0$, a simple solution to the above is $u_p(x)$ a constant with $E = (\hbar k)^2/(2m)$. Then the Bloch-Floquet wave becomes a simple traveling wave. Otherwise, we can rewrite (7.2.7) as

$$\left\{-\frac{\hbar^2}{2m}\left[\frac{\mathrm{d}^2}{\mathrm{d}x^2} + 2ik\frac{\mathrm{d}}{\mathrm{d}x} - k^2\right] + V_p(x)\right\}u_p(x) = Eu_p(x) \tag{7.2.8}$$

with $u_p(x)$ satisfying the periodic boundary condition within period a.

The above is a simple eigenvalue problem with eigenvector $u_p(x)$ and eigenvalue E. We can use matrix method to convert the above into a matrix equation

$$\sum_n H_{mn}a_n = E\sum_n B_{mn}a_n \tag{7.2.9}$$

by letting $u_p(x) = \sum_{n=1}^{N} a_n\phi_n(x)$, and testing with $\phi_m(x)$. One also ensures that the choice of basis is such that the periodic boundary condition is satisfied. Here, $\phi_n(x)$ are not necessarily orthogonal. In the above,

$$H_{mn} = \langle\phi_m|\hat{H}_p|\phi_n\rangle, \quad B_{mn} = \langle\phi_m|\phi_n\rangle \tag{7.2.10}$$

where

$$\hat{H}_p = -\frac{\hbar^2}{2m}\left[\frac{\mathrm{d}^2}{\mathrm{d}x^2} + 2ik\frac{\mathrm{d}}{\mathrm{d}x} - k^2\right] + V_p(x) \tag{7.2.11}$$

Notice that $\hat{H}_p$ is a function of k, and so H_{mn} is a function of k. It is to be noted that $\hat{H}$ is a Hermitian operator since we have shown that $-i\hbar d/dx$ is a Hermitian operator in (5.2.51) in Chapter 2, so is d^2/dx^2.

Therefore, (7.2.9) can be written as a generalized eigenvalue problem

$$\overline{\mathbf{H}}(k)\cdot\mathbf{a} = E(k)\overline{\mathbf{B}}\cdot\mathbf{a} \tag{7.2.12}$$

with real eigenvalues. In the above, we have explicitly denoted that $\overline{\mathbf{H}}(k)$ is a function of k. Hence, we have to fix k in order to solve the eigenvalue problem (7.2.8), or the generalized matrix eigenvalue problem (7.2.12). Therefore, the eigenvalue E obtained is also a function of k, the wavenumber in the Bloch-Floquet wave (7.2.1). In general, a $E(k)$ plot will look like that in Figure 7.2.
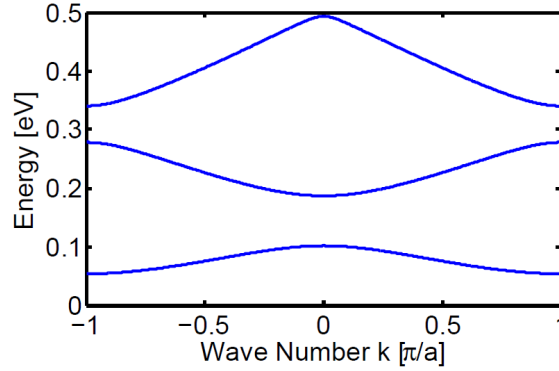
Figure 7.2: The E-k diagram or band structure diagram of a Bloch-Floquet wave propagating in a 1D periodic structure.

7.2.1 Periodicity of $E(k)$

For each k value, there are many possible values of E. Also, in (7.2.1), if we let $k \Rightarrow k + \frac{2p\pi}{a}$, where p is an integer value, then, (7.2.1) becomes

$$\psi(k, x) \Rightarrow \psi(k + 2p\pi/a, x) = e^{ikx} u_p(x) e^{\frac{2ip\pi}{a}x} = e^{ikx} \tilde{u}_p(x) \tag{7.2.13}$$

Notice that $e^{\frac{2ip\pi}{a}x}$ is a periodic function; hence, it can be lumped with $u_p(x)$ to form a new $\tilde{u}_p(x)$ periodic function. They both satisfy the same equation (7.2.8) with the same k value. Therefore, $u_p(x)$ and $\tilde{u}_p(x)$ share the same set of eigenvalues. Consequently, $E(k)$ is a periodic function in k with period $\frac{2\pi}{a}$. Each of this period is called the Brillouin zone (named after Leon Brillouin (1884-1969)). The zone that is centered about the origin is called the first Brillouin zone.

7.2.2 Symmetry of $E(k)$ with respect to k

Furthermore, since every term in the operator of (7.2.2) is real, $\psi^*(k, x)$ is also a solution of it whenever $\psi(k, x)$ is a solution. But

$$\psi^*(k, x) = e^{-ikx} u_p^*(x) \tag{7.2.14}$$

is also a Bloch-Floquet wave with $-k$ wavenumber. In fact, $u_p^*(x)$ remains a periodic function that satisfies an equation that is the conjugate of (7.2.8) but with the same eigenvalue E. Therefore, $E(-k) = E(k)$.

Another way to see this is that if we were to set $k = -k$ in (7.2.12), then assuming that $\phi_n(x)$ is real without loss of generality, then we have

$$\overline{\mathbf{H}}(-k) \cdot \mathbf{a} = \overline{\mathbf{H}}^*(k) \cdot \mathbf{a} = E(-k)\overline{\mathbf{B}} \cdot \mathbf{a} \tag{7.2.15}$$

since $E(-k)$ is real. It is easy to see that this equation is just the complex conjugation of the original equation (7.2.12), and hence, share the same set of eigenvalues. Therefore, $E(-k) = E(k)$.

7.3   Bloch-Floquet Theorem for 3D

In three dimensions, the Bloch-Floquet wave looks like

$$\psi(\mathbf{k},\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}}u_p(\mathbf{r}) \tag{7.3.1}$$

with the property that

$$u_p(\mathbf{r}+\mathbf{R}_L) = u_p(\mathbf{r}) \tag{7.3.2}$$

It is to be noted that $u_p(\mathbf{r})$ is a function of $\mathbf{r}$ as well. We can expand $u_p(\mathbf{r})$ as a generalized Fourier series

$$u_p(\mathbf{r}) = \sum_{\mathbf{G}} a_{\mathbf{G}} e^{i\mathbf{G}\cdot\mathbf{r}} \tag{7.3.3}$$

where $\mathbf{G}$ represents points on the reciprocal lattice denoted by

$$\mathbf{G} = l_1\mathbf{b}_1 + l_2\mathbf{b}_2 + l_3\mathbf{b}_3 \tag{7.3.4}$$

with the property that

$$e^{i\mathbf{G}\cdot\mathbf{R}_L} = 1 \tag{7.3.5}$$

where

$$\mathbf{R}_L = n_1\mathbf{a}_1 + n_2\mathbf{a}_2 + n_3\mathbf{a}_3 \tag{7.3.6}$$

In the above, l_i and n_i, $i = 1, 2, 3$ are integers. The above property indicates that $u_p(\mathbf{r})$ in (7.3.3) is periodic as indicated by (7.3.2). The summation in (7.3.2) is over all possible values of $\mathbf{G}$ or l_1, l_2, and l_3.

One can solve for $\mathbf{b}_1$, $\mathbf{b}_2$ and $\mathbf{b}_3$ to yield

$$\mathbf{b}_1 = \frac{\mathbf{a}_2 \times \mathbf{a}_3}{\mathbf{a}_1 \cdot (\mathbf{a}_2 \times \mathbf{a}_3)}2\pi, \quad \mathbf{b}_2 = \frac{\mathbf{a}_3 \times \mathbf{a}_1}{\mathbf{a}_1 \cdot (\mathbf{a}_2 \times \mathbf{a}_3)}2\pi, \quad \mathbf{b}_3 = \frac{\mathbf{a}_1 \times \mathbf{a}_2}{\mathbf{a}_1 \cdot (\mathbf{a}_2 \times \mathbf{a}_3)}2\pi \tag{7.3.7}$$

for this $\mathbf{G}$. Then by back substitution, (7.3.5) is satisfied. Futhermore,

$$\psi(\mathbf{k}+\mathbf{G},\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}}e^{i\mathbf{G}\cdot\mathbf{r}}u_p(\mathbf{r}) \tag{7.3.8}$$

where

$$e^{i\mathbf{G}\cdot(\mathbf{r}+\mathbf{R}_L)} = e^{i\mathbf{G}\cdot\mathbf{r}} \tag{7.3.9}$$

Since it is a periodic function on the lattice, the $e^{i\mathbf{G}\cdot\mathbf{r}}$ term can be lumped with the unit cell function $u_p(\mathbf{r})$; therefore, the new Bloch-Floquet wave is

$$\psi(\mathbf{k}+\mathbf{G},\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}}\tilde{u}_p(\mathbf{r}) \tag{7.3.10}$$

It is of the same form as (7.3.1); and hence will yield the same set of eigenvalues as $\psi(\mathbf{k},\mathbf{r})$. Consequently, the lattice points are defined by the lattice vector $\mathbf{R}_L$, and the reciprocal lattice points are defined by the reciprocal lattice vector $\mathbf{G}$. The Brillouin zone is defined by the
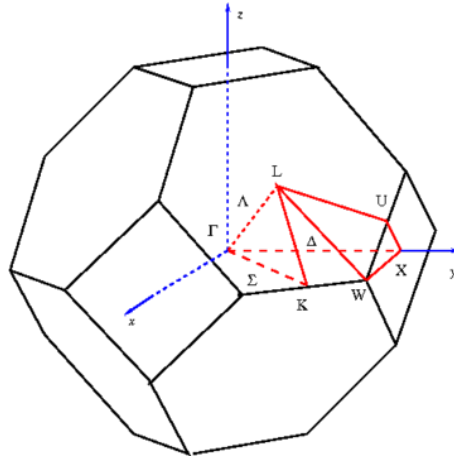
Figure 7.1: The first Brillouin zone of the FCC (face centered cubic) lattice (from Wikipedia).
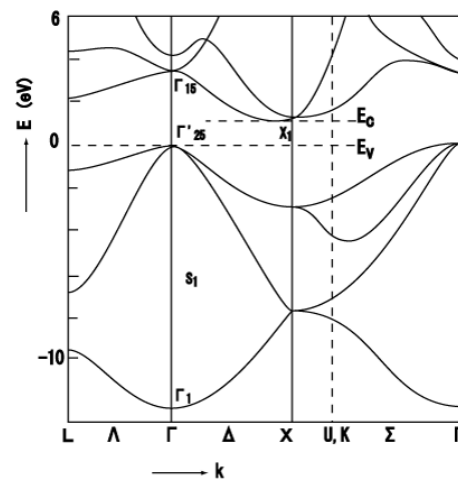
Figure 7.2: The band structure diagram of Silicon along prescribed lines in the first Brillouin zone (from Warwich Physics Dept.).

reciprocal lattice points. The first Brillouin zone is the first unique zone centered around the origin. These zones in 3D can be quite complicated as shown in Figure 7.1.

The band structure diagram is rather complicated and is usually plotted with E versus $\mathbf{k}$ vector along prescribed lines in the first Brillouin zone as shown in Figure 7.2.

We can take yet another viewpoint of the existence of band structure in a crystal. We have studied the tight-binding model of two identical quantum wells. This can be extended to two identical atoms as shown in Figure 7.3. The figure shows how the energy of the trapped modes of the atoms are split versus the spacing of the atoms (as we have found on our tight binding model in the previous chapter). The closer the spacing, the stronger is the coupling between the modes, and the larger is the split in the levels of the degenerate modes.
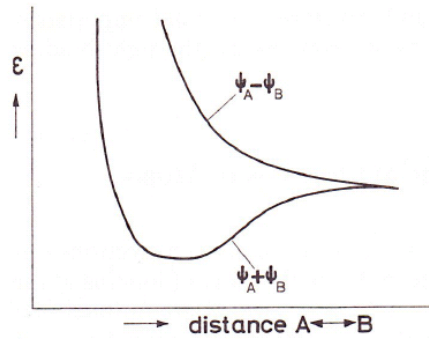
Figure 7.3: The energy levels of two atoms versus the atomic spacing. (from Seeger, Semiconductor Physics).

If we have a system with billions of atoms, each atom has its energy levels due to trapped modes in the atom. If the inter-atomic distance is large, there is no interaction between the atoms, and energy levels will be many billion times degenerate. However, when the interatomic spacing becomes smaller, the wavefunctions from one atom will overlap with the neighboring atoms through tunneling. The interatomic coupling, like the example of the tight-binding model, will split the energy levels and make them into a continuous band of energy. This give rise to energy band where energy states exist and energy band where no states exist, yielding the band structures we have seen.

As we have seen in the simple case of the tight-binding model study of two quantum wells, the degenerate modes split into even and odd mode coupling between them. In one case, the modes are in phase, and the other case, the modes are out of phase with respect to each other. When multitude of mode coupling prevails, there could be infinitely many possible phase relations between the atomic modes, giving rise to Bloch-Floquet wave with different $\exp(i\mathbf{k} \cdot \mathbf{r})$ phase variations across the lattice. The Bloch-Floquet wave establishes the phase relationship between different trapped modes of the atoms as they couple to each other.
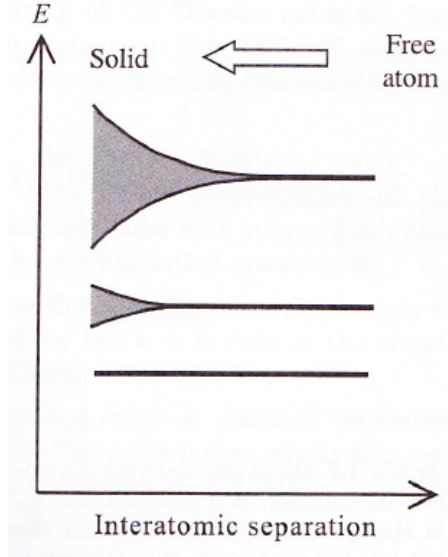
Figure 7.4: The genesis of band structure as the interatomic spacing of the atoms of the lattice becomes smaller and smaller. (From M. Fox, Quantum Optics.)

7.4   Fermi-Dirac Distribution Function

Now that we know how band structures come about, it will be prudent to introduce the Fermi-Dirac distribution function. We will first present this function and derive it later in the course. The Fermi-Dirac distribution function, also called the Fermi-Dirac function, is derived based on Boltzmann law and the Pauli exclusion principle for fermions since electrons are fermions. The Pauli exclusion principle states that no two identical electrons can be in the same state simultaneously. But they are non identical if they have different spins. Hence, each state can admit a spin-up electron, and a spin-down electron.

The Fermi-Dirac distribution function is given by

$$f(E) = \frac{1}{1 + e^{(E-E_f)/(k_B T)}} \tag{7.4.1}$$

where E_f is the Fermi level also called the chemical potential. E_f is actually a function of temperature, but if we assume it to be a constant, $f(E)$ for different T's looks like those shown in Figure 7.1. In the above, $k_B = 1.380650 \times 10^{-23}$ joules/Kelvin, the Boltzmann constant. The value $k_B T = 25.6$ meV at room temperature of $25°$ C or $298°$ K.

The Fermi-Dirac function gives the probability that a state with energy E is being occupied. Hence, $0 < f(E) < 1$. For $T = 0$, it has a binary value of 1 for $E < E_f$ and 0 for $E > E_f$. So all energy states below E_f are occupied and unoccupied for $E > E_f$. For $T > 0$, due to thermal agitation, some electrons are removed from $E < E_f$ and moved to $E > E_f$ states, giving rise to the distribution as shown. Therefore, as electrons are added to the crystalline material, the low energy states will be filled first, followed by high energy
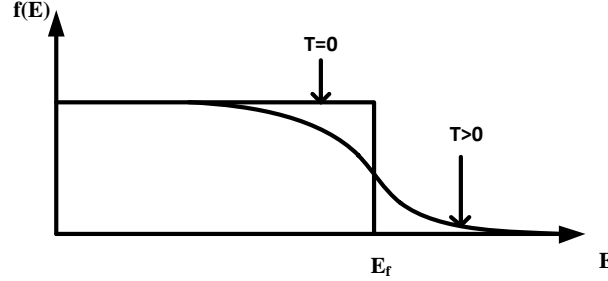
Figure 7.1: The Fermi-Dirac distribution function at $T = 0$ and at $T > 0$.

states. By the Pauli exclusion principle, each energy state can admit at most two electrons, one with spin up, and the other with spin down. Their probability of occupying the state is in accordance with Fermi-Dirac distribution function. The Fermi energy, as we shall see later, is defined to be the Fermi level at zero temperature.

7.4.1   Semiconductor, Metal, and Insulator
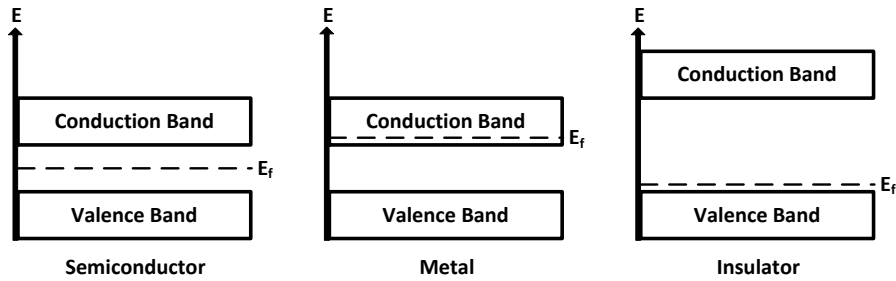
Figure 7.2: The relationship of the Fermi level to the conduction and valence bands for semiconductor, metal, and insulator.

In a semiconductor, E_f is sandwiched between the conduction band and the valence band. At $T = 0$, the valence band is completely occupied, and there are no electrons in the conduction band. Therefore, no conduction is possible. For $T > 0$, or at room temperature, the band gap is not large (typical 1 eV) compared to $k_B T$ so that some electrons can jump from the valence band to the conduction band according to Fermi-Dirac distribution function. The electrons in the conduction band start to conduct. Moreover, empty states are left behind in the valence band known as holes, and they can also conduct electricity.

For a metal, the Fermi level is in the conduction band. So even if $T = 0$, there are electrons in the conduction band allowing the flow of electrons. For an insulator, the band gap is relatively large (typically 4 eV) compared to $k_B T$ at room temperature, and the Fermi level is far from the conduction band, giving zero probability that an electron is in the conduction band. Hence, it cannot conduct electricity.

In semiconductors, the Fermi level moves toward the conduction band when the material is doped with n type impurity since more electron carriers are generated. Conversely, the Fermi level moves toward the valence band when the material is doped with p type impurity since more holes are generated.

7.4.2 Why Do Electrons and Holes Conduct Electricity?
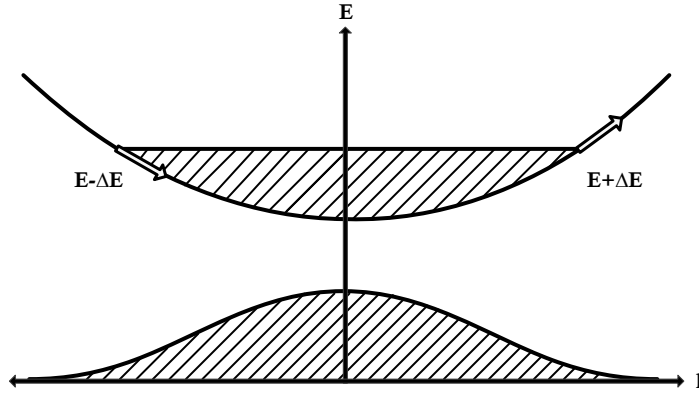
Figure 7.3: When the quantum system of a crystalline material is in equilibrium, the Bloch-Floquet modes correspond to $+k$ are equally likely compared to $-k$. Therefore, there is no net momentum, and hence, no net movement of electrons. However, in the partially filled conduction band, the electrons can readjust themselves so that some have larger $+k$ values compared to $-k$ values, giving rise to a net momentum, and hence movement of the electrons.

In some courses, we are told that electrons conduct electricity because there are empty sites so that electrons can move about and hence gives rise to conduction. But the wave picture of an electron implies that it can be simultaneously everywhere. From a wave picture viewpoint, this story can be quite different. In the valence band, for every electron that has a $\hbar k$ momentum, there is a state with a $-\hbar k$ momentum. The net momentum is always zero, and a fully filled valence band cannot conduct electricity. But in the conduction band, when an electric field is applied, some electrons can move to a higher energy state $E + \Delta E$ with $k + \Delta k$, while others can move to $E - \Delta E$ with $-k + \Delta k$. Hence, there is a net gain of $2\hbar\Delta k$ in the total momentum of the electrons giving rise to electron flow. Similarly, if empty states are created in the valence band, the states can rearrange themselves so that there is a net momentum for the electrons, and hence electricity flows. This is the wave picture of how electrons and holes conduct electricty.

7.5   Effective Mass Schrödinger Equation

The effective mass Schrödinger equation is yet another approximate picture of an electron in a crystalline lattice. It has been used with astounding success for physical phenomena prediction, and gives rise to many new technologies.

The bottom of many conduction band is parabolic. In the vicinity of $k = 0$, we can write the E-k relation as

$$E_k = \frac{\hbar^2 k^2}{2m_e} + V \tag{7.5.1}$$

where m_e is the effective mass chosen to fit the curvature of the parabola, and V is the energy at the bottom of the conduction band. It is to be emphasized that this V is very different from the V_p that has been discussed earlier. Here, V is slowly varying over the lattice structure while V_p varies on the length scales of the lattice constants.

We can write a wave packet as a superposition of waves with different energies or k values, namely

$$\psi(\mathbf{r}, t) = \sum_k c_k u_k(\mathbf{r}) e^{i\mathbf{k}\cdot\mathbf{r}} e^{-i\omega_k t}, \quad \hbar\omega_k = E_k \tag{7.5.2}$$

We are interested in electrons near the bottom of the conduction band where $k \simeq 0$. Then we can approximate

$$u_k(\mathbf{r}) \simeq u_0(\mathbf{r}) \tag{7.5.3}$$

This can be seen from (7.2.8) also when $k \to 0$; namely, $u_k(\mathbf{r})$ is weakly dependent on k when k is small. Then

$$\psi(\mathbf{r}, t) \simeq u_0(\mathbf{r}) \sum_k c_k e^{i\mathbf{k}\cdot\mathbf{r}} e^{-i\omega_k t}$$

$$= u_0(\mathbf{r})\psi_e(\mathbf{r}, t) \tag{7.5.4}$$

where $\psi_e(\mathbf{r}, t)$ is the slowly varying envelope function. In other words, it changes little over many lattice constants. One defines

$$\psi_e(\mathbf{r}, t) = \sum_k c_k e^{i\mathbf{k}\cdot\mathbf{r}} e^{-i\omega_k t} \tag{7.5.5}$$

We can show, using $E_k = \hbar\omega_k$ and (7.5.1), that

$$i\hbar\frac{\partial \psi_e(\mathbf{r}, t)}{\partial t} = \sum_k c_k E_k e^{i\mathbf{k}\cdot\mathbf{r} - i\omega_k t}$$

$$= \sum_k c_k \left(\frac{\hbar^2 k^2}{2m_e} + V \right) e^{i\mathbf{k}\cdot\mathbf{r} - i\omega_k t}$$

$$= \left(-\frac{\hbar^2 \nabla^2}{2m_e} + V \right) \sum_k c_k e^{i\mathbf{k}\cdot\mathbf{r} - i\omega_k t} \tag{7.5.6}$$

or

$$i\hbar\frac{\partial\psi_e(\mathbf{r},t)}{\partial t} = \left(-\frac{\hbar^2\nabla^2}{2m_e} + V\right)\psi_e(\mathbf{r},t) \tag{7.5.7}$$

The above is the effective mass Schrödinger equation. It is valid when k is small, or $e^{i\mathbf{k}\cdot\mathbf{r}}$ is slowly varying compared to the lattice spacing a. In the above, we can assume that $V(\mathbf{r})$ is slowly varying, and that the effective mass $m_e(\mathbf{r})$ is also a slowly varying function of $\mathbf{r}$. Then the above can be correctly written as

$$-\frac{\hbar^2}{2}\nabla\cdot\frac{1}{m_e(\mathbf{r})}\nabla\psi_e(\mathbf{r},t) + V(\mathbf{r})\psi_e(\mathbf{r},t) = i\hbar\frac{\partial}{\partial t}\psi_e(\mathbf{r},t) \tag{7.5.8}$$

The essence of the above approximation is that if the material properties are slowly varying with respect to the lattice spacing, an electron "sees" locally an environment of an infinite lattice. Consequently, if there are two domains with different V and m_e, the above equation induces the boundary conditions at the interface of the domains as

$$\psi_{e1}(\mathbf{r}) = \psi_{e2}(\mathbf{r}), \quad \mathbf{r}\in S \tag{7.5.9}$$

$$\frac{1}{m_{e1}}\frac{\partial}{\partial n}\psi_{e1}(\mathbf{r}) = \frac{1}{m_{e2}}\frac{\partial}{\partial n}\psi_{e2}(\mathbf{r}), \quad \mathbf{r}\in S \tag{7.5.10}$$

where S is the interface separating the two domains and $\partial/\partial n$ refers to normal derivative. Even though $V(\mathbf{r})$ and $m_e(\mathbf{r})$ have jump discontinuities at such an interface, we assume that the discontinuities are still slowly varying compared to the lattice spacings.

It is to be noted that the particle current that we have defined earlier in (4.4.27), Chapter 4, is proportional to

$$\frac{1}{m_e(\mathbf{r})}\psi_e^*(\mathbf{r},t)\nabla\psi_e(\mathbf{r},t) \tag{7.5.11}$$

Hence, the boundary conditions (7.5.10) and (7.5.10) correspond to the continuity of the normal component of the particle current across an interface. The normal component of the particle current is given by

$$\frac{1}{m_e(\mathbf{r})}\psi_e^*(\mathbf{r},t)\hat{n}\cdot\nabla\psi_e(\mathbf{r},t) \tag{7.5.12}$$

which is proportional to the terms in (7.5.10) since $\hat{n}\cdot\nabla = \partial/\partial n$.

7.6 Heterojunctions and Quantum Wells

When two identical semiconductor crystalline materials are joint together, but doped differently, the junction is known as a homojunction. However, when the two crystalline materials are different, we have a heterojunction. We now know that an electron propagating in a semiconductor crystalline lattice resembles an electron in vacuum, except that it has a different effective mass, and it sees a potential that is the bottom of the conduction band. Hence, when
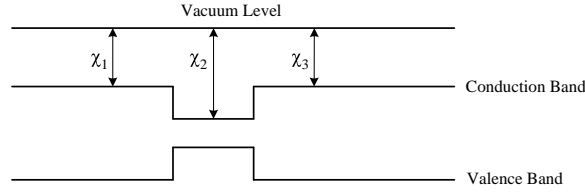
Figure 7.1: Heterojunctions where the energy levels can be aligned using Anderson's rule. Multiple heterojunctions can be used to make quantum wells.

two different materials are joint together to form a junction, we can use Anderson's rule to line up the energy level. Anderson's rule requires that the vacuum levels of the two materials be aligned. But the conduction band of the two materials need not be aligned, since the two materials can have different electron affinity χ. The electron affinity χ is the energy needed to lift an electron from the bottom of the conduction band to the vacuum level.

Because the electron follows the bottom of the conduction band, it will see a quantum well when more than one heterojunctions are used. However, Anderson's rule is not perfect, as the effective mass model is only good for infinite lattice and imperfections can exist at the heterojunctions. Therefore, the rule is used as a guide.

In order to grow one material on top of another material without defects, the lattice constants of the two materials have to be matched. A popular alloy compound for such growth is alluminum gallium arsenide, $Al_xGa_{1-x}As$, where x denotes the percentage of Al in the alloy. The lattice constant remains invariant as x is varied, but the bandgap of the compound changes.

In the early days, vapour-phase epitaxy (VPE) and liquid-phase epitaxy (LPE) were used to grow semiconductor crystals. Later, metal-organic chemical vapor deposition (MOCVD), and molecular beam epitaxy (MBE) were used. The latter technologies allow the crystalline structure grown layer by layer. Many new technologies ensue in this area: heterojunction bipolar transistor (HBT), high electron mobility transistor (HEMT), multiple quantum well (MQW) lasers, as well as light emitting diode (LED) and laser diode (LD). Nick Holonyak (1928-) plays a major role in the development of LED and heterojunction semiconductor laser. He also recently developed the transistor laser with Milton Feng (1950-).

The use of heterojunctions confines the electrons and photons to a potential well, allowing for the making of the first room temperature semiconductor laser. A Nobel Prize was awarded to Herbert Kroemern (1928-) and Zhores Alferov (1930-) in 2000 for their contributions in this area.

7.7   Density of States (DOS)

The density of states is a function of energy E that gives the number of states per unit energy interval. It is important for studying the optical transition probability and optical absorption spectral of materials. As can be seen, when tight binding model is used to model the energy levels of a cluster of atoms, many more energy levels are generated due to inter-

atomic coupling. When one has an infinite number of atoms, a continuum of energy levels are generated. Then the DOS becomes a continuum function of the energy. The density of states $g(E)$ can be neatly written as

$$g(E) = \sum_\alpha \delta(E - E_\alpha) \tag{7.7.1}$$

The above resembles the spectrum of say the hydrogen atom with energy levels for different E_α, when the density of states is very sparse. These energy levels are discrete and countable at this point. The above is valid even when the number of levels per unit energy is very high. Moreover, the number of states between E_0 and $E_0 + \Delta E$ is given by

$$N_{\Delta E} = \int_{E_0}^{E_0 + \Delta E} g(E) dE \tag{7.7.2}$$

The above fits the definition of DOS.

To find the DOS a bulk material, its energy levels can be first studied as the energy levels of an electron bounded in a cuboid box. We can use periodic boundary condition to discretize the k vector in the Bloch-Floquet wave so that it is countable, or that

$$\mathbf{k}_L = l_1 \frac{2\pi}{L_1} \hat{a}_1 + l_2 \frac{2\pi}{L_2} \hat{a}_2 + l_3 \frac{2\pi}{L_3} \hat{a}_3 \tag{7.7.3}$$

In 1D, it is

$$k_l = l \frac{2\pi}{L} = l\Delta k \tag{7.7.4}$$

where $\Delta k = \frac{2\pi}{L}$. When we count the number of discrete modes, we have 1 mode per Δk on the k_l real line. Or the density of states is

$$g_{1D}(k) = 1/\Delta k = L/2\pi \tag{7.7.5}$$

The number of states is proportional to the length of L. In 2D, the density of states can be easily derived to be

$$g_{2D}(k) = L_1 L_2 / (2\pi)^2 = A/(2\pi)^2 \tag{7.7.6}$$

In 3D, it becomes

$$g_{3D}(k) = L_1 L_2 L_3 / (2\pi)^3 = V/(2\pi)^3 \tag{7.7.7}$$

We can normalize the density of states with respect to length L, area A, and volume V for 1D, 2D, and 3D, respectively. Then

$$g_{1D}(k) = 1/2\pi, \; g_{2D}(k) = 1/(2\pi)^2, \; g_{3D}(k) = 1/(2\pi)^3 \tag{7.7.8}$$

We may want to count the density of states per unit energy instead as shown by the above E-k diagram. Even though the states are evenly spaced on the k line, they are not evenly spread on the E line, thereby, altering the density of states per ΔE. In general, we can write, in $3D$,

$$\begin{aligned} g_{3D}(E)\, dE &= 2g_{3D}(k)\, d^3k = 2g_{3D}(k)\, 4\pi k^2 dk \\ &= 8\pi g_{3D}(k)\, k^2 \frac{dk}{dE} dE \end{aligned} \tag{7.7.9}$$
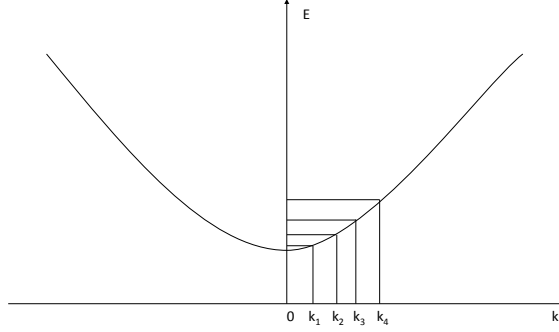
Figure 7.1: If every state on the k line corresponds to a state on the E line, then the density of states per unit length on the E line is different from that on the k line.

where $g_{3D}(E)$ is the density of states per unit energy, while $g_{3D}(k)$ is the density of states per unit wavenumber. We put a factor of 2 for two electron states per energy state.

Therefore, the relation between DOS on the E line and the k line is

$$g_{3D}(E) = 8\pi g_{3D}(k)k^2 \frac{dk}{dE} \qquad (7.7.10)$$

Since

$$E = \frac{\hbar^2 k^2}{2m_e} + V \qquad (7.7.11)$$

for a free electron in 3D, we have

$$k = \sqrt{E - V}\frac{\sqrt{2m_e}}{\hbar} \qquad (7.7.12)$$

$$\frac{dk}{dE} = \sqrt{\frac{m_e}{2\hbar^2}}\frac{1}{\sqrt{E - V}} \qquad (7.7.13)$$

and

$$g_{3D}(E) = \frac{1}{2\pi^2}\left(\frac{2m_e}{\hbar^2}\right)^{\frac{3}{2}}(E - V)^{\frac{1}{2}} \qquad (7.7.14)$$

The above is the DOS in the conduction band of a bulk material. A plot of this DOS is given as solid line in Figure 7.3.

7.7.0.1 Fermi Level and Fermi Energy–An Application of DOS

The electron density in a bulk material is given by

$$N = \int_0^\infty g_{3D}(E)f(E)dE \qquad (7.7.15)$$

where $f(E)$ is the Fermi-Dirac distribution function given in (7.4.1). When T changes, the electron density remains constant, but $f(E)$ changes in shape. Therefore, E_f in (7.4.1) has to change to maintain the constancy of N: E_f is a function of temperature T. The value of E_f at $T = 0$ is known as the Fermi energy E_F. In this case, the integral above can be truncated abruptly at $E = E_F$, we can rewrite it as

$$N = \int_0^{E_F} g_{3D}(E)f(E)dE \qquad (7.7.16)$$

From knowing N which can be ascertained by counting the number of conduction electrons, we can obtain the value of E_F, the Fermi energy, from the above equation.

7.7.1 DOS in a Quantum Well

In a 1D quantum well, the potential varies as a function of z, but is independent of x and y. Then the effective mass Schrödinger equation becomes

$$-\frac{\hbar^2}{2}\nabla\frac{1}{m_e(z)}\nabla\psi(\mathbf{r}) + V(z)\psi(\mathbf{r}) = E\psi(\mathbf{r}) \qquad (7.7.17)$$

We can further separate the z variation from the x and y variations to get

$$-\frac{\hbar^2}{2m_e}\nabla_s^2\psi(\mathbf{r}) - \frac{\hbar^2}{2}\frac{\partial}{\partial z}\frac{1}{m_e}\frac{\partial}{\partial z}\psi(\mathbf{r}) + V(z)\psi(\mathbf{r}) = E\psi(\mathbf{r}) \qquad (7.7.18)$$

where $\nabla_s^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$.

By separation of variables, we let

$$\psi(\mathbf{r}) = \psi_n(z)\psi_s(\mathbf{r}_s) \qquad (7.7.19)$$

where $\mathbf{r}_s = \hat{x}x + \hat{y}y$. Then, we can show assuming m_e is constant, that[3]

$$-\frac{\hbar^2}{2m_e}\nabla_s^2\psi_s(\mathbf{r}_s) = E_s\psi_s(\mathbf{r}_s) \qquad (7.7.20)$$

and

$$-\frac{\hbar^2}{2}\frac{d}{dz}\frac{1}{m_e}\frac{d}{dz}\psi_n(z) + V(z)\psi_n(z) = E_n\psi_n(z) \qquad (7.7.21)$$

---

[3]When m_e is not a constant, the problem is still separable, but the solution is more complicated.

In the above

$$E = E_s + E_n \tag{7.7.22}$$

Equation (7.7.21) can be solved for the eigenvalues E_n. For example, it can be the 1-D finite potential well problem, or the infinite potential well problem. Alternatively, it can have a complicated $V(z)$ and an $m_e(z)$ that are amenable to numerical methods only.

In the x-y direction, we let

$$\psi(\mathbf{r}_s) \propto e^{i\mathbf{k}_s \cdot \mathbf{r}_s} \tag{7.7.23}$$

where $\mathbf{k}_s = \hat{x}k_x + \hat{y}k_y$. Then,

$$E_s = \frac{\hbar^2 k_s^2}{2m_e} \tag{7.7.24}$$

where $k_s^2 = k_x^2 + k_y^2$. To make the solution simple, it is better to assume that m_e is constant. Therefore, the total E-k_s diagram looks like that shown in Figure 7.2
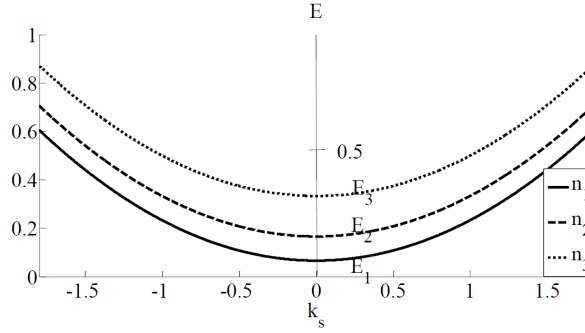
Figure 7.2: The subband E-k_s relation of a quantum well.

In general, $g_{2D}(k_s) = \frac{1}{(2\pi)^2}$. We need to transfer this from the variable k_s to variable E_s using (7.7.24). And we have

$$g_{2D}(E_s)dE_s = 2g_{2D}(k_s)2\pi k_s \frac{dk_s}{dE_s}dE_s \tag{7.7.25}$$

Finally, we get the DOS in terms of E_s as

$$g_{2D}(E_s) = \frac{m_e}{\pi\hbar^2} \tag{7.7.26}$$

which is a constant independent of E_s. But the total $E - k_s$ diagram is shown in Figure 7.2. We can then express the DOS in terms of variable E.

Hence, if we march along the E-line from 0 upward, the DOS as a function of total E is shown in Figure 7.3. We pick up density of states contribution from subband n as $E > E_n$. Hence, the total DOS can be written as

$$g_{2D}(E) = \frac{m_e}{\pi\hbar^2} \sum_n \theta(E - E_n) \tag{7.7.27}$$
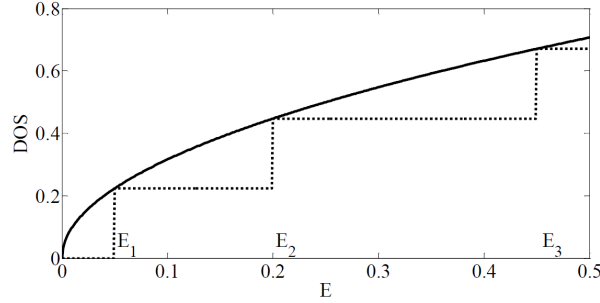
Figure 7.3: The density of states of a 2D quantum well compared to that of a 3D bulk material.

If the 1D quantum well is an infinite potential well, the energy is given by

$$E_n = \frac{\hbar^2}{2m_e} \left(\frac{n\pi}{L}\right)^2 \tag{7.7.28}$$

But density of states for a 3D bulk material is

$$g_{3D}(E) = \frac{1}{2\pi^2} \left(\frac{2m_e}{\hbar^2}\right)^{\frac{3}{2}} E^{\frac{1}{2}} \tag{7.7.29}$$

If we evaluate this at $E = E_n$, we have

$$g_{3D}(E_n) = \frac{m_e}{\pi\hbar^2} \left(\frac{n}{L}\right) = g_{2D}\frac{n}{L} \tag{7.7.30}$$

where g_{2D} here refers to the value in (7.7.26). Therefore, when we plot $g_{2D}(E)/L$ versus E, where $g_{2D}(E)$ is as given in (7.7.27), it touches the g_{3D} at E_n values as shown in Figure 7.3. In other words, every time a new subband is encounter when E is increasing, the value jumps by g_{2D}/L.

When L becomes large, so that the quantization level in the quantum well becomes finer, the DOS plot versus total E is shown in Figure 7.4. The DOS for a quantum well resembles that of a bulk material.

7.7.2   Quantum Wires

Quantum wires can be made by etching a quantum well until the electron is confined in a rod-like region, like a dielectric rod waveguide. In this case, the wavefunction satisfies the following equation

$$-\frac{\hbar^2}{2} \nabla_s \cdot \frac{1}{m_e} \nabla_s \psi(\mathbf{r}) - \frac{\hbar^2}{2m_e} \frac{\partial^2}{\partial z^2} \psi(\mathbf{r}) + V(x, y)\psi(\mathbf{r}) = E\psi(\mathbf{r}) \tag{7.7.31}$$

Similar to separation of variables, we assume that

$$\psi(\mathbf{r}) = e^{ik_z z}\psi_{mn}(x, y) \tag{7.7.32}$$
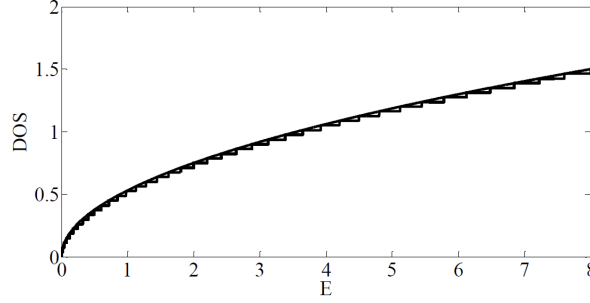
Figure 7.4: The density of states of a 2D quantum well compared to that of a 3D bulk material when the width of the quantum well L is large.

Substituting (7.7.32) into (7.7.31), we have

$$-\frac{\hbar^2}{2}\nabla_s \cdot \frac{1}{m_e}\nabla_s \psi(\mathbf{r_s}) + V(\mathbf{r_s})\psi_{mn}(\mathbf{r_s}) = E_{mn}\psi_{mn}(\mathbf{r}) \tag{7.7.33}$$

$$E_{mn} = E - \frac{\hbar^2 k_z^2}{2m_e} \tag{7.7.34}$$

We can assume m_e to be constant to make the solution simple. The above form gives us

$$k_z = \frac{\sqrt{2m_e}}{\hbar}\sqrt{E - E_{mn}} \tag{7.7.35}$$

For infinite rectangular potential well case, an E_{mn} can be found in closed form yielding

$$E_{mn} = \frac{\hbar^2}{2m_e}\left[\left(\frac{m\pi}{L_1}\right)^2 + \left(\frac{n\pi}{L_2}\right)^2\right] \tag{7.7.36}$$

where L_1 and L_2 are the two sides of the rectangular rod. We can get the density of states as

$$g_{1D}(E)dE = 4g_{1D}(k_z)\frac{dk_z}{dE}dE = \frac{1}{\pi}\frac{\sqrt{2m_e}}{\hbar}\frac{1}{\sqrt{E - E_{mn}}}dE \tag{7.7.37}$$

where $g_{1D} = 1/(2\pi)$. The factor of 4 after the first equality comes about because for every energy state, two electrons can occupy it, and that there are two k_z states, corresponding to $\pm k_z$ states, for every energy state. Consequently,

$$g_{1D}(E) = \frac{1}{\pi}\frac{\sqrt{2m_e}}{\hbar}\frac{1}{\sqrt{E - E_{mn}}} \tag{7.7.38}$$

which is singular at $E = E_{mn}$. The reason is that the $E - k_z$ relationship in (7.7.34) is parabolic, and the mapping at the bottom of the parabola compresses the density of states into a high value.
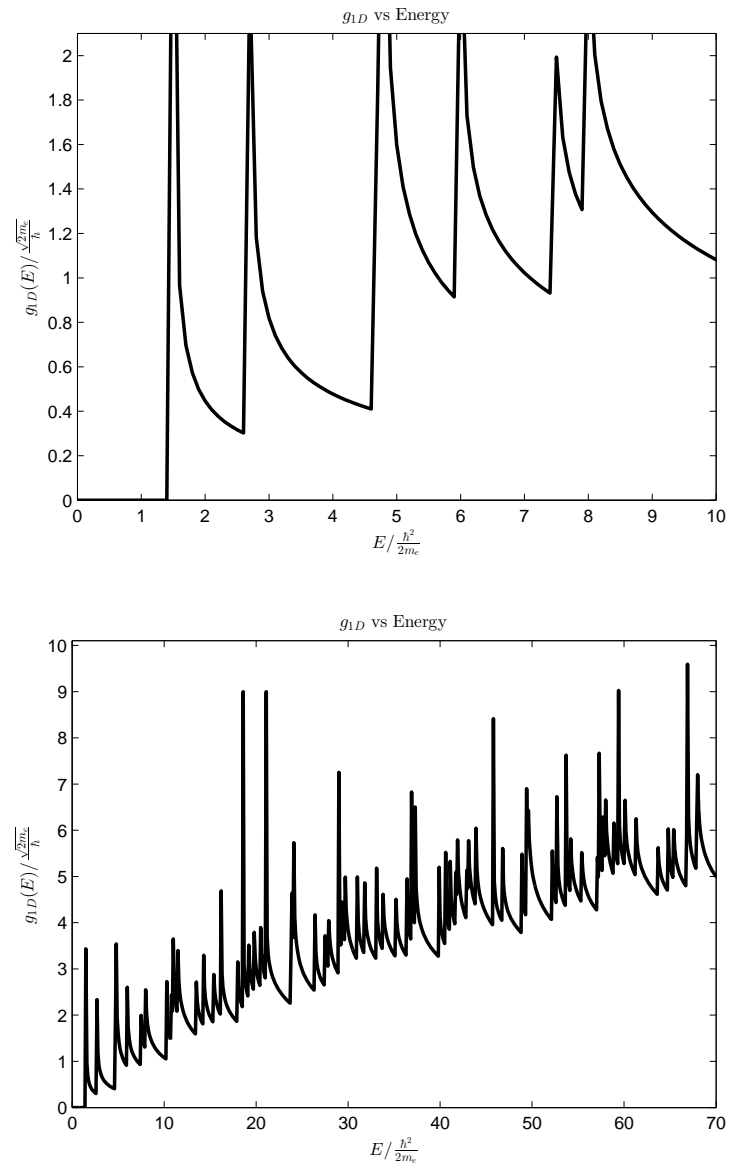
Figure 7.5: The density of states of a quantum wire showing the effect of the subbands plotted on different scales.

When many subbands are included the DOS is

$$g_{1D}(E) = \frac{1}{\pi} \frac{\sqrt{2m_e}}{\hbar} \sum_{m,n} \frac{1}{\sqrt{E - E_{mn}}} \theta(E - E_{mn}) \tag{7.7.39}$$

where θ is a unit step function. The plot of density of states versus E is shown in Figure 7.5. When the size of the quantum wire becomes large, the DOS resembles that of the bulk material.

When we have quantum dots, the density of states will have spikes or delta functions at the resonant energy of the dots. It makes quantum dots behave like artificial atoms.

The density of states is important for determining the absorption and conduction properties of semiconductor materials. The availability of electron states indicate the number of electrons that can participate in electricity conduction. For instance, the number of electron in the conduction band is given by

$$N = \int_{E_c}^{\infty} g(E) f(E) dE \tag{7.7.40}$$

where E_c is the bottom of the conduction band, $g(E)$ is the DOS, and $f(E)$ is the Fermi-Dirac distribution.

The availability of states in different subbands affects the optical absorption property of quantum wells as shown in Figure 7.6. The first peak corresponds to the absorption from the first subband, and the second peak corresponds to the second subband absorption.

In this chapter, when we discuss the propagation of an electron in a crystalline lattice, we have ignored the many-body effect, or the inter-electron effect. Each electron has a Coulomb potential around itself, and it will affect the nearby electrons. But this many body problem is a very difficult to solve, and is usually handled with the density function theory, introduced by Kohn and Sham. Walter Kohn was awarded the Nobel Prize in 1998 with John Pople on this work.
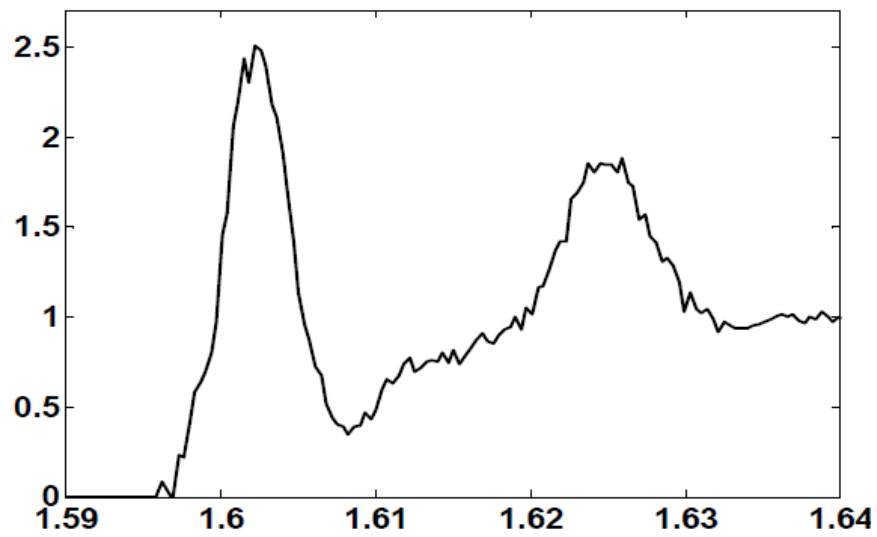
Figure 7.6: The density of states of a quantum well affects the intersubband optical absorption of a quantum well.

Chapter 8

Angular Momentum

8.1 Introduction

In quantum mechanics, we learn that travelling wave e^{ikx} carries a linear momentum proportional to $\hbar k$. Hence, the more oscillatory the wave is, the higher the momentum it carries. However, when an electron is trapped inside the potential of an atom, the wave is not a linearly travelling wave, but a wave that swirls around the nucleus of the atom. Hence, the wave carries mostly angular momentum. The magnitude of this angular momentum, intuitively, is proportional to the angular variation of the wavefunction.

If we take Schrödinger's Equation in free space, it is given by

$$-\frac{\hbar^2}{2m_0}\nabla^2\psi(\mathbf{r}) = E\psi(\mathbf{r}) \tag{8.1.1}$$

We can rewrite this as

$$\left(\nabla^2 + k_0^2\right)\psi(\mathbf{r}) = 0 \tag{8.1.2}$$

where $k_0^2 = 2m_0 E/\hbar^2$. The above is the Helmholtz wave equation (named after Herman von Helmholtz, 1821-1894) whose solution is well known. In cylindrical coordinates, it is

$$\left(\frac{1}{\rho}\frac{\partial}{\partial\rho}\rho\frac{\partial}{\partial\rho} + \frac{1}{\rho^2}\frac{\partial^2}{\partial\phi^2} + \frac{\partial^2}{\partial z^2} + k_0^2\right)\psi(\rho, \phi, z) = 0 \tag{8.1.3}$$

where $\rho = \sqrt{x^2 + y^2}$, $\phi = \tan^{-1}(y/x)$. In spherical coordinates, it is

$$\left(\frac{1}{r^2}\frac{\partial}{\partial r}r^2\frac{\partial}{\partial r} + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\sin\theta\frac{\partial}{\partial\theta} + \frac{1}{r^2\sin^2\theta}\frac{\partial^2}{\partial\phi^2} + k_0^2\right)\psi(r, \theta, \phi) = 0 \tag{8.1.4}$$

where $r = \sqrt{x^2 + y^2 + z^2}$, $\theta = \cos^{-1}(z/r)$, $\phi = \tan^{-1}(y/x)$.

8.1.1   Electron Trapped in a Pill Box

We will first solve the above equation in cylindrical coordinates to gain physical insight into the wave-angular momentum relationship. Hence, we study the case of an electron trapped inside a pill box. The solutions to Helmholtz equation can be found by the separation of variables. For example, to solve (8.1.3), we let

$$\psi(\rho, \phi, z) = R(\rho)\Phi(\phi)Z(z) \tag{8.1.5}$$

Using (8.1.5) in (8.1.3) yields

$$\Phi(\phi)Z(z)\frac{1}{\rho}\frac{\partial}{\partial\rho}\rho\frac{\partial}{\partial\rho}R(\rho) + R(\rho)Z(z)\frac{1}{\rho^2}\frac{\partial^2}{\partial\phi^2}\Phi(\phi)$$
$$+ R(\rho)\Phi(\phi)\frac{\partial^2}{\partial z^2}Z(z) + k_0^2 R(\rho)\Phi(\phi)Z(z) = 0 \tag{8.1.6}$$

To solve the above, we propose two eigenvalue problems,

$$\frac{\partial^2}{\partial z^2}Z(z) = -k_z^2 Z(z) \tag{8.1.7}$$

$$\frac{\partial^2}{\partial\phi^2}\Phi(\phi) = -n^2\Phi(\phi) \tag{8.1.8}$$

The solutions to the above are easily found:

$$Z(z) = e^{\pm ik_z z} \tag{8.1.9}$$

$$\Phi(\phi) = e^{\pm in\phi} \tag{8.1.10}$$

By so doing, Equation (8.1.6) becomes

$$\frac{1}{\rho}\frac{d}{d\rho}\rho\frac{d}{d\rho}R(\rho) - \frac{n^2}{\rho^2}R(\rho) + k_\rho^2 R(\rho) = 0 \tag{8.1.11}$$

where

$$k_\rho^2 = k_0^2 - k_z^2 \tag{8.1.12}$$

Equation (8.1.11) is the Bessel equation (named after Friedrich Bessel, 1784-1846) where solutions are Bessel function $J_n(k_\rho\rho)$ and Neumann function $Y_n(k_\rho\rho)$ (named after Carl Gottfried Neumann, 1832-1925). Here, n is the order of these functions. So the general solution (8.1.5) becomes

$$\psi(\rho, \phi, z) = \left[AJ_n(k_\rho\rho) + BY_n(k_\rho\rho)\right]e^{\pm in\phi}e^{\pm ik_z z} \tag{8.1.13}$$

We can now put an electron in a pill box whose wall is an infinite potential as shown in Figure 8.1. Since $Y_n(k_\rho\rho) \to \infty$, when $\rho \to 0$, we set $B = 0$ to eliminate the Neumann function. Also, $\psi(\rho, \phi, z)$ has to be regular and finite inside the pill box. Matching boundary conditions at the two ends of the pill box, we have

$$\psi(\rho, \phi, z) = AJ_n(k_\rho\rho)\,e^{\pm in\phi}\sin\left(\frac{p\pi}{L}z\right) \tag{8.1.14}$$
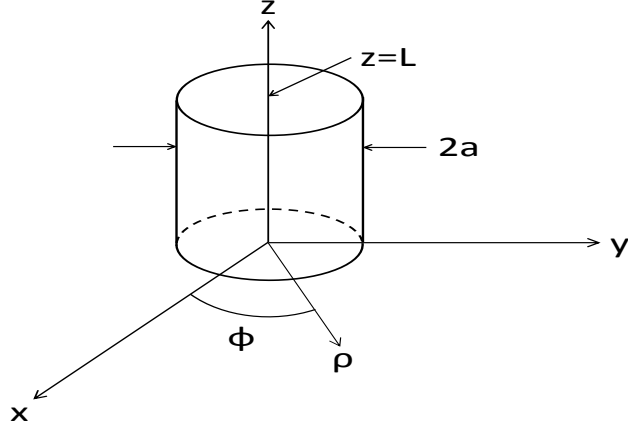
Figure 8.1: The geometry of a cylindrical pill box where an electron is trapped.

Table 8.1: The first few zeros of Bessel functions:

m	ζ_{0m}	ζ_{1m}	ζ_{2m}	ζ_{3m}	ζ_{4m}	ζ_{5m}
1	2.4048	3.8317	5.1356	6.3802	7.5883	8.7715
2	5.5201	7.0156	8.4172	9.7610	11.0647	12.3386
3	8.6537	10.1735	11.6198	13.0152	14.3725	15.7002
4	11.7915	13.3237	14.7960	16.2235	17.6160	18.9801
5	14.9309	16.4706	17.9598	19.4094	20.8269	22.2178

where $k_z = p\pi/L$, for all integer p.

The boundary condition at $\rho = a$ is that $\psi(a, \phi, z) = 0$, or that

$$J_n(k_\rho a) = 0 \tag{8.1.15}$$

The zeros of Bessel functions are tabulated and given by

$$J_n(\zeta_{nm}) = 0 \tag{8.1.16}$$

The first few zeros are given in Table 8.1. Hence, we obtain that

$$k_\rho = \zeta_{nm}/a \tag{8.1.17}$$

So (8.1.14) becomes

$$\psi(\rho, \phi, z) = A J_n\left(\zeta_{nm}\frac{\rho}{a}\right) e^{\pm in\phi} \sin\left(\frac{p\pi}{L}z\right) \tag{8.1.18}$$

In the above, we see a standing wave $\sin(p\pi z/L)$ in the z direction, a standing wave represented by $J_n\left(\zeta_{nm}\frac{\rho}{a}\right)$ in the ρ direction, and a travelling wave $e^{\pm in\phi}$ in the ϕ direction. The
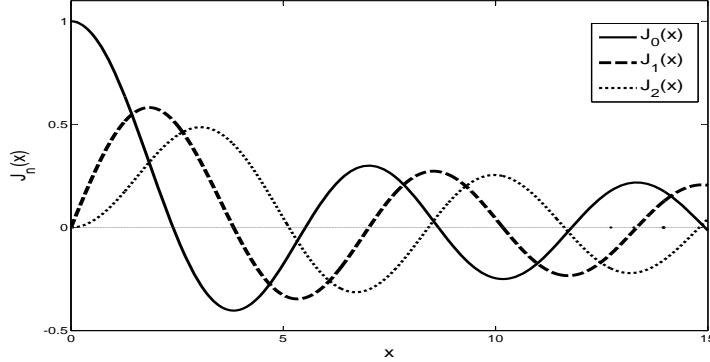
Figure 8.2: The Bessel function of different order as a function of its argument. Notice that the higher the order, the smaller the Bessel function is at the origin.

larger n is, the more rapidly varying is the travelling wave as a function of ϕ, and the more angular momentum it carries.

In fact, plots of $J_n(x)$ versus x for different n is shown in Figure 8.2.[1] We notice that the larger n is, the more rapidly varying $e^{\pm in\phi}$ is, and the more the Bessel function $J_n(x)$ is pulled away from the origin. This is characteristic of a particle swirling around in a circle. The centrifugal force is keeping the particle away from the origin.

From the fact that

$$k_0^2 = k_\rho^2 + k_z^2 = \left(\frac{\zeta_{nm}}{a}\right)^2 + \left(\frac{p\pi}{L}\right)^2 = \frac{2m_0 E}{\hbar^2} \qquad (8.1.19)$$

we can derive that the energy levels of the trapped electron are given by

$$E_{nmp} = \frac{\hbar^2}{2m_0}\left[\left(\frac{\zeta_{nm}}{a}\right)^2 + \left(\frac{p\pi}{L}\right)^2\right] \qquad (8.1.20)$$

Hence, the energy levels of the trapped electron assume quantized values. This is the characteristic of the wave nature of the electron.

8.1.2 Electron Trapped in a Spherical Box

To obtain different physical insight with different wavefunctions, next, we place the electron in a spherical box with infinite potential at the wall. We also solve this problem by separation of variables by letting

$$\psi(r,\theta,\phi) = R(r)\,\Theta(\theta)\,\Phi(\phi) \qquad (8.1.21)$$

---

[1]Ways to compute $J_n(x)$ is well documented and is available in programming toolbox like Matlab.
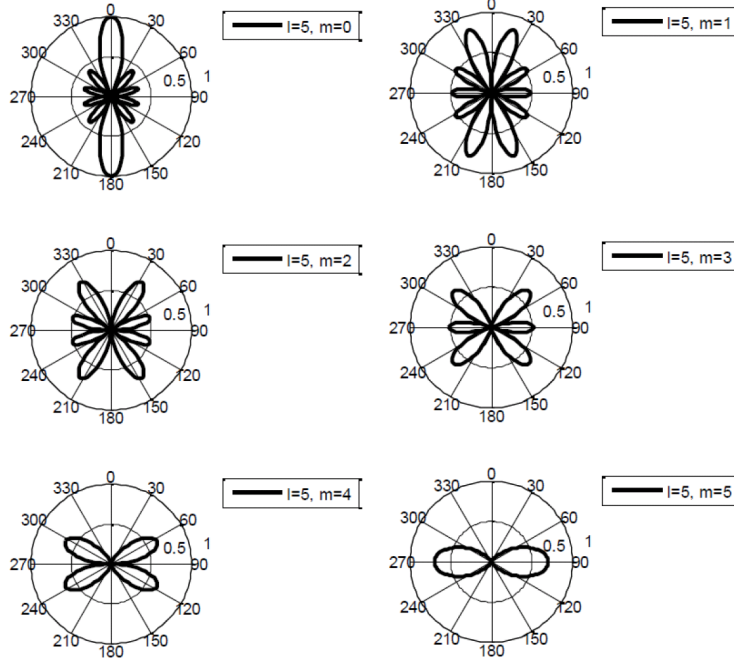
Figure 8.3: Legendre Polynomial with argument $\cos\theta$ plotted for $l = 5$ and different m values. Notice that as m increases, the z component of the angular increases, and the wavefunction spirals away from the z axis, due to the centra-fugal force of the electron. Also, the θ variation of the wavefunction slows down, indicating a lesser component of the angular momentum in the transverse direction.

Using (8.1.21) in (8.1.4), we have

$$\Theta\left(\theta\right)\Phi\left(\phi\right)\frac{1}{r^2}\frac{\partial}{\partial r}r^2\frac{\partial}{\partial r}R\left(r\right)+R\left(r\right)\Phi\left(\phi\right)\frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\Theta\left(\theta\right)\right)$$
$$+R\left(r\right)\Theta\left(\theta\right)\frac{1}{r^2\sin^2\theta}\frac{\partial^2}{\partial\phi^2}\Phi\left(\phi\right)+k_0^2R\left(r\right)\Theta\left(\theta\right)\Phi\left(\phi\right)=0 \tag{8.1.22}$$

We propose another eigenvalue problem such that

$$\frac{\partial^2}{\partial\phi^2}\Phi\left(\phi\right)=-m^2\Phi\left(\phi\right) \tag{8.1.23}$$

Using the above in (8.1.22), $\Phi\left(\phi\right)$ can be canceled in the equation. Then we propose another eigenvalue problem such that

$$\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\sin\theta\frac{\partial}{\partial\theta}\Theta\left(\theta\right)-\frac{m^2}{\sin^2\theta}\Theta\left(\theta\right)=-l\left(l+1\right)\Theta\left(\theta\right) \tag{8.1.24}$$

The solution to (8.1.23) is $e^{\pm im\phi}$ while the solution to (8.1.24) is in terms of associated Legendre polynomials, $P_l^m(x)$ (named after Adrien-Marie Legendre, 1752-1833), or in terms of θ,[2]

$$\Theta(\theta) = P_l^m(\cos\theta), \quad -l \leqslant m \leqslant l \tag{8.1.25}$$

Eventually, the equation for $R(r)$ is

$$\frac{1}{r^2}\frac{d}{dr}r^2\frac{d}{dr}R(r) + \left[k_0^2 - \frac{l(l+1)}{r^2}\right]R(r) = 0 \tag{8.1.26}$$

The last equation is the spherical Bessel equation whose solution is

$$R(r) = Aj_l(k_0 r) + Bh_l^{(1)}(k_0 r) \tag{8.1.27}$$

In the above, $j_l(x)$ represents a spherical Bessel function of order l, which is regular at $x = 0$, while $h_l^{(1)}(x)$ represents a spherical Hankel function of the first kind of order l (named after Herman Hankel, 1839-1873), which is singular at $x = 0$.

When we put this solution inside a spherical box, we set $B = 0$ since $h_l^{(1)}(k_0 r)$ is singular at $r = 0$. Then the general solution to (8.1.4) inside a box is

$$\psi(r,\theta,\phi) = Aj_l(k_0 r)P_l^m(\cos\theta)e^{\pm im\phi}, \qquad -l \leqslant m \leqslant l \tag{8.1.28}$$

Assuming the radius of the sphere is a, to satisfy the boundary condition, we require that

$$j_l(k_0 a) = 0, \tag{8.1.29}$$

Since $j_l(\zeta_{lp}) = 0$, we deduce that $k_0 = \zeta_{lp}/a$. The first few zeros of the spherical Bessel functions are given in Table 8.2.

Table 8.2: The first few zeros of the spherical Bessel functions:

p	ζ_{0p}	ζ_{1p}	ζ_{2p}	ζ_{3p}	ζ_{4p}
1	3.142	4.493	5.763	6.988	8.183
2	6.283	7.725	9.095	10.417	11.705
3	9.425	10.904	12.323	13.698	15.040
4	12.566	14.066	15.515	16.924	18.301

Therefore, the general solution becomes

$$\psi(r,\theta,\phi) = Aj_l\left(\zeta_{lp}\frac{r}{a}\right)P_l^m(\cos\theta)e^{\pm im\phi}, \qquad -l \leqslant m \leqslant l \tag{8.1.30}$$

In the above, $e^{\pm im\phi}$ represents a traveling wave in the ϕ direction, $P_l^m(\cos\theta)$ represents a standing wave in the θ direction. When l becomes large, $P_l^m(x)$ is a higher order polynomial

---

[2]The other solution to (8.1.24) is the associate Legendre function of the second kind, $Q_l^n(\cos\theta)$ but this function is singular for $0 \leqslant \theta \leqslant \pi$, and hence, is not admissible as a solution.

and $P_l^m(\cos\theta)$ is a rapidly varying function of θ. In this case, $e^{\pm im\phi}$, $-l \leqslant m \leqslant l$ could also be a rapidly varying function of ϕ implying the presence of high angular momentum. From Figure 8.3, it is seen that the larger m is, the more slowly varying is the associated Legendre polynomial. The variations from $P_l^m(\cos\theta)$ and $e^{\pm im\phi}$ compensate each other for a fixed l so as to keep the "sum" of the angular momentum a constant when l is fixed.

The function $j_l(x)$ is well known and is plotted in Figure 8.1, showing that when n is large, the centrifugal force of the particle pulls it away from the origin as in the pill-box case in cylindrical coordinates. From the fact that

$$k_0^2 = \left(\frac{\zeta_{lp}}{a}\right)^2 = \frac{2m_0 E}{\hbar^2} \tag{8.1.31}$$

we deduce that the energy levels of the trapped electron are

$$E_{lp} = \frac{\hbar^2}{2m_0}\left(\frac{\zeta_{lp}}{a}\right)^2 \tag{8.1.32}$$

This is again quantized due to the wave nature of the electron.

The above cases illustrate the trapping of an electron by a pill box and a spherical box. In nature, the positive charge of a nucleus forms a potential that can trap an electron. In the case of the hydrogen atom, closed form solutions can be obtained for the trapping of an electron by the Coulomb potential of the positive nucleus. This has been documented in many quantum mechanics books.

8.2   Mathematics of Angular Momentum
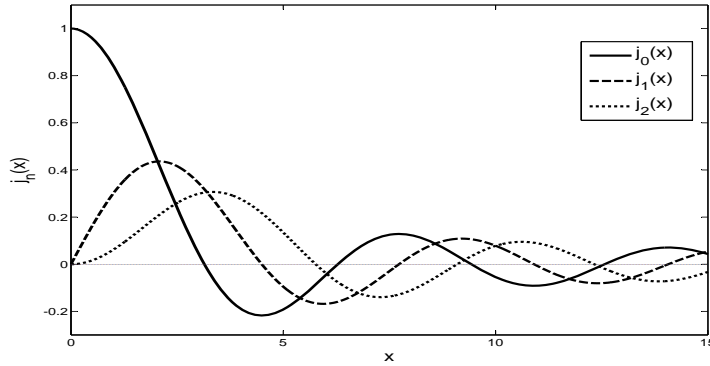
Figure 8.1: Spherical Bessel function of different order versus its argument. The higher order functions are smaller around the origin.

We have shown previously that an electron trapped inside a pill box or a spherical box does display angular momentum. Next, we derive the operator that is related to angular momentum. In classical mechanics, the angular momentum is given by

$$\mathbf{L} = \mathbf{r} \times \mathbf{p} \tag{8.2.1}$$

where $\mathbf{p}$ is the linear momentum. Notice that by taking the cross product of $\mathbf{p}$ with $\mathbf{r}$, only the circumferential component of the linear momentum is extracted. Written explicitly, it is

$$L_x = yp_z - zp_y, \quad L_y = zp_x - xp_z, \quad L_z = xp_y - yp_x \tag{8.2.2}$$

We raise these to operators, and write them as, in coordinate space representation,

$$\hat{L}_x = -i\hbar \left(y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y} \right), \hat{L}_y = -i\hbar \left(z\frac{\partial}{\partial x} - x\frac{\partial}{\partial z} \right), \hat{L}_z = -i\hbar \left(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x} \right), \tag{8.2.3}$$

x, y, z can be raised to be operators as well, but in coordinate space representation, they are just scalars. Actually, we can write (8.2.1) as the operator

$$\hat{\mathbf{L}} = \hat{\mathbf{r}} \times \hat{\mathbf{p}} \tag{8.2.4}$$

The hat signs above indicate that they are operators and not unit vectors. The coordinate space representation of the $\hat{\mathbf{r}}$ operator is just $\mathbf{r}$. The coordinate space representation representation of the momentum operator $\hat{\mathbf{p}}$ has been derived before. Consequently, the above operators can be shown to satisfy the following commutation relations

$$\left[\hat{L}_x, \hat{L}_y \right] = i\hbar\hat{L}_z, \quad \left[\hat{L}_y, \hat{L}_z \right] = i\hbar\hat{L}_x, \quad \left[\hat{L}_z, \hat{L}_x \right] = i\hbar\hat{L}_y \tag{8.2.5}$$

The $\hat{L}^2$ operator is defined as

$$\hat{L}^2 = \hat{\mathbf{L}} \cdot \hat{\mathbf{L}} = \hat{L}_x^2 + \hat{L}_y^2 + \hat{L}_z^2 \tag{8.2.6}$$

8.2.1   Transforming to Spherical Coordinates

The above can be transformed to spherical coordinate system. To show this, we write (8.2.4) in coordinate representation yielding [3]

$$\hat{\mathbf{L}} = -i\hbar\mathbf{r} \times \nabla = -i\hbar \left(\hat{\phi}\frac{\partial}{\partial \theta} - \hat{\theta}\frac{1}{\sin\theta}\frac{\partial}{\partial \phi} \right) \tag{8.2.7}$$

The last equality can be obtained by writing gradient operator ∇ in spherical coordinates. Then

$$\hat{L}_x = \hat{x} \cdot \hat{\mathbf{L}} = -i\hbar \left(\hat{x} \cdot \hat{\phi}\frac{\partial}{\partial \theta} - \hat{x} \cdot \hat{\theta}\frac{1}{\sin\theta}\frac{\partial}{\partial \phi} \right) \tag{8.2.8}$$

But $\hat{x} \cdot \hat{\phi} = -\sin\phi$, $\hat{x} \cdot \hat{\theta} = \hat{x} \cdot \hat{\rho}\cos\theta = \cos\phi\cos\theta$. Then (8.2.8) becomes

$$\hat{L}_x = i\hbar \left(\sin\phi\frac{\partial}{\partial \theta} + \cot\theta\cos\phi\frac{\partial}{\partial \phi} \right) \tag{8.2.9}$$

---

[3]We will use the $\hat{x}, \hat{y}, \hat{z}, \hat{r}, \hat{\theta}, \hat{\phi}$ to denote unit vectors in the respective directions.

Similarly,

$$\hat{L}_y = \hat{y} \cdot \hat{\mathbf{L}} = -i\hbar \left(\hat{y} \cdot \hat{\phi} \frac{\partial}{\partial \theta} - \hat{y} \cdot \hat{\theta} \frac{1}{\sin \theta} \frac{\partial}{\partial \phi} \right) \tag{8.2.10}$$

with $\hat{y} \cdot \hat{\phi} = \cos \phi$, $\hat{y} \cdot \hat{\theta} = \sin \phi \cos \theta$, we have

$$\hat{L}_y = i\hbar \left(-\cos \phi \frac{\partial}{\partial \theta} + \cot \theta \cos \phi \frac{\partial}{\partial \phi} \right) \tag{8.2.11}$$

Similarly,

$$\hat{L}_z = i\hbar \hat{z} \cdot \hat{\theta} \frac{1}{\sin \theta} \frac{\partial}{\partial \phi} = -i\hbar \frac{\partial}{\partial \phi} \tag{8.2.12}$$

The eigenfunction of the $\hat{L}_z$ operator is $\Phi = e^{\pm im\phi}$ so that

$$\hat{L}_z \Phi(\phi) = m\hbar \Phi(\phi) \tag{8.2.13}$$

For $\hat{L}^2$ operator, we have

$$\hat{L}^2 = -\hbar^2 (\mathbf{r} \times \nabla) \cdot (\mathbf{r} \times \nabla) \tag{8.2.14}$$

Using a vector identity that $(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c} = \mathbf{a} \cdot \mathbf{b} \times \mathbf{c}$, we rewrite the above as

$$\hat{L}^2 = -\hbar^2 \mathbf{r} \cdot \nabla \times (\mathbf{r} \times \nabla) \tag{8.2.15}$$

By expressing the $\nabla \times$ operator in spherical coordinates, and that $\mathbf{r} \times \nabla$ can be expressed as in (8.2.7), the above becomes

$$\hat{L}^2 = -\hbar^2 \frac{1}{\sin \theta} \left(\frac{\partial}{\partial \theta} \sin \theta \frac{\partial}{\partial \theta} + \frac{\partial}{\partial \phi} \frac{1}{\sin \theta} \frac{\partial}{\partial \phi} \right) \tag{8.2.16}$$

The above is just the circumferential part of the Laplacian operator in (8.1.4). It is quite easy to show that the above commutes with $\hat{L}_z$. Hence, $\hat{L}^2$ commutes with $\hat{L}_z$. By symmetry, $\hat{L}^2$ also commutes with $\hat{L}_x$ and $\hat{L}_y$.

We can find the eigenfunctions of the $\hat{L}^2$ operator. Traditionally, this eigenfunction and eigenvalue problem is expressed as

$$\hat{L}^2 Y_{lm}(\theta, \phi) = \hbar^2 l(l+1) Y_{lm}(\theta, \phi), \qquad -l \leqslant m \leqslant l \tag{8.2.17}$$

where l is the orbital angular momentum quantum number and m is the azimuthal angular momentum quantum number. Its solution is similar to the method of solving for the solution of Helmholtz wave equation in spherical coordinates. By the separation of variables, the solution is

$$Y_{lm}(\theta, \phi) = C P_l^m(\cos \theta) e^{im\phi}, \qquad -l \leqslant m \leqslant l \tag{8.2.18}$$

which can be normalized to yield

$$Y_{lm}(\theta, \phi) = (-1)^m \sqrt{\frac{(2l+1)(l-m)!}{4\pi(l+m)!}} P_l^m(\cos \theta) e^{im\phi}, \qquad -l \leqslant m \leqslant l \tag{8.2.19}$$

In Hilbert space, this eigenfunction is denoted as $|l, m\rangle$. We can define another operator

$$\hat{L}_s^2 = \hat{L}^2 - \hat{L}_z^2 = \hat{L}_x^2 + \hat{L}_y^2 \tag{8.2.20}$$

It is quite clear that $\hat{L}_z$ commutes with $\hat{L}_s^2$, and $\hat{L}^2$ commutes with $\hat{L}_s^2$. Hence, $\hat{L}_z$, $\hat{L}_s^2$ and $\hat{L}^2$ are mutually commutative, and share the same eigenfunctions $|l, m\rangle$. Therefore,

$$\hat{L}_s^2|l, m\rangle = \hbar^2[l(l+1) - m^2]|l, m\rangle \tag{8.2.21}$$

A pictorial illustration of this concept is given in Figure 8.2. It is to be noted that this picture can be drawn because $\hat{L}_z$, $\hat{L}_s^2$, and $\hat{L}^2$ are mutually commutative. For the common eigenfunction $|l, m\rangle$, their eigenvalues can be "observed" uniquely and simultaneously.[4] The corresponding increase of L_z and decrease of L_s also correlate with the wavefunctions shown in Figure 8.3.

However, $\hat{L}_z$ does not commute with $\hat{L}_x$ or $\hat{L}_y$. Therefore, for an eigenfunction $|l, m\rangle$ that gives the value of $\langle \hat{L}_z \rangle$ precisely, the value of $\langle \hat{L}_x \rangle$ $\langle \hat{L}_y \rangle$ are indeterminate. In fact, from the picture of $|Y_{lm}|^2$, the function is completely indeterminate in the ϕ direction.

8.3   Field-Atom Interaction–A Simple Introduction

Previously in the section on time-dependent perturbation theory, it was seen that an external electric field applied to a quantum well can change the potential distribution, and hence, gives rise to a change or a perturbation of the Hamiltonian. It can be shown that when electromagnetic field is present, the classical Hamiltonian is

$$H = \frac{1}{2m}(\mathbf{p} - q\mathbf{A})^2 + q\phi \tag{8.3.1}$$

where $\mathbf{A}$ and ϕ are electric vector and scalar potentials, respectively, and the corresponding electric field $\mathbf{E} = -\dot{\mathbf{A}} - \nabla\phi$, and magnetic field $\mathbf{B} = \nabla \times \mathbf{A}$. The presence of the electric potential gives rise to a potential in the Hamiltonian $V(\mathbf{r}) = q\phi(\mathbf{r})$. In general, $V(\mathbf{r}) = V_0(\mathbf{r}) + q\phi(\mathbf{r})$ where $V_0(\mathbf{r})$ could be the Coulomb potential from the nucleus of the atom. Then the classical Hamiltonian becomes

$$H = \frac{1}{2m}(\mathbf{p} - q\mathbf{A})^2 + V(\mathbf{r}) \tag{8.3.2}$$

One can expand above squared term to obtain

$$H = \frac{1}{2m}(\mathbf{p}^2 - 2q\mathbf{p} \cdot \mathbf{A} + q^2\mathbf{A} \cdot \mathbf{A}) + V(\mathbf{r}) \tag{8.3.3}$$

The above can be written as $H_0 + H_p$ where H_0 is the unperturbed Hamiltonian when the external field is zero, or

$$H_0 = \frac{1}{2m}\mathbf{p}^2 + V_0(\mathbf{r}) \tag{8.3.4}$$

---

[4]We are actually drawing the eigenvalues of $\hat{L}_s = \sqrt{\hat{L}_x^2 + \hat{L}_y^2}$ and $\hat{L} = \sqrt{\hat{L}_x^2 + \hat{L}_y^2 + \hat{L}_z^2}$, but they are still mutually commutative.

L$_z$ axis

L$_s$

$+2\hbar$

L$_z$ L

$+\hbar$

L$_s$ axis

$-\hbar$

$-2\hbar$
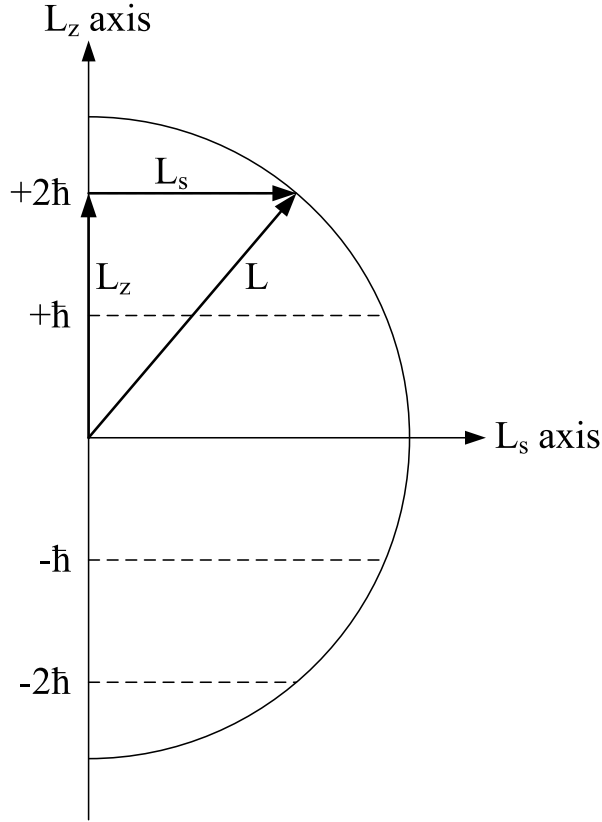
Figure 8.2: A picture showing the angular momentum components, L_z and L_s, for different tilt angles of the total angular momentum L.

and H_p is the perturbing Hamiltonian given by

$$H_p = \frac{1}{2m}(-2q\mathbf{p} \cdot \mathbf{A} + q^2 \mathbf{A} \cdot \mathbf{A}) + q\phi \qquad (8.3.5)$$

Notice that $H_p = 0$ if the field is not turned on. When the field is weak, the quadratic term $\mathbf{A} \cdot \mathbf{A}$ can be ignored, and

$$H_p \approx -\frac{q}{m}\mathbf{p} \cdot \mathbf{A} + q\phi \qquad (8.3.6)$$

For the case of light-atom interaction, one can generate a plane-wave light field such that $\nabla \cdot \mathbf{A} = 0$ with $\phi = 0$. In that case, the above becomes

$$H_p \approx -\frac{q}{m}\mathbf{p} \cdot \mathbf{A} \qquad (8.3.7)$$

In the quantum world, the the momentum $\mathbf{p}$ is elevated to be a momentum operator $\hat{\mathbf{p}}$, and the corresponding Hamiltonian is

$$\hat{H}_p \approx -\frac{q}{m}\hat{\mathbf{p}} \cdot \mathbf{A}(t) = -\frac{q}{m}\mathbf{A}(t) \cdot \hat{\mathbf{p}} \qquad (8.3.8)$$

where $\hat{\mathbf{p}} = -i\hbar\nabla$ in coordinate space. Moreover, since $\nabla \cdot \mathbf{A} = 0$, the $\hat{\mathbf{p}}$ operator commutes with $\mathbf{A}(\mathbf{r})$ and therefore the last equality in the above. In the above, we did not elevate $\mathbf{A}$ to be a quantum operator, and had left it as a classical field.[5] This is also known as the semi-classical approximation.

For a plane wave, without loss of generality,

$$\mathbf{A}(t) = \mathbf{A}_0 \cos(\mathbf{k} \cdot \mathbf{r} - \omega t) \approx \mathbf{A}_0 \cos(\omega t) \qquad (8.3.9)$$

The last approximation above follows if the size of the atom is much smaller than the wavelength in question so that $\mathbf{k} \cdot \mathbf{r}$ is a small number. Then above Hamiltonian can be used in the time-dependent perturbation theory to ascertain the radiative transitions in an atom when it is excited by a light field.

8.3.1   Magnetic Field and Zeeman Effect

When an atom interacts with a low-frequency or static magnetic field, the vector potential can be expressed as

$$\mathbf{A} = \frac{1}{2}\mathbf{B} \times \mathbf{r} \qquad (8.3.10)$$

It can be shown that, if $\mathbf{B}$ is a constant, that $\nabla \times \mathbf{A} = \mathbf{B}$ and $\nabla \cdot \mathbf{A} = 0$ satisfying the Coulomb gauge. Using the above in (8.3.7), one gets[6]

$$H_p \approx -\frac{q}{2m_0}\mathbf{B} \times \mathbf{r} \cdot \mathbf{p} = -\frac{q}{2m}\mathbf{B} \cdot \mathbf{r} \times \mathbf{p} = -\frac{q}{2m_0}\mathbf{B} \cdot \mathbf{L} \qquad (8.3.11)$$

Again, in the quantum world, the above can be elevated to be a quantum operator yielding

$$\hat{H}_p \approx -\frac{q}{2m_0}\mathbf{B} \cdot \hat{\mathbf{L}} \qquad (8.3.12)$$

For a static or low-frequency magnetic field, the number of photons involved is infinitely or almost infinitely large. Hence, the graininess of the field is unimportant. For a magnetic field pointing in the z direction, namely, $\mathbf{B} = \hat{z}B_z$, the above becomes

$$\hat{H}_p \approx -\frac{q}{2m_0}B_z\hat{L}_z \qquad (8.3.13)$$

---

[5]The treatment of $\mathbf{A}$ as a quantum observable is important when the "graininess" of electromagnetic field is indispensable, so that the photon picture or the particle picture of electromagnetic field is important. This is the case when the frequency of the field is high so that $\hbar\omega$ is not trivial, such as in optics, or when the number of photons involved is a small number.

[6]Here, m_0 is used to denote the mass of the electron to distinguish it from m, the azimuthal angular momentum quantum number to be introduced later.

The Hamiltonian is then

$$\hat{H}\psi(\mathbf{r}) = (\hat{H}_0 + \hat{H}_p)\psi(\mathbf{r}) = E\psi(\mathbf{r}) \tag{8.3.14}$$

where $\hat{H}_0 = \frac{\hat{\mathbf{p}}^2}{2m_0} + V_0$. If $\hat{H}_0$ is axially symmetric, and since $\hat{H}_p \sim \frac{\partial}{\partial\phi}$, then $[\hat{H}_0, \hat{H}_p] = 0$. Then it is possible that $\psi(\mathbf{r})$ is both the eigenfunction of both $\hat{H}_0$ and $\hat{H}_p$. Therefore,

$$-\frac{q}{2m_0}B_z\hat{L}_z\psi(\mathbf{r}) = -\frac{q}{2m_0}B_z m\hbar\psi(\mathbf{r}) \tag{8.3.15}$$

If the unperturbed Hamiltonian yields eigen-energy E_0, then the the perturbed Hamiltonian yields

$$E = E_0 - \frac{q}{2m_0}B_z m\hbar \tag{8.3.16}$$

The shift in energy level is due to the coupling of the orbital angular momentum to the perturbing magnetic field. The shift is proportional to the azimuthal angular momentum quantum number m. In the unperturbed Hamiltonian, these eigenfunctions with different m's have the same eigen-energies (eigenvalues), and hence they are degenerate. After the application of the magnetic field, the eigen-energies split and become non-degenerate. This is known as the Zeeman effect, named after Pieter Zeeman (1865-1943).

Appendix

A Infinitesimal Translation

From Taylor series expansion, we have

$$
\begin{aligned}
f(x+a) &= f(x) + af(x) + \frac{a^2}{2!}f^2(x) + \cdots \\
&= \left[1 + a\frac{d}{dx} + \frac{a^2}{2!}\frac{d^2}{dx^2} + \cdots\right]f(x) \\
&= e^{a\frac{d}{dx}}f(x) \tag{A1}
\end{aligned}
$$

The exponential to an operator is interpreted as a Taylor series when it needs to be evaluated. The above can be written with a momentum operator (assuming that $\hbar = 1$)

$$f(x+a) = e^{ia\hat{p}}f(x) \tag{A2}$$

or in Dirac notation

$$|f_a\rangle = e^{ia\hat{p}}|f\rangle \tag{A3}$$

where $|f_a\rangle$ is the state vector representation of the function $f(x+a)$. If a Hamiltonian is translational invariant, it will commute with the translation operator, namely

$$\hat{H}e^{ia\hat{p}}|f\rangle = e^{ia\hat{p}}\hat{H}|f\rangle \tag{A4}$$

or

$$\left[e^{ia\hat{p}}, \hat{H}\right] = 0 \tag{A5}$$

For example, if the Hamiltonian is such that $-\frac{1}{2}d^2/dx^2$, then it is quite clear that the left-hand side of (A4) is

$$-\frac{1}{2}\frac{d^2}{dx^2}e^{ia\hat{p}}f(x) = -\frac{1}{2}\frac{d^2}{dx^2}f(x+a) = -\frac{1}{2}f''(x+a) \tag{A6}$$

The right-hand side of (A4) is

$$e^{ia\hat{p}}\left[-\frac{1}{2}\frac{d^2}{dx^2}f(x)\right] = e^{ia\hat{p}}\left[-\frac{1}{2}f''(x)\right] = -\frac{1}{2}f''(x+a) \tag{A7}$$

which is the same as the left-hand side.

By assuming that $e^{ia\hat{p}} \approx 1 + ia\hat{p} + \cdots$, the above also means that

$$\hat{H}\hat{p} = \hat{p}\hat{H}, \quad [\hat{H}, \hat{p}] = 0 \tag{A8}$$

This means that $\hat{p}$ is a constant of motion as shown in Chapter 5, Appendix C, or

$$\frac{d\langle\hat{p}\rangle}{dt} = 0 \tag{A9}$$

In other words, momentum is conserved in a system where the Hamiltonian is translational invariant.

B   Infinitesimal Rotation

Similarly, the L_z operator is $-i\frac{\partial}{\partial\phi}$ which is a generator of rotation,[7]

$$f(\phi + \alpha) = e^{\alpha\frac{d}{d\phi}}f(\phi) = e^{i\alpha\hat{L}_z}f(\phi) \tag{B1}$$

Going through the derivation as we have before, for a Hamiltonian that is rotationally symmetric about the z axis, then it commutes with $e^{i\alpha\hat{L}_z}$. In this case

$$[e^{i\alpha\hat{L}_z}, \hat{H}] = 0 \tag{B2}$$

or Taylor series expanding the above, we derive that

$$[\hat{L}_z, \hat{H}] = 0 \tag{B3}$$

Hence, the $\hat{z}$ component of the angular momentum is conserved for a system that is rotationally symmetric about the z axis.

The above has been motivated by $\hat{L}_z$ that follows from the orbital angular momentum, whose form has been motivated by the classical angular momentum. What if the angular momentum has no classical analogue like the spin angular momentum? Or the state vector

---

[7]Assume again that $\hbar = 1$.

may not be written using wavefunctions at all. In this case, we can postulate a generalized generator of rotation of the form

$$e^{i\alpha\hat{J}_z}|j\rangle \tag{B4}$$

The above will take the angular momentum state $|j\rangle$ and generate rotated state about the z axis. We postulate the above form for three reasons:

1. The generator of rotation for orbital angular momentum is already of this form. It must be the special case of the above form.

2. If the above generator commutes with the Hamiltonian with similar rotational symmetry, this component of angular momentum will be conserved.

3. The rotation will cause the expectation value of the angular momentum vector, an observable, to rotate according to a coordinate rotation about the z axis.

It is easy to see that the expectation value of the $\hat{J}_z$ operator remains unchanged under this rotation. We can easily show that

$$\langle j|e^{-i\alpha\hat{J}_z}\hat{J}_z e^{i\alpha\hat{J}_z}|j\rangle = \langle j|\hat{J}_z|j\rangle \tag{B5}$$

In the above, functions of operator $\hat{J}_z$ commute with each other, a fact that can be easily proved by expanding the functions as a power series. The conjugate transpose of $e^{i\alpha\hat{J}_z}$ is $e^{-i\alpha\hat{J}_z}$ since $\hat{J}_z$ is Hermitian because it represents an observable. Hence, the last equality in (B5) follows.

C   Derivation of Commutation Relations

In general, we can define the angular momentum operator to be

$$\hat{\mathbf{J}} = \mathbf{i}\hat{J}_x + \mathbf{j}\hat{J}_y + \mathbf{k}\hat{J}_z \tag{C1}$$

The above is an observable so are the components of the operator in the x, y, z directions. Hence, the expectation value of the above with respect to the state $|j\rangle$ gives rise to

$$\langle\hat{\mathbf{J}}\rangle = \mathbf{i}\langle\hat{J}_x\rangle + \mathbf{j}\langle\hat{J}_y\rangle + \mathbf{k}\langle\hat{J}_z\rangle = \mathbf{i}J_x + \mathbf{j}J_y + \mathbf{k}J_z = \mathbf{J} \tag{C2}$$

where we denote the expectation values of $\hat{J}_i$ with scalar number J_i, and that of $\hat{\mathbf{J}}$ with $\mathbf{J}$. We will test point 3 above with respect to this vector $\mathbf{J}$. This vector will have to rotate according to coordinate rotation as the state vector is rotated according to (B4).

If $e^{i\alpha\hat{J}_z}$ is a generator of rotation about the z axis, it will leave J_z unchanged as shown above. But it will not leave J_x and J_y components unchanged. Now, we can find the expectation value of $\hat{J}_x$ under rotation about z axis, or that

$$J_x = \langle j|\hat{J}_x|j\rangle, \qquad J_x' = \langle j|e^{-i\alpha\hat{J}_z}\hat{J}_x e^{i\alpha\hat{J}_z}|j\rangle \tag{C3}$$

Before rotation, this would have represented the expectation values of the x component of $\mathbf{J}$, the angular momentum. After rotation,

$$\mathbf{J} = \mathbf{i}' J_x' + \mathbf{j}' J_y' + \mathbf{k} J_z \tag{C4}$$

and J_x' will contain both the x and y component of $\mathbf{J}$. When we rotate the state vector by angle α, the expectation value of the vector $\mathbf{J}$ will rotate by α in the same direction. But from the Figure C1, it is clear that

$$J_x' = J_x \cos\alpha + J_y \sin\alpha \tag{C5}$$

The equivalence of rotations in (C3) and (C5) can be proved by lengthy algebra. To simplify the algebra, we can show the equivalence for infinitesimal rotation. To this end, we assume that α is small so that we keep only the first two terms of the Taylor expansion for $e^{i\alpha\hat{J}_z}$ or $e^{i\alpha\hat{J}_z} \approx 1 + i\alpha\hat{J}_z$. Then,

$$J_x' = \langle j | \hat{J}_x - i\alpha\hat{J}_z\hat{J}_x + i\alpha\hat{J}_x\hat{J}_z + \cdots | j \rangle \tag{C6}$$

For small α, (C5) becomes

$$J_x' \cong J_x + \alpha J_y \tag{C7}$$

But (C6) can be written as

$$J_x' = \langle j | \hat{J}_x | j \rangle - i\alpha\langle j | \hat{J}_z\hat{J}_x - \hat{J}_x\hat{J}_z | j \rangle \tag{C8}$$
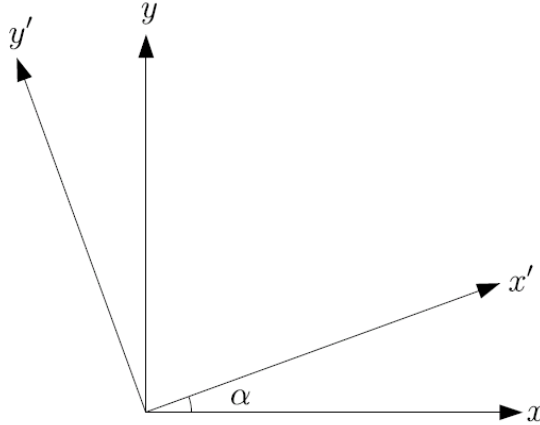
Figure C1: Coordinate rotation of the xy plane about the z axis by angle α.

Comparing (C7) and (C8), we have

$$\hat{J}_z\hat{J}_x - \hat{J}_x\hat{J}_z = i\hat{J}_y \tag{C9}$$

We can let $\hbar \neq 1$ to arrive back at the previously postulated commutator relations. The above is arrived at using only rotational symmetry argument that the angular momentum operator is a generator of rotation. The other commutation relations for angular operators can be derived by similar arguments. In general, for all operators that represent angular momentum, we have

$$\left[\hat{J}_x, \hat{J}_y\right] = i\hbar\hat{J}_z, \qquad \left[\hat{J}_y, \hat{J}_z\right] = i\hbar\hat{J}_x, \qquad \left[\hat{J}_z, \hat{J}_x\right] = i\hbar\hat{J}_y \tag{C10}$$

From the above, we can define raising and lowering operators and use the ladder approach to derive the properties of the eigenstates of angular momentum operators.

Chapter 9

Spin

9.1 Introduction

Spin is a special property of atomic or subatomic particles that has no classical analogue. Electron has spin. We can think of it as being due to the self spinning of the electron, but we should not let our imagination run further than that. Spin of an electron gives it a spin angular momentum in addition to the orbital angular momentum it possesses. The spin also endows an electron with a magnetic dipole moment that causes it to interact with a magnetic field.

   The spin of some particles is found to have binary values of "spin up" and "spin down" experimentally by the famous Stern-Gerlach experiment in 1922 by Otto Stern (1888-1969) and Walter Gerlach (1889-1979). This binary nature, as we shall see, fits nicely in the mathematical structure of angular momentum in quantum mechanics, but it cannot be described by a wavefunction or wave mechanics. Instead, it can be represented by matrix mechanics.

9.2 Spin Operators

We have seen that the z component of the orbital angular momentum, represented by the operator $\hat{L}_z$, is quantized to be $m\hbar$ where $-l \leqslant m \leqslant l$, l being an integer related to the total angular momentum square operator $\hat{L}^2$ with eigenvalue $l(l+1)\hbar^2$.

   It can be shown that the relationship between the total angular momentum number l and the z-component of the angular number m is not restricted to orbital angular momenta. It can be established for all quantum mechanical angular momenta, as is shown in Appendix A. A generalized angular momentum can be $\mathbf{J} = \mathbf{L} + \mathbf{S}$, $\mathbf{J} = \mathbf{L}_1 + \mathbf{L}_2$, or $\mathbf{J} = \mathbf{S}_1 + \mathbf{S}_2$. Like linear momentum, angular momentum can be added like vectors. Also, they are conserved quantities. In the quantum world, these observables will be elevated to become operators.

   A more general framework for angular momentum is that for $\hat{J}^2 = \hat{J}_x^2 + \hat{J}_y^2 + \hat{J}_z^2$, an operator that represents the square of the total angular momentum, and $\hat{J}_x$, $\hat{J}_y$, $\hat{J}_z$, operators that represent the x, y, and z components of angular momenta, then

$$\hat{J}^2 \left| J, M \right\rangle = J\left(J+1\right)\hbar^2 \left| J, M \right\rangle \tag{9.2.1}$$

$$\hat{J}_z \left| J, M \right\rangle = M\hbar^2 \left| J, M \right\rangle, \quad -J \leqslant M \leqslant J \tag{9.2.2}$$

We have proved the above results for orbital angular momentum by using wave mechanics and wavefunctions, but they can be proven for general angular momentum by using rotational symmetry of 3D coordinate space, and mathematics of raising and lowering operators (see Exercise at the end of the Chapter).

Spin angular momentum operators also fit under the framework of general angular momentum operator, and can be thought of as a special case of the above framework. Hence, the above can also be applied to spin $\frac{1}{2}$ angular momentum where $J = \frac{1}{2}$, $-\frac{1}{2} \leqslant M \leqslant \frac{1}{2}$. The values of M has to be spaced by value one apart, hence, $M = \pm\frac{1}{2}$.

For spins, we let $\hat{S}$ represent the total angular momentum operator, while $\hat{S}_z$ represents the z component of the spin angular momentum. Applying (9.2.1) and (9.2.2) to spin $\frac{1}{2}$ angular momentum, we have

$$\hat{S}^2 \left| \tfrac{1}{2}, \pm\tfrac{1}{2} \right\rangle = \tfrac{3}{4}\hbar^2 \left| \tfrac{1}{2}, \pm\tfrac{1}{2} \right\rangle, \quad J = \tfrac{1}{2}, \ M = \pm\tfrac{1}{2} \tag{9.2.3}$$

$$\hat{S}_z \left| \tfrac{1}{2}, \pm\tfrac{1}{2} \right\rangle = \pm\tfrac{1}{2}\hbar \left| \tfrac{1}{2}, \pm\tfrac{1}{2} \right\rangle \tag{9.2.4}$$

As a result, the corresponding z component of the spin angular momentum, represented by the operator $\hat{S}_z$, has only two eigenvalues and two eigenstates: an up state with angular momentum of $\frac{1}{2}\hbar$ and a down state with angular momentum of $-\frac{1}{2}\hbar$.[1]

The corresponding x and y components of the spin angular momentum can be represented by operators $\hat{S}_x$ and $\hat{S}_y$. Together with $\hat{S}_z$, they satisfy the following commutation relations,

$$\left[\hat{S}_x, \hat{S}_y\right] = i\hbar\hat{S}_z, \qquad \left[\hat{S}_y, \hat{S}_z\right] = i\hbar\hat{S}_x, \qquad \left[\hat{S}_z, \hat{S}_x\right] = i\hbar\hat{S}_y \tag{9.2.5}$$

The above is similar to the commutation relations satisfied by $\hat{L}_x$, $\hat{L}_y$, and $\hat{L}_z$, where they have been motivated by wave mechanics. However, as has been shown in Appendix A, that if an operator is to represent an angular momentum, then their x, y, and z components have to satisfy the above commutation relations by rotational symmetry of the 3D coordinate space.

It is expedient at this point to define Pauli spin matrices, introduced by Wolfgang Pauli (1900-1958), and given by

$$\hat{\sigma}_x = \frac{2}{\hbar}\hat{S}_x, \qquad \hat{\sigma}_y = \frac{2}{\hbar}\hat{S}_y, \qquad \hat{\sigma}_z = \frac{2}{\hbar}\hat{S}_z \tag{9.2.6}$$

with the commutation relation

$$[\hat{\sigma}_x, \hat{\sigma}_y] = 2i\hat{\sigma}_z, \qquad [\hat{\sigma}_y, \hat{\sigma}_z] = 2i\hat{\sigma}_x, \qquad [\hat{\sigma}_z, \hat{\sigma}_x] = 2i\hat{\sigma}_y \tag{9.2.7}$$

Since there are only two spin states, they can be represented by a two dimensional column vector. Hence, we have

$$\left|\uparrow\right\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \qquad \left|\downarrow\right\rangle = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \tag{9.2.8}$$

---

[1] This was discovered in silver atoms by the famous Stern-Gerlach experiment.

The spin operators can in turn be represented by 2×2 matrices. Then, the matrix representation of $\hat{\sigma}_z$ is

$$\hat{\sigma}_z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \tag{9.2.9}$$

such that

$$\hat{\sigma}_z \left| \uparrow \right\rangle = \left| \uparrow \right\rangle, \qquad \hat{\sigma}_z \left| \downarrow \right\rangle = - \left| \downarrow \right\rangle \tag{9.2.10}$$

The raising and lowering Pauli matrices can be easily constructed as

$$\hat{\sigma}_+ = \begin{bmatrix} 0 & 2 \\ 0 & 0 \end{bmatrix}, \qquad \hat{\sigma}_- = \begin{bmatrix} 0 & 0 \\ 2 & 0 \end{bmatrix} \tag{9.2.11}$$

$$\frac{1}{2}\hat{\sigma}_+ \left| \downarrow \right\rangle = \left| \uparrow \right\rangle, \qquad \frac{1}{2}\hat{\sigma}_- \left| \uparrow \right\rangle = \left| \downarrow \right\rangle \tag{9.2.12}$$

By using the the relationship that $\hat{\sigma}_\pm = \hat{\sigma}_x \pm i\hat{\sigma}_y$, the other Pauli matrices can be found to be

$$\hat{\sigma}_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \qquad \hat{\sigma}_y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \tag{9.2.13}$$

It can be shown that any 2×2 matrix can be expanded in terms of the Pauli matrices and the identity matrix; namely,

$$\overline{\mathbf{A}} = a_1\hat{\sigma}_x + a_2\hat{\sigma}_y + a_3\hat{\sigma}_z + a_4\overline{\mathbf{I}}$$

Furthermore, it can be shown that

$$\hat{\sigma}_x^2 = \hat{\sigma}_y^2 = \hat{\sigma}_z^2 = \overline{\mathbf{I}} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{9.2.14}$$

Hence, we can define a $\hat{\sigma}^2$ operator such that

$$\hat{\sigma}^2 = \hat{\sigma}_x^2 + \hat{\sigma}_y^2 + \hat{\sigma}_z^2 = 3\overline{\mathbf{I}} \tag{9.2.15}$$

By the same token, the $\hat{S}^2$ operator, from (9.2.6) and the above, can be evaluated to be

$$\hat{S}^2 = \hat{S}_x^2 + \hat{S}_y^2 + \hat{S}_z^2 = \frac{\hbar^2}{4}\hat{\sigma}^2 = \frac{3}{4}\hbar\overline{\mathbf{I}} \tag{9.2.16}$$

It can be shown that the eigenvector of the $\hat{\sigma}_x$ operator is

$$\left| x \right\rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \frac{1}{\sqrt{2}} [\left| \uparrow \right\rangle + \left| \downarrow \right\rangle]$$

while that for the $\hat{\sigma}_y$ operator is

$$\left| y \right\rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ i \end{bmatrix} = \frac{1}{\sqrt{2}} [\left| \uparrow \right\rangle + i\left| \downarrow \right\rangle]$$

It indicates that the state vector that represents a spin pointing in the x or y direction is a linear supposition of an up spin state vector and a down spin state vector with different phases.

In the above, the quantum mechanics of spin half particle such as an electron is described by matrix mechanics.

9.3   The Bloch Sphere

We have seen that the eigenstates of $\hat{\sigma}_z$ are

$$|\uparrow\rangle = \left[\begin{array}{c} 1 \\ 0 \end{array} \right], \quad |\downarrow\rangle = \left[\begin{array}{c} 0 \\ 1 \end{array} \right] \tag{9.3.1}$$

In the two dimensional space in which spin states are represented, these two states are orthonormal and complete. Hence, the eigenstates of the $\hat{\sigma}_x$ and $\hat{\sigma}_y$ operators can be expanded as well in terms of $|\uparrow\rangle$ and $|\downarrow\rangle$ states. Consequently, a general spin state can be written as

$$|s\rangle = a_\uparrow |\uparrow\rangle + a_\downarrow |\downarrow\rangle = \cos\left(\frac{\theta}{2}\right) |\uparrow\rangle + e^{i\phi} \sin\left(\frac{\theta}{2}\right) |\downarrow\rangle \tag{9.3.2}$$

We have chosen $a_\uparrow$ and $a_\downarrow$ judiciously so that the wavefunction is clearly normalized, namely, $|a_\uparrow|^2 + |a_\downarrow|^2 = 1$. The relative phase between $a_\uparrow$ and $a_\downarrow$ is in $e^{i\phi}$ as absolute phase is unimportant. Let

$$\boldsymbol{\sigma} = \boldsymbol{\imath}\hat{\sigma}_x + \boldsymbol{\jmath}\hat{\sigma}_y + \mathbf{k}\hat{\sigma}_z \tag{9.3.3}$$

be an operator that represents a vector of the spin angular momentum, where $\boldsymbol{\imath}$, $\boldsymbol{\jmath}$, and $\mathbf{k}$ are unit vectors in the x, y, and z directions. Then the expectation value of this operator for a given quantum state $|s\rangle$ in (9.3.2) should point in a vectorial direction. This is a basic tenet of quantum mechanics since $\boldsymbol{\sigma}$ represents angular momentum, which is an observable. Consequently, we find the vector

$$\mathbf{P}_s = \langle s|\boldsymbol{\sigma}|s\rangle = \boldsymbol{\imath}\langle s|\hat{\sigma}_x|s\rangle + \boldsymbol{\jmath}\langle s|\hat{\sigma}_y|s\rangle + \mathbf{k}\langle s|\hat{\sigma}_z|s\rangle \tag{9.3.4}$$

After substituting $|s\rangle$ from (9.3.2) into the above, we can show that

$$\mathbf{P}_s = \boldsymbol{\imath}\sin\theta\cos\phi + \boldsymbol{\jmath}\sin\theta\sin\phi + \mathbf{k}\cos\theta \tag{9.3.5}$$

The vector $\mathbf{P}_s$ maps out a sphere called the Bloch sphere[2] with corresponding θ and ϕ values as shown in Figure 9.1.

The above analysis shows that by superposing the up and down spin states judiciously, the observed spin vector, which is the expectation of the vector spin operator, can be made to point in any directions in the 3D space. This feature has been of great interest in constructing a quantum computer and storing quantum information.

---

[2]Named after Felix Bloch (1905-1983).
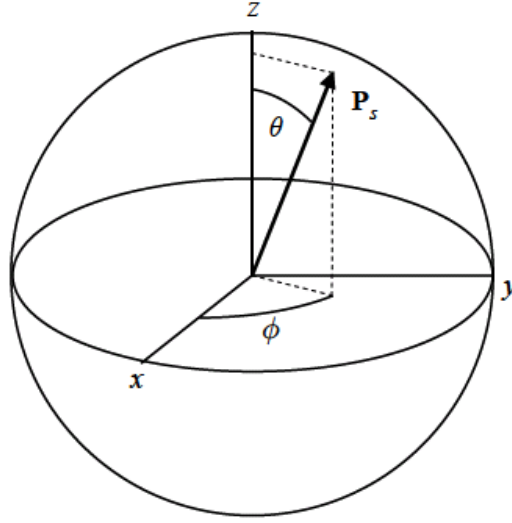
Figure 9.1: The Bloch sphere showing the vector $\mathbf{P}_s$.

9.4   Spinor

In general, we need a two component vector to describe both the wavefunction and the spin state of an electron, namely,

$$|\psi\rangle = \left[\begin{array}{c} |\psi_{up}\rangle \\ |\psi_{down}\rangle \end{array} \right] = |\psi_{up}\rangle \left[\begin{array}{c} 1 \\ 0 \end{array} \right] + |\psi_{down}\rangle \left[\begin{array}{c} 0 \\ 1 \end{array} \right] \tag{9.4.1}$$

These wave functions with two components are known as spinors. The vector space for spinors consists of the vector space for the electron wave function multiplying with the two-dimensional vector space for the spin. This is known as a direct product space.

9.5   Pauli Equation

Before deriving the Pauli equation, it is useful to derive the Hamiltonian for the Zeeman effect using heuristic arguments. The Zeeman effect is due to the coupling of the magnetic field to the orbital angular momentum while the Pauli equation accounts of the coupling of the magnetic field to the spins.

The classical angular momentum of a particle moving in a circle is

$$L = m_0 vr \tag{9.5.1}$$

which can also be written as a vector

$$\mathbf{L} = \mathbf{r} \times \mathbf{p} = \mathbf{r} \times m_0 \mathbf{v} \tag{9.5.2}$$

If this particle is an electron, it will complete an orbit in $2\pi r/v$ time, and we can denote the current due to this orbiting electron as

$$I = -\frac{ev}{2\pi r} \tag{9.5.3}$$

A circulating current loop produces a magnetic dipole moment

$$\boldsymbol{\mu}_d = I\mathbf{a} \tag{9.5.4}$$

where $|\mathbf{a}|$ is the area of the current loop, the vector $\mathbf{a}$ points in the direction normal to the surface $|\mathbf{a}|$. For circulating electron, this moment is

$$|\boldsymbol{\mu}_d| = -\frac{ev}{2\pi r}\pi r^2 = -\frac{evr}{2} \tag{9.5.5}$$

We can then write

$$\boldsymbol{\mu}_d = -\frac{e\mathbf{r} \times \mathbf{v}}{2} = -\frac{e\mathbf{L}}{2m_0} \tag{9.5.6}$$

The interaction energy of a magnetic dipole and a background magnetic field $\mathbf{B}$ is given by

$$E_u = -\boldsymbol{\mu}_d \cdot \mathbf{B} \tag{9.5.7}$$

Assume that $\mathbf{B} = \hat{z}B_z$, then

$$E_u = \frac{e}{2m_0}B_z L_z = \frac{e}{2m_0}B_z m\hbar \tag{9.5.8}$$

since the eigenvalue of $\hat{L}_z$ is $m\hbar$ and the expectation value of $\hat{L}_z$ can be precisely $m\hbar$. Equation (9.5.8) can also be written as

$$E_u = m\mu_B B_z \tag{9.5.9}$$

where μ_B is a Bohr magneton defined to be

$$\mu_B = \frac{e\hbar}{2m_0} \tag{9.5.10}$$

The Hamiltonian of quantum mechanics is inspired by the classical Hamiltonian mechanics. The Hamiltonian represents the total energy of the system. If an applied external magnetic field contribution to a small change in the total energy, we can add a perturbation Hamiltonian in accordance to (9.5.8) and let

$$\hat{H}_p = \frac{eB_z}{2m_0}\hat{L}_z \tag{9.5.11}$$

Originally, for angular momentum with number l, the z component of the angular momentum is $m\hbar$ with $(2l+1)$ values for m that are degenerate. The applied B field will split this degeneracy into $2l+1$ distinct energy values. This is known as the Zeeman effect.

Equation (9.5.11) can be written as

$$\hat{H}_p = \frac{e}{2m_0} \hat{\mathbf{L}} \cdot \mathbf{B} \tag{9.5.12}$$

When the spin angular momentum is added, the perturbing Hamiltonian becomes

$$\hat{H}_p = \frac{e}{2m_0} (\hat{\mathbf{L}} + g\hat{\mathbf{S}}) \cdot \mathbf{B} \tag{9.5.13}$$

where g is the g-factor which is approximately 2. It can be determined from Dirac's equation, but quantum electrodynamic correction has determined this factor to extraordinary precision.[3]

In general, the Hamiltonian for an electron in the presence of a magnetic field is

$$\hat{H} = \frac{1}{2m_0} \hat{\mathbf{p}}^2 + V_0 + \frac{e}{2m_0} (\hat{\mathbf{L}} + g\hat{\mathbf{S}}) \cdot \mathbf{B} \tag{9.5.14}$$

The above Hamiltonian is again motivated by the Hamiltonian for classical mechanics for an electron in the presence of an electromagnetic field, which is

$$\hat{H} = \frac{1}{2m_0} \left[\hat{\mathbf{p}} + e\mathbf{A}(\mathbf{r}) \right]^2 + V_0(\mathbf{r}) + \frac{e}{2m_0} g\hat{\mathbf{S}} \cdot \mathbf{B} \tag{9.5.15}$$

The above is also known as the Pauli equation. The Hamiltonian of the Pauli equation can also be expressed as

$$\hat{H} = \frac{1}{2m_0} \left(\hat{\boldsymbol{\sigma}} \cdot (\hat{\mathbf{p}} + e\mathbf{A}) \right)^2 + V_0(\mathbf{r}) \tag{9.5.16}$$

9.5.1   Splitting of Degenerate Energy Level

The interaction of spins with magnetic field can cause the degenerate energy levels of a quantum system to split. If no magnetic field is present, the spin up state and the spin down state can share the same wave function with the same energy level, and hence, they are degenerate. In the presence of a magnetic field, the state where the spin is parallel with the magnetic field is of lower energy compared to the case where the spin is antiparallel to the magnetic field. This causes the splitting of a degenerat energy level as shown in Figure 9.1. This fact can be used to design interesting spin devices.

9.6   Spintronics

Some materials are highly magnetic because there are many unpaired electrons in these materials. These unpaired electrons give these materials magnetic dipole moments, such as

---

[3]In 1985, Richard P. Feynman (1918-1988) wrote in "QED, The Strange Theory of Light and Matter," that this factor has been determined to the precision of a few hair lines if one were to err from the travel from New York to Los Angeles. In 2012, Daniel F. Styer (1955-) wrote that the precision of this constant has been improved to a few hair lines from the distance of the Earth to the Moon.
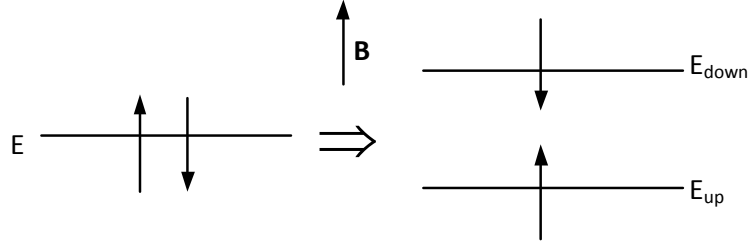
Figure 9.1: The splitting of degenerate energy level when the spin of the electron interacts with an ambient magnetic field.

ferromagnetic materials. In ferromagnets, the magnetic dipoles of the same orientation cluster together to form microscopic domains that are random on a macroscopic scale. These domains can be aligned macroscopically by magnetization, making these materials into magnets. Fe, Co, Ni and their alloys have this property.

The conduction property of these materials can also be affected by remnant magnetic field in the domain, or externally applied magnetic field. This can be explained by the DOS diagram. The electrons in a magnetic material see a self field that affects their energy levels: electrons that are aligned with the magnetic field are in the lower energy states compared to the anti-parallel electrons. Hence the DOS of a ferromagnets is different for up and down spins compared to normal metals.
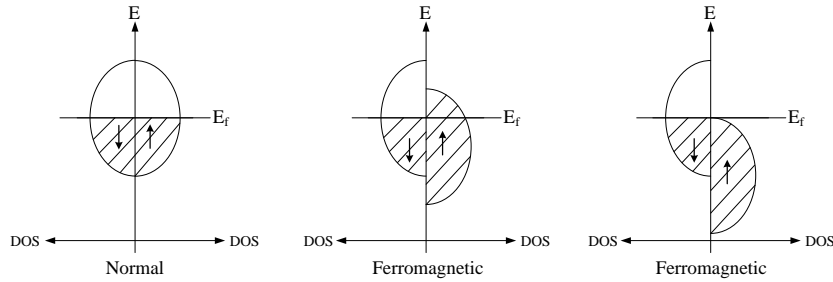
Figure 9.1: The shifting of the energy levels of the up and down spins states in the DOS plot due to different ambient or external magnetic field. The left half is for DOS of down spins and conversely for the right half.

Now, the conduction electrons available for up spins do not exist, but occur in abundance for the down spins. (The skewness in DOS can be further affected by an externally applied magnetic field.) This gives rise to conductivity that is spin dependent, or spin polarized

transport. Also, electrons that pass through a ferromagnets become spin polarized. Devices of high resistance and low resistance can be made by sandwiching three layers together of a normal metal between two ferromagnets layers as shown in Figure 9.2. Such phenomenon is known as giant magneto-resistance (GMR).

Moreover, the external magnetic field can alter the resistance of the device using it to sense magnetic field or store information in magnetic field. Hence, such devices can be used as magnetic read head or non-volatile memory as shown in Figure 9.4.

Exercise 2

In the Appendix, we have derived the commutation relation for general angular momentum in equation (B.3.9). Use the same technique, derive the other commutation relations for angular momentum in (B.3.10).

1. The raising and lowering operators in angular momentum are defined as

$$J_{\pm} = J_x \pm i J_y$$

Using the commutation relations in (B.3.10), show that

$$J_{\mp} J_{\pm} = J_x^2 + J_y^2 \pm i[J_x, J_y]$$

$$\mathbf{J}^2 = J_x^2 + J_y^2 + J_z^2 = J_{\mp} J_{\pm} + J_z^2 \pm \hbar J_z$$

$$[J_{\pm}, J_z] = \mp \hbar J_{\pm}$$

$$[\mathbf{J}^2, J_z] = 0$$

$$[\mathbf{J}^2, J_{\pm}] = 0$$

The above means that $J_{\pm}$ shares the same eigenfunctions with $\mathbf{J}^2$ but not with J_z. However, $\mathbf{J}^2$ shares the same set of eigenfunctions with J_z.

2. Denoting the eigenfunctions of $\mathbf{J}^2$ and J_z as

$$|J, M\rangle$$

where

$$\mathbf{J}^2 |J, M\rangle = J(J+1)\hbar^2 |J, M\rangle$$

$$J_z |J, M\rangle = \hbar M |J, M\rangle$$

and J and M are integers. Show that

$$J_z J_{\pm} |J, M\rangle = \hbar(M \pm 1) J_{\pm} |J, M\rangle$$

3. By using the fact that

$$\langle J, M | J_{\mp} J_{\pm} | J, M\rangle \geq 0$$

show that

$$-J \leq M \leq J$$

The wonder of the above is that it could be derived from commutation relations alone, and the explicit form of $|J, M\rangle$ need not be known.
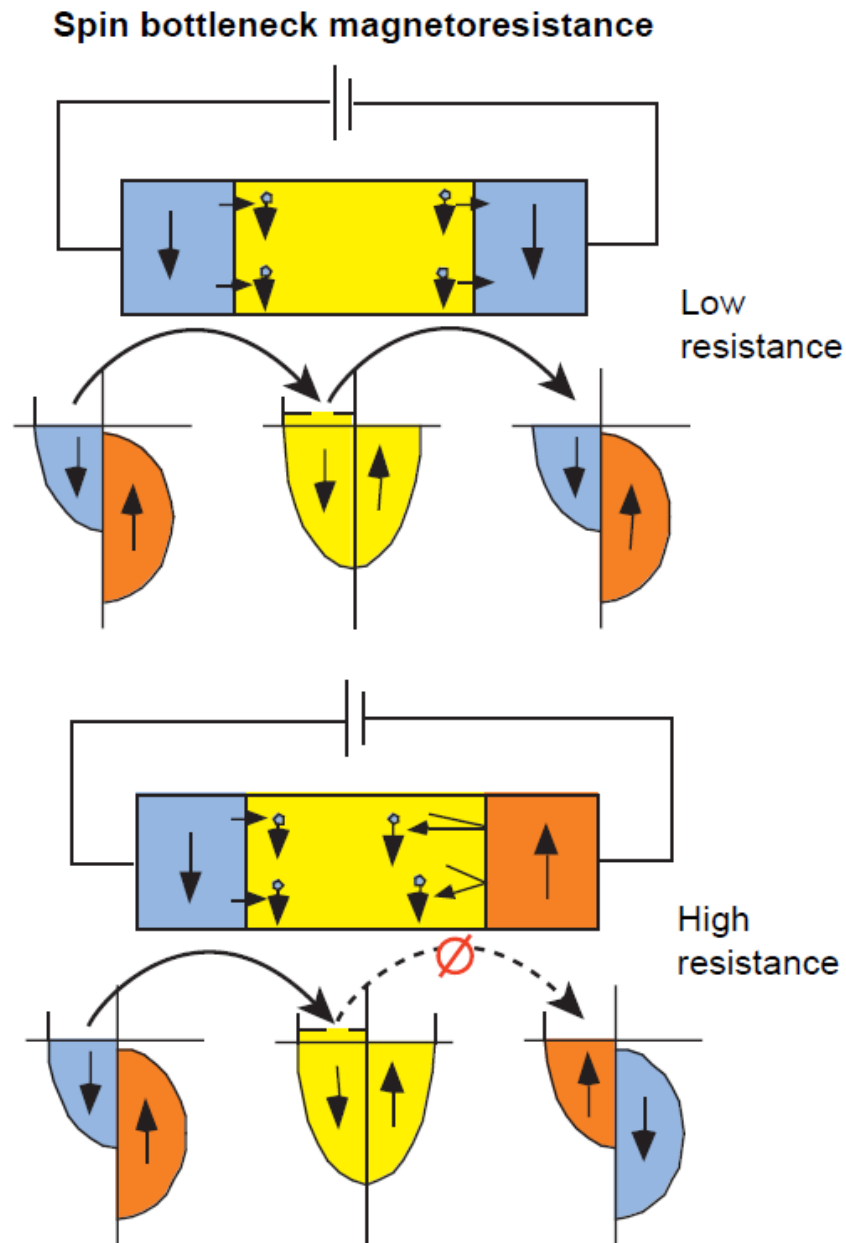
Figure 9.2: The sandwiching of a normal metal between two ferromagnets can give rise to high or low resistance depending on their relative orientations (from G.A. Prinz, Science, v. 282, 1660, 1998).

High resistance
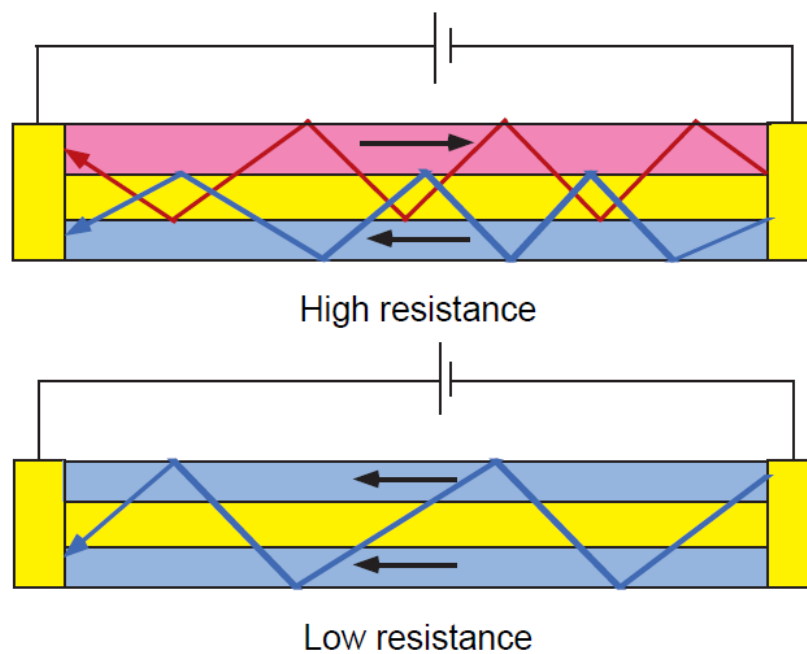
Low resistance

Figure 9.3: Schematic representations of channel transport that is parallel to the plane of a layered magnetic metal sandwich structure for aligned (low resistance) and antialigned (high resistance) orientation. The resistance of the channel can be altered by changing the magnetization of the sandwiching layers (from G.A. Prinz, Science, v. 282, 1660, 1998).
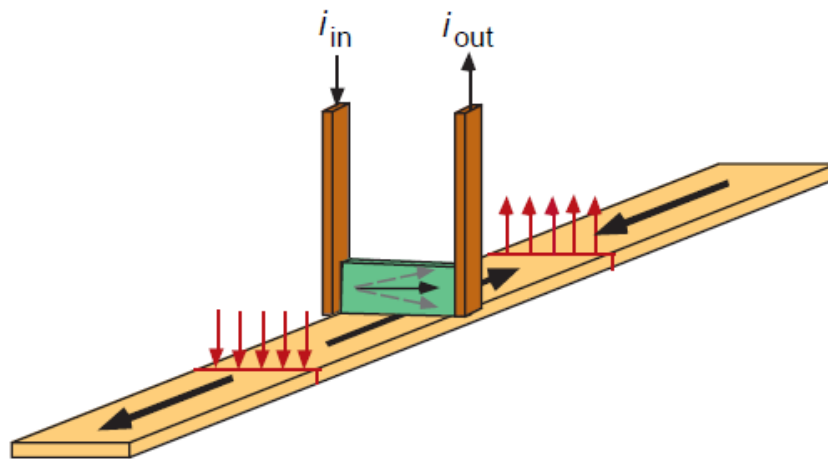
Figure 9.4: When a GMR read head passes over magnetic domains of different polarization in a storage media, the leakage field in between the domains can be detected, as the resistance of the GMR is affected by such leakage field (Gary A. Prinz, Science v. 282, 1660, 1998).

Chapter 10

Identical Particles*

10.1 Introduction

The physics of identical particles is another intellectual triumph in quantum mechanics where the phenomenon is not observed in the classical world. When particles are identical or indistinguishable, their wavefunctions have to assume a certain form to satisfy a certain symmetry when the positions of the two particles are swapped. In quantum mechanics, it is the magnitude squared of a wavefunction that has physical meaning. If the particles are indistinguishable, the magnitude squared of a wavefunction does not change when the two particles are swapped.

The many-particle problem cannot be solved in closed form, but we can use a variational method such as the Rayleigh-Ritz method to find an approximate eigenvalue of the system as described in Section 6.3.1. To this end, we have to construct basis functions for identical particles. These basis functions can be constructed from the single-particle solutions when the particles are non-interacting with and isolated from each other.

First, we can consider the particles to be non-interacting and isolated. If there are two non-interacting particles, we can write down two independent equations for them, namely

$$\hat{H}_1\psi_a(\mathbf{r}_1) = \left[-\frac{\hbar^2}{2m}\nabla_1^2 + V(\mathbf{r}_1)\right]\psi_a(\mathbf{r}_1) = E_a\psi_a(\mathbf{r}_1) \qquad (10.1.1)$$

$$\hat{H}_2\psi_b(\mathbf{r}_2) = \left[-\frac{\hbar^2}{2m}\nabla_2^2 + V(\mathbf{r}_2)\right]\psi_b(\mathbf{r}_2) = E_b\psi_b(\mathbf{r}_2) \qquad (10.1.2)$$

In the above, the two particles share the same potential function V, but are in different worlds. Alternatively, we can write down the equation for the entire quantum system which is the sum of the two quantum systems, namely

$$(\hat{H}_1 + \hat{H}_2)\psi_a(\mathbf{r}_1)\psi_b(\mathbf{r}_2) = (E_a + E_b)\psi_a(\mathbf{r}_1)\psi_b(\mathbf{r}_2) \qquad (10.1.3)$$

We use a direct product space consisting of $\psi_a(\mathbf{r}_1)\psi_b(\mathbf{r}_2)$ to expand the eigenfunction of the system. This is alright if the particles are non-interacting, but if they are interacting, the

---

*This chapter can be skipped on first reading.

141

issue of identical particles and their indistinguishability emerges. The wave function has to be constructed with caution.

In general, we can write the general wavefunction for two particles as

$$\psi_{tp}(\mathbf{r}_1, \mathbf{r}_2) = c_{12}\psi_a(\mathbf{r}_1)\psi_b(\mathbf{r}_2) + c_{21}\psi_a(\mathbf{r}_2)\psi_b(\mathbf{r}_1) \tag{10.1.4}$$

where $\psi_\eta(\mathbf{r}_i)$ is a one-particle wavefunction for a particle in state η described by coordinate $\mathbf{r}_i$.

But when $\mathbf{r}_1 \rightleftharpoons \mathbf{r}_2$, the assumption is that the magnitude squared of the wavefunction remains unchanged, since these particles are indistinguishable. Namely,

$$|\psi_{tp}(\mathbf{r}_1, \mathbf{r}_2)|^2 = |\psi_{tp}(\mathbf{r}_2, \mathbf{r}_1)|^2 \tag{10.1.5}$$

which means that

$$\psi_{tp}(\mathbf{r}_1, \mathbf{r}_2) = \gamma\psi_{tp}(\mathbf{r}_2, \mathbf{r}_1) \tag{10.1.6}$$

where $\gamma = e^{i\theta}$. Using (10.1.4) in (10.1.6), we have

$$c_{12}\psi_a(\mathbf{r}_1)\psi_b(\mathbf{r}_2) + c_{21}\psi_a(\mathbf{r}_2)\psi_b(\mathbf{r}_1) = \gamma\left[c_{12}\psi_a(\mathbf{r}_2)\psi_b(\mathbf{r}_1) + c_{21}\psi_a(\mathbf{r}_1)\psi_b(\mathbf{r}_2)\right] \tag{10.1.7}$$

From the above, we get

$$c_{21} = \gamma c_{12}, \qquad c_{12} = \gamma c_{21} \tag{10.1.8}$$

Consequently,

$$c_{12} = \gamma c_{21} = \gamma^2 c_{12} \tag{10.1.9}$$

or $\gamma^2 = 1$, $\gamma = \pm 1$, $c_{21} = \pm c_{12}$, and

$$\psi_{tp}(\mathbf{r}_1, \mathbf{r}_2) = c_{12}\left[\psi_a(\mathbf{r}_1)\psi_b(\mathbf{r}_2) \pm \psi_a(\mathbf{r}_2)\psi_b(\mathbf{r}_1)\right] \tag{10.1.10}$$

When the "+" sign is chosen in (10.1.10), it corresponds to the wavefunction of two identical bosons. When the "−" sign is chosen, it corresponds to the wavefunction of two identical fermions.

10.2 Pauli Exclusion Principle

For fermions, the two-particle wavefunction is given by

$$\psi_{tp}(\mathbf{r}_1, \mathbf{r}_2) = c\left[\psi_a(\mathbf{r}_1)\psi_b(\mathbf{r}_2) - \psi_a(\mathbf{r}_2)\psi_b(\mathbf{r}_1)\right] \tag{10.2.1}$$

when $a = b$, the above vanishes, meaning that two fermions cannot be in the same state. The above vanishes when $\mathbf{r}_1 = \mathbf{r}_2$, meaning that two fermions cannot be in the same position simultaneously. This known as the Pauli exclusion principle. No two identical electrons can be in the same state nor the same position simultaneously. However, two electrons with different spins, up spin and down spin, are considered different, and they can be in the same state, like in the same orbital in an atom. Also, the sign of the total wavefunction, $\psi_{tp}(\mathbf{r}_1, \mathbf{r}_2)$, changes sign when the positions of the two particles are swapped.

10.3 Exchange Energy

When we have two electrons, each will see the Coulomb potential of the other electron. Hence, the Hamiltonian is given by

$$\hat{H} = -\frac{\hbar^2}{2m_e} \left(\nabla_1^2 + \nabla_2^2 + V_1 + V_2\right) + \frac{e^2}{4\pi\epsilon_0 |\mathbf{r}_1 - \mathbf{r}_2|} = \hat{H}_1 + \hat{H}_2 + \frac{e^2}{4\pi\epsilon_0 |\mathbf{r}_1 - \mathbf{r}_2|} \quad (10.3.1)$$

where $V_1 = V(\mathbf{r}_1)$, $V_2 = V(\mathbf{r}_2)$ and ∇_i^2 is the Laplacian operator expressed in the i-th coordinate. The two particles are in the same environment and their wavefunctions can overlap. The two-particle fermion wavefunction can be written as

$$|\psi_{tp}\rangle = \frac{1}{\sqrt{2}} \left[|1, a\rangle |2, b\rangle - |2, a\rangle |1, b\rangle\right] \quad (10.3.2)$$

where $\frac{1}{\sqrt{2}}$ ensures that $|\psi_{tp}\rangle$ is also normalized. In the above, $|i, \eta\rangle$ is the Hilbert space representation of the wavefunction $\psi_\eta(\mathbf{r}_i)$. The above does not solve the Schrödinger equation corresponding to the above Hamiltonian, but it can be used to estimate the eigenvalue or the energy of the Hamiltonian in the spirit of the variational method described previously, using the Rayleigh quotient.

To find the corresponding expectation value of the energy, we evaluate

$$\langle E \rangle = \langle \psi_{tp}| \hat{H} |\psi_{tp}\rangle \quad (10.3.3)$$

Since the function is normalized, the above is also the Rayleigh quotient for the energy. When expanded to four terms, gives

$$\langle E \rangle = \frac{1}{2} \Big[\langle 1, a| \langle 2, b| \hat{H} |1, a\rangle |2, b\rangle + \langle 2, a| \langle 1, b| \hat{H} |2, a\rangle |1, b\rangle$$

$$- \langle 1, a| \langle 2, b| \hat{H} |2, a\rangle |1, b\rangle - \langle 2, a| \langle 1, b| \hat{H} |1, a\rangle |2, b\rangle \Big] \quad (10.3.4)$$

The first two terms above are equal to each other. Specifically,

$$\langle 1, a| \langle 2, b| \hat{H} |1, a\rangle |2, b\rangle$$
$$= \langle 1, a| \langle 2, b| \hat{H}_1 |1, a\rangle |2, b\rangle + \langle 1, a| \langle 2, b| \hat{H}_2 |1, a\rangle |2, b\rangle$$
$$+ \langle 1, a| \langle 2, b| \frac{e^2}{4\pi\epsilon_0 |\mathbf{r}_1 - \mathbf{r}_2|} |1, a\rangle |2, b\rangle$$
$$= \langle 1, a| \hat{H}_1 |1, a\rangle + \langle 2, b| \hat{H}_2 |2, b\rangle$$
$$+ \langle 1, a| \langle 2, b| \frac{e^2}{4\pi\epsilon_0 |\mathbf{r}_1 - \mathbf{r}_2|} |1, a\rangle |2, b\rangle$$
$$= E_a + E_b + E_{IE,ab} \quad (10.3.5)$$

E_a is the energy of particle 1 in state a, while E_b is the energy of particle 2 in state b. The $E_{IE,ab}$ term is the interaction energy that can be derived explicitly to yield

$$E_{IE,ab} = e^2 \int d\mathbf{r} d\mathbf{r}'' \frac{|\psi_a(\mathbf{r})|^2 \; |\psi_b(\mathbf{r}')|^2}{4\pi\epsilon_0 |\mathbf{r} - \mathbf{r}'|} \quad (10.3.6)$$

This is exactly the same as the potential energy from the second term on the right-hand side of (10.3.4) if we were to write it explicitly. Consequently, we have

$$E_a + E_b + E_{IE,ab} = \frac{1}{2} \left[\langle 1,a| \langle 2,b| \hat{H} |1,a\rangle |2,b\rangle + \langle 2,a| \langle 1,b| \hat{H} |2,a\rangle |1,b\rangle \right] \qquad (10.3.7)$$

However, the third and fourth terms on the right-hand side of (10.3.4) have no classical analog. They are there because of the presence of the second term in (10.1.10) be it for bosons or fermions. The second term is needed because the particles are indistinguishable or identical if we exchange their positions.

Due to the Hermiticity of the Hamiltonian operator, the fourth term on the right-hand side of (10.3.4) is the complex conjugate of the third term. Consequently, these two terms can be written as

$$E_{EX,ab} = -\frac{1}{2} \left[\langle 1,a| \langle 2,b| \hat{H} |2,a\rangle |1,b\rangle + \left(\langle 1,a| \langle 2,b| \hat{H} |2,a\rangle |1,b\rangle \right)^* \right] \qquad (10.3.8)$$

$$= -\Re e \left[\int d\mathbf{r} d\mathbf{r}' \psi_a^* (\mathbf{r}) \psi_b^* (\mathbf{r}') \hat{H} \psi_a (\mathbf{r}') \psi_b (\mathbf{r}) \right] \qquad (10.3.9)$$

The exchange energy is proportional to the overlap between the wavefunctions $\psi_a (\mathbf{r})$ and $\psi_b (\mathbf{r})$. This term is non-zero even if the Coulomb interaction is absent, but the wavefunctions overlap and are non-orthogonal. When their wavefunctions overlap, the issue of indistinguishability arises, and the exchange term can be non-zero even for non-interacting particles.

10.4   Extension to More Than Two Particles

In the following, we will discuss methods to construct basis functions for different kinds of particles. These basis functions does not solve the Schrödinger equation yet, but they have to satisfy certain symmetry conditions depending on the kind of particles they represent. The basis functions can be used in the Rayleigh-Ritz procedure to estimate the energy of the system. We can use the eigenstates of the isolated particles to construct basis functions.

1. **Non-identical Particle Case:**

   Let us assume that we have N particles, and M modes to fit this N particles. We can construct a state for non-identical particles that looks like

   $$|\psi_{\text{diff}}\rangle = |1,a\rangle |2,b\rangle |3,c\rangle ... |N,n\rangle \qquad (10.4.1)$$

   In terms of basis function, we may express the above as

   $$|\psi_{ab\cdots n}\rangle = |1,a\rangle |2,b\rangle |3,c\rangle \cdots |N,n\rangle \qquad (10.4.2)$$

   or

   $$\psi_{ab\cdots n}(\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_N) = \psi_a(\mathbf{r}_1)\psi_b(\mathbf{r}_2) \cdots \psi_n(\mathbf{r}_N) \qquad (10.4.3)$$

   We can fit the N particles in M modes, and these M modes can be repeating or non-repeating. In other words, $a, b, c, \ldots, n$ are chosen from M eigenstates with energy

values $E_1, E_2, E_3, \ldots, E_M$ with possibility of repetition. For non-repeating case, it is necessary for $M \geqslant N$.

However, the above wavefunction cannot be used for bosons and fermions, as we will get a new wavefunction when we swap the positions of two particles. But bosons and fermions are indistinguishable particles. We will consider them separately.

2. **Boson Case:**

For the N boson particle case, we can write the legitimate wavefunction, which can be used as a basis function, as

$$|\psi_{\text{identical-bosons}}\rangle \propto \sum_{\hat{P}} \hat{P}|1, a\rangle|2, b\rangle|3, c\rangle \cdots |N, n\rangle \qquad (10.4.4)$$

where $\hat{P}$ is a permutation operator, and the above summation is over all possible permutations of the coordinate $\mathbf{r}_i$ over the one-particle eigenstates $a, b, c, \cdots, n$. Repetition of these energy eigenstates is allowed since more than one particle can be placed in one energy state. The above wavefunction remains unchange when we permute the positions of two particles, because for every $|1, a\rangle \cdots |i, l\rangle \cdots |j, p\rangle \cdots |N, n\rangle$, there is a $|1, a\rangle \cdots |j, l\rangle \cdots |i, p\rangle \cdots |N, n\rangle$ in the above summation. Hence, swapping of i and j will not change the sign of the above wavefunction. The above can also be written as a basis function as

$$|\psi_{ab\cdots n}\rangle \propto \sum_{\hat{P}} \hat{P}|1, a\rangle|2, b\rangle|3, c\rangle \cdots |N, n\rangle \qquad (10.4.5)$$

3. **Fermion Case:**

For the N fermion case, we can write the wavefunction, which can be used as a basis function, as

$$|\psi_{\text{identical-fermion}}\rangle = \frac{1}{\sqrt{N!}} \sum_{\hat{P}=1}^{N!} \pm\hat{P}|1, a\rangle|2, b\rangle|3, c\rangle \cdots |N, n\rangle \qquad (10.4.6)$$

where the "+" sign is chosen for even permutation while the "−" sign is chosen for odd permutation. A permutation involves a unique pairwise exchange of two particles . The permutation is even or odd depending on the number of pairwise exchanges that have taken place.

Therefore, given a term $|1, a\rangle \cdots |i, l\rangle \cdots |j, p\rangle \cdots |N, n\rangle$, there always exists another term: $-|1, a\rangle \cdots |j, l\rangle \cdots |i, p\rangle \cdots |N, n\rangle$ in the above summation since they differ by one permutation. If $i = j$, implying that $\mathbf{r}_i = \mathbf{r}_j$, the two terms cancel each other implying that they cannot be in the same position. Likewise all the terms in the sum cancel each other since every term that contains i and j can be paired up with every other terms in the sum. Moreover, If $l = p$, all terms in the summation above cancel as well implying that they cannot be in the same mode or state. Therefore, the above is a legitimate basis function that represents the fermions as it obeys Pauli exclusion principle. Also, there is a sign change when the position of two particles are swapped.

Another way of writing the fermion case is

$$
|\psi_{\text{identical-fermion}}\rangle = \frac{1}{\sqrt{N!}}
\begin{vmatrix}
|1,a\rangle & |2,a\rangle & \cdots & |N,a\rangle \\
|1,b\rangle & |2,b\rangle & \cdots & |N,b\rangle) \\
\vdots & \vdots & \ddots & \vdots \\
|1,n\rangle & |2,n\rangle & \cdots & |N,n\rangle
\end{vmatrix}
\tag{10.4.7}
$$

The above is known as the Slater determinant. When two particles are in the same position, two columns are the same, and the above determinant is zero. When two states are the same, two rows are the same, and the determinant of the above is again zero, implying that these two cases are not allowed. Also when two columns are swapped, the sign of the determinant changes, because it corresponds to two particles exchanging positions. In terms of basis function, we can express the above as

$$
\begin{aligned}
|\psi_{ab\cdots n}\rangle &= \frac{1}{\sqrt{N!}} \sum_{\hat{P}=1}^{N!} \pm \hat{P}|1,a\rangle|2,b\rangle|3,c\rangle\cdots|N,n\rangle \\
&= \frac{1}{\sqrt{N!}}
\begin{vmatrix}
|1,a\rangle & |2,a\rangle & \cdots & |N,a\rangle \\
|1,b\rangle & |2,b\rangle & \cdots & |N,b\rangle) \\
\vdots & \vdots & \ddots & \vdots \\
|1,n\rangle & |2,n\rangle & \cdots & |N,n\rangle
\end{vmatrix}
\end{aligned}
\tag{10.4.8}
$$

10.5 Counting the Number of Basis States

Usually, the one particle eigenfucntion has infinitely many possible states. In the finite basis method, we may want to choose a subset of this eigenstates, say M states. At this point, if we have N particles and M states to put this N particles in, it may be prudent to count how many possible basis states there are. Such counting scheme will also be used in the appendix to derive the thermal distribution functions.

1. **Non-identical Particles:**

   The first particle has M state to choose from in (10.4.2), the second particle also has M states to choose from, and eventually, there are M^N states possible for (10.4.2), since repetition of the states is allowed.

   $$
   N_{basis} = M^N
   \tag{10.5.1}
   $$

2. **Identical Bosons:**

   In this case, the number of particles that can be put into one state is not limited. Since we need to permute the state (10.4.2) to arrive at new states to be linear superposed in the manner of (10.4.5), the ordering of the particles within the states is not important.

   We will first derive the number of permutations when ordering is important. One can imaging the M states to be equivalent to M bins. The first particle has M possible bins to fit into. However, once the particle is put into one of the bins, since ordering

is important, repetition is allowed, there are two ways to put in the second particle in the same bin, before or after the first particle. In other words, the first particle splits the bin into two compartments, allowing two ways to put two particles in the same bin. Together with the other $M-1$ bins, there are $M+1$ ways to insert the second particle into the bins. In other words, adding a new particle always adds a new compartment for the next particle to fit in. By the same token, the third particle has $M+2$ ways to fit into the bins etc. Finally the number of possible ways is

$$M(M+1)(M+2)...(M+N-1) \tag{10.5.2}$$

One can verify that the above gives the correct answer of $N!$ if only one bin is available for N particles. Since ordering is unimportant, we have

$$N_{basis} = \frac{(M+N-1)!}{N!(M-1)!} \tag{10.5.3}$$

3. **Identical Fermions:**

For fermions, the first particle has M states to choose from. Since each state can only admit one fermion particle, it is necessary that $M \geq N$. The second particle has $M-1$ states to choose from, since repetition is not allowed. Also, ordering is not important since all permutations are used in the summation (10.4.8). Consequently,

$$N_{basis} = \frac{M!}{N!(M-N)!} \tag{10.5.4}$$

10.6   Examples

1. **Non-identical Particles:**

Say if we have two electrons with different spins, the distinct states are

$$|1,a\rangle|2,a\rangle, |1,b\rangle|2,b\rangle, |1,a\rangle|2,b\rangle, |1,b\rangle|2,a\rangle \tag{10.6.1}$$

The above is in agreement with $M=2$, $N=2$, $M^N = 2^2 = 4$.

2. **Identical Bosons:**

Consider the ^4He (helium four) atoms which are bosons. Then the possible boson states are

$$|1,a\rangle|2,a\rangle, |1,b\rangle|2,b\rangle, \frac{1}{\sqrt{2}}(|1,a\rangle|2,b\rangle + |1,b\rangle|2,a\rangle) \tag{10.6.2}$$

Again this is in agreement with $M=2$, $N=2$, $\frac{(M+N-1)!}{N!(M-1)!} = \frac{3!}{2!1!} = 3$.

3. **Identical Fermions:**

If we have two identical fermions, the only state is

$$\frac{1}{\sqrt{2}}(|1,a\rangle|2,b\rangle - |1,b\rangle|2,a\rangle) \tag{10.6.3}$$

Again this is in agreement with $M=2$, $N=2$, $\frac{M!}{N!(M-N)!} = \frac{2!}{2!0!} = 1$.

In the above simple examples, we see that for the boson case, there are more states with particles in the same state compared to the other cases. In the boson case, there are two out of three states where the particles are in the same state (see (10.6.2)). In the non-identical particle case, there are two out of four such states. In the fermion case, there is none. This is indicative of the fact that bosons like to cluster together.

10.7   Thermal Distribution Functions

There are some important thermal distribution functions in understanding semiconductor physics. These are the Maxwell-Boltzmann distribution, the Fermi-Dirac distribution, and the Bose-Einstein distribution. These distribution functions are derived via statistical mechanics, and the results will just quoted here. The first-principles derivations are given in Appendix 10.7.

1. **Maxwell-Boltzmann Distribution:**

   For this distribution, the number of particles $N(E)$ in a given energy state is given by

   $$N(E) = e^{\frac{-(E-\mu)}{k_B T}} \tag{10.7.1}$$

   where μ is the chemical potential, k_B is the Boltzmann constant, and T is temperature in Kelvin.

2. **Fermi-Dirac Distribution:**

   This is given by

   $$N(E) = \frac{1}{1 + e^{\frac{E-\mu}{k_B T}}} \tag{10.7.2}$$

   This is the most important distribution function for semiconductors. The chemical potential μ is loosely called the Fermi level. To be precise, μ is the same as the Fermi level at absolute zero temperature. When T is zero, the distribution function looks like as shown in Figure 10.1. The electrons are frozen into the ground state for $E < \mu$. When $T > 0$, some of the electrons from the states where $E < \mu$ are dislodged by thermal agitation into states where $E > \mu$. This phenomenon is more pronounced as T increases. This distribution also explains the physical character of semiconductors.

   In semiconductors, the Fermi level is midway in between the valence band and the conduction band. When $T = 0$, all the electrons are frozen in the valence band, and the semiconductor cannot conduct electricity, as there are no free electrons in the conduction band. When $T > 0$, some electrons in the valence band are dislodged into the conduction band. This gives rise to electrons in the conduction band, and holes in the valence band. They contribute to the flow of electric current and the semiconductor starts to conduct. The conductivity of the semi-conductor material increases with increasing temperature.

3. **Bose-Einstein Distribution:**
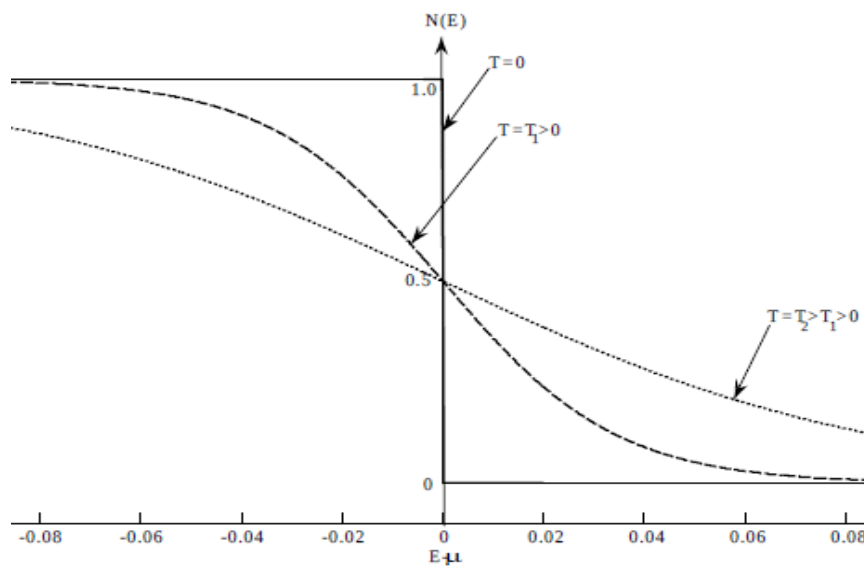
Figure 10.1: Fermi-Dirac Distribution for changing temperature T.

The Bose-Einstein distribution is for bosons, and it is given by

$$N(E) = \frac{1}{e^{(E-\mu)/(k_B T)} - 1} \tag{10.7.3}$$

This distribution has a divergence when $E = \mu$. It also reflects the fact that bosons like to cluster together. When the temperature is low, they condense to around the chemical potential μ.
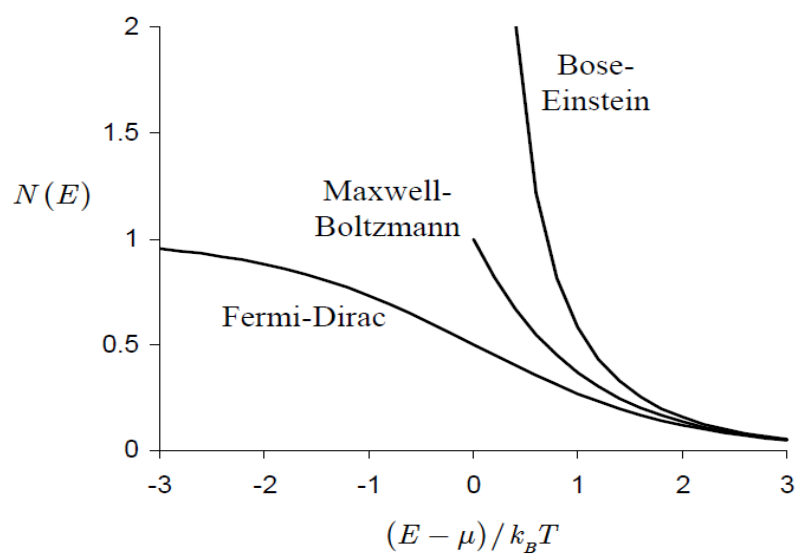
Figure 10.2: Comparison of different thermal distribution functions. For high energy states, they are similar to each other (from DAB Miller).

Appendix

A   Quantum Statistical Mechanics

We would like to derive Maxwell-Boltzmann, Fermi-Dirac, and Bose-Einstein distributions. It is based on statistical mechanics that assumes that all identical energy levels have equal likelihood of being occupied. Only the counting is based on quantum mechanics.

We consider a quantum system with energy levels E_1, E_2, E_3, $\cdots$. Each level E_i could have degeneracy of d_i, $i = 1$, 2, 3, $\cdots$, and is filled by N_i, $i = 1$, 2, 3, $\cdots$ particles. When the system is coupled to its environment, energy will be exchanged with its environment via collision, vibration, and radiation. At thermal equilibrium, there is no net energy flowing into and out of the system. We expect that

$$E = \sum_{n=1}^{\infty} E_n N_n \tag{A1}$$

to be a constant due to energy conservation. Also we expect the number of particles to remain constant, or that

$$N = \sum_{n=1}^{\infty} N_n \tag{A2}$$

due to particle conservation.

The particles will acquire energy from their environment due to energy exchange, and distribute themselves across the energy levels subject to the above constraints, but also according to the availability of energy levels. Energy levels with higher degeneracies are more likely to be filled. We will consider three different cases: (i) when the particles are distinguishable, (ii) when they are identical fermions, and (iii) when they are identical bosons.

A1   Distinguishable Particles

We will find the configuration function $Q(N_1, N_2, N_3, \cdots)$ whose value equals the number of ways that there are N_1 particles in E_1, N_2 particles in E_2, N_3 particles in E_3 and so on. In order to fill energy level E_1 with N_1 particles, the number of ways is

$$\binom{N}{N_1} = \frac{N!}{N_1! \, (N - N_1)!} \tag{A3}$$

If E_1 has degeneracy d_1, all the degenerate levels are equally likely to be filled. Then there are $d_1^{N_1}$ ways that the N_1 particles can go into E_1 level (see Section 10.5). Hence, the number of ways that E_1 can be filled is

$$Q_{E_1} = \frac{N! d_1^{N_1}}{N_1! \, (N - N_1)!} \tag{A4}$$
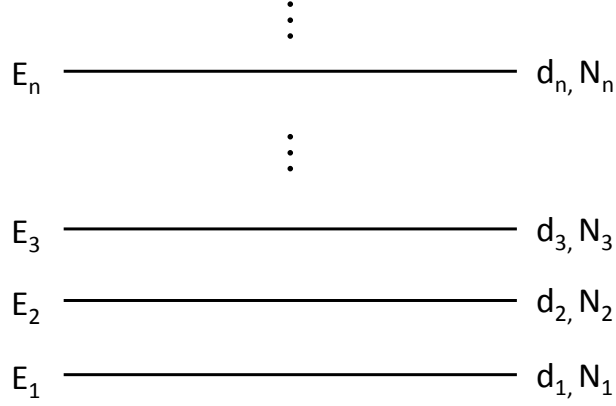
Figure A1: Particles are distributed among different energy levels according to quantum statistics. An energy level is denoted with E_i with degeneracy d_i and N_i particles occupying it.

The number of ways that E_2 can be filled is

$$Q_{E_2} = \frac{(N - N_1)! d_2^{N_2}}{N_2! (N - N_1 - N_2)!} \tag{A5}$$

Then the total number of ways

$$
\begin{aligned}
Q(N_1, N_2, N_3, \cdots) &= Q_{E_1} Q_{E_2} Q_{E_3} \cdots \\
&= \frac{N! d_1^{N_1}}{N_1! (N - N_1)!} \frac{(N - N_1)! d_2^{N_2}}{N_2! (N - N_1 - N_2)!} \frac{(N - N_1 - N_2)! d_3^{N_3}}{N_3! (N - N_1 - N_2 - N_3)!} \cdots \\
&= N! \frac{d_1^{N_1} d_2^{N_2} d_3^{N_3} \cdots}{N_1! N_2! N_3! \cdots} = N! \prod_{n=1}^{\infty} \frac{d_n^{N_n}}{N_n!}
\end{aligned} \tag{A6}
$$

Notice that the above is independent of the order in which the levels are filled, as expected. The above can also be interpreted in a different light: There are $d_n^{N_n}$ ways that N_n particles can fit into the E_n energy level. But order is unimportant and a division by $N_n!$ is necessary. But there are $N!$ ways that these distinguishable particles can be selected in an ordered way, and hence, a prefactor of $N!$ is needed in the above.

A2   Identical Fermions

Due to Pauli exclusion principle, each energy level can admit only one particle. In this case, the number of ways energy E_n can be filled is (see Section 10.5)

$$Q_{E_n} = \frac{d_n!}{N_n! \, (d_n - N_n)!}, \ \ d_n \geq N_n \tag{A7}$$

(The above has value 1 when $d_n = 1$.) The total number of ways is then

$$Q\left(N_1, N_2, N_3, ...\right) = \prod_{n=1}^{\infty} \frac{d_n!}{N_n! \, (d_n - N_n)!} \tag{A8}$$

A3   Identical Bosons

For bosons, repetition for filling a given state is allowed, but the particles are indistinguishable from each other. Then when N_n particles are picked to fill the d_n degenerate E_n levels, the first particle from N_n particles has d_n slots to fill in, while the second particle has $d_n + 1$ slots to fill in, and the third particle has $d_n + 2$ slots to fill in and so on, since repetition is allowed. That is the new particle can take the position of the old particle as well. So the number of ways that E_n with d_n degeneracy can be filled is (see Section 10.5)

$$Q_{E_n} = \frac{d_n \, (d_n + 1) \, (d_n + 2) \, ... \, (d_n + N_n - 1)}{N_n!} \tag{A9}$$

$$= \frac{(N_n + d_n - 1)!}{N_n! \, (d_n - 1)!} \tag{A10}$$

Therefore

$$Q\left(N_1, N_2, N_3, ...\right) = \prod_{n=1}^{\infty} \frac{(N_n + d_n - 1)!}{N_n! \, (d_n - 1)!} \tag{A11}$$

Unlike the distinguishable particle case, no prefactor of $N!$ is needed for the identical particle case, since the order with which they are selected is unimportant.

B   Most Probable Configuration

A way of filling the energy levels $\{E_1, E_2, E_3, ...\}$ with $\{N_1, N_2, N_3, ...\}$ is called a configuration. The most likely configuration to be filled is the one with the largest $Q\left(N_1, N_2, N_3, ...\right)$. At statistical equilibrium, the system will gravitate toward this configuration. Hence, to find this configuration, we need to maximize Q with respect to different $\{N_1, N_2, N_3, ...\}$ subject to the energy conservation constraint (A1) and particle conservation constraint (A2).

We use the Lagrange multiplier technique to find the optimal Q subject to constraints (A1) and (A2). We define a function

$$G = \ln Q + \alpha \left[N - \sum_{n=1}^{\infty} N_n \right] + \beta \left[E - \sum_{n=1}^{\infty} N_n E_n \right] \tag{B1}$$

The optimal Q value is obtained by solving

$$\frac{\partial G}{\partial N_n} = 0, \; n = 1, 2, 3, \cdots, \quad \frac{\partial G}{\partial \alpha} = \left[N - \sum_{n=1}^{\infty} N_n \right] = 0, \quad \frac{\partial G}{\partial \beta} = \left[E - \sum_{n=1}^{\infty} N_n E_n \right] = 0 \quad \text{(B2)}$$

Notice that the constraint conditions for particle and energy conservations are imposed in the second and third equations above.

B1   Distinguishable Particles

In this case,

$$G = \ln N! + \sum_{n=1}^{\infty} \left[N_n \ln d_n - \ln N_n! \right] \tag{B3}$$

$$+ \alpha \left[N - \sum_{n=1}^{\infty} N_n \right] + \beta \left[E - \sum_{n=1}^{\infty} N_n E_n \right] \tag{B4}$$

We use Stirling's formula that

$$\ln (z!) \approx z \ln z - z \tag{B5}$$

Consequently,

$$G = \sum_{n=1}^{\infty} \left[N_n \ln d_n - N_n \ln N_n + N_n - \alpha N_n - \beta E_n N_n \right] \tag{B6}$$

$$+ \ln N! + \alpha N + \beta E \tag{B7}$$

From the above

$$\frac{\partial G}{\partial N_n} = \ln d_n - \ln N_n - \alpha - \beta E_n = 0 \tag{B8}$$

As a result, we have

$$N_n = d_n e^{-(\alpha + \beta E_n)} \tag{B9}$$

The above is the precursor to the Maxwell-Boltzmann distribution from first principles.

B2   Identical Fermions

In this case

$$G = \sum_{n=1}^{\infty} \left[\ln(d_n!) - \ln(N_n!) - \ln((d_n - N_n)!) \right] \tag{B10}$$

$$+ \alpha \left[N - \sum_{n=1}^{\infty} N_n \right] + \beta \left[E - \sum_{n=1}^{\infty} N_n E_n \right] \tag{B11}$$

After applying Stirling's formula to N_n dependent terms, we have

$$G = \sum_{n=1}^{\infty} [\ln d_n! - N_n \ln N_n + N_n - (d_n - N_n) \ln(d_n - N_n) + (d_n - N_n) - \alpha N_n - \beta N_n E_n] \tag{B12}$$

$$+ \alpha N + \beta E \tag{B13}$$

From the above,

$$\frac{\partial G}{\partial N_n} = -\ln N_n + \ln(d_n - N_n) - \alpha - \beta E_n = 0 \tag{B14}$$

or

$$N_n = \frac{d_n}{1 + e^{\alpha + \beta E_n}} \tag{B15}$$

In the above derivation, we have assumed that d_n is large so that N_n is large since $N_n \leq d_n$ in (A7). The above is the precursor to the Fermi-Dirac distribution.

B3   Identical Bosons

In this case

$$G = \sum_{n=1}^{\infty} [\ln((N_n + d_n - 1)!) - \ln(N_n!) - \ln((d_n - 1)!)] \tag{B16}$$

$$+ \alpha \left[N - \sum_{n=1}^{\infty} N_n \right] + \beta \left[E - \sum_{n=1}^{\infty} N_n E_n \right] \tag{B17}$$

With Stirling's approximation to the N_n dependent terms,

$$G = \sum_{n=1}^{\infty} [(N_n + d_n - 1) \ln(N_n + d_n - 1) - (N_n + d_n - 1) - N_n \ln N_n + N_n \tag{B18}$$

$$- \ln((d_n - 1)!) - \alpha N_n - \beta N_n E_n] + \alpha N + \beta E \tag{B19}$$

Then

$$\frac{\partial G}{\partial N_n} = \ln(N_n + d_n - 1) - \ln(N_n) - \alpha - \beta E_n = 0 \tag{B20}$$

yielding

$$N_n = \frac{d_n - 1}{e^{\alpha + \beta E_n} - 1} \simeq \frac{d_n}{e^{\alpha + \beta E_n} - 1} \tag{B21}$$

where we assume that $d_n \gg 1$. The above is the precursor to the Bose-Einstein distribution.

C   The Meaning of α and β

We will apply the above to a simple quantum system in order to infer what α and β should be. To find them, we need to solve the constraint equations (A1) and (A2) that follow from the optimization of (B1). To this end, we need to find the degeneracy per energy state, d_k of a quantum system.

We consider N electrons inside a bulk material describable by a single electron in an effective mass approximation. In such a case, the kinetic energy of the electron is described by

$$E_k = \frac{\hbar^2 k^2}{2m} \tag{C1}$$

If we assume periodic boundary conditions on a box of lengths L_x, L_y, and L_z, then

$$k^2 = \left(\frac{2n_x \pi}{L_x}\right)^2 + \left(\frac{2n_y \pi}{L_y}\right)^2 + \left(\frac{2n_z \pi}{L_z}\right)^2 \tag{C2}$$

In the k space, there is a state associated with a unit box of volume

$$\Delta V_k = \frac{(2\pi)^3}{L_x L_y L_z} = \frac{8\pi^3}{V} \tag{C3}$$

where V is the volume of the box of bulk material. In k space, in a spherical shell of thickness Δk, the number of states in the neighborhood of k is

$$d_k = \frac{4\pi k^2 \Delta k}{\Delta V_k} = \frac{V}{2\pi^2} k^2 \Delta k \tag{C4}$$

Then the total number of particles is, using (B9) for Maxwell-Boltzmann,

$$N = \sum_k d_k e^{-(\alpha + \beta E_k)} = \frac{V}{2\pi^2} e^{-\alpha} \sum_k k^2 \Delta k e^{-\beta E_k}$$
$$= \frac{V}{2\pi^2} e^{-\alpha} \int_0^\infty e^{-\beta \hbar^2 k^2/(2m)} k^2 dk \tag{C5}$$

By using the fact that

$$I_1 = \int_0^\infty e^{-Ak^2} k^2 dk = -\frac{d}{dA} \int_0^\infty e^{-Ak^2} dk = -\frac{d}{dA} \frac{1}{2}\sqrt{\frac{\pi}{A}} = \frac{\sqrt{\pi}}{4} A^{-\frac{3}{2}} \tag{C6}$$

the above integrates to

$$N = V e^{-\alpha} \left(\frac{m}{2\pi\beta\hbar^2}\right)^{\frac{3}{2}} \tag{C7}$$

Similarly, the total E is given by

$$E = \sum_k E_k N_k = \sum_k E_k d_k e^{-(\alpha + \beta E_k)} = -\frac{d}{d\beta} \sum_k d_k e^{-(\alpha + \beta E_k)} = -\frac{d}{d\beta} N(\beta) \tag{C8}$$

It can be shown that

$$E = \frac{3N}{2\beta}$$

From the equipartition theorem of energy from statistical mechanics we know that

$$\frac{E}{N} = \frac{3}{2}k_B T \tag{C9}$$

Hence, we conclude that

$$\beta = \frac{1}{k_B T} \tag{C10}$$

It is customary to write

$$\alpha = -\frac{\mu(T)}{k_B T} \tag{C11}$$

so that for Maxwell-Boltzmann distribution,

$$N_n = d_n e^{-(E_n - \mu)/(k_B T)} \tag{C12}$$

where $\mu(T)$ is the chemical potential. When normalized with respect to degeneracy, we have

$$n_{mb} = e^{-(E_n - \mu)/(k_B T)}, \qquad \text{Maxwell-Boltzmann} \tag{C13}$$

Fermi-Dirac and Bose-Einstein distribution become Maxwell-Boltzmann when $(\beta E_n + \alpha) \gg 1$. So they have the same α and β. Therefore,

$$n_{fd} = \frac{1}{e^{(E_n - \mu)/(k_B T)} + 1}, \qquad \text{Fermi-Dirac} \tag{C14}$$

$$n_{be} = \frac{1}{e^{(E_n - \mu)/(k_B T)} - 1}, \qquad \text{Bose-Einstein} \tag{C15}$$

When $(E_n - \mu) \gg k_B T$ the above distributions resemble the Maxwell-Boltzmann distribution. This is because n becomes small per energy level, and the indistinguishability of the particles plays a less important role. When $(E_n - \mu) < 0$, the Fermi-Dirac distribution "freezes" to 1 for all energy levels below μ, since the Pauli exclusion principle only allows one fermion per level. When $(E_n - \mu) \ll k_B T$, the Bose-Einstein distribution diverges, implying that the particles condense to an energy level close to μ. This condensation is more pronounced if $k_B T$ is small.

For photons in a box, the number of photons need not be conserved as photons can freely the box than entering the box. Hence, we can set α or μ, the chemical potential, to zero. In this case, the Bose-Einstein distribution becomes the Planck distribution.

$$n_{pl} = \frac{1}{e^{E_n/(k_B T)} - 1}, \qquad \text{Planck} \tag{C16}$$

Chapter 11

Density Matrix

11.1 Pure and Mixed States

Given a quantum state described by

$$|\psi\rangle = a_1|\psi_1\rangle + a_2|\psi_2\rangle \tag{11.1.1}$$

which is a linear superposition of two states, the interpretation of quantum mechanics is that the particle is in a linear superposition of the two states, and one does not know what state the particle is in until the measurement is done. The measurement collapses the particle into one of the two quantum states. For instance, the two states could be the spin states of an atom. The famous Stern-Gerlach experiment separates the spin states into the up state and the down state. Such a state indicated in (11.1.1) is known as a pure quantum state. The phase relationship between a_1 and a_2 is maintained precisely. When the quantum state is in such a state, we say that the two states are coherent.

The above interpretation of quantum mechanics is known as the Copenhagen school of thought, lead by Niels Bohr. The probabilistic interpretation of quantum mechanics was fraught with controversies in its early days. Many physicists, including Einstein, cannot bring themselves to term with it. "God does not play dice!" was one famous saying of Einstein. He posited that "a quantum system is already in a known state before the measurement" with the hidden variable description. However, experimental findings are with the Copenhagen school for coherent quantum states. For incoherent quantum states, they are with Einstein's position.

Nevertheless, it is difficult for a particle to be in a pure quantum state if it involves a linear superposition of many quantum states unless the quantum system is completely isolated. Coupling to the thermal bath can destroy the coherence between these states. Hence, there are states for which partial coherence still exists, and the linear superposition of quantum states de-coheres with respect to each other, before the linear superposition of states becomes entirely a mixed state. Coherence lifetime is used to measure when a linear superposition of quantum states, which starts out being a coherent sum, becomes an incoherent superposition.

However, there are states where the particle has already collapsed into state 1 or state 2 even before the measurement. Or these two states are entirely uncorrelated or incoherent.

The particle is either in state 1 or state 2. The probability of finding the particle in state 1 is $|a_1|^2$ while that for state 2 is $|a_2|^2$. The phase relationship between a_1 and a_2 is completely random or incoherent. States for which there is no coherence between a_1 and a_2 are known as mixed states. If we take an ensemble average of the measurement outcomes of the mixed states, they satisfy the aforementioned probability property.

The probabilistic interpretation of quantum mechanics cannot be verified by a single experiment, but an ensemble of measurements whose setups are identical to each other. An ensemble average* of an experimental outcome is the average over these outcomes over the identical ensemble of measurements. It will be prudent to find a way to describe a quantum system that is related to its ensemble average and its experimental outcome.

11.2   Density Operator

An elegant way to represent a quantum state that is either in the pure state, mixed state, or partial coherent state, is via the density operator

$$\hat{\rho} = |\psi\rangle\langle\psi| \tag{11.2.1}$$

We can easily show that the expectation of the operator $\hat{A}$ in this state is

$$\langle\hat{A}\rangle = \langle\psi|\hat{A}|\psi\rangle = \text{tr}\left(\hat{A}\hat{\rho}\right) \tag{11.2.2}$$

This has been shown in Section 5.2.3. Therefore, knowing the density operator is equivalent to knowing the quantum state of a system which is denoted by state $|\psi\rangle$. We can calculate the trace of an operator in terms of the sum of the diagonal elements of its matrix representation. Hence,

$$\text{tr}\left(\hat{A}\hat{\rho}\right) = \sum_n \langle\phi_n|\hat{A}\hat{\rho}|\phi_n\rangle \tag{11.2.3}$$

We show that the above is basis independent by inserting the identity operator $\hat{I} = \sum_m |\phi_m\rangle\langle\phi_m|$ twice in (11.2.2), to yield

$$\langle\psi|\hat{A}|\psi\rangle = \sum_{n,m} \langle\psi|\phi_n\rangle\langle\phi_n|\hat{A}|\phi_m\rangle\langle\phi_m|\psi\rangle \tag{11.2.4}$$

Since the factors in the summand are scalars, we can rearrange them to give

$$\langle\psi|\hat{A}|\psi\rangle = \sum_{n,m} \langle\phi_m|\psi\rangle\langle\psi|\phi_n\rangle\langle\phi_n|\hat{A}|\phi_m\rangle = \sum_{n,m} \rho_{mn}A_{nm} = \text{tr}\left(\overline{\boldsymbol{\rho}} \cdot \overline{\mathbf{A}}\right) \tag{11.2.5}$$

where ρ_{mn} is the matrix representation of $\hat{\rho}$ while A_{nm} is the matrix representation of $\hat{A}$. The above is clearly independent of the orthonormal basis we choose.

---

*An ensemble average of a random variable is the average of the variable over its different outcomes for the same system or identical systems.

Say if we start with (11.1.1) for a quantum state, using it in (11.2.1) will yield

$$\hat{\rho} = |a_1|^2|\psi_1\rangle\langle\psi_1| + |a_2|^2|\psi_2\rangle\langle\psi_2| + a_1 a_2^*|\psi_1\rangle\langle\psi_2| + a_2 a_1^*|\psi_2\rangle\langle\psi_1| \qquad (11.2.6)$$

Notice that only relative phases between a_1 and a_2 are needed to form the above. Absolute phase has no meaning in quantum mechanics.

For a mixed state, we assume that the off diagonal terms involving $a_1^* a_2$ and $a_2^* a_1$ will either time average or ensemble average to zero since they are incoherent.[†] Hence, we suppose that

$$\begin{aligned}\hat{\rho} &= |a_1|^2|\psi_1\rangle\langle\psi_1| + |a_2|^2|\psi_2\rangle\langle\psi_2| \\ &= p_1|\psi_1\rangle\langle\psi_1| + p_2|\psi_2\rangle\langle\psi_2| \end{aligned} \qquad (11.2.7)$$

The form (11.2.7) can be thought of as having been ensemble averaged or time averaged. For a general mixed state we can further write the density operator as

$$\hat{\rho} = \sum_j p_j|\psi_j\rangle\langle\psi_j| \qquad (11.2.8)$$

where $|\psi_j\rangle$ are pure quantum states but not necessarily stationary states nor orthogonal, and p_j is the probability of finding the quantum system in state $|\psi_j\rangle$. Hence, $\sum_j p_j = 1$.

When we find the expectation value of an operator $\hat{A}$ with a mixed state, it can be thought of as having been ensemble averaged. For example,

$$\begin{aligned}\langle\hat{A}\rangle = \text{tr}(\hat{A}\hat{\rho}) &= \sum_j p_j \text{tr}(\hat{A}|\psi_j\rangle\langle\psi_j|) \\ &= \sum_j p_j \langle\psi_j|\hat{A}|\psi_j\rangle \end{aligned} \qquad (11.2.9)$$

The trace of the matrix representation of an operator is independent of the basis states chosen to represent the operator as explained in Chapter 5. It will be interesting to see how the density operator changes under the change of basis state. To this end, we let

$$|\psi_j\rangle = \sum_n c_n^{(j)}(t)|\phi_n\rangle \qquad (11.2.10)$$

When the quantum state is a function of time, but the relevant operator is time independent, this picture is known as the Schrödinger picture of quantum mechanics. The basis state $|\phi_n\rangle$ however, can be time independent. Using this in (11.2.8), we have

$$\begin{aligned}\hat{\rho} &= \sum_j p_j \sum_{n,m} c_n^{(j)}|\phi_n\rangle\langle\phi_m|(c_m^{(j)})^* \\ &= \sum_{n,m} |\phi_n\rangle\langle\phi_m| \sum_j p_j c_n^{(j)}(c_m^{(j)})^* \end{aligned} \qquad (11.2.11)$$

[†]Statistical processes whose time average and ensemble average are equivalent to each other are known as ergodic processes.

We can find the uv element of the matrix representation of $\hat{\rho}$ as

$$\rho_{uv} = \langle \phi_u | \hat{\rho} | \phi_v \rangle = \sum_j p_j c_u^{(j)} [c_v^{(j)}]^* = \overline{c_u c_v^*} \tag{11.2.12}$$

where the overbar stands for ensemble average. From the above, it is quite clear that

$$\rho_{uv} = \rho_{vu}^* \tag{11.2.13}$$

or that the density matrix is Hermitian. The Hermiticity property of the density operator is obvious from (11.1.1) and (11.2.8). Furthermore,

$$\mathrm{tr}\,(\hat{\rho}) = \sum_u \rho_{uu} = \sum_j p_j \sum_u |c_u^{(j)}|^2 = \sum_j p_j = 1 \tag{11.2.14}$$

The second last equality follows from the normality of $|\psi_j\rangle$ in (11.2.10). Since a trace of an operator is independent of the basis state in which it is represented, the trace of the density operator is defined as

$$\mathrm{tr}(\hat{\rho}) = \sum_u \langle \phi_u | \hat{\rho} | \phi_u \rangle \tag{11.2.15}$$

where ϕ_u constitutes an orthonormal basis.

For a pure quantum state,

$$\hat{\rho}^2 = |\psi\rangle\langle\psi|\psi\rangle\langle\psi| = |\psi\rangle\langle\psi| = \hat{\rho} \tag{11.2.16}$$

Therefore,

$$\mathrm{tr}(\hat{\rho}^2) = \mathrm{tr}(\hat{\rho}) = 1 \tag{11.2.17}$$

The trace of the above is 1 is obviated by setting $\hat{A}$ to 1 in Equation (11.2.2). The above shows that the trace of a density operator is always one irrespective of it is for a pure-state or a mixed-state quantum system. But this is not true for the trace of the $\hat{\rho}^2$ operator.

For a mixed state

$$\hat{\rho}^2 = \sum_i \sum_j p_i p_j |\psi_i\rangle\langle\psi_i|\psi_j\rangle\langle\psi_j| \tag{11.2.18}$$

Taking the trace of the above, we have

$$\mathrm{tr}(\hat{\rho}^2) = \sum_n \langle \phi_n | \hat{\rho}^2 | \phi_n \rangle = \sum_n \sum_i \sum_j p_i p_j \langle \phi_n | \psi_i \rangle \langle \psi_i | \psi_j \rangle \langle \psi_j | \phi_n \rangle$$

$$= \sum_i \sum_j p_i p_j |\langle \psi_i | \psi_j \rangle|^2 \leq \sum_i \sum_j p_i p_j = \left(\sum_i p_i \right)^2 = 1 \tag{11.2.19}$$

In the above, we did not assume that $|\psi_j\rangle$ are orthogonal, but they are normalized. The inequality follows from that $|\langle\psi_i|\psi_j\rangle|^2 \leq 1$. Hence, the trace of the density operator has the property that

$$\mathrm{tr}(\hat{\rho}^2) < 1, \qquad \text{mixed state} \tag{11.2.20}$$

$$\mathrm{tr}(\hat{\rho}^2) = 1, \qquad \text{pure state} \tag{11.2.21}$$

11.3   Time Evolution of the Matrix Element of an Operator

The matrix representation of an operator is given by

$$A_{mn}(t) = \langle \psi_m(t) | \hat{A} | \psi_n(t) \rangle \tag{11.3.1}$$

Taking the time derivative of the above, assuming that $\hat{A}$ is time independent, we have

$$\partial_t A_{mn}(t) = \{\partial_t \langle \psi_m(t) | \} \hat{A} | \psi_n(t) \rangle + \langle \psi_m(t) | \hat{A} \partial_t | \psi_n(t) \rangle \tag{11.3.2}$$

Now, using the fact that

$$i\hbar \partial_t | \psi_m(t) \rangle = \hat{H} | \psi_m(t) \rangle, \qquad i\hbar \partial_t \langle \psi_m(t) | = \langle \psi_m(t) | \hat{H} \tag{11.3.3}$$

we have

$$\partial_t A_{mn}(t) = \frac{i}{\hbar} \langle \psi_m(t) | \hat{H} \hat{A} | \psi_n(t) \rangle - \frac{i}{\hbar} \langle \psi_m(t) | \hat{A} \hat{H} | \psi_n(t) \rangle$$
$$= \frac{i}{\hbar} \langle \psi_m(t) | \hat{H} \hat{A} - \hat{A} \hat{H} | \psi_n(t) \rangle \tag{11.3.4}$$

or

$$i\hbar \partial_t A_{mn}(t) = (\hat{A} \hat{H} - \hat{H} \hat{A})_{mn} = \left[\hat{A}, \hat{H} \right]_{mn} \tag{11.3.5}$$

The algebra above applies to the expectation value of $\hat{A}$ as well, as shown in Chapter 5, where

$$\langle \hat{A} \rangle = \langle \psi(t) | \hat{A} | \psi(t) \rangle \tag{11.3.6}$$

and

$$i\hbar \partial_t \langle \hat{A} \rangle = \langle \hat{A} \hat{H} - \hat{H} \hat{A} \rangle = \left\langle \left[\hat{A}, \hat{H} \right] \right\rangle \tag{11.3.7}$$

In other words, if $\hat{A}$ commutes with $\hat{H}$, its expectation value is time independent. In the above derivation, we have assumed that $\hat{A}$ is independent of time, but the functions $\psi_m(t)$ are functions of time. This is known as the Schrodinger picture in quantum mechanics.

   The time evolution of the density operator can be derived, similar to the time evolution of the expectation value of an operator. However, the density operator is in general time dependent, whereas operators considered previously are time independent; and hence, the derivation is slightly different. Given

$$\hat{\rho} = | \psi \rangle \langle \psi | \tag{11.3.8}$$

then

$$\partial_t \hat{\rho} = (\partial_t | \psi \rangle) \langle \psi | + | \psi \rangle \partial_t \langle \psi | \tag{11.3.9}$$

Using (11.3.3), we have

$$\partial_t \hat{\rho} = -\frac{i}{\hbar} \hat{H} |\psi\rangle\langle\psi| + \frac{i}{\hbar} |\psi\rangle\langle\psi| \hat{H} = \frac{i}{\hbar} \left[\hat{\rho}\hat{H} - \hat{H}\hat{\rho} \right] = \frac{i}{\hbar} \left[\hat{\rho}, \hat{H} \right] \qquad (11.3.10)$$

There is a sign difference between (11.3.7) and (11.3.10). The above holds if $\hat{\rho}$ is in a mixed state as well

$$\hat{\rho} = \sum_j p_j |\psi_j\rangle\langle\psi_j| \qquad (11.3.11)$$

This follows from the linearity of the above equations and that p_j is time-independent.

Equation (11.3.10) can be easily generalized to matrix elements when the basis functions used for seeking the matrix representation are from time-independent basis. Or the matrix representation of (11.3.10) can be directly derived. Then,

$$\partial_t \rho_{mn} = \frac{i}{\hbar} \left[\hat{\rho}\hat{H} - \hat{H}\hat{\rho} \right]_{mn} = \frac{i}{\hbar} \left[\hat{\rho}, \hat{H} \right]_{mn} \qquad (11.3.12)$$

where $A_{mn}(t) = \langle\phi_m|\hat{A}(t)|\phi_n\rangle$, and ϕ_m is from a time-independent basis. By inserting identity operator defined using the same basis set, we can convert the above into wholly a matrix form:

$$\partial_t \overline{\boldsymbol{\rho}} = \frac{i}{\hbar} \left[\overline{\boldsymbol{\rho}}, \overline{\mathbf{H}} \right] \qquad (11.3.13)$$

where $\overline{\boldsymbol{\rho}}$ and $\overline{\mathbf{H}}$ are the matrix representations of $\hat{\rho}$ and $\hat{H}$, respectively. We often do not distinguished the matrix representations $\overline{\boldsymbol{\rho}}$ and $\overline{\mathbf{H}}$, and operator representations $\hat{\rho}$ and $\hat{H}$, as mathematically, they are the same. The above equation replaces the Schrödinger equation as the equation of motion for the quantum system.

11.4 Two-Level Quantum Systems

Two-level systems are encountered in nuclear magnetic resonance when the spin states of a particle interact with a magnetic field. They can also be used to study the Josephson junction effect in superconductive devices. Many more complex systems, for simplicity, can be approximated by a two-level system. When an electric field (optical field) interacts with an atom and causes an atomic transition between two energy levels, while the other energy levels are far away, the system can be approximated by a simpler two-level system. Also, approximate two-level systems are often used to represent the value "0" and "1" in quantum computing as we shall see later. One can start with Schrödinger equation which is the equation of motion for a quantum system:

$$i\hbar\partial_t|\psi\rangle = \hat{H}|\psi\rangle \qquad (11.4.1)$$

By making a finite basis approximation, where the state function is approximated by

$$|\psi\rangle = a_1(t)|\psi_1\rangle + a_2(t)|\psi_2\rangle \qquad (11.4.2)$$

where $|\psi_1\rangle$ and $|\psi_2\rangle$ are time independent. Then one can show the equation of motion can be approximated by a much smaller system. A simple two-level system can be described with something that resembles coupled-mode equation;

$$i\hbar\partial_t a_1(t) = E_1 a_1(t) + K a_2(t) \tag{11.4.3}$$

$$i\hbar\partial_t a_2(t) = E_2 a_2(t) + K a_1(t) \tag{11.4.4}$$

where if $K = 0$, the two states will be stationary states evolving independently of each other. When $K \neq 0$, their time evolution will be different evolving according to the coupled-mode equation. We will discuss how to find K.

11.4.1 Interaction of Light with Two-Level Systems

For an optical system, one assumes that before it is perturbed, the atom (or quantum well) is described by a simple two-level system with two stationary states bearing energies E_1 and E_2. The unperturbed Hamiltonian can be described by

$$\hat{H}_0 = \begin{bmatrix} E_1 & 0 \\ 0 & E_2 \end{bmatrix} \tag{11.4.5}$$

In the above, we assume that the Hamiltonian operator and its matrix representation are the same. In this case, the perturbing Hamiltonian due to an electric field pointing in the z direction with strength E_z is assumed to be[‡]

$$\hat{H}_p = ezE_z = -E_z\hat{\mu} \tag{11.4.6}$$

where $\hat{\mu} = -e\hat{z}$.[§] The matrix representation of the dipole moment in terms of two stationary states before perturbation is given by[¶]

$$\mu_{mn} = -e\langle\psi_m|\hat{z}|\psi_n\rangle \tag{11.4.7}$$

where $\hat{z}$ is the position operator, and

$$\left(\hat{H}_p\right)_{mn} = H_{p,mn} = -E_z\mu_{mn} \tag{11.4.8}$$

In the above $\mu_{11} = \mu_{22} = 0$, so are $H_{p11} = H_{p22} = 0$, because ψ_m and ψ_n have different parities. In other words, if one of them has odd parity, the other one has even parity. Then in the coordinate representation, the function $|\psi_i|^2 z$ will have odd parity, and its integral evaluates to zero. But for the off diagonal elements, the integrands will have even parity, and their integrals evaluate to nonzero values. The matrix elements above are formed with

---

[‡]This can be derived using the length gauge (see Gerry and Knight, *Introductory Quantum Optics*, p. 76).
[§]The coordinate representation of $\hat{z}$ is just z.
[¶]This is an approximation since after perturbation, the quantum state is not necessarily describable by just two stationary states. We saw that in the time-dependent perturbation theory, all eigenstates of the system was necessary to approximate the perturbed system.

stationary states, and μ_{12} can be made pure real so that $\mu_{12} = \mu_{21} = \mu_d$. Hence, the perturbing Hamiltonian becomes

$$\hat{H}_p = \begin{bmatrix} 0 & -E_z\mu_d \\ -E_z\mu_d & 0 \end{bmatrix} \tag{11.4.9}$$

The above can be simplified to, and approximated by a two-level system with the use of simple matrix mechanics. This is also very much in the spirit of a finite-basis method or reduced-basis method, where a small number of basis can form a good approximation of the quantum system. The unperturbed Hamiltonian, $\hat{H}_0$ is assumed to have eigenvalues E_1 and E_2 with respect to these two stationary states. The total Hamiltonian is then

$$\hat{H} = \hat{H}_0 + \hat{H}_p = \begin{bmatrix} E_1 & -E_z\mu_d \\ -E_z\mu_d & E_2 \end{bmatrix} \tag{11.4.10}$$

The equation of motion according to Schrödinger equation is

$$i\hbar\partial_t|\psi\rangle = \hat{H}|\psi\rangle \tag{11.4.11}$$

In the above, the perturbing Hamiltonian introduces the off diagonal terms which are the cross-coupling terms in coupled-mode theory. They are responsible for causing the transition of eigenstates between state 1 and state 2. If the off diagonal terms are absent, the stationary states will remain in their respective states without transition. In the above finite basis approximation to the Hamiltonian, the basis used consists of the two eigenstates of the original unperturbed problem. This is also a subspace approximation method, whereby an infinite dimensional Hilbert space has been replaced by a two-dimensional subspace.

The density matrix which is the matrix representation of the density operator is[||]

$$\hat{\rho} = \begin{bmatrix} \rho_{11} & \rho_{12} \\ \rho_{21} & \rho_{22} \end{bmatrix} \tag{11.4.12}$$

Equation (11.2.6) expresses the density operator as operator. But by finding its matrix representation in terms of the two states assuming that they are orthonormal, the density matrix is

$$\hat{\rho} = \begin{bmatrix} |a_1(t)|^2 & a_1(t)a_2^*(t) \\ a_2(t)a_1^*(t) & |a_2(t)|^2 \end{bmatrix} \tag{11.4.13}$$

In (11.2.6), we assume here that the two states are stationary states (they need not be), and the time variations of the states have been put into $a_1(t)$ and $a_2(t)$. Hence, the off diagonal terms can be rapidly varying functions of time due to the interference between the two stationary states.

Next, we derive the equation of motion for the density matrix. Since

$$\begin{aligned} \frac{d\hat{\rho}}{dt} &= \frac{i}{\hbar}\left[\hat{\rho}\hat{H} - \hat{H}\hat{\rho}\right] \\ &= \frac{i}{\hbar}\left(\begin{bmatrix} \rho_{11} & \rho_{12} \\ \rho_{21} & \rho_{22} \end{bmatrix}\begin{bmatrix} E_1 & -E_z\mu_d \\ -E_z\mu_d & E_2 \end{bmatrix} - \begin{bmatrix} E_1 & -E_z\mu_d \\ -E_z\mu_d & E_2 \end{bmatrix}\begin{bmatrix} \rho_{11} & \rho_{12} \\ \rho_{21} & \rho_{22} \end{bmatrix}\right) \\ &= \frac{i}{\hbar}\begin{bmatrix} -E_z\mu_d(\rho_{12} - \rho_{21}) & -E_z\mu_d(\rho_{11} - \rho_{22}) + (E_2 - E_1)\rho_{12} \\ -E_z\mu_d(\rho_{22} - \rho_{11}) + (E_1 - E_2)\rho_{21} & -E_z\mu_d(\rho_{21} - \rho_{12}) \end{bmatrix} \end{aligned} \tag{11.4.14}$$

[||]Again, we ignore the difference between operator and its matrix representation and use the same notation.

From the above

$$\frac{\mathrm{d}\rho_{21}}{\mathrm{d}t} = \frac{i}{\hbar}\left[(\rho_{11} - \rho_{22})E_z\mu_d - (E_2 - E_1)\rho_{21}\right] = -i\omega_{21}\rho_{21} + i\frac{\mu_d E}{\hbar}\Delta\rho \qquad (11.4.15)$$

where $\hbar\omega_{21} = E_2 - E_1$ is the same as the resonant frequency of the two-level system according to the Fermi's golden rule, and $\Delta\rho = \rho_{11} - \rho_{22}$ where $\Delta\rho$ is the proportional to the population difference between the two states. We need not write down the equation for ρ_{12} since $\rho_{12} = \rho_{21}^*$.

From (11.4.14)

$$\frac{\mathrm{d}\rho_{11}}{\mathrm{d}t} = -\frac{i}{\hbar}E_z\mu_d(\rho_{12} - \rho_{21}), \qquad \frac{\mathrm{d}\rho_{22}}{\mathrm{d}t} = -\frac{i}{\hbar}E_z\mu_d(\rho_{21} - \rho_{12}) \qquad (11.4.16)$$

From them, we get

$$\frac{d\Delta\rho}{dt} = \frac{d}{dt}(\rho_{11} - \rho_{22}) = -\frac{i}{\hbar}E_z\mu_d(\rho_{12} - \rho_{21} - \rho_{21} + \rho_{12}) = \frac{2i}{\hbar}E_z\mu_d(\rho_{21} - \rho_{21}^*)$$
$$= \frac{-4}{\hbar}E_z\mu_d\Im m(\rho_{21}) \qquad (11.4.17)$$

since $\rho_{12} = \rho_{21}^*$.

The density matrix method can be used to describe an ensemble of identical atoms. In this case, we implicitly assume that the density matrix has been ensemble averaged. If this system models an ensemble of identical atoms, $\Delta\rho$ is proportional to the population difference of atoms in states 1 and 2. The above equations, (11.4.15) and (11.4.17), form two coupled equations from which the unknowns, $\Delta\rho = \rho_{11} - \rho_{22}$ and ρ_{12} can be solved for.

It is noted that if no external field is applied, there is no change of the population densities in the two states. Moreover, if the external field is switched off, then (11.4.15) evolves in time according to $\exp(-i\omega_{21}t)$. Since these are stationary states, their phase relationship for an ideal quantum system remains locked in this manner. The driving term, the last term in (11.4.15), due to the electric field being on will cause it to veer away from this phase-locked relationship.

At his point, we are going to add phenomenological terms to the above equations to make them agree with experimental observations. The above two level system, so far, has been isolated from a thermal bath or interaction with other atoms. It more aptly describes the two level system of a single atom in isolation. In actuality, there will be a collection of atoms that will be similarly excited by the electric field. These atoms will be coupled to other thermal sources such as by collision with other atoms or with the wall of the container.

As mentioned before, the density matrix can be used to represent the ensemble averaged state of the collection of these atoms. Here, for a collection of atoms, $\rho_{11} - \rho_{22}$ is the fractional population difference in electrons of the atoms between state 1 and state 2. But when coupled to a thermal bath, there will be some electrons excited to level 2 even in the absence of an external electric field. Hence, we modify (11.4.17) to account for the coupling to a thermal bath to become

$$\frac{d}{dt}\Delta\rho = \frac{2i}{\hbar}E_z\mu_d\left(\rho_{21} - \rho_{21}^*\right) - \frac{\Delta\rho - \Delta\rho_o}{T_1} \qquad (11.4.18)$$

where $\Delta\rho_o = (\rho_{11} - \rho_{22})_o$ is the quiescent steady state value of $\Delta\rho$ in the absence of external driving field. The transient value $\Delta\rho$ is assumed to relax to the quiescent value $\Delta\rho_o$ in time T_1. The last term above is added phenomenologically in order to match experimental data.

Also, $\rho_{21} = C\, e^{-i\omega_{21}t}$ in the absence of external driving field and if we have a pure quantum state which is a linear supposition of two quantum states in quantum coherence, or in locked phase with respect to each other. But if the two level system is coupled to an external heat bath, the phase coherence between the two energy states will be lost and average to zero. We can describe this by adding a term to (11.4.15) to account for dephasing, or

$$\frac{d\rho_{21}}{dt} = -i\omega_{21}\rho_{21} + i\frac{\mu_d}{\hbar}E_z\, \Delta\rho - \frac{\rho_{21}}{T_2} \tag{11.4.19}$$

T_2 is the dephasing time which is usually shorter than T_1. The first term on the right-hand side of the above gives rise to a resonant solution if the other terms are absent. The second term is a driving term due to the presence of the electric field and the population difference between states 1 and 2. The last term is an attenuation due to the decoherence between the two quantum states. If the two states are entirely incoherent or uncorrelated, $\rho_{21} = 0$. The above implies that if the exciting field E_z is zero, the coherence between the two states will eventually become zero or lost due to dephasing and coupling to a thermal bath (noise bath).

At this point, it is prudent to discuss the role of the exciting field, which is of the form

$$E_z(t) = E_o \cos\omega t = \frac{E_o}{2}\left[e^{i\omega t} + e^{-i\omega t}\right] \tag{11.4.20}$$

When $\Delta\rho$ is slowly varying in (11.4.19), the driving term, which is the second term on the right-hand side in (11.4.19) has two rotating signals $e^{i\omega t}$ and $e^{-i\omega t}$. The response, ρ_{21} will have two rotating signals as well. Hence, $\rho_{21} = C_1\, e^{-i\omega t} + C_2\, e^{i\omega t}$ when driven by a sinusoidal source (11.4.20). In the above, (11.4.19) has a resonant frequency in the vicinity of ω_{21}. If $\omega \approx \omega_{21}$, then $C_1 \gg C_2$ because C_1 corresponds to the amplitude of the resonance solution which becomes very large under sinusoidal excitation at a frequency close to resonance. Therefore, we define

$$\rho_{21}(t)\, e^{i\omega t} = C_1 + C_2 e^{2i\omega t} \approx C_1 = \beta_{21}(t) \tag{11.4.21}$$

where C_1 is slowly varying so that $\beta_{21}(t)$ is slowly varying when we ignore the rapid term. Using this in (11.4.18) and making us of (11.4.20), and keeping only the slowly varying terms, we have

$$\frac{d}{dt}\Delta\rho = \frac{i}{\hbar}E_o\mu_d\left(\beta_{21} - \beta_{12}^*\right) - \frac{\Delta\rho - \Delta\rho_o}{T_1} \tag{11.4.22}$$

In the above, we have kept only the slowest varying term of the first term on the right-hand side. The rapid term is assumed to only result in a small response in $\Delta\rho$. This is known as the rotating wave approximation.

From (11.4.20), we deduce that

$$\frac{d\rho_{21}(t)}{dt} = \left[\frac{d}{dt}\beta_{21}(t)\right]e^{-i\omega t} - i\omega\beta_{21}(t)\, e^{-i\omega t} \tag{11.4.23}$$

Consequently, (11.4.19) becomes

$$\frac{d}{dt}\beta_{21} = i\,\Delta\omega\,\beta_{21} + i\frac{\mu_d E_o}{2\hbar}\Delta\rho - \frac{\beta_{21}}{T_2} \tag{11.4.24}$$

where $\Delta\omega = \omega - \omega_{21}$, and we have kept only the slowly varying terms.

In steady state, $d\Delta\rho/dt = 0$ in (11.4.22), or $\Delta\rho \to \Delta\rho_o$ in (11.4.24), or it is a constant. Hence, (11.4.24) has a particular solution and a homogeneous solution. The homogeneous solution is transient and will die out in the steady-state limit. The particular solution is driven by a constant term, and hence, is a constant of time. Therefore, β_{21} must tend to a constant in steady state with $d\beta_{21}/dt = 0$.

Next, the steady state solution of the above equations can be sought. Defining $\Omega = \mu_d E_o / (2\hbar)$, where E_o is the electric field amplitude, we have from (11.4.24) and (11.4.22) that

$$0 = i\Delta\omega\beta_{21} + i\Omega\Delta\rho - \beta_{21}/T_2 \tag{11.4.25}$$

$$0 = -4\Omega\Im m\left(\beta_{21}\right) - \left(\Delta\rho - \Delta\rho_o\right)/T_1 \tag{11.4.26}$$

Taking the real and imaginary parts of (11.4.25), we arrive at

$$0 = -\Delta\omega\Im m\left(\beta_{21}\right) - \Re e\left(\beta_{21}\right)/T_2 \tag{11.4.27}$$

$$0 = \Delta\omega\Re e\left(\beta_{21}\right) + \Omega\Delta\rho - \Im m\left(\beta_{21}\right)/T_2 \tag{11.4.28}$$

Equations (11.4.26)-(11.4.28) constitute three equations with three unknowns $\Delta\rho$, $\Im m\left(\beta_{21}\right)$, $\Re e\left(\beta_{21}\right)$, from which they can be solved for. Therefore, in the steady state,

$$\Delta\rho = \Delta\rho_o \frac{1 + \Delta\omega^2 T_2^2}{1 + \Delta\omega^2 T_2^2 + 4\Omega^2 T_1 T_2} \tag{11.4.29}$$

$$\Re e\left(\beta_{21}\right) = -\frac{\Delta\omega\Omega T_2^2 \Delta\rho_o}{1 + \Delta\omega^2 T_2^2 + 4\Omega^2 T_1 T_2} \tag{11.4.30}$$

$$\Im m\left(\beta_{21}\right) = -\frac{\Omega T_2 \Delta\rho_o}{1 + \Delta\omega^2 T_2^2 + 4\Omega^2 T_1 T_2} \tag{11.4.31}$$

If there are N atoms, then we define

$$\Delta N = N\Delta\rho, \quad \Delta N_o = N\Delta\rho_o \tag{11.4.32}$$

and

$$\Delta N = \Delta N_o \frac{1 + \Delta\omega^2 T_2^2}{1 + \Delta\omega^2 T_2^2 + 4\Omega^2 T_1 T_2} \tag{11.4.33}$$

We would eventually like to use the above calculation to ascertain the effect of the electric field on the ensemble of atoms. The result is to polarize each atom yielding a dipole moment producing a polarization current. To this end, we will find the dipole moment produced. The expectation value of the dipole moment is

$$\langle\hat{\mu}\rangle = \text{tr}(\hat{\rho}\hat{\mu}) \tag{11.4.34}$$

where

$$\hat{\rho}\hat{\mu} = \begin{bmatrix} \rho_{11} & \rho_{12} \\ \rho_{21} & \rho_{22} \end{bmatrix} \begin{bmatrix} 0 & \mu_d \\ \mu_d & 0 \end{bmatrix} = \begin{bmatrix} \rho_{12}\mu_d & \rho_{11}\mu_d \\ \rho_{22}\mu_d & \rho_{21}\mu_d \end{bmatrix} \tag{11.4.35}$$

Hence

$$\langle \hat{\mu} \rangle = \mu_d (\rho_{12} + \rho_{21}) = 2\mu_d \Re e(\rho_{12}) \tag{11.4.36}$$

because $\rho_{12} = \rho_{21}^*$.

After making the rotating wave approximation that $\rho_{12} \doteq \beta_{12} e^{i\omega t}$, $\rho_{21} \doteq \beta_{21} e^{-i\omega t}$ we arrive at

$$\langle \hat{\mu} \rangle = \mu_d (\rho_{12} + \rho_{21}) \doteq \mu_d \left(\beta_{12} e^{i\omega t} + \beta_{21} e^{-i\omega t} \right) \tag{11.4.37}$$

after using (11.4.21). Also, $\langle \hat{\mu} \rangle$ has to be a measurable, observable, real-valued quantity, and that is what the above shows. Therefore,

$$\langle \hat{\mu} \rangle = 2\mu_d \left[\Re e \left(\beta_{21} \right) \cos \omega t + \Im m \left(\beta_{21} \right) \sin \omega t \right] \tag{11.4.38}$$

since $\beta_{21} = \beta_{12}^*$. In general, it is not in phase with the exciting electric field, having in-phase and quadrature components.

In electromagnetics, it is customary to write a relationship in the frequency domain using phasors, or we have

$$\tilde{P} = \epsilon_o \tilde{\chi} \tilde{E} \tag{11.4.39}$$

where $\tilde{A}$ indicates that A is a phasor quantity, and $\tilde{P}$, $\tilde{\chi}$, and $\tilde{E}$ are their phasor representations in frequency domain. The corresponding relationhsip in the time domain is

$$\begin{aligned} P(t) &= \Re e \left[\epsilon_o \left(\chi' + i\chi'' \right) E_o e^{-i\omega t} \right] \\ &= \epsilon_o \left(\chi' \cos \omega t + \chi'' \sin \omega t \right) E_o \end{aligned} \tag{11.4.40}$$

But also, from the above calculation using density matrix method,

$$P(t) = N \langle \hat{\mu} \rangle \tag{11.4.41}$$

From (11.4.38), we have

$$P(t) = 2N\mu_d \left[\Re e \left(\beta_{21} \right) \cos \omega t + \Im m \left(\beta_{21} \right) \sin \omega t \right] \tag{11.4.42}$$

Comparing (11.4.40) and (11.4.42), we have

$$\begin{aligned} \chi' &= \frac{2N\mu_d}{\epsilon_o E_o} \Re e \left(\beta_{21} \right) \\ &= -\frac{\mu_d^2 T_2 \Delta N_o}{\epsilon_o \hbar} \frac{\Delta \omega T_2}{1 + (\Delta \omega)^2 T_2^2 + 4\Omega^2 T_1 T_2} \end{aligned} \tag{11.4.43}$$

$$\begin{aligned} \chi'' &= \frac{2N\mu_d}{\epsilon_o E_o} \Im m \left(\beta_{21} \right) \\ &= \frac{\mu_d^2 T_2 \Delta N_o}{\epsilon_o \hbar} \frac{1}{1 + (\Delta \omega)^2 T_2^2 + 4\Omega^2 T_1 T_2} \end{aligned} \tag{11.4.44}$$

where Ω is proportional to the strength of the externally applied electric field. When the applied electric field is small, we can ignore the $4\Omega^2 T_1 T_2$ term to get

$$\chi'(\omega) = \frac{\mu_d^2 T_2 \Delta N_o}{\epsilon_o \hbar} \frac{(\omega_{21} - \omega) T_2}{1 + (\omega - \omega_{21})^2 T_2^2} \tag{11.4.45}$$

$$\chi''(\omega) = \frac{\mu_d^2 T_2 \Delta N_o}{\epsilon_o \hbar} \frac{1}{1 + (\omega - \omega_{21})^2 T_2^2} \tag{11.4.46}$$
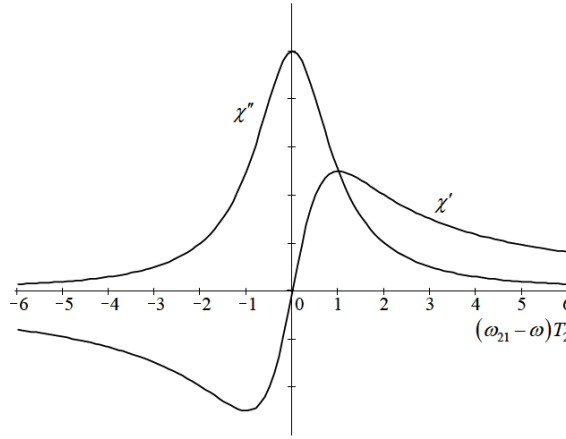
Figure 11.1: Plots of χ' and χ'' according to (11.4.45) and (11.4.46) (from DAB Miller).

Note that in this analysis, the perturbing field need not be small. The approximation we have made is the two-state approximation. This is a decent approximation if the other eigenstates are far away from these two states in terms of energy levels. The other approximation is the rotating wave approximation which is good if the exciting frequency of the electric field is high and is much faster than the relaxation times T_1 and T_2, and that it is close to ω_{21}.

Since this analysis is not a small perturbation analysis, Ω above, which is proportional to the exciting electric field, can be large. When it is large, we note from (11.4.29) and (11.4.33) that the fractional population difference approaches zero, or that $\rho_{11} = \rho_{22}$. In this case, the absorption transition is equal to the stimulated emission. From (11.4.44), when $\Delta \omega = 0$, or $\omega = \omega_{21}$ we can express the denominator as

$$1 + 4\Omega^2 T_1 T_2 = 1 + \frac{I}{I_s} \tag{11.4.47}$$

where I is the field intensity, and I_s is the saturation intensity, since I is proportional to the square of the electric field. The saturation intensity is defined to be the intensity at which the absorption peak of χ'' will drop to half of its value compared to the weak excitation case where I is very small.

Equations (11.4.22) and (11.4.24) are also known as the optical Bloch equations. These equations were first used to analyze another two-state system, the nuclear magnetic resonance

system. With a magnetic field pointing in the z direction, when the nuclear spins are pointing upward and aligned with the magnetic field, they are in the lower energy state. However, when they are pointing anti-parallel to the magnetic field, they are in a higher energy state. RF (radio frequency) field can be used to flip the spin states, and similar equations as above can be derived to describe the spin dynamics.

Exercise 3

A Josephson junction can also be described by a two-level quantum system where a Cooper pair (quasi-particle) is trapped on either side of a thin insulating junction. The Hamilton operator of the system can be described by

$$\hat{H} = \begin{bmatrix} E_1 & K \\ K & E_2 \end{bmatrix} \tag{11.4.48}$$

where E_1 and E_2 are the potential energy at the bottom of the conduction band, according to the effective mass theory for such quasi-particle. By connecting the two regions of the superconductor to a voltage source, a difference in the potential energy between the regions can be created, yielding $E_2 - E_1 = -qV$ where $q = -2e$ for Cooper pairs, and V is the voltage of the source.

Tunnelling of Cooper Pair's
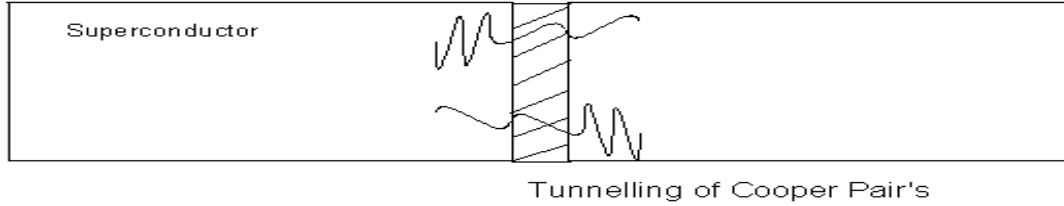
Use the two-level quantum system theory, where

$$\rho_{11} = |a_1|^2, \rho_{22} = |a_2|^2, \rho_{12} = a_1 a_2^* = |a_1||a_2|e^{i(\theta_1 - \theta_2)} = \rho_{21}^* \tag{11.4.49}$$

Show that

$$\rho_{12} = \sqrt{\rho_{11}\rho_{22}}\, e^{-i\delta} \tag{11.4.50}$$

where $\delta = \theta_2 - \theta_1$. Next show that the current flow through the Josephson junction is equal to

$$J = \frac{d\rho_{11}}{dt} = \frac{2K}{\hbar}\sqrt{\rho_{11}\rho_{22}}\sin\delta \tag{11.4.51}$$

Hence, the current flow between the two regions is proportional to the sine of the phase difference between the wave functions of the two regions. Furthermore, show that

$$\frac{d}{dt}\sqrt{\rho_{11}\rho_{22}} = 0 \tag{11.4.52}$$

and that

$$\frac{d(\rho_{12} + \rho_{21})}{dt} = i\omega_{21}(\rho_{12} - \rho_{21}) \tag{11.4.53}$$

From the above, show that the phase difference can be affected by the applied voltage, and that the rate of change of the phase difference is given by

$$\dot{\delta} = -(E_2 - E_1)/\hbar = qV/\hbar \tag{11.4.54}$$

Deduce that

$$\delta(t) = \delta_0 + \frac{q}{\hbar}\int_0^t V(t')dt' \tag{11.4.55}$$

If

$$V = V_0 + v\cos(\omega t) \tag{11.4.56}$$

then

$$\delta(t) = \delta_0 + \frac{q}{\hbar}V_0 t + \frac{q}{\hbar}\frac{v}{\omega}\sin(\omega t) \tag{11.4.57}$$

Since $\hbar$ is small, the above gives rise to rapidly varying phases, and the associated current cannot be measured, or averages to zero. Using

$$\sin(x + \Delta x) \approx \sin(x) + \Delta x\cos(x) \tag{11.4.58}$$

Show that the current is now

$$J \approx J_0[\sin(\delta_0 + \frac{q}{\hbar}V_0 t) + \frac{q}{\hbar}\frac{v}{\omega}\sin(\omega t)\cos(\delta_0 + \frac{q}{\hbar}V_0 t)] \tag{11.4.59}$$

Explain how you would pick the frequency ω so that a measurable DC current ensues.

Because of the sensitivity of the Josephson junction to an applied EMF, it can be used to sense magnetic field in a SQUID (superconducting quantum interference detector). The magnetic field through the ring in the above figure generates a difference in the EMF and therefore, a phase difference on the two arms of the ring superconductor with Josephson junctions in them. The interference pattern of the output current is very sensitive to the magnetic field, and hence, can be used to detect small field differences.
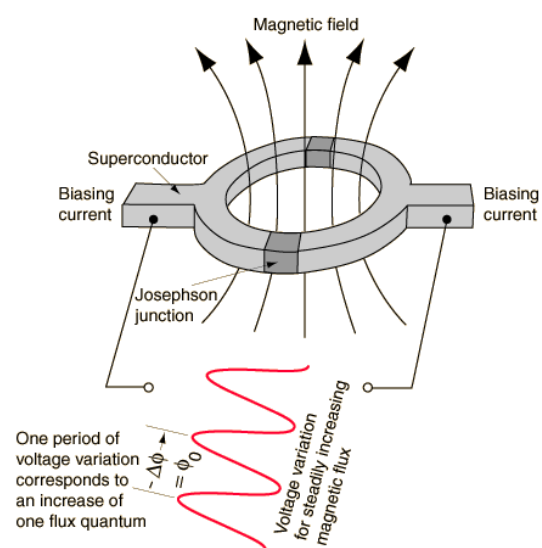
Chapter 12

Quantization of Classical Fields

12.1 Introduction

The quantum harmonic oscillator is extremely important for the quantization of classical fields, be it electromagnetic, acoustic, or elastic in nature. Classical fields can be thought of as due to a set of coupled classical harmonic oscillators. In the case of electromagnetics field, these classical harmonic oscillators are dipoles in space polarized by the electric field. In the case of acoustic or elastic field, they are due to coupled atomic or molecular vibrations in a medium. Many experimental evidence suggest that there are photon particles associated with electromagnetic field, and phonon particles associated with vibrations of atoms or molecules in crystalline solids.

When the classical harmonic oscillators are replaced by quantum harmonic oscillators, the classical field becomes quantized as shall be seen. The energy that can propagate and associated with the field is quantized, as in the case of the quantum harmonic oscillator. In classical electromagnetics when we calculate energy, e.g. as stored energies

$$W_e = \frac{1}{2}\epsilon \left|\mathbf{E}\right|^2, \qquad W_h = \frac{1}{2}\mu \left|\mathbf{H}\right|^2 \tag{12.1.1}$$

we think of W_e and W_h as a continuum, capable of assuming all values from zero to a large value. Also, the power flow by a plane wave

$$S = \frac{1}{2\eta_0} \left|\mathbf{E}\right|^2 \tag{12.1.2}$$

assumes continuous values in classical electromagnetics. As shall be shown, this cannot be the case if the electromagnetic field is due to a set of coupled quantum harmonic oscillators. The energy carried by a plane wave is also quantized corresponding to packets of energy.

The quantized nature of electromagnetic radiation has been observed historically. This was first suggested for explaining black-body radiation and deriving the Planck's radiation law in 1900. Later, it was used by Einstein to explain experimental findings in the photoelectric effect. From the experimental data, it was deduced that the packet of energy associated with

electromagnetic field is

$$E_{photon} = \hbar\omega \tag{12.1.3}$$

This quantized energy is very small when ω is small but sizeable in the optical regime. The grainy nature of electromagnetic field is unimportant at radio frequency (RF), and microwave frequencies, but is important at optical frequencies. We shall show that this is the case when the electromagnetic field is quantized when the classical harmonic oscillators are replaced by quantum harmonic oscillators.

To introduce the quantized nature of electromagnetic field, one way is to regard the field in a resonant cavity or a box with periodic boundary condition. The electromagnetic field in such a system oscillates harmonically. We can replace the classical harmonic oscillators associated with such a system with the quantum harmonic oscillator; and hence, quantize the amplitude and the associated energy of the electromagnetic field.

We will first revisit the quantum harmonic oscillator because of its importance here. Next we study the wave on a linear atomic chain, first as a classical wave, and then as a set of coupled quantum harmonic oscillators. Then we study the quantization of electromagnetic wave field along similar spirit.

12.2   The Quantum Harmonic Oscillator Revisited

The governing equation for the quantum harmonic oscillator is

$$\hat{H}\psi = \left(-\frac{\hbar^2}{2m}\frac{d^2}{dq^2} + \frac{1}{2}m\omega^2 q^2\right)\psi = E\psi \tag{12.2.1}$$

where q has been used to denote the displacement from the mean of this simple harmonic oscillator. We divide the above by $\hbar\omega$ to make the operator and the eigenvalue dimensionless. Letting $\xi = \sqrt{\frac{m\omega}{\hbar}}q$, a dimensionless quantity, the above becomes

$$\frac{1}{2}\left(-\frac{d^2}{d\xi^2} + \xi^2\right)\psi = \frac{E}{\hbar\omega}\psi \tag{12.2.2}$$

The above looks almost like $A^2 - B^2$ with the exception that operators are involved instead of scalars. One can show that

$$\frac{1}{\sqrt{2}}\left(-\frac{d}{d\xi} + \xi\right)\frac{1}{\sqrt{2}}\left(\frac{d}{d\xi} + \xi\right) = \frac{1}{2}\left(-\frac{d^2}{d\xi^2} + \xi^2\right) - \frac{1}{2}\left(\frac{d}{d\xi}\xi - \xi\frac{d}{d\xi}\right) \tag{12.2.3}$$

Furthermore, it can be easily shown that

$$\left(\frac{d}{d\xi}\xi - \xi\frac{d}{d\xi}\right) = \hat{I} \tag{12.2.4}$$

Hence,

$$\frac{1}{2}\left(-\frac{d^2}{d\xi^2} + \xi^2\right) = \frac{1}{\sqrt{2}}\left(-\frac{d}{d\xi} + \xi\right)\frac{1}{\sqrt{2}}\left(\frac{d}{d\xi} + \xi\right) + \frac{1}{2} \tag{12.2.5}$$

We define

$$\hat{a}^{\dagger} = \frac{1}{\sqrt{2}} \left(-\frac{d}{d\xi} + \xi \right), \qquad \text{creation or raising operator} \tag{12.2.6}$$

$$\hat{a} = \frac{1}{\sqrt{2}} \left(\frac{d}{d\xi} + \xi \right), \qquad \text{annihilation or lowering operator} \tag{12.2.7}$$

It can easily be shown by integration by parts that $\hat{a}^{\dagger}$ is in fact the conjugate transpose operator of $\hat{a}$.

Consequently, from (12.2.1) and (12.2.2), we have

$$\left(\hat{a}^{\dagger} \hat{a} + \frac{1}{2} \right) \psi = \frac{E}{\hbar \omega} \psi \tag{12.2.8}$$

and

$$\hat{H} = \hbar \omega \left(\hat{a}^{\dagger} \hat{a} + \frac{1}{2} \right) \tag{12.2.9}$$

Since

$$E_n = \left(n + \frac{1}{2} \right) \hbar \omega \tag{12.2.10}$$

from our previous knowledge, we conclude that

$$\hat{a}^{\dagger} \hat{a} \left| \psi_n \right\rangle = n \left| \psi_n \right\rangle \tag{12.2.11}$$

In other words, if $\left| \psi_n \right\rangle$ is an eigenstate of (12.2.1) or (12.2.8), the above must be true. We define a number operator

$$\hat{n} = \hat{a}^{\dagger} \hat{a} \tag{12.2.12}$$

such that

$$\hat{n} \left| \psi_n \right\rangle = n \left| \psi_n \right\rangle \tag{12.2.13}$$

One can also show by direct substitution using (12.2.6) and (12.2.7) that

$$\left[\hat{a}, \hat{a}^{\dagger} \right] = \hat{a} \hat{a}^{\dagger} - \hat{a}^{\dagger} \hat{a} = \hat{I} \tag{12.2.14}$$

One can easily show, using the commutation relation, that

$$\hat{a}^{\dagger} \hat{a} \hat{a} \left| \psi_n \right\rangle = \left(\hat{a} \hat{a}^{\dagger} - 1 \right) \hat{a} \left| \psi_n \right\rangle = \hat{a} \left(\hat{a}^{\dagger} \hat{a} - 1 \right) \left| \psi_n \right\rangle = (n-1) \hat{a} \left| \psi_n \right\rangle \tag{12.2.15}$$

Therefore

$$\hat{a} \left| \psi_n \right\rangle = A_n \left| \psi_{n-1} \right\rangle \tag{12.2.16}$$

Hence, $\hat{a}$ is a lowering operator. Similarly, using the commutation relation,

$$\hat{a}^{\dagger} \hat{a} \hat{a}^{\dagger} \left| \psi_n \right\rangle = \hat{a}^{\dagger} \left(1 + \hat{a}^{\dagger} \hat{a} \right) \left| \psi_n \right\rangle = (n+1) \hat{a}^{\dagger} \left| \psi_n \right\rangle \tag{12.2.17}$$

Therefore

$$\hat{a}^{\dagger} \left| \psi_n \right\rangle = B_{n+1} \left| \psi_{n+1} \right\rangle \tag{12.2.18}$$

Hence, $\hat{a}^\dagger$ is a raising operator. Testing (12.2.16) with $\langle\psi_{n-1}|$, and (12.2.18) with $\langle\psi_{n+1}|$, we have

$$\langle\psi_{n-1}|\,\hat{a}\,|\psi_n\rangle = A_n \ , \quad \langle\psi_{n+1}|\,\hat{a}^\dagger\,|\psi_n\rangle = B_{n+1} \tag{12.2.19}$$

By comparing the two equations in (12.2.19), taking the transpose of the second equation, and let $n+1 \to n$, we deduce that

$$\langle\psi_{n-1}|\,\hat{a}\,|\psi_n\rangle = B_n^* \tag{12.2.20}$$

It is clear that

$$A_n = B_n^* \tag{12.2.21}$$

Since

$$\hat{a}^\dagger\hat{a}\,|\psi_n\rangle = A_n\hat{a}^\dagger\,|\psi_{n-1}\rangle = A_n B_n\,|\psi_n\rangle \tag{12.2.22}$$

$$= |A_n|^2\,|\psi_n\rangle = n\,|\psi_n\rangle \tag{12.2.23}$$

We conclude that

$$A_n = \sqrt{n} \tag{12.2.24}$$

to within an arbitrary phase factor. But the phase factor can be absorbed into the definition of the eigenfunction $|\psi_n\rangle$ since it is defined to within a arbitrary phase factor. Therefore, for simplicity, we assume that A_n and B_n are both real. Therefore

$$\hat{a}\,|\psi_n\rangle = \sqrt{n}\,|\psi_{n-1}\rangle \tag{12.2.25}$$

$$\hat{a}^\dagger\,|\psi_n\rangle = \sqrt{n+1}\,|\psi_{n+1}\rangle \tag{12.2.26}$$

12.2.1 Eigenfunction by the Ladder Approach

With the raising operators, we can construct the requisite eigenfunctions if we only know the ground state eigenfunction. The ground state eigenfunction satisfies

$$\hat{a}\,|\psi_0\rangle = 0 \tag{12.2.27}$$

Written explicitly in coordinate space representation we have

$$\frac{1}{\sqrt{2}}\left(\frac{d}{d\xi} + \xi\right)\psi_0\left(\xi\right) = 0 \tag{12.2.28}$$

We can solve the above ordinary differential equation to get

$$\psi_0\left(\xi\right) = \frac{1}{\pi^{1/4}}e^{-\xi^2/2} \tag{12.2.29}$$

We can also verify that $\psi_0\left(\xi\right)$ satisfies (12.2.28) by direct substitution. Consequently, one can show that

$$\left(\hat{a}^\dagger\right)^n|\psi_0\rangle = \sqrt{n!}\,|\psi_n\rangle \tag{12.2.30}$$

or that the n-th eigenstate is

$$|\psi_n\rangle = \frac{1}{\sqrt{n!}}\left(\hat{a}^\dagger\right)^n|\psi_0\rangle \tag{12.2.31}$$

12.3   Schrödinger Picture versus Heisenberg Picture

At this juncture, it is prudent to discuss the difference between the Schrödinger picture versus the Heisenberg picture of quantum mechanics. In the Schrödinger picture, the time dependence is in the wavefunction or the state vector $|\psi\rangle = |\psi(t)\rangle$; whereas in the Heisenberg picture, the time dependence is in the operator that represents an observable which is a measurable quantity. As shall be seen, the Heisenberg picture is closer to the classical picture.

The Schrödinger equation is given by

$$\hat{H}|\psi\rangle = i\hbar\partial_t|\psi\rangle \tag{12.3.1}$$

The formal solution to the above in operator form can be written as

$$|\psi(t)\rangle = e^{-\frac{i}{\hbar}\hat{H}t}|\psi_0\rangle \tag{12.3.2}$$

where $|\psi_0\rangle = |\psi(0)\rangle$ is the state vector at $t = 0$ which is independent of time, and $\hat{H}$ in (12.3.1) is time independent. The expectation value of an operator $\hat{O}$, which represents an observable or a measurable quantity, is

$$\langle\hat{O}\rangle = \langle\psi(t)|\hat{O}|\psi(t)\rangle = \langle\psi_0|e^{\frac{i}{\hbar}\hat{H}t}\hat{O}\ e^{-\frac{i}{\hbar}\hat{H}t}|\psi_0\rangle = \langle\psi_0|\hat{O}_H(t)|\psi_0\rangle \tag{12.3.3}$$

where

$$\hat{O}_H = e^{\frac{i}{\hbar}\hat{H}t}\hat{O}\ e^{-\frac{i}{\hbar}\hat{H}t} = \hat{U}^\dagger(t)\hat{O}\hat{U}(t) \tag{12.3.4}$$

and

$$\hat{U}(t) = e^{-\frac{i}{\hbar}\hat{H}t} \tag{12.3.5}$$

Here, $\hat{U}(t)$ is the propagator or the time evolution operator. It is clearly unitary since $\hat{U}^\dagger(t)\hat{U}(t) = \hat{I}$.

In the above, $\hat{O}$ is time independent in the Schrödinger picture, but $\hat{O}_H(t)$ is time dependent in the Heisenberg picture. Hence, in general, the relationship between an operator in the Schrödinger picture and in the Heisenberg picture is given by

$$\hat{O}_H(t) = e^{\frac{i}{\hbar}\hat{H}t}\hat{O}_S e^{-\frac{i}{\hbar}\hat{H}t} \tag{12.3.6}$$

where the subscripts "H" and "S" mean the Heisenberg picture and the Schrödinger picture respectively. Clearly, $\hat{O}_S = \hat{O}_H(t = 0)$. It can be shown easily that

$$\frac{d\hat{O}_H}{dt} = \frac{i}{\hbar}\left(\hat{H}\hat{O}_H - \hat{O}_H\hat{H}\right) = \frac{i}{\hbar}\left[\hat{H}, \hat{O}_H\right] \tag{12.3.7}$$

This is the Heisenberg equation of motion for quantum operators.

For the simple quantum harmonic oscillator,

$$\hat{H} = \hbar\omega\left(\hat{a}^\dagger\hat{a} + \frac{1}{2}\right) \tag{12.3.8}$$

It can be shown easily that

$$\left[\hat{H}, \hat{a}\right] = -\hbar\omega\hat{a} \tag{12.3.9}$$

or that

$$\frac{d\hat{a}}{dt} = -i\omega\hat{a} \tag{12.3.10}$$

after using the commutation relation for that $\hat{a}\hat{a}^{\dagger}\hat{a} = (\hat{I} - \hat{a}^{\dagger}\hat{a})\hat{a}$. Consequently, solving give

$$\hat{a}(t) = \hat{a}(0)e^{-i\omega t} \tag{12.3.11}$$

Similarly, one can show that

$$\hat{a}^{\dagger}(t) = \hat{a}^{\dagger}(0)e^{i\omega t} \tag{12.3.12}$$

The above are the annihilation and creation operators in the Heisenberg picture.

12.3.1   Quantum Hamiltonian Mechanics

We can derive the quantum version of Hamiltonian mechanics.* In the Heisenberg picture, the operators which represent observables are functions of time, and the equations of motion for the observable operators in Heisenberg picture evolve as

$$\frac{\partial\hat{q}}{\partial t} = -\frac{i}{\hbar}\left[\hat{q}, \hat{H}\right], \qquad \frac{\partial\hat{p}}{\partial t} = -\frac{i}{\hbar}\left[\hat{p}, \hat{H}\right] \tag{12.3.13}$$

The basic commutation relation is

$$[\hat{q}, \hat{p}] = \hat{q}\hat{p} - \hat{p}\hat{q} = i\hbar\hat{I} \tag{12.3.14}$$

Equation (12.3.14) above can be easily verified in the Schrödinger picture because the coordinate representations of $\hat{p}$ and $\hat{q}$ are

$$\hat{p} = -i\hbar\partial_q = -i\hbar\frac{\partial}{\partial q}, \qquad \hat{q} = q \tag{12.3.15}$$

Hence, (12.3.14) can be easily verified by substituting the above into it. In other words, they follow from de Broglie and Schrödinger postulates, which are the fundamental postulates leading to quantum mechanics. In contrast, some schools assume that the commutation relation $[\hat{q}, \hat{p}] = i\hbar\hat{I}$ as the fundamental quantum postulate, and that de Broglie and Schrödinger postulates being derivable from it.† Some authors refer to this as canonical quantization, where the classical variables p and q as canonical variables that are elevated to be quantum operators, and the commutation relation between them as canonical commutation. The two views

---

*This section is extracted from Chew, Liu, Salazer-Lazarus, and Wei, "Quantum Electromagnetics: A New Look," to be published in *IEEE J. Multiscale Multiphysics Computation*.

†This is a view promulgated by Dirac also.

are largely equivalent, but we favor the historical development over the formal commutation relation.

Having verified (12.3.14) in Schrödinger picture, it can be easily verified in the Heisenberg picture as well, namely

$$[\hat{q}(t), \hat{p}(t)] = i\hbar \hat{I} \tag{12.3.16}$$

The above is also known as the equal time commutator.

It can be shown by the repeated application of the commutator in (12.3.16) that

$$[\hat{p}, \hat{q}^n] = -in\hat{q}^{n-1}\hbar \tag{12.3.17}$$

Note that the above is derived without applying calculus, but only using (12.3.16). Nevertheless we can borrow the calculus notation and rewrite the above as

$$[\hat{p}, \hat{q}^n] = -in\hat{q}^{n-1}\hbar = -i\hbar \left(\frac{\partial}{\partial \hat{q}} \hat{q}^n \right) \tag{12.3.18}$$

As has been seen before, the derivative with respect to an operator has no meaning unless the operator acts on its eigenvector. Consequently, the above can be rewritten as

$$[\hat{p}, \hat{q}^n]|q\rangle = -in\hat{q}^{n-1}\hbar|q\rangle = -inq^{n-1}\hbar|q\rangle = -i\hbar \left(\frac{\partial}{\partial q} q^n \right) |q\rangle = -i\hbar \left(\frac{\partial}{\partial \hat{q}} \hat{q}^n \right) |q\rangle \tag{12.3.19}$$

In the above $|q\rangle$ is the eigenvector of the position operator $\hat{q}$ with eigenvalue q, namely, that $\hat{q}|q\rangle = q|q\rangle$. It is to be noted that the $\partial_{\hat{q}}$ operator above acts only on $\hat{q}^n$, and nothing beyond to its right.

One can expand

$$\hat{H}(\hat{p}, \hat{q}) = \hat{H}_0(\hat{p}, 0) + \hat{H}_1(\hat{p}, 0)\hat{q} + \hat{H}_2(\hat{p}, 0)\hat{q}^2 + \hat{H}_3(\hat{p}, 0)\hat{q}^3 + \cdots \tag{12.3.20}$$

then it is clear that

$$\left[\hat{p}, \hat{H} \right] = -i\hbar \frac{\partial}{\partial \hat{q}} \hat{H}(\hat{p}, \hat{q}) \tag{12.3.21}$$

Similarly, one can show that

$$\left[\hat{q}, \hat{H} \right] = i\hbar \frac{\partial}{\partial \hat{p}} \hat{H}(\hat{p}, \hat{q}) \tag{12.3.22}$$

Hence, (12.3.13) can then be rewritten as

$$\frac{\partial \hat{q}}{\partial t} = \frac{\partial \hat{H}(\hat{p}, \hat{q})}{\partial \hat{p}}, \qquad \frac{\partial \hat{p}}{\partial t} = -\frac{\partial \hat{H}(\hat{p}, \hat{q})}{\partial \hat{q}} \tag{12.3.23}$$

The above are just the Hamilton equations for a quantum system: they bear striking similarities to the classical Hamilton equations of motion. It is to be noted that the above equations (12.3.23) do not involve $\hbar$, but $\hbar$ surfaces in the coordinate representation of the operator $\hat{p}$

in (12.3.15). In addition, one needs to take the expectation value of (12.3.23) to arrive at their classical analogue.

Therefore, the procedure for obtaining the quantum equations of motion is clear. First, the classical Hamiltonian for the system is derived. Then the conjugate variables, in this case, position q and momentum p, are elevated to become quantum operators $\hat{q}$ and $\hat{p}$. In turn, the Hamiltonian becomes a quantum operator as well. Then the quantum state of the system evolves according to (12.3.1) and (12.3.2).

More important, in the Heisenberg picture, the quantum equations of motion have the same algebra as the classical equations of motion as in (12.3.23). A summary of this quantization procedure is given in Figure 12.1.
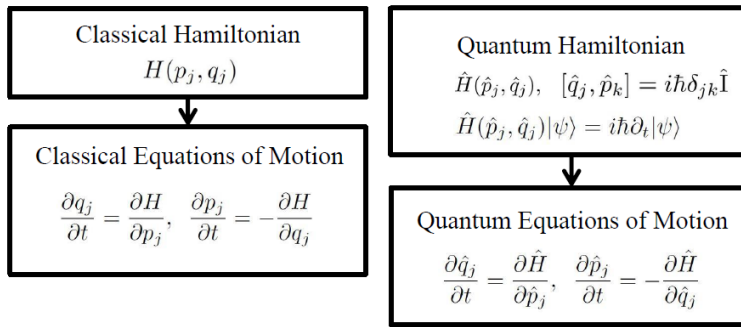
Figure 12.1: (Left) The classical Hamiltonian and the classical equations of motion. (Right) The quantum Hamiltonian and the quantum equations of motion. The quantum Hamiltonian is obtained from the classical Hamiltonian by elevating the conjugate variables p_j and q_j to become quantum operator, and endowing them with a commutation relation. The commutator induces derivative operator, making the quantum equations of motion strikingly similar to the classical ones. The figure is for many particle system, but it can be specialized to one particle system by letting $i = j$.

12.4 Quantization of Waves on a Linear Atomic Chain–Phonons

When a particle is in simple harmonic motion, quantum mechanics allows us to quantize its motion using Schrödinger equation for the harmonic oscillator. When we have an atomic chain, there will be coupling forces between the atoms. When one of the atoms is set into simple harmonic oscillation, the whole chain will be in simple harmonic oscillation due to coupling between them. Since the energy of one single particle is quantized in the quantum realm, we expect that harmonic oscillation of the whole chain to be quantized in energy as well. This is a much easier notion to accept if we accept that the energy of one particle is quantized.

Before proceeding further, it will be prudent to visit the physics of coupled LC tank

circuits. A lone tank circuit has only one resonant frequency. As soon as we put two tank circuits next to each other, the coupled system has two resonant frequencies. A chain of tank circuits coupled together will have many resonant frequencies or modes. A continuum of tank circuits forms the transmission line which has infinitely many propagating modes (see Figure 12.1).
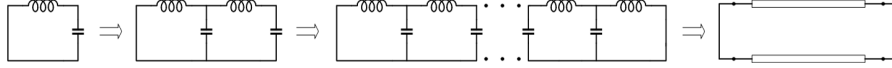
Figure 12.1: A transmission line of infinite length allows the propagation of a traveling wave of any frequency on it. It can be considered as due to a chain of coupled harmonic oscillators.
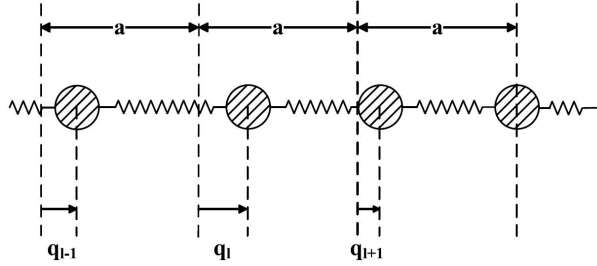
Figure 12.2: The linear atomic chain (redrawn from Haken).

For atoms on a linear atomic chain, assuming only nearest neighbor forces between the particles, the equation of motion, according to Newton's law, is

$$m_0\ddot{q}_\ell(t) = f\left[q_{\ell+1}(t) - q_\ell(t)\right] - f\left[q_\ell(t) - q_{\ell-1}(t)\right]$$
$$= f\left[q_{\ell+1}(t) - 2q_\ell(t) + q_{\ell-1}(t)\right] \tag{12.4.1}$$

The left-hand side is the inertial force needed to accelerate the atom at position ℓ. The right-hand side is the force exerted on the same atom in accordance to Hooke's law.

The above equation of motion can also be derived from the Hamiltonian of the system. First, one notes that the kinetic energy of the system is

$$T = \sum_{\ell=1}^{N} \frac{m_0}{2}\dot{q}_\ell^2 = \sum_{\ell=1}^{N} \frac{p_\ell^2}{2m_0} \tag{12.4.2}$$

while the potential energy

$$V = \frac{1}{2}f\sum_{\ell=1}^{N}\left(q_\ell - q_{\ell+1}\right)^2 \tag{12.4.3}$$

such that the force on the ℓ-th atom is

$$F_\ell = -\frac{\partial}{\partial q_\ell}V \tag{12.4.4}$$

The classical Hamiltonian of the system is then

$$H = T + V = \sum_{\ell=1}^{N} \frac{p_\ell^2}{2m_0} + \frac{1}{2} f \sum_{\ell=1}^{N} (q_\ell - q_{\ell+1})^2 \qquad (12.4.5)$$

Furthermore, we impose the periodic boundary condition so that

$$q_\ell = q_{\ell+N}, \qquad p_\ell = p_{\ell+N} \qquad (12.4.6)$$

so that the above summations need only be over one period of the linear atomic chain.

The equation of motion (12.4.1) can also be derived from the above using Hamilton equations of motion derived previously, and reproduced here as

$$\dot{q}_{\ell'} = \partial H / \partial p_{\ell'}, \qquad \dot{p}_{\ell'} = -\partial H / \partial q_{\ell'} \qquad (12.4.7)$$

It is to be noted that $H(q_1, q_2, \ldots, q_N; p_1, p_2, \ldots, p_N)$ is a functional.[‡] Moreover, since the conjugate variables for different atoms are independent of each other, the sifting property ensues, namely,

$$\partial p_\ell / \partial p_{\ell'} = \delta_{\ell, \ell'} \qquad (12.4.8)$$

Taking the first variation of the Hamiltonian with respect to the first variation of the conjugate variables p_ℓ and q_ℓ, one has

$$\delta H = \sum_{\ell=1}^{N} \frac{p_\ell}{m_0} \delta p_\ell + f \sum_{\ell=1}^{N} (q_\ell - q_{\ell+1})(\delta q_\ell - \delta q_{\ell+1}) \qquad (12.4.9)$$

Dividing the above by $\delta p_{\ell'}$ and $\delta q_{\ell'}$ respectively and letting these variations tend to zero, then

$$\frac{\partial H}{\partial p_{\ell'}} = \sum_{\ell} \left\{ \frac{p_\ell(t)}{m_0} \frac{\partial p_\ell}{\partial p_{\ell'}} + f (q_\ell - q_{\ell+1}) \left(\frac{\partial q_\ell}{\partial p_{\ell'}} - \frac{\partial q_{\ell+1}}{\partial p_{\ell'}} \right) \right\} \qquad (12.4.10)$$

$$\frac{\partial H}{\partial q_{\ell'}} = \sum_{\ell} \left\{ \frac{p_\ell(t)}{m_0} \frac{\partial p_\ell}{\partial q_{\ell'}} + f (q_\ell - q_{\ell+1}) \left(\frac{\partial q_\ell}{\partial q_{\ell'}} - \frac{\partial q_{\ell+1}}{\partial q_{\ell'}} \right) \right\} \qquad (12.4.11)$$

It is clear that due to the independence of q_ℓ and p_ℓ, and that these are independent particles,

$$\frac{\partial p_\ell}{\partial p_{\ell'}} = \frac{\partial q_\ell}{\partial q_{\ell'}} = \delta_{\ell\ell'} \qquad \frac{\partial q_\ell}{\partial p_{\ell'}} = \frac{\partial p_\ell}{\partial q_{\ell'}} = 0 \qquad (12.4.12)$$

where $\delta_{\ell\ell'}$ is the Kronecker delta function. Upon carefully applying the sifting property of the Kronecker delta function, and the Hamilton equations of motion (12.4.7), one has

$$\dot{q}_{\ell'}(t) = \frac{1}{m_0} p_{\ell'}(t), \quad \dot{p}_{\ell'}(t) = f [q_{\ell'+1}(t) - 2q_{\ell'}(t) + q_{\ell'-1}(t)] \qquad (12.4.13)$$

Combining the above, (12.4.1) is recovered.

[‡]A functional is a function of functions. The functions here are the vectors $\mathbf{q} = [q_1, q_2, \ldots, q_N]$, $\mathbf{p} = [p_1, p_2, \ldots, p_N]$.

12.4.1   Quantization: A Direct Approach

In the direct quantization approach, first, the conjugate variables, p_ℓ and q_ℓ are first elevated to be operators, $\hat{p}_\ell$ and $\hat{q}_\ell$. Then the classical Hamiltonian is elevated to be a quantum Hamiltonian, namely,

$$\hat{H} = \hat{T} + \hat{V} = \sum_{\ell=1}^{N} \frac{\hat{p}_\ell^2}{2m_0} + \frac{1}{2} f \sum_{\ell=1}^{N} \left(\hat{q}_\ell - \hat{q}_{\ell+1} \right)^2 \tag{12.4.14}$$

Furthermore, $\hat{p}_\ell$ and $\hat{q}_\ell$ are now endowed with a commutator

$$[\hat{q}_\ell, \hat{p}_{\ell'}] = i\hbar \hat{I} \delta_{\ell,\ell'} \tag{12.4.15}$$

The above equal-time commutator induces similar algebra as in the lone (single) quantum harmonic oscillator case, but with the presence of the Kronecker delta function δ_{ij} implying further that the oscillators i and j are independent of each other. Similar to the lone harmonic oscillator case, it can be shown that

$$[\hat{p}_{\ell'}, \hat{q}_\ell^n] = -in\delta_{\ell\ell'}\hat{q}_\ell^{n-1}\hbar = -i\hbar\frac{\partial}{\partial \hat{q}_{\ell'}}\hat{q}_\ell^n \Rightarrow \left[\hat{p}_{\ell'}, \hat{H}\right] = -i\hbar\frac{\partial}{\partial \hat{q}_{\ell'}}\hat{H} = i\hbar\dot{\hat{p}}_{\ell'}(t) \tag{12.4.16}$$

$$[\hat{q}_{\ell'}, \hat{p}_\ell^n] = in\delta_{\ell\ell'}\hat{p}_\ell^{n-1}\hbar = i\hbar\frac{\partial}{\partial \hat{p}_{\ell'}}\hat{p}_\ell^n \Rightarrow \left[\hat{q}_{\ell'}, \hat{H}\right] = i\hbar\frac{\partial}{\partial \hat{p}_{\ell'}}\hat{H} = i\hbar\dot{\hat{q}}_{\ell'}(t) \tag{12.4.17}$$

As is seen, the above yields the same algebra as the classical Hamiltonian case, yielding the quantum Hamilton equations of motion very similar to (12.4.7), but with the conjugate variables replaced by operators. Going through the same algebra, the quantum equation of motion can be derived:

$$m_0 \ddot{\hat{q}}_\ell(t) = f \left[\hat{q}_{\ell+1}(t) - 2\hat{q}_\ell(t) + \hat{q}_{\ell-1}(t) \right] \tag{12.4.18}$$

The above operators are in the Heisenberg picture, and they operate on a quantum state $|\psi_0\rangle$. Better still, they can be converted to observable quantities if their expectation values are taken with respect to a quantum state $|\psi_0\rangle$.

12.4.2   Quantization: Mode Decomposition Approach

In the spirit of separation of variables, an eigenmode of (12.4.1) or (12.4.13) is of the form

$$q_\ell(t) = q_{\nu,\ell} B_\nu(t) \tag{12.4.19}$$

where

$$\ddot{B}_\nu(t) = -\omega_\nu^2 B_\nu(t) \tag{12.4.20}$$

which is the equation of motion of a simple harmonic oscillator. Upon substituting the above into (12.4.1) gives rise to

$$-m_0 \omega_\nu^2 q_{\nu,\ell} = f \left[q_{\nu,\ell+1}(t) - 2q_{\nu,\ell}(t) + q_{\nu,\ell-1}(t) \right] \tag{12.4.21}$$

The above is a finite difference equation, which is an approximation to $-k^2 q_k(x) = \frac{d^2}{dx^2} q_k(x)$. Hence, (12.4.18) admits traveling wave or eigen solution of the form

$$q_{\nu,\ell} = \frac{1}{\sqrt{N}} e^{ik_\nu \ell a} \tag{12.4.22}$$

Furthermore, the periodic boundary condition is imposed so that

$$q_{\nu,\ell} = q_{\nu,\ell+N} \tag{12.4.23}$$

Consequently, due to this boundary condition,

$$k_\nu = \frac{2\nu\pi}{Na} \tag{12.4.24}$$

where ν is an integer ranging from $-\infty$ to ∞. Since k_ν is countable, we will replace the subscript ν with k as an index. In other words, $\nu \to k$ and $k_\nu \to k$. Furthermore, upon substituting the eigenmode solution (12.4.22) into (12.4.21) gives

$$-m_0 \omega_k^2 q_{k,\ell} = f\left(e^{ika} - 2 + e^{-ika}\right) q_{k,\ell} = -4f \sin^2\left(\frac{ka}{2}\right) q_{k,\ell} \tag{12.4.25}$$

yielding the dispersion relation, a relation between frequency ω_k and the wave number k, as

$$\omega_k^2 = \frac{4f}{m_0} \sin^2\left(\frac{ka}{2}\right) \tag{12.4.26}$$

If the inner product between two eigenmodes is defined to be

$$\langle q_k, q_{k'} \rangle = \sum_{\ell=1}^{N} q_{k,l}^* q_{k',l} \tag{12.4.27}$$

then they are orthonormal as

$$\langle q_k, q_{k'} \rangle = \sum_{\ell=1}^{N} q_{k,l}^* q_{k',l} = \sum_{\ell=1}^{N} \frac{1}{N} e^{i(k'-k)la} = \delta_{k,k'} \tag{12.4.28}$$

If $q_\ell(t)$ is expanded in a single mode only, it can be shown that

$$q_\ell(t) - q_{\ell+1}(t) = \left(q_{k,\ell} - q_{k,\ell+1}\right) B_k(t) = \left(1 - e^{ika}\right) q_{k,\ell} B_k(t) = -2i e^{ika/2} \sin(ka/2) q_{k,\ell} B_k(t) \tag{12.4.29}$$

Now, if q_ℓ is expanded in terms of sum of the eigenmodes, or

$$q_\ell(t) = \sum_k q_{k,\ell} B_k(t) = \sum_k q_{k,\ell}^* B_k^*(t) \tag{12.4.30}$$

where the second equality follows because $q_\ell(t)$ is real. Then the potential energy part of the Hamiltonian, after using (12.4.3), (12.4.29), and (12.4.30) is

$$
\begin{aligned}
V =& \frac{f}{2} \sum_{\ell=1}^{N} (q_\ell - q_{\ell+1})^2 \\
=& \frac{f}{2} \sum_{\ell=1}^{N} \left(\sum_{k} -2ie^{ika/2} \sin(ka/2) q_{k,l} B_k(t) \right) \left(\sum_{k'} 2ie^{-ik'a/2} \sin(k'a/2) q_{k',l}^* B_{k'}^*(t) \right) \\
=& \frac{f}{2} \sum_{k,k'} 4e^{ika/2} \sin(ka/2) e^{-ik'a/2} \sin(k'a/2) \left(\sum_{\ell=1}^{N} q_{k,l} q_{k',l}^* \right) B_k(t) B_{k'}^*(t) \\
=& \frac{f}{2} \sum_{k,k'} 4e^{ika/2} \sin(ka/2) e^{-ik'a/2} \sin(k'a/2) \delta_{kk'} B_k(t) B_{k'}^*(t) \\
=& \frac{f}{2} \sum_{k} 4 \sin^2(ka/2) |B_k(t)|^2 = \frac{m_0}{2} \sum_{k} \omega_k^2 |B_k(t)|^2
\end{aligned}
\tag{12.4.31}
$$

where (12.4.28) has been used above. Similarly, for the kinetic energy part of the Hamiltonian, it is

$$
\begin{aligned}
T &= \sum_{\ell=1}^{N} \frac{1}{2} m_0 \dot{q}_\ell^2 = \sum_{\ell=1}^{N} \frac{1}{2} m_0 \sum_{k,k'} q_{k,\ell} \dot{B}_k(t) q_{k',\ell}^* \dot{B}_{k'}^*(t) \\
&= \frac{1}{2} m_0 \sum_{k,k'} \dot{B}_k(t) \dot{B}_{k'}^*(t) \delta_{kk'} = \frac{1}{2} m_0 \sum_{k} |\dot{B}_k(t)|^2
\end{aligned}
\tag{12.4.32}
$$

Consequently,

$$
H = \frac{1}{2} m_0 \sum_{k} \left[|\dot{B}_k(t)|^2 + \omega_k^2 |B_k(t)|^2 \right]
\tag{12.4.33}
$$

where the first term in (12.4.33) corresponds to the kinetic energy while the second term corresponds to the potential energy of the k-th mode. A closer inspection of the above indicates that each of the modes behaves like a simple harmonic oscillator. Hence, the Hamiltonian of each mode resembles that of a simple harmonic oscillator. The total Hamiltonian is just the sum of the Hamiltonian of each simple harmonic oscillator coming from each mode.

12.4.2.1 More on Traveling Wave Modes

The general solution to (12.4.20) is

$$
B_k(t) = B_{k,+} e^{-i\omega_k t} + B_{k,-} e^{i\omega_k t}
\tag{12.4.34}
$$

Thus, from (12.4.30), after using (12.4.20)

$$
q_\ell(t) = \sum_{k} q_{k,\ell} B_k(t) = \frac{1}{\sqrt{N}} \sum_{k} \left(B_{k,+} e^{-i\omega_k t} e^{ik\ell a} + B_{k,-} e^{i\omega_k t} e^{ik\ell a} \right)
\tag{12.4.35}
$$

Since $\omega_{-k} = \omega_k$ from (12.4.26), the first and second sums physically represent plane wave modes propagating in all directions, and hence, only one of the two sums suffices to express the modal expansion of a general traveling wave solution. Therefore, when only one sum is needed to capture this physics, and it can be expressed as

$$q_\ell(t) = \frac{1}{\sqrt{N}} \sum_k B_k(0) e^{-i\omega_k t} e^{ik\ell a} \tag{12.4.36}$$

Alternatively, the above can be rewritten as

$$q_\ell(t) = \frac{1}{2\sqrt{N}} \sum_k \left[B_k(0) e^{-i\omega_k t} e^{ik\ell a} + B_k^*(0) e^{i\omega_k t} e^{-ik\ell a} \right] \tag{12.4.37}$$

The second term comes from that the left-hand side of (12.4.36) is real-valued, hence taking its complex conjugation yields the same value. The factor of $1/2$ can be removed if we only sum over positive k's. The last form is preferred so as to keep each mode real-valued. So for a traveling wave mode, the amplitude $B_k(t) = B_k(0) e^{-i\omega_k t}$ is a rotating wave. Moreover, using the above gives $|\dot{B}_k|^2 = \omega_k^2 |B_k|^2$ and hence, the Hamiltonian is always constants of time. Notice that the kinetic energy and the potential energy are equal to each other, a mark of an simple harmonic oscillator.

In this case, under the rotating wave picture, where $B_k(t) \sim e^{-i\omega_k t}$, the Hamiltonian of the k-th mode is

$$H_k = m_0 \omega_k^2 |B_k(t)|^2 = m_0 \omega_k^2 |B_k(0)|^2 \tag{12.4.38}$$

where $|B_k(t)|$, and hence the Hamiltonian, is time independent. Using Euler's identity for $e^{i\omega_k t} = \cos(\omega_k t) + i\sin(\omega_k t)$, one lets

$$\sqrt{2m_0\omega_k} B_k(t) = p_k(t) \pm i q_k(t)$$

or

$$2m_0\omega_k |B_k(t)|^2 = p_k^2(t) + q_k^2(t) \tag{12.4.39}$$

then the Hamiltonian for the k-th mode is

$$H_k = \frac{\omega_k}{2} \left[p_k^2(t) + q_k^2(t) \right] \tag{12.4.40}$$

In the above, $p_k(t)$ and $q_k(t)$ are sinusoidal functions in quadrature phase. The above is identical to the Hamiltonian of a classical harmonic oscillator for normalized coordinates and momenta. We can think of $p_k(t)$ and $q_k(t)$ to be the in phase and quadrature component of the displacement of the harmonic oscillator. Each of them is time dependent, but the sum of their squares is not time-dependent.

The reason why the above simple form ensues is:

1. By imposing periodic boundary condition on the linear atomic chain, only countable, discrete modes with wavenumber k are allowed to propagate on the linear atomic chain;

2. Only a single mode is allowed to propagate on the linear atomic chain so as to make the solution and the corresponding Hamiltonian very simple.

The physical picture we shall have of the above is that because only one mode exists, all the atoms on the chain are oscillating at the same frequency and in unison except for a phase lag. So each of them is in a simple harmonic motion. Since we have learned how to convert a classical harmonic oscillator to a quantum harmonic oscillator, the same principles can be applied to turn the above into a chain of coupled quantum harmonic oscillators.

To convert the above into a quantum harmonic oscillator, we elevate the above amplitudes to operators, namely

$$\hat{H}_k = \frac{\omega_k}{2} \left(\hat{p}_k^2 + \hat{q}_k^2 \right) \tag{12.4.41}$$

where

$$\hat{p}_k = -i\hbar \frac{d}{dq_k}, \qquad \hat{q}_k^2 = q_k^2 \tag{12.4.42}$$

in coordinate space representation. We can further define dimensionless entities

$$\xi_k = q_k/\sqrt{\hbar}, \qquad \hat{\Pi}_k = \hat{p}_k/\sqrt{\hbar} = -i\frac{d}{d\xi_k} \tag{12.4.43}$$

so that

$$\hat{H}_k = \frac{\hbar\omega_k}{2} \left(-\frac{d^2}{d\xi_k^2} + \xi_k^2 \right) = \frac{\hbar\omega_k}{2} (\hat{\Pi}_k^2 + \xi_k^2) \tag{12.4.44}$$

$$\tag{12.4.45}$$

The above is the quantum analogue of (12.4.40). Furthermore, we can show that

$$\hat{a}_k^\dagger = \frac{1}{\sqrt{2}} \left(-i\hat{\Pi}_k + \hat{\xi}_k \right)$$

$$\hat{a}_k(t) = \frac{1}{\sqrt{2}} \left(i\hat{\Pi}_k(t) + \hat{\xi}_k(t) \right) \tag{12.4.46}$$

$$\left(\hat{\Pi}_k(t), \hat{\xi}_k(t) \right) \tag{12.4.47}$$

Hence, by direct use of the above, we can show that

$$\hat{H}_k = \frac{\hbar\omega_k}{2} \left(\hat{a}_k^\dagger \hat{a}_k + \hat{a}_k \hat{a}_k^\dagger \right) \tag{12.4.48}$$

The above can be transformed to the Heisenberg picture using the unitary transformation of the time-evolution operator yielding

$$\hat{H}_k = \frac{\hbar\omega_k}{2} \left[\hat{a}_k^\dagger(t)\hat{a}_k(t) + \hat{a}_k(t)\hat{a}_k^\dagger(t) \right] \tag{12.4.49}$$

In the above, the operators are in the rotating wave picture, with time dependence proportional to $e^{\pm i\omega_k t}$. Hence, $\hat{a}_k^\dagger(t)$ and $\hat{a}_k(t)$ are the quantum analogue of $B_k(t)$ and $B_k^*(t)$ which are also expressed in the rotating wave picture. Therefore, we can rewrite (12.4.38) more suggestively as

$$H_k = \frac{1}{2}m_0\omega_k^2\left[B_k^*(t)B_k(t) + B_k(t)B_k^*(t)\right] \tag{12.4.50}$$

The quantum analogue of the above is to elevate the B_k and B_k^* to operators or

$$\hat{H}_k = \frac{1}{2}m_0\omega_k^2\left[\hat{B}_k^\dagger(t)\hat{B}_k(t) + \hat{B}_k(t)\hat{B}_k^\dagger(t)\right] \tag{12.4.51}$$

We can identify that

$$\hat{B}_k(t) = \sqrt{\frac{\hbar}{m_0\omega_k}}\,\hat{a}_k(t)$$

$$\hat{B}_k^\dagger(t) = \sqrt{\frac{\hbar}{m_0\omega_k}}\,\hat{a}_k^\dagger(t) \tag{12.4.52}$$

by comparing (12.4.51) with (12.4.48).

Had we assumed that the solutions are summation of different k modes, then (12.4.37), after elevation to quantum operators, becomes

$$\hat{q}_\ell(t) = \sum_k \frac{1}{2}\sqrt{\frac{\hbar}{m_0\omega_k N}}\,e^{ik\ell a}\hat{a}_k(t) + \text{c.c.} \tag{12.4.53}$$

$$\hat{p}_\ell(t) = \sum_k -i\frac{1}{2}\sqrt{\frac{\hbar m_0\omega_k}{N}}\,e^{ik\ell a}\hat{a}_k(t) + \text{c.c.} \tag{12.4.54}$$

The annihilation operator for the k-th mode is shown to have the time dependence of $\exp(-i\omega t)$ in (12.3.11). Hence, the above can be rewritten as

$$\hat{q}_\ell(t) = \sum_k \frac{1}{2}\sqrt{\frac{\hbar}{m_0\omega_k N}}\,e^{ik\ell a - i\omega_k t}\hat{a}_k + \text{c.c.} \tag{12.4.55}$$

$$\hat{p}_\ell(t) = \sum_k -i\frac{1}{2}\sqrt{\frac{\hbar m_0\omega_k}{N}}\,e^{ik\ell a - i\omega_k t}\hat{a}_k + \text{c.c.} \tag{12.4.56}$$

where $\hat{a}_k = \hat{a}_k(0)$.

The above can also be directly derived from the quantum equation of motion, (12.4.18). It also represents a multitude of propagating modes with different k wavenumbers. They can come together to form a wave packet that more realistically represents a wave that propagates on an infinite linear atomic chain. A packet of energy then satisfies locality and causality if it is associated with a wave packet.

12.5 The Continuum Limit

In the continuum limit, we let

$$m_0 = \rho a \tag{12.5.1}$$

Then (12.4.1) becomes

$$\rho a \ddot{q}(x,t) = f a^2 \frac{\partial^2}{\partial x^2} q(x,t) \tag{12.5.2}$$

or

$$\rho \ddot{q}(x,t) = g \frac{\partial^2}{\partial x^2} q(x,t) \tag{12.5.3}$$

where $g = fa$. Defining the momentum density to be

$$\pi(x,t) = \rho \dot{q}(x,t) \tag{12.5.4}$$

then a Hamiltonian that will lead to (12.5.3) is

$$H = \frac{1}{2\rho} \int_0^L \pi^2(x,t) dx + \frac{g}{2} \int_0^L \left[\frac{\partial q(x,t)}{\partial x} \right]^2 dx \tag{12.5.5}$$

The Hamilton equations of motion in the continuum limit has to be expressed with functional derivatives. Analogous to the discrete case expounded in (12.4.7), they are

$$\frac{\partial q(x,t)}{\partial t} = \frac{\delta H}{\delta \pi(x,t)}, \qquad \frac{\partial \pi(x,t)}{\partial t} = -\frac{\delta H}{\delta q(x,t)} \tag{12.5.6}$$

The above are the Hamilton equations in the continuum limit expressed in terms of functional derivatives. Furthermore, analogous to (12.4.12) for the discrete case, the continuum case becomes

$$\frac{\delta q(x,t)}{\delta q(x',t)} = \frac{\delta \pi(x,t)}{\delta \pi(x,t)} = \delta(x - x'), \qquad \frac{\delta q(x,t)}{\delta \pi(x',t)} = \frac{\delta \pi(x,t)}{\delta q(x',t)} = 0 \tag{12.5.7}$$

To obtain the functional derivatives, one takes the first variation of (12.5.5) to get

$$\delta H = \frac{1}{\rho} \int_0^L dx \pi(x,t) \delta \pi(x,t) + g \int_0^L dx \frac{\partial q(x,t)}{\partial x} \frac{\partial \delta q(x,t)}{\partial x} \tag{12.5.8}$$

After using integration by parts then

$$\delta H = \frac{1}{\rho} \int_0^L dx \pi(x,t) \delta \pi(x,t) - g \int_0^L dx \frac{\partial^2 q(x,t)}{\partial x^2} \delta q(x,t) \tag{12.5.9}$$

Dividing the above by the first variations in $\pi(x,t)$ and $q(x,t)$, respectively gives rise to

$$\frac{\delta H}{\delta \pi(x',t)} = \frac{1}{\rho} \int_0^L dx \pi(x,t) \frac{\delta \pi(x,t)}{\delta \pi(x',t)} - g \int_0^L dx \frac{\partial^2 q(x,t)}{\partial x^2} \frac{\delta q(x,t)}{\delta \pi(x',t)} \tag{12.5.10}$$

$$\frac{\delta H}{\delta q(x',t)} = \frac{1}{\rho} \int_0^L dx \pi(x,t) \frac{\delta \pi(x,t)}{\delta q(x',t)} - g \int_0^L dx \frac{\partial^2 q(x,t)}{\partial x^2} \frac{\delta q(x,t)}{\delta q(x',t)} \qquad (12.5.11)$$

Making use of (12.5.6) and (12.5.7) yields

$$\frac{\partial q(x,t)}{\partial t} = \frac{\pi(x,t)}{\rho}, \qquad \frac{\partial \pi(x,t)}{\partial t} = g \frac{\partial^2 q(x,t)}{\partial x^2} \qquad (12.5.12)$$

Combining the above gives rise to (12.5.3).

12.5.1   Quantization in the Continuum Limit

In quantization, the conjugate variables π and q are first elevated to be quantum operators. Then the Hamiltonian also become a quantum operator, namely,

$$\hat{H} = \frac{1}{2\rho} \int_0^L \hat{\pi}^2(x,t) dx + \frac{g}{2} \int_0^L \left(\frac{\partial \hat{q}(x,t)}{\partial x} \right)^2 dx \qquad (12.5.13)$$

The conjugate variables are endowed with the commutator:

$$[\hat{\pi}(x,t), \hat{q}(x',t)] = i\hbar \delta(x-x') \hat{I} \qquad (12.5.14)$$

where the Kronecker delta function becomes a Dirac delta function in the continuum case. With this commutation relation, similar to the discrete case in (12.4.16) and (12.4.17), it is quite easy to show that

$$[\hat{\pi}(x',t), \hat{q}^n(x,t)] = -i\hbar \delta(x-x') n \hat{q}^{n-1}(x,t) = -i\hbar \frac{\delta \hat{q}^n(x,t)}{\delta \hat{q}(x',t)}$$

$$\Rightarrow \left[\hat{\pi}(x',t), \hat{H} \right] = -i\hbar \frac{\delta \hat{H}}{\delta \hat{q}(x',t)} = i\hbar \dot{\hat{\pi}}(x',t)$$

$$(12.5.15)$$

$$[\hat{q}(x',t), \hat{\pi}^n(x,t)] = i\hbar \delta(x-x') n \hat{\pi}^{n-1}(x,t) = i\hbar \frac{\delta \hat{\pi}^n(x,t)}{\delta \hat{\pi}(x',t)}$$

$$\Rightarrow \left[\hat{q}(x',t), \hat{H} \right] = i\hbar \frac{\delta \hat{H}}{\delta \hat{\pi}(x',t)} = i\hbar \dot{\hat{q}}(x',t)$$

$$(12.5.16)$$

where the functional derivatives of operators have meaning only when they act on the eigenvectors of these operators. With the use of the above, one arrives at the quantum equation of motion:

$$\rho \ddot{\hat{q}}(x,t) = g \frac{\partial^2}{\partial x^2} \hat{q}(x,t) \qquad (12.5.17)$$

12.5.1.1 Expansion in terms of traveling wave modes

As before, the field can be expanded in terms of traveling wave modes, and the solution of the quantized fields can be written as

$$\hat{q}(x,t) = \sum_k \frac{1}{2} \sqrt{\frac{\hbar}{\rho \omega_k L}} e^{ikx - i\omega_k t} \hat{a}_k + \text{c.c.} \tag{12.5.18}$$

$$\hat{\pi}(x,t) = -i \sum_k \frac{1}{2} \sqrt{\frac{\hbar \omega_k \rho}{L}} e^{ikx - i\omega_k t} \hat{a}_k + \text{c.c.} \tag{12.5.19}$$

To obtain the above from the discrete case, one needs to replace $m_0 = \rho a$, $N = L/a$, and $\ell a = x$.

12.6 Quantization of Electromagnetic Field

For electromagnetic field, the harmonic oscillators are dipoles in a medium that are polarized by an electric field. In general, the electric flux in a medium is given by

$$\mathbf{D} = \epsilon_0 \mathbf{E} + \mathbf{P} \tag{12.6.1}$$

The first term is the contribution to the electric flux $\mathbf{D}$ due to vacuum, while the second term is the polarization density contributed from dipoles polarized in a material medium by an electric field. It is customary to write $\mathbf{P} = \epsilon_0 \chi \mathbf{E}$ to indicate that the polarization density is proportional to the electric field.

The time variation of the electric flux, $\partial_t \mathbf{D}$, is instrumental in giving rise to displacement current, and hence, yielding a propagating wave. This electric flux exists even in vacuum. Hence, we can imagine that even vacuum is polarized by an electric field to produce dipole density.[§] In other words, the oscillating electric dipole will produce a magnetic field creating an inductive effect. Together, they can form a resonating system behaving like a harmonic oscillator. They resemble a tiny LC tank circuit.

Maxwell's equations are

$$\nabla \times \mathbf{E} = -\partial_t \mathbf{B}, \qquad \frac{1}{\mu} \nabla \times \mathbf{B} = \partial_t \mathbf{D} + \mathbf{J} \tag{12.6.2}$$

$$\nabla \cdot \mathbf{B} = 0, \qquad \nabla \cdot \mathbf{D} = \varrho \tag{12.6.3}$$

where $\mathbf{B} = \mu_0 \mathbf{H}$ and $\mathbf{D} = \epsilon \mathbf{E}$. We can assume that ϵ is a constant representing a homogeneous medium, as the mathematics is not any more difficult. With the potential approach, one lets

$$\mathbf{B} = \nabla \times \mathbf{A}, \qquad \mathbf{E} = -\dot{\mathbf{A}} - \nabla \Phi \tag{12.6.4}$$

[§]It is debatable as to what gives rise to these dipoles in vacuum. It could be electron-positron pairs that are embedded in vacuum. When vacuum is bombarded with an energetic photon, it is known to produce an electron-position pair.

When the first equation above is substituted into the first equation in (12.6.2), the second equation in (12.6.4) above for $\mathbf{E}$ is obtained. The above can be substituted into the second equation in (12.6.2), and with the use of Lorenz gauge that

$$\nabla \cdot \mathbf{A} = -\mu_0 \epsilon \dot{\Phi} \tag{12.6.5}$$

then one arrives at

$$\nabla \times \nabla \times \mathbf{A}(\mathbf{r}, t) - \nabla(\nabla \cdot \mathbf{A}(\mathbf{r}, t)) + \mu_0 \epsilon \partial_t^2 \mathbf{A}(\mathbf{r}, t) = \mu_0 \mathbf{J}(\mathbf{r}, t) \tag{12.6.6}$$

With the second equation in (12.6.4) substituted into the second equation in (12.6.3) plus the use of the Lorenz gauge, then

$$\mu_0 \epsilon \partial_t^2 \Phi(\mathbf{r}, t) - \nabla^2 \Phi(\mathbf{r}, t) = \varrho(\mathbf{r}, t)/\epsilon \tag{12.6.7}$$

In the case where the sources are at infinity, the $\Phi = 0$ gauge can be used. In this case, the Lorenz gauge is the same as the Coulomb gauge, namely, $\nabla \cdot \mathbf{A} = 0$.

12.6.1   Hamiltonian

Before quantizing Maxwell's equations, the Hamiltonian of the system should be derived. In general, the Hamiltonian represents the total energy of the system, and it is given by

$$H = \int d\mathbf{r} \mathscr{H}(\mathbf{r}) \tag{12.6.8}$$

where

$$\mathscr{H}(\mathbf{r}) = \frac{1}{2} \left[\epsilon \mathbf{E}^2 + \frac{1}{\mu_0} \mathbf{B}^2 \right] \tag{12.6.9}$$

In the potential formulation, the above becomes

$$\mathscr{H}(\mathbf{r}) = \frac{1}{2} \left[\epsilon (\dot{\mathbf{A}})^2 + \frac{1}{\mu_0} (\nabla \times \mathbf{A})^2 \right] \tag{12.6.10}$$

As in the linear atomic chain case, we will treat a one dimensional equivalence of the above to arrive at

$$H = \int_0^L dx \frac{1}{2} \left[\epsilon (\dot{A}_z)^2 + \frac{1}{\mu_0} (\partial_x A_z)^2 \right] \tag{12.6.11}$$

The conjugate momentum can be defined to be $\Pi_z = \epsilon \dot{A}_z$, and the above becomes

$$H = \int_0^L dx \frac{1}{2} \left[\frac{1}{\epsilon} (\Pi_z)^2 + \frac{1}{\mu_0} (\partial_x A_z)^2 \right] \tag{12.6.12}$$

With the mapping

$$\epsilon \Rightarrow \rho, \quad \frac{1}{\mu_0} \Rightarrow g, \quad \Pi_z \Rightarrow \pi, \quad A_z \Rightarrow q \tag{12.6.13}$$

the above is entirely analogous to the linear atomic chain case. The equation of motion for A_z can be derived as

$$\partial_x^2 A_z = \mu_0 \epsilon \ddot{A}_z \tag{12.6.14}$$

Hence, quantization follows the same path as before, and one obtains

$$\hat{A}_z(x,t) = \sum_k \frac{1}{2} \sqrt{\frac{\hbar}{\omega_k L \epsilon_0}} \hat{a}_k e^{ikx - i\omega_k t} + \text{c.c.} \tag{12.6.15}$$

Then the electric field E_z can be derived yielding

$$\hat{E}_z(x,t) = \frac{i}{2} \sqrt{\frac{\hbar \omega_k}{L \epsilon_0}} \hat{a}_k(t) e^{ikx} + \text{c.c.} \tag{12.6.16}$$

Similarly, the B_z can be derived giving

$$\hat{B}_y(x,t) = \sum_k \frac{i}{2} \sqrt{\frac{\hbar \omega_k}{L \epsilon_0}} \frac{1}{c} \hat{a}_k e^{ikx - i\omega_k t} + \text{c.c.} \tag{12.6.17}$$

The above can be generalized to a plane wave mode propagating in arbitrary direction, namely,

$$\hat{\mathbf{E}}(\mathbf{r},t) = \sum_{\mathbf{k},s} \frac{i}{2} \sqrt{\frac{\hbar \omega_k}{V \epsilon_0}} \mathbf{e}_s \hat{a}_{\mathbf{k},s} e^{i\mathbf{k}\cdot\mathbf{r} - i\omega_k t} + \text{c.c.} \tag{12.6.18}$$

$$\hat{\mathbf{B}}(\mathbf{r},t) = \sum_{\mathbf{k},s} \frac{i}{2} \sqrt{\frac{\hbar \omega_k}{V \epsilon_0}} \frac{1}{c} \mathbf{e_k} \times \mathbf{e}_s \hat{a}_{\mathbf{k},s} e^{i\mathbf{k}\cdot\mathbf{r} - i\omega_k t} + \text{c.c.} \tag{12.6.19}$$

where $\mathbf{e}_s$ is a unit vector denoting the polarization of the electric field, and $\mathbf{e_k}$ is a unit vector pointing in the $\mathbf{k}$ direction. Hence, the above is orthogonal to the $\mathbf{k}$ vector. It can be either linearly polarized or circularly polarized. In the linearly polarized case, $\mathbf{e}_s$ is either vertical or horizontal. In the circularly polarized case, $\mathbf{e}_s$ is either right-hand or left-hand circularly polarized. In the above, we have defined the Hamiltonian as an integral over a volume V, and hence, the normalization above is with respect to V. The above equations are constructed so that if their expectation values are found with respect to a quantum state $|\psi_0\rangle$, they yield real value fields. Hence, these field operators are Hermitian operators.

The multitude of modes above allows a photon that is emitted by an atom, say by spontaneous emission, to form a wave packet that is localized in time and space. This makes the photon emission a causal event. Moreover, this wave packet propagates through space with the speed of light. Also, this wave packet zips through space with the aid of the coupled quantum harmonic oscillators.

It is seen that if the photon is purely monochromatic, it will be non-causal, and cannot represent the emission from an atomic transition. Hence, it is more appropriate to think of a photon as quasi-monochromatic with a linear superposition of modes with different wavenumbers k, allowing a localized wave packet to be formed. The fact that photons are

particles cannot be demonstrated by the mathematics expressed in the above equations. It is experimental evidence that has demonstrated that photons are particles, as much as electrons are particles.

A photon can be thought of as a "particle" possessing an expected or average packet of energy $\hbar\omega_e$, but it exists as a linear superposition of different modes with different k's, each of which zips through space with the velocity of light. Each mode can have a energy $\hbar\omega_k$ associated with it, but as the particle is a linear superposition of different modes, $\hbar\omega_e$ is the expectation value of the photon energy when the photon is in a superposition of different quantum modes.

It is interesting to note that even though both electrons and photons display wave-particle duality, but the discovery of their wave-particle duality nature follows a different path. Electrons were thought to be particles very early in time, but compelling experimental evidence showed that they possessed wave-like properties. On the other hand, photons, first manifested as light, were first thought to be wave, but compelling experimental evidence showed that they possessed particle-like properties. Now photon is accepted as one of the elementary particles, or quantum particle, displaying quantum properties.

From the above equations, it can be shown easily that

$$\nabla \times \hat{\mathbf{E}}(\mathbf{r}, t) = -\partial_t \hat{\mathbf{B}}(\mathbf{r}, t), \qquad \frac{1}{\mu_0} \nabla \times \hat{\mathbf{B}}(\mathbf{r}, t) = \partial_t \hat{\mathbf{D}}(\mathbf{r}, t) \qquad (12.6.20)$$

where $\hat{\mathbf{D}}(\mathbf{r}, t) = \epsilon_0 \hat{\mathbf{E}}(\mathbf{r}, t)$. The above are quantum Maxwell's equations where the classical fields are replaced with field operators.

12.6.1.1　Quantized Maxwell's equations to photon versus Schrödinger equation to electron

¶It is often asked if Maxwell's equations to photon is the same as Schrödinger equation to electron. They are not exactly identical, or in a sense, this is like comparing apples and oranges. Both Maxwell's equations and Schrödinger equation endow their corresponding particles with wave nature. Both equations associate the particles with a wave field which displays wave physics. Both particles, photon and electron, are quantum particles subject to quantum interpretation.

The field operator, for example, $\hat{\mathbf{E}}$, is associated with a mean and a standard deviation or fluctuation. The mean, for example, is obtained by taking the expectation value of the field operator by $\langle\psi_0|\hat{\mathbf{E}}|\psi_0\rangle$ where $|\psi_0\rangle$ is the quantum state of the photon For non-classical photons, such as those in the photon number state, this value is always zero. To obtain non-zero value, the photon can be assumed to be in the coherent state. On the other hand, the Schrödinger field of an electron does not have a mean or a standard deviation.

An electron can be detected by its scattering or interaction with a Coulomb potential for instance. But a photon is unaffected by a Coulomb potential, and its detection is by its interaction with other dipoles, such as a dielectric medium or an atom. The probability of detection of an electron at location $\mathbf{r}$ can be associated with $|\psi_0(\mathbf{r})|^2$. By the same token,

---

¶This section is written with input from E. Kudeki and A.Y. Liu.

the probability of detecting a photon at $\mathbf{r}$ by dipole interaction is proportional to $\langle \psi_i | \hat{\mathbf{E}}^\dagger(\mathbf{r}) \cdot \hat{\mathbf{E}}(\mathbf{r}) | \psi_i \rangle$ where $|\psi_i\rangle$ is the quantum state of the photon (see Gerry and Knight, Chapter 5).

The wave nature of their detection will emerge due to the nature of the associated wave fields, e.g., in the double slit interference experiment. Otherwise, these are quite different wave fields. Moreover, when the photon wave field operator acts on a quantum state $\hat{\mathbf{E}}|\Psi_0\rangle$, it gives rise to an infinite dimensional state vector which still bears quantum information.

The quantization of electromagnetic field, also called quantum electrodynamics (QED), stimulated the growth of quantum field theory and particle physics. The Schrödinger theory of electron and the pertinent quantum mechanics is called first quantization theory, while the quantum field theory of electron is called second quantization theory. Hence, QED can be thought of as a second quantized version of photon theory, and there has been interest in finding the first quantized version of photon theory, very much akin to Schrödinger theory is for electron (see O. Keller, *Quantum Theory of Near Field Electrodynamics*).

12.6.2 Multimode Case and Fock State

When polarization is considered, we may write the Hamiltonian for a single mode photon as

$$\hat{H}_{\mathbf{k}} = \sum_s \hat{H}_{\mathbf{k},s} = \sum_s \hbar\omega_k \left(\hat{a}^\dagger_{\mathbf{k},s} \hat{a}_{\mathbf{k},s} + \frac{1}{2} \right) \tag{12.6.21}$$

where s stands for either horizontal or vertical polarization. The above can be easily derived by considering the electric field with both polarizations present.

In the above, we can denote eigenstate of the Hamiltonian as[||]

$$|\psi\rangle = |n_v\rangle|n_h\rangle \tag{12.6.22}$$

or better still

$$|\psi\rangle = |n_v, n_h\rangle \tag{12.6.23}$$

where v stands for vertical polarization and h stands for horizontal polarization. In the above, there are n_v photons in the vertical polarization state and there are n_h photons in the horizontal polarization state.

When many modes are considered simultaneously, it can be shown that the corresponding Hamiltonian is

$$\hat{H} = \sum_{\mathbf{k},s} \hat{H}_{\mathbf{k},s} = \sum_{\mathbf{k},s} \hbar\omega_{\mathbf{k}} \left(\hat{a}^\dagger_{\mathbf{k},s} \hat{a}_{\mathbf{k},s} + \frac{1}{2} \right) \tag{12.6.24}$$

The eigenstate of the above Hamiltonian can be written as

$$|\psi\rangle = \prod_{\mathbf{k},s} |n_{\mathbf{k},s}\rangle \tag{12.6.25}$$

[||]When the Hamiltonian consists of sum of the Hamiltonians of two quantum systems, the states of the Hamiltonian operator can be represented by the product space of the states of the individual quantum system.

The above can be written alternatively as

$$|\psi\rangle = \prod_s |n_{\mathbf{k}_1,s}\rangle|n_{\mathbf{k}_2,s}\rangle|n_{\mathbf{k}_3,s}\rangle\cdots = \prod_s |n_{\mathbf{k}_1,s}, n_{\mathbf{k}_2,s}, n_{\mathbf{k}_3,s}n_{\mathbf{k}_4,s}\cdots\rangle = \prod_s |\{n_{\mathbf{k},s}\}\rangle \quad (12.6.26)$$

or better still

$$|\psi\rangle = |\{n_{\boldsymbol{\kappa}}\}\rangle \tag{12.6.27}$$

where $\boldsymbol{\kappa}$ stands for all possible combinations of $\mathbf{k}, s$. The above are eigenstates of the corresponding Hamiltonian operator. A general quantum state can be expanded in terms of the eigenstates of the system as

$$|\psi\rangle = \sum_{\boldsymbol{\kappa}} a_{\boldsymbol{\kappa}}|\{n_{\boldsymbol{\kappa}}\}\rangle \tag{12.6.28}$$

The previously mentioned eigenstate is known as a Fock state or the occupational number state. It is customary to leave out the states where the occupational number is zero, listing only the states with non-zero number of photons. For example

$$|\psi\rangle = |3_{\mathbf{k}_1,v}, 2_{\mathbf{k}_2,h}, 1_{\mathbf{k}_3,v}\rangle \tag{12.6.29}$$

indicates a quantum eigenstate with three V (vertical) polarized photons in $\mathbf{k}_1$ mode, two H (horizontal) polarized photons in $\mathbf{k}_2$ mode, and one V polarized photon in $\mathbf{k}_3$ mode.

12.6.3   One-Photon State

The understanding of the one-photon state is important because of its usefulness in quantum information, quantum communication, and quantum cryptography. It is prudent to elaborate on it further. For example, when only one $\mathbf{k}$ mode is present, to denote a photon in the vertical polarization state, and no photons in the horizontal polarization state, the state vector should be

$$|\psi\rangle = |1_v\rangle|0_h\rangle \tag{12.6.30}$$

or just

$$|\psi\rangle = |1_v, 0_h\rangle \tag{12.6.31}$$

Often, this is just written as

$$|\psi\rangle = |1_v\rangle \tag{12.6.32}$$

where the second part is understood. For a photon in arbitrary polarization, it is written as

$$|\psi\rangle = a_v|1_v\rangle + a_h|1_h\rangle \tag{12.6.33}$$

The above denotes a one-photon state where the photon is in the linear superposition of two orthonormal one-photon states. Hence,

$$|a_v|^2 + |a_h|^2 = 1 \tag{12.6.34}$$

Since a photon is a package of energy propagating through space via coupled quantum harmonic oscillators, the above is subject to quantum interpretation. The photon is in the linear superposition of two quantum states $|1_v\rangle$ and $|1_h\rangle$.

Often, photons are generated by atomic transitions whereby the electron changes from a high energy state to a low energy state giving rise to one or more photons. Since this is a causal event, the wave field associated with a photon is a localized wave in space time. The localized wave is formed by linear superposing wave field with different $\mathbf{k}$ values. For instance, a localized one-photon, vertically polarized state is a linear superposition of a number of one-photon states with different $\mathbf{k}$ values as follows:

$$|\psi\rangle = \sum_{\mathbf{k}} a_{\mathbf{k},v}|1_{\mathbf{k},v}\rangle \tag{12.6.35}$$

The wave field associated with the above state can be localized in space time, and the wave field satisfies causality.

It is to be noted that (12.6.24) is a quantum Hamiltonian involving the sum over individual Hamiltonians of different quantum systems. The quantum system arises from a set of spatially coupled quantum harmonic oscillators. The choice of appropriate modes "diagonalizes" the system, giving rise to apparently uncoupled systems. When a photon is in a quantum state, it can be in a linear superposition of different quantum states of these different quantum systems. Therefore, one-photon states (12.6.33) and (12.6.35) should be thought of as a linear superposition of different quantum states, subject to the quantum interpretation of quantum mechanics. The particle is in a linear superposition of states before the measurement, and it collapses to one of the states after the measurement.

Even though a photon is associated with a packet of energy, when it is detected with a polarizer (that detects it either in the vertical or horizontal polarized state), it is found either in one state or the other. The packet of energy is never split between the two states indicated by (12.6.33). Hence, experiment evidence suggests that a photon is a quantum particle in the sense that an electron is a quantum particle at the quantum level. The same interpretation applies to (12.6.35). The subject of quantum interpretation will be discussed later.

12.6.4   Coherent State Revisited

The expectation value of the annihilation and creation operators with respect to the photon number states is always zero. Therefore, the expectation value of the field operators, which are proportional to the annihilation and creation operators, is always zero. Hence, the photon number states are non-classical. To arrive at an expectation value of field operators that are non-zero, and resemble a classical field, one has to work with the coherent states.

We have studied the coherent state previously. It is defined to be the eigenstate of the annihilation operator. Namely, if $|\alpha\rangle$ represents a coherent state, then

$$\hat{a}|\alpha\rangle = \alpha|\alpha\rangle \tag{12.6.36}$$

where we have used the eigenvalue α to index the coherent state. Since the number state $|n\rangle$

is complete, we can expand the coherent state in terms of the number state, or

$$|\alpha\rangle = \sum_{n=0}^{\infty} C_n |n\rangle \tag{12.6.37}$$

When the annihilation operator is applied to both sides, we obtain

$$\hat{a}|\alpha\rangle = \sum_{n=0}^{\infty} C_n \hat{a}|n\rangle = \sum_{n=1}^{\infty} C_n \hat{a}|n\rangle = \sum_{n=1}^{\infty} C_n \sqrt{n}|n-1\rangle = \sum_{n=0}^{\infty} C_{n+1}\sqrt{n+1}|n\rangle \tag{12.6.38}$$

where we have used $\hat{a}|0\rangle = 0$, $\hat{a}|n\rangle = \sqrt{n}|n-1\rangle$. Equating the above with $\alpha|\alpha\rangle$, we have

$$\sum_{n=0}^{\infty} C_{n+1}\sqrt{n+1}|n\rangle = \alpha \sum_{n=0}^{\infty} C_n|n\rangle \tag{12.6.39}$$

By the orthogonality of the number states, matching the coefficients, we have $C_{n+1} = \alpha C_n/\sqrt{n+1}$, or

$$C_n = C_{n-1}\alpha/\sqrt{n} = C_{n-2}\alpha^2/\sqrt{(n(n-1))} = C_0\alpha^n/\sqrt{n!} \tag{12.6.40}$$

Consequently,

$$|\alpha\rangle = C_0 \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}}|n\rangle \tag{12.6.41}$$

The above can be normalized to show that $C_0 = \exp(-|\alpha|^2/2)$.

The photon number operator is $\hat{n} = \hat{a}^\dagger \hat{a}$. It can be shown that the average number of photons associated with a coherent state is given by

$$\langle \hat{n} \rangle = \bar{n} = \langle \alpha|\hat{n}|\alpha\rangle = \langle \alpha|\hat{a}^\dagger \hat{a}|\alpha\rangle = |\alpha|^2 \tag{12.6.42}$$

where $\langle \alpha|\hat{a}^\dagger = \langle \alpha|\alpha^*$ is used. Moreover

$$\langle \hat{n}^2 \rangle = \langle \alpha|\hat{n}^2|\alpha\rangle = \langle \alpha|\hat{a}^\dagger \hat{a}\hat{a}^\dagger \hat{a}|\alpha\rangle = \langle \alpha|\hat{a}^\dagger \hat{a}^\dagger \hat{a}\hat{a} + \hat{a}^\dagger \hat{a}|\alpha\rangle = |\alpha|^4 + |\alpha|^2 = \bar{n}^2 + \bar{n} \tag{12.6.43}$$

where we have used $\hat{a}\hat{a}^\dagger - \hat{a}^\dagger \hat{a} = 1$. Then

$$\Delta n = \sqrt{\langle \hat{n}^2 \rangle - \langle \hat{n} \rangle^2} = \bar{n}^{1/2} \tag{12.6.44}$$

The above is characteristic of a Poisson process. The probability of detecting n photons in a coherent state is

$$P_n = |\langle n|\alpha\rangle|^2 = e^{-|\alpha|^2}\frac{|\alpha|^{2n}}{n!} = e^{-\bar{n}}\frac{\bar{n}^n}{n!} \tag{12.6.45}$$

typical of a Poisson distribution.** The fractional uncertainty in the photon number is

$$\frac{\Delta n}{\bar{n}} = \frac{1}{\sqrt{\bar{n}}} \tag{12.6.46}$$
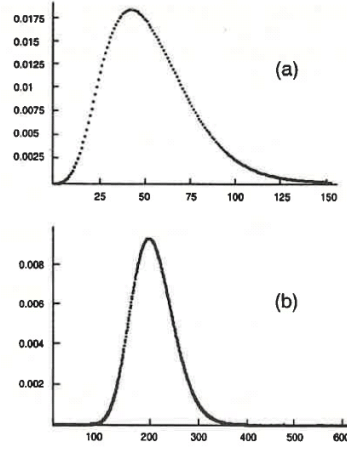
Figure 12.1: Two typical Poisson distributions. Case (a) has $\bar{n} = 50$ while case (b) has $\bar{n} = 200$ (from Haus).

On the other hand, for a photon number state $|n'\rangle$, it can be shown that

$$P_n = |\langle n|n'\rangle|^2 = \delta_{nn'}$$

The above distribution has zero standard deviation. A light with photon distribution narrower than a Poisson distribution is known as sub-Poissonian. It illustrates the quantum nature of light. On the other hand, light with photon distribution broader than a Poisson distribution is known as super-Poissonian. It illustrates the incoherent nature of light since coherent light has Poissonian distribution.

12.6.4.1 Time Evolution of the Coherent State

In the Schrödinger picture, the time dependence is with the eigenfunctions; hence, the time dependence of the coherent state is

$$|\alpha, t\rangle = e^{-\frac{1}{2}|\alpha|^2} \sum_{n=0}^{\infty} \frac{\alpha^n e^{-i(n+1/2)\omega t}}{\sqrt{n!}} |n\rangle \tag{12.6.47}$$

The above follows from that the time dependence of the photon number state, which is an eigenstate, is $\exp(-iE_n t/\hbar)$ with $E_n = \hbar\omega(n + 1/2)$. The above can be rewritten as

$$|\alpha, t\rangle = e^{-i\omega t/2} e^{-\frac{1}{2}|\alpha|^2} \sum_{n=0}^{\infty} \frac{(\alpha e^{-i\omega t})^n}{\sqrt{n!}} |n\rangle = e^{-i\omega t/2} |\alpha e^{-i\omega t}\rangle \tag{12.6.48}$$

**An example of a Poisson distribution is a shot noise source, producing $\bar{n}$ photons in interval τ. The probability of find n photons in this time interval follows the Poisson distribution.

The above can be derived in an alternate way. In the Heisenberg picture, the coherent state is time independent, and time-dependent eigenstate can be obtained by applying the time evolution operator

$$
\begin{aligned}
|\alpha, t\rangle = e^{-i\hat{H}t/\hbar}|\alpha\rangle &= e^{-\frac{1}{2}|\alpha|^2} \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}} e^{-i\hat{H}t/\hbar}|n\rangle \\
&= e^{-\frac{1}{2}|\alpha|^2} \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}} e^{-i(n+1/2)\omega t}|n\rangle \\
&= e^{-i\omega t/2} e^{-\frac{1}{2}|\alpha|^2} \sum_{n=0}^{\infty} \frac{(\alpha e^{-i\omega t})^n}{\sqrt{n!}}|n\rangle = e^{-i\omega t/2}|\alpha e^{-i\omega t}\rangle \qquad (12.6.49)
\end{aligned}
$$

It can be further shown that the above can be summed in terms of a "Gaussian pulse"[††] indicating that this pulse evolves in time without shape distortion.
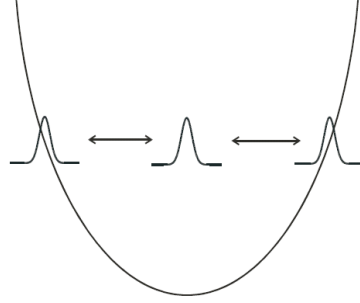
Figure 12.2: The time evolution of the coherent state. It follows the motion of a classical pendulum or harmonic oscillator (From Gerry and Knight).

The expectation value of the field operator in the Heisenberg picture with respect to the coherent state is then

$$
\begin{aligned}
\mathbf{E}(\mathbf{r}, t) = \langle\alpha|\hat{\mathbf{E}}(\mathbf{r}, t)|\alpha\rangle &= \sum_{\mathbf{k},s} \frac{1}{2} \sqrt{\frac{\hbar\omega_k}{V\epsilon_0}} \mathbf{e}_s \langle\alpha|\hat{a}_{\mathbf{k},s}|\alpha\rangle e^{i\mathbf{k}\cdot\mathbf{r} - i\omega_k t} + \text{c.c.} \\
&= \sum_{\mathbf{k},s} \frac{1}{2} \sqrt{\frac{\hbar\omega_k}{V\epsilon_0}} \mathbf{e}_s \alpha_{\mathbf{k},s} e^{i\mathbf{k}\cdot\mathbf{r} - i\omega_k t} + \text{c.c.} \qquad (12.6.50)
\end{aligned}
$$

In the above, it is assumed that the coherent state is a product state of the coherent states of each individual mode, namely,

$$
|\alpha\rangle = \prod_{\mathbf{k},s} |\alpha_{\mathbf{k},s}\rangle \qquad (12.6.51)
$$

[††]It is to be noted that the Gaussian pulse we refer to is with respect to the displacement of the quantum harmonic oscillator, not with respect to the direction of propagation of the photon.
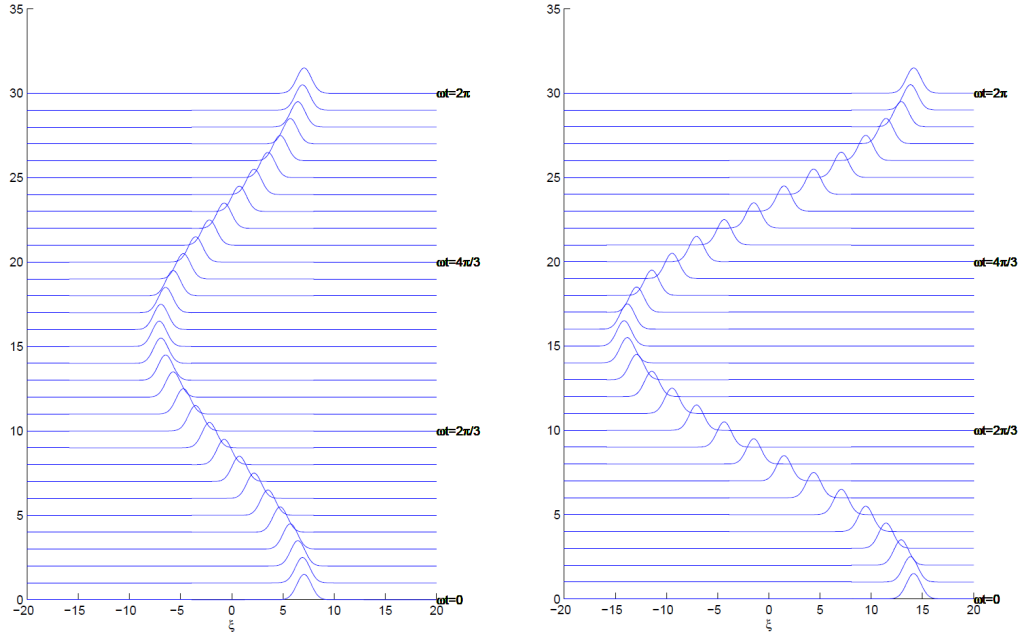
Figure 12.3: Evolution of the coherent state as a function of time. The left figure is for $\alpha = 5$ while the right figure is for $\alpha = 10$.

so that

$$\langle \alpha | \hat{a}_{\mathbf{k},s} | \alpha \rangle = \langle \alpha_{\mathbf{k},s} | \hat{a}_{\mathbf{k},s} | \alpha_{\mathbf{k},s} \rangle = \alpha_{\mathbf{k},s} \tag{12.6.52}$$

Similarly,

$$\mathbf{B}(\mathbf{r}, t) = \langle \alpha | \hat{\mathbf{B}}(\mathbf{r}, t) | \alpha \rangle = \sum_{\mathbf{k},s} \frac{1}{2} \sqrt{\frac{\hbar \omega_k}{V \epsilon_0}} \frac{1}{c} \mathbf{e}_{\mathbf{k}} \times \mathbf{e}_s \alpha_{k,s} e^{i\mathbf{k}\cdot\mathbf{r} - i\omega_k t} + \text{c.c.} \tag{12.6.53}$$

The above certainly look like a classical field where $\alpha_{k,s}$ are complex numbers.

Appendix

A   Thermal Light and Black-Body Radiation

Coherent state is a state where all the photon number states are in phase and in quantum coherence. Hence the density representation of the coherent state is

$$\hat{\rho} = |\alpha\rangle\langle\alpha| = e^{-|\alpha|^2} \sum_{n=0}^{\infty} \sum_{n'=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}} \frac{\alpha^{*n'}}{\sqrt{n'!}} |n\rangle\langle n'| \tag{A1}$$

This is indicated by that the off-diagonal elements of the density operator multiply $|n\rangle\langle n'|$ where $n \neq n'$ are non-zero due to coherence.

When a quantum harmonic oscillator is in thermal equilibrium with its environment, it absorbs energy from its environment causing it to be in the linear superposition of $|n\rangle$ photon number states. The state vector can be expressed as

$$|\psi\rangle = \sum_{n=0}^{\infty} a_n |n\rangle \tag{A2}$$

But these states are incoherent with each other since they are randomly excited by a thermal bath. This is the nature of thermal photons. The density operator for such a state can be written as

$$\hat{\rho}_T = |\psi\rangle\langle\psi| = \sum_{n=0}^{\infty} |a_n|^2 |n\rangle\langle n| \tag{A3}$$

where the off diagonal terms ensemble average to zero.

According to Boltzmann's law, at thermal equilibrium, the probability of finding a system to be energy state E_n is

$$P(n) = \frac{e^{-E_n/k_B T}}{Z} \tag{A4}$$

where

$$Z = \sum_{n=0}^{\infty} e^{-E_n/k_B T} \tag{A5}$$

and Z is the partition function. It is clear that $\sum_{n=0}^{\infty} P(n) = 1$.

For photon, $E_n = \left(n + \frac{1}{2}\right)\hbar\omega$. Hence,

$$Z = e^{-\frac{1}{2}\hbar\omega/k_B T} \sum_{n=0}^{\infty} e^{-n\hbar\omega/k_B T} = e^{-\frac{1}{2}\hbar\omega/k_B T} \sum_{n=0}^{\infty} x^n = \frac{e^{-\frac{1}{2}\hbar\omega/k_B T}}{1 - e^{-\hbar\omega/k_B T}} \tag{A6}$$

where $x = e^{-\hbar\omega/k_B T}$ and from (A4)

$$P(n) = e^{-n\hbar\omega/k_B T}(1 - e^{-\hbar\omega/k_B T}) \tag{A7}$$

The mean photon number is

$$\bar{n} = \sum_{n=0}^{\infty} n P(n) = \sum_{n=0}^{\infty} n x^n (1 - x) = \frac{x}{1 - x} = \frac{1}{e^{\hbar\omega/k_B T} - 1} \tag{A8}$$

The probability distribution given by (A7) definitely has a broader spread compared to the case for coherent state, which has a Poissonian probability distribution. Light of this nature is termed super-Poissonian.

We can use the above to explain black-body radiation. From the above, we deduce that the average energy of a quantum harmonic oscillator is

$$\langle E \rangle = \bar{n}\hbar\omega = \frac{\hbar\omega}{e^{\hbar\omega/k_B T} - 1} \tag{A9}$$

In the limit when $\hbar \to 0$, the above becomes $\langle E \rangle = k_B T$ as predicated by the equi-partition of energy theorem that each degree of freedom of a particle absorbs energy equal to $k_B T/2$ when the particle is at thermal equilibrium with its environment. A classical harmonic oscillator has two degrees of freedom, one for kinetic energy, and one for potential energy.

To understand black-body radiation, consider a cuboid of dimensions $L_x \times L_y \times L_z$. Applying periodic boundary condition, discrete plane-wave modes exist in the cuboid. In the k space, each mode occupies a volume of

$$(2\pi)^3/(L_x L_y L_z) = (2\pi)^3/V$$

where V is the volume of the cuboid. Hence, in a spherical shell in the k space, the number of modes is

$$\Delta N = 2\frac{V 4\pi k^2 \Delta k}{(2\pi)^3} \tag{A10}$$

The 2 factor is due to two polarizations per mode. For photons, $k = \omega/c$, and the above can be rewritten as density of states per unit volume by dividing the above by V,

$$\Delta\mathcal{N} = \frac{\omega^2 \Delta\omega}{\pi^2 c^3} \tag{A11}$$

We can convert the above into energy density since we know the energy per each quantum harmonic oscillator, or

$$\Delta E = \Delta\mathcal{N} \bar{n} \hbar\omega = \frac{\omega^2 \Delta\omega}{\pi^2 c^3} \frac{\hbar\omega}{e^{\hbar\omega/k_B T} - 1} \tag{A12}$$

Consequently, the energy density per unit frequency is

$$U(\omega) = \frac{dE}{d\omega} = \frac{\omega^2}{\pi^2 c^3} \frac{\hbar\omega}{e^{\hbar\omega/k_B T} - 1} \tag{A13}$$

Since a black-body is in thermal equilibrium, the energy it radiates is proportional to the energy density. Hence, the black-body radiation spectrum is proportional to the above expression. It indicates that the radiation from a black-body is small when the frequency is small, and increases as the frequency increases, but diminishes again when the frequency is high, in agreement with experimental observation. The initial increase with ω is because the density of states increases with ω. However, as $\hbar\omega$ increases, the quantization energy level increases in the quantum harmonic oscillator. According to Boltzmann's law, only the low photon number states are likely to be occupied, decreasing the average number of photons with high ω. This gives rise to the decreased radiation with increasing frequency. Moreover, the frequency where the peak radiation occurs also becomes higher (shorter wavelength) as the temperature increases.

When $\hbar \to 0$, the above becomes

$$U(\omega) = \frac{\omega^2}{\pi^2 c^3} k_B T$$

The above is Rayleigh-Jeans law of radiation. It predict large radiation with increasing frequency called the ultra-violet catastrophe. This problem is remedied by the discovery of Planck's law.
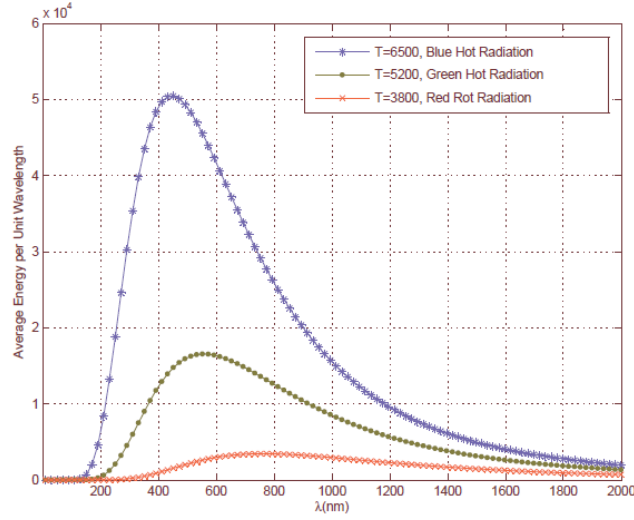
Figure A1: According to Planck's radiation law, body of different temperature will radiate with different colors with increasing order of hotness from red to green to blue.

When $\hbar\omega \gg k_B T$, or when the temperature is low,

$$U(\omega) \approx \frac{\hbar\omega^3}{\pi^2 c^3} e^{-\hbar\omega/k_B T} \tag{A14}$$

This is Wien's law.

The average energy per unit volume is obtained by integrating over all frequencies

$$\overline{U} = \int_0^\infty U(\omega)d\omega = \frac{\hbar}{\pi^2 c^3} \int_0^\infty \frac{\omega^3 d\omega}{e^{\hbar\omega/k_B T} - 1}$$
$$= \frac{\hbar}{\pi^2 c^3} \left(\frac{k_B T}{\hbar}\right)^4 \int_0^\infty \frac{x^3 dx}{e^x - 1} \tag{A15}$$

It can be shown that (Gradshteyn and Ryzhik, p. 326, Eq. 17)

$$\int_0^\infty \frac{x^3 dx}{e^x - 1} = \frac{\pi^4}{15} \tag{A16}$$

Then

$$\overline{U} = \frac{\pi^2}{15c^3} \frac{(k_B T)^4}{\hbar^3} \tag{A17}$$

The above is the Stefan-Boltzmann law.
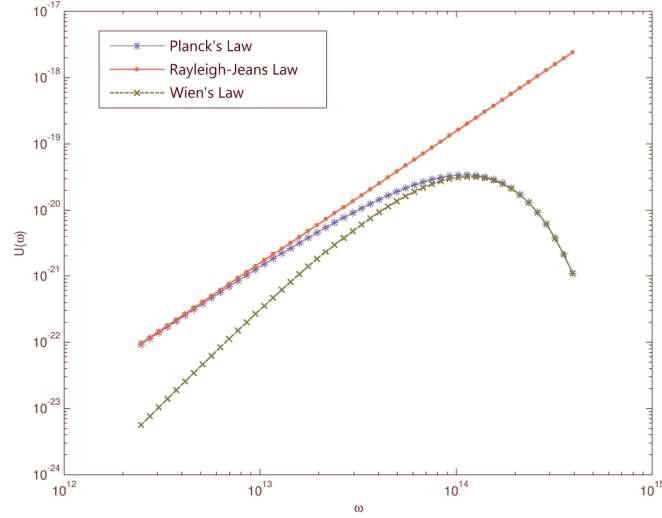
Figure A2: Different radiation laws are the special case of the Planck's radiation law, which takes into account of quantum mechanics.

B Vacuum Fluctuation and Casimir Force

The energy of a harmonic oscillator is given by

$$E_n = \hbar\omega \left(n + \frac{1}{2} \right) \tag{B1}$$

The vacuum state corresponds to $n = 0$ is equivalent to the absence of electric field or photons. It can be shown that even though the mean of electric field is zero for the vacuum state, the mean square of the deviation of the electric field is non-zero. This is known as vacuum fluctuation. Even in vacuum with "no" electric field or photon, there is always a fluctuating electric field!

This fluctuating electric field between two parallel plates gives rise to a force known as the Casimir force. Consider a parallel plate resonator making up of two square plates of length L separated by distance d and then terminated by conducting walls on the side. The resonant frequencies are given by

$$\omega_{mnp} = c\left[\left(\frac{m\pi}{L}\right)^2 + \left(\frac{n\pi}{L}\right)^2 + \left(\frac{p\pi}{d}\right)^2 \right]^{1/2} \tag{B2}$$
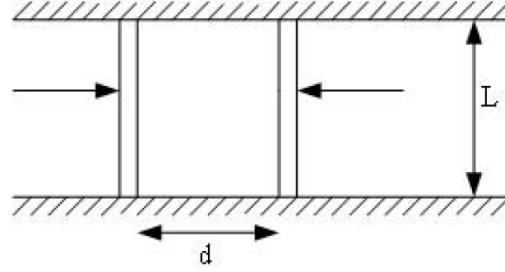
Figure B1: In the model for the calculation of Casimir force, we assume two square metallic plates enclosed by a square $L \times L$ cylinder. The two plates are freely movable within the square cylinder, and we calculate the work done needed to bring the two metallic plates from infinite separation to a separation d.

The zero point energy (ZPE) of the modes in the parallel plate resonator is given by

$$E_0(d) = \sum_{m,n,p}{}' (2)\frac{1}{2}\hbar\omega_{mnp} \tag{B3}$$

$$= \hbar c \sum_{m,n,p}{}' \left[\left(\frac{m\pi}{L}\right)^2 + \left(\frac{n\pi}{L}\right)^2 + \left(\frac{p\pi}{d}\right)^2\right]^{1/2} \tag{B4}$$

$$\approx \hbar c \left(\frac{L}{\pi}\right)^2 \sum_{p=0}^{\infty}{}' \int_0^{\infty} dk_x \int_0^{\infty} dk_y \left[k_x^2 + k_y^2 + \left(\frac{p\pi}{d}\right)^2\right]^{1/2} \tag{B5}$$

When $L \to \infty$, the modes in the m and n directions become indenumerable, the summation over m and n can be replaced by integrals. The above has a physical picture of a plane wave bouncing between two infinitely large parallel plates. The modes can be decomposed into TE and TM to z modes. But when $p = 0$, the TE modes does not exist as it is shorted out. Then only TM mode exists. For all the other p's, both TE and TM modes exist. Hence, there are two polarizations for each mode except for the first mode. That explains the factor of 2 in the above summation. Therefore, the prime in the above indicates that the first term is weighted by a factor of $1/2$. If the plates are very far apart, we have

$$E_0(\infty) = \hbar c \left(\frac{L}{\pi}\right)^2 \frac{d}{\pi} \int_0^{\infty} dk_x \int_0^{\infty} dk_y \int_0^{\infty} dk_z \, [k_x^2 + k_y^2 + k_z^2]^{1/2} \tag{B6}$$

The potential energy of the system is the energy needed to move the plates from infinity to

separation d. Hence, the potential energy

$$
\begin{aligned}
U(d) &= E_0(d) - E_0(\infty) \\
&= \hbar c \left(\frac{L}{\pi}\right)^2 \left\{ \sum_{p=0}^{\infty}{}' \int_0^{\infty} dk_x \int_0^{\infty} dk_y \left[k_x^2 + k_y^2 + \left(\frac{p\pi}{d}\right)^2\right]^{1/2} \right. \\
&\quad \left. -\frac{d}{\pi} \int_0^{\infty} dk_x \int_0^{\infty} dk_y \int_0^{\infty} dk_z \, [k_x^2 + k_y^2 + k_z^2]^{1/2} \right\}
\end{aligned}
\tag{B7}
$$

Converting the above into cylindrical coordinates, and noting that we are integrating only over a quadrant of the plane, we have

$$
U(d) = \frac{\pi}{2} \hbar c \left(\frac{L}{\pi}\right)^2 \left\{ \sum_{p=0}^{\infty}{}' \int_0^{\infty} d\eta \, \eta \left[\eta^2 + \left(\frac{p\pi}{d}\right)^2\right]^{1/2} - \frac{d}{\pi} \int_0^{\infty} d\zeta' \int_0^{\infty} d\eta \, \eta (\eta^2 + \zeta'^2)^{1/2} \right\} \tag{B8}
$$

By letting $\eta^2 = (\frac{\pi}{d})^2 w$, and $\zeta' = \zeta\pi/d$, the above becomes

$$
\begin{aligned}
U(d) &= \frac{\pi}{4} \hbar c \left(\frac{L}{\pi}\right)^2 \left(\frac{\pi}{d}\right)^3 \left[\sum_{p=0}^{\infty}{}' \int_0^{\infty} dw (w + p^2)^{1/2} - \int_0^{\infty} d\zeta \int_0^{\infty} dw (w + \zeta^2)^{1/2} \right] \\
&= \frac{\pi}{4} \hbar c \left(\frac{L}{\pi}\right)^2 \left(\frac{\pi}{d}\right)^3 \left[\sum_{p=0}^{\infty}{}' F(p) - \int_0^{\infty} d\zeta F(\zeta) \right]
\end{aligned}
\tag{B9}
$$

where

$$
F(\zeta) = \int_0^{\infty} dw (w + \zeta^2)^{1/2} \tag{B10}
$$

The above derivation follows that of Milonni and Shih, and also discussed in many text books including Milonni, and Gerry and Knight. The above consists of divergent integrals and summation. But the integrals and summation can be tapered by windowing so that they will converge. The window can be chosen so that the values of $F(\zeta)$ and its derivatives are not altered so that the Euler-Maclaurin formula can be used to estimate the difference in the above.

In general, the Euler-Maclaurin formula estimates the difference between an integral and its trapezoidal rule approximation. This difference is given by

$$
\Delta = \left\{ \frac{1}{2} \left[F(0) + F(N)\right] + \sum_{p=1}^{N-1} F(p) - \int_0^N d\zeta \, F(\zeta) \right\} \tag{B11}
$$

$$
= \sum_{k=1}^{K} \frac{B_{2k}}{2k!} \left[F^{(2k-1)}(N) - F^{(2k-1)}(0)\right] + R \tag{B12}
$$

When the integrand vanishes at the upper limit of integration, it simplifies to

$$\Delta = \left\{ \frac{1}{2}\Big[F(0)\Big] + \sum_{p=1}^{\infty} F(p) - \int_0^{\infty} d\zeta \, F(\zeta) \right\} \tag{B13}$$

$$= \sum_{k=1}^{K} \frac{B_{2k}}{2k!}\Big[-F^{(2k-1)}(0)\Big] + R \tag{B14}$$

where R is the residual error which can be made small by choosing K properly, and $B_1 = -\frac{1}{2}$, $B_2 = \frac{1}{6}$, $B_3 = 0$, $B_4 = -\frac{1}{30}$, $\cdots$. The superscript $2k - 1$ indicates derivatives with respect to the argument of the function. Consequently,

$$\Delta = -\frac{1}{12}F'(0) + \frac{1}{720}F'''(0) \tag{B15}$$

With $F'(0) = 0$ and $F'''(0) = -4$, then

$$\Delta = \frac{4B_4}{4!} = -\frac{1}{180} \tag{B16}$$

and

$$U(d) = -\frac{\pi^2 \hbar c}{720 d^3}L^2 \tag{B17}$$

Since this is the work done in bringing two plates from infinity to a separation d apart, the force per unit area on the plate is

$$F(d) = U'(d) = -\frac{\pi^2 \hbar c}{240 d^4} \tag{B18}$$

which is the Casimir force due to zero point energy or vacuum fluctuation.

Chapter 13

Schrödinger Wave Fields*

13.1 Introduction

We notice that electromagnetic field can be viewed as a collection of photons each with an field attached. These photons collectively form the coherent state that produces a wave field that is analogous to the classical electromagnetic wave field. Photons are bosons whose wave field satisfies Maxwell's equations.

Similarly, if a boson wavefunction satisfies Schrödinger equation, a collection of bosons gives rise to a wave field that satisfies Schrödinger equation. Similarly, a collection of fermions gives rise to a wave field that satisfies Schrödinger equation if each individual particle wavefunction satisfies the same equation. We will call such fields Schrödinger wave fields.

We have seen that the number of bosons in an electromagnetic wave field can be tracked by the annihilation and creation operators; similar operators are needed for tracking the particle of a Schrödinger wave field. The wave field can be a collection of bosons or a collection of fermions. Eventually, the bookkeeping of the many-particle system becomes simpler.

For Schrödinger equation, this is the second quantization. The first quantization was the discovery that electrons can be associated with a wavefunction, and the trapped modes in potential well are quantized. The second quantization refers to the fact that the Schrödinger wave field, which can be treated as a continuum, can be made granular or quantized. This is the advanced way of expressing the wave-particle duality concept inherent in a wave field.*

13.2 Fock Space for Fermions

For fermions, each eigenmode or eigenstate is either occupied or unoccupied. In general, there will be infinitely many modes that the fermions can occupy. We could denote a two-particle state as

$$|\psi_{2p}\rangle = |\cdots, 0, 1_m, 0, \cdots, 0, 1_v, 0, \cdots\rangle \qquad (13.2.1)$$

---

*This chapter can be skipped on first reading.
*The quantization of classical fields is also known as first quantization for them.

where only the m and v modes are occupied and the rest of the modes are unoccupied. The notation above is cumbersome, and it is common to denote the two-particle fermion Fock state as

$$|\psi_{2p}\rangle = |m, v\rangle \tag{13.2.2}$$

where for fermions, it is necessary that

$$|m, v\rangle = -|v, m\rangle \tag{13.2.3}$$

We define a fermion creation operator such that

$$\hat{b}_k^\dagger |m, n\rangle = |k, m, n\rangle \tag{13.2.4}$$

The newly created particle always occupies the first space in the Fock state vector. By definition then

$$\hat{b}_l^\dagger \hat{b}_k^\dagger |\cdots\rangle = |l, k, \cdots\rangle \tag{13.2.5}$$

$$\hat{b}_k^\dagger \hat{b}_l^\dagger |\cdots\rangle = |k, l, \cdots\rangle = -|l, k, \cdots\rangle = -\hat{b}_l^\dagger \hat{b}_k^\dagger |\cdots\rangle \tag{13.2.6}$$

Therefore

$$\hat{b}_k^\dagger \hat{b}_l^\dagger + \hat{b}_l^\dagger \hat{b}_k^\dagger = 0 \tag{13.2.7}$$

Similarly, we can show that

$$\hat{b}_k \hat{b}_l + \hat{b}_l \hat{b}_k = 0 \tag{13.2.8}$$

or that the above is the Hermitian conjugate of (13.2.7). To show (13.2.8) more rigorously, we start with a fermion Fock state

$$|\psi\rangle = |\cdots, l, k, \cdots\rangle \tag{13.2.9}$$

This state acquires a minus sign when we permute two of the particles. Finally, we can have a state such that

$$|\psi_F\rangle = |l, k, \cdots\rangle = -|k, l, \cdots\rangle \tag{13.2.10}$$

An annihilation operator can be defined such that

$$\hat{b}_l |l, k, \cdots\rangle = |k, \cdots\rangle \tag{13.2.11}$$

It can be shown that

$$\hat{b}_k \hat{b}_l |\psi_F\rangle = \hat{b}_k \hat{b}_l |l, k, \cdots\rangle = \hat{b}_k |k, \cdots\rangle = |\cdots\rangle \tag{13.2.12}$$

Similarly,

$$-\hat{b}_l \hat{b}_k |\psi_F\rangle = \hat{b}_l \hat{b}_k |k, l, \cdots\rangle = \hat{b}_l |l, \cdots\rangle = |\cdots\rangle \tag{13.2.13}$$

Therefore,

$$\hat{b}_k \hat{b}_l + \hat{b}_l \hat{b}_k = 0 \tag{13.2.14}$$

The above equations (13.2.7) and (13.2.14) imply that if $k = l$,

$$\hat{b}_k^\dagger \hat{b}_k^\dagger = 0 \tag{13.2.15}$$

$$\hat{b}_k \hat{b}_k = 0 \tag{13.2.16}$$

It means that no two particles can be created in the same mode k or no two particles can be annihilated from the same mode k.

By the same token, we can show that for $k \neq l$,

$$\hat{b}_k^\dagger \hat{b}_l + \hat{b}_l \hat{b}_k^\dagger = 0 \tag{13.2.17}$$

When $k = l$, we have a state $|k, \cdots\rangle$ and then

$$
\begin{aligned}
(\hat{b}_k^\dagger \hat{b}_k + \hat{b}_k \hat{b}_k^\dagger)|k, \cdots\rangle &= \hat{b}_k^\dagger \hat{b}_k |k, \cdots\rangle + \hat{b}_k \hat{b}_k^\dagger |k, \cdots\rangle \\
&= \hat{b}_k^\dagger |\cdots\rangle + 0|k, \cdots\rangle = \hat{b}_k^\dagger |\cdots\rangle = |k, \cdots\rangle \\
&= \hat{I}|k, \cdots\rangle
\end{aligned}
\tag{13.2.18}
$$

In conclusion,

$$\hat{b}_k^\dagger \hat{b}_l + \hat{b}_l \hat{b}_k^\dagger = \delta_{kl} \hat{I} \tag{13.2.19}$$

In summary,

$$\left[\hat{b}_k^\dagger, \hat{b}_l\right]_+ = \hat{b}_k^\dagger \hat{b}_l + \hat{b}_l \hat{b}_k^\dagger = \delta_{kl} \hat{I} \tag{13.2.20}$$

$$\left[\hat{b}_k^\dagger, \hat{b}_l^\dagger\right]_+ = \hat{b}_k^\dagger \hat{b}_l^\dagger + \hat{b}_l^\dagger \hat{b}_k^\dagger = 0 \tag{13.2.21}$$

$$\left[\hat{b}_k, \hat{b}_l\right]_+ = \hat{b}_k \hat{b}_l + \hat{b}_l \hat{b}_k = 0 \tag{13.2.22}$$

where $[A, B]_+ = AB + BA$. These commutation relations, as we shall see, do wonders in simplifying the book-keeping of many particle problems.

13.3 Field Operators

Just as in the case of electromagnetic field operators that represent many photons which are bosons, we can define field operators for fermions and bosons in general. A one-particle Schrödinger field operator is defined such that

$$\hat{\psi}(\mathbf{r}) = \sum_j \hat{b}_j \phi_j(\mathbf{r}) \tag{13.3.1}$$

where ϕ_j is an eigenmode from a complete set of orthonormal functions. It may or may not be the j-th eigenstate of the quantum system governed by the one-particle Schrödinger equation. The above field operator acts on a state vector in the Fock space. For instance,

$$
\begin{aligned}
\hat{\psi}(\mathbf{r})|m\rangle &= \hat{\psi}(\mathbf{r})\hat{b}_m^\dagger|0\rangle \\
&= \sum_j \phi_j(\mathbf{r})\hat{b}_j\hat{b}_m^\dagger|0\rangle \\
&= \sum_j \phi_j(\mathbf{r})(\delta_{jm} - \hat{b}_m^\dagger\hat{b}_j)|0\rangle \\
&= \phi_m(\mathbf{r})|0\rangle
\end{aligned}
\tag{13.3.2}
$$

where the commutation relation (13.2.20) has been used. Notice that the above field operator, when operating on a Fock vector with a m-th mode occupied, produces a vector tagged with the spatial dependence of the m-th eigenmode.

After using the completeness property of the orthonormal basis, plus the use of commutation relations for the annihilation and creation operators, the field operator can be shown to satisfy the following commutation relations:

$$
\left[\hat{\psi}^\dagger(\mathbf{r}), \hat{\psi}(\mathbf{r}')\right]_+ = \hat{\psi}^\dagger(\mathbf{r})\hat{\psi}(\mathbf{r}') + \hat{\psi}(\mathbf{r}')\hat{\psi}^\dagger(\mathbf{r}) = \delta(\mathbf{r} - \mathbf{r}')
\tag{13.3.3}
$$

$$
\left[\hat{\psi}^\dagger(\mathbf{r}), \hat{\psi}^\dagger(\mathbf{r}')\right]_+ = \hat{\psi}^\dagger(\mathbf{r})\hat{\psi}^\dagger(\mathbf{r}') + \hat{\psi}^\dagger(\mathbf{r}')\hat{\psi}^\dagger(\mathbf{r}) = 0
\tag{13.3.4}
$$

$$
\left[\hat{\psi}(\mathbf{r}), \hat{\psi}(\mathbf{r}')\right]_+ = \hat{\psi}(\mathbf{r})\hat{\psi}(\mathbf{r}') + \hat{\psi}(\mathbf{r}')\hat{\psi}(\mathbf{r}) = 0
\tag{13.3.5}
$$

As an extension, one can define two-particle field operator

$$
\hat{\psi}(\mathbf{r}_1, \mathbf{r}_2) = \frac{1}{\sqrt{2}} \sum_{n_1, n_2} \hat{b}_{n_2}\hat{b}_{n_1}\phi_{n_1}(\mathbf{r}_1)\phi_{n_2}(\mathbf{r}_2)
\tag{13.3.6}
$$

where the factor $1/\sqrt{2}$ is needed for normalization. Here, $\phi_{n_1}(\mathbf{r}_1)$ and $\phi_{n_2}(\mathbf{r}_2)$ may or may not be eigenstates of the one-particle Schrödinger equation. The requirements on them are that they are complete and orthonormal. In the above, the $n_1 = n_2$ terms vanish because $\hat{b}_n\hat{b}_n = 0$. It can be shown easily that

$$
\begin{aligned}
\hat{\psi}(\mathbf{r}_1, \mathbf{r}_2)|l, m\rangle = \hat{\psi}(\mathbf{r}_1, \mathbf{r}_2)\hat{b}_l^\dagger\hat{b}_m^\dagger|0\rangle &= \frac{1}{\sqrt{2}}\left[\phi_l(\mathbf{r}_1)\phi_m(\mathbf{r}_2) - \phi_l(\mathbf{r}_2)\phi_m(\mathbf{r}_1)\right]|0\rangle \\
&= \phi_{lm}(\mathbf{r}_1, \mathbf{r}_2)|0\rangle
\end{aligned}
\tag{13.3.7}
$$

where $\phi_{lm}(\mathbf{r}_1, \mathbf{r}_2)$ is a two-particle wave function satisfying the symmetry requirements for fermion particles. It is to be noted that the above is zero when $\mathbf{r}_1 = \mathbf{r}_2$. It is a consequence of the field operator operating on the two-particle Fock state, and the commutation relations for fermions. A three-particle field operator is defined to be

$$
\hat{\psi}(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3) = \frac{1}{\sqrt{3!}} \sum_{n_1, n_2, n_3} \hat{b}_{n_3}\hat{b}_{n_2}\hat{b}_{n_1}\phi_{n_1}(\mathbf{r}_1)\phi_{n_2}(\mathbf{r}_2)\phi_{n_3}(\mathbf{r}_3)
\tag{13.3.8}
$$

In the above, none of the n_1, n_2, n_3 can be repeated for the reasons of (13.2.15) and (13.2.16). Then, using the commutation relations derived in the previous section, (13.2.20) to (13.2.22), it can be shown that

$$
\begin{aligned}
\hat{\psi}\left(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3\right)|l, m, n\rangle &= \hat{\psi}\left(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3\right) \hat{b}_l^\dagger \hat{b}_m^\dagger \hat{b}_n^\dagger|0\rangle \\
&= \frac{1}{\sqrt{3!}}\left[\phi_l\left(\mathbf{r}_1\right) \phi_m\left(\mathbf{r}_2\right) \phi_n\left(\mathbf{r}_3\right)+\phi_l\left(\mathbf{r}_2\right) \phi_m\left(\mathbf{r}_3\right) \phi_n\left(\mathbf{r}_1\right)\right. \\
&\qquad +\phi_l\left(\mathbf{r}_3\right) \phi_m\left(\mathbf{r}_1\right) \phi_n\left(\mathbf{r}_2\right)-\phi_l\left(\mathbf{r}_1\right) \phi_m\left(\mathbf{r}_3\right) \phi_n\left(\mathbf{r}_2\right) \\
&\qquad \left.-\phi_l\left(\mathbf{r}_2\right) \phi_m\left(\mathbf{r}_1\right) \phi_n\left(\mathbf{r}_3\right)-\phi_l\left(\mathbf{r}_3\right) \phi_m\left(\mathbf{r}_2\right) \phi_n\left(\mathbf{r}_1\right)\right]|0\rangle \\
&= \phi_{lmn}(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3)|0\rangle \qquad\qquad\qquad\qquad\qquad (13.3.9)
\end{aligned}
$$

where $\phi_{lmn}(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3)$ is the three-particle wave function with the requisite symmetry requirements. In general, an N-particle field operator is

$$
\hat{\psi}\left(\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_N\right)=\frac{1}{\sqrt{N!}} \sum_{n_1, n_2, \cdots, n_N} \hat{b}_{n_N} \cdots \hat{b}_{n_2} \hat{b}_{n_1} \phi_{n_1}\left(\mathbf{r}_1\right) \phi_{n_2}\left(\mathbf{r}_2\right) \cdots \phi_{n_N}\left(\mathbf{r}_N\right) \quad (13.3.10)
$$

In the above sum, only terms where $n_1, n_2, \cdots, n_N$ are distinct are contributing for the same reason given in (13.3.8). It is seen that

$$
\hat{\psi}_{NP}\left(\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_N\right)=\frac{1}{\sqrt{N!}} \hat{\psi}_{1P}\left(\mathbf{r}_N\right) \cdots \hat{\psi}_{1P}\left(\mathbf{r}_2\right) \hat{\psi}_{1P}\left(\mathbf{r}_1\right) \qquad (13.3.11)
$$

where $\hat{\psi}_{1P}$ is a one-particle field operator while $\hat{\psi}_{NP}$ is the N-particle field operator. However, the used of such factorized form has to be handled with caution, as each one-particle operator has to operate on distinct modes in order for the product of factors to be non-zero.

A general Fock state for N fermion particles can be written as

$$
|\Phi_N\rangle=\hat{b}_{n_1}^\dagger \hat{b}_{n_2}^\dagger \hat{b}_{n_3}^\dagger \cdots \hat{b}_{n_N}^\dagger|0\rangle=|n_1, n_2, n_3, \cdots, n_N\rangle \qquad (13.3.12)
$$

In the above, the state vector changes sign when any two of the annihilation operators swap position, and that the state vector is zero if any two of the modes are identical.

When the Fock state is operated on by the N-particle field operator, the coordinate space representation of the N-particle fermion state is obtained as shown by (13.3.7) and (13.3.9). This approach avoids the cumbersome use of Slater determinant in the book-keeping of the N-particle fermion states. In general

$$
\hat{\psi}_{NP}\left(\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_N\right)|\Phi_N\rangle=\phi_N\left(\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_N\right)|0\rangle \qquad (13.3.13)
$$

The above is an elegant and compact way to express an N-particle fermion wave function.

13.4 Similarity Transform

It is prudent to note from the above that one can change between coordinate space basis and Fock space basis via the algebra shown. This allows us to effect a change of basis for a

complex quantum system. Some quantum systems, when expressed in the Fock space basis, is a lot simpler than the coordinate space basis. This was first shown for electromagnetic field. It can be shown for other quantum fields. Hence, it is worthwhile to review the mathematics of similarity transform next.

We can view the change of basis as a similarity transform when we change from eigenfunction space or coordinate space representation to Fock space representation. Given a matrix equation

$$\overline{\mathbf{A}} \cdot \mathbf{x} = \lambda \mathbf{x} \tag{13.4.1}$$

we can define a new representation for the unknowns as

$$\mathbf{x} = \overline{\mathbf{S}} \cdot \mathbf{y} \tag{13.4.2}$$

In the above, $\overline{\mathbf{S}}$ is unitary since the length of the vectors does not change. The above equation becomes

$$\overline{\mathbf{A}} \cdot \overline{\mathbf{S}} \cdot \mathbf{y} = \lambda \overline{\mathbf{S}} \cdot \mathbf{y} \tag{13.4.3}$$

Multiplying the above by $\overline{\mathbf{S}}^{\dagger}$, we have

$$\overline{\mathbf{S}}^{\dagger} \cdot \overline{\mathbf{A}} \cdot \overline{\mathbf{S}} \cdot \mathbf{y} = \lambda \overline{\mathbf{S}}^{\dagger} \cdot \overline{\mathbf{S}} \cdot \mathbf{y} = \lambda \mathbf{y} \tag{13.4.4}$$

where the unitary property of the $\overline{\mathbf{S}}$ operator has been used. Hence,

$$\overline{\mathbf{A}}_s \cdot \mathbf{y} = \lambda \mathbf{y} \tag{13.4.5}$$

where

$$\overline{\mathbf{A}}_s = \overline{\mathbf{S}}^{\dagger} \cdot \overline{\mathbf{A}} \cdot \overline{\mathbf{S}} \tag{13.4.6}$$

Equation (13.4.6) is in the form of a similarity transform.

13.5   Additive One-Particle Operator

We can consider a simple identical many particle Hamiltonian where the particles do not interact with each other, say via the Coulomb potential. For instance, for fermions, the only way they interact is via Pauli's exclusion principle. An example of such a Hamiltonian for N particles, in coordinate space representation, is

$$\hat{H}_r = \sum_{i=1}^{N} \hat{H}_{ri} \tag{13.5.1}$$

where

$$\hat{H}_{ri} = -\frac{\hbar^2}{2m} \nabla_i^2 + V(\mathbf{r}_i) \tag{13.5.2}$$

As an example, consider the three particle case. The three-particle wavefunction in coordinate representation is

$$
\begin{aligned}
\phi_{lmn}(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3) = \frac{1}{\sqrt{3!}} \big[& \phi_l(\mathbf{r}_1)\phi_m(\mathbf{r}_2)\phi_n(\mathbf{r}_3) + \phi_l(\mathbf{r}_2)\phi_m(\mathbf{r}_3)\phi_n(\mathbf{r}_1) \\
& + \phi_l(\mathbf{r}_3)\phi_m(\mathbf{r}_1)\phi_n(\mathbf{r}_2) - \phi_l(\mathbf{r}_1)\phi_m(\mathbf{r}_3)\phi_n(\mathbf{r}_2) \\
& - \phi_l(\mathbf{r}_2)\phi_m(\mathbf{r}_1)\phi_n(\mathbf{r}_3) - \phi_l(\mathbf{r}_3)\phi_m(\mathbf{r}_2)\phi_n(\mathbf{r}_1) \big]
\end{aligned} \tag{13.5.3}
$$

If the eigenmodes above are the eigenmodes of the $\hat{H}_{ri}$ operator, then

$$
\hat{H}_r \phi_{lmn}(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3) = \sum_{i=1}^{N} \hat{H}_{ri} \phi_{lmn}(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3) = (E_l + E_m + E_n)\phi_{lmn}(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3) \tag{13.5.4}
$$

The above approach gets unwieldy as the number of particles increases. However, if we were to test the above equation with $\phi^*_{l'm'n'}(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3)$ and integrate over space, we have

$$
\int d\mathbf{r}_1 d\mathbf{r}_2 d\mathbf{r}_3 \phi^*_{lmn}(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3)\hat{H}_r \phi_{lmn}(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3) = (E_l + E_m + E_n)\delta_{l'l}\delta_{m'm}\delta_{n'n} \tag{13.5.5}
$$

The above indicates that the matrix representation of $\hat{H}_r$ is diagonal and very simple despite the complicated nature of the wave functions. It signals an alternative simpler representation of the N-particle Hamilton. This simplication can be achieved by the change of basis using the identity in (13.3.9) to obtain that

$$
\langle n', m', l'|\hat{H}|l, m, n\rangle = (E_l + E_m + E_n)\delta_{l'l}\delta_{m'm}\delta_{n'n} = \langle n', m', l'|(E_l + E_m + E_n)|l, m, n\rangle \tag{13.5.6}
$$

where

$$
\hat{H} = \int d\mathbf{r}_1 d\mathbf{r}_2 d\mathbf{r}_3 \hat{\psi}^\dagger(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3)\hat{H}_r \hat{\psi}(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3) \tag{13.5.7}
$$

By inspection, if we let

$$
\hat{H} = \sum_p E_p \hat{b}_p^\dagger \hat{b}_p \tag{13.5.8}
$$

it will have the same matrix representation as (13.5.6) above. But we can also derive the above by a lengthier exercise starting with (13.5.7).

In general, for a simpler approach, we can perform a change of basis or similarity transform on the original Hamiltonian. The new Hamiltonian is

$$
\hat{H} = \int \sum_{i=1}^{N} \hat{\psi}^\dagger_{NP}(\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_N)\hat{H}_{ri}\hat{\psi}_{NP}(\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_N)\, d\mathbf{r}_1 d\mathbf{r}_2 \cdots d\mathbf{r}_N \tag{13.5.9}
$$

The operator above now acts on a vector in the Fock space, and transforms it to another vector in the same space.

Equation (13.5.9) can be written, after using (13.3.11), as

$$\hat{H} = \frac{1}{N!} \sum_{i=1}^{N} \int \hat{\psi}_{1P}^{\dagger}(\mathbf{r}_N) \cdots \sim \cdots \hat{\psi}_{1P}^{\dagger}(\mathbf{r}_1) \left[\int d\mathbf{r}_i \hat{\psi}_{1P}^{\dagger}(\mathbf{r}_i) \hat{H}_{ri} \hat{\psi}_{1P}(\mathbf{r}_i) \right]$$

$$\hat{\psi}_{1P}(\mathbf{r}_1) \cdots \sim \cdots \hat{\psi}_{1P}(\mathbf{r}_N) d\mathbf{r}_1 \cdots \sim \cdots d\mathbf{r}_N \tag{13.5.10}$$

where the $\sim$ sign implies that $\mathbf{r}_i$ is excluded from the sequence. We can move $\hat{\psi}_{1P}^{\dagger}(\mathbf{r}_i)$ and $\hat{\psi}_{1P}(\mathbf{r}_i)$ as there are an even number of sign change as we anti-commute the field operators inward.

13.5.1 Three-Particle Case

It is quite complex to sort out the algebra of the above system. Much insight, however, can be gotten by studying a simpler three-particle case. In this case, a typical term of the transformed Hamiltonian is

$$\hat{H} = \frac{1}{3!} \int \hat{\psi}_{1P}^{\dagger}(\mathbf{r}_3) \hat{\psi}_{1P}^{\dagger}(\mathbf{r}_2) \left[\int d\mathbf{r}_1 \hat{\psi}_{1P}^{\dagger}(\mathbf{r}_1) \hat{H}_{r1} \hat{\psi}_{1P}(\mathbf{r}_1) \right]$$

$$\hat{\psi}_{1P}(\mathbf{r}_2) \hat{\psi}_{1P}(\mathbf{r}_3) d\mathbf{r}_2 d\mathbf{r}_3 + \cdots \tag{13.5.11}$$

where the $+ \cdots$ above refers to other two terms where $\mathbf{r}_i = \mathbf{r}_2$ and $\mathbf{r}_i = \mathbf{r}_3$, and $\mathbf{r}_i$ here refers to the coordinates for the term inside the square brackets.

The space-dependent parts involving $\phi_{n_i}(\mathbf{r}_i)$ in the field operators can be grouped together and the spatial integrations can be performed first. These modes are orthonormal giving rise to $\delta_{n_2 n_2'}$ and $\delta_{n_3 n_3'}$ which can be used to reduce double summations into single summations. Finally, one obtains

$$\hat{H} = \frac{1}{6} \sum_{n_2, n_3} \hat{b}_{n_3}^{\dagger} \hat{b}_{n_2}^{\dagger} \left[\sum_{n_1', n_1} \hat{b}_{n_1'}^{\dagger} H_{n_1', n_1}^{(1)} \hat{b}_{n_1} \right] \hat{b}_{n_2} \hat{b}_{n_3} + \cdots \tag{13.5.12}$$

The general case where $\phi_{n_1}(\mathbf{r}_1)$ is not an eigenstate of the Hamiltonian $\hat{H}_{r1}$ is assumed here. Hence,

$$H_{n_1', n_1}^{(1)} = \langle \phi_{n_1'} | \hat{H}_{r1} | \phi_{n_1} \rangle \tag{13.5.13}$$

In the event that $\phi_{n_1}(\mathbf{r}_1)$ is an eigenstate of $\hat{H}_{r1}$, the above becomes

$$H_{n_1', n_1}^{(1)} = E_{n_1} \delta_{n_1' n_1} \tag{13.5.14}$$

To evaluate the action of the above operator in (13.5.12) on the three-fermion-particle state $|l, m, n\rangle = \hat{b}_l^{\dagger} \hat{b}_m^{\dagger} \hat{b}_n^{\dagger} |0\rangle$, one needs to first show that

$$\hat{b}_{n_3} |l, m, n\rangle = \hat{b}_{n_3} \hat{b}_l^{\dagger} \hat{b}_m^{\dagger} \hat{b}_n^{\dagger} |0\rangle$$

$$= \left(\delta_{n n_3} \hat{b}_l^{\dagger} \hat{b}_m^{\dagger} + \delta_{l n_3} \hat{b}_m^{\dagger} \hat{b}_n^{\dagger} + \delta_{m n_3} \hat{b}_n^{\dagger} \hat{b}_l^{\dagger} \right) |0\rangle \tag{13.5.15}$$
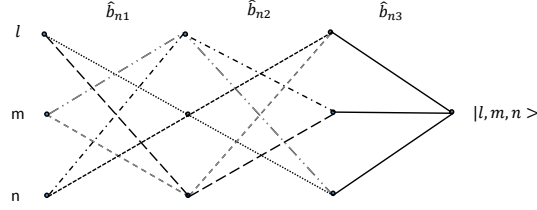
Figure 13.1: The three annihilation operators $\hat{b}_{n_1}\hat{b}_{n_2}\hat{b}_{n_3}$ acting on the three-particle state $|l,m,n\rangle$ produces six threads that eventually gives rise to six terms.

Next, one can show that

$$\hat{b}_{n_2}\hat{b}_{n_3}|l,m,n\rangle = \left(\delta_{ln_3}\delta_{mn_2}\hat{b}_n^\dagger - \delta_{ln_3}\hat{b}_m^\dagger\delta_{nn_2} - \delta_{ln_2}\delta_{mn_3}\hat{b}_n^\dagger \right.$$
$$\left. +\hat{b}_l^\dagger\delta_{mn_3}\delta_{nn_2} + \delta_{ln_2}\hat{b}_m^\dagger\delta_{nn_3} - \hat{b}_l^\dagger\delta_{mn_2}\delta_{nn_3}\right)|0\rangle \qquad (13.5.16)$$

Finally,

$$\hat{b}_{n_1}\hat{b}_{n_2}\hat{b}_{n_3}|l,m,n\rangle = (\delta_{ln_3}\delta_{mn_2}\delta_{nn_1} - \delta_{ln_3}\delta_{mn_1}\delta_{nn_2} - \delta_{ln_2}\delta_{mn_3}\delta_{nn_1}$$
$$+\delta_{ln_1}\delta_{mn_3}\delta_{nn_2} + \delta_{ln_2}\delta_{mn_1}\delta_{nn_3} - \delta_{ln_1}\delta_{mn_2}\delta_{nn_3})|0\rangle \qquad (13.5.17)$$

A total of six terms is found. In the above $\hat{b}_{n_3}$, operating on the three-particle state, produces three terms each of which is a two-particle state as shown in Figure 13.1. Next, $\hat{b}_{n_2}$ operating on them will produce two terms per each term of the two-particle state, producing a net of six terms of one-particle states. The final operation $\hat{b}_{n_1}$ produces one term for each particle. In general, if we start with an N-particle state, the operation of $\hat{b}_{n_1}\cdots\hat{b}_{n_N}$ on it will produce $N!$ terms.

The $\hat{b}^\dagger$ operators in (13.5.12) anti-commute with each other. Hence, $\hat{b}_{n_3}^\dagger\hat{b}_{n_2}^\dagger$ can be moved to the right of the $H_{n_1',n_1}^{(1)}$ element, together with the summations. Therefore, we finally need to evaluate

$$\sum_{n_2,n_3}\hat{b}_{n_3}^\dagger\hat{b}_{n_2}^\dagger\hat{b}_{n_1}\hat{b}_{n_2}\hat{b}_{n_3}|l,m,n\rangle \qquad (13.5.18)$$

The above can be shown to evaluate to

$$\sum_{n_2,n_3}\hat{b}_{n_3}^\dagger\hat{b}_{n_2}^\dagger\hat{b}_{n_1}\hat{b}_{n_2}\hat{b}_{n_3}|l,m,n\rangle = 2\left(\delta_{nn_1}\hat{b}_l^\dagger\hat{b}_m^\dagger + \delta_{ln_1}\hat{b}_m^\dagger\hat{b}_n^\dagger + \delta_{mn_1}\hat{b}_n^\dagger\hat{b}_l^\dagger\right)|0\rangle$$
$$= 2\hat{b}_{n_1}\hat{b}_l^\dagger\hat{b}_m^\dagger\hat{b}_n^\dagger|0\rangle = 2\hat{b}_{n_1}|l,m,n\rangle \qquad (13.5.19)$$

The first equality above can be obtained by using (13.5.17). The second equality follows from

(13.5.15). The above can be generalized to the N-particle case yielding[†]

$$\sum_{n_2,\cdots,n_N} \hat{b}^\dagger_{n_N}\cdots\hat{b}^\dagger_{n_2}\hat{b}_{n_1}\hat{b}_{n_2}\cdots\hat{b}_{n_N}|l_1,\cdots,l_N\rangle = (N-1)!\hat{b}_{n_1}|l_1,\cdots,l_N\rangle \qquad (13.5.20)$$

Consequently,

$$\hat{H}|l,m,n\rangle = \frac{1}{3}\sum_{n'_1,n_1} \hat{b}^\dagger_{n'_1} H^{(1)}_{n'_1,n_1}\hat{b}_{n_1}|l,m,n\rangle + \cdots \qquad (13.5.21)$$

The other two terms indicated by the $+\cdots$ would contribute to exactly the same expression as they are from indistinguishable particles. Finally, we have

$$\hat{H}|l,m,n\rangle = \sum_{n'_1,n_1} \hat{b}^\dagger_{n'_1} H^{(1)}_{n'_1,n_1}\hat{b}_{n_1}|l,m,n\rangle \qquad (13.5.22)$$

In general, the additive one-particle operator in Fock space is

$$\hat{H} = \sum_{n',n} \hat{b}^\dagger_{n'} H^{(1)}_{n',n}\hat{b}_n \qquad (13.5.23)$$

The above can be used for the N-particle case as long as they are indistinguishable. It can be shown to yield the same matrix representation using Fock state vectors, compared to the case in coordinate space representation. It can be easily validated using the two-particle or three-particle state vectors.

For the case when the one-particle eigenstates are also the eigenstates of the $\hat{H}_{r_i}$ operator, the above matrix becomes diagonal yielding,

$$\hat{H} = \sum_n E_n \hat{b}^\dagger_n \hat{b}_n \qquad (13.5.24)$$

The above means that if we have an N-particle fermion field, it can be represented by the physics of one-particle Hamiltonian if the eigenstates chosen for the similarity transform are also the eigenstates of the one-particle Hamiltonian. Also, the above Hamiltonian is very similar to the Hamiltonian for photons derived in the previous chapter, except that photons are bosons and the above derivation is for fermions. A similar derivation for bosons shows that the Hamiltonian is similar to the above.

The above Hamiltonian (13.5.24) does not distinguish between one particle or N particles. This difference of the quantum systems is expressed by the Fock states of the particles.

13.6   Additive Two-Particle Operator

In general, the N particles in a Schrödinger wave field will interact with each other. They may interact pair-wise, for instance, via Coulomb potential other than just the Pauli's exclusion

---

[†]For the fermion case, since $n_1,\cdots,n_N$ are distinct, we can anti-commute the operators to arrange them in pairs of $\hat{b}^\dagger\hat{b}$ and arrive at (13.5.19) and (13.5.20) more succinctly.

principle. An example of an additive two particle operator in coordinate space is

$$V(\mathbf{r}_1, \mathbf{r}_2) = \frac{e^2}{4\pi\epsilon_0 |\mathbf{r}_1 - \mathbf{r}_2|} \tag{13.6.1}$$

This will be part of a two-particle Hamiltonian

$$\hat{H}_r(\mathbf{r}_1, \mathbf{r}_2) = -\frac{\hbar^2}{2m} \left(\nabla_1^2 + \nabla_2^2 \right) + V(\mathbf{r}_1, \mathbf{r}_2) \tag{13.6.2}$$

When N-particle Hamiltonian is considered, the Coulomb interaction appears as

$$V = \sum_{i<j} V(\mathbf{r}_i, \mathbf{r}_j) = \frac{1}{2} \sum_{i\neq j} V(\mathbf{r}_i, \mathbf{r}_j) \tag{13.6.3}$$

For N particles, there are $N(N-1)$ terms in the above summations. Again, we can transform the above from coordinate space representation to Fock-space representation as before arriving at

$$\hat{V} = \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1, j\neq i}^{N} \frac{1}{N!} \int \hat{\psi}_{1P}^\dagger(\mathbf{r}_N) \cdots \sim \cdots \hat{\psi}_{1P}^\dagger(\mathbf{r}_1)$$

$$\left[\int d\mathbf{r}_i d\mathbf{r}_j \hat{\psi}_{1P}^\dagger(\mathbf{r}_i) \hat{\psi}_{1P}^\dagger(\mathbf{r}_j) V(\mathbf{r}_i, \mathbf{r}_j) \hat{\psi}_{1P}(\mathbf{r}_i) \hat{\psi}_{1P}(\mathbf{r}_j) \right]$$

$$\hat{\psi}_{1P}(\mathbf{r}_1) \cdots \sim \cdots \hat{\psi}_{1P}(\mathbf{r}_N) \, d\mathbf{r}_1 \cdots \sim \cdots d\mathbf{r}_N \tag{13.6.4}$$

where the $\sim$ indicates that $\mathbf{r}_i$ and $\mathbf{r}_j$ are excluded from the sequence. Without loss of generality, we can focus on the term where $\mathbf{r}_i = \mathbf{r}_1$ and $\mathbf{r}_j = \mathbf{r}_2$. Then, the above becomes

$$\hat{V} = \frac{1}{2N!} \int \hat{\psi}_{1P}^\dagger(\mathbf{r}_N) \cdots \hat{\psi}_{1P}^\dagger(\mathbf{r}_3) \left[\int d\mathbf{r}_1 d\mathbf{r}_2 \hat{\psi}_{1P}^\dagger(\mathbf{r}_2) \hat{\psi}_{1P}^\dagger(\mathbf{r}_1) V(\mathbf{r}_1, \mathbf{r}_2) \hat{\psi}_{1P}(\mathbf{r}_1) \hat{\psi}_{1P}(\mathbf{r}_2) \right]$$

$$\hat{\psi}_{1P}(\mathbf{r}_3) \cdots \hat{\psi}_{1P}(\mathbf{r}_N) \, d\mathbf{r}_1 \cdots d\mathbf{r}_N + \cdots \tag{13.6.5}$$

where $+ \cdots$ above implies the rest of the terms; there are altogether $N(N-1)$ terms. Performing the $d\mathbf{r}_1 \cdots d\mathbf{r}_N$ integrations first, making use of mode orthonormality, we arrive at

$$\hat{V} = \frac{1}{2N!} \sum_{n_3, \cdots, n_N} \hat{b}_{n_N}^\dagger \cdots \hat{b}_{n_3}^\dagger$$

$$\left[\sum_{n_1', n_1, n_2', n_2} \hat{b}_{n_2'}^\dagger \hat{b}_{n_1'}^\dagger V_{n_1', n_2', n_1, n_2}^{(1,2)} \hat{b}_{n_1} \right] \hat{b}_{n_2} \cdots \hat{b}_{n_N} + \cdots \tag{13.6.6}$$

where

$$V_{n_1', n_2', n_1, n_2}^{(1,2)} = \int d\mathbf{r}_1 d\mathbf{r}_2 \phi_{n_1'}^*(\mathbf{r}_1) \phi_{n_2'}^*(\mathbf{r}_2) V(\mathbf{r}_1, \mathbf{r}_2) \phi_{n_1}(\mathbf{r}_1) \phi_{n_2}(\mathbf{r}_2) \tag{13.6.7}$$

The above needs to operate on an N-particle Fock state, namely,

$$\hat{V}|l_1,\cdots,l_N\rangle = \frac{1}{2N!} \sum_{n_3,\cdots,n_N} \hat{b}^\dagger_{n_N}\cdots\hat{b}^\dagger_{n_3}$$

$$\left[\sum_{n'_1,n_1,n'_2,n_2} \hat{b}^\dagger_{n'_2}\hat{b}^\dagger_{n'_1} V^{(1,2)}_{n'_1,n'_2,n_1,n_2} \hat{b}_{n_1}\hat{b}_{n_2}\right] \hat{b}_{n_3}\cdots\hat{b}_{n_N}|l_1,\cdots,l_N\rangle + \cdots$$

$$(13.6.8)$$

The right-most term that acts on the Fock state is of the form

$$\hat{b}_{n_1}\hat{b}_{n_2}\cdots\hat{b}_{n_N}|l_1,\cdots,l_N\rangle \qquad (13.6.9)$$

The above yields $N!$ terms similar to (13.5.17). But $\hat{b}^\dagger_{n_N}\cdots\hat{b}^\dagger_{n_3}$ commutes with $\hat{b}^\dagger_{n_2}\hat{b}^\dagger_{n_1}$ in (13.6.6) above. We can then move $\hat{b}^\dagger_{n_N}\cdots\hat{b}^\dagger_{n_3}$ to the right of $V^{(1,2)}_{n'_1,n'_2,n_1,n_2}$. Then we need to evaluate

$$\sum_{n_3,\cdots,n_N} \hat{b}^\dagger_{n_N}\cdots\hat{b}^\dagger_{n_3}\hat{b}_{n_1}\hat{b}_{n_2}\cdots\hat{b}_{n_N}|l_1,\cdots,l_N\rangle = (N-2)!\hat{b}_{n_1}\hat{b}_{n_2}|l_1,\cdots,l_N\rangle \qquad (13.6.10)$$

The above equality can be proved by induction from the three-particle case or from equations (13.5.19) and (13.5.20). As a result,

$$\hat{V}|l_1,\cdots,l_N\rangle = \frac{1}{2N(N-1)} \sum_{n'_1,n_1,n'_2,n_2} \hat{b}^\dagger_{n'_2}\hat{b}^\dagger_{n'_1} V^{(1,2)}_{n'_1,n'_2,n_1,n_2} \hat{b}_{n_1}\hat{b}_{n_2}|l_1,\cdots,l_N\rangle + \cdots \qquad (13.6.11)$$

There are $N(N-1)$ terms of the similar kind in the $+\cdots$ above. Summing them up, we have

$$\hat{V}|l_1,\cdots,l_N\rangle = \frac{1}{2} \sum_{n'_1,n_1,n'_2,n_2} \hat{b}^\dagger_{n'_2}\hat{b}^\dagger_{n'_1} V^{(1,2)}_{n'_1,n'_2,n_1,n_2} \hat{b}_{n_1}\hat{b}_{n_2}|l_1,\cdots,l_N\rangle \qquad (13.6.12)$$

In general, the additive two-particle operator in an N-particle fermion field is given by

$$\hat{V} = \frac{1}{2} \sum_{n'_1,n_1,n'_2,n_2} \hat{b}^\dagger_{n'_2}\hat{b}^\dagger_{n'_1} V^{(1,2)}_{n'_1,n'_2,n_1,n_2} \hat{b}_{n_1}\hat{b}_{n_2} \qquad (13.6.13)$$

The above can be easily validated to produce the same matrix representation compared to the coordinate representation using the two-particle or three-particle state vectors.

The above derivations for fermions can be repeated for the boson case. In this case, we will have a field of bosons. It is pleasing that by starting out with the N-particle coordinate space representation of the Schrödinger equation, one arrives at a much simpler Fock space representation of the same equation. It allows one to treat N-particle problems more succinctly using such a representation.

13.7   More on Field Operators

The above shows that the one-particle additive operator is independent of the number of particles. So what distinguishes one-particle quantum system from a two-particle one is in the Fock state: the first one will have a Fock state for one particle while the latter will have a Fock state for two particles.

A general one-particle Fock state is

$$|\psi\rangle = \sum_{\mu} c_{\mu} \hat{b}_{\mu}^{\dagger}|0\rangle \tag{13.7.1}$$

When the one-particle field operator acts on the above, it projects the coordinate space representation of the one-particle wavefunction. Namely,

$$\hat{\psi}(\mathbf{r})|\psi\rangle = \sum_{\mu} c_{\mu} \phi_m(\mathbf{r})|0\rangle = \psi(\mathbf{r})|0\rangle \tag{13.7.2}$$

The above one-particle state is in a linear superposition of different eigenstates. Hence, it can constitute a wave packet.[‡]

Also, from (13.3.1), we can easily show that

$$\hat{b}_{\mu}^{\dagger} = \int \phi_{\mu}(\mathbf{r})\hat{\psi}^{\dagger}(\mathbf{r})d\mathbf{r} \tag{13.7.3}$$

On combining with (13.7.3), we have

$$|\psi\rangle = \sum_{\mu} c_{\mu} \int \phi_{\mu}(\mathbf{r})\hat{\psi}^{\dagger}(\mathbf{r})d\mathbf{r}|0\rangle$$

$$= \int \left[\sum_{\mu} c_{\mu} \phi_{\mu}(\mathbf{r})\right] \hat{\psi}^{\dagger}(\mathbf{r})d\mathbf{r}|0\rangle$$

$$= \int f(\mathbf{r})\hat{\psi}^{\dagger}(\mathbf{r})d\mathbf{r}|0\rangle \tag{13.7.4}$$

Therefore, a general one-particle state can be written as a linear superposition of the one-particle field operator $\hat{\psi}^{\dagger}(\mathbf{r})$ weighted by $f(\mathbf{r})$. It is clear that $f(\mathbf{r})$ satisfies the one-particle Schrödinger equation.

For a general two-particle state in Fock space,

$$|\psi\rangle = \sum_{\mu_1,\mu_2} c_{\mu_1,\mu_2} \hat{b}_{\mu_1}^{\dagger} \hat{b}_{\mu_2}^{\dagger}|0\rangle \tag{13.7.5}$$

upon substituting (13.7.3) into the above, we have

$$|\psi\rangle = \int \left(\sum_{\mu_1\mu_2} \phi_{\mu_1}(\mathbf{r}_1)\phi_{\mu_2}(\mathbf{r}_2)c_{\mu_1,\mu_2}\right) \hat{\psi}^{\dagger}(\mathbf{r}_1)\hat{\psi}^{\dagger}(\mathbf{r}_2)d\mathbf{r}_1 d\mathbf{r}_2|0\rangle$$

$$= \int f(\mathbf{r}_1,\mathbf{r}_2)\hat{\psi}^{\dagger}(\mathbf{r}_1)\hat{\psi}^{\dagger}(\mathbf{r}_2)d\mathbf{r}_1 d\mathbf{r}_2|0\rangle \tag{13.7.6}$$

[‡]See the appendix of Chapter 4 for a discussion of wave packets.

where

$$f(\mathbf{r}_1, \mathbf{r}_2) = \sum_{\mu_1, \mu_2} \phi_{\mu_1}(\mathbf{r}_1)\phi_{\mu_2}(\mathbf{r}_2)c_{\mu_1, \mu_2}. \tag{13.7.7}$$

For fermions, it can be easily shown, using the commutation relations for the field operators, that

$$f(\mathbf{r}_1, \mathbf{r}_2) = -f(\mathbf{r}_2, \mathbf{r}_1) \tag{13.7.8}$$

The above implies that $c_{\mu_1, \mu_2} = -c_{\mu_2, \mu_1}$. The above satisfies the two-particle Schrödinger equation. It can be generalized to N particles, yielding

$$|\psi\rangle = \int f(\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_N)\, \hat{\psi}^\dagger(\mathbf{r}_1)\, \hat{\psi}^\dagger(\mathbf{r}_2) \cdots \hat{\psi}^\dagger(\mathbf{r}_N)\, d\mathbf{r}_1 d\mathbf{r}_2 \cdots d\mathbf{r}_N\, |0\rangle \tag{13.7.9}$$

13.8 Boson Wave Field

The results for fermion wave field in the previous section can also be derived for boson wave field. The difference is that more than one boson can occupy a given state: The Pauli exclusion principle does not apply to bosons. A boson state involving two modes can be denoted as:

$$|\psi\rangle = |n_m, n_v\rangle \tag{13.8.1}$$

where there are n_m particles in mode m, and n_v particles in mode v. Since ordering is unimportant, it is necessary that

$$|n_m, n_v\rangle = |n_v, n_m\rangle \tag{13.8.2}$$

A boson creation operator can be defined such that

$$\hat{a}_k^\dagger |n_k, n_m, n_v\rangle = C_{n_k+1}|n_k + 1, n_m, n_v\rangle \tag{13.8.3}$$

It raises the number of bosons in state k by one. By definition,

$$\hat{a}_l^\dagger \hat{a}_k^\dagger |n_k, n_l, \cdots\rangle = C_{n_k+1}C_{n_l+1}|n_k + 1, n_l + 1, \cdots\rangle \tag{13.8.4}$$

$$\hat{a}_k^\dagger \hat{a}_l^\dagger |n_l, n_k, \cdots\rangle = C_{n_l+1}C_{n_k+1}|n_l + 1, n_k + 1, \cdots\rangle = \hat{a}_l^\dagger \hat{a}_k^\dagger |n_k, n_l, \cdots\rangle \tag{13.8.5}$$

The last equality follows since ordering is unimportant. Consequently,

$$\hat{a}_k^\dagger \hat{a}_l^\dagger - \hat{a}_l^\dagger \hat{a}_k^\dagger = 0 \tag{13.8.6}$$

Similarly, one can define annihilation operators for bosons that have the opposite effect as the creation operators. Hence, going through a similar process, one also has

$$\hat{a}_k \hat{a}_l - \hat{a}_l \hat{a}_k = 0 \tag{13.8.7}$$

From definition,

$$\hat{a}^\dagger |n\rangle = C_{n+1}|n + 1\rangle \tag{13.8.8}$$

$$\hat{a}|n\rangle = B_n |n - 1\rangle \tag{13.8.9}$$

Furthermore, from (13.8.9), we have

$$\langle n+1|\hat{a}^\dagger|n\rangle = C_{n+1}, \qquad C_{n+1}^* = \langle n|\hat{a}|n+1\rangle = B_{n+1} \tag{13.8.10}$$

Then, we can show that

$$\hat{a}_l^\dagger \hat{a}_k |n_k, n_l, \cdots\rangle = C_{n_k}^* C_{n_l+1} |n_k - 1, n_l + 1, \cdots\rangle \tag{13.8.11}$$

$$\hat{a}_k \hat{a}_l^\dagger |n_l, n_k, \cdots\rangle = C_{n_l+1}^* C_{n_k} |n_l + 1, n_k - 1, \cdots\rangle \tag{13.8.12}$$

If C_n is a real number, then when $l \neq k$,

$$\hat{a}_k \hat{a}_l^\dagger - \hat{a}_l^\dagger \hat{a}_k = 0 \tag{13.8.13}$$

For the case when $l = k$, we can drop the subscripts k and l and look at

$$\hat{a}^\dagger \hat{a} |n\rangle = C_n^2 |n\rangle \tag{13.8.14}$$

$$\hat{a}\hat{a}^\dagger |n\rangle = C_{n+1}^2 |n\rangle \tag{13.8.15}$$

If $C_{n+1}^2 - C_n^2 = 1$, then

$$\hat{a}\hat{a}^\dagger - \hat{a}^\dagger \hat{a} = \hat{I} \tag{13.8.16}$$

Then above implies that $C_n^2 = n$, since $C_0 = 0$. Also, $C_n = \sqrt{n}$. In summary, for bosons,

$$\hat{a}_k^\dagger \hat{a}_l^\dagger - \hat{a}_l^\dagger \hat{a}_k^\dagger = 0 \tag{13.8.17}$$

$$\hat{a}_k \hat{a}_l - \hat{a}_l \hat{a}_k = 0 \tag{13.8.18}$$

$$\hat{a}_k \hat{a}_l^\dagger - \hat{a}_l^\dagger \hat{a}_k = \hat{I}\delta_{kl} \tag{13.8.19}$$

The above is derived without resorting to the use of the quantum harmonic oscillator, but the assumption that C_n is real and that $C_n^2 = n$. One can further conclude that

$$\hat{a}^\dagger \hat{a} |n\rangle = n|n\rangle \tag{13.8.20}$$

which is the number operator, which has been derived using a different approach.

13.9 Boson Field Operators

Similar to fermions, we can define field operators for bosons, so that

$$\hat{\psi}(\mathbf{r}) = \sum_j \hat{a}_j \phi_j(\mathbf{r}) \tag{13.9.1}$$

In gereral, for N particles,

$$\hat{\psi}(\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_N) = \frac{1}{\sqrt{N!}} \sum_{n_1, n_2, \cdots, n_N} \hat{a}_{n_N} \cdots \hat{a}_{n_1} \phi_{n_1}(\mathbf{r}_1)\phi_{n_2}(\mathbf{r}_2) \cdots \phi_{n_N}(\mathbf{r}_N) \tag{13.9.2}$$

We can illustrate with a three-particle field operator:

$$\hat{\psi}(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3) = \frac{1}{\sqrt{3!}} \sum_{n_1, n_2, n_3} \hat{a}_{n_3} \hat{a}_{n_3} \hat{a}_{n_1} \hat{\phi}(\mathbf{r}_1) \phi(\mathbf{r}_1) \phi(\mathbf{r}_2) \phi(\mathbf{r}_3) \tag{13.9.3}$$

When this operates on $|1_l, 1_m, 1_n\rangle$ where the three particles are in different modes, we get (13.3.9) except that $-$ signs are now replaced by $+$ signs. If we have a state denoted by

$$|2_l, 1_m\rangle = \frac{1}{\sqrt{2!}} \left(\hat{a}_l^\dagger \right)^2 a_m |0\rangle \tag{13.9.4}$$

then

$$\hat{\psi}(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3)|2_l, 1_m\rangle$$
$$= \frac{1}{\sqrt{2!3!}} \left[2\phi_l(\mathbf{r}_1)\phi_m(\mathbf{r}_2)\phi_l(\mathbf{r}_3) + 2\phi_l(\mathbf{r}_2)\phi_m(\mathbf{r}_3)\phi_l(\mathbf{r}_1) + 2\phi_l(\mathbf{r}_3)\phi_m(\mathbf{r}_1)\phi_l(\mathbf{r}_2) \right] |0\rangle \tag{13.9.5}$$

It can be shown that the above is the correctly normalized wavefunction. If the state is

$$|3_l\rangle = \frac{1}{\sqrt{3!}} \left(\hat{a}_l^\dagger \right)^3 |0\rangle \tag{13.9.6}$$

then

$$\hat{\psi}(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3)|3_l\rangle = \frac{6}{3!} \phi_l(\mathbf{r}_1)\phi_l(\mathbf{r}_2)\phi_l(\mathbf{r}_3)|0\rangle \tag{13.9.7}$$

The above is in fact normalized.

13.10   Additive One-Particle Operator

In this case, we can illustrate with the three-particle case for bosons, arriving at an expression similar to (13.5.12)

$$\hat{H} = \frac{1}{6} \sum_{n_2, n_3} \hat{a}_{n_3}^\dagger \hat{a}_{n_2}^\dagger \left[\sum_{n_1', n_1} \hat{a}_{n_1'}^\dagger H_{n_1', n_1}^{(1)} \hat{a}_{n_1} \right] \hat{a}_{n_2} \hat{a}_{n_3} + \cdots \tag{13.10.1}$$

where

$$H_{n_1', n_1}^{(1)} = \langle \phi_{n_1'} | \hat{H}_{r_1} | \phi_{n_1} \rangle \tag{13.10.2}$$

When the three particles are in the $|1_l, 1_m, 1_n\rangle$ state, (13.5.17) follows for bosons except that we replace $-$ signs with $+$ signs. Equation (13.5.19) follows similarly for bosons, so does (13.5.21) and (13.5.22). When two of the particles are in the same state as (13.9.4), we have

$$\hat{a}_{n_1} \hat{a}_{n_2} \hat{a}_{n_3} |2_l, 1_m\rangle = \sqrt{2!} (\delta_{ln_3} \delta_{mn_2} \delta_{ln_1} + \delta_{ln_3} \delta_{mn_1} \delta_{ln_2} + \delta_{ln_2} \delta_{mn_3} \delta_{ln_1}) |0\rangle \tag{13.10.3}$$

Similar to (13.5.19), we need to evaluate

$$\sum_{n_2,n_3} \hat{a}_{n_3}^\dagger \hat{a}_{n_2}^\dagger \hat{a}_{n_1} \hat{a}_{n_2} \hat{a}_{n_3} |2_l, 1_m\rangle$$

$$= \sqrt{2!} \left(2\delta_{ln_1} \hat{a}_l^\dagger \hat{a}_m^\dagger + \delta_{mn_1} \hat{a}_l^\dagger \hat{a}_l^\dagger \right) |0\rangle$$

$$= \sqrt{2!}\hat{a}_{n_1} \left(\hat{a}_l^\dagger \right)^2 \hat{a}_m^\dagger |0\rangle = (2!)\hat{a}_{n_1}|2_l, 1_m\rangle \qquad (13.10.4)$$

The first equality is established using (13.10.3), while the second equality is obtained by working backward using commutation relations to expand $\hat{a}_{n_1} \left(\hat{a}_l^\dagger \right)^2 \hat{a}_m^\dagger |0\rangle$. When the three particles are in the same state as indicated by (13.9.6), we have

$$\hat{a}_{n_1}\hat{a}_{n_2}\hat{a}_{n_3}|3_l\rangle = (3!)^{3/2}\delta_{ln_3}\delta_{ln_2}\delta_{ln_1}|0\rangle \qquad (13.10.5)$$

and

$$\sum_{n_2 n_3} \hat{a}_{n_3}^\dagger \hat{a}_{n_2}^\dagger \hat{a}_{n_1} \hat{a}_{n_2} \hat{a}_{n_3} |3_l\rangle$$

$$= (3!)^{3/2} \delta_{ln_1} \left(\hat{a}_l^\dagger \right)^2 |0\rangle = \frac{1}{3}(3!)^{3/2}\hat{a}_{n_1} \left(\hat{a}_l^\dagger \right)^3 |0\rangle = (2!)\hat{a}_{n_1}|3_l\rangle \qquad (13.10.6)$$

In general,

$$\sum_{n_2,n_3} \hat{a}_{n_3}^\dagger \hat{a}_{n_2}^\dagger \hat{a}_{n_1} \hat{a}_{n_2} \hat{a}_{n_3} |\psi_{3p}\rangle = (2!) \, \hat{a}_{n_1} |\psi_{3p}\rangle \qquad (13.10.7)$$

where $|\psi_{3p}\rangle$ is the three-particle Fock state of either

$$|1_l, 1_m, 1_n\rangle, \quad |2_l, 1_m\rangle, \quad |3_l\rangle \qquad (13.10.8)$$

or other combinations. The proof for the rest follows that of fermion particles of (13.5.21) to (13.5.24). In general, the Hamiltonian for N bosons is given by

$$\hat{H} = \sum_{n',n} H_{n',n}\hat{a}_{n'}^\dagger \hat{a}_n \qquad (13.10.9)$$

For the diagonal case, it becomes

$$\hat{H} = \sum_n E_n \hat{a}_n^\dagger \hat{a}_n \qquad (13.10.10)$$

Chapter 14

Interaction of Different Particles*

14.1   Introduction

In the previous chapter, it was shown that the Hamiltonian of a many particle system can be written succinctly using Fock space representation. This opens up the study of complex systems, such as excitons that involve the interaction of electron-hole pair, as well as superconductivity that involves the interaction of electrons with phonons. In this chapter, we will study the interaction of the electron of an atom with photons of a cavity using the many-particle formalism.

14.2   Interaction of Particles

Say if we have two quantum systems that are initially non-interacting, incoherent, and independent, their individual quantum systems can be described by eigenstates that satisfy

$$\hat{H}_1 \left| \psi_1 \right\rangle = E_1 \left| \psi_1 \right\rangle \tag{14.2.1}$$

$$\hat{H}_2 \left| \psi_2 \right\rangle = E_2 \left| \psi_2 \right\rangle \tag{14.2.2}$$

Even though the two systems are entirely non-interacting with each other, nevertheless, we can combine the two systems together and write

$$\hat{H}_0 \left| \psi \right\rangle = \left(\hat{H}_1 + \hat{H}_2 \right) \left| \psi_1 \right\rangle \left| \psi_2 \right\rangle = (E_1 + E_2) \left| \psi_1 \right\rangle \left| \psi_2 \right\rangle = E \left| \psi \right\rangle \tag{14.2.3}$$

where $\left| \psi_1 \right\rangle \left| \psi_2 \right\rangle$ represents a state in the direct product space. It can be used to represent the eigenstates of the combined system. Equation (14.2.3) is entirely equivalent to (14.2.1) and (14.2.2). In the above, $\hat{H}_1$ acts only on $\left| \psi_1 \right\rangle$ and $\hat{H}_2$ acts on $\left| \psi_2 \right\rangle$. The above allows

---

*This chapter can be skipped on first reading.

us to add an interaction term between system 1 and system 2, and the same direct product space can be used to span the solution space. By using Fock space representation, the above Hamiltonians can be written in terms of annihilation and creation operators for a system consisting of fermions and bosons,

$$\hat{H}_1 = \sum_j E_j \hat{b}_j^\dagger \hat{b}_j, \qquad \hat{H}_2 = \sum_\lambda \hbar\omega_\lambda \hat{a}_\lambda^\dagger \hat{a}_\lambda \qquad (14.2.4)$$

where $\hat{H}_1$ can be the Hamiltonian for electrons, and $\hat{H}_2$ can be the Hamiltonian for photons where the $\frac{1}{2}\hbar\omega_\lambda$ term or the zero-point energy has been ignored. This term will just introduce a phase shift in the solution.

When electric dipole interaction exists between the electron of an atom and the electromagnetic field, the interaction Hamiltonian may be written as

$$\hat{H}_{ed,\mathbf{r}} = e\mathbf{E} \cdot \mathbf{r} \qquad (14.2.5)$$

The above has been added as a perturbation to the unperturbed equation (14.2.3) where the electric field is treated classically in a previous chapter. We have enough knowledge now to treat both $\mathbf{E}$ and $\mathbf{r}$ quantum mechanically. Here, $\mathbf{r}$ represents the position of the electron. For the electron in the i-th atom, the above becomes

$$\hat{H}_{ed,\mathbf{r}_i} = ie \sum_\lambda \left(\hat{a}_\lambda - \hat{a}_\lambda^\dagger\right) \sqrt{\frac{\hbar\omega}{2\epsilon_0}} \mathbf{u}_\lambda(\mathbf{r}_i) \cdot \mathbf{r}_i \qquad (14.2.6)$$

We have used a quantized cavity mode electric field in (14.2.5) to arrive at the above. Also, $\mathbf{r}$ in coordinate space representation remains unchanged. When N atoms are present,

$$\hat{H}_{ed,\mathbf{r}} = \sum_{i=1}^N ie \sum_\lambda \left(\hat{a}_\lambda - \hat{a}_\lambda^\dagger\right) \sqrt{\frac{\hbar\omega}{2\epsilon_0}} \mathbf{u}_\lambda(\mathbf{r}_i) \cdot \mathbf{r}_i \qquad (14.2.7)$$

A similarity transform of the above Hamiltonian can be performed with the N particle field operator to yield

$$\hat{H}_{ed} = \int \hat{\psi}_{NP}^\dagger(\mathbf{r}_1, \cdots, \mathbf{r}_N) \hat{H}_{ed,\mathbf{r}} \hat{\psi}_{NP}(\mathbf{r}_1, \cdots, \mathbf{r}_N) d\mathbf{r}_1 \cdots d\mathbf{r}_N \qquad (14.2.8)$$

As mentioned in the previous chapter, since only one-particle interaction is involved, the above can be transformed with the one-particle field operator yielding

$$\hat{H}_{ed} = \sum_{j,k,\lambda} H_{ed,\lambda,j,k} \hat{b}_j^\dagger \hat{b}_k \left(\hat{a}_\lambda - \hat{a}_\lambda^\dagger\right) \qquad (14.2.9)$$

where

$$H_{ed,\lambda,j,k} = ie\sqrt{\frac{\hbar\omega}{2\epsilon_0}} \int d\mathbf{r}\, \phi_j^*(\mathbf{r}) \mathbf{u}_\lambda(\mathbf{r}) \cdot \mathbf{r}\, \phi_k(\mathbf{r}) \qquad (14.2.10)$$

In the above, j, k represents the orbital wave functions of the atomic energy states with energy E_j and E_k, respectively, and λ denotes the electromagnetic mode.

14.3 Time-Dependent Perturbation Theory

In the time-dependent perturbation theory, we first seek the solution of the unperturbed system. The eigenstate of the unperturbed system is defined as

$$\hat{H}_0 \left| N_{fm}; N_{bm} \right\rangle = E_m \left| N_{fm}; N_{bm} \right\rangle \tag{14.3.1}$$

where $\left| N_{fm}; N_{bm} \right\rangle$ stands for a vector in the direct product space of fermions and bosons: N_{fm} stands for the m-th state of the fermions, while N_{bm} stands for the m-th state of the bosons.

When the system is perturbed, we denote the solution $\left| \psi \right\rangle$ as a linear superposition of the eigenstates of the unperturbed solution. That is

$$\left| \psi \right\rangle = \sum_m c_m \left(t \right) e^{-i\omega_m t} \left| N_{fm}; N_{bm} \right\rangle \tag{14.3.2}$$

where $\hbar\omega_m = E_m$. The state $\left| \psi \right\rangle$ evolves in time according to the equation

$$i\hbar \frac{\partial}{\partial t} \left| \psi \right\rangle = \left(\hat{H}_0 + \hat{H}_p \right) \left| \psi \right\rangle \tag{14.3.3}$$

Applying the time-dependent perturbation method as we have done previously, and testing the equation with $\left\langle N_{fq}; N_{bq} \right|$ gives

$$\dot{c}_q^{(1)}(t) = \frac{1}{i\hbar} \sum_m c_m^{(0)} e^{-i(\omega_m - \omega_q)t} \left\langle N_{fq}; N_{bq} \right| \hat{H}_p \left| N_{fm}; N_{bm} \right\rangle \tag{14.3.4}$$

If the starting state is assumed such that $c_s^{(0)} = 1$, and all other states are zero, then the above becomes

$$\dot{c}_q^{(1)}(t) = \frac{1}{i\hbar} e^{i(\omega_q - \omega_s)t} \left\langle N_{fq}; N_{bq} \right| \hat{H}_p \left| N_{fs}; N_{bs} \right\rangle \tag{14.3.5}$$

with

$$\hat{H}_p = \hat{H}_{ed} = \sum_{j,k,\lambda} H_{ed,\lambda,j,k} \hat{b}_j^\dagger \hat{b}_k \left(\hat{a}_\lambda - \hat{a}_\lambda^\dagger \right) \tag{14.3.6}$$

14.3.1 Absorption

The electron in an atom is assumed to have two states: a lower energy state E_1, and an upper energy state E_2. We assume a photon in mode λ_1 in the cavity. Then the starting state can be written as

$$\left| N_{fs}; N_{bs} \right\rangle = \hat{b}_1^\dagger \hat{a}_{\lambda_1}^\dagger \left| 0 \right\rangle \tag{14.3.7}$$

with energy

$$E_s = E_1 + \hbar\omega_{\lambda_1} \tag{14.3.8}$$

Consequently,

$$\hat{H}_p \left| N_{fs}; N_{bs} \right\rangle = \sum_{j,k,\lambda} H_{ed,\lambda,j,k} \hat{b}_j^\dagger \hat{b}_k \left(\hat{a}_\lambda - \hat{a}_\lambda^\dagger \right) \hat{b}_1^\dagger \hat{a}_{\lambda_1}^\dagger \left| 0 \right\rangle \tag{14.3.9}$$

It can be shown that

$$\hat{b}_j^\dagger \hat{b}_k \left(\hat{a}_\lambda - \hat{a}_\lambda^\dagger \right) \hat{b}_1^\dagger \hat{a}_{\lambda_1}^\dagger |0\rangle = \delta_{k1}\delta_{\lambda\lambda_1}\hat{b}_j^\dagger |0\rangle - \delta_{k1}\hat{b}_j^\dagger \hat{a}_\lambda^\dagger \hat{a}_{\lambda_1}^\dagger |0\rangle \tag{14.3.10}$$

In order for $\langle N_{fq}; N_{bq}| \hat{H}_p |N_{fs}; N_{bs}\rangle$ to be nonzero, it is necessary that $|N_{fq}; N_{bq}\rangle$ contains $\hat{b}_j^\dagger |0\rangle$ or $\hat{b}_j^\dagger \hat{a}_\lambda^\dagger \hat{a}_{\lambda_1}^\dagger |0\rangle$. If

$$|N_{fq}; N_{bq}\rangle = \hat{b}_j^\dagger |0\rangle \tag{14.3.11}$$

then $E_q = E_j$, and

$$\dot{c}_q^{(1)}(t) = \frac{1}{i\hbar} e^{i(\omega_j - \omega_1 - \omega_{\lambda_1})t} \sum_{j,k,\lambda} H_{ed,\lambda,j,k}\delta_{k1}\delta_{\lambda\lambda_1} \langle 0| \hat{b}_j \hat{b}_j^\dagger |0\rangle \tag{14.3.12}$$

$$= \frac{1}{i\hbar} \sum_j e^{i(\omega_j - \omega_1 - \omega_{\lambda_1})t} H_{ed,\lambda_1,j,1} \tag{14.3.13}$$

where $\hbar\omega_j = E_j$. The above is an oscillatory function of t unless $\omega_j - \omega_1 - \omega_{\lambda_1} = 0$.

When this happens, $\dot{c}_q^{(1)}(t)$ will integrate to a large value, indicating the high likelihood of transition to his eigenstate. Therefore, for this to happen, we need

$$E_2 - E_1 = \hbar\omega_{\lambda_1} \tag{14.3.14}$$

In this process, we start with the electron in E_1 and a photon in the cavity, and end up with the electron in E_2 and no photon as indicated by (14.3.11).

If we integrate the above equation, a sinc function results, which can be approximated by a delta function for long integration time, as was done previously. One can show that

$$|c_q^{(1)}|^2 \approx \frac{2\pi}{\hbar} t_0 |H_{ed\lambda_1,2,1}|^2 \delta(E_2 - E_1 - \hbar\omega_{\lambda_1}), \qquad t_o \to \infty \tag{14.3.15}$$

One can derive the transition rate as before to arrive at

$$w_q = \frac{2\pi}{\hbar} | H_{ed\lambda_1,2,1} |^2 \delta(E_2 - E_1 - \hbar\omega_{\lambda_1}) \tag{14.3.16}$$

In the above, the other possibility is for the final state to be

$$|N_{fq}; N_{bq}\rangle = \hat{b}_j^\dagger \hat{a}_\lambda^\dagger \hat{a}_{\lambda_1}^\dagger |0\rangle \tag{14.3.17}$$

with energy

$$E_q = E_j + \hbar\omega_\lambda + \hbar\omega_{\lambda_1} \tag{14.3.18}$$

Hence

$$E_q - E_s = E_j - E_1 + \hbar\omega_\lambda \tag{14.3.19}$$

and $E_q - E_s = 0$ in order for transition to occur or $c_q^{(1)}$ to be large. However, (14.3.19) cannot be zero. If $E_j = E_2$, $E_q - E_s > 0$ always. If $E_j = E_1$, then $\hbar\omega_\lambda > 0$. What this means is that it is not possible to start with an electron in the ground state and a photon in the cavity to end up with the electron in the excited state with the emission of a photon.

14.4   Spontaneous Emission

In the case of spontaneous emission, we assume that the starting state of the electron is in the excited state with energy E_2. Then, $E_s = E_2$ initially and

$$|N_{fs}; N_{bs}\rangle = \hat{b}_2^\dagger |0\rangle \tag{14.4.1}$$

Then

$$\hat{H}_p |N_{fs}; N_{bs}\rangle = \sum_{j,k,\lambda} H_{ed,\lambda,j,k} \; \hat{b}_j^\dagger \hat{b}_k \left(\hat{a}_\lambda - \hat{a}_\lambda^\dagger\right) \hat{b}_2^\dagger |0\rangle \tag{14.4.2}$$

The above can be reduced to

$$\hat{H}_p |N_{fs}; N_{bs}\rangle = -\sum_{j,k,\lambda} H_{ed,\lambda,j,k} \; \delta_{2k} \hat{b}_j^\dagger \hat{a}_\lambda^\dagger |0\rangle \tag{14.4.3}$$

Therefore, for $\langle N_{fq}; N_{bq} | \hat{H}_p | N_{fs}; N_{bs} \rangle$ to be nonzero, we need

$$|N_{fq}; N_{bq}\rangle = \hat{b}_j^\dagger \hat{a}_\lambda^\dagger |0\rangle \tag{14.4.4}$$

with

$$E_q = E_j + \hbar \omega_\lambda \tag{14.4.5}$$

Therefore

$$E_q - E_s = E_j - E_2 + \hbar \omega_\lambda \tag{14.4.6}$$

In order for a sizeable $c_q^{(1)}$, we need $E_j = E_1$, and then

$$E_2 - E_1 = \hbar \omega_\lambda \tag{14.4.7}$$

Then electron starts with the excited state E_2, spontaneously emits a photon, and drops to lower state E_1. Then energy of the emitted photon satisfies (14.4.7) by energy conservation.

14.5   Stimulated Emission

In this case, the electron in an atom is in the excited state. The presence of a photon in the cavity stimulates the emission of another photon from the electron. The initial state is

$$|N_{fs}; N_{bs}\rangle = \hat{b}_2^\dagger \hat{a}_{\lambda_1}^\dagger |0\rangle \tag{14.5.1}$$

with

$$E_s = E_2 + \hbar \omega_{\lambda_1} \tag{14.5.2}$$

We can show that

$$\hat{H}_p|N_{fs}; N_{bs}\rangle = \sum_{j,k,\lambda} H_{ed,\lambda,j,k} \; \hat{b}_j^\dagger \hat{b}_k \left(\hat{a}_\lambda - \hat{a}_\lambda^\dagger\right) \hat{b}_2^\dagger \hat{a}_{\lambda_1}^\dagger |0\rangle \tag{14.5.3}$$

$$= \sum_{j,k,\lambda} H_{ed,\lambda,j,k} \; \delta_{k2} \left(\delta_{\lambda\lambda_1} \hat{b}_j^\dagger - \hat{b}_j^\dagger \hat{a}_\lambda^\dagger \hat{a}_{\lambda_1}^\dagger\right) |0\rangle \tag{14.5.4}$$

In order for the above to transition to the final state, or one requires a non-zero result for

$$\langle N_{fq}; N_{bq}|\hat{H}_p|N_{fs}; N_{bs}\rangle$$

it is necessary that

$$|N_{fq}; N_{bq}\rangle = \hat{b}_j^\dagger \hat{a}_\lambda^\dagger \hat{a}_{\lambda_1}^\dagger |0\rangle \tag{14.5.5}$$

with energy

$$E_q = E_j + \hbar\omega_\lambda + \hbar\omega_{\lambda_1} \tag{14.5.6}$$

or

$$E_q - E_s = E_j - E_2 + \hbar\omega_\lambda = 0 \tag{14.5.7}$$

In other words, the only possibility is for $E_j = E_1$, yielding

$$E_2 - E_1 = \hbar\omega_\lambda \tag{14.5.8}$$

The above is just the spontaneous emission of a photon with the above energy, regardless if we already have a photon with energy $\hbar\omega_{\lambda_1}$ in the cavity. So it is not the most interesting case.

Next we consider the case when $\lambda = \lambda_1$. Then

$$|N_{fq}; N_{bq}\rangle = \frac{1}{\sqrt{2!}} \hat{b}_j^\dagger \left(\hat{a}_{\lambda_1}^\dagger\right)^2 |0\rangle \tag{14.5.9}$$

with

$$E_q = E_j + 2\hbar\omega_{\lambda_1} \tag{14.5.10}$$

Consequently,

$$\langle N_{fq}; N_{bq}|\hat{H}_p|N_{fs}; N_{bs}\rangle = H_{ed,\lambda_1,j,2}\langle 0|\frac{1}{\sqrt{2!}}(\hat{a}_{\lambda_1})^2 \hat{b}_j \hat{b}_j^\dagger \left(\hat{a}_{\lambda_1}^\dagger\right)^2 |0\rangle$$

$$= \sqrt{2!} H_{ed,\lambda_1,j,2}\langle 0|\frac{1}{\sqrt{2!}}(\hat{a}_{\lambda_1})^2 \hat{b}_j \frac{1}{\sqrt{2!}}\hat{b}_j^\dagger \left(\hat{a}_{\lambda_1}^\dagger\right)^2 |0\rangle$$

$$= \sqrt{2} H_{ed,\lambda_1,j,2} \tag{14.5.11}$$

The $\sqrt{2}$ factor is important implying that the transition is two times more likely to occur compared to the previous case. This is peculiar to stimulated emission where the emission of a photon is enhanced by the presence of a photon of the same frequency.

14.6 Multi-photon Case

In the multi-photon case, the transition rate for stimulated emission can be shown to be

$$w_q = \frac{2\pi}{\hbar}(n_{\lambda_1} + 1)|H_{ed,\lambda_1,1,2}|^2 \delta(E_1 - E_2 + \hbar\omega_{\lambda_1}) \tag{14.6.1}$$

implying that the presence of n_{λ_1} photons in the cavity enhances the emission by $(n_{\lambda_1} + 1)$ times.

The spontaneous emission, however, is not affected by the presence of photons of other frequencies in the cavity. For the absorption case, it can be shown that the formula is

$$w_q = \frac{2\pi}{\hbar}n_{\lambda_1}|H_{ed,\lambda_1,1,2}|^2 \delta(E_2 - E_1 - \hbar\omega_{\lambda_1}) \tag{14.6.2}$$

14.7 Total Spontaneous Emission Rate

When an electron emits a photon into the cavity, there are many modes with the same frequency that the emission can occur. In general, the total spontaneous emission rate is

$$W_{\text{spon}} = \sum_q w_q = \frac{2\pi}{\hbar}\sum_\lambda |H_{ed,\lambda,1,2}|^2 \delta(E_1 - E_2 + \hbar\omega_\lambda) \tag{14.7.1}$$

where in the above summation, we have replaced q with λ since λ is the index for the photon mode that the quantum state transition to in the spontaneous emission of one photon. Furthermore,

$$H_{ed,\lambda,1,2} = ie\sqrt{\frac{\hbar\omega_\lambda}{2\epsilon_0}} \int \phi_j^*(\mathbf{r})[\mathbf{u}_\lambda(\mathbf{r}) \cdot \mathbf{r}]\phi_k(\mathbf{r})d\mathbf{r}$$

$$\simeq ie\sqrt{\frac{\hbar\omega_\lambda}{2\epsilon_0}}\mathbf{u}_\lambda(\mathbf{r}_0) \cdot \mathbf{r}_{jk} \tag{14.7.2}$$

where

$$\mathbf{r}_{jk} = \int \phi_j^*(\mathbf{r})\mathbf{r}\phi_k(\mathbf{r})d\mathbf{r} \tag{14.7.3}$$

We have assumed here that $\mathbf{u}_\lambda(\mathbf{r}_0)$ is slowly varying compared to $\phi_l(\mathbf{r})$, $l = j,k$, and that $\phi_l(\mathbf{r})$ is highly localized around an atom.

The modes in the cavity can be made countable by imposing a periodic boundary condition arriving at

$$\mathbf{u}_\lambda(\mathbf{r}) = \mathbf{e}\frac{1}{\sqrt{V_b}}e^{i\mathbf{k}_\lambda \cdot \mathbf{r}} \tag{14.7.4}$$

where we have denoted the index of $\mathbf{k}$ by λ where ordinarily, it is omitted or implied.

The summation over different electromagnetic modes then becomes

$$\sum_\lambda \rightarrow \sum_{\text{pol}}\sum_{\mathbf{k}} \rightarrow \sum_{\text{pol}}\sum_{\mathbf{k}}\frac{V_b}{(2\pi)^3}\Delta\mathbf{k} = \sum_{\text{pol}}\int \frac{V_b}{(2\pi)^3}d\mathbf{k} \tag{14.7.5}$$
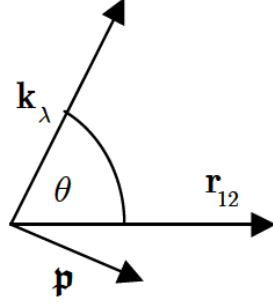
Figure 14.1: The vectors in the polarization of the emitted photon are aligned to simplify the calculation (from DAB Miller).

In the k space, the spacings between the modes are given by $(2\pi/L_x, 2\pi/L_y, 2\pi/L_z)$. Hence, each mode occupies a volume of $(2\pi)^3/V_b$ where $V_b = L_x L_y L_z$, and the number of modes in $d\mathbf{k}$ is $dV_b/(2\pi)^3 d\mathbf{k}$.

We can also pick the polarization such that one of them is always orthogonal to $\mathbf{r}_{12}$, so that

$$\mathbf{u}_\lambda(\mathbf{r}_0) \cdot \mathbf{r}_{12} = u_\lambda(\mathbf{r}_0)r_{12}\sin\theta \tag{14.7.6}$$

The above is also equivalent to making $\mathbf{r}_{12}$ to coincide with the z-axis of a spherical coordinate system, so that the radiation problem is axi-symmetric.

Consequently,

$$W_{\text{spon}} = \frac{2\pi}{\hbar} \int \frac{V_b}{(2\pi)^3} \left| ie\sqrt{\frac{\hbar\omega_k}{2\epsilon_0}} \frac{1}{\sqrt{V_b}} e^{i\mathbf{k}\cdot\mathbf{r}_0} r_{12}\sin\theta \right|^2 \delta(E_1 - E_2 + \hbar\omega_k)d\mathbf{k} \tag{14.7.7}$$

or

$$\begin{aligned}
W_{\text{spon}} &= \frac{e^2|r_{12}|^2}{8\pi^2\epsilon_0} \int \omega_k \sin^2\theta\delta(E_1 - E_2 + \hbar\omega_k)d\mathbf{k} \\
&= \frac{e^2|r_{12}|^2}{8\pi^2\epsilon_0} \int_{k=0}^\infty \int_{\theta=0}^\pi \omega_k\delta(E_1 - E_2 + \hbar\omega_k)2\pi\sin^3\theta k^2 d\theta dk
\end{aligned} \tag{14.7.8}$$

In the above, $\omega_k = ck$, or $\hbar ck = \hbar\omega_k$. Then

$$W_{\text{spon}} = \frac{e^2|r_{12}|^2}{4\pi\epsilon_0 c^3\hbar^4} \int_{k=0}^\infty \int_{\theta=0}^\pi (\hbar\omega_k)\delta(E_1 - E_2 + \hbar\omega_k)\sin^3\theta d\theta(\hbar\omega_k)^2 d(\hbar\omega_k) \tag{14.7.9}$$

Since

$$\int_{\theta=0}^\pi \sin^3\theta d\theta = -\int_{-1}^1 (1 - \cos^2\theta)d\cos\theta = \frac{4}{3} \tag{14.7.10}$$

then,

$$W_{\text{spon}} = \frac{e^2 |r_{12}|^2 \omega_{12}^3}{3\pi\epsilon_0 \hbar c^3} \tag{14.7.11}$$

where $\hbar\omega_{12} = E_2 - E_1$. The life time of a state is then $\tau = W_{\text{spon}}^{-1}$.

14.8 More on Atom-Field Interaction

It is shown that when an electron is in the presence of a field, the classical Hamiltonian for an atom is

$$H_A = \frac{1}{2m} \left[\mathbf{p} + e\mathbf{A}(\mathbf{r},t)\right]^2 - e\Phi(\mathbf{r},t) + V(\mathbf{r}) \tag{14.8.1}$$

Lorentz force law can be derived from the above. For quantum mechanics, we elevate $\mathbf{p}$ to $\hat{\mathbf{p}} = -i\hbar\nabla$ to become an operator arriving at

$$\hat{H}_A = \frac{1}{2m} \left[\hat{\mathbf{p}} + e\mathbf{A}\right]^2 - e\Phi + V(\mathbf{r}) \tag{14.8.2}$$

The scalar potential Φ and vector potential $\mathbf{A}$ describe an electromagnetic field, where

$$\mathbf{E}(\mathbf{r},t) = -\nabla\Phi - \partial_t\mathbf{A}(\mathbf{r},t) \tag{14.8.3}$$

$$\mathbf{B}(\mathbf{r},t) = \nabla \times \mathbf{A}(\mathbf{r},t) \tag{14.8.4}$$

The definitions of $\mathbf{A}$ and Φ are not unique. It is common to apply the radiation gauge where $\nabla \cdot \mathbf{A} = 0, \Phi = 0$, where $\mathbf{E} = -\partial_t\mathbf{A}(\mathbf{r},t)$.* In this gauge, $\mathbf{E}$ and $\mathbf{A}$ are linearly related, $\nabla \cdot \mathbf{E} = 0$, and it works for a source-free medium. Consequently,

$$\hat{H}_A = \frac{1}{2m} \left[\hat{\mathbf{p}} + e\mathbf{A}\right]^2 + V(\mathbf{r}) \tag{14.8.5}$$

In the long wavelength approximation, we let $\mathbf{A}(\mathbf{r},t) = \mathbf{A}(t)$ be independent of $\mathbf{r}$. In this case we can show that

$$(\hat{\mathbf{p}} + e\mathbf{A}) \, e^{-\frac{ie}{\hbar}\mathbf{A}(t)\cdot\mathbf{r}} \phi(\mathbf{r},t) = e^{-\frac{ie}{\hbar}\mathbf{A}(t)\cdot\mathbf{r}} \hat{\mathbf{p}}\phi(\mathbf{r},t) \tag{14.8.6}$$

$$(\hat{\mathbf{p}} + e\mathbf{A})^2 \, e^{-\frac{ie}{\hbar}\mathbf{A}(t)\cdot\mathbf{r}} \phi(\mathbf{r},t) = e^{-\frac{ie}{\hbar}\mathbf{A}(t)\cdot\mathbf{r}} \hat{\mathbf{p}}^2\phi(\mathbf{r},t) \tag{14.8.7}$$

But

$$\frac{\partial}{\partial t} e^{-\frac{ie}{\hbar}\mathbf{A}(t)\cdot\mathbf{r}} \phi(\mathbf{r},t) = -\frac{ie}{\hbar}\frac{\partial\mathbf{A}(t)}{\partial t} \cdot \mathbf{r}\phi(\mathbf{r},t) + e^{-\frac{ie}{\hbar}\mathbf{A}(t)\cdot\mathbf{r}}\partial_t\phi(\mathbf{r},t) \tag{14.8.8}$$

---

*This is Coulomb gauge with $\Phi = 0$.

Given the original Schrödinger equation of the form

$$\hat{H}_A \phi'(\mathbf{r}, t) = \left(\frac{1}{2m} \left[\hat{\mathbf{p}} + e\mathbf{A} \right]^2 + V(\mathbf{r}) \right) \phi'(\mathbf{r}, t) = i\hbar \frac{\partial}{\partial t} \phi'(\mathbf{r}, t) \qquad (14.8.9)$$

with the following transformation

$$\phi'(\mathbf{r}, t) = e^{-\frac{ie}{\hbar} \mathbf{A}(t) \cdot \mathbf{r}} \phi(\mathbf{r}, t) \qquad (14.8.10)$$

and making use of (14.8.6) to (14.8.8), it becomes

$$\left[\frac{1}{2m} \hat{\mathbf{p}}^2 + V(\mathbf{r}) - e\partial_t \mathbf{A}(t) \cdot \mathbf{r} \right] \phi(\mathbf{r}, t) = i\hbar \partial_t \phi(\mathbf{r}, t) \qquad (14.8.11)$$

Using the fact that $\mathbf{E} = -\partial_t \mathbf{A}$, the above becomes

$$\left[\frac{1}{2m} \hat{\mathbf{p}}^2 + V(\mathbf{r}) + e\mathbf{E} \cdot \mathbf{r} \right] \phi(\mathbf{r}, t) = i\hbar \partial_t \phi(\mathbf{r}, t) \qquad (14.8.12)$$

The above is equivalent to performing a similarity transform to the Hamiltonian $\hat{H}_A$ as

$$\hat{H}_A' = e^{\frac{ie}{\hbar} \mathbf{A}(t) \cdot \mathbf{r}} \hat{H}_A e^{-\frac{ie}{\hbar} \mathbf{A}(t) \cdot \mathbf{r}} \qquad (14.8.13)$$

or a change of basis from

$$\phi'(\mathbf{r}, t) = e^{-\frac{ie}{\hbar} \mathbf{A}(t) \cdot \mathbf{r}} \phi(\mathbf{r}, t) \qquad (14.8.14)$$

However, the total Hamiltonian of the atom-field system is

$$\hat{H} = \hat{H}_A + \hat{H}_F \qquad (14.8.15)$$

where

$$\hat{H}_F = \sum_\lambda \frac{1}{2} \hbar \omega_\lambda [\hat{a}_\lambda^\dagger \hat{a}_\lambda + \hat{a}_\lambda \hat{a}_\lambda^\dagger] \qquad (14.8.16)$$

In a fully quantum system $\mathbf{A} \to \hat{\mathbf{A}}$ or fields are elevated to operators in $\hat{H}_A$ or in (14.8.5). It can be shown that the above similarity transform works ever if we include $\hat{H}_F$ part in the Hamiltonian. After the similarity transform, the above becomes

$$\hat{H}' = \hat{H}_{A0} + \hat{H}_F' + \hat{H}_I' \qquad (14.8.17)$$

where

$$\hat{H}_{A0} = \frac{\hat{\mathbf{p}}^2}{2m} + V(\mathbf{r}) \qquad (14.8.18)$$

and

$$\hat{H}_I' = e\hat{\mathbf{E}} \cdot \mathbf{r} \qquad (14.8.19)$$

H_F' retains the form similar to H_F.[†] The above manipulation is also known as the electric field gauge or E-gauge.

[†]See C. Cohen-Tannoudji, J. Dupont-Roc and G. Grynberg, *Photons and Atoms: Introduction to Quantum Electrodynamics*, Wiley Professional, Feb 1997.

Another way to derive the dipole interaction term is to rewrite (14.8.5) as

$$\hat{H}_A = \frac{\hat{\mathbf{p}}^2}{2m} + \frac{e\mathbf{A} \cdot \hat{\mathbf{p}}}{m} + \frac{e^2 \mathbf{A} \cdot \mathbf{A}}{2m} + V(\mathbf{r}) \tag{14.8.20}$$

where we have assumed that $\hat{\mathbf{p}} \cdot \mathbf{A} = 0$ because $\nabla \cdot \mathbf{A} = 0$. Since $\hat{\mathbf{p}}$ acts on the atomic orbital wave functions, which are rapidly varying, we can assume that $\hat{\mathbf{p}} \gg e\mathbf{A}$. Then (14.8.20) becomes

$$\hat{H}_A \cong \frac{\hat{\mathbf{p}}^2}{2m} + V(\mathbf{r}) + \frac{e\mathbf{A} \cdot \hat{\mathbf{p}}}{m} \tag{14.8.21}$$

The above requires no similarity transform. When the field Hamiltonian is added, and all fields are elevated to operators, we can write the total Hamiltonian as

$$\hat{H} = \hat{H}_{A0} + \hat{H}_F + \hat{H}_I \tag{14.8.22}$$

where $\hat{H}_{A0}$ is as defined before in (14.8.18), and

$$\hat{H}_I = \frac{e\hat{\mathbf{A}} \cdot \hat{\mathbf{p}}}{m} \tag{14.8.23}$$

14.8.1 Interaction Picture

If the quantum system is initially described by a Hamiltonian $\hat{H}_0$, and with the addition of the interaction Hamiltonian $\hat{H}_I$, we can write the exact solution of the quantum system in terms of the interaction picture. For Schrödinger equation

$$i\hbar \frac{\partial}{\partial t} |\psi\rangle = \left(\hat{H}_0 + \hat{H}_I \right) |\psi\rangle \tag{14.8.24}$$

we first let

$$|\psi\rangle = e^{-\frac{i}{\hbar} \hat{H}_0 t} |\phi\rangle \tag{14.8.25}$$

Then

$$i\hbar \frac{\partial}{\partial t} |\psi\rangle = \hat{H}_0 |\psi + e^{-\frac{i}{\hbar} \hat{H}_0 t} \frac{\partial}{\partial t} |\phi\rangle \tag{14.8.26}$$

Using the above in (14.8.24), we have

$$i\hbar \frac{\partial}{\partial t} |\phi\rangle = e^{\frac{i}{\hbar} \hat{H}_0 t} \hat{H}_I e^{-\frac{i}{\hbar} \hat{H}_0 t} |\phi\rangle \tag{14.8.27}$$

The above equation can be integrated formally to yield

$$|\phi(t)\rangle = \exp\left(-\frac{i}{\hbar} \int_0^t dt' e^{\frac{i}{\hbar} \hat{H}_0 t'} \hat{H}_I e^{-\frac{i}{\hbar} \hat{H}_0 t'} \right) |\phi(0)\rangle \tag{14.8.28}$$

If instead we use a perturbation concept and let

$$|\phi\rangle \approx \sum_m \left[c_m^{(0)} + c_m^{(1)}(t) \right] |\phi_m\rangle \tag{14.8.29}$$

Using the fact that $\partial_t c_m^{(0)} = 0$, $\partial_t |\phi_m\rangle = 0$, because $|\phi_m\rangle$ is a stationary state with no time dependence, from (14.8.27), we have

$$i\hbar \sum_m \dot{c}_m^{(1)} |\phi_m\rangle = \sum_m c_m^{(0)} e^{\frac{i}{\hbar} \hat{H}_0 t} \hat{H}_I e^{-\frac{i}{\hbar} \hat{H}_0 t} |\phi_m\rangle \tag{14.8.30}$$

Assuming that $c_m^{(0)} = \delta_{sm}$, i.e., the quantum system is in only one initial state, then testing (14.8.30) with $|\phi_q\rangle$ where $|\phi_q\rangle$ is a final stationary state, we have

$$i\hbar \dot{c}_q^{(1)} = e^{i(\omega_q - \omega_s)t} \langle \phi_q | \hat{H}_I | \phi_s \rangle \tag{14.8.31}$$

The above is the same as the one derived from time-dependent perturbation theory.

14.8.2 $\mathbf{E} \cdot \mathbf{r}$ or $\mathbf{A} \cdot \mathbf{p}$ Interaction

To find the time evolution of the quantum system due to interaction between the atom and the field, we need to evaluate

$$\langle \phi_q | \hat{H}_I | \phi_s \rangle \tag{14.8.32}$$

However, the interaction Hamiltonian can be written as

$$\hat{H}_I = e \hat{\mathbf{E}} \cdot \mathbf{r} \tag{14.8.33}$$

in one case, and

$$\hat{H}'_I = e \hat{\mathbf{A}} \cdot \mathbf{p}/m \tag{14.8.34}$$

in another case. In order to reconcile the difference, we note that in (14.8.14), ϕ' and ϕ are related by a transform. When this transform is accounted for, the difference disappears.[‡]

[‡]For details, see M.O. Scully and M. Suhail Zubairy, *Quantum Optics*, CUP, 1997.

Chapter 15

Quantum Interpretation and Quantum Information

15.1 Introduction

Before the 1980s, the interpretation of quantum mechanics has been fraught with controversies. The two main schools of thought was one that followed the Copenhagen school led by Niels Bohr (1885-1962), and the other one that followed Einstein's school led by Albert Enstein (1879-1955). It was not until 1982 that Alain Aspect (1947-) did some conclusive experiments in favor of the Copenhagen school. Single photon sources were not created until 1974 by John F. Clauser (1942-). It was not until then that experiment could confirm that photon behaves like a quantum particle. Then photon can be used to test the tenet of quantum interpretation.

One important tenet of quantum mechanics is that one does not know what the state of the system is until one performs a measurement. Before the measurement, the quantum system is described by a state that can be in a linear superposition of different states. After the measurement, the system collapses to the state that is "discovered" by the measurement. It is the existence as a linear superposition of states that greatly enriches the information content of a quantum system. The possibility of a system to be simultaneously in different states is peculiar to quantum mechanics. Objects in the classical world cannot be in such a state. It puts quantum systems in the realm of "ghosts" and "angels" in fairy tales and ghost stories of different cultures; these ghosts and angels can be simultaneously in many places or in different states. This "spookiness" of quantum mechanics is only recently confirmed by experiments in the late 70's and early 80's. It is because of these "ghost-angel" states, that we can have quantum cryptography, quantum communication, quantum circuits, and quantum computing.

An uneasiness about the philosophical interpretation of quantum mechanics is that one does not know what state the quantum system is in before the measurement. This has prompted Einstein to ask,"Is the moon there if you don't look at it?" The uncertainty of the state applied to quantum mechanics is only true for a linear superposition of coherent

quantum states, which I term the ghost-angel state. This state has not been found to exist for macroscopic objects.

However, if one insists that, "One does not know if the moon is there before one looks at it." as a true statement, it cannot be refuted nor confirmed by experiments. The mere act of an experiment already means that we have "looked" at the moon. The same claim goes that "If I saw a fallen tree in the forest, it did not necessary follow from the act of falling before I arrived there." Alternatively, "If we found dinosaur bones, it did not necessary mean that dinosaurs roamed the earth over 200 million years ago." We believe that the moon is there even if we do not look at it, the tree fell because it went through the act of falling, and that dinosaur bones were found because they roamed the earth 200 million years ago, because we believe in the realism of the world we live in. This realism cannot be proved but is generally accepted by those who live in this world. Hence, it is this surreal interpretation of quantum mechanics that causes the uneasiness among many physicists. But the interpretation of quantum mechanics is slightly better than the above: a pure quantum state is in a linear superposition of states, the precise one of which we are not sure of until a measurement is performed. However, this surrealism of this ghost-angel state exists in our minds in fairy tales and ghost stories of many cultures. Experimental effort has agreed with the surreal interpretation of quantum mechanics in terms of the Bell's theorem, that will be discussed.

The ghost-angel state of a quantum system is what enriches the information in it. However, for a quantum system to be in such a state, the linear superposition of states must be coherent with each other. Quantum coherence is the largest stumbling block to the construction of quantum computers; however, rapid advances are being made, and one day, it can be a reality.

15.2   Interpretation of Quantum Mechanics

As mentioned before, quantum mechanics has the basic tenet that a quantum state is in a linear superposition of states before a measurement. A measurement projects a quantum state into one of the states. The ability of a quantum state to be in a linear superposition of states is surreal, and it has bothered a great many physicists. In the classical world, a system can only be in one state or the other, but not in a linear superposition of states. Only the world of ghosts and angels can we imagine that an object is in a linear superposition of states.

In the coordinate space, an electron, represented by its wavefunction, can be simultaneously at all locations where the wavefunction is non-zero. In the Young's double slit experiment, the electron, represented by its wavefunction, can go through both slits simultaneously like a wave.*

When the ghost-angel state concept is extended to classical objects, such as a cat, it gives rise to the ludicrous result: the story of the Schrödinger cat. The Schrödinger cat is a linear superposition of a dead cat and a live cat. To understand why the Schrödinger cat does not exist, we need to understand the concept of quantum coherence.

Two states are in quantum coherence if the phase relationships between them are deterministic and not random. When this coherence is lost, the phase relationship between them is lost. The quantum system has already collapsed into one of the two states. Hence, in

---

*Or the apparition of the ghost-angel state.

practice, a measurement is not always necessary before the quantum system collapses into one or more of the quantum states. The interaction of quantum system with its environment can cause such a collapse. These are called mixed states in the language of density matrix.

From a statistical physics viewpoint, it is impossible for a quantum system to be completely isolated. Almost all systems are in a thermal bath of the universe with which they are seeking equilibrium. Macroscopic objects cannot be in a pure quantum state which has the characteristics of the ghost-angel state. It is impossible for the huge number of atoms in the Schrödinger cat to be coherent with respect to each other.

The density matrix is a nice way of representing a state of a quantum system where the physics of quantum coherence surfaces explicitly. This concept is expressed in the off-diagonal terms of the density matrix. If one allows time average or ensemble average[†] to the density matrix, when the system is expressed by quantum states that are not coherent, the off-diagonal elements will average to zero. The system is in a mixed state rather than a pure quantum state. The system is similar to the local hidden variable theory: the state of the quantum system is already predetermined before the measurement.

15.3   Entangled States

A photon has to be a causal wave packet since it could be emitted by say, spontaneous emission from an atomic transition. The spontaneous emission gives rise to a packet of energy zipping through space at the speed of light. Since this packet of energy has a beginning and an end, it can be described by a wave packet. The particle that is associated with this wave packet is a photon, carrying a quantum of energy equal to $\hbar\omega$. Since a photon is necessary localized, the wave packet cannot be purely monochromatic, but it has to be quasi-monochromatic, or it consists of many modes.

A multi-mode photon can be described by the following field operator

$$\hat{\mathbf{E}}(\mathbf{r},t) = \sum_{\mathbf{k},s} \frac{1}{2}\sqrt{\frac{\hbar\omega_k}{V\epsilon_0}}\mathbf{e}_s \hat{a}_{\mathbf{k},s} e^{i\mathbf{k}\cdot\mathbf{r}-i\omega t} + \text{c.c} \tag{15.3.1}$$

with the corresponding Hamiltonian given by

$$\hat{H} = \sum_{k,s} \hat{H}_{k,s} \tag{15.3.2}$$

with the Fock state given by

$$|\psi\rangle = \prod_{k,s} |n_{k,s}\rangle \tag{15.3.3}$$

For a quasi-monochromatic case, when only one frequency component is important, the dominant mode above can be kept, and the electric field operator can be approximated by

$$\hat{\mathbf{E}}(\mathbf{r},t) = \sum_{s} \frac{1}{2}\sqrt{\frac{\hbar\omega_k}{V\epsilon_0}}\mathbf{e}_s \hat{a}_{\mathbf{k},s} e^{i\mathbf{k}\cdot\mathbf{r}-i\omega t} + \text{c.c} \tag{15.3.4}$$

[†]Processes for which time average is the same as ensemble average are known as ergodic processes.

with the corresponding Hamiltonian being

$$\hat{H} = \sum_s \hat{H}_{k,s} = \hat{H}_{k,v} + \hat{H}_{k,h} = \hat{H}_{k,v'} + \hat{H}_{k,h'} \tag{15.3.5}$$

Since any two orthogonal polarizations represented by

$$\hat{\mathbf{E}}(\mathbf{r}, t) \sim \mathbf{e}_v \hat{a}_{\mathbf{k},v} + \mathbf{e}_h \hat{a}_{\mathbf{k},h} = \mathbf{e}_{v'} \hat{a}_{\mathbf{k},v'} + \mathbf{e}_{h'} \hat{a}_{\mathbf{k},h'} \tag{15.3.6}$$

are complete (ignoring the c.c. part) to represent the field, one can pick any two orthogonal polarizations. If only one photon is present, then it is in the linear superposition of these two orthogonal states as a quantum particle. The corresponding Fock state is

$$|\psi\rangle = \prod_s |n_{k,s}\rangle \tag{15.3.7}$$

Hence, the single photon Fock state that the field operator acts on can be denoted by

$$|\psi\rangle = C_v |1_{\mathbf{k},v}\rangle |0_{\mathbf{k},h}\rangle + C_h |0_{\mathbf{k},v}\rangle |1_{\mathbf{k},h}\rangle = C_{v'} |1_{\mathbf{k},v'}\rangle |0_{\mathbf{k},h'}\rangle + C_{h'} |0_{\mathbf{k},v'}\rangle |1_{\mathbf{k},h'}\rangle \tag{15.3.8}$$

For simplicity, the unoccupied number state is usually omitted, and the above can be written as

$$|\psi\rangle = C_v |1_{\mathbf{k},v}\rangle + C_h |1_{\mathbf{k},h}\rangle = C_{v'} |1_{\mathbf{k},v'}\rangle + C_{h'} |1_{\mathbf{k},h'}\rangle \tag{15.3.9}$$

The above denotes a photon in a purely monochromatic $\mathbf{k}$ state with a linear superposition of two orthogonal polarizations. It has a packet of energy $E = \hbar \omega_k$. However, a photon with a purely monochromatic $\mathbf{k}$ state is not localized. But a photon generated by a source like an atomic transition must be causal, and hence, localized. A localized wave packet describing this photon field can be constructed by using a linear superposition of wavenumber $\mathbf{k}$ or frequencies. For high frequency photons, this localized state can have a center frequency with a small spread of frequencies around the center frequency. The single-photon Fock state can hence be written as

$$|\psi\rangle = \sum_{\mathbf{k}} C_{\mathbf{k},v} |1_{\mathbf{k},v}\rangle + C_{\mathbf{k},h} |1_{\mathbf{k},h}\rangle \tag{15.3.10}$$

For quasi-monochromatic photons, the above will be dominated by one frequency term and we can denote this photon approximately with the state vector (15.3.8).

With the above picture in mind, we can think of two localized, quasi-monochromatic photons traveling in different directions. A direct product space can be used to represent the eigenstate of these two photons:

$$|\psi\rangle_{ab} = |1_{\mathbf{k}_a,v}\rangle_a |1_{\mathbf{k}_b,v}\rangle_b \tag{15.3.11}$$

For simplicity, we denote these two-photon eigen state as

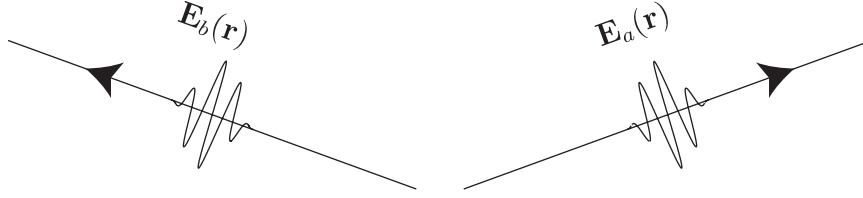$$|\psi\rangle_{12} = |V\rangle_1 |V\rangle_2 \tag{15.3.12}$$

Figure 15.1: Two photons traveling in different directions.

for subsequent discussion representing two photons zipping in different directions.

If the two photons are generated from the same source, entangled photon states may result. Entangled two-particle states are those that cannot be written as a product (outer product or tensor product) of simpler states. An example is the EPR (Einstein, Podolsky, and Rosen) pair

$$|\Phi^+\rangle_{12} = \frac{1}{\sqrt{2}}\big(|H\rangle_1|H\rangle_2 + |V\rangle_1|V\rangle_2\big) \tag{15.3.13}$$

Other entangled states are

$$|\Phi^-\rangle_{12} = \frac{1}{\sqrt{2}}\big(|H\rangle_1|H\rangle_2 - |V\rangle_1|V\rangle_2\big) \tag{15.3.14}$$

$$|\Psi^+\rangle_{12} = \frac{1}{\sqrt{2}}\big(|H\rangle_1|V\rangle_2 + |V\rangle_1|H\rangle_2\big) \tag{15.3.15}$$

$$|\Psi^-\rangle_{12} = \frac{1}{\sqrt{2}}\big(|H\rangle_1|V\rangle_2 - |V\rangle_1|H\rangle_2\big) \tag{15.3.16}$$

The above four states are also called the Bell states. They are usually generated due to the particle pair needing to satisfy conservation of angular momentum. For instance, two photons are generated by atomic transitions where the initial angular momentum of the system is zero.

An angular momentum conserving state with two counter-propagating photon is

$$|\Phi\rangle = \frac{1}{\sqrt{2}}\big(|R\rangle_1|R\rangle_2 + |L\rangle_1|L\rangle_2\big) \tag{15.3.17}$$

where R and L stand for right-handed and left-handed circular polarizations, respectively. By letting

$$|R\rangle_1 = (|H\rangle_1 + i|V\rangle_1)\,, \quad |R\rangle_2 = (|H\rangle_2 + i|V\rangle_2) \tag{15.3.18}$$
$$|L\rangle_1 = (|H\rangle_1 - i|V\rangle_1)\,, \quad |L\rangle_2 = (|H\rangle_2 - i|V\rangle_2) \tag{15.3.19}$$

which is a change of basis, the above becomes

$$|\Phi\rangle = -\frac{1}{\sqrt{2}}\big(|H\rangle_1|H\rangle_2 - |V\rangle_1|V\rangle_2\big) \tag{15.3.20}$$

one of the Bell states.

The entangled states are bewildering because it means that for two counter-propagating photons, if one measures one photon is in $|H\rangle$ state, the other photon immediately collapses to an $|H\rangle$ state as well, regardless of how far apart the two photons are. Whereas before the measurement, the photons are in a linear superposition of a $|H\rangle$ and $|V\rangle$ states.

15.3.1   EPR Paradox

The interpretation of quantum mechanics went through difficult times. The fact that a particle can be in a superposition of many states before a quantum measurement, and the probabilistic interpretation of a quantum measurement behooves the challenge by many great physicists, especially Einstein. The most severe challenge of quantum mechanics and its interpretation comes from the EPR (Einstein, Podolsky and Rosen) paradox. To describe it simply, we imagine a π meson that decays into an electron-positron pair:

$$\pi^o \rightarrow e^- + e^+ \tag{15.3.21}$$

The π meson originally has spin zero. So for conservation of angular momentum, the electron-positron pair will have opposite spins: spin up and spin down. Since the total angular momentum is zero, they are in the singlet state which has total spin of zero, or

$$|\Psi\rangle = \frac{1}{\sqrt{2}} \left(|\uparrow_-\downarrow_+\rangle - |\downarrow_-\uparrow_+\rangle \right) \tag{15.3.22}$$

The electron-positron pair is in the linear superposition of two states, but the electron and positron are flying in opposite directions. According to the interpretation of quantum mechanics, one does not know the spin state of the electron nor the positron before the measurement. After one measures the spin state of, say the electron, irrespective of how far the positron is away from the electron, we immediately know the spin state of the positron according to the above equation. The spins of the two particles are always opposite to each other. This notion is unpalatable to many physicists, and hence, is called "spooky action at a distance" by Einstein. Information cannot travel faster than the speed of light. How could the state of one particle be immediately determined after a measurement is made at another particle far away? This paradox attempts to prove that quantum mechanics is incomplete by *reductio ad absurdum*.

15.4   Bell's Theorem[‡]

Quantum measurements are known to be random, and the data can only be interpreted probabilistically. If one were to measure the spin of one of the particle, it is equally likely to be in the spin up or spin down state randomly according to (15.3.22). In the hidden variable theory, it is suggested that the outcome of the experiment is already predetermined even before the measurement. The outcome is determined by a hidden random variable λ. It is the randomness of this variable that gives rise to the randomness of the outcome in quantum measurements. This is contrary to the now accepted quantum interpretation[§] where

---

[‡]This section is written with important input from Y. H. Lo and Q. Dai.

[§]This was espoused by the Copenhagen school.

one does not know what state a quantum system is until after a measurement is performed. Before the measurement, the quantum system can be in a linear superposition of states. The measurement collapses the quantum system into the state "discovered" by the measurement.

Many hidden variable theories were proposed shortly after the EPR paradox was published. In 1964, John Steward Bell (1928-1990), in the spirit of proving the correctness of the hidden variable theory, came up with an inequality that showed the incompatibility of quantum mechanics and the hidden variable theory. If hidden variable theory is correct, the inequality will be satisfied, but if quantum mechanics is correct, the inequality is violated. This is known as Bell's theorem.

We can discuss the derivation of the Bell's theorem in the context of the two-photon experiment, since the experiment that verifies the theorem has been done using photons. The actual experiment is quite complex, but we will reduce it to a simplified case. The simplified experiment involves a photon source that produces an entangled photon pair, each of which is traveling in opposite directions. The photon pair is in one of the Bell state, say the EPR pair:

$$|\Psi\rangle = \frac{1}{\sqrt{2}} \left(|V_1 V_2\rangle + |H_1 H_2\rangle \right) \tag{15.4.1}$$

In the above state, if one of the photons is measured to be V (vertical) polarized, the other photon must be V polarized. However, if one photon is measured to be H (horizontal) polarized, the other photon must be H polarized. We will detect the photon state with a simple polarizer. In the above state, which is a linear superposition of two states, the photons are equally likely to be found in the first state, $|V_1 V_2\rangle$, or the second state, $|H_1 H_2\rangle$. The polarizer will detect an H or a V polarization with equal likelihood, but the moment that one photon is determined to be H polarized, the other photon is immediately known to be H polarized, and vice versa. This is the "spookiness" of quantum interpretation.

Imagine an atomic source that generates two photons propagating in opposite directions. The atom initially has zero angular momentum, so that the two photons are either both horizontally polarized or both vertically polarized. Hence, the photon can be described in one of the Bell states or an entangled state as shown in (15.4.1).
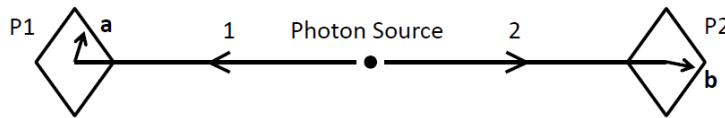

Figure 15.1: Experimental verification of Bell's theorem

15.4.1   Prediction by Quantum Mechanics

First, we will predict the experimental outcome by invoking the quantum measurement hypothesis and quantum interpretation. Photon 1 is measured with a polarizer P_1 with vertical

polarization pointing in the $\mathbf{a}$ direction. If P_1 measures a V polarization, we set $A(\mathbf{a}) = 1$, and if P_1 measures an H polarization, we set $A(\mathbf{a}) = -1$. Here, $A(\mathbf{a})$ denotes the measurement outcome, and it is completely random according to quantum mechanics.

Similarly, polarizer P_2 has its vertical polarization oriented in the $\mathbf{b}$ direction. When it measures a V polarization, we set $B(\mathbf{b}) = 1$, and when it measures an H polarization, it sets $B(\mathbf{b}) = -1$. Again, $B(\mathbf{b})$ is completely random. In the above, we set A and B to be functions of $\mathbf{a}$ and $\mathbf{b}$ respectively, as the experimental outcomes are expected to be functions of the orientation of the polarizers.

If $\mathbf{a} = \mathbf{b}$, we expect that

$$\langle A(\mathbf{a})B(\mathbf{b})\rangle = E(\mathbf{a}, \mathbf{b}) = AB = 1 \tag{15.4.2}$$

If $\mathbf{a} \perp \mathbf{b}$, if P_1 measures a V polarization, P_2 will measure a H polarization, we expect that

$$\langle A(\mathbf{a})B(\mathbf{b})\rangle = E(\mathbf{a}, \mathbf{b}) = AB = -1 \tag{15.4.3}$$

Even though A and B are random, their products are deterministic in the above two cases.

An interesting case ensues if $\mathbf{a}$ and $\mathbf{b}$ are at an incline with respect to each other. This shows the distinguishing feature of a quantum particle subject to quantum interpretation. In this case, we know by quantum mechanics that

$$|V\rangle_a = \cos\theta|V\rangle_b - \sin\theta|H\rangle_b \tag{15.4.4}$$

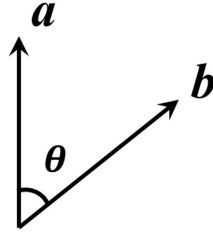$$|H\rangle_a = \sin\theta|V\rangle_b + \cos\theta|H\rangle_b \tag{15.4.5}$$

Figure 15.2: The case when the two polarizers are at an incline with respect to each other for proving the Bell's theorem.

If P_1 measures an outcome with $A(\mathbf{a}) = 1$, then the photon that propagates to P_2 must be polarized in the $\mathbf{a}$ direction. However, according to (15.4.4), for P_2, such a photon has the probability of $\cos^2\theta$ being detected in the V polarization, and the probability of $\sin^2\theta$ being detected in the H polarization. Hence, the expectation value of B, or $\langle B\rangle_{A=1} = \cos^2\theta - \sin^2\theta$. The above can also be written as

$$\begin{aligned}
\langle B\rangle_{A=1} &= \sum_B P(B|A=1)B \\
&= P(B=1|A=1) - P(B=-1|A=1) \\
&= \cos^2\theta - \sin^2\theta
\end{aligned} \tag{15.4.6}$$

where $B = \pm 1$, $P(B = 1|A = 1) = \cos^2 \theta$, and $P(B = -1|A = 1) = \sin^2 \theta$.

If P_1 finds that $A = -1$, by similar arguments, the expectation value of B or $\langle B \rangle_{A=-1} = \sin^2 \theta - \cos^2 \theta$, or similar to before,

$$\langle B \rangle_{A=-1} = \sum_B P(B|A = -1)B$$
$$= P(B = 1|A = -1) - P(B = -1|A = -1)$$
$$= \sin^2 \theta - \cos^2 \theta \qquad (15.4.7)$$

where $P(B = 1|A = -1) = \sin^2 \theta$, and $P(B = -1|A = -1) = \cos^2 \theta$. Then using Bayes' theorem, it can be shown that

$$\langle AB \rangle = \sum_{A,B} P(A, B)AB = \sum_{A,B} P(B|A)P(A)AB = \sum_A \left(\sum_B P(B|A)B \right) AP(A)$$
$$= P(A = 1) \sum_B P(B|A = 1)B - P(A = -1) \sum_B P(B|A = -1)B$$
$$= P(A = 1)\langle B \rangle_{A=1} - P(A = -1)\langle B \rangle_{A=-1} \qquad (15.4.8)$$

where (15.4.6) and (15.4.7) have been used in the above to evaluate to

$$E(\mathbf{a}, \mathbf{b}) = \langle A(\mathbf{a})B(\mathbf{b}) \rangle = [P(A = 1) + P(A = -1)] \left(\cos^2 \theta - \sin^2 \theta \right)$$
$$= \cos^2 \theta - \sin^2 \theta = \cos(2\theta) \qquad (15.4.9)$$

since $P(A = 1) + P(A = -1) = 1$. Notice that the above reduces to the special cases of: (i) when the polarizers P_1 and P_2 are aligned, namely when $\theta = 0°$, as in (15.4.2), and (ii) when the polarizers are perpendicular to each other, with $\theta = 90°$, as in (15.4.3).

15.4.2 Prediction by Hidden Variable Theory

In the hidden variable theory derivation, a particle is assumed to be already in a polarization state even before a measurement. A and B are random, but they are predetermined by a hidden random variable λ. We let

$$A(\mathbf{a}, \lambda) = \pm 1 \qquad (15.4.10)$$

$$B(\mathbf{b}, \lambda) = \pm 1 \qquad (15.4.11)$$

Here, A and B hence flip-flop between ± 1 entirely due to randomness of the variable λ. This theory predicts the randomness of quantum mechanics experiments nicely. The expectation value of AB then is

$$E(\mathbf{a}, \mathbf{b}) = \langle AB \rangle = \int \rho(\lambda)A(\mathbf{a}, \lambda)B(\mathbf{b}, \lambda)d\lambda \qquad (15.4.12)$$

where $\rho(\lambda)$ is the probability distribution of the random variable λ. The above is a very general representation of the hidden variable theory where we have not explicitly stated the functions $\rho(\lambda)$, $A(\mathbf{a}, \lambda)$ nor $B(\mathbf{b}, \lambda)$.

If $\mathbf{a} = \mathbf{b}$, then

$$A(\mathbf{a}, \lambda) = B(\mathbf{a}, \lambda) \tag{15.4.13}$$

and the above becomes

$$E(\mathbf{a}, \mathbf{a}) = \int \rho(\lambda) A^2(\mathbf{a}, \lambda) d\lambda = \int \rho(\lambda) d\lambda = 1 \tag{15.4.14}$$

Same as we would have found in (15.4.2).

If $\mathbf{a} \perp \mathbf{b}$, then

$$A(\mathbf{a}, \lambda) = -B(\mathbf{b}, \lambda) \tag{15.4.15}$$

and (15.4.12) becomes

$$E(\mathbf{a}, \mathbf{b}) = -\int \rho(\lambda) A^2(\mathbf{a}, \lambda) d\lambda = -\int \rho(\lambda) d\lambda = -1 \tag{15.4.16}$$

as would have been found in (15.4.3). Hence, hidden variable theory is in good agreement with quantum mechanics interpretation of (15.4.2) and (15.4.3).

So far, everything is fine and dainty until when $\mathbf{a}$ and $\mathbf{b}$ do not belong to any of the above category, but in general are inclined with respect to each other. If $\mathbf{a} \neq \mathbf{b}$ in general, then the hidden variable generates random A and B in such a manner that

$$|E(\mathbf{a}, \mathbf{b})| \leq 1 \tag{15.4.17}$$

The above is less than one because now $\mathbf{a}$ and $\mathbf{b}$ are not exactly correlated or anti-correlated. This is quite different from why (15.4.9) is less than one when $\theta \neq 0$. As we shall see, this gives rise to quite a different nature for $E(\mathbf{a}, \mathbf{b})$ for hidden variable theory.

To this end, we can show that

$$E(\mathbf{a}, \mathbf{b}) - E(\mathbf{a}, \mathbf{c}) = \int [A(\mathbf{a}, \lambda) B(\mathbf{b}, \lambda) - A(\mathbf{a}, \lambda) B(\mathbf{c}, \lambda)] \rho(\lambda) d\lambda \tag{15.4.18}$$

Since $A(\mathbf{a}, \lambda) B(\mathbf{a}, \lambda)$, $B(\mathbf{a}, \lambda)$ can be removed from the above and it can be rewritten as

$$E(\mathbf{a}, \mathbf{b}) - E(\mathbf{a}, \mathbf{c}) = \int [A(\mathbf{a}, \lambda) A(\mathbf{b}, \lambda) - A(\mathbf{a}, \lambda) A(\mathbf{c}, \lambda)] \rho(\lambda) d\lambda$$

$$= \int A(\mathbf{a}, \lambda) A(\mathbf{b}, \lambda) [1 - A(\mathbf{b}, \lambda) A(\mathbf{c}, \lambda)] \rho(\lambda) d\lambda \tag{15.4.19}$$

A factor $A^2(\mathbf{b}, \lambda) = 1$ has been added to the above so that it can be factorized. Also, it can be seen that

$$-1 \geq A(\mathbf{b}, \lambda) A(\mathbf{c}, \lambda) \leq 1 \Rightarrow 1 - A(\mathbf{b}, \lambda) A(\mathbf{c}, \lambda) \geq 0 \tag{15.4.20}$$

$$|E(\mathbf{a}, \mathbf{b}) - E(\mathbf{a}, \mathbf{c})| \le \int |A(\mathbf{a}, \lambda)B(\mathbf{c}, \lambda)| \left[1 - A(\mathbf{b}, \lambda)A(\mathbf{c}, \lambda)\right] \rho(\lambda)d\lambda$$

$$= \int \left[1 - A(\mathbf{b}, \lambda)A(\mathbf{c}, \lambda)\right] \rho(\lambda)d\lambda \tag{15.4.21}$$

where $|A(\mathbf{a}, \lambda)B(\mathbf{c}, \lambda)| = 1$ has been used. The above implies that

$$|E(\mathbf{a}, \mathbf{b}) - E(\mathbf{a}, \mathbf{c})| \le 1 - E(\mathbf{b}, \mathbf{c}) \tag{15.4.22}$$

which is the Bell's inequality. A more elaborate CHSH inequality is derived in the Appendix. It can be easily shown that it cannot be satisfied by quantum mechanics that predicts (15.4.9).
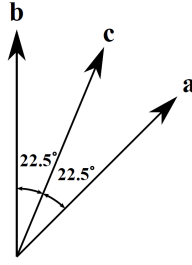

Figure 15.3: The case for $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{c}$ where the quantum mechanics prediction violates the Bell's inequality.

Say if we pick $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{c}$ as shown in the Figure 15.3, then according to quantum mechanics or (15.4.9),

$$E(\mathbf{a}, \mathbf{b}) = \cos(90°) = 0 \tag{15.4.23}$$

$$E(\mathbf{a}, \mathbf{c}) = \cos(45°) = \frac{1}{\sqrt{2}} \tag{15.4.24}$$

$$E(\mathbf{b}, \mathbf{c}) = \cos(45°) = \frac{1}{\sqrt{2}} \tag{15.4.25}$$

Then in (15.4.22), we have

$$\left|0 - \frac{1}{\sqrt{2}}\right| = 0.707 \nleq 1 - \frac{1}{\sqrt{2}} = 0.293 \tag{15.4.26}$$

In experimental tests of Bell's theorem, quantum mechanics triumphs over hidden variable theory so far.[¶] Hence, the spookiness of ghost-angel states will prevail in quantum mechanics.

[¶] In 1982, Alain Aspect (1947-) performed an experiment to confirm this.

15.5   Quantum Cryptography

Quantum cryptography can be used for secure quantum communication. It is secure because a quantum state cannot be replicated without disturbing the quantum state. Hence, whoever wants to replicate a quantum state to steal the data will be easily detected. The property of non-replication follows from the no-cloning theorem.

15.5.1   A Simple Quantum Encryption Algorithm

Let us assume that Alice and Bob communicate by the use of simple photon polarizers. When Alice sends out a $|V\rangle$ state photon, it represents a "1" and similarly, an $|H\rangle$ state photon represents a "0". If Bob receives with a similarly aligned polarizer, he receives the information correctly. (We call this the VH mode.) If Alice aligns her polarizer at 45° and Bob follows suit, he continues to receive the information correctly. In this case, the $+45°$ state represents a "1" and $-45°$ represents a "0". A plausible experimental set up is indicated in Figure 15.1.
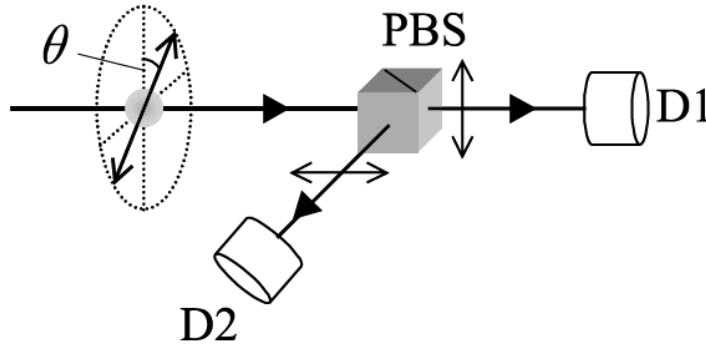

Figure 15.1: Polarization of the incident photon can be detected with a polarizing beam splitter (PBS) and single photon detectors (from Quantum Optics, M. Fox).

However, if Alice is in the VH mode and aligns her polarizer vertically, meaning that Alice will send out a $|V\rangle$ state photon to represent a "1", and Bob is in the 45° mode, and aligns his polarizer at $+45°$, the bit information received by him is correct only 50% of the time. This is because a $|V\rangle$ photon

$$|V\rangle = \frac{1}{\sqrt{2}}(|+45\rangle - |-45\rangle) \tag{15.5.1}$$

which can be written as a quantum linear superposition of two quantum states in the $|+45\rangle$ and $|-45\rangle$ states. Similarly, if Alice transmit an $|H\rangle$ state photon to represent a "0",

$$|H\rangle = \frac{1}{\sqrt{2}}(|+45\rangle + |-45\rangle) \tag{15.5.2}$$

Bob receives the bit of information correctly only 50% of the time.
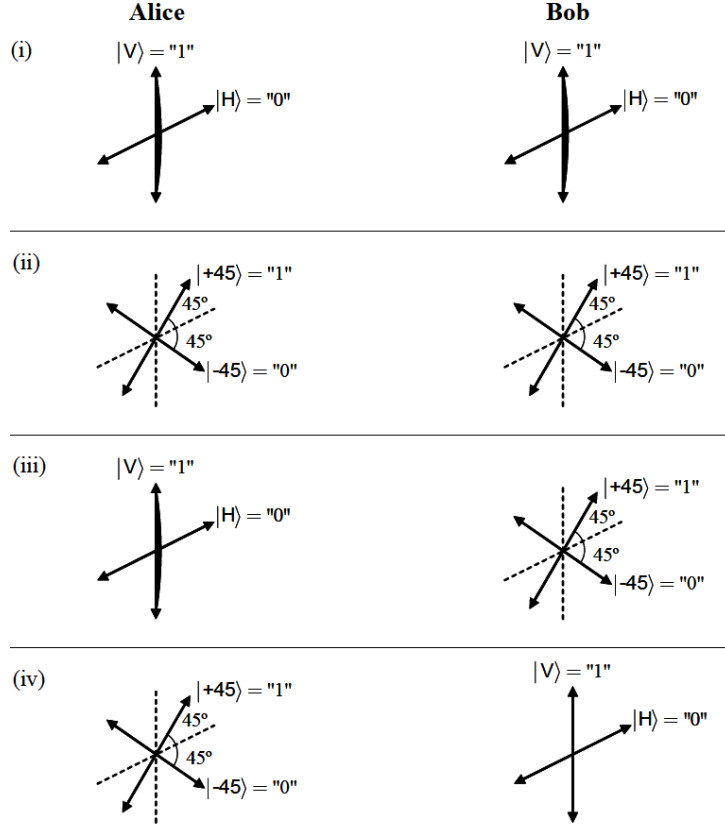
Figure 15.2: Communication between Alice and Bob using single-photon source and simplified polarizer measurement schemes (from DAB Miller).

Similarly, if Alice is in the 45° mode, and transmits with her polarizer aligned in the ±45° angle, and Bob receives with polarizers in the VH mode with H and V polarization, he receives the polarization of $|H\rangle$ and $|V\rangle$ with equal likelihood (50% chance) regardless of what polarization Alice sends. This is because the $|+45\rangle$ state and $|-45\rangle$ state are expressible as a linear superposition of the $|H\rangle$ and $|V\rangle$ states.

$$|+45\rangle = \frac{1}{\sqrt{2}}(|H\rangle + |V\rangle) \tag{15.5.3}$$

$$|-45\rangle = \frac{1}{\sqrt{2}}(|H\rangle - |V\rangle) \tag{15.5.4}$$

By the tenet of quantum measurement, these states are measured with equal likelihood of being $|H\rangle$ or $|V\rangle$.

Now if Alice decides to use her polarizer randomly, and Bob receives with the polarizer randomly so that they are equally likely to use VH mode or 45° mode. The probability that their polarizers are aligned is correctly 50%. During this time, they communicate with no error. The other 50% time, when their polarizers are misaligned, they communicate with 50% error. Hence, 25% of the data are wrong.

After a preliminary quantum communication, Alice and Bob can communicate the information about the alignments of their polarizers, say, by a phone call. Bob will then retain only the data when their polarizers are correctly aligned and discard the rest. For the preliminary communication, they can compare their data over the aligned case, and there should be error free in principle.

If an eavesdropper, Eve, attempts to steal the information, she does so by using a polarizer to intercept the photon. If Eve knows that Alice is using the VH mode, Eve aligns her polarizer in the VH mode. After Eve has received the data, she can retransmit the data to Bob, thus stealing the data. However, if the polarization used by Alice is random, Eve's polarizer is not aligned with Alice's half the time. Eve would have corrupted her data making the wrong transmission 50% of the time. This would increase the error in transmission of the data from Alice to Bob making it wrong 25% of the time. If Alice and Bob communicate by a phone call to check the security of their data transmission and found that it is wrong approximately 25% of the time, they would have suspected an eavesdropper.

Notice that Eve has to collapse the state of the photon sent out by Alice into one of the two states of Eve's polarizer before she can duplicate the photon and send it to Bob. Because of the no-cloning theorem to be discussed next, Eve cannot duplicate the state of the photon that Alice has sent without measuring it.

In a secure communication system, Alice and Bob do not send the real message in preliminary testing of the security of time channel. Alice will first send Bob the secret key and test if the channel is secure. If it is a secure channel, then she would send the rest of the information. The above is known as the BB84 protocol, attributed to Bennett and Brassard's work in 1984.

Also, notice that the above secure communication system does not work in a classical optical communication channel where a bunch of photons is sent. If a bunch of photon is sent by Alice to denote a V or an H polarization to send "1" and "0", when Eve eavesdrops with her misaligned polarizer by 45°, she would have noticed that equal number of photons are emerging from her two orthogonal polarizations. By checking the phase of the two streams of photons, she can easily duplicate a classical photon bunch and send it to Bob, meanwhile stealing the data off the communication channel. Hence, the security of the quantum communication channel comes from the interpretation of quantum mechanics: a particle is in the linear superposition of quantum states before the measurement. A measurement collapses the quantum state into one of the states "discovered" by the measurement.

The above discussion of a secure channel is based on ideal single-photon sources. In practice, non-ideality will give rise to communication errors. Quantum error correction schemes have been devised to minimize the errors in a quantum communication channel.

15.5.2   No-cloning Theorem

First we assume that a quantum operator can be designed such that it can clone a quantum state without measuring it. But such a capability will violate the principle of linear superposition, as we shall show. Therefore, such an operator cannot be designed. This is a proof by contradiction.

A quantum system is described by a quantum state that evolves from an initial state to a final state following the laws of quantum mechanics. An example of such an evolutionary operator is

$$\hat{T} = e^{-i\frac{\hat{H}}{\hbar}t} \tag{15.5.5}$$

The above quantum operator is a unitary operator as well as a linear operator. First, assume that it has the capability of replicating a quantum state in system 2 to be identical to the state in system 1 after acting on such a quantum system. It does so without altering the quantum state of system 1, e.g., by a measurement. We denote this by

$$\hat{T}|\Psi_s\rangle_2|\Psi_a\rangle_1 = |\Psi_a\rangle_2|\Psi_a\rangle_1 \tag{15.5.6}$$

By the same token, it should replicate

$$\hat{T}|\Psi_s\rangle_2|\Psi_b\rangle_1 = |\Psi_b\rangle_2|\Psi_b\rangle_1 \tag{15.5.7}$$

Now if the state to be replicated is

$$|\Psi_c\rangle_1 = \frac{1}{\sqrt{2}}\left[|\Psi_a\rangle_1 + |\Psi_b\rangle_1\right] \tag{15.5.8}$$

Then

$$\hat{T}|\Psi_s\rangle_2|\Psi_c\rangle_1 = \frac{1}{\sqrt{2}}\left[\hat{T}|\Psi_s\rangle_2|\Psi_a\rangle_1 + \hat{T}|\Psi_s\rangle_2|\Psi_b\rangle_1\right]$$

$$= \frac{1}{\sqrt{2}}\left[|\Psi_a\rangle_2|\Psi_a\rangle_1 + |\Psi_b\rangle_2|\Psi_b\rangle_1\right] \neq |\Psi_c\rangle_2|\Psi_c\rangle_1 \tag{15.5.9}$$

Clearly, the above violates the principle of linear superposition if the last equality is true. Hence, such a cloning operator is impossible due to the violation of the principle of linear superposition. The above proves the no-cloning theorem.

15.6   Quantum Computing

The distinguishing feature of quantum computing is quantum parallelism. Again, this follows from the tenet of quantum measurement. A quantum state can be in a linear superposition of many states before the measurement. After the measurement, the quantum state collapses to one of the quantum states. The prowess of quantum computing, as mentioned before, comes from the "ghost-angel" state of a quantum system.

15.6.1 Quantum Bits (Qubits)

A quantum bit or a qubit is a bit in a quantum state that is in the linear superposition of two eigen states representing the $|0\rangle$ bit and the $|1\rangle$ bit. Namely,

$$|\psi\rangle = C_0 |0\rangle + C_1 |1\rangle \tag{15.6.1}$$

where $|C_0|^2 + |C_1|^2 = 1$. The two states $|0\rangle$ and $|1\rangle$ can be the vertical and horizontal polarization of a photon. Alternatively, it can be the up and down state of a spin, or any two energy levels of a multi-level system.

The richness of quantum information is already manifested in this very simple example. Unlike classical bits in classical computers, which can only have binary values, a qubit can have multitudes of possible values depending on the values of C_0 and C_1. A two-level quantum system, as has been shown in the spin case, can be represented by a Bloch sphere. Every point on the Bloch sphere represents a possible quantum state due to the linear superposition of two eigen states, depending on C_0 and C_1. Hence, there could be infinitely many states that follow.

15.6.2 Quantum Gates

Analogous to classical logic gates, there are quantum gates that manipulate the $|0\rangle$ and $|1\rangle$ states of a qubit. A qubit as indicated by (15.6.1) can be represented by a column vector of length two. For example, the qubit in (15.6.1) can be represented by $[C_0, C_1]^t$, which is the vector representation of the quantum state. A quantum gate transforms the quantum state $[C_0, C_1]^t$ to another state $[C_0', C_1']^t$. Such a matrix

$$\begin{bmatrix} C_0' \\ C_1' \end{bmatrix} = \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix} \begin{bmatrix} C_0 \\ C_1 \end{bmatrix} = \overline{\mathbf{M}} \cdot \begin{bmatrix} C_0 \\ C_1 \end{bmatrix} \tag{15.6.2}$$

has to be unitary. The matrix $\overline{\mathbf{M}}$ can be regarded as the two-state representation of the time-evolution operator described more abstractly by

$$\hat{M} = e^{-i\frac{\hat{H}}{\hbar}t} \tag{15.6.3}$$

In the above, the Hilbert space is replaced by that of a two-level system, so by subspace projection, the operators become two-by-two matrix operators. Moreover, they are easily proved to be unitary, a must for an operator constructed from the time-evolution operator.

Examples of quantum gates, expressed in their matrix representations, are

$$\overline{\mathbf{X}} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \tag{15.6.4}$$

$$\overline{\mathbf{Z}} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \tag{15.6.5}$$

$$\overline{\mathbf{H}} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \tag{15.6.6}$$

$$C_0\,|0\rangle + C_1\,|1\rangle \quad\boxed{\hat{X}}\quad C_1\,|0\rangle + C_0\,|1\rangle$$

$$C_0\,|0\rangle + C_1\,|1\rangle \quad\boxed{\hat{Z}}\quad C_0\,|0\rangle - C_1\,|1\rangle$$

$$C_0\,|0\rangle + C_1\,|1\rangle \quad\boxed{\hat{H}}\quad C_0\frac{|0\rangle+|1\rangle}{\sqrt{2}} + C_1\frac{|0\rangle-|1\rangle}{\sqrt{2}}$$

Figure 15.1: Some single qubit gates showing their input and output states.

$\overline{\mathbf{X}}$ represents the NOT gate while $\overline{\mathbf{Z}}$ represents one that flips the sign of $|1\rangle$ bit. $\overline{\mathbf{H}}$ is called the Hadamard gate that is almost like a "square root" gate. The above matrices can be easily shown to be unitary.

When these quantum gates operate on the qubit denoted by (15.6.1), the results are as follows:

$$|\psi\rangle = C_0|0\rangle + C_1|1\rangle \tag{15.6.7}$$

$$\hat{X}|\psi\rangle = C_1|0\rangle + C_0|1\rangle \tag{15.6.8}$$

$$\hat{Z}|\psi\rangle = C_0|0\rangle - C_1|1\rangle \tag{15.6.9}$$

$$\hat{H}|\psi\rangle = \frac{C_0}{\sqrt{2}}\left(|0\rangle + |1\rangle\right) + \frac{C_1}{\sqrt{2}}\left(|0\rangle - |1\rangle\right) \tag{15.6.10}$$

Moreover, one can show that

$$\hat{X}^2 = \hat{Z}^2 = \hat{H}^2 = \hat{I} \tag{15.6.11}$$

When expressed in terms of matrix algebra,

$$\overline{\mathbf{X}}\begin{bmatrix} C_0 \\ C_1 \end{bmatrix} = \begin{bmatrix} C_1 \\ C_0 \end{bmatrix} \tag{15.6.12}$$

$$\overline{\mathbf{Z}}\begin{bmatrix} C_0 \\ C_1 \end{bmatrix} = \begin{bmatrix} C_0 \\ -C_1 \end{bmatrix} \tag{15.6.13}$$

$$\overline{\mathbf{H}}\begin{bmatrix} C_0 \\ C_1 \end{bmatrix} = \begin{bmatrix} \frac{C_0+C_1}{\sqrt{2}} \\ \frac{C_0-C_1}{\sqrt{2}} \end{bmatrix} \tag{15.6.14}$$

In addition to one qubit gate, there are also two qubit gates. In this case, there are two input bits for these gates. They can be arranged as $|0,0\rangle$, $|0,1\rangle$, $|1,0\rangle$, and $|1,1\rangle$ as the four input and output states. A very important one is the CNOT gate shown in Figure 15.2. The upper line represents the first bit which is also the control bit, while the lower line represents the second bit, whose output is the target bit. The $\oplus$ symbol represents addition modulus 2.

Hence, its transformation matrix is given by

$$\overline{\mathbf{U}}_{\text{CNOT}} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} \overline{\mathbf{I}} & \overline{\mathbf{0}} \\ \overline{\mathbf{0}} & \overline{\mathbf{Z}} \end{bmatrix} \tag{15.6.15}$$

The above can also be implemented with a unitary transform.

$$|A\rangle \quad\bullet\quad |A\rangle$$
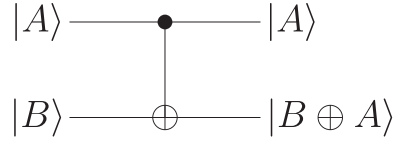$$|B\rangle \quad\oplus\quad |B \oplus A\rangle$$

Figure 15.2: A quantum circuit of a two-qubit gate representing the CNOT gate. It takes two input streams, and has two output streams. The $\oplus$ symbol represents an *exclusive or* operation, or addition modulus 2 (redrawn from Quantum Optics, M. Fox).

15.6.3 Quantum Computing Algorithms

As mentioned before, the most important aspect of quantum computing algorithm is quantum parallelism. We will illustrate this with the simple Deutsch algorithm. This algorithm assumes a two-qubit system, so that it has a two-qubit input and two-qubit output. It can be implemented with the quantum circuit shown below:

Input   Output

$$|q_1\rangle = |0\rangle \quad \hat{H} \quad |x\rangle \qquad |x\rangle \quad \hat{H}$$
$$\hat{u}_f$$
$$|q_2\rangle = |1\rangle \quad \hat{H} \quad |y\rangle \qquad |y \oplus f(x)\rangle$$

$$|\psi_0\rangle \qquad |\psi_1\rangle \qquad\qquad |\psi_2\rangle \qquad |\psi_3\rangle$$
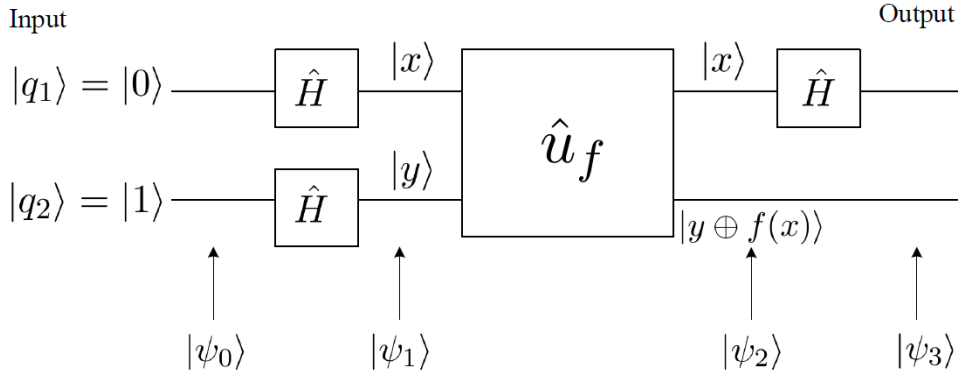
Figure 15.3: The implementation of the Deutsch algorithm with quantum gates and circuits (redrawn from Quantum Optics, M. Fox).

We start with a state representing the two qubits in the input as

$$|\psi_0\rangle = |0\rangle|1\rangle = |0, 1\rangle \tag{15.6.16}$$

At stage $|\psi_1\rangle$, we have

$$|x\rangle = \hat{H}|q_1\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \tag{15.6.17}$$

$$|y\rangle = \hat{H}|q_2\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) \tag{15.6.18}$$

Hence, the quantum state representing the two-qubit system at this stage is

$$|\psi_1\rangle = |x\rangle|y\rangle = \frac{1}{2}(|0\rangle + |1\rangle)(|0\rangle - |1\rangle) = \frac{1}{2}(|0,0\rangle - |0,1\rangle + |1,0\rangle - |1,1\rangle) \tag{15.6.19}$$

The unitary operator $\hat{u}_f$ has no effect on the x qubit, but performs the operation $y \oplus f(x)$ on the y qubit. The function $f(x)$ takes input x which is either 0 or 1. It produces an output which is either constant or balanced, but the output is either 0 or 1.

Hence, according to the rules above, the quantum state $|\psi_2\rangle$ of the two qubit system is

$$|\psi_2\rangle = |x\rangle|y \oplus f(x)\rangle \tag{15.6.20}$$

Substituting the state $|x\rangle$ and $|y\rangle$ from (15.6.17) and (15.6.18) gives rise to

$$|\psi_2\rangle = |x\rangle|y \oplus f(x)\rangle = \frac{1}{2}(|0\rangle + |1\rangle)(|0 \oplus f(x)\rangle - |1 \oplus f(x)\rangle) \tag{15.6.21}$$

Multiplying out the above, with the rule that $f(x)$ always takes its argument from the $|x\rangle$ bit, then

$$|\psi_2\rangle = \frac{1}{2}(|0, f(0)\rangle - |0, 1 \oplus f(0)\rangle + |1, f(1)\rangle - |1, 1 \oplus f(1)\rangle) \tag{15.6.22}$$

with the further rule that $0 \oplus f(x) = f(x)$ has been used.

If $f(x)$ is a constant function, then $f(0) = f(1)$, and the above becomes

$$\begin{aligned}
|\psi_2\rangle_{\text{const}} &= \frac{1}{2}(|0, f(0)\rangle - |0, 1 \oplus f(0)\rangle + |1, f(0)\rangle - |1, 1 \oplus f(0)\rangle) \\
&= \frac{1}{2}(|0\rangle + |1\rangle)(|f(0)\rangle - |1 \oplus f(0)\rangle)
\end{aligned} \tag{15.6.23}$$

With the Hadamard operation on the upper qubit, we have

$$|\psi_3\rangle_{\text{const}} = |0\rangle \frac{1}{\sqrt{2}}(|f(0)\rangle - |1 \oplus f(0)\rangle) \tag{15.6.24}$$

If for the balanced case, $f(0) \neq f(1)$, hence $f(1) = 1 \oplus f(0)$. Then

$$\begin{aligned}
|\psi_2\rangle_{\text{bal}} &= \frac{1}{2}(|0, f(0)\rangle - |0, 1 \oplus f(0)\rangle + |1, 1 \oplus f(0)\rangle - |1, f(0)\rangle) \\
&= \frac{1}{2}(|0\rangle - |1\rangle)(|f(0)\rangle - |1 \oplus f(0)\rangle)
\end{aligned} \tag{15.6.25}$$

After the H gate operation, we have

$$|\psi_3\rangle_{\text{bal}} = |1\rangle \frac{1}{\sqrt{2}} \left(|f(0)\rangle - |1 \oplus f(0)\rangle \right) \tag{15.6.26}$$

The above shows the prowess of quantum parallelism with just one operation, one can determine if a function $f(x)$ is balanced or constant by observing the $|q_1\rangle$ value of the output qubits. A constant coin can represent a fake coin, while a balanced coin represents a real coin. With one operation, one can determine if the coin is real or fake without having to check both sides of the coin.

More sophisticated algorithms exploiting quantum parallelism, such as the Shor's algorithm and the Grover's algorithm, have been devised. The Shor's algorithm can perform a Fourier transform in $(\log N)^2$ operations rather than the classical $N \log N$ operations. The Grover's algorithm can search a data base with N data in $\sqrt{N}$ operations rather than the classical N operations. Because of quantum parallelism, quantum computer can also perform quantum simulation of quantum system which is not possible on classical computers.

15.7 Quantum Teleportation

Quantum teleportation is the idea of Alice being able to send a photon of unknown state to Bob, without having to perform a measurement on this photon, nor disturb its state. We denote the photon, called photon 1, in the unknown state by

$$|\psi\rangle_1 = C_0 |0\rangle_1 + C_1 |1\rangle_1 \tag{15.7.1}$$

In the beginning, this photon is only accessible to Alice. Alice, however, is accessible to another photon, called photon 2, of an entangled Bell state. The other photon of the Bell state, photon 3, is not accessible to Alice but is accessible to Bob.

We can define the state of the three photons by the direct product state

$$|\Psi\rangle_{123} = \frac{1}{\sqrt{2}} \left(C_0 |0\rangle_1 + C_1 |1\rangle_1 \right) \left(|0\rangle_2 |1\rangle_3 - |1\rangle_2 |0\rangle_3 \right) \tag{15.7.2}$$

where the Bell state is assumed to be

$$\left| \Psi^- \right\rangle_{23} = \frac{1}{\sqrt{2}} \left(|0\rangle_2 |1\rangle_3 - |1\rangle_2 |0\rangle_3 \right) \tag{15.7.3}$$

Expanding (15.7.2) gives rise to

$$|\Psi\rangle_{123} = \frac{1}{\sqrt{2}} \left(C_0 |0\rangle_1 |0\rangle_2 |1\rangle_3 - C_0 |0\rangle_1 |1\rangle_2 |0\rangle_3 \right.$$
$$\left. + C_1 |1\rangle_1 |0\rangle_2 |1\rangle_3 - C_1 |1\rangle_1 |1\rangle_2 |0\rangle_3 \right) \tag{15.7.4}$$

The four Bell states are complete and orthogonal, and the states of photon 1 and photon 2 can be expanded in the four Bell states; namely

$$\left| \Phi^+ \right\rangle_{12} = \frac{1}{\sqrt{2}} \left(|0\rangle_1 |0\rangle_2 + |1\rangle_1 |1\rangle_2 \right) \tag{15.7.5}$$
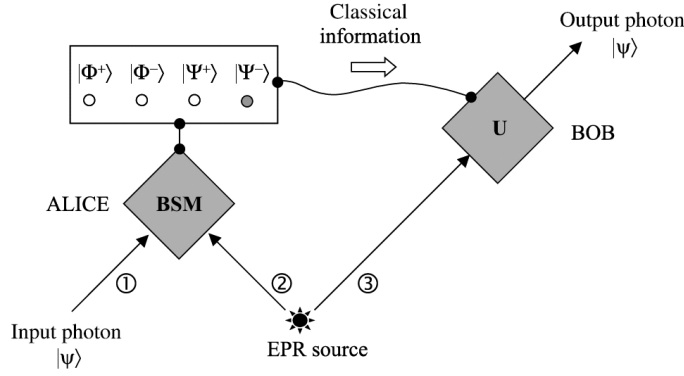
Figure 15.1: A teleportation experiment setup with photons. Alice makes measurement on the two photons accessible to her using the Bell state measurement device, and communicate the outcome to Bob via a classical channel. Bob then performs a unitary transformation on his photon to obtain the input photon state. (From M. Fox, Quantum Optics.)

$$\left|\Phi^{-}\right\rangle_{12} = \frac{1}{\sqrt{2}}\left(\left|0\right\rangle_1 \left|0\right\rangle_2 - \left|1\right\rangle_1 \left|1\right\rangle_2\right) \tag{15.7.6}$$

$$\left|\Psi^{+}\right\rangle_{12} = \frac{1}{\sqrt{2}}\left(\left|0\right\rangle_1 \left|1\right\rangle_2 + \left|1\right\rangle_1 \left|0\right\rangle_2\right) \tag{15.7.7}$$

$$\left|\Psi^{-}\right\rangle_{12} = \frac{1}{\sqrt{2}}\left(\left|0\right\rangle_1 \left|1\right\rangle_2 - \left|1\right\rangle_1 \left|0\right\rangle_2\right) \tag{15.7.8}$$

Projecting (15.7.4) onto the four Bell states, and subsequently expanding (15.7.4) in terms of them, we have

$$\begin{aligned}
\left|\Psi\right\rangle_{123} = \frac{1}{2}\big[&\left|\Phi^{+}\right\rangle_{12}\left(C_0 \left|1\right\rangle_3 - C_1 \left|0\right\rangle_3\right) \\
&+ \left|\Phi^{-}\right\rangle_{12}\left(C_0 \left|1\right\rangle_3 + C_1 \left|0\right\rangle_3\right) \\
&+ \left|\Psi^{+}\right\rangle_{12}\left(-C_0 \left|0\right\rangle_3 + C_1 \left|1\right\rangle_3\right) \\
&- \left|\Psi^{-}\right\rangle_{12}\left(C_0 \left|0\right\rangle_3 + C_1 \left|1\right\rangle_3\right)\big]
\end{aligned} \tag{15.7.9}$$

In the above, the first two photons, photon 1 and photon 2, are grouped into different Bell states. Moreover, the state of photon 3 resembles the state of the original photon 1. The quantum system now is in a linear superposition of different Bell states.

The Bell state measurement device projects the first two photons onto a Bell state. The Bell state measurement device collapses the quantum system into one of the four Bell states. For example, when Alice finds that the first two photons are in the Bell state $\left|\Phi^{+}\right\rangle_{12}$, then the third photon must be in the state

$$\left|\psi\right\rangle_3 = C_0 \left|1\right\rangle_3 - C_1 \left|0\right\rangle_3 \tag{15.7.10}$$

Bob can apply a unitary operator, similar to qubit gate operators described in the previous section, to obtain the original state of photon 1.

The above does not violate the no-cloning theorem, because the original state of the photon 1 is destroyed, and its semblance is reproduced in photon 3.

15.8   A Final Word on Quantum Parallelism

Quantum interpretation gives rise to the spookiness of quantum mechanics, in a way giving it the capability that empowers ghosts and angels. Quantum parallelism is such an empowerment. It reminds me of a novel that I have read when I was young on the Monkey King. This Monkey King is of Indian origin, but has permeated Chinese culture to take on a different persona. He is known as Hanuman in India. According to the story in The Journey to the West, this Monkey King had unusual capabilities. One of them was that he could pull a strand of hair from his body, and with a puff of air, he could turn it into many duplicates of himself. He could then fight his enemies from all angles and all sides. He thus made himself invincible.

He was arrogant, mischievous, and wreaked havoc in the Heavenly Palace where the other gods lived. Finally, he could only be tamed and subdued by Lord Buddha, summoning him to accompany and protect the monk Xuan Zang in his treacherous journey to collect the Buddhist Sutra from the West (in this case India).

Oh Lord, if we were to empower ourselves with the capabilities of ghosts and angels, who is there to curb our power!

Appendix

A   Derivation of the Clauser-Horne-Shimon-Holt Inequality

In addition to the Bell's inequality, a more elaborate Clauser-Horne-Shimony-Holt (CHSH) inequality can be derived. To this end, we can show that

$$E(\mathbf{a}, \mathbf{b}) - E(\mathbf{a}, \mathbf{b}') = \int [A(\mathbf{a}, \lambda)B(\mathbf{b}, \lambda) - A(\mathbf{a}, \lambda)B(\mathbf{b}', \lambda)]\rho(\lambda)d\lambda \tag{A1}$$

The above can be rewritten as

$$E(\mathbf{a}, \mathbf{b}) - E(\mathbf{a}, \mathbf{b}') = \int A(\mathbf{a}, \lambda)B(\mathbf{b}, \lambda)[1 \pm A(\mathbf{a}', \lambda)B(\mathbf{b}', \lambda)]\rho(\lambda)d\lambda$$

$$- \int A(\mathbf{a}, \lambda)B(\mathbf{b}', \lambda)[1 \pm A(\mathbf{a}', \lambda)B(\mathbf{b}, \lambda)]\rho(\lambda)d\lambda \tag{A2}$$

We have just added and subtracted identical terms in the above. After using the triangular inequality

$$|E(\mathbf{a}, \mathbf{b}) - E(\mathbf{a}, \mathbf{b}')| \leq \int |A(\mathbf{a}, \lambda)B(\mathbf{b}, \lambda)||1 \pm A(\mathbf{a}', \lambda)B(\mathbf{b}', \lambda)|\rho(\lambda)d\lambda$$
$$+ \int |A(\mathbf{a}, \lambda)B(\mathbf{b}', \lambda)||1 \pm A(\mathbf{a}', \lambda)B(\mathbf{b}, \lambda)|\rho(\lambda)d\lambda \qquad \text{(A3)}$$

Using the fact that

$$|AB| = 1, \qquad 1 \pm AB \geq 0 \qquad \text{(A4)}$$

we have

$$|E(\mathbf{a}, \mathbf{b}) - E(\mathbf{a}, \mathbf{b}')| \leq \int [1 \pm A(\mathbf{a}', \lambda)B(\mathbf{b}', \lambda)]\rho(\lambda)d\lambda$$
$$+ \int [1 \pm A(\mathbf{a}', \lambda)B(\mathbf{b}, \lambda)]\rho(\lambda)d\lambda \qquad \text{(A5)}$$

The above is the same as

$$|E(\mathbf{a}, \mathbf{b}) - E(\mathbf{a}, \mathbf{b}')| \leq 2 \pm [E(\mathbf{a}', \mathbf{b}') + E(\mathbf{a}', \mathbf{b})] \qquad \text{(A6)}$$

It is of the form

$$|X| \leq 2 \pm Y \qquad \text{(A7)}$$

which implies that

$$|X| \pm Y \leq 2 \qquad \text{(A8)}$$

or

$$|X + Y| \leq |X| + |Y| \leq 2 \qquad \text{(A9)}$$

Consequently, we have

$$|E(\mathbf{a}, \mathbf{b}) - E(\mathbf{a}, \mathbf{b}') + E(\mathbf{a}', \mathbf{b}') + E(\mathbf{a}', \mathbf{b})| \leq 2 \qquad \text{(A10)}$$

which is the Clauser-Horne-Shimony-Holt (CHSH) inequality. In (A6), if $\mathbf{a}' = \mathbf{b}' = \mathbf{c}$, then (A6) becomes

$$|E(\mathbf{a}, \mathbf{b}) - E(\mathbf{a}, \mathbf{c})| \leq 2 \pm [E(\mathbf{c}, \mathbf{c}) + E(\mathbf{b}, \mathbf{c})] \qquad \text{(A11)}$$

From (15.4.14), $E(\mathbf{c}, \mathbf{c}) = 1$. We pick the smaller of the right-hand side of (A11) and arrive at

$$|E(\mathbf{a}, \mathbf{b}) - E(\mathbf{a}, \mathbf{c})| \leq 1 - E(\mathbf{b}, \mathbf{c}) \qquad \text{(A12)}$$

The above is the previously derived Bell's inequality.

Figure 15.1: The Monkey King sending replicas of himself against his enemy, defeating everyone in his path.