PH2201

INTRODUCTION TO QUANTUM MECHANICS

LECTURE NOTES

Instructor : Dr. Siddhartha Lal

Email : slal@iiserkol.ac.in

Notes prepared by : Mounica Mahankali, Abhirup Mukherjee and Som Dev Bishoyi

SPRING SEMESTER, 2020

INDIAN INSTITUTE OF SCIENCE EDUCATION & RESEARCH, KOLKATA

IISER KOLKATA

Contents

Foreword

These are notes of lectures I have been giving at the Department of Physical Sciences (DPS), IISER-Kolkata for the course on introductory Quantum Mechanics (PH2201) over the years. I am very thankful to Ms. Mounica Mahankali, Mr. Abhirup Mukherjee and Mr. Som Dev Bishoyi for kindly typing up the first version of the lectures notes (along with drawing and including several figures) so meticulously. Any errors, typographical or otherwise, can be laid at my door. Please do email me with corrections.

The intention of the lectures is to provide a (hopefully gentle!) introduction to the theoretical machinery of quantum mechanics. Even as attention has given towards covering the phenomenology of quantum mechanics, an effort has also been made to provide adequately the mathematical underpinnings of the subject. Given their venerable status, almost all the topics covered within these areas are well-known and covered in detail in several places. I have tried to bring things together in my own way. The material has been separated into first covering the basics and leads to a presentation of some advanced topics. The classification of topics into basic and advanced is certainly subjective, and I trust that upon strengthening her/his fundamentals, the reader will find passage through some of the advanced topics easier as well as informative.

The topics covered are in keeping with the syllabus that has evolved at IISER Kolkata over the years. Certainly, there are many intereting topics that I have not been able to cover in the lectures, primarily as they are the ambit of more advanced courses in Quantum Mechanics that are taught at IISER Kolkata. I do not expect that advanced students and researchers in theoretical physics will find the lectures particularly enlightening. However, I hope that my lecture notes will prove to be a good companion to the several excellent textbooks in Quantum Mechanics (of which there are more than stars in the night sky!) for study by undergraduates (and perhaps even early graduate students) keen to strengthen their foundations by learning some of the material I have covered. In this sense, I hope that these lecture notes can act as a springboard on which students can launch deeper explorations of the fascinating world of Quantum Mechanics.

A word on prerequisites. Even though these lectures are introductory in nature, the reader will need a sound grasp of the basics of vectors, matrices, determinants and ordinary differential equations in making quick progress through the material covered. For readers unfamiliar with these prerequisites, there are several excellent textbooks that impart the basics of these topics in mathematical physics. I would like to acknowledge many discussions with my graduate students Santanu Pal, Anirban Mukherjee and Siddhartha Patra, as well as many colleagues and former students at IISER-Kolkata whose insight into the subject I have benefited from greatly. Thanks are also due to the students who attended the lectures, and offered valuable feedback. Unless mentioned specifically, no claims of originality are being made in either the content or presentation of material covered in these lectures. Indeed, they were prepared from a thorough consultation of a selection of excellent textbooks already available on the topics I covered, including:

- Robert Scherrer, “Quantum Mechanics: an accessible introduction”, Pearson Addison-Wesley, 2006
- D. J. Griffiths, “Introduction to Quantum Mechanics”, 2nd Edition, Pearson Prentice Hall, 2005
- R. Shankar, “Principles of Quantum Mechanics”, 2nd Edition, Springer, 1994
- E. Merzbacher, “Quantum Mechanics”, 3rd Edition, Wiley, 2004
- R. P. Feynman, R. B. Leighton and M. Sands, “The Feynman Lectures on Physics Vol. III”, Addison-Wesley, 1965
- F. Schwabl, “Quantum Mechanics”, 4th Edition, Springer, 2007
- J.J. Sakurai, “Modern Quantum Mechanics”, Revised Edition, Addison-Wesley, 1999
- L. E. Ballentine, “Quantum Mechanics: a modern development”, 2nd Edition, World Scientific, 2015

Several pictures have been taken from various sources on the internet, and I am deeply grateful to colleagues from all over the world for such excellent figures! I offer my sincere apologies for not citing them specifically in every instance.

Finally, I invite you to begin your journey into these lectures with a quote from a favourite fictional character of mine:

To ∞ and beyond!

Siddhartha Lal
Mohanpur, West Bengal, India
January, 2020

To ∞ and Beyond! (Source: The internet.)

Chapter 1

Introduction

Through the course of our journey through these lectures, we are going to depart from our familiar classical world and learn about the fascinating world of quantum phenomena. This is not an easy passage, as the world around us does not easily show us quantum phenomena. Our intuition is, instead, built from our observations of the "classical world" around us. I will, therefore, try to introduce you here to some surprising differences between the classical and quantum worlds.

1.1 Getting started

Features of the Classical World
1. Our everyday observations tell us that the physical universe is deterministic.
By this we mean that with enough information about a given system, we can predict its evolution in time precisely!

2. Light consists of waves, while matter consists of particles.
Our understanding of Maxwell's theory for classical electromagnetism, together with the field of classical optics, confirms that light is made of waves. On the other hand, the view of the ancient Hindu and Greek philosophers that all matter is constructed from some indivisible units — the atomistic view of matter — is confirmed by experiments starting with the Brownian motion of particles. All of chemistry is contingent on this view.

3. Physical quantities are continuous variables.
Consider examples of quantities you have measured in the laboratory thus far: energy, linear momentum, angular momentum, position etc.

4. There exists an objective reality independent of any observer.
In making an observation of a classical system, we never affect the system in the process of making the measurement.

NONE OF THESE ARE COMPLETELY ACCURATE IN THE WORLD OF

Features of the Quantum World

1. The physical universe is not determistic!

(a) In the quantum world, the various outcomes of any process have probabilities associated with them: we cannot predict the exact result with certainty.

(b) Uncertainty is an intrinsic property of the quantum world, and has nothing to do with our ability to make accurate observations of the world around us.

1.2 The Schrödinger's Cat gedanken

For an illustration of both these (seemingly mysterious) concepts, consider the famous example of the Schrödinger's Cat *gedanken* (or "thought experiment" in the German language).

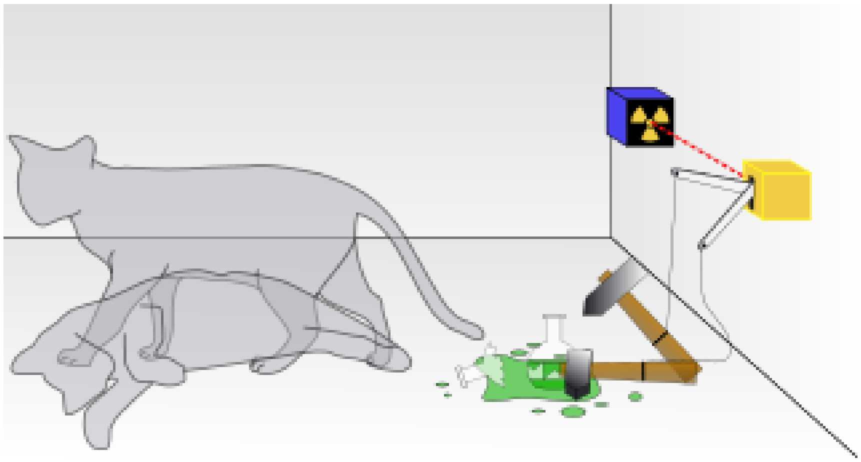


Figure 1.1: A schematic diagram of the Schrödinger's Cat thought experiment. See discussion in text. Disclaimer: No such experiment has ever been carried out with a real cat! Source: The internet.

The gedanken is straight-forward. A cat is placed in an opaque box together with a radioactive source, a Geiger counter that can detect the radioactivity and a vial of poison gas. Any radioactive source has a well-defined lifetime, i.e., a definite fraction of the population of the atoms will have undergone radioactive decay within the lifetime. The decay process is, however, probabilistic; this means that at any given time, we can associate a probability that a given atom will have undergone radioactive decay. Now, if during the time the cat is kept in the box, the radioactive source undergoes decay, the Geiger counter will sense this and break the vial of poison gas, killing the cat. Thus, upon opening the lid of the box, we will find the cat to be dead. On the other hand, if the source does not decay, the cat will be alive upon observation.

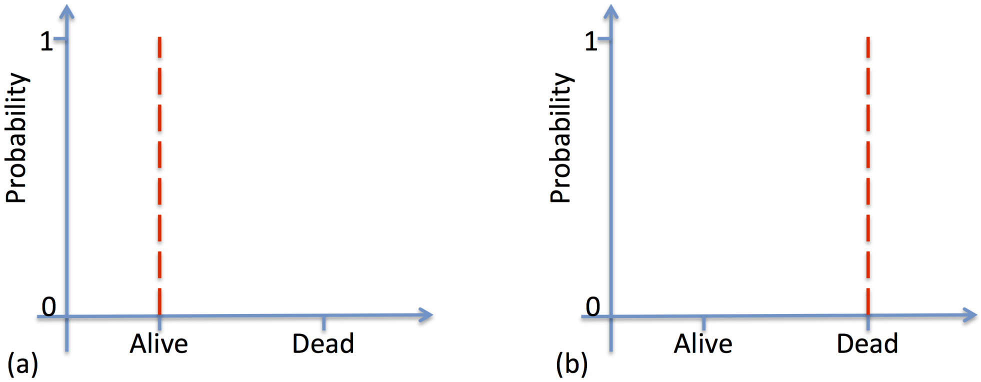


Figure 1.2: A schematic diagram of the probability for observations of the “Dead” and “Alive” states in the Schrödinger’s Cat thought experiment by continuous classical observations. See discussion in text.

Now this is where things become strange. While in the classical world, the cat would have been either Dead or Alive with perfect precision, it exists in a strange admixture of both possibilities within the quantum world (i.e., with the box closed!). For instance, if we carry out the gedanken with a large number (say, a million) of identical set-ups, each of which is being continuously observed by a classical observer for a time T (i.e., with the lid of the box open) and the distribution of all million outcomes then plotted in a histogram, we should either find a histogram of the kind shown in Fig.1.2(a) or Fig.1.2(b) depending on whether the class is alive or dead at time T. This is completely consistent with our knowledge of the classical states of the cat.

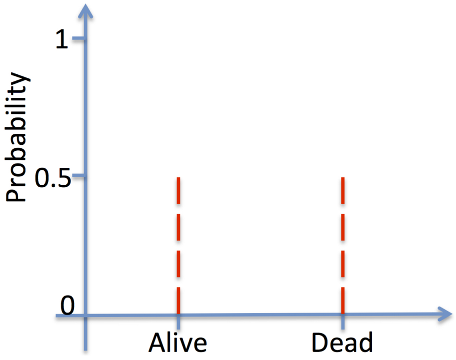


Figure 1.3: A schematic diagram of the probability for observations of the “Dead” and “Alive” states in the Schrödinger’s Cat thought experiment by a classical observer who only observes the state of the cat after a certain time interval T. See discussion in text.

However, as mentioned above, the cat can exist in a linear superposition of the two classical probabilities in the quantum world. For instance, for an equal probability admixture of the classical states of "Alive" and "Dead", we can think of the "Quantum Mechanical State" of the cat within the closed box as being:

$$|\psi_{\text{Quantum Cat}}\rangle = \frac{1}{\sqrt{2}}(|\text{Alive}\rangle + |\text{Dead}\rangle) \ , \tag{1.1}$$

$$P_{\text{Alive}} = (\frac{1}{\sqrt{2}})^2 = \frac{1}{2} = P_{\text{Dead}} \ , \tag{1.2}$$

where the Ps are the probabilities for the classical states of "Alive" and "Dead" (see Fig.1.3). Such an admixture tells us that (i) the cat is dead or alive with well-defined probabilities. As with the gedanken protocol set out above, these probabilities will be revealed if we were to carry out similar measurements on a large number of identical setups of the gedanken, and only if the observations are taken after a time interval T has elapsed with the lid of the box closed (allowing for the quantum state of the cat to be realised). Further, the uncertainty in our knowledge of the state of the cat (i.e., whether it is alive or dead) arises from its state being an admixture of the two, and has nothing to do with our ability to observe it once we open the box. Thus, the passage from the quantum to the classical world can be charted when the probability for a certain outcome becomes perfect (i.e., 1).

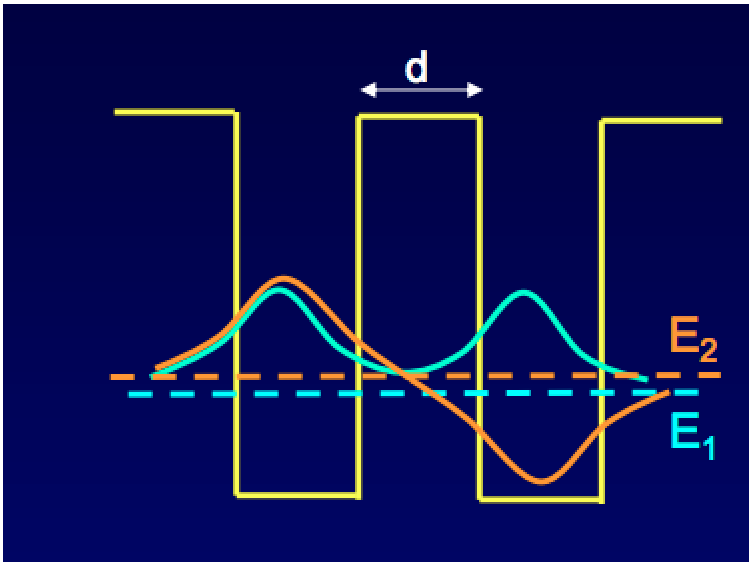


Figure 1.4: The wavefunctions for the symmetric and anti-symmetric states of a quantum particle placed in a double well potential. The symmetric state is lower in energy than the anti-symmetric one due to quantum tunneling. Source: The internet.

A disclaimer: this gedanken has never been implemented in the laboratory with cats (and hopefully, never will be!). Further, as my 11 year old daughter pointed out, a real cat would have died of suffocation (or even boredom!) first. (Yes, we physicists often miss the obvious truths!) However, versions of this experiment with atoms in a superposition of two classical states have been conducted, e.g., quantum particle in a double well potential.

2. $\boxed{\text{Both light and matter displays properties of waves as well as particles!}}$

This is called “wave-particle duality”. For instance, the Schrödinger’s Cat gedanken gives a “wave-like” linear superposition of possible classical states to the quantum state of a creature clearly composed of particles. It is important to note, however, that any given experiment shows either a wave-like property or a particle-like property, but never both simultaneously. Nevertheless, it is amazing to consider the fact that different experiments show very different features of the same system!

3. $\boxed{\text{Physical quantities can have discrete values, i.e., they may be quantised.}}$
The word “quanta” comes from the ancient Greek word for a discrete unit. The phenomenon of quantisation is at the heart of all quantum phenomena: it helps us understand blackbody radiation, the photoelectric effect, the stability of the atom and so much more!

4. $\boxed{\text{A System and its Observer are entangled with one another in the quantum world:}}$
the observer affects the system through the process of measurement. Words like “entanglement” and “measurement” need careful explanation, and I will dwell on this in more detail in one of the final chapters.

1.3 A historical background

If all of this has left you confused, worried and/or excited, then thats fine. Everything that I mentioned above is not only counter-intuitive, it is plain outrageous! Then why believe it? Because it works! And it has taken us around 120 years to come to terms with the intrinsic “strangeness” of quantum mechanics.

Indeed, around 1900, most physicists were convinced that almost everything could be understood using the conceptual pillars of classical mechanics, electromagnetism and thermodynamics. Very few experimental puzzles remained to be understood, e.g., Blackbody radiation, the Michelson-Morley experiment, the photoelectric effect, Brownian motion and diffusion etc. It is quite sobering to consider that it was these same experiments that went on to shake the very foundations of physics, and gave birth to the fields of quantum mechanics and statistical mechanics.

Quantum mechanics took roughly three decades to be formalised, with a lot of the action in the 1920s. Some of the greats involved in this effort are shown below from the famous Solvay Conference in 1927. Please note that among all the Nobel prize-winning luminaries in that photograph, none is more distinguished than Madame Curie (who happens to be the only double Nobel prize winner among them all!). While the epicentre of the action was in Europe (and mostly Germany and France), people from all over the world contributed to this effort. Noteworthy contributors from India included Prof. Satyendranath Bose and Prof. C. V. Raman. While the latter won the Nobel prize for his contribution (the “Raman effect”), the former richly deserved one (but was never awarded it) for the particles (“bosons”) that bear his name. Some small comfort can perhaps be taken from the fact that since all “force” particles are classified as bosons, his legacy extends far beyond the narrow purview of any award.

Importantly, Quantum mechanics led to the discovery of new fields of study in atomic,

Figure 1.5: A photograph of the participants at the famous Solvay Conference in 1927. Many of the founding greats of quantum mechanics were present. Source: The internet.

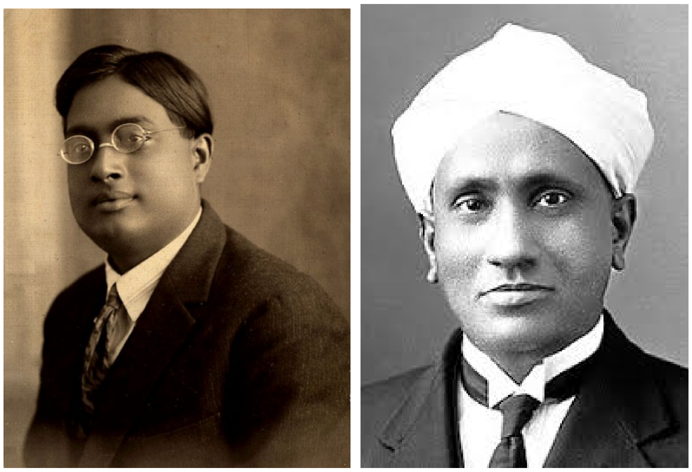


Figure 1.6: Left: Prof. Satyendranath Bose. Right: Prof. C. V. Raman Source: The internet.

nuclear, particle and condensed matter physics! This journey has revolutionised not only our view of the universe around us, but also led to a surge in technological progress over the past century. In a photo below, representative examples of our progress involve the creation of a quantum computer by Google in 2019, and the experimental finding of a single-sheet of Carbon atoms, Graphene, in 2005 by Novoselov and Geim. Also, marvel at the picture given below of a superconducting sample "floating" above a magnet due to the Meissner-Ochsenfeld effect (and the sumo wrestler standing on such a levitating superconducting slab in the accompanying picture) at very low temperatures. Superconductivity is an excellent example of quantum mechanics at the macroscale. A challenge would be to obtain a superconductor at room temperatures! Even more importantly, physics is at its heart an empirical science:

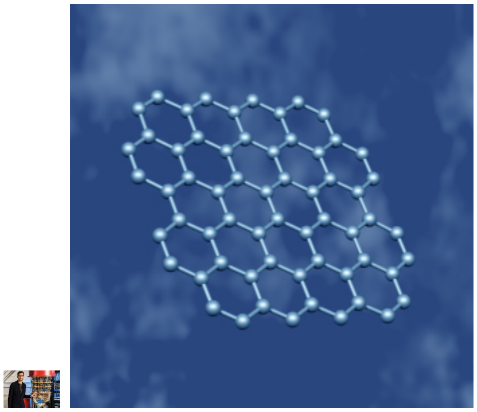


Figure 1.7: Left: Google CEO Sundar Pichai standing beside the quantum computer built in 2019 by Google. Right: Graphene, a single sheet of Carbon atoms, is a wonder material. Source: The internet.

Figure 1.8: (Left) The Meissner-Ochsenfeld effect for the expulsion of magnetic flux from a superconductor, observed spectacularly in the form of a block of a superconductor floating above a permanent magnet. (Right) A Sumo wrestler on a levitating superconductor.

experiments often lead the way in uncovering new phenomena and theories are then built to understand them. So it was with the birth of quantum mechanics. Thus, in beginning our journey into the world of quantum mechanics, it is well worth exploring the experimental puzzles that heralded the quantum revolution, and understanding their resolution.

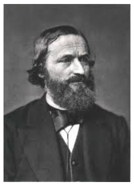
Chapter 2

The Origins of Quantum Theory

2.1 Solving the Blackbody Radiation Spectrum Puzzle

2.1.1 Blackbody Radiation

Though it does not provide the most compelling evidence for quantum mechanics, it is of historical importance. The perfectly "black body" absorbs all radiation incident on it and reflects nothing. There is more: the blackbody is special in that it emits radiation of all frequencies.



Kirchoff

Kirchoff's law: Conservation of heat current in a system which is in thermal equilibrium with its surroundings ⇒ $\boxed{\text{the rate of heat absorption} = \text{the rate of emission}}$.

Thermal equilibrium ⇒ $\boxed{\text{Temperature of System} = \text{Temperature of surroundings}}$.

This law says nothing about the frequency (ν) of the radiation absorbed and emitted, i.e., it is independent of ν. Without this concept, a body would spontaneously heat up or cool down! A practical example of a blackbody is a cavity with a tiny aperture. Light inside the cavity is absorbed, reflected & emitted by its blackened walls and spends a very long time within the cavity before escaping. In this way, the cavity is filled with blackbody radiation, and the aperture acts like the blackbody absorber and emitter.

Stefan

Boltzmann

The total power emitted by a black body follows the Stefan-Boltzmann law

$$P = \sigma A T^4 \ , \tag{2.1}$$

where $\sigma = 5.67 \times 10^{-8}\mathrm{Js^{-1}m^{-2}K^{-4}}$ is the Stefan-Boltzmann's constant, A is the surface area of the blackbody and T is its temperature. As we will soon see, Boltzmann showed how the T^4 could be obtained from thermodynamic considerations. Total energy density of radiation

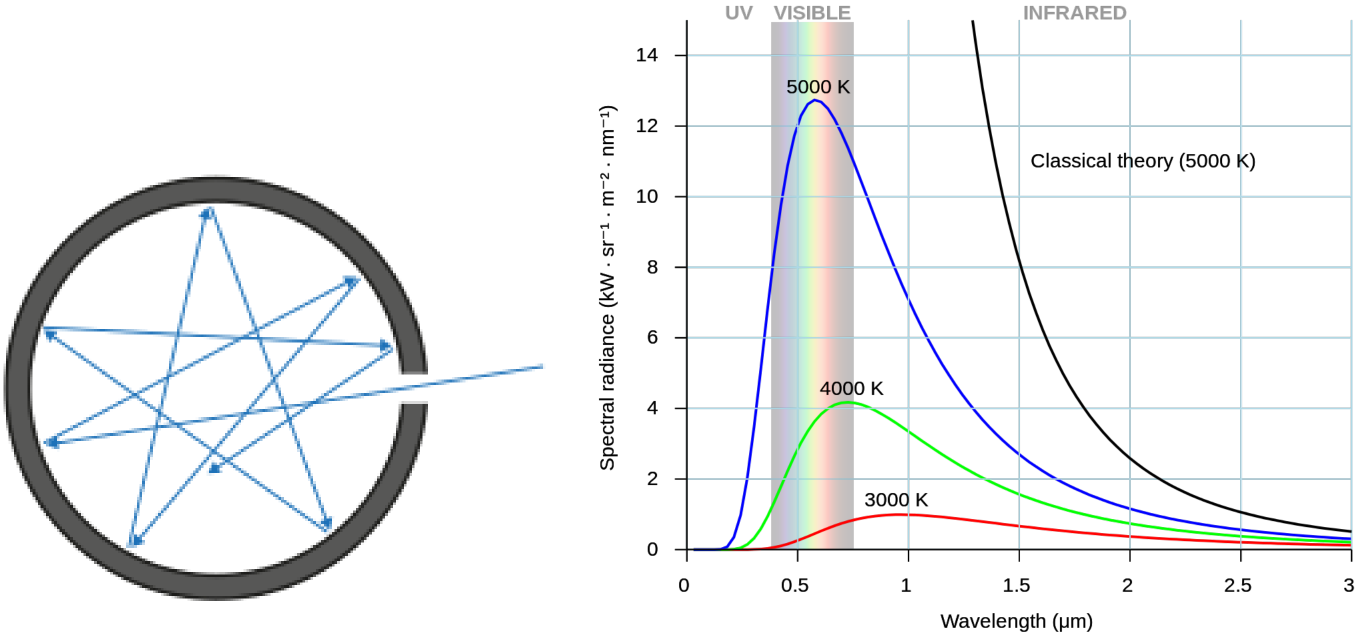


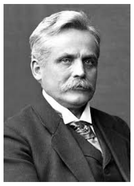
Figure 2.1: Left: A realisation of a blackbody cavity. Right: A plot of Spectral Radiance vs. wavelength for blackbody radiation. Source: The internet.

inside the blackbody

$$\rho = \frac{4P}{cA} = \frac{4\sigma}{c}T^4 = aT^4 \tag{2.2}$$

$$= \int_0^\infty \rho(\nu)d\nu \tag{2.3}$$

where $\rho(\nu)$ is the energy density in a given interval of the frequency of emitted radiation ranging from ν to dν. The constant $a = \frac{4\sigma}{c} = 7.56\times10^{-16}\mathrm{Jm}^{-3}\mathrm{K}^{-4}$. The spectral density $\rho(\nu)$ is the quantity which tells us how the radiated energy is distributed over the electromagnetic spectrum, and the total area under the curve gives us the Stefan-Boltzmann law.



Wien

A plot of $\rho(\lambda)$ versus wavelength is shown above in the right panel of Fig.2.1. The curves are for different values of temperature T, and show a non-monotonic behaviour (i.e., with a peak). Equivalent plots can also be obtained for $\rho(\nu)$ versus frequency ν (recall that $\nu \propto \frac{1}{\lambda}$). Wien's empirical observation was that the frequency of the peak was proportional to the temperature: $\nu_{peak} \propto T$.

Since $\nu \propto \frac{1}{\lambda}$, we obtain Wien's displacement law as

$$\lambda_{peak} = \frac{w}{T} \quad , \quad w = 2.9 \times 10^{-3}\mathrm{mK} \ . \tag{2.4}$$

Consequences:
(i) At T $\sim 300K$, λ_{peak} is in IR range ⇒ Thermal imaging glasses useful!

(ii) At T $\sim 5000K$, λ_{peak} shifts into the optical range ⇒, the temperature within a flame increases in going from the red part to the white part and finally the blue part!

Figure 2.2: Left: A view of wildlife through thermal imaging glasses, showing the thermal/heat map of IR radiation emitted by the body. Right: Pictures of a lit bunsen burner displaying various parts of the flame. Source: The internet.

2.1.2 Classical theory (heuristic derivation)

Consider a collection of E-M waves inside a blackbody cavity at temperature T, i.e., a superposition of harmonic waves corresponding to the various normal modes of oscillation of the cavity.

Energy density of Radiation with frequency ν is

$$\rho(\nu) = \overline{E}\ n(\nu)\ , \tag{2.5}$$

where $n(\nu)$ is the number density of wave modes within the cavity with frequency ν and $\overline{E}$ corresponds to the average energy content of the radiation.

From the classical law of equipartition of energy,

$$\overline{E} = k_B T\ , \tag{2.6}$$

where $k_B = 1.38 \times 10^{-23}\ \mathrm{JK}^{-1}$.

Now, from the fact that the wavenumber for modes within the cavity is given by

$$\vec{k} = \frac{2\pi}{L}\vec{n} \quad , \quad \vec{n} = (n_x, n_y, n_z) \quad , \quad n_x \in \mathcal{Z}, n_y \in \mathcal{Z}, n_z \in \mathcal{Z}\ , \tag{2.7}$$

where $\vec{n}$ is the mode index (not to be confused with $n(\nu)$!) and L is the spatial extent of the system along any one of the dimensions of the cavity (assumed to be a very large 3D cube), we use the relation

$$\nu = \frac{c}{\lambda} = \frac{c|\vec{k}|}{2\pi} = \frac{c}{L}n \quad , \quad n = \sqrt{n_x^2 + n_y^2 + n_z^2} \tag{2.8}$$

for e-m waves, together with the differential volume element for the case of $n >> 1$ (cavity with a large number of wave modes, such that we can treat the integer n as a continuum variable) being $dV = 4\pi n^2 dn$ to obtain

$$\nu d\nu = \frac{8\pi}{c^3}\nu^2 d\nu\ . \tag{2.9}$$

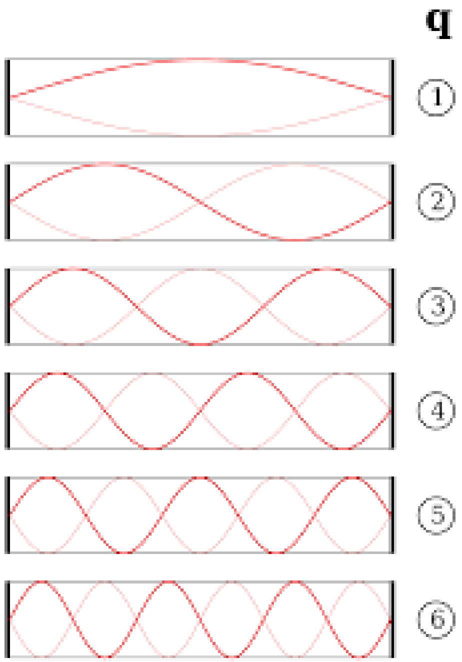


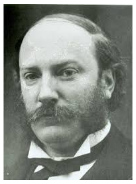
Figure 2.3: Normal modes/harmonics of a linear cavity. The topmost figure is the fundamental or first harmonic, $n = 1$, the next is the first higher harmonic ($n = 2$) and so on. Source: The internet.

In turn, this leads to

$$\rho(\nu)d\nu = \overline{E}.n(\nu)d\nu = \frac{8\pi k_B T}{c^3}\nu^2 d\nu \ . \tag{2.10}$$

The above formula is also called Rayleigh-Jeans formula, which as we shall see shortly, works only in the limit $\nu \to 0$! Taking this expression for $\rho(\nu)$ at face value, we see that it is divergent: $\rho(\nu) \to \infty$ as $\nu >> 1$, indicating a catastrophe in the UV part of the e-m spectrum! Clearly, something is wrong.

Jeans



Rayleigh

2.1.3 Planck's Idea

The average energy is computed by using the probability distribution for the energy in the range E to $E + dE$

$$P(E) = \frac{e^{-\frac{E}{k_B T}}}{k_B T} \ , \tag{2.11}$$

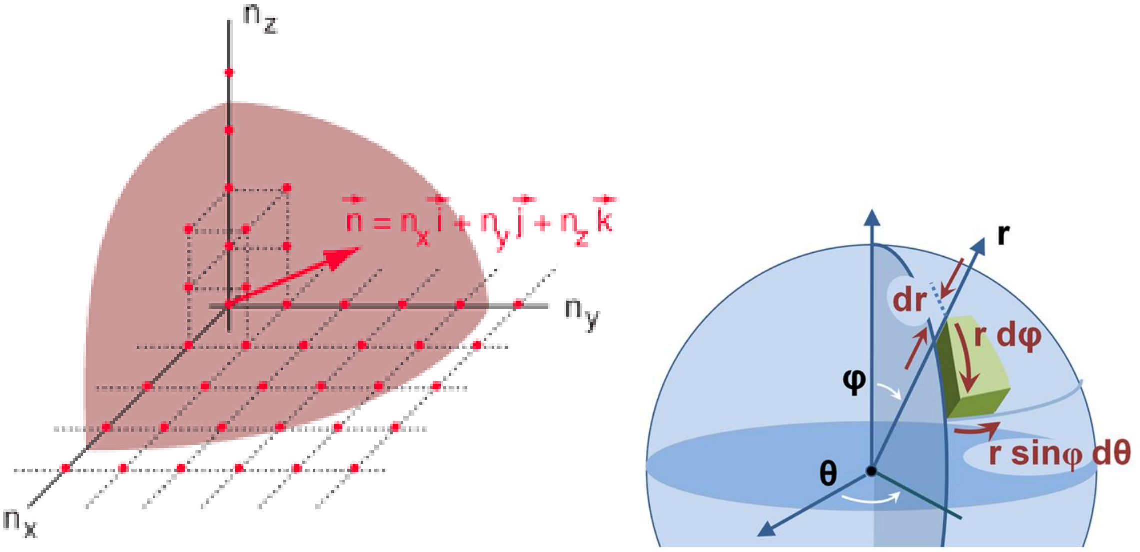


Figure 2.4: (Left) Normal modes of the cavity distributed in (n_x, n_y, n_z) space of integers. (Right) The Volume element in Spherical Polar Coordinates. Source: The internet.

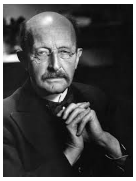
which is the Boltzmann distribution. This gives us

Boltzmann

$$\begin{aligned}
\overline{E} = \frac{\int_0^\infty EP(E)dE}{\int_0^\infty P(E)dE} &= k_BT\,\frac{\int_0^\infty d(\frac{E}{k_BT})\frac{E}{k_BT}e^{\frac{-E}{k_BT}}}{\int_0^\infty d(\frac{E}{k_BT})e^{\frac{-E}{k_BT}}} && (2.12)\\
&= k_BT\,\frac{\int_0^\infty dx\ xe^{-x}}{\int_0^\infty dx\ e^{-x}} && (2.13)\\
&= k_BT\ . && (2.14)
\end{aligned}$$

Thus, we see that the Boltzmann distribution leads to the Rayleigh Jeans formula. Note that in the above, we used the following relation (recall Gamma function integrals)

$$\int_0^\infty dx\ x^{2n}e^{-x^2/a^2} = \sqrt{\pi}\frac{(2n)!}{n!}(\frac{a}{2})^{2n+1}\ . \qquad (2.15)$$



Planck

Planck started the quantum revolution in physics in the year 1900 by quantizing the energy scale in quanta of radiation frequency ν i.e. E $= 0, h\nu, 2h\nu, 3h\nu, ...$, where h is the Planck constant. Then, recomputing the average energy for blackbody radiation, we obtain

$$\begin{aligned}
\overline{E} &= \frac{\sum_{E=0,h\nu,2h\nu,...} EP(E)}{\sum_E P(E)} && (2.16)\\
&= \frac{\sum_{n=0}^\infty \frac{nh\nu}{k_BT}e^{-nh\nu/k_BT}}{\sum_{n=0}^\infty \frac{e^{-nh\nu/k_BT}}{k_BT}} \\
&= \frac{h\nu}{e^{nh\nu/k_BT} - 1}\ . && (2.17)
\end{aligned}$$

We can see that this answer for $\overline{E}$ is quite different from that obtained from classical arguments. In obtaining the answer, we used the following relations

$$1 + x + x^2 + x^3 + \ldots = \frac{1}{1-x} \tag{2.18}$$

$$1 + 2x + 3x^2 + 4x^3 + \ldots = \frac{1}{(1-x)^2} \ . \tag{2.19}$$

This gives us

$$\begin{aligned} \rho(\nu)d\nu &= \overline{E}n(\nu)d\nu & (2.20)\\ &= \frac{8\pi}{c^3}\nu^2 d\nu \frac{h\nu}{e^{nh\nu/k_BT}-1} \\ &= \frac{8\pi h}{c^3}\frac{\nu^3}{e^{nh\nu/k_BT}-1}d\nu & (2.21)\\ &\to \frac{8\pi k_B T}{c^3}\nu^2 d\nu \quad \text{for } h\nu << 1 \ \ \text{(Rayleigh-Jeans limit)} & (2.22)\\ &\to \frac{8\pi h\nu^3}{c^3}e^{-nh\nu/k_BT} \quad \text{for } h\nu >> 1 \ \ \text{(Wien limit)} \ . & (2.23) \end{aligned}$$

Thus, we find the total energy density ρ as

$$\begin{aligned} \rho &= \int_0^\infty d\nu \rho(\nu) & (2.24)\\ &= \frac{8\pi}{c^3h^3}(k_BT)^4 \int_{x=0}^\infty dx \frac{x^3}{e^x-1} \qquad (x = \frac{h\nu}{k_BT}) \\ &= \frac{8\pi^5 k_B^4}{15c^3h^3}\, T^4 \quad \text{ as } \int_{x=0}^\infty dx\frac{x^3}{e^x-1} = \frac{\pi^4}{15} \ , & (2.25)\\ &= a\, T^4 \quad \text{where } a = \frac{8\pi^5k_B^4}{15c^3h^3} \ . & (2.26) \end{aligned}$$

This expression for ρ gives us the correct T^4 dependence of the Stefan-Boltzmann law, and gives the Stefan-Boltzmann constant as

$$\begin{aligned} \sigma &= a \times \frac{c}{4} = \frac{8\pi^5k_B^4}{15c^3h^3} \times \frac{c}{4} & (2.27)\\ &= \frac{2\pi^4k_B^4}{15c^2h^3} \ . & (2.28) \end{aligned}$$

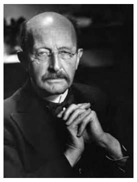
Using the values for $\sigma = 5.67 \times 10^{-8}\text{Js}^{-1}\text{m}^{-2}\text{K}^{-4}$, $c = 3 \times 10^8\text{ms}^{-1}$ and $k_B = 1.38 \times 10^{-23}$ JK^{-1}, we obtain $h = 6.6626 \times 10^{-34}\text{Js}$.

Finally, from the expression for $\rho(\nu)$, we can find the ν_{max} for which $\rho(\nu)$ is a maxima

$$\begin{aligned} \frac{d\rho(\nu)}{d\nu}|_{\nu_{max}} &= 0 \\ &= (3 - \frac{h\nu_{max}}{k_BT})e^{h\nu_{max}/k_BT} - 3 \ , & (2.29) \end{aligned}$$

which is numerically found to be

$$\nu_{max} \approx 2.8\frac{k_B T}{h} \quad \text{(Wien's Law)}. \tag{2.30}$$



Planck

Putting in the values of k_B, T and ν_{max} obtained from the experiments, we again obtain $h = 6.6626 \times 10^{-34}$Js. In this way, Planck offered a comprehensive explanation of the nature of the blackbody spectrum, i.e., he resolved any fears of the UV catastrophe by explaining the origin of the non-monotonic behaviour of the curve for $\rho(\nu)$ plotted versus ν. His courageous step involved quantising the energy variable for e-m radiation. This was a first step taken towards the theory of quantum mechanics, and thus the proportionality constant h bears his name.

2.2 The Dual Nature of Light

2.2.1 Wave-like: Diffraction and Interference

Classical optics relies on the wave nature of light. For instance, take the phenomena of diffraction and interference.

Precisely the same can be done with water waves.

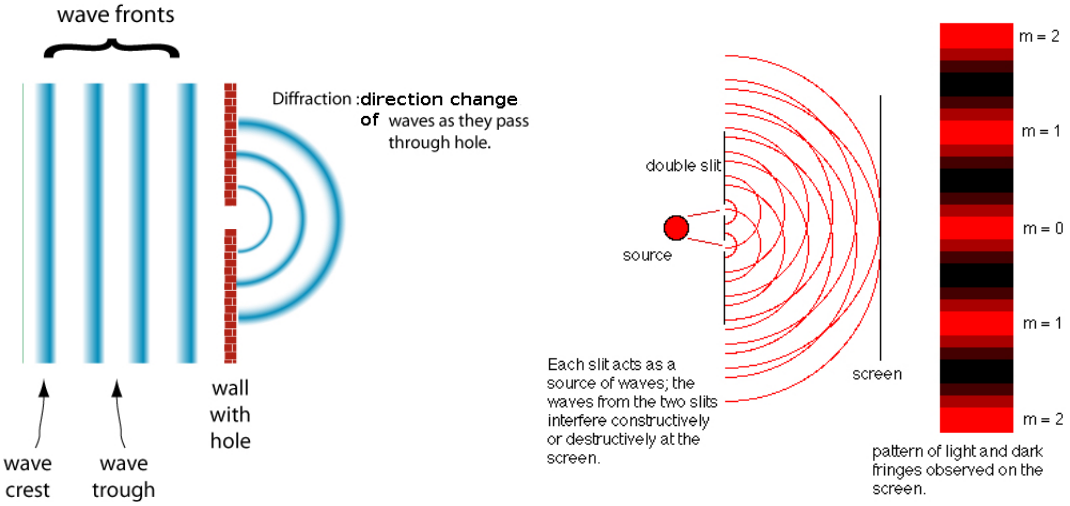


Figure 2.5: Left: Diffraction of waves from a single slit. Right: Inteference of waves from a double-slit setup. Source: The internet.

2.2.2 Particle-like: the Photoelectric effect

Maxwell's equations correctly predict the wave nature of light as the propagation of oscillating Electric and magnetic fields. They are backed up by experiments; actually they are derived from there. But this is not a completely consistent picture of light.

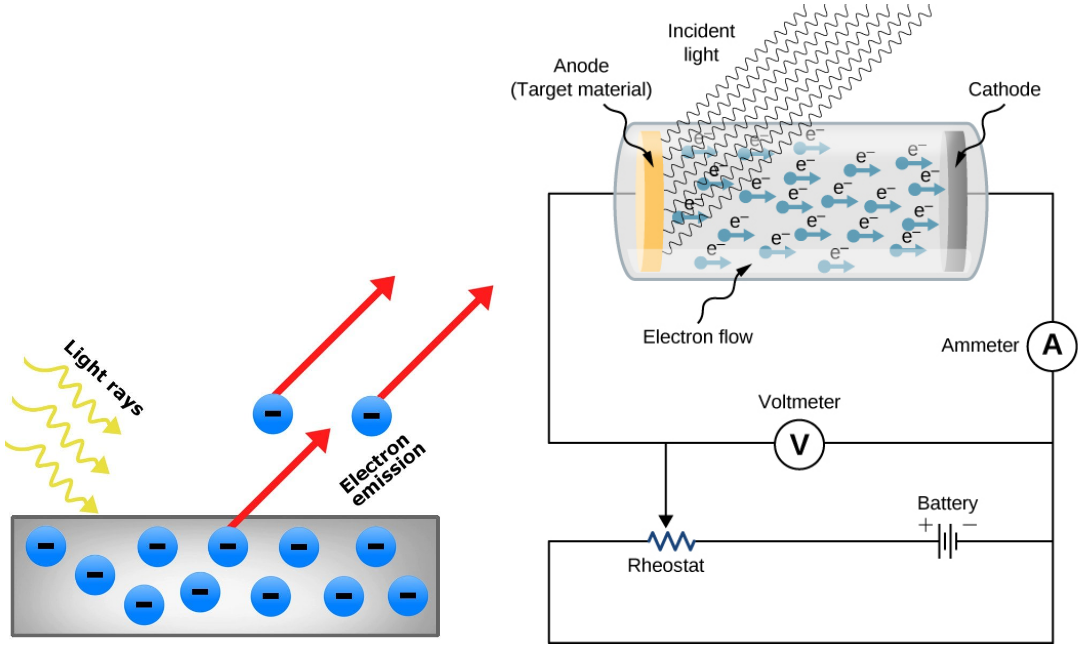


Figure 2.6: Left: The Photoelectric Effect. Right: A schematic diagram of the apparatus used for observing the Photoelectric effect. Source: The internet.

In the Photoelectric effect, shining light on a metal surface produces a current. We vary the frequency (ν) and intensity (I) of the incident light and in the process measure the current (i) and maximum energy $E_{max} = eV_0$ of the emitted electrons. This current is called the "photo-induced current" or "photocurrent", and V_0 is the "stopping potential" required to cause the current flow to stop.

Typical observations include

$$(i)\ E_{max} = eV_0 \propto \nu$$

and with a universal slope for all metals given by the Planck constant h, and

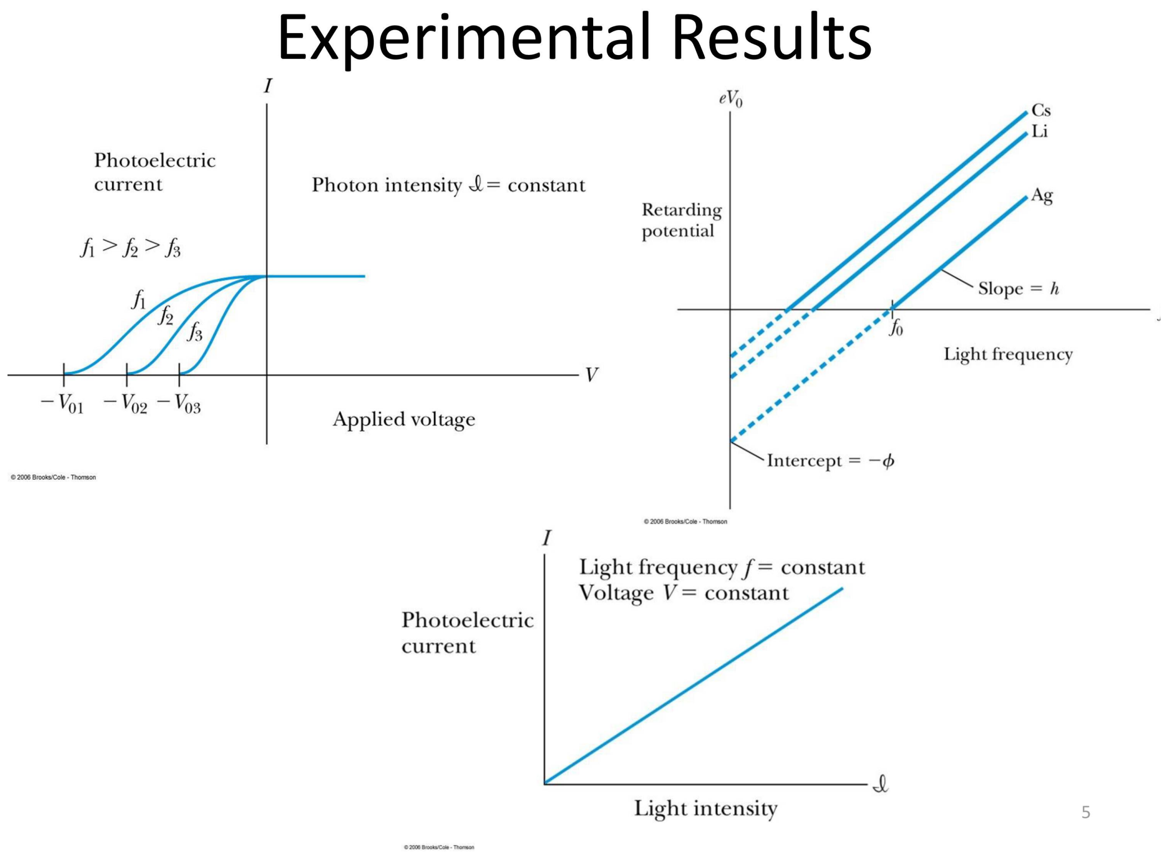
$$(ii)\ \text{photocurrent } i \propto I\ ,\ \text{ and }\ I = \frac{W}{A}$$

where i is the photoelectric current, I is the intensity of light, W is the power of the incident radiation and A is the surface area of the metal on which radiation is incident.

(iii) non-universal intercept of i vs. ν line suggests a minimum (or threshold) energy $-eV_{0,min}$ that must be paid by the incident radiation $h\nu_{min} = -eV_{0,min}$ for a finite photocurrent to be observed.

Point (ii) is certainly logical, as we expect that the higher the intensity of the incident light, the greater the number of electrons will be liberated. But why is $E_{max} \propto \nu$?

From classical E-M theory, $E_{max} \propto I$ (or the energy of radiation). Einstein suggested in 1905 that light consists of a collection of particles called PHOTONS, carrying energy $h\nu$.



Experimental Results

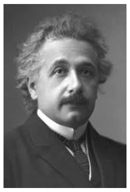
Figure 2.7: Typical experimental observations made for the Photoelectric Effect. Source: The internet.

Furthermore, he postulated that

$$E_B + E_{max} = h\nu \ ,$$

where the electrons are bound to the metal surface with a binding energy/work function E_B. E_B is therefore the minimum energy required to liberate the electron, and the rest can be transferred to the electrons as its kinetic energy (E_{max}). While the intercept (E_B) can vary between metals, the slope (h, Planck's constant) of the line is universal!

This gives another independent measurement measurement of h, suggesting its fundamental importance. It also establishes the particle-like nature of light. This explanation won Einstein the Nobel Prize in 1921.



Einstein

2.2.3 de Broglie's Idea

de
Broglie

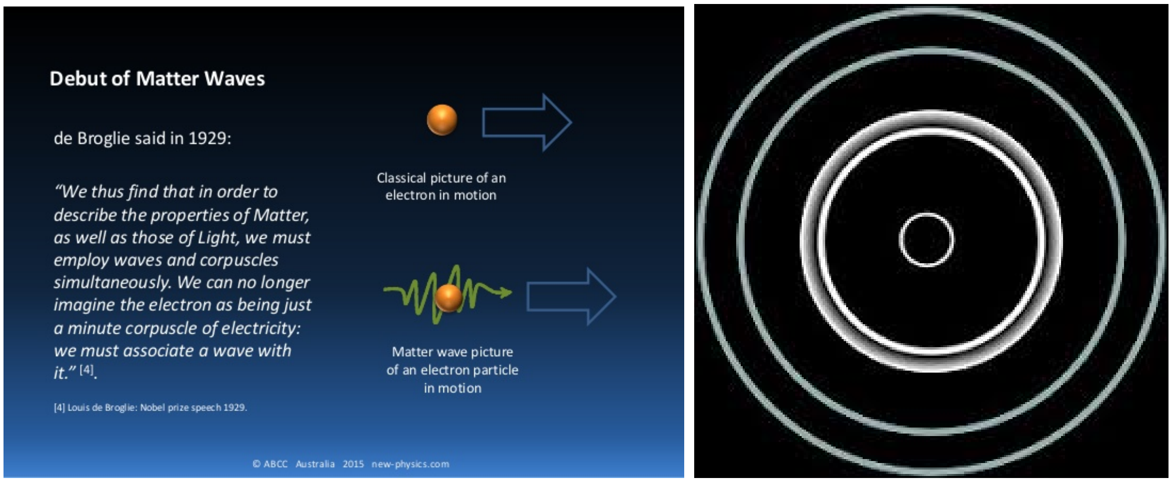
Louis de Broglie postulated that all matter particles have associated with them a wave train whose wavelength is

$$\lambda_{matter} = \frac{h}{p} \ ,$$

Davisson

where p is the particle momentum and h is the Planck constant. This suggests that h is truly a fundamental constant of nature. This can be observed in the Davisson-Germer experiment on diffraction of electrons.


Heisenberg


Debut of Matter Waves

de Broglie said in 1929:

"We thus find that in order to describe the properties of Matter, as well as those of Light, we must employ waves and corpuscles simultaneously. We can no longer imagine the electron as being just a minute corpuscle of electricity: we must associate a wave with it." [4].

[4] Louis de Broglie: Nobel prize speech 1929.

Classical picture of an electron in motion

Matter wave picture of an electron particle in motion

© ABCC Australia 2015 new-physics.com

Figure 2.8: Left: de Broglie "Matter Waves". Right: The characterisitic rings observed in the electron diffraction experiment of Davisson and Germer. Source: The internet.

Heisenberg's uncertainty principle can also be glimpsed here: for a particle possessing a definite momentum (i.e, with a vanishingly small variance), while its associated waveform is everywhere in space! This will render a large variance in its position. Contrast this with Classical Mechanics, where we can enumerate both its position and momentum with arbitrary accuracy!

Quantum tunneling through a barrier can also be understood with de Broglie's idea. Recall that for $E < V$, the particle is forbidden classically to pass through the barrier. However, by replacing the particle by a wave in quantum mechanics, the waveform can exist under the barrier, as well as "leak through" the barrier to the other side! Tunnel diodes and scanning tunneling microscopy (STM) are all applications of quantum tunneling.


Germer

2.3 The Dual Nature of Matter

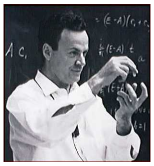
We now turn to the discussion of Feynman on the double-slit interferometer with electrons (Feynman Lecture on Physics, Vol.III).

2.3.1 Double-slit experiment with pellets

We begin by investigating the case of $1mm$ sized pellets being sprayed from a gun at a double slit in a wall with a screen made of wood and a collector which is movable.

Assume that the pellets do not break in their passage through the slits. Also, that the pellets arrive in groups at the screen (this is similar to the assumption that the slits act as a source


Young


Feynman

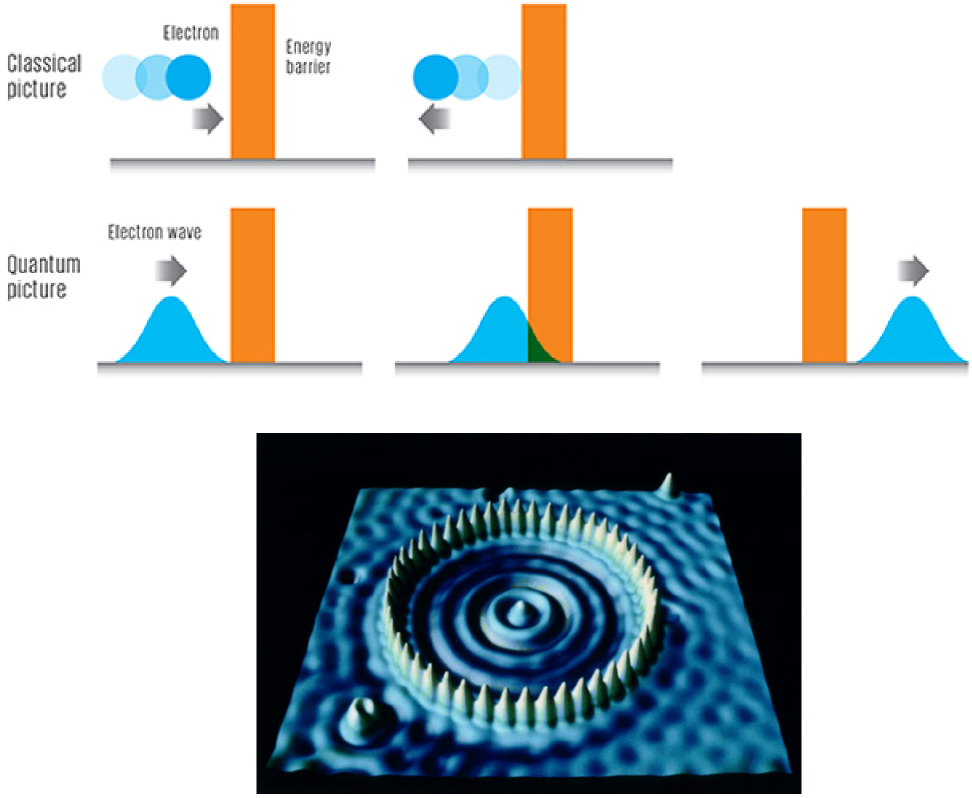


Figure 2.9: Above: The phenomenon of tunneling. Below: A “quantum corral” observed via STM measurements of Iron atoms placed on a Cu surface. Source: The internet.

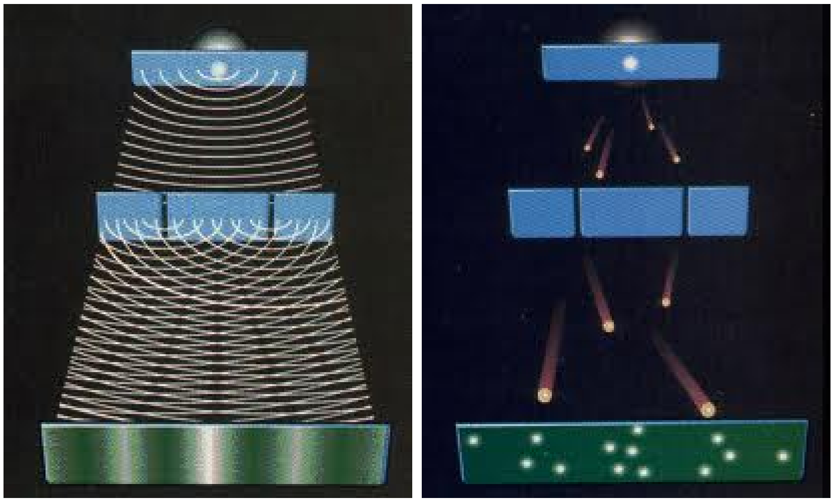


Figure 2.10: (Left) Double-slit Interference Experiment with Light waves. (Right) Double-slit Interference Experiment with Matter particles. Source: The internet.

of pellets); the size of these groups is independent of the rate of firing of the gun. Pellets are caught in the detector one at a time.

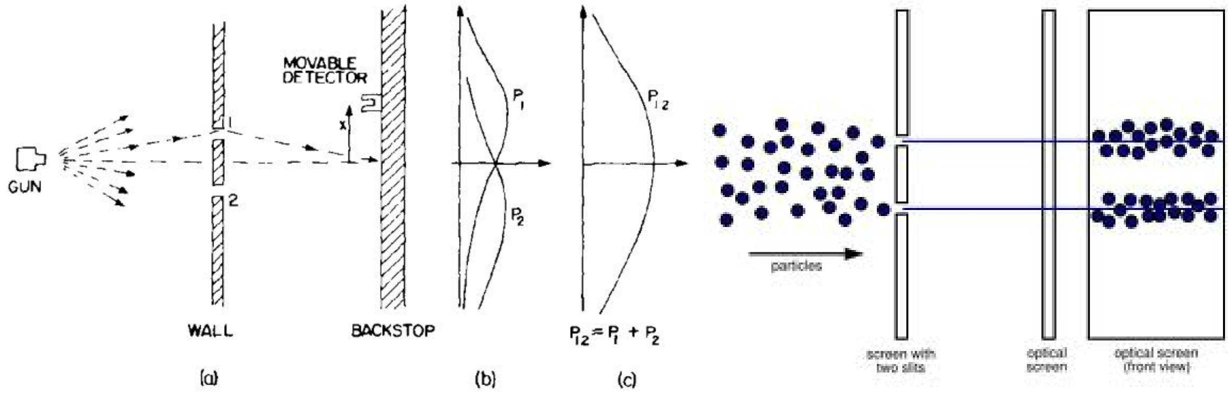


Figure 2.11: Double-slit experiment with pellets. Source: Feynman Lectures Vol.III and the Internet.

Q: What is the probability that a pellet that passes through the holes in the wall will arrive at the screen at a distance "x" from the center of the screen?

Probability means the chance that the pellet will reach the detector. It can be measured by counting the number that arrive at the detector within a particular interval of time, and then taking the ratio of this number to the total number that hit the screen within this interval. Or, for a fixed rate of firing, just the number that reached the detector in a given time interval. When either of hole 1 or 2 are closed, we get the probability distributions given by P_1 and P_2, and with both both holes open we get

$$P_{12} = P_1 + P_2$$

which is the simple addition of probabilities, i.e., no interference as in the case of Young's double slit experiment with a light source.

2.3.2 Double-slit experiment with electrons

Use an electron gun as the source (e.g., a tungsten wire heated by a electric current) surrounded by a metallic box with a pinhole. When the box is at a positive voltage with respect to the wire, electrons that are emitted from the wire will be accelerated towards the walls of the box; some will escape through the pinhole.

The electrons passing through the pinhole will be assumed to have the same energy. The double-slit wall can be a metallic sheet with two small holes, and the screen being another plate with the collector being an electron multiplier. Analogously, the entire screen can be an array of electron multipliers synchronised via a computer so as to be able to take simultaneous measurements.

Measurements can take the form of the impulse imparted on the detector: a "click", say. Note that all clicks are sharp, distinguishable but identical in, say, the amplitude of the impulse imparted. Note also that the clicks appear in no fixed sequence; they can be erratic/random.

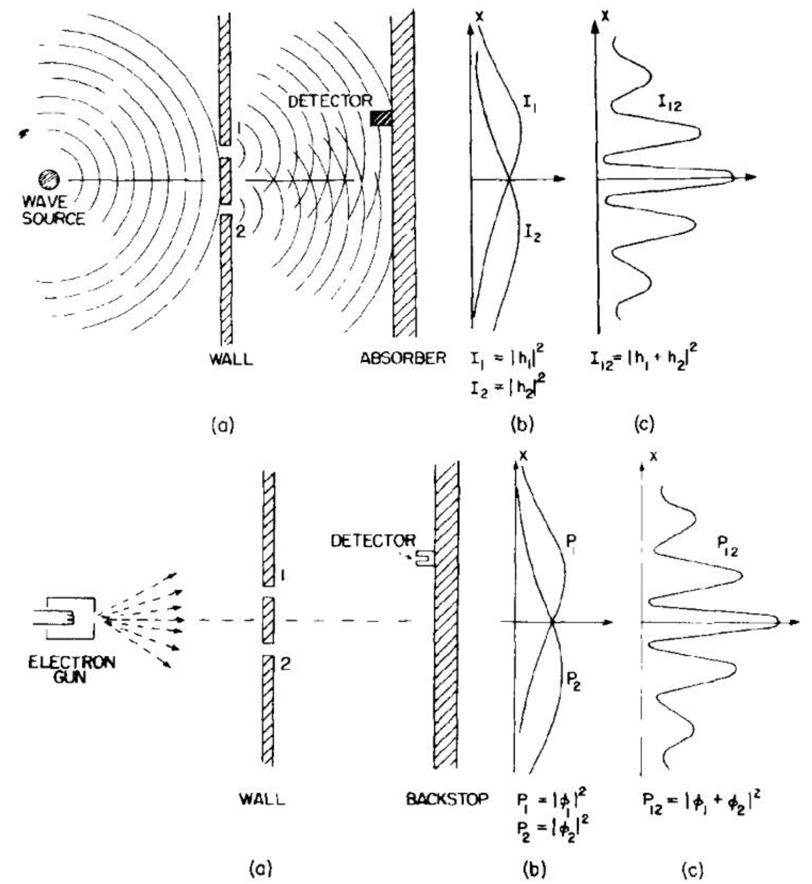


Figure 2.12: Above: Double-slit experiment with light waves. Below: Double-slit experiment with electron "matter waves". Source: Feynman Lectures Vol.III.

Several counts made over identical time periods lead to similar results for the collected clicks: this makes a time-averaged rate of clicks helpful to keep track of.

Now, as the detector is moved up and down the screen, the rate at which the clicks appear can speed up or slow down, but the amplitude of each click remains unchanged. Also, changing the temperature of the wire gun slows or speeds up the rate of clicks, but doesn't change the amplitude. Further, with two detectors in place, only a single one receives a click at any point in time (within our ability to resolve).

So far, everything suggests that the electrons are showing particle like behaviour, just as in the case of the double-slit experiment conducted with pellets.

Q: What is the probability that an electron arrives on the array at a distance "x" from the centre?

Keeping the operation of the gun fixed, measure the average rate of clicks at a distance "x" from the centre. Measurements of P_1 and P_2 with one of the slits closed appear quite reasonable!. But what about P_{12} (i.e., when both slits are open)?

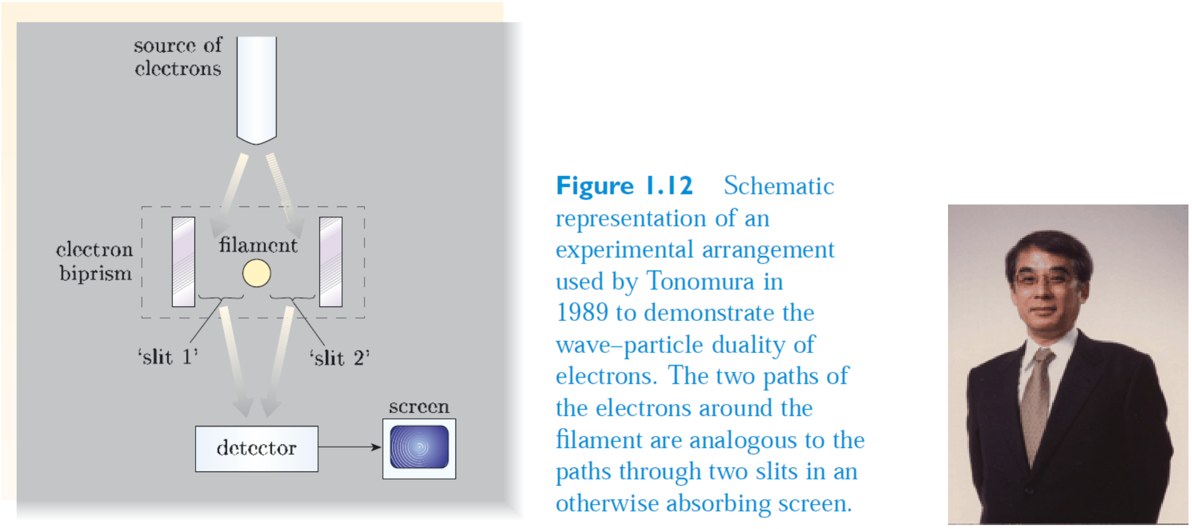


Figure 1.12 Schematic representation of an experimental arrangement used by Tonomura in 1989 to demonstrate the wave–particle duality of electrons. The two paths of the electrons around the filament are analogous to the paths through two slits in an otherwise absorbing screen.

Figure 2.13: Left: Double-slit experiment with electron "matter waves". Right: Tonomura Source: The Internet.

The outcome of the experiment:

$$P_{12} \neq P_1 + P_2!$$

The electrons shows interference when both slits are kept open! Not convinced? Take a look at Tonomura's video showing a realisation of the experiment.

$$\boxed{\text{Proposition: Each electron goes through either hole 1 or hole 2.}}$$

Can this be right if P_{12} shows signature of interference?

Note that

(i) electrons can't split in two!

(ii) the electrons are highly unlikely to take more complex paths (in through, reverse and out of the two holes)

This can be ruled out by noting that

(i) the interference pattern shows zeros where either P_1 or P_2 would have shown a finite value. Thus, closing one hole suggests the mystery that the number through the other hole has somehow increased.

(ii) at the center of the pattern, $P_{12} > 2(P_1 + P_2)$. This suggests that closing one hole somehow decreased the number through the other!

Both of the above cannot be true; i.e, they are mutually inconsistent. Thus, complex paths are ruled out.

Instead, let us define two complex numbers as A_1 and A_2 which we call Probability Amplitudes, such that the probabilities are

$$P_1 = |A_1|^2, P_2 = |A_2|^2 \text{ and } P_{12} = |A_1 + A_2|^2 \ . \tag{2.31}$$

$\boxed{\textbf{Conclusion:}}$ **The electrons arrive like particles at the screen and the probability of arrival of the particles is distributed just like the distribution of intensity of a wave!**

Since $P_{12} \neq P_1 + P_2$, the proposition given above is not verified: the experiment does not yield the information on whether the electron goes through either hole 1 or hole 2! This means that we cannot decide on which slit the electron goes through.

2.4 Spying on the Electrons

Let us try to watch on the electrons as they pass through the slits. We can do this by putting a light source in between the slits and the screen : on its way to the detector, an electron will scatter some light towards us. By placing the light source suitably, we will hopefully be able to distinguish the slit the electron passes through in terms of a flash of light in the vicinity of that slit.

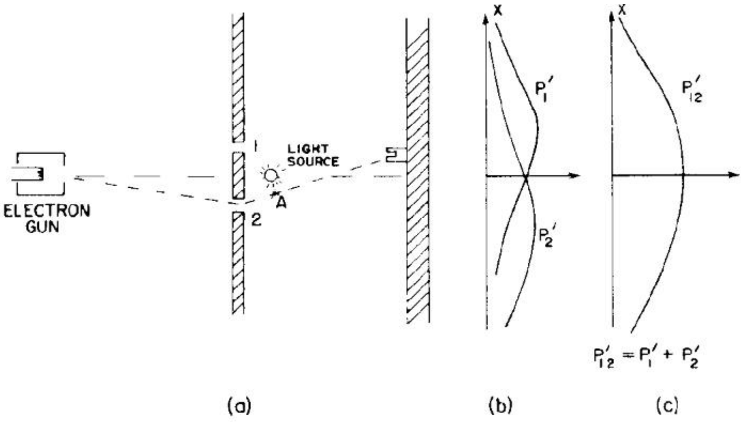


Figure 2.14: Double-slit experiment with electron "matter waves" and a light source by which to spy on which slit the electrons pass through. Source: Feynman Lectures Vol.III.

The result of this gedanken is that every time we find a "click" on the detector, we see a flash at either A (hole 1) or at B (hole 2), but never both at once! This is irrespective of the detector position. This leads to the belief that the electrons pass through either of the two slits, and the proposition holds true (in the presence of a light source) yielding a classical addition of probabilities (say, P_{12}').

This means that watching the passage of the electrons seems to remove the quantum interference we noted earlier. However, if we switch off the light source and just look at the screen, we get back the P_{12} with the interference! Why is this happening?

Is the light source disturbing the electrons by imparting momentum to them, changing their trajectories so as to give us $P_{12}^{'}$ instead of P_{12}? Lets check this out by repeating the gedanken, but by dimming the light source. As we do this, we should observe that the flashes caused by the scattering do not get weaker; rather, they become intermittent, i.e., there are occasions when the detector "clicks" but we get no flash.

This happens because some electrons are passing by unnoticed. This coincides with the reappearance of the P_{12} interference pattern! Further, this is happening because dimming the light source reduces the number of the photons that comprise the light source, leaving their energy unchanged ($\propto h\nu$, circa the photoelectric effect). Thus, some electrons escape without meeting any photons on the way.

We're a little stuck here: since the flash amplitude is independent of the light source intensity, the only electrons we can observe through flashes are ones that we are disturbing. So, if we now plot separately (i) the flashes near point A, (ii) those near point B and (iii) those that pass by unobserved (but are accounted for by the detector clicks), we will get P_1, P_2 and P_{12} respectively for them!

$\boxed{\textbf{Conclusion:}}$ **the electrons can they retain their ability to show quantum interference when they are not beign observed.**

We note the fact that since the photon momentum

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

if we use a very large λ, $p \to 0$ and the momentum imparted will be almost negligible So, lets use a very large λ light source ($\to$ red end of the spectrum) for spying on the electrons. As we gradually increase λ, we will still get P_1, P_2 and $P_{12}^{'}$ for the 3 kinds of observations. Then at some point when $\lambda \gtrsim d$ (the slit separation), we lose the ability to resolve between which slit the electron passed through, i.e., flashes for A and B are unresolvable and the flashes become fuzzy! Remarkably, it is for around this λ that the interference pattern emerges, i.e., $P_{12}^{'} \to P_{12}$, indicating that the disturbance from scattering off the photons becomes negligible enough for the interference pattern to emerge.

$\boxed{\textbf{Conclusion:}}$ **We cannot find a way to tell which slit the electron went through, and yet not disturb the interference pattern.**

This relates to a fundamental uncertainty of any quantum system and is not a limitation of the measuring apparatus: the Heisenberg uncertainty principle. So, what about our proposition: is it true or false? We can only say that it appears to be true when we are seeking to determine which hole/ slit the electron went through. But if we are not trying to do this (and is not doing so, are not disturbing the system in any way), we must conclude that the proposition cannot be true, i.e., we cannot determine which slit the electrons pass through! The inherent or intrinsic indeterminacy of the system has to be taken into account when making any analysis and any conclusion/deductions.

For larger sized objects, the associated de Broglie wave form has an extremely short λ, i.e., oscillations and interference patterns become very fine and the maxima and minima are very hard to distinguish. Any detection typically involves length scales which encompass many wavelengths of the probability distribution curve. Thus, such detection will average over the rapid oscillations, sensing only the envelope. That's why macroscopic objects such as humans will always have a well-defined choice for which of the two slits they will go through, even though the electron doesn't!

"Once you eliminate the impossible, whatever remains, no matter how improbable, must be the truth." Sherlock Holmes.

There is no mathematical theorem that renders quantum mechanics impossible.

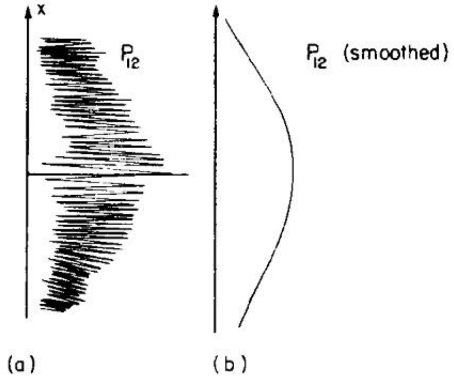


Figure 2.15: Smoothened out "envelope" of the rapid oscillations related to the matter waves. Source: Feynman Lectures Vol.III.

2.5 Fundamental Postulates of Quantum mechanics

(1) Associated with every classical outcome of a quantum experiment/setup is a probability amplitude ψ. Remember that ψ is a complex valued quantity, $\psi \in \mathcal{C}$. Also, ψ is not directly observable, nevertheless ...

(2) The probability distribution given by P = $|\psi|^2$ is directly observable

(3) For the case of different classical outcomes, the correct way to take account of all of them is to write the probability amplitude for the entire system as a linear superposition of all possible outcomes:

$$\psi_{system} = \psi_1 + \psi_2 + ...$$

Then, the probability distribution shows signature of quantum interference

$$P = |\psi_{system}|^2 = |\psi_1|^2 + |\psi_2|^2 + |\psi_3|^2 + ... + \text{ Quantum Interference terms}$$

eg. for just $\psi_1 = |\psi_1| e^{i\delta_1}$ and $\psi_2 = |\psi_2| e^{i\delta_2}$

$$P = |\psi_1|^2 + |\psi_2|^2 + 2\psi_1^* \psi_2 \tag{2.32}$$

$$= P_1 + P_2 + 2\sqrt{P_1 P_2}\cos\delta \tag{2.33}$$

where $|\psi_1| = \sqrt{P_1}$, $|\psi_2| = \sqrt{P_2}$ and $\delta = \delta_2 - \delta_1$ is the difference between the phases of the waveforms ψ_1 and ψ_2 and leads to the quantum interference. For classical systems, averaging over rapidly oscillating δ removes the quantum interference terms.

Conclusion for the double-slit experiment for electrons:

Every electron passing through the slits has a matter-wave associated with it. If the wavelength is of the order of the slit-width, the matter waves will undergo *quantum interference*, and involve a probability distribution that has peaks and troughs. However, the observation of a single electron on the screen is classical, i.e., it appears as a bright dot in Tonomura's experiment. The quantum outcome of interference fringes is nevertheless manifested when we overlay the outcome from repeating this experiment with a million electrons! This is equivalent to recording the probabilistic pattern in the Schroedinger's Cat gedanken with a large ensemble of identical setups being observed by identical observers (as discussed earlier).

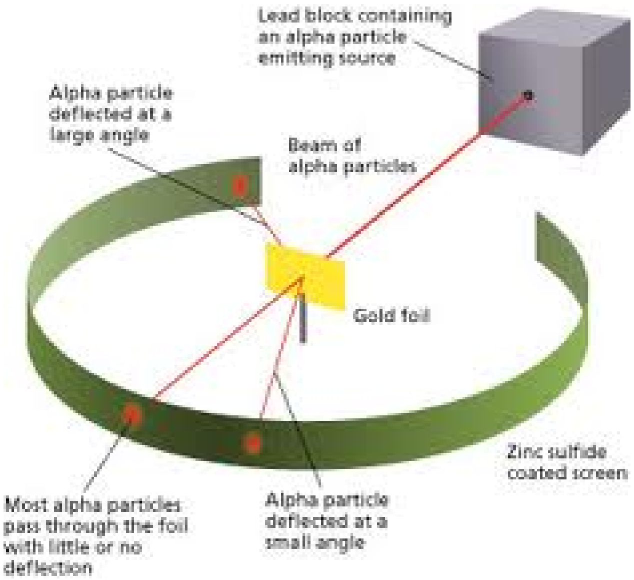
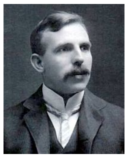


Figure 2.16: A schematic diagram of the experimental setup for the Rutherford gold foil scattering experiment, in which positively charged α particles (that emerge from the radioactive decay process of the source) are directed at a target thin gold foil. The observation of large-angle scattering provides strong evidence for the existence of a positively charged nuclear core within the atom. Source: The internet.

2.6 Bohr's Atom



Rutherford's experiment, in which positively charged α particles were scattered off a gold foil, had already helped establish the picture of a positively charged nucleus surrounded by negatively charged electrons (that are roaming in orbits).

However, this picture has two problems:

(i) the electrons must undergo centripetal acceleration in their orbits. According to electro-magnetic theory, an accelerating charge must radiate e-m radiation and lose energy in doing so. Such an orbit will then become unstable and the electron should collapse into the positively charged necleus! But no such radiation is observed; the electron orbits are stable. How can this be?

Rydberg

(ii) an accelerated electron should radiate e-m radiation in a wide range of frequencies. Instead, radiation of certain discrete wavelengths is observed experimentally when H_2 gas is heated through electrical discharge:

$$\frac{1}{\lambda} = R(\frac{1}{m^2} - \frac{1}{n^2}) \ , \tag{2.34}$$

where $m = 1, 2, 3, \ldots$, $n = 2, 3, 4, \ldots$, $n > m$ and the Rydberg constant $R = 1.097 \times 10^7 \mathrm{m}^{-1}$. Note that this relation was first obtained empirically! The sequence of lines for $m = 1$ are called the Lyman series, $m = 2$ the Balmer series, $m = 3$ the Paschen series etc.

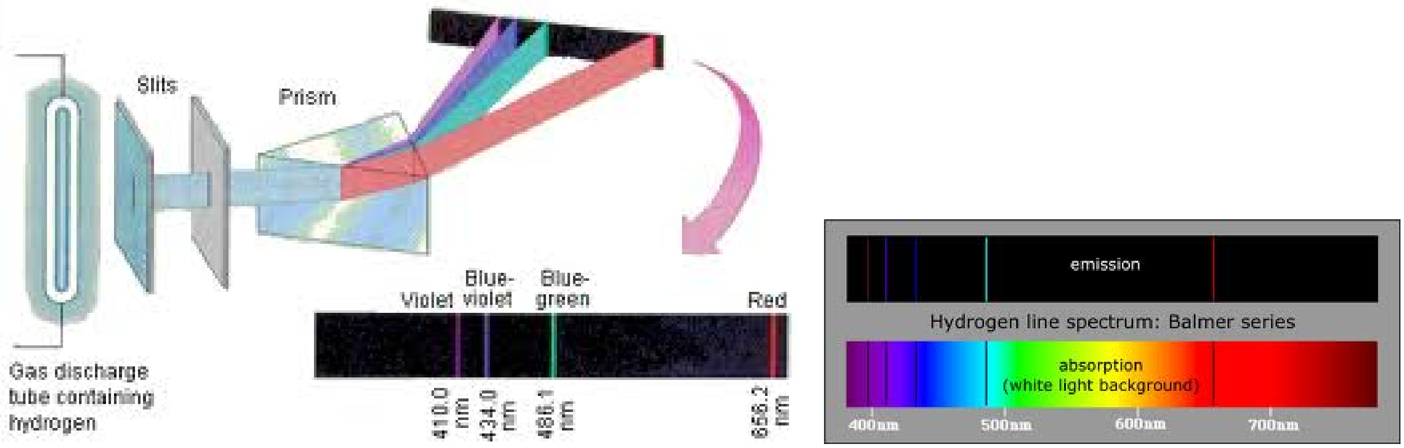


Figure 2.17: Left: A schematic diagram of the experimental setup for the observation of atomic spectra. Right: The atomic spectra of the Balmer series of the Hydrogen atom. Source: The internet.

What is the correct physical picture for the stable atom, and the radiation that is emitted? Enter Niels Bohr. He assumed that the **angular momentum of the electrons in their orbits is quantised**

$$L = pr = mvr = n\frac{h}{2\pi} = n\hbar \ , \quad n = 1, 2, 3, \ldots \tag{2.35}$$

Bohr

This quantisation can be understood from de Broglie's hypothesis as follows: assume that

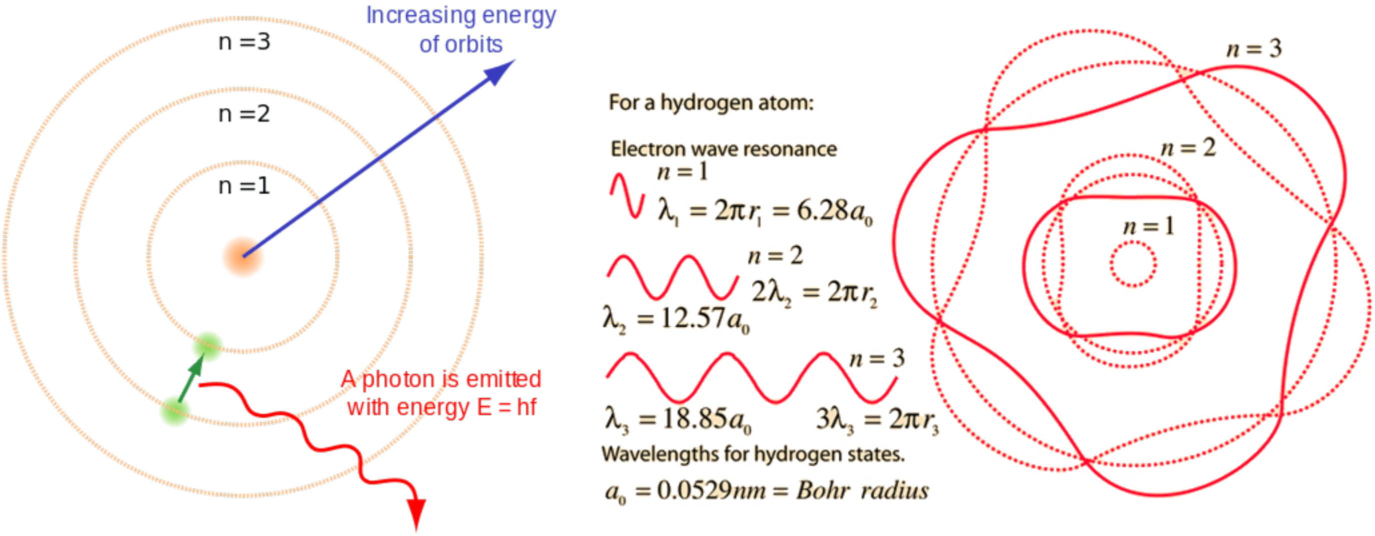


Figure 2.18: Left: A schematic diagram of the experimental setup for the observation of atomic spectra. Right: The atomic spectra of the Balmer series of the Hydrogen atom. Source: The internet.

the electron forms a standing wave in its orbit around the nucleus, with a circumference equal to an integer multiple of the wavelength λ

$$\begin{aligned} 2\pi r &= n\lambda = \frac{nh}{p} \\ L &= pr = \frac{nh}{2\pi}\ . \end{aligned} \tag{2.36}$$

The rest of the way is purely classical physics. For a classical orbit, we must equate Coulomb and centripetal forces

$$\begin{aligned} \frac{e^2}{4\pi\epsilon_0 r^2} &= \frac{mv^2}{r} \\ \Rightarrow r &= \frac{e^2}{4\pi\epsilon_0}\frac{1}{mv^2} \\ \text{together with}\ \ v &= \frac{nh}{2\pi m r}\ . \end{aligned} \tag{2.37}$$

Solving these two relations together gives

$$\begin{aligned} r &= \frac{4\pi\epsilon_0\hbar^2}{me^2}n^2 = a_0 n^2\ \ ,\ \ \text{Bohr radius}\ a_0 = \frac{4\pi\epsilon_0\hbar^2}{me^2} = 5.29\times 10^{-11}\text{m} \\ v &= \frac{e^2}{4\pi\epsilon_0\hbar}\frac{1}{n}\ . \end{aligned} \tag{2.38}$$

Putting these expressions for r and v in the total energy for the various orbits gives

Balmer

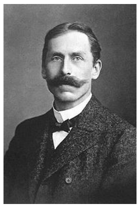
Lyman

$$\begin{aligned} E_n &= \frac{1}{2}mv^2 - \frac{e^2}{4\pi\epsilon_0 r} \\ &= -\frac{m}{2\hbar^2}\left(\frac{e^2}{4\pi\epsilon_0}\right)^2\frac{1}{n^2} \\ &= -\frac{13.6}{n^2}\text{eV} \quad , \ n = 1, 2, 3, \dots \ , \end{aligned} \tag{2.39}$$

and $1\text{eV} = 1.6 \times 10^{-19}\text{J}$. The experimentally observed spectral lines are then given by (using Planck's formula)

$$h\nu = E_{n_1} - E_{n_2} \quad , \ n_1 > n_2 \ , \tag{2.40}$$

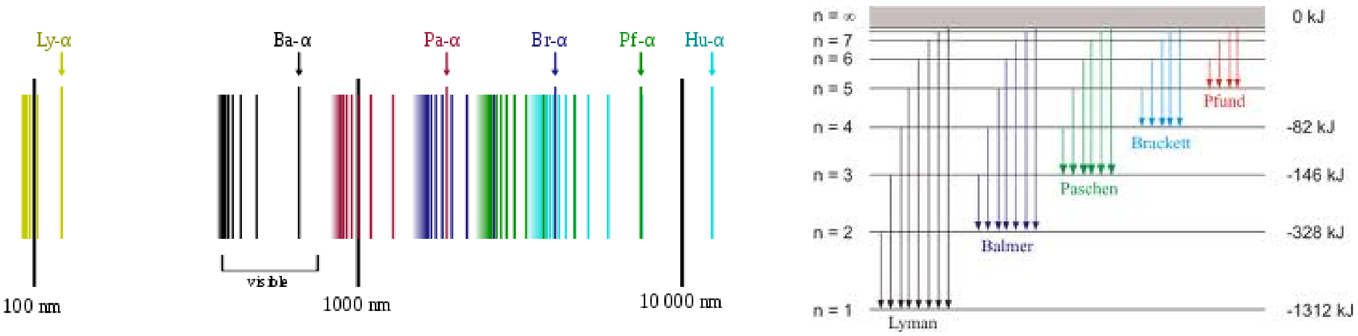
Pfund



Paschen

and n_1 and n_2 have the meaning of the initial and final energy levels of the H-atom respectively, and ν is the frequency of the emitted photon. Written in terms of the wavelength λ, the above relation yields the Rydberg constant as $R = 13.6\text{eV}/hc = 1.097 \times 10^7\text{m}^{-1}$.

In fact, by using this model, we can understand the various spectral series mentioned earlier, and their corresponding wavelengths. Also, since $n = 1$ is the lowest possible energy state, the electron cannot lose any more energy by emitting radiation. This explains the stability of the H-atom. Nevertheless, this derivation uses too many notions of classical mechanics in a regime where they are not applicable: for instance, the electron does not have a well defined orbit radius or velocity. However, as we will see towards the end of these lectures, a fully quantum mechanical treatment of the H-atom gives the same spectrum. Clearly, Bohr's assumption hits the bulls eye!



Brackett

Figure 2.19: Left: Various Spectral Series in the e-m spectrum. Right: The electronic transitions in the Hydrogen atom associated with the respective spectral series. Source: The Internet.

2.7 Conclusions

We have, in this chapter, toured several of the impressive original demonstrations of the quantum world in the form of blackbody radiation, the photoelectric effect, the electron double-slit interference experiment and atomic spectra. It is worth noting that the explanations provided for each of these demonstrations needed the quantisation of an observable (i.e., an experimentally measurable quantity) that would have taken any value in the classical world! In itself, this sets the stage for our entry into the world of the quantum.

Further, by carefully noting the duality between the particle and wave aspects of both electromagnetic radiation as well as matter, we confronted (say) the beautiful demonstration of interference of matter waves in a double-slit experiment. This led us to formulate certain postulates for the quantum world. We are now in a position to take our journey yet deeper by establishing the foundations of quantum mechanics on a firmer mathemetical footing.

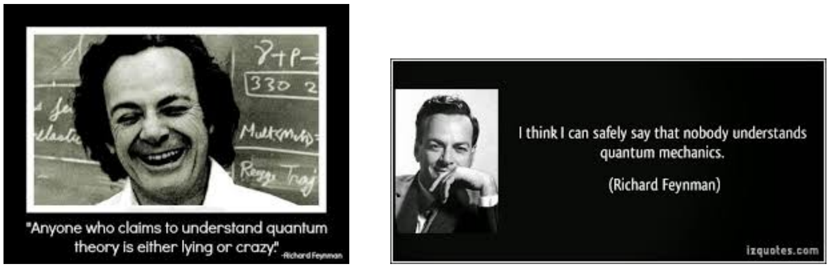


Figure 2.20: Source: The internet.

Chapter 3

The Foundations of Quantum Mechanics

3.1 Recap: Fundamental Postulates of Quantum mechanics

(1) Associated with every classical outcome of a quantum experiment/setup is a probability amplitude ψ. Remember that ψ is a complex valued quantity, $\psi \in \mathcal{C}$. Also, ψ is not directly observable, nevertheless ...

(2) The probability distribution given by P $= |\psi|^2$ is directly observable

(3) For the case of different classical outcomes, the correct way to take account of all of them is to write the probability amplitude for the entire system as a linear superposition of all possible outcomes:

$$\psi_{system} = \psi_1 + \psi_2 + ...$$

Then, the probability distribution shows signature of quantum interference

$$P = |\psi_{system}|^2 = |\psi_1|^2 + |\psi_2|^2 + |\psi_3|^2 + ... \ + \text{ Quantum Interference terms}$$

eg. for just $\psi_1 = |\psi_1|e^{i\delta_1}$ and $\psi_2 = |\psi_2|e^{i\delta_2}$

$$\begin{aligned} P &= |\psi_1|^2 + |\psi_2|^2 + 2\psi_1^*\psi_2 & (3.1) \\ &= P_1 + P_2 + 2\sqrt{P_1 P_2}\cos\delta & (3.2) \end{aligned}$$

where $|\psi_1| = \sqrt{P_1}$, $|\psi_2| = \sqrt{P_2}$ and $\delta = \delta_2 - \delta_1$ is the difference between the phases of the waveforms ψ_1 and ψ_2 and leads to the quantum interference. For classical systems, averaging over rapidly oscillating δ removes the Q.I terms.

3.1.1 Heuristic derivation of the Schrödinger equation

Promulgated by Erwin Schrödinger in 1926, this equation gives a description for the behaviour of particles in the quantum world. Following the ideas of wave-particle duality of matter, the Schrödinger equation (SE) describes the mechanics and dynamics of matter waves.

Born

Schrödinger

Unlike classical mechanics, a particle in quantum mechanics is no longer associated with a well defined trajectory $\vec{r}(t)$. Instead, we will deal with the wave amplitude for the particle as a function of ($\vec{r}$,t).

What then is $\psi(\vec{r},t)$? The straight forward answer, thanks to Max Born,is that

$$\int_V d\vec{r}|\psi(\vec{r},t)|^2 \equiv \text{Probability of finding the particle within volume V at time t.}$$

$$\text{Also, } \int_{-\infty}^{\infty} d\vec{r}|\psi|^2 = 1 \ .$$

The second relation is called the “normalisation condition” (or “square integrability” condition) for ψ, and all physically realizable solutions ψ must obey this condition. Further, all other physical observables of the particle are also related to $\psi(\vec{r},t)$.

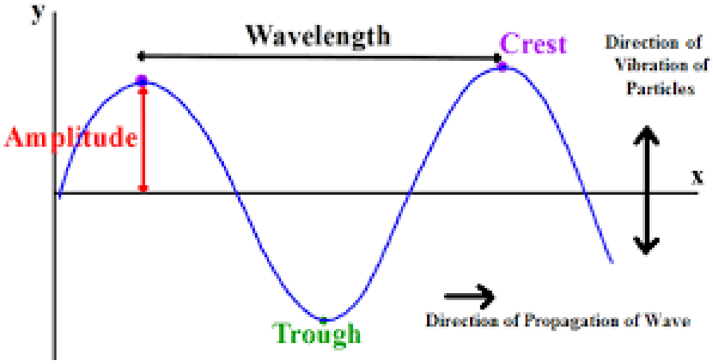


Figure 3.1: A simple waveform in 1D: $y = A\sin\left(\frac{2\pi x}{\lambda}\right)$. Source: The internet.

The derivation : Consider a wave in 1D

$$\psi(x,t) = A\cos\left(\frac{2\pi x}{\lambda} - 2\pi\nu t\right)$$

Denote $k = \frac{2\pi}{\lambda}$ (wave number) , $\omega = 2\pi\nu$ (angular frequency)

$$\psi(x,t) = A\cos\left(kx - \omega t\right)$$

This is a wave moving to the right with a phase velocity $v = \omega/k$ and group velocity $v_g = d\omega/dk$. In general, we could have written

$$\psi(\vec{r}, t) = A\cos\left(\vec{k}\cdot\vec{r} - \omega t\right) + B\sin\left(\vec{k}\cdot\vec{r} - \omega t\right) \ ,$$

or even

$$\psi(\vec{r}, t) = C\exp\Big\{i(\vec{k}\cdot\vec{r} - \omega t)\Big\} \ , \ C \in \mathcal{C}$$

A quick recap on phase and group velocities. The phase velocity v is the velocity of propagation of points of constant phase $\phi = kx - \omega t$ for *a single harmonic component* of the full wavefunction ψ, i.e., stationarity of phase ϕ with respect to time t

$$\frac{d\phi}{dt} = k\frac{dx}{dt} - \omega = 0 \to \frac{dx}{dt} \equiv v = \frac{\omega}{k} \ . \tag{3.3}$$

On the other hand, in the neighbourhood of a point within a wavepacket, the phase $\phi = kx - \omega t$ is stationary with respect to the wavevector k, i.e., various k components of the wavepacket stand in constant phase relative to one another. Then, the group velocity (v_g) given by

$$\frac{d\phi}{dk} = x - \frac{d\omega}{dk}t = 0 \to \frac{x}{t} \equiv v_g = \frac{d\omega}{dk} \ , \tag{3.4}$$

and defines the motion of all such points that are in constant relative phase to one another.

Applying the properties of e-m waves to matter-waves, we already know that

$$(i)\ E = h\nu = \hbar\omega \quad \text{(Planck)}$$
$$(ii)\ p = \frac{h}{\lambda} = \hbar k \quad \text{(de Broglie)} \ ,$$

where $\hbar = h/2\pi$. For a non-relativistic particle with mass m: $E = \frac{p^2}{2m}$.

Now, we can check that acting the following derivative on ψ

$$i\hbar\frac{\partial\psi(\vec{r}, t)}{\partial t} = i\hbar(-i\omega)C\exp\Big\{i(\vec{k}\cdot\vec{r} - \omega t)\Big\} = E\psi(\vec{r}, t)$$

Note: in the language of linear algebra, we say that ψ is the eigenfunction of the linear operator $i\hbar\frac{\partial\psi(\vec{r},t)}{\partial t}$ with eigenvalue E.

Now, for $\vec{k}\cdot\vec{r} = k_x x + k_y y + k_z z$, we can also see that $\psi(\vec{r}, t)$ is the eigenfunction of the linear operator $-i\hbar\frac{\partial}{\partial x}$ with eigenvalue $\hbar k_x$

$$-i\hbar\frac{\partial\psi(\vec{r}, t)}{\partial x} = -i\hbar(ik_x)C\exp\Big\{i(\vec{k}\cdot\vec{r} - \omega t)\Big\} = \hbar k_x\psi(\vec{r}, t) \ .$$

Similarly, $-i\hbar\frac{\partial}{\partial y}$ and $-i\hbar\frac{\partial}{\partial z}$ give eigenvalues $\hbar k_y$ and $\hbar k_z$ respectively.

$$
\begin{aligned}
\therefore -i\hbar\vec{\nabla}\psi(\vec{r},t) &= -i\hbar\left(\hat{x}\frac{\partial}{\partial x}+\hat{y}\frac{\partial}{\partial y}+\hat{z}\frac{\partial}{\partial z}\right)\psi(\vec{r},t) && (3.5)\\
&= \hbar\left(\hat{x}k_x+\hat{y}k_x+\hat{z}k_z\right)\psi(\vec{r},t) && (3.6)\\
&= \hbar\vec{k}\psi(\vec{r},t) && (3.7)\\
\Rightarrow -i\hbar\vec{\nabla}\psi &= \vec{p}\psi && (3.8)
\end{aligned}
$$

and

$$
\begin{aligned}
(-i\hbar\vec{\nabla}\psi)\cdot(-i\hbar\vec{\nabla}\psi) &= \vec{p}\cdot\vec{p}\;\psi && (3.9)\\
\Rightarrow -\hbar^2\vec{\nabla}^2\psi &= |\vec{p}|^2\psi \;=\; p^2\psi\;. && (3.10)
\end{aligned}
$$

This gives us

$$
\begin{aligned}
\frac{-\hbar^2}{2m}\nabla^2\psi &= \frac{p^2}{2m}\psi\;, && (3.11)\\
i\hbar\frac{\partial\psi(\vec{r},t)}{\partial t} &= E\psi\;. && (3.12)
\end{aligned}
$$

Since $E=\frac{p^2}{2m}$, it is entirely plausible that

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

Now assume that, in the presence of a spatially dependent potential energy $V(\vec{r})$, the extension to the above relation is

$$\hat{H}\psi \equiv -\frac{\hbar^2}{2m}\nabla^2\psi + V(\vec{r})\psi = i\hbar\frac{\partial\psi(\vec{r},t)}{\partial t}\;,$$

where we have defined the Hamiltonian operator

$$\hat{H} \equiv -\frac{\hbar^2}{2m}\nabla^2 + V(\vec{r}) = i\hbar\frac{\partial}{\partial t}\;. \tag{3.13}$$

Hamilton

Clearly, for $V(\vec{r})\neq 0$, the solution $\psi \simeq \exp\left\{i(\vec{k}\vec{r}-\omega t)\right\}$ is not guaranteed to work (though it worked very well for $V=0$).

Postulate: The above equation gives the correct $\psi(\vec{r},t)$ for any potential $V(\vec{r})$.

This has been checked and verified rigorously in many cases for the past century! Thus, we can say with some comfort that the postulate holds, and that the Schrödinger equation correctly describes the dynamics of matter waves in various situations (i.e.. described by different $V(\vec{r})$).

3.2 Physical Meaning of Ψ

Born

$\Psi(\vec{r}, t)$ is a complex quantity in Quantum Mechanics, and thus it cannot represent a physically measurable quantity. Born argued that the probability density is, however, measurable:

$$|\psi|^2 = \psi^* \psi \ . \tag{3.14}$$

Now, just as ψ is a solution of the SE, so is $c\psi$ (where c is an arbitrary complex number).In this sense, the value of the complex number c cannot be determined from just the SE. Instead, we need an additional requirement called the "Normalisation Condition" on the probability density to determine c

$$\int_{\text{all space}} d^3x \ \psi^* \psi = 1 \ . \tag{3.15}$$

As an example, consider the case of a particle in an infinite square well of extent a, with the ground state wavefunction given by

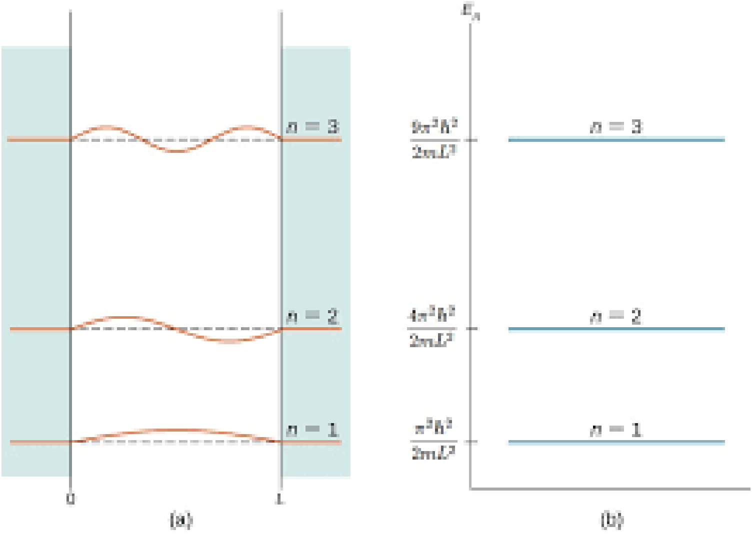


Figure 3.2: Wavefunctions (Left) and Energy Spectrum (Right) for the lowest three eigenstates of the particle in a box problem.

$$\begin{aligned} \psi &= A \sin\left(\frac{\pi x}{a}\right) \exp\left\{\frac{-i\hbar\pi^2}{2ma^2} t\right\} \ 0 \leq x \leq a \\ &= 0 \text{ for x } < 0 \text{ and } x > a \end{aligned} \tag{3.16}$$

From the normalisation condition

$$\int_0^a dx |A|^2 \sin\left(\frac{\pi x}{a}\right)^2 = 1$$

$$|A|^2 \frac{a}{2} = 1 \implies A = \sqrt{\frac{2}{a}}$$

$$\implies \psi = \sqrt{\frac{2}{a}} \sin\left(\frac{\pi x}{a}\right) e^{\frac{-i\hbar\pi^2}{2ma^2} t}$$

We could have chosen $A = -\sqrt{\frac{2}{a}}$ or even $A = \exp\{i\phi\}\sqrt{\frac{2}{a}}$ where the phase ϕ is arbitrary. This choice doesn't matter, as all observables depend on the probability density ($\propto |A|^2$) and not on the probability amplitude.

$\boxed{\text{Convention:}}$ We usually take A to be real and positive.

3.3 Observables, Operators and Expectation values

For a particle in state ψ, the expectation value of the position operator $\hat{x}$ is defined as

$$\langle\hat{x}\rangle = \int_{-\infty}^{\infty} dx\ \psi^* \hat{x} \psi = \int_{-\infty}^{\infty} dx\ x \times |\psi(x)|^2$$

$\boxed{\text{Note:}}$ It is worth stressing that the quantity $\langle\hat{x}\rangle$ is the *average* (or *mean*) position obtained from measurements made from an entire ensemble of systems, all prepared in the state ψ. It is not the average obtained by taking repeated measurements on any one system.

Similarly, we have already seen linear operators for momentum and energy $\hat{p}_x = -i\hbar\frac{\partial}{\partial x}, E = i\hbar\frac{\partial}{\partial t}$, such that we can write the expectation value for $\hat{p}_x$ as

$$\langle\hat{p}_x\rangle = \int_{-\infty}^{\infty} dx\ \psi^*(x,t)(-i\hbar\frac{\partial}{\partial x})\psi(x,t)\ .$$

These are all physically measurable quantities and are called **observables**. Very generally, in quantum mechanics,

$\boxed{\text{OBSERVABLES are associated with LINEAR OPERATORS}}$.

For any linear operator $\hat{\theta}$, its expectation value is

$$\langle\hat{\theta}\rangle = \int_v d\vec{r}\psi^*(\vec{r},t)\hat{\theta}\psi(\vec{r},t) \tag{3.17}$$

$\boxed{\text{Note:}}$ One obtains definite results for the action of a linear operator on a state only if it is an eigenstate of that operator, i.e.,

$$\hat{\theta}\psi = \theta\psi$$

In this case, it is easy to see that

$$\begin{aligned}\langle\hat{\theta}\rangle &= \int_v d\vec{r}\psi^*(\vec{r},t)\hat{\theta}\psi(\vec{r},t)\\ &= \theta\int_v \psi^*(\vec{r},t)\psi(\vec{r},t)\\ &= \theta \qquad (\text{as} \int_v \psi^*(\vec{r},t)\psi(\vec{r},t) = 1)\ .\end{aligned} \tag{3.18}$$

Conversely, if ψ is not an eigenstate of $\hat{\theta}$, acting with $\hat{\theta}$ on ψ will not give a value which will always be the same for repeated measurements on an ensemble with state ψ, i.e., we cannot predict the outcome of measuring θ on ψ with complete certainty. Instead, in such cases, the expectation value θ is always a well-defined quantity.

In fact, these **expectation values obey classical laws**. For instance,

Ehrenfest

$$\langle\hat{v}\rangle = \frac{d\langle\hat{x}\rangle}{dt} \ ,$$

$$\langle\hat{p}_x\rangle = m\langle\hat{v}\rangle = m\frac{d\langle\hat{x}\rangle}{dt} \ ,$$

$$\frac{d\langle\hat{p}_x\rangle}{dt} = \Big\langle -\frac{d\hat{V}}{dx}\Big\rangle \ .$$

These classical relations are examples of **Ehrenfest's theorem**, and are an expression of the correspondence principle for matter waves. Lets see how the third identity can be arrived at.

$$\begin{aligned}
\frac{d\langle\hat{p}_x\rangle}{dt} &= \int dx\Big(\frac{\partial\psi^*}{\partial t}\Big)\hat{p}_x\psi + \int dx\psi^*\frac{\partial\hat{p}_x}{\partial t}\psi + \int dx\psi^*\hat{p}_x\frac{\partial\psi}{\partial t} \\
&= \frac{-1}{i\hbar}\int dx(\hat{H}\psi)^*\hat{p}_x\psi + \Big\langle\frac{\partial\hat{p}_x}{\partial t}\Big\rangle + \frac{1}{i\hbar}\int dx\psi^*\hat{p}_x(\hat{H}\psi) \quad \text{(using S.E: } \hat{H}\psi = i\hbar\frac{\partial\psi}{\partial t}\text{)} \\
&= \Big\langle\frac{\partial\hat{p}_x}{\partial t}\Big\rangle + \frac{1}{i\hbar}\int dx\psi^*\Big(\hat{p}_x\hat{H} - \hat{H}^*\hat{p}_x\Big)\psi \\
&= \frac{1}{i\hbar}\int dx\psi^*\Big(\hat{p}_x\hat{H} - \hat{H}\hat{p}_x\Big)\psi \qquad \text{(as } \Big\langle\frac{\partial\hat{p}_x}{\partial t}\Big\rangle = 0\ ,\ \hat{H}^* = \hat{H}\ ,\ (\hat{H}\psi)^* = \psi^*\hat{H}^*\text{)} \\
&= \frac{1}{i\hbar}\int dx\psi^*\Big(\hat{p}_x\hat{V} - \hat{V}\hat{p}_x\Big)\psi \qquad \text{(as the kinetic energy part vanishes)} \\
&= \frac{1}{i\hbar}\times -i\hbar\int dx\psi^*\Big(\frac{\partial\hat{V}}{\partial x}\Big)\psi \qquad \text{(using the chain rule)} \\
&= \Big\langle -\frac{d\hat{V}}{dx}\Big\rangle \ . \qquad (3.19)
\end{aligned}$$

Note that the 4th line above involves the commutator of the two operators $\hat{p}_x$ and $\hat{H}$, i.e.,

$$\begin{aligned}
\hat{p}_x\hat{H} - \hat{H}\hat{p}_x &\equiv [\hat{p}_x, \hat{H}] \\
&= -\frac{\hbar^2}{2m}[\hat{p}_x, \hat{p}_x^2] + [\hat{p}_x, \hat{V}(\vec{r})] \\
&= [\hat{p}_x, \hat{V}(\vec{r})] \ , \qquad (3.20)
\end{aligned}$$

where we have used the fact that $[\hat{p}_x, \hat{p}_x^2] = 0$, i.e., the kinetic energy term commutes with $\hat{p}_x$. The commutator for two linear operators $\hat{\theta}_1$ and $\hat{\theta}_2$ is defined as

$$[\hat{\theta}_1, \hat{\theta}_2] = \hat{\theta}_1\hat{\theta}_2 - \hat{\theta}_2\hat{\theta}_1 \ . \qquad (3.21)$$

The fact that $[\hat{p}_x, \hat{p}_x^2] = 0$ can be easily established by taking $\hat{p}_x = -i\hbar\partial/\partial x$ and evaluating $[\hat{p}_x, \hat{p}_x^2]f(x)$ where $f(x)$ is some smooth, differentiable function of x. Also, note that $\hat{H}^* = \hat{H}$ relates to the Hermitian nature of the linear operator $\hat{H}$, i.e., it must have a real eigenvalue. Further, $(\hat{H}\psi)^* = \psi^*\hat{H}^*$ is the Hermitian adjoint of $\hat{H}\psi$. We will see these relations in more detail in a later lecture.

The above Ehrenfest relation is nothing but an expression of Newton's second law for matter waves!

I leave it to you to prove some of the other examples of Eherenfest's theorem as an exercise. Also, we have only been considering the expectation value (or "mean") of experimentally measurable quantities (observables) thus far. What about the fact that the such quantities will also have a variance (and therefore standard deviation), skew, kurtosis etc.? We will turn to this next.

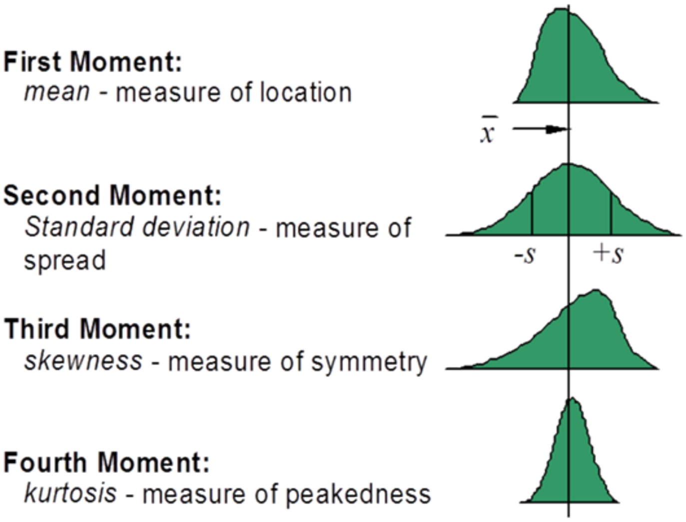


Figure 3.3: First four moments of a distribution. Source: The Internet.

Chapter 4

Heisenberg's Uncertainty principle

4.1 The case of a wavepacket

In quantum mechanics, a particle is described by a wave packet (which itself is composed of many different waves!). The wave packet surrounds the expectation value of the position, and its "centre of gravity" moves with the group velocity along the classical trajectory. As we will see, such a wave packet has an intrinsic indeterminacy in its position and momentum.

Heisenberg

Consider the wave packet through the Fourier integral at $t = 0$

$$\psi(x, t=0) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-infty}^{\infty} dp\ a(p)\ e^{\frac{ipx}{\hbar}} \quad \text{where} \quad a(p) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-infty}^{\infty} dx \psi(x) e^{\frac{-ipx}{\hbar}}$$

Using this relation, we can construct a wave packet localised within a region of width Δx and height $\frac{1}{\sqrt{\Delta x}}$, i.e., is a rectangular waveform $\psi(x)$ that is nonzero within a region of width Δx, and zero everywhere else.

The Fourier coefficient for this waveform is (for $p = \hbar k$)

$$a(p) = \sqrt{\frac{\Delta x}{2\pi\hbar}} \frac{\sin\frac{p\Delta x}{2\hbar}}{\frac{p\Delta x}{2\hbar}}$$

and $a(p)^2 \propto$ Probability that the particle has momentum p. Plotting $a(p)$ as a function of p shows maximum at $p = 0$, and the maximum contribution to the middle part comes from the narrow interval around the origin with a width

$$\Delta p \equiv \frac{\hbar}{\Delta x}$$

$\Rightarrow \Delta p \Delta x \equiv \hbar$ and this is irrespective of the shape of the wavepacket

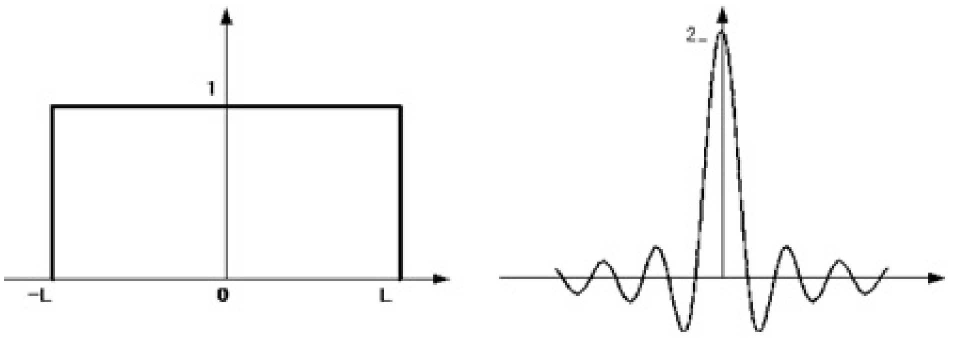


Figure 4.1: The square wavepacket and its Fourier transform (the sinc function, also known as the "Fraunhofer diffraction pattern"). Source: The internet.

It is not possible to build into the wavepacket more exact information on both the position and the momentum of the particle. This reflects the intrinsic indeterminacy of the wavepacket in terms of conjugate variables like (x,p) etc. This is the content of the Heisenberg uncertainty principle (HUP).

4.2 The γ-ray microscope gedanken

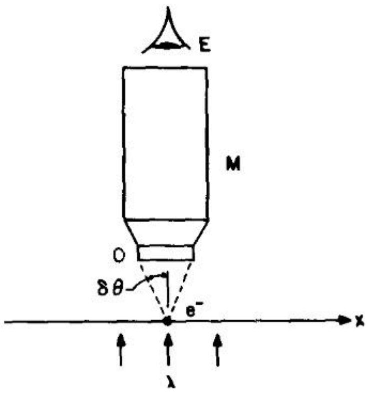


Figure 4.2: The γ-ray microscope gedanken. See text for discussion. Source: R. Shankar's text "Principles of Quantum Mechanics".

The Heisenberg Uncertainty Principle (HUP) sets a limit to the precision with which certain pairs of conjugate dynamical variables (e.g., linear position x and linear momentum p_x, angular momentum L and angular position θ etc.) can be defined *simultaneously.* There is no theoretical limit on the accuracy with which a single variable related to the wavepacket can be defined. The statement is that increasing the accuracy in the definition of a given variable makes manifest a decreasing accuracy in the definition of a variable conjugate to it, such that the uncertainty relation written in terms of both holds. It is important to note that **measurements make the intrinsic uncertainty manifest, but do not define them**.

A gedanken that highlights the intrinsic uncertainty in conjugate quantities is the γ-ray microscope. Here, light of wavelength λ shines on (i.e., scatters off) an electron (e^- in Fig.4.2 above) and then enters the objective lens (O) of a microscope (M) such that it finally reaches our eye (E). The aperture width of the lens is a, and the electron is placed at the focal distance f such that the gathering angle its position makes with respect to the lens is $2\delta\theta$ (i.e., the angle of the cone of light that enters O after scattering off the e^-). Ensure that the beam of electrons contains monoenergetic e^-s, such that we know the e^--momentum precisely ($p = \sqrt{2mE}$) before its interaction with the light. According to the HUP, we can say nothing about the position related to the matter waves associated with such e^-s. Can we use a microscope to obtain some information on their position?

Since the classical resolving power of a microscope is inversely proportional to the wavelength (λ) of light used, why not try very short-wavelength λ, i.e., γ-rays? From physical optics, the accuracy of such a measurement is

$$\Delta x = \frac{\lambda}{\sin\delta\theta} \simeq \frac{\lambda f}{a} \ . \tag{4.1}$$

Now recall that quantum mechanically, the light being used is a flux of photons with momentum $p = h/\lambda$. Indeed, for us to be able to observe an e^- in M, a photon must be scattered such that the x-component of its momentum must lie in a range

$$0 \leq p_x \leq p\sin\delta\theta \simeq \frac{ha}{\lambda f} \quad \Rightarrow \Delta p_x \simeq \frac{ha}{\lambda f}. \tag{4.2}$$

Thus, we can see that the product of Δx and Δp_x gives

$$\Delta x \cdot \Delta p_x \simeq h \ , \tag{4.3}$$

offering the HUP. This relation tells us that if we "soften the blow" imparted to the e^- by the photon by increasing λ or decreasing a (i.e., narrowing the objective lens width) such that $\Delta p_x \to 0$, we lose out on the resolution of the position of the e^-. Indeed, there is no way around the HUP: more elaborate schemes that attempt to take into account the recoil of the photon turn out to be equally futile. The appearance of the Planck constant h instead is telling us that the source of the HUP is intrinsic to the quantum world. Wave-particle duality is only another manifestation of this.

The only way out of this is to make measurement of macroscopic (i.e., classical) objects: for them, measurements typically have $h \to 0$ such that we can simultaneously have both $\Delta x \to 0$ as well as $\Delta p_x \to 0$. In this way, we can justify how the arbitrary precision of Newtonian mechanics emerges from the Heisenberg uncertainty principle quantum mechanics. Note, though, that taking the limit of $h \to 0$ is a heuristic tool ... we don't really understand at present how this emergence takes place!

4.3 A derivation of the HUP for wavepackets

Definition of Uncertainty in expectation value of an operator $\hat{\theta}$ is

$$\Delta\theta = \sqrt{\langle(\theta - \langle\theta\rangle)^2\rangle} \text{ akin to the standard deviation in statistics}$$

Now,

$$\langle x\rangle = \int_{-\infty}^{\infty} dx\ \psi^* x\psi = \int_{-\infty}^{\infty} dx\ x|\psi(x)|^2$$
$$\langle \hat{p}_x\rangle = \int_{-\infty}^{\infty} dx\ \psi^*(x,t)\big(-i\hbar\frac{\partial}{\partial x}\big)\psi(x,t)\ .$$

Note that we are dropping the ˆ (hat) symbol above all operators for the sake of brevity; we assume henceforth that the reader will understand their presence from the context. In order to reach the HUP, let us start with

$$\begin{aligned}\int dx\ i\hbar\frac{d\psi^*}{dx}x\psi &= -i\hbar\int dx\ \psi^*\psi - i\hbar\int dx\frac{d\psi}{dx}x\psi^* \quad (\text{where } \psi^*x\psi|_{-\infty}^{\infty}\to 0 \text{ as } \psi,\psi^*|_{-\infty}^{\infty}\to 0)\\ &= -i\hbar\int dx\ \psi^*\psi + \Big[i\hbar\int\ dx\ \frac{d\psi^*}{dx}x\psi\Big]^* \quad (\text{as } x^*=x)\ . \qquad (4.4)\end{aligned}$$

Since $c - c^* = 2i\text{Im}[c]\ ,\ c\in\mathcal{C}$, we can rewrite the above as

$$i\hbar\int dx|\psi|^2 = 2i\text{Im}\Big[\int i\hbar dx\frac{d\psi^*}{dx}x\psi\Big] \qquad (4.5)$$
$$\therefore \hbar^2\Big|\int\ dx|\psi|^2\Big|^2 = 4\Big|Im\int\ dx\ i\hbar\frac{d\psi^*}{dx}x\psi\Big|^2 \le 4\Big|i\hbar\int dx\frac{d\psi^*}{dx}x\psi\Big|^2\ , \qquad (4.6)$$

where in the second line, we have used the relation $|\text{Im}[c]|^2 \le |c|^2\ ,\ c\in\mathcal{C}$. Now, using the Schwarz inequality $|c_1^*c_2|^2 \le |c_1|^2\times|c_2|^2$, we obtain

$$\Big|\int\ dx\ i\hbar\frac{d\psi^*}{dx}x\psi\Big|^2 \le \int\ dx\ x\psi^*x\psi\int\ dx\ \Big(i\hbar\frac{\partial\psi^*}{\partial x}\Big)\Big(-i\hbar\frac{\partial\psi}{\partial x}\Big)$$
$$\Big|\int\ dx\ i\hbar\frac{d\psi^*}{dx}x\psi\Big|^2 \le \int\ dx\ x^2|\psi|^2\int\ dx\ \Big|-i\hbar\frac{\partial\psi}{\partial x}\Big|^2\ . \qquad (4.7)$$

Thus, the inequality relation eq.(4.6) becomes

$$\frac{\hbar^2}{4}\underbrace{\Big|\int\ dx\ |\psi|^2\Big|}_{=1} \le \underbrace{\int\ dx\ x^2|\psi|^2}_{\langle x^2\rangle} \times \underbrace{\int\ dx\ \Big(i\hbar\frac{\partial\psi^*}{\partial x}\Big)\Big(-i\hbar\frac{\partial\psi}{\partial x}\Big)}_{\langle p^2\rangle}$$
$$\Rightarrow \frac{\hbar^2}{4} \le \langle x^2\rangle\langle p^2\rangle\ . \qquad (4.8)$$

Recall that the relations for the standard deviations in x and p are

$$\Delta x = \sqrt{\langle(x-\langle x\rangle)^2\rangle}\ ,\ \Delta p = \sqrt{\langle(p-\langle p\rangle)^2\rangle}\ , \qquad (4.9)$$

we replace x and p by their generalisations $x-\langle x\rangle$ and $p-\langle p\rangle$ respectively in eq.(4.8) to obtain

$$\frac{\hbar^2}{4} \le \langle(x-\langle x\rangle)^2\rangle\langle(p-\langle p\rangle)^2\rangle \equiv \Big(\Delta x\Big)^2\Big(\Delta p\Big)^2 \qquad (4.10)$$
$$\Rightarrow \frac{\hbar}{2} \le \Delta x\Delta p\ . \qquad (4.11)$$

This is the celebrated form of the Heisenberg Uncertainty Principle (HUP) relation between x and p. It shows clearly the existence of a lower bound on the precisions that can be obtained on simultaneous measurement of x and its conjugate momentum p.

$\boxed{\text{Note:}}$ As we have said earlier, there are several other examples of similar uncertainty relations in quantum mechanics, e.g., angular momentum and angular position etc. One such is that between energy E of a system E and time t: $\Delta E \Delta t \geq \hbar/2$. This relation cannot be derived from the above approach as t is not a dynamical variable like x and p; rather, t is a parameter in quantum mechanics. Thus, the content of this uncertainty relation is different from those that involve only dynamical variables. The uncertainty relation for E and t has the following (rough) meaning:

The energy E of a system that has only been in existence for a time (i.e., its "lifetime") Δt has an uncertainty (or spread about its mean value) of atleast ΔE, such that the above quoted relation is satisfied. Clearly, for a stationary energy eigenstate, $\Delta E = 0$, and its lifetime $\Delta t \to \infty$. On the other hand, a state with a finite lifetime will undergo decay; such states have a finite and non-zero ΔE. The origin of this relation will be revealed in further courses up ahead. For now, suffice it to say that there are several contexts in which such decay processes can show up, two being:

(a) an open quantum system, i.e., a quantum system connected to its environment. Here, if the system is prepared in one (or a superposition) of the eigenstates of the isolated system, it will "decay" into the true eigenstates of the coupled system, and

(b) an interacting quantum system, i.e., a system with inter-constituent interactions. In such a many-body problem, if the system is prepared in one (or a superposition) of the eigenstates of the non-interacting system, it will "decay" into the true eigenstates of the full interacting system.

Note that in both cases, the true eigenstates are very difficult to learn in general. Most of our theories for such complex quantum systems typically work with those that are almost isolated and weakly interacting.

Chapter 5

The Time Dependent and Independent Schrödinger Equations (TDSE & TISE)

The Schrödinger equation in its most general form is

$$\hat{H}\Psi(\vec{r},t) \equiv \frac{-\hbar^2}{2m}\nabla^2\Psi(\vec{r},t) + V(\vec{r},t)\Psi(\vec{r},t) = i\hbar\frac{\partial\Psi}{\partial t} \ ,$$

where the various cases we are interested in studying are characterized by different functional forms of $V(\vec{r},t)$. This is the time dependent Schrödinger equation (TDSE).

Schrödinger

For a special case of potentials $V(\vec{r})$ (i.e., no explicit time dependence in the potential), we have

$$\hat{H}\Psi(\vec{r},t) = \Big[\frac{-\hbar^2}{2m}\nabla^2 + V(\vec{r})\Big]\Psi(\vec{r},t) = i\hbar\frac{\partial\Psi(\vec{r},t)}{\partial t}$$

Several cases of such potentials will be encountered in the days ahead. Concrete examples include the potential felt by an electron in an atom or in a crystal lattice.
We require that $\Psi(\vec{r},t)$ be an eigenfunction of the energy operator $i\hbar\frac{\partial}{\partial t}$ with energy eigenvalue E

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

such that

$$\hat{H}\Psi(\vec{r},t) = \Big[\frac{-\hbar^2}{2m}\nabla^2 + V(\vec{r})\Big]\Psi(\vec{r},t) = E\Psi(\vec{r},t) \ .$$

This is called the time independent Schrödinger equation (TISE). We can solve the equation above using separation of variables

$$\Psi(\vec{r},t) = \psi(\vec{r})\chi(t) \ ,$$

such that

$$
\begin{aligned}
i\hbar\psi(\vec{r})\frac{\partial\chi(t)}{\partial t} &= E\psi(\vec{r})\chi(t) && (5.1)\\
\implies \dot{\chi(t)} &= \frac{-iE}{\hbar}\chi(t)\ , && (5.2)\\
\chi(t) &= e^{\frac{-iEt}{\hbar}}\ , && (5.3)
\end{aligned}
$$

choosing $\chi(t=0)=1$ as our normalisation choice.
Similarly,

$$
\begin{aligned}
\left[\frac{-\hbar^2}{2m}\nabla^2 + V(\vec{r})\right]\psi(\vec{r})\chi(t) &= E\psi(\vec{r})\chi(t)\\
\text{or, } \left[\frac{-\hbar^2}{2m}\nabla^2 + V(\vec{r})\right]\psi(\vec{r}) &= E\psi(\vec{r}) && (5.4)\\
& && (5.5)
\end{aligned}
$$

such that

$$\Psi(\vec{r},t) = \psi(\vec{r})e^{\frac{-iEt}{\hbar}} \tag{5.6}$$

is the full wavefunction. We can also define a time evolution operator $U = e^{-\frac{i\hat{H}t}{\hbar}}$, such that

$$
\begin{aligned}
U\Psi(\vec{r},t) &= e^{-\frac{i\hat{H}t}{\hbar}}\Psi(\vec{r},t)\\
&= e^{-\frac{iEt}{\hbar}}\Psi(\vec{r},t)\ , && (5.7)
\end{aligned}
$$

where E is the eigenvalue of the Hamiltonian $\hat{H}$ for the state $\Psi(\vec{r},t)$. We will see later that the Hamiltonian corresponds to a **Hermitian** operator with real-valued eigenvalues $E \in \mathcal{R}$. In such cases, the unitary time evolution operator $U = e^{-\frac{i\hat{H}t}{\hbar}}$ (i.e., the exponentiation of a Hermitian operator) corresponds to a **Unitary** operator, whose eigenvalue is simply the phase factor $\chi(t) = e^{\frac{-iEt}{\hbar}}$.

Further, it is easily seen that

$$
\begin{aligned}
\text{Probability Density}: \ \Psi^*(\vec{r},t)\Psi(\vec{r},t) &= \psi^*(\vec{r})e^{\frac{iEt}{\hbar}}\psi(\vec{r})e^{\frac{-iEt}{\hbar}} = \psi^*(\vec{r})\psi(\vec{r}) \ \text{ and } && (5.8)\\
\text{Expectation Value}: \ \langle\hat{\theta}\rangle &= \int_v d\vec{r}\psi^*(\vec{r})e^{\frac{iEt}{\hbar}}\hat{\theta}\psi(\vec{r})e^{\frac{-iEt}{\hbar}}\\
&= \int_v d\vec{r}\psi^*(\vec{r})\hat{\theta}\psi(\vec{r})\ . && (5.9)
\end{aligned}
$$

From these two relations for the Probability density and the Expectation value, we learn that the phase accrued from time evolution $\chi(t) = e^{\frac{-iEt}{\hbar}}$ does not change the total probability (or "norm") associated with the state $\Psi(\vec{r})$ in time. We denote such a probability preserving time evolution as "**unitary**".

As mentioned earlier, the time independent Schrödinger equation (TISE, eq.(5.5)) also helps us define the Hamiltonian operator

$$
\begin{aligned}
\hat{H} &= \frac{-\hbar^2}{2m}\nabla^2 + V(\vec{r}) \ \text{ whose eigenvalue is } E\ , && (5.10)\\
\hat{H}\psi(\vec{r}) &= E\psi(\vec{r})\ . && (5.11)
\end{aligned}
$$

Further, time evolution of a system is neatly encoded by the (energy) eigenstates of the Hamiltonian operator. This makes them special.

Degenerate Eigenstates: Also, note that if

$$i\hbar\frac{\partial\Psi_1}{\partial t} = E\Psi_1\ ,\ \Psi_1 = \psi_1(\vec{r})\chi_1(t) \qquad \& \qquad i\hbar\frac{\partial\Psi_2}{\partial t} = E\Psi_2\ ,\ \Psi_2 = \psi_2(\vec{r})\chi_2(t)\ , \tag{5.12}$$

$$\text{i.e.,}\ \ \hat{H}\psi_1 = E\psi_1 \quad , \quad \hat{H}\psi_2 = E\psi_2\ , \tag{5.13}$$

then

$$\begin{aligned} \hat{H}\Big(c_1\psi_1 + c_2\psi_2\Big) &= Ec_1\psi_1 + Ec_2\psi_2 \ \ (c_1, c_2 \in \mathcal{C})\ , \\ &= E\Big(c_1\psi_1 + c_2\psi_2\Big)\ , \end{aligned} \tag{5.14}$$

i.e., $c_1\psi_1 + c_2\psi_2$ is also an eigenfunction of $\hat{H}$ with the same energy eigenvalue E. In this case, the wavefunctions ψ_1 and ψ_2 are said to be **degenerate with respect to one other**.

5.1 Qualitative Solutions and the origin of Quantisation

We will now try to reach some qualitative conclusions on how "proper" wavefunctions ψ should look:

1. As $x \to \pm\infty$, $\psi(x) \to 0$; else, ψ is not normalisable (or square integrable). This is especially important in distinguishing the physically sensible wavefunctions in parts of the potential $V(x)$ that are classically forbidden.
2. The wavefunction ψ cannot be discontinuous. Physically, this makes no sense, as we would need extra information with which to make sense of missing probabilities $|\psi|^2$.
3. The spatial derivative $\frac{d\psi}{dx}$ is continuous for all x, unless the potential $V(x)$ is very peculiar (such that $\psi(x)$ has kinks). Can you think of such a peculiar $V(x)$?

For a wavefunction $\psi(x)$ that satisfies all these three conditions, the particle can exist classically in regions of x where the potential $V(x) < E$, while it is classically forbidden in regions where $E < V(x)$. In the classical world, a particle is bound to regions where $V(x) < E$. However, in the quantum world, perfectly legal $\psi(x)$ can be found for both $V(x) < E$ as well as $E < V(x)$. We will encounter this with bound states (e.g., in the simple harmonic oscillator problem) as well as with scattering (i.e., spatiall extended) states (e.g., in the barrier tunneling problem).

The origin of the phenomenon of quantisation (or discretisation of the energy spectrum/ allowed values of the eigenvalues E) lies in the fact that for states that are classically bound, the TISE has solutions ψ that satisfy all of the three criteria above *only at certain discrete values of E*. For instance, for a particle inside the infinitely deep well, it is easily seen that the simplest plausible $\psi(x)$ that satisfies curvature function $-\frac{\hbar^2}{2m}\vec{\nabla}^2$ as well as all three

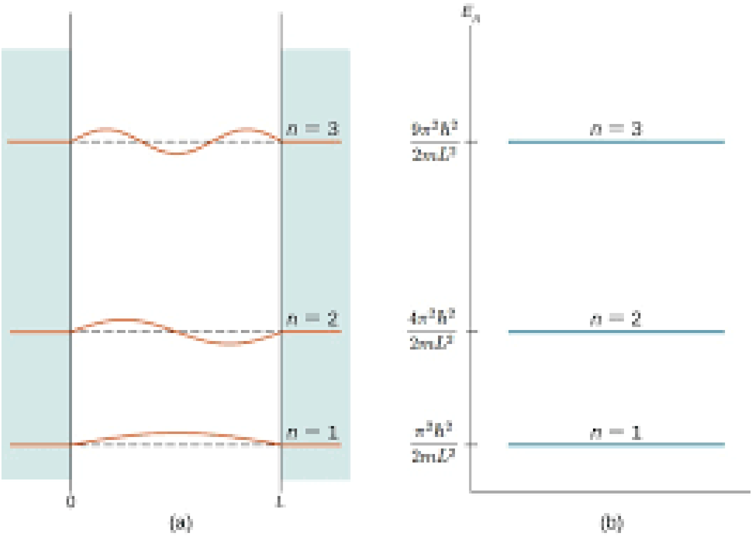
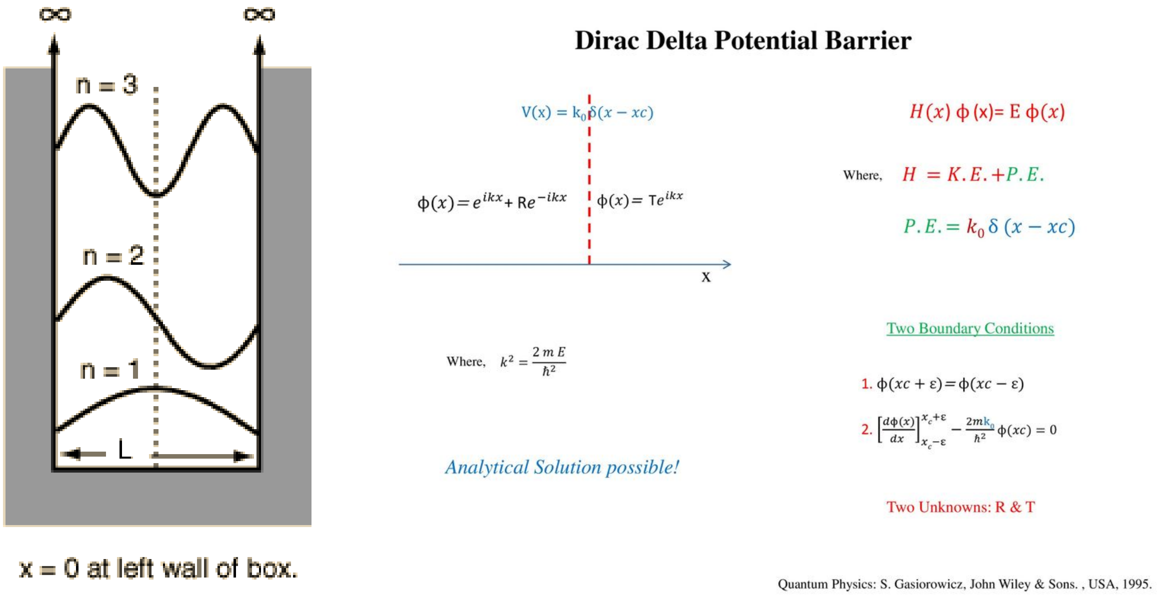


Figure 5.1: Wavefunctions (Left) and Energy Spectrum (Right) for the lowest three eigenstates of the particle in a box problem.

criteria given above is the one that characterises **the ground state** $\psi_{n=0}$: it has the minimum curvature (and thence the minimum kinetic energy) and vanishes at the walls of the potential (as $V \to \infty$ there). Tuning the energy E continuously does not reveal another solution $\psi_{n=1}$ vanishingly close by in energy to $\psi_{n=0}$. However, one finds other solutions at particular discrete values. Clearly, these will all have greater kinetic energy than the ground state, simply as they have greater curvature $-\frac{\hbar^2}{2m}\vec{\nabla}^2$.

Indeed, the next simplest and plausible ψ, $\psi_{n=1}$ is one that changes the curvature $\frac{d^2\psi(x)}{dx^2}$ qualitatively in comparison to that for $\psi_{n=0}$: $\psi_{n=1}$ crosses the x-axis once and is called **the first excited state**. A general result for excited bound states is that the nth excited state ($n \in \mathcal{Z}$) crosses the x-axis n times. Thus, the discrete bound states ψ_n have an energy E_n arising from an ever-increasing curvature due to n crossings of the x-axis. Also, if the potential $V(x)$ has a centre of symmetry, the eigenstates will be alternating even parity ($\psi(-x) = \psi(x)$) and odd parity ($\psi(-x) = -\psi(x)$) functions about that centre of symmetry.

For cases where $E > V(x)$ for all x, only unbound (or "scattering") states exist. These states do not have a tendency to diverge at $x \to \pm\infty$ (though you have to be a little careful in showing that!) and their curvature $\frac{d^2\psi(x)}{dx^2}$ can vary continuously as E is varied. Thus, these unbound states form a continuous (rather than discrete) spectrum. For instance, the free particle with energy $E = p^2/2m$, $p \in \mathcal{R}$ and $V(x) = 0$ (or any constant with respect to which the kinetic energy can be defined) is a case of such a system with unbound states forming a continuous spectrum.



Dirac Delta Potential Barrier

$V(x) = k_0 \delta(x - xc)$

$\phi(x) = e^{ikx} + Re^{-ikx}$

$\phi(x) = Te^{ikx}$

Where, $k^2 = \frac{2\,m\,E}{\hbar^2}$

Analytical Solution possible!

$H(x)\,\phi(x) = E\,\phi(x)$

Where, $H = K.E. + P.E.$

$P.E. = k_0\,\delta\,(x - xc)$

Two Boundary Conditions

1. $\phi(xc + \varepsilon) = \phi(xc - \varepsilon)$
2. $\left[\frac{d\phi(x)}{dx}\right]_{x_c - \varepsilon}^{x_c + \varepsilon} - \frac{2mk_0}{\hbar^2}\phi(xc) = 0$

Two Unknowns: R & T

Quantum Physics: S. Gasiorowicz, John Wiley & Sons. , USA, 1995.

Figure 5.2: (Left) Wavefunctions for the lowest three eigenstates of the particle in a box problem. (Right) Barrier Tunneling problem with a Dirac-Delta Function Potential. Source: The internet.

Chapter 6

Formalism I : Operators, Eigenfunctions and Eigenvalues

6.1 Operators, Eigenfunctions and Eigenvalues

6.1.1 Definition of an operator

Recall that a Function is simply a rule for taking a number and changing it to another number, e.g., $f(x) = x^2$.

An operator is a rule for changing one function into another

$$D[f(x)] \equiv \frac{df(x)}{dx}$$

$$D[3x^2] = 6x$$

$$D[\cos x] = -\sin x$$

We are interested in a special class of operators called *linear* operators (L) such that:

$$\begin{aligned} (i) \qquad L[f(x) + g(x)] &= L[f(x)] + L[g(x)] && (6.1) \\ (ii) \qquad L[cf(x)] &= cL[f(x)] \;\; , \quad c \in \mathcal{R} \text{ or } c \in \mathcal{C} \; . && (6.2) \end{aligned}$$

Note that the differential operator D satisfies the criteria for being a Linear Operator

$$D[f(x) + g(x)] = D[f(x)] + D[g(x)]$$

$$D[cf(x)] = cD[f(x)] \; ,$$

but the operator $A[f(x)+g(x)] = [f(x)+g(x)]^n$ does not

$$A[f(x)+g(x)] = [f(x)+g(x)]^n \neq f(x)^n + g(x)^n \quad \text{for } n > 1 \; .$$

6.1.2 Eigenfunctions and Eigenvalues

For a linear operator L and a function $f(x)$ such that

$$L[f(x)] = cf(x) \; , \quad c \in \mathcal{R} \text{ or } c \in \mathcal{C} \; ,$$

we call $f(x)$ an eigenfunction of L with eigenvalue c. The function $f(x)$ and the eigenvalue c will clearly depend on the form of the operator L. For instance, for the case of the differential operator D, we have for

$$D[f(x)] = cf(x) \Rightarrow f(x) = A\exp\{cx\}$$

is the general solution, where A is arbitrary constant. On the other hand, functions like $\ln x$, x^n and $\sin x$ are not eigenfunctions of the linear operator D.

In Quantum Mechanics, we postulate that linear operators associated with **all observables (i.e., experimentally measurable quantities) have real eigenvalues**. Such operators are called **Hermitian**

$$\begin{aligned} L[f(x)] &= cf(x) \;\; , \;\; c \in \mathcal{R} \;\text{ and} && (6.3)\\ \langle L \rangle &= \langle L^\dagger \rangle \; \in \mathcal{R} \; , && (6.4) \end{aligned}$$

where the † operation is defined as

$$\begin{aligned} \langle A \rangle &= \int_{-\infty}^{\infty} dx \; \psi^* A \psi \quad , \quad \langle A^\dagger \rangle = \int_{-\infty}^{\infty} dx \; \psi^* A^\dagger \psi \\ \therefore \langle A \rangle &= \langle A^\dagger \rangle \\ \Rightarrow \langle A \rangle = \int_{-\infty}^{\infty} dx \; \psi^* A \psi &= \int_{-\infty}^{\infty} dx \; \underbrace{(A\psi)^*}_{=\psi^* A^\dagger} \psi = \langle A^\dagger \rangle \; . \end{aligned}$$

In the matrix formulation of quantum mechanics, the operator A is a $N \times N$ matrix, the state vector (wavefunction) ψ is a $1 \times N$ column vector and the state vector (wavefunction) ψ^* is a $N \times 1$ row vector, such that

$$\langle A \rangle = \int_V d^3r \; (\psi_1 \psi_2 \ldots \psi_N)^* \begin{pmatrix} A_{11} & A_{12} & \ldots & A_{1N} \\ A_{21} & A_{22} & \ldots & A_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ A_{N1} & A_{N2} & \ldots & A_{NN} \end{pmatrix} \begin{pmatrix} \psi_1 \\ \psi_2 \\ \vdots \\ \psi_N \end{pmatrix} \; . \tag{6.5}$$

Then, it becomes clear that the † (i.e., Hermitian conjugation) operation involves taking both the complex conjugation *as well as* a transpose operation (i.e., such that A can be transferred from its action on a column vector to a row vector). Recall that the complex conjugation for the matrix must be done for every one of its elements.

For the special case of the row and column vector becoming a single state, say ψ_n^* and ψ_N respectively, and the operator A becoming a single element, say A_{NN}, there is now no longer any need to take the transpose action.

6.1.3 Commutators

As two linear operator $\hat{A}$ and $\hat{B}$ are not generally commutative under multiplication

$$\hat{A}\hat{B} - \hat{B}\hat{A} \neq 0$$

$$[\hat{A}, \hat{B}] = \hat{A}\hat{B} - \hat{B}\hat{A} \neq 0 \ .$$

Why does this matter? For the case that $[\hat{A}, \hat{B}] = 0$,

$$\begin{aligned} (i) \ \ \hat{A}\psi &= a\psi \ , && (6.6) \\ (ii) \ \ \hat{B}\psi &= b\psi \ . && (6.7) \end{aligned}$$

Now from (i)

$$\hat{B}(\hat{A}\psi) = \hat{B}a\psi = a\hat{B}\psi \ . \tag{6.8}$$

But since $\hat{A}\hat{B} = \hat{B}\hat{A}$, we can write the above as

$$\hat{B}(\hat{A}\psi) = \hat{A}(\hat{B}\psi) = a(\hat{B}\psi) \ , \tag{6.9}$$

i.e., $\hat{B}\psi$ is also an eigenfunction of $\hat{A}$ with eigenvalue a. This indicates that $\hat{B}\psi$ must satisfy the eigenvalue equation: $\hat{B}\psi = b\psi$.
Note: For $[\hat{A}, \hat{B}] = 0$, we see that the state ψ is an eigenstate of both $\hat{A}$ and $\hat{B}$, and we can simultaneously measure eigenvalues for both observables related to $\hat{A}$ and $\hat{B}$.

On the other hand, for $[\hat{A}, \hat{B}] \neq 0$, it is clear that the operators $\hat{A}$ and $\hat{B}$ do not share the same eigenbasis, and we cannot simultaneously measure well-defined eigenvalues for both. For instance, consider the example of the conjugate observables $\hat{p} \equiv -i\hbar\frac{d}{dx}$ and $\hat{x} \equiv x$ (the spatial variable defined on $\mathcal{R}$) acting on wavefunctions defined on a wavefunction chosen to be in the x-representation, $\psi(x)$:

$$\begin{aligned} [\hat{p}, \hat{x}]\psi(x) &= (-i\hbar\frac{d}{dx})(x\psi(x)) - x(-i\hbar\frac{d}{dx})\psi(x) \\ &= -i\hbar x(\frac{d\psi(x)}{dx}) - i\hbar\psi + i\hbar x(\frac{d\psi(x)}{dx}) \\ &= -i\hbar\psi \\ \Rightarrow [\hat{p}, \hat{x}] &= -i\hbar \quad \text{(dropping the } \psi \text{ from both sides)} \ , && (6.10) \end{aligned}$$

i.e., confirming the fact $\hat{p}$ and $\hat{x}$ do not commute with one another. Note that you could have chosen to carry out this calculation with $\hat{p} \equiv p$ (the momentum variable defined on $\mathcal{R}$), $\hat{x} \equiv i\hbar\frac{d}{dp}$ and wavefunctions in the p-representation, $\psi(p)$. You will again obtain precisely the same result (check this for yourself!). This shows the representation independence of the relation $[\hat{p}, \hat{x}] = -i\hbar$.

This relation tells us that $\hat{p}$ and $\hat{x}$ do not share the same eigenbasis, and we cannot simultaneously measure well-defined eigenvalues for both $\hat{p}$ and $\hat{x}$. This is basically a restatement of the Heisenberg uncertainty relation for $\hat{p}$ and $\hat{x}$. Further, the appearance of the $i\hbar$ is actually signalling the wave-particle duality at the heart of quantum mechanics (we are playing with matter waves after all!). As you will learn in advanced courses up ahead, the $i\hbar$ factor is also an indication of the fact that the quantum state (or Hilbert) space can be seen as the quantisation of the associated phase space for an equivalent classical mechanical system in terms of sub-blocks of side-length $\hbar$. If you don't know what I mean at this point, don't worry about it!

Further, the above commutator relation between conjugate operators such as $\hat{p}$ and $\hat{x}$ is itself a special case of a more generalised relation: $[\hat{\mathcal{O}}_1, \hat{\mathcal{O}}_2] \propto \hat{\mathcal{O}}_3$, i.e., the commutation relation between two operators will typically lead to a third operator. We will encounter an example of this when we study the orbital angular momentum problem (in terms of the various components (L_x, L_y, L_z) of the orbital angular momentum operator).

Finally, a note on **symmetries**. the fact that an operator $\hat{\theta}$ commutes with the Hamiltonian operator $\hat{H}$, $[\hat{H}, \hat{\theta}] = 0$, implies that the corresponding physical observable satisfies a conservation law (i.e., an invariance in time) and reflects on a corresponding symmetry of the quantum system. Further, this symmetry (and its corresponding conservation law) will be preserved in time. This can be seen simply from the fact that since the unitary time evolution operator is given by $U = e^{-\frac{i\hat{H}t}{\hbar}}$ (i.e., U is a function of the Hamiltonian $\hat{H}$),

$$[\hat{H}, \hat{\theta}] = 0 \implies [U, \hat{\theta}] = 0 \ . \tag{6.11}$$

Let us consider a couple of simple examples to clarify these point.

(a) For a free particle on the 1D line, $\hat{H} = \hat{p}^2/2m$ and $[\hat{H}, \hat{p}] = 0$, implying that the linear momentum $\hat{p}$ is conserved in time and its eigenvalue is a "good" quantum number. This reflects on the fact that $\hat{H}$ and $\hat{p}$ have the same basis, and that there is a continuous translation symmetry in the problem, such that for the eigenfunction $\psi(x) = \mathcal{N} e^{-i\hat{p}x}$ (and where $\mathcal{N}$ is the normalisation factor), we have

$$T_{\delta x}\psi(x) = e^{-i\hat{p}\delta x/\hbar}\psi(x) = \psi(x + \delta x) \ , \tag{6.12}$$

where $T_{\delta x} = e^{-i\hat{p}\delta x/\hbar}$ is the operator for a translation in real space along the 1D line by an amount δx. Thus, the linear momentum $\hat{p}$ is the generator of infinitesmal translation δx.

(b) For a quantum system in which the (say, 1D) potential has inversion symmetry $V(\hat{x}) = V(-\hat{x})$, the parity operation $\hat{P}$ is such that its action on any eigenfunction ψ gives

$$\begin{aligned} \hat{P}\psi_0(x) &= \pm\psi_0(-x) \\ \text{This arises from the fact that } \hat{P}^2\psi_0(x) &= \psi_0(x) \\ \Rightarrow \psi_0(-x) &= \pm\psi_0(x) \ , \end{aligned} \tag{6.13}$$

i.e., eigenvalues of $\hat{P}$ are ± 1, such that all eigenstates have either even (+1) or odd (-1) parity eigenvalue. For such a parity symmetric system, $[\hat{H}, \hat{P}] = 0$. Consider the case of the wavefunctions of the particle in a box problem shown below. Can you think of any other well known potential in 1D that possesses parity symmetry?

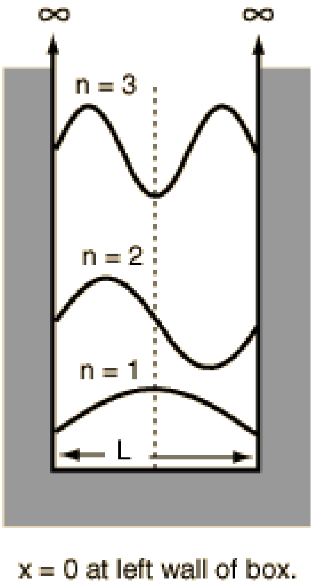


Figure 6.1: Wavefunctions for the lowest three eigenstates of the particle in a box problem. Source: The internet.

6.2 Important properties of Eigenstates

(1) $\boxed{\text{Eigenvalues of Hermitian Operators are real-valued.}}$
Recall that this is important because Hermitian operators are associated with physical observables/ experimentally measurable quantities. Take a Hermitian Operator $\hat{A}$ such that $\hat{A} = \hat{A}^\dagger = (\hat{A}^*)^T$, where $*$ and T refer, as discussed earlier, to complex conjugation and transpose operations respectively. Let

$$\hat{A}\psi_n = a_n\psi_n \quad , \quad n = 1, 2, 3, \ldots \tag{6.14}$$

Now, by computing the expectation value

$$\begin{aligned}
\langle\hat{A}\rangle &= \int_{\infty}^{-\infty} dx\ \psi_n^*\hat{A}\psi_n = \int_{\infty}^{-\infty} dx\ \underbrace{(\hat{A}\psi_n)^*}_{=\psi_n^*\hat{A}^\dagger}\psi_n \\
&= \int_{-\infty}^{\infty} dx\ \psi_n^* a_n\psi_n = \int_{-\infty}^{\infty} dx (a_n\psi_n)^*\psi_n \\
&= a_n\int_{-\infty}^{\infty} dx\ \psi_n^*\psi_n = a_n^*\int_{-\infty}^{\infty} dx\ \psi_n^*\psi_n \\
\langle\hat{A}\rangle &= a_n = a_n^* \\
\text{or, } a_N &\in \mathcal{R} \quad \text{(QED)} \qquad\qquad \tag{6.15}
\end{aligned}$$

(2) $\boxed{\text{Different eigenstates of the same potential are orthogonal.}}$

Note that the **Inner Product** is defined as

$$\int_{-\infty}^{\infty} dx\ \psi_n^* \psi_m = \delta_{n,m} = \begin{cases} 1, & n = m \\ 0, & \text{otherwise} \end{cases} \tag{6.16}$$

For a Hermitian operator $\hat{A}$ with eigenstates ψ_n and corresponding eigenvalues a_n, such that $\hat{A}\psi_n = a_n\psi_n$ and $\hat{A}\psi_m = a_m\psi_m$. Now, since $\hat{A} = \hat{A}^\dagger$,

$$\begin{aligned} \int_{-\infty}^{\infty} dx\ \psi_m^* \hat{A} \psi_n &= \int_{-\infty}^{\infty} dx\ \underbrace{(\hat{A}\psi_m)^*}_{=\psi_n^* \hat{A}^\dagger} \psi_n \\ \int_{-\infty}^{\infty} dx\ \psi_m^* a_n \psi_n &= \int_{-\infty}^{\infty} dx\ (a_m\psi_m)^* \psi_n \\ a_n \int_{-\infty}^{\infty} dx\ \psi_m^* \psi_n &= a_m^* \int_{-\infty}^{\infty} dx\ \psi_m^* \psi_n = a_m \int_{-\infty}^{\infty} dx\ \psi_m^* \psi_n \quad (\text{as } \hat{A} \text{ is Hermitian}) \\ (a_n - a_m) \int_{-\infty}^{\infty} dx\ \psi_m^* \psi_n &= 0\ . \end{aligned} \tag{6.17}$$

Now, for $a_n \neq a_m$ for $n \neq m$, $\int_{-\infty}^{\infty} dx\ \psi_m^* \psi_n = 0$.
On the other hand, for $n = m$, $\int_{-\infty}^{\infty} dx\ \psi_n^* \psi_n = 1 = \int_{-\infty}^{\infty} dx\ \psi_m^* \psi_m$. (QED)
What about the case of degenerate eigenstates? The above relation cannot distinguish between them, and we need to find the appropriate linear combinations of such degenerate eigenstates that are orthogonal to one another.

(3) $\boxed{\text{Postulate of quantum mechanics: eigenstates of the TISE form a complete set of states}}$, i.e, any other function $f(x)$ can be expressed as a linear combination of the eigenstates ψ_n

$$f(x) = \sum_{n=1} c_n \psi_n(x)\ . \tag{6.18}$$

The states $\psi_n(x)$ are then said to "span the vector space" or "form a suitable basis set".

The superposition principle of quantum mechanics says that such linear combinations are valid solutions of the time-dependent Schrodinger's equation as well

$$\Psi(x,t) = \sum_n C_n \psi_n(x) \exp\left\{\frac{-iE_n t}{\hbar}\right\} = \sum_n C_n \phi_n(x,t)\ ,$$

where $\phi_n(x,t)$ forms a complete basis of eigenstates.

Note that for such a linear combination of eigenstates with different E_n ,

$$H\Psi \neq E\Psi\ \ .$$

What does the superposition mean physically?

It means that QM allows for the possibility that a particle in a given potential can be simultaneously in a mixture of different eigenstates (with different energy eigenvalues).

In order to make sense of any measurement of a physically observable quantity, the measurement process is often referred to as the collapse of the wavefunction $\Psi(x,t)$ into one of its constituent eigenstates $\phi_n(x,t)$. Further, the weight factors C_n refer to the probabilities that the measurement will lead to the eigenvalues of particular eigenstates ϕ_n. (At this point, recall our discussion of the Schrödinger cat gedanken, or the Feynman double slit experiment with electrons!)

Finally, if one knows the solution to the eigenvalue problem $H\psi_n = E_n\psi_n$, the superposition provides a means by which to determine the time dependence of any wavefunction Ψ. For this, we assume that at $t=0$, $\Psi(x,t=0) = \sum_n C_n\psi_n(x)$; then, we need to determine the coefficients C_n

$$\int_{-\infty}^{\infty} dx\ \psi_m^*\Psi(x) = \sum_n C_n \int_{-\infty}^{\infty} dx\ \psi_m^*\psi_n = \sum_n C_n\delta_{n,m} = C_m\ , \tag{6.19}$$

where we have used the orthogonality property of eigenstates we discussed earlier. Therefore, the coefficient $C_m = \int_{-\infty}^{\infty} dx\ \psi_m^*\Psi(x)$ gives the weight factors. This relation can be seen as implementing the "collapse" of the wavefunction in terms of a measurement of the coefficient C_m, and appears to be projective in the sense that it picks out one member of the entire Hilbert space from the rest. Such projections are typically non-unitary, i.e., they do not preserve the total probability. Nevertheless, this is one way in which to learn the C_m coefficients. It is not clear how to build in such processes involving measurement and wavefunction collapse within the Schrödinger equation formalism. We will discuss this in more detail towards the end of these lectures.

To continue, with the coefficients C_n having been computed, we can always obtain the full time-evolved state as

$$\Psi(x,t) = \sum_n C_n\psi_n(x)e^{-iE_nt/\hbar}\ . \tag{6.20}$$

It is safe to say that it is this "linear superposition" property of the wavefunction in the quantum world that scientists see as the power of a quantum computer. In some sense, such a computer will have all the answers (in a linear superposition) to a particular question you may ask of it (an oracle?) ... when you finally want the answer, a measurement causes the computer's wavefunction to collapse onto your answer! Interesting, isn't it? In this sense, the Schrödinger cat gedanken and the Feynman double slit experiment with electrons are simple quantum computation setups in which the cat and the electron can have only two possibilites that can exist in a linear superposition.

Similarly, the ammonia molecule lives in a linear superposition of the two positions of the Nitrogen atom (above and below the plane of three Hydrogen atoms), the Benzene molecule lives in a superposition (called the "resonance") of the two configurations for the three alternating double bond structures. They can be thought of as molecular quantum computers ... but can we do some useful computation with them (or an array of such molecules)?

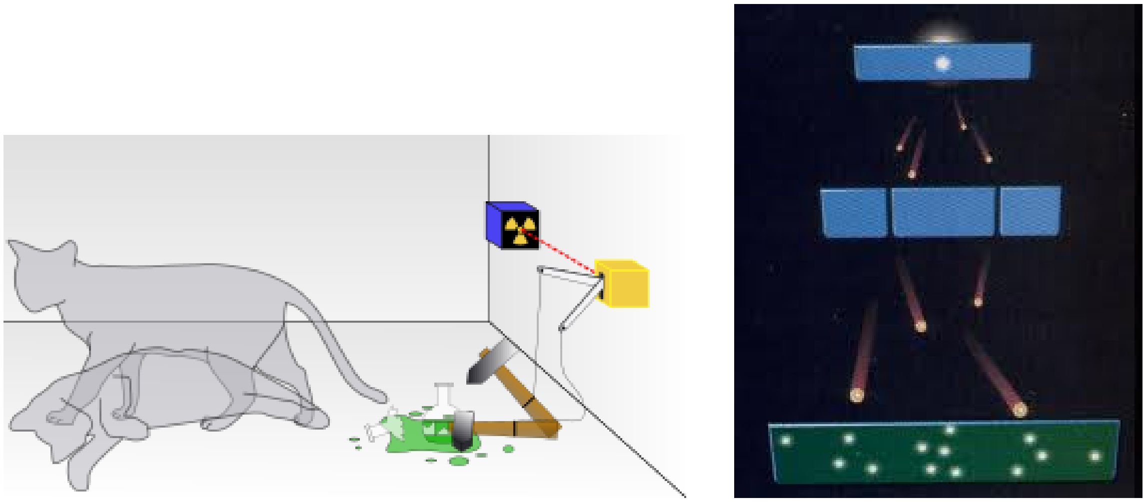


Figure 6.2: (Left) The Schrödinger Cat gedanken. (Right) Double-slit Interference Experiment with Matter particles. Source: The internet.

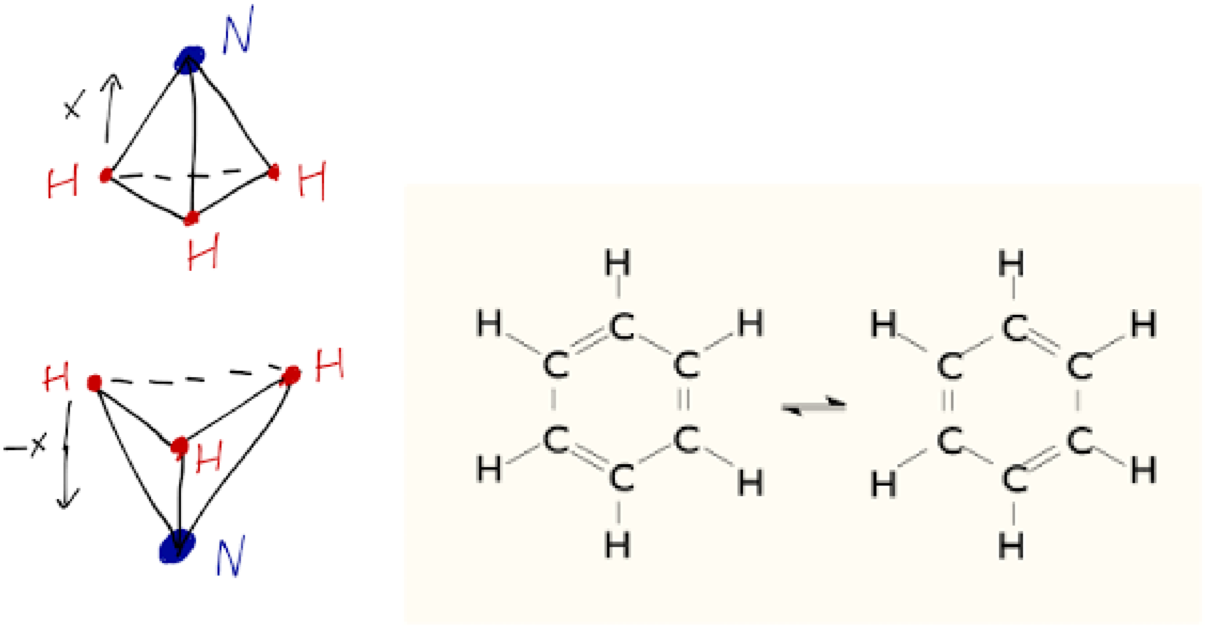


Figure 6.3: (Left) The two configurations of the Ammonia molecule. (Right) The two configurations of the Benzene molecule. Source: The internet.

Chapter 7

Schrödinger Wave Mechanics: Bound states

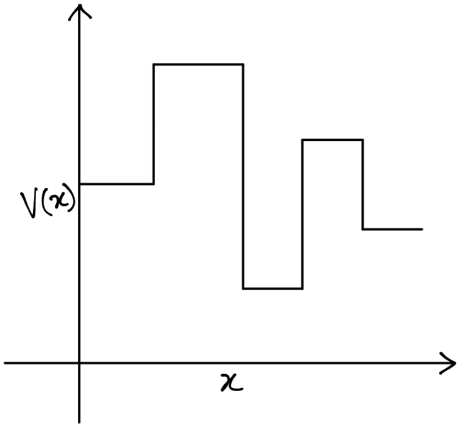


Figure 7.1: Schematic diagram for a piecewise constant potential in 1D.

For the case of piecewise constant potentials (i.e., potentials with step-like discontinuities as shown in figure 7.1), we consider a range of x over which the potential $V(x)$ is a constant then the TISE can be rewritten as

$$-\frac{\hbar^2}{2m}\frac{d^2\psi}{dx^2}+V_0\psi=E\psi \tag{7.1}$$

$$\implies \frac{d^2\psi}{dx^2}+\frac{2m}{\hbar^2}\left(E-V_0\right)\psi=0 \tag{7.2}$$

Trying a guess (or *ansatz*) solution $\psi(x) = Ce^{nx}$ in Eq. 7.2, we get

$$\begin{aligned} n^2 Ce^{nx} + \frac{2m}{\hbar^2}(E - V_0)\, Ce^{nx} &= 0 \\ n^2 &= \frac{2m}{\hbar^2}(V_0 - E) \\ \therefore n &= \pm\sqrt{\frac{2m}{\hbar^2}(V_0 - E)} \qquad (7.3) \end{aligned}$$

For $E < V_0$, $n \in \mathbb{R}$ (a classically forbidden solution!),

$$\psi(x) = C_1 e^{nx} + C_2 e^{-nx} \qquad (7.4)$$

For $E > V_0$, $n = \pm i\bar{n}$ where $\bar{n} \in \mathbb{R}$ (a classically permitted solution!) and

$$\begin{aligned} \psi(x) &= C_3 e^{i\bar{n}x} + C_4 e^{-i\bar{n}x} \qquad (7.5) \\ &= D_1 \cos(\bar{n}x) + D_2 \sin(\bar{n}x) \qquad (7.6) \end{aligned}$$

Here, C_1, C_2, C_3, C_4, D_1, D_2 are constants that are determined from boundary conditions appropriate to the problem at hand.

For the full wavefunction over several such piecewise constant potentials, we must "stitch together" the ψs obtained from each section by demanding continuity in ψ and $d\psi/dx$, i.e., matching the wavefunction and its spatial derivative at various interfaces where the potential jumps. The continuity is ψ leads to the continuity in probability density, while that for $d\psi/dx$ leads to the continuity in probability current. This will give us the complete wavefunction for all x. It also tells us that ψ can be non-zero even in regions of x that are classically forbidden, e.g., inside a potential barrier. As we will see later, this is responsible for the quantum mechanical phenomenon of tunneling.

Chapter 8

Infinite square well/ Particle in a Box

Consider a potential as shown in figure 8.1.

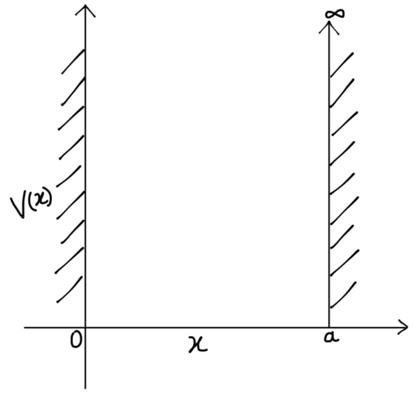


Figure 8.1: Schematic Diagram of the potential for the particle in a 1D Box.

$$V(x) = \begin{cases} 0 & \text{if } 0 < x < a \\ \infty & \text{otherwise} \end{cases} \tag{8.1}$$

The hard walls at $r = 0$ make the probability of finding the particle outside well to be zero i.e., $\psi(x) = 0$.

Inside the well, i.e., $0 < x < a$,

$$\frac{d^2\psi}{dx^2} = -k^2\psi\ , \quad k = \frac{\sqrt{2mE}}{\hbar} \quad (E > 0) \tag{8.2}$$

Here, $E < 0$ does not give normalizable ψ. The general solutions are

$$\psi(x) = A\sin(kx) + B\cos(kx) \tag{8.3}$$

Boundary Conditions

In general $\psi(x)$ & $\frac{d\psi(x)}{dx}$ must be continuous but where $V \to \infty$, only the first applies. Therefore,

$$\psi(0) = 0 = \psi(a) \ . \tag{8.4}$$

Using these two boundary conditions gives

$$\begin{aligned} \psi(0) = 0 \implies A\sin(0) + B\cos(0) &= 0 \\ B &= 0 \\ \therefore \psi(x) &= A\sin(kx) \qquad (8.5) \\ \psi(a) = 0 \implies A\sin(ka) &= 0 \\ \therefore k_n &= \frac{n\pi}{a}, \quad n \in \mathbb{Z} \qquad (8.6) \end{aligned}$$

From Eq. 8.2,

$$\therefore E_n = \frac{\hbar^2 k_n^2}{2m} = \frac{n^2\pi^2\hbar^2}{2ma^2}, \quad n \in \mathbb{Z} = 1, 2, 3, \dots\dots \tag{8.7}$$

- E_n is discrete i.e., levels exist only for special values that are quantised in units of $\pi^2\hbar^2/2ma^2$

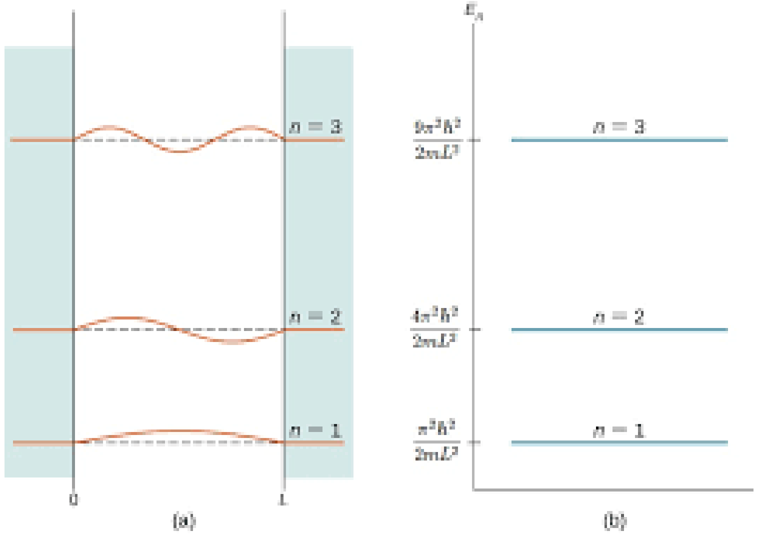


Figure 8.2: Wavefunctions (Left) and Energy Spectrum (Right) for the lowest three eigenstates of the particle in a box problem. Source: the internet.

Normalization

The normalizing condition given by Eq. 8.8 gives

$$\int_{-\infty}^{\infty} |\psi_n|^2 \, dx = 1 \tag{8.8}$$

$$\implies |A|^2 \int_0^a dx \;\sin^2 (k_n x) = 1$$

$$\therefore A = \sqrt{\frac{2}{a}}$$

$$\psi_n (x) = \sqrt{\frac{2}{a}} \sin \left(\frac{n\pi x}{a}\right) \tag{8.9}$$

Discussion

An important point to note about this problem is the fact that the ground state energy is non-zero (while classically it should have been zero)! This, so called, **zero-point energy** arises from the fact that the particle is not sitting still in the quantum mechanical ground state. This can be seen from the fact that the ground state wavefunction, even if peaked at the centre of the well (as expected classically) is actually spread out all over the well. This "spreading out" of matter waves is typical of waves, who (unlike particles) hate being confined. Indeed, the ground state energy can be shown to arise completely from the Heisenberg uncertainty principle (and is left as an exercise to the interested reader).

Chapter 9

Simple Harmonic Oscillator

A spring block system is a classic oscillator problem in which force is linearly proportional to displacement. Note, however, that there is no such thing as a perfect oscillator and especially when applying forces beyond Hooke's law. However, for small forces applied such that the amplitude of motion is small, the motion can be taken to be harmonic. Put differently, about any local minima of any complicated potential, the parabolic approximation is decent as shown in figure 9.1

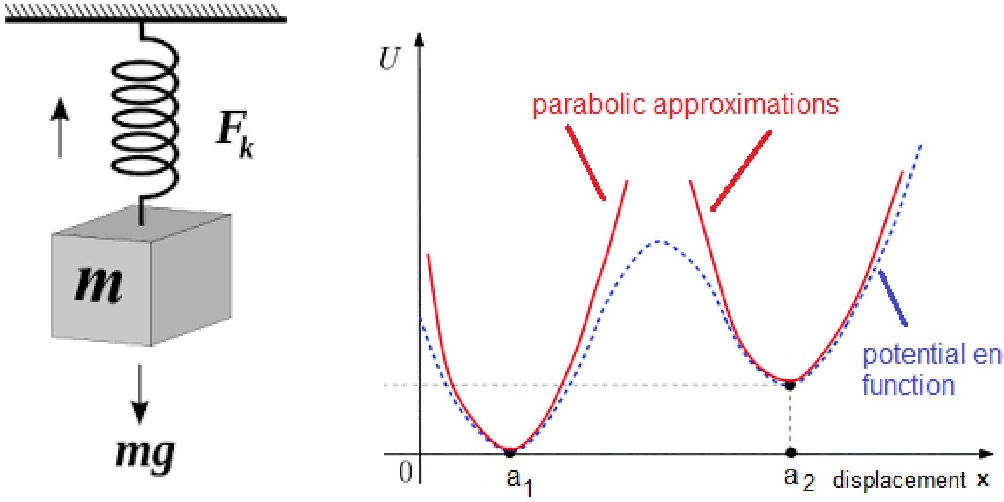


Figure 9.1: (Left) A classical mass spring oscillator system. (Right) Parabolic approximation at a minima

Let x_0 be the local minima of a potential $V(x)$ (i.e., $V'(x_0) = 0$, $V''(x_0) > 0$). Expanding $V(x)$ using taylor expansion about x_0 and looking at the close neighbourhood of x_0 (i.e.,

$|x-x_0| << 1$), we get

$$V(x) = V(x_0) + \underbrace{V'}_{=0}(x_0)(x-x_0) + \frac{V''(x_0)}{2!}(x-x_0)^2 + \mathscr{O}\left((x-x_0)^3\right)$$
$$\approx V(x_0) + \frac{1}{2}V''(x_0)(x-x_0)^2 \tag{9.1}$$

By shifting our axes so that $V(x_0) = 0$, $x_0 = 0$ and taking $V''(x_0) = m\omega^2$ (the spring stiffness), we see that $V(x)$ in the close neighbourhood of x_0 behaves like a simple harmonic oscillator as seen in Eq. 9.2

$$V(x) = \frac{1}{2}m\omega^2 x^2 \tag{9.2}$$

The time independent Schrödinger's equation can then be written as

$$-\frac{\hbar^2}{2m}\frac{d^2\psi(x)}{dx^2} + \frac{1}{2}m\omega^2x^2\psi(x) = E\psi(x) \tag{9.3}$$

9.1 Algebraic Method

Eq. 9.3 could be written as

$$\frac{1}{2m}\left[p^2 + (m\omega x)^2\right]\psi = E\psi \tag{9.4}$$

Inspired by the following relation between two classical variables $(u,v) \in \mathcal{C}$

$$\left(u^2+v^2\right) = (iu+v)(-iu+v)\ , \tag{9.5}$$

we rewrite

$$p^2 + (m\omega x)^2 = (ip+m\omega x)(-ip+m\omega x) + im\omega\underbrace{[x,p]}_{\neq 0}\ .$$

Raising & lowering operators

Defining new operators a_+, a_-

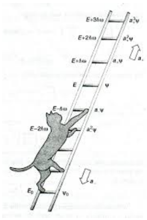
$$a_\pm = \frac{1}{\sqrt{2m\omega\hbar}}(\mp ip + m\omega x) \tag{9.6}$$

$$[a_+, a_-] = \frac{1}{2m\omega\hbar}\left[(-ip+m\omega x)(ip+m\omega x) - (ip+m\omega x)(-ip+m\omega x)\right]$$
$$= \frac{1}{2m\omega\hbar}\left[p^2 - p^2 + m\omega^2 x(1-1) + 2im\omega(xp-px)\right]$$
$$= \frac{2im\omega}{2m\omega\hbar}[x,p]$$
$$= -1 \tag{9.7}$$
$$\&\ [a_-, a_+] = 1 \tag{9.8}$$

Here, $a_\pm$ don't commute with each other as x, p don't commute. The Hamiltonian H can be written in terms of $a_\pm$ as given by Eqs. 9.10, 9.11.

$$\begin{aligned}
a_- a_+ &= \frac{1}{2m\hbar\omega}\left(ip + m\omega x\right)\left(-ip + m\omega x\right) \\
&= \frac{1}{2m\hbar\omega}\left(p^2 + m^2\omega^2 x^2 - im\omega\left[x,p\right]\right) \\
&= \frac{1}{2\hbar m\omega}\left(p^2 + m^2\omega^2 x^2\right) - \frac{i}{2\hbar}\left[x,p\right] \\
&= \frac{1}{\hbar\omega}H + \frac{1}{2} \qquad (9.9) \\
\therefore H &= \hbar\omega\left(a_- a_+ - \frac{1}{2}\right) \qquad (9.10) \\
&= \hbar\omega\left(a_+ a_- + \frac{1}{2}\right) \qquad \text{using } \left[a_-, a_+\right] = 1\ . \qquad (9.11)
\end{aligned}$$

If ψ is an eigenstate of H with eigenvalue E, then $a_+\psi$ is also an eigenstate with eigenvalue $(E + \hbar\omega)$ and $a_-\psi$ is an eigenstate with eigenvalue $(E - \hbar\omega)$.



SHO spectrum.

Consider

$$\begin{aligned}
H\left(a_+\psi\right) &= \hbar\omega\left(a_+ a_- + \frac{1}{2}\right)\left(a_+\psi\right) \\
&= \hbar\omega\left[a_+ a_- a_+ + \frac{1}{2}a_+\right]\psi \\
&= \hbar\omega a_+\left[a_- a_+ + \frac{1}{2}\right]\psi \\
&= \hbar\omega a_+\left[a_- a_+ - \frac{1}{2} + 1\right]\psi \\
&= a_+\left(H + \hbar\omega\right)\psi \\
&= \left(E + \hbar\omega\right)\left(a_+\psi\right) \qquad (9.12)
\end{aligned}$$

Similarly,

$$\begin{aligned}
H\left(a_-\psi\right) &= \hbar\omega\left(a_- a_+ - \frac{1}{2}\right)\left(a_-\psi\right) \\
&= \hbar\omega\left[a_- a_+ a_- - \frac{1}{2}a_-\right]\psi \\
&= \hbar\omega a_-\left[a_+ a_- - \frac{1}{2}\right]\psi \\
&= \hbar\omega a_-\left[a_+ a_- + \frac{1}{2} - 1\right]\psi \\
&= a_-\left(H - \hbar\omega\right)\psi \\
&= \left(E - \hbar\omega\right)\left(a_-\psi\right) \qquad (9.13)
\end{aligned}$$

From Eqs. 9.12, 9.13, a_+ is called a **raising operator** while a_- is called a **lowering operator**.

Note that

$$a_+^\dagger = a_- \quad , \quad a_-^\dagger \;= a_+ \tag{9.14}$$

Thus, even though the operators $a_\pm$ are not Hermitian (and therefore cannot be associated with experimental observables), we have seen that the Hamiltonian can be written in terms of the product a_+a_- (or a_-a_+).

Ground state

As the classical global minimum energy must be zero (bottom of the potential well) for the stability of such a confining (parabolic) potential, in order to get a normalizable ("legal") $\psi(x)$, there must be a eigenstate $\psi_0(x)$ with lowest possible energy. As there are no other states below this state, it should vanish upon the action of lowering operator a_-. Thus, we impose the condition

$$\begin{aligned} a_-\psi_0(x) &= 0 \qquad (9.15)\\ \implies \frac{1}{\sqrt{2\hbar m\omega}}\left(\hbar\frac{d}{dx} + m\omega x\right)\psi_0(x) &= 0\\ \frac{d\psi_0}{dx} &= -\frac{m\omega}{\hbar}x\psi_0\\ \int\frac{d\psi_0}{\psi_0} &= -\frac{m\omega}{\hbar}\int xdx\\ \ln\psi_0 &= -\frac{m\omega}{2\hbar}x^2 + c\\ \psi_0 = C_1 e^{-m\omega x^2/2\hbar} \;, &\quad C_1 = \ln c \;. \qquad (9.16) \end{aligned}$$

Here, C_1 can be calculated using the normalizing condition

$$\begin{aligned} \int_{-\infty}^{\infty} dx \;\; |\psi_0|^2 &= 1\\ \implies |C_1|^2\int_{-\infty}^{\infty} e^{-m\omega x^2/\hbar} &= 1\\ \Rightarrow |C_1|^2\sqrt{\frac{\pi\hbar}{m\omega}} &= 1 \quad (\text{using } \int_{-\infty}^{\infty} dx e^{-x^2/a^2} = \sqrt{\frac{\pi}{a}}) \;,\\ \therefore C_1 &= \left(\frac{m\omega}{\pi\hbar}\right)^{1/4} \qquad (9.17)\\ \implies \psi_0 &= \left(\frac{m\omega}{\pi\hbar}\right)^{1/4} e^{-m\omega x^2/2\hbar} \qquad (9.18) \end{aligned}$$

ψ_0 thus takes the shape of a Gaussian function in the spatial variable x, $\psi \propto e^{-ax^2}$. Let E_0

be the energy of ψ_0,

$$H\psi_0 = E_0\psi_0 \tag{9.19}$$

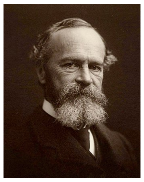
$$\hbar\omega\left(a_+a_- + \frac{1}{2}\right)\psi_0 = E_0\psi_0$$

$$\hbar\omega a_+\left(a_-\psi_0\right) + \frac{1}{2}\hbar\omega\psi_0 = E_0\psi_0$$

$$\implies \frac{1}{2}\hbar\omega\psi_0 = E_0\psi_0 \quad (\text{using} \;\; a_-\psi_0 = 0)$$

$$\therefore E_0 = \frac{1}{2}\hbar\omega \tag{9.20}$$

Here, E_0 is called as the **zero-point energy**. An important point to note about this problem is the fact that the ground state energy is non-zero (while classically it should have been zero)! In common with the particle in a box problem, this zero-point energy too arises from the fact that the particle is not sitting still in the quantum mechanical ground state. This can be seen from the fact that the ground state wavefunction, even if peaked at the centre of the well (as expected classically) is actually spread out all over the well. This "spreading out" of matter waves is typical of waves, who (unlike particles) hate being confined. Indeed, the ground state energy can be shown to arise completely from the Heisenberg uncertainty principle (and is left as an exercise to the interested reader). Next, we will see that applying a_+ on ψ_0 recursively will give us other eigenstates with higher energy.



James

Turtles all the way down!

A small humorous diversion on the stability of the system. Here is an anecdote that tells us why this concept is important. The philosopher and psychologist William James (1842-1910) was once giving a seminar on cosmology and the solar system, and was confronted by an elderly woman saying "Everything you've said is nonsense. It's well known that the Earth is held up on the back of seven elephants, themself standing on the back of a giant turtle swimming through space!" Perplexed, James countered "But what, Madam, is the turtle standing on?" Without a thought, the old lady shot back, "Clever question, Professor James. The answer is simple: it's turtles all the way down!" Perhaps the same could have been said of our search for the ground state of the harmonic oscillator in the absence of the condition imposed in eq.(9.15).

Figure 9.2: Another view of the universe. Source: the internet.

The energy spectrum and excited states

Obtaining the energy spectrum is actually quite straightforward now. Given that we now have the ground state ψ_0 with energy eigenvalue $\frac{1}{2}\hbar\omega$, we can obtain an infinite tower of excited states by acting repeatedly on ψ_0 with the raising operator a_+. From what we learnt earlier, the first excited eigenstate's energy eigenvalue will be $\hbar\omega$ greater than that of the ground state, i.e., $E_1 = \frac{3}{2}\hbar\omega$, $E_2 = \frac{5}{2}\hbar\omega$ and so on. We will now see this through an explicit calculation.

Thus, let us work out the energy eigenvalue of the first excited state $n = 1$ using $\psi_1 = A_1 a_+ \psi_0$ (where A_1 is a normalisation constant)

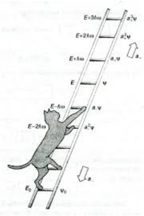
$$\begin{aligned}
H\psi_1 &= (a_+a_- + \frac{1}{2})\hbar\omega \; A_1 a_+ \psi_0 \\
E_1\psi_1 &= A_1\hbar\omega(a_+a_-a_+)\psi_0 + \frac{1}{2}\hbar\omega\psi_1 \\
&= A_1\hbar\omega a_+(1 + a_+a_-)\psi_0 + \frac{1}{2}\hbar\omega\psi_1 \qquad (\text{using } [a_-, a_+] = 1) \;, \\
&= A_1\hbar\omega a_+\psi_0 + \frac{1}{2}\hbar\omega\psi_1 \qquad (\text{as } a_-\psi_0 = 0) \;, \\
&= (1 + \frac{1}{2})\hbar\omega A_1 a_+\psi_0 \\
&= \frac{3}{2}\hbar\omega\psi_1 \\
\therefore E_1 &= \frac{3}{2}\hbar\omega \;. \qquad (9.21)
\end{aligned}$$

The entire spectrum can now be obtained by proceeding with the same strategy in an iterative fashion. To be precise, now that we know ψ_1 and its eigenvalue $E_1 = 3\hbar\omega/2$, we can now define $\psi_2 = A_2 a_+ \psi_1$ and proceed exactly as given above. This will show us that $E_2 = (2 + 1/2)\hbar\omega = 5\hbar\omega/2$:

Turtles all the way down!

$$\begin{aligned}
H\psi_2 &= (a_+a_- + \frac{1}{2})\hbar\omega \; A_2 a_+ \psi_1 \\
E_2\psi_2 &= A_2\hbar\omega(a_+a_-a_+)\psi_1 + \frac{1}{2}\hbar\omega\psi_2 \\
&= A_2\hbar\omega a_+(1 + a_+a_-)\psi_1 + \frac{1}{2}\hbar\omega\psi_2 \qquad (\text{using } [a_-, a_+] = 1) \;, \\
&= A_2\hbar\omega a_+(1 + 1)\psi_1 + \frac{1}{2}\hbar\omega\psi_2 \qquad (\text{as } a_+a_-\psi_1 = (E_1 - \frac{\hbar\omega}{2})\psi_1 = \hbar\omega\psi_1) \;, \\
&= (2 + \frac{1}{2})\hbar\omega A_2 a_+\psi_1 \\
&= \frac{5}{2}\hbar\omega\psi_2 \\
\therefore E_2 &= \frac{5}{2}\hbar\omega \;. \qquad (9.22)
\end{aligned}$$

Continuing in this way till the nth eigenstate ψ_n, using $\psi_n = A_n a_+ \psi_{n-1}$ and $E_{n-1} = (n - \frac{\hbar\omega}{2})$, we find $E_n = (n+1/2)\hbar\omega$:



SHO spectrum.

$$
\begin{aligned}
H\psi_n &= (a_+a_- + \frac{1}{2})\hbar\omega\ A_n a_+\psi_{n-1} \\
E_n\psi_n &= A_n\hbar\omega(a_+a_-a_+)\psi_{n-1} + \frac{1}{2}\hbar\omega\psi_n \\
&= A_2\hbar\omega a_+(1+a_+a_-)\psi_{n-1} + \frac{1}{2}\hbar\omega\psi_n \qquad (\text{using } [a_-, a_+] = 1)\ , \\
&= A_2\hbar\omega a_+(n-1+1)\psi_{n-1} + \frac{1}{2}\hbar\omega\psi_n \qquad (\text{as } a_+a_-\psi_{n-1} = (E_{n-1} - \frac{\hbar\omega}{2})\psi_{n-1} = (n-1)\hbar\omega\psi_{n-1})\ , \\
&= (n+\frac{1}{2})\hbar\omega A_n a_+\psi_{n-1} \\
&= (n+\frac{1}{2})\hbar\omega\psi_n \\
\therefore E_n &= (n+\frac{1}{2})\hbar\omega\ . \qquad (9.23)
\end{aligned}
$$

In this way, we find the entire spectrum for $n \in \mathcal{Z}$ and $n \geq 0$.

Now, the form of $\psi_n = A_n(a_+)^n\psi_0$ is exact in principle (while getting the exact form of $(a_+)^n\psi_0$ is quite tedious!), and all we need to do is to compute the normalisation constant A_n to be able to say that we have computed all the wavefunctions precisely. However, calculating A_n directly by performing the action of a_+ on ψ_0 n times is far too tedious. Instead, we compute A_n in a different way. Let us take

$$a_+\psi_n = c_n\psi_{n+1}\ , \qquad (9.24a)$$

$$a_-\psi_n = d_n\psi_{n-1}\ . \qquad (9.24b)$$

Consider,

$$
\begin{aligned}
\int_{-\infty}^{\infty} (a_+\psi_n)^*\,(a_+\psi_n)\,dx &= \int_{-\infty}^{\infty} \psi_n^* a_- a_+\psi_n dx \qquad (\text{as } (a_+\psi_n)^* = \psi_n^* a_-, a_+^\dagger = a_-)\ , \\
|c_n|^2 \int_{-\infty}^{\infty} \psi_{n+1}^*\psi_{n+1} dx &= \int_{-\infty}^{\infty} \psi_n^* \left(\frac{1}{\hbar\omega}H + \frac{1}{2}\right)\psi_n dx \qquad (\text{using } H = (a_-a_+ - 1/2)\hbar\omega)\ , \\
|c_n|^2 &= \frac{1}{\hbar\omega}\int_{-\infty}^{\infty} \psi_n^* H\psi_n dx + \frac{1}{2}\int_{-\infty}^{\infty} \psi_n^*\psi_n dx \\
&= \left(n+\frac{1}{2}\right) + \frac{1}{2} \qquad (\text{using } \int_{-\infty}^{\infty} \psi_n^*\psi_n dx = 1)\ , \\
\Rightarrow |c_n|^2 &= n+1 \\
\therefore c_n &= \sqrt{n+1} \qquad (9.25)
\end{aligned}
$$

Similarly consider,

$$\int_{-\infty}^{\infty} (a_-\psi_n)^* (a_-\psi_n)\, dx = \int_{-\infty}^{\infty} \psi_n^* a_+ a_- \psi_n dx \qquad (\text{as } (a_-\psi_n)^* = \psi_n^* a_+, a_+^\dagger = a_-)\ ,$$

$$|d_n|^2 \int_{-\infty}^{\infty} \psi_{n-1}^* \psi_{n-1} dx = \int_{-\infty}^{\infty} \psi_n^* \left(\frac{1}{\hbar\omega}H - \frac{1}{2}\right) \psi_n dx \qquad (\text{using } H = (a_+a_- + 1/2)\hbar\omega)\ ,$$

$$|d_n|^2 = \frac{1}{\hbar\omega} \int_{-\infty}^{\infty} \psi_n^* H \psi_n dx - \frac{1}{2} \int_{-\infty}^{\infty} \psi_n^* \psi_n dx$$

$$= \left(n + \frac{1}{2}\right) - \frac{1}{2} \qquad (\text{using } \int_{-\infty}^{\infty} \psi_n^* \psi_n dx = 1)\ ,$$

$$\Rightarrow |d_n|^2 = n$$

$$\therefore d_n = \sqrt{n} \tag{9.26}$$

So, finally Eqs. 9.24a, 9.24b can be written as

$$\psi_{n+1} = (n+1)^{-1/2}\ a_+\psi_n \tag{9.27a}$$

$$\psi_{n-1} = n^{-1/2}\ a_-\psi_n \tag{9.27b}$$

Using Eqs. 9.27, ψ_n can be written in terms of ψ_0 as

$$\psi_n = A_n\, (a_+)^n\, \psi_0\, (x) = \frac{(a_+)^n}{\sqrt{n!}} \psi_0 \tag{9.28}$$

$$\Longrightarrow A_n = \frac{1}{\sqrt{n!}} \tag{9.29}$$

We can also check orthogonality of ψ_n by using the following integral,

$$\int_{-\infty}^{\infty} (a_-\psi_m)^* a_-\psi_n = \int_{-\infty}^{\infty} \psi_m^* a_+ a_- \psi_n = \int_{-\infty}^{\infty} \psi_m^* \left(\frac{1}{\hbar\omega}H - \frac{1}{2}\right)\psi_n$$

$$\int_{-\infty}^{\infty} (a_-\psi_m)^* a_-\psi_n = n\int_{-\infty}^{\infty} \psi_m^*\psi_n dx$$

$$\int_{-\infty}^{\infty} (a_+ a_-\psi_m)^* \psi_n = n\int_{-\infty}^{\infty} \psi_m^*\psi_n dx \qquad (\text{as } (a_-\psi_m)^* a_- = (a_+ a_-\psi_m)^*)$$

$$\int_{-\infty}^{\infty} \left(\left(\frac{1}{\hbar\omega}H - \frac{1}{2}\right)\psi_m\right)^* \psi_n = m\int_{-\infty}^{\infty} \psi_m^*\psi_n dx = n\int_{-\infty}^{\infty} \psi_m^*\psi_n dx$$

$$\therefore (m-n)\int_{-\infty}^{\infty} \psi_m^*\psi_n dx = 0$$

$$\implies \int_{-\infty}^{\infty} \psi_m^*\psi_n dx = \delta_{mn} \tag{9.30}$$

Figure 9.3 shows the first five energy levels of the harmonic oscillator [2], their wavefunctions (ψ) and their probability densities ($|\psi|^2$). Note that all the wavefunctions of the harmonic oscillator problem extend well into the classically forbidden region (i.e., outside the potential, where $V > E$); as expected, the ψs decay exponentially to zero as $x \to \pm\infty$. This is a striking departure from our classical expectations. Further, the parity symmetry of the harmonic potential (in common with the particle in a box problem!) means that the ground state has even parity (i.e., is reflection symmetric), the first excited state has odd parity (i.e., is reflection anti-symmetric) and so on. Further, the excited states show the appearance of an increasing number of nodes (i.e., zeros of the wavefunction) as n increases. The fact that the nodes appear within the region where $E > V$ (i.e., inside the potential) reflects the fact that the kinetic energy (or curvature of the ψ) must increase substantially in the passage between neighbouring eigenstates resulting in the quantisation phenomenon marked by the quantum number n.

For $n >> 1$, the "envelope" of the quantum probability distribution begins to resemble the large amplitude oscillations observed classically (as shown as the dashed line in Fig.9.4) [1]. Also, note that very little of the probability distribution of $|\psi_{100}|^2$ is in the classically forbidden region. Together, these observations are called the **Correspondence Principle**, and suggest how a quantum system will turn classical. We do not, however, have a well understood framework for understanding this passage between the two worlds as yet.

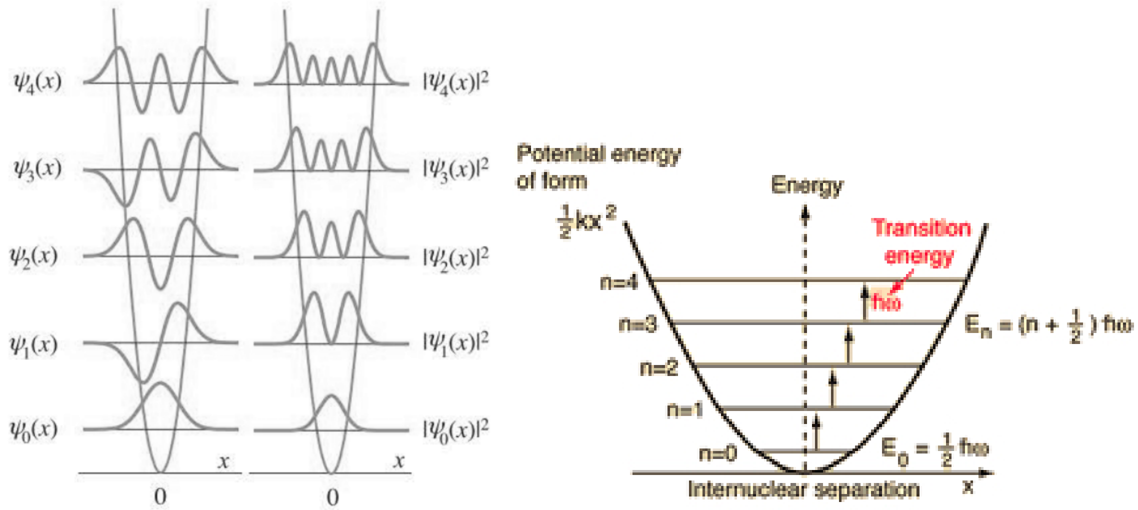


Figure 9.3: First few eigenstates (Left) and Energy Spectrum (Right) of the quantum mechanical harmonic oscillator.

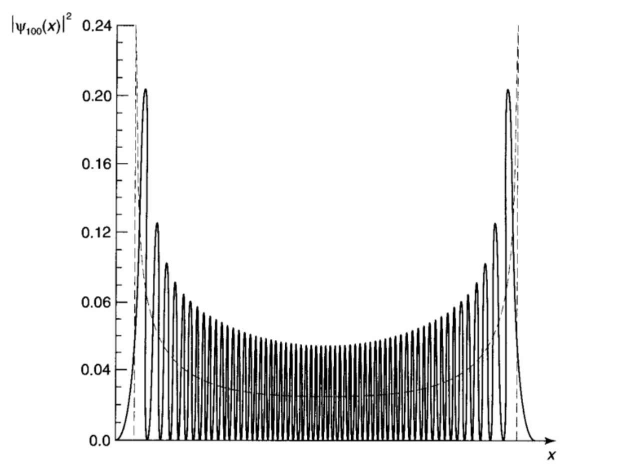


Figure 9.4: Probability distribution of high energy eigenstate ψ_{100}.

9.2 Analytic Method

We solve the TISE differential equation for the simple harmonic oscillator via a series solution. Recall that the starting point is:

$$-\frac{\hbar^2}{2m}\frac{d^2\psi(x)}{dx^2}+\frac{1}{2}m\omega^2x^2\psi(x)=E\psi(x) \tag{9.31}$$

Choosing $\xi=\sqrt{\frac{m\omega}{\hbar}}x$, $K=2E/\hbar\omega$, Eq. 9.31 can be written as

$$\frac{d^2\psi}{d\xi^2}=\left(\xi^2-K\right)\psi \tag{9.32}$$

For $\xi^2 >> K$, we can approximate the TISE as

$$\begin{aligned}\frac{d^2\psi}{d\xi^2}&\approx\xi^2\psi\\ \implies \psi(\xi)&\approx Ae^{-\xi^2/2}+Be^{\xi^2/2} \qquad (9.33)\\ \implies \psi(\xi)&\approx Ae^{-\xi^2/2} \qquad (9.34)\end{aligned}$$

Here, $B=0$ in Eq. 9.33, else $e^{\xi^2/2}\to\infty$ in the limit $x\to\pm\infty$ will cause ψ to diverge. Let us consider the following ansatz solution:

$$\psi(\xi)=h(\xi)\,e^{-\xi^2/2} \tag{9.35}$$

Note that we have already found the Gaussian part of the solution for ψ (as observed earlier from the algebraic solution).
Now, using Eq. 9.35, we obtain

$$\begin{aligned}\frac{d\psi}{d\xi}&=\left(\frac{dh(\xi)}{d\xi}-\xi h(\xi)\right)e^{-\xi^2/2}\\ \frac{d^2\psi}{d\xi^2}&=\left(\frac{d^2h(\xi)}{d\xi^2}-2\xi\frac{dh(\xi)}{d\xi}+\left(\xi^2-1\right)h(\xi)\right)e^{-\xi^2/2} \qquad (9.36)\end{aligned}$$

Using Eq. 9.36 in Eq. 9.32,

$$\begin{aligned}\left(\frac{d^2h(\xi)}{d\xi^2}-2\xi\frac{dh(\xi)}{d\xi}+\left(\xi^2-1\right)h(\xi)\right)e^{-\xi^2/2}&=\left(\xi^2-K\right)h(\xi)\,e^{-\xi^2/2}\\ \Rightarrow\frac{d^2h(\xi)}{d\xi^2}-2\xi\frac{dh(\xi)}{d\xi}+(K-1)\,h(\xi)&=0 \qquad (9.37)\end{aligned}$$

Solving Eq. 9.37 by the Frobenius method of a power series solution:

$$h(\xi) = a_0 + a_1\xi + a_2\xi^2 + \cdots = \sum_{j=0}^{\infty} a_j \xi^j \tag{9.38}$$

$$\frac{dh(\xi)}{d\xi} = a_1 + 2a_2\xi + 3a_3\xi^2 + \cdots = \sum_{j=0}^{\infty} j a_j \xi^{j-1} \tag{9.39}$$

$$\frac{d^2h(\xi)}{d\xi^2} = 2a_2 + 6a_3\xi + \cdots = \sum_{j=0}^{\infty} (j+1)(j+2)\, a_{j+2}\xi^j \tag{9.40}$$

Putting Eqs. 9.38, 9.39, 9.40 in Eq. 9.37, we get

$$\sum_{j=0}^{\infty} \left[(j+1)(j+2)\, a_{j+2} - 2j a_j + (K-1)\, a_j\right] \xi^j = 0 \tag{9.41}$$

From the uniqueness of the power series expansion,

$$(j+1)(j+2)\, a_{j+2} - 2j a_j + (K-1)\, a_j = 0 \tag{9.42}$$

leading to the **recursion relation**

$$\boxed{a_{j+2} = \frac{(2j+1-K)}{(j+1)(j+2)}\, a_j}$$

Note that this recursion relation will relate the odd number indexed coefficients among themselves, and the even number indexed coefficients among themselves. This means that we can write $h(\xi) = h_{even}(\xi) + h_{odd}(\xi)$ where

$$h_{even}(\xi) = a_0 + a_2\xi^2 + a_4\xi^4 + \cdots\ , \tag{9.43a}$$

$$h_{odd}(\xi) = a_1\xi + a_3\xi^3 + a_5\xi^5 + \cdots\ , \tag{9.43b}$$

the recursion formula builds everything on 2 arbitrary constants, a_0 and a_1, as would be expected from a 2^{nd} order differential equation.

Now, for $j >> 1$, K, we can approximate the recursion relation as

$$a_{j+2} \approx \frac{2}{j} a_j\ .$$

The relation $a_{j+2} = 2a_j/j$ has an approximate solution $a_j \approx c/(j/2)!$, where c is a constant. Then, for $\xi >> 1$,

$$h(\xi) \approx c \sum_j \frac{1}{(j/2)!}\xi^j \approx c \sum_j \frac{1}{j!}\xi^{2j} \approx c e^{\xi^2}\ .$$

But since $\psi = h(\xi)e^{-\xi^2/2}$, we can see that with the $h(\xi)$ found above, $\psi \sim e^{\xi^2/2} \to \infty$ as $\xi \to \pm\infty$! The only way out of this divergent solution is if the power series expansion terminates, i.e., if $a_{n+2} = 0$, which truncates either the even or the odd series, while the other series is zero right from the start:

$$\begin{aligned} a_{n+2} = 0 \quad \text{and} \quad & a_0 = 0 \quad \text{if } n \text{ is odd} \\ \text{or} \quad & a_1 = 0 \quad \text{if } n \text{ is even} \end{aligned}$$

Put another way, for a truly ∞ series, the polynomial part of ψ always dominates and leads to a solution that diverges as $\xi \to \pm\infty$. On the other hand, for any finite polynomial series, the $e^{-\xi^2/2}$ part of ψ dominates and gives a convergent, normalizable solution.

Thus, from $a_{n+2} = 0$ (but $a_n \neq 0$), we get

$$\begin{aligned} \frac{2n+1-K}{(n+1)(n+2)} &= 0 \\ \implies K = 2n+1 &= \frac{2E}{\hbar\omega} \\ \therefore E &= \left(n + \frac{1}{2}\right)\hbar\omega \end{aligned} \tag{9.44}$$

It appears amazing that the quantization of the energy eigenvalue, E, should arise from a technical detail in finding the solutions to the TISE for the SHO, but the fact is that normalizable solutions (i.e., that satisfy the boundary conditions $\psi(x) \to 0$ for $x \to \pm\infty$) only appear for certain E values (all the other values of E give solutions but these diverge to $\pm\infty$ as $x \to \pm\infty$).

Thus for the allowed values of K,

$$a_{j+2} = -\frac{2(n-j)}{(j+1)(j+2)}\, a_j\ . \tag{9.45}$$

For the ground state $n = 0$, we take $a_1 = 0$ to kill all terms in h_{odd}, while $j = 0$ gives $a_2 = 0$ (and hence all a_4, a_6 etc), such that

$$\begin{aligned} h_0(\xi) &= a_0 \qquad (9.46) \\ \implies \psi_0(\xi) &= a_0 e^{-\xi^2/2}\ . \end{aligned}$$

For the first excited state $n = 1$, $a_0 = 0$ (such that all terms in h_{even} vanish) together with $j = 1$ (such that $a_3 = 0 = a_5 = \cdots$) gives

$$\begin{aligned} h_1(\xi) &= a_1\xi \qquad (9.47) \\ \psi_1(\xi) &= a_1\xi e^{-\xi^2/2}\ . \end{aligned}$$

For the second excited state $n = 2$, take $a_1 = 0$ (such that all terms in h_{odd} vanish) and $a_2 = -2a_0$, and $j = 2$ (such that $a_4 = 0 = a_6 = \cdots$) gives

$$\begin{aligned} h_2(\xi) &= a_0\left(1 - 2\xi^2\right) \qquad (9.48) \\ \psi_2(\xi) &= a_0\left(1 - 2\xi^2\right)e^{-\xi^2/2}\ . \end{aligned}$$

In general, $h_n(\xi)$ will be a polynomial in ξ of degree n, involving only even powers when n is even and only odd powers when n is odd. Apart from constants a_0, a_1, these polynomials form are the so-called **Hermite Polynomials** $H_n(\xi)$:

$$
\begin{aligned}
H_0(\xi) &= 1 \\
H_1(\xi) &= 2\xi \\
H_2(\xi) &= 4\xi^2 - 2 \\
H_3(\xi) &= 8\xi^3 - 12\xi \\
H_4(\xi) &= 16\xi^4 - 48\xi^2 + 12 \\
H_5(\xi) &= 32\xi^5 - 160\xi^3 + 120\xi
\end{aligned}
$$

where an arbitrary multiplicative factor is chosen such that the coefficient of the highest power of ξ is 2^n.
Thus, we obtain the eigenfunctions for the S.H.O. as

$$\psi_n(\xi) = \left(\frac{m\omega}{\pi\hbar}\right)^{1/4} \frac{1}{\sqrt{2^n n!}} H_n(\xi)\, e^{-\xi^2/2} \tag{9.49}$$

where the $(m\omega/\pi\hbar)^{1/4}$ factor comes from normalising ψ_0 and ψ_1 to learn the constants a_0 and a_1 respectively, and the $(n!)^{-1/2}$ factor we had already learnt from the algebraic solution.

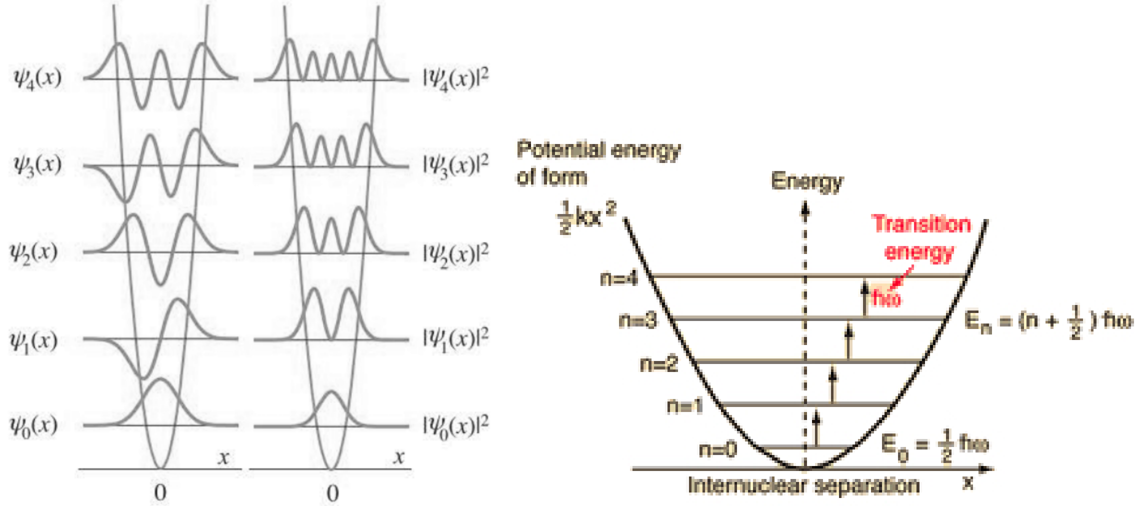


Figure 9.5: First few eigenstates (Left) and Energy Spectrum (Right) of the quantum mechanical harmonic oscillator.

It is important to note that the additional zeros ("nodes") of the excited wavefunctions ψ_n $(n \geq 0)$ (over and above the nodes at $x \to \pm\infty$ imposed by the boundary conditions) arises from the zeros of the Hermite polynomials $H_n(\xi)$. Further, the $H_n(\xi)$ are odd and even under the parity (reflection) transformation for $n \in$ odd and even respectively.

Chapter 10

Free particle

For a free particle, $V(x) = 0$ and the TISE is:

$$-\frac{\hbar^2}{2m}\frac{d^2\psi}{dx^2} = E\psi \tag{10.1}$$

$$\psi'' = -k^2\psi \quad \text{where } k = \sqrt{\frac{2mE}{\hbar^2}},\ E \geq 0 \tag{10.2}$$

The general solution is:

$$\begin{aligned}\Psi(x,t) &= \psi(x)\, e^{-iEt/\hbar} \\ &= \left(Ae^{ikx} + Be^{-ikx}\right) e^{-iEt/\hbar} \\ &= Ae^{ik\left(x-\frac{\hbar k}{2m}t\right)} + Be^{-ik\left(x+\frac{\hbar k}{2m}t\right)} \end{aligned} \tag{10.3}$$

Eq. 10.3 shows a superposition of right going wave $(x - vt)$ and a left going wave $(x + vt)$ moving with speed $v = \hbar k/2m = \sqrt{E/2m}$, momentum $p = \hbar k$.

Note that the classical velocity of the particle $v_{cl} = \sqrt{2E/m}$. This paradox is resolved in the following way: The wavefunction given by Eq. 10.3 is not normalizable i.e., it doesn't die out at $\pm\infty$. This tells us that such separable solutions do not represent physically reasonable states for a free particle. Instead, a solution can be found by taking a linear combination of Eq. 10.3 (shown in Eq. 10.4) with different velocities so that they are normalizable. Indeed, this localizes the wave-function giving rise to the idea of "wave-packet".

$$\Psi(x,t) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} \phi(k)\, e^{i\left(kx-\frac{\hbar k^2}{2m}t\right)} dk \tag{10.4}$$

This $\psi(x,t)$ can be normalized for appropriate $\phi(k)$. Note that it contains an entire range of k's – hence the wavepacket. The simpler problem is to find the $\phi(k)$ for $\psi(x, t = 0)$ (the initial wavefunction).

$$\Psi\left(x,0\right)=\int\limits_{-\infty}^{\infty}\frac{\phi\left(k\right)}{\sqrt{2\pi}}e^{ikx}dk \tag{10.5}$$

At this point, we can define the Fourier transform $F\left(k\right)$ of a function $f\left(x\right)$ (as well as the and inverse Fourier transform) as

$$F\left(k\right)=\frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{\infty}f\left(x\right)e^{-ikx}dx \tag{10.6}$$

$$f\left(x\right)=\frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{\infty}F\left(k\right)e^{ikx}dk \tag{10.7}$$

Any normalizable $\psi\left(x,0\right)$ will have a valid Fourier transform. So,

$$\phi\left(k\right)=\frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{\infty}e^{-ikx}\Psi\left(x,0\right)dx \tag{10.8}$$

Now the wave packet has a **phase velocity** $v_{ph}=\omega/k$ and a **group velocity** $v_g=d\omega/dk$. Taking $\omega=E/\hbar$,

$$\omega=\frac{\hbar k^2}{2m} \tag{10.9}$$

$$v_{ph}=\frac{\hbar k}{2m} \tag{10.10}$$

$$v_{gr}=\frac{\hbar k}{m}=2v_{ph} \tag{10.11}$$

(Revise your concepts of phase and group velocity!)

Chapter 11

Scattering Processes and Quantum Tunneling

Having learnt the basics of the solution for the free particle problem in quantum mechanics, we will now study a couple of problems that involve us understand how matter waves can be scattered from some simple potentials. In doing so, we will be using the simple $e^{\pm ikx}$ type waveforms that we encountered in the previous chapter, and not the somewhat more complex matter wavepacket. This is more for ease of convenience, as the maths is somewhat harder for the wavepackets than it is for the simple $e^{\pm ikx}$ waveforms. For those who are worried about normalisability of the wavefunctions, you can view the wavepacket as being composed out of a huge number of such $e^{\pm ikx}$ waves. Thus, we can always learn the physics of scattering from the simpler situations, and then use the idea of wavepacket construction to create the mathematically correct solutions.

11.1 Step potential

Consider a step potential as shown in figure 11.1 given by

$$V(x) = \begin{cases} 0 & \text{if } x < 0 \quad \text{(region I)} \\ V_0 & \text{if } x \geq 0 \quad \text{(region II)} \end{cases} \tag{11.1}$$

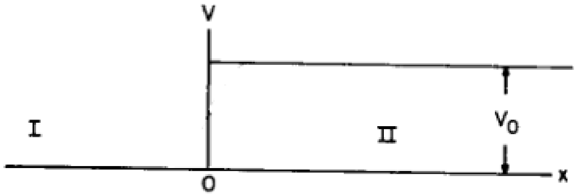


Figure 11.1: A schematic diagram of $V(x)$

Clasically, an incoming particle from the left will bounce back from the step if $E < V_0$ and will be able to move to region II if $E > V_0$ with a changed momentum. Let's see what quantum mechanics tells us.

The TISE is:

$$\frac{d^2\psi}{dx^2} + \frac{2mE}{\hbar^2}\psi = 0 \quad x < 0 \;\; \text{(Region I)}$$
$$\frac{d^2\psi}{dx^2} + \frac{2m}{\hbar^2}\left(E - V_0\right)\psi = 0 \quad x \geq 0 \;\; \text{(Region II)}$$

Boundary conditions:

$$(i) \;\; \psi\left(x \to 0^-\right) = \psi\left(x \to 0^+\right) \tag{11.2a}$$
$$(ii) \;\; \left.\frac{d\psi}{dx}\right|_{x\to 0^-} = \left.\frac{d\psi}{dx}\right|_{x\to 0^+} \tag{11.2b}$$

For $E > V_0$, since $E > 0$ & $(E - V_0) > 0$, we can use the "free-particle" e^{ikx} type ψ everywhere.

$$\psi\left(x\right) = \begin{cases} Ae^{ik_1x} + Be^{-ik_1x} & x < 0, \, k_1 = \sqrt{\frac{2mE}{\hbar^2}} \\ Ce^{ik_2x} + De^{-ik_2x} & x \geq 0, \, k_2 = \sqrt{\frac{2m(E-V_0)}{\hbar^2}} \end{cases} \tag{11.3}$$

For a particle coming towards the step from the left, $D = 0$ on physical grounds. Thus the two boundary conditions gives us:

$$(i) \;\; A + B = C \tag{11.4}$$
$$(ii) \;\; ik_1\left(A - B\right) = ik_2C \tag{11.5}$$

On general grounds, we can choose to measure all fluxes(or amplitudes) of particle relative to the incoming flux which is equivalent to taking $A = 1$. In this case

$$(i) \;\; 1 + B = C \tag{11.6}$$
$$(ii) \;\; k_1\left(1 - B\right) = k_2C \tag{11.7}$$
$$\implies B = \frac{k_1 - k_2}{k_1 + k_2} \quad \& \quad C = \frac{2k_1}{k_1 + k_2} \tag{11.8}$$

Then the coefficients B, C can be interpreted as the **reflection amplitude** r and the **transmission amplitude** t respectively. Do these relations for r and t appear familiar from something you may have seen in wave optics? The corresponding reflection and transmission coefficients R, T are given by

$$\begin{aligned} R &= |r|^2 = B^2 \\ &= \frac{(k_1-k_2)^2}{(k_1+k_2)^2} = \frac{\left(\sqrt{E}-\sqrt{E-V_0}\right)^2}{\left(\sqrt{E}+\sqrt{E-V_0}\right)^2} \qquad (11.9) \\ T &= |t|^2 = C^2 \\ &= \frac{4k_1^2}{(k_1+k_2)^2} = \frac{4E}{\left(\sqrt{E}+\sqrt{E-V_0}\right)^2} \qquad (11.10) \end{aligned}$$

Also, note that

$$R + \frac{k_2}{k_1}T = 1 \tag{11.11}$$

In order to see such scattering, we have to take the length scale over which the step exists is much smaller compared to de Broglie λ of the quantum particle. In the opposite regime, we will see only classical results.

For $E < V_0$,

$$\psi(x) = \begin{cases} Ae^{ik_1x} + Be^{-ik_1x} & x < 0,\ k_1 = \sqrt{\frac{2mE}{\hbar^2}} \\ Ce^{k_2x} + De^{-k_2x} & x \geq 0,\ k_2 = \sqrt{\frac{2m(V_0-E)}{\hbar^2}} \end{cases} \tag{11.12}$$

Unless $C = 0$, $\psi \to \infty$ as $x \to \infty$. So, $C = 0$ and $A = 1$ following the same logic as in the previous case. Applying boundary conditions given in Eq. 11.2 we get

$$1 + B = D \tag{11.13}$$

$$ik_1(1-B) = -k_2D \tag{11.14}$$

$$\implies B = \frac{k_1 - ik_2}{k_1 + ik_2} \quad \& \quad D = \frac{2k_1}{k_1 + ik_2} \tag{11.15}$$

The reflection, transmission amplitudes will be

$$\begin{aligned} R &= |r|^2 = |B|^2 \\ &= \left|\frac{k_1 - ik_2}{k_1 + ik_2}\right|^2 = 1 \qquad (11.16) \end{aligned}$$

$$\implies T = \frac{k_1}{k_2}(1-R) = 0 \tag{11.17}$$

Even if the transmission amplitude T is zero, $\psi(x) \neq 0$ at $x \geq 0$. This means there is a non-zero probability to find the particle under the step (even though this may be experimentally quite challenging!). Note: only when $V_0 \to \infty$, $\psi(x) \to 0$ for $x \geq 0$. This phenomenon has no classical particle analog and is responsible for **tunneling** when we have a barrier of finite extent (the next case we will study!). However, an analogy can be drawn to the phenomenon of evanescent waves in total internal reflection of e-m waves at a boundary between dense and rare media. For this, revise Fresnel's theory for reflection and refraction of e-m waves.

11.2 Potential Barrier & Tunneling

Instead of the step potential shown in Figure 11.1, consider a barrier potential given by 11.18 and shown in Figure 11.2

$$V(x) = \begin{cases} 0 & \text{if } x < 0 & \text{(region I)} \\ V_0 & \text{if } 0 \leq x \leq a & \text{(region II)} \\ 0 & \text{if } x > a & \text{(region III)} \end{cases} \tag{11.18}$$

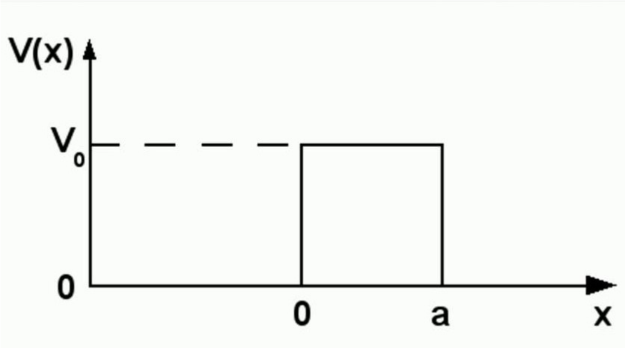


Figure 11.2: A potential barrier

Solving the TISE in each of the three regions individually means that we can take $\psi(x)$ as

$$\psi(x) = \begin{cases} e^{ik_1x} + re^{-ik_1x} & x < 0,\ k_1 = \sqrt{\frac{2mE}{\hbar^2}} \\ Ae^{ik_2x} + Be^{-ik_2x} & 0 \leq x \leq a,\ k_2 = \sqrt{\frac{2m(E-V_0)}{\hbar^2}} \\ te^{ik_1x} & x > a,\ k_1 = \sqrt{\frac{2mE}{\hbar^2}} \end{cases} \tag{11.19}$$

Note that we have already used the boundary condition of an incoming particle from the left with a probability amplitude set to 1, while r and t give the probability amplitudes for reflection and transmission respectively. Further, E is the energy eigenvalue of the Hamiltonian for the full problem. It is also identical to the kinetic energy of the particle in regions I and III. To evaluate the four unknowns r, A, B, t, we use the continuity conditions on $\psi(x)$ and $d\psi/dx$ at both $x = 0$ & $x = a$, giving

$$\psi(x \to 0-) = \psi(x \to 0+) \Rightarrow 1 + r = A + B \tag{11.20a}$$

$$\psi'(x \to 0-) = \psi'(x \to 0+) \Rightarrow ik_1(1 - r) = ik_2(A - B) \tag{11.20b}$$

$$\psi(x \to a-) = \psi(x \to a+) \Rightarrow te^{ik_1a} = Ae^{ik_2a} + Be^{-ik_2a} \tag{11.20c}$$

$$\psi'(x \to a-) = \psi'(x \to a+) \Rightarrow ik_1te^{ik_1a} = ik_2\left(Ae^{ik_2a} - Be^{-ik_2a}\right) \tag{11.20d}$$

We set $\mu = k_2/k_1 = \sqrt{1 - V_0/E}$. Then solving the set of Eqs. 11.20 after some lengthy algebra, we get

$$r = \frac{(1-\mu^2)\sin(k_2 a)}{(1+\mu^2)\sin k_2 a + 2i\mu\cos k_2 a} \tag{11.21a}$$

$$t = \frac{2i\mu e^{-ik_1 a}}{(1+\mu^2)\sin k_2 a + 2i\mu\cos k_2 a} \tag{11.21b}$$

$$A = \frac{i(1+\mu)e^{-ik_2 a}}{(1+\mu^2)\sin k_2 a + 2i\mu\cos k_2 a} \tag{11.21c}$$

$$B = \frac{-i(1-\mu)e^{ik_2 a}}{(1+\mu^2)\sin k_2 a + 2i\mu\cos k_2 a} \tag{11.21d}$$

$\boxed{\text{For } E > 0 \text{ and } E \geq V_0,}$ k_1, k_2 and $\mu \in \mathcal{R}$. Then, the transmission coefficient is given by

$$\begin{aligned} T = |t|^2 &= \frac{4\mu^2}{(1+\mu^2)^2\sin^2 k_2 a + 4\mu^2\cos^2 k_2 a} \\ &= \frac{4\mu^2}{4\mu^2 + (1-\mu^2)^2\sin^2 k_2 a} = \frac{1}{1+\frac{1}{4}\left(\frac{1-\mu^2}{\mu}\right)^2\sin^2 k_2 a} \\ &= \frac{1}{1+\frac{1}{4}\left(\frac{k_1^2-k_2^2}{k_1 k_2}\right)^2\sin^2 k_2 a} \end{aligned} \tag{11.22}$$

Note that the reflection coefficient $R = 1 - T$. Both R and T are plotted in Fig.11.3 as a function of E/V_0. Whenever $k_2 a = n\pi$ the system is in resonance and there will be perfect

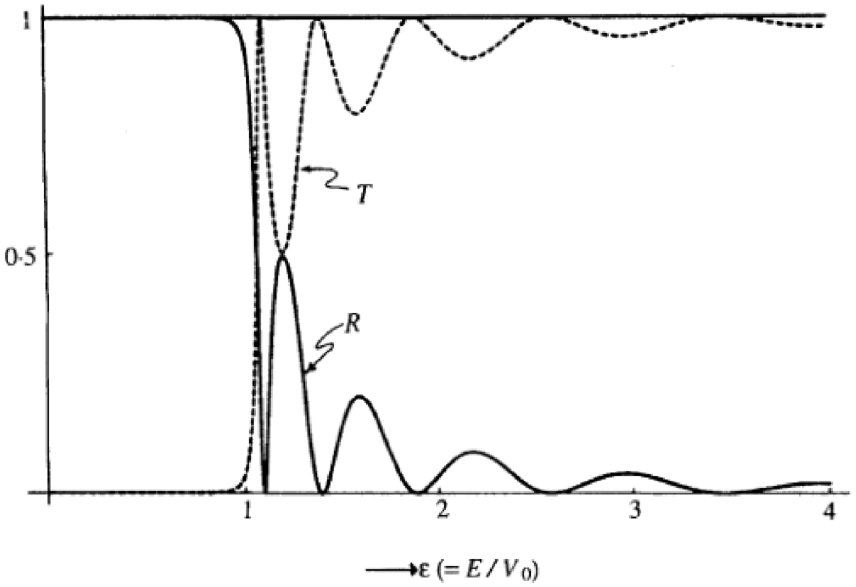


Figure 11.3: Transmission, Reflection coefficients as a function of $\epsilon = E/V_0$

transmission $T = 1$, $R = 0$ (see Fig.11.3). We also see that T oscillates and finally becomes asymptotically equal to 1 for $E >> V_0$; this is seen by taking $\mu \to 1$ in the expression for T in eq.(11.22). The criterion for the nth Resonance is given by the energy E_n

$$\begin{aligned} k_2 a &= \sqrt{\frac{2m\left(E - V_0\right)}{\hbar^2}}\, a = n\pi \\ E_n &= V_0 + \frac{n^2\pi^2\hbar^2}{2ma^2} \qquad (11.23) \\ &= V_0\left(1 + \frac{n^2\pi^2\hbar^2}{2mV_0a^2}\right) \qquad (11.24) \end{aligned}$$

Where do these resonances arise from? A simple way to picture them is to recall the Fabry-Perot resonances that arise within a cavity: light bouncing back and forth between the two "leaky" mirrors of the F-P cavity leads to a sequence of reflections and transmissions at the two mirrors, giving rise to standing waves within the cavity. A similar phenomenon is happening here for the wavefunction of the electron when over the barrier: there is a sequence of back and forth reflections and transmissions that happen at the two interfaces (at the two ends of the barrier) that fit together to form a perfect standing wave pattern at resonance. This is also indicated by the energy $n^2\pi^2\hbar^2/2ma^2$ (i.e., the eigenenergy of the standing waves formed in the particle in a box problem!) appearing in eq.(11.23). It is from this standing wave pattern that the perfect $T = 1$ arises from adding up the entire sequence of transmission probability amplitudes coherently. Similarly, $R = 0$ arises from adding up the entire sequence of reflection probability amplitudes coherently.

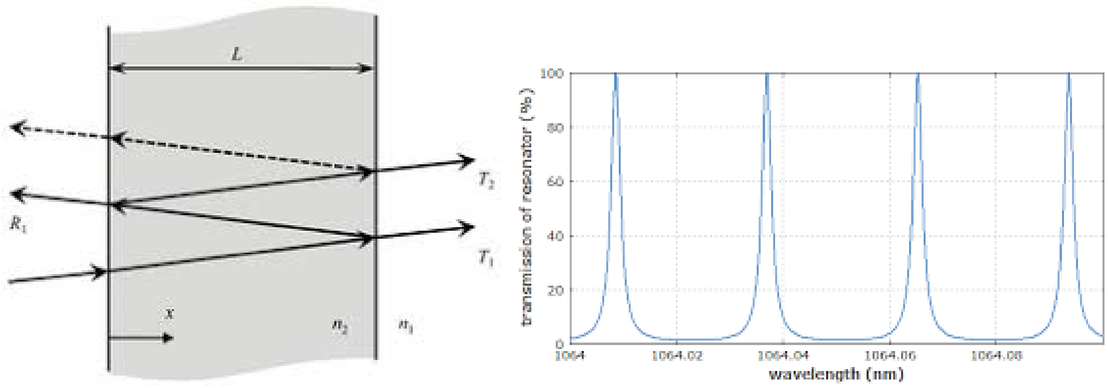


Figure 11.4: (Left) The Fabry-Perot Cavity, displaying the sequence of reflections and transmissions of e-m waves at the two leaky mirrors and (Right) the Transmission coefficient of the F-P cavity as a function the e-m wavelength. Source: The internet.

At the limit $E_0 \to V_0^+$, $k_2 \to 0 \implies \mu \to 0^+$,

$$
\begin{aligned}
\frac{k_1^2 - k_2^2}{k_1 k_2} &\approx \frac{k_1^2}{k_1 k_2} \approx \frac{k_1}{k_2} \\
\implies T &\approx \left(1 + \frac{1}{4}\left(\frac{k_1}{k_2}\right)^2 k_2^2 a^2\right)^{-1} \\
&\approx 1 - \frac{1}{4} k_1^2 a^2 \approx 1 - \frac{1}{4}\left(\frac{2mV_0}{\hbar^2}\right) a^2 \\
&\approx 1 - \frac{mV_0 a^2}{2\hbar^2}
\end{aligned}
\tag{11.25}
$$

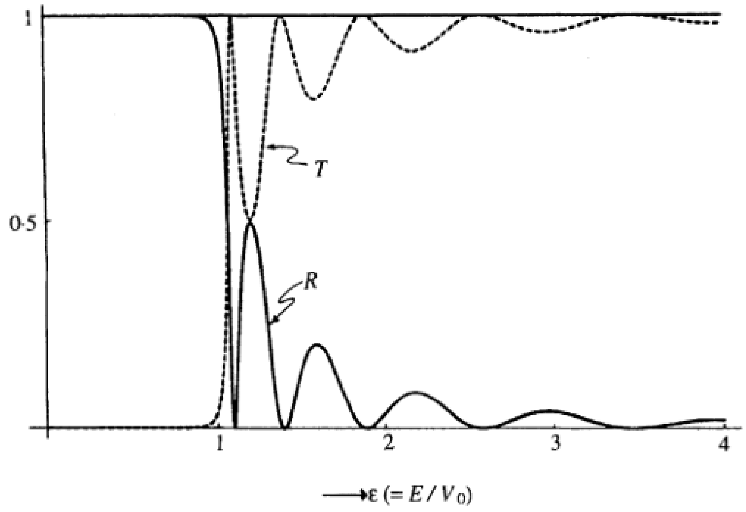


Figure 11.5: Transmission, Reflection coefficients as a function of Energy ϵ

$\boxed{\text{For } E < V_0,}$ k_2 becomes imaginary

$$
k_2 = iK = \frac{i}{\hbar}\sqrt{2m\left(V_0 - E\right)}
$$

$$
\implies T = |t|^2 = \frac{\left(2k_1/K\right)^2}{\left(1 - \frac{k_1^2}{K^2}\right)^2 \sinh^2\left(Ka\right) + \left(\frac{2k_1}{K}\right)^2 \cosh^2\left(Ka\right)} \tag{11.26}
$$

Thus, we can see that $T \to 0$ and $R = 1 - T \to 1$ as $E \to 0^+$, joining smoothly onto the expressions for T and R given by (11.25) as $E \to V_0^-$ (see Fig.11.3). An application of the phenomenon of quantum tunneling is realised in the scanning tunneling microscope (STM) shown below in Fig.11.6.

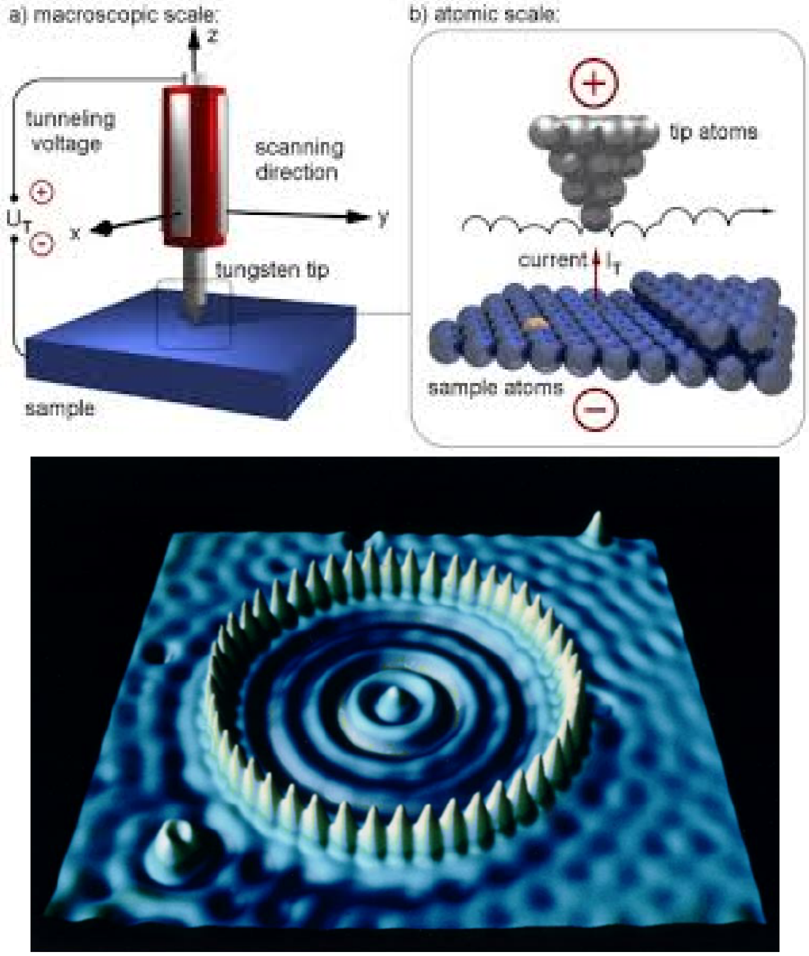


Figure 11.6: (Above) The physics of the scanning tunneling microscope (STM) is based on the phenomenon of quantum tunneling of electrons between the sample surface and the atoms of the STM tip and across a vacuum "barrier". The tunnel current IT depends on how easily sample surface electrons are available for such tunneling events, giving an indication of the nature of the atoms on the sample surface in terms of a picture. (Below) A "quantum corral" observed via STM measurements of Iron atoms placed on a Cu surface. Source: The internet.

Chapter 12

Time Independent Schrödinger Equation (TISE) in 3D

In 3D, the time dependant, independant Schrödinger equations is given by

$$-\frac{\hbar^2}{2m}\nabla^2\psi\left(\mathbf{r},t\right)+V\left(\mathbf{r}\right)\psi\left(\mathbf{r},t\right)=i\hbar\frac{\partial\psi\left(\mathbf{r},t\right)}{\partial t} \tag{12.1}$$

$$-\frac{\hbar^2}{2m}\nabla^2\psi\left(\mathbf{r}\right)+V\left(\mathbf{r}\right)\psi\left(\mathbf{r}\right)=E\psi\left(\mathbf{r}\right) \tag{12.2}$$

In solving the TISE in 3D, we need to choose a system of coordinates: rectangular, cylindrical & spherical, which are best chosen by noting any symmetries of the potential $V\left(\mathbf{r}\right)$, e.g., for a central force $F\left(\mathbf{r}\right)$ which arises from a $V\left(|\mathbf{r}|\right)$ is most approximately (and easily) solved using spherical coordinates.

12.1 The Particle in a 3D Box

As a warm up, lets first consider the simpler problem of the infinite square well in 3D, which can be solved using rectangular coordinates.

$$V\left(\mathbf{r}\right)=\begin{cases}0 & \text{when} \quad 0<x<a \\ & \qquad\qquad 0<y<b \\ & \qquad\qquad 0<z<c \\ \infty & \text{otherwise}\end{cases} \tag{12.3}$$

Thus inside the box,

$$-\frac{\hbar^2}{2m}\nabla^2\psi\left(\mathbf{r}\right)=E\psi\left(\mathbf{r}\right)$$

$$\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2}\right)\psi=-\frac{2mE}{\hbar^2}\psi \tag{12.4}$$

Using the idea of separation of variables, we take the trial wavefunction

$$\psi\left(x, y, z\right) = \psi_1\left(x\right)\psi_2\left(y\right)\psi_3\left(z\right) \tag{12.5}$$

Putting Eq. 12.5 in Eq. 12.4,

$$\frac{1}{\psi_1\left(x\right)}\frac{\partial^2\psi_1\left(x\right)}{\partial x^2} + \frac{1}{\psi_2\left(y\right)}\frac{\partial^2\psi_2\left(y\right)}{\partial y^2} + \frac{1}{\psi_3\left(z\right)}\frac{\partial^2\psi_3\left(z\right)}{\partial z^2} = -\frac{2mE}{\hbar^2}$$

Since the LHS has nicely factorized out into three separate equations & the RHS is a constant independant of (x, y, z), we can take

$$\frac{1}{\psi_1}\frac{\partial^2\psi_1}{\partial x^2} = -\frac{2mE_x}{\hbar^2} \quad \frac{1}{\psi_2}\frac{\partial^2\psi_2}{\partial y^2} = -\frac{2mE_y}{\hbar^2} \quad \frac{1}{\psi_3}\frac{\partial^2\psi_3}{\partial z^2} = -\frac{2mE_z}{\hbar^2}$$

Here, $E = E_x + E_y + E_z$. We thus have 3 independent 1D infinite square well problems with

$$\psi_1\left(x\right) = A_1 \sin\frac{n_x\pi x}{a}, \quad \psi_2\left(y\right) = A_2 \sin\frac{n_y\pi y}{a}, \quad \psi_3\left(z\right) = A_3 \sin\frac{n_z\pi z}{a} \tag{12.6}$$

$$E_x = \frac{\hbar^2\pi^2}{2ma^2}n_x^2, \quad E_y = \frac{\hbar^2\pi^2}{2mb^2}n_y^2, \quad E_z = \frac{\hbar^2\pi^2}{2mc^2}n_z^2 \tag{12.7}$$

Then,

$$\psi\left(\mathbf{r}\right) = A\sin\left(\frac{n_x\pi x}{a}\right)\sin\left(\frac{n_y\pi y}{b}\right)\sin\left(\frac{n_z\pi z}{c}\right) \tag{12.8}$$

$$\begin{aligned} E &= E_x + E_y + E_z \\ &= \frac{\hbar^2\pi^2}{2m}\left(\frac{n_x^2}{a^2} + \frac{n_y^2}{b^2} + \frac{n_z^2}{c^2}\right) \end{aligned} \tag{12.9}$$

where $A = A_1A_2A_3$. For the special case of a cube $(a = b = c)$

$$\psi\left(\mathbf{r}\right) = A\sin\left(\frac{n_x\pi x}{a}\right)\sin\left(\frac{n_y\pi y}{a}\right)\sin\left(\frac{n_z\pi z}{a}\right) \tag{12.10}$$

$$E = \frac{\hbar^2\pi^2}{2ma^2}\left(n_x^2 + n_y^2 + n_z^2\right) \tag{12.11}$$

The ground state is given by $(n_x = 1, n_y = 1, n_z = 1)$, as $\psi = 0$ for any one of these three indices being zero. The corresponding ground state energy is $E_{(1,1,1)} = 3\hbar^2\pi^2/2ma^2$. Further, we can see that the three states $(n_x, n_y, n_z) = (1, 2, 1)$, $(1, 1, 2)$, $(2, 1, 1)$ all have the same energy $E_{(1,2,1} = 6\hbar^2\pi^2/2ma^2 = E_{(1,1,2)} = E_{(2,1,1)}$, and correspond to a triplet of the lowest lying excited states. This is the phenomenon of "degeneracy" (when 2 or more eigenstates with distinct ψ have the same E). Here, the wavefunctions are related by pair-wise interchanges of the x, y & z axes (interchanges one face of the cube with another) which leaves the potential unchanged (the symmetry of the cube). However, such a degeneracy needs the tuning of atleast two parameters here (say, b and c with respect to a) so as to meet $a = b = c$. This means that this degeneracy is "accidental" (i.e., fragile), and is easily lifted by moving even slightly away from the condition $a = b = c$.

Chapter 13

Formalism II : Vector Spaces in Quantum Mechanics

13.1 Vector Spaces

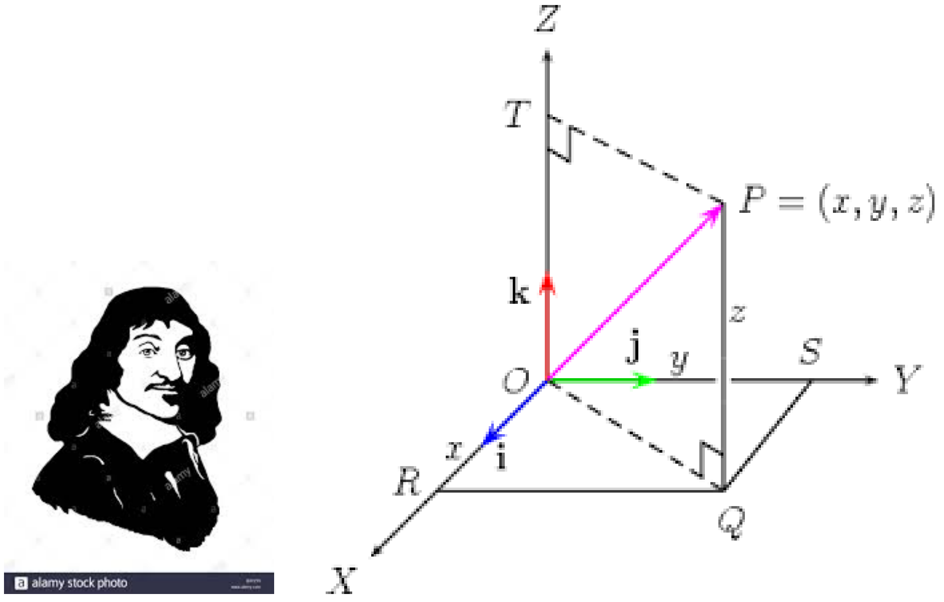


Figure 13.1: Left: Renée Descartes. Right: A vector in Cartesian coordinatestt. Source: The Internet.

We are already familiar with 2D and 3D spaces that can accommodate vectors. Generalising this to $n > 3$ dimensions means that for

(i) $\vec{u} = (u_1, u_2, u_3, \ldots, u_n)$, $\vec{v} = (v_1, v_2, v_3, \ldots, v_n)$, we need a "rule for vector addition"

$$\vec{u} + \vec{v} = (u_1 + v_1, u_2 + v_2, u_3 + v_3, \ldots, u_n + v_n) = \vec{w} \ , \tag{13.1}$$

as well as

(ii) a “rule for the multiplication of a vector by a scalar”

$$\vec{r} = c\vec{u} \ , \tag{13.2}$$

where $c \in \mathcal{R}$ for a vector space of real-valued vectors and $c \in \mathcal{C}$ for a vector space of complex-valued vectors (as in quantum mechanics).

Finally, we also need a “rule for the dot product” between two vectors that leads to a scalar quantity

$$\vec{u} \cdot \vec{v} = u_1 v_1 + u_2 v_2 + \ldots + u_n v_n \ . \tag{13.3}$$

Examples of vector spaces includes

(a) the set of all real numbers, where each real number is a one-component vector and the entire set forms a 1D vector space, and

(b) the set of (suitably defined) functions $f(x)$ is an infinite-dimensional vector space, since $f(x) + g(x) = h(x)$ and $g(x) = cf(x)$ are well-defined operations.

13.1.1 Dimension of a vector space:

The dimension of a vector space is the maximum number of linearly independent vectors the space can have (also known as “the basis”). This number can be infinite. For a N-dimensional vector space with a basis given by $(\phi_1, \phi_2, \ldots, \phi_N)$ such that any vector ψ can be written as

$$\psi = \sum_{i=1}^{N} a_i \phi_i \ , \tag{13.4}$$

i.e., a vector ψ can be written in terms of a linear superposition of the ϕ_is, where the ϕ_is form a “complete set which spans the entire set”. The coefficients (or weight factors) of the various components of the vector ψ, a_i, are the projections of ψ onto the various “orthogonal” (or “normal”) directions given by ϕ_i. To understand this better, we need to define the “dot”. “scalar” or “inner product” $\langle \phi_i | \phi_j \rangle$ such that

$$a_i = \langle \psi | \phi_i \rangle \ , \tag{13.5}$$

and for a set of orthonormal basis vectors $\{\phi_i\}$

$$\langle \phi_i | \phi_j \rangle = \delta_{ij} \ , \tag{13.6}$$

where the Kronecker delta is defined as $\delta_{ij} = 1$ for $i = j$, and 0 otherwise.

13.1.2 More on the Inner Product

Defining a “dot” product for a finite dimensional vector space is easy, but what about for an infinite-dimensional vector space (e.g., the space of functions)? For such cases, we need to construct a more abstract concept called the “Inner Product”. Just as the dot product of two finite dimensional vectors gives a real number, by analogy, we consider the Inner

Product of two vectors, ψ and ϕ, of an abstract vector spaceto be a function which uses ψ and ϕ as inputs and whose result is (in general) a complex number

$$\langle\psi|\phi\rangle = c \ , \ c \in \mathcal{C} \ . \tag{13.7}$$

The Inner Product has the following properties:

$$\begin{aligned} (i) \ \ \langle\psi + \phi|\theta\rangle &= \langle\psi|\theta\rangle + \langle\phi|\theta\rangle \\ (ii) \ \ \langle\psi|c\phi\rangle &= c\langle\psi|\phi\rangle \\ (iii) \ \ \langle\psi|\phi\rangle &= \langle\phi|\psi\rangle^* \\ (iv) \ \ \langle\psi|\psi\rangle &\geq 0 \ . \end{aligned}$$

Using relations (ii) and (iii) from above, we can see that

$$\begin{aligned} \langle\phi|c\psi\rangle^* &= c^*\langle\phi|\phi\rangle^* \\ &= c^*\langle\psi|\phi\rangle \\ \text{But} \ \ \langle\phi|c\psi\rangle^* &= \langle c\psi|\phi\rangle \\ \Rightarrow \ \ \langle c\psi|\phi\rangle &= c^*\langle\psi|\phi\rangle \ . \end{aligned}$$

The inner product for the set of complex-valued functions defined in 3D is

$$\langle f(\vec{r})|g(\vec{r})\rangle = \int_{-\infty}^{\infty} dxdydz f(\vec{r})^* g(\vec{r}) \ . \tag{13.8}$$

In 1D, this becomes $\langle f(x)|g(x)\rangle = \int_{-\infty}^{\infty} dx f^*(x)g(x)$, such that for

$$\begin{aligned} f(x) = \sin x \quad &, \quad g(x) = x\sin x \\ \langle f(x)|g(x)\rangle &= \int_{-\infty}^{\infty} dx f^*(x)g(x) \\ &= \int_{-\infty}^{\infty} x\sin^2 x \\ &\rightarrow \ 0 \ , \end{aligned}$$

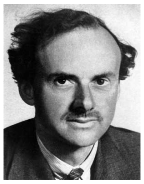
as the argument $x\sin^2 x$ is an odd function in x.
Finally, the requirement that the wavefunction $\psi(x)$ be normalised can now be recast as

$$\langle\psi|\psi\rangle = 1 \ , \tag{13.9}$$

i.e., the "length" of ψ is set of 1. Further, the Expectation value of an operator $\hat{\theta}$ can be defined as

$$\langle\hat{\theta}\rangle = \langle\psi|\hat{\theta}|\psi\rangle \equiv \langle\psi|\hat{\theta}\psi\rangle \ . \tag{13.10}$$

13.2 The Dirac Bra-Ket Notation



Dirac

Hilbert

The vector space of function that we're interested in Quantum Mechanics is that of wavefunctions such that

$$\langle\psi|\psi\rangle = \int d^3x|\psi(r)|^2 = 1 \ , \tag{13.11}$$

i.e., the ψs are both square-integrable and normalisable. Such a vector space is an example of a linear vector space called a **Hilbert space**. Thus, the physical state of a system is represented by elements of a Hilbert space called "state vectors". We do have freedom in how to choose to represent these state vectors by means of an expansion using different functions. This is called the "choice of basis", and is analogous to being able to representing a 3D coordinate system using the (x, y, z), (r, θ, ϕ) or (ρ, ϕ, z) coordinates.

It is, however, important to note that the meaning of the state of a system is independent of the basis of functions we choose. To highlight the usefulness of adopting the concept of an abstract vector space (from which all physical representations of the state vectors could be obtained), the physicist Dirac introduced the "bra-ket" notation of Quantum mechanics:

Ket: $|\psi\rangle$ belongs to (i.e, is an element of) an abstract Hilbert state vector space

Bra: $\langle\psi|$ belongs to an abstract Hilbert space dual to that whose elements are the Kets $|\psi\rangle$

Scalar/Inner Product: defined between members of the two dual Hilbert spaces $\langle\phi|\psi\rangle$

Expectation Value of an operator: $\langle\hat{\theta}\rangle = \langle\psi|\hat{\theta}|\psi\rangle \equiv \langle\psi|\hat{\theta}\psi\rangle$

Basis choice: $\psi(\vec{r}, t) = \langle\vec{r}, t|\psi\rangle$ (Position Representation) , $\psi(\vec{p}, t) = \langle\vec{p}, t|\psi\rangle$ (Momentum Representation) .

Thus, while the Kets are independent of any particular basis choice, they represent the system completely, and hence knowing the Hilbert space $\{|\psi\rangle\}$ means knowing all its amplitudes (read wavefunctions) in all possible representations.

13.3 Properties of $|\psi\rangle$, $\langle\psi|$ and $\langle\phi|\psi\rangle$

1. To every $|\psi\rangle$, there exists a unique $\langle\psi|$ and vice versa:

$$|\psi\rangle \leftrightarrow \langle\psi| \tag{13.12}$$

There is a one-to-one correspondence between bras and kets

$$a|\psi\rangle + b|\phi\rangle \quad \leftrightarrow \quad a^*\langle\psi| + b^*\langle\phi| \quad a, b \in \mathcal{C} \tag{13.13}$$

$$|a\psi\rangle = a|\psi\rangle \quad \& \quad \langle a\psi| = a^*\langle\psi| \ . \tag{13.14}$$

2. In Quantum Mechanics, as the scalar product is a complex number in general, the ordering matters, i.e., $\langle\psi|\phi\rangle$ is not necessarily the same as $\langle\phi|\psi\rangle$. Then,

$$\begin{aligned}\langle\phi|\psi\rangle^* = (\int d\vec{r}\phi^*(\vec{r},t)\psi(\vec{r},t))^* &= \int d\vec{r}\psi^*(\vec{r},t)\phi(\vec{r},t) = \langle\psi|\phi\rangle \\ \text{i.e., } \langle\phi|\psi\rangle^* &= \langle\psi|\phi\rangle \ . \qquad (13.15)\end{aligned}$$

For the special case of $|\psi\rangle$, $|\phi\rangle \in \mathcal{R}$, $\langle\phi|\psi\rangle^* = \langle\phi|\psi\rangle = \langle\psi|\phi\rangle$.
Further,

$$\begin{aligned}\langle\psi|a_1\psi_1 + a_2\psi_2\rangle &= a_1\langle\psi|\psi_1\rangle + a_2\langle\psi|\psi_2\rangle \qquad (13.16)\\ \langle a_1\phi_1 + a_2\phi_2|\psi\rangle &= a_1^*\langle\phi_1|\psi\rangle + a_2^*\langle\phi_2|\psi\rangle \qquad (13.17)\\ \langle a_1\phi_1 + a_2\phi_2|b_1\psi_1 + b_2\psi_2\rangle &= a_1^*b_1\langle\phi_1|\psi_1\rangle + a_1^*b_2\langle\phi_1|\psi_2\rangle + a_2^*b_1\langle\phi_2|\psi_1\rangle + a_2^*b_2\langle\phi_2|\psi_2\rangle (13.18)\end{aligned}$$

3. For any state vector $|\psi\rangle$ of a Hilbert space $\mathcal{H}$, the norm ("length") $\langle\psi|\psi\rangle$ is real and positive definite

$$\langle\psi|\psi\rangle \geq 0 \ \ , \ \ \langle\psi|\psi\rangle \in \mathcal{R} \ . \qquad (13.19)$$

Further, is the state $|\psi\rangle$ is normalised, $\langle\psi|\psi\rangle = 1$.

4. For two states $|\psi\rangle$ and $|\phi\rangle$ such that $\langle\psi|\phi\rangle = 0$, the two states are said to be "orthogonal".

If in addition, $\langle\psi|\psi\rangle = 1 = \langle\phi|\phi\rangle$, the two states are said to be "orthonormal".

5. Schwarz Inequality: $|\langle\psi|\phi\rangle|^2 \leq \langle\psi|\psi\rangle\langle\phi|\phi\rangle$ (equality holds for $|\psi\rangle = a|\phi\rangle$, $a \in \mathcal{C}$).
Analogy with relation between vectors in Euclidean space: $|\vec{A}\cdot\vec{B}|^2 \leq |\vec{A}|^2|\vec{B}|^2$.

6. Triangle Inequality: $\sqrt{\langle\psi+\phi|\psi+\phi\rangle} \leq \sqrt{\langle\psi|\psi\rangle} + \sqrt{\langle\phi|\phi\rangle}$ (equality for $|\psi\rangle = a|\phi\rangle$, $a \in \mathcal{R}$ and $a > 0$).
Analogy with relation between vectors in Euclidean space: $|\vec{A}+\vec{B}| \leq |\vec{A}| + |\vec{B}|$.

Note that products like $|\psi\rangle|\phi\rangle$ and $\langle\psi|\langle\phi|$ are meaningless if $|\psi\rangle$ and $|\phi\rangle$ belong to the same Hilbert space $\mathcal{H}$. However, if $|\psi\rangle$ and $|\phi\rangle$ belong to different Hilbert spaces (e.g., real space wavefunctions and Spin angular momentum states), then products such as $|\psi\rangle|\phi\rangle$ and $\langle\psi|\langle\phi|$ are meaningful and are called "Direct Product".

Physical meaning of $\langle\psi|\phi\rangle$: First, $\langle\psi|\phi\rangle$ represents the projection of the state vector $|\psi\rangle$ onto the state vector $|\phi\rangle$. Second, for the states $|\psi\rangle$ and $|\phi\rangle$ being normalised, and from Born's probabilistic interpretation of Quantum Mechanics, the quantity $\langle\psi|\phi\rangle$ represents the probability amplitude that the system's initial state $|\phi\rangle$ will, after measurement, be projected onto another state $|\psi\rangle$.

13.4 Operators

1. Operators are mathematical rules such that

$$\hat{A}|\psi\rangle = |\psi'\rangle \ \ , \ \ \langle\phi|\hat{A} = \langle\phi'| \ . \qquad (13.20)$$

2. Products of Operators: In general, the product

$$\begin{aligned}
[\hat{A},\hat{B}] = \hat{A}\hat{B} - \hat{B}\hat{A} &\neq 0 && (13.21)\\
\Rightarrow \hat{A}\hat{B} &\neq \hat{B}\hat{A} \ . && (13.22)\\
\text{Further,}\ \ \hat{A}\hat{B}\hat{C} &= \hat{A}(\hat{B}\hat{C}) = (\hat{A}\hat{B})\hat{C} && (13.23)\\
\hat{A}^m\hat{A}^n &= \hat{A}^{m+n} && (13.24)\\
\hat{A}\hat{B}\hat{C}|\psi\rangle &= \hat{A}\hat{B}(\hat{C}|\psi\rangle) = \hat{A}\{\hat{B}(\hat{C}|\psi\rangle)\} \ . && (13.25)
\end{aligned}$$

3. Linearity:

$$\begin{aligned}
\hat{A}(a_1|\psi_1\rangle + a_2|\psi_2\rangle) &= a_1\hat{A}|\psi_1\rangle + a_2\hat{A}|\psi_2\rangle && (13.26)\\
(a_1\langle\psi_1| + a_2\langle\psi_2|)\hat{A} &= a_1\langle\psi_1|\hat{A} + a_2\langle\psi_2|\hat{A} \ . && (13.27)
\end{aligned}$$

4. Expectation Value:

$$\begin{aligned}
\langle\hat{A}\rangle = \frac{\langle\psi|\hat{A}|\psi\rangle}{\langle\psi|\psi\rangle} &\equiv \langle\psi|\hat{A}|\psi\rangle \ \ \text{(if } |\psi\rangle \text{ is normalised)} && (13.28)\\
&= \langle\psi|\hat{A}\psi\rangle \ . && (13.29)
\end{aligned}$$

5. Projection Operators: $|\phi\rangle\langle\psi|$ is a linear operator

$$\begin{aligned}
|\phi\rangle\langle\psi|\psi'\rangle &= \langle\psi|\psi'\rangle|\phi\rangle \ , && (13.30)\\
|\psi\rangle\langle\psi|\psi'\rangle &= \langle\psi|\psi\rangle|\psi\rangle \ , && (13.31)
\end{aligned}$$

where $\langle\psi|\psi'\rangle \in \mathcal{C}$ is the probability amplitude arising from the projection.

6. Operations such as $|\psi\rangle\hat{A}$ and $\hat{A}\langle\psi|$ are not sensible.

7. The operator "adjoint" to $\hat{\theta}$ is $\hat{\theta}^\dagger$ such that

$$\langle\phi|\hat{\theta}\psi\rangle = \langle\hat{\theta}^\dagger\phi|\psi\rangle \ . \qquad (13.32)$$

For example, let us work out the adjoint of the differential operator $\hat{D} \equiv \frac{d}{dx}$, i.e., $\hat{D}^\dagger$.

$$\begin{aligned}
\langle\phi|\hat{D}\psi\rangle &= \int_{-\infty}^{\infty} dx\ \phi^*(x)\frac{d\psi(x)}{dx} \quad \text{(in position basis)} && (13.33)\\
&= [\phi^*(x)\psi(x)]_{-\infty}^{\infty} - \int_{-\infty}^{\infty} dx\ \frac{d\phi^*(x)}{dx}\psi(x) && (13.34)\\
&= -\int_{-\infty}^{\infty} dx\ \frac{d\phi^*(x)}{dx}\psi(x) \quad (\text{as } \phi(x), \psi(x) \to 0 \text{ as } x \to \pm\infty) && (13.35)\\
\Rightarrow \langle\phi|\hat{D}\psi\rangle &= -\int_{-\infty}^{\infty}(-\hat{D}\phi^*(x))\psi(x) && (13.36)\\
&= \langle -\hat{D}\phi|\psi\rangle \equiv \langle\hat{D}^\dagger\phi|\psi\rangle && (13.37)\\
\Rightarrow \hat{D}^\dagger &= -\hat{D} \ . && (13.38)
\end{aligned}$$

Adjoint Operators have the following properties:

$$
\begin{aligned}
(i)\ \ (c\hat{P})^\dagger &= c^*\hat{P}^\dagger \quad (c \in \mathcal{C}) && (13.39)\\
(ii)\ \ (\hat{P}+\hat{Q})^\dagger &= \hat{P}^\dagger + \hat{Q}^\dagger && (13.40)\\
(iii)\ \ (\hat{P}\hat{Q})^\dagger &= \hat{Q}^\dagger\hat{P}^\dagger && (13.41)\\
(iv)\ \ (\hat{P}^\dagger)^\dagger &= \hat{P}\ . && (13.42)
\end{aligned}
$$

Relation (iii) can also be seen as follows:

$$
\begin{aligned}
\langle\phi|\hat{P}(\hat{Q}\psi)\rangle &= \langle\hat{P}^\dagger\phi|\hat{Q}\psi\rangle && (13.43)\\
&= \langle\hat{Q}\hat{P}\phi|\psi\rangle\ . && (13.44)
\end{aligned}
$$

8. Self-adjoint or "Hermitian" Operators: $\hat{\theta}^\dagger = \hat{\theta}$!
For example, the position operator $\hat{x}$ is Hermitian: $\hat{x}^\dagger = \hat{x}$.
The momentum operator $\hat{p} = -i\hbar\frac{d}{dx}$ is also Hermitian:

$$
\hat{p}^\dagger = -\hat{p}^* = -(-i\hbar\frac{d}{dx})^* = \hat{p}\ . \qquad (13.45)
$$

Hermitian Operators are special, as their expectation values are real-valued:

$$
\begin{aligned}
\langle\hat{Q}\rangle &= \langle\psi|\hat{Q}\psi\rangle && (13.46)\\
&= \langle\psi|\hat{Q}^\dagger\psi\rangle \quad (\text{for } \hat{Q} = \hat{Q}^\dagger) && (13.47)\\
&= \langle\hat{Q}\psi|\psi\rangle && (13.48)\\
&= \langle\hat{Q}\rangle^*\ , && (13.49)\\
\Rightarrow \langle\hat{Q}\rangle &\in \mathcal{R}\ . && (13.50)
\end{aligned}
$$

Similarly, if $|\psi\rangle$ is an eigenfunction of a Hermitian operator $\hat{Q}$ with eigenvalue q, i.e., $\hat{Q}|\psi\rangle = q|\psi\rangle$,

$$
\begin{aligned}
\langle\psi|\hat{Q}\psi\rangle &= q\langle\psi|\psi\rangle = q && (13.51)\\
\text{But }\ \langle\psi|\hat{Q}\psi\rangle &= \langle\hat{Q}^\dagger\psi|\psi\rangle && (13.52)\\
&= \langle\hat{Q}\psi|\psi\rangle = q^*\langle\psi|\psi\rangle = q^* && (13.53)\\
\Rightarrow q &\in \mathcal{R}\ . && (13.54)
\end{aligned}
$$

Finally, writing

$$
\begin{aligned}
\langle\hat{Q}\rangle &= \langle\psi|\hat{Q}|\psi\rangle \ \ (\text{where the states } |\psi\rangle \text{ are not eigenstates of } \hat{Q}) && (13.55)\\
&= \sum_{m,n}\langle\psi|\xi\rangle_{mm}\langle\xi|\hat{Q}|\xi\rangle_{nn}\langle\xi|\psi\rangle \ (\text{where we've introduced the identity operator} && (13.56)\\
&\qquad \mathrm{I} = \sum_n |\xi\rangle_{nn}\langle\xi| \ \ \text{in terms of eigenstates of } \hat{Q},\ |\xi\rangle_n) && (13.57)\\
&= \sum_{m,n}\int\int dxdy\ \xi_m^*(x)\delta_{m,n}Q_n\xi_n(y)\ (\text{as } \hat{Q}|\xi\rangle_n = Q_n\xi\rangle_n = Q_n) && (13.58)\\
& && (13.59)
\end{aligned}
$$

Chapter 14

The Angular Momentum problem: first passage

14.1 Introduction

For a classical particle, the angular momentum vector is given by

$$\mathbf{L} = \mathbf{r} \times \mathbf{p} \tag{14.1}$$

In quantum mechanics, the variables $\mathbf{r}$, $\mathbf{p}$ are replaced by operators $\mathbf{R}$, $\mathbf{p} = -i\hbar\nabla$.

$$\begin{aligned} \mathbf{L} &= \mathbf{r} \times \mathbf{p} \\ &= L_x\hat{x} + L_y\hat{y} + L_z\hat{z} \end{aligned} \tag{14.2}$$

Here,

$$\begin{aligned} L_x &= yp_z - zp_y \\ L_y &= zp_x - xp_z \\ L_z &= xp_y - yp_x \end{aligned}$$

Are the components of $\mathbf{L}$ *Hermitian?*

$$\begin{aligned} L_x^\dagger &= (yp_z)^\dagger - (zp_y)^\dagger = p_z^\dagger y^\dagger - p_y^\dagger z^\dagger \\ &= p_z y - p_y z \\ &= yp_z - zp_y = L_x \\ \mathbf{L}^\dagger &= \mathbf{L} \end{aligned} \tag{14.3}$$

Let us try computing $[L_x, L_y]$ now. First note that

$$[R_\alpha, p_\beta] = ih\delta_{\alpha\beta} \tag{14.4}$$

$$[A + B, C] = [A, C] + [B, C] \tag{14.5}$$

$$[AB, C] = A\,[B, C] + [A, C]\,B \tag{14.6}$$

Thus,

$$\begin{aligned}[L_x, L_y] &= [yp_z - zp_y, zp_x - xp_z] \\ &= [yp_z, zp_x] - [zp_y, zp_x] - [yp_z, xp_z] + [zp_y, xp_z] \\ &= y\,[p_z, zp_x] + [z, xp_z]\,p_y \quad \text{(as the second and third commutators vanish identically)} \\ &= y\,[p_z, z]\,p_x + x\,[z, p_z]\,p_y \\ &= i\hbar\,(xp_y - yp_x) = i\hbar L_z \qquad (14.7)\end{aligned}$$

Similarly, we can show that

$$[L_y, L_z] = i\hbar L_x \ , \qquad (14.8)$$

$$[L_z, L_x] = i\hbar L_y \ . \qquad (14.9)$$

Thus, simultaneous eigenstates of L_x, L_y, L_z(or even any pair from these 3) do not exist! This means that we can neither measure all three components precisely at once, nor any two components simultaneously. We can only make precise measurements for any one component of the angular momentum at a given instant.

We usually choose L_z for convenience, but this is still a matter of choice. Also, note that these three commutation relations show that L_x, L_y and L_z form, say, a right-handed triad (recall the right-hand corkscrew rule!) so as to remember the commutation relations as a mnemonic: a commutation relation that involves an anti-clockwise motion on the triad comes with a $+i\hbar$ factor in product with the third member of the triad, while a commutation relation that involves a clockwise motion on the triad comes with a $-i\hbar$ in product with the third member of the triad.

However, note that

$$\begin{aligned}\left[L^2, L_z\right] &= \left[L_x^2 + L_y^2 + L_z^2, L_z\right] \\ &= \left[L_x^2, L_z\right] + \left[L_y^2, L_z\right] \\ &= L_x\,[L_x, L_z] + [L_x, L_z]\,L_x + L_y\,[L_y, L_z] + [L_y, L_z]\,L_y \\ &= -i\hbar L_x L_y - i\hbar L_y L_x + i\hbar L_y L_x + i\hbar L_x L_y = 0 \qquad (14.10)\end{aligned}$$

L^2 is the square of the “length” of the angular momentum vector. Given that L^2 and L_z operators commute, we can define simultaneous eigenfunctions for them. In this way, we can simultaneously measure them both precisely, and these eigenvalues are “good quantum numbers”. Then what are the eigenvalues related to these operators? For this, we will fall back to on the “ladder operators” formalism for angular momentum. (Recall that we encountered such objects in our discussion of the algebraic solution for the simple harmonic oscillator problem a few chapters ago.) By the way, it can be easily shown that $[L^2, L_x] = 0$ and $[L^2, L_y] = 0$, so that we could have equivalently chosen the simultaneous eigenstates for the pair (L^2, L_x) or (L^2, L_y).

14.2 Total Angular Momentum

First, we should clarify that in any atomic system, the total angular momentum **J** is given by the vector sum

$$\begin{aligned} \mathbf{J} &= \mathbf{L} + \mathbf{S} \qquad (14.11)\\ \mathbf{L} &= \text{Orbital Angular Momentum}\\ \mathbf{S} &= \text{Spin Angular Momentum} \end{aligned}$$

Here, **L** literally relates to the classical orbital motion of a particle while **S** is an internal degree of freedom and is quite different from the classical picture of rotational motion. Spin angular momentum arises from the internal structure of the state (i.e., Hilbert) space for the Lorentz transformation invariant relativistic Dirac equation. Thus, a more complete discussion of spin lies well beyond the purview of these lectures. We will, therefore, have to be content for now with a discussion of spin as simply a source of angular momentum that, unlike orbital angular momentum, cannot be visualised in terms of rotations in real space.

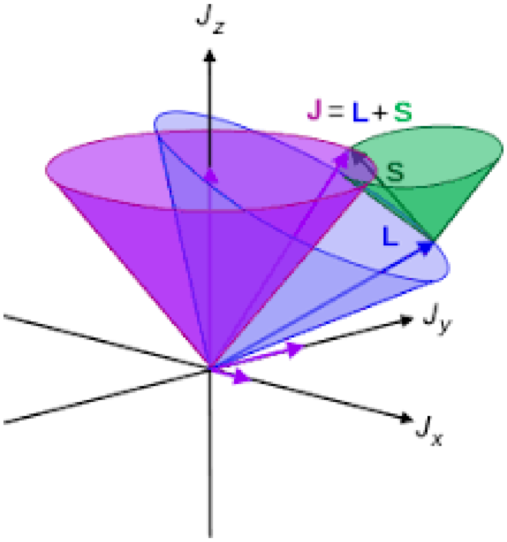


Figure 14.1: The angular momentum in quantum mechanics. Source: the internet.

Generally, in the absence of any physics that couples the **L** and **S** angular momenta, the commutation relations given earlier for the components of **L** go through, as well as a similar algebra is found for **S**. Thus, we may write

$$[J_x, J_y] = i\hbar J_z \qquad (14.12)$$

$$[J_y, J_z] = i\hbar J_x \qquad (14.13)$$

$$[J_z, J_x] = i\hbar J_y \qquad (14.14)$$

$$\left[J^2, J_\alpha\right] = 0 \qquad \text{for } \alpha = x,\ y,\ z \qquad (14.15)$$

We are going to work here onwards with the eigenstates of J^2 and J_z. For them, we define the ladder operators as

$$J_+ = J_x + iJ_y \qquad \text{Creation (raising) operator} \tag{14.16a}$$

$$J_- = J_x - iJ_y \qquad \text{Annihilation (lowering) operator} \tag{14.16b}$$

$$[J_+, J_-] = -i\,[J_x, J_y] + i\,[J_y, J_x] = 2\hbar J_z\ . \tag{14.16c}$$

Note that $J_+^\dagger = J_-$, $J_-^\dagger = J_+$ which means that J_+, J_- are not Hermitian, i.e., they do not correspond to observable quantities. But they can be used to go between different eigenfunctions (recall the ladder operators of the harmonic oscillator!).

For instance, starting with $J^2\psi = \alpha\psi$, $J_z\psi = \beta\psi$, assuming $J_+\psi = \phi$ (where ϕ is another normalised eigenstate of J^2 and J_z), let us try calculating $J^2\phi$, $J_z\phi$. Note that

$$\left[J^2, J_\pm\right] = 0 \qquad \text{as} \qquad \left[J^2, J_x\right] = 0 = \left[J^2, J_y\right]\ . \tag{14.17}$$

Thus,

$$J^2\phi = J^2\left(J_+\psi\right) = J_+\left(J^2\psi\right) = \alpha J_+\psi = \alpha\phi\ .$$

This shows that ϕ is an eigenstate of J^2 with the same eigenvalue α as that of ψ, i.e., the action of J_+ on ψ does not affect the eigenvalue of J^2 ! However, $[J_z, J_+] \neq 0$. In fact,

$$\begin{aligned} [J_z, J_+] &= [J_z, J_x + iJ_y] \\ &= [J_z, J_x] + i\,[J_z, J_y] \\ &= i\hbar J_y + i\,(-i\hbar)\,J_x \\ &= \hbar\,(J_x + iJ_y) = \hbar J_+ \qquad (14.18) \\ [J_z, J_-] &= \hbar J_- \qquad (14.19) \end{aligned}$$

Calculating $J_z\phi$:

$$\begin{aligned} J_z\phi = J_zJ_+\psi &= J_+J_z\psi + \hbar J_+\psi \\ &= \beta J_+\psi + \hbar J_+\psi \\ &= (\beta + \hbar)\,\phi \end{aligned} \tag{14.20}$$

This means J_+ transforms ψ (with eigenvalues α, β for J^2, J_z operators) into a new eigenfunction ϕ with eigenvalues α, $\beta + \hbar$. One more operation with J_+ will further increase the eigenvalue of J_z to $\beta + 2\hbar$ and so on.

Can we do this indefinitely (as done in simple harmonic oscillator)? No! This is because β^2 must have an upper bound that is defined in some way by α (the square of the length of the total angular momentum vector). This is simply seen from the fact that even classically, the z component of a vector is bounded by the magnitude of vector itself, i.e., $J_z^2 \leq J^2$.

More can be learnt about this upper bound on β by taking an eigenfunction ψ_{max} such that

$$J_+\psi_{max} = 0 \tag{14.21}$$

In the same way, we can argue that there must exist a ψ_{min} such that

$$J_{-}\psi_{min} = 0 \tag{14.22}$$

First, note that we can write J^2 in terms of $J_\pm$ and J_z as follows

$$\begin{aligned} J^2 &= J_x^2 + J_y^2 + J_z^2 && (14.23a)\\ &= J_-J_+ + J_z^2 + \hbar J_z && (14.23b)\\ &= J_+J_- + J_z^2 - \hbar J_z \ , && (14.23c) \end{aligned}$$

where we have used the fact that $J_x^2 + J_y^2 = J_+J_- - \hbar J_z = J_-J_+ + \hbar J_z$.

Eqs. 14.23 are obtained by using the definition and commutation relations of $J_\pm$. Now, we can compute the values of α and β as follows

$$\begin{aligned} J^2\psi_{max} &= \left(J_-J_+ + J_z^2 + \hbar J_z\right)\psi_{max}\\ \alpha\psi_{max} &= 0 + \beta_{max}^2\psi_{max} + \hbar\beta_{max}\psi_{max}\\ \therefore \alpha &= \beta_{max}^2 + \hbar\beta_{max} && (14.24) \end{aligned}$$

Similarly, we find by operating $J^2\psi_{min}$

$$\alpha = \beta_{min}^2 - \hbar\beta_{min} \tag{14.25}$$

From Eqs. 14.24, 14.25,

$$\begin{aligned} \beta_{min}^2 - \hbar\beta_{min} &= \beta_{max}^2 + \hbar\beta_{max}\\ \beta_{max}^2 - \beta_{min}^2 + \hbar\left(\beta_{max} + \beta_{min}\right) &= 0\\ \left(\beta_{max} - \beta_{min} + \hbar\right)\left(\beta_{max} + \beta_{min}\right) &= 0\\ \implies \beta_{max} &= -\beta_{min} && (14.26) \end{aligned}$$

as the other solution $\beta_{max} = \beta_{min} - \hbar$ does not make physical sense. Also, from the idea of ladder operators for J_z,

$$\beta_{max} = \beta_{min} + n\hbar, \quad n \in \mathbb{Z} \tag{14.27}$$

$$\therefore \beta_{max} = \frac{n\hbar}{2} \tag{14.28}$$

The first line of the above equation is simply telling us that applying the raising operator an integer n number of times should take us from the state ψ_{min} to the state ψ_{max}. Now, defining $j = n/2$, $n \in \mathbb{Z}$, from Eq. 14.24, we obtain

$$\alpha = \hbar^2 j\left(j+1\right), \tag{14.29}$$

$$\beta = m_j\hbar, \qquad m_j = -j, -j+1, \cdots, j-1, j \tag{14.30}$$

Therefore, we can write

$$\begin{aligned} J^2\psi &= \hbar^2 j\left(j+1\right)\psi \ , && \text{where } \ j = \frac{n}{2},\ n \in \mathbb{Z} \ , && (14.31)\\ J_z\psi &= m_j\hbar\psi \ , && \text{where } \ m_j = -j, -j+1, \cdots, j-1, j \ . && (14.32) \end{aligned}$$

These relations are visualised in Fig.14.2 below. It is important to note that the integer parameter n has not been determined as yet. It turns out that nature appears to divide the odd valued integers and even valued integers for n into the world of Fermions (all matter particles) and Bosons (all force particles) respectively!

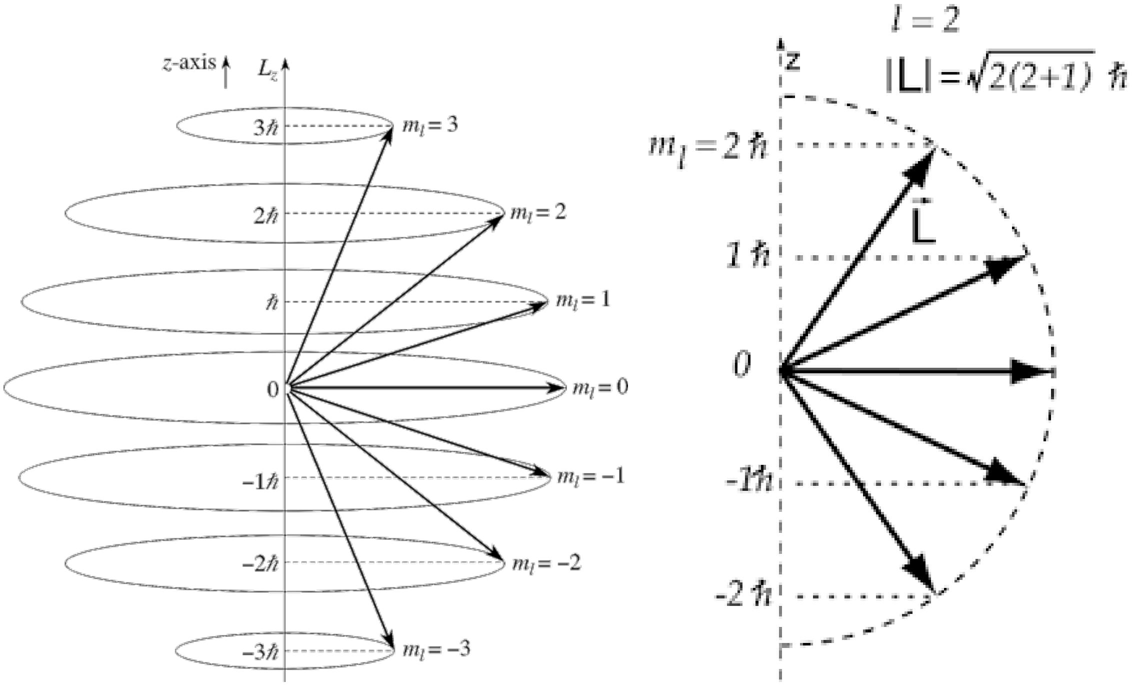


Figure 14.2: The angular momentum in quantum mechanics for the case of $n = 6$, $l = 3$ (left) and $n = 4$, $l = 2$ (right). Source: the internet.

Finally, note that for $J_\pm \psi_{j,m_j} = C_{j,m_j} \psi_{j,m_j \pm 1}$, the constant C_{j,m_j} can be obtained by using the relations eq.(14.23b) and eq.(14.23c)

$$C_{j,m_j} = \hbar \sqrt{j(j+1) - m_j(m_j \pm 1)}\ . \tag{14.33}$$

Chapter 15

Orbital Angular Momentum Eigenstates

As rotational symmetry in three dimensions is best visualised in spherical polar coordinates, we write the TISE in that coordinate system to obtain the eigenfunctions and eigenvalues of orbital angular momentum.

First, recall that the gradient in spherical coordinates is given by

$$\nabla = \hat{e}_r \frac{\partial}{\partial r} + \hat{e}_\theta \frac{1}{r} \frac{\partial}{\partial \theta} + \hat{e}_\phi \frac{1}{r \sin\theta} \frac{\partial}{\partial \phi} \tag{15.1}$$

! Revise the transformation of the gradient from Cartesian to spherical coordinates.

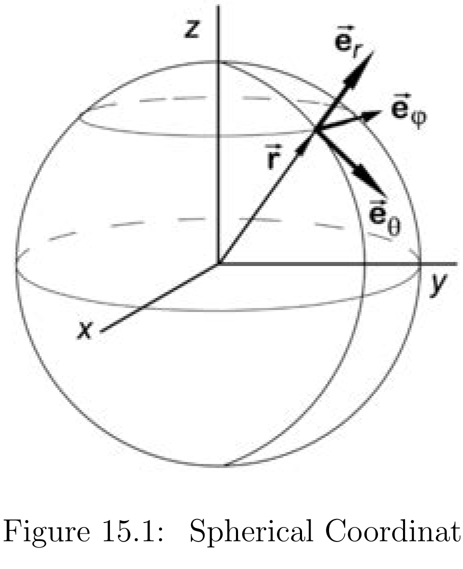


Figure 15.1: Spherical Coordinates

Then one can compute the components of the orbital angular momentum L_x, L_y, L_z

$$L_x = i\hbar\left(\sin\phi\frac{\partial}{\partial\theta} + \cot\theta\cos\phi\frac{\partial}{\partial\phi}\right) \ , \tag{15.2}$$

$$L_y = i\hbar\left(-\cos\phi\frac{\partial}{\partial\theta} + \cot\theta\sin\phi\frac{\partial}{\partial\phi}\right) \ , \tag{15.3}$$

$$L_z = -i\hbar\frac{\partial}{\partial\phi} \ . \tag{15.4}$$

! Work these out for yourself carefully, following the definitions for the components of $\vec{L}$ given in the previous chapter.

This gives us the norm of, and the raising and lowering operators for, the orbital angular momentum as

$$L^2 = -\hbar^2\left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2}\right] \ , \tag{15.5}$$

$$L_\pm = \hbar e^{\pm i\phi}\left(\pm\frac{\partial}{\partial\theta} + i\cot\frac{\partial}{\partial\phi}\right) \ . \tag{15.6}$$

Thus, from our ladder operator analysis presented in the last chapter, we know that we can define eigenfunctions that are labelled by two good quantum l, m such that they satisfy the following eigenvalue equations

$$L^2\psi_{lm}\left(\theta,\phi\right) = \hbar^2 l\left(l+1\right)\psi_{lm}\left(\theta,\phi\right) \ , \tag{15.7a}$$

$$L_z\psi_{lm}\left(\theta,\phi\right) = m\hbar\psi_{lm}\left(\theta,\phi\right) \ . \tag{15.7b}$$

This can be written as

$$\left[\frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2} + \frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right)\right]\psi_{lm} = -l\left(l+1\right)\psi_{lm} \ , \tag{15.8}$$

$$\frac{\partial}{\partial\phi}\psi_{lm} = im\psi_{lm} \ . \tag{15.9}$$

The above prompts us to try the following trial solution using separation of variables: $\psi_{lm}\left(\theta,\phi\right) = \Theta\left(\theta\right)\Phi\left(\phi\right)$.

From eq.(15.9), it is easily seen that

$$\Phi\left(\phi\right) = e^{im\phi} \tag{15.10}$$

! Verify this by plugging it into eq.(15.9). Importantly, the solution for $\Phi(\phi)$ is simply a phase that can only be known modulo 2π. This means that since ϕ is an angular variable, the continuity of ψ_{lm} demands

$$\Phi\left(\phi+2\pi\right) = \Phi\left(\phi\right) \quad \text{(Periodic Boundary Condition)} \tag{15.11}$$

$$\text{i.e., } e^{im2\pi} = 1 \implies m \in \mathbb{Z} \tag{15.12}$$

! Note how the imposition of periodic boundary conditions leads immediately to the quantisation of the quantum number m. Further, it indicates that the orbital angular momentum orbits shown in Fig.14.2 must have a quantised period (m) for the matter waves. Recall that this was the assumption made by Bohr in formulating his solution to the Hydrogen atom problem.

From eq.(14.24), we know that $l = 0, 1, 2, \cdots$. Then, putting eq.(15.10) in qq.(15.8) gives the "Legendre" differential equation

$$\left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) - \frac{m^2}{\sin^2\theta} + l\left(l+1\right)\right]\Theta\left(\theta\right) = 0 \ . \tag{15.13}$$

The solutions to the above Legendre differential equation are the well-known **Legendre polynomials**. In this way, we find that the $\psi_{lm}\left(\theta,\phi\right)$ are the spherical harmonics $Y_{lm}\left(\theta,\phi\right)$:

$$\psi_{lm}\left(\theta,\phi\right) = Y_{lm}\left(\theta,\phi\right) = \left(-1\right)^{(m+|m|)/2}\left[\frac{2l+1}{4\pi}\frac{(l-|m|)!}{(l+|m|)!}\right]^{1/2} \times e^{im\phi} \times P_{l|m|}\left(\cos\theta\right) \ , \tag{15.14}$$

where the $P_{lm}\left(\xi\right)$ ($m > 0$) are the Legendre functions (given below in eq.(15.15)) in terms of Legendre polynomials (given below in eq.(15.16)).

$$\text{Legendre Functions: } P_{lm}\left(\xi\right) = \left(1-\xi^2\right)^{m/2}\frac{d^m}{d\xi^m}P_l\left(\xi\right) \ , \quad \left(\xi = \cos\theta\right) \tag{15.15}$$

$$\text{Legendre Polynomials: } P_l\left(\xi\right) = \frac{1}{2^l l!}\frac{d^l}{d\xi^l}\left(\xi^2-1\right)^l \ . \tag{15.16}$$

The $P_l\left(\xi\right)$ are l^{th} order polynomials in ξ; the $P_{lm}\left(\xi\right)$ are thus $\left(l-m\right)^{th}$ order polynomials in ξ, and are multiplied by ξ^m. Further, the $P_{lm}\left(\xi\right)$ have $\left(l-m\right)$ nodes in $-1 < \xi < 1$. Some of the Legendre polynomials are given by

$$\begin{aligned}
P_0 &= 1 \ , \\
P_1 &= \xi \ , \\
P_2 &= \frac{1}{2}\left(3\xi^2 - 1\right) \ , \\
P_3 &= \frac{1}{2}\left(5\xi^3 - 3\xi\right) \ , \text{ and so on.}
\end{aligned}$$

Thus, we see that the $P_l(\xi)$ are either odd order or even order polynomials in ξ. Recall that we saw a similar phenomenon for the Hermite polynomial solutions for the wavefunctions of the harmonic oscillator problem as well. While the mathematical details given in the rest of this chapter are not essential for proceeding towards the Hydrogen atom problem, they are worth knowing for the overall completeness in our understanding of the angular momentum problem.

The Legendre polynomials satisfy the following recursion relation

$$(l+1)\,P_{l+1}(\xi) = (2l+1)\,\xi P_l(\xi) - lP_{l-1}(\xi) \ , \tag{15.17}$$

$$\left(1-\xi^2\right)\frac{dP_l}{d\xi} = l\left(P_{l-1} - \xi P_l\right) \ . \tag{15.18}$$

The associated Legendre functions satisfy

$$\text{Legendre differential equation :} \quad \left[\left(1-\xi^2\right)\frac{d^2}{d\xi^2} - 2\xi\frac{d}{d\xi} + l\,(l+1) - \frac{m^2}{\left(1-\xi^2\right)}\right] P_{lm}(\xi) = 0 \tag{15.19}$$

$$\text{Parity transformation :} \quad P_{lm}(-\xi) = (-1)^{l+m}\,P_{lm}(\xi) \tag{15.20}$$

$$\text{Orthogonality relation :} \quad \int\limits_{-1}^{-1} P_{lm}(\xi)\,P_{l'm}(\xi)\,d\xi = \frac{2}{(2l+1)}\frac{(l-m)!}{(l+m)!}\delta_{ll'} \tag{15.21}$$

In particular, $P_{l0}(\xi) = P_l(\xi)$ and $P_{ll}(\xi) = (2l-1)!!\,(1-\xi^2)^{1/2}$, where $(2l-1)!! = (2l-1)\,(2l-3)\cdots 1$. From these relations, we get the following properties of the spherical harmonics $Y_{lm}(\theta,\phi)$

$$\text{Orthogonality relation :} \quad \int\limits_0^{2\pi} d\phi \int\limits_0^{\pi} d\theta \ \sin\theta \ Y^*_{lm}(\theta,\phi)\,Y_{l'm'}(\theta,\phi) = \delta_{ll'}\delta_{mm'} \ , \tag{15.22}$$

$$\text{Completeness relation :} \quad \sum_{l=0}^{\infty}\sum_{m=-l}^{l} Y_{lm}(\theta,\phi)\,Y^*_{lm}(\theta',\phi') = (\sin\theta)^{-1}\,\delta(\theta-\theta')\,\delta(\phi-\phi') \ . \tag{15.23}$$

Eqs.(15.22) and (15.23) prove the orthogonality and completeness of the spherical harmonics $Y_{lm}(\theta,\phi)$, thus making them suitable candidate for an eigenbasis. Eq.(15.24) below is called the **addition theorem**, which is a generalization of $\cos(\theta-\theta') = \cos\theta'\cos\theta + \sin\theta'\sin\theta$:

$$\sum_{m=-l}^{l} Y_{lm}(\theta,\phi)\,Y^*_{lm}(\theta',\phi') = \frac{(2l+1)}{4\pi}P_l(\cos\theta) \ , \tag{15.24}$$

$$Y_{l,-m}(\theta,\phi) = (-1)^m\,Y^*_{lm}(\theta,\phi) \ . \tag{15.25}$$

For the unit vectors $\hat{a}$, $\hat{b}$, it can be shown that

$$P_l\left(\hat{a}\cdot\hat{b}\right) = \frac{4\pi}{2l+1}\sum_{m=-l}^{l} Y^*_{lm}(\theta',\phi')\,Y_{lm}(\theta,\phi) \ ,$$

$$\text{such that for } \hat{x} = \hat{y} \implies \sum_{m=-l}^{l} Y^*_{lm}(\theta',\phi')\,Y_{lm}(\theta,\phi) = \frac{2l+1}{4\pi} \ . \tag{15.26}$$

Eq.(15.26) is a generalization of $\cos^2\theta + \sin^2\theta = 1$ to 2D. Some of the spherical harmonics are given by

$$\begin{aligned}
Y_{00} &= \frac{1}{\sqrt{4\pi}} \ , \\
Y_{10} &= \sqrt{\frac{3}{4\pi}} \cos\theta \ , \\
Y_{11} &= -\sqrt{\frac{3}{8\pi}} \sin\theta e^{i\phi} \ , \\
Y_{20} &= \sqrt{\frac{5}{16\pi}} \left(3\cos^2\theta - 1\right) \ , \\
Y_{21} &= -\sqrt{\frac{15}{8\pi}} \sin\theta \cos\theta e^{i\phi} \ , \\
Y_{22} &= \sqrt{\frac{15}{32\pi}} \sin^2\theta e^{i2\phi} \ , \text{ and so on.}
\end{aligned}$$

Under the action of parity operator $\mathcal{P}\psi(\mathbf{r}) = \psi(-\mathbf{r})$,

$$\begin{aligned}
\mathcal{P}Y_{lm}(\theta,\phi) &= Y_{lm}(\pi-\theta, \pi+\phi) \\
&= e^{im\pi}(-1)^{l+|m|} Y_{lm}(\theta,\phi) \ , \qquad (15.27) \\
\therefore \ \mathcal{P}Y_{lm}(\theta,\phi) &= (-1)^l Y_{lm}(\theta,\phi) \ . \qquad (15.28)
\end{aligned}$$

This means that Y_{lm} is even for even l and odd for odd l. Figure 15.2 shows the spherical harmonics $Y_{lm}(\theta,\phi)$ for $l = 0, 1, 2, 3$.

States with $l = 0$: s-orbitals
$l = 1$: p-orbitals
$l = 2$: d-orbitals
$l = 3$: f-orbitals
and so on ...

Below, in Fig.15.2, we see plots of the first few Spherical Harmonics. You may well recall having seen these somewhere as the orbitals for the Hydrogen atom.

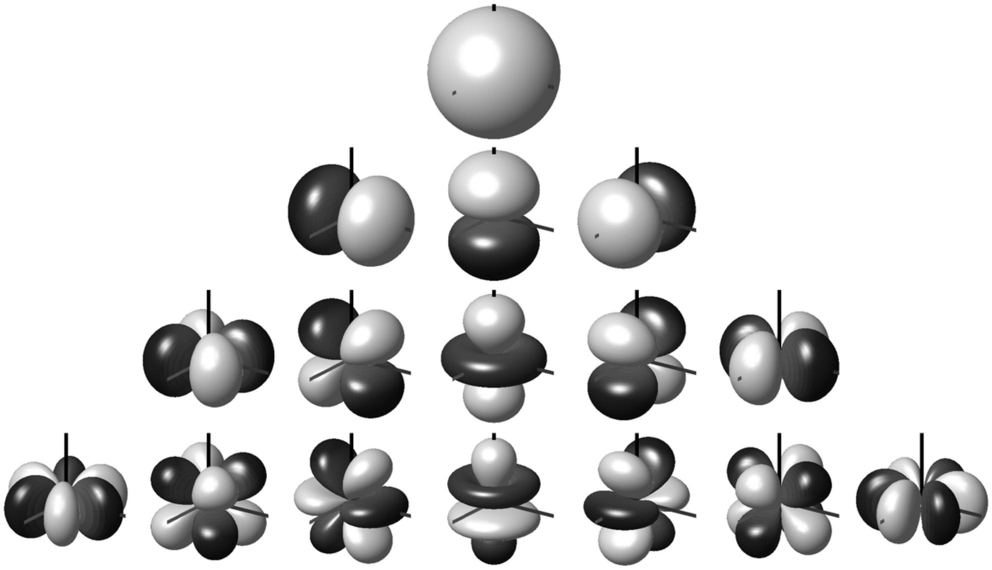


Figure 15.2: Spherical Harmonics

Chapter 16

Central Potential Problem

For a central force problem, $V(\mathbf{r}) = V(r)$ i.e., the potential only depends on the radial distance. Thus, the TISE will then be

$$H = \frac{1}{2m}\vec{p}^2 + V(r) \tag{16.1}$$

Now as $\mathbf{L} = \mathbf{r} \times \mathbf{p}$ such that $L_k = \epsilon_{ijk} x_i p_k$, the norm of the orbital angular momentum L^2 will be ($\mathbf{r} = x_1\hat{x} + x_2\hat{y} + x_3\hat{z}$)

$$\begin{aligned} L^2 = \sum_{k=x,y,z} L_k^2 &= \left(\epsilon_{ijk} x_i p_j\right)\left(\epsilon_{mnk} x_m p_n\right) \\ &= \epsilon_{ijk}\epsilon_{mnk} x_i p_j x_m p_n \\ &= \left(\delta_{im}\delta_{jn} - \delta_{in}\delta_{jm}\right) x_i p_j x_m p_n \\ &= x_i p_j x_i p_j - x_i p_j x_j p_i \\ &= x_i \left(x_i p_j - i\hbar\delta_{ij}\right) p_j - x_i p_j \left(p_i x_j + i\hbar\delta_{ij}\right) \\ &= x_i^2 p_j^2 - i\hbar x_i p_i - x_i p_i p_j x_j - i\hbar x_i p_i \\ &= x_i^2 p_j^2 - 2i\hbar x_i p_i - x_i p_i \left(x_j p_j - i\hbar\right) \\ &= x_i^2 p_j^2 - x_i p_i x_j p_j - i\hbar x_i p_i \\ &= r^2 p^2 - \left(\mathbf{r}\cdot\mathbf{p}\right)^2 - i\hbar\mathbf{r}\cdot\mathbf{p} \end{aligned} \tag{16.2}$$

where ϵ_{ijk} is the totally anti-symmetric tensor of rank 3 (i, j, k can each take the values $1, 2, 3$). ϵ_{ijk} has the following properties: (i) its value is zero for any two components being the same, all $\epsilon_{ijk} = 1$ for (i, j, k) arranged (say) in a clockwise manner and $\epsilon_{jik} = -\epsilon_{ijk} = \ldots$ (i.e., the value is -1 for any two indices being swapped).
Further, the product $\epsilon_{ijk}\epsilon_{mnk} = \left(\delta_{im}\delta_{jn} - \delta_{in}\delta_{jm}\right)$. Note that following the Eintein convention, we are summing over all repeated indices.
Finally, we have used the identity $(x_i p_j - p_j x_i) = i\hbar\delta_{ij}$. The last term on the RHS of Eq.

16.2 arises purely due to the non-commutativity of $\mathbf{r}$, $\mathbf{p}$. For $\mathbf{r} = r\hat{e}_r$ as $\hat{p} = i\hbar\boldsymbol{\nabla}$

$$\boldsymbol{\nabla} = \hat{e}_r\frac{\partial}{\partial r} + \hat{e}_\theta\frac{\partial}{\partial\theta} + \hat{e}_\phi\frac{1}{r\sin\theta}\frac{\partial}{\partial\phi} \tag{16.3}$$

$$\implies \mathbf{r}\cdot\mathbf{p} = -i\hbar\mathbf{r}\cdot\boldsymbol{\nabla} = -i\hbar r\frac{\partial}{\partial r}$$

From Eq. 16.2,

$$\begin{aligned}
\hat{p}^2 &= \frac{L^2}{r^2} - \frac{\hbar^2}{r^2}\left(\left(r\frac{\partial}{\partial r}\right)^2 + r\frac{\partial}{\partial r}\right)\\
&= \frac{L^2}{r^2} - \frac{\hbar^2}{r}\frac{\partial}{\partial r} - \frac{\hbar^2}{r}\frac{\partial}{\partial r}\left(r\frac{\partial}{\partial r}\right)\\
&= \frac{L^2}{r^2} - \hbar^2\left(\frac{2}{r}\frac{\partial}{\partial r} + \frac{\partial^2}{\partial r^2}\right) && (16.4)\\
\hat{p}^2 &= \frac{L^2}{r^2} + \hat{p}_r^2 && (16.5)
\end{aligned}$$

where $\hat{p}_r$ is the radial momentum given by

$$\begin{aligned}
\hat{p}_r &= -i\frac{\hbar}{r}\frac{\partial}{\partial r}\cdot r = -i\hbar\left(\frac{1}{r} + \frac{\partial}{\partial r}\right) && (16.6)\\
[r, \hat{p}_r] &= \left[r, -i\hbar\left(\frac{\partial}{\partial r} + \frac{1}{r}\right)\right]\\
&= i\hbar r\frac{\partial}{\partial r} - i\hbar + i\hbar\left(\frac{\partial}{\partial r} + \frac{1}{r}\right)r\\
&= i\hbar && (16.7)\\
& && (16.8)
\end{aligned}$$

Note that you have to be a little careful in carrying out the commutation relation here, so as not to miss out any terms. The best way to do this is to explicitly put in a wavefunction ψ on the right of the commutator, carry out the entire calculation, and remove it at the end.

$\boxed{\text{Verify that the operator } \hat{p}_r \text{ is Hermitian.}}$

Putting Eq. 16.4 in Eq. 16.1,

$$\left[-\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial r^2} + \frac{2}{r}\frac{\partial}{\partial r}\right) + \frac{1}{2mr^2}L^2 + V(r)\right]\psi(r,\theta,\phi) = E\psi(r,\theta,\phi) \tag{16.9}$$

For the case of a central potential, $V(r)$ has rotational symmetry. Then, $[H, L^2] = 0$, which means that $\psi(r,\theta,\phi)$ is a simultaneous eigenvalue of H, L^2 and L_z. So, we can take $\psi(r,\theta,\phi)$ as

$$\psi(r,\theta,\phi) = R(r)\,Y_{lm}(\theta,\phi) \quad \text{where} \quad L^2Y_{lm}(\theta,\phi) = l(l+1)\hbar^2Y_{lm}(\theta,\phi)\ , \tag{16.10}$$

$$\Rightarrow \left[-\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial r^2} + \frac{2}{r}\frac{\partial}{\partial r}\right) + \frac{\hbar^2}{2mr^2}l(l+1) + V(r)\right]R(r) = ER(r) \tag{16.11}$$

Taking $R\left(r\right) = u\left(r\right)/r$,

$$\left[-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial r^2} + \frac{\hbar^2}{2mr^2}l\left(l+1\right) + V\left(r\right)\right]u\left(r\right) = Eu\left(r\right) \tag{16.12}$$

$$\implies V_{eff}\left(r\right) = V\left(r\right) + \frac{\hbar^2}{2mr^2}l\left(l+1\right) \tag{16.13}$$

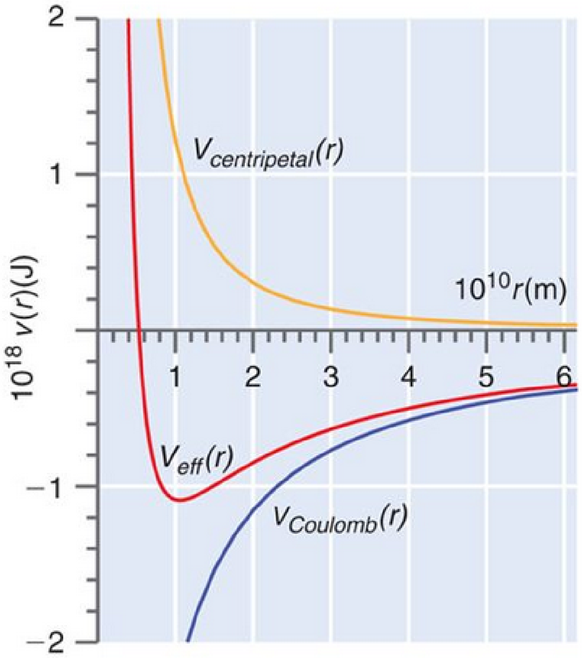


Figure 16.1: Effective potential for an atom with Coulomb and Centripetal potentials.

The plot in Fig.16.1 shows the effective potential V_{eff} for the Coulomb potential $V_{Coulomb}(r) \propto -1/r$ and the centripetal potential $V_{Centripetal} = l(l+1)/2mr^2$.

The normalizability condition of ψ tells us that

$$\int d^3x \left|\psi\right|^2 = \text{constant} \times \int\limits_0^\infty drr^2 \frac{\left|u\left(r\right)\right|^2}{r^2} < \infty$$

$$\implies \int\limits_0^\infty dr \left|u\left(r\right)\right|^2 < \infty$$

$$\implies \lim_{r\to\infty} \left|u\left(r\right)\right| \leq \frac{a}{r^{1/2+\epsilon}} \quad \text{where} \quad a = \text{constant}, \epsilon > 0$$

where the finite constant arise from the angular integrals $\int d\theta \int d\phi sin^2\theta |Y_{lm}\left(\theta,\phi\right)|^2$. Thus, $u\left(r\right)$ must fall away faster than $1/\sqrt{r}$ as $r \to \infty$. Also, for the case of $\epsilon = 0$, you can easily see a logarithmic divergence of the integral with r in the limit of $r \to \infty$.

In the limit of $r \to 0$, we must have $u(r \to 0) \to 0$ faster than $r \to 0$ for the overall stability of the atomic system. Certainly, this is easily seen for $l \neq 0$ due to the overall repulsive nature of $V_{eff}(r \to 0)$. As

$$u\,(0) = 0 \;\text{ in the limit }\; r \to 0 \;, \tag{16.14}$$

we must have (for $l = 0$) a hard-wall boundary condition at $r = 0$

$$V_{eff}\,(r) = \begin{cases} V(r) & r > 0 \;, \\ \infty & r = 0 \;. \end{cases} \tag{16.15}$$

Do bound states exist in 3D for the case of the Coulomb potential problem?

Having set the stage for solving the Hydrogen atom problem, we now turn to this task in the next chapter.

Chapter 17

The Hydrogen Atom

For the Coulomb potential $V(r) = -e^2/4\pi\epsilon_0 r$ (note: a central potential!), let us first consider the two-body problem of the electron orbiting around the positively charged nucleus. For a classical system of two bodies of masses m_1, m_2, at positions $\mathbf{r}_1$, $\mathbf{r}_2$, the problem can always be decomposed into center of mass and relative variables. The center of mass variables are given by

$$\begin{aligned}
\mathbf{r}_{CM} &= \frac{m_1\mathbf{r}_1 + m_2\mathbf{r}_2}{m_1+m_2} \ , \\
\mathbf{p}_{CM} &= \mathbf{p}_1 + \mathbf{p}_2 \ , \\
M &= m_1 + m_2 \ .
\end{aligned}$$

The relative variables are given by

$$\begin{aligned}
\mathbf{r} &= \mathbf{r}_1 - \mathbf{r}_2 \ , && (17.1) \\
\mathbf{p} &= \mu\left(\mathbf{v}_1 - \mathbf{v}_2\right) \ , \quad \text{Reduced mass } \mu = \left[\frac{1}{m_1} + \frac{1}{m_2}\right]^{-1} = \frac{m_1 m_2}{m_1+m_2} \ , \\
&= \frac{m_1 m_2}{m_1+m_2}\left(\frac{\mathbf{p}_1}{m_1} - \frac{\mathbf{p}_2}{m_2}\right) \ , \\
&= \frac{m_2\mathbf{p}_1 - m_1\mathbf{p}_2}{m_1+m_2} \ . && (17.2)
\end{aligned}$$

We now turn to the quantum mechanical case, where the position and momenta variables have been promoted to operators. Thus, denoting the particle index by μ, ν and the Cartesian components by i, j, we can write

$$\begin{aligned}
&\left[r_{\nu i}, p_{\mu j}\right] = i\hbar\delta_{ij}\delta_{\mu\nu} \ , && (17.3) \\
\implies &\left[\left(\mathbf{r}_{CM}\right)_i, \left(\mathbf{p}_{CM}\right)_j\right] = i\hbar\delta_{ij} = \left[r_i, p_j\right] \ . && (17.4)
\end{aligned}$$

Thus, we can simplify the TISE for the 2-body system of the H-atom using

$$\begin{aligned} \mathbf{p}_{CM} &= -i\hbar\boldsymbol{\nabla}_{CM} \ , \\ \mathbf{p}_r &= -i\hbar\boldsymbol{\nabla}_r \ , \\ \frac{p_1^2}{2m_1} + \frac{p_2^2}{2m_2} &= \frac{p_{CM}^2}{M} + \frac{p^2}{2\mu} \ , \\ \implies H\Psi\left(\mathbf{r}_{CM},\mathbf{r}\right) &= \left[\frac{p_{CM}^2}{M} + \frac{p^2}{2\mu} + V\left(r\right)\right]\Psi\left(\mathbf{r}_{CM},\mathbf{r}\right) = E_{total}\Psi\left(\mathbf{r}_{CM},\mathbf{r}\right) \ . \qquad (17.5) \end{aligned}$$

Now, given that there is no (external) potential experienced by the the 2-body system as a whole (i.e., there is no $V_{ext}(\mathbf{r}_{CM})$), the center of mass degree of freedom behaves effectively like a free particle, such that we can decompose the total wavefunction into $\Psi(\mathbf{r}_{CM},\mathbf{r}) = \phi(\mathbf{r}_{CM})\psi_{\mathbf{r}}$ such that

$$\begin{aligned} E_{CM} &= \frac{\hbar^2 k_{CM}^2}{2M} \ , \quad \phi\left(\mathbf{r}_{CM}\right) = e^{i\mathbf{k}_{CM}\cdot\mathbf{r}_{CM}} \ , \\ \left[\frac{p^2}{2\mu} + V\left(r\right)\right]\psi\left(\mathbf{r}\right) &= E_{rel}\psi\left(\mathbf{r}\right) \quad \text{where } E_{rel} = E_{total} - \frac{\hbar^2 k_{CM}^2}{2M} \ , \qquad (17.6) \end{aligned}$$

where $m_e = 9.109 \times 10^{-31}$ kg, $m_p = 1.672 \times 10^{-27}$ kg and $\mu = m_e m_p/M = 0.9995 m_e$ for the H-atom problem. We have already studied the case of a TISE in the radial coordinate for a general central potential in the previous chapter. Thus, the equivalent 1D TISE for the H-atom is

$$\left[-\frac{\hbar^2}{2\mu}\frac{d^2}{dr^2} + \frac{\hbar^2 l\left(l+1\right)}{2\mu r^2} - \frac{1}{4\pi\epsilon_0}\frac{e^2}{r}\right]u\left(r\right) = E_{rel}u\left(r\right) \ . \qquad (17.7)$$

Note that for $l = 0$,

$$V_{eff}\left(r\right) = \frac{\hbar^2 l\left(l+1\right)}{2\mu r^2} - \frac{1}{4\pi\epsilon_0}\frac{e^2}{r} \to -\frac{1}{4\pi\epsilon_0}\frac{e^2}{r} < 0 \ .$$

Given that the Coulomb potential is overall confining, any bound state (if at all possible!) must exist for $E_{rel} < 0$ at $l = 0$. Thus, let us take $\varepsilon = -E_{rel}$ and simplify the TISE to

$$\frac{d^2 u\left(r\right)}{dr^2} + \frac{2\mu e^2}{4\pi\epsilon_0\hbar^2}\frac{u\left(r\right)}{r} - \frac{l\left(l+1\right)}{r^2}u\left(r\right) = \frac{2\mu}{\hbar^2}\varepsilon u\left(r\right) \quad \text{where } \varepsilon > 0 \ . \qquad (17.8)$$

At $r \to \infty$ limit, we can drop the V_{eff} term and take the approximate form $u(r) \sim u_{app}(r)$

$$\begin{aligned} \frac{d^2 u_{app}}{dr^2} &= \frac{2\mu}{\hbar^2}\varepsilon u_{app} \\ \implies u_{app}\left(r\right) &\propto e^{-\sqrt{\frac{2\mu\varepsilon}{\hbar^2}}r} \ , \qquad (17.9) \end{aligned}$$

where we have thrown away the possibile solution of $u_{app}\left(r\right) \propto e^{\sqrt{\frac{2\mu\varepsilon}{\hbar^2}}r}$ on grounds of normalisability (as it diverges in the limit of $r \to \infty$). With this in mind, we consider the trial solution

$$u\left(r\right) = v\left(r\right) e^{-\sqrt{\frac{2\mu\varepsilon}{\hbar^2}}r} \ , \tag{17.10}$$

$$v\left(r\right) = \sum_{p=1}^{\infty} A_p r^p \ . \tag{17.11}$$

Here, A_p are constants, and $A_0 = 0$ (from the constraint that $u(r \to 0) \to 0$, eq.16.14). Substituting Eq. 17.10 in Eq. 17.8 we get

$$\frac{d^2 v\left(r\right)}{dr^2} - \frac{2\sqrt{2\mu\varepsilon}}{\hbar}\frac{dv\left(r\right)}{dr} + \frac{2\mu e^2}{4\pi\epsilon_0\hbar^2}\frac{v\left(r\right)}{r} - \frac{l\left(l+1\right)}{r^2}v\left(r\right) = 0 \tag{17.12}$$

Substituting Eq. 17.11 in Eq. 17.12 in seeking the Frobenius power-series solution and equating the coefficient of r^p to be zero, we obtain the recursion relation

$$\left[p\left(p+1\right) - l\left(l+1\right)\right] A_{p+1} = \left[\frac{2p\sqrt{2\mu\varepsilon}}{\hbar} - \frac{2\mu e^2}{\hbar^2 4\pi\epsilon_0}\right] A_p \ . \tag{17.13}$$

For $p = l$, $A_p = 0 \implies A_{p-1} = 0 = A_{p-2} = A_{p-3} = \cdots$. Thus, the only non-zero coefficients have $p > l$. Also, as $r \to \infty$, $u\left(r\right) \to 0$ (for a normalisable solution to the TISE). This means that the power series $v\left(r\right)$ must terminate at a finite $p(> l)$, so that the exponential part can dominate as $r \to \infty$.

Thus, for some $p = n > l$, $A_{p+1} = 0$

$$\implies \frac{2n\sqrt{2\mu\varepsilon}}{\hbar} = \frac{2\mu e^2}{\hbar^2 4\pi\epsilon_0}$$

$$\therefore E^n_{rel} = -\varepsilon = -\frac{\mu e^4}{\left(4\pi\epsilon_0\right)^2 2\hbar^2 n^2} \tag{17.14}$$

$$= -\frac{13.6}{n^2}\text{eV} \ , \quad n \geq l+1 \tag{17.15}$$

$$R\left(r\right) = \frac{v\left(r\right)}{r} e^{-\frac{\sqrt{2\mu\varepsilon}}{\hbar}r} \tag{17.16}$$

$$= R_{nl}\left(r\right)$$

Bohr

Thus, after such a long calculation, we do get back the Bohr formula (eq.(2.39))! The various $R_{nl}\left(r\right)$ can be found from getting various A_p from Eq. 17.13 , and in terms of the characteristic lengthscale for the atom (the Bohr radius a_0):

$$R_{10}\left(r\right) = \left(\frac{1}{a_0}\right)^{3/2} 2e^{-r/a_0} \ ,$$

$$R_{20}\left(r\right) = \left(\frac{1}{2a_0}\right)^{3/2} 2\left(1 - \frac{r}{2a_0}\right) e^{-r/2a_0} \ ,$$

$$R_{21}\left(r\right) = \left(\frac{1}{2a_0}\right)^{3/2} \frac{1}{\sqrt{3}}\left(\frac{r}{a_0}\right) e^{-r/2a_0} \ ,$$

and so on. Here, a_0 is the Bohr radius given by

$$a_0 = \frac{4\pi\epsilon_0\hbar^2}{\mu e^2} = 5.3\times 10^{-11}\,\mathrm{m} \quad , \quad E^n_{rel} = -\frac{\hbar^2}{2\mu n^2 a_0^2} \ ,$$

and the full wavefunction of the H-atom is given by

$$\psi_{nlm}\left(r,\theta,\phi\right) = R_{nl}\left(r\right) Y_{lm}\left(\theta,\phi\right) \tag{17.17}$$

where the good quantum numbers are given by

$$\begin{aligned} n &= 1,2,3,\cdots && \text{(Principal quantum number)} \ , \\ l &= 0,1,\cdots,n-1 && \text{(Orbital quantum number)} \ , \text{ and} \\ m &= -l,-l+1,\cdots,l-1,l && \text{(Magnetic quantum number)} \ . \end{aligned}$$

While $R_{nl}\left(r\right)$ is dependent on both n & l, the energy E^n_{rel} is dependent on n alone. There must, therefore, be denegeracies for all the levels with different values of l and m that correspond to a given value of n. This degeneracy is, however, true only for the Coulomb central potential, and reflects a case of an "accidental" degeneracy. Quantifying this degeneracy, any n has $(n-1)$ possible values for l, and each l has $2l+1$ possibles values of m. Thus, the total degeneracy for E_n will be

$$\begin{aligned} \text{Total degeneracy of } E^n_{rel} &= \sum_{l=0}^{n-1}\left(2l+1\right) \ , \\ &= 2\frac{n\left(n-1\right)}{2} + n \ , \\ &= n^2 \ . \end{aligned} \tag{17.18}$$

Taking account of the spin internal degree of freedom (with its double degeneracy), the total degeneracy for E^n_{rel} will be $2n^2$. While the energy levels thus obtained match the Bohr model exactly, our earlier understanding of quantised electron orbits with a fixed trajectory is replaced by a probability distribution $P(r)$ obtained from $\psi_{nlm}\left(r,\theta,\phi\right)$

$$P(r) = \int_0^r dr r^2 \int_0^\pi d\theta \sin\theta \int_0^{2\pi} d\phi \left|\psi_{nlm}\left(r,\theta,\phi\right)\right|^2 \quad , \quad lim_{r\to\infty}P(r) = 1 \ . \tag{17.19}$$

We can also define radial probability density $P_{nl}\left(r\right)$ such that it gives the probability of finding the electron in a small interval dr at a distance r

$$P_{nlm}\left(r\right) = \int_0^\pi \int_0^{2\pi} r^2 \sin\theta d\phi d\theta \left|\psi_{nlm}\left(r,\theta,\phi\right)\right|^2 \ , \tag{17.20}$$

$$\begin{aligned} &= r^2\left|R_{nl}\left(r\right)\right|^2 \int_0^\pi \int_0^{2\pi} \left|Y_{lm}\left(\theta,\phi\right)\right|^2 d\phi d\theta \ , \\ &= r^2\left|R_{nl}\left(r\right)\right|^2 \times I_{lm} \end{aligned} \tag{17.21}$$

where $I_{lm} = \int\limits_0^{\pi} \int\limits_0^{2\pi} |Y_{lm}(\theta, \phi)|^2 \, d\phi d\theta$.

$$\therefore \quad P_{nlm}(r) = r^2 \left| R_{nl}(r) \right|^2 \times I_{lm} \ , \tag{17.22}$$

$$\text{and} \quad P(r) = \int\limits_0^r P_{nlm}(r) \, dr \ . \tag{17.23}$$

Some plots of the $P_{nlm}(r)$ (as a function of r/a_0) are shown in Fig.17.1 below. The classical picture of orbits can (very roughly speaking) be understood as tracking the peaks of the radial probability density $P_{nlm}(r)$. Also, we display once again the solutions of the angular momentum problem in Fig.17.2 in order to visualise the "orbitals" of the Hydrogen atom.

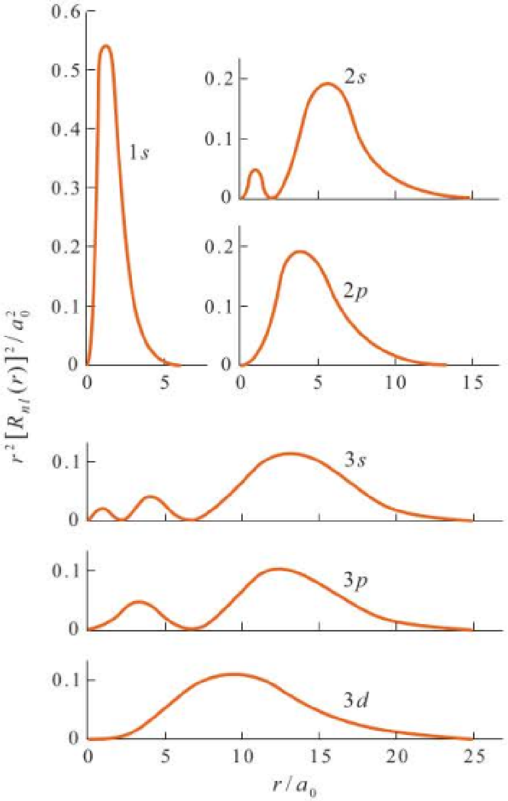


Figure 17.1: Radial Probability density function

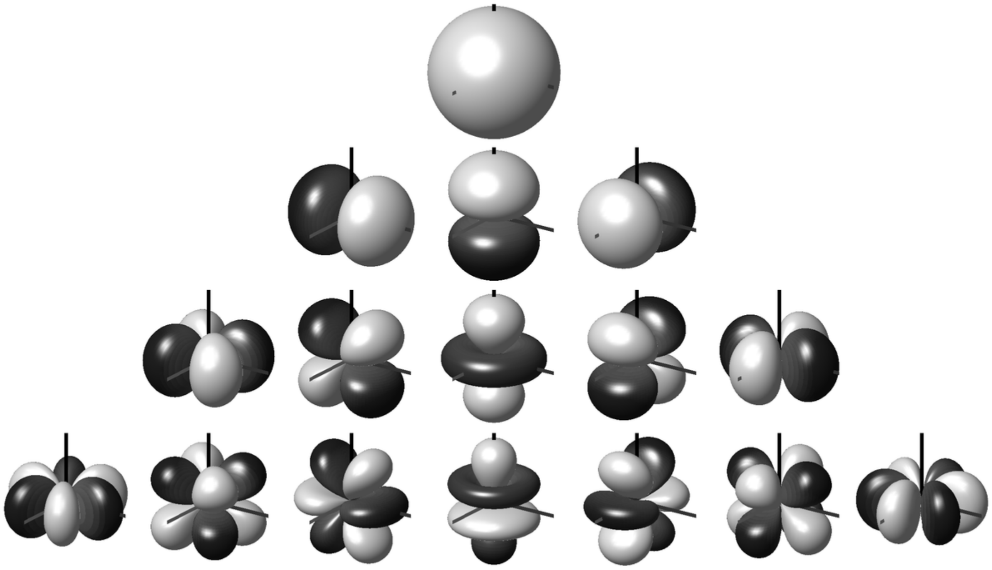


Figure 17.2: Spherical Harmonics

Chapter 18

Spin Angular Momentum

18.1 Evidence for spin

18.1.1 The Zeeman Effect

For a Hydrogen atom in an external $B-$field aligned along the z-axis:

$$H = H_0 - \frac{e}{2m}\vec{B}\cdot\vec{L}_{eff} = H_0 - \frac{e}{2m}BL^z_{eff} \tag{18.1}$$

where $\vec{B} = B\hat{z}$ is the magnetic field vector, H_0 is the Hamiltonian of the Hydrogen atom and $\vec{L}_{eff}$ is the effective angular momentum of the H-atom observed when placed in the external B-field. The Schrödinger equation can then be written as

$$H\Psi_{nlm} = \left(-\frac{13.6eV}{n^2} - \frac{e\hbar B}{2m}m_{l_{eff}}\right)\Psi_{nlm} \implies E_{nlm} = -\frac{13.6eV}{n^2} - \hbar\omega_L m_{l_{eff}}\ , \tag{18.2}$$

where $\omega_L = \frac{eB}{2m}$ is the Larmor frequency. We see that the coupling with the magnetic field removes ("lifts") the $2l+1$-fold degeneracy of the levels which had energy $E_n = -\frac{13.6}{n^2}eV$. The splitting is given by $\hbar\omega_L = \frac{e\hbar}{2m}B$, for $\Delta m_{l_{eff}} = 1$. The coupling to the external magnetic field can also be written in terms of an effective magnetic moment $\vec{\mu}_L$

$$H = \vec{\mu}_L \cdot \vec{B} \tag{18.3}$$

where $\vec{\mu}_L = -\frac{g_l\mu_B}{\hbar}\vec{L}$ and $\mu_b = \frac{e\hbar}{2m} = 5.788\times 10^{-5}eV/T$ is the Bohr magneton. For orbital angular momentum, the Lande g factor is $g_l = 1$.

However, in the H-atom, the splitting is different and an even number of levels is observed, as if $L_{\text{eff}} \in \frac{n}{2}$ and $n \in$ odd integer. Further, $n = 1$ for the H-atom with the electron in the $l = 0$ state, such that $L_{eff} = \frac{1}{2}$. Also, the splitting is different for different levels. This points towards the existence of another source of angular momentum other than the orbital angular momentum we learnt of earlier. In offering an explanation for this finding, Uhlenbeck and Goudsmit (1925) called it *spin* angular momemtum. It's origin was learnt a little later from Dirac's theory for relativistic quantum mechanics (1929).

18.1.2 The Stern-Gerlach Experiment

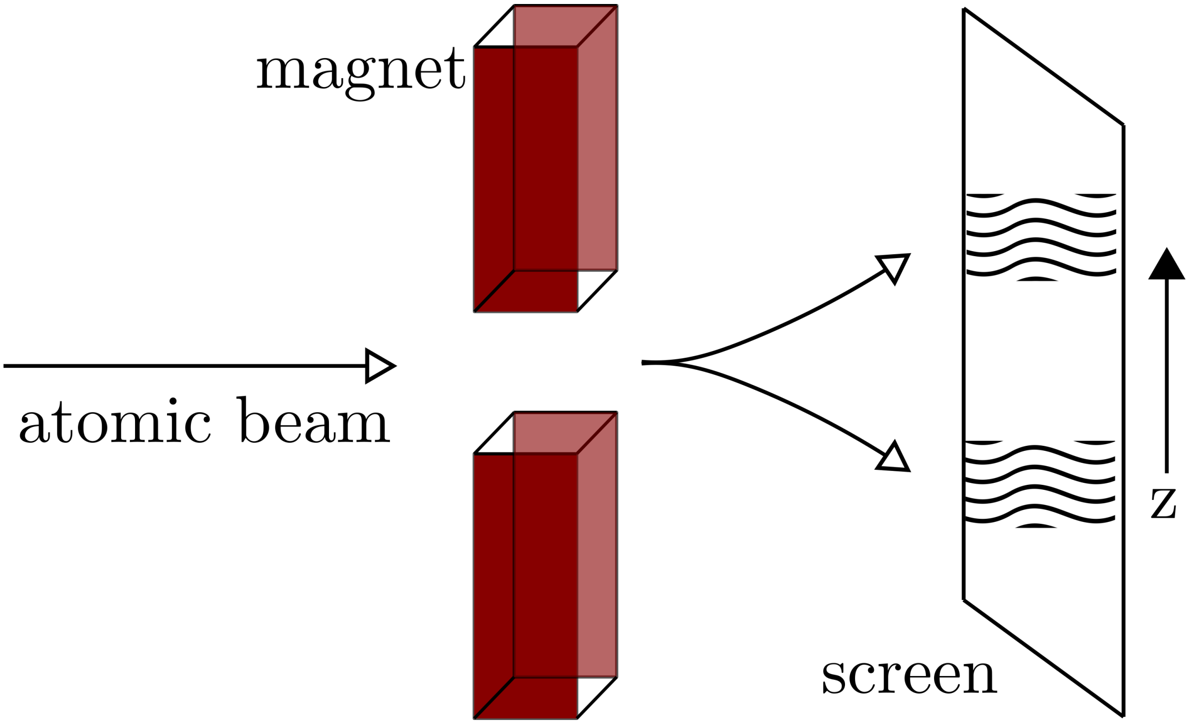


Figure 18.1: Splitting of the beam of particles into two patches when subjected to a non-uniform magnetic field

Here, an atomic beam traverses an inhomogeneous $B-$field such that there is a force on the atom (see fig. 18.1), given by

$$\vec{F} \propto \vec{\nabla}\left(\vec{\mu} \cdot \vec{B}(\vec{r})\right) \approxeq \mu_z \frac{\partial B_z(r)}{\partial z} \hat{e}_z \tag{18.4}$$

From the earlier, we would naively expect the beam to be split into an odd number $(2l+1)$ of beams. Stern and Gerlach, in that experiment, took silver atoms having a single 5s electron in the outermost shell. The only possible configuration, then, was $l = 0$. This means there should be no splitting.

The result is that the beam split into two beams! This suggests that this outermost $5s$ electron possesses an internal angular momentum, called *spin*, with the spin magnetic moment given by

$$\vec{\mu}_S = -\frac{g_S \mu_B}{\hbar} \vec{S} \tag{18.5}$$

where $\vec{S} = S_x \hat{x} + S_y \hat{y} + S_z \hat{z}$ is the spin angular momentum vector (a quantum mechanical operator!) and its z-component has eigenvalues $S_z = \pm\frac{1}{2}$. The force on the electron is thus given by

$$\vec{F} \propto \hat{z} \frac{\partial B_z}{\partial z} \left(-\frac{g_S \mu_B}{\hbar} S_z\right) \tag{18.6}$$

and the value of the corresponding Lande g-factor g_S is 2.0023193. The relativistic theory of Dirac gives $g_S = 2$, but Quantum Electrodynamics (QED) gives corrections to this.

A small historical aside. Uhlenbeck and Goudsmit were graduate students of Ehrenfest, and working for their PhDs in Leiden (Netherlands). In 1925, their advisor Ehrenfest communicated their proposal for spin angular momentum to the journal Naturwissenschaften. In the meanwhile, unknown to them, Ehrenfest had a discussion with Lorentz on the interesting (supposedly "very witty") idea of his young students. Lorentz was critical, pointing out that their idea of a "spinning electron" was incompatible with classical electrodynamics. Upon telling his students the criticism offered by Lorentz, they requested Ehrenfest not to submit their work. Enhrenfest replied that he had already done so, and that they were "both young enough to be able to afford a stupidity"! Thanks to Dirac's explanation of spin, things did not turn out quite so badly for them after all.

18.2 Eigenstates and Eigenvalues of the Spin Operators

The formalism developed earlier for orbital angular momentum $\vec{L}$ can be carried over to the case of spin angular momentum. For example, we have the following algebra:

$$[S_i, S_j] = i\hbar\epsilon_{ijk}S_k, \text{for } i, j, k \in \{x, y, z\} \tag{18.7}$$

In terms of the components, these relations look like

$$[S_x, S_y] = i\hbar S_z, \quad [S_y, S_z] = i\hbar S_x. \quad [S_z, S_x] = i\hbar S_y \tag{18.8}$$

From these commutation relations, we can write down the simultaneous eigenstates $|s, m\rangle$ of S^2 and S_z such that

$$\begin{aligned} S^2\,|s, m_s\rangle &= s(s+1)\hbar^2\,|s, m_s\rangle \\ S_z\,|s, m_s\rangle &= m\hbar\,|s, m_s\rangle \ . \end{aligned} \tag{18.9}$$

At this point, we need a brief interlude on the states $|s, m\rangle$ given above. Following Dirac, the $|.\rangle$, referred to as a **ket**, denotes a vector in the space of all states in the Hilbert space of the quantum mechanical system. It is an abstract object, and it differs from the wavefunction ψ in that the latter is a representation of an abstract ket. For example, the ket corresponding to the ground state of the simple harmonic oscillator can be written as $|n = 0\rangle$ (where n is the good quantum number labelling the eigenspectrum, and is the eigenvalue of the operator $a^\dagger a$ we studied in the chapter on the harmonic oscialltor), and its corresponding wavefunction in the position representation is

$$\psi_{n=0}(x) = \langle x|n = 0\rangle \tag{18.10}$$

where $\langle x|$ refers to the basis states comprising all functions defined in terms of the real-space coordinate x (to be discussed below), $\langle|\rangle$ refers to a projection of the state $|\rangle$ onto the state $\langle|$ called the Inner product (also discussed below), and $\psi(x)$ is the familiar real-space representation of the state we are familiar with.

Similarly, the $\langle .|$, referred to as a **bra**, is a member of the linear vector space *dual* to that of the kets. They are defined by introducing the notion of an inner product, and the bras are then linear maps that map the space of kets to that of the complex numbers via inner products. In other words, a bra is an object that when acted on the ket gives a complex number:

$$\langle a| : |b\rangle \to \mathbb{C} \tag{18.11}$$

The inner product is a generalisation of the scalar product we are familiar with from our studies of (real-valued) vectors in Cartesian space.

Further, in keeping with the example given above for the ket corresponding to the harmonic oscillator ground state and its wavefunction written in the position representation, the same can be constructed for the bra as well

$$\psi^*_{n=0}(x) = \langle n=0|x\rangle \ . \tag{18.12}$$

Returning to the spin problem, along with the spin operators along the three directions, we also define the spin creation and annihilation operators,

$$S_\pm = S_x \pm iS_y \ \ , \ \ S_+^\dagger = S_- \ , \ S_- = S_+^\dagger \ . \tag{18.13}$$

Their action on the simultaneous eigenstates can be shown to give (as shown in eq.(14.33) earlier)

$$S_\pm |s, m_s\rangle = \hbar\sqrt{s(s+1) - m_s(m_s \pm 1)}\, |s, m_s \pm 1\rangle \tag{18.14}$$

18.3 Matrix Representation of Spin

While the eigenstates of the orbital angular momentum can be written as functions of position (for eq. $Y_{lm}(\theta,\phi)$), the eigenstates of the spin angular momentum cannot. Instead, we need another way to represent these internal degrees of freedom. We first consider the simplest case of $S = \frac{1}{2}$. As observed from, e.g., the Stern-Gerlach experiment, m_s can then take just two values, $\pm\frac{1}{2}$. S_z thus has two eigenstates, which we label $|\uparrow\rangle$ and $|\downarrow\rangle$, such that

$$\begin{aligned} S_z|\uparrow\rangle &= \frac{\hbar}{2}|\uparrow\rangle\,, \quad & S^2|\uparrow\rangle = s(s+1)\hbar^2|\uparrow\rangle = \frac{3\hbar^2}{4}|\uparrow\rangle \\ S_z|\downarrow\rangle &= -\frac{\hbar}{2}|\downarrow\rangle\,, \quad & S^2|\downarrow\rangle = s(s+1)\hbar^2|\uparrow\rangle = \frac{3\hbar^2}{4}|\downarrow\rangle \end{aligned} \tag{18.15}$$

Since we know that eigenstates of a Hermitian operator are orthogonal and they span the entire Hilbert space, we can choose any two orthogonal vectors in the 2-dimensional Hilbert space of the spin-half to represent these two eigenstates. For convenience, we choose the simplest vectors (also known as *spinors*)

$$|\uparrow\rangle \equiv \left|s=\frac{1}{2}, m_s=\frac{1}{2}\right\rangle = \begin{pmatrix}1\\0\end{pmatrix}, \quad |\downarrow\rangle \equiv \left|s=\frac{1}{2}, m_s=-\frac{1}{2}\right\rangle = \begin{pmatrix}0\\1\end{pmatrix} \ , \tag{18.16}$$

$$\langle\uparrow| \equiv \left\langle s=\frac{1}{2}, m_s=\frac{1}{2}\right| = (1\ 0) \ , \quad \langle\downarrow| \equiv \left\langle s=\frac{1}{2}, m_s=-\frac{1}{2}\right| = (0\ 1) \ . \tag{18.17}$$

The two vectors are of course orthogonal, as required, and moreover, orthonormal.

$$\begin{aligned}
\langle\uparrow\mid\downarrow\rangle &= \begin{pmatrix}1 & 0\end{pmatrix}\begin{pmatrix}0\\1\end{pmatrix} = 0\\
\langle\downarrow\mid\downarrow\rangle &= \begin{pmatrix}0 & 1\end{pmatrix}\begin{pmatrix}0\\1\end{pmatrix} = 1\\
\langle\uparrow\mid\uparrow\rangle &= \begin{pmatrix}1 & 0\end{pmatrix}\begin{pmatrix}1\\0\end{pmatrix} = 1
\end{aligned} \tag{18.18}$$

While we will focus on $s=\frac{1}{2}$, this representation can be easily extended to $s=1$:

$$|s=1,m_s=1\rangle = \begin{pmatrix}1\\0\\0\end{pmatrix}, \quad |s=1,m_s=0\rangle = \begin{pmatrix}0\\1\\0\end{pmatrix}, \quad |s=1,m_s=-1\rangle = \begin{pmatrix}0\\0\\1\end{pmatrix} \tag{18.19}$$

Combing back to $s=\frac{1}{2}$, since the operators S_i act on these two-component vectors, they themselves must be 2×2 matrices. We can construct them by simply calculating the 4 matrix elements.

$$\begin{aligned}
\langle\uparrow|S_z|\uparrow\rangle &= \frac{\hbar}{2}\\
\langle\uparrow|S_z|\downarrow\rangle &= 0 = \langle\downarrow|S_z|\uparrow\rangle\\
\langle\downarrow|S_z|\downarrow\rangle &= -\frac{\hbar}{2}
\end{aligned} \tag{18.20}$$

We can thus write

$$S_z = \frac{\hbar}{2}\begin{pmatrix}\langle\uparrow|S_z|\uparrow\rangle & \langle\uparrow|S_z|\downarrow\rangle\\ \langle\downarrow|S_z|\uparrow\rangle & \langle\downarrow|S_z|\downarrow\rangle\end{pmatrix} = \frac{\hbar}{2}\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix} \tag{18.21}$$

By definition, the states $|\uparrow\rangle$ and $|\downarrow\rangle$ are eigenstates of S_z with $m_s=\frac{1}{2}$ and $m_s=-\frac{1}{2}$ respectively.

We will next find the matrices $S_\pm$. From their operation, we know that $S_\pm$ increases/decreases the eigenvalue m_s. From the theory of angular momentum, we know that m_s can extend from $-s$ to s. For spin-half, the highest value is $\frac{1}{2}$, which is the state $|\uparrow\rangle$. This implies that there is no state with higher m_s, and S_+ should give 0 when acting on the state $|\uparrow\rangle$. Similarly, since $|\downarrow\rangle$ has the lowest value of m_s for a spin-half, S_- should give 0 when acting on it.

$$S_+|\uparrow\rangle = 0 = S_-|\downarrow\rangle \tag{18.22}$$

To find the other actions, we recall eq. 18.14.

$$\begin{aligned}
S_+|\downarrow\rangle &= \hbar|\uparrow\rangle\\
S_-|\uparrow\rangle &= \hbar|\downarrow\rangle
\end{aligned} \tag{18.23}$$

These four equations are sufficient to write down all the matrix elements.

$$\begin{aligned}
\langle\uparrow|S_+|\uparrow\rangle = \langle\downarrow|S_+|\downarrow\rangle = \langle\uparrow|S_-|\uparrow\rangle = \langle\downarrow|S_-|\downarrow\rangle = 0\\
\langle\downarrow|S_+|\uparrow\rangle = \langle\uparrow|S_-|\downarrow\rangle = 0\\
\langle\uparrow|S_+|\downarrow\rangle = \langle\downarrow|S_-|\uparrow\rangle = \hbar
\end{aligned} \tag{18.24}$$

We can now construct the matrices.

$$\begin{aligned} S_+ &= \frac{\hbar}{2}\begin{pmatrix} \langle\uparrow|S_+|\uparrow\rangle & \langle\uparrow|S_+|\downarrow\rangle \\ \langle\downarrow|S_+|\uparrow\rangle & \langle\downarrow|S_+|\downarrow\rangle \end{pmatrix} = \hbar\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \ , \\ S_- &= \frac{\hbar}{2}\begin{pmatrix} \langle\uparrow|S_-|\uparrow\rangle & \langle\uparrow|S_-|\downarrow\rangle \\ \langle\downarrow|S_-|\uparrow\rangle & \langle\downarrow|S_-|\downarrow\rangle \end{pmatrix} = \hbar\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \ . \end{aligned} \tag{18.25}$$

Using the relations

$$\begin{aligned} S_x &= \frac{1}{2}\left(S_+ + S_-\right) \ , \\ S_y &= \frac{1}{2i}\left(S_+ - S_-\right) \ , \end{aligned} \tag{18.26}$$

we can construct the S_x and S_y matrices.

$$\begin{aligned} S_x &= \frac{\hbar}{2}\begin{pmatrix} \langle\uparrow|S_x|\uparrow\rangle & \langle\uparrow|S_x|\downarrow\rangle \\ \langle\downarrow|S_x|\uparrow\rangle & \langle\downarrow|S_x|\downarrow\rangle \end{pmatrix} = \frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \ , \\ S_y &= \frac{\hbar}{2}\begin{pmatrix} \langle\uparrow|S_y|\uparrow\rangle & \langle\uparrow|S_y|\downarrow\rangle \\ \langle\downarrow|S_y|\uparrow\rangle & \langle\downarrow|S_y|\downarrow\rangle \end{pmatrix} = \frac{\hbar}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \ . \end{aligned} \tag{18.27}$$

The three spin operators can be written as $S_i = \frac{\hbar}{2}\sigma_i$, where σ_i are the Pauli spin matrices.

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{18.28}$$

18.4 Properties of Pauli matrices

The Pauli matrices satisfy the following identities.

- $\sigma_x^2 = \mathbb{I} = \sigma_y^2 = \sigma_z^2 \ \ , \ \ \mathbb{I} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$
- $[\sigma_x, \sigma_y] = 2i\sigma_z$, and its cyclic permutations
- $\{\sigma_x, \sigma_y\} = \sigma_x\sigma_y + \sigma_y\sigma_x = 0$, and its cyclic permutations
- the previous two points imply $\sigma_x\sigma_y = -\sigma_y\sigma_x = i\sigma_z$
- $\sigma_x\sigma_y\sigma_z = i\mathbb{I}$
- $\text{Trace}\left(\sigma_x\right) = \text{Trace}\left(\sigma_y\right) = \text{Trace}\left(\sigma_z\right) = 0 \longrightarrow$ traceless matrices
- $\text{Det}\left(\sigma_x\right) = \text{Det}\left(\sigma_y\right) = \text{Det}\left(\sigma_z\right) = -1$

The first three properties can be combined into

$$\sigma_i\sigma_j = \delta_{ij} + i\epsilon_{ijk}\sigma_k \tag{18.29}$$

Often we will talk about a vector of Pauli matrices. That is defined as

$$\vec{\sigma} = \sigma_x\hat{x} + \sigma_y\hat{y} + \sigma_z\hat{z} \tag{18.30}$$

For arbitrary vector operators $\vec{a} = \sum_i a_i \hat{i}$ and $\vec{b} = \sum_i b_i \hat{i}$ ($i \in \{x, y, z\}$) that commute with the Pauli matrices ($[a_i, \sigma_j] = 0 = [b_i, \sigma_j]$), we have the following identity:

$$(\vec{\sigma} \cdot \vec{a}) \left(\vec{\sigma} \cdot \vec{b}\right) = \mathbb{I}\vec{a} \cdot \vec{b} + i\vec{\sigma} \cdot \left(\vec{a} \times \vec{b}\right) \tag{18.31}$$

Finally, it is worth noting that any arbitrary 2 × 2 matrix (say A) with complex-valued matrix elements can be written as a linear combination the three Pauli matrices and the 2 × 2 identity matrix

$$A = a\sigma_x + b\sigma_y + c\sigma_z + d\mathbb{I} \ , \tag{18.32}$$

where $a, b, c, d \in \mathcal{C}$ in general (and $\in \mathcal{R}$ in particular).

18.5 Eigenstates of S_x

We have already seen the eigenstates and eigenvalues of S_z:

$$|{\uparrow}\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, |{\downarrow}\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$
$$S_z |{\uparrow}\rangle = \frac{\hbar}{2} |{\uparrow}\rangle , S_z |{\downarrow}\rangle = -\frac{\hbar}{2} |{\downarrow}\rangle$$

Now consider S_x and assume that it has an eigenvector

$$\begin{pmatrix} \psi_1 & \psi_2 \end{pmatrix} \tag{18.33}$$

with eigenvalue c:

$$\underbrace{\frac{\hbar}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}}_{S_x} \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix} = c \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix} \tag{18.34}$$

Solving the eigenvalue problem leads to the determinant equation:

$$\begin{vmatrix} -c & \frac{\hbar}{2} \\ \frac{\hbar}{2} & c \end{vmatrix} = 0 \implies c^2 - \left(\frac{\hbar}{2}\right)^2 = 0 \implies c = \pm\frac{\hbar}{2} \tag{18.35}$$

So, the eigenvalues of S_x (and S_y, if you check them out!) are identical to that of S_z! But there is no surprise in this really: there is no preferred spin direction and we have the freedom of rotating the x-axis into the z-direction. So, why would any measurable quantity look different?

Solving for the eigenvectors of S_x corresponding to the eigenvalues $\pm\frac{\hbar}{2}$ gives

$$\begin{aligned} \text{eigenvalue } \frac{\hbar}{2} &\leftrightarrow |{\rightarrow}\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \frac{1}{\sqrt{2}} \left(|{\uparrow}\rangle + |{\downarrow}\rangle\right) \\ \text{eigenvalue } -\frac{\hbar}{2} &\leftrightarrow |{\leftarrow}\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \frac{1}{\sqrt{2}} \left(|{\uparrow}\rangle - |{\downarrow}\rangle\right) \end{aligned} \tag{18.36}$$

while for S_y, we get

$$\begin{aligned} \text{eigenvalue } \frac{\hbar}{2} &\leftrightarrow |\nearrow\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\ i\end{pmatrix} = \frac{1}{\sqrt{2}}\left(|\uparrow\rangle + i\,|\downarrow\rangle\right) \\ \text{eigenvalue } -\frac{\hbar}{2} &\leftrightarrow |\swarrow\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\ -i\end{pmatrix} = \frac{1}{\sqrt{2}}\left(|\uparrow\rangle - i\,|\downarrow\rangle\right) \end{aligned} \tag{18.37}$$

Thus, we see that a spin in the $x(-x)$ or $y(-y)$ directions are actually linear combinations of states in which the spin is aligned along the z and $-z$ directions! This counter-intuitive finding is purely an outcome of quantum mechanics and has no classical analogue.

In fact, just to test these findings, we can run an initially unpolarised beam of H-atoms through a succession of three Stern-Gerlach experiments with inhomogeneous B-fields along z-direction for the first, the x-direction for the second, and the z-direction for the third, and in the following sequence

- B-field along z-direction : the initially unpolarised beam first splits into $|\uparrow\rangle$ and $|\downarrow\rangle$ states,
- B-field along x-direction next : putting (say) $|\downarrow\rangle$ through second S-G gives $|\leftarrow\rangle$ and $|\rightarrow\rangle$ state beams, and
- B-field along z-direction again : putting (say) $|\rightarrow\rangle$ through third S-G produces both $|\uparrow\rangle$ and $|\downarrow\rangle$ beams again!

Interesting, isn't it?

18.6 Spin Precession

Classically, a magnetic dipole $\vec{\mu}$ (initially at rest) when placed in an external magnetic field $\vec{B}$ feels a torque $\vec{\mu} \times \vec{B}$, causing it to align with the field. If, in addition, it has an angular momentum, the applied torque will cause the rotating body to precess about a direction given by $\vec{\mu} \times \vec{B}$.

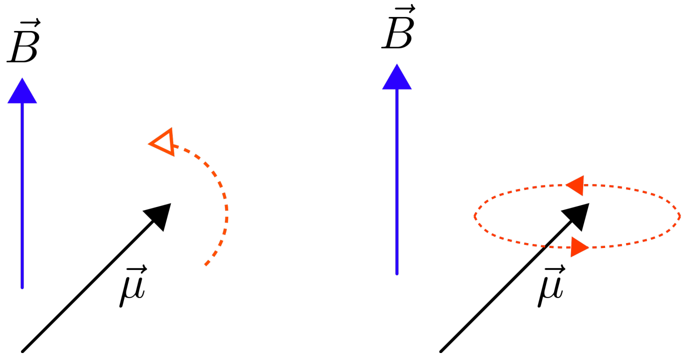


Figure 18.2: Left: Torque that aligns the spin along $\vec{B}$. Right: Precession of the spin about $\vec{B}$

What happens to a quantum mechanical spin? Consider an electron with magnetic moment $\vec{\mu}$ at rest in a magnetic field $\vec{B}$. We need to study the time evolution of the spin, using the TDSE

$$H\left|\Psi\right\rangle = i\hbar\frac{\partial}{\partial t}\left|\Psi\right\rangle \tag{18.38}$$

What is the Hamiltonian H for this problem? If we assume the particle is at rest, we get

$$H = -\vec{\mu}\cdot\vec{B} \tag{18.39}$$

The magnetic moment $\vec{\mu}$ can be related to the spin angular momentum using $\vec{\mu} = -\frac{2\mu_B}{\hbar}\vec{S}$. Assuming the applied magnetic field is in the direction of $\hat{z}$, we get

$$H = \frac{2\mu_B}{\hbar}B\hat{z}\cdot\vec{S} = \frac{2\mu_B}{\hbar}BS_z = \mu_B B\sigma_z \tag{18.40}$$

where $\sigma_z = \frac{2}{\hbar}S_z$ is the Pauli matrix along z direction. Since the spin angular momentum operators belong to a two-dimensional Hilbert space, the state vector $\left|\Psi\right\rangle$ will also have two components

$$\left|\Psi\right\rangle = \begin{pmatrix}\Psi_+\\ \Psi_-\end{pmatrix}\ . \tag{18.41}$$

The TDSE becomes

$$\mu_B B\sigma_z\left|\Psi\right\rangle = i\hbar\frac{\partial}{\partial t}\left|\Psi\right\rangle \implies \mu_B B\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\begin{pmatrix}\Psi_+\\ \Psi_-\end{pmatrix} = i\hbar\frac{\partial}{\partial t}\begin{pmatrix}\Psi_+\\ \Psi_-\end{pmatrix} \tag{18.42}$$

Comparing the matrix elements gives

$$\pm\mu_B B\Psi_\pm = i\hbar\frac{\partial}{\partial t}\Psi_\pm \tag{18.43}$$

The solution of these equations are

$$\Psi_\pm = A_\pm\exp\left\{\mp i\left(\mu_B B\frac{t}{\hbar}\right)\right\} \tag{18.44}$$

where the coefficients $A_\pm$ specify the initial condition: $\left|\Psi(t=0)\right\rangle = \begin{pmatrix}A_+\\ A_-\end{pmatrix}$. We now consider two specific initial conditions.

18.6.1 Spin is along +z at t=0

The initial state will be

$$\left|\Psi(t=0)\right\rangle = \begin{pmatrix}1\\ 0\end{pmatrix}\ , \tag{18.45}$$

hence $A_+ = 1, A_- = 0$. Plugging these into the solutions gives $\Psi_+(t) = \exp\left\{-i\left(\mu_B B\frac{t}{\hbar}\right)\right\}$ and $\Psi_-(t) = 0$.

$$\left|\Psi(t)\right\rangle = \begin{pmatrix}\exp\left\{-i\left(\mu_B B\frac{t}{\hbar}\right)\right\}\\ 0\end{pmatrix} \tag{18.46}$$

The probability of finding the spin along $+z$, as a function of time, is

$$P(t) = |\langle\uparrow|\Psi(t)\rangle|^2 = \left|\begin{pmatrix}1 & 0\end{pmatrix}\begin{pmatrix}\exp\left\{-i\left(\mu_B B\frac{t}{\hbar}\right)\right\}\\ 0\end{pmatrix}\right|^2 = \left|\exp\left\{-i\left(\mu_B B\frac{t}{\hbar}\right)\right\}\right|^2 = 1 \quad (18.47)$$

The probability is always equal to unity, so the time evolution for this initial condition is trivial (i.e., there is no precession of the spin).

18.6.2 Spin is along +x at t=0

$$|\Psi(t=0)\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix} \implies A_\pm = \frac{1}{\sqrt{2}} \quad (18.48)$$

$$|\Psi(t)\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}\exp\left\{-i\left(\mu_B B\frac{t}{\hbar}\right)\right\}\\ \exp\left\{i\left(\mu_B B\frac{t}{\hbar}\right)\right\}\end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix}e^{-i\omega\frac{t}{2}}\\ e^{i\omega\frac{t}{2}}\end{pmatrix} \quad (18.49)$$

where $\omega \equiv \frac{2\mu_B B}{\hbar}$ is the Larmor frequency. We will now look at the state $|\Psi(t)\rangle$ for various values of $t > 0$.

At a later time $t = \frac{2\pi}{\omega}$ (i.e., a full time period),

$$|\Psi\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}e^{-i\pi}\\ e^{i\pi}\end{pmatrix} = -\frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix} \quad (18.50)$$

This is equal to $\frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix}$ up to a phase factor ($e^{i\pi} = -1$), so at $t = \frac{2\pi}{\omega}$ (i.e., a full time period), the spin is pointing towards $+x$ once again.

At precisely half this time period, i.e., $t = \frac{\pi}{\omega}$,

$$|\Psi\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}e^{-i\frac{\pi}{2}}\\ e^{i\frac{\pi}{2}}\end{pmatrix} = -\frac{i}{\sqrt{2}}\begin{pmatrix}1\\-1\end{pmatrix} \quad (18.51)$$

This is a spin pointing along $-x$ (upto the phase $e^{i\pi/2} = -i$).

At a quarter of this time period, i.e., $t = \frac{\pi}{2\omega}$,

$$|\Psi\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}e^{-i\frac{\pi}{4}}\\ e^{i\frac{\pi}{4}}\end{pmatrix} = \frac{1-i}{\sqrt{2}}\begin{pmatrix}1\\i\end{pmatrix} \quad (18.52)$$

This is a spin pointing along $+y$ (upto a phase $e^{i\pi/4}$).

At three quarters of this time period, i.e., $t = \frac{3\pi}{2\omega}$,

$$|\Psi\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}e^{-i\frac{3\pi}{4}}\\ e^{i\frac{3\pi}{4}}\end{pmatrix} = \frac{1+i}{\sqrt{2}}\begin{pmatrix}-1\\i\end{pmatrix} \quad (18.53)$$

This is a spin pointing along $-y$ (upto a phase $e^{i3\pi/4}$).

Thus we can see that in keeping with our classical idea of the electron precessing in the $x-y$ plane, the direction of the spin vector is also rotating in the counter-clockwise direction in the $x-y$ plane (the "precession"). The angular frequency of the precession is $\omega = \frac{2\mu_B B}{\hbar}$, and the time period is $\tau = \frac{2\pi}{\omega} = \frac{\pi\hbar}{\mu_B B}$. This phenomenon is the basis of MRI: the Hydrogen atoms (protons actually) precess when in the presence of external magnetic fields, and emit radiation in the mega-hertz range upon relaxing upon being excited. The detection of this radiation can be used to map the emitting protons. All of this falls within the field of nuclear magnetic resonance (NMR), which happens to be one of the most spectacular discoveries of the 20th century simply in terms of its wide usage in the natural sciences! Naturally, the discovery was awarded the Nobel prize.

Note that you can always carry out the same calculation with the state initially being in some general direction; the maths will be a little more tedious, but you will still see the precession.

18.7 Spin and spatial degrees of freedom

As mentioned in an earlier chapter, when there exists no physics that can couple the orbital ($\vec{L}$) and spin ($\vec{S}$) angular momenta

$$\left[\vec{S},\vec{r}\right] = 0, \quad \left[\vec{S},\vec{p}\right] = 0, \quad \left[\vec{S},\vec{L}\right] = 0 \tag{18.54}$$

we can define eigenstates simultaneously for $\vec{S}$ and any one of the other three. For example, we can construct a basis from $|\vec{r}\rangle\,|\uparrow\rangle$ and $|\vec{r}\rangle\,|\downarrow\rangle$ (outer or direct product of vector spaces), such that

$$|\Psi\rangle = \int d^3\vec{r}\,[\psi_+(\vec{r})\,|\vec{r}\rangle\,|\uparrow\rangle + \psi_-(\vec{r})\,|\vec{r}\rangle\,|\downarrow\rangle] \tag{18.55}$$

where $\psi_\pm(\vec{r})$ are the amplitudes associated with finding a particle at position $\vec{r}$ and spin $\uparrow$ and $\downarrow$ respectively, with the time-dependent Schrödinger equation

$$i\hbar\frac{\partial}{\partial t}\begin{pmatrix}\psi_+\\ \psi_-\end{pmatrix} = \left[-\frac{\hbar^2}{2m}\vec{\nabla}^2 + V(\vec{r}) + \mu_B\left(\vec{L}+\vec{S}\right)\cdot\vec{B}\right]\begin{pmatrix}\psi_+(\vec{r},t)\\ \psi_-(\vec{r},t)\end{pmatrix}\ , \tag{18.56}$$

where $\vec{L}+\vec{S}=\vec{J}$ gives the total angular momentum.

18.8 Addition of two angular momenta

For two angular momentum operators $\vec{J_1}$ and $\vec{J_2}$ that do not interact with each other in any way (for example, two electrons which have no spin-dependent coupling among themselves), we can write $\left[\vec{J_1},\vec{J_2}\right] = 0$.

For the total angular momentum operator $\vec{J} = \vec{J_1}+\vec{J_2}$, we can see from the above, and using

$$[J_{\nu,i}, J_{\mu,j}] = i\hbar\delta_{\nu\mu}\epsilon_{ijk}J_{\mu k}, \quad \nu,\mu\in\{1,2\} \tag{18.57}$$

such that

$$[J_i, J_j] = i\hbar\epsilon_{ijk}J_k \ . \tag{18.58}$$

Thus, all properties of angular momentum operators and their eigenstates hold for $\vec{J}$. What do these eigenstates look like?

Clearly, $|j_1, m_1; j_2, m_2\rangle = |j_1, m_1\rangle\, |j_2, m_2\rangle$, which are eigenstates of J_1^2, J_{1z}, J_2^2 and J_{2z}, will not work. This is for the following reason. While

$$J_z\, |j_1, m_1; j_2, m_2\rangle = (J_{1z} + J_{2z})\, |j_1, m_1; j_2, m_2\rangle = \hbar\,(m_1 + m_2)\, |j_1, m_1; j_2, m_2\rangle \ , \tag{18.59}$$

these eigenstates are not eigenstates of J^2 as

$$\left[J^2, J_{iz}\right] \neq 0 \quad , \quad i = 1, 2 \ . \tag{18.60}$$

since

$$\left[(J_{1x} + J_{2x})^2, J_{\alpha z}\right] \neq 0 \ , \ \ \left[(J_{1y} + J_{2y})^2, J_{\alpha z}\right] \neq 0 \ \text{ for } \alpha = 1, 2 \ . \tag{18.61}$$

But since

$$\left[J^2, J_\alpha\right] = 0 \ \ (\alpha = x, y, z) \quad , \quad \left[J^2, J_i^2\right] = 0 \quad , \quad i = 1, 2 \ , \tag{18.62}$$

we need eigenstates of J^2, J_z, J_1^2, andJ_2^2. That is, we need $|j, m_j, j_1, j_2\rangle$. We next investigate the addition of two spin-half angular momenta operators.

18.8.1 Addition of two spin-half operators

Given two spin-half operators S_1 and S_2, we can create a total spin operator

$$\vec{S} = \vec{S}_1 + \vec{S}_2 \tag{18.63}$$

such that we have four states $|\uparrow, \uparrow\rangle, |\uparrow, \downarrow\rangle, |\downarrow, \uparrow\rangle, |\downarrow, \downarrow\rangle$ (where we are using the notation that the first member denotes the z component of the first spin operator, and the second component the same for the second spin operator).

$$S_z\, |\uparrow\uparrow\rangle = \hbar\, |\uparrow\uparrow\rangle \, , S_z\, |\downarrow\downarrow\rangle = -\hbar\, |\downarrow\downarrow\rangle \, , S_z\, |\uparrow\downarrow\rangle = 0 = S_z\, |\downarrow\uparrow\rangle \tag{18.64}$$

where $S_z = S_{1z} + S_{2z}$. Thus, given that $S^2|S, S_z\rangle = S(S+1)\hbar^2|S, S_z\rangle$, it is reasonable to suppose that $S = 0, 1$ here.

Note that

$$\begin{aligned} S^2 &= S_1^2 + S_2^2 + \vec{S_1} \cdot \vec{S_2} + \vec{S_2} \cdot \vec{S_1} \\ &= S_1^2 + S_2^2 + 2\vec{S_1} \cdot \vec{S_2} \qquad\qquad \left[\because \left[\vec{S_1}, \vec{S_2}\right] = 0\right] \\ &= \frac{3}{2}\hbar^2 + 2S_{1z}S_{2z} + S_{1+}S_{2-} + S_{1-}S_{2+} \ , \end{aligned} \tag{18.65}$$

as

$$S_i^2 = \frac{3\hbar^2}{4} \ , \tag{18.66}$$

$$\vec{S_1} \cdot \vec{S_2} = S_{1z}S_{2z} + S_{1x}S_{2x} + S_{1y}S_{2y} \ , \tag{18.67}$$

$$= S_{1z}S_{2z} + \frac{1}{2}(S_{1+}S_{2-} + S_{1-}S_{2+}) \ , \tag{18.68}$$

and we have used $S_{\nu\pm} = S_{\nu x} \pm iS_{\nu y}$ ($\nu \in 1, 2$). Then, we can write

$$\begin{aligned} &S^2 \left|\uparrow\uparrow\right\rangle = \left[\frac{3}{2}\hbar^2 + 2\left(\frac{\hbar}{2}\right)^2\right] \left|\uparrow\uparrow\right\rangle = 2\hbar^2 \left|\uparrow\uparrow\right\rangle \quad , \quad S_z \left|\uparrow\uparrow\right\rangle = \hbar \left|\uparrow\uparrow\right\rangle \ , \\ &S^2 \left|\downarrow\downarrow\right\rangle = 2\hbar^2 \left|\downarrow\downarrow\right\rangle \quad , \quad S_z \left|\downarrow\downarrow\right\rangle = -\hbar \left|\downarrow\downarrow\right\rangle \ . \end{aligned} \tag{18.69}$$

As $S^2 \left|\Psi\right\rangle = S(S+1)\hbar^2 \left|\Psi\right\rangle$, we get from the above relations $S = 1, S_z = \pm 1$ for $\Psi = \left|\uparrow\uparrow\right\rangle , \left|\downarrow\downarrow\right\rangle$. The third state is obtained by operating $S_- = S_{1-} + S_{2-}$.

$$\begin{aligned} \frac{S_-}{\sqrt{2}\hbar} \left|\uparrow\uparrow\right\rangle &= \frac{S_{1-} + S_-}{\sqrt{2}\hbar} \left|\uparrow\uparrow\right\rangle \\ &= \frac{1}{\sqrt{2}} \left(\left|\uparrow\downarrow\right\rangle + \left|\downarrow\uparrow\right\rangle\right) \quad [\because S_- \left|\uparrow\right\rangle = \hbar \left|\downarrow\right\rangle , S_+ \left|\downarrow\right\rangle = \hbar \left|\uparrow\right\rangle] \ , \end{aligned} \tag{18.70}$$

and the $1/\sqrt{2}$ factor is a normalisation factor. Further, we can easily see that

$$S^2 \frac{1}{\sqrt{2}} \left(\left|\uparrow\downarrow\right\rangle + \left|\downarrow\uparrow\right\rangle\right) = \hbar^2 (\frac{3}{2} - \frac{2}{4} + 1) \frac{1}{\sqrt{2}} \left(\left|\uparrow\downarrow\right\rangle + \left|\downarrow\uparrow\right\rangle\right) = 2\hbar^2 \frac{1}{\sqrt{2}} \left(\left|\uparrow\downarrow\right\rangle + \left|\downarrow\uparrow\right\rangle\right) \tag{18.71}$$

$$S_z \frac{1}{\sqrt{2}} \left(\left|\uparrow\downarrow\right\rangle + \left|\downarrow\uparrow\right\rangle\right) = 0 \ , \tag{18.72}$$

i.e., $S = 1 \ , \ S_z = 0$ for the third state given by $\frac{1}{\sqrt{2}} \left(\left|\uparrow\downarrow\right\rangle + \left|\downarrow\uparrow\right\rangle\right)$.

There is another normalised state, $\frac{1}{\sqrt{2}} \left(\left|\uparrow\downarrow\right\rangle - \left|\downarrow\uparrow\right\rangle\right)$, which is orthogonal to all these three eigenstates (Check this!), such that

$$S_z \frac{1}{\sqrt{2}} \left(\left|\uparrow\downarrow\right\rangle - \left|\downarrow\uparrow\right\rangle\right) = 0 \tag{18.73}$$

and

$$S^2 \frac{1}{\sqrt{2}} \left(\left|\uparrow\downarrow\right\rangle - \left|\downarrow\uparrow\right\rangle\right) = \left(\frac{3}{2}\hbar^2 - 2\frac{\hbar^2}{4} - \hbar^2\right) \frac{1}{\sqrt{2}} \left(\left|\uparrow\downarrow\right\rangle - \left|\downarrow\uparrow\right\rangle\right) = 0 \tag{18.74}$$

Therefore, this fourth eigenstate is characterized by $S = 0$ and $S_z = 0$.

Together, we have the following eigenstates:

$$\text{triplet(3) states} \quad : \quad \begin{cases} \left|S = 1, S_z = 1\right\rangle & \rightarrow \quad \left|\uparrow\uparrow\right\rangle \\ \left|S = 1, S_z = 0\right\rangle & \rightarrow \quad \frac{1}{\sqrt{2}} \left(\left|\uparrow\downarrow\right\rangle + \left|\downarrow\uparrow\right\rangle\right) \\ \left|S = 1, S_z = -1\right\rangle & \rightarrow \quad \left|\downarrow\downarrow\right\rangle \end{cases} \tag{18.75}$$

$$\text{singlet(1) state} \quad : \quad \left\{ \left|S = -0, S_z = 0\right\rangle \quad \rightarrow \quad \frac{1}{\sqrt{2}} \left(\left|\uparrow\downarrow\right\rangle - \left|\downarrow\uparrow\right\rangle\right) \right.$$

Note the projection operators

$$\begin{aligned} P_1 &= \frac{3}{4} + \frac{1}{\hbar^2} \vec{S_1} \cdot \vec{S_2} = \frac{3}{4} + \frac{1}{2\hbar^2} \left(S^2 - S_1^2 - S_2^2\right) \\ P_2 &= \frac{1}{4} - \frac{1}{\hbar^2} \vec{S_1} \cdot \vec{S_2} = \frac{1}{4} - \frac{1}{2\hbar^2} \left(S^2 - S_1^2 - S_2^2\right) \end{aligned} \tag{18.76}$$

which project onto the triplet and singlet spaces respectively:

$$\begin{aligned} P_1 \left|1, m\right\rangle &= \left(\frac{3}{4} + \frac{1}{\hbar^2}\vec{S_1}\cdot\vec{S_2}\right)\left|1, m\right\rangle \\ &= \left[\frac{3}{4} + \frac{1}{2\hbar^2}\left(S^2 - S_1^2 - S_2^2\right)\right]\left|1, m\right\rangle \\ &= \left[\frac{3}{4} + \frac{1}{2\hbar^2}\left(1\times 2 - \frac{3}{4} - \frac{3}{4}\right)\right]\left|1, m\right\rangle \\ &= \left|1, m\right\rangle \end{aligned} \tag{18.77}$$

$$P_1 \left|0, 0\right\rangle = \left[\frac{3}{4} + \frac{1}{2}\left(0 - \frac{3}{4} - \frac{3}{4}\right)\right]\left|0, 0\right\rangle = 0 \tag{18.78}$$

$$P_0 \left|1, m\right\rangle = (1 - P_1)\left|1, m\right\rangle = 0 \tag{18.79}$$

$$P_0 \left|0, m\right\rangle = (1 - P_1)\left|0, 0\right\rangle = \left|0, 0\right\rangle \tag{18.80}$$

We will see in the next chapter the elementary problem for which these are the eigenstates.

Chapter 19

A system of Interacting Spins

To see a World in a Grain of Sand
And a Heaven in a Wild Flower
Hold Infinity in the palm of your hand
And Eternity in an hour
William Blake, “Auguries of Innocence”.

We are now going to embark on an exciting journey. Using the Heisenberg spin exchange problem of two spin-1/2 degrees of freedom as a template, we will unveil a wide variety of important concepts and ideas in quantum mechanics. Buckle up!

19.1 Eigenspectrum of the Heisenberg spin exchange problem

How do we enumerate the eigenstates of a system of coupled, i.e., interacting, spins? Consider the simplest case - the Heisenberg spin exchange Hamiltonian:

$$H = \lambda \vec{S}_1 \cdot \vec{S}_2 \quad , \quad \lambda > 0 \ , \tag{19.1}$$

where λ is the spin exchange coupling. Consider also that the particles carrying the spin are localised in space and $V(r) = 0$. Now, note that since

$$H \propto \vec{S}_i \quad , \quad [H, S_{iz}] \neq 0 \quad , \quad (i = 1, 2) \ . \tag{19.2}$$

Hence, states of definite S_{iz} cannot be eigenstates of the above Hamiltonian H. However,

$$\left[H, S^2\right] = 0, \quad \text{where } S^2 = \left(\vec{S}_1 + \vec{S}_2\right)^2 \tag{19.3}$$

This is because

$$S^2 = S_1^2 + S_2^2 + 2\vec{S}_1 \cdot \vec{S}_2 \implies \vec{S}_1 \cdot \vec{S}_2 = \frac{1}{2}\left(S^2 - S_1^2 - S_2^2\right) \tag{19.4}$$

Clearly, $[H, S^2] = [H, S_1^2] = [H, S_2^2] = 0$. This means that the eigenstates can be labelled by S, m_S, S_1 and S_2 - states like $|S, m_S; S_1, S_2\rangle$. The eigenvalues are

$$\begin{aligned} H\,|S, m_S; S_1, S_2\rangle &= \frac{\lambda}{2}\hbar^2\left[S(S+1) - S_1(S_1+1) - S_2(S_2+1)\right]|S, m_S; S_1, S_2\rangle \\ &= \frac{\lambda}{2}\hbar^2\left[S(S+1) - \frac{3}{2}\right]|s, m_S; S_1, S_2\rangle \quad , \end{aligned} \tag{19.5}$$

as $S_i(S_i+1)\,|S, m_S; S_1, S_2\rangle = \frac{3\hbar^2}{4}\,|S, m_S; S_1, S_2\rangle$. The energy eigenvalue is clearly independent of m_S, and the contributions of the $S_i(S_i+1)\;\;(i = 1, 2)$ terms has become a constant factor $3\hbar^2/4$. All states with the same S but different m_S are thus degenerate:

$$\begin{aligned} \text{unique ground state: } E_{S=0} &= -\frac{3}{4}\lambda\hbar^2 \qquad [\text{singlet} \rightarrow S = 0, m_S = 0] \\ \text{triply degenerate excited states: } E_{S=1} &= \frac{1}{4}\lambda\hbar^2 \qquad [\text{triplets} \rightarrow S + 1, m_S = -1, 0, 1] \end{aligned} \tag{19.6}$$

The gap between the ground state and the group of excited states is

$$\Delta E = \lambda\hbar^2 \tag{19.7}$$

The form of the eigenstates have already been mentioned in eq. 18.75. We thus get the following spectrum for the problem:

E	$\|\Psi\rangle$
$-\frac{3}{4}\lambda\hbar^2$	$\|S=0, S_z=0\rangle = \frac{1}{\sqrt{2}}(\|\uparrow,\downarrow\rangle - \|\downarrow,\uparrow\rangle)$
$\frac{1}{4}\lambda\hbar^2$	$\|S=1, S_z=0\rangle = \frac{1}{\sqrt{2}}(\|\uparrow,\downarrow\rangle + \|\downarrow,\uparrow\rangle)$ $\|S=1, S_z=1\rangle = \|\uparrow,\uparrow\rangle$ $\|S=1, S_z=-1\rangle = \|\downarrow,\downarrow\rangle$

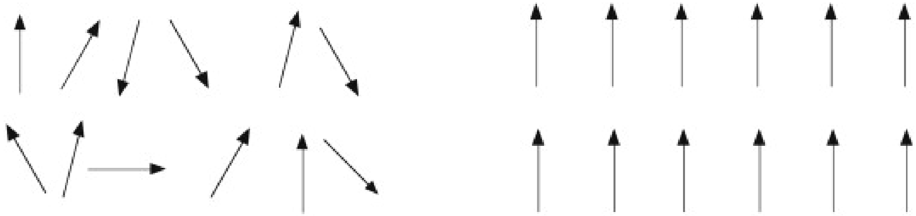
Table 19.1: Spectrum of Heisenberg spin-exchange Hamiltonian

19.2 Symmetries, symmetry breaking and quantum fluctuations

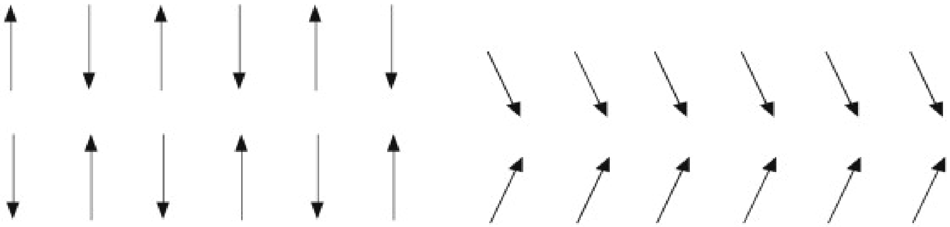
At this point, we can have a discussion on the importance of symmetries for a quantum mechanical system and the consequences that arise from their breaking. We will, of course, use the Heisenberg spin exchange problem of two interacting spin-1/2 degrees of freedom discussed just above as an illustrative example.

Why are symmetries so important in Quantum Mechanics?
Recall that we have often encountered the fact that when a Hamiltonian possesses certain symmetries, this reflected in the commutation of the Hamiltonian with certain operators

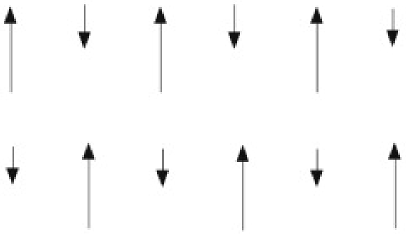
that carried out the symmetry related transformations. The commutation means that simultaneous eigenstates of the Hamiltonian and these operators, and reflected an invariance of members of the Hilbert space, especially the ground state, under those transformations. For instance, remember that in the case of the particle in a box and the simple harmonic oscillator problems, we found the potentials (and hence the entire Hamiltonian) to enjoy parity symmetry in the one dimensional space on which these problems are defined. As a consequence, the Hamiltonian commuted with the parity operator and the ground state wavefunction was parity symmetric (i.e., had even parity under the parity transformation, of eigenvalue $\hat{P} = +1$). Further, the rest of the eigenstates alternated in the parity eigenvalue (between odd and even parity). Similarly, the free particle problem had continuous translation symmetry, such that the Hamiltonian commuted with the linear momentum operator and all the eigenstates $\psi \propto e^{ikx}$ possess continuous translation invariance on the open interval $-\infty \leq x \leq \infty$.



(a) Paramagnet with random spins **(b)** Ferromagnetic ordering of spins



(c) Antiferromagnetic ordering **(d)** Weak ferromagnet (or Canted antiferromagnet)



(e) Ferrimagnetic ordering

Figure 19.1: Various kinds of magnetism observed in nature. Source: The internet.

Now, let us return to the problem of two interacting spin-$1/2$s we studied above. First, note that for the spin exchange strength (or "coupling") $\lambda > 0$, we have an *antiferromagnetic* exchange interaction between the two spins. By this, we mean that such an interaction will naturally favour the anti-alignment of the two spins within the ground state configuration (such the the ground state energy is negative). Without having worked out the eigenstates and eigenvalues to the Hamiltonian, you could (as a first guess) have chosen the state

$|\uparrow_1\downarrow_2\rangle$ as the ground state, based on the reasoning that $\lambda S_{1z}S_{2z}|\uparrow_1\downarrow_2\rangle = -\frac{\lambda\hbar^2}{4}|\uparrow_1\downarrow_2\rangle$ (i.e., a lowering of the energy beyond 0 has thus been achieved). However, this reasoning is too naive: we are after all working with a Hamiltonian $\lambda\vec{S}_1\cdot\vec{S}_2$, i.e., other than the $S_{1z}S_{2z}$ term, it also contains the $S_{1x}S_{2x}+S_{1y}S_{2y}=\frac{1}{2}(S_{1+}S_{2-}+S_{1-}S_{2+})$ terms. Unless there is a miracle, these terms will also have a role to play in determining the nature of the ground state as well as its eigenvalue. Now, if you look back at the solution, you can easily see that this is indeed the case! The true ground state is neither $|\uparrow_1\downarrow_2\rangle$ nor $|\downarrow_1\uparrow_2\rangle$. Instead, it is the singlet state $\frac{1}{\sqrt{2}}(|\uparrow_1,\downarrow_2\rangle - |\downarrow_1,\uparrow_2\rangle)$, a linear superposition of $|\uparrow_1\downarrow_2\rangle$ and $|\downarrow_1\uparrow_2\rangle$. We will see soon below that this singlet state possesses a property called *entanglement* that is quite special to quantum mechanics.

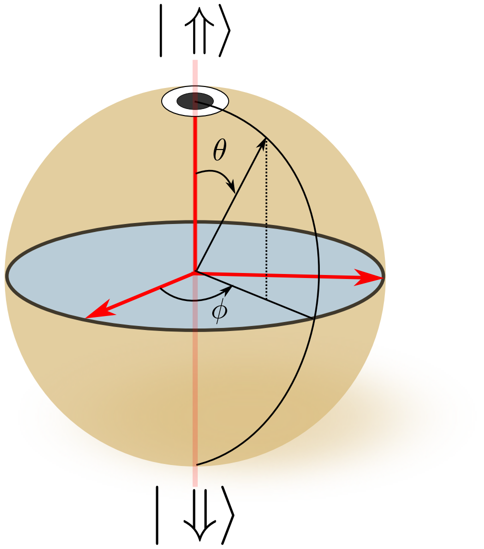


Figure 19.2: The spin-S degree of freedom represented on the Bloch sphere. The north and south poles represent the $|1\rangle \equiv |\uparrow\rangle$ and $|0\rangle \equiv |\downarrow\rangle$ "classical" states. The vector $|\psi\rangle$ has length $\sqrt{S(S+1)}\hbar$, and lies on the surface of a sphere, and defined by the polar angles (θ,ϕ).

But why is the singlet state ($\frac{1}{\sqrt{2}}(|\uparrow_1,\downarrow_2\rangle - |\downarrow_1,\uparrow_2\rangle)$) chosen instead of its components ($|\uparrow_1\downarrow_2\rangle$ and $|\downarrow_1\uparrow_2\rangle$)?

Well, an answer lies in looking at the symmetries of the Heisenberg spin exchange Hamiltonian

$$H = \frac{\lambda}{2}\hbar^2\left[S\,(S+1) - \frac{3}{2}\right]\ . \tag{19.8}$$

An inspection of the Hamiltonian reveals that it possesses symmetry under a common rotation of both spins $\vec{S}_1$ and $\vec{S}_2$, amounting to the fact that the total spin vector $\vec{S}=\vec{S}_1+\vec{S}_2$ can point in *any* direction in the (x,y,z) Cartesian coordinate system on the surface of a sphere of radius $\sqrt{S(S+1)}\hbar$. But when we make the specific choice of $|\uparrow_1\downarrow_2\rangle$ and $|\downarrow_1\uparrow_2\rangle$, we are explicitly "breaking" (or lowering/violating) the spherical (i.e., rotational) symmetry of $\vec{S}$. Further, $|\uparrow_1\downarrow_2\rangle$ and $|\downarrow_1\uparrow_2\rangle$ are not even the eigenstates of the Hamiltonian H given above, as can be seen from the fact that even though

$$S_z(= S_{1z} + S_{2z})|\uparrow_1\downarrow_2\rangle = 0 = S_z|\downarrow_1\uparrow_2\rangle\ , \tag{19.9}$$

i.e., it appears as if $|\uparrow_1\downarrow_2\rangle$ and $|\downarrow_1\uparrow_2\rangle$ are eigenstates of S_z, we find that

$$S_{1-}S_{2+}|\uparrow_1\downarrow_2\rangle = |\downarrow_1\uparrow_2\rangle\ \ ,\ \ S_{1+}S_{2-}|\downarrow_1\uparrow_2\rangle = |\uparrow_1\downarrow_2\rangle\ . \tag{19.10}$$

This means that $|\uparrow_1\downarrow_2\rangle$ and $|\downarrow_1\uparrow_2\rangle$ are not eigenstates of S^2. In fact, we can see that the two linear superposition states $\frac{1}{\sqrt{2}}(|\uparrow_1\downarrow_2\rangle \pm |\downarrow_1\uparrow_2\rangle)$ arise from the diagonalisation of the following simple 2×2 matrix H_{split}, obtained by taking the matrix elements of the Hamiltonian $H = \lambda(S_{1z}S_{2z} + \frac{1}{2}(S_{1+}S_{2-} + S_{1-}S_{2+}))$ written in the basis of the $|\uparrow_1\downarrow_2\rangle$ and $|\downarrow_1\uparrow_2\rangle$ states

$$H_{split} = \begin{pmatrix} \langle\uparrow_1\downarrow_2|H|\uparrow_1\downarrow_2\rangle & \langle\uparrow_1\downarrow_2|H|\downarrow_1\uparrow_2\rangle \\ \langle\downarrow_1\uparrow_2|H|\uparrow_1\downarrow_2\rangle & \langle\downarrow_1\uparrow_2|H|\downarrow_1\uparrow_2\rangle \end{pmatrix} = \frac{\lambda\hbar^2}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} - \frac{\lambda\hbar^2}{4}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \frac{\lambda\hbar^2}{2}\sigma_x - \frac{\lambda\hbar^2}{4}\mathcal{I}\,. \tag{19.11}$$

The diagonalisation of H_{split} is quite straighforward: we can easily see that the spin single and triplet zero states ($\frac{1}{\sqrt{2}}(|\uparrow_1\downarrow_2\rangle \pm |\downarrow_1\uparrow_2\rangle)$) are the true quantum eigenstates formed from the "classical" symmetry broken choices of $|\uparrow_1\downarrow_2\rangle$ and $|\downarrow_1\uparrow_2\rangle$. Further, this is clearly the outcome of the action of the $S_{1x}S_{2x} + S_{1y}S_{2y} = \frac{1}{2}(S_{1+}S_{2-} + S_{1-}S_{2+})$ terms, and leads to a "lifting" of the degeneracy (in the energy contribution from only the $S_{1z}S_{2z}$ term of the Hamiltonian, as seen above) of the $|\uparrow_1\downarrow_2\rangle$ and $|\downarrow_1\uparrow_2\rangle$ states. From the viewpoint of the "classical" (or non-entangled) states $|\uparrow_1\downarrow_2\rangle$ and $|\downarrow_1\uparrow_2\rangle$, these two terms can be thought of as the *quantum fluctuation* content in the Heisenberg spin exchange Hamiltonian, while the $S_{1z}S_{2z}$ term can be thought of as a *classical potential energy cost*. In this way, we can see that the singlet ground state $\frac{1}{\sqrt{2}}(|\uparrow_1\downarrow_2\rangle - |\downarrow_1\uparrow_2\rangle)$ and the triplet zero excited state $\frac{1}{\sqrt{2}}(|\uparrow_1\downarrow_2\rangle + |\downarrow_1\uparrow_2\rangle)$ possess the full rotational symmetry of the Heisenberg Hamiltonian (reflected in $S_z = 0$ for these two states), and this is achieved through the action of the quantum fluctuation terms.

A closer look at quantum fluctuations.
For another way in which to see where the terminology "quantum fluctuations" arises from, note that applying the Schrödinger equation to the two "classical" states $|\psi_{Cl}\rangle = (|\uparrow_1\downarrow_2\rangle\ \ |\downarrow_1\uparrow_2\rangle)^\dagger$, we obtain

$$\begin{aligned} i\hbar\frac{\partial|\psi_{Cl}\rangle}{\partial t} &= H|\psi_{Cl}\rangle \\ &= \lambda(S_{1z}S_{2z} + \frac{1}{2}(S_{1+}S_{2-} + S_{1-}S_{2+}))|\psi_{Cl} \\ i\hbar\frac{\partial}{\partial t}(|\uparrow_1\downarrow_2\rangle\ \ |\downarrow_1\uparrow_2\rangle)^\dagger &= \lambda(S_{1z}S_{2z} + \frac{1}{2}(S_{1+}S_{2-} + S_{1-}S_{2+}))(|\uparrow_1\downarrow_2\rangle\ \ |\downarrow_1\uparrow_2\rangle)^\dagger \\ &= (\frac{\lambda\hbar^2}{2}\sigma_x - \frac{\lambda\hbar^2}{4}\mathcal{I})(|\uparrow_1\downarrow_2\rangle\ \ |\downarrow_1\uparrow_2\rangle)^\dagger\ . \end{aligned} \tag{19.12}$$

The presence of the off-diagonal pieces within the σ_x part of the Hamiltonian reflects the fact that the classical states $|\uparrow_1\downarrow_2\rangle$ and $|\downarrow_1\uparrow_2\rangle$ are not eigenstates of H. Indeed, this equation of motion shows that if we start with a state $|\uparrow_1\downarrow_2\rangle$, then there exists a "transition amplitude" $\langle\downarrow_1\uparrow_2|H|\uparrow_1\downarrow_2\rangle = \lambda\hbar^2/2$, i.e., the Hamiltonian causes a transition from the state $|\uparrow_1\downarrow_2\rangle$ to the state $|\downarrow_1\uparrow_2\rangle$. A similar transition amplitude is found if we start instead with the state $|\downarrow_1\uparrow_2\rangle$, with the system transition to the state $|\uparrow_1\downarrow_2\rangle$. These non-zero transition amplitudes are the quantum fluctuations between the two classical states. Instead, if we were to start with the spin singlet and spin-triplet-zero eigenstates of H, then obviously the solution to the time-dependent Schrödinger equation will yield only a trivial phase of $e^{iEt/\hbar}$ (where E corresponds to the energy eigenvalues $E_s = -3J/4$ and $E_{t0} = J/4$ of the singlet and triplet-zero eigenstates respectively). Thus, we can write the time evolution of the state $|\uparrow_1\downarrow_2\rangle$ as

$$|\uparrow_1\downarrow_2\rangle(t) = \frac{e^{iE_st/\hbar}}{\sqrt{2}} \times \frac{1}{\sqrt{2}}(|\uparrow_1\downarrow_2\rangle - |\downarrow_1\uparrow_2\rangle) + \frac{e^{iE_{t0}t/\hbar}}{\sqrt{2}} \times \frac{1}{\sqrt{2}}(|\uparrow_1\downarrow_2\rangle + |\downarrow_1\uparrow_2\rangle) \ . \tag{19.13}$$

This indicates that the state $|\uparrow_1\downarrow_2\rangle$ evolves in time t in the form of an equal amplitude (i.e., with equal amplitudes $1/\sqrt{2}$ for the two pieces) superposition of the two eigenstates $|\psi_s\rangle = \frac{1}{\sqrt{2}}(|\uparrow_1\downarrow_2\rangle - |\downarrow_1\uparrow_2\rangle)$ and $|\psi_{t0}\rangle = \frac{1}{\sqrt{2}}(|\uparrow_1\downarrow_2\rangle + |\downarrow_1\uparrow_2\rangle)$. Then, at any given time point t, the probability of the state $|\uparrow_1\downarrow_2\rangle(t)$ being in either of the two eigenstates is equal (and independent of time), as can be seen from

$$\begin{aligned} P_{\langle\psi_s|\uparrow_1\downarrow_2\rangle} &= |\langle\psi_s|\uparrow_1\downarrow_2\rangle(t)|^2 = |\frac{e^{iE_st/\hbar}}{\sqrt{2}}|^2 = \frac{1}{2} \ , \\ P_{\langle\psi_{t0}|\uparrow_1\downarrow_2\rangle} &= |\langle\psi_{t0}|\uparrow_1\downarrow_2\rangle(t)|^2 = |\frac{e^{iE_{t0}t/\hbar}}{\sqrt{2}}|^2 = \frac{1}{2} \ . \end{aligned} \tag{19.14}$$

Further, if we ask for the probability for the state $|\uparrow_1\downarrow_2\rangle$ to remain in the same state at time t, we find

$$\begin{aligned} P_{\langle\uparrow_1\downarrow_2(0)|\uparrow_1\downarrow_2\rangle}(t) &= |\frac{1}{\sqrt{2}}(\frac{1}{\sqrt{2}}(\langle\uparrow_1\downarrow_2| - \langle\downarrow_1\uparrow_2|) + \frac{1}{\sqrt{2}}(\langle\uparrow_1\downarrow_2| - \langle\downarrow_1\uparrow_2|)) \times \\ &\quad \frac{1}{\sqrt{2}}(\frac{e^{iE_st/\hbar}}{\sqrt{2}}(|\uparrow_1\downarrow_2\rangle - |\downarrow_1\uparrow_2\rangle) + \frac{e^{iE_{t0}t/\hbar}}{\sqrt{2}}(|\uparrow_1\downarrow_2\rangle - |\downarrow_1\uparrow_2\rangle))|^2 \\ &= |\frac{e^{iE_st/\hbar}}{2} + \frac{e^{iE_{t0}t/\hbar}}{2}|^2 \\ &= \frac{1}{2}(1 + \cos\left(\frac{(E_{t0} - E_s)t}{\hbar}\right)) \ . \end{aligned} \tag{19.15}$$

This reveals the fact that the probability $P_{\langle\psi_s|\uparrow_1\downarrow_2\rangle}(t)$ varies periodically in time about the mean value of $1/2$, and ranges between $0 \leq P_{\langle\psi_s|\uparrow_1\downarrow_2\rangle}(t) \leq 1$. Similarly, we find that the probability for the state $|\uparrow_1\downarrow_2\rangle$ to transition to the same state $|\downarrow_1\uparrow_2\rangle$ time t is given by

$$\begin{aligned}
P_{\langle\uparrow_1\downarrow_2(0)|\downarrow_1\uparrow_2\rangle}(t) &= |\frac{1}{\sqrt{2}}(\frac{1}{\sqrt{2}}(\langle\uparrow_1\downarrow_2| - \langle\downarrow_1\uparrow_2|) + \frac{1}{\sqrt{2}}(\langle\uparrow_1\downarrow_2| - \langle\downarrow_1\uparrow_2|)) \times \\
&\quad \frac{1}{\sqrt{2}}(\frac{e^{iE_st/\hbar}}{\sqrt{2}}(|\uparrow_1\downarrow_2\rangle - |\downarrow_1\uparrow_2\rangle) - \frac{e^{iE_{t0}t/\hbar}}{\sqrt{2}}(|\uparrow_1\downarrow_2\rangle - |\downarrow_1\uparrow_2\rangle))|^2 \\
&= |\frac{e^{iE_st/\hbar}}{2} - \frac{e^{iE_{t0}t/\hbar}}{2}|^2 \\
&= \frac{1}{2}(1 - \cos\left(\frac{(E_{t0} - E_s)t}{\hbar}\right)) \ . \qquad (19.16)
\end{aligned}$$

Thus, the transition probability $P_{\langle\uparrow_1\downarrow_2(0)|\downarrow_1\uparrow_2\rangle}(t)$ also has a similar periodic variation in time about the mean value of 1/2, and ranges between $0 \leq P_{\langle\psi_s|\uparrow_1\downarrow_2\rangle}(t) \leq 1$. However, the two probabilities $P_{\langle\uparrow_1\downarrow_2(0)|\uparrow_1\downarrow_2\rangle}(t)$ and $P_{\langle\uparrow_1\downarrow_2(0)|\downarrow_1\uparrow_2\rangle}(t)$ are constrained by the requirements of unitarity

$$P_{\langle\uparrow_1\downarrow_2(0)|\uparrow_1\downarrow_2\rangle}(t) + P_{\langle\uparrow_1\downarrow_2(0)|\downarrow_1\uparrow_2\rangle}(t) = 1 \ . \qquad (19.17)$$

In this way, we see quantum fluctuations reflect the fact that the time evolution of classical states involve the simultaneous time evolution of multiple quantum eigenstates, as dictated by quantum mechanics. Further, quantum fluctuations clearly show up in the fact that the probability to stay in the same classical state, as well as the probability to transition into another classical state, are both time dependent functions that are overall constrained by the requirement by unitarity. Clearly, there can be no quantum fluctuations if we are working with the true quantum eigenstates of the Hamiltonian. As we will see below, while the classical states do not possess the property known as entanglement, quantum fluctuations lead to a finite entanglement in the quantum eigenstates (and which can be quantified).

Ferromagnetism and Spontaneous Symmetry Breaking.
However, the $|\uparrow_1\uparrow_2\rangle$ and $|\downarrow_1\downarrow_2\rangle$ eigenstates do not possess the full rotational symmetry on the sphere, as can be seen from the fact that (i) we have defined the z-axis from among an infinite number of possible choices on the full sphere, and (ii) $S_z = \pm 1$ for these two eigenstates. However, they are excited eigenstates of the full Hamiltonian ($H \propto S(S+1)$) for $\lambda > 0$.

However, what would happen if they (or even one of them) would become the ground state(s)? This can actually happen for the case of a *ferromagnetic* spin exchange coupling $\lambda < 0$; in this case, it is easy to see that the energy is lowered by choosing the spins to be aligned. Indeed, the eigenspectrum for the problem with $\lambda < 0$ is simply obtained from that for $\lambda > 0$ by reversing the sign of λ in the expressions shown above. Thus, we have a three-fold degenerate ground state manifold in the problem now: $|\uparrow_1\uparrow_2\rangle$, $|\downarrow_1\downarrow_2\rangle$ and $\frac{1}{\sqrt{2}}(|\uparrow_1\downarrow_2\rangle + |\downarrow_1\uparrow_2\rangle)$. The S_z values for these three states is $+1, -1$ and 0 respectively. Furthermore, only the last of these three possesses the rotational symmetry of the Hamiltonian. It turns out, though, that this degenerate manifold of ground states is highly susceptible towards the lifting of the degeneracy.

Consider adding a Zeeman field-spin coupling term $B(S_{1z}+S_{2z})$ to the ferromagnetic Heisenberg spin exchange Hamiltonian, and with $\lambda < 0$. Then, it is easily seen that $B \neq 0$ immediately lifts the triple degeneracy: such a coupling to an external field immediately affects the energy of the $|\uparrow_1\uparrow_2\rangle$ and $|\downarrow_1\downarrow_2\rangle$ in *opposite* ways, while it does not affect the triplet zero state $\frac{1}{\sqrt{2}}|\uparrow_1\downarrow_2 + \downarrow_1\uparrow_2\rangle$ state at all!

Thus, $|\uparrow_1\uparrow_2\rangle$ is the unique ground state for $B \to 0-$, while $|\downarrow_1\downarrow_2\rangle$ is the unique ground state for $B \to 0+$. These are the two ferromagnetic (or spin aligned) ground states. Thus, it would appear that taking the limit of $B = 0$ is singular, as the nature of the ground state reverses completely from $|\uparrow_1\uparrow_2\rangle$ to $|\downarrow_1\downarrow_2\rangle$ across it. In this sense, we can say that even an infinitesmally small B field picks out a unique ferromagnetic (spin aligned) ground state for the case of $\lambda < 0$. This strange occurrence is termed as *spontaneous symmetry breaking* for a thermodynamically large system of spin-1/2 degrees of freedom (i.e., with 10^{23} no. of constituents) that are interacting with the same Heisenberg spin exchange Hamiltonian. Indeed, this mechanism is understood to be the origin of ferromagnetism in materials such as Iron. However, as will become clear in an upcoming lecture, there is absolutely no entanglement content in such symmetry broken (or "classical") ground states.

The antiferromagnetic problem can be different.
We will see in a later section of this chapter that such a breaking of the rotational symmetry of the antiferromagnetic Heisenberg Hamiltonian's singlet ground state ($\frac{1}{\sqrt{2}}(|\uparrow_1\downarrow_2\rangle - |\downarrow_1\uparrow_2\rangle)$, for $\lambda > 0$) is not achieved for $B \to 0\pm$. Instead, one needs a critical value of the external $|B| > 2\lambda$ field to replace the singlet state with either the $|\uparrow_1\uparrow_2\rangle$ or $|\downarrow_1\downarrow_2\rangle$ states. As we will discuss there, this is a classic signature of a level-crossing phenomenon that is at the heart of phase transitions (i.e., changes of state of matter) at zero temperature driven purely by quantum fluctuations.

The overarching lesson to take from here.
Quantum fluctuations prefer eigenstates (and especially ground states) that preserve all symmetries of the Hamiltonian; such eigenstates typically possess a rich entanglement content. Thus, they naturally tend to compete with the tendency for a system to undergo spontaneous symmetry breaking and end up in classical ground states that possess very little (if any) entanglement.

The Mermin-Wagner-Hohenberg theorem states that quantum fluctuations always win over symmetry breaking tendencies in systems that exist in one spatial dimension (or less), there is adequate evidence for the fact that three dimensional quantum systems undergo symmetry breaking quite easily. This makes two dimensional quantum systems the ideal playground for the competition between quantum fluctuations and symmetry breaking, and physicists have been devoting considerable attention to them for the past four decades (including the Nobel Prize being awarded to Thouless, Kosterlitz and Haldane in 2016).

19.3 Entanglement of the Ground State: preliminaries

We will now see that the singlet ground state ($\frac{1}{\sqrt{2}}(|\uparrow_1\downarrow_2\rangle - |\downarrow_1\uparrow_2\rangle)$) of the antiferromagnetic Heisenberg spin exchange problem for two interacting spin-1/2s is a so-called *maximally entangled* state of the two spins. Entanglement here refers to the phenomenon that we cannot describe the state of any one of the spins without specifying that of the other. In other words, measuring the spin angular momentum of either spin would also determine the other spin - the two results will be *anti-correlated.* If S_{1z} is measured to be up (upon acting with S_{1z} on the singlet state), we know that S_{2z} is down, and vice versa. This property persists irrespective of spatial distance between the two spins. If we produce a pair of spins in such a singlet state, and distribute them over a large distance without affecting their spin angular momenta, they will remain locked in this state. If the first measurement on S_{1z} is done on the moon and the result is up (i.e., $\langle S_{1z}\rangle = \hbar/2$), a second measurement done infinitesimally later on the earth on S_{2z} will give down (i.e., $\langle S_{2z}\rangle = -\hbar/2$). We will return to this discussion slightly later, in the form of the Einstein-Podolsky-Rosen (EPR) gedanken.

This is in contrast to a *separable* state, where measuring one subsystem (say, spin 1) does not affect the other (say, spin 2). This is an example of a separable state (though un-normalized) :

$$|\uparrow\rangle_1 \otimes (|\uparrow\rangle + |\downarrow\rangle)_2 \ , \tag{19.18}$$

where it is clearly seen that the state of spin 1 is in a Direct Product ($\otimes$) with that of spin 2. Such a separation is clearly not possible for the singlet state $\frac{1}{\sqrt{2}}(|\uparrow_1\downarrow_2\rangle - |\downarrow_1\uparrow_2\rangle)$.

One way of quantifying quantum entanglement is by calculating the von Neumann entropy (VNE) of the reduced density matrix of the state (referred to more popularly as the Entanglement entropy (EE)). First we will define what a density matrix is. Density matrices are an alternate formulation of quantum mechanics, where we describe the system by an operator ρ instead of the wavefunction ψ. This is more general compared to the wavefunction approach, because there are systems which cannot be represented by a wavefunction. Consider, for example, a system of particles (with a spectrum $\{E_1, ..., E_n\}$) coupled with a thermal bath at a temperature T. The system could be in any of its eigenstates E_i, with a probability $\exp\{-\beta E_i\}$ (where $\beta = 1/k_B T$, and k_B is the Boltzmann constant). This cannot be represented by a wavefunction. (Note that the wavefunction $|\Psi\rangle = \sum_i \exp\{-\beta E_i\} |\Psi_i\rangle$ is not a correct choice, because this is still a single state, whereas we know from statistical mechanics that the system has to be in a probabilistic mixture of various eigenstates.) Such a probabilistic system can instead be represented using a density matrix.

19.4 A quick introduction to Density Matrices

A density matrix ρ is mathematically defined as a positive semi-definite Hermitian operator in the Hilbert space of the complete system (i.e., its eigenvalues as non-negative), such that the expectation value of any operator $\hat{O}$ is given by $\left\langle \hat{O} \right\rangle = \text{Tr}\left(\rho\, \hat{O}\right)$. Being Hermitian, the density matrix can be expanded in a complete orthonormal basis:

$$\rho = \sum_i C_i \,|i\rangle \langle i| \tag{19.19}$$

Note that ρ is an operator, as $|i\rangle\langle i|$ corresponds to a matrix that acts on the state vectors $|i\rangle$and $\langle i|$. To see this in a simple example, we can use $\langle i|$ as the state row vectors $(1\ 0), (0\ 1)$

of a 2-level system and $|i\rangle$ as the corresponding column vectors. Then

$$\begin{aligned}
\rho &= \sum_i C_i \; |i\rangle \langle i| = C_1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} \; (1 \; 0) + C_2 \begin{pmatrix} 0 \\ 1 \end{pmatrix} \; (0 \; 1) \; . \\
&= C_1 \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + C_2 \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \\
&= \begin{pmatrix} C_1 & 0 \\ 0 & C_2 \end{pmatrix} \; . \qquad (19.20)
\end{aligned}$$

Products such as $|i\rangle \langle i|$ are called *Outer Products*, and correspond to operators. On the other hand, terms like $\langle i|i\rangle$ is called the *Inner Products* and give rise to scalar quantities (i.e., number, either real or complex).

Since ρ is positive semi-definite (i.e., its eigenvalues as non-negative), we will have $C_i \geq 0$. If we take the special case of $\hat{O} = \mathbb{I}$, we can write

$$\begin{aligned}
1 &= \langle \mathbb{I} \rangle = \mathrm{Tr}\,(\rho \mathbb{I}) = \sum_j \langle j| \, \rho \, |j\rangle \\
&= \sum_{i,j} \; C_i \; \langle j|i\rangle \; \langle i|j\rangle \quad (\text{as } \rho = \sum_i C_i |i\rangle\langle i|) \\
&= \sum_j C_j \quad [\because \langle j|i\rangle = \delta_{ij}] \; , \qquad (19.21)
\end{aligned}$$

which gives

$$\mathrm{Tr}(\rho) = \sum_j C_j = 1 \; . \qquad (19.22)$$

This indicates that the C_is can be thought of as the quantum mechanical probabilities associated with the basis of states $\{|i\rangle\}$ indexed by i.

All states that can be written as a density matrix can be classified into two classes: *pure* states and *mixed* states. If the ρ is defined by $C_i = \delta_{ij}$ such that $\rho = |j\rangle \langle j|$, we call this a pure state. Such a state is exactly equivalent to saying that the system can be described by a single ket $|j\rangle$, and the density matrix has only one diagonal element (say C_i) of value 1 and all other elements (diagonal as well as off-diagonal) being 0. Working with wavefunctions means we are restricting our systems to just pure states. Mixed states, on the other hand, refer to any system which is not in a pure state. Such a state cannot be represented by a wavefunction, and we must resort to working with density matrices. In such density matrices, more than one diagonal element is non-zero; additionally, there may be off-diagonal elements that are non-zero as well. We will now see that the presence of several non-zero diagonal (and maybe also off-diagonal) elements encodes a special property called entanglement. It is also important to note that a pure state can also be entangled, and studied via the notion of a so-called reduced density matrix (to be introduced below).

The motivation for defining such an object can be noted in the following manner. Recall that if we have an isolated quantum system (which is what we have been discussing for most of these lectures!), we can always write a general superposition state of the system as $|\Psi\rangle = \sum_i C_i |\psi_i\rangle$ (where the C_i are the superposition coefficients and the $\{|\psi_i\rangle\}$ is the eigenbasis), then the expectation value of an operator $\hat{O}$ is simply obtained as

$$\langle\hat{O}\rangle = \langle\Psi|\,\hat{O}\,|\Psi\rangle = \sum_{ij} C_i C_j^* \,\langle\psi_j|\,\hat{O}\,|\psi_i\rangle \ . \tag{19.23}$$

In the above, the product $C_i C_j^*$ defines the probability distribution for the states within the state $|\Psi\rangle$.

On the other hand, let us consider an open quantum system, i.e., a quantum system that is coupled to an environment. In such a two-component system, a general state of the complete system + environment would be $|\Psi\rangle = \sum_{ij} K_{ij} |\psi_i\rangle |\phi_j\rangle$, where ψ represents the state of one component (say, the quantum system) and ϕ that of the other (say, the environment to which the system is coupled). An operator $\hat{O}$ that acts only on component ψ (i.e., the quantum system) would have an expectation value

$$\begin{aligned}\left\langle\hat{O}\right\rangle &= \langle\Psi|\,\hat{O}\,|\Psi\rangle = \sum_{ijmn} K_{ij}K_{mn}\langle\phi_n|\phi_j\rangle\,\langle\psi_m|\,\hat{O}\,|\psi_i\rangle \\ &= \sum_{ijm} K_{ij}K_{mj}\hat{O}_{mi} \quad (\text{as } \langle\phi_j|\phi_i\rangle = \delta_{ij}) \\ &= \sum_{im} \rho_{mi}\hat{O}_{mi} \quad (\text{where } \rho_{mi} = \sum_j K_{ij}K_{mj} \ \ , \ \ O_{mi} = \langle\psi_m|\,\hat{O}\,|\psi_i\rangle) \\ &= \mathrm{Tr}(\rho\hat{O}) \ , \end{aligned} \tag{19.24}$$

where the matrix elements $\rho_{mi} = \sum_j K_{ij}K_{mj}$ constitute the density matrix ρ. The goal here was to be able to write down the expectation value of a $\psi-$ (or system-)only operator purely in terms of $\psi-$basis states, such that the details of the subsystem ϕ (i.e., the environment) have been put into the density matrix elements ρ_{mi}.

In the natural (i.e., diagonal) basis of ρ, the expectation value can be written as

$$\left\langle\hat{O}\right\rangle = \sum_i \rho_i \hat{O}_{ii} \ . \tag{19.25}$$

This tells us that the diagonal matrix element ρ_i is like a probability for the system to be in the state $|i\rangle$, and the expectation value is a sum weighted over these probabilities. This is analogous to the expression for thermal average of a function $g(E)$: $\langle g\rangle = \sum_E f(E)g(E)$, $f(E) = \exp(-E/k_BT)$ being the Boltzmann probability distribution. This opens a link between the worlds of quantum mechanics and statistical mechanics, and I hope that you will learn more about this in advanced courses.

Finally, the density matrix ρ describing a mixed state has an equation of motion equivalent to the Schrodinger equation called the *Liouville-von Neumann equation* :

$$\begin{aligned}\frac{\mathrm{d}\rho}{\mathrm{d}t} &= \sum_i C_i \left[\frac{\mathrm{d}\,|i\rangle}{\mathrm{d}t} \langle i| + |i\rangle \frac{\mathrm{d}\,\langle i|}{\mathrm{d}t} \right] \quad (\text{using } \ \rho = \sum_i C_i |i(t)\rangle\langle i(t)|) \ , \\ &= \frac{1}{i\hbar} \sum_i C_i \left[H\,|i\rangle\,\langle i| - |i\rangle\,\langle i|\,H \right] \quad (\text{using } \ H|i\rangle = i\hbar \frac{d|i\rangle}{dt} \ , \ \langle i|H = -i\hbar \frac{d\langle i|}{dt}) \ , \\ &= \frac{1}{i\hbar}(H\rho - \rho H) \ , \\ &= \frac{1}{i\hbar}\left[H, \rho\right] \ . \end{aligned} \tag{19.26}$$

19.5 Reduced Density Matrices and von-Neumann Entropy

We can now formulate a measure for quantum entanglement. Consider a system consisting of two subsystems A and B, each with its own basis $\left\{ \left| i^A \right\rangle \right\}$ and $\left\{ \left| i^B \right\rangle \right\}$, described by the density matrix ρ^{AB}. Further, the *reduced density matrix of subsystem A* is defined as

$$\rho^A \equiv \mathrm{Tr}^B \left(\rho^{AB} \right) = \sum_i \left\langle i^B \right| \rho^{AB} \left| i^B \right\rangle \ , \tag{19.27}$$

and is obtained by tracing the density matrix over the states of subsystem B. The von-Neumann entropy (VNE) of a density matrix ρ^{AB} (i.e., for the complete system AB) is defined as

$$\begin{aligned} S &= -\mathrm{Tr}\left(\rho^{AB} \ln \rho^{AB} \right) \\ &= -\sum_i \rho_i^{AB} \ln \rho_i^{AB} \\ &= -\sum_i C_i \ln C_i \ , \end{aligned} \tag{19.28}$$

where we have evaluated the trace in the natural (or diagonal) basis of $\rho^{AB} = \sum_i C_i \, |i\rangle \, \langle i|$. Note the similarity of this expression with the well-known expression (due to Boltzmann) for thermal entropy of a quantum system in thermal equilibrium with a reservoir at temperature T: $S_{thermal} = -k_B \sum_i w_i \ln(w_i)$, where w_i are the Boltzmann probability weights associated with the states i of the system. It is this similarity that led von Neumann to denote the above quantity S as an *entropy*; it is, however, a purely quantum mechanical phenomenon and has nothing to do with our understanding of thermal entropy.

Similarly, the von-Neumann entropy for the reduced density matrix of A is known as the *Entanglement entropy* (EE) and defined as

$$S^A \equiv -\mathrm{Tr}\left[\rho^A \ln \rho^A \right] = -\sum_i \rho_i^A \ln \rho_i^A \tag{19.29}$$

where we evaluated the trace in the natural (or diagonal) basis of ρ^A. This quantity S_A is a measure of the entanglement between the subsystems A and B, and is non-zero only when the complete state of the system *cannot* be written as a *separable state* (i.e., a direct product of the states) of the subsystems A and B. In this way, S^A encodes some kind of a connectivity between the subsystems A and B within the state of the complete system, and its non-zero value arises when we attempt to (somewhat artificially) partition the full system into subsystems (here A and B).

We could also have equivalently carried this out from the perspective of B, and it can be shown that $S_A = S_B$. Below, we will demonstrate this explicitly for the case of two spin-1/2s locked in a singlet state, but it can be shown more generally with a little more hard work using something called the Schmidt decomposition. I will demonstrate this now.

Let us begin by considering two Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$ of dimensions N_A and N_B. The Schmidt decomposition theorem of linear algebra states that for a pure state $|\psi\rangle \in \mathcal{H}_{AB}$ (where $\mathcal{H}_{AB}$ is the tensor product space $\mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B$), there exist two orthonormal sets of states $\{|\psi_{iA}\rangle \in \mathcal{H}_A\}$ and $\{|\phi_{iB}\rangle \in \mathcal{H}_B\}$ with $i = 1, \ldots, N_S$ where $N_S = \min(N_A, N_B)$ such that

$$|\psi\rangle = \sum_{i=1}^{N_S} K_i \, |\psi_{iA}\rangle \, |\phi_{iB}\rangle \quad , \quad 0 \leq K_i \leq 1 \ , \quad \sum_{i=1}^{N_S} K_i^2 = 1 \ . \tag{19.30}$$

The coefficients K_i are called the Schmidt coefficients, N_S the Schmidt number, $|\psi_{iA}\rangle$ and $|\phi_{iB}\rangle$ the Schmidt basis functions and K_i^2 are the probabilities associated with the states $|\psi_{iA}\rangle$ and $|\phi_{iB}\rangle$. The proof of the Schmidt decomposition theorem uses the singular value decomposition theorem, and lies beyond the purview of this course.

Using the $|\psi\rangle$ given above, we obtain the density matrix ρ^{AB} as

$$\rho^{AB} = |\psi\rangle \langle\psi| = \sum_{i=1}^{N_S} K_i \, |\psi_{iA}\rangle \, |\phi_{iB}\rangle \sum_{j=1}^{N_S} K_j \, \langle\psi_{jA}| \, \langle\phi_{jB}| \ , \tag{19.31}$$

such that taking a partial trace over the subsystem B yields

$$\rho^A = \mathrm{Tr}^B \ (\rho^{AB}) = \sum_{i=1}^{N_S} \langle\phi_{iB}| \, \rho^{AB} \, |\phi_{iB}\rangle = \sum_{i=1}^{N_S} K_i^2 \, |\psi_{iA}\rangle \, \langle\psi_{iA}| \ , \tag{19.32}$$

where we have used the fact that $\langle\phi_{iB}|\phi_{jB}\rangle = \delta_{ij}$. Clearly, ρ^A represents a mixed state of the subsystem A. Similarly, by taking a partial trace of ρ^{AB} over subsystem A yields the reduced density matrix for subsystem B as

$$\rho^B = \mathrm{Tr}^A \ (\rho^{AB}) = \sum_{i=1}^{N_S} \langle\psi_{iA}| \, \rho^{AB} \, |\psi_{iA}\rangle = \sum_{i=1}^{N_S} K_i^2 \, |\phi_{iB}\rangle \, \langle\phi_{iB}| \ , \tag{19.33}$$

using $\langle\psi_{jA}|\psi_{iA}\rangle = \delta_{ij}$, and obtains a mixed state for subsystem B. Thus, given that we have now obtained the diagonal representations of both ρ^A and ρ^B, we find that the set of non-trivial eigenvalues $\{K_i^2\}$ are obtained for both ρ^A and ρ^B, i.e.,

$$\{\rho_i^A\} = \{K_i^2\} = \{\rho_i^B\} \ . \tag{19.34}$$

In this way, we have obtained the *spectral equality* of the two reduced density matrices ρ^A and ρ^B. It is now straightforward to see that the entanglement entropies derived from their eigenvalues is the same:

$$S_A = S_B = -\sum_{i=1}^{N_S} K_i^2 \ln K_i^2 \ . \tag{19.35}$$

This is a striking result, as it shows that the EE is not dependent on the details of the individual subsystems A and B. Rather, the EE is dependent on something that is common to them (and which I referred to as their connectivity earlier).

Further, it can be shown that the von Neumann entropy of the complete system (S_{AB}) is less than or equal to the sum of the entanglement entropies of the two subsystems A and B,

$$S_{AB} \leq S_A + S_B \implies S_{AB} \leq 2S_A \quad (\text{as } S_A = S_B) \ . \tag{19.36}$$

This is referred to as the *subadditivity property* of entanglement entropy. The equality is attained for the case of two independent system A and B, such that $\rho^{AB} = \rho^A \otimes \rho^B$. This is because we can write $\rho^{AB} = |\psi\rangle \otimes \langle\psi|$ for such a case, where $|\psi\rangle = \left|\psi^A\right\rangle \otimes \left|\psi^B\right\rangle$.

As an exercise, show that $S_{AB} = 0 = S_A = S_B$ for $\rho^{AB} = \rho^A \otimes \rho^B$.

Let us try to understand why the von Neumann entropy (VNE, S^{AB}) and the entanglement entropy (EE, S^A) should be indicators for the quantum entanglement. First note that if a general density matrix ρ represents a pure state $|i\rangle$, then the corresponding von Neumann entropy (VNE) will be zero. This is simply seen since for a pure state

$$\begin{aligned} \rho = |i\rangle\langle i| = \delta_{ij} \implies \rho_j &= \delta_{ij} \quad (\text{as } \rho_i = 1 \text{ for some one value of } i \text{ and zero otherwise}) \\ \implies S &= -\sum_j \rho_j \ln \rho_j = -\rho_i \ln \rho_i = 0 \ . \end{aligned} \tag{19.37}$$

Notice that the zero appeared because the $0\ln(0)$ terms evaluate to 0 (by l'Hospital's rule) and $1\ln(1) = 0$. Thus, we see that *the VNE for a pure state is zero.* Next, note that since the function $\ln(x)$ is concave (as its derivative, $1/x$, is a strictly decreasing function on the entire interval $(0,\infty)$), we can use Jensen's inequality which states that for a concave function $f(x)$ and a set of positive weights λ_i, the following inequality holds:

$$f\left(\frac{\sum_i \lambda_i x_i}{\sum_i \lambda_i}\right) \geq \frac{\sum_i \lambda_i f(x_i)}{\sum_i \lambda_i} \ , \tag{19.38}$$

which roughly amounts to saying that the function of the weighted arguments must be greater than or equal to the sum of the weighted function of the arguments. Applying this inequality to the VNE by mapping $\lambda_i \equiv \rho_i$, $x_i \equiv 1/\rho_i$, and the function $f(x_i) \equiv \ln(1/\rho_i)$ gives

$$S = -\sum_i \rho_i \ln \rho_i = \sum_i \rho_i \ln \frac{1}{\rho_i} \leq \ln \frac{\sum_i \rho_i \frac{1}{\rho_i}}{\sum_i \rho_i} = \ln N \implies S \leq \ln N \ , \tag{19.39}$$

where we used $\sum_i \rho_i = 1$ and $\sum_i = N$, N being the dimension of the Hilbert space of ρ. We learn the important result that the maximum value of the VNE is $\ln N$. It is now easy to see that this maximum value comes about when $\rho_i = \frac{1}{N} \forall i$, which is a maximally mixed state. To see this, note that

$$\begin{aligned} S(\rho_i = \frac{1}{N}) &= -\sum_{i=1}^{N} \frac{1}{N} \ln\left(\frac{1}{N}\right) \\ &= \sum_{i=1}^{N} \frac{1}{N} \ln(N) \\ &= \frac{N}{N} \ln(N) = \ln(N) \ . \end{aligned} \tag{19.40}$$

A maximally mixed state is, thus, one in which all states appear with equal probability. We can now infer that the more mixed a density matrix is, the more will be its VNE. On the other hand, because a pure state is not mixed at all, we have seen that it's VNE is zero. Thus, you may be tempted into thinking that its entanglement entropy (EE) is 0 as well. However, you need to be cautious here, as even a pure state can be entangled, and therefore can possess a finite and non-zero EE obtained from a reduced density matrix for a subsystem. We will see precisely this with the singlet state for the two spin-1/2 system in the next section.

For now, we will go further with making connecting pure and mixed states with the idea of entanglement, and especially entanglement entropy (EE). As an example, consider a *separable* state for two quantum systems A and B that are both 2 level systems (i.e., both can have two states $|1\rangle$ and $|2\rangle$)

$$|\psi\rangle = \left|1^A\right\rangle \otimes \left(\sqrt{\rho}_1 \left|1^B\right\rangle + \sqrt{\rho}_2 \left|2^B\right\rangle\right) \ , \tag{19.41}$$

with the corresponding density matrix ρ^{AB} being:

$$\rho^{AB} = |\psi\rangle\langle\psi| = \left|1^A\right\rangle\left\langle 1^A\right| \otimes \left(\rho_1 \left|1^B\right\rangle\left\langle 1^B\right| + \rho_2 \left|2^B\right\rangle\left\langle 2^B\right| + \sqrt{\rho_1\rho_2}\left|1^B\right\rangle\left\langle 2^B\right| + \sqrt{\rho_1\rho_2}\left|2^B\right\rangle\left\langle 1^B\right|\right) \ . \tag{19.42}$$

Then, tracing out subsystem B will leave us with a pure state density matrix

$$\begin{aligned} \rho^A &= \left\langle 1^B\right| \rho^{AB} \left|1^B\right\rangle + \left\langle 2^B\right| \rho^{AB} \left|2^B\right\rangle \\ &= (\rho_1 + \rho_2)\left|1^A\right\rangle\left\langle 1^A\right| = |\chi\rangle\langle\chi| = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \end{aligned} \tag{19.43}$$

$$\tag{19.44}$$

where $|\chi\rangle = \sqrt{\rho_1 + \rho_2}\left|1^A\right\rangle$, and the cross terms ($\propto \sqrt{\rho_1\rho_2}$) vanish because of the orthogonality condition $\langle \alpha^i | \beta^j \rangle = \delta_{\alpha\beta}\delta_{ij}$ for $(\alpha, \beta) = (1, 2)$ and $(i, j) = (A, B)$. It is now easy to obtain the entanglement entropy as

$$S_A = -\sum_j \rho_j \ln \rho_j = -1 \times \ln(1) - 0 \times \ln(0) = 0 \ . \tag{19.45}$$

Thus, *separable states lead to pure state density matrices upon reduction, and the entangled entropy from such pure state density matrices is zero.*

Instead, if we start with a *non-separable* state $|\psi\rangle = \sqrt{\rho}_1 \left|1^A\right\rangle \left|1^B\right\rangle + \sqrt{\rho}_2 \left|2^A\right\rangle \left|2^B\right\rangle$, the corresponding density matrix obtained is

$$\begin{aligned} \rho^{AB} &= |\psi\rangle\langle\psi| \\ &= \rho_1 \left|1^A\right\rangle \left\langle 1^A \middle| 1^B \right\rangle \left\langle 1^B\right| + \rho_2 \left|2^A\right\rangle \left\langle 2^A \middle| 2^B \right\rangle \left\langle 2^B\right| \\ &\quad + \sqrt{\rho_1\rho_2}\left(\left|1^A\right\rangle \left\langle 2^A \middle| 1^B \right\rangle \left\langle 2^B\right| + \left|2^A\right\rangle \left\langle 1^A \middle| 2^B \right\rangle \left\langle 1^B\right|\right) . \end{aligned} \tag{19.46}$$

Then tracing out subsystem B would leave us with:

$$\begin{aligned} \rho^A &= \left\langle 1^B\right| \rho_{AB} \left|1^B\right\rangle + \left\langle 2^B\right| \rho_{AB} \left|2^B\right\rangle \\ &= \rho_1 \left|1^A\right\rangle \left\langle 1^A\right| + \rho_2 \left|2^A\right\rangle \left\langle 2^A\right| = \begin{pmatrix} \rho_1 & 0 \\ 0 & \rho_2 \end{pmatrix} , \quad (\rho_1, \rho_2 > 0) \end{aligned} \tag{19.47}$$

where the cross terms ($\propto \sqrt{\rho_1\rho_2}$) vanish because of the orthogonality condition $\langle \alpha^i | \beta^j \rangle = \delta_{\alpha\beta}\delta_{ij}$ for $(\alpha, \beta) = (1, 2)$ and $(i, j) = (A, B)$. This is clearly a mixed state, and leads to an EE given by

$$S_A = -\sum_j \rho_j \ln \rho_j = -\rho_1 \ln \rho_1 - \rho_2 \ln \rho_2 \neq 0 \ , \quad (\rho_1, \rho_2 > 0) \ . \tag{19.48}$$

Thus, *non-separable, or entangled, states lead to mixed state density matrices upon reduction; further, such mixed state density matrices obtain non-zero entanglement entropy.* Also, the entanglement entropy obtained from a reduced density matrix must also be bounded by 0 and $\ln(N)$ (where N is the dimension of the Hilbert space of the reduced density matrix), just as we had observed earlier for the von Neumann entropy.

This establishes the connection between the notions and entanglement and separability of states with the ideas of pure and mixed states. Further, it shows that entanglement reflects on some kind of *connectivity between different parts of a system within the eigenstates.* In this way, entanglement can be quantified by the entanglement entropy (EE). However, it is worth noting that EE is only one measure of entanglement; there are many other quantitative measures of entanglement that we will not discuss here.

The take home message: Entangled states will lead to mixed states upon reduction, and the EE of such states will be non-zero. Separable states, on the other hand will produce pure states upon reduction, and lead to vanishing EE.

19.6 Entanglement Entropy of the Singlet, and what it tells us about Schrödinger's Cat

We can now calculate the EE of the singlet ground state of the Heisenberg Hamiltonian for two spins A and B, and hence quantify its entanglement content. The eigenstate is

$|\psi\rangle_{AB} = \frac{1}{\sqrt{2}}\left(|\uparrow_A\rangle\,|\downarrow_B\rangle - |\downarrow_A\rangle\,|\uparrow_B\rangle\right)$. The corresponding density matrix is given by

$$\rho^{AB} = \frac{1}{2}\left(|\uparrow_A\rangle\,|\downarrow_B\rangle - |\downarrow_A\rangle\,|\uparrow_B\rangle\right)\left(\langle\uparrow_A|\,\langle\downarrow_B| - \langle\downarrow_A|\,\langle\uparrow_B|\right)~. \tag{19.49}$$

The reduced density matrix for spin A (obtained upon tracing out spin B) is

$$\rho^{A} = \langle\uparrow_B|\,\rho_{AB}\,|\uparrow_B\rangle + \langle\downarrow_B|\,\rho_{AB}\,|\downarrow_B\rangle = \frac{1}{2}\left(|\uparrow_A\rangle\,\langle\uparrow_A| + |\downarrow_A\rangle\,\langle\downarrow_A|\right) = \begin{pmatrix}\frac{1}{2} & 0\\ 0 & \frac{1}{2}\end{pmatrix}~, \tag{19.50}$$

where the other terms identically vanish as $\langle\uparrow_B\mid\downarrow_B\rangle = 0 = \langle\downarrow_B\mid\uparrow_B\rangle$. The reduced density matrix ρ^A clearly corresponds to a maximally mixed denisty matrix whose weights are given by $\rho^A_i = 1/N$, with $N=2$ being the dimension of the $SU(2)$ Hilbert space (i.e., the Hilbert space of the spin-1/2 degree of freedom). The EE is then found to be

$$S_A = -\sum_i \rho_i \ln\rho_i = 2\times\frac{1}{2}\ln 2 = \ln 2~. \tag{19.51}$$

In this way, we see that the singlet has the maximum possible entanglement, i.e., it is *maximally entangled*. Also, it is quite obvious from the construction of the singlet state that precisely the same result would be obtained if we were to trace out spin A, and obtain the EE for spin B: $S_B = \ln 2 = S_A$.

Furthermore, we have learnt **another important lesson**: *interactions between the constituents of a system give rise to the entanglement of the many-constituent wavefunction.* We saw this here for the simplest case of two subsystems (i.e., the spins) that are interacting with one another through the Heisenberg Hamiltonian.

We can now bring up an old gedanken to highlight the consequences what we have just seen. Recall the Schrödinger's cat gedanken we encountered a long time ago in these lectures.

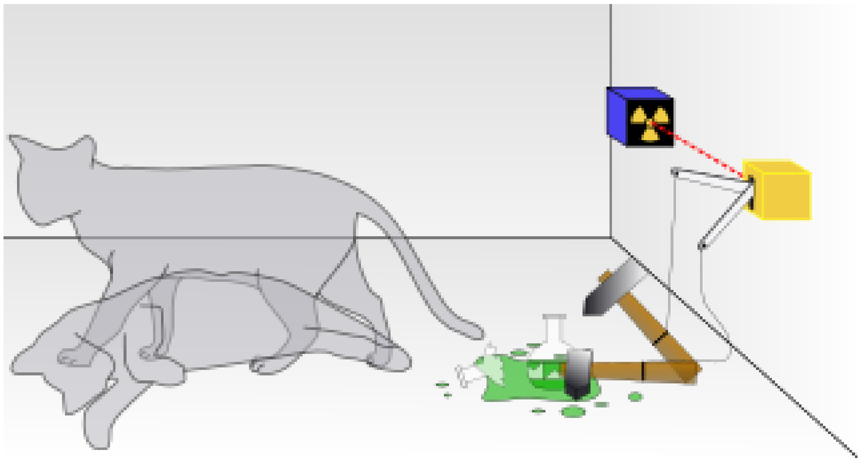


Figure 19.3: A schematic diagram of the Schrödinger's Cat thought experiment. See discussion in text. Disclaimer: No such experiment has ever been carried out with a real cat! Source: The internet.

A quick recapitulation. A cat is placed in an opaque box together with a radioactive source, a Geiger counter that can detect the radioactivity and a vial of poison gas. Any radioactive source has a well-defined lifetime, i.e., a definite fraction of the population of the atoms will have undergone radioactive decay within the lifetime. The decay process is, however, probabilistic; this means that at any given time, we can associate a probability that a given atom will have undergone radioactive decay. Now, if during the time the cat is kept in the box, the radioactive source undergoes decay, the Geiger counter will sense this and break the vial of poison gas, killing the cat. Thus, upon opening the lid of the box, we will find the cat to be dead. On the other hand, if the source does not decay, the cat will be alive upon observation.

Now, it is easy to see that the state of the cat and that of the atom are entangled with one another. Since the cat being alive or dead was contingent upon the radioactive state of an atom (i.e., on whether it "has decayed" or "has not decayed"). Thus, the state of the full "Cat + Atom" system can be written as

$$|\psi_{\text{Cat+Atom}}\rangle = \frac{1}{\sqrt{2}}(|\text{Cat Alive}\rangle|\text{Atom not decayed}\rangle + |\text{Cat Dead}\rangle|\text{Atom decayed}\rangle) \tag{19.52}$$

As both the Cat and the Atom have two states each, we can replace "Cat" by Spin 1 and "Atom" by Spin 2. Then, the above is nothing but the triplet zero state

$$|S=1, S_z=0\rangle = \frac{1}{\sqrt{2}}(|\uparrow_1\rangle|\downarrow_2\rangle + |\downarrow_1\rangle|\uparrow_2\rangle) \ , \tag{19.53}$$

which is equally entangled a state as the singlet we discussed just above. By this we mean that if we compute the reduced density matrix for the Cat by tracing over the two states of the Atom, we will again receive the reduced density matrix given above in eq.(19.50) and the entanglement entropy as shown in eq.(19.51). Thus, we can see that as long as the lid of the box is closed, the cat and atom exist in an entangled state. We will discuss at a later point what happens when we open the lid and make a measurement of the cat's state.

19.7 The EPR Paradox

Some interesting history and related physics can be brought up. Schrödinger coined the word "entanglement" for the phenomenon we have just observed, and noted that this was something very intrinsic to quantum mechanics. However, this very feature was also dubbed as a "spooky action at a distance" type correlation between subsystems of a quantum mechanical system by Einstein, Podolsky and Rosenberg (EPR). Thus, they came up with a gedanken to highlight what they thought was a paradox. That is, they wanted to show just how weird entanglement was, and proposed that nature just could not be this way. In keeping with our discussion above of the entanglement within the singlet state for two interacting spin-1/2 degrees of freedom, we present David Bohm's version of the EPR paradox below.

Consider the following. A system of two spin-1/2 angular momenta are coupled to one another, such that they are initially placed in the singlet state ($\frac{1}{\sqrt{2}}(|\uparrow_1, \downarrow_2\rangle - |\downarrow_1, \uparrow_2\rangle)$). Now, send one of the spins (spin 1) off to (say) an observer A placed at $x \to \infty$ and the other (spin

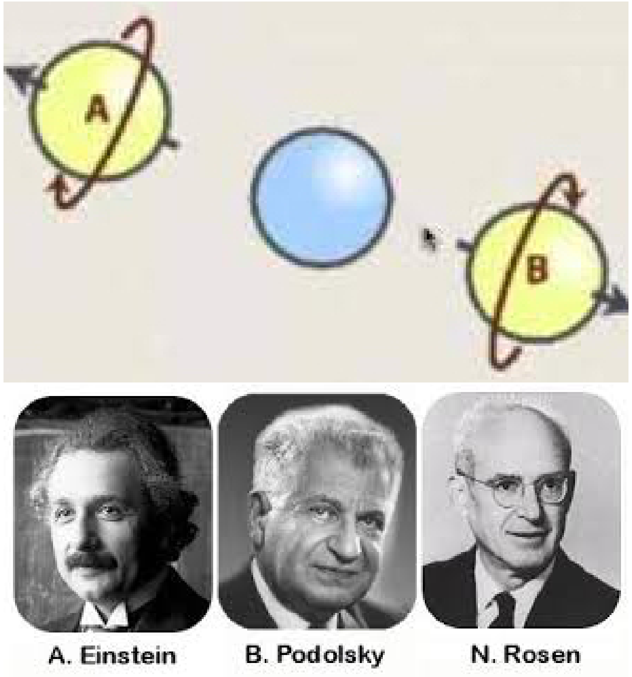


Figure 19.4: (Above) The EPR paradox with a spin singlet state. See text for discussion. (Below) From left to right: Einstein, Podolsky and Rosen.

2) to an observer B placed at $x \to -\infty$. Remember that we are assuming that, even as the spins fly apart from one another, they remain in the entangled singlet state. (This amounts to saying that there exists no mechanism that can spoil their being entangled through any form of measurement, decoherence etc. even as they are flying apart.)

Now, let's say that observer A receives her spin first (i.e., even before B has received his) and takes a single measurement of the value of S_{1z}. Given that the singlet state is not an eigenstate of the individual S_{1z} and S_{2z}, she has a 50% chance of obtaining $\langle S_{1z} \rangle = \hbar/2$ and 50% chance of obtaining $\langle S_{1z} \rangle = -\hbar/2$. The same would have been true if observer B were to obtain spin 2 first and make a measurement of $\langle S_{2z} \rangle$. Further, this would also be true if either of them were to receive their spin first and measure, say, $\langle S_{1x} \rangle$ or $\langle S_{2x} \rangle$.

However, any single measurement (rather than the expectation value) will cause the wavefunction of the coupled spin system to *collapse* onto one of its components, i.e., the initial state $\frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle)$ will collapse to either $|\uparrow\downarrow\rangle$ or $|\downarrow\uparrow\rangle$. (We will offer more insight into measurement and wavefunction collapse in a following section.) If both observers A and B knew that the initial state of the two spins was a singlet, then A has received *instantaneously* from her measurement not only the state of the spin she received (spin 1, here $|\uparrow\rangle$), but also the state of the spin she did not have (spin 2, here $|\downarrow\rangle$)!

EPR noted that this was strange: A's measurement of S_{1z} would cause the collapse of not only spin 1 onto (say) $|\uparrow\rangle$, but also simultaneously (i.e., instantaneously) also cause the collapse of spin 2 onto $|\downarrow\rangle$ even though it was infinitely far away! Thus, it would appear that information has travelled faster than the speed of light between the two spins so as to

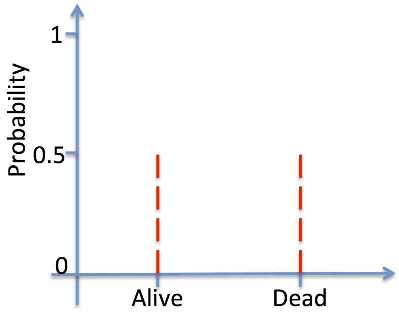


Figure 19.5: A schematic diagram of the probability for observations of the "Dead" and "Alive" states in the Schrödinger's Cat thought experiment by a classical observer who only observes the state of the cat after a certain time interval T. In the EPR gedanken, replace "Dead" by $\langle S_{1z}\rangle = \hbar/2$ and "Alive" by $\langle S_{1z}\rangle = -\hbar/2$ for measurements of $\langle S_{1z}\rangle$ by observer A on spin 1 prior to observer B receiving his spin.

collapse spin 2 into $|\downarrow\rangle$ as the same instant as spin 1's state collapsed into $|\uparrow\rangle$. This, for them, this showed the *spooky action at a distance* nature inherent within quantum mechanics. Well, it got worse. As we shall see, the collapse of the state by observer A will severely affect the measurements of observer B on spin 2.

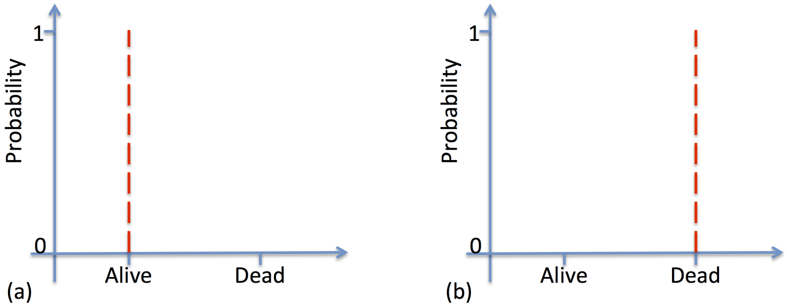


Figure 19.6: A schematic diagram of the probability for observations of the "Dead" and "Alive" states in the Schrödinger's Cat thought experiment by continuous classical observations. In the EPR gedanken, replace "Dead" by $\langle S_{2z}\rangle = \hbar/2$ and "Alive" by $\langle S_{2z}\rangle = -\hbar/2$ for measurements of $\langle S_{2z}\rangle$ by observer B on spin 2 after observer A has already measured spin 1 (and observes $\langle S_{1z}\rangle = -\hbar/2$ and $\langle S_{1z}\rangle = \hbar/2$ respectively).

Imagine that observer B decides to measure either S_{2z} or S_{2x} (but only one of them at any instant). Given that spin 2 has already collapsed onto $|\downarrow\rangle$, then B will find $\langle S_{2z}\rangle = -\hbar/2$ will absolute certainty (and not the half-half chance of finding $\pm\hbar/2$ if A had not already made her measurement)! Thus, it appears that information has travelled faster than light between the two spins (and which are infinitely far away) so as to fix observer B's spin (spin 2) into a definite value (say, $-\hbar/2$) along an appropriately chosen axis (say, the z-axis). Further, if B measures $\langle S_{2x}\rangle$, then he has equal (i.e., half-half) chances of getting $|\rightarrow\rangle$ and $|\leftarrow\rangle$. On the other hand, if A had not measured her spin, then B's measurements of either $\langle S_{2z}\rangle$ or $\langle S_{2x}\rangle$ would have yielded equal (i.e., half-half) chances of getting $|\uparrow\rangle$, $|\downarrow\rangle$ or $|\rightarrow\rangle$, $|\leftarrow\rangle$ respectively. Further, EPR also worried about the fact that observer A could, having already obtained the information of the state of spin 2, convey the information to observer B *even before* he could receive and measure it for himself!

How could all of this possibly be? Surely this defies the belief from the special theory of relativity that information could not be conveyed between two systems (and their observers) faster than the speed of light. Well, this effect has been tested in the laboratory between photons (and over increasing distances, e.g., between the surface of the Earth and an orbiting satellite) several times, and each time the entanglement has shown up! There have been several further theoretical developments in quantum mechanics, e.g., Bell's inequalities, towards understanding whether this entanglement could be arising from a more deterministic theory (the so called "hidden variables" theory) that remains unknown to us. However, all of these theoretical proposals have also been tested experimentally, and they all show (i) no signs of any hidden underlying theory and (ii) that the entanglement inherent in quantum mechanics is very much present and real. Indeed, the Nobel prize in Physics for 2022 was awarded to Alain Aspect, John F. Clauser and Anton Zeilinger for "for experiments with entangled photons, establishing the violation of Bell inequalities and pioneering quantum information science".

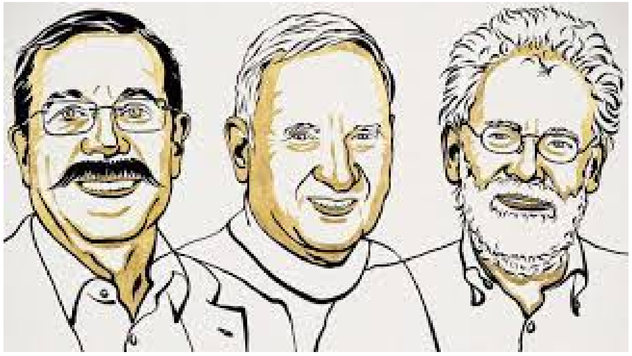


Figure 19.7: (Above) From left to right: Alain Aspect, John F. Clauser and Anton Zeilinger. They were awarded the Nobel prize for Physics in 2022 for their pioneering work on quantum entanglement.

We also understand that it is not possible for observers A and B to communicate among themselves the states of each other spins even before the second of them receives theirs, i.e., the foreknowledge of the state of both spins by the first of the observers cannot be shared with the second prior to that person receiving their spin. There is thus no violation of the special theory of relativity. I am sure that you will learn more about this in further courses in quantum mechanics.

19.8 The 2-spins in the Presence of a Magnetic Field: a toy model for a quantum phase transition

A few words on phase transitions.

You may have studied about phase transitions somewhere earlier, with water being used as a typical example to show the changes in the state of matter between solid (ice), liquid (water) and gas (vapour) as temperature is raised from below $T < 273K$ to above $T > 373K$. Recall that these drastic changes in the phase (or state of matter) were driven by the increasingly large affects of thermal energy ($k_B T$) being given to the water molecules as the temperature T was raised, i.e., such "phase transitions" are driven by thermal fluctuations in the system. The question we ask here is: can there be counterparts of such phase transitions in a quantum system at $T = 0$? Certainly, such transitions would have to be observed by changing some other parameter in the problem (instead of T), and would be driven by quantum (rather than thermal) fluctuations.

Recall that we have given a physical description to quantum fluctuations earlier in these lectures in terms of the zero-point energy that we observed in simple quantum systems such as the particle in a box and the simple harmonic oscillator; in both of those cases, quantum fluctuations/zero point energy was responsible for the non-zero ground state energy. In a quantum system with a thermodynamically large number of constituents who are interacting one another, such zero point energy (or quantum fluctuations) can even drive a phase transition at $T = 0$. What is the typical signature of such a quantum phase transition?b This is easily seen from the following argument. We note that all thermal phase transitions can be captured by noting the existence of some kind of singular behaviour in the thermodynamic free energy F

$$F = E + TS, \tag{19.54}$$

where E is the internal energy of the system (described by it's Hamiltonian's ground state eigenvalue), S is the thermal entropy and T is the temperature. The singular behaviour is typically observed in the discontinuous behaviour of either the first or some higher derivative of the free energy F with respect to T or some other parameter (such as an external magnetic field for the case of magnetism).

At $T = 0$, the free energy is simply $F \equiv E$, and we can easily see that a quantum phase transition must be signalled in some form of singular behaviour in E. The simplest form of a quantum phase transition corresponds to a discontinuous behaviour in E (or one of its derivatives) as some parameter in the quantum Hamiltonian is varied. We will now study such behaviour in a simple toy model.

The toy model.

Recall that we had studied earlier the problem of two spin-1/2 degrees of freedom interacting via an antiferromagnetic Heisenberg spin exchange. The more general Heisenberg problem consists of two interacting spins placed in a magnetic field

$$H = \lambda \vec{S_1} \cdot \vec{S_2} + B\left(S_{1z} + S_{2z}\right), \tag{19.55}$$

where we have assumed that $\lambda > 0$ (antiferromagnetic spin exchange coupling) and the magnetic field (B) points along the z-direction (but its sign can be either positive or negative). The magnetic term corresponds to a Zeeman coupling between the total magnetisation of the system ($(S_{1z} + S_{2z})$) and the external B field. This Hamiltonian can again be diagonalised easily. Using $S_\pm = S_x \pm iS_y$, the Hamiltonian can be written as

$$\vec{S_1} \cdot \vec{S_2} = \sum_{i=\{x,y,z\}} S_{1i} S_{2i} = S_{1z} S_{2z} + \frac{1}{2}\left(S_{1+} S_{2-} + S_{1-} S_{2+}\right) \tag{19.56}$$

We can now write the matrix elements in the basis of S_{1z}, S_{2z}.

$$H = \begin{array}{cccc} |{\uparrow_1}, {\uparrow_2}\rangle & |{\uparrow_1}, {\downarrow_2}\rangle & |{\downarrow_1}, {\uparrow_2}\rangle & |{\downarrow_1}, {\downarrow_2}\rangle \\ \end{array}$$

$$H = \left(\begin{array}{cccc} \frac{1}{4}\lambda + B & 0 & 0 & 0 \\ 0 & -\frac{1}{4}\lambda & \frac{1}{2}\lambda & 0 \\ 0 & \frac{1}{2}\lambda & -\frac{1}{4}\lambda & 0 \\ 0 & 0 & 0 & \frac{1}{4}\lambda - B \end{array} \right) \tag{19.57}$$

The Hamiltonian has three decoupled blocks; the top left and bottom right form two diagonal blocks, while the middle 2×2 matrix forms the third block which is internally off-diagonal. Since the top left and bottom right blocks are diagonal, we can immediately write down two of the four eigenstates and eigenvalues as:

$$\begin{aligned} H\,|{\uparrow_1}, {\uparrow_2}\rangle &= \left(\frac{1}{4}\lambda + B\right) |{\uparrow_1}, {\uparrow_2}\rangle \\ H\,|{\downarrow_1}, {\downarrow_2}\rangle &= \left(\frac{1}{4}\lambda - B\right) |{\downarrow_1}, {\downarrow_2}\rangle \end{aligned} \tag{19.58}$$

The middle block can be easily diagonalised (and has been done earlier). The eigenstates are

$$\begin{aligned} H\frac{1}{\sqrt{2}}\left(|{\uparrow_1}, {\downarrow_2}\rangle + |{\downarrow_1}, {\uparrow_2}\rangle\right) &= \frac{1}{4}\lambda\frac{1}{\sqrt{2}}\left(|{\uparrow_1}, {\downarrow_2}\rangle + |{\downarrow_1}, {\uparrow_2}\rangle\right) \\ H\frac{1}{\sqrt{2}}\left(|{\uparrow_1}, {\downarrow_2}\rangle - |{\downarrow_1}, {\uparrow_2}\rangle\right) &= -\frac{3}{4}\lambda\frac{1}{\sqrt{2}}\left(|{\uparrow_1}, {\downarrow_2}\rangle - |{\downarrow_1}, {\uparrow_2}\rangle\right) \end{aligned} \tag{19.59}$$

$E(\hbar = 1)$	$\|\Psi\rangle$	S	S_z
$-\frac{3}{4}\lambda$	$\frac{1}{\sqrt{2}}\left(\|{\uparrow_1}, {\downarrow_2}\rangle - \|{\downarrow_1}, {\uparrow_2}\rangle\right)$	0	0
$\frac{1}{4}\lambda$	$\frac{1}{\sqrt{2}}\left(\|{\uparrow_1}, {\downarrow_2}\rangle + \|{\downarrow_1}, {\uparrow_2}\rangle\right)$	1	0
$\frac{1}{4}\lambda + B$	$\|{\uparrow_1}, {\uparrow_2}\rangle$	1	1
$\frac{1}{4}\lambda - B$	$\|{\downarrow_1}, {\downarrow_2}\rangle$	1	-1

Table 19.2: Spectrum of Heisenberg Hamiltonian in presence of $\vec{B}$

The first distinction from the $B = 0$ case is that the *magnetic field breaks the degeneracy of the triplet states* (recall that we have also encountered this earlier in our discussion of the ferromagnetic Heisenberg spin exchange problem). The more important effect of the field, however, is that we can now tune the ground state of the system by varying the value of B. For concreteness, let us consider first the case of B being positive (and comment on the case of $B < 0$ later). For $B > 0$, the two lowest eigenvalues are $-\frac{3}{4}\lambda$ corresponding to the singlet ($\frac{1}{\sqrt{2}}\left(|{\uparrow_1}, {\downarrow_2}\rangle - |{\downarrow_1}, {\uparrow_2}\rangle\right)$), and $\frac{1}{4}\lambda - B$ corresponding to $|{\downarrow_1}, {\downarrow_2}\rangle$. The energy difference between these two states

$$\Delta E_{gap} = E_{|{\downarrow_1}{\downarrow_2}\rangle} - E_{\frac{1}{\sqrt{2}}(|{\uparrow_1}, {\downarrow_2}\rangle - |{\downarrow_1}, {\uparrow_2}\rangle)} = \lambda - B\ . \tag{19.60}$$

As shown in Fig.19.8, for $B < B^* \equiv \lambda$, the singlet ($\frac{1}{\sqrt{2}}(|\uparrow_1, \downarrow_2\rangle - |\downarrow_1, \uparrow_2\rangle)$) is still the ground state (with energy $E_g = -3\lambda/4$) and the gap is finite and positive, $\Delta E_{gap} > 0$. However, if we tune the magnetic field above this value, then *the ground state suddenly changes to the* $|\downarrow_1, \downarrow_2\rangle$ *state* (with $E_g = \lambda/4 - B$). This level-crossing between the $|\downarrow_1\downarrow_2\rangle$ and $\frac{1}{\sqrt{2}}(|\uparrow_1, \downarrow_2\rangle - |\downarrow_1, \uparrow_2\rangle)$ states marks a zero-temperature zero-dimensional quantum transition between the singlet and the triplet down states, with there being a degeneracy (i.e., $\Delta E_{gap} = 0$) precisely at the transition. While the ground state energy (E_g) changes smoothly across $B = B^*$, the derivative $(\partial E_g/\partial B)|_{B=B^*}$ undergoes an abrupt change across this point (from $(\partial E_g/\partial B)|_{B<B^*} = 0$ to $(\partial E_g/\partial B)|_{B>B^*} = -1$). In this way, the observed level crossing here is a simple toy model for the quantum phase transitions that are studied by experimentalists and theorists in more complex problems. Note that for $B < 0$, precisely the same level crossing phenomenon is encountered between the $|\uparrow_1\uparrow_2\rangle$ and $\frac{1}{\sqrt{2}}(|\uparrow_2, \downarrow_2\rangle - |\downarrow_1, \uparrow_2\rangle)$ states at $B^* = -\lambda$. This is also shown in in Fig.19.8.

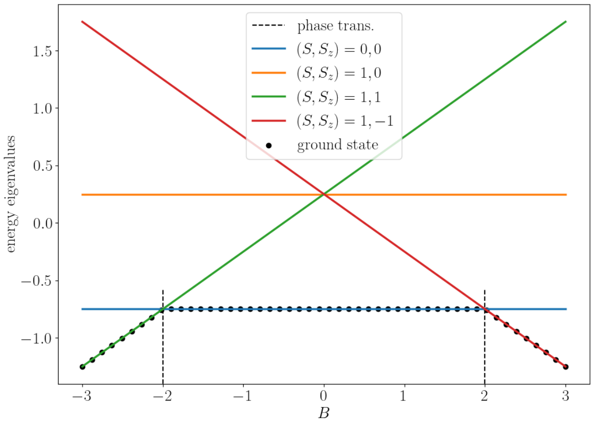


Figure 19.8: Variation of all four eigenvalues with the magnetic field, for $\lambda = 2$. There is a quantum transition at the critical value $B^* = \lambda$, where the ground state changes from $S = 0, S_z = 0$ to $S = 1, S_z = -1$.

We can also track the transition by monitoring the entanglement of the ground state. For $B < B^* \equiv \lambda$, the ground state is a singlet, so the entanglement entropy (EE) of the reduced density matrix (with respect to either the first or the second spin) will be $\ln 2$. As we cross the critical magnetic field, the new ground state ($|\downarrow_1, \downarrow_2\rangle$ is a separable state and will hence have zero EE.

$$S_1 = \mathrm{Tr}\left[\rho_1 \ln \rho_1\right] = \begin{cases} \ln 2, & B < B^* \equiv \lambda \\ 0, & B > B^* \equiv \lambda \end{cases} \tag{19.61}$$

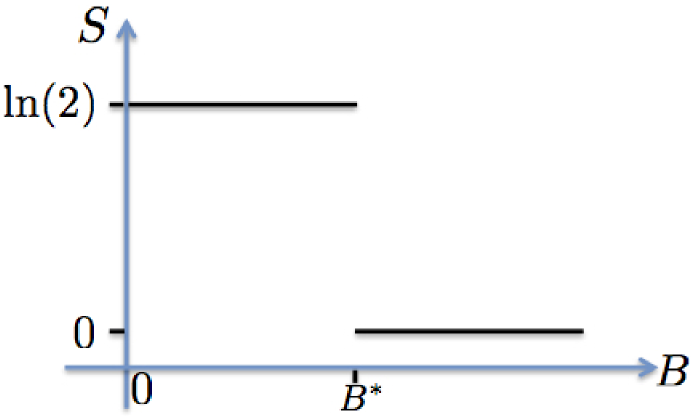


Figure 19.9: Variation of entanglement entropy S_1 with the external magnetic field B. The transition at $B = B^*$ is visible as S_1 undergoes a sudden lowering in value from $\ln 2$ for $B < B^*$ to 0 for $B > B^*$.

An effective theory for the transition.
We can even go one step further, and obtain an *effective* understanding of the transition in terms of only the states that undergo the level crossing transition. For this, we first define the effective subspace of the singlet and triplet down states undergoing the transition as follows

$$|\tilde{\uparrow}\rangle = \frac{1}{\sqrt{2}}(|\uparrow_1\downarrow_2\rangle - |\downarrow_1\uparrow_2\rangle) \quad , \quad |\tilde{\downarrow}\rangle = |\downarrow_1\downarrow_2\rangle \quad , \quad \langle\tilde{\uparrow}|\tilde{\downarrow}\rangle = 0 \ . \tag{19.62}$$

Then, we define the effective spin-1/2 operators that act on only this subspace of two states as follows

$$\tilde{S}_z = (S_{1,z} + S_{2,z} + \hbar\frac{\mathcal{I}}{2}) \ , \ \tilde{S}_+ = \frac{1}{\sqrt{2}}(S_{1,+} - S_{2,+}) \ , \ \tilde{S}_- = \frac{1}{\sqrt{2}}(S_{1,-} - S_{2,-}) \ . \tag{19.63}$$

We can easily see that

$$\begin{aligned} \tilde{S}_z|\tilde{\uparrow}\rangle &= \frac{\hbar}{2}|\tilde{\uparrow}\rangle \ , \ \tilde{S}_z|\tilde{\downarrow}\rangle = -\frac{\hbar}{2}|\tilde{\downarrow}\rangle \ , & (19.64)\\ \tilde{S}_+|\tilde{\downarrow}\rangle &= \hbar|\tilde{\uparrow}\rangle \ , \ \tilde{S}_-|\tilde{\uparrow}\rangle = \hbar|\tilde{\downarrow}\rangle \ , & (19.65)\\ \tilde{S}_+|\tilde{\uparrow}\rangle &= 0 = \tilde{S}_-|\tilde{\downarrow}\rangle \ . & (19.66) \end{aligned}$$

Now, we can construct the effective Hamiltonian H_{eff} governing the quantum phase transition from the following four matrix elements

$$\begin{aligned} \langle\tilde{\uparrow}|H|\tilde{\uparrow}\rangle &= -\frac{3\lambda}{4} \ , \ \langle\tilde{\downarrow}|H|\tilde{\downarrow}\rangle = (\frac{\lambda}{4} - B) \ , & (19.67)\\ \langle\tilde{\uparrow}|H|\tilde{\downarrow}\rangle &= 0 = \langle\tilde{\downarrow}|H|\tilde{\uparrow}\rangle \ , & (19.68) \end{aligned}$$

to write

$$H_{eff} = \begin{pmatrix} -\frac{3\lambda}{4} & 0 \\ 0 & (\frac{\lambda}{4} - B) \end{pmatrix} = a\tilde{\sigma}_z + b\mathcal{I} \ , \tag{19.69}$$

where a, b are constants to be determined, $\tilde{\sigma}_z = \frac{2}{\hbar}\tilde{S}_z$ and $\mathcal{I}$ is the 2×2 identity matrix. Solving for a and b, we obtain

$$a = -\frac{1}{2}(\lambda - B) \equiv -\frac{1}{2}\Delta E_{gap} \;\; , \;\; b = -\frac{1}{4}(\lambda + B) \; . \tag{19.70}$$

Thus, we can write the effective Hamiltonian as

$$H_{eff} = -\frac{\Delta E_{gap}}{\hbar}\tilde{S}_z - \frac{1}{4}(\lambda + B)\mathcal{I} \; . \tag{19.71}$$

The effective Hamiltonian can be seen to take the form of a simple Zeeman-like coupling between an effective external field (ΔE_{gap}) and the effective spin-1/2 degree of freedom given by $\tilde{S}_z$. From this effective Hamiltonian, we can see that tuning the external B-field leads to the vanishing of ΔE_{gap} at the quantum critical point $B \equiv B^* = \lambda$, such that the effective Hamiltonian for the two degenerate levels ($|\tilde{\uparrow}\rangle$ and $|\tilde{\downarrow}\rangle$) at the critical point is simply a constant given by

$$H_{eff}^{QCP} = H_{eff}|_{B=\lambda} = -\frac{1}{4}(\lambda + B)\mathcal{I} \; . \tag{19.72}$$

In this way, we have been able to obtain *the effective low-energy theory for the quantum transition* in this simple model. It is worth noting that physicists working in the fields of statistical physics, condensed matter physics, high energy physics, cosmology and even string theory spend a fair bit of their time obtaining the effective low-energy theory for the quantum phase transitions that appear in more complicated quantum systems with many interacting constituents. They have to use techniques (e.g., the renormalisation group) from quantum field theory, many-body theory and quantum statistical field theory to make progress with their goal, but their aim is pretty much the same: find a way to obtain the effective theory that governs the physics of the relevant degrees of freedom within the system that undergo the quantum phase transition. Furthermore, they also want to obtain from their effective theories insights into the emergence of novel collective states of quantum matter from the inter-constituent interactions, e.g., magnetism, superconductors, insulators of various kinds, protons, neutrons and much more!

19.9 Wavefunction Collapse, Measurement and Projection

We can use our discussion of the 2-spin problem to make some remarks about making measurements and wavefunction collapse. Recall that we saw earlier in the Schrödinger cat gedanken that the state of the full "Cat + Atom" system can be written as

$$|\psi_{\text{Cat+Atom}}\rangle \;\; = \;\; \frac{1}{\sqrt{2}}(|\text{Cat Alive}\rangle|\text{Atom not decayed}\rangle + |\text{Cat Dead}\rangle|\text{Atom decayed}\rangle) \tag{19.73}$$

As both the Cat and the Atom have two states each, we can replace "Cat" by Spin 1 and "Atom" by Spin 2. Then, the above is nothing but the triplet zero state

$$|S = 1, S_z = 0\rangle = \frac{1}{\sqrt{2}}(|\uparrow_1\rangle|\downarrow_2\rangle + |\downarrow_1\rangle|\uparrow_2\rangle) \; , \tag{19.74}$$

which is a maximally entangled state, i.e., tracing over all states of the Atom gives us a reduced density matrix that shows maximal mixed nature. Now, consider that you have a switch by which to either (i) stop altogether the Atom from decaying or (ii) force the Atom to decay instantaneously. It is clear that if we flip the switch to the first case (i.e., stop altogether the Atom from decaying), then the Cat is alive for all times. On the other hand, if we flip the switch to the second case (i.e., force the Atom to decay instantaneously), the Cat is then dead thereafter for all times. But surely, each of these events has "collapsed the state" of the Cat and Atom system from the above superposition, $\frac{1}{\sqrt{2}}(|\uparrow_1\rangle|\downarrow_2\rangle + |\downarrow_1\rangle|\uparrow_2\rangle))$, to either $|\uparrow_1\rangle|\downarrow_2\rangle$ or $|\downarrow_1\rangle|\uparrow_2\rangle$.

But how do we study such a collapse of the state? The simplest way to model it is to define the following two projection operators (in terms of outer products) of the orthonormal states $|\uparrow_1\rangle|\downarrow_2\rangle$ and $|\downarrow_1\rangle|\uparrow_2\rangle$

$$P_1 = |\uparrow_1\rangle|\downarrow_2\rangle\langle\uparrow_1|\langle\downarrow_2| \quad , \quad P_2 = |\downarrow_1\rangle|\uparrow_2\rangle\langle\downarrow_1|\langle\uparrow_2| \ , \tag{19.75}$$

such that, by acting on the normalised eigenstate $\frac{1}{\sqrt{2}}(|\uparrow_1\rangle|\downarrow_2\rangle + |\downarrow_1\rangle|\uparrow_2\rangle)$ with P_1 and P_2, we can obtain

$$|\psi_1\rangle = P_1\frac{1}{\sqrt{2}}(|\uparrow_1\rangle|\downarrow_2\rangle + |\downarrow_1\rangle|\uparrow_2\rangle) \tag{19.76}$$

$$= |\uparrow_1\rangle|\downarrow_2\rangle\langle\uparrow_1|\langle\downarrow_2|\frac{1}{\sqrt{2}}(|\uparrow_1\rangle|\downarrow_2\rangle + |\downarrow_1\rangle|\uparrow_2\rangle) \tag{19.77}$$

$$|\psi_1\rangle = \frac{1}{\sqrt{2}}|\uparrow_1\rangle|\downarrow_2\rangle \quad (\text{as } \ \langle\uparrow_1|\langle\downarrow_2|\downarrow_1\rangle|\uparrow_2\rangle = 0) \ , \tag{19.78}$$

$$|\psi_2\rangle = P_2\frac{1}{\sqrt{2}}(|\uparrow_1\rangle|\downarrow_2\rangle + |\downarrow_1\rangle|\uparrow_2\rangle) \tag{19.79}$$

$$= |\downarrow_1\rangle|\uparrow_2\rangle\langle\downarrow_1\ \langle\uparrow_2|\frac{1}{\sqrt{2}}(|\uparrow_1\rangle|\downarrow_2\rangle + |\downarrow_1\rangle|\uparrow_2\rangle) \tag{19.80}$$

$$|\psi_2\rangle = \frac{1}{\sqrt{2}}|\downarrow_1\rangle|\uparrow_2\rangle \quad (\text{as } \ \langle\downarrow_1|\langle\uparrow_2|\uparrow_1\rangle|\downarrow_2\rangle = 0) \ . \tag{19.81}$$

Thus, as shown in Fig.19.10, we have obtained the collapse of the wavefunction from the triplet zero state to the individual components, but at a price. Note that the norm of the wavefunction is affected during the projection process. This can be seen by noting that

$$\langle\psi_1|\psi_1\rangle = \langle\psi_2|\psi_2\rangle = \frac{1}{2} < 1 \ , \tag{19.82}$$

even though the triplet zero state we started from was normalised. This indicates that the application of the projection is a non-unitary process, i.e., this transformation process does not preserve the norm of the state, and cannot be described by any unitary evolution of the original state we started with (see discussions of eqs.(5.7) - (5.9) in Chapter 5).

Once the projection has been carried out using either P_1 or P_2, then it can be shown that

$$\begin{aligned} P_1|\psi_1\rangle &= |\psi_1\rangle = P_1^2|\psi_1\rangle \ , \\ P_1|\psi_2\rangle &= 0 = P_2|\psi_1\rangle \ , \\ P_2|\psi_2\rangle &= |\psi_2\rangle = P_2^2|\psi_2\rangle \ . \end{aligned} \tag{19.83}$$

These relations indicate that the operators P_1 and P_2 keep us within the subspace of states $|\psi_1\rangle$ and $|\psi_2\rangle$, and with eigenvalues 1 and 0.

Further, it is easily seen that the addition of probabilities obtained from the projections P_1 and P_2 add up to 1. This is simply because we can define a "total" projection operator $P = P_1 + P_2$ such that

$$\begin{aligned}
P\frac{1}{\sqrt{2}}(|\uparrow_1\rangle|\downarrow_2\rangle \pm |\downarrow_1\rangle|\uparrow_2\rangle) &= (P_1+P_2)\frac{1}{\sqrt{2}}(|\uparrow_1\rangle|\downarrow_2\rangle \pm |\downarrow_1\rangle|\uparrow_2\rangle) \\
&= (|\uparrow_1\rangle|\downarrow_2\rangle\langle\uparrow_1|\langle\downarrow_2| + |\downarrow_1\rangle|\uparrow_2\rangle\langle\downarrow_1|\langle\uparrow_2|)\frac{1}{\sqrt{2}}(|\uparrow_1\rangle|\downarrow_2\rangle \pm |\downarrow_1\rangle|\uparrow_2\rangle) \\
&= \frac{1}{\sqrt{2}}(|\uparrow_1\rangle|\downarrow_2\rangle \pm |\downarrow_1\rangle|\uparrow_2\rangle)\ . \qquad (19.84)
\end{aligned}$$

Thus, we see that the eigenvalue of P acting on the triplet zero and singlet states is 1, i.e., it preserves the state and its norm. On the other hand,

$$\begin{aligned}
P|\uparrow_1\uparrow_2\rangle &= (P_1+P_2)|\uparrow_1\uparrow_2\rangle \\
&= (|\uparrow_1\rangle|\downarrow_2\rangle\langle\uparrow_1|\langle\downarrow_2| + |\downarrow_1\rangle|\uparrow_2\rangle\langle\downarrow_1|\langle\uparrow_2|)|\uparrow_1\uparrow_2\rangle \\
&= 0 \ \ (\text{as } \langle\uparrow_1|\langle\downarrow_2|\uparrow_1\uparrow_2\rangle = 0 = \langle\downarrow_1|\langle\uparrow_2|\uparrow_1\uparrow_2\rangle) \qquad (19.85) \\
&= P|\downarrow_1\downarrow_2\rangle\ . \qquad (19.86)
\end{aligned}$$

Thus, for the states $|\uparrow_1\uparrow_2\rangle$ and $|\downarrow_1\downarrow_2\rangle$ (i.e., the states that are orthogonal to the states from which P is constructed), the eigenvalue of P is 0. In this way, we may think of P as acting on the complete Hilbert space of $|\uparrow_1\uparrow_2\rangle$ and $|\downarrow_1\downarrow_2\rangle$, $\frac{1}{\sqrt{2}}(|\uparrow_1\rangle|\downarrow_2\rangle + |\downarrow_1\rangle|\uparrow_2\rangle)$ and $\frac{1}{\sqrt{2}}(|\uparrow_1\rangle|\downarrow_2\rangle - |\downarrow_1\rangle|\uparrow_2\rangle)$, so as to project onto the subspace of the states $\frac{1}{\sqrt{2}}(|\uparrow_1\rangle|\downarrow_2\rangle + |\downarrow_1\rangle|\uparrow_2\rangle)$ and $\frac{1}{\sqrt{2}}(|\uparrow_1\rangle|\downarrow_2\rangle - |\downarrow_1\rangle|\uparrow_2\rangle)$.

Very generally, it can be shown that the eigenvalues of a total projection operator such as P can only be 1 and 0: the former for all the sub-basis of (say N) orthogonal states $(\{|i\rangle\}\,,\ i \in 1,\ldots,N)$ from which it is composed, and the latter for all other states orthogonal to that sub-basis. The eignevalue of 1 is simply telling us that, upon using P, the total probability to find any one of the orthogonal states within the sub-basis in 1, and the probability to find any other state lying outside the sub-basis 0. P_i, where $(i \in (1,\ldots,N)$ are the sub-parts of $P = \sum_{i=1}^{N} P_i$, give us the individual states lying within $\{|i\rangle\}$.

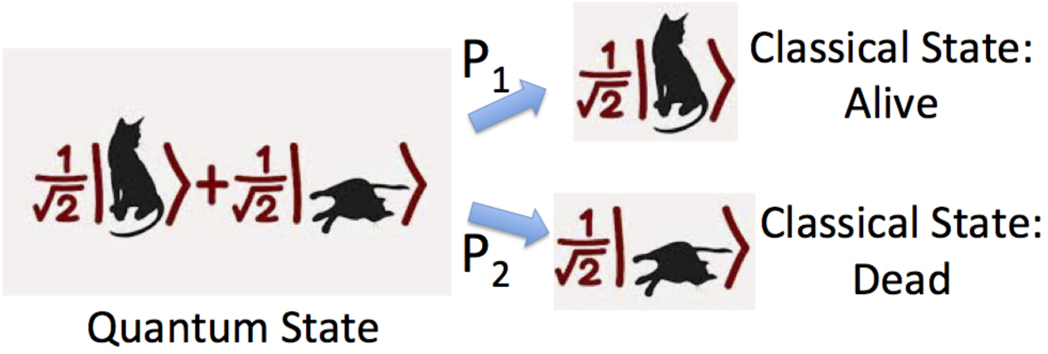


Figure 19.10: The collapse of the quantum Wavefunction of the Cat+Atom system to the classical states of the Cat via projection operators P_1 and P_2. Source: the internet.

Thus, the classical states of the Cat obtained by projection correspond to a non-unitary process applied to the original quantum state of the Cat + Atom system. Certainly, all entanglement encoded between the states of the Cat and the Atom in the original quantum state has been lost during the projection operation.

Note that had we started with a direct product state $|\psi\rangle = |\psi\rangle_1 \otimes |\psi\rangle_2$, then using the projection operators $P_1 = |\psi\rangle_1 \langle\psi|_1$ and $P_2 = |\psi\rangle_2 \langle\psi|_2$ leads to

$$\begin{aligned} |\Psi\rangle_1 &= P_1 |\psi\rangle \\ &= |\psi\rangle_1 \langle\psi|_1 |\psi\rangle_1 \otimes |\psi\rangle_2 \\ &= |\psi\rangle_1 \otimes |\psi\rangle_2 \ , \qquad (19.87) \\ |\Psi\rangle_2 &= P_2 |\psi\rangle \\ &= |\psi\rangle_2 \langle\psi|_2 |\psi\rangle_2 \otimes |\psi\rangle_1 \\ &= |\psi\rangle_1 \otimes |\psi\rangle_2 \ . \qquad (19.88) \end{aligned}$$

Thus, we see that $|\Psi\rangle_1 = |\psi\rangle = |\Psi\rangle_2$, i.e., the action of the projection operators P_1 and P_2 does not change the original state in any way! Thus, projections on Hilbert spaces that are already independent does not lead to the generation of loss of unitarity.

It is worth noting that the switch we developed towards forcing the quantum states of the Schrödinger cat gedanken into classical states is just a *thought device* for illuminating a more general concept dealing with measurements that we make on quantum systems. In the original Schrödinger's cat gedanken, the measurement was made by opening the lid of the box and noting the state of the cat by an observer; here, we have replaced this process by the activation of a switch. If we consider that there are many copies of the gedanken, in which each copy has its own switch and each switch has an equal probability to take either case (i) or case (ii), we will again obtain a probability distribution for the cat being alive or dead being 1/2 and 1/2 for dead and alive. In this way, the discussion of the outcomes obtained from the gedanken with the switch is completely equivalent to that where we need to open the lid and measure the state of the cat.

We believe, in all generality, that the process of making measurements on quantum systems leads to the collapse of the quantum wavefunction (typically, a linear superposition of many classical states/outcomes) onto a particular classical state/outcome. Further, we believe that such measurements, and the wavefunction collapses they enforce via projection operations, are always non-unitary in nature. Thus, encoding them via the unitary evolution in time of a Hamiltonian may not, in general, be possible. Further, such a non-unitary projection is *irreversible* (i.e., the measurement cannot be reversed by doing something to the classical states, and the original quantum state reversed). This is, thus, one way of visualising how the classical world "emerges" from the quantum world.

19.9.1 Measurement, Symmetry breaking and non-unitarity

A final word on wavefunction collapse and symmetries. Recall that we have shown earlier (see eq.(6.11)) that all symmetries of a Hamiltonian (and therefore its eigenstates) are preserved in time under unitary time evolution (i.e., using the operator $U = e^{-\frac{i\hat{H}t}{\hbar}}$). Thus, we should not expect that symmetries will necessarily be preserved under the application of projection operators that are non-unitary, i.e., such operations may violate the symmetries of the system.

In order to see how this happens, let us consider a variation of an earlier problem we have studied. Take the Heisenberg problem of two interacting spin 1/2s in a staggered magnetic field

$$H = \lambda \vec{S_1} \cdot \vec{S_2} + \tilde{B}\left(S_{1z} - S_{2z}\right), \tag{19.89}$$

where we have assumed that $\lambda > 0$ (antiferromagnetic spin exchange coupling) and the staggered magnetic field ($\tilde{B}$) points along the z-direction (but its sign can be either positive or negative). The magnetic term corresponds to a Zeeman coupling between the total staggered magnetisation of the system (($S_{1z} - S_{2z}$)) and the external $\tilde{B}$ staggered field. This Hamiltonian can again be diagonalised easily. Using $S_\pm = S_x \pm iS_y$, the Hamiltonian can be written as

$$\vec{S_1} \cdot \vec{S_2} = \sum_{i=\{x,y,z\}} S_{1i}S_{2i} = S_{1z}S_{2z} + \frac{1}{2}\left(S_{1+}S_{2-} + S_{1-}S_{2+}\right) \tag{19.90}$$

We can now write the matrix elements in the basis of S_{1z}, S_{2z}.

$$H = \begin{array}{c} \begin{array}{cccc} |\uparrow_1,\uparrow_2\rangle & |\uparrow_1,\downarrow_2\rangle & |\downarrow_1,\uparrow_2\rangle & |\downarrow_1,\downarrow_2\rangle \end{array} \\ \begin{pmatrix} \frac{1}{4}\lambda & 0 & 0 & 0 \\ 0 & -\frac{1}{4}+\tilde{B}\lambda & \frac{1}{2}\lambda & 0 \\ 0 & \frac{1}{2}\lambda & -\frac{1}{4}+\tilde{B}\lambda & 0 \\ 0 & 0 & 0 & \frac{1}{4}\lambda \end{pmatrix} \end{array} \tag{19.91}$$

The Hamiltonian has three decoupled blocks; the top left and bottom right form two diagonal blocks, while the middle 2×2 matrix forms the third block which is internally off-diagonal. Since the top left and bottom right blocks are diagonal, we can immediately write down two of the four eigenstates and eigenvalues as:

$$\begin{aligned} H\left|\uparrow_1,\uparrow_2\right\rangle &= \frac{1}{4}\lambda\left|\uparrow_1,\uparrow_2\right\rangle \\ H\left|\downarrow_1,\downarrow_2\right\rangle &= \frac{1}{4}\lambda\left|\downarrow_1,\downarrow_2\right\rangle \end{aligned} \tag{19.92}$$

The middle block can be written as

$$\begin{aligned} \tilde{H} &= \begin{pmatrix} \langle\uparrow_1\downarrow_2|H|\uparrow_1\downarrow_2\rangle & \langle\uparrow_1\downarrow_2|H|\downarrow_1\uparrow_2\rangle \\ \langle\downarrow_1\uparrow_2|H|\uparrow_1\downarrow_2\rangle & \langle\downarrow_1\uparrow_2|H|\downarrow_1\uparrow_2\rangle \end{pmatrix} \\ &= \frac{\lambda\hbar^2}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} - \frac{\lambda\hbar^2}{4}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \tilde{B}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \\ &= \frac{\lambda\hbar^2}{2}\sigma_x - \frac{\lambda\hbar^2}{4}\mathcal{I} + \tilde{B}\sigma_z\ . \end{aligned} \tag{19.93}$$

The diagonalisation of $\tilde{H}$ gives us

$$\begin{aligned} E_+ &= -\frac{\lambda\hbar^2}{4} + \sqrt{\tilde{B}^2 + \frac{\lambda^2\hbar^4}{4}} \quad , \quad |\psi\rangle_+ = \mathcal{N}_+(\frac{\lambda\hbar^2}{2}|\uparrow_1\downarrow_2\rangle + (-\frac{\lambda\hbar^2}{4} + \tilde{B} - E_+)|\downarrow_1\uparrow_2\rangle) \\ E_- &= -\frac{\lambda\hbar^2}{4} - \sqrt{\tilde{B}^2 + \frac{\lambda^2\hbar^4}{4}} \quad , \quad |\psi\rangle_- = \mathcal{N}_-((-\frac{\lambda\hbar^2}{4} + \tilde{B} - E_-)|\uparrow_1\downarrow_2\rangle + \frac{\lambda\hbar^2}{2}|\downarrow_1\uparrow_2\rangle) \end{aligned} \tag{19.94}$$

where $N_\pm = \left[(\frac{\lambda\hbar^2}{2})^2 + (-\frac{\lambda\hbar^2}{4} + \tilde{B} - E_\pm)^2\right]^{-1/2}$. This can be summarised in tabular form as follows

$E(\hbar = 1)$	$\|\Psi\rangle$
$E_+ = -\frac{\lambda\hbar^2}{4} + \sqrt{\tilde{B}^2 + \frac{\lambda^2\hbar^4}{4}}$	$\|\psi\rangle_+ = \mathcal{N}_+(\frac{\lambda\hbar^2}{2}\|\uparrow_1\downarrow_2\rangle + (-\frac{\lambda\hbar^2}{4} + \tilde{B} - E_+)\|\downarrow_1\uparrow_2\rangle)$
$E_- = -\frac{\lambda\hbar^2}{4} - \sqrt{\tilde{B}^2 + \frac{\lambda^2\hbar^4}{4}}$	$\|\psi\rangle_- = \mathcal{N}_-((-\frac{\lambda\hbar^2}{4} + \tilde{B} - E_-)\|\uparrow_1\downarrow_2\rangle + \frac{\lambda\hbar^2}{2}\|\downarrow_1\uparrow_2\rangle)$
$\frac{1}{4}\lambda$	$\|\uparrow_1, \uparrow_2\rangle$
$\frac{1}{4}\lambda$	$\|\downarrow_1, \downarrow_2\rangle$

Table 19.3: Spectrum of Heisenberg Hamiltonian in presence of a staggered field $\tilde{B}$

$\boxed{\text{Check:}}$ that for the case of $\tilde{B} = 0$, we get back the eigenspectrum of the simple two interacting spin-1/2 problem we had studied earlier.

Now, it can be seen that for any $|\tilde{B}| > 0$, the spin rotational symmetry of the singlet and triplet zero states $\frac{1}{\sqrt{2}}(|\uparrow_1\downarrow_2\rangle \mp |\downarrow_1\uparrow_2\rangle)$ (eigenstates of $S_z = S_{1z} + S_{2z}$ with eigenvalue 0) is **manifestly broken**. This is seen simply from the fact that the weight factors in front of the $|\uparrow_1\downarrow_2\rangle$ and $|\downarrow_1\uparrow_2\rangle$ components of $|\psi\rangle_\pm$ are not equal any more, meaning that these two states are no longer eigenstates of $S_z = S_{1z} + S_{2z}$. In fact, as we tune $|\tilde{B}|$, either $|\uparrow_1\downarrow_2\rangle$ or $|\downarrow_1\uparrow_2\rangle$ are increasingly weighted over another.

The connection with measurement via (non-unitary) projection operators is observed by taking the limit of $|\tilde{B}| \to \infty$. For $\tilde{B} \to \infty$, the state $|\downarrow_1\uparrow_2\rangle$ becomes the eigenstate, while for $\tilde{B} \to -\infty$, $|\uparrow_1\downarrow_2\rangle$ is the eigenstate. We can thus mimic the projection mechanism (i.e., the application of the projection operators $P_1 = |\uparrow_1\rangle|\downarrow_2\rangle\langle\uparrow_1|\langle\downarrow_2|$ and $P_2 = |\downarrow_1\rangle|\uparrow_2\rangle\langle\downarrow_1|\langle\uparrow_2|$ by taking the limit of $|\tilde{B}| \to \infty$. The non-unitary nature of the projection mechanism lies in the observation that there exists no unitary transformation can map the eigenstates $|\psi\rangle_\pm$ (observed for $|\tilde{B}| \neq 0$) onto the eigenstates at $\tilde{B} = 0$ ($\frac{1}{\sqrt{2}}(|\uparrow_1\downarrow_2\rangle \mp |\downarrow_1\uparrow_2\rangle)$). Another way of saying this is: there is no way connect the symmetry broken Hamiltonian (i.e., with $|\tilde{B}| \neq 0$) to the symmetry preserved Hamiltonian at $\tilde{B} = 0$ via unitary transformations. The "classical" states $|\downarrow_1\uparrow_2\rangle$ and $|\uparrow_1\downarrow_2\rangle$ are achieved upon taking the limit of $|\tilde{B}| \to \infty$; the journey towards the limit is the "emergence of the classical world" from the world of quantum mechanics.

19.10 The Qubit and quantum information processing

A final quick word on the basic building block of quantum information science: the qubit (or quantum mechanical bit). This is actually nothing but the two-level (or spin-1/2) system we have been studying all along! Recall that the most general state for the spin-1/2 system we can write down is

$$|\psi\rangle = a|\uparrow\rangle + b|\downarrow\rangle \ , \tag{19.95}$$

where $a, b \in \mathcal{C}$ and $|a|^2 + |b|^2 = 1$. Indeed, if we think of $|\uparrow\rangle \equiv |1\rangle$ and $|\downarrow\rangle \equiv |0\rangle$ (where $|0\rangle$ and $|1\rangle$ are the states of a bit, the elementary unit of classical information), we can see that $|\psi\rangle$ is already a generalisation of the classical bit. A vectorial representation can be given in terms of the Bloch sphere (see Fig.19.11). This is done by taking $a = \cos(\theta/2)$, $b = \sin(\theta/2)e^{i\phi}$ (in terms of the polar angles (θ, ϕ) made by the vector $|\psi\rangle$ anywhere on the surface of a sphere of unit radius), such that

$$|\psi\rangle = \cos(\theta/2)|\uparrow\rangle + \sin(\theta/2)e^{i\phi}|\downarrow\rangle \ . \tag{19.96}$$

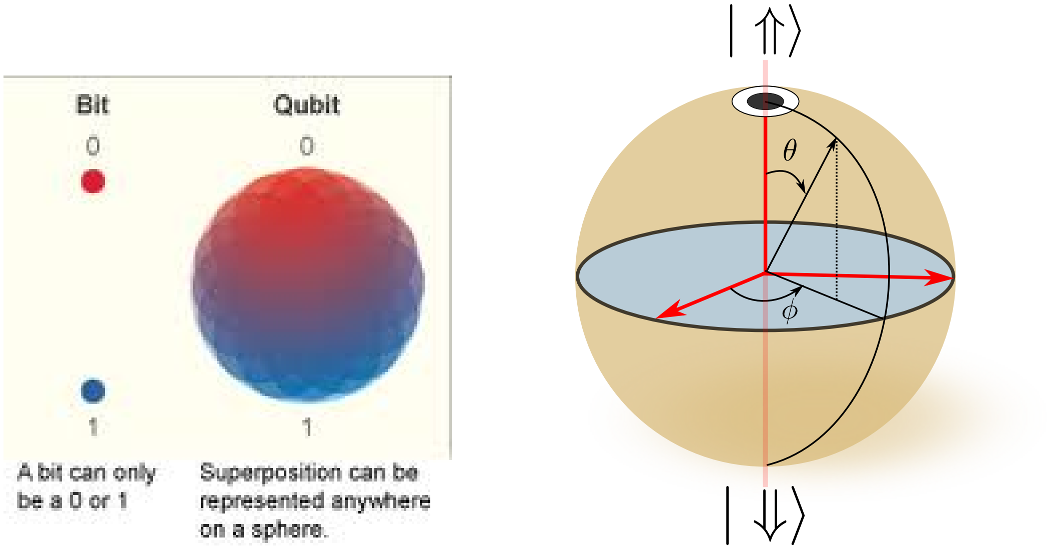


Figure 19.11: The qubit represented on the Bloch sphere. The north and south poles represent the $|1\rangle \equiv |\uparrow\rangle$ and $|0\rangle \equiv |\downarrow\rangle$ elements of the classical bit. The vector $|\psi\rangle$ has length $\sqrt{3/4}\hbar$, and lies on the surface of a sphere, and defined by the polar angles $(0 \leq \theta \leq \pi \ , \ 0 \leq \phi \leq 2\pi)$. Source: the internet.

A quantum computer (or more appropriately, a quantum information processor) is made out of a large number of such two-state systems. For a system of N qubits, the Hilbert space is comprised of 2^N vectors; note that such a Hilbert space grows exponentially with N. Very generally, the state of such a quantum computer can be written as

$$\begin{aligned} |\Psi_{Nqubits}\rangle &= \otimes_{j=1}^{N}(a_j|\uparrow_j\rangle + b_j|\downarrow_j\rangle) \\ &= (a_1 a_2 \dots a_N)\,|\uparrow\rangle_1 \otimes |\uparrow\rangle_2 \dots |\uparrow\rangle_N \\ &\quad + (a_1 a_2 \dots a_{N-1} b_N)\,|\uparrow\rangle_1 \otimes |\uparrow\rangle_2 \dots |\uparrow\rangle_{N-1}\,|\downarrow\rangle_N + \dots \ , \end{aligned} \tag{19.97}$$

where $\{a_j, b_j\} \in \mathcal{C}$, and the $\otimes$ symbol denotes a direct product of the N linear combinations of single-qubit superposition states. This product form demonstrates the power of the parallelism involved in the way the quantum computer works: as can seen from the second line above, any algorithm working on such a quantum processor will running on a huge number of initial starting "classical" combinations of $|\uparrow\rangle$s and $|\downarrow\rangle$ bits (e.g., $|\uparrow\rangle_1 \otimes |\uparrow\rangle_2 \ldots |\uparrow\rangle_{N-1} |\downarrow\rangle_N$ etc.) given by the set of coefficients $\{a_j, b_j\}$ (e.g., $a_1 a_2 \ldots a_{N-1} b_N$ etc.). Thus, a huge number of "classical" outcomes are generated during the operation of the algorithm, and the final measurement gives one of them as the output to the task.

This *oracular* (or "know it all") nature of a quantum computer is what gives it a likely advantage over a classical computer at certain tasks, e.g., factorisation of a huge number its into prime factors. Indeed, this particular example is known to be a very difficult task for classical computers (i.e., it takes a huge amount of time to get the prime factors), and the reason why the RSA algorithm for encryption of data passed over the internet works so well. Shor's algorithm for prime factorisation on a quantum computer offers a massive advantage, and its implementation on large scale quantum computers could well change our outlook on information technology, encryption etc.

Chapter 20

Quantum Particle On a Ring (QPOR) Problem

20.1 The importance of boundary conditions

At several points in these lectures, we have encountered the important role played by boundary conditions in determining the nature of the eigenvectors and eigenvalues for a given problem in quantum mechanics. Recall, for instance, the particle in a box problem. There, the imposition of the boundary conditions that the wavefunction must vanish at the box walls (i.e., *clamped* boundary conditions), say $\psi(x=0) = 0 = \psi(x=a)$, dictated the fact that we had standing wave type solutions analogous to the harmonics of a pipe. Further, recall that in the simple harmonic oscillator problem, even though the potential is overall confining, the fact that we had open boundary conditions meant that $\psi(x \to \pm\infty) \to 0$. This immediately led to the fact that the quantum wavefunction penetrated into the (classically forbidden) barrier regions. Also, recall that in the scattering problems, the absence of any confining potentials meant that we were working with travelling wave solutions $\psi \propto e^{\pm ikx}$. For open boundary conditions, this meant that the normalisation of such solutions would have to be considered carefully, but we are always going to have to work with travelling waves. Certainly, it becomes clear that the boundary conditions determine the nature of the eigenspectrum.

In carrying this idea further forward, in this chapter we are going to study the problem of a quantum particle whose motion is confined to a circle. Here, we will encounter another type of boundary condition: the periodic boundary condition. Actually, we already saw an example of this when we encountered the $\Phi(\phi)$ angular part of the wavefunction for the H-atom problem. Now, we will see how the entire eigenspectrum of the problem can be sensitive to small changes in boundary conditions, i.e., the phenomenon of *spectral flow*. But before doing so, we are going to show another way in which boundary conditions can have a crucial role. Consider the linear momentum operator, $\hat{p}_x = -i\hbar\partial/\partial x$. We believe that this operator is Hermitian, and it should be, as it corresponds to an experimentally observable quantity. Let us then verify that this operator is indeed Hermitian. For this, we check

$$\begin{aligned}
\int_{x_1}^{x_2} dx\psi_2^*(x)\hat{p}_x\psi_1(x) &= \int_{x_1}^{x_2} dx\psi_2^*(x) - i\hbar\frac{\partial}{\partial x}\psi_1(x)\ , && (20.1)\\
&= -i\hbar\psi_2^*(x)\psi_1(x)|_{x_1}^{x_2} + i\hbar\int_{x_1}^{x_2} dx\frac{\partial\psi_2^*(x)}{\partial x}\psi_1(x)\ ,\ \text{(int. by parts)} && (20.2)\\
&= i\hbar\int_{x_1}^{x_2} dx\frac{\partial\psi_2^*(x)}{\partial x}\psi_1(x)\ , && (20.3)\\
&= \Big\{\int_{x_1}^{x_2} dx\psi_1^*(x)\Big[-i\hbar\frac{\partial\psi_2(x)}{\partial x}\Big]\Big\}^*\ , && (20.4)
\end{aligned}$$

where we have assumed that

$$-i\hbar\psi_2^*(x)\psi_1(x)|_{x_1}^{x_2} = 0\ , \tag{20.5}$$

in order to establish the Hermiticity of the linear momentum operator $\hat{p}_x = -i\hbar\partial/\partial x$. Thus, $\hat{p}_x$ is Hermitian if its eigenvector space consists of functions for whom the surface term eq.(20.5) vanishes. Certainly, wavefuctions arising from clamped boundary conditions satisfy this criteria, as do those wavefunctions $\psi(x \to \pm\infty) \to 0$ arising from open boundary conditions (e.g., in overall confining potentials). For scattering state wavefunctions of the kind $\psi \propto e^{\pm ikx}$, the criteria can only be met if

$$e^{\pm ik_1x} \times e^{-ik_2x}|_{-\infty}^{\infty} = 0\ . \tag{20.6}$$

While this is clear for the case of $k_1 = k_2$, the product oscillates for $k_1 \neq k_2$ and so it is unclear what happens as $x \to \pm\infty$. However, there is a way to obtain a definite (and satisfactory) answer to this conundrum. It is called "regularisation", and depends on a prescription which says that for such functions, the limit of $|x| \to \infty$ is defined to be the average over a large interval:

$$\lim_{x\to\infty} e^{\pm ik_1x} \times e^{-ik_2x} = \lim_{L\to\infty,\Delta\to\infty}\frac{1}{\Delta}\int_L^{L+\Delta} dx\ e^{i(k_1-k_2)x} = 0\ \ ,\ \ \text{if}\ \ k_1 \neq k_2\ . \tag{20.7}$$

This ensures that the linear momentum operator $\hat{p}_x$ is again Hermitian on this space of wavefunctions.

Another important case is that of the wavefunctions arising from periodic boundary conditions, for which the boundary terms in eq.(20.5) vanish by construction. This is what we will now study.

20.2 Hamiltonian, Eigenstates and Eigenvalues

Imagine a particle of mass M constrained to move on the circumference of a circle of radius R. The motion can be described completely by the angular coordinate ϕ on the ring, where $\phi \in [0, 2\pi]$. The periodic boundary conditions on the eigenstates of this problem $\Psi(\phi)$ dictate that

$$\Psi(\phi) = \Psi(\phi + 2\pi) \tag{20.8}$$

That is, $\Psi_m(\phi)$ is required to be single-valued on the ring. Further, the angular momentum conjugate to ϕ is

$$\hat{p}_\phi = -i\hbar \frac{\mathrm{d}}{\mathrm{d}\phi} \tag{20.9}$$

such that $[\phi, p_\phi] = i\hbar$. Note that the angular momentum operator $\hat{p}_\phi$ is clearly Hermitian, as the surface term discussed above naturally vanishes in this problem

$$-i\hbar\Psi_2^*\Psi_1|_0^{2\pi} = -i\hbar[\Psi_2^*(2\pi)\Psi_1(2\pi) - \Psi_2^*(0)\Psi_1(0)] = 0 \quad (\text{as} \quad \Psi(0) = \Psi(2\pi)) \ . \tag{20.10}$$

The Hamiltonian for this problem is simply

$$\mathcal{H} = \underbrace{-\frac{\hbar^2}{2MR^2}}_{\text{moment of inertia}} \frac{\mathrm{d}^2}{\mathrm{d}\phi^2} \tag{20.11}$$

The time-independent Schrödinger equation is thus

$$\mathcal{H}\Psi\left(\phi\right) = -\frac{\hbar^2}{2MR^2}\frac{\mathrm{d}^2\Psi}{\mathrm{d}\phi^2} = E\Psi\left(\phi\right) \tag{20.12}$$

The eigenvalues and eigenfunctions are

$$\begin{aligned} \Psi_m\left(\phi\right) &= \frac{1}{\sqrt{2\pi}}e^{im\phi}, \quad m = 0, \pm 1, \pm 2, ... \\ E_m &= \frac{\hbar^2 m^2}{2MR^2} \end{aligned} \tag{20.13}$$

The quantum number m is the eigenvalue of the angular momentum p_ϕ, which makes sense because the Hamiltonian is simply p_ϕ^2, and we would expect m to be a good quantum number because the Hamiltonian commutes with p_ϕ. The factor of $\sqrt{2\pi}$ is a normalization constant.

$$\int_0^{2\pi} d\phi\ |\Psi_m|^2 = \frac{1}{2\pi}\int_0^{2\pi} d\phi = 1 \tag{20.14}$$

Classically, $m > 0$ corresponds to clockwise rotation and $m < 0$ to anti-clockwise rotation. The expression for E_m is very similar to the infinitely deep 1D square well, but the boundary conditions on the eigenstates $\Psi_m\left(\phi\right)$ are different in this problem. Note that the number of nodes in the real part of $\Psi_m\left(\phi\right)$ (as well as $|\Psi_m\left(\phi\right)|^2$) increase by 2 for $\Delta m = 1$.

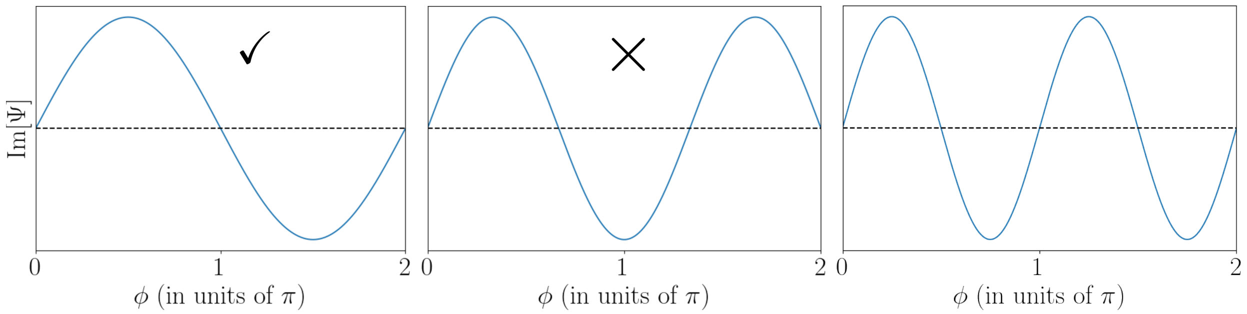
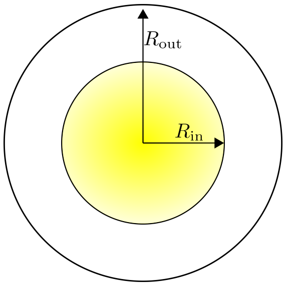


Figure 20.1: Plots of $Im\left[\Psi\right] = \sin\phi$ vs ϕ and $\frac{\mathrm{d}Im[\Psi]}{\mathrm{d}\phi}$ vs ϕ

Note also the non-trivial topology of the system; there are two ways to go form any point on the ring to any other. This is simply because the circle corresponds to a multiply-connected geometry, and a surface of genus one (as it effectively encloses a hole). The latter statement means that we cannot shrink a closed curve bigger than the size of the circle to a point, as it will stop when it hits the circle.

20.3 Effect of Flux Insertion and Spectral Flow



At this point, we introduce a solenoidal flux through the "hole" of the ring, as follows.

$$\vec{B} = \vec{\nabla} \times \vec{A}, \qquad B_z = \begin{cases} B, & r \leq R_{\text{in}} \\ 0, & r > R_{\text{in}} \end{cases} \tag{20.15}$$

This is satisfied by the vector potential

$$\begin{aligned} \vec{A} &= A_r \hat{e}_r + A_\phi \hat{e}_\phi + A_z \hat{e}_z, \\ \text{with } A_r &= A_z = 0, \text{ and} \\ A_\phi &= \begin{cases} \frac{Br}{2}, & r \leq R_{\text{in}} \\ \frac{BR_{\text{in}}^2}{2r} = \frac{\Phi}{2\pi r}, & r > R_{\text{in}} \end{cases} \end{aligned} \tag{20.16}$$

Note that while the charged particle will not feel the effects of the B-field on the ring ($R > R_{\text{in}}$), it will nevertheless feel the vector potential $\vec{A}$ (as this extends everywhere). Thus, the Hamiltonian for the charged particle with this non-trivial $\vec{A}$ is given by

$$\begin{aligned} H &= \frac{1}{2M}\left(p_\phi - \frac{e}{c}\vec{A}\right)^2 \\ &= \frac{1}{2M}\left(-i\hbar\frac{1}{R}\frac{\mathrm{d}}{\mathrm{d}\phi} - \frac{e}{c}\frac{\Phi}{2\pi R}\right)^2 \\ &= \frac{\hbar^2}{2M}\left(-i\frac{\mathrm{d}}{\mathrm{d}\phi} - \frac{\Phi}{\hbar\frac{2\pi c}{e}}\right)^2 \\ &= \frac{\hbar^2}{2M}\left(-i\frac{\mathrm{d}}{\mathrm{d}\phi} - \frac{\Phi}{\Phi_0}\right)^2 \end{aligned} \tag{20.17}$$

[Rule for minimal coupling: Peierls substitution- $\vec{p} \to \vec{p} - \frac{e}{c}\vec{A}$, seen in Maxwell's laws itself with test charges, and is a result of gauge invariance.]

($\Phi_0 = \frac{\hbar 2\pi c}{e} = \frac{hc}{e}$ is the magnetic flux quantum.)

The solutions for this Hamiltonian are as follows. The wavefunction retains the same form:

$$\Psi_m(\phi) = e^{im\phi}, \quad m = 0, \pm1, \pm2, ... \tag{20.18}$$

while the energy eigenvalues are

$$E_m(\phi) = \frac{\hbar^2}{2MR^2}\left(m - \frac{\Phi}{\Phi_0}\right)^2 \tag{20.19}$$

That is, $E_m(\Phi)$ is now a function of the ratio of the total magnetic flux Φ and the flux quantum Φ_0.

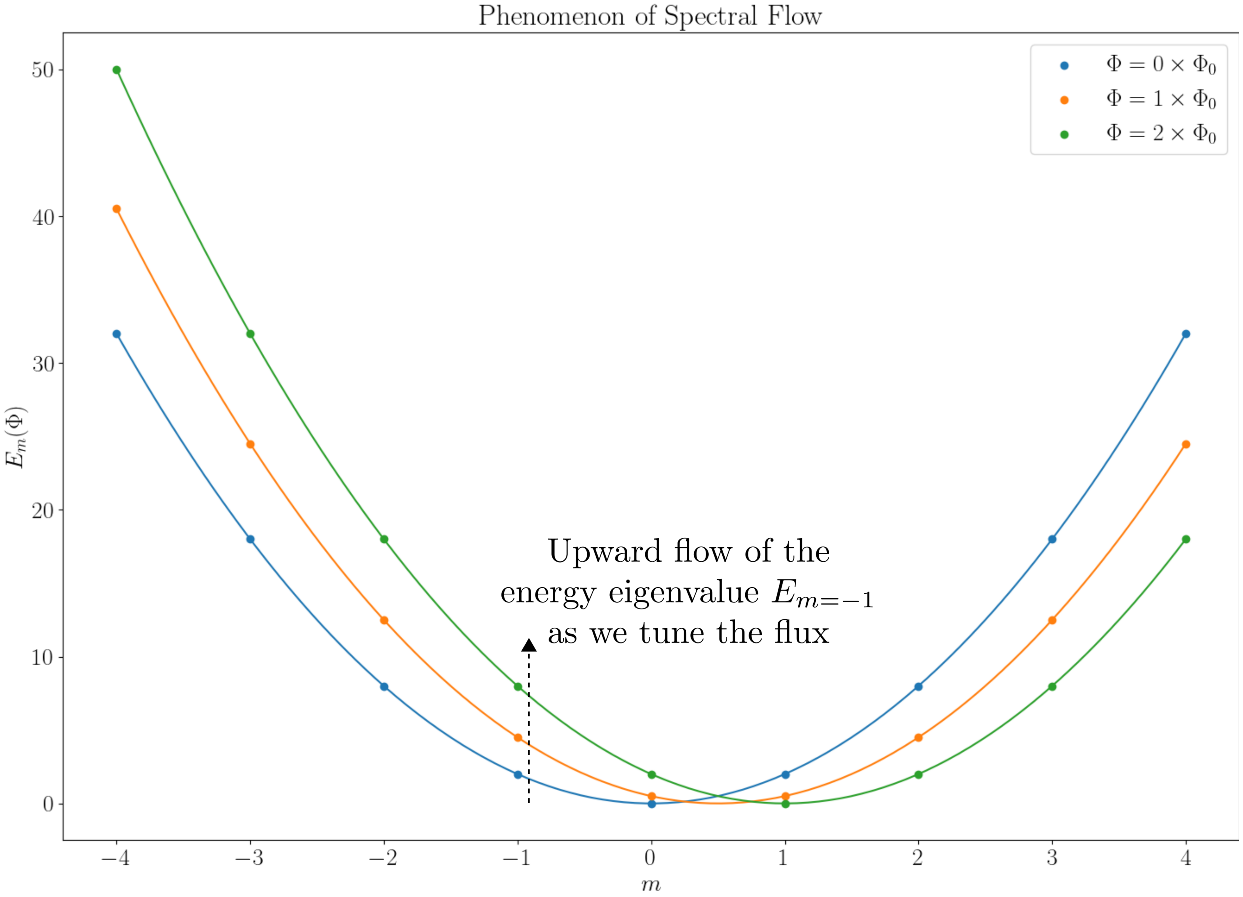


Figure 20.2: Differently colored lines represent the spectrum for different values of the flux. As we tune the flux ratio $\frac{\Phi}{\Phi_0}$, the energy eigenvalues for a particular value of m move to different values of energy. Since this occurs for all value of m, the entire spectrum appears to "flow" forward. This is called *spectral flow*.

Note the extra parity symmetry at $\Phi = \frac{n}{2}\Phi_0, n\Phi_0$ and not otherwise. Also, note the doubly-degenerate ground state at $\Phi = \frac{n}{2}\Phi_0$. To see the current flow in the ground state, we first define the current density operator as $\vec{j} = \frac{c}{e}\frac{\partial H}{\partial \vec{A}}$, where $\vec{A} = \frac{\Phi}{2\pi R}\hat{e}_\phi$. Therefore,

$$\begin{aligned}-c\frac{\partial E}{\partial \Phi} = -c\langle\Psi\left|\frac{\partial H}{\partial \Phi}\right|\Psi\rangle = -c\langle\Psi\left|\frac{\partial H}{\partial \vec{A}}\frac{\partial \vec{A}}{\partial \Phi}\right|\Psi\rangle = -\langle\Psi\left|\frac{c}{e}\frac{\partial H}{\partial \vec{A}}\frac{e}{2\pi R}\hat{e}_\phi\right|\Psi\rangle = -\frac{e}{2\pi R}\langle\psi|j_\phi|\Psi\rangle \\ = -\frac{e}{2\pi R}\int_0^{2\pi} d\phi j_\phi \equiv I\end{aligned} \tag{20.20}$$

What this shows is that the measured current is proportional to the change in energy ΔE of the quantum system to a small change $\Delta\Phi$ of the external flux (response to an electromagnetic stimulus can arise from some "stiffness" of the system). This current for the ground state, for example, is called the persistent current of the ring.

For the particle on a ring problem, the current turns out to be

$$I = -c\frac{\partial E_m}{\partial \Phi} = \frac{c\hbar^2}{\Phi_0 M R^2}\left(m - \frac{\Phi}{\Phi_0}\right) \tag{20.21}$$

For $\Delta\Phi$ in the neighborhood of $\Phi = 0$, the ground state current is $I_{m=0} = 0$.

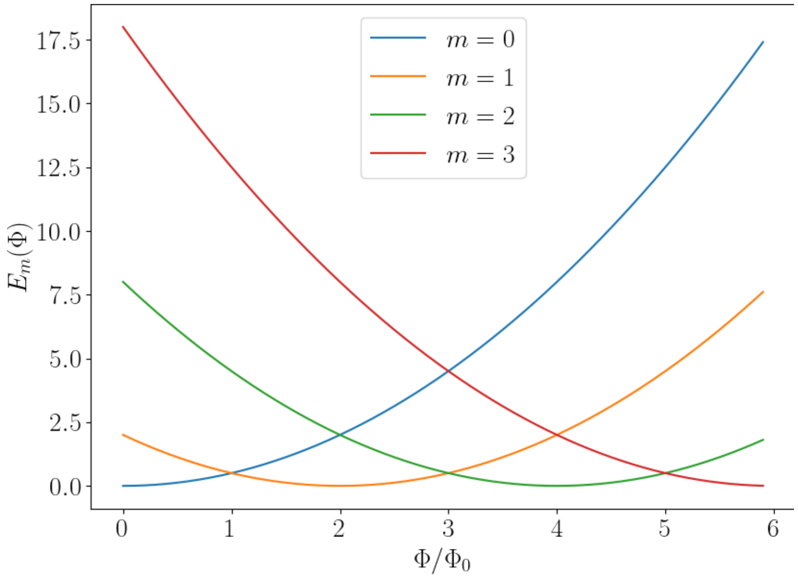


Figure 20.3: The evolution of the ground state upon tuning the external flux Φ: The ground state can have a non-zero m at finite Φ, suggesting that the ground state can have a non-zero current.

20.4 The Field-Free Picture

Quantum mechanics allows for a unitary transformation on all observables and states:

$$\begin{aligned} U &: \hat{\Omega} \to \hat{\Omega}' = U\hat{\Omega}U^{-1} \\ U &: |\Psi\rangle \to |\Psi\rangle' = U|\Psi\rangle \end{aligned} \tag{20.22}$$

The last definition implies $\langle\Phi| \to \langle\Phi| U^{-1}$. The transformation is such that matrix elements (and hence expectation values) are left unchanged.

$$\langle\Phi'|\hat{\Omega}'|\Psi'\rangle = \langle\Phi| U^{-1}U\hat{\Omega}U^{-1}U|\Psi\rangle = \langle\Phi|\hat{\Omega}|\Psi\rangle \tag{20.23}$$

With this in mind, we can construct a unitary transformation to remove (”gauge out”) the $\vec{A}$ from the Hamiltonian:

$$U(\phi) = \exp\left(-\frac{ie}{c\hbar}\int_0^\Phi \vec{A}\cdot d\vec{l}\right) = \exp\left(-i\frac{\Phi}{\Phi_0}\phi\right) \tag{20.24}$$

such that the Hamiltonian transforms as

$$H \to H' = UHU^{-1} = \exp\left(-i\frac{\Phi}{\Phi_0}\phi\right)\frac{\hbar^2}{2MR^2}\left(-i\frac{\mathrm{d}}{\mathrm{d}\phi} - \frac{\Phi}{\Phi_0}\right)^2 \exp\left(i\frac{\Phi}{\Phi_0}\phi\right) \tag{20.25}$$

Because

$$\exp\left(-i\frac{\Phi}{\Phi_0}\phi\right)\frac{1}{\hbar}p_\phi\exp\left(i\frac{\Phi}{\Phi_0}\phi\right)=\frac{1}{\hbar}p_\phi+\frac{\Phi}{\Phi_0} \tag{20.26}$$

, the transformed Hamiltonian becomes

$$H'=\frac{\hbar^2}{2MR^2}\left(-i\frac{\mathrm{d}}{\mathrm{d}\phi}+\frac{\Phi}{\Phi_0}-\frac{\Phi}{\Phi_0}\right)=-i\frac{\hbar^2}{2MR^2}\frac{\mathrm{d}}{\mathrm{d}\phi} \tag{20.27}$$

However, recall that a unitary transformation cannot change the energy eigenvalues:

$$H'\left|\Psi'\right\rangle=UHU^{-1}U\left|\Psi\right\rangle=UH\left|\Psi\right\rangle=UE\left|\Psi\right\rangle=E\left|\Psi'\right\rangle \tag{20.28}$$

What this means, is that even though we have removed the vector potential from the Hamiltonian, it still exists in the energy eigenvalues. This is because, while the transformation has removed $\vec{A}$ from the Hamiltonian, it has also changed the boundary conditions of the problem. The original boundary condition was

$$\Psi(\phi+2\pi)=\Psi(\phi) \tag{20.29}$$

This has now become

$$\begin{aligned}\Psi'\left(\phi+2\phi\right)=U(\phi+2\pi)\Psi(\phi+2\pi)=U(\phi+2\pi)\Psi(\phi)&=\exp\left(-2\pi i\frac{\Phi}{\Phi_0}\right)U(\phi)\Psi(\phi)\\&=\exp\left(-2\pi i\frac{\Phi}{\Phi_0}\right)\Psi'(\phi)\end{aligned} \tag{20.30}$$

We call these the *twisted boundary conditions*. Consider the Möbius strip shown in Fig.20.4 below for how twisted boundary conditions affect the states Ψ.

Figure 20.4: The Möbius strip. Source: the internet.

Thus, the $\Psi'(\phi)$ consistent with these boundary conditions are

$$\Psi'_m(\phi)=\exp\left[i\left(m-\frac{\Phi}{\Phi_0}\right)\phi\right],\quad m\in\mathbb{Z} \tag{20.31}$$

because,

$$\Psi'\left(\phi+2\pi\right)=\exp\left[i\left(m-\frac{\Phi}{\Phi_0}\right)\phi+2\pi im-2\pi i\frac{\Phi}{\Phi_0}\right]=\exp\left[-2\pi i\frac{\Phi}{\Phi_0}\right]\Psi'(\phi) \tag{20.32}$$

Clearly,

$$-\frac{\hbar^2}{2MR^2}\frac{\mathrm{d}^2}{\mathrm{d}\phi^2}\Psi'_m\left(\phi\right)=\frac{\hbar^2}{2MR^2}\left(m-\frac{\Phi}{\Phi_0}\right)^2\Psi'_m\left(\phi\right) \tag{20.33}$$

Thus, while we can remove the $\vec{A}$ from the Hamiltonian, it reappears in the boundary conditions, and its (topological) effects on the energy spectrum (i.e., spectral flow) and other observables remain unaffected.

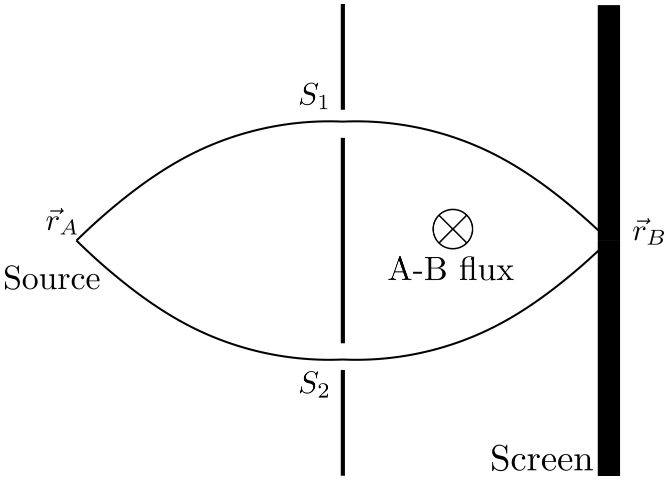
Writing $k = m - \frac{\Phi}{\Phi_0}, \Psi_k(\phi) = \exp(ik\phi)$, we see that the phase picked up in a closed path is given by

$$\exp\left(-2\pi i \frac{\Phi}{\Phi_0}\right) \tag{20.34}$$

while the phase picked up moving on the circle from ϕ_1 to ϕ_2 is

$$\Delta = -\frac{e}{c\hbar}\int_{\phi_1}^{\phi_2} \vec{A}\cdot\vec{dl} = -\frac{2\pi}{\Phi_0}\int_{\phi_1}^{\phi_2} \frac{\Phi R}{2\pi R}d\phi = -\frac{\Phi}{\Phi_0}(\phi_2 - \phi_1) \tag{20.35}$$

This extra phase is called the *Aharonov-Bohm phase.* It is an example of a geometric ("Berry") phase, and is purely dependent on the non-trivial topology. The "particle on a ring" problem is also referred to as the bound state Aharonov-Bohm problem. In the original A-B exposition, a solenoidal flux was placed in a Feynman double-slit interference experiment, with electrons in between the slit and the screen. It was predicted that the interference pattern would change on changing the A-B flux.



For an electron through slit 1 (S_1) (that is, above the solenoid), the probability amplitude is

$$|\Psi_1\rangle \to |\Psi_1'\rangle = \exp\left(-\frac{ie}{c\hbar}\int_{\vec{r}_A}^{\vec{r}_B} \vec{A}\cdot\vec{dl}\right)|\Psi_1\rangle \tag{20.36}$$

Similarly, for an electron through S_2, the probability amplitude is

$$|\Psi_2\rangle \to |\Psi_2'\rangle = \exp\left(-\frac{ie}{c\hbar}\int_{\vec{r}_A}^{\vec{r}_B} \vec{A}\cdot\vec{dl}\right)|\Psi_2\rangle \tag{20.37}$$

Thus, the total probability between the source ($\vec{r}_A$) and a point on the screen ($\vec{r}_B$) is

$$\begin{aligned}
&(\langle\Psi_2'| + \langle\Psi_1'|)\,(|\Psi_2'\rangle + |\Psi_1'\rangle) \\
&= \langle\Psi_1'|\Psi_1'\rangle + \langle\Psi_2'|\Psi_2'\rangle + \overbrace{\langle\Psi_1'|\Psi_2'\rangle + \langle\Psi_2'|\Psi_1'\rangle}^{\text{quantum interference terms}} \\
&= \langle\Psi_1|\Psi_1\rangle + \langle\Psi_2|\Psi_2\rangle + 2\mathrm{Re}\ \{\langle\Psi_2'|\Psi_1'\rangle\} \\
&= \langle\Psi_1|\Psi_1\rangle + \langle\Psi_2|\Psi_2\rangle + 2\mathrm{Re}\ \left\{\langle\Psi_2|\Psi_1\rangle \exp\left(-\frac{ie}{c\hbar}\int_{S_1}\vec{A}\cdot d\vec{l} + \frac{ie}{c\hbar}\int_{S_2}\vec{A}\cdot d\vec{l}\right)\right\} \qquad (20.38) \\
&= \langle\Psi_1|\Psi_1\rangle + \langle\Psi_2|\Psi_2\rangle + 2\mathrm{Re}\ \left\{\langle\Psi_2|\Psi_1\rangle \exp\left(-\frac{ie}{c\hbar}\oint\vec{A}\cdot d\vec{l}\right)\right\} \\
&= \langle\Psi_1|\Psi_1\rangle + \langle\Psi_2|\Psi_2\rangle + 2\mathrm{Re}\ \left\{\langle\Psi_2|\Psi_1\rangle \exp\left(-2\pi i\frac{\Phi}{\Phi_0}\right)\right\}
\end{aligned}$$

Thus, the quantum interference is now modulated by the A-B flux Φ. That is, tuning the flux should turn a dark fringe at $\Phi = 0$ to a bright fringe at $\frac{\Phi_0}{2}$, and vice-versa. Note that at $\Phi = n\Phi_0, n \in \mathbb{Z}$, the interference pattern remains invariant.

× ——— ×

The Aharonov-Bohm phase as an example of a Berry (geometric) phase Consider the bound state A-B effect/QPOR in which we get twisted boundary conditions:

$$\Psi(\phi + 2\pi) = \exp\left(-2\pi i\frac{\Phi}{\Phi_0}\right)\psi(\phi) \qquad (20.39)$$

This shows that upon completing a closed path in real phase, the particle has picked up a non-trivial phase which is directly related to the AB flux. This is a special case of a so called *geometric phase.* This was discovered by Pancharatnam in quantum optics, and later by Berry more generally in quantum physics. The topic of geometric phases has become increasingly popular in physics since its discovery, and appears to play an important role in all kinds of contexts.

Chapter 21

Particle Tunneling in a Double Well

The problem of a quantum particle placed in a double well, and tunneling between the two wells (whereas clasically, you would have expected it to sit still in any one of the two wells) is a protoypical problem in quantum mechanics. This problem appears in the quantum mechanical description of the possible conformations of some simple molecules such as ammonia (NH_3, where the Nitrogen atom can lie either above or below the plane of the three Hydrogen atoms), various simple problems in quantum optics (e.g., lasing), nuclear magnetic resonance, the physics of semiconductors (e.g., the tunnel diode) etc. Below, we set up the simplest version of the problem and learn its solution.

21.1 The symmetric double well

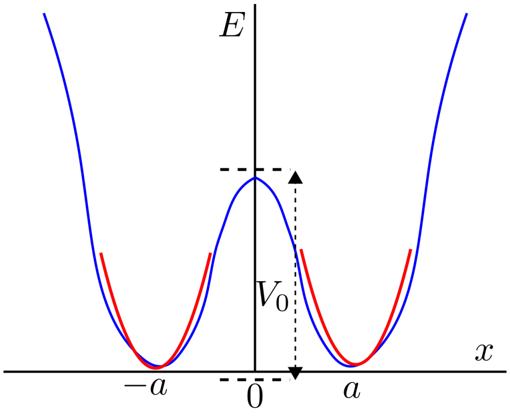


Figure 21.1: Double well potential

Consider a 1D potential of the form in fig. 21.1, with parity symmetry about the $y-$axis ($x = 0$). Deep inside either of the two wells, we can expect a particle to be bound (classically, at the bottom) with the quantum mechanical bound states that can be obtained by approximating the well by a harmonic potential. This potential is represented by a red curve in the figure. However, this is only true at the zeroth level (that is, for two isolated wells).

We assume that, for the case of symmetric wells, the eigenspectrum for the harmonic approximation will give us a set of discrete states. Further, the eigenspectra of these two wells are identical to one another, and therefore degenerate as well. We now consider only the lowest state in each of the wells (which are also degenerate with one another, at an energy given by $\hbar\omega$) and proceed by considering the fact that these two wells are really not disconnected from one another, i.e., the potential barrier between them is finite. Also, note that the symmetric nature of potential reflects on its parity (i.e., right-left symmetry).

Any quantum-mechanical tunneling between the symmetrically placed degenerate energy levels and across the potential barrier (of height V_0) will certainly lift this degeneracy into two states.

$$\begin{aligned} |\Psi_s\rangle &= \frac{1}{\sqrt{2}}\left(|\Psi(a)\rangle + |\Psi(-a)\rangle\right); E_s = \frac{\hbar\omega}{2} + \frac{\Delta}{2} \\ |\Psi_a\rangle &= \frac{1}{\sqrt{2}}\left(|\Psi(a)\rangle - |\Psi(-a)\rangle\right); E_a = \frac{\hbar\omega}{2} - \frac{\Delta}{2} \end{aligned} \tag{21.1}$$

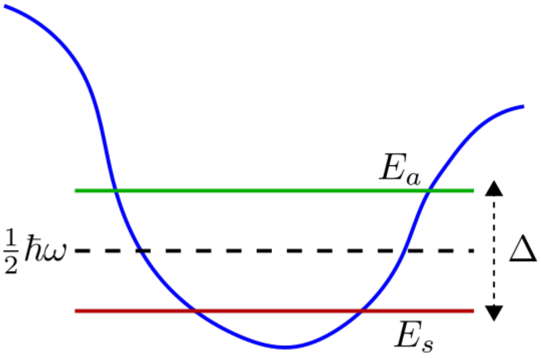


Figure 21.2: Splitting of the degenerate levels. Green-asymmetric level, red-symmetric level

The symmetric state $|\Psi_s\rangle$ has its energy lowered by $\frac{\Delta}{2}$, while the anti-symmetric state $|\Psi_a\rangle$ has its energy raised by the same amount. The ground state has the same symmetry as the potential, the parity symmetry (recall the infinitely deep 1D square well). The two states are shown in fig. 21.3.

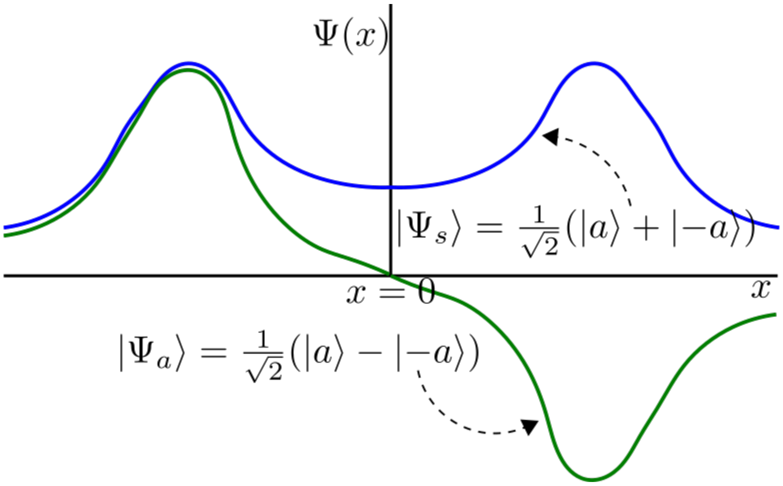


Figure 21.3: Symmetric and asymmetric wavefunctions produced under splitting

How can we estimate Δ? The quantum-mechanical tunneling amplitude between the two initially degenerate levels can be written as

$$\tilde{t} = C_1 e^{-\Gamma\frac{t}{\hbar}} \tag{21.2}$$

Γ is the decay rate of particle in either of the two levels. Thus, with a degenerate energy level $E_0 = \frac{\hbar\omega}{2}$ (which acts as the inverse lifetime of the state), we can write down the constant in the tunneling amplitude as $C_1 \in \mathbb{C}$. Written in the basis of these lowest states in the two wells, the form of the effective Hamiltonian matrix which acts on these degenerate levels is

$$H_{eff} = \begin{pmatrix} E_0 & \tilde{t} \\ \tilde{t}^* & E_0 \end{pmatrix} , \tag{21.3}$$

where $E_0 \equiv \hbar\omega$ (as discussed above), and the $\tilde{t}$ tunneling matrix element connects the two degenerate states. C_1 and Γ can be computed using advanced methods like WKB approximation and instanton analysis (which you will hopefully learn in advanced courses in quantum mechanics). Diagonalizing the Hamiltonian matrix gives

$$\begin{aligned}(E_0-\lambda)^2 &= |t|^2 = C_1^2 e^{-2\Gamma\frac{t}{\hbar}}\\ \implies E_0-\lambda_{1,2} &= \pm|C_1|e^{-\Gamma\frac{t}{\hbar}}\\ \implies \lambda_1 = E_0-|C_1|e^{-\Gamma\frac{t}{\hbar}}, &\; \lambda_2 = E_0+|C_1|e^{-\Gamma\frac{t}{\hbar}}\end{aligned} \tag{21.4}$$

From eq. 21.1, we know that Δ is equal to the difference in the two energy eigenvalues. Therefore,

$$\Delta = \lambda_2-\lambda_1 = 2|C_1|e^{-\Gamma\frac{t}{\hbar}} \tag{21.5}$$

This is the splitting in the energy levels induced by the tunneling Γ. Note that we had actually solved this problem above in eq.(19.11) in section 18.2.

21.2 The asymmetric double well

For the case of asymmetric double well, the parity symmetry is explicitly lifted. This will not immediately forbid tunnelling, but will instead make it asymmetric and strongly favour the transitions towards the lower well. In this way, an increasing asymmetry will gradually suppress the tunnelling. Thus, the stable ground state configuration is now of the particle in the lower well. And what of the particle in the higher well? The particle, if placed there, can certainly stay there, but there is a strong probability that it will tunnel out of this state and towards the true ground state. This state is thus called *metastable*, and has a finite lifetime, which comes about from the ability of the particle to tunnel out of it.

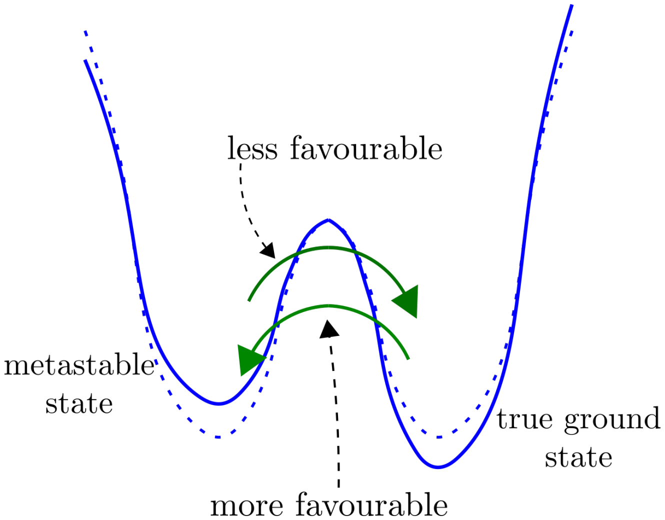


Figure 21.4: Asymmetric double well potential. Dotted line is the symmetric well. Green lines are the transitions between the two wells.

We have even solved the problem of the asymmetric double well earlier. The Hamiltonian is given by eq.(19.93) in section 18.9.1., and can be written as

$$H_{eff} = A\mathcal{I} + b\sigma_x + c\sigma_z \ , \tag{21.6}$$

where A denotes an overall constant shift in the double well potential, b the tunneling amplitude and c the asymmetry between the two states. Please revisit the entire solution given there, and you will immediately see the connection with how the aysmmetry between the wells suppresses the tunneling between them.

21.3 Double Well on the Circle

Thus far, we have considered the case of tunneling for a system defined on an open spatial interval. Recall that other than quantum tunneling, another manifestly quantum mechanical phenomenon (from the perspective of particle mechanics) is quantum interference. While we have encountered this in our discussion of the Feynman double slit gedanken (and it's experimental verification), we have also seen that such interference exists for particle dynamics on the multiply connected manifold such as the circle. As shown in the previous chapter, this is what led to appearance of the Aharanov-Bohm phase in that problem. We can now go one step further and combine the phenomena of quantum tunneling and quantum interference. In order to do so, we consider the problem of the double well potential on the circle. To do so, we impose the potential $V(\phi) = V_0 \cos(2\phi)$, where $0 \leq \phi \leq 2\pi$ is the angular degree of freedom on the circle. Note that $V(\phi) = V(\phi + \pi)$ and $V(-\phi) = V(\phi)$. This is clearly a symmetric double well on the circle, with the two minima at $\phi = \pi/2, 3\pi/2$ and maxima at $\phi = 0, \pi$.

Further, we consider an Aharonov-Bohm flux Φ threading the cicle on which the particle's dynamics is confined. Thus, we can now see that the tunneling amplitude factor in going between the two wells (i.e., half-way across the circle, such that $\phi_{final} - \phi_{initial} = \pm\pi$ for the two oppositely directed tunneling pathways) is given by

$$\tilde{t}_{\Phi} = \tilde{t}(e^{i\frac{\pi\Phi}{\Phi_0}} + e^{-i\frac{\pi\Phi}{\Phi_0}}) = 2\tilde{t}\cos\left(\frac{\pi\Phi}{\Phi_0}\right) , \tag{21.7}$$

where $\Phi_0 = hc/e$ is the magnetic flux quantum and $\tilde{t}$ has already been defined above. The effective Hamiltonian for this problem is then given by

$$H_{eff} = \begin{pmatrix} E_0 & \tilde{t}_{\Phi} \\ \tilde{t}^*_{\Phi} & E_0 \end{pmatrix} . \tag{21.8}$$

The tunnel splitting is then given by

$$\Delta = 2\tilde{t}_{\Phi} = 4\tilde{t}\cos\left(\frac{\pi\Phi}{\Phi_0}\right) . \tag{21.9}$$

It is then easily seen that for the case of the AB flux $\Phi = \frac{(2n+1)}{2}\Phi_0$, $n \in \mathcal{Z}$, the tunnel splitting vanishes, $\Delta \to 0$! This is due to the AB-flux aided destructive quantum interference between the two possible ways of tunneling between the two wells (i.e., along oppositely directed pathways). Similarly, for the case of the AB flux $\Phi = n\Phi_0$, $n \in \mathcal{Z}$, the tunnel splitting is maximised, $\Delta = 4\tilde{t}$; this corresponds to constructive quantum interference ! In this way, we can see that non-trivial quantum interference can either nullify or amplify the effects of quantum mechanical tunneling.

Chapter 22

The Landau Problem: 2D Electrons in a perpendicular Magnetic field

22.1 The Eigenspectrum for the problem

Consider an electron in 2D placed in a perpendicular magnetic field. The Hamiltonian in its most general form is

$$H = \frac{1}{2m}\left(\vec{p} - \frac{e}{c}\vec{A}\right)^2 - \mu\vec{S}\cdot\vec{B} + u\left(\vec{r}\right) \tag{22.1}$$

where $\vec{B} = B\hat{z}$ and $u\left(\vec{r}\right)$ is some confining potential. Since $[H, S_z] = 0$, the spin and spatial parts of the wavefunction will be separable. For the time being, we ignore the effects of $u(\vec{r})$, and focus on solving what remains of the spatial part. We can write the Hamiltonian as

$$\begin{aligned} H_0 &= \frac{1}{2m}\left[p^2 - \frac{e}{c}\left(\vec{A}\cdot\vec{p} + \vec{p}\cdot\vec{A}\right) + \frac{e^2}{c^2}A^2\right] \qquad \left[\vec{B} = \boldsymbol{\nabla}\times\vec{A}\right] \\ &= \frac{1}{2m}\left[p^2 - \frac{e}{c}\left(i\hbar\boldsymbol{\nabla}\cdot A + 2\vec{p}\cdot\vec{A}\right) + \frac{e^2}{c^2}A^2\right] \quad \left[\because \left[\vec{A}, \vec{p}\right] = i\hbar\boldsymbol{\nabla}\cdot A\right] \end{aligned} \tag{22.2}$$

Now if we choose $\vec{A}$ such that $\boldsymbol{\nabla}\cdot A = 0$, this simplifies matters considerably (Landau). Thus, following Landau, we choose

$$\vec{A} = -By\hat{x}, \text{ such that } \boldsymbol{\nabla}\cdot A = 0 \tag{22.3}$$

The Hamiltonian becomes

$$H_0 = \frac{1}{2m}\left[\left(p_x + \frac{eB}{c}y\right)^2 + p_y^2 + p_z^2\right] \tag{22.4}$$

Because H_0 is a function of p_x, p_y, p_z, y but not of x or z, we get

$$[H_0, p_x] = [H_0, p_z] = 0 \tag{22.5}$$

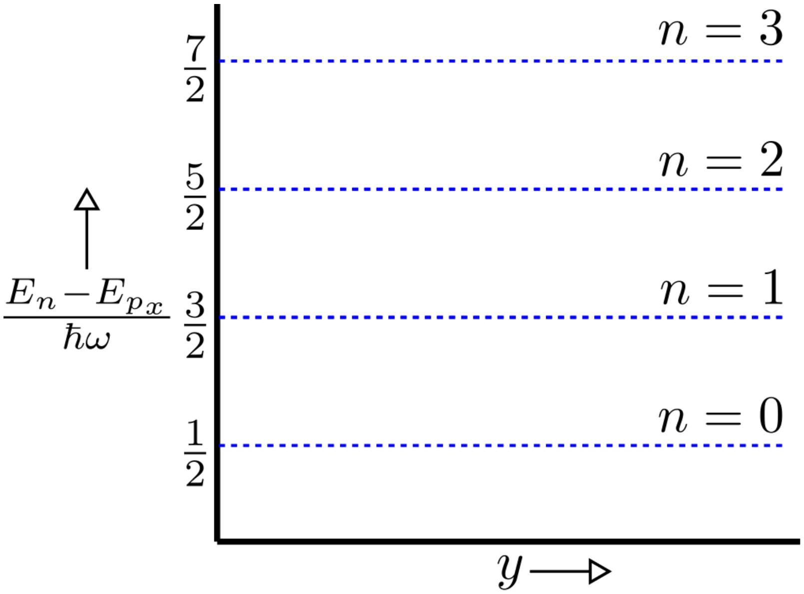


Figure 22.1: The first four Landau levels. Each level has a huge degeneracy.

That is, the particle is free in the x and z directions. Since the motion is in 2D, $p_z = 0$. In x, we have $\psi(x) \sim e^{ip_x \frac{x}{\hbar}}$. In the y-direction, by replacing p_x with its eigenvalue and by defining $y_0 = \frac{p_x c}{|e|B}$ and $\omega = \frac{|e|B}{mc}$, we get

$$H_0 = \frac{1}{2m}p_y^2 + \frac{1}{2}m\omega^2 \left(y - y_0\right)^2 \tag{22.6}$$

which is just a S.H.O in the y-direction. We can immediately write down the eigenvalues and eigenfunctions.

$$\begin{aligned} E(p_x, n) &= \frac{p_x^2}{2m} + \left(n + \frac{1}{2}\right)\hbar\omega \\ \psi_n(x,y) &= e^{i\frac{p_x x}{\hbar}}\chi_n(y) \end{aligned} \tag{22.7}$$

where $\chi_n(y) = c_n e^{-\frac{(y-y_0)^2}{2\xi}} H_n\left(\frac{y-y_0}{\xi}\right), \xi = \sqrt{\frac{\hbar c}{m\omega}}$. H_n is the Hermite polynomial of degree n. The index n characterises the Landau levels. The first four Landau levels are shown in fig. 22.1. Some wavefunctions are shown schematically in fig. 22.2.

To calculate the degeneracy of each level, we consider a finite sized 2D system of large but finite area $A = L_x L_y$, such that $l_B = \sqrt{\frac{\hbar c}{|e|B}} \ll L_x, L_y$. l_B is the magnetic length. The number of possible values of p_x (that is, states of the free eletron in the x-direction), in the momentum interval Δp_x, is given by the phase space $\frac{\Delta x \Delta p_x}{\hbar}$. For $\Delta x = L_x$, number of states is

$$\frac{L_x \Delta p_x^{\max}}{\hbar} \tag{22.8}$$

Now, $p_x = \frac{|e|By_0}{x}$, and $\Delta p_x = \frac{|e|B\Delta y_0}{c}$ with $0 \leq \Delta y_0 \leq L_y$, which means $\Delta p_x^{\max} = \frac{|e|BL_y}{c}$.

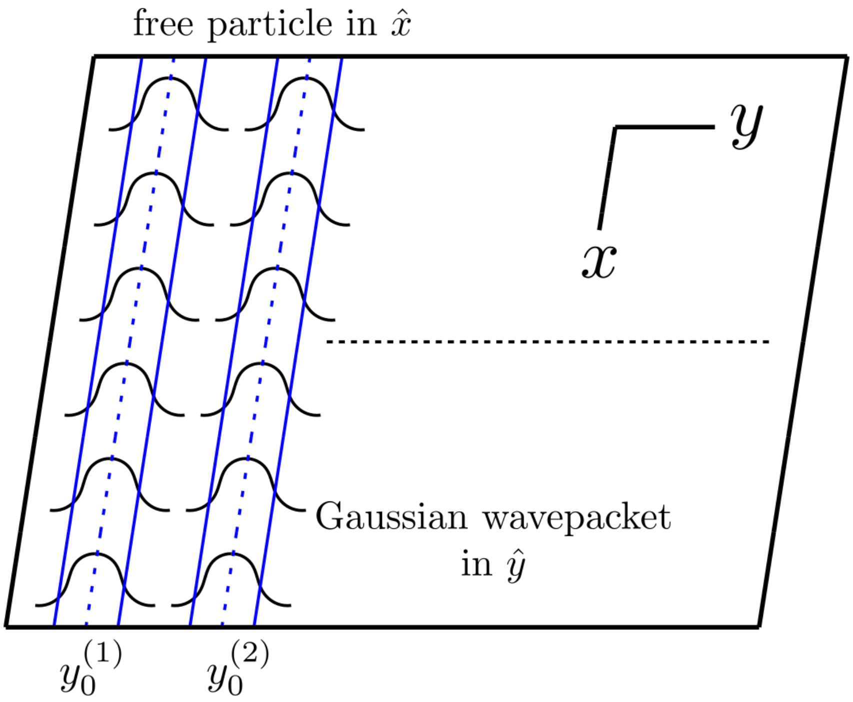


Figure 22.2: Wavefunctions in the $n = 0$ Landau level. $y_0^{(1)} = \frac{cp_x^{(1)}}{|e|B}, y_0^{(2)} = \frac{cp_x^{(2)}}{|e|B}, |p_x^{(2)}| > |p_x^{(1)}|$

Therefore, number of states is

$$\frac{L_x|e|BL_y}{\hbar c} = \frac{L_x L_y}{\frac{2\pi\hbar c}{|e|B}} = \frac{A}{2\pi l_B^2} = N\Phi \tag{22.9}$$

The degeneracy per unit area is $\frac{1}{2\pi l_B^2}$.

22.2 Magnetic Translation Operators and non-commutativity

Defining $\vec{\Pi} = \vec{p} - \frac{e}{c}\vec{A}$, we see that

$$H_0 = \frac{1}{2m}\Pi^2 \tag{22.10}$$

It is worth noting that

$$\begin{aligned}
[\Pi_x, \Pi_y] &= \left[p_x - \frac{e}{c}\vec{A}.p_y - \frac{e}{c}A_y\right] \\
&= \frac{e}{c}\left[p_y, A_x\right] - \frac{e}{c}\left[p_x, A_y\right] \\
&= i\hbar\frac{e}{c}\underbrace{(\nabla_x A_y - \nabla_y A_x)}_{=B} \\
&= i\hbar\frac{eB}{c} \\
&= -i\frac{\hbar^2}{l_B^2}(\neq 0)
\end{aligned} \tag{22.11}$$

where $l_B^2 = \frac{\hbar c}{|e|B}$. Now,

$$\begin{aligned}
[\Pi_x, p_x] &= \left[p_x - \frac{e}{c}A_x, p_x\right] \\
&= \frac{e}{c}\left[p_x, A_x\right] = 0 \text{ for perpendicular B-field}
\end{aligned} \tag{22.12}$$

but

$$\begin{aligned}
[\Pi_x, p_y] &= \left[p_x - \frac{e}{c}A_x, p_y\right] \\
&= \frac{e}{c}\left[p_y, A_x\right] \\
&= -i\hbar\frac{e}{c}\nabla_y A_x \neq 0 \text{ for perpendicular B-field}
\end{aligned} \tag{22.13}$$

Similarly,

$$[\Pi_y, p_y] = 0 \text{ for perpendicular B-field} \tag{22.14}$$

and

$$[\Pi_y, p_x] = i\hbar\frac{e}{c}\nabla_x A_y \neq 0 \text{ for perpendicular B-field} \tag{22.15}$$

In fact, $\left[H_0, \vec{\Pi}\right] \neq 0$ either, from above. Thus, neither $\hat{p}$ nor $\hat{\Pi}$ make good translation operators. In fact, the generator of translations in the presence of a magnetic field is

$$\begin{aligned}
\hat{K} &= \hat{\Pi} + \frac{e}{c}\vec{B} \times \vec{r} \\
&= \vec{p} - \frac{e}{c}\vec{A} + \frac{e}{c}B\hat{k} \times \left(x\hat{i} + y\hat{j} + z\hat{k}\right) \\
&= \vec{p} - \frac{e}{c}\vec{A} + \frac{e}{c}B\left(-y\hat{i} + x\hat{j}\right)
\end{aligned} \tag{22.16}$$

We can now check that $\left[\hat{\Pi}, \hat{K}\right] = 0$, as follows:

$$\begin{aligned}
[\Pi_x, K_x] &= \left[\Pi_x, \Pi_x - \frac{eB}{c}y\right] \\
&= \left[\Pi_x, -\frac{eBy}{c}\right] \\
&= \left[p_x - \frac{eA}{c}x, -\frac{eB}{c}y\right] \\
&= 0 \\
&= [\Pi_y, K_y]
\end{aligned} \tag{22.17}$$

and

$$\begin{aligned}
[\Pi_x, K_y] &= \left[\Pi_x, \Pi_y + \frac{eB}{c}x\right] \\
&= [\Pi_x, \Pi_y] + \left[\Pi_x, \frac{eB}{c}x\right] \\
&= i\hbar\frac{eB}{c} - \frac{eB}{c}\underbrace{[x, p_x]}_{=i\hbar} \\
&= 0 \\
&= [\Pi_y, K_x]
\end{aligned} \tag{22.18}$$

These four commutators imply

$$\left[\vec{\Pi}, K\right] = 0 = \left[H_0, \hat{K}\right] \tag{22.19}$$

such that translation by a vector $\vec{\delta}$ is given by

$$\hat{t}(\delta) = e^{\frac{i}{\hbar}\vec{\delta}\cdot\vec{K}} \tag{22.20}$$

This operator is called the magnetic translation operator, because it causes translations in the magnetic Brillouin zone. Note that

$$\begin{aligned}
[K_x, K_y] &= \left[\Pi_x - \frac{eB}{c}y, \Pi_y - \frac{eB}{c}x\right] \\
&= [\Pi_x, \Pi_y] - \frac{eB}{c}\left([x, p_X] + [y, p_y]\right) \\
&= -i\frac{\hbar^2}{l_B^2} - 2ie\frac{\hbar B}{c} \\
&= -i\frac{\hbar^2}{l_B^2} + 2i\frac{\hbar^2}{l_B^2} \qquad\qquad \left[l_B^2 = \frac{\hbar c}{|e|B}\right] \\
&= \frac{i\hbar^2}{l_B^2} \neq 0
\end{aligned} \tag{22.21}$$

Then, we can show that

$$\begin{aligned}
\hat{t}(a)\hat{t}(b)\hat{t}(a)^{(-1)}\hat{t}(b)^{(-1)} &= e^{-\frac{i}{\hbar}\vec{a}\cdot\hat{K}}e^{-\frac{i}{\hbar}\vec{b}\cdot\hat{K}}e^{\frac{i}{\hbar}\vec{a}\cdot\hat{K}}e^{\frac{i}{\hbar}\vec{b}\cdot\hat{K}} \\
&= \exp\left[-\frac{i}{\hbar}\vec{a}\cdot\hat{K} - \frac{i}{\hbar}\vec{b}\cdot\hat{K} + \frac{i}{\hbar}\vec{b}\cdot\hat{K} + \frac{i}{\hbar}\vec{b}\cdot\hat{K} - \frac{2}{\hbar^2}\left[\vec{a}\cdot\hat{K}, \vec{b}\cdot\hat{K}\right]\right] \\
&= \exp\left[-2\left(a_ib_j - a_jb_i\right)\frac{K_iK_j}{\hbar^2}\right] \\
&= \exp\left[-\left(a_ib_j - a_jb_i\right)\frac{1}{\hbar^2}\left(K_iK_j - K_jK_i\right)\right]
\end{aligned} \tag{22.22}$$

In the second step, we used the BCH relation. But since $[K_x, K_y] = i\frac{\hbar^2}{l_B^2}$, we can write

$$\hat{t}(a)\hat{t}(b)\hat{t}(a)^{(-1)}\hat{t}(b)^{(-1)} = \exp\left[-i\frac{1}{l_B^2}\left(\vec{a}\times\vec{b}\right)\cdot\hat{z}\right] \tag{22.23}$$

as $\hat{t}(\vec{a})\hat{t}(\vec{b}) = \exp\left[-\frac{i}{l_B^2}\left(\vec{a}\times\vec{b}\right)\cdot\hat{z}\right]\hat{t}(\vec{b})\hat{t}(\vec{a})$ → magnetic algebra. Furthermore, the translation operators $\hat{t}(\vec{a})$ form a group in the sense that they obey the following composition law:

$$\hat{t}(\vec{a})\hat{t}(\vec{b}) = \exp\left[\frac{i}{2l_B^2}\left(\vec{a}\times\vec{b}\right)\cdot\hat{z}\right]\hat{t}(\vec{b}+\vec{a}) \tag{22.24}$$

The above relations tell us that in going around (translating) in a closed cycle, we have picked up a phase:

$$e^{-\frac{i}{l_B^2}\left(\vec{a}\times\vec{b}\right)\cdot\hat{z}} = e^{-2\pi i\frac{\Phi_{ab}}{\Phi_0}} \tag{22.25}$$

where $\Phi_{ab} = B\left(\vec{a}\times\vec{b}\right)\cdot\hat{z}$ is the flux enclosed in the plaquette of $\vec{a}, \vec{b}, -\vec{a}$ and $-\vec{b}$. This is the familiar Aharanov-Bohm phase of the particle on a circle with an enclosed solenoidal flux! Further, for $N_\Phi = \frac{\Phi_{ab}}{\Phi_0} \in \mathbb{Z}$ (an integer),

$$\left[\hat{t}(\vec{a}), \hat{t}(\hat{b})\right] = 0 \tag{22.26}$$

That is, the magnetic translation operators in the x and y directions commute with one another. GIven that $\left[\hat{t}(\delta), H_0\right] = 0$, we can diagonalize $H_0, \hat{t}(x)$ and $\hat{t}(y)$ simultaneously. Thus, the generalised (twisted) boundary conditions now are

$$\begin{aligned}\Psi_{\Phi_1\Phi_2}\left(..., x_j + L_1, y_j, ...\right) &= \exp\left[-\frac{2\pi i}{L_2}N_\Phi y_j\right]\Psi_{\Phi_1\Phi_2}\left(..., x_j, y_j, ...\right)\\ \Psi_{\Phi_1\Phi_2}\left(..., x_j, y_j + L_2, ...\right) &= \Psi_{\Phi_1\Phi_2}\left(..., x_j, y_j, ...\right)\end{aligned} \tag{22.27}$$

22.2.1 The Landau Problem for N interacting electrons

Discuss the Tao-Haldane initial steps that lead to the COM and relative degrees of freedom.

- $$\begin{aligned}\left[e^{ia\Pi_x^c}, \Pi_y^{c\,2}\right] &= \Pi_y^c\left[e^{ia\Pi_x^c}, \Pi_y^{c\,2}\right] + \left[e^{ia\Pi_x^c}, \Pi_y^{c\,2}\right]\Pi_y^c\\ &= \Pi_y^c \times \left[-aeB\hbar e^{ia\Pi_x^c}\right] + \left[-aeB\hbar e^{ia\Pi_x^c}\right]\times\Pi_y^c \quad \left[\Pi_\alpha^c = \sum_j \Pi_{j\alpha}, \alpha = x, y\right]\end{aligned} \tag{22.28}$$

- $$\begin{aligned}H_c &= \frac{1}{2mN_e}\left[\left(\Pi_x^c + \frac{N_e\hbar\alpha_1}{L_1}\right)^2 + \left(\Pi_y^c + \frac{N_e\hbar\alpha_2}{L_2}\right)^2\right]\\ \Pi_x^e &= -i\hbar\frac{\partial}{\partial x}\ ,\ \Pi_y^c = -i\hbar\frac{\partial}{\partial y} + eBxN_e \qquad \left[\vec{A} = [0, eBx, 0]\right]\end{aligned} \tag{22.29}$$

- Compute

$$\left[\Pi_x^c, \Pi_y^c\right] = \left[-i\hbar\frac{\partial}{\partial x}, N_e eBx\right] = -i\hbar eBN_e, \quad \left[[x, p_x] = i\hbar\right] \tag{22.30}$$

- Define pseudo-momenta:

$$\begin{aligned} K_x &= \Pi_x^c + eByN_e - i\hbar\frac{\partial}{\partial x} + eByN_e \\ K_y &= \Pi_y^c - eBxN_e = -i\hbar\frac{\partial}{\partial y} \end{aligned} \tag{22.31}$$

  Compute

$$\underbrace{\left[\Pi_x^c, K_x\right]}_{=0}, \quad \underbrace{\left[\Pi_x^c, K_y\right]}_{=0}, \quad \underbrace{\left[\Pi_y^c, K_x\right]}_{\text{check}}, \quad \underbrace{\left[\Pi_y^c, K_y\right]}_{=0} \tag{22.32}$$

$$\begin{aligned} \left[\Pi_y^c, K_x\right] &= -\left[i\hbar\frac{\partial}{\partial y} + N_e eBx, -i\hbar\frac{\partial}{\partial x} + N_e eBy\right] \\ &= \left[-i\hbar\frac{\partial}{\partial y}, N_e eBy\right] + \left[N_e eBx, -i\hbar\frac{\partial}{\partial x}\right] \\ &= -i\hbar N_e eB + i\hbar eN_e B \\ &= 0 \end{aligned} \tag{22.33}$$

  Now compute $[K_x, K_y]$.

$$\begin{aligned} \left[K_x, K_y\right] &= \left[-i\hbar\frac{\partial}{\partial x} + N_e eBy, -i\hbar\frac{\partial}{\partial y}\right] \\ &= N_e eB\left[y, -i\hbar\frac{\partial}{\partial y}\right] \\ &= -i\hbar eBN_e \\ &= -\left[\Pi_x^c, \Pi_y^c\right] \end{aligned} \tag{22.34}$$

- Compute $[K_x, H_c] = 0, [K_y, H_c] = 0$
- Define

$$T_1 = \exp\left\{i\frac{L_1 K_x}{N_s\hbar}\right\}, \quad T_2 = \exp\left\{i\frac{L_2 K_y}{N_s\hbar}\right\}, \quad \frac{\hbar N_s}{e} = BL_1L_2\ . \tag{22.35}$$

- Compute $T_1^{-1}T_2^{-1}T_1T_2$. Use the fact that $e^{-A}e^{-B}e^{A}e^{B} = e^{-[A,B]}$ if $[A,[A,B]] = 0 = [B,[A,B]]$ and the BCH relation $e^A e^B = \exp\left[A + B + \frac{1}{2}[A,B]\right]$. Then,

$$\begin{aligned} T_1^{-1}T_2^{-1}T_1T_2 &= \exp\left[-\frac{L_1L_2}{N_s^2\hbar^2}\left[K_x, K_y\right]\right] \\ &= \exp\left[-i\frac{\hbar eBL_1L_2}{N_s^2\hbar^2}N_e\right] \qquad \left[BL_1L_2 = \frac{h}{e}N_s\right] \\ &= \exp\left[-i\frac{e\frac{h}{e}N_s}{N_s^2\hbar}N_e\right] \qquad \left[N_s = N_\Phi\right] \\ &= \exp\left[-i2\pi\frac{N_e}{N_s}\right] \end{aligned} \tag{22.36}$$

- Define

$$\gamma = -\frac{1}{2\pi}\mathrm{Im}\left[\ln\left(T_1^{-1}T_2^{-1}T_1T_2\right)\right] \tag{22.37}$$

Compute

$$\begin{aligned}\gamma &= -\frac{1}{2\pi}\mathrm{Im}\left[\ln\exp\left(-i2\pi\frac{N_e}{N_s}\right)\right]\\ &= -\frac{1}{2\pi}\left(-2\pi\right)\frac{N_e}{N_s}\\ &= \frac{N_e}{N_s}\end{aligned} \tag{22.38}$$

In the QHE, $\sigma_H^{2D} = \gamma\frac{e^2}{h}$. Thus we can see that

$$\begin{aligned}T_1\left|\psi_0^c\right\rangle &= T_1\psi_0^c\left(x_1, y_1, ..., x_j y_j, ...\right)\\ &= T_1\psi_0^c\left(x_1+L, y_1, ..., x_j+L, y_j, ...\right)\end{aligned} \tag{22.39}$$

Bibliography

[1] David J. Griffiths. *Introduction to Quantum Mechanics (2nd Edition)*. Pearson Prentice Hall, 2004.

[2] Donald A. Mcquarrie. *Quantum Chemistry, 2nd edition*. University Science Books,U.S., 2 edition, 2007.