

PH3102

QUANTUM MECHANICS

LECTURE NOTES

Instructor : Dr. Siddhartha Lal

Email : slal@iiserkol.ac.in

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

AUTUMN SEMESTER, 2024

INDIAN INSTITUTE OF SCIENCE EDUCATION & RESEARCH,
KOLKATA



Contents

Contents	4
1 Introduction	8
1.1 Getting started	8
1.2 The Schrödinger's Cat gedanken	9
1.3 A historical background	12
2 The Origins of Quantum Theory	15
2.1 Solving the Blackbody Radiation Spectrum Puzzle	15
2.1.1 Blackbody Radiation	15
2.1.2 Classical theory (heuristic derivation)	17
2.1.3 Planck's Idea	18
2.2 The Dual Nature of Light	21
2.2.1 Wave-like: Diffraction and Interference	21
2.2.2 Particle-like: the Photoelectric effect	21
2.2.3 de Broglie's Idea	23
2.3 The Dual Nature of Matter	24
2.3.1 Double-slit experiment with pellets	24
2.3.2 Double-slit experiment with electrons	26
2.4 Spying on the Electrons	29
2.5 Fundamental Postulates of Quantum mechanics	31
2.6 Bohr's Atom	32
2.7 Conclusions	35
3 The Foundations of Quantum Mechanics	37
3.1 Recap: Fundamental Postulates of Quantum mechanics	37
3.1.1 Heuristic derivation of the Schrödinger equation	38
3.2 Physical Meaning of Ψ	41
3.3 Observables, Operators and Expectation values	42
4 Formalism I : Operators, Eigenfunctions and Eigenvalues	45
4.1 Operators, Eigenfunctions and Eigenvalues	45
4.1.1 Definition of an operator	45
4.1.2 Eigenfunctions and Eigenvalues	46

4.1.3	Commutators	47
4.2	Important properties of Eigenstates	49
5	Formalism II : Vector Spaces in Quantum Mechanics	53
5.1	Vector Spaces	53
5.1.1	Dimension of a vector space:	54
5.1.2	More on the Inner Product	54
5.2	The Dirac Bra-Ket Notation	56
5.3	Properties of $ \psi\rangle$, $\langle\psi $ and $\langle\phi \psi\rangle$	56
5.4	Operators	57
5.5	Discrete & Continuous Basis	60
5.6	Position and Momentum Basis	60
5.6.1	Momentum operator in Position Basis	63
5.7	Change of Basis	64
5.8	Matrix representation of Operators & Matrix mechanics	67
6	Heisenberg's Uncertainty principle	68
6.1	The case of a wavepacket	68
6.2	The γ -ray microscope gedanken	69
6.3	A derivation of the HUP for wavepackets	70
6.3.1	A generalised Uncertainty principle	72
6.3.2	The Energy-time uncertainty relation	73
7	The Time Dependent and Independent Schrödinger Equations (TDSE & TISE)	76
7.1	Qualitative Solutions and the origin of Quantisation	78
8	Quantum dynamics: the Schrödinger, Heisenberg and interaction pictures	81
8.0.1	The interaction picture	83
9	Schrödinger Wave Mechanics: Bound states	87
10	Infinite square well/ Particle in a Box	89
11	Simple Harmonic Oscillator	92
11.1	Algebraic Method	93
11.2	Analytic Method	102
12	Free particle	106
13	Scattering Processes and Quantum Tunneling	108
13.1	Step potential	108
13.2	Potential Barrier & Tunneling	111
14	Time Independent Schrödinger Equation (TISE) in 3D	116
14.1	The Particle in a 3D Box	116

15 The Angular Momentum problem: first passage	118
15.1 Introduction	118
15.2 Total Angular Momentum	120
16 Orbital Angular Momentum Eigenstates	125
17 Central Potential Problem	131
18 The Hydrogen Atom	135
19 Spin Angular Momentum	141
19.1 Evidence for spin	141
19.1.1 The Zeeman Effect	141
19.1.2 The Stern-Gerlach Experiment	142
19.2 Eigenstates and Eigenvalues of the Spin Operators	143
19.3 Matrix Representation of Spin	143
19.4 Properties of Pauli matrices	146
19.5 Eigenstates of S_x	146
19.6 Spin Precession	148
19.6.1 Spin is along +z at t=0	149
19.6.2 Spin is along +x at t=0	149
19.6.3 The Heisenberg equation of motion approach to spin precession	151
19.7 Spin and spatial degrees of freedom	152
20 Addition of two angular momenta	153
20.1 The general problem	153
20.1.1 Addition of two spin-1/2 operators	154
20.2 Adding two total angular momenta \vec{J}_1 and \vec{J}_2	157
20.2.1 $j_1 = 1, j_2 = 1$	160
20.2.2 $j_1 = 1, j_2 = 1/2$	162
20.3 Adding Orbital and Spin Angular Momenta	164
21 Symmetries in Quantum Mechanics	169
21.1 Symmetries, Conservation laws and Degeneracies	169
21.2 A few important points on Unitary Operators	171
21.3 Continuous Symmetries: Translation and Rotation symmetry	172
21.3.1 Spatial translation invariance	172
21.3.2 Time Translation invariance	175
21.3.3 Rotational invariance	175
21.4 Discrete Symmetries: Parity, Time-reversal and all that	176
21.4.1 Parity	176
21.4.2 Time Reversal	178
Time-reversal properties of particles with non-zero spin	180
21.4.3 Discrete translation invariance and Bloch's theorem	181
21.4.4 Exchange Symmetry	187

21.5 Total number conservation and global phase rotation symmetry of the many-particle wavefunction	190
21.5.1 In the presence of a gauge field.	191
22 Approximate Methods I: Time-independent Perturbation Theory	195
22.1 Non-degenerate Perturbation theory	195
22.1.1 Wavefunction Renormalisation	201
22.2 The anharmonic oscillator with a quartic perturbation potential	204
22.3 Selection Rules	211
22.4 The Quadratic Stark Effect	212
22.5 Sum Rules	216
22.6 Degenerate Perturbation theory	221
22.7 Linear Stark effect: atom with a permanent intrinsic dipole moment	222
22.8 Fine structure and Spin-Orbit Coupling	224
22.9 The Zeeman Effect	226
22.9.1 Weak Zeeman effect	226
22.10 Hyperfine splitting	227
Bibliography	230

Foreword

These are notes partly based on 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. The rest of the matter covers topics in the next course (PH3012). 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 been 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 interesting 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 existing 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 grad-

uate students Santanu Pal, Anirban Mukherjee, Siddhartha Patra and Abhirup Mukherjee 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
- Y. B. Band and Y. Avishai, “Quantum Mechanics with Applications to Nanotechnology and Information Science”, 1st Edition, Academic Press, 2013 .

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

QUANTUM PHYSICS!

Features of the Quantum World

1. The physical universe is not deterministic!
 - (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).

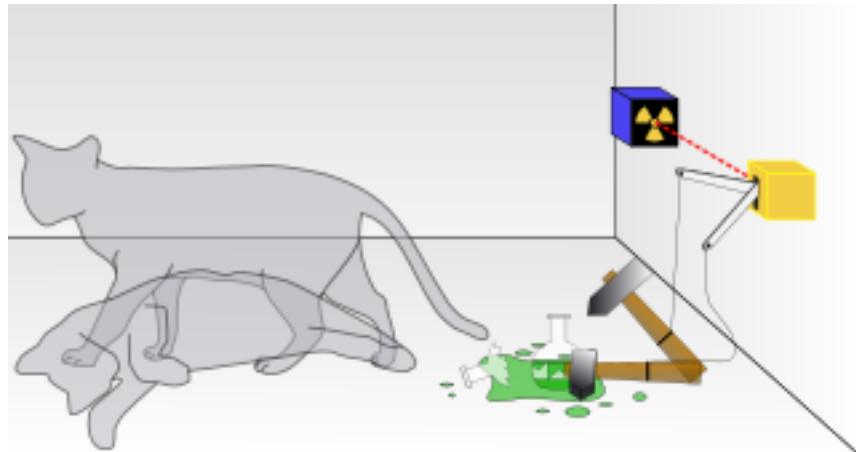


Figure 1.1: A schematic diagram of the Schrödinger's Cat thought experiment. See discussion in text. Disclaimer: No such experiment has ever been carried out with a real cat! Source: The internet.

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.

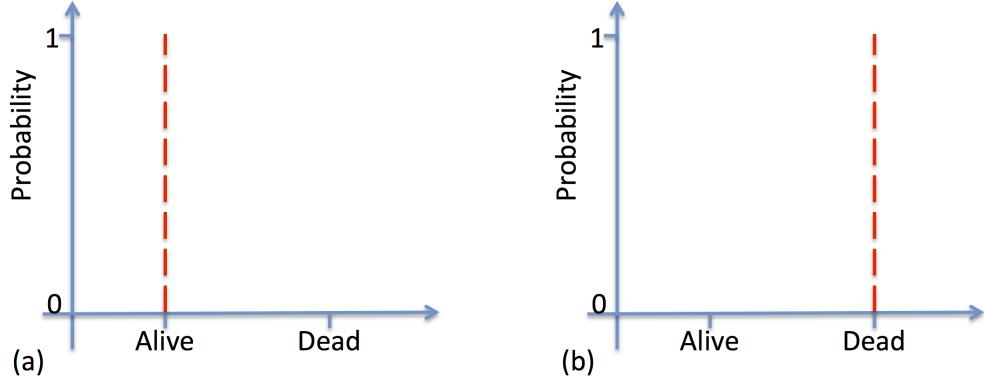


Figure 1.2: A schematic diagram of the probability for observations of the “Dead” and “Alive” states in the Schrödinger’s Cat thought experiment by continuous classical observations. See discussion in text.

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.

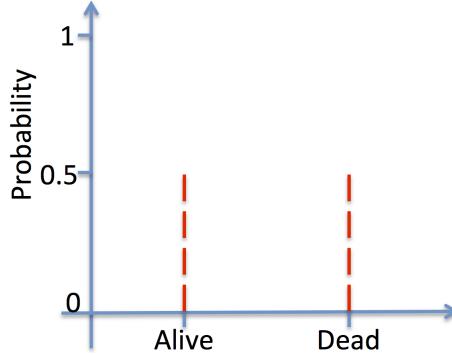


Figure 1.3: A schematic diagram of the probability for observations of the “Dead” and “Alive” states in the Schrödinger’s Cat thought experiment by a classical observer who only observes the state of the cat after a certain time interval T . See discussion in text.

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:

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

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

where the P s 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).

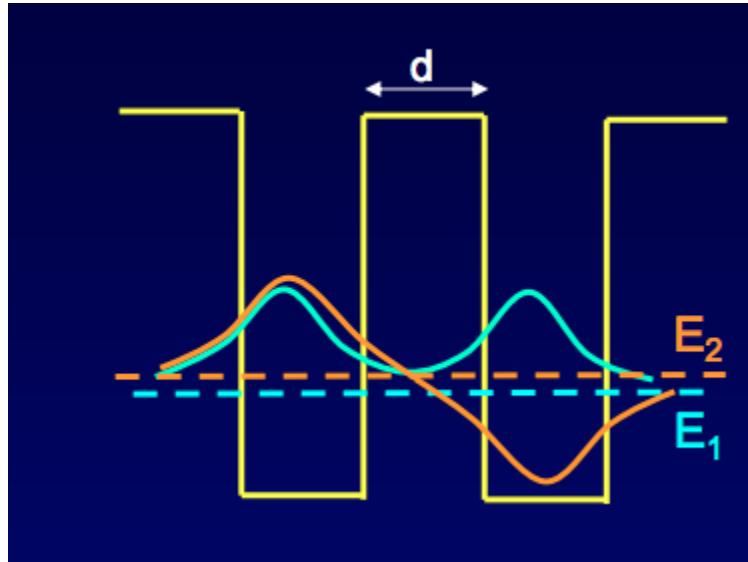


Figure 1.4: The wavefunctions for the symmetric and anti-symmetric states of a quantum particle placed in a double well potential. The symmetric state is lower in energy than the anti-symmetric one due to quantum tunneling. Source: The internet.

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. 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. 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. 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 that’s 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,



SOLVAY CONFERENCE 1927

colourised by pastincolour.com

A. PICARD	E. HENRIOT	P. EHRENFEST	Ed. HERSEN	Th. DE DONDER	E. SCHRÖDINGER	E. VERSCHAFFELT	W. PAULI	W. HEISENBERG	R.H FOWLER	L. BRILLOUIN
P. DEBYE	M. KNUDSEN	W.L. BRAGG	H.A. KRAMERS	P.A.M. DIRAC	A.H. COMPTON	L. de BROGLIE	M. BORN		N. BOHR	
I. LANGMUIR	M. PLANCK	Mme CURIE	H.A. LORENTZ	A. EINSTEIN	P. LANGEVIN	Ch.E. GUYE	C.T.R. WILSON	O.W. RICHARDSON		

Absents : Sir W.H. BRAGG, H. DESLANDRES et E. VAN AUBEL

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.

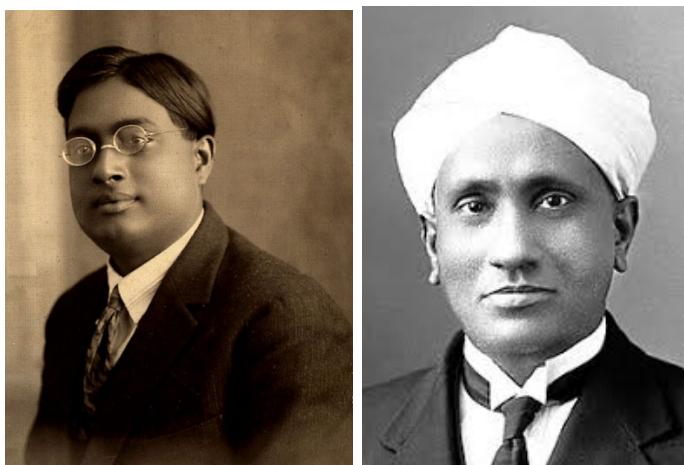


Figure 1.6: Left: Prof. Satyendranath Bose. Right: Prof. C. V. Raman Source: The internet.

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's law: Conservation of heat current in a system which is in thermal equilibrium with its surroundings \Rightarrow the rate of heat absorption = the rate of emission.

Thermal equilibrium \Rightarrow Temperature of System = Temperature of surroundings.

Kirchoff

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.



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

$$P = \sigma AT^4 , \quad (2.1)$$

Stefan

where $\sigma = 5.67 \times 10^{-8} \text{ Js}^{-1} \text{ m}^{-2} \text{ K}^{-4}$ 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



Boltzmann

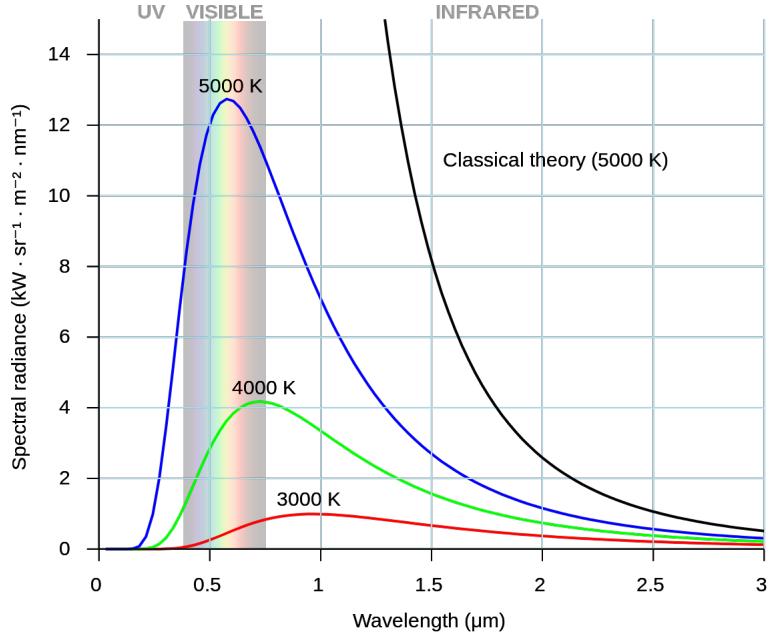
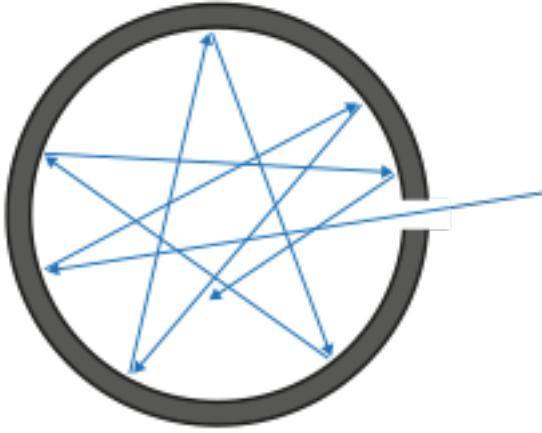


Figure 2.1: Left: A realisation of a blackbody cavity. Right: A plot of Spectral Radiance vs. wavelength for blackbody radiation. Source: The internet.

inside the blackbody

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

$$= \int_0^\infty \rho(\nu) d\nu \quad (2.3)$$

where $\rho(\nu)$ is the energy density in a given interval of the frequency of emitted radiation ranging from ν to $d\nu$. The constant $a = \frac{4\sigma}{c} = 7.56 \times 10^{-16} \text{ Jm}^{-3} \text{ K}^{-4}$. 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.



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: $\nu_{peak} \propto T$.

Wien

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

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

Consequences:

- (i) At $T \sim 300K$, λ_{peak} is in IR range \Rightarrow Thermal imaging glasses useful!
- (ii) At $T \sim 5000K$, λ_{peak} shifts into the optical range \Rightarrow , 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

$$\rho(\nu) = \bar{E} n(\nu) , \quad (2.5)$$

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

From the classical law of equipartition of energy,

$$\bar{E} = k_B T , \quad (2.6)$$

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

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

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

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

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

for e-m waves, together with the differential volume element for the case of $n \gg 1$ (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

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

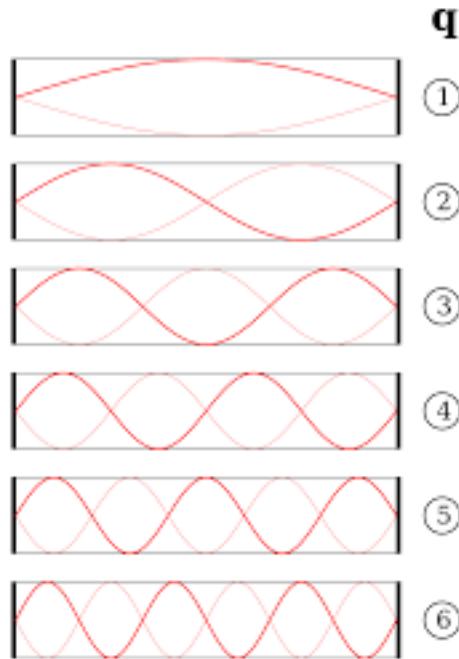


Figure 2.3: Normal modes/harmonics of a linear cavity. The topmost figure is the fundamental or first harmonic, $n = 1$, the next is the first higher harmonic ($n = 2$) and so on.
Source: The internet.

In turn, this leads to

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

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

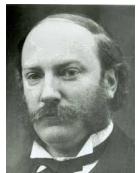


Jeans

2.1.3 Planck's Idea

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

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



Rayleigh

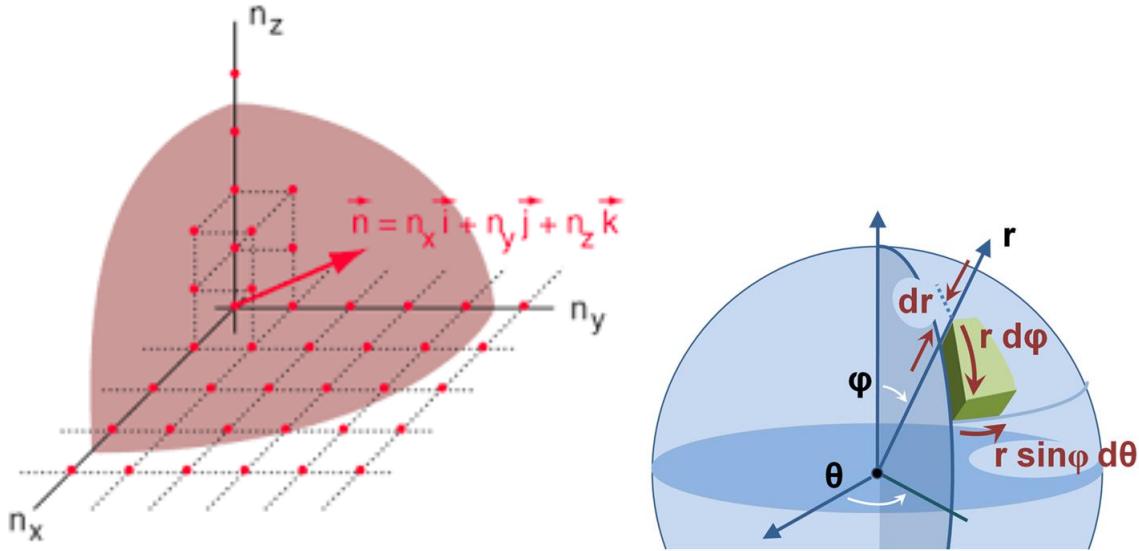


Figure 2.4: (Left) Normal modes of the cavity distributed in (n_x, n_y, n_z) space of integers. (Right) The Volume element in Spherical Polar Coordinates. Source: The internet.

which is the Boltzmann distribution. This gives us

$$\bar{E} = \frac{\int_0^\infty EP(E)dE}{\int_0^\infty P(E)dE} = k_B T \frac{\int_0^\infty d(\frac{E}{k_B T}) \frac{E}{k_B T} e^{-\frac{E}{k_B T}}}{\int_0^\infty d(\frac{E}{k_B T}) e^{-\frac{E}{k_B T}}} \quad (2.12)$$

$$= k_B T \frac{\int_0^\infty dx x e^{-x}}{\int_0^\infty dx e^{-x}} \quad (2.13)$$

$$= k_B T . \quad (2.14)$$



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)

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



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

$$\bar{E} = \frac{\sum_{E=0, h\nu, 2h\nu, \dots} EP(E)}{\sum_E P(E)} \quad (2.16)$$

Planck

$$= \frac{\sum_{n=0}^\infty \frac{n h \nu}{k_B T} e^{-n h \nu / k_B T}}{\sum_{n=0}^\infty \frac{e^{-n h \nu / k_B T}}{k_B T}} \quad (2.17)$$

$$= \frac{h \nu}{e^{n h \nu / k_B T} - 1} .$$

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

$$1 + x + x^2 + x^3 + \dots = \frac{1}{1-x} \quad (2.18)$$

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

This gives us

$$\rho(\nu)d\nu = \bar{E}n(\nu)d\nu \quad (2.20)$$

$$\begin{aligned} &= \frac{8\pi}{c^3} \nu^2 d\nu \frac{h\nu}{e^{nh\nu/k_B T} - 1} \\ &= \frac{8\pi h}{c^3} \frac{\nu^3}{e^{nh\nu/k_B T} - 1} d\nu \end{aligned} \quad (2.21)$$

$$\rightarrow \frac{8\pi k_B T}{c^3} \nu^2 d\nu \quad \text{for } h\nu \ll 1 \quad (\text{Rayleigh-Jeans limit}) \quad (2.22)$$

$$\rightarrow \frac{8\pi h\nu^3}{c^3} e^{-nh\nu/k_B T} \quad \text{for } h\nu \gg 1 \quad (\text{Wien limit}) . \quad (2.23)$$

Thus, we find the total energy density ρ as

$$\rho = \int_0^\infty d\nu \rho(\nu) \quad (2.24)$$

$$\begin{aligned} &= \frac{8\pi}{c^3 h^3} (k_B T)^4 \int_{x=0}^\infty dx \frac{x^3}{e^x - 1} \quad (x = \frac{h\nu}{k_B T}) \\ &= \frac{8\pi^5 k_B^4}{15 c^3 h^3} T^4 \quad \text{as } \int_{x=0}^\infty dx \frac{x^3}{e^x - 1} = \frac{\pi^4}{15} , \end{aligned} \quad (2.25)$$

$$= a T^4 \quad \text{where } a = \frac{8\pi^5 k_B^4}{15 c^3 h^3} . \quad (2.26)$$

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

$$\sigma = a \times \frac{c}{4} = \frac{8\pi^5 k_B^4}{15 c^3 h^3} \times \frac{c}{4} \quad (2.27)$$

$$= \frac{2\pi^4 k_B^4}{15 c^2 h^3} . \quad (2.28)$$

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 $k_B = 1.38 \times 10^{-23} \text{ J K}^{-1}$, 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

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

which is numerically found to be

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

Putting in the values of k_B , T and ν_{max} obtained from the experiments, we again obtain $h = 6.6626 \times 10^{-34} \text{ Js}$. 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.



Planck

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.

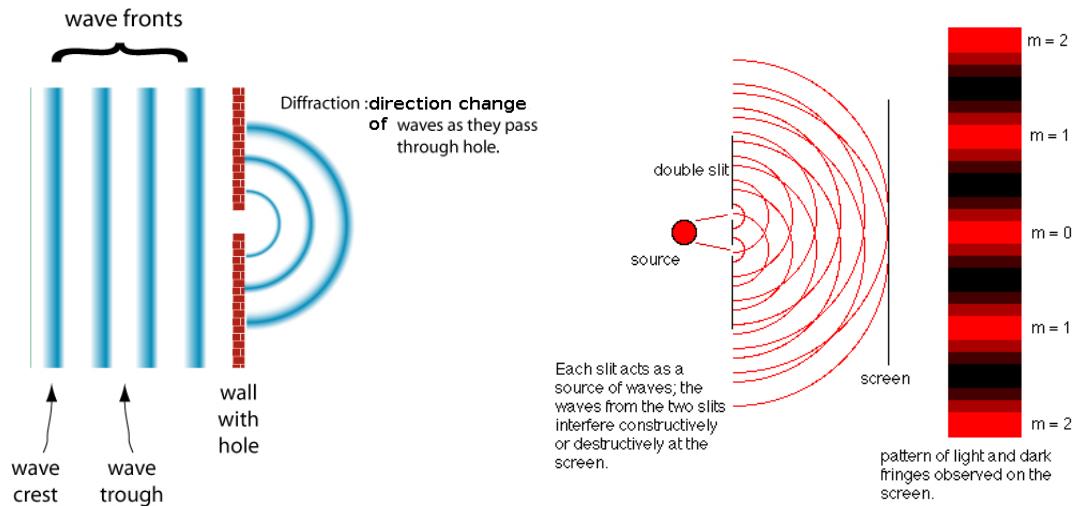


Figure 2.5: Left: Diffraction of waves from a single slit. Right: Interference of waves from a double-slit setup. Source: The internet.

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.

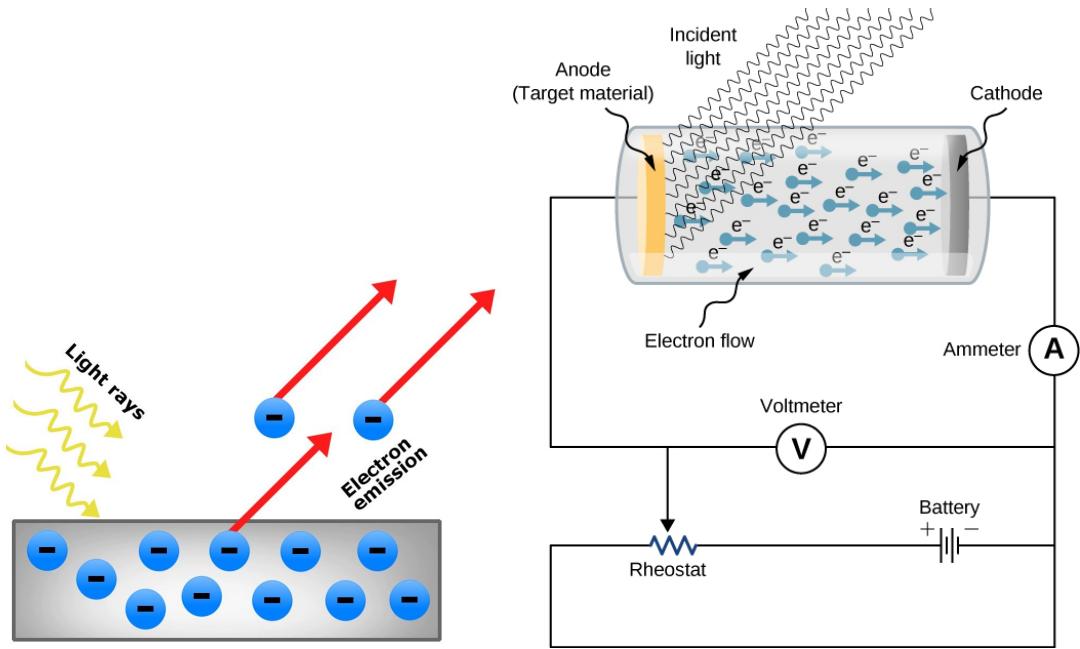


Figure 2.6: Left: The Photoelectric Effect. Right: A schematic diagram of the apparatus used for observing the Photoelectric effect. Source: The internet.

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

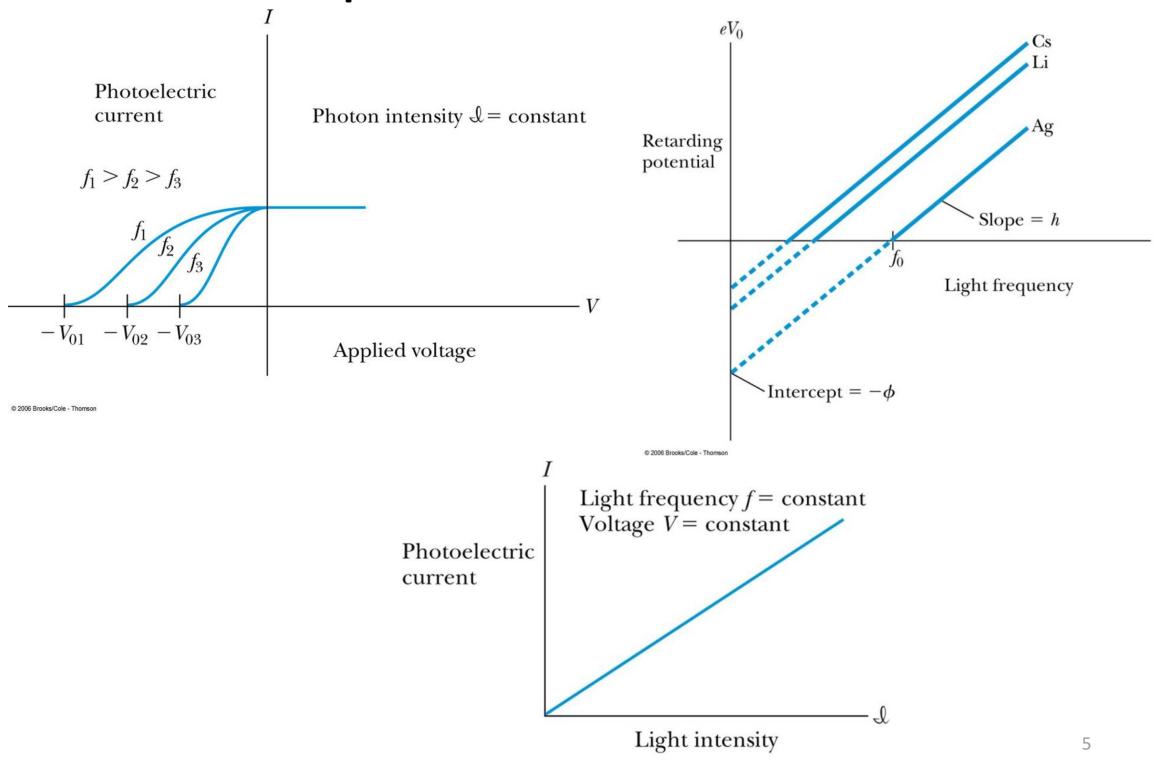


Figure 2.7: Typical experimental observations made for the Photoelectric Effect. Source: The internet.

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.



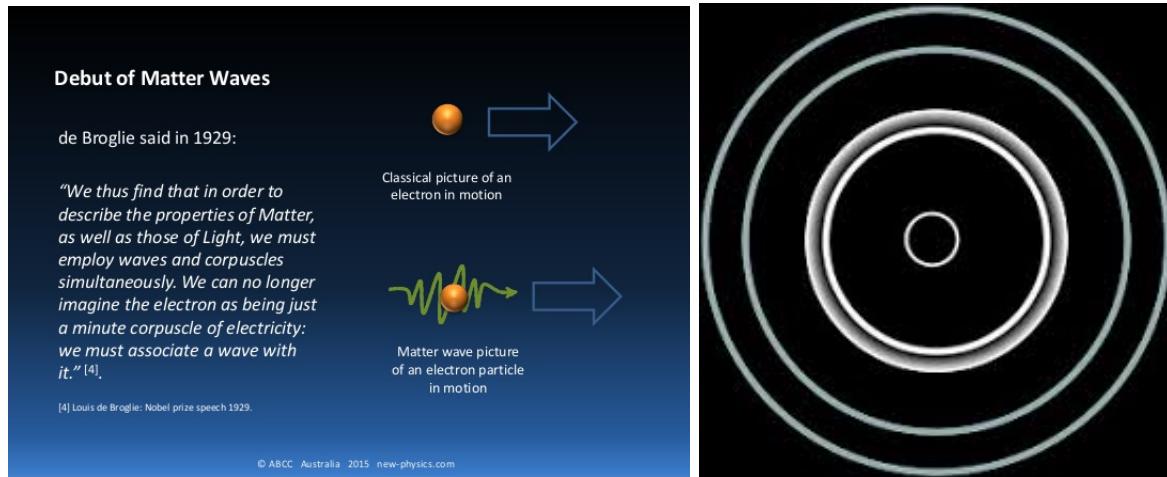
Einstein 2.2.3 de Broglie's Idea

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

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



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

Figure 2.8: Left: de Broglie “Matter Waves”. Right: The characteristic 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.



Born

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).

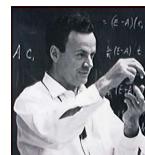
2.3.1 Double-slit experiment with pellets

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



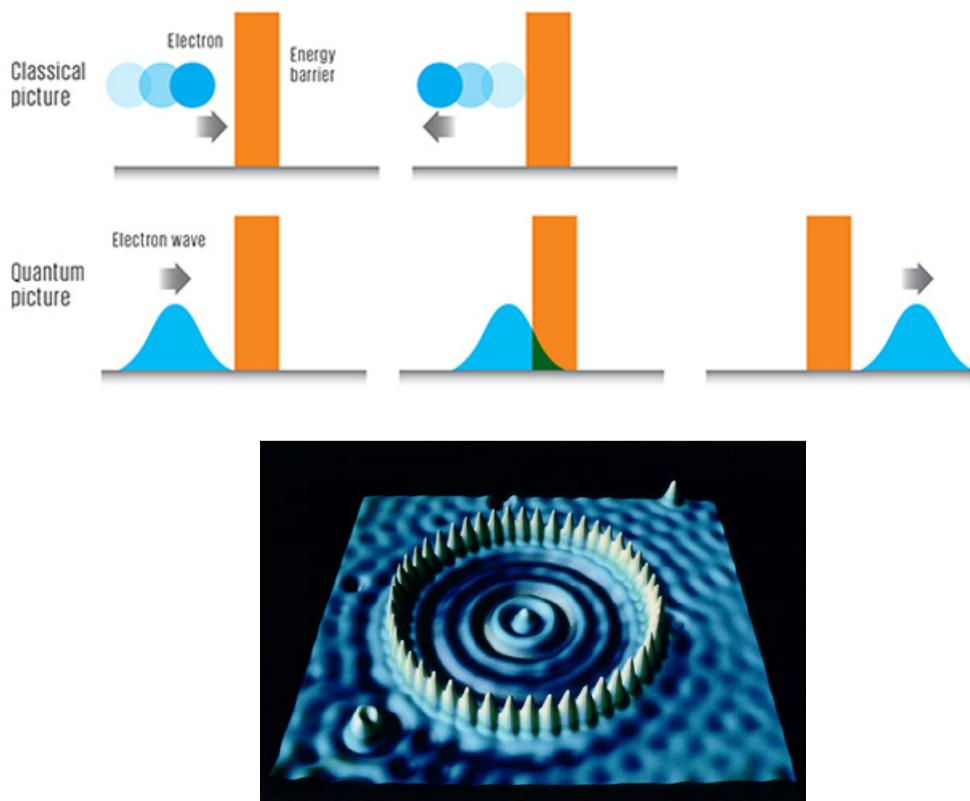


Figure 2.9: Above: The phenomenon of tunneling. Below: A “quantum corral” observed via STM measurements of Iron atoms placed on a Cu surface. Source: The internet.

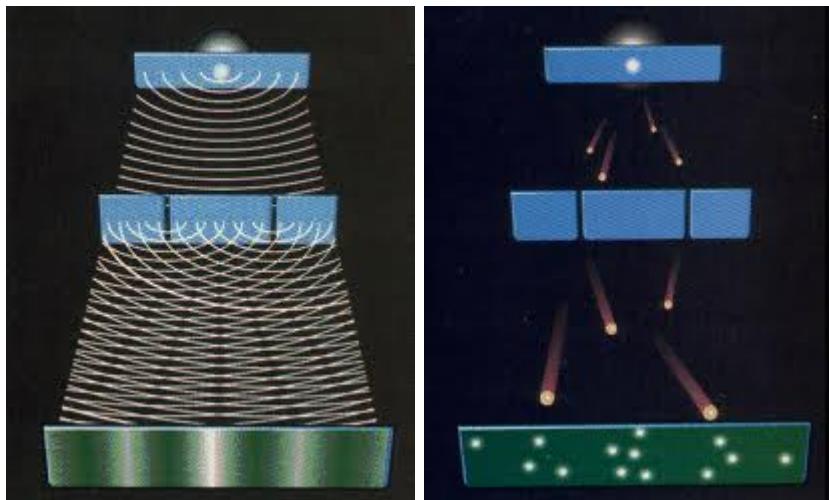


Figure 2.10: (Left) Double-slit Interference Experiment with Light waves. (Right) Double-slit Interference Experiment with Matter particles. Source: The internet.

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.

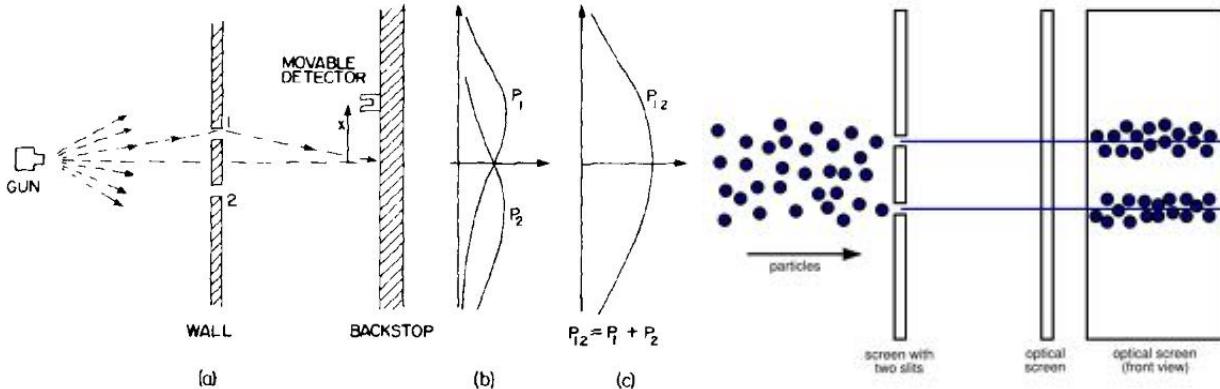


Figure 2.11: Double-slit experiment with pellets. Source: Feynman Lectures Vol.III and the Internet.

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.

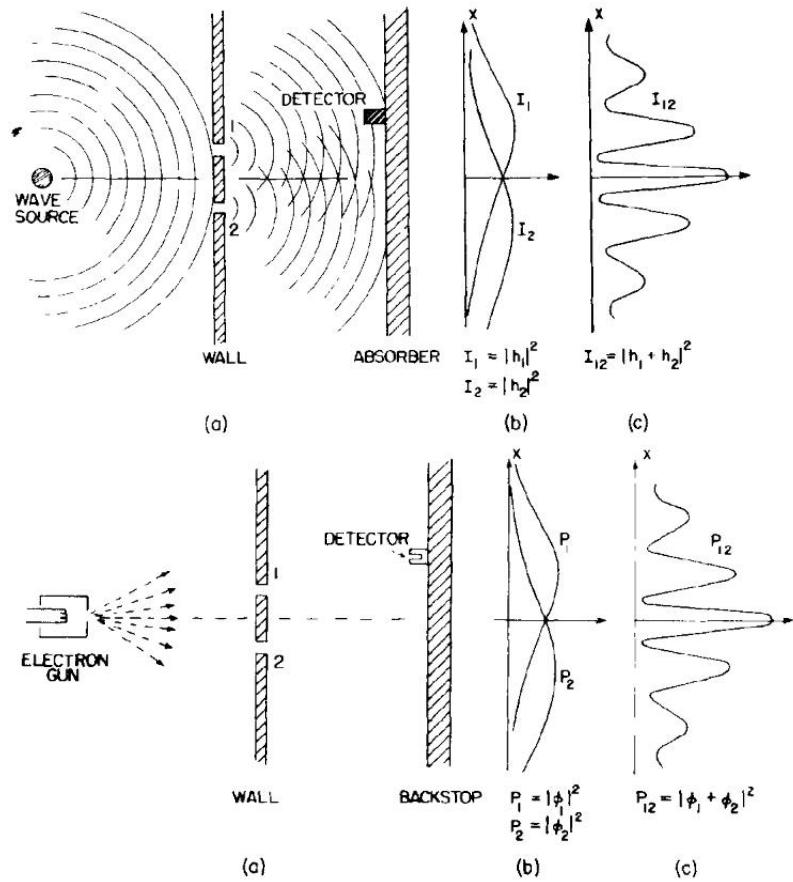


Figure 2.12: Above: Double-slit experiment with light waves. Below: Double-slit experiment with electron “matter waves”. Source: Feynman Lectures Vol.III.

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)?

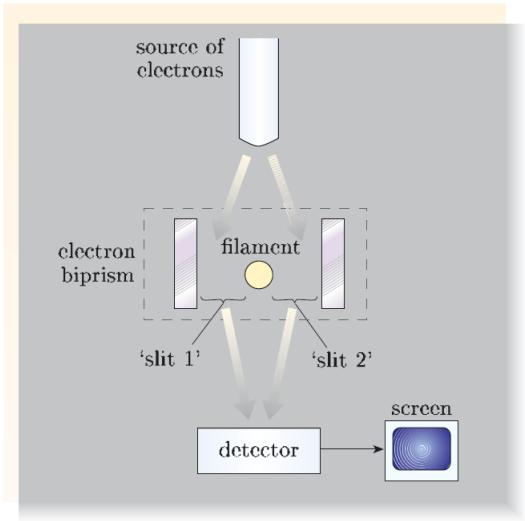


Figure 1.12 Schematic representation of an experimental arrangement used by Tonomura in 1989 to demonstrate the wave–particle duality of electrons. The two paths of the electrons around the filament are analogous to the paths through two slits in an otherwise absorbing screen.



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.

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

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

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.

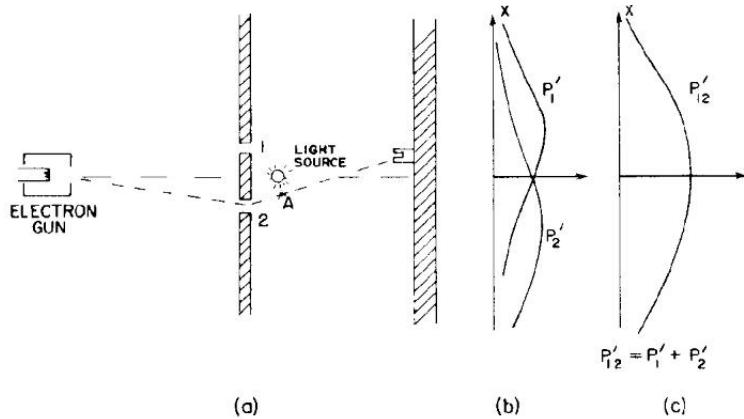


Figure 2.14: Double-slit experiment with electron “matter waves” and a light source by which to spy on which slit the electrons pass through. Source: Feynman Lectures Vol.III.

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!

Conclusion: the electrons can retain their ability to show quantum interference when they are not being observed.

We note the fact that since the photon momentum

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

if we use a very large λ , $p \rightarrow 0$ and the momentum imparted will be almost negligible. So, lets use a very large λ light source (\rightarrow 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., $P'_{12} \rightarrow P_{12}$, indicating that the disturbance from scattering off the photons becomes negligible enough for the interference pattern to emerge.

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.

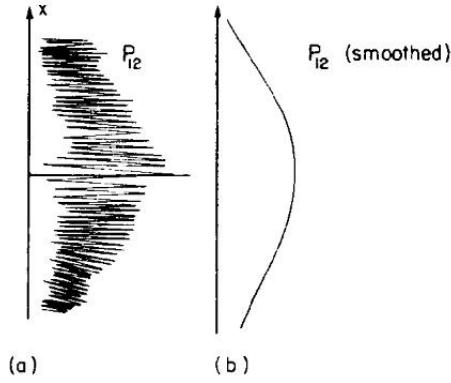


Figure 2.15: Smoothed out “envelope” of the rapid oscillations related to the matter waves.
Source: Feynman Lectures Vol.III.

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:

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

Then, the probability distribution shows signature of quantum interference

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

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

$$P = |\psi_1|^2 + |\psi_2|^2 + 2\psi_1^*\psi_2 \quad (2.32)$$

$$= P_1 + P_2 + 2\sqrt{P_1 P_2} \cos \delta \quad (2.33)$$

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).

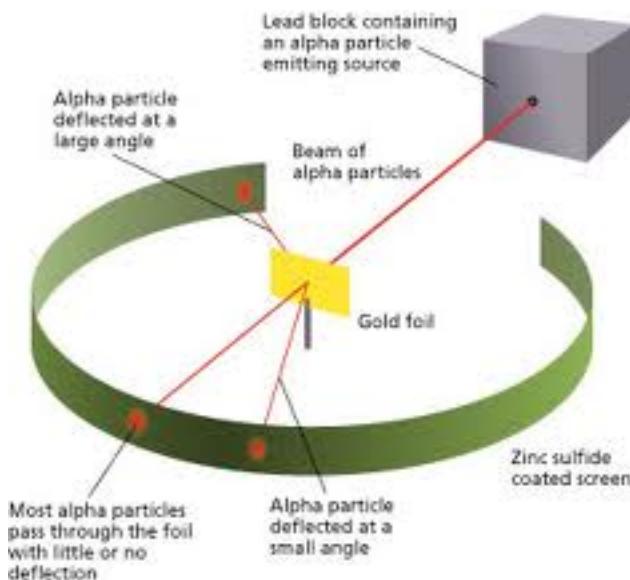
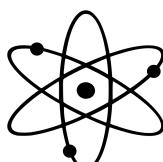


Figure 2.16: A schematic diagram of the experimental setup for the Rutherford gold foil scattering experiment, in which positively charged α particles (that emerge from the radioactive decay process of the source) are directed at a target thin gold foil. The observation of large-angle scattering provides 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 nucleus! But no such radiation is observed; the electron orbits are stable. How can this be?
- (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 H₂ gas is heated through electrical discharge:



Rydberg

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

where $m = 1, 2, 3, \dots$, $n = 2, 3, 4, \dots$, $n > m$ and the Rydberg constant $R = 1.097 \times 10^7 \text{ m}^{-1}$. 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.

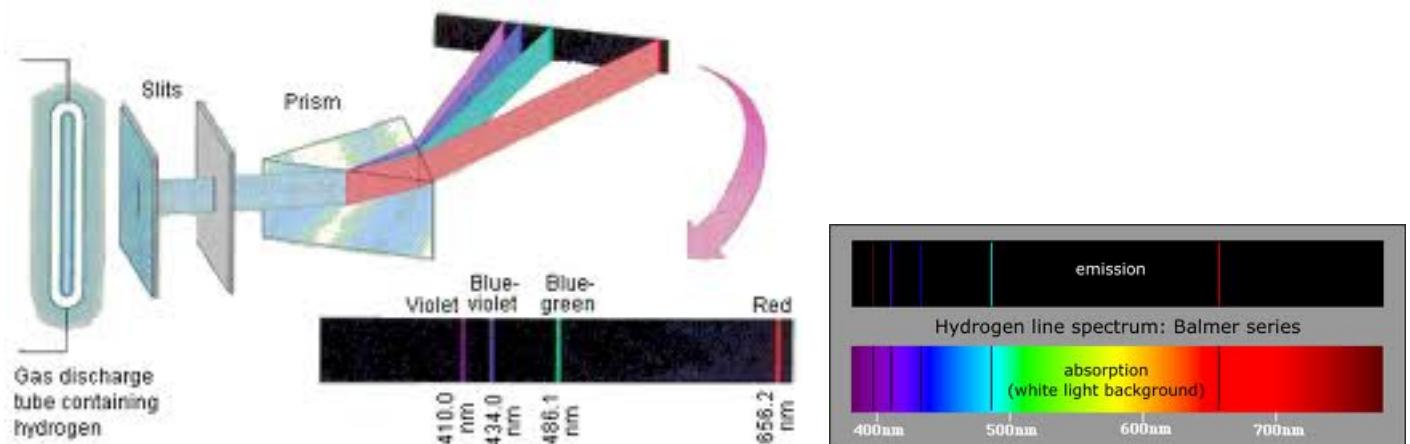


Figure 2.17: Left: A schematic diagram of the experimental setup for the observation of atomic spectra. Right: The atomic spectra of the Balmer series of the Hydrogen atom. Source: The internet.

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**



$$L = pr = mvr = n \frac{h}{2\pi} = n\hbar, \quad n = 1, 2, 3, \dots \quad (2.35)$$

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

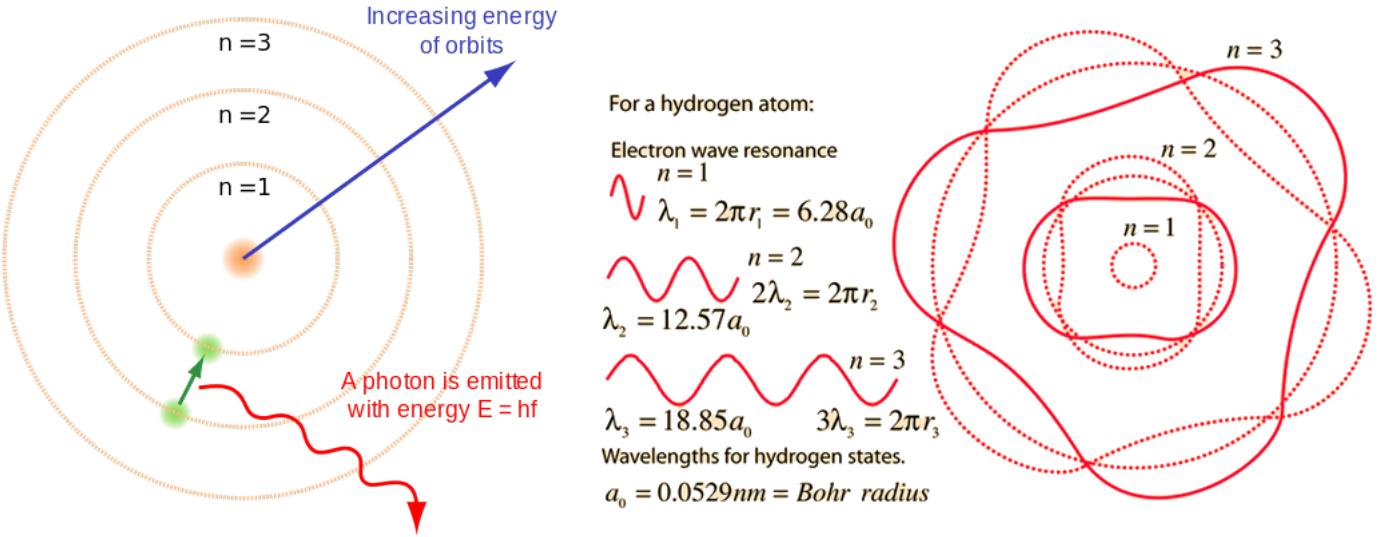


Figure 2.18: Left: A schematic diagram of the experimental setup for the observation of atomic spectra. Right: The atomic spectra of the Balmer series of the Hydrogen atom. Source: The internet.

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

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

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

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

Solving these two relations together gives

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

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

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



Balmer



Lyman



Pfund



Paschen

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

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

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!

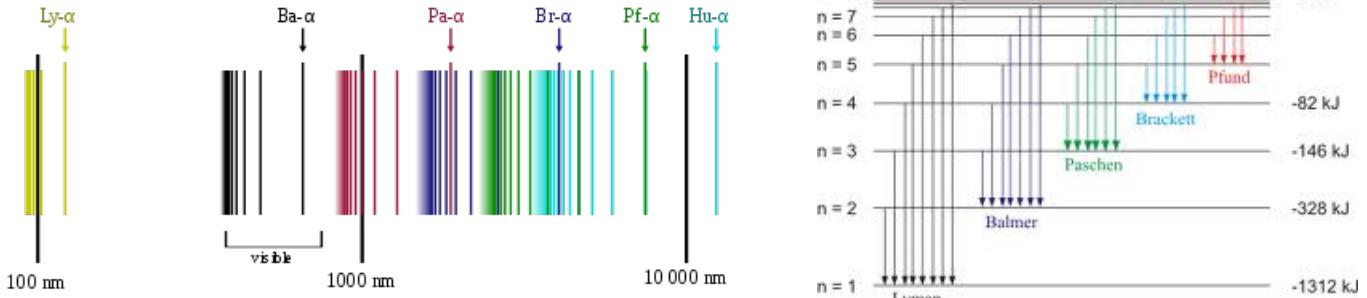


Figure 2.19: Left: Various Spectral Series in the e-m spectrum. Right: The electronic transitions in the Hydrogen atom associated with the respective spectral series. Source: The Internet.

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 mathematical footing.

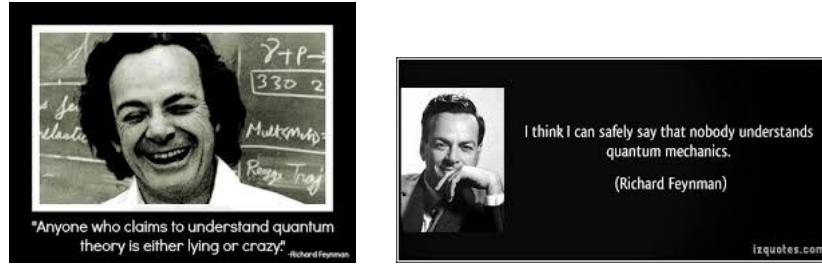


Figure 2.20: Source: The internet.

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:

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

Then, the probability distribution shows signature of quantum interference

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

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

$$P = |\psi_1|^2 + |\psi_2|^2 + 2\psi_1^*\psi_2 \quad (3.1)$$

$$= P_1 + P_2 + 2\sqrt{P_1 P_2} \cos \delta \quad (3.2)$$

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

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 Schrödinger as a function of (\vec{r}, t) .

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

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

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 $\psi(\vec{r}, t)$.

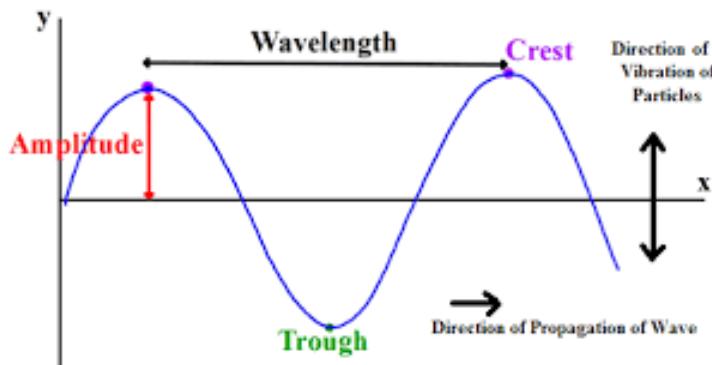


Figure 3.1: A simple waveform in 1D: $y = A \sin\left(\frac{2\pi x}{\lambda}\right)$. Source: The internet.

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)

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

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

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

or even

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

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

$$\frac{d\phi}{dt} = k \frac{dx}{dt} - \omega = 0 \rightarrow \frac{dx}{dt} \equiv v = \frac{\omega}{k}. \quad (3.3)$$

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

$$\frac{d\phi}{dk} = x - \frac{d\omega}{dk}t = 0 \rightarrow \frac{x}{t} \equiv v_g = \frac{d\omega}{dk}, \quad (3.4)$$

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) $E = h\nu = \hbar\omega$ (Planck)
- (ii) $p = \frac{h}{\lambda} = \hbar k$ (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

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

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 $\hbar k_x$

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

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

$$\therefore -i\hbar\vec{\nabla}\psi(\vec{r}, t) = -i\hbar\left(\hat{x}\frac{\partial}{\partial x} + \hat{y}\frac{\partial}{\partial y} + \hat{z}\frac{\partial}{\partial z}\right)\psi(\vec{r}, t) \quad (3.5)$$

$$= \hbar\left(\hat{x}k_x + \hat{y}k_y + \hat{z}k_z\right)\psi(\vec{r}, t) \quad (3.6)$$

$$= \hbar\vec{k}\psi(\vec{r}, t) \quad (3.7)$$

$$\Rightarrow -i\hbar\vec{\nabla}\psi = \vec{p}\psi \quad (3.8)$$

and

$$(-i\hbar\vec{\nabla}\psi) \cdot (-i\hbar\vec{\nabla}\psi) = \vec{p} \cdot \vec{p} \psi \quad (3.9)$$

$$\Rightarrow -\hbar^2\vec{\nabla}^2\psi = |\vec{p}|^2\psi = p^2\psi. \quad (3.10)$$

This gives us

$$\frac{-\hbar^2}{2m}\nabla^2\psi = \frac{p^2}{2m}\psi, \quad (3.11)$$

$$i\hbar\frac{\partial\psi(\vec{r}, t)}{\partial t} = E\psi. \quad (3.12)$$

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

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

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 $V(\vec{r})$.



Hamilton

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



$\Psi(\vec{r}, t)$ 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:

$$|\psi|^2 = \psi^* \psi . \quad (3.14)$$

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,

Born we need an additional requirement called the “Normalisation Condition” on the probability density to determine c

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

As an example, consider the case of a particle in an infinite square well of extent a , with

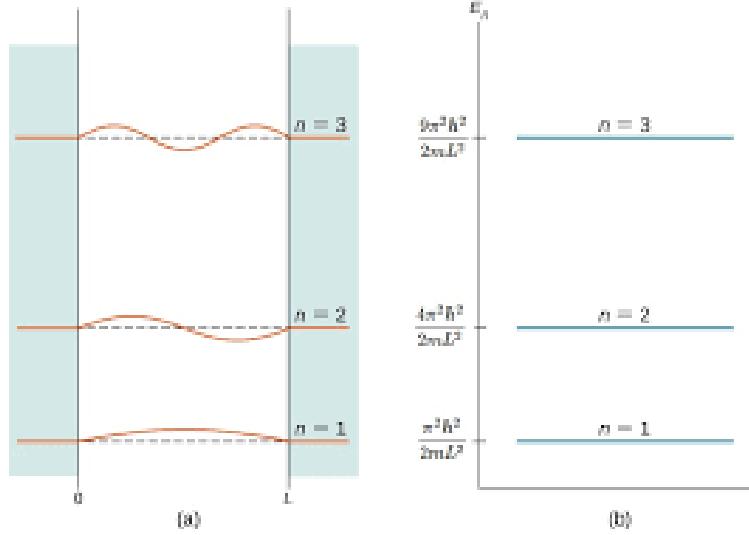


Figure 3.2: Wavefunctions (Left) and Energy Spectrum (Right) for the lowest three eigenstates of the particle in a box problem.

the ground state wavefunction given by

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

From the normalisation condition

$$\begin{aligned} \int_0^a dx |A|^2 \sin^2\left(\frac{\pi x}{a}\right) &= 1 \\ |A|^2 \frac{a}{2} &= 1 \implies A = \sqrt{\frac{2}{a}} \\ \implies \psi &= \sqrt{\frac{2}{a}} \sin\left(\frac{\pi x}{a}\right) e^{\frac{-i\hbar\pi^2}{2ma^2}t} \end{aligned}$$

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.

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

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

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,

OBSERVABLES are associated with LINEAR OPERATORS].

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

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

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

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

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,



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

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

Ehrenfest

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

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.

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

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

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

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

$$[\hat{\theta}_1, \hat{\theta}_2] = \hat{\theta}_1 \hat{\theta}_2 - \hat{\theta}_2 \hat{\theta}_1. \tag{3.21}$$

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 Ehrenfest's theorem as an exercise. Note that the form of the third type of Ehrenfest relation for an operator A will take

$$\frac{d\langle A(t) \rangle}{dt} = \frac{1}{i\hbar} \langle [H, A(t)] \rangle , \quad (3.22)$$

where $A(t) = U_t A U_t^\dagger$ and $U_t = e^{itH/\hbar}$. (Note that we have not assumed any explicit dependence of A on time t . 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.)

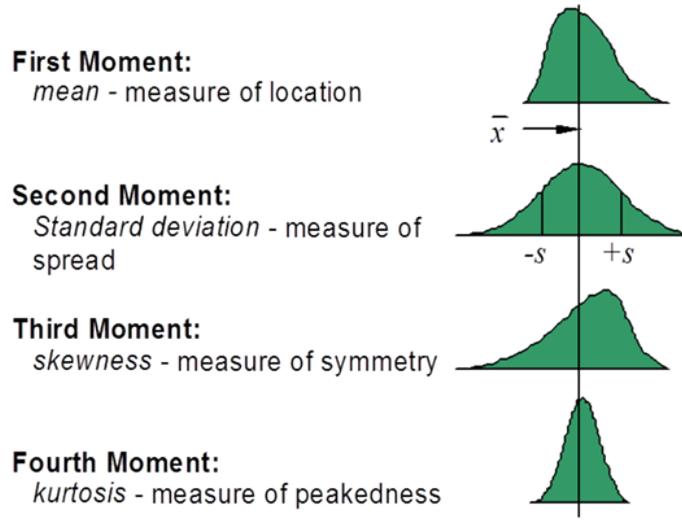


Figure 3.3: First four moments of a distribution. Source: The Internet.

Chapter 4

Formalism I : Operators, Eigenfunctions and Eigenvalues

4.1 Operators, Eigenfunctions and Eigenvalues

4.1.1 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:

$$(i) \quad L[f(x) + g(x)] = L[f(x)] + L[g(x)] \quad (4.1)$$

$$(ii) \quad L[cf(x)] = cL[f(x)] , \quad c \in \mathcal{R} \text{ or } c \in \mathcal{C} . \quad (4.2)$$

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

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

4.1.2 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**

$$L[f(x)] = cf(x), \quad c \in \mathcal{R} \quad \text{and} \quad (4.3)$$

$$\langle L \rangle = \langle L^\dagger \rangle \in \mathcal{R}, \quad (4.4)$$

where the \dagger operation is defined as

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

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

$$\langle A \rangle = \int_V d^3r (\psi_1 \psi_2 \dots \psi_N)^* \begin{pmatrix} A_{11} & A_{12} & \dots & A_{1N} \\ A_{21} & A_{22} & \dots & A_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ A_{N1} & A_{N2} & \dots & A_{NN} \end{pmatrix} \begin{pmatrix} \psi_1 \\ \psi_2 \\ \vdots \\ \psi_N \end{pmatrix}. \quad (4.5)$$

Then, it becomes clear that the \dagger (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.

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

$$(i) \quad \hat{A}\psi = a\psi , \tag{4.6}$$

$$(ii) \quad \hat{B}\psi = b\psi . \tag{4.7}$$

Now from (i)

$$\hat{B}(\hat{A}\psi) = \hat{B}a\psi = a\hat{B}\psi . \tag{4.8}$$

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

$$\hat{B}(\hat{A}\psi) = \hat{A}(\hat{B}\psi) = a(\hat{B}\psi) , \tag{4.9}$$

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, $\psi(x)$:

$$\begin{aligned} [\hat{p}, \hat{x}]\psi(x) &= (-i\hbar \frac{d}{dx})(x\psi(x)) - x(-i\hbar \frac{d}{dx})\psi(x) \\ &= -i\hbar x(\frac{d\psi(x)}{dx}) - i\hbar\psi + i\hbar x(\frac{d\psi(x)}{dx}) \\ &= -i\hbar\psi \\ \Rightarrow [\hat{p}, \hat{x}] &= -i\hbar \quad (\text{dropping the } \psi \text{ from both sides}) , \end{aligned} \tag{4.10}$$

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: $[\hat{O}_1, \hat{O}_2] \propto \hat{O}_3$, 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}),

$$[\hat{H}, \hat{\theta}] = 0 \implies [U, \hat{\theta}] = 0 . \quad (4.11)$$

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

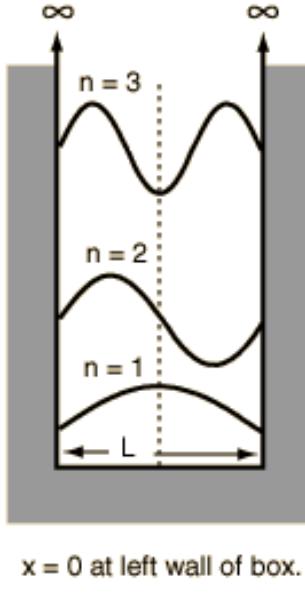
$$T_{\delta x}\psi(x) = e^{-i\hat{p}\delta x/\hbar}\psi(x) = \psi(x + \delta x) , \quad (4.12)$$

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 infinitesimal 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

$$\begin{aligned} \hat{P}\psi_0(x) &= \pm\psi_0(-x) \\ \text{This arises from the fact that } \hat{P}^2\psi_0(x) &= \psi_0(x) \\ \Rightarrow \psi_0(-x) &= \pm\psi_0(x) , \end{aligned} \quad (4.13)$$

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?



$x = 0$ at left wall of box.

Figure 4.1: Wavefunctions for the lowest three eigenstates of the particle in a box problem.
Source: The internet.

4.2 Important properties of Eigenstates

(1) 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

$$\hat{A}\psi_n = a_n\psi_n , \quad n = 1, 2, 3, \dots \quad (4.14)$$

Now, by computing the expectation value

$$\begin{aligned} \langle \hat{A} \rangle &= \int_{-\infty}^{\infty} dx \psi_n^* \hat{A} \psi_n = \int_{-\infty}^{\infty} dx \underbrace{(\hat{A} \psi_n)^*}_{=\psi_n^* \hat{A}^\dagger} \psi_n \\ &= \int_{-\infty}^{\infty} dx \psi_n^* a_n \psi_n = \int_{-\infty}^{\infty} dx (a_n \psi_n)^* \psi_n \\ &= a_n \int_{-\infty}^{\infty} dx \psi_n^* \psi_n = a_n^* \int_{-\infty}^{\infty} dx \psi_n^* \psi_n \\ \langle \hat{A} \rangle &= a_n = a_n^* \\ \text{or, } a_N &\in \mathcal{R} \quad (\text{QED}) \end{aligned} \quad (4.15)$$

(2) Different eigenstates of the same potential are orthogonal.

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

$$\int_{-\infty}^{\infty} dx \psi_n^* \psi_m = \delta_{n,m} = \begin{cases} 1, & n = m \\ 0, & \text{otherwise} \end{cases} \quad (4.16)$$

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 $\hat{A} = \hat{A}^\dagger$,

$$\begin{aligned} \int_{-\infty}^{\infty} dx \psi_m^* \hat{A}\psi_n &= \int_{-\infty}^{\infty} dx \underbrace{(\hat{A}\psi_m)^*}_{=\psi_n^* \hat{A}^\dagger} \psi_n \\ \int_{-\infty}^{\infty} dx \psi_m^* a_n \psi_n &= \int_{-\infty}^{\infty} dx (a_m \psi_m)^* \psi_n \\ a_n \int_{-\infty}^{\infty} dx \psi_m^* \psi_n &= a_m^* \int_{-\infty}^{\infty} dx \psi_m^* \psi_n = a_m \int_{-\infty}^{\infty} dx \psi_m^* \psi_n \quad (\text{as } \hat{A} \text{ is Hermitian}) \\ (a_n - a_m) \int_{-\infty}^{\infty} dx \psi_m^* \psi_n &= 0 . \end{aligned} \quad (4.17)$$

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) 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

$$f(x) = \sum_{n=1} c_n \psi_n(x) . \quad (4.18)$$

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

$$\int_{-\infty}^{\infty} dx \psi_m^* \Psi(x) = \sum_n C_n \int_{-\infty}^{\infty} dx \psi_m^* \psi_n = \sum_n C_n \delta_{n,m} = C_m , \quad (4.19)$$

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

$$\Psi(x, t) = \sum_n C_n \psi_n(x) e^{-iE_n t / \hbar} . \quad (4.20)$$

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 possibilities 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)?

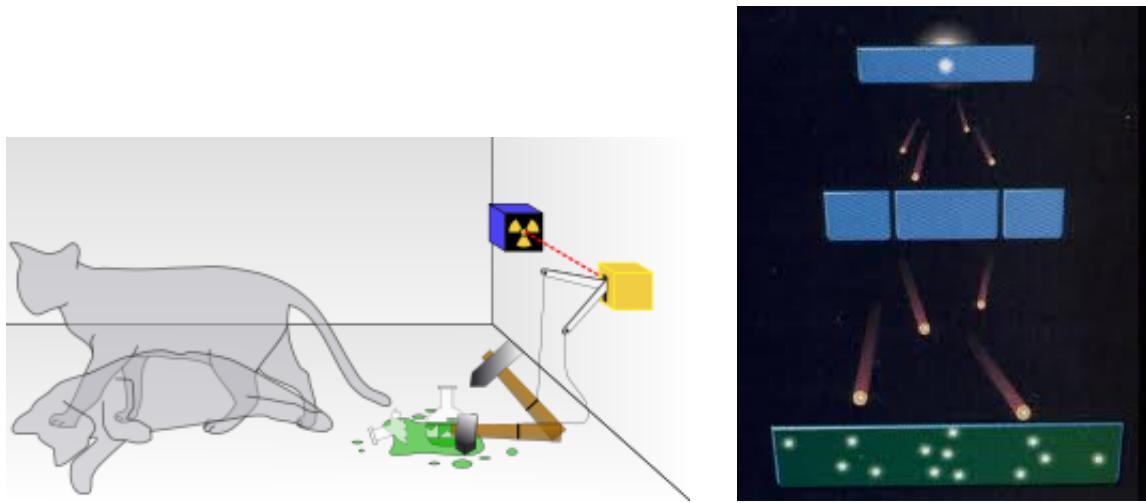


Figure 4.2: (Left) The Schrödinger Cat gedanken. (Right) Double-slit Interference Experiment with Matter particles. Source: The internet.

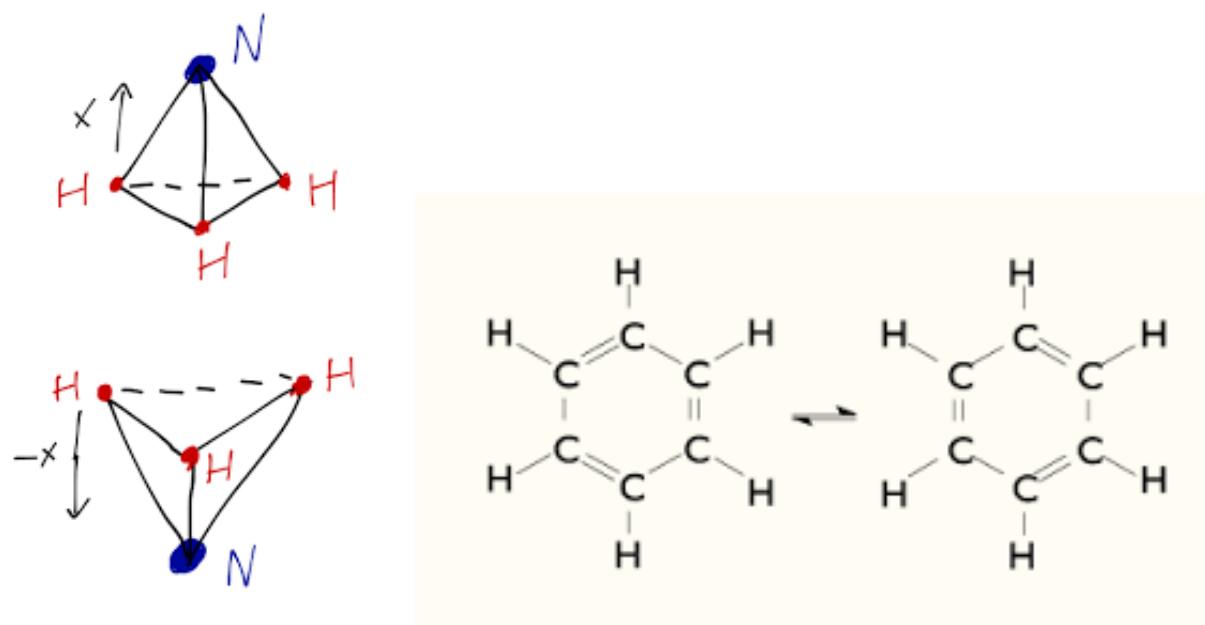


Figure 4.3: (Left) The two configurations of the Ammonia molecule. (Right) The two configurations of the Benzene molecule. Source: The internet.

Chapter 5

Formalism II : Vector Spaces in Quantum Mechanics

5.1 Vector Spaces

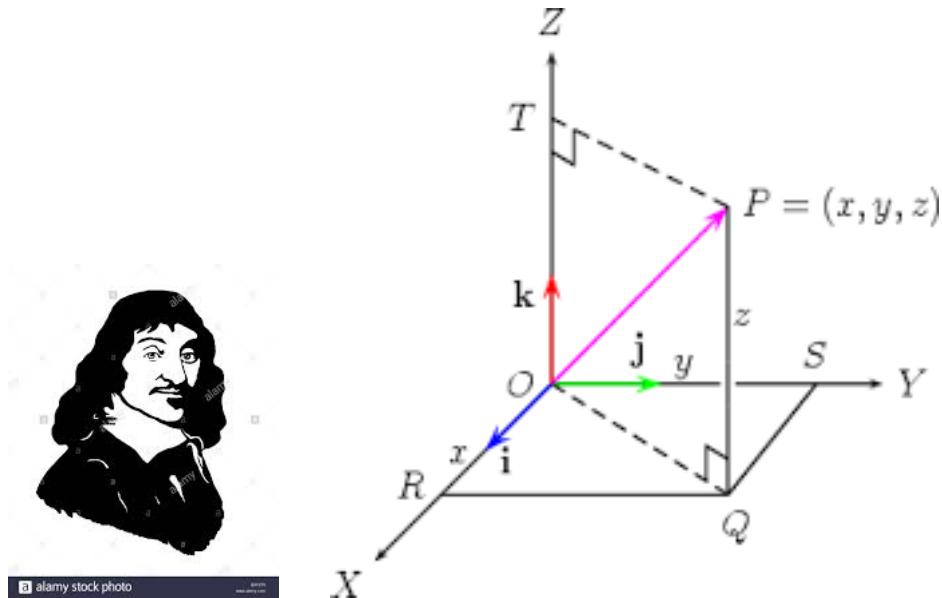


Figure 5.1: Left: Renée Descartes. Right: A vector in Cartesian coordinates. Source: The Internet.

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

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

$$\vec{u} + \vec{v} = (u_1 + v_1, u_2 + v_2, u_3 + v_3, \dots, u_n + v_n) = \vec{w}, \quad (5.1)$$

as well as

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

$$\vec{r} = c\vec{u} , \quad (5.2)$$

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

$$\vec{u} \cdot \vec{v} = u_1 v_1 + u_2 v_2 + \dots + u_n v_n . \quad (5.3)$$

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.

5.1.1 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 $(|\phi_1\rangle, |\phi_2\rangle, \dots, |\phi_N\rangle)$ such that any vector $|\psi\rangle$ can be written as

$$|\psi\rangle = \sum_{i=1}^N a_i |\phi_i\rangle , \quad (5.4)$$

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

$$a_i = \langle\psi|\phi_i\rangle , \quad (5.5)$$

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

$$\langle\phi_i|\phi_j\rangle = \delta_{ij} , \quad (5.6)$$

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

5.1.2 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, $|\psi\rangle$ and $|\phi\rangle$, of an abstract vector space to be a function which uses $|\psi\rangle$ and $|\phi\rangle$ as inputs and whose result is (in general) a complex number

$$\langle\psi|\phi\rangle = c, \quad c \in \mathcal{C}. \quad (5.7)$$

The Inner Product has the following properties:

- (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) $\langle\psi|\psi\rangle \geq 0$.

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

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

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

$$\langle f(\vec{r})|g(\vec{r})\rangle = \int_{-\infty}^{\infty} dx dy dz f(\vec{r})^* g(\vec{r}). \quad (5.8)$$

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

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

as the argument $x \sin^2 x$ is an odd function in x .

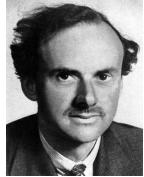
Finally, the requirement that the wavefunction $\psi(x)$ be normalised can now be recast as

$$\langle\psi|\psi\rangle = 1, \quad (5.9)$$

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

$$\langle\hat{\theta}\rangle = \langle\psi|\hat{\theta}|\psi\rangle \equiv \langle\psi|\hat{\theta}\psi\rangle. \quad (5.10)$$

5.2 The Dirac Bra-Ket Notation



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

$$\langle \psi | \psi \rangle = \int d^3x |\psi(r)|^2 = 1 , \quad (5.11)$$

i.e., the ψ s 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.



Hilbert

Dirac

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.

5.3 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:

$$|\psi\rangle \leftrightarrow \langle \psi | \quad (5.12)$$

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

$$a|\psi\rangle + b|\phi\rangle \leftrightarrow a^* \langle \psi | + b^* \langle \phi | \quad a, b \in \mathcal{C} \quad (5.13)$$

$$|a\psi\rangle = a|\psi\rangle \quad \& \quad \langle a\psi | = a^* \langle \psi | . \quad (5.14)$$

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,

$$\begin{aligned} \langle \phi | \psi \rangle^* &= (\int d\vec{r} \phi^*(\vec{r}, t) \psi(\vec{r}, t))^* = \int d\vec{r} \psi^*(\vec{r}, t) \phi(\vec{r}, t) = \langle \psi | \phi \rangle \\ \text{i.e., } \langle \phi | \psi \rangle^* &= \langle \psi | \phi \rangle . \end{aligned} \quad (5.15)$$

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,

$$\langle \psi | a_1 \psi_1 + a_2 \psi_2 \rangle = a_1 \langle \psi | \psi_1 \rangle + a_2 \langle \psi | \psi_2 \rangle \quad (5.16)$$

$$\langle a_1 \phi_1 + a_2 \phi_2 | \psi \rangle = a_1^* \langle \phi_1 | \psi \rangle + a_2^* \langle \phi_2 | \psi \rangle \quad (5.17)$$

$$\langle a_1 \phi_1 + a_2 \phi_2 | b_1 \psi_1 + b_2 \psi_2 \rangle = a_1^* b_1 \langle \phi_1 | \psi_1 \rangle + a_1^* b_2 \langle \phi_1 | \psi_2 \rangle + a_2^* b_1 \langle \phi_2 | \psi_1 \rangle + a_2^* b_2 \langle \phi_2 | \psi_2 \rangle . \quad (5.18)$$

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

$$\langle \psi | \psi \rangle \geq 0 , \quad \langle \psi | \psi \rangle \in \mathcal{R} . \quad (5.19)$$

Further, if 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 $|\psi\rangle$.

5.4 Operators

1. Operators are mathematical rules such that

$$\hat{A}|\psi\rangle = |\psi'\rangle , \quad \langle \phi | \hat{A} = \langle \phi' | . \quad (5.20)$$

2. Products of Operators: In general, the product

$$[\hat{A}, \hat{B}] = \hat{A}\hat{B} - \hat{B}\hat{A} \neq 0 \quad (5.21)$$

$$\Rightarrow \hat{A}\hat{B} \neq \hat{B}\hat{A}. \quad (5.22)$$

$$\text{Further, } \hat{A}\hat{B}\hat{C} = \hat{A}(\hat{B}\hat{C}) = (\hat{A}\hat{B})\hat{C} \quad (5.23)$$

$$\hat{A}^m\hat{A}^n = \hat{A}^{m+n} \quad (5.24)$$

$$\hat{A}\hat{B}\hat{C}|\psi\rangle = \hat{A}\hat{B}(\hat{C}|\psi\rangle) = \hat{A}\{\hat{B}(\hat{C}|\psi\rangle)\}. \quad (5.25)$$

3. Linearity:

$$\hat{A}(a_1|\psi_1\rangle + a_2|\psi_2\rangle) = a_1\hat{A}|\psi_1\rangle + a_2\hat{A}|\psi_2\rangle \quad (5.26)$$

$$(a_1\langle\psi_1| + a_2\langle\psi_2|)\hat{A} = a_1\langle\psi_1|\hat{A} + a_2\langle\psi_2|\hat{A}. \quad (5.27)$$

4. Expectation Value:

$$\langle\hat{A}\rangle = \frac{\langle\psi|\hat{A}|\psi\rangle}{\langle\psi|\psi\rangle} \equiv \langle\psi|\hat{A}|\psi\rangle \quad (\text{if } |\psi\rangle \text{ is normalised}) \quad (5.28)$$

$$= \langle\psi|\hat{A}\psi\rangle. \quad (5.29)$$

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

$$|\phi\rangle\langle\psi|\psi'\rangle = \langle\psi|\psi'\rangle|\phi\rangle = c|\phi\rangle, \quad (5.30)$$

$$|\psi\rangle\langle\psi|\psi'\rangle = \langle\psi|\psi'\rangle|\psi\rangle = c|\psi\rangle, \quad (5.31)$$

where $c \equiv \langle\psi|\psi'\rangle \in \mathcal{C}$ 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

$$\langle\phi|\hat{\theta}\psi\rangle = \langle\hat{\theta}^\dagger\phi|\psi\rangle. \quad (5.32)$$

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

$$\langle\phi|\hat{D}\psi\rangle = \int_{-\infty}^{\infty} dx \phi^*(x) \frac{d\psi(x)}{dx} \quad (\text{in position basis}) \quad (5.33)$$

$$= [\phi^*(x)\psi(x)]_{-\infty}^{\infty} - \int_{-\infty}^{\infty} dx \frac{d\phi^*(x)}{dx} \psi(x) \quad (5.34)$$

$$= - \int_{-\infty}^{\infty} dx \frac{d\phi^*(x)}{dx} \psi(x) \quad (\text{as } \phi(x), \psi(x) \rightarrow 0 \text{ as } x \rightarrow \pm\infty) \quad (5.35)$$

$$\Rightarrow \langle\phi|\hat{D}\psi\rangle = \int_{-\infty}^{\infty} (-\hat{D}\phi^*(x))\psi(x) \quad (5.36)$$

$$= \langle -\hat{D}\phi|\psi\rangle \equiv \langle \hat{D}^\dagger\phi|\psi\rangle \quad (5.37)$$

$$\Rightarrow \hat{D}^\dagger = -\hat{D}. \quad (5.38)$$

Adjoint Operators have the following properties:

$$(i) \quad (c\hat{P})^\dagger = c^*\hat{P}^\dagger \quad (c \in \mathcal{C}) \quad (5.39)$$

$$(ii) \quad (\hat{P} + \hat{Q})^\dagger = \hat{P}^\dagger + \hat{Q}^\dagger \quad (5.40)$$

$$(iii) \quad (\hat{P}\hat{Q})^\dagger = \hat{Q}^\dagger\hat{P}^\dagger \quad (5.41)$$

$$(iv) \quad (\hat{P}^\dagger)^\dagger = \hat{P} . \quad (5.42)$$

Relation (iii) can also be seen as follows:

$$\langle \phi | \hat{P}(\hat{Q}\psi) \rangle = \langle \hat{P}^\dagger \phi | \hat{Q}\psi \rangle \quad (5.43)$$

$$= \langle \hat{Q}^\dagger \hat{P}^\dagger \phi | \psi \rangle . \quad (5.44)$$

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:

$$\hat{p}^\dagger = -\hat{p}^* = -(-i\hbar \frac{d}{dx})^* = \hat{p} . \quad (5.45)$$

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

$$\langle \hat{Q} \rangle = \langle \psi | \hat{Q} \psi \rangle \quad (5.46)$$

$$= \langle \psi | \hat{Q}^\dagger \psi \rangle \quad (\text{for } \hat{Q} = \hat{Q}^\dagger) \quad (5.47)$$

$$= \langle \hat{Q}\psi | \psi \rangle \quad (5.48)$$

$$= \langle \hat{Q} \rangle^* , \quad (5.49)$$

$$\Rightarrow \langle \hat{Q} \rangle \in \mathcal{R} . \quad (5.50)$$

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

$$\langle \psi | \hat{Q} \psi \rangle = q \langle \psi | \psi \rangle = q \quad (5.51)$$

$$\text{But } \langle \psi | \hat{Q} \psi \rangle = \langle \hat{Q}^\dagger \psi | \psi \rangle \quad (5.52)$$

$$= \langle \hat{Q}\psi | \psi \rangle = q^* \langle \psi | \psi \rangle = q^* \quad (5.53)$$

$$\Rightarrow q \in \mathcal{R} . \quad (5.54)$$

Finally, writing

$$\langle \hat{Q} \rangle = \langle \psi | \hat{Q} | \psi \rangle \quad (\text{where the states } |\psi\rangle \text{ are not eigenstates of } \hat{Q}) \quad (5.55)$$

$$= \sum_{m,n} \langle \psi | \xi \rangle_{mm} \langle \xi | \hat{Q} | \xi \rangle_{nn} \langle \xi | \psi \rangle \quad (\text{where we've introduced the identity operator}) \quad (5.56)$$

$$I = \sum_n |\xi\rangle_{nn} \langle \xi| \quad \text{in terms of eigenstates of } \hat{Q}, \quad |\xi\rangle_n \quad (5.57)$$

$$= \sum_{m,n} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} dx dy \xi_m(x) \delta_{m,n} Q_n \xi_n^*(y) \quad (\text{as } \hat{Q}|\xi\rangle_n = Q_n|\xi\rangle_n) \quad (5.58)$$

where we have used $\langle \psi | \xi \rangle_m = \xi_m(x)$, $\langle \xi | \psi \rangle = \xi_n^*(y)$ and $\langle \xi | \hat{Q} | \xi \rangle_n = Q_n \langle \xi | \xi \rangle_n = Q_n \delta_{mn}$ in the last line.

5.5 Discrete & Continuous Basis

We have noted earlier that the quantum world is unlike the classical world: in the latter, all observables can take a continuum of values whereas in the former one can find examples of observables whose spectra can either be continuous or discrete. We have stressed considerably the quantisation phenomena of several observables that arises in many quantum systems in earlier chapters. Most generally, an identical form of the eigenvalue equation is obtained for operators \hat{A} and $\hat{\Gamma}$ acting on their eigenstates $|A\rangle$ and $|\psi\rangle$ so to obtain eigenvalues A and Γ that are discrete and continuous respectively

$$\hat{A}|A\rangle = A|A\rangle \leftrightarrow \hat{\Gamma}|\psi\rangle = \Gamma|\psi\rangle . \quad (5.59)$$

The inner product structure for the discrete ($|A\rangle$) and continuous ($|\psi\rangle$) states is, however, slightly different so as to signify their nature

$$(\text{discrete}) \quad \langle A|A' \rangle = \delta_{AA'} \leftrightarrow \langle \psi|\psi' \rangle = \delta(\psi - \psi') \quad (\text{continuous}) , \quad (5.60)$$

$$(\text{discrete}) \quad \sum_A |A\rangle \langle A| = 1 \leftrightarrow \int d\psi |\psi\rangle \langle \psi| = 1 \quad (\text{continuous}) , \quad (5.61)$$

$$(\text{discrete}) \quad |B\rangle = \sum_A |A\rangle \langle A|B\rangle \leftrightarrow |\phi\rangle = \int d\psi |\psi\rangle \langle \psi|\phi\rangle \quad (\text{continuous}) , \quad (5.62)$$

$$(\text{discrete}) \quad \langle B|B \rangle = \sum_A |\langle A|B \rangle|^2 = 1 \leftrightarrow \int d\psi |\langle \psi|\phi \rangle|^2 = 1 \quad (\text{continuous}) , \quad (5.63)$$

$$(\text{discrete}) \quad \langle C|B \rangle = \sum_A \langle C|A \rangle \langle A|B \rangle \leftrightarrow \langle \xi|\phi \rangle = \int d\psi \langle \xi|\psi \rangle \langle \psi|\phi \rangle \quad (\text{continuous}) \quad (5.64)$$

$$(\text{discrete}) \quad \langle A'|\hat{A}|A \rangle = A\delta_{AA'} \leftrightarrow \langle \psi'|\hat{\Gamma}|\psi \rangle = \Gamma\delta(\psi - \psi') \quad (\text{continuous}) \quad (5.65)$$

where we have used eq.(5.61) to obtain eq.(5.62) and eq.(5.64). Please note that the Dirac delta-functions $\delta(\psi - \psi')$ are mathematically well defined objects on the abstract Hilbert space $\{|\psi\rangle\}$. They are, however, better understood by projecting them onto some concrete basis related to some observable (such as position $\{|x\rangle\}$, momentum $\{|p\rangle\}$ etc.) in terms of which the Dirac delta-functions transform into their counterparts in the position (x), momentum (p) etc. variables.

5.6 Position and Momentum Basis

The eigenstates of the position operator \hat{x} for a one-dimensional quantum system satisfy

$$\hat{x}|x\rangle = x|x\rangle , \quad \langle x'|x\rangle = \delta(x' - x) , \quad (5.66)$$

where x is the eigenvalue obtained from the action of \hat{x} on $|x\rangle$, and has dimensions of length. An arbitrary physical state $|\alpha\rangle$ can always be written in terms of the set of eigenstates $\{|x\rangle\}$

$$|\alpha\rangle = \int_{-\infty}^{\infty} dx |x\rangle \langle x|\alpha\rangle = \int_{-\infty}^{\infty} dx \psi_{\alpha}(x) |x\rangle , \quad \psi_{\alpha}(x) = \langle x|\alpha\rangle , \quad (5.67)$$

$$\langle\alpha|\alpha\rangle = 1 \implies \int_{-\infty}^{\infty} dx \langle\alpha|x\rangle \langle x|\alpha\rangle = \int_{-\infty}^{\infty} dx |\langle\alpha|x\rangle|^2 = \int_{-\infty}^{\infty} dx |\psi_{\alpha}(x)|^2 = 1 , \quad (5.68)$$

$$\langle\beta|\alpha\rangle = \int_{-\infty}^{\infty} dx \langle\beta|x\rangle \langle x|\alpha\rangle = \int_{-\infty}^{\infty} dx \psi_{\beta}^{*}(x) \psi_{\alpha}(x) , \quad (5.69)$$

and where $\psi_{\alpha}(x) \equiv \langle x|\alpha\rangle$ is the real-space wavefunction for the physical state represented by $|\alpha\rangle$.

Extending this to three dimensions is straightforward. We now assume that the set of position eigenstates $|\vec{r}\rangle \equiv |x, y, z\rangle$ is a complete basis (i.e., where $\vec{r} = (x, y, z)$), such that we can write the physical state $|\psi\rangle$ in the real-space basis $\{|\vec{r}\rangle\}$ as

$$|\alpha\rangle = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} d^3x |\vec{r}\rangle \langle\vec{r}|\alpha\rangle . \quad (5.70)$$

Note that we are assuming that $|\vec{r}\rangle \equiv |x, y, z\rangle$ is a simultaneous eigenstate of the observables \hat{x}, \hat{y} and \hat{z}

$$\hat{x}|\vec{r}\rangle = x|\vec{r}\rangle , \quad \hat{y}|\vec{r}\rangle = y|\vec{r}\rangle , \quad \hat{z}|\vec{r}\rangle = z|\vec{r}\rangle . \quad (5.71)$$

Naturally, this means that the operators \hat{x}_i ($i = 1, 2, 3$) corresponds to the operators $(\hat{x}, \hat{y}, \hat{z})$ satisfy the following trivial commutation relations

$$[\hat{x}_i, \hat{x}_j] = 0 . \quad (5.72)$$

Starting from some arbitrary state $|\alpha\rangle$ that has been expanded in some basis $\{|A\rangle\}$ other than the position,

$$|\alpha\rangle = \sum_A |A\rangle \langle A|\alpha\rangle = \sum_A c_A |A\rangle , \quad \text{where } c_A = \langle A|\alpha\rangle \quad (5.73)$$

$$\langle x|\alpha\rangle = \sum_A \langle x|A\rangle \langle A|\alpha\rangle = \sum_A c_A \langle x|A\rangle ,$$

$$\implies \psi_{\alpha}(x) = \sum_A c_A \psi_A(x) , \quad \text{where } \psi_A(x) = \langle x|A\rangle . \quad (5.74)$$

Computing matrix elements of an operator \hat{A} is also found to be

$$\begin{aligned} \langle\beta|\hat{A}|\alpha\rangle &= \int_{-\infty}^{\infty} dx' \int_{-\infty}^{\infty} dx \langle\beta|x'\rangle \langle x'|\hat{A}|x\rangle \langle x|\alpha\rangle , \\ &= \int_{-\infty}^{\infty} dx' \int_{-\infty}^{\infty} dx \psi_{\beta}^{*}(x') \langle x'|\hat{A}|x\rangle \psi_{\alpha}(x) , \end{aligned} \quad (5.75)$$

such that computing the matrix element $\langle\beta|\hat{A}|\alpha\rangle$ requires the knowledge of the matrix element $\langle x'|\hat{A}|x\rangle$ (which is a function of x and x'). For the operator $\hat{A} \equiv \hat{f}(\hat{x})$ (where $\hat{f}(\hat{x})$

corresponds to an operator that is some function of \hat{x} , $f(\hat{x})$), we have

$$\begin{aligned}\langle x' | \hat{A} | x \rangle &= \langle x' | \hat{f}(\hat{x}) | x \rangle = f(x) \langle x' | x \rangle = f(x) \delta(x' - x) , \\ \implies \langle \beta | \hat{A} | \alpha \rangle &= \langle \beta | \hat{f}(\hat{x}) | \alpha \rangle ,\end{aligned}\quad (5.76)$$

$$\begin{aligned}&= \int_{-\infty}^{\infty} dx' \int_{-\infty}^{\infty} dx \langle \beta | x' \rangle \langle x' | \hat{f}(\hat{x}) | x \rangle \langle x | \alpha \rangle , \\ &= \int_{-\infty}^{\infty} dx' \int_{-\infty}^{\infty} dx \psi_{\beta}^{*}(x') \langle x' | \hat{f}(\hat{x}) | x \rangle \psi_{\alpha}(x) , \\ &= \int_{-\infty}^{\infty} dx' \int_{-\infty}^{\infty} dx \psi_{\beta}^{*}(x') f(x) \delta(x' - x) \psi_{\alpha}(x) , \\ &= \int_{-\infty}^{\infty} dx \psi_{\beta}^{*}(x) f(x) \psi_{\alpha}(x) .\end{aligned}\quad (5.77)$$

The momentum, or p , basis is defined in the same way as observed above for the position (or x) basis

$$\hat{p} | p \rangle = p | p \rangle , \quad \langle p' | p \rangle = \delta(p' - p) , \quad (5.78)$$

where p is the eigenvalue obtained from the action of \hat{p} on $|p\rangle$, and has dimensions of momentum. An arbitrary physical state $|\alpha\rangle$ can always be written in terms of the set of eigenstates $\{|p\rangle\}$

$$|\alpha\rangle = \int_{-\infty}^{\infty} dp |p\rangle \langle p | \alpha \rangle = \int_{-\infty}^{\infty} dp \psi_{\alpha}(p) |p\rangle , \quad \psi_{\alpha}(p) = \langle p | \alpha \rangle , \quad (5.79)$$

$$\langle \alpha | \alpha \rangle = 1 \implies \int_{-\infty}^{\infty} dp \langle \alpha | p \rangle \langle p | \alpha \rangle = \int_{-\infty}^{\infty} dx |\langle \alpha | p \rangle|^2 = \int_{-\infty}^{\infty} dx |\psi_{\alpha}(p)|^2 = 1 , \quad (5.80)$$

$$\langle \beta | \alpha \rangle = \int_{-\infty}^{\infty} dp \langle \beta | p \rangle \langle p | \alpha \rangle = \int_{-\infty}^{\infty} dp \psi_{\beta}^{*}(p) \psi_{\alpha}(p) , \quad (5.81)$$

and where $\psi_{\alpha}(p) \equiv \langle p | \alpha \rangle$ is the momentum-space wavefunction for the physical state represented by $|\alpha\rangle$.

For the operator $\hat{A} \equiv \hat{f}(\hat{p})$ (where $\hat{f}(\hat{p})$ corresponds to an operator that is some function of \hat{p} , $f(\hat{p})$), we have

$$\begin{aligned}\langle p' | \hat{A} | p \rangle &= \langle p' | \hat{f}(\hat{p}) | p \rangle = f(p) \langle p' | p \rangle = f(p) \delta(p' - p) , \\ \implies \langle \beta | \hat{A} | \alpha \rangle &= \langle \beta | \hat{f}(\hat{p}) | \alpha \rangle ,\end{aligned}\quad (5.82)$$

$$\begin{aligned}&= \int_{-\infty}^{\infty} dp' \int_{-\infty}^{\infty} dp \langle \beta | p' \rangle \langle p' | \hat{f}(\hat{p}) | p \rangle \langle p | \alpha \rangle , \\ &= \int_{-\infty}^{\infty} dp' \int_{-\infty}^{\infty} dp \psi_{\beta}^{*}(p') \langle p' | \hat{f}(\hat{p}) | p \rangle \psi_{\alpha}(p) , \\ &= \int_{-\infty}^{\infty} dp' \int_{-\infty}^{\infty} dp \psi_{\beta}^{*}(p') f(p) \delta(p' - p) \psi_{\alpha}(p) , \\ &= \int_{-\infty}^{\infty} dp \psi_{\beta}^{*}(p) f(p) \psi_{\alpha}(p) .\end{aligned}\quad (5.83)$$

5.6.1 Momentum operator in Position Basis

Now, using the fact that the momentum operator \hat{p}_x in the position representation $\hat{p}_x \equiv -i\hbar \frac{\partial}{\partial x}$, we obtain

$$\begin{aligned}\hat{p}|\alpha\rangle &= \hat{p} \int_{-\infty}^{\infty} dx |x\rangle \langle x|\alpha\rangle = \hat{p} \int_{-\infty}^{\infty} dx \psi_{\alpha}(x) |x\rangle , \quad \psi_{\alpha}(x) = \langle x|\alpha\rangle , \\ &= \int_{-\infty}^{\infty} dx (\hat{p}\psi_{\alpha}(x)) |x\rangle , \\ &= \int_{-\infty}^{\infty} dx \left(-i\hbar \frac{\partial}{\partial x} \langle x|\alpha\rangle \right) |x\rangle ,\end{aligned}\tag{5.84}$$

$$\implies \langle x|\hat{p}|\alpha\rangle = -i\hbar \frac{\partial}{\partial x} \langle x|\alpha\rangle = -i\hbar \frac{\partial}{\partial x} \psi_{\alpha}(x) ,\tag{5.85}$$

using the fact that $\langle x'|x\rangle = \delta(x' - x)$. In general, the matrix element for \hat{p} in the x representation is given by

$$\langle x'|\hat{p}|x\rangle = -i\hbar \frac{\partial}{\partial x} \delta(x' - x) ,\tag{5.86}$$

and

$$\begin{aligned}\langle \beta|\hat{p}|\alpha\rangle &= \int_{-\infty}^{\infty} dx \langle \beta|x\rangle \left(-i\hbar \frac{\partial}{\partial x} \langle x|\alpha\rangle \right) , \\ &= \int_{-\infty}^{\infty} dx \psi_{\beta}^{*}(x) \left(-i\hbar \frac{\partial}{\partial x} \right) \psi_{\alpha}(x) .\end{aligned}\tag{5.87}$$

We can now also obtain the wavefunction for the momentum eigenstate in the position representation, using

$$\langle x|\hat{p}|p\rangle = -i\hbar \frac{\partial}{\partial x} \langle x|p\rangle ,\tag{5.88}$$

$$\implies p \langle x|p\rangle = -i\hbar \frac{\partial}{\partial x} \langle x|p\rangle ,\tag{5.89}$$

$$\implies \langle x|p\rangle = \mathcal{N} e^{\frac{ipx}{\hbar}} ,\tag{5.90}$$

where the normalisation factor \mathcal{N} is obtained from

$$\langle x'|x\rangle = \delta(x' - x) = \int_{-\infty}^{\infty} dp \langle x'|p\rangle \langle p|x\rangle ,\tag{5.91}$$

$$\begin{aligned}&= |\mathcal{N}|^2 \int_{-\infty}^{\infty} dp e^{\frac{ip(x'-x)}{\hbar}} , \\ &= 2\pi\hbar |\mathcal{N}|^2 \delta(x' - x) , \\ \implies \mathcal{N} &= \frac{1}{\sqrt{2\pi\hbar}} ,\end{aligned}\tag{5.92}$$

where we have used the relation for the Dirac delta function

$$\delta(x' - x) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} dp e^{\frac{ip(x'-x)}{\hbar}} .\tag{5.93}$$

This gives us the so-called “transformation function” $\langle x|p\rangle$ as

$$\langle x|p\rangle = \frac{1}{\sqrt{2\pi\hbar}} e^{\frac{ipx}{\hbar}} , \quad (5.94)$$

such that we can write the position-space ($\psi_\alpha(x)$) and momentum-space ($\psi_\alpha(p)$) wavefunctions for a general state $|\alpha\rangle$ as

$$\langle x|\alpha\rangle = \int_{-\infty}^{\infty} dp \langle x|p\rangle \langle p|\alpha\rangle , \quad (5.95)$$

$$\text{or, } \psi_\alpha(x) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} dpe^{\frac{ipx}{\hbar}} \psi_\alpha(p) , \quad (5.96)$$

$$\langle p|\alpha\rangle = \int_{-\infty}^{\infty} dx \langle p|x\rangle \langle x|\alpha\rangle , \quad (5.97)$$

$$\text{or, } \psi_\alpha(p) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} dxe^{-\frac{ipx}{\hbar}} \psi_\alpha(x) . \quad (5.98)$$

Indeed, we can see that $\psi_\alpha(x)$ and $\psi_\alpha(p)$ are connected to one another through a Fourier transformation.

5.7 Change of Basis

Very generally, we can thinking of a change in basis (or representation) as a “generalised rotation of the old basis eigenfunctions into the new ones” through the application of a unitary operator U_Θ

$$|\phi_1\rangle = U_\Theta |\psi_1\rangle , \dots , |\phi_N\rangle = U_\Theta |\psi_N\rangle , \quad (5.99)$$

where Θ refers to the operator that acts as the “generator” of the transformation by an amount ϵ : $U_\Theta = e^{-i\epsilon\Theta/\hbar}$. But why unitary operators? The answer lies in recognising the fact that transformations such as U_Θ which act on members of the Hilbert space $U_\Theta |\psi\rangle = |\psi'\rangle$ must preserve all inner products between them

$$\begin{aligned} \langle\phi|\psi\rangle \rightarrow \langle\phi'|\psi'\rangle &= (\langle\phi|U_\Theta^\dagger)(U_\Theta|\psi\rangle) = \langle\phi|\psi\rangle , \\ \implies U_\Theta^\dagger U_\Theta &= 1 , \\ \implies U_\Theta^\dagger &= U_\Theta^{-1} , \end{aligned} \quad (5.100)$$

indicating that U_Θ is a unitary operator. Indeed, transformation operators in quantum mechanics must be either unitary or anti-unitary (i.e., an operator U whose action on ψ leads to $U(c|\psi\rangle) = c^*U|\psi\rangle$, e.g., the time-reversal transformation operation).

Here are a few more interesting properties of unitary operators. First, for a Hermitian operator Θ , we can always form a unitary operator $U_\Theta(\lambda) = e^{-i\lambda\Theta}$ where λ is real-valued:

$$U_\Theta^\dagger = e^{i\lambda\Theta} , \quad U_\Theta^\dagger U_\Theta = 1 . \quad (5.101)$$

Further, the eigenvectors of Θ correspond to the eigenvectors of U_Θ , and the eigenvalues of Θ are related to those of U_Θ :

$$\begin{aligned}\hat{\Theta}|\psi\rangle &= \Theta|\psi\rangle , \\ U_\Theta|\psi\rangle &= e^{-i\lambda\hat{\Theta}}|\psi\rangle = e^{-i\lambda\Theta}|\psi\rangle ,\end{aligned}\quad (5.102)$$

where the eigenvalues of U_Θ are phases given by $e^{-i\lambda\Theta}$.

An important relation that will often prove useful is the following. For two Hermitian operators Θ and A , such that the unitary operator $U_\Theta = e^{-i\lambda\Theta}$, we have the Baker-Campbell-Hausdorff relation (which we will not prove here)

$$\begin{aligned}U_\Theta^\dagger A U_\Theta &= e^{i\lambda\Theta} A e^{-i\lambda\Theta} , \\ &= A + i\lambda[\Theta, A] + \frac{(i\lambda)^2}{2!}[\Theta, [\Theta, A]] + \dots + \frac{(i\lambda)^n}{n!}[\Theta, [\Theta, [\Theta, \dots [\Theta, A]]] \dots]\end{aligned}\quad (5.103)$$

For the special case of $[\Theta, A] = 0$, all the commutators on the right hand side vanish identically, and the above simplifies to

$$U_\Theta^\dagger A U_\Theta = e^{i\lambda\Theta} A e^{-i\lambda\Theta} = A . \quad (5.104)$$

As we will now see, this relation indicates that, for an operator \hat{A} acting on its eigenbasis $\{|A\rangle\}$, there also exists an operator $\hat{B} \equiv U_\Theta \hat{A} U_\Theta^\dagger$ (i.e., $\hat{A} = U_\Theta^\dagger \hat{B} U_\Theta$) acting on its eigenbasis $\{|B\rangle\} \equiv \{U_\Theta |A\rangle\}$ with an identical set of eigenvalues

$$\begin{aligned}\hat{A}|A\rangle &= A|A\rangle , \\ \implies U_\Theta \hat{A} &= U_\Theta A|A\rangle , \\ \text{or, } U_\Theta \hat{A} U_\Theta^\dagger U_\Theta |A\rangle &= AU_\Theta|A\rangle , \\ \text{or, } \hat{B}|B\rangle &= A|B\rangle ,\end{aligned}\quad (5.105)$$

where we have identified \hat{B} and $|B\rangle$ above. Thus, we say that \hat{A} and $\hat{B} \equiv U_\Theta \hat{A} U_\Theta^\dagger$ are unitarily equivalent observables with identical spectra (i.e., the set of eigenvalues $\{A\}$).

Now, using the fact that the matrix elements of the U_Θ transformation matrix between the two basis sets are given by $\langle A|B\rangle = \langle A|U_\Theta|A\rangle$, we can expand a given state $|\alpha\rangle$ as

$$|\alpha\rangle = \sum_A |A\rangle \langle A|\alpha\rangle , \quad (5.106)$$

$$\implies \langle B|\alpha\rangle = \sum_A \langle B|A\rangle \langle A|\alpha\rangle , \quad (5.107)$$

$$= \sum_A \langle A|U_\Theta^\dagger|A\rangle \langle A|\alpha\rangle = \sum_A \langle A|U_\Theta^{-1}|A\rangle \langle A|\alpha\rangle . \quad (5.108)$$

Similarly, we can relate the matrix elements of any operator \hat{X} in the new basis $|B\rangle$ to those in the old basis $|A\rangle$ as follows

$$\begin{aligned}\langle B_i|\hat{X}|B_j\rangle &= \sum_m \sum_n \langle B_i|A_m\rangle \langle A_m|\hat{X}|A_n\rangle \langle A_n|B_j\rangle , \\ &= \sum_m \sum_n \langle A_i|U_\Theta^\dagger|A_m\rangle \langle A_m|\hat{X}|A_n\rangle \langle A_n|U_\Theta|A_j\rangle .\end{aligned}\quad (5.109)$$

Defining the trace of an operator \hat{X} as the sum of its diagonal matrix elements computed in a basis, say $\{|A\rangle\}$

$$\text{Tr}_{|A\rangle} \hat{X} = \sum_A \langle A | \hat{X} | A \rangle , \quad (5.110)$$

we find that the trace is invariant under the transformations brought about by U_Θ

$$\text{Tr}_{|A\rangle} \hat{X} = \sum_A \langle A | \hat{X} | A \rangle = \sum_A \sum_B \sum_{B'} \langle A | B \rangle \langle B | \hat{X} | B' \rangle \langle B' | A \rangle , \quad (5.111)$$

$$= \sum_A \sum_B \sum_{B'} \langle B' | A \rangle \langle A | B \rangle \langle B | \hat{X} | B' \rangle ,$$

$$= \sum_B \sum_{B'} \langle B' | B \rangle \langle B | \hat{X} | B' \rangle \quad \text{as } \sum_A |A\rangle \langle A| = 1 ,$$

$$= \sum_{B'} \langle B' | \hat{X} | B' \rangle \quad \text{as } \sum_B |B\rangle \langle B| = 1 ,$$

$$= \sum_B \langle B | \hat{X} | B \rangle = \text{Tr}_{|B\rangle} \hat{X} . \quad (5.112)$$

Further, for $A(\lambda) = U_\Theta^\dagger A U_\Theta = e^{i\lambda\Theta} A e^{-i\lambda\Theta}$, it can be seen that

$$\frac{dA(\lambda)}{d\lambda} = i [\Theta, A(\lambda)] . \quad (5.113)$$

This can be easily derived by using the Baker-Campbell-Hausdorff relation for $A(\lambda)$ given above:

$$\begin{aligned} \frac{dA(\lambda)}{d\lambda} &= \frac{d}{d\lambda} (U_\Theta^\dagger A U_\Theta) = \frac{d}{d\lambda} (e^{i\lambda\Theta} A e^{-i\lambda\Theta}) , \\ &= \frac{d}{d\lambda} (A + i\lambda [\Theta, A] + \frac{(i\lambda)^2}{2!} [\Theta, [\Theta, A]] + \dots + \frac{(i\lambda)^n}{n!} [\Theta, [\Theta, [\Theta, \dots [\Theta, A]]] \dots] + \dots) , \\ &= i [\Theta, A] + \frac{2i^2\lambda}{2!} [\Theta, [\Theta, A]] + \frac{3i^3\lambda^2}{3!} [\Theta, [\Theta, [\Theta, A]]] + \dots , \\ &= i [\Theta, A] + i(i\lambda) [\Theta, [\Theta, A]] + i \frac{(i\lambda)^2}{2!} [\Theta, [\Theta, [\Theta, A]]] + \dots , \\ &= i \left[\Theta, (A + i\lambda [\Theta, A] + \frac{(i\lambda)^2}{2!} [\Theta, [\Theta, A]] + \frac{(i\lambda)^3}{3!} [\Theta, [\Theta, [\Theta, A]]] + \dots) \right] , \\ \Rightarrow \frac{dA(\lambda)}{d\lambda} &= i [\Theta, A(\lambda)] . \end{aligned} \quad (5.114)$$

A special case of the above relation corresponds to $\hat{\Theta}$ being the Hamiltonian H , $\hat{\Theta} \equiv H$, and $\lambda \equiv t$, such that $U_\Theta(t) \equiv U_t = e^{-itH/\hbar}$ corresponds to the time (t) evolution operator, the time evolved operator $A(t)$ is given by

$$A(t) = U_t^\dagger A U_t = e^{itH/\hbar} A e^{-itH/\hbar} , \text{ and} \quad (5.115)$$

$$\implies i\hbar \frac{dA(t)}{dt} = [A(t), H] \quad (5.116)$$

is the Heisenberg equation of motion for the operator $A(t)$. We will explore the origin of this equation of motion for the operators such as A shortly.

5.8 Matrix representation of Operators & Matrix mechanics

From the above discussion, we can see that the matrix representation of the operator \hat{X} in the basis $\{|A\rangle\}$ is given by

$$\hat{X} = \sum_{|A'\rangle} \sum_{|A\rangle} |A'\rangle \langle A'| \hat{X} |A\rangle \langle A| , \quad (5.117)$$

where the matrix \tilde{X} whose elements are given by $\langle A' | \hat{X} | A \rangle$ can be written explicitly as

$$\tilde{X} = \begin{pmatrix} \langle A_1 | \hat{X} | A_1 \rangle & \langle A_1 | \hat{X} | A_2 \rangle & \dots \\ \langle A_2 | \hat{X} | A_1 \rangle & \langle A_2 | \hat{X} | A_2 \rangle & \dots \\ \vdots & \vdots & \ddots \\ \vdots & \vdots & \ddots \end{pmatrix} . \quad (5.118)$$

The matrix representation of \hat{X} becomes simple if we use the eigenstates ($|X\rangle$) of \hat{X} to compute the matrix elements

$$\hat{X} = \sum_{|X'\rangle} \sum_{|X\rangle} |X'\rangle \langle X' | \hat{X} |X\rangle \langle X| , \quad (5.119)$$

$$= \sum_{|X'\rangle} \sum_{|X\rangle} |X'\rangle (\langle X | \hat{X} | X \rangle \delta_{XX'}) \langle X | , \quad (5.120)$$

$$= \sum_{|X'\rangle} \sum_{|X\rangle} |X'\rangle X \delta_{XX'} \langle X | , \quad (5.121)$$

$$= \sum_{|X\rangle} X |X\rangle \langle X| = \sum_{|X\rangle} X P_{|X\rangle} , \quad (5.122)$$

where $P_{|X\rangle} \equiv |X\rangle \langle X|$ is the projector operator onto the state $|X\rangle$. Thus, we find that the operator in its diagonal basis can be written in terms of projectors onto its eigenstates.

Thus, we can use the above result to write the time-independent Schrödinger equation in the energy (or Hamiltonian) eigenbasis $\{|\psi\rangle\}$ as follows

$$H |\psi\rangle = \left(\sum_{|\psi'\rangle} E_{\psi'} |\psi'\rangle \langle \psi'| \right) |\psi\rangle , \quad (5.123)$$

$$= \sum_{|\psi'\rangle} E_{\psi'} |\psi'\rangle \langle \psi'| \psi \rangle , \quad (5.124)$$

$$= \sum_{|\psi'\rangle} E_{\psi'} |\psi'\rangle \delta_{\psi\psi'} \text{ as } \langle \psi' | \psi \rangle = \delta_{\psi\psi'} , \quad (5.125)$$

$$= E_\psi |\psi\rangle , \quad (5.126)$$

which gives the energy eigenvalues as the set $\{E_\psi\}$.

Chapter 6

Heisenberg's Uncertainty principle

6.1 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.



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

Heisenberg

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

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

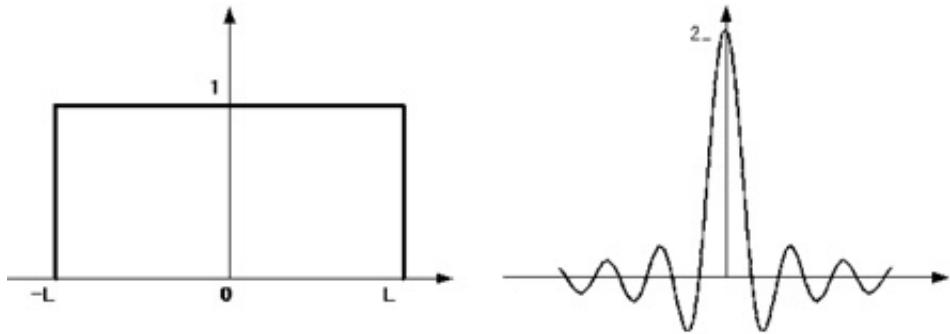


Figure 6.1: The square wavepacket and its Fourier transform (the sinc function, also known as the “Fraunhofer diffraction pattern”). Source: The internet.

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).

6.2 The γ -ray microscope gedanken

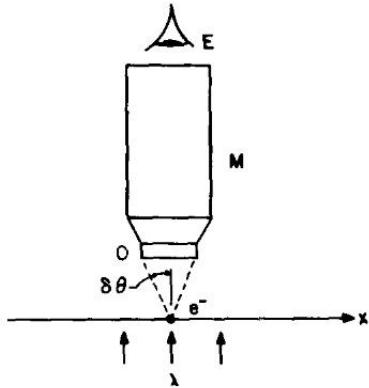


Figure 6.2: The γ -ray microscope gedanken. See text for discussion. Source: R. Shankar’s text “Principles of Quantum Mechanics”.

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.6.2 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

$$\Delta x = \frac{\lambda}{\sin \delta\theta} \simeq \frac{\lambda f}{a}. \quad (6.1)$$

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

$$0 \leq p_x \leq p \sin \delta\theta \simeq \frac{ha}{\lambda f} \Rightarrow \Delta p_x \simeq \frac{ha}{\lambda f}. \quad (6.2)$$

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

$$\Delta x \cdot \Delta p_x \simeq h, \quad (6.3)$$

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 $\Delta p_x \rightarrow 0$, 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 $h \rightarrow 0$ such that we can simultaneously have both $\Delta x \rightarrow 0$ as well as $\Delta p_x \rightarrow 0$. 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 $h \rightarrow 0$ is a heuristic tool ... we don't really understand at present how this emergence takes place!

6.3 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,

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

Note that we are dropping the $\hat{}$ (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

$$\begin{aligned}\int dx i\hbar \frac{d\psi^*}{dx} x \psi &= -i\hbar \int dx \psi^* \psi - i\hbar \int dx \frac{d\psi}{dx} x \psi^* \quad (\text{where } \psi^* x \psi|_{-\infty}^{\infty} \rightarrow 0 \text{ as } \psi, \psi^*|_{-\infty}^{\infty} \rightarrow 0) \\ &= -i\hbar \int dx \psi^* \psi + \left[i\hbar \int dx \frac{d\psi^*}{dx} x \psi \right]^* \quad (\text{as } x^* = x) .\end{aligned}\tag{6.4}$$

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

$$i\hbar \int dx |\psi|^2 = 2i\text{Im} \left[\int i\hbar dx \frac{d\psi^*}{dx} x \psi \right] \tag{6.5}$$

$$\therefore \hbar^2 \left| \int dx |\psi|^2 \right|^2 = 4 \left| \text{Im} \int dx i\hbar \frac{d\psi^*}{dx} x \psi \right|^2 \leq 4 \left| i\hbar \int dx \frac{d\psi^*}{dx} x \psi \right|^2 , \tag{6.6}$$

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

$$\begin{aligned}\left| \int dx i\hbar \frac{d\psi^*}{dx} x \psi \right|^2 &\leq \int dx x \psi^* x \psi \int dx \left(i\hbar \frac{\partial \psi^*}{\partial x} \right) \left(-i\hbar \frac{\partial \psi}{\partial x} \right) \\ \left| \int dx i\hbar \frac{d\psi^*}{dx} x \psi \right|^2 &\leq \int dx x^2 |\psi|^2 \int dx \left| -i\hbar \frac{\partial \psi}{\partial x} \right|^2 .\end{aligned}\tag{6.7}$$

Thus, the inequality relation eq.(6.6) becomes

$$\begin{aligned}\underbrace{\frac{\hbar^2}{4} \left| \int dx |\psi|^2 \right|^2}_{=1} &\leq \underbrace{\int dx x^2 |\psi|^2}_{\langle x^2 \rangle} \times \underbrace{\int dx \left(i\hbar \frac{\partial \psi^*}{\partial x} \right) \left(-i\hbar \frac{\partial \psi}{\partial x} \right)}_{\langle p^2 \rangle} \\ \Rightarrow \frac{\hbar^2}{4} &\leq \langle x^2 \rangle \langle p^2 \rangle .\end{aligned}\tag{6.8}$$

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

$$\Delta x = \sqrt{\langle (x - \langle x \rangle)^2 \rangle}, \quad \Delta p = \sqrt{\langle (p - \langle p \rangle)^2 \rangle} , \tag{6.9}$$

we replace x and p by their generalisations $x - \langle x \rangle$ and $p - \langle p \rangle$ respectively in eq.(6.8) to obtain

$$\frac{\hbar^2}{4} \leq \langle (x - \langle x \rangle)^2 \rangle \langle (p - \langle p \rangle)^2 \rangle \equiv \left(\Delta x \right)^2 \left(\Delta p \right)^2 \tag{6.10}$$

$$\Rightarrow \frac{\hbar}{2} \leq \Delta x \Delta p . \tag{6.11}$$

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 .

6.3.1 A generalised Uncertainty principle

Let us consider a state $|\psi\rangle$ and two Hermitian operators $A \equiv A^\dagger$ and $B \equiv B^\dagger$, such that

$$A|\psi\rangle = |X\rangle , \quad iB|\psi\rangle = |Y\rangle . \quad (6.12)$$

Now, we define the norms of $|X\rangle, |Y\rangle$ and $|X\rangle + |Y\rangle$ as

$$|X| = \sqrt{\langle X|X \rangle} , \quad |Y| = \sqrt{\langle Y|Y \rangle} , \quad |X+Y| = \sqrt{(\langle X| + \langle Y|)(|X\rangle + |Y\rangle)} . \quad (6.13)$$

Then, we start from the triangle inequality

$$|X| + |Y| \geq |X+Y| , \quad (6.14)$$

square it and simplify to get

$$\begin{aligned} (|X| + |Y|)^2 &\geq |X+Y|^2 , \\ \text{or, } |X|^2 + |Y|^2 + 2|X||Y| &\geq (\sqrt{(\langle X| + \langle Y|)(|X\rangle + |Y\rangle)})^2 , \\ &\geq (\langle X| + \langle Y|)(|X\rangle + |Y\rangle) , \\ &\geq |\langle X|X \rangle + \langle Y|Y \rangle + \langle X|Y \rangle + \langle Y|X \rangle| , \\ \implies 2|X||Y| &\geq |\langle X|Y \rangle + \langle Y|X \rangle| , \end{aligned} \quad (6.15)$$

which is the form of the Cauchy-Schwarz inequality we will use to obtain the uncertainty principle.

Substituting eq.(6.12) in eq.(6.15), we use

$$\begin{aligned} \langle Y| &= (|Y\rangle)^\dagger = (iB|\psi\rangle)^\dagger = -i\langle\psi|B^\dagger = -i\langle\psi|B , \\ \langle X| &= (|X\rangle)^\dagger = (A|\psi\rangle)^\dagger = \langle\psi|A^\dagger = \langle\psi|A , \end{aligned} \quad (6.16)$$

to obtain

$$2\sqrt{\langle A^2 \rangle \langle B^2 \rangle} \geq |\langle\psi|AB|\psi\rangle - \langle\psi|BA|\psi\rangle| , \quad (\text{removing overall } i \text{ factor from ||}) \quad (6.17)$$

Now, for the case $\langle A \rangle = 0 = \langle B \rangle$, $\langle A^2 \rangle = (\Delta A)^2, \langle B^2 \rangle = (\Delta B)^2$, this gives us

$$\begin{aligned} 2\Delta A \Delta B &\geq |\langle\psi|[A, B]|\psi\rangle| , \\ \implies \Delta A \Delta B &\geq \frac{1}{2}|\langle\psi|[A, B]|\psi\rangle| . \end{aligned} \quad (6.18)$$

This is the generalised uncertainty relation for the case of $\langle A \rangle = 0 = \langle B \rangle$. For the case of $\langle A \rangle \neq 0, \langle B \rangle \neq 0$, we can obtain the generalised uncertainty relation by defining two new operators

$$\bar{A} = A - \langle A \rangle , \quad \bar{B} = B - \langle B \rangle . \quad (6.19)$$

Then, using the fact that the uncertainties in A and B are defined as $(\Delta A)^2 = \langle \bar{A}^2 \rangle$, $(\Delta B)^2 = \langle \bar{B}^2 \rangle$, and the fact that $[\bar{A}, \bar{B}] = [A, B]$, we obtain the generalised uncertainty relation as

$$\Delta A \Delta B \geq \frac{1}{2} |\langle \psi | [\bar{A}, \bar{B}] | \psi \rangle| . \quad (6.20)$$

Putting in $A \equiv x$ and $B \equiv p_x$, we obtain the results of the previous subsection: $\Delta x \Delta p_x \geq \frac{\hbar}{2}$, as the commutator $[x, p_x] = i\hbar$.

Why did we choose the operator iB above, rather than B (even though B itself is Hermitian, i.e., $B^\dagger = B$)? Recall that the commutator of two Hermitian operators A and B , $[A, B]$ is actually anti-Hermitian in nature

$$([A, B])^\dagger = (AB - BA)^\dagger = B^\dagger A^\dagger - A^\dagger B^\dagger , \quad (6.21)$$

$$= BA - AB = -(AB - BA) = -[A, B] . \quad (6.22)$$

Such an anti-Hermitian operator has purely imaginary values: for an operator $C^\dagger = -C$

$$\langle \psi_n | C | \psi_n \rangle = c_n \langle \psi_n | \psi_n \rangle , \quad (6.23)$$

$$= (\langle \psi_n | -C^\dagger) | \psi_n \rangle = -c_n^* \langle \psi_n | \psi_n \rangle , \quad (6.24)$$

$$\implies c_n = -c_n^* , \quad (6.25)$$

which means that the eigenvalue c_N of the anti-Hermitian operator C is purely imaginary. Thus, we can see that the eigenvalue of the anti-Hermitian commutator $[A, B]$ is purely imaginary. Thus, in order to obtain a commutator that is Hermitian, we consider the anti-Hermitian operator iB , $(iB)^\dagger = -iB$ and its commutator with A

$$([A, iB])^\dagger = (A \times iB - iB \times A)^\dagger = (iB)^\dagger A^\dagger - A^\dagger (iB)^\dagger , \quad (6.26)$$

$$= (-iB) \times A - A \times (-iB) = i(AB - BA) = [A, iB] . \quad (6.27)$$

Thus, we now have a Hermitian commutator $[A, iB]$. Further, as we can see above, it yields real valued lower bound for the product of the uncertainties.

6.3.2 The Energy-time uncertainty relation

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 that particular eigenstate 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 $\Delta t \rightarrow \infty$. On the other hand, a non-stationary state (i.e., one with a finite lifetime) will undergo decay; such states have a finite and non-zero ΔE . More details on 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. This can give us an estimate (in the form of a lower bound) for time evolution between states, as was first pursued by Mandelstam and Tamm in 1945 (J. Phys. USSR **9**, 249-254 (1945))! They reasoned as follows.

Recall that the generalised uncertainty relation was given by

$$\Delta A \Delta B \geq \frac{1}{2} |\langle \psi | [A, B] | \psi \rangle| . \quad (6.28)$$

Further, the Ehrenfest relations for the time evolution of the expectation values of the operators A and B are given by

$$\frac{d\langle A(t) \rangle}{dt} = \frac{1}{i\hbar} \langle [H, A(t)] \rangle , \quad \frac{d\langle B(t) \rangle}{dt} = \frac{1}{i\hbar} \langle [H, B(t)] \rangle , \quad (6.29)$$

where $A(t) = U_t A U_t^\dagger$, $B(t) = U_t B U_t^\dagger$ and $U_t = e^{itH/\hbar}$. Taking $A \equiv H$ (the Hamiltonian), and using the Ehrenfest relation for B in the generalised uncertainty relation, we get the product of the uncertainties in the energy E and B

$$\begin{aligned} \Delta E \Delta B &\geq \frac{1}{2} |\langle \psi | [H, B] | \psi \rangle| = \frac{1}{2} |\langle \psi | i\hbar \frac{dB(t)}{dt} | \psi \rangle| , \\ \implies \Delta E \Delta B &\geq \frac{\hbar}{2} \left| \langle \psi | \frac{dB(t)}{dt} | \psi \rangle \right| . \end{aligned} \quad (6.30)$$

We now integrate this relation over a time period from t to $t + \Delta t$. For this, we note that (i) we are integrating over a total derivative in t on the right hand side, (ii) ΔE does not change over the time interval Δt and that (iii) ΔB is the average value of the variance in B computed over Δt . Thus, we obtain

$$\begin{aligned} \Delta E \times \Delta B &\geq \frac{\hbar}{2} \left| \frac{\langle B_{t+\Delta t} \rangle - \langle B_t \rangle}{\Delta t} \right| , \\ \implies \Delta E \Delta t &\geq \frac{\hbar}{2} \left| \frac{\langle B_{t+\Delta t} \rangle - \langle B_t \rangle}{\Delta B} \right| , \end{aligned} \quad (6.31)$$

for ΔB computed over the time interval Δt . Now, for the time interval ΔT such that $\langle B_{t+\Delta T} \rangle - \langle B_t \rangle = \Delta B$, we have

$$\Delta E \Delta T \geq \frac{\hbar}{2} . \quad (6.32)$$

In this light, ΔT is referred to as the evolution time for the state $|\psi(t=0)\rangle$ to evolve to $|\psi(t=T)\rangle$ such that the condition $\langle B_{t+\Delta T} \rangle - \langle B_t \rangle = \Delta B$ is satisfied. Note that for the energy-time uncertainty relation to be satisfied, it is important that both $\Delta E \neq 0$ as well as $\Delta B \neq 0$ (as long as $\Delta E \neq \infty$). This means that the state $|\psi\rangle$ whose evolution we are considering can be neither an eigenstate of H nor of B .

Finally, we note that in both open as well as interacting quantum systems, 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. It is then natural to expect that such states are neither likely eigenstates of the fully interacting theory's Hamiltonian nor an arbitrary observable's operator.

Chapter 7

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

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

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

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

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

$$i\hbar\psi(\vec{r})\frac{\partial\chi(t)}{\partial t} = E\psi(\vec{r})\chi(t) \quad (7.1)$$

$$\implies \dot{\chi}(t) = \frac{-iE}{\hbar}\chi(t), \quad (7.2)$$

$$\chi(t) = e^{\frac{-iEt}{\hbar}}, \quad (7.3)$$

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

Similarly,

$$\left[\frac{-\hbar^2}{2m}\nabla^2 + V(\vec{r})\right]\psi(\vec{r})\chi(t) = E\psi(\vec{r})\chi(t)$$

$$\text{or, } \left[\frac{-\hbar^2}{2m}\nabla^2 + V(\vec{r})\right]\psi(\vec{r}) = E\psi(\vec{r}) \quad (7.4)$$

$$(7.5)$$

such that

$$\Psi(\vec{r}, t) = \psi(\vec{r})e^{\frac{-iEt}{\hbar}} \quad (7.6)$$

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

$$\begin{aligned} U\Psi(\vec{r}, t) &= e^{-\frac{i\hat{H}t}{\hbar}}\Psi(\vec{r}, t) \\ &= e^{-\frac{iEt}{\hbar}}\Psi(\vec{r}, t), \end{aligned} \quad (7.7)$$

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

$$\text{Probability Density : } \Psi^*(\vec{r}, t)\Psi(\vec{r}, t) = \psi^*(\vec{r})e^{\frac{iEt}{\hbar}}\psi(\vec{r})e^{\frac{-iEt}{\hbar}} = \psi^*(\vec{r})\psi(\vec{r}) \text{ and} \quad (7.8)$$

$$\begin{aligned} \text{Expectation Value : } \langle \hat{\theta} \rangle &= \int_v d\vec{r} \psi^*(\vec{r})e^{\frac{iEt}{\hbar}}\hat{\theta}\psi(\vec{r})e^{\frac{-iEt}{\hbar}} \\ &= \int_v d\vec{r} \psi^*(\vec{r})\hat{\theta}\psi(\vec{r}). \end{aligned} \quad (7.9)$$

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.(7.5)) also helps us define the Hamiltonian operator

$$\hat{H} = \frac{-\hbar^2}{2m}\nabla^2 + V(\vec{r}) \text{ whose eigenvalue is } E, \quad (7.10)$$

$$\hat{H}\psi(\vec{r}) = E\psi(\vec{r}). \quad (7.11)$$

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

$$i\hbar \frac{\partial \Psi_1}{\partial t} = E\Psi_1, \quad \Psi_1 = \psi_1(\vec{r})\chi_1(t) \quad \& \quad i\hbar \frac{\partial \Psi_2}{\partial t} = E\Psi_2, \quad \Psi_2 = \psi_2(\vec{r})\chi_2(t), \quad (7.12)$$

$$\text{i.e., } \hat{H}\psi_1 = E\psi_1, \quad , \quad \hat{H}\psi_2 = E\psi_2, \quad (7.13)$$

then

$$\begin{aligned} \hat{H}(c_1\psi_1 + c_2\psi_2) &= Ec_1\psi_1 + Ec_2\psi_2 \quad (c_1, c_2 \in \mathcal{C}), \\ &= E(c_1\psi_1 + c_2\psi_2), \end{aligned} \quad (7.14)$$

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**.

7.1 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 $x \rightarrow \pm\infty$, $\psi(x) \rightarrow 0$; 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., spatially 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

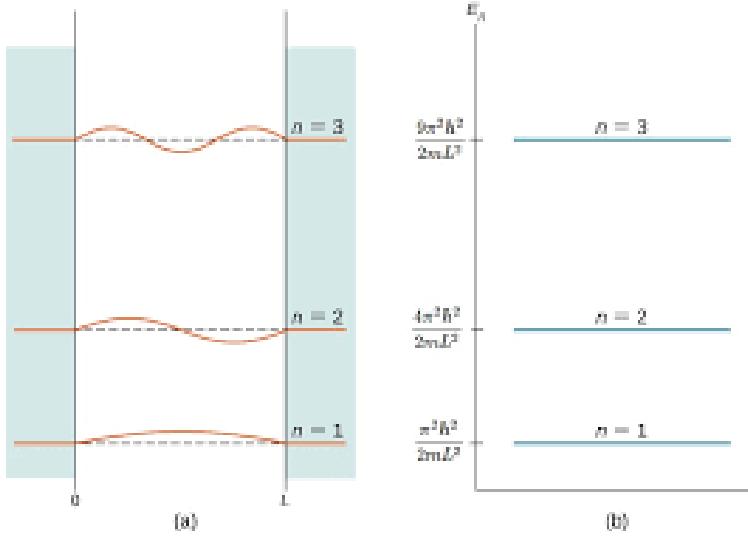
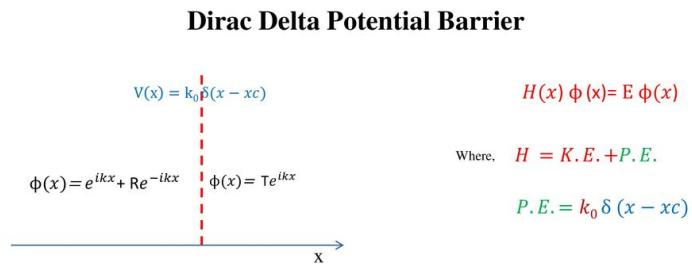
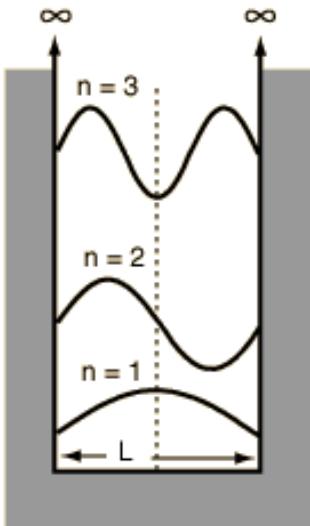


Figure 7.1: Wavefunctions (Left) and Energy Spectrum (Right) for the lowest three eigenstates of the particle in a box problem.

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 $V \rightarrow \infty$ 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 n th excited state ($n \in \mathbb{Z}$) 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 $x \rightarrow \pm\infty$ (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.



Two Boundary Conditions

$$1. \phi(xc + \epsilon) = \phi(xc - \epsilon)$$

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

Analytical Solution possible!

Two Unknowns: R & T

$x = 0$ at left wall of box.

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

Figure 7.2: (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 8

Quantum dynamics: the Schrödinger, Heisenberg and interaction pictures

We can now bring some more clarity to the time evolution (or dynamics) of the quantum mechanical universe. Recall that under time evolution, the state vectors transformed as $|\psi\rangle \rightarrow |\psi(t)\rangle = U_t |\psi\rangle$ where $U_t = e^{-iHt/\hbar}$, but preserved the inner product (and norms) of the associated Hilbert space:

$$\langle\phi|\psi\rangle \rightarrow \langle\phi(t)|\psi(t)\rangle = (\langle\phi|U_t^\dagger)(U_t|\psi\rangle) = \langle\phi|\psi\rangle . \quad (8.1)$$

From what you have studied in earlier chapters, you will recall that we assumed that while state (vectors) such as $|\psi\rangle$ underwent time evolution as given above, operators acting on these states did not. Thus, for the time evolution of matrix elements such as $\langle\phi|\hat{A}|\psi\rangle$, we find that

$$\begin{aligned} \langle\phi|\hat{A}|\psi\rangle \rightarrow \langle\phi(t)|\hat{A}|\psi(t)\rangle &= (\langle\phi|U_t)\hat{A}(U_t^\dagger|\psi\rangle) , \\ &= \langle\phi|(U_t^\dagger\hat{A}U_t)|\psi\rangle , \\ &= \langle\phi|\hat{A}(t)|\psi\rangle , \end{aligned} \quad (8.2)$$

where $\hat{A}(t) = U_t^\dagger \hat{A} U_t$. If we ponder on this for a moment, this relation tells us something really interesting: computing the matrix element of the time-unevolved operator \hat{A} with time-evolved states $|\psi(t)\rangle$ and $|\phi(t)\rangle$ is equivalent to computing the matrix element of the time-evolved operator $\hat{A}(t)$ defined above but on the time-unevolved states $|\psi\rangle$ and $|\phi\rangle$. The first approach, where the states $|\psi\rangle$ evolve in time but the operators \hat{A} do not, is referred to as the Schrödinger picture of quantum mechanics. On the other hand, the second approach — where the states $|\psi\rangle$ are unchanged in time but the operators \hat{A} evolve in time, is referred to as the Heisenberg picture of quantum mechanics.

The earlier chapters dealt with the Schrödinger picture, where the only time evolution we considered was for problems where the Hamiltonian (or, the potential operator \hat{V}) had no explicit time dependence ($\hat{V} \equiv V(r)$). There, we studied the eigenstates $\{|\psi\rangle\}$ of the Hamiltonian H , and whose time evolution was simply given by $\{|\psi\rangle\} \rightarrow \{|\psi(t)\rangle\} = \{U_t|\psi\rangle\}$ where $U_t = e^{-iHt/\hbar}$. On the other hand, in the Heisenberg picture, we would keep the states

unchanged (i.e., precisely the same) with time, but allow the operators to evolve according to the Heisenberg equation of motion we derived earlier:

$$\begin{aligned}\frac{d\hat{A}(t)}{dt} &= \frac{d}{dt} \left(U_t^\dagger \hat{A} U_t \right) , \\ &= \frac{i}{\hbar} [H, A(t)] , \\ \implies i\hbar \frac{d\hat{A}(t)}{dt} &= [A(t), H] .\end{aligned}\tag{8.3}$$

Clearly, for the case of U_t (or equivalently, the Hamiltonian H) commuting with \hat{A} , $[U_t, \hat{A}] = 0$, the equation of motion tells us that the eigenvalues of \hat{A} are conserved in time.

It is also instructive to compare the Heisenberg equation of motion with the classical equation of motion for a quantity $A \equiv A(p, q)$ (where p and q are the generalised classical momenta and coordinates respectively) that does not have any explicit time dependence, written in the Poisson bracket form

$$\frac{dA}{dt} = [A, H]_{\text{Classical}} ,\tag{8.4}$$

where H is the Hamiltonian of the classical system. It can be argued that the the classical equation of motion can be obtained from it's quantum counterpart (the Heisenberg equation of motion) via the ansatz

$$\frac{[A, H]}{i\hbar} \rightarrow [A, H]_{\text{Classical}} .\tag{8.5}$$

This ansatz is believed to be one way in which to see the correspondence between the dynamics of the classical and quantum worlds. However, we must deal with the complication that the quantum world deals with experimentally observed quantities that arise from the action of operators on state vectors, while this is not true in the classical world. Thus, another way by which to see the correspondence is through the Ehrenfest theorem, which involves looking at the equation of motion for the expectation value of an operator $\langle \hat{A} \rangle$

$$i\hbar \frac{d\langle \hat{A}(t) \rangle}{dt} = \langle [A(t), H] \rangle \rightarrow [A, H]_{\text{Classical}} ,\tag{8.6}$$

as leading to the equivalent classical experimentally observable quantity.

A word on “transition amplitudes” is due at this point. Imagine a system at time $t = 0$ prepared in an eigenstate of observable \hat{A} with eigenvalue a , $|a\rangle$. We can ask: what is the probability for the system to be found in an eigenstate of an observable B with eigenvalue b , $|b\rangle$, at a later time t ? This is known as the “transition amplitude” to go from $|a\rangle$ at time $t = 0$ to $|b\rangle$ at time t . In the Schrödinger picture, as $|a\rangle \rightarrow U_t |a\rangle$, this transition amplitude is given simply by the inner product $\langle b| (U_t |a\rangle)$. By contrast, in the Heisenberg picture $|a\rangle$ remains unchanged under time evolution. However, the transition amplitude in going from $|a\rangle$ to $|b\rangle$ with time cannot change dependent on the picture we have chosen for the time evolution. Thus, the invariance of the transition amplitude dictates that, in the Heisenberg picture, we interpret the transition amplitude as $(\langle b| U_t) |a\rangle$, i.e., the final state $|b\rangle$ in the

Heisenberg picture undergoes a time evolution opposite to that of the initial state $|a\rangle$ in the Schrödinger picture. The same is true of the so-called “survival probability” $|\langle b| U_t |a\rangle|^2$, i.e., the probability corresponding to the transition amplitude.

8.0.1 The interaction picture

For the case when $\hat{V} \equiv V(r, t)$, we would have to treat the time-dependent Schrödinger equation

$$\hat{H}\Psi(\vec{r}, t) = \left[\frac{-\hbar^2}{2m} \nabla^2 + V(\vec{r}, t) \right] \Psi(\vec{r}, t) = i\hbar \frac{\partial\Psi(\vec{r}, t)}{\partial t} . \quad (8.7)$$

But how would we go about this? As we will now see, this involves working with neither the Schrödinger nor the Heisenberg pictures, but rather with a picture called the “Interaction” (or Dirac) picture which is intermediate between the Schrödinger and Heisenberg pictures. Consider a Hamiltonian H which can be split into two parts, one of which is time-independent (H_0) while the other is time-dependent ($V(t)$): $H = H_0 + V(t)$. We assume that we know the eigenstates ($|n\rangle$) and energy eigenvalues (E_n) corresponding to H_0 : $H_0|n\rangle = E_n|n\rangle$.

The complication in dealing with $V(t)$ arises from the fact that we cannot presume that the time evolution of the eigenstates of the complete Hamiltonian H is simply a phase factor $e^{-iHt/\hbar}$, as H is now (through $V(t)$) itself changing in time t . Even if we start from the eigenstates of H_0 , $\{|n\rangle\}$, at time $t = 0$, the application of $V(t)$ for $t > 0$ can very generally lead to transitions between the states $\{|n\rangle\}$. How then would we determine how an arbitrary state vector $|\alpha\rangle$ evolves in time? Suppose at $t = 0$, we start from

$$|\alpha\rangle = \sum_n c_n(t=0)|n\rangle , \quad (8.8)$$

we wish to find $c_n(t)$ for $t > 0$ such that

$$|\alpha, t_0 = 0; t\rangle_S = \sum_n c_n(t)e^{-iE_nt/\hbar}|n\rangle , \quad (8.9)$$

where $|\alpha, t_0 = 0; t\rangle_S$ corresponds to a state vector in the Schrödinger picture at time t of a system whose state at time $t = 0$ is given by $|\alpha\rangle$ (see above). Note that in eq.(8.9), we have already explicitly accounted for the time evolution phase factor $e^{-iE_nt/\hbar}$ of the eigenstates $|n\rangle$ of the Hamiltonian H_0 (as this is present even if V is absent). In this way, we ensure that the time evolution of the coefficients $c_n \equiv c_n(t)$ arise purely from the presence of $V(t)$ and such that $c_n \equiv c_n(t=0)$ if $V(t)$ were set to zero. Clearly, the probability of finding $|n\rangle$ in $|\alpha, t_0 = 0; t\rangle$ is given by $|c_n(t)|^2$, and $\sum_n |c_n(t)|^2 = 1$ (i.e., unitarity is guaranteed). We will comment on the validity (or correctness) of the form of $|\alpha, t_0 = 0; t\rangle_S$ given in eq.(8.9) above at the end of this section.

We define the time dependent state in the interaction picture, $|\alpha, t_0; t\rangle_I$ (where the I subscript denotes the interaction picture), in terms of the state in the Schrödinger picture $|\alpha, t_0; t\rangle_S$

$$|\alpha, t_0; t\rangle_I = e^{iH_0t/\hbar} |\alpha, t_0; t\rangle_S . \quad (8.10)$$

Clearly, at $t = 0$, $|\alpha, t_0; t\rangle_I = |\alpha, t_0; t\rangle_S$. The operators related to observables are defined in the interaction picture as

$$\begin{aligned}\hat{A}_I &= \equiv e^{iH_0t/\hbar} \hat{A}_S e^{-iH_0t/\hbar}, \\ \implies V_I &= \equiv e^{iH_0t/\hbar} V(t) e^{-iH_0t/\hbar},\end{aligned}\quad (8.11)$$

where the potential $V(t)$ is understood to be the time-dependent potential in the Schrödinger picture. Recalling that the Heisenberg picture is connected to the Schrödinger picture as follows

$$\begin{aligned}|\alpha\rangle_H &= e^{iHt/\hbar} |\alpha, t_0; t\rangle_S, \\ \hat{A}_H &= e^{iHt/\hbar} \hat{A}_S e^{-iHt/\hbar},\end{aligned}\quad (8.12)$$

we see that the basic difference between the way in which the interaction picture is connected to the Schrödinger picture vis-a-vis the connection between the Heisenberg and Schrödinger pictures is given that the former connection relies on H_0 while the latter relies on H .

In order to obtain the differential equation that governs the time evolution of a state in the interaction picture, we take the time derivative of eq.(8.10)

$$\begin{aligned}i\hbar \frac{\partial}{\partial t} |\alpha, t_0; t\rangle_I &= i\hbar \frac{\partial}{\partial t} (e^{iH_0t/\hbar} |\alpha, t_0; t\rangle_S), \\ &= -H_0 e^{iH_0t/\hbar} |\alpha, t_0; t\rangle_S + e^{iH_0t/\hbar} (i\hbar \frac{\partial}{\partial t}) |\alpha, t_0; t\rangle_S, \\ &= -H_0 e^{iH_0t/\hbar} |\alpha, t_0; t\rangle_S + e^{iH_0t/\hbar} (H_0 + V) |\alpha, t_0; t\rangle_S, \text{ using } i\hbar \frac{\partial}{\partial t} |\psi\rangle_S = H |\psi\rangle_S \\ &= e^{iH_0t/\hbar} V e^{-iH_0t/\hbar} |\alpha, t_0; t\rangle_I, \\ \implies i\hbar \frac{\partial}{\partial t} |\alpha, t_0; t\rangle_I &= V_I |\alpha, t_0; t\rangle_I.\end{aligned}\quad (8.13)$$

Clearly, this is a Schrödinger-like equation where the evolution of the state $|\alpha, t_0; t\rangle_I$ is dictated by V_I (instead of the total H), such that the state $|\alpha, t_0; t\rangle_I$ does not evolve if V_I vanishes. Similarly, if we take a time derivative of eq.(8.11) for an operator \hat{A} that does not contain time t explicitly, we find

$$\frac{d\hat{A}_I}{dt} = \frac{i}{\hbar} [H_0, \hat{A}_I] = \frac{1}{i\hbar} [\hat{A}_I, H_0], \quad (8.14)$$

which is a Heisenberg-like equation for \hat{A}_I with H being replaced by H_0 . An explicit time dependence within \hat{A} will lead to an additional $\frac{\partial \hat{A}}{\partial t}$ term in the above expression.

Now, we continue to use the eigenstates of the H_0 , $|n\rangle$, as our state vectors:

$$|\alpha, t_0; t\rangle_I = \sum_n c_n(t) |n\rangle. \quad (8.15)$$

Thus, we can act with $e^{-iH_0t/\hbar}$ on both sides, and use $H_0|n\rangle = E_n|n\rangle$ to get

$$\begin{aligned} e^{-iH_0t/\hbar}|\alpha, t_0 = 0; t\rangle_I &= e^{-iH_0t/\hbar} \sum_n c_n(t)|n\rangle , \\ \implies |\alpha, t_0 = 0; t\rangle_S &= \sum_n c_n(t)e^{-iH_0t/\hbar}|n\rangle , \\ &= \sum_n c_n(t)e^{-iE_n t/\hbar}|n\rangle , \end{aligned} \quad (8.16)$$

giving us back eq.(8.9) once again. Thus, we see that eq.(8.15) is just another way of writing eq.(8.9); indeed, eq.(8.15) has a simpler form (as all the phase factors of $e^{-iE_n t/\hbar}$ are missing). We can now derive the differential equation for the coefficient $c_n(t)$ by acting on both side of the Schrödinger-like equation in the interaction picture (eq.8.13) with $\langle n|$ to obtain

$$i\hbar \frac{\partial}{\partial t} \langle n|\alpha, t_0; t\rangle_I = \sum_m \langle n| V_I |m\rangle \langle m|\alpha, t_0; t\rangle_I . \quad (8.17)$$

$$\begin{aligned} \text{But } \langle n| V_I |m\rangle &= \langle n| e^{iH_0t/\hbar} V(t) e^{-iH_0t/\hbar} |m\rangle , \\ &= e^{i(E_n - E_m)t/\hbar} \langle m| V(t) |n\rangle , \\ &= V_{nm}(t) e^{i(E_n - E_m)t/\hbar} . \end{aligned} \quad (8.18)$$

$$\text{And } c_n(t) = \langle n|\alpha, t_0; t\rangle_I . \quad (8.19)$$

$$\therefore i\hbar \frac{dc_n(t)}{dt} = \sum_m V_{nm}(t) e^{i\omega_{nm}t} c_m(t) , \quad (8.20)$$

where $\omega_{nm} = \frac{E_n - E_m}{\hbar}$. This is the basic differential equation for the time evolution of $c_n(t)$ that arises from the fact that $V(t)$ causes couplings with various other coefficients $c_m(t)$. Explicitly, this can be written out as

$$i\hbar \frac{d}{dt} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ \vdots \\ \vdots \end{pmatrix} = \begin{pmatrix} V_{11} & V_{12}e^{i\omega_{12}t} & V_{13}e^{i\omega_{13}t} & \dots \\ V_{21}e^{-i\omega_{12}t} & V_{22} & V_{23}e^{i\omega_{23}t} & \dots \\ V_{31}e^{-i\omega_{13}t} & V_{32}e^{-i\omega_{23}t} & V_{33} & \dots \\ \vdots & \vdots & \vdots & \ddots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ \vdots \\ \vdots \end{pmatrix} . \quad (8.21)$$

The frequency ω_{mn} corresponds the energy cost of transitions between $|n\rangle$ and $|m\rangle$ brought about by the time-varying potential $V(t)$, i.e., $\langle m| V(t) |n\rangle \neq 0$. In this way, we have now arrived at an equation for the time evolution of the coefficients $c_n(t)$ in terms of the action of $V(t)$.

Finally, we end by asking: when is it safe to assume that eq.(8.9) for the time-evolved state in the Schrödinger picture is valid, i.e., when is the expansion of $|\alpha, t_0; t\rangle_S$ in terms of the eigenbasis of H_0 ($\{|n\rangle\}$) valid? Put another way, when can we be sure that the separation of the complete Hamiltonian into two separate parts, H_0 and $V(t)$, remains valid? The answer is that as long as the time-dependent potential $V(t)$ does not cause the ground state of H_0 , $|0\rangle$, to become orthogonal to itself. If this should happen, the eigenspectrum will undergo a

huge shake-up and the application of the interaction picture in terms of H_0 and $V(t)$ is no longer valid. This is precisely what happens at a quantum phase transition in a many-body system comprised of a thermodynamically large number of interacting constituents. Thus, we are assuming that we are nowhere near any such interesting physics!

Chapter 9

Schrödinger Wave Mechanics: Bound states

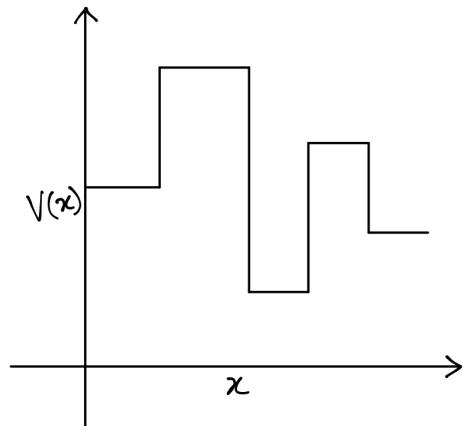


Figure 9.1: Schematic diagram for a piecewise constant potential in 1D.

For the case of piecewise constant potentials (i.e., potentials with step-like discontinuities as shown in figure 9.1), we consider a range of x over which the potential $V(x)$ is a constant then the TISE can be rewritten as

$$-\frac{\hbar^2}{2m} \frac{d^2\psi}{dx^2} + V_0\psi = E\psi \quad (9.1)$$

$$\implies \frac{d^2\psi}{dx^2} + \frac{2m}{\hbar^2} (E - V_0) \psi = 0 \quad (9.2)$$

Trying a guess (or *ansatz*) solution $\psi(x) = Ce^{nx}$ in Eq. 9.2, we get

$$\begin{aligned} n^2Ce^{nx} + \frac{2m}{\hbar^2}(E - V_0)Ce^{nx} &= 0 \\ n^2 &= \frac{2m}{\hbar^2}(V_0 - E) \\ \therefore n &= \pm \sqrt{\frac{2m}{\hbar^2}(V_0 - E)} \end{aligned} \tag{9.3}$$

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

$$\psi(x) = C_1 e^{nx} + C_2 e^{-nx} \tag{9.4}$$

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

$$\psi(x) = C_3 e^{i\bar{n}x} + C_4 e^{-i\bar{n}x} \tag{9.5}$$

$$= D_1 \cos(\bar{n}x) + D_2 \sin(\bar{n}x) \tag{9.6}$$

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 ψ s 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 in 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 10

Infinite square well/ Particle in a Box

Consider a potential as shown in figure 10.1.

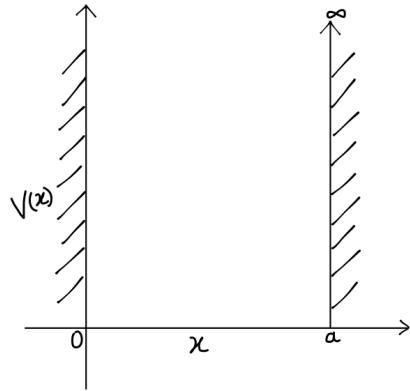


Figure 10.1: Schematic Diagram of the potential for the particle in a 1D Box.

$$V(x) = \begin{cases} 0 & \text{if } 0 < x < a \\ \infty & \text{otherwise} \end{cases} \quad (10.1)$$

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., $0 < x < a$,

$$\frac{d^2\psi}{dx^2} = -k^2\psi, \quad k = \frac{\sqrt{2mE}}{\hbar} \quad (E > 0) \quad (10.2)$$

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

$$\psi(x) = A \sin(kx) + B \cos(kx) \quad (10.3)$$

Boundary Conditions

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

$$\psi(0) = 0 = \psi(a) . \quad (10.4)$$

Using these two boundary conditions gives

$$\begin{aligned} \psi(0) = 0 \implies A \sin(0) + B \cos(0) &= 0 \\ B &= 0 \\ \therefore \psi(x) &= A \sin(kx) \end{aligned} \quad (10.5)$$

$$\begin{aligned} \psi(a) = 0 \implies A \sin(ka) &= 0 \\ \therefore k_n &= \frac{n\pi}{a}, \quad n \in \mathbb{Z} \end{aligned} \quad (10.6)$$

From Eq. 10.2,

$$\therefore E_n = \frac{\hbar^2 k_n^2}{2m} = \frac{n^2 \pi^2 \hbar^2}{2ma^2}, \quad n \in \mathbb{Z} = 1, 2, 3, \dots \dots \quad (10.7)$$

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

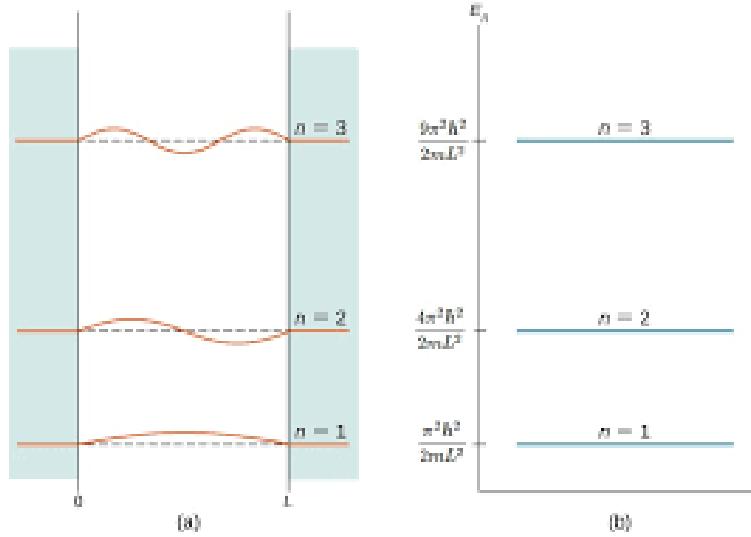


Figure 10.2: Wavefunctions (Left) and Energy Spectrum (Right) for the lowest three eigenstates of the particle in a box problem. Source: the internet.

Normalization

The normalizing condition given by Eq. 10.8 gives

$$\begin{aligned} & \int_{-\infty}^{\infty} |\psi_n|^2 dx = 1 \\ \implies & |A|^2 \int_0^a dx \sin^2(k_n x) = 1 \\ \therefore & A = \sqrt{\frac{2}{a}} \\ \psi_n(x) &= \sqrt{\frac{2}{a}} \sin\left(\frac{n\pi x}{a}\right) \end{aligned} \tag{10.8}$$

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 11

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 11.1

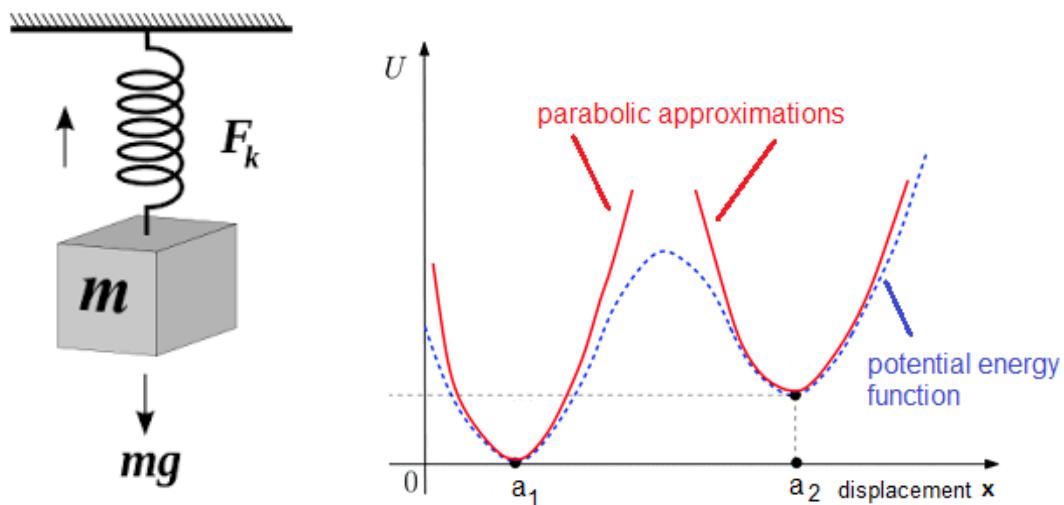


Figure 11.1: (Left) A classical mass spring oscillator system. (Right) Parabolic approximation at a minima

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.,

$|x - x_0| \ll 1$), we get

$$\begin{aligned} V(x) &= V(x_0) + \underbrace{V'(x_0)}_{=0}(x - x_0) + \frac{V''(x_0)}{2!}(x - x_0)^2 + \mathcal{O}((x - x_0)^3) \\ &\approx V(x_0) + \frac{1}{2}V''(x_0)(x - x_0)^2 \end{aligned} \quad (11.1)$$

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. 11.2

$$V(x) = \frac{1}{2}m\omega^2x^2 \quad (11.2)$$

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

$$-\frac{\hbar^2}{2m}\frac{d^2\psi(x)}{dx^2} + \frac{1}{2}m\omega^2x^2\psi(x) = E\psi(x) \quad (11.3)$$

11.1 Algebraic Method

Eq. 11.3 could be written as

$$\frac{1}{2m}[p^2 + (m\omega x)^2]\psi = E\psi \quad (11.4)$$

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

$$(u^2 + v^2) = (iu + v)(-iu + v) , \quad (11.5)$$

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_-

$$a_{\pm} = \frac{1}{\sqrt{2m\omega\hbar}}(\mp ip + m\omega x) \quad (11.6)$$

$$\begin{aligned} [a_+, a_-] &= \frac{1}{2m\omega\hbar} [(-ip + m\omega x)(ip + m\omega x) - (ip + m\omega x)(-ip + m\omega x)] \\ &= \frac{1}{2m\omega\hbar} [p^2 - p^2 + m\omega^2x(1 - 1) + 2im\omega(xp - px)] \\ &= \frac{2im\omega}{2m\omega\hbar} [x, p] \\ &= -1 \end{aligned} \quad (11.7)$$

$$\& [a_-, a_+] = 1 \quad (11.8)$$

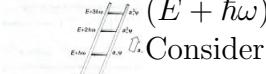
Here, a_{\pm} don't commute with each other as x, p don't commute. The Hamiltonian H can be written in terms of a_{\pm} as given by Eqs. 11.10, 11.11.

$$\begin{aligned} a_- a_+ &= \frac{1}{2m\hbar\omega} (ip + m\omega x)(-ip + m\omega x) \\ &= \frac{1}{2m\hbar\omega} (p^2 + m^2\omega^2 x^2 - im\omega [x, p]) \\ &= \frac{1}{2\hbar m\omega} (p^2 + m^2\omega^2 x^2) - \frac{i}{2\hbar} [x, p] \\ &= \frac{1}{\hbar\omega} H + \frac{1}{2} \end{aligned} \quad (11.9)$$

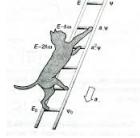
$$\therefore H = \hbar\omega \left(a_- a_+ - \frac{1}{2} \right) \quad (11.10)$$

$$= \hbar\omega \left(a_+ a_- + \frac{1}{2} \right) \quad \text{using } [a_-, a_+] = 1. \quad (11.11)$$

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 $(E - \hbar\omega)$.



Consider



SHO
spec-
trum.

$$\begin{aligned} H(a_+\psi) &= \hbar\omega \left(a_+ a_- + \frac{1}{2} \right) (a_+\psi) \\ &= \hbar\omega \left[a_+ a_- a_+ + \frac{1}{2} a_+ \right] \psi \\ &= \hbar\omega a_+ \left[a_- a_+ + \frac{1}{2} \right] \psi \\ &= \hbar\omega a_+ \left[a_- a_+ - \frac{1}{2} + 1 \right] \psi \\ &= \hbar\omega a_+ (H + \hbar\omega) \psi \\ &= (E + \hbar\omega) (a_+\psi) \end{aligned} \quad (11.12)$$

Similarly,

$$\begin{aligned} H(a_-\psi) &= \hbar\omega \left(a_- a_+ - \frac{1}{2} \right) (a_-\psi) \\ &= \hbar\omega \left[a_- a_+ a_- - \frac{1}{2} a_- \right] \psi \\ &= \hbar\omega a_- \left[a_+ a_- - \frac{1}{2} \right] \psi \\ &= \hbar\omega a_- \left[a_+ a_- + \frac{1}{2} - 1 \right] \psi \\ &= \hbar\omega a_- (H - \hbar\omega) \psi \\ &= (E - \hbar\omega) (a_-\psi) \end{aligned} \quad (11.13)$$

From Eqs. 11.12, 11.13, a_+ is called a **raising operator** while a_- is called a **lowering operator**.

Note that

$$a_+^\dagger = a_- \quad , \quad a_-^\dagger = a_+ \quad (11.14)$$

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

$$\begin{aligned} a_-\psi_0(x) &= 0 \\ \Rightarrow \frac{1}{\sqrt{2\hbar m\omega}} \left(\hbar \frac{d}{dx} + m\omega x \right) \psi_0(x) &= 0 \\ \frac{d\psi_0}{dx} &= -\frac{m\omega}{\hbar} x \psi_0 \\ \int \frac{d\psi_0}{\psi_0} &= -\frac{m\omega}{\hbar} \int x dx \\ \ln \psi_0 &= -\frac{m\omega}{2\hbar} x^2 + c \\ \psi_0 &= C_1 e^{-m\omega x^2/2\hbar} , \quad C_1 = \ln c . \end{aligned} \quad (11.16)$$

Here, C_1 can be calculated using the normalizing condition

$$\begin{aligned} \int_{-\infty}^{\infty} dx |\psi_0|^2 &= 1 \\ \Rightarrow |C_1|^2 \int_{-\infty}^{\infty} e^{-m\omega x^2/\hbar} &= 1 \\ \Rightarrow |C_1|^2 \sqrt{\frac{\pi\hbar}{m\omega}} &= 1 \quad (\text{using } \int_{-\infty}^{\infty} dx e^{-x^2/a^2} = \sqrt{\frac{\pi}{a}}) , \\ \therefore C_1 &= \left(\frac{m\omega}{\pi\hbar} \right)^{1/4} \end{aligned} \quad (11.17)$$

$$\Rightarrow \psi_0 = \left(\frac{m\omega}{\pi\hbar} \right)^{1/4} e^{-m\omega x^2/2\hbar} \quad (11.18)$$

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 ,

$$H\psi_0 = E_0\psi_0 \quad (11.19)$$

$$\begin{aligned} \hbar\omega \left(a_+ a_- + \frac{1}{2} \right) \psi_0 &= E_0 \psi_0 \\ \hbar\omega a_+ (a_- \psi_0) + \frac{1}{2} \hbar\omega \psi_0 &= E_0 \psi_0 \\ \implies \frac{1}{2} \hbar\omega \psi_0 &= E_0 \psi_0 \quad (\text{using } a_- \psi_0 = 0) \\ \therefore E_0 &= \frac{1}{2} \hbar\omega \end{aligned} \quad (11.20)$$

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

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.(11.15).



Turtles
all the
way
down!



Figure 11.2: 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)

$$\begin{aligned}
 H\psi_1 &= (a_+ a_- + \frac{1}{2})\hbar\omega A_1 a_+ \psi_0 \\
 E_1 \psi_1 &= A_1 \hbar\omega (a_+ a_- a_+) \psi_0 + \frac{1}{2} \hbar\omega \psi_1 \\
 &= A_1 \hbar\omega a_+ (1 + a_+ a_-) \psi_0 + \frac{1}{2} \hbar\omega \psi_1 \quad (\text{using } [a_-, a_+] = 1) , \\
 &= A_1 \hbar\omega a_+ \psi_0 + \frac{1}{2} \hbar\omega \psi_1 \quad (\text{as } a_- \psi_0 = 0) , \\
 &= (1 + \frac{1}{2}) \hbar\omega A_1 a_+ \psi_0 \\
 &= \frac{3}{2} \hbar\omega \psi_1 \\
 \therefore E_1 &= \frac{3}{2} \hbar\omega . \tag{11.21}
 \end{aligned}$$

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 $E_2 = (2 + 1/2)\hbar\omega = 5\hbar\omega/2$:

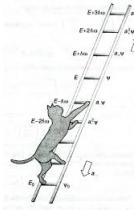


Turtles
all the
way
down!

$$\begin{aligned}
 H\psi_2 &= (a_+ a_- + \frac{1}{2})\hbar\omega A_2 a_+ \psi_1 \\
 E_2 \psi_2 &= A_2 \hbar\omega (a_+ a_- a_+) \psi_1 + \frac{1}{2} \hbar\omega \psi_2 \\
 &= A_2 \hbar\omega a_+ (1 + a_+ a_-) \psi_1 + \frac{1}{2} \hbar\omega \psi_2 \quad (\text{using } [a_-, a_+] = 1) , \\
 &= A_2 \hbar\omega a_+ (1 + 1) \psi_1 + \frac{1}{2} \hbar\omega \psi_2 \quad (\text{as } a_+ a_- \psi_1 = (E_1 - \frac{\hbar\omega}{2}) \psi_1 = \hbar\omega \psi_1) , \\
 &= (2 + \frac{1}{2}) \hbar\omega A_2 a_+ \psi_1 \\
 &= \frac{5}{2} \hbar\omega \psi_2 \\
 \therefore E_2 &= \frac{5}{2} \hbar\omega . \tag{11.22}
 \end{aligned}$$

Continuing in this way till the n th 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
spec-
trum.

$$\begin{aligned}
 H\psi_n &= (a_+a_- + \frac{1}{2})\hbar\omega A_n a_+ \psi_{n-1} \\
 E_n \psi_n &= A_n \hbar\omega (a_+a_- a_+) \psi_{n-1} + \frac{1}{2}\hbar\omega \psi_n \\
 &= A_2 \hbar\omega a_+ (1 + a_+ a_-) \psi_{n-1} + \frac{1}{2}\hbar\omega \psi_n \quad (\text{using } [a_-, a_+] = 1) , \\
 &= A_2 \hbar\omega a_+ (n - 1 + 1) \psi_{n-1} + \frac{1}{2}\hbar\omega \psi_n \quad (\text{as } a_+ a_- \psi_{n-1} = (E_{n-1} - \frac{\hbar\omega}{2}) \psi_{n-1} = (n - 1) \hbar\omega \psi_{n-1}) , \\
 &= (n + \frac{1}{2}) \hbar\omega A_n a_+ \psi_{n-1} \\
 &= (n + \frac{1}{2}) \hbar\omega \psi_n \\
 \therefore E_n &= (n + \frac{1}{2}) \hbar\omega . \tag{11.23}
 \end{aligned}$$

In this way, we find the entire spectrum for $n \in \mathbb{Z}$ 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

$$a_+ \psi_n = c_n \psi_{n+1} , \tag{11.24a}$$

$$a_- \psi_n = d_n \psi_{n-1} . \tag{11.24b}$$

Consider,

$$\begin{aligned}
 \int_{-\infty}^{\infty} (a_+ \psi_n)^* (a_+ \psi_n) dx &= \int_{-\infty}^{\infty} \psi_n^* a_- a_+ \psi_n dx \quad (\text{as } (a_+ \psi_n)^* = \psi_n^* a_-, a_+^\dagger = a_-) , \\
 |c_n|^2 \int_{-\infty}^{\infty} \psi_{n+1}^* \psi_{n+1} dx &= \int_{-\infty}^{\infty} \psi_n^* \left(\frac{1}{\hbar\omega} H + \frac{1}{2} \right) \psi_n dx \quad (\text{using } H = (a_- a_+ - 1/2) \hbar\omega) , \\
 |c_n|^2 &= \frac{1}{\hbar\omega} \int_{-\infty}^{\infty} \psi_n^* H \psi_n dx + \frac{1}{2} \int_{-\infty}^{\infty} \psi_n^* \psi_n dx \\
 &= \left(n + \frac{1}{2} \right) + \frac{1}{2} \quad (\text{using } \int_{-\infty}^{\infty} \psi_n^* \psi_n dx = 1) , \\
 \Rightarrow |c_n|^2 &= n + 1 \\
 \therefore c_n &= \sqrt{n + 1} \tag{11.25}
 \end{aligned}$$

Similarly consider,

$$\begin{aligned}
\int_{-\infty}^{\infty} (a_- \psi_n)^* (a_- \psi_n) dx &= \int_{-\infty}^{\infty} \psi_n^* a_+ a_- \psi_n dx \quad (\text{as } (a_- \psi_n)^* = \psi_n^* a_+, a_+^\dagger = a_-) , \\
|d_n|^2 \int_{-\infty}^{\infty} \psi_{n-1}^* \psi_{n-1} dx &= \int_{-\infty}^{\infty} \psi_n^* \left(\frac{1}{\hbar\omega} H - \frac{1}{2} \right) \psi_n dx \quad (\text{using } H = (a_+ a_- + 1/2)\hbar\omega) , \\
|d_n|^2 &= \frac{1}{\hbar\omega} \int_{-\infty}^{\infty} \psi_n^* H \psi_n dx - \frac{1}{2} \int_{-\infty}^{\infty} \psi_n^* \psi_n dx \\
&= \left(n + \frac{1}{2} \right) - \frac{1}{2} \quad (\text{using } \int_{-\infty}^{\infty} \psi_n^* \psi_n dx = 1) , \\
\Rightarrow |d_n|^2 &= n \\
\therefore d_n &= \sqrt{n} \tag{11.26}
\end{aligned}$$

So, finally Eqs. 11.24a, 11.24b can be written as

$$\psi_{n+1} = (n+1)^{-1/2} a_+ \psi_n \tag{11.27a}$$

$$\psi_{n-1} = n^{-1/2} a_- \psi_n \tag{11.27b}$$

Using Eqs. 11.27, ψ_n can be written in terms of ψ_0 as

$$\psi_n = A_n (a_+)^n \psi_0 (x) = \frac{(a_+)^n}{\sqrt{n!}} \psi_0 \tag{11.28}$$

$$\Rightarrow A_n = \frac{1}{\sqrt{n!}} \tag{11.29}$$

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

$$\begin{aligned}
\int_{-\infty}^{\infty} (a_{-}\psi_m)^* a_{-}\psi_n &= \int_{-\infty}^{\infty} \psi_m^* a_{+} a_{-}\psi_n = \int_{-\infty}^{\infty} \psi_m^* \left(\frac{1}{\hbar\omega} H - \frac{1}{2} \right) \psi_n \\
\int_{-\infty}^{\infty} (a_{-}\psi_m)^* a_{-}\psi_n &= n \int_{-\infty}^{\infty} \psi_m^* \psi_n dx \\
\int_{-\infty}^{\infty} (a_{+}a_{-}\psi_m)^* \psi_n &= n \int_{-\infty}^{\infty} \psi_m^* \psi_n dx \quad (\text{as } (a_{-}\psi_m)^* a_{-} = (a_{+}a_{-}\psi_m)^*) \\
\int_{-\infty}^{\infty} \left(\left(\frac{1}{\hbar\omega} H - \frac{1}{2} \right) \psi_m \right)^* \psi_n &= m \int_{-\infty}^{\infty} \psi_m^* \psi_n dx = n \int_{-\infty}^{\infty} \psi_m^* \psi_n dx \\
\therefore (m-n) \int_{-\infty}^{\infty} \psi_m^* \psi_n dx &= 0 \\
\Rightarrow \int_{-\infty}^{\infty} \psi_m^* \psi_n dx &= \delta_{mn} \tag{11.30}
\end{aligned}$$

Figure 11.3 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 ψ s decay exponentially to zero as $x \rightarrow \pm\infty$. 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 $n \gg 1$, the “envelope” of the quantum probability distribution begins to resemble the large amplitude oscillations observed classically (as shown as the dashed line in Fig.11.4) [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.

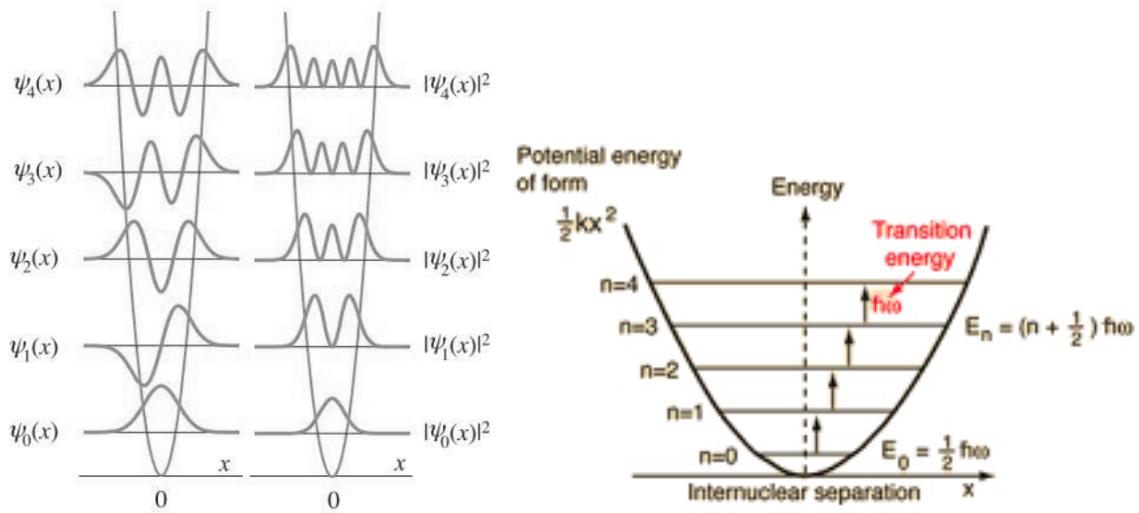


Figure 11.3: First few eigenstates (Left) and Energy Spectrum (Right) of the quantum mechanical harmonic oscillator.

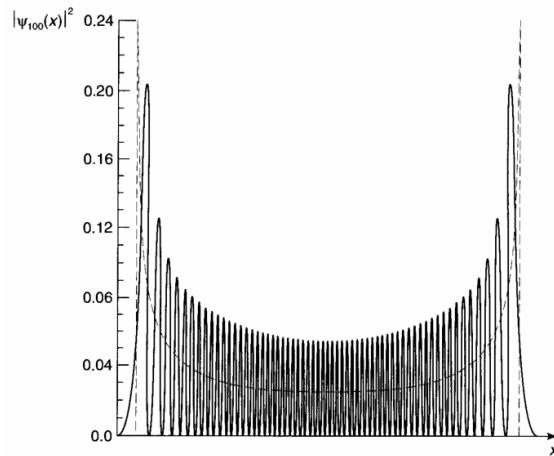


Figure 11.4: Probability distribution of high energy eigenstate ψ_{100} .

11.2 Analytic Method

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

$$-\frac{\hbar^2}{2m} \frac{d^2\psi(x)}{dx^2} + \frac{1}{2}m\omega^2x^2\psi(x) = E\psi(x) \quad (11.31)$$

Choosing $\xi = \sqrt{\frac{m\omega}{\hbar}}x$, $K = 2E/\hbar\omega$, Eq. 11.31 can be written as

$$\frac{d^2\psi}{d\xi^2} = (\xi^2 - K)\psi \quad (11.32)$$

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

$$\begin{aligned} \frac{d^2\psi}{d\xi^2} &\approx \xi^2\psi \\ \implies \psi(\xi) &\approx Ae^{-\xi^2/2} + Be^{\xi^2/2} \end{aligned} \quad (11.33)$$

$$\implies \psi(\xi) \approx Ae^{-\xi^2/2} \quad (11.34)$$

Here, $B = 0$ in Eq. 11.33, else $e^{\xi^2/2} \rightarrow \infty$ in the limit $x \rightarrow \pm\infty$ will cause ψ to diverge. Let us consider the following ansatz solution:

$$\psi(\xi) = h(\xi)e^{-\xi^2/2} \quad (11.35)$$

Note that we have already found the Gaussian part of the solution for ψ (as observed earlier from the algebraic solution).

Now, using Eq. 11.35, we obtain

$$\begin{aligned} \frac{d\psi}{d\xi} &= \left(\frac{dh(\xi)}{d\xi} - \xi h(\xi) \right) e^{-\xi^2/2} \\ \frac{d^2\psi}{d\xi^2} &= \left(\frac{d^2h(\xi)}{d\xi^2} - 2\xi \frac{dh(\xi)}{d\xi} + (\xi^2 - 1)h(\xi) \right) e^{-\xi^2/2} \end{aligned} \quad (11.36)$$

Using Eq. 11.36 in Eq. 11.32,

$$\begin{aligned} \left(\frac{d^2h(\xi)}{d\xi^2} - 2\xi \frac{dh(\xi)}{d\xi} + (\xi^2 - 1)h(\xi) \right) e^{-\xi^2/2} &= (\xi^2 - K)h(\xi)e^{-\xi^2/2} \\ \Rightarrow \frac{d^2h(\xi)}{d\xi^2} - 2\xi \frac{dh(\xi)}{d\xi} + (K - 1)h(\xi) &= 0 \end{aligned} \quad (11.37)$$

Solving Eq. 11.37 by the Frobenius method of a power series solution:

$$h(\xi) = a_0 + a_1\xi + a_2\xi^2 + \cdots = \sum_{j=0}^{\infty} a_j \xi^j \quad (11.38)$$

$$\frac{dh(\xi)}{d\xi} = a_1 + 2a_2\xi + 3a_3\xi^2 + \cdots = \sum_{j=0}^{\infty} j a_j \xi^{j-1} \quad (11.39)$$

$$\frac{d^2h(\xi)}{d\xi^2} = 2a_2 + 6a_3\xi + \cdots = \sum_{j=0}^{\infty} (j+1)(j+2) a_{j+2} \xi^j \quad (11.40)$$

Putting Eqs. 11.38, 11.39, 11.40 in Eq. 11.37, we get

$$\sum_{j=0}^{\infty} [(j+1)(j+2) a_{j+2} - 2ja_j + (K-1)a_j] \xi^j = 0 \quad (11.41)$$

From the uniqueness of the power series expansion,

$$(j+1)(j+2) a_{j+2} - 2ja_j + (K-1)a_j = 0 \quad (11.42)$$

leading to the **recursion relation**

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

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

$$h_{even}(\xi) = a_0 + a_2\xi^2 + a_4\xi^4 + \cdots, \quad (11.43a)$$

$$h_{odd}(\xi) = a_1\xi + a_3\xi^3 + a_5\xi^5 + \cdots, \quad (11.43b)$$

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

Now, for $j \gg 1, K$, 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 $\xi \gg 1$,

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

But since $\psi = h(\xi)e^{-\xi^2/2}$, we can see that with the $h(\xi)$ found above, $\psi \sim e^{\xi^2/2} \rightarrow \infty$ as $\xi \rightarrow \pm\infty$! 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:

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

Put another way, for a truly ∞ series, the polynomial part of ψ always dominates and leads to a solution that diverges as $\xi \rightarrow \pm\infty$. 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

$$\begin{aligned} \frac{2n+1-K}{(n+1)(n+2)} &= 0 \\ \implies K &= 2n+1 = \frac{2E}{\hbar\omega} \\ \therefore E &= \left(n + \frac{1}{2}\right)\hbar\omega \end{aligned} \tag{11.44}$$

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 $\psi(x) \rightarrow 0$ for $x \rightarrow \pm\infty$) only appear for certain E values (all the other values of E give solutions but these diverge to $\pm\infty$ as $x \rightarrow \pm\infty$).

Thus for the allowed values of K ,

$$a_{j+2} = -\frac{2(n-j)}{(j+1)(j+2)} a_j . \tag{11.45}$$

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

$$\begin{aligned} h_0(\xi) &= a_0 \\ \implies \psi_0(\xi) &= a_0 e^{-\xi^2/2} . \end{aligned} \tag{11.46}$$

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

$$\begin{aligned} h_1(\xi) &= a_1 \xi \\ \psi_1(\xi) &= a_1 \xi e^{-\xi^2/2} . \end{aligned} \tag{11.47}$$

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 $a_4 = 0 = a_6 = \dots$) gives

$$\begin{aligned} h_2(\xi) &= a_0 (1 - 2\xi^2) \\ \psi_2(\xi) &= a_0 (1 - 2\xi^2) e^{-\xi^2/2} . \end{aligned} \tag{11.48}$$

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

$$\psi_n(\xi) = \left(\frac{m\omega}{\pi\hbar}\right)^{1/4} \frac{1}{\sqrt{2^n n!}} H_n(\xi) e^{-\xi^2/2} \quad (11.49)$$

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.

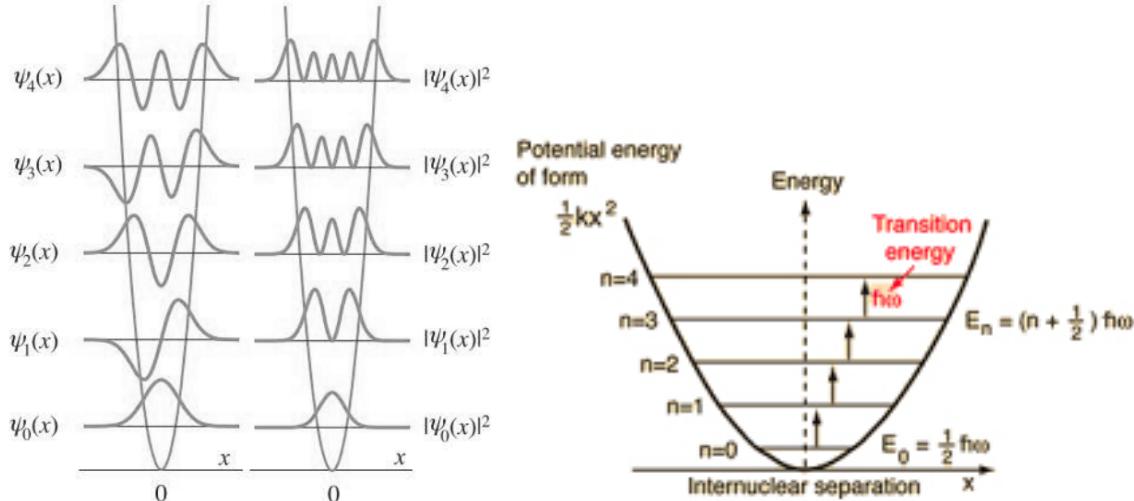


Figure 11.5: First few eigenstates (Left) and Energy Spectrum (Right) of the quantum mechanical harmonic oscillator.

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 $x \rightarrow \pm\infty$ 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 12

Free particle

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

$$-\frac{\hbar^2}{2m} \frac{d^2\psi}{dx^2} = E\psi \quad (12.1)$$

$$\psi'' = -k^2\psi \quad \text{where } k = \sqrt{\frac{2mE}{\hbar^2}}, E \geq 0 \quad (12.2)$$

The general solution is:

$$\begin{aligned} \Psi(x, t) &= \psi(x) e^{-iEt/\hbar} \\ &= (Ae^{ikx} + Be^{-ikx}) e^{-iEt/\hbar} \\ &= Ae^{ik(x - \frac{\hbar k}{2m}t)} + Be^{-ik(x + \frac{\hbar k}{2m}t)} \end{aligned} \quad (12.3)$$

Eq. 12.3 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. 12.3 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. 12.3 (shown in Eq. 12.4) with different velocities so that they are normalizable. Indeed, this localizes the wave-function giving rise to the idea of “wave-packet”.

$$\Psi(x, t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \phi(k) e^{i(kx - \frac{\hbar k^2}{2m}t)} dk \quad (12.4)$$

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).

$$\Psi(x, 0) = \int_{-\infty}^{\infty} \frac{\phi(k)}{\sqrt{2\pi}} e^{ikx} dk \quad (12.5)$$

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

$$F(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x) e^{-ikx} dx \quad (12.6)$$

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} F(k) e^{ikx} dk \quad (12.7)$$

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

$$\phi(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-ikx} \Psi(x, 0) dx \quad (12.8)$$

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

$$\omega = \frac{\hbar k^2}{2m} \quad (12.9)$$

$$v_{ph} = \frac{\hbar k}{2m} \quad (12.10)$$

$$v_{gr} = \frac{\hbar k}{m} = 2v_{ph} \quad (12.11)$$

(Revise your concepts of phase and group velocity!)

Chapter 13

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.

13.1 Step potential

Consider a step potential as shown in figure 13.1 given by

$$V(x) = \begin{cases} 0 & \text{if } x < 0 \quad (\text{region I}) \\ V_0 & \text{if } x \geq 0 \quad (\text{region II}) \end{cases} \quad (13.1)$$

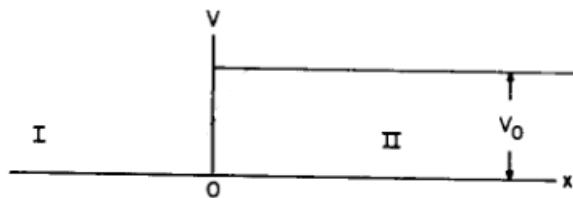


Figure 13.1: A schematic diagram of $V(x)$

Classically, 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:

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

Boundary conditions:

$$(i) \quad \psi(x \rightarrow 0^-) = \psi(x \rightarrow 0^+) \quad (13.2a)$$

$$(ii) \quad \left. \frac{d\psi}{dx} \right|_{x \rightarrow 0^-} = \left. \frac{d\psi}{dx} \right|_{x \rightarrow 0^+} \quad (13.2b)$$

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

$$\psi(x) = \begin{cases} Ae^{ik_1 x} + Be^{-ik_1 x} & x < 0, k_1 = \sqrt{\frac{2mE}{\hbar^2}} \\ Ce^{ik_2 x} + De^{-ik_2 x} & x \geq 0, k_2 = \sqrt{\frac{2m(E-V_0)}{\hbar^2}} \end{cases} \quad (13.3)$$

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

$$(i) \quad A + B = C \quad (13.4)$$

$$(ii) \quad ik_1(A - B) = ik_2C \quad (13.5)$$

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

$$(i) \quad 1 + B = C \quad (13.6)$$

$$(ii) \quad k_1(1 - B) = k_2C \quad (13.7)$$

$$\implies B = \frac{k_1 - k_2}{k_1 + k_2} \quad \& \quad C = \frac{2k_1}{k_1 + k_2} \quad (13.8)$$

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

$$\begin{aligned} R &= |r|^2 = B^2 \\ &= \frac{(k_1 - k_2)^2}{(k_1 + k_2)^2} = \frac{\left(\sqrt{E} - \sqrt{E - V_0}\right)^2}{\left(\sqrt{E} + \sqrt{E - V_0}\right)^2} \end{aligned} \quad (13.9)$$

$$\begin{aligned} T &= |t|^2 = C^2 \\ &= \frac{4k_1^2}{(k_1 + k_2)^2} = \frac{4E}{\left(\sqrt{E} + \sqrt{E - V_0}\right)^2} \end{aligned} \quad (13.10)$$

Also, note that

$$R + \frac{k_2}{k_1}T = 1 \quad (13.11)$$

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 $E < V_0$,

$$\psi(x) = \begin{cases} Ae^{ik_1x} + Be^{-ik_1x} & x < 0, k_1 = \sqrt{\frac{2mE}{\hbar^2}} \\ Ce^{k_2x} + De^{-k_2x} & x \geq 0, k_2 = \sqrt{\frac{2m(V_0-E)}{\hbar^2}} \end{cases} \quad (13.12)$$

Unless $C = 0$, $\psi \rightarrow \infty$ as $x \rightarrow \infty$. So, $C = 0$ and $A = 1$ following the same logic as in the previous case. Applying boundary conditions given in Eq. 13.2 we get

$$1 + B = D \quad (13.13)$$

$$ik_1(1 - B) = -k_2D \quad (13.14)$$

$$\implies B = \frac{k_1 - ik_2}{k_1 + ik_2} \quad \& \quad D = \frac{2k_1}{k_1 + ik_2} \quad (13.15)$$

The reflection, transmission amplitudes will be

$$\begin{aligned} R &= |r|^2 = |B|^2 \\ &= \left| \frac{k_1 - ik_2}{k_1 + ik_2} \right|^2 = 1 \end{aligned} \quad (13.16)$$

$$\implies T = \frac{k_1}{k_2}(1 - R) = 0 \quad (13.17)$$

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 $V_0 \rightarrow \infty$, $\psi(x) \rightarrow 0$ 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.

13.2 Potential Barrier & Tunneling

Instead of the step potential shown in Figure 13.1, consider a barrier potential given by 13.18 and shown in Figure 13.2

$$V(x) = \begin{cases} 0 & \text{if } x < 0 \quad (\text{region I}) \\ V_0 & \text{if } 0 \leq x \leq a \quad (\text{region II}) \\ 0 & \text{if } x > a \quad (\text{region III}) \end{cases} \quad (13.18)$$

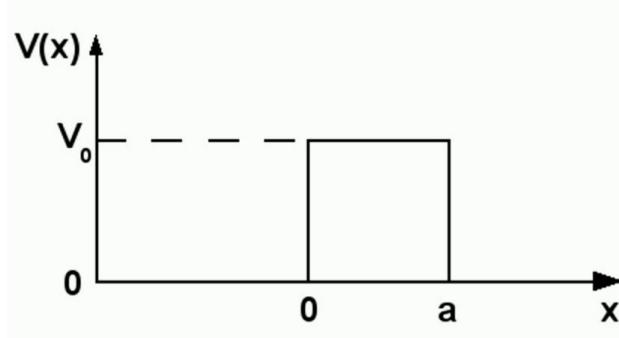


Figure 13.2: A potential barrier

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

$$\psi(x) = \begin{cases} e^{ik_1 x} + r e^{-ik_1 x} & x < 0, k_1 = \sqrt{\frac{2mE}{\hbar^2}} \\ A e^{ik_2 x} + B e^{-ik_2 x} & 0 \leq x \leq a, k_2 = \sqrt{\frac{2m(E-V_0)}{\hbar^2}} \\ t e^{ik_1 x} & x > a, k_1 = \sqrt{\frac{2mE}{\hbar^2}} \end{cases} \quad (13.19)$$

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

$$\psi(x \rightarrow 0-) = \psi(x \rightarrow 0+) \Rightarrow 1 + r = A + B \quad (13.20a)$$

$$\psi'(x \rightarrow 0-) = \psi'(x \rightarrow 0+) \Rightarrow i k_1 (1 - r) = i k_2 (A - B) \quad (13.20b)$$

$$\psi(x \rightarrow a-) = \psi(x \rightarrow a+) \Rightarrow t e^{ik_1 a} = A e^{ik_2 a} + B e^{-ik_2 a} \quad (13.20c)$$

$$\psi'(x \rightarrow a-) = \psi'(x \rightarrow a+) \Rightarrow i k_1 t e^{ik_1 a} = i k_2 (A e^{ik_2 a} - B e^{-ik_2 a}) \quad (13.20d)$$

We set $\mu = k_2/k_1 = \sqrt{1 - V_0/E}$. Then solving the set of Eqs. 13.20 after some lengthy

algebra, we get

$$r = \frac{(1 - \mu^2) \sin(k_2 a)}{(1 + \mu^2) \sin k_2 a + 2i\mu \cos k_2 a} \quad (13.21a)$$

$$t = \frac{2i\mu e^{-ik_1 a}}{(1 + \mu^2) \sin k_2 a + 2i\mu \cos k_2 a} \quad (13.21b)$$

$$A = \frac{i(1 + \mu) e^{-ik_2 a}}{(1 + \mu^2) \sin k_2 a + 2i\mu \cos k_2 a} \quad (13.21c)$$

$$B = \frac{-i(1 - \mu) e^{ik_2 a}}{(1 + \mu^2) \sin k_2 a + 2i\mu \cos k_2 a} \quad (13.21d)$$

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

$$\begin{aligned} T = |t|^2 &= \frac{4\mu^2}{(1 + \mu^2)^2 \sin^2 k_2 a + 4\mu^2 \cos^2 k_2 a} \\ &= \frac{4\mu^2}{4\mu^2 + (1 - \mu^2)^2 \sin^2 k_2 a} = \frac{1}{1 + \frac{1}{4} \left(\frac{1-\mu^2}{\mu}\right)^2 \sin^2 k_2 a} \\ &= \frac{1}{1 + \frac{1}{4} \left(\frac{k_1^2 - k_2^2}{k_1 k_2}\right)^2 \sin^2 k_2 a} \end{aligned} \quad (13.22)$$

Note that the reflection coefficient $R = 1 - T$. Both R and T are plotted in Fig.13.3 as a function of E/V_0 . Whenever $k_2 a = n\pi$ the system is in resonance and there will be perfect

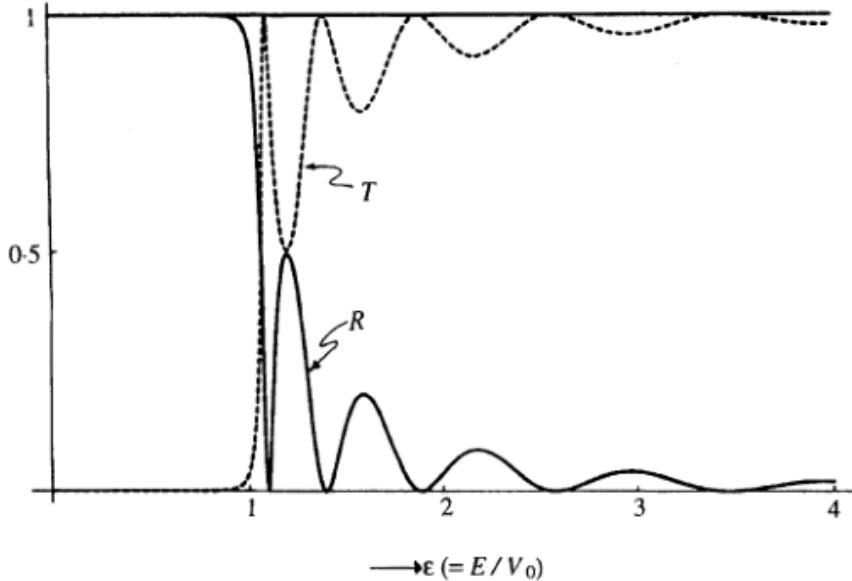


Figure 13.3: Transmission, Reflection coefficients as a function of $\epsilon = E/V_0$

transmission $T = 1$, $R = 0$ (see Fig.13.3). We also see that T oscillates and finally becomes

asymptotically equal to 1 for $E \gg V_0$; this is seen by taking $\mu \rightarrow 1$ in the expression for T in eq.(13.22). The criterion for the n th Resonance is given by the energy E_n

$$k_2 a = \sqrt{\frac{2m(E - V_0)}{\hbar^2}} a = n\pi \quad (13.23)$$

$$E_n = V_0 + \frac{n^2\pi^2\hbar^2}{2ma^2} \quad (13.24)$$

$$= V_0 \left(1 + \frac{n^2\pi^2\hbar^2}{2mV_0a^2} \right)$$

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.(13.23). 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.

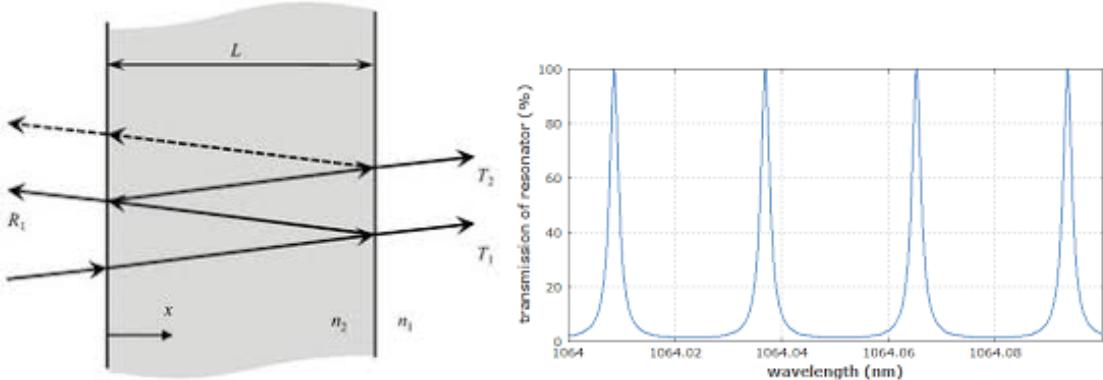


Figure 13.4: (Left) The Fabry-Perot Cavity, displaying the sequence of reflections and transmissions of e-m waves at the two leaky mirrors and (Right) the Transmission coefficient of the F-P cavity as a function the e-m wavelength. Source: The internet.

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

$$\begin{aligned}
\frac{k_1^2 - k_2^2}{k_1 k_2} &\approx \frac{k_1^2}{k_1 k_2} \approx \frac{k_1}{k_2} \\
\implies T &\approx \left(1 + \frac{1}{4} \left(\frac{k_1}{k_2} \right)^2 k_2^2 a^2 \right)^{-1} \\
&\approx 1 - \frac{1}{4} k_1^2 a^2 \approx 1 - \frac{1}{4} \left(\frac{2mV_0}{\hbar^2} \right) a^2 \\
&\approx 1 - \frac{mV_0 a^2}{2\hbar^2}
\end{aligned} \tag{13.25}$$

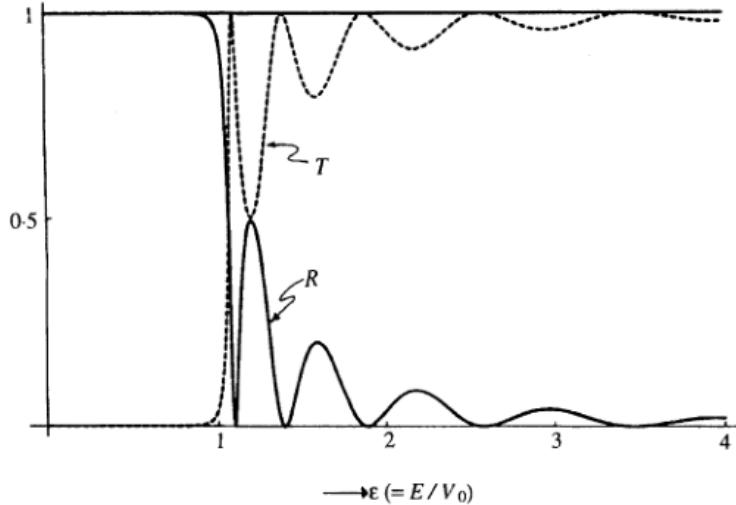


Figure 13.5: Transmission, Reflection coefficients as a function of Energy ϵ

For $E < V_0$, k_2 becomes imaginary

$$\begin{aligned}
k_2 &= iK = \frac{i}{\hbar} \sqrt{2m(V_0 - E)} \\
\implies T = |t|^2 &= \frac{(2k_1/K)^2}{\left(1 - \frac{k_1^2}{K^2} \right)^2 \sinh^2(Ka) + \left(\frac{2k_1}{K} \right)^2 \cosh^2(Ka)}
\end{aligned} \tag{13.26}$$

Thus, we can see that $T \rightarrow 0$ and $R = 1 - T \rightarrow 1$ as $E \rightarrow 0^+$, joining smoothly onto the expressions for T and R given by (13.25) as $E \rightarrow V_0^-$ (see Fig.13.3). An application of the phenomenon of quantum tunneling is realised in the scanning tunneling microscope (STM) shown below in Fig.13.6.

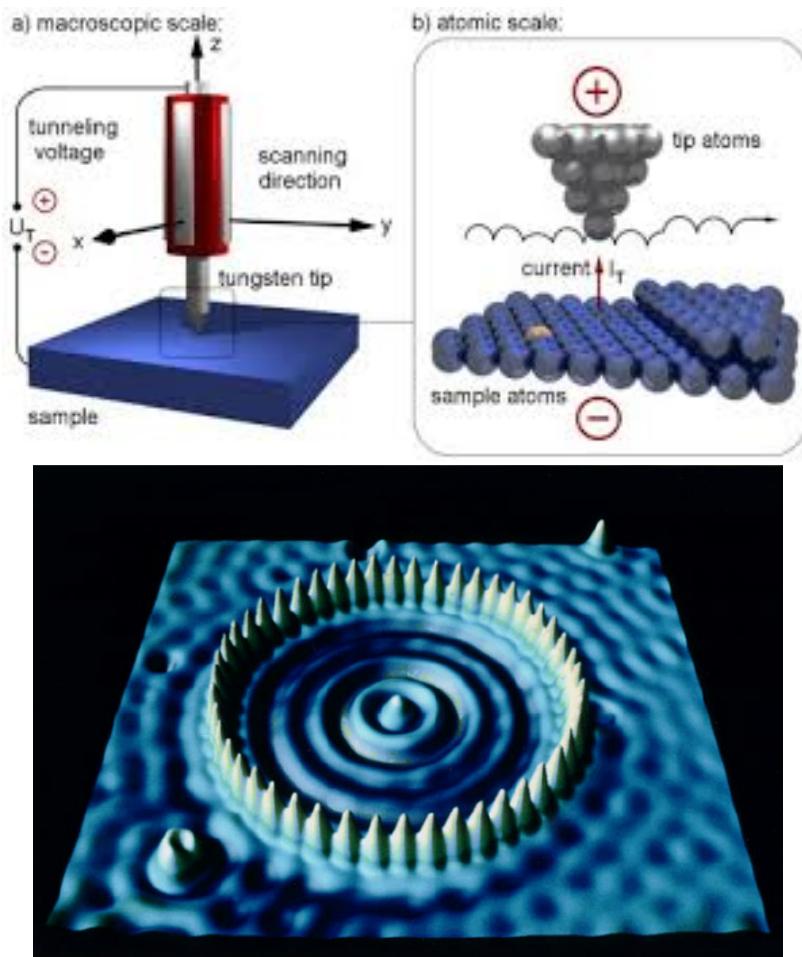


Figure 13.6: (Above) The physics of the scanning tunneling microscope (STM) is based on the phenomenon of quantum tunneling of electrons between the sample surface and the atoms of the STM tip and across a vacuum “barrier”. The tunnel current I_T depends on how easily sample surface electrons are available 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 14

Time Independent Schrödinger Equation (TISE) in 3D

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

$$-\frac{\hbar^2}{2m}\nabla^2\psi(\mathbf{r},t) + V(\mathbf{r})\psi(\mathbf{r},t) = i\hbar\frac{\partial\psi(\mathbf{r},t)}{\partial t} \quad (14.1)$$

$$-\frac{\hbar^2}{2m}\nabla^2\psi(\mathbf{r}) + V(\mathbf{r})\psi(\mathbf{r}) = E\psi(\mathbf{r}) \quad (14.2)$$

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 $V(\mathbf{r})$, e.g., for a central force $F(\mathbf{r})$ which arises from a $V(|\mathbf{r}|)$ is most approximately (and easily) solved using spherical coordinates.

14.1 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.

$$V(\mathbf{r}) = \begin{cases} 0 & \text{when } 0 < x < a \\ & 0 < y < b \\ & 0 < z < c \\ \infty & \text{otherwise} \end{cases} \quad (14.3)$$

Thus inside the box,

$$\begin{aligned} -\frac{\hbar^2}{2m}\nabla^2\psi(\mathbf{r}) &= E\psi(\mathbf{r}) \\ \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) \psi &= -\frac{2mE}{\hbar^2}\psi \end{aligned} \quad (14.4)$$

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

$$\psi(x, y, z) = \psi_1(x)\psi_2(y)\psi_3(z) \quad (14.5)$$

Putting Eq. 14.5 in Eq. 14.4,

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

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

$$\psi_1(x) = A_1 \sin \frac{n_x \pi x}{a}, \quad \psi_2(y) = A_2 \sin \frac{n_y \pi y}{a}, \quad \psi_3(z) = A_3 \sin \frac{n_z \pi z}{a} \quad (14.6)$$

$$E_x = \frac{\hbar^2 \pi^2}{2ma^2} n_x^2, \quad E_y = \frac{\hbar^2 \pi^2}{2mb^2} n_y^2, \quad E_z = \frac{\hbar^2 \pi^2}{2mc^2} n_z^2 \quad (14.7)$$

Then,

$$\psi(\mathbf{r}) = A \sin \left(\frac{n_x \pi x}{a} \right) \sin \left(\frac{n_y \pi y}{b} \right) \sin \left(\frac{n_z \pi z}{c} \right) \quad (14.8)$$

$$\begin{aligned} E &= E_x + E_y + E_z \\ &= \frac{\hbar^2 \pi^2}{2m} \left(\frac{n_x^2}{a^2} + \frac{n_y^2}{b^2} + \frac{n_z^2}{c^2} \right) \end{aligned} \quad (14.9)$$

where $A = A_1 A_2 A_3$. For the special case of a cube ($a = b = c$)

$$\psi(\mathbf{r}) = A \sin \left(\frac{n_x \pi x}{a} \right) \sin \left(\frac{n_y \pi y}{a} \right) \sin \left(\frac{n_z \pi z}{a} \right) \quad (14.10)$$

$$E = \frac{\hbar^2 \pi^2}{2ma^2} (n_x^2 + n_y^2 + n_z^2) \quad (14.11)$$

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 pairwise 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 15

The Angular Momentum problem: first passage

15.1 Introduction

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

$$\mathbf{L} = \mathbf{r} \times \mathbf{p} \quad (15.1)$$

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

$$\begin{aligned} \mathbf{L} &= \mathbf{r} \times \mathbf{p} \\ &= L_x \hat{x} + L_y \hat{y} + L_z \hat{z} \end{aligned} \quad (15.2)$$

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?

$$\begin{aligned} L_x^\dagger &= (yp_z)^\dagger - (zp_y)^\dagger = p_z^\dagger y^\dagger - p_y^\dagger z^\dagger \\ &= p_z y - p_y z = yp_z - zp_y = L_x \\ L_y^\dagger &= L_y, \quad L_z^\dagger = L_z \\ \Rightarrow \mathbf{L}^\dagger &= \mathbf{L} \end{aligned} \quad (15.3)$$

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

$$[R_\alpha, p_\beta] = i\hbar\delta_{\alpha\beta} \quad (15.4)$$

$$[A + B, C] = [A, C] + [B, C] \quad (15.5)$$

$$[AB, C] = A[B, C] + [A, C]B \quad (15.6)$$

Thus,

$$\begin{aligned}
[L_x, L_y] &= [yp_z - zp_y, zp_x - xp_z] \\
&= [yp_z, zp_x] - [zp_y, zp_x] - [yp_z, xp_z] + [zp_y, xp_z] \\
&= y [p_z, zp_x] + [z, xp_z] p_y \quad (\text{as the second and third commutators vanish identically}) \\
&= y [p_z, z] p_x + x [z, p_z] p_y \\
&= i\hbar (xp_y - yp_x) = i\hbar L_z
\end{aligned} \tag{15.7}$$

Similarly, we can show that

$$[L_y, L_z] = i\hbar L_x , \tag{15.8}$$

$$[L_z, L_x] = i\hbar L_y . \tag{15.9}$$

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

$$\begin{aligned}
[L^2, L_z] &= [L_x^2 + L_y^2 + L_z^2, L_z] \\
&= [L_x^2, L_z] + [L_y^2, L_z] \\
&= L_x [L_x, L_z] + [L_x, L_z] L_x + L_y [L_y, L_z] + [L_y, L_z] L_y \\
&= -i\hbar L_x L_y - i\hbar L_y L_x + i\hbar L_y L_x + i\hbar L_x L_y = 0 , \\
[L^2, L_x] &= 0 , \quad [L^2, L_y] = 0 .
\end{aligned} \tag{15.10}$$

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 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, given that $[L^2, L_x] = 0$ and $[L^2, L_y] = 0$, we could have equivalently chosen the simultaneous eigenstates for the pair (L^2, L_x) or (L^2, L_y) instead of for the pair (L^2, L_z) .

15.2 Total Angular Momentum

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

$$\mathbf{J} = \mathbf{L} + \mathbf{S} \quad (15.11)$$

\mathbf{L} = Orbital Angular Momentum

\mathbf{S} = Spin Angular Momentum

Here, \mathbf{L} literally relates to the classical orbital motion of a particle while \mathbf{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, and you will surely encounter it somewhere on your journey into quantum mechanics. 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.

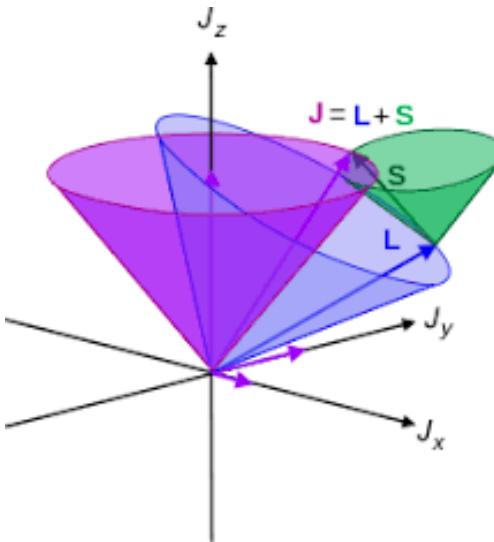


Figure 15.1: The angular momentum in quantum mechanics. Source: the internet.

Generally, in the absence of any physics that couples the \mathbf{L} and \mathbf{S} angular momenta, the commutation relations given earlier for the components of \mathbf{L} go through as given, and a similar algebra is found to separately exist for \mathbf{S} as well. Thus, we may write the commutator algebra for the components of the total angular momentum vector operator \vec{J} as

$$[J_x, J_y] = i\hbar J_z \quad (15.12)$$

$$[J_y, J_z] = i\hbar J_x \quad (15.13)$$

$$[J_z, J_x] = i\hbar J_y \quad (15.14)$$

$$[J^2, J_\alpha] = 0 \quad \text{for } \alpha = x, y, z \quad (15.15)$$

We are going to work here onwards with the eigenstates of J^2 and J_z . In order to discover these eigenstates, we define the ladder operators as

$$J_+ = J_x + iJ_y \quad \text{Creation (raising) operator} \quad (15.16a)$$

$$J_- = J_x - iJ_y \quad \text{Annihilation (lowering) operator} \quad (15.16b)$$

$$[J_+, J_-] = -i[J_x, J_y] + i[J_y, J_x] = 2\hbar J_z . \quad (15.16c)$$

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 $J^2 |\psi\rangle = \alpha |\psi\rangle$, $J_z |\psi\rangle = \beta |\psi\rangle$, assuming $J_+ |\psi\rangle = |\phi\rangle$ (where $|\phi\rangle$ is another normalised eigenstate of J^2 and J_z), let us try calculating $J^2 |\phi\rangle$ and $J_z |\phi\rangle$. Note that

$$[J^2, J_\pm] = 0 \quad \text{as} \quad [J^2, J_x] = 0 = [J^2, J_y] . \quad (15.17)$$

Thus,

$$J^2 |\phi\rangle = J^2 (J_+ |\psi\rangle) = J_+ (J^2 |\psi\rangle) = \alpha J_+ |\psi\rangle = \alpha |\phi\rangle .$$

This shows that $|\phi\rangle$ is an eigenstate of J^2 with the same eigenvalue α as that of the state $|\psi\rangle$, i.e., the action of J_+ on $|\psi\rangle$ does not affect the eigenvalue of J^2 ! However, $[J_z, J_+] \neq 0$. In fact,

$$\begin{aligned} [J_z, J_+] &= [J_z, J_x + iJ_y] \\ &= [J_z, J_x] + i[J_z, J_y] \\ &= i\hbar J_y + i(-i\hbar) J_x \\ &= \hbar (J_x + iJ_y) = \hbar J_+ \end{aligned} \quad (15.18)$$

$$[J_z, J_-] = \hbar J_- \quad (15.19)$$

Calculating $J_z |\phi\rangle$:

$$\begin{aligned} J_z |\phi\rangle &= J_z J_+ |\psi\rangle = J_+ J_z |\psi\rangle + \hbar J_+ |\psi\rangle \\ &= \beta J_+ |\psi\rangle + \hbar J_+ |\psi\rangle \\ &= (\beta + \hbar) |\phi\rangle \end{aligned} \quad (15.20)$$

This means J_+ transforms the state $|\psi\rangle$ (with eigenvalues α , β for J^2 , J_z operators) into a new eigenstate $|\phi\rangle$ 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 $|\psi_{max}\rangle$ such that

$$J_+ |\psi_{max}\rangle = 0 \quad (15.21)$$

In the same way, we can argue that there must exist a $|\psi_{min}\rangle$ such that

$$J_- |\psi_{min}\rangle = 0 \quad (15.22)$$

First, note that we can write J^2 in terms of J_\pm and J_z as follows

$$J^2 = J_x^2 + J_y^2 + J_z^2 \quad (15.23a)$$

$$= J_- J_+ + \hbar J_z + J_z^2 \quad (15.23b)$$

$$= J_+ J_- - \hbar J_z + J_z^2, \quad (15.23c)$$

where we have used the fact that

$$\begin{aligned} J_+ J_- &= (J_x + iJ_y)(J_x - iJ_y), \\ &= Jx^2 + J_y^2 - i(J_x J_y - J_y J_x), \\ &= Jx^2 + J_y^2 + \hbar J_z, \\ \Rightarrow Jx^2 + J_y^2 &= J_+ J_- - \hbar J_z. \end{aligned} \quad (15.24)$$

$$\begin{aligned} \text{Also, } J_- J_+ &= (J_x - iJ_y)(J_x + iJ_y), \\ &= Jx^2 + J_y^2 + i(J_x J_y - J_y J_x), \\ &= Jx^2 + J_y^2 - \hbar J_z, \\ \Rightarrow Jx^2 + J_y^2 &= J_- J_+ + \hbar J_z. \end{aligned} \quad (15.25)$$

$$(15.26)$$

Now, we can compute the values of α and β as follows

$$\begin{aligned} J^2 |\psi_{max}\rangle &= (J_- J_+ + J_z^2 + \hbar J_z) |\psi_{max}\rangle \\ \alpha |\psi_{max}\rangle &= 0 + \beta_{max}^2 |\psi_{max}\rangle + \hbar \beta_{max} |\psi_{max}\rangle \\ \therefore \alpha &= \beta_{max}^2 + \hbar \beta_{max} \end{aligned} \quad (15.27)$$

Similarly, we find by operating $J^2 |\psi_{min}\rangle$

$$\alpha = \beta_{min}^2 - \hbar \beta_{min} \quad (15.28)$$

From Eqs. (15.27), (15.28),

$$\begin{aligned} \beta_{min}^2 - \hbar \beta_{min} &= \beta_{max}^2 + \hbar \beta_{max} \\ \beta_{max}^2 - \beta_{min}^2 + \hbar (\beta_{max} + \beta_{min}) &= 0 \\ (\beta_{max} - \beta_{min} + \hbar)(\beta_{max} + \beta_{min}) &= 0 \\ \implies \beta_{max} &= -\beta_{min} \end{aligned} \quad (15.29)$$

as the other solution $\beta_{max} = \beta_{min} - \hbar$ does not make physical sense. Also, from the idea of ladder operators for J_z ,

$$\beta_{max} = \beta_{min} + n\hbar, \quad n \in \mathbb{Z} \quad (15.30)$$

$$\therefore \beta_{max} = \frac{n\hbar}{2} \quad (15.31)$$

The first line of the above equation tells us that applying the raising operator an integer n number of times should take us from the state $|\psi_{min}\rangle$ to the state $|\psi_{max}\rangle$. Now, defining $j = n/2$, $n \in \mathbb{Z}$, from Eq. (15.27), we obtain

$$\alpha = \hbar^2 j(j+1), \quad (15.32)$$

$$\beta = m_j \hbar, \quad m_j = -j, -j+1, \dots, j-1, j \quad (15.33)$$

Therefore, we can write

$$J^2 |\psi\rangle = \hbar^2 j(j+1) |\psi\rangle, \quad \text{where } j = \frac{n}{2}, n \in \mathbb{Z}, \quad (15.34)$$

$$J_z |\psi\rangle = m_j \hbar |\psi\rangle, \quad \text{where } m_j = -j, -j+1, \dots, j-1, j. \quad (15.35)$$

These relations are visualised in Fig.15.2 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!

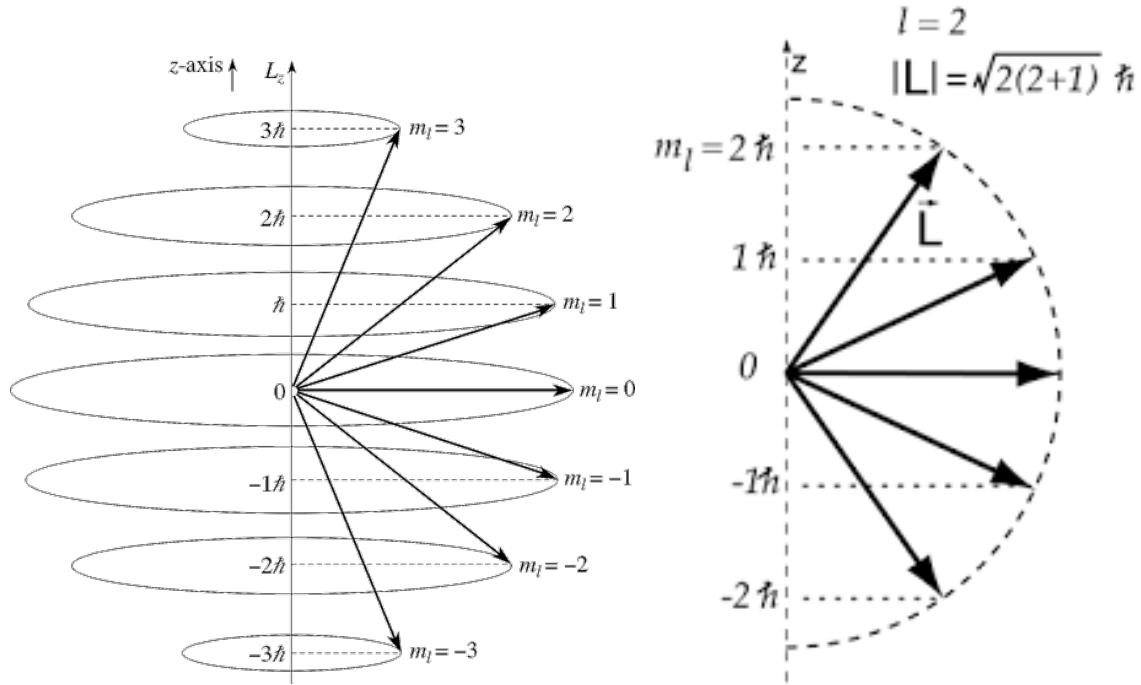


Figure 15.2: The angular momentum in quantum mechanics for the case of $n = 6$, $l = 3$ (left) and $n = 4$, $l = 2$ (right). Source: the internet.

Thus, we are now in a position to denote the common eigenstates of J^2 and J_z are denoted by the good quantum numbers j and m_j : $|\psi\rangle \equiv |j, m_j\rangle$.

Finally, note that we have found that for $|\psi\rangle = |j, m_j\rangle$, we must have $|\phi\rangle \propto |j, m_j \pm 1\rangle$. This then means that we can write

$$J_{\pm} |j, m_j\rangle = C_{j, m_j}^{\pm} |j, m_j \pm 1\rangle, \quad (15.36)$$

where the constants C_{j,m_j}^\pm can be obtained by using the relations eq.(15.23b) and eq.(15.23c)

$$C_{j,m_j}^\pm = \hbar \sqrt{j(j+1) - m_j(m_j \pm 1)} . \quad (15.37)$$

To see how this form of C_{j,m_j}^\pm is obtained, let us proceed as follows. Note that

$$\begin{aligned} J_+ |j, m_j\rangle &= C_{j,m_j}^+ |j, m_j + 1\rangle , \\ \implies (J_+ |j, m_j\rangle)^\dagger &= \langle j, m_j | (J_+)^\dagger = \langle j, m_j | J_- = C_{j,m_j}^{+\ast} \langle j, m_j + 1 | . \end{aligned} \quad (15.38)$$

Therefore, by using relation $J^2 = J_- J_+ + \hbar J_z + J_z^2$ (eq.(15.23b)), we have

$$\begin{aligned} \langle j, m_j | J_- J_+ | j, m_j \rangle &= |C_{j,m_j}^+|^2 \langle j, m_j + 1 | j, m_j + 1 \rangle = |C_{j,m_j}^+|^2 , \\ \implies \langle j, m_j | J^2 - J_z^2 - \hbar J_z | j, m_j \rangle &= |C_{j,m_j}^+|^2 , \\ \implies |C_{j,m_j}^+|^2 &= [j(j+1) - m_j(m_j + 1)] \hbar^2 , \\ \implies C_{j,m_j}^+ &= \sqrt{j(j+1) - m_j(m_j + 1)} \hbar . \end{aligned} \quad (15.39)$$

Similarly, we can start from

$$\begin{aligned} J_- |j, m_j\rangle &= C_{j,m_j}^- |j, m_j - 1\rangle , \\ \implies (J_- |j, m_j\rangle)^\dagger &= \langle j, m_j | (J_-)^\dagger = \langle j, m_j | J_+ = C_{j,m_j}^{-\ast} \langle j, m_j - 1 | . \end{aligned} \quad (15.40)$$

Therefore, by using relation $J^2 = J_+ J_- - \hbar J_z + J_z^2$ (eq.(15.23c)), we have

$$\begin{aligned} \langle j, m_j | J_+ J_- | j, m_j \rangle &= |C_{j,m_j}^-|^2 \langle j, m_j - 1 | j, m_j + 1 \rangle = |C_{j,m_j}^-|^2 , \\ \implies \langle j, m_j | J^2 - J_z^2 + \hbar J_z | j, m_j \rangle &= |C_{j,m_j}^-|^2 , \\ \implies |C_{j,m_j}^-|^2 &= [j(j+1) - m_j(m_j - 1)] \hbar^2 , \\ \implies C_{j,m_j}^- &= \sqrt{j(j+1) - m_j(m_j - 1)} \hbar . \end{aligned} \quad (15.41)$$

These relations will prove to be handy in our discussions on the addition of angular momenta.

Chapter 16

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

$$\nabla = \hat{e}_r \frac{\partial}{\partial r} + \hat{e}_\theta \frac{1}{r} \frac{\partial}{\partial \theta} + \hat{e}_\phi \frac{1}{r \sin \theta} \frac{\partial}{\partial \phi} \quad (16.1)$$

! Revise the transformation of the gradient from Cartesian to spherical coordinates.

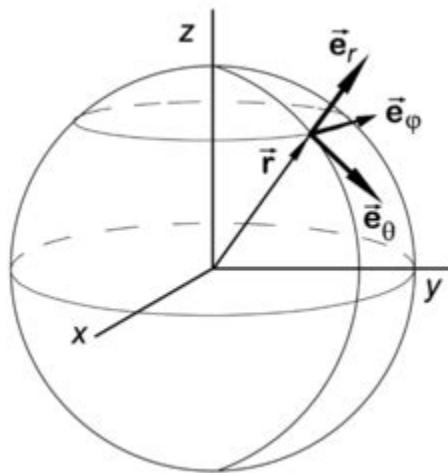


Figure 16.1: Spherical Coordinates

Then one can compute the components of the orbital angular momentum L_x , L_y , L_z

$$L_x = i\hbar \left(\sin \phi \frac{\partial}{\partial \theta} + \cot \theta \cos \phi \frac{\partial}{\partial \phi} \right) , \quad (16.2)$$

$$L_y = i\hbar \left(-\cos \phi \frac{\partial}{\partial \theta} + \cot \theta \sin \phi \frac{\partial}{\partial \phi} \right) , \quad (16.3)$$

$$L_z = -i\hbar \frac{\partial}{\partial \phi} . \quad (16.4)$$

! 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

$$L^2 = -\hbar^2 \left[\frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \phi^2} \right] , \quad (16.5)$$

$$L_{\pm} = \hbar e^{\pm i\phi} \left(\pm \frac{\partial}{\partial \theta} + i \cot \theta \frac{\partial}{\partial \phi} \right) . \quad (16.6)$$

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

$$L^2 \psi_{lm}(\theta, \phi) = \hbar^2 l(l+1) \psi_{lm}(\theta, \phi) , \quad (16.7a)$$

$$L_z \psi_{lm}(\theta, \phi) = m\hbar \psi_{lm}(\theta, \phi) . \quad (16.7b)$$

This can be written as

$$\left[\frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \phi^2} + \frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) \right] \psi_{lm} = -l(l+1) \psi_{lm} , \quad (16.8)$$

$$\frac{\partial}{\partial \phi} \psi_{lm} = im \psi_{lm} . \quad (16.9)$$

The above prompts us to try the following trial solution using separation of variables: $\psi_{lm}(\theta, \phi) = \Theta(\theta) \Phi(\phi)$.

From eq.(16.9), it is easily seen that

$$\Phi(\phi) = e^{im\phi} \quad (16.10)$$

! Verify this by plugging it into eq.(16.9). 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

$$\Phi(\phi + 2\pi) = \Phi(\phi) \quad (\text{Periodic Boundary Condition}) \quad (16.11)$$

$$\text{i.e., } e^{im2\pi} = 1 \implies m \in \mathbb{Z} \quad (16.12)$$

! 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.15.2 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.(15.27), we know that $l = 0, 1, 2, \dots$. Then, putting eq.(16.10) in qq.(16.8) gives the “Legendre” differential equation

$$\left[\frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) - \frac{m^2}{\sin^2 \theta} + l(l+1) \right] \Theta(\theta) = 0 . \quad (16.13)$$

The solutions to the above Legendre differential equation are the well-known **Legendre polynomials**. In this way, we find that the $\psi_{lm}(\theta, \phi)$ are the spherical harmonics $Y_{lm}(\theta, \phi)$:

$$\psi_{lm}(\theta, \phi) = Y_{lm}(\theta, \phi) = (-1)^{(m+|m|)/2} \left[\frac{2l+1}{4\pi} \frac{(l-|m|)!}{(l+|m|)!} \right]^{1/2} \times e^{im\phi} \times P_{l|m|}(\cos \theta) , \quad (16.14)$$

where the $P_{lm}(\xi)$ ($m > 0$) are the Legendre functions (given below in eq.(16.15)) in terms of Legendre polynomials (given below in eq.(16.16)).

$$\text{Legendre Functions: } P_{lm}(\xi) = (1 - \xi^2)^{m/2} \frac{d^m}{d\xi^m} P_l(\xi) , \quad (\xi = \cos \theta) \quad (16.15)$$

$$\text{Legendre Polynomials: } P_l(\xi) = \frac{1}{2^l l!} \frac{d^l}{d\xi^l} (\xi^2 - 1)^l . \quad (16.16)$$

The $P_l(\xi)$ are l^{th} order polynomials in ξ ; the $P_{lm}(\xi)$ are thus $(l-m)^{th}$ order polynomials in ξ , and are multiplied by ξ^m . Further, the $P_{lm}(\xi)$ have $(l-m)$ nodes in $-1 < \xi < 1$. Some of the Legendre polynomials are given by

$$\begin{aligned} P_0 &= 1 , \\ P_1 &= \xi , \\ P_2 &= \frac{1}{2} (3\xi^2 - 1) , \\ P_3 &= \frac{1}{2} (5\xi^3 - 3\xi) , \text{ and so on.} \end{aligned}$$

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

$$(l+1) P_{l+1}(\xi) = (2l+1) \xi P_l(\xi) - l P_{l-1}(\xi) , \quad (16.17)$$

$$(1-\xi^2) \frac{dP_l}{d\xi} = l (P_{l-1} - \xi P_l) . \quad (16.18)$$

The associated Legendre functions satisfy

$$\text{Legendre differential equation : } \left[(1-\xi^2) \frac{d^2}{d\xi^2} - 2\xi \frac{d}{d\xi} + l(l+1) - \frac{m^2}{(1-\xi^2)} \right] P_{lm}(\xi) = 0 \quad (16.19)$$

$$\text{Parity transformation : } P_{lm}(-\xi) = (-1)^{l+m} P_{lm}(\xi) \quad (16.20)$$

$$\text{Orthogonality relation : } \int_{-1}^1 P_{lm}(\xi) P_{l'm'}(\xi) d\xi = \frac{2}{(2l+1)(l+m)!} \frac{(l-m)!}{(l+m)!} \delta_{ll'} \quad (16.21)$$

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 $Y_{lm}(\theta, \phi)$

$$\text{Orthogonality relation : } \int_0^{2\pi} d\phi \int_0^\pi d\theta \sin \theta Y_{lm}^*(\theta, \phi) Y_{l'm'}(\theta, \phi) = \delta_{ll'} \delta_{mm'} , \quad (16.22)$$

$$\text{Completeness relation : } \sum_{l=0}^{\infty} \sum_{m=-l}^l Y_{lm}(\theta, \phi) Y_{lm}^*(\theta', \phi') = (\sin \theta)^{-1} \delta(\theta - \theta') \delta(\phi - \phi') . \quad (16.23)$$

Eqs.(16.22) and (16.23) prove the orthogonality and completeness of the spherical harmonics $Y_{lm}(\theta, \phi)$, thus making them suitable candidate for an eigenbasis. Eq.(16.24) below is called the **addition theorem**, which is a generalization of $\cos(\theta - \theta') = \cos \theta' \cos \theta + \sin \theta' \sin \theta$:

$$\sum_{m=-l}^l Y_{lm}(\theta, \phi) Y_{lm}^*(\theta', \phi') = \frac{(2l+1)}{4\pi} P_l(\cos \theta) , \quad (16.24)$$

$$Y_{l,-m}(\theta, \phi) = (-1)^m Y_{lm}^*(\theta, \phi) . \quad (16.25)$$

For the unit vectors \hat{a}, \hat{b} , it can be shown that

$$P_l(\hat{a} \cdot \hat{b}) = \frac{4\pi}{2l+1} \sum_{m=-l}^l Y_{lm}^*(\theta', \phi') Y_{lm}(\theta, \phi) ,$$

such that for $\hat{x} = \hat{y} \implies \sum_{m=-l}^l Y_{lm}^*(\theta', \phi') Y_{lm}(\theta, \phi) = \frac{2l+1}{4\pi} .$ (16.26)

Eq.(16.26) is a generalization of $\cos^2 \theta + \sin^2 \theta = 1$ to 2D. Some of the spherical harmonics are given by

$$\begin{aligned} Y_{00} &= \frac{1}{\sqrt{4\pi}} , \\ Y_{10} &= \sqrt{\frac{3}{4\pi}} \cos \theta , \\ Y_{11} &= -\sqrt{\frac{3}{8\pi}} \sin \theta e^{i\phi} , \\ Y_{20} &= \sqrt{\frac{5}{16\pi}} (3 \cos^2 \theta - 1) , \\ Y_{21} &= -\sqrt{\frac{15}{8\pi}} \sin \theta \cos \theta e^{i\phi} , \\ Y_{22} &= \sqrt{\frac{15}{32\pi}} \sin^2 \theta e^{i2\phi} , \text{ and so on.} \end{aligned}$$

Under the action of parity operator $\mathcal{P}\psi(\mathbf{r}) = \psi(-\mathbf{r})$,

$$\begin{aligned} \mathcal{P}Y_{lm}(\theta, \phi) &= Y_{lm}(\pi - \theta, \pi + \phi) \\ &= e^{im\pi} (-1)^{l+|m|} Y_{lm}(\theta, \phi) , \end{aligned} \quad (16.27)$$

$$\therefore \mathcal{P}Y_{lm}(\theta, \phi) = (-1)^l Y_{lm}(\theta, \phi) . \quad (16.28)$$

This means that Y_{lm} is even for even l and odd for odd l . Figure 16.2 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.16.2, we see plots of the first few Spherical Harmonics. You may well recall having seen these somewhere as the orbitals for the Hydrogen atom.

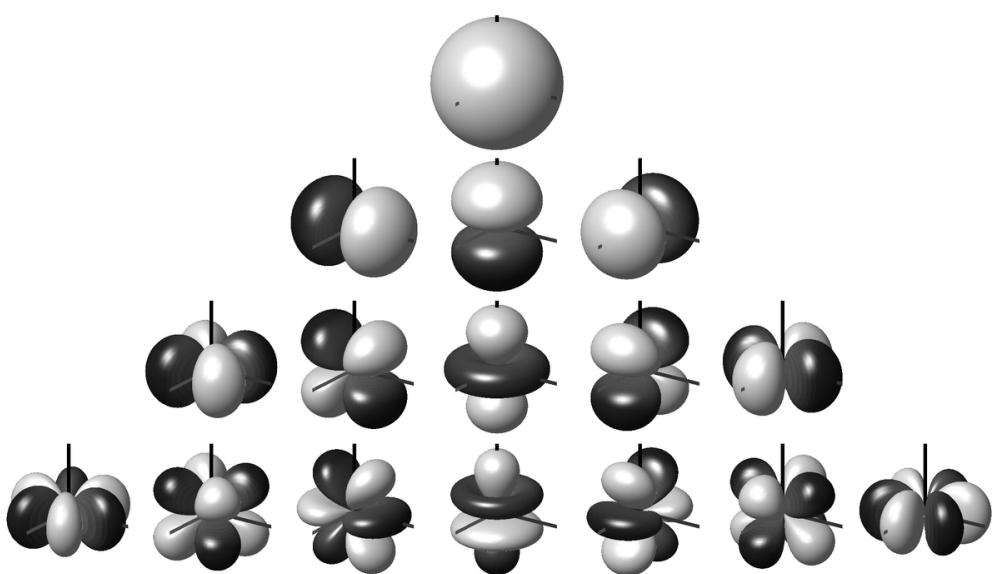


Figure 16.2: Spherical Harmonics

Chapter 17

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

$$H = \frac{1}{2m}\vec{p}^2 + V(r) \quad (17.1)$$

Now as $\mathbf{L} = \mathbf{r} \times \mathbf{p}$ such that $L_k = \epsilon_{ijk}x_ip_k$, the norm of the orbital angular momentum L^2 will be ($\mathbf{r} = x_1\hat{x} + x_2\hat{y} + x_3\hat{z}$)

$$\begin{aligned} L^2 &= \sum_{k=x,y,z} L_k^2 = (\epsilon_{ijk}x_ip_j)(\epsilon_{mnk}x_mp_n) \\ &= \epsilon_{ijk}\epsilon_{mnk}x_ip_jx_mp_n \\ &= (\delta_{im}\delta_{jn} - \delta_{in}\delta_{jm})x_ip_jx_mp_n \\ &= x_ip_jx_ip_j - x_ip_jx_jp_i \\ &= x_i(x_ip_j - i\hbar\delta_{ij})p_j - x_ip_j(p_ix_j + i\hbar\delta_{ij}) \\ &= x_i^2p_j^2 - i\hbar x_ip_i - x_ip_ip_jx_j - i\hbar x_ip_i \\ &= x_i^2p_j^2 - 2i\hbar x_ip_i - x_ip_i(x_jp_j - i\hbar) \\ &= x_i^2p_j^2 - x_ip_ix_jp_j - i\hbar x_ip_i \\ &= r^2p^2 - (\mathbf{r} \cdot \mathbf{p})^2 - i\hbar\mathbf{r} \cdot \mathbf{p} \end{aligned} \quad (17.2)$$

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 $\epsilon_{jik} = -\epsilon_{ijk} = \dots$ (i.e., the value is -1 for any two indices being swapped).

Further, the product $\epsilon_{ijk}\epsilon_{mnk} = (\delta_{im}\delta_{jn} - \delta_{in}\delta_{jm})$. Note that following the Einstein convention, we are summing over all repeated indices.

Finally, we have used the identity $(x_ip_j - p_jx_i) = i\hbar\delta_{ij}$. The last term on the RHS of Eq.

17.2 arises purely due to the non-commutativity of \mathbf{r} , \mathbf{p} . For $\mathbf{r} = r\hat{e}_r$ as $\hat{p} = i\hbar\nabla$

$$\begin{aligned}\nabla &= \hat{e}_r \frac{\partial}{\partial r} + \hat{e}_\theta \frac{\partial}{\partial \theta} + \hat{e}_\phi \frac{1}{r \sin \theta} \frac{\partial}{\partial \phi} \\ \implies \mathbf{r} \cdot \mathbf{p} &= -i\hbar \mathbf{r} \cdot \nabla = -i\hbar r \frac{\partial}{\partial r}\end{aligned}\quad (17.3)$$

From Eq. 17.2,

$$\begin{aligned}\hat{p}^2 &= \frac{L^2}{r^2} - \frac{\hbar^2}{r^2} \left(\left(r \frac{\partial}{\partial r} \right)^2 + r \frac{\partial}{\partial r} \right) \\ &= \frac{L^2}{r^2} - \frac{\hbar^2}{r} \frac{\partial}{\partial r} - \frac{\hbar^2}{r} \frac{\partial}{\partial r} \left(r \frac{\partial}{\partial r} \right) \\ &= \frac{L^2}{r^2} - \hbar^2 \left(\frac{2}{r} \frac{\partial}{\partial r} + \frac{\partial^2}{\partial r^2} \right)\end{aligned}\quad (17.4)$$

$$\hat{p}^2 = \frac{L^2}{r^2} + \hat{p}_r^2 \quad (17.5)$$

where \hat{p}_r is the radial momentum given by

$$\hat{p}_r = -i \frac{\hbar}{r} \frac{\partial}{\partial r} \cdot r = -i\hbar \left(\frac{1}{r} + \frac{\partial}{\partial r} \right) \quad (17.6)$$

$$\begin{aligned}[r, \hat{p}_r] &= \left[r, -i\hbar \left(\frac{\partial}{\partial r} + \frac{1}{r} \right) \right] \\ &= i\hbar r \frac{\partial}{\partial r} - i\hbar + i\hbar \left(\frac{\partial}{\partial r} + \frac{1}{r} \right) r \\ &= i\hbar\end{aligned}\quad (17.7)$$

$$(17.8)$$

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.

Verify that the operator \hat{p}_r is Hermitian.

Putting Eq. 17.4 in Eq. 17.1,

$$\left[-\frac{\hbar^2}{2m} \left(\frac{\partial^2}{\partial r^2} + \frac{2}{r} \frac{\partial}{\partial r} \right) + \frac{1}{2mr^2} L^2 + V(r) \right] \psi(r, \theta, \phi) = E\psi(r, \theta, \phi) \quad (17.9)$$

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

$$\psi(r, \theta, \phi) = R(r) Y_{lm}(\theta, \phi) \quad \text{where } L^2 Y_{lm}(\theta, \phi) = l(l+1)\hbar^2 Y_{lm}(\theta, \phi), \quad (17.10)$$

$$\Rightarrow \left[-\frac{\hbar^2}{2m} \left(\frac{\partial^2}{\partial r^2} + \frac{2}{r} \frac{\partial}{\partial r} \right) + \frac{\hbar^2}{2mr^2} l(l+1) + V(r) \right] R(r) = ER(r) \quad (17.11)$$

Taking $R(r) = u(r)/r$,

$$\left[-\frac{\hbar^2}{2m} \frac{\partial^2}{\partial r^2} + \frac{\hbar^2}{2mr^2} l(l+1) + V(r) \right] u(r) = Eu(r) \quad (17.12)$$

$$\implies V_{eff}(r) = V(r) + \frac{\hbar^2}{2mr^2} l(l+1) \quad (17.13)$$

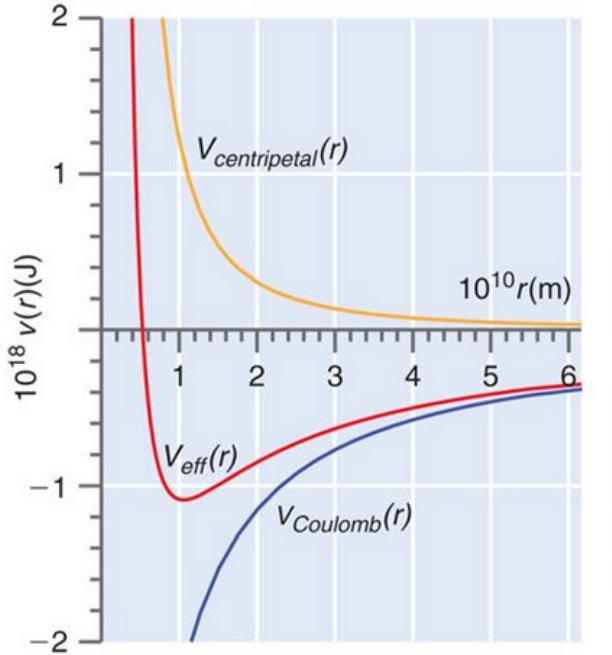


Figure 17.1: Effective potential for an atom with Coulomb and Centripetal potentials.

The plot in Fig.17.1 shows the effective potential V_{eff} for the Coulomb potential $V_{Coulomb}(r) \propto -1/r$ and the centripetal potential $V_{Centrifugal} = l(l+1)/2mr^2$.

The normalizability condition of ψ tells us that

$$\begin{aligned} \int d^3x |\psi|^2 &= \text{constant} \times \int_0^\infty dr r^2 \frac{|u(r)|^2}{r^2} < \infty \\ \implies \int_0^\infty dr |u(r)|^2 &< \infty \\ \implies \lim_{r \rightarrow \infty} |u(r)| &\leq \frac{a}{r^{1/2+\epsilon}} \quad \text{where } a = \text{constant}, \epsilon > 0 \end{aligned}$$

where the finite constant arise from the angular integrals $\int d\theta \int d\phi \sin^2\theta |Y_{lm}(\theta, \phi)|^2$. Thus, $u(r)$ must fall away faster than $1/\sqrt{r}$ as $r \rightarrow \infty$. Also, for the case of $\epsilon = 0$, you can easily see a logarithmic divergence of the integral with r in the limit of $r \rightarrow \infty$.

In the limit of $r \rightarrow 0$, we must have $u(r \rightarrow 0) \rightarrow 0$ faster than $r \rightarrow 0$ for the overall stability of the atomic system. Certainly, this is easily seen for $l \neq 0$ due to the overall repulsive nature of $V_{eff}(r \rightarrow 0)$. As

$$u(0) = 0 \text{ in the limit } r \rightarrow 0 , \quad (17.14)$$

we must have (for $l = 0$) a hard-wall boundary condition at $r = 0$

$$V_{eff}(r) = \begin{cases} V(r) & r > 0 , \\ \infty & r = 0 . \end{cases} \quad (17.15)$$

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 18

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

$$\mathbf{r} = \mathbf{r}_1 - \mathbf{r}_2 , \tag{18.1}$$

$$\begin{aligned}\mathbf{p} &= \mu(\mathbf{v}_1 - \mathbf{v}_2) , \quad \text{Reduced mass } \mu = \left[\frac{1}{m_1} + \frac{1}{m_2} \right]^{-1} = \frac{m_1 m_2}{m_1 + m_2} , \\ &= \frac{m_1 m_2}{m_1 + m_2} \left(\frac{\mathbf{p}_1}{m_1} - \frac{\mathbf{p}_2}{m_2} \right) , \\ &= \frac{m_2 \mathbf{p}_1 - m_1 \mathbf{p}_2}{m_1 + m_2} .\end{aligned} \tag{18.2}$$

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

$$[r_{\nu i}, p_{\mu j}] = i\hbar\delta_{ij}\delta_{\mu\nu} , \tag{18.3}$$

$$\implies [(\mathbf{r}_{CM})_i, (\mathbf{p}_{CM})_j] = i\hbar\delta_{ij} = [r_i, p_j] . \tag{18.4}$$

Thus, we can simplify the TISE for the 2-body system of the H-atom using

$$\begin{aligned} \mathbf{p}_{CM} &= -i\hbar\nabla_{CM}, \\ \mathbf{p}_r &= -i\hbar\nabla_r, \\ \frac{p_1^2}{2m_1} + \frac{p_2^2}{2m_2} &= \frac{p_{CM}^2}{M} + \frac{p^2}{2\mu}, \\ \implies H\Psi(\mathbf{r}_{CM}, \mathbf{r}) &= \left[\frac{p_{CM}^2}{M} + \frac{p^2}{2\mu} + V(r) \right] \Psi(\mathbf{r}_{CM}, \mathbf{r}) = E_{total}\Psi(\mathbf{r}_{CM}, \mathbf{r}). \end{aligned} \quad (18.5)$$

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 $\Psi(\mathbf{r}_{CM}, \mathbf{r}) = \phi(\mathbf{r}_{CM})\psi_r$ such that

$$\begin{aligned} E_{CM} &= \frac{\hbar^2 k_{CM}^2}{2M}, \quad \phi(\mathbf{r}_{CM}) = e^{i\mathbf{k}_{CM} \cdot \mathbf{r}_{CM}}, \\ \left[\frac{p^2}{2\mu} + V(r) \right] \psi(\mathbf{r}) &= E_{rel}\psi(\mathbf{r}) \quad \text{where } E_{rel} = E_{total} - \frac{\hbar^2 k_{CM}^2}{2M}, \end{aligned} \quad (18.6)$$

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

$$\left[-\frac{\hbar^2}{2\mu} \frac{d^2}{dr^2} + \frac{\hbar^2 l(l+1)}{2\mu r^2} - \frac{1}{4\pi\epsilon_0} \frac{e^2}{r} \right] u(r) = E_{rel}u(r). \quad (18.7)$$

Note that for $l = 0$,

$$V_{eff}(r) = \frac{\hbar^2 l(l+1)}{2\mu r^2} - \frac{1}{4\pi\epsilon_0} \frac{e^2}{r} \rightarrow -\frac{1}{4\pi\epsilon_0} \frac{e^2}{r} < 0.$$

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

$$\frac{d^2u(r)}{dr^2} + \frac{2\mu e^2}{4\pi\epsilon_0 \hbar^2} \frac{u(r)}{r} - \frac{l(l+1)}{r^2} u(r) = \frac{2\mu}{\hbar^2} \varepsilon u(r) \quad \text{where } \varepsilon > 0. \quad (18.8)$$

At $r \rightarrow \infty$ limit, we can drop the V_{eff} term and take the approximate form $u(r) \sim u_{app}(r)$

$$\begin{aligned} \frac{d^2u_{app}}{dr^2} &= \frac{2\mu}{\hbar^2} \varepsilon u_{app} \\ \implies u_{app}(r) &\propto e^{-\sqrt{\frac{2\mu\varepsilon}{\hbar^2}}r}, \end{aligned} \quad (18.9)$$

where we have thrown away the possible solution of $u_{app}(r) \propto e^{\sqrt{\frac{2\mu\varepsilon}{\hbar^2}}r}$ on grounds of normalisability (as it diverges in the limit of $r \rightarrow \infty$). With this in mind, we consider the trial

solution

$$u(r) = v(r) e^{-\sqrt{\frac{2\mu\varepsilon}{\hbar^2}}r}, \quad (18.10)$$

$$v(r) = \sum_{p=1}^{\infty} A_p r^p. \quad (18.11)$$

Here, A_p are constants, and $A_0 = 0$ (from the constraint that $u(r \rightarrow 0) \rightarrow 0$, eq.17.14). Substituting Eq. 18.10 in Eq. 18.8 we get

$$\frac{d^2v(r)}{dr^2} - \frac{2\sqrt{2\mu\varepsilon}}{\hbar} \frac{dv(r)}{dr} + \frac{2\mu e^2}{4\pi\epsilon_0\hbar^2} \frac{v(r)}{r} - \frac{l(l+1)}{r^2} v(r) = 0 \quad (18.12)$$

Substituting Eq. 18.11 in Eq. 18.12 in seeking the Frobenius power-series solution and equating the coefficient of r^p to be zero, we obtain the recursion relation

$$[p(p+1) - l(l+1)] A_{p+1} = \left[\frac{2p\sqrt{2\mu\varepsilon}}{\hbar} - \frac{2\mu e^2}{\hbar^2 4\pi\epsilon_0} \right] A_p. \quad (18.13)$$

For $p = l$, $A_p = 0 \implies A_{p-1} = A_{p-2} = A_{p-3} = \dots$. Thus, the only non-zero coefficients have $p > l$. Also, as $r \rightarrow \infty$, $u(r) \rightarrow 0$ (for a normalisable solution to the TISE). This means that the power series $v(r)$ must terminate at a finite $p(> l)$, so that the exponential part can dominate as $r \rightarrow \infty$.

Thus, for some $p = n > l$, $A_{p+1} = 0$

$$\implies \frac{2n\sqrt{2\mu\varepsilon}}{\hbar} = \frac{2\mu e^2}{\hbar^2 4\pi\epsilon_0} \quad (18.14)$$

$$\therefore E_{rel}^n = -\varepsilon = -\frac{\mu e^4}{(4\pi\epsilon_0)^2 2\hbar^2 n^2} \quad (18.15)$$

$$= -\frac{13.6}{n^2} \text{eV}, \quad n \geq l+1 \quad (18.15)$$

$$R(r) = \frac{v(r)}{r} e^{-\frac{\sqrt{2\mu\varepsilon}}{\hbar}r} \quad (18.16)$$

$$= R_{nl}(r)$$

Thus, after such a long calculation, we do get back the Bohr formula (eq.(2.39))! The various $R_{nl}(r)$ can be found from getting various A_p from Eq. 18.13, and in terms of the characteristic lengthscale for the atom (the Bohr radius a_0):

$$R_{10}(r) = \left(\frac{1}{a_0} \right)^{3/2} 2e^{-r/a_0},$$

$$R_{20}(r) = \left(\frac{1}{2a_0} \right)^{3/2} 2 \left(1 - \frac{r}{2a_0} \right) e^{-r/2a_0},$$

$$R_{21}(r) = \left(\frac{1}{2a_0} \right)^{3/2} \frac{1}{\sqrt{3}} \left(\frac{r}{a_0} \right) e^{-r/2a_0},$$



Bohr

and so on. Here, a_0 is the Bohr radius given by

$$a_0 = \frac{4\pi\epsilon_0\hbar^2}{\mu e^2} = 5.3 \times 10^{-11} \text{ m} , \quad E_{rel}^n = -\frac{\hbar^2}{2\mu n^2 a_0^2} ,$$

and the full wavefunction of the H-atom is given by

$$\psi_{nlm}(r, \theta, \phi) = R_{nl}(r) Y_{lm}(\theta, \phi) \quad (18.17)$$

where the good quantum numbers are given by

$$\begin{aligned} n &= 1, 2, 3, \dots && \text{(Principal quantum number)} , \\ l &= 0, 1, \dots, n-1 && \text{(Orbital quantum number)} , \text{ and} \\ m &= -l, -l+1, \dots, l-1, l && \text{(Magnetic quantum number)} . \end{aligned}$$

While $R_{nl}(r)$ is dependent on both n & l , the energy E_{rel}^n is dependent on n alone. There must, therefore, be degeneracies 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 $2l+1$ possible values of m . Thus, the total degeneracy for E_n will be

$$\begin{aligned} \text{Total degeneracy of } E_{rel}^n &= \sum_{l=0}^{n-1} (2l+1) , \\ &= 2 \frac{n(n-1)}{2} + n , \\ &= n^2 . \end{aligned} \quad (18.18)$$

Taking account of the spin internal degree of freedom (with its double degeneracy), the total degeneracy for E_{rel}^n 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 $\psi_{nlm}(r, \theta, \phi)$

$$P(r) = \int_0^r dr r^2 \int_0^\pi d\theta \sin \theta \int_0^{2\pi} d\phi |\psi_{nlm}(r, \theta, \phi)|^2 , \quad \lim_{r \rightarrow \infty} P(r) = 1 . \quad (18.19)$$

We can also define radial probability density $P_{nl}(r)$ such that it gives the probability of finding the electron in a small interval dr at a distance r

$$P_{nlm}(r) = \int_0^\pi \int_0^{2\pi} r^2 \sin \theta d\phi d\theta |\psi_{nlm}(r, \theta, \phi)|^2 , \quad (18.20)$$

$$\begin{aligned} &= r^2 |R_{nl}(r)|^2 \int_0^\pi \int_0^{2\pi} |Y_{lm}(\theta, \phi)|^2 d\phi d\theta , \\ &= r^2 |R_{nl}(r)|^2 \times I_{lm} \end{aligned} \quad (18.21)$$

where $I_{lm} = \int_0^{\pi} \int_0^{2\pi} |Y_{lm}(\theta, \phi)|^2 d\phi d\theta$.

$$\therefore P_{nlm}(r) = r^2 |R_{nl}(r)|^2 \times I_{lm}, \quad (18.22)$$

$$\text{and } P(r) = \int_0^r P_{nlm}(r) dr. \quad (18.23)$$

Some plots of the $P_{nlm}(r)$ (as a function of r/a_0) are shown in Fig.18.1 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.18.2 in order to visualise the “orbitals” of the Hydrogen atom.

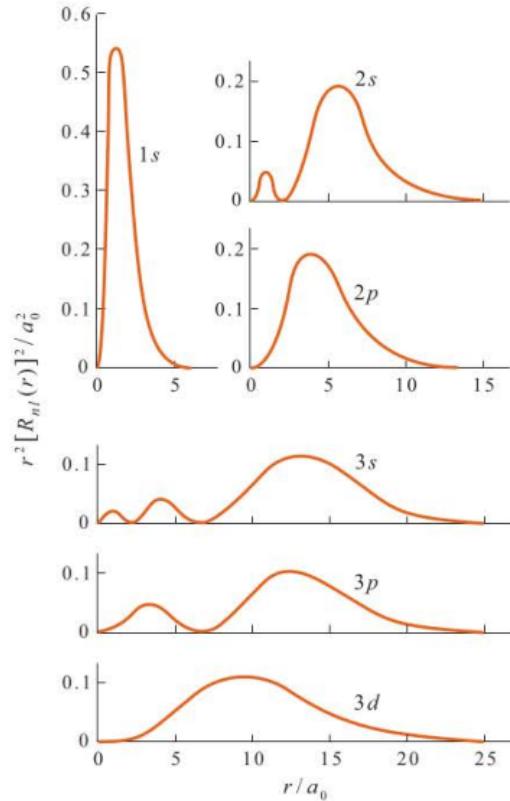


Figure 18.1: Radial Probability density function

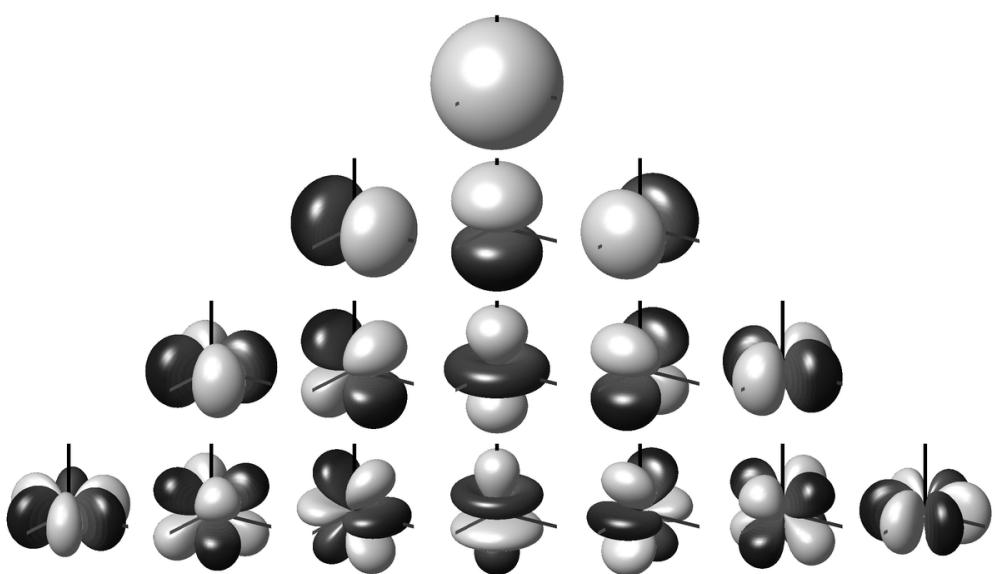


Figure 18.2: Spherical Harmonics

Chapter 19

Spin Angular Momentum

19.1 Evidence for spin

19.1.1 The Zeeman Effect

A Hydrogen atom placed in an external B -field aligned along the z-axis acquires an extra energy term in it's Hamiltonian

$$H = H_0 - \frac{e}{2m} \vec{B} \cdot \vec{L}_{eff} = H_0 - \frac{e}{2m} B L_{eff}^z \quad (19.1)$$

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

$$H\Psi_{nlm} = \left(-\frac{13.6eV}{n^2} - \hbar \frac{eB}{2m} m_{l_{eff}} \right) \Psi_{nlm} \implies E_{nlm} = -\frac{13.6eV}{n^2} - \hbar\omega_L m_{l_{eff}}, \quad (19.2)$$

where $\omega_L = \frac{eB}{2m}$ is the Larmor frequency and $m_{l_{eff}} \in \mathcal{Z}, -l_{eff} \leq m_{l_{eff}} \leq l_{eff}$. 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 $\vec{\mu}_L$

$$H = \vec{\mu}_L \cdot \vec{B} \quad (19.3)$$

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 Landé 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 $L_{eff} \in \frac{n}{2}$ 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).

19.1.2 The Stern-Gerlach Experiment

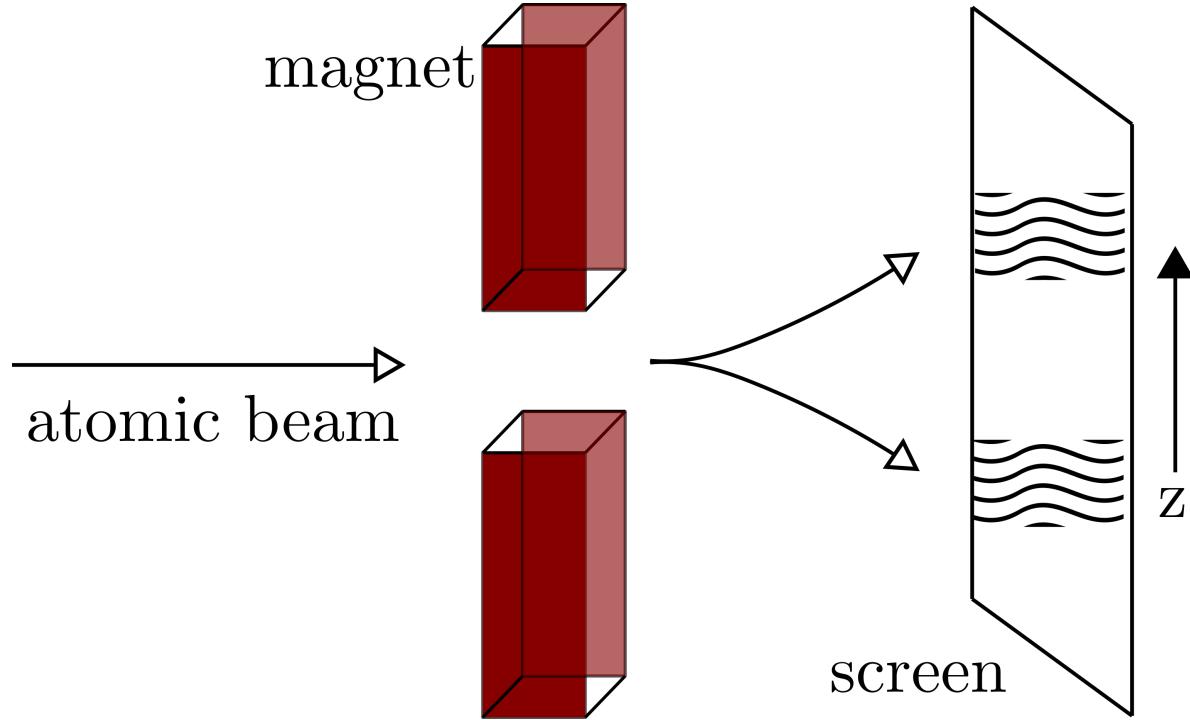


Figure 19.1: Splitting of the beam of particles into two patches when subjected to a non-uniform magnetic field

Here, an atomic beam traverses an inhomogeneous B -field such that there is a force on the atom (see fig. 19.1), given by

$$\vec{F} \propto \vec{\nabla} (\vec{\mu} \cdot \vec{B}(\vec{r})) \approx \mu_z \frac{\partial B_z(r)}{\partial z} \hat{e}_z \quad (19.4)$$

From the earlier discussion, 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

$$\vec{\mu}_S = -\frac{g_S \mu_B}{\hbar} \vec{S} \quad (19.5)$$

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

$$\vec{F} \propto \hat{z} \frac{\partial B_z}{\partial z} \left(-\frac{g_S \mu_B}{\hbar} S_z \right) \quad (19.6)$$

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. Ehrenfest 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.

19.2 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:

$$[S_i, S_j] = i\hbar\epsilon_{ijk}S_k \quad , \quad [S^2, S_i] = 0 \quad \text{for } i, j, k \in \{x, y, z\}. \quad (19.7)$$

Written out explicitly in terms of the components, these relations are

$$[S_x, S_y] = i\hbar S_z, \quad [S_y, S_z] = i\hbar S_x, \quad [S_z, S_x] = i\hbar S_y, \quad (19.8)$$

$$[S^2, S_x] = 0 = [S^2, S_y] = [S^2, S_z]. \quad (19.9)$$

From these commutation relations, we can write down the simultaneous eigenstates $|s, m\rangle$ of S^2 and S_z such that

$$\begin{aligned} S^2 |s, m_s\rangle &= s(s+1)\hbar^2 |s, m_s\rangle \\ S_z |s, m_s\rangle &= m_s \hbar |s, m_s\rangle . \end{aligned} \quad (19.10)$$

Along with the spin operators along the three directions, we also define the spin creation and annihilation operators,

$$S_{\pm} = S_x \pm iS_y, \quad S_+^\dagger = S_-, \quad S_- = S_+^\dagger. \quad (19.11)$$

Their action on the simultaneous eigenstates can be shown to give (as shown in eq.(15.37) earlier)

$$S_{\pm} |s, m_s\rangle = \hbar \sqrt{s(s+1) - m_s(m_s \pm 1)} |s, m_s \pm 1\rangle \quad (19.12)$$

19.3 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

$$\begin{aligned} S_z |\uparrow\rangle &= \frac{\hbar}{2} |\uparrow\rangle, & S^2 |\uparrow\rangle &= s(s+1)\hbar^2 |\uparrow\rangle = \frac{1}{2}(\frac{1}{2}+1)\hbar^2 = \frac{3\hbar^2}{4} |\uparrow\rangle \\ S_z |\downarrow\rangle &= -\frac{\hbar}{2} |\downarrow\rangle, & S^2 |\downarrow\rangle &= s(s+1)\hbar^2 |\downarrow\rangle = \frac{3\hbar^2}{4} |\downarrow\rangle \end{aligned} \quad (19.13)$$

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*)

$$|\uparrow\rangle \equiv \left| s = \frac{1}{2}, m_s = \frac{1}{2} \right\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad |\downarrow\rangle \equiv \left| s = \frac{1}{2}, m_s = -\frac{1}{2} \right\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad (19.14)$$

$$\langle \uparrow | \equiv \left\langle s = \frac{1}{2}, m_s = \frac{1}{2} \right| = (1 \ 0), \quad \langle \downarrow | \equiv \left\langle s = \frac{1}{2}, m_s = -\frac{1}{2} \right| = (0 \ 1). \quad (19.15)$$

The two vectors are of course orthogonal, as required, and moreover, orthonormal.

$$\begin{aligned} \langle \uparrow | \downarrow \rangle &= (1 \ 0) \begin{pmatrix} 0 \\ 1 \end{pmatrix} = 0 \\ \langle \downarrow | \downarrow \rangle &= (0 \ 1) \begin{pmatrix} 0 \\ 1 \end{pmatrix} = 1 \\ \langle \uparrow | \uparrow \rangle &= (1 \ 0) \begin{pmatrix} 1 \\ 0 \end{pmatrix} = 1 \end{aligned} \quad (19.16)$$

While we will focus on $s = \frac{1}{2}$, this representation can be easily extended to $s = 1$:

$$|s = 1, m_s = 1\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad |s = 1, m_s = 0\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad |s = 1, m_s = -1\rangle = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \quad (19.17)$$

Returning 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.

$$\begin{aligned} \langle \uparrow | S_z | \uparrow \rangle &= \frac{\hbar}{2} \\ \langle \uparrow | S_z | \downarrow \rangle &= 0 = \langle \downarrow | S_z | \uparrow \rangle \\ \langle \downarrow | S_z | \downarrow \rangle &= -\frac{\hbar}{2} \end{aligned} \quad (19.18)$$

We can thus write

$$S_z = \frac{\hbar}{2} \begin{pmatrix} \langle \uparrow | S_z | \uparrow \rangle & \langle \uparrow | S_z | \downarrow \rangle \\ \langle \downarrow | S_z | \uparrow \rangle & \langle \downarrow | S_z | \downarrow \rangle \end{pmatrix} = \frac{\hbar}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad (19.19)$$

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.

$$S_+ |\uparrow\rangle = 0 = S_- |\downarrow\rangle \quad (19.20)$$

To find the other actions, we recall that $S_{\pm} |s, m_s\rangle = \hbar \sqrt{s(s+1) - m_s(m_s \pm 1)} |s, m_s \pm 1\rangle$ (eq. 19.12). Then, it is easily seen that

$$\begin{aligned} S_+ |\downarrow\rangle &= \hbar \sqrt{\frac{3}{4} - (-\frac{1}{2})(-\frac{1}{2} + 1)} |\uparrow\rangle = \hbar \sqrt{\frac{3}{4} + \frac{1}{4}} |\uparrow\rangle = \hbar |\uparrow\rangle \\ S_- |\uparrow\rangle &= \hbar \sqrt{\frac{3}{4} - (\frac{1}{2})(\frac{1}{2} - 1)} |\downarrow\rangle = \hbar \sqrt{\frac{3}{4} + \frac{1}{4}} |\downarrow\rangle = \hbar |\downarrow\rangle \end{aligned} \quad (19.21)$$

These four equations are sufficient to write down all the matrix elements.

$$\begin{aligned} \langle \uparrow | S_+ | \uparrow \rangle &= \langle \downarrow | S_+ | \downarrow \rangle = \langle \uparrow | S_- | \uparrow \rangle = \langle \downarrow | S_- | \downarrow \rangle = 0 \\ \langle \downarrow | S_+ | \uparrow \rangle &= \langle \uparrow | S_- | \downarrow \rangle = 0 \\ \langle \uparrow | S_+ | \downarrow \rangle &= \langle \downarrow | S_- | \uparrow \rangle = \hbar \end{aligned} \quad (19.22)$$

We can now construct the matrices

$$\begin{aligned} S_+ &= \frac{\hbar}{2} \begin{pmatrix} \langle \uparrow | S_+ | \uparrow \rangle & \langle \uparrow | S_+ | \downarrow \rangle \\ \langle \downarrow | S_+ | \uparrow \rangle & \langle \downarrow | S_+ | \downarrow \rangle \end{pmatrix} = \hbar \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \\ S_- &= \frac{\hbar}{2} \begin{pmatrix} \langle \uparrow | S_- | \uparrow \rangle & \langle \uparrow | S_- | \downarrow \rangle \\ \langle \downarrow | S_- | \uparrow \rangle & \langle \downarrow | S_- | \downarrow \rangle \end{pmatrix} = \hbar \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}. \end{aligned} \quad (19.23)$$

Using the relations

$$\begin{aligned} S_x &= \frac{1}{2} (S_+ + S_-), \\ S_y &= \frac{1}{2i} (S_+ - S_-), \end{aligned} \quad (19.24)$$

we can construct the S_x and S_y matrices

$$\begin{aligned} S_x &= \frac{\hbar}{2} \begin{pmatrix} \langle \uparrow | S_x | \uparrow \rangle & \langle \uparrow | S_x | \downarrow \rangle \\ \langle \downarrow | S_x | \uparrow \rangle & \langle \downarrow | S_x | \downarrow \rangle \end{pmatrix} = \frac{\hbar}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \\ S_y &= \frac{\hbar}{2} \begin{pmatrix} \langle \uparrow | S_y | \uparrow \rangle & \langle \uparrow | S_y | \downarrow \rangle \\ \langle \downarrow | S_y | \uparrow \rangle & \langle \downarrow | S_y | \downarrow \rangle \end{pmatrix} = \frac{\hbar}{2} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}. \end{aligned} \quad (19.25)$$

The three spin operators can be written as $S_i = \frac{\hbar}{2} \sigma_i$, where σ_i are the Pauli spin matrices.

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad (19.26)$$

19.4 Properties of Pauli matrices

The Pauli matrices satisfy the following identities.

- $\sigma_x^2 = \mathbb{I} = \sigma_y^2 = \sigma_z^2$, $\mathbb{I} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$
- $[\sigma_x, \sigma_y] = \sigma_x \sigma_y - \sigma_y \sigma_x = 2i\sigma_z$, 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 $\sigma_x \sigma_y = -\sigma_y \sigma_x = i\sigma_z$
- $\sigma_x \sigma_y \sigma_z = i\mathbb{I}$
- $\text{Trace}(\sigma_x) = \text{Trace}(\sigma_y) = \text{Trace}(\sigma_z) = 0 \longrightarrow \text{traceless matrices}$
- $\text{Det}(\sigma_x) = \text{Det}(\sigma_y) = \text{Det}(\sigma_z) = -1$

The first three properties can be combined into

$$\sigma_i \sigma_j = \delta_{ij} + i\epsilon_{ijk}\sigma_k \quad (19.27)$$

Often we will talk about a vector of Pauli matrices. That is defined as

$$\vec{\sigma} = \sigma_x \hat{x} + \sigma_y \hat{y} + \sigma_z \hat{z} \quad (19.28)$$

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:

$$(\vec{\sigma} \cdot \vec{a}) (\vec{\sigma} \cdot \vec{b}) = (\vec{a} \cdot \vec{b}) \mathbb{I} + i\vec{\sigma} \cdot (\vec{a} \times \vec{b}) \quad (19.29)$$

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

$$A = a\sigma_x + b\sigma_y + c\sigma_z + d\mathbb{I}, \quad (19.30)$$

where $a, b, c, d \in \mathcal{C}$ in general (and $\in \mathcal{R}$ in particular).

19.5 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

$$(\psi_1 \quad \psi_2) \quad (19.31)$$

with eigenvalue c :

$$\underbrace{\frac{\hbar}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}}_{S_x} \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix} = c \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix} \quad (19.32)$$

Solving the eigenvalue problem leads to the determinant equation:

$$\begin{vmatrix} -c & \frac{\hbar}{2} \\ \frac{\hbar}{2} & c \end{vmatrix} = 0 \implies c^2 - \left(\frac{\hbar}{2}\right)^2 = 0 \implies c = \pm \frac{\hbar}{2} \quad (19.33)$$

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

$$\begin{aligned} \text{eigenvalue } \frac{\hbar}{2} : | \rightarrow \rangle &= \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \frac{1}{\sqrt{2}} (| \uparrow \rangle + | \downarrow \rangle) \\ \text{eigenvalue } -\frac{\hbar}{2} : | \leftarrow \rangle &= \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \frac{1}{\sqrt{2}} (| \uparrow \rangle - | \downarrow \rangle) \end{aligned} \quad (19.34)$$

while for S_y , we get

$$\begin{aligned} \text{eigenvalue } \frac{\hbar}{2} : | \nearrow \rangle &= \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \end{pmatrix} = \frac{1}{\sqrt{2}} (| \uparrow \rangle + i | \downarrow \rangle) \\ \text{eigenvalue } -\frac{\hbar}{2} : | \swarrow \rangle &= \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -i \end{pmatrix} = \frac{1}{\sqrt{2}} (| \uparrow \rangle - i | \downarrow \rangle) \end{aligned} \quad (19.35)$$

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?

19.6 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 $\vec{\mu} \times \vec{B}$.

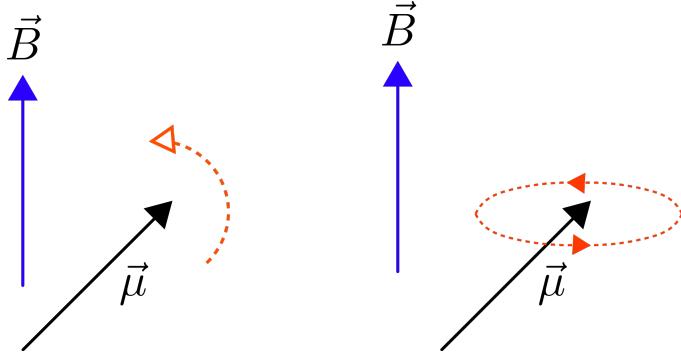


Figure 19.2: Left: Torque that aligns the spin along \vec{B} . Right: Precession of the spin about \vec{B}

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

$$H |\Psi\rangle = i\hbar \frac{\partial}{\partial t} |\Psi\rangle \quad (19.36)$$

What is the Hamiltonian H for this problem? If we assume the particle is at rest (i.e., no orbital motion), we get the Zeeman Hamiltonian

$$H = -\vec{\mu} \cdot \vec{B} , \quad (19.37)$$

where 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

$$H = \frac{2\mu_B}{\hbar} B \hat{z} \cdot \vec{S} = \frac{2\mu_B}{\hbar} B S_z = \mu_B B \sigma_z \quad (19.38)$$

where $\sigma_z = \frac{2}{\hbar} S_z$ is the Pauli matrix along z direction. The solution for the eigenstates and eigenvalues of H are simple: $|\uparrow\rangle$ with $E_{|\uparrow\rangle} = \mu_B B$ and $|\downarrow\rangle$ with $E_{|\downarrow\rangle} = -\mu_B B$. The difference between the energies of the two states, $\Delta E = 2\mu_B B$ in this case is called the Zeeman splitting. Now, given that the spin angular momentum operators belong to a two-dimensional Hilbert space, a general state vector $|\Psi\rangle$ will also have two components

$$|\Psi\rangle = \begin{pmatrix} \Psi_+ \\ \Psi_- \end{pmatrix} . \quad (19.39)$$

The TDSE becomes

$$\mu_B B \sigma_z |\Psi\rangle = i\hbar \frac{\partial}{\partial t} |\Psi\rangle \implies \mu_B B \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} \Psi_+ \\ \Psi_- \end{pmatrix} = i\hbar \frac{\partial}{\partial t} \begin{pmatrix} \Psi_+ \\ \Psi_- \end{pmatrix} \quad (19.40)$$

Comparing the matrix elements gives

$$\pm \mu_B B \Psi_{\pm} = i\hbar \frac{\partial}{\partial t} \Psi_{\pm} \quad (19.41)$$

The solution of these equations are

$$\Psi_{\pm} = A_{\pm} \exp \left\{ \mp i \left(\mu_B B \frac{t}{\hbar} \right) \right\} \quad (19.42)$$

where the coefficients A_{\pm} specify the initial condition: $|\Psi(t=0)\rangle = \begin{pmatrix} A_+ \\ A_- \end{pmatrix}$. We now consider two specific initial conditions.

19.6.1 Spin is along +z at t=0

The initial state with the spin aligned along the z -direction at time $t = 0$ is given by

$$|\Psi(t=0)\rangle = |\uparrow\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad (19.43)$$

hence $A_+ = 1, A_- = 0$. Plugging these into the solutions gives $\Psi_+(t) = \exp\{-i(\mu_B B \frac{t}{\hbar})\}$ and $\Psi_-(t) = 0$, we obtain

$$|\Psi(t)\rangle = \begin{pmatrix} \exp\{-i(\mu_B B \frac{t}{\hbar})\} \\ 0 \end{pmatrix} = \begin{pmatrix} e^{-i\omega \frac{t}{2}} \\ 0 \end{pmatrix} = e^{-i\omega \frac{t}{2}} |\uparrow\rangle, \quad (19.44)$$

where $\omega \equiv \frac{2\mu_B B}{\hbar}$ is the Larmor frequency. The probability of finding the spin along $+z$ is, as a function of time, given by

$$P(t) = |\langle \uparrow | \Psi(t) \rangle|^2 = \left| \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} \exp\{-i(\mu_B B \frac{t}{\hbar})\} \\ 0 \end{pmatrix} \right|^2 = \left| \exp\left\{-i\left(\mu_B B \frac{t}{\hbar}\right)\right\} \right|^2 = 1. \quad (19.45)$$

We find that the probability is always equal to unity, as the time evolution for this initial condition is trivial, i.e., there is no precession of the spin as it is always aligned with the field.

19.6.2 Spin is along +x at t=0

We now begin with the spin aligned along the x -direction at time $t = 0$

$$|\Psi(t=0)\rangle = |\rightarrow\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \implies A_{\pm} = \frac{1}{\sqrt{2}} \quad (19.46)$$

$$|\Psi(t)\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} \exp\{-i(\mu_B B \frac{t}{\hbar})\} \\ \exp\{i(\mu_B B \frac{t}{\hbar})\} \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} e^{-i\omega \frac{t}{2}} \\ e^{i\omega \frac{t}{2}} \end{pmatrix} \quad (19.47)$$

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),

$$|\Psi\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} e^{-i\pi} \\ e^{i\pi} \end{pmatrix} = -\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = e^{i\pi} |\rightarrow\rangle . \quad (19.48)$$

Thus, the state is equal to $\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = |\rightarrow\rangle$ 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. It is worth noting that had we take $t = 4\pi/\omega$, the dynamical phase factor would have been $e^{i2\pi} = 1$ (i.e., trivial). The question is, why did it need two full rotations (i.e., by 4π) for the state to return to itself with a trivial phase factor, while a single complete rotation (i.e., 2π) led to the state returning to itself but multiplied by a phase factor corresponding to -1 ? This is related to the so-called “double covering” structure of the $SU(2)$ symmetry group that determines the algebra of a spin-1/2 degree of freedom. We visualise the quantum dynamics of the spin-1/2 as the motion of a vector of fixed length ($3/4\hbar^2$) with its vertex fixed to the surface of the so-called “Bloch sphere”. The $SU(2)$ group has a peculiar structure that it involves a double covering of the Bloch sphere associated with the $SO(3)$ group, i.e., wrapping around the sphere once does not cover it completely, but twice does!

At precisely half this time period, i.e., $t = \frac{\pi}{\omega}$,

$$|\Psi\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} e^{-i\frac{\pi}{2}} \\ e^{i\frac{\pi}{2}} \end{pmatrix} = -\frac{i}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix} = e^{i\pi/2} |\leftarrow\rangle , \quad (19.49)$$

i.e., this is a spin pointing along $-x$ (upto the phase $e^{i\pi/2} = -i$).

At a quarter of this time period, i.e., $t = \frac{\pi}{2\omega}$,

$$|\Psi\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} e^{-i\frac{\pi}{4}} \\ e^{i\frac{\pi}{4}} \end{pmatrix} = \frac{1-i}{\sqrt{2}} \begin{pmatrix} 1 \\ i \end{pmatrix} = e^{i\pi/4} |\nearrow\rangle , \quad (19.50)$$

i.e., this is a spin pointing along $+y$ (upto a phase $e^{i\pi/4}$).

At three quarters of this time period, i.e., $t = \frac{3\pi}{2\omega}$,

$$|\Psi\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} e^{-i\frac{3\pi}{4}} \\ e^{i\frac{3\pi}{4}} \end{pmatrix} = \frac{1+i}{\sqrt{2}} \begin{pmatrix} -1 \\ i \end{pmatrix} = e^{i3\pi/4} |\swarrow\rangle , \quad (19.51)$$

i.e., 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 after having being excited with a pulse like perturbation. 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.

19.6.3 The Heisenberg equation of motion approach to spin precession

Recall that in the Heisenberg picture, the operator undergoes the time evolution while the state vector does not. This is the very opposite of the Schrödinger picture. Spin precession dynamics represents an excellent example in which to visualise both pictures of quantum dynamics. We have already seen the Schrödinger dynamics for spin precession just above, and we will now see it in the Heisenberg picture. Recall that the Hamiltonian for the problem at hand is given by

$$H = -\vec{\mu} \cdot \vec{B} = -\frac{g_s \mu_B}{\hbar} \vec{S} \cdot \vec{B}. \quad (19.52)$$

With this Hamiltonian, the Heisenberg equation of motion is given by

$$\frac{d\vec{S}}{dt} = \frac{i}{\hbar} [H, \vec{S}] , \quad (19.53)$$

$$= -\frac{ig_s \mu_B}{\hbar^2} [\vec{S} \cdot \vec{B}, \vec{S}] . \quad (19.54)$$

We now evaluate this component-wise for \vec{S} . Thus,

$$\hat{i} \frac{dS_x}{dt} = -\frac{ig_s \mu_B}{\hbar^2} [S_x B_x, S_y \hat{j} + S_z \hat{k}] , \quad (19.55)$$

$$= -\frac{ig_s \mu_B}{\hbar^2} [S_x, S_y] B_x \hat{j} - \frac{ig_s \mu_B}{\hbar^2} [S_x, S_z] B_x \hat{k} ,$$

$$= -\frac{ig_s \mu_B}{\hbar^2} (i\hbar S_z) B_x \hat{j} - \frac{ig_s \mu_B}{\hbar^2} (-i\hbar S_y) B_x \hat{k} ,$$

$$= \frac{g_s \mu_B}{\hbar} (S_y B_x \hat{j} - S_y B_x \hat{k}) . \quad (19.56)$$

Similarly, we obtain

$$\hat{j} \frac{dS_y}{dt} = \frac{g_s \mu_B}{\hbar} (-S_z B_y \hat{i} + S_x B_y \hat{k}) , \quad (19.57)$$

$$\hat{k} \frac{dS_z}{dt} = \frac{g_s \mu_B}{\hbar} (S_y B_z \hat{i} - S_x B_z \hat{j}) . \quad (19.58)$$

Bringing everything together, we obtain

$$\frac{d\vec{S}}{dt} = \frac{g_s\mu_B}{\hbar} \left((S_y B_z - S_z B_y) \hat{i} + (S_z B_x - S_x B_z) \hat{j} + (S_x B_y - S_y B_x) \hat{k} \right), \quad (19.59)$$

$$\begin{aligned} &= \frac{g_s\mu_B}{\hbar} (\vec{S} \times \vec{B}), \\ &= \frac{2e\hbar}{2m\hbar} (\vec{S} \times \vec{B}) \quad (\text{using } g_s = 2, \mu_B = \frac{e\hbar}{2m}), \\ \Rightarrow \frac{d\vec{S}}{dt} &= \frac{e}{m} (\vec{S} \times \vec{B}). \end{aligned} \quad (19.60)$$

The final expression for the equation of motion of the spin angular momentum operator \vec{S} (known as the Landau-Lifshitz equation for spin torque) clearly displays it's precession with time about an axis given by \vec{B} . Importantly, the Larmor frequency $\omega_B = eB/m$ clearly appears in the equation of motion.

19.7 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

$$[\vec{S}, \vec{r}] = 0, \quad [\vec{S}, \vec{p}] = 0, \quad [\vec{S}, \vec{L}] = 0 \quad (19.61)$$

we can define eigenstates simultaneously for \vec{S} and any one of the other three. For example, we can construct a basis from $|\vec{r}\rangle \otimes |\uparrow\rangle$ and $|\vec{r}\rangle \otimes |\downarrow\rangle$ (where the \otimes indicates a direct product of vector spaces), such that

$$|\Psi\rangle = \int d^3\vec{r} [\psi_+(\vec{r}) |\vec{r}\rangle |\uparrow\rangle + \psi_-(\vec{r}) |\vec{r}\rangle |\downarrow\rangle] \quad (19.62)$$

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

$$i\hbar \frac{\partial}{\partial t} \begin{pmatrix} \psi_+ \\ \psi_- \end{pmatrix} = \left[-\frac{\hbar^2}{2m} \vec{\nabla}^2 + V(\vec{r}) + \mu_B (\vec{L} + \vec{S}) \cdot \vec{B} \right] \begin{pmatrix} \psi_+(\vec{r}, t) \\ \psi_-(\vec{r}, t) \end{pmatrix}, \quad (19.63)$$

where $\vec{L} + \vec{S} = \vec{J}$ gives the total angular momentum.

Chapter 20

Addition of two angular momenta

20.1 The general problem

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 $[\vec{J}_1, \vec{J}_2] = 0$. We can always consider the eigenstates for such systems to be $|j_1, m_1; j_2, m_2\rangle = |j_1, m_1\rangle |j_2, m_2\rangle$, as these are simultaneous eigenstates of J_1^2 , J_{1z} , J_2^2 and J_{2z} , i.e., direct product combinations of the separate Hilbert spaces for \vec{J}_1 and \vec{J}_2 .

There are, however, many quantum systems in which operators such as \vec{J}_1 and \vec{J}_2 do interact with one another (or are coupled to one another). In such systems, it is often the case that the total angular momentum operator $\vec{J} = \vec{J}_1 + \vec{J}_2$ is an important observable, e.g., systems whose Hamiltonians have couplings between J_1 and J_2 . Clearly, using

$$[J_{\nu,i}, J_{\mu,j}] = i\hbar\delta_{\nu\mu}\epsilon_{ijk}J_{\mu k}, \quad \nu, \mu \in \{1, 2\} \quad (20.1)$$

such that

$$[J_i, J_j] = i\hbar\epsilon_{ijk}J_k, \quad (20.2)$$

all properties of angular momentum operators and their eigenstates hold also for \vec{J} . The question we would like to ask is: what do the eigenstates of \vec{J} 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

$$J_z |j_1, m_1; j_2, m_2\rangle = (J_{1z} + J_{2z}) |j_1, m_1; j_2, m_2\rangle = \hbar(m_1 + m_2) |j_1, m_1; j_2, m_2\rangle, \quad (20.3)$$

these states are not eigenstates of J^2 as

$$[J^2, J_{iz}] \neq 0, \quad i = 1, 2. \quad (20.4)$$

Since

$$[(J_{1x} + J_{2x})^2, J_{\alpha z}] \neq 0, \quad [(J_{1y} + J_{2y})^2, J_{\alpha z}] \neq 0 \quad \text{for } \alpha = 1, 2, \quad (20.5)$$

and we also have the relations

$$[J^2, J_\alpha] = 0 \quad (\alpha = x, y, z) , \quad [J^2, J_i^2] = 0 , \quad i = 1, 2 , \quad (20.6)$$

we need eigenstates of J^2 , J_z , J_1^2 and J_2^2 . That is, we need $|j, m_j, j_1, j_2\rangle$. Importantly, J_{1z} and J_{2z} are no longer good quantum numbers; rather, their sum $J_z = J_{1z} + J_{2z}$ is. In order to see how this works out in a simple example, we first investigate the addition of two spin-half angular momenta operators.

20.1.1 Addition of two spin-1/2 operators

Given two spin-half operators S_1 and S_2 , we can create a total spin operator

$$\vec{S} = \vec{S}_1 + \vec{S}_2 , \quad (20.7)$$

and we are looking for eigenstates of S^2 and $S_z = S_{1z} + S_{2z}$, $|S, S_z\rangle$ such that

$$S^2 |S, S_z\rangle = S(S+1)\hbar^2 |S, S_z\rangle , \quad S = 0, 1 \quad \text{and} \quad (20.8)$$

$$S_z |S, S_z\rangle = S_z \hbar |S, S_z\rangle , \quad S_z \in [-S, S] , \quad (20.9)$$

i.e., we have two sectors of S given by $S = 1/2 + 1/2 = 1$ and $S = 1/2 - 1/2 = 0$, and the two sectors have a range of S_z given by $S_z = [-1, 0, 1]$ and $S_z = 0$ respectively.

To proceed, we note that we can write down four direct product states in the basis of the spins S_1 and S_2 : $|\uparrow, \uparrow\rangle, |\uparrow, \downarrow\rangle, |\downarrow, \uparrow\rangle, |\downarrow, \downarrow\rangle$ (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). Further, the action of $S_z = S_{1z} + S_{2z}$ on these states is straightforward to compute

$$S_z |\uparrow\uparrow\rangle = \hbar |\uparrow\uparrow\rangle , S_z |\downarrow\downarrow\rangle = -\hbar |\downarrow\downarrow\rangle , S_z |\uparrow\downarrow\rangle = 0 = S_z |\downarrow\uparrow\rangle . \quad (20.10)$$

Note that

$$\begin{aligned} S^2 &= S_1^2 + S_2^2 + \vec{S}_1 \cdot \vec{S}_2 + \vec{S}_2 \cdot \vec{S}_1 \\ &= S_1^2 + S_2^2 + 2\vec{S}_1 \cdot \vec{S}_2 \quad \left[\because [\vec{S}_1, \vec{S}_2] = 0 \right] \\ &= \frac{3}{2}\hbar^2 + 2S_{1z}S_{2z} + S_{1+}S_{2-} + S_{1-}S_{2+} , \end{aligned} \quad (20.11)$$

where we have used

$$S_i^2 = \frac{3\hbar^2}{4} , \quad (20.12)$$

$$\vec{S}_1 \cdot \vec{S}_2 = S_{1z}S_{2z} + S_{1x}S_{2x} + S_{1y}S_{2y} , \quad \text{and} \quad (20.13)$$

$$= S_{1z}S_{2z} + \frac{1}{2}(S_{1+}S_{2-} + S_{1-}S_{2+}) , \quad (20.14)$$

and $S_{\nu\pm} = S_{\nu x} \pm iS_{\nu y}$ ($\nu \in 1, 2$). Then, we can write

$$S^2 |\uparrow\uparrow\rangle = \left[\frac{3}{2}\hbar^2 + 2\left(\frac{\hbar}{2}\right)^2 \right] |\uparrow\uparrow\rangle = 2\hbar^2 |\uparrow\uparrow\rangle , \quad S_z |\uparrow\uparrow\rangle = \hbar |\uparrow\uparrow\rangle , \quad (20.15)$$

$$S^2 |\downarrow\downarrow\rangle = 2\hbar^2 |\downarrow\downarrow\rangle , \quad S_z |\downarrow\downarrow\rangle = -\hbar |\downarrow\downarrow\rangle .$$

As $S^2 |S, S_z\rangle = S(S+1)\hbar^2 |S, S_z\rangle$, we get from the above relations $S = 1, S_z = \pm 1$ for the states $|S, S_z\rangle = |\uparrow\uparrow\rangle$ and $|\downarrow\downarrow\rangle$ respectively. The third state is obtained by operating with the operator $S_- = S_{1-} + S_{2-}$ on the state $|\uparrow\uparrow\rangle$

$$\begin{aligned} \frac{S_-}{\sqrt{2}\hbar} |\uparrow\uparrow\rangle &= \frac{S_{1-} + S_{2-}}{\sqrt{2}\hbar} |\uparrow\uparrow\rangle \\ &= \frac{1}{\sqrt{2}} (|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle) \quad [\because S_- |\uparrow\rangle = \hbar |\downarrow\rangle, S_+ |\downarrow\rangle = \hbar |\uparrow\rangle] , \end{aligned} \quad (20.16)$$

and the $1/\sqrt{2}$ factor is a normalisation factor. Just in case you are failing to see how $S_- |\uparrow\rangle = \hbar |\downarrow\rangle, S_+ |\downarrow\rangle = \hbar |\uparrow\rangle$, recall that

$$S_\pm |s, m_s\rangle = \hbar \sqrt{s(s+1) - m_s(m_s \pm 1)} |s, m_s \pm 1\rangle \quad (20.17)$$

such that

$$\begin{aligned} S_- |\uparrow\rangle &= S_- \left| \frac{1}{2}, \frac{1}{2} \right\rangle = \hbar \sqrt{\frac{1}{2}(\frac{1}{2}+1) - \frac{1}{2}(\frac{1}{2}-1)} \left| \frac{1}{2}, -\frac{1}{2} \right\rangle , \\ &= \hbar \sqrt{\frac{3}{4} + \frac{1}{4}} \left| \frac{1}{2}, -\frac{1}{2} \right\rangle = \hbar \left| \frac{1}{2}, -\frac{1}{2} \right\rangle = \hbar |\downarrow\rangle , \end{aligned} \quad (20.18)$$

$$\begin{aligned} S_+ |\downarrow\rangle &= S_+ \left| \frac{1}{2}, -\frac{1}{2} \right\rangle = \hbar \sqrt{\frac{1}{2}(\frac{1}{2}+1) - (-\frac{1}{2})(-\frac{1}{2}+1)} \left| \frac{1}{2}, \frac{1}{2} \right\rangle , \\ &= \hbar \sqrt{\frac{3}{4} + \frac{1}{4}} \left| \frac{1}{2}, \frac{1}{2} \right\rangle = \hbar \left| \frac{1}{2}, \frac{1}{2} \right\rangle = \hbar |\uparrow\rangle . \end{aligned} \quad (20.19)$$

Further, we can easily see that

$$\begin{aligned} S^2 \frac{1}{\sqrt{2}} (|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle) &= (\frac{3}{2}\hbar^2 + 2S_{1z}S_{2z} + S_{1+}S_{2-} + S_{1-}S_{2+}) \frac{1}{\sqrt{2}} (|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle) , \\ &= \hbar^2 (\frac{3}{2} - \frac{2}{4} + 1) \frac{1}{\sqrt{2}} (|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle) = 2\hbar^2 \frac{1}{\sqrt{2}} (|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle) \end{aligned} \quad (20.20)$$

$$S_z \frac{1}{\sqrt{2}} (|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle) = 0 , \quad (20.21)$$

i.e., $S = 1, S_z = 0$ for the third state given by $\frac{1}{\sqrt{2}} (|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle)$.

There is another normalised state, $\frac{1}{\sqrt{2}} (|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle)$, which is orthogonal to all these three eigenstates (Check this!), such that

$$S_z \frac{1}{\sqrt{2}} (|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle) = (S_{1z} + S_{2z}) \frac{1}{\sqrt{2}} (|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle) = 0 \quad (20.22)$$

and

$$\begin{aligned} S^2 \frac{1}{\sqrt{2}} (|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle) &= (\frac{3}{2}\hbar^2 + 2S_{1z}S_{2z} + S_{1+}S_{2-} + S_{1-}S_{2+}) \frac{1}{\sqrt{2}} (|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle) , \\ &= \left(\frac{3}{2}\hbar^2 - 2\frac{\hbar^2}{4} - \hbar^2 \right) \frac{1}{\sqrt{2}} (|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle) = 0 . \end{aligned} \quad (20.23)$$

Therefore, this fourth eigenstate is characterized by $S = 0$ and $S_z = 0$. Another way to obtain this state is to consider the state

$$|S = 0, S_z = 0\rangle = c_1 |\uparrow\downarrow\rangle + c_2 |\downarrow\uparrow\rangle , \quad |c_1|^2 + |c_2|^2 = 1 . \quad (20.24)$$

Now, noting that this state must be orthogonal to the state $|S = 1, S_z = 0\rangle$, we obtain

$$\begin{aligned} \langle 1, 0 | 0, 0 \rangle &= \frac{1}{\sqrt{2}} (\langle \uparrow\downarrow | + \langle \downarrow\uparrow |) (c_1 |\uparrow\downarrow\rangle + c_2 |\downarrow\uparrow\rangle) = 0 , \\ \implies \frac{1}{\sqrt{2}} (c_1 + c_2) &= 0 , \\ \implies c_2 &= -c_1 = \frac{1}{\sqrt{2}} , \end{aligned} \quad (20.25)$$

where we have also used the normalisation condition $|c_1|^2 + |c_2|^2 = 1$. Thus, we see that

$$|S = 1, S_z = 0\rangle = \frac{1}{\sqrt{2}} (|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle) . \quad (20.26)$$

Together, we have the following eigenstates:

$$\text{triplet(3) states : } \begin{cases} |S = 1, S_z = 1\rangle & \rightarrow |\uparrow\uparrow\rangle \\ |S = 1, S_z = 0\rangle & \rightarrow \frac{1}{\sqrt{2}} (|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle) \\ |S = 1, S_z = -1\rangle & \rightarrow |\downarrow\downarrow\rangle \end{cases} \quad (20.27)$$

$$\text{singlet(1) state : } \begin{cases} |S = 0, S_z = 0\rangle & \rightarrow \frac{1}{\sqrt{2}} (|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle) \end{cases}$$

Note the projection operators P_1 and P_2 which project onto the triplet and singlet spaces respectively

$$\begin{aligned} P_1 &= \frac{3}{4} + \frac{1}{\hbar^2} \vec{S}_1 \cdot \vec{S}_2 = \frac{3}{4} + \frac{1}{2\hbar^2} (S^2 - S_1^2 - S_2^2) \\ P_2 &= \frac{1}{4} - \frac{1}{\hbar^2} \vec{S}_1 \cdot \vec{S}_2 = \frac{1}{4} - \frac{1}{2\hbar^2} (S^2 - S_1^2 - S_2^2) \\ P_1 + P_2 &= 1 , \end{aligned} \quad (20.28)$$

such that

$$\begin{aligned} P_1 |1, m\rangle &= \left(\frac{3}{4} + \frac{1}{\hbar^2} \vec{S}_1 \cdot \vec{S}_2 \right) |1, m\rangle \\ &= \left[\frac{3}{4} + \frac{1}{2\hbar^2} (S^2 - S_1^2 - S_2^2) \right] |1, m\rangle \\ &= \left[\frac{3}{4} + \frac{1}{2} \left(1 \times 2 - \frac{3}{4} - \frac{3}{4} \right) \right] |1, m\rangle \\ &= |1, m\rangle \end{aligned} \quad (20.29)$$

$$P_1 |0, 0\rangle = \left[\frac{3}{4} + \frac{1}{2} \left(0 - \frac{3}{4} - \frac{3}{4} \right) \right] |0, 0\rangle = 0 \quad (20.30)$$

$$P_2 |1, m\rangle = (1 - P_1) |1, m\rangle = 0 \quad (20.31)$$

$$P_2 |0, 0\rangle = (1 - P_1) |0, 0\rangle = |0, 0\rangle . \quad (20.32)$$

We will see in another chapter the elementary problem for which these are the eigenstates.

We end this section by pointing out the another significance of what we have learnt here. The $\{S, S_z\}$ basis of states can be written in terms of the $\{S_{1z}, S_{2z}\}$ basis of states (i.e., $|\uparrow, \uparrow\rangle$, $|\uparrow, \downarrow\rangle$, $|\downarrow, \uparrow\rangle$ and $|\downarrow, \downarrow\rangle$) : the three triplet states $|S = 1, S_z = 1\rangle = |\uparrow, \uparrow\rangle$, $|S = 1, S_z = -1\rangle = |\downarrow, \downarrow\rangle$ & $|S = 1, S_z = 0\rangle = \frac{1}{\sqrt{2}}(|\uparrow, \downarrow\rangle + |\downarrow, \uparrow\rangle)$, and the lone singlet state $|S = 0, S_z = 0\rangle = \frac{1}{\sqrt{2}}(|\uparrow, \downarrow\rangle - |\downarrow, \uparrow\rangle)$ can be written as

$$\begin{bmatrix} |1, 1\rangle \\ |1, -1\rangle \\ |1, 0\rangle \\ |0, 0\rangle \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 0 & 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} |\uparrow, \uparrow\rangle \\ |\downarrow, \downarrow\rangle \\ |\uparrow, \downarrow\rangle \\ |\downarrow, \uparrow\rangle \end{bmatrix} . \quad (20.33)$$

The 4×4 matrix on the right hand side of the above equation corresponds to the transformation matrix that connects the $\{S_{1z}, S_{2z}\}$ basis to the $\{S, S_z\}$ basis (i.e., a “rotation” from one basis to another). It is clear that this matrix is not diagonal, as the operators S_{1+} , S_{1-} , S_{2+} and S_{2-} connect between the states $|\uparrow, \downarrow\rangle$ and $|\downarrow, \uparrow\rangle$. The non-trivial elements of this transformation matrix are called the “Clebsch-Gordan” coefficients; in this problem, the Clebsch-Gordan coefficients are given by the 1 , $1/\sqrt{2}$ and $-1/\sqrt{2}$ factors within the above matrix. We will learn more about them below.

It is also worth pointing out that the triplet-zero state $|S = 1, S_z = 0\rangle = \frac{1}{\sqrt{2}}(|\uparrow, \downarrow\rangle + |\downarrow, \uparrow\rangle)$ and the singlet state $|S = 0, S_z = 0\rangle = \frac{1}{\sqrt{2}}(|\uparrow, \downarrow\rangle - |\downarrow, \uparrow\rangle)$ are special: unlike the $|S = 1, S_z = 1\rangle = |\uparrow, \uparrow\rangle$ and $|S = 1, S_z = -1\rangle = |\downarrow, \downarrow\rangle$ states, $|S = 1, S_z = 0\rangle$ and $|S = 0, S_z = 0\rangle$ cannot be represented as direct product states in terms of Hilbert spaces for S_1 and S_2 . Instead, the latter two are the simplest examples of what are known as tensor product states, i.e., they are linear combinations of direct product states. As a result, we say that the states $|S = 0, S_z = 0\rangle$ and $|S = 1, S_z = 0\rangle$ possess an “entanglement” between the Hilbert spaces of \vec{S}_1 and \vec{S}_2 . This notion of entanglement can be quantified, and we will delve into this in a later chapter. For now, we will end by noting that these states are an outcome of the fact that it is the total $(\vec{S})^2$ and \vec{S}_z operators whose simultaneous eigenstates we are interested in; the constraint acting on \vec{S}_1 and \vec{S}_2 and their Hilbert spaces in determining the eigenstates of $(\vec{S})^2$ and \vec{S}_z is what leads to the entanglement.

20.2 Adding two total angular momenta \vec{J}_1 and \vec{J}_2

In general, upon adding two angular momenta \vec{J}_1 and \vec{J}_2 , we obtain tensor product states that can be written as

$$|j, m_j; j_1, j_2\rangle = \sum_{m_j=m_1+m_2} \langle j_1, m_1; j_2, m_2 | j, m; j_1, j_2 \rangle |j_1, m_1; j_2, m_2\rangle , \quad (20.34)$$

where the Clebsch-Gordan (or C-G) coefficients connect the $|j_1, m_1; j_2, m_2\rangle$ direct product basis states to the $|j, m_j; j_1, j_2\rangle$ eigenstates, and are given by the inner product $\langle j_1, m_1; j_2, m_2 | j, m; j_1, j_2 \rangle$. For a simple visualisation, recall the $\pm \frac{1}{\sqrt{2}}$ coefficients of the triplet-zero state $|S = 1, S_z = 0\rangle = \frac{1}{\sqrt{2}}(|\uparrow, \downarrow\rangle + |\downarrow, \uparrow\rangle)$ and the singlet state $|S = 0, S_z = 0\rangle = \frac{1}{\sqrt{2}}(|\uparrow, \downarrow\rangle - |\downarrow, \uparrow\rangle)$.

The C-G coefficients have a few properties that are worth noting:

- the C-G coefficients $\langle j_1, m_1; j_2, m_2 | j, m; j_1, j_2 \rangle \neq 0$ only if $m_1 + m_2 = m$.
To see this, we first note that

$$\begin{aligned} J_z |j, m; j_1, j_2\rangle &= (J_{1z} + J_{2z}) |j, m; j_1, j_2\rangle , \\ \implies m |j, m; j_1, j_2\rangle &= (m_1 + m_2) |j, m; j_1, j_2\rangle , \\ \implies m \langle j_1, m_1; j_2, m_2 | j, m; j_1, j_2 \rangle &= (m_1 + m_2) \langle j_1, m_1; j_2, m_2 | j, m; j_1, j_2 \rangle , \\ \implies (m - m_1 - m_2) \langle j_1, m_1; j_2, m_2 | j, m; j_1, j_2 \rangle &= 0 , \end{aligned} \quad (20.35)$$

such that $m = m_1 + m_2$ unless the C-G coefficient $\langle j_1, m_1; j_2, m_2 | j, m; j_1, j_2 \rangle = 0$.

We can see how this works out in the two spin-1/2s problem as follows. By noting that

$$\begin{aligned} |\uparrow, \uparrow\rangle &= |S_1 = 1/2, S_{1z} = 1/2; S_2 = 1/2, S_{2z} = 1/2\rangle , \\ |\downarrow, \downarrow\rangle &= |S_1 = 1/2, S_{1z} = -1/2; S_2 = 1/2, S_{2z} = -1/2\rangle , \\ |\uparrow, \downarrow\rangle &= |S_1 = 1/2, S_{1z} = 1/2; S_2 = 1/2, S_{2z} = -1/2\rangle , \\ |\downarrow, \uparrow\rangle &= |S_1 = 1/2, S_{1z} = -1/2; S_2 = 1/2, S_{2z} = 1/2\rangle , \end{aligned}$$

the above relationship in eq.(20.35) is easily seen here as

$$\begin{aligned} \langle 1/2, 1/2; 1/2, -1/2 | S = 1, S_z = 0 \rangle &\neq 0 , \quad \langle 1/2, -1/2; 1/2, 1/2 | S = 1, S_z = 0 \rangle \neq 0 , \\ \langle 1/2, 1/2; 1/2, 1/2 | S = 1, S_z = 0 \rangle &= 0 = \langle 1/2, -1/2; 1/2, -1/2 | S = 1, S_z = 0 \rangle , \\ \langle 1/2, 1/2; 1/2, 1/2 | S = 1, S_z = 1 \rangle &\neq 0 , \quad \langle 1/2, -1/2; 1/2, -1/2 | S = 1, S_z = -1 \rangle \neq 0 , \\ \langle 1/2, 1/2; 1/2, -1/2 | S = 1, S_z = 1 \rangle &= 0 = \langle 1/2, -1/2; 1/2, 1/2 | S = 1, S_z = 1 \rangle \neq 0 , \\ \langle 1/2, 1/2; 1/2, -1/2 | S = 1, S_z = -1 \rangle &= 0 = \langle 1/2, -1/2; 1/2, 1/2 | S = 1, S_z = -1 \rangle \neq 0 , \\ \langle 1/2, 1/2; 1/2, -1/2 | S = 0, S_z = 0 \rangle &\neq 0 , \quad \langle 1/2, -1/2; 1/2, 1/2 | S = 0, S_z = 0 \rangle \neq 0 , \\ \langle 1/2, 1/2; 1/2, 1/2 | S = 0, S_z = 0 \rangle &= 0 = \langle 1/2, -1/2; 1/2, -1/2 | S = 0, S_z = 0 \rangle \neq 0 . \end{aligned}$$

- the C-G coefficients $\langle j_1, m_1; j_2, m_2 | j, m; j_1, j_2 \rangle \neq 0$ only if $|j_1 - j_2| \leq j \leq j_1 + j_2$.
This is called the “triangle inequality” as it indicates a geometrical triangle with sides j , j_1 and j_2 .

In the two spin-1/2s problem, this is seen in the fact that there are only two classes of eigenstates of S : the three triplet states with $S = 1/2 + 1/2 = 1$ and the lone singlet state with $S = 1/2 - 1/2 = 0$.

- the C-G coefficients are taken to be real by convention. Further, we take the coefficient $\langle j_1, m_1 = j_1; j_2, m_2 = j - j_1 | j, m = j; j_1, j_2 \rangle$ to be positive by convention.

In the two spin-1/2s problem, this is seen from the fact that

$$\langle S = 1, S_z = 1 | S_1 = 1, S_{1z} = 1/2; S_2 = 1, S_{2z} = 1 - 1/2 = 1/2 \rangle \equiv \langle S = 1, S_z = 1 | \uparrow, \uparrow \rangle = 1 , \quad (20.36)$$

and all the Clebsch-Gordan coefficients are either 1 or $\pm \frac{1}{\sqrt{2}}$ (i.e., real-valued).

- $\langle j_1, m_1; j_2, m_2 | j, m; j_1, j_2 \rangle = (-1)^{j_1+j_2-j} \langle j_1, -m_1; j_2, -m_2 | j, -m; j_1, j_2 \rangle .$

This relation tells us that we need to compute only the C-G coefficients for the states with $m \geq 0$, i.e., by starting from the state $|j, m = j; j_1, j_2\rangle$ and work our way down to the state $|j, m = 0; j_1, j_2\rangle$. The coefficients for the states with $m < 0$ are determined by the above relation.

In the two spin-1/2s problem, this is seen from the fact that

$$\langle S = 1, S_z = 1 | \uparrow, \uparrow \rangle = (-1)^{1/2+1/2-1} \langle S = 1, S_z = -1 | \downarrow, \downarrow \rangle . \quad (20.37)$$

This indicates that if we know the C-G coefficient for the $|S = 1, S_z = 1\rangle$ state, we immediately know that for the $|S = 1, S_z = -1\rangle$ state as well.

In order to implement a calculation for the case of the addition of two angular momenta \vec{J}_1 and \vec{J}_2 , the strategy is as follows

- start with the sector $j = j_1 + j_2$, with the state of total angular momentum \vec{J} whose norm is given by $J^2 |j, m_j\rangle = j(j+1)\hbar^2 |j, m_j\rangle$, and whose z -component is given by $J_z |j, m_j\rangle = m_j \hbar |j, m_j\rangle$) such that the state has the largest value of $j = j_1 + j_2 \equiv m_j$, as it is easiest to visualise in the direct product basis:

$$|j = j_1 + j_2, m_j = j_1 + j_2\rangle = |j_1, m_1 = j_1; j_2, m_2 = j_2\rangle \equiv |j_1, m_1 = j_1\rangle \otimes |j_2, m_1 = j_2\rangle . \quad (20.38)$$

Then act with $J_- = J_{1-} + J_{2-}$ on this state to obtain a relation of the type:

$$J_- |j_1 + j_2, j_1 + j_2\rangle = c_1 |j_1, m_1 = j_1 - 1; j_2, m_2 = j_2\rangle + c_2 |j_1, m_1 = j_1; j_2, m_2 = j_2 - 1\rangle , \quad (20.39)$$

where c_1 and c_2 are Clebsch-Gordan coefficients to be determined. This can be done by using the normalisation condition, as well as appropriately chosen orthogonality relations (e.g., with $|j = j_1 + j_2, m_j = j_1 + j_2\rangle$).

- Keep reiterating the action with $J_- = J_{1-} + J_{2-}$ until all states that can be reached from $|j = j_1 + j_2, m_j = j_1 + j_2\rangle$ have their C-G coefficients determined.
- Then, start with the sector $j = j_1 + j_2 - 1$, with the state $|j = j_1 + j_2 - 1, m_j = j_1 + j_2 - 1\rangle$ and proceed as above.
- Continue until all sectors of j and m_j have been mapped out.

20.2.1 $j_1 = 1, j_2 = 1$

Let us now carry out an explicit calculation for the case of $j_1 = 1 = j_2$. First, recall that

$$\begin{aligned} J_{\pm} |j, m_j\rangle &= \hbar \sqrt{j(j+1) - m_j(m_j \pm 1)} |j, m_j \pm 1\rangle , \\ J_{1-} |j_1, m_{j_1}; j_2, m_{j_2}\rangle &= \hbar \sqrt{j_1(j_1+1) - m_{j_1}(m_{j_1}-1)} |j_1, m_{j_1}-1; j_2, m_{j_2}\rangle , \\ J_{2-} |j_1, m_{j_1}; j_2, m_{j_2}\rangle &= \hbar \sqrt{j_2(j_2+1) - m_{j_2}(m_{j_2}-1)} |j_1, m_{j_1}; j_2, m_{j_2}-1\rangle . \end{aligned} \quad (20.40)$$

We begin with the $J = 2$ sector, and consider the normalised eigenstate $|2, 2\rangle = |1, 1; 1, 1\rangle \equiv |1, 1\rangle \otimes |1, 1\rangle$; this state is very easy to guess, as it's the only way we can obtain $m = m_1 + m_2$ for $m = 2$ and is hence simply equivalent to the direct product state $|1, 1\rangle \otimes |1, 1\rangle$. We now act with $J_- = J_{1-} + J_{2-}$ on both sides of the expression for this state to obtain

$$\begin{aligned} J_- |2, 2\rangle &= (J_{1-} + J_{2-}) |1, 1; 1, 1\rangle , \\ \hbar \sqrt{6-2} |2, 1\rangle &= \hbar \sqrt{2} |1, 0; 1, 1\rangle + \hbar \sqrt{2} |1, 1; 1, 0\rangle , \\ \implies |2, 1\rangle &= \frac{1}{\sqrt{2}} (|1, 0; 1, 1\rangle + |1, 1; 1, 0\rangle) . \end{aligned} \quad (20.41)$$

The state $|2, 1\rangle$ is easily seen to be normalised, as well as orthogonal to $|2, 2\rangle$. Also, note that the direct product states that comprise the linear combination leading to $|2, 1\rangle$ are actually easy to guess. We are looking for all possible ways in which to get $m = 1$, and this can only happen in two ways: either $m_1 = 1$ in combination with $m_2 = 0$ (i.e., $|1, 1; 1, 0\rangle$) or $m_1 = 0$ in combination with $m_2 = 1$ (i.e., $|1, 0; 1, 1\rangle$). The C-G coefficients for the two states need to be determined, however, and we have done so above through the action of J_- on $|2, 2\rangle$.

Acting with J_- once more, we obtain

$$\begin{aligned} \sqrt{6} |2, 0\rangle &= \frac{1}{\sqrt{2}} \left[\sqrt{2} |1, -1; 1, 1\rangle + 2\sqrt{2} |1, 0; 1, 0\rangle + \sqrt{2} |1, 1; 1, -1\rangle \right] , \\ \implies |2, 0\rangle &= \frac{1}{\sqrt{6}} |1, -1; 1, 1\rangle + \sqrt{\frac{2}{3}} |1, 0; 1, 0\rangle + \frac{1}{\sqrt{6}} |1, 1; 1, -1\rangle . \end{aligned} \quad (20.42)$$

Again, $|2, 0\rangle$ is normalised as well as orthogonal to $|2, 2\rangle$ and $|2, 1\rangle$. Again, the direct product states that comprise the linear combination leading to $|2, 0\rangle$ are easily guessed. We are looking for all possible ways in which to get $m = 0$, and this can only happen in three ways: $m_1 = 1$ in combination with $m_2 = -1$ (i.e., $|1, 1; 1, -1\rangle$), $m_1 = -1$ in combination with $m_2 = 1$ (i.e., $|1, -1; 1, 1\rangle$) and $m_1 = 0$ with $m_2 = 0$ (i.e., $|1, 0; 1, 0\rangle$). The C-G coefficients for the three states need to be determined, however, and we have done so above through the action of J_- on $|2, 1\rangle$.

Acting with J_- once more, we obtain

$$\begin{aligned} \sqrt{6} |2, -1\rangle &= \sqrt{\frac{2}{6}} |1, -1; 1, 0\rangle + \sqrt{\frac{2}{6}} |1, 0; 1, -1\rangle + 2\sqrt{\frac{2}{6}} |1, -1; 1, 0\rangle + 2\sqrt{\frac{2}{6}} |1, 0; 1, -1\rangle , \\ &= 3\sqrt{\frac{2}{6}} [|1, -1; 1, 0\rangle + |1, 0; 1, -1\rangle] , \\ \implies |2, -1\rangle &= \frac{1}{\sqrt{2}} [|1, -1; 1, 0\rangle + |1, 0; 1, -1\rangle] . \end{aligned} \quad (20.43)$$

Finally, acting with J_- once more, we obtain

$$\begin{aligned}\sqrt{2 \times 3 - 2} |2, -2\rangle &= 2\sqrt{\frac{2}{2}} |1, -1; 1, -1\rangle , \\ \implies |2, -2\rangle &= |1, -1; 1, -1\rangle .\end{aligned}\quad (20.44)$$

In this way, we have obtained the C-G coefficients for all the states that can be obtained for $J = 2$. Note that while we did everything explicitly, we have verified the (earlier stated) fact that the C-G coefficients for the state $|2, 1\rangle$ and $|2, -1\rangle$ are identical to one another, as are those for the states $|2, 2\rangle$ and $|2, -2\rangle$. We could, therefore, have simply written down the states $|2, -2\rangle$ and $|2, -1\rangle$ by noting carefully the form of the states $|2, 2\rangle$ and $|2, 1\rangle$ respectively.

Now, we begin the $J = 1$ sector by considering the state

$$|1, 1\rangle = c_1 |1, 1; 1, 0\rangle + c_2 |1, 0; 1, 1\rangle , \quad c_1^2 + c_2^2 = 1 , \quad (20.45)$$

and c_1 and c_2 are real-valued. Now, we consider the orthogonality of $|1, 1\rangle$ with the state $|2, 1\rangle$ (i.e., a known state that has an identical value of m but a different value of j)

$$\begin{aligned}\langle 2, 1 | 1, 1 \rangle &= \frac{1}{\sqrt{2}} [\langle 1, 1; 1, 0 | + \langle 1, 0; 1, 1 |] [c_1 |1, 1; 1, 0\rangle + c_2 |1, 0; 1, 1\rangle] , \\ &= \frac{1}{\sqrt{2}} (c_1 + c_2) = 0 , \\ \implies c_2 &= -c_1 .\end{aligned}\quad (20.46)$$

Putting this into the normalisation relation above, we obtain $c_1 = 1/\sqrt{2} = -c_2$. Thus,

$$|1, 1\rangle = \frac{1}{\sqrt{2}} [|1, 1; 1, 0\rangle - |1, 0; 1, 1\rangle] . \quad (20.47)$$

Now, acting with J_- on this state, we obtain

$$\begin{aligned}\sqrt{2} |1, 0\rangle &= \frac{1}{\sqrt{2}} [\sqrt{2} |1, 0; 1, 0\rangle + \sqrt{2} |1, 1; 1, -1\rangle - \sqrt{2} |1, -1; 1, 1\rangle - \sqrt{2} |1, 0; 1, 0\rangle] , \\ \implies |1, 0\rangle &= \frac{1}{\sqrt{2}} [|1, 1; 1, -1\rangle - |1, -1; 1, 1\rangle] .\end{aligned}\quad (20.48)$$

Acting with J_- on this state, we obtain

$$\begin{aligned}\sqrt{2} |1, -1\rangle &= \frac{1}{\sqrt{2}} [\sqrt{2} |1, 0; 1, -1\rangle - \sqrt{2} |1, -1; 1, 0\rangle] , \\ \implies |1, -1\rangle &= \frac{1}{\sqrt{2}} [|1, 0; 1, -1\rangle - |1, -1; 1, 0\rangle] .\end{aligned}\quad (20.49)$$

In this way, we have obtained the C-G coefficients for all the states that can be obtained for $J = 1$. Again, the C-G coefficients of $|1, -1\rangle$ are observed to be identical to those of $|1, 1\rangle$; we could simply have written down the state $|1, -1\rangle$ by noting carefully the form of $|1, 1\rangle$.

Finally, we begin the $J = 0$ sector by considering the state

$$|0, 0\rangle = c_1 |1, 1; 1, -1\rangle + c_2 |1, 0; 1, 0\rangle + c_3 |1, -1; 1, 1\rangle \quad , \quad c_1^2 + c_2^2 + c_3^2 = 1 \quad , \quad (20.50)$$

and c_1, c_2, c_3 are real-valued. Now, we consider the orthogonality of $|0, 0\rangle$ with the states $|2, 0\rangle$ and $|1, 0\rangle$ (i.e., both known state that have an identical value of m but different values of j)

$$\begin{aligned} \langle 2, 0 | 0, 0 \rangle &= \frac{c_1}{\sqrt{6}} + \frac{c_3}{\sqrt{6}} + \sqrt{\frac{2}{3}} c_2 = 0 \ , \\ \langle 1, 0 | 0, 0 \rangle &= \frac{c_1}{\sqrt{2}} - \frac{c_3}{\sqrt{2}} = 0 \implies c_1 = c_3 \ , \\ \implies \frac{2c_1}{\sqrt{6}} + \sqrt{\frac{2}{3}} c_2 &= 0 \implies c_2 = -c_1 \ . \end{aligned} \quad (20.51)$$

Putting this into the normalisation condition above gives: $c_1 = -c_2 = c_3 = 1/\sqrt{3}$, and the state $|0, 0\rangle$ as

$$|0, 0\rangle = \frac{1}{\sqrt{3}} [|1, 1; 1, -1\rangle - |1, 0; 1, 0\rangle + |1, -1; 1, 1\rangle] \ . \quad (20.52)$$

In this way, we have determined the C-G coefficients for all the states that can be obtained for all three sectors $J = 2, 1, 0$. It is easily seen that only the $|2, 2\rangle$ and $|2, -2\rangle$ states are direct product in nature. All other states are tensor products, and therefore possess the same property of “entanglement” that was briefly mentioned earlier for the $|1, 0\rangle$ and $|0, 0\rangle$ states of the two spin-1/2s problem.

20.2.2 $j_1 = 1, j_2 = 1/2$

As an another example, we will work out all the C-G coefficients for the case of $j_1 = 1$ and $j_2 = 1/2$. For the $j = 3/2$ sector, we begin from the state $|3/2, 3/2\rangle = |1, 1; 1/2, 1/2\rangle \equiv |1, 1\rangle \otimes |1/2, 1/2\rangle$ and act with $J_- = J_{1-} + J_{2-}$ on both sides

$$\begin{aligned} J_- |3/2, 3/2\rangle &= (J_{1-} + J_{2-}) |1, 1; 1/2, 1/2\rangle \ , \\ \hbar\sqrt{3} |3/2, 1/2\rangle &= \hbar\sqrt{2} |1, 0; 1/2, 1/2\rangle + \hbar |1, 1; 1/2, -1/2\rangle \ , \\ \implies |3/2, 1/2\rangle &= \sqrt{\frac{2}{3}} |1, 0; 1/2, 1/2\rangle + \frac{1}{\sqrt{3}} |1, 1; 1/2, -1/2\rangle \ . \end{aligned} \quad (20.53)$$

Acting with J_- once more, we obtain

$$\begin{aligned} \sqrt{\frac{15+1}{4}} |3/2, -1/2\rangle &= \sqrt{\frac{2}{3}} \sqrt{2} |1, -1; 1/2, 1/2\rangle + \sqrt{\frac{2}{3}} |1, 0; 1/2, -1/2\rangle + \sqrt{\frac{2}{3}} |1, 0; 1/2, -1/2\rangle \ , \\ \text{or, } 2 |3/2, -1/2\rangle &= \frac{2}{\sqrt{3}} |1, -1; 1/2, 1/2\rangle + 2\sqrt{\frac{2}{3}} |1, 0; 1/2, -1/2\rangle \ , \\ \implies |3/2, -1/2\rangle &= \frac{1}{\sqrt{3}} |1, -1; 1/2, 1/2\rangle + \sqrt{\frac{2}{3}} |1, 0; 1/2, -1/2\rangle \ . \end{aligned} \quad (20.54)$$

Acting with J_- once more, we obtain

$$\begin{aligned} \sqrt{\frac{15-3}{4}} |3/2, -3/2\rangle &= \sqrt{\frac{2}{3}}\sqrt{2}|1, -1; 1/2, -1, 2\rangle + \frac{1}{\sqrt{3}}|1, -1; 1/2, -1/2\rangle , \\ \text{or, } \sqrt{3}|3/2, -3/2\rangle &= \frac{2+1}{\sqrt{3}}|1, -1; 1/2, -1/2\rangle , \\ \implies |3/2, -3/2\rangle &= |1, -1; 1/2, -1/2\rangle . \end{aligned} \quad (20.55)$$

In this way, we have obtained all the C-G coefficients for the sector $j = 3/2$. As discussed earlier, we could have guessed the $|3/2, -1/2\rangle$ and $|3/2, -3/2\rangle$ eigenstates by noting carefully the form of the $|3/2, 1/2\rangle$ and $|3/2, 3/2\rangle$ states respectively as their C-G coefficients are identical to one another.

Now, for the case of the $j = 1/2$ sector, we begin with the state

$$|1/2, 1/2\rangle = c_1|1, 0; 1/2, 1/2\rangle + c_2|1, 1; 1/2, -1/2\rangle , \quad c_1^2 + c_2^2 = 1 . \quad (20.56)$$

Considering the orthogonality of $|1/2, 1/2\rangle$ with $|3/2, 1/2\rangle$, we find

$$\begin{aligned} \langle 3/2, 1/2 | 1/2, 1/2 \rangle &= \sqrt{\frac{2}{3}}c_1 + \frac{1}{\sqrt{3}}c_2 = 0 \implies c_2 = -\sqrt{2}c_1 , \\ \implies 3c_1^2 &= 1 \implies c_1 = \frac{1}{\sqrt{3}} , \quad c_2 = -\sqrt{\frac{2}{3}} , \end{aligned} \quad (20.57)$$

and

$$|1/2, 1/2\rangle = \frac{1}{\sqrt{3}}|1, 0; 1/2, 1/2\rangle - \sqrt{\frac{2}{3}}|1, 1; 1/2, -1/2\rangle . \quad (20.58)$$

Acting with J_- on $|1/2, 1/2\rangle$, we obtain

$$\begin{aligned} |1/2, -1/2\rangle &= \sqrt{\frac{2}{3}}|1, -1; 1/2, 1/2\rangle + \frac{1}{\sqrt{3}}|1, 0; 1/2, -1/2\rangle - \sqrt{2}\sqrt{\frac{2}{3}}|1, 0; 1/2, -1/2\rangle , \\ &= \frac{1-2}{\sqrt{3}}|1, 0; 1/2, -1/2\rangle + \sqrt{\frac{2}{3}}|1, -1; 1/2, 1/2\rangle , \\ \implies |1/2, -1/2\rangle &= -\frac{1}{\sqrt{3}}|1, 0; 1/2, -1/2\rangle + \sqrt{\frac{2}{3}}|1, -1; 1/2, 1/2\rangle . \end{aligned} \quad (20.59)$$

In this way, we have now computed all the C-G coefficients for the $j = 1/2$ sector as well, and thence the complete problem. As discussed earlier, we could have guessed the $|1/2, -1/2\rangle$ eigenstate by noting carefully the form of the $|1/2, 1/2\rangle$ state as their C-G coefficients are identical.

It is clear that, armed with a lot more patience and determination, we can compute systematically the C-G coefficients for larger values of j_1 and j_2 .

20.3 Adding Orbital and Spin Angular Momenta

Recall our earlier definitions of total angular momentum for the problem of an electron with spin $S = 1/2$ and an arbitrary orbital angular momentum \vec{L}

$$\vec{J} = \vec{L} + \vec{S}, \quad \left(S^z = \pm \frac{1}{2} \right) . \quad (20.60)$$

We want eigenstates of J^2, J_z, L^2 & S^2 . From our earlier discussions, it is now clear that the direct product (or separable) states $|l, m_l\rangle |\uparrow\rangle$ and $|l, m_l\rangle |\downarrow\rangle$ do not necessarily qualify, as they aren't necessarily eigenstates of J^2 . Presumably, j (the eigenvalue of J^2) has only two values $l + \frac{1}{2}$ and $l - \frac{1}{2}$, and each of these two sectors has states labelled by m_j as lying in the range $[-(l + 1/2), \dots, -1/2, 1/2, \dots, (l + 1/2)]$ and $[-(l - 1/2), \dots, -1/2, 1/2, \dots, (l - 1/2)]$ respectively. The total number of states is obtained as $2(2l + 1) = 4(l + 1/2) = 4j_{max}$. Accordingly, let us denote the simultaneous eigenstates of J^2, J_z, L^2 & S^2 as $|j, m_j; l, s\rangle$.

We will first work with the $j = l + 1/2$ sector. In order to obtain the states for $j = l + 1/2, m_j = l + \frac{1}{2}, \dots, -(l + \frac{1}{2})$, we proceed as follows. First, note that we can start our journey from the state

$$\left| l + \frac{1}{2}, l + \frac{1}{2}; l, \frac{1}{2} \right\rangle \equiv |l, m_l = l\rangle \otimes |\uparrow\rangle , \quad (20.61)$$

as this state with the largest $m_j = l + 1/2$ is one of the few eigenstates of J^2 and J_z that can be easily written down in the direct product basis. Indeed, it is the analog of the $|\uparrow, \uparrow\rangle$ (triplet-one) state in the two spin-1/2s problem. Now,

$$\begin{aligned} J_z \left| l + \frac{1}{2}, l + \frac{1}{2}; l, \frac{1}{2} \right\rangle &= (L_z + S_z) \left| l + \frac{1}{2}, l + \frac{1}{2}; l, \frac{1}{2} \right\rangle \\ &= \hbar \left(l + \frac{1}{2} \right) \left| l + \frac{1}{2}, l + \frac{1}{2}; l, \frac{1}{2} \right\rangle . \end{aligned} \quad (20.62)$$

Further, recall that $J^2 = L^2 + S^2 + 2L_zS_z + L_+S_- + L_-S_+$, where

$$L_{\pm} |l, m_l\rangle = \hbar \sqrt{[(l \pm m_l + 1)(l \mp m_l)]} |l, m_l \pm 1\rangle , \quad S_+ |\downarrow\rangle = \hbar |\uparrow\rangle , \quad S_- |\uparrow\rangle = \hbar |\downarrow\rangle . \quad (20.63)$$

Therefore,

$$\begin{aligned} J^2 \left| l + \frac{1}{2}, l + \frac{1}{2}; l, \frac{1}{2} \right\rangle &= J^2 |l, l\rangle |\uparrow\rangle \\ &= \hbar^2 \left[l(l+1) + \frac{3}{4} + 2l\frac{1}{2} \right] |l, l\rangle |\uparrow\rangle \\ &= \hbar^2 \left(l + \frac{1}{2} \right) \left(l + \frac{3}{2} \right) |l, l\rangle |\uparrow\rangle \\ &= \hbar^2 \left(l + \frac{1}{2} \right) \left(l + \frac{1}{2} + 1 \right) |l, l\rangle |\uparrow\rangle \end{aligned} \quad (20.64)$$

In order to obtain all other $|l + \frac{1}{2}, m + \frac{1}{2}; l, s\rangle$ ($m < l$), we need to apply $J_- = L_- + S_-$ systematically to the above state:

$$J_- \left| l + \frac{1}{2}, l + \frac{1}{2}; l, \frac{1}{2} \right\rangle = \sqrt{2l} \hbar |l, m_l = l-1\rangle |\uparrow\rangle + \hbar |l, m_l = l\rangle |\downarrow\rangle \quad (20.65)$$

which upon using the fact that

$$J_{\pm} |j, m_j\rangle = \hbar \sqrt{j(j+1) - m_j(m_j \pm 1)} |j, m_j \pm 1\rangle , \quad (20.66)$$

gives

$$\begin{aligned} \sqrt{2l+1} \left| l + \frac{1}{2}, l - \frac{1}{2}; l, \frac{1}{2} \right\rangle &= \sqrt{2l} |l, l-1\rangle |\uparrow\rangle + |l, l\rangle |\downarrow\rangle , \\ \implies \left| l + \frac{1}{2}, l - \frac{1}{2}; l, \frac{1}{2} \right\rangle &= \sqrt{\frac{2l}{2l+1}} |l, l-1\rangle |\uparrow\rangle + \sqrt{\frac{1}{2l+1}} |l, l\rangle |\downarrow\rangle . \end{aligned} \quad (20.67)$$

which can be written as

$$\left| l + \frac{1}{2}, l - \frac{1}{2}; l, \frac{1}{2} \right\rangle = \sqrt{\frac{l + (l - \frac{1}{2}) + \frac{1}{2}}{2l+1}} |l, l-1\rangle |\uparrow\rangle + \sqrt{\frac{l - (l - \frac{1}{2}) + \frac{1}{2}}{2l+1}} |l, l\rangle |\downarrow\rangle \quad (20.68)$$

Applying J_- once more gives

$$\begin{aligned} J_- \left| l + \frac{1}{2}, l - \frac{1}{2}; l, \frac{1}{2} \right\rangle &= \sqrt{\left(l + \frac{1}{2} - l + \frac{1}{2} + 1 \right) \times \left(l + \frac{1}{2} + l - \frac{1}{2} \right)} \hbar \left| l + \frac{1}{2}, l - \frac{3}{2}; l, \frac{1}{2} \right\rangle \\ &= \sqrt{\frac{2l}{2l+1}} (L_- + S_-) |l, l-1\rangle |\uparrow\rangle + \sqrt{\frac{1}{2l+1}} (L_- + S_-) |l, l\rangle |\downarrow\rangle \\ &= \sqrt{\frac{2l}{2l+1}} \sqrt{(l - (l-1) + 1)(l + l-1)} \hbar |l, l-2\rangle |\uparrow\rangle + \sqrt{\frac{2l}{2l+1}} \hbar |l, l-1\rangle |\downarrow\rangle \\ &\quad + \sqrt{\frac{1}{2l+1}} \hbar \sqrt{(l - l + 1)(l + l)} |l, l-1\rangle |\downarrow\rangle \\ \implies \sqrt{2.2l} \left| l + \frac{1}{2}, l - \frac{3}{2}; l, \frac{1}{2} \right\rangle &= \sqrt{\frac{2l}{2l+1}} \sqrt{2(2l-1)} |l, l-2\rangle |\uparrow\rangle + 2\sqrt{\frac{2l}{2l+1}} |l, l-1\rangle |\downarrow\rangle \\ \implies \left| l + \frac{1}{2}, l - \frac{3}{2}; l, \frac{1}{2} \right\rangle &= \sqrt{\frac{2l-1}{2l+1}} |l, l-2\rangle |\uparrow\rangle + \sqrt{\frac{2}{2l+1}} |l, l-1\rangle |\downarrow\rangle \end{aligned} \quad (20.69)$$

which can be written as

$$\left| l + \frac{1}{2}, l - \frac{3}{2}; l, \frac{1}{2} \right\rangle = \sqrt{\frac{l + (l - \frac{3}{2}) + \frac{1}{2}}{2l+1}} |l, l-2\rangle |\uparrow\rangle + \sqrt{\frac{l - (l - \frac{3}{2}) + \frac{1}{2}}{2l+1}} |l, l-1\rangle |\downarrow\rangle \quad (20.70)$$

In general, applying J_- in sequence leads to the state

$$\begin{aligned} \left| l + \frac{1}{2}, m_j; l, \frac{1}{2} \right\rangle &= \sqrt{\frac{l + m_j + \frac{1}{2}}{2l+1}} |l, m_l = m_j - 1/2\rangle |\uparrow\rangle + \sqrt{\frac{l - m_j + \frac{1}{2}}{2l+1}} |l, m_l = m_j + 1/2\rangle |\downarrow\rangle \\ &= \sqrt{\frac{j_{max} + m_j}{2j_{max}}} |l, m_l = m_j - 1/2\rangle |\uparrow\rangle + \sqrt{\frac{j_{max} - m_j}{2j_{max}}} |l, m_l = m_j + 1/2\rangle |\downarrow\rangle \end{aligned}$$

where $m_j = l + \frac{1}{2}, l - \frac{1}{2}, l - \frac{3}{2}, \dots, -(l + \frac{1}{2})$ and $j_{max} = l + \frac{1}{2}$. Note that the state $|l + \frac{1}{2}, m_j; l, \frac{1}{2}\rangle$ can only be comprised of two direct product combination of states from the l and s Hilbert spaces: either $m_l = m_j - 1/2$ and $m_s = 1/2$ (i.e., $|l, m_l = m_j - 1/2\rangle |\uparrow\rangle$) or $m_l = m_j + 1/2$ and $m_s = -1/2$ (i.e., $|l, m_l = m_j + 1/2\rangle |\downarrow\rangle$). In that sense, these states are analogous to the $\frac{1}{\sqrt{2}}(|\uparrow, \downarrow\rangle + |\downarrow, \uparrow\rangle)$ (triplet-zero) state in the two spin-1/2s problem. Thus, the action of J_- on the $|l + \frac{1}{2}, m_j + 1; l, \frac{1}{2}\rangle$ has revealed the C-G coefficients.

Finally, note that the state $|j = l + \frac{1}{2}, m_j = -l - \frac{1}{2}; l, \frac{1}{2}\rangle$ (i.e., the state with lowest m_j eigenvalue that can be reached through repeated application of J_-) is equivalent to a direct product state that can be visualised easily

$$\left| j = l + \frac{1}{2}, m_j = -l - \frac{1}{2}; l, \frac{1}{2} \right\rangle \equiv |l, m_l = -l\rangle \otimes |\downarrow\rangle , \quad (20.73)$$

and is analogous to the $|\downarrow, \downarrow\rangle$ (triplet-minus one) state of the two spin-1/2s problem.

We now turn to the $j = l - 1/2$ sector. In order to get the states for $j = l - \frac{1}{2}, m_j = l - \frac{1}{2}, \dots, -(l - \frac{1}{2})$, let us proceed as follows. We note that the state $|l - \frac{1}{2}, l - \frac{1}{2}; l, \frac{1}{2}\rangle$ can be written as

$$\left| l - \frac{1}{2}, l - \frac{1}{2}; l, \frac{1}{2} \right\rangle = c_1 |l, m_l = l\rangle \otimes |\downarrow\rangle + c_2 |l, m_l = l - 1\rangle \otimes |\uparrow\rangle , \quad (20.74)$$

such that it's orthogonality with the state $|j = l + 1/2, m_j = l - 1/2; l, 1/2\rangle = \sqrt{\frac{2l}{2l+1}} |l, l - 1\rangle |\uparrow\rangle + \sqrt{\frac{1}{2l+1}} |l, l\rangle |\downarrow\rangle$ obtains

$$\begin{aligned} \langle l + 1/2, l - 1/2; l, 1/2 | l - 1/2, l - 1/2; l, 1/2 \rangle &= c_2 \sqrt{\frac{2l}{2l+1}} + c_1 \sqrt{\frac{1}{2l+1}} = 0 , \\ \implies \sqrt{2l} c_2 &= -c_1 . \end{aligned} \quad (20.75)$$

Putting this together with the normalisation condition $c_1^2 + c_2^2 = 1$, we get

$$c_1 = -\sqrt{\frac{2l}{2l+1}} , \quad c_2 = \sqrt{\frac{1}{2l+1}} , \quad (20.76)$$

and

$$\left| l - \frac{1}{2}, l - \frac{1}{2}; l, \frac{1}{2} \right\rangle = -\sqrt{\frac{2l}{2l+1}} |l, l\rangle |\downarrow\rangle + \sqrt{\frac{1}{2l+1}} |l, l - 1\rangle |\uparrow\rangle , \quad (20.77)$$

which can be written as

$$\left| l - \frac{1}{2}, l - \frac{1}{2}; l, \frac{1}{2} \right\rangle = -\sqrt{\frac{l + (l - \frac{1}{2}) + \frac{1}{2}}{2l+1}} |l, l\rangle |\downarrow\rangle + \sqrt{\frac{l - (l - \frac{1}{2}) + \frac{1}{2}}{2l+1}} |l, l - 1\rangle |\uparrow\rangle \quad (20.78)$$

Note that the state $|l - 1/2, l - 1/2; l, 1/2\rangle$ looks somewhat similar to $|l + 1/2, l - 1/2; l, 1/2\rangle$.

Acting with J_- on this state, we find

$$\begin{aligned}
J_- |l - 1/2, l - 1/2; l, 1/2\rangle &= (L_- + S_-) \left(-\sqrt{\frac{2l}{2l+1}} |l, l\rangle |\downarrow\rangle + \sqrt{\frac{1}{2l+1}} |l, l-1\rangle |\uparrow\rangle \right) , \\
\hbar\sqrt{2l-1} |l - 1/2, l - 3/2; l, 1/2\rangle &= -\frac{2l-1}{\sqrt{2l+1}} \hbar |l, l-1\rangle |\downarrow\rangle + \sqrt{\frac{2(2l-1)}{2l+1}} \hbar |l, l-2\rangle |\uparrow\rangle , \\
|l - 1/2, l - 3/2; l, 1/2\rangle &= -\sqrt{\frac{2l-1}{2l+1}} |l, l-1\rangle |\downarrow\rangle + \sqrt{\frac{2}{2l+1}} |l, l-2\rangle |\uparrow\rangle , \\
&= -\sqrt{\frac{l+(l-\frac{3}{2})+\frac{1}{2}}{2l+1}} |l, l-1\rangle |\downarrow\rangle + \sqrt{\frac{l-(l-\frac{3}{2})+\frac{1}{2}}{2l+1}} |l, l-2\rangle |\uparrow\rangle ,
\end{aligned} \tag{20.79}$$

Note that the state $|l - 1/2, l - 3/2; l, 1/2\rangle$ looks somewhat similar to $|l + 1/2, l - 3/2; l, 1/2\rangle$. Indeed, we can also obtain $|l - 1/2, l - 3/2; l, 1/2\rangle$ by writing it as

$$|l - 1/2, l - 3/2; l, 1/2\rangle = c_1 |l, l-1\rangle |\downarrow\rangle + c_2 |l, l-2\rangle |\uparrow\rangle , \tag{20.81}$$

and considering the orthogonality of this state with the state $|l + 1/2, l - 3/2; l, 1/2\rangle = \sqrt{\frac{2l-1}{2l+1}} |l, l-2\rangle |\uparrow\rangle + \sqrt{\frac{2}{2l+1}} |l, l-1\rangle |\downarrow\rangle$ obtains

$$\begin{aligned}
\langle l + 1/2, l - 3/2; l, 1/2 | l - 1/2, l - 3/2; l, 1/2 \rangle &= \sqrt{\frac{2}{2l+1}} c_1 + \sqrt{\frac{2l-1}{2l+1}} c_2 = 0 , \\
\Rightarrow \sqrt{2} c_1 &= -\sqrt{2l-1} c_2 ,
\end{aligned} \tag{20.82}$$

which together with the normalisation condition $c_1^2 + c_2^2 = 1$ gives

$$c_1 = -\sqrt{\frac{2l-1}{2l+1}} , \quad c_2 = \sqrt{\frac{2}{2l+1}} \tag{20.83}$$

precisely as that obtained above.

To proceed, one can obtain all the states $|l - 1/2, m_j; l, 1/2\rangle$ either by acting with J_- repeatedly, or to construct states orthogonal to those given above earlier for $|l + 1/2, m_j; l, 1/2\rangle$:

$$|l - \frac{1}{2}, m_j; l, \frac{1}{2}\rangle = -\sqrt{\frac{l-m_j+\frac{1}{2}}{2l+1}} |l, m_l = m_j - 1/2\rangle |\uparrow\rangle + \sqrt{\frac{l+m_j+\frac{1}{2}}{2l+1}} |l, m_l = m_j + 1/2\rangle |\downarrow\rangle \tag{20.84}$$

$$= -\sqrt{\frac{j_{max}-m_j}{2j_{max}}} |l, m_l = m_j - 1/2\rangle |\uparrow\rangle + \sqrt{\frac{j_{max}+m_j}{2j_{max}}} |l, m_l = m_j + 1/2\rangle |\downarrow\rangle \tag{20.85}$$

where $m_j = l - \frac{1}{2}, l - \frac{3}{2}, \dots, -(l - \frac{1}{2})$ and $j_{max} = l + \frac{1}{2}$. Note that, once again, the state $|l - \frac{1}{2}, m_j; l, \frac{1}{2}\rangle$ can only be comprised of two direct product combination of states from the l and s Hilbert spaces: either $m_l = m_j - 1/2$ and $m_s = 1/2$ (i.e., $|l, m_l = m_j - 1/2\rangle |\uparrow\rangle$) or $m_l = m_j + 1/2$ and $m_s = -1/2$ (i.e., $|l, m_l = m_j + 1/2\rangle |\downarrow\rangle$). In that sense, these states are analogous to the $\frac{1}{\sqrt{2}}(|\uparrow, \downarrow\rangle - |\downarrow, \uparrow\rangle)$ (singlet) state in the two spin-1/2s problem. Thus, the action of J_- on the $|l + \frac{1}{2}, m_j + 1; l, \frac{1}{2}\rangle$ has revealed the C-G coefficients.

Thus, we can bring together the two relations for $|l + \frac{1}{2}, m_j; l, \frac{1}{2}\rangle$ (eq.(20.72)) and $|l - \frac{1}{2}, m_j; l, \frac{1}{2}\rangle$ (eq.(20.85)) as

$$\left|l \pm \frac{1}{2}, m_j; l, \frac{1}{2}\right\rangle = \alpha_{\pm} |l, m_l = m_j - 1/2\rangle |\uparrow\rangle + \beta_{\pm} |l, m_l = m_j + 1/2\rangle |\downarrow\rangle \quad (20.86)$$

$$\alpha_{\pm} = \pm \sqrt{\frac{j_{max} \pm m_j}{2j_{max}}} = \pm \beta_{\mp} \implies \frac{\beta_+}{\alpha_+} = \sqrt{\frac{j_{max} - m_j}{j_{max} + m_j}} = -\frac{\alpha_-}{\beta_-}, \quad (20.87)$$

$$\alpha_+^2 + \beta_+^2 = 1 = \alpha_-^2 + \beta_-^2, \quad (20.88)$$

$$\alpha_+ \alpha_- + \beta_+ \beta_- = 0. \quad (20.89)$$

α_{\pm} and β_{\pm} are examples of the so-called “Clebsch-Gordan” coefficients that connect the $|l, m_l; s, m_s\rangle$ basis kets to the $|j, m_j; l, s\rangle$ basis kets. It is quite satisfying to be able to compute the C-G coefficients, and the related eigenstates, for an arbitrary \vec{L} .

As we will see elsewhere, these specific $|j, m_j; l, s\rangle$ basis states are important in dealing with the problem of spin-orbit interaction (which involves the $\vec{L} \cdot \vec{S}$ operator). As observed earlier with the triplet zero ($|1, 0\rangle = \frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle)$) and singlet ($|0, 0\rangle = \frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle)$) states of the two spin-1/2s problem, the tensor product nature of the states in eq.(20.86) indicates the presence of “entanglement” between the (real-space) orbital momentum Hilbert space of states and the spin angular momentum Hilbert space of states. Indeed, the relationship between the states $|l + \frac{1}{2}, m_j; l, \frac{1}{2}\rangle$ and $|l - \frac{1}{2}, m_j; l, \frac{1}{2}\rangle$ appear to be akin to that between the triplet-zero ($|1, 0\rangle = \frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle)$) and singlet ($|0, 0\rangle = \frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle)$) states in the two spin-1/2s problem.

Finally, we can extend these results to the case of the addition of a general orbital angular momentum \vec{L} and a general spin angular momentum \vec{S} . The computation of the Clebsch-Gordan coefficients will be much more tedious, of course, but it can be done.

Chapter 21

Symmetries in Quantum Mechanics

21.1 Symmetries, Conservation laws and Degeneracies

In an earlier chapter, we noted that symmetries of a quantum system were tied to conservation laws that were obeyed by the system. Recall that if a Hermitian operator $\hat{\Theta}$ commutes with the Hamiltonian operator \hat{H} , $[\hat{H}, \hat{\Theta}] = 0$, both \hat{H} and Θ possess a simultaneous eigenbasis, say, $|\psi\rangle$, with the eigenvalues of both operators corresponding to good quantum numbers. Further, the trivial commutation relation implies that the corresponding physical observable satisfies a conservation law, i.e., an invariance in time. This can be seen simply from the fact that since the unitary time evolution operator is given by $U_t = e^{-i\hat{H}t/\hbar}$ (i.e., U_t is a function of the Hamiltonian \hat{H}),

$$\begin{aligned} [\hat{H}, \hat{\Theta}] &= 0, \\ \implies [U_t, \hat{\Theta}] &= 0, \\ \implies U_t \hat{\Theta} &= \hat{\Theta} U_t, \\ \implies U_t^\dagger U_t \hat{\Theta} &= U_t^\dagger \hat{\Theta} U_t, \\ \implies \hat{\Theta} &= U_t^\dagger \hat{\Theta} U_t, \quad (\text{as } U_t^\dagger U_t = 1) \end{aligned} \tag{21.1}$$

i.e., the $\hat{\Theta}$ operator does not evolve in time, and thus the eigenvalue of this time-independent $\hat{\Theta}$ operator cannot change in time either. This can also be seen by noting that the Heisenberg equation of motion (which we will derive below) for Θ is given by

$$i\hbar \frac{d\Theta}{dt} = [\hat{\Theta}, \hat{H}] = 0. \tag{21.3}$$

Furthermore, the operator Θ characterises a corresponding symmetry of the quantum system. In order to see this, let us consider the unitary operator $U_\Theta = e^{-i\epsilon\Theta/\hbar}$, where ϵ is a parameter corresponding to the change (or transformation) that is affected by acting the operator U_Θ on a state $|\psi\rangle$:

$$U_\Theta |\psi(\epsilon = 0)\rangle = |\psi(\epsilon)\rangle. \tag{21.4}$$

In this way, the operator Θ is observed to be the “generator” of the transformation brought about by the action of U_Θ . As we will see shortly, for $\Theta \equiv \hat{p}$ (the linear momentum), ϵ

corresponds to a translation of the position eigenstate $\psi(x)$ to a new position eigenstate $\psi'(x) \equiv \psi(x - \epsilon)$:

$$\psi(x) \rightarrow \psi'(x) \equiv \langle x | \psi' \rangle = \langle x | U_{\hat{p}} | \psi \rangle = \psi(x - \epsilon) . \quad (21.5)$$

Further, as $U_{\Theta} = e^{-i\epsilon\Theta/\hbar}$, the commutator $[\hat{H}, \hat{\Theta}] = 0$ leads to

$$\begin{aligned} [\hat{H}, U_{\Theta}] &= 0 , \\ \implies \hat{H}U_{\Theta} &= U_{\Theta}\hat{H} , \\ \implies U_{\Theta}^{\dagger}\hat{H}U_{\Theta} &= U_{\Theta}^{\dagger}U_{\Theta}\hat{H} , \\ U_{\Theta}^{\dagger}HU_{\Theta} &= H . \quad (\text{as } U_{\Theta}^{\dagger}U_{\Theta} = 1) \end{aligned} \quad (21.6)$$

This shows that the Hamiltonian H is unchanged, i.e., *invariant*, under the application of the transformation U_{Θ} . The invariance reflects a symmetry of H , such that its eigenvalues (E) are unchanged under the application of U_{Θ} :

$$\begin{aligned} U_{\Theta}H|\psi\rangle &= U_{\Theta}E|\psi\rangle , \\ \implies HU_{\Theta}|\psi\rangle &= EU_{\Theta}|\psi\rangle , \end{aligned} \quad (21.8)$$

which shows that, if $|\psi\rangle$ and $U_{\Theta}|\psi\rangle$ are distinct states, both have identical energy eigenvalues (even if they possess different quantum number eigenvalues for Θ !). Thus, symmetries (and their corresponding conservation laws) can lead to degeneracies in the energy eigenspectrum. It is worth noting that the mathematical connection between symmetries and conservation laws in classical mechanics and classical field theory was first pointed out by Emmy Noether; these will be encountered by you in courses on those subjects. In quantum many-particle systems comprised of constituents that are interacting with one another, similar results were reached by Ward and Takahashi, and are often taught in advanced courses in quantum field theory.

Finally, we end by enquiring on what happens when multiple transformations are applied in sequence. For instance, for

$$\begin{aligned} |\psi'\rangle &= U(\Theta_1)U(\Theta_2)|\psi\rangle , \quad |\tilde{\psi}\rangle = U(\Theta_2)U(\Theta_1)|\psi\rangle , \\ \text{is } |\psi'\rangle &= |\tilde{\psi}\rangle \end{aligned} \quad (21.9)$$

upto an overall phase? If this is indeed the case, i.e., $|\psi'\rangle = |\tilde{\psi}\rangle$, then clearly

$$\begin{aligned} U(\Theta_1)U(\Theta_2) &= U(\Theta_2)U(\Theta_1) , \\ \implies [U(\Theta_1), U(\Theta_2)] &= 0 , \end{aligned} \quad (21.10)$$

i.e., the generators $U(\Theta_1)$ and $U(\Theta_2)$ of the transformation commute with one another. In such cases, we say that the mathematical structure to which all such $U(\Theta_i)$ operators — the “group” — is Abelian. We will see that the group of linear transformations in 3D is Abelian. On the other hand, if $[U(\Theta_1), U(\Theta_2)] \neq 0$, we say that the corresponding group is non-Abelian. The group of rotational transformations in 3D corresponds to one such example.

21.2 A few important points on Unitary Operators

We saw above that unitary operators such as U_Θ appear to be very important in our discussion of symmetries, conservation laws and degeneracies. But why unitary operators? The answer lies in recognising the fact transformations such as U_Θ which act on members of the Hilbert space $U_\Theta |\psi\rangle = |\psi'\rangle$ must preserve all inner products between them

$$\begin{aligned} \langle\phi|\psi\rangle \rightarrow \langle\phi'|\psi'\rangle &= (\langle\phi| U_\Theta^\dagger) U_\Theta |\psi\rangle = \langle\phi|\psi\rangle , \\ \implies U_\Theta^\dagger U_\Theta &= 1 , \\ \implies U_\Theta^\dagger &= U_\Theta^{-1} , \end{aligned} \quad (21.11)$$

indicating that U_Θ is a unitary operator. Indeed, transformation operators in quantum mechanics must be either unitary or anti-unitary (i.e., an operator U whose action on ψ leads to $U(c|\psi\rangle) = c^*U|\psi\rangle$, e.g., the time-reversal transformation operation).

Here are a few more interesting properties of unitary operators. First, for a Hermitian operator Θ , we can always form a unitary operator $U_\Theta(\lambda) = e^{-i\lambda\Theta}$ where λ is real-valued:

$$U_\Theta^\dagger = e^{i\lambda\Theta} , \quad U_\Theta^\dagger U_\Theta = 1 . \quad (21.12)$$

Further, the eigenvectors of Θ correspond to the eigenvectors of U_Θ , and the eigenvalues of Θ are related to those of U_Θ :

$$\begin{aligned} \hat{\Theta} |\psi\rangle &= \Theta |\psi\rangle , \\ U_\Theta |\psi\rangle &= e^{-i\lambda\hat{\Theta}} |\psi\rangle = e^{-i\lambda\Theta} |\psi\rangle , \end{aligned} \quad (21.13)$$

where the eigenvalues of U_Θ are phases given by $e^{-i\lambda\Theta}$.

An important relation that will often prove useful is the following. For two Hermitian operators Θ and A , such that the unitary operator $U_\Theta = e^{-i\lambda\Theta}$, we have the Baker-Campbell-Hausdorff relation (which we will not prove here)

$$\begin{aligned} U_\Theta^\dagger A U_\Theta &= e^{i\lambda\Theta} A e^{-i\lambda\Theta} , \\ &= A + i\lambda [\Theta, A] + \frac{(i\lambda)^2}{2!} [\Theta, [\Theta, A]] + \dots + \frac{(i\lambda)^n}{n!} [\Theta, [\Theta, [\Theta, \dots [\Theta, A]]] \dots] \end{aligned} \quad (21.14)$$

For the special case of $[\Theta, A] = 0$, all the commutators on the right hand side vanish identically, and the above simplifies to

$$U_\Theta^\dagger A U_\Theta = e^{i\lambda\Theta} A e^{-i\lambda\Theta} = A . \quad (21.15)$$

That this is clearly the case for any operator Θ that commutes with the Hamiltonian H was shown above, and observed to lead to the invariance of H under the transformation U_Θ . Finally, for $A(\lambda) = U_\Theta^\dagger A U_\Theta = e^{i\lambda\Theta} A e^{-i\lambda\Theta}$, it can be seen that

$$\frac{dA(\lambda)}{d\lambda} = i [\Theta, A(\lambda)] . \quad (21.16)$$

This can be easily derived by using the Baker-Campbell-Hausdorff relation for $A(\lambda)$ given above:

$$\begin{aligned}
\frac{dA(\lambda)}{d\lambda} &= \frac{d}{d\lambda} (U_\Theta^\dagger A U_\Theta) = \frac{d}{d\lambda} (e^{i\lambda\Theta} A e^{-i\lambda\Theta}) , \\
&= \frac{d}{d\lambda} (A + i\lambda [\Theta, A] + \frac{(i\lambda)^2}{2!} [\Theta, [\Theta, A]] + \dots + \frac{(i\lambda)^n}{n!} [\Theta, [\Theta, [\Theta, \dots [\Theta, A]]] \dots] + \dots) , \\
&= i [\Theta, A] + \frac{2i^2\lambda}{2!} [\Theta, [\Theta, A]] + \frac{3i^3\lambda^2}{3!} [\Theta, [\Theta, [\Theta, A]]] + \dots , \\
&= i [\Theta, A] + i(i\lambda) [\Theta, [\Theta, A]] + i \frac{(i\lambda)^2}{2!} [\Theta, [\Theta, [\Theta, A]]] + \dots , \\
&= i \left[\Theta, (A + i\lambda [\Theta, A] + \frac{(i\lambda)^2}{2!} [\Theta, [\Theta, A]] + \frac{(i\lambda)^3}{3!} [\Theta, [\Theta, [\Theta, A]]] + \dots) \right] , \\
&= i [\Theta, A(\lambda)] .
\end{aligned} \tag{21.17}$$

A special case of the above relation corresponds to $\hat{\Theta}$ being the Hamiltonian H , $\hat{\Theta} \equiv H$, and $\lambda \equiv t$, such that $U_\Theta(t) \equiv U_t = e^{-itH/\hbar}$ corresponds to the time (t) evolution operator and $A(t) = U_t^\dagger A U_t = e^{itH/\hbar} A e^{-itH/\hbar}$ and

$$\frac{dA(t)}{dt} = \frac{i}{\hbar} [H, A(t)] , \tag{21.18}$$

$$\begin{aligned}
\implies -i\hbar \frac{dA(t)}{dt} &= [H, A(t)] , \\
\implies i\hbar \frac{dA(t)}{dt} &= [A(t), H] ,
\end{aligned} \tag{21.19}$$

which is the Heisenberg equation of motion for the operator $A(t)$. We will explore the origin of this equation of motion for the operators such as A shortly.

21.3 Continuous Symmetries: Translation and Rotation symmetry

21.3.1 Spatial translation invariance

Let us consider a three dimensional system with a single particle in a position basis eigenket $|\vec{r}\rangle$ and apply an infinitesimal translation of the coordinates $\vec{r} \rightarrow \vec{r} + \delta\vec{r}$ through a unitary operator $U(\delta\vec{r})$:

$$|\vec{r} + \delta\vec{r}\rangle = U(\delta\vec{r}) |\vec{r}\rangle . \tag{21.20}$$

Noting that the real space representation of the wavefunction is given by

$$\psi(\vec{r}) = \langle \vec{r} | \psi \rangle , \tag{21.21}$$

we can write the action of the linear translation transformation $U(\delta\vec{r})$ on a general state $|\psi\rangle$ as

$$|\psi_{\delta\vec{r}}\rangle = U(\delta\vec{r})|\psi\rangle , \quad (21.22)$$

$$= U(\delta\vec{r}) \int_{-\infty}^{\infty} d\vec{r} |\vec{r}\rangle \langle \vec{r}|\psi\rangle , \quad (21.23)$$

$$= \int_{-\infty}^{\infty} d\vec{r} U(\delta\vec{r}) |\vec{r}\rangle \langle \vec{r}|\psi\rangle , \quad (21.24)$$

$$= \int_{-\infty}^{\infty} d\vec{r} |\vec{r} + \delta\vec{r}\rangle \langle \vec{r}|\psi\rangle , \quad (21.25)$$

$$= \int_{-\infty}^{\infty} d\vec{r}' |\vec{r}'\rangle \langle \vec{r}' - \delta\vec{r}|\psi\rangle , \quad (\vec{r}' = \vec{r} + \delta\vec{r}) \quad (21.26)$$

$$\implies \langle \vec{r}|\psi_{\delta\vec{r}}\rangle = \langle \vec{r}|U(\delta\vec{r})|\psi\rangle , \quad (21.27)$$

$$= \langle \vec{r}| \int_{-\infty}^{\infty} d\vec{r}' |\vec{r}'\rangle \langle \vec{r}' - \delta\vec{r}|\psi\rangle , \quad (21.28)$$

$$= \int_{-\infty}^{\infty} d\vec{r}' \langle \vec{r}|\vec{r}'\rangle \langle \vec{r}' - \delta\vec{r}|\psi\rangle , \quad (21.29)$$

$$= \int_{-\infty}^{\infty} d\vec{r}' \delta(\vec{r} - \vec{r}') \langle \vec{r}' - \delta\vec{r}|\psi\rangle , \quad (\langle \vec{r}|\vec{r}'\rangle = \delta(\vec{r} - \vec{r}'))$$

$$= \langle \vec{r} - \delta\vec{r}|\psi\rangle ,$$

$$= \psi(\vec{r} - \delta\vec{r}) . \quad (21.30)$$

This indicates that the application of the translation operator $U(\delta\vec{r})$ has translated the real space wavefunction from $\psi(\vec{r})$ by an amount $\delta\vec{r}$ to $\psi(\vec{r} - \delta\vec{r})$ without distorting it in any way. Here's an example to help you visualise the action of the translation: consider the original wavefunction to be a 1D gaussian with its center at $x = 0$, $\psi(x) \sim e^{-x^2}$, and the translated wavefunction to be a gaussian with its center shifted to $x = \delta x$, $\psi(x - \delta x) \sim e^{-(x - \delta x)^2}$. We can also check that the translation indeed causes a shift in the position \vec{r}

$$\begin{aligned} \int_{-\infty}^{\infty} d\vec{r} \psi^*(\vec{r} - \delta\vec{r}) \vec{r} \psi(\vec{r} - \delta\vec{r}) &= \int_{-\infty}^{\infty} d\vec{r}' \psi^*(\vec{r}') (\vec{r}' + \delta\vec{r}) \psi(\vec{r}') , \quad (\vec{r}' = \vec{r} - \delta\vec{r}) \\ &= \int_{-\infty}^{\infty} d\vec{r}' \psi^*(\vec{r}') \vec{r}' \psi(\vec{r}') + \delta\vec{r} \int_{-\infty}^{\infty} d\vec{r}' \psi^*(\vec{r}') \psi(\vec{r}') , \\ \langle \vec{r} \rangle &\xrightarrow{U_{\delta\vec{r}}} \langle \vec{r} \rangle + \delta\vec{r} . \end{aligned} \quad (21.31)$$

Further, one can verify that the momentum remains unchanged under the translation

$$\begin{aligned} \int_{-\infty}^{\infty} d\vec{r} \psi^*(\vec{r} - \delta\vec{r}) (-i\hbar \vec{\nabla}_{\vec{r}}) \psi(\vec{r} - \delta\vec{r}) &= \int_{-\infty}^{\infty} d\vec{r}' \psi^*(\vec{r}') (-i\hbar \vec{\nabla}_{\vec{r}'}) \psi(\vec{r}') , \quad (\vec{r}' = \vec{r} - \delta\vec{r}) \\ \implies \langle \vec{p} \rangle_{\vec{r}} &\xrightarrow{U_{\delta\vec{r}}} \langle \vec{p} \rangle_{\vec{r}'} . \end{aligned} \quad (21.32)$$

In order to learn the form of $U(\delta\vec{r})$, we can Taylor expand the state $\psi(\vec{r})$ about \vec{r} to obtain

$$\psi(\vec{r} - \delta\vec{r}) = \psi(\vec{r}) - \left(\vec{\nabla} \psi(\vec{r}) \right) \cdot \delta\vec{r} + \dots = (1 - \delta\vec{r} \cdot \vec{\nabla} + \dots) \psi(\vec{r}) . \quad (21.33)$$

From what we have learnt above, this identifies the linear translation operator as

$$\begin{aligned} U(\delta\vec{r}) &= 1 - \delta\vec{r} \cdot \vec{\nabla} + \dots, \\ &= (1 - \frac{i}{\hbar} \delta\vec{r} \cdot (-i\hbar\vec{\nabla}) + \dots), \\ &= e^{-\frac{i}{\hbar} \delta\vec{r} \cdot \vec{p}_{\vec{r}}}, \end{aligned} \quad (21.34)$$

where $\vec{p}_{\vec{r}} = -i\hbar\vec{\nabla}$. Written in terms of its components, the translation operator (a unitary operator!) is given by

$$U(\delta x) = e^{-i\delta x \hat{p}_x / \hbar}, \quad U(\delta p) = e^{-i\delta y \hat{p}_y / \hbar}, \quad U(\delta z) = e^{-i\delta z \hat{p}_z / \hbar}. \quad (21.35)$$

Given that

$$\begin{aligned} [\hat{p}_i, \hat{p}_j] &= 0, \quad (i, j) = (x, y, z) \\ \implies [U(\delta i), U(\delta j)] &= 0, \quad \text{and} \quad [\hat{p}_i, U(\delta j)] = 0. \end{aligned} \quad (21.36)$$

Further, if for the same system, we have a Hamiltonian H such that

$$[p_i, H] = 0, \quad i = (x, y, z) \implies [U(\delta i), H] = 0, \quad i = (x, y, z), \quad (21.37)$$

we can see that the momentum of the system is conserved in time:

$$\frac{dp_i}{dt} = 0 \quad \text{for } i = (x, y, z), \quad (21.38)$$

i.e., there is no external imbalanced force acting on the system.

Similarly, we can argue that if we have a system comprised of N particles, translating the entire system (i.e., for each of the N particles) from $\vec{r} \rightarrow \vec{r} + \delta\vec{r}$, it is the total momentum $\vec{P}_{tot} = \sum_{i=1}^N \vec{p}_i$ that is conserved in time in the absence of an external imbalanced force acting on the system

$$\begin{aligned} [\vec{P}_{tot}, H] &= 0, \quad [U(\delta\vec{r}), H] = 0 \\ \implies \frac{d\vec{P}_{tot}}{dt} &= 0, \end{aligned} \quad (21.39)$$

where $U(\delta\vec{r}) = e^{-i\delta\vec{r} \cdot \vec{P}_{tot} / \hbar}$.

At the level of operators, the translation $\vec{r} \rightarrow \vec{r} + \delta\vec{r}$ affects the position coordinates $(x, y, z)_n$, momenta $(p_x, p_y, p_z)_n$ for the n th particle and the Hamiltonian (H) as follows:

$$\begin{aligned} x_n - \delta r_x &= U^\dagger(\delta\vec{r}) x_n U(\delta\vec{r}), \quad y_n - \delta r_y = U^\dagger(\delta\vec{r}) y_n U(\delta\vec{r}), \quad z_n - \delta r_z = U^\dagger(\delta\vec{r}) z_n U(\delta\vec{r}), \\ p_{x,n} &= U^\dagger(\delta\vec{r}) p_{x,n} U(\delta\vec{r}), \quad p_{y,n} = U^\dagger(\delta\vec{r}) p_{y,n} U(\delta\vec{r}), \quad p_{z,n} = U^\dagger(\delta\vec{r}) p_{z,n} U(\delta\vec{r}), \\ H(\vec{R}_{tot} - \delta\vec{r}, \vec{P}_{tot}) &= U^\dagger(\delta\vec{r}) H(\vec{R}_{tot}, \vec{P}_{tot}) U(\delta\vec{r}), \end{aligned} \quad (21.40)$$

where $\vec{R}_{tot} = \sum_{i=1}^N \vec{r}_i$ is the center of mass position coordinate (upto a factor of N^{-1}) and is the operator conjugate to the total momentum \vec{P}_{tot} .

We can now see that if, along with \vec{P}_{tot} , the Hamiltonian is invariant under translation, we must have

$$H(\vec{R}_{tot} - \delta\vec{r}, \vec{P}_{tot}) = U^\dagger(\delta\vec{r})H(\vec{R}_{tot}, \vec{P}_{tot})U(\delta\vec{r}) = H(\vec{R}_{tot}, \vec{P}_{tot}). \quad (21.41)$$

This indicates that the Hamiltonian H (or rather the potential term V , which is a function of \vec{r}_i) must be a function of only the coordinate differences (i.e., $\vec{r}_i - \vec{r}_j$ for all $j \neq i$). Thus, we can see that this is telling us that translation invariance is a property of all composite systems whose constituents are interacting with one another but nothing external. Every known fundamental interaction between matter particles — gravitational, electromagnetic, weak and strong nuclear forces — all have this property of translation invariance. It is clear from our arguments that such inter-particle interactions cannot affect \vec{P}_{tot} , i.e., the momentum of the system as a whole is conserved in the presence of such translationally invariant inter-particle interactions.

Further, translation invariance also implies that local observers will obtain the same results for identical experiments that are repeated at different places, and these results will be distinct from one another by only a spatial translation factor that is constant in time. Indeed, the translation invariance of natural laws reflects the homogeneity (or uniformity) of space: the physical laws should not change depending on where we choose to study them. This will allow us to compare, corroborate and correct our results obtained on the mechanics and dynamics of identical quantum systems but in different parts of the universe.

21.3.2 Time Translation invariance

We have seen above that $U_t = e^{-itH/\hbar}$ corresponds to the time (t) evolution operator and $A(t) = U_t^\dagger A U_t = e^{itH/\hbar} A e^{-itH/\hbar}$, such that

$$i\hbar \frac{dA(t)}{dt} = [A(t), H] \quad (21.42)$$

is the Heisenberg equation of motion for the operator $A(t)$. For the special case of the operator $A \equiv H$ being the time-independent Hamiltonian of a system, the above equation of motion gives $\frac{dH}{dt} = 0$. When H acts on its (stationary) eigenstates $H|\psi\rangle = E|\psi\rangle$ (the Schrödinger equation!) and E corresponds to the energy eigenvalues, the above time-independence of the Hamiltonian H translates to the law of conservation of energy $\frac{dE}{dt} = 0$. Since the energy eigenspectrum determines all the properties of a quantum system, the invariance of the energy spectrum with time indicates a homogeneity of the natural laws in time, i.e., the same experiment performed at two different times on the same system will lead to the same result.

21.3.3 Rotational invariance

Consider an infinitesimal rotation $\delta\hat{\phi}$ by an angle $\delta\phi$ about an axis of rotation $\hat{\phi}$. As the change in the coordinate vector $\delta\vec{r}$ arising from such a rotation $\delta\hat{\phi}$ is given by $\mathcal{R}_{\delta\phi}\vec{r}$: $\delta\vec{r} = \delta\hat{\phi} \times \vec{r}$,

the change in the state $|\psi\rangle$ is given by

$$\begin{aligned}
|\psi(r - \delta\vec{r})\rangle &= \psi(r) - (\vec{\nabla}\psi(r)) \cdot \delta\vec{r} + \dots = \psi(r) - \delta\vec{\phi} \times \vec{r} \cdot (\vec{\nabla}\psi(r)) + \dots , \\
&= (1 - \delta\vec{\phi} \times \vec{r} \cdot \vec{\nabla} + \dots) |\psi(r)\rangle , \\
&= (1 - \frac{i\delta\vec{\phi}}{\hbar} \cdot (\vec{r} \times -i\hbar\vec{\nabla}) + \dots) |\psi(r)\rangle , \\
&= (1 - \frac{i\delta\vec{\phi}}{\hbar} \cdot \vec{L} + \dots) |\psi(r)\rangle , \quad \text{where } \vec{L} = \vec{r} \times -i\hbar\vec{\nabla} = \vec{r} \times \vec{p} \\
\implies \psi(\mathcal{R}_{\delta\phi}\vec{r}) &= U(\delta\vec{\phi}) |\psi(\vec{r})\rangle = e^{-i\delta\vec{\phi}\cdot\vec{L}/\hbar} |\psi(\vec{r})\rangle ,
\end{aligned} \tag{21.43}$$

where (as given above) $\mathcal{R}_{\delta\phi}\vec{r}$ corresponds to a rotation of the coordinate \vec{r} by an angle $\delta\phi$ about an axis of rotation $\hat{\phi}$. The unitary transformation operator for the infinitesimal rotation by $\delta\phi$, $U(\delta\vec{\phi}) = e^{-i\delta\vec{\phi}\cdot\vec{L}/\hbar}$ shows that the components of the angular momentum operator \vec{L} along $\hat{\phi}$ (arising from $\hat{\phi} \cdot \vec{L}$) act as the generators of the rotation. Clearly, as different components of \vec{L} do not commute with one another (recall what we have learnt earlier!), $[L_i, L_j] = i\hbar\epsilon_{ijk}L_k$, we can see that $U_x(\delta\phi_x)U_y(\delta\phi_y) \neq U_y(\delta\phi_y)U_x(\delta\phi_x)$. This means that the group of rotations in three dimensions is non-Abelian, i.e., the order (or sequence) in which the rotations are carried out about different axes leads to different states.

Now, for the case where the angular momentum commutes with the Hamiltonian, $[\vec{L}, H] = 0$, such that the eigenvalues of L^2 and L_z are good quantum numbers that label the common eigenstates (along with the energy eigenvalues E), we see that

$$[\vec{L}, H] = 0 = \frac{d\vec{L}}{dt} \implies [U(\delta\vec{\phi}), H] = 0 , \tag{21.44}$$

such that the conservation of angular momentum reflects that the Hamiltonian is rotation symmetric, i.e., it is invariant under a rotation by $\delta\vec{\phi}$

$$H(\vec{\phi} + \delta\vec{\phi}) = U^\dagger(\delta\vec{\phi})H(\vec{\phi})U(\delta\vec{\phi}) = H(\vec{\phi}) . \tag{21.45}$$

This indicates that the potential V in the Hamiltonian must be a central force potential, i.e., V must be a function of the radial coordinates alone $V \equiv V(r)$ (recall the Hydrogen atom for an example!). Further, this means that certain key properties of the system, e.g., the energy eigenvalues, cannot be dependent on the angular coordinates in any way.

21.4 Discrete Symmetries: Parity, Time-reversal and all that

21.4.1 Parity

We have seen above that continuous spatial translation, time translation and rotational symmetries (or invariance of the system with respect to linear displacements (in space and time) and rotations) obtains the conservation of linear momentum, energy and angular momentum respectively. Similarly, we will now observe that an invariance under space inversion,

i.e., a simultaneous change in the sign of all the spatial coordinates of the constituents of a system, obtains the parity symmetry. Note, however, that while translation and rotational symmetries were continuous in nature, parity is discrete. In quantum mechanics, the action of the parity operator \mathcal{P} is defined by

$$\mathcal{P} |\psi(\vec{r})\rangle = |\psi(-\vec{r})\rangle , \quad (21.46)$$

and corresponds to an inversion of the system about the origin. The parity operation corresponds to a symmetry operation if

$$i\hbar \frac{d\mathcal{P}}{dt} = [\mathcal{P}, H] = 0 , \quad (21.47)$$

such that the eigenvalues of \mathcal{P} are conserved in time. But what are these eigenvalues (say, η_P)? For this, we reason that

$$\begin{aligned} \mathcal{P} |\psi(\vec{r})\rangle &= |\psi(-\vec{r})\rangle = \eta_P |\psi(\vec{r})\rangle , \\ \mathcal{P}^2 |\psi(\vec{r})\rangle &= \mathcal{P} |\psi(-\vec{r})\rangle = \eta_P^2 |\psi(\vec{r})\rangle = |\psi(\vec{r})\rangle , \\ \implies \eta_P^2 &= 1 \implies \eta_P = \pm 1 , \end{aligned} \quad (21.48)$$

where we have used the fact that $\mathcal{P}^2 |\psi(\vec{r})\rangle \equiv |\psi(\vec{r})\rangle$ (i.e., the action of two consecutive parity operations must be trivial). Eigenstates with $\eta_P = 1$ are said to be even (or, invariant) under parity, and those with $\eta_P = -1$ are said to be odd (as they change by a sign of -1). Further, since the eigenvalues $\eta_P = \pm 1$, it is easily seen that the parity operator is somewhat special in the sense that

$$\mathcal{P}^\dagger = \mathcal{P}^{-1} = \mathcal{P} . \quad (21.49)$$

The vector operators $\hat{\vec{r}}$, $\hat{\vec{p}}$ and $\hat{\vec{L}}$ transform under inversion as follows:

$$\mathcal{P}^\dagger \hat{\vec{r}} \mathcal{P} = -\hat{\vec{r}} , \quad \mathcal{P}^\dagger \hat{\vec{p}} \mathcal{P} = -\hat{\vec{p}} , \quad \mathcal{P}^\dagger \hat{\vec{L}} \mathcal{P} = \mathcal{P}^\dagger (\hat{\vec{r}} \times \hat{\vec{p}}) \mathcal{P} = \hat{\vec{L}} . \quad (21.50)$$

In order to see how the first of these relations works out, we note that

$$\begin{aligned} \hat{\vec{r}} |\psi(\vec{r})\rangle &= \vec{r} |\psi(\vec{r})\rangle , \\ \hat{\vec{r}} \mathcal{P} |\psi(\vec{r})\rangle &= \hat{\vec{r}} |\psi(-\vec{r})\rangle = -\vec{r} |\psi(-\vec{r})\rangle , \\ \implies \mathcal{P}^\dagger \hat{\vec{r}} \mathcal{P} |\psi(\vec{r})\rangle &= -\vec{r} \mathcal{P}^\dagger |\psi(-\vec{r})\rangle , \\ &= -\vec{r} \mathcal{P} |\psi(-\vec{r})\rangle , \quad (\text{as } P^\dagger = P) \\ \implies \mathcal{P}^\dagger \hat{\vec{r}} \mathcal{P} |\psi(\vec{r})\rangle &= -\vec{r} |\psi(\vec{r})\rangle = -\hat{\vec{r}} |\psi(\vec{r})\rangle , \end{aligned} \quad (21.51)$$

leading to

$$\mathcal{P}^\dagger \hat{\vec{r}} \mathcal{P} = -\hat{\vec{r}} . \quad (21.52)$$

In precisely the same way, the transformation of $\hat{\vec{p}}$ and $\hat{\vec{L}}$ under parity \mathcal{P} can be obtained. Further, the spin angular momentum (\vec{S}) and total angular momentum $\vec{J} = \vec{L} + \vec{S}$ are also even under parity: $\mathcal{P}^\dagger \vec{S} \mathcal{P} = \vec{S}$, $\mathcal{P}^\dagger \vec{J} \mathcal{P} = \vec{J}$.

21.4.2 Time Reversal

The time-reversal transformation (\mathcal{T}) transforms $t \rightarrow -t$; more precisely, we mean $\mathcal{T}t\mathcal{T}^\dagger = -t$. Thus, it reverses the velocity (and hence the momentum) and the angular momentum, but does not affect the position:

$$\mathcal{T}\vec{r}\mathcal{T}^\dagger = \vec{r} \quad , \quad \mathcal{T}\vec{p}\mathcal{T}^\dagger = -\vec{p} \quad , \quad \mathcal{T}\vec{L}\mathcal{T}^\dagger = \mathcal{T}(\vec{r} \times \vec{p})\mathcal{T}^\dagger = -\vec{L} \quad . \quad (21.53)$$

Perhaps it is more appropriate to call this the “reversal of motion” transformation, akin to the case of “rewinding the movie”. Now, for a system that is time-reversal invariant,

$$[\mathcal{T}, H] = 0 \implies H = \mathcal{T}H\mathcal{T}^\dagger \quad . \quad (21.54)$$

We can then use the fact that

$$\mathcal{T}t\mathcal{T}^\dagger = -t \implies \mathcal{T}(i\hbar \frac{\partial}{\partial t})\mathcal{T}^\dagger = -(i\hbar \frac{\partial}{\partial t}) \quad , \quad (21.55)$$

to rewrite the Schrödinger equation of motion

$$\begin{aligned} H|\psi\rangle &= i\hbar \frac{\partial |\psi\rangle}{\partial t} \quad , \\ \implies \mathcal{T}H|\psi\rangle &= \mathcal{T}i\hbar \frac{\partial |\psi\rangle}{\partial t} \quad , \\ \implies \mathcal{T}H\mathcal{T}^\dagger\mathcal{T}|\psi\rangle &= \mathcal{T}(i\hbar \frac{\partial}{\partial t})\mathcal{T}^\dagger\mathcal{T}|\psi\rangle \quad , \quad (\text{as } \mathcal{T}^\dagger\mathcal{T} = 1) \\ \text{or, } H(\mathcal{T}|\psi\rangle) &= -i\hbar \frac{\partial (\mathcal{T}|\psi\rangle)}{\partial t} \quad . \end{aligned} \quad (21.56)$$

Now, if we take the complex conjugate of the Schrödinger equation, we obtain

$$\begin{aligned} (H|\psi\rangle)^* &= (i\hbar \frac{\partial |\psi\rangle}{\partial t})^* \quad , \\ \implies H^*|\psi^*\rangle &= H|\psi^*\rangle = -i\hbar \frac{\partial |\psi^*\rangle}{\partial t} \quad . \end{aligned} \quad (21.57)$$

Comparing the Schrödinger equation for $|\psi^*\rangle$ with that obtained for $\mathcal{T}|\psi\rangle$ above yields

$$\mathcal{T}|\psi\rangle = |\psi^*\rangle \quad , \quad (21.58)$$

i.e., the action of the time-reversal transformation on the state $|\psi\rangle$ is equivalent to that of the complex conjugation of the state. It turns out that this is true only for the case of systems that do not have any half-integer spin degrees of freedom.

At this point, we note that the complex conjugation operation \mathcal{K} is an antilinear operator, i.e.,

$$\mathcal{K}(c_1|\psi_1\rangle + c_2|\psi_2\rangle) = c_1^*\mathcal{K}|\psi_1\rangle + c_2^*\mathcal{K}|\psi_2\rangle \quad , \quad (21.59)$$

for any two states $|\psi_1\rangle$ and $|\psi_2\rangle$ and complex numbers c_1 and c_2 . Also, the product of two antilinear operators \mathcal{K}_1 and \mathcal{K}_2 , such that $(\mathcal{K}_1\mathcal{K}_2)|\psi\rangle = \mathcal{K}_1(\mathcal{K}_2|\psi\rangle)$, obtains a linear operator as the complex conjugation of the two antilinear operators negate one another.

Further, the complex conjugation operation \mathcal{K} is an antiunitary operator. By this we mean that \mathcal{K} is antiunitary if (i) \mathcal{K} is antilinear, (ii) its inverse (\mathcal{K}^{-1}) exists with $\mathcal{K}^{-1} = \mathcal{K}$, and the norm of $\mathcal{K}|\psi\rangle$ (i.e., $\|\mathcal{K}|\psi\rangle\|$) is equal to that of $|\psi\rangle$ for all $|\psi\rangle$, i.e.,

$$\|\mathcal{K}|\psi\rangle\| = \||\psi\rangle\| \quad \forall |\psi\rangle . \quad (21.60)$$

This definition of an antiunitary operator implies that if

$$\begin{aligned} \mathcal{K}|\psi_1\rangle &= |\phi_1\rangle , \quad \mathcal{K}|\psi_2\rangle = |\phi_2\rangle , \\ \implies \langle\phi_1|\phi_2\rangle &= \langle\psi_1|\psi_2\rangle^* . \end{aligned} \quad (21.61)$$

Indeed, the time-reversal transformation operator \mathcal{T} is also an antiunitary operator. All antiunitary operators \mathcal{A} can be written as a product of a unitary operator (U) and the complex conjugation operator \mathcal{K} . Thus, we can also write $\mathcal{T} = U\mathcal{K}$, where \mathcal{K} is the complex conjugation transformation and U is a unitary operator. As we shall see shortly, $U = 1$ for the case of spin zero particles, while U is a $(2S + 1) \times (2S + 1)$ matrix for particles with non-zero spin S .

For a time-reversal invariant system ($[\mathcal{T}, H] = 0$), any stationary state $|\psi\rangle$ can have a time-reversed state $\mathcal{T}|\psi\rangle$ that is degenerate with it

$$\begin{aligned} H|\psi\rangle &= E|\psi\rangle , \\ \implies H(\mathcal{T}|\psi\rangle) &= \mathcal{T}(H|\psi\rangle) = E(\mathcal{T}|\psi\rangle) . \end{aligned} \quad (21.62)$$

This leaves us with two possibilities.

1. $|\psi\rangle$ and $\mathcal{T}|\psi\rangle$ are proportional to one another and describe the same state, or
2. $|\psi\rangle$ and $\mathcal{T}|\psi\rangle$ are linearly independent and describe degenerate states.

For the first case, i.e., of a nondegenerate state, we can take $\mathcal{T}|\psi\rangle = \tau|\psi\rangle$, where τ is the eigenvalue of \mathcal{T} . Then, we have

$$\mathcal{T}^2|\psi\rangle = \mathcal{T}(\mathcal{T}|\psi\rangle) = \mathcal{T}(\tau|\psi\rangle) = \tau^*(\mathcal{T}|\psi\rangle) = \tau^*\tau|\psi\rangle = |\tau|^2|\psi\rangle . \quad (21.63)$$

Further, reversing a nondegenerate state twice should leave it unchanged

$$\mathcal{T}^2|\psi\rangle = |\psi\rangle \implies \mathcal{T}^2 = 1 \implies \tau = e^{i\theta} , \quad (21.64)$$

where θ is a constant phase. Further, a nondegenerate energy eigenstate is real, such that

$$\mathcal{T}|\psi(E)\rangle = |\psi^*(E)\rangle = e^{i\theta}|\psi(E)\rangle , \quad (21.65)$$

where θ is an arbitrary constant phase factor. We will see below that this corresponds to the case of a system with integer spin, while the second case (i.e., of degenerate states) corresponds to that for a system with half-integer spin. Indeed, we will see below that for the second case, $\mathcal{T}^2|\psi\rangle = -|\psi\rangle$.

Time-reversal properties of particles with non-zero spin

Recall that

$$\mathcal{T}\vec{S}\mathcal{T}^\dagger = -\vec{S}. \quad (21.66)$$

Further, we recall that $\mathcal{T} = U\mathcal{K}$, and that

$$\mathcal{K}S_x\mathcal{K}^{-1} = S_x, \quad \mathcal{K}S_y\mathcal{K}^{-1} = -S_y, \quad \mathcal{K}S_z\mathcal{K}^{-1} = S_z, \quad (21.67)$$

as S_x and S_z are real matrices while S_y is purely imaginary (recall the Pauli matrices for $S = 1/2!$). Thus, in order to satisfy the requirement of eq.(21.66), the operation U must satisfy

$$US_xU^{-1} = -S_x, \quad US_yU^{-1} = S_y, \quad US_zU^{-1} = -S_z. \quad (21.68)$$

Indeed, it is not hard to see from eq.(21.68) that U corresponds to a rotation by 180° about the y -axis:

$$U = e^{-i\pi S_y/\hbar}, \quad (21.69)$$

such that S_y and hence U are matrices of size $(2S+1) \times (2S+1)$. Further, U can affect only the spin degrees of freedom as the correct transformation properties of \vec{r} and \vec{p} are already produced by the complex conjugation operation \mathcal{K} . Thus, the time-reversal transformation is given by

$$\mathcal{T} = U\mathcal{K} = e^{-i\pi S_y/\hbar}\mathcal{K}. \quad (21.70)$$

For the case of $S = 1/2$, using the fact that

$$U_{\hat{n}}(\theta) = e^{-i\theta\hat{n}\cdot\vec{S}/\hbar} = e^{-i\frac{\theta}{2}\hat{n}\cdot\sigma} = \cos(\theta/2)\mathbf{1} - i\sin(\theta/2)\hat{n}\cdot\sigma, \quad (21.71)$$

we find that

$$\begin{aligned} U_{\hat{n}_y}(\pi) &= \cos(\pi/2)\mathbf{1} - i\sin(\pi/2)\sigma_y = -i\sigma_y, \\ -i\sigma_y|\uparrow\rangle &= +|\downarrow\rangle, \quad -i\sigma_y|\downarrow\rangle = -|\uparrow\rangle, \end{aligned} \quad (21.72)$$

and the time-reversal transformation operation for a spin-1/2 particle is given by

$$\begin{aligned} \mathcal{T} &= -i\sigma_y\mathcal{K}, \\ \mathcal{T}^2|\psi_n\rangle &= -|\psi_n\rangle \implies \mathcal{T}^2 = -1, \end{aligned} \quad (21.73)$$

where we have used the fact that $\sigma_y^2 = 1 = \mathcal{K}^2$. This shows clearly that the states $|\psi_n\rangle$ and $\mathcal{T}|\psi_n\rangle$ are distinct and degenerate. This is unlike what we saw for spin-zero particles above: this degeneracy is referred to as the Kramers degeneracy. Very generally, any system with an odd number of electrons must have two-fold degeneracy in all the eigenstates in its energy spectrum.

Any such spin-1/2 system will have its degeneracy lifted when placed in an external magnetic field (due to the Zeeman splitting); the time reversal symmetry is explicitly broken. This is simply because the Zeeman term $H_{Zeeman} = -\vec{\mu}\cdot\vec{B}$ changes sign under \mathcal{T} as the magnetic moment $\vec{\mu} \propto \vec{S}$

$$\mathcal{T}H_{Zeeman}\mathcal{T}^\dagger = H_{Zeeman}. \quad (21.74)$$

It is worth noting that in H_{Zeeman} , reversing the direction of $\vec{\mu}$ (thereby changing its sign) has the same effect as reversing the direction of \vec{B} but keeping the direction of $\vec{\mu}$ unchanged. In this sense, we can see that

$$\mathcal{T}H_{Zeeman}(\vec{B})\mathcal{T}^\dagger = H_{Zeeman}(-\vec{B}) . \quad (21.75)$$

Bear in mind, though, that \mathcal{T} acts on $\vec{\mu}$ and not on \vec{B} .

Generalising these results to the case of an arbitrary spin S system, we note that

$$\mathcal{T} = U\mathcal{K} = e^{-i\pi S_y/\hbar}\mathcal{K} \implies \mathcal{T}^2 \propto e^{-i2\pi S_y/\hbar} . \quad (21.76)$$

Now, since

$$e^{-i2\pi S_y/\hbar} |s, m_s\rangle = (e^{-i\pi S_y/\hbar})^2 |s, m_s\rangle = (-1)^{2s} |s, m_s\rangle , \quad (21.77)$$

we obtain the eigenvalue of \mathcal{T}^2 is given by $(-1)^{2s}$, such that

$$\mathcal{T}^2 |j \in 1/2 - \text{integer}\rangle = -|j \in 1/2 - \text{integer}\rangle , \quad (21.78)$$

$$\mathcal{T}^2 |j \in \text{integer}\rangle = |j \in \text{integer}\rangle . \quad (21.79)$$

Indeed, this value of j could correspond to either an atom or a many-body system with an odd number of electrons. Finally, note that our conclusions on time reversal invariance of dynamics are very generally backed up by experiments on small systems at very low temperatures that exhibit quantum mechanical properties. The outstanding question then is: why do macroscopic systems not show this behaviour? How does the arrow of time arise for classical systems and what does this have to do with the notion of temperature (which quantum mechanics does not admit!) and the second law of thermodynamics? These are active research questions that are being investigated at present. With these questions, we conclude our discussion on time reversal invariance.

21.4.3 Discrete translation invariance and Bloch's theorem

Consider a one-dimensional system that possesses a periodic potential $V(x \pm a) = V(x)$ where a is the period of the potential. An example involves considering the dynamics of an electron confined to live on a chain that contains regularly spaced positive ions. It is now clear that the Hamiltonian H is not invariant under continuous translation. Rather, given that the potential possesses the property of being the same when translated by a discrete amount a , $T(a)$, we have

$$T^\dagger(a)V(x)T(a) = V(x + a) = V(x) . \quad (21.80)$$

Given the fact that the kinetic energy term in H is invariant under any translation, we know that the entire Hamiltonian is invariant under translation by a

$$T^\dagger(a)HT(a) = H . \quad (21.81)$$

Recall that since the unitary translation operator, here $T(x = a)$, commutes with the Hamiltonian

$$[H, T(a)] = 0 , \quad (21.82)$$

we have simultaneous eigenstates for H and $T(a)$.

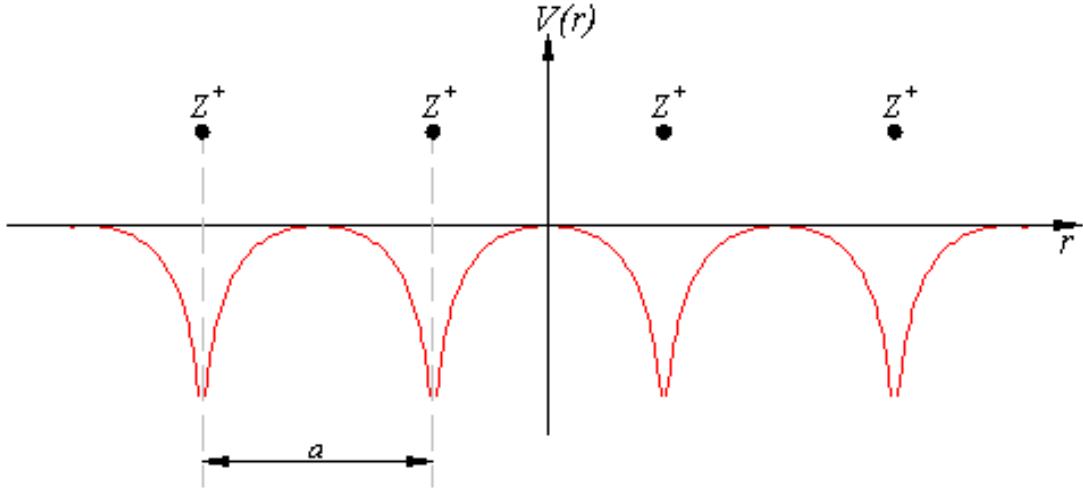


Figure 21.1: The periodic potential for a one-dimensional lattice. The deep wells correspond to the locations of the nuclear cores of the atoms from which the loosely bound outermost electrons form the delocalised states. a is the period of the lattice potential. Source: Wikipedia.

For a moment, let us consider the case of the atomic limit of this problem, i.e., the case when the amplitude of the periodic potential is taken to be infinitely large and each of the potential minima are completely cut off from one another (think of them as a periodic sequence of particle in a box problems!). We call these isolated islands the “sites of a one-dimensional lattice”. Then, H simplifies to solving for the discrete levels (denoted by the symbol i) of the purely local problem for, say the n site: the states $|\psi^i\rangle_n$ and energy eigenvalues E_n^i . Let us now assume that the particle (whose dynamics we wish to understand) is localised on the lowest eigenstate of, say, the n th site: $|\psi^0\rangle_n \equiv |n\rangle$ with energy E_0 : $H|n\rangle = E_0|n\rangle$.

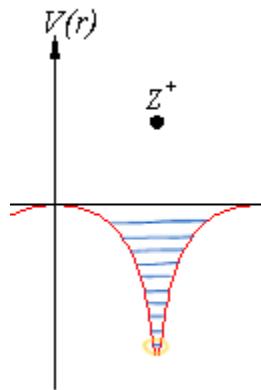


Figure 21.2: The bound state energy eigenstates (blue lines) for a single, infinitely deep, potential well. This corresponds to the physics of a single site of the 1D lattice in the atomic limit. The lowest energy bound state is encircled in yellow. Source: Wikipedia.

Clearly, there are a infinite number of sites for a one-dimensional lattice with an infinite number of site; on each site, we consider a similar localised state (labelled by the site number) $|n\rangle$ and an energy identical to E_0 . This indicates that the atomic limit has a huge degeneracy. Once we lower the potential amplitude from ∞ , the quantum particle will surely delocalise from the n site and explore the rest of the system; the infinite degeneracy of the lowest states will be lifted. Now, $|n\rangle$ is not an eigenstate of $T(a)$: $T(a)|n\rangle = |n+1\rangle$, $T(-a)|n\rangle = |n-1\rangle$. Thus, we need to find the simultaneous eigenstates of H and the discrete translation operator $T(a)$.

Given the infinite degeneracy of the local eigenstates $|n\rangle$, we choose a linear combination as a candidate eigenstate of H and $T(a)$

$$|\theta\rangle = \mathcal{N} \sum_{n=-\infty}^{\infty} e^{in\theta} |n\rangle , \quad (21.83)$$

where θ is a real parameter lying in the range $-\pi \leq \theta \leq \pi$ and \mathcal{N} is a normalisation factor. Acting with $T(a)$ on $|\theta\rangle$ gives

$$\begin{aligned} T(a)|\theta\rangle &= T(a)(\mathcal{N} \sum_{n=-\infty}^{\infty} e^{in\theta} |n\rangle) = \mathcal{N} \sum_{n=-\infty}^{\infty} e^{in\theta} T(a)|n\rangle , \\ &= \mathcal{N} \sum_{n=-\infty}^{\infty} e^{in\theta} |n+1\rangle = \mathcal{N} \sum_{n=-\infty}^{\infty} e^{i(n-1)\theta} |n\rangle , \\ &= e^{-i\theta} |\theta\rangle , \end{aligned} \quad (21.84)$$

$$\begin{aligned} T(-a)|\theta\rangle &= \mathcal{N} \sum_{n=-\infty}^{\infty} e^{in\theta} |n-1\rangle = \mathcal{N} \sum_{n=-\infty}^{\infty} e^{i(n+1)\theta} |n\rangle , \\ &= e^{i\theta} |\theta\rangle . \end{aligned} \quad (21.85)$$

In order to attain a solution to the problem, we assume that lowering the potential amplitude slightly from ∞ allows for the particle to delocalise between just the nearest neighbour minima (this is called the “tight-binding (or Wannier) approximation”)

$$\langle n \pm 1 | H | n \rangle = -t , \quad (21.86)$$

where the so-called hopping amplitude parameter t is taken to be independent of n due to the discrete translation invariance of the problem. The parameter t arises from the physics of quantum tunneling through barriers, and reflects a finite kinetic energy for the matter waves. In the tight-binding regime, we assume that the typical kinetic energy of the particle ($\sim \mathcal{O}(t)$) is much smaller than the on-site potential energy for the particle to be localised within the n -th potential well ($\langle n | H | n \rangle = E_0$). Thus, we obtain

$$H|n\rangle = E_0|n\rangle - t|n+1\rangle - t|n-1\rangle , \quad (21.87)$$

$$\implies H = E_0|n\rangle\langle n| - t|n+1\rangle\langle n| - t|n-1\rangle\langle n| , \quad (21.88)$$

clearly indicating that the $|n\rangle$ states are no longer eigenstates of the Hamiltonian H .

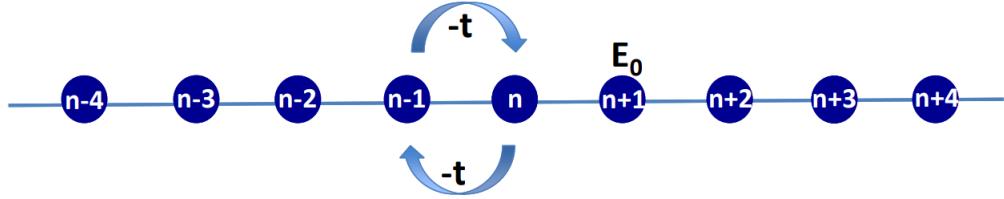


Figure 21.3: The one-dimensional tight-binding lattice with hopping amplitude $-t$ between nearest neighbour lattice sites and on-site energy E_0 .

Instead, let us try the state $|\theta\rangle = \sum_n e^{in\theta} |n\rangle$ we considered above

$$H|\theta\rangle = H\mathcal{N} \sum_{n=-\infty}^{\infty} e^{in\theta} |n\rangle , \quad (21.89)$$

$$\begin{aligned} &= E_0\mathcal{N} \sum_n e^{in\theta} |n\rangle - t\mathcal{N} \sum_n e^{in\theta} |n-1\rangle - t\mathcal{N} \sum_n e^{in\theta} |n+1\rangle , \\ &= E_0\mathcal{N} \sum_n e^{in\theta} |n\rangle - t(\mathcal{N} \sum_n e^{i(n+1)\theta} |n\rangle + \mathcal{N} \sum_n e^{i(n-1)\theta} |n\rangle) , \\ &= E_0\mathcal{N} \sum_n e^{in\theta} |n\rangle - t(e^{i\theta} + e^{-i\theta})\mathcal{N} \sum_n e^{in\theta} |n\rangle , \end{aligned} \quad (21.90)$$

$$\begin{aligned} &= (E_0 - 2t \cos \theta)\mathcal{N} \sum_n e^{in\theta} |n\rangle , \\ &= (E_0 - 2t \cos \theta)|\theta\rangle \quad (\text{for } -\pi \leq \theta \leq \pi). \end{aligned} \quad (21.91)$$

This reveals that $|\theta\rangle$ is indeed an eigenstate of H with energy eigenvalues now a function of the real parameter θ , $E(\theta) = E_0 - 2t \cos \theta$ where $-\pi \leq \theta \leq \pi$. It is important to note that the infinitely degenerate ground state manifold of the lowest energy bound states in the isolated wells has now been replaced by a “band” of states labelled by the continuously varying parameter θ lying in the range $-\pi \leq \theta \leq \pi$, i.e., a continuous spectrum of non-degenerate energy eigenstates lying within the energy range (or “bandwidth”) $E_0 - 2t \leq E(\theta) \leq E_0 + 2t$.

An important subtlety worth mentioning is that we have thus far been considering the case of an infinitely long one-dimensional lattice: $-\infty \leq n \leq \infty$; in this case, shifting the index n to $n \pm 1$ had no effect on the summations. For the case of a chain of finite size, the same results are obtained only if we impose periodic boundary conditions, i.e., we picture the system as being laid out on a circle such that quantum tunneling links the well with largest number index to the very first one. In that case, however, we will obtain a finite number of values of θ and energy eigenstates within each band (corresponding to the number of wells in the chain), i.e., both the parameter θ and the energy spectrum $E(\theta)$ form a discrete set in this case. Further, a study of the finite chain reveals the importance of boundary conditions, and this is something we will deal with in a later chapter.

In order to learn the physical significance of θ , we consider the real-space wavefunction

$\psi_\theta(x) = \langle x|\theta\rangle$ for the state $T(a)|\theta\rangle$

$$\langle x|T(a)|\theta\rangle = \langle x-a|\theta\rangle \text{ for } T(a) \text{ acting on } \langle x| , \quad (21.92)$$

$$= e^{-i\theta}\langle x|\theta\rangle \text{ for } T(a) \text{ acting on } |\theta\rangle , \quad (21.93)$$

$$\implies \langle x-a|\theta\rangle = e^{-i\theta}\langle x|\theta\rangle , \quad (21.94)$$

$$\text{Similarly, } \langle x+a|\theta\rangle = e^{i\theta}\langle x|\theta\rangle . \quad (21.95)$$

The final two relations can be solved by setting

$$\psi_\theta(x) = \langle x|\theta\rangle = \mathcal{N}e^{ikx} u_k(x) \equiv \psi_k(x) , \quad (21.96)$$

where we have set the dimensionless parameter $\theta = ka$, a is the lattice spacing, k is a momentum and $u_k(x)$ is a periodic function in x with period a and \mathcal{N} is a normalisation constant. The periodic nature of $u_k(x)$ can be verified explicitly by putting eq.(21.96) in eq.(21.94):

$$\begin{aligned} e^{ik(x-a)} u_k(x-a) &= e^{-ika} e^{ikx} u_k(x) , \\ \implies u_k(x-a) &= u_k(x) . \end{aligned} \quad (21.97)$$

Equation (21.96) is referred to as the **Floquet-Bloch theorem** (because Floquet found it first almost 40 years before Bloch in a different context!): the wavefunction $\psi_k(x) = \langle x|k\rangle = e^{ikx} u_k(x)$ is a simultaneous eigenstate of H and T , and can be written as a product of a plane wave phase factor e^{ikx} (where k is the propagation wavevector/momenta of the extended eigenstate) and a periodic function $u_k(x)$ with periodicity a . It is important to note that all we needed to reach this conclusion was the fact that $|k\rangle$ is an eigenstate of $T(\pm a)$ with eigenvalue $e^{\mp ika}$; indeed, our conclusions will continue to hold even beyond the validity of the tight-binding approximation we made earlier, e.g., the theorem continues to hold even in the opposite regime, i.e., for the eigenstates of an electron placed in a periodic potential whose amplitude is very small compared to its typical kinetic energy.

We can now understand the energy eigenvalue we obtained earlier: $E(\theta) = E_0 - 2t \cos \theta$. Given that $\theta = ka$, we see that as θ varies in the range $\pi \leq \theta \leq \pi$, the propagation wavevector k varies in the range $-\pi/a \leq k \leq \pi/a$ (the so-called “Brillouin zone”). Then, we can write the energy eigenvalue E in terms of the propagation wavevector k , $E \equiv E(k)$ (the so-called “dispersion relation”)

$$E(k) = E_0 - 2t \cos(ka) . \quad (21.98)$$

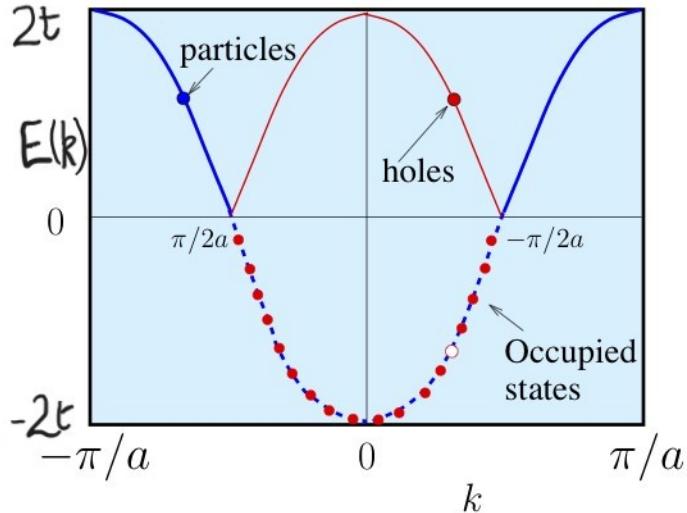


Figure 21.4: Energy-momentum (dispersion) relation for the periodic potential with $-\frac{\pi}{a} \leq k \leq \frac{\pi}{a}$ for the lattice spacing a and on-site energy $E_0 = 0$. The lowest (blue) dispersion curve represents the band of delocalised states corresponding to the tight-binding model for only the lowest energy level in any isolated well, while the next (green) dispersion curve represents the band of delocalised states corresponding to the first excited energy level in any isolated well. Source: the internet.

In this way, we see that the delocalisation of the particle between the local eigenstates $|n\rangle$ in order to obtain the true energy eigenstate $|k\rangle$ leads to a complete lifting of the infinite degeneracy we encountered in the atomic limit. Indeed, we now have a “band” of eigenstates for the propagating particle with energies lying in the range $E_0 - 2t \leq E(k) \leq E_0 + 2t$. Note that while the macroscopic degeneracy of the eigenstates of the isolated wells has been lifted in creating the eigenstates of the band, these band states still possess a two-fold degeneracy arising from time-reversal symmetry $k \leftrightarrow -k$. Further, for the case of electrons with spin occupying these states, these eigenstates are two-fold degenerate due to the spin degeneracy arising from the Pauli exclusion principle.

This is the so-called “band theory” for electrons propagating in the periodic lattice of a crystal, and forms the cornerstone of our understanding of the quantum dynamics of metals, insulator and semiconductors. In any material, the electronic states can exist in any one of several bands that comprise the energy eigenspectrum (see Fig. 21.5). There is, of course, much more to learn here (e.g., the periodic potential applied to a nearly free electron gas, the physics that gives rise to the gaps between various energy bands etc.) and I hope you will do so in a later course on basic condensed matter physics.

I close this section by raising a question for the reader to think about. Extending our attention to systems in three dimensions, Bloch’s theorem states that the eigenstate ψ for an electron in a periodic potential, i.e., $V(\vec{r} + \vec{R}) = V(\vec{r})$ for all \vec{R} being the lattice translation vectors in a so-called Bravais lattice, and satisfying the Schrödinger equation with the Hamiltonian $H = -\frac{\hbar^2}{2m} \nabla^2 + V(\vec{r})$ is given by

$$\psi_{n,\vec{k}}(\vec{r}) = e^{i\vec{k}\cdot\vec{r}} u_{n,\vec{k}}(\vec{r}), \quad (21.99)$$

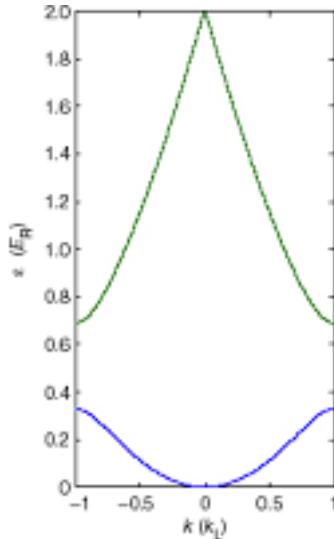


Figure 21.5: Energy-momentum (dispersion) relation for the periodic potential with $-1 \leq k/\pi \leq 1$, $a = 1$, $t = 0.075$, $E_0 = 0.15$. The lowest (blue) dispersion curve represents the band of delocalised states corresponding to the tight-binding model for only the lowest energy level in any isolated well, while the next (green) dispersion curve represents the band of delocalised states corresponding to the first excited energy level in any isolated well. Source: the internet.

where $u_{n,\vec{k}}(\vec{r})$ has the same periodicity as the underlying potential of the system and n is the “band” index, i.e., specifies which band the eigenstate belongs to. To check whether such a wavefunction is an eigenstate of the momentum operator $\vec{p} \equiv -i\hbar\nabla$, we simply act with the momentum operator on it

$$-i\hbar\nabla\psi_{n,\vec{k}}(\vec{r}) = \hbar\vec{k}e^{i\vec{k}\cdot\vec{r}}u_{n,\vec{k}}(\vec{r}) - i\hbar e^{i\vec{k}\cdot\vec{r}}\nabla u_{n,\vec{k}}(\vec{r}) = \hbar\vec{k}\psi_{n,\vec{k}}(\vec{r}) - i\hbar e^{i\vec{k}\cdot\vec{r}}\nabla u_{n,\vec{k}}(\vec{r}) . \quad (21.100)$$

The second term, i.e., $-i\hbar e^{i\vec{k}\cdot\vec{r}}\nabla u_{n,\vec{k}}(\vec{r})$, has no reason a priori to be zero. Thus, we can clearly see that the Bloch wavefunction is not an eigenstate of the momentum operator $\vec{p} \equiv -i\hbar\nabla$, and that \vec{k} cannot be related to the momentum of the electron in the usual sense. The question then is: what kind of momentum does \vec{k} represent? We will answer this question in later chapters.

21.4.4 Exchange Symmetry

We briefly touch here upon the subject of exchange symmetry, which you will hopefully learn in more depth when you study the quantum mechanics of systems comprised of identical quantum particles. For now, we shall simply state a few results. It turns out that the eigenstates of a system of bosons (or integer spin particles in general) is invariant (i.e., unchanged) under the exchange of identical particles: this requires the symmetrization of the wavefunctions:

$$\Psi_{many-boson}(x_1, \dots, x_i, \dots, x_j, \dots, x_N) = \Psi_{many-boson}(x_1, \dots, x_j, \dots, x_i, \dots, x_N) . \quad (21.101)$$

On the other hand, eigenstates of a system of fermions (or half-integer spin particles in general) acquire negative signs under the exchange of any two identical fermions: this requires the antisymmetrization of the wavefunctions. The latter is an outcome of the Pauli exclusion principle (formulated in 1925) which states that only one electron can occupy a given state (characterised by a set of good quantum numbers). This leads to the fact that any many-fermion wavefunction has the form of a determinant (the so-called Slater determinant), with the following property:

$$\Psi_{many-fermion}(x_1, \dots, x_i, \dots, x_j, \dots, x_N) = -\Psi_{many-fermion}(x_1, \dots, x_j, \dots, x_i, \dots, x_N) . \quad (21.102)$$

This symmetrization of the many-boson wavefunction, and the antisymmetrization of the many-fermion wavefunction, is implemented by taking careful account of the operators responsible for permutation symmetry (antisymmetry) under exchange of identical particles. I present here a brief exposition of the many-fermion antisymmetrised wavefunction. For a system of N non-interacting fermions, the Hamiltonian (H) is given by the sum of N independent Hamiltonians (i.e., one for each of the fermions, H_i^0): $H = \sum_{i=1}^N H_i^0$. Given that the Hamiltonian is separable, the N -fermion wavefunction is a linear combination of N wavefunctions of the single particle Hamiltonian H^0 that has been suitably antisymmetrised in the shape of the Slater determinant:

$$\Psi(\vec{r}_1\sigma_1, \dots, \vec{r}_N\sigma_N) = \langle \vec{r}_1\sigma_1, \dots, \vec{r}_N\sigma_N | \psi_N \rangle = \frac{1}{\sqrt{N!}} \begin{vmatrix} \psi_{\vec{k}_1}(\vec{r}_1, \sigma_1) & \cdots & \psi_{\vec{k}_N}(\vec{r}_1, \sigma_1) \\ \vdots & \ddots & \vdots \\ \psi_{\vec{k}_1}(\vec{r}_N, \sigma_N) & \cdots & \psi_{\vec{k}_N}(\vec{r}_N, \sigma_N) \end{vmatrix} \quad (21.103)$$

where $\psi_{\vec{k}}(\vec{r}, \sigma) = \frac{1}{\sqrt{V}} e^{i\vec{k}\cdot\vec{r}} \otimes \xi_\sigma$ is a product state composed of a direct product between a plane-wave single-particle state and a spin eigenstate ξ_σ , and whose energy is given by $\epsilon_{\vec{k}}^{(0)} = \frac{\hbar^2 |\vec{k}|^2}{2m_e}$. As noted above, the determinant form of $\Psi(\vec{r}_1\sigma_1, \dots, \vec{r}_N\sigma_N)$ arises from the need for anti-symmetrisation under exchange (or permutation) of any two electrons (i.e., columns of rows), as dictated by the Pauli exclusion principle.

It is also useful to consider the states that are occupied by the N particles among all possible states of H^0 (the former are typically far smaller than the latter)

$$|\psi_N\rangle = |\alpha, \beta, \dots, \mu\rangle = |\{N_\mu\}\rangle , \quad (21.104)$$

where $\{N_\mu\} \equiv \alpha, \beta, \dots, \mu$ are N states chosen among the eigenstates of H^0 . Adding another particle to this state will require antisymmetrisation over all coordinates, and involves specifying the phase of the $N + 1$ particle state relative to that of the N -particle wavefunction. Thus, we can obtain the $N + 1$ -particle state $|\psi_{N+1}\rangle$ by acting on the state $|\psi_N\rangle$ with the creation operator a_ν^\dagger

$$|\psi_{N+1}\rangle = (-1)^N a_\nu^\dagger |\psi_N\rangle = (-1)^N a_\nu^\dagger |\{N_\mu\}\rangle = (-1)^N |\alpha, \beta, \dots, \mu, \nu\rangle , \quad (21.105)$$

which describes $N + 1$ particles in the states $\alpha, \beta, \dots, \mu, \nu \equiv \{N\} + \nu$, and where the phase $(-1)^N$ arises from the antisymmetrisation of all the coordinates. Similarly, the N -particle

state can be obtained from the action of the destruction operator a_ν on the $N + 1$ -particle state as follows

$$|\psi_N\rangle = |\{N_\mu\}\rangle = (-1)^N a_\nu |\{N_\mu\}, \nu\rangle = (-1)^N a_\nu |\psi_{N+1}\rangle . \quad (21.106)$$

In this way, a completely antisymmetrised N -fermion state can be obtained from the vacuum ($|0\rangle$) by the action of N creation operators ($a_i^\dagger, i = 1, \dots, N$)

$$|\psi_N\rangle = (-1)^P \prod_{i \in \{N_\mu\}} a_i^\dagger |0\rangle , \quad (21.107)$$

where $(-1)^P$ is a phase factor that arises from the antisymmetrisation procedure (due to all possible pairwise permutations of the N particles). For example, the wavefunctions for one and two fermions are

$$\langle \vec{r}_1 | a_\alpha^\dagger | 0 \rangle = \langle \vec{r}_1 | \alpha \rangle = \psi_\alpha(\vec{r}_1) , \quad (21.108)$$

$$\langle \vec{r}_1, \vec{r}_2 | a_\beta^\dagger a_\alpha^\dagger | 0 \rangle = \langle \vec{r}_1, \vec{r}_2 | \beta, \alpha \rangle = \psi_\alpha(\vec{r}_2) \psi_\beta(\vec{r}_1) - \psi_\beta(\vec{r}_2) \psi_\alpha(\vec{r}_1) . \quad (21.109)$$

The N -fermion Slater determinant ground state is called the Fermi volume, and obtained by choosing the set of states $\{N_\nu\}$ as the N lowest energy states of H^0 (including spin degeneracy). The creation and destruction operators satisfy the following algebra

$$\{a_\mu, a_\nu^\dagger\} = a_\mu a_\nu^\dagger - a_\nu^\dagger a_\mu = \delta_{\mu\nu} , \quad (21.110)$$

$$\{a_\mu, a_\nu\} = 0 = \{a_\mu^\dagger, a_\nu^\dagger\} . \quad (21.111)$$

This is the algebraic manner by which the Pauli exclusion principle (exchange antisymmetry) is enforced. The number (often also called occupation number) operator is given by $n_\nu = a_\nu^\dagger a_\nu$ and

$$n_\nu |\{N_\mu\}\rangle = 0 \text{ if } \nu \notin \{N_\mu\} , \quad (21.112)$$

$$= 1 \text{ if } \nu \in \{N_\mu\} , \quad (21.113)$$

such that the N -fermion Hamiltonian H can be written in terms of the fermionic occupation number operator as follows

$$H = \sum_{\vec{k}, \sigma} \epsilon_{\vec{k}}^{(0)} n_{\vec{k}, \sigma} = \sum_{\vec{k}, \sigma} \frac{\hbar^2 |\vec{k}|^2}{2m_e} n_{\vec{k}, \sigma} . \quad (21.114)$$

The symmetrisation of the bosonic state can be obtained in precisely the same manner, but in terms of the bosonic creation (b_ν^\dagger) and destruction (b_ν) operators with the commutation relation $[b_\mu, b_\nu^\dagger] = \delta_{\mu\nu}$ and dropping all the $(-1)^N$ phase factors in the expressions given above.

Anyons. An important final point on the matter of changes in the N -particle wavefunction upon the exchange of any two particles: the factors of $+1$ and -1 for bosons and fermions respectively can always be thought of as phases of 2π and π that are accrued by the bosonic

and fermionic wavefunction under the exchange to two bosons and two fermions respectively. One may then ask: are these the only two particle exchange phases that can possibly exist? It turns out that there exists the special possibility for interacting many-particle systems in two dimensions that one can find special ground states where exchanging any two of the particles can give phases other than π and 2π . Such particles were dubbed as “anyons” by Wilczek, and have been experimentally shown to exist in the ground states of fractional quantum Hall systems. Please note, however, that anyons are not fundamental particles such as electrons and photons; they are “emergent” in certain systems which are placed in special circumstances (e.g., extremely strong magnetic fields, very low temperatures etc.).

21.5 Total number conservation and global phase rotation symmetry of the many-particle wavefunction.

Note that the total number operator $\hat{N} = \sum_{\vec{k},\sigma} n_{\vec{k},\sigma}$ obtains the total particle number of N upon acting on the N -fermion wavefunction

$$\hat{N} |\psi_N\rangle = N |\psi_N\rangle . \quad (21.115)$$

If the N -fermion Hamiltonian H describes an isolated quantum system with a fixed number of constituents (i.e., the eigenvalue N is fixed), it is clear that

$$[\hat{N}, H] = 0 \implies \frac{d\hat{N}}{dt} = i\hbar [\hat{N}, H] = 0, \quad (21.116)$$

i.e., the total particle number N is conserved in time. As we have discussed above, all such conservation laws are related to a symmetry. Which symmetry relates to the invariance of N in time (for a closed quantum system)? Recall that the N -fermion wavefunction is a $N \times N$ Slater determinant: multiplying every element of this determinant by an identical arbitrary phase factor $e^{i\theta}$ leads a global arbitrary phase factor of $e^{iN\theta}$ that multiplies the entire N -fermion wavefunction (there are N elements in each term of the expanded determinant, and each element contributes a factor of $e^{i\theta}$). It is well known that such an arbitrary phase does not affect the expectation value of any observable, and thus has no physical consequences: it simply reflects a $U(1)$ global phase rotation symmetry of the N -fermion wavefunction. Given that such an arbitrary phase θ is known only modulo 2π , i.e., the transformation $\theta \rightarrow \theta + 2\pi$ should leave the N -fermion wavefunction invariant, we find that this leaves us with the fact that $e^{i2\pi N} = 1$, i.e., N must be an integer $N \in \mathbb{Z}$. Interestingly, this indicates that the total particle number N acts as a winding number (a topological quantity!) as far as the global phase of the N -particle wavefunction is concerned.

What transformation would lead to a global phase such as $e^{iN\theta}$? Consider translating the position of every particle in an N -particle system by an identical and arbitrary amount $\delta\vec{r}$: $\vec{r}_i \rightarrow \vec{r}_i + \delta\vec{r}$, $i = 1, \dots, N$. Clearly, this amounts to a global translation of the center of mass coordinate of the system, $\vec{r}_{CM} \rightarrow \vec{r}_{CM} + \delta\vec{r}$, and will involve the center of mass translation operator $T_{CM} = e^{-iN\delta\vec{r}\cdot\hat{\vec{p}}_{CM}/\hbar}$ (where the center of mass momentum operator is given by $\hat{\vec{p}}_{CM} = \frac{1}{N} \sum_i \hat{\vec{p}}_i$) on the state $|\psi(\vec{r})\rangle$

$$T_{CM} |\psi(\vec{r})\rangle = e^{-iN\delta\vec{r}\cdot\hat{\vec{p}}_{CM}/\hbar} |\psi(\vec{r})\rangle = |\psi(\vec{r} - \delta\vec{r})\rangle . \quad (21.117)$$

Clearly, the operator T_{CM} commutes with the Hamiltonian for a system of N non-interacting fermions which are not placed in any global confining potential, $T_{CM}^\dagger H T_{CM} = H$, as the center of mass momentum $\hat{\vec{p}}_{CM}$ operator commutes with the Hamiltonian for such a system

$$[\hat{\vec{p}}_{CM}, H] = 0 \implies \frac{d\hat{\vec{p}}_{CM}}{dt} = i\hbar [\hat{\vec{p}}_{CM}, H] = 0. \quad (21.118)$$

This is simply a statement of the continuous real-space translation symmetry of the system taken as a whole when there is no global confining potential. This obtains the arbitrary constant global phase θ as $\theta = \delta\vec{r}\langle\hat{\vec{p}}_{CM}\rangle/\hbar$, where $\langle\hat{\vec{p}}_{CM}\rangle$ is a real-valued constant. The upshot of is that for a state with a fixed number (N) of particles, there exists a large uncertainty in its (global) phase (of the order of 2π). This reflects the global $U(1)$ phase rotation symmetry of the system, and arises from the continuous real-space translation symmetry of the system taken as a whole. Coherent states, on the other hand, are states in which there exists uncertainties in both the total number of particles as well as the overall phase.

Importantly, this $U(1)$ global phase rotation symmetry indicates a good quantum number of all metallic states of matter, i.e., the total particle number N corresponding to the filled number of states within the Fermi volume at temperature $T = 0$. Further, this quantum number is found to be remain good even upon the adiabatic inclusion of inter-particle interactions as long as they do not lead to a quantum phase transition that affects the sharp Fermi surface that bounds the Fermi volume at $T = 0$. This invariance is known as Luttinger's theorem, and arises from the fact that the total particle number N (corresponding to the filled number of states within the Fermi volume at temperature $T = 0$) is a topological invariant of the N -particle wavefunction (as mentioned above). Further, the $U(1)$ global phase rotation symmetry is observed to be broken spontaneously for the case of superconducting states of matter emergent from metals that undergo a quantum phase transition. While a deeper discussion of superconductivity and how it is obtained lies well beyond the purview of this course, suffice it to say that the superconducting ground state is a phenomenon where the low-energy description is that of a condensate of “Cooper pair” bound states of pairs of electrons whose total number is not conserved in time. Naturally, the associated $U(1)$ global phase rotation symmetry no longer holds for this state of matter, and the superconducting ground state corresponds to a coherent state of Cooper pairs.

21.5.1 In the presence of a gauge field.

In the presence of an external gauge field given by a vector potential $\vec{A}(\vec{r})$, the total Hamiltonian is given by

$$H = \sum_{i=1}^N \frac{1}{2m} \left[\vec{p}_i - \frac{e}{c} \vec{A}(\vec{r}) \right]^2. \quad (21.119)$$

Given a spatially dependent vector potential $\vec{A}(\vec{r})$, the Hamiltonian no longer commutes with T_{CM} as

$$T_{CM}^\dagger \sum_{i=1}^N \frac{1}{2m} \left[\vec{p}_i - \frac{e}{c} \vec{A}(\vec{r}) \right]^2 T_{CM} = \sum_{i=1}^N \frac{1}{2m} \left[\vec{p}_i - \frac{e}{c} \vec{A}(\vec{r} - \delta\vec{r}) \right]^2 \neq H. \quad (21.120)$$

Now, we recall that the Peierls substitution $\hat{\vec{p}}_i \rightarrow \hat{\vec{p}}_i + \frac{e\vec{A}(\vec{r})}{c}$ gives the correct momentum that is gauge invariant for the case of a system with a magnetic field, i.e., the vector potential invariant under the gauge transformation of the vector potential $\vec{A}(\vec{r})$ by an arbitrary scalar function λ : $\vec{A}(\vec{r}) \rightarrow \vec{A}(\vec{r} + \delta\vec{r}) = \vec{A}(\vec{r}) + \vec{\nabla}\lambda_{\delta\vec{r}}(\vec{r})$ such that $\vec{\nabla} \times \vec{A}(\vec{r} + \delta\vec{r}) = \vec{\nabla} \times \vec{A}(\vec{r}) = \vec{B}(\vec{r})$. Thus, in order to keep maintain the continuous translation invariance of the Hamiltonian in the presence of such a magnetic field, the gauge invariant momentum $\hat{\vec{p}}_i + \frac{e\vec{A}}{c} + \vec{\nabla}\lambda_{\delta\vec{r}}(\vec{r})$, we must perform a gauge transformation given by $\delta T_{CM} = \exp\left(-iN\frac{e}{c\hbar} \int_0^r d\vec{r} \cdot \vec{\nabla}\lambda_{\delta\vec{r}}(\vec{r})\right) = \exp\left(iN\frac{e}{c\hbar}\lambda_{\delta\vec{r}}(\vec{r})\right)$ on the Hamiltonian

$$\delta T_{CM}^\dagger H(\hat{\vec{p}}_i + \frac{e}{c}(\vec{A}(\vec{r}) + \vec{\nabla}\lambda_{\delta\vec{r}}(\vec{r})))\delta T_{CM} = H(\hat{\vec{p}}_i + \frac{e\vec{A}(\vec{r})}{c}) , \quad (21.121)$$

$$\begin{aligned} &\Rightarrow \exp\left(iN\frac{e}{c\hbar}\lambda_{\delta\vec{r}}(\vec{r})\right)H(\hat{\vec{p}}_i + \frac{e}{c}(\vec{A}(\vec{r}) + \vec{\nabla}\lambda_{\delta\vec{r}}(\vec{r})))\exp\left(-iN\frac{e}{c\hbar}\lambda_{\delta\vec{r}}(\vec{r})\right) \\ &= H(\hat{\vec{p}}_i + \frac{e\vec{A}(\vec{r})}{c}) , \end{aligned} \quad (21.122)$$

together with

$$\delta T_{CM} |\psi(\vec{r})\rangle = e^{-iN\frac{e}{c\hbar}\lambda_{\delta\vec{r}}(\vec{r})} |\psi(\vec{r})\rangle = |\psi(\vec{r} - \delta\vec{r})\rangle . \quad (21.123)$$

This gauge transformation guarantees the continuous translation symmetry of the Hamiltonian $H = \sum_{i=1}^N \frac{1}{2m} \left[\vec{p}_i - \frac{e}{c} \vec{A}(\vec{r}) \right]^2$ under the translation operation $\tilde{T}_{CM} = \delta T_{CM} \times T_{CM} = e^{-iN\frac{e}{c\hbar}\lambda_{\delta\vec{r}}(\vec{r})} \times e^{-i\delta\vec{r}\hat{\vec{p}}_{CM}/\hbar}$

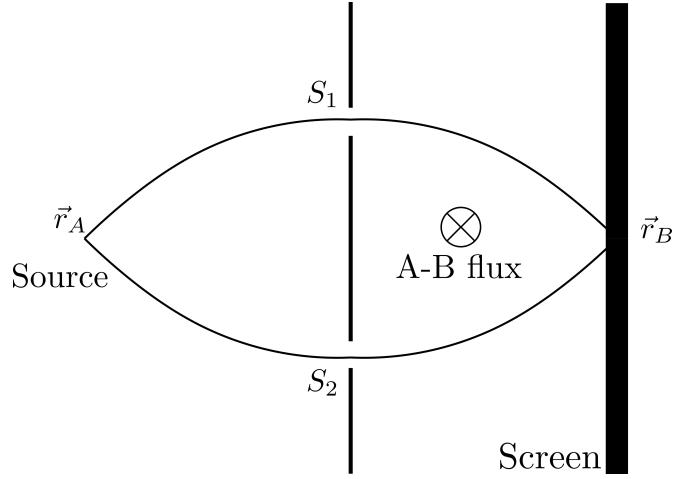
$$[\tilde{T}_{CM}, H] = 0 , \quad (21.124)$$

$$\Rightarrow \tilde{T}_{CM}^\dagger H \tilde{T}_{CM} = H , \quad (21.125)$$

$$\begin{aligned} &\Rightarrow e^{iN\frac{e}{c\hbar}\lambda_{\delta\vec{r}}(\vec{r})} \times e^{i\delta\vec{r}\hat{\vec{p}}_{CM}/\hbar} \sum_{i=1}^N \frac{1}{2m} \left[\vec{p}_i - \frac{e}{c} \vec{A}(\vec{r}) \right]^2 e^{-iN\frac{e}{c\hbar}\lambda_{\delta\vec{r}}(\vec{r})} \times e^{-i\delta\vec{r}\hat{\vec{p}}_{CM}/\hbar} \\ &= \sum_{i=1}^N \frac{1}{2m} \left[\vec{p}_i - \frac{e}{c} \vec{A}(\vec{r}) \right]^2 . \end{aligned} \quad (21.126)$$

In this way, we obtain from eq.(21.123) an arbitrary spatially dependent phase $\phi_{\delta\vec{r}}(\vec{r})$ considered above to be that arising from the invariance of the $\hat{\vec{p}}_{CM}$ under a gauge transformation, $\phi_{\delta\vec{r}}(\vec{r}) = \frac{e}{c\hbar}\lambda_{\delta\vec{r}}(\vec{r}) = 2\pi\frac{\lambda_{\delta\vec{r}}(\vec{r})}{\Phi_0}$ where $\Phi_0 = hc/e$ is the flux quantum. This phase corresponds to a spatially dependent variation of the center of mass coordinates within the many-particle wavefunction due to the gauge transformation, and can give rise to a non-trivial geometric phase called the Aharonov-Bohm (A-B) phase upon closing a topologically non-trivial circuit in the real-space trajectory corresponding to $\delta\vec{r}$ (e.g., a closed loop, a torus etc.).

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: the pattern of dark and bright fringes would move collectively up (or down) the screen.



For an electron through slit 1 (S_1) (that is, above the solenoid), the probability amplitude is

$$|\Psi_1\rangle \rightarrow |\Psi'_1\rangle = \exp\left(-\frac{ie}{c\hbar} \int_{\vec{r}_A}^{\vec{r}_B} \vec{A} \cdot d\vec{l}\right) |\Psi_1\rangle \quad (21.127)$$

Similarly, for an electron through S_2 , the probability amplitude is

$$|\Psi_2\rangle \rightarrow |\Psi'_2\rangle = \exp\left(-\frac{ie}{c\hbar} \int_{\vec{r}_A}^{\vec{r}_B} \vec{A} \cdot d\vec{l}\right) |\Psi_2\rangle \quad (21.128)$$

Thus, the total probability between the source (\vec{r}_A) and a point on the screen (\vec{r}_B) is

$$\begin{aligned} & (\langle \Psi'_2 | + \langle \Psi'_1 |) (|\Psi'_2\rangle + |\Psi'_1\rangle) \\ & \qquad \qquad \qquad \text{quantum interference terms} \\ & = \langle \Psi'_1 | \Psi'_1 \rangle + \langle \Psi'_2 | \Psi'_2 \rangle + \overbrace{\langle \Psi'_1 | \Psi'_2 \rangle + \langle \Psi'_2 | \Psi'_1 \rangle} \\ & = \langle \Psi_1 | \Psi_1 \rangle + \langle \Psi_2 | \Psi_2 \rangle + 2\text{Re} \{ \langle \Psi'_2 | \Psi'_1 \rangle \} \\ & = \langle \Psi_1 | \Psi_1 \rangle + \langle \Psi_2 | \Psi_2 \rangle + 2\text{Re} \left\{ \langle \Psi_2 | \Psi_1 \rangle \exp\left(-\frac{ie}{c\hbar} \int_{S_1} \vec{A} \cdot d\vec{l} + \frac{ie}{c\hbar} \int_{S_2} \vec{A} \cdot d\vec{l}\right) \right\} \quad (21.129) \\ & = \langle \Psi_1 | \Psi_1 \rangle + \langle \Psi_2 | \Psi_2 \rangle + 2\text{Re} \left\{ \langle \Psi_2 | \Psi_1 \rangle \exp\left(-\frac{ie}{c\hbar} \oint \vec{A} \cdot d\vec{l}\right) \right\} \\ & = \langle \Psi_1 | \Psi_1 \rangle + \langle \Psi_2 | \Psi_2 \rangle + 2\text{Re} \left\{ \langle \Psi_2 | \Psi_1 \rangle \exp\left(-2\pi i \frac{\Phi}{\Phi_0}\right) \right\} \end{aligned}$$

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. In the particle on a ring, the wavefunction picks up a seemingly trivial phase $e^{-2\pi ni}$ for $\Phi = n\Phi_0, n \in \mathbb{Z}$ even though the ground state has changed its quantum number to $m \equiv n$ due to spectral flow. The spectrum at $\Phi = n\Phi_0, n \in \mathbb{Z}$ clearly ‘looks the same’ as that at $\Phi = 0$ (modulo a reshuffle of the ordering of the states). This similarity of the spectrum is referred to an ‘invariance of the system under a large gauge transformation’, i.e., by a change in the vector potential that is not small (of the kind we usually consider) but large (and leading to $\Phi = n\Phi_0$).

The Aharonov-Bohm phase as an example of a Berry (geometric) phase. In the A-B effect, the twisted boundary conditions

$$\Psi(\phi + 2\pi) = \exp\left(-2\pi i \frac{\Phi}{\Phi_0}\right) \psi(\phi) \quad (21.130)$$

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.

Interestingly, this idea has been used by Oshikawa to obtain the Luttinger volume, i.e., the filled number of states within the Fermi volume at temperature $T = 0$ (corresponding to the total particle number N , and known to be a topological invariant of the N -particle wavefunction from Luttinger's theorem) upon a large gauge transformation/ changing the external magnetic field by a flux proportional to $N\Phi_0$. Further, the phase $\phi_{\delta\vec{r}}(\vec{r})$ is clearly distinct from the arbitrary constant phase θ considered above (and which was related to the conservation of the total number of particles N).

Chapter 22

Approximate Methods I: Time-independent Perturbation Theory

Till now, we have been very lucky in obtaining exact, closed form expressions for the eigenvalues and eigenvectors of various problems we have dealt with. However, it turns out that these belong to a small subset of problems that we encounter in real life. In the absence of exact analytical solutions, we have to typically rely on methods that have one form of approximation or another. While they typically have their own domain of validity, they often bring valuable qualitative insight on the nature of the solution. And with some luck, our answers may also prove to be reasonably accurate when compared quantitatively with experimental realisations of the problem at hand. Finally, we can check our results by comparing against an exact numerical diagonalisation of the problem.

Here, we will consider a particular approximate method called (Rayleigh-Schrödinger) perturbation theory. The approach involves considering a system characterised by a Hamiltonian that can be decomposed into two parts: $H = H_0 + V$. The first part, H_0 , corresponds to a Hamiltonian whose eigenvalues and eigenvectors are known (or can be determined exactly in some way). The second part V , which is called the “perturbation”, refers to the fact that in learning the eigenstates and eigenvalues of the full Hamiltonian H , the effects of V on the eigenstates and eigenvalues of H_0 are to be determined in some systematic manner. As we shall see, the important assumption being made in taking this approach is that the effects caused by V are small.

22.1 Non-degenerate Perturbation theory

Here, we will study the case of non-degenerate perturbation theory for a time-independent Hamiltonian $H = H_0 + V$, and where we are interested in finding the eigenstates and eigenvalues of H

$$H |\psi_n\rangle = (H_0 + V) |\psi_n\rangle = E_n |\psi_n\rangle . \quad (22.1)$$

Instead of the above, however, we will solve

$$(H_0 + \lambda V) |\psi_n\rangle = E_n |\psi_n\rangle , \quad (22.2)$$

where $0 \leq \lambda \leq 1$ is a continuous real parameter that controls the strength of the perturbation, and can be varied continuously between $\lambda = 0$ (the unperturbed problem H_0) and $\lambda = 1$ (the H given above). Indeed, we will finally set $\lambda = 1$ in order to receive the answers we sought for H .

The perturbation theory approach involves an expansion of the energy eigenvalues E_n and eigenstates $|\psi_n\rangle$ in powers of λ , and with the expectation that the expansion will lead to a smooth interpolation between the eigenstates and eigenvalues of H_0 ($|\psi^{(0)}\rangle$ and $E_n^{(0)}$) and those of H ($|\psi\rangle$ and E_n). As we shall see, λ keeps track of the number of times the perturbation is introduced within each term of the expansion. Further, we are assuming the analyticity of $|\psi\rangle$ and E_n in the complex λ plane around the value of $\lambda = 0$; when the converse happens, the perturbation theory expansion will reflect this. Also, while we typically expect that the first few orders of the perturbation should be sufficient to yield good approximations (if the approach is valid at all) that can be quantified and benchmarked against experimental results, it is altogether remarkable to note that there are instances where perturbation theory to the 12th order in the expansion has been carried out and is found to compare quite favourably with experimental observations!

Thus, we start from the formal expressions for $|\psi_n\rangle$ and E_n as power-series in λ

$$|\psi_n\rangle = |\psi_n^{(0)}\rangle + \lambda |\psi_n^{(1)}\rangle + \lambda^2 |\psi_n^{(2)}\rangle + \dots , \quad (22.3)$$

$$E_n = E_n^{(0)} + \lambda E_n^{(1)} + \lambda^2 E_n^{(2)} + \dots . \quad (22.4)$$

Note that we have assumed here that $\langle \psi_n | \psi_n^{(0)} \rangle = 1$ rather than the usual normalisation condition $\langle \psi_n | \psi_n \rangle = 1$. This unusual normalisation condition means that there exists an overall multiplicative normalisation factor that we not determined as yet. We will proceed with $|\psi_n\rangle$ as defined above for now, and carry out the normalisation in the next section.

We now write the Schrödinger equation $H |\psi_n\rangle = (H_0 + \lambda V) |\psi_n\rangle = E_n |\psi_n\rangle$ as

$$(H_0 + \lambda V)(|\psi_n^{(0)}\rangle + \lambda |\psi_n^{(1)}\rangle + \lambda^2 |\psi_n^{(2)}\rangle + \dots) = (E_n^{(0)} + \lambda E_n^{(1)} + \lambda^2 E_n^{(2)} + \dots)(|\psi_n^{(0)}\rangle + \lambda |\psi_n^{(1)}\rangle + \lambda^2 |\psi_n^{(2)}\rangle + \dots) . \quad (22.5)$$

Equating terms with equal powers of λ on both sides of the above, we get the following equations from the first upto the j th power of λ

$$\mathcal{O}(\lambda^0) : H_0 |\psi_n^{(0)}\rangle = E_n^{(0)} |\psi_n^{(0)}\rangle ,$$

$$\mathcal{O}(\lambda^1) : H_0 |\psi_n^{(1)}\rangle + V |\psi_n^{(0)}\rangle = E_n^{(0)} |\psi_n^{(1)}\rangle + E_n^{(1)} |\psi_n^{(0)}\rangle , \quad (22.6)$$

$$\mathcal{O}(\lambda^2) : H_0 |\psi_n^{(2)}\rangle + V |\psi_n^{(1)}\rangle = E_n^{(0)} |\psi_n^{(2)}\rangle + E_n^{(1)} |\psi_n^{(1)}\rangle + E_n^{(2)} |\psi_n^{(0)}\rangle , \quad (22.7)$$

$$\mathcal{O}(\lambda^j) : H_0 |\psi_n^{(j)}\rangle + V |\psi_n^{(j-1)}\rangle = E_n^{(0)} |\psi_n^{(j)}\rangle + E_n^{(1)} |\psi_n^{(j-1)}\rangle + \dots + E_n^{(j)} |\psi_n^{(0)}\rangle \quad (22.8)$$

We can also write the state $|\psi_n^{(j)}\rangle$ in the expansion for $|\psi_n\rangle$ given above in terms of the

complete basis of zeroth-order eigenstates ($\{|\psi_n^{(0)}\rangle\}$) of H_0

$$|\psi_n^{(j)}\rangle = \sum_{k \neq n} c_{nk}^{(j)} |\psi_k^{(0)}\rangle , \quad (22.9)$$

where $|\psi_n^{(0)}\rangle$ has been explicitly excluded as it is already present in the first term of the expansion for $|\psi_n\rangle$ given above. By taking such a state, we are assuming that the perturbing potential V is simply admixing the eigenstates of the H_0 , and that the new states obtained are not orthogonal to those of H_0 . The coefficients $c_{nk}^{(j)}$ and the corrections to the energy $E_n^{(j)}$ are to be determined.

Now, taking the inner product of eq.(22.6) with $\langle\psi_n^{(0)}|$ yields the first order correction to the energy $E_n^{(1)}$

$$E_n^{(1)} = \langle\psi_n^{(0)}| V |\psi_n^{(0)}\rangle , \quad (22.10)$$

and where we have used the orthogonality relation $\langle\psi_k^{(0)}|\psi_n^{(0)}\rangle = \delta_{kn}$ and the fact that $|\psi_n^{(1)}\rangle = \sum_{k \neq n} c_{nk}^{(1)} |\psi_k^{(0)}\rangle$. Thus, the first order energy correction is very simply obtained from the expectation value of the perturbing potential V with respect to the zeroth-order eigenstates.

Similarly, we use the expansion for $|\psi_n^{(1)}\rangle = \sum_{k \neq n} c_{nk}^{(1)} |\psi_k^{(0)}\rangle$ given above and take the inner product of eq.(22.6) with $\langle\psi_k^{(0)}|$ (as well as using the orthogonality relation $\langle\psi_k^{(0)}|\psi_n^{(0)}\rangle = \delta_{kn}$) to obtain the coefficient $c_{nk}^{(1)}$

$$\begin{aligned} \langle\psi_k^{(0)}| \left[H_0 \left(\sum_{k \neq n} c_{nk}^{(1)} |\psi_k^{(0)}\rangle \right) + V |\psi_n^{(0)}\rangle \right] &= \langle\psi_k^{(0)}| \left[E_n^{(0)} \left(\sum_{k \neq n} c_{nk}^{(1)} |\psi_k^{(0)}\rangle \right) + E_n^{(1)} |\psi_n^{(0)}\rangle \right] , \\ \implies c_{nk}^{(1)} \langle\psi_k^{(0)}| H_0 |\psi_k^{(0)}\rangle + \langle\psi_k^{(0)}| V |\psi_n^{(0)}\rangle &= c_{nk}^{(1)} E_n^{(0)} \langle\psi_k^{(0)}|\psi_k^{(0)}\rangle , \\ \implies (E_n^{(0)} - E_k^{(0)}) c_{nk}^{(1)} &= \langle\psi_k^{(0)}| V |\psi_n^{(0)}\rangle , \\ \implies c_{nk}^{(1)} &= \frac{\langle\psi_k^{(0)}| V |\psi_n^{(0)}\rangle}{(E_n^{(0)} - E_k^{(0)})} . \end{aligned} \quad (22.11)$$

The above relation indicates that the “mixing coefficient” $c_{nk}^{(1)}$ arises from the overlap between the zeroth eigenstates $|\psi_n^{(0)}\rangle$ and $|\psi_k^{(0)}\rangle$ due to the perturbing potential V (i.e., the transition caused between $|\psi_n^{(0)}\rangle$ and $|\psi_k^{(0)}\rangle$ by V) and divided by the zeroth energy difference between them $E_n^{(0)} - E_k^{(0)}$ (i.e., the zeroth transition energy cost). It is obvious that for a general state $|\psi_n^{(0)}\rangle$, $|\psi_k^{(0)}\rangle$ can be any other state (i.e., $k \neq n$) that is connected to $|\psi_n^{(0)}\rangle$ via V ; for the special case where n corresponds to the ground state, the set of states $\{|\psi_k^{(0)}\rangle\}$ corresponds to all the excited states to which the ground state is connected via V .

In this way, we see that the first order correction $|\psi_n^{(1)}\rangle$ is given by

$$|\psi_n^{(1)}\rangle = \sum_{k \neq n} c_{nk}^{(1)} |\psi_k^{(0)}\rangle = \sum_{k \neq n} \frac{\langle \psi_k^{(0)} | V | \psi_n^{(0)} \rangle}{(E_n^{(0)} - E_k^{(0)})} |\psi_k^{(0)}\rangle . \quad (22.12)$$

We can now proceed further by taking the inner product of eq.(22.7) with $\langle \psi_n^{(0)} |$ to obtain the second order correction to the energy $E_n^{(2)}$

$$E_n^{(2)} = \langle \psi_n^{(0)} | V | \psi_n^{(1)} \rangle = \langle \psi_n^{(0)} | V \left(\sum_{k \neq n} c_{nk}^{(1)} |\psi_k^{(0)}\rangle \right) \rangle = \sum_{k \neq n} \frac{|(\langle \psi_k^{(0)} | V | \psi_n^{(0)} \rangle)|^2}{(E_n^{(0)} - E_k^{(0)})} , \quad (22.13)$$

where we have taken $|\psi_n^{(2)}\rangle = \sum_{k \neq n} c_{nk}^{(2)} |\psi_k^{(0)}\rangle$. The above relation indicates that the second order correction $E_n^{(2)}$ also arises from the $| |^2$ of the overlap between the zeroth eigenstates $|\psi_n^{(0)}\rangle$ and $|\psi_k^{(0)}\rangle$ due to the perturbing potential V (i.e., the $| |^2$ of the transition amplitude between $|\psi_n^{(0)}\rangle$ and $|\psi_k^{(0)}\rangle$ caused by V) and divided by the zeroth energy difference between them $E_n^{(0)} - E_k^{(0)}$ (i.e., the zeroth transition energy cost). It is easily seen that

$$\frac{|(\langle \psi_k^{(0)} | V | \psi_n^{(0)} \rangle)|^2}{(E_n^{(0)} - E_k^{(0)})} = (E_n^{(0)} - E_k^{(0)}) \times \frac{|(\langle \psi_k^{(0)} | V | \psi_n^{(0)} \rangle)|^2}{(E_n^{(0)} - E_k^{(0)})^2} = (E_n^{(0)} - E_k^{(0)}) |c_{nk}^{(1)}|^2 \quad (22.14)$$

$$\implies E_n^{(2)} = \sum_{k \neq n} (E_n^{(0)} - E_k^{(0)}) |c_{nk}^{(1)}|^2 , \quad (22.15)$$

indicating that the second order correction $E_n^{(2)}$ is simply the average of the transition energies $(E_n^{(0)} - E_k^{(0)})$ weighted by their respective transition probabilities $|c_{nk}^{(1)}|^2$. As before, it is obvious that for a general state $|\psi_n^{(0)}\rangle$, $|\psi_k^{(0)}\rangle$ can be any other state (i.e., $k \neq n$) that is connected to $|\psi_n^{(0)}\rangle$ via V ; for the special case where n corresponds to the ground state, the set of states $\{|\psi_k^{(0)}\rangle\}$ corresponds to all the excited states to which the ground state is connected via V .

Similarly, we take the inner product of eq.(22.7) with $\langle \psi_k^{(0)} |$ to obtain the second order

correction coefficient $c_{nk}^{(2)}$

$$\begin{aligned}
\left\langle \psi_k^{(0)} \right| [E_n^{(0)} |\psi_n^{(2)}\rangle - H_0 |\psi_n^{(2)}\rangle] &= \left\langle \psi_k^{(0)} \right| [V |\psi_n^{(1)}\rangle - E_n^{(1)} |\psi_n^{(1)}\rangle] , \\
\implies (E_n^{(0)} - E_k^{(0)}) c_{nk}^{(2)} &= \left\langle \psi_k^{(0)} \right| \left[V \sum_{j \neq n} \frac{\left\langle \psi_j^{(0)} \right| V \left| \psi_n^{(0)} \right\rangle}{E_n^{(0)} - E_j^{(0)}} \left| \psi_j^{(0)} \right\rangle \right. \\
&\quad \left. - \left\langle \psi_n^{(0)} \right| V \left| \psi_n^{(0)} \right\rangle \sum_{k \neq n} \frac{\left\langle \psi_k^{(0)} \right| V \left| \psi_n^{(0)} \right\rangle}{E_n^{(0)} - E_k^{(0)}} \left| \psi_k^{(0)} \right\rangle \right] , \\
c_{nk}^{(2)} &= \sum_{j \neq n} \frac{\left\langle \psi_k^{(0)} \right| V \left| \psi_j^{(0)} \right\rangle \left\langle \psi_j^{(0)} \right| V \left| \psi_n^{(0)} \right\rangle}{(E_n^{(0)} - E_k^{(0)})(E_n^{(0)} - E_j^{(0)})} \\
&\quad - \frac{\left\langle \psi_k^{(0)} \right| V \left| \psi_n^{(0)} \right\rangle \left\langle \psi_n^{(0)} \right| V \left| \psi_n^{(0)} \right\rangle}{(E_n^{(0)} - E_k^{(0)})^2} , \tag{22.16}
\end{aligned}$$

$$= \sum_{j \neq n} \frac{(E_j^{(0)} - E_k^{(0)})}{(E_n^{(0)} - E_k^{(0)})} c_{jk}^{(1)} c_{nj}^{(1)} - \frac{\left\langle \psi_n^{(0)} \right| V \left| \psi_n^{(0)} \right\rangle}{(E_n^{(0)} - E_k^{(0)})} c_{nk}^{(1)} \tag{22.17}$$

$$= \sum_{j \neq n} \frac{(E_j^{(0)} - E_k^{(0)})}{(E_n^{(0)} - E_k^{(0)})} c_{jk}^{(1)} c_{nj}^{(1)} - \frac{E_n^{(1)} c_{nk}^{(1)}}{(E_n^{(0)} - E_k^{(0)})} . \tag{22.18}$$

The first term in the above indicates that there are two transitions involved (from $n \rightarrow j$ and then from $j \rightarrow k$) while the second term indicates only one transition (from $n \rightarrow k$) weighted by the ratio of the first order correction and the transition energy.

This gives us the perturbed eigenstate to second order in the perturbation as

$$\begin{aligned}
|\psi_n\rangle &= |\psi_n^{(0)}\rangle + \lambda |\psi_n^{(1)}\rangle + \lambda^2 |\psi_n^{(2)}\rangle , \\
&= |\psi_n^{(0)}\rangle + \lambda \sum_{k \neq n} \frac{\left\langle \psi_k^{(0)} \right| V \left| \psi_n^{(0)} \right\rangle}{(E_n^{(0)} - E_k^{(0)})} \left| \psi_k^{(0)} \right\rangle \\
&\quad + \lambda^2 \left(\sum_{k \neq n} \sum_{j \neq n} \frac{\left\langle \psi_k^{(0)} \right| V \left| \psi_j^{(0)} \right\rangle \left\langle \psi_j^{(0)} \right| V \left| \psi_n^{(0)} \right\rangle}{(E_n^{(0)} - E_k^{(0)})(E_n^{(0)} - E_j^{(0)})} \left| \psi_k^{(0)} \right\rangle \right. \\
&\quad \left. - \sum_{k \neq n} \frac{\left\langle \psi_k^{(0)} \right| V \left| \psi_n^{(0)} \right\rangle \left\langle \psi_n^{(0)} \right| V \left| \psi_n^{(0)} \right\rangle}{(E_n^{(0)} - E_k^{(0)})^2} \left| \psi_k^{(0)} \right\rangle \right) . \tag{22.19}
\end{aligned}$$

It is clear from this expression that the n th level mixes with various other unperturbed energy eigenstates of H_0 through the perturbation V , i.e., it acquires components along other unperturbed energy eigenstates.

Further, the energy shift $\Delta_n \equiv E_n - E_n^{(0)}$ to second order in perturbation theory is given by

$$\begin{aligned}\Delta_n \equiv E_n - E_n^{(0)} &= \lambda E_n^{(1)} + \lambda^2 E_n^{(2)}, \\ &= \lambda \langle \psi_n^{(0)} | V | \psi_n^{(0)} \rangle + \lambda^2 \sum_{k \neq n} \frac{|(\langle \psi_k^{(0)} | V | \psi_n^{(0)} \rangle)|^2}{(E_n^{(0)} - E_k^{(0)})}, \\ &= \lambda \langle \psi_n^{(0)} | V | \psi_n^{(0)} \rangle + \lambda^2 \sum_{k \neq n} (E_n^{(0)} - E_k^{(0)}) |c_{nk}^{(1)}|^2.\end{aligned}\quad (22.20)$$

The second term in the expression for the shift Δ_n tells us that any two levels, say the n th and m th, will repel one another (i.e., are pushed apart) upon being connected by the perturbation matrix element $\langle \psi_m^{(0)} | V | \psi_n^{(0)} \rangle$: the lower one (say the n th) is pushed down by the amount $(\langle \psi_m^{(0)} | V | \psi_n^{(0)} \rangle)^2 / (E_m^{(0)} - E_n^{(0)})$ while the upper one (say the m th) is raised up by the same amount. Another consequence is that if the lower state (say the n th) corresponds to the ground state, the second order term is always found to be negative, i.e., the perturbation always lowers the ground state energy.

Finally, it is worth noting that the perturbation expansions for the energy splitting Δ_n and the eigenstate $|\psi_n\rangle$ will converge only if the quantity $|\langle \psi_k^{(0)} | V | \psi_n^{(0)} \rangle| / (E_n^{(0)} - E_k^{(0)})$ is sufficiently small, i.e.,

$$|\langle \psi_k^{(0)} | V | \psi_n^{(0)} \rangle| \ll |(E_n^{(0)} - E_k^{(0)})|. \quad (22.21)$$

This is the mathematical criterion that justifies the perturbation theory formalism.

We end with a small aside. Very generally, if we write the perturbation in the Hamiltonian as

$$H(g) = H_0(g_0) + (g - g_0)V(g_0) = H_0(g_0) + (g - g_0)\frac{\partial H(g)}{\partial g}|_{g=g_0} \quad (22.22)$$

as a Taylor series expansion in the small parameter $g - g_0$ about the Hamiltonian $H_0(g_0)$, we obtain the first order correction to the energy n th level as

$$E_n = E_n^{(0)} + (g - g_0)E_n^{(1)} = E_n^{(0)} + (g - g_0)\frac{\partial E_n}{\partial g}|_{g=g_0}. \quad (22.23)$$

It is then easily seen that

$$\frac{\partial E_n}{\partial g} \equiv E_n^{(1)} = \langle \psi_n^{(0)} | V(g_0) | \psi_n^{(0)} \rangle, \quad (22.24)$$

$$\Rightarrow \frac{\partial E_n}{\partial g} = \langle \psi_n^{(0)} | \frac{\partial H(g)}{\partial g} | \psi_n^{(0)} \rangle|_{g=g_0}. \quad (22.25)$$

This is known as the Hellmann-Feynman theorem. While we have used the Taylor expansion of $H(g)$ to obtain it, one can derive it on very general grounds. We start by noting that

$$E_n(g) = \langle \psi_n(g) | H(g) | \psi_n(g) \rangle. \quad (22.26)$$

Then, taking a derivative of this expression with respect to the parameter g , we obtain

$$\begin{aligned}
\frac{\partial E_n(g)}{\partial g} &= \left\langle \frac{\partial \psi_n(g)}{\partial g} \middle| H(g) |\psi_n(g)\rangle + \langle \psi_n(g)| \frac{\partial H(g)}{\partial g} |\psi_n(g)\rangle + \langle \psi_n(g)| H(g) \left| \frac{\partial \psi_n}{\partial g} \right\rangle \right\rangle_{22.27} \\
&= E_n \left[\left\langle \frac{\partial \psi_n(g)}{\partial g} \middle| \psi_n \right\rangle + \left\langle \psi_n(g) \middle| \frac{\partial \psi_n}{\partial g} \right\rangle \right] + \langle \psi_n(g)| \frac{\partial H(g)}{\partial g} |\psi_n(g)\rangle , \\
&= E_n \frac{\partial}{\partial g} [\langle \psi_n(g)| \psi_n(g)\rangle] + \langle \psi_n(g)| \frac{\partial H(g)}{\partial g} |\psi_n(g)\rangle , \\
&= \langle \psi_n(g)| \frac{\partial H(g)}{\partial g} |\psi_n(g)\rangle \quad (\text{as } \langle \psi_n(g)| \psi_n(g)\rangle = 1) . \tag{22.28}
\end{aligned}$$

22.1.1 Wavefunction Renormalisation

An important aspect remains to be addressed before we can start applying what we have learnt to a few problems. As mentioned earlier, a careful look at the wavefunction $|\psi_n\rangle$ written above in eq.(22.3)

$$|\psi_n\rangle = |\psi_n^{(0)}\rangle + \lambda |\psi_n^{(1)}\rangle + \lambda^2 |\psi_n^{(2)}\rangle + \dots ,$$

will show you that it is not normalised as presently defined. Indeed, we started by assuming that $\langle \psi_n | \psi_n^{(0)} \rangle = 1$ rather than the usual normalisation condition $\langle \psi_n | \psi_n \rangle = 1$. This unusual normalisation condition means that there exists an overall multiplicative normalisation factor that we will now determine. However, we first discuss briefly the importance for why the correct normalisation needs to be restored. The perturbed expansion of $|\psi_n\rangle$ shows, as we have mentioned above, that the perturbation V causes the unperturbed eigenstate $|\psi_n^{(0)}\rangle$ to “mix” with (or “leak into”) other unperturbed eigenstates $|\psi_m^{(0)}\rangle$, $m \neq n$. Therefore, in contradiction with our earlier assumption, the inner product $\langle \psi_n^{(0)} | \psi_n \rangle \neq 1$, and shows the existence of non-zero probabilities $|\langle \psi_n^{(j)} | \psi_n \rangle|^2$, $j \neq 0$. Thus, in order to correctly account for the shifting around of the probability weight factors (i.e., spectral weight transfer) caused by the perturbation V , it is important to normalise the perturbed eigenstate $|\psi_n\rangle$.

Let us define the normalised perturbed eigenstate as

$$|\psi_n\rangle_N = Z_n^{1/2} |\psi_n\rangle \quad \text{such that } {}_N \langle \psi_n | \psi_n \rangle_N = 1 , \tag{22.29}$$

and $Z_n^{1/2}$ is the overall normalisation factor to be determined. Then, taking an inner product of $|\psi_n\rangle_N$ with $\langle \psi_n^{(0)} |$ gives

$$\begin{aligned}
\langle \psi_n^{(0)} | \psi_n \rangle_N &= \langle \psi_n^{(0)} | Z_n^{1/2} |\psi_n\rangle , \\
&= \langle \psi_n^{(0)} | Z_n^{1/2} [|\psi_n^{(0)}\rangle + \lambda |\psi_n^{(1)}\rangle + \lambda^2 |\psi_n^{(2)}\rangle + \dots] , \\
&= Z_n^{1/2} ,
\end{aligned} \tag{22.30}$$

as $|\psi_n^{(k)}\rangle$ ($k \neq 0$) comprises of states $|\psi_m^{(0)}\rangle$ ($m \neq n$), and $\langle \psi_n^{(0)} | \psi_m^{(0)} \rangle = 0$ for $m \neq n$. This displays clearly that Z_n corresponds to the probability for the perturbed eigenstate $|\psi_n\rangle_N$

to be found in the related unperturbed eigenstate $|\psi_n^{(0)}\rangle$. Clearly, the rest of the probability will go into the mixing with other unperturbed eigenstates $|\psi_m^{(0)}\rangle$, $m \neq n$. In the language of quantum information processing, the quantity $Z_n^{1/2} = \langle \psi_n^{(0)} | \psi_n \rangle_N$ is simply the “fidelity” of the normalised perturbed state $|\psi_n\rangle_N$ with the corresponding unperturbed state $|\psi_n^{(0)}\rangle$, i.e., a measure of how close they are to one another.

We can now proceed to compute Z_n by using the fact that

$$_N \langle \psi_n | \psi_n \rangle_N = Z_n \langle \psi_n | \psi_n \rangle = 1 . \quad (22.31)$$

Thus, we can compute

$$\begin{aligned} Z_n^{-1} &= \langle \psi_n | \psi_n \rangle = (\langle \psi_n^{(0)} | + \lambda \langle \psi_n^{(1)} | + \lambda^2 \langle \psi_n^{(2)} | + \dots)(\langle \psi_n^{(0)} | + \lambda \langle \psi_n^{(1)} | + \lambda^2 \langle \psi_n^{(2)} | + \dots) , \\ &= \langle \psi_n^{(0)} | \psi_n^{(0)} \rangle + \lambda^2 \langle \psi_n^{(1)} | \psi_n^{(1)} \rangle + \mathcal{O}(\lambda^3) \quad (\text{as } \langle \psi_n^{(1)} | \psi_n^{(0)} \rangle = 0) , \\ &= 1 + \lambda^2 \sum_{k \neq n} \frac{|\langle \psi_k^{(0)} | V | \psi_n^{(0)} \rangle|^2}{(E_n^{(0)} - E_k^{(0)})^2} + \mathcal{O}(\lambda^3) = 1 + \lambda^2 \sum_{k \neq n} |c_{nk}^{(1)}|^2 + \mathcal{O}(\lambda^3) , \end{aligned} \quad (22.32)$$

where we have used eq.(22.12) to rewrite the second term (i.e., of order $\mathcal{O}(\lambda^2)$). We can now invert the above to obtain Z_n as

$$Z_n = 1 - \lambda^2 \sum_{k \neq n} \frac{|\langle \psi_k^{(0)} | V | \psi_n^{(0)} \rangle|^2}{(E_n^{(0)} - E_k^{(0)})^2} + \mathcal{O}(\lambda^3) = 1 - \lambda^2 \sum_{k \neq n} |c_{nk}^{(1)}|^2 + \mathcal{O}(\lambda^3). \quad (22.33)$$

The second term in Z_n clearly shows that the mixing between unperturbed eigenstates due to the perturbation V (given by finite and non-zero matrix elements $\langle \psi_k^{(0)} | V | \psi_n^{(0)} \rangle$) is the “leakage” in the probability to unperturbed eigenstates other than $|\psi_n^{(0)}\rangle$ (i.e., reduction in the perfect probability for $|\psi_n\rangle$ to remain $|\psi_n^{(0)}\rangle$). Clearly, $0 \leq Z_n \leq 1$.

Note that the inversion of the relation for Z_n^{-1} is only valid when the quantity $\sum_{k \neq n} |c_{nk}^{(1)}|^2 \ll 1$ and $0 \leq \lambda \leq 1$. The physics of this can be understood as follows. Recall that perturbation theory works best when

$$\begin{aligned} |\langle \psi_k^{(0)} | V | \psi_n^{(0)} \rangle| &<< |(E_n^{(0)} - E_k^{(0)})| , \\ \implies c_{nk}^{(1)} &<< 1 , \\ \implies |c_{nk}^{(1)}|^2 &= \frac{|\langle \psi_k^{(0)} | V | \psi_n^{(0)} \rangle|^2}{(E_n^{(0)} - E_k^{(0)})^2} << 1 , \end{aligned} \quad (22.34)$$

indicating that the contribution of spectral weight within $|\psi_n\rangle$ from the states $|\psi_m^{(0)}\rangle$ ($m \neq n$) is sufficiently small such that $Z \sim \mathcal{O}(1)$, consistent with the argument presented just above

for the validity of the inversion of the relation for Z_n^{-1} . Further, the case $Z_n \rightarrow 0$ indicates $\langle \psi_n^{(0)} | \psi_n \rangle_N \rightarrow 0$, i.e., the n th eigenstate of H becomes orthogonal to the corresponding eigenstate of H_0 . Under such circumstances, the perturbation theory approach is rendered invalid. We will comment on this shortly.

Finally, noting that

$$\begin{aligned} E_n &= E_n^{(0)} + \lambda E_n^{(1)} + \lambda^2 E_n^{(2)} + \mathcal{O}(\lambda^3) = E_n^{(0)} + \Delta_n + \mathcal{O}(\lambda^3), \\ &= E_n^{(0)} + \lambda \langle \psi_n^{(0)} | V | \psi_n^{(0)} \rangle + \lambda^2 \sum_{k \neq n} \frac{|(\langle \psi_k^{(0)} | V | \psi_n^{(0)} \rangle)|^2}{(E_n^{(0)} - E_k^{(0)})} + \mathcal{O}(\lambda^3), \end{aligned} \quad (22.35)$$

we can relate the expression for Z_n obtained above to E_n (as well as the energy shift Δ_n) as

$$\begin{aligned} Z_n &= 1 - \lambda^2 \sum_{k \neq n} \frac{|(\langle \psi_k^{(0)} | V | \psi_n^{(0)} \rangle)|^2}{(E_n^{(0)} - E_k^{(0)})^2} + \mathcal{O}(\lambda^3), \\ &= \frac{\partial E_n}{\partial E_n^{(0)}} = 1 + \frac{\partial \Delta_n}{\partial E_n^{(0)}}, \end{aligned} \quad (22.36)$$

where we have assumed that the matrix elements $\langle \psi_k^{(0)} | V | \psi_n^{(0)} \rangle$ are fixed while taking the partial derivative with respect to $E_n^{(0)}$. While we have quoted the above relation based on a calculation to $\mathcal{O}(\lambda^2)$, it has been asserted (see, e.g., the text by Sakurai) that the relation holds beyond second order perturbation theory (i.e., it is a non-perturbative statement); presumably, this is obtained by computing the energy shift $\Delta_n(E_n^{(0)})$ in an exact manner.

We end by pointing out that the idea of the wavefunction renormalisation plays a very important conceptual role in many-body physics: the excitations of an interacting gas of electrons in a metal were shown by Lev Landau to be essentially identical to those of the non-interacting system (upon introducing inter-particle interactions adiabatically) by computing the wavefunction renormalisation coefficient for the lowest lying excitations above the ground state (with the excitations of the non-interacting electron gas as the reference) and showing that it was non-zero. Landau called it the “quasiparticle residue”, but its essentially the wavefunction renormalisation. This is a milestone result in the field of condensed matter physics, and referred to as the “adiabatic continuity” at the heart of Landau’s Fermi liquid paradigm for metals. This line of thinking has also been extended in recent times to show that the wavefunction renormalisation of the ground state vanishes at a quantum phase transition, indicating the orthogonality between the quantum states of matter separated by the transition. It also indicates that the passage through a quantum phase transition cannot be captured by a perturbation theory approach. A similar orthogonality mechanism of the lowest lying excitations is also the cause of the breakdown of Landau’s Fermi liquid paradigm of the metal, and its replacement by new types of metallic behaviours.

22.2 The anharmonic oscillator with a quartic perturbation potential

We will now apply the non-degenerate perturbation theory technique to obtain results for the problem of a harmonic oscillator with a quartic perturbation potential:

$$H = H_0 + V = \frac{1}{2m}(p^2 + m^2\omega^2x^2) + \lambda gx^4 , \quad (22.37)$$

where the first two terms indicate the unperturbed harmonic oscillator Hamiltonian (H_0) and the third term corresponds to the quartic perturbation potential $V = \lambda gx^4$. The parameter $0 \leq \lambda \leq 1$ is a dimensionless quantity, while the strength of the perturbation potential (also called the “coupling” in fields such as quantum field theory and many-body theory!) g is chosen in terms of the ground state energy eigenvalue ($\hbar\omega/2$) and the width of the Gaussian ground state wavefunction ($\psi_0(x) = \langle x|0\rangle = (\frac{m\omega}{\pi\hbar})^{1/4} e^{-x^2/x_0^2}$, $x_0 = \sqrt{2\hbar/m\omega}$) such that the dimensions of V are that of energy,

$$g = \frac{\hbar\omega}{2x_0^4} = \frac{m^2\omega^3}{8\hbar} . \quad (22.38)$$

We will use the algebraic (or ladder operator) approach towards solving this problem. Recall the following relations for the raising and lowering (“ladder”) operators and their action on the number operator (\hat{n}) eigenstates of the harmonic oscillator

$$\begin{aligned} x &= \sqrt{\frac{\hbar}{2m\omega}}(a_+ + a_-) , \quad p = i\sqrt{\frac{m\omega\hbar}{2}}(a_+ - a_-) , \\ a_{\pm} &= \frac{1}{\sqrt{2m\omega\hbar}}(\mp ip + m\omega x) , \quad [a_-, a_+] = 1 \\ a_+|n\rangle &= \sqrt{n+1}|n+1\rangle , \quad a_-|n\rangle = \sqrt{n}|n-1\rangle , \quad a_-|n=0\rangle = 0 = \langle n=0|a_+ , \\ \hat{n} &\equiv a_+a_- , \quad \hat{n}|n\rangle = n|n\rangle , \quad \hat{n}|n=0\rangle = 0 , \quad |n\rangle = \frac{(a_+)^n}{\sqrt{n!}}|0\rangle . \end{aligned} \quad (22.39)$$

Thus, we can rewrite the problem as

$$\begin{aligned} H &= H_0 + V , \\ H_0 &= (\hat{n} + \frac{1}{2})\hbar\omega , \end{aligned} \quad (22.40)$$

$$\begin{aligned} V &= \lambda gx^4 = \lambda \frac{\hbar\omega}{2} \left(\frac{x}{x_0}\right)^4 , \\ &= \lambda \frac{m^2\omega^3}{8\hbar} \left(\frac{\hbar}{2m\omega}\right)^2 (a_+ + a_-)^4 = \lambda \frac{\hbar\omega}{32} (a_+ + a_-)^4 . \end{aligned} \quad (22.41)$$

An exact solution evades us as H_0 and V clearly do not commute with one another. Then, the first order correction to the ground state eigenvalue due to the quartic perturbation potential is given by

$$E_0^{(1)} = \lambda \frac{\hbar\omega}{32} \langle 0| (a_+ + a_-)^4 |0\rangle . \quad (22.42)$$

Now, in order to evaluate the matrix element in the above expression, we can use the operator algebra for the ladder and number operators given earlier. First, it is important to note that opening the expression $(a_+ + a_-)^4$ will give a large number of terms, each of which contains a product strings of a_+ and a_- operators that total four in number. In dealing with these, there are two simplifying rules that we must remember

- only expressions that have an equal number of a_+ and a_- operators can contribute to the matrix element $\langle 0| (a_+ + a_-)^4 |0\rangle$. This is simply because an unequal number of will always finally lead to an inner product between two orthogonal states $\langle m|n\rangle$ ($m \neq n$) and will therefore vanish. In essence, only an operator string that leads finally to an inner product $\langle m|n\rangle$ ($m = n$) can contribute, and
- all expressions that have an a_- acting on the ground state ket eigenstate $|0\rangle$ or an a_+ acting on the ground state bra eigenstate $\langle 0|$ will annihilate the ground state (as given above). Thus, all product strings that start with an a_+ and end with an a_- will naturally vanish upon computing the ground state matrix element.

Using the first condition, we can eliminate all terms of the kind $a_+a_+a_+a_+$, $a_+a_+a_+a_-$, $a_+a_+a_-a_+$, and their conjugates. This already simplifies our search radically. Further, the second condition also eliminates the term $a_+a_-a_+a_-$, even though it has an equal number of a_+ s and a_- s. Finally, we are left with just two terms: $a_-a_-a_+a_+$ and $a_-a_+a_-a_+$. We can now simplify these two as follows:

$$\begin{aligned} a_-a_-a_+a_+ + a_-a_+a_-a_+ &= a_-a_+ + 2a_-a_+a_-a_+ , \quad (\text{using } [a_-, a_+] = 1) \\ &= 1 + a_+a_- + 2(1 + a_+a_-)(1 + a_+a_-) , \\ &= 1 + \hat{n} + 2(1 + \hat{n})(1 + \hat{n}) , \quad (\text{using } a_+a_- \equiv \hat{n}) \\ &= 3 + 5\hat{n} + 2\hat{n}^2 . \end{aligned} \tag{22.43}$$

Thus, we can write

$$\begin{aligned} \langle n| (a_-a_-a_+a_+ + a_-a_+a_-a_+) |n\rangle &= \langle n| (3 + 5\hat{n} + 2\hat{n}^2) |n\rangle , \\ &= 3 + 5\langle n| \hat{n} |n\rangle + 2\langle n| \hat{n}^2 |n\rangle , \\ &= 3 + 5n + 2n^2 , \quad (\text{using } \hat{n}|n\rangle = n|n\rangle , \hat{n}^2|n\rangle = n^2|n\rangle) , \\ &= 3 \quad \text{for } n = 0 \quad (\text{ground state of harmonic oscillator}) \end{aligned} \tag{22.44}$$

Finally, we obtain

$$\begin{aligned} E_0^{(1)} &= \lambda \frac{\hbar\omega}{32} \langle 0| (a_+ + a_-)^4 |0\rangle , \\ &= \frac{3}{32} \lambda \hbar\omega . \end{aligned} \tag{22.45}$$

We can easily extend this to the case of a general state $|n\rangle$ by noting that for excited states (i.e., $n \geq 1$), the second restriction above no longer always holds, even though the first does.

Thus, there are six matrix elements that contribute in general to the state $|n\rangle$

$$\begin{aligned}
\langle n | (a_+ + a_-)^4 | n \rangle &= \langle n | (a_- a_- a_+ a_+ + a_- a_+ a_- a_+ + a_- a_+ a_+ a_-) | n \rangle + \\
&\quad + \langle n | a_+ a_- a_- a_+ + a_+ a_- a_+ a_- + a_+ a_+ a_- a_- | n \rangle , \\
&= (\sqrt{(n+1)(n+2)})^2 + (\sqrt{n+1})^4 + (\sqrt{n(n+1)})^2 \\
&\quad + (\sqrt{n(n+1)})^2 + (\sqrt{n})^4 + (\sqrt{n(n-1)})^2 , \\
&= (n+1)(n+2) + (n+1)^2 + 2n(n+1) + n^2 + n(n-1) , \\
&= 6n^2 + 6n + 3 . \tag{22.46}
\end{aligned}$$

Thus, the first order correction for $|n\rangle$ is given by

$$\begin{aligned}
E_n^{(1)} &= \lambda \frac{\hbar\omega}{32} \langle n | (a_+ + a_-)^4 | n \rangle , \\
&= \frac{1}{32} \lambda \hbar\omega (6n^2 + 6n + 3) , \tag{22.47}
\end{aligned}$$

which is easily seen to obtain the result for $n = 0$ given above. Further, one can easily see that the correction grows as $\frac{3}{16}(n^2 + n + 1/2)$, which is small in comparison with the zeroth energy $E_n^{(0)} = (n+1/2)\hbar\omega$ only for small values of n . However, the correction is destined to become greater than $E_n^{(0)}$ for some finite value of n (indeed, this happens at $n = 5$). Even the spacing between successive levels is $3(n+1)\hbar\omega/8$ (at first order in perturbation theory); this outstrips the $\hbar\omega$ level spacing of the harmonic oscillator for $n = 2$. Thus, these first order shifts grow substantially quickly with n . This also means that the best results from a perturbation theory approach can be guaranteed for the ground state $|0\rangle$ (and perhaps the first excited state). Thus, we will restrict our focus to the perturbative corrections of the ground state for now.

Thus, we proceed to compute the second order correction for the ground state, which is given by

$$\begin{aligned}
E_0^{(2)} &= -\lambda^2 \sum_{k \neq 0}^{\infty} \frac{|\langle \psi_k^{(0)} | V | \psi_n^{(0)} \rangle|^2}{(E_k^{(0)} - E_0^{(0)})} , \\
&= -\left(\frac{\lambda \hbar\omega}{32}\right)^2 \sum_{k \neq 0}^{\infty} \frac{|\langle k | (a_+ + a_-)^4 | 0 \rangle|^2}{(E_k^{(0)} - E_0^{(0)})} ,
\end{aligned}$$

Now, $(a_+ + a_-)^4 |0\rangle = 3|0\rangle + 6\sqrt{2}|2\rangle + \sqrt{4!}|4\rangle$, (algebra and logic! see below)

$$\begin{aligned}
\implies E_0^{(2)} &= -\left(\frac{\lambda \hbar\omega}{32}\right)^2 \left[\frac{|\langle 2 | (a_+ + a_-)^4 | 0 \rangle|^2}{E_2^{(0)} - E_0^{(0)}} + \frac{|\langle 4 | (a_+ + a_-)^4 | 0 \rangle|^2}{E_4^{(0)} - E_0^{(0)}} \right] , \\
&= -\left(\frac{\lambda \hbar\omega}{32}\right)^2 \left[\frac{(6\sqrt{2})^2}{2\hbar\omega} + \frac{(\sqrt{4!})^2}{4\hbar\omega} \right] , \\
&= -\frac{21\lambda^2\hbar\omega}{512} . \tag{22.48}
\end{aligned}$$

You may wonder how we get such numbers as $6\sqrt{2}$ and $\sqrt{4!}$ as coefficients of $|2\rangle$ and $|4\rangle$ respectively in eq.(22.48)? First of all, the coefficient of 3 for $|0\rangle$ is precisely for the same reason as the 3 we obtained above in calculating the first order correction, i.e., it is an outcome of all operator product strings that start from $|0\rangle$ and end at $|0\rangle$. The $\sqrt{4!}$ coefficient of $|4\rangle$ is obtained from the action of $a_+a_+a_+a_+$ on $|0\rangle$, as is easily seen from the fact that each of the four a_+ takes us progressively from $|0\rangle$ to $|1\rangle$, $|2\rangle$, $|3\rangle$ and $|4\rangle$ and the fact that $a_+|n\rangle = \sqrt{n}|n+1\rangle$ (and this is the one and only term in $(a_+ + a_-)^4|0\rangle$ that can do this!). Thus,

$$\begin{aligned} (a_+a_+a_+a_+)|0\rangle &= \sqrt{1}(a_+a_+a_+)|1\rangle , \\ &= \sqrt{1 \times 2}(a_+a_+)|2\rangle , \\ &= \sqrt{1 \times 2 \times 3}(a_+)|3\rangle , \\ &= \sqrt{1 \times 2 \times 3 \times 4}|4\rangle = \sqrt{4!}|4\rangle . \end{aligned} \quad (22.49)$$

Further, we can see that

$$\begin{aligned} (a_+a_+a_-a_+)|0\rangle &= \sqrt{1}(a_+a_+a_-)|1\rangle , \\ &= \sqrt{1 \times 1}(a_+a_+)|0\rangle , \\ &= \sqrt{1 \times 1 \times 1}(a_+)|1\rangle , \\ &= \sqrt{1 \times 1 \times 1 \times 2}|2\rangle = \sqrt{2}|2\rangle . \end{aligned} \quad (22.50)$$

Noting that there are 4 operator product strings in $(a_+ + a_-)^4|0\rangle$ that have an overall imbalance of two between the a_+ s and the a_- s such that there are three a_+ s and only one a_- , there are 4 ways in which we can go from $|0\rangle$ to $|2\rangle$. These are the terms $a_+a_+a_-a_+$, $a_+a_+a_+a_-$, $a_+a_-a_+a_+$ and $a_-a_+a_+a_+$. While we have already seen the contribution from the first of these to be $\sqrt{2}$ just above, those from the rest are 0 (the rightmost a_- annihilates the ground state!), $2\sqrt{2}$ and $3\sqrt{2}$ respectively. Adding up the contribution from each of these gives a factor of $6\sqrt{2}$ as the coefficient. Thus, we obtain the $6\sqrt{2}$ coefficient for $|2\rangle$ in eq.(22.48). Importantly, as both the quartic potential perturbation and the harmonic potential within H_0 are invariant under a parity transformation, we expect the eigenstates of the complete problem to be parity eigenstates as well. It is then quite clear that, in leading to the perturbed ground state $|0\rangle'$, the even parity ground state $|0\rangle$ of H_0 can only mix with other even parity excited states such as $|2\rangle$ and $|4\rangle$. We will discuss this matter further in a following section.

Taking everything together gives the ground state energy of the problem to second order in the perturbation as

$$\begin{aligned} E_0(\lambda) &= \frac{\hbar\omega}{2} + \frac{3}{32}\lambda\hbar\omega - \frac{21\lambda^2\hbar\omega}{512} , \\ &= \left(1 + \frac{3\lambda}{16} - \frac{21\lambda^2}{256}\right)\frac{\hbar\omega}{2} + \mathcal{O}(\lambda^3) , \\ \implies E_0(\lambda = 1) &= \left(1 + \frac{3}{16} - \frac{21}{256}\right)\frac{\hbar\omega}{2} . \end{aligned} \quad (22.51)$$

We now proceed to compute the first order correction to the ground state wavefunction using

$$\begin{aligned}
|\psi_n^{(1)}\rangle &= \sum_{k \neq n} c_{nk}^{(1)} |\psi_k^{(0)}\rangle = \sum_{k \neq n} \frac{\langle \psi_k^{(0)} | V | \psi_n^{(0)} \rangle}{(E_n^{(0)} - E_k^{(0)})} |\psi_k^{(0)}\rangle , \\
\implies |0^{(1)}\rangle &= \frac{\lambda \hbar \omega}{32} \left[\frac{\langle 2 | (a_+ + a_-)^4 | 0 \rangle}{E_2^{(0)} - E_0^{(0)}} |2\rangle + \frac{\langle 4 | (a_+ + a_-)^4 | 0 \rangle}{E_4^{(0)} - E_0^{(0)}} |4\rangle \right] , \\
&= \frac{\lambda \hbar \omega}{32} \left[\frac{6\sqrt{2}}{2\hbar\omega} |2\rangle + \frac{\sqrt{4!}}{4\hbar\omega} |4\rangle \right] ,
\end{aligned} \tag{22.52}$$

rendering the mixing of $|0\rangle$ with $|2\rangle$ and $|4\rangle$ amply clear. Thus, we get the perturbed ground state ($|0\rangle'(\lambda)$) to first order in the perturbation as

$$\begin{aligned}
|0\rangle'(\lambda) &= |0\rangle + \lambda |0^{(1)}\rangle + \mathcal{O}(\lambda^2) , \\
&= |0\rangle + \frac{\lambda \hbar \omega}{32} \left[\frac{6\sqrt{2}}{2\hbar\omega} |2\rangle + \frac{\sqrt{4!}}{4\hbar\omega} |4\rangle \right] + \mathcal{O}(\lambda^2) , \\
\implies |0\rangle'(\lambda = 1) &= |0\rangle + \frac{\hbar \omega}{32} \left[\frac{6\sqrt{2}}{2\hbar\omega} |2\rangle + \frac{\sqrt{4!}}{4\hbar\omega} |4\rangle \right] .
\end{aligned} \tag{22.53}$$

It is clear that the second order correction to the ground state wavefunction will be somewhat more tedious to compute. We shall leave it to the interested reader as an exercise. However, it is possible to compute the wavefunction renormalisation factor without too much difficulty

$$\begin{aligned}
Z_0(\lambda) &= 1 - \lambda^2 \sum_{k \neq n} \frac{|\langle \psi_k^{(0)} | V | \psi_n^{(0)} \rangle|^2}{(E_n^{(0)} - E_k^{(0)})^2} + \mathcal{O}(\lambda^3) , \\
&= 1 - \left(\frac{\lambda \hbar \omega}{32} \right)^2 \left[\left(\frac{6\sqrt{2}}{2\hbar\omega} \right)^2 + \left(\frac{\sqrt{4!}}{4\hbar\omega} \right)^2 \right] + \mathcal{O}(\lambda^3) , \\
&= 1 - \frac{39\lambda^2}{2048} + \mathcal{O}(\lambda^3) .
\end{aligned} \tag{22.54}$$

A quick comparison with eq.(22.48) immediately reveals that the correction to the factor of 1 in Z_0 follows from precisely the same line of arguments that led to $E_0^{(2)}$, i.e., the same matrix elements and energy difference denominators are used in both $E_0^{(2)}$ and Z_0 (and one extra energy difference denominator appears within each term of Z_0). Thus, the normalised ground state is given by

$$\begin{aligned}
|0\rangle'_N(\lambda) &= Z_0^{1/2} |0\rangle'(\lambda) , \\
&= \left(1 - \frac{39\lambda^2}{2048} \right)^{1/2} \left(|0\rangle + \frac{(6|2\rangle + \sqrt{3}|4\rangle)\lambda}{32\sqrt{2}} \right) , \\
|0\rangle'_N(\lambda = 1) &= \left(1 - \frac{39}{2048} \right)^{1/2} \left(|0\rangle + \frac{(6|2\rangle + \sqrt{3}|4\rangle)}{32\sqrt{2}} \right) .
\end{aligned} \tag{22.55}$$

We stress that all the corrections we have computed thus far are very small for the ground state, and even for the case of $\lambda = 1$. This justifies our use of perturbation theory for understanding the ground state of this problem.

You may remain curious on how difficult this calculation can get for the n th eigenstate. Thus, we now compute the second order contribution to the n th eigenstate $|n\rangle$ by noting that the only non-vanishing matrix elements arising from the quartic perturbation are those that connect $|n\rangle$ with $|n \pm 2\rangle$ and $|n \pm 4\rangle$. We will first compute those contributions ($I_n^{(2)}$) that connect $|n\rangle$ with $|n + 2\rangle$ and $|n + 4\rangle$

$$\begin{aligned}
I_n^{(2)} &= -\left(\frac{\lambda\hbar\omega}{32}\right)^2 \left[\frac{|\langle n+2| (a_+ + a_-)^4 |n\rangle|^2}{E_{n+2}^{(0)} - E_n^{(0)}} + \frac{|\langle n+4| (a_+ + a_-)^4 |n\rangle|^2}{E_{n+4}^{(0)} - E_n^{(0)}} \right], \\
&= -\left(\frac{\lambda\hbar\omega}{32}\right)^2 \left[\frac{|\langle n+2| (a_+ a_+ a_- a_+ + a_- a_+ a_+ a_- + a_+ a_+ a_+ a_- + a_+ a_- a_+ a_-) |n\rangle|^2}{2\hbar\omega} \right] \\
&\quad - \left(\frac{\lambda\hbar\omega}{32}\right)^2 \left[\frac{|\langle n+4| a_+ a_+ a_+ a_- |n\rangle|^2}{4\hbar\omega} \right], \\
&= -\left(\frac{\lambda\hbar\omega}{32}\right)^2 \left[\frac{\left[(n+1+n+3+n+n+2)^2\sqrt{(n+1)(n+2)}\right]^2}{2\hbar\omega} \right] \\
&\quad - \left(\frac{\lambda\hbar\omega}{32}\right)^2 \left[\frac{(\sqrt{(n+1)(n+2)(n+3)(n+4)})^2}{4\hbar\omega} \right], \\
&= -\frac{\lambda^2\hbar\omega}{4096} [33n^4 + 202n^3 + 459n^2 + 458n + 168]. \tag{22.56}
\end{aligned}$$

We will now compute those contributions ($J_n^{(2)}$) that connect $|n\rangle$ with $|n-2\rangle$ and $|n-4\rangle$

$$\begin{aligned}
J_n^{(2)} &= -\left(\frac{\lambda\hbar\omega}{32}\right)^2 \left[\frac{|\langle n-2| (a_+ + a_-)^4 |n\rangle|^2}{E_{n-2}^{(0)} - E_n^{(0)}} + \frac{|\langle n-4| (a_+ + a_-)^4 |n\rangle|^2}{E_{n-4}^{(0)} - E_n^{(0)}} \right], \\
&= -\left(\frac{\lambda\hbar\omega}{32}\right)^2 \left[\frac{|\langle n-2| (a_- a_- a_+ a_- + a_+ a_- a_- a_- + a_- a_- a_+ a_- + a_- a_+ a_- a_-) |n\rangle|^2}{-2\hbar\omega} \right] \\
&\quad - \left(\frac{\lambda\hbar\omega}{32}\right)^2 \left[\frac{|\langle n-4| a_- a_- a_- a_- |n\rangle|^2}{-4\hbar\omega} \right], \\
&= -\left(\frac{\lambda\hbar\omega}{32}\right)^2 \left[\frac{\left[(n+n+1+n-1+n-2)^2\sqrt{n(n-1)}\right]^2}{-2\hbar\omega} \right] \\
&\quad - \left(\frac{\lambda\hbar\omega}{32}\right)^2 \left[\frac{(\sqrt{n(n-1)(n-2)(n-3)})^2}{-4\hbar\omega} \right], \\
&= +\frac{\lambda^2\hbar\omega}{4096} [33n^4 - 70n^3 + 51n^2 - 14n]. \tag{22.57}
\end{aligned}$$

Thus, the energy correction to the second order for the $|n\rangle$ is given by

$$\begin{aligned} E_n^{(2)}(\lambda) &= I_n^{(2)} + J_n^{(2)}, \\ &= -\frac{\lambda^2 \hbar \omega}{4096} [33n^4 + 202n^3 + 459n^2 + 458n + 168] + \frac{\lambda^2 \hbar \omega}{4096} [33n^4 - 70n^3 + 51n^2 - 14n], \\ &= -\frac{\lambda^2 \hbar \omega}{512} [34n^3 + 51n^2 + 59n + 21]. \end{aligned} \quad (22.58)$$

Note the exact cancellation of the $\mathcal{O}(n^4)$ term in the expression for $E_n^{(2)}(\lambda)$. Putting this together with the expression for $E_n^{(1)}(\lambda)$ obtained earlier gives the perturbed energy for the n th eigenstate (E_n) (including upto the second order perturbation corrections) as

$$\begin{aligned} E_n &= (n + \frac{1}{2})\hbar\omega + \frac{\lambda\hbar\omega}{32}(6n^2 + 6n + 3) \\ &\quad - \frac{\lambda^2 \hbar \omega}{512} [34n^3 + 51n^2 + 59n + 21] + \mathcal{O}(\lambda^3), \\ \implies E_n(\lambda = 1) &= (n + \frac{1}{2})\hbar\omega + \frac{\hbar\omega}{32}(6n^2 + 6n + 3) \\ &\quad - \frac{\hbar\omega}{512} [34n^3 + 51n^2 + 59n + 21] + \mathcal{O}(\lambda^3), \end{aligned} \quad (22.59)$$

and once more, it can be easily checked that the answer obtained from here for $|0\rangle$ is precisely that obtained above. Indeed, our expressions are precisely the equations 25 and 26 found in Adali & Nutku, Physics and Astronomy Reports (PAR) **1**, 93 (2023) (doi:10.26650/PAR.2023.00011, Turkish University Press), upto the difference that the authors of this work have assumed the coupling $g = 1$, while we have taken $g = m^2\omega^3/8\hbar$ here.

Thus, we are now in a position to compute the energy shift $\Delta_n = E_n - E_n^{(0)}$ to second order as

$$\begin{aligned} \Delta_n(\lambda) &= \frac{\lambda\hbar\omega}{32}(6n^2 + 6n + 3) - \frac{\lambda^2 \hbar \omega}{512} [34n^3 + 51n^2 + 59n + 21] + \mathcal{O}(\lambda^3), \\ \Delta_n(\lambda = 1) &= \frac{\hbar\omega}{32}(6n^2 + 6n + 3) - \frac{\hbar\omega}{512} [34n^3 + 51n^2 + 59n + 21] + \mathcal{O}(\lambda^3), \end{aligned} \quad (22.60)$$

such that the energy shift for the ground state is given by

$$\Delta_0(\lambda = 1) = \frac{3\hbar\omega}{32} - \frac{21\hbar\omega}{512} + \mathcal{O}(\lambda^3). \quad (22.61)$$

It is clear how brutal these calculations will become if we go to $\mathcal{O}(\lambda^3)$. It is also clear that Δ_n grows as $\mathcal{O}(n^3)$ (due to the second order corrections) and will become as big as n (in units of $\hbar\omega$) for even $|4\rangle$; note that this is already worse than the $|5\rangle$ threshold we found above from the first order corrections $E_n^{(1)}$. This reinforces our earlier conclusion that the quartic perturbation is best understood using perturbation theory for only the ground (and at most the first excited) state.

22.3 Selection Rules

Perturbation theory can be tedious. However, sometimes the labour involved can be lessened through the use of selection rules, such that we can be sure that certain matrix elements of the perturbation V are zero even without computing them explicitly. This is obtained from the following argument. If the perturbation V commutes with some operator Ω

$$[V, \Omega] = 0 , \quad (22.62)$$

then the action of V on the eigenstates of Ω , $\Omega |\omega\rangle = \omega |\omega\rangle$, is qualified

$$\langle \omega_2 | V | \omega_1 \rangle = 0 \text{ unless } \omega_1 = \omega_2 . \quad (22.63)$$

Then, acting on the eigenstate $|\omega\rangle$ with the operator combination ΩV gives

$$\Omega(V|\omega\rangle) = V(\Omega|\omega\rangle) = V\omega|\omega\rangle = \omega(V|\omega\rangle) , \quad (22.64)$$

such that the action of V on $|\omega\rangle$ has no effect with regards to the eigenvalue (ω) of the operator Ω . (Actually, we have seen this earlier during our discussion of symmetries.) To understand the origin of the selection rule given above, consider

$$\begin{aligned} 0 &= \langle \omega_2 | [V, \Omega] | \omega_1 \rangle , \\ &= \langle \omega_2 | V\Omega - \Omega V | \omega_1 \rangle , \\ &= (\omega_1 - \omega_2) \langle \omega_2 | V | \omega_1 \rangle . \end{aligned} \quad (22.65)$$

It is now clear that unless $\omega_1 = \omega_2$, $\langle \omega_2 | V | \omega_1 \rangle = 0$. Indeed, we can see that the selection rule simply reflects the orthogonality of eigenstates with different values of ω with regards to the action of the operator V , i.e., V cannot mix/cause a transition between two eigenstates with different values of ω as $V|\omega_1\rangle$ is orthogonal to $|\omega_2\rangle$. Thus, V can only mix states within the degenerate subspace of states labelled by $\omega_1 = \omega_2$, $|\omega_1\rangle$ and $|\omega_2\rangle = V|\omega_1\rangle$.

Let us consider a couple of simple examples. First, let us take a perturbation $V = \lambda z$, which is invariant under rotations around the z -axis, for a H_0 that corresponds to a system with finite orbital angular momentum (i.e., whose Hilbert space consist of angular momentum eigenstates). Clearly, this V commutes with the orbital angular momentum operator about the z -axis (L_z): $[L_z, \lambda z] = 0$. Further, for $L_z|m\rangle = m|m\rangle$, we have

$$\langle m_2 | \lambda z | m_1 \rangle = 0 \text{ unless } m_1 = m_2 . \quad (22.66)$$

Similarly, if the perturbation V is invariant under parity transformations, $V = \lambda z^2$ (i.e., clearly invariant under $z \rightarrow -z$), $[\hat{P}, V] = 0$ (where \hat{P} is the parity operator). Given that parity has only two eigenvalues, $P = \pm 1$, such a V cannot mix eigenstates of opposite parity, i.e., all matrix elements for V between eigenstates of opposite parity must vanish. This is another way to understand why the corrections to the ground state of the harmonic oscillator acquired corrections from the second and fourth excited states under perturbation from a quartic potential (which is invariant under parity; see previous section), but not the first and third; the former have the same (even) parity as that of the ground state, while the latter

have the opposite (odd) parity. Similarly, consider operators that are odd under parity (i.e., not parity invariant/ even), e.g., x and x^3 , such that $P^\dagger x P = -x$ and $P^\dagger x^3 P = -x^3$. Acting with x and x^3 on a state of definite parity clearly changes the parity of the state. Thus, the matrix element of x and x^3 between eigenstates of parity vanishes unless they have opposite parity; this is, for instance, easily seen to be the case when a cubic perturbation potential is applied to the one-dimensional simple harmonic oscillator problem.

These ideas can be extended to cases where the perturbation carries a definite amount of the eigenvalue of the operator Ω . Consider the fact that the operator z can change the orbital angular momentum eigenvalue l by ± 1 (as z does not commute with the L_x^2 and L_y^2 terms with the L^2 operator), but cannot change the value of the z component of the orbital angular momentum m (as z commutes with L_z). Similarly, x and y change l by ± 1 (each of them does not commute with two components of L^2) as well as change m by ± 1 (neither of them commutes with L_z). As orbital angular momentum eigenstates have definite parity ($(-1)^l$), the parity selection rule causes all matrix elements of operators such as x , y and z between eigenstates with the same l value to vanish. Hence, we can combine these conditions into:

$$\begin{aligned} \langle l_2, m_2 | z | l_1, m_1 \rangle &= 0 , \text{ unless } l_2 = l_1 \pm 1, m_2 = m_1 , \\ \langle l_2, m_2 | x | l_1, m_1 \rangle &= 0 = \langle l_2, m_2 | y | l_1, m_1 \rangle , \text{ unless } l_2 = l_1 \pm 1, m_2 = m_1 \pm 1 \end{aligned} \quad (22.67)$$

22.4 The Quadratic Stark Effect

Another example of nondegenerate perturbation theory is the so-called “quadratic Stark effect”. Consider the Hydrogen atom in its ground state, with a single electron in the state $|n = 1, l = 0, m = 0\rangle \otimes |\uparrow\rangle$; the Stark effect involves understanding the effects of an external electric field \vec{E} applied to this system. While we’ve chosen the up spin state, we could have equally well chosen the down state. What matters more is that it turns out that the spin degree of freedom is irrelevant to the problem. We can therefore safely ignore the spin degree of freedom of the electron, and treat the ground state as effectively non-degenerate. Note that even if we could ignore spin, the excited states of the Hydrogen atom (i.e., $n = 2$ onwards) have a degeneracy arising from the orbital angular momentum degree of freedom l . Thus, we cannot make the same assumption of non-degeneracy when treating the Stark effect for those states. For the ground state, however, non-degenerate perturbation theory is fine.

As always with problems treated in perturbation theory, we separate the Hamiltonian into two parts: the first (H_0) corresponding to the Hydrogen atom problem that we have already solved in an earlier chapter, and a second being the perturbation of an applied external electric field

$$V = -\vec{\mu}_e \cdot \vec{E} = -eEz \quad (22.68)$$

where $\vec{\mu}_e = -e\vec{r}$ corresponds to the electric dipole moment of the H-atom, and we have chosen the external electric field to be aligned along the z -axis $\vec{E} = E\hat{z}$. We assume that the eigenstates and eigenvalues of H_0 (i.e., the Hydrogen atom) are known exactly (in case you’ve forgotten the solution, please consult the chapter on the H-atom given earlier in these notes!). Recall that the ground state of the H-atom $|n = 1, l = 0, m = 0\rangle$ is spherically symmetric (we

drop the spin component henceforth as it is irrelevant) as it is parity even, and therefore must have a vanishing intrinsic permanent electric dipole moment: $\langle 1, 0, 0 | \vec{\mu} | 1, 0, 0 \rangle = 0$. Therefore, the energy shift to first order must also vanish:

$$E_{1,0,0}^{(1)} = \langle 1, 0, 0 | -\vec{\mu} \cdot \vec{E} | 1, 0, 0 \rangle = -\langle 1, 0, 0 | \vec{\mu} | 1, 0, 0 \rangle \cdot \vec{E} = 0. \quad (22.69)$$

Put another way, given that the ground state is effectively unique (as we have suppressed the spin degree of freedom) and V is parity odd, there is no degenerate eigenstate of opposite parity available which can contribute to the first order change in energy. Thus, the first order matrix element must vanish via the parity selection rule discussed earlier.

However, at second order, the electric field will *induce* a dipole moment (by distorting or “polarising” the initially spherically symmetric wavefunction in the direction of the electric field) $\vec{\mu}_{in} = \tilde{\alpha} \vec{E}$ and where $\tilde{\alpha}$ is the so-called polarisability tensor (a rank two symmetric tensor). A few quick words on tensors. Tensors are mathematical objects that can be used to describe physical properties, just like scalars (quantities that are pure numbers) and vectors (quantities that are a combination of numbers and directions). Indeed, tensors are a generalisation of scalars and vectors: a scalar is a rank zero tensor, a vector is a rank one tensor and a matrix is a rank two tensor. The rank (or order) of a tensor is defined by the number of directions (and hence the dimensionality of the array) required to describe it. For example, all properties of a system that require one direction (vector, and hence rank one) can be fully described by a 3×1 column vector, while properties that require atleast two directions (second rank tensors) can be described by 9 numbers and correspond to a 3×3 matrix. In general, a n th rank tensor can be described by $3n$ coefficients (in 3 spatial dimensions) and is clearly higher dimensional than the matrix! As seen in the example of the dipole moment induced by an externally applied electric field in 3 dimensions, the need for second rank tensors is clear: it arises as we need to consider more than one direction to describe the polarisability (a physical property!) along any of the three directions. Thus, for our present problem, we can visualise the 3×3 polarisability matrix as

$$\tilde{\alpha} = \begin{bmatrix} \tilde{\alpha}_{11} & \tilde{\alpha}_{12} & \tilde{\alpha}_{13} \\ \tilde{\alpha}_{21} & \tilde{\alpha}_{22} & \tilde{\alpha}_{23} \\ \tilde{\alpha}_{31} & \tilde{\alpha}_{32} & \tilde{\alpha}_{33} \end{bmatrix}, \quad (22.70)$$

such that the induced perturbation is given by

$$V = -\frac{1}{2} \vec{\mu}_{in} \cdot \vec{E} = -\frac{1}{2} E_i \tilde{\alpha}_{ij} E_j, \quad (22.71)$$

and the components of the induced dipole moment $\vec{\mu}_{in}$ can be explicitly written out as

$$\begin{aligned} \mu_1 &= \tilde{\alpha}_{11} E_1 + \tilde{\alpha}_{12} E_2 + \tilde{\alpha}_{13} E_3, \\ \mu_2 &= \tilde{\alpha}_{21} E_1 + \tilde{\alpha}_{22} E_2 + \tilde{\alpha}_{23} E_3, \\ \mu_3 &= \tilde{\alpha}_{31} E_1 + \tilde{\alpha}_{32} E_2 + \tilde{\alpha}_{33} E_3. \end{aligned} \quad (22.72)$$

It is now clear that $\tilde{\alpha}_{12}$, for instance, is the polarisation along the spatial direction 1 (the x-axis) due to a field along direction 2 (the y-axis) and so on. Given that the state $|1, 0, 0\rangle$ is

spherically symmetric, we expect the polarisability tensor $\tilde{\alpha}$ to be a symmetric matrix with $\tilde{\alpha}_{ij} = \tilde{\alpha}_{ji}$ for $j \neq i$ and $\tilde{\alpha}_{ii}$ a constant $\forall i$.

To proceed, our task is now to learn the elements within the induced polarisability tensor $\tilde{\alpha}$ and dipole moment $\vec{\mu}_{in}$. For this, we will proceed with computing the second order change in the ground state energy due to the perturbation $V = -eEz$

$$\begin{aligned} E_{1,0,0}^{(2)} &= -(eE)^2 \sum'_{n,l,m} \frac{|\langle nlm | z | 100 \rangle|^2}{E_{n,l,m}^{(0)} - E_{1,0,0}^{(0)}} , \\ &= -(eE)^2 \sum_{n=2}^{\infty} \frac{|\langle n10 | z | 100 \rangle|^2}{E_{n,1,0}^{(0)} - E_{1,0,0}^{(0)}} , \end{aligned} \quad (22.73)$$

where the ' on the sum means that the sum involves all excited states not excluded by selection rules, we have used the dipole selection rule (eq.(22.67)) given earlier to restrict the sum to only those excited states with $n \geq 1$, $l = 1$ and $m = 0$. Finally, note that

$$E_{n,l,m}^{(0)} - E_{1,0,0}^{(0)} = E_{Ry}(1 - \frac{1}{n^2}) = E_{Ry}(\frac{n^2 - 1}{n^2}) , \quad (22.74)$$

where the Rydberg energy is given by $E^{Ry} = \frac{\hbar^2}{2\mu a_0^2}$, the Bohr radius by $a_0 = \frac{4\pi\epsilon_0\hbar^2}{\mu e^2} = 5.3 \times 10^{-11}$ m , and μ is the reduced effective mass of the H-atom.

To proceed further, let us first take a step back and consider the general problem of evaluating the term

$$E_n^{(2)} = \sum_m' \frac{\langle n^{(0)} | V | m^{(0)} \rangle \langle m^{(0)} | V | n^{(0)} \rangle}{E_n^{(0)} - E_m^{(0)}} . \quad (22.75)$$

Had the energy difference denominator not been present, we could have used the completeness relation to carry out the summation (by adding and subtracting the contribution from the $m = n$ term). Indeed, there exists a way to eliminate the energy denominator such that this can be carried out (see A. Dalgarno and J. T. Lewis, Proc. Roy. Soc. **A233**, 70 (1955)). Suppose we identify an operator Ω such that $V = [\Omega, H_0]$. Then, we can use this relation

to simplify the expression for the second order correction as follows

$$\begin{aligned}
E_n^{(2)} &= \sum_m' \frac{\langle n^{(0)} | V | m^{(0)} \rangle \langle m^{(0)} | V | n^{(0)} \rangle}{E_n^{(0)} - E_m^{(0)}} , \\
&= \sum_m' \frac{\langle n^{(0)} | V | m^{(0)} \rangle \langle m^{(0)} | (\Omega H_0 - H_0 \Omega) | n^{(0)} \rangle}{E_n^{(0)} - E_m^{(0)}} , \\
&= \sum_m' \frac{\langle n^{(0)} | V | m^{(0)} \rangle (E_n^{(0)} - E_m^{(0)}) \langle m^{(0)} | \Omega | n^{(0)} \rangle}{E_n^{(0)} - E_m^{(0)}} , \\
&= \sum_m' \langle n^{(0)} | V | m^{(0)} \rangle \langle m^{(0)} | \Omega | n^{(0)} \rangle , \\
&= \sum_m \langle n^{(0)} | V | m^{(0)} \rangle \langle m^{(0)} | \Omega | n^{(0)} \rangle - \langle n^{(0)} | V | n^{(0)} \rangle \langle n^{(0)} | \Omega | n^{(0)} \rangle , \\
&= \langle n^{(0)} | V \Omega | n^{(0)} \rangle - \langle n^{(0)} | V | n^{(0)} \rangle \langle n^{(0)} | \Omega | n^{(0)} \rangle ,
\end{aligned} \tag{22.76}$$

where we have added and subtracted the $m = n$ term in the third from last line to obtain the second last line, and used the completeness relation for the sum over the full set of eigenstates $|m^{(0)}\rangle$, $m = 1, \dots, \infty$ to obtain the final line. The above expression needs only the evaluation of three matrix elements ($V\Omega$, V , and Ω) in the state $|n^{(0)}\rangle$.

It is, however, not an easy task for find the right Ω for any given problem. A more tractable problem, it turns out, is to find Ω such that

$$V |n^{(0)}\rangle = [\Omega, H_0] |n^{(0)}\rangle \tag{22.77}$$

for a given $|n^{(0)}\rangle$. Indeed, this is all that is needed in eq.(22.76). For the given problem, following the work by Dalgarno and Lewis mentioned above, we take

$$\Omega = \frac{ma_0eE}{\hbar^2} \left(\frac{rz}{2} + a_0z \right) , \tag{22.78}$$

It can be shown that

$$\begin{aligned}
[\Omega, H_0] &= \left[\Omega, \frac{\vec{p}^2}{2m} \right] , \\
&= \frac{ma_0eE}{2m\hbar^2} \left[\left(\frac{rz}{2} + a_0z \right), \vec{p}^2 \right] , \\
&= \frac{a_0eE}{2\hbar^2} \left[\left(\frac{rz}{2} + a_0z \right), \vec{p}^2 \right] , \\
\implies [\Omega, H_0] |1, 0, 0\rangle &= \frac{a_0eE}{2\hbar^2} \left[\left(\frac{rz}{2} + a_0z \right), \vec{p}^2 \right] |1, 0, 0\rangle , \\
&= -eEz |1, 0, 0\rangle \equiv V |1, 0, 0\rangle ,
\end{aligned} \tag{22.79}$$

where we have left the commutation relation in the second last step above as an exercise for the interested reader. Thus, noting that $\langle 1, 0, 0 | V | 1, 0, 0 \rangle = 0$ (as found earlier!), the second

term in eq.(22.76) vanishes and the problem at hand reduces to

$$\begin{aligned}
E_{1,0,0}^{(2)} &= \langle 1, 0, 0 | V\Omega | 1, 0, 0 \rangle , \\
&= -\frac{ma_0eE^2}{\hbar^2} \langle 1, 0, 0 | z(\frac{rz}{2} + a_0z) | 1, 0, 0 \rangle , \\
&= -\frac{ma_0eE^2}{2\hbar^2} \langle 1, 0, 0 | rz^2 | 1, 0, 0 \rangle - \frac{ma_0^2eE^2}{\hbar^2} \langle 1, 0, 0 | z^2 | 1, 0, 0 \rangle . \quad (22.80)
\end{aligned}$$

It is worth noting that the computation of potentially infinite number of matrix elements has been reduced to simply one! For a spherically symmetric system, the following ground state expectation values are obtained

$$\begin{aligned}
\langle x^2 \rangle &= \langle y^2 \rangle = \langle z^2 \rangle = \frac{1}{3} \langle r^2 \rangle , \\
\langle rz^2 \rangle &= \frac{1}{3} \langle r^3 \rangle \text{ and ,} \quad (22.81)
\end{aligned}$$

and we recall that $\langle r^2 \rangle \sim a_0^2$ and $\langle r^3 \rangle \sim a_0^3$. Computing $\langle r^2 \rangle$ and $\langle r^3 \rangle$ and putting them into the expression for $E_{1,0,0}^{(2)}$ above, we obtain

$$E_{1,0,0}^{(2)} = -\frac{9}{4}a_0^3E^2 , \quad (22.82)$$

and the induced dipole moment $\vec{\mu}_{in}$ defined earlier as

$$\vec{\mu}_{in} = -\frac{9}{2}a_0^3\vec{E} . \quad (22.83)$$

Thus, the diagonal component of the polarisability tensor ($\tilde{\alpha}$) is found to be

$$\tilde{\alpha} = \frac{9}{2}a_0^3 = 0.67 \times 10^{-33}m^3 , \quad (22.84)$$

which agrees quite well with experiments. Finding the correct Ω for the Dalgarno-Lewis method clearly requires ingenuity.

22.5 Sum Rules

We briefly investigate the concept of a sum rule, and display how the second order perturbation theory correction to the energy $E_n^{(2)}$ is related to the components of the wavefunction renormalisation Z_n through a sum rule. A well known sum rule in quantum mechanics is the Thomas-Reiche-Kuhn (TRK) sum rule:

$$\sum_{k \neq n} (E_k - E_n) |\langle k | x | n \rangle|^2 = \sum_{k \neq n} (E_k - E_n) \langle n | x | k \rangle \langle k | x | n \rangle = \frac{\hbar^2}{2m} , \quad (22.85)$$

where $|n\rangle$ and $|k\rangle$ are eigenstates of $H = p^2/2m + V(x)$. To see how this is obtained, we write

$$\begin{aligned}
\sum_{k \neq n} (E_k - E_n) |\langle k | x | n \rangle|^2 &= \sum_{k \neq n} (E_k - E_n) \langle n | x | k \rangle \langle k | x | n \rangle , \\
&= \frac{1}{2} \left(\sum_k \langle n | x | k \rangle \langle k | [Hx - xH] | n \rangle - \langle n | [Hx - xH] | k \rangle \langle k | x | n \rangle \right) , \\
&= \frac{1}{2} \left(\sum_k \langle n | x | k \rangle \langle k | [H, x] | n \rangle - \langle n | [H, x] | k \rangle \langle k | x | n \rangle \right) , \\
&= \frac{1}{2} \left(\sum_k \langle n | x | k \rangle \langle k | \left[\frac{p^2}{2m}, x \right] | n \rangle - \langle n | \left[\frac{p^2}{2m}, x \right] | k \rangle \langle k | x | n \rangle \right) , \\
&= \frac{1}{2} \left(\sum_k \langle n | x | k \rangle \langle k | \frac{-2i\hbar p}{2m} | n \rangle - \langle n | \frac{-2i\hbar p}{2m} | k \rangle \langle k | x | n \rangle \right) , \\
&= \frac{-i\hbar}{2m} \left(\sum_k \langle n | x | k \rangle \langle k | p | n \rangle - \langle n | p | k \rangle \langle k | x | n \rangle \right) , \\
&= \frac{-i\hbar}{2m} \langle n | (xp - px) | n \rangle = \frac{-i\hbar}{2m} \langle n | [x, p] | n \rangle , \\
&= \frac{\hbar^2}{2m} \quad (\text{as } [x, p] = i\hbar) . \tag{22.86}
\end{aligned}$$

We can test the TRK sum rule on $|n\rangle$ of the simple harmonic oscillator. Here, we see that

$$\begin{aligned}
\langle k | x | n \rangle &= \sqrt{\frac{\hbar}{2m\omega}} \langle k | (a_+ + a_-) | n \rangle = \sqrt{\frac{\hbar}{2m\omega}} \left(\sqrt{n+1} \delta_{k,n+1} + \sqrt{n} \delta_{k,n-1} \right) , \\
\implies |\langle k | x | n \rangle|^2 &= \frac{\hbar}{2m\omega} ((n+1)\delta_{k,n+1} + n\delta_{k,n-1}) , \tag{22.87}
\end{aligned}$$

such that

$$\begin{aligned}
\sum_k (E_k - E_n) |\langle k | x | n \rangle|^2 &= \sum_k (E_k - E_n) \frac{\hbar}{2m\omega} ((n+1)\delta_{k,n+1} + n\delta_{k,n-1}) , \\
&= \frac{\hbar}{2m\omega} (\hbar\omega(n+1) - \hbar\omega(n)) = \frac{\hbar^2}{2m} . \quad \text{QED} \tag{22.88}
\end{aligned}$$

What is the physical meaning of the TRK sum rule? If we define the dimensionless ‘spectral weight’ factor f_{nk} that quantifies the probability weight fraction carried by the excitation between $|n\rangle$ and $|k\rangle$ due to the perturbation x in terms of an energy-difference $(E_k - E_n)$ weighted sum of matrix elements $|\langle k | x | n \rangle|^2$

$$f_{nk} = \frac{2m}{\hbar^2} (E_k - E_n) |\langle k | x | n \rangle|^2 , \tag{22.89}$$

the TRK sum rule simply says

$$\sum_{k \neq n} f_{kn} = 1 , \tag{22.90}$$

i.e., the sum of all spectral (or, probability) weight fractions due to transitions from a nondegenerate state $|n\rangle$ to all other states under perturbation is unity. Strictly speaking, a definition in terms of a probability can only be carried out strictly for the initial state $|n\rangle$ being the ground state, such that $E_k - E_n > 0$ for $k \neq n$; this is not true for all transitions from any given excited state. We may, however, continue to refer to f_{nk} as a spectral weight fraction. Also, a sum over potentially infinite number of transition processes (i.e., between $|n\rangle$ and $|k\rangle$, $k \neq n$) give rise to a finite number ($\frac{\hbar^2}{2m}$).

Some more simple examples of sum rules can also be found as follows. Recall that for any one-dimensional system with a Hamiltonian of the type $H = p^2/2m + V(x)$, we have

$$[x, H] = \left[x, \frac{p^2}{2m} \right] = \frac{i\hbar p}{m} . \quad (22.91)$$

Recall the Ehrenfest relation

$$m \frac{d}{dt} (\langle n | x | n \rangle) = \frac{-im}{\hbar} [x, H] = \langle n | p | n \rangle , \quad (22.92)$$

showing us that, at the level of expectation values on the state $|n\rangle$, the product of the mass and the “velocity” ($m \langle n | (dx/dt) | n \rangle$) is given by the linear momentum ($\langle n | p | n \rangle$). Then, consider the following off-diagonal matrix element

$$\langle n | p | n' \rangle = \frac{im}{\hbar} \langle n | [H, x] | n' \rangle , \quad (22.93)$$

$$= \frac{im}{\hbar} (E_n - E_{n'}) \langle n | x | n' \rangle ,$$

$$\Rightarrow \langle n | p^2 | n \rangle = \sum_{n'} \langle n | p | n' \rangle \langle n' | p | n \rangle , \text{ introducing } 1 \equiv \sum_{n'} |n'\rangle \langle n'| \quad (22.94)$$

$$= \sum_{n'} \frac{im}{\hbar} (E_n - E_{n'}) \langle n | x | n' \rangle \times \frac{im}{\hbar} (E_{n'} - E_n) \langle n' | x | n \rangle ,$$

$$= \frac{m^2}{\hbar^2} \sum_{n'} |\langle n | x | n' \rangle|^2 (E_n - E_{n'})^2 . \quad (22.95)$$

Thus, we obtain another sum rule

$$\sum_{n'} |\langle n | x | n' \rangle|^2 (E_n - E_{n'})^2 = \frac{\hbar^2}{m^2} \langle n | p^2 | n \rangle = \frac{2\hbar^2}{m} \langle n | \hat{T} | n \rangle , \quad (22.96)$$

where $\hat{T} = p^2/2m$ is the kinetic energy operator. The above relation in eq.(22.96) gives us a sum rule that obtains the expectation value of the kinetic energy in the state $|n\rangle$ in terms of the product of matrix elements such as $|\langle n | x | n' \rangle|^2$ and the square of transition energies $(E_n - E_{n'})^2$.

Similarly, for the same one-dimensional Hamiltonian $H = p^2/2m + V(x)$, we have

$$[p, H] = [p, V(x)] = -i\hbar \frac{\partial V(x)}{\partial x} = -i\hbar V'(x) . \quad (22.97)$$

At this point, we recall the Ehrenfest relation

$$\frac{d}{dt}(\langle n| p |n\rangle) = \langle n| (-V'(x)) |n\rangle , \quad (22.98)$$

showing us that, at the level of expectation values on the state $|n\rangle$, the “effective force” ($\langle n| V'(x) |n\rangle$) is given by the rate of change of the linear momentum ($\langle n| p |n\rangle$). Now, consider the following off-diagonal matrix element

$$\begin{aligned} \langle n| V'(x) |n'\rangle &= \frac{i}{\hbar} \langle n| [p, H] |n'\rangle , \\ &= \frac{i}{\hbar} (E_{n'} - E_n) \langle n| p |n'\rangle , \\ \implies \langle n| (V'(x))^2 |n\rangle &= \sum_{n'} \langle n| V'(x) |n'\rangle \langle n'| V'(x) |n\rangle , \text{ introducing } 1 \equiv \sum_{n'} |n\rangle \langle n'| \\ &= \sum_{n'} \frac{i}{\hbar} (E_{n'} - E_n) \langle n| p |n'\rangle \times \frac{i}{\hbar} (E_n - E_{n'}) \langle n'| p |n\rangle , \\ &= \frac{1}{\hbar^2} \sum_{n'} |\langle n| p |n'\rangle|^2 (E_{n'} - E_n)^2 . \end{aligned} \quad (22.101)$$

Thus, we obtain another sum rule

$$\sum_{n'} |\langle n| p |n'\rangle|^2 (E_{n'} - E_n)^2 = \hbar^2 \langle n| (V'(x))^2 |n\rangle . \quad (22.102)$$

The above relation in eq.(22.102) gives us a sum rule that obtains the expectation value of the square of the “effective force” ($V'(x)$) in the state $|n\rangle$ in terms of the product of matrix elements such as $|\langle n| p |n'\rangle|^2$ and the square of transition energies $(E_n - E_{n'})^2$.

It is amusing to note that the above two sum rules for the diagonal matrix elements (expectation values) $\langle n| \hat{T} |n\rangle$ and $\langle n| (V'(x))^2 |n\rangle$ are observed to arise from taking summations of certain off-diagonal matrix elements (e.g., $|\langle n| x |n'\rangle|^2$ and $|\langle n| p |n'\rangle|^2$ respectively). Further, the ultimate source of these sum rules originates from the non-commutativity of the conjugate operators p and x (which transfers into the non-commutativity of the kinetic and potential energy parts of the Hamiltonian). We recall that the same non-commutativity has earlier been observed to give rise to the Heisenberg uncertainty principle (in terms of the uncertainties in expectation values $\langle n| p |n\rangle$ and $\langle n| x |n\rangle$). And, as observed above, the same non-commutativity gives rise to the Ehrenfest equations of motion for the expectation values $\langle n| p |n\rangle$ and $\langle n| x |n\rangle$.

Finally, we will see that we can see that the second order perturbation correction to the energy $E_n^{(2)}$ and the components of the wavefunction renormalisation Z_n can be related to one another through a similar sum rule. For this, we first define the wavefunction renormalisation factor as follows

$$\begin{aligned} Z_n &= 1 - \lambda^2 \sum_{k \neq n} \frac{|\langle \psi_k^{(0)} | V | \psi_n^{(0)} \rangle|^2}{(E_n^{(0)} - E_k^{(0)})^2} + \mathcal{O}(\lambda^3) , \\ &= 1 - \sum_{k \neq n} Z_{nk} , \end{aligned} \quad (22.103)$$

where Z_{nk} corresponds to the contribution to Z_n arising from the transition between $|n\rangle$ and $|k\rangle$

$$Z_{nk} = \lambda^2 \frac{|\langle \psi_k^{(0)} | V | \psi_n^{(0)} \rangle|^2}{(E_n^{(0)} - E_k^{(0)})^2} = \frac{\partial \Delta_{nk}^{(2)}}{\partial E_n^{(0)}} , \quad (22.104)$$

and $\Delta_{nk}^{(2)}$ is the second order energy shift of the state $|n\rangle$ due to the transition to $|k\rangle$ ($k \neq n$) . Now, we can write the second order correction to the energy as

$$\begin{aligned} E_n^{(2)} &= \lambda^2 \sum_{k \neq n} \frac{|(\langle \psi_k^{(0)} | V | \psi_n^{(0)} \rangle)|^2}{(E_n^{(0)} - E_k^{(0)})} , \\ &= \lambda^2 \sum_{k \neq n} (E_n^{(0)} - E_k^{(0)}) \frac{|(\langle \psi_k^{(0)} | V | \psi_n^{(0)} \rangle)|^2}{(E_n^{(0)} - E_k^{(0)})^2} , \\ &= \sum_{k \neq n} (E_n^{(0)} - E_k^{(0)}) \frac{\partial \Delta_{nk}^{(2)}}{\partial E_n^{(0)}} , \\ \implies E_n^{(2)} &= \sum_{k \neq n} (E_n^{(0)} - E_k^{(0)}) Z_{nk} , \\ \implies 1 &= \sum_{k \neq n} \frac{(E_n^{(0)} - E_k^{(0)})}{E_n^{(2)}} Z_{nk} , \\ \text{or, } 1 &= \sum_{k \neq n} f_{nk} , \end{aligned} \quad (22.105)$$

where

$$f_{nk} = \frac{(E_n^{(0)} - E_k^{(0)})}{E_n^{(2)}} Z_{nk} . \quad (22.106)$$

Thus, we see that the dimensionless quantity f_{nk} corresponds to the appropriately defined spectral weight fraction for transitions between $|n\rangle$ and $|k\rangle$ that leads to the second order perturbation correction $E_n^{(2)}$, i.e., f_{nk} is a product of the energy fraction $\frac{(E_n^{(0)} - E_k^{(0)})}{E_n^{(2)}}$ weighted by the appropriate wavefunction renormalisation fraction Z_{nk} . As we can see, with this definition for f_{nk} , we obtain the same sum rule as TRK, i.e., $\sum_{k \neq n} f_{nk} = 1$. As with the TRK, the remarkable fact arising from this sum rule is precisely how a sum over potentially infinite number of transition processes (i.e., between $|n\rangle$ and $|k\rangle$, $(k \neq n)$) give rise to a finite number ($E_n^{(2)}$). Given that these sum rules are governed by unitarity (a non-perturbative statement on the conservation of probabilities), one suspects the existence of similar relations at higher orders as well; it would be interesting to investigate whether they exist. We end by noting that similar sum rules, by the name of the f -sum rule and the Friedel sum rule, do appear in the physics of interacting many-particle quantum systems when studied via advanced formalisms such as quantum field theory and many-body theory.

22.6 Degenerate Perturbation theory

When there are degenerate subspaces in the Hilbert space, the perturbation theory formalism we set up previously will obviously fail, as certain energy denominators ($E_k - E_n$) are going to diverge. The new strategy to be adopted is as follows

- Identify the degenerate unperturbed subspace, say g -fold, and construct the perturbation matrix V in that basis (corresponding to a $g \times g$ matrix). For this, we note that there must exist some other operator \hat{A} (corresponding to certain symmetry of the full H) such that

$$[H_0, \hat{A}] = 0 , \quad (22.107)$$

such that the eigenvalues of \hat{A} are also good quantum numbers. Given that V does not commute with H_0 (which is why we need to do the perturbation theory in the first place!), the action of V must first be treated within the degenerate subspace of the eigenstates of H_0 (which are labelled by different eigenvalues of \hat{A}).

- Diagonalise the $g \times g$ perturbation matrix in the g -fold degenerate subspace by solving the appropriate secular equation. This will yield the correct zeroth order nondegenerate eigenbasis for the problem, as the diagonalisation procedure will involve the lifting of all degeneracies by the perturbation V . The non-degenerate eigenstates obtained from here will be linear superpositions of the (erstwhile degenerate) eigenstates of H_0 and \hat{A} ; this is the way in which the diagonalisation procedure mixes the degenerate states of H_0 . Diagonalisation of V means that in the new basis, all off-diagonal matrix elements of V must vanish, $\langle n' | V | n \rangle = 0$ ($n' \neq n$) .
- Identify the eigenvalues of the diagonalised perturbation matrix as the first order energy shifts.
- Use the basis that diagonalises the perturbation matrix within the standard nondegenerate perturbation theory formalism to obtain higher order corrections. Be careful to exclude all contributions from new unperturbed states within the diagonal basis from all summations in the nondegenerate perturbation theory formalism.

A few points are worth noting. First, for all practical purposes, the degenerate subspaces are usually finite sized, and therefore should be a tractable problem (its hard to diagonalise big matrices in general!). Second, the diagonalisation procedure at the heart of degenerate perturbation theory is formally exact and, unlike non-degenerate perturbation theory, does not involve any expansion in terms of a small parameter. Third, the diagonalisation process essentially involves finding a unitary transformation (\tilde{U}) for the degenerate subspace such that $\tilde{H} = \tilde{U}^\dagger (H_0 + V) \tilde{U}$ takes a diagonal form within that subspace.

With these steps laid out, we first carry out a transformation (basis rotation) on the n -fold degenerate subspace

$$|n^{(0)}\rangle \rightarrow |n'^{(0)}\rangle = \sum_{m=1}^n \alpha_{nm} |m^{(0)}\rangle , \quad (22.108)$$

such that we diagonalise the perturbation matrix

$$\left\langle m'^{(0)} \middle| V \middle| n'^{(0)} \right\rangle = 0 \text{ if } m \neq n . \quad (22.109)$$

Then,

$$\begin{aligned} E_n^{(1)} &= \left\langle n'^{(0)} \middle| V \middle| n'^{(0)} \right\rangle , \\ E_n^{(2)} &= \sum_{p \neq m,n} \frac{|\left\langle p'^{(0)} \middle| V \middle| n'^{(0)} \right\rangle|^2}{E_n^{(0)} - E_p^{(0)}} . \end{aligned} \quad (22.110)$$

We have already seen an example of degenerate perturbation theory in dealing with the problem of two spin-1/2 degrees of freedom interacting with one another in an earlier chapter. We will now see another example in the form of the linear Stark effect.

22.7 Linear Stark effect: atom with a permanent intrinsic dipole moment

As mentioned earlier, unlike the ground state ($|n = 1, l = 0, m = 0\rangle$), the excited states of the H-atom can well have an intrinsic permanent dipole moment. However, the excited states of the H-atom have a non-trivial degeneracy arising from both orbital and spin angular momenta. For the case of the Stark effect, $V = -ezE$; while we can ignore the spin degree of freedom altogether, we must address the degeneracy arising from the orbital degrees of freedom. Specifically, the L^2 and L_z operators label the degenerate subspace: for $n = 2$, we have an $l = 0$ ($m = 0$) state ($2s$) and three $l = 1$ ($m = 0, \pm 1$) states ($2p$) all of whom have the same energy ($-e^2/8a_0$, where a_0 is the Bohr radius). Given that the perturbation V at hand causes transitions between states with different parity (and for $l_2 = l_1 \pm 1$), here they are between the $l = 0$ and $l = 1$ states. However, as z cannot change the m quantum number (which makes the L_z angular momentum operator the operator that commutes with V and H_0), the only states that this perturbation can mix are the $|n = 2, l = 0, m = 0\rangle$ and $|n = 2, l = 1, m = 0\rangle$ states:

$$\langle 2, 0, 0 | V | 2, 1, 0 \rangle = \langle 2, 1, 0 | V | 2, 0, 0 \rangle = 3ea_0E . \quad (22.111)$$

To see how this is obtained, we note that

$$\begin{aligned} \psi_{2,0,0} = \langle r, \theta, \phi | 2, 0, 0 \rangle &= \sqrt{\frac{1}{2\pi a_0}} \frac{1}{2a_0} \left(1 - \frac{r}{2a_0}\right) e^{-r/2a_0} , \\ \psi_{2,1,0} = \langle r, \theta, \phi | 2, 1, 0 \rangle &= \sqrt{\frac{1}{2\pi a_0}} \frac{r \cos \theta}{4a_0^2} e^{-r/2a_0} . \end{aligned} \quad (22.112)$$

Now, the matrix element above can be computed by noting that $z = r \cos \theta$ and

$$\begin{aligned}
\langle 2, 0, 0 | V | 2, 1, 0 \rangle &= -eE \langle 2, 0, 0 | r \cos \theta | 2, 1, 0 \rangle , \\
&= -eE \int d\Omega \int dr \frac{1}{16\pi a_0^4} r^4 \cos^2 \theta (1 - \frac{r}{2a_0}) e^{-r/a_0} , \\
&= -eE \frac{2\pi}{16\pi a_0^4} \left(\int_{-1}^1 d(\cos \theta) \cos^2 \theta \right) \left(\int_0^\infty dr r^4 (1 - \frac{r}{2a_0}) e^{-r/a_0} \right) , \\
&= -\frac{ea_0 E}{12} \left(\int_0^\infty d(\frac{r}{a_0}) (\frac{r}{a_0})^4 (1 - \frac{r}{2a_0}) e^{-r/a_0} \right) , \\
&= -\frac{ea_0 E}{12} \left(\Gamma(5) - \frac{\Gamma(6)}{2} \right) , \\
&= 3ea_0 E ,
\end{aligned} \tag{22.113}$$

where we have used the fact that $\Gamma(n) = \int_0^\infty dt t^{n-1} e^{-t} = (n-1)!$ to write $\Gamma(5) - \frac{\Gamma(6)}{2} = 24 - 60 = -36$. Thus, the zeroth problem to solve lies in the degenerate subspace of $|2, 0, 0\rangle$ and $|2, 1, 0\rangle$

$$\begin{aligned}
V &= \begin{bmatrix} 0 & \langle 2, 0, 0 | V | 2, 1, 0 \rangle \\ \langle 2, 1, 0 | V | 2, 0, 0 \rangle & 0 \end{bmatrix} , \\
&= \begin{bmatrix} 0 & 3ea_0 E \\ 3ea_0 E & 0 \end{bmatrix} = 3ea_0 E \sigma_x .
\end{aligned} \tag{22.114}$$

The diagonal eigenbasis is simply obtained as

$$|\pm\rangle = \frac{1}{\sqrt{2}}(|2, 0, 0\rangle \pm |2, 1, 0\rangle) , \tag{22.115}$$

i.e., singlet-like and triplet-zero like combinations of the states $|2, 0, 0\rangle$ and $|2, 1, 0\rangle$. The corresponding energies (upon including the first order shifts) are

$$\begin{aligned}
E_+ &= -\frac{e^2}{8a_0} - 3ea_0 E , \\
E_- &= -\frac{e^2}{8a_0} + 3ea_0 E .
\end{aligned} \tag{22.116}$$

The energy shifts $E_\pm - E_0$ being linear in E , we call this the linear Stark effect. Thus, the eigenstates in the diagonal basis $|\pm\rangle$ are not eigenstates of parity, and we learn from the energy shift that they have an intrinsic permanent dipole moment given by $\mu_e = 3ea_0$. It is clear that we treat the effects of the perturbation potential within this degenerate subspace exactly during the diagonalisation; the first order shifts given above are therefore exact in nature. The diagonalisation corresponds to a rotation about $-\hat{j}$ by $\pi/2$ (a unitary transformation!). Finally, remember that the two states $|2, 1, \pm 1\rangle$ are stable under the perturbation V and remain degenerate at energy $-e^2/8a_0$; they must be excluded from any further computations of higher order corrections to the energy E_2 under non-degenerate perturbation theory. Also, as pointed out earlier, we had ignored the spin magnetic moment of the electron in our discussion of the linear and quadratic Stark effects. In what follows, we will take the spin moment into account and find that there are subtle perturbative corrections to the energies of the Hydrogen atom arising from it.

22.8 Fine structure and Spin-Orbit Coupling

Our earlier study of the H-atom involved accounting for the Coulomb attraction between the positively charged nucleus and the negatively charged electron orbiting the nucleus. However, it turns out that this is not the only interaction that must be accounted for (though it is the most important!): there are very small corrections of order α^2 with regards to the Bohr energies that are called “fine structure” (where α is the so-called fine structure constant, $\alpha = e^2/(4\pi\epsilon_0\hbar c) \simeq 1/137$). One of the sources of this fine structure is spin-orbit coupling. This arises as follows. In the reference frame of the electron orbiting the nucleus, the positively charged proton is circling around it. This orbiting positive charge creates an effective internal magnetic field (\vec{B}_{int}) in the electron’s rest frame, and this magnetic field exerts a torque on the electron and tries to align its magnetic moment ($\vec{\Sigma}_e$) along the direction of the field

$$H_{so}^0 = -\vec{\Sigma}_e \cdot \vec{B}_{int} . \quad (22.117)$$

To determine B_{int} , we use the Biot-Savart law $B_{int} = \tilde{\mu}_0 I / 2r$ for an effective current $I = e/T$, $\tilde{\mu}_0$ is the permeability of the vacuum and T is the orbit’s time period. Now, we recall that T can be obtained from the fact that the orbital angular momentum of the electron (in the rest frame of the nucleus) is given by $L = mvr = mr(\times 2\pi r/T) = 2\pi mr^2/T$. Eliminating T in favour of L , and noting that \vec{B}_{int} and \vec{L} point in the same direction (say, up), we obtain

$$\vec{B}_{int} = \frac{e}{4\pi\epsilon_0 mc^2 r^3} \vec{L} , \quad (22.118)$$

where we have used $c^2 = 1/\epsilon_0\tilde{\mu}_0$ to replace $\tilde{\mu}_0$ by ϵ_0 .

The magnetic moment of the electron is given by

$$\vec{\Sigma}_e = -g\mu_0 \vec{S} = -\frac{ge}{2m} \vec{S} = -\frac{ge\hbar}{4m} \vec{\sigma} , \quad (22.119)$$

where the classical magnetic moment $\mu_0 = e/2m$, $g = 2$ is the Landé g -factor of the electron (as obtained from Dirac’s relativistic quantum mechanical theory for the electron) and $\vec{\sigma}$ denote the Pauli representation of the spin $\vec{S} = \frac{\hbar}{2}\vec{\sigma}$. There is now one more fact that we must account for before obtain the spin-orbit Hamiltonian. Recall that we worked in the rest frame of the electron, but this is not an inertial frame as the electron is accelerating as it orbits the electron, i.e., it is continually moving from one inertial system to another. The cumulative effect of all the associated Lorentz transformations involved in moving between the inertial frames of reference can be take into account, and is known as the Thomas precession factor (equal to $1/2$, see L. H. Thomas, Nature **117**, 574 (1926)) that must be multiplied to the Hamiltonian H_{so}^0 to obtain the final spin-orbit Hamiltonian

$$\begin{aligned} H_{so} &= \frac{1}{2} H_{so}^0 = \frac{1}{2} \frac{e^2}{4\pi\epsilon_0 m^2 c^2 r^3} \frac{g}{\hbar} \vec{S} \cdot \vec{L} , \\ &= \frac{e^2}{4\pi\epsilon_0 m^2 c^2 r^3} \vec{S} \cdot \vec{L} , \end{aligned} \quad (22.120)$$

where the Landé g -factor and the correction factor from the Thomas precession are observed to cancel one another.

In taking care of this spin-orbit coupling, we see complete Hamiltonian (i.e., that of the H-atom obtained earlier together with H_{so} from above) no longer commutes with \vec{L} and \vec{S} : neither the orbital nor the spin angular momenta are separately conserved. However, H_{so} commutes with L^2 , S^2 , the total angular momentum (which we have encountered in earlier chapters) $\vec{J} = \vec{L} + \vec{S}$ and its z -component J_z , which means that it is these four quantities that are conserved. This means that we can now label the eigenstates by the new quantum numbers $|n, l^2, s^2, j^2, j_z\rangle$, and use these in perturbation theory. Further,

$$\begin{aligned}\vec{J}^2 &= (\vec{L} + \vec{S}) \cdot (\vec{L} + \vec{S}) = L^2 + S^2 + 2\vec{L} \cdot \vec{S}, \\ \implies \vec{L} \cdot \vec{S} &= \frac{1}{2}(J^2 - L^2 - S^2), \\ \implies H_{so} &= \frac{e^2}{4\pi\epsilon_0 m^2 c^2 r^3} \vec{L} \cdot \vec{S} = \frac{e^2}{8\pi\epsilon_0 m^2 c^2 r^3} (J^2 - L^2 - S^2),\end{aligned}\quad (22.121)$$

which tells us immediately that the eigenvalues of $\vec{L} \cdot \vec{S}$ are

$$\frac{\hbar^2}{2} [j(j+1) - l(l+1) - s(s+1)]. \quad (22.122)$$

Thus, the first order correction that arises due to the perturbation H_{so} in the eigenbasis $|n, j, m; l, s = 1/2\rangle$ is given by

$$\begin{aligned}E_{so}^{(1)} &= \langle n', j', m'; l', 1/2 | H_{so} | n, j, m; l, s = 1/2 \rangle \\ &= \delta_{nn'} \delta_{jj'} \delta_{mm'} \delta_{ll'} \frac{e^2 \hbar^2}{8\pi\epsilon_0 m^2 c^2} (j(j+1) - l(l+1) - \frac{3}{4}) \langle \frac{1}{r^3} \rangle_{nl}, \\ &= \frac{e^2 \hbar^2}{8\pi\epsilon_0 m^2 c^2} (j(j+1) - l(l+1) - \frac{3}{4}) \frac{1}{a_0^3 n^3 l(l+1/2)(l+1)}, \\ &= \frac{e^2 \hbar^2}{8\pi\epsilon_0 m^2 c^2 a_0^3 n^3} \frac{j(j+1) - l(l+1) - \frac{3}{4}}{l(l+1/2)(l+1)}, \\ &= \frac{13.6 \text{ eV}}{n^2} \frac{\alpha^2}{n} \frac{j(j+1) - l(l+1) - \frac{3}{4}}{l(l+1/2)(l+1)}, \\ &= -E_n \frac{\alpha^2}{n} \frac{j(j+1) - l(l+1) - \frac{3}{4}}{l(l+1/2)(l+1)},\end{aligned}\quad (22.123)$$

where we have used the expressions for α , a_0 (the Bohr radius), the fact that $\hbar^2/(2ma_0^2) = 13.6 \text{ eV}$, the zeroth spectrum of the H-atom is given by $E_n = -\frac{13.6 \text{ eV}}{n^2}$ and the result that $\langle r^{-3} \rangle_{nl} = (a_0^3 n^3 l(l+1/2)(l+1))^{-1}$ (see Shankar, Ex.17.3.4). Thus, as mentioned at the beginning of this section, we can see that the correction due to spin-orbit coupling is depressed by a factor of α^2 . Finally, we see that the first order corrected energy of the H-atom is

$$E_n = -\frac{13.6 \text{ eV}}{n^2} \left[1 - \frac{\alpha^2}{n} \frac{j(j+1) - l(l+1) - \frac{3}{4}}{l(l+1/2)(l+1)} \right]. \quad (22.124)$$

Note that the correction we have obtained from H_{so} is well defined only for the excited states $n \neq 1, l = 0$, as $\langle r^{-3} \rangle$ diverges and $\langle \vec{L} \cdot \vec{S} \rangle$ vanishes for $n = 1, l = 0$. However, it turns out that there is a fine structure correction to the ground state that arises from a relativistic treatment via the Dirac equation. We will not treat this case here.

22.9 The Zeeman Effect

We will now consider the physics of a H-atom in a weak uniform magnetic field along the z -axis given by

$$\vec{B}_{ext} = B\hat{z} = \vec{\nabla} \times \vec{A} , \text{ where } \vec{A} = \frac{B}{2}(y\hat{x} - x\hat{y}) . \quad (22.125)$$

The Zeeman Hamiltonian perturbation is then given by

$$H_Z = -(\vec{\mu}_L + \vec{\mu}_S) \cdot \vec{B}_{ext} = \mu_0(\vec{L} + g\vec{S}) \cdot \vec{B}_{ext} = \mu_0(\vec{L} + 2\vec{S}) \cdot \vec{B}_{ext} , \quad (22.126)$$

where $\mu_0 = e/2m$, $\vec{\mu}_L = -\mu_0\vec{L}$, $\vec{\mu}_S = -g\mu_0\vec{S}$ and $g \equiv 2$ is the Lande g -factor. The effects of the Zeeman splitting depends critically on how the external field $|B_{ext}| \equiv B$ compares with the internal field B_{int} that gives rise to spin-orbit coupling. For the case of the weak Zeeman effect, $B \ll B_{int}$: the spin-orbit coupling dominates and the associated (fine structure) Hamiltonian (H_{so}) must be treated as part of the zeroth Hamiltonian H_0 whose eigenstates and energy eigenvalues must first be determined prior to a perturbation theoretic analysis of the effects of the external field B in the Zeeman Hamiltonian H_Z . On the other hand, for $B \gg B_{int}$, we are in the regime of the strong Zeeman effect. Here, we must treat the Zeeman term H_Z as part of H_0 and the spin-orbit coupling H_{so} as the perturbation. Finally, when $B \sim B_{int}$, both H_Z and H_{so} must be treated on par and degenerate perturbation theory must be employed. Below, we will focus on only the case of the weak Zeeman effect. The interested reader is requested to consult the excellent texts by Griffiths and Sakurai for details on the other two cases.

22.9.1 Weak Zeeman effect

As mentioned above, in this regime $B \ll B_{int}$, and we treat H_{so} within H_0 first in order to learn the eigenstates. This was done in the previous section on spin-orbit coupling, and we learnt that the relevant states are $|n, j, m; l, s = 1/2\rangle$. Thus, the first order correction due to H_Z is

$$\begin{aligned} E_Z^{(1)} &= \langle n', j', m'; l', 1/2 | H_Z | n, j, m; l, s = 1/2 \rangle \\ &= \mu_0 B \langle n', j', m'; l', 1/2 | (\vec{L} + 2\vec{S}) | n, j, m; l, s = 1/2 \rangle , \\ &= \mu_0 B \langle n', j', m'; l', 1/2 | (\vec{J} + \vec{S}) | n, j, m; l, s = 1/2 \rangle , \end{aligned} \quad (22.127)$$

where we have used the fact that $\vec{L} + 2\vec{S} = \vec{J} + \vec{S}$. While \vec{J} is conserved in time, \vec{S} is not; indeed, both \vec{L} and \vec{S} precess about \vec{J} . Thus, the expectation value of \vec{S} can be obtained

from the expectation value of its projection onto \vec{J}

$$\begin{aligned}
\langle \vec{S} \rangle &= \langle \frac{\vec{S} \cdot \vec{J}}{J^2} \vec{J} \rangle , \\
\implies E_Z^{(1)} = \langle H_z \rangle &= \mu_0 B \langle \vec{J} + \vec{S} \rangle , \\
&= \mu_0 B \langle \left(1 + \frac{\vec{S} \cdot \vec{J}}{J^2} \vec{J} \right) \rangle , \\
&= \mu_0 B \left(1 + \frac{\langle \vec{S} \cdot \vec{J} \rangle}{j(j+1)\hbar^2} \right) \langle \vec{J} \rangle , \\
&= \mu_0 B \left(1 + \frac{\langle \vec{S} \cdot \vec{J} \rangle}{j(j+1)\hbar^2} \right) m\hbar ,
\end{aligned} \tag{22.128}$$

as $\langle \vec{J} \rangle = m\hbar$ for the external field being aligned along the z -axis. Now, recall that

$$\begin{aligned}
\vec{L} &= \vec{J} - \vec{S} \implies L^2 = J^2 + S^2 - 2\vec{J} \cdot \vec{S} , \\
\text{or } \vec{S} \cdot \vec{J} &= \frac{1}{2}(J^2 + S^2 - L^2) , \\
\implies \langle \vec{S} \cdot \vec{J} \rangle &= \frac{\hbar^2}{2}(j(j+1) + s(s+1) - l(l+1)) = \frac{\hbar^2}{2}(j(j+1) - l(l+1) + \frac{3}{4}) ,
\end{aligned} \tag{22.129}$$

Substituting the value of $\langle \vec{S} \cdot \vec{J} \rangle$ into the expression for $E_Z^{(1)}$ above, we obtain

$$E_Z^{(1)} = \left(1 + \frac{j(j+1) - l(l+1) + \frac{3}{4}}{2j(j+1)} \right) \mu_B B m = g_J \mu_B B m , \tag{22.130}$$

where the Bohr magneton $\mu_B = \mu_0 \hbar = \frac{e\hbar}{2m} = 5.788 \times 10^{-5}$ eV/T and $g_J = \left(1 + \frac{j(j+1) - l(l+1) + \frac{3}{4}}{2j(j+1)} \right)$ is the effective Landé g -factor. The total energy is the sum of the contributions obtained from the spin-orbit coupling above and the Zeeman coupling. For the ground state $n = 1, l = 0, j = 1/2, m = \pm 1/2$ and $g_J = 1 + \frac{2 \times 3/4}{2 \times 3/4} = 2$, the Zeeman term splits (i.e., lifts the degeneracy in m) to give

$$E_0 = -13.6 \text{ eV} \left(1 + \frac{\alpha^2}{4} \right) \pm \frac{g_J \mu_B B}{2} = -13.6 \text{ eV} \left(1 + \frac{\alpha^2}{4} \right) \pm \mu_B B , \tag{22.131}$$

where the $-13.6 \text{ eV} \frac{\alpha^2}{4}$ correction to E_0 is of relativistic origin (as was alluded to at the end of the previous section). The negative correction from the Zeeman term arises from the $m = -1/2$ state, while the positive correction arises from the $m = 1/2$ state.

22.10 Hyperfine splitting

This is an order $\frac{m_e}{m_p} \alpha^2$ correction to the ground state energy (E_1) of the H-atom, and arises from an interaction between the magnetic dipole moments of the electron and proton

$$\mu_e = -g_e \mu_0 \vec{S}_e , \quad \mu_p = -g_p \mu_0 \vec{S}_p , \tag{22.132}$$

where the Landé g -factors for the electron and proton are $g_e = 2$ and $g_p = 5.59$. The factor of m_e/m_p leads to a huge suppression of this correction to E_1 compared to the correction arising from the fine structure (i.e., spin-orbit coupling) term ($\sim \alpha^2$). This is why this correction factor is called the “hyperfine” splitting. According to classical electrodynamics, the magnetic field created by a dipole $\vec{\mu}$ is given by

$$\vec{B}_{dip} = \frac{\tilde{\mu}_0}{4\pi r^3} [3(\vec{\mu} \cdot \hat{r})\hat{r} - \vec{\mu}] + \frac{2\tilde{\mu}_0}{3}\vec{\mu}\delta^3(\vec{r}), \quad (22.133)$$

where $\tilde{\mu}_0$ is the permeability of the vacuum. For further insight into the origin of the second term in the above expression, i.e., with the $\delta^3(\vec{r})$, note that it has an influence only at the origin (see D. J. Griffiths, Am. J. Phys. **50**, 698 (1982)). Thus, the “hyperfine” Hamiltonian for the electron’s magnetic dipole moment interacting with \vec{B}_{dip} is given by

$$H_{hyp} = \frac{\tilde{\mu}_0 g_p g_e \mu_0^2}{8\pi} \frac{[(\vec{S}_e \cdot \hat{r})(\vec{S}_p \cdot \hat{r}) - \vec{S}_e \cdot \vec{S}_p]}{r^3} + \frac{2\tilde{\mu}_0 g_p g_e \mu_0^2}{3} \vec{S}_e \cdot \vec{S}_p \delta^3(\vec{r}). \quad (22.134)$$

Thus, the first order correction due to the hyperfine perturbation is given by

$$E_{hyp}^{(1)} = \frac{\tilde{\mu}_0 g_p g_e \mu_0^2}{8\pi} \left\langle \frac{[(\vec{S}_e \cdot \hat{r})(\vec{S}_p \cdot \hat{r}) - \vec{S}_e \cdot \vec{S}_p]}{r^3} \right\rangle + \frac{2\tilde{\mu}_0 g_p g_e \mu_0^2}{3} \langle \vec{S}_e \cdot \vec{S}_p \rangle |\psi_{n,l,m}(\vec{r}=0)|^2 \quad (22.135)$$

As the ground state of the H-atom (as well as for any state with $l = 0$) is spherically symmetric, it can be shown that the first term in the expression for $E_{hyp}^{(1)}$ vanishes identically. Further, using $|\psi_{1,0,0}|^2 = 1/(\pi a_0^3)$, we obtain $E_{hyp}^{(1)}(n=0)$ as

$$\begin{aligned} E_{hyp}^{(1)}(n=0) &= \frac{2\tilde{\mu}_0 g_p g_e \mu_0^2}{3\pi a_0^3} \langle \vec{S}_e \cdot \vec{S}_p \rangle, \\ &= \frac{2\tilde{\mu}_0 g_p g_e \mu_0^2}{3\pi a_0^3} \frac{1}{2} \langle (S_{Tot}^2 - S_e^2 - S_p^2) \rangle. \end{aligned} \quad (22.136)$$

Noting that both the electron and proton have spin-1/2, $S_e^2 = (3/4)\hbar^2 = S_p^2$. Adding the two spin-1/2 degrees of freedom gives rise to 3 triplet states with $S_{Tot} = 1$ (and $S_{Tot}^z = 0, \pm 1$) and a singlet state with $S_{Tot} = 0$ (and $S_{Tot}^z = 0$). Thus, we find

$$\begin{aligned} \frac{1}{2} \langle (S_{Tot}^2 - S_e^2 - S_p^2) \rangle &= \frac{\hbar^2}{4} \text{ for the triplets,} \\ &= -\frac{3\hbar^2}{4} \text{ for the singlet.} \end{aligned} \quad (22.137)$$

This leads to the following first order corrections for the triplet and singlet states

$$\begin{aligned} E_{hyp}^{(1)}(n=0, S_{Tot} = 1) &= \frac{2\tilde{\mu}_0 g_p g_e \mu_0^2}{3\pi a_0^3} \times \frac{\hbar^2}{4}, \\ E_{hyp}^{(1)}(n=0, S_{Tot} = 0) &= -\frac{2\tilde{\mu}_0 g_p g_e \mu_0^2}{3\pi a_0^3} \times \frac{3\hbar^2}{4}, \end{aligned} \quad (22.138)$$

such that the energy split between these two sets of states is given

$$\begin{aligned}\Delta E_{hyp}^{(1)}(n=0) &= \frac{2\tilde{\mu}_0 g_p g_e \mu_0^2 \hbar^2}{3\pi a_0^3}, \\ &= \frac{2E_1}{3} \frac{g_p g_e}{(4\pi\epsilon_0)^2} \frac{m_e}{m_p} \alpha^2,\end{aligned}\quad (22.139)$$

where $E_1 = \hbar^2/2m_e a_0^2 = 13.6\text{eV}$ is the modulus of the ground state energy of the H-atom. From this expression, it is now clear that $\Delta E_{hyp}^{(1)}(n=0) \sim \frac{m_e}{m_p} \alpha^2$ smaller than the ground state energy E_1 , and a factor $m_e/m_p \sim 10^{-7}$ smaller than the correction arising from the fine structure (i.e., spin-orbit) interaction ($\sim \alpha^2$). Putting all the numbers in, we find that $\Delta E_{hyp}^{(1)}(n=0) = 5.88 \times 10^{-6} \text{ eV}$, such that the frequency and wavelength of the radiation emitted in undergoing the transition between the triplet and singlet states is given by

$$\nu = \frac{\Delta E_{hyp}^{(1)}(n=0)}{h} = 1420 \text{ MHz}, \quad \lambda = \frac{c}{\nu} = 21 \text{ cm}. \quad (22.140)$$

This wavelength falls within the microwave range of wavelengths, and is amongst the most common form of radiation observed in radio astronomy. Indeed, astronomers (such as those based at the GMRT in Pune, India) use this radiation to determine how far away various celestial objects are from us via the Doppler effect.

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.