My Notes on Quantum Mechanics

Salah Nasri

^aUnited Arab Emirates University, Al-Ain, UAE

E-mail: snasri@uaeu.ac.ae

Abstract. These are my notes on Quantum Mechanics for the course PHY 355 that I have taught at UAEU during the academic years **2012**, **2013**, **2015**, and **2016**. All figures and photos presented in these notes are referenced.

- 1. In the first three chapters I discuss the **historical developments** of the quantum theory. I find it important for Physics students to appreciate how the Quantum theory evolved to its current form as it is presented in textbooks. For that I relied on the original papers written by the founders of the Quantum Theory and a number of excellent books such as [1], [2] and [3], given in the list of references.
- 2. I tried hard to avoid the statement "it can be shown that ...". For that, I gave a **detailed derivations** to most of the important results presented in these notes.
- 3. Topics with **superscript** * **are for advanced** undergraduate students. These can be discussed at the end of the course if time permits.
- 4. I have included a lot of footnotes (more than 80).
- 5. I suggested a **set of exercises**, and provided the final answers to most of them. Most of these are straightforward, but some of them require both the thinking and technical skills.
- 6. I added a number of **Appendices** which I felt they can be useful to the students who want to know more about the topics discussed in theses notes.

Contents

1	Cracks in the Foundations of Classical Physics		2
	1.1	Black Body Radiation [1859 – 1901]	2
		1.1.1 Kirchhoff's law of radiation [1859]	3
		1.1.2 Stefan-Boltzmann law $[1879/1884]$	4
		1.1.3 Wien's Displacement Law [1893]	6
		1.1.4 Wien's Energy Distribution for a Black Body [1896]	9
		1.1.5 Rayleigh-Jeans law and Ultra-Violet Catastrophe [1900/1902]	10
		1.1.6 Planck Distribution $[1900 - 1901]$	12
	1.2	Photoelectric Effect $[1887 - 1905]$	21
		1.2.1 Experimental Observations	21
		1.2.2 Predictions of the Classical Physics	23
		1.2.3 Einstein explanation for the photoelectric effect	24
	1.3	Compton Effect [1923]	26
2	Old	Quantum Mechanics: [1913 – 1925]	29
	2.1	Atomic Spectral Lines	29
	2.2	Instability of Atoms	31
	2.3	Bohr's Model [1913]	33
	2.4	Sommerfeld Extension of Bohr's Model $[1915 - 1919]$	38
	2.5	Einstein Coefficients [1916]	42
3	Wa	ve Mechanics [1923 – 1927]	45
	3.1	Wave-Particle Duality and Davisson-Germer Experiment [1923 – 1927]	45
	3.2	Schrodinger Equation [1926]	48
	3.3	Interpretation of the Wave Function [1926]	50
1			00
-	Exa	mples of One Dimensional Potentials	52
т	Exa 4.1	Examples of One Dimensional Potentials Free Particle in 1-dimension and Wave Packet	52 52
т	Exa 4.1 4.2	Imples of One Dimensional Potentials Free Particle in 1-dimension and Wave Packet Infinite Square Potential	52 52 57
т	Exa 4.1 4.2 4.3	Imples of One Dimensional Potentials Free Particle in 1-dimension and Wave Packet Infinite Square Potential Symmetric Finite Potential: Bound and Unbound States	52 52 57 57
т	Exa 4.1 4.2 4.3 4.4	Imples of One Dimensional Potentials Free Particle in 1-dimension and Wave Packet Infinite Square Potential Symmetric Finite Potential: Bound and Unbound States Particle moving through a Step Potential	52 52 57 57 57
T	Exa 4.1 4.2 4.3 4.4 4.5	Examples of One Dimensional Potentials Free Particle in 1-dimension and Wave Packet Infinite Square Potential Symmetric Finite Potential: Bound and Unbound States Particle moving through a Step Potential Harmonic Oscillator	52 52 57 57 57 61
T	Exa 4.1 4.2 4.3 4.4 4.5 4.6	Imples of One Dimensional Potentials Free Particle in 1-dimension and Wave Packet Infinite Square Potential Symmetric Finite Potential: Bound and Unbound States Particle moving through a Step Potential Harmonic Oscillator Quantum Tunneling and the WKB approximation	52 52 57 57 57 61 63
5	Exa 4.1 4.2 4.3 4.4 4.5 4.6 Mat	Interpretation of the trace random [19926]Imples of One Dimensional PotentialsFree Particle in 1-dimension and Wave PacketInfinite Square PotentialSymmetric Finite Potential: Bound and Unbound StatesParticle moving through a Step PotentialHarmonic OscillatorQuantum Tunneling and the WKB approximationtrix Mechanics * [1925]	52 52 57 57 57 61 63 64
5	Exa 4.1 4.2 4.3 4.4 4.5 4.6 Ma 5.1	 Imples of One Dimensional Potentials Free Particle in 1-dimension and Wave Packet Infinite Square Potential Symmetric Finite Potential: Bound and Unbound States Particle moving through a Step Potential Harmonic Oscillator Quantum Tunneling and the WKB approximation trix Mechanics * [1925] Heisenberg's Re-interpretation of the Position and New Multiplication 	 52 52 57 57 61 63 64
5	Exa 4.1 4.2 4.3 4.4 4.5 4.6 Ma 5.1	Interpretation of the trace random [19926] Imples of One Dimensional Potentials Free Particle in 1-dimension and Wave Packet Infinite Square Potential Symmetric Finite Potential: Bound and Unbound States Particle moving through a Step Potential Harmonic Oscillator Quantum Tunneling and the WKB approximation trix Mechanics * [1925] Heisenberg's Re-interpretation of the Position and New Multiplication Rules	52 52 57 57 57 61 63 64 64
5	Exa 4.1 4.2 4.3 4.4 4.5 4.6 Ma 5.1 5.2	 Interpretation of the trace random [19926] Imples of One Dimensional Potentials Free Particle in 1-dimension and Wave Packet Infinite Square Potential Symmetric Finite Potential: Bound and Unbound States Particle moving through a Step Potential Harmonic Oscillator Quantum Tunneling and the WKB approximation trix Mechanics * [1925] Heisenberg's Re-interpretation of the Position and New Multiplication Rules Heisenberg's Quantum Condition and The non linear Oscillator 	52 52 57 57 57 61 63 64 64 66
5	Exa 4.1 4.2 4.3 4.4 4.5 4.6 Ma 5.1 5.2 5.3	 Interpretation of the Water Function [1990] Imples of One Dimensional Potentials Free Particle in 1-dimension and Wave Packet Infinite Square Potential Symmetric Finite Potential: Bound and Unbound States Particle moving through a Step Potential Harmonic Oscillator Quantum Tunneling and the WKB approximation trix Mechanics * [1925] Heisenberg's Re-interpretation of the Position and New Multiplication Rules Heisenberg's Quantum Condition and The non linear Oscillator Born and Jordan's Matrix Mechanics 	52 52 57 57 61 63 64 64 66 72

6	The Mathematical Formalism of Quantum Mechanics	81
	6.1 Pre-Hilbert Space	81
	6.2 Dual Space and Bras	83
	6.3 Hilbert Space	84
	6.4 Linear Operators	85
	6.5 Hermitian and Anti-Hermitian	89
	6.6 Isometric and Unitary Operators	91
	6.7 Projection Operator	91
	6.8 Matrix Representation	92
	6.9 Tensor Products	94
7	Principles of Quantum Mechanics	95
	7.1 The Postulates of Quantum Mechanics	95
	7.2 Position and Momentum representations	99
	7.2.1 Position representation	99
	7.2.2 Momentum representation	103
	7.3 Compatible Operators and Simultaneous Measurement	106
	7.4 Heisenberg Uncertainty Relation	107
	7.5 Schrodinger, Heisenberg and Interaction Picture *	108
8	Harmonic Oscillator Using Ladder Operators	110
	8.1 Raising and Lowering Operators	110
	8.2 Energy Spectrum and Eigenstates	111
	8.3 The Wave Function of the Harmonic Oscillator	112
	8.4 The Matrix Representation of the Harmonic Oscillator	114
	8.5 Coherent States	115
9	Theory of Angular Momentum	116
	9.1 Angular Momentum Algebra	116
	9.2 Quantization of Angular momentum	118
	9.3 Angular Momentum and Rotations in \mathbb{R}^3	121
	9.4 Representation of Rotation in Hilbert State	125
10	The Hydrogen Atom	126
11	Approximation Methods	127
	11.1 Time Independent Perturbation Theory	127
	11.1.1 Non Degenerate Spectrum	127
	11.1.2 Degenerate Spectrum	131
	11.2 variational Methods	131
	11.3 Time-Dependent Perturbation Theory	131

12 Formal Theory of Scattering		
12.1 Møller Operator	132	
12.2 Properties of the Møller Operator	133	
12.3 Møller Operator and the Resolvant of \hat{H}	136	
12.4 The Transition Operator	138	
12.5 The Case of System with Bound States	139	
12.6 Unitarity Relations	140	
12.7 The \hat{S} Operator	141	
12.8 Relating the S- Matrix and the T-Matrix	143	
12.9 A heuristic Derivation of the Cross Section	144	
13 Interpretation of Quantum Mechanics *	146	
13.1 Mixed States and the Density Matrix	146	
13.2 Measurement and Interpretation	146	
13.3 Schrodinger's Cat	146	
13.4 Einstein-Podolsky-Rosen (EPR) Argument	146	
14 Exercices	147	
15 Answers	148	
16 Exams with Solutions	149	
17 Appendices	179	
17.1 Specific intensity and Flux	180	
17.2 The Hamilton-Jacobi Equations	186	
17.3 Ehrenfest's Adiabatic Principle	186	
17.4 Discovery of the Electron Spin	186	
17.5 Clebsch Gordon Coefficients	186	
17.6 Derivation of Cross Section using Wave Packets	186	

1 Cracks in the Foundations of Classical Physics

1.1 Black Body Radiation [1859 – 1901]

In the middle of the 19th century, Physicists studied thermal radiation radiation, i.e. the emission and absorption of radiation by an object at some temperature T. It was known then that an object whose temperature is higher than the surrounding, it will emit more radiation than it absorbs from its surrounding. Conversely, a cooler object absorbs more radiation than it emits. Moreover, objects made of different materials and kept at different temperatures emit different amount of radiation.

An object which absorbs all the incident radiation for all wavelengths is called **ideal black body**, or for short a black body. It is an object that allows the whole of the incident radiation to pass into it and absorbs within itself the whole incident radiation¹. The reason for such name is because those bodies that absorbs all the incident visible light seem black to the human eye². A body that does not absorb all incident radiation is sometimes known as a **grey body**.

In practice, a very good approximation of a black body can be realized by considering as a hollow cavity with a tiny hole, as shown in fig.1, so that practically no radiation can escape from it. In this case any ray of radiation that penetrates the hole will be reflected by the cavity's walls and eventually absorbed by them. On the other hand, the atoms/molecules of the wall emit also radiation due to their chaotic motion. After sufficiently long time, the amount of the emitted and absorbed radiation of a given wave length will be equal and isotropic. That is the radiation is in thermal equilibrium with the walls of the cavity, and so the only parameters that can characterize the radiation are the temperature of the cavity and the wavelength of the radiation.



Figure 1: A schematic of a black body. Taken from ref..

¹The concept of black-body was introduced by Gustav Robert Kirchhoff in 1860.

 $^{^{2}}$ The term black is purely conventional. For instance, our sun is almost a black body within a wide band of the electromagnetic radiation wavelengths, however, surely, we don't characterize it as a black object. Another example of radiation that is approximately black body radiation is the one emitted by a metal heated by a welding torch.

1.1.1 Kirchhoff's law of radiation [1859]

In 1859, Gustav Kirchhoff, a German physicist derived a law which gives the relation between the emission and absorption intensities of any object in thermal equilibrium with its surrounding. But before discussing Kirchhoff's law, we will first introduce few definitions. Let dE_{ν} be the amount of energy of radiation in the frequency band between ν and $\nu + d\nu$ which propagates inside an infinitesimal solid angle $d\Omega$ in the direction Ω and passes through elementary area dA at position r during the time interval from t to (t+dt). We denote by θ the angle between the unit vector \hat{n} normal to the element of area d A and the Ω . If $I(\nu, T, \Omega)$ is the spectral intensity radiated by the object (see the appendix for the definition and details) and $I_{BB}(\nu, T)$ is that of black-body at the same temperature and the same wavelength, then the ratio

$$\kappa(\nu, T, \mathbf{\Omega}) = \frac{I^{emis}(\nu, T, \mathbf{\Omega})}{I_{BB}(\nu, T)}$$
(1.1)

is called **the directional emissivity** (or the directional emissive ability) of an object. Similarly, we define **the directional spectral absorptivity** (or the absorbing ability) of an object as

$$\alpha(\nu, T, \mathbf{\Omega}) = \frac{I^{abs}(\nu, T, \mathbf{\Omega})}{I_{inc}(\nu, T, \mathbf{\Omega})}$$
(1.2)

where $I^{abs}(\nu, T, \Omega)$ is the spectral intensity of absorbed radiation, and $I_{inc}(\nu, T, \Omega)$ is the one of the incident radiation.

Now, let us assume that an object at temperature T is held inside an isothermal, ideally black cavity. After sufficiently long time both the walls of the cavity and the object will have the same temperature T. The intensity of radiation incident on an element d A of the object's surface in the direction Ω will be equal to $I_{BB}(\nu, T)$ since for an ideal black-body cavity the radiation intensity is isotropic. Since the object and the cavity walls came into equilibrium, then the fluxes of the absorbed and the emitted radiation should be equal³, and therefore we get

$$\kappa(\nu, T, \mathbf{\Omega}) = \alpha(\nu, T, \mathbf{\Omega}) \tag{1.3}$$

The above equality is called **Kirchhoff's law**. It states that at equilibrium the emissivity of a body is equal to its absorptivity and is valid for all bodies in a state of thermal equilibrium with their surroundings. It also implies that a good absorber of a particular frequency is a good emitter at the same frequency; similarly weak absorber is weak emitter.

Another formulation of Kirchhoff's law follows from (1.1), (1.2), and (1.3), which yield

³Otherwise, the body will heat up or cool down in time, contrary to the assumption of thermal equilibrium.

$$\frac{I(\lambda, T, \mathbf{\Omega})}{\alpha(\lambda, T, \mathbf{\Omega})} = I_{BB}(\lambda, T)$$
(1.4)

So the ratio of the radiation intensity of a body, at temperature T, to its absorptivity is a universal function of temperature and wave length, but it is independent of the material and the shape of the object.

It is straight forward to show that

$$I_{BB}(\lambda, T) = \frac{c}{4\pi} u(\lambda, T) \tag{1.5}$$

Here c is the speed of light, and $u(\lambda, T)$ is the spectral energy density, i.e. energy density per unit wavelength, of a black-body radiation. By integrating equation (1.5) over all wavelengths, yields

$$I_{BB}(T) = \int_0^\infty I_{BB}(\lambda, T) d\lambda = \frac{c}{4\pi} u(T)$$
(1.6)

where $I_{BB}(T)$ is the radiant energy emitted per unit time per unit area by a black-body, and u(T) is its total energy density, which depend only on the temperature of the body.

In 1859, very little was known about $u(\lambda, T)$, and Kirchhoff had challenged physicists to find it. Indeed, it was one of the great experimental challenges of the next four decades which attracted the interest of many scientists.

1.1.2 Stefan-Boltzmann law [1879/1884]

In 1879, Joseph Stefan, after analyzing the experimental data published by Tyndall in an 1865 book on the radiation emitted by a platinum strip heated electrically to different temperatures, deduced that the total power radiated per unit area by the heated body is proportional to the fourth power of its temperature. Later in 1884, Ludwig Boltzmann, who was an ex-student of Stefan, provided a theoretical derivation of this law using thermodynamics considerations (the first and the second law) and Maxwell's equations of electromagnetism. Below I present his derivation.

Imagine a cylinder of volume V filled only with electromagnetic radiation enclosed by a piston so that the gas of radiation can be compressed or expanded as the piston moves, which allows the radiation to do work on its external environment and vice versa. The walls of the cylinder are assumed to be perfectly reflecting and maintained at constant temperature T by being in contact with an external reservoir. If some heat ΔQ is added to the system, the total internal energy increases by an amount dU and work is done on the piston as the radiation gas exerts a pressure P on it, increasing the volume of the cylinder by dV. From the conservation of energy one can write,

$$\delta Q = dU + PdV \tag{1.7}$$

or, equivalently,

$$dU = TdS - PdV \tag{1.8}$$

with $dS = \delta Q/T$ is the entropy increase of the gas. The above equation can re-written in terms of energy and entropy densities u = U/V and s = S/v, as

$$du = Tds + \frac{1}{V} \left(Ts - u - P\right) dV \tag{1.9}$$

At the time when Boltzmann set out to prove Stefan formula, it was known that the radiation pressure was derived almost a decade earlier by **James Clerk Maxwell** in his **Treatise on electricity and magnetism**, and it is given by⁴

$$P = \frac{1}{3}u\tag{1.10}$$

Substituting this expression in (1.9), we get

$$du = Tds + \frac{1}{V}\left(Ts - \frac{4}{3}u\right)dV \tag{1.11}$$

But u and s depend only on temperature, not volume, and so

$$Ts = \frac{4}{3}u; \qquad du = Tds \tag{1.12}$$

Combining the two equations above, yields

$$\frac{du}{u} = 4\frac{dT}{T} \tag{1.13}$$

So, we obtain

⁴Another way to get this result is given in the appendix.

$$u = aT^4 \tag{1.14}$$

where a is called **the radiation density constant**. The emergent flux density from a black-body is⁵

$$\mathcal{F}_{BB} = \int_{\cos\theta>0} I\cos\theta \ d\Omega = \pi I_{BB}(T) \tag{1.15}$$

Using (??), we obtain

$$\mathcal{F}_{BB} = \sigma \ T^4 \tag{1.16}$$

which is known as the **Stefan-Boltzmann** law for the total radiant energy. The constant of proportionality $\sigma = ca/4$, called the **Stefan's constant**, is measured to be approximately $5.67 \times 10^{-8} W/m^2 K^4$ in the SI unit system⁶.

For a body which is not a black body, emits less radiation than given by equation (1.16). It is, however, proportional to T^4 and can be written as

$$\mathcal{F} = \kappa \ \sigma \ T^4 \tag{1.17}$$

where the dimensionless constant κ is the emissivity of the body which is less than unity.

1.1.3 Wien's Displacement Law [1893]

In 1893, based on a combination of thermodynamics and electromagnetism considerations, Wilhelm Wien showed that the wavelength of maximum emission of any body is inversely proportional to its absolute temperature (measured in Kelvin). Below I will show how he derived his formula.

Wien considered the following thought experiment. He imagined a perfectly reflecting spherical cavity of radius r and volume V, filled with black-body radiation with temperature T. Assume that the sphere expands at a speed v_w that is vanishingly small

$$a = \begin{cases} 7.56 \times 10^{-16} \ W/m^3 K^4, & \text{SI unit system} \\ \\ 7.56 \times 10^{-15} \ erg/s.cm^3.K^4, & \text{cgs unit system.} \end{cases}$$

⁵See appendix for the relation between the radiant flux density and specific intensity.

⁶In the cgs unit system $\sigma = 5.67 \times 10^{-5} \ erg/s.cm^2.K^4$. The value of the radiation density is

relative to the speed of light so that the process is adiabatic. Then, with U = uV and P = u/3, we have

$$d(uV) + \frac{1}{3}udV = 0 \Rightarrow \frac{du}{u} = -\frac{4}{3}\frac{dV}{V}$$
(1.18)

which after integration yields

$$u \propto V^{-4/3} \tag{1.19}$$

Using (1.14) for the energy density, we obtain

$$T \propto V^{-1/3} \propto r^{-1} \tag{1.20}$$

Now, in this adiabatic process, where there is a change in volume, we expect that the a shift in the wave length of the radiation reflected on the wall of the cavity. Let ν be the frequency of the incident radiation as seen by an observer at rest, and K_w the reference frame attached to an element of the wall on which a ray of radiation strikes. Then, in K_w reference frame, the source of the radiation looks that is receding with a speed v_w , and so the radiation get redshifted in this frame, with frequency $\nu_{inc}^{(K_w)}$, given by

$$\nu_{inc}^{(K_w)} = \nu \frac{\sqrt{1 - \frac{v_w^2}{c^2}}}{1 + \frac{v_w}{c} \cos \theta_i}$$
(1.21)

where θ_i is the angle of incidence on a point at the wall. Note that only the component of the of the velocity parallel to the direction of observation that contribute to the shift in the frequency. Then the radiation gets reflected by an angle θ_r but with the same frequency from K_w viewpoint since the element of the wall is at rest in this frame, i.e. $\nu_{ref}^{K_w} = \nu_{inc}^{K_w}$. So, the reflected radiation will be seen by an observer at rest to have frequency

$$\nu' = \nu_{ref}^{(K_w)} \frac{\sqrt{1 - \frac{v_w^2}{c^2}}}{1 + \frac{v_w}{c} \cos \theta_r}$$

$$= \nu_{inc}^{(K_w)} \frac{\sqrt{1 - \frac{v_w^2}{c^2}}}{1 + \frac{v_w}{c} \cos \theta_r}$$

$$= \nu \frac{1 - \frac{v_w^2}{c^2}}{\left(1 + \frac{v_w}{c} \cos \theta_i\right) \left(1 + \frac{v_w}{c} \cos \theta_r\right)}$$
(1.22)

or, equivalently, in terms of wavelengths

$$\frac{\lambda'}{\lambda} = \frac{1 + \frac{v_w}{c} \left(\cos\theta_i + \cos\theta_r\right) + \frac{v^2}{c^2} \cos\theta_i \cos\theta_r}{1 - \frac{v^2}{c^2}}$$
(1.23)

The angle reflection θ_r can be expressed in terms of θ_i as⁷

$$\cos \theta_r = \frac{-2\frac{v_w}{c} + \left(1 + \frac{v_w^2}{c^2}\right)\cos \theta_i}{1 - 2\frac{2v_w}{c}\cos \theta_i + \frac{v_w^2}{c^2}}$$
(1.24)

which for $v_w/c \ll 1$ can be approximated by

$$\cos\theta_r \simeq \cos\theta_i$$
 (1.25)

Substituting this result in (1.23), we obtain⁸

$$\frac{\Delta\lambda}{\lambda} = \frac{2v_w}{c}\cos\theta_i \tag{1.26}$$

This the shift in the wavelength during the time $\Delta t = 2r \cos \theta_i/c$ between reflections. In the infinitesimal time interval d t, the the radius of the sphere expands by $dr = v_w dt$, and the change in the wave length during this time is $d\lambda = (dt/\Delta T) \Delta \lambda$. Therefore, we get

$$\frac{d\lambda}{\lambda} = \frac{dr}{r} \Rightarrow \lambda \propto r \tag{1.27}$$

Combining this result with the one in (1.20), we find⁹

$$\lambda T = b \tag{1.28}$$

where b is a constant, called Wien's displacement constant. Note that if we consider the whole spectrum of blackbody radiation, or some interval of it, at some given temperature, then b is not a constant, instead it changes with the wavelength¹⁰. So, for a radiation with a particular wavelength and subject it to an adiabatic expansion or compression, its wavelength is inversely proportional to the temperature. That it is the radiation wavelength get displaced as the temperature changes. As a result, as the temperature rises, the maximum of the radiant energy shifts toward the shorter wavelength end of the spectrum, i.e.

⁷For the derivation see the appendix.

⁸Actually, this result could also be obtained as if the image were moving away from the observer at a speed $2v_w \cos \theta / (1 + v_w^2/c^2)$.

 $^{^9 {\}rm The}$ value of the constant b was determined by Lummer and Pringsheim in 1897 to be $\simeq 0.29 \ cm.K.$

¹⁰W. Wien, Sitzungsberichte der preussischer Akademie, pp. 55-62 (1893).

$$\lambda_{max} = \frac{b}{T} \tag{1.29}$$

where λ_{max} is the wavelength at the peak of the radiation intensity spectrum¹¹.

1.1.4 Wien's Energy Distribution for a Black Body [1896]

In 1896, based on simple thermodynamics considerations, Wien was able to derive a constrain on the form of spectral energy distribution of a black-body radiation. He found that $u_{\lambda}\lambda^5$ is a constant. To show this, consider a blackbody radiation, approximately monochromatic, with wavelength between λ and $\lambda + d\lambda$ at temperature T in a cavity of volume V which is expanded adiabatically. So during the expansion, the volume changes by ΔV , and the work done by the radiation on wall (negative work) of the cavity is equal to the decrease of the energy within the cavity, i.e.

$$\frac{1}{3}(u_{\lambda} \ d\lambda)\Delta V = -\Delta \left(u_{\lambda} d\lambda V\right) \tag{1.30}$$

or, equivalently,

$$\frac{4}{3}\frac{\Delta V}{V} + \frac{\Delta u_{\lambda}}{u_{\lambda}} + \frac{\Delta(d\lambda)}{d\lambda} = 0$$
(1.31)

where we have used the fact that the radiation pressure is one third the energy density of the radiation. Since $d\lambda$ is the interval within which the wavelengths are included, we have

$$\frac{\Delta(d\lambda)}{d\lambda} = \frac{\Delta\lambda}{\lambda} \tag{1.32}$$

so,

$$\frac{4}{3}\frac{\Delta V}{V} + \frac{\Delta u_{\lambda}}{u_{\lambda}} + \frac{\Delta \lambda}{\lambda} = 0 \tag{1.33}$$

But, according to (1.27), $\Delta V/V = 3\Delta\lambda/\lambda$, and the above equation reads

$$\frac{\Delta u_{\lambda}}{u_{\lambda}} = -5\frac{\Delta\lambda}{\lambda} \tag{1.34}$$

which implies that

$$u_{\lambda} \ \lambda^5 = Constant \tag{1.35}$$

Now, the only combination of wavelength and temperature that is constant is λT . So, in general, the spectral energy density has the form

$$u_{\lambda} = \lambda^{-5} \Phi(\lambda T) \tag{1.36}$$

¹¹ In some textbooks, the above expression is called Win's displacement law.

where Φ is an arbitrary function of a single variable λT . The above formula is called the general **Wien's displacement law**. Note that integrating (1.36) over the wavelength gives

$$u_{\lambda}(T) = \int_0^\infty u_{\lambda} d\lambda = \int_0^\infty \lambda^{-5} \Phi(\lambda T) \ d\lambda = T^4 \int_0^\infty x^{-5} \Phi(x) \ dx \tag{1.37}$$

which is proportional to T^4 , in agreement with the Stefan-Boltzmann law. Note that for the integration to be finite, $\Phi(x)$ must approach zero faster than x^5 when $x \to 0$ and $x \to \infty$.

To determine the form of the function Φ , Wien made made the assumption that the blackbody radiation was emitted by molecules obeying Maxwell's velocity distribution, where each one emits a radiation of a single wavelength only depending on the velocity of the molecule. He also supposed the spectral energy density u_{λ} , in the wavelength interval $[\lambda, \lambda + d\lambda]$, is proportional to the number of vibrating molecules with periods corresponding to wavelengths in the same range. Based on these assumptions he found that¹²:

$$u_{\lambda}(T) = \frac{\alpha}{\lambda^5} exp\left[-\frac{\beta}{\lambda T}\right]$$
(1.38)

where α and β are constants to be determined empirically. In terms of frequency, Eq (1.38) reads

$$u_{\nu}(T) = C_1 \nu^3 exp\left[-\frac{C_2 \nu}{T}\right]$$
(1.39)

where we used the equality $u_{\lambda}|d\lambda| = u_{\nu}|d\nu|$, and the relation $\lambda = c/\nu$, with c being the speed of light. Here $C_1 = \alpha/c^4$, and $C_2 = \beta/c$. In 1897, Lummer and Pringsheim¹³ had shown by a very careful measurements that Stefan-Boltzmann and Wien's radiation formula were correct for short wave lengths. Based on their data they determined $b \simeq 0.29 \ cm$. K. However, soon after that, they carried out experiments for long wavelengths and showed that Wein's radiation formula was incorrect for shorter wavelengths¹⁴.

1.1.5 Rayleigh-Jeans law and Ultra-Violet Catastrophe [1900/1902]

In 1900, Lord Rayleigh published a paper which suggests a modification of Wien's distribution law, which appeared to give a good fit to the the data for low frequency.

¹²W. Wien, Ann. Phys. Ser. 3, 58, 662-669 (1896). There is also an English version of this paper published in "The London, Edinburg, and Dublin Philosophical Magazine and Journal of Science, Series 5, Vol. 43, No. 262, p 214- 220 (1897)".

¹³O. Lummer and E. Pringsheim, Wied. Ann. 63, 395 (1897).

¹⁴ O. Lummer and E. Pringsheim, Verh. Deutsch. Phys. Gesell. 1, 226 (1899); O. Lummer and E. Pringsheim, Verh. Deutsch. Phys. Gesell. 2, 163 (1900).

His idea was to think of the black-body radiation inside a cavity as a superposition of standing waves, which satisfy the wave equation

$$\vec{\nabla}^2 \psi - \frac{1}{c^2} \frac{\partial^2 \psi}{\partial t^2} = 0 \tag{1.40}$$

For a cubic cavity of size L filled with black body radiation, the energy density of waves, with frequency between ν and $(\nu + d\nu)$, is given by

$$u_{\nu}d\nu = \frac{(Number \ of \ standing \ waves)}{L^3} \ \bar{\epsilon} \ d\nu \equiv \bar{\epsilon}\mathcal{N}(\nu) \ d\nu \tag{1.41}$$

where $\mathcal{N}(\nu)$ is the number of standing waves per unit frequency and unit volume, and $\bar{\epsilon}$ is the average energy of each standing wave ψ . Note that the total energy is independent on the size of the cavity. However, for a standing wave, it should have zero amplitude at the walls. This implies that

$$\psi = C \sin\left(\frac{n_x \pi x}{L}\right) \sin\left(\frac{n_y \pi y}{L}\right) \sin\left(\frac{n_z \pi z}{L}\right) \cos\left(\omega t\right)$$
(1.42)

Here n_x, n_y and n_z are positive integers. So for each standing wave is a harmonic oscillator of frequency $\nu = \omega/2\pi$. If we write $n^2 = (n_x^2 + n_y^2 + n_z^2)$, we obtain the dispersion relation

$$n^2 = \frac{4L^2}{2c^2} \nu^2 \tag{1.43}$$

We can regard the triplets (n_x, n_y, n_z) as points forming a lattice in a 3 dimensional space. For larger modes, the lattice becomes essentially continuous, and then (1.43) represents the equation of a sphere of radius $2L\nu/c$. In this case the number of modes with frequency in the interval $[\nu - \nu + d\nu]$ is given by the surface area of a shell in octant of a concentric sphere with inner and outer radii n and (n + dn), respectively:

$$\mathcal{N}(\nu) \ d\nu = \frac{1}{8} \left[\frac{4\pi}{3} (n+dn)^3 - \frac{4\pi}{3} n^3 \right]$$

$$\simeq \frac{\pi}{2} n^2 \frac{dn}{d\nu} \ d\nu = \frac{4\pi}{c^3} \nu^2 \ d\nu$$
(1.44)

For electromagnetic waves, there are two independent polarization states for a given value of the wave vector $\vec{k} = \pi/L \vec{n}$. Therefore, we need to multiply the above result by a factor of two, and get

$$\mathcal{N}(\nu) \ d\nu = \frac{8\pi}{c^3} \nu^2 \ d\nu \tag{1.45}$$

Now we need to calculate the average energy of each oscillator. This is can be obtained from the equipartition theorem in statistical mechanics, which states that for a system in equilibrium at temperature T, every degree of freedom (or generalized coordinate) that appears only quadratically in the total energy has an average energy of $1/2 k_B T$, where $k_B = 1.38 \times 10^{-23} J/K$ is the Boltzmann constant¹⁵. Since the electromagnetic energy density is $\mathcal{E}_{EM} \propto (\vec{E}^2 + \vec{B}^2)$, one might think that there are 6 degrees of freedom: 3 for the electric field and 3 for the magnetic field. However, \vec{E} and \vec{B} are not independent, they are related are constrained by Maxwell's equations (4 equations), which reduces the degrees of freedom from 4 to (6-4) = 2. So the energy per mode is K_BT . This implies that the energy density per unit frequency of a black body radiation is¹⁶

$$u_{\nu}(T) = \frac{8\pi\nu^2}{c^3} k_B T$$
 (1.46)

The factor 8 on the right hand side of (1.46) was 64 in Rayleigh's original calculation¹⁷, and five years later was corrected by Jeans and showed that the correct result is one-eighth of the amount found by Rayleigh, as given in the above equation. That is why the above distribution is known as the **Rayleigh-Jeans law**¹⁸.

The Rayleigh-Jeans distribution agrees quite well with the experimental results at long wavelengths, but it has an obvious problem in that the irradiance energy increases without bound at high frequencies, leading to the "ultra-violet catastrophe"¹⁹. In other word, the total energy energy density emitted by a black-body, is infinite regardless of the temperature T^{20} .

1.1.6 Planck Distribution [1900 – 1901]

We have seen two theoretical relations for the specific intensity of a black-body radiation given in (1.38) and (1.46), respectively. Wien's law was in disagreement with the data at long wavelengths (IR) and the Rayleigh-jean's law, although it provides a good fit at long wavelength, gives nonsensical result in the UV. This was a signal of a fundamental flaw in classical physics.

$$u_{\nu}(T) = \frac{8\pi}{\lambda^4} \ k_B T$$

¹⁹A name that wasn't coined until 1911 by the Austrian physicist Paul Ehrenfest for the divergence of the Rayleigh-Jeans distribution at high frequencies.

²⁰If one reads Rayleigh's paper given in the footnote above, will see that in the fifth paragraph of his paper he introduces empirically an exponential factor $exp(-\nu/k_BT)$ so that the radiation spectrum would converge at high frequencies. However, his exponential distribution is not consistent with Wien's displacement law given in Eq (1.29).

¹⁵For example, the equipartition theorem for the case of a harmonic oscillator requires assigning $k_B T/2$ of kinetic energy term plus another $k_B T/2$ for potential energy.

¹⁶In terms of wavelength, the energy density reads

 $^{^{17}\}mathrm{J.}$ W. Strutt (Lord Rayleigh), Philos. Mag. 49, 539 (1900).

¹⁸J. H. Jeans, Philos. Mag. 10, 91 (1905).

In October, 1900, Max Planck, found an empirical formula for u_{ν} that fit the black-body radiation data²¹:

$$u_{\nu} d\nu = \frac{8\pi h\nu^3}{c^3} \frac{1}{e^{\frac{h\nu}{k_B T}} - 1} d\nu$$
(1.47)

However, Planck wanted to derive and interpret $(1.47)^{22}$, and that lead him to the discovery of quantum mechanics. To understand Planck's derivation of the black-body sdistribution we need first to discuss the equilibrium radiation spectrum of a harmonic oscillator.

As we described earlier, the radiation field in a cavity in thermal equilibrium with the walls. The radiation field then exerts forces on the electrons making up the molecules and that cause them to be in motion and as a result radiate. So, the electron can be modeled by a harmonic oscillator which radiates when undergoing oscillation. Let us recall that the rate of energy loss of radiation by an accelerated electron in its instantaneous rest frame is given by the Larmor's formula ²³

$$-\left(\frac{d\mathcal{E}}{dt}\right) = \frac{|\mathbf{\ddot{p}}|}{6\pi\epsilon_0 c^3} = \frac{e^2|\mathbf{\ddot{r}}|}{6\pi\epsilon_0 c^3} \tag{1.48}$$

Here $\mathbf{p} = |e|\vec{r}$ is the dipole moment of the electron with respect to some origin, and e is its electric charge. For the case of a one dimensional harmonic oscillator, with amplitude $x = x_0 \cos \omega_0 t$, we get

$$-\left(\frac{d\mathcal{E}}{dt}\right)_{rad} = \frac{e^2\omega^2 x_0^2}{6\pi\epsilon_0 c^3}\cos^2\omega_0 t \tag{1.49}$$

which, after averaging over time yilds

$$\mathcal{P}_{rad} = < -\left(\frac{d\mathcal{E}}{dt}\right) >_{rad} = \frac{e^2 \omega_0^2 x_0^2}{12\pi\epsilon_0 c^3} \tag{1.50}$$

Note that the time averaged of a harmonic oscillator is

$$\langle \mathcal{E} \rangle = \langle \frac{1}{2}m\dot{x}^2 \rangle + \langle \frac{1}{2}m\omega_0^2 x^2 \rangle = \frac{1}{2}mx_0^2\omega_0^2$$
 (1.51)

where we used the fact that for a harmonic oscillator the spring constant is $k = m\omega_0^2$. Defining $\gamma = e^2 \omega_0^2 / 6m \pi \epsilon_0 c^3$, we re-write (1.50) as

$$\mathcal{P}_{rad} = \gamma < \mathcal{E} > \tag{1.52}$$

²¹M. Planck, Verh. Deutsch. Phys. Gesell, 202(1900).

²²As he stated it in one of his letters to William Robert Wood in 1931: "....A theoretical interpretation therefore had to be found at any cost, no matter how high...".

²³For a particle of charge q, we just substitute e^2 by q^2 in the equation (1.48)

The oscillating charge will experience a damping force, F, also called "radiation resistance", due to the radiation it emits. So, the power radiated away by the electron integrated over time must be equal to the mechanical power lost by the oscillator :

$$\int_{t_1}^{t_2} F v dt = -\int_{t_1}^{t_2} \mathcal{P}_{rad} dt \qquad (1.53)$$
$$= \frac{e^2}{6\pi\epsilon_0 c^3} \left[v \frac{dv}{dt} \right]_{t_1}^{t_2} - \frac{e^2}{6\pi\epsilon_0 c^3} \int \frac{d^2v}{dt^2} v dt$$

If the time integration is over a full cycle of oscillation, v dv/dt takes the same value at the equivalent points of the cycle. Thus, we deduce that

$$F = -\frac{e^2}{6\pi\epsilon_0 c^3} \frac{d^3x}{dt^3}$$
(1.54)

Thus, the emitted radiation can be modeled by damped oscillator with equation of motion

$$m\ddot{x} + kx - \frac{e^2}{6\pi\epsilon_0 c^3} \frac{d^3x}{dt^3} = 0$$
 (1.55)

If the damping effect very small, the natural frequency of oscillation is $\omega_0^2 = k/m$. This implies that

$$\frac{d^3x}{dt^3} \simeq \frac{d}{dt} \left(-\omega_0^2 x \right) = -\omega_0^2 \frac{dx}{dt}$$
(1.56)

In this approximation, Eq (1.55) reads

$$\ddot{x} + \omega_0^2 x + \gamma \ \dot{x} = 0 \tag{1.57}$$

with

$$\gamma = \frac{e^2}{6\pi\epsilon_0 c^3} \frac{\omega_0^2}{m} \tag{1.58}$$

Now, if the system is a cavity, with a perfectly reflecting walls, then the re-emitted radiation by the electron remain inside the cavity undergoing multiple bounces on the wall and interact with the oscillators. So, although the radiation is represented by a damped oscillator, the total energy of the system is not lost. Planck called this "conservative damping". So, if the electron is accelerated by an incident wave, with electric field $E_x = eE_{0x} \cos \omega t$, the equation of motion reads

$$\ddot{x} + \omega_0^2 x + \gamma \ \dot{x} = \frac{eE_{0x}}{m} \cos \omega t \tag{1.59}$$

where we assumed that the damping coefficient is very small so that $k/m \simeq \omega_0^2$, with ω_0 is the frequency of free oscillator. The generalization to three dimensions is obtained by simply replacing x by the \vec{r} , and E by the corresponding electric field in three

dimensions. The steady state solution to Eq (1.59) can be written as $x = \mathcal{R}e(x_0e^{i\omega t})$, which after substituting into the equation above yields

$$x_{0} = \frac{eE_{0x}/m}{(\omega_{0}^{2} - \omega^{2} + i\gamma\omega)}$$
(1.60)

Note that x_0 is complex which means that the oscillator does not vibrate in phase with the incident electric field. Hence according to Eq (1.48), the rate of radiation of the oscillator due to the incident wave is given by

$$-\left(\frac{d\mathcal{E}}{dt}\right) = \frac{e^2\omega^4 |x_0|^2}{6\pi\epsilon_0 c^3} \cos^2\left(\omega t - \phi\right)$$
(1.61)

where ϕ is the argument of the complex amplitude x_0 . Taking the average over time gives

$$-\left(\frac{d\mathcal{E}}{dt}\right) = \frac{1}{12\pi\epsilon_0 c^3 m^2} \frac{e^4 \omega^4 E_{0x}^2}{(\omega^2 - \omega_0^2)^2 + \omega^2 \gamma^2}$$
(1.62)

The energy E of the oscillator is $m\omega_0^2 x_{max}^2/2$ and one get

$$\mathcal{E}_{osc}(\omega) = \frac{e^2 E_{0x}^2}{2m} \frac{\omega_0^2}{(\omega^2 - \omega_0^2)^2 + \omega^2 \gamma^2}$$
(1.63)

For small damping, the only driving frequencies that really matter are those close to resonance frequency, i.e. $\omega \simeq \omega_0$, since only these give rise to a large amplitude of oscillations. Thus,

$$\mathcal{E}_{osc}(\omega) \simeq \left(\frac{e^2 E_{0x}^2}{2m}\right) \frac{1}{(\omega - \omega_0)^2 + \gamma^2} \tag{1.64}$$

This energy is for one single frequency ω of an incident radiation. However, in reality, the electric field consists of a superposition of waves with frequency in the interval ω and $(\omega + \Delta \omega)$. For black-body radiation at temperature T we can make the following replacement²⁴

$$\frac{1}{2}\epsilon_0 E_{0x}^2 \to u^{(x)}(\omega, T) \ d\omega \tag{1.65}$$

Hence the total energy of the oscillator can be obtained by integrating $\mathcal{E}_{osc}(\omega)d\omega$ over the above all possible frequencies of the incident radiation:

$$\mathcal{E}_{osc} = \left(\frac{e^2}{\epsilon_0 m}\right) \int_0^\infty \frac{u(\omega, T)}{(\omega - \omega_0)^2 + \gamma^2} d\omega$$
(1.66)

²⁴Form the electromagnetism course you know that the energy density in an electric field E is $u_E = \frac{1}{2} \epsilon_0 E^2$.

Since for γ small, the contribution at the resonance is the most dominant one, we can replace $u(\omega, T)$ by $u(\omega, T)$ in the above integral. This yields²⁵

$$u^{(x)}(\omega_0, T) = \frac{\omega_0^2}{3\pi^2 c^3} \mathcal{E}_{osc}$$
(1.68)

However, this This is the spectral energy density for a single component of the electric field along the direction of the oscillation, say, the x-axis. If the system is homogeneous and isotropic, the other two components of the field and the directions of oscillations will give the same result. So to account for the other directions we simply multiply by three the righthand side of (1.68) and we get

$$u(\omega_0, T) = \frac{\omega_0^2}{\pi^2 c^3} \mathcal{E}_{osc}$$
(1.69)

or, equivalently, in terms of frequency 26

$$u(\nu,T) = \frac{8\pi\nu^2}{\pi^2 c^3} \mathcal{E}_{osc}$$
(1.71)

Here ν is the oscillator characteristic frequency. It is remarquable that this result does not depend on the damping constant or other detailed properties of the resonator²⁷. Moreover, it says that if we, somehow, can manage to find the energy of a given oscillator of frequency ν in an enclosure at temperature T, we will automatically know the spectrum of the black-body radiation. So, the natural thing to do now is to find an expression for $\mathcal{E}_{osc}(T,\nu)$.

In classical thermodynamics, one can use the equipartition theorem where an average energy of $k_B T/2$ is associated to each of the square terms \dot{x}^2 and x^2 in the expression of the energy of the harmonic oscillator, i.e. $\mathcal{E}_{osc} = k_B T$. Thus, we find that

$$u(\nu,T) = \frac{8\pi\nu^2}{\pi^2 c^3} k_B T$$
(1.72)

which is exactly the Rayleigh-Jeans law we saw earlier. As we discussed it in the previous subsection, Eq (1.72) leads to non physical result at high frequencies. It is interesting to note that, although Planck could derive the Rayleigh-Jeans formula, but

$$\int_{-\infty}^{+\infty} \frac{dx}{x^2 + a^2} = \frac{\pi}{a}$$
(1.67)

 26 We used the fact that

$$u(\nu) \ d\nu = u(\omega) \ d\omega \Rightarrow u(\nu) = 2\pi u(\omega = 2\pi\nu) \tag{1.70}$$

²⁷M. Planck, Physikalische Abhandlungen und Vortrage, 1, 560- 600 (1899).

 $^{^{25}}$ We used the integral

he did not. The reason for that is that he did not accept the equipartition theorem as fundamental.

On October, 1900, Rubens, based on the data of the measurements on the blackbody spectrum he and Kurlbaum have taken, informed Planck that at low frequencies the intensity of radiation was proportional to the temperature. For Planck, this means that in the limit $\nu/T \to 0$, the oscillator energy must proportional to the temperature since $u(\nu, T) \propto \mathcal{E}_{osc}$. Using the thermodynamics definition of the absolute temperature $\frac{1}{T} = \frac{\partial S}{\partial \mathcal{E}_{osc}}$, with S being the entropy of the oscillator entropy, he concluded that

$$\lim_{\nu/T \to 0} \frac{\partial^2 S}{\partial \mathcal{E}_{osc}^2} \propto \frac{1}{\mathcal{E}_{osc}^2}$$
(1.73)

For larger values of ν/T , Planck knew that Wien's formula (1.38) is in good agreement with data, which after equating it with his expression for the spectral energy density in (1.47) yields

$$\mathcal{E}_{osc} = \alpha \nu e^{-\beta \nu/T} \tag{1.74}$$

or, equivalently,

$$\frac{1}{T} = -\frac{1}{\beta\nu} \ln\left(\frac{\mathcal{E}_{osc}}{\alpha\nu}\right) = \frac{\partial S}{\partial E}$$
(1.75)

Taking the derivative with respect to the energy of the oscillator gives

$$\lim_{\nu/T \to \infty} \frac{\partial^2 S}{\partial \mathcal{E}_{osc}^2} \propto \frac{1}{\mathcal{E}_{osc}}$$
(1.76)

So, for \mathcal{E}_{osc} between small and large values, the change in entropy satisfies an equation of the form

$$\frac{\partial^2 S}{\partial \mathcal{E}_{osc}^2} = -\frac{a}{\mathcal{E}_{osc} \left(b + \mathcal{E}_{osc}\right)} \tag{1.77}$$

where a and b are undetermined coefficients. Now, integrating the above equation with respect to \mathcal{E}_{osc} gives the inverse of the temperature, i.e.

$$\frac{\partial S}{\partial \mathcal{E}_{osc}} = \frac{1}{T} = \frac{a}{b} \ln \left(\frac{\mathcal{E}_{osc} + b}{\mathcal{E}_{osc}} \right)$$
(1.78)

Solving for \mathcal{E}_{osc} we obtain

$$\mathcal{E}_{osc} = \frac{b}{e^{b/aT} - 1} \tag{1.79}$$

Now substituting the above expression in (1.71), we get

$$u(\nu,T) = \frac{8\pi\nu^2}{c^3} \frac{b}{e^{b/aT} - 1}$$
(1.80)

In the limit where ν/T is very large, the above formula for the energy density must reduce to Wien's expression in (1.39). This requires the coefficient *b* to be proportional to the frequency. Therefore, Planck expression of the energy density reads²⁸

$$u(\nu,T) = \frac{8\pi\nu^3}{c^3} \frac{C}{e^{k\nu/T} - 1}$$
(1.81)

where both C and k are constants that don't depend on the frequency. Planck also calculated the entropy of the oscillator by integrating (1.78), which gives

$$S = a \left[\left(1 + \frac{\mathcal{E}_{osc}}{b} \right) \ln \left(1 + \frac{\mathcal{E}_{osc}}{b} \right) - \frac{\mathcal{E}_{osc}}{b} \ln \left(\frac{\mathcal{E}_{osc}}{b} \right) \right]$$
(1.82)

Although, it did not take long to show that Planck's formula agreed well with the experimental data, it was lacking the theoretical justification. Planck was convinced that thermodynamics was not sufficient to tackle the problem, and so he decided to use statistical mechanics which was developed few years earlier by Boltzmann²⁹. In 1901, Planck provided his approach to the derivation of the spectral distribution in (1.81), based on Boltzmann definition of entropy.

He considered N oscillator with total energy E_N to be shared among them. For that he introduced an element of energy ϵ so that there are $p = E/\epsilon$ to be divided among the N oscillators. So the number of possible ways that one can distribute p energy elements over the N oscillators is ³⁰

$$\Omega(E_N, N) = \frac{(N+p-1)!}{p! (N-1)!}$$
(1.83)

Then, Planck, made the following two statements: (i) entropy is proportional to the logarithm of probability of a state, (ii) the probability of a state is proportional to

²⁹Ludwig Boltzmann, Wiener Berichte 76 (1877), 373D435, reprinted in Wissenschaftlich Abhandlungen, 3 Vols. (Leipzig: Barth, 1909), Vol. 2, pp. 164-223.

³⁰This can be seen as follow. Consider a sequence of symbols

$$\epsilon ... \epsilon | \epsilon ... \epsilon | ... | \epsilon ... \epsilon$$

where n_i is the number of energy elements assigned to the i^{th} oscillator, i.e it is the occupation number of the i^{th} cell, such that $\sum_{i=1}^{N} n_i = p$, and with (N-1) symbols "|" separating the N spaces. There are (p + N - 1)! permutations of the (p + N - 1) symbols in all. However, suffling the ϵ 's among themselves, or the |'s among themselves, do not change the occupation numbers. In other word, the ϵ 's are indistinguishable and so are the | symbols. Thus, in order not to have an over-counting, we must divide the number of permutations by p!(N-1)!, and we obtain the result in (1.83). The general expression of $\Omega(E, N)$ given above was first given (with no proof) by Boltzmann in 1877. In statistical Physics language, the quantity $\Omega(E, N)$ is called the total **number of macrostates**.

To see this explicitly, suppose there are two kids (N = 2), A and B and two identical chocolate bars (p = 2). Then we can distribute the chocolate bars as follows: A gets 2 bars and B get nothing, or B gets two bars and A nothing, or A get one bar and B get one bar. So the number of possible ways to distribute the two bars of chocolate is 3. If we use the formula (1.83) we get the same result.

²⁸Max Planck, Verhandlungen der Deutshchen Physikalischen Gesellschaft, 2, pp. 202-204. An english translation can be found in: H. Kangro, Planck's Original Papers Quantum Physics, pp. 38-45 (1972).

the total number of configurations in which the state can be realized, and (iii) all configurations are equally probable. So,

$$S_N = k \ln \left(\Omega(E_N, N) \right) = k \ln \left(\frac{(N+p-1)!}{p! (N-1)!} \right)$$
(1.84)

where k is a constant, known as **Boltzmann constant**. For very large values of N and p, we can use the Stirling's formula, $N! \simeq N^N$ and $p! \simeq p^p$, and obtain

$$S_N \simeq k \left[(N+p) \ln(N+p) - p \ln p - N \ln N \right]$$

= $kN \left[(N+p) \ln(N+p) - p \ln p - N \ln N \right]$ (1.85)

Thus, for an oscillator of average energy $\mathcal{E}_{osc} = E_N/N$, the average entropy per oscillator, $S = S_N/N$, is given by

$$S = k_B \left[\left(1 + \frac{\mathcal{E}_{osc}}{\epsilon} \right) \ln \left(1 + \frac{\mathcal{E}_{osc}}{\epsilon} \right) - \frac{\mathcal{E}_{osc}}{\epsilon} \ln \left(\frac{\mathcal{E}_{osc}}{\epsilon} \right) \right]$$
(1.86)

Comparing the above expression to the one in Eq (1.82) implies that ϵ must be proportional to the frequency:

$$\epsilon = h\nu \tag{1.87}$$

The constant h is known as **Planck's constant**. So, the expression of the spectral energy density of a black body radiation is³¹

$$u(\nu,T) = \frac{8\pi h\nu^3}{c^3} \frac{1}{e^{h\nu/k_B T} - 1}$$
(1.88)

or, equivalently, in terms of wavelength

$$u(\lambda,T) = \frac{8\pi c \ h}{\lambda^5} \frac{1}{e^{c \ h/\lambda k_B T} - 1}$$
(1.89)

The total energy density for a black-body radiation is

$$u(T) = \int_0^\infty u(\nu, T) d\nu = \frac{8\pi}{c^3} \frac{\nu^3}{e^{h\nu/k_B T} - 1}$$
(1.90)
= $\frac{8\pi k_B^4}{h^3 c^3} T^4 \int_0^\infty \frac{1}{e^x - 1}$

Which has the same form as the Boltzmann formula in Eq (1.14) for the energy density, with

$$a = \frac{48\pi k_B^4}{c^3 h^3} = \frac{4}{c}\sigma$$
(1.91)

³¹Max Planck, Ann. Phys. 4, pp. 553-563 (1901), Ann. Phys. 4 (1901), 564-566.

To compare with Wien's displacement law, Planck differentiated his expression for the spectral energy density (1.89) with respect to the wavelength and set it to zero to obtain

$$\left(1 - \frac{c h}{5k_B \lambda_{max}T}\right) \exp\left(\frac{c h}{k_B \lambda_{max}T}\right) = 1$$
(1.92)

By solving the above transcendental equation, yields

$$\lambda_{max}T = \frac{c\ h}{4.965\ k_B}\tag{1.93}$$

Now, the expression of the energy density that Planck has derived contains two constants: k_B^{32} and h, which were inferred from the experimental data. Kurlbaum had measured the total heat radiated from a 1 cm^2 per second of a black-body at $T = 0^{(0)}$ Celsius (273 K) and at $T = 100^{(0)}$ C (373 K) and found that the difference was $7.31 \times 10^5 \ erg/cm^2/s = 0.0731 \ Watt/cm^2$. So, with these experimental data, and using Stefan-Boltzmann law, Planck calculated the energy density of a black-body radiation at T = 1 K:

$$u(T = 1 \ K) = \frac{4}{c} \sigma \ (1 \ K)^4 = \frac{4}{3 \times 10^{10} \ cm/s} \left(\frac{7.31 \times 10^5 \ erg.cm^{-2}.\ s^{-1}}{[(373 \ K)^4 - (100 \ K)^4]} \right) (1 \ K)^4$$
$$= 7 \times 10^{-15} \ erg. \ cm^{-3}.\ s^{-1} = 7 \times 10^{22} \ Watt. \ cm^{-3}$$
(1.95)

Now, using his formula (1.91) he deduced that

$$\frac{k_B^4}{h^3} = 1.168 \times 10^{15} \left[\frac{J}{K^4. s^3} \right]$$
(1.96)

Next, he made use of the experimental result on the value of $\lambda_{max}T$ reported by Lummer and Pringsheim., which they found to be approximately 0.29 cm.K. Replacing this value in (??) he got

$$\frac{h}{k_B} = 4.866 \times 10^{-11} \left[\frac{K}{s}\right]$$
(1.97)

$$\Delta S = k_B \ln p_2 - k_B \ln p_1 = k_B \ln p_2 / p_1 = k_B \ln 2^{\mathcal{N}_A} = k_B \mathcal{N}_A \ln 2^{\mathcal{N}_A}$$

However, in principle, you have seen in your thermodynamics course that the change of entropy of a perfect gas undergoing Joule expansion, is given by $\Delta S = R \ln \left(\frac{V_2}{V_1}\right) = R \ln 2$, where $R = 8.314 \frac{J}{mol.K}$ is the universal gas constant. Therefore we conclude that

$$k_B = \frac{R}{\mathcal{N}_A} \tag{1.94}$$

So, the Boltzmann constant is the number of molecules per kilogram molecules.

³²The Boltzmann constant is related to the universal constant of gases and the Avogadro number. This can be shown as follows. Consider a container of volume V_1 filled with one mole of a perfect gas. The the probability for a molecule to occupy only half of the container is 1/2. So, the probability of for the $\mathcal{N}_A \simeq 6 \times 10^{23}$ to occupy half of the container is $p_1 = (1/2)_A^{\mathcal{N}}$, which is, of course, extremely tiny. If we double the volume of the container, the probability change from p_1 to $p_2 = p_1(V_2/V_1)^{\mathcal{N}_A}$ (because the number of arrangements for each molecule got doubled). So, the change in the entropy of the system when the volume was doubled is

Combining (1.96) and (1.97), gave Planck his universal constants

$$h = 6.55 \times 10^{-27} \ erg.s = 6.55 \times 10^{-34} \ J.s$$

$$k_B = 1.35 \times 10^{-16} \ erg.K^{-1} = 1.35 \times 10^{-23} \ J.K^{-1}$$
(1.98)

The best current measured values of these constants are³³

$$h = 6.62606957 \times 10^{-34} J.s$$

$$k_B = 1.3806488 \times 10^{-23} J.K^{-1}$$
(1.99)

Max Planck was awarded the Noble Prize in Physics in 1918 for "his discovery of energy quanta".

1.2 Photoelectric Effect [1887 – 1905]

The photoelectric effect occurs when a metal surface is illuminated with light of certain frequency and electrons are emitted by the surface. The explanation of this effect allowed the understanding of the nature of light which was a crucial step forward in the development of modern Physics.

1.2.1 Experimental Observations

It was first observed in 1887 by Heinrich Hertz^{34} by accident when he was working on an experiment trying to test the existence of electromagnetic waves. He noticed that when the electrode in his experimental setup is shined with ultraviolet light, the intensity of the spark increased. Hertz documented his observation of this phenomenon, but he could not explain why it was happening.

In 1888, Wilhelm Hallwachs considered a circular plate of zinc mounted on an insulating stand and attached by a wire to a gold leaf electroscope, which he charged it negatively. The electroscope loses its energy very slowly. However, when the zinc plate was exposed to ultraviolet light charge leaked away more quickly. He also observed that if the plate was positively charged, there was no fast discharge.

During the period 1900 to 1902, Philipp Lenard, performed a series of experiment to study the newly discovered photoelectric effect by Hertz. A sketch of the experimental setup used by Lenard is similar to the one shown in Fig.3, which consists of a vacuumed glass tube with two electrodes and a window allowing light to shine on a metal surface of the cathode. He found that a current (photocurrent) was induced

 $^{^{33}{\}rm For}$ the latest measurement of planck and Boltzmann constants see the following link at the National Institute of Standards and Technology:

http://physics.nist.gov/cuu/Constants/index.html.

³⁴This is not to be confused with his nephew Gustav Hertz who won the Nobel prize of Physics in 1925.



Figure 2: A schematic of Hallwachs experiment.

as a result of the charges ejected from from the electrode, and had shown that they were identical to the cathode rays discovered by Thomson in 1897, that is, they were electrons³⁵.



Figure 3: A schematic of Lenard's experiment on the photoelectric effect.

By studying the relationship between the kinetic energy energy of the emitted electrons and the intensity and the frequency of incident light he found that:

 $^{^{35}}$ Actually, in 1899, J. J. Thomson showed that the charges emitted from a an illuminated metal (he used zinc plate) were identical to his previously "discovered" electrons. He published his result in Philosophical Magazine 48, 547-567 (1899).

- 1. The number of emitted photoelectrons per unit time is directly proportional to the intensity of the incident light.
- 2. The maximum kinetic energy of the photoelectrons is independent of the intensity of the incident radiation and increases with increasing the frequency.
- 3. There exist a certain minimum frequency for the photoelectric effect to occur. The value of this frequency ν_0 , called threshold frequency, is different for different materials.
- 4. The emission of photoelectrons was almost instantly (less than 10^{-9} s), even when the incident radiation is made exceedingly dim³⁶.

Philipp Lenard was awarded the Nobel Prize in Physics 1905 for " his work on cathode rays".

1.2.2 Predictions of the Classical Physics

In classical Physics, It was believed that when a radiation is shined on a metal, the electrons in atoms are being shaken by the oscillating electric field of the incident light wave, and eventually some of them would would be ejected. So, according to this classical picture, for Lenard's experiment, we expect that:

³⁶One can derive an estimate of the lower limit on the amount of time it should take for an electron at the surface of a metal to absorb enough energy of EM radiation to be able to escape from the metal. For that let us assume that there is one free electron per atom in the metal and imagine that it occupies a square of size of about the atomic spacing d. Then each electron will absorb the energy that falls on its square with the maximum rate of absorption given by

$$\left(\frac{dE}{dt}\right)_{max} = \mathcal{I} \ d^2$$

where \mathcal{I} is the intensity of the incident light, and d^2 is the area of the square. Thus, the maximum amount of energy that an electron can absorb in a time interval Δt is

$$E_{max} = \mathcal{I} \ d^2 \Delta t$$

If the atoms are arranged in a simple cubic array with spacing d, with number density $N_{atom} = d^{-3}$, then

$$d = \left(\frac{\mu}{\rho}\right)^{1/3}$$

with μ is the mass of each atom and ρ is the mass density of the metal. The time it takes before it gets emitted, t_{em} , is the time needed to absorb an amount of energy to overcome its bindingenergy, E_{bin} , in the metal. Thus, setting $E_{max} = E_{bin}$, and solving for $\Delta t = t_{em}$, we obtain

$$t_{em} = \frac{E_{bin}}{\mathcal{I}} \left(\frac{\rho}{\mu}\right)^{2/3}$$

- 1. Increasing the intensity of light would shake the electrons more violently, so there would be more electrons emitted, and on average, they will come out of the cathode at greater speed.
- 2. Increasing the frequency of the incident radiation would shake the electrons faster, which causes the electrons to be emitted faster.
- 3. For very dim light, i.e. low intensity, it would take some time for an electron to reach certain amplitude of vibration before it gets loose and come out of the cathode.

Comparing with the experimental observations, we see that classical physics fails to explain the photoelectric effect.

1.2.3 Einstein explanation for the photoelectric effect

In 1905, Albert Einstein proposed a simple and radical explanation to Lenard's experimental results of the photoelectric effect phenomenon. He extended Planck's concept of quantization (see the subsection on black-body radiation) to electromagnetic waves, by considering radiation as a quanta, which were later called **photons**, of energy $\mathcal{E} = h \nu$, where is h is Planck's constant and ν the frequency of light. In this picture, the entire energy $h\nu$ of the photon is absorbed by an electron. If this energy is less than the minimum energy, ϕ , needed for the electron to escape, called the **work function** (binding energy), from the metal surface, then then no photoelectron can be emitted. Whereas, if the photon energy exceeds the work function, the electron is emitted with a maximum kinetic energy, K_{max} , given by

$$K_{max} = h\nu - \phi \tag{1.100}$$

This equation is known as "Einstein equation of the photoelectric effect". In table 2, we list the work function of some metals. From this proposal we can draw the following conclusions:

- 1. The intensity of light for a given frequency is determined by the number of photons incident per second. So, increasing the intensity will increase the number of photons incident per second, and as a result will increase the number of emitted electrons.
- 2. K_{max} depends linearly on ν because in this picture, the photoelectric effect arises from the absorption of a single photon by a single electron. So Increasing the frequency will increase the kinetic energy of the emitted electrons. Moreover, K_{max} is independent of the intensity of the incident light.
- 3. Since K_{max} is non-negative, Eq (1.100) implies that photoelectric effect is possible only if $hv > \phi$, i.e. $\nu > \phi/h \equiv \nu_0$, where ν_0 is the threshold frequency needed to to emit an electron from the metal surface.

Metal	$\phi(eV)$
Na	2.46
Al	4.08
Cu	4.70
Zn	4.031
Ag	4.73
Pb	4.14
Fe	4.50

Albert Einstein was awarded the Nobel Prize in Physics 1921 for " his services to Theoretical Physics, and especially for his discovery of the law of the photoelectric effect".

Robert Millikan, an American experimental Physicist, did not believe Einstein theory for the photoelectric effect, because it contradicted the wave theory of light, so he set up an experiment to disprove it. To his surprise, he found a linear relationship between the stoping potential, and so the kinetic energy, and the frequency of the incident light, which confirms Einstein theory. This is shown in the fig.4. The slope of the line should be h/e. So, with his previous knowledge of the electron charge³⁷, he was able to measure Planck's constant to within 0.5% precision level.



Figure 4: Experimental results of Millikan experiment on the photoelectric effect. Figure taken from ref [15].

Robert Millikan was awarded the Nobel Prize in Physics in 1923 "for his work on the elementary charge of electricity and on the photoelectric effect".

³⁷Robert Millikan is famous for measuring the charge of the electron in 1909.

1.3 Compton Effect [1923]

Pre-1923, experiments scattering X-rays ³⁸ from a crystal reveled their wave-like properties of construction and destruction interference. From the experimental study of the scattering of X-ray off electrons that are weakly bound to the atom (see Fig. 5), Compton showed that X-rays can act not only as a wave but also as a particles, i.e. photons, which was an important leap in the understanding of the quantum nature of light.

In 1923, Arthur Compton carried out a careful experiment where he scattered a monochromatic beam of X-rays³⁹ off a thin layer of graphite (carbon). He measured the intensity of the scattered X-ray as a function of its wavelength and found that there are were two components in the scattering direction (see Fig. 6). One component had a wavelength λ_0 identical to the one of the incident radiation, whereas the second component had a wavelength λ larger than λ_0 . Furthermore, he found that the shift $\Delta \lambda \equiv (\lambda - \lambda_0)$ has the following properties:

- It depends on the scattering angle θ .
- It increases rapidly at large scattering angles.
- It is independent on the wavelength of the incident radiation.
- It is independent of the scattering material.



Figure 5: Compton scattering process. Figure taken from ref [16]

Now, according the classical Physics, the carbons in the graphite will oscillate at the frequency ν_0 of the incident light, and radiate electromagnetic wave at the

³⁸X rays were discovered in 1895 by William Roentgen while experimenting with a cathode radiation. Roentgen was the first Physicist to be awarded the Nobel Prize in Physics in 1901.

³⁹The energy range of an X-ray is between 0.1 keV to 100 keV. A keV is an energy unit equal to the kinetic energy an electron would gain by being accelerated through a voltage difference of 10³ Volts.

same frequency. So, the effect that Compton observed in his experiment can not be explained with the classical theory. This, lead him to propose that a photon, denoted by γ , behave like a particle so that its scattering off other particles is similar to the collision of billiard balls⁴⁰. From the of special relativity, discovered by Einstein in 1905, the energy and the components of velocity of a relativistic particle are given by

$$E = \sqrt{p^2 c^2 + m^2 c^4}; \qquad v^i = \frac{\partial E}{\partial p_i} = \frac{p^i c^2}{\sqrt{p^2 c^2 + m^2 c^4}}$$
 (1.101)

where the index *i* takes values 1, 2, and 3 corresponding to the x, y, and z components and p^i is the *i*th component of the particle's momentum. For a photon the magnitude of the velocity is the speed of light, i.e. $v = c \simeq 3 \times 10^8 \ m/s$, from which we deduce that

$$m_{\gamma} = 0 \tag{1.102}$$

Hence the photon is massless, and its energy is given by E = pc. Now, since the photon energy i a quanta of energy $E = h\nu$, its momentum is given by

$$p_{\gamma} = \frac{h\nu}{c} = \frac{\hbar\omega}{c} = \hbar\vec{k} \tag{1.103}$$

where $\omega = 2\pi\nu$, $\hbar = h/2\pi$ is called the reduced Planck constant, and \vec{k} is the wavenumber of the electromagnetic radiation of magnitude ω/c .

Let us by $(\vec{p}_{\gamma,i} = h\nu_i, E_{\gamma,i} = p_{\gamma,i}c)$ and $(\vec{p} = h\nu_f, E_{\gamma,i} = p_{\gamma,i}c)$ the energy and momentum of the incident and the scattered photons, respectively. Then, energy and momentum conservation read:

• Energy conservation:

$$h\nu_i + \sqrt{p_{ie}^2 c^2 + m_e c^4} = h\nu_f + \sqrt{p_{e,f}^2 c^2 + m_e^2 c^4}$$
(1.104)

where $p_{e,i}$ and $p_{e,f}$ represent the momenta of the electrons before and after the scattering. Since the energy of an X-ray photon ($\sim keV$) is much larger than the binding energy of an electron it was assumed that the electron (few eV to about 10 eV), we can approximate its energy by its rest mass. So, (1.104) becomes

$$h\nu_i + m_e c^2 = h\nu_f + \sqrt{p_{e,f}^2 c^2 + m_e^2 c^4}$$
(1.105)

It can also re-written as

$$(h\nu_i - h\nu_f + m_e c^2)^2 = p_{e,f}^2 c^2 + m_e^2 c^4$$
(1.106)

⁴⁰A. H. Compton, Phys. Rev. 21, 207 (1923); Phys. Rev. 21, 483 (1923).



Figure 6: A schematic of of Compton experiment and its results: In the plots, the Y-axis represents the intensity and the x-axis is the wavelength of scattered X-rays for different scattering angles [18].

• Momentum conservation:

$$\vec{p}_{\gamma}, i = \vec{p}_{\gamma,f} + \vec{p}_{e,f}$$
 (1.107)

Taking $\vec{p}_{\gamma,f}$ to the left hand-side of (1.107), then squaring both sides of the equation yields

$$p_{e,f}^{2} = p_{\gamma,i}^{2} + p_{\gamma,f}^{2} - 2p_{\gamma,i}p_{\gamma,i}\cos\theta$$
(1.108)

or, equivalently,

$$p_{e,f}^{2} = (h\nu_{i})^{2} + (h\nu_{f})^{2} - 2h^{2}\nu_{i}\nu_{f}\cos\theta + m_{e}^{2}c^{4}$$
(1.109)

Here θ is the angle between the incident and the scattered photon. Substituting the above expression into (1.106), we obtain

$$(\nu_i - \nu_f)m_e c^2 - 2h^2 \nu_i \nu_f = -2h\nu_i \nu_f \cos\theta$$
(1.110)

dividing both sides of the equation above by $\nu_i \nu_f$, and using the relation $\nu = c/\lambda$, we find that the shift in the wavelength is

$$\Delta \lambda \equiv \lambda_f - \lambda_i = \frac{h}{m_e c} \left(1 - \cos \theta \right) \tag{1.111}$$

Thus, the wavelength of the scattered photon has increased (or the frequency has decreased) compared to the wavelength of the incident radiation, and the shift is directly related to the scattering angle. It is convenient to define

$$\lambda_C = \frac{h}{mc} \tag{1.112}$$

called the **Compton wavelength** of particle of mass m, which for the electron is about $2.4 \times 10^{-10} \ cm = 2.4 \ A^{(0)}$. Equation (1.111) was found to be in good agreement with the measured values of $\Delta \lambda$ for different scattering angles in Compton experiment. Since λ_C is very small, one needs high energy radiation, such as X-ray, in order to be able to observe the effect. For instance, an incident light of wavelength $\lambda = 10^{--8} \ cm$ (in the lowest energy region of the X-ray), scattered at right angle, one expect to see about 5% shift in the initial wavelength.

The above calculation correspond to the shifted line in the intensity of the scattered X-rays. The reason for the existence of an un-shifted line is due to the fact that not all X-rays scattered off a free electron. An X-ray photon passing through graphite target could strike valence electrons (a tightly bound electron to its carbon atom), or even the nucleus . So, in the above calculations one should use the effective mass of the electron which will be much larger than the mass of free electron, and in the case of scattering off nucleus one should use mass of the entire the carbon atom . In this case the predicted shift will have identical form as in (1.111) but with m_e replaced by a much larger effective mass. For example, the scattering off the carbon nucleus, the shift in the X-ray wavelength will be more than four order of magnitude smaller ($M_{carbon} = 2 \times 10^4 m_e$) than the shift caused by scattering off free electron. Thus, the scattering of X-rays off a tightly bound electron or the nucleus of the atom is too small to observe, and so the scattered X-ray off such target will seem to come out with the same initial wavelength.

//

2 Old Quantum Mechanics: [1913 - 1925]

2.1 Atomic Spectral Lines

In 1814, **Joseph von Fraunhofer**, made a detailed studies of the solar spectrum and, to his surprise, he found a multitude of dark lines (over 570 lines), indicating that

a certain colors (i.e. light wave lengths) are missing in solar light's spectrum⁴¹. he labelled the 10 strongest lines in the (solar) spectrum by the letters A, a, B, C, D, E, b, F, G and H⁴². These dark lines are now referred to as Fraunhofer lines. However, the origin of these spectral lines were not understood at the time. In 1859, **Robert Bunsen** and **Gustav Kirchhoff** established that a heated low pressure gas emits a spectrum of light with discrete lines at wavelengths characteristic of the gas; and a cool low pressure gas in front of a hot source absorbs at those same characteristic wavelengths. So, they realized that, for sun light's spectrum, the heated interior of the sun emits a continuous (Planck) spectrum and the cooler photosphere does the line absorbing. This means that the atoms of the gas in the sun's atmosphere has acted as a filter for the light emitted from the sun's interior and absorbed its own characteristic wavelengths⁴³. The study of the absorption lines in the sun's spectrum lead scientists to deduce its chemical composition ⁴⁴, which turns out to be mostly hydrogen and after that mostly helium.



Figure 7: Fraunhofer's solar spectrum. The continuous line above the spectrum shows, the approximate the solar continuous intensity, as estimated by Frauhofer [19].

In 1885, J. Balmer, a lecturer at a girls college in Switzerland,.... formula for

 $^{^{41}}$ These lines were observed earlier by **William Wollaston** in 1802, but his interest was in the colors themselves, and not the dark lines, and so he did not provide any significance to them.

⁴²J. Von Fraunhofer, Denkschriften der Koniglichen Akademie der Wissenschaften zu Munchen, 1814-1815, 5, 193-226.

⁴³Kirchhoff showed that the D line (of wavelength 589 nm) in the sun's spectrum was due to the absorption of solar radiation by sodium atoms.

⁴⁴For instance, Pierre Janssen, a french astronomer, along with Norman Lockyer, an English astronomer, detected during a solar eclipse an unknown yellow spectral line signature in sunlight. The line had a wavelength of about 588 nm, slightly less than the so-called D-lines of sodium. Norman Lockyer suggested that this line is associated with an absorption line of some unknown solar element (no known element at that time gave this line) and he named it " Helium" from the greek word Helios for sun. Note that Helium was not observed on Earth until 1895.

the visible region of the atomic hydrogen spectrum, denoted by $H_{\alpha}, H_{\beta}, H_{\gamma}, \dots$

$$\frac{1}{\lambda} = \mathcal{R}_H \left(\frac{1}{2^2} - \frac{1}{n^2}\right); \quad n = 3, 4, 5, \dots$$
 (2.1)

where n = 1, 2, 3, ... label the lines $H_{\alpha}, H_{\beta}, H_{\gamma}, ...,$ and \mathcal{R}_H is a constant which has a value 10.97 μm^{-1} , which was later called the Rydberg constant for hydrogen. Few years later similar series , known as **Lyman**, **Paschen**, **Brackett** and **Pfund** series, and many others, were discovered. By 1900, **Johannes Rydberg** was able to describe all wave numbers of the lines of hydrogen in each of of theses series, including the Blamer series, by the formula

$$\frac{1}{\lambda} = \mathcal{R} \left(\frac{1}{n_1^2} - \frac{1}{n_2^2} \right); \quad n_2 = n_1 + 1, n_1 + 2, n_1 + 3, \dots$$
(2.2)

The constant $\mathcal{R} \simeq \mathcal{R}_H$, and it is also called Rydberg constant⁴⁵. Here $n_1 = 1, 2, 3, 4$, and 3 correspond to the Lyman, Balmer, Paschen, Brackett and Pfund series, respectively.

The relations (2.1) and (2.2) are empirical, and no theoretical model at that time could their origin. In order to make any progress in the understanding of the origin and the characteristic spectra of elements, required a model of the atom. Unfortunately, all the existing models pre-1913 were unsuccessful in explaining the origin of these lines.

2.2 Instability of Atoms

After the discovery of the electron⁴⁶ in 1897 by J. J. Thomson in 1897, it became clear that atoms are not the fundamental building blocks of nature. The neutrality of the atom indicated that it should contain an equal amount of positive charge as that of the electrons in it. This lead Thomson to propose that atom is a sphere where a positive charge is spread through it, like a "pudding" in which in which negative electron are suspended like "plum". This model is sometimes dubbed a "**plum pudding model of Thomson**" ⁴⁷.

in 1911 Ernest Rutherford tested Thomson's proposal by firing a beam of alpha particles⁴⁸ (of energy 7.68 MeV) from a Radium source through a thin (0.6 μm) piece

⁴⁵Although the value of \mathcal{R} is not exactly the same for different elements, the variation is small compared to \mathcal{R}_H . Now in modern physics, we know that this constant measures the strength of the binding between electrons and nuclei in atoms.

 $^{^{46}}$ the measurement of the size of its charge by Millikan in 1909. Combining Millikan results with the measured value of e/m by J. Thomson it was possible to deduce that the electron mass is about 1/2000 the mass of the ion of hydrogen.

⁴⁷Thomson model was proposed in 1904.

⁴⁸Alpha particle is the helium nucleus. It was discovered along with the β particle (which is the electron) by Ernest Rutherford in 1899.

of gold foil. If Thomson's model was correct, then the positive mass in the gold foil would be relatively diffuse and would allow the alpha particles to pass through the foil with only minor scattering. However, the results from that experiment showed most of the alpha particles passed through the gold foil without changing direction much as expected, but some of the alpha particles bounced back in the opposite direction, as though they had struck something dense and solid in the gold foil. So Rutherford came to the conclusion that an atom is mostly empty space and that its positive charge is concentrated at its core, forming a nucleus. He also noticed that in order to overcome the electrons must be moving around the nucleus like planets move around the sun. This is why the atomic model of Rutherford is sometimes called the **planetary model for the atom**⁴⁹.

However, the planetary model of Rutherford does not explain the spectral lines in the spectrum of atoms. Furthermore, the electrons are unstable in this model. Indeed, according to classical electromagnetic theory, an electron orbiting around the nucleus will emit radiation, and thus depletes its kinetic energy. As a result, the electron will spiral inward toward the nucleus and eventually collapse on it. For instance, if we find that the time of the collapse is large the age of the universe, ~ 14 billion years, then that might not be a serious problem. So let us see how long it would take an electron before it crashes on a nucleus. For that suppose that the electron is moving in a circular orbit of radius r around the a nucleus of charge Z e. Then according to Newton's second law we have

$$\frac{Ze^2}{4\pi\epsilon_0 r^2} = m_e \left| \frac{d^2 \vec{r}}{dt^2} \right| = \frac{m_e v^2}{r}$$
(2.3)

Using Eq (1.48), and the expression of the acceleration from the above equation, we find that the rate at which the electron loses energy by radiation reads

$$\frac{dE}{dt} = -\frac{2}{3} \frac{e^6}{\left(4\pi\epsilon_0 c\right)^3 m_e^2}$$
(2.4)

The left hand side of the above equation can also be re-written as

$$\frac{dE}{dt} = \frac{dE}{dr}\frac{dr}{dt} = \frac{dr}{dt}\frac{d}{dr}\left(\frac{1}{2}m_ev^2\right)$$

$$= \frac{dr}{dt}\frac{d}{dr}\left(\frac{Ze^2}{8\pi\epsilon_0 r}\right)$$

$$= \frac{Ze^2}{8\pi\epsilon_0}\frac{1}{r^2}\frac{dr}{dt}$$
(2.5)

⁴⁹It was until 1932, when **James Chadwick** demonstrated that inside the nucleus, besides the proton there are neutral particles. He used the Beryllium that is radioactive, the emanation from which is electrically neutral and was thought to be an electromagnetic radiation. Chadwick aimed this emanation at various materials, including paraffin (containing hydrogen), helium and nitrogen and studied the energies of the recoiling nuclei from different targets. He was able to conclude that this emanation which came from beryllium was a neutral particle with mass of the proton, which was called **neutron**.

Hence, we obtain

$$\frac{dr}{r^2} = -\frac{4}{3} \frac{e^4}{\left(4\pi\epsilon_0\right)^2 c^3 m_e^2} \frac{4}{3r^2} dt$$
(2.6)

Integrating both sides of the above equation gives

$$t = \frac{(4\pi\epsilon_0)^2 c^3 m_e^2}{4e^4} \left(r_0^3 - r^3\right)$$
(2.7)

where r_0 is the radius of the orbit at a the time t = 0. If we assume that $r_0 \sim 10^{-10} m$, we get the time it takes for the electron to crash on the nucleus is

$$T_{crash} = \frac{(4\pi\epsilon_0)^2 c^3 m_e^2}{4e^4} r_0^3 \simeq 10^{-10} \ s \ !!!$$
(2.8)

which is clearly in contradiction with the observed stability of the atoms.

2.3 Bohr's Model [1913]

In 1913, Bohr constructed a model for hydrogen-like atom, described by an electron in circular orbit atom with one electron in circular orbit around a nucleus with positive charge Z e^{50} . As we saw, earlier, the acceleration of the electron (circular orbit) will cause it to lose energy by electromagnetic radiation and spiral toward the nucleus. Bohr resolved this problem by assuming that the electron can move only in certain orbits, which he called stationary states. The total energy of the electron in a circular orbit of radius r is given by the sum of its kinetic and potential energy:

$$E_{tot} = \frac{1}{2}m_e v^2 + \frac{(-Ze^2)}{r^2} = -\frac{Ze^2}{2r} = -W$$
(2.9)

where we made used of Eq (2.3), and W represents the necessary energy to extract the electron from its orbit. The frequency of revolution can be expressed in terms the total energy of the electron as⁵¹

$$\nu_{rev} = \frac{\sqrt{2}}{\pi Z e^2} \frac{W^{3/2}}{\sqrt{m_e}}$$
(2.10)

So, in order for the electron to remain in a certain stationary state, it must have a certain corresponding energy. Then, Bohr, Inspired by Planck hypothesis that a

$$\nu_{rev} = \frac{4\epsilon_0\sqrt{2}}{\pi e^2} \frac{W^{3/2}}{\sqrt{m_e}}$$

 $^{^{50}\}rm{N}.$ Bohr, Philos. Mag. 26, 1 (1913). This is really a nice paper that I advise every student who is learning Quantum mechanics to read it.

 $^{^{51}\}mathrm{In}$ SI units, it reads
charged oscillator can emit or absorb energy only as a multiple $h\nu$, with ν is the oscillation frequency of the oscillator ,postulated that

$$W = \frac{n}{2}h\nu_{rev} \tag{2.11}$$

But why there is a factor of 1/2 in the above equation, instead of 1 as in **Planck's formula?** Bohr explains his choice of the factor of 1/2 as follows⁵². First, he assumes that the emitted energy is proportional to some, yet unkown, function f(n), with n being an integer :

$$W_n = f(n)h\nu_{rev} \tag{2.12}$$

where n now labels a stationary state. Substituting the expression of ν_{rev} in (2.10) into the above equation, we obtain

$$W_n = \frac{\pi^2 m_e Z^2 e^4}{2h^2 f^2(n)}; \qquad \nu_{rev}^{(n)} = \frac{\pi^2 Z^2 e^4 m_e}{2h^3 f^3(n)}$$
(2.13)

The system can not radiate energy while remaining in the stationary. However, Bohr assumed that when the electron moves from one stationary state to another, it emits radiation. The frequency of the emitted light from the transition from state n_1 to a state $n_2 > n_1$, is given by the formula

$$\nu = \frac{W_{n_2} - W_{n_1}}{h} = \frac{\pi^2 Z^2 e^4 m_e}{2h^3} \left[\frac{1}{f^2(n_2)} - \frac{1}{f^2(n_1)} \right]$$
(2.14)

He then noted that in order for the above expression to have the same form as the Balmer series, one must have f(n) = C n, where c is some constant. To determine c, he considered the transition between the state $n_1 = N$ and $n_2 = (N-1)$, which yields

$$\nu = \frac{\pi^2 Z^2 e^4 m_e}{2h^3 C^2} \left(\frac{2N-1}{N^2 (N-1)^2}\right)$$
(2.15)

Using (2.13), we find that the frequency of revolution of the electron before and after the emission read

$$\nu_{rev}^{(N)} = \frac{\pi^2 Z^2 e^4 m_e}{2C^3 h^3 N^3}; \qquad \nu_{rev}^{(N)} = \frac{\pi^2 Z^2 e^4 m_e}{2C^3 h^3 (N-1)^3}$$
(2.16)

which for very large value of N, the ratio of the between the two frequencies approaches 1. So, in this limit, according to ordinary electrodynamics, the ratio between the frequency of radiation ν and the the frequency of revolution $\nu_{rev}^{(N)}$ is also very close to 1. This implies that we must have C = 1/2, which justify the factor of 1/2 in the expression of W_n in (2.17). This implies that

$$W_n = \frac{2\pi^2 m_e Z^2 e^4}{h^2 n^2}; \qquad \nu_{rev}^{(n)} = \frac{4\pi^2 Z^2 e^4 m_e}{n^3 h^3}$$
(2.17)

 $^{^{52}}$ See page 13 in his 1913-paper given in the footnote.

Applying this result too the hydrogen atom, the frequency of radiation when the electron make the transition from state n_1 to n_2 is

$$\nu(n_1 \to n_2) = \frac{2\pi^2 e^4 m_e}{h^3} \left(\frac{1}{n_2^2} - \frac{1}{n_1^2}\right)$$
(2.18)

Then by setting $n_2 = 2$, and allowing n_1 to vary over larger values of integers, he reproduced the experimental Balmer formula for the lines in the spectrum of hydrogen. He also point out that $n_2 = 3$ (with n_1 taking values 4, 5, ...ect) corresponds to the infrared spectrum observed by Paschen in 1908, and for $n_2 = 1$ (or $n_2 = 4, 5...$) while letting n_1 takes larger values gives a spectral lines in the extreme ultraviolet (extreme infrared), which were not yet observed at that time.

For $n_1 >> (n_1 - n_2) = \Delta n$, then we can Taylor expand the term $1/(n_2)^2$ in (2.18) to the first order in Δn and find that

$$\nu(n_2, \Delta n) = \frac{4\pi^2 e^4 m_e}{h^3 n_2^2} \Delta n$$
(2.19)

Which for $\Delta n = 1$ gives the orbital frequency, i.e. $\nu = \nu_{rev}$. Thus, in this limit, the orbital frequency, which in classical theory gives rise to radiation, coincide with the frequency associated with the difference of the quantum states of the electron. For transitions with $\Delta n > 1$ between states with large values of n such that $n >> \Delta n$, he noted that since the electron motion is periodic, its position can be represented by a Fourier series as

$$x(t) = \sum_{\tau=1}^{\infty} C_{\tau} \cos \tau \omega(n)$$
(2.20)

where each of these terms in the sum is known as a harmonic, and the τ^{th} harmonic has amplitude C_{τ} and a frequency $\tau\omega(n) \equiv \nu(n, \tau)$, with ω is the fundamental frequency. Then, according to classical electrodynamics, the frequencies the radiation emitted by the atom is given by the frequencies in the harmonics, i.e. $\omega, 2\omega, 3\omega, \dots \text{ ect}^{53}$. So, Bohr considered that the frequency associated with the transition from state n to n', is equal to the frequency of the τ^{th} harmonic of the classical motion in the n' stationary, i.e.

$$\nu_{quantum} \equiv \nu(n, n - \tau) = \nu(n, \tau) \equiv \nu_{classical}; \quad \tau = n - n'$$
(2.21)

where $here^{54}$

$$\nu(n, n - \tau) = \frac{1}{h} \left[E(n) - E(n - \tau) \right]; \quad \tau = 1, 2, \dots$$
 (2.22)

 $^{^{53}}$ As it was pointed out in [20], strictly speaking the motion of an electron emitting radiation is not periodic; but rather is a spiral toward the nucleus. However, if the initial energy of the electron is very large compared to the energy lost, so that the energy radiated is negligible, then it can be well approximated as being periodic.

⁵⁴Here E_n is the binding energy of the stationary state with the quantum number n, i.e. $E_n = -W_n < 0$.

Bohr called theses correspondences the **correspondence principle**, which are valid for large n and small τ . As we will see later that this principle can be used to to determine the intensity and polarization of the emitted radiation in the quantum theory associated with the transition from one state to another.

The quantum-theoretic frequency in (2.22) can be re-written as

$$\nu(n, n - \tau) = \frac{1}{h} \left[H[J(n)] - H[(J(n - \tau))] \right]; \quad \tau = 1, 2, \dots$$
 (2.23)

where H is the Hamiltonian of the system which is represents the total energy, and J is the action variable (see the Appendix), which in Bohr's theory is equal to nh. For n very large and τ small, we can Taylor expand the right-hand side of the above expression, and obtain

$$\nu(n, n-\tau) = \frac{1}{h} \frac{dH}{dJ} \Delta J = \tau \frac{dH}{dJ}; \qquad n >> \tau$$
(2.24)

Now, in the classical theory, for a periodic motion the frequency of the α^{th} harmonic can be calculated using the Hamiltonian formalism, and it is given by

$$\nu_{\alpha} = \alpha \frac{dH}{dJ} \tag{2.25}$$

where $\nu_{\alpha} = \alpha \nu(n)$. Thus, in the classical theory the formula of the frequencies became the same as in the quantum theories in the limit where n is very large, which is consistent with Bohr's principle. Furthermore, based on (2.23), (2.23), and (2.23), Born suggested a prescription to translate from classical formulas into their corresponding quantum-theoretic analogues by making the following replacement:

$$\tau \frac{\Phi(n)}{dn} \longleftrightarrow \Phi(n) - \Phi(n-\tau)$$
 (2.26)

where $\Phi(n)$ can be any arbitrary differentiable function defined for a stationary state n. The above recipe, which is sometimes referred to as **Born's correspondence rule** [2], and plays an important role in Heisenberg's discovery of the matrix mechanics. It has been generalized by Kramer and Heisenberg to the form⁵⁵

$$\tau \frac{\Phi(n,\tau)}{dn} \longleftrightarrow \Phi(n+\tau,n) - \Phi(n,n-\tau)$$
(2.27)

 $^{^{55}\}mathrm{H.}$ Kramers and W. Heisenberg, "Uber die Streuung durch Atome ("On the dispersion of radiation by atoms"), Zeitschrift fur Physik, **31**, 681- 708 (1925).

Bohr was also able to deduce the value of the Rydberg constant in terms of the measured values of e, m_e , and h^{56} :

$$\mathcal{R}_H = \frac{2\pi^2 e^4 m_e}{c \ h^3} \tag{2.28}$$

Here c denotes the speed of light. Using the measured values of e, e/m and h at that time, gives $\mathcal{R}_H = 1.097 \times 10^7 \ m^{-1}$, in agreement with the experimental value. Furthermore, from Eq (2.17), he calculated the energy of electron in the n = 1 state to be 13 eV, not very different from the precise value of the energy of the lowest energy state of hydrogen (13.6 eV). So, the energy of the n^{th} stationary state is⁵⁷

$$E_n = -W = -\frac{13.6 \ eV}{n^2}$$
(2.29)

Bohr's condition for the stationary states in (2.17), can be translated into a quantization rule for the angular momentum. This can be seen by writing the angular momentum L as

$$L = \frac{m_e v^2 / 2}{\pi \nu_{rev}} = \frac{W}{\pi \nu_{rev}}$$
(2.30)

Since the kinetic energy is equal to W, the above expression reads

$$L = n \hbar \tag{2.31}$$

where we have defined the constant $\hbar = \frac{h}{2\pi} = 1.05457 \times 10^{-34} J.s.$ So, we see that for the electron to arrange itself in some stationary states, requires the quantization of the orbital angular momentum⁵⁸.

$$\mathcal{R}_H = \frac{\pi^2 e^4 m_e}{8\epsilon_0^2 c \ h^3}$$

⁵⁷For a hydrogen-like atom in which the nucleus contains Z protons, the energy of the n^{th} state is

$$E_n = -W = -\frac{13.6\ Z^2\ eV}{n^2}$$

⁵⁸So, in his original paper, Bohr arrived at the quantization of the angular momentum using the relation $W = \frac{n}{2} h \nu_{rev}$; not as a starting assumption. It is interesting to note that the quantization of angular momentum was used first in 1912 by J. Nicholson who was to trying to understand some set of spectral lines observed in nebula and in the sun's corona. He noticed that given that the angular momentum has the same dimension as Planck constant, and suggested that the quantization of angular momentum. This was acknowledged by Bohr in his 1913 -paper.

⁵⁶To get the expression of \mathcal{R}_H written in SI units one just replace e^2 by $e^2/4\pi$ in the (2.28), and we get

In addition to the Balmer, it was known that there is a different type of spectral series called **Pickering series**, found in 1896 by Edward Pickering in the star ζ Puppis. These lines can arranged in a form analogue to Balmer's series using half-integers instead of integers, and were, erroneously, thought to be a second spectrum of hydrogen. In 1912, **Alfred Fowler** showed that similar lines can be produced in a laboratory mixture of hydrogen and helium gas. Bohr noticed that they have the same spectrum of spectral lines as of hydrogen but with wavelength four times shorter⁵⁹ and argued that they could correspond to the spectral lines of **ionized helium**⁶⁰. In Bohr's theory this means that the corresponding Rydberg constant for the helium, \mathcal{R}_{He} , is four times \mathcal{R}_H for the hydrogen atom. However, Fowler was not convinced and sent a letter to Nature⁶¹ where he pointed out that the ratio $\mathcal{R}_{He}/\mathcal{R}_H$ is not simply a factor of 4, but instead is 4.0016. Bohr understood that this small discrepancy is due to the fact that one neglected the effect of the finite mass of the nucleus. So, in the above expressions of W_n and ν_{rev} , the electron mass must be replaced by the reduced mass of the atom or the ion:

$$\mu_N = \frac{m_e M_N}{M_N + m_e} \tag{2.32}$$

where N denotes the nucleus, and M_N its mass. This implies that

$$\frac{\mathcal{R}_{He}}{\mathcal{R}_H} = \frac{\mu_{He}}{\mu_H} = 4\left(\frac{1 + \frac{m_e}{M_H}}{1 + \frac{m_e}{M_{He}}}\right) = 4.00160 \tag{2.33}$$

In a very good agreement with the experimental data. Impressive !!⁶²

2.4 Sommerfeld Extension of Bohr's Model [1915 – 1919]

Although, Bohr's model provided an explanation for the origin of the spectral lines observed in hydrogen and hydrogen like atoms, it failed when applied to systems with more than one electron. For instance, it could not account for the spectrum of **neutral helium** atom. Even for the hydrogen atom, it was known at time⁶³ that the H_{α} and H_{β} lines of the Balmer series were composed of more than one line, i.e. they displayed some narrow line splittings which is incompatible with Bohr's theory⁶⁴.

⁵⁹This can be seen by substituting in the expression (2.18)substitute e^4 by Z^2e^2 , and which for helium Z = 2.

 $^{^{60}}$ In 1913, while Bohr still developing his theory, E. Evans showed that Pickering lines can be detected using a pure helium sample. He reported his results in a letter to Nature: E. Evans, "The spectra of helium and hydrogen", Nature **92**: 5 (1913).

⁶¹A. Fowler, "The spectra of helium and hydrogen", Nature **92**: 95 (1913). Note the paper was published in same month (september) as Evan's paper and with the same title.

⁶² It is said that when Einstein was told about this result, he responded: "This is an enormous achievement. The theory of Bohr must be right".

⁶³Michelson, Philosophical Magazine (4), **31**, 338- 346(1891).

⁶⁴ This problem, although recognized by the time Bohr's announced his theory, but was not regarded as a very strong argument against Bohr's theory in view of its remarkable success and the fact the splittings was very small.

Arnold Sommerfeld, suspected that is due to the fact that Bohr's model was an approximation, as the quantization condition of the angular momentum involved a single quantum number. So, in 1915, Sommerfeld proposed a generalization of Bohr's model to two degrees of freedom where the electron orbit is elliptical rather than circular, and hence, now there are two independent parameters r and ϕ . He and Bohr also realized that the angular momentum quantization (2.31) could be also written as

$$\int_0^{2\pi} p_\phi \, d\phi = nh \tag{2.34}$$

where ϕ is the angular coordinate describing the position of the electron in its circular orbit, and p_{ϕ} is the angular momentum, which in Hamiltonian mechanics represents the canonical conjugate momentum of ϕ . This lead Sommerfeld to postulate that stationary states of a periodic system with f degrees of freedom are determined by the condition that **phase integral for every coordinate is an integral multiple of the quantum action**⁶⁵, i.e.

$$\oint p_k \ dq_k = n_k h; \qquad k = 1, 2, \dots f$$
(2.35)

where q_k is one of the coordinate describing the position of the electron, p_k is the momentum conjugate associated with that coordinate⁶⁶, n_k is a non-negative integer, and the integration is taken over one period of the coordinate.

For the case of elliptical orbits, we have two quantization conditions

$$\oint p_{\phi} d\phi = n_{\phi}h; \qquad \oint p_r d\phi = n_r h \tag{2.36}$$

This system of electron in the potential V(r) of a nucleus is described by the Lagrangian

$$L = \frac{m_e}{2}v^2 + V(r)$$
(2.37)
= $\frac{m_e}{2}\left(\dot{r}^2 + r^2\dot{\phi}^2\right) + V(r)$

from which we can calculate the conjugate momenta of the coordinates ϕ and r (see the footnote) and find

$$p_{\phi} = m_e r^2 \dot{\phi}; \qquad p_r = m\dot{r} \tag{2.38}$$

which corresponds to the azimuthal angular momentum and the radial momentum of the electron, respectively. Since p_{ϕ} is a constant of motion⁶⁷, the quantization condition

$$\frac{\partial L}{\partial \phi} - \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{\phi}} \right) = 0 \tag{2.39}$$

⁶⁵A. Sommerfeld, Munchener Berichte, pp. 425- 458 (1915); Annalen der Physik 51, 1- 94 (1916).

⁶⁶For a system that is described by a Lagrangian $L(q_k, \dot{q}_k)$, the momentum conjugate associated with the coordinate q_k is $p_k = \frac{\partial L}{\partial \dot{q}_k}$.

 $^{^{67}\}mathrm{Euler}\text{-}\mathrm{Lagrange}$ equation of motion for the coordinate ϕ reads

for the azimuthal angle is

$$p_{\phi} = n_{\phi}\hbar \tag{2.41}$$

which is exactly the same as the quantization condition given by Bohr for circular orbit.

Now, we would like to derive the quantization condition p_r . For that, we first recall that the equation for an ellipse with a semi-major axis a is given by

$$r = \frac{a\left(1 - \epsilon^2\right)}{1 + \epsilon \cos\phi} \equiv \frac{A}{1 + \epsilon \cos\phi}$$
(2.42)

where r is the distance from a focus to a point on the ellipse, ϵ is the eccentricity of the ellipse, and the quantity $A = a (1 - \epsilon^2)$ is called the semi-latus rectum. With the help of the expression of p_{ϕ} in (2.38), we can write the integrand $p_r dr$ in (2.36) as

$$p_r dr = \left(m_e \frac{dr}{d\phi} \dot{\phi}\right) \frac{dr}{d\phi} d\phi = p_\phi \left(\frac{1}{r} \frac{dr}{d\phi}\right)^2 d\phi \qquad (2.43)$$

Taking the logarithm on both sides of Eq(2.42), then differentiating with respect to ϕ gives

$$\frac{1}{r}\frac{dr}{d\phi} = \frac{\epsilon\sin\phi}{1+\epsilon\cos\phi} \tag{2.44}$$

After substituting the above expression into in (2.43), the the quantization condition reads

$$\oint p_r \, dr = p_\phi \epsilon^2 \int_0^{2\pi} \frac{\epsilon \sin^2 \phi}{\left(1 + \epsilon \cos \phi\right)^2} d\phi = n_r h \tag{2.45}$$

or, equivalently,

$$\frac{n_r}{n_\phi} = \frac{\epsilon^2}{2\pi} \int_0^{2\pi} \frac{\epsilon \sin^2 \phi}{\left(1 + \epsilon \cos \phi\right)^2} d\phi \tag{2.46}$$

where we made use of the quantization condition on p_{ϕ} given in Eq(??). Evaluating the integral, yields

$$1 - \epsilon^2 = \frac{n_{\phi}^2}{\left(n_{\phi}^2 + n_r^2\right)}$$
(2.47)

Since L is independent of ϕ (in this case we say that ϕ is a cyclic coordinate), we get

$$\frac{d}{dt}\left(p_{\phi}\right) = 0\tag{2.40}$$

So, p_{ϕ} is a constant of motion.

which shows that not all eccentricities are allowed. Since the coordinate r is function of ϕ , we can express the total energy of the electron solely in terms of angle ϕ , and get

$$E_{tot} = \frac{p_{\phi}^2}{m_e a^2 \left(1 - \epsilon^2\right)^2} \left[\frac{1 + \epsilon^2}{2} + \epsilon \cos\phi\right] - \frac{Ze^2}{4\pi\epsilon_0} \frac{1 + \epsilon\cos\phi}{(1 - \epsilon^2)}$$
(2.48)

However, we know that for time-independent potential, the energy of a system is a constant of motion, and therefore the therms in front of $\cos \phi$ must add up to zero. This implies that the semi-major axis is given by

$$a = \frac{4\pi\epsilon_0 p_\phi^2}{Ze^2 \left(1 - \epsilon^2\right)} \tag{2.49}$$

or, equivalently,

$$a_n = \frac{4\pi\epsilon_0\hbar^2}{Ze^2m_e} (n_\phi + n_r)^2 \equiv \frac{4\pi\epsilon_0\hbar^2}{Ze^2m_e} n^2$$
(2.50)

which has exactly the same form as the energy derived by Bohr for circular orbit but with the difference that for elliptical orbits the quantum number n is replaced by $(n_{\phi} + n_r)$. So, now the emitted photon in the transition between two quantum states represented by $(n_{\phi} + n_r)$ and $(n'_{\phi} + n'_r) > (n_{\phi} + n_r)$, has frequency

$$\nu = \frac{Ze^4 m_e}{8\epsilon_0^2 h^3} \left[\frac{1}{(n_\phi + n_r)^2} - \frac{1}{(n'_\phi + n'_r)^2} \right]$$
(2.51)

The state with $n_{\phi} = 0$ gives to $\epsilon = 1$, which is a straight line joining the two foci of the ellipse. This means the electron will pass through the nucleus, therefore this possibility was excluded. So, the lowest energy state corresponds to $(n_{\phi}, n_r) = (1, 0)$, which yields $\epsilon = 0$, i.e. in this state the electron will have a circular orbit. For the first excited energy state $(n_{\phi} + n_r) = 2$, there are two possible orbits: circle $((n_{\phi}, n_r) = (2, 0))$, and an ellipse $((n_{\phi}, n_r) = (1, 1))$. In the second excited state n = 3, the electron can be in one of the three possible orbits: circle $((n_{\phi}, n_r) = (3, 0))$, and three ellipses with different ellipticity $((n_{\phi}, n_r) = (2, 1), (1, 2))$. So for Z = 1, Eq (2.51) gives the the spectral series of the hydrogen atom, but with the number of ways, also called degeneracy, in which the transition can be produced has been greatly increased. For example, for hydrogen (Z = 1), the transition between the states n = 2 and n = 3, which corresponds to the line H_{α} has a degeneracy $3 \times 2 = 6$. The degeneracy is even higher for the lines $H_{\beta}, H_{\gamma}, H_{\delta}$, and so on. Since not all of these transitions are realized in nature, Sommerfeld suggested some empirical selection rules on the possible spectral lines.

Finally, one of the important results of Sommerfeld extension of the Bohr's model is the calculation of the relativistic expression of the energy levels of hydrogen-like atoms.

2.5 Einstein Coefficients [1916]

In 1916, Einstein studied the transition and absorption of radiation by matter using statistical mechanics. He considered a quantum system consisting of a large number of molecules which occupy a discrete set of states which, in his original paper⁶⁸, he denoted them by $Z_1, Z_2, ...$, ect, with the corresponding energies $\epsilon_1, \epsilon_2, ...$ ect⁶⁹, and postulated that the number of transitions, in time dt, from state Z_m to a state Z_n will consist of two components. The first component corresponds to the **spontaneous emission**⁷⁰, which can occur without the need for an external field. The number of transitions for this kind of emission is given by

$$dN^{(spont)} = A_m^n N_m dt \tag{2.52}$$

where N_m denotes the number of molecules. The second component. The second component arises from the **stimulated emission**, produced when when a particle in an excited state interacts with an incident photon and make the transition from a higher to a lower energy state⁷¹. The number of transitions in stimulated emission reads

$$dN^{(stim)} = B_m^n N_m u \ dt \tag{2.53}$$

where u is the energy density of the incident radiation⁷². The presence of this radiation will also induce jumps from the state Z_n to the state Z_m if a molecule in the lower state absorbs a photon of frequency $\nu = (\epsilon_m - \epsilon_n)/h$. In this case e the number of transitions as a result of photon absorption is

$$dN^{(abs)} = B_n^m N_n u \ dt \tag{2.54}$$

The coefficients A_m^n, B_m^n and B_n^m , called Einstein coefficients, are assumed to depend only on the properties of the molecules, and not the radiation. A schematic of the three types of emissions is shown in Fig. 8.

If the system is in a thermodynamic equilibrium at temperature T, then the energy density of the radiation is a function of temperature, i.e. u = u(T). Also, in equilibrium, the rate at which molecules make a transitions from the state Z_n to Z_m must be equal to the transition rate from Z_m to Z_n :

$$N_m \left[A_m^n + B_m^n u(T) \right] = N_n B_n^m u(T)$$
(2.55)

⁶⁸A. Einstein, "Strahlungs-Emission und -Absorption n ach der Quantentheorie" [Emission and Absorption of Radiation in Quantum Theory]. Verhandl. D. Deutch. Phys. Ges., Vol 18, pp. 318-323 (1916).

⁶⁹These are internal energies that a molecule can occupy apart from its translational and rotational motion.

⁷⁰Note Einstein never used the term "spontaneous emission" in his paper. It was Bohr who interpreted such transition as a spontaneous in the sense of "acausal".

⁷¹It's like the photon "convinces" the particle in some excited state to jump to a lower energy state. This process is accompanied with the emission of a photon which will have a frequency at or near the frequency of the transition, i.e. $h\nu \simeq (\epsilon_m - \epsilon_n)$.

⁷²The incident radiation is not necessarily a black-body radiation.



Figure 8: A schematic of the types of radiation emitted between two energy states (taken from [17]).

The number of molecules that occupy a state Z_j is proportional to the probability \mathcal{W}_j , which according to classical statistical mechanics is given by the Boltzmann distribution

$$\mathcal{W}_j = g_j \exp\left(-\frac{\epsilon_j}{k_B T}\right) \tag{2.56}$$

where the factor g_j is the degeneracy of the quantum state Z_j^{73} . This implies that the equilibrium relation (2.55) becomes

$$g_m \ e^{-\epsilon_m/k_B T} \left[A_m^n + B_m^n \ u(T) \right] = g_n \ e^{-\epsilon_n/k_B T} B_n^m \ u(T)$$
(2.57)

Assuming that when $T \to \infty$, the energy density of the radiation goes to infinity, one can neglect the coefficient A_m^n compared to the $B_m^n u$, and obtain

$$g_m \ B_m^n = g_n \ B_n^m \tag{2.58}$$

Using this relation in (2.55), Einstein deduced that the spectral energy density u is given by

$$u(T) = \frac{A_m^n / B_m^n}{e^{(\epsilon_m - \epsilon_n)/k_B T} - 1}$$

$$(2.59)$$

In the high frequency range, i.e. $h\nu >> k_B T$, the above expression reduces to Wien radiation law in (1.39), whereas for low frequencies, i.e. $h\nu << k_B T$, it and produces

⁷³It is also called the statistical weight of the state Z_j .

the In the Rayleigh- Jeans formula in (1.46). So Einstein concluded that

$$E_m - E_n = h\nu \tag{2.60}$$

which corresponds to the Bohr's frequency condition, and that

$$A_m^n = \frac{8\pi h\nu^3}{c^3} B_m^n \tag{2.61}$$

Substituting (2.60) and (6.54) into (2.59), gives the Planck's energy density distribution. The above relation can be used to

Despite the success of Bohr's model and its extension by Sommerfeld, it failed to explain the spectrum of atoms with more than one electron, such as the helium atom, and the anomalous Zeeman effect. As suspected by Pauli and Heisenberg and Born, the root of the failure of the Bohr's theory was the orbit paradigm. This lead to the development of the matrix mechanics by Heisenberg in 1925 and the wave mechanics by Schrodinger in 1926 which naturally explain the quantization of the atomic energy levels as well the fine structure of the spectral lines. Furthermore, as will be discussed later, both the Heisenberg and the Schrodinger formulations of the quantum theory turn out to be equivalent.

3 Wave Mechanics [1923 – 1927]

3.1 Wave-Particle Duality and Davisson-Germer Experiment [1923 – 1927]

In the previous chapter, we have seen that the photoelelectric and Compton effect provided a clear evidence that light manifests not only as a wave, but also as a stream of particles, called photons. For an electromagnetic wave with frequency ν , the photon energy is $E_{\gamma} = h\nu = \hbar\omega$.

In 1923, Louis de Broglie, suggested that a particle of mass m, such as the electron, can be regarded as a wave with frequency and wave number given by⁷⁴

$$\omega = \frac{E}{\hbar}; \qquad \vec{k} = \frac{\vec{p}}{\hbar} = \frac{m\vec{v}}{\hbar}$$
(3.1)

which means that to each particle one associates a de Broglie wavelength

$$\lambda_d = \frac{h}{p} = \frac{h}{m\upsilon} \tag{3.2}$$

In using p = mv it was assumed that the particle is non-relativistic. A generalization to the case of a relativistic particle is straightforward; one just uses $p = mv/\sqrt{1 - v^2/c^2}$ for the momentum of the particle in the above equation. In the table below, we show the de Broglie wavelengths for electrons, and α (Helium nuclei) particles for different kinetic energies.

Now, if particles such as electrons can manifest not only as a corpuscule but also as a wave, they should exhibit diffraction effects. Indeed, that what has been shown in 1927 in a scattering experiment by C. J. Davisson and L. H. Germer ⁷⁵ at the Bell Laboratories in USA⁷⁶. They directed a beam of low energy electrons ($E_e = 54 \ eV$)

⁷⁴de Broglie, Louis, "The Beginnings of Wave Mechanics" in "Wave mechanics the first fifty years" edited by Price, William; Chissick, Seymour; Ravensdale, Tom . Halsted Press (Wiley) 1973. p. 13.

⁷⁵C. Davisson and L. Germer, Phys. Rev, Volume 30, 6, 705-740 (1927).

 $^{^{76}}$ In the same year, few months after Davisson-Germer experiment, G. P. Thomson (the son of J. J. Thomson) and his assistant Reid also observed that by passing electrons with kinetic energy 50 keV through very thin gold foil (they also did similar experiment with Aluminum), they exhibit diffraction effect.

onto a nickel crystal , and measured the intensity of the scattered electron at some angle that can be varied using a movable detector. They found that there was a peak in the intensity of the outgoing electron beam at an angle $\phi = 50^{(0)}$ and a minimum at $\phi = 35^{\circ}$. By varying the the kinetic energy of the incident electrons⁷⁷, the y could vary the electron wavelength and produce maxima and minima at different locations for the intensity. This was similar to the diffraction pattern predicted by W. Bragg for x-ray scattering from a crystal. .A schematic of Davisson-Germer experiment is shown in Fig. 6. So the existence of peaks in the intensity can be explained as due to constructive interference of waves scattered from atoms in different planes of the crystal. The condition for a maximum constructive interference was established by W. Bragg, and it is given by

$$2d\sin\theta = n\ \lambda; \quad n = 1, 2, \dots \tag{3.3}$$

where θ is the angle of incidence ... In this experiment, only the first order diffraction n = 1 is observed. Thus, using the lattice spacing of the atoms in nickel crystal $d = 0.9 \ A^0$, and the measured scattering angle $\phi = 50^0 = (\pi - 2\theta)$ (see the fig. 6) where the peak occurs, they calculated the wavelength that could produce such a maximum and found

$$\lambda = 2 \times 0.92 \ A^0 \times \sin 65^0 = 1.67 \ A^0 \tag{3.4}$$

Now, the momentum of the electron when it leaves can be obtained from the equation of conservation of energy t

$$\frac{p_e^2}{2m_e} = |e|U_e \tag{3.5}$$

where U_e is the accelerating potential. So, according to de Broglie, the wavelength of the electron is

$$\lambda_d = \frac{h}{\sqrt{2|e|m_e V_e}} = \frac{12.26 \ A^0}{\sqrt{U_e/Volt}}$$
(3.6)

Using $U_e = 54 V$ in the above expression gives $\lambda_d = 1.668 A^0$, which is in excellent agreement with the value inferred from observation and provides a convincing evidence of the wave nature of electrons as postulated by de Broglie.

Clinton Joseph Davisson and George Paget Thomson⁷⁸ shared the Nobel Prize in Physics in 1937 for " their experimental discovery of the diffraction of electrons by crystals"[8].

A modern device that relies on the wave property of electrons is the **transmission** electron microscope (TEM). It is similar to the "optical microscope", but has much

 $^{^{77}\}mathrm{By}$ applying different accelerating voltages from 44 V to 68 V.

⁷⁸It is interesting to note that J. J. Thomson discovered the electron as a "particle" in 1897, while his son George Thomson confirmed the wave character of the electron in 1927.



Figure 9: A schematic of of Davisson-Germer experiment (up) and the intensity of the outgoing electron as a function of the scattering angle (down). Figure taken from [18].

Kinetic Energy	Electron	Neutron	α -particle
$3 imes 10^{-2} {\rm eV}$	$71 A^0$	$1.65 \ A^0$	$0.83 \ A^0$
1 eV	$12.3 A^0$	$0.28 \ A^0$	$0.14 \ A^0$
10 keV	$0.12 A^0$	$3 \times 10^{-3} A^0$	$10^{-3} A^0$

higher resolving power⁷⁹. This is because the electrons can be accelerated by an electric potential to a very kinetic energies, getting a short wavelengths which can be as small as the x-ray radiation. The speed of these electrons approaches light velocity, and relativistic effects must be added. If U_e denotes the accelerating electric potential, then the de Broglie wavelength including the relativistic effects is given by

$$\lambda_d = \frac{h}{\sqrt{2m_e |e|U_e \left(1 + eU_e/2m_e c^2\right)}}$$
(3.7)

In a 100 k V accelerating potential (i.e electron kinetic energy of 100 keV), the wavelength of electron is of the order 0.3 pm. Today, TEM have reached resolution of better

 $^{^{79}{\}rm The}$ first TEM was built in 1931 by Ernst Ruska at the university of Berlin, Germany. He was awarded the Nobel Prize for Physics in 1986.

than 0.05 nm, more than 40,000 times better than a typical optical microscope⁸⁰.

3.2 Schrodinger Equation [1926]

In 1925, Peter Debye invited Erwin Schrodinger to give a seminar at Zurich university on the de Broglie wave-particle duality. After the talk has finished, Debye told to Schrodinger that this wave-particle duality approach is somehow "childish" and to deal properly with waves one "had to have a wave equation". Perhaps, this is what stimulated Schrodinger and lead him in to the discovery of the equation of the de Broglie waves in 1926, now known as **Schrodinger equation**⁸¹. The original derivation of Schrodinger equation used the analogy between the Hamiltonian formulation of classical mechanics and optics, which I give the details in the appendix. Here we give a more simpler derivation, but non rigorous means.

The energy of a free non-relativistic particle of mass m, momentum \vec{p} is given by

$$E = \frac{p^2}{2m} \tag{3.8}$$

The above equation does not change by multiplying its both sides by a wave function ψ :

$$E \ \psi(\vec{x},t) = \frac{p^2}{2m}\psi \tag{3.9}$$

If we assume that the plane wave $\psi_{PW}(\vec{x},t) = e^{i(\vec{k}\cdot\vec{x}-\omega t)}$ is the wave function for a particle of momentum $\vec{p} = \hbar \vec{k}$ and energy $E = \hbar \omega$, then we have

$$i\hbar \frac{\partial}{\partial t} \psi_{PW}(\vec{x}, t) = E \ \psi_{PW}(\vec{x}, t)$$

$$-i\hbar \vec{\nabla} \ \psi_{PW}(\vec{x}, t) = \vec{p} \ \psi_{PW}(\vec{x}, t)$$

$$(3.10)$$

⁸⁰The resolution, δ , in a perfect optical system can be calculated using Abbe's equation:

$$\delta = \frac{0.612 \ \lambda_d}{n \sin \alpha}$$

where n is index of refraction of medium between point source and lens relative to free space, which is approximately equal to unity, and α is half aperture angle in radians. For TEM, the angle α is very small, of the order 10^{-2} radians. Using the expression of λ_d in (3.6) (if we neglect the relativistic effects which is a good approximation for $U_e < 100kV$), and using the approximation $\sin \alpha \simeq \alpha$, we find that

$$\delta = \frac{0.753}{\alpha \sqrt{U_e/Volt} \ nm}$$

For an accelerating potential $U_e \sim 100 kV$, we get a resolution $d \simeq 0.2 nm$.

⁸¹ Erwin Shcrodinger, Phys. Review, Vol. 28, N0. 6, p (1927). This paper is based on a series of papers that Schrodinger has written in German in 1926.

Substituting theses equation in (3.9), we get⁸²

$$i\hbar\frac{\partial}{\partial t}\psi_{PW}(\vec{x},t) = -\frac{\hbar^2}{2m}\nabla^2\psi_{PW}(\vec{x},t)$$
(3.11)

However, for a particle in a potential $U(\vec{x})$, the wave function will no longer be a simple plane wave, since the wavelength, determined through the de Broglie relation, varies with the potential. It turns out that the appropriate wave function is given by

$$i\hbar \frac{\partial}{\partial t} \psi(\vec{x}, t) = \mathcal{H} \psi(\vec{x}, t)$$
 (3.12)

where

$$\mathcal{H} = \left(-\frac{\hbar^2}{2m}\nabla^2 + U(\vec{x})\right) \tag{3.13}$$

is the Hamiltonian operator of a particle of mass in a potential U. Equation (3.12) is known as the **time-dependent Schrodinger equation**, or the **TDSE** for short. Note that because of the presence of "i" in the TDSE, the solution, $\psi(\vec{x}, t)$, is generally a complex function.

If the potential time-independent, we can use separation of variables, and write

$$\psi(\vec{x},t) = e^{-i\omega t} \phi(\vec{x}) \tag{3.14}$$

which after substituting into (3.12), we obtain

$$\left(-\frac{\hbar^2}{2m}\nabla^2 + U(\vec{x})\right)\psi(\vec{x},t) = E \ \psi(\vec{x},t)$$
(3.15)

This equation is called the **time-independent Schrodinger equation** (TISE).

To test his equation, Schrodinger applied it to the case of a particle in a Coulomb potential, such as the electron in the hydrogen atom. He computed the hydrogen spectral series by treating the representing the electron by a wave function moving in the Coulomb potential created by the proton. He found that the allowed values of the energy and momentum were the same as those of the Bohr model which was an evidence that it was the correct wave equation⁸³.

Erwin Schrodinger shared with Paul Dirac the Nobel Prize in Physics in 1933 .

$$\nabla^2 \psi - \frac{1}{v^2} \frac{\partial^2}{\partial t^2} \psi = 0$$

 $^{^{82}\}mathrm{Note}$ that this is different from a normal wave equation, such as of a vibrating string, or the electromagnetic wave, given by

where v is the speed at which the wave travels. In this case, one needs to specify both ψ and $\partial \psi / \partial t$ as a boundary conditions.

⁸³Erwin Schrodinger, Annalen der Physik: 361-377 (1926).

3.3 Interpretation of the Wave Function [1926]

in 1926, Max Born suggested to interpret the square of the wave function associated with a particle, $|\psi(\vec{x},t)|^2 \equiv \rho$, as the probability density to find the particle at location \vec{x} at time t⁸⁴. In other words, it is the probability density distribution. For instance, the probability of finding a particle at a position between the points \vec{x} and $(\vec{x} + d\vec{x})$ is given by

$$dP(\vec{x}) = \psi(\vec{x}, t)\psi^*(\vec{x}, t) \ d^3\vec{x}$$
(3.16)

where $d^3\vec{x}$ is the element of volume centered around the point \vec{x} . In a two-slit experiment, it was found that this interpretation correctly predicts the probability distribution of electrons hitting a screen after passing through the two slits.

So, with this probabilistic interpretation of the wave function, the following properties must be satisfied:

- 1. $\phi(\vec{x})$ and its first derivative must be **finite**, **single valued**, and **continuous** every where. However, if the potential is not a function but a distributions, such as δ distribution, the derivative does not have to be continuous.
- 2. ϕ must be square integrable, i.e. $\int |\phi(\vec{x})| d^3x = 1$, which is equivalent to saying that for a particle in three dimensional space, there is 100% chance to find it some where within the whole space. This is called the **normalization** condition. However, it is important to note that as long as the integral is a finite constant N, one can always normalize it to unity by defining the wave function to be ϕ/\sqrt{N} .
- 3. If $\phi_1(\vec{x})$ and $\phi(\vec{x})$ are two normalized independent solutions to the TISE, then the **linear** combination $(c_1\phi_1 + c_2\phi_2)$ with $|c_1|^2 + |c_2|^2 = 1$ is also a solution. This is because the TISE is a linear differential equation.

Now, taking the time derivative of the probability density distribution $\rho = \psi^*(\vec{x}, t)\psi(\vec{x}, t)$ of a particle of mass m in a potential $U(\vec{x})$, and using the TDSE (3.12) gives

$$\frac{\partial}{\partial t} \rho = \frac{i}{\hbar} \left[(\mathcal{H}\psi^*)\psi - \psi^*(\mathcal{H}\leftarrow) \right]$$

$$= -\frac{i\hbar}{2m} \left[(\Delta\psi^*)\psi - \psi^*\Delta\psi \right] + \frac{i}{\hbar} \left[U\psi^*\psi - \psi^*U\psi \right]$$

$$= -\frac{i\hbar}{2m} \vec{\nabla} \left[(\vec{\nabla}\psi^*)\psi - \psi^*\vec{\nabla}\psi \right]$$
(3.17)

If we now define the probability current density distribution as

$$\vec{j} = -\frac{i\hbar}{2m} \left[\psi^* \vec{\nabla} \psi - (\vec{\nabla} \psi^*) \psi \right]$$
(3.18)

⁸⁴The term Quantum Mechanics was first coined by Max Born in his 1924 paper, entitled "Uber quantenmechanik" (On quantum mechanics), published in Zeitschrft fur Physik, **26**, 379- 395 (1924).

we obtain

$$\frac{\partial}{\partial t}\rho + \vec{\nabla}\vec{j} = 0 \tag{3.19}$$

which is known as the **continuity equation**, similar to the equation one encounters in fluid mechanics or electrodynamics. If we integrate the above equation over a region of volume \mathcal{V} bounded by a surface $\mathcal{S} = \partial \mathcal{V}$, we get

$$\frac{d}{dt} \int \rho(\vec{x}, t) \ dV = -\oint_{\mathcal{S}} \vec{j}(\vec{x}, t) \ \vec{n} \ dS =$$
(3.20)

Here \vec{n} is a unit vector normal to the surface at the point \vec{x} . In the above equation, changed the order of the integration and the differentiation on the left hand side and we used the Gauss- Ostrogradsky theorem to obtain the last term on right hand side. So, Eq (3.20) means that the change in probability within \mathcal{V} is equal to the outward probability flux through the the surface \mathcal{S} . If we tak the volume \mathcal{V} to be the whole space, then the $\vec{j}(|\vec{x}| \to \infty) = 0$, and thus we get

$$\frac{d}{dt} \int_{All-space} |\psi(\vec{x},t)|^2 d^3 \vec{x} = 0 \Rightarrow \int_{All-space} |\psi(\vec{x},t)|^2 d^3 \vec{x} = Constant \qquad (3.21)$$

In other words, the total probability is conserved. So, if the wave function was initially normalized to one, it remains equal to unity at all time.

With this probabilistic description, we can no longer speak with certainty about the position of a particle, but, instead, we can compute the probability of finding it at some particular position. However, given a probability density distribution $|\psi|^2$, we can calculate the average position of the particle, say of the coordinate $\langle x_i \rangle$, given by

$$\langle x_i \rangle = \frac{\int x_i \ \rho(\vec{x}, t) \ d^3 \vec{x}}{\int \rho(\vec{x}, t) \ d^3 \vec{x}} = \int x_i \ |\psi_{norm}(\vec{x}, t)|^2 \ d^3 \vec{x}$$
 (3.22)

where

$$\psi_{nom}(\vec{x},t) = \frac{\psi(\vec{x},t)}{\int |\psi(\vec{x},t)|^2 \ d^3\vec{x}}$$
(3.23)

is normalized wave function. In general, the expectation value of any function of the position, $f(\vec{x})$ is

$$\langle f(\vec{x}) \rangle = \int f(\vec{x}) |\psi_{norm}(\vec{x}, t)|^2 d^3 \vec{x}$$
 (3.24)

We can use (3.24) to calculate the standard deviation in the coordinate x_i (also called the variance), denoted by Δx_i , defined as

$$(\Delta x_i)^2 = \langle (x_i - \langle x_i \rangle)^2 \rangle = \langle x_i^2 \rangle - \langle x_i \rangle^2$$
(3.25)

where we used the fact that $\langle x_i \langle x_i \rangle > = \langle x_i \rangle^2$. Hence, we can write

$$(\Delta x_i)^2 = \int x_i^2 \ \rho(\vec{x}, t) \ d^3 \vec{x} - \left[\int x_i \ \rho(\vec{x}, t) \ d^3 \vec{x} \right]^2 \tag{3.26}$$

where $\rho(\vec{x}, t)$ is the normalized probability density.

Max born was awarded the Nobel Prize in Physics in 1954 for "his fundamental research in quantum mechanics, especially for his statistical interpretation of the wave function".

4 Examples of One Dimensional Potentials

4.1 Free Particle in 1-dimension and Wave Packet

The simplest system is free particle in one dimension, i.e U(x) = 0, for which equation (3.15) becomes

$$\frac{d^2\phi(x)}{dx^2} + k^2\phi(x) = 0$$
(4.1)

with

$$|k| = \sqrt{\frac{2mE}{\hbar^2}} \tag{4.2}$$

The general solution of (4.1) can be written as a combination of plane waves

$$\phi(x) = A \ e^{i|k|x} + B \ e^{-i|k|x} \tag{4.3}$$

Note that there are no boundary conditions on the above wave function, that is why it represents a "free" particle. The wave function at instant, can be found using (3.14), which gives

$$\psi(x,t) = A \ e^{i|k|\left(x - \frac{\hbar|k|}{2m}t\right)} + B \ e^{-i|k|\left(x + \frac{\hbar k}{2m}t\right)}$$
(4.4)

note the that A-term corresponds to a wave propagating in the positive direction of the x-axis, and the B-term is a propagating wave in the opposite direction. So, we can simply write

$$\psi_k(x,t) = A_k e^{ik\left(x - \frac{\hbar k}{2m}t\right)} \tag{4.5}$$

where k can be |k| or -|k|, depending if the wave is moving to the right or to the left, respectively, and A_k is a constant.

However, there are two problems when representing a free particle by a plane wave:

1. Plane waves are not square integrable (i.e. not renormalizable):

$$\int_{-\infty}^{\infty} |\psi_k(x,t)|^2 \, dx = |A_k|^2 \tag{4.6}$$

2. The speed of a plane wave is

$$v_{wave} = \frac{\hbar k}{2m} = \sqrt{\frac{E}{2m}} \tag{4.7}$$

Whereas, a free particle has a speed

$$v_{particle} = \sqrt{\frac{2E}{m}} = v_{wave} \tag{4.8}$$

So the speed of the particle is different from the speed of the wave function that it represents.

This means that there can be no free particle with definite momentum. The plane wave representation of free particle has a definite momenta but it is extended over the whole space. We can construct a wave function that has a finite spread in space, however, with a finite spread in momentum. This can be achieved by taking a linear combination of an infinite number of plane waves, called the wave packet. In general, a wave packet has the following form

$$\psi(x,t) = \frac{1}{\sqrt{2\pi}} \int \tilde{\phi}(k) \ e^{i(kx - \omega_k t)} \ dk \tag{4.9}$$

where $\tilde{\phi}(k)$ is some smooth function peaked at the wave number $k = k_0$, and ω_k is the frequency of oscillation which is a function of k. At the instant t = 0, we have

$$\psi(x,0) = \frac{1}{\sqrt{2\pi}} \int dk \; \tilde{\phi}(k) \; e^{ikx} \tag{4.10}$$

which means that $\tilde{\phi}(k)$ is just the Fourier transform of the wave packet at t = 0. The normalization of $\tilde{\phi}(k)$ follows from the normalization of $\psi(x, 0)$:

$$\int_{-\infty}^{\infty} |\psi(x,0)|^2 dx = \frac{1}{2\pi} \int \tilde{\phi}(k_1) \tilde{\phi}(k_2) e^{i(k_1-k_2)x} dk_1 dk_2 dx = 1$$
(4.11)

Using the identity

$$\int_{-\infty}^{\infty} e^{i(k_1 - k_2)x} dx = 2\pi \ \delta(k_1 - k_2) \tag{4.12}$$

we obtain

$$\int |\tilde{\phi}(k)|^2 \, dk = 1 \tag{4.13}$$

Hence, $|\tilde{\phi}(k)|^2 dk$ can be interpreted as the probability for a particle to have wave number k in an interval of width dk around some value of k.

By Taylor expanding ω_k around k_0 and neglecting the quadratic and higher order terms, we have

$$\omega_k = \omega_0 + \omega'_0 (k - k_0) \tag{4.14}$$

with $\omega_0 = \omega(k_0), \, \omega'_0) = \frac{d\omega}{dk}(k_0)$. This gives

$$\psi(x,t) \simeq e^{i(k_0x-\omega_0t)} \int \frac{\tilde{\phi}(k)}{\sqrt{2\pi}} e^{i(k-k_0)\left(x-\omega_0't\right)} dk \qquad (4.15)$$
$$\equiv e^{i(k_0x-\omega_0t)} \mathcal{F}\left(x-\omega'(k_0)t\right)$$

where \mathcal{F} is the envelope function for the plane wave $e^{i(k_0x-\omega_0t)}$. So the wave packet is traveling with a speed ω'_o . This can also be shown by re-writing the above expression of ψ as

$$\psi(x,t) \simeq e^{-[it(\omega_0 - k_0\omega'(k_0))]} \int \frac{\tilde{\phi}(k)}{\sqrt{2\pi}} e^{ik(x - \omega'(k_0)t)} dk$$
(4.16)

from which it follows that

$$|\psi(x,t)|^2 \simeq |\psi((x-\omega'(k_0)t),0)|^2$$
(4.17)

This means that if the wave packet was initially concentrated at some position x = 0, at time became concentrated around $x = \omega'(k_0)t$. Thus, it moved with a speed ⁸⁵

$$\upsilon_g = \frac{\partial \omega_k}{\partial k}(k_0) \tag{4.19}$$

which is called the **group velocity** of the wave packet. On, the other hand, the individual plane waves that constitute the wave packet have phase velocity

$$v_p = \frac{\omega_k}{k} \tag{4.20}$$

For a very narrow wave packet that is peaked around k_0 , the phase speed can be approximated by $v_p \simeq \omega_0/k_0$, which is not necessarily the same as the the group

$$\frac{\partial}{\partial k} \left[k \left(x - \omega'(k_0) t \right) \right] = 0 \Rightarrow x = \omega'(k_0) t \tag{4.18}$$

which implies that the envelope moves with a speed $v_g = \omega'(k_0)$.

⁸⁵Yet another way to determine the group velocity, is by noting that for arbitrarily chosen values of k, the exponential in (4.9) oscillates rapidly, resulting in the vanishing of ψ except where the phase is constant, i.e. when

velocity. But , If the system is dispersionless, i.e. ω_k is linear in k, then all the wave with the same speed and it same as the group velocity. For the case of a free particle, $\omega_k = \hbar k^2/2m$ which yields

$$v_p \simeq \frac{\hbar k_0}{2m} = v_{wave}; \qquad v_g = \frac{\hbar k_0}{m} = \frac{p_0}{m} = v_{particle}$$
 (4.21)

Thus, the group velocity represents the classical velocity of the particle.

Now let us consider a wave packet where $\tilde{\phi}(k)$ has a Gaussian profile centered at $k = k_0$

$$\tilde{\phi}(k) = A_k \ e^{-(k-k_0)^2/4\Delta_k^2}$$
(4.22)

which is a bell-shaped curve. Here A_k is a normalization constant and Δ_k is the width of the Gaussian wave packet in k-space⁸⁶. Using the normalization condition (4.13) we get

$$A_k = \frac{1}{\left(2\pi\Delta_k^2\right)^{1/4}} \tag{4.23}$$

With this profile, the wave function reads ⁸⁷

$$\psi(x,t) = \frac{A_k}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{\left[-\frac{(k-k_0)^2}{4\Delta_k^2} - i\frac{\hbar k^2 t}{2m} + ikx\right]} dk$$

$$= \frac{A_k}{\sqrt{2\pi}} e^{i\left(k_0 x - \frac{\hbar k_0^2 t}{2m}\right)} \int_{-\infty}^{\infty} e^{\left[-\left(\Delta_x^2 + i\frac{\hbar t}{2m}\right)(k-k_0)^2 + i(k-k_0)(x-v_0t)\right]} dk$$
(4.24)

where we have replaced $1/2\Delta_k$ by Δ_x which has dimension of length, and $v_0 = \hbar k_0/m$ is the classical velocity of the center of the wave packet associated with the free particle.

⁸⁶ If we can compute the average square of the deviation of k from k_0 , we get

$$<(\Delta k)^2>=\frac{\int_{-\infty}^{\infty}(k-k_0)^2e^{-(k-k_0)^2/4\Delta_k^2}\,dk}{e^{-(k-k_0)^2/4\Delta_k^2}\,dk}=\Delta_k^2$$

So, Δ_k is the width of the wave packet.

⁸⁷We made use of the Gaussian integral

$$\int_{-\infty}^{\infty} e^{-\alpha x^2} \, dx = \sqrt{\frac{\pi}{\alpha}}$$

In general, for n integer, we have

$$\int_{-\infty}^{\infty} x^n \ e^{-\alpha x^2} \ dx = \frac{1.3.5....(2n+1)\sqrt{\pi}}{2^{n/2}\alpha^{(n+1)/2}}$$

The above integral can be brought into the form of a Gaussian integral, and we obtain⁸⁸

$$\psi(x,t) = e^{i\left(k_0 x - \frac{\hbar k_0^2 t}{2m}\right)} \frac{\sqrt{\Delta_x}}{(2\pi)^{1/4} \left(\Delta_x^2 + i\frac{\hbar t}{2m}\right)^{1/2}} e^{-\left[\frac{(x-v_0t)}{4(\Delta_x^2 + i\frac{\hbar t}{2m})}\right]}$$
(4.25)

Thus, the probability density is given by

$$|\psi(x,t)|^{2} = \frac{1}{\sqrt{2\pi}\Delta_{x}(t)} exp\left[-\frac{(x-v_{0}t)}{2\Delta_{x}^{2}(t)}\right]$$
(4.26)

where

$$\Delta_x(t) = \left(\Delta_x^2 + \frac{\hbar^2 t^2}{4m^2 \Delta_x^2}\right) \tag{4.27}$$

So, the maximum of the Gaussian wave packet moves with the group velocity $v_g = k_0/m = \frac{\partial \omega_k}{k_0}$. Note this expression seems to be not consistent with Eq (4.17). This is due to the fact that in deriving (4.17) we have considered ω_k to be linear in the wave number, which is not true for the case of free particle where $\omega_k = \hbar k^2/2m$, i.e. quadratic in the wave number. Another point worth noticing is that the Fourier transform of a Gaussian is also a Gaussian.

The quantity $\Delta_x(t)$ looks like the width of the Gaussian packet in the position space. This suggest that it might be related to the average or the variance of the particle at time t. Using the expression of $|\psi(x,t)|^2$ above and the definition of the average a function of a position given in (3.24), we have

$$< x^{2}(t) > = \frac{1}{\sqrt{2\pi}\Delta_{x}(t)} \int_{-\infty}^{\infty} x^{2}(t) \exp\left[-\frac{(x-v_{0}t)}{2\Delta_{x}^{2}(t)}\right]$$

$$= \frac{1}{\sqrt{2\pi}\Delta_{x}(t)} \left[v_{0}^{2}t^{2} \left(2\pi\Delta_{x}^{2}(t)\right)^{1/2} + \frac{\sqrt{\pi}}{2} \left(2\Delta_{x}^{2}(t)\right)^{3/2}\right]$$

$$= v_{0}^{2}t^{2} + \Delta_{x}^{2}(t)$$
(4.28)

and

$$\langle x(t) \rangle = \frac{1}{\sqrt{2\pi}\Delta_x(t)} \int_{-\infty}^{\infty} x(t) \exp\left[-\frac{(x-v_0t)}{2\Delta_x^2(t)}\right]$$

$$= \frac{1}{\sqrt{2\pi}\Delta_x(t)} v_0 t \left(2\pi\Delta_x^2(t)\right)^{1/2}$$

$$= v_0 t$$
(4.29)

 $^{88}\mathrm{We}$ have used the fact that

$$I(a,b) = \int_{-\infty}^{\infty} e^{-(ax^2 + ibx)} dx = e^{-b^2/4a} \int_{-\infty}^{\infty} e^{-a(x - i\frac{b}{2a})^2} dx = \sqrt{\frac{\pi}{a}} e^{-b^2/4a}$$

Thus, the variance in the position at time t is

$$\Delta x(t) \equiv \langle x^2(t) \rangle - (\langle x(t) \rangle)^2 = \Delta_x^2(t)$$
(4.30)

So, the quantity $\Delta_x(t)$ is the variance of the position of the particle at time t, which grows linearly with time for $t >> t_c = \frac{2m\Delta_x^2}{\hbar}$. This means that the probability density spreads out as the time goes on, and eventually it will not be localized.

4.2 Infinite Square Potential

4.3 Symmetric Finite Potential: Bound and Unbound States

4.4 Particle moving through a Step Potential

In this subsection, we will study a particle initially traveling in one dimension from in a region of constant potential $V_1 \ge 0$ to a region with different potential $V_2 > V_1$. This means that there is a potential barrier separating the two regions. We assume that the particle has a constant energy $E > V_1$. Let x = 0 be the transition point from one region to the other. So, the potential reads

$$V(x) = \begin{cases} V_1, & x < 0; \\ V_2, & x \ge 0. \end{cases}$$
(4.31)

So the TISE reads

$$-\frac{\hbar^2}{2m}\frac{d^2\phi(x)}{dx^2} + V_1\phi(x) = 0; \qquad x < 0$$

$$-\frac{\hbar^2}{2m}\frac{d^2\phi(x)}{dx^2} + V_2\phi(x) = 0; \qquad x > 0$$
(4.32)

There are two cases to consider depending if the energy of the particle is larger or smaller compared to V_2 .

1. Case I: $E > V_2$

In the region x < 0, the solution has the form

$$\phi_{<}(x) \equiv \phi(x < 0) = A_1 e^{ik_1 x} + B_1 e^{-ik_1 x}$$
(4.33)

where A and B are constants to be determined, which correspond represent the amplitudes of the incident and the reflected waves, respectively, and k_1 is the wave number, given by

$$k_1 = \frac{\sqrt{2m(E - V_1)}}{\hbar}$$
(4.34)

Similarly, in the region x > 0, the wave function reads

$$\phi_{>}(x) \equiv \phi(x>0) = C_1 e^{ik_2 x} + D_1 e^{-ik_2 x}$$
(4.35)

with the coefficients C_1 and D_1 to be determined, and

$$k_2 = \frac{\sqrt{2m(E - V_2)}}{\hbar} \tag{4.36}$$

Since in the region x > 0 the particle is always traveling in the x > 0 direction, then $D_1 = 0$. So the wave function in the x > 0-region is given by

$$\phi_{>} = C_1 e^{ik_2 x} \tag{4.37}$$

Comparing both expressions of k_1 and k_2 , we note that $k_2 < k_1$, and thus in moving from a region of lower potential (x < 0) to higher potential (x > 0), the particle wave length increases.

Now, we use the continuity of the wave function and its first derivative⁸⁹ at the transition point x = 0, i.e

$$\phi_{<}(x=0) = \phi_{>}(x=0)$$

$$\frac{d\phi_{<}}{dx}(x=0) = \frac{d\phi_{>}}{dx}(x=0)$$
(4.38)

which yield

$$A_1 + B_1 = C$$

$$A_1 - B_1 = \frac{k_2}{k_1} C_1$$
(4.39)

This is a system of two equation with three unkowns A_1, B_1 and C_1 , so we can solve for B_1 and C_1 in terms of the coefficient A_1 of the incident wave, and we obtain

$$B_{1} = \left(\frac{1 - k_{2}/k_{1}}{1 + k_{2}/k_{1}}\right) A_{1}$$

$$C_{1} = \left(\frac{2}{1 + k_{2}/k_{1}}\right) A_{1}$$
(4.40)

The specific values of A_1 and B_1 can be determined using the normalization condition of the wave function in each region. However, for the purpose of this discussion we will not need their explicit expressions.

⁸⁹Even though V(x) is a step potential, the derivative of $\phi(x)$ is continuous. This can be seen as follows:

$$\left(\frac{d\phi}{dx}\right)_{x=\epsilon} - \left(\frac{d\phi}{dx}\right)_{x=-\epsilon} = \int_{-\epsilon}^{\epsilon} dx \, \frac{d}{dx} \left(\frac{d\phi}{dx}\right)$$
$$= -\frac{2m}{\hbar^2} \int_{-\epsilon}^{\epsilon} dx \, \left[E - V(x)\right] \phi(x)$$

where ϵ is a positive number. Taking the limit $\epsilon \to 0$, the right-hand side of the above equation vanishes.

Given the wave functions in the two regions we can calculate the current probability in each region. In the region x < 0, we have

$$j_{<} = i \frac{\hbar}{2m} \left(\frac{d\phi_{<}^{*}}{dx} \phi_{<} - \phi_{<}^{*} \frac{d\phi_{<}}{dx} \right)$$

$$= \frac{\hbar k_{1}}{m} \left(|A_{1}|^{2} - |B_{1}|^{2} \right)$$

$$= j_{inc} - j_{ref}$$

$$(4.41)$$

where j_{inc} and j_{ref} are the incident and reflected currents, respectively, given by

$$j_{inc} = \frac{\hbar k_1}{m} |A_1|^2$$

$$j_{ref} = \frac{\hbar k_1}{m} |B_1|^2$$
(4.42)

In the region x > 0, we have the transmitted current density

$$j_{trans} = i \frac{\hbar}{2m} \left(\frac{d\phi_{>}^{*}}{dx} \phi_{>} - \phi_{>}^{*} \frac{d\phi_{>}}{dx} \right)$$

$$= \frac{\hbar k_{2}}{m} |C_{1}|^{2}$$

$$(4.43)$$

We define the reflection and transmission coefficients \mathcal{R} and \mathcal{T} , respectively, as

$$\mathcal{R} := |\frac{j_{ref}}{j_{inc}}|; \qquad \mathcal{T} := |\frac{j_{trans}}{j_{inc}}|$$
(4.44)

Substituting the expression (4.42) and (4.45) in the above equations, we obtain

$$\mathcal{R} := \left(\frac{1 - k_2/k_1}{1 + k_2/k_1}\right)^2; \qquad \mathcal{T} := \frac{4k_2/k_1}{\left(1 + k_2/k_1\right)^2} \tag{4.45}$$

Note that $\mathcal{R} + \mathcal{T} = 1$, which means that the current density probability is conserved, or equivalently, the number of particles is conserved. It is interesting to note that, despite that the particle has energy E is larger than the potential barrier V_2 , the reflection coefficient is not zero. Classically we expect that there will be no reflection at all. Furthermore, the expression of the reflection coefficient has similar form as the one find in the light waves reflected normally at the boundary between two media, say 1 and 2, with refractive indices $n_1 = c/v_1$ and $n_2 = c/v_2$, respectively, where v_i is the speed of the light in the i^{th} medium⁹⁰. This shows the quantum mechanical manifestation of a particle as a wave.

$$\mathcal{R}^{(light)} = \frac{(n_1 - n_2)^2}{(n_1 + n_2)^2} = \frac{(1 - \nu_2/\nu_1)^2}{(1 + \nu_2/\nu_1)^2}$$

 $^{^{90}}$ The reflection coefficient for light incident at right angle at the boundary between two media is

2. Case II: $V_1 < E < V_2$

Classically, we know what would happen to a stream particles coming in from the x < 0 region with $E < V_2$: they will bounce against the step and we get a complete reflection. Let us see if this is also the case in the quantum theory.

In the region x < 0, the expression of the wave function takes the same form as in (4.33):

$$\phi_{<}(x) = A_2 e^{ik_1 x} + B_2 e^{-ik_1 x} \tag{4.46}$$

with A_2 , B_2 are constants to be determined by the use of boundary conditions and normalization of the wave functions. The wave number k_1 is given by Eq (4.34). As we noted in case I, the A and B terms represent the incident and the reflected waves respectively.

In the x > 0 region, we have $(E - V_2)$ is negative, the wave function is given in terms of exponentials

$$\phi_{>}(x) = C_2 e^{q_2 x} + D_2 e^{-q_2 x} \tag{4.47}$$

with the wave number q_2 is given by

$$q_2 = \frac{\sqrt{2m(V_2 - E)}}{\hbar} \tag{4.48}$$

The coefficient C_2 must be zero otherwise the wave function is not square integrable in this region. Now, requiring the continuity of the wave function and its first derivative at x = 0 we get

$$B_2 = \frac{1 - iq_2/k_1}{1 + iq_2/k_1} A_2; \qquad D_2 = \frac{2}{1 + iq_2/k_1} A_2 \tag{4.49}$$

So, the reflection coefficient reads

$$\mathcal{R} = \frac{|B_2|^2}{|A_2|^2} = \frac{|1 - iq_2/k_1|^2}{|1 + iq_2/k_1|^2} = 1$$
(4.50)

So all the particles incident on the potential barrier get reflected, which is consistent with what one expect from classically. However, the ratio B/A is complex, which means that there is a phase shift on reflection from the barrier, which is not what one would expect from classical theory. Furthermore, the probability for a particle to penetrate into the x > 0-region is not zero since $|\psi(x > 0)|^2 \neq 0$. But in this region, the particle does not propagate since the current $j_>$ vanishes, and instead there is an exponentially decaying wave. We can define the depth δx into this classically forbidden region, to be the distance from x = 0 to the point in the x > 0 region at which the probability density drops to 1/e, i.e.

$$\frac{|\phi_{>}(\delta x)|^{2}}{|\phi_{>}(0)|^{2}} = e^{-2q_{2}\delta x} = \frac{1}{e}$$
(4.51)

which yields

$$\delta x = \frac{1}{2q_2} = \frac{\hbar}{2\sqrt{2m(V_2 - E)}}$$
(4.52)

Hence, the penetration decreases for increasing the barrier height $(V_2 - E)$.

4.5 Harmonic Oscillator

Consider a particle subject to a potential V(x) and let x_0 be the point at which the particle can be in stable equilibrium. This means that one can Taylor-expand V(x) around x_0 and write

$$V(x) = V(x_0) + \frac{1}{2} (x - x_0)^2 \frac{d^2 V}{dx^2}|_{(x=x_0)} + \dots$$
(4.53)

For small deviations from the equilibrium point we can approximate the potential by keeping only the the constant and the quadratic terms. Also, one can always choose coordinates that such that $V(x_0) = 0$. Defining $m\omega^2 = \frac{d^2V}{dx^2}|_{(x=x_0)}$, the potential reads

$$V(x) = \frac{1}{2}m\omega^2 x^2$$
 (4.54)

This system represents simple harmonic oscillator (SHO) where the particle the motion of particle corresponds to harmonic oscillations around the equilibrium point. Classically, its energy is a continuous function which can take values from E = 0 to $E = \frac{1}{2}m\omega^2 x_0^2$.

In Quantum mechanics, this system is described by the Hamiltonian operator

$$\hat{\mathcal{H}} = -\hbar^2 \frac{d^2}{dx^2} + \frac{1}{2}m\omega^2 x^2$$
(4.55)

So if ψ_E represents the wave function of the SHO with energy E, the TISE $\hat{\mathcal{H}}\psi_E = E\psi_E$ can be written as

$$\frac{d^2\psi_E}{dx^2} + \frac{2m}{\hbar^2}m\left(E - \frac{1}{2}\omega^2 x^2\right)\psi_E = 0$$
(4.56)

If we make the following change of variables

$$y = \sqrt{\frac{m\omega}{\hbar}}x$$
; $\xi = \frac{E}{\hbar\omega}$ (4.57)

the wave equation will take the form

$$\frac{d^2\psi_E}{dy^2} + (2\xi - y^2)\,\psi_E = 0 \tag{4.58}$$

To get an idea about the form of the solution to the above equation, we analyze the asymptotic behavior of the solution by considering the limiting case $\xi \to \infty$. In this limit (4.58) becomes

$$\frac{d^2\psi_E}{dy^2} - y^2\psi_E = 0 , \quad y \to \infty$$
(4.59)

which has the solution of the form

$$\psi_E \propto y^n \ e^{-y^2/2}, \ y \to \infty$$
 (4.60)

note that we omitted a second solution with a positive sign in the argument of the exponential since it is not square integrable. With the above asymptotic behavior of ψ_E , we look for the general solution of the form

$$\psi_E(\xi) = H(\xi) \ e^{-y^2/2} \tag{4.61}$$

where $H(\xi)$ is a function of ξ . Inserting Eq (4.61) into Eq (4.58) yields

$$\frac{d^2 H(y)}{dy^2} - 2y \frac{dH(y)}{dy} + (2\xi - 1) H(y) = 0$$
(4.62)

Now, we make the following power law Ansatz for the function $H(\xi)$:

$$H(y) = \sum_{n=0}^{\infty} C_n y^n \tag{4.63}$$

By substituting into (4.69) yields

$$0 = \sum_{n=0}^{\infty} C_n \left[n(n-1)y^{n-2} + (-2n+2\xi-1)y^n \right]$$

$$= \sum_{n=0}^{\infty} \left[C_{n+2}(n+2)(n+1) + C_n(2\xi-1-2n) \right] y^n$$
(4.64)

Since the above equality has to hold for any y, the coefficients in the bracket must vanish at all order in n, which gives the recursion relation

$$C_{n+2} = -\frac{(2\xi - 1 - 2n)}{(n+2)(n+1)}C_n \tag{4.65}$$

So, given C_0 and C_1 we can determine all the C_n in (4.63), with the coefficients of even powers of y being dependent upon a_0 , and the coefficients of odd powers depending upon a_1 . So, in general the solution can be written as

$$H_n(y) = C_0 H_n^{(even)}(y) + C_1 H_n^{(odd)}(y)$$
(4.66)

where $H_n^{(even)}(y)$ and $H_n^{(odd)}(y)$ are series in only even and odd powers, respectively. In order to have a Physically acceptable solutions the series (4.63) must terminate at a certain power n. So requiring that C_{n+2} vanishes for arbitrary C_n implies that

$$\xi = n + \frac{1}{2} \tag{4.67}$$

or, equivalently,

$$E_n = \hbar\omega(n + \frac{1}{2}), \quad n = 0, 1, 2, ..$$
 (4.68)

Thus, unlike in the classical theory, the energy of a simple harmonic oscillator is quantized.

Now, setting $\xi = n + \frac{1}{2}$ into equation (4.69), yields

$$\frac{d^2 H_n(y)}{dy^2} - 2y \frac{dH_n(y)}{dy} + 2nH_n(y) = 0$$
(4.69)

which is the differential equation for the so-called **Hermite polynomials**. They can be easily calculated using the **Rodrigues formula**:

$$H_n(y) = (-1)^n e^{y^2} \left(\frac{d^n}{dy^n} e^{-y^2}\right)$$
(4.70)

Note that $H_n(y)$ involve even powers only if n is even integer, and odd powers if n is odd integer. The first few examples of Hermite polynomials are given in the table below.

$$H_0(y) = 1$$

$$H_1(y) = 2y$$

$$H_2(y) = 4y^2 - 1$$

$$H_3(y) = 8y^3 - 12y$$

$$H_4(y) = 16y^4 - 48y^2 + 12$$

$$H_5(y) = 32y^5 - 160y^3 + 120y.$$
(4.71)

so, the wave function of the harmonic oscillator of energy level n is

$$\psi_n(x) = \frac{1}{\sqrt{n! \, 2^n x_0}} \frac{1}{\pi^{1/4}} H_n\left(\frac{x}{x_0}\right) e^{-\frac{x^2}{2x_0^2}} \tag{4.72}$$

where $x_0 = \sqrt{\hbar/m\omega}$ is a constant that has the dimension of length.

4.6 Quantum Tunneling and the WKB approximation

5 Matrix Mechanics * [1925]

5.1 Heisenberg's Re-interpretation of the Position and New Multiplication Rules

On July 29, 1925, Heisenberg wrote a ground breaking paper, which its translation to English, entitled "Quantum- theoretical interpretation of kinematic and mechanical relations" ⁹¹, in which he gives a new formulation to quantum mechanics. He believed that the position and the period of electron in an atom are not measurable, and instead one should seek quantum theoretic relations between observable quantities such the transition probabilities between stationary states and their associated frequencies.

He considered an oscillator as a simple 1-dim model of an electron bound to an atom, for which a stationary state n has a frequency $\omega(n)$, and coordinate x(n,t). Since the electron is undergoing periodic motion, we can represent x(n,t) by Fourier series:

$$x(n,t) = \sum_{\alpha = -\infty}^{\alpha = +\infty} X_{\alpha}(n) \ e^{i\alpha\omega(n)t}$$
(5.1)

where α is an integer, and the factor $X_{\alpha}(n)$ represents the amplitude of the α^{th} harmonic. The fact that coordinate x(n,t) is real, requires that $X_{-\alpha}(n) = X_{\alpha}^*(n)$. The power radiated by this oscillator is given by the Larmor formula

_

$$-\frac{dE}{dt} = \left(\frac{e^2}{6\pi\epsilon_0 c^3}\right) \ddot{x}^2 \tag{5.2}$$

Heisenberg realized that according to the correspondence principle, the α^{th} harmonics of the classical motion in the state n corresponds to the quantum mechanical transition from the state n to $(n - \alpha)$, which lead him to re-write (5.1) in terms of the quantum quantities $\omega(n, n - \alpha)$ and $X(n, n - \alpha)$, i.e.

$$\alpha\omega(n) \to \omega(n, n-\alpha); \quad X_{\alpha}(n) \to X(\alpha, n-\alpha)$$
 (5.3)

So the power radiated in a transition corresponding to α^{th} harmonic $\omega(n, \alpha)$ now reads

$$-\frac{dE}{dt} = \left(\frac{e^2}{12\pi\epsilon_0 c^3}\right) \left[\omega(n, n-\alpha)\right]^4 |X(n, n-\alpha)|^2$$
(5.4)

Furthermore, he replaced the left hand side of the above equation by the spontaneous transition probability A_m^n times the emitted energy $\hbar\omega(n,m)$. So the transition rate from a state n to state m is

$$A_m^n = \left(\frac{e^2}{12\pi\epsilon_0 c^3\hbar}\right) \left[\omega(n,m)\right]^3 |X_{nm}|^2$$
(5.5)

⁹¹W. Heisenberg, Zeitschrift fur Physik, **33**, 879- 893 (1925).

where we have defined 92

$$x(t) = \sum_{m} [x]_{nm}; \quad [x]_{nm} = X(n,m) \ e^{i\omega(n,m)t}$$
(5.6)

with complex number x_{mn} corresponding to the component of the transition $n \to m$. Note that the above formula also holds when $E_n < E_m$, in which case the transition $m \to n$ is represented by $x_{nm} = x_{mn}^*$.

Knowing how to represent x(t) in terms of quantum theoretical quantities $\omega(n, m)$ and X(n, m), what about the quantum equivalent of $x^2(t)$?. Classically, using (5.1), we have

$$x^{2}(t) = \left[\sum_{\alpha} X_{\alpha}(n) \ e^{\alpha \omega(n)t}\right] \times \left[\sum_{\beta} X_{\beta}(n) \ e^{\beta \omega(n)t}\right]$$
(5.7)

$$= \left[\sum_{\alpha} X_{\alpha-\beta}(n) \ e^{(\beta-\alpha)\omega(n)t}\right] \times \left[\sum_{\beta} X_{\beta}(n) \ e^{\beta\omega(n)t}\right]$$
(5.8)

where in the last sum we replaced α by $(\alpha - \beta)$ since the sum extends from $-\infty$ to ∞ , and so all the terms in the summation will be accounted for. So, $x^2(t)$ we can be written in the form

$$x^{2}(t) = \sum_{\alpha} X_{\alpha}^{(2)}(n) \ e^{\alpha \omega(n)t}$$
(5.9)

where

$$X_{\alpha}^{(2)}(n) = \sum_{\beta} X_{\beta}(n) \ X_{(\alpha-\beta)}(n)$$
 (5.10)

and the frequency $\alpha\omega(n)$ is given by the simple combination

$$\alpha\omega(n) = \beta\omega(n) + (\alpha - \beta)\,\omega(n) \tag{5.11}$$

For the quantum analogue of $x^2(t)$ in (5.12) one is tempted to write

$$x^{2}(t) = \left[\sum_{\beta} X(n, n-\beta) \ e^{\omega(n, n-\beta)t}\right] \times \left[\sum_{\alpha} X(n, n-\alpha-\beta) \ e^{\omega(n, n-\alpha-\beta)t}\right]$$
(5.12)

However according the so called **Ritz's combination principle**, the combination of frequencies is given by :

$$\omega(n, n - \beta) + \omega(n - \beta, n - \beta - \alpha) = \omega(n, n - \beta - \alpha)$$
(5.13)

$$\omega(n-\alpha, n-\beta-\alpha) + \omega(n, n-\alpha) = \omega(n, n-\beta-\alpha)$$
(5.14)

 $^{^{92}}$ It is worth mentioning that this is not the notation used by Heisenberg in his paper. The reason we changed to the above notation so that it becomes transparent, later, to the reader how matrices arises in Heisenberg formulation of the quantum theory.

which satisfy the Bohr's frequency relation $\omega(n, n-\alpha) = \frac{1}{h} (E_n - E_{(n-\alpha)})$. This shows that the combination $[\omega(n, n-\beta) + \omega(n, n-\alpha-\beta)]$ does not result in $\omega(n, n-\alpha)$. Therefore, Heisenberg suggested that the quantum-theoretic $x^2(t)$ reads

$$x^{2}(t) = \left[\sum_{\beta} X(n, n-\beta) e^{\omega(n, n-\beta)t}\right] \times \left[\sum_{\alpha} X(n-\alpha, n-\beta) e^{\omega(n-\alpha, n-\beta)t}\right] (5.15)$$
$$-\sum_{\alpha} \sum_{\alpha} X(n, n-\beta) X(n-\beta, n-\alpha) e^{\omega(n, n-\alpha)t}$$
(5.16)

$$= \sum_{\alpha} \sum_{\beta} X(n, n-\beta) X(n-\beta, n-\alpha) e^{\omega(n, n-\alpha)t}$$
(5.16)

or, equivalently,

$$x^{2}(n,t) = \sum_{m} [x^{2}]_{nm} ; \qquad [x^{2}]_{nm} = \sum_{j=-\infty}^{\infty} [x]_{nj} [x]_{jm}$$
(5.17)

Similar form can be found for the quantum analogue of $x^n(t)$, and for any given function f(x). For instance, the quantum theoretic representation of $x^3(t)$ is given by

$$x^{3}(n,t) = \sum_{m} [x^{3}]_{nm} ; \qquad [x^{3}]_{nm} = \sum_{j,k} [x]_{nj} [x]_{jk} [x]_{km}$$
(5.18)

Furthermore, Heisenberg noticed that if one consider two quantities $x(t) = \sum_{m} [x_{nm}]$ and $y(t) = \sum_{m} [y_{nm}]$, then

$$x(t)y(t) = \sum_{m=-\infty}^{\infty} [x_{nm}] [y_{mn}] ; \quad y(t)x(t) = \sum_{m=-\infty}^{\infty} [y_{nm}] [x_{mn}]$$
(5.19)

So, unlike in classical theory where x(t)y(t) and y(t)x(t) are always equal, this is not in general the case in the quantum theory, and one expect that $x(t)y(t) \neq y(t)x(t)$ for generic x(t) and y(t). This was strange to Heisenberg so that he refers to it as "difficulty".

5.2 Heisenberg's Quantum Condition and The non linear Oscillator

In the second section of his paper, Heisenberg Consider the dynamics of the motion of a system which is classically described by Newton's second law of the form

$$\ddot{x} + f(x) = 0 \tag{5.20}$$

In the old quantum theory, to determine the quantum-theoretical solution of this system, one has to integrate the equation of motion (5.20) and then uses the Bohr-Sommerfeld quantization condition

$$\oint p \, dq = \oint m\dot{x}^2 \, dt = nh \tag{5.21}$$

- 66 -

In terms of the classical components of the Fourier series for x(n, t), the above condition takes the form

$$nh = 2\pi m \sum_{\alpha = -\infty}^{\infty} X_{\alpha}(n) X_{-\alpha}(n) \alpha^{2} \omega(n)$$

$$= 2\pi m \sum_{\alpha = -\infty}^{\infty} |X_{\alpha}(n)|^{2} \alpha^{2} \omega(n)$$
(5.22)

where we used the fact that $X_{-\alpha}(n) = X^*_{\alpha}(n)$. Heisenberg felt that the presence of the integer n in the above condition was arbitrary and argued that one should replace it by a condition about the transition between stationary states. In other words, he believed that what matters is the difference between $\oint pdq$ evaluated between for two neighboring states n and (n - 1), i.e. $[\oint pdq]_n - [\oint pdq]_{(n - 1)}$. For that he took the derivative of (5.22) with respect to n and obtained

$$h = 2\pi m \sum_{\alpha = -\infty}^{\infty} \alpha \frac{d}{dn} \left[|X_{\alpha}(n)|^{2} \alpha \omega(n) \right]$$

$$= 4\pi m \sum_{\alpha = 0}^{\infty} \alpha \frac{d}{dn} \left[|X_{\alpha}(n)|^{2} \alpha \omega(n) \right]$$
(5.23)

Using the generalized Born's correspondence rule in (2.27), he replaced the above differentiation by

$$h = 4\pi m \sum_{\alpha=0}^{\infty} \left[|X(n+\alpha,n)|^2 \omega(n+\alpha,n) - |X(n,n-\alpha)|^2 \omega(n,n-\alpha) \right]$$
(5.24)

This condition relates the amplitudes of different transitions between stationary states, whereas Bohr-Sommerfeld condition relates the frequencies. By requiring that no radiation should be emitted in the ground state, one can determine the amplitudes a in (5.24) completely, i.e. with no undetermined constant. So, in principle, using equations (5.20) and (5.24) one can solve for the amplitudes X(n,m) and the frequencies $\omega(n,m)$. We can also re-write (5.24)as

$$h = 4\pi m \sum_{m} [|[x]_{nm}|^2 \omega(n,m)$$
(5.25)

where now m in the summation runs over positive integers. The terms in the sum give positive contribution if the state m has energy higher than the state n, and negative if m has lower energy. In the case of Bohr model of the atom, the positive terms correspond to m < n, whereas the negative ones for m > n. a an application to his new formalism, studied a non linear oscillator with a potential $V(x) = \lambda/3x^3$. The equation of motion reads

$$\ddot{x} + \omega_0^2 x + \lambda x^2 = 0 \tag{5.26}$$

where t λ is a small real number. Classically, the solution of (5.26) can be expressed as an expansion in terms of powers of the parameter λ , i.e.

$$x(n,t) = \lambda a_0(n) + \sum_{\alpha=1}^{\infty} \lambda^{\alpha-1} a_\alpha(n) \cos\left(\alpha \ \omega(n)t\right)$$
(5.27)

The coefficients a_i and the frequency ω are also to be expanded in powers of λ :

$$a_{i}(n) = \sum_{k=0}^{\infty} \lambda^{k} a_{i}^{(k)}(n) ; \qquad i = 0, 1, 2, ..$$

$$\omega(n) = \sum_{k=0}^{\infty} \lambda^{k} \omega^{(k)}(n)$$
(5.28)

Substituting (5.27) into the equation of motion (5.20), yields

$$\begin{split} \lambda[\omega_0^2 a_0(n) + \frac{1}{2}a_1^2(n) + \lambda^2(a_0^2(n) + \frac{1}{2}a_2^2(n)) + ..] \\ + \left[(-\omega^2(n) + \omega_0^2(n))a_1 + \lambda^2(a_1(n)a_2(n) + 2a_0(n)a_1(n)) + ..\right]\cos\omega t \\ + \lambda[(-4\omega^2(n) + \omega_0^2)a_2(n) + \frac{1}{2}a_1^2(n) + \lambda^2(a_1(n)a_3(n) + 2a_0(n)a_2(n)) + ..]\cos 2\omega(n)t \\ + \lambda^2[(-9\omega^2(n) + \omega_0^2)a_3(n) + a_1(n)a_2(n) + \lambda^2(a_1(n)a_4(n) + 2a_0(n)a_3(n)) + ..]\cos 3\omega(n)t + ... = 0 \end{split}$$

where the dots correspond to powers of λ higher than two. Now, by equating to zero the constant term and the terms that multiply $\cos \omega t$, $\cos 2\omega t$, ..ect, one can find a recursion formulas for the coefficients $a_i(n)$ at each order of the expansion parameter λ .

In the quantum case, as we saw in the previous subsection, the position is represented by $x = \sum_{\alpha} X(n, n - \alpha)$, and so after substituting into (5.26), the transition amplitude $X(n, n - \alpha)$ satisfy the equation

$$\left[\omega_0^2 - \omega^2(n, n - \alpha)\right] X(n, n - \alpha) + \lambda \sum_{\beta} X(n, n - \beta) X(n - \beta, n - \alpha) = 0 \quad (5.29)$$

where the $X(n, n - \alpha) = X^*(n - \alpha, n)$ because x(n, t) is real. However, no general solution can be found to the above equation. For this reason Heisenberg suggested to represent x(n, t) with similar form as of (5.27) but replacing the coefficients $a_i(n), i = 0, 1, 2, ...$ by the amplitudes a(n, n - i), and $\alpha \omega(n)$ by $\omega(n, n - \alpha)$:

$$x(n,t) = \lambda a(n,n) + \sum_{\alpha=1}^{\infty} \lambda^{\alpha-1} a(n,n-\alpha) \cos\left(\omega(n,n-\alpha)t\right)$$
(5.30)

where again, as in (5.28), the coefficients $a(n, n - \alpha)$ and the frequencies $\omega(n, n - \alpha)$ can be expanded as

$$a(n, n - \alpha) = \sum_{k=0}^{\infty} \lambda^k \ a^{(k)}(n, n - \alpha) ; \qquad i = 0, 1, 2, \dots$$

$$\omega(n, n - \alpha) = \sum_{k=0}^{\infty} \lambda^k \ \omega^{(k)}(n, n - \alpha)$$
(5.31)

In order to be able to use the above expansion in terms of cosnines, one can choose the amplitudes $X(n, n - \alpha)$ to be real, so that one has

$$X(n, n - \alpha) = X(n - \alpha, n)$$
(5.32)

and

$$X(n,n) = \lambda a(n,n) ; \quad X(n,n-\alpha) = \frac{\lambda^{\alpha-1}}{2}a(n,n-\alpha) ; \quad \alpha > 0$$
 (5.33)

For general value of α , positive or negative, we have

$$X(n, n - \alpha) = \frac{\lambda^{|\alpha| - 1}}{2} a(n, n - \alpha) ; \quad \alpha \neq 0$$
(5.34)

In terms of $a(n, n - \alpha)$ the quantization condition (5.24) reads

$$h = \pi m \sum_{\alpha=1}^{\infty} \lambda^{2\alpha-2} \left[a(n+\alpha,n)^2 \omega(n+\alpha,n) - a(n,n-\alpha)^2 \omega(n,n-\alpha) \right]$$
(5.35)

Substituting the expression of $X(n, n - \alpha)$ in (5.33) into the equation of motion (5.29), and keeping terms up to order λ^2 , we get

• For $\alpha = 0$:

$$\lambda \{ \omega_0^2 a(n,n) + \frac{1}{4} \left[a^2(n+1,n) + a^2(n,n-1) \right]$$

+ $\lambda^2 [a^2(n,n) + \frac{1}{4} \left(a^2(n+2,n) + a^2(n,n-2) \right)] \} = 0$ (5.36)

• For $\alpha = 1$:

$$\begin{bmatrix} -\omega^2(n, n-1) + \omega_0^2 \end{bmatrix} a(n, n-1) + \lambda^2 \{ a(n, n)a(n, n-1)1a(n, n-1)a(n-1, n-1) \\+ \frac{1}{2} [a(n, n+1)a(n+1, n-1)] \} = 0$$
(5.37)
• For $\alpha = 2$:

$$\lambda\{[-\omega^{2}(n, n-2) + \omega_{0}^{2}]a(n, n-2) + \frac{1}{2}a(n, n-1)a(n-1, n-2) + \lambda^{2}[a(n, n)a(n, n-2) + a(n, n-2)a(n-2, n-2) + \frac{1}{2}a(n, n+1)a(n+1, n-2) + \frac{1}{2}a(n, n-3)a(n-3, n-2)]\} = 0$$
(5.38)

• For $\alpha = 3$:

$$\lambda^{2} \{ [-\omega^{2}(n, n-3) + \omega_{0}^{2}]a(n, n-3) + \frac{1}{2}[a(n, n-1)a(n-1, n-3) + a(n, n-2)a(n-2, n-3)] + \lambda^{2}[a(n, n)a(n, n-3) + a(n, n-3)a(n-3, n-3) + \frac{1}{2}a(n, n+1)a(n+1, n-3) + \frac{1}{2}a(n, n-4)a(n-4, n-3)] \} = 0$$
(5.39)

Let us first consider the lowest order solution to the amplitudes and frequencies by ignoring the terms proportional to λ^2 , and replacing the a's and $\omega's$ by their corresponding zeroth order terms. In this case Eq (5.37) gives (provided that $a^{(0)}(n, n-1) \neq 0$)

$$\omega^{(0)}(n,n-1) = \omega_0 ; \quad \forall n$$
(5.40)

By substituting the above equation into (5.41), the quantum condition reads

$$h = \pi m \omega_0 \{ [a^{(0)}(n+1,n)]^2 - [a^{(0)}(n,n-1)]^2 \}$$
(5.41)

which yields the solution

$$[a^{(0)}(n,n-1)]^2 = \frac{h}{\pi m \omega_0} \left(n + contant\right)$$
(5.42)

To determine the value of the constant in the above expression Heisenberg used the requirement that in the ground state there should be no transition to lower lower state, i.e.

$$a^{(0)}(0,-1) = 0 (5.43)$$

This implies that the constant in (5.42) is zero, and we get

$$a^{(0)}(n,n-1) = C\sqrt{n} \tag{5.44}$$

where

$$C = \sqrt{\frac{h}{\pi m \omega_0}} \tag{5.45}$$

With these results, Eq (5.36) gives

$$a^{(0)}(n,n) = -\frac{C^2}{4\omega_0^2}(2n+1)$$
(5.46)

A similar treatment of Eq (5.38) leads to

$$\left(-\left[\omega^{(0)}(n,n-2)\right]^2 + \omega_0^2\right)a^{(0)}(n,n-2) + \frac{1}{2}a^{(0)}(n,n-1)a^{(0)}(n-1,n-2) = 0(5.47)$$

Since the combination rule of frequencies in (5.13) must also hold for the lowest-order frequencies, we have

$$\omega^{(0)}(n, n-2) = \omega^{(0)}(n, n-1) + \omega^{(0)}(n-1, n-2) = 2\omega_0$$
(5.48)

where in the second equality we used (5.40). By repeatedly applying the frequencies combination rule, we obtain

$$\omega^{(0)}(n, n - \alpha) = \alpha \omega_0 \tag{5.49}$$

Now, by substituting (5.44) and (5.48) into Eq (5.47), we get

$$a^{(0)}(n, n-2) = -\frac{C^2}{6\omega_0^2}\sqrt{n(n-1)}$$
(5.50)

in general, the lowest-order transition amplitude $a^{(0)}(n, n - \alpha)$ will have the form

$$a^{(0)}(n, n - \alpha) = A_{\alpha} \frac{C^{\alpha}}{6 \,\omega_0^{2(\alpha - 1)}} \sqrt{\frac{n!}{(n - \alpha)}} ; \qquad \alpha = 1, 2, \dots$$
 (5.51)

where A_{α} is a numerical factor which depend on the α .

The total energy of this oscillator is the sum of the kinetic and potential energies⁹³:

$$W = \frac{1}{2}m\dot{x}^2 + \frac{1}{2}m\omega_0^2 x^2 + \frac{\lambda}{3}x^3$$
(5.52)

Now according to Heisenberg formalism, the quantities x^2, x^3 , and \dot{x}^2 in the expression of W can be commuted using the multiplication rule (5.19), which implies that the total energy will have the form

$$W = \sum_{\alpha} W(n, n - \alpha) \ e^{i\omega(n, n - \alpha)}$$
(5.53)

$$\frac{d}{dt}\left[\frac{1}{2}m\dot{x}^2 + \frac{1}{2}m\omega_0^2x^2 + \frac{\lambda}{3}x^3\right]$$

Hence, the energy W as defined in (5.52) is conserved.

⁹³This is a conservative system, i.e. the total energy is conserved. This can be seen as follows. By multiplying the equation of motion (5.26) by \dot{x} we get

However, since the total energy of this system is conserved, i.e. it means it is independent the time, the matrix elements with $\alpha \neq 0$ must vanish,

$$W(n, n - \alpha) = 0 ; \qquad \alpha \neq 0$$
(5.54)

If we ignore terms of order λ and higher, then this system corresponds to a harmonic oscillator with energy given by ⁹⁴

$$W \equiv W(n,n) = \frac{1}{2}m\dot{x}^{2} + \frac{1}{2}m\omega_{0}^{2}x^{2}$$

$$= \frac{1}{2}m\sum_{\beta}\omega^{2}(n,n-\beta)x(n,-\beta)x(n-\beta,n) + \frac{1}{2}m\omega_{0}^{2}\sum x(n,-\beta)x(n-\beta,n)$$
(5.55)

where we used the fact that $\omega(n, n-\beta) = -\omega(n-\beta, n)$. The only terms which survive in the above summation are for $\beta = \pm 1$ (see Eq (5.40)). So using (5.34) and (5.32), we obtain

$$W = \frac{1}{4} \{ [a^{(0)}(n, n-1)]^2 + [a^{(0)}(n+1, n)]^2 \}$$
(5.56)

which after substituting the expression of $a^{(0)}(n, n-1)$ in (5.44) (and deducing from it the expression of $a^{(0)}(n+1, n)$), we find

$$W = \frac{h}{2\pi}\omega_0 \left(n + \frac{1}{2}\right) \tag{5.57}$$

A similar analysis can be carried for higher order in λ . We will not give the derivation here, however, the interested reader can find the details in [4].

5.3 Born and Jordan's Matrix Mechanics

Born Realized that Heisenberg multiplication rule in (5.19) was nothing but the rule for multiplying matrices, which, in general, is non commutative. So, for example, the quantities $[x^2]_{nm}$ represent the elements of the square of the matrix [x] which has matrix elements $[x]_{nm}$.

Born considered a dynamical system with one degree of freedom and represented the classical coordinate q (which in the previous subsections was denoted by x) and the momentum p by a matrix q and p, respectively,

$$\mathbf{q} = \sum_{m} q(n,m) e^{i\omega_{nm}t} = \sum_{m} [q_{nm}]$$

$$\mathbf{p} = \sum_{m} p(n,m) e^{i\omega_{nm}t} = \sum_{m} [p_{nm}]$$
(5.58)

⁹⁴ With the use of $x = \sum_{\alpha} x(n, n - \alpha)$, a naive inspection of the contribution of the terms x^2 and \dot{x}^2 to the expression of W in (5.53) seems to suggest that in addition to the matrix element W(n, n) (which is consistent with energy conservation), the matrix elements W(n, n-2) and W(n, n+2) are also present. However, in [4], the authors showed explicitly (see the Appendix B in their paper) that $W(n, n \pm 2) = 0$.

where $[q_{nm}]^* = [q_{mn}]$, and $[p_{nm}]^* = [p_{mn}]$. Now, from the definition $p = m\dot{x}$, the momentum matrix element reads

$$[p_{mn}] = m[\dot{x}]_{mn} = im\omega(n,m)[x]_{nm}$$
(5.59)

So the matrix elements of [px] and [xp] read

$$[px]_{nm} = im \sum_{k} \omega(n,k) [x]_{nk} [x]_{km}$$

$$[xp]_{nm} = im \sum_{k} \omega(k,m) [x]_{nk} [x]_{km}$$
(5.60)

and so,

$$[xp - px]_{nm} = im \sum_{k} \left(\omega(k, m) - \omega(n, k)\right) [x]_{nk} [x]_{km}$$

$$= 2im \sum_{k} \omega(k, m) [x]_{nk} [x]_{km} theor$$
(5.61)

Now let us consider only the the diagonal elements of matrix [xp - px]. We have

$$[xp - px]_{nn} = -2im \sum_{k} \omega(n,k) |[x]_{kn}|^2$$
(5.62)

where the property $[x]_{nk} = [x]_{kn}^*$ has been used. Now by comparing the left hand side of the above equation with the quantum condition (5.35), yields

$$[xp - px]_{nn} = i\frac{h}{2\pi} = i\hbar$$
(5.63)

Born also convinced him self that the only reasonable value of the off-diagonal elements of [xp - px] should be zero. But as Born said it in his words "But this was only a guess, and all my attempts to prove it failed". Indeed, his guess was correct, and two months after Heisenberg's paper was sent for publication, he and his assistant Pascual Jordan wrote a paper in which they demonstrate that [xp - px] is diagonal⁹⁵.

They first postulate that the quantum-theoretic frequencies ω_{nm} associated with the transitions between states described by the quantum numbers n and m are required to satisfy Ritz rule, i.e

$$\omega(i,j) + \omega(j,k) + \omega(k,i) = 0 \tag{5.64}$$

which implies that $\omega(i, j) = -\omega(j, i)$. The reason that they imposed this postulate is because the Ritz rule explained the relations of the spectral lines in atomic spectroscopy, and played a crucial role in Heisenberg's discovery of the multiplication rule (5.19). This requirement suggests that there exists quantities W_n such that

$$\omega(n,m) = \frac{2\pi}{h} \left(W_n - W_m \right) \tag{5.65}$$

⁹⁵M. Born and P. Jordan, "Zur Quantenmechanik", Z. Phys, **34**, 858-888, (1925). An English translation can be found in the reference [3].

where, at this point, the quantity W_n is unrelated to the energy E_n of the state n. The above equation insures that the nm element of any function of **q** and **p**, $[f(q,p)]_{nm}$ oscillates with the same frequency $\omega(n,m)$ as $[q]_{nm}$ and $[p]_{nm}$. This means that f(q,p)takes the form

$$\mathbf{f} = f(\mathbf{q}, \mathbf{p}) = \sum_{m} f(n, m) e^{i\omega(n, m)t} \equiv \sum_{m} [f]_{nm}$$
(5.66)

where

$$f(n,m) = f\left(\sum_{k} q(n,k)q(k,m)\right)$$
(5.67)

Then, the derivative with respect to time of the matrix \mathbf{f} gives

$$\dot{\mathbf{f}} = \sum_{m} i\omega(n,m) f(n,m) e^{i\omega(n,m)t} = \sum_{m} [\dot{\mathbf{f}}]_{nm}$$
(5.68)

Now if $\dot{\mathbf{f}} = 0$, then $\omega(n,m)f(n,m) = 0$ for all values of (n,m). This is because $\omega(n,m) \neq 0$ when $n \neq m$, which is a consequence of Eq (5.64). Thus, we conclude that if $\dot{f}(\mathbf{q},\mathbf{p}) = 0$, then \mathbf{f} is a diagonal matrix, i.e. $[\mathbf{f}]_{nm} = \delta_{nm}[\mathbf{f}]_{nn}$.

For a general function $f(\mathbf{qp})$, then it can be expressed a linear aggregate of terms of the form

$$z = \prod_{j=1}^{k} p^{s_j} q^{r_j} \tag{5.69}$$

Then Born and Jordan showed that⁹⁶

$$\left(q\frac{\partial z}{\partial q} - \frac{\partial z}{\partial q}q\right) + \left(p\frac{\partial z}{\partial p} - \frac{\partial z}{\partial p}p\right) = 0$$
(5.70)

Since the matrix $f(\mathbf{qp})$ is linear in z, the above equation also holds for the matrix \mathbf{f} , and so

$$\left(q\frac{\partial \mathbf{f}}{\partial q} - \frac{\partial \mathbf{f}}{\partial q}q\right) + \left(p\frac{\partial \mathbf{f}}{\partial p} - \frac{\partial \mathbf{f}}{\partial p}p\right) = 0$$
(5.71)

⁹⁶For that they used the definition of the partial derivative of the product of matrices. For instance if $Y = \prod_{m=1}^{s} x_{l_m}$, then the partial derivative of the matrix Y with respect to, say, x_k is

$$\frac{\partial Y}{\partial x_k} = \sum_{r=1}^s \delta_{lk} \prod_{m=r+1}^s x_{l_m} \prod_{m=1}^{m-r-1} x_{l_m}$$

For example,

$$\frac{\partial \left(x_1^2 x_2 x_1 x_3\right)}{\partial x_1} = x_1 x_2 x_1 x_3 + x_2 x_1 x_3 x_1 + x_3 x_1^2 x_2$$

To describe the dynamics, they generalized Heisenberg's Newtonian approach using the Hamiltonian formalism where the energy of the system is given by the Hamilton function

$$\mathbf{H} = \frac{\mathbf{p}}{2m} + U(\mathbf{q}) \tag{5.72}$$

and they assumed, as Heisenberg did, that the equations of motion for \mathbf{q} and \mathbf{p} have the same form as in the classical theory, where the time evolution of \mathbf{q} and \mathbf{p} are given by teh Hamilton's equations ⁹⁷

$$\dot{\mathbf{q}} = \frac{\partial \mathbf{H}}{\partial \mathbf{p}} = \frac{1}{m} \mathbf{p}$$

$$\dot{\mathbf{p}} = -\frac{\partial \mathbf{H}}{\partial \mathbf{q}} = -\frac{\partial U}{\partial \mathbf{q}}$$
(5.73)

Now, since **H** is a function of **qp**, one can take $\mathbf{f} = \mathbf{H}$ in Eq (5.71) and uses the Hamilton's equations to obtain

$$\dot{\mathbf{p}}q - p\dot{\mathbf{q}} - \dot{\mathbf{q}}p - q\dot{\mathbf{p}} = 0 \tag{5.74}$$

or, equivalently

$$\frac{d}{dt}\left(\mathbf{pq} + \mathbf{qp}\right) = 0\tag{5.75}$$

which, as we discussed above, implies that $(\mathbf{pq} + \mathbf{qp})$ is a diagonal matrix. Therefore, with the Born's result of (5.63), we get

$$[\mathbf{q}, \mathbf{p}] = i \frac{h}{2\pi} \mathbf{I}$$
(5.76)

where **I** is the unit matrix, and $[\mathbf{A}, \mathbf{B}] := \mathbf{A}\mathbf{B} - \mathbf{B}\mathbf{A}$, called the commutator of the matrices **A** and **B**. The above expression is a consequence of the correspondence principle (2.35), and is one of the most fundamental relations in the quantum mechanics⁹⁸. Now, with the use of (5.76), by induction we obtain

$$[\mathbf{q}\mathbf{p}^n - \mathbf{p}^n\mathbf{q}] = ni\frac{h}{2\pi} \mathbf{p}^{n-1}; \quad [\mathbf{p}\mathbf{q}^n - \mathbf{q}^n\mathbf{p}] = -ni\frac{h}{2\pi} \mathbf{q}^{n-1}$$
(5.77)

Considering the Hamiltonian to have the form

$$\mathbf{H} = H_1(\mathbf{p}) + H_2(\mathbf{q}) \tag{5.78}$$

⁹⁷Heisenberg, Born and Jordan realized that the problem with classical mechanics i not the form of the equations of motion (dynamics), but rather the interpretation of the position and momentum (kinematics).

 $^{^{98}}$ In their paper, Born and Jordan refer to the relation (5.76) as "versharft Quantenbedingung ", which translates to English as "sharpened quantum condition".

where $H_1(\mathbf{p})$ and $H_2(\mathbf{q})$ can be represented by power series of \mathbf{p} and \mathbf{q} , respectively. Then, using the formulae in (5.77) one finds

$$\mathbf{Hq} - \mathbf{qH} = -i\frac{\hbar}{2\pi}\frac{\partial\mathbf{H}}{\partial\mathbf{p}},$$

$$\mathbf{Hp} - \mathbf{pH} = i\frac{\hbar}{2\pi}\frac{\partial\mathbf{H}}{\partial\mathbf{q}}$$
(5.79)

which after comparing with the Hamilton's equations of motion yields

$$\dot{\mathbf{q}} = \frac{2\pi i}{h} \left(\mathbf{H}\mathbf{q} - \mathbf{q}\mathbf{H} \right) \equiv \frac{2\pi i}{h} \left[\mathbf{H}, \mathbf{q} \right], \qquad (5.80)$$
$$\dot{\mathbf{p}} = \frac{2\pi i}{h} \left(\mathbf{H}\mathbf{p} - \mathbf{p}\mathbf{H} \right) \equiv \frac{2\pi i}{h} \left[\mathbf{H}, \mathbf{p} \right]$$

This implied that the time derivative of the product \mathbf{pq} is given by

$$\frac{d}{dt}(\mathbf{pq}) = \frac{2\pi i}{h} \left[\mathbf{p}(\mathbf{Hq} - \mathbf{qH}) + (\mathbf{Hp} - \mathbf{pH})\mathbf{q} \right]$$

$$= \frac{2\pi i}{h} \left(\mathbf{Hpq} - \mathbf{pqH} \right) \equiv \frac{2\pi i}{h} \left[\mathbf{H}, \mathbf{pq} \right]$$
(5.81)

So, in general, for a matrix $f(\mathbf{pq})$, we have

$$\dot{\mathbf{f}} = \frac{2\pi i}{h} \left[\mathbf{H}, \mathbf{f} \right] \tag{5.82}$$

If we set $\mathbf{f} = \mathbf{H}$ in the above relation, we get

$$\dot{\mathbf{H}} = 0 \tag{5.83}$$

which according to our previous discussion means that **H** is a diagonal matrix.

In 1926, Heisenberg joined Born and Jordan and they wrote an important paper, known as the "three men's paper", which generalized the above results to systems with more than one degrees of freedom. Namely, for a system with coordinates $\{\mathbf{q}_1, \mathbf{q}_2, ... \mathbf{q}_N\}$ and their corresponding canonically-conjugate momenta $\{\mathbf{p}_1, \mathbf{p}_2, ... \mathbf{p}_N\}$, the \mathbf{q}' and \mathbf{p} 's satisfy the commutation relations

$$[\mathbf{q}_i, \mathbf{p}_j] = i\hbar\delta_{ij} ; \quad i, j = 1, 2, ..N$$
 (5.84)

They also noted that for any quantum-theoretic quantity, \mathbf{f} , one can always write

$$[W\mathbf{f} - \mathbf{f}W]_{nm} = W_n[\mathbf{f}]_{nm} - [\mathbf{f}]_{nm}W_m$$

$$= h\omega(n,m)[\mathbf{f}]_{nm} = \left(\frac{h}{2\pi i}\right)[\dot{\mathbf{f}}]_{nm}$$
(5.85)

or, equivalently,

$$[\dot{\mathbf{f}}] = \frac{2i\pi}{h} [\mathbf{W}, \mathbf{f}]$$
(5.86)

where **W** is a diagonal matrix with elements $[W_{nm}] = \delta_{nm}W_n$. Now, in the above equation if we set $\mathbf{f} = \mathbf{q}$ and use the Hamilton's equation of motions (5.73), we get

$$\frac{\partial \mathbf{H}}{\partial \mathbf{p}} = \frac{2\pi i}{h} \left[\mathbf{W}, \mathbf{q} \right] \tag{5.87}$$

By taking $f = \mathbf{H}$ in (5.82), we can rewrite the above equations as

$$Wq - qW = Hq - qH$$
(5.88)

or, equivalently

$$q(n,m) (H_n - H_m) = q(n,m) (W_n - W_m)$$
(5.89)

where we have used the fact that \mathbf{H} is a diagonal matrix as we showed above. The above relation implies that

$$\omega(n,m) = \frac{(H_n - H_m)}{\hbar}$$
(5.90)

Note that the above relation between $(E_n - E_m)$ and ω is derived based on the Born's equation (5.76), unlike in the old quantum theory where it was postulated.

In the fall of 1925, Paul Dirac, still a PhD student, wrote a very important paper where he gave an algebraic version of quantum mechanics, known as q-number algebra, which reproduced many of the results obtained by Born, Heisenberg, and Jordan.

5.4 Equivalence between Wave Mechanics and Matrix mechanics

So far we have seen two formulations of the quantum mechanics: Schrodinger wave mechanics and Heisenberg matrix mechanics. Even though they look different from one another, both succeeded in explaining the same physical phenomena, such as the spectral lines of the hydrogen, and the quantization of the energies of the harmonic oscillator⁹⁹. This lead Schrodinger to seek a reconciliation of the two approaches, which resulted in a paper in March of 1926¹⁰⁰.

⁹⁹One should point out that these two founders of Quantum mechanics was unhappy with the other's approach. In a letter **Heisenberg** wrote to Pauli : "The more I ponder about the physical part of Schrodinger's theory, the more disgusting it seems to me.". Similarly, **Shcrodinger** stated that "I was discouraged, if not repelled, by what appeared to me rather difficult method of transcendental algebra, defying any visualization.".

¹⁰⁰E. Schrodinger, "Uber das Verhaltnis der Heisenberg-Born- Jordan Quanten-mechanik zu der meinen", Annalen der Physik (4), **79**, 734-756.

Schrödinger noted that the relation (5.76) can be realized in terms of operators where for each coordinate q_i the momentum conjugate p_i is to be replaced by the operator $K\partial/\partial q_i$, with K is a constant to be determined later. For brevity, in what follows we denote the set of variables $(q_1, q_2, ..., q_N)$ and $(p_1, p_2, ..., p_N)$ by simply q and p, respectively. He then considered a well-ordered operator $\mathcal{F}(q, p)$, function of q's and p's of the the form

$$\mathcal{F}(q,p) = f(q)p_r p_s p_t g(q) p_{r'} h(q) p_{r''} p_{s''}.$$
(5.91)

and to this he assigns

$$\hat{\mathcal{F}} = f(q)K^3 \frac{\partial^3}{\partial q_r q_s q_t} g(q)K \frac{\partial}{\partial q_{r'}} h(q)K^2 \frac{\partial^3}{\partial q_{r''} q_{s''}}.$$
(5.92)

where the differential operators act on all the factors on the right. So, the action of $\hat{\mathcal{F}}$ on some function u(q) yields another function of q's. Now let us chose a set of complete orthonormal functions in the space of the coordinates:

$$u_1(q)\sqrt{\rho(q)}, \ u_2(q)\sqrt{\rho(q)}, \ u_3(q)\sqrt{\rho(q)}, \dots ect$$
 (5.93)

So, we have

$$\int \rho(q)u_a(q)u_b(q)dq = \delta_{ab}$$
(5.94)

Here $\int dq$ is a short notation for the integration over the whole space of the q's. The density function $\rho(q)$ is was introduced to guarantee that ua(q) are orthonormal and the integrands are self-adjoint. Now to the operator $\hat{\mathcal{F}}$ one associates the following matrix¹⁰¹

$$\mathcal{F}^{ab} = \int \rho(q) u_a(q) \left[\hat{\mathcal{F}} u_b(q) \right] dq \tag{5.95}$$

In addition, Schrodinger defines another operator $\overline{\hat{\mathcal{F}}}$ by¹⁰²

$$\overline{\hat{\mathcal{F}}} = (-1)^{\tau} \dots K^2 \frac{\partial^3}{\partial q_{r''} q_{s''}} h(q) K \frac{\partial}{\partial q_{r'}} K^3 \frac{\partial^3}{\partial q_r q_s q_t} g(q) f(q) \dots$$
(5.96)

where τ is the number of derivatives in the operator $\hat{\mathcal{F}}$. Assuming that the functions $u_a(q)$ and their derivatives vanish at the boundaries of the q-space, the expression (??) can be written as

$$\mathcal{F}^{ab} = \int u_b(q) \left[\overline{\hat{\mathcal{F}}} \rho(q) u_a(q) \right] dq$$
(5.97)

¹⁰¹In his paper, Schrödinger, denotes the operator associated $\mathcal{F}(q,p)$ by \mathcal{F} and the new function obtained upon the action of \mathcal{F} on u(q) by $[\mathcal{F}, u_a]$. To avoid confusion with the symbol [,] for the commutator, I used instead $\hat{\mathcal{F}}$ for operator and denoted its action on u(q) by $[\hat{\mathcal{F}}u(q)]$.

¹⁰²Schrodinger called this operator "gewwallze", which in English means "rolled over operator".

Now, if $\hat{\mathcal{G}}$ is another operator, then by using the above equation, we can write

$$\sum_{b} \mathcal{F}^{ab} G^{bc} = \sum_{b} \int u_{b}(q) \left[\overline{\hat{\mathcal{F}}} \rho(q) u_{a}(q) \right] dq \times \int \rho(q') u_{b}(q') \left[\hat{\mathcal{G}} u_{c}(q') \right] dq' \quad (5.98)$$
$$= \int \left[\overline{\hat{\mathcal{F}}} \rho(q) u_{a}(q) \right] \left[\mathcal{G} u_{c}(q) \right] dq$$

where in obtaining the last equality we used the relation $\sum_{b} \rho(q_1) u_b(q_1) u_b(q) = \delta^{(N)}(q_1 - q)$, which is simply the relation of completeness of $u'_b s$, and then integrated over the variables q_1 . Now, transforming from $\overline{\hat{\mathcal{F}}}$ to $\hat{\mathcal{F}}$ we get

$$\sum_{b} \mathcal{F}^{ab} G^{bc} = \int \rho(q) u_a(q) \left[\hat{\mathcal{F}} \left[\hat{\mathcal{G}} u_c(q) \right] \right] dq$$
(5.99)

or, equivalently,

$$\left(\hat{\mathcal{F}}\hat{\mathcal{G}}\right)^{ac} = \sum_{b} \mathcal{F}^{ab}Gbc \tag{5.100}$$

This results shows that the quantities \mathcal{F}^{ab} and \mathcal{G}^{ab} defined by (5.95) satisfy the rule of matrix multiplication.

So, now we consider the matrix representation of the operators \hat{q} and \hat{p} . We have

$$(q_i)^{ab} = \int \rho(q) u_a(q) [\hat{q}_i u_b(q)] dq = \int q_i \rho(q) u_a(q) u_b(q) dq$$

$$(p_i)^{ab} = \int \rho(q) u_a(q) [\hat{p}_i u_b(q)] dq = K \int \rho(q) u_a(q) \frac{\partial u_b(q)}{\partial q_i} dq$$
(5.101)

Taking $K = -i\hbar$, the commutation relation $(\hat{p}_i \hat{q}_j - \hat{q}_j \hat{p}_j)$ corresponds to the matrix

$$\left(p_i q_j - q_j p_j\right)^{ab} = -i\hbar \int \rho(q) u_a(q) u_b(q) dq = -i\hbar \delta^{ab}$$
(5.102)

which is exactly the Born-Heisenberg quantization (5.76).

Now, if we choose as a set of orthonormal functions, the eigenfunctions of the Hamiltonian operator¹⁰³, i.e. $\hat{\mathcal{H}}u_a(q) = E_a u_a(q)$, then we have

$$\mathcal{H}^{ab} = \int \rho(q) u_a(q) [\hat{\mathcal{H}} u_b(q)] dq = E_b \delta_{ab}$$
(5.103)

where we used the orthonormality of the functions $\{u_a(q)\}$. So, the matrix \mathcal{H}^{ab} is diagonal and we have

$$\left[(q_i \mathcal{H}) - (\mathcal{H}q_i) \right]^{ab} = \left[\sum_c (q_i)^{ac} \mathcal{H}^{cb} - \sum_c \mathcal{H}^{ac} (q_i)^{cb} \right]$$

$$= E_b(q_i)^{ab} - E_a(q_i)_{ab}$$

$$= -i\hbar \left(\frac{dq_i}{dt} \right)^{ab}$$
(5.104)

¹⁰³It is very easy to show that the eigenfunctions of $\hat{\mathcal{H}}$ form a complete set of orthonormal functions.

or,

$$-i\hbar \left(\frac{dq_i}{dt}\right)^{ab} = [\hat{\mathcal{H}}, q]^{ab}$$
(5.105)

which is equivalent to the Hamilton equations for the matrix $(q_i)^{ab}$ in (5.80) derived by Born, Heisenberg and Jordan. Similarly, we have for the momentum

$$-i\hbar\left(\frac{dp_i}{dt}\right)^{ab} = [\hat{\mathcal{H}}, p]^{ab}$$
(5.106)

6 The Mathematical Formalism of Quantum Mechanics

In this subsection, we will introduce the notions of vector space, Hilbert space.

6.1 Pre-Hilbert Space

(a) Linear Vector Space and Kets :

A complex linear vector space¹⁰⁴, denoted $\mathcal{V}(\mathbb{C})$, is a set of elements called vectors, which satisfy the following axioms:

- (a) There is an operation "+" such that $(\mathcal{V}, +)$ is an abelian group¹⁰⁵.
- (b) The product of any complex number a, with any vector v_i is a vector, i.e. $av \in \mathbb{C}$.
- (c) For any complex numbers $a, b \in \mathbb{C}$ and vectors $v, \omega \in \mathcal{V}$, the following must be satisfied
 - $a(bv) = (a \times b)v;$
 - 1v = v;
 - $a(v+\omega) = av + a\omega;$
 - (a+b)v = av + bv.

where \times is the multiplication law of complex numbers, 1 is the identity element in \mathbb{C} . A real vector space has satisfy the same axioms except that a, b are numbers in \mathbb{R} , not in \mathbb{C} .

Throughout these notes, we will use Dirac notation to denote the elements of \mathcal{V} where a vector α is represented by the symbol $|\alpha\rangle$, called the **ket** α ", and the vector $(-\alpha)$ by $|-\alpha\rangle$. However, I will use $\vec{0}$ to denote the null vector. This notation, besides it avoids confusing a number with a scalar, proved to be useful in quantum mechanics.

- (a) **Closure**: $\forall g_i, g_j; g_i.g_j \in G$
- (b) Associativity: $\forall g_i, g_j, g_k; (g_i.g_j).g_k = g_i.(g_j.g_k)$
- (c) Existence of the identity element : $\exists e \in G; g.e = e.g = g, \forall g \in G$
- (d) Every element has an inverse: $\forall g_i \in G, \exists g_i^{-1}; g_i.g_i^{-1} = g_i^{-1}.g_i = e.$

¹⁰⁴In general a vector space is defined over a **field**, \mathbb{F} . A field is a set \mathbb{F} equipped with two binary operations, one called addition " + " and the other multiplication " \cdot " such that :

⁽i) $(\mathbb{F}, +)$ is abelian group with identity element $0_{\mathbb{F}}$,

⁽*ii*) $(\mathbb{F} - \{0_{\mathbb{F}}\}, \cdot)$ is abelian group with identity element $1_{\mathbb{F}}$,

⁽iii) The multiplication " \cdot " is distributive on the addition " + ",

 $⁽iv) \ 0_{\mathbb{F}} \neq 1_{\mathbb{F}}.$

¹⁰⁵A group is a set of elements $G = \{g_1, g_2, ..\}$ with a multiplication law "." such that the following axioms are satisfied:

If for any two elements g_i, g_j of a group G, their **commutator** $[g_i, g_j] \equiv g_i g_j g_i^{-1} g_j^{-1}$ is equal to identity element, then G is called **abelian**. For such groups, we denote the multiplication law by "+", the identity element by "0", and the inverse of an element g_i by " $-g_i$ ".

The set of vectors $\{[v_1 >, |v_2 >, ..., |v_n >\} \in \mathcal{V}$ are said to be linearly independent, if $\sum_{i=1}^{n} \alpha_i | v_i \rangle \ge \vec{0}$ only if the coefficients $\alpha_1 = \alpha_2 = ... = \alpha_n = 0$. If there exist at most N linearly independent vectors, then N is called the dimension of \mathcal{V} , and we write dimV = N. In this case, a set $\{|e_1 >, |e_2 > ..., |e_N >\}$ of linearly independent vectors is called a basis of the vector space. This implies that in a given a basis, any vector |v > can be written in a form

$$|v\rangle = \sum_{i=1}^{n} a_i |e_i\rangle$$
 (6.1)

where the coefficients a_i represent the components of the vector $|v\rangle$ and they are uniquely determined. These definition generalize to the case of infinite dimensional vector space. For finite dimensional vector spaces one has the following result:

Theorem 1:

Any n-dimensional vector space over a field \mathbb{K} is linearly isomorphic to \mathbb{K}^n .

(b) Inner Product space:

An inner product, or sometimes called the scalar product, in vector space is a map from $\mathcal{V} \times \mathcal{V}$ to \mathbb{C}^{106} which for each ordered pair of vectors $|\psi\rangle$ and $|\phi\rangle$, associates a number in \mathbb{C} , denoted by (.,.):

$$(.,.): \mathcal{V} \times \mathcal{V} \to \mathbb{C} \tag{6.2}$$

such that for any vectors $|\psi\rangle$, $|\phi\rangle$ and χ in \mathcal{V} and any scalars a and $b \in \mathbb{C}$, the following properties must be satisfied:

- (a) $(|\psi\rangle, a|\phi\rangle + b|\chi\rangle) = a(|\psi\rangle, \phi\rangle) + b(|\psi\rangle, \chi\rangle)$ (linear in the second argument);
- (b) $(|\psi|>, \phi>) = (|\phi>, |\psi>)^*$ (Hermitian symmetric);
- (c) $(|\psi\rangle, |\psi\rangle) \ge 0$ (non-negative);
- (d) $(|\psi \rangle, |\psi \rangle) = 0$ if and only if $|\psi \rangle$ is a null vector (positive definite).

Note that the inner product is antiinear in \mathbb{C} in the first argument, where as linear in the second argument¹⁰⁷. We say that two vectors $|\psi\rangle$ and $|\phi\rangle$ are said to be orthogonal if their inner product vanishes, namely $(|\psi\rangle, |\psi\rangle) = 0$. A vector space endowed with an inner product is said to be an inner product

 $^{^{106} \}mathrm{for}$ real vector space instead of $\mathbb C$ we have $\mathbb R$

¹⁰⁷For a vector space over real number, the inner product is linear.

space or, sometimes it is called **pre-Hilbert space**.

Another important concept in vector spaces is the notion of **normed linear space**. A norm on a vector space is a non-negative real function such that, if $|\psi\rangle$, $|\phi\rangle$ are vectors, the norm of $|\psi\rangle$, denoted as $||\psi||$, satisfying:

- (a) $||\psi|| \ge 0$; $||\psi|| = 0$ if and only if $|\psi\rangle$ is a null vector,
- (b) $||a\psi|| = |a|.||\psi||,$
- (c) $||\psi + \phi|| \le ||\psi|| + ||\phi||.$

So, given a scalar product, it is natural define the norm of a vector $|\psi\rangle$ by

$$||\psi|| = \sqrt{(|\psi, |\psi\rangle)} \tag{6.3}$$

Moreover, the above definition of $||\psi||$ satisfies the so called **Cauchy- Schwarz** inequality which states that for any two vectors $|\psi\rangle$ and $|\phi\rangle$ we have¹⁰⁸

$$|(|\psi>,|\phi>)| \le ||\psi|| \ ||\phi|| \tag{6.4}$$

The above inequality is crucial for ||.|| as defined in (6.3) to be a norm.

Theorem 2:

Any finite dimensional vector space can be equipped with a norm.

6.2 Dual Space and Bras

The dual space is the space of complex-linear valued maps from \mathcal{V} to \mathbb{C} , denoted by < . | (mathematicians call it 1-form):

$$< . |: \mathcal{V} \to \mathbb{C}$$
 (6.5)

so that for a given (ket) vector $|\phi\rangle$ it associates a number $\langle \phi|(|\psi\rangle) :\equiv \langle \phi|\psi\rangle$. The linearity of this mapping means that, for any two vectors $|\psi\rangle$ and $|\chi\rangle$, we have

$$<\phi|(a\psi>+b|\chi>) = a < \phi|\psi>+b < \phi|\chi>$$
(6.6)

Furthermore, we have

$$(a < \phi|) (\psi >) = a < \phi|\psi >, \forall a \in \mathcal{V}$$

$$(<\phi|+<\alpha|) (\psi >) = <\phi|\psi > + <\alpha|\psi >$$
(6.7)

Hence, the set of all linear mappings as defined above forms a vector space, which we denote by \mathcal{V}^* . A vector $\langle \psi |$ in the dual space is called **bra**. Moreover, if \mathcal{V} is a

¹⁰⁸The proof is as follows:

finite-dimensional, then \mathcal{V}^* is also, and with the same dimension as \mathcal{V} .

Now, given a an inner product, we can define a dual correspondence, denoted DC, between \mathcal{V} and \mathcal{V}^* :

$$DC: \quad \mathcal{V} \to \mathcal{V}^* \tag{6.8}$$
$$|\psi > \to <\psi|$$

such that the action of the bra $\langle \psi |$ on arbitrary ket $|\phi \rangle$ is given by

$$\langle \psi | \phi \rangle = (|\psi \rangle, |\phi \rangle) \tag{6.9}$$

Thus, from the properties of the inner product we deduce that the bra associated with of the ket $|\psi\rangle = a|\psi_1\rangle + b|\psi_2\rangle$ is

$$<\psi|=a^*<\psi_1|+b^*<\psi_2|$$
(6.10)

6.3 Hilbert Space

(a) Cauchy Sequence, Convergent Sequence and Complete Space:

Let \mathcal{V} be a normed space, and let $\{|\phi_n \rangle\}_{n=1}^{\infty}$, with $n \in \mathbb{N}$, be a sequence of elements of \mathcal{V} . The sequence is called **Cauchy sequence** if,

$$\forall \epsilon > 0, \ \exists N \in \mathbb{N}, \ \forall n, m \ge N, \ ||\phi_n - \phi_m|| < \epsilon \tag{6.11}$$

A sequence is said to be **convergent sequence** to $|\phi\rangle \in \mathcal{V}$ if

$$\forall \epsilon > 0, \ \exists N \in \mathbb{N}, \ \forall n \ge N, \ ||\phi - \phi_n|| < \epsilon \tag{6.12}$$

or, equivalently, $\lim_{n\to\infty} ||\phi - \phi_n|| < \epsilon$. Any convergent sequence in a linear normed space is a Cauchy sequence, whereas the converse is not generally true. Furthermore, every Cauchy sequence is bounded, i.e. there is some $|\phi\rangle \in \mathcal{V}$ and some $\alpha > 0$ such that $||\phi_n - \phi|| < \alpha$ all values of n. However, a bounded sequence in \mathcal{V} need not to be a Cauchy sequence¹⁰⁹.

(b) Banach Space:

A normed linear space that in which all Cauchy sequences are convergent is said to be **complete**, and it is also called a **Banach space**.

The following are examples of Banach space:

¹⁰⁹For example, $\{x_n = (-1)^n\}$ is a bounded but it is not Cauchy sequence.

•
$$\mathbb{R}^n = \{ |X\rangle = (x_1, x_2, ..., x_n), x_i \in \mathbb{R}, i = 1, 2, ..., n \}, \quad ||X|| := \sqrt{\sum_{i=1}^n x_i^2}.$$

•
$$\mathbb{C}^n = \{ |X\rangle = (x_1, x_2, ..., x_n), x_i \in \mathbb{C}, i = 1, 2, ..., n \}, \quad ||X|| := \sqrt{\sum_{i=1}^n |x_i|^2}.$$

•
$$l^2(\mathbb{N}) = \{ |X\rangle = \{x_n\}_{n=1}^{\infty}, x_n \in \mathbb{C} : \sum_{n=1}^{\infty} |x_n|^2 \}, ||X|| := \sqrt{\sum_{i=1}^{\infty} |x_i|^2}.$$

• $\mathcal{L}^{p}(\mathbb{R}) = \{ |f\rangle = f : \mathbb{R} \to \mathbb{C} : \int_{\mathbb{R}} |f(x)|^{p} dx < \infty \}, \quad ||f|| := \left(\int_{\mathbb{R}} |f(x)|^{p} dx \right)^{1/p}.$

A finite-dimensional vector space is complete in any possible norm. In infinite dimension, completeness generally depends on the norm.

(c) Hilbert Space:

Hilbert space, denoted by \mathcal{H} , is a complete vector space endowed with an inner product. In other words, a Hilbert space is a Banach space whose norm is determined by an inner product.

The following are examples of Hilbert space:

•
$$l^{2}(\mathbb{N}) = \{|X\rangle = \{x_{n}\}_{n=1}^{\infty}, x_{n} \in \mathbb{C} : \sum_{n=1}^{\infty} |x_{n}|^{2}\}, \text{ with}$$

 $< X|Y\rangle := \sum_{i=1}^{\infty} x_{i}^{*}y_{i}.$
• $\mathcal{L}^{2}(\mathbb{R}^{d}) = \{|f\rangle = f: \mathbb{R} \to \mathbb{C} : \int_{\mathbb{C}} |f(x_{1}, x_{2}, ...x_{d})|^{2} dx < \infty\}, \text{ with}$
 $< f|g\rangle := \int_{\mathbb{C}} f^{*}(x_{1}, x_{2}, ...x_{d})g(x_{1}, x_{2}, ..., x_{d}) dx_{1}dx_{2}..dx_{d}.$ (6.13)

Theorem:

Every infinite dimensional Hilbert space is isomorphic to $l^2(\mathbb{N})$.

6.4 Linear Operators

Operators in Hilbert space is a map which to each vector associate another vector. In other word, if $\hat{\mathcal{O}}$ is an operator, and a ket $|\phi\rangle$ of the Hilbert space \mathcal{V} , then we have

$$\hat{\mathcal{O}}|\phi\rangle = |\phi'\rangle \tag{6.14}$$

In particular, the **identity operator** (or the unit operator), denoted by **I**, has the property that for every vector $|\phi\rangle \in \mathcal{V}$:

$$\mathbf{I}|\phi\rangle = |\phi\rangle \tag{6.15}$$

Two operators $\hat{\mathcal{O}}_1$ and $\hat{\mathcal{O}}_2$ are said to be equal if

$$\hat{\mathcal{O}}_1 |\phi\rangle = \hat{\mathcal{O}}_2 |\phi\rangle, \quad \forall |\phi\rangle \in \mathcal{V}$$
(6.16)

An operator $\hat{\mathcal{L}}$ is said to be a **linear** on a vector space if for any vectors $|\phi_1 \rangle$ and $|\phi_2 \rangle$ in \mathcal{V} , we have

$$\hat{\mathcal{L}}(c_1|\phi_1 > +c_2|\phi_2 >) = c_1\hat{\mathcal{L}}|\phi_1 > +c_2\hat{\mathcal{L}}|\phi_2 >$$
(6.17)

Whereas, an operator $\hat{\mathcal{A}}$ is said to be **anti-linear** on a vector space if

$$\hat{\mathcal{A}}(c_1|\phi_1 > +c_2|\phi_2 >) = c_1^* \hat{\mathcal{A}}|\phi_1 > +c_2^* \hat{\mathcal{A}}|\phi_2 >$$
(6.18)

In quantum mechanics we will be mostly interested in linear operators, and so in these notes we will consider only linear operators, unless explicitly stated.

The set of linear operators form a group under the operation of addition. Also, multiplying a linear operator by a complex number is also a linear operator. In fact, the set of linear operators form a complex vector space in their own. Furthermore, we can define the product of two linear operators \mathcal{L}_1 and \mathcal{L}_2 , denoted by $\mathcal{L}_1\mathcal{L}_2$, as

$$(\mathcal{L}_1 \mathcal{L}_2) |\phi\rangle := \mathcal{L}_1 (\mathcal{L}_2 |\phi\rangle) \tag{6.19}$$

This multiplication of operators is associative, that is for any linear operators $\mathcal{L}_1, \mathcal{L}_2$ and \mathcal{L}_3 , we have

$$\left(\mathcal{L}_{1}\mathcal{L}_{2}\right)\mathcal{L}_{3}=\mathcal{L}_{1}\left(\mathcal{L}_{2}\mathcal{L}_{3}\right) \tag{6.20}$$

However, the product of operators is not commutative, i.e., in general, $\mathcal{A}_1 \mathcal{A}_2 \neq \mathcal{A}_2 \mathcal{A}_2$. The non-commutativity of two operators \mathcal{A}_1 and \mathcal{A}_2 is quantified by their commutator

$$[\mathcal{A}_1, \mathcal{A}_2] := \mathcal{L}_1 \mathcal{A}_2 - \mathcal{A}_2 \mathcal{A}_1 \tag{6.21}$$

which has the following properties

1. It is **Linear**

$$[a_1\mathcal{A}_1 + a_2\mathcal{A}_2, \mathcal{B}] = a_1 [\mathcal{A}_1, \mathcal{B}] + a_2 [\mathcal{A}_2, \mathcal{B}], \qquad (6.22)$$
$$[\mathcal{A}, b_1\mathcal{B}_1 + b_2\mathcal{B}_2] = b_1 [\mathcal{A}, \mathcal{B}_1] + b_2 [\mathcal{A}, \mathcal{B}_2],$$

2. It is antisymmetric

$$[\mathcal{A}, \mathcal{B}] = -[\mathcal{A}, \mathcal{B}] \tag{6.23}$$

3. It obeys the **Jacobi identity**

$$[\mathcal{A}_1, [\mathcal{A}_2, \mathcal{A}_3]] + [\mathcal{A}_3, [\mathcal{A}_1, \mathcal{A}_2]] + [\mathcal{A}_2, [\mathcal{A}_1, \mathcal{A}_1]] = 0$$
(6.24)

where \mathcal{A} 's and $\mathcal{B}'s$ are not necessarily linear operators, and the *a*'s and *b*'s are arbitrary complex numbers. With these properties, we say that the vector space of operators with the commutation relation (6.21) is a Lie algebra¹¹⁰. Another property that the commutator satisfies is the so called Leibniz rule of derivation

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

$$[\mathcal{A}, \mathcal{BC}] = \mathcal{B} [\mathcal{A}, \mathcal{C}] + [\mathcal{A}, \mathcal{B}] \mathcal{C}$$

$$(6.25)$$

• The action of Operator on a Bra :

As we saw, by definition an operators map kets in a vector space into other kets. However, we can also define their action on a bras to produce other bras. For an operator \mathcal{A} , we denote by $\langle \phi |$ by $\langle \phi | \mathcal{A}$ the action of \mathcal{A} on $\langle \phi |$ which results in other bra. The new bra $\langle \phi | \mathcal{A}$ is specified by it action on an arbitrary ket $|\psi \rangle$, which by definition is given by

$$(\langle \phi | \mathcal{A}) (|\psi \rangle) := \langle \phi | (\mathcal{A} | \psi \rangle) \equiv \langle \phi | \mathcal{A} | \psi \rangle$$
(6.26)

Hence, we can think of \mathcal{A} as acting either to the left (on the bra) or to the right (on the ket).

A special linear operator is the one denoted by $|\phi\rangle \langle \psi|$, which is defined by its action on an arbitrary ket $|\chi\rangle$:

$$(|\phi\rangle \langle \psi|) |\chi\rangle := |\phi\rangle \langle \psi||\chi\rangle \in \mathcal{V}$$
(6.27)

From the above definition it is not difficult to show that the action of $|\phi\rangle \langle \psi|$ on a bra $\langle \xi|$ is given by

$$\langle \xi | (|\phi \rangle \langle \psi|) := \langle \xi | \phi \rangle \langle \psi| \in \mathcal{V}^*$$
(6.28)

The operator $|\phi\rangle < \psi|$ is called the **outer product** of $|\phi\rangle$ and $\langle\psi\rangle$.

 $^{110}\mathrm{In}$ general, a Lie algebra is a vector space L over a field $\mathbb F$ together with a multiplication

$$L \times L \to L$$
$$(\mathcal{A}, \mathcal{B}) \to \mathcal{A} \odot \mathcal{B}$$

such for any $\mathcal{A}, \mathcal{B}, \mathcal{C}$ in L and for all $\lambda \in \mathbb{F}$ the following axioms are satisfied

- (a) $(\mathcal{A} + \mathcal{B}) \odot \mathcal{C} = \mathcal{A} \odot \mathcal{C} + \mathcal{B} \odot \mathcal{C}$, and $\mathcal{A} \odot (\mathcal{B} + \mathcal{B}) = \mathcal{A} \odot \mathcal{B} + \mathcal{A} \odot \mathcal{C}$,
- (b) $\lambda (\mathcal{A} \odot \mathcal{B}) = (\lambda \mathcal{A}) \odot \mathcal{B} = \mathcal{A} \odot (\lambda \mathcal{B}),$
- (c) $\mathcal{A} \odot \mathcal{B} = -\mathcal{B} \odot \mathcal{A}$ (or, equivalently $\mathcal{A} \odot \mathcal{A} = 0$ which leads to the antisymmetric property of \odot),
- (d) $(\mathcal{A} \odot \mathcal{B}) \odot \mathcal{C} + (\mathcal{C} \odot \mathcal{A}) \odot \mathcal{B} + (\mathcal{B} \odot \mathcal{C}) \odot \mathcal{A} = 0.$

The last property is called **Jacobi identity**. A vector space with a multiplication that satisfy only the properties (a) and (b) forms an **algebra**. If in addition to the above axioms $\mathcal{A} \odot (\mathcal{B} \odot \mathcal{C}) = (\mathcal{A} \odot \mathcal{B}) \odot \mathcal{C}$, we have a Lie algebra of an associative algebra.

• Spectrum of an Operator:

If $|\psi_a\rangle$ is a vector such that the action of an operator $\hat{\mathcal{O}}$ on it gives

$$\hat{\mathcal{O}}|\psi_{\lambda}\rangle = \lambda|\psi_{\lambda}\rangle \tag{6.29}$$

where λ is, in general, a complex number, then we say $|\psi_{\lambda}\rangle$ is **eigenstate** (or eigenket) of $\hat{\mathcal{O}}$ with **right- eigenvalue** λ . Similarly, if there exists a complex number λ such that the action of $\hat{\mathcal{O}}$ on a non zero vector $\langle \psi |$ of the dual vector space \mathcal{V}^* gives

$$\langle \psi_{\lambda} | \hat{\mathcal{O}} = \lambda \langle \psi_{\lambda} |$$
 (6.30)

then we say that $\langle \psi_{\lambda} |$ is **eigenstate** (or eigenbra) of $\hat{\mathcal{O}}$ with **left- eigenvalue** λ . In a finite dimensional vector space, every right-eigenvalue is also a left - eigenvalue. However, in infinite dimension vector space this is n general not the case.

The set of eigenvalues of an operator \mathcal{A} is called the **spectrum of** A. In finite dimensional vector space, this is a set of discrete points in the complex plane.

• Bounded Operator and Continuous Operator:

An operator is said to be **bounded operator** for any vector $|\psi\rangle$ in a vector space \mathcal{V} there exists a positive number k such that

$$||\mathcal{A}\psi|| \le k||\psi|| \tag{6.31}$$

For a given bounded operator \mathcal{A} , we define the supremum of \mathcal{A} to be the smallest number k for which the above inequality is satisfied for any vector $|\psi\rangle \in \mathcal{V}$, and it is denoted by

$$||\mathcal{A}|| := \sup \left\{ \frac{||\mathcal{A}\psi||}{||\psi||}, \ |\psi \rangle \neq 0 \right\}$$
 (6.32)

Hence we obtain

$$||\mathcal{A}|| \le ||\mathcal{A}||||\psi|| \tag{6.33}$$

Furthermore, we have

$$||\mathcal{A} + \mathcal{B}|| = \sup \left\{ \frac{||(\mathcal{A} + \mathcal{B})\psi||}{||\psi||}, |\psi\rangle \neq 0 \right\} = \sup \left\{ \frac{||\mathcal{A}\psi + \mathcal{B}\psi||}{||\psi||}, |\psi\rangle \neq 0 \right\}$$
$$\leq \left\{ \frac{||\mathcal{A}\psi||}{||\psi||}, |\psi\rangle \neq 0 \right\} + \left\{ \frac{||\mathcal{B}\psi||}{||\psi||}, |\psi\rangle \neq 0 \right\} = ||\mathcal{A}|| + ||\mathcal{B}|| \quad (6.34)$$

Now let us consider a convergent sequence $\{..., \psi_m, \psi_{m+1}, ..., \psi_n, ...\}$. If \mathcal{A} is a bounded operator, then $\{..., \mathcal{A}\psi_m, \mathcal{A}\psi_{m+1}, ..., \mathcal{A}\psi_n, ...\}$ is a convergent sequence. This is because

$$\lim_{n \to \infty} ||\mathcal{A}\psi_n - \mathcal{A}\psi|| \le ||\mathcal{A}|| \lim_{n \to \infty} ||\psi_n - \psi|| = 0$$
(6.35)

and we say that the bounded operator \mathcal{A} is a **continuous operator**.

• Inverse Operator:

Let \mathcal{A} be an operator. If there exists an operator \mathcal{B}_R such that $\mathcal{AB}_R = I$, then we say that \mathcal{B}_R is the **right inverse** of \mathcal{A} . Similarly, if there exists an operator \mathcal{B}_L such that $\mathcal{B}_L \mathcal{A} = I$, then \mathcal{B}_L is said to be the **left inverse** of \mathcal{A} . However, if there exist both a left and right inverses of \mathcal{A} , then we have

$$\mathcal{B}_L = \mathcal{B}_L I = \mathcal{B}_L \left(\mathcal{A} \mathcal{B}_R \right) = \left(\mathcal{B}_L \mathcal{A} \right) \mathcal{B}_R = \mathcal{B}_R \tag{6.36}$$

Therefore, operator of a given operator is unique and it is denoted by \mathcal{A}^{-1} , where

$$\mathcal{A}\mathcal{A}^{-1} = \mathcal{A}^{-1}\mathcal{A} \tag{6.37}$$

Also note that

$$(\mathcal{AB}) \mathcal{B}^{-1} \mathcal{A}^{-1} = \mathcal{A} (\mathcal{BB}^{-1}) \mathcal{A}^{-1}$$

= $\mathcal{AA}^{-1} = I$ (6.38)

where we used the property of associativity of the product of operators. Thus, the above equation implies that $(\mathcal{AB})^{-1} = \mathcal{B}^{-1}\mathcal{A}^{-1}$.

• Adjoint Operator:

The adjoint of an operator \mathcal{A} , denoted by \mathcal{A}^{\dagger} , is an operator acting on the dual space such that

$$\left(\langle \psi | \mathcal{A}^{\dagger} \right)^{\dagger} := \mathcal{A} | \psi \rangle \tag{6.39}$$

From with this definition we note that if \mathcal{A} is a linear operator, so is the operator \mathcal{A}^{\dagger} . Furthermore, as a consequence of (6.39), we have the following properties

$$\langle \phi | \mathcal{A}^{\dagger} | \psi \rangle = \langle \psi | \mathcal{A}^{\dagger} | \phi \rangle^{*}$$
 (6.40)

$$(\mathcal{A})^{\dagger} = \mathcal{A} \tag{6.41}$$

$$\left(\mathcal{A}^{\dagger}\right)^{-1} = \left(\mathcal{A}^{-1}\right)^{\dagger} \tag{6.42}$$

$$(a\mathcal{A} + b\mathcal{B})^{\dagger} = a^*\mathcal{A}^{\dagger} + b^*\mathcal{B}^{\dagger}$$
(6.43)

$$(\mathcal{A}\mathcal{B}) = \mathcal{B}^{\dagger}\mathcal{A}^{\dagger} \tag{6.44}$$

$$(|\phi\rangle \langle \psi|)^{\dagger} = |\psi\rangle \langle \phi| \tag{6.45}$$

6.5 Hermitian and Anti-Hermitian

An operator \mathcal{A} is said to be **Hermitian** if $\mathcal{A}^{\dagger} = \mathcal{A}$, i.e. for all $|\phi\rangle$ and $|\psi\rangle$ in the vector space, we have

$$\langle \phi | \hat{\mathcal{A}} | \psi \rangle = \langle \psi | \hat{\mathcal{A}} | \phi \rangle^*$$
 (6.46)

where we used (6.40) of the definition of adjoint operator. Thus, if \mathcal{A} is hermitian $\langle \psi | \mathcal{A} | \psi \rangle$ is a real number, for all vectors $| \psi \rangle$. On the other hand, if every

expectation value of \mathcal{A} is real, then \mathcal{A} is hermitian. We say that a Hermitian operator A is **positive definite** if

$$\langle \psi | \hat{\mathcal{A}} | \psi \rangle > 0, \quad \forall | \psi \rangle$$

$$(6.47)$$

and it is said to be non negative if $\langle \psi | \hat{\mathcal{A}} | \psi \rangle \geq 0$. In quantum mechanics (QM), Hermitian operator, say \mathcal{A} , is associated with an observables, denoted by A, and often one is interested in the mean value or the **expectation value** of A. As we shall see, in QM a sate of the system is represented by a ket vector in a Hilbert space, say $|\psi\rangle$, and one define the **expectation value** of A in this state by

$$\langle A \rangle_{\psi} := \langle \psi | \hat{\mathcal{A}} | \psi \rangle$$
 (6.48)

For example, the Hamiltonian of a Physical system must be Hermitian since it is associated with its energy which is an observable quantity. Another quantity that characterizes an observable A is the **root-mean-square deviation**, $(\Delta A)_{\psi}$, defined as

$$(\Delta A)_{\psi} = \langle A^2 \rangle_{\psi} - \langle A \rangle_{\psi}^2 \tag{6.49}$$

Thus, the mean -square root deviation of an observable A in the state $|\psi\rangle$ is zero if $|\psi\rangle$ is an eigenstate of its associated operator \hat{A} .

Similarly, an operator is said to be **anti-Hermitian** if $\hat{\mathcal{A}}^{\dagger} = -\hat{\mathcal{A}}$. Consequently, for every state $|\psi\rangle$ we can write $\langle \psi | \hat{\mathcal{A}} | \psi \rangle = \left(\langle \psi | \hat{\mathcal{A}}^{\dagger} | \psi \rangle \right)^* = -\left(\langle \psi | \hat{\mathcal{A}} | \psi \rangle \right)^*$. Thus, the expectation value of an anti-Hermitian operator must be imaginary. Also, multiplying a Hermitian operator by an imaginary results in an anti-Hermitian operator.

Given an arbitrary operator A we can decompose it into a sum of a Hermitian and anti-Hermitian operator as:

$$\hat{\mathcal{A}} = \left(\frac{\hat{\mathcal{A}} + \hat{\mathcal{A}}^{\dagger}}{2}\right) + \left(\frac{\hat{\mathcal{A}} - \hat{\mathcal{A}}^{\dagger}}{2}\right) \tag{6.50}$$

Another property of Hermitian operators is that the product of two Hermitian operators is a Hermitian operator if they commute.

An important feature of Hermitian operators is that its eigenvalues are all real and any two eigenvectors are orthogonal to each other if the corresponding if the corresponding eigenvalues are not equal. To show this let $|\phi_a\rangle$ be a eigenstate of a Hermitian operator \mathcal{A} with a corresponding eigenvalue a, i.e.

$$\hat{\mathcal{A}}|\phi_a\rangle = a|\phi_a\rangle \tag{6.51}$$

Then we have

$$\langle \phi_a | \hat{\mathcal{A}} | \phi_a \rangle = a ||\phi_a|| \tag{6.52}$$

However since \mathcal{A} is Hermitian we can write

$$\langle \phi_a | \hat{\mathcal{A}} | \phi_a \rangle = (\hat{\mathcal{A}} | \phi_a \rangle)^{\dagger} (|\phi_a \rangle) = a^* ||\phi_a||$$
(6.53)

Comparing (6.52) and (6.53), we get $a = a^*$, and so the eigenvalue is real.

Now suppose that $|\phi_a\rangle$ and $|\phi_b\rangle$ are two the eigenstates of \mathcal{A} with the eigenvalues a and b, respectively:

$$\hat{\mathcal{A}}|\phi_a\rangle = a|\phi_a\rangle \quad , \quad \mathcal{A}|\phi_b\rangle = b|\phi_a\rangle \tag{6.54}$$

Since \mathcal{A} is Hermitian we can write

$$(b-a) < \phi_a | \phi_b >= 0 \tag{6.55}$$

This implies that as long as the eigenvalues a and b are not equal, we have $\langle \phi_a | \phi_b \rangle = 0$, that is the corresponding eigenstates are orthogonal to each other.

6.6 Isometric and Unitary Operators

An operator Ω is said to be an **isometric** operator if

$$\hat{\Omega}^{\dagger}\hat{\Omega} = I \tag{6.56}$$

It also means that for all $|\phi\rangle$ and $|\psi\rangle$ in the vector space, if $|\psi\rangle \rightarrow \hat{\Omega}|\psi\rangle$, and $|\phi\rangle \rightarrow \hat{\Omega}|\phi\rangle$, then we have

$$(\langle \psi | \hat{\Omega}^{\dagger}) \left(\hat{\Omega} | \phi \rangle \right) = \langle \psi | \hat{\Omega}^{\dagger} \hat{\Omega} | \phi \rangle = \langle \psi | \phi \rangle$$
(6.57)

Thus, unitary operator preserves the inner product of any two vectors.

An operator \mathcal{U} is said to be unitary if

$$\hat{\mathcal{U}}\hat{\mathcal{U}}^{\dagger} = \hat{\mathcal{U}}^{\dagger}\hat{\mathcal{U}} = I \tag{6.58}$$

or, equivalently $\hat{\mathcal{U}}^{-1} = \hat{\mathcal{U}}^{\dagger}$. In finite dimensional vector spaces, an isometric operator is unitary. However, in infinite dimensional vector spaces this is in general not the case.

An important result of unitary operator, which I leave it as an exercise for the reader to prove it, is that its eigenvalues are pure phases and the eigenstates of two distinct eigenvalues are orthogonal to one another.

6.7 Projection Operator

Let \mathcal{H} be a Hilbert space and $\mathcal{M} \subset \mathcal{H}$ be an arbitrary subset of \mathcal{H} . Then the orthogonal of complement space \mathcal{M}^{\perp} of \mathcal{M} is defined as

$$\mathcal{M}^{\perp} = \{ |\phi_{\mathcal{M}^{\perp}} \rangle \in \mathcal{H} : \langle \psi_{\mathcal{M}} | \phi_{\mathcal{M}^{\perp}} \rangle = 0, \ \forall \ |\psi_{\mathcal{M}} \rangle \in \mathcal{M} \}$$
(6.59)

The orthogonal complement \mathcal{M}^{\perp} of a subset \mathcal{M} of a Hilbert space \mathcal{H} is Hilbert subspace. So, any vector in \mathcal{H} can be decomposed in a unique way into a vector $|\psi_{\mathcal{M}}\rangle$ in \mathcal{M} and a vector $|\phi_{\mathcal{M}^{\perp}}\rangle$ in \mathcal{M}^{\perp} such that

$$|\psi\rangle = |\psi_{\mathcal{M}}\rangle + |\phi_{\mathcal{M}^{\perp}}\rangle \tag{6.60}$$

We define the **projection** operator $\hat{\mathcal{P}}_{\mathcal{M}}$ on a subspace \mathcal{M} to be such that when acting on a vector $|\psi\rangle \in \mathcal{H}$ results in a vector in the Hilbert subspace \mathcal{M} , i.e.

$$\hat{\mathcal{P}}_{\mathcal{M}}|\psi\rangle = |\psi_{\mathcal{M}}\rangle \tag{6.61}$$

As a result, $\hat{\mathcal{P}}_{\mathcal{M}}|\psi_{\mathcal{M}}\rangle = |\psi_{\mathcal{M}}\rangle$ and $\hat{\mathcal{P}}_{\mathcal{M}}|\psi_{\mathcal{M}^{\perp}}\rangle = 0$. Furthermore, we have the following

$$<\psi|\hat{\mathcal{P}}_{\mathcal{M}}|\phi> = <\psi|\phi_{\mathcal{M}}> = (<\psi_{\mathcal{M}}|+<\psi_{\mathcal{M}^{\perp}}|)\phi_{\mathcal{M}}> = <\psi_{\mathcal{M}}|\phi_{\mathcal{M}}> \qquad (6.62)$$
$$= <\psi_{\mathcal{M}}|\phi> = (\hat{\mathcal{P}}|\psi>)^{\dagger}|\phi_{\mathcal{M}}> = <\psi|\hat{\mathcal{P}}_{\mathcal{M}}^{\dagger}|\phi>$$

which implies that

$$\hat{\mathcal{P}}_{\mathcal{M}}^{\dagger} = \hat{\mathcal{P}}_{\mathcal{M}} \tag{6.63}$$

Thus the projector operator $\hat{\mathcal{P}}_{\mathcal{M}}$ is Hermitian. Similarly, we define $\hat{\mathcal{P}}_{\mathcal{M}^{\perp}}$ to be such that when on a vector $|\psi\rangle \in \mathcal{H}$ results in a vector in the Hilbert subspace \mathcal{M}^{\perp} , i.e.

$$\hat{\mathcal{P}}_{\mathcal{M}^{\perp}}|\psi\rangle = |\psi_{\mathcal{M}^{\perp}}\rangle \tag{6.64}$$

and which is also Hermitian. From the definitions of $\hat{\mathcal{P}}_{\mathcal{M}}$ and $\hat{\mathcal{P}}_{\mathcal{M}^{\perp}}$, we have $\hat{\mathcal{P}}_{\mathcal{M}} + \hat{\mathcal{P}}_{\mathcal{M}^{\perp}} = I$. Furthermore, they have the property that

$$\hat{\mathcal{P}}_{\mathcal{M}}^2 = \hat{\mathcal{P}}_{\mathcal{M}}, \quad \hat{\mathcal{P}}_{\mathcal{M}^{\perp}}^2 = \hat{\mathcal{P}}_{\mathcal{M}^{\perp}} \tag{6.65}$$

This means that the projection operators are **idempotent** operators.

6.8 Matrix Representation

Now, let us consider a finite dimensional vector space \mathcal{V} of dimension N, and let the set of vectors $\{|e_1\rangle, |e_2\rangle, |e_3\rangle, ..., |e_N\rangle\}$ be a basis in \mathcal{V} . Then, as we discussed previously, any vector ψ can be written uniquely as $|\psi\rangle = \sum_{i=1}^{N} a_i |e_n\rangle$. This also means that in the basis $\{|e_i\rangle\}$ we can represent the state $|\psi\rangle$ by a column vector in \mathbb{C}^N with components $\{a_i, i = 1, ..N\}$:

$$|\psi\rangle \longrightarrow \begin{pmatrix} a_1\\a_2\\\vdots\\a_n \end{pmatrix} = \begin{pmatrix} \langle e_1|\psi\rangle\\\langle e_2|\psi\rangle\\\vdots\\\langle e_n|\psi\rangle \end{pmatrix}$$
(6.66)

Similarly, a bra vector $\langle \psi | = (|\psi \rangle)^{\dagger}$ can be represented in \mathbb{C}^N by a row vector in \mathbb{C}^N with components $\{a_i^*, i = 1, ...N\}$:

$$\langle \psi | \longrightarrow \left(a_1^*, a_2^*, \cdots, a_n^* \right) = \left(\langle \psi | e_1 \rangle, \langle \psi | e_2 \rangle, \cdots, \langle \psi | e_n \rangle \right)$$
(6.67)

Now, consider an operator $\hat{\mathcal{A}}$ acting on the ket $|\psi\rangle$ such that

$$\hat{\mathcal{A}}|\psi\rangle = |\psi'\rangle \tag{6.68}$$

where $|\psi\rangle$ and $|\psi'\rangle$ are represented by column vectors with components $\{a_i\}$ and $\{a'_i\}$, respectively. Then, by multiplying both sides of the above equation by the bra $\langle e_n|$ and inserting the identity operator $\mathbb{I} = \sum_m |e_m\rangle \langle e_m|$ between $\hat{\mathcal{A}}$ and $|\psi\rangle$, we obtain

$$\sum_{m} \hat{\mathcal{A}}_{nm} a_m = a'_n \tag{6.69}$$

where

$$\hat{\mathcal{A}}_{nm} = \langle e_n | \hat{\mathcal{A}} | e_m \rangle \tag{6.70}$$

which is the (nm) element of the matrix $[\mathcal{A}]$. Hence, in a given basis, say $\{e_i\}$, an operator $\hat{\mathcal{A}}$ can be represented by a matrix with elements given in

$$\hat{\mathcal{A}} \longrightarrow [\hat{\mathcal{A}}] = \begin{pmatrix} \langle e_1 | \hat{\mathcal{A}} | e_1 \rangle \langle e_1 | \hat{\mathcal{A}} | e_2 \rangle \cdots \langle e_1 | \hat{\mathcal{A}} | e_n \rangle \\ \langle e_2 | \hat{\mathcal{A}} | e_1 \rangle \langle e_2 | \hat{\mathcal{A}} | e_2 \rangle \cdots \langle e_2 | \hat{\mathcal{A}} | e_n \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle e_n | \hat{\mathcal{A}} | e_1 \rangle \langle e_n | \hat{\mathcal{A}} | e_2 \rangle \cdots \langle e_n | \hat{\mathcal{A}} | e_n \rangle \end{pmatrix}$$
(6.71)

As particular case, consider \mathcal{A} to be Hermitian operator with eigenstates $\{|\phi_n \rangle\}$ and real eigenvalues $\{a_n\}$, with n = 1, 2, ... If all the eigenvalues are distinct, then, as we showed earlier, $\{|\phi_n \rangle, n = 1, 2, ...\}$ form an orthonormal basis. In this basis, we have

$$\langle \phi_n | \hat{\mathcal{A}} | \phi_m \rangle = a_m \langle \phi_n | \phi_m \rangle = \delta_{nm} a_m$$
 (6.72)

So, the matrix representation of the Hermitian operator \mathcal{A} is given by

$$\hat{\mathcal{A}} \longrightarrow [\hat{\mathcal{A}}] = \begin{pmatrix} a_1 & 0 & \cdots & 0 \\ 0 & a_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_n \end{pmatrix}$$
(6.73)

Thus, in the basis formed by its eigenstates, a Hermitian operator is represented by a diagonal matrix with the diagonal elements given by its eigenvalues. Such representation is often called the **A- representation**.

The above discussion applies also to the case of an infinite vector space in which a basis vectors is labeled by some continuous parameter. In such case the discrete sum is replaced by an integral. For instance, suppose that $\hat{\mathcal{A}}$ is some operator with a continuous set of eigenvalues $\{a \in \mathcal{I}\}\$ and with corresponding eigenstates $\{|a\rangle\}$. Then, the orthonormality of the eigenstates and the completeness relation read

$$\langle a|a' \rangle = \delta(a-a'), \qquad \int_{a \in \mathcal{I}} |a \rangle \langle a| \ da = \mathbb{I}$$
 (6.74)

where $\delta(a - a')$ is the Dirac delta function¹¹¹. Thus, an arbitrary state $|\phi\rangle$ can be expanded in the $\{|a\rangle\}$ basis as

$$|\phi\rangle = \mathbb{I} |\phi\rangle = \int_{a\in\mathcal{I}} |a\rangle \langle a|\phi\rangle \ da \tag{6.75}$$

In this continuous basis, the matrix representation of the operator $\hat{\mathcal{A}}$ is given by

$$[\hat{\mathcal{A}}](a,a') = \langle a|\hat{\mathcal{A}}|a'\rangle = a\delta(a-a') \tag{6.76}$$

6.9 Tensor Products

•

$$\int_{x_1}^{x_2} f(x)\delta(x-x_0) = f(x_0), \text{ with } x_0 \in [x_1, x_2]$$

¹¹¹Actually $\delta(a - a')$ is not a function; it is a distribution which maps nice well-behaved functions f to a some (complex) number given by

7 Principles of Quantum Mechanics

7.1 The Postulates of Quantum Mechanics

(a) Postulate 1: The State of the System

Every physical system has an associated Hilbert space \mathcal{H} of some dimension $D_{\mathcal{H}}$.

Every physical state of a quantum system is described by a \mathbf{ray} in \mathcal{H} .

by definition, a **ray**, denoted by \mathcal{R}_{ψ} , is a set of vectors which differ from some vector $|\psi\rangle$ by multiplication by non zero complex number, ie

$$\mathcal{R}_{\psi} = \{ |\phi\rangle, |\phi\rangle \sim |\psi\rangle \Rightarrow |\psi\rangle = a |\psi\rangle, \text{where} a \in \mathbb{C} - \{0\} \}$$
(7.1)

So a ray is an equivalence class of vectors which correspond to the same physical state; the normalization and phase of a vector are of no physical significance. Using this equivalence we chose a representative of the ray to be a **state** vector $|\psi(t)\rangle$ with unit norm $\langle \psi | \psi \rangle = 1$, which completely describe the system¹¹².

(b) Postulate 2: Observables

Every observable A of a physical system is represented by a Hermitian operator \hat{A} on \mathcal{H} whose eigenstates form a complete basis.

(c) Postulate 3: Measurement

Suppose the system is in some arbitrary state $|\psi_{\text{before}} \rangle$ in \mathcal{H} . If \mathcal{A} is a Hermitian operator representing some observable A, then

1. Every measurement of the observable A yields only one of the eigenvalues $\hat{\mathcal{A}}$.

2. For a discrete spectrum, with eigenvalues $\{a_i\}$ the probability that the measurement of A yields a particular eigenvalue a_n , is given by

$$P(a_n) = \frac{\langle \psi_{\text{before}} | \Pi_n | \psi_{\text{before}} \rangle}{\langle \psi_{\text{before}} | \psi_{\text{before}} \rangle}$$
(7.2)

where $\hat{\Pi}_n$ is the projection operator onto the eigenspace $\mathcal{E}_n \subset \mathcal{H}$ corresponding to the eigenvalue a_n .

¹¹²The state $|\psi\rangle$ is the direction of the ray.

3. If $\hat{\mathcal{A}}$ has a continuous spectrum of eigenvalues $\{a(\alpha)\}$ with α some continuous parameter, then the probability that the measurement of A yields a value lying in some interval \mathcal{I} of the spectrum is given by

$$P(A \in \mathcal{I}) = \frac{\langle \psi_{\text{before}} | \hat{\Pi}_{\mathcal{I}} | \psi_{\text{before}} \rangle}{\langle \psi_{\text{before}} | \psi_{\text{before}} \rangle}$$
(7.3)

where $\Pi_{\mathcal{I}}$ is the projection operator onto the eigenspace $\mathcal{E}_{\mathcal{I}} \subset \mathcal{H}$ corresponding to the interval \mathcal{I} , i.e.

$$\hat{\Pi}_{\mathcal{I}} = \int_{\mathcal{I}} d\alpha |\phi(\alpha)\rangle \langle \phi(\alpha)|, \qquad (7.4)$$

with $|\phi(\alpha)\rangle$ is the eigenstate corresponding to the eigenvalue $a(\alpha)$.

4. Immediately after the measurement, the system will be in the state

$$|\psi_{\text{after}} \rangle = \begin{cases} \frac{\hat{\Pi}_n |\psi_{\text{before}}\rangle}{\sqrt{\langle\psi|\psi\rangle}} & \text{if the spectrum is discrete} \\ \\ \frac{\hat{\Pi}_{\mathcal{I}} |\psi_{\text{before}}\rangle}{\sqrt{\langle\psi|\psi\rangle}} & \text{if the spectrum is continuous} \end{cases}$$

For instance, if a_n is not degenerate with the eigenstate $|\phi_n\rangle$, then $\hat{\Pi}_n = |\phi_n\rangle \langle \phi_n|$, and in this case we have

$$P(a_n) = \frac{|\langle \phi_n | \psi_{\text{before}} \rangle|^2}{\langle \psi_{\text{before}} | \psi_{\text{before}} \rangle}$$

$$|\psi_{\text{after}} \rangle = \frac{\langle \phi_n | \psi_{\text{before}} \rangle}{\sqrt{\langle \psi_{\text{before}} | \psi_{\text{before}} \rangle}} |\phi_n \rangle$$
(7.5)

If a_n is d_n -degenerate with the corresponding eigenstates $\{|\phi_n^{(k)}\rangle, k = 1, 2, ..d_n\}$, then the projection operator is $\hat{\Pi}_n = \sum_{k=1}^{d_n} |\phi_n^{(k)}\rangle \langle \phi_n^{(k)}|$. In this case

$$P(a_n) = \frac{\sum_{k=1}^{d_n} |\langle \phi_n^{(k)} | \psi_{\text{before}} \rangle|^2}{\langle \psi_{\text{before}} | \psi_{\text{before}} \rangle}$$

$$|\psi_{\text{after}} \rangle = \sum_{k=1}^{d_n} \frac{\langle \phi_n^{(k)} | \psi_{\text{before}} \rangle}{\sqrt{\langle \psi_{\text{before}} | \psi_{\text{before}} \rangle}} |\phi_n^{(k)} \rangle$$

$$(7.6)$$

In the case of a continuous spectrum, we have

$$P(A \in \mathcal{I}) = \frac{\int_{\mathcal{I}} |\langle \phi(\alpha) | \psi_{\text{before}} \rangle |^2 d\alpha}{\langle \psi_{\text{before}} | \psi_{\text{before}} \rangle}$$

$$|\psi_{\text{after}} \rangle = \int_{\mathcal{I}} \frac{\langle \phi(\alpha) | \psi_{\text{before}} \rangle}{\sqrt{\langle \psi_{\text{before}} | \psi_{\text{before}} \rangle}} |\phi(\alpha) \rangle d\alpha$$

$$(7.7)$$

From the discussion above we can say that an observable A with possible outcomes $\{a\}$ (discrete or continuous) has the **spectral decomposition**¹¹³

$$\hat{\mathcal{A}} = \sum_{a} a \,\hat{\Pi}_a \tag{7.8}$$

where a is the eigenvalue of the operator $\hat{\mathcal{A}}$ representing the observable A and Π_a is the projection onto the eigenspace corresponding to a. Thus, when a measurement of an observable A is made, the system is driven to one of the eigenstate of A. However, once the measurement is repeated the measurement we obtain the same value of observable and with same eigenstate with a 100% probability. This is known as the **collapse of the wave function**.

(d) Postulate 4: Time Evolution of a System

The time evolution of a state $|\psi\rangle$ of a system is given by the Schrödinger equation

$$i\hbar \frac{d}{dt}|\psi(t)\rangle = \hat{\mathcal{H}}(t)|\psi(t)\rangle$$
(7.9)

where $\hat{\mathcal{H}}(t)$ is the Hermitian Hamiltonian operator of the system which describes the energy of the system.

Note that the equation of the bra $\langle \psi(t) |$ is

$$-i\hbar\frac{d}{dt} < \psi(t)| = <\psi(t)|\hat{\mathcal{H}}$$
(7.10)

Then, we have

$$\frac{d}{dt}\left(\langle\psi(t)|\psi(t)\rangle\right) = \left(\frac{d}{dt}\langle\psi(t)|\right)|\psi(t)\rangle + \langle\psi(t)|\left(\frac{d}{dt}|\psi(t)\rangle\right)$$
(7.11)
$$= -\hbar \langle\psi(t)|\hat{\mathcal{H}}|\psi(t)\rangle + \hbar \langle\psi(t)|\hat{\mathcal{H}}|\psi(t)\rangle = 0$$

Thus, $\langle \psi(t) | \psi(t) \rangle$ remains unchanged during the evolution of the state of the system.

To find the state of a system $|\psi(t)\rangle$ that has evolved from some initial state $|\psi(t_0)\rangle$, we define an evolution operator $\hat{U}(t, t_0)$:

$$|\psi(t)\rangle = \hat{U}(t,t_0)|\psi(t_0)\rangle$$
 (7.12)

¹¹³In fact the spectral decomposition applies for any **normal** operator, i.e operators which satisfy $\hat{A}\hat{A}^{\dagger} = \hat{A}^{\dagger}\hat{A}$. Of course Hermitian operators are normal operators.

with $\hat{U}(t_0, t_0) = \mathbb{I}$. This allows us to write

$$\psi(t) >= \hat{U}(t, t_1) |\psi(t_1) >= \hat{U}(t, t_1) \hat{U}(t_1, t_0) |\psi(t_0) >$$
(7.13)

which implies that

$$\hat{U}(t,t_0) = \hat{U}(t,t_1)\hat{U}(t_1,t_0)$$
(7.14)

Taking $t = t_0$ in the above equation implies that

$$\hat{U}(t_0, t_1)\hat{U}(t_1, t_0) = \mathbb{I}, \quad \forall t_0 \tag{7.15}$$

Since $<\psi(t)|\psi(t)>$ is invariant during the time evolution of the system, we can write

$$<\psi(t_0)|\psi(t_0)> = <\psi(t)|\psi(t)> = <\psi(t_0)|\hat{U}^{\dagger}(t,t_0)\hat{U}(t,t_0)|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,\quad\forall|\psi(t_0)>,$$

Thus,

$$\hat{U}^{\dagger}(t,t_0)\hat{U}(t,t_0) = \mathbb{I}$$
(7.17)

So, \hat{U} is a unitary operator. Furthermore, the relation (7.15) yields

$$\hat{U}(t_0, t) = \hat{U}^{\dagger}(t, t_0)$$
(7.18)

Now let us consider the evolution of the system from some t to $(t + \delta t)$, with Δt very small. Then, according to Schrödinger equation we have

$$|\psi(t + \Delta t)\rangle = \left(1 - \frac{i}{\hbar}\hat{\mathcal{H}}\Delta t\right)|\psi(t)\rangle$$
(7.19)

For an isolated system, the Hamiltonian is time-independent, and in this case we have

$$\hat{U}(t,t_0) = \exp\left(-\frac{\hat{\mathcal{H}}(t_2 - t_1)}{\hbar}\right)$$
(7.20)

Hence, the state of the system at an instant t is

$$|\psi(t)\rangle = \exp\left(-\frac{\hat{\mathcal{H}}t)}{\hbar}\right)|\psi(0)\rangle$$
 (7.21)

We can always choose a basis formed by the eigenvectors $\{\phi_k\}$ of $\hat{\mathcal{H}}$ with eigenvalues E_k , i.e. $\hat{\mathcal{H}}|\phi_k\rangle = E_k|\phi_k\rangle$, and write

$$|\psi(t)\rangle = \sum_{k} C_k(t) |\phi_k\rangle, \quad \hat{\mathcal{H}} = \sum_{k} E_k |\phi_k\rangle \langle \phi_k|$$
(7.22)

where we take $\sum_{k} |\alpha_k|^2 = 1$ so that $|\psi(t)\rangle$ is normalized to unity. With the use of (7.22), Schrödinger equation reads

$$\frac{dC_k(t)}{dt} = -i\frac{E_k}{\hbar}C_k(t) \tag{7.23}$$

which has the solution

$$C_k(t) = exp\left(-i\frac{E_k}{\hbar}\right) C_k(0)$$
(7.24)

This can be understood if each eigenstate of $\hat{\mathcal{H}}$ undergoes a phase rotation with the argument given by its corresponding energy.

7.2 Position and Momentum representations

In the previous chapter, we have seen that any vector of a Hilbert space \mathcal{H} can be expanded as a linear combination of some basis-vectors. In particular, one can expand the vectors of in terms of the eigenstates of some Hermitian operator since the later form a complete basis. In this section, we will consider the representation of states and operators in the position basis, then in the momentum basis.

7.2.1 Position representation

In quantum mechanics , the position of a particle is associated with a Hermitian operator operator \hat{X} . Its eigenvalues consist of a set of continuous real numbers $\{x\}$, with corresponding eigenstates $\{|x>\}$:

$$\hat{X}|x\rangle = x |x\rangle \tag{7.25}$$

The set $\{|x\rangle\}$ forms an orthonormal basis, i.e.

$$\langle x|x' \rangle = \delta(x - x'), \qquad \int_{\text{all space}} |x\rangle \langle x| \ dx = \mathbb{I}$$
 (7.26)

So, if the particle is in some state $|\psi\rangle$, it can be written in the position basis as

$$|\psi\rangle = \int_{\text{all space}} |x\rangle \langle x|\psi\rangle \ dx \tag{7.27}$$

Given two states $|\phi\rangle$ and $|\psi\rangle$, their inner product is

$$|\psi\rangle = \int_{\text{all space}} \langle \phi | x \rangle \langle x | \psi \rangle \, dx \tag{7.28}$$

The expansion coefficient $< x | \psi >$ is called the wave function of the particle at position x, and it is denoted as

$$\psi(x) := \langle x | \psi \rangle \tag{7.29}$$

and with its complex conjugate given by

$$\psi^*(x) := \langle x | \psi \rangle^* = \langle \psi | x \rangle \tag{7.30}$$

The wave function can be can be interpreted, as the probability amplitude to find the particle at position x. That is the probability that a measurement of the position of the particle yields a value between x and (x + dx) is

$$dP(x) = |\psi(x)|^2 dx$$
(7.31)

which is in agreement with the interpretation suggested by Born in 1926.

Now that we have give the state representation in the position basis, what about the representation of operator?. For that consider

$$<\phi|\hat{\mathcal{A}}|\psi> = <\phi|\mathbb{I}\ \hat{\mathcal{A}}\ \mathbb{I}|\psi>$$

$$= \int_{-\infty}^{+\infty} dx' \int_{-\infty}^{\infty} dx <\phi|x'> < x'|\hat{\mathcal{A}}|x> < x|\psi>$$

$$= \int_{-\infty}^{+\infty} dx' \int_{-\infty}^{\infty} dx \phi^{*}(x')A(x',x)\psi(x)$$
(7.32)

where

$$A(x',x) = \langle x'|\hat{\mathcal{A}}|x\rangle \tag{7.33}$$

is the matrix element of the operator $\hat{\mathcal{A}}$ in the position representation.

As a particular case, consider $\hat{\mathcal{A}}$ to be a function only of the position operator, i.e. $\hat{\mathcal{A}} = \mathcal{A}(\hat{X})$. Then

$$A(x',x) = \langle x' | \mathcal{A}(\hat{X}) | x \rangle = \mathcal{A}(x)\delta(x-x')$$
(7.34)

Substituting into Eq (7.32) and using the property of the Dirac distribution, yields

$$\langle \phi | \mathcal{A}(\hat{X}) | \psi \rangle = \int_{-\infty}^{+\infty} dx \ \phi^*(x) \ \mathcal{A}(x) \ \psi(x)$$
 (7.35)

We define an operator $\hat{\mathcal{T}}(\alpha)$ that translates the system a distance α as

$$\hat{\mathcal{T}}(\alpha) | x \rangle := | x + \alpha \rangle \tag{7.36}$$

Note that for arbitrary α , we have

$$\hat{\mathcal{T}}(\alpha)\hat{\mathcal{T}}(-\alpha) |x\rangle = \hat{\mathcal{T}}(\alpha) |x-\alpha\rangle = |x\rangle$$
(7.37)

which shows that $\hat{\mathcal{T}}(\alpha)$ is unitary operator, with $\hat{\mathcal{T}}^{\dagger}(\alpha) = \hat{\mathcal{T}}(-\alpha)$. This implies that

$$\langle x \mid \hat{\mathcal{T}}(\alpha) = \langle x - \alpha \mid$$

$$\tag{7.38}$$

In general, acting on a state $|\psi\rangle$ in the Hilbert space by $\hat{\mathcal{T}}(\alpha)$ gives a new state $|\psi'\rangle$:

$$\hat{\mathcal{T}}(\alpha) |\psi\rangle := |\psi'\rangle \tag{7.39}$$

Writing $|\psi\rangle$ and $|\psi'\rangle$ in the position representation yields

$$\int dy \psi'(y) |y\rangle = \int dy' \psi(y') U(\alpha) |xy\rangle$$

$$= \int dy' \psi(y') |y' + \alpha\rangle = \int dy \psi(y - \alpha) |y\rangle$$
(7.40)

Multiplying both sides of the above equation by the bra $\langle x|$, gives¹¹⁴

$$\psi'(x) = \psi(x - \alpha) \tag{7.41}$$

Since the translation operator is unitary, it can be written as 115

$$\hat{\mathcal{T}}(\alpha) = e^{-iK\alpha} \tag{7.42}$$

with \hat{K} being a Hermitian operator. In this way of writing the translation operator, \hat{K} is called the "generator" of the translation operator. If we denote by [k >the eigenstate of \hat{K} , then

$$\hat{\mathcal{T}}(\alpha)|k\rangle = e^{-i\hat{k}\alpha}|k\rangle \tag{7.43}$$

Now, the components of $\hat{\mathcal{T}}(\alpha)|k\rangle$ in the position basis read

$$\langle y|\hat{\mathcal{T}}(\alpha)|k\rangle = e^{-i\hat{k}\alpha} \langle y|k\rangle$$

= $e^{-i\hat{k}\alpha} \psi_k(y)$ (7.44)

where we defined $\langle y|k \rangle = \psi_k(y)$. Using the fact that $\langle y|\hat{\mathcal{T}}(\alpha) = \langle y-\alpha|$, the above equation yields

$$\psi_k(y - \alpha) = e^{-i\hat{k}\alpha} \ \psi_k(y) \tag{7.45}$$

¹¹⁴This result could also be obtained by simply using (7.38) as

$$\psi'(x) = \langle x | \hat{\mathcal{T}}(\alpha) | \psi \rangle$$

= $\langle x - a | \hat{\mathcal{T}}(\alpha) | \psi \rangle$
= $\psi(x - \alpha)$

 $^{115}\mathrm{This}$ can be generalized to the case of three dimension as

$$\hat{\mathcal{T}}(\vec{\alpha}) = e^{-i\vec{K}.\vec{\alpha}}$$

Setting y = 0 and and re-naming α as x, we get

$$\psi_k(x) = e^{-ikx} \ \psi_k(0) \tag{7.46}$$

Hence, $\psi_k(x)$ represents a plane wave with a wave number k, which according to de Broglie it is related to the momentum \hat{p} of the associated particle by $\vec{p} = \hbar \vec{p}$. This implies that

$$\hat{P} = \hbar \hat{K} \tag{7.47}$$

or, equivalently,

$$-\frac{i}{\hbar}\hat{P} = \frac{\partial}{\partial\alpha}\hat{\mathcal{T}}(\alpha)|_{\alpha=0}$$
(7.48)

With the above expression of \hat{P} , we can write

$$\langle x'|\hat{P}|x\rangle = i\hbar \frac{\partial}{\partial \alpha} \langle x'|\hat{\mathcal{T}}(\alpha)|x\rangle|_{\alpha=0}$$

$$= i\hbar \frac{\partial}{\partial \alpha} \delta(x+a-x')|_{\alpha=0}$$

$$(7.49)$$

which is equivalent to

$$\langle x'|\hat{P}|x\rangle = i\hbar \,\,\delta'(x-x') \tag{7.50}$$

Here, for ease of notation, $\delta'(x)$ denotes the derivative of $\delta(x)$ with respect to x. To find the momentum representation of \hat{P} , we consider the quantity $\langle x|\hat{P}|\psi\rangle$, with $|\psi\rangle$ some arbitrary state. We have

$$\langle x|\hat{P}|\psi\rangle = \int \langle x|\hat{P}|x'\rangle \langle x'|\psi\rangle dx'$$

$$= -i\hbar \int_{-\infty}^{+\infty} \delta'(x-x') \ \psi(x')dx'$$

$$(7.51)$$

Using the property that¹¹⁶ $\int_{-\infty}^{+\infty} f(x)\delta'(x-x_0)dx = -f'(x_0)$, we obtain $\boxed{\langle x|\hat{P}|\psi \rangle = -i\hbar\frac{\partial}{\partial x}\psi(x)}$

$$\langle x|\hat{P}|\psi\rangle = -i\hbar \frac{\partial}{\partial x}\psi(x)$$
(7.52)

¹¹⁶This can be shown as follows:

$$\int_{-\infty}^{+\infty} f(x)\delta'(x-x_0)dx = -\int_{-\infty}^{+\infty} f'(x)\delta(x-x_0)dx + f(x)\delta(x-x_0)|_{-\infty}^{+\infty}$$

where we used integration by parts. Now, since $\delta(-\infty) = \delta(+\infty) = 0$, and assuming that f(x) is finite, the second term vanishes. Hence, we get

$$\int_{-\infty}^{+\infty} f(x)\delta'(x-x_0)dx = -f'(x_0).$$

So, in the position basis¹¹⁷, the momentum operator can be represented by $(-i\hbar\partial/\partial x)$.

7.2.2 Momentum representation

Since the momentum operator is Hermitian, its eigenstates form a complete set of basis vector |p>, with

$$\hat{P} \mid p \ge p \mid p \ge \tag{7.55}$$

with

$$\langle p|p' \rangle = \delta(p-p'), \qquad \int_{\text{all momenta}} |p\rangle \langle p| \ dp = \mathbb{I}$$
 (7.56)

Multiplying both sides of (7.55) by the bra $\langle x|$, yields the differential equation

$$-i\frac{\partial}{\partial x} < x|p\rangle = p < x|p\rangle \tag{7.57}$$

which has as a solution $\langle x|p \rangle = c(p)e^{i\frac{p}{h}x}$, where c (p) is an arbitrary constant. We will see that c is fixed by the normalization of the eigenstates $|p\rangle$. For that let us calculate the inner product between two moment eigenstates $|p\rangle$ and $|p'\rangle$

$$< p'|p> = \int dx < p'|x> < x|p> = c^{2}(p) \int e^{i\frac{(p'-p)}{\hbar}x} dx$$
 (7.58)
= $2\pi c(p)^{2} \delta(p'-p)$

where we used the fact that $\int dx e^{ikx} = 2\pi\delta(x)$. So to normalize the inner product, we choose $c = 1/\sqrt{2\pi}$, i.e

$$\langle x|p\rangle = \frac{1}{\sqrt{2\pi\hbar}} e^{i\frac{p}{\hbar}x} \tag{7.59}$$

with

$$\langle p'|p\rangle = \delta(p'-p) \tag{7.60}$$

¹¹⁷In this representation, the commutator $[\hat{X}, \hat{P}]$ can be calculated as follows:

$$< x | [\hat{X}, \hat{P}] | \psi > = < x | \hat{X} \hat{P} | \psi > - < x | \hat{P} \hat{X} | \psi >$$

$$= x \left(-i\hbar \frac{\partial}{\partial x} \right) < x | \psi > - \left(-i\hbar \frac{\partial}{\partial x} \right) (x < x | \psi >)$$

$$= i < x | \psi >$$
(7.53)

Since this equality holds for any state $|\psi\rangle$, we conclude that

$$\left[\hat{X},\hat{P}\right] = i\hbar \tag{7.54}$$

which is the Born-Heisenberg quantization rule. It is also important to note that the canonical quantization rule $\left[\hat{X}, \hat{P}\right] = i\hbar$ holds for any representation chosen.

If the system is some state $|\psi\rangle$, the wave function in the $|p\rangle$ representation is

$$\tilde{\psi}(p) = \langle p | \psi \rangle \tag{7.61}$$

which can be interpreted as amplitude for the particle to have momentum p. So, for two arbitrary states $|\psi\rangle$ and $|\phi\rangle$, the inner product can be written in momentum representation as

$$\langle \phi | \psi \rangle = \int dp \langle \phi | p \rangle \langle p | \psi \rangle = \int dp \tilde{\phi}^*(p) \tilde{\psi}(p)$$
 (7.62)

In particular, if $|\psi\rangle$ is a normalized state, then

$$\int dp |\tilde{\psi}|^2(p) = 1 \tag{7.63}$$

Moreover, the probability that the measurement of the momentum of the system yields a value between p_1 and p_2 is given by

$$P\{p \in [p_1, p_2]\} = \int_{p_1}^{p_2} dp |\tilde{\psi}|^2(p)$$
(7.64)

The momentum space wave function $\tilde{\psi}(p)$ can be re-written as

$$\tilde{\psi}(p) = \int \langle p|x \rangle \langle x|\psi \rangle \ dx = \frac{1}{\sqrt{2\pi\hbar}} e^{-i\frac{p}{\hbar}x} \ \psi(x) \ dx \tag{7.65}$$

On the other hand, the wave function $\psi(x)$ can expressed as

$$\psi(x) = \int \langle x|p \rangle \langle p|\psi \rangle \ dx = \frac{1}{\sqrt{2\pi\hbar}} e^{i\frac{p}{\hbar}x} \tilde{\psi}(p) \ dp \tag{7.66}$$

Thus, by comparing the expressions (7.65) and (7.66), we see that $\tilde{\psi}(p)$ and $\psi(x)$ are Fourier transforms of each other.

Just as we found the representation of \hat{P} in the $|x\rangle$ basis, we can derive the representation of the position operator \hat{X} in the momentum basis. For that, let $|\psi\rangle$ an arbitrary state and consider

$$< p|\hat{X}|\psi > = \int < p|x' > < x'|\hat{X}|\psi > dx'$$

$$= \frac{1}{\sqrt{2\pi\hbar}} \int x' e^{-i\frac{px'}{\hbar}} \psi(x') dx'$$

$$= (i\hbar) \frac{1}{\sqrt{2\pi\hbar}} \frac{\partial}{\partial p} \left[\int e^{-i\frac{px'}{\hbar}} \psi(x') dx' \right]$$
(7.67)

or, equivalently,

$$|\langle x|\hat{P}|\psi\rangle = i\hbar\frac{\partial}{\partial p}\psi(\tilde{p})$$
(7.68)

Hence, in the momentum basis, the position operator can be represented by $-i\hbar\partial/\partial p$.

For the case of three dimensional space, the expressions (7.26), (7.52), (7.59), (7.60), and (7.68) become

$$<\vec{r'}|\vec{r}> = \delta^{(3)}(\vec{r'}-\vec{r})$$
 (7.69)

$$\int_{\text{all space}} |\vec{r}\rangle \langle \vec{r}| \ d^3 \vec{r} = \mathbb{I}$$
(7.70)

$$\langle \vec{r} | \hat{\vec{P}} | \psi \rangle = -i\hbar \vec{\nabla} \psi(x)$$
 (7.71)

$$<\vec{r}|\vec{p}> = \frac{1}{(2\pi\hbar)^{3/2}} e^{i\frac{\vec{p}}{\hbar}.\vec{r}}$$
(7.72)

$$< \vec{p} | \vec{r} > = \frac{1}{\left(2\pi\hbar\right)^{3/2}} e^{-i\frac{\vec{p}}{\hbar}.\vec{r}}$$
(7.73)

$$<\vec{p'}|\vec{p}> = \delta^{(3)}(\vec{p'}-\vec{p})$$
 (7.74)

$$\int_{\text{all momenta}} |\vec{p}\rangle \langle \vec{p}| \ d^3 \vec{p} = \mathbb{I}$$
(7.75)

$$\langle \vec{r} | \hat{\vec{P}} | \psi \rangle = i\hbar \vec{\nabla}_{\vec{p}} \; \tilde{\psi}(\vec{p})$$
 (7.76)

where $\vec{\nabla} = (\partial/\partial x, \partial/\partial y, \partial/\partial z)$ and $\vec{\nabla}_{\vec{p}} = (\partial/\partial p_x, \partial/\partial p_y, \partial/\partial p_z)$ are the gradient operator in the the position and momentum space, respectively. Here the states labeled by a vector quantity, such as $|\vec{r}\rangle$ is the tensor product of $|x\rangle$, $|y\rangle$, and $|z\rangle$. For instance,

$$\int |\vec{r}| < \vec{r}| \ d^{3}\vec{r} := \int dx dy dz \left(|x > \otimes|y > \otimes|z >\right) \left(< x| \otimes < y| \otimes < z||\right)$$
$$= \int dx dy dz (|x > < x|) \otimes \left(|y > < y|\right) \otimes \left(|z > < z|\right) = \mathbb{I}_{x} \mathbb{I}_{y} \mathbb{I}_{F}$$
For momentum space wave function in three dimension, we have

$$<\vec{p}|\psi> = \int <\vec{p}|\vec{r}> <\vec{r}|\psi> d^{3}\vec{r}$$

$$= \int <\vec{r}|\psi> d^{3}\vec{r}$$

$$= \frac{1}{(2\pi\hbar)^{3/2}} \int e^{-i\frac{\vec{p}}{\hbar}\cdot\vec{r}} \psi(\vec{r}) d^{3}\vec{r}$$
(7.77)

7.3 Compatible Operators and Simultaneous Measurement

In classical physics, one can make simultaneous measurement of two observables, say position and momentum, or the x and the y components of the orbital angular momentum. however, as we saw in the third postulate above, right after the measurement the wave function collapse. So, in quantum theory it matters which measurement is made first. In other words, the first measurement might destroy some of the information regarding the second observable. However, there is a situation when the order of the measurements does not matter. That is when the state of the system is a common an eigenstate to the operators associated with observables to be measured.

Suppose that $\{|\phi_i\rangle, i = 1, 2, ..\}$ are eigenstates of both Hermitian operators $\hat{\mathcal{A}}$ and $\hat{\mathcal{B}}$, which represent the observables A and B. Let a_i and b_i are the eigenvalues of $\hat{\mathcal{A}}$ and $\hat{\mathcal{B}}$, respectively, i.e.

$$\hat{\mathcal{A}}|\phi_i\rangle = a_i|\phi_i\rangle, \quad \hat{\mathcal{B}}|\phi_i\rangle = b_i|\phi_i\rangle$$

$$(7.78)$$

This implies that

$$\left[\hat{\mathcal{A}}, \ \hat{\mathcal{B}}\right] |\phi_i\rangle = (b_i a_i - a_i b_i) |\phi_i\rangle = 0, \ \forall |\phi_i\rangle$$
(7.79)

Since the eigenstate of Hermitian operators form a basis of a Hilbert space \mathcal{H} , then any state in \mathcal{H} can be written as a linear combination of $\{|\phi_i\rangle, i = 1, 2, ..\}$. Thus,

$$\left[\hat{\mathcal{A}}, \ \hat{\mathcal{B}}\right] |\psi_i \rangle = 0, \ \forall |\psi_\rangle \in \mathcal{H} \Longrightarrow \left[\hat{\mathcal{A}}, \ \hat{\mathcal{B}}\right] = 0$$
(7.80)

If there exist a complete set of linearly independent states that are eigenstates of, two or more Hermitian operators, then their corresponding observables are said to be **compatible**.

Now, suppose that $\hat{\mathcal{A}}$ and $\hat{\mathcal{B}}$ are two commuting Hermitian operators and the question is: "Do they share a complete set of linearly independent states?. The answer turns out to be "yes" they do. Here I will give a proof for the special case where at least one of the operators, say $\hat{\mathcal{A}}$, is non-degenerate, i.e. it has a set of eigenstates $\{|a_i\rangle\}$ with distinct eigenvalues a_i . Acting with the operator product $\hat{\mathcal{B}}\hat{\mathcal{A}}$ on $|a_i\rangle$ gives

$$\hat{\mathcal{B}}\hat{\mathcal{A}}|a_i\rangle = a_i\hat{\mathcal{B}}|a_i\rangle \tag{7.81}$$

The action of $\hat{\mathcal{A}}\hat{\mathcal{B}}$ on $|a_i\rangle$ gives

$$\hat{\mathcal{A}}\hat{\mathcal{B}}|a_i\rangle = a_i\hat{\mathcal{B}}|a_i\rangle \tag{7.82}$$

where we have used the assumption that $\hat{\mathcal{A}}$ and $\hat{\mathcal{B}}$ commute. The above equation means that $\mathcal{B}|a_i >$ is also an eigenstate of $\hat{\mathcal{A}}$ with eigenvalue a_i . But since $\hat{\mathcal{A}}$ is nondegenerate, $\mathcal{B}|a_i >$ must be related to the eigenstate $|a_i >$ by a multiplicative complex number, i.e.

$$\mathcal{B}|a_i\rangle = C|a_i\rangle, \quad C \in \mathbb{C} \tag{7.83}$$

However, this equation shows that $|a_i\rangle$ is an eigenstate of \mathcal{B} with eigenvalue C. Because \mathcal{B} is Hermitian, the constant C is real. Since every eigenstate of \mathcal{A} is an eigenstate of \mathcal{B} , this means that the set $\{|a_i\rangle\}$ forms a complete basis for both of them¹¹⁸. Now that $|a_i\rangle$ is an eigenstate of \mathcal{B} , but with eigenvalue b_i , we can label the common (i.e. simultaneous) eigenstates of the two operators by $|a_i, b_i\rangle$.

7.4 Heisenberg Uncertainty Relation

Let $\hat{\mathcal{A}}$ and $\hat{\mathcal{B}}$ be two Hermitian operators and let $|\psi\rangle$ be a normalized state. We would like to calculate the product of the variances of the observables A and B in the state $|\psi\rangle$. For that let us define the Hermitian operators

$$\Delta \hat{\mathcal{A}} := \hat{\mathcal{A}} - \langle A \rangle$$

$$\Delta \hat{\mathcal{B}} := \hat{\mathcal{B}} - \langle B \rangle$$
(7.84)

Operating $|\psi\rangle$ with the uncertainty operators $\Delta \hat{\mathcal{A}}$ and $\Delta \hat{\mathcal{B}}$, yield new states

$$|\phi_A\rangle = \Delta \hat{\mathcal{A}} |\psi\rangle, \quad |\phi_B\rangle = \Delta \hat{\mathcal{B}} |\psi\rangle$$
(7.85)

with

$$\langle \phi_A | \phi_A \rangle = \langle \psi | (\hat{\mathcal{A}} - \langle A \rangle)^2 | \psi \rangle = (\Delta A)^2$$

$$\langle \phi_B | \phi_B \rangle = \langle \psi | (\hat{\mathcal{B}} - \langle B \rangle)^2 | \psi \rangle = (\Delta B)^2$$

(7.86)

Using Schwartz inequality given in (6.4), we have

$$(\Delta A)^2 (\Delta B)^2 \ge |<\Delta \hat{\mathcal{A}} \Delta \hat{\mathcal{B}}>|^2 \tag{7.87}$$

Next, we rewrite the product $\Delta \hat{\mathcal{A}} \Delta \hat{\mathcal{B}}$ as a sum of of a commutator and anticommutator:

$$\Delta \hat{\mathcal{A}} \Delta \hat{\mathcal{B}} = \frac{1}{2} \left[\Delta \hat{\mathcal{A}}, \Delta \hat{\mathcal{B}} \right] + \frac{1}{2} \{ \Delta \hat{\mathcal{A}}, \Delta \hat{\mathcal{B}} \}$$
(7.88)

¹¹⁸For this special case where $\hat{\mathcal{A}}$ is non-degenerate, there is only one such common basis.

Howerver,

$$\left(\left[\Delta\hat{\mathcal{A}},\Delta\hat{\mathcal{B}}\right]\right)^{\dagger} = -\left[\Delta\hat{\mathcal{A}},\Delta\hat{\mathcal{B}}\right], \quad \left(\left\{\Delta\hat{\mathcal{A}},\Delta\hat{\mathcal{B}}\right\}\right)^{\dagger} = \left(\left\{\Delta\hat{\mathcal{A}},\Delta\hat{\mathcal{B}}\right\}\right)$$
(7.89)

As a result, the expectation value of the commutator is an imaginary number, whereas the one of the anti-commutator is a real. Thus, we have

$$(\Delta A)^2 (\Delta B)^2 \ge \frac{1}{4} | < \left[\Delta \hat{\mathcal{A}}, \Delta \hat{\mathcal{B}} \right] > |^2 + \frac{1}{4} | < \{ \Delta \hat{\mathcal{A}}, \Delta \hat{\mathcal{B}} \} > |^2$$
(7.90)

Since the last term is greater or equal to zero, we conclude that

$$(\Delta A)^2 (\Delta B)^2 \ge \frac{1}{4} | < \left[\Delta \hat{\mathcal{A}}, \Delta \hat{\mathcal{B}} \right] > |^2$$
(7.91)

Applying the above inequality to the operators \hat{X} and \hat{P} , and using the commutation relation $[\hat{X}, \hat{P}] = i\hbar$, we find

$$\Delta x \Delta p \ge \frac{\hbar}{2} \tag{7.92}$$

which is the celebrated **Heisenberg uncertainty relation**. What this inequality is saying is that if the position wave function $\psi(x)$, that is the probability amplitude for the particle to be found at position x, has a narrow distribution, then the probability amplitude $\tilde{\psi}(p)$ for the momentum will have a spread out distribution and vice versa.

7.5 Schrodinger, Heisenberg and Interaction Picture *

The first and the third postulates of quantum mechanics assumes that the state of the system depends on time and it's evolution is given by Schrödinger equation. However, this is not the only way to describe the evolution of the quantum system. This is can be seen from the fact that observables are expectations values of the corresponding Hermitian operators in some state of the system. So, if the system is in some state

 $|\psi(t)\rangle$, then the expectation value of an observable A is¹¹⁹

$$< A > = <\psi(t)|\hat{\mathcal{A}}|\psi(t)> = \operatorname{Tr}\left[|\psi(t)> <\psi(t)|\hat{\mathcal{A}}\right]$$

$$= \operatorname{Tr}\left[\hat{\mathcal{U}}(t)|\psi(0)> <\psi(0)|\hat{\mathcal{U}}^{\dagger}(t)\hat{\mathcal{A}}\right]$$

$$= \operatorname{Tr}\left[|\psi(0)> <\psi(0)|\hat{\mathcal{U}}^{\dagger}(t)\hat{\mathcal{A}}\hat{\mathcal{U}}(t)\right]$$
(7.93)

where in the third equality we used the cyclic property of the trace.

$$\operatorname{Tr}\left[|\psi(t)\rangle \langle \psi(t)|\hat{\mathcal{A}}\right] = \sum_{i,j,k} C_j(t) C_k^*(t) \langle \alpha_i | \phi_j \rangle \langle \phi_k | \hat{\mathcal{A}} | \alpha_i \rangle$$
$$= \sum_{i,j,k} C_j(t) C_k^*(t) a_k \langle \alpha_i | \phi_j \rangle \langle \phi_k | \alpha_i \rangle = \langle \psi(t) | \hat{\mathcal{A}} | \psi(t) \rangle$$

¹¹⁹The expression of $\langle A \rangle$ as a trace can be shown as follows. In some basis $\{\alpha_i\}$, the trace of the operator $\left[|\psi(t)\rangle\langle\psi(t)|\hat{\mathcal{A}}\right]$ reads $\operatorname{Tr}\left[|\psi(t)\rangle\langle\psi(t)|\hat{\mathcal{A}}\right] = \sum_i \langle \alpha_i|\psi(t)\rangle\langle\psi(t)|\hat{\mathcal{A}}|\alpha_i\rangle$. Since $\hat{\mathcal{A}}$ is hermitian operator, its eigenstates $\{|\phi_j\rangle\}$ form an orthonormal basis, and we can write $|\psi(t)\rangle = \sum_j C_j(t)|\phi_j\rangle$. So,

8 Harmonic Oscillator Using Ladder Operators

8.1 Raising and Lowering Operators

As we discussed in section 3, a particle of mass m oscillating with a frequency ω can be described by the Hamiltonian of a simple harmonic oscillator given by

$$\hat{\mathcal{H}} = \frac{\hat{P}^2}{2m} + \frac{1}{2}m\omega^2 \hat{X}^2$$
(8.1)

It can be re-written as follows

$$\hat{\mathcal{H}} = \hbar\omega \left[\frac{1}{2} \left(\sqrt{\frac{m\omega}{\hbar}} \hat{X} \right)^2 + \frac{1}{2} \left(\sqrt{\frac{1}{m\omega\hbar}} \hat{P} \right)^2 \right]$$

$$= \hbar\omega \left[\frac{\left(\sqrt{\frac{m\omega}{\hbar}} \hat{X} - i\sqrt{\frac{1}{m\omega\hbar}} \hat{P} \right)}{\sqrt{2}} \frac{\left(\sqrt{\frac{m\omega}{\hbar}} \hat{X} + i\sqrt{\frac{1}{m\omega\hbar}} \hat{P} \right)}{\sqrt{2}} - \frac{i}{2\hbar} [\hat{X}, \hat{P}] \right]$$
(8.2)

This suggests that we define the following operators

$$\hat{a} = \frac{1}{\sqrt{2}} \left(\sqrt{\frac{m\omega}{\hbar}} \hat{X} + i\sqrt{\frac{1}{m\omega\hbar}} \hat{P} \right)$$

$$\hat{a}^{\dagger} = \frac{1}{\sqrt{2}} \left(\sqrt{\frac{m\omega}{\hbar}} \hat{X} - i\sqrt{\frac{1}{m\omega\hbar}} \hat{P} \right)$$
(8.3)

so that the Hamiltonian reads

$$\mathcal{H} = \hbar\omega \left(\hat{a}^{\dagger} \hat{a} + \frac{1}{2} \right) \tag{8.4}$$

Using the expressions of \hat{a} and \hat{a}^{\dagger} in (8.3), we have

$$\left[\hat{a}, \hat{a}^{\dagger}\right] = 1 \tag{8.5}$$

Furthermore, they satisfy the following commutation relations with the hamiltonian,

$$\begin{bmatrix} \hat{\mathcal{H}}, \hat{a} \end{bmatrix} = -\hbar\omega$$

$$\begin{bmatrix} \hat{\mathcal{H}}, \hat{a}^{\dagger} \end{bmatrix} = +\hbar\omega$$
(8.6)

We denote the eigenstate of the \mathcal{H} by $|n\rangle$ with its corresponding energy E_n , i.e.¹²⁰

$$\mathcal{H}|n\rangle = E_n|n\rangle \tag{8.7}$$

¹²⁰The reason for denoting the eigenstate of the Hamiltonian by just one label representing its energy is because there is no other combination of position and momentum operators that commute with \mathcal{H} except functions of \mathcal{H} .

Applying \mathcal{H} on the states $\hat{a}|n > \text{and } \hat{a}^{\dagger}|n > \text{yields}$

$$\hat{\mathcal{H}}\hat{a}|n\rangle = \left(\hat{a}H + [\hat{\mathcal{H}}, \hat{a}]\right)|n\rangle = (E_n - \hbar\omega)\,\hat{a}|n\rangle$$

$$\hat{\mathcal{H}}\hat{a}^{\dagger}|n\rangle = \left(\hat{a}^{\dagger}\hat{\mathcal{H}} + [\hat{\mathcal{H}}, \hat{a}^{\dagger}]\right)|n\rangle = (E_n + \hbar\omega)\,\hat{a}^{\dagger}|n\rangle$$
(8.8)

Thus the operators \hat{a} and \hat{a}^{\dagger} lower and raise the energy of a state by $\hbar\omega$, respectively. Hence we will refer to them as the **lowering** and **raising** operators or sometime people called them the **ladder** operators.

8.2 Energy Spectrum and Eigenstates

For an arbitrary normalized state $|\psi\rangle$, the expectation of \mathcal{H} is

$$<\psi|\hat{\mathcal{H}}|\psi> = \frac{1}{2}\hbar\omega + \hbar\omega <\psi|\hat{a}^{\dagger}\hat{a}|\psi>$$

$$= \frac{1}{2}\hbar\omega + \hbar\omega||a\psi||^{2} \ge \frac{1}{2}\hbar\omega$$
(8.9)

Thus all the eigenvalues of H are greater than or equal to $\hbar\omega/2$, i.e.

$$E_n \ge \frac{1}{2}\hbar\omega, \quad \forall n.$$
 (8.10)

Since the operator \hat{a} lowers the energy of $|n\rangle$ by $\hbar\omega$, it means that $a^k|n\rangle$ is a state with energy $(E_n - k\hbar\omega)$. However, as we noted above, the energy of a state can not be lower than $\hbar\omega/2$. Hence, for each n, there exists an integer k_n such that $\hat{a}^{k_n}|n\rangle = 0$. This means that the spectrum of $\hat{\mathcal{H}}$ must contain a state which is annihilated by \hat{a} . Such a state is called the ground state of the system. and it is denoted by $|0\rangle$. So, we have

$$\hat{a}|0\rangle = 0 \tag{8.11}$$

Putting $|\psi\rangle = |0\rangle$ in (8.9) and using the fact that $\hat{\mathcal{H}}|0\rangle = E_0|0\rangle$, yields

$$E_0 = \frac{\hbar\omega}{2} \tag{8.12}$$

The higher energy states can be obtained by the multiple application of the raising operator on the ground state, i. e.

$$|n\rangle = \mathcal{N}_n(\hat{a}^{\dagger})^n |0\rangle \tag{8.13}$$

where \mathcal{N}_n is a normalization constant. By acting with $\hat{\mathcal{H}}$ on both side of the above equation, we find that

$$E_n = \hbar \omega \left(n + \frac{1}{2} \right), \quad n = 0, 1, 2, \dots$$
(8.14)

Since $\hat{\mathcal{H}}$ is Hermitian and its energy eigenvalues $\{E_n, n = 0, 1, ..\}$ are non-degenerate, their corresponding eigenstates are mutually orthogonal. They also be normalized by requiring that

$$1 = |\mathcal{N}_{n}|^{2} < 0|\hat{a}^{n}(\hat{a}^{\dagger})^{n}| >$$

$$= |\mathcal{N}_{n}|^{2} < 0|\hat{a}^{(n-1)}(\hat{a}\hat{a}^{\dagger})(\hat{a}^{\dagger})^{(n-1)}| 0 >$$

$$= \frac{|\mathcal{N}_{n}|^{2}}{|\mathcal{N}_{n-1}|^{2}} < n - 1|\hat{a}\hat{a}^{\dagger}|n - 1 >$$
(8.15)

Using the commutation relation between \hat{a} and \hat{a}^{\dagger} given in Eq (8.5), we can write

$$1 = \frac{|\mathcal{N}_n|^2}{|\mathcal{N}_{n-1}|^2} < n - 1 | \left(\frac{\hat{\mathcal{H}}}{\hbar\omega} + \frac{1}{2}\right) | n - 1 >$$
(8.16)

from which it follows that

$$|\mathcal{N}_n| = \frac{1}{\sqrt{n}} |\mathcal{N}_{n-1}| = \frac{1}{\sqrt{n(n-1)}} |\mathcal{N}_{n-2}| = \dots = \frac{1}{\sqrt{n(n-1)\dots 1}} |\mathcal{N}_0|$$
(8.17)

Since the ground state $|0\rangle$ was assumed to be normalized, i.e. $\mathcal{N}_0 = 1$, then we get¹²¹

$$|n\rangle = \frac{1}{\sqrt{n!}} (\hat{a}^{\dagger})^{n} |0\rangle$$
 (8.18)

8.3 The Wave Function of the Harmonic Oscillator

First, we will determine the wave function of the ground state. For that we use Eq (8.11), which in the position basis $\{|x >\}$ reads

$$\langle x|\hat{a}|0\rangle = \frac{1}{\sqrt{2m\hbar\omega}} \langle x|\left(m\omega\hat{X} + i\hat{P}\right)|0\rangle = 0$$
(8.19)

In the position basis $\{|x\rangle\}$, the position and the momentum operators of a system in some state $|\psi\rangle$ are given by

$$\langle x|\hat{X}|\psi\rangle = x\psi(x), \qquad \langle x|\hat{P}|\psi\rangle = -i\hbar\frac{\partial}{\partial x}\psi(x)$$
(8.20)

¹²¹Actually, the most general solution to (8.17) is of the form

$$\frac{1}{\sqrt{n!}}e^{i\theta}$$

where θ is an arbitrary phase. However, two states that differ by an multiplicative complex number, in particular a phase belong to the same ray, and hence yield the same physics. thus, one can choose the eigenstates of $\hat{\mathcal{H}}$ whose relative are all zero.

where $\psi(x) = \langle x | \psi \rangle$ is the wave function of the system. Thus, the wave function ground state $\psi_0(x)$ is obeys of the differential equation

$$\left(\frac{\partial}{\partial x} + \frac{m\omega}{\hbar}x\right)\psi_0(x) = 0 \tag{8.21}$$

which has the solution $\psi_0(x) \propto \exp\{-\frac{m\omega}{2\hbar}x^2\}$. Normalizing $\psi_0(x)$ gives

$$\psi_0(x) = \left(\frac{m\omega}{\pi\hbar}\right)^{1/4} \exp\left(-\frac{m\omega}{2\hbar}x^2\right)$$
(8.22)

and so the wave the wave function of the ground state is Gaussian.

In the $|x\rangle$ representation, the raising and lowering operators take the following simple differential form

$$\hat{a} = \frac{1}{\sqrt{2}} \left(z + \frac{\partial}{\partial z} \right), \quad \hat{a}^{\dagger} = \frac{1}{\sqrt{2}} \left(z - \frac{\partial}{\partial z} \right)$$

$$(8.23)$$

where we defined a dimensionless coordinate $z = x\sqrt{m\omega/\hbar}$. So, the wave function of the energy states $|n\rangle$, is given by

$$\psi_n(x) = \frac{1}{\sqrt{n!}} \langle x | (\hat{a}^{\dagger})^n | 0 \rangle = \frac{1}{\sqrt{n! \, 2^n}} \left(z - \frac{\partial}{\partial z} \right)^n \, \psi_0(z) \tag{8.24}$$

This also provides an alternative definition of the Hermite polynomials¹²²

$$\left(z - \frac{\partial}{\partial z}\right)^n e^{-z^2/2} = H_n(z)e^{-z^2/2}$$
(8.25)

 122 Using the definitions of the ladder operators we can derive a recurrence relation between these polynomial as follows. First, we use the fact that

$$< z|(\hat{a}^{\dagger})^{n+1}|0> = \frac{1}{\sqrt{n+1}} < z|\hat{a}^{\dagger}|n>$$

or, equivalently,

$$\psi_{n+1}(z) = \frac{1}{2\sqrt{n+1}} \left(z - \frac{\partial}{\partial z} \right) \psi_n(z)$$

This gives

$$H_{n+1}(z) = 2zH_n(z) - \frac{dH_n(z)}{dz}$$

On the other hand, we have

$$\psi_{n-1}(z) = \frac{1}{\sqrt{n}} \langle z | \hat{a} | n \rangle = \frac{1}{2\sqrt{n}} \left(z + \frac{\partial}{\partial z} \right) \psi_n(z)$$

which gives

$$\frac{dH_n(z)}{dz} = 2nH_{n-1}(z)$$

so that

$$\psi_n(x) = \frac{1}{\sqrt{n! \, 2^n}} \frac{1}{\pi^{1/4}} H_n(z) e^{-z^2/2} \tag{8.26}$$

Note that the harmonic wave functions are also eigenstates of the parity operator \hat{P}

$$\hat{P}\psi_n(y) := \psi_n(-y) = (-1)^n \psi_n(y)$$
(8.27)

This should not be surprising since the parity operator commute with the Hamiltonian of the simple harmonic oscillator.

8.4 The Matrix Representation of the Harmonic Oscillator

Applying the annihilation and creation operators on the states $\{|n\rangle\}$ given in (8.18), we find that

$$\hat{a}|n\rangle = \sqrt{n} |n-1\rangle, \quad \hat{a}^{\dagger}|n\rangle = \sqrt{n+1} |n+1\rangle$$
(8.28)

Thus, in the $|n\rangle$ basis, \hat{a} and \hat{a}^{\dagger} have the following matrix representations:

$$[\hat{a}^{\dagger}] = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & \cdots & \cdots & 0 \\ 1 & 0 & 0 & 0 & 0 & \cdots & \cdots & 0 \\ 0 & \sqrt{2} & 0 & 0 & 0 & \cdots & \cdots & 0 \\ 0 & 0 & \sqrt{3} & 0 & 0 & \cdots & \cdots & 0 \\ 0 & 0 & 0 & \sqrt{4} & 0 & \cdots & \cdots & 0 \\ \vdots & \vdots \end{pmatrix}, \quad [\hat{a}] = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & \cdots & \cdots & 0 \\ 0 & 0 & \sqrt{2} & 0 & 0 & \cdots & \cdots & 0 \\ 0 & 0 & 0 & \sqrt{3} & 0 & \cdots & \cdots & 0 \\ 0 & 0 & 0 & 0 & \sqrt{4} & \cdots & \cdots & 0 \\ 0 & 0 & 0 & 0 & \sqrt{4} & \cdots & \cdots & 0 \\ \vdots & \vdots \end{pmatrix}$$
(8.29)

Note that the matrices $[\hat{a}]$ and \hat{a}^{\dagger} are Hermitian conjugate to each other, as expected. The Hamiltonian operator is represented by a diagonal matrix since $|n\rangle$ are its eigenstates, and it is given by

$$[\hat{\mathcal{H}}] = \hbar \omega \begin{pmatrix} 1/2 & 0 & 0 & 0 \cdots & 0 \\ 0 & 3/2 & 0 & 0 \cdots & 0 \\ 0 & 0 & 5/2 & 0 \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix}$$
(8.30)

The matrix element representing the position operator \hat{X} is given by

$$< n |\hat{X}|m > = \sqrt{\frac{\hbar}{2m\omega}} \left(< n |\hat{a}^{\dagger}|m > + < n |\hat{a}|m > \right)$$

$$= \sqrt{\frac{\hbar}{2m\omega}} \left(\sqrt{n} \ \delta_{n,m+1} + \sqrt{n+1} \ \delta_{n,m-1} \right)$$

$$(8.31)$$

Substituting the expression of the derivative of $H_n(y)$ into the relation of H_{n+1} , yields

$$H_{n+1}(z) = 2zH_n(z) - 2nH_{n-1}(z)$$

Thus, \hat{X} can be represented by the matrix¹²³

$$[\hat{X}] = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 & \cdots & \cdots \\ 1 & 0 & \sqrt{2} & 0 & 0 & 0 & \cdots & \cdots \\ 0 & \sqrt{2} & 0 & \sqrt{3} & 0 & 0 & \cdots & \cdots \\ 0 & 0 & \sqrt{3} & 0 & \sqrt{4} & 0 & \cdots & \cdots \\ 0 & 0 & 0 & \sqrt{4} & 0 & \sqrt{5} & \cdots & \cdots \\ \vdots & \vdots \end{pmatrix}$$
(8.32)

For the operator \hat{X}^2 , the matrix element reads

$$< n|\hat{X}^{2}|m> = \frac{\hbar}{2m\omega} \left(< n|\hat{a}^{\dagger}\hat{a}^{\dagger}|m> + < n|\hat{a}\hat{a}|m> + \frac{2}{\hbar\omega} < n|\hat{\mathcal{H}}|m> \right)$$

$$= \frac{\hbar}{2m\omega} \left(\sqrt{n(n-1)} \,\delta_{n,m+2} + \sqrt{(n+1)(n+2)} \,\delta_{n,m-2} + \frac{(2n+1)}{\hbar\omega} \delta_{nm} \right)$$
(8.33)

So, the matrix representation of \hat{X}^2 is¹²⁴

$$[\hat{X}^{2}] = \frac{\hbar}{2m\omega} \begin{pmatrix} 1 & 0 & \sqrt{2} & 0 & 0 & \cdots & \cdots \\ 0 & 3 & 0 & \sqrt{6} & 0 & \cdots & \cdots \\ \sqrt{2} & 0 & 5 & 0 & \sqrt{12} & \cdots & \cdots \\ 0 & \sqrt{6} & 0 & 7 & 0 & \cdots & \cdots \\ 0 & 0 & \sqrt{12} & 0 & 9 & \cdots & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix}$$
(8.34)

Note that the diagonal element $[\hat{X}^2]_{nn}$ represents the expectation values of x^2 , in the state $|n\rangle$. So, from the above equation we have

$$\langle \hat{X}^2 \rangle_n = \frac{(2n+1)}{2m\omega^2} \hbar\omega = \frac{E_n}{m\omega^2}$$

$$(8.35)$$

Thus, the expectation value of the potential of the harmonic oscillator in this state is

$$\langle V(\hat{X}) \rangle_n = \frac{1}{2}m\omega^2 \langle \hat{X}^2 \rangle = \frac{E_n}{2}$$
 (8.36)

A similar analysis can be carried out for the momentum operator \hat{P} and the kinetic energy operator $\hat{T} = \hat{P}^2/2m$. I will leave to you as an exercise to show that $\langle \hat{T} \rangle_n = \langle V(\hat{X}) \rangle = E_n/2^{125}$.

8.5 Coherent States

¹²³We could obtain $[\hat{X}]$ by simply using $[\hat{X}] = \sqrt{\frac{\hbar}{2m\omega}} \left([\hat{a}] + [\hat{a}^{\dagger}] \right)$.

 $^{^{124}}$ It can also be obtained by matrix multiplication of $[\hat{X}]$ with itself.

 $^{^{125}{\}rm The}$ same relationship also holds for a classical harmonic oscillator when one takes the average over a full period of oscillation.

9 Theory of Angular Momentum

9.1 Angular Momentum Algebra

The orbital angular momentum of a single particle of mass m, located at a position \vec{r} with respect to a reference point P_0 at \vec{r}_0 is defined by

$$\vec{L} = (\vec{r} - \vec{r_0}) \times \vec{p} \tag{9.1}$$

where $\vec{p} = m\vec{v}$ is the momentum vector of the particle. For simplicity we choose the origin of system coordinates at P_0 so that the angular momentum vector takes the simple form

$$\vec{L} = \vec{r} \times \vec{p} \tag{9.2}$$

In terms of components, we have

$$L_x = yp_z - zp_y, \quad L_y = zp_x - xp_z, \quad L_z = xp_y - yp_x.$$
 (9.3)

which can be written in a compact form as

$$L_{i} = \sum_{j,k=1}^{3} \epsilon_{ijk} x_{j} p_{k}, \quad i = 1, 2, 3$$
(9.4)

with the index 1 corresponds to the x component, 2 to the y component and 3 to the z component. The symbol ϵ_{ijk} is the totally antisymmetric tensor, called the Levi-Civita tensor, defined as

$$\epsilon_{ijk} = \begin{cases} 0, & \text{if two or more indices are equal,} \\ 1, & \text{if ijk are an even permutation of 123,} \\ -1, & \text{if ijk are an odd permutation of 123.} \end{cases}$$
(9.5)

In quantum mechanics, to find angular momentum operator, one replaces the position and the momentum by their corresponding operators, i.e.

$$\hat{L}_{i} = \sum_{j,k=1}^{3} \epsilon_{ijk} \hat{X}_{j} \hat{P}_{k}, \quad i = 1, 2, 3$$
(9.6)

So, in the position representation (9.6) reads

$$\hat{L}_i = -i\hbar \sum_{j,k=1}^3 \epsilon_{ijk} x_j \frac{\partial}{\partial x_k}, \quad i = 1, 2, 3$$
(9.7)

or, more explicitly,

$$\hat{L}_{x} = -i\hbar \left(y \frac{\partial}{\partial z} - z \frac{\partial}{\partial y} \right)
\hat{L}_{y} = -i\hbar \left(z \frac{\partial}{\partial x} - x \frac{\partial}{\partial z} \right)
\hat{L}_{z} = -i\hbar \left(x \frac{\partial}{\partial y} - y \frac{\partial}{\partial x} \right)$$
(9.8)

Note that \hat{L}_i is Hermitian since

$$\left(\hat{L}_{i}\right)^{\dagger} = \sum_{j,k=1}^{3} \epsilon_{ijk} \left(\hat{P}_{k}\right)^{\dagger} \left(\hat{X}_{j}\right)^{\dagger} = \sum_{j,k=1}^{3} \epsilon_{ijk} \hat{P}_{k} \hat{X}_{j}$$

$$= \sum_{j,k=1}^{3} \epsilon_{ijk} \hat{X}_{j} \hat{P}_{k} - i\hbar \sum_{j,k=1}^{3} \epsilon_{ijk} \delta_{ij}$$

$$= \hat{L}_{i}$$

$$(9.9)$$

where the term $\sum_{j,k=1}^{3} \epsilon_{ijk} \delta_{ij}$ = since the Levi-Cevita tensor vanishes whenever two indices are equal. Hence, the components of $\hat{\vec{L}}$ corresponds to observables. So, the question now is can one measure all the components of the angular momentum simultaneously?. The answer turns out to be "NO" and that is because they do not commute with each other. For example, we have

$$[L_x, L_y] = [\hat{Y}\hat{P}_z - \hat{Z}\hat{P}_y, \hat{Z}\hat{P}_x - \hat{X}\hat{P}_z]$$

$$= \hat{Y}\hat{P}_x[\hat{P}_z, \hat{Z}] + \hat{P}_y\hat{X}[\hat{Z}, \hat{P}_z]$$

$$= i\left(\hat{X}\hat{P}_y - \hat{Y}\hat{P}_x\right) = i\hbar\hat{L}_z$$
(9.10)

Similarly, we can calculate the other commutators and find

$$[L_y, L_z] = i\hbar \hat{L}_x, \qquad [L_z, L_x] = i\hbar \hat{L}_y \tag{9.11}$$

which shows that the commutations relations between different components of the angular momentum operator can be obtained by a simple cyclic permutation of x, y, and z. In compact notation, we can write¹²⁶

$$[L_i, L_j] = i\hbar \sum_k \epsilon_{ijk} \hat{L}_k$$
(9.12)

$$\sum_{i,j} \epsilon_{ijk} \hat{L}_i \hat{L}_j = i\hbar \hat{L}_k$$

 $^{^{126}}$ In fact, we can also express the left hand side of (9.12) in terms of the Levi-Cevita tensor and write

which is just the cross product of the operator angular momentum with itself. So, in quantum mechanics, the cross product of an operator with itself is not necessarily zero.

It is important to note that the above commutation relations have been derived without referring to a specific representation for the position and momentum operators. Using the above result, we can calculate the commutator of the square of the angular momentum operator, \hat{L}^2 , with one of its component, \hat{L}_i , i.e.

$$\begin{bmatrix} \hat{L}^2, \hat{L}_i \end{bmatrix} = \sum_j \begin{bmatrix} \hat{L}_j \hat{L}_j, \hat{L}_i \end{bmatrix} = \sum_j \hat{L}_j \begin{bmatrix} \hat{L}_j, \hat{L}_i \end{bmatrix} + \sum_j \begin{bmatrix} \hat{L}_j, \hat{L}_i \end{bmatrix} \hat{L}_j$$
(9.13)
$$= i \sum_{j,k} \epsilon_{jik} \left(\hat{L}_j \hat{L}_k + \hat{L}_k \hat{L}_j \right)$$

However, the term in bracket is completely symmetric under the exchange of the j and k indices, and so its multiplication by the totally antisymmetric tensor ϵ_{ijk} vanishes when summing over the indices j and k. Therefore, we obtain

$$[\hat{L}^2, \hat{L}_i] = 0$$
 (9.14)

In general, a Hermitian operator $\hat{\vec{J}} = (\hat{J}_1, \hat{J}_2, \hat{J}_3)$ is said to be an angular momentum if it satisfies the following commutation relations

$$[J_i, J_j] = i \sum_k \epsilon_{ijk} \hat{J}_k \tag{9.15}$$

which defines the **algebra** of angular momentum. Furthermore, as we will show below, angular momentum is intimately related to the group of rotational transformations. Similar to the orbital angular momentum, the square of the angular momentum operator, \hat{J}^2 , commutes with all the components \hat{J}_i , i.e.

$$[\hat{J}^2, \hat{J}_i] = 0 \tag{9.16}$$

9.2 Quantization of Angular momentum

The commutation relations between the components of the angular momentum in (9.15) have a far-reaching consequences. As we will see it leads to the quantization of the angular momentum.

The fact that $[\hat{J}^2, \hat{J}_i] = 0$, and the \hat{J}_i do not commute between themselves, implies that one can form set of two commuting operators formed by \hat{J}^2 and one of the components of the angular momentum. We choose this set to be $\{\hat{J}^2, \hat{J}_3\}$, and denote the common eigenstates by $|\lambda, m\rangle$ with

$$\hat{J}^{2}|\lambda, m\rangle = \lambda \ \hbar^{2}|\lambda, m\rangle, \qquad (9.17)$$
$$\hat{J}_{3}|\lambda, m\rangle = m \ \hbar|\lambda, m\rangle$$

Here λ and m are dimensionless real numbers. The explicit appearance of \hbar^2 and \hbar is to match the dimension the \hat{J}^2 and \hat{J}_z , respectively. Now, the average of the operator $(\hat{J}^2 - \hat{J}_z)$ in the state $|\lambda, m\rangle$ reads

$$\langle \lambda, m | \hat{J}^2 - \hat{J}_z^2 | \lambda, m \rangle = (\lambda - m^2)\hbar^2$$

$$(9.18)$$

where we assumed that $|\lambda, m \rangle$ are orthornormal. However, the left hand side of the equation above can be written as

$$<\lambda, m |\hat{J}^{2} - \hat{J}_{z}^{2}|\lambda, m > = <\lambda, m |\hat{J}_{x}^{2}|\lambda, m > + <\lambda, m |\hat{J}_{y}^{2}|\lambda, m >$$

$$= ||\hat{J}_{x}|\lambda, m > ||^{2} + ||\hat{J}_{y}|\lambda, m > ||^{2} \ge 0$$
(9.19)

which implies that the operator $(\hat{J}^2 - \hat{J}_3^2)$ is positive definite, and hence

$$\lambda \ge m^2 \tag{9.20}$$

We introduce the ladder operators, defined as

$$\hat{J}_{\pm} = \hat{J}_1 \pm i\hat{J}_2 \tag{9.21}$$

Note that these operators are the adjoints of one another. By expressing \hat{J}_x and \hat{J}_y in terms of \hat{J}_{\pm} , we can rewriting \hat{J}^2 in the following equivalent forms

$$\hat{J}^2 = \frac{1}{2}(\hat{J}_+\hat{J}_- + \hat{J}_-\hat{J}_+ + \hat{J}_z^2)$$
(9.22)

$$= \dot{J}_{+}\dot{J}_{-} + \dot{J}_{z}^{2} - \hbar\dot{J}_{z}$$
(9.23)

$$= J_{-}J_{+} + J_{z}^{2} + \hbar J_{z}$$
(9.24)

With the use of the commutation relations between the \hat{J}_i in (9.15), we find that

$$[\hat{J}_z, \hat{J}_{\pm}] = \pm \hbar \ \hat{J}_{\pm}, \qquad [\hat{J}_+, \hat{J}_-] = 2\hbar \hat{J}_z \tag{9.25}$$

Now, applying \hat{J}_z on the states $\hat{J}_{\pm}|\lambda, m >$ gives

$$\hat{J}_{z}\hat{J}_{\pm}|\lambda,m\rangle = [\hat{J}_{z},\hat{J}_{\pm}]|\lambda,m\rangle + \hat{J}_{\pm}\hat{J}_{z}|\lambda,m\rangle$$

$$= \pm\hbar\hat{J}_{\pm}|\lambda,m\rangle + m\hbar\hat{J}_{\pm}|\lambda,m\rangle$$

$$= (m\pm 1)\hbar |\lambda,m\rangle$$
(9.26)

Thus, $\hat{J}_{\pm}|\lambda, m\rangle$ is eigenstates of \hat{J}_z with eigenvalues $(m \pm 1)$, without changing the the value of λ . Therefore, by repeatedly applying \hat{J}_+ to eigenstates of \hat{J}_z leads to a tower of states with larger values of m. However, this sequence of states has to stop at some point otherwise the bound in Eq (9.20) gets violated. That is there should be a state for some value $m = m_+$ such that the state is annihilated by \hat{J}_+ , i.e.

$$\hat{J}_+|\lambda, m_+\rangle = 0 \tag{9.27}$$

This implies that

$$0 = <\lambda, m_{+} |\hat{J}_{-} \hat{J}_{+} |\lambda, m_{+} >$$

= <\lambda, m_{+} |(\beta^{2} - \beta_{z}^{2} - \beta \beta_{z})|\lambda, m_{+} >
= \beta^{2} (\lambda - m_{+}^{2} - m_{+}) (9.28)

Similarly, by successive application of \hat{J}_{-} one can obtain arbitrarily small negative value m which will violate the upper bound on m^2 , and so there must be some value $m = m_{-}$ for which the state $|\lambda, m_{-}\rangle$ is annihilated by \hat{J}_{-} , i.e.

$$\hat{J}_{-}|\lambda, m_{-}\rangle = 0 \tag{9.29}$$

which implies that

$$0 = \langle \lambda, m_{-} | \hat{J}_{+} \hat{J}_{-} | \lambda, m_{-} \rangle$$

= $\langle \lambda, m_{-} | (\hat{J}^{2} - \hat{J}_{z}^{2} - \hbar \hat{J}_{z}) | \lambda, m_{-} \rangle$
= $\hbar^{2} (\lambda - m_{-}^{2} + m_{-})$ (9.30)

Now, by a successive application of \hat{J}_+ on the state $|\lambda, m_-\rangle$ one must end up at the state $|\lambda, m_+\rangle$, otherwise the sequence would increase indefinitely. Therefore, we have

$$m_{+} - m_{-} = n \tag{9.31}$$

where n is a positive integer which represents the number of times that the operator \hat{J}_+ has been applied in order to reach the state $|\lambda, m_+\rangle$. Moreover, by combining Eq (9.28) and Eq (9.30), we can eliminate λ and get an equation in terms of m_+ and m_- :

$$(m_{+} + m_{-})(m_{+} - m_{-} + 1) = 0 (9.32)$$

Since the term in the second parenthesis is positive definite, it follows that $m_{+} = -m_{-}$. Defining $j = m_{+}$, and using Eq (9.31) gives

$$j = \frac{n}{2} \tag{9.33}$$

which shows that j can be either integer or half integer:

$$j = \{0, \frac{1}{2}, 1, \frac{3}{2}, \dots\}$$
(9.34)

Substituting the relation $m_{+} = -m_{-} \equiv j$ into Eq (9.28), yields

$$\lambda = j(j+1) \tag{9.35}$$

We shall from now change the label λ of the eigenstates of \hat{J}_z and \hat{J}_z by j. Thus, the action of the angular momentum operators on a state $|j, m\rangle$ are given by

$$\hat{J}^{2}|j,m\rangle = j(j+1)\hbar^{2}|j,m\rangle$$
(9.36)

$$\hat{J}_z|j,m\rangle = m\hbar|j,m\rangle \tag{9.37}$$

$$\hat{J}_{\pm}|j,m\rangle = \sqrt{j(j+1) \pm (m\pm 1)}\hbar|j,m\pm 1\rangle$$
(9.38)

Note that for each j, the quantum number m takes the values

$$m = \{-j, -j+1, -j+2, \dots, j-1, j\}$$
(9.39)

Therefore, there are (2j+1) states for a given j. One can show that the space of states, defined by

$$\mathcal{H}^{(j)} = \{ |\psi\rangle = \sum_{m=-j}^{+j} a_m |j, m\rangle, \text{ where } a_m \in \mathbb{C} \}$$
(9.40)

is a Hilbert space of dimension (2j + 1) and with $\{|j, m \rangle, m = -j, -j + 1, ..., j\}$ forming an orthonormal basis.

9.3 Angular Momentum and Rotations in \mathbb{R}^3

A rotation in 3-dimension (not necessarily in the position space) is characterized by an axis of rotation, defined by a unit vector \vec{n} , and an angle of rotation. We will denote it by $\mathcal{R}_{\vec{n}}(\theta)$, or sometimes by $\mathcal{R}(\vec{\theta})$, where $\vec{\theta} = \theta \ \vec{n}$. Moreover, the set of these rotations forms a non-abelian group, called SO(3), parametrized by three independent parameters. So, a rotation $\mathcal{R}(\vec{\theta})$ transform a vector \vec{V} into a new vector

$$\vec{V'} = \mathcal{R}(\vec{\theta}) \ \vec{V} \tag{9.41}$$

In cartesian coordinates, the above transformation can be expressed as

$$V_i' = \sum_{j=1}^3 \mathcal{R}(\vec{\theta})_{ij} \ V_j \tag{9.42}$$

The rotation transformation leave the scalar product between any two vectors unchanged, and so we can write

$$\sum_{i=1}^{3} V_i' V_i' = \sum_{i,j,k} \mathcal{R}(\vec{\theta})_{ij} \mathcal{R}(\vec{\theta})_{ik} \ V_j V_k = \sum_{i=1}^{3} V_i V_i = \sum_{j,k} \delta_{jk} V_j V_k \tag{9.43}$$

from which it follows that

$$\sum_{i} \mathcal{R}(\vec{\theta})_{ij} \mathcal{R}(\vec{\theta})_{ik} = \delta_{jk}$$
(9.44)

Using the fact that $\mathcal{R}(\vec{\theta})_{ik} = [\mathcal{R}(\vec{\theta})^T]_{ki}$, the above equation reads in matrix form as

$$\mathcal{R}(\vec{\theta})^T \mathcal{R}(\vec{\theta}) = \mathbb{I}_{3 \times 3} \tag{9.45}$$

Taking the determinant of both sides of the equation, yields

$$\det \mathcal{R}(\theta) = \pm 1 \tag{9.46}$$

The set of transformations that have $\det \mathcal{R}(\vec{\theta}) = -1$, do not form a group since the identity transformation has $\det(\mathbb{I}) = +1$, and are called **improper rotations**. This set of transformation represents an inversion or a combination of rotation and inversion. Thus, the group SO(3) is the group of **proper rotations** defined as¹²⁷

$$SO(3) := \{\mathcal{R}, \ \mathcal{R}^T \mathcal{R} = \mathbb{I}, \text{and } \det(\mathcal{R}) = 1\}$$

$$(9.47)$$

For instance, the transformation describing the rotation around the x axis, defined by the unit vector \vec{e}_1 , by an angle α is given by

$$\mathcal{R}_{\vec{e}_1}(\alpha) = \begin{pmatrix} 1 & 0 & 0\\ 0 \cos \alpha & -\sin \alpha\\ 0 & \sin \alpha & \cos \alpha \end{pmatrix}.$$
 (9.48)

Similarly, rotations around the y and z axis defined by the unit vectors \vec{e}_2 and \vec{e}_3 , respectively, are given by

$$\mathcal{R}_{\vec{e}_2}(\beta) = \begin{pmatrix} \cos\beta & 0 \sin\beta \\ 0 & 1 & 0 \\ -\sin\beta & 0 \cos\beta \end{pmatrix}, \qquad \mathcal{R}_{\vec{e}_3}(\gamma) = \begin{pmatrix} \cos\gamma & -\sin\gamma & 0 \\ -\sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{pmatrix}$$
(9.49)

¹²⁷The Letters "SO" in the group SO(3) stands for "special orthogonal". It is special because $\det \mathcal{R}(\vec{\theta}) = +1$, and orthogonal because $\mathcal{R}(\vec{\theta})^T \mathcal{R}(\vec{\theta}) = \mathbb{I}_{3\times 3}$. It is a very simple exercise to check that the set of matrices as defined in (9.47) for a group under the operation of matrix multiplication.

The expression of a rotation about an arbitrary axis defined by a unit vector \vec{n} can be found as follows. Any vector \vec{V} can be decomposed into two vectors; one parallel to the axis of rotation, \vec{V}_{\parallel} and the second one \vec{V}_{\perp} , that is perpendicular to it, with

$$\vec{V}_{\parallel} = (\vec{V}.\vec{n})\vec{n}, \quad \vec{V}_{\perp} = \vec{V} - (\vec{V}.\vec{n})\vec{n}$$
(9.50)

Then, the rotation of \vec{V} by an angle ψ around the direction \vec{n} can be written as

$$\mathcal{R}_{\vec{n}}\vec{V} = \mathcal{R}_{\vec{n}}\vec{V}_{\parallel} + \mathcal{R}_{\vec{n}}\vec{V}_{\perp}$$

$$= \vec{V}_{\parallel} + \cos\psi\vec{V}_{\perp} + \sin\psi(\vec{n}\times\vec{V})$$

$$= \cos\psi\vec{V} + (1 - \cos\psi)(\vec{V}.\vec{n})\vec{n} + \sin\psi\vec{n}\times\vec{V}$$
(9.51)

In the second equality above we used the fact that \vec{V}_{\parallel} remains unchanged under the action of $\mathcal{R}_{\vec{n}}$ since it is parallel to the axis of rotation. In terms of components, the above equation reads

$$\sum_{j=1}^{3} \mathcal{R}_{\vec{n}}(\psi)_{ij} \vec{V}_{j} = \sum_{j=1}^{3} \left[\cos \psi \ \delta_{ij} + (1 - \cos \psi) n_{i} n_{j} - \sin \psi n_{k} \epsilon_{ijk} n_{k} \right] V_{j}$$
(9.52)

where the ϵ_{ijk} comes from the definition of the scalar product. Since the above relation holds for any vector \vec{V} , then it implies that

$$\mathcal{R}_{\vec{n}}(\psi)_{ij} = \cos\psi \,\,\delta_{ij} + (1 - \cos\psi) \,\,n_i n_j - \sin\psi \,\,\epsilon_{ijk} n_k \tag{9.53}$$

This is known as **Rodrigues' rotation formula**¹²⁸.

$$\mathcal{R}_{\vec{n}}(\psi)_{ij} = A(\psi) \ \delta_{ij} + B(\psi) \ n_i n_j + C(\psi) \ \epsilon_{ijk} \ n_k$$

Second, we use the property that the action of a rotation about \vec{n} does not affect \vec{n} or any vector parallel to it, i.e.

$$\mathcal{R}_{\vec{n}}(\psi) \ \vec{n} = \vec{n}$$

which after substituting the above expression of $\mathcal{R}_{\vec{n}}(\psi)_{ij}$ yields

$$A(\psi) + B(\psi) = 1$$

Finally, we consider $\vec{n} = \vec{k}$, for which $\mathcal{R}_{\vec{k}}(\psi)_{11} = \cos \psi$ and $\mathcal{R}_{\vec{k}}(\psi)_{12} = -\sin \psi$. This leads to

$$A(\psi) = \cos \psi, \quad C(\psi) = -\sin \psi, \quad B(\psi) = 1 - \cos \psi$$

Hence, we obtain the expression of $\mathcal{R}_{\vec{n}}(\psi)_{ij}$ given in (9.53).

¹²⁸Another way of deriving the Rodrigues is as follows. First, we note that $\mathcal{R}_{\vec{n}}(\psi)_{ij}$ is an element of a tensor of rank two. Thus, given at our disposal the symmetric tensor δ_{ij} of rank 2, the totally antisymmetric tensor ϵ_{ijk} of rank 3, and the vector n_i , we can express $\mathcal{R}_{\vec{n}}(\psi)_{ij}$ in the form

Let us now consider an infinitesimal rotation associated with an angle $\delta \psi \ll 1$ by Taylor expanding $\mathcal{R}_{\vec{n}}(\psi)_{ij}$ in terms of $\delta \psi$ and keeping only lowest order. We obtain

$$\mathcal{R}(\vec{\delta\psi})_{ij} = \delta_{ij} - \delta\psi \ \epsilon_{ijk} \ n_k + \dots \tag{9.54}$$

Defining $\delta \psi_k = \delta \psi \ n_k$, we write the rotation matrix $\mathcal{R}(\vec{\delta \psi})$ in the form

$$\mathcal{R}(\vec{\delta\psi}) = \mathbb{I} - i \sum_{k} \delta\psi_k \ \mathcal{M}^k + \dots$$
(9.55)

where $\{\mathcal{M}^k, k = 1, 2, 3\}$ are matrices 3×3 , and they are called the generators of the transformation along the direction \vec{e}_k^{129} , with elements given by

$$M_{ij}^k = -i\epsilon_{kij}, \quad k = 1, 2, 3.$$
 (9.56)

So, in matrix form, the generators of rotations along the x, y, and z axis read

$$\mathcal{M}^{1} = i \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \mathcal{M}^{2} = i \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix} \quad \mathcal{M}^{3} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}$$
(9.57)

Defining $\mathcal{M}^i = \frac{J_i}{\hbar}$, it is straightforward to check that

$$[J^i, J^j] = i\epsilon_{ijk} \ J^k. \tag{9.58}$$

Thus, the generators of a rotation satisfy the algebra of angular momentum.

According to (9.56), under an infinitesimal rotation $\mathcal{R}(\vec{\delta\psi})$, an arbitrary differentiable function $F(x_i)$ of the the coordinate x^i , transforms as

$$F(x'_{i}) \simeq F(x_{i} - \delta\psi_{k} \ \epsilon_{kij} \ x_{j}) = F(x_{i}) - \delta\psi_{k} \ \epsilon_{kij} \ x_{j} \frac{\partial F}{\partial x_{i}}$$

$$\simeq F(x_{i}) - \delta\psi_{k} \ \frac{1}{2} \epsilon_{kij} \left(x_{i} \frac{\partial}{x_{j}} - x_{j} \frac{\partial}{x_{i}} \right) \ F(x_{i})$$
(9.59)

from which we infer that the generator of this rotation in the space of functions of coordinates is represented by the operator

$$\mathcal{M}^{k} = -\frac{i}{2} \epsilon_{kij} \left(x_{i} \frac{\partial}{x_{j}} - x_{j} \frac{\partial}{x_{i}} \right), \quad k = 1, 2, 3$$
(9.60)

which, after rescaling by an \hbar , correspond to the same expressions of the orbital angular momentum given in (9.8). Hence, in the position representation, the generators of

 $^{^{129}\}mathrm{This}$ is similar to the momentum operator which, as we saw before, is the generator of the translation.

rotation correspond to the orbital angular momenta.

Now, to go from infinitesimal transformation to a full rotation we make use of the identity

$$\lim_{N \to \infty} \left(1 - i \frac{\psi \mathcal{A}}{N} \right)^N = e^{-i\psi \vec{\mathcal{A}}}$$
(9.61)

for some arbitrary matrix \mathcal{A} . Therefore, a rotation with a finite angle ψ can be represented by an exponential

$$\mathcal{R}(\vec{\psi}) = e^{-\frac{i}{\hbar}\vec{\psi}\vec{J}} \tag{9.62}$$

where J^i are the angular momentum operators which in the position space correspond to the orbital angular momentum.

9.4 Representation of Rotation in Hilbert State

10 The Hydrogen Atom

11 Approximation Methods

Exact solution to the Schrödinger's equation exists only for a few idealized systems. Thus, in general cases, one must resort to approximate methods. A variety of methods have been developed and each has its area and range of applicability. In this chapter, we will consider approximation methods for stationary, i.e, corresponding to time-independent Hamiltonians.

11.1 Time Independent Perturbation Theory

This method is most suitable when the Hamiltonian of the system, \hat{H} , is close to a Hamiltonian \hat{H}_0 that can be solved exactly. In this case, we can write

$$\hat{H} = \hat{H}_0 + \delta \hat{H} \tag{11.1}$$

where $\delta \hat{H}$ is very small compared to \hat{H}_0 . As a result, the effect of $\delta \hat{H}$ on the spectrum of \hat{H}_0 will be small, and hence $\delta \hat{H}$ is called perturbation. We can introduce a parameter λ which is much smaller than one, and write $\delta \hat{H}$ in the form

$$\delta \hat{H} = \lambda \hat{V} \tag{11.2}$$

Thus, the eigenvalue problem of the system becomes

$$\left(\hat{H}_0 + \lambda \hat{V}\right) |\Psi_n\rangle = E_n |\Psi_n\rangle \tag{11.3}$$

There are two cases to consider depending on whether the spectrum of \hat{H}_0 is degenerate or non-degenerate.

11.1.1 Non Degenerate Spectrum

If H_0 has no degenerate eigenvalues, we can write

$$\hat{H}_0 |\phi_n\rangle = E_n^{(0)} |\phi_n\rangle \tag{11.4}$$

where the exact eigenvalues $E_n^{(0)}$ and the exact eigenstates $|\phi_n\rangle$ are known. Thus, we expand $|\Psi_n\rangle$ and E_n in power series of the parameter λ^{130} :

$$E_n = E_n^{(0)} + \lambda E_n^{(1)} + \lambda^2 E_n^{(2)} + \dots$$

$$|\Psi_n \rangle = |\phi_n \rangle + \lambda |\Psi_n^{(1)} \rangle + \lambda^2 |\Psi_n^{(2)} \rangle + \dots$$
(11.5)

The aim of perturbation theory is to find $E_n^{(1)}, E_n^{(2)}, ...,$ and $|\Psi_n^{(1)}\rangle, |\Psi_n^{(2)}\rangle, ...$ For $\lambda \ll 1$, the first few terms provide a reliable description of the system. Hence, we will be concerned only with the first two terms in the above expansion. Then, we obtain

$$\left(\hat{H}_{0}+\lambda\hat{W}\right)\left(|\phi_{n}\rangle+\lambda|\Psi_{n}^{(1)}\rangle+\lambda^{2}|\Psi_{n}^{(2)}\rangle\right)=\left(E_{n}^{(0)}+\lambda E_{n}^{(1)}+\lambda^{2}E_{n}^{(2)}\right)\left(|\phi_{n}\rangle+\lambda|\Psi_{n}^{(1)}\rangle+\lambda^{2}|\Psi_{n}^{(2)}\rangle\right)$$

¹³⁰Even if the parameter of parameter of the perturbation is sufficiently small, it does not necessarily mean that the above expansion exists.

By matching the powers of λ on both sides of this equation we obtain

$$\lambda^0 \to \hat{H}_0 | \phi_n \rangle = E_n^{(0)} | \phi_n \rangle \tag{11.6}$$

$$\lambda^{1} \to \hat{H}_{0} |\psi_{n}^{(1)} \rangle + \hat{V} |\phi_{n} \rangle = E_{n}^{(0)} |\psi_{n}^{(1)} \rangle + E_{n}^{(1)} |\phi_{n} \rangle$$
(11.7)

$$\lambda^2 \to \hat{H}_0 |\psi_n^{(2)} > + \hat{V} |\psi_n^{(1)} \rangle = E_n^{(0)} |\psi_n^{(2)} \rangle + E_n^{(1)} |\psi_n^{(1)} \rangle + E_n^{(2)} |\phi_n \rangle$$
(11.8)

Similarly, we can expand the normalization equation $\langle \psi_n | \psi_n \rangle$ in powers of λ , and require that all terms of the same power in λ cancel independently, giving

$$\lambda^0 \to \langle \phi_n | \phi_n \rangle = 1 \tag{11.9}$$

$$\lambda^{1} \to \langle \psi_{n}^{(1)} | \phi_{n} \rangle + \langle \phi_{n} | \psi_{n}^{(1)} \rangle = 0$$
(11.10)

$$\lambda^2 \to \langle \psi_n^{(2)} | \phi_n \rangle + \langle \psi_n^{(1)} | \psi_n^{(1)} \rangle + \langle \phi_n | \psi_n^{(2)} \rangle = 0$$
(11.11)

Without lose of generality, we can choose the phase of the full eigenstates $|\psi_n\rangle$ to be such that $\langle \phi_n | \psi_n \rangle$ is real-valued. In terms of powers of λ , we have

$$<\phi_n|\psi_n>=1+\lambda<\phi_n|\psi_n^{(1)}>+\lambda<\phi_n|\psi_n^{(2)}>+..$$
 (11.12)

In order for the right hand side to be real for any real λ requires that each term be independently real-valued, i.e

$$<\phi_n|\psi_n^{(k)}> = <\psi_n^{(k)}|\phi_n>$$
 (11.13)

To obtain the j^{th} correction to the n^{th} energy eigenvalue, we multiply the equation in the j^{th} power of λ by $\langle \phi_n |$ and solve for $E_n^{(j)}$, giving

$$E_n^{(j)} = \langle \phi_n | \hat{V} | \psi_n^{(j)} \rangle - \sum_{k=1}^{j-1} E_n^{(k)} \langle \phi_n | \psi_n^{(j-k)} \rangle$$
(11.14)

where in the summation j > k, and the quantity $\langle \phi_n | \psi_n^{(m)} \rangle$ is obtained from the normalization condition (12), and it is given by

$$<\phi_n|\psi_n^{(m)}> = -\frac{1}{2}\sum_{k=1}^{j-1} <\psi_n^{(m-k)}|\psi_n^{(m)}>$$
 (11.15)

which must be solved iteratively. The j^{th} corrections to the n^{th} eigenstate can be obtained by expanding $|\psi_n\rangle$ in the basis $|\phi_n\rangle$,

$$|\psi_n \rangle = \sum_m \langle \phi_m | \psi_n \rangle | \phi_m \rangle$$

$$= |\phi_n \rangle \sum_{j=1} \lambda^j \langle \phi_n | \psi_n^{(j)} \rangle + \sum_{m \neq 0} |\phi_m \rangle \sum_{j=1} \lambda^j \langle \phi_m | \psi_n^{(j)} \rangle$$
(11.16)

where $\langle \phi_m | \psi_n^{(j)} \rangle$ can be obtained from the normalization condition of the state $|\psi_n \rangle$, which gives

$$<\phi_m|\psi_n^{(j)}> = -\frac{<\phi_m|V|\psi_n^{(j-1)}>}{E_m^{(0)} - E_n^{(0)}} + \sum_{k=1}^{j-1} \frac{E_n^{(k)} < \phi_m|\psi_n^{(j-k)}>}{E_m^{(0)} - E_n^{(0)}}$$
(11.17)

Let us introduce the notation $V_{mn} \equiv \langle \phi_n | V | \phi_n \rangle$ and $E_{mn} \equiv E_n^{(0)} - E_n^{(0)}$. Then we have

• First order approximation

$$E_n = E_n^{(0)} + E_n^{(1)}$$

$$= E_n^{(0)} + \lambda < \phi_n |V| \phi_n > = E_n^{(0)} + \lambda V_{nn}$$
(11.18)

and

$$|\psi_n\rangle = |\phi_n\rangle + \lambda \sum_{m \neq 0} |\phi_m\rangle < \phi_m |\psi_n^{(1)}\rangle$$

$$= |\phi_n\rangle - \lambda \sum_{m \neq n} \frac{V_{mn}}{E_{mn}} |\phi_m\rangle$$
(11.19)

• Second order approximation

Setting j = 2 in equations (14), (16), and (17), we get

$$E_n = E_n^{(0)} + E_n^{(1)} + E_n^{(2)}$$

$$= E_n^{(0)} + \lambda V_{nn} - \lambda^2 \sum_{m \neq n} \frac{|V_{mn}|^2}{E_{mn}}$$
(11.20)

and

$$|\psi_{n}\rangle = \left(1 - \frac{\lambda^{2}}{2} \sum_{m \neq n} \frac{|V_{mn}|^{2}}{E_{mn}^{2}}\right) |\phi_{n}\rangle +$$

$$\sum_{m \neq 0} |\phi_{m}\rangle \left[-\lambda \sum_{m \neq n} \frac{V_{mn}}{E_{mn}} + \lambda^{2} \left(\sum_{m' \neq n} \frac{V_{mm'V_{m'n}}}{E_{mn}E_{m'n}} - \frac{V_{mn}V_{nn}}{E_{mn}^{2}}\right)\right]$$
(11.21)

There is another way to derive the expressions of the energy eigenvalues for the unperturbed Hamiltonian. We write the Schrödinger equation as

$$(H_0 + \lambda V) |\psi_n\rangle = E_n |\psi_n\rangle \Rightarrow (\lambda V - \Delta_n) |\psi_n\rangle = (E_n^{(0)} - H_0) |\psi_n\rangle \qquad (11.22)$$

where $\Delta_n = E_n - E_n^{(0)}$ Let us introduce the projector operator

$$Q_n = 1 - |\phi_n| > <\phi_n| = \sum_{m \neq n} |\phi_m| > <\phi_m|$$
(11.23)

and note that $\langle \phi_n | (\lambda V - \Delta_n) | \psi_n \rangle = 0$. Thus we can have

$$\left(E_n^{(0)} - H_0\right)|\psi_n\rangle = Q_n\left(\lambda V - \Delta_n\right)|\psi_n\rangle$$
(11.24)

With the presence of the projector operator, It is now safe to write

$$|\psi_n\rangle = G_n \left(\lambda V - \Delta_n\right) |\psi_n\rangle \tag{11.25}$$

(11.26)

where

$$G_n = \frac{1}{E_n^{(0)} - H_0} Q_n \tag{11.27}$$

However, in the limit $\lambda \to 0$, the above equation gives $|\psi_n\rangle = 0$ instead of $|\psi_n\rangle \propto |\phi_n\rangle$. This is because the above solution is just a particular solution and one needs to add to that the solution to the homogeneous equation. The later are the unperturbed states. Thus, the general solution reads

$$|\psi_n\rangle = G_n \left(\lambda V - \Delta_n\right) |\psi_n\rangle + f_n(\lambda) |\phi_n\rangle$$
(11.28)

where $f(\lambda)$ is an arbitrary function with well define limit as $\lambda \to 0$. We $f_n(\lambda) = 1$ for any λ , and have

$$|\psi_n\rangle = |\phi_n\rangle + G_n \left(\lambda V - \Delta_n\right) |\psi_n\rangle \tag{11.29}$$

which results in $\langle \psi_n \rangle \neq 1$. However, this should not be a problem because we can always renormalize the states (wave function renormalization). With this choice we have

$$\langle \psi_n | \phi_n \rangle = 1 \tag{11.30}$$

Combining the above equation with the fact that the state $(\lambda V - \Delta_n) |\psi_n\rangle$ is perpendicular on $|\phi_n\rangle$, we deduce that

$$\Delta_n = \lambda < \phi_n | V | \psi_n > \tag{11.31}$$

We expands the states $|\psi_n\rangle$ and the shifts in the energies in powers of λ ,

$$|\psi_n\rangle = |\phi_n\rangle + \sum_{j=1} \lambda^j |\psi_n^{(j)}\rangle$$

$$\Delta_n = \sum_{j=1} \lambda^j \Delta_n^{(j)}$$
(11.32)

Thus, we have

$$E_n^{(j)} = E_n^{(0)} + \lambda V_{nn} + \sum_{j=2} \lambda^j < \phi_n |V| \psi_n^{(j-1)} >$$
(11.33)

and

$$\begin{split} |\psi_{n}\rangle &= |\phi_{n}\rangle - \lambda \sum_{m \neq n} \frac{<\phi_{m}|V|\phi_{n}\rangle}{E_{mn}} |m\rangle - \sum_{j=1} \lambda^{j+1} \sum_{m \neq n} \frac{<\phi_{m}|V|\psi_{n}^{(j)}\rangle}{E_{mn}} |m(\mathbb{M}.34) \\ &+ \sum_{j=1} \sum_{k=1} \sum_{m \neq n} \lambda^{j+k} \Delta_{n}^{(k)} \frac{<\phi_{m}|\psi^{(j)}\rangle}{E_{mn}} |m\rangle \\ &= |\phi_{n}\rangle - \lambda \sum_{m \neq n} \frac{<\phi_{m}|V|\phi_{n}\rangle}{E_{mn}} |m\rangle + \\ &\sum_{j=2} \lambda^{j} \sum_{m \neq n} \left[\sum_{k=1}^{(j-1)} \Delta_{n}^{(k)} \frac{<\phi_{m}|\psi^{(j-k)}\rangle}{E_{mn}} - \frac{<\phi_{m}|V|\psi_{n}^{(j-1)}\rangle}{E_{mn}} \right] |m\rangle \end{split}$$

By stetting j = 2 in the above equations, we find that the energy and the state of the system at the second order in perturbation theory, are given by

$$E_n = E_n^{(0)} + \lambda V_{nn} - \lambda^2 \sum_{m \neq n} \frac{|V_{mn}|^2}{E_{mn}}$$
(11.35)

$$|\psi_n \rangle = |\phi_n \rangle + \sum_{m \neq 0} |\phi_m \rangle \left[-\lambda \sum_{m \neq n} \frac{V_{mn}}{E_{mn}} + \lambda^2 \left(\sum_{m' \neq n} \frac{V_{mm'V_{m'n}}}{E_{mn}E_{m'n}} - \frac{V_{mn}V_{nn}}{E_{mn}^2} \right) \right]$$

We can define a state $|\psi_n\rangle_N = \mathcal{Z}_N^{1/2} |\phi_n\rangle$ such that $\langle_N \psi |\psi_n\rangle_N = 1$. This implies that

$$\begin{aligned} \mathcal{Z}_N^{-1} &= \left[<\phi_n | +\lambda < \psi_n^{(1)} | +\lambda^2 < \psi_n^{(2)} | + \dots \right] \left[|\phi_n > +\lambda | \psi_n^{(1)} > +\lambda^2 | \psi_n^{(2)} > + \dots \right] \\ &= 1 + \sum_{j,k=1} \lambda^{j+k} < \psi_n^{(k)} | \psi_n^{(j)} > \end{aligned}$$

where we used the fact that $\langle \phi_n | \psi_n \rangle = 1$. If we truncate the expansion to the second order in power of λ , we get

$$\mathcal{Z}^{-1} = \left(1 + \lambda^2 \sum_{m \neq n} \frac{|V_{mn}|^2}{E_{mn}^2}\right)$$

$$= \left(1 - \lambda^2 \sum_{m \neq n} \frac{|V_{mn}|^2}{E_{mn}^2}\right)^{-1}$$
(11.37)

Thus, to the second order in perturbation theory, the normalized state reads

$$|\psi_{n}\rangle = \left(1 - \frac{\lambda^{2}}{2} \sum_{m \neq n} \frac{|V_{mn}|^{2}}{E_{mn}^{2}}\right) |\phi_{n}\rangle + \sum_{m \neq 0} |\phi_{m}\rangle \left[-\lambda \sum_{m \neq n} \frac{V_{mn}}{E_{mn}} + \lambda^{2} \left(\sum_{m' \neq n} \frac{V_{mm'V_{m'n}}}{E_{mn}E_{m'n}} - \frac{V_{mn}V_{nn}}{E_{mn}^{2}}\right)\right]$$
(11.38)

11.1.2 Degenerate Spectrum

- 11.2 variational Methods
- 11.3 Time-Dependent Perturbation Theory

12 Formal Theory of Scattering

12.1 Møller Operator

The solution to the time dependent Schrodinger equation

$$i\hbar \frac{\partial}{\partial t} |\psi(t)\rangle = \hat{H} |\psi(t)\rangle$$
 (12.1)

is given by

$$|\psi(t)\rangle = e^{-i\frac{\hat{H}}{\hbar}}|\psi(0)\rangle \equiv U(t)|\psi(0)\rangle$$
 (12.2)

where $|\psi(0)\rangle$ is an arbitrary initial eigenstate of $\hat{H} = \hat{H}_0 + V$, with the potential V is supposed to represent the interaction of the projectile with the target, where as \hat{H}_0 describes the time development of the projectile state outside the interaction region. In general, \hat{H}_0 may include part of the interaction, but through out this chapter we will assume that \hat{H}_0 is simply the kinetic energy operator. We denote by $|\phi_a(t)\rangle$ the free wave packet state at time t, and which satisfy the equation

$$i\hbar\frac{\partial}{\partial t}|\phi_a(t)\rangle = \hat{H}_0|\phi_a(t)\rangle$$
(12.3)

The index *a* denotes an arbitrary initial distribution of momenta in the wave packet. The solution of Eq.(12.3) can be written in terms of an arbitrary initial eigenstate of \hat{H}_0 , $|\phi_a(0)\rangle$ as:

$$|\phi_a(t)\rangle = e^{-i\frac{\hat{H}_0}{\hbar}}|\phi_a(0)\rangle \equiv U_0(t)|\phi_a(0)\rangle$$
 (12.4)

A reasonable requirement that (12.2) is a scattering state will be to demand that long before the scattering process $|\psi(t)\rangle$ behaves as a free wave packet $|\phi_a(t)\rangle$. Formally, we require that

$$\lim |||\psi_a^{(+)}(t) > -|\phi_a(t) > || \to 0$$
(12.5)

$$t \to -\infty$$
 (12.6)

or, equivalently

$$\lim ||U(t)|\psi_a^{(+)}(0) > -U_0(t)|\phi_a(0) > || \to 0$$
(12.7)

$$t \to -\infty$$
 (12.8)

Since U(t) is a unitary operator, we can write the asymptotic condition

$$|\psi_a^{(+)}(0) > -\Omega_+ |\phi_a(0) > \tag{12.9}$$

where

$$\Omega_{+} = \Omega_{+}(\hat{H}, \hat{H}_{0}) = \lim U^{\dagger}(t)U_{0}(t) \equiv \lim W(t)$$
(12.10)

$$t \to -\infty$$
 (12.11)

which is called the MØller Operator. In order to show the existence of this operator, we have to prove that $limW(t)|\phi_a(0) >$ exists for any $|\psi_a(0) >$ in the Hilbert space \mathcal{H} . For that, let $t_2 < t_1 < 0$ and consider

$$||(W(t_{2}) - W(t_{1}))|\phi_{a} > || = ||\int_{t_{1}}^{t_{2}} dt \frac{dW}{dt}|\phi_{a} > ||$$

$$\leq \int_{t_{1}}^{t_{2}} dt ||dt \frac{dW}{dt}|\phi_{a} > ||$$

$$= \int_{t_{1}}^{t_{2}} dt ||U^{\dagger} \frac{1}{\hbar} \left(\hat{H} - \hat{H}_{0}\right) \phi_{a} > ||$$

$$= \int_{t_{1}}^{t_{2}} dt 1\hbar ||V\phi_{a} > ||$$

The existence of the integral for $t_2 \to -\infty$ means that the norm || has to fall off at least as $1/t^{1+\epsilon}$, with $\epsilon > 0$. We will show the convergence of (??existence) for a Gaussian momentum distribution in the wave packet. The reason is the fact that any $|\phi_a\rangle \in \mathcal{H}$ can be arbitrarily well approximated by a finite superposition of Gaussians¹³¹. For Gaussian wave packet, the wave function is bounded by

$$|\psi_G(\mathbf{r},t)| \le \frac{constant}{\left(\Delta^2 + t^2/\mu^2\right)^{3/2}}$$
 (12.13)

where μ is the mass of the particle (or the reduced mass for a system of two particles), and Δ is the width of the wave packet. If we assume that the potential is square integrable, it follows that

$$||V|\phi_G(t) > ||^2 \le \left(\int d^3 \mathbf{x} |V|^2\right) \times Sup_{(\mathbf{x})} ||\phi_G(t) > ||^2 = \frac{constant}{(c_2 + t^2)^{3/2}}$$
(12.14)

Thus, we proved that for a square integrable potential, i.e $V(\mathbf{x}) \in \mathcal{L}^2(\mathcal{R}^3)$, the M \emptyset ller Operator exists. It can actually be shown [?] that for a potential, V(r), with a range R, the operator Ω_+ exists

$$|V(r)| \le \frac{constant}{r^{1+\epsilon}}; \quad \epsilon > 0 \tag{12.15}$$

for $r \geq R$, i.e V(r) falls off for larger r faster than the Coulomb potential.

12.2 Properties of the Møller Operator

• Let us consider $\lim ||W(t)|\phi > || = ||\Omega_+|\phi > ||$ $t \to -\infty$. Since W(t) is unitary, it follows that $||\Omega_+|\phi > || = |||\phi > ||$, or equivalently

$$<\phi|\phi> = <\Omega_+\phi|\Omega_+\phi> = <\phi|\Omega_+^{\dagger}\Omega_+|\phi>$$
(12.16)

¹³¹The proof of this statement is given in the appendix

Since $|\phi\rangle$ is an arbitrary state, then we deduce that

$$\Omega_{+}^{\dagger}\Omega_{+} = 1 \tag{12.17}$$

This shows that Ω_+ is an isometry operator, i.e, it preserves the norm; however, it is not a unitary operator, because the left inverse does not exist in general. Only in a finite dimensional space an isometry is necessary unitary. In other words, an isometry can be a unitary mapping from the Hilbert space \mathcal{H} onto a subspace of \mathcal{H} .

• An important property of the Møller Operator is the inter-wining relation:

$$H\Omega_{+}(H, H_0) = \Omega_{+}(H, H_0)H_0 \tag{12.18}$$

To prove this relation consider

$$U(\tau)\Omega_{+}(H,H_{0}) = \lim U(\tau)U^{\dagger}(t)U_{0}(t)$$

$$t \to -\infty$$
(12.19)

$$= \lim U^{\dagger}(t-\tau)U_{0}(t-\tau)U_{0}(\tau) \qquad (12.20)$$

$$= \lim_{t' \to -\infty} U^{\dagger}(t')U_0(t')U_0(\tau) \qquad (12.21)$$

$$= \Omega_{H,H_0} U_0(\tau) \tag{12.22}$$

where $t' = t - \tau$. If we differentiate this result with respect to τ and set $\tau = 0$, we obtain the inter-wining relation in Eq (??inter).

=

Another way to see the inter-wining relation is to re-write it using the fact that $\Omega^{\dagger}_{+} = 1$ as follow

$$\Omega_{+}^{\dagger}H\Omega_{+} = H_0 \tag{12.23}$$

This means that Ω_+ can not be unitary unless H has no bound state, in which case H and H_0 would have the same spectrum.

As a consequence of the inter-wining relation we have $U(t)\Omega_+|\phi_a(0)\rangle = \Omega_+U_0(t)|\phi_a(0)\rangle$, or equivalently

$$|\psi_a^{(+)}(t)\rangle = \Omega_+ |\phi_a(t)\rangle$$
 (12.24)

which shows that the wave operators relate the free and the interacting states at all time, not just the initial time.

Another important result that follows from the inter-wining relation is that the ranges of the wave operators are orthogonal to the bound states of \hat{H} . To see this let $|E_a\rangle$ be a bound state of \hat{H} with energy $|E_0\rangle$ be an energy eigenstate of \hat{H}_0 , with energy $E_0 \geq 0$. It follows that

$$0 = \langle E_b | \left(\hat{H} - \hat{H}_0 \Omega_+ \right) | E_0 \rangle$$

= $(E_0 - E_b) \langle E_b | \Omega_+ | E_0 \rangle$ (12.25)

Since $(E_0 - E_b)$ is positive we conclude that

$$\langle E_b | \Omega_+ | E_0 \rangle = 0$$
 (12.26)

This means that the scattering eigenstates of \hat{H} are necessarily orthogonal to the bound states. Thus, if we define \mathcal{H}_+ to be the subspace of \mathcal{H} spanned by the states of the form $\Omega_+ | \phi \rangle$; $| \phi \rangle \in \mathcal{H}_0$, and \mathcal{H}_b to be the subspace spanned by the bound states of the Hamiltonian \hat{H} , then $\mathcal{H} = \mathcal{H}_b \oplus \mathcal{H}_+$

• The next property of the wave operator is the chain rule. Let three Hamiltonian operators, \mathcal{H}_1 , \mathcal{H}_2 and \mathcal{H}_3 , such that non of them has a bound state. If we assume that $\Omega_+(\mathcal{H}_1, \mathcal{H}_2)$ and $\Omega_+(\mathcal{H}_2, \mathcal{H}_3)$ exist, the $\Omega_+(\mathcal{H}_1, \mathcal{H}_3)$ also exists and is equal to

$$\Omega_{+}(\mathcal{H}_{1},\mathcal{H}_{3}) = \Omega_{+}(\mathcal{H}_{1},\mathcal{H}_{2})\Omega_{+}(\mathcal{H}_{2},\mathcal{H}_{3})$$
(12.27)

To prove this relation we note that

$$\begin{aligned} &||\{U_{1}^{+}U_{3} - \Omega_{+}(H_{1}, H_{2})\Omega_{+}(H_{2}, H_{3})\}|\phi > || \\ &= ||\{U_{1}^{+}U_{2}U_{2}^{+}U_{3} - \Omega_{+}(H_{1}, H_{2})\Omega_{+}(H_{2}, H_{3})\}|\phi > || \\ &= ||\{U_{1}^{+}U_{2}\left[U_{2}^{+}U_{3} - \Omega_{+}(H_{2}, H_{3}) + \Omega_{+}(H_{2}, H_{3})\right] - \Omega_{+}(H_{1}, H_{2})\Omega_{+}(H_{2}, H_{3})\}|\phi > || \\ &\leq ||\{U_{2}^{+}U_{3} - \Omega_{+}(H_{2}, H_{3})\}|\phi > || + \\ &||\{U_{1}^{+}U_{2} - \Omega_{+}(H_{1}, H_{2})\Omega_{+}(H_{2}, H_{3})\}|\phi > || \end{aligned}$$

However, in the limit $t \to -\infty$, the right-hand side vanishes provided that $\Omega_+(H_1, H_2)$ and $\Omega_+(H_2, H_3)$ exist. In this case we obtain the chain rule relation in Eq(?? chain).

In deriving Eq(??chain) we assumed that there are no bound states. When there are bound states the chain rule holds if we generalize the definition of the Møller Operator as

$$\Omega_{+}(H_i, H_j) = \lim U_i^{\dagger} U_j \mathcal{P}_j^{(c)}$$
(12.29)

$$t \to -\infty \tag{12.30}$$

where $U_k \equiv \exp \frac{iH_k t}{\hbar}$ and $\mathcal{P}_j^{(c)}$ is the projector operator on the subspace orthogonal to the bound states \hat{H}_j .

As a simple application of the chain rule, consider $\hat{H}_3 = \hat{H}_1 \equiv \hat{H}_0$ and $\hat{H}_2 = \hat{H}$. Then, it follows that

$$\Omega_{+}(\mathcal{H}_{0},\mathcal{H})\Omega_{+}(\mathcal{H},\mathcal{H}_{0}) = I \tag{12.31}$$

which implies that the subspace \mathcal{H}_+ is perpendicular to \mathcal{H}_b . Now, if we consider $\hat{H}_3 = \hat{H}_1 \equiv \hat{H}$ and $\hat{H}_2 = \hat{H}_0$, then the chain rule gives

$$\Omega_{+}(\mathcal{H},\mathcal{H}_{0})\Omega_{+}(\mathcal{H}_{0},\mathcal{H}) = \mathcal{P}^{(c)} \equiv \mathcal{P}_{+}$$
(12.32)

Here \mathcal{P}_+ is the projector operator into the subspace $\mathcal{R}_+ \equiv Ran \ (\Omega_+) \subset \mathcal{H}$. This equation means that any vector that is orthogonal to the bound states is necessary in the range of $\Omega_+(H, H_0)$, that is $\mathcal{H}_b^{\perp} \subset \mathcal{H}_+$. Therefore, from Eq(??orth) and Eq (??sub) we deduce that $\mathcal{H}_+ = \mathcal{H}_b^{\perp}$

12.3 Møller Operator and the Resolvant of \hat{H}

Let us first prove the following statement:

$$\lim f(t) = -\lim \int_0^{-\infty} dt' \epsilon e^{\epsilon t'} f(t')$$
(12.33)

$$t \to -\infty$$
 (12.34)

The integral on the right hand side of Eq(??integral) can be written as

$$-\int_{0}^{-\infty} dt' \epsilon e^{\epsilon t'} f(t') = -\int_{0}^{-\infty} \frac{d}{dt'} \left(e^{\epsilon t'} f(t') \right)$$
(12.35)
= $f(0) + -\int_{0}^{-\infty} dt' e^{\epsilon t'} \frac{f(t')}{dt'}$

The limit and the integral in Eq(??integral) can be interchanged if both converge separately. Thus, we have

$$\lim(-\epsilon)\int_0^{-\infty} dt' e^{\epsilon t'} f(t') = f(-\infty)$$
(12.36)

$$\epsilon \to 0 \tag{12.37}$$

So this result states that the limit $t \to -\infty$ can be replaced by a limit $\epsilon \to 0$. Now we consider the operator

$$\Omega_{+} \equiv \lim W(t) = -\lim \int_{0}^{-\infty} \epsilon e^{\epsilon t} W(t) dt \qquad (12.38)$$

$$t \to -\infty$$

$$= -\lim \int_{0}^{-\infty} \epsilon e^{\epsilon t} e^{\frac{i}{\hbar} H t} e^{\frac{i}{\hbar} H_{0} t} dt$$

$$\epsilon \to 0$$

$$= -\lim \int_{0}^{-\infty} \epsilon e^{\frac{i}{\hbar} H t} e^{\frac{-i}{\hbar} (H_{0} + i\epsilon) t} dt$$

$$\epsilon \to 0 \qquad (12.39)$$

Applying Ω_+ on plane wave $|\mathbf{p}\rangle$ yields

$$\Omega_{+}|\mathbf{p}\rangle = -\lim \epsilon \int_{0}^{-\infty} dt e^{\frac{i}{\hbar}(H-E_{\mathbf{p}}-i\epsilon)t}|\mathbf{p}\rangle$$

$$\epsilon \to 0$$

$$= \lim i\epsilon \frac{\hbar}{(E_{\mathbf{p}}+i\epsilon-H)}|\mathbf{p}\rangle$$

$$\epsilon \to 0$$

$$\equiv i\epsilon \hat{G}(E_{\mathbf{p}}+i\epsilon)|\mathbf{p}\rangle \equiv \hat{G}^{(+)}(E_{\mathbf{p}})|\mathbf{p}\rangle$$

$$\epsilon \to 0$$
(12.40)

where $G(z) = \frac{1}{(z-\hat{H})}$ is the resolvant of \hat{H} . Similarly, we define $G_0(z) = \frac{1}{(z-\hat{H}_0)}$ to be the resolvant of the free Hamiltonian.

Let us consider the relation

$$\hat{G}_0(z) \left[\hat{G}_0^{-1}(z) - \hat{G}^{-1}(z) \right] \hat{G}(z) = \hat{G}_0(z) \hat{V} \hat{G}(z)$$
(12.41)

from which it follows

$$\hat{G}(z) = \hat{G}_0(z) + \hat{G}_0(z)\hat{V}\hat{G}(z)$$
(12.42)

called the Hilbert identity or the second resolvant equation. If now we consider

$$\hat{G}(z) \left[\hat{G}_0^{-1}(z) - \hat{G}^{-1}(z) \right] \hat{G}_0(z) = \hat{G}(z) \hat{V} \hat{G}_0(z)$$
(12.43)

then, we get

$$\hat{G}(z) = \hat{G}_0(z) + \hat{G}(z)\hat{V}\hat{G}_0(z)$$
(12.44)

By comparing Eq (??) and Eq (12.44), we find that

$$\hat{G}_0(z)\hat{V}\hat{G}(z) = \hat{G}_0(z)\hat{V}\hat{G}_0(z)$$
(12.45)

Now, let us apply the Møller operator on a free wave packet $|\phi\rangle$ and use Eq (12.40) to write

$$|\psi^{(+)}\rangle \equiv \Omega_{+}|\phi\rangle = \int d^{3}\mathbf{p}\Omega_{+}|\mathbf{p}\rangle \langle \mathbf{p}|\phi\rangle$$

$$= \lim \int d^{3}\mathbf{p}i\epsilon \hat{G}(z_{\mathbf{p}})|\mathbf{p}\rangle \langle \mathbf{p}|\phi\rangle$$

$$\epsilon \to 0$$

$$= \lim \int d^{3}\mathbf{p}i\epsilon \hat{G}_{0}(z_{\mathbf{p}})\hat{V}\hat{G}(z_{\mathbf{p}})|\mathbf{p}\rangle \langle \mathbf{p}|\phi\rangle$$

$$\epsilon \to 0$$

$$\epsilon \to 0$$

$$(12.46)$$

By using the relation in Eq (12.45), and the fact that

$$\lim_{\epsilon \to 0} i\epsilon \, \hat{G}_0(z_{\mathbf{p}}) |\mathbf{p}\rangle = \lim_{\epsilon \to 0} \frac{i\epsilon}{E_{\mathbf{p}} + i\epsilon - E_{\mathbf{p}}} |\mathbf{p}\rangle = |\mathbf{p}\rangle$$
(12.47)
$$\epsilon \to 0 \qquad \epsilon \to 0$$

we obtain

$$\Omega_{+} = \mathbf{1} + \lim_{\epsilon \to 0} \int d^{3} \mathbf{p}' \hat{G}_{0}(E_{\mathbf{p}'} + i\epsilon) \hat{V} \hat{G} |\mathbf{p}'\rangle \langle \mathbf{p}'| \qquad (12.48)$$

or in a short notation, I will write it as

$$\Omega_{+} = \mathbf{1} + \lim_{\epsilon \to 0} \hat{G}_{0}(E + i\epsilon) \hat{V}\hat{G}$$
(12.49)

If we define the state $|\mathbf{p}\rangle^{(+)} := \Omega_+ |\mathbf{p}\rangle$, we can write

$$|\mathbf{p}\rangle^{(+)} = |\mathbf{p}\rangle + lim \ \hat{G}_0(E_{\mathbf{p}} + i\epsilon)\hat{V}|\mathbf{p}\rangle^{(+)}$$

$$\epsilon \to 0$$
(12.50)

which is just the Lippmann-Schwinger equation in the momentum representation.

12.4 The Transition Operator

Let us multiply $\Omega_+|\phi\rangle \equiv |\psi^{(+)}\rangle$ with \hat{V} and define $|\psi^{(+)}\rangle := \hat{T}|\phi\rangle$. Then, it follows that

$$\hat{T}|\phi\rangle = \left[\hat{V} + \lim \int d^{3}\mathbf{p} \ \hat{V}\hat{G}(z_{\mathbf{p}})\hat{V}|\mathbf{p}\rangle < \mathbf{p}|\right]\phi\rangle$$

$$\epsilon \to 0$$
(12.51)

or, as operator equation we can write

$$\hat{T} = \hat{V} + \lim \int d^{3}\mathbf{p} \ \hat{V}\hat{G}(z_{\mathbf{p}})\hat{V}|\mathbf{p}\rangle \langle \mathbf{p}| \qquad (12.52)$$

$$\epsilon \to 0$$

$$\equiv \hat{V} + \lim \hat{V}\hat{G}(z_{\mathbf{p}})\hat{V}$$

$$\epsilon \to 0$$

where \hat{T} is called the T-matrix or the transition operator. When this expression is multiplied on the left by $G_0^{(+)}(z)$ and moreover relation Eq(12.42) substituted, then we find ¹³²

$$\hat{G}_0(z)\hat{T}(z) = \hat{G}(z)\hat{V}$$
 (12.54)

By substituting (12.52) into (12.53), we obtain

$$\hat{T}(z) = \hat{V} + \hat{V}\hat{G}_0(z)\hat{T}(z)$$
(12.55)

$$\hat{T}(z)\hat{G}_0(z) = \hat{V}\hat{G}(z)$$
 (12.53)

¹³²similarly, when expression (12.51) is multiplied on the right by $\hat{G}_0(z)$ and moreover relation (12.44) substituted, then we find

which is the Lippmann-Schwinger Equation for the transition operator. In the momentum representation, (12.55) reads

$$\langle \mathbf{p}'|\hat{T}|\mathbf{p}\rangle = \langle \mathbf{p}'|\hat{V}|\mathbf{p}\rangle + \int d^3k \langle \mathbf{p}'|\hat{V}|\mathbf{k}\rangle \frac{1}{E_{\mathbf{p}} + i\epsilon - \frac{k^2}{2\mu}} \langle \mathbf{k}|\hat{V}|\mathbf{p}\rangle \quad (12.56)$$

Now, let us consider the matrix element of the free-resolvant operator $\hat{G}_0(z)$ in the momentum representation:

$$<\mathbf{p}'|\hat{G}_{0}|\mathbf{p}> = \frac{\delta^{(3)}(\mathbf{p}'-\mathbf{p})}{E_{\mathbf{p}}+i\epsilon-\frac{\mathbf{p}'^{2}}{2\mu}}$$

$$= \delta^{(3)}(\mathbf{p}'-\mathbf{p})\left[\mathcal{P}\left(\frac{1}{E_{\mathbf{p}}+i\epsilon-\frac{\mathbf{p}'^{2}}{2\mu}}\right) - i\pi\delta(E_{\mathbf{p}}-\frac{\mathbf{p}'^{2}}{2\mu})\right]$$
(12.57)

where \mathcal{P} stands for the principal value¹³³. This we can write

$$\hat{G}_0(E_{\mathbf{p}} + i0^+) = \mathcal{P}\hat{G}_0(E_{\mathbf{p}}) - i\pi\delta(E_{\mathbf{p}} - \hat{H}_0)$$
(12.58)

which separates the real and the imaginary parts of the non-interacting resolvant operator. Using Eq(12.58) in the Lippmann-Schwinger expression in Eq(12.55) gives

$$\hat{T}(E+i0^{+}) = \hat{V} + \hat{V} \left[\mathcal{P}\hat{G}_{0}(E) - i\pi\delta(E-\hat{H}_{0}) \right] \hat{T}(E+i0^{+})$$
(12.59)

This is equivalent to the following pair of equations

$$\hat{K}(E) = \hat{V} + \hat{V}\mathcal{P}\hat{G}_0(E)\hat{K}(E)$$

$$\hat{T}(E+i0^+) = \hat{K}(E) - i\pi\hat{K}(E)\delta(E-\hat{H}_0)\hat{T}(E+i0^+)$$
(12.60)

The operator $\hat{K}(E)$ is called the **K** – **operator** and its elements are the K-matrix elements. The equation for the K-matrix has the advantage that the integral is real.

12.5 The Case of System with Bound States

Let us assume that the Hamiltonian of the system, \hat{H} , has a bound state $|\psi_b\rangle$ at $E = E_b$. Then the Schrödinger equation reads

$$\left(\hat{H}_0 - E_b\right)|\psi_b\rangle = -V|\psi_b\rangle$$
(12.61)

Since E_b is negative, there is no regular solution for the case V = 0 and we can write

$$|\psi_b\rangle = \frac{1}{E_b - \hat{H}_0} V |\psi_b\rangle$$
 (12.62)

¹³³By definition $\frac{1}{x+i\epsilon} = \mathcal{P}(\frac{1}{x}) - i\pi\delta(x).$

which is the homogenous Lippmann-Schwinger equation for $|\psi_b\rangle$. Evaluating the Green's function for $E = E_b < 0$, we find that the wave function for the bound state, $\psi_b(\mathbf{x}) \equiv \langle \mathbf{x} | \psi_b \rangle$ has the correct exponential behavior, namely

$$\psi_b(\mathbf{x}) \propto \int d^3 \mathbf{x}' \frac{e^{-\sqrt{2\mu E_b}(|\mathbf{x}| - |\mathbf{x}'|)}}{|\mathbf{x}| - |\mathbf{x}'|} V(\mathbf{x}') \psi_b(\mathbf{x}')$$
(12.63)

So, if the Hamiltonian has bound states and continuous spectrum, then we can write the spectral decomposition of $\hat{G}^{(\pm)}(z)$ as ¹³⁴

$$\frac{1}{E \pm i\epsilon - \hat{H}} = \sum_{b} |\psi_{b}\rangle = \frac{1}{E \pm i\epsilon - \hat{H}} < \psi_{b}| + \int d^{3}\mathbf{p} |\psi_{\mathbf{p}}\rangle^{(\pm)} = \frac{1}{E \pm i\epsilon - \hat{H}} < \langle \mathbf{M}_{\mathbf{p}} \rangle |64\rangle$$

where the asymptotic state $|\psi\rangle^{(-)}$ is defined as

$$\lim ||U(t)|\psi_a^{(-)}(0) > -U_0(t)|\phi_a(0) > || \to 0$$
(12.65)

$$t \to +\infty \tag{12.66}$$

Note that the resolvant $\hat{G}(z)$ has poles at the binding energies E_b . Inserting the spectral decomposition of $\hat{G}(z)$ into Eq(12.52), we obtain

$$<\mathbf{p}'|\hat{T}|\mathbf{p}> = <\mathbf{p}'|\hat{V}|\mathbf{p}> + \sum_{b} \frac{<\mathbf{p}'|\hat{V}|\psi_{b}><\psi_{b}|\hat{V}|\mathbf{p}>}{E-E_{b}}$$
(12.67)
$$+ \int d^{3}\mathbf{k} \frac{<\mathbf{p}'|\hat{T}^{(+)}|\mathbf{k}><\mathbf{k}|\hat{T}^{(-)}|\mathbf{p}>}{E\pm i\epsilon - E_{\mathbf{k}}}$$

If E is positive, then there are no bound states and the discrete sum over b vanishes. On the other hand, if E < 0 is allowed, then $\langle \mathbf{p}' | \hat{T} | \mathbf{p} \rangle$ has poles at $E = E_b$, and the residue is separable¹³⁵. The residue can be written as:

$$<\mathbf{p}'|\hat{V}|\psi_b> = <\mathbf{p}'|(\hat{H}-\hat{H}_0)|\psi_b> = (E-E_{\mathbf{p}})<\mathbf{p}'|\psi_b>$$
(12.68)
= $(E-E_{\mathbf{p}})\psi_b(\mathbf{p}')$

where $\psi_b(\mathbf{p}')$ is the wave function corresponding to the energy $E = E_b$. Thus, if E < 0 is close to E_b , then the *T*-matrix is dominated by pole term and the residue is characterized by the corresponding bound state wave function. So the matrix element $\langle \mathbf{p}' | \hat{T} | \mathbf{p} \rangle$ has poles for the bound states and a cut for the continuous spectrum.

12.6 Unitarity Relations

Let us consider the difference between $\hat{T}^{(+)}$ and $\hat{T}^{(-)}$ as obtained from Eq (??*Tplus*) under the assumption that \hat{V} is hermetian:

$$<\mathbf{p}'|\left[\hat{T}^{(+)}(E) - \hat{T}^{(-)}(E)\right]|\mathbf{p}> = \int d^3\mathbf{k} < \mathbf{p}'|\hat{T}^{(+)}(E)|\mathbf{k}> \left[\frac{1}{E - E_{\mathbf{k}} + i\epsilon} - c(\mathbf{d})\right] .69)$$

¹³⁴The superscript \pm refers to $\pm i\epsilon$ in $\hat{G}(E \pm i\epsilon)$.

¹³⁵A function $f(q_1, q_2)$ is said separable if $f(q_1, q_2) = f_1(q_1)f_2(q_2)$.

Using the Cauchy principal value, we obtain

$$<\mathbf{p}'|\left[\hat{T}^{(+)}(E) - \hat{T}^{(-)}(E)\right]|\mathbf{p}> = -2\pi i \int d^3\mathbf{k} < \mathbf{p}'|\hat{T}^{(+)}(E_{\mathbf{k}})|\mathbf{k}>\delta(E-E_{\mathbf{k}}) < \mathbf{k}|\hat{T}^{(+)}(E_{\mathbf{k}})|\mathbf{p}.50\rangle$$

or, equivalently,

$$<\mathbf{p}'|\Big[\hat{T}^{(+)}(E) - \hat{T}^{(-)}(E)\Big]|\mathbf{p}> = -2i\pi\mu\int dkk\int d\Omega_k <\mathbf{p}'|\hat{T}^{(+)}(E_{\mathbf{k}})|\mathbf{k}>\delta(p-k) < \mathbf{k}|\hat{T}^{(+)}(E_{\mathbf{k}})|\mathbf{p}|.$$

where in the last relation we used $\delta(ax) = \delta(x)/a$, $\delta(f(x)) = \sum_i \delta(x-a_i)/f'(a_i)$, where a_i 's are the roots of the equation f(x) = 0, and $E = p^2/2\mu$. Eq (12.71) is called the **off-shell unitary relation** since in general $|\mathbf{p}| \neq |\mathbf{p}'| \neq |\mathbf{k}|$. For elastic scattering, i.e, $|\mathbf{p}| = |\mathbf{p}'|$, Eq (??off-shell-unit) simplifies to

$$<\mathbf{p}'|\left[\hat{T}^{(+)}(E) - \hat{T}^{(-)}(E)\right]|\mathbf{p}> = -2i\pi\mu p \int d\Omega_k <\mathbf{p}'|\hat{T}^{(+)}(E_{\mathbf{p}})|\mathbf{k}> <\mathbf{k}|\hat{T}^{(+)}(E_{\mathbf{p}})|\mathbf{k}> <\mathbf{k}|\hat{T}^{(+)}(E_{\mathbf{p}})|\mathbf{k}>$$

which is called the **on-shell unitary relation**. If we restrict our selves to the forward direction, i.e, $\mathbf{p}' = \mathbf{p}$, or equivalently $\theta = 0$, with θ being the angle between \mathbf{p}' and \mathbf{p} we obtain

$$\mathcal{I}m < \mathbf{p}|\hat{T}^{(+)}(E)|\mathbf{p}\rangle = -\pi\mu p \int d\Omega_k |<\mathbf{p}|\hat{T}^{(+)}(E)|\mathbf{k}\rangle|^2$$
(12.73)

This is referred as the **optical theorem**, which is a non-linear relation between the imaginary part of \hat{T} in the forward direction and its absolute value integrated over all angles.

12.7 The \hat{S} Operator

In order to completely characterize the scattering process, we need to have the behavior of $|\psi_a\rangle^{(+)}$ for large positive t, i.e,

$$|\psi_a(t)\rangle^{(+)} = e^{-i\hat{H}t} |\psi_a\rangle^{(+)} = e^{-i\hat{H}t} \Omega_+ |\phi_a\rangle$$
 (12.74)

Long after the scattering, the state $|\psi_a(t)\rangle^{(+)}$ should behave as a free state. Thus, we shall consider the amplitude $\langle \phi_b(t)\psi_a(t)\rangle^{(+)}$ for large time t. Here $\phi_b(t)\rangle^{(+)}$ is a free state. We define

$$S_{ab} := \lim_{a \to \infty} \langle \phi_b(t) | \psi_a(t) \rangle^{(+)}$$

$$t \to -\infty$$

$$= \lim_{a \to \infty} \langle e^{-i\hat{H}_0 t} \phi_b | e^{-i\hat{H}t} \psi_a \rangle^{(+)}$$

$$t \to -\infty$$

$$= \lim_{a \to \infty} \langle e^{i\hat{H}t} e^{-i\hat{H}_0 t} \phi_b | \psi_a \rangle^{(+)}$$

$$t \to -\infty$$

$$\equiv \lim_{a \to \infty} \langle \Omega_- \phi_b | \psi_a \rangle^{(+)} \equiv^{(-)} \langle \psi_b | \psi_a \rangle^{(+)}$$
(12.75)
(12.75)
where we defined

$$\Omega_{-} := lime^{i\hat{H}t}e^{-i\hat{H}_{0}t}$$

$$t \to -\infty$$
(12.76)

and

$$|\psi_b >^{(-)} := \Omega_- |\phi_b >$$
 (12.77)

which characterizes the outgoing state. The existence of the operator Ω_{-} can be shown in a similar way as we proved it for Ω_{+} . Thus the probability transition from an incoming state is given by

$$\mathcal{P}_{a\to b} = |S_{ab}|^2 = |^{(-)} < \psi_b |\psi_a >^{(+)}|^2 = |^{(-)} < \phi_b |\hat{S}|\phi_a >^{(+)}|^2$$
(12.78)

where the \hat{S} operator is defined by

$$\hat{S} := \Omega_{-}^{\dagger} \Omega_{+} \tag{12.79}$$

The existence of Ω_{-}^{\dagger} and Ω_{+} means that a scattering state $|\psi\rangle = \Omega_{+}|\phi_{in}\rangle = \Omega_{-}|\phi_{out}\rangle$, from which it follows that

$$|\phi_{out}\rangle = \hat{S}|\phi_{in}\rangle \tag{12.80}$$

The scattering operator \hat{S} contains all informations of experimental interest about scattering process. In fact, we will later find that, through the "magic" of analytic continuation, that the *S*-matrix contains all the information about the bound states as well. The scattering operator \hat{S} has the following properties:

• \hat{S} is unitary :

Let us consider

$$S^{+}S = \Omega_{+}^{\dagger}\Omega_{-}\Omega_{-}^{\dagger}\Omega_{+} = \Omega_{+}^{\dagger}\mathcal{P}_{-}\Omega_{+}$$

$$S^{+}S = \Omega_{-}^{\dagger}\Omega_{+}\Omega_{+}^{\dagger}\Omega_{-} = \Omega_{-}^{\dagger}\mathcal{P}_{+}\Omega_{-}$$
(12.81)

where \mathcal{P}_{\pm} are the projectors into the subspaces $\mathcal{R}_{\pm} \in \mathcal{H}$. If Ω_{\pm} have the same range, i.e, $\mathcal{P}_{+} = \mathcal{P}_{-} = \mathcal{P}^{(c)}$, then

$$\mathcal{P}_{-}\Omega_{(+)} = \Omega_{(+)}$$

$$\mathcal{P}_{+}\Omega_{(-)} = \Omega_{(-)}$$
(12.82)

which leads to

$$\hat{S}^{\dagger}\hat{S} = \hat{S}\hat{S}^{\dagger} = I \tag{12.83}$$

Now consider the probability of scattering into all possible final states

$$\int d^{3}\mathbf{p} |\langle \mathbf{p} | \hat{S} | \phi_{a} \rangle |^{2} = \int d^{3}\mathbf{p} \langle \mathbf{p} | \hat{S} | \phi_{a} \rangle^{*} \langle \mathbf{p} | \hat{S} | \phi_{a} \rangle \qquad (12.84)$$
$$= \int d^{3}\mathbf{p} \langle \phi_{a} | \hat{S}^{\dagger} | \mathbf{p} \rangle \langle \mathbf{p} | \hat{S} | \phi_{a} \rangle$$
$$= \langle \phi_{a} | \hat{S}^{\dagger} \hat{S} | \phi_{a} \rangle = || \phi_{a} \rangle ||$$

Thus, the unitarity of \hat{S} is just the conservation of probability, i.e, what goes in, must go out.

• \hat{S}^{\dagger} commutes with \hat{H}_0 :

Let us consider

$$\hat{S}\hat{H}_{0} = \Omega^{\dagger}_{(-)}\Omega_{(+)}\hat{H}_{0} = \Omega^{\dagger}_{(-)}\hat{H}\Omega_{(+)}$$

$$= \hat{H}_{0}\Omega^{\dagger}_{(-)}\Omega_{(+)} = \hat{H}\hat{S}$$
(12.85)

which shows that $[\hat{S}, \hat{H}_0] = 0$. Consequently, we have

$$\hat{H}_0 \hat{S} |\mathbf{p}\rangle = E_{\mathbf{p}} \hat{S} |\mathbf{p}\rangle \tag{12.86}$$

which means that $|\mathbf{p}\rangle$ as well $\hat{S}|\mathbf{p}\rangle$ are eigenstates of \hat{H}_0 and with the same energy. As consequence of this property, the expectation value of the free Hamiltonian between the post-scattering (final) and the pre-scattering (initial) states are equal.

 \hat{H}_0 has a complete spectrum of non-renormalizable eigenstates $|\mathbf{p}\rangle$, and thus \hat{S} can be represented in these states., and we have

$$0 = \langle \mathbf{p}' | \left[\hat{H}_0, \hat{S} \right] | \mathbf{p} \rangle = \langle \mathbf{p}' | \hat{H}_0 \hat{S} | \mathbf{p} \rangle - \langle \mathbf{p}' | \hat{S} \hat{H}_0 | \mathbf{p} \rangle$$

$$= (E_{\mathbf{p}'} - E_{\mathbf{p}}) \langle \mathbf{p}' | \hat{S} | \mathbf{p} \rangle$$
(12.87)

from which it follows that

$$\langle \mathbf{p}' | \hat{S} | \mathbf{p} \rangle \neq 0; \qquad iff \ E_{\mathbf{p}'} = E_{\mathbf{p}}$$

$$(12.88)$$

which means that $S_{\mathbf{p'p}}$ is defined off-shell.

12.8 Relating the S- Matrix and the T-Matrix

The S-matrix element $\langle \psi_b^- | \psi_a^+ \rangle$ can be written as

$$S_{ba} = \langle \phi_b | \psi_a^{(+)} \rangle + \langle \hat{G}(E_b - i\epsilon) \hat{V} \phi_b | \psi_a^{(+)} \rangle$$

$$= \langle \phi_b | \psi_a^{(+)} \rangle + \langle \phi_b | \hat{V} \hat{G}(E_b - i\epsilon) | \psi_a^{(+)} \rangle$$

$$= \langle \phi_b | \psi_a^{(+)} \rangle + \frac{1}{E_b - E_a + i\epsilon} \langle \phi_b | \hat{V} | \psi_a^{(+)} \rangle$$
(12.89)

where we used the fact that

$$\hat{G}^{\dagger}(E_{b} - i\epsilon)|\psi_{a}^{(+)}\rangle = \left(\frac{1}{E_{b} - i\epsilon - \hat{H}}\right)^{\dagger}|\psi_{a}^{(+)}\rangle$$

$$= \frac{1}{E_{b} - E_{a} - i\epsilon}|\psi_{a}^{(+)}\rangle$$
(12.90)

By using Eq.12.50, we can write

$$<\phi_{b}|\psi_{a}^{(+)}> = \delta_{ab} + <\phi_{b}|\hat{G}_{0}(E_{b} - i\epsilon)\hat{V}|\psi_{a}^{(+)}>$$

$$= \delta_{ab} + \frac{1}{E_{b} - E_{a} - i\epsilon} <\phi_{b}|\hat{V}|\psi_{a}^{(+)}>$$
(12.91)

Thus,

$$S_{ba} = \delta_{ab} + \left[\frac{1}{E_b - E_a + i\epsilon} - \frac{1}{E_b - E_a - i\epsilon}\right] < \phi_b |\hat{V}|\psi_a^{(+)} >$$
(12.92)

and which, after using the Cauchy principal value, becomes

$$S_{ba} = \delta_{ab} - 2\pi \tau_{ba}^{(+)} \tag{12.93}$$

where

$$\tau_{ba}^{(+)} = \delta \left(E_a - E_b \right) < \phi_b |\hat{T}| \psi_a^{(+)} >$$

$$= \delta \left(E_a - E_b \right) T_{ba}^{(+)}$$
(12.94)

The operator $\tau^{(+)}$ describes the scattering process, and hence will appear in the differential cross-section.

The transition matrix $T^{(+)}$ can be expressed in terms of the K-operator (see Eq. 12.60) as

$$T^{(+)}(E) = \frac{1}{1 + i\pi\delta(E - \hat{H}_0)\hat{K}(E)}$$
(12.95)

Inserting the above expression of \hat{T} in Eq. 12.93, we find ¹³⁶

$$\hat{S} = \frac{1 - i\pi \hat{K}(E)\delta(E - H_0)}{1 + i\pi \hat{K}(E)\delta(E - H_0)}$$
(12.96)

which shows that the scattering operator can be obtained directly from the K-matrix. We also note that, since \hat{K} is hermitian, the scattering operator is unitary.

12.9 A heuristic Derivation of the Cross Section

Suppose that the system is in a box with periodic boundary conditions. For a cubic box of size L, the momentum operator has eigenstates¹³⁷

$$\langle \mathbf{x} | \mathbf{p} \rangle = \frac{1}{\sqrt{V}} e^{i\mathbf{p}.\mathbf{x}}; \qquad \mathbf{p} = \frac{2\pi}{L} \left(n_x, n_y, n_z \right)$$
(12.97)

If the system is in a box, the S-matrix **does not exist**, because we can not define asymptotic state. Instead, we define

$$S_T := e^{i\hat{H}_0 \frac{T}{2}} e^{-i\hat{H} \frac{T}{2}} e^{-i\hat{H} \frac{T}{2}} e^{i\hat{H}_0 \frac{T}{2}}$$
(12.98)

and take $T \to \infty$ limit as we the let the volume V goes to infinity.

¹³⁶This expression of \hat{S} shows that the scattering operator is a Cayley transform of $\hat{K}(E)\delta(E-H_0)$. ¹³⁷with the conventional discrete normalization:....

Consider the scattering in the non-forward direction, i.e, $\mathbf{p} \neq \mathbf{p}$:

$$<\mathbf{p}'|\left(\hat{S}_T-1\right)|\mathbf{p}>=-2\pi i\delta_T\left(E'_{\mathbf{p}}-E_{\mathbf{p}}\right)T^{(Box)}\left(\mathbf{p}',\mathbf{p}\right)\times\frac{\left(2\pi\right)^3}{V}$$
(12.99)

Here $\frac{(2\pi)^3}{V}$ is a normalization factor, inserted so that

$$\lim T^{(Box)}(\mathbf{p}', \mathbf{p}) \to T^{(Box)}(\mathbf{p}', \mathbf{p}) \equiv <\mathbf{p}'|\hat{T}^{(+)}(E)|\mathbf{p}>$$
(12.100)
$$V, T \to \infty$$

and δ_T is an approximation of the δ -function with a width of order 1/T, i.e.,

$$\delta_T \left(E'_{\mathbf{p}} - E_{\mathbf{p}} \right) = \frac{1}{2\pi} \int_{-\frac{T}{2}}^{\frac{T}{2}} dt e^{i \left(E_{\mathbf{p}'} - E_{\mathbf{p}} \right)} = \frac{1}{\pi} \frac{\sin \left(E_{\mathbf{p}'} - E_{\mathbf{p}} \right) \frac{T}{2}}{\left(E'_{\mathbf{p}} - E_{\mathbf{p}} \right)}$$
(12.101)

The transition probability is given by

$$\mathcal{P}_{\mathbf{p}',\mathbf{p}} = \frac{(2\pi)^8}{V^2} \cdot \left[\delta_T \left(E_{\mathbf{p}'} - E_{\mathbf{p}} \right) \right]^2 \cdot |T^{(Box)} \left(\mathbf{p}', \mathbf{p} \right)|^2$$
(12.102)

As $T \to \infty$, $\left[\delta_T \left(E_{\mathbf{p}'} - E_{\mathbf{p}}\right)\right]^2 \to \frac{T}{2\pi} \delta \left(E_{p'} - E_p\right)$. and hence

$$\mathcal{P}_{\mathbf{p}',\mathbf{p}} = \frac{(2\pi)^{7}}{V^{2}} \frac{T}{\bar{p}} \mu \,\,\delta\left(p'-p\right) |T\left(\mathbf{p}',\mathbf{p}\right)|^{2} \tag{12.103}$$

Recall that the density of states in momentum space is $dn = \frac{V}{(2\pi)^3} d^3 \mathbf{p}' = \frac{V}{(2\pi)^3} p'^2 dp' d\Omega$, and thus, the probability of transition to a state in solid angle $d\Omega$ is

$$\mathcal{P}(\bar{p}) = \int \mathcal{P}_{\mathbf{p}',\mathbf{p}} dn = \frac{(2\pi)^7}{V^2} \frac{T}{\bar{p}} \mu \int dp' {p'}^2 \delta\left(p'-p\right) \frac{V}{(2\pi)^3} d\Omega |T\left(\mathbf{p}',\mathbf{p}\right)|^2 (12.104)$$
$$= (2\pi)^4 \cdot \frac{T}{V} \cdot \bar{p} \cdot \mu |T\left(\mathbf{p}',\mathbf{p}\right)|^2 d\Omega$$

where $\mathbf{p}' = \mathbf{p}$. The transition rate is given by

$$\Gamma = \frac{\mathcal{P}(\bar{p})}{T} = (2\pi)^4 \cdot \frac{\bar{p}}{V} \cdot \mu |T\left(\mathbf{p}',\mathbf{p}\right)|^2 d\Omega \qquad (12.105)$$

To find the cross section, we must divide the transition rate by the incident flux associated with the state $|\mathbf{p}\rangle \simeq |\mathbf{\bar{p}}\rangle$, given by

$$j_{inc} = \frac{\bar{p}/\mu}{V} \tag{12.106}$$

Thus, the differential cross-section reads

$$\frac{d\sigma}{d\Omega} = |(2\pi)^2 \mu \ T\left(\mathbf{p}',\mathbf{p}\right)|^2 \equiv |f\left(\mathbf{p}',\mathbf{p}\right)|^2$$
(12.107)

13 Interpretation of Quantum Mechanics *

- 13.1 Mixed States and the Density Matrix
- 13.2 Measurement and Interpretation
- 13.3 Schrodinger's Cat
- 13.4 Einstein-Podolsky-Rosen (EPR) Argument

14 Exercices

• Exercise 1

The wave function of the bound state of a particle of mass m in the onedimensional attractive delta-function potential $V(x) = -\alpha \ \delta(x)$ can be written as $\phi(x) = N \ e^{-a|x|}$, where N is the normalization constant.

(a) Find the constant a and the energy eigenvalue E in terms of α and m.

(b) Find the variances in position and momentum in terms of α and m. Compare the product $\Delta x \ \Delta p$ to the reduced Planck constant.

• Exercise 2

- Exercise 3
- Exercise 4
- Exercise 5

Consider a particle of mass moving in one dimension starting from far negative x-axis toward the positive direction in a triple-delta potential:

$$U(x) = \alpha \left[\delta(x-a) + \delta(x-b) + \delta(x-c) \right]$$
(14.1)

where α , a, b, and c are real constants. Assume that the particle has energy $E >> m\alpha/2\hbar^2$.

What are the relative positions of the potential spikes (a, b, c) that maximize the reflection coefficients.

- 15 Answers
 - Exercise 1 (a)

(b)

- Exercise 2
- Exercise 3
- Exercise 4
- Exercise 5

The relative positions of the potential spikes (a, b, c) that maximize the reflection coefficient are:

$$(b-a) = \frac{2\pi n}{k}; \qquad (c-a) = \frac{2\pi m}{k}$$

where n, and m are integers.

16 Exams with Solutions

United Arab Emirates University

Midterm Exam

Quantum Physics (PHYS 525)

Semester: Fall 2016

Date: 10/25/2016 Time: 3:00 - 4:00 pm

Student Name/ID:

Abstract:

- 1. This exam consists of of **Two Problems**.
- 2. There is one **BONUS** question in Problem 2.
- 3. You need to solve both problems.
- 4. Please **SHOW YOUR WORK**.

I am not telling you it is going to be easy – I am telling you it is going to be worth it !.

Art Williams

Problem One [10 points] (Measurement of Spin-1/2 Particle)

Consider a spin 1/2 particle (e.g. electron), with a magnetic moment. At time t = 0, the state of the particle os $|\psi(t = 0)\rangle = |+\rangle$.

1. If the observable S_x is measured at time t = 0, what are the possible results and the probabilities of those results ?

2. If instead of performing the above measurement, the system is allowed to evolve in a uniform magnetic field $\vec{B} = B_0 \hat{y}$. Calculate the state of the system at time t = T using the S_z basis¹³⁸.

3. Suppose at time t = T, the observable S_x is measured, what is the probability that a value $+\hbar/2$ will be found ?

¹³⁸i.e. express your answer in terms of the states |+> and |->.

Answer Problem 1

1. The eigenvalues of the operator \hat{S}_x are $+\hbar/2$ and $-\hbar/2$ with the corresponding eigenstates $|+\rangle_x = (|+\rangle + |-\rangle)/\sqrt{2}$ and $|-\rangle_x = (|+\rangle - |-\rangle)/\sqrt{2}$, respectively¹³⁹. So the outcome of the measurement of the observable S_x will be either $+\hbar/2$ or $-\hbar/2$. The probability for each measurement is

$$\mathcal{P}^{(+\hbar/2)} = |_{x} < +|+>|^{2} = \frac{1}{2}$$
$$\mathcal{P}^{(-\hbar/2)} = |_{x} < -|+>|^{2} = \frac{1}{2}$$

2. The Hamiltonian of this system is

$$H = \vec{\mu}.\vec{B} = \frac{e}{m}\vec{S}.\vec{B} \equiv \omega_0 \hat{S}_y$$

where $\omega_0 = eB_0/m$ and which has the unit of frequency. So the eigenstates of the above Hamiltonian are $|+\rangle_y$ and $|-\rangle_y$ which can be expressed in terms of the states $|\pm\rangle = as^{140}$

$$|+>_{y} = \frac{1}{\sqrt{2}} (|+>+i|->)$$
$$|->_{y} = \frac{1}{\sqrt{2}} (|+>-i|->)$$

From the the expressions above, we infer that the state $|+\rangle$ reads in the basis $|\pm\rangle_y$ (which are the eigenstates of the Hamiltonian) as

$$|+>=\frac{1}{\sqrt{2}}(|+>_{y}+|->_{y})$$

¹³⁹This can be shown easily from the definition $S_x = \frac{\hbar}{2}\hat{\sigma}_x$, where $\hat{\sigma}_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$. ¹⁴⁰You can show this from the definition $S_y = \frac{\hbar}{2}\hat{\sigma}_y$, where $\hat{\sigma}_y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}$.

Thus, at the instant T, the state $|\psi(t=0)\rangle$ evolves into the state

$$|\psi(T)\rangle = \frac{1}{\sqrt{2}}e^{-i\frac{\omega_0}{2}T}(|+\rangle_y + \frac{1}{\sqrt{2}}e^{+i\frac{\omega_0}{2}T}|-\rangle_y$$
$$= \cos\left(\frac{\omega_0}{2}T\right)(|+\rangle_+ \sin\left(\frac{\omega_0}{2}T\right)|-\rangle$$

3. The idea is similar to question 1, that is the probability is given by the square of the component of $|\psi(T)\rangle$ along the state $|+\rangle_x$, i.e

$$\mathcal{P}^{(+\hbar/2)}(T) = |_{x} < +|\psi(T) > |^{2}$$

= $|\frac{1}{\sqrt{2}} (<+|+<-|) \left(\cos\left(\frac{\omega_{0}}{2}T\right)(|+>_{+}\sin\left(\frac{\omega_{0}}{2}T\right)|->\right)|^{2}$
= $\frac{1}{2} \left(1 + 2\cos\left(\frac{\omega_{0}}{2}T\right)\sin\left(\frac{\omega_{0}}{2}T\right)\right)$
= $\frac{1}{2} (1 + \sin(\omega_{0}T))$

Problem Two [20 points] (Harmonic Oscillators from A to Z)

We have seen in class that the Hamiltonian of a particle in a onedimensional simple Harmonic oscillator potential can be expressed in term of the creation and annihilation operators \hat{a}^{\dagger} and \hat{a} , respectively, as:

$$\hat{H}_0 = \hbar\omega(a^{\dagger}a + \frac{1}{2})$$

with $\{|n\rangle, n=0,1,..\}$ are the n^{th} eigenstates of the above Hamiltonian.

• Part A

A.1. Show that the energy levels of a simple harmonic oscillator are given by

$$E_n = \hbar\omega(n + \frac{1}{2})$$

A.2. Calculate the expectation value of the operators \hat{x} in a state $|n\rangle$.

A.3. Calculate the expectation value of \hat{x}^2 and \hat{p}^2 in a state $|n\rangle$.

A.4. Show that the expectation value of \hat{H}_0 in an arbitrary normalized state $|\psi\rangle$ is always greater or equal than the energy of the ground state.

A.5. A coherent state is defined as

$$|z>=e^{-\frac{|z|^2}{2}}\sum_{n=0}^{\infty}rac{z^n}{\sqrt{n!}}\;|n>$$

where z is a complex number. Show that $|z\rangle$ is an eigenstate of the annihilation operator and find its corresponding eigenvalue.

• <u>Part B</u>

Now we consider two independent simple harmonic oscillators with the same frequency ω . This is equivalent of a one simple harmonic oscillator in two dimensions¹⁴¹, described by the Hamiltonian

$$\hat{\mathcal{H}} = \frac{\hat{p}_x^2}{2} + \frac{m\omega^2}{2}x^2 + \frac{\hat{p}_y^2}{2} + \frac{m\omega^2}{2}y^2$$

B.1. What are the energies of the first three lowest-lying states?

B.2. Make a table where you show the quantum numbers (n_x, n_y) and the degeneracy for the first five lowest-lying states? Can you **deduce** the degeneracy of state $|n_x, n_y >$?

Bonus Question : [2 points]

For a three dimensional harmonic oscillator, derive the degeneracy of state of energy $|n_x, n_y, n_z > ?$

¹⁴¹It is also known as the isotropic 2-dimensional simple harmonic oscillator.

Answer Problem 2

• Part A

A.1. By just applying the definition of how the annihilation and the creation operators act on a state $|n\rangle$, we get

$$\hat{H}|n\rangle = \frac{\hbar}{2} \left(\hat{a}^{\dagger} \sqrt{n} |n-1\rangle + |n\rangle \right)$$
$$= \frac{\hbar}{2} \left(\sqrt{n} \sqrt{n} |n\rangle + |n\rangle \right) = \frac{\hbar}{2} \left(\sqrt{n} + 1 \right) |n\rangle$$

Thus, the energy levels of the a simple harmonic are given by

$$\mathbf{E_n} = \hbar\omega(n+\frac{1}{2})$$

A.2. the expectation of the position operator is

$$\langle \hat{x} \rangle = \langle n|\hat{x}|n \rangle = \sqrt{\frac{\hbar}{2m\omega}} \left(\langle n|\hat{a}^{\dagger}|n \rangle + \langle n|\hat{a}|n \rangle \right)$$
$$= \sqrt{\frac{\hbar}{2m\omega}} \left(\sqrt{n+1} \langle n||n+1 \rangle + \sqrt{n} \langle n||n-1 \rangle \right) = 0$$

From the orthogonality of $|n - 1 \rangle$, $|n \rangle$, and $|n + 1 \rangle$, we deduce that¹⁴²

$$<\mathbf{x}>=\mathbf{0}$$

A.3. To evaluate $\langle \hat{x}^2 \rangle$ and $\langle \hat{p}^2 \rangle$, we first express of the operator \hat{x}^2 and \hat{p}^2 in terms of the annihilation and creation operators:

$$\hat{x}^{2} = \frac{\hbar}{2m\omega} \left(\hat{a}^{2} + \hat{a}^{\dagger^{2}} + \hat{a}\hat{a}^{\dagger} + \hat{a}^{\dagger}\hat{a} \right) = \frac{\hbar}{2m\omega} \left(\hat{a}^{2} + \hat{a}^{\dagger^{2}} + 2\hat{a}^{\dagger}\hat{a} + 1 \right)$$
$$\hat{p}^{2} = -\frac{m\hbar\omega}{2} \left(\hat{a}^{2} + \hat{a}^{\dagger^{2}} - \hat{a}\hat{a}^{\dagger} - \hat{a}^{\dagger}\hat{a} \right) = -\frac{m\hbar\omega}{2} \left(\hat{a}^{2} + \hat{a}^{\dagger^{2}} - 2\hat{a}^{\dagger}\hat{a} - 1 \right)$$

¹⁴²This result can be deduced by symmetry argument since the Hamiltonian of the simple harmonics is invariant under the transformation $x \to -x$.

Since the operator the expectation values of the operators a^2 and \hat{a}^{\dagger^2} are zero, we have

$$\begin{aligned} <\hat{x}^2> &= \frac{\hbar}{2m\omega} < n|2\hat{a}^{\dagger}\hat{a}+1|n> \\ <\hat{p}^2> &= \frac{m\hbar\omega}{2} < n|2\hat{a}^{\dagger}\hat{a}+1|n> \end{aligned}$$

Therefore we obtain

$$\langle \hat{x}^2 \rangle = \frac{\hbar}{2m\omega} (2n+1)$$

 $\langle \hat{p}^2 \rangle = \frac{m\hbar\omega}{2} (2n+1)$

A.4. Since $|n\rangle$, n = 0, 1, ... form a basis, any state $|\psi\rangle$ can be written as a linear combination of these states, i.e.

$$|\psi> = \sum_{n=0}^{\infty} c_n |n>$$

For normalized state we have $\sum_{n} |c_n|^2 = 1$. then, the expectation value of \hat{H}_0 in a normalized state $|\psi\rangle$ reads

$$<\psi|\hat{H}_{0}|\psi> = \sum_{n,m} c_{m}^{*}c_{n} < m|\hat{H}_{0}|n> = \hbar\omega\sum_{n,m} c_{m}^{*}c_{n}\left(n+\frac{1}{2}\right)\delta_{nm}$$
$$= \frac{\hbar\omega}{2}\sum_{n} |c_{n}|^{2} + \sum_{n} |c_{n}|^{2}n$$

The first sum is equal to unity, and so the first term is just the energy of the ground state. The second term is greater or equal to zero. This implies that

$$<\psi|\hat{H}_0|\psi>\geq rac{\hbar}{2}$$

A.5. The action of the annihilation operator on the state $|z\rangle$ is given by

$$\begin{aligned} \hat{a}|z &> = e^{-\frac{|z|^2}{2}} \sum_{n=0}^{\infty} \frac{z^n}{\sqrt{n!}} \hat{a} \ |n > = e^{-\frac{|z|^2}{2}} \sum_{n=0}^{\infty} \frac{z^n}{\sqrt{n!}} \sqrt{n} \ |n-1 > \\ &= e^{-\frac{|z|^2}{2}} \sum_{n=1}^{\infty} \frac{z^n}{\sqrt{(n-1)!}} \ |n-1 > = e^{-\frac{|z|^2}{2}} \sum_{n=0}^{\infty} \frac{z^{n+1}}{\sqrt{n!}} \ |n > \end{aligned}$$

which implies that

$$\hat{\mathbf{a}}|\mathbf{z}>=\mathbf{z}|\mathbf{z}> \longrightarrow |\mathbf{z}>$$
 is an eigenstate of $\hat{\mathbf{a}}$

• Part B

B.1. As given in the problem above, the Hamiltonian of the system can be written as

$$\mathcal{H} = \mathcal{H}_x + \mathcal{H}_y$$

where $\hat{\mathcal{H}}_x$ and $\hat{\mathcal{H}}_x$ are the Hamiltonian of the simple harmonic oscillators. Since the two set of observables \hat{x}, \hat{p}_x and \hat{y}, \hat{p}_y commute, then eigenstates of \mathcal{H} are the simultaneous eigenstates of \mathcal{H}_x and \mathcal{H}_y , i.e. $|n_x, n_y \rangle$, where

$$\hat{\mathcal{H}}|n_x, n_y\rangle = E_{n_x} + E_{n_y} = \hbar\omega\left(n_x + \frac{1}{2}\right) + \hbar\omega\left(n_y + \frac{1}{2}\right)$$

Thus, the energy eigenstates of the Hamiltonian of a simple harmonic oscillator in two dimensions (or equivalently two independent 1-dimensional simple harmonic oscillators) are given by

$$E_n = \hbar\omega (n+1), \quad n = n_x + n_y$$

B.2. In the table below d_n denotes the degeneracy of the n^{th} energy level.

Level (n)	States (n_x, n_y)	Energy of the Level (E_n)	Degeneracy (d_n)
0	$(0,\!0)$	$\hbar\omega$	1 = 0 + 1
1	(1,0), (0,1)	$2\hbar\omega$	2 = 1 + 1
2	(2,0), (1,1), (0,2)	$3\hbar\omega$	3 = 2 + 1
3	(3,0), (2,1), (1, 2), (0,3)	$4\hbar\omega$	4 = 3 + 1
4	(4,0), (3,1), (2, 2), (1, 3), (0,4)	$5\hbar\omega$	5 = 4 + 1

 Table 1: Nonlinear Model Results

To deduce the degeneracy of the n^{th} level, we notice that the degeneracy of each level in the table above is equal to the level number plus one. Thus,

$$d_n=n+1$$

We can also "derive" the degeneracy of the n^{th} level without resorting into the table above. So for that we fix E_n , which equivalently amounts to fixing n, and count how many states have the same n. In other words, how many ways we can partition n into two nonnegative integers?. The set of states are (n_x, n_y) ; $n = n_x + n_y :=$ $(0, n), (1, n), \dots, (n - 1, 1), (n, 0)$. Thus the degeneracy of the n^{th} level is

$$d_n = \underbrace{1+1+1+\ldots+1}_{\mathrm{n}+1 \, \mathrm{terms.}} = n+1$$

Bonus Question :

In three dimensions, the energy state of an isotropic harmonic oscillator is labeled by three integers n_x, n_y , and n_z , with energy

$$E_{n_x,n_y,n_z} = \hbar\omega\left(n_x + \frac{1}{2}\right) + \hbar\omega\left(n_y + \frac{1}{2}\right) + \hbar\omega\left(n_y + \frac{1}{2}\right)$$

Thus, the energy of a level n is given by

$$E_n = \hbar\omega\left(n + \frac{3}{2}\right), \qquad n = n_x + n_y + n_z$$

The number of states with with the same energy level n, can be obtained as follows. For each fixed n_x , the number of states correspond to the degeneracy of a 2-dimensional isotropic harmonic oscillator, that we derived in the previous question:

$$d_n^{(n_x = \text{fixed})} = (n_y + 1) = (n - n_x + 1)$$

Thus, to find degeneracy of the 3-dimensional isotropic harmonic oscillator at energy level n, we sum over all possible allowed values of n_x , i.e.

$$d_n = \sum_{n_x=0}^n (n - n_x + 1) = (n+1) \sum_{n_x=0}^n 1 - \sum_{n_x=0}^n n_x$$
$$= (n+1)^2 - \frac{n(n+1)}{2}$$

from which we obtain

$$\mathbf{d_n} = \frac{(\mathbf{n+1})(\mathbf{n+2})}{\mathbf{2}}$$

United Arab Emirates University

Final Exam

Quantum Physics (PHYS 525)

Semester: Fall 2016

Date: 14/12/2016 Time: 2:00 - 4:00 pm

Student Name/ID:

Abstract:

- 1. This exam consists of of **4 Problems**.
- 2. You need to solve all of the problems.
- 3. The last page is an **Appendix** with some formulas.
- 4. Please **SHOW YOUR WORK**.

I am not telling you it is going to be easy – I am telling you it is going to be worth it !.

Art Williams

Problem One [10 points] (System of Two Spin 1/2 Particles)

1. The hamiltonian describing the interaction of two unperturbed spin 1/2 particles¹⁴³ is given by¹⁴⁴

$$\hat{H}_0 = \frac{A}{\hbar^2} \,\hat{\mathbf{S}}_{(1)}.\hat{\mathbf{S}}_{(2)}$$

where A is a constant with dimension of energy. Let $\hat{\mathbf{J}} = \hat{\mathbf{S}}_{(1)} + \hat{\mathbf{S}}_{(2)}$ be the total angular momentum operator of the system.

1.1. Argue that with this perturbation term it is convenient to work in the basis $\{|j, M >\}$ of the total angular momentum instead of the basis $\{|j_1, j_2; m_1, m_2 >\}$.

1.2. Show that the above Hamiltonian can be written as

$$\hat{H}_0 = A\left(\frac{\hat{\mathbf{J}}^2}{2\hbar^2} - \frac{3}{4}\right)$$

and that it has the eigenvalues -3A/4 and A/4 with degeneracies 1 and 3, respectively.

2. Now we add to \hat{H}_0 the perturbation term

$$\delta \hat{H} = \frac{\Delta}{\hbar} \left(\hat{\mathbf{S}}_{(1)} - \hat{\mathbf{S}}_{(2)} \right) \cdot \mathbf{B}, \qquad \Delta < 1$$

with **B** is a magnetic field applied to the system, which we take to be along the z-axis.

Determine the energies of this system and show that the degeneracy of the unperturbed state of energy A/4 gets reduced¹⁴⁵.

 $^{^{143}}$ For example, this could be a system of electron-positron (like the bound state of Positronium), or an electron-proton (like in the hydrogen atom).

 $^{^{144}\}mathrm{Here}\ \mathrm{I}$ denote a vector quantity by boldface character.

 $^{^{145}{\}rm For}$ the degeneracy It will be enough you draw energy levels diagram showing the amount by which the relevant state is lifted.

Answer For Problem One

1.

1.1. We note that \hat{H}_0 commutes with $\hat{\mathbf{J}}^2$, J_z , $\hat{\mathbf{S}}^2_{(1)}$ and $\hat{\mathbf{S}}^2_{(2)}$, where as it does not commute with any of the components of $\hat{\mathbf{S}}_{(1)}$ and $\hat{\mathbf{S}}_{(2)}$. This means that the eigenstates of \hat{H}_0 can be labelled with the quantum numbers J, M associated with the total angular momentum operator \hat{J} and its component along the z-axis, respectively. Thus, it is convenient to work in the basis $\{|j, M >\}$.

1.2. Using the simple relation $\hat{J}^2 = \hat{S}_1^2 + \hat{S}_2^2 + 2\hat{S}_1.\hat{S}_2$, we can write

$$\hat{H}_0 = \frac{A}{2\hbar^2} \left(\hat{\mathbf{J}}^2 - \hat{\mathbf{S}}_1^2 - \hat{\mathbf{S}}_2^2 \right)$$

Since spin 1/2 particles $\hat{\mathbf{S}}_1^2 = \hat{\mathbf{S}}_1^2 = 3/4$ Id, the above hamiltonian reads

$$\hat{H}_0 = A\left(\frac{\hat{\mathbf{J}}^2}{2\hbar^2} - \frac{3}{4}\right)$$

Thus, the eigenstates and the corresponding eigenvalues of this system are given by

$$\hat{H}_0|j,M> = \frac{A}{2}\left[j(j+1) - \frac{3}{2}\right]$$

where the possible values of j are 0 and 1. Hence, we have¹⁴⁶

$$\begin{cases} \text{Singlet} \equiv {}^{1}S_{0}: & \hat{H}_{0}|0, 0 > = -\frac{3A}{4}|0, 0 > \Rightarrow E_{|0,0>} = -\frac{3A}{4} \\ \text{Triplet} \equiv {}^{3}S_{1}: & \hat{H}_{0}|1, M > = \frac{A}{4}|1, M >, \Rightarrow E_{|1,M>} = \frac{A}{4} \end{cases}$$

We notice that the state j = 1 has three fold degeneracy.

¹⁴⁶Recall the spectroscopic notation $^{2j+1}X_j$.

Answer For Problem One (Continued):

2. It is easy to see that $\delta \hat{H}$ commutes with $\hat{\mathbf{S}}_{(1)}^2$, $\hat{\mathbf{S}}_{(2)}^2$, $\hat{S}_{(1)}^{(z)}$ and $\hat{S}_{(2)}^{(z)}$, but does not commute with $\hat{\mathbf{J}}^2$. We need to express the state $|0, 0\rangle$ in terms of the states $|1/2, 1/2; m_1, m_2\rangle \equiv |m_1, m_2\rangle$, with $m_i = \pm 1/2$ which I will represent them by the symbols \uparrow and \downarrow , respectively. We have shown in class¹⁴⁷ that

$$\begin{aligned} |1,1\rangle &= |\uparrow;\uparrow\rangle \\ |1,0\rangle &= \frac{1}{\sqrt{2}} \left(|\uparrow;\downarrow\rangle + |\downarrow;\uparrow\rangle\right) \\ |1,-1\rangle &= |\downarrow;\downarrow\rangle \\ |0,0\rangle &= \frac{1}{\sqrt{2}} \left(|\uparrow;\downarrow\rangle - |\downarrow;\uparrow\rangle\right) \end{aligned}$$

Below, we will calculate the correction to the energies for the singlet and the triplets to the lowest order in perturbation theory.

• Singlet State $({}^{1}S_{0})$:

$$\delta E_{|0,0\rangle}^{(1)} = \frac{B\Delta}{\hbar} < 0, 0 |\hat{S}_{(1)}^{(z)} - \hat{S}_{(2)}^{(z)}|0,0\rangle$$

We first evaluate $\hat{S}_{(1)}^{(z)} - \hat{S}_{(2)}^{(z)}|0, 0>$ using the expression of |0, 0> in terms of the spin 1/2 states given above. We have

$$\begin{bmatrix} \hat{S}_{(1)}^{(z)} - \hat{S}_{(2)}^{(z)} \end{bmatrix} |0,0\rangle = \frac{1}{\sqrt{2}} \hat{S}_{(1)}^{(z)} (|\uparrow;\downarrow\rangle - |\downarrow;\uparrow\rangle) - \frac{1}{\sqrt{2}} \hat{S}_{(2)}^{(z)} (|\uparrow;\downarrow\rangle - |\downarrow;\uparrow\rangle) = \frac{\hbar}{\sqrt{2}} (|\uparrow;\downarrow\rangle + |\downarrow;\uparrow\rangle) = \hbar |1,0\rangle$$
(16.-2)

Since the states |0, 0 > and 1, 0 > are orthogonal¹⁴⁸, it follows that

$$\delta E^{(1)}_{|0,0>} = 0$$

 $^{^{147}\}mathrm{It}$ was also part of a homework. Otherwise, you can also re-derive it easily:).

¹⁴⁸Since they the eigenstates |j, M > form an orthonormal basis

Answer For Problem One (Continued):

• Triplet State $({}^{3}S_{1})$:

We already showed that this state is degenerate, and hence we need to apply the procedure that we learnt in class for degenerate perturbation theory. For that we need to evaluate the matrix elements $< 1, M' | \delta \hat{H} | 1, M >$, with M and M' takes the values in the set $\{1, 0, -1\}$. Using the expressions of |1, M > in terms of the spin-1/2 states, have

$$\left[\hat{S}_{(1)}^{(z)} - \hat{S}_{(2)}^{(z)}\right]|1, 1\rangle = \left[\hat{S}_{(1)}^{(z)} - \hat{S}_{(2)}^{(z)}\right]|\uparrow, \uparrow\rangle = 0, \quad (16.-1)$$

$$\left[\hat{S}_{(1)}^{(z)} - \hat{S}_{(2)}^{(z)}\right]|1, -1> = \left[\hat{S}_{(1)}^{(z)} - \hat{S}_{(2)}^{(z)}\right]|\downarrow, \downarrow> = 0, \quad (16.0)$$

and

$$\begin{bmatrix} \hat{S}_{(1)}^{(z)} - \hat{S}_{(2)}^{(z)} \end{bmatrix} |1,0\rangle = \frac{1}{\sqrt{2}} \hat{S}_{(1)}^{(z)} (|\uparrow;\downarrow\rangle + |\downarrow;\uparrow\rangle) - \frac{1}{\sqrt{2}} \hat{S}_{(2)}^{(z)} (|\uparrow;\downarrow\rangle + |\downarrow;\uparrow\rangle) = \frac{\hbar}{\sqrt{2}} (|\uparrow;\downarrow\rangle - |\downarrow;\uparrow\rangle) = \hbar |0,0\rangle$$
(16.1)

Since $\langle 1, M' | 1, M \rangle = \delta_{MM'}$, it follows that the first-order energy corrections for the states $|1, M \rangle$ vanish.

Hence, to see the effect of the perturbation on these states we need to go to higher order, i.e. the second order of perturbation, given by

$$\delta E_{|0,0\rangle}^{(2)} = \langle 0, 0 | \delta \hat{H} | 0, 0 \rangle^{(1)}$$

$$\delta E_{|1,M\rangle}^{(2)} = \langle 1, M | \delta \hat{H} | 1, M \rangle^{(1)}$$

where the $|0, 0\rangle^{(1)}$ and $|1, M\rangle^{(1)}$ are the first order corrected states to $|0, 0\rangle$ and $|1, M\rangle$, respectively.

Answer For Problem One (Continued)

With the help of the formula given in the Appendix we we have

$$\begin{split} |0,0>^{(1)} &= \sum_{M=0,\pm} \frac{<1, M |\delta \hat{H}|0, 0>}{E_{|0,0>} - E_{|1,M>}} |1, M> \\ |1,1>^{(1)} &= \quad \frac{<0, 0 |\delta \hat{H}|1, 1>}{E_{|1,1>} - E_{|0,0>}} |0, 0> \\ |1,-1>^{(1)} &= \quad \frac{<0, 0 |\delta \hat{H}|1, -1>}{E_{|1,-1>} - E_{|0,0>}} |0, 0> \\ |1,0>^{(1)} &= \quad \frac{<0, 0 |\delta \hat{H}|1, 0>}{E_{|1,0>} - E_{|0,0>}} |0, 0> \end{split}$$

With the use of equations (16.-1), (16.0) and (16.1), we obtain

$$|0,0\rangle^{(1)} = -\frac{B\Delta}{A} |1,0\rangle$$

$$|1,1\rangle^{(1)} = 0, \quad |1,-1\rangle^{(1)} = 0, \quad |1,0\rangle^{(1)} = \frac{B\Delta}{A} |0,0\rangle$$

We note that there is no second order energy shift for the states $|1, 1 \rangle$ and $|1, -1 \rangle$, whereas the states $|0, 0 \rangle$ and $|1, 0 \rangle$ receive correction given by

$$\begin{split} \delta E^{(2)}_{|0,0>} &= <0,0 |\delta \hat{H}|0,0>^{(1)} = -\frac{B\Delta}{A} <0,0 |\delta \hat{H}|1,0>\\ \delta E^{(2)}_{|1,0>} &= <1,0 |\delta \hat{H}|1,0>^{(1)} = \frac{B\Delta}{A} <1,0 |\delta \hat{H}|0,0> \end{split}$$

Using Eq (16.-2), we get

$$\delta E_{|0,0\rangle}^{(2)} = -\frac{(B\Delta)^2}{A} \\ \delta E_{|1,0\rangle}^{(2)} = +\frac{(B\Delta)^2}{A}$$

Problem Two [10 points] (Perturbed Harmonic Oscillator)

1. Estimate the energy of the ground state of a 1-dimensional simple harmonic oscillator using the trial function

$$\psi_{\alpha}(x) = \begin{cases} C \cos \alpha x, & -\pi/2 < \alpha x < \pi/2 \\ 0, & \text{elsewhere} \end{cases}$$

where α is the variational parameter, and C is a normalization constant.

2. We add to the simple oscillator a small perturbation so that

$$\hat{H} = \hat{H}_0 + \delta \hat{V}$$

where $\delta \hat{V}$ is Hermitian. Suppose that $\delta \hat{V} = \lambda \hat{x}^4$, with $\lambda \ll 1$. Using perturbation theory **find** the correction to the energy of the nth-level of the simple harmonic oscillator up to the first order in the parameter λ .

3. Suppose that $\delta \hat{V} = \delta \hat{x}$, with δ a real number smaller than unity.

3.1. Compute the energy of the non-zero correction to the energy of the ground state of the simple harmonic oscillator to the lowest order in perturbation.

3.2. Show that the energy levels of this system can be solved exactly, and then **Compare** your result you obtained in **B.2.1**.

3.3. Can you **think** of a physical phenomenon for which the system is subject to a linear perturbation ? What does the parameter δ represent?

4. Now let us act on the unperturbed simple harmonic by a time dependent force^{149}

$$F(t) = \frac{F_0}{\omega} \frac{\tau}{\tau^2 + t^2}$$

and we assume that at $t \to -\infty$ the oscillator was in ground state. Using the time-dependent perturbation theory to first order, **calculate** the probability that the oscillator is found in the first excited state at $t \to +\infty$.

¹⁴⁹Note that this a force, not a potential.

Answer For Problem Two

1. We need first to compute the average energy

$$H(\alpha) := \frac{\langle \psi_{\alpha}(x) | H | \psi_{\alpha}(x) \rangle}{\langle \psi_{\alpha} | \psi_{\alpha} \rangle}$$

and then find its minimum value. For our system, we have

$$H(\alpha) = \frac{1}{\langle \psi_{\alpha} | \psi_{\alpha} \rangle} \times \left[-\frac{\hbar^2}{2m} \langle \psi_{\alpha} | \frac{d^2}{dx^2} | \psi_{\alpha} \rangle + \frac{1}{2}m\omega^2 \langle \psi_{\alpha} | x^2 | \psi_{\alpha} \rangle \right]$$
$$= \frac{1}{\langle \psi_{\alpha} | \psi_{\alpha} \rangle} \times \left[-\frac{\hbar^2}{2m} \int_{-\frac{\pi}{2\alpha}}^{\frac{\pi}{2\alpha}} \psi_{\alpha}^*(x) \frac{d^2}{dx^2} \psi_{\alpha}(x) + \frac{1}{2}m\omega^2 \int_{-\frac{\pi}{2\alpha}}^{\frac{\pi}{2\alpha}} \psi_{\alpha}^*(x) x^2 \psi_{\alpha}(x) \right]$$

For our trial wave function $\psi_{\alpha}(x)$ in the range $|\alpha x| < \pi/2$, we have

$$H(\alpha) = \frac{1}{\int_{-\frac{\pi}{2\alpha}}^{\frac{\pi}{2\alpha}} \cos^2\left(\alpha x\right) \, dx} \times \left(\frac{\hbar^2 \alpha^2}{2m} \int_{-\frac{\pi}{2\alpha}}^{\frac{\pi}{2\alpha}} \cos^2\left(\alpha x\right) \, dx + \frac{1}{2}m\omega^2 \int_{-\frac{\pi}{2\alpha}}^{\frac{\pi}{2\alpha}} x^2 \cos^2\left(\alpha x\right) \, dx\right)$$

$$=\frac{\hbar^2 \alpha^2}{2m} + \frac{\alpha m \omega^2}{2\pi} \left[\int_{-\frac{\pi}{2\alpha}}^{\frac{\pi}{2\alpha}} x^2 \, dx - \frac{1}{\alpha^3} \left(\frac{\partial^2}{\partial \beta^2} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos 2\beta \theta \, d\theta \right)_{\beta=2} \right]$$

which yields

$$H(\alpha) = \frac{\hbar^2 \alpha^2}{2m} + \frac{m\omega^2 \left(\pi^2 - 6\right)}{12\alpha^2}$$

Setting the derivative of $H(\alpha)$ with respect to the parameter α to zero we

find that it is minimum at the value

$$\alpha_*^2 = \left(\frac{\pi^2 - 6}{12}\right)^{1/2} \frac{m\omega}{\hbar}$$

Substituting this value into the above expression of $H(\alpha)$ we obtain

$$H(\alpha_*) = \left(\frac{\pi^2 - 6}{12}\right)^{1/2} \hbar\omega \simeq 0.57 \ \hbar\omega$$

which is a about 12% larger than the exact value $\hbar\omega/2$ for the ground state.

Answer For Problem Two (Continued)

2. To first order in (non-degenrate) perturbation theory, the correction of the energy of the n^{th} level is given by

$$\begin{split} \delta E_n &= \langle n | \delta \hat{V} | n \rangle = \lambda \langle n | \delta \hat{x}^4 | n \rangle = \lambda \langle n | \left(\hat{a} + \hat{a}^\dagger \right)^4 | n \rangle \\ &= \lambda \left(\frac{\hbar}{2m\omega} \right)^2 \langle n | \hat{a}^2 \hat{a}^{\dagger 2} + \hat{a} \hat{a}^\dagger \hat{a} \hat{a}^\dagger + \hat{a} \hat{a}^\dagger \hat{a} \hat{a}^\dagger \hat{a} + \hat{a}^\dagger \hat{a} \hat{a} \hat{a}^\dagger + \hat{a}^\dagger \hat{a} \hat{a} \hat{a}^\dagger \hat{a} + \hat{a}^\dagger \hat{a} \hat{a} \hat{a}^\dagger \hat{a} + \hat{a}^\dagger \hat{a} \hat{a} \hat{a}^\dagger \hat{a} + \hat{a}^\dagger \hat{a} \hat{a} \hat{a} \hat{a} + \hat{a} + \hat{a}^\dagger \hat{a} \hat{a} + \hat{a} + \hat{a} \hat{a} \hat{a} + \hat{a} + \hat{a} + \hat{a} \hat{a} + \hat{a}$$

where the average of the terms \hat{a}^4 and $\hat{a}^{\dagger 4}$ vanish¹⁵⁰. The action of the the different products of the annihilation and creation operators on the state $|n\rangle$ can be calculated in straight forward way using the definition of the lowering and raising operators given in the appendix, and you find

$$\begin{aligned} \hat{a}^{2} \hat{a}^{\dagger 2} |n\rangle &= (n+1) (n+2) |n\rangle \\ \hat{a} \hat{a}^{\dagger} \hat{a} \hat{a}^{\dagger} |n\rangle &= (n+1)^{2} |n\rangle \\ \hat{a}^{\dagger} \hat{a} \hat{a} \hat{a}^{\dagger} |n\rangle &= n (n+1) |n\rangle \\ \hat{a} \hat{a}^{\dagger} \hat{a}^{\dagger} \hat{a} |n\rangle &= n (n+1) |n\rangle \\ \hat{a}^{\dagger} \hat{a} \hat{a}^{\dagger} \hat{a} |n\rangle &= n^{2} |n\rangle \\ \hat{a}^{\dagger 2} \hat{a}^{2} |n\rangle &= n (n-1) |n\rangle \end{aligned}$$

Substituting in the above expression of δE_n we obtain

$$\delta E_n = 3\lambda \left(\frac{\hbar}{2m\omega}\right)^2 \left(2n^2 + 2n + 1\right)$$

3.

3.1. Using the explicit expression of the position operator in terms of the annihilation and creation operators, we can easily show that

$$< m |\hat{x}|n > = \sqrt{\frac{\hbar}{2m\omega}} \left(\sqrt{n+1}\delta_{m,n+1} + \sqrt{n}\delta_{m,n+1}\right)$$

¹⁵⁰Since $< n|\hat{a}^4|n > \propto < n|n-4 > = 0$ and $< n|\hat{a}^4|n > \propto < n|n+4 > = 0$.

Answer For Problem Two (Continued)

This implies that at the first-order of perturbation theory the correction to the energy of any unperturbed state vanishes. So, we need to consider the the second-order correction, given by

$$\delta E_n^{(2)} = \lambda^{(0)} < n |\hat{x}| n >^{(1)} = \lambda^2 \sum_{m \neq n} \frac{|\langle n | x | m \rangle|^2}{E_n^{(0)} - E_m^{(0)}}$$
$$= \frac{\lambda^2}{2m\omega^2} \sum_{m \neq n} \frac{(m+1)\delta_{n,m+1} + m\delta_{n,m-1}}{n - m}$$

which gives

$$\delta E_n^{(2)} = -\frac{\lambda^2}{2m\omega^2}$$

3.2. This system can be solved exactly as follows:

$$\hat{H} = -\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + \frac{1}{2}m\omega^2 x^2 + \lambda x$$
$$= -\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + \frac{1}{2}m\omega^2 \left(x + \frac{\lambda}{m\omega^2}\right)^2 - \frac{\lambda^2}{2m\omega^2}$$

Now by make the change of variable $z = x - \frac{\lambda}{2m\omega^2}$, the above Hamiltonian reads

$$\hat{H} = -\frac{\hbar^2}{2m}\frac{d^2}{dz^2} + \frac{1}{2}m\omega^2 z^2 - \frac{\lambda^2}{2m\omega^2}$$

Hence, $\left(\hat{H} + \frac{\lambda^2}{2m\omega^2}\right)$ is the energy operator of a simple harmonic oscillator with the same frequency as the unperturbed one. Therefore.

$$E_n = \hbar\omega\left(n + \frac{1}{2}\right) - \frac{\lambda^2}{2m\omega^2}$$

Notice that the above result is the same as the one we derived using the lowest order perturbation theory. Thus, we except that all correction at order higher than the second order vanish.

Answer For Problem Two

3.3. We saw in class a system that exhibit similar phenomenon, which is the **Stark Effect**, where δ is proportional to the applied **electric field**.

4. In one dimension (this case), the force is the minus of the derivative of the potential with respect to x, which implies that the perturbation operator reads

$$\hat{V}(x,t) = -\frac{F_0}{\omega} \frac{\tau}{\tau^2 + t^2} \hat{x}$$

The ground state, denoted by $|0\rangle$, energy is $E_0 = \hbar \omega/2$ and the one of the first excited state, denoted by $|1\rangle$, is $E_1 = 3\hbar \omega/2$, and so transition frequency is

$$\omega_{10} = \frac{E_1 - E_0}{\hbar} = \omega$$

Hence, for this system, the transition probability amplitude to the first excited state (see the appendix) is given by

$$c_{0\to1}^{(1)}(\infty) = \frac{i}{\hbar} \frac{F_0 \tau}{\omega} < 1 |\hat{x}|_0 > \int_{-\infty}^{+\infty} \frac{e^{i\omega t}}{\tau^2 + t^2} dt$$
$$= \frac{i}{\hbar} \frac{F_0 \tau}{\omega} \sqrt{\frac{\hbar}{2m\omega}} \int_{-\infty}^{+\infty} \frac{e^{i\omega t}}{\tau^2 + t^2} dt$$
(16.2)

where in the second equality we used the fact that $\hat{x}|0\rangle = \sqrt{\hbar/2m\omega} \hat{a}^{\dagger}|0\rangle = \sqrt{\hbar/2m\omega} |1\rangle$ and $\langle 1|1\rangle$ is normalized to unity. The time integral above can performed using the residue theorem. For that we chose the closed contour to be the real axis plus the half circle in the upper plane enclosing the pole $t = i\tau$ with its radius R taken to infinity:

$$\lim_{R \to \infty} \oint \frac{e^{i\omega z}}{(z+i\tau)(z-i\tau)} dz = \int_{-\infty}^{+\infty} \frac{e^{i\omega t}}{\tau^2 + t^2} dt + \lim_{R \to \infty} \int_{\text{half circle in upper plane}} \frac{e^{i\omega z}}{(z+i\tau)(z-i\tau)} dz$$
$$= \int_{-\infty}^{+\infty} \frac{e^{i\omega t}}{\tau^2 + t^2} dt$$

The left hand side of the above equation is equal to $2\pi i$ times the residue at the pole $z = i\tau$, which yields

$$\int_{-\infty}^{+\infty} \frac{e^{i\omega t}}{\tau^2 + t^2} dt = 2\pi i \frac{e^{-\omega}}{2i\tau} = \frac{\pi e^{-\omega}}{\tau}$$

Substituting the value of this integral into Eq. (16.2), we obtain

$$c_{0\to1}^{(1)}(\infty) = \frac{i}{\hbar} \frac{F_0 \tau}{\omega} \sqrt{\frac{\hbar}{2m\omega}} \frac{\pi e^{-\omega}}{\tau}$$

Thus, the transition probability amplitude for the oscillator is in the first excited state reads

$$P(0 \to 1) = |c_{0 \to 1}^{(1)}(\infty)|^2 = \frac{\pi^2 F_0^2}{2m\hbar\omega^3} e^{-2\omega\tau}$$

Problem Three [6 points] (Landau Levels)

Consider an electron moving on the x-y plane in a constant magnetic field \mathbf{B} directed along the z-axis. As we have seen in class, the Hamiltonian of such system is given by¹⁵¹

$$\hat{H} = \frac{\hat{\Pi}_x^2 + \hat{\Pi}_y^2}{2m}$$

where $\hat{\Pi}_i = \hat{p}_i - e\hat{A}_i$, and \hat{A}_i 's are the components of the vector potential.

1. Calculate the commutator $\left[\hat{\Pi}_x, \hat{\Pi}_y\right]$.

2. **Deduce** that the operators $\hat{a} = \frac{1}{\sqrt{2e\hbar B}} \left(\hat{\Pi}_x + i\hat{\Pi}_y \right)$, and $\hat{a}^{\dagger} = \frac{1}{\sqrt{2e\hbar B}} \left(\hat{\Pi}_x - i\hat{\Pi}_y \right)$ satisfy the commutation relation $[\hat{a}, \hat{a}^{\dagger}] = 1$.

3. Show that this system is equivalent to a simple harmonic oscillator with frequency $\omega = eB/m$.

3. Let us choose the so-called the symmetric gauge $\mathbf{A} = (-By/2, Bx/2, 0)$.

3.1. Show that the ground state wave function is a solution to the differential equation¹⁵²

$$\left(\bar{\partial} + \frac{eB}{4\hbar}z\right)\psi(z,\bar{z}) = 0,$$

3.2. Show that a wave function of the form

$$\psi(z,\bar{z}) = \Phi(z) e^{-eBz\bar{z}/4\hbar}$$

is a solution to the above differential equation.

$$\partial \equiv \frac{\partial}{\partial z} = \frac{1}{2} \left(\frac{\partial}{\partial x} - i \frac{\partial}{\partial y} \right), \quad \bar{\partial} \equiv \frac{\partial}{\partial \bar{z}} = \frac{1}{2} \left(\frac{\partial}{\partial x} + i \frac{\partial}{\partial y} \right),$$

where z = x + iy is a complex number, and \bar{z} its complex conjugate. The above derivatives satisfy the following identities $\partial \bar{z} = \bar{\partial} z = 0$, and $\partial z = \bar{\partial} \bar{z} = 1$.

¹⁵¹This is part of a problem first treated in 1930 by Lev Landau (Nobel Prize in Physics in 1962) which turns out to be the setup for one of the most profound phenomena in condensed matter physics called the Quantum Hall Effect.

 $^{^{152}}$ Here the partial derivative is in the complex plane, defined as

Answer For Problem Three

1.

$$[\Pi_x, \Pi_y] = [p_x - eA_x, p_y - eA_y] = -e[p_x, A_y] + e[p_y, A_x]$$
$$= ie\hbar \left(\frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y}\right) = ie\hbar (\nabla \times \mathbf{A})_z = ie\hbar B$$

where in the second equality we used the fact that $[p_i, f(x_i)] = -i\hbar \partial f(x_i)/\partial x_i$.

2.

$$\begin{bmatrix} \hat{a}, \hat{a}^{\dagger} \end{bmatrix} = \frac{1}{2e\hbar B} [\Pi_x + i \Pi_y, \Pi_x - i \Pi_y] \\ = -2i \left(\frac{1}{2e\hbar B}\right) [\Pi_x, \Pi_y] = \mathbf{1}$$

3. We will express $\left(\hat{\Pi}_x^2 + \hat{\Pi}_y^2\right)$ in terms of \hat{a} and \hat{a}^{\dagger} as follows:

$$\begin{pmatrix} \hat{\Pi}_x^2 + \hat{\Pi}_y^2 \end{pmatrix} = (\Pi_x - i \ \Pi_y) (\Pi_x + i \ \Pi_y) - i[\Pi_x, \Pi_y]$$
$$= 2e\hbar B \hat{a}^{\dagger} \hat{a} - i (ie\hbar B)$$
$$= 2e\hbar B \left(\hat{a}^{\dagger} \hat{a} + \frac{1}{2} \right)$$

Thus, the Hamiltonian of this system can be written as

$$\hat{H} = \hbar \frac{eB}{m} \left(\hat{a}^{\dagger} \hat{a} + \frac{1}{2} \right)$$

which is represents a simple harmonic oscillator with a frequency $\omega = eB/m$, i.e. the cyclotron frequency.

Answer For Problem Three (Continued)

BONUS [6 points]:

B.1. By definition, the operator \hat{a} annihilates the ground state of a simple harmonic oscillator. Hence, the corresponding wave function satisfy the equation¹⁵³

$$\langle x|\hat{a}|0\rangle = 0 \implies \frac{1}{\sqrt{2e\hbar B}} [-i\hbar\partial_x - eA_x + i(-i\hbar\partial_y - eA_y)]\psi(x,y) = 0$$

where $\psi(x, y)$ is the ground state wave function. In the symmetric gauge, the above equation reads

$$\left[\left(\partial_x + i\partial_y\right) + \frac{eB}{2\hbar}\left(x + iy\right)\right]\psi(x, y) = 0$$

or, equivalently, in terms of the complex variable z can be written as¹⁵⁴

$$\left(\bar{\partial} + \frac{eB}{4\hbar}z\right) \ \psi(z,\bar{z}) = 0, \qquad \qquad \bar{\partial} := \frac{\partial}{\partial\bar{z}}$$

B.2.

$$\begin{split} \left(\bar{\partial} + \frac{eB}{4\hbar}z\right) \ \Phi(z) e^{-eBz\bar{z}/4\hbar} &= \Phi(z)\bar{\partial}\left(e^{-eBz\bar{z}/4\hbar}\right) + \frac{eB}{4\hbar}z \ \Phi(z) e^{-eBz\bar{z}/4\hbar} \\ &= \Phi(z)\left(-\frac{eB}{4\hbar}z\right) \ e^{-eBz\bar{z}/4\hbar} + \frac{eB}{4\hbar}z \ \Phi(z) e^{-eBz\bar{z}/4\hbar} \\ &= 0 \end{split}$$

¹⁵³For the ease of the notation I will denote $\frac{\partial}{\partial x_i}$ by ∂_i .

¹⁵⁴A point in a plane with coordinates (x, y), can be equivalently represented by a complex number, z = x + iy. The technical word used in mathematics for such one to one correspondence is called **isomorphism**.

Problem Four [4 points] (Exciting the Hydrogen Atom)

Consider a hydrogen atom in a uniform (does not depend on position) but time dependent electric field along the z-direction, such that

$$\delta V(t) = \begin{cases} 0, & \text{for } t < 0\\ |e| z \mathcal{E}_0 \exp\left(-t/\tau\right), & \text{for } t \ge 0 \end{cases}$$

where |e| is the magnitude of the electron's electric charge, and \mathcal{E}_0 is a constant. At time t = 0, the atom is in the ground 1s state.

Show that to lowest order in perturbation theory, the probability that as $t \to \infty$, the atom is in the 2p state is given by

$$\mathcal{P}_{1s \to 2p} = \frac{2^{15}}{3^{10}} \frac{\left(|e|\mathcal{E}_0 a_0\right)^2}{\left(E_{ep} - E_{1s}\right)^2 + \left(\hbar/\tau\right)^2}$$

where a_0 is the Bohr radius.

Answer For Problem Four

The transition probability amplitude for this system at instant t is

$$c_{1s\to 2p}^{(1)}(t) = -\frac{i}{\hbar} \int_0^t \langle 210|\delta \hat{V}(t)|100\rangle e^{i\omega t'} dt', \qquad \omega = \frac{E_{2p} - E_{1s}}{\hbar}$$

Using the wave functions given in the appendix, the matrix element inside the integral reads

$$<210|\delta\hat{V}(t)|100> = |e|\mathcal{E}_0 \exp\left(-t/\tau\right) < 210|r\cos\theta|100>$$
$$= \frac{|e|\mathcal{E}_0 \exp\left(-t/\tau\right)}{4\pi a_0^4 \sqrt{2}} \left[\int_0^{2\pi} d\phi\right] \left[\int_0^{\pi} \cos^2\theta \sin\theta \ d\theta\right] \left[\int_0^{\infty} r^4 \exp\left(-3r/2a_0\right) \ dr\right]$$

With the use of the integrals given in the appendix, we obtain

$$< 210|\delta \hat{V}(t)|100> = \frac{2^{15/2}}{3^5}|e|\mathcal{E}_0 a_0 \exp(-t/\tau)$$

So,

$$c_{1s\to 2p}^{(1)}(t) = -\frac{i}{\hbar} \frac{2^{15/2}}{3^5} |e| \mathcal{E}_0 a_0 \int_0^t e^{-\frac{t'}{\tau} + i\omega t'} dt' = -\frac{i}{\hbar} \frac{2^{15/2}}{3^5} |e| \mathcal{E}_0 a_0 \left[\frac{e^{i\omega t - \frac{t'}{\tau} - 1}}{i\omega - 1/\tau} \right]$$

Therefore, the transition probability is given by

$$\mathcal{P}_{1s \to 2p} = |c_{1s \to 2p}^{(1)}(\infty)|^2 = \frac{2^{15}}{3^{10}} \frac{\left(|e|\mathcal{E}_0 a_0\right)^2}{\left(E_{ep} - E_{1s}\right)^2 + \left(\hbar/\tau\right)^2}$$
Appendix

• Simple Harmonic Oscillator

$$\hat{x} = \sqrt{\frac{\hbar}{2m\omega}} \left(\hat{a}^{\dagger} + \hat{a} \right), \qquad \hat{p} = i\sqrt{\frac{m\omega\hbar}{2}} \left(\hat{a}^{\dagger} - \hat{a} \right)$$
$$\hat{a}|n\rangle = \sqrt{n} |n-1\rangle, \qquad \hat{a}^{\dagger}|n\rangle = \sqrt{n+1} |n+1\rangle$$

• Angular Momentum

$$\hat{J}_{\pm} = \hat{J}_x \pm i\hat{J}_y$$

$$\hat{J}_{\pm} | j, m \rangle = \hbar \sqrt{j(j+1) - m(m\pm 1)} | j, m\pm 1 \rangle$$

• Time-independent Perturbation: Non-Degenerate Case

$$|n\rangle = |n^{0}\rangle + \sum_{k \neq n} |k^{0}\rangle \frac{V_{kn}}{E_{n}^{0} - E_{k}^{0}} + \dots$$
$$\Delta_{n} \equiv E_{n} - E_{n}^{0} = V_{nn} + \sum_{k \neq n} \frac{|V_{kn}|^{2}}{E_{n}^{0} - E_{k}^{0}} + \dots$$

• Time-dependent Perturbation

$$C_n^0(t) = \delta_{ni}$$

$$C_n^{(1)}(t) = -\frac{i}{\hbar} \int_0^t e^{i\omega_{ni}} V_{ni}(t') dt'; \qquad \omega_{ni} = \frac{E_n - E_i}{\hbar}$$

• Hydrogen Atom

state 1 s:
$$\psi_{100} = (\pi a_0^3)^{-1/2} \exp(-r/a_0)$$

state 2 p: $\psi_{210} = (8\pi a_0^3)^{-1/2} \cos\theta \frac{r}{2a_0} \exp(-r/2a_0)$

• Integrals

$$\int_0^\infty r^4 \exp\left(-\frac{3r}{2a_0}\right) dr = \left(\frac{2a_0}{3}\right)^5 4!$$
$$\int \cos^2\theta \ d\theta = \frac{2}{3}$$

17 Appendices

17.1 Specific intensity and Flux

The intensity, flux, energy density, and radiation pressure are very important concepts in studying the emission and absorption of radiation. Below, we give their definitions.

1. Specific Intensity (I_{ν}) :

the specific intensity, also called the **surface brightness**, is the rate of energy flowing through an infinitesimal element of surface of area dA^{155} , around a point of position \vec{r} , per unit frequency interval $d\nu$, per unit solid angle $d\Omega$ in the direction of radiation propagation vector $\vec{\Omega}$, and per unit time interval dt. It can be expressed mathematically as

$$I_{\nu}(\vec{r},\vec{\Omega},t) = \frac{dE_{\nu}}{dA\cos\theta dt d\nu d\Omega} \left[W.m^{-2}.Hz^{-1}.sr^{-1} \right]$$
(17.1)

where θ is the angle between the vector normal to the surface dAand $\vec{\Omega}$, as shown in the figure below. So, $dA \cos \theta$ is the area of the projection of the surface $d\vec{A}$ on the plane perpendicular to the direction $\vec{\Omega}$. The specific intensity gives the amount of energy in the **direction** of the propagation of the radiation ray. It is almost a complete description of of the radiant energy. As we will see quantities like the flux, and the radiant energy density, radiation pressure can be all obtained from the specific intensity. The only thing that I_{ν} does not include is the information about the polarization.

We can transform from "per unit frequency" to "per unit wavelength" by noting that, for a frequency -dependent quantity, Q_{ν} , we can write

$$\mathcal{Q}_{\nu}|d\nu| = \mathcal{Q}_{\lambda}|d\lambda| \tag{17.2}$$

from which it follows that

$$Q_{\lambda} = \frac{c}{\lambda^2} Q_{\nu} \tag{17.3}$$

where we used the relation $\lambda = c/\nu$ between the wave length and frequency. So, specific intensity in "per unit wavelength" reads

$$I_{\lambda} = \frac{c}{\lambda^2} I_{\nu} \, d\nu \quad \left[W.m^{-2}.m^{-1}.sr^{-1} \right] \tag{17.4}$$



Figure 10

Consider two infinitesimal elements of surfaces dA_1 and dA_2 at a distance r apart. Let a ray of radiation passing through dA_1 and dA_2 at their centers p_1 and p_1 , respectively. if there is no radiation reflection or absorption, i.e. in free space, the radiant energy passing through both surfaces must be the same. Thus, we have

$$(I_{\nu})_1 \cos\theta_1 dA_1 dt d\nu d\Omega_1 = (I_{\nu})_2 \cos\theta_2 dA_2 dt d\nu d\Omega_2 \qquad (17.5)$$

where $d\Omega_1$ is the subtended by dA_2^{156} as seen from p_1 , and $d\Omega_2$ is the subtended by dA_1 as seen from p_2 . Then,

$$d\Omega_1 = \frac{dA_2 \cos \theta_2}{r^2}; \qquad d\Omega_2 = \frac{dA_1 \cos \theta_2}{r^2}$$
(17.6)

Substituting the expressions of the solid angles into (17.5), yields

$$(I_{\nu})_1 = (I_{\nu})_2 \tag{17.7}$$

So, in the absence of sources and sinks, the surface brightness of an object is independent of its distance from the detector¹⁵⁷.

The total intensity can be obtained by integrating over the frequency (or, equivalently over wavelengths if one uses I_{λ}):

$$\mathcal{I} = \int_0^\infty I_\nu \, d\nu = \int_0^\infty I_\lambda \, d\lambda \quad \left[W.m^{-2}.m^{-1}.sr^{-1} \right] \quad (17.8)$$

 $^{^{155}}dA$ could be the physical surface of a detector, the source, or an imaginary surface anywhere along the ray of radiation.

¹⁵⁶Yes, it's dA_2 .

 $^{^{157}}$ In cosmology, this is not true, since there will be a red shift in the frequency of the emitted radiation. We neglect this effect here.

It is often useful to average the specific intensity over the solid angle, and define the define the **mean intensity**, J_{ν} , as

$$J_{\nu} = \frac{\int_{\Omega} I_{\nu} \, d\Omega}{\int \Omega} = \frac{1}{4\pi} \int_{\Omega} I_{\nu} \, d\Omega \tag{17.9}$$

or, equivalently,

$$J_{\nu} = \frac{1}{4\pi} \int_{0}^{2\pi} \int_{-1}^{1} I_{\nu}(\mu, \phi) \ d\mu \ d\phi \tag{17.10}$$

where we introduced the variable $\mu = \cos \theta$. In particular, if the radiation is independent of the azimuthal angle ϕ but not θ , then

$$J_{\nu} = \frac{1}{2} \int_{-1}^{1} I_{\nu}(\mu, \phi) \ d\mu \tag{17.11}$$

Obviously, if I_{ν} is independent of both ϕ and θ , i.e. isotropic, then we have

$$J_{\nu}^{(isotropic)} = I_{\nu}^{(isotropic)} \tag{17.12}$$

2. Radiation Flux Density (F_{ν}) :

The radiation flux density, or simply radiation flux, is the rate of energy flowing across unit area (e.g. at a detector), from all directions, per unit time, and per unit frequency. Mathematically it reads

$$F_{\nu} = \frac{\int_{\Omega} dE_{\nu}}{dAdtd\nu} = \int I_{\nu} \cos\theta d\Omega \quad \left[W.m^{-2}.Hz^{-1}\right]$$
(17.13)

This is the quantity that is measured in the detector¹⁵⁸. Equation (17.13) can expressed as

$$F_{\nu} = \int_{0}^{2\pi} \int_{-1}^{1} I_{\nu}(\mu, \phi) \ \mu \ d\mu \ d\phi \quad \left[W.m^{-2}.Hz^{-1} \right] \qquad (17.14)$$

¹⁵⁸ In astronomy, the unit $W.m^{-2}.Hz^{-1}$ is much too big for the measured flux density. Instead a new unit called **Jansky** and denoted by Jy was introduced, where $1 Jy = 10^{-26} W.m^{-2}.Hz^{-1}$. For example, the measured flux of the cosmic microwave background radiation (CMB) is 0.23 Jy. The unit was named after Karl Jansky, an American physicist and radio engineer who in August 1931 first discovered radio waves emanating from the Milky Way.

In the case where the intensity is independent of the azimuthal angle, we have

$$F_{\nu} = 2\pi \int_{-1}^{1} I_{\nu}(\mu) \ \mu \ d\mu \tag{17.15}$$

Often, the radiation impinges an area from above or below. Hence, it is convenient to divide the contributions to the flux into:

(a) Upward-directed (i.e., emitted) from a surface $(0 \le \theta \le \pi/2)$:

$$F_{\nu}^{(+)} = \int_{0}^{2\pi} \int_{0}^{\pi/2} I_{\nu} \cos\theta \ d\theta d\phi \qquad (17.16)$$

and (a) Downward-directed (i.e., incident) on a surface $(\pi/2 \le \theta \le 0)$:

$$F_{\nu}^{(-)} = \int_{2\pi}^{0} \int_{0}^{\pi/2} I_{\nu} \cos\theta \ d\theta d\phi \qquad (17.17)$$

with the net flux being

$$F_{\nu} = F_{\nu}^{(+)} - F_{\nu}^{(-)} \tag{17.18}$$

If the specific intensity is independent of ϕ , then

$$F_{\nu} = F_{\nu}^{(+)} = 2\pi \int_{0}^{1} I_{\nu}(\mu) \ \mu \ d\mu \qquad (17.19)$$

If in addition, there is no θ dependence over the range $0 \le \theta \le \pi/2$, then the flux of radiation from the emitting source is

$$F_{\nu} = \pi I_{\nu} \tag{17.20}$$

As an example, consider a detector placed at a distance D from a star of spherical shape with radius R and uniform brightness B. Clearly such a star is an isotropic source of radiation. At the position of the detector, the specific intensity is I_{ν} if the ray intersects the star and zero otherwise, i.e.

$$I_{\nu}^{(detector)} = \begin{cases} B, & 0 \le \theta_*; \\ 0, & \text{otherwise.} \end{cases}$$
(17.21)

where $\theta_* = \sin^{-1} R/D$ is the angle at which a ray from the detector is tangent to the star. Then,

$$F_{\nu}^{(observed)} = 2\pi B \int_{0}^{\theta_{*}} \cos\theta \sin\theta \ d\theta \qquad (17.22)$$

$$=\pi B \sin^2 \theta_* = \pi B \left(\frac{R}{D}\right)^2 \tag{17.23}$$

or, equivalently,

$$F_{\nu}^{(observed)} = \pi B \left(\frac{R}{D}\right)^2 F_{\nu}^{(star)}$$
(17.24)

where $F_{\nu}^{(star)} = F_{\nu}^{(+)}$ is the emitted flux of radiation from the star. Another way to derive the above expression is as follows. The projected area of the annulus of radius r at angle θ as

$$dA = 2\pi r dr = 2\pi R^2 \ \mu \ d\mu \tag{17.25}$$

Thus, the annulus subtends an element of solid angle

$$d\Omega = \frac{dA}{D^2} = 2\pi \left(\frac{R}{D}\right)^2 \ \mu \ d\mu \tag{17.26}$$

So, the total observed flux at the detector is

$$F_{\nu}^{(observed)} = 2\pi \left(\frac{R}{D}\right)^2 \int_0^1 B \ \mu \ d\mu \qquad (17.27)$$
$$= \left(\frac{R}{D}\right)^2 F_{\nu}^{(star)}$$

Thus, while the specific intensity is the same at both the surface of the star and the location of the detector, the flux depends on the distance between them and it follows the inverse square law.

3. Specific Radiation Energy Density (u_{ν}) :

The energy dE_{ν} flowing in an element of a cylindrical volume dV, of area dA and length dl, with a frequency between ν and $(\nu + d\nu)$ in a solid angle $d\Omega$ is

$$dE_{\nu} = u_{\nu}(\Omega) \ dV \ d\Omega \ d\nu$$

$$= u_{\nu} \ dA \ dl \ d\Omega \ d\nu$$
(17.28)

where $u_{\nu}(\Omega)$ is the energy density per unit solid angle. The element of length is the distance travelled by the radiation during the time interval d t,. Hence we can write

$$dE_{\nu} = c \ u_{\nu}(\Omega) \ dA \ d\Omega \ d\nu \ dt \tag{17.29}$$

In terms of specific intensity, the energy dE_{ν} reads $dE_{\nu} = I_{\nu} dA d\Omega d\nu dt$. Thus, we deduce that

$$u_{\nu}(\Omega) = \frac{I_{\nu}}{c} \tag{17.30}$$

Integrating $u_{\nu}(\Omega)$ over solid angle, we obtain the specific energy density u_{ν} as

$$u_{\nu} = \frac{1}{c} \int I_{\nu} \ d\Omega = \frac{4\pi}{c} \ J_{\nu} \tag{17.31}$$

4. Radiation Pressure

Consider radiation with energy ΔE bouncing back and forth between two parallel perfect mirrors of area A and separation d. Each time the radiation bounces off the mirror, at an angle θ it imparts a momentum Δp_{\perp} normal to the mirror, given by

$$\Delta p_{\perp} = \frac{2\Delta E}{c} \cos \theta \tag{17.32}$$

in an interval of time $\Delta t = 2L \cos \theta / c$. So, the contribution to pressure is given by

$$\Delta P = \frac{\Delta p / \Delta t}{\Delta A} = \Delta u \, \cos^2 \theta \qquad (17.33)$$
$$= u_{\nu}(\Omega) \, \Delta \Omega \, \Delta \nu \, \cos^2 \theta$$

where $\Delta u = \Delta E/AL$ is the energy density of the radiation. Using (17.30), and integrating over the frequency and the solid angle, we get

$$P = \frac{1}{c} \int I_{\nu} \, d\nu \, \cos^2 \theta d\Omega \tag{17.34}$$

So, the radiation pressure is related to the second -order momentum of the intensity¹⁵⁹. For isotropic radiation field, we obtain

$$P = \int \frac{4\pi}{c} J_{\nu} \, d\nu = \frac{1}{3} \, u \tag{17.35}$$

- 17.2 The Hamilton-Jacobi Equations
- 17.3 Ehrenfest's Adiabatic Principle
- 17.4 Discovery of the Electron Spin
- 17.5 Clebsch Gordon Coefficients

17.6 Derivation of Cross Section using Wave Packets

The S-matrix element $\langle \phi_b | \hat{S} | \phi_a \rangle$ represents the probability amplitude that a state asymptotic to $|\phi_a\rangle$ in the distant past evolves into a state asymptotic to $|\phi_b\rangle$. The scattering experiments measure the number of particles scattered into an element of solid angle $d\Omega$:

$$\mathcal{N}_{scat}(d\Omega) = \frac{d\sigma}{d\Omega} . n_{inc} \tag{17.36}$$

where n_{inc} is the number of incident particles per unit area in a direction perpendicular to the initial momentum. For a given initial wave packet, the probability of a scattering into a solid a solid angle $d\Omega$ is

$$d\mathcal{P}(|\psi^{(in)}\rangle) = d\Omega \int_0^\infty {p'}^3 dp' |< \mathbf{p}' |\hat{S}|\psi^{(in)}\rangle |^2$$
(17.37)
$$= d\Omega \int_0^\infty {p'}^3 dp' |\int d^3 \mathbf{p} < \mathbf{p}' |\hat{S}| \mathbf{p} > < \mathbf{p} |\psi^{(in)}\rangle |^2$$

The total number of particles scattering into a given cone with solid angle $d\Omega$ will be the sum over the number of incident particles times the

$$M_{\nu}^{(n)} = \frac{1}{2} \int_{-1}^{1} I_{\nu}(\mu) \mu^{n} \ d\mu$$

For instance, the first-order moment is related to the flux and is known as Eddington flux given by

$$H_{\nu} = \frac{1}{2} \int_{-1}^{1} I_{\nu}(\mu) \mu \ d\mu = \frac{F_{\nu}}{4\pi}$$

¹⁵⁹The n-th order momentum of the intensity is defined as

probabilities of scattering into a given cone with solid angle $d\Omega$:

$$d\mathcal{N}_{scat} = \sum_{i} d\mathcal{P}(|\psi_i^{(in)}\rangle) \tag{17.38}$$

Here the index *i* labels the particles in the incident beam. The wave packet $|\psi_i^{(in)}\rangle$ will have an average momentum approximately equal to the beam momentum \bar{P} . However, they differ by impact parameter **b**, which is a 2-dimensional vector in the plane perpendicular to the direction of the incident beam, say *z*-axis. An initial state shifted by an amount **b** relative to a reference state $\langle \mathbf{p} | \psi_0^{(in)} \rangle$ is

$$\langle \mathbf{p} | \psi_{\mathbf{b}}^{(in)} \rangle = e^{i\mathbf{b}\cdot\mathbf{p}} \langle \mathbf{p} | \psi_{0}^{(in)} \rangle$$
(17.39)

We will assume that the impact parameters are distributed uniformly over the cross-sectional area subtended by the target with density n_{inc} . The sum over the states in ?? sum can be replaced by

$$d\mathcal{N}_{scat} = \int d^2 \mathbf{b} \ n_{inc} \ d\mathcal{P}(|\psi_{\mathbf{b}}^{(in)} >)$$
(17.40)

Thus, the differential cross section is given by

$$\frac{d\sigma}{d\Omega} = \int d^{3}\mathbf{b} \int_{0}^{\infty} p'^{2} dp' | \int d^{3}\mathbf{p} \left[-2\pi i \delta(E_{p'} - E_{p}) \right] e^{i\mathbf{b}\cdot\mathbf{p}} < \mathbf{p}' |T^{(+)}(E)|\mathbf{p}|^{2} < \mathbf{p} |\psi_{0}^{(in)} > \mathbf{p}|^{2} < \mathbf{p} |\psi_{0}^{(in)} < \mathbf{p}|^$$

Expanding every thing out gives

$$\frac{d\sigma}{d\Omega} = \int d^{3}\mathbf{b} \int_{0}^{\infty} p'^{2} dp' \int d^{3}\mathbf{p} d^{3}\mathbf{k} 4\pi^{2} \delta(E_{k} - E_{p'}) \delta(E_{p} - E_{p'})$$
(17.41)
$$< \mathbf{p}' |T^{(+)}(E)|\mathbf{p} > < \mathbf{k} |T^{(-)}(E)|\mathbf{p}' > < \mathbf{p} |\psi_{0}^{(in)} > e^{i\mathbf{b}.(\mathbf{p}-\mathbf{k})} < \psi_{0}^{(in)}|\mathbf{k} >$$

The integral over the impact parameter gives

$$\int d^3 \mathbf{b} e^{i\mathbf{b}.(\mathbf{p}-\mathbf{k})} = (2\pi)^2 \delta \left(\mathbf{p}_\perp - \mathbf{k}_\perp\right)$$
(17.42)

The product of the energy δ -functions times $\delta (\mathbf{p}_{\perp} - \mathbf{k}_{\perp})$ reads:

$$\delta (E_k - E_{p'}) \,\delta (E_p - E_{p'}) \,\delta (\mathbf{p}_\perp - \mathbf{k}_\perp) =$$

$$= \delta (E_k - E_p) \,\delta (E_p - E_{p'}) \,\delta (\mathbf{p}_\perp - \mathbf{k}_\perp)$$

$$= 2\mu \delta \left(k_3^2 - p_3^2\right) \,\delta (E_p - E_{p'}) \,\delta (\mathbf{p}_\perp - \mathbf{k}_\perp)$$

$$\simeq \frac{\mu}{\bar{p}_3} \delta^{(3)} \left(\mathbf{k} - \mathbf{p}\right) \,\delta (E_p - E_{p'})$$
(17.43)

where $\bar{\mathbf{p}}$ denotes the average momentum of the incident particles . With these identities, the differential scattering cross-section can be written as

$$\frac{d\sigma}{d\Omega} = \frac{16\pi^2 \mu}{\bar{p}_3} \int {p'}^2 dp' \int d^3 \mathbf{p} \ \delta \left(E_p - E_{p'}\right)$$
(17.44)

$$\times |<\mathbf{p}'| T^{(+)}(E) |\mathbf{p}>|^2 |<\mathbf{p}| \psi_0^{(in)}>|^2$$

For sufficiently sharply peaked wave packet, the transition matrix elements are approximately constant on the region where the momentum of the wave packets are non-vanishing. In this case we have:

$$\frac{d\sigma}{d\Omega} = \frac{16\pi^2 \mu}{\bar{p}_3} |<\bar{\mathbf{p}'}|T^{(+)}(E_{\bar{p}_3})|\bar{\mathbf{p}}>|^2 \int d^3\mathbf{p} |<\mathbf{p}|\psi_0^{(in)}>|^2 \int p'^2 dp' \ \delta \left(E_p \left(1E_p 45\right)\right) d^3\mathbf{p} |<\mathbf{p}|^2 dp' \ \delta \left(E_p 45\right) d^3\mathbf{p} |<\mathbf{p}|^2 dp' \ \delta$$

where $|\bar{\mathbf{p}'}| = |\bar{\mathbf{p}}|$. Using the fact that the wave function $\langle \mathbf{p}|\psi_0^{(in)} \rangle$ is normalizable, and the identity $\int p'^2 dp' \,\delta\left(p^2/2\mu - p'^2/2\mu\right) = 2\pi\mu\bar{p}$, we find that

$$\frac{d\sigma}{d\Omega} = 16\pi^4 \mu^2 \frac{\bar{p}}{\bar{p}_3} |\langle \bar{\mathbf{p}'} | T^{(+)}(E_{\bar{p}_3}) | \bar{\mathbf{p}} \rangle |^2$$
(17.46)

The assumption that the beam is in the z-direction means that $\bar{p}_3 = \bar{p}$, which leads to

$$\frac{d\sigma}{d\Omega} = |(2\pi)^2 \,\mu < \bar{\mathbf{p}}'|T^{(+)}(E_{\bar{p}_3})|\bar{\mathbf{p}} > |^2 \equiv |f(\bar{\mathbf{p}}',\bar{\mathbf{p}})|^2 \qquad (17.47)$$

which gives the on-shell relation between the T-matrix and the scattering amplitude, that is

$$f(\bar{\mathbf{p}}', \bar{\mathbf{p}}) = -\mu \left(2\pi\right)^2 < \bar{\mathbf{p}}' |T^{(+)}(E_{\bar{p}_3})|\bar{\mathbf{p}} >$$
(17.48)

Finally, we notice that the expression of the differential cross-section ??diff can also be cast in the form

$$\frac{d\sigma}{d\Omega} = \frac{(2\pi)^4}{v_{inc}} |<\bar{\mathbf{p}}'|T^{(+)}(E_{\bar{p}_3})|\bar{\mathbf{p}}>|^2 p^2 \frac{dp}{dE}$$
(17.49)

where $v_{inv} = \bar{p}_3/\mu = \bar{p}/\mu$ is the mean speed of the incident beam. A key observation is that as long as the wave packet is sharply peaked, the differential cross section is independent of the shape of the wave packet.

References

- [1] *Quantum Concepts in Physics*, Malcolm Longair; Cambridge University Press; 2013 edition.
- [2] The Conceptual Development of Quantum Mechanics, Max Jammer; Thomas Publishers; Volume 12, 1989.
- [3] Sources of Quantum Mechanics, B. L. Van Der Waerden; NORTH-HOLLAND publishing company; 1966.
- [4] I. Aitchison, D. MacManus, and T. Snyder, Am. J. Phys. 72 (2004).
- [5] Ludwig Boltzmann: The Man Who Trusted Atoms, Carlo Cercignani; Oxford University Press; 2006 edition.
- [6] Passive Microwave Remote Sensing of the Earth: Physical Foundations, Eugene A. Sharkov; Springer; 2003 edition.
- [7] Planck, the Quantum, and the Historians, Clayton Gearhart; Phys. perspect. 4, 170-215 (2002).
- [8] The Official Web Site of the Nobel Prize. Available via: [www.nobelprize.org/nobel_prizes/physics/laureates].
- [9] The Physics of Stars, Ian Howarth. Available via: [http://zuserver2.star.ucl.ac.uk/ idh/].
- [10] Modern Physics, Michael Fowler. Available via: [http://galileo.phys.virginia.edu/classes/252/home.html].
- [11] The Photoelectric Effect, M. Brandl. Available via: [http://www.physnet.org/].
- [12] Thermal Radiation: Kinetic Theory and Thermodynamics, R. Madivanane. Available via: [http://aadhi2ani.synthasite.com/notes-2y.php].
- [13] Friedrich Reinert and Stefan Hufner, New J. Phys. 7, 97 (2005).
- [14] Course on Quantum mechanics, Reinhold Bertlmann, summer semester 2008.
 Available via: [http://homepage.univie.ac.at/reinhold.bertlmann/lectures.html].
- [15] The electron and light-quant from the experimental point of view, Robert Millikan, Nobel Lecture, May 23, 1924.
- [16] Revealing the Magic of Everyday Life, M. K. Yip. Available via: [http://www.physics.hku.hk/ phys0607/lectures/chap06.html].
- [17] http://chemwiki.ucdavis.edu/.
- [18] Compton scattering data, HyperPhysics. Available via: [http://hyperphysics.phy-astr.gsu.edu/hbase/quantum/compdat.html].
- [20] W. Fedak and J. Prentis, "Quantum Jumps and classical harmonics", American Journal of Physics, 70: 332-344 (2002).
- [21] Notes on Quantum Mechanics, K. Schulten, University of Illinois at Urbana Champaign 2000.