

Lecture Notes on Quantum Mechanics - Part I

Yunbo Zhang

Institute of Theoretical Physics, Shanxi University

This is the first part of my lecture notes. I mainly introduce some basic concepts and fundamental axioms in quantum theory. One should know what we are going to do with Quantum Mechanics - solving the Schrödinger Equation.

Contents

I. Introduction: Matter Wave and Its Motion	1
A. de Broglie's hypothesis	2
B. Stationary Schrödinger equation	3
C. Conclusion	4
II. Statistical Interpretation of Wave Mechanics	5
A. Pose of the problem	5
B. Wave packet - a possible way out?	6
C. Born's statistical interpretation	7
D. Probability	8
1. Example of discrete variables	8
2. Example of continuous variables	9
E. Normalization	11
III. Momentum and Uncertainty Relation	12
A. Expectation value of dynamical quantities	12
B. Examples of uncertainty relation	14
IV. Principle of Superposition of states	15
A. Superposition of 2-states	15
B. Superposition of more than 2 states	16
C. Measurement of state and probability amplitude	16
1. Measurement of state	16
2. Measurement of coordinate	17
3. Measurement of momentum	17
D. Expectation value of dynamical quantity	18
V. Schrödinger Equation	19
A. The quest for a basic equation of quantum mechanics	19
B. Probability current and probability conservation	20
C. Stationary Schrödinger equation	22
VI. Review on basic concepts in quantum mechanics	25

I. INTRODUCTION: MATTER WAVE AND ITS MOTION

The emergence and development of quantum mechanics began in early years of the previous century and accomplished at the end of the twentieth years of the same century. We will not trace the historical steps since it is a long story. Here we try to access the theory by a way that seems to be more "natural" and more easily conceivable.

A. de Broglie's hypothesis

Inspiration: Parallelism between light and matter

Wave: frequency ν, ω , wavelength λ , wave vector $k \dots \dots$

Particle: velocity v , momentum p , energy $\varepsilon \dots \dots$

- Light is traditionally considered to be a typical case of wave. Yet, it also shows (possesses) a corpuscle nature - light photon. For *monochromatic* light wave

$$\begin{aligned}\varepsilon &= h\nu = \hbar\omega \\ p &= \frac{h\nu}{c} = \frac{h}{\lambda} = \hbar k, \quad \left(k = \frac{2\pi}{\lambda}\right)\end{aligned}$$

- Matter particles should also possess another side of nature - the wave nature

$$\begin{aligned}\varepsilon &= h\nu = \hbar\omega \\ p &= \frac{h}{\lambda} = \hbar k\end{aligned}$$

This is called de Broglie's Hypothesis and is verified by all experiments. In the case of non-relativistic theory, the de Broglie wavelength for a free particle with mass m and energy ε is given by

$$\lambda = h/p = h/\sqrt{2m\varepsilon}$$

The state of (micro)-particle should be described by a **wave function**. Here are some examples of state functions:

1. Free particle of definite momentum and energy is described by a monochromatic traveling wave of definite wave vector and frequency

$$\begin{aligned}\psi_1 &= A' \cos\left(\frac{2\pi}{\lambda}x - 2\pi\nu t + \varphi_0\right) \\ &= A' \cos(kx - \omega t + \varphi_0) \\ &= A' \cos\left(\frac{1}{\hbar}px - \frac{1}{\hbar}\varepsilon t + \varphi_0\right)\end{aligned}$$

Replenish an imaginary part

$$\psi_2 = A' \sin\left(\frac{1}{\hbar}px - \frac{1}{\hbar}\varepsilon t + \varphi_0\right),$$

we get the final form of wave function

$$\psi = \psi_1 + i\psi_2 = A'e^{i\varphi_0}e^{\frac{i}{\hbar}px}e^{-\frac{i}{\hbar}\varepsilon t} = Ae^{\frac{i}{\hbar}px}e^{-\frac{i}{\hbar}\varepsilon t}$$

which is the wave picture of motion of a free particle. In 3D we have

$$\psi = Ae^{\frac{i}{\hbar}\vec{p}\cdot\vec{r}}e^{-\frac{i}{\hbar}\varepsilon t}$$

2. Hydrogen atom in the ground state will be shown later to be

$$\psi(\mathbf{r}, t) = \left(\frac{1}{\pi a_0^3}\right)^{1/2} e^{-\frac{r}{a_0}} e^{-\frac{i}{\hbar}E_1 t}$$

This wave function shows that the motion of electron is in a "standing wave" state. This is the wave picture of above state.

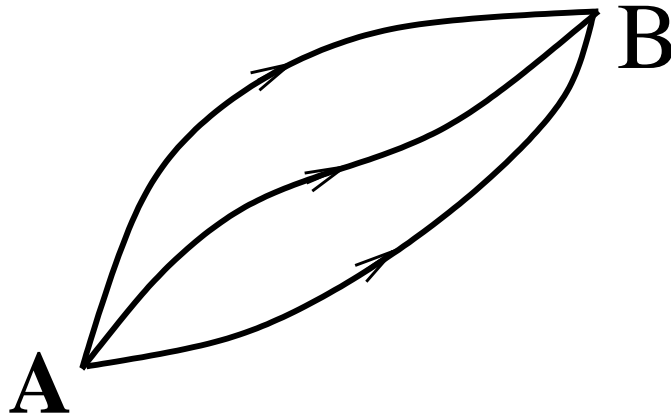
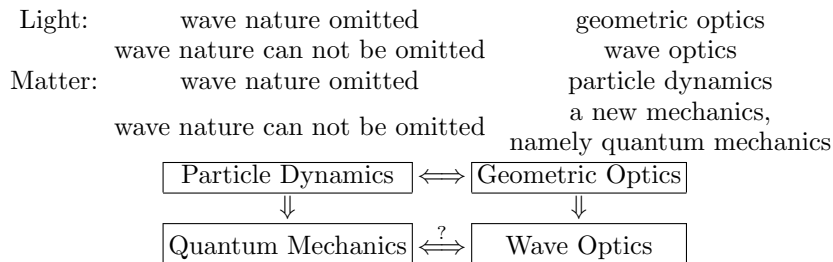


FIG. 1: Maupertuis' Principle.

B. Stationary Schrödinger equation

The parallelism between light and matter can go further



Here we make comparison between light propagation of monochromatic light wave and wave propagation of monochromatic matter wave

Light wave	geometric optics	Fermat's principle
	wave propagation	Helmholtz equation
Matter wave	particle dynamics	Principle of least action
	wave propagation	Presently unknown

Now consider light wave propagation in a non-homogeneous medium

$$light\ path = \int_A^B n(\vec{r}) ds$$

Fermat's principle

$$\delta \int_A^B n(\vec{r}) ds = 0$$

For a particle moving in a potential field $V(\vec{r})$, the principle of least action reads

$$\delta \int_A^B \sqrt{2mT} ds = \delta \int_A^B \sqrt{2m(E - V(\vec{r}))} ds = 0$$

The corresponding light wave equation (Helmholtz equation) is

$$\left(\nabla^2 - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \right) u(\vec{r}, t) = 0$$

which after the separation of variables reduces to

$$\nabla^2\psi + \frac{n^2\omega^2}{c^2}\psi = 0.$$

Here we note that ω is a constant. Thus we arrived at a result of comparison as follows

$$\begin{array}{ccc} \boxed{\delta \int_A^B \sqrt{2m(E - V(\vec{r}))} ds = 0} & \iff & \boxed{\delta \int_A^B n(\vec{r}) ds = 0} \\ \downarrow & & \downarrow \\ \boxed{\text{Presently unknown}} & \stackrel{?}{\iff} & \boxed{\nabla^2\psi + \frac{n^2\omega^2}{c^2}\psi = 0} \\ \downarrow & & \\ \boxed{\text{Presumed to be of the form}} & & \\ \nabla^2\psi + An^2\psi = 0 & & \end{array}$$

Here A is an unknown constant, and the expression $\sqrt{2mT}$ plays the role of "index of refraction" for the propagation of matter waves. The unknown equation now can be written as

$$\nabla^2\psi + A[2m(E - V(\vec{r}))]\psi = 0$$

Substitute the known free particle solution

$$\begin{aligned} \psi &= e^{\frac{i}{\hbar}\vec{p}\cdot\vec{r}} \\ E &= \frac{1}{2m}\vec{p}^2 \\ V(\vec{r}) &= 0 \end{aligned}$$

into the above equation. We find the unknown constant A equals to $1/\hbar^2$, therefore we have

$$\nabla^2\psi + \frac{2m}{\hbar^2}[E - V(\vec{r})]\psi = 0$$

for the general case. It is often written in a form

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

and bears the name "**Stationary Schrödinger Equation**"

C. Conclusion

By a way of comparison, we obtained the equation of motion for a particle with *definite* energy moving in an external potential field $V(\vec{r})$ - the Stationary Schrödinger Equation. From the procedure we stated above, here we stressed on the "wave propagation" side of the motion of micro-particle. $\sqrt{2mT} = \sqrt{2m(E - V(\vec{r}))}$ is treated as "refraction index" of matter waves.

It may happen for many cases that in some spatial districts E is less than $V(\vec{r})$, i.e., $E < V(\vec{r})$. What will happen in such cases?

- Classical mechanics: particles with total energy E can not arrive at places with $E < V(\vec{r})$.
- Matter waves (Q.M.): matter wave can propagate into districts $E < V(\vec{r})$, but in that cases, the refraction index becomes imaginary.

About Imaginary refraction index: For light propagation, imaginary refraction index means dissipation, light wave will be attenuated in its course of propagation. For matter waves, no meaning of dissipation, but matter wave will be attenuated in its course of propagation.

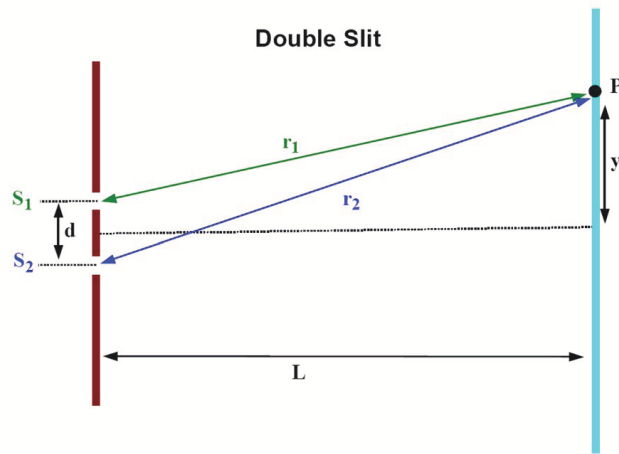


FIG. 2: Double slit experiments.

II. STATISTICAL INTERPRETATION OF WAVE MECHANICS

People tried hard to confirm the wave nature of micro-particles, and electron waves were first demonstrated by measuring diffraction from crystals in an experiment by Davison and Germer in 1925. They scattered electrons off a Nickel crystal which is the first experiment to show matter waves 3 years after de Broglie made his hypothesis. Series of other experiments provided more evidences, such as the double slit experiments using different particle beams: photons, electrons, neutrons, etc. and X-ray (a type of electromagnetic radiation with wavelengths of around 10^{-10} meters). Diffraction off polycrystalline material gives concentric rings instead of spots when scattered off single crystal.

Wave function is a complex function of its variables

$$\psi(x, t) = Ae^{\frac{i}{\hbar}(px - Et)}$$

$$\psi(r, \theta, \phi, t) = \frac{1}{\sqrt{\pi a_0^3}} e^{-\frac{r}{a_0}} e^{-\frac{i}{\hbar} E_1 t}$$

1. Dynamical equation governing the motion of micro-particle is by itself a equation containing imaginary number
2. The wave function describing the state of micro-particle must fit the general theory frame of quantum theory (operator formalism) - requirement of homogeneity of space. This means, the symmetry under a translation in space $\mathbf{r} \rightarrow \mathbf{r} + \mathbf{a}$, where \mathbf{a} is a constant vector, is applicable in all isolated systems. Every region of space is equivalent to every other, or physical phenomena must be reproducible from one location to another.

A. Pose of the problem

What kind of wave it is?

- Optics: Electromagnetic wave

$$\mathbf{E}(x, t) = \hat{y}_0 E_0 e^{i(\frac{2\pi}{\lambda}x - 2\pi\nu t)} = \hat{y}_0 E_0 \overbrace{e^{i(kx - \omega t)}}^{\text{wave propagating}}$$

E_0 – amplitude \rightarrow field strength

Intensity $E_0^2 \rightarrow$ energy density

- Acoustic wave

$$\mathbf{U}(x, t) = \hat{y}_0 U_0 e^{i(\frac{2\pi}{\lambda}x - 2\pi\nu t)}$$

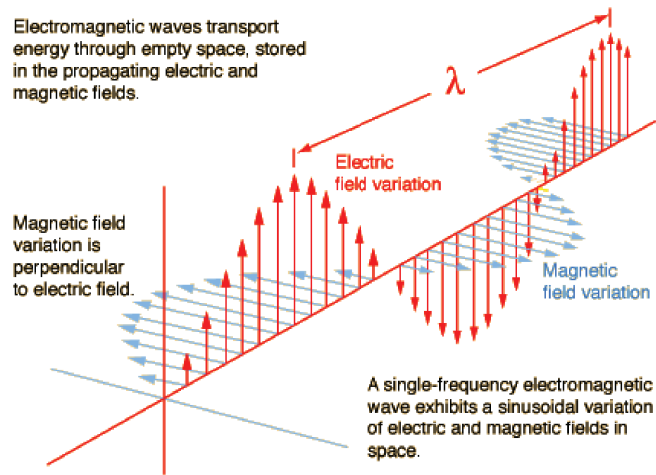


FIG. 3: Electromagnetic Wave Propagation.

U_0 – amplitude \rightarrow mechanical displacement

Intensity $U_0^2 \rightarrow$ energy density

- Wave function

$$\psi(x, t) = Ae^{i\left(\frac{2\pi}{\lambda}x - 2\pi\nu t\right)} = Ae^{\frac{i}{\hbar}(px - Et)}$$

Scalar wave, Amplitude - $A \rightarrow ?$

Intensity - $A^2 \rightarrow ?$

Early attempt: Intensity \rightarrow material density, particle mass distributed in wave. But wave is endless in space, how can it fit the idea of a particle which is local.

B. Wave packet - a possible way out?

particle = wave packet - rain drop

e^{ik_0x} – an endless train, How can it be connected with particle picture?

Two examples of localized wave packets

- Superposition of waves with wave number between $(k_0 - \Delta k)$ and $(k_0 + \Delta k)$ - square packet

$$\varphi(k) = \begin{cases} A, & k_0 - \Delta k < k < k_0 + \Delta k \\ 0, & \text{elsewhere} \end{cases}$$

$$\begin{aligned} \psi(x) &= \frac{1}{\sqrt{2\pi}} \int \varphi(k) e^{ikx} dk \\ &= \frac{1}{\sqrt{2\pi}} \int_{k_0 - \Delta k}^{k_0 + \Delta k} A e^{ikx} dk \\ &= \frac{1}{\sqrt{2\pi}} 2A \underbrace{\frac{\sin(\Delta kx)}{x}}_{\text{Amplitude}} e^{ik_0x} \end{aligned}$$

Both wave number and spatial position have a spread - uncertainty relation

$$\Delta x \sim \pi / \Delta k,$$

$$\Delta x \cdot \Delta k \approx \pi$$

$$\Delta x \cdot \Delta p \approx \pi \hbar$$

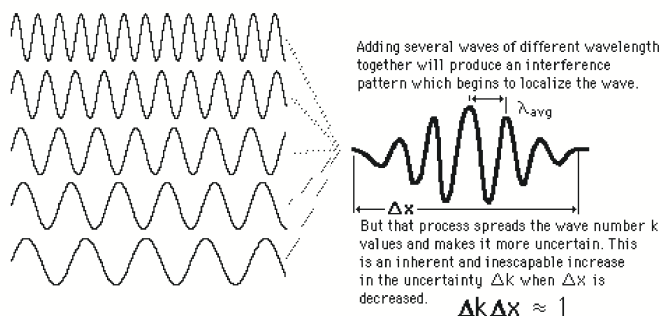


FIG. 4: Wave packet formed by plane waves with different frequency.

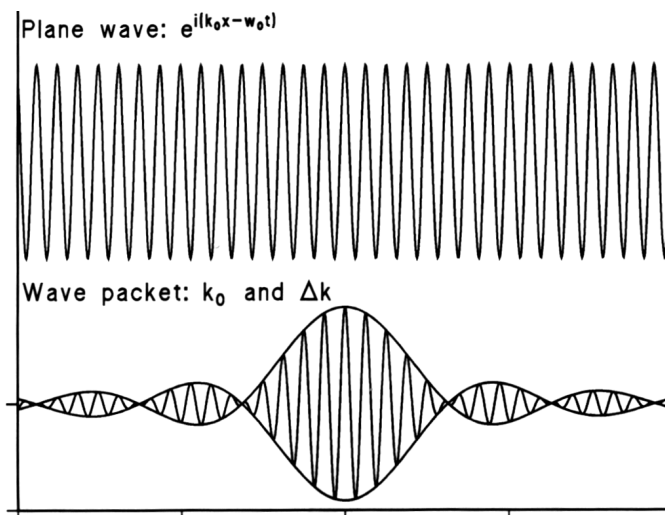


FIG. 5: Spread of a wave packet.

- Superposition of waves with wave number in a Gaussian packet

$$\varphi(k) = \left(\frac{2\alpha}{\pi}\right)^{1/4} e^{-\alpha(k-k_0)^2}$$

Problem 1 The Fourier transformation of $\varphi(k)$ is also Gaussian. It is the best one can do to localize a particle in position and momentum spaces at the same time. Find the root mean square (RMS) deviation $\Delta x \cdot \Delta p = ?$

Problem: Wave packet is unstable - The waves have different phase velocity and the wave components of a wave packet will disperse even in vacuum. But electrons are localized in the atom (1\AA). This viewpoint emphasize the wave nature of micro-particle while killing its particle nature.

Another extreme viewpoint on wave particle duality is that wave is composed by large amount of particles (as waves in air). However experiments show clearly a single electron possesses wave nature. This viewpoint over-stressed the particle nature.

C. Born's statistical interpretation

A particle by its nature is localized at a point, whereas the wave function is spread out in space. How can such an object represent the state of a particle? The answer is provided by Born's Statistical Interpretation of the wave function, which says that in 1D case $|\psi(x, t)|^2$ gives the probability of finding the particle at point x , at time t - or, more precisely

$$\int_a^b |\psi(x, t)|^2 dx = \left\{ \begin{array}{l} \text{Probability of finding the particle} \\ \text{between } a \text{ and } b, \text{ at time } t. \end{array} \right\}$$

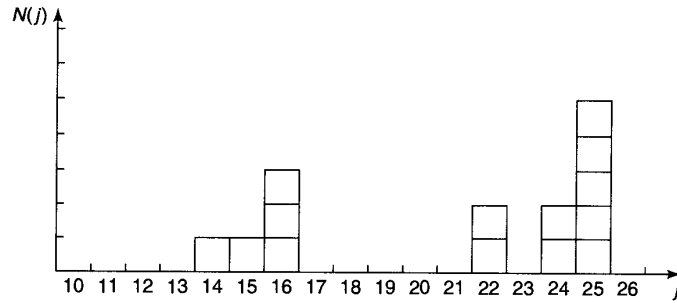


FIG. 6: Histogram for the distribution.

In 3D case, the relative probability for the particle appearing in a volume element $d^3\vec{r}$ is proportional to the intensity of wave in that volume element

$$dw \propto |\psi(\vec{r}, t)|^2 d^3\vec{r} = \underbrace{\psi^*(\vec{r}, t)}_{\text{Probability density}} \overbrace{\psi(\vec{r}, t)}^{\text{Probability amplitude}} d^3\vec{r}$$

D. Probability

Because of the statistical interpretation, probability plays a central role quantum mechanics. So we introduce some notation and terminology.

1. Example of discrete variables

14 people in a room. Let $N(j)$ represent the number of people of age j .

$$N(14) = 1, N(15) = 1, N(16) = 3, N(22) = 2, N(24) = 2, N(25) = 5 \quad (1)$$

while $N(17)$, for instance, is zero. The total number of people in the room is

$$N = \sum_{j=0}^{\infty} N(j)$$

Questions about the distribution

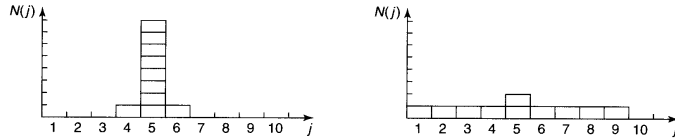
1. If you selected one individual at random from this group, what is the **probability** that this person's age would be 15? *Answer:* One chance in 14, since there are 14 possible choices, all equally likely, of whom only one has this particular age. In general the probability of getting age j is

$$P(j) = \frac{N(j)}{N}$$

Notice that the probability of getting *either 14 or 15* is the sum of the individual probabilities (in this case, $1/7$). In particular, the sum of all the probabilities is 1 - you're certain to get some age:

$$\sum_{j=0}^{\infty} P(j) = 1$$

2. What is the **most probable** age? *Answer:* 25. In general the most probable j is the j for which $P(j)$ is a maximum.
3. What is the **median** age? *Answer:* 23, for 7 people are younger than 23, and 7 are older. In general, the median is that value of j such that the probability of getting a larger result is the same as the probability of getting a smaller result.

FIG. 7: Two histogram with different σ .

4. What is the **average** (or **mean**) age? *Answer:*

$$\frac{(14) + (15) + 3(16) + 2(22) + 2(24) + 5(25)}{14} = \frac{294}{14} = 21$$

In general the average value of j (which we shall write thus: $\langle j \rangle$) is

$$\langle j \rangle = \frac{\sum jN(j)}{N} = \sum_{j=0}^{\infty} jP(j)$$

In quantum mechanics it is called the expectation value. Nevertheless the value is not necessarily the one you can expect if you made a single measurement. In the above example, you will never get 21.

5. What is the **average of the squares of the ages**? *Answer:* You could get $14^2 = 196$, with probability $1/14$, or $15^2 = 225$, with probability $1/14$, or $16^2 = 256$, with probability $3/14$, and so on. The average is

$$\langle j^2 \rangle = \sum_{j=0}^{\infty} j^2 P(j)$$

In general, the average value of some *function* of j is given by

$$\langle f(j) \rangle = \sum_{j=0}^{\infty} f(j)P(j)$$

Beware $\langle j^2 \rangle \neq \langle j \rangle^2$.

Now, there is a conspicuous difference between the following two histograms, even though they have the same median, the same average, the same most probable value, and the same number of elements: The first is sharply peaked about the average value, whereas the second is broad and flat. (The first might represent the age profile for students in a big-city classroom, and the second, perhaps, a rural one-room schoolhouse.) We need a numerical measure of the amount of "spread" in a distribution, with respect to the average. The most effective way to do this is compute a quantity known as the **standard deviation** of the distribution

$$\sigma \equiv \sqrt{\langle (\Delta j)^2 \rangle} = \sqrt{\langle j^2 \rangle - \langle j \rangle^2}$$

$$\Delta j = j - \langle j \rangle$$

For the two distributions in the above figure, we have

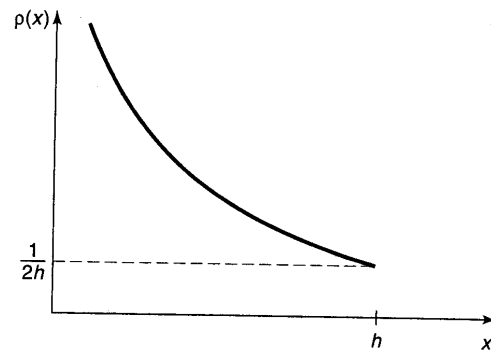
$$\sigma_1 = \sqrt{25.2 - 25} = \sqrt{0.2}$$

$$\sigma_2 = \sqrt{31.0 - 25} = \sqrt{6}$$

2. Example of continuous variables

It is simple enough to generalize to continuous distribution. Technically we need "infinitesimal intervals". Thus

$$\left\{ \begin{array}{l} \text{Probability that an individual (chosen} \\ \text{at random) lies between } x \text{ and } (x + dx) \end{array} \right\} = \rho(x)dx$$

FIG. 8: The probability density in the Example $\rho(x)$.

$\rho(x)$ is the probability of getting x , or probability density. The probability that x lies between a and b (a finite interval) is given by the integral of $\rho(x)$

$$P_{ab} = \int_a^b \rho(x) dx$$

and the rules we deduced for discrete distributions translate in the obvious way:

$$\begin{aligned} 1 &= \int_{-\infty}^{+\infty} \rho(x) dx \\ \langle x \rangle &= \int_{-\infty}^{+\infty} x \rho(x) dx \\ \langle f(x) \rangle &= \int_{-\infty}^{+\infty} f(x) \rho(x) dx \\ \sigma^2 &\equiv \langle (\Delta x)^2 \rangle = \langle x^2 \rangle - \langle x \rangle^2 \end{aligned}$$

Example: Suppose I drop a rock off a cliff of height h . As it falls, I snap a million photographs, at random intervals. On each picture I measure the distance the rock has fallen. *Question:* What is the *average* of all these distance? That is to say, what is the *time average* of the distance traveled?

Solution: The rock starts out at rest, and picks up speed as it falls; it spends more time near the top, so the average distance must be less than $h/2$. Ignoring air resistance, the distance x at time t is

$$x(t) = \frac{1}{2}gt^2$$

The velocity is $dx/dt = gt$, and the total flight time is $T = \sqrt{2h/g}$. The probability that the camera flashes in the interval dt is dt/T , so the probability that a given photograph shows a distance in the corresponding range dx is

$$\frac{dt}{T} = \frac{dx}{gt} \sqrt{\frac{g}{2h}} = \frac{1}{2\sqrt{hx}} dx$$

Evidently the probability density is

$$\rho(x) = \frac{1}{2\sqrt{hx}}, (0 \leq x \leq h)$$

(outside this range, of course, the probability density is zero.)

We can check the normalization of this result

$$\int_0^h \frac{1}{2\sqrt{hx}} dx = \frac{1}{2\sqrt{h}} \left(2x^{1/2} \right)_0^h = 1$$

The average distance is

$$\langle x \rangle = \int_0^h x \frac{1}{2\sqrt{hx}} dx = \frac{1}{2\sqrt{h}} \left(\frac{2}{3} x^{3/2} \right)_0^h = h/3$$

which is somewhat less than $h/2$, as anticipated. Figure shows the graph of $\rho(x)$. Notice that a probability density can be infinite, though probability itself (the integral of ρ) must of course be finite (indeed, less than or equal to 1).

Problem 2 Griffiths, page 12, Problem 1.2, 1.3

E. Normalization

We return now to the statistical interpretation of the wave function, which says that $|\psi(x, t)|^2$ is the probability density for finding the particle at point x , at time t . It follows that the integral of $|\psi(x, t)|^2$ must be 1 (the particle's got to be somewhere)

$$\int_{-\infty}^{+\infty} |\psi(x, t)|^2 dx = 1 \quad (2)$$

If $\psi(x, t)$ is a solution of Schrödinger equation, so too are $A\psi(x, t)$ and $e^{i\alpha}\psi(x, t)$. Here A is any complex constant, α is a phase factor of the wave function. What we must do, then, is pick this undetermined multiplicative factor so as to ensure that Equation (2) is satisfied. This process is called **normalizing** the wave function. For some solutions to the Schrödinger equation, the integral is infinite; in that case no multiplicative factor is going to make it 1. The same goes for the trivial solution $\psi(x, t) = 0$. Such non-normalizable solutions cannot represent particles, and must be rejected. Physically realizable states correspond to the "square-integrable" solutions to Schrödinger's equation.

Example: Problem 1.4 on page 14 of Griffiths.

At time $t = 0$, a particle is represented by the wave function

$$\psi(x, 0) = \begin{cases} A \frac{x}{a}, & \text{if } 0 \leq x \leq a \\ A \frac{b-x}{b-a}, & \text{if } a \leq x \leq b \\ 0, & \text{otherwise} \end{cases}$$

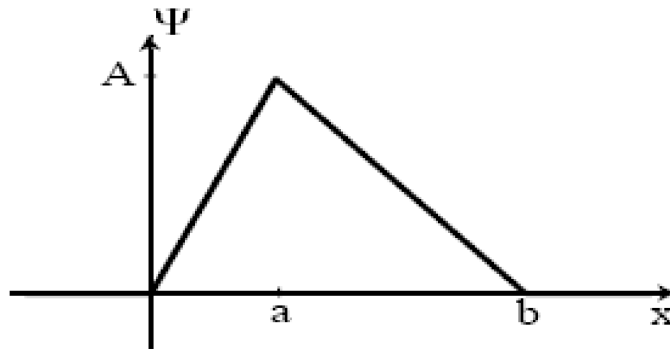
where A, a and b are constants.

- Normalize ψ (that is, find A , in terms of a and b).
- Sketch $\psi(x, 0)$ as a function of x .
- Where is the particle most likely to be found, at $t = 0$?
- What is the probability of finding the particle to the left of a ? Check your result in the limiting cases $b = a$ and $b = 2a$.
- What is the expectation value of x ?

Solution:

-

$$\begin{aligned} 1 &= \frac{|A|^2}{a^2} \int_0^a x^2 dx + \frac{|A|^2}{(b-a)^2} \int_a^b (b-x)^2 dx \\ &= |A|^2 \left\{ \frac{1}{a^2} \left(\frac{x^3}{3} \right)_0^a + \frac{1}{(b-a)^2} \left(-\frac{(b-x)^3}{3} \right)_a^b \right\} \\ &= |A|^2 \left[\frac{a}{3} + \frac{b-a}{3} \right] = |A|^2 \frac{b}{3} \\ \Rightarrow A &= \sqrt{\frac{3}{b}} \end{aligned}$$

FIG. 9: Sketch of $\psi(x, 0)$ as a function of x .(c) At $x = a$.

(d)

$$P = \int_0^a |\psi(x, 0)|^2 dx = \frac{|A|^2}{a^2} \int_0^a x^2 dx = |A|^2 \frac{a}{3} = \frac{a}{b} \begin{cases} P = 1 \text{ if } b = a \\ P = 1/2 \text{ if } b = 2a \end{cases} \checkmark$$

(e)

$$\begin{aligned} \langle x \rangle &= \int x |\psi|^2 dx = |A|^2 \left\{ \frac{1}{a^2} \int_0^a x^3 dx + \frac{1}{(b-a)^2} \int_a^b x (b-x)^2 dx \right\} \\ &= \frac{3}{b} \left\{ \frac{1}{a^2} \left(\frac{x^4}{4} \right)_0^a + \frac{1}{(b-a)^2} \left(b^2 \frac{x^2}{2} - 2b \frac{x^3}{3} + \frac{x^4}{4} \right)_a^b \right\} \\ &= \dots = \frac{2a+b}{4} \end{aligned}$$

Problem 3 Griffiths, page 14, Problem 1.5

III. MOMENTUM AND UNCERTAINTY RELATION

In classical mechanics, we learn that position and momentum are canonical variables to each other. In 1D case, this can be seen from the Fourier transformation and its reverse

$$\begin{aligned} \psi(x, t) &= \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{+\infty} \varphi(p, t) e^{\frac{i}{\hbar} px} dp \\ \varphi(p, t) &= \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{+\infty} \psi(x, t) e^{-\frac{i}{\hbar} px} dx \end{aligned}$$

A. Expectation value of dynamical quantities

For a particle in state $\psi(x)$, the **expectation/mean/average** value of x is

$$\langle x \rangle = \int_{-\infty}^{+\infty} x |\psi(x, t)|^2 dx = \int_{-\infty}^{+\infty} \psi^*(x, t) x \psi(x, t) dx$$

What exactly does this mean? It emphatically does not mean that if you measure the position of one particle over and over again, $\int x |\psi|^2 dx$ is the average of the results you'll get. On the contrary, the first measurement (whose outcome is indeterminate) will collapse the wave function to a spike at the value actually obtained, and the subsequent measurements (if they're performed quickly) will simply repeat that same result. Rather, $\langle x \rangle$ is the

average of measurements performed on particles all in the state ψ , which means that either you must find some way of returning the particle to its original state after each measurement, or else you prepare a whole **ensemble** of particles, each in the same state ψ , and measure the positions of all of them: $\langle x \rangle$ is the average of these results.

In short, *the expectation value is the average of repeated measurements on an ensemble of identically prepared systems*, not the average of repeated measurements on one and the same system.

Similarly one can get the expectation value of the potential energy $V(x)$

$$\langle V(x) \rangle = \int_{-\infty}^{+\infty} V(x) |\psi(x, t)|^2 dx = \int_{-\infty}^{+\infty} \psi^*(x, t) V(x) \psi(x, t) dx$$

But for the momentum p , this is not true

$$\langle p \rangle \neq \int_{-\infty}^{+\infty} \psi^*(x, t) \psi(x, t) p dx$$

Instead, by means of the Fourier transformation of the wave function, $\varphi(p, t)$, we have

$$\begin{aligned} \langle p \rangle &= \int_{-\infty}^{+\infty} \varphi^*(p, t) \varphi(p, t) p dp = \int_{-\infty}^{+\infty} \varphi^*(p, t) p \varphi(p, t) dp \\ &= \int_{-\infty}^{+\infty} dp \left(\int_{-\infty}^{+\infty} dx \frac{1}{\sqrt{2\pi\hbar}} \psi^*(x, t) e^{\frac{i}{\hbar} px} \right) p \varphi(p, t) \\ &= \frac{1}{\sqrt{2\pi\hbar}} \int \int_{-\infty}^{+\infty} dp dx \psi^*(x, t) \underline{e^{\frac{i}{\hbar} px} p} \varphi(p, t) \\ &= \int \int_{-\infty}^{+\infty} dx dp \psi^*(x, t) \left(-i\hbar \frac{d}{dx} \right) \underline{\frac{1}{\sqrt{2\pi\hbar}} e^{\frac{i}{\hbar} px} \varphi(p, t)} \end{aligned}$$

Collecting the underlined terms into $\psi(x, t)$, we have

$$\langle p \rangle = \int_{-\infty}^{+\infty} dx \psi^*(x, t) \left(-i\hbar \frac{d}{dx} \right) \psi(x, t)$$

Definition: For 1D case, the momentum operator is defined as

$$\hat{p} = -i\hbar \frac{d}{dx}$$

and

$$\begin{aligned} \langle p \rangle &= \int_{-\infty}^{+\infty} \psi^*(x, t) \hat{p} \psi(x, t) dx \\ &\neq \int_{-\infty}^{+\infty} \psi^*(x, t) \psi(x, t) \hat{p} dx \end{aligned}$$

It is straightforward to generalize the Fourier transformation result to 3D case

$$\begin{aligned} \psi(\mathbf{r}, t) &= \frac{1}{(2\pi\hbar)^{3/2}} \iiint_{-\infty}^{+\infty} \varphi(\mathbf{p}, t) e^{\frac{i}{\hbar} \mathbf{p} \cdot \mathbf{r}} d^3 \mathbf{p} \\ \varphi(\mathbf{p}, t) &= \frac{1}{(2\pi\hbar)^{3/2}} \iiint_{-\infty}^{+\infty} \psi(\mathbf{r}, t) e^{-\frac{i}{\hbar} \mathbf{p} \cdot \mathbf{r}} d^3 \mathbf{r} \end{aligned}$$

The corresponding momentum operators are defined as

$$\hat{\mathbf{p}} = -i\hbar \nabla$$

with the components

$$\hat{p}_x = -i\hbar \frac{\partial}{\partial x}, \hat{p}_y = -i\hbar \frac{\partial}{\partial y}, \hat{p}_z = -i\hbar \frac{\partial}{\partial z}$$

For simplicity we introduce

$$(\psi, \psi) \equiv \int d\tau \psi^* \psi = \int d\tau |\psi|^2$$

where $d\tau$ means dx for 1D, $dx dy dz$ for 3D, and $dx_1 dy_1 dz_1 \cdots dx_N dy_N dz_N$ for N particles. The normalization takes a very simple form

$$(\psi, \psi) = 1$$

The mean value of a dynamical quantity A is

$$\langle A \rangle = \iiint_{-\infty}^{+\infty} \psi^*(\mathbf{r}, t) \hat{A} \psi(\mathbf{r}, t) d^3 \mathbf{r} = (\psi, \hat{A} \psi)$$

If the wave function is not normalized we should compute it as

$$\langle A \rangle = \frac{(\psi, \hat{A} \psi)}{(\psi, \psi)}$$

Problem 4 Consider a particle in one dimensional system. Both the wave function $\psi(x, t)$ and its Fourier transformation $\varphi(p, t)$ can represent the state of the particle.

*$\psi(x, t)$ is the wave function (or state function, probability amplitude) in position representation
 $\varphi(p, t)$ represents the same state in momentum representation, just like a vector can be expressed in different coordinate systems*

If we use $\varphi(p, t)$ to represent the state, what are the operators x and p ? p is just p itself, but what about x ?

B. Examples of uncertainty relation

Let us examine again the example of Gaussian wave packets

$$\varphi(k) = \left(\frac{2\alpha}{\pi}\right)^{1/4} e^{-\alpha(k-k_0)^2}.$$

It is localized in momentum about $p = \hbar k_0$. We can check the normalization

$$\int_{-\infty}^{+\infty} |\varphi(k)|^2 dk = \sqrt{\frac{2\alpha}{\pi}} \int_{-\infty}^{+\infty} e^{-2\alpha(k-k_0)^2} dk = \sqrt{\frac{2\alpha}{\pi}} \sqrt{\frac{\pi}{2\alpha}} = 1$$

because

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

The Fourier transformation

$$\psi(x) = \left(\frac{1}{2\pi\alpha}\right)^{1/4} e^{ik_0 x} e^{-\frac{x^2}{4\alpha}}$$

The probability density in position space

$$|\psi(x)|^2 = \frac{1}{\sqrt{2\pi\alpha}} e^{-\frac{x^2}{2\alpha}}$$

while in wave vector space

$$|\varphi(k)|^2 = \sqrt{\frac{2\alpha}{\pi}} e^{-2\alpha(k-k_0)^2}$$

From the standard definition of Gaussian distribution

$$P(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-x)^2}{2\sigma^2}}$$

we easily identify the standard deviation in x and k spaces, respectively

$$\begin{aligned}\sigma_x &= \sqrt{\alpha} = \Delta x \\ \sigma_k &= \frac{1}{\sqrt{4\alpha}} = \Delta k\end{aligned}$$

We thus have

$$\begin{aligned}\Delta x \cdot \Delta k &= 1/2 \\ \Delta x \cdot \Delta p &= \hbar/2\end{aligned}$$

More strict calculations show that

$$\Delta x \cdot \Delta p \geq \hbar/2$$

which is Heisenberg's Uncertainty Principle. So the Gaussian wave packets seem to saturate the bound! We need in near future some constants to evaluate the experiments related quantities

$$\begin{aligned}1eV &= 1.602 \times 10^{-12} \text{erg} = 1.602 \times 10^{-19} \text{Joule} \\ \hbar c &= 1973eV \cdot \text{\AA}, \quad a_0 = 0.53\text{\AA}, \quad m_e = 0.51MeV/c^2\end{aligned}$$

and finally the fine structure constant $\alpha = e^2/\hbar c = 1/137$.

IV. PRINCIPLE OF SUPERPOSITION OF STATES

Superposition is a common property of all kinds of linear waves. The most striking characters, the interference and the diffraction, are results of superposition of waves. The distinction between wave motion and particle dynamics lies on whether there exists superposition. The wave nature of micro-particle is recognized and confirmed just through experimental observation of interference of matter waves. It is not surprised that the wave functions (the states) obey superposition principle.

A. Superposition of 2-states

Let ψ_1 and ψ_2 represent two different states of one micro-particle. The principle of superposition asserts that

$$\psi = a_1\psi_1 + a_2\psi_2$$

is also a possible state of the micro-particle - the superposition state.

Here we show an interesting animation of the double slit experiment for bullets, water waves and electrons which is from

<http://www.upscale.utoronto.ca/GeneralInterest/Harrison/DoubleSlit/DoubleSlit.html>

We summarize the main conclusions of the experiments.

1. The probability of an event in an ideal experiment is given by the square of the absolute value of a complex number ψ which is called the probability amplitude

$$\begin{aligned}P &= \text{probability} \\ \psi &= \text{probability amplitude} \\ P &= |\psi|^2\end{aligned}$$

2. When an event can occur in several alternative ways, the probability amplitude for the event is the sum of the probability amplitude for each way considered separately. There is interference:

$$\begin{aligned}\psi &= \psi_1 + \psi_2 \\ P &= |\psi_1 + \psi_2|^2\end{aligned}$$

3. If an experiment is performed which is capable of determining whether one or another alternative is actually taken, the probability of the event is the sum of the probabilities for each alternative. The interference is lost:

$$P = P_1 + P_2$$

The more precisely you know through which slit the electron passes (particle nature), the more unlikely you will see the interference pattern (wave nature). This is the manifestation of Heisenberg uncertainty principle.

B. Superposition of more than 2 states

We have in this case

$$\psi = a_1\psi_1 + a_2\psi_2 + \dots = \sum_i a_i\psi_i$$

where the coefficients a_i determine the relative magnitudes and relative phases among the participating waves.

A special case, yet the most important one, is the superposition of monochromatic plane waves (with one-dimensional case as example). The superposition of several monochromatic plane waves of different p generates a superposition state

$$\psi(x) = c_1e^{ip_1x/\hbar} + c_2e^{ip_2x/\hbar} + \dots$$

Generalization: the value p of monochromatic plane wave $e^{ipx/\hbar}$ changes continuously from $-\infty$ to $+\infty$, the coefficients now are written as a function of p and denoted by $\varphi(p, t)$

$$\psi(x, t) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{+\infty} \varphi(p, t)e^{ipx/\hbar} dp$$

which is nothing but the Fourier transformation.

C. Measurement of state and probability amplitude

1. Measurement of state

The measurement of state is a rather subtle subject. Let us analyze it as follows:

1. There is a microscopic system in a particular state ψ that is to be measured
2. Measurement means that we try to understand the state ψ by means of a process called measurement. This measurement is certainly a macroscopic process and the measured quantity is some kind of physical quantities such as position, momentum, energy, angular momentum, etc.
3. It can be anticipated that the measurement results will be a statistical one.

We will discuss here only measurement of two basic physical quantities, the coordinate x and the momentum p .

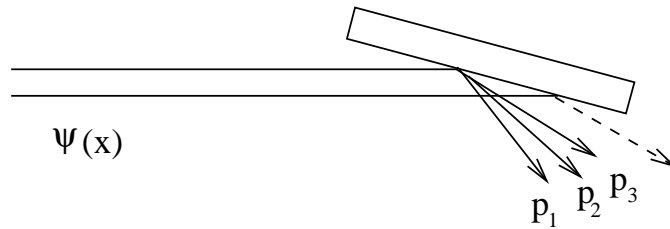


FIG. 10: Bragg diffraction element: Different momentum components are diffracted into different directions.

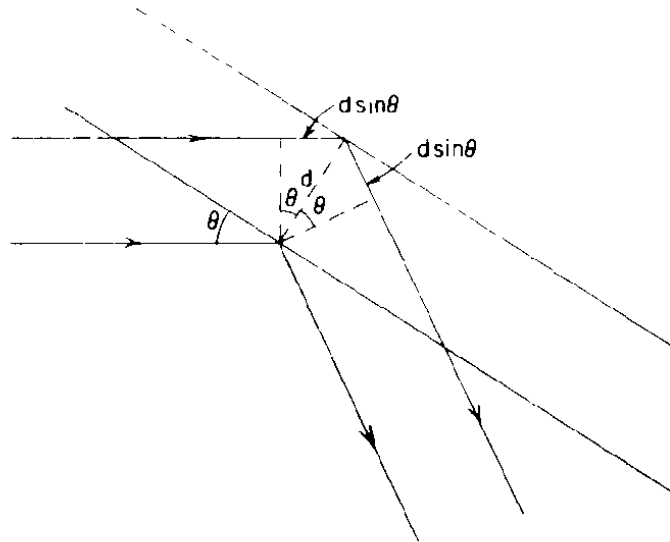


FIG. 11: Scattering of waves by crystal planes.

2. Measurement of coordinate

To simplify the discussion, we restrict our discussion to one-dimensional case. In fact, this is just the Born's statistical interpretation. The probability for the measuring result to be between x and $x + dx$ is

$$P(x)dx = \psi^*(x)\psi(x)dx$$

The generalization to 3D case is straightforward. The probability for the measuring result to be between (x, y, z) and $(x + dx, y + dy, z + dz)$ is $\psi^*(\mathbf{r})\psi(\mathbf{r})d^3\mathbf{r}$.

A new light for comprehending wave function:

Intensity of matter wave described by $|\psi|^2$ at certain point \mathbf{r} is the *probability density* of finding particle there

Amplitude of the matter wave (probability wave) described by $\psi(\mathbf{r}, t)$ at certain point \mathbf{r} is the *probability amplitude* of finding the particle there.

3. Measurement of momentum

In this case, we consider the case that the state $\psi(x)$ is a superposition of monochromatic waves of different but discrete p values

$$\psi(x) = c_1 e^{ip_1 x/\hbar} + c_2 e^{ip_2 x/\hbar} + \dots$$

The apparatus is an idealized equipment that can analyze momenta exactly and diffracted them in different directions, e.g., an idealized Bragg diffraction element described by Bragg formula

$$n\lambda = 2d \sin \theta$$

The probability for the outcome of p is proportional to the intensity of the i -th component $c_i^* c_i = |c_i|^2$. Generalizing to arbitrary 1D state with continuous p values we have

$$\psi(x, t) = \frac{1}{\sqrt{2\pi\hbar}} \int \varphi(p, t) e^{\frac{i}{\hbar} p x} dp$$

where $\varphi(p, t)$ plays the role of coefficients c_i in the case of continuous p . Hence the probability of finding the particle to have its momentum between p and $p + dp$

$$P(p)dp = \varphi^*(p, t)\varphi(p, t)dp = |\varphi(p, t)|^2 dp$$

Further generalization to 3D case

$$\psi(\mathbf{r}, t) = \frac{1}{(2\pi\hbar)^{3/2}} \iiint \varphi(\mathbf{p}, t) e^{\frac{i}{\hbar} \mathbf{p} \cdot \mathbf{r}} d^3\mathbf{p}$$

and the probability for the particle to have its momentum between (p_x, p_y, p_z) and $(p_x + dp_x, p_y + dp_y, p_z + dp_z)$ is

$$P(\mathbf{p})d^3\mathbf{p} = \varphi^*(\mathbf{p}, t)\varphi(\mathbf{p}, t)dp_x dp_y dp_z$$

This can be visualized in momentum space as the probability of \mathbf{p} vector to have its endpoint in volume element $d^3\mathbf{p}$.

Note:

- Once the wave function $\psi(\mathbf{r}, t)$ is known, $\varphi(\mathbf{p}, t)$ is uniquely defined through Fourier transformation. Thus, the probability distribution of \mathbf{p} is known.
- $\psi(\mathbf{r}, t)$ and $\varphi(\mathbf{p}, t)$ have the same function in representing a state in this sense, i.e., there is no priority of $\psi(\mathbf{r}, t)$ over $\varphi(\mathbf{p}, t)$, $\psi(\mathbf{r}, t)$ is called coordinate representation of the state, while $\varphi(\mathbf{p}, t)$ is called the momentum representation of the same state.
- Just as in coordinate representation, $\varphi(\mathbf{p}, t)$ represents a *probability wave* and the numerical value of $\varphi(p, t)$ is the probability amplitude at momentum \mathbf{p} .

D. Expectation value of dynamical quantity

In 1D case we say that the operator x represents position, and the operator $-i\hbar\frac{\partial}{\partial x}$ represents momentum, in quantum mechanics; to calculate expectation values, we **sandwich** the appropriate operator between ψ^* and ψ , and integrate

$$\begin{aligned} \langle x \rangle &= \int_{-\infty}^{+\infty} dx \psi^*(x, t) (x) \psi(x, t) \\ \langle p_x \rangle &= \int_{-\infty}^{+\infty} dx \psi^*(x, t) \left(-i\hbar \frac{\partial}{\partial x} \right) \psi(x, t) \end{aligned}$$

That's cute, but what about other dynamical variables? The fact is, all such quantities can be written in terms of position and momentum. Kinetic energy, for example, is

$$T = \frac{1}{2}mv^2 = \frac{p^2}{2m}$$

and the angular momentum is

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

(the latter, of course, does not occur for motion in one dimension). To calculate the expectation value of *any* such quantity, $Q(x, p)$, we simply replace every p by $-i\hbar\frac{\partial}{\partial x}$, insert the resulting operator between ψ^* and ψ , and integrate

$$\langle Q(x, p) \rangle = \int_{-\infty}^{+\infty} dx \psi^*(x, t) Q\left(x, -i\hbar\frac{\partial}{\partial x}\right) \psi(x, t) \quad (3)$$

Extension to 3D case is straightforward.

There are two cornerstones of Quantum Mechanics: The *state of motion* is described by wave functions while *dynamical quantities* are represented as operators. They cooperate each other into a complete theory. Thus all physical quantities must be operatorized. The consequences of operatorization of physical quantities have profound significance in quantum theory. We also notice the importance of cartesian coordinate in establishing operator system.

V. SCHRÖDINGER EQUATION

A. The quest for a basic equation of quantum mechanics

Premise: The state of system is described by a wave function $\psi(\mathbf{r}, t)$

Quest: Basic equation for the development of $\psi(\mathbf{r}, t)$

What is the rule that $\psi(\mathbf{r}, t)$ must obey and what is the rule that governs the development and evolution of the wave function?

Answer: Ingenious conjecture

Our way: Try to approach the result through a simple example - Free particle, whose wave function is already known

$$\psi(\mathbf{r}, t) = e^{\frac{i}{\hbar}(\mathbf{p}\cdot\mathbf{r}-Et)}$$

We calculate the first derivatives of the wave function with respect to its variables. The equation should be of the form of a partial differential equation

$$\begin{aligned}\frac{\partial\psi}{\partial t} &= -\frac{i}{\hbar}E\psi \Rightarrow i\hbar\frac{\partial\psi}{\partial t} = E\psi \\ \frac{\partial\psi}{\partial x} &= \frac{i}{\hbar}p_x, \quad \frac{\partial^2\psi}{\partial x^2} = -\frac{1}{\hbar^2}p_x^2\psi \\ \frac{\partial\psi}{\partial y} &= \frac{i}{\hbar}p_y, \quad \frac{\partial^2\psi}{\partial y^2} = -\frac{1}{\hbar^2}p_y^2\psi \\ \frac{\partial\psi}{\partial z} &= \frac{i}{\hbar}p_z, \quad \frac{\partial^2\psi}{\partial z^2} = -\frac{1}{\hbar^2}p_z^2\psi\end{aligned}$$

$$\nabla^2\psi = -\frac{1}{\hbar^2}(p_x^2 + p_y^2 + p_z^2)\psi = -\frac{2m}{\hbar^2}E\psi$$

We have for free particle

$$i\hbar\frac{\partial}{\partial t}\psi = E\psi = -\frac{\hbar^2}{2m}\nabla^2\psi$$

For a particle in potential $V(\mathbf{r})$ we replace E in the above equation with

$$E = \frac{1}{2m}\mathbf{p}^2 + V(\mathbf{r})$$

and get a generalized equation

$$i\hbar\frac{\partial\psi(\mathbf{r}, t)}{\partial t} = \left(-\frac{\hbar^2}{2m}\nabla^2 + V(\mathbf{r})\right)\psi(\mathbf{r}, t) \quad (4)$$

Usually it is written in a more compact form

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

with the Hamiltonian operator of the system (the operator of energy of the system) defined as

$$\begin{aligned}\hat{H} &= \frac{1}{2m}\hat{\mathbf{p}}^2 + V(x, y, z) \\ &= \frac{1}{2m}(\hat{p}_x^2 + \hat{p}_y^2 + \hat{p}_z^2) + V(x, y, z)\end{aligned}$$

Equation (4) is called the Schrödinger Equation, one of the basic axioms of quantum mechanics.

We noted that:

- The Schrödinger equation (4) is not deduced, but conjectured. Its correctness is qualified by the accordance of results deduced from it with experimental facts.
- Notice the special role played by Cartesian coordinate system in obtaining the Hamiltonian.
- The discussion can be generalized directly to many-body system

$$i\hbar \frac{\partial}{\partial t} \psi(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \dots) = \hat{H} \psi(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \dots)$$

$$\hat{H} = \sum_i \frac{1}{2m_i} \hat{\mathbf{p}}_i^2 + V(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \dots) + \sum_{ij} U(\mathbf{r}_{ij})$$

where $U(\mathbf{r}_{ij})$ characterize the interaction between particles.

B. Probability current and probability conservation

The physical meaning of $\rho(\mathbf{r}, t) = \psi^*(\mathbf{r}, t)\psi(\mathbf{r}, t)$ is the probability density. From the Schrödinger equation (4) and its complex conjugate

$$\begin{aligned} i\hbar \frac{\partial \psi(\mathbf{r}, t)}{\partial t} &= \left(-\frac{\hbar^2}{2m} \nabla^2 + V(\mathbf{r}) \right) \psi(\mathbf{r}, t) \\ -i\hbar \frac{\partial \psi^*(\mathbf{r}, t)}{\partial t} &= \left(-\frac{\hbar^2}{2m} \nabla^2 + V(\mathbf{r}) \right) \psi^*(\mathbf{r}, t) \end{aligned}$$

we can calculate the time derivative of the probability density

$$\begin{aligned} i\hbar \frac{\partial}{\partial t} (\psi^*(\mathbf{r}, t)\psi(\mathbf{r}, t)) &= -\frac{\hbar^2}{2m} (\psi^* \nabla^2 \psi - \psi \nabla^2 \psi^*) \\ &= -\frac{\hbar^2}{2m} \nabla \cdot (\psi^* \nabla \psi - \psi \nabla \psi^*) \end{aligned}$$

An integration over a volume V gives

$$i\hbar \frac{d}{dt} \iiint_V \rho(\mathbf{r}, t) d^3\mathbf{r} = -\frac{\hbar^2}{2m} \iiint_V \nabla \cdot (\psi^* \nabla \psi - \psi \nabla \psi^*) d^3\mathbf{r}$$

which can be transferred into an integration over the surface \mathfrak{s} enclosed by V

$$i\hbar \frac{d}{dt} \iiint_V \rho(\mathbf{r}, t) d^3\mathbf{r} = -\frac{\hbar^2}{2m} \oint_{\mathfrak{s}} (\psi^* \nabla \psi - \psi \nabla \psi^*) \cdot d\mathfrak{s}$$

and

$$\frac{d}{dt} \iiint_V \rho(\mathbf{r}, t) d^3\mathbf{r} + \oint_{\mathfrak{s}} \left(-\frac{i\hbar}{2m} \right) (\psi^* \nabla \psi - \psi \nabla \psi^*) \cdot d\mathfrak{s} = 0$$

Define now the *current density of probability* as

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

we thus arrive at the continuation equation of probability density

$$\frac{d}{dt} \iiint_V \rho(\mathbf{r}, t) d^3\mathbf{r} + \oint_{\mathfrak{s}} \mathbf{j} \cdot d\mathfrak{s} = 0$$

and the *differential and integral equation of probability conservation*

$$\begin{aligned} \frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{j} &= 0 \\ \frac{d}{dt} \iiint_V \rho(\mathbf{r}, t) d^3\mathbf{r} &= - \oint_{\mathfrak{s}} \mathbf{j} \cdot d\mathfrak{s} \end{aligned}$$

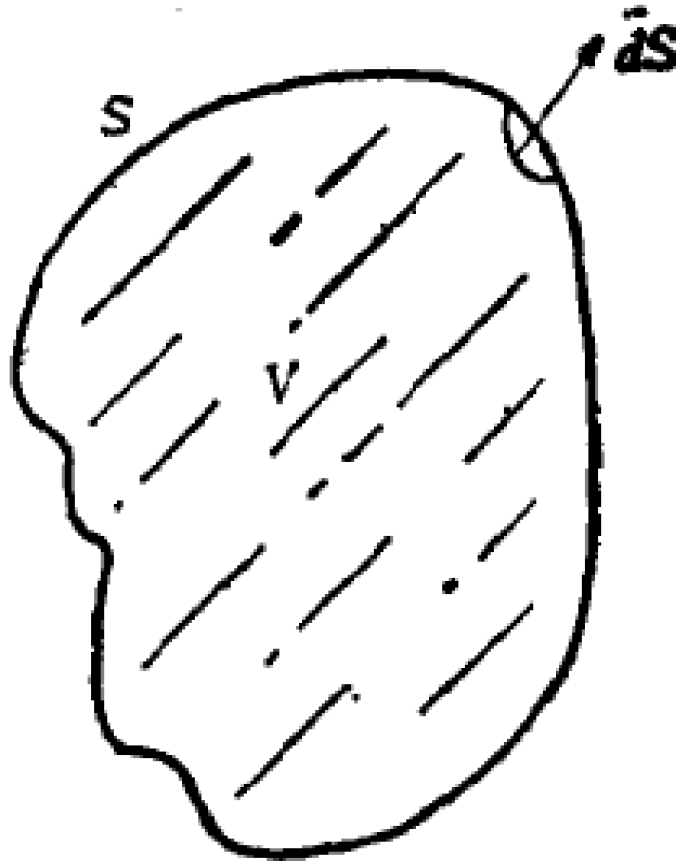


FIG. 12: Current density of probability

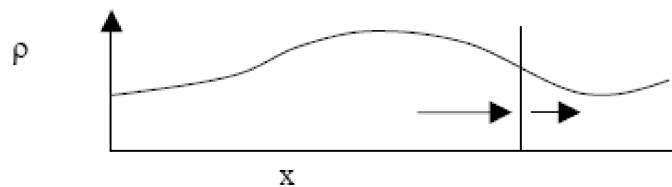


FIG. 13: Conservation of probability.

This equation is telling us that the rate of change of density (the amount of a substance at a point) is equal to (-ve of) the gradient (grad) of the current density. If we write this in one dimension it's perhaps a little easier to see why this relationship is true:

$$\frac{\partial \rho}{\partial t} = -\frac{\partial \mathbf{j}}{\partial x}$$

Essentially, $\partial \mathbf{j} / \partial x$ at a point is the difference between the amount of “stuff” flowing *into* that point and the amount flowing *out*. Since *probability* is conserved, if there is any difference, this must be the rate of change of the amount of stuff at that point i.e. $\partial \rho / \partial t$.

We take again the free particle as an example. The wave function propagating along x

$$\psi = e^{i(kx - \omega t)}$$

leads immediately to a constant probability density

$$\rho = 1 = \text{const.}$$

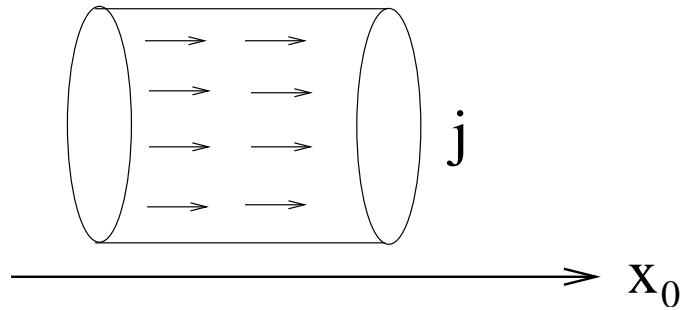


FIG. 14: Example: Free particle

The current density

$$\begin{aligned}\mathbf{j} &= -\frac{i\hbar}{2m} \left(e^{-i(kx-\omega t)} \nabla e^{i(kx-\omega t)} - e^{i(kx-\omega t)} \nabla e^{-i(kx-\omega t)} \right) \\ &= -\frac{i\hbar}{2m} 2ik\mathbf{x}_0 = \frac{\hbar k}{m} \mathbf{x}_0 = v\mathbf{x}_0.\end{aligned}$$

The velocity of classical particle v is just the current density and we have $\nabla \cdot \mathbf{j} = 0$ which is consistent with the result $\partial\rho/\partial t = 0$.

We can extend the region of interest to the entire space. In this case the wave function ψ must go to zero in the infinity - otherwise it would not be normalizable, so must the current \mathbf{j}

$$\lim_{r \rightarrow \pm\infty} \psi(\mathbf{r}, t) = 0, \quad \lim_{r \rightarrow \pm\infty} \mathbf{j} = 0$$

We thus have

$$\frac{d}{dt} \iiint_{-\infty}^{+\infty} \rho(\mathbf{r}, t) d^3\mathbf{r} = 0$$

which means the conservation of total probability in the entire space.

C. Stationary Schrödinger equation

For many physical models the potential $V(\mathbf{r})$ is independent of time t . In those cases the Schrödinger Equation (we denote the time-dependent wave function by Ψ and stationary solution by ψ in the following)

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

can be solved by the method of **separation of variables** (the physicist's first line of attack on any partial differential equation): We look for solutions that are simple products,

$$\Psi(\mathbf{r}, t) = \psi(\mathbf{r})f(t)$$

where $\psi(\mathbf{r})$ (*lower-case*) is a function of \mathbf{r} alone, and $f(t)$ is a function of t alone. On its face, this is an absurd restriction, and we cannot hope to get more than a tiny subset of all solutions in this way. But hang on, because the solutions we do obtain turn out to be of great interest. Moreover, as is typically the case with separation of variables, we will be able at the end to patch together the separable solutions in such a way as to construct the most general solution.

The Schrödinger Equation reads

$$\left[i\hbar \frac{d}{dt} f(t) \right] \psi(\mathbf{r}) = \left\{ \left(-\frac{\hbar^2}{2m} \nabla^2 + V(\mathbf{r}) \right) \psi(\mathbf{r}) \right\} f(t)$$

Or dividing through by $\psi(\mathbf{r})f(t)$

$$\frac{i\hbar \frac{d}{dt} f(t)}{f(t)} = \frac{\left(-\frac{\hbar^2}{2m} \nabla^2 + V(\mathbf{r}) \right) \psi(\mathbf{r})}{\psi(\mathbf{r})}$$

Now the left side is a function of t alone, and the right side is a function of \mathbf{r} alone. The only way this can possibly be true is if both sides are in fact constant - otherwise, by varying t , I could change the left side without touching the right side, and the two would no longer be equal. The separation constant E is introduced in the process of separation of variable and recognized to be the energy of the system. Then we have

$$i\hbar \frac{\partial}{\partial t} f(t) = f(t)E$$

with solution

$$f(t) = e^{-\frac{i}{\hbar}Et}$$

and

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

This is the **time-independent Schrödinger Equation**, or **Stationary Schrödinger Equation**. The natural boundary conditions should be imposed on: first of all the wave function should be single-valued and limited. Furthermore $\psi(\mathbf{r})$ and its spatial derivations should be continuous (there may be exceptions in the above conditions for some artificially simplified $V(\mathbf{r})$)

The following chapters will be devoted to solving the time-independent Schrödinger equation, for a variety of simple potentials. But before we get to that you have every right to ask:

What's so great about separable solutions?

After all, most solutions to the (time-dependent) Schrödinger equation do not take the form $\psi(\mathbf{r})f(t)$. I offer three answers - two of them physical and one mathematical:

- (i) They are **stationary states**. Although the wave function itself, $\Psi(\mathbf{r}, t) = \psi(\mathbf{r})e^{-\frac{i}{\hbar}Et}$ does (obviously) depend on t , the *probability density* $|\Psi(\mathbf{r}, t)|^2 = |\psi(\mathbf{r})|^2$ does not - the time dependence cancels out. The same thing happens in calculating the expectation value of any dynamical variable; Equation (3) reduces to

$$\langle Q(x, p) \rangle = \int_{-\infty}^{+\infty} dx \psi^*(x) Q\left(x, -i\hbar \frac{\partial}{\partial x}\right) \psi(x)$$

Summarized in one word, stationary state is stable to all the test, i.e. the result doesn't change with time. If a particle is initially put in a state with energy E , it will stay there all the time, i.e.

$$\Psi(\mathbf{r}, 0) = \psi_E(\mathbf{r}) \Rightarrow \Psi(\mathbf{r}, t) = \psi_E(\mathbf{r})e^{-\frac{i}{\hbar}Et}$$

and moreover

$$\begin{aligned} \rho(\mathbf{r}, t) &= \left(\psi_E(\mathbf{r})e^{-\frac{i}{\hbar}Et} \right)^* \psi_E(\mathbf{r})e^{-\frac{i}{\hbar}Et} \\ &= \psi_E^*(\mathbf{r})\psi_E(\mathbf{r}) = \rho(\mathbf{r}) \end{aligned}$$

$$\begin{aligned} \mathbf{j}(\mathbf{r}, t) &= -\frac{i\hbar}{2m} \left\{ \begin{array}{l} \psi_E^*(\mathbf{r})e^{\frac{i}{\hbar}Et} \nabla \left(\psi_E(\mathbf{r})e^{-\frac{i}{\hbar}Et} \right) \\ -\psi_E(\mathbf{r})e^{-\frac{i}{\hbar}Et} \nabla \left(\psi_E^*(\mathbf{r})e^{\frac{i}{\hbar}Et} \right) \end{array} \right\} \\ &= -\frac{i\hbar}{2m} \{ \psi_E^*(\mathbf{r})\nabla \psi_E(\mathbf{r}) - \psi_E(\mathbf{r})\nabla \psi_E^*(\mathbf{r}) \} = \mathbf{j}(\mathbf{r}) \end{aligned}$$

$$\begin{aligned} \langle p_x \rangle &= \iiint \left(\psi_E^*(\mathbf{r})e^{\frac{i}{\hbar}Et} \right) \left(-i\hbar \frac{\partial}{\partial x} \right) \left(\psi_E(\mathbf{r})e^{-\frac{i}{\hbar}Et} \right) d^3\mathbf{r} \\ &= \iiint \psi_E^*(\mathbf{r}) \left(-i\hbar \frac{\partial}{\partial x} \right) \psi_E(\mathbf{r}) d^3\mathbf{r} \end{aligned}$$

Problem 5 Show that the expectation value of the velocity is equal to the time derivative of the expectation value of position

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

And it is customary to work with momentum ($p = mv$)

$$\langle p \rangle = m \frac{d\langle x \rangle}{dt}$$

For stationary states, every expectation value is constant in time. In particular, $\langle x \rangle$ is constant, and hence $\langle p \rangle = 0$. Nothing ever happens in a stationary state. (page 16 Griffiths)

- (ii) They are states of *definite total energy*. In classical mechanics, the total energy (kinetic plus potential) is called the **Hamiltonian**:

$$H(x, p) = \frac{p^2}{2m} + V(x)$$

The corresponding Hamiltonian operator, obtained by the canonical substitution $p \rightarrow -i\hbar \frac{\partial}{\partial x}$, is therefore

$$\hat{H} = -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} + V(x)$$

Thus the time-independent Schrödinger equation can be written as the eigenvalue equation of operator \hat{H}

$$\hat{H}\psi = E\psi$$

which is the same as what we met in the introduction. And the expectation value of the total energy is

$$\langle H \rangle = \int \psi^* \hat{H} \psi dx = E \int |\psi|^2 dx = E \int |\Psi|^2 dx = E$$

(Notice that the normalization of Ψ entails the normalization of ψ .) Moreover,

$$\hat{H}^2 \psi = \hat{H} (\hat{H} \psi) = \hat{H} (E\psi) = E (\hat{H} \psi) = E^2 \psi$$

and hence

$$\langle H^2 \rangle = \int \psi^* \hat{H}^2 \psi dx = E^2 \int |\psi|^2 dx = E^2$$

So the variance of H is

$$\sigma_H^2 = \langle H^2 \rangle - \langle H \rangle^2 = E^2 - E^2 = 0$$

But remember, if $\sigma = 0$, then every member of the sample must share the same value (the distribution has zero spread). *Conclusion: A separable solution has the property that every measurement of the total energy is certain to return the value E .* (That's why I chose that letter for the separation constant.)

- (iii) The general solution is a **linear combination** of separable solutions. As we're about to discover, the stationary Schrödinger equation yields an infinite collection of solutions, each with its associated value of the separation constant

$$\begin{aligned} &\psi_1(\mathbf{r}), \psi_2(\mathbf{r}), \psi_3(\mathbf{r}) \dots \\ &E_1, E_2, E_3 \dots \end{aligned}$$

thus there is a different wave function for each allowed energy

$$\begin{aligned} \Psi_1(\mathbf{r}, t) &= \psi_1(\mathbf{r}) e^{-iE_1 t/\hbar}, \\ \Psi_2(\mathbf{r}, t) &= \psi_2(\mathbf{r}) e^{-iE_2 t/\hbar}, \dots \end{aligned}$$

Superposition of two or more states with different energies is not a stationary state. A much more general solution can be constructed by a linear combination of the separable solutions

$$\Psi(\mathbf{r}, t) = \sum_{n=1}^{\infty} c_n \psi_n(\mathbf{r}) e^{-iE_n t/\hbar}$$

which means essentially the principle of superposition of states. The main point is this: Once you've solved the time-independent Schrödinger equation, you're essentially done; getting from them to the general solution of the time-dependent Schrödinger equation is simple and straightforward.

Problem 6 *Introduction to QM (2ed) by Zeng Jinyan, page 47-48, 2, 3, 5 (homework2.doc)*

Problem 7 *Write down the stationary Schrödinger Equation in momentum representation (1D and 2D).*

VI. REVIEW ON BASIC CONCEPTS IN QUANTUM MECHANICS

7 topics, 5 minutes presentation for each group

- State described by wave function
- Dynamical variables \Rightarrow operators
- Measurement and uncertainty principle
- Schrödinger equation
- Conservation of probability density
- Wave-particle duality
- Average value and standard deviation

Lecture Notes on Quantum Mechanics - Part II

Yunbo Zhang

Institute of Theoretical Physics, Shanxi University

In this chapter we solve the stationary Schrödinger equation for various types of one dimensional potentials, including square potential well, harmonic oscillator, scattering of potential barrier, and δ -potential.

Contents

I. Square Potential Well	2
A. An idealized case - Infinite square potential well	2
1. Solution	2
2. Energy quantization	2
3. Wave function normalization	3
4. Shift of origin of coordinate	4
B. Square potential well of finite depth	5
1. Solution	5
2. Even parity case	5
3. Odd parity case	6
4. Graphical solution of the energy defining equation	7
II. Generic Problem of Quantum Mechanics	8
A. Properties of wave functions	9
B. Construction of the most general solution	10
C. An example	10
III. The Harmonic Oscillator	11
A. Algebraic method: ladder operator	12
1. Commutator	13
2. Ladder operators	14
3. Normalization algebraically: find A_n	16
B. Analytic method: differential equation	18
1. Discussion on the energy quantization	20
2. Discussion on the wave function	20
IV. The Free Particle: Revisited	24
V. Penetration of Potential Barrier	29
A. Case I: Barrier penetration $E < V_0$	29
1. Boundary conditions	30
B. Case II: Barrier penetration $E > V_0$	32
C. Case III: Bound states in finite potential well $-V_0 < E < 0$	33
D. Case IV: Scattering from square well $-V_0 < 0 < E$	33
VI. The δ-function potential	33
A. Bound states and scattering states	33
B. From square potential to δ -potential	35
C. Bound states in δ potential well ($E < 0$)	36
D. Scattering states in δ -potential well ($E > 0$)	38
E. δ -function barrier	39
VII. Summary on Part II	39

I. SQUARE POTENTIAL WELL

We consider a particle moving in a square potential well (Figure 1)

$$V(x) = \begin{cases} 0, & 0 < x < a \\ V_0, & x < 0, x > a \end{cases}$$

and in this case the stationary Schrödinger equation

$$-\frac{\hbar^2}{2m} \frac{d^2}{dx^2} \psi(x) + V(x)\psi(x) = E\psi(x)$$

can be solved exactly.

The physics here is the one-dimensional material-wave propagation in a potential well. We will confine our discussions to cases with energy E less than V_0 .

A. An idealized case - Infinite square potential well

First we solve the idealized model of above potential, that is when $V_0 \rightarrow +\infty$, a infinite square potential well

1. Solution

For the potential

$$V(x) = \begin{cases} 0, & 0 < x < a \\ +\infty, & x < 0, x > a \end{cases}$$

it is easy to know

$$\begin{aligned} -\frac{\hbar^2}{2m} \frac{d^2}{dx^2} \psi(x) &= E\psi(x) \quad \text{for } 0 < x < a \\ \psi(x) &= 0 \quad \text{for } x < 0, x > a \end{aligned}$$

Let $k^2 = 2mE/\hbar^2$, we have

$$\begin{aligned} \frac{d^2}{dx^2} \psi(x) + k^2 \psi(x) &= 0 \\ \Rightarrow \psi(x) &= A \sin(kx + \delta) \end{aligned}$$

A and δ are two constants to be determined shortly.

2. Energy quantization

According to the natural boundary conditions, $\psi(x)$ should be continuous at $x = 0$ and $x = a$. So

$$\psi(0) = \psi(a) = 0$$

which tells us $\delta = 0$, and $ka = n\pi$, with $n = 1, 2, 3 \dots$ (because $A = 0$ and $n = 0$ give trivial, non-normalizable solution $\psi(x) = 0$. Furthermore $n = -1, -2, -3 \dots$ give nothing new since we can absorb the minus sign into A). So the distinct solutions are

$$k_n = \frac{n\pi}{a}, \quad \text{with } n = 1, 2, 3 \dots$$

and

$$\psi_n(x) = A \sin\left(\frac{n\pi}{a}x\right)$$

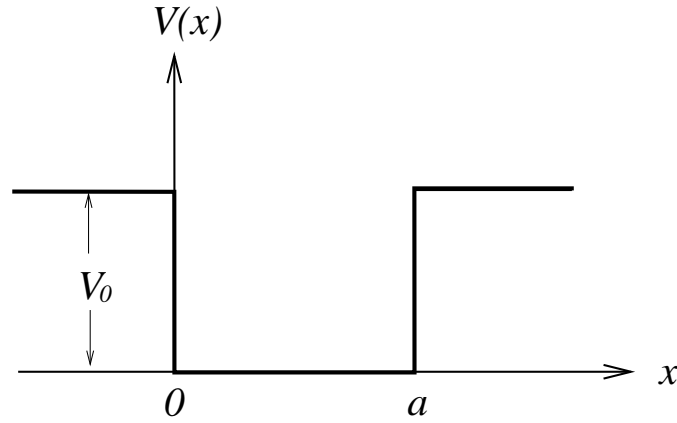


FIG. 1: One dimensional square potential well.

Curiously, the boundary condition at $x = a$ does not determine the constant A , but rather the constant k , and hence the possible values of E :

$$E_n = \frac{\hbar^2 k_n^2}{2m} = n^2 \frac{\pi^2 \hbar^2}{2ma^2}$$

In sharp contrast to the classical case, a quantum particle in the infinite square well cannot have just any old energy—only these special allowed values. *The energy of system (particle) is quantized.*

Note: For infinitely high potential well, waves can't penetrate into the forbidden regions. They are fully reflected and standing waves is established. (Figure 2)

3. Wave function normalization

Well, how do we fix the constant A ? *Answer:* We normalize ψ :

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

This only determines the magnitude of A , but it is simplest to pick the positive real root:

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

(the phase of A carries no physical significance anyway). Inside the well, then, the solutions are

$$\psi_n(x) = \sqrt{\frac{2}{a}} \sin\left(\frac{n\pi}{a}x\right), \quad n = 1, 2, 3 \dots$$

As promised, the time-independent Schrödinger equation has delivered an infinite set of solutions, one for each integer n . The first few of these are plotted in Figure 2; they look just like the standing waves on a string of length a . ψ_1 , which carries the lowest energy, is called the **ground state**:

$$E_1 = \frac{\pi^2 \hbar^2}{2ma^2} \neq 0$$

$$\Psi_1(x, t) = \sqrt{\frac{2}{a}} \sin\left(\frac{\pi x}{a}\right) \exp\left(-i \frac{\pi^2 \hbar}{2ma^2} t\right)$$

the others, whose energies increase in proportion to n^2 , are called **excited states**.

Heisenberg uncertainty relation can be easily checked on the lowest energy state. We know the standard deviation (uncertainty) in position space is obviously $\Delta x \sim a$ and it is also clear $\Delta p \sim \hbar/a$ from the ground state energy $E = p^2/2m \sim \frac{\hbar^2}{2ma^2}$. So $\Delta x \cdot \Delta p \sim \hbar$.

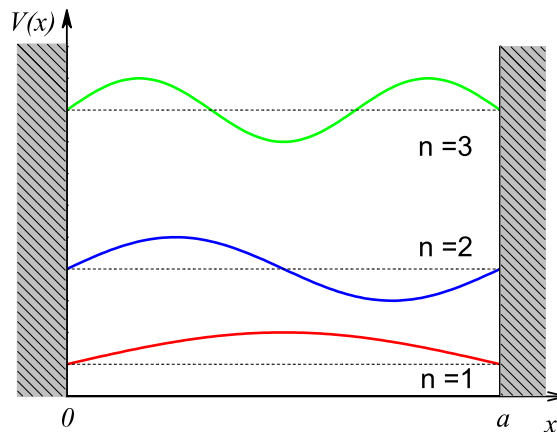


FIG. 2: First 3 wave functions and the corresponding energy levels of infinite one dimensional square potential well.

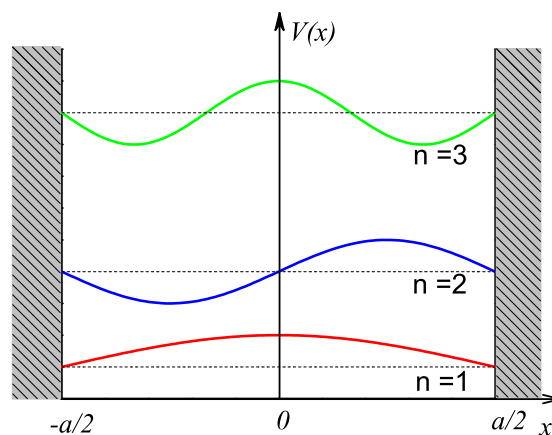


FIG. 3: First 3 wave functions and the corresponding energy levels of infinite one dimensional square potential well - shifted coordinate system.

4. Shift of origin of coordinate

We now shift the origin of coordinate so that the potential takes the form

$$V(x) = \begin{cases} 0, & -\frac{a}{2} < x < \frac{a}{2} \\ +\infty, & |x| > \frac{a}{2} \end{cases}$$

The wave function in the shifted coordinate system are

$$\begin{aligned} \psi_n(x) &= \sqrt{\frac{2}{a}} \sin\left(\frac{n\pi}{a}x - \frac{n\pi}{a}\frac{a}{2}\right) \\ &= \begin{cases} \sqrt{\frac{2}{a}} \cos\left(\frac{n\pi}{a}x\right), & \text{for } n = 1, 3, 5 \dots \\ \sqrt{\frac{2}{a}} \sin\left(\frac{n\pi}{a}x\right), & \text{for } n = 2, 4, 6 \dots \end{cases} \end{aligned}$$

which show a regular symmetry about the origin $x = 0$, called the parity. For $n = 1, 3, 5 \dots$, wave functions are symmetric (even functions) about $x = 0$, and are said possessing even parity. For $n = 2, 4, 6 \dots$, wave functions are

anti-symmetric (odd function) about $x = 0$, they have odd parity. There is a minus sign in the wave functions for the 1st and 2nd excited states compared with the unshifted system, which however, does not lead to big problem because it gives a trivial phase of π .

B. Square potential well of finite depth

Now we consider the situation that the depth of the barrier is finite. We use here the shifted coordinate system.

1. Solution

The potential is now

$$V(x) = \begin{cases} 0, & -\frac{a}{2} < x < \frac{a}{2} \\ V_0, & |x| > \frac{a}{2} \end{cases}$$

$\psi(x)$ for a particle with energy E less than V_0 ($E < V_0$) differs from zero for all x values.

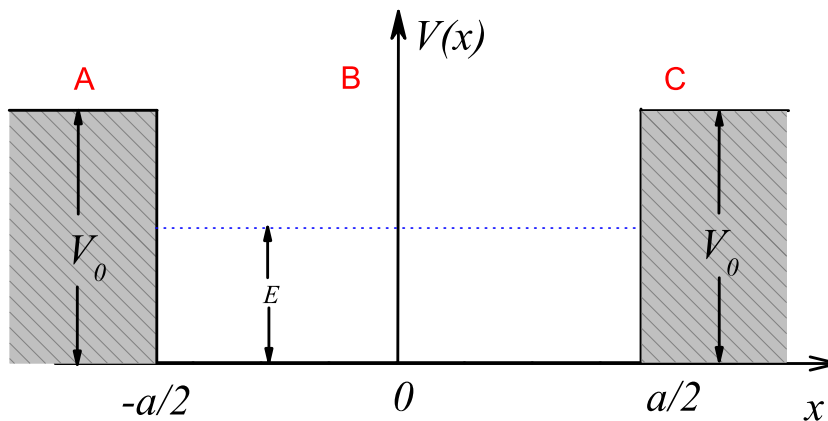


FIG. 4: One dimensional square potential well of finite depth.

Region A or C	Region B
$-\frac{\hbar^2}{2m} \frac{d^2}{dx^2} \psi(x) + V_0 \psi(x) = E \psi(x)$	$-\frac{\hbar^2}{2m} \frac{d^2}{dx^2} \psi(x) + 0 \cdot \psi(x) = E \psi(x)$
$\frac{d^2}{dx^2} \psi(x) - \frac{2m}{\hbar^2} (V_0 - E) \psi(x) = 0$	$\frac{d^2}{dx^2} \psi(x) + \frac{2mE}{\hbar^2} \psi(x) = 0$
$\frac{2m}{\hbar^2} (V_0 - E) = \beta^2$	$\frac{2mE}{\hbar^2} = k^2$
$\frac{d^2}{dx^2} \psi(x) - \beta^2 \psi(x) = 0$	$\frac{d^2}{dx^2} \psi(x) + k^2 \psi(x) = 0$
$\psi(x): \exp(\beta x) \text{ or } \exp(-\beta x)$	$\cos(kx) \text{ or } \sin(kx)$

Requirements of general rules on determining the correct wave functions

- Naturel boundary conditions: wave function must be limited, especially for $x \rightarrow \pm\infty$.
- The potential $V(x)$ is symmetric about origin, the wave functions must be of definite parity.

2. Even parity case

From the continuation of wave function and its derivative (Figure 5), we have

$$A \cos\left(\frac{ka}{2}\right) = B \exp\left(-\beta \frac{a}{2}\right)$$

$$-\frac{k}{2} A \sin\left(\frac{ka}{2}\right) = -\frac{\beta}{2} B \exp\left(-\beta \frac{a}{2}\right)$$

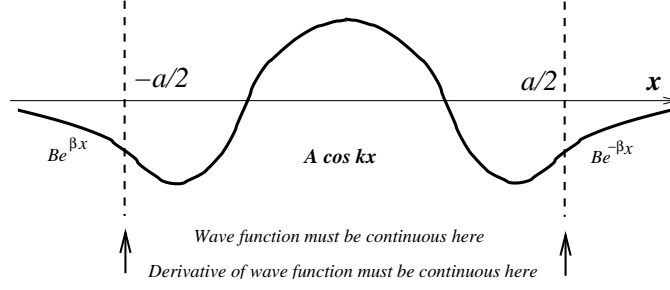


FIG. 5: Even parity solution.

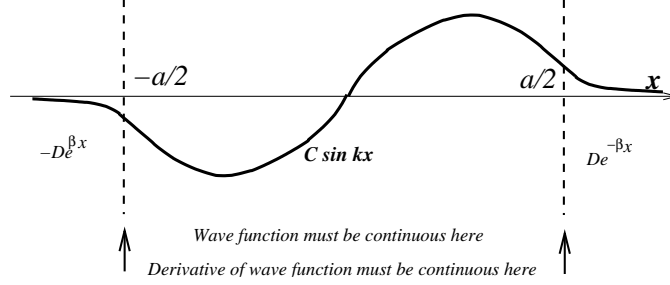


FIG. 6: Odd parity solution.

which gives

$$\frac{k}{2} \tan\left(\frac{ka}{2}\right) = \frac{\beta}{2} \quad (1)$$

This is the limitation on wave function set by natural boundary conditions. It is actually a transcendental equation of the form

$$\frac{1}{2} \sqrt{\frac{2mE}{\hbar^2}} \tan\left(\sqrt{\frac{2mE}{\hbar^2}} \frac{a}{2}\right) = \frac{1}{2} \sqrt{\frac{2m(V_0 - E)}{\hbar^2}} \quad (2)$$

Since values of m , a , and V_0 are given, it is a transcendental equation of E . Only E values that satisfy the above equation can be the energy value of that particle.

3. Odd parity case

The natural boundary conditions lead to the following equations

$$\begin{aligned} C \sin\left(\frac{ka}{2}\right) &= D \exp\left(-\beta \frac{a}{2}\right) \\ \frac{k}{2} C \cos\left(\frac{ka}{2}\right) &= -\frac{\beta}{2} D \exp\left(-\beta \frac{a}{2}\right) \end{aligned}$$

which gives

$$\frac{k}{2} \cot\left(\frac{ka}{2}\right) = -\frac{\beta}{2} \quad (3)$$

This is another limitation on wave functions set by natural boundary conditions which is equivalent to the transcendental equation

$$\frac{1}{2} \sqrt{\frac{2mE}{\hbar^2}} \cot\left(\sqrt{\frac{2mE}{\hbar^2}} \frac{a}{2}\right) = -\frac{1}{2} \sqrt{\frac{2m(V_0 - E)}{\hbar^2}} \quad (4)$$

It is the equation submitted by energy values E of odd parity states.

4. Graphical solution of the energy defining equation

The energy value E must satisfy one of the two equations

$$\begin{aligned}\frac{k}{2} \tan\left(\frac{ka}{2}\right) &= \frac{\beta}{2} \\ \frac{k}{2} \cot\left(\frac{ka}{2}\right) &= -\frac{\beta}{2}\end{aligned}$$

Multiplied by a , they become

$$\begin{aligned}\left(\frac{ka}{2}\right) \tan\left(\frac{ka}{2}\right) &= \frac{\beta a}{2} \\ \left(\frac{ka}{2}\right) \cot\left(\frac{ka}{2}\right) &= -\left(\frac{\beta a}{2}\right)\end{aligned}$$

Let

$$\begin{aligned}\xi &= \frac{1}{2}ka = \frac{1}{2}\sqrt{\frac{2mE}{\hbar^2}}a \\ \eta &= \frac{1}{2}\beta a = \frac{1}{2}\sqrt{\frac{2m(V_0 - E)}{\hbar^2}}a\end{aligned}$$

with ξ and η larger than 0. We notice that ξ and η are not independent each other, rather, they are linked by

$$\xi^2 + \eta^2 = \frac{m}{2\hbar^2}V_0a^2$$

Thus the energy defining relations can be decomposed into two set simultaneous equations

$$\begin{aligned}\xi \tan \xi &= \eta \\ \xi^2 + \eta^2 &= \frac{m}{2\hbar^2}V_0a^2\end{aligned}$$

and

$$\begin{aligned}-\xi \cot \xi &= \eta \\ \xi^2 + \eta^2 &= \frac{m}{2\hbar^2}V_0a^2\end{aligned}$$

For a given set of values m , V_0 , a , the energy E can be evaluated by finding out the intersection points of above two sets of simultaneous equations. Group of $\xi^2 + \eta^2 = \frac{m}{2\hbar^2}V_0a^2$ with different (V_0a^2) values are shown in Figure 7. Our main results are:

1. For a given potential well (V_0a^2) , numbers of intersection points (bound states) are limited

value of $\left(\frac{m}{2\hbar^2}V_0a^2\right)^{\frac{1}{2}}$:	$(0, \frac{\pi}{2})$	$(\frac{\pi}{2}, \pi)$	$(\pi, \frac{3\pi}{2})$	$(\frac{3\pi}{2}, 2\pi)$	\dots
number of intersection points:	1	2	3	4	\dots

2. Bound states have definite parities. As the energy increases, even parity states and odd parity states appear alternatively with the leading (lowest energy) one to be even parity.

lowest state with even parity always exist

$$\begin{aligned}\frac{m}{2\hbar^2}V_0a^2 \geq \frac{\pi^2}{4}, i.e. V_0a^2 \geq \frac{\pi^2\hbar^2}{2m}, & \quad \text{lowest state with odd parity can exist} \\ \frac{m}{2\hbar^2}V_0a^2 \geq \pi^2, i.e. V_0a^2 \geq \frac{2\pi^2\hbar^2}{m}, & \quad \text{1st excited state with even parity will appear}\end{aligned}$$

3. No odd parity bound state for δ potential well. In the symmetric finite potential well, let the width $a \rightarrow 0$, the depth $V_0 \rightarrow +\infty$ and keep $V_0a = \gamma = \text{constant}$ The potential becomes δ potential

$$V(x) = -\gamma\delta(x)$$

It is obviously not possible for V_0a^2 to satisfy the condition for lowest state with odd parity.

Problem 1 Zen Jinyan, Textbook, page 80-84, Problem 1,3,4,12

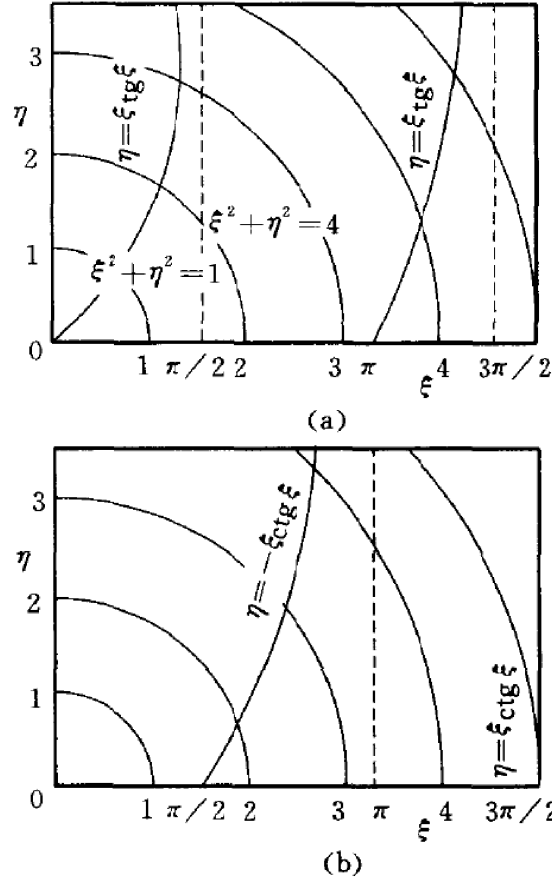


FIG. 7: Group of $\xi^2 + \eta^2 = \frac{m}{2\hbar^2} V_0 a^2$ with different $(V_0 a^2)$ values

II. GENERIC PROBLEM OF QUANTUM MECHANICS

A lot has happened in the last several sections, so let us recapitulate, from a somewhat different perspective. Here is the generic problem of Quantum Mechanics.

$$\left. \begin{array}{l} V(x) \text{ time-independent potential} \\ \Psi(x,0) \text{ starting wave function} \end{array} \right\} \Rightarrow \Psi(x,t) \text{ for any subsequent time } t$$

To do this you must solve

$$i\hbar \frac{\partial \Psi}{\partial t} = -\frac{\hbar^2}{2m} \frac{\partial^2 \Psi}{\partial x^2} + V(x) \Psi$$

$$\Psi(x,t) = \psi(x) f(t) = \psi(x) \exp(-iEt/\hbar)$$

The strategy is first to solve the stationary (time-independent) Schrödinger equation

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

this yields, in general, an infinite set of solutions

$$\begin{array}{cccc} \psi_1(x) & \psi_2(x) & \psi_3(x) & \cdots \\ E_1 & E_2 & E_3 & \cdots \end{array}$$

To fit $\Psi(x,0)$ you write down the general linear combination of these solutions

$$\Psi(x,0) = \sum_{n=1}^{\infty} c_n \psi_n(x)$$

the miracle is that you can *always* match the special initial state by appropriate choice of the constants c_1, c_2, c_3, \dots . To construct $\Psi(x, t)$ you simply tack to each term its characteristic time dependence $\exp(-iE_n t/\hbar)$

$$\Psi(x, t) = \sum_{n=1}^{\infty} c_n \psi_n(x) \exp(-iE_n t/\hbar) = \sum_{n=1}^{\infty} c_n \Psi_n(x, t) \quad (5)$$

The separable solutions themselves

$$\Psi_n(x, t) = \psi_n(x) \exp(-iE_n t/\hbar)$$

are stationary states, in the sense that all probabilities and expectation values are independent of time. This property, however, is emphatically **NOT** shared by the general solution (5); the energies are different, for different stationary states, and the exponentials do not cancel, when you calculate $|\Psi(x, t)|^2$.

A. Properties of wave functions

In the case of the infinite square well, the solutions

$$\psi_n(x) = \sqrt{\frac{2}{a}} \sin\left(\frac{n\pi x}{a}\right), \quad n = 1, 2, 3, \dots$$

have some interesting and important properties:

1. They are alternately even and odd (for symmetric $V(x)$)
2. ψ_n has $(n - 1)$ nodes (zero-crossing) (universal)
3. Orthogonality (quite general)

$$\int \psi_m^*(x) \psi_n(x) dx = 0 \quad m \neq n$$

Normalization

$$\int \psi_m^*(x) \psi_m(x) dx = 1$$

In fact we can combine the orthogonality and normalization into a single statement

$$\int \psi_m^*(x) \psi_n(x) dx = \delta_{mn} = \begin{cases} 0 & \text{if } m \neq n \\ 1 & \text{if } m = n \end{cases}$$

We say that the ψ 's are **orthonormal**.

4. Completeness (holds for most potentials)

Any function $f(x)$ can be expressed as a linear combination of ψ_n

$$f(x) = \sum_{n=1}^{\infty} c_n \psi_n(x) = \sqrt{\frac{2}{a}} \sum_{n=1}^{\infty} c_n \sin\left(\frac{n\pi}{a}x\right) \quad (6)$$

We recognize that Equation (6) is nothing but the Fourier series for $f(x)$, and the fact that "any" function can be expanded in this way is sometimes called *Dirichlet's theorem*. The expansion coefficients c_n can be evaluated—for a given $f(x)$ —by a method I call **Fourier's trick**, which beautifully exploits the orthonormality of $\{\psi_n(x)\}$: Multiply both sides of Equation (6) by $\psi_n^*(x)$, and integrate

$$c_n = \int \psi_n^*(x) f(x) dx$$

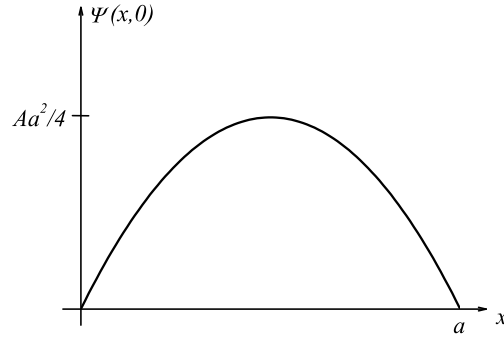


FIG. 8: The starting wave function in the example.

B. Construction of the most general solution

The stationary states of infinite square well are evidently

$$\Psi_n(x, t) = \sqrt{\frac{2}{a}} \sin\left(\frac{n\pi}{a}x\right) e^{-i(n^2\pi^2\hbar/2ma^2)t}$$

We claimed that the most general solution to the time dependent Schrödinger equation is a linear combination of stationary states

$$\Psi(x, t) = \sum_{n=1}^{\infty} c_n \sqrt{\frac{2}{a}} \sin\left(\frac{n\pi}{a}x\right) e^{-i(n^2\pi^2\hbar/2ma^2)t}$$

(Don't believe? Check it!) Any prescribed initial wave function $\Psi(x, 0)$ can be expanded on $\{\psi_n(x)\}$ by appropriate choice of the coefficients c_n

$$\begin{aligned} \Psi(x, 0) &= \sum_{n=1}^{\infty} c_n \psi_n(x) && \text{completeness} \\ c_n &= \sqrt{\frac{2}{a}} \int_0^a \sin\left(\frac{n\pi}{a}x\right) \Psi(x, 0) dx && \text{orthonormality} \end{aligned}$$

Here is a procedure:

$$\Psi(x, 0) \rightarrow c_n \rightarrow \Psi(x, t) \rightarrow \text{dynamical quantities}$$

C. An example

A particle in the infinite square well has the initial wave function

$$\Psi(x, 0) = Ax(a-x), \quad (0 \leq x \leq a)$$

outside the well $\Psi = 0$. Find $\Psi(x, t)$.

Solution: First we determine A by normalization

$$\begin{aligned} 1 &= \int_0^a |\Psi(x, 0)|^2 dx = |A|^2 \int_0^a x^2(a-x)^2 dx = |A|^2 \frac{a^5}{30} \\ \therefore A &= \sqrt{\frac{30}{a^5}} \end{aligned}$$

$$c_n = \sqrt{\frac{2}{a}} \int_0^a \sin\left(\frac{n\pi}{a}x\right) \sqrt{\frac{30}{a^5}} x(a-x) dx = \begin{cases} 0, & \text{if } n \text{ is even} \\ 8\sqrt{15}/(n\pi)^3, & \text{if } n \text{ is odd} \end{cases}$$

$$\Psi(x, t) = \sqrt{\frac{30}{a}} \left(\frac{2}{\pi}\right)^3 \sum_{n=1,3,5,\dots} \frac{1}{n^3} \sin\left(\frac{n\pi}{a}x\right) \exp(-in^2\pi^2\hbar t/2ma^2)$$

Loosely speaking, c_n tells you the “amount of ψ_n that is contained in Ψ ”. Some people like to say that $|c_n|^2$ is the “probability of finding the particle in the n -th stationary state”, but this is bad language; the particle is in the state Ψ , not Ψ_n , and anyhow, in the laboratory you don’t “find a particle to be in a particular state” - you measure some *observable*, and what you get is a *number*. As we will see later, what $|c_n|^2$ tells you is the *probability that a measurement of the energy would yield the value E_n* . The sum of these probability should be 1, $\sum_{n=1}^{\infty} |c_n|^2 = 1$, and the expectation value of the energy

$$\langle H \rangle = \sum_{n=1}^{\infty} |c_n|^2 E_n$$

is independent of time which is a manifestation of **conservation of energy** in Quantum Mechanics.

The starting wave function closely resembles the ground state ψ_1 . This suggests that $|c_1|^2$ should dominate, and in fact

$$|c_1|^2 = \left(\frac{8\sqrt{15}}{\pi^3}\right)^2 = 0.998555\dots$$

The rest of the coefficients make up the difference

$$\sum_{n=1}^{\infty} |c_n|^2 = \left(\frac{8\sqrt{15}}{\pi^3}\right)^2 \sum_{n=1,3,5,\dots} \frac{1}{n^6} = 1$$

The expectation value of the energy in this example is

$$\langle H \rangle = \sum_{n=1,3,5,\dots}^{\infty} \left(\frac{8\sqrt{15}}{\pi^3}\right)^2 \frac{1}{n^6} \frac{n^2\pi^2\hbar^2}{2ma^2} = \frac{480\hbar^2}{\pi^4 ma^2} \sum_{n=1,3,5,\dots}^{\infty} \frac{1}{n^4} = \frac{5\hbar^2}{ma^2}$$

where we have used

$$\frac{1}{1^6} + \frac{1}{3^6} + \frac{1}{5^6} + \dots = \frac{\pi^6}{960}$$

$$\frac{1}{1^4} + \frac{1}{3^4} + \frac{1}{5^4} + \dots = \frac{\pi^4}{96}$$

As one might expect, it is very close to $E_1 = \frac{\pi^2\hbar^2}{2ma^2}$ - slightly larger, because of the admixture of excited states.

Problem 2 Griffiths, page 38, 2.4, 2.5, 2.6, (for advanced exercise)

III. THE HARMONIC OSCILLATOR

The classical harmonic oscillator is governed by Hooke’s Law

$$F = -kx = m \frac{d^2x}{dt^2}$$

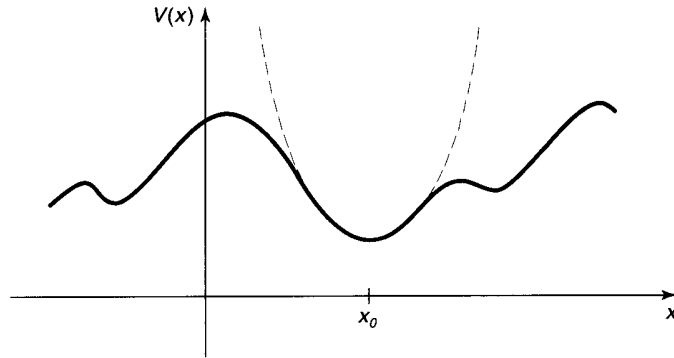


FIG. 9: Parabolic approximation (dashed curve) to an arbitrary potential, in the neighborhood of a local minimum.

where m is the mass of the particle and k is the force constant of the spring. The solution is

$$x(t) = A \sin(\omega t) + B \cos(\omega t)$$

with the frequency of oscillation $\omega \equiv \sqrt{\frac{k}{m}}$. The potential energy is a parabola

$$V(x) = \frac{1}{2} k x^2.$$

Practically any potential is approximately parabolic, in the neighborhood of a local minimum (Figure 9). Formally, if we expand $V(x)$ in a Taylor series about the minimum

$$\begin{aligned} V(x) &= V(x_0) + V'(x_0)(x - x_0) + \frac{1}{2} V''(x_0)(x - x_0)^2 + \dots \\ V(x) &\cong \frac{1}{2} V''(x_0)(x - x_0)^2 \end{aligned}$$

which describes a simple harmonic oscillator about point x_0 with an effective spring constant $k = V''(x_0)$. The simple harmonic oscillator is so important: virtually any oscillatory motion is approximately simple harmonic, as long as the amplitude is small. The quantum problem is to solve the Schrödinger equation for the potential

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

As we have seen, it suffices to solve the time-independent/stationary equation

$$-\frac{\hbar^2}{2m} \frac{d^2 \psi}{dx^2} + \frac{1}{2} m \omega^2 x^2 \psi = E \psi \quad (7)$$

A. Algebraic method: ladder operator

To begin with, let's rewrite Equation (7) in a more suggestive form

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

because of

$$\hat{p} \equiv \frac{\hbar}{i} \frac{d}{dx}, \quad \hat{H} = \frac{1}{2m} \left[p^2 + (m\omega x)^2 \right]$$

1. Commutator

The idea is to *factor* the term in square brackets. If these were numbers, it would be easy

$$u^2 + v^2 = (iu + v)(-iu + v)$$

However, u and v are operators and do not commute

$$\hat{x}\hat{p} \neq \hat{p}\hat{x}$$

(If I forget the hat for operators, forgive me). Still, this does motivate us to examine the quantities

$$\hat{a}_{\pm} \equiv \frac{1}{\sqrt{2\hbar m\omega}} (\mp i\hat{p} + m\omega\hat{x})$$

We see that

$$\begin{aligned} \hat{a}_-\hat{a}_+ &= \frac{1}{2\hbar m\omega} (i\hat{p} + m\omega\hat{x})(-i\hat{p} + m\omega\hat{x}) \\ &= \frac{1}{2\hbar m\omega} (\hat{p}^2 + (m\omega\hat{x})^2 - im\omega(\hat{x}\hat{p} - \hat{p}\hat{x})) \end{aligned}$$

As anticipated, there is an extra term $(\hat{x}\hat{p} - \hat{p}\hat{x})$, which is called the **commutator of \hat{x} and \hat{p}** , it is a measure of how badly they fail to commute. In general, the commutator of operators \hat{A} and \hat{B} is

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

In this notation,

$$\hat{a}_-\hat{a}_+ = \frac{1}{2\hbar m\omega} [\hat{p}^2 + (m\omega\hat{x})^2] - \frac{i}{2\hbar} [\hat{x}, \hat{p}].$$

To figure out $[\hat{x}, \hat{p}]$, we need give it a test function $f(x)$ to act on

$$\begin{aligned} [\hat{x}, \hat{p}]f(x) &= \left[x \frac{\hbar}{i} \frac{d}{dx} (f) - \frac{\hbar}{i} \frac{d}{dx} (xf) \right] \\ &= \frac{\hbar}{i} \left(x \frac{df}{dx} - x \frac{df}{dx} - f \right) = i\hbar f(x) \end{aligned}$$

Dropping the test function, which has served its purpose

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

This lovely and ubiquitous result is known as the **canonical commutation relation**. With this,

$$\hat{a}_-\hat{a}_+ = \frac{1}{\hbar\omega} \hat{H} + \frac{1}{2}$$

or

$$\hat{H} = \hbar\omega \left(\hat{a}_-\hat{a}_+ - \frac{1}{2} \right)$$

The Hamiltonian does not factor perfectly - extra term $-\frac{1}{2}$. The ordering of $\hat{a}_-\hat{a}_+$ is important

$$\hat{a}_+\hat{a}_- = \frac{1}{\hbar\omega} \hat{H} - \frac{1}{2}$$

we thus have

$$[\hat{a}_-, \hat{a}_+] = 1$$

and

$$\hat{H} = \hbar\omega \left(\hat{a}_+\hat{a}_- + \frac{1}{2} \right)$$

The Schrödinger equation $\hat{H}\psi = E\psi$ for the harmonic oscillator takes the form

$$\hbar\omega \left(\hat{a}_{\pm}\hat{a}_{\mp} \pm \frac{1}{2} \right) \psi = E\psi$$

2. Ladder operators

Now, here comes the crucial step: I claim that if $\hat{H}\psi = E\psi$ then

$$\begin{aligned}\hat{H}(\hat{a}_+\psi) &= (E + \hbar\omega)(\hat{a}_+\psi) \\ \hat{H}(\hat{a}_-\psi) &= (E - \hbar\omega)(\hat{a}_-\psi)\end{aligned}$$

Proof:

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

By the same token, $\hat{a}_-\psi$ is a solution with energy $(E - \hbar\omega)$.

$$\hat{H}(\hat{a}_-\psi) = (E - \hbar\omega)(\hat{a}_-\psi)$$

Here, then, is a wonderful machine for generating new solutions, with higher and lower energies - if we could just find one solution, to get started! We call \hat{a}_\pm **ladder operators** or **raising/lowering operators**, because they allow us to climb up and down in energy. The ladder of states is illustrated in Figure 10.

But wait! What if I apply the lowering operator repeatedly? Eventually I'm going to reach a state with energy less than zero, which does not exist! At some point the machine must fail. How can that happen? We know that $\hat{a}_-\psi$ is a new solution to the Schrödinger equation, but there is no guarantee that it will be normalizable—it might be zero, or its square integral might be infinite. In practice it is the former: There occurs a “lowest rung” (let's call it ψ_0) such that

$$\hat{a}_-\psi_0 = 0$$

i.e

$$\frac{1}{\sqrt{2\hbar m\omega}}\left(\hbar\frac{d}{dx} + m\omega x\right)\psi_0 = 0$$

We can use this to determine $\psi_0(x)$

$$\begin{aligned}\frac{d\psi_0}{dx} &= -\frac{m\omega}{\hbar}x\psi_0 \\ \int \frac{d\psi_0}{\psi_0} &= -\frac{m\omega}{\hbar} \int x dx \\ \ln \psi_0 &= -\frac{m\omega}{2\hbar}x^2 + \text{const.} \\ \psi_0 &= Ae^{-\frac{m\omega}{2\hbar}x^2}\end{aligned}$$

The constant A is given by the normalization

$$1 = |A|^2 \int_{-\infty}^{+\infty} \exp\left(-\frac{m\omega}{\hbar}x^2\right) dx = |A|^2 \sqrt{\frac{\hbar\pi}{m\omega}}$$

so

$$A^2 = \sqrt{\frac{m\omega}{\pi\hbar}}$$

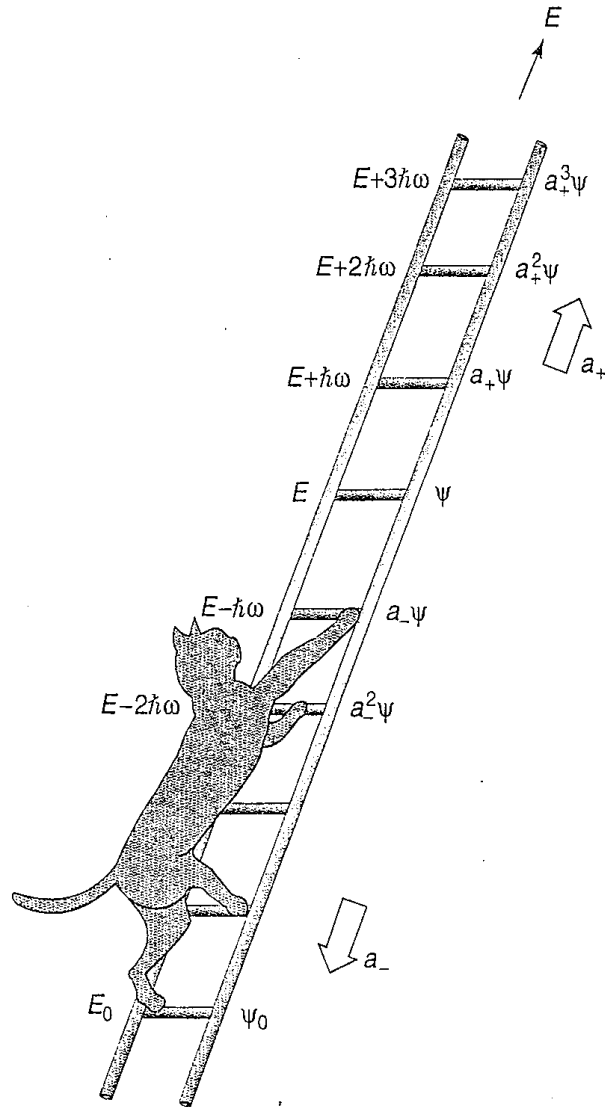


FIG. 10: The ladder of states for the harmonic oscillator.

and hence

$$\psi_0 = \left(\frac{m\omega}{\pi\hbar}\right)^{1/4} \exp\left(-\frac{m\omega}{2\hbar}x^2\right) \quad (8)$$

To determine the energy of this state, we plug it into the Schrödinger equation

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

and exploit the fact that $\hat{a}_-\psi_0 = 0$

$$E_0 = \frac{1}{2}\hbar\omega \quad (9)$$

It is customary for harmonic oscillator that the states start with $n = 0$, instead of $n = 1$. We then apply \hat{a}_+ repeatedly to generate the excited states, increasing the energy by $\hbar\omega$ with each step

$$\psi_n(x) = A_n(\hat{a}_+)^n\psi_0(x) \quad (10)$$

with

$$E_n = \left(n + \frac{1}{2} \right) \hbar\omega \quad (11)$$

Example: Find the first excited state of the harmonic oscillator.

Solution: Using equation (10)

$$\begin{aligned} \psi_1(x) &= A_1 \hat{a}_+ \psi_0(x) = \frac{A_1}{\sqrt{2\hbar m\omega}} \left(-\hbar \frac{d}{dx} + m\omega x \right) \left(\frac{m\omega}{\pi\hbar} \right)^{1/4} \exp\left(-\frac{m\omega}{2\hbar} x^2\right) \\ &= A_1 \left(\frac{m\omega}{\pi\hbar} \right)^{1/4} \sqrt{\frac{2m\omega}{\hbar}} x \exp\left(-\frac{m\omega}{2\hbar} x^2\right) \end{aligned}$$

We can normalize it

$$|A_1|^2 \sqrt{\frac{m\omega}{\pi\hbar}} \left(\frac{2m\omega}{\hbar} \right) \int_{-\infty}^{+\infty} x^2 \exp\left(-\frac{m\omega}{\hbar} x^2\right) dx = |A_1|^2 = 1$$

We often use the following Gaussian integrals

$$\begin{aligned} \int_0^{\infty} x^{2n} \exp\left(-\frac{x^2}{a^2}\right) dx &= \sqrt{\pi} \frac{(2n)!}{n!} \left(\frac{a}{2}\right)^{2n+1} \\ \int_0^{\infty} x^{2n+1} \exp\left(-\frac{x^2}{a^2}\right) dx &= \frac{n!}{2} a^{2n+2} \end{aligned}$$

3. Normalization algebraically: find A_n

It seems not easy to get the normalization factor for ψ_{50} . So we do it algebraically. We know that $\hat{a}_{\pm}\psi_n$ is proportional to $\psi_{n\pm 1}$

$$\begin{aligned} \hat{a}_+ \psi_n &= c_n \psi_{n+1} \\ \hat{a}_- \psi_n &= d_n \psi_{n-1} \end{aligned}$$

but what are the proportional factors, c_n and d_n ? First note that

- For any functions $f(x)$ and $g(x)$

$$\int_{-\infty}^{+\infty} f^*(\hat{a}_{\pm}g) dx = \int_{-\infty}^{+\infty} (\hat{a}_{\mp}f)^* g dx$$

\hat{a}_{\mp} is the hermitian conjugate of \hat{a}_{\pm} . Proof:

$$left = \frac{1}{\sqrt{2\hbar m\omega}} \int_{-\infty}^{+\infty} f^* \left(\mp \hbar \frac{d}{dx} + m\omega x \right) g dx$$

Integration by part gives

$$\int f^* \frac{dg}{dx} dx \rightarrow - \int g \left(\frac{df}{dx} \right)^* dx$$

so

$$left = \frac{1}{\sqrt{2\hbar m\omega}} \int_{-\infty}^{+\infty} \left[\left(\pm \hbar \frac{d}{dx} + m\omega x \right) f \right]^* g dx = right, QED$$

- From Schrödinger equation

$$\begin{aligned} \hbar\omega \left(\hat{a}_+ \hat{a}_- + \frac{1}{2} \right) \psi_n &= E_n \psi_n = \hbar\omega \left(\hat{n} + \frac{1}{2} \right) \psi_n \\ \hbar\omega \left(\hat{a}_- \hat{a}_+ - \frac{1}{2} \right) \psi_n &= E_n \psi_n = \hbar\omega \left(\hat{n} + 1 - \frac{1}{2} \right) \psi_n \end{aligned}$$

we clearly see

$$\hat{a}_+ \hat{a}_- \psi_n = n \psi_n \quad \hat{a}_- \hat{a}_+ \psi_n = (n+1) \psi_n$$

In particular

$$\begin{aligned} \int_{-\infty}^{+\infty} (\hat{a}_+ \psi_n)^* (\hat{a}_+ \psi_n) dx &= |c_n|^2 \int_{-\infty}^{+\infty} |\psi_{n+1}|^2 dx = |c_n|^2 \\ &= \int_{-\infty}^{+\infty} (\hat{a}_- \hat{a}_+ \psi_n)^* \psi_n dx = (n+1) \int_{-\infty}^{+\infty} |\psi_n|^2 dx \\ &= (n+1) \end{aligned}$$

similarly

$$\begin{aligned} \int_{-\infty}^{+\infty} (\hat{a}_- \psi_n)^* (\hat{a}_- \psi_n) dx &= |d_n|^2 \int_{-\infty}^{+\infty} |\psi_{n-1}|^2 dx = |d_n|^2 \\ &= \int_{-\infty}^{+\infty} (\hat{a}_+ \hat{a}_- \psi_n)^* \psi_n dx = n \int_{-\infty}^{+\infty} |\psi_n|^2 dx \\ &= n \end{aligned}$$

$$\therefore \hat{a}_+ \psi_n = \sqrt{n+1} \psi_{n+1} \quad \hat{a}_- \psi_n = \sqrt{n} \psi_{n-1}$$

Thus

$$\begin{aligned} \psi_1 &= \hat{a}_+ \psi_0 \\ \psi_2 &= \frac{1}{\sqrt{2}} \hat{a}_+ \psi_1 = \frac{1}{\sqrt{2}} (\hat{a}_+)^2 \psi_0 \\ \psi_3 &= \frac{1}{\sqrt{3}} \hat{a}_+ \psi_2 = \frac{1}{\sqrt{3 \cdot 2}} (\hat{a}_+)^3 \psi_0 \\ &\dots \\ \psi_n &= \frac{1}{\sqrt{n!}} (\hat{a}_+)^n \psi_0 \quad \left(A_n = \frac{1}{\sqrt{n!}} \right) \end{aligned}$$

- As in the case of the infinite square well, the stationary states of the harmonic oscillator are orthogonal

$$\int_{-\infty}^{+\infty} \psi_m^* \psi_n dx = \delta_{mn}$$

Proof:

$$\begin{aligned} \int_{-\infty}^{+\infty} \psi_m^* (\hat{a}_+ \hat{a}_-) \psi_n dx &= n \int_{-\infty}^{+\infty} \psi_m^* \psi_n dx = \int_{-\infty}^{+\infty} (\hat{a}_- \psi_m)^* \hat{a}_- \psi_n dx \\ &= \int_{-\infty}^{+\infty} (\hat{a}_+ \hat{a}_- \psi_m)^* \psi_n dx = m \int_{-\infty}^{+\infty} \psi_m^* \psi_n dx \end{aligned}$$

Unless $m = n$, then $\int_{-\infty}^{+\infty} \psi_m^* \psi_n dx$ must be zero. Orthonormality means that we can again use Fourier's trick to evaluate the coefficients, when we expand $\Psi(x, 0)$ as a linear combination of stationary states, and $|c_n|^2$ is again the probability that a measurement of the energy would yield the value E_n .

Example: Find the expectation value of the potential energy in the n -th state of the harmonic oscillator.

Solution:

$$\langle \hat{V} \rangle = \left\langle \frac{1}{2} m \omega^2 \hat{x}^2 \right\rangle = \frac{1}{2} m \omega^2 \int_{-\infty}^{+\infty} \psi_n^* \hat{x}^2 \psi_n dx$$

$$\hat{x} = \sqrt{\frac{\hbar}{2m\omega}} (\hat{a}_+ + \hat{a}_-)$$

$$\hat{x}^2 = \frac{\hbar}{2m\omega} [\hat{a}_+^2 + \hat{a}_+ \hat{a}_- + \hat{a}_- \hat{a}_+ + \hat{a}_-^2]$$

$$\begin{aligned} \langle \hat{V} \rangle &= \frac{\hbar\omega}{4} \int \psi_n^* [\hat{a}_+^2 + \hat{a}_+ \hat{a}_- + \hat{a}_- \hat{a}_+ + \hat{a}_-^2] \psi_n dx \\ &= \frac{\hbar\omega}{4} (n + n + 1) = \frac{1}{2} \hbar\omega (n + \frac{1}{2}) \\ &= \frac{1}{2} E_n \end{aligned}$$

B. Analytic method: differential equation

We return now to the Schrödinger equation for the harmonic oscillator

$$-\frac{\hbar^2}{2m} \frac{d^2\psi}{dx^2} + \frac{1}{2} m\omega^2 x^2 \psi = E\psi.$$

Things look a little clearer if we introduce the dimensionless variable

$$\xi \equiv \sqrt{\frac{m\omega}{\hbar}} x = \alpha x;$$

in terms of ξ the Schrödinger equation reads

$$\frac{d^2\psi}{d\xi^2} = (\xi^2 - K) \psi \quad \text{with } K \equiv \frac{2E}{\hbar\omega}$$

Our problem is to solve the differential equation, and in the process obtain the allowed values of K (or E).

To begin with, note the behavior of ψ at $\xi \rightarrow \pm\infty$. We are only interested in the bound states, i.e., $\lim_{\xi \rightarrow \pm\infty} \psi(\xi) \rightarrow 0$.

$$\begin{aligned} \frac{d^2\psi}{d\xi^2} &\approx \xi^2 \psi \\ \psi(\xi) &\approx A \exp(-\frac{\xi^2}{2}) + B \exp(\frac{\xi^2}{2}) \end{aligned}$$

The physically acceptable solutions, then, have the asymptotic form

$$\psi(\xi) \rightarrow () \exp(-\frac{\xi^2}{2}) \quad \text{at large } \xi$$

We define

$$\psi(\xi) = \exp(-\frac{\xi^2}{2}) h(\xi)$$

where $h(\xi)$ should have a simpler functional form than $\psi(\xi)$. Differentiating $\psi(\xi)$

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

Schrödinger Equation becomes Hermite equation

$$\frac{d^2h}{d\xi^2} - 2\xi \frac{dh}{d\xi} + (K - 1) h = 0 \tag{12}$$

We look for solutions in the form of power series in ξ , i.e. solutions by method of Frobenius method

$$\begin{aligned} h(\xi) &= \sum_{j=0}^{\infty} a_j \xi^j = a_0 + a_1 \xi + a_2 \xi^2 + \dots \\ \frac{dh}{d\xi} &= \sum_{j=0}^{\infty} j a_j \xi^{j-1} = a_1 + 2a_2 \xi + 3a_3 \xi^2 + \dots \\ \frac{d^2h}{d\xi^2} &= \sum_{j=0}^{\infty} j(j-1) a_j \xi^{j-2} = \sum_{j=0}^{\infty} (j+1)(j+2) a_{j+2} \xi^j \\ &= 2a_2 + 2 \cdot 3a_3 \xi + 3 \cdot 4a_4 \xi^2 + \dots \end{aligned}$$

Putting these into equation (12), we find

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

It follows that the coefficients satisfy the **recursion formula**

$$a_{j+2} = \frac{2j+1-K}{(j+1)(j+2)} a_j. \quad (13)$$

Starting with a_0 , it generate all the even-numbered coefficients

$$a_2 = \frac{1-K}{2} a_0, \quad a_4 = \frac{5-K}{12} a_2 = \frac{(1-K)(5-K)}{24} a_0, \dots$$

and starting with a_1 , it generate all the odd-numbered coefficients

$$a_3 = \frac{3-K}{6} a_1, \quad a_5 = \frac{7-K}{20} a_3 = \frac{(3-K)(7-K)}{120} a_1, \dots$$

The complete solution is written as

$$h(\xi) = h_{even}(\xi) + h_{odd}(\xi)$$

$$h_{even}(\xi) = a_0 + a_2 \xi^2 + a_4 \xi^4 + \dots$$

$$h_{odd}(\xi) = a_1 \xi + a_3 \xi^3 + a_5 \xi^5 + \dots$$

where two free parameters a_0 and a_1 are just what we would expect for a second-order differential equation.

However, not all the solution so obtained are normalizable. For at very large j , the recursion formula becomes

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

1)

$$j = 2m \quad \frac{a_{2m+2}}{a_{2m}} = \frac{2}{2m} = \frac{1}{m}$$

which are the same coefficients as e^{ξ^2} in Taylor series

$$e^{\xi^2} = \sum_{m=0}^{\infty} \frac{\xi^{2m}}{m!} = 1 + \xi^2 + \frac{\xi^4}{2} + \frac{\xi^6}{6} + \frac{\xi^8}{24} + \dots$$

$$|\xi| \rightarrow \infty \quad h_{even}(\xi) \sim e^{\xi^2}$$

2)

$$j = 2m + 1 \quad \text{similarly} \quad h_{\text{odd}}(\xi) \sim \xi e^{\xi^2}$$

We have

$$\psi(\xi) \sim \begin{cases} e^{-\xi^2/2} h_{\text{even}}(\xi) \sim e^{\xi^2/2} \\ e^{-\xi^2/2} h_{\text{odd}}(\xi) \sim \xi e^{\xi^2/2} \end{cases} \quad \text{diverges at } |\xi| \rightarrow \infty$$

This is precisely the asymptotic behavior we didn't want. The series must be terminated. There must occur some "highest" j (call it n), such that the recursion formula spits out $a_{n+2} = 0$. Then, equation (13) requires that

$$K = 2n + 1, \quad n = 0, 1, 2, \dots \quad (14)$$

(This will truncate *either* the series h_{even} *or* h_{odd} ; the *other one* must be zero from the start

$$\begin{aligned} a_1 &= 0 & \text{if } n \text{ is even} \\ a_0 &= 0 & \text{if } n \text{ is odd} \end{aligned}$$

or remains infinite series which should be discarded.)

$$a_n \neq 0, \quad a_{n+2} = a_{n+4} = \dots = 0$$

In each case, we find a solution $H_n(\xi)$.

1. Discussion on the energy quantization

The energy must be (referring to equation (14))

$$E = \left(n + \frac{1}{2} \right) \hbar\omega, \quad n = 0, 1, 2, \dots$$

Thus we recover, by a completely different method, the fundamental quantization condition we found algebraically in equation (11). The levels in harmonic oscillator are equally distributed.

The Schrödinger equation has solutions for any value of E but almost all of these solutions blow up exponentially at large x , and hence are not normalizable. Figure shows the solutions for 3 different values of E . It is clear that only the solution corresponding to $E = 0.5\hbar\omega$ is physically acceptable and hence survives.

2. Discussion on the wave function

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

- if $n = 0$ there is only one term in the series (we must pick $a_1 = 0$ to kill h_{odd} , and $j = 0$ yields $a_2 = 0$)

$$h_0(\xi) = a_0$$

hence

$$\psi_0(\xi) = a_0 e^{-\xi^2/2}$$

- if $n = 1$ we take $a_0 = 0$, and $j = 1$ yields $a_3 = 0$

$$h_1(\xi) = a_1 \xi$$

hence

$$\psi_1(\xi) = a_1 \xi e^{-\xi^2/2}$$

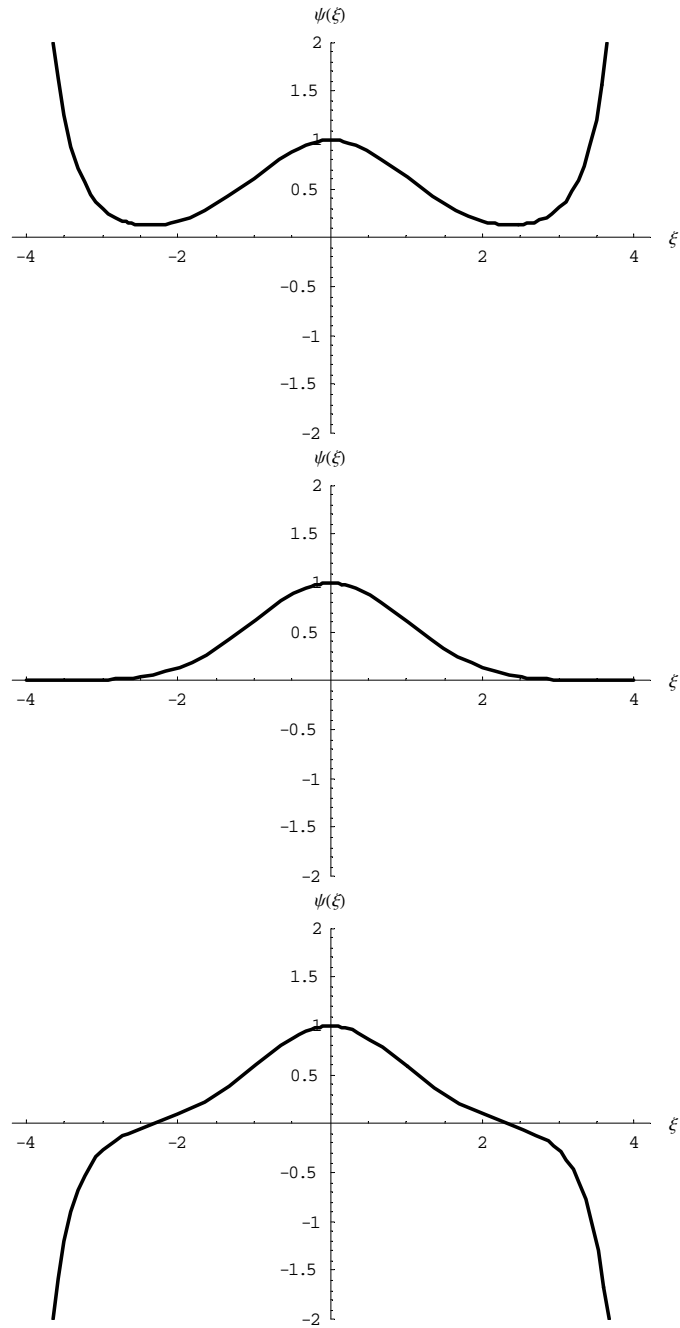


FIG. 11: Solutions to the Schrödinger equation for $E = 0.49\hbar\omega$, $0.5\hbar\omega$, and $0.51\hbar\omega$.

- if $n = 2$ $j = 0$ yields $a_2 = -2a_0$, $j = 2$ yields $a_4 = 0$

$$h_2(\xi) = a_0(1 - 2\xi^2)$$

hence

$$\psi_2(\xi) = a_0(1 - 2\xi^2)e^{-\xi^2/2}$$

and so on

In general, $h_n(\xi)$ will be a polynomial of degree n in ξ , involving even powers only, if n is an even integer, and odd powers only, if n is an odd integer. Apart from the overall factor (a_0 or a_1) they are the so-called **Hermite**

polynomials, $H_n(\xi)$. The first few of them are listed here

$$\begin{aligned} H_0(\xi) &= 1 \\ H_1(\xi) &= 2\xi \\ H_2(\xi) &= 4\xi^2 - 2 \\ H_3(\xi) &= 8\xi^3 - 12\xi \\ &\vdots \end{aligned}$$

By tradition, the arbitrary multiplicative factor is chosen so that the coefficient of the highest power of is 2^n . The orthonormality and recursion formula of $H_n(\xi)$ are

$$\int_{-\infty}^{+\infty} H_m(\xi)H_n(\xi)e^{-\xi^2} d\xi = \sqrt{\pi}2^n \cdot n!\delta_{mn}$$

$$H_{n+1}(\xi) - 2\xi H_n(\xi) + 2nH_{n-1}(\xi) = 0$$

$$\frac{\partial H_n(\xi)}{\partial \xi} = 2nH_{n-1}(\xi)$$

With this polynomials, the stationary states for the harmonic oscillator are

$$\psi_n(x) = N_n e^{-\frac{1}{2}\alpha^2 x^2} H_n(\alpha x)$$

We normalize it by

$$\int_{-\infty}^{+\infty} \psi_m^*(x)\psi_n(x)dx = \delta_{mn},$$

and find

$$N_n = \left(\frac{\alpha}{\pi^{1/2}2^n n!}\right)^{1/2} = \left(\frac{1}{2^n n!}\right)^{1/2} \left(\frac{m\omega}{\pi\hbar}\right)^{1/4}$$

Therefore

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

We see that the potential is symmetric about $x = 0$, so the wave functions have either even or odd parity.

$$\psi_n(-x) = (-1)^n \psi_n(x)$$

The ground state with $E_0 = \frac{1}{2}\hbar\omega$ is easily identified

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

which is the same as what we have got by algebraic method equation (8).

In Figure 12 I have plotted $\psi_n(x)$ for the first few n 's. The quantum oscillator is strikingly different from its classical counterpart – not only are the energies quantized, but the position distributions have some bizarre features. For instance, the probability of finding the particle outside the classically allowed range (that is, with x greater than the classical amplitude for the energy in question) is not zero (see the following example), and in all odd states the probability of finding the particle at the center of the potential well is zero. Only at relatively large n do we begin to see some resemblance to the classical case. In Figure 13 I have superimposed the classical position distribution on the quantum one (for $n = 100$); if you smoothed out the bumps in the latter, the two would fit pretty well (however, in the classical case we are talking about the distribution of positions over time for one oscillator, whereas in the quantum case we are talking about the distribution over an ensemble of identically-prepared systems).

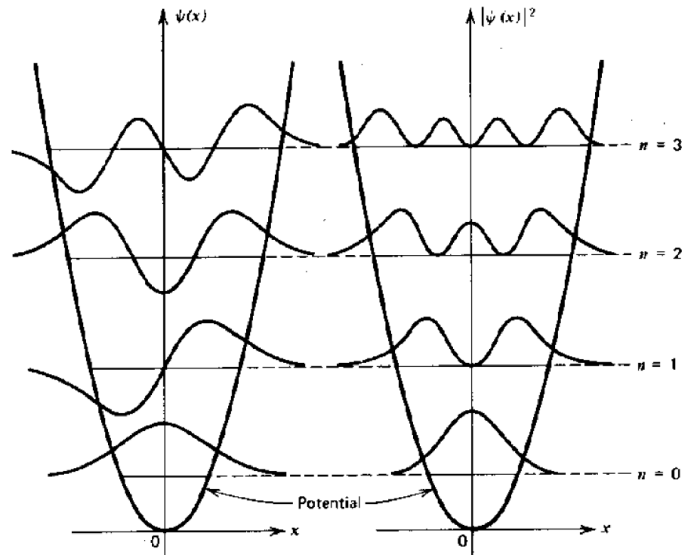


FIG. 12: The first four stationary states of the harmonic oscillator.

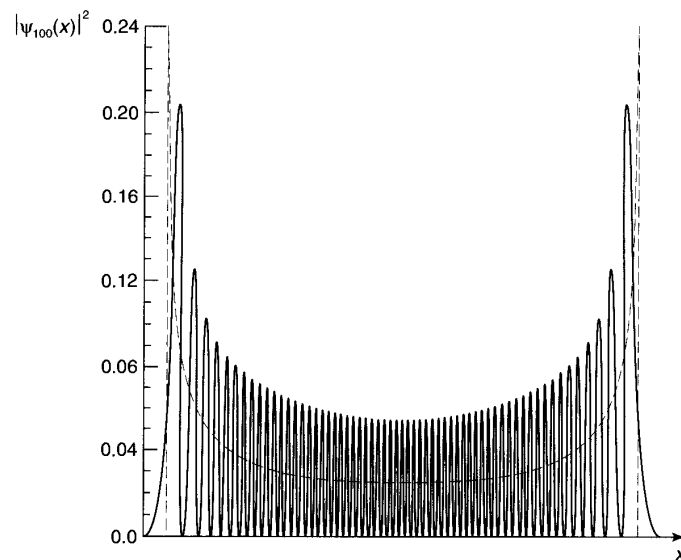


FIG. 13: Graph of $|\psi_{100}|^2$ with the classical distribution (dashed curve) superimposed.

Example: In the ground state of the harmonic oscillator, what is the probability of finding the particle outside the classically allowed region?

Solution: Classically allowed region extends out to

$$\frac{1}{2}m\omega^2 x_0^2 = E_0 = \frac{1}{2}\hbar\omega$$

or

$$x_0 = \sqrt{\frac{\hbar}{m\omega}}$$

so $\xi_0 = 1$ where the velocity of the particle is zero. Quantum mechanics tells us the probability of finding the particle outside the classically allowed region is

$$\left(\int_1^\infty e^{-\xi^2} d\xi\right) / \left(\int_0^\infty e^{-\xi^2} d\xi\right) \simeq 15.7\%$$

Problem 3 Griffiths, page 86, 2.41

Problem 4 Zeng Jinyan, Textbook, page 80-84, 8,9,11

IV. THE FREE PARTICLE: REVISITED

We turn next to what *should* have been the simplest case of all: the free particle [$V(x) = 0$ everywhere]. Classically this would just mean motion at constant velocity, but in quantum mechanics, the free particle is in fact a surprisingly subtle and tricky example. The time-independent Schrödinger equation reads

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

or

$$\frac{d^2\psi}{dx^2} = -k^2\psi, \quad k \equiv \sqrt{\frac{2mE}{\hbar^2}}$$

As in the infinite potential well,

$$\psi(x) = Ae^{ikx} + Be^{-ikx}$$

We have no boundary condition to restrict E , free particle thus can carry *any* (positive) energy. Tacking on the standard time dependence, $\exp(-iEt/\hbar)$, we have

$$\Psi(x, t) = Ae^{ik(x - \frac{\hbar k}{2m}t)} + Be^{-ik(x + \frac{\hbar k}{2m}t)}$$

The first term in the above equation represents a wave traveling to the right, and the second term represents a wave (of the same energy) going to the left. By the way, since they only differ by the sign in front of k , we might as well write

$$\Psi_k(x, t) = Ae^{ik(x - \frac{\hbar k}{2m}t)}$$

and let k runs from negative to positive

$$k \equiv \pm \sqrt{\frac{2mE}{\hbar^2}}, \quad \begin{cases} k > 0 \Rightarrow \text{traveling to the right} \\ k < 0 \Rightarrow \text{traveling to the left} \end{cases}$$

The “stationary states” of the free particle are propagating waves with wave length $\lambda = 2\pi/|k|$, momentum $\vec{p} = \hbar\vec{k}$. The speed of these waves (the coefficient of t over that of x)

$$v_{\text{quantum}} = \frac{\hbar|k|}{2m} = \sqrt{\frac{E}{2m}}$$

On the other hand, the classical speed of a free particle with energy E is given by $E = \frac{1}{2}mv^2$ (pure kinetic, since $V = 0$)

$$v_{\text{classical}} = \sqrt{\frac{2E}{m}} = 2v_{\text{quantum}}$$

Evidently the quantum mechanical wave function travels at *half* the speed of the particle it is supposed to represent! We'll return to this **paradox** in a moment – there is an even more serious problem we need to confront first: This wave function $\Psi_k(x, t)$ is not normalizable

$$\int_{-\infty}^{+\infty} \Psi_k^* \Psi_k dx = |A|^2 \int_{-\infty}^{+\infty} dx = |A|^2 \infty$$

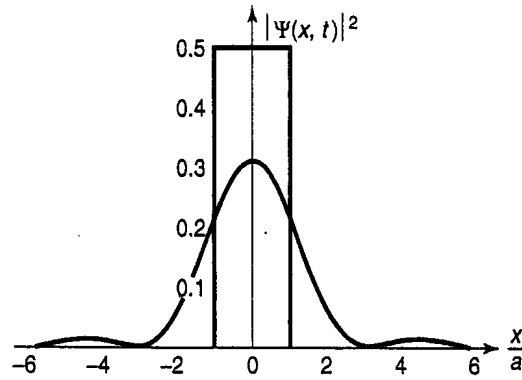


FIG. 14: Graph of $|\Psi(x, t)|^2$ at $t = 0$ (the rectangle) and at $t = ma^2/\hbar$ (the curve).

The separable solutions do not represent physically realizable states: There is no such thing as a free particle with a definite energy. They play a mathematical role that is entirely independent of their physical interpretation.

The general solution to the time-dependent Schrödinger equation is still a linear combination of separable solutions (only this time it's an integral over the continuous variable k , instead of a sum over the discrete index n)

$$\Psi(x, t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \phi(k) e^{ik(x - \frac{\hbar k}{2m}t)} dk \quad (15)$$

$$c_n \leftrightarrow \frac{1}{\sqrt{2\pi}} \phi(k) dk$$

Now this wave function can be normalized (for appropriate $\phi(k)$). But it necessarily carries a range of k 's, and hence a range of energies and speeds. We call it a wave packet.

Generic quantum problem

$$\Psi(x, 0) \rightarrow \Psi(x, t)$$

For infinite potential well

$$\begin{aligned} \Psi(x, 0) &= \sum_{n=1}^{\infty} c_n \psi_n(x), \\ c_n &= \sqrt{\frac{2}{a}} \int_0^a \sin\left(\frac{n\pi}{a}x\right) \Psi(x, 0) dx \end{aligned}$$

For free particle

$$\begin{aligned} \Psi(x, 0) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \phi(k) e^{ikx} dk \\ \phi(k) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \Psi(x, 0) e^{-ikx} dx \end{aligned}$$

Example: A free particle, initially localized in $[-a, a]$ is released at time $t = 0$

$$\Psi(x, 0) = \begin{cases} A, & -a < x < a \\ 0, & \text{otherwise} \end{cases}$$

A and a are positive real constant. Find $\Psi(x, t)$

Solution: Normalize $\Psi(x, 0)$, $A = 1/\sqrt{2a}$

$$\begin{aligned}\phi(k) &= \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{2a}} \int_{-a}^a e^{-ikx} dx \\ &= \frac{1}{2\sqrt{\pi a}} \frac{e^{-ikx}}{-ik} \Big|_{-a}^a = \frac{1}{k\sqrt{\pi a}} \left(\frac{e^{ika} - e^{-ika}}{2i} \right) \\ &= \frac{1}{\sqrt{\pi a}} \frac{\sin(ka)}{k}\end{aligned}$$

Finally, we plug this back into equation (15)

$$\Psi(x, t) = \frac{1}{\pi\sqrt{2a}} \int_{-\infty}^{+\infty} \frac{\sin(ka)}{k} e^{i(kx - \frac{\hbar k}{2m}t)} dk$$

Numerical integration gives figure 14.

It is illuminating to explore the two limiting cases

- In the limiting case $a \rightarrow 0$, we use the small angle approximation

$$\phi(k) = \frac{1}{\sqrt{\pi a}} \frac{\sin(ka)}{k} \approx \frac{1}{\sqrt{\pi a}} a = \sqrt{\frac{a}{\pi}}$$

it is flat. This is an example of uncertainty principle: if the spread in position is small, the spread in momentum must be large. (Figure 15)

- At the other extreme $a \rightarrow \infty$, the spread in position is broad

$$\phi(k) = \sqrt{\frac{a}{\pi}} \frac{\sin(ka)}{ka}$$

it is a sharp spike about $k = 0$. (Figure 16)

I return now to the paradox noted earlier—the fact that the separable solution $\Psi_k(x, t)$ travels at the “wrong” speed for the particle it ostensibly represents. The essential idea is this: A wave packet is a sinusoidal function whose amplitude is modulated by (Figure 17); it consists of “ripples” contained within an “envelope”. What corresponds to the particle velocity is not the speed of the individual ripples (the so-called phase velocity v_p), but rather the speed of the envelope (the group velocity v_g)—which, depending on the nature of the waves, can be greater than, less than, or equal to the velocity of the ripples that go to make it up. What I need to show is that for the wave function of a free particle in quantum mechanics the group velocity is twice the phase velocity—just right to represent the classical particle speed.

$$\begin{aligned}v_{\text{quantum}} &= \text{phase velocity} \\ v_{\text{classical}} &= \text{group velocity}\end{aligned}$$

For a monochromatic plane wave,

$$\Psi_k(x, t) = Ae^{i(kx - \omega t)}$$

the phase plane is given by

$$kx - \omega t = \text{constant}$$

The velocity is the **phase velocity** $v_p = \frac{\omega}{k}$. The problem, then, is to determine the group velocity of a wave packet with the

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

The discussion here applies to all kinds of dispersion relation - the formula for ω as a function of k .

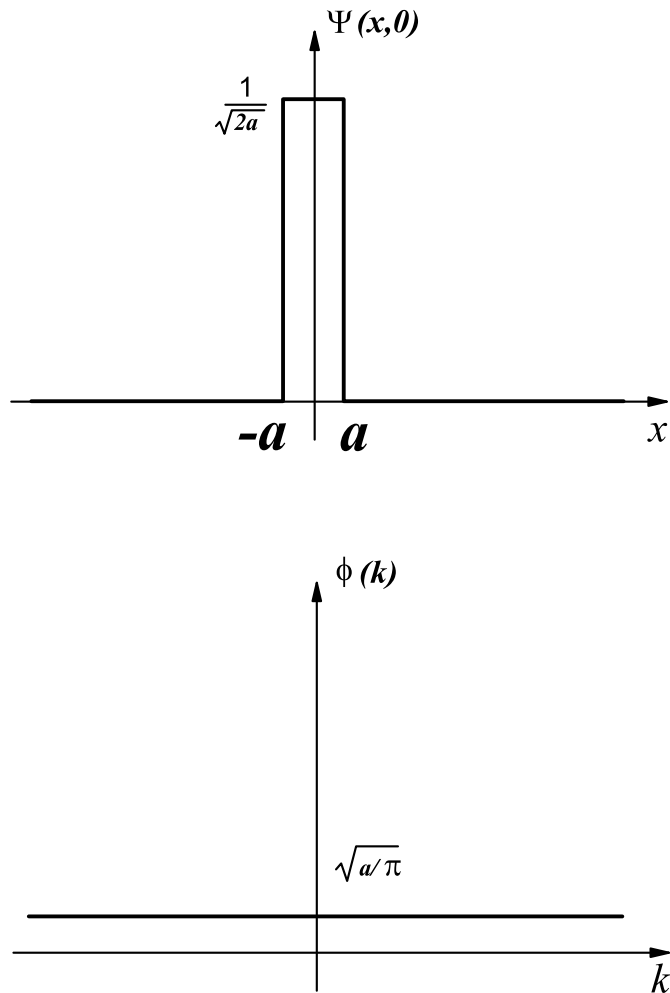


FIG. 15: For small a , graphs of $\Psi(x, 0)$ and $\phi(k)$.

The center of the wave packet is determined by taking the extreme value of the phase angle $\varphi = kx - \omega t$

$$\frac{\partial \varphi}{\partial k} = 0 \rightarrow x - \left(\frac{d\omega}{dk}\right)t = 0$$

So the center of the wave packet is at

$$x = x_c = \left(\frac{d\omega}{dk}\right)t$$

its velocity is called the **group velocity**

$$v_g = \frac{dx}{dt} = \frac{d\omega}{dk}$$

In the case of free particle $\omega = \frac{\hbar k^2}{2m}$, so

$$v_g = \frac{\hbar k}{m}$$

$$v_p = \frac{\hbar k}{2m}$$

$$v_g = 2v_p$$

v_g is twice as great as v_p . This confirms that it is the group velocity of the wave packet, not the phase velocity of the stationary states, that matches the classical particle velocity

$$v_{\text{classical}} = v_{\text{group}} = 2v_{\text{quantum}} = 2v_{\text{phase}}$$

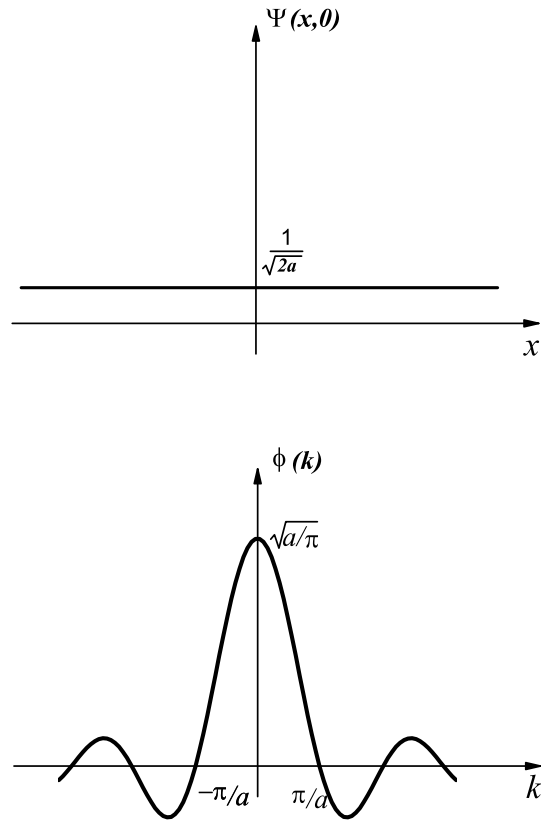


FIG. 16: For large a , graphs of $\Psi(x, 0)$ and $\phi(k)$.

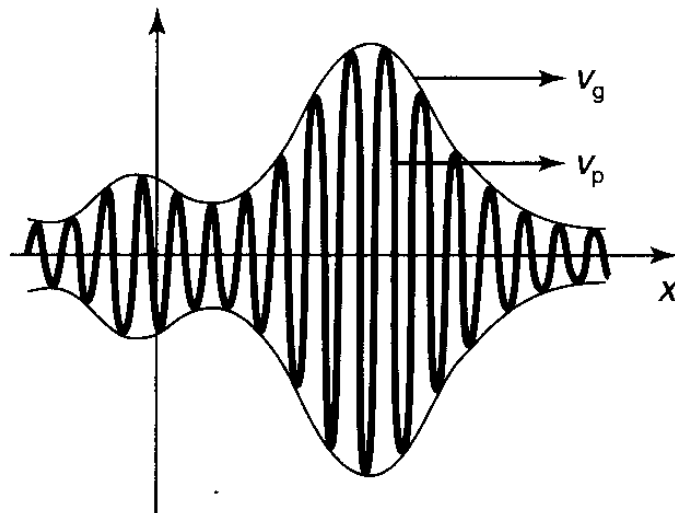


FIG. 17: A wave packet. The “envelop” travels at the group velocity; the “ripples” travel at the phase velocity.

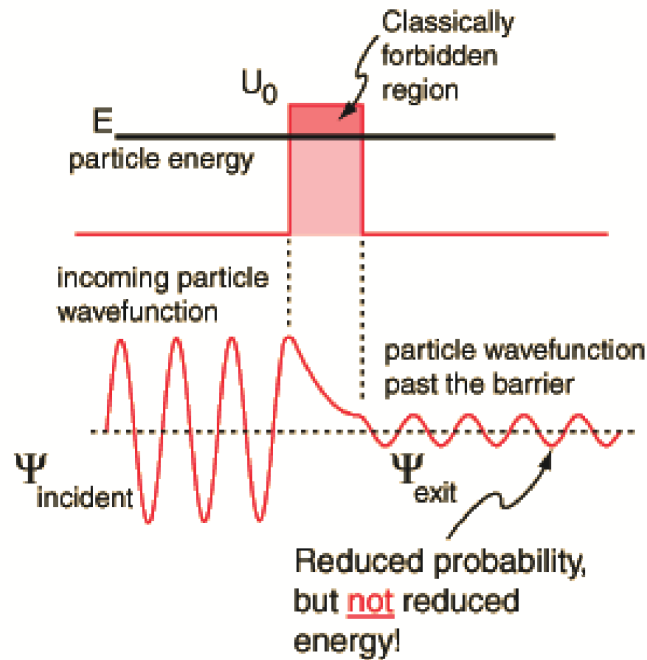


FIG. 18: Barrier penetration model for alpha decay.

Problem 5 *Griffiths, page 67, 2.22. page 86, 2.43*

V. PENETRATION OF POTENTIAL BARRIER

According to classical physics, a particle of energy E less than the height V_0 of a barrier could not penetrate - the region inside the barrier is classically forbidden. But the wave function associated with a free particle must be continuous at the barrier and will show an exponential decay inside the barrier. The wave function must also be continuous on the far side of the barrier, so there is a finite probability that the particle will tunnel through the barrier.

A. Case I: Barrier penetration $E < V_0$

Composed of two protons and two neutrons, the alpha particle is a nucleus of the element helium. As a typical 1D scattering problem, alpha particle emission is modeled as a barrier penetration process. Suppose a particle with energy $E < V_0$ approaches the barrier

$$V(x) = \begin{cases} V_0, & 0 < x < a \\ 0, & x < 0, x > a \end{cases}$$

from the left along the positive x -axis, it is described by a free particle wave function. The Schrödinger equation

$$-\frac{\hbar^2}{2m} \frac{d^2}{dx^2} \psi(x) + V(x)\psi(x) = E\psi(x)$$

has different solutions in different regions. Outside the barrier, the solution is

$$\psi(x) \sim e^{\pm ikx}, \quad k = \sqrt{2mE}/\hbar$$

For $x < 0$, we have the incident wave and reflected wave, while for $x > a$ we only have the transmitted wave.

$$\psi(x) = \begin{cases} e^{ikx} + Re^{-ikx}, & x < 0 \\ Se^{ikx}, & x > a \end{cases}$$

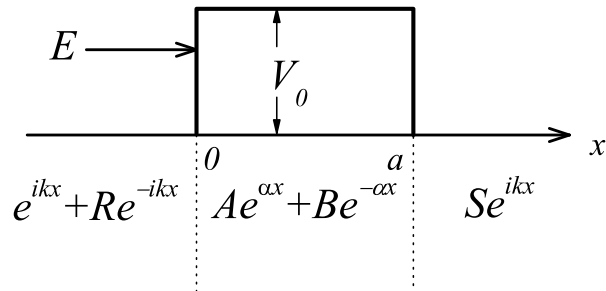


FIG. 19: Scattering from a potential barrier.

We define the current density of probability for the incoming wave function

$$j_i = \frac{-i\hbar}{2m} \left(e^{-ikx} \frac{\partial}{\partial x} e^{ikx} - c.c. \right) = \frac{\hbar k}{m} = v$$

while for the reflected wave and the transmitted wave

$$j_r = |R|^2 v$$

$$j_t = |S|^2 v$$

Therefore the reflection coefficient and transmission coefficient are

$$j_r/j_i = |R|^2$$

$$j_t/j_i = |S|^2$$

Inside the barrier, i.e. in the classically forbidden region $0 < x < a$, the solution is

$$\psi(x) = Ae^{\alpha x} + Be^{-\alpha x}$$

with

$$\alpha = \frac{\sqrt{2m(V_0 - E)}}{\hbar} \quad (16)$$

1. Boundary conditions

There are four boundary conditions: continuity of $\psi(x)$ and $\psi'(x)$ at 0 says

$$1 + R = A + B$$

$$ik - ikR = \alpha A - \alpha B$$

continuity of $\psi(x)$ and $\psi'(x)$ at a gives

$$Ae^{\alpha a} + Be^{-\alpha a} = Se^{ika}$$

$$\alpha Ae^{\alpha a} - \alpha Be^{-\alpha a} = ikSe^{ika}$$

We can solve the equations for A, B, R, S as following

$$1 + R = A + B$$

$$\frac{ik}{\alpha} (1 - R) = A - B$$

$$A = \frac{1}{2} \left(1 + \frac{ik}{\alpha} \right) + \frac{1}{2} \left(1 - \frac{ik}{\alpha} \right) R$$

$$B = \frac{1}{2} \left(1 - \frac{ik}{\alpha} \right) + \frac{1}{2} \left(1 + \frac{ik}{\alpha} \right) R$$

$$Ae^{\alpha a} + Be^{-\alpha a} = Se^{ika}$$

$$Ae^{\alpha a} - Be^{-\alpha a} = \frac{ik}{\alpha} Se^{ika}$$

$$A = \frac{1}{2} \left(1 + \frac{ik}{\alpha} \right) Se^{ika} e^{-\alpha a}$$

$$B = \frac{1}{2} \left(1 - \frac{ik}{\alpha} \right) Se^{ika} e^{\alpha a}$$

we thus have

$$\begin{aligned} \left(1 + \frac{ik}{\alpha} \right) + \left(1 - \frac{ik}{\alpha} \right) R &= \left(1 + \frac{ik}{\alpha} \right) Se^{ika} e^{-\alpha a} \\ \left(1 - \frac{ik}{\alpha} \right) + \left(1 + \frac{ik}{\alpha} \right) R &= \left(1 - \frac{ik}{\alpha} \right) Se^{ika} e^{\alpha a} \end{aligned}$$

$$1 + \frac{1 - ik/\alpha}{1 + ik/\alpha} R = Se^{ika} e^{-\alpha a}$$

$$1 + \frac{1 + ik/\alpha}{1 - ik/\alpha} R = Se^{ika} e^{\alpha a}$$

The equation for transmission amplitude S is

$$\frac{Se^{ika} e^{-\alpha a} - 1}{Se^{ika} e^{\alpha a} - 1} = \left(\frac{1 - ik/\alpha}{1 + ik/\alpha} \right)^2$$

$$Se^{ika} e^{-\alpha a} - 1 = Se^{ika} e^{\alpha a} \left(\frac{1 - ik/\alpha}{1 + ik/\alpha} \right)^2 - \left(\frac{1 - ik/\alpha}{1 + ik/\alpha} \right)^2$$

$$\begin{aligned} Se^{ika} &= \frac{1 - \left(\frac{1 - ik/\alpha}{1 + ik/\alpha} \right)^2}{e^{-\alpha a} - e^{\alpha a} \left(\frac{1 - ik/\alpha}{1 + ik/\alpha} \right)^2} = \frac{(1 + ik/\alpha)^2 - (1 - ik/\alpha)^2}{e^{-\alpha a} (1 + ik/\alpha)^2 - e^{\alpha a} (1 - ik/\alpha)^2} \\ &= \frac{-4ik/\alpha}{e^{\alpha a} (1 - 2ik/\alpha - k^2/\alpha^2) - e^{-\alpha a} (1 + 2ik/\alpha - k^2/\alpha^2)} \\ &= \frac{-4ik/\alpha}{(e^{\alpha a} - e^{-\alpha a}) (1 - k^2/\alpha^2) - 2ik/\alpha (e^{\alpha a} + e^{-\alpha a})} \\ &= \frac{-2ik/\alpha}{(1 - k^2/\alpha^2) \sinh \alpha a - 2ik/\alpha \cosh \alpha a} \\ &= \frac{-2ik\alpha}{(\alpha^2 - k^2) \sinh \alpha a - 2ik\alpha \cosh \alpha a} \end{aligned}$$

The transmission coefficient is

$$\begin{aligned} T = |S|^2 &= \frac{4k^2\alpha^2}{(\alpha^2 - k^2)^2 \sinh^2 \alpha a + 4k^2\alpha^2 \cosh^2 \alpha a} \\ &= \frac{4k^2\alpha^2}{(\alpha^2 + k^2)^2 \sinh^2 \alpha a + 4k^2\alpha^2} \end{aligned}$$

where we have used

$$\cosh^2 x - \sinh^2 x = 1$$

Inserting $\alpha = \sqrt{2m(V_0 - E)}/\hbar$ and $k = \sqrt{2mE}/\hbar$, we can express T in a standard form

$$T^{-1} = 1 + \frac{(\alpha^2 + k^2)^2}{4k^2\alpha^2} \sinh^2 \alpha a$$

$$T^{-1} = 1 + \frac{V_0^2}{4E(V_0 - E)} \sinh^2 \left(\frac{a}{\hbar} \sqrt{2m(V_0 - E)} \right) \quad (17)$$

In this case $E < V_0$, classically the particle can not penetrate through the barrier - it will be reflected completely. Quantum mechanics tells us that the particle has a probability to go through the barrier, the transmission coefficient is nonzero. The process by which a particle can penetrate a classically forbidden region of space is called **Quantum Tunnelling (or Tunneling) Effect**.

For very high and wide barrier, i.e., suppose that $\alpha a = \sqrt{\frac{2m(V_0 - E)}{\hbar^2}} a \gg 1$, we have

$$\sinh \alpha a = \frac{1}{2} (e^{\alpha a} - e^{-\alpha a}) = \frac{1}{2} e^{\alpha a} \gg 1.$$

The transmission coefficient becomes

$$T = \frac{4k^2\alpha^2}{(\alpha^2 + k^2)^2 \frac{1}{4} e^{2\alpha a}} = \frac{16k^2\alpha^2}{(\alpha^2 + k^2)^2} e^{-2\alpha a} \sim e^{-2\alpha a} = \exp \left(-2 \sqrt{\frac{2m(V_0 - E)}{\hbar^2}} a \right)$$

It depends on the mass of the particle m , the barrier width a , and the barrier height V_0 sensitively. Gamow explained the alpha decay phenomena by means of quantum tunneling effect (See Figure (18)). It is easy to generalize our calculation to more complicated potential, the transmission coefficient, i.e., the ratio of probability current density of transmitted wave to that of incident wave, is

$$T = \exp \left(-\frac{2}{\hbar} \int_a^b \sqrt{2m(V(x) - E)} dx \right)$$

B. Case II: Barrier penetration $E > V_0$

In the case of $E > V_0$, we need to modify the definition (16) into

$$\alpha = \frac{\sqrt{2m(V_0 - E)}}{\hbar} = ik'$$

$$k' = \frac{\sqrt{2m(E - V_0)}}{\hbar}$$

so that

$$\sinh \alpha a = \sinh ik'a = i \sin k'a$$

Correspondingly, we have (note $k' \leq k$)

$$T = \frac{4k^2 k'^2}{(k^2 - k'^2)^2 \sin^2 k'a + 4k^2 k'^2}$$

$$T^{-1} = 1 + \frac{V_0^2}{4E(E - V_0)} \sin^2 \left(\frac{a}{\hbar} \sqrt{2m(E - V_0)} \right) \quad (18)$$

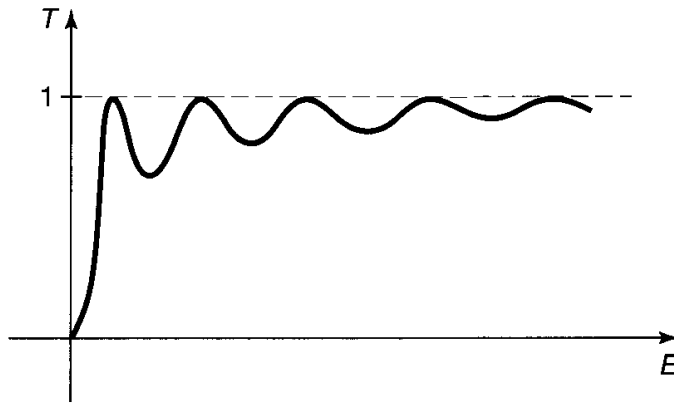


FIG. 20: Transmission coefficient as a function of energy.

C. Case III: Bound states in finite potential well $-V_0 < E < 0$

This is what we have learnt in Section V where we should define now $\beta = \frac{\sqrt{2m(E+V_0)}}{\hbar}$.

D. Case IV: Scattering from square well $-V_0 < 0 < E$

A potential well can scatter the particle and there is a finite probability for the particle to be reflected by the potential well. This can not be understood in classical physics. In this case we simply change V_0 to $-V_0$ and

$$k' = \frac{\sqrt{2m(E+V_0)}}{\hbar} \geq k = \frac{\sqrt{2mE}}{\hbar}.$$

The transmission coefficient is

$$T^{-1} = 1 + \frac{V_0^2}{4E(E+V_0)} \sin^2\left(\frac{a}{\hbar}\sqrt{2m(E+V_0)}\right) \quad (19)$$

It is clearly seen that for $V_0 = 0$, i.e., $k = k'$, then $T = 1$, we have the perfect transmission as expected. But for $V_0 \neq 0$ we generally have $T < 1$. This is essentially a quantum effect. For given potential well, the transmission coefficient depends on the energy of the incoming particle as shown in Figure 20. Notice that $T = 1$ (the well becomes transparent) whenever the sine is zero, which is to say, when

$$\frac{a}{\hbar}\sqrt{2m(E_n+V_0)} = n\pi$$

where n is an integer. The energies for perfect transmission, then, are given by

$$E_n = -V_0 + \frac{n^2\pi^2\hbar^2}{2ma^2}$$

which happen to be precisely the allowed energies for the infinite square well. The discussion about resonant transmission here applies to Case II equally.

We summarize this section in Figure 21.

Problem 6 Zeng Jinyan, *Textbook*, page, 80-84, 5, 10

VI. THE δ -FUNCTION POTENTIAL

A. Bound states and scattering states

We have encountered two very different kinds of solutions to the time-independent Schrödinger equation:

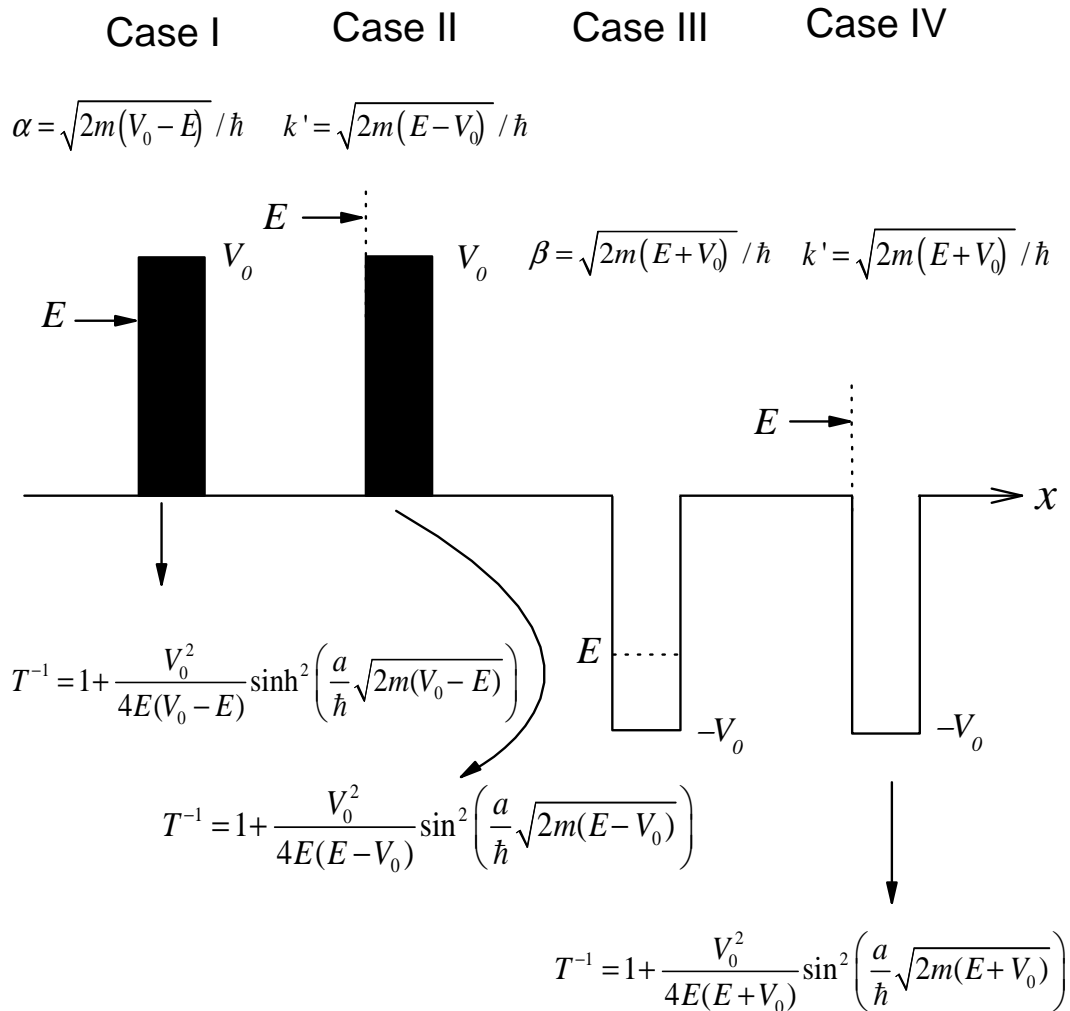


FIG. 21: Barrier or well penetration: Transmission coefficients.

infinite square well	normalizable	discrete index n	physically	general solution
harmonic oscillator			realizable states	sum over n
free particle	non-normalizable	continuous variable k	non-realizable	integral over k

In classical mechanics a one-dimensional time-independent potential can give rise to two rather different kinds of motion. If $V(x)$ rises higher than E on either side (Figure 22a), then the particle is "stuck" in the potential well—it rocks back and forth between the turning points, but it cannot escape. We call this a **bound state**. If, on the other hand, E exceeds $V(x)$ on one side (or both), then the particle comes in from "infinity", slows down or speeds up under the influence of the potential, and returns to infinity (Figure 22b). We call this a **scattering state**. Some potentials admit only bound states (the harmonic oscillator, infinite square well); some allow only scattering states (a potential barrier, free particle); some permit both kinds, depending on the energy of the particle (finite square well, δ -function potential well).

As you have probably guessed, the two kinds of solutions to the Schrödinger equation correspond precisely to bound and scattering states. The distinction is even cleaner in the quantum domain, because the phenomenon of tunneling allows the particle to "leak" through any finite potential barrier, so the only thing that matters is the potential at infinity

$$E < V(-\infty) \text{ and } V(+\infty) \text{ bound state}$$

$$E > V(-\infty) \text{ or } V(+\infty) \text{ scattering state}$$

In the case of a cubic potential (Figure 22c), the classical bound state has got a finite lifetime (the eigenenergy is $E = E_0 + i\Gamma$) because of the tunneling through the barrier and quantum mechanically it is a scattering state (though

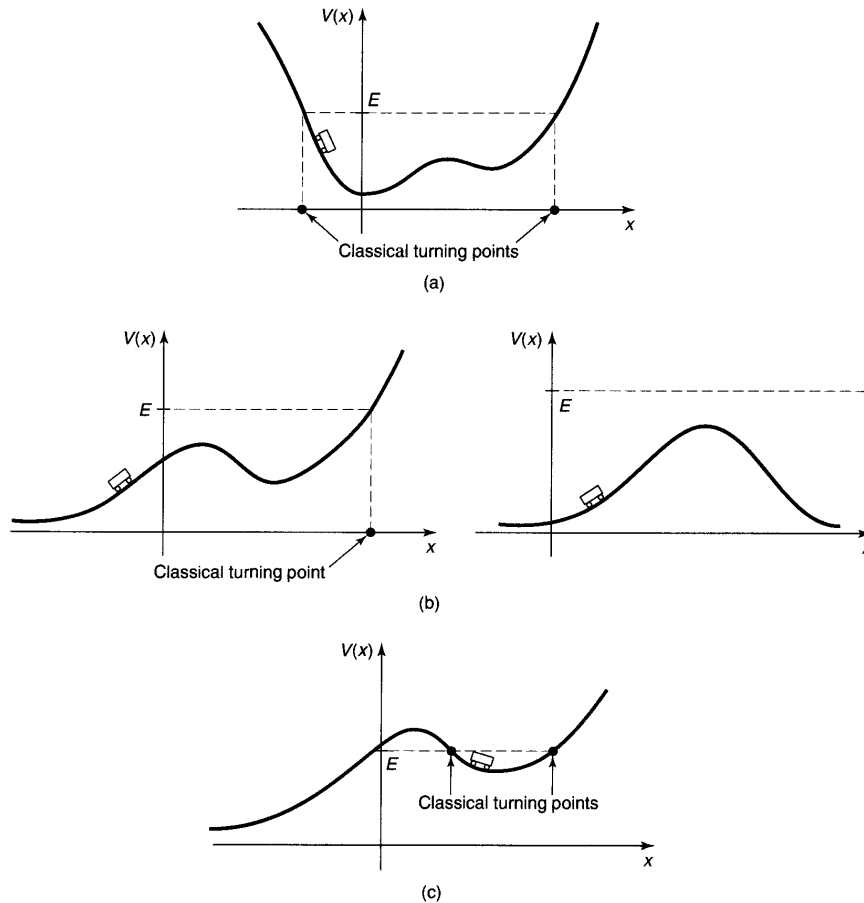


FIG. 22: (a) A bound state. (b) Scattering states. (c) A classical bound state, but a quantum scattering state.

we sometimes still call it a bound state).

B. From square potential to δ -potential

Basically the square potential well or barrier is some kind of idealization of practical potentials. We can further squeeze the width of the well/barrier to very small and depth to very high.

square well	depth $\rightarrow -\infty$	$V(x) = -V_0\delta(x - a)$
	width $\rightarrow 0$	
square barrier	depth $\rightarrow +\infty$	$V(x) = V_0\delta(x - a)$
	width $\rightarrow 0$	

The Dirac delta function, $\delta(x)$, is defined informally as follows:

$$\delta(x) = \begin{cases} 0, & \text{if } x \neq 0 \\ \infty, & \text{if } x = 0 \end{cases}$$

It is an infinitely high, infinitesimally narrow spike at the origin, whose area is 1. In particular

$$\int_{-\infty}^{+\infty} f(x)\delta(x - a)dx = f(a). \quad (20)$$

That's the most important property of the delta function: Under the integral sign it serves to “pick out” the value of $f(x)$ at the point a .

What happens in a δ -potential (e.g. potential well)? Let's consider a potential of the form

$$V(x) = -V_0\delta(x - a)$$

The Schrödinger equation reads

$$-\frac{\hbar^2}{2m} \frac{d^2}{dx^2} \psi(x) - V_0\delta(x - a)\psi(x) = E\psi(x)$$

$$\frac{d^2}{dx^2} \psi(x) = -\frac{2m}{\hbar^2} (V_0\delta(x - a) + E) \psi(x)$$

We integrate on both sides in a small range of x near a , in which case we can safely drop E because V goes to infinite here

$$\int_{a-\epsilon}^{a+\epsilon} \psi''(x) dx = \int_{a-\epsilon}^{a+\epsilon} \left(-\frac{2m}{\hbar^2} V_0\delta(x - a)\psi(x) \right) dx$$

Use equation (20) we have

$$\psi'(a + \epsilon) - \psi'(a - \epsilon) = -\frac{2mV_0}{\hbar^2} \psi(a) \quad (21)$$

Thus there is a sudden change in the derivative of the wave function.

On the other hand the wave function is continuous at a

$$\begin{aligned} \psi(a + \epsilon) - \psi(a - \epsilon) &= \int_{a-\epsilon}^{a+\epsilon} \psi'(x) dx = \int_{a-\epsilon}^a \psi'(a - \epsilon) dx + \int_a^{a+\epsilon} \psi'(a + \epsilon) dx \\ &= \psi'(a - \epsilon) \cdot \epsilon + \psi'(a + \epsilon) \cdot \epsilon \underset{\epsilon \rightarrow 0}{=} 0 \end{aligned}$$

We notice that the boundary conditions in the δ -potential model are

1. ψ is always continuous, and
2. $d\psi/dx$ is continuous except at points where the potential is infinite.

Introduction of δ -potential \Rightarrow two interfaces combined into one \Rightarrow greatly simplified the mathematical procedure \Rightarrow more easily to be handled \Rightarrow on the other hand, the information provided by this model is also limited

The Schrödinger equation for the δ -function well yields both bound states $E < 0$ and scattering states $E > 0$.

C. Bound states in δ potential well ($E < 0$)

The equation is now

$$-\frac{\hbar^2}{2m} \frac{d^2}{dx^2} \psi(x) + V(x)\psi(x) = E\psi(x)$$

where $V(x) = -V_0\delta(x)$ and $E < 0$. In the region $x \neq 0$

$$\begin{aligned} -\frac{\hbar^2}{2m} \frac{d^2}{dx^2} \psi(x) &= E\psi(x) \\ \frac{d^2}{dx^2} \psi(x) - \left(-\frac{2mE}{\hbar^2} \right) \psi(x) &= 0 \end{aligned}$$

We define the parameter $\alpha = \sqrt{\frac{-2mE}{\hbar^2}}$, so the wave function can be $\psi(x) \sim \exp(\alpha x)$ or $\exp(-\alpha x)$

$$\begin{aligned} x < 0, \quad \psi(x) &= A \exp(\alpha x) \\ x > 0, \quad \psi(x) &= B \exp(-\alpha x) \end{aligned}$$

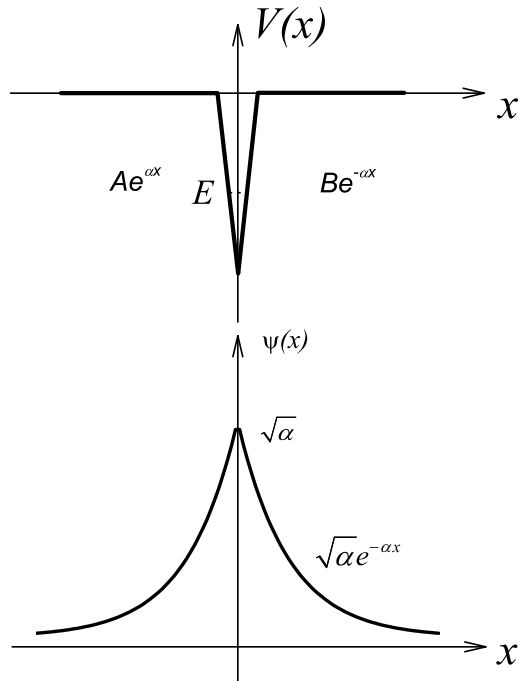


FIG. 23: Bound states in δ potential well ($E < 0$).

The boundary conditions say

$$\begin{aligned} A &= B \\ -\alpha B - \alpha A &= -\frac{2mV_0}{\hbar^2}A \end{aligned}$$

which gives $\alpha = \frac{mV_0}{\hbar^2}$, i.e.

$$-\frac{2mE}{\hbar^2} = \frac{m^2V_0^2}{\hbar^4}$$

The energy may only take the value

$$E = -\frac{mV_0^2}{2\hbar^2} \quad (22)$$

The wave function

$$\psi(x) = \begin{cases} A \exp(\alpha x), & \text{for } x < 0 \\ A \exp(-\alpha x), & \text{for } x > 0 \end{cases}$$

Finally we normalize ψ

$$\begin{aligned} \int_{-\infty}^{+\infty} |\psi(x)|^2 dx &= |A|^2 \left(\int_{-\infty}^0 e^{2\alpha x} dx + \int_0^{+\infty} e^{-2\alpha x} dx \right) \\ &= 2|A|^2 \int_0^{+\infty} e^{-2\alpha x} dx = \frac{|A|^2}{\alpha} = 1 \end{aligned}$$

so

$$A = \sqrt{\alpha}$$

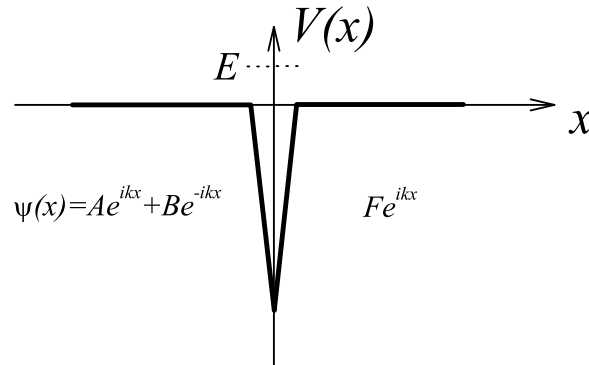


FIG. 24: Scattering states on δ potential well ($E > 0$).

Evidently the delta function well, regardless of its "strength" V_0 , has exactly one bound state

$$\psi(x) = \frac{\sqrt{mV_0}}{\hbar} \exp(-mV_0|x|/\hbar^2) \quad (23)$$

which is illustrated in Figure 23. We notice that $\psi(-x) = \psi(x)$, thus the wave function has even parity. This also excludes the odd parity bound states for delta function potential well.

D. Scattering states in δ -potential well ($E > 0$)

What about scattering states, with $E > 0$? The general solution for $x < 0$ is

$$\psi(x) = Ae^{ikx} + Be^{-ikx}$$

where $k = \sqrt{\frac{2mE}{\hbar^2}}$ is real and positive. For $x > 0$, we have

$$\psi(x) = Fe^{ikx}$$

The boundary conditions that ψ is always continuous while $d\psi/dx$ has a jump at $x = 0$ require that

$$A + B = F$$

and

$$\frac{d\psi}{dx}|_+ = ikF, \quad \frac{d\psi}{dx}|_- = ik(A - B)$$

$$\psi'(\epsilon) - \psi'(-\epsilon) = \frac{d\psi}{dx}|_+ - \frac{d\psi}{dx}|_- = -\frac{2mV_0}{\hbar^2}\psi(0)$$

which means

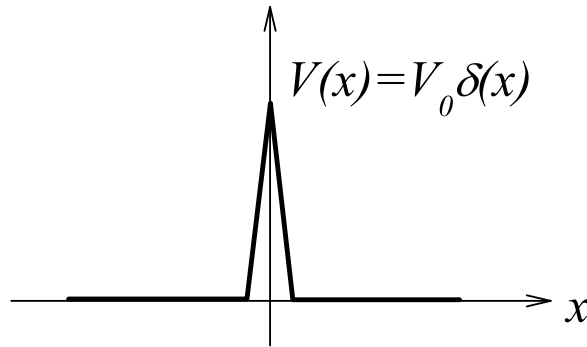
$$ik(F - A + B) = -\frac{2mV_0}{\hbar^2}F$$

$$F - A = i\beta F$$

where we have defined $\beta = \frac{mV_0}{\hbar^2 k}$. Finally

$$F = \frac{1}{1 - i\beta}A$$

$$B = \frac{i\beta}{1 - i\beta}A$$

FIG. 25: δ potential barrier.

The reflection and transmission coefficients are

$$R = \frac{|B|^2}{|A|^2} = \frac{\beta^2}{1 + \beta^2}$$

$$T = \frac{|F|^2}{|A|^2} = \frac{1}{1 + \beta^2}$$

Clearly $R + T = 1$. In terms of the original variables, we have

$$R = \frac{1}{1 + 2\hbar^2 E/mV_0^2}$$

$$T = \frac{1}{1 + mV_0^2/2\hbar^2 E}$$

Obviously the higher the energy, the greater the probability of transmission.

The scattering wave functions are not normalizable, so they don't actually represent possible particle states. The solution lies in that we must construct normalizable linear combination of the stationary states, i.e., wave packets, just as we did for the free particles. Tunneling of wave packets through a potential barrier is known as the Hartman effect. Many numerical simulations can be found on the web.

E. δ -function barrier

As long as we've got the relevant equations on the table, let's look briefly at the case of a delta-function barrier

$$V(x) = V_0 \delta(x - a).$$

Formally, all we have to do is change the sign of V_0 . This kills the bound state, of course. On the other hand, the reflection and transmission coefficients, which depend only on V_0^2 , are unchanged. Strange to say, the particle is just as likely to pass through the barrier as to cross over the well!

Problem 7 *Griffiths, page 87, 2.44*

VII. SUMMARY ON PART II

In very few cases, we can solve the stationary Schrödinger equation exactly and some examples are given in this part.

- Square potential well of width a : eigenvalues

$$E_n = \frac{n^2 \pi^2 \hbar^2}{2ma^2}$$

inside the well, wavefunction has definite parity

$$\psi_n(x) = \begin{cases} \sqrt{\frac{2}{a}} \cos\left(\frac{n\pi}{a}x\right), & \text{for } n = 1, 3, 5 \dots \\ \sqrt{\frac{2}{a}} \sin\left(\frac{n\pi}{a}x\right), & \text{for } n = 2, 4, 6 \dots \end{cases}$$

- Simple harmonic oscillator: eigenvalues

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

wavefunction

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

- Penetration of potential barrier of width a : transmission coefficient

$$T \sim \exp\left(-2\sqrt{\frac{2m(V_0 - E)}{\hbar^2}}a\right)$$

- δ function potential: only one bound state

$$\psi(x) = \frac{\sqrt{mV_0}}{\hbar} \exp(-mV_0|x|/\hbar^2)$$

transmission coefficient for δ function potential barrier

$$T = \frac{1}{1 + mV_0^2/2\hbar^2 E}$$

For finite square potential, one usually can find the graphical solution. The basic procedure in solving the one dimensional potential is: assume reasonable solutions in different regions and connect them through boundary conditions, i.e., the wavefunctions ψ is always continuous, and $d\psi/dx$ is continuous except at points where the potential is infinite. There exists a jump there

$$\psi'(a + \epsilon) - \psi'(a - \epsilon) = -\frac{2mV_0}{\hbar^2}\psi(a).$$

The generic problem of quantum mechanics is: Given the time-independent potential $V(x)$ and the initial wave function $\Psi(x, 0)$, find $\Psi(x, t)$ for any subsequent time t . We first solve the stationary (time-independent) Schrödinger equation

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

which yields, in general, an infinite set of solutions

$$\begin{array}{cccc} \psi_1(x) & \psi_2(x) & \psi_3(x) & \dots \\ E_1 & E_2 & E_3 & \dots \end{array}$$

We then fit $\Psi(x, 0)$ by writing down the general linear combination of these solutions

$$\Psi(x, 0) = \sum_{n=1}^{\infty} c_n \psi_n(x)$$

and match the special initial state by appropriate choice of the constants c_1, c_2, c_3, \dots . The general solution is given by simply attaching the time evolution factor

$$\Psi(x, t) = \sum_{n=1}^{\infty} c_n \psi_n(x) \exp(-iE_n t/\hbar)$$

which solve the time-dependent Schrödinger equation

$$i\hbar \frac{\partial \Psi}{\partial t} = -\frac{\hbar^2}{2m} \frac{\partial^2 \Psi}{\partial x^2} + V(x) \Psi.$$

Lecture Notes on Quantum Mechanics - Part III

Yunbo Zhang

Instituter of Theoretical Physics, Shanxi University

We introduce in this part operators for dynamical variables such as coordinate, momentum, etc. Especially we deal with the angular momentum operator, and then find the quantized energy levels for the simplest atom - hydrogen, by solving the stationary Schrödinger equation for a typical three dimensional system.

Contents

I. Operator Representing Physical Quantities	2
A. Operators	2
B. The Hermitian operators	3
C. The eigenvalue and eigenfunction of a Hermitian operator	5
D. The momentum operator	5
II. Eigenfunctions of Orbital Angular Momentum	7
A. Angular momentum operator in spherical coordinates	7
B. Eigenfunctions of orbital angular momentum	9
1. Series solution for $m = 0$ case	11
2. Series solution for $m \neq 0$ case	12
C. Eigenfunction of \hat{L}_z and \hat{L}^2 , the spherical harmonics	14
III. Central Force Field	16
A. General properties of central force field problem	16
1. Separation of variables in central force field problem	16
2. Effective potential – the centrifugal potential	17
3. Change of variables	17
B. Reduced mass	18
C. Spherical square potential	19
1. Infinite spherical well	19
2. Finite spherical well	22
IV. The Hydrogen Atom	24
A. Radial equation and its solution	24
1. Asymptotic behavior of $u(\rho)$ at singular points	25
2. Solution of the radial equation	26
B. The property of eigenstates	27
1. Energy quantization	27
2. Angular momentum and degeneracy	28
3. The wave function	29
4. Summary on the quantum numbers	32
V. Summary on Part III	33

I. OPERATOR REPRESENTING PHYSICAL QUANTITIES

A. Operators

By dynamical variables we mean physical quantities that we are familiar in classical mechanics (Newtonian Mechanics), such as coordinate, momentum, energy, orbital angular momentum, etc. The states of microsystems have been proved to be wave motions and are described by wave functions Ψ . What are the fate of these physical quantities?

- The formalism of wave function system and its statistical interpretation will inevitably lead to statistical answers on the information of physical quantities, such as probability distributions, expectation values, standard deviations, etc.
- The expectation values of coordinates and momenta have been deduced and are expressed as

$$\langle x \rangle = \iiint \psi^*(x, y, z) x \psi(x, y, z) dx dy dz,$$

$$\begin{aligned} \langle p_x \rangle &= \iiint \phi^*(p_x, p_y, p_z) p_x \phi(p_x, p_y, p_z) dp_x dp_y dp_z \\ &= \iiint \psi^*(x, y, z) \left(-i\hbar \frac{\partial}{\partial x} \right) \psi(x, y, z) dx dy dz. \end{aligned}$$

Thus, the idea of operators are introduced. The operator is a mathematical symbol to represent certain mathematical operation. If we want to know the expectation value of a physical quantity, we substitute the corresponding *differential operator* into the expression of expectation values. The operators of coordinate and momentum are

$$\hat{x} = x, \quad \hat{p}_x = -i\hbar \frac{\partial}{\partial x},$$

respectively.

- Generalization: For any physical quantity F , called dynamical variable, there corresponds an operator \hat{F} such that

$$\langle F \rangle = \iiint \psi^*(x, y, z) \hat{F} \psi(x, y, z) dx dy dz.$$

Here is how the operator \hat{F} is obtained: write out the explicit form of physical quantity F in Cartesian coordinates, change the coordinate and momenta into corresponding operators. Thus the operator for energy is

$$\begin{aligned} H &= \frac{1}{2m} (p_x^2 + p_y^2 + p_z^2) + V(x, y, z) \\ &\implies \frac{1}{2m} \left[\left(-i\hbar \frac{\partial}{\partial x} \right)^2 + \left(-i\hbar \frac{\partial}{\partial y} \right)^2 + \left(-i\hbar \frac{\partial}{\partial z} \right)^2 \right] + V(x, y, z) \\ &= -\frac{\hbar^2}{2m} \nabla^2 + V(\mathbf{r}) = \hat{H} \end{aligned}$$

The orbital angular momentum

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

and its components

$$L_x = yp_z - zp_y, \quad L_y = zp_x - xp_z, \quad L_z = xp_y - yp_x,$$

are changed into operators

$$\hat{\mathbf{L}} = \hat{\mathbf{r}} \times \hat{\mathbf{p}}$$

$$\hat{L}_x = -i\hbar \left(y \frac{\partial}{\partial z} - z \frac{\partial}{\partial y} \right),$$

$$\hat{L}_y = -i\hbar \left(z \frac{\partial}{\partial x} - x \frac{\partial}{\partial z} \right),$$

$$\hat{L}_z = -i\hbar \left(x \frac{\partial}{\partial y} - y \frac{\partial}{\partial x} \right).$$

This, actually, constitutes the second Basic Postulate of quantum mechanics: *All physical quantities are changed into and are expressed by Hermitian operators.* The introduction of operators to represent physical quantities in quantum mechanics is of primary importance and has profound consequences.

Wave functions for states
Operators for physical quantities } cornerstones of quantum mechanics

B. The Hermitian operators

The basic premise here is that the expectation value of any real physical quantity must be real. Thus, it is required that

$$\langle F \rangle = \langle F \rangle^*$$

so that,

$$\begin{aligned} \int \psi^* \hat{F} \psi d^3 \mathbf{r} &= \left[\int \psi^* \hat{F} \psi d^3 \mathbf{r} \right]^* \\ &= \int (\hat{F} \psi)^* \psi d^3 \mathbf{r} \end{aligned}$$

thus this equation or equivalently in our short hand notation

$$(\psi, \hat{F} \psi) = (\hat{F} \psi, \psi)$$

serves as a criterion to justify whether an operator may represent a physical quantity. Those that satisfy the above criterion are called **Hermitian or Hermitean operators**, in the name of French mathematician Hermite.

Proposition 1 *General criterion for Hermitian operators. If*

$$\int \psi^* \hat{A} \psi d^3 \mathbf{r} = \int (\hat{A} \psi)^* \psi d^3 \mathbf{r} \longrightarrow (\psi, \hat{A} \psi) = (\hat{A} \psi, \psi) \quad (1)$$

for any state ψ , then there must be

$$\int \psi_1^* \hat{A} \psi_2 d^3 \mathbf{r} = \int (\hat{A} \psi_1)^* \psi_2 d^3 \mathbf{r} \longrightarrow (\psi_1, \hat{A} \psi_2) = (\hat{A} \psi_1, \psi_2). \quad (2)$$

Proof. First let $\psi = \psi_1 + \psi_2$, Eq. (1) gives

$$((\psi_1 + \psi_2), \hat{A}(\psi_1 + \psi_2)) = (\hat{A}(\psi_1 + \psi_2), (\psi_1 + \psi_2))$$

$$\begin{aligned} &(\psi_1, \hat{A} \psi_1) + (\psi_1, \hat{A} \psi_2) + (\psi_2, \hat{A} \psi_1) + (\psi_2, \hat{A} \psi_2) \\ &= (\hat{A} \psi_1, \psi_1) + (\hat{A} \psi_1, \psi_2) + (\hat{A} \psi_2, \psi_1) + (\hat{A} \psi_2, \psi_2) \end{aligned}$$

and

$$\left(\psi_1, \hat{A}\psi_2\right) + \left(\psi_2, \hat{A}\psi_1\right) = \left(\hat{A}\psi_1, \psi_2\right) + \left(\hat{A}\psi_2, \psi_1\right) \quad (3)$$

Then we set $\psi = \psi_1 + i\psi_2$

$$\left((\psi_1 + i\psi_2), \hat{A}(\psi_1 + i\psi_2)\right) = \left(\hat{A}(\psi_1 + i\psi_2), (\psi_1 + i\psi_2)\right)$$

which leads to (i before the comma contributes a minus sign from the conjugation)

$$\left(\psi_1, \hat{A}\psi_2\right) - \left(\psi_2, \hat{A}\psi_1\right) = \left(\hat{A}\psi_1, \psi_2\right) - \left(\hat{A}\psi_2, \psi_1\right). \quad (4)$$

The combination of (3) and (4) gives

$$\left(\psi_1, \hat{A}\psi_2\right) = \left(\hat{A}\psi_1, \psi_2\right).$$

This will be viewed as a general criterion for Hermitian operators. ■

Example 2 $\hat{x} = x$ and $\hat{p}_x = -i\hbar\frac{d}{dx}$ are Hermitian operators while the differential operator $\frac{d}{dx}$ is not.

Solution 3 The coordinate x is real, obviously

$$\left(\psi_1, x\psi_2\right) = \left(x\psi_1, \psi_2\right)$$

For \hat{p}_x , we have

$$\begin{aligned} \int_{-\infty}^{\infty} \psi_1^* \left(-i\hbar\frac{d}{dx}\right) \psi_2 dx &= -i\hbar \int_{-\infty}^{\infty} \psi_1^* d\psi_2 \\ &= -i\hbar \psi_1^* \psi_2 \Big|_{-\infty}^{\infty} - (-i\hbar) \int_{-\infty}^{\infty} (d\psi_1^*) \psi_2 \\ &= i\hbar \int_{-\infty}^{\infty} \left(\frac{d\psi_1}{dx}\right)^* \psi_2 dx \\ &= \int_{-\infty}^{\infty} \left(-i\hbar\frac{d}{dx}\psi_1\right)^* \psi_2 dx. \end{aligned}$$

where we have used the fact that both ψ_1 and ψ_2 are bound states so $\psi_1^* \psi_2 \Big|_{-\infty}^{\infty} = 0$. In short

$$\left(\psi_1, \hat{p}_x\psi_2\right) = \left(\hat{p}_x\psi_1, \psi_2\right).$$

On the other hand, $\hat{A} = \frac{d}{dx}$ is not a Hermitian operator

$$\begin{aligned} \int \psi_1^* \hat{A}\psi_2 dx &= \int_{-\infty}^{\infty} \psi_1^* \frac{d}{dx} \psi_2 dx = \int_{-\infty}^{\infty} \psi_1^* d\psi_2 \\ &= \psi_1^* \psi_2 \Big|_{-\infty}^{\infty} - \int_{-\infty}^{\infty} \left(\frac{d\psi_1}{dx}\right)^* \psi_2 dx \\ &= \int_{-\infty}^{\infty} \left(-\frac{d}{dx}\psi_1\right)^* \psi_2 dx \\ &\neq \int_{-\infty}^{\infty} \left(\frac{d}{dx}\psi_1\right)^* \psi_2 dx. \end{aligned}$$

In this case, let $\hat{A} = \frac{d}{dx}$, $\hat{B} = -\frac{d}{dx}$, the above relation can be expressed in a form

$$\int_{-\infty}^{\infty} \psi_1^* \hat{A}\psi_2 dx = \int_{-\infty}^{\infty} (\hat{B}\psi_1)^* \psi_2 dx$$

$$\left(\psi_1, \hat{A}\psi_2\right) = \left(\hat{B}\psi_1, \psi_2\right)$$

\hat{A} and \hat{B} are called conjugate operators. $\hat{A} = \frac{d}{dx}$ is not a Hermitian operator, it has a conjugate operator $\hat{B} = -\frac{d}{dx}$. To show the conjugate relation more clearly, \hat{B} is written in a new form \hat{A}^\dagger . For momentum operator, $\hat{A} = \hat{p}_x = -i\hbar\frac{d}{dx}$, the conjugate operator is itself $\hat{p}_x^\dagger = \hat{p}_x$, \hat{p}_x is called a self-conjugate operator. All Hermitian operators are self-conjugate operators.

C. The eigenvalue and eigenfunction of a Hermitian operator

For a statistical problem, the first information we can draw from it are the expectation value

$$\langle F \rangle = \int \psi^* \hat{F} \psi d^3 \mathbf{r}$$

and the variance

$$\langle (\Delta F)^2 \rangle = \int \psi^* (\hat{F} - \langle F \rangle)^2 \psi d^3 \mathbf{r} \geq 0.$$

There exists a special case where the variance or standard deviation is exactly zero, i.e. all measurements of \hat{F} give the same result $\langle F \rangle$. Using the Hermitian property of \hat{F} , we have

$$\begin{aligned} \langle (\Delta F)^2 \rangle &= \int \psi^* (\hat{F} - \langle F \rangle)^2 \psi d^3 \mathbf{r} \\ &= \int \psi^* (\hat{F} - \langle F \rangle) (\hat{F} - \langle F \rangle) \psi d^3 \mathbf{r} \\ &= \int |(\hat{F} - \langle F \rangle) \psi|^2 d^3 \mathbf{r} = 0. \end{aligned}$$

The integrand $(\hat{F} - \langle F \rangle) \psi$ must be identical to zero, i.e.

$$(\hat{F} - \langle F \rangle) \psi(\mathbf{r}) = 0.$$

for all \mathbf{r} values.

In the language of quantum mechanics, if it is wanted that the a physical quantity has definite value, the system must be in a state ψ that satisfies the equation

$$\hat{F} \psi = \langle F \rangle \psi.$$

the above equation is called eigenvalue equation of operator \hat{F} , while ψ and $\langle F \rangle$ are called eigenfunction and eigenvalue of operator \hat{F} , respectively.

Generally, given an operator \hat{F} , there is a series of eigenfunctions and eigenvalues

$$\hat{F} \psi_n = f_n \psi_n$$

Eigenfunctions and eigenvalues are one of the most important points in quantum mechanics. Up to now, we only know that the eigenvalues of Hermitian operators are real. An immediate example is that the Schrödinger equation is essentially the eigenequation for Hamiltonian \hat{H}

$$\hat{H} \psi_n = E_n \psi_n.$$

D. The momentum operator

We already know that the momentum operator $\hat{\mathbf{p}}$

$$\begin{aligned} \hat{\mathbf{p}} &= \mathbf{x}_0 \hat{p}_x + \mathbf{y}_0 \hat{p}_y + \mathbf{z}_0 \hat{p}_z \\ &= \mathbf{x}_0 \left(-i\hbar \frac{\partial}{\partial x} \right) + \mathbf{y}_0 \left(-i\hbar \frac{\partial}{\partial y} \right) + \mathbf{z}_0 \left(-i\hbar \frac{\partial}{\partial z} \right) \\ &= -i\hbar \nabla \end{aligned}$$

is a Hermitian operator. We are interested here in the eigenequation of the momentum operator. It is easily to show that the eigenfunctions are nothing but the monochromatic plane waves

$$\psi = A e^{\frac{i}{\hbar} \mathbf{P} \cdot \mathbf{r}} = A e^{\frac{i}{\hbar} (p_x x + p_y y + p_z z)}$$

and it is easy to check

$$\hat{\mathbf{p}}\psi = (\mathbf{x}_0 p_x + \mathbf{y}_0 p_y + \mathbf{z}_0 p_z)\psi = \mathbf{p}\psi$$

The eigenfunctions are free states, not bound states. So there exists a question of how to normalize them. It is important here to distinguish the definitions of Kronecker delta δ_{mn} for discrete numbers and Dirac delta $\delta(x - y)$ for continuous variables.

- δ -normalization. We choose

$$A = \frac{1}{(2\pi\hbar)^{3/2}},$$

so

$$\psi = \frac{1}{(2\pi\hbar)^{3/2}} e^{\frac{i}{\hbar}(p_x x + p_y y + p_z z)}$$

and

$$\begin{aligned} \int \psi_{p'_x p'_y p'_z}^* \psi_{p_x p_y p_z} dx dy dz &= \frac{1}{(2\pi\hbar)^3} \iiint e^{\frac{i}{\hbar}(p_x - p'_x)x} e^{\frac{i}{\hbar}(p_y - p'_y)y} e^{\frac{i}{\hbar}(p_z - p'_z)z} dx dy dz \\ &= \delta(p_x - p'_x) \delta(p_y - p'_y) \delta(p_z - p'_z) \end{aligned}$$

The eigenvalues p_x, p_y, p_z can change from $-\infty$ to $+\infty$ and the state is innumerable.

- Box-normalization: Consider a box of L^3 in space. Only waves confined in this box is considered. The wave functions obey periodic boundary conditions

$$\begin{aligned} \psi(x, y, z) &= \psi(x + L, y, z) \\ &= \psi(x, y + L, z) \\ &= \psi(x, y, z + L) \end{aligned}$$

For example

$$A e^{\frac{i}{\hbar}(p_x x + p_y y + p_z z)} = A e^{\frac{i}{\hbar}(p_x x + p_x L + p_y y + p_z z)}$$

which gives

$$\frac{p_x L}{\hbar} = n_x 2\pi, \quad \frac{p_y L}{\hbar} = n_y 2\pi, \quad \frac{p_z L}{\hbar} = n_z 2\pi$$

The eigenvalues are therefore discrete

$$p_x = n_x \frac{2\pi\hbar}{L}, \quad p_y = n_y \frac{2\pi\hbar}{L}, \quad p_z = n_z \frac{2\pi\hbar}{L}$$

where n_x, n_y, n_z are integers $0, \pm 1, \pm 2 \dots$. Each separate state occupies a volume $(\frac{2\pi\hbar}{L})^3$ in momentum space. For a volume $\Delta p_x \Delta p_y \Delta p_z$, the number of possible independent momentum eigenstate is

$$\frac{\Delta p_x \Delta p_y \Delta p_z}{(\frac{2\pi\hbar}{L})^3} = \frac{\Delta p_x \Delta p_y \Delta p_z L^3}{(2\pi\hbar)^3} = \frac{\Delta p_x \Delta p_y \Delta p_z \Delta x \Delta y \Delta z}{h^3}$$

which means that *one quantum state occupies a volume h^3 in phase space*. Thus, the eigenstates of momentum become numerable. The normalization factor in this case is

$$A = \frac{1}{\sqrt{L^3}}$$

The eigenfunction now has an explicit expression

$$\psi_{\mathbf{p}}(\mathbf{r}) = \frac{1}{\sqrt{L^3}} e^{\frac{i}{\hbar}\mathbf{p}\cdot\mathbf{r}}.$$

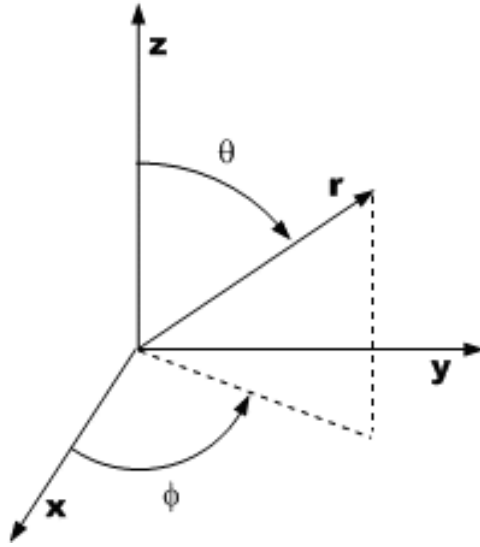


FIG. 1: The spherical coordinates: radius r , polar angle θ and azimuthal angle φ .

which satisfy the normalization relation

$$\iiint_{\{L^3\}} \psi_{\mathbf{p}}^*(\mathbf{r}) \psi_{\mathbf{p}'}(\mathbf{r}) d\mathbf{r}^3 = \delta_{p_x p'_x} \delta_{p_y p'_y} \delta_{p_z p'_z}.$$

$$\frac{1}{L} \int_0^L e^{\frac{i}{\hbar}(m-n)\frac{2\pi\hbar}{L}x} dx = \frac{1}{L} \int_0^L e^{i2\pi(m-n)x/L} dx = \delta_{mn}$$

In the final result of calculation we will let $L \rightarrow +\infty$. If we discuss the normalization using the wave functions in momentum space, the result should be

$$\iiint_{\{P^3\}} \psi_{\mathbf{r}}^*(\mathbf{p}) \psi_{\mathbf{r}'}(\mathbf{p}) d\mathbf{p}^3 = \delta_{xx'} \delta_{yy'} \delta_{zz'}.$$

The integration is limited in a box in momentum space.

II. EIGENFUNCTIONS OF ORBITAL ANGULAR MOMENTUM

In this section we discuss the eigenfunctions and eigenvalues of another often used operator, the orbital angular momentum.

A. Angular momentum operator in spherical coordinates

We already know that the classically defined orbital angular momentum $\mathbf{L} = \mathbf{r} \times \mathbf{p}$ and its components are changed into operators as

$$\begin{aligned} \hat{\mathbf{L}} &= \hat{\mathbf{r}} \times \hat{\mathbf{p}} \\ \hat{L}_x &= \hat{y}\hat{p}_z - \hat{z}\hat{p}_y \\ \hat{L}_y &= \hat{z}\hat{p}_x - \hat{x}\hat{p}_z \\ \hat{L}_z &= \hat{x}\hat{p}_y - \hat{y}\hat{p}_x \end{aligned}$$

We define a new operator

$$\hat{L}^2 = \hat{L}_x^2 + \hat{L}_y^2 + \hat{L}_z^2.$$

The angular momentum is defined in Cartesian coordinates, but is extensively used in spherical coordinates. Here we first transform these angular momentum operators into spherical coordinates (Figure 1)

$$\begin{aligned} x &= r \sin \theta \cos \varphi \\ y &= r \sin \theta \sin \varphi \\ z &= r \cos \theta \end{aligned}$$

We notice that

$$\begin{aligned} \frac{\partial}{\partial \varphi} &= \frac{\partial x}{\partial \varphi} \frac{\partial}{\partial x} + \frac{\partial y}{\partial \varphi} \frac{\partial}{\partial y} + \frac{\partial z}{\partial \varphi} \frac{\partial}{\partial z} \\ &= -r \sin \theta \sin \varphi \frac{\partial}{\partial x} + r \sin \theta \cos \varphi \frac{\partial}{\partial y} \\ &= -y \frac{\partial}{\partial x} + x \frac{\partial}{\partial y} \end{aligned}$$

which gives immediately the z -component of the angular momentum

$$\begin{aligned} -i\hbar \frac{\partial}{\partial \varphi} &= x(-i\hbar \frac{\partial}{\partial y}) - y(-i\hbar \frac{\partial}{\partial x}) \\ &= \hat{x}\hat{p}_y - \hat{y}\hat{p}_x = \hat{L}_z, \end{aligned}$$

i.e.

$$\hat{L}_z = -i\hbar \frac{\partial}{\partial \varphi}. \quad (5)$$

The combination of the derivative with respect to θ

$$\begin{aligned} \frac{\partial}{\partial \theta} &= \frac{\partial x}{\partial \theta} \frac{\partial}{\partial x} + \frac{\partial y}{\partial \theta} \frac{\partial}{\partial y} + \frac{\partial z}{\partial \theta} \frac{\partial}{\partial z} \\ &= r \cos \theta \cos \varphi \frac{\partial}{\partial x} + r \cos \theta \sin \varphi \frac{\partial}{\partial y} - r \sin \theta \frac{\partial}{\partial z} \end{aligned}$$

and that with respect to φ , i.e.

$$\frac{\partial}{\partial \varphi} = -r \sin \theta \sin \varphi \frac{\partial}{\partial x} + r \sin \theta \cos \varphi \frac{\partial}{\partial y}$$

leads to

$$\begin{aligned} \cos \varphi \frac{\partial}{\partial \theta} - \cot \theta \sin \varphi \frac{\partial}{\partial \varphi} &= r \cos \theta \frac{\partial}{\partial x} - r \sin \theta \cos \varphi \frac{\partial}{\partial z} \\ &= z \frac{\partial}{\partial x} - x \frac{\partial}{\partial z}. \end{aligned}$$

Appending the factor $-i\hbar$ we thus have

$$\hat{L}_y = -i\hbar \left(\cos \varphi \frac{\partial}{\partial \theta} - \cot \theta \sin \varphi \frac{\partial}{\partial \varphi} \right). \quad (6)$$

By same token

$$\hat{L}_x = -i\hbar \left(-\sin \varphi \frac{\partial}{\partial \theta} - \cot \theta \cos \varphi \frac{\partial}{\partial \varphi} \right). \quad (7)$$

The operator \hat{L}^2 can be derived from the square of the 3 components

$$\begin{aligned}\hat{L}_x^2 &= -\hbar^2 \left\{ \left(\sin \varphi \frac{\partial}{\partial \theta} + \cot \theta \cos \varphi \frac{\partial}{\partial \varphi} \right) \left(\sin \varphi \frac{\partial}{\partial \theta} + \cot \theta \cos \varphi \frac{\partial}{\partial \varphi} \right) \right\} \\ &= -\hbar^2 \left\{ \sin^2 \varphi \frac{\partial^2}{\partial \theta^2} - \frac{1}{\sin^2 \theta} \sin \varphi \cos \varphi \frac{\partial}{\partial \varphi} + 2 \cot \theta \sin \varphi \cos \varphi \frac{\partial^2}{\partial \theta \partial \varphi} \right. \\ &\quad \left. + \cot \theta \cos^2 \varphi \frac{\partial}{\partial \theta} - \cot^2 \theta \sin \varphi \cos \varphi \frac{\partial}{\partial \varphi} + \cot^2 \theta \cos^2 \varphi \frac{\partial^2}{\partial \varphi^2} \right\}\end{aligned}$$

$$\begin{aligned}\hat{L}_y^2 &= -\hbar^2 \left\{ \left(\cos \varphi \frac{\partial}{\partial \theta} - \cot \theta \sin \varphi \frac{\partial}{\partial \varphi} \right) \left(\cos \varphi \frac{\partial}{\partial \theta} - \cot \theta \sin \varphi \frac{\partial}{\partial \varphi} \right) \right\} \\ &= -\hbar^2 \left\{ \cos^2 \varphi \frac{\partial^2}{\partial \theta^2} + \frac{1}{\sin^2 \theta} \sin \varphi \cos \varphi \frac{\partial}{\partial \varphi} - 2 \cot \theta \sin \varphi \cos \varphi \frac{\partial^2}{\partial \theta \partial \varphi} \right. \\ &\quad \left. + \cot \theta \sin^2 \varphi \frac{\partial}{\partial \theta} + \cot^2 \theta \sin \varphi \cos \varphi \frac{\partial}{\partial \varphi} + \cot^2 \theta \sin^2 \varphi \frac{\partial^2}{\partial \varphi^2} \right\}\end{aligned}$$

$$\hat{L}_z^2 = -\hbar^2 \frac{\partial^2}{\partial \varphi^2}$$

one finally finds

$$\begin{aligned}\hat{L}^2 &= \hat{L}_x^2 + \hat{L}_y^2 + \hat{L}_z^2 \\ &= -\hbar^2 \left(\frac{\partial^2}{\partial \theta^2} + \cot \theta \frac{\partial}{\partial \theta} + \cot^2 \theta \frac{\partial^2}{\partial \varphi^2} + \frac{\partial^2}{\partial \varphi^2} \right)\end{aligned}$$

i.e.

$$\hat{L}^2 = -\hbar^2 \left(\frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \varphi^2} \right) \quad (8)$$

The equations (5,6,7,8) are the spherical coordinate expression for orbital angular momentum.

B. Eigenfunctions of orbital angular momentum

Here we try to find the common eigenfunction of operators \hat{L}^2 and \hat{L}_z . The operators have the differential operator form expressed in spherical coordinates. The eigenfunctions $Y(\theta, \varphi)$ are also expressed in spherical coordinates with their arguments θ and φ changing in the domain

$$\begin{aligned}\theta &: 0 \rightarrow \pi \\ \varphi &: 0 \rightarrow 2\pi\end{aligned}$$

The eigenequations are

$$\begin{aligned}\hat{L}^2 Y(\theta, \varphi) &= \lambda \hbar^2 Y(\theta, \varphi) \\ \hat{L}_z Y(\theta, \varphi) &= m \hbar Y(\theta, \varphi)\end{aligned}$$

where λ, m are some constants to be found. As always, we try again the separation of variables

$$Y(\theta, \varphi) = \Theta(\theta) \Phi(\varphi),$$

then the equation for \hat{L}_z

$$\hat{L}_z \Theta(\theta) \Phi(\varphi) = -i\hbar \frac{\partial}{\partial \varphi} \Theta(\theta) \Phi(\varphi) = m\hbar \Theta(\theta) \Phi(\varphi)$$

gives simply

$$\Phi(\varphi) = \exp(im\varphi). \quad (9)$$

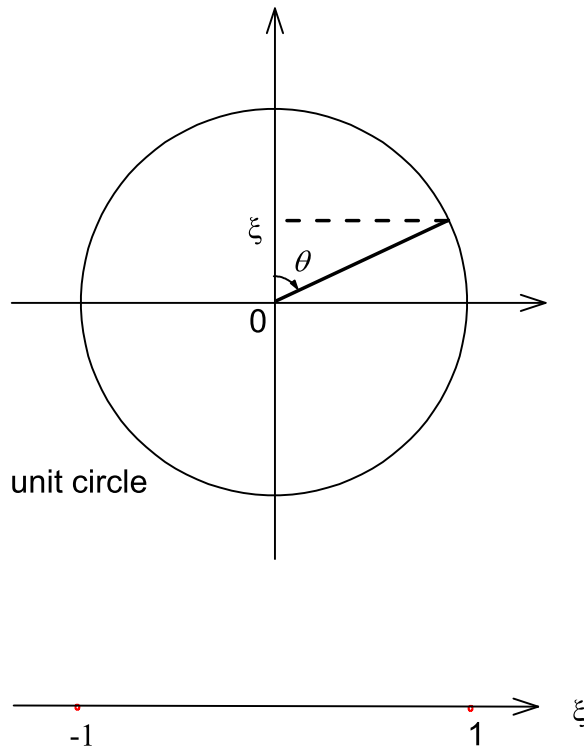


FIG. 2: Change of variables $\xi = \cos \theta$.

There could also be a constant factor in front, but we might as well absorb that into $\Theta(\theta)$. As the azimuthal angle φ advances by 2π , we return to the same point in space (see Figure 1), so it is natural to require that

$$\Phi(\varphi + 2\pi) = \Phi(\varphi)$$

In other words, $\exp(im\varphi + im2\pi) = \exp(im\varphi)$, or $\exp(2\pi im) = 1$. From this it follows that m must be an integer

$$m = 0, \pm 1, \pm 2, \dots$$

For the eigenequation of \hat{L}^2 , we substitute the explicit operator form into the equation

$$-h^2 \left\{ \frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} Y(\theta, \varphi) \right) + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \varphi^2} Y(\theta, \varphi) \right\} = \lambda h^2 Y(\theta, \varphi)$$

The second term in the equation can be replaced by

$$\frac{\partial^2}{\partial \varphi^2} Y(\theta, \varphi) = \frac{\partial^2}{\partial \varphi^2} (\Theta(\theta) \exp(im\varphi)) = -m^2 \Theta(\theta) \Phi(\varphi),$$

which reduces the original equation into

$$- \left\{ \frac{1}{\sin \theta} \frac{d}{d\theta} \left(\sin \theta \frac{d\Theta(\theta)}{d\theta} \right) - \frac{m^2}{\sin^2 \theta} \Theta(\theta) \right\} = \lambda \Theta(\theta)$$

If you are familiar with the special functions, it is easy to recognize the solution of this equation is the **associated Legendre function** $P_l^m(\cos \theta)$. Here we assume not very much knowledge about this special function and start from the beginning. First we change the independent variable from θ to ξ by $\xi = \cos \theta$. There is one-to-one correspondence in this transformation (See Figure 2)

$$\begin{aligned} \theta : 0 &\rightarrow \pi \\ \xi : 1 &\rightarrow -1 \end{aligned}$$

The derivative is

$$\frac{d}{d\theta} = \frac{d\xi}{d\theta} \frac{d}{d\xi} = -\sin\theta \frac{d}{d\xi}$$

Correspondingly, the unknown function $\Theta(\theta)$ changes to $P(\xi)$, and the equation becomes

$$\frac{d}{d\xi} \left[(1 - \xi^2) \frac{dP}{d\xi} \right] + \left(\lambda - \frac{m^2}{1 - \xi^2} \right) P(\xi) = 0$$

or

$$(1 - \xi^2) \frac{d^2P}{d\xi^2} - 2\xi \frac{dP}{d\xi} + \left(\lambda - \frac{m^2}{1 - \xi^2} \right) P = 0$$

$P(\xi)$ is defined in the interval $(-1, 1)$ with the two end points $\xi = \pm 1$ being singular points.

1. Series solution for $m = 0$ case

Let us first check the case of $m = 0$

$$(1 - \xi^2) \frac{d^2P}{d\xi^2} - 2\xi \frac{dP}{d\xi} + \lambda P = 0$$

We seek the series solution of the equation, i.e., presume $P(\xi)$ can be expressed by a series of ξ

$$P(\xi) = \sum_{\nu=0}^{+\infty} c_{\nu} \xi^{\nu}$$

It is easy to find the recursion formula

$$\begin{aligned} (1 - \xi^2) \sum_{\nu=0}^{+\infty} \nu(\nu - 1) c_{\nu} \xi^{\nu-2} - 2\xi \sum_{\nu=0}^{+\infty} \nu c_{\nu} \xi^{\nu-1} + \lambda \sum_{\nu=0}^{+\infty} c_{\nu} \xi^{\nu} &= 0 \\ \sum_{\nu=0}^{+\infty} \nu(\nu - 1) c_{\nu} \xi^{\nu-2} - \sum_{\nu=0}^{+\infty} \nu(\nu - 1) c_{\nu} \xi^{\nu} - 2 \sum_{\nu=0}^{+\infty} \nu c_{\nu} \xi^{\nu} + \lambda \sum_{\nu=0}^{+\infty} c_{\nu} \xi^{\nu} &= 0 \\ (\nu + 2)(\nu + 1) c_{\nu+2} - \nu(\nu - 1) c_{\nu} - 2\nu c_{\nu} + \lambda c_{\nu} &= 0 \\ (\nu + 2)(\nu + 1) c_{\nu+2} &= (\nu(\nu - 1) + 2\nu - \lambda) c_{\nu} \end{aligned}$$

$$c_{\nu+2} = \frac{\nu(\nu + 1) - \lambda}{(\nu + 2)(\nu + 1)} c_{\nu}$$

i.e., knowing the two initial coefficients c_0 and c_1 , we can deduce the rest of them,

$$\begin{aligned} c_0 &\rightarrow c_2 \rightarrow c_4 \rightarrow \dots \\ c_1 &\rightarrow c_3 \rightarrow c_5 \rightarrow \dots \end{aligned}$$

The infinite series $P(\xi)$ diverges at $\xi \rightarrow \pm 1$, so we have to **truncate** it. Let $\lambda = l(l + 1)$, then the coefficient $c_l \neq 0$ while $c_{l+2} = c_{l+4} = \dots = 0$. If l is even, we choose $c_1 = 0$, if l is odd, we choose $c_0 = 0$. This will kill the other series. Thus we have found a polynomial solution of the equation

$$P(\xi) = \sum_{\nu=0}^l c_{\nu} \xi^{\nu}$$

It is a result of truncation of a power series. The reason of truncation is the requirement that $P(\xi)$ should be limited at $\xi = \pm 1$, or, in spherical coordinate, the wave function should be limited at the polar points (Figure 3).

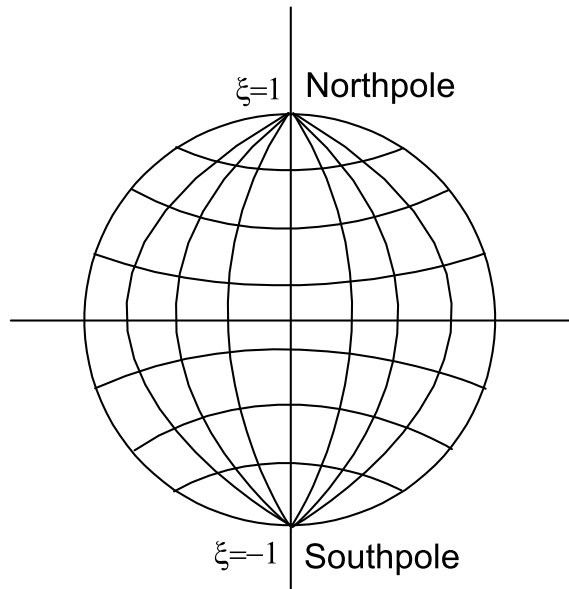


FIG. 3: Spherical coordinate: North pole and south pole.

The polynomial is called **Legendre Polynomial**, the general expression of it is (Rodriguez Formula)

$$P_l(\xi) = \frac{1}{2^l \cdot l!} \frac{d^l}{d\xi^l} (\xi^2 - 1)^l, \quad l = 0, 1, 2, \dots$$

with

$$\begin{aligned} P_0(\xi) &= 1 \\ P_1(\xi) &= \xi \\ P_2(\xi) &= \frac{1}{2}(3\xi^2 - 1) \\ P_3(\xi) &= \frac{1}{2}(5\xi^3 - 3\xi) \end{aligned}$$

The first few Legendre polynomials are illustrated in Figure 4.

2. Series solution for $m \neq 0$ case

The equation for $m \neq 0$ case reads

$$\frac{d}{d\xi} \left((1 - \xi^2) \frac{dP}{d\xi} \right) + \left(l(l+1) - \frac{m^2}{1 - \xi^2} \right) P(\xi) = 0 \quad (10)$$

We try to tackle this equation from two sides.

- Let $P(\xi) = (1 - \xi^2)^{\frac{m}{2}} Q(\xi)$, substitute it into the above equation

$$\frac{d}{d\xi} \left\{ (1 - \xi^2) \frac{d}{d\xi} \left[(1 - \xi^2)^{\frac{m}{2}} Q(\xi) \right] \right\} + \left(l(l+1) - \frac{m^2}{1 - \xi^2} \right) (1 - \xi^2)^{\frac{m}{2}} Q(\xi) = 0$$

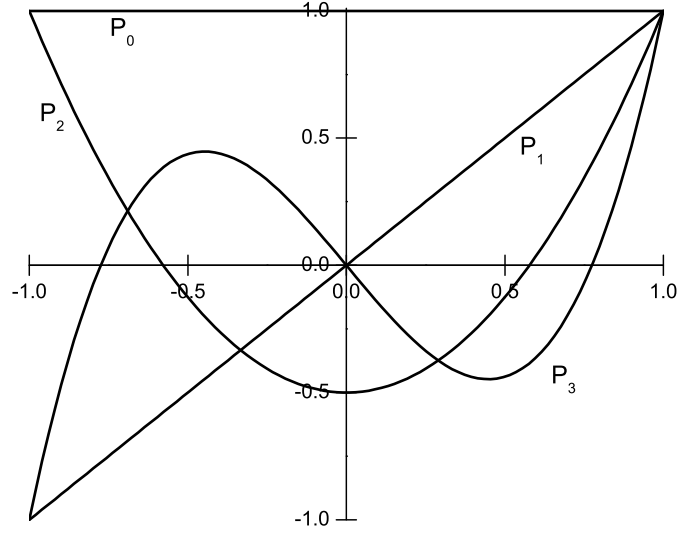


FIG. 4: The first few Legendre polynomials.

A detailed calculation of the first term in the equation gives

$$\begin{aligned}
& \frac{d}{d\xi} \left\{ (1 - \xi^2) \frac{d}{d\xi} \left[(1 - \xi^2)^{\frac{m}{2}} Q(\xi) \right] \right\} \\
&= \frac{d}{d\xi} \left\{ (1 - \xi^2) \left[(1 - \xi^2)^{\frac{m}{2}} \frac{dQ}{d\xi} + \frac{m}{2} (1 - \xi^2)^{\frac{m}{2}-1} (-2\xi) Q \right] \right\} \\
&= \frac{d}{d\xi} \left\{ (1 - \xi^2)^{\frac{m}{2}+1} \frac{dQ}{d\xi} - m\xi (1 - \xi^2)^{\frac{m}{2}} Q \right\} \\
&= (1 - \xi^2)^{\frac{m}{2}+1} \frac{d^2Q}{d\xi^2} + \frac{m+2}{2} (1 - \xi^2)^{\frac{m}{2}} (-2\xi) \frac{dQ}{d\xi} \\
&\quad - m(1 - \xi^2)^{\frac{m}{2}} Q - m\xi \frac{m}{2} (1 - \xi^2)^{\frac{m}{2}-1} (-2\xi) Q - m\xi (1 - \xi^2)^{\frac{m}{2}} \frac{dQ}{d\xi} \\
&= (1 - \xi^2)^{\frac{m}{2}} \left\{ (1 - \xi^2) \frac{d^2Q}{d\xi^2} - (m+2)\xi \frac{dQ}{d\xi} - m\xi \frac{dQ}{d\xi} - mQ + \frac{m^2\xi^2}{(1 - \xi^2)} Q \right\} \\
&= (1 - \xi^2)^{\frac{m}{2}} \left\{ (1 - \xi^2) \frac{d^2Q}{d\xi^2} - 2(m+1)\xi \frac{dQ}{d\xi} - mQ + \frac{m^2\xi^2}{(1 - \xi^2)} Q \right\}.
\end{aligned}$$

In the equation we drop the factor $(1 - \xi^2)^{\frac{m}{2}}$ on both sides

$$(1 - \xi^2) \frac{d^2Q}{d\xi^2} - 2(m+1)\xi \frac{dQ}{d\xi} + l(l+1)Q + \frac{m^2\xi^2 - m^2}{(1 - \xi^2)} Q - mQ = 0$$

and after simplification one arrives at

$$(1 - \xi^2) \frac{d^2Q}{d\xi^2} - 2(m+1)\xi \frac{dQ}{d\xi} + [l(l+1) - m(m+1)]Q = 0 \quad (11)$$

- For the case $m = 0$, the solution has been found to be the Legendre Polynomials $P_l(\xi)$

$$(1 - \xi^2) \frac{d^2}{d\xi^2} P_l(\xi) - 2\xi \frac{d}{d\xi} P_l(\xi) + l(l+1)P_l(\xi) = 0$$

Differentiating this equation with respect to ξ for m times (Leibnitz formula), we get

$$\begin{aligned}
& (1 - \xi^2) \frac{d^2}{d\xi^2} \frac{d^m}{d\xi^m} P_l(\xi) + m \left[\frac{d}{d\xi} (1 - \xi^2) \right] \frac{d^2}{d\xi^2} \frac{d^{m-1}}{d\xi^{m-1}} P_l(\xi) \\
& + \frac{1}{2} m(m-1) \left[\frac{d^2}{d\xi^2} (1 - \xi^2) \right] \frac{d^2}{d\xi^2} \frac{d^{m-2}}{d\xi^{m-2}} P_l(\xi) \\
& - 2\xi \frac{d}{d\xi} \frac{d^m}{d\xi^m} P_l(\xi) + m \left[\frac{d}{d\xi} (-2\xi) \right] \frac{d}{d\xi} \frac{d^{m-1}}{d\xi^{m-1}} P_l(\xi) + l(l+1) \frac{d^m}{d\xi^m} P_l(\xi) \\
& = 0
\end{aligned}$$

After simplification, we arrived at

$$\begin{aligned}
& (1 - \xi^2) \frac{d^2}{d\xi^2} \left[\frac{d^m}{d\xi^m} P_l(\xi) \right] - 2(m+1)\xi \frac{d}{d\xi} \left[\frac{d^m}{d\xi^m} P_l(\xi) \right] \\
& + [l(l+1) - m(m+1)] \left[\frac{d^m}{d\xi^m} P_l(\xi) \right] = 0
\end{aligned} \tag{12}$$

Eq. (11) and (12) are exactly the same, their solutions should be the same too. Thus, the unknown function $Q(\xi)$ are found to be

$$\begin{aligned}
Q(\xi) &= \frac{d^m}{d\xi^m} P_l(\xi) = \frac{d^m}{d\xi^m} \left[\frac{1}{2^l l!} \frac{d^l}{d\xi^l} (\xi^2 - 1)^l \right] \\
&= \frac{1}{2^l l!} \frac{d^{l+m}}{d\xi^{l+m}} [(\xi^2 - 1)^l]
\end{aligned}$$

The solution of Eq. (10) is the **associated Legendre function** (it is not polynomial for odd m)

$$\begin{aligned}
P(\xi) &= (1 - \xi^2)^{\frac{m}{2}} \frac{d^m}{d\xi^m} P_l(\xi) \\
&= \frac{1}{2^l l!} (1 - \xi^2)^{\frac{m}{2}} \frac{d^{l+m}}{d\xi^{l+m}} [(\xi^2 - 1)^l]
\end{aligned}$$

where m can only take values $0 \leq l + m \leq 2l$, i.e. $-l \leq m \leq l$ because the $2l$ times derivative of ξ in the above equation will give zero.

C. Eigenfunction of \hat{L}_z and \hat{L}^2 , the spherical harmonics

We begin with the task of finding wave functions in which the dynamical variables \hat{L}_z and \hat{L}^2 have definite values

$$\begin{aligned}
\hat{L}^2 Y(\theta, \varphi) &= \lambda \hbar^2 Y(\theta, \varphi) \\
\hat{L}_z Y(\theta, \varphi) &= m \hbar Y(\theta, \varphi)
\end{aligned}$$

The steps are

(i) Separation of variables $Y(\theta, \varphi) = \Theta(\theta)\Phi(\varphi)$ where

$$\Phi(\varphi) = \exp(im\varphi), \quad m = \dots - 2, -1, 0, 1, 2 \dots$$

(ii) $\xi = \cos \theta$, $\Theta(\theta) \implies P(\xi)$. For $m = 0$ case, it is found $\lambda = l(l+1)$

$$P(\xi) = P_l(\xi), \quad l = 0, 1, 2 \dots$$

(iii) For $m \neq 0$ case, it is found $\lambda = l(l+1)$ unchanged,

$$P(\xi) = (1 - \xi^2)^{\frac{m}{2}} \frac{d^m}{d\xi^m} P_l(\xi)$$

Summing the above results up, the eigenfunction has the form

$$Y(\theta, \varphi) = \Theta(\theta)\Phi(\varphi) = P(\cos\theta) \exp(im\varphi)$$

$$Y(\theta, \varphi) = (1 - \cos^2\theta)^{\frac{m}{2}} \frac{1}{2^l l!} \frac{d^{l+m}}{d(\cos\theta)^{l+m}} [(\cos^2\theta - 1)^l] \exp(im\varphi)$$

Adding suitable normalization factor, the solution got a new name **spherical harmonics**

$$Y_{lm}(\theta, \varphi) = (-1)^m \sqrt{\frac{(l-m)!}{(l+m)!}} \sqrt{\frac{2l+1}{4\pi}} \frac{1}{2^l l!} \sin^m\theta \frac{d^{l+m}}{d(\cos\theta)^{l+m}} [(\cos^2\theta - 1)^l] \exp(im\varphi)$$

We discuss the eigenvalues and eigenfunctions below.

(i) Eigenvalues

$$\begin{aligned} \hat{L}^2 Y_{lm}(\theta, \varphi) &= l(l+1)\hbar^2 Y_{lm}(\theta, \varphi), & l &= 0, 1, 2, \dots \\ \hat{L}_z Y_{lm}(\theta, \varphi) &= m\hbar Y_{lm}(\theta, \varphi), & m &= -l, -l+1, \dots, 0, 1, \dots, l \end{aligned}$$

Thus the angular momentum is quantized. l is called the *angular quantum number*, while m is called the *magnetic quantum number*. The relation of them can be seen from the definition of the spherical harmonics. They can take the following values

$$\begin{aligned} l &= 0, m = 0 \\ l &= 1, m = -1, 0, 1 \\ l &= 2, m = -2, -1, 0, 1, 2 \end{aligned}$$

(ii) Eigenfunctions. The eigenfunctions of the orbital angular momentum are spherical harmonics $Y_{lm}(\theta, \varphi)$, the first few of which are listed below

$$\begin{aligned} Y_{0,0}(\theta, \varphi) &= \frac{1}{\sqrt{4\pi}} \\ Y_{1,1}(\theta, \varphi) &= -\sqrt{\frac{3}{8\pi}} \sin\theta \exp(i\phi) \\ Y_{1,-1}(\theta, \varphi) &= \sqrt{\frac{3}{8\pi}} \sin\theta \exp(-i\phi) \\ Y_{1,0}(\theta, \varphi) &= \sqrt{\frac{3}{4\pi}} \cos\theta \\ Y_{2,2}(\theta, \varphi) &= \sqrt{\frac{15}{32\pi}} \sin^2\theta \exp(2i\phi) \\ Y_{2,1}(\theta, \varphi) &= -\sqrt{\frac{15}{8\pi}} \sin\theta \cos\theta \exp(i\phi) \\ Y_{2,0}(\theta, \varphi) &= \sqrt{\frac{5}{16\pi}} (3\cos^2\theta - 1) \\ Y_{2,-1}(\theta, \varphi) &= \sqrt{\frac{15}{8\pi}} \sin\theta \cos\theta \exp(-i\phi) \\ Y_{2,-2}(\theta, \varphi) &= \sqrt{\frac{15}{32\pi}} \sin^2\theta \exp(-2i\phi) \end{aligned}$$

The spherical harmonics has already been normalized

$$\iint Y_{lm}^*(\theta, \varphi) Y_{lm}(\theta, \varphi) d\Omega = \iint Y_{lm}^*(\theta, \varphi) Y_{lm}(\theta, \varphi) \sin\theta d\theta d\varphi = 1$$

The spherical harmonics determines the angular distribution of the particle. For a function $Y(\theta, \varphi)$, the probability of finding the particle in certain direction is $|Y(\theta, \varphi)|^2 d\Omega$.

III. CENTRAL FORCE FIELD

A particle moves under the influence of a central-force field $F(\mathbf{r}) = F(r)\hat{\mathbf{r}}$ if the force on the particle is independent of the *angular position* of the particle about the center of force and depends only on its *distance* r from the center of force. Here, the magnitude $F(r)$ (which is positive for a repulsive force and negative for an attractive force) is defined in terms of the central potential $V(r)$ as $F(r) = -V'(r)$.

A. General properties of central force field problem

The Schrödinger equation $\hat{H}\psi = E\psi$ is now

$$-\frac{\hbar^2}{2m}\nabla^2\psi + V(r)\psi = E\psi \quad (13)$$

where the potential energy V is a function of r only $V = V(r)$ and m is the mass of the particle. This equation will be solved in spherical coordinate.

1. Separation of variables in central force field problem

Starting from $\nabla = \hat{r}\frac{\partial}{\partial r} + \hat{\theta}\frac{1}{r}\frac{\partial}{\partial\theta} + \hat{\varphi}\frac{1}{r\sin\theta}\frac{\partial}{\partial\varphi}$, you will find it is not quite easy to show that the Laplacian or Laplace operator in spherical coordinate reads as

$$\nabla^2 = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + \frac{1}{r^2}\left\{\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\varphi^2}\right\}.$$

The equation (13) can be written as

$$-\frac{\hbar^2}{2mr^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\psi\right) + \frac{\hat{L}^2}{2mr^2}\psi + V(r)\psi = E\psi$$

by means of the definition of \hat{L}^2 , eq. (8).

We begin by looking for solutions that are separable into products $\psi(r, \theta, \varphi) = R(r)Z(\theta, \varphi)$. Inserting this into the above equation leads us to

$$\begin{aligned} -Z\frac{\hbar^2}{2mr^2}\frac{d}{dr}\left(r^2\frac{d}{dr}R\right) + R\frac{\hat{L}^2}{2mr^2}Z &= (E - V)RZ \\ Z\hbar^2\frac{d}{dr}\left(r^2\frac{d}{dr}R\right) + 2mr^2(E - V)RZ &= R\hat{L}^2Z \\ \frac{\hbar^2\frac{d}{dr}\left(r^2\frac{d}{dr}R\right) + 2mr^2(E - V)R}{R} &= \frac{\hat{L}^2Z}{Z} \end{aligned}$$

Clearly it should be a constant because the *lhs* depends on r only and the *rhs* depends on θ, φ . This constant should be $l(l+1)\hbar^2$ and Z is nothing but the spherical harmonics $Y_{lm}(\theta, \varphi)$ obtained in the above section, i.e. $\hat{L}^2Y_{lm}(\theta, \varphi) = l(l+1)\hbar^2Y_{lm}(\theta, \varphi)$. For the radial part we have

$$\begin{aligned} \hbar^2\frac{d}{dr}\left(r^2\frac{d}{dr}R\right) + 2mr^2(E - V)R &= l(l+1)\hbar^2R(r) \\ -\frac{\hbar^2}{2mr^2}\frac{d}{dr}\left(r^2\frac{d}{dr}R\right) + \frac{l(l+1)\hbar^2}{2mr^2}R(r) &= (E - V)R \end{aligned}$$

Thus we have successfully separate the variables as $\psi(r, \theta, \varphi) = R(r)Y_{lm}(\theta, \varphi)$ and are left with *the radial equation*

$$-\frac{\hbar^2}{2mr^2}\frac{d}{dr}\left(r^2\frac{dR}{dr}\right) + \left(\frac{l(l+1)\hbar^2}{2mr^2} + V(r)\right)R(r) = ER(r)$$

with the independent variable r changes from 0 to infinity.

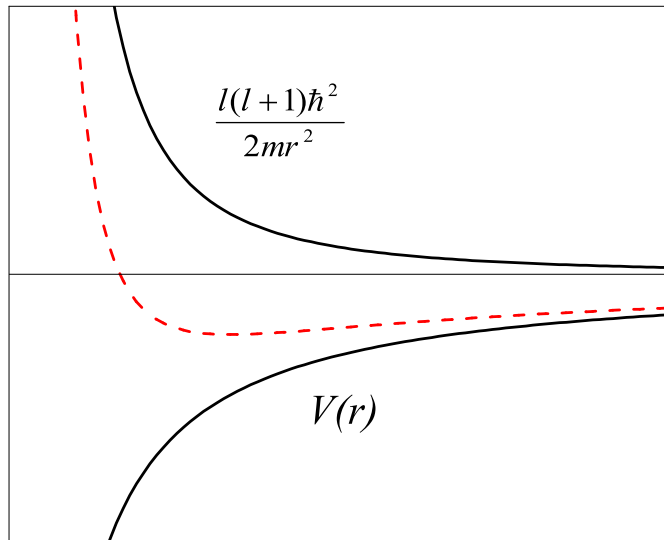


FIG. 5: The effective potential and the centrifugal potential.

2. Effective potential – the centrifugal potential

The radial equation is identical in form to the one dimensional Schrödinger equation, except that the effective potential

$$\frac{l(l+1)\hbar^2}{2mr^2} + V(r)$$

contains an extra piece, the so-called *centrifugal potential* $l(l+1)\hbar^2/2mr^2$. It tends to throw the particle outward (away from the origin), and the same term also appears in classical mechanics (see Figure 5).

Meanwhile it is proportional to the square of angular momentum $l(l+1)\hbar^2$ so that for $l = 0$, there is no centrifugal potential. It only exists for $l \neq 0$. Particle moving with different angular momentum experiences different centrifugal potential and has *different* radial equation. Thus it can be anticipated that particles with different angular momenta have different solutions.

The solution of Schrödinger equation has definite angular momentum

$$\begin{aligned}\psi &= R(r) Y_{lm}(\theta, \varphi) \\ \hat{L}^2 \psi &= l(l+1)\hbar^2 \psi \\ \hat{L}_z \psi &= m\hbar \psi\end{aligned}$$

We solve $R(r)$ in the next subsection.

3. Change of variables

Let $R(r) = u(r)/r$, so that

$$\begin{aligned}\frac{d}{dr} \left[r^2 \frac{d}{dr} \left(\frac{u(r)}{r} \right) \right] &= \frac{d}{dr} \left[r^2 \frac{u'(r)}{r} - r^2 \frac{u(r)}{r^2} \right] \\ &= \frac{d}{dr} [ru'(r) - u(r)] = ru''(r)\end{aligned}$$

and hence

$$\begin{aligned}-\frac{\hbar^2}{2mr^2} ru''(r) + \left(\frac{l(l+1)\hbar^2}{2mr^2} + V(r) \right) \frac{u(r)}{r} &= E \frac{u(r)}{r} \\ -\frac{\hbar^2}{2m} \frac{d^2}{dr^2} u(r) + \left(\frac{l(l+1)\hbar^2}{2mr^2} + V(r) \right) u(r) &= Eu(r)\end{aligned}$$

This is also called *the radial equation*.

The resemblance of this equation with 1D Schrödinger equation is more evident. However, care should be taken on the difference between them.

- For the radial wave equation the interval for the independent variable r is $[0, +\infty]$
- The radial wave function has the relation $R(r) = u(r)/r$, when $r \rightarrow 0$, $u(r)$ must tend to zero more quickly than r .

B. Reduced mass

For a single particle, the wave function $\psi(\mathbf{r}, t)$ is a function of the spatial coordinates \mathbf{r} and the time t . The wave function for a two-particle system is a function of the coordinates of particle one (\mathbf{r}_1), the coordinates of particle two (\mathbf{r}_2), and the time: $\psi(\mathbf{r}_1, \mathbf{r}_2, t)$. Hydrogen atom is practically a two-particle problem, just as in classical mechanics. The two-particle problem can be transformed into a single particle problem by introducing the reduced mass.

The spatial wave function $\psi(\mathbf{r}_1, \mathbf{r}_2)$ of hydrogen atom satisfies the time-independent Schrödinger equation

$$\left(-\frac{\hbar^2}{2m_1} \nabla_1^2 - \frac{\hbar^2}{2m_2} \nabla_2^2\right) \psi(\mathbf{r}_1, \mathbf{r}_2) + V(\mathbf{r}_1, \mathbf{r}_2) \psi(\mathbf{r}_1, \mathbf{r}_2) = E\psi(\mathbf{r}_1, \mathbf{r}_2)$$

where m_1 and m_2 are the masses of the electron and the nucleon, respectively. We introduce *the centre of mass coordinate*

$$\mathbf{R} = \frac{1}{m_1 + m_2} (m_1 \mathbf{r}_1 + m_2 \mathbf{r}_2)$$

and *the relative coordinate* $\mathbf{r} = \mathbf{r}_2 - \mathbf{r}_1$. It is easy to check that

$$\begin{aligned} \mathbf{r}_1 &= \mathbf{R} - \frac{m_2}{m_1 + m_2} \cdot \mathbf{r}, & x_1 &= X - \frac{m_2}{m_1 + m_2} x, & \text{etc.} \\ \mathbf{r}_2 &= \mathbf{R} + \frac{m_1}{m_1 + m_2} \cdot \mathbf{r}, & x_2 &= X + \frac{m_1}{m_1 + m_2} x, & \text{etc.} \end{aligned}$$

In terms of the new coordinates the derivatives are

$$\begin{aligned} \frac{\partial}{\partial x_1} &= \frac{\partial X}{\partial x_1} \frac{\partial}{\partial X} + \frac{\partial x}{\partial x_1} \frac{\partial}{\partial x} = \frac{m_1}{m_1 + m_2} \frac{\partial}{\partial X} - \frac{\partial}{\partial x} \\ \frac{\partial}{\partial x_2} &= \frac{\partial X}{\partial x_2} \frac{\partial}{\partial X} + \frac{\partial x}{\partial x_2} \frac{\partial}{\partial x} = \frac{m_2}{m_1 + m_2} \frac{\partial}{\partial X} + \frac{\partial}{\partial x} \end{aligned}$$

$$\begin{aligned} \left\{ \frac{\partial^2}{\partial x_1^2} = \left(\frac{m_1}{m_1 + m_2} \right)^2 \frac{\partial^2}{\partial X^2} - 2 \frac{m_1}{m_1 + m_2} \frac{\partial^2}{\partial X \partial x} + \frac{\partial^2}{\partial x^2} \right\} &\left(-\frac{\hbar^2}{2m_1} \right) \\ \left\{ \frac{\partial^2}{\partial x_2^2} = \left(\frac{m_2}{m_1 + m_2} \right)^2 \frac{\partial^2}{\partial X^2} + 2 \frac{m_2}{m_1 + m_2} \frac{\partial^2}{\partial X \partial x} + \frac{\partial^2}{\partial x^2} \right\} &\left(-\frac{\hbar^2}{2m_2} \right) \end{aligned}$$

The kinetic terms are therefore expressed as

$$\begin{aligned} &-\frac{\hbar^2}{2m_1} \frac{\partial^2}{\partial x_1^2} - \frac{\hbar^2}{2m_2} \frac{\partial^2}{\partial x_2^2} \\ &= -\frac{m_1 + m_2}{2(m_1 + m_2)^2} \hbar^2 \frac{\partial^2}{\partial X^2} + 0 + \left(-\frac{\hbar^2}{2} \right) \left(\frac{1}{m_1} + \frac{1}{m_2} \right) \frac{\partial^2}{\partial x^2} \\ &= -\frac{\hbar^2}{2M} \frac{\partial^2}{\partial X^2} - \frac{\hbar^2}{2} \frac{m_1 + m_2}{m_1 m_2} \frac{\partial^2}{\partial x^2} \\ &= -\frac{\hbar^2}{2M} \frac{\partial^2}{\partial X^2} - \frac{\hbar^2}{2\mu} \frac{\partial^2}{\partial x^2} \end{aligned}$$

The same thing happens for the other two components, in summary we have

$$-\frac{\hbar^2}{2m_1} \nabla_1^2 - \frac{\hbar^2}{2m_2} \nabla_2^2 = -\frac{\hbar^2}{2M} \nabla_R^2 - \frac{\hbar^2}{2\mu} \nabla_r^2$$

where we have defined the *reduced mass* and the *total mass* as

$$\mu = \frac{m_1 m_2}{m_1 + m_2}, \quad M = m_1 + m_2$$

Typically, the interaction potential depends only on the vector $\mathbf{r} = \mathbf{r}_2 - \mathbf{r}_1$ between the two particles. In that case the Schrödinger equation in the new coordinate system

$$\left(-\frac{\hbar^2}{2M} \nabla_R^2 - \frac{\hbar^2}{2\mu} \nabla_r^2 \right) \psi(\mathbf{R}, \mathbf{r}) + V(\mathbf{r})\psi(\mathbf{R}, \mathbf{r}) = E\psi(\mathbf{R}, \mathbf{r})$$

can be solved by the method of separation of variables, i.e. we suppose

$$\psi(\mathbf{R}, \mathbf{r}) = \psi_R(\mathbf{R})\psi_r(\mathbf{r})$$

Inserting this into the equation we have

$$\left[-\frac{\hbar^2}{2M} \nabla_R^2 \psi_R(\mathbf{R}) \right] \psi_r(\mathbf{r}) + \left[-\frac{\hbar^2}{2\mu} \nabla_r^2 \psi_r(\mathbf{r}) + V(\mathbf{r})\psi_r(\mathbf{r}) \right] \psi_R(\mathbf{R}) = E\psi_r(\mathbf{r})\psi_R(\mathbf{R})$$

$$\frac{-\frac{\hbar^2}{2M} \nabla_R^2 \psi_R(\mathbf{R})}{\psi_R(\mathbf{R})} = E - \frac{-\frac{\hbar^2}{2\mu} \nabla_r^2 \psi_r(\mathbf{r}) + V(\mathbf{r})\psi_r(\mathbf{r})}{\psi_r(\mathbf{r})} = E_R$$

Obviously $\psi_R(\mathbf{R})$ satisfies the one particle Schrödinger equation

$$-\frac{\hbar^2}{2M} \nabla_R^2 \psi_R(\mathbf{R}) = E_R \psi_R(\mathbf{R})$$

with the total mass $M = m_1 + m_2$ in place of m , potential zero, and energy E_R , while $\psi_r(\mathbf{r})$ satisfies the one particle Schrödinger equation

$$-\frac{\hbar^2}{2\mu} \nabla_r^2 \psi_r(\mathbf{r}) + V(\mathbf{r})\psi_r(\mathbf{r}) = E_r \psi_r(\mathbf{r})$$

with the reduced mass in place of m , potential $V(\mathbf{r})$, and energy E_r . The total energy is the sum

$$E = E_R + E_r$$

What this tells us is that the center of mass moves like a free particle, and the relative motion (that is, the motion of particle 2 with respect to particle 1) is the same as if we had a *single* particle with the *reduced mass*, subject to the potential V . Exactly the same separation occurs in classical mechanics; it reduces the two-body problem to an equivalent one-body problem.

C. Spherical square potential

1. Infinite spherical well

First we consider the infinite spherical well

$$V(r) = \begin{cases} 0, & r < a \\ +\infty, & r > a \end{cases}$$

The wave function is separable

$$\psi = R(r) Y_{lm}(\theta, \varphi) = \frac{u(r)}{r} Y_{lm}(\theta, \varphi).$$

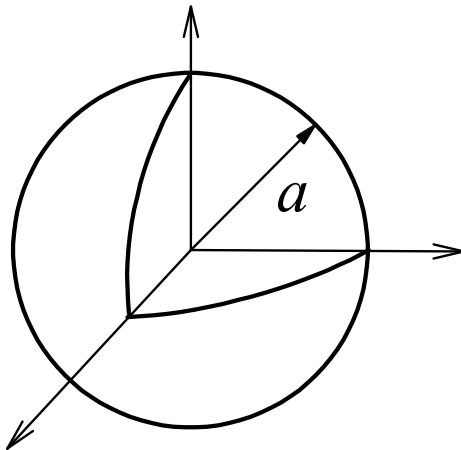


FIG. 6: The infinite spherical well.

Outside the well the wave function is zero; inside the well the radial equation says

$$-\frac{\hbar^2}{2m} \frac{d^2}{dr^2} u(r) + \left(\frac{l(l+1)\hbar^2}{2mr^2} + V(\mathbf{r}) \right) u(r) = Eu(r)$$

where $V(\mathbf{r})$ equals 0 inside the sphere, and only the centrifugal potential $l(l+1)\hbar^2/2mr^2$ survives here. Thus

$$\frac{d^2 u}{dr^2} = \left[\frac{l(l+1)}{r^2} - k^2 \right] u \quad (14)$$

where

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

as usual. Our problem is to solve this equation, subject to the boundary condition $u(a) = 0$. The case $l = 0$ (s -state) is easy

$$\frac{d^2 u}{dr^2} = -k^2 u \Rightarrow u(r) = A \sin(kr) + B \cos(kr)$$

But remember, the actual radial wave function is $R(r) = u(r)/r$, and $\cos(kr)/r$ blows up as $r \rightarrow 0$. So we must choose $B = 0$. The boundary condition then requires $\sin(ka) = 0$, and hence $ka = n\pi$, for some integer n . The allowed energies are evidently

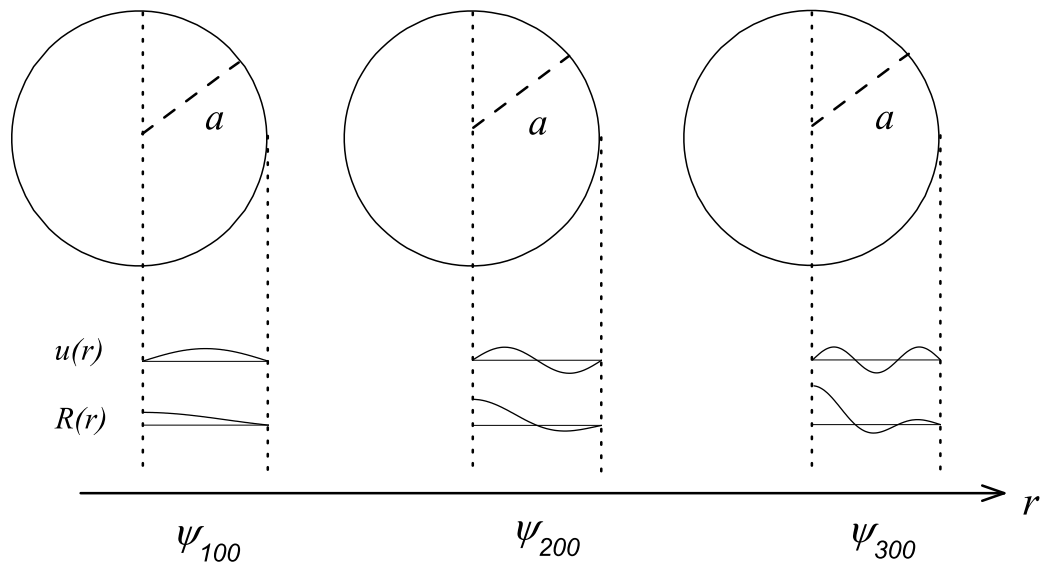
$$E_{n0} = \frac{n^2 \pi^2 \hbar^2}{2ma^2}, n = 1, 2, 3 \dots$$

the same as for the one-dimensional infinite square well. Inclusion of the angular part (constant, in this instance, since $Y_{00} = 1/\sqrt{4\pi}$), we conclude that

$$\psi = A \frac{\sin(kr)}{r} Y_{00}$$

The normalization gives

$$\iiint \psi^* \psi r^2 dr \sin \theta d\theta d\varphi = 1$$

FIG. 7: The first three ψ_{n00} .

$$\begin{aligned} & \iiint A^2 \left(\frac{\sin(kr)}{r} \right)^2 Y_{00}^* Y_{00} r^2 dr \sin \theta d\theta d\varphi \\ &= \int_0^a A^2 \sin^2(kr) dr = A^2 \frac{a}{2} = 1 \end{aligned}$$

so $A = \sqrt{2/a}$ and

$$\psi_{n00} = \frac{1}{\sqrt{2\pi a}} \frac{\sin(n\pi r/a)}{r}$$

Notice that the stationary states are labeled by three quantum numbers, n, l , and m : $\psi_{nlm}(r, \theta, \phi)$. The energy, however, depends only on n and l : E_{n0} . We show the first three ψ_{n00} in Figure 7.

The general solution to equation (14) (for an arbitrary integer l) is not so familiar

$$u(r) = Arj_l(kr) + Brn_l(kr),$$

where $j_l(x)$ is the *spherical Bessel function* of order l , and $n_l(x)$ is the *spherical Neumann function* of order l . They are defined as follows

$$\begin{aligned} j_l(x) &= (-x)^l \left(\frac{1}{x} \frac{d}{dx} \right)^l \frac{\sin x}{x} \\ n_l(x) &= -(-x)^l \left(\frac{1}{x} \frac{d}{dx} \right)^l \frac{\cos x}{x} \end{aligned}$$

For example,

$$\begin{aligned} j_0(x) &= \frac{\sin x}{x}; n_0(x) = -\frac{\cos x}{x} \\ j_1(x) &= (-x) \frac{1}{x} \frac{d}{dx} \left(\frac{\sin x}{x} \right) = \frac{\sin x}{x^2} - \frac{\cos x}{x} \\ n_1(x) &= -(-x) \frac{1}{x} \frac{d}{dx} \left(\frac{\cos x}{x} \right) = -\frac{\cos x}{x^2} - \frac{\sin x}{x} \end{aligned}$$

and so on. Notice that for small x (where $\sin x = x - x^3/3! + x^5/5! \dots$ and $\cos x = 1 - x^2/2! + x^4/4! - \dots$)

$$j_0(x) \approx 1, n_0(x) \approx -\frac{1}{x}$$

$$j_1(x) \approx \frac{x}{3}, n_1(x) \approx -\frac{1}{x^2}$$

etc. The point is that the Bessel functions are finite at the origin, but the Neumann functions blow up at the origin. Accordingly, we must have $B = 0$, and hence

$$R(r) = A j_l(kr)$$

There remains the boundary condition, $R(a) = 0$. Evidently k must be chosen such that

$$j_l(ka) = 0$$

that is, ka is a zero of the l^{th} order spherical Bessel function. Now the Bessel functions are oscillatory; each one has an infinite number of zeros. But (unfortunately, for us) they are not located at nice sensible points (such as n , or $n\pi$, or something); they have to be computed numerically. At any rate, the boundary condition requires that

$$k = \frac{1}{a} \beta_{nl}$$

where β_{nl} is the n^{th} zero of the l^{th} spherical Bessel function. The allowed energies, then, are given by

$$E_{nl} = \frac{\hbar^2}{2ma^2} \beta_{nl}^2$$

and the wave functions are

$$\psi_{nlm}(r, \theta, \phi) = A_{nl} j_l(\beta_{nl} r/a) Y_{lm}(\theta, \phi)$$

with the constant A_{nl} to be determined by normalization. Each energy level is $(2l + 1)$ -fold degenerate, since there are $(2l + 1)$ different values of m for each value of l .

2. Finite spherical well

Consider now a particle of mass m is placed in a finite spherical well

$$V(r) = \begin{cases} -V_0, & r < a \\ 0, & r > a \end{cases}$$

We try to solve the radial equation with $l = 0$ and find the ground state with energy $-V_0 < E < 0$. The wave function is again decomposed into

$$\psi = R(r) Y_{lm} = \frac{u(r)}{r} Y_{00}$$

where the radial part satisfies

$$-\frac{\hbar^2}{2m} \frac{d^2}{dr^2} u(r) + V(r)u(r) = Eu(r)$$

In different regions we have

$$\frac{d^2 u}{dr^2} + k^2 u = 0, \quad r < a$$

$$\frac{d^2 u}{dr^2} - \beta^2 u = 0, \quad r > a$$

where we have defined

$$k = \frac{\sqrt{2m(V_0 + E)}}{\hbar}, \quad \beta = \frac{\sqrt{2m(-E)}}{\hbar}.$$

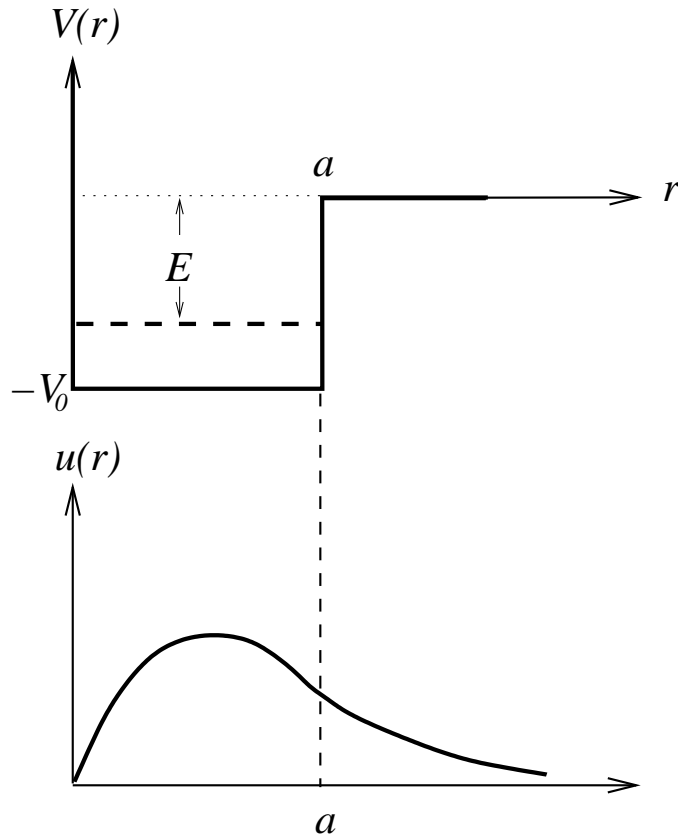


FIG. 8: Finite spherical well and the continuity of $u(r)$.

For $r < a$, $u(r) = A \sin(kr)$ (no \cos as in the infinite case), while for $r > a$, $u(r) = D e^{-\beta r}$ (no $e^{\beta r}$ term which blows up as $r \rightarrow \infty$). The boundary condition at $r = a$ gives

$$\begin{aligned} A \sin ka &= D \exp(-\beta a) \\ kA \cos ka &= -\beta D \exp(-\beta a) \end{aligned}$$

which is

$$ka \cot ka = -\beta a.$$

Let $\xi = ka$, $\eta = \beta a$, so

$$-\xi \cot \xi = \eta,$$

together with the restriction

$$\xi^2 + \eta^2 = \frac{2mV_0}{\hbar^2} a^2 = r_0^2$$

which are exactly the same transcendental equation we encountered in 1D finite potential well. There is no solution if $r_0 < \pi/2$, which is to say, if $2mV_0 a^2 / \hbar^2 < \pi^2 / 4$, or $V_0 a^2 < \pi^2 \hbar^2 / 8m$. Otherwise, the ground state energy occurs somewhere between $\xi = \pi/2$ and π :

$$E + V_0 = \frac{\hbar^2 k^2 a^2}{2ma^2} = \frac{\hbar^2}{2ma^2} \xi^2$$

so

$$\frac{\hbar^2 \pi^2}{8ma^2} < (E_0 + V_0) < \frac{\hbar^2 \pi^2}{2ma^2}$$

Remark 4 *Advanced study material: 3D harmonic oscillator, Zeng Jinyan, Textbook, page 184, section 6.4 or problem 4.38 and 4.39 in Griffiths' book.*

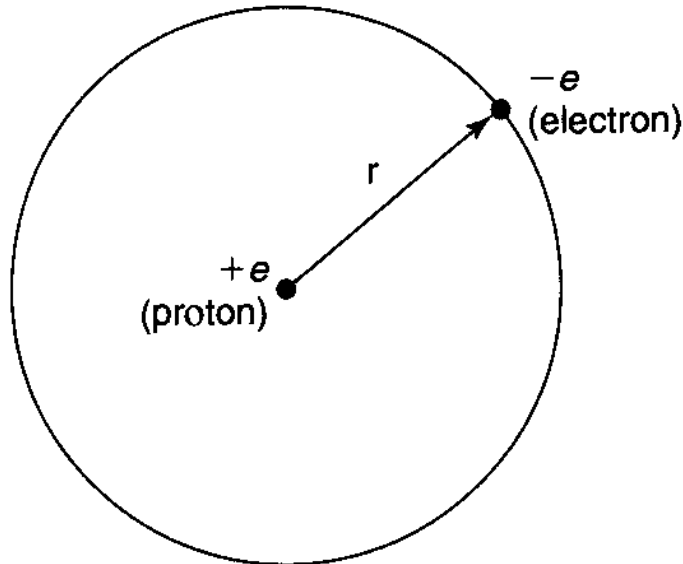


FIG. 9: The hydrogen atom.

IV. THE HYDROGEN ATOM

The hydrogen atom consists of a heavy, essentially motionless proton (we may as well put it at the origin) of charge e , together with a much lighter electron (charge $-e$) that circles around it, held in orbit by the mutual attraction of opposite charges (see Figure 9).

A. Radial equation and its solution

The hydrogen atom, or a *hydrogenic atom* or *hydrogen-like ion* with nucleus with charge Ze , is a typical central force field problem. From Coulomb's law, the potential energy (in SI units) is

$$V(r) = -\frac{1}{4\pi\epsilon_0} \frac{Ze^2}{r} \quad (15)$$

The factor $1/4\pi\epsilon_0$ will disappear in Gaussian unit system (For simplicity we use Gaussian unit in this section). The radial equation is thus

$$-\frac{\hbar^2}{2m} \frac{d^2u}{dr^2} + \left(-\frac{Ze^2}{r} + \frac{l(l+1)\hbar^2}{2mr^2} \right) u = Eu$$

Our problem is to solve this equation for $u(r)$ and determine the allowed electron energies E . The hydrogen atom is such an important case that we'll work them out in detail by the method we used in the analytical solution to the harmonic oscillator. Incidentally, the Coulomb potential (Equation (15)) admits continuum states (with $E > 0$), describing electron-proton scattering, as well as discrete bound states (with $E < 0$), representing the hydrogen atom, but we shall confine our attention to the latter (Figure 10).

Our first task is tidy up the notation. Let us make the equation dimensionless

$$\left(\frac{1}{4(-E)} \right) \left\{ \frac{\hbar^2}{2m} \frac{d^2u}{dr^2} + \left\{ E + \frac{Ze^2}{r} - \frac{\hbar^2}{2m} \frac{l(l+1)}{r^2} \right\} u = 0 \right\}$$

$$\frac{\hbar^2}{8m(-E)} \frac{d^2}{dr^2} u(r) + \left\{ -\frac{1}{4} + \frac{1}{4(-E)} \frac{Ze^2}{r} - \frac{\hbar^2}{8m(-E)} \frac{l(l+1)}{r^2} \right\} u(r) = 0$$

This suggests us to introduce

$$\alpha = \sqrt{\frac{8m(-E)}{\hbar^2}}, \quad \alpha r = \rho$$

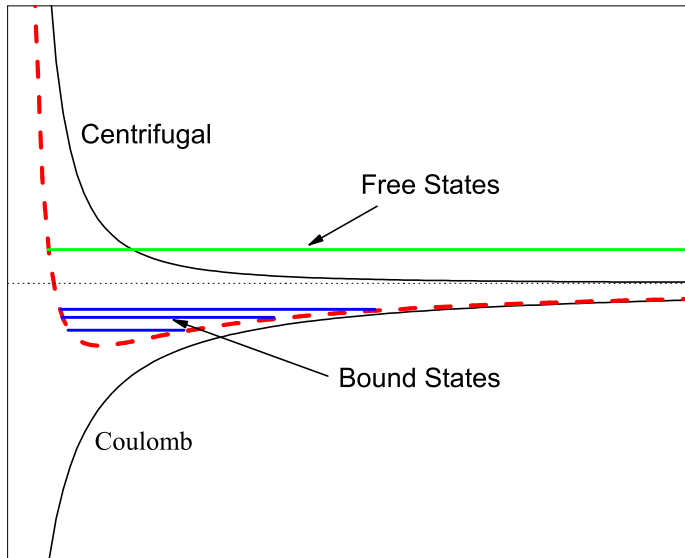


FIG. 10: The bound states and free states in Coulomb potential.

(α is real for bound states $E < 0$) so that we get the dimensionless radial equation of hydrogen atom

$$\frac{d^2}{d\rho^2}u(\rho) + \left\{ -\frac{1}{4} + \frac{\beta}{\rho} - \frac{l(l+1)}{\rho^2} \right\} u(\rho) = 0 \quad (16)$$

where

$$\beta = \frac{Ze^2}{4(-E)} \left(\frac{8m(-E)}{\hbar^2} \right)^{1/2} = \frac{Ze^2}{\hbar} \left(\frac{m}{2(-E)} \right)^{1/2} \quad (17)$$

1. Asymptotic behavior of $u(\rho)$ at singular points

- Next we examine the asymptotic form of the solutions. As $\rho \rightarrow +\infty$, the constant term in the brackets dominates, so (approximately)

$$\frac{d^2}{d\rho^2}u(\rho) = \frac{1}{4}u(\rho)$$

The general solution is

$$u(\rho) = Ae^{-\rho/2} + Be^{\rho/2}$$

but $e^{\rho/2}$ blows up as $\rho \rightarrow +\infty$, so $B = 0$. Evidently

$$u(\rho) \sim Ae^{-\rho/2}$$

for large ρ .

- On the other hand, as $\rho \rightarrow 0$ the centrifugal term dominates; approximately, then

$$\frac{d^2u}{d\rho^2} = \frac{l(l+1)}{\rho^2}u$$

The general solution (check it!) is

$$u(\rho) = C\rho^{l+1} + D\rho^{-l}$$

but ρ^{-l} blows up as $\rho \rightarrow 0$, so $D = 0$. Thus

$$u(\rho) \sim C\rho^{l+1}$$

for small ρ .

The next step is to peel off the asymptotic behavior, introducing the the new function $f(\rho)$

$$u(\rho) = \rho^{l+1} e^{-\rho/2} f(\rho)$$

in the hope that $f(\rho)$ will turn out to be simpler than $u(\rho)$. We assume the solution, $f(\rho)$, can be expressed as a power series in ρ

$$f(\rho) = \sum_{\gamma=0}^{\infty} b_{\gamma} \rho^{\gamma}$$

and

$$u(\rho) = e^{-\rho/2} \sum_{\gamma=0}^{\infty} b_{\gamma} \rho^{\gamma+l+1}$$

2. Solution of the radial equation

We now solve the radial equation for hydrogen atom. Differentiating term by term

$$\begin{aligned} \frac{d^2 u}{d\rho^2} &= \frac{d^2}{d\rho^2} \left\{ e^{-\rho/2} \sum_{\gamma=0}^{\infty} b_{\gamma} \rho^{\gamma+l+1} \right\} \\ &= \left(\frac{d^2}{d\rho^2} e^{-\rho/2} \right) \sum_{\gamma=0}^{\infty} b_{\gamma} \rho^{\gamma+l+1} + 2 \left(\frac{d}{d\rho} e^{-\rho/2} \right) \left(\frac{d}{d\rho} \sum_{\gamma=0}^{\infty} b_{\gamma} \rho^{\gamma+l+1} \right) + e^{-\rho/2} \frac{d^2}{d\rho^2} \sum_{\gamma=0}^{\infty} b_{\gamma} \rho^{\gamma+l+1} \end{aligned}$$

The first item gives $u(\rho)/4$, which cancels the $-u(\rho)/4$ in equation (16). We are left with

$$\begin{aligned} &- e^{-\rho/2} \sum_{\gamma=0}^{\infty} (\gamma + l + 1) b_{\gamma} \rho^{\gamma+l} + e^{-\rho/2} \sum_{\gamma=0}^{\infty} (\gamma + l + 1) (\gamma + l) b_{\gamma} \rho^{\gamma+l-1} \\ &+ \beta e^{-\rho/2} \sum_{\gamma=0}^{\infty} b_{\gamma} \rho^{\gamma+l} - l(l+1) e^{-\rho/2} \sum_{\gamma=0}^{\infty} b_{\gamma} \rho^{\gamma+l-1} = 0 \end{aligned}$$

Equating the coefficients of $(\gamma + l)$ powers yields

$$b_{\gamma} (\beta - \gamma - l - 1) + b_{\gamma+1} [(\gamma + l + 2) (\gamma + l + 1) - l(l + 1)] = 0$$

or

$$b_{\gamma+1} = \frac{(\gamma + l + 1 - \beta)}{(\gamma + l + 2) (\gamma + l + 1) - l(l + 1)} b_{\gamma} \quad (18)$$

This recursion formula determines the coefficients, and hence the function $f(\rho)$: We start with $b_0 = A$ (this becomes an overall constant, to be fixed eventually by normalization), and Equation (18) gives us b_1 ; putting this back in, we obtain b_2 , and so on.

Now let's see what the coefficients look like for large γ (this corresponds to large ρ , where the higher powers dominate). In this regime the recursion formula says

$$\frac{b_{\gamma+1}}{b_\gamma} \rightarrow \frac{1}{\gamma}$$

This means that the infinite series will converge to e^ρ that blows up for large ρ and is not permitted by quantum mechanics.

There is only one way out of this dilemma: *The series must truncate*. There must occur some maximal integer, n_γ , such that

$$b_{n_\gamma+1} = 0$$

(and beyond which all coefficients vanish automatically).

$$b_0 \rightarrow b_1 \rightarrow b_2 \rightarrow \cdots \rightarrow b_{n_\gamma} \xrightarrow{\uparrow} b_{n_\gamma+1} \rightarrow \cdots$$

truncate here (n_γ may be $0, 1, 2, \dots$)

From equation (18) we evidently have a compulsory condition for β

$$n_\gamma + l + 1 = \beta$$

Defining

$$n = n_\gamma + l + 1$$

(the so-called **principle quantum number**) we have

$$\beta = n$$

i.e., β must take positive integer values to truncate the power series!

Originally, we have no mind of the value of β , it is dimensionless but the value of it is not known in advance. Finally, the radial wave equation is expressed as

$$u(\rho) = e^{-\rho/2} \sum_{\gamma=0}^{n_\gamma} b_\gamma \rho^{\gamma+l+1}$$

while b_γ 's satisfy the recursion relation

$$b_{\gamma+1} = \frac{(\gamma + l + 1 - n)}{(\gamma + l + 2)(\gamma + l + 1) - l(l + 1)} b_\gamma$$

B. The property of eigenstates

1. Energy quantization

β determines E through equation (17)

$$\beta = \frac{Ze^2}{\hbar} \left(\frac{m}{2(-E)} \right)^{1/2} = n$$

thus

$$E = -\frac{Z^2}{n^2} \frac{1}{2} \frac{e^2 m e^2}{\hbar^2} = -\frac{Z^2}{n^2} \frac{1}{2} \frac{e^2}{\hbar^2 / m e^2}$$

For hydrogen

$$E = -\frac{e^2}{2a_0} \frac{1}{n^2} \tag{19}$$

where $a_0 = \hbar^2/mc^2 = 0.529\text{\AA}$ is the **Bohr radius**. This is the famous **Bohr formula** – by any measure the most important result in all of quantum mechanics. Bohr obtained it in 1913 by a serendipitous mixture of inapplicable classical physics and premature quantum theory (the Schrödinger equation did not come until 1924). The ground state (that is, the state of lowest energy) is the case $n = 1$; putting in the accepted values for the physical constants, we get

$$E_1 = -\frac{e^2}{2a_0} = -13.6 \text{ eV}$$

Evidently the binding energy of hydrogen (the amount of energy you would have to impart to the electron in order to ionize the atom) is 13.6eV . The energy level scheme for hydrogen atom is determined solely by principal quantum number n

$$E_n = \frac{E_1}{n^2}.$$

For $n = 2$, the energy for the first excited state is

$$E_2 = -\frac{13.6}{4}\text{eV} = -3.4\text{eV}.$$

2. Angular momentum and degeneracy

The spatial wave function can be explicitly written as

$$\psi = R_{nl}(r)Y_{lm}(\theta, \varphi)$$

with the radial function $R_{nl}(r)$ is dependent on both n and l . For the spherical harmonics $Y_{lm}(\theta, \varphi)$ we already know that the eigenvalues of the angular momentum are quantized.

$$\begin{aligned}\hat{L}^2 Y_{lm}(\theta, \varphi) &= l(l+1)\hbar^2 Y_{lm}(\theta, \varphi) \\ \hat{L}_z Y_{lm}(\theta, \varphi) &= m\hbar Y_{lm}(\theta, \varphi)\end{aligned}$$

However, we notice the relation between the quantum numbers

$$n = n_\gamma + l + 1 \geq l + 1$$

For a given n , the possible values of l are

$$l = 0, 1, 2, \dots, n-1,$$

(the electrons are known as *s, p, d, f, ...* etc.) and for each l , there are $(2l+1)$ possible values of m ,

$$m = -l, -l+1, \dots, l$$

The spectrum is said to be *degenerate* if two or more linearly independent eigenfunctions share the same eigenvalue. The number of linearly independent eigenfunctions is the *degeneracy*. Evidently in the case of hydrogen atom, the energy levels depend only on the quantum number n and the other two quantum numbers l and m are irrelevant. The total degeneracy of the energy level E_n is thus

$$d(n) = \sum_{l=0}^{n-1} (2l+1) = n^2$$

The degeneracy will be removed if the spin orbital coupling be considered, which will be discussed later in this course. Here we show the energy levels and transitions in the spectrum of hydrogen atom in Figure 11. Transitions to the ground state, first excited state, and the second excited state are known as **Lyman**, **Balmer**, and **Paschen series**, respectively.

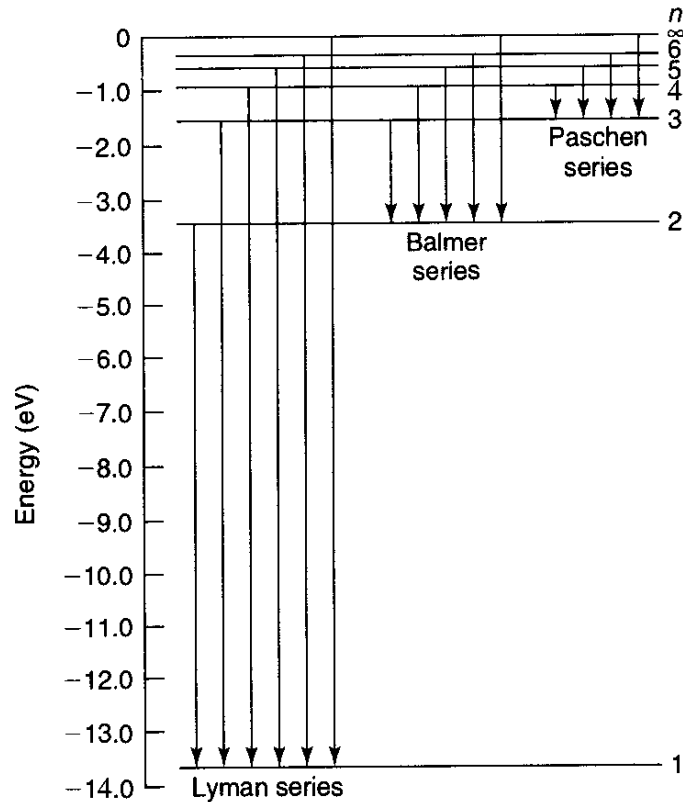


FIG. 11: The energy levels and transitions in the spectrum of hydrogen atom.

3. The wave function

We are dealing with the Schrödinger equation of hydrogen atom

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

Evidently the spatial wave functions for hydrogen are labeled by three quantum numbers (n , l , and m):

$$\begin{aligned} \psi &= R_{nl}(r)Y_{lm}(\theta, \varphi) = \frac{1}{r}e^{-\rho/2}\rho^{l+1} \sum_{\gamma=0}^{n_\gamma} b_\gamma \rho^\gamma Y_{lm}(\theta, \varphi) \\ &= \frac{1}{r}e^{-\rho/2}\rho^{l+1} (b_0 + b_1\rho + \dots + b_{n_\gamma}\rho^{n_\gamma}) Y_{lm}(\theta, \varphi) \end{aligned}$$

with

$$b_{\gamma+1} = \frac{(\gamma + l + 1 - n)}{(\gamma + l + 2)(\gamma + l + 1) - l(l + 1)} b_\gamma$$

It is clear that the lowest power of ρ in R_{nl} is l , while the highest power is $l + n_\gamma = n - 1$. The recursion formula depends on both n and l . We go back to our original notation. Remember that $\rho = \alpha r$ thus

$$\alpha = \left(\frac{8m(-E)}{\hbar^2} \right)^{1/2} = \left(\frac{8m}{\hbar^2} \frac{e^2}{2a_0} \frac{1}{n^2} \right)^{1/2} = \left(\frac{4}{a_0^2} \frac{1}{n^2} \right)^{1/2} = \frac{2}{na_0}$$

which gives

$$\rho = \alpha r = \frac{2r}{na_0}$$

and

$$R_{nl}(r) = \frac{2}{na_0} e^{-\frac{r}{na_0}} \sum_{\gamma=0}^{n-l-1} b_{\gamma} \left(\frac{2r}{na_0} \right)^{l+\gamma}$$

Now we discuss the first few states.

- The ground state $n = 1, l = 0$, i.e. $1s$ electron. We have

$$R_{10} = \frac{2b_0}{a_0} e^{-\frac{r}{a_0}}$$

and the constant b_0 is found by normalization

$$\int_0^{\infty} R_{10}^2 r^2 dr = 1 \Rightarrow \frac{a_0}{2} b_0^2 \int_0^{\infty} e^{-r} r^2 dr = a_0 b_0^2 = 1 \Rightarrow b_0 = 1/\sqrt{a_0}$$

so

$$R_{10} = \left(\frac{1}{a_0} \right)^{3/2} 2e^{-\frac{r}{a_0}}$$

It forces $l = 0$, whence also $m = 0$

$$\psi_{100} = R_{10} Y_{00} = \left(\frac{1}{a_0^3} \right)^{1/2} 2e^{-\frac{r}{a_0}} \frac{1}{\sqrt{4\pi}}$$

Finally we obtain *the ground state spatial wave function of hydrogen atom*

$$\psi_{100} = \frac{1}{\sqrt{\pi a_0^3}} e^{-\frac{r}{a_0}}$$

This wave function has the following features.

1. It has spherical symmetry;
2. The amplitude exponentially decays with increasing r ;
3. The probability for the electron to appear in district $r \rightarrow r + dr$ is

$$\begin{aligned} P(r)dr &= |\psi_{100}|^2 4\pi r^2 dr = \frac{1}{\pi a_0^3} e^{-\frac{2r}{a_0}} 4\pi r^2 dr \\ &= \frac{4}{a_0^3} e^{-\frac{2r}{a_0}} r^2 dr \end{aligned}$$

The most probable value of r occurs at r_m determined by (the peak position in the distribution)

$$\frac{\partial}{\partial r} \left(e^{-\frac{2r}{a_0}} r^2 \right) = 0 \Rightarrow r_m = a_0$$

(Generally for R_{nl} with $l = n - 1$ we have $r_m = n^2 a_0$, $n = 1, 2, 3, \dots$)

4. Normalization: Both R_{10} and Y_{00} are already normalized, so

$$\iiint \psi_{100}^* \psi_{100} d^3r = \int_0^{\infty} R_{10}^2(r) r^2 dr \iint Y_{00}^* Y_{00} d\Omega = 1.$$

- The degenerate first excited state $n = 2, l = 0$, i.e. $2s$ electron. The radial part

$$R_{20}(r) = \left(\frac{1}{2a_0} \right)^{3/2} \left(2 - \frac{r}{a_0} \right) e^{-\frac{r}{2a_0}}$$

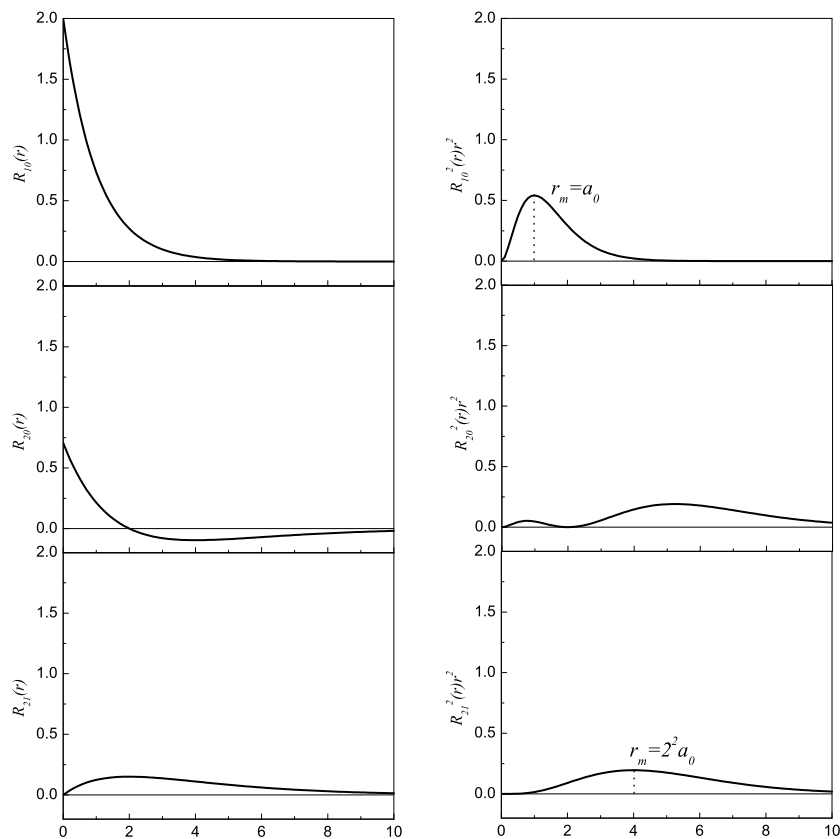


FIG. 12: Graphs of the first few radial functions and the corresponding radial distributions for hydrogen atom.

forces $l = 0$ and $m = 0$, so

$$Y_{00}(\theta, \varphi) = \frac{1}{\sqrt{4\pi}}.$$

These two gives the spatial wave function

$$\psi_{200} = \frac{1}{\sqrt{4\pi}} \left(\frac{1}{2a_0} \right)^{3/2} \left(2 - \frac{r}{a_0} \right) e^{-\frac{r}{2a_0}}$$

The probability for electron to appears in district $r \rightarrow r + dr$ reads

$$P(r)dr = |\psi_{200}|^2 4\pi r^2 dr = \frac{r^2}{8a_0^3} \left(2 - \frac{r}{a_0} \right)^2 e^{-\frac{r}{a_0}} dr$$

The normalization is again automatically done

$$\iiint R_{20}^*(r) Y_{00}^* R_{20}(r) Y_{00} r^2 dr d\Omega = \int_0^\infty R_{20}^2(r) r^2 dr \iint Y_{00}^* Y_{00} \sin \theta d\theta d\varphi = 1$$

- The degenerate first excited state $n = 2$, $l = 1$, in which case $m = 1, 0, -1$, i.e. $2p$ electron.

$$R_{21}(r) = \left(\frac{1}{2a_0}\right)^{3/2} \frac{r}{\sqrt{3}a_0} e^{-\frac{r}{2a_0}}$$

Then we have three wave functions

$$\psi_{21m} = R_{21}(r)Y_{1m}(\theta, \varphi)$$

with

$$Y_{1,\pm 1} = \mp \sqrt{\frac{3}{8\pi}} \sin \theta \exp(\pm i\varphi)$$

$$Y_{1,0} = \sqrt{\frac{3}{4\pi}} \cos \theta$$

We discuss here about the probability distribution in the case of hydrogen atom. We should be very careful about the probability, which is not simply the square of the wave function. The *radial distribution* is defined as the probability for electron to appear in shell $r \rightarrow r + dr$ (any direction)

$$P(r)dr = |R_{nl}(r)|^2 r^2 dr = |u_{nl}(r)|^2 dr$$

while the *directional distribution* is the probability for electron to appear in the region $\theta \rightarrow \theta + d\theta$ and $\varphi \rightarrow \varphi + d\varphi$, i.e. $d\Omega = \sin \theta d\theta d\varphi$

$$|Y_{lm}(\theta, \varphi)|^2 d\Omega \propto |P_l^m(\cos \theta)|^2 d\Omega$$

The normalization condition is always satisfied

$$\iiint R_{nl}^*(r) Y_{lm}^* R_{nl}(r) Y_{lm} r^2 dr \sin \theta d\theta d\varphi$$

$$= \int_0^\infty R_{nl}^2(r) r^2 dr \int_0^\pi \int_0^{2\pi} Y_{lm}^*(\theta, \varphi) Y_{lm}(\theta, \varphi) \sin \theta d\theta d\varphi = 1$$

We show the first few wave functions and the corresponding radial distributions in Figure 12.

4. Summary on the quantum numbers

The Bohr model was a one-dimensional model that used one quantum number to describe the distribution of electrons in the atom. The only information that was important was the size of the orbit, which was described by the n quantum number. Schrödinger's model allowed the electron to occupy three-dimensional space. It therefore required three coordinates, or three quantum numbers, to describe the orbitals in which electrons can be found.

The three coordinates that come from Schrödinger's wave equations are the principal (n), angular (l), and magnetic (m) quantum numbers. These quantum numbers describe the size, shape, and orientation in space of the orbitals on an atom. For the election of the hydrogen atom, its energy eigenstate is expressed by the wave function ψ_{nlm}

$$\psi_{nlm} = R_{nl}(r)Y_{lm}(\theta, \varphi)$$

- The **principal quantum number** (n) describes the *size* of the orbital. Orbitals for which $n = 2$ are larger than those for which $n = 1$, for example. Because they have opposite electrical charges, electrons are attracted to the nucleus of the atom. Energy must therefore be absorbed to excite an electron from an orbital in which the electron is close to the nucleus ($n = 1$) into an orbital in which it is further from the nucleus ($n = 2$). The principal quantum number therefore indirectly describes the energy of an orbital.

$$E_n = -\frac{e^2}{2a_0} \frac{1}{n^2}$$

- The **angular quantum number** (l) describes the *shape* of the orbital. Orbitals have shapes that are best described as spherical ($l = 0$), polar ($l = 1$), or cloverleaf ($l = 2$) (Figure 13). They can even take on more complex shapes as the value of the angular quantum number becomes larger.

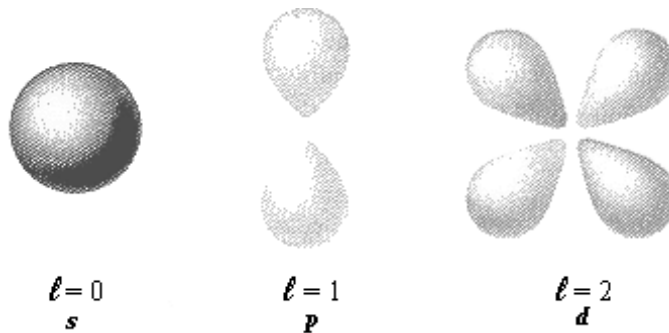


FIG. 13: The angular quantum number describes the shape of the orbital.

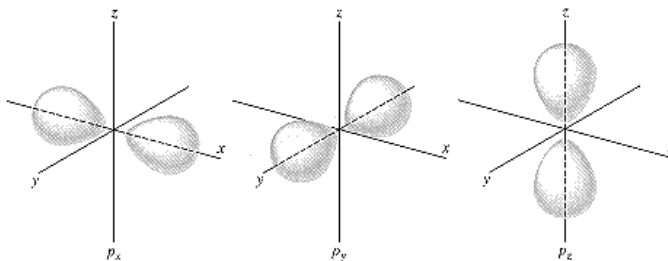


FIG. 14: The magnetic quantum number describes the orientation of the orbital.

- There is only one way in which a sphere ($l = 0$) can be oriented in space. Orbitals that have polar ($l = 1$) or cloverleaf ($l = 2$) shapes, however, can point in different directions. We therefore need a third quantum number, known as the **magnetic quantum number** (m), to describe the *orientation* in space of a particular orbital. (It is called the magnetic quantum number because the effect of different orientations of orbitals was first observed in the presence of a magnetic field). For example, there is only one orbital in the $2s$ subshell. But, there are three orbitals in the $2p$ subshell because there are three directions in which a p orbital can point. One of these orbitals is oriented along the X axis, another along the Y axis, and the third along the Z axis of a coordinate system, as shown in Figure 14. These orbitals are therefore known as the $2p_x$, $2p_y$, and $2p_z$ orbitals.

Problem 5 Zeng Jinyan, *Textbook*, page 190, 4,5,6

V. SUMMARY ON PART III

The state of a particle in quantum mechanics is described by a wavefunction ψ , while the physical quantities/observables prove themselves differential operators acting on the wavefunction. Writing out the explicit form of physical quantity F in Cartesian coordinates, changing the coordinate x and momentum p into corresponding operators \hat{x} and $\hat{p}_x = -i\hbar\partial/\partial x$, we get the operator $\hat{F}(x, -i\hbar\partial/\partial x)$ with expectation value

$$\langle F \rangle = \int \psi^* \hat{F} \psi d^3\mathbf{r}$$

All physical quantities are changed into and are expressed by Hermitian operators, for which there is a general criterion

$$(\psi_1, \hat{A}\psi_2) = (\hat{A}\psi_1, \psi_2)$$

Given an operator \hat{F} , there is a series of eigenfunctions and eigenvalues

$$\hat{F}\psi_n = f_n\psi_n$$

Typical examples of Hermitian operators are the position \hat{x} , the momentum operator $\hat{\mathbf{p}}$, the Hamiltonian operator \hat{H} , and orbital angular momentum $\hat{\mathbf{L}} = \hat{\mathbf{r}} \times \hat{\mathbf{p}}$. The eigenfunctions of \hat{x} and $\hat{\mathbf{p}}$ are $g_y(x) = \delta(x - y)$ plane waves $\psi = Ae^{\frac{i}{\hbar}\mathbf{p}\cdot\mathbf{r}}$ and the latter can only be normalized by δ -normalization or in a box.

The quantum mechanics in three dimensional depends heavily on the properties of angular momentum $\hat{\mathbf{L}}$. The operators \hat{L}^2 and \hat{L}_z share common eigenfunctions

$$\begin{aligned}\hat{L}^2 Y_{lm}(\theta, \varphi) &= l(l+1)\hbar^2 Y_{lm}(\theta, \varphi), & l &= 0, 1, 2 \dots \\ \hat{L}_z Y_{lm}(\theta, \varphi) &= m\hbar Y_{lm}(\theta, \varphi), & m &= -l, -l+1, \dots, 0, 1, \dots, l\end{aligned}$$

where $Y_{lm}(\theta, \varphi)$ are spherical harmonics. For $m = 0$, they reduce to the Legendre polynomials $Y_{l0}(\theta, \varphi) = P^l(\cos \theta)$. In a central force field, the solutions are separable into products of a radial part and an angular part $\psi(r, \theta, \varphi) = R(r) Y_{lm}(\theta, \varphi)$. The radial wavefunctions satisfy the radial wave equation

$$-\frac{\hbar^2}{2m} \frac{d^2}{dr^2} u(r) + \left(\frac{l(l+1)\hbar^2}{2mr^2} + V(r) \right) u(r) = Eu(r)$$

where we have changed the variables such that $R(r) = u(r)/r$ and the additional term $l(l+1)\hbar^2/2mr^2$ in the effective potential is known as centrifugal potential. We can solve the equation for spherical square potential with infinite or finite barrier height. The solution for infinite spherical well for arbitrary l contains spherical Bessel function

$$\psi_{nlm}(r, \theta, \phi) = A_{nl} j_l(\beta_{nl} r/a) Y_{lm}(\theta, \phi)$$

For finite potential well, we got exactly the same transcendental equation as we encountered in 1D finite potential well, which permits no solution if $V_0 a^2 < \pi^2 \hbar^2 / 8m$.

The stationary states in Coulomb potential

$$V(r) = -\frac{1}{4\pi\epsilon_0} \frac{Ze^2}{r}$$

gives us the energy levels of hydrogen and hydrogen-like atoms. For the election of the hydrogen atom, its energy eigenstate is expressed by the wave function

$$\psi_{nlm} = R_{nl}(r) Y_{lm}(\theta, \varphi)$$

Solving the radial equation

$$-\frac{\hbar^2}{2m} \frac{d^2 u}{dr^2} + \left(-\frac{Ze^2}{r} + \frac{l(l+1)\hbar^2}{2mr^2} \right) u = Eu$$

by truncating the series to finite terms, we get the energy quantization condition - the Bohr's formula for hydrogen atom

$$E = -\frac{e^2}{2a_0} \frac{1}{n^2}$$

The ground state of hydrogen lies $-13.6eV$ beneath the zero potential point and the wavefunction is

$$\psi_{100} = \frac{1}{\sqrt{\pi a_0^3}} e^{-\frac{r}{a_0}}, \quad a_0 = 0.529 \text{ \AA}$$

The degeneracy of the energy level E_n is n^2 . The principal (n), angular (l), and magnetic (m) quantum numbers describe the size, shape, and orientation in space of the orbitals on an atom.

Lecture Notes on Quantum Mechanics - Part IV

Yunbo Zhang

Instituter of Theoretical Physics, Shanxi University

In this part we introduce the other formalism of quantum mechanics - the matrix mechanics developed by Heisenberg. To do this we need some basic knowledge on linear algebra and operator arithmetical rules. Three essential properties of the eigenfunctions of the Hermitian operators (reality, orthogonality, and completeness) constitute the frame of matrix mechanics.

Contents

I. Operators and its Arithmetical Rules	3
A. Hermitian operators	3
B. Multiplication of operators - the commutator	3
II. General Properties of Eigenvalues and Eigenfunctions of Hermitian Operators	6
A. Background	6
B. Reality of eigenvalues	7
C. Orthogonality of eigenfunctions	7
1. Examples of orthogonality	8
2. Degenerate case	9
D. Completeness of eigenfunctions	11
1. Completeness of the eigenfunctions of momentum operator	11
2. Completeness of spherical harmonics	12
3. Completeness of eigenfunctions of discrete spectrum	12
4. Summary	13
III. Generalized Statistical Interpretation and Uncertainty Principle	13
A. Generalization	13
B. The hypothesis of measurement in quantum mechanics	15
C. Reproduce of the original statistical interpretation	15
D. Commutative operators and common eigenfunctions	16
E. A rigorous proof of the uncertainty relation	17
IV. Linear Space - Mathematical Preparation	20
A. 3D ordinary space	20
B. Operation onto space	21
C. Basis transformation	22
D. Transformation of operations	23
V. Matrix Formulation of Quantum Mechanics - Representation Theory	24
A. The idea of representation	24
B. Representation of state	24
C. Representation of operators	25
D. Property of the matrix of a Hermitian operator	26
E. Matrix formulation of rules of quantum mechanics	28
VI. Unitary Transformation between two Representations	29
A. Two sets of basis and their relations	29
B. Transformation of state from representation \hat{A} to \hat{B}	30
C. Transformation of operators	31
VII. Dirac's Notation	36
A. States	36
B. Rules	37
C. $ n\rangle$	38

D. Application	39
E. Projection operators	40
VIII. Conservation Law and Symmetry	41
A. Constant of motion	41
B. Several typical cases of constant of motion	42
C. Conservation law and symmetry	45
1. Translation invariance and momentum conservation	45
2. Rotation invariance and orbital angular momentum conservation	46
3. Conservation laws and symmetry in general	48
IX. Summary on Part IV	49

I. OPERATORS AND ITS ARITHMETICAL RULES

A. Hermitian operators

Operators are introduced in quantum mechanics to represent dynamical variables. There exists one-to-one correspondence between dynamical variables (physical quantities) F and operators \hat{F} . Here we review some basic concepts by raising some questions.

- How is the operator obtained? Answer: Any physical quantity can be expressed in Cartesian coordinate $F(x, y, z, p_x, p_y, p_z)$, one gets the operator by simply replacing the position and momentum by the corresponding operators $\hat{F}(\hat{x}, \hat{y}, \hat{z}, \hat{p}_x, \hat{p}_y, \hat{p}_z)$.
- What is its function? Answer: Given the state of the system ψ , the expectation value of the physical quantity can be obtained through $\langle F \rangle = \int \psi^* \hat{F} \psi d^3\mathbf{r}$.
- What is a Hermitian operator? It means that $\langle F \rangle$ must be real. There is a general criterion for Hermitian operator $(\psi_1, \hat{F}\psi_2) = (\hat{F}\psi_1, \psi_2)$.
- Can we know more than the expectation value? Yes, For a special kind of states which satisfies the equation $\hat{F}\psi_n = f_n\psi_n$, we have $\langle (\Delta F)^2 \rangle = 0$. This is the eigenequation of \hat{F} .
- Are only Hermitian operators prevail in quantum mechanics? No, in quantum theory different kinds of operators are used, many of them are not Hermitian (but intimately related to Hermitian operators). A typical example is $\hat{A} = d/dx$.

In summary, there is a special class of operators which are called Hermitian operators. They are of particular importance in quantum mechanics because they have the property that all of their eigenvalues are real (proved later). This is convenient for the measurement outcome of any experiment must be a real number. There are non-Hermitian operators, but they do not correspond to observable properties. All observable properties are represented by Hermitian operators (but not all Hermitian operators correspond to an observable property).

B. Multiplication of operators - the commutator

Here we define the algebraic manipulation of operators. The meaning of multiplication of operators \hat{A} and \hat{B} is that two operators act successively on a wave function

$$\hat{A}\hat{B}\psi = \hat{A}(\hat{B}\psi)$$

in this sense the multiplication is order-concerned! *The commutation relation is the difference of acting results.* Here are some examples:

The most important and fundamental one is the commutator between \hat{x} and \hat{p}_x . For arbitrary state ψ we have

$$\begin{aligned}\hat{x}\hat{p}_x\psi &= x \left(-i\hbar \frac{\partial}{\partial x} \right) \psi \\ \hat{p}_x\hat{x}\psi &= \left(-i\hbar \frac{\partial}{\partial x} \right) (x\psi) = -i\hbar\psi + x \left(-i\hbar \frac{\partial}{\partial x} \right) \psi.\end{aligned}$$

The difference of acting results

$$(\hat{x}\hat{p}_x - \hat{p}_x\hat{x})\psi = i\hbar\psi$$

is irrelevant with the choice of state, thus

$$\hat{x}\hat{p}_x - \hat{p}_x\hat{x} = i\hbar$$

We say that operators \hat{x} and \hat{p}_x do not commute (are non-commutative) and obey the above commutation relation, or, in other words, the commutator between \hat{x} and \hat{p}_x reads

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

Take another example, \hat{x} and \hat{p}_y . Clearly

$$\begin{aligned}\hat{x}\hat{p}_y\psi &= x\left(-i\hbar\frac{\partial}{\partial y}\right)\psi \\ \hat{p}_y\hat{x}\psi &= \left(-i\hbar\frac{\partial}{\partial y}\right)(x\psi) = x\left(-i\hbar\frac{\partial}{\partial y}\right)\psi. \\ (\hat{x}\hat{p}_y - \hat{p}_y\hat{x})\psi &= 0\end{aligned}$$

or

$$[\hat{x}, \hat{p}_y] = 0$$

We say that operators \hat{x} and \hat{p}_y commute each other (are commutative).

For any two operators \hat{A} and \hat{B} , we can calculate the commutation relation between them $[\hat{A}, \hat{B}]$. For example,

$$\begin{aligned}[\hat{x}, \hat{L}_y] &= \hat{x}\hat{L}_y - \hat{L}_y\hat{x} = \hat{x}(\hat{z}\hat{p}_x - \hat{x}\hat{p}_z) - (\hat{z}\hat{p}_x - \hat{x}\hat{p}_z)\hat{x} \\ &= \hat{z}\hat{x}\hat{p}_x - \hat{x}^2\hat{p}_z - \hat{z}\hat{p}_x\hat{x} + \hat{x}\hat{p}_z\hat{x} = \hat{z}(\hat{x}\hat{p}_x - \hat{p}_x\hat{x}) = i\hbar\hat{z}\end{aligned}$$

\hat{x} and \hat{L}_y are non-commutative. Generally, we can see that the operators obey a non-commutative algebra.

There is another question about the Hermiticity of the product of two operators.

Proposition 1 *The product of two Hermitian operators are generally not Hermitian, unless they commute each other.*

Proof. The Hermitian operator \hat{A} satisfies

$$(\psi_1, \hat{A}\psi_2) = (\hat{A}\psi_1, \psi_2) \text{ or } \hat{A}^\dagger = \hat{A}$$

The product of \hat{A} and \hat{B} , however

$$\begin{aligned}(\psi_1, \hat{A}\hat{B}\psi_2) &= \int \psi_1^* \hat{A}(\hat{B}\psi_2) d^3\mathbf{r} = \int (\hat{A}^\dagger\psi_1)^* (\hat{B}\psi_2) d^3\mathbf{r} \\ &= \int (\hat{B}^\dagger\hat{A}^\dagger\psi_1)^* \psi_2 d^3\mathbf{r} = (\hat{B}^\dagger\hat{A}^\dagger\psi_1, \psi_2)\end{aligned}$$

i.e.

$$(\hat{A}\hat{B})^\dagger = \hat{B}^\dagger\hat{A}^\dagger = \hat{B}\hat{A}$$

That means, the Hermiticity of \hat{A} and \hat{B} can not assure the Hermiticity of their product $\hat{A}\hat{B}$. The product operator $\hat{A}\hat{B}$ is Hermitian only if

$$\hat{B}\hat{A} = \hat{A}\hat{B}$$

i.e. $[\hat{A}, \hat{B}] = 0$. The operator $\hat{L}_x = \hat{y}\hat{p}_z - \hat{z}\hat{p}_y$ is composed of 4 Hermitian operators, all of them are Hermitian, and in the products \hat{y} and \hat{p}_z , \hat{z} and \hat{p}_y are commutative, respectively. Thus \hat{L}_x is Hermitian. By the same token, $\hat{L}^2 = \hat{L}_x^2 + \hat{L}_y^2 + \hat{L}_z^2$ is Hermitian, too. ■

Now we calculate the commutation relation between the components of the angular momentum. Till now we have met several vector operators, $\hat{\mathbf{x}}$, $\hat{\mathbf{p}}$ and $\hat{\mathbf{L}}$. The components of $\hat{\mathbf{x}}$, $\hat{\mathbf{p}}$ are commutative evidently

$$\begin{aligned}[\hat{x}, \hat{y}] &= 0, \text{ etc.} \\ [\hat{p}_x, \hat{p}_y] &= 0, \text{ etc.}\end{aligned}$$

The operators \hat{L}_x and \hat{L}_y do not commute, in fact

$$\begin{aligned}[\hat{L}_x, \hat{L}_y] &= (\hat{y}\hat{p}_z - \hat{z}\hat{p}_y)(\hat{z}\hat{p}_x - \hat{x}\hat{p}_z) - (\hat{z}\hat{p}_x - \hat{x}\hat{p}_z)(\hat{y}\hat{p}_z - \hat{z}\hat{p}_y) \\ &= \hat{y}\hat{p}_z\hat{z}\hat{p}_x - \hat{y}\hat{p}_z\hat{x}\hat{p}_z - \hat{z}\hat{p}_y\hat{z}\hat{p}_x + \hat{z}\hat{p}_y\hat{x}\hat{p}_z \\ &\quad - \hat{z}\hat{p}_x\hat{y}\hat{p}_z + \hat{x}\hat{p}_z\hat{y}\hat{p}_z + \hat{z}\hat{p}_x\hat{z}\hat{p}_y - \hat{x}\hat{p}_z\hat{z}\hat{p}_y \\ &= \hat{y}\hat{p}_z[\hat{p}_z, \hat{z}] + \hat{x}\hat{p}_y[\hat{z}, \hat{p}_z] = i\hbar(\hat{x}\hat{p}_y - \hat{y}\hat{p}_x) = i\hbar\hat{L}_z\end{aligned}$$

Of course, we could have started out with $[\hat{L}_y, \hat{L}_z]$ or $[\hat{L}_x, \hat{L}_z]$, but there is no need to calculate these separately - we can get them immediately by cyclic permutation of the indices ($\hat{x} \rightarrow \hat{y}, \hat{y} \rightarrow \hat{z}, \hat{z} \rightarrow \hat{x}$)

$$\begin{aligned} [\hat{L}_x, \hat{L}_y] &= i\hbar\hat{L}_z \\ [\hat{L}_y, \hat{L}_z] &= i\hbar\hat{L}_x \\ [\hat{L}_z, \hat{L}_x] &= i\hbar\hat{L}_y \end{aligned}$$

or

$$[\hat{L}_\alpha, \hat{L}_\beta] = i\hbar\epsilon_{\alpha\beta\gamma}\hat{L}_\gamma$$

with $\epsilon_{\alpha\beta\gamma}$ the Levi-Civita permutation symbol. It is an anti-symmetric tensor rank 3 defined as

$$\begin{aligned} \epsilon_{\alpha\beta\gamma} &= -\epsilon_{\beta\alpha\gamma} = -\epsilon_{\alpha\gamma\beta} \\ \epsilon_{xyz} &= 1, \epsilon_{xxz} = 0, \text{etc} \end{aligned}$$

They are the fundamental commutation relations for angular momentum; everything else follows from them.

On the other hand, the square of the total angular momentum

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

does commute with \hat{L}_x

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

where we have used the result of Problem in the following. It follows that

$$[\hat{L}^2, \hat{L}_\alpha] = 0, \quad \alpha = x, y, z$$

Problem 2 Define

$$\hat{L}_\pm = \hat{L}_x \pm i\hat{L}_y$$

Show that

$$\begin{aligned} [\hat{L}_z, \hat{L}_\pm] &= \pm\hbar\hat{L}_\pm \\ [\hat{L}_+, \hat{L}_-] &= 2\hbar\hat{L}_z \end{aligned}$$

One can manipulate the operators in many ways, such as summation, product and more generally, given a function $F(x)$, the power series converge if the derivatives are well defined

$$F(x) = \sum_{n=0}^{+\infty} \frac{F^{(n)}(0)}{n!} x^n$$

the function of operator \hat{A} is expressed as

$$F(\hat{A}) = \sum_{n=0}^{+\infty} \frac{F^{(n)}(0)}{n!} \hat{A}^n$$

Specially for $F(x) = e^{\alpha x}$ one may define

$$F\left(\frac{d}{dx}\right) = e^{\alpha \frac{d}{dx}} = \sum_{n=0}^{+\infty} \frac{\alpha^n}{n!} \frac{d^n}{dx^n}.$$

Problem 3 Show that the expectation values of \hat{L}_x and \hat{L}_y in any eigenstate of \hat{L}_z are zero. (Zeng, p133, 13)

Problem 4 Prove the following commutator identity

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

Show that

$$[\hat{x}^n, \hat{p}] = i\hbar n \hat{x}^{n-1}$$

and more generally that

$$[f(\hat{x}), \hat{p}] = i\hbar \frac{df}{dx}$$

for any function $f(x)$.

II. GENERAL PROPERTIES OF EIGENVALUES AND EIGENFUNCTIONS OF HERMITIAN OPERATORS

A. Background

We have learnt many eigenfunctions of Hermitian operators:

- For 1D momentum, the eigenfunction is

$$\hat{p}_x \psi = -i\hbar \frac{d}{dx} \psi = p_x \psi$$

where p_x is the eigenvalue and the eigenfunctions are chosen such that they can be δ -normalized

$$\psi_{p_x}(x) = \frac{1}{\sqrt{2\pi\hbar}} e^{\frac{i}{\hbar} p_x x}$$

where $-\infty < p_x < \infty$ i.e. the eigenfunctions are innumerable. The 3D momentum eigenfunction takes a similar form

$$\hat{\mathbf{p}}\psi = -i\hbar \nabla \psi = (\mathbf{x}_0 p_x + \mathbf{y}_0 p_y + \mathbf{z}_0 p_z) \psi = \mathbf{p}\psi$$

and the eigenfunctions are

$$\psi = \frac{1}{(2\pi\hbar)^{3/2}} e^{\frac{i}{\hbar} (p_x x + p_y y + p_z z)}.$$

- We have also solved the eigenfunctions for the square of angular momentum and its z -component

$$\hat{L}^2 Y_{lm} = l(l+1)\hbar^2 Y_{lm}$$

$$\hat{L}_z Y_{lm} = m\hbar Y_{lm}$$

with the eigenvalues $l = 0, 1, 2, \dots$ and $m = -l, -l+1, \dots, l$ and the eigenfunctions are spherical harmonics $Y_{lm}(\theta, \varphi)$.

- The other example of eigenvalue problem is the eigenequation of Hamiltonian operator, the stationary Schrödinger equation

$$\hat{H}\psi_n = E_n \psi_n$$

Typical exactly solvable models include the simple harmonic oscillator, square potential well, hydrogen atom etc.

These examples provide us the information of micro-particle in concrete conditions and play important role in the construction of the theoretical frame of quantum mechanics. In this section we discuss the general property of eigenvalues and eigenfunctions of Hermitian operators.

The spectrum of eigenvalues of Hermitian operators fall into two categories: If the spectrum is **discrete** (i.e., the eigenvalues are separated from one another) then the eigenfunctions constitute physically realizable states. If the spectrum **continuous**, (i.e., the eigenvalues fill out an entire range) then the eigenfunctions are not normalizable, and they do not represent possible wave functions (though linear combinations of them - involving necessarily a spread in eigenvalues - may be normalizable).

- a Some operators have a discrete spectrum only (for example, the angular momentum, the Hamiltonian for the harmonic oscillator, square potential well of infinite depth),
- b some have only a continuous spectrum (for example, the free particle Hamiltonian, the 1D and 3D momentum operators),
- c and some have both a discrete part and a continuous part (for example, the Hamiltonian for a finite square well, the spectrum is discrete for $E < V_0$ and continuous for $E > V_0$; the Hamiltonian for the system of nucleus plus electron, the energy spectrum for hydrogen atom is discrete for $E < 0$ and continuous for the scattering of electrons $E > 0$)

The essence of theory to be developed is the same for the three cases, but more tedious for the third case. We will presume that the eigenvalues of the system would be entirely discrete here and discuss the situation of continuous spectrum later through some examples.

B. Reality of eigenvalues

Mathematically, the normalizable eigenfunctions of a Hermitian operator have the following properties:

Theorem 5 *Their eigenvalues are real.*

Proof. Suppose

$$\hat{F}\psi_m = f_m\psi_m$$

Multiplying ψ_m^* from the left, and integrating over entire space, we see

$$\begin{aligned} \int \psi_m^* (\hat{F}\psi_m = f_m\psi_m) d^3\mathbf{r} \\ f_m = \int \psi_m^* \hat{F}\psi_m d^3\mathbf{r} = \int (\hat{F}\psi_m)^* \psi_m d^3\mathbf{r} \\ = \int (f_m\psi_m)^* \psi_m d^3\mathbf{r} = f_m^* \end{aligned}$$

where we have used the property that \hat{F} is Hermitian and ψ_m is normalized. Evidently f_m are real for all Hermitian operators. ■

C. Orthogonality of eigenfunctions

Theorem 6 *Two eigenfunctions of the same Hermitian operator with different eigenvalues are orthogonal. Suppose*

$$\begin{aligned} \hat{F}\psi_m &= f_m\psi_m \\ \hat{F}\psi_n &= f_n\psi_n \end{aligned}$$

where $f_m \neq f_n$. Orthogonality means

$$\int \psi_m^* \psi_n d^3\mathbf{r} = 0$$

Proof. \hat{F} is a Hermitian operator, so

$$\int \psi_m^* \hat{F} \psi_n d^3 \mathbf{r} = \int \psi_m^* \hat{F} \psi_n d^3 \mathbf{r}$$

gives

$$f_n \int \psi_m^* \psi_n d^3 \mathbf{r} = f_m \int \psi_m^* \psi_n d^3 \mathbf{r}$$

i.e.

$$(f_m - f_n) \int \psi_m^* \psi_n d^3 \mathbf{r} = 0$$

Because $(f_m - f_n) \neq 0$, it must be that $\int \psi_m^* \psi_n d^3 \mathbf{r} = 0$. ■

1. Examples of orthogonality

Two stationary states with different energy values of the Schrödinger equation are orthogonal, i.e.

$$\begin{aligned} \hat{H} \psi_m &= E_m \psi_m \\ \hat{H} \psi_n &= E_n \psi_n \\ \iiint \psi_m^* \psi_n d^3 \mathbf{r} &= 0 \end{aligned}$$

Spherical harmonics with different indices are orthogonal, i.e.

$$\begin{aligned} \iint Y_{lm}^* Y_{l'm} d\Omega &= 0, l \neq l' \\ \iint Y_{lm}^* Y_{lm'} d\Omega &= 0, m \neq m' \end{aligned}$$

Orthogonality is for different eigenvalue states, while normalization is for the same state. The combination of them is called orthonormality.

$$\int \psi_m^*(\mathbf{r}) \psi_n(\mathbf{r}) d^3 \mathbf{r} = \delta_{mn} \quad (1)$$

For example we can write the orthogonality and normalization of spherical harmonics into

$$\iint Y_{lm}^*(\theta, \varphi) Y_{l'm'}(\theta, \varphi) d\Omega = \delta_{ll'} \delta_{mm'}$$

If the spectrum of a hermitian operator is continuous, the eigenfunctions are not normalizable, and the proofs of the above two Theorems fail, because the wavefunctions are not square integrable. For example, the eigenvalues of 1D momentum are real and the eigenfunctions satisfy

$$\int \left(\frac{1}{\sqrt{2\pi\hbar}} e^{\frac{i}{\hbar} p x} \right)^* \left(\frac{1}{\sqrt{2\pi\hbar}} e^{\frac{i}{\hbar} p' x} \right) dx = \delta(p - p')$$

i.e.

$$\int \psi_p^*(x) \psi_{p'}(x) dx = \delta(p - p') \quad (2)$$

which is strikingly reminiscent of true orthonormality equation (1) - the indices are now continuous variables, and the Kronecker delta has become a Dirac delta, but otherwise it looks just the same. Equation (2) is called **Dirac orthonormality**.

2. Degenerate case

We here discuss more about the degeneracy of the energy spectrum of Schrödinger equation. If two (or more) distinct solutions to the (time-independent) Schrödinger equation have the same energy E , these states are said to be *degenerate*. For example, the free particle states are doubly degenerate – one solution representing motion to the right, and the other motion to the left. But we have encountered no *normalizable* degenerate solutions, and this is not an accident. We have the following theorem:

Theorem 7 *In one dimension there are no degenerate bound states.*

Proof. Suppose there are two solutions, ψ_1 and ψ_2 , with the same energy E .

$$\begin{aligned} -\frac{\hbar^2}{2m} \frac{d^2\psi_1}{dx^2} + V\psi_1 &= E\psi_1 \\ -\frac{\hbar^2}{2m} \frac{d^2\psi_2}{dx^2} + V\psi_2 &= E\psi_2 \end{aligned}$$

Multiply the Schrödinger equation for ψ_1 by ψ_2 , and the Schrödinger equation for ψ_2 by ψ_1 ,

$$\begin{aligned} -\frac{\hbar^2}{2m} \psi_2 \frac{d^2\psi_1}{dx^2} + V\psi_2\psi_1 &= E\psi_2\psi_1 \\ -\frac{\hbar^2}{2m} \psi_1 \frac{d^2\psi_2}{dx^2} + V\psi_1\psi_2 &= E\psi_1\psi_2 \end{aligned}$$

and subtract, we have

$$-\frac{\hbar^2}{2m} \left(\psi_2 \frac{d^2\psi_1}{dx^2} - \psi_1 \frac{d^2\psi_2}{dx^2} \right) = 0$$

But

$$\frac{d}{dx} \left(\psi_2 \frac{d\psi_1}{dx} - \psi_1 \frac{d\psi_2}{dx} \right) = \psi_2 \frac{d^2\psi_1}{dx^2} - \psi_1 \frac{d^2\psi_2}{dx^2}$$

it follows that

$$\psi_2 \frac{d\psi_1}{dx} - \psi_1 \frac{d\psi_2}{dx} = K$$

where K is a constant. For normalizable solutions $\psi \rightarrow 0$ at $x \rightarrow \infty$, this constant is in fact zero. Thus

$$\psi_2 \frac{d\psi_1}{dx} = \psi_1 \frac{d\psi_2}{dx} \Rightarrow \frac{1}{\psi_1} \frac{d\psi_1}{dx} = \frac{1}{\psi_2} \frac{d\psi_2}{dx}$$

so

$$\ln \psi_1 = \ln \psi_2 + \text{const.} \Rightarrow \psi_1 = (\text{constant}) \psi_2$$

and hence that the two solutions are not distinct. ■

Imagine now a bead of mass m that slides frictionlessly around a circular wire ring of circumference L . This is just like a free particle, except that $\psi(x) = \psi(x + L)$, since $x + L$ is the same point as x . We try to find the stationary states (with appropriate normalization) and the corresponding allowed energies. From the equation

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

where x is measured around the circumference, we immediately know the solution is

$$\psi(x) = Ae^{ikx} + Be^{-ikx}, \quad k = \sqrt{2mE/\hbar^2}.$$

The boundary condition gives

$$Ae^{ikx} e^{ikL} + Be^{-ikx} e^{-ikL} = Ae^{ikx} + Be^{-ikx},$$

and this is true for all x . In particular, for $x = 0$

$$Ae^{ikL} + Be^{-ikL} = A + B \quad (3)$$

And for $x = \pi/2k$

$$\begin{aligned} Ae^{i\pi/2}e^{ikL} + Be^{-i\pi/2}e^{-ikL} &= Ae^{i\pi/2} + Be^{-i\pi/2} \\ Ae^{ikL} + Be^{-i\pi}e^{-ikL} &= A + Be^{-i\pi} \end{aligned}$$

or

$$Ae^{ikL} - Be^{-ikL} = A - B \quad (4)$$

Add (3) and (4)

$$Ae^{ikL} = A$$

Either $A = 0$, or else $e^{ikL} = 1$, in which case $kL = 2n\pi$ ($n = 0, \pm 1, \pm 2, \dots$). But if $A = 0$, then $Be^{-ikL} = B$, leading to the same conclusion. So for every positive n there are two solutions:

$$\begin{aligned} \psi^+(x) &= Ae^{i(2n\pi x/L)} \\ \psi^-(x) &= Be^{-i(2n\pi x/L)} \end{aligned}$$

The two independent solutions for each energy E_n correspond to clockwise and counterclockwise circulation ($n = 0$ is ok too, but in that case there is just one solution). Normalizing

$$\int_0^L |\psi^+(x)|^2 dx = 1 \Rightarrow A = B = \frac{1}{\sqrt{L}}$$

Any other solution (with the same energy) is a linear combination of these

$$\begin{aligned} \psi_n^\pm &= \frac{1}{\sqrt{L}} e^{\pm i(2n\pi x/L)} \\ E_n &= \frac{\hbar^2 k^2}{2m} = \frac{2n^2 \pi^2 \hbar^2}{mL^2}, n = 0, 1, 2, \dots \end{aligned}$$

How do we account for this degeneracy, in view of the theorem – that is, why does the theorem fail in this case? The theorem fails because here ψ does not go to zero at ∞ ; x is restricted to a finite range, and we are unable to determine the constant K .

Problem 8 Consider the operators

$$\hat{Q}_1 \equiv i \frac{d}{d\phi}, \hat{Q}_2 \equiv \frac{d^2}{d\phi^2}$$

where ϕ is the usual polar coordinate in two dimensions. (These two operators might arise in a physical context if we were studying the bead-on-a-ring.) Are \hat{Q}_1 and \hat{Q}_2 Hermitian? Find the eigenfunctions and eigenvalues. Are the spectra degenerate?

It is always possible to reshuffle these degenerate states so that they are normalized by themselves while orthogonal to other (the rest) states. Here is an example on how to construct the orthogonal eigenfunctions from the degenerate states.

Example 9 Suppose that $f(x)$ and $g(x)$ are two eigenfunctions of an operator \hat{Q} , with the same eigenvalue q

$$\begin{aligned} \hat{Q}f &= qf \\ \hat{Q}g &= qg \end{aligned}$$

It is easy to show that any linear combination of f and g

$$h(x) = af(x) + bg(x)$$

(for arbitrary constants a and b) is itself an eigenfunction of \hat{Q} , with eigenvalue q

$$\hat{Q}h = \hat{Q}(af + bg) = a\hat{Q}f + b\hat{Q}g = aqf + bqg = qh.$$

For example, $f(x) = e^x$ and $g(x) = e^{-x}$ are eigenfunctions of the operator d^2/dx^2 , with the same eigenvalue 1

$$\begin{aligned}\frac{d^2}{dx^2}f &= \frac{d^2}{dx^2}e^x = e^x = f \\ \frac{d^2}{dx^2}g &= \frac{d^2}{dx^2}e^{-x} = (-1)^2 e^{-x} = g\end{aligned}$$

The simplest orthogonal linear combinations are

$$\begin{aligned}\sinh x &= \frac{1}{2}(e^x - e^{-x}) = \frac{1}{2}(f - g) \\ \cosh x &= \frac{1}{2}(e^x + e^{-x}) = \frac{1}{2}(f + g)\end{aligned}$$

They are clearly orthogonal, since $\sinh x$ is odd while $\cosh x$ is even.

D. Completeness of eigenfunctions

For a 3D space, three basic vectors \mathbf{x}_0 , \mathbf{y}_0 and \mathbf{z}_0 form an orthonormal basis. Any vector can be expressed as a linear combination of the them. They are “complete”. If one of the basic vectors is missed, we can no more express all the vectors through the basis. They are “incomplete”.

A periodic function $f(x)$ of period L can be expanded into a Fourier series

$$f(x) = \sum_{n=0}^{\infty} a_n \cos \frac{2n\pi}{L}x + \sum_{n=1}^{\infty} b_n \sin \frac{2n\pi}{L}x$$

thus the function set $\{\cos \frac{2n\pi}{L}x, \sin \frac{2n\pi}{L}x\}$ serve as a basis and are used to express periodic functions. They are complete. If one of the above basis, say $\cos \frac{2\pi}{L}x$, is missed, they are incomplete and can not express arbitrary periodic functions at will.

1. Completeness of the eigenfunctions of momentum operator

As an example, we discuss here first the completeness of the eigenfunctions of 1D momentum operator $\hat{p}_x = -i\hbar \frac{d}{dx}$,

$$\psi_p(x) = \frac{1}{\sqrt{2\pi\hbar}} e^{\frac{i}{\hbar}px}.$$

The set $\{\psi_p(x)\}$ with $-\infty < p < +\infty$ are complete, i.e. any wave function can be expanded as a superposition of $\psi_p(x)$'s.

$$\psi(x) = \frac{1}{\sqrt{2\pi\hbar}} \int \phi(p) e^{\frac{i}{\hbar}px} dp = \int \phi(p) \psi_p(x) dp$$

This is nothing but the Fourier transformation between x and p spaces. Inversely, we have

$$\phi(p) = \frac{1}{\sqrt{2\pi\hbar}} \int \psi(x') e^{-\frac{i}{\hbar}px'} dx' = \int \psi_p^*(x') \psi(x') dx'$$

so

$$\begin{aligned}\psi(x) &= \int dp \phi(p) \psi_p(x) \\ &= \int dp \left(\int dx' \psi_p^*(x') \psi(x') \right) \psi_p(x) \\ &= \int dx' \left\{ \int dp \psi_p^*(x') \psi_p(x) \right\} \psi(x') \\ &= \int dx' \delta(x' - x) \psi(x')\end{aligned}$$

The completeness of $\psi_p(x)$ with p a parameter can be expressed as

$$\int dp \psi_p^*(x') \psi_p(x) = \delta(x' - x)$$

i.e.

$$\begin{aligned} \int \frac{1}{\sqrt{2\pi\hbar}} e^{-\frac{i}{\hbar} p x'} \frac{1}{\sqrt{2\pi\hbar}} e^{\frac{i}{\hbar} p x} dp &= \frac{1}{2\pi\hbar} \int e^{\frac{i}{\hbar} p(x-x')} dp \\ &= \delta(x' - x). \end{aligned}$$

2. Completeness of spherical harmonics

The spherical harmonics $\{Y_{lm}(\theta, \varphi)\}$ are complete. Any function $F(\theta, \varphi)$ can be expressed as a superposition of spherical harmonics

$$F(\theta, \varphi) = \sum_{l,m} c_{l,m} Y_{lm}(\theta, \varphi)$$

The coefficients are

$$c_{l,m} = \iint Y_{l,m}^*(\theta', \varphi') F(\theta', \varphi') d\Omega'$$

Put it back we have

$$\begin{aligned} F(\theta, \varphi) &= \iint \sum_{l,m} Y_{l,m}^*(\theta', \varphi') Y_{l,m}(\theta, \varphi) F(\theta', \varphi') d\Omega' \\ &= \iint \delta(\theta - \theta') \delta(\varphi - \varphi') F(\theta', \varphi') d\Omega' \end{aligned}$$

therefore the completeness of spherical harmonics is expressed as

$$\sum_{l,m} Y_{l,m}^*(\theta', \varphi') Y_{l,m}(\theta, \varphi) = \delta(\theta - \theta') \delta(\varphi - \varphi')$$

3. Completeness of eigenfunctions of discrete spectrum

Let us check the completeness of eigenfunctions for a simple harmonic oscillator

$$\psi_n(x) = N_n e^{-\frac{1}{2}\alpha^2 x^2} H_n(\alpha x)$$

with $\alpha = \sqrt{\frac{m\omega}{\hbar}}$. $\{\psi_n\}$ is complete. Any state of simple harmonic oscillator can be expressed as a superposition of ψ_n 's

$$\psi(x) = \sum_n a_n \psi_n(x)$$

and

$$a_n = \int \psi_n^*(x') \psi(x') dx'$$

We have

$$\begin{aligned} \psi(x) &= \sum_n \int \psi_n^*(x') \psi(x') dx' \psi_n(x) \\ &= \int \sum_n \psi_n^*(x') \psi_n(x) \psi(x') dx' \\ &= \int \delta(x' - x) \psi(x') dx' \end{aligned}$$

which gives us

$$\sum_n \psi_n^*(x') \psi_n(x) = \delta(x' - x)$$

This can be generalized to the complete set of eigenfunctions $\{u_n(x)\}$ of a certain operators \hat{F}

$$\hat{F}u_n(x) = f_n u_n(x)$$

$\{u_n(x)\}$ is complete in that any function of the same kind can be developed upon $\{u_n(x)\}$

$$\psi(x) = \sum_n c_n u_n(x)$$

with

$$c_n = \int u_n^*(x') \psi(x') dx'$$

The completeness of $\{u_n(x)\}$ is expressed as

$$\sum_n u_n^*(x') u_n(x) = \delta(x' - x)$$

4. Summary

The property of completeness is essential to the internal consistency of quantum mechanics, so (following Dirac) we will take it as an axiom (or, more precisely, as a restriction on the class of Hermitian operators that can represent observables/dynamical variables)

Axiom 10 *The eigenfunctions of an observable operator are complete: Any function can be expressed as a linear combination of them.*

III. GENERALIZED STATISTICAL INTERPRETATION AND UNCERTAINTY PRINCIPLE

We have learnt how to calculate the probability that a particle would be found in a particular location

$$|\psi(x)|^2 dx \text{ probability between } x \text{ and } x + dx$$

(*position measurement*), and how to determine the expectation value of any dynamical variables. In the section about generic problem in quantum mechanics, we have also learnt how to find the possible outcomes of an *energy measurement* and their probabilities

$$\hat{H}\psi_n = E_n \psi_n, \quad \langle H \rangle = \sum_{n=1}^{+\infty} |c_n|^2 E_n$$

probability of getting $E_n \rightarrow |c_n|^2$

We are now to state the **generalized statistical interpretation**, which subsumes all of this and enables you to figure out the possible results of any measurement and their probabilities.

A. Generalization

If you measure a physical quantity $F(x, p)$ on a particle in the state $\psi(x)$ which is normalized

$$\int \psi^*(x) \psi(x) dx = 1,$$

you are certain to get one of the eigenvalues of the hermitian operator $\hat{F}(x, -i\hbar d/dx)$.

- If the spectrum of \hat{F} is discrete

$$\hat{F}u_n(x) = f_n u_n(x),$$

and $\{u_n(x)\}$ is complete, the probability of getting the particular eigenvalue f_n associated with orthonormalized eigenfunction $u_n(x)$ is

$$|c_n|^2, \text{ where } c_n = \int u_n^*(x) \psi(x) dx$$

- If the spectrum is continuous,

$$\hat{F}u_z(x) = f(z)u_z(x)$$

with real eigenvalues $f(z)$ and associated Dirac-orthonormalized, complete eigenfunctions $u_z(x)$, the probability of getting a result in the range dz is

$$|c(z)|^2 dz, \text{ where } c(z) = \int u_z^*(x) \psi(x) dx$$

Upon measurement, the wavefunction “collapses” to the corresponding eigenstate.

In both cases, the coefficients are calculated by the Fourier’s trick. One can easily check the normalization of the coefficients and the expectation value. Normalization of ψ now means the total probability (summed over all possible outcomes) has got to be one

$$\sum_n |c_n|^2 = 1$$

and sure enough, this follows from the normalization of the eigenfunctions

$$\begin{aligned} 1 &= \int \psi^*(x) \psi(x) dx \\ &= \int \left(\sum_m c_m^* u_m^*(x) \right) \left(\sum_n c_n u_n(x) \right) dx \\ &= \sum_n |c_n|^2 \end{aligned}$$

Similarly, the expectation value of F should be sum over all possible outcomes of the eigenvalues times the probability of getting that eigenvalue

$$\langle F \rangle = \sum_n |c_n|^2 f_n$$

Indeed,

$$\begin{aligned} \langle F \rangle &= \int \psi^*(x) \hat{F} \psi(x) dx \\ &= \int \left(\sum_m c_m^* u_m^*(x) \right) \hat{F} \left(\sum_n c_n u_n(x) \right) dx \\ &= \sum_m \sum_n c_m^* c_n f_n \int u_m^*(x) u_n(x) dx \\ &= \sum_n |c_n|^2 f_n \end{aligned}$$

So far at least, everything looks consistent.

B. The hypothesis of measurement in quantum mechanics

In summary we have the hypothesis: For a measurement of physical quantity \hat{F} on systems described by a wave function $\psi(x)$

- I The measured result is a statistical one
- II The outcome of a single measurement could be nothing but one of the eigenvalues of \hat{F} , say f_n
- III The probability of obtaining certain eigenvalue f_n is $|c_n|^2 = |\int u_n^*(x)\psi(x) dx|^2$.
- IV Mechanism: the system, having interacted with apparatus, has changed its state from original $\psi(x)$ to one of the eigenstate $u_n(x)$ of \hat{F} with a probability $|c_n|^2$ and gives a measuring outcome f_n .

C. Reproduce of the original statistical interpretation

Can we reproduce, in this language, the original statistical interpretation for position measurements? Sure - it is real overkill, but worth checking.

A measurement of x on a particle in state ψ must return one of the eigenvalues of \hat{x} . But what are the eigenfunctions and eigenvalues of \hat{x} ? Let $g_y(x)$ be the eigenfunction and y the eigenvalue

$$xg_y(x) = yg_y(x)$$

Here y is a fixed number (for any given $g_y(x)$), but x is a continuous variable. What function of x has the property that multiplying it by x is the same as multiplying it by the constant y ? Obviously, it's got to be zero, except at the one point $x = y$; it is nothing but the Dirac delta function

$$g_y(x) = A\delta(x - y)$$

This time the eigenvalue has to be real; the eigenfunctions are not square-integrable, but again admit Dirac orthonormality

$$\begin{aligned} \int_{-\infty}^{+\infty} g_{y'}^*(x) g_y(x) dx &= \int_{-\infty}^{+\infty} |A|^2 \delta(x - y') \delta(x - y) dx \\ &= |A|^2 \delta(y - y') \end{aligned}$$

If we pick $A = 1$, so

$$g_y(x) = \delta(x - y).$$

We see these eigenfunctions are also complete, any wavefunction $\psi(x)$ can be expanded as

$$\psi(x) = \int_{-\infty}^{+\infty} c(y) g_y(x) dy = \int_{-\infty}^{+\infty} c(y) \delta(x - y) dy$$

with

$$c(y) = \psi(y)$$

so the probability of getting a result in the range dy is

$$|c(y)|^2 dy = |\psi(y)|^2 dy$$

not familiar? Write it as $|\psi(x)|^2 dx$, which is precisely the original statistical interpretation.

What about momentum? We have found that the eigenfunctions of the 1D momentum operator

$$\psi_p(x) = \frac{1}{\sqrt{2\pi\hbar}} e^{\frac{i}{\hbar}px}$$

which is also complete and any wavefunction can be expanded as

$$\begin{aligned}\psi(x) &= \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{+\infty} e^{\frac{i}{\hbar}px} \phi(p) dp \\ &= \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{+\infty} e^{\frac{i}{\hbar}px} c(p) dp\end{aligned}$$

where the coefficients is given by the inverse Fourier transformation

$$c(p) = \phi(p) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{+\infty} e^{-\frac{i}{\hbar}px} \psi(x) dx$$

and $|\phi(p)|^2 dp$ is the probability that a measurement of momentum would yield a result in the range dp .

In momentum space, the position operator is $\hat{x} = i\hbar\partial/\partial p$. More generally

$$\langle F(x, p) \rangle = \begin{cases} \int \psi^*(x) \hat{F}(x, -i\hbar\frac{\partial}{\partial x}) \psi(x) dx, & \text{in position space} \\ \int \phi^*(p) \hat{F}(i\hbar\frac{\partial}{\partial p}, p) \phi(p) dp, & \text{in momentum space} \end{cases}$$

In principle you can do all calculations in momentum space just as well (though not always as *easily*) as in position space.

Problem 11 *A particle of mass m is bound in the delta function well*

$$V(x) = -\alpha\delta(x).$$

What is the probability that a measurement of its momentum would yield a value greater than $p_0 = m\alpha/\hbar$? (Griffiths Example 3.4)

D. Commutative operators and common eigenfunctions

In earlier sections we have met some situations where two or more operators may share eigenfunctions. For example, the square of angular momentum \hat{L}^2 and its z component \hat{L}_z do admit complete sets of simultaneous eigenfunctions, the spherical harmonics $Y_{lm}(\theta, \varphi)$. In the problem of hydrogen atom, the Hamiltonian, the magnitude of the angular momentum, and the z component of angular momentum are mutually commutative operators, and one can construct simultaneous eigenfunctions of all three, labeled by their respective eigenvalues n, l, m . In general we have the following theorem:

Theorem 12 *If operators \hat{F} and \hat{G} have common eigenfunctions set $\{\phi_n\}$, they are commutative.*

Proof. Suppose

$$\begin{aligned}\hat{F}\phi_n &= f_n\phi_n \\ \hat{G}\phi_n &= g_n\phi_n\end{aligned}$$

The set $\{\phi_n\}$ is complete, thus any wavefunction can be developed on it

$$\psi = \sum_n a_n \phi_n$$

Acting the commutator of \hat{F} and \hat{G} on ψ gives

$$\begin{aligned}(\hat{F}\hat{G} - \hat{G}\hat{F})\psi &= (\hat{F}\hat{G} - \hat{G}\hat{F}) \sum_n a_n \phi_n \\ &= \sum_n a_n (\hat{F}\hat{G} - \hat{G}\hat{F}) \phi_n \\ &= \sum_n a_n (f_n g_n - g_n f_n) \phi_n = 0\end{aligned}$$

therefore

$$\hat{F}\hat{G} - \hat{G}\hat{F} = [\hat{F}, \hat{G}] = 0$$

■

Let us check the reverse proposition:

Theorem 13 *If \hat{F} and \hat{G} are commutative and the eigenfunctions of \hat{F} in the set $\{\phi_n\}$ are non-degenerate, $\{\phi_n\}$ must be eigenfunctions of \hat{G} .*

Proof. Now the premise is

$$\hat{F}\hat{G} - \hat{G}\hat{F} = 0$$

or more precisely

$$\hat{F}\hat{G}\phi_n - \hat{G}\hat{F}\phi_n = 0$$

which means

$$\hat{F}(\hat{G}\phi_n) = \hat{G}f_n\phi_n = f_n(\hat{G}\phi_n)$$

$(\hat{G}\phi_n)$ is thus an eigenfunction of \hat{F} with eigenvalue f_n . On the other hand, $\{\phi_n\}$ as eigenfunctions of \hat{F} are non-degenerate, ϕ_n and $(\hat{G}\phi_n)$ must be the same states - the difference between them is a constant factor which we shall call it g_n

$$\hat{G}\phi_n = g_n\phi_n$$

It turns out that $\{\phi_n\}$ are also eigenfunctions of \hat{G} , i.e., \hat{F} and \hat{G} have common eigenfunctions. ■

The case is a little complicated if the eigenfunctions of \hat{F} or \hat{G} are degenerate, we can nevertheless arrive at the same conclusion. For example,

$$[\hat{L}^2, \hat{L}_z] = 0$$

the combinations of spherical harmonics Y_{lm}

$$\begin{aligned}\psi_1 &= a_1 Y_{11} + b_1 Y_{10} + c_1 Y_{1,-1} \\ \psi_2 &= a_2 Y_{11} + b_2 Y_{10} + c_2 Y_{1,-1} \\ \psi_3 &= a_3 Y_{11} + b_3 Y_{10} + c_3 Y_{1,-1}\end{aligned}$$

are eigenfunctions of \hat{L}^2

$$\hat{L}^2\psi_i = 1(1+1)\hbar^2\psi_i$$

for $i = 1, 2, 3$ (three-fold degeneracy). Evidently they are not eigenfunctions of \hat{L}_z . It is however always possible to reshuffle them to the eigenfunctions of \hat{L}_z , i.e. Y_{11} can always be expressed as a combination of ψ_i . The general proof will be omitted here.

Problem 14 $\hat{L}_\alpha (\alpha = x, y, z)$ are non commutative, do they share common eigenstates? If so, what are they?

E. A rigorous proof of the uncertainty relation

Proposition 15 *If two Hermitian operators \hat{F} and \hat{G} do not commute, $[\hat{F}, \hat{G}] \neq 0$, they have no common eigenfunction set and their variances obey an uncertainty relation*

$$\sigma_F^2 \cdot \sigma_G^2 \geq \left(\frac{1}{2i} \langle [\hat{F}, \hat{G}] \rangle \right)^2$$

Proof. We define

$$\begin{aligned}\Delta\hat{F} &= \hat{F} - \langle F \rangle, \\ \Delta\hat{G} &= \hat{G} - \langle G \rangle\end{aligned}$$

Evidently $\Delta\hat{F}$ and $\Delta\hat{G}$ are also Hermitian because $\langle F \rangle$ and $\langle G \rangle$ are real. Consider now a positive definite integration

$$I(\xi) = \int \left| \left(\xi \Delta\hat{F} - i \Delta\hat{G} \right) \psi \right|^2 d^3\mathbf{r} \geq 0$$

where ξ is a real constant parameter. We have

$$\begin{aligned}I(\xi) &= \int \left(\xi (\Delta\hat{F}\psi)^* + i (\Delta\hat{G}\psi)^* \right) \left(\xi (\Delta\hat{F}\psi) - i (\Delta\hat{G}\psi) \right) d^3\mathbf{r} \\ &= \xi^2 \int (\Delta\hat{F}\psi)^* (\Delta\hat{F}\psi) d^3\mathbf{r} + \int (\Delta\hat{G}\psi)^* (\Delta\hat{G}\psi) d^3\mathbf{r} \\ &\quad - \xi i \int \left\{ (\Delta\hat{F}\psi)^* (\Delta\hat{G}\psi) - (\Delta\hat{G}\psi)^* (\Delta\hat{F}\psi) \right\} d^3\mathbf{r} \\ &= \xi^2 \int \psi^* (\Delta\hat{F})^2 \psi d^3\mathbf{r} + \int \psi^* (\Delta\hat{G})^2 \psi d^3\mathbf{r} \\ &\quad - \xi i \int \psi^* (\Delta\hat{F}\Delta\hat{G} - \Delta\hat{G}\Delta\hat{F}) \psi d^3\mathbf{r}\end{aligned}$$

Define

$$\begin{aligned}a &= \int \psi^* (\Delta\hat{F})^2 \psi d^3\mathbf{r} = \left\langle (\Delta\hat{F})^2 \right\rangle = \sigma_F^2 \geq 0 \\ c &= \int \psi^* (\Delta\hat{G})^2 \psi d^3\mathbf{r} = \left\langle (\Delta\hat{G})^2 \right\rangle = \sigma_G^2 \geq 0 \\ b &= i \int \psi^* (\Delta\hat{F}\Delta\hat{G} - \Delta\hat{G}\Delta\hat{F}) \psi d^3\mathbf{r} = \left\langle i [\Delta\hat{F}, \Delta\hat{G}] \right\rangle\end{aligned}$$

It is easy to show that the commutator $[\Delta\hat{F}, \Delta\hat{G}]$ is the same as $[\hat{F}, \hat{G}]$. Further more $i [\hat{F}, \hat{G}]$ is Hermitian operator (Check it!) so its expectation value b is real. The integral becomes

$$I(\xi) = a\xi^2 - b\xi + c \geq 0.$$

There is no real root for ξ (the parabola doesn't intersect with the ξ axis), so we must have

$$b^2 - 4ac \leq 0$$

i.e.

$$\sigma_F^2 \cdot \sigma_G^2 \geq \left(\frac{1}{2i} \left\langle [\hat{F}, \hat{G}] \right\rangle \right)^2 \quad (5)$$

or in the notation of standard deviation

$$\begin{aligned}\sigma_F \cdot \sigma_G &\geq \frac{1}{2} \left\langle \left| [\hat{F}, \hat{G}] \right| \right\rangle \\ \Delta F \cdot \Delta G &\geq \frac{1}{2} \left\langle \left| [\hat{F}, \hat{G}] \right| \right\rangle\end{aligned}$$

This is the uncertainty principle in its most general form. (You might think the i makes it trivial – isn't the right side of equation (5) negative? No, for the commutator carries its own factor of i , and the two cancel out.) ■

For example, suppose the first observable is position ($\hat{F} = \hat{x}$), and the second is momentum ($\hat{G} = \hat{p} = -i\hbar d/dx$). We have already known

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

Accordingly

$$\sigma_x^2 \cdot \sigma_p^2 \geq \left(\frac{1}{2i}\hbar\right)^2 = \left(\frac{\hbar}{2}\right)^2$$

or, since standard deviations are by their nature positive,

$$\Delta x \cdot \Delta p \geq \hbar/2$$

That proves the original Heisenberg uncertainty principle, but we now see that it is just one application of a far more general theorem: There will be an "uncertainty principle" for *any pair of observables whose corresponding operators do not commute*. We call them **incompatible observables**. Evidently, incompatible observables do not have shared eigenvectors – at least, they cannot have a complete set of common eigenvectors. Their standard deviations can not be zero at the same time. Later we will see matrices representing incompatible observables cannot be simultaneously diagonalized (that is, they cannot both be brought to diagonal form by the same similarity transformation). On the other hand, compatible observables (whose operators do commute) share a complete set of eigenvectors, and the corresponding matrices can be simultaneously diagonalized.

As an application we can estimate the ground state energy of a simple harmonic oscillator

$$\hat{H} = \frac{\hat{p}^2}{2m} + \frac{1}{2}m\omega^2\hat{x}^2$$

Apparently we have the expectation values for \hat{x}, \hat{p}

$$\langle \hat{x} \rangle = 0, \langle \hat{p} \rangle = 0$$

so

$$\langle (\Delta \hat{p})^2 \rangle = \langle \hat{p}^2 \rangle$$

$$\langle (\Delta \hat{x})^2 \rangle = \langle \hat{x}^2 \rangle$$

and

$$\begin{aligned} \langle H \rangle &= \frac{1}{2m} \langle (\Delta \hat{p})^2 \rangle + \frac{1}{2}m\omega^2 \langle (\Delta \hat{x})^2 \rangle \\ &\geq \frac{1}{2m} \frac{\hbar^2}{4 \langle (\Delta \hat{x})^2 \rangle} + \frac{1}{2}m\omega^2 \langle (\Delta \hat{x})^2 \rangle \end{aligned}$$

The minimum value is determined by

$$\frac{\partial \langle H \rangle}{\partial \langle (\Delta \hat{x})^2 \rangle} = -\frac{\hbar^2}{8m} \frac{1}{\left(\langle (\Delta \hat{x})^2 \rangle\right)^2} + \frac{1}{2}m\omega^2 = 0$$

which gives

$$\left(\langle (\Delta \hat{x})^2 \rangle_{\min}\right)^2 = \frac{\hbar^2}{4m^2\omega^2}$$

so

$$\langle (\Delta \hat{x})^2 \rangle_{\min} = \frac{\hbar}{2m\omega}$$

Inserting back to the expectation value of Hamiltonian we have

$$\langle H \rangle_{\min} = \frac{1}{2m} \frac{\hbar^2}{4} \frac{2m\omega}{\hbar} + \frac{1}{2}m\omega^2 \frac{\hbar}{2m\omega} = \frac{1}{2}\hbar\omega$$

which is exactly the ground state energy we have got in the 1D harmonic potential.

For angular momentum we know their components are generally non-commutative. Fortunately we have

$$\left[\hat{L}^2, \hat{L}_\alpha\right] = 0, \quad \alpha = x, y, z$$

This is why we could have common eigenfunctions of \hat{L}^2 and one of the components, say, \hat{L}_z , i.e., the spherical harmonics $Y_{lm}(\theta, \varphi)$ are chosen to represent the angular momentum states.

Problem 16 Zeng Jinyan, page 133, 11, 12, 14.

IV. LINEAR SPACE - MATHEMATICAL PREPARATION

The purpose of this chapter is to develop the alternative formalism of quantum mechanics – *matrix mechanics*. Quantum theory is based on two constructs: wavefunctions and operators. The state of a system is represented by its wavefunction, observables are represented by operators. Mathematically, wavefunctions satisfy the defining conditions for abstract **vectors**, and operators act on them as **linear transformation**. So the natural language of quantum mechanics is *linear algebra*. I begin with a brief survey of linear algebra. Linear algebra abstracts and generalizes the arithmetic of ordinary vectors, as we encounter them in first-year physics. The generalization is in two directions: (1) We allow the scalars to be *complex*, and (2) we do not restrict ourselves to three dimensions (indeed, we shall be working with vectors that live in spaces of *infinite* dimension).

A. 3D ordinary space

We are already familiar with 3D space. We will rely upon it and do not repeat the sophisticated definition of linear space. 3D space has three basis vectors **i**, **j** and **k** and one has defined scalar products of vectors as

$$\begin{aligned}\mathbf{R} &= a\mathbf{i} + b\mathbf{j} + c\mathbf{k} \\ \mathbf{i} \cdot \mathbf{i} &= 1, \quad \mathbf{i} \cdot \mathbf{R} = a, \quad \text{etc.}\end{aligned}$$

Here we will give the theory with slight formal modification. First, we define the *scalar product* anew. By new stipulations on dot-products we can discriminate the two participants, one as $\langle \mathbf{x}_0 |$, the other as $|\mathbf{x}_0\rangle$

$$\mathbf{i} \cdot \mathbf{i} = 1 \Rightarrow \langle \mathbf{x}_0 | \mathbf{x}_0 \rangle = 1$$

and the vector is now expressed as

$$\begin{aligned}|\mathbf{R}\rangle &= a|\mathbf{x}_0\rangle + b|\mathbf{y}_0\rangle + c|\mathbf{z}_0\rangle \\ a &= \langle \mathbf{x}_0 | \mathbf{R} \rangle, \quad \text{etc.}\end{aligned}$$

The square of *norm* of $|\mathbf{R}\rangle$ is

$$\begin{aligned}\langle \mathbf{R} | \mathbf{R} \rangle &= (a \langle \mathbf{x}_0 | + b \langle \mathbf{y}_0 | + c \langle \mathbf{z}_0 |) (a |\mathbf{x}_0\rangle + b |\mathbf{y}_0\rangle + c |\mathbf{z}_0\rangle) \\ &= a^2 + b^2 + c^2\end{aligned}$$

and similarly, the *inner product* of $|\mathbf{R}_1\rangle$ and $|\mathbf{R}_2\rangle$ is

$$\begin{aligned}\langle \mathbf{R}_1 | \mathbf{R}_2 \rangle &= (a_1 \langle \mathbf{x}_0 | + b_1 \langle \mathbf{y}_0 | + c_1 \langle \mathbf{z}_0 |) (a_2 |\mathbf{x}_0\rangle + b_2 |\mathbf{y}_0\rangle + c_2 |\mathbf{z}_0\rangle) \\ &= a_1 a_2 + b_1 b_2 + c_1 c_2.\end{aligned}$$

The real linear space with the definition of inner product is called the *Euclidean space* or inner product space. Secondly, we use matrix to represent the vectors. For example, the three basis vectors are expressed as column matrices

$$|\mathbf{x}_0\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, |\mathbf{y}_0\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, |\mathbf{z}_0\rangle = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

and the vector

$$|\mathbf{R}\rangle = a|\mathbf{x}_0\rangle + b|\mathbf{y}_0\rangle + c|\mathbf{z}_0\rangle = \begin{pmatrix} a \\ b \\ c \end{pmatrix}$$

Accordingly the left vectors are represented as row matrices

$$\begin{aligned}\langle \mathbf{x}_0 | &= (1, 0, 0), \langle \mathbf{y}_0 | = (0, 1, 0), \langle \mathbf{z}_0 | = (0, 0, 1) \\ \langle \mathbf{R} | &= (a, b, c)\end{aligned}$$

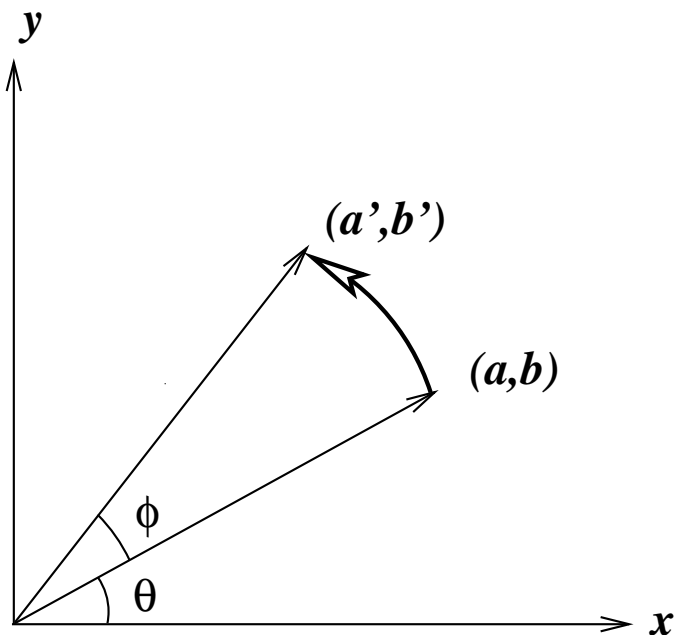


FIG. 1: An example of operation.

thus we have the inner products

$$\langle \mathbf{y}_0 | \mathbf{R} \rangle = (0, 1, 0) \begin{pmatrix} a \\ b \\ c \end{pmatrix} = b$$

$$\langle \mathbf{R}_1 | \mathbf{R}_2 \rangle = (a_1, b_1, c_1) \begin{pmatrix} a_2 \\ b_2 \\ c_2 \end{pmatrix} = a_1 a_2 + b_1 b_2 + c_1 c_2$$

B. Operation onto space

Operation is defined as a regular, ordered and linear deformation of the system with the basis of space untouched. Here is an example of operation: We rotate the system by an angle φ about z axis, followed by a compression along z axis by a factor α . The operation changed vector $|\mathbf{R}\rangle$ into $|\mathbf{T}\rangle$. Suppose

$$|\mathbf{R}\rangle = \begin{pmatrix} a \\ b \\ c \end{pmatrix}, |\mathbf{T}\rangle = \begin{pmatrix} a' \\ b' \\ c' \end{pmatrix}$$

From Figure (1) we know

$$a = l \cos \theta, \quad b = l \sin \theta$$

and

$$\begin{aligned} a' &= l \cos(\theta + \varphi) = l \cos \theta \cos \varphi - l \sin \theta \sin \varphi = a \cos \varphi - b \sin \varphi \\ b' &= l \sin(\theta + \varphi) = l \cos \theta \sin \varphi + l \sin \theta \cos \varphi = a \sin \varphi + b \cos \varphi \\ c' &= \alpha c \end{aligned}$$

Write it in matrix form

$$\begin{pmatrix} a' \\ b' \\ c' \end{pmatrix} = \begin{pmatrix} \cos \varphi & -\sin \varphi & 0 \\ \sin \varphi & \cos \varphi & 0 \\ 0 & 0 & \alpha \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix}$$

In general the deformation can be expressed as a square matrix

$$|\mathbf{T}\rangle = F |\mathbf{R}\rangle, \quad F = \begin{pmatrix} F_{xx} & F_{xy} & F_{xz} \\ F_{yx} & F_{yy} & F_{yz} \\ F_{zx} & F_{zy} & F_{zz} \end{pmatrix}.$$

C. Basis transformation

The same linear space may be expressed by two different orthogonal coordinates/basis sets. A change of basis is mostly described by a matrix by which the new basis vectors are given as linear combinations of the old basis vectors. Suppose we have two basis sets

$$\begin{aligned} \text{basis 1 : } & |\mathbf{x}_0\rangle, |\mathbf{y}_0\rangle, |\mathbf{z}_0\rangle \\ \text{basis 2 : } & |\alpha_0\rangle, |\beta_0\rangle, |\gamma_0\rangle \end{aligned}$$

A vector can be developed upon both

$$\begin{aligned} |\mathbf{R}\rangle_x &= a |\mathbf{x}_0\rangle + b |\mathbf{y}_0\rangle + c |\mathbf{z}_0\rangle = \begin{pmatrix} a \\ b \\ c \end{pmatrix} \\ |\mathbf{R}\rangle_\alpha &= \rho |\alpha_0\rangle + \sigma |\beta_0\rangle + \tau |\gamma_0\rangle = \begin{pmatrix} \rho \\ \sigma \\ \tau \end{pmatrix} \end{aligned}$$

and one must have

$$a |\mathbf{x}_0\rangle + b |\mathbf{y}_0\rangle + c |\mathbf{z}_0\rangle = \rho |\alpha_0\rangle + \sigma |\beta_0\rangle + \tau |\gamma_0\rangle$$

that is the change of basis will not affect the vector. Multiplying $\langle\alpha_0|$ from the left on the above equation we have

$$a \langle\alpha_0|\mathbf{x}_0\rangle + b \langle\alpha_0|\mathbf{y}_0\rangle + c \langle\alpha_0|\mathbf{z}_0\rangle = \rho$$

similarly

$$\begin{aligned} a \langle\beta_0|\mathbf{x}_0\rangle + b \langle\beta_0|\mathbf{y}_0\rangle + c \langle\beta_0|\mathbf{z}_0\rangle &= \sigma \\ a \langle\gamma_0|\mathbf{x}_0\rangle + b \langle\gamma_0|\mathbf{y}_0\rangle + c \langle\gamma_0|\mathbf{z}_0\rangle &= \tau \end{aligned}$$

Define

$$S_{\alpha x} = \langle\alpha_0|\mathbf{x}_0\rangle, S_{\alpha y} = \langle\alpha_0|\mathbf{y}_0\rangle, S_{\alpha z} = \langle\alpha_0|\mathbf{z}_0\rangle, \text{ etc.}$$

we get

$$\begin{pmatrix} \rho \\ \sigma \\ \tau \end{pmatrix} = \begin{pmatrix} S_{\alpha x} & S_{\alpha y} & S_{\alpha z} \\ S_{\beta x} & S_{\beta y} & S_{\beta z} \\ S_{\gamma x} & S_{\gamma y} & S_{\gamma z} \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix}$$

or

$$|\mathbf{R}\rangle_\alpha = S |\mathbf{R}\rangle_x$$

where S is the matrix for the basis transformation. The norm of a vector should not depend on the choice of the basis sets, which means

$$\alpha \langle\mathbf{R}|\mathbf{R}\rangle_\alpha = {}_x \langle\mathbf{R}|\mathbf{R}\rangle_x$$

or

$$(\rho, \sigma, \tau) \begin{pmatrix} \rho \\ \sigma \\ \tau \end{pmatrix} = (a, b, c) \begin{pmatrix} a \\ b \\ c \end{pmatrix}$$

So

$$\left[S \begin{pmatrix} a \\ b \\ c \end{pmatrix} \right]^T S \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} a \\ b \\ c \end{pmatrix}^T \begin{pmatrix} a \\ b \\ c \end{pmatrix}$$

According to $(AB)^T = B^T A^T$ we have

$$(a, b, c) S^T S \begin{pmatrix} a \\ b \\ c \end{pmatrix} = (a, b, c) \begin{pmatrix} a \\ b \\ c \end{pmatrix}$$

The matrix S is fixed by the transformation, while the vector (a, b, c) is arbitrary. We thus have

$$S^T S = I$$

i.e., S is an *orthogonal* matrix.

D. Transformation of operations

Operation F changes vector $|\mathbf{R}\rangle$ into $|\mathbf{T}\rangle = F|\mathbf{R}\rangle$, and this relation should be maintained under a basis transformation. In basis 1, this change is expressed by

$$|\mathbf{T}\rangle_x = F_x |\mathbf{R}\rangle_x \quad (6)$$

or more explicitly

$$\begin{pmatrix} a' \\ b' \\ c' \end{pmatrix} = \begin{pmatrix} F_{xx} & F_{xy} & F_{xz} \\ F_{yx} & F_{yy} & F_{yz} \\ F_{zx} & F_{zy} & F_{zz} \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix}$$

In basis 2, it should read

$$|\mathbf{T}\rangle_\alpha = F_\alpha |\mathbf{R}\rangle_\alpha$$

or

$$\begin{pmatrix} \rho' \\ \sigma' \\ \tau' \end{pmatrix} = \begin{pmatrix} F_{\alpha\alpha} & F_{\alpha\beta} & F_{\alpha\gamma} \\ F_{\beta\alpha} & F_{\beta\beta} & F_{\beta\gamma} \\ F_{\gamma\alpha} & F_{\gamma\beta} & F_{\gamma\gamma} \end{pmatrix} \begin{pmatrix} \rho \\ \sigma \\ \tau \end{pmatrix}$$

Acting the matrix for the basis transformation S onto both sides of equation (6), we have

$$\begin{aligned} S |\mathbf{T}\rangle_x &= S F_x S^T S |\mathbf{R}\rangle_x = F_\alpha |\mathbf{R}\rangle_\alpha \\ &= |\mathbf{T}\rangle_\alpha \end{aligned}$$

we immediately know

$$F_\alpha = S F_x S^T$$

or

$$\begin{pmatrix} F_{\alpha\alpha} & F_{\alpha\beta} & F_{\alpha\gamma} \\ F_{\beta\alpha} & F_{\beta\beta} & F_{\beta\gamma} \\ F_{\gamma\alpha} & F_{\gamma\beta} & F_{\gamma\gamma} \end{pmatrix} = \begin{pmatrix} S_{\alpha x} & S_{\alpha y} & S_{\alpha z} \\ S_{\beta x} & S_{\beta y} & S_{\beta z} \\ S_{\gamma x} & S_{\gamma y} & S_{\gamma z} \end{pmatrix} \begin{pmatrix} F_{xx} & F_{xy} & F_{xz} \\ F_{yx} & F_{yy} & F_{yz} \\ F_{zx} & F_{zy} & F_{zz} \end{pmatrix} \begin{pmatrix} S_{\alpha x} & S_{\beta z} & S_{\gamma x} \\ S_{\alpha y} & S_{\beta y} & S_{\gamma y} \\ S_{\alpha z} & S_{\beta z} & S_{\gamma z} \end{pmatrix}$$

V. MATRIX FORMULATION OF QUANTUM MECHANICS - REPRESENTATION THEORY

A. The idea of representation

Till now, we are used to express the state of micro-system by wave functions $\psi(\mathbf{r})$, and to express the operators by some mathematical terms $\hat{p}_x = -i\hbar\frac{\partial}{\partial x}$, etc. Is this the unique way to express the quantum phenomena, quantum rules and laws? In other words, whether all our knowledge of quantum mechanics could only be expressed through wave functions $\psi(\mathbf{r})$'s and operators like $\hat{p}_x = -i\hbar\frac{\partial}{\partial x}$?

The answer is no. $\psi(\mathbf{r})$'s and operators like $\hat{p}_x = -i\hbar\frac{\partial}{\partial x}$ are only one way (style) to express (represent) the quantum theory. They are called *coordinate representation* (*x-representation*) of quantum mechanics. We have other ways to represent quantum theory.

We have already familiarized with the Fourier transformation of $\psi(\mathbf{r})$, i.e., $\phi(\mathbf{p})$, which can describe the state precisely just as $\psi(\mathbf{r})$ do. In the $\phi(\mathbf{p})$ frame, operators are expressed by $\hat{p}_x = p_x$, $\hat{x} = i\hbar\frac{\partial}{\partial p_x}$. $\phi(\mathbf{p})$ and operators like $\hat{x} = i\hbar\frac{\partial}{\partial p_x}$ could also express the quantum theory effectively and completely just as $\psi(\mathbf{r})$ and $\hat{p}_x = -i\hbar\frac{\partial}{\partial x}$ do (perhaps not so conveniently). They are called *momentum representation* (*p-representation*) of quantum theory.

There are other representations that we are going to discuss in this section.

B. Representation of state

Let \hat{Q} be an Hermitian operator, its eigenfunctions $\{u_n(x)\}$ is a complete orthonormal set. Quantum theory can be expressed through $\{u_n(x)\}$ set and this formulation is called *Q-representation*.

Premise: $\{u_n(x)\}$ set is known and forms a complete orthonormal set. Any wavefunction $\psi(x)$ can be expanded upon $\{u_n(x)\}$ set

$$\psi(x) = \sum_n c_n u_n(x)$$

with

$$c_n = \int u_n^*(x') \psi(x') dx'$$

State function $\psi(x)$ and the set of coefficients $\{c_n\}$ are uniquely connected. That means, given the wave function $\psi(x)$, the coefficients $\{c_n\}$ are completely known. On the other hand, knowing the coefficients $\{c_n\}$, the state function is uniquely determined.

*The coefficients $\{c_n\}$ uniquely determine the state of the system just as the wavefunction $\psi(x)$ does. The coefficients $\{c_n\}$ is called *Q-representation of the state*.*

Formally the *Q-representation* is expressed as a "column matrix", i.e. the *N*-tuple of the component of a vector $\{c_n\}$, with respect to a specified orthonormal basis

$$\psi = \begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ \vdots \end{pmatrix}$$

the Hermitian conjugate of ψ is denoted by ψ^\dagger

$$\psi^\dagger = \tilde{\psi}^* = (c_1^*, c_2^*, c_3^*, \dots)$$

and normalization of $\psi(x)$

$$\int \psi^*(x) \psi(x) dx = \int \sum_m c_m^* u_m^*(x) \sum_n c_n u_n(x) dx = \sum_n c_n^* c_n = 1$$

is expressed in *Q-representation*

$$\psi^\dagger \psi = (c_1^*, c_2^*, c_3^*, \dots) \begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ \vdots \end{pmatrix} = 1.$$

Similarly inner product of ψ and another vector ϕ

$$\phi = \sum_n a_n u_n(x) = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \end{pmatrix},$$

generalizing the dot product in 3D, is the complex number

$$\int \psi^* \phi dx = \psi^\dagger \phi = (c_1^*, c_2^*, c_3^*, \dots) \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \end{pmatrix} = \sum_n c_n^* a_n.$$

The set of all square integrable functions constitutes a vector space, which physicists call it Hilbert space after David Hilbert, who studied linear spaces in infinite dimensions. In quantum mechanics, then,

Wave functions live in Hilbert Space.

C. Representation of operators

In Q -representation, the states are expressed by column (row) matrix, how would an operators \hat{F} be expressed in Q -representation?

The action of an operators \hat{F} upon a certain wave function ψ changes it into another function $\phi(x)$.

$$\phi(x) = \hat{F}\psi(x).$$

Both of them can be expanded upon the complete set $\{u_n(x)\}$ in Q -representation

$$\phi(x) = \sum_n b_n u_n(x), \quad \psi(x) = \sum_n a_n u_n(x)$$

we thus have

$$\sum_n b_n u_n(x) = \hat{F} \sum_n a_n u_n(x)$$

Multiplying $u_m^*(x)$ and integrating

$$\int_{-\infty}^{+\infty} u_m^*(x) \left\{ \sum_n b_n u_n(x) = \hat{F} \sum_n a_n u_n(x) \right\} dx$$

by means of the orthogonal properties of the eigenfunctions, we get

$$b_m = \sum_n \int_{-\infty}^{\infty} u_m^*(x) \hat{F} u_n(x) a_n dx = \sum_n F_{mn} a_n.$$

In matrix form, it is

$$\begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ \dots \end{pmatrix} = \begin{pmatrix} F_{11} & F_{12} & F_{13} & \dots \\ F_{21} & F_{22} & F_{23} & \dots \\ F_{31} & F_{32} & F_{33} & \dots \\ \dots & \dots & \dots & \dots \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ \dots \end{pmatrix}$$

or

$$\phi = F\psi.$$

The operator \hat{F} is expressed by a square matrix in Q -representation

$$F = \begin{pmatrix} F_{11} & F_{12} & F_{13} & \cdots \\ F_{21} & F_{22} & F_{23} & \cdots \\ F_{31} & F_{32} & F_{33} & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{pmatrix}$$

with the elements F_{mn} 's defined by following integration

$$F_{mn} = \int u_m^*(x) \hat{F} u_n(x) dx.$$

D. Property of the matrix of a Hermitian operator

The matrix element F_{mn} and its conjugate

$$F_{mn}^* = \int u_m (\hat{F} u_n)^* dx = \int (\hat{F} u_n)^* u_m dx = \int u_n^* \hat{F} u_m dx$$

satisfy

$$F_{mn}^* = F_{nm}$$

Thus the matrix F representing a Hermitian operator \hat{F} is a Hermitian matrix

$$F^\dagger = F.$$

Indeed

$$\begin{aligned} F^\dagger &= \begin{pmatrix} F_{11} & F_{12} & F_{13} & \cdots \\ F_{21} & F_{22} & F_{23} & \cdots \\ F_{31} & F_{32} & F_{33} & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{pmatrix}^\dagger = \begin{pmatrix} F_{11} & F_{21} & F_{31} & \cdots \\ F_{12} & F_{22} & F_{32} & \cdots \\ F_{13} & F_{23} & F_{33} & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{pmatrix}^* \\ &= \begin{pmatrix} F_{11}^* & F_{12}^* & F_{13}^* & \cdots \\ F_{21}^* & F_{22}^* & F_{23}^* & \cdots \\ F_{31}^* & F_{32}^* & F_{33}^* & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{pmatrix} = \begin{pmatrix} F_{11} & F_{12} & F_{13} & \cdots \\ F_{21} & F_{22} & F_{23} & \cdots \\ F_{31} & F_{32} & F_{33} & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{pmatrix} = F \end{aligned}$$

Example 17 As an example, let us try to find the matrix form of angular momentum operator \hat{L}_y

$$\hat{L}_y = -i\hbar \cos \varphi \frac{\partial}{\partial \theta} + i\hbar \cot \theta \sin \varphi \frac{\partial}{\partial \varphi}$$

We choose the common eigenfunctions of (\hat{L}^2, \hat{L}_z) for $l = 1$ as our basis and denote them as (in this case we have $m = 1, 0, -1$)

$$u_1 : Y_{1,1} = -\sqrt{\frac{3}{8\pi}} \sin \theta e^{i\varphi}$$

$$u_2 : Y_{1,0} = \sqrt{\frac{3}{4\pi}} \cos \theta$$

$$u_3 : Y_{1,-1} = \sqrt{\frac{3}{8\pi}} \sin \theta e^{-i\varphi}$$

The matrix of \hat{L}_y is accordingly a 3×3 matrix. First let us operate \hat{L}_y upon $Y_{1,1}$

$$\begin{aligned} \hat{L}_y u_1 &= \hat{L}_y Y_{1,1} = i\hbar \sqrt{\frac{3}{8\pi}} \cos \varphi \cos \theta e^{i\varphi} - i\hbar \sqrt{\frac{3}{8\pi}} \cot \theta \sin \varphi \sin \theta e^{i\varphi} \\ &= i\hbar \sqrt{\frac{3}{8\pi}} \cos \theta (\cos \varphi - i \sin \varphi) e^{i\varphi} \\ &= \frac{i\hbar}{\sqrt{2}} \sqrt{\frac{3}{4\pi}} \cos \theta = \frac{i\hbar}{\sqrt{2}} Y_{1,0} \end{aligned}$$

we have

$$\begin{aligned}\iint u_1^* \hat{L}_y u_1 d\Omega &= \iint Y_{1,1}^* \hat{L}_y Y_{1,1} (\sin \theta d\theta d\varphi) = 0 \\ \iint u_2^* \hat{L}_y u_1 d\Omega &= \iint Y_{1,0}^* \hat{L}_y Y_{1,1} (\sin \theta d\theta d\varphi) = \frac{i\hbar}{\sqrt{2}} \\ \iint u_3^* \hat{L}_y u_1 d\Omega &= \iint Y_{1,-1}^* \hat{L}_y Y_{1,1} (\sin \theta d\theta d\varphi) = 0\end{aligned}$$

The result of $\hat{L}_y Y_{1,0}$ gives

$$\begin{aligned}\hat{L}_y u_2 &= \hat{L}_y Y_{1,0} = -i\hbar \cos \varphi \sqrt{\frac{3}{4\pi}} (-\sin \theta) = i\hbar \sqrt{\frac{3}{4\pi}} \sin \theta \cos \varphi \\ &= i\hbar \sqrt{\frac{3}{4\pi}} \sin \theta \left(\frac{e^{i\varphi} + e^{-i\varphi}}{2} \right) \\ &= \frac{i\hbar}{\sqrt{2}} \sqrt{\frac{3}{8\pi}} \sin \theta e^{i\varphi} + \frac{i\hbar}{\sqrt{2}} \sqrt{\frac{3}{8\pi}} \sin \theta e^{-i\varphi} \\ &= -\frac{i\hbar}{\sqrt{2}} Y_{1,1} + \frac{i\hbar}{\sqrt{2}} Y_{1,-1}.\end{aligned}$$

and

$$\begin{aligned}\iint u_1^* \hat{L}_y u_2 d\Omega &= \iint Y_{1,1}^* \hat{L}_y Y_{1,0} (\sin \theta d\theta d\varphi) = -\frac{i\hbar}{\sqrt{2}}, \\ \iint u_2^* \hat{L}_y u_2 d\Omega &= \iint Y_{1,0}^* \hat{L}_y Y_{1,0} (\sin \theta d\theta d\varphi) = 0 \\ \iint u_3^* \hat{L}_y u_2 d\Omega &= \iint Y_{1,-1}^* \hat{L}_y Y_{1,0} (\sin \theta d\theta d\varphi) = \frac{i\hbar}{\sqrt{2}},\end{aligned}$$

In summary the matrix form of \hat{L}_y is

$$L_y = \begin{pmatrix} 0 & -\frac{i\hbar}{\sqrt{2}} & 0 \\ \frac{i\hbar}{\sqrt{2}} & 0 & -\frac{i\hbar}{\sqrt{2}} \\ 0 & \frac{i\hbar}{\sqrt{2}} & 0 \end{pmatrix}.$$

Operator in its own-representation takes a diagonal form with diagonal elements its eigenvalues. Suppose

$$\hat{Q}u_n(x) = q_n u_n(x)$$

then the matrix element

$$\begin{aligned}Q_{mn} &= \int u_m^*(x) \hat{Q}u_n(x) dx \\ &= q_n \int u_m^*(x) u_n(x) dx \\ &= q_n \delta_{mn}\end{aligned}$$

which means

$$Q = \begin{pmatrix} q_1 & 0 & 0 & \dots \\ 0 & q_2 & 0 & \dots \\ 0 & 0 & q_3 & \dots \\ \dots & \dots & \dots & \dots \end{pmatrix}.$$

In the above example, we find

$$L_z = \begin{pmatrix} \hbar & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -\hbar \end{pmatrix}, L^2 = \begin{pmatrix} 2\hbar^2 & 0 & 0 \\ 0 & 2\hbar^2 & 0 \\ 0 & 0 & 2\hbar^2 \end{pmatrix}.$$

E. Matrix formulation of rules of quantum mechanics

Here we state the rules of quantum mechanics in matrix formulation. First, in Q -representation

$$\psi = \sum_n c_n u_n(x)$$

the expectation value of observable F

$$\begin{aligned} \langle F \rangle &= \int \psi^* \hat{F} \psi dx \\ &= \int \sum_m c_m^* u_m^*(x) \hat{F} \sum_n c_n u_n(x) dx \\ &= \sum_{mn} c_m^* F_{mn} c_n \end{aligned}$$

This is expressed in matrix form

$$(c_1^*, c_2^*, c_3^*, \dots) \begin{pmatrix} F_{11} & F_{12} & F_{13} & \dots \\ F_{21} & F_{22} & F_{23} & \dots \\ F_{31} & F_{32} & F_{33} & \dots \\ \dots & \dots & \dots & \dots \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ \dots \end{pmatrix} = \psi^\dagger F \psi,$$

In short the expectation value of an observable F is

$$\langle F \rangle = \int \psi^* \hat{F} \psi dx = \psi^\dagger F \psi.$$

Secondly, we may formulate the eigenvalue equation of operator \hat{F} in the language of matrix. Starting from the

$$\hat{F} \psi = \lambda \psi$$

we multiply $u_m^*(x)$ and do the integration

$$\int u_m^*(x) \left\{ \hat{F} \sum_n c_n u_n(x) = \lambda \sum_n c_n u_n(x) \right\} dx$$

immediately we have

$$\sum_n c_n \int u_m^* \hat{F} u_n dx = \lambda c_m$$

or

$$\sum_n F_{mn} c_n = \lambda c_m$$

In matrix form it reads

$$\begin{pmatrix} F_{11} & F_{12} & F_{13} & \dots \\ F_{21} & F_{22} & F_{23} & \dots \\ F_{31} & F_{32} & F_{33} & \dots \\ \dots & \dots & \dots & \dots \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ \dots \end{pmatrix} = \lambda \begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ \dots \end{pmatrix}.$$

Thus, eigenequation of operator \hat{F} is now expressed to find the eigenvalues and eigenvectors of an Hermitian matrix.

An example that finds large amount of utilization is the case that F is a square matrix of finite rows and columns

$$\begin{pmatrix} F_{11} & F_{12} & \dots & F_{1k} \\ F_{21} & F_{22} & \dots & F_{2k} \\ \dots & \dots & \dots & \dots \\ F_{k1} & F_{k2} & \dots & F_{kk} \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ \dots \\ c_k \end{pmatrix} = \lambda \begin{pmatrix} c_1 \\ c_2 \\ \dots \\ c_k \end{pmatrix},$$

The condition that nontrivial solutions exist requires that

$$\begin{vmatrix} F_{11} - \lambda & F_{12} & \cdots & F_{1k} \\ F_{21} & F_{22} - \lambda & \cdots & F_{2k} \\ \cdots & \cdots & \cdots & \cdots \\ F_{k1} & F_{k2} & \cdots & F_{kk} - \lambda \end{vmatrix} = 0,$$

which gives an algebraic equation of λ of k -th power, and by solving it, we can find the eigenvalues of F to be the roots of the algebraic equation

$$\lambda_1, \lambda_2, \cdots \lambda_k.$$

VI. UNITARY TRANSFORMATION BETWEEN TWO REPRESENTATIONS

In the section Linear Space, we introduced the basic idea of linear algebra, including basis, operation, and the transformation of them. The *vectors* transform as

$$|\mathbf{R}\rangle_\alpha = S |\mathbf{R}\rangle_x$$

while the *operations* change according to

$$F_\alpha = S F_x S^T.$$

In quantum mechanics, the eigenfunctions of Hermitian operators play the role of basis in the function space. It is naturally to consider the generalization of 3D vectors in two directions as we mentioned earlier: (1) We allow the components of the vector to be *complex* in order to represent the wavefunctions. This turns the transformation matrix S , which is orthogonal for 3D (real) vector space, into a unitary matrix. The transformation is therefore a **unitary** one, and (2) instead of working in 3D, our vectors live in spaces of *infinite* dimension. We will show in this section a similar transformation of *eigenvectors* and *operators* arises naturally in quantum mechanics.

A. Two sets of basis and their relations

Suppose we have two Hermitian operators \hat{A} and \hat{B} . Both of them have their own complete orthonormal eigenfunction sets

$$\begin{array}{lll} \hat{A}, & \{u_n(x)\} & u_1(x), u_2(x), \cdots u_n(x) \cdots \\ \hat{B}, & \{\phi_\nu(x)\} & \phi_1(x), \phi_2(x), \cdots \phi_\nu(x) \cdots \end{array}$$

and form two independent representations, \hat{A} representation and \hat{B} representation. The relation between them can be established through the relation of the two basis sets $\{u_n(x)\}$ and $\{\phi_\nu(x)\}$.

Here we use a slightly different method to introduce the transformation matrix (actually it is more often used in linear algebra textbook), that is, we obtain the matrix element from the transformation of *basis vectors*. The new basis vectors in 3D are given as linear combinations of the old ones. Likely, the *eigenfunction* in the set $\{u_n(x)\}$ can be expanded upon $\{\phi_\nu(x)\}$

$$u_n(x) = \sum_\nu S_{\nu n} \phi_\nu(x) \tag{7}$$

where the coefficients are inner product of two of the basis vectors

$$\int \phi_\mu^*(x) u_n(x) dx = \int \phi_\mu^*(x) \sum_\nu S_{\nu n} \phi_\nu(x) dx = S_{\mu n}$$

The elements of the transformation matrix are thus

$$S_{\nu n} = (\phi_\nu, u_n)$$

This matrix changes the basis vector in representation \hat{A} to representation \hat{B} . The orthonormality of the basis vectors gives

$$\begin{aligned} \int u_m^*(x)u_n(x)dx &= \int \sum_{\mu} S_{\mu m}^* \phi_{\mu}^*(x) \sum_{\nu} S_{\nu n} \phi_{\nu}(x) dx = \sum_{\mu} \sum_{\nu} S_{\mu m}^* S_{\nu n} \delta_{\mu\nu} \\ &= \sum_{\mu} S_{\mu m}^* S_{\mu n} = \sum_{\mu} (S^{\dagger})_{m\mu} S_{\mu n} = (S^{\dagger}S)_{mn} = \delta_{mn} \end{aligned}$$

i.e.

$$S^{\dagger}S = I.$$

Here S^{\dagger} is the hermitian conjugate of S and I is unit matrix. It is too early to say the matrix S is unitary - we need prove $SS^{\dagger} = I$ (one may think it trivial, however the situation does occur that $SS^{\dagger} \neq S^{\dagger}S$). For this purpose we calculate

$$\begin{aligned} \sum_n S_{\mu n} S_{\nu n}^* &= \sum_n S_{\mu n} (S^{\dagger})_{n\nu} = (SS^{\dagger})_{\mu\nu} \\ &= \sum_n \int \phi_{\mu}^*(x)u_n(x)dx \int \phi_{\nu}(x')u_n^*(x')dx' \end{aligned} \quad (8)$$

On the other hand, the eigenfunction in the set $\{\phi_{\nu}(x)\}$ can be expanded upon $\{u_n(x)\}$

$$\phi_{\nu}(x') = \sum_m c_m u_m(x')$$

with

$$c_m = \int u_m^*(x')\phi_{\nu}(x')dx' = (u_m, \phi_{\nu}) = S_{\nu m}^* = (S^{\dagger})_{m\nu}$$

Inserting back to the right side of (8), we have

$$\begin{aligned} (SS^{\dagger})_{\mu\nu} &= \sum_n \int \phi_{\mu}^*(x)u_n(x)dx \int \sum_m c_m u_m(x')u_n^*(x')dx' \\ &= \sum_n \int \phi_{\mu}^*(x)u_n(x)dx \sum_m c_m \delta_{mn} \\ &= \int \phi_{\mu}^*(x) \sum_n c_n u_n(x)dx = \int \phi_{\mu}^*(x)\phi_{\nu}(x)dx = \delta_{\mu\nu} \end{aligned}$$

i.e.

$$SS^{\dagger} = I.$$

The inverse matrix of S is thus

$$S^{-1} = S^{\dagger}$$

The matrix satisfying this is called *unitary matrix*, the corresponding transformation is called *unitary transformation*. We conclude that the transformation from one representation to another is unitary.

B. Transformation of state from representation \hat{A} to \hat{B}

We now check the transformation of the state between two representations. An arbitrary state function $\psi(x)$ has its \hat{A} representation expression as follows

$$\psi(x) = \sum_n a_n u_n(x) \left. \begin{array}{l} \\ \\ \\ \end{array} \right\} (A)$$

In representation \hat{B} , the same state is depicted as

$$\psi(x) = \sum_{\nu} b_{\nu} \phi_{\nu}(x) \left. \vphantom{\sum_{\nu}} \right\} (B)$$

$$\psi = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \end{pmatrix}$$

Now, starting from representation \hat{A} , we want to know the expression of ψ in representation \hat{B} . By means of equation (7)

$$\begin{aligned} \psi(x) &= \sum_n a_n u_n(x) = \sum_n a_n \left(\sum_{\nu} S_{\nu n} \phi_{\nu}(x) \right) \\ &= \sum_{\nu} \left(\sum_n (S_{\nu n} a_n) \right) \phi_{\nu}(x) = \sum_{\nu} b_{\nu} \phi_{\nu}(x) \end{aligned}$$

which gives the relation of representation of states in different representations

$$b_{\nu} = \sum_n S_{\nu n} a_n$$

$$\psi_B = S \psi_A$$

or, in the matrix form

$$\begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ \dots \end{pmatrix} = \begin{pmatrix} S_{11} & S_{12} & S_{13} & \dots \\ S_{21} & S_{22} & S_{23} & \dots \\ S_{31} & S_{32} & S_{33} & \dots \\ \dots & \dots & \dots & \dots \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ \dots \end{pmatrix}$$

Inversely, starting from representation \hat{B} , we know the expression of ψ in representation \hat{A}

$$a_n = \sum_{\nu} S_{n\nu}^{\dagger} b_{\nu}$$

$$\psi_A = S^{\dagger} \psi_B$$

or

$$\begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ \dots \end{pmatrix} = \begin{pmatrix} S_{11}^* & S_{21}^* & S_{31}^* & \dots \\ S_{12}^* & S_{22}^* & S_{32}^* & \dots \\ S_{13}^* & S_{23}^* & S_{33}^* & \dots \\ \dots & \dots & \dots & \dots \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ \dots \end{pmatrix}$$

We notice that the matrix elements are now complex (wavefunction is in nature complex) and the matrix itself is unitary (in ordinary 3D vector space it is real and orthonormal).

C. Transformation of operators

In representation \hat{A} , an operator \hat{F} is expressed by a square matrix

$$F_A = \begin{pmatrix} F_{11} & F_{12} & F_{13} & \dots \\ F_{21} & F_{22} & F_{23} & \dots \\ F_{31} & F_{32} & F_{33} & \dots \\ \dots & \dots & \dots & \dots \end{pmatrix}$$

with the elements

$$F_{mn} = \int u_m^*(x) \hat{F} u_n(x) dx.$$

In representation \hat{B} the matrix is

$$F_B = \begin{pmatrix} F'_{11} & F'_{12} & F'_{13} & \cdots \\ F'_{21} & F'_{22} & F'_{23} & \cdots \\ F'_{31} & F'_{32} & F'_{33} & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{pmatrix}$$

and the elements are defined as

$$F'_{\mu\nu} = \int \phi_\mu^*(x) \hat{F} \phi_\nu(x) dx.$$

Inserting back the expansion

$$\begin{aligned} \phi_\nu(x) &= \sum_n (S^\dagger)_{n\nu} u_n(x) = \sum_n S_{\nu n}^* u_n(x) \\ \phi_\mu^*(x) &= \sum_m S_{\mu m} u_m^*(x) \end{aligned}$$

we have

$$\begin{aligned} F'_{\mu\nu} &= \int \sum_m S_{\mu m} u_m^*(x) \hat{F} \sum_n S_{\nu n}^* u_n(x) dx \\ &= \sum_{mn} S_{\mu m} \left(\int u_m^*(x) \hat{F} u_n(x) dx \right) S_{\nu n}^* \\ &= \sum_{mn} S_{\mu m} F_{mn} (S^\dagger)_{n\nu} \end{aligned}$$

The operator then changes according to

$$F_B = S F_A S^\dagger = S F_A S^{-1}$$

which is reminiscent of the transformation of operation in 3D space, in that case the transformation matrix S is orthonormal. We immediately have the transformation from representation \hat{B} to representation \hat{A}

$$F_A = S^\dagger F_B S = S^{-1} F_B S$$

Example 18 Suppose we have a microsystem with orbital angular momentum with eigenvalue $1(1+1)\hbar^2$. We choose the common eigenfunctions of \hat{L}^2, \hat{L}_z as representation \hat{A}

$$\begin{aligned} u_1 &= Y_{1,1}(\theta, \varphi) \\ u_2 &= Y_{1,0}(\theta, \varphi) \\ u_3 &= Y_{1,-1}(\theta, \varphi) \end{aligned}$$

The operator \hat{L}_y in this representation is a 3×3 matrix

$$L_y = \begin{pmatrix} 0 & -\frac{i\hbar}{\sqrt{2}} & 0 \\ \frac{i\hbar}{\sqrt{2}} & 0 & -\frac{i\hbar}{\sqrt{2}} \\ 0 & \frac{i\hbar}{\sqrt{2}} & 0 \end{pmatrix}$$

The representation \hat{B} is chosen as the common eigenfunctions of \hat{L}^2, \hat{L}_y . We now try to find the matrix form of \hat{L}_y in representation \hat{B} . Obviously, we should have already know the result - \hat{L}_y in its own representation is a diagonal matrix with the diagonal elements its eigenvalues.

Solution 19 Step 1. Solve the eigenvalue problem of \hat{L}_y in \hat{A} representation. Eigenvalues of \hat{L}_y are determined by

$$\begin{vmatrix} 0 - \lambda & -\frac{i\hbar}{\sqrt{2}} & 0 \\ \frac{i\hbar}{\sqrt{2}} & 0 - \lambda & -\frac{i\hbar}{\sqrt{2}} \\ 0 & \frac{i\hbar}{\sqrt{2}} & 0 - \lambda \end{vmatrix} = 0$$

which gives

$$\begin{aligned} -\lambda^3 - \frac{i^2\hbar^2}{2}\lambda - \frac{i^2\hbar^2}{2}\lambda &= 0 \\ \lambda^3 - \hbar^2\lambda &= 0 \end{aligned}$$

The three eigenvalues are $\hbar, 0, -\hbar$ respectively.

- The eigenvector corresponding to \hbar is obtained by solving the equations

$$\frac{i\hbar}{\sqrt{2}} \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \hbar \begin{pmatrix} a \\ b \\ c \end{pmatrix}$$

which are reduced to

$$\begin{aligned} \frac{-i}{\sqrt{2}}b &= a \\ \frac{i}{\sqrt{2}}(a - c) &= b \\ \frac{i}{\sqrt{2}}b &= c \end{aligned}$$

with solutions

$$a = -c, \quad b = \sqrt{2}ia$$

The normalization of the eigenvector gives

$$aa^* + bb^* + cc^* = 1$$

or

$$\begin{aligned} aa^* + 2aa^* + aa^* &= 1 \\ aa^* &= \frac{1}{4} \end{aligned}$$

thus

$$a = \frac{1}{2}e^{i\varphi}$$

We take $\varphi = 0$

$$a = \frac{1}{2}, b = \frac{i}{\sqrt{2}}, c = -\frac{1}{2}$$

and obtain the eigenvector

$$\phi_1 = \begin{pmatrix} \frac{1}{2} \\ \frac{i}{\sqrt{2}} \\ -\frac{1}{2} \end{pmatrix}$$

- The equations corresponding to eigenvalue 0 are

$$\frac{i\hbar}{\sqrt{2}} \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = 0 \begin{pmatrix} a \\ b \\ c \end{pmatrix}$$

So

$$b = 0, \quad a - c = 0$$

The normalization gives

$$\begin{aligned} aa^* + bb^* + cc^* &= 1 \\ 2aa^* &= 1 \end{aligned}$$

We thus have

$$a = \frac{1}{\sqrt{2}} e^{i\varphi}$$

again we take $\varphi = 0$, and

$$a = c = \frac{1}{\sqrt{2}}$$

The corresponding eigenvector is

$$\phi_2 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ \frac{1}{\sqrt{2}} \end{pmatrix}$$

- For the eigenvalue $-\hbar$

$$\frac{i\hbar}{\sqrt{2}} \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = -\hbar \begin{pmatrix} a \\ b \\ c \end{pmatrix}$$

We have

$$\begin{aligned} \frac{-i}{\sqrt{2}} b &= -a \\ \frac{i}{\sqrt{2}} (a - c) &= -b \\ \frac{i}{\sqrt{2}} b &= -c \end{aligned}$$

$$a = -c, \quad b = -\sqrt{2}ia$$

It is easy to repeat the procedure above and find the eigenvector

$$\phi_3 = \begin{pmatrix} \frac{1}{2} \\ -\frac{i}{\sqrt{2}} \\ -\frac{1}{2} \end{pmatrix}$$

Step 2. Determine the new basis and transformation coefficients. In representation \hat{B} the new basis are chosen as ϕ_1, ϕ_2, ϕ_3 , and they can be expanded as $\phi_\nu(x) = \sum_n (S^\dagger)_{n\nu} u_n(x)$

$$\begin{aligned} \phi_1 &= S_{11}^\dagger u_1 + S_{21}^\dagger u_2 + S_{31}^\dagger u_3 \\ &= \frac{1}{2} u_1 + \frac{i}{\sqrt{2}} u_2 - \frac{1}{2} u_3 \end{aligned}$$

We easily obtain the transformation matrix

$$S^\dagger = \begin{pmatrix} S_{11}^\dagger & S_{12}^\dagger & S_{13}^\dagger \\ S_{21}^\dagger & S_{22}^\dagger & S_{23}^\dagger \\ S_{31}^\dagger & S_{32}^\dagger & S_{33}^\dagger \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & \frac{1}{\sqrt{2}} & \frac{1}{2} \\ \frac{i}{\sqrt{2}} & 0 & -\frac{i}{\sqrt{2}} \\ -\frac{1}{2} & \frac{1}{\sqrt{2}} & -\frac{1}{2} \end{pmatrix}$$

which is just writing the three basis to form one square matrix with its hermitian conjugate

$$S = \begin{pmatrix} \frac{1}{2} & -\frac{i}{\sqrt{2}} & -\frac{1}{2} \\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{2} & \frac{i}{\sqrt{2}} & -\frac{1}{2} \end{pmatrix}.$$

i.e. $S_{\nu n} = (\phi_\nu, u_n)$.

Step 3. Running the transformation through. The states change as

$$\psi_B = S\psi_A$$

so in representation \hat{B} the eigenvectors are as expected

$$\begin{aligned} (\phi_1)_B &= \begin{pmatrix} \frac{1}{2} & -\frac{i}{\sqrt{2}} & -\frac{1}{2} \\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{2} & \frac{i}{\sqrt{2}} & -\frac{1}{2} \end{pmatrix} \begin{pmatrix} \frac{1}{2} \\ \frac{i}{\sqrt{2}} \\ -\frac{1}{2} \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \\ (\phi_2)_B &= \begin{pmatrix} \frac{1}{2} & -\frac{i}{\sqrt{2}} & -\frac{1}{2} \\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{2} & \frac{i}{\sqrt{2}} & -\frac{1}{2} \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ \frac{1}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \\ (\phi_3)_B &= \begin{pmatrix} \frac{1}{2} & -\frac{i}{\sqrt{2}} & -\frac{1}{2} \\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{2} & \frac{i}{\sqrt{2}} & -\frac{1}{2} \end{pmatrix} \begin{pmatrix} \frac{1}{2} \\ -\frac{i}{\sqrt{2}} \\ -\frac{1}{2} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \end{aligned}$$

And the operator \hat{L}_y transforms as

$$\begin{aligned} (\hat{L}_y)_B &= S(\hat{L}_y)_A S^{-1} \\ &= \begin{pmatrix} \frac{1}{2} & -\frac{i}{\sqrt{2}} & -\frac{1}{2} \\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{2} & \frac{i}{\sqrt{2}} & -\frac{1}{2} \end{pmatrix} \begin{pmatrix} 0 & -\frac{i\hbar}{\sqrt{2}} & 0 \\ \frac{i\hbar}{\sqrt{2}} & 0 & -\frac{i\hbar}{\sqrt{2}} \\ 0 & \frac{i\hbar}{\sqrt{2}} & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{2} & \frac{1}{\sqrt{2}} & \frac{1}{2} \\ \frac{i}{\sqrt{2}} & 0 & -\frac{i}{\sqrt{2}} \\ -\frac{1}{2} & \frac{1}{\sqrt{2}} & -\frac{1}{2} \end{pmatrix} \\ &= \begin{pmatrix} \frac{1}{2} & -\frac{i}{\sqrt{2}} & -\frac{1}{2} \\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{2} & \frac{i}{\sqrt{2}} & -\frac{1}{2} \end{pmatrix} \begin{pmatrix} \frac{\hbar}{2} & 0 & -\frac{\hbar}{2} \\ \frac{i\hbar}{\sqrt{2}} & 0 & \frac{i\hbar}{\sqrt{2}} \\ -\frac{\hbar}{2} & 0 & \frac{\hbar}{2} \end{pmatrix} = \begin{pmatrix} \hbar & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -\hbar \end{pmatrix} \end{aligned}$$

We summarize the result of representation transformation in the following table.

	Representation $\hat{A} : \{u_n(x)\}$	Representation $\hat{B} : \{\phi_\nu(x)\}$
State ψ	$\psi_A = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ \dots \end{pmatrix}, a_k = (u_k, \psi)$	$\psi_B = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ \dots \end{pmatrix}, b_\nu = (\phi_\nu, \psi)$
Operator \hat{F}	$F_A = \begin{pmatrix} F_{11} & F_{12} & \dots \\ F_{21} & F_{22} & \dots \\ \dots & \dots & \dots \end{pmatrix},$ $F_{mn} = (u_m, \hat{F}u_n)$	$F_B = \begin{pmatrix} F'_{11} & F'_{12} & \dots \\ F'_{21} & F'_{22} & \dots \\ \dots & \dots & \dots \end{pmatrix},$ $F_{\mu\nu} = (\phi_\mu, \hat{F}\phi_\nu)$
State transformation	$\psi_A = S^\dagger \psi_B$	$\psi_B = S\psi_A$
Operator transformation	$F_A = S^\dagger F_B S = S^{-1} F_B S$	$F_B = S F_A S^\dagger = S F_A S^{-1}$
Transformation matrix	$S^\dagger = \begin{pmatrix} S_{11}^\dagger & S_{12}^\dagger & \dots \\ S_{21}^\dagger & S_{22}^\dagger & \dots \\ \vdots & \vdots & \ddots \end{pmatrix},$ $S_{n\mu}^\dagger = (u_n, \phi_\mu)$	$S = \begin{pmatrix} S_{11} & S_{12} & \dots \\ S_{21} & S_{22} & \dots \\ \dots & \dots & \dots \end{pmatrix},$ $S_{\mu n} = (\phi_\mu, u_n)$

Problem 20 Zeng Jinyan, textbook, page 133, 15.

VII. DIRAC'S NOTATION

The theory of linear vector spaces had, of course, been known to mathematicians prior to the birth of quantum mechanics, but Dirac's way of introducing vector spaces has many advantages, especially from the physicist's point of view. Like a vector in 3D lives "out there in space", independent of anybody's choice of coordinates, the state of a system in quantum mechanics is represented by a vector, that lives "out there in Hilbert space". We can express the wavefunction by means of $\psi(x)$ in position space, $\phi(p)$ in momentum space, or c_n in \hat{Q} -representation with bases $\{u_n(x)\}$. But it's all the same state, $\psi(x)$, $\phi(p)$, and c_n contain exactly the same information.

A. States

Detached from concrete representations, Dirac's notation for a physical state represented by a state vector in a complex vector space is called a **ket** and denoted by $|\alpha\rangle$. The vector space is a ket space. We can also introduce the notion of a **bra** space, a vector space dual to the ket space. We postulate that corresponding to every ket $|\alpha\rangle$ there exists a bra, denoted by $\langle\alpha|$, in this dual, or bra, space. In \hat{Q} -representation the ket vector $|\alpha\rangle$ is a column matrix

$$|\alpha\rangle = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ \dots \end{pmatrix}$$

while the bra vector is a row matrix (the hermitian conjugate of the column matrix)

$$\langle\alpha| = (a_1^*, a_2^*, a_3^*, \dots)$$

The inner product of a bra and a ket is written as a bra standing on the left and a ket standing on the right, for example,

$$\langle\beta|\alpha\rangle = \underset{\text{bra}}{\langle\beta|} \cdot \underset{\text{ket}}{(|\alpha\rangle)} \underset{(c)}{}$$

This product is, in general, a complex number

$$\langle\beta|\alpha\rangle = b_1^* a_1 + b_2^* a_2 + b_3^* a_3 + \dots + b_n^* a_n + \dots = \sum_n b_n^* a_n$$

Evidently $\langle\beta|\alpha\rangle$ and $\langle\alpha|\beta\rangle$ are complex conjugates of each other

$$\langle\alpha|\beta\rangle = \langle\beta|\alpha\rangle^*$$

We immediately deduce that $\langle\alpha|\alpha\rangle$ must be a real number. Furthermore $\langle\alpha|\alpha\rangle \geq 0$, where the equality sign holds only if $|\alpha\rangle$ is a null ket. This is sometimes known as the postulate of positive definite metric, which is essential for the probabilistic interpretation of wavefunctions. Two kets $|\alpha\rangle$ and $|\beta\rangle$ are said to be orthogonal if

$$\langle\alpha|\beta\rangle = 0$$

Given a ket which is not a null ket, we can form a normalized ket $|\tilde{\alpha}\rangle$, where

$$|\tilde{\alpha}\rangle = \left(\frac{1}{\sqrt{\langle\alpha|\alpha\rangle}} \right) |\alpha\rangle$$

with the property

$$\langle\tilde{\alpha}|\tilde{\alpha}\rangle = 1$$

Quite generally, $\sqrt{\langle\alpha|\alpha\rangle}$ is known as the norm of $|\alpha\rangle$, analogous to the magnitude of a vector in Euclidean vector space.

B. Rules

In Dirac's notation, the basic rules of quantum mechanics look much simpler. Suppose we have discrete eigenvalue spectrum $\{u_n(x)\}$ or in Dirac's notation $\{|n\rangle\}$. The orthonormality reads

$$\langle m|n\rangle = \delta_{mn}$$

The completeness of the spectrum means that any ket $|\psi\rangle$ can be expanded as

$$|\psi\rangle = \sum_n c_n |n\rangle$$

where

$$c_n = \langle n|\psi\rangle$$

Inserting this back we find

$$|\psi\rangle = \sum_n |n\rangle \langle n|\psi\rangle$$

so

$$\sum_n |n\rangle \langle n| = 1 \tag{9}$$

which is the property of completeness in Dirac's notation. For continuous spectrum $\{u_z(x)\}$ (or $\{|z\rangle\}$) the orthonormality and completeness are expressed as

$$\begin{aligned} \langle z|z'\rangle &= \delta(z - z') \\ \int dz |z\rangle \langle z| &= 1 \end{aligned}$$

The eigen-problem of operator \hat{F} is written as

$$\hat{F} |n\rangle = f_n |n\rangle$$

The time-dependent Schrödinger equation takes the form of

$$i\hbar \frac{d}{dt} |\Psi\rangle = \hat{H} |\Psi\rangle$$

whereas the stationary equation is

$$\hat{H} |n\rangle = E_n |n\rangle.$$

As an example of application of the completeness, we notice that equation (9) is often inserted into appropriate position to simplify the calculation. Suppose the operator \hat{F} transforms ket $|\psi\rangle$ into $|\phi\rangle$

$$|\phi\rangle = \hat{F} |\psi\rangle$$

Inserting the completeness of $\{|n\rangle\}$ into the above equation twice

$$\sum_n |n\rangle \langle n|\phi\rangle = \sum_n \hat{F} |n\rangle \langle n|\psi\rangle$$

Multiply the bra $\langle m|$ from the left and notice $\langle m|n\rangle = \delta_{mn}$, we have

$$\langle m|\phi\rangle = \sum_n \langle m|\hat{F}|n\rangle \langle n|\psi\rangle$$

Here $\langle m|\hat{F}|n\rangle$ is the Dirac notation of the matrix elements F_{mn} .

C. $|n\rangle$

The occupation-number representation of the simple harmonic oscillator is very useful in representing quantum theory. We change the notation for *raising and lowering* operators into \hat{a}^\dagger and \hat{a} , which are known as the *creation and annihilation* operator for a particle with energy unit $\hbar\omega$. Let $|n\rangle$ be the eigenvector of Hamiltonian

$$\hat{H} = \hbar\omega \left(\hat{a}^\dagger \hat{a} + \frac{1}{2} \right) = \hbar\omega \left(\hat{N} + \frac{1}{2} \right)$$

with eigenvalue

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

The physical meaning of $|n\rangle$ is thus there are n particles in the state $|n\rangle$ and this representation is known as the *occupation-number representation*. The operator \hat{a} (\hat{a}^\dagger) annihilate (create) one particle in the state

$$\begin{aligned} \hat{a} |n\rangle &= \sqrt{n} |n-1\rangle \\ \hat{a}^\dagger |n\rangle &= \sqrt{n+1} |n+1\rangle \end{aligned}$$

For the ground state $|0\rangle$

$$\hat{a} |0\rangle = 0$$

so the first excited state can be generated from it

$$|1\rangle = \hat{a}^\dagger |0\rangle$$

and we can generate any state $|n\rangle$ from $|0\rangle$

$$|n\rangle = \frac{1}{\sqrt{n!}} (\hat{a}^\dagger)^n |0\rangle$$

Simple harmonic oscillator provides us an example of representing wavefunctions and operators. The basis in this example is the eigenvectors $\{|n\rangle\}$ and the matrix elements of operators are given by

$$F_{mn} = \langle m | \hat{F} | n \rangle$$

For example, the matrix elements for the annihilation operator are

$$\langle n' | \hat{a} | n \rangle = \sqrt{n} \delta_{n', n-1}$$

while for creation operator

$$\langle n' | \hat{a}^\dagger | n \rangle = \sqrt{n+1} \delta_{n', n+1}$$

In matrix form they are

$$\begin{aligned} a &= \begin{bmatrix} 0 & \sqrt{1} & 0 & 0 & \dots \\ 0 & 0 & \sqrt{2} & 0 & \dots \\ 0 & 0 & 0 & \sqrt{3} & \dots \\ 0 & 0 & 0 & 0 & \dots \\ \dots & \dots & \dots & \dots & \dots \end{bmatrix} \\ a^\dagger &= \begin{bmatrix} 0 & 0 & 0 & 0 & \dots \\ \sqrt{1} & 0 & 0 & 0 & \dots \\ 0 & \sqrt{2} & 0 & 0 & \dots \\ 0 & 0 & \sqrt{3} & 0 & \dots \\ \dots & \dots & \dots & \dots & \dots \end{bmatrix} \end{aligned}$$

and the number operator $\hat{N} = \hat{a}^\dagger \hat{a}$ is

$$N = \begin{bmatrix} 0 & 0 & 0 & 0 & \cdots \\ 0 & 1 & 0 & 0 & \cdots \\ 0 & 0 & 2 & 0 & \cdots \\ 0 & 0 & 0 & 3 & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \end{bmatrix}.$$

D. Application

We discuss here a crude model for (among other things) neutrino oscillations.

Example 21 *Imagine a system in which there are just two linearly independent states:*

$$|1\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \text{and} \quad |2\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

The most general state is a normalized linear combination

$$|\Psi\rangle = a|1\rangle + b|2\rangle = \begin{pmatrix} a \\ b \end{pmatrix}, \quad \text{with } a^2 + b^2 = 1.$$

The Hamiltonian can be expressed as a hermitian matrix; suppose it has the specific form

$$H = \begin{pmatrix} h & g \\ g & h \end{pmatrix}$$

where g and h are real constants. If the system starts out (at $t = 0$) in state $|1\rangle$, what is its state at time t ?

Solution 22 *The time-dependent Schrödinger equation says*

$$i\hbar \frac{d}{dt} |\Psi\rangle = \hat{H} |\Psi\rangle$$

as always, we begin by solving the time-independent Schrödinger equation

$$\hat{H} |\psi\rangle = E |\psi\rangle$$

that is, we look for the eigenvectors and eigenvalues of \hat{H} . The characteristic equation determines the eigenvalues

$$\begin{vmatrix} h - E & g \\ g & h - E \end{vmatrix} = 0 \Rightarrow (h - E)^2 - g^2 = 0 \Rightarrow h - E = \mp g \Rightarrow E_{\pm} = h \pm g$$

Evidently the allowed energies are $(h + g)$ and $(h - g)$. To determine the eigenvectors, we write

$$\begin{pmatrix} h & g \\ g & h \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = (h \pm g) \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \Rightarrow h\alpha + g\beta = (h \pm g)\alpha \Rightarrow \beta = \pm\alpha$$

so the normalized eigenvectors are

$$|\psi_{\pm}\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ \pm 1 \end{pmatrix}$$

Next we expand the initial state as a linear combination of eigenvectors of the Hamiltonian

$$|\Psi(0)\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \frac{1}{\sqrt{2}} (|\psi_+\rangle + |\psi_-\rangle)$$

Finally we tack on the standard time-dependence $\exp(-iE_n t/\hbar)$

$$\begin{aligned} |\Psi(t)\rangle &= \frac{1}{\sqrt{2}} \left(e^{-i(h+g)t/\hbar} |\psi_+\rangle + e^{-i(h-g)t/\hbar} |\psi_-\rangle \right) \\ &= \frac{1}{2} \left(e^{-i(h+g)t/\hbar} \begin{pmatrix} 1 \\ 1 \end{pmatrix} + e^{-i(h-g)t/\hbar} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right) \\ &= \frac{1}{2} e^{-iht/\hbar} \begin{pmatrix} e^{-igt/\hbar} + e^{igt/\hbar} \\ e^{-igt/\hbar} - e^{igt/\hbar} \end{pmatrix} \\ &= e^{-iht/\hbar} \begin{pmatrix} \cos(gt/\hbar) \\ -i \sin(gt/\hbar) \end{pmatrix} \end{aligned}$$

In the model of neutrino oscillation $|1\rangle$ represents the electron neutrino, and $|2\rangle$ the muon neutrino; if the Hamiltonian has a non-vanishing off-diagonal term g , then in the course of time the electron neutrino will turn into a muon neutrino, and back again. In 2002 Nobel Prize in Physics was awarded with one half jointly to: Raymond Davis Jr, Department of Physics and Astronomy, University of Pennsylvania, Philadelphia, USA, and Masatoshi Koshiba, International Center for Elementary Particle Physics, University of Tokyo, Japan, “for pioneering contributions to astrophysics, in particular for the detection of cosmic neutrinos”. In order to increase sensitivity to cosmic neutrinos, Koshiba constructed a larger detector, Super Kamiokande, which came into operation in 1996. This experiment has recently observed effects of neutrinos produced within the atmosphere, indicating a completely new phenomenon, neutrino oscillations, in which one kind of neutrino can change to another type. This implies that neutrinos have a non-zero mass, which is of great significance for the Standard Model of elementary particles and also for the role that neutrinos play in the universe.

E. Projection operators

The license to treat bras as separate entities in their own right allows for some powerful and pretty notation. For example, if $|\alpha\rangle$ is a normalized vector, the operator

$$\hat{P} \equiv |\alpha\rangle \langle\alpha|$$

picks out the component of any other vector that “lies along” $|\alpha\rangle$

$$\hat{P} |\beta\rangle = \langle\alpha|\beta\rangle |\alpha\rangle$$

we call it the *projection operator* onto the one-dimensional subspace spanned by $|\alpha\rangle$. It is easy to show that projection operators are idempotent: $\hat{P}^2 = \hat{P}$

$$\hat{P}^2 |\beta\rangle = \hat{P} (\hat{P} |\beta\rangle) = \hat{P} \langle\alpha|\beta\rangle |\alpha\rangle = \langle\alpha|\beta\rangle (\hat{P} |\alpha\rangle) = \langle\alpha|\beta\rangle \langle\alpha|\alpha\rangle |\alpha\rangle = \langle\alpha|\beta\rangle |\alpha\rangle = \hat{P} |\beta\rangle$$

To say two operators are equal means that they have the same effect on all vectors. Since $\hat{P}^2 |\beta\rangle = \hat{P} |\beta\rangle$ for any vector $|\beta\rangle$, $\hat{P}^2 = \hat{P}$.

One may also determine the eigenvalues and eigenvectors of \hat{P} . If $|\gamma\rangle$ is an eigenvector of \hat{P} with eigenvalue λ , then

$$\hat{P} |\gamma\rangle = \lambda |\gamma\rangle$$

and it follows that

$$\hat{P}^2 |\gamma\rangle = \lambda \hat{P} |\gamma\rangle = \lambda^2 |\gamma\rangle$$

But $\hat{P}^2 = \hat{P}$, and $|\gamma\rangle \neq 0$, so $\lambda^2 = \lambda$, and hence the eigenvalues of \hat{P} are 0 and 1. Any (complex) multiple of $|\alpha\rangle$ is an eigenvector of \hat{P} , with eigenvalue 1; any vector orthogonal to $|\alpha\rangle$ is an eigenvector of \hat{P} , with eigenvalue 0.

Let \hat{Q} be an operator with a complete set of orthonormal eigenvectors $\hat{Q} |n\rangle = q_n |n\rangle$. We can show that \hat{Q} can be written in terms of its spectral decomposition

$$\hat{Q} = \sum_n q_n |n\rangle \langle n|$$

An operator is characterized by its action on all possible vectors, so what we must show is that for any vector

$$\hat{Q}|\alpha\rangle = \left\{ \sum_n q_n |n\rangle \langle n| \right\} |\alpha\rangle$$

Indeed we expand $|\alpha\rangle$ as $|\alpha\rangle = \sum_n c_n |n\rangle$ with $c_n = \langle n|\alpha\rangle$,

$$\hat{Q}|\alpha\rangle = \sum_n \hat{Q}c_n |n\rangle = \sum_n \langle n|\alpha\rangle q_n |n\rangle = \left(\sum_n q_n |n\rangle \langle n| \right) |\alpha\rangle$$

so

$$\hat{Q} = \sum_n q_n |n\rangle \langle n|.$$

Problem 23 *Griffiths, page 123, 3.22, 3.23, 3.26, 3.35 (example), 3.37.*

VIII. CONSERVATION LAW AND SYMMETRY

In classical mechanics, there exist many conservation laws of dynamical variables, such as energy conservation, momentum conservation, etc. These conservation laws are closely connected with symmetries of the system. In quantum mechanics, the corresponding conservation law still holds and the relation between conservations and symmetries remains the same.

A. Constant of motion

Let us first check how the expectation value of operator changes with time. Consider the mean value of a certain dynamical variable \hat{F}

$$\langle F \rangle = \int \Psi^* \hat{F} \Psi d^3\mathbf{r}$$

on a state $\Psi(\mathbf{r}, t)$ which satisfies the time dependent Schrödinger equation

$$i\hbar \frac{\partial \Psi}{\partial t} = \hat{H} \Psi$$

The time derivative is calculated as

$$\begin{aligned} \frac{d}{dt} \langle F \rangle &= \frac{d}{dt} \int \Psi^* \hat{F} \Psi d^3\mathbf{r} \\ &= \int \left(\frac{\partial \Psi}{\partial t} \right)^* \hat{F} \Psi d^3r + \int \Psi^* \left(\frac{\partial \hat{F}}{\partial t} \right) \Psi d^3r + \int \Psi^* \hat{F} \left(\frac{\partial \Psi}{\partial t} \right) d^3r \\ &= \int \Psi^* \left(\frac{\partial \hat{F}}{\partial t} \right) \Psi d^3r + \int \left(\frac{1}{i\hbar} \hat{H} \Psi \right)^* \hat{F} \Psi d^3r + \int \Psi^* \hat{F} \left(\frac{1}{i\hbar} \hat{H} \Psi \right) d^3r \\ &= \int \Psi^* \left(\frac{\partial \hat{F}}{\partial t} + \frac{1}{i\hbar} [\hat{F}, \hat{H}] \right) \Psi d^3r \end{aligned}$$

If operator \hat{F} satisfy the operator relation

$$\frac{\partial \hat{F}}{\partial t} + \frac{1}{i\hbar} [\hat{F}, \hat{H}] = 0$$

Then $d\langle F \rangle / dt = 0$ and \hat{F} is called a *constant of motion*. Generally, the operator doesn't rely on time, $\partial \hat{F} / \partial t = 0$, so the criterion for a dynamical variable to be a constant of motion is

$$[\hat{F}, \hat{H}] = 0$$

That is, if $[\hat{F}, \hat{H}] = 0$, the expectation value of \hat{F} on any state does not depend on time. Furthermore, we can show the probability distribution of finding eigenvalues of \hat{F} is independent of time. We can choose the common eigenfunction set of \hat{F} and \hat{H} (this is possible because they commute each other) $\{u_{n,\mu}\}$

$$\begin{aligned}\hat{H}u_{n,\mu} &= E_n u_{n,\mu} \\ \hat{F}u_{n,\mu} &= f_\mu u_{n,\mu}.\end{aligned}$$

as our complete and orthonormal basis, with

$$\int u_{n,\mu}^* u_{m,\nu} d^3\mathbf{r} = \delta_{mn} \delta_{\mu\nu}$$

Thus any state can be expanded upon them

$$\begin{aligned}\psi(\mathbf{r}) &= \sum_{n,\mu} c_{n,\mu} u_{n,\mu} \\ \Psi(\mathbf{r}, t) &= \sum_{n,\mu} c_{n,\mu} u_{n,\mu} e^{-\frac{i}{\hbar} E_n t},\end{aligned}$$

and the expectation value of \hat{F} on state $\Psi(\mathbf{r}, t)$ is

$$\begin{aligned}\langle F \rangle &= \int \Psi^*(\mathbf{r}, t) \hat{F} \Psi(\mathbf{r}, t) d^3\mathbf{r} \\ &= \int \sum_{n,\mu} c_{n,\mu}^* u_{n,\mu}^* e^{\frac{i}{\hbar} E_n t} \sum_{m,\nu} f_\nu c_{m,\nu} u_{m,\nu} e^{-\frac{i}{\hbar} E_m t} d^3\mathbf{r} \\ &= \sum_{n,\mu} c_{n,\mu}^* c_{n,\mu} f_\mu = \sum_{n,\mu} |c_{n,\mu}|^2 f_\mu\end{aligned}$$

The probability of a measurement giving result f_μ , that is $|c_{n,\mu}|^2$, is therefore independent of time for given n .

B. Several typical cases of constant of motion

- (a) Momentum conservation: If the space is *homogeneous* (*free particle*), or the particle feels no potential, the momentum is conserved

$$\hat{H} = \frac{1}{2m} \hat{\mathbf{p}}^2$$

obviously

$$[\hat{\mathbf{p}}, \hat{H}] = 0.$$

- (b) Angular momentum conservation: If the space is *isotropic* (*central force field*), e.g. the particle feels the same potential in any direction

$$\begin{aligned}\hat{H} &= \frac{1}{2m} \hat{\mathbf{p}}^2 + V(r) \\ &= -\frac{\hbar^2}{2m} \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \right) + \frac{\hat{L}^2}{2mr^2} + V(r)\end{aligned}$$

with

$$\hat{L}^2 = -\hbar^2 \left\{ \frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \varphi^2} \right\}$$

The angular momentum operator and its components contain only θ and φ , and has nothing to do with r . Thus

$$[\hat{H}, \hat{L}^2] = 0$$

and

$$[\hat{H}, \hat{L}_x] = [\hat{H}, \hat{L}_y] = [\hat{H}, \hat{L}_z] = 0$$

i.e. the angular momentum is conserved, including its three components.

(c) Energy conservation: If the Hamiltonian itself \hat{H} doesn't depend on t , then evidently

$$[\hat{H}, \hat{H}] = 0$$

which is the conservation of energy in quantum mechanics.

(d) Parity conservation: Define an operation of space inversion $\mathbf{r} \rightarrow -\mathbf{r}$

$$\hat{P}\Psi(\mathbf{r}, t) = \Psi(-\mathbf{r}, t)$$

\hat{P} is known as the *parity operator*. We have

$$\hat{P}^2\Psi(\mathbf{r}, t) = \hat{P}\Psi(-\mathbf{r}, t) = \Psi(\mathbf{r}, t)$$

i.e. the eigenvalues of \hat{P}^2 is 1 and the eigenvalues for \hat{P} is ± 1 . So

$$\begin{aligned}\hat{P}\Psi_1 &= \Psi_1 \\ \hat{P}\Psi_2 &= -\Psi_2\end{aligned}$$

The eigenfunction Ψ_1 (Ψ_2) is known as the even (odd) parity state of the system. If the Hamiltonian is unchanged under space inversion

$$\hat{H}(-\mathbf{r}) = \hat{H}(\mathbf{r})$$

then

$$\begin{aligned}\hat{P}(\hat{H}\Psi(\mathbf{r}, t)) &= \hat{H}(-\mathbf{r})\hat{P}\Psi(\mathbf{r}, t) \\ &= \hat{H}(\mathbf{r})\hat{P}\Psi(\mathbf{r}, t)\end{aligned}$$

i.e.

$$\hat{P}\hat{H} = \hat{H}\hat{P}$$

Thus the parity is a constant of motion and \hat{P} and \hat{H} share common eigenfunction set, which means that the eigenfunctions of the system have definite parity and do not change with time. This is the conservation of parity in quantum mechanics.

Constant of motion is an important point in quantum mechanics. It should not be confused with the concept of **stationary state**. We emphasize here the distinction between them. Explicitly, *the stationary state is a particular eigenstate of \hat{H}* . It is the result of separation of variables from the Schrödinger equation

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

which gives the stationary equation

$$\hat{H}\psi(\mathbf{r}) = E\psi(\mathbf{r})$$

and the time-dependent factor

$$i\hbar \frac{d}{dt}f(t) = Ef(t)$$

and

$$\Psi(\mathbf{r}, t) = \psi(\mathbf{r})e^{-\frac{i}{\hbar}Et}.$$

The expectation value of *any* dynamical variable \hat{F} (constant of motion or not) on a stationary state

$$\begin{aligned}\langle F \rangle &= \int (\psi(\mathbf{r})e^{-\frac{i}{\hbar}Et})^* \hat{F}(\psi(\mathbf{r})e^{-\frac{i}{\hbar}Et}) d^3\mathbf{r} \\ &= \int \psi^*(\mathbf{r}) \hat{F} \psi(\mathbf{r}) d^3\mathbf{r}\end{aligned}$$

does not change with time. On the other hand, *the constant of motion is a particular observable* represented by an operator \hat{F} that commutes with \hat{H}

$$[\hat{F}, \hat{H}] = 0.$$

The expectation value of a constant of motion on *any* state (stationary or not)

$$\langle F \rangle = \int \Psi^*(\mathbf{r}, t) \hat{F} \Psi(\mathbf{r}, t) d^3\mathbf{r}$$

does not change with time.

Here is an additional notes on expectation value of conserved quantities. *The conserved quantities do not necessarily take definite values, unless the system is in one of the eigenstates of that operator.* For example, in the case of free particles, the constant of motion or the conserved quantity is the momentum \mathbf{p} , whose eigenstates are plane waves $e^{i\mathbf{p}\cdot\mathbf{x}/\hbar}$. The state of a free particle is generally a wave packet, i.e. the superposition of plane waves. The momentum does not take definite value - there is a probability distribution. Another example is the particle in central force field. The constant of motion or conserved quantity is the angular momentum $\hat{\mathbf{L}}$

$$\begin{aligned}\hat{L}^2 Y_{lm}(\theta, \varphi) &= l(l+1)\hbar^2 Y_{lm}(\theta, \varphi) \\ \hat{L}_z Y_{lm}(\theta, \varphi) &= m\hbar Y_{lm}(\theta, \varphi)\end{aligned}$$

The system may be in a state like $\psi = c_1 Y_{11}(\theta, \varphi) + c_2 Y_{20}(\theta, \varphi)$, for example or more generally

$$\psi = \sum_{lm} c_{lm} Y_{lm}(\theta, \varphi)$$

which is not an eigenstate of \hat{L}^2 and \hat{L}_z at all.

Generally the initial condition determines whether the system would stay in the eigenstate of the conserved quantity at time t . If the system is prepared initially ($t = 0$) in ψ_k , one of the eigenstate of conserved quantity \hat{F} , it will stay there forever. On the contrary, or more practically, if the system is initially in a combination state (not an eigenstate of \hat{F})

$$\Psi(\mathbf{r}, 0) = \sum_k a_k(0) \psi_k$$

at time t it is

$$\Psi(\mathbf{r}, t) = \sum_k a_k(t) \psi_k$$

and the probability of finding the particle in one of the eigenstates does not change with time

$$\frac{d}{dt} |a_k(t)|^2 = 0 \text{ or } |a_k(t)|^2 = |a_k(0)|^2$$

The quantum numbers of conserved quantity \hat{F} are known as *good quantum numbers*.

C. Conservation law and symmetry

Noether's theorem is an amazing result which lets physicists get conserved quantities from symmetries of the laws of nature. Time translation symmetry gives conservation of energy; space translation symmetry gives conservation of momentum; rotation symmetry gives conservation of angular momentum, and so on.

This result, proved in 1915 by Emmy Noether shortly after she first arrived in Göttingen, was praised by Einstein as a piece of "penetrating mathematical thinking". It's now a standard workhorse in theoretical physics.

If someone claims Noether's theorem says "every symmetry gives a conserved quantity", they are telling a half-truth. The theorem only applies to certain classes of theories. In its original version it applies to theories described by a *Lagrangian*, and the Lagrangian formalism does most of the work in proving the theorem. There is also a version which applies to theories described by a *Hamiltonian*. Luckily, almost all the theories studied in physics are described by both a Lagrangian and a Hamiltonian.

1. Translation invariance and momentum conservation

A translation operation is defined as a displacement of the system bodily along x -direction by an infinitesimal segment a so that $x \rightarrow x' = x + a$. Suppose $\psi(x)$ is the original wave function, $\psi'(x)$ is the translated one, $\hat{D}(a)$ is the translation operator, then

$$\psi \rightarrow \psi'(x) = \hat{D}(a)\psi(x)$$

Evidently translation invariance means that

$$\psi'(x') = \psi(x)$$

i.e.

$$\hat{D}(a)\psi(x+a) = \psi(x)$$

Replacing x with $x - a$ we have

$$\begin{aligned} \hat{D}(a)\psi(x) &= \psi(x-a) = \psi(x) - a\frac{d\psi}{dx} + \frac{1}{2!}(-a)^2\frac{d^2\psi}{dx^2} + \dots \\ &= \sum_{n=0}^{\infty} \frac{1}{n!}(-a)^n \frac{d^n}{dx^n}\psi(x) = e^{-a\frac{d}{dx}}\psi(x) \end{aligned}$$

The translation operator along x by a is thus expressed as

$$\hat{D}(a) = e^{-a\frac{d}{dx}} = e^{-\frac{i}{\hbar}a\hat{p}_x}$$

where the infinitesimal operator is nothing but the momentum operator

$$\hat{p}_x = -i\hbar\frac{d}{dx}$$

Suppose the space is homogeneous, the Hamiltonian $\hat{H} = \frac{1}{2m}\hat{\mathbf{p}}^2$ is invariant by a translation $\hat{D}(a)$

$$\hat{D}\left(\hat{H}\psi(x)\right) = \hat{H}\hat{D}\psi(x)$$

that is

$$\hat{D}\hat{H} = \hat{H}\hat{D} \Rightarrow \left[\hat{D}, \hat{H}\right] = 0 \Rightarrow \left[\hat{p}_x, \hat{H}\right] = 0$$

which is exactly the condition for conservation of momentum. We conclude that *homogeneity means conservation of momentum*.

It is easy to show translation operator $\hat{D}(a) = e^{-\frac{i}{\hbar}a\hat{p}_x}$ is unitary. Indeed, the Hermitian conjugate of $\hat{D}(a)$ is

$$\hat{D}(a)^\dagger = e^{\frac{i}{\hbar}a\hat{p}_x} = \hat{D}(a)^{-1} = \hat{D}(-a)$$

so

$$\hat{D}(a)\hat{D}(a)^\dagger = 1$$

Translation realized by $\hat{D}(a)$ is thus a unitary transformation. Operators are transformed as

$$\hat{F}' = \hat{D}(a)\hat{F}\hat{D}(a)^\dagger.$$

Generalization to 3D is straightforward.

2. Rotation invariance and orbital angular momentum conservation

Rotate the system bodily around z -axis by an angle α , the wavefunction $\psi(r, \theta, \varphi)$ changes as

$$\psi \rightarrow \psi'(r, \theta, \varphi) = \hat{R}(\alpha)\psi(r, \theta, \varphi)$$

Similarly rotation invariance means that

$$\begin{aligned}\psi'(r, \theta, \varphi') &= \psi(r, \theta, \varphi) \\ \hat{R}(\alpha)\psi(r, \theta, \varphi + \alpha) &= \psi(r, \theta, \varphi)\end{aligned}$$

Replacing φ with $\varphi - \alpha$ we have

$$\begin{aligned}\hat{R}(\alpha)\psi(r, \theta, \varphi) &= \psi(r, \theta, \varphi - \alpha) \\ &= \psi(r, \theta, \varphi) + (-\alpha)\frac{\partial}{\partial\varphi}\psi(r, \theta, \varphi) + \frac{1}{2!}(-\alpha)^2\frac{\partial^2}{\partial\varphi^2}\psi(r, \theta, \varphi) + \dots \\ &= \sum_{n=0}^{\infty} \frac{1}{n!}(-\alpha\frac{\partial}{\partial\varphi})^n\psi(r, \theta, \varphi) \\ &= e^{-\alpha\frac{\partial}{\partial\varphi}}\psi(r, \theta, \varphi)\end{aligned}$$

The rotation operator around z -axis by α is thus expressed as

$$\hat{R}(\alpha) = e^{-\alpha\frac{\partial}{\partial\varphi}} = e^{-\frac{i}{\hbar}\alpha\hat{L}_z}$$

where the infinitesimal operator is the z -component of the angular momentum operator

$$\hat{L}_z = -i\hbar\frac{\partial}{\partial\varphi}$$

Now we consider an infinitesimal rotation around an arbitrary axis \mathbf{n} by α

$$\begin{aligned}R(\mathbf{n}, \alpha)\psi(\mathbf{r}) &= \psi(\mathbf{r} - \Delta\mathbf{r}) \\ &= \psi(\mathbf{r}) - \Delta x\frac{\partial}{\partial x}\psi(\mathbf{r}) - \Delta y\frac{\partial}{\partial y}\psi(\mathbf{r}) - \Delta z\frac{\partial}{\partial z}\psi(\mathbf{r}) + \dots \\ &= \sum_{n=0}^{\infty} \frac{1}{n!}(-\Delta\mathbf{r} \cdot \nabla)^n\psi(\mathbf{r})\end{aligned}$$

As Figure 2 shows, $\Delta\mathbf{r} = \alpha(\mathbf{n} \times \mathbf{r})$

$$R(\mathbf{n}, \alpha)\psi(\mathbf{r}) = \sum_{n=0}^{\infty} \frac{1}{n!}(-\alpha)^n((\mathbf{n} \times \mathbf{r}) \cdot \nabla)^n\psi(\mathbf{r})$$

By means of the relation

$$(\mathbf{A} \times \mathbf{B}) \cdot \mathbf{C} = \mathbf{A} \cdot (\mathbf{B} \times \mathbf{C})$$

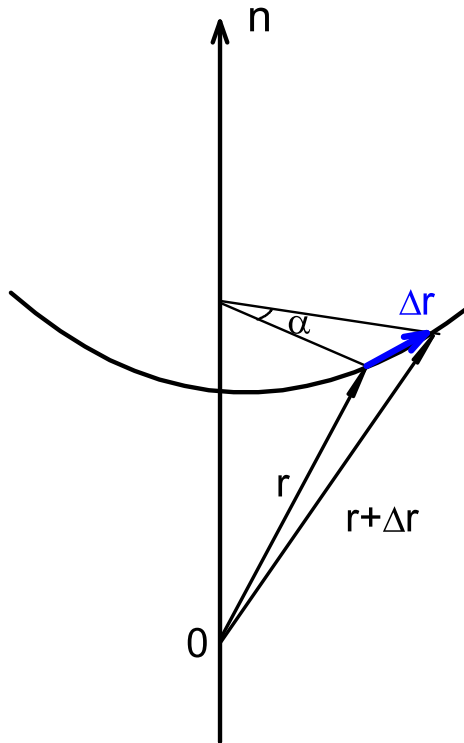


FIG. 2: A rotation around an arbitrary axis \mathbf{n} by α .

we know

$$(\mathbf{n} \times \mathbf{r}) \cdot \nabla = \mathbf{n} \cdot (\mathbf{r} \times \nabla)$$

so

$$\begin{aligned} R(\mathbf{n}, \alpha)\psi(\mathbf{r}) &= \sum_{n=0}^{\infty} \frac{1}{n!} [(-\alpha)\mathbf{n} \cdot (\mathbf{r} \times \nabla)]^n \psi(\mathbf{r}) \\ &= \sum_{n=0}^{\infty} \frac{1}{n!} \left[-\frac{i}{\hbar} \alpha \mathbf{n} \cdot (\mathbf{r} \times \mathbf{p}) \right]^n \psi(\mathbf{r}) \\ &= e^{-\frac{i}{\hbar} \alpha (\mathbf{n} \cdot \hat{\mathbf{L}})} \psi(\mathbf{r}) \end{aligned}$$

The rotation operator around an arbitrary axis \mathbf{n} by α is thus expressed as

$$R(\mathbf{n}, \alpha) = e^{-\frac{i}{\hbar} \alpha (\mathbf{n} \cdot \hat{\mathbf{L}})}$$

More explicitly

$$\begin{aligned} \mathbf{n} \Rightarrow \mathbf{x}_0, R(\mathbf{x}_0, \alpha) &= e^{-\frac{i}{\hbar} \alpha \hat{L}_x} \\ \mathbf{n} \Rightarrow \mathbf{y}_0, R(\mathbf{y}_0, \alpha) &= e^{-\frac{i}{\hbar} \alpha \hat{L}_y} \\ \mathbf{n} \Rightarrow \mathbf{z}_0, R(\mathbf{z}_0, \alpha) &= e^{-\frac{i}{\hbar} \alpha \hat{L}_z} \end{aligned}$$

Suppose the Hamiltonian is isotropic or the potential has spherical symmetry (central force field)

$$\hat{H} = \frac{1}{2m}\hat{\mathbf{p}}^2 + V(r)$$

we thus have

$$\begin{aligned} [R(\mathbf{x}_0, \alpha), \hat{H}] &= [R(\mathbf{y}_0, \alpha), \hat{H}] = [R(\mathbf{z}_0, \alpha), \hat{H}] = 0 \\ [\hat{L}_x, \hat{H}] &= [\hat{L}_y, \hat{H}] = [\hat{L}_z, \hat{H}] = 0 \end{aligned}$$

i.e. $\hat{L}_x, \hat{L}_y, \hat{L}_z$ are conservatives, so is operator $\hat{L}^2 = \hat{L}_x^2 + \hat{L}_y^2 + \hat{L}_z^2$. *Isotropy thus means the conservation of momentum.* In cases the potential V is invariant under rotation around certain direction, the component of angular momentum on that direction is conservative. For example, in the case of a cylindric symmetry potential $V = \frac{1}{2}m\omega_{\perp}^2(x^2 + y^2) + \frac{1}{2}m\omega_z^2z^2$, \hat{L}_z is conservative.

Rotation operator is also a unitary operator and transformation accomplished by such an operator is an unitary transformation. Indeed

$$\begin{aligned} R(\mathbf{n}, \alpha) &= e^{-\frac{i}{\hbar}\alpha(\mathbf{n}\cdot\hat{\mathbf{L}})} \\ R^{\dagger}(\mathbf{n}, \alpha) &= e^{\frac{i}{\hbar}\alpha(\mathbf{n}\cdot\hat{\mathbf{L}})} = R^{-1}(\mathbf{n}, \alpha) \end{aligned}$$

and wavefunctions and operators are transformed according to the following rule

$$\begin{aligned} \psi(\mathbf{r}) &\Rightarrow \psi'(\mathbf{r}) = R(\mathbf{n}, \alpha)\psi(\mathbf{r}) = e^{-\frac{i}{\hbar}\alpha(\mathbf{n}\cdot\hat{\mathbf{L}})}\psi(\mathbf{r}) \\ \hat{F} &\Rightarrow \hat{F}' = R(\mathbf{n}, \alpha)\hat{F}R^{\dagger}(\mathbf{n}, \alpha) = e^{-\frac{i}{\hbar}\alpha(\mathbf{n}\cdot\hat{\mathbf{L}})}\hat{F}e^{\frac{i}{\hbar}\alpha(\mathbf{n}\cdot\hat{\mathbf{L}})} \end{aligned}$$

3. Conservation laws and symmetry in general

More generally, we consider a linear transformation \hat{Q} , which transforms Ψ into $\Psi' = \hat{Q}\Psi$, the Schrödinger equation is invariant under this transformation, i.e.

$$i\hbar\frac{\partial}{\partial t}\Psi = \hat{H}\Psi \quad \text{and} \quad i\hbar\frac{\partial}{\partial t}\Psi' = \hat{H}\Psi'$$

We have

$$i\hbar\frac{\partial}{\partial t}\hat{Q}\Psi = \hat{H}\hat{Q}\Psi$$

Applying \hat{Q}^{-1} gives

$$i\hbar\frac{\partial}{\partial t}\Psi = \hat{Q}^{-1}\hat{H}\hat{Q}\Psi = \hat{H}\Psi$$

then

$$\hat{Q}^{-1}\hat{H}\hat{Q} = \hat{H} \Rightarrow \hat{H}\hat{Q} = \hat{Q}\hat{H} \Rightarrow [\hat{Q}, \hat{H}] = 0$$

The symmetric transformation \hat{Q} forms a symmetric group and the conservation of probability requires that \hat{Q} is unitary, i.e. $\hat{Q}\hat{Q}^{\dagger} = \hat{Q}^{\dagger}\hat{Q} = I$.

For symmetry transformations that differ infinitesimally from the identity transformation I , we can write

$$\hat{Q} = I - i\varepsilon\hat{F}$$

where ε is an infinitesimal parameter. The unitary of \hat{Q} means

$$\begin{aligned} \hat{Q}^{\dagger}\hat{Q} &= (I + i\varepsilon\hat{F}^{\dagger})(I - i\varepsilon\hat{F}) \\ &= I - i\varepsilon(\hat{F} - \hat{F}^{\dagger}) + O(\varepsilon^2) \\ &= I \end{aligned}$$

so

$$\hat{F} = \hat{F}^\dagger$$

i.e. \hat{F} is Hermitian operator, known as the Hermitian generator of \hat{Q} . Then

$$[\hat{Q}, \hat{H}] = 0 \implies [\hat{F}, \hat{H}] = 0 \implies \frac{d}{dt} \langle F \rangle = 0$$

In general, *there exists a constant of motion \hat{F} corresponding to symmetry transformation \hat{Q} .*

The following table summarizes the main result in this section.

Symmetry	Conserved Quantity
Homogenous Space/Translational Invariance	Momentum
Isotropic Space/Rotational Invariance	Angular Momentum
Space Inversion Invariance	Parity
Time Translation Invariance	Energy

IX. SUMMARY ON PART IV

Two cornerstones in quantum mechanics are wavefunctions and operators. In wave mechanics, wavefunctions are mathematical functions representing wave nature of the micro-particle, while operators are mathematical derivatives acting on the wavefunctions. There exists a special class of operators in quantum theory - Hermitian operators. They usually obey a non-commutative algebra and their eigenvalues are real, the eigenfunctions are orthonormal and complete.

We can nevertheless describe the theory by means of another totally different language - matrix. In matrix mechanics, wavefunctions are column or row matrices, whereas operators are square matrices. The eigenvalue problem is reduced to find the eigenvalues and eigenfunctions of a Hermitian matrix.

The completeness of the eigenvectors of a Hermitian operator enables us to represent the quantum theory using the eigenvectors as basis just like we use three basic vectors to express any vectors in 3D space. In this sense, wavefunctions live in Hilbert space, which is complex and infinite dimensions. Moreover, we can jump between two different bases - the wavefunctions and operators are connected by a unitary matrix.

Dirac invented much simpler notations to express quantum theory. In this new picture, states are distinguished as bras and kets. An especially useful representation is the occupation number representation of harmonic oscillator.

For a measurement of physical quantity \hat{F} on a wave function $\psi(x)$ the result could be nothing but one of the eigenvalues of \hat{F} , say f_n . And probability of obtaining f_n is $|c_n|^2$. If two operators are commutative, one can find the common eigenfunctions set and measure them simultaneously. On the other hand, if they fail to commute, there always exists a limit for their variances - the uncertainty relation.

These conservation of physical quantities are closely connected with the symmetries of system. According to Noether's theorem, there exists a constant of motion \hat{F} corresponding to symmetry transformation \hat{Q} .

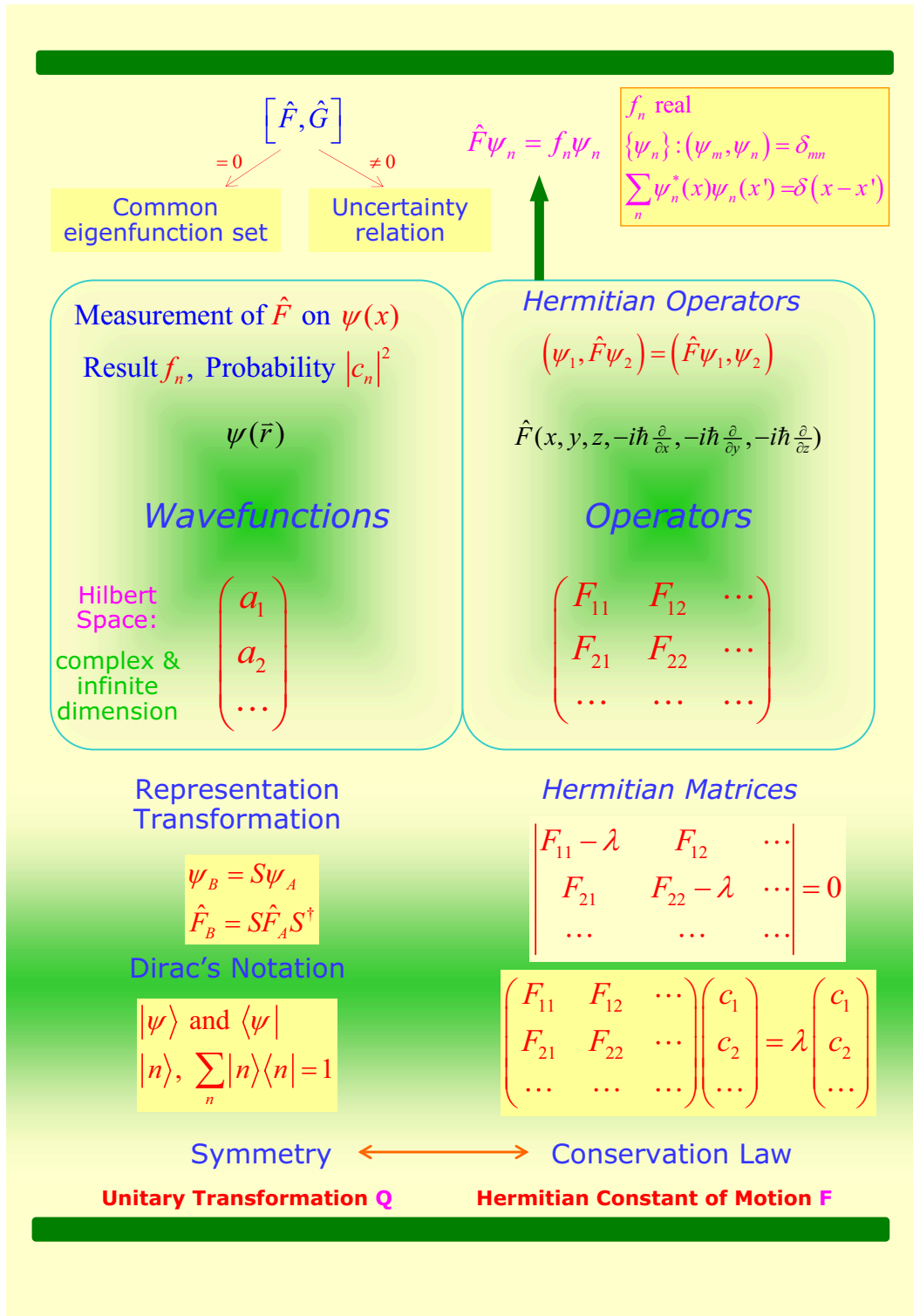


FIG. 3: Summary on Part IV.

Lecture Notes on Quantum Mechanics - Part V

Yunbo Zhang

Instituter of Theoretical Physics, Shanxi University

Few problems in quantum mechanics - with either time-independent or time dependent Hamiltonians - can be solved exactly. This part is devoted to a systematic discussion of the approximation methods used in quantum mechanics bound state problems, including the time-(in)dependent perturbation theory and variational approach. Within a semiclassical treatment, we shall address the basic issues of interaction between light and atom. Some modern developments on adiabatic approximation and Berry's phase will be discussed.

Contents

I. Stationary Perturbation Theory: Non-Degenerate Case	2
A. Order of approximation	2
B. First order approximation	3
C. Second order approximation	4
D. Examples	5
II. Stationary Perturbation Theory: Degenerate Case	10
A. Degenerate perturbation theory	10
B. Stark effect of hydrogen atom	12
III. Variational Method	14
A. Ground state energy of 1D infinite square well	15
B. Ground state energy of Helium atom	17
1. Hamiltonian and trial wave function	17
2. The expectation value	19
3. Optimization	21
4. Perturbation result	22
IV. Time-Dependent Perturbation Theory	23
A. Time evolution of quantum state	23
B. Time-dependent perturbation theory - quantum transition probability	23
C. Constant perturbation	26
D. Sinusoidal perturbation	28
E. Elastic scattering	29
V. Semiclassical Treatment of Emission and Absorption of Radiation	30
A. Absorption and stimulated emission of light	31
B. Einstein's treatment of spontaneous emission	33
VI. Schrödinger Picture, Heisenberg Picture and Interaction Picture	34
A. Schrödinger picture	34
B. Heisenberg picture	36
C. Interaction picture	37
VII. Adiabatic Approximation and Berry's Phase	40
A. Proof of the adiabatic theorem	40
B. Nonholonomic Processes	43
C. Berry's Phase	44
VIII. Summary on Part V	46

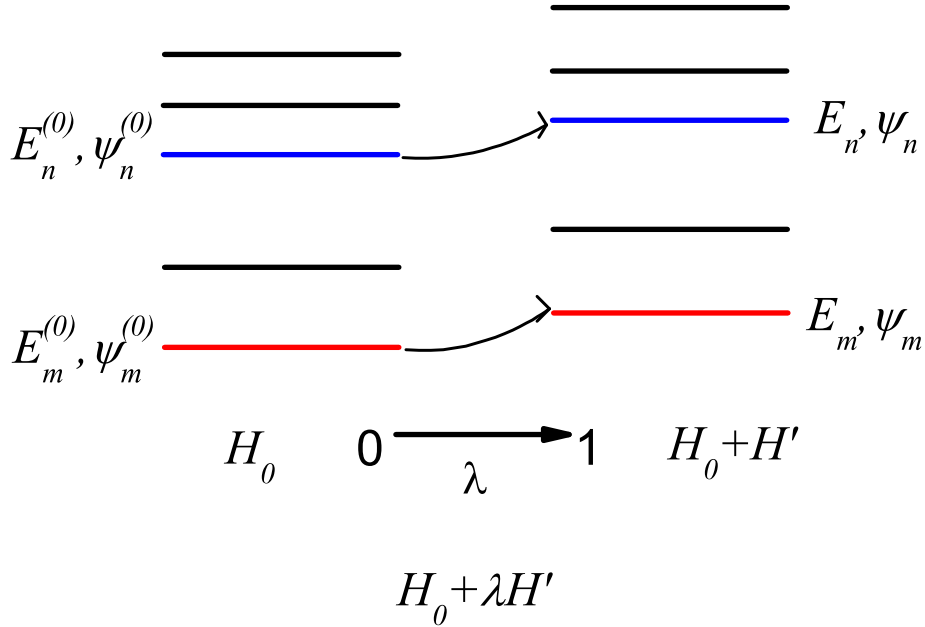


FIG. 1: Perturbation theory: A small perturbation modifies the wave functions and energies of the system.

I. STATIONARY PERTURBATION THEORY: NON-DEGENERATE CASE

One of the main problems in quantum mechanics is to solve the stationary Schrödinger Equation

$$\hat{H}\psi = E\psi \quad (1)$$

Cases that the equation can be solved exactly are comparatively rare. In most cases, the Hamiltonian becomes too complicated to be solved exactly. Thus we had to turn ourselves to invoke help of approximation method. The most frequently used method is perturbation method.

The premise of perturbation theory is that the Hamiltonian \hat{H} can be separated into two parts, \hat{H}_0 and \hat{H}'

$$\hat{H} = \hat{H}_0 + \hat{H}'$$

in which \hat{H}_0 is the main part and \hat{H}' is the perturbation.

\hat{H}_0 is the principal part of the Hamiltonian that defines the system. Suppose we have solved the (time-independent) Schrödinger equation for some potential (say, the one-dimensional infinite square well)

$$\hat{H}_0 \psi_n^{(0)} = E_n^{(0)} \psi_n^{(0)} \quad (2)$$

obtaining a complete set of orthonormal eigenfunctions $\psi_n^{(0)}$ and the corresponding eigenvalues $E_n^{(0)}$, which give the wave motion and the corresponding energy values of the system under unperturbed condition. But unless we are very lucky, we're unlikely to be able to solve the Schrödinger equation (1) exactly, for more complicated potential.

\hat{H}' is the perturbation. Its effects are much less than \hat{H}_0 . It only “modifies” the motion (wave function) and system energy to some extent. Perturbation theory is a systematic procedure to calculate these modifications.

A. Order of approximation

In order to show the order of approximation, we add a parameter λ to \hat{H}' so that

$$\hat{H} = \hat{H}_0 + \lambda \hat{H}'. \quad (3)$$

For the moment we'll take λ to be a small number; later we'll crank it up to 1, and \hat{H} will be the true, exact Hamiltonian. Writing E_n and ψ_n as power series in λ , we have

$$E_n = E_n^{(0)} + \lambda E_n^{(1)} + \lambda^2 E_n^{(2)} + \dots \quad (4)$$

$$\psi_n = \psi_n^{(0)} + \lambda \psi_n^{(1)} + \lambda^2 \psi_n^{(2)} + \dots \quad (5)$$

where $E_n^{(1)}$, $E_n^{(2)}$ are the first- and second-order perturbations to the n^{th} eigenvalue, and $\psi_n^{(1)}$, $\psi_n^{(2)}$ are the first- and second-order perturbations to the n^{th} eigenfunctions, etc. Plugging Equations (4), (5), and (3) into Equation (1), we have

$$\begin{aligned} & \left(\hat{H}_0 + \lambda \hat{H}' \right) \left(\psi_n^{(0)} + \lambda \psi_n^{(1)} + \lambda^2 \psi_n^{(2)} + \dots \right) \\ & = \left(E_n^{(0)} + \lambda E_n^{(1)} + \lambda^2 E_n^{(2)} + \dots \right) \left(\psi_n^{(0)} + \lambda \psi_n^{(1)} + \lambda^2 \psi_n^{(2)} + \dots \right) \end{aligned}$$

To the zeroth order (λ^0) this yields

$$\hat{H}_0 \psi_n^{(0)} = E_n^{(0)} \psi_n^{(0)}$$

which is nothing new just equation (2). To first order (λ^1)

$$\hat{H}_0 \psi_n^{(1)} + \hat{H}' \psi_n^{(0)} = E_n^{(0)} \psi_n^{(1)} + E_n^{(1)} \psi_n^{(0)} \quad (6)$$

To second order (λ^2)

$$\hat{H}_0 \psi_n^{(2)} + \hat{H}' \psi_n^{(1)} = E_n^{(0)} \psi_n^{(2)} + E_n^{(1)} \psi_n^{(1)} + E_n^{(2)} \psi_n^{(0)} \quad (7)$$

and so on. (I'm done with λ , now it was just a device to keep track of the different orders - so crank it up to 1.)

B. First order approximation

Taking the inner product of equation (6), that is, multiplying by $(\psi_n^{(0)})^*$ and integrating,

$$\int \psi_n^{(0)*} \left\{ \hat{H}_0 \psi_n^{(1)} + \hat{H}' \psi_n^{(0)} = E_n^{(0)} \psi_n^{(1)} + E_n^{(1)} \psi_n^{(0)} \right\} d^3 \mathbf{r}$$

But \hat{H}_0 is Hermitian, so

$$\int \psi_n^{(0)*} \hat{H}_0 \psi_n^{(1)} d^3 \mathbf{r} = E_n^{(0)} \int \psi_n^{(0)*} \psi_n^{(1)} d^3 \mathbf{r}$$

which cancels the first term on the right. So

$$E_n^{(1)} = \int \psi_n^{(0)*} \hat{H}' \psi_n^{(0)} d^3 \mathbf{r} = \hat{H}'_{nn} \quad (8)$$

or in Dirac's notation

$$E_n^{(1)} = \left\langle \psi_n^{(0)} \left| \hat{H}' \right| \psi_n^{(0)} \right\rangle. \quad (9)$$

This is the fundamental result of first-order perturbation theory; as a *practical* matter, it may well be the most important equation in quantum mechanics. It says that the **first-order correction** to the energy is the *expectation value* of the perturbation in the *unperturbed state*.

To find the first order approximation of wave function $\psi_n^{(1)}$, we recognized that the zeroth order wave function $\{\psi_n^{(0)}\}$ is a complete set. The unknown $\psi_n^{(1)}$ may be developed upon them,

$$\psi_n^{(1)} = \sum_{l \neq n} a_l^{(1)} \psi_l^{(0)}.$$

If we could determine the coefficients $a_l^{(1)}$, we'd be done. First we can show that $a_n^{(1)} = 0$, that is, there is no need to include $l = n$ in the sum. A glance at Equation (5) reveals that any $\psi_n^{(0)}$ component in $\psi_n^{(1)}$ might as well be pulled out and combined with the first term. We are only concerned, for the moment, with solving the Schrödinger equation (Equation (1)), and the ψ_n we get will not, in general, be normalized. Equation (6) then becomes

$$\hat{H}_0 \sum_{l \neq n} a_l^{(1)} \psi_l^{(0)} + \hat{H}' \psi_n^{(0)} = E_n^{(0)} \sum_{l \neq n} a_l^{(1)} \psi_l^{(0)} + E_n^{(1)} \psi_n^{(0)}. \quad (10)$$

Taking the inner product with $\psi_n^{(0)*}$ gives exactly (9), while that with $\psi_m^{(0)*}$ ($m \neq n$) leads to

$$E_m^{(0)} a_m^{(1)} + H'_{mn} = E_n^{(0)} a_m^{(1)}$$

or

$$a_m^{(1)} = \frac{H'_{mn}}{E_n^{(0)} - E_m^{(0)}}.$$

The first order approximation of wave function $\psi_n^{(1)}$ is thus

$$\psi_n^{(1)} = \sum_{m \neq n} \frac{H'_{mn}}{E_n^{(0)} - E_m^{(0)}} \psi_m^{(0)} \quad (11)$$

or in Dirac's notation

$$\psi_n^{(1)} = \sum_{m \neq n} \frac{\langle \psi_m^{(0)} | \hat{H}' | \psi_n^{(0)} \rangle}{E_n^{(0)} - E_m^{(0)}} \psi_m^{(0)} \quad (12)$$

The first order approximation $\psi_n^{(1)}$ does not involve $\psi_n^{(0)}$ itself - $\psi_n^{(0)}$ only appears in second-order approximation. Notice that the denominator is safe, since there is no coefficient with $m = n$, as long as the unperturbed energy spectrum is non-degenerate. But if two different unperturbed states share the same energy, we're in serious trouble (we divided by zero to get Equation 11); in that case we need degenerate perturbation theory, which will be discussed in next section.

That completes first-order perturbation theory: $E_n^{(1)}$ is given by Equation (8), and $\psi_n^{(1)}$ is given by Equation (11). I should warn you that whereas perturbation theory often yields surprisingly accurate energies (that is, $E_n^{(0)} + E_n^{(1)}$ is quite close to the exact value E), the wave functions are notoriously poor.

C. Second order approximation

Now we turn to equation (7) for second-order approximation. Proceeding as before, we insert the first-order wavefunction into the second-order equation and take the inner product with $\psi_n^{(0)}$

$$\int \psi_n^{(0)*} \left\{ \hat{H}_0 \psi_n^{(2)} + \hat{H}' \sum_{m \neq n} a_m^{(1)} \psi_m^{(0)} = E_n^{(0)} \psi_n^{(2)} + E_n^{(1)} \sum_{m \neq n} a_m^{(1)} \psi_m^{(0)} + E_n^{(2)} \psi_n^{(0)} \right\} d^3 \mathbf{r}.$$

Again, we exploit the Hermiticity of \hat{H}_0

$$\int \psi_n^{(0)*} \hat{H}_0 \psi_n^{(2)} d^3 \mathbf{r} = \int \psi_n^{(0)*} E_n^{(0)} \psi_n^{(2)} d^3 \mathbf{r}$$

so the first term on the left cancels the first term on the right. Moreover

$$E_n^{(1)} \int \psi_n^{(0)*} \sum_{m \neq n} a_m^{(1)} \psi_m^{(0)} d^3 \mathbf{r} = 0$$

and we are left with

$$\begin{aligned} E_n^{(2)} &= \sum_{m \neq n} a_m^{(1)} \int \psi_n^{(0)*} \hat{H}' \psi_m^{(0)} d^3\mathbf{r} \\ &= \sum_{m \neq n} a_m^{(1)} H'_{nm} \\ &= \sum_{m \neq n} \frac{H'_{mn} H'_{nm}}{E_n^{(0)} - E_m^{(0)}} \end{aligned}$$

or finally

$$E_n^{(2)} = \sum_{m \neq n} \frac{|H'_{mn}|^2}{E_n^{(0)} - E_m^{(0)}}. \quad (13)$$

This is the fundamental result of **second-order perturbation** theory. We could proceed to calculate the second-order correction to the wave function ($\psi_n^{(2)}$), the third-order correction to the energy, and so on, but in practice Equation (13) is ordinarily as high as it is useful to pursue this method.

In summary we have

$$\begin{aligned} \psi_n &= \psi_n^{(0)} + \psi_n^{(1)} + \dots \\ &= \psi_n^{(0)} + \sum_{m \neq n} \frac{H'_{mn}}{E_n^{(0)} - E_m^{(0)}} \psi_m^{(0)} + \dots \\ E_n &= E_n^{(0)} + E_n^{(1)} + \dots \\ &= E_n^{(0)} + H'_{nn} + \sum_{m \neq n} \frac{|H'_{mn}|^2}{E_n^{(0)} - E_m^{(0)}} + \dots \end{aligned}$$

The condition for validity of this perturbation theory is that the perturbation Hamiltonian \hat{H}' should be small such that the matrix elements H'_{mn} , H'_{nn} be small, i.e.

$$H'_{mn} \ll |E_m^{(0)} - E_n^{(0)}|.$$

D. Examples

Example 1 *Charged Oscillator in Electric Field.* Suppose we have an Harmonic oscillator

$$\hat{H}_0 = \frac{1}{2m} \hat{p}^2 + \frac{1}{2} m \omega^2 \hat{x}^2,$$

which carries electric charge e . A weak electric field ε applies along the $+x$ direction with static electric potential $-\varepsilon x$. The interaction of the oscillator with electric field is described by

$$\hat{H}' = -e\varepsilon \hat{x}.$$

We have already known the zeroth-order states and energy

$$\begin{aligned} E_n^{(0)} &= \left(n + \frac{1}{2} \right) \hbar \omega \\ \psi_n^{(0)} &= |n\rangle. \end{aligned}$$

The first order perturbation to the energy is evidently zero

$$E_n^{(1)} = \langle n | \hat{H}' | n \rangle = 0$$

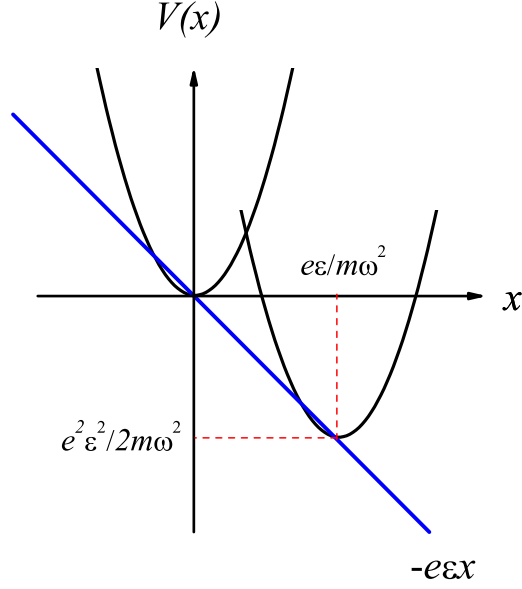


FIG. 2: Charged harmonic oscillator in electric field: Perturbation and exact result.

As for the first order correction to the wave function, we need calculate the matrix element

$$\begin{aligned}
 \hat{H}'_{mn} &= \langle m | (-e\varepsilon\hat{x}) | n \rangle \\
 &= -e\varepsilon\sqrt{\frac{\hbar}{2m\omega}} \langle m | (\hat{a}^\dagger + \hat{a}) | n \rangle \\
 &= \begin{cases} -e\varepsilon\sqrt{\frac{\hbar}{2m\omega}}\sqrt{n+1} & m = n+1 \\ -e\varepsilon\sqrt{\frac{\hbar}{2m\omega}}\sqrt{n} & m = n-1 \\ 0 & \text{others} \end{cases} .
 \end{aligned}$$

thus

$$\begin{aligned}
 \psi_n^{(1)} &= \sum_{m \neq n} \frac{H'_{mn}}{E_n^{(0)} - E_m^{(0)}} \psi_m^{(0)} \\
 &= \frac{H'_{n+1,n}}{E_n^{(0)} - E_{n+1}^{(0)}} \psi_{n+1}^{(0)} + \frac{H'_{n-1,n}}{E_n^{(0)} - E_{n-1}^{(0)}} \psi_{n-1}^{(0)} \\
 &= \frac{e\varepsilon}{\hbar\omega} \sqrt{\frac{\hbar}{2m\omega}} \sqrt{n+1} \psi_{n+1}^{(0)} - \frac{e\varepsilon}{\hbar\omega} \sqrt{\frac{\hbar}{2m\omega}} \sqrt{n} \psi_{n-1}^{(0)}.
 \end{aligned}$$

where we have used

$$E_n^{(0)} - E_{n\pm 1}^{(0)} = \mp \hbar\omega$$

The second order correction to the energy is

$$\begin{aligned}
 E_n^{(2)} &= \frac{|H'_{n+1,n}|^2}{E_n^{(0)} - E_{n+1}^{(0)}} + \frac{|H'_{n-1,n}|^2}{E_n^{(0)} - E_{n-1}^{(0)}} \\
 &= -\frac{e^2\varepsilon^2}{\hbar\omega} (n+1) \frac{\hbar}{2m\omega} + \frac{e^2\varepsilon^2}{\hbar\omega} n \frac{\hbar}{2m\omega} = -\frac{e^2\varepsilon^2}{2m\omega^2}
 \end{aligned}$$

which means the perturbation always lower the energy by an amount $\frac{e^2 \epsilon^2}{2m\omega^2}$. On the other hand this example can be solved exactly. The potential energy can be rewritten as

$$\frac{1}{2}m\omega^2 x^2 - e\epsilon x = \frac{1}{2}m\omega^2 \left(x - \frac{e\epsilon}{m\omega^2}\right)^2 - \frac{e^2 \epsilon^2}{2m\omega^2}$$

which is just a shift of the harmonic potential as shown in Figure 2.

Example 2 Perturbation in infinite square well. The unperturbed wave functions for the infinite square well are

$$\psi_n^{(0)}(x) = \sqrt{\frac{2}{a}} \sin\left(\frac{n\pi}{a}x\right), \quad E_n^{(0)} = \frac{n^2 \pi^2 \hbar^2}{2ma^2}$$

which satisfy

$$H_0 \psi_n^{(0)} = E_n^{(0)} \psi_n^{(0)}$$

We perturb the system in three different ways (Figure 3) and solve the equation

$$H \psi_n(x) = E_n \psi_n(x).$$

with

$$H = H_0 + H'.$$

- Suppose first that we perturb the system by simply raising the “floor” of the well by a constant amount V_0 (Figure 3). In that case $H' = V_0$ and the first-order correction to the energy of the n -th state is

$$\begin{aligned} E_n^{(1)} &= \left\langle \psi_n^{(0)}(x) \left| V_0 \right| \psi_n^{(0)}(x) \right\rangle \\ &= V_0 \left\langle \psi_n^{(0)}(x) \left| \psi_n^{(0)}(x) \right\rangle = V_0 \end{aligned}$$

The corrected energy levels, then, are

$$E_n = E_n^{(0)} + V_0.$$

they are simply lifted by the amount V_0 . The only surprising thing is that in this case the first-order theory yields the exact answer. For a constant perturbation all the higher corrections vanish.

- If the perturbation extends only half-way across the well (Figure 3) then

$$\begin{aligned} E_n^{(1)} &= \int_0^{\frac{a}{2}} \psi_n^{(0)*}(x) V_0 \psi_n^{(0)}(x) dx \\ &= \frac{2V_0}{a} \int_0^{\frac{a}{2}} \sin^2\left(\frac{n\pi}{a}x\right) dx = \frac{V_0}{2}. \end{aligned}$$

In this case every energy level is lifted by $V_0/2$. That's not the exact result, presumably, but it does seem reasonable, as a first order approximation.

- Finally we put a δ -function bump in the center of the well

$$H' = \alpha \delta\left(x - \frac{a}{2}\right)$$

where α is a constant. We have

$$\begin{aligned} E_n^{(1)} &= \left\langle \psi_n^{(0)}(x) \left| H' \right| \psi_n^{(0)}(x) \right\rangle = \frac{2}{a} \alpha \int_0^a \sin^2\left(\frac{n\pi}{a}x\right) \delta\left(x - \frac{a}{2}\right) dx \\ &= \frac{2}{a} \alpha \sin^2\left(\frac{n\pi}{2}\right) \\ &= \begin{cases} 0 & \text{if } n \text{ is even} \\ \frac{2}{a} \alpha & \text{if } n \text{ is odd} \end{cases} \end{aligned}$$

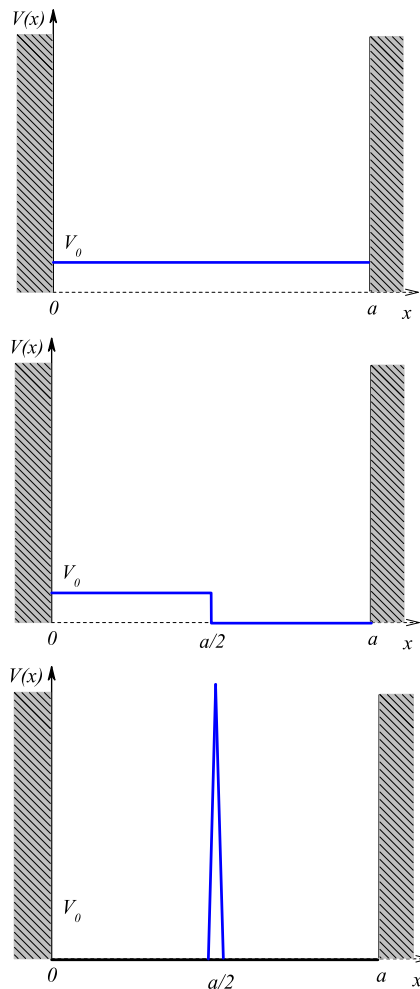


FIG. 3: Three ways of perturbation in infinite square well.

For even n the wave function is zero at the location of the perturbation ($x = \frac{a}{2}$), so it never “feels” H' . That is why the energies are not perturbed for even n . The correction to the ground state is

$$\psi_1^{(1)} = \sum_{m \neq 1} \frac{H'_{m1}}{E_1^{(0)} - E_m^{(0)}} \psi_m^{(0)},$$

To find this we need

$$\begin{aligned} H'_{m1} &= \langle \psi_m^{(0)} | H' | \psi_1^{(0)} \rangle \\ &= \frac{2}{a} \alpha \int_0^a \sin\left(\frac{m\pi}{a}x\right) \delta\left(x - \frac{a}{2}\right) \sin\left(\frac{\pi}{a}x\right) dx \\ &= \frac{2}{a} \alpha \sin\left(\frac{m\pi}{2}\right) \\ &= \begin{cases} \text{zero} & \text{if } m \text{ is even} \\ \text{non-zero} & \text{if } m \text{ is odd} \end{cases} \quad (m \neq n = 1) \end{aligned}$$

The first 3 nonzero terms will be $m = 3, 5, 7$. Meanwhile

$$E_1^{(0)} - E_m^{(0)} = \frac{\pi^2 \hbar^2}{2ma^2} (1 - m^2)$$

so

$$\begin{aligned} \psi_1^{(1)} &= \sum_{m=3,5,7,\dots} \frac{(2\alpha/a) \sin\left(\frac{m\pi}{2}\right)}{E_1^{(0)} - E_m^{(0)}} \psi_m^{(0)} \\ &= \frac{2}{a} \frac{2ma^2}{\pi^2 \hbar^2} \left(\frac{-1}{1-9} \psi_3^{(0)} + \frac{1}{1-25} \psi_5^{(0)} + \frac{-1}{1-49} \psi_7^{(0)} + \dots \right) \\ &= \frac{4m\alpha a}{\pi^2 \hbar^2} \sqrt{\frac{2}{a}} \left(\frac{1}{8} \sin\left(\frac{3\pi}{a}x\right) - \frac{1}{24} \sin\left(\frac{5\pi}{a}x\right) + \frac{1}{48} \sin\left(\frac{7\pi}{a}x\right) + \dots \right) \\ &= \frac{m\alpha}{\pi^2 \hbar^2} \sqrt{\frac{a}{2}} \left(\sin\left(\frac{3\pi}{a}x\right) - \frac{1}{3} \sin\left(\frac{5\pi}{a}x\right) + \frac{1}{6} \sin\left(\frac{7\pi}{a}x\right) + \dots \right) \end{aligned}$$

The second order corrections $E_n^{(2)}$ can be calculated by explicitly summing the series. We know

$$E_n^{(2)} = \sum_{m \neq n} \frac{|H'_{mn}|^2}{E_n^{(0)} - E_m^{(0)}}$$

with

$$\begin{aligned} H'_{mn} &= \left\langle \psi_m^{(0)} \left| H' \right| \psi_n^{(0)} \right\rangle \\ &= \frac{2}{a} \alpha \int_0^a \sin\left(\frac{m\pi}{a}x\right) \delta\left(x - \frac{a}{2}\right) \sin\left(\frac{n\pi}{a}x\right) dx \\ &= \frac{2}{a} \alpha \sin\left(\frac{m\pi}{2}\right) \sin\left(\frac{n\pi}{2}\right) \end{aligned}$$

which is 0 unless both m and n are odd in which case it is $\pm \frac{2\alpha}{a}$

$$\begin{aligned} E_n^{(2)} &= \sum_{m \neq n, \text{ odd}} \left(\frac{2\alpha}{a}\right)^2 \frac{1}{E_n^{(0)} - E_m^{(0)}} \\ &= \begin{cases} 0 & \text{if } n \text{ is even} \\ 2ma^2 \left(\frac{2\alpha}{a}\right)^2 \frac{1}{\pi^2 \hbar^2} \sum_{m \neq n, \text{ odd}} \frac{1}{n^2 - m^2}, & \text{if } n \text{ is odd} \end{cases} \end{aligned}$$

To sum the series, note that

$$\frac{1}{n^2 - m^2} = \frac{1}{2n} \left(\frac{1}{m+n} - \frac{1}{m-n} \right).$$

For $n = 1$

$$\begin{aligned} \sum &= \frac{1}{2} \sum_{3,5,7,\dots} \left(\frac{1}{m+1} - \frac{1}{m-1} \right) \\ &= \frac{1}{2} \left(\frac{1}{4} + \frac{1}{6} + \frac{1}{8} + \dots - \frac{1}{2} - \frac{1}{4} - \frac{1}{6} - \frac{1}{8} \dots \right) \\ &= \frac{1}{2} \left(-\frac{1}{2} \right) = -\frac{1}{4}. \end{aligned}$$

For $n = 3$

$$\begin{aligned} \sum &= \frac{1}{6} \sum_{1,5,7,\dots} \left(\frac{1}{m+3} - \frac{1}{m-3} \right) \\ &= \frac{1}{6} \left(\frac{1}{4} + \frac{1}{8} + \frac{1}{10} + \dots + \frac{1}{2} - \frac{1}{2} - \frac{1}{4} - \frac{1}{6} - \frac{1}{8} - \frac{1}{10} \dots \right) \\ &= \frac{1}{6} \left(-\frac{1}{6} \right) = -\frac{1}{36}. \end{aligned}$$

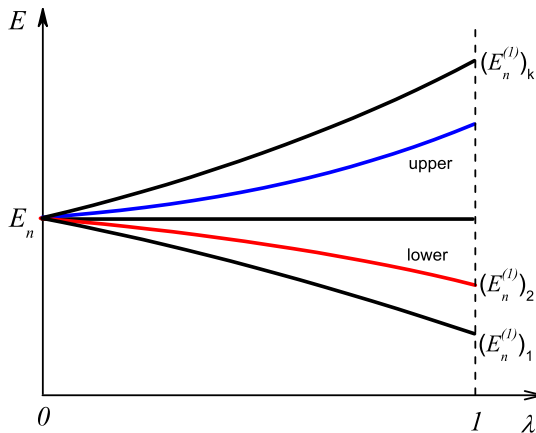


FIG. 4: Perturbation theory in degenerate case.

In general, there is perfect cancellation except for the “missing” term $\frac{1}{2n}$ in the first sum, so the total is $\frac{1}{2n} \left(-\frac{1}{2n}\right) = -\frac{1}{(2n)^2}$. Therefore

$$E_n^{(2)} = \begin{cases} 0 & \text{if } n \text{ is even} \\ -2m \left(\frac{\alpha}{\pi \hbar n}\right)^2 & \text{if } n \text{ is odd} \end{cases} .$$

Problem 3 Zeng Jinyan, p309-310, 1,4

II. STATIONARY PERTURBATION THEORY: DEGENERATE CASE

A. Degenerate perturbation theory

The perturbation theory says that, in the absence of perturbation ($\lambda = 0$), the system has its eigenfunctions and eigenvalues, $\psi_n^{(0)}$, $E_n^{(0)}$. If the perturbation is added ($\lambda = 1$), the wave function is modified $\psi_n^{(0)} \rightarrow \psi_n^{(0)} + \psi_n^{(1)} + \dots$, and the energy level is shifted from its original value $E_n^{(0)}$ to $E_n^{(0)} + E_n^{(1)} + \dots$. This should also be the case when the energy levels of \hat{H}_0 are degenerate.

Consider a degenerate \hat{H}_0 under perturbation. $\phi_1, \phi_2, \phi_3, \dots, \phi_k$ are orthonormal eigenstates of \hat{H}_0 corresponding to the same energy $E_n^{(0)}$. Typically, the perturbation \hat{H}' will “break” the degeneracy: As we increase λ (from 0 to 1), the common unperturbed energy E_n splits into k (Figure 4).

Let us focus on two states, the upper state and lower state. The essential problem is this: When we turn off the perturbation, the “upper” state reduces down to one linear combination of ϕ_i , and the “lower” state reduce to some other linear combination of ϕ_i , but we don’t know a priori what these “good” linear combinations will be. For this reason we can’t even calculate the first-order energy (Equation 8) because we don’t know what unperturbed states to use.

For the moment, therefore, let’s just write the “good” unperturbed states (the zeroth-order wave function) in the general form, that is, a certain superposition of ϕ_i

$$\psi_n^{(0)} = \sum_{i=1}^k c_i^{(0)} \phi_i$$

keeping the coefficients $c_i^{(0)}$ adjustable. The first order perturbation equation is now changed into a new form

$$\hat{H}_0 \psi_n^{(1)} + \hat{H}' \psi_n^{(0)} = E_n^{(0)} \psi_n^{(1)} + E_n^{(1)} \psi_n^{(0)}$$

$$\hat{H}_0 \psi_n^{(1)} + \hat{H}' \sum_{i=1}^k c_i^{(0)} \phi_i = E_n^{(0)} \psi_n^{(1)} + E_n^{(1)} \sum_{i=1}^k c_i^{(0)} \phi_i$$

Left multiplying the above equation by ϕ_1^* and integrating over the entire space,

$$\int \phi_1^* \left\{ \hat{H}_0 \psi_n^{(1)} + \hat{H}' \sum_{i=1}^k c_i^{(0)} \phi_i = E_n^{(0)} \psi_n^{(1)} + E_n^{(1)} \sum_{i=1}^k c_i^{(0)} \phi_i \right\} d^3 \mathbf{r}$$

and noticing the hermiticity of \hat{H}_0

$$\int \phi_1^* \hat{H}_0 \psi_n^{(1)} d^3 \mathbf{r} = E_n^{(0)} \int \phi_1^* \psi_n^{(1)} d^3 \mathbf{r}$$

which cancels the first term on the left and that on the right. We get

$$H'_{11} c_1^{(0)} + H'_{12} c_2^{(0)} + \dots + H'_{1k} c_k^{(0)} = E_n^{(1)} c_1^{(0)}$$

with

$$H'_{ij} = \int \phi_i^* \hat{H}' \phi_j d^3 \mathbf{r}.$$

Operating the above equation successively with ϕ_2, ϕ_3, \dots and ϕ_k , We get k equations

$$\begin{aligned} H'_{11} c_1^{(0)} + H'_{12} c_2^{(0)} + \dots + H'_{1k} c_k^{(0)} &= E_n^{(1)} c_1^{(0)} \\ H'_{21} c_1^{(0)} + H'_{22} c_2^{(0)} + \dots + H'_{2k} c_k^{(0)} &= E_n^{(1)} c_2^{(0)} \\ &\dots \dots \dots \\ H'_{k1} c_1^{(0)} + H'_{k2} c_2^{(0)} + \dots + H'_{kk} c_k^{(0)} &= E_n^{(1)} c_k^{(0)} \end{aligned}$$

The condition for the existence of non-trivial solutions of $\{c_i^{(0)}\}$ is

$$\begin{vmatrix} H'_{11} - E_n^{(1)} & H'_{12} & \dots & H'_{1k} \\ H'_{21} & H'_{22} - E_n^{(1)} & \dots & H'_{2k} \\ \dots & \dots & \dots & \dots \\ H'_{k1} & H'_{k2} & \dots & H'_{kk} - E_n^{(1)} \end{vmatrix} = 0$$

This equation has got a name **secular equation** after celestial mechanics. The k roots of the above algebraic equation give the k possible first order perturbation energy $E_n^{(1)}$ and the corresponding solutions $\{c_i^{(0)}\}$ give the corresponding zeroth-order wave functions $\psi_n^{(0)}$. Thus by solving the eigenvalue problem, *we obtain in one stroke both the first order energy shifts and the correct zeroth-order wavefunctions.*

The k degenerate eigenstates are reshuffled under the interaction of perturbation

$$\begin{aligned} (E_n^{(1)})_1, \quad (\psi_n^{(0)})_1 &= \sum_{i=1}^k c_{1i}^{(0)} \phi_i \\ (E_n^{(1)})_2, \quad (\psi_n^{(0)})_2 &= \sum_{i=1}^k c_{2i}^{(0)} \phi_i \\ &\dots \dots \dots \\ (E_n^{(1)})_k, \quad (\psi_n^{(0)})_k &= \sum_{i=1}^k c_{ki}^{(0)} \phi_i \end{aligned}$$

The above are reshuffled zeroth order wave function and first-order perturbation energy.

If higher order perturbation is needed, these reshuffled wave functions can be used as appropriate zeroth order wave functions, in this case, each reshuffled states can be considered as independent and results of non-degenerate perturbation theory can be used to it despite the existence of others degenerate states.

B. Stark effect of hydrogen atom

As an example of degenerate perturbation theory, let us study the effect of a uniform electric field on excited states of the hydrogen atom. As is well known, the bound state energy of the hydrogen atom depends only on the principal quantum number n . This leads to degeneracy for all but the ground state. To be specific, for the $n = 2$ level, there is an $l = 0$ state called $2s$ and three $l = 1$ ($m = 0, \pm 1$) states called $2p$, all with energy $-e^2/8a_0$. We label the four states as

$$\begin{aligned}\phi_1 &= \psi_{200} = R_{20}Y_{00} = R_{20}(r) \frac{1}{\sqrt{4\pi}} \\ \phi_2 &= \psi_{210} = R_{21}Y_{10} = R_{21}(r) \sqrt{\frac{3}{4\pi}} \cos \theta \\ \phi_3 &= \psi_{211} = R_{21}Y_{11} = -R_{21}(r) \sqrt{\frac{3}{8\pi}} \sin \theta e^{i\varphi} \\ \phi_4 &= \psi_{21-1} = R_{21}Y_{1-1} = R_{21}(r) \sqrt{\frac{3}{8\pi}} \sin \theta e^{-i\varphi}\end{aligned}$$

As we applied a uniform electric field ε in the z -direction, the Hamiltonian of the system is divided into two parts

$$\hat{H} = \hat{H}_0 + \hat{H}'$$

where

$$\hat{H}_0 = -\frac{\hbar^2}{2m} \nabla^2 + V(r)$$

and the appropriate perturbation operator is given by

$$\begin{aligned}\hat{H}' &= -\mathbf{D} \cdot \boldsymbol{\varepsilon}, \quad \mathbf{D} = -e\mathbf{r} \\ \hat{H}' &= e\mathbf{r} \cdot \boldsymbol{\varepsilon} = e\varepsilon r \cos \theta\end{aligned}$$

The matrix elements

$$\begin{aligned}H'_{ii} &= e\varepsilon \int \phi_i^* r \cos \theta \phi_i d^3\mathbf{r} = e\varepsilon \int R_{nl}(r) Y_{lm}^*(r \cos \theta) R_{nl}(r) Y_{lm} r^2 dr d\Omega \\ &= e\varepsilon \int_0^\infty R_{nl}^2(r) r^3 dr \iint Y_{lm}^*(\cos \theta) Y_{lm} d\Omega\end{aligned}$$

Now we notice that the integration should be unchanged under the reverse of the coordinate $\mathbf{r} \rightarrow -\mathbf{r}$, which corresponds to $\theta \rightarrow \pi - \theta$, $\varphi \rightarrow \pi - \varphi$, $r \rightarrow r$

$$\begin{aligned}\iint Y_{lm}^* Y_{lm} \cos \theta d\Omega &= \iint Y_{lm}^* Y_{lm} \cos(\pi - \theta) d\Omega \\ &= -\iint Y_{lm}^* Y_{lm} \cos \theta d\Omega\end{aligned}$$

So $H'_{ii} = 0$ for all i . Moreover, owing to the orthogonality of $e^{im\varphi}$, only $H'_{12} = H'_{21}^*$ is different from zero. To find this non-zero matrix element, we calculate $\langle \phi_1 | e\varepsilon r \cos \theta | \phi_2 \rangle$ with

$$\begin{aligned}\phi_1 &= \psi_{200} = \frac{1}{4\sqrt{2\pi}} \left(\frac{1}{a_0}\right)^{\frac{3}{2}} \left(2 - \frac{r}{a_0}\right) e^{-\frac{r}{2a_0}} \\ \phi_2 &= \psi_{210} = \frac{1}{4\sqrt{2\pi}} \left(\frac{1}{a_0}\right)^{\frac{3}{2}} \left(\frac{r}{a_0}\right) e^{-\frac{r}{2a_0}} \cos \theta\end{aligned}$$

and get

$$\begin{aligned}H'_{12} &= \frac{e\varepsilon}{32\pi} \frac{1}{a_0^3} \int_0^\infty \left(2 - \frac{r}{a_0}\right) \left(\frac{r}{a_0}\right) e^{-\frac{r}{a_0}} r^3 dr \iint \cos^2 \theta \sin \theta d\theta d\varphi \\ &= \frac{e\varepsilon}{32\pi} (-72a_0) \frac{4\pi}{3} = -3e\varepsilon a_0\end{aligned}$$

(You might think H'_{34} is non-zero, but $H'_{34} \propto \int_0^\pi \sin^2 \theta \cos \theta \sin \theta d\theta = 0$). The secular equation for this system is

$$\begin{vmatrix} 0 - E_2^{(1)} & -3e\epsilon a_0 & 0 & 0 \\ -3e\epsilon a_0 & 0 - E_2^{(1)} & 0 & 0 \\ 0 & 0 & 0 - E_2^{(1)} & 0 \\ 0 & 0 & 0 & 0 - E_2^{(1)} \end{vmatrix} = 0$$

The four eigenvalues for $E_2^{(1)}$ are $3e\epsilon a_0, -3e\epsilon a_0, 0, 0$, respectively. It is straightforward to find the corresponding eigenfunctions by solving the equation

$$\begin{pmatrix} -E_2^{(1)} & -3e\epsilon a_0 & 0 & 0 \\ -3e\epsilon a_0 & -E_2^{(1)} & 0 & 0 \\ 0 & 0 & -E_2^{(1)} & 0 \\ 0 & 0 & 0 & -E_2^{(1)} \end{pmatrix} \begin{pmatrix} c_1^{(0)} \\ c_2^{(0)} \\ c_3^{(0)} \\ c_4^{(0)} \end{pmatrix} = 0$$

For $E_2^{(0)} + 3e\epsilon a_0$, the normalized eigenvector is

$$\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix}$$

which gives the new zeroth order wavefunction

$$(\psi_2^{(0)})_1 = \frac{1}{\sqrt{2}}\psi_{200} - \frac{1}{\sqrt{2}}\psi_{210}$$

For $E_2^{(0)} - 3e\epsilon a_0$, the normalized eigenvector is

$$\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}$$

with the new zeroth order wavefunction

$$(\psi_2^{(0)})_2 = \frac{1}{\sqrt{2}}\psi_{200} + \frac{1}{\sqrt{2}}\psi_{210}$$

For the two-fold degenerate eigenvalue $E_2^{(0)} + 0$, we have

$$\begin{pmatrix} 0 \\ 0 \\ x \\ y \end{pmatrix}$$

while x, y are undetermined. We thus take the original wavefunctions

$$\begin{aligned} (\psi_2^{(0)})_3 &= \psi_{21,1} = R_{21}Y_{1,1} \\ (\psi_2^{(0)})_4 &= \psi_{21,-1} = R_{21}Y_{1,-1} \end{aligned}$$

The presence of the electric field removes only partially the degeneracy. Schematically the energy levels are split into three as shown in Figure 5. Notice the shift is linear in the applied electric field strength, which is known as the **linear Stark effect**.

Problem 4 *Griffiths, page, 266, Problem 6.9*

Problem 5 *Zeng Jinyan, p309-310, 2,5*

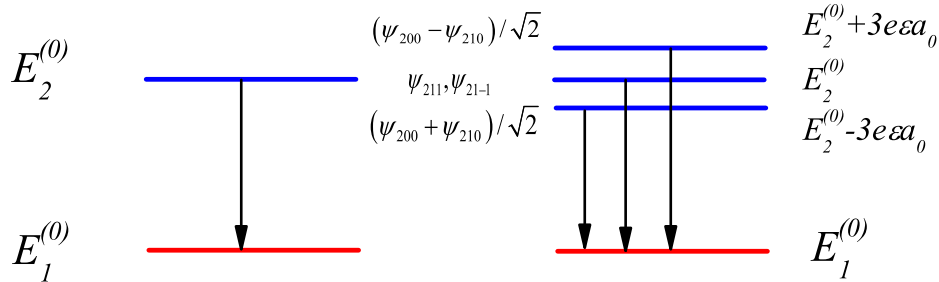


FIG. 5: Linear Stark effect of hydrogen atom.

III. VARIATIONAL METHOD

The perturbation theory developed in the previous section is, of course, of no help unless we already know exact solution to a problem described by \hat{H}_0 . The variational method we now discuss is very useful for estimating the ground state energy E_0 when such exact solutions are not available. We should keep in mind the drastic difference between variational method here and variational principle in classical mechanics.

Variational principle in classical mechanics is a fundamental principle in classical mechanics. Variational method in quantum mechanics is mere an approximation method in finding out the eigenvalue of \hat{H} in some instances. We never bestow it too much importance.

The *basic idea* is try to conjecture a wave function ψ_0 for a given system \hat{H} , calculate the expectation value of \hat{H} using this wave function, and finally optimize the calculated result.

The physical principles here are: (1) the given \hat{H} should have its own eigenvalues and eigenfunctions (not known due to practical difficulties)

$$E_1, E_2, \dots E_n \dots$$

$$\psi_1, \psi_2, \dots \psi_n \dots$$

(2) the conjectured wave function ψ_0 can be expanded upon $\{\psi_n\}$, i.e. $\psi_0 = \sum_n c_n \psi_n$, and (3) the expectation value calculated using ψ_0 is always larger than E_1

$$\langle H \rangle = \sum_n |c_n|^2 E_n \geq E_1$$

The variational method does not tell us what kind of trial wavefunctions are to be used to estimate the ground state energy. Quite often we must appeal to physical intuition - for example, the asymptotic behavior of wavefunction at large distances. What we do in practice is to characterize trial wavefunctions by one or more parameters $\psi_0(\mathbf{r}, \lambda_1, \lambda_2 \dots)$ and compute $\langle H \rangle$ as a function of $\lambda_1, \lambda_2 \dots$.

$$\langle H(\lambda_1, \lambda_2 \dots) \rangle = \frac{\int \psi_0^*(\mathbf{r}, \lambda_1, \lambda_2 \dots) \hat{H} \psi_0(\mathbf{r}, \lambda_1, \lambda_2 \dots) d^3 \mathbf{r}}{\int \psi_0^*(\mathbf{r}, \lambda_1, \lambda_2 \dots) \psi_0(\mathbf{r}, \lambda_1, \lambda_2 \dots) d^3 \mathbf{r}}$$

We then minimize $\langle H(\lambda_1, \lambda_2 \dots) \rangle$ by (1) setting the derivative with respect to the parameters all zero, namely

$$\frac{\partial \langle H(\lambda_1, \lambda_2 \dots) \rangle}{\partial \lambda_1} = 0, \frac{\partial \langle H(\lambda_1, \lambda_2 \dots) \rangle}{\partial \lambda_2} = 0, \dots$$

(2) determining the optimum values of $\lambda_1, \lambda_2 \dots$, and (3) substituting them back to the expression for $\langle H(\lambda_1, \lambda_2 \dots) \rangle$, which is the upper bound of the ground state energy of the system.

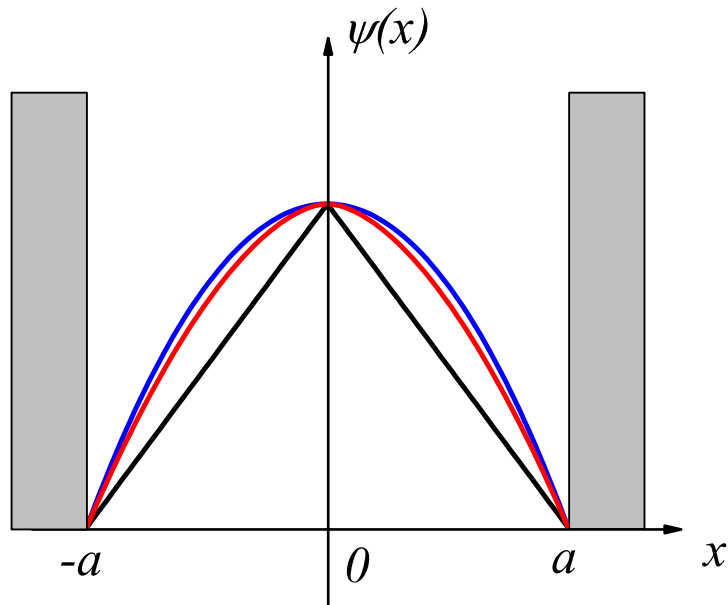


FIG. 6: Three trial functions for the variational method: Triangular (Black), Parabola (Blue), and more sophisticated one (Red).

A. Ground state energy of 1D infinite square well

As an example, we attempt to estimate the ground state energy of one dimensional infinite square well

$$V(x) = \begin{cases} 0, & \text{for } -a < x < a \\ +\infty, & \text{others} \end{cases}$$

using three different trial functions. The exact results are, of course, well known

$$\psi_1 = \sqrt{\frac{1}{a}} \cos\left(\frac{\pi x}{2a}\right)$$

$$E_1 = \frac{\pi^2 \hbar^2}{8ma^2}$$

But suppose we did not know these. Evidently the wavefunction must vanish at $x = \pm a$. Furthermore for ground state the wavefunction cannot have any wiggles. We try three functions that satisfy both of these two requirements.

- Triangular trial wavefunction (Figure 6)

$$\psi(x) = \begin{cases} A(a+x), & \text{if } -a \leq x \leq 0 \\ A(a-x), & \text{if } 0 \leq x \leq a \\ 0, & \text{others} \end{cases}$$

where A is determined by normalization

$$1 = |A|^2 \left[\int_{-a}^0 (a+x)^2 dx + \int_0^a (a-x)^2 dx \right] = |A|^2 \frac{2a^3}{3}$$

$$A = \sqrt{\frac{3}{2a^3}}$$

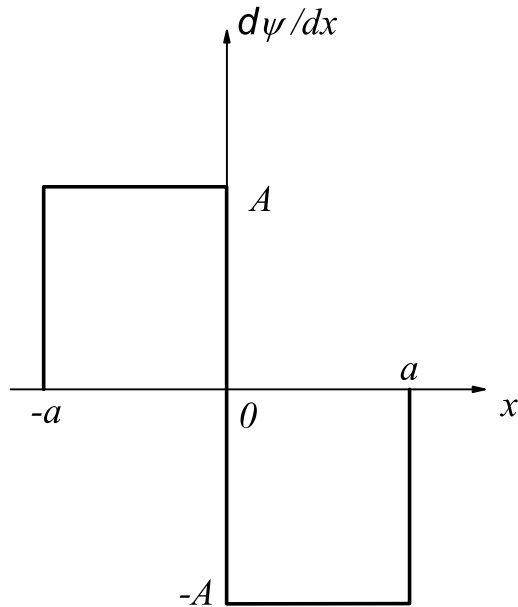


FIG. 7: The derivative of the triangular function.

In this case

$$\frac{d\psi}{dx} = \begin{cases} A, & \text{if } -a \leq x \leq 0 \\ -A, & \text{if } 0 \leq x \leq a \\ 0, & \text{others} \end{cases}$$

as indicated in Figure 7. Now the derivative of a step function is a delta function

$$\frac{d^2\psi}{dx^2} = A\delta(x+a) - 2A\delta(x) + A\delta(x-a)$$

and hence

$$\begin{aligned} \langle H \rangle &= -\frac{\hbar^2 A}{2m} \int_{-a}^a (\delta(x+a) - 2\delta(x) + \delta(x-a)) \psi(x) dx \\ &= -\frac{\hbar^2 A}{2m} (\psi(-a) - 2\psi(0) + \psi(a)) \\ &= \frac{\hbar^2 A}{m} \psi(0) = \frac{\hbar^2 A^2 a}{m} \\ &= \frac{3\hbar^2}{2ma^2} = \frac{12}{\pi^2} \frac{\pi^2 \hbar^2}{8ma^2} = 1.2159 E_1 \end{aligned}$$

which is larger than the exact ground state energy.

- Parabola trial wavefunction going through $x = \pm a$

$$\psi(x) = a^2 - x^2$$

where we have not bothered to normalize it. Here there is no variational parameter. We can compute $\langle H \rangle$ as

follows

$$\begin{aligned}\langle H \rangle &= \frac{-\frac{\hbar^2}{2m} \int_{-a}^a (a^2 - x^2) \frac{d^2}{dx^2} (a^2 - x^2) dx}{\int_{-a}^a (a^2 - x^2)^2 dx} \\ &= -\frac{\hbar^2}{2m} \frac{-2\frac{4}{3}a^3}{\frac{16}{15}a^5} = \frac{5\hbar^2}{4ma^2} = \frac{10\hbar^2}{8ma^2} \\ &= \frac{10}{\pi^2} \frac{\pi^2 \hbar^2}{8ma^2} = 1.0132E_1\end{aligned}$$

It is remarkable that with such a simple trial function we can come within 1.3% of the true ground state energy.

- A much better result can be obtained if we use a more sophisticated trial function. We try

$$\psi(x) = |a|^\lambda - |x|^\lambda \quad (14)$$

where λ is now regarded as a variational parameter. Straightforward algebra gives

$$\langle H \rangle = \frac{(\lambda + 1)(2\lambda + 1)}{2\lambda - 1} \left(\frac{\hbar^2}{4ma^2} \right)$$

which has a minimum at

$$\begin{aligned}\frac{\partial \langle H \rangle}{\partial \lambda} &= \frac{\partial}{\partial \lambda} \left(\frac{(\lambda + 1)(2\lambda + 1)}{2\lambda - 1} \right) \\ &= \frac{4\lambda^2 - 4\lambda - 5}{(2\lambda - 1)^2} = 0 \\ &\rightarrow \lambda = \frac{1}{2} (1 + \sqrt{6}) \approx 1.72\end{aligned}$$

not far from $n = 2$ (parabola) considered earlier. This gives

$$\begin{aligned}\langle H \rangle_{\min} &= \frac{(3 + \sqrt{6})(2 + \sqrt{6})}{\sqrt{6}} \left(\frac{\hbar^2}{8ma^2} \right) \\ &= \frac{5 + 2\sqrt{6}}{\pi^2} E_1 \approx 1.00298E_1\end{aligned}$$

So the variational method with (14) gives the correct ground state energy with in 0.3% - a fantastic result considering the simplicity of the trial function used.

B. Ground state energy of Helium atom

1. Hamiltonian and trial wave function

The helium atom (Figure 8) consists of two electrons in orbit around a nucleus containing two protons (also some neutrons, which are irrelevant to our purpose). The Hamiltonian for this system (ignoring fine structure and smaller corrections) is

$$\hat{H} = \frac{1}{2m} \hat{\mathbf{p}}_1^2 + \frac{1}{2m} \hat{\mathbf{p}}_2^2 + V(r_1) + V(r_2) + \frac{e^2}{|\mathbf{r}_1 - \mathbf{r}_2|}.$$

Our problem is to calculate the ground-state energy, E_g - the amount of energy it would take to strip off the two electrons. (Given E_g it is easy to figure out the "ionization energy" required to remove a single electron.) E_g has been measured very accurately in the laboratory

$$E_g = -2.904 \frac{e^2}{a_0}$$

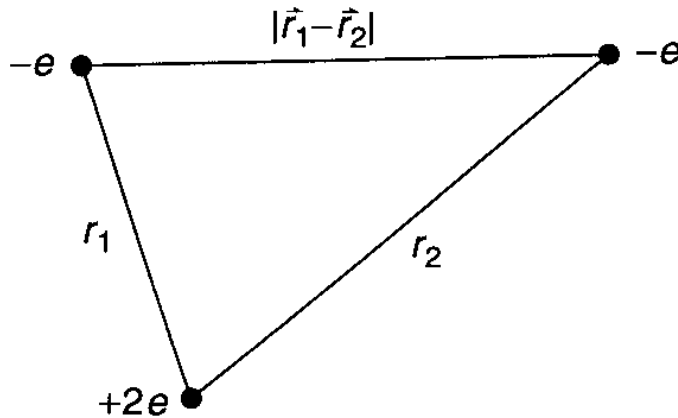


FIG. 8: The Helium atom.

This is the number we would like to reproduce theoretically. The Schrödinger equation reads

$$\hat{H}\psi(\mathbf{r}_1, \mathbf{r}_2) = E\psi(\mathbf{r}_1, \mathbf{r}_2).$$

It is curious that such a simple and important problem has no known exact solution. The trouble comes from the electron-electron repulsion,

$$V_{12} = \frac{e^2}{|\mathbf{r}_1 - \mathbf{r}_2|}.$$

If we ignore this term altogether, \hat{H} splits into two independent hydrogen Hamiltonians (only with a nuclear charge of $2e$, instead of e); the exact solution is just the product of hydrogenic wave functions

$$\psi(\mathbf{r}_1, \mathbf{r}_2) = \psi(\mathbf{r}_1)\phi(\mathbf{r}_2)$$

with energy

$$E = E_1 + E_2$$

which satisfy

$$\begin{aligned} \left(\frac{1}{2m}\hat{\mathbf{p}}_1^2 + V(r_1)\right)\psi(\mathbf{r}_1) &= E_1\psi(\mathbf{r}_1), & \hat{H}_1\psi(\mathbf{r}_1) &= E_1\psi(\mathbf{r}_1) \\ \left(\frac{1}{2m}\hat{\mathbf{p}}_2^2 + V(r_2)\right)\phi(\mathbf{r}_2) &= E_2\phi(\mathbf{r}_2), & \hat{H}_2\phi(\mathbf{r}_2) &= E_2\phi(\mathbf{r}_2) \end{aligned}$$

and

$$V(r) = -\frac{Ze^2}{r}.$$

Explicitly they are

$$\begin{aligned} \psi(\mathbf{r}_1) &= \sqrt{\frac{Z^3}{\pi a_0^3}} e^{-\frac{Zr_1}{a_0}}, & \phi(\mathbf{r}_2) &= \sqrt{\frac{Z^3}{\pi a_0^3}} e^{-\frac{Zr_2}{a_0}} \\ E_1 = E_2 &= -\frac{1}{2} \frac{Z^2 m e^4}{\hbar^2} = -\frac{Z^2 e^2}{2a_0}, & a_0 &= \frac{\hbar^2}{m e^2} \end{aligned}$$

As a starting point, we take the above wave function as the trial function

$$\psi(\mathbf{r}_1, \mathbf{r}_2) = \sqrt{\frac{Z^3}{\pi a_0^3}} e^{-\frac{Z}{a_0} r_1} \sqrt{\frac{Z^3}{\pi a_0^3}} e^{-\frac{Z}{a_0} r_2} = \frac{Z^3}{\pi a_0^3} e^{-\frac{Z}{a_0} (r_1 + r_2)}$$

Screening effect due to the existence of another electron will effectively decrease the charge Z . But we are not sure in advance for how much. It is thus reasonable to regard Z as a variational parameter λ . The trial function is thus

$$\psi(\mathbf{r}_1, \mathbf{r}_2) = \frac{\lambda^3}{\pi a_0^3} e^{-\frac{\lambda}{a_0}(r_1+r_2)}$$

The constant $\lambda^3/\pi a_0^3$ is still kept here to make sure the trial wave function normalized.

2. The expectation value

The expectation value of Hamiltonian is calculated as

$$\begin{aligned} \langle H(\lambda) \rangle &= \iiint \iiint \frac{\lambda^3}{\pi a_0^3} e^{-\frac{\lambda}{a_0}(r_1+r_2)} \left(-\frac{\hbar^2}{2m} \nabla_1^2 - \frac{\hbar^2}{2m} \nabla_2^2 - \frac{Ze^2}{r_1} - \frac{Ze^2}{r_2} \right. \\ &\quad \left. + \frac{e^2}{|\mathbf{r}_1 - \mathbf{r}_2|} \right) \frac{\lambda^3}{\pi a_0^3} e^{-\frac{\lambda}{a_0}(r_1+r_2)} d^3\mathbf{r}_1 d^3\mathbf{r}_2 \end{aligned}$$

The kinetic energy term for electron 1 is

$$\begin{aligned} &\int \dots \int \frac{\lambda^3}{\pi a_0^3} e^{-\frac{\lambda}{a_0}(r_1+r_2)} \left(-\frac{\hbar^2}{2m} \nabla_1^2 \right) \frac{\lambda^3}{\pi a_0^3} e^{-\frac{\lambda}{a_0}(r_1+r_2)} d^3\mathbf{r}_1 d^3\mathbf{r}_2 \\ &= \iiint \frac{\lambda^3}{\pi a_0^3} e^{-\frac{\lambda}{a_0}r_1} \left(-\frac{\hbar^2}{2m} \nabla_1^2 \right) e^{-\frac{\lambda}{a_0}r_1} d^3\mathbf{r}_1 \\ &= \frac{\lambda^3}{\pi a_0^3} \iiint e^{-\frac{\lambda}{a_0}r_1} \left\{ -\frac{\hbar^2}{2m} \left(\frac{2}{r_1} \left(-\frac{\lambda}{a_0} \right) + \frac{\lambda^2}{a_0^2} \right) e^{-\frac{\lambda}{a_0}r_1} \right\} r_1^2 dr_1 d\Omega \\ &= -\frac{\hbar^2}{2m} \frac{4\lambda^2}{a_0^2} \int_0^\infty e^{-2x} (-2x + x^2) dx = \frac{\hbar^2}{2m} \frac{\lambda^2}{a_0^2} \end{aligned}$$

where we have used

$$\begin{aligned} x &= \frac{\lambda}{a_0} r_1, & \nabla_1^2 &= \frac{1}{r_1^2} \frac{d}{dr_1} \left(r_1^2 \frac{d}{dr_1} \right) \\ \int_0^{+\infty} x e^{-2x} dx &= e^{-2x} \left(-\frac{1}{4} - \frac{x}{2} \right)_0^{+\infty} = \frac{1}{4} \\ \int_0^{+\infty} x^2 e^{-2x} dx &= -\frac{1}{4} e^{-2x} (1 + 2x + 2x^2)_0^\infty = \frac{1}{4} \end{aligned}$$

so

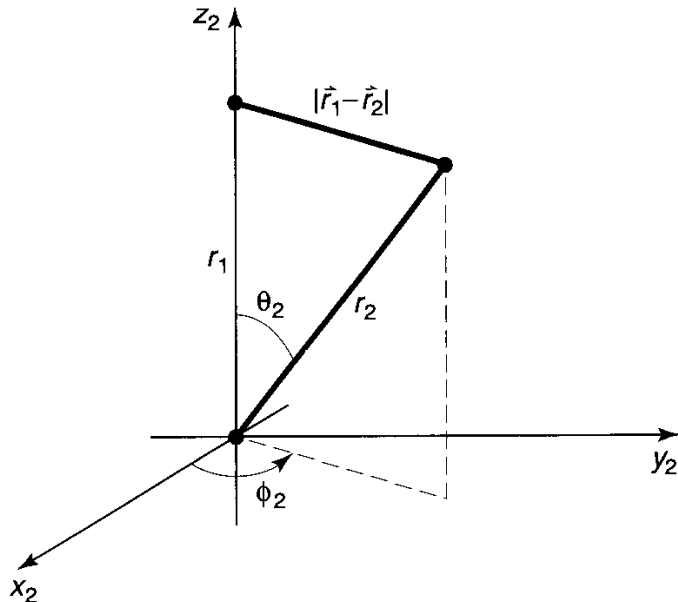
$$\langle H(\lambda) \rangle_k = \frac{\hbar^2}{m} \frac{\lambda^2}{a_0^2} = \frac{\hbar^2}{m} \frac{\lambda^2}{a_0} \frac{1}{\hbar^2/m e^2} = \frac{\lambda^2 e^2}{a_0}. \quad (15)$$

The potential energy term between electron 1 and nucleus is

$$\begin{aligned} &\int \dots \int \frac{\lambda^3}{\pi a_0^3} e^{-\frac{\lambda}{a_0}(r_1+r_2)} \left(-\frac{Ze^2}{r_1} \right) \frac{\lambda^3}{\pi a_0^3} e^{-\frac{\lambda}{a_0}(r_1+r_2)} d^3\mathbf{r}_1 d^3\mathbf{r}_2 \\ &= \frac{\lambda^3}{\pi a_0^3} \iiint e^{-\frac{\lambda}{a_0}r_1} \left(-\frac{Ze^2}{r_1} \right) e^{-\frac{\lambda}{a_0}r_1} r_1^2 dr_1 d\Omega \\ &= (-Ze^2) \frac{4\lambda}{a_0} \int_0^\infty e^{-2x} x dx = -\frac{\lambda Ze^2}{a_0} \end{aligned}$$

so

$$\langle H(\lambda) \rangle_p = -\frac{2\lambda Ze^2}{a_0} \quad (16)$$

FIG. 9: Choice of coordinates for the r_2 -integral.

Finally we calculate the electrostatic interaction energy between two electrons

$$\langle H(\lambda) \rangle_{int} = \iiint \iiint \left(\frac{\lambda^3}{\pi a_0^3} \right)^2 e^{-\frac{2\lambda}{a_0}(r_1+r_2)} \frac{e^2}{|\mathbf{r}_1 - \mathbf{r}_2|} d^3\mathbf{r}_1 d^2\mathbf{r}_2$$

We shall do the \mathbf{r}_2 integral first, for this purpose \mathbf{r}_1 is fixed, and we may as well orient the \mathbf{r}_2 coordinate system so that the polar axis lies along \mathbf{r}_1 (see Figure 9). By the law of cosines,

$$|\mathbf{r}_1 - \mathbf{r}_2| = \sqrt{r_1^2 + r_2^2 - 2r_1r_2 \cos \theta_2}$$

and hence

$$\begin{aligned} I_2 &= \iiint \frac{e^{-\frac{2\lambda}{a_0}r_2}}{|\mathbf{r}_1 - \mathbf{r}_2|} d^3\mathbf{r}_2 \\ &= \iiint \frac{e^{-\frac{2\lambda}{a_0}r_2}}{\sqrt{r_1^2 + r_2^2 - 2r_1r_2 \cos \theta_2}} r_2^2 \sin \theta_2 dr_2 d\theta_2 d\phi_2 \end{aligned}$$

The ϕ_2 integral is trivial (2π); the θ_2 integral is

$$\begin{aligned} \int_0^\pi \frac{\sin \theta_2}{\sqrt{r_1^2 + r_2^2 - 2r_1r_2 \cos \theta_2}} d\theta_2 &= \frac{\sqrt{r_1^2 + r_2^2 - 2r_1r_2 \cos \theta_2}}{r_1r_2} \Big|_0^\pi \\ &= \frac{\sqrt{r_1^2 + r_2^2 + 2r_1r_2} - \sqrt{r_1^2 + r_2^2 - 2r_1r_2}}{r_1r_2} \\ &= \frac{1}{r_1r_2} [(r_1 + r_2) - |r_1 - r_2|] = \begin{cases} \frac{2}{r_1}, & \text{if } r_2 < r_1 \\ \frac{2}{r_2}, & \text{if } r_2 > r_1 \end{cases} \end{aligned}$$

Thus

$$\begin{aligned} I_2 &= 4\pi \left(\frac{1}{r_1} \int_0^{r_1} e^{-\frac{2\lambda}{a_0}r_2} r_2^2 dr_2 + \int_{r_1}^{+\infty} e^{-\frac{2\lambda}{a_0}r_2} r_2 dr_2 \right) \\ &= 4\pi \frac{a_0^2}{\lambda^2} \left(\frac{1}{x} \int_0^x e^{-2y} y^2 dy + \int_x^{+\infty} e^{-2y} y dy \right) \end{aligned}$$

with

$$x = \frac{\lambda}{a_0} r_1, \quad y = \frac{\lambda}{a_0} r_2$$

We have

$$\begin{aligned} \frac{1}{x} \int_0^x e^{-2y} y^2 dy &= \frac{1}{x} \left(\frac{1}{4} - \frac{1}{2} x e^{-2x} - \frac{1}{2} x^2 e^{-2x} - \frac{1}{4} e^{-2x} \right) \\ \int_x^{+\infty} e^{-2y} y dy &= \frac{1}{4} e^{-2x} + \frac{1}{2} x e^{-2x} \end{aligned}$$

then

$$\begin{aligned} I_2 &= 4\pi \frac{a_0^2}{\lambda^2} \left[\frac{1}{4x} - \frac{1}{4} e^{-2x} - \frac{1}{4x} e^{-2x} \right] \\ &= \frac{\pi a_0^2}{\lambda^2} \left[\frac{1}{x} - e^{-2x} - \frac{1}{x} e^{-2x} \right] \\ &= \frac{\pi a_0^3}{\lambda^3 r_1} \left[1 - \left(1 + \frac{\lambda r_1}{a_0} \right) e^{-2 \frac{\lambda}{a_0} r_1} \right] \end{aligned}$$

It follows that $\langle H(\lambda) \rangle_{int}$ is equal to

$$\begin{aligned} &\left(\frac{\lambda^3}{\pi a_0^3} \right)^2 e^2 \iiint e^{-\frac{2\lambda}{a_0} r_1} I_2 d^3 \mathbf{r}_1 \\ &= \frac{\lambda^3 e^2}{\pi a_0^3} \iiint e^{-\frac{2\lambda}{a_0} r_1} \left[1 - \left(1 + \frac{\lambda r_1}{a_0} \right) e^{-2 \frac{\lambda}{a_0} r_1} \right] r_1 dr_1 \sin \theta_1 d\theta_1 d\phi_1 \\ &= \frac{4\lambda^3 e^2}{a_0^3} \frac{a_0^2}{\lambda^2} \int_0^{+\infty} e^{-2x} (1 - (1+x) e^{-2x}) x dx \end{aligned}$$

The integral is 5/32 and we obtain

$$\langle H(\lambda) \rangle_{int} = \frac{5\lambda e^2}{8a_0} \quad (17)$$

The expectation value of \hat{H} is evidently the summation of equation (15), (16) and (17)

$$\begin{aligned} \langle H(\lambda) \rangle &= \frac{\lambda^2 e^2}{a_0} - 2\lambda Z \frac{e^2}{a_0} + \frac{5}{8} \lambda \frac{e^2}{a_0} \\ &= \left(\lambda^2 - 2\lambda Z + \frac{5}{8} \lambda \right) \frac{e^2}{a_0} \end{aligned}$$

3. Optimization

According to the variational principle, this quantity exceeds E_g for any value of Z . The lowest upper bound occurs when $\langle H(\lambda) \rangle$ is minimized

$$\frac{\partial \langle H(\lambda) \rangle}{\partial \lambda} = 2\lambda - 2Z + \frac{5}{8} = 0$$

from which it follows that

$$\lambda = Z - \frac{5}{16} \approx 1.69$$

This is a reasonable result; it tells us that the other electron partially screens the nucleus, reducing its effective charge from 2 down to 1.69. Putting in this value for Z , we find

$$\begin{aligned} \langle H(\lambda) \rangle_{\min} &= \left(\left(Z - \frac{5}{16} \right)^2 - 2Z \left(Z - \frac{5}{16} \right) + \frac{5}{8} \left(Z - \frac{5}{16} \right) \right) \frac{e^2}{a_0} \\ &= \left(-Z^2 + \frac{5}{8} Z - \frac{25}{256} \right) \frac{e^2}{a_0} = -2.85 \frac{e^2}{a_0} \end{aligned}$$

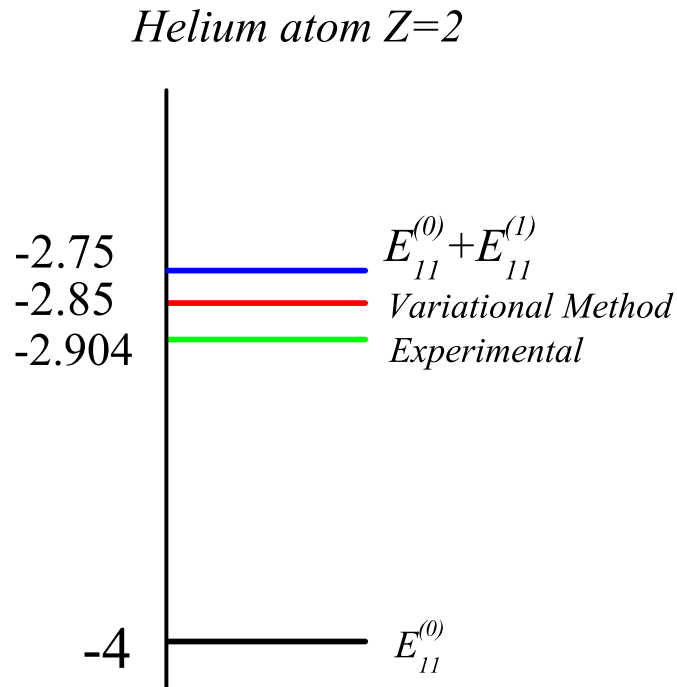


FIG. 10: Ground state of helium atom: unperturbed (Black), experimental result (Green), perturbation result (Blue), and variational result (Red).

4. Perturbation result

Suppose the interaction between the two electrons vanish, the zeroth order (unperturbed) energy is the sum of two ground state energies of a single electron in the field of the nucleus Ze

$$E_{11}^{(0)} = -Z^2 \frac{1}{2} \frac{e^2}{a_0} - Z^2 \frac{1}{2} \frac{e^2}{a_0} = -Z^2 \frac{e^2}{a_0} = -4 \frac{e^2}{a_0}$$

One can treat the interaction term $e^2/|\mathbf{r}_1 - \mathbf{r}_2|$ as perturbation \hat{H}' , and the first order correction result gives

$$E_{11}^{(1)} = \iiint \iiint \left(\frac{Z^3}{\pi a_0^3} \right)^2 e^{-\frac{2Z}{a_0}(r_1+r_2)} \frac{e^2}{|\mathbf{r}_1 - \mathbf{r}_2|} d^3\mathbf{r}_1 d^3\mathbf{r}_2 = \frac{5Z}{8} \frac{e^2}{a_0}$$

so

$$E_{11} = E_{11}^{(0)} + E_{11}^{(1)} = \left(-Z^2 + \frac{5Z}{8} \right) \frac{e^2}{a_0} = -2.75 \frac{e^2}{a_0}$$

These results are compared in Figure 10. The perturbation result is poor because the interaction between two electrons is not necessarily small compared with the Coulomb attraction between the nucleus and the electrons.

Problem 6 Zeng Jinyan, page 311, 7 (Griffiths, page 294, Example 7.1), 9(Griffiths, page 310, Problem 7.13)), 10

Problem 7 (a) Use the variational principle to prove that first-order non-degenerate perturbation theory always overestimates (or at any rate never underestimates) the ground state energy. (b) In view of (a), you would expect that the second-order correction to the ground state is always negative. Confirm that this is indeed the case, by examining Equation (13).

IV. TIME-DEPENDENT PERTURBATION THEORY

A. Time evolution of quantum state

So far we have been concerned with Hamiltonians that do not contain time explicitly (essentially the potential energy $V(\mathbf{r}, t) = V(\mathbf{r})$), which might properly be called **quantum statics**. In that case the law governing the evolution of the system is the time-dependent Schrödinger equation

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

which can be treated by separation of variables

$$\Psi(\mathbf{r}, t) = \psi(\mathbf{r})f(t).$$

We find the stationary solution ψ by solving the time-independent Schrödinger equation

$$\hat{H}\psi_n = E_n\psi_n,$$

then attach the time evolution factor

$$f(t) = e^{-\frac{i}{\hbar}E_n t}$$

and the general solution is a linear combination

$$\Psi(\mathbf{r}, t) = \sum_n a_n \psi_n(\mathbf{r}) e^{-\frac{i}{\hbar}E_n t}.$$

The time dependence of here cancels out when we construct the physically relevant quantity $|\Psi|^2$, all probabilities and expectation values are constant in time. For example, suppose we have an initial state for the harmonic oscillator

$$\psi(x) = \frac{1}{\sqrt{2}}\psi_0(x) + \frac{1}{\sqrt{2}}\psi_1(x)$$

we immediately know the state at time t

$$\Psi(x, t) = \frac{1}{\sqrt{2}}\psi_0(x)e^{-i\omega t/2} + \frac{1}{\sqrt{2}}\psi_1(x)e^{-i3\omega t/2}.$$

Similarly the initial state for hydrogen atom

$$\psi(r, \theta, \varphi) = \frac{1}{\sqrt{2}}\psi_{100} + \frac{1}{\sqrt{2}}\psi_{211}$$

will evolve with time as

$$\Psi(\mathbf{r}, t) = \frac{1}{\sqrt{2}}\psi_{100}e^{\frac{i}{\hbar}E_1 t} + \frac{1}{\sqrt{2}}\psi_{211}e^{\frac{i}{\hbar}E_2 t}.$$

B. Time-dependent perturbation theory - quantum transition probability

If we want to allow for *transitions* (quantum jumps, as they are sometimes called) between one energy level and another, we must introduce a *time-dependent* Hamiltonian (**quantum dynamics**). There are precious few exactly solvable problems in quantum dynamics. However, if the time-dependent portion of the Hamiltonian is small compared to the time-independent part, it can be treated as a perturbation. The purpose of this section is to develop time-dependent perturbation theory, and study its most important application: the emission or absorption of radiation by an atom.

We consider a Hamiltonian \hat{H} which can be split into two parts

$$\hat{H} = \hat{H}_0 + \hat{H}'(t)$$

where \hat{H}_0 does not contain time explicitly. Assume the problem for $\hat{H}'(t) = 0$ has already been solved in the sense that the energy eigenstates and the energy eigenvalues are completely known

$$\hat{H}_0\psi_n = E_n\psi_n \quad (18)$$

and the general solution for \hat{H}_0 is given by

$$\Psi^{(0)}(\mathbf{r}, t) = \sum_n a_n^{(0)}\psi_n(\mathbf{r})e^{-\frac{i}{\hbar}E_n t}$$

We notice that in this case the coefficients $a_n^{(0)}$ are *independent of time* because the time dependence of the system is completely governed by the exponential factor.

We are interested in situations where initially $t = 0$ only one of the eigenstates of \hat{H}_0 , say ψ_k , is populated

$$\Psi(\mathbf{r}, 0) = \psi_k$$

As time goes on, however, states other than ψ_k are populated because with $\hat{H}'(t) \neq 0$ we are no longer dealing with “stationary” problems; the time evolution is (no longer as simple as $e^{-\frac{i}{\hbar}E_k t}$) governed by

$$i\hbar\frac{\partial}{\partial t}\Psi(\mathbf{r}, t) = (\hat{H}_0 + \hat{H}')\Psi(\mathbf{r}, t). \quad (19)$$

Quite generally the time-dependent perturbation $\hat{H}'(t)$ can cause transitions to states other than ψ_k . The essential question here is: What is the probability as a function of time for the system to be found in ψ_m with $m \neq k$?

The general wavefunction can still be expressed as a linear combination of eigenstates ψ_n (because they constitute a complete set). The only difference is that the coefficients are now *time-dependent*, i.e.

$$\Psi(\mathbf{r}, t) = \sum_n a_n(t)\psi_n(\mathbf{r})e^{-\frac{i}{\hbar}E_n t}$$

Substitute this into the Schrödinger equation (19)

$$i\hbar\frac{\partial}{\partial t}\sum_n a_n(t)\psi_n(\mathbf{r})e^{-\frac{i}{\hbar}E_n t} = (\hat{H}_0 + \hat{H}')\sum_n a_n(t)\psi_n(\mathbf{r})e^{-\frac{i}{\hbar}E_n t}$$

performing the differentiation, and using the eigenvalue equation (18), we obtain

$$\begin{aligned} & i\hbar\sum_n \dot{a}_n(t)\psi_n(\mathbf{r})e^{-\frac{i}{\hbar}E_n t} + \sum_n E_n a_n(t)\psi_n(\mathbf{r})e^{-\frac{i}{\hbar}E_n t} \\ &= \hat{H}_0\sum_n a_n(t)\psi_n(\mathbf{r})e^{-\frac{i}{\hbar}E_n t} + \sum_n a_n(t)\hat{H}'\psi_n(\mathbf{r})e^{-\frac{i}{\hbar}E_n t} \end{aligned}$$

where the second and third terms cancel. Multiplying $\psi_m^*(\mathbf{r})e^{-\frac{i}{\hbar}E_m t}$ and integrating over the entire space, the orthonormality of the basis vector ψ_n leads to an equation for the coefficients

$$i\hbar\dot{a}_m(t) = \sum_n a_n(t)H'_{mn}e^{i\omega_{mn}t} \quad (20)$$

where

$$\begin{aligned} H'_{mn} &= \int \psi_m^*(\mathbf{r})\hat{H}'\psi_n(\mathbf{r})d^3\mathbf{r} = \langle \psi_m | \hat{H}' | \psi_n \rangle \\ \omega_{mn} &= (E_m - E_n) / \hbar \end{aligned}$$

Equation (20) is the matrix form of the time dependent Schrödinger equation in the \hat{H}_0 representation. Its solutions rely on the initial condition at $t = 0$ and the perturbation Hamiltonian.

For a very weak perturbation $\hat{H}' \ll \hat{H}_0$, we can solve equation (20) by a process of **successive approximations**, as follows.

- Zeroth Order: If there were no perturbation \hat{H}' at all, then

$$i\hbar\dot{a}_m^{(0)}(t) = 0 \Rightarrow a_m^{(0)}(t) = \text{const.}$$

(We use a superscript in parentheses to indicate the order of the approximation) Suppose initially the system is prepared in one of the eigenstates of \hat{H}_0 , say ψ_k

$$a_m(0) = \delta_{mk}$$

then it would stay this way forever

$$a_m^{(0)}(t) = \delta_{mk} \quad (21)$$

- First Order: Insert the zeroth-order values on the right side of (20)

$$i\hbar\dot{a}_m^{(1)}(t) = \sum_n a_n^{(0)}(t)H'_{mn}e^{i\omega_{mn}t} = \sum_n \delta_{nk}H'_{mn}e^{i\omega_{mn}t} = H'_{mk}e^{i\omega_{mk}t}$$

we get the first order approximation

$$a_m^{(1)}(t) = \frac{1}{i\hbar} \int_0^t H'_{mk}(t_1)e^{i\omega_{mk}t_1} dt_1 \quad (22)$$

- Second Order: Now we insert this expression on the right

$$i\hbar\dot{a}_m^{(2)}(t) = \sum_n a_n^{(1)}(t)H'_{mn}e^{i\omega_{mn}t} = \frac{1}{i\hbar} \sum_n \int_0^t H'_{nk}(t_1)e^{i\omega_{nk}t_1} dt_1 H'_{mn}e^{i\omega_{mn}t}$$

to obtain the second-order approximation (change t_1 to t_2 and t to t_1 , the second integration is for t_1)

$$a_m^{(2)}(t) = \left(\frac{1}{i\hbar}\right)^2 \sum_n \int_0^t dt_1 \int_0^{t_1} dt_2 H'_{mn}(t_1)e^{i\omega_{mn}t_1} H'_{nk}(t_2)e^{i\omega_{nk}t_2} \quad (23)$$

The physical process corresponding to the second order application involves an intermediate state n , i.e. the transition starts from initial k to n , and then from n to the final state m , which is known as *indirect transition* or *two-photon process*. The probability is much smaller than that for the direct transition.

In principle, we could continue this process indefinitely, always inserting the n th-order approximation into the right side of Equation (20) and solving for the $(n+1)$ th order. This is known as **Dyson series** after Freeman J. Dyson, who applied this method to covariant electrodynamics (QED). For small perturbation we need only consider the first order approximation. According to the statistical interpretation of the coefficients $a_m(t)$, during the action time of perturbation t , the **transition probability** from the initial state ψ_k to the final state ψ_m is

$$P_{mk}(t) = \left| a_m^{(1)}(t) \right|^2 = \frac{1}{\hbar^2} \left| \int_0^t H'_{mk}(t_1)e^{i\omega_{mk}t_1} dt_1 \right|^2 \quad (24)$$

Example 8 Suppose the 1D harmonic oscillator is perturbed by a time-dependent potential

$$V(x, t) = \frac{P_0}{\sqrt{\pi}} e^{-(t/\tau)^2} x$$

where τ is the characteristic interaction time. Initially at $t \rightarrow -\infty$ the system is in the ground state. Let us check the probability of the system can be found in n -th excited states

$$\begin{aligned} P_{n0} &= \left| a_n^{(1)}(t = +\infty) \right|^2 = \frac{1}{\hbar^2} \left| \int_{-\infty}^{+\infty} \frac{P_0}{\sqrt{\pi}} \langle n|x|0 \rangle e^{-(t/\tau)^2} e^{in\omega t} dt \right|^2 \\ &= \frac{1}{\hbar^2} \left| \frac{P_0}{\sqrt{\pi}} \langle n|x|0 \rangle \int_{-\infty}^{+\infty} \exp\left(-\left(\frac{t}{\tau} - \frac{i n \omega \tau}{2}\right)^2\right) \exp(-n^2 \omega^2 \tau^2 / 4) dt \right|^2 \\ &= \frac{1}{\hbar^2} \left| \frac{P_0}{\sqrt{\pi}} \langle n|x|0 \rangle \tau \sqrt{\pi} \exp(-n^2 \omega^2 \tau^2 / 4) \right|^2 \\ &= \frac{P_0^2}{\hbar^2} |\langle n|x|0 \rangle|^2 \tau^2 \exp(-n^2 \omega^2 \tau^2 / 2) \end{aligned}$$

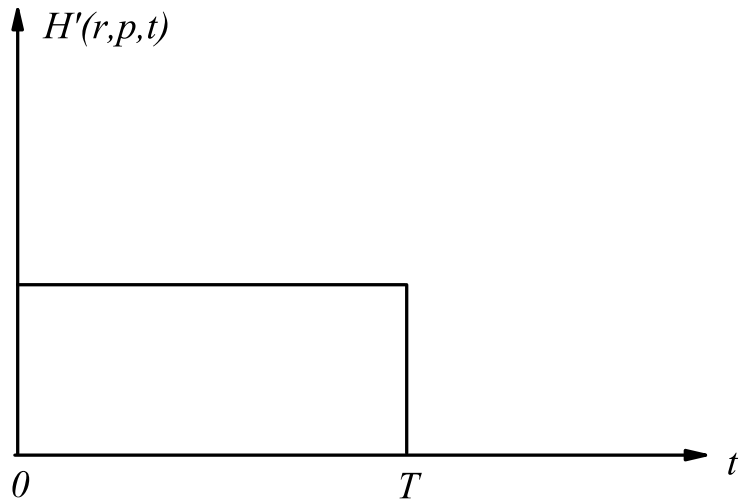


FIG. 11: Constant perturbation.

Evidently for very large τ

$$P_{n0} \rightarrow 0$$

The system remains in the initially populated state at all times whenever the perturbation is added very slowly. This is the so-called **adiabatic approximation**. We will come back to this in next section. On the other hand, for a very short perturbation duration τ , $P_{n0} \rightarrow 0$, the system again won't be changed, which is known as **sudden perturbation**.

C. Constant perturbation

Further analysis is possible only if we choose a specific form for the time dependence of the perturbation. As an application of equation (24), let us consider a constant perturbation turned on at time $t = 0$ for a duration T

$$\hat{H}'(\mathbf{r}, \mathbf{p}, t) = \begin{cases} H'(\mathbf{r}, \mathbf{p}), & 0 < t < T \\ 0, & t > T \end{cases}$$

(See Figure 11.) To first order we have

$$a_m^{(1)}(T) = \frac{1}{i\hbar} H'_{mk} \int_0^T e^{i\omega_{mk}t} dt = \frac{1}{i\hbar} H'_{mk} \frac{e^{i\omega_{mk}T} - 1}{i\omega_{mk}}$$

so that the quantum transition probability from state k to state m is given by

$$\begin{aligned} P_{mk}(T) &= \left| a_m^{(1)}(T) \right|^2 = \frac{|H'_{mk}|^2}{\hbar^2} \left| \frac{e^{i\omega_{mk}T/2} (e^{i\omega_{mk}T/2} - e^{-i\omega_{mk}T/2})}{i\omega_{mk}} \right|^2 \\ &= \frac{|H'_{mk}|^2}{\hbar^2} \frac{\sin^2(\omega_{mk}T/2)}{(\omega_{mk}/2)^2} \end{aligned}$$

This is shown in Figure 12. We notice that there exists a peak around $\omega_{mk} = 0$, i.e. $E_m = E_k$ with the peak value proportional to

$$\lim_{\omega_{mk} \rightarrow 0} \frac{\sin^2(\omega_{mk}T/2)}{(\omega_{mk}/2)^2} = T^2$$

In a certain range of E_m the transition probability is nonzero, which means that the system may jump to many states in the neighborhood of E_k . Furthermore the curve shrinks with increasing duration T , and the area under

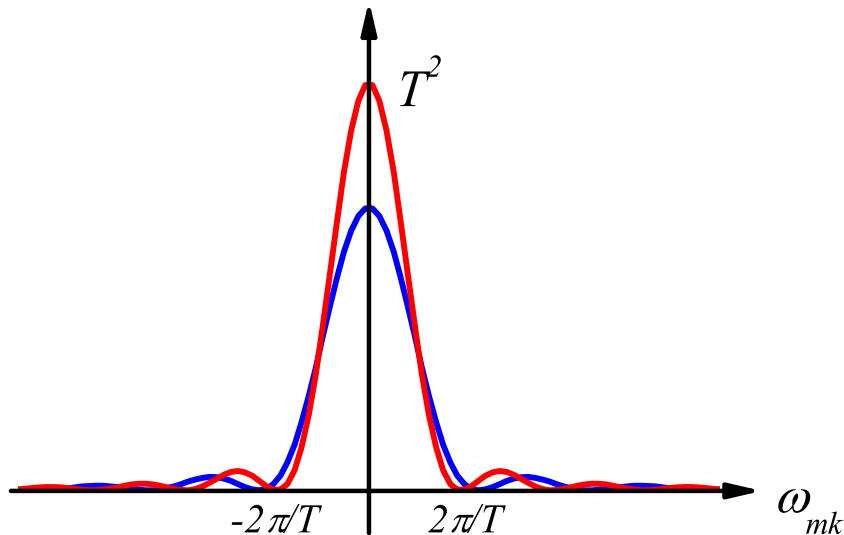


FIG. 12: Transition probability under constant perturbation for different duration T .

the curve increases with T . If there were only one final state, we would get an unreasonable result: the probability that the system remains in the initial state k is quadratic (T^2), not linear, in the time interval, while transition to other states changes irregularly. *The only explanation is that the final state forms a continuous energy spectrum in the neighborhood of E_k .* We should sum the probability over all possible final states, which is nothing but the area under the curve proportional to perturbation duration T . In this sense, the transition probability per unit time is a constant.

The total probability is thus summation over all final states

$$P = \sum_m \frac{|H'_{mk}|^2}{\hbar^2} \frac{\sin^2(\omega_{mk}T/2)}{(\omega_{mk}/2)^2}$$

It is customary to define the density of the final states as the number of states within energy interval $(E, E + dE)$ as $\rho(E)dE$, which changes the summation into an integral for continuous spectrum

$$P = \int \frac{|H'_{mk}|^2}{\hbar^2} \frac{\sin^2(\omega_{mk}T/2)}{(\omega_{mk}/2)^2} \rho(E_m) dE_m$$

As $T \rightarrow \infty$, we may take advantage of

$$\lim_{\alpha \rightarrow \infty} \frac{1}{\pi} \frac{\sin^2 \alpha x}{\alpha x^2} = \delta(x)$$

It follows that

$$\lim_{T \rightarrow \infty} \frac{\sin^2(\omega_{mk}T/2)}{(\omega_{mk}/2)^2} = \pi T \delta(\omega_{mk}/2) = 2\hbar\pi T \delta(E_m - E_k) \quad (25)$$

It is convenient to consider the **transition rate**, that is, the transition probability per unit time

$$w_{mk} = P_{mk}/T = \frac{|H'_{mk}|^2}{\hbar^2} \frac{\sin^2(\omega_{mk}T/2)}{T(\omega_{mk}/2)^2} = \frac{2\pi}{\hbar} |H'_{mk}|^2 \delta(E_m - E_k)$$

the delta function here essentially means the energy conservation in quantum transition. The total transition rate to all final states is given by the integration

$$w = P/T = \int \frac{|H'_{mk}|^2}{\hbar^2} 2\hbar\pi \delta(E_m - E_k) \rho(E_m) dE_m$$

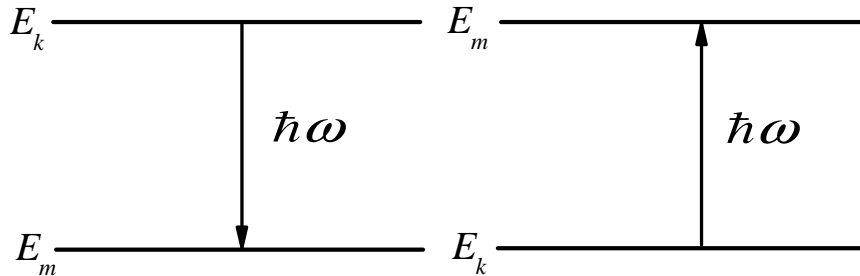


FIG. 13: Stimulated emission: Quantum system gives up $\hbar\omega$ to V - possible only if initial state is excited; Absorption: Quantum system receives $\hbar\omega$ from V and ends up as an excited state.

For small energy interval of the final states, suppose H'_{mk} and $\rho(E_m)$ change smoothly with E_m , one has

$$w = \frac{2\pi}{\hbar} |H'_{mk}|^2 \rho(E_m) \quad (26)$$

This formula is of great practical importance; it is called the **Fermi's golden rule** even though the basic formalism of time dependent theory is due to Dirac.

D. Sinusoidal perturbation

We shall now specialize to a sinusoidal time dependence of the perturbation, since it is often encountered in practice

$$\hat{H}'(\mathbf{r}, \mathbf{p}, t) = \hat{A}(\mathbf{r}, \mathbf{p}) \cos(\omega t) = \hat{F}(\mathbf{r}, \mathbf{p}) (e^{i\omega t} + e^{-i\omega t})$$

so that

$$H'_{mk} = F_{mk} (e^{i\omega t} + e^{-i\omega t})$$

where

$$F_{mk} = \langle \psi_m | \hat{F} | \psi_k \rangle$$

Substitute this into the first-order formula (22), at time T we have

$$\begin{aligned} a_m^{(1)}(T) &= \frac{1}{i\hbar} F_{mk} \int_0^T e^{i\omega_{mk}t} (e^{i\omega t} + e^{-i\omega t}) dt \\ &= \frac{1}{i\hbar} F_{mk} \left(\frac{e^{i(\omega_{mk}+\omega)T} - 1}{i(\omega_{mk} + \omega)} + \frac{e^{i(\omega_{mk}-\omega)T} - 1}{i(\omega_{mk} - \omega)} \right) \\ &= \frac{F_{mk}}{\hbar} \left(\frac{1 - e^{i(\omega_{mk}+\omega)T}}{\omega_{mk} + \omega} + \frac{1 - e^{i(\omega_{mk}-\omega)T}}{\omega_{mk} - \omega} \right) \end{aligned}$$

Let us now consider the behavior of this amplitude in the limit $|\omega T| \gg 1$. Provided the denominators do not vanish, this amplitude remains bounded as T increases. But if $\omega_{mk} + \omega \rightarrow 0$ the first term will grow in proportion to T , and if $\omega_{mk} - \omega \rightarrow 0$ the second term will grow in proportion to T . These are both conditions for *resonance*. Because $\omega > 0$, the first of them, $\hbar\omega = E_k - E_m$, is the condition for *resonant emission* of energy; and the second, $\hbar\omega = E_m - E_k$, is the condition for *resonant absorption* of energy by the system. Near a resonance it is permissible to retain only the dominant term. The validity of perturbation theory at resonance is assured only if F_{mk} are small.

We consider the case $E_m - E_k > 0$, and retain only the resonant absorption term. Then the absorption probability is given by

$$\left| a_m^{(1)}(T) \right|^2 = \frac{|F_{mk}|^2}{\hbar^2} \frac{|1 - e^{i(\omega_{mk}-\omega)T}|^2}{(\omega_{mk} - \omega)^2} = \frac{|F_{mk}|^2}{\hbar^2} \frac{\sin^2\left(\frac{(\omega_{mk}-\omega)T}{2}\right)}{\left(\frac{\omega_{mk}-\omega}{2}\right)^2}, E_m \simeq E_k + \hbar\omega$$

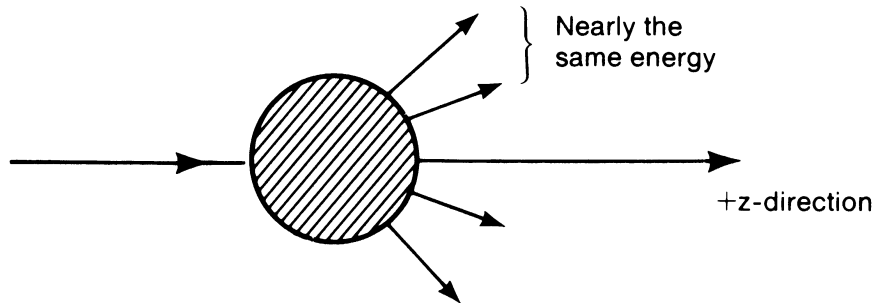


FIG. 14: Elastic scattering of plane wave by some finite range potential.

By the same token we have the emission probability

$$\left| a_m^{(1)}(T) \right|^2 = \frac{|F_{mk}|^2}{\hbar^2} \frac{|1 - e^{i(\omega_{mk} + \omega)T}|^2}{(\omega_{mk} + \omega)^2} = \frac{|F_{mk}|^2}{\hbar^2} \frac{\sin^2\left(\frac{(\omega_{mk} + \omega)T}{2}\right)}{\left(\frac{\omega_{mk} + \omega}{2}\right)^2}, E_m \simeq E_k - \hbar\omega$$

Formally passing to the limit $T \rightarrow \infty$, the transition rates are given by means of equation (25)

$$w_{mk} = \lim_{T \rightarrow \infty} \frac{\left| a_m^{(1)}(T) \right|^2}{T} = \frac{2\pi}{\hbar} |F_{mk}|^2 \delta(E_m - E_k \pm \hbar\omega)$$

Similarly we need a continuous factor to get a reasonable interpretation. This time the initial state k is again fixed so the continuous factor may come from that the final state E_m is in continuum, like the situation of photo-electric effect. Or, suppose both E_k and E_m are fixed, the perturbation itself is continuous, which means $\hbar\omega$ covers a range - it is not monochromatic. The Fermi's golden rule in harmonic perturbation reads

$$w = \frac{2\pi}{\hbar} |F_{mk}|^2 \rho(E_m \pm \hbar\omega). \quad (27)$$

To summarize, for constant perturbation, we obtain appreciable transition probability for $k \rightarrow m$ only if $E_m \simeq E_k$. In contrast, for harmonic perturbation we have appreciable transition probability only if $E_m \simeq E_k - \hbar\omega$ (stimulated emission) or $E_m \simeq E_k + \hbar\omega$ (absorption) (see Figure 13)

E. Elastic scattering

We take the example of elastic scattering as an application of the Fermi's golden rule. In 1D the particles are scattered by a potential barrier - there are only two possibilities, reflection or penetration. We can define the corresponding coefficients $R = j_r/j_i$, $T = j_t/j_i$ with $T + R = 1$. In 3D, however, the incident particles along z -axis may be scattered to different directions and the probability exhibits an angular distribution, see Figure 14. For elastic scattering, the momentum is conserved, i.e. $|\mathbf{p}| = |\mathbf{p}_0|$. Take the incidental particle as a box-normalized plane wave

$$\psi_i = \frac{1}{\sqrt{L^3}} e^{i\mathbf{p}_0 \cdot \mathbf{r}/\hbar},$$

and regard the target particle as perturbation $V(\mathbf{r})$, the scattered wave will be in a state

$$\psi_s = \frac{1}{\sqrt{L^3}} e^{i\mathbf{p} \cdot \mathbf{r}/\hbar}.$$

The matrix elements are calculated as

$$\begin{aligned} H'_{mk} &= \int \frac{1}{\sqrt{L^3}} e^{-i\mathbf{p} \cdot \mathbf{r}/\hbar} V(\mathbf{r}) \frac{1}{\sqrt{L^3}} e^{i\mathbf{p}_0 \cdot \mathbf{r}/\hbar} d^3\mathbf{r} \\ &= \frac{1}{L^3} \int V(\mathbf{r}) e^{i(\mathbf{p}_0 - \mathbf{p}) \cdot \mathbf{r}/\hbar} d^3\mathbf{r} \end{aligned}$$

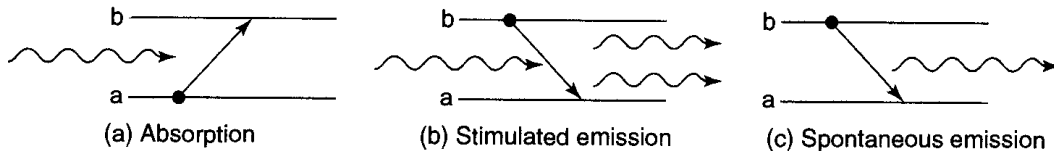


FIG. 15: Three ways in which light interacts with atoms.

Suppose the density of states for the outgoing particle in a small phase volume is ρ , we know

$$\rho dE = \Delta x \Delta y \Delta z \Delta p_x \Delta p_y \Delta p_z / h^3 = L^3 p^2 dp d\Omega / h^3 = L^3 m p d\Omega dE / h^3$$

where we have used for free particle

$$E = p^2/2m, \quad dE = p dp/m$$

The Fermi's golden rule gives the transition rate

$$\begin{aligned} w &= \frac{2\pi}{\hbar} |H'_{mk}|^2 \rho(E_m) \\ &= \frac{2\pi}{\hbar} \frac{1}{L^6} \left| \int V(\mathbf{r}) e^{i(\mathbf{p}_0 - \mathbf{p}) \cdot \mathbf{r} / \hbar} d^3 \mathbf{r} \right|^2 L^3 m p d\Omega / h^3 \end{aligned}$$

The angular distribution of scattered particles in a particular process is described in terms of a *differential cross section*. Suppose that a flux of j_i particles per unit area per unit time is incident on the target. The number of particles per unit time scattered into a narrow cone of solid angle $d\Omega$ centered about the direction whose polar angles with respect to the incident flux are θ and φ , will be proportional to the incident flux j_i and to the angular opening $d\Omega$ of the cone. Hence it may be written as $w = j_i \sigma(\theta, \varphi) d\Omega$. The proportionality factor $\sigma(\theta, \varphi)$ is known as the differential cross section. In our case the incident flux is

$$\begin{aligned} j_i &= -\frac{i\hbar}{2m} (\psi_i^* \nabla \psi_i - \psi_i \nabla \psi_i^*) \\ &= \frac{1}{L^3} \frac{p}{m} \end{aligned}$$

which gives

$$\sigma(\theta, \varphi) = \frac{w}{j_i d\Omega} = \frac{m^2}{(2\pi\hbar^2)^2} \left| \int V(\mathbf{r}) e^{i(\mathbf{p}_0 - \mathbf{p}) \cdot \mathbf{r} / \hbar} d^3 \mathbf{r} \right|^2.$$

Problem 9 Zeng Jinyan, page 341, 2,3 (Solution, page 404, 10.2,10.4 (constant perturbation))

V. SEMICLASSICAL TREATMENT OF EMISSION AND ABSORPTION OF RADIATION

Our basic knowledge about the structure of atoms comes directly from the study of interaction between light (which is essentially one kind of electromagnetic wave) and atoms. An atom in one of its lower states can absorb a passing photon of the right energy and be raised to a higher energy state - a process called **absorption**. Or, the same atom in one of its higher levels can absorb a passing photon and get to a lower level by emitting two photons - a process called **stimulated radiation**. Furthermore, even in the absence of a perturbing external electromagnetic field, the atom in one of its higher levels can still get to a lower level by emitting a photon - a process called **spontaneous radiation**. These three processes are shown in figure 15. Two important observables are the *frequency* and the *relative strength* of the spectral line, which are related to the energy difference and the transition rate. For a more rigorous treatment we need quantum electrodynamics to quantize both the electric field and atom. Here we adopt a semiclassical method to discuss the absorption and stimulated emission, that is, the electric field is treated as a classical field, while the atom is a quantum mechanical system. This method fails for the spontaneous radiation. Interestingly it may be explained by a concept called detailed balance according to Einstein.

A. Absorption and stimulated emission of light

An electromagnetic wave (I'll refer to it as "light", though it could be infrared, ultraviolet, microwave, X-ray, etc.; these differ only in their frequencies) consists of transverse (and mutually perpendicular) oscillating electric and magnetic fields

$$\begin{aligned}\mathbf{E} &= \mathbf{E}_0 \cos(\omega t - \mathbf{k} \cdot \mathbf{r}) \\ \mathbf{B} &= \mathbf{B}_0 \cos(\omega t - \mathbf{k} \cdot \mathbf{r}).\end{aligned}$$

For a Gaussian system, $|\mathbf{E}_0| = |\mathbf{B}_0|$. The electric part of the interaction scales as $e\mathbf{E}_0$, while the magnetic part as $e(\mathbf{v} \times \mathbf{B}_0)/c$ with $v/c \ll 1$. An atom, in the presence of a passing light wave, responds primarily to the electric component. The atom, then, is exposed to a sinusoidally oscillating electric field with the perturbation Hamiltonian

$$\hat{H}' = -\mathbf{D} \cdot \mathbf{E} = -\mathbf{D} \cdot \mathbf{E}_0 \cos(\omega t - \mathbf{k} \cdot \mathbf{r}),$$

with the electric dipole moment $\mathbf{D} = -e\mathbf{r}$. The transition induced by this dipole moment is called **electric dipole moment approximation**. If the wavelength is long compared to the size of the atom (For visible light, the wavelength $\lambda \sim 4000 - 7000 \text{ \AA}$, which is much larger than the Bohr radius a_0), we can ignore the spatial variation in the field, i.e. $\mathbf{k} \cdot \mathbf{r} \sim a/\lambda \ll 1$. The electric field is applied homogeneously on the atom, so

$$\hat{H}' = -\mathbf{D} \cdot \mathbf{E}_0 \cos(\omega t) = \frac{W}{2} (e^{i\omega t} + e^{-i\omega t})$$

with $W = -\mathbf{D} \cdot \mathbf{E}_0$. The first-order transition rate for sinusoidal time dependence of the perturbation gives

$$\begin{aligned}w_{mk} &= \frac{2\pi}{\hbar} \left| \frac{W_{mk}}{2} \right|^2 \delta(E_m - E_k \pm \hbar\omega) \\ &= \frac{\pi}{2\hbar^2} |W_{mk}|^2 \delta(\omega_{mk} \pm \omega) \\ &= \frac{\pi}{2\hbar^2} |\mathbf{D}_{mk}|^2 E_0^2 \cos^2 \theta \delta(\omega_{mk} \pm \omega)\end{aligned}$$

where θ is the angle between \mathbf{D} and \mathbf{E}_0 . We shall be interested in the case of an atom bathed in radiation coming from all directions, and with all possible polarizations. What we need is the average of $\cos^2 \theta$ in all directions

$$\begin{aligned}\overline{\cos^2 \theta} &= \frac{1}{4\pi} \iint \cos^2 \theta d\Omega = \frac{1}{4\pi} \iint \cos^2 \theta \sin \theta d\theta d\varphi \\ &= \frac{1}{2} \int_0^\pi \cos^2 \theta \sin \theta d\theta = \frac{1}{3}\end{aligned}$$

so

$$w_{mk} = \frac{\pi}{6\hbar^2} |\mathbf{D}_{mk}|^2 E_0^2 \delta(\omega_{mk} \pm \omega)$$

But this is for a *monochromatic* perturbation, consisting of a single frequency ω . In many applications the system is exposed to electromagnetic waves at a whole range of frequencies; in that case we need integrate over a frequency range $d\omega$ with density of states

$$\begin{aligned}\rho(\omega) &= \frac{1}{8\pi} \overline{(\mathbf{E}^2 + \mathbf{B}^2)} = \frac{1}{4\pi} \overline{\mathbf{E}^2} \\ &= \frac{1}{4\pi} \overline{\mathbf{E}_0^2 \cos^2 \omega t} = \frac{E_0^2}{4\pi} \frac{1}{T} \int_0^T dt \cos^2 \omega t = \frac{1}{8\pi} E_0^2\end{aligned}$$

where the average is taken in a period $T = 2\pi/\omega$. The energy density E_0^2 should then be replaced by $\int d\omega 8\pi \rho(\omega)$. We have the total transition rate induced by incoherent, non-polarized natural light (to be more specific, for absorption)

$$\begin{aligned}w_{mk} &= \frac{4\pi^2}{3\hbar^2} |\mathbf{D}_{mk}|^2 \rho(\omega_{mk}) \\ &= \frac{4\pi^2 e^2}{3\hbar^2} |\mathbf{r}_{mk}|^2 \rho(\omega_{mk})\end{aligned}$$

Evidently the transition rate is proportional to the energy intensity in the field $\rho(\omega_{mk})$ and the matrix element $|\mathbf{r}_{mk}|$. This leads to the **selection rules** in quantum mechanics.

Suppose we are interested in systems like hydrogen, for which the Hamiltonian is spherically symmetrical. In that case we may specify the states with the usual quantum numbers n, l , and m , and the matrix elements are

$$\langle n'l'm' | \mathbf{r} | nlm \rangle$$

Clever exploitation of the angular momentum commutation relations and the hermiticity of the angular momentum operators yields a set of powerful constraints on this quantity.

- Selection rules involving m and m' : Consider first the commutators of \hat{L}_z with \hat{x} , \hat{y} , and \hat{z} , which we worked out earlier

$$\begin{aligned} [\hat{L}_z, \hat{x}] &= i\hbar\hat{y}, \\ [\hat{L}_z, \hat{y}] &= -i\hbar\hat{x}, \\ [\hat{L}_z, \hat{z}] &= 0. \end{aligned}$$

From the third of these it follows that

$$0 = \langle n'l'm' | [\hat{L}_z, \hat{z}] | nlm \rangle = \langle n'l'm' | (\hat{L}_z\hat{z} - \hat{z}\hat{L}_z) | nlm \rangle = (m' - m)\hbar \langle n'l'm' | \hat{z} | nlm \rangle$$

So unless $m' = m$, the matrix elements of \hat{z} are always zero. Meanwhile, from the commutator of \hat{L}_z with \hat{x} and \hat{y} we get

$$\begin{aligned} \langle n'l'm' | [\hat{L}_z, \hat{x}] | nlm \rangle &= \langle n'l'm' | (\hat{L}_z\hat{x} - \hat{x}\hat{L}_z) | nlm \rangle \\ &= (m' - m)\hbar \langle n'l'm' | \hat{x} | nlm \rangle = i\hbar \langle n'l'm' | \hat{y} | nlm \rangle \\ \langle n'l'm' | [\hat{L}_z, \hat{y}] | nlm \rangle &= \langle n'l'm' | (\hat{L}_z\hat{y} - \hat{y}\hat{L}_z) | nlm \rangle \\ &= (m' - m)\hbar \langle n'l'm' | \hat{y} | nlm \rangle = -i\hbar \langle n'l'm' | \hat{x} | nlm \rangle \end{aligned}$$

so

$$(m' - m)^2 \langle n'l'm' | \hat{x} | nlm \rangle = \langle n'l'm' | \hat{x} | nlm \rangle$$

and hence either $(m' - m)^2 = 1$ or the matrix elements of \hat{x} are zero. So we obtain the selection rule for m

$$\text{No transition occurs unless } \Delta m = \pm 1, 0$$

This is an easy result to understand if you remember that the photon carries spin 1, and hence its value of m is 1, 0, -1; conservation of (the z -component of) angular momentum requires that the atom give up whatever the photon takes away.

- Selection rules involving l and l' : From the following commutation relation (Check it!)

$$[\hat{L}^2, [\hat{L}^2, \mathbf{r}]] = 2\hbar^2(\mathbf{r}\hat{L}^2 + \hat{L}^2\mathbf{r})$$

we can get the selection rule for l

$$\text{No transition occurs unless } \Delta l = \pm 1$$

Again, this result (though scarcely trivial to derive) is easy to interpret: The photon carries spin 1, so the rules for addition of angular momentum would allow $l' = l + 1, l, l - 1$. For electric dipole radiation the middle possibility – though permitted by conservation of angular momentum – does not occur.

Evidently not all transitions to lower-energy states can proceed; some are forbidden by the selection rules. The scheme of allowed transitions for the first four Bohr levels in hydrogen is shown in Figure 16. Note that the $2S$ state ψ_{200} is "stuck": It cannot decay, because there is no lower-energy state with $l = 1$. It is called a metastable state, and its lifetime is indeed much longer than that of, for example, the $2P$ states $\psi_{211}, \psi_{210}, \psi_{21-1}$. Metastable states do eventually decay, either by collisions or by **magnetic dipole** or **electric quadrupole** transitions, or by **multi-photon emission**.

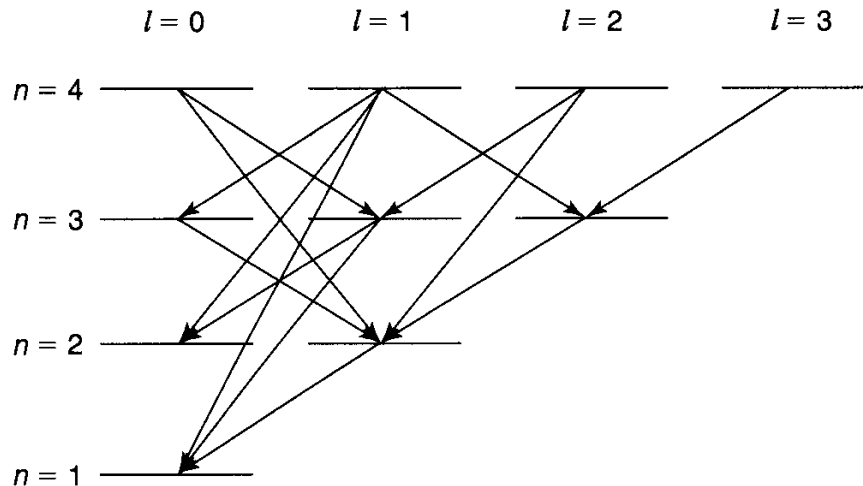


FIG. 16: Allowed decays for the first four Bohr levels in hydrogen.

B. Einstein's treatment of spontaneous emission

It is well known that an atom in an excited state will spontaneously emit radiation and return to its ground state. That phenomenon is not predicted by this version of the theory, in which only the atom is treated as a quantum-mechanical system, but the radiation is treated as an external classical field. A more rigorous calculation requires a quantum theory of the electromagnetic field. Nevertheless, Einstein was able to deduce some of the most important features of spontaneous emission in 1917, when most of the quantum theory was unknown. The argument below, derived from Einstein's ideas, is based on the principle that the radiation mechanism must preserve the statistical equilibrium among the excited states of the atoms.

We first define the absorption and stimulated emission coefficients. Let the number of atoms in state k be n_k , and consider transitions between states k and m involving the emission and absorption of radiation at the frequency $\omega_{mk} = (E_m - E_k)/\hbar > 0$. We have calculated the transition rate for an atom to absorb radiation. It has the form

$$w_{mk} = B_{mk}\rho(\omega_{mk})$$

where

$$B_{mk} = \frac{4\pi^2 e^2}{3\hbar^2} |\mathbf{r}_{mk}|^2$$

and $\rho(\omega_{mk})$ is the energy density of the radiation at the angular frequency ω_{mk} . Therefore the rate of excitation of atoms by absorption of radiation will be

$$B_{mk}\rho(\omega_{mk})n_k.$$

Einstein assumed that the rate of de-excitation of atoms by emitting radiation is of the form

$$B_{km}\rho(\omega_{mk})n_m + A_{km}n_m.$$

The first term corresponds to *stimulated emission*, which we have shown how to calculate. The second term, which is independent of the presence of any radiation, describes *spontaneous emission*. At equilibrium the rates of excitation and de-excitation must balance, so we must have

$$B_{km}\rho(\omega_{mk})n_m + A_{km}n_m = B_{mk}\rho(\omega_{mk})n_k$$

Invoking the principle of *detailed balance*, Einstein assumed that the probabilities of induced emission and absorption should be equal: $B_{mk} = B_{km}$. This relation was confirmed by our quantum-mechanical calculation above. Furthermore, we know that in thermodynamic equilibrium the relative occupation of the various atomic states is given by the Boltzmann distribution, so we must have

$$n_k/n_m = e^{(E_m - E_k)/kT} = e^{\hbar\omega_{mk}/kT}$$

with k the Boltzmann constant and T is the temperature. Therefore we may solve for energy density of the radiation field at equilibrium

$$\begin{aligned}\rho(\omega_{mk}) &= \frac{A_{km}}{B_{km}} \frac{1}{n_k/n_m - 1} \\ &= \frac{A_{km}}{B_{km}} \frac{1}{e^{\hbar\omega_{mk}/kT} - 1}\end{aligned}$$

But Planck's black body formula tells us the energy density of thermal radiation

$$\rho(\omega_{mk}) = \frac{\hbar}{\pi^2 c^3} \frac{\omega_{mk}^3}{e^{\hbar\omega_{mk}/kT} - 1}$$

Comparing the two expressions, we conclude that the *spontaneous emission rate* is

$$A_{km} = \frac{\hbar\omega_{mk}^3}{\pi^2 c^3} B_{km}$$

which relates the spontaneous emission probability to the induced emission probability that has already been calculated. This relation, derived before most of quantum mechanics had been formulated, remains valid in modern quantum electrodynamics.

VI. SCHRÖDINGER PICTURE, HEISENBERG PICTURE AND INTERACTION PICTURE

In the historical development there appeared two different views on the time evolution of the system:

- Schrödinger Picture: It is considered that it is the state of the system that is evolved, i.e., the state evolves from $|\Psi(\mathbf{r},0)\rangle$ into $|\Psi(\mathbf{r},t)\rangle$ as time goes from 0 to t . As consequences of the state evolution, all results we can draw from the system evolve along with it.
- Heisenberg Picture: It is considered that the state of the system does not change with time and will stay as it was. It is the dynamical variables (operators) that evolve with time.

These two points of view are equal in explaining the evolution of the system and can be transformed into each other through suitable transformation.

A. Schrödinger picture

In an arbitrary representation, state $|\Psi(\mathbf{r},0)\rangle$ is expressed as a vector. As time flies, the vector evolves, at time t , the vector reaches $|\Psi(\mathbf{r},t)\rangle$. We can define the *time evolution operator* $U(t,0)$ as

$$|\Psi(\mathbf{r},t)\rangle = U(t,0) |\Psi(\mathbf{r},0)\rangle$$

The Schrödinger equation is now

$$i\hbar \frac{d}{dt} (U(t,0) |\Psi(\mathbf{r},0)\rangle) = \hat{H} (U(t,0) |\Psi(\mathbf{r},0)\rangle)$$

The initial state is arbitrary, so we have the Schrödinger equation for time evolution operator

$$i\hbar \frac{d}{dt} U(t,0) = \hat{H} U(t,0) \quad (28)$$

Our first task is to derive formal solutions. There are three cases to be treated separately.

Case I The Hamiltonian operator \hat{H} is independent of time t . The Hamiltonian for a spin magnetic moment interacting with a time-independent magnetic field is an example of this: $\hat{H} \sim \mathbf{S} \cdot \mathbf{B}$. The solution in such a case is given by

$$U(t,0) = e^{-\frac{i}{\hbar} \hat{H} t}. \quad (29)$$

For simplicity we take as an example the harmonic oscillator initially in its ground state

$$|\Psi(x,0)\rangle = \psi_0(x) = \left(\frac{m\omega}{\pi\hbar}\right)^{1/4} e^{-\frac{1}{2}\frac{m\omega}{\hbar}x^2}$$

$$\hat{H} = \frac{p^2}{2m} + \frac{1}{2}m\omega^2x^2$$

Action of time evolution operator upon $|\Psi(x,0)\rangle$ gives

$$U(t,0)|\Psi(x,0)\rangle = e^{-\frac{i}{\hbar}\hat{H}t}\psi_0(x) = e^{-\frac{i}{\hbar}E_0t}\psi_0(x) = \Psi(x,t)$$

which is just the solution of time-dependent Schrödinger equation. Similarly, for a superposition state of the ground state and the first excited state

$$|\Psi(x,0)\rangle = \frac{1}{\sqrt{2}}\psi_0(x) + \frac{1}{\sqrt{2}}\psi_1(x)$$

at time t we get

$$U(t,0)|\Psi(x,0)\rangle = \frac{1}{\sqrt{2}}e^{-\frac{i}{\hbar}E_0t}\psi_0(x) + \frac{1}{\sqrt{2}}e^{-\frac{i}{\hbar}E_1t}\psi_1(x) = \Psi(x,t)$$

Case II The Hamiltonian operator \hat{H} is time dependent but the \hat{H} 's at different times commute. As an example, let us consider the spin magnetic moment subjected to a magnetic field whose strength varies with time but whose direction is always unchanged. The formal solution in this case is

$$U(t,0) = e^{-\frac{i}{\hbar}\int_0^t dt' \hat{H}(t')}. \quad (30)$$

Case III The \hat{H} 's at different times do *not* commute. Continuing with the example involving spin magnetic momentum, we suppose, this time, that the magnetic field direction also changes with time: at $t = t_1$ in the x -direction, at $t = t_2$ in the y -direction, and so forth. Because S_x and S_y do not commute, $H(t_1)$ and $H(t_2)$, which go like $\mathbf{S} \cdot \mathbf{B}$, do not commute either. The formal solution in such a situation is given by

$$U(t,0) = 1 + \sum_{n=1}^{\infty} \left(-\frac{i}{\hbar}\right)^n \int_0^t dt_1 \int_0^{t_1} dt_2 \cdots \int_0^{t_{n-1}} dt_n \hat{H}(t_1)\hat{H}(t_2)\cdots\hat{H}(t_n) \quad (31)$$

which is an alternative form of Dyson series.

What are some properties we would like to ascribe to the *time evolution operator*? The first important property is the unitary requirement for $U(t,0)$ that follows from probability conservation (See Sakurai, Modern Quantum Mechanics, page 70)

$$U^\dagger(t,0) = U(t,0)^{-1} \text{ or } U^\dagger(t,0)U(t,0) = U(t,0)U^\dagger(t,0) = 1$$

This assures that the state must remain normalized to unity at all later times if it is initially normalized to unity

$$\langle\Psi(x,0|\Psi(x,0)\rangle = 1 \Rightarrow \langle\Psi(x,t|\Psi(x,t)\rangle = \langle\Psi(x,0|U^\dagger(t,0)U(t,0)|\Psi(x,0)\rangle = 1.$$

The expectation value of an observable \hat{A} on a non-stationary state may change with time, i.e.

$$\langle A(0)\rangle = \langle\Psi(x,0|\hat{A}|\Psi(x,0)\rangle$$

and

$$\langle A(t)\rangle = \langle\Psi(x,t|\hat{A}|\Psi(x,t)\rangle$$

may differ from each other. Another feature we require of the U operator is the composition property

$$U(t_2,t_0) = U(t_2,t_1)U(t_1,t_0) \quad (32)$$

It says that if we are interested in obtaining time evolution from t_0 to t_2 , then we can obtain the same result by first considering time evolution from t_0 to t_1 , then from t_1 to t_2 , a reasonable requirement. Note that we read equation (32) from right to left.

B. Heisenberg picture

There is another formulation of quantum dynamics where observables, rather than state kets, vary with time, known as Heisenberg picture. It is important to keep in mind that under a unitary transformation that changes the state kets, the inner product of a state bra $|\alpha\rangle$ and a state ket $|\beta\rangle$ remains unchanged

$$\langle\beta|\alpha\rangle \rightarrow \langle\beta|U^\dagger U|\alpha\rangle = \langle\beta|\alpha\rangle$$

We can infer how $\langle\beta|\hat{F}|\alpha\rangle$ must change

$$\langle\beta|\hat{F}|\alpha\rangle \rightarrow \{\langle\beta|U^\dagger\} \hat{F} \{U|\alpha\rangle\} = \langle\beta|(U^\dagger \hat{F} U)|\alpha\rangle$$

This suggests two approaches to unitary transformations

Approach 1 $|\alpha\rangle \rightarrow U|\alpha\rangle$, with operators unchanged

Approach 2 $\hat{F} \rightarrow U^\dagger \hat{F} U$, with state kets unchanged

Motivated by the Approach 2, we define the Heisenberg picture observable by

$$\hat{A}^H(t) = U^\dagger(t, 0) \hat{A}^S U(t, 0) \quad (33)$$

where the superscripts H and S stand for Heisenberg and Schrödinger, respectively. At $t = 0$, the Heisenberg picture observable and the corresponding Schrödinger picture observable coincide

$$\hat{A}^H(0) = \hat{A}^S$$

The state kets also coincide between the two pictures at $t = 0$; at later time t the Heisenberg picture state ket is frozen to what it was at $t = 0$

$$|\Psi(\mathbf{r}, t)\rangle_H = |\Psi(\mathbf{r}, 0)\rangle$$

independent of t . This is in dramatic contrast with the Schrödinger picture state kets

$$|\Psi(\mathbf{r}, t)\rangle_S = U(t, 0) |\Psi(\mathbf{r}, 0)\rangle$$

The experimental measurement must give the same result in these two pictures. In fact, the expectation value $\langle A(t) \rangle$ is obviously the same in both pictures

$$\begin{aligned} {}_S \langle \Psi(\mathbf{r}, t) | \hat{A}^S | \Psi(\mathbf{r}, t) \rangle_S &= \langle \Psi(\mathbf{r}, 0) | U^\dagger \hat{A}^S U | \Psi(\mathbf{r}, 0) \rangle \\ &= {}_H \langle \Psi(\mathbf{r}, t) | \hat{A}^H(t) | \Psi(\mathbf{r}, t) \rangle_H \end{aligned}$$

or

$$\langle A \rangle_S = \langle A \rangle_H$$

We now derive the fundamental equation of motion in the Heisenberg picture. Assuming that \hat{A}_S does not depend explicitly on time, which is the case in most physical situations of interest, we obtain by differentiating equation (33)

$$\begin{aligned} \frac{d\hat{A}^H(t)}{dt} &= \frac{\partial U^\dagger(t, 0)}{\partial t} \hat{A}^S U(t, 0) + U^\dagger(t, 0) \hat{A}^S \frac{\partial U(t, 0)}{\partial t} \\ &= -\frac{1}{i\hbar} U^\dagger(t, 0) \hat{H} \hat{A}^S U(t, 0) + \frac{1}{i\hbar} U^\dagger(t, 0) \hat{A}^S \hat{H} U(t, 0) \\ &= -\frac{1}{i\hbar} (U^\dagger \hat{H} U) (U^\dagger \hat{A}^S U) + \frac{1}{i\hbar} (U^\dagger \hat{A}^S U) (U^\dagger \hat{H} U) \\ &= \frac{1}{i\hbar} [\hat{A}^H, \hat{H}^H] \end{aligned}$$

where we have used equation (28) and its conjugate. The Hamiltonian was originally defined in Schrödinger picture. In elementary applications (Case I), $U(t, 0)$ depends solely on \hat{H} as $e^{-\frac{i}{\hbar}\hat{H}t}$ and obviously commutes with \hat{H} , as a result, the Hamiltonian operator has the same form in both pictures

$$\hat{H}^H = U^\dagger \hat{H} U = \hat{H}^S$$

so it is all right to write the equation of motion for operator as

$$i\hbar \frac{d}{dt} \hat{A}^H(t) = [\hat{A}^H, \hat{H}] \quad (34)$$

Historically, this equation was first written by P. A. M. Dirac, who - with his characteristic modesty - call it **Heisenberg equation of motion**. Note that for Heisenberg picture, there is only Heisenberg equation for operators and no Schrödinger equation for wavefunctions. And if an operator \hat{A}^S commute with \hat{H} in Schrödinger picture, it also does in Heisenberg picture. Indeed

$$\begin{aligned} i\hbar \frac{d}{dt} \hat{A}^H(t) &= \hat{A}^H \hat{H} - \hat{H} \hat{A}^H \\ &= U^\dagger \hat{A}^S U \hat{H} - \hat{H} U^\dagger \hat{A}^S U \\ &= U^\dagger \hat{A}^S \hat{H} U - U^\dagger \hat{H} \hat{A}^S U \\ &= U^\dagger (\hat{A}^S \hat{H} - \hat{H} \hat{A}^S) U \end{aligned}$$

i.e. if \hat{A}^S is a conservative quantity in Schrödinger picture, it also does in Heisenberg picture. In discussing the conservative quantities in Schrödinger picture, we arrived at the result that

$$i\hbar \frac{d}{dt} \langle \hat{A} \rangle = \langle \hat{A} \hat{H} - \hat{H} \hat{A} \rangle$$

if \hat{A} does not explicitly depend on time. This should not be confused with Heisenberg equation here because they are discussed in different pictures.

To conclude let us summarize the differences between the Schrödinger picture and the Heisenberg picture in the following table.

	Schrödinger picture	Heisenberg picture
Wavefunctions	$ \Psi(\mathbf{r}, t)\rangle_S = U(t, 0) \Psi(\mathbf{r}, 0)\rangle$	$ \Psi(\mathbf{r}, t)\rangle_H = \Psi(\mathbf{r}, 0)\rangle$
Dynamical equation	$i\hbar \frac{d}{dt} \Psi(\mathbf{r}, t)\rangle_S = \hat{H} \Psi(\mathbf{r}, t)\rangle_S$	unchanged
Operators	\hat{A}^S	$\hat{A}^H(t) = U^\dagger(t, 0) \hat{A}^S U(t, 0)$
Dynamical equation	unchanged	$i\hbar \frac{d}{dt} \hat{A}^H(t) = [\hat{A}^H, \hat{H}]$
	$\langle A \rangle_S = \langle A \rangle_H$	

C. Interaction picture

In dealing with the problem of quantum transition, it is more convenient to use the interaction picture, which is in between the Schrödinger picture and the Heisenberg picture. The premise for this is that our Hamiltonian \hat{H} can be divided into two terms $\hat{H} = \hat{H}_0 + \hat{H}'$, where \hat{H}_0 is a stationary Hamiltonian and does not depend on time. The perturbation \hat{H}' is time dependent. The effects of \hat{H}_0 and \hat{H}' are different in inducing transitions. Initially the state of the system

$$\Psi(\mathbf{r}, 0) = \sum_n a_n \psi_n$$

where ψ_n 's are the eigenfunctions of \hat{H}_0 . If only \hat{H}_0 exists, that means, no perturbation present, the evolution of the state would be

$$\Psi(\mathbf{r}, t) = \sum_n a_n \psi_n e^{-\frac{i}{\hbar} E_n t}$$

The probability of finding the system in state ψ_n is $\left| a_n e^{-\frac{i}{\hbar} E_n t} \right|^2$, which does not change with time. Thus, \hat{H}_0 has no effect in causing transitions. Only the perturbation \hat{H}' induce transitions. The interaction picture is a modification to Schrödinger picture that makes the effect of \hat{H}' more prominent and is very powerful in treating the transition problems.

Definition of state in interaction picture: In Schrödinger picture the state of the system at any time t is expressed by

$$|\Psi(\mathbf{r}, t)\rangle_S = U(t, 0) |\Psi(\mathbf{r}, 0)\rangle$$

where $U(t, 0)$ comprises both the effect of \hat{H}_0 and \hat{H}' . The interaction picture “deducts” the effect of \hat{H}_0 in the final state by defining the interaction picture state vector $|\Psi(\mathbf{r}, t)\rangle_I$ as follows

$$|\Psi(\mathbf{r}, t)\rangle_I = U_0^{-1}(t, 0) |\Psi(\mathbf{r}, t)\rangle_S = U_0^{-1}(t, 0) U(t, 0) |\Psi(\mathbf{r}, 0)\rangle$$

where

$$U_0^{-1}(t, 0) = e^{\frac{i}{\hbar} \hat{H}_0 t}$$

is the inverse of evolution operator of \hat{H}_0 . So

$$|\Psi(\mathbf{r}, t)\rangle_I = e^{\frac{i}{\hbar} \hat{H}_0 t} |\Psi(\mathbf{r}, t)\rangle_S \quad (35)$$

Definition of operator in interaction picture: We define

$$\begin{aligned} \hat{A}_I &= U_0^{-1}(t, 0) \hat{A}_S U_0(t, 0) \\ &= e^{\frac{i}{\hbar} \hat{H}_0 t} \hat{A}_S e^{-\frac{i}{\hbar} \hat{H}_0 t} \end{aligned} \quad (36)$$

This is a natural result as the expectation values obtained in two pictures must equal each other

$$\begin{aligned} \langle A \rangle &= {}_I \langle \Psi(\mathbf{r}, t) | \hat{A}^I | \Psi(\mathbf{r}, t) \rangle_I \\ &= {}_S \langle \Psi(\mathbf{r}, t) | e^{-\frac{i}{\hbar} \hat{H}_0 t} \left(e^{\frac{i}{\hbar} \hat{H}_0 t} \hat{A}_S e^{-\frac{i}{\hbar} \hat{H}_0 t} \right) e^{\frac{i}{\hbar} \hat{H}_0 t} | \Psi(\mathbf{r}, t) \rangle_S \\ &= {}_S \langle \Psi(\mathbf{r}, t) | \hat{A}_S | \Psi(\mathbf{r}, t) \rangle_S \end{aligned}$$

Dynamical equation in interaction picture: We now derive the fundamental differential equation that characterize the time evolution of a state ket in the interaction picture. Take the time derivative of equation (35)

$$\begin{aligned} i\hbar \frac{\partial}{\partial t} |\Psi(\mathbf{r}, t)\rangle_I &= i\hbar \frac{\partial}{\partial t} \left(e^{\frac{i}{\hbar} \hat{H}_0 t} |\Psi(\mathbf{r}, t)\rangle_S \right) \\ &= -\hat{H}_0 e^{\frac{i}{\hbar} \hat{H}_0 t} |\Psi(\mathbf{r}, t)\rangle_S + e^{\frac{i}{\hbar} \hat{H}_0 t} \left(i\hbar \frac{\partial}{\partial t} |\Psi(\mathbf{r}, t)\rangle_S \right) \\ &= -\hat{H}_0 e^{\frac{i}{\hbar} \hat{H}_0 t} |\Psi(\mathbf{r}, t)\rangle_S + e^{\frac{i}{\hbar} \hat{H}_0 t} \left(\hat{H}_0 + \hat{H}'_S \right) |\Psi(\mathbf{r}, t)\rangle_S \\ &= e^{\frac{i}{\hbar} \hat{H}_0 t} \hat{H}'_S |\Psi(\mathbf{r}, t)\rangle_S \\ &= \left(e^{\frac{i}{\hbar} \hat{H}_0 t} \hat{H}'_S e^{-\frac{i}{\hbar} \hat{H}_0 t} \right) \left(e^{\frac{i}{\hbar} \hat{H}_0 t} |\Psi(\mathbf{r}, t)\rangle_S \right) \end{aligned}$$

we have

$$i\hbar \frac{\partial}{\partial t} |\Psi(\mathbf{r}, t)\rangle_I = \hat{H}'_I |\Psi(\mathbf{r}, t)\rangle_I. \quad (37)$$

This is the Schrödinger-like equation with the total Hamiltonian replaced by \hat{H}'_I

$$\hat{H}'_I = e^{\frac{i}{\hbar} \hat{H}_0 t} \hat{H}'_S e^{-\frac{i}{\hbar} \hat{H}_0 t}$$

which is the perturbation in interaction picture. The change of $|\Psi(\mathbf{r}, t)\rangle_I$ with time is solely decided by \hat{H}'_I . The initial states $|\Psi(\mathbf{r}, 0)\rangle_I$ and $|\Psi(\mathbf{r}, 0)\rangle_S$ are the same and will be denoted by $|\Psi(\mathbf{r}, 0)\rangle$ hereafter. We can also show for an observable \hat{A} that does not contain time explicitly in Schrödinger picture

$$i\hbar \frac{d}{dt} \hat{A}^I(t) = \left[\hat{A}^I, \hat{H}_0 \right] \quad (38)$$

which is the Heisenberg-like equation with the total Hamiltonian replaced by \hat{H}_0 . In many aspects, the interaction picture, or *Dirac picture*, is intermediate between the Schrödinger picture and the Heisenberg picture. This should be evident from the following table.

	Schrödinger picture	Interaction picture	Heisenberg picture
Wavefunctions	Evolution determined by \hat{H}	Evolution determined by \hat{H}'	No change
Operators	No change	Evolution determined by \hat{H}_0	Evolution determined by \hat{H}

Integration of the Schrödinger-like equation by iteration: Consider the evolution of the system from time $t = 0$ to arbitrary time t . Integration the equation directly gives

$$i\hbar (|\Psi(\mathbf{r}, t)\rangle_I - |\Psi(\mathbf{r}, 0)\rangle_I) = \int_0^t \hat{H}'_I(t') |\Psi(\mathbf{r}, t')\rangle_I dt'$$

or

$$|\Psi(\mathbf{r}, t)\rangle_I = |\Psi(\mathbf{r}, 0)\rangle_I + \frac{1}{i\hbar} \int_0^t \hat{H}'_I(t') |\Psi(\mathbf{r}, t')\rangle_I dt'$$

The $|\Psi(\mathbf{r}, t')\rangle_I$ appeared in the integrand is itself an unknown state vector to be searched. Repeat the above procedure by substituting $|\Psi(\mathbf{r}, t')\rangle_I$ with the above formula gives the following result

$$\begin{aligned} |\Psi(\mathbf{r}, t)\rangle_I &= |\Psi(\mathbf{r}, 0)\rangle_I + \frac{1}{i\hbar} \int_0^t \hat{H}'_I(t') \left\{ |\Psi(\mathbf{r}, 0)\rangle_I + \frac{1}{i\hbar} \int_0^{t'} \hat{H}'_I(t'') |\Psi(\mathbf{r}, t'')\rangle_I dt'' \right\} dt' \\ &= |\Psi(\mathbf{r}, 0)\rangle_I + \frac{1}{i\hbar} \int_0^t \hat{H}'_I(t') |\Psi(\mathbf{r}, 0)\rangle_I dt' \\ &\quad + \left(\frac{1}{i\hbar} \right)^2 \int_0^t dt' \int_0^{t'} dt'' \hat{H}'_I(t') \hat{H}'_I(t'') |\Psi(\mathbf{r}, t'')\rangle_I \end{aligned}$$

The above argument can be repeated once and once (iteration) to give the solution $|\Psi(\mathbf{r}, t)\rangle_I$ in series form

$$\begin{aligned} |\Psi(\mathbf{r}, t)\rangle_I &= |\Psi(\mathbf{r}, 0)\rangle_I && \text{0th-order - no perturbation} \\ &+ \frac{1}{i\hbar} \int_0^t dt_1 \hat{H}'_I(t_1) |\Psi(\mathbf{r}, 0)\rangle_I && \text{1st-order - } \hat{H}'_I \text{ acted at } t_1 \\ &+ \left(\frac{1}{i\hbar} \right)^2 \int_0^t dt_1 \int_0^{t_1} dt_2 \hat{H}'_I(t_1) \hat{H}'_I(t_2) |\Psi(\mathbf{r}, 0)\rangle_I && \text{2nd-order - } \hat{H}'_I \text{ acted at } t_2 \text{ and then } t_1 \\ &+ \left(\frac{1}{i\hbar} \right)^3 \int_0^t dt_1 \int_0^{t_1} dt_2 \int_0^{t_2} dt_3 \hat{H}'_I(t_1) \hat{H}'_I(t_2) \hat{H}'_I(t_3) |\Psi(\mathbf{r}, 0)\rangle_I + \dots \end{aligned}$$

Transition amplitude: We reconsider the problem of quantum transition in interaction picture. The system starts out initially in one of the eigenstate $|\psi_k\rangle$ at time $t = 0$, i.e $|\Psi(\mathbf{r}, 0)\rangle_I = |\psi_k\rangle$. We want to know the transition rate to another eigenstate $|\psi_m\rangle$ under the perturbation \hat{H}' . We have the zeroth order result

$$\langle \psi_m | \Psi(\mathbf{r}, 0) \rangle_I = \langle \psi_m | \psi_k \rangle = \delta_{mk}$$

and the first order transition amplitude

$$\begin{aligned} \langle \psi_m | \frac{1}{i\hbar} \int_0^t dt_1 \hat{H}'_I(t_1) |\psi_k\rangle &= \frac{1}{i\hbar} \int_0^t dt_1 \langle \psi_m | \hat{H}'_I(t_1) |\psi_k\rangle \\ &= \frac{1}{i\hbar} \int_0^t dt_1 \langle \psi_m | e^{\frac{i}{\hbar} \hat{H}_0 t_1} \hat{H}'_S(t_1) e^{-\frac{i}{\hbar} \hat{H}_0 t_1} |\psi_k\rangle \\ &= \frac{1}{i\hbar} \int_0^t dt_1 e^{\frac{i}{\hbar} E_m t_1} e^{-\frac{i}{\hbar} E_k t_1} \langle \psi_m | \hat{H}'_S(t_1) |\psi_k\rangle \\ &= \frac{1}{i\hbar} \int_0^t H'_{mk}(t_1) e^{i\omega_{mk} t_1} dt_1 \end{aligned}$$

We arrive at the same result as in time-dependent perturbation theory equation (22). The second-order transition amplitude appears to be

$$\begin{aligned} \langle \psi_m | \left(\frac{1}{i\hbar} \right)^2 \int_0^t dt_1 \int_0^{t_1} dt_2 \hat{H}'_I(t_1) \hat{H}'_I(t_2) | \psi_k \rangle &= \left(\frac{1}{i\hbar} \right)^2 \int_0^t dt_1 \int_0^{t_1} dt_2 \langle \psi_m | \hat{H}'_I(t_1) \hat{H}'_I(t_2) | \psi_k \rangle \\ &= \left(\frac{1}{i\hbar} \right)^2 \int_0^t dt_1 \int_0^{t_1} dt_2 \langle \psi_m | \hat{H}'_I(t_1) \sum_n | \psi_n \rangle \langle \psi_n | \hat{H}'_I(t_2) | \psi_k \rangle \\ &= \left(\frac{1}{i\hbar} \right)^2 \sum_n \int_0^t dt_1 \int_0^{t_1} dt_2 \hat{H}'_{mn}(t_1) e^{i\omega_{mn}t_1} \hat{H}'_{nk}(t_2) e^{i\omega_{nk}t_2} \end{aligned}$$

where we have inserted the complete relation of $|\psi_n\rangle$. This is nothing but the second-order result $a_m^{(2)}(t)$ in (23).

VII. ADIABATIC APPROXIMATION AND BERRY'S PHASE

Imagine a perfect pendulum, with no friction or air resistance, oscillating back and forth in a vertical plane. If one grabs the support and shake it in a jerky manner, the bob will swing around in a wild chaotic fashion. But if I very gently and steadily move the support, the pendulum will continue to swing in a nice, smooth way, in the same plane (or one parallel to it) with the same amplitude. This gradual change in the external conditions characterizes an adiabatic process. Notice that there are two characteristic times involved: T_i , the "internal" time, representing the motion of the system itself (in this case the period of the pendulum's oscillations), and T_e , the "external" time, over which the parameters of the system change appreciably (if the pendulum were mounted on an oscillating platform, for example, T_e would be the period of the platform's motion). An adiabatic process is one for which $T_e \gg T_i$.

In quantum mechanics, the essential content of the adiabatic approximation can be cast in the form of a theorem. Suppose that the Hamiltonian changes gradually from some initial form H_i to some final form H_f . The **adiabatic theorem** states that if the particle was initially in the n th eigenstate of H_i , it will be carried into the n th eigenstate of H_f if we assume that the spectrum is discrete and non-degenerate throughout the transition from H_i to H_f .

For example, suppose we prepare a particle in the ground state of the infinite square well (Figure 17)

$$\psi^i(x) = \sqrt{\frac{2}{a}} \sin\left(\frac{\pi}{a}x\right)$$

If we now gradually move the right wall out to $2a$, the adiabatic theorem says that the particle will end up in the ground state of the expanded well

$$\psi^f(x) = \sqrt{\frac{1}{a}} \sin\left(\frac{\pi}{2a}x\right)$$

(apart, perhaps, from a phase factor). Notice that we're not talking about a *small* change in the Hamiltonian (as in perturbation theory) – this one is a *huge* change. All we require is that it happen *slowly*. Energy is not conserved here: whoever is moving the wall is extracting energy from the system, just like the piston on a slowly expanding cylinder of gas. By contrast, if the well expands *suddenly*, the resulting state is still $\psi^i(x)$, which is a complicated linear combination of eigenstates of the new Hamiltonian. In this case energy is conserved; as in free expansion of a gas into the vacuum when the barrier is suddenly removed, no work is done.

A. Proof of the adiabatic theorem

The adiabatic theorem is simple to state, and it sounds plausible, but it is not easy to prove. Suppose the the Hamiltonian is independent of time, then a particle which starts out in the n th eigenstate ψ_n

$$\hat{H}\psi_n = E_n\psi_n$$

remains in the n th eigenstate, simply picking up a phase factor

$$\Psi_n(t) = \psi_n e^{-iE_n t/\hbar}.$$

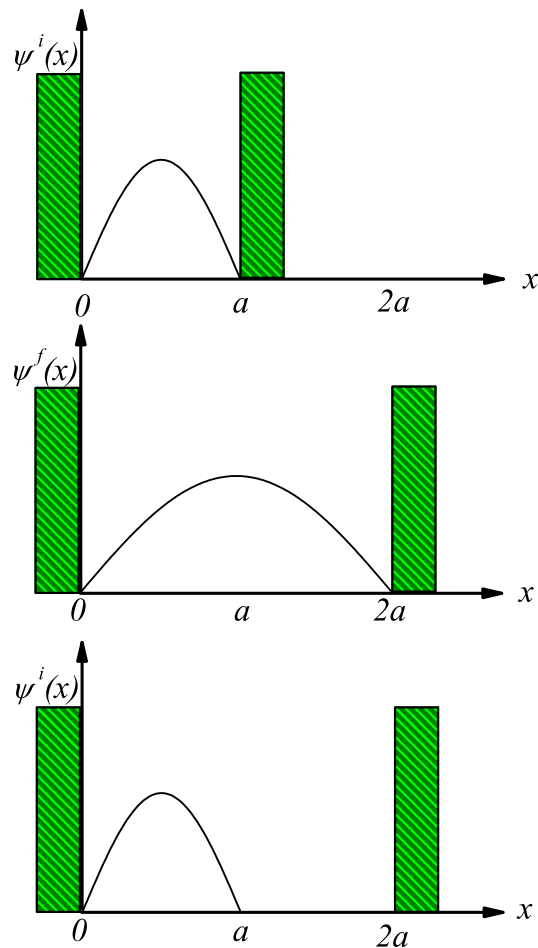


FIG. 17: Particle starts out in the ground state of the infinite square well. If the wall moves slowly, the particle remains in the ground state. If the wall moves rapidly, the particle is left (momentarily) in its initial state.

If the Hamiltonian changes with time, then the eigenfunctions and eigenvalues are themselves time dependent and we have the **instantaneous eigenequations**

$$\hat{H}(t)\psi_n(t) = E_n(t)\psi_n(t) \quad (39)$$

At any particular instant the eigenfunctions still constitute an orthonormal set

$$\langle \psi_n(t) | \psi_m(t) \rangle = \delta_{nm}$$

and they are complete, so the general solution to the time dependent Schrödinger equation

$$i\hbar \frac{\partial}{\partial t} \Psi(t) = \hat{H}(t) \Psi(t) \quad (40)$$

can be expressed as a linear combination of them

$$\Psi(t) = \sum_n c_n(t) \psi_n(t) e^{i\theta_n(t)} \quad (41)$$

where

$$\theta_n(t) \equiv -\frac{1}{\hbar} \int_0^t E_n(t') dt' \quad (42)$$

generalizes the “standard” phase factor to the case where E_n varies with time. Substitute equation (41) into (40) we obtain

$$i\hbar \sum_n \left(\dot{c}_n \psi_n + c_n \dot{\psi}_n + i c_n \psi_n \dot{\theta}_n \right) e^{i\theta_n} = \sum_n c_n \left(\hat{H} \psi_n \right) e^{i\theta_n}$$

(dot denotes the time derivative) In view of equations (39) and (42), the last two terms cancel, leaving

$$\sum_n \dot{c}_n \psi_n e^{i\theta_n} = - \sum_n c_n \dot{\psi}_n e^{i\theta_n}$$

Taking the inner product with ψ_m , and invoking the orthonormality of the instantaneous eigenequations

$$\sum_n \dot{c}_n \delta_{mn} e^{i\theta_n} = - \sum_n c_n \langle \psi_m | \dot{\psi}_n \rangle e^{i\theta_n}$$

or

$$\dot{c}_m(t) = - \sum_n c_n \langle \psi_m | \dot{\psi}_n \rangle e^{i(\theta_n - \theta_m)} \quad (43)$$

Now differentiating the instantaneous eigenequations (39) with respect to time yields

$$\dot{H} \psi_n + \hat{H} \dot{\psi}_n = \dot{E}_n \psi_n + E_n \dot{\psi}_n$$

and hence (again taking the inner product with ψ_m)

$$\langle \psi_m | \dot{H} | \psi_n \rangle + \langle \psi_m | \hat{H} | \dot{\psi}_n \rangle = \dot{E}_n \delta_{mn} + E_n \langle \psi_m | \dot{\psi}_n \rangle$$

Exploiting the hermiticity of \hat{H} to write $\langle \psi_m | \hat{H} | \dot{\psi}_n \rangle = E_m \langle \psi_m | \dot{\psi}_n \rangle$, it follows that for $n \neq m$

$$\langle \psi_m | \dot{H} | \psi_n \rangle = (E_n - E_m) \langle \psi_m | \dot{\psi}_n \rangle$$

Putting this into equation (43) and assuming the energies are non-degenerate we conclude that

$$\dot{c}_m(t) = -c_m \langle \psi_m | \dot{\psi}_m \rangle - \sum_{n \neq m} c_n \frac{\langle \psi_m | \dot{H} | \psi_n \rangle}{E_n - E_m} \exp \left(-\frac{i}{\hbar} \int_0^t (E_n(t') - E_m(t')) dt' \right)$$

This result is exact. Now comes the adiabatic approximation. Assume that \dot{H} is extremely small, and drop the second term, leaving

$$\dot{c}_m(t) = -c_m \langle \psi_m | \dot{\psi}_m \rangle$$

with the solution

$$c_m(t) = c_m(0) e^{i\gamma_m(t)}$$

where

$$\gamma_m(t) \equiv i \int_0^t \left\langle \psi_m(t') \left| \frac{\partial}{\partial t'} \psi_m(t') \right. \right\rangle dt'$$

In particular, if the particle starts out in the n th eigenstate, i.e. $c_n(0) = 1$ and $c_m(0) = 0$ for $m \neq n$, then at time t ,

$$\Psi_n(t) = e^{i\gamma_n(t)} e^{i\theta_n(t)} \psi_n(t) \quad (44)$$

it will remain in the n th eigenstate of the evolving Hamiltonian, picking up only a couple of phase factors.

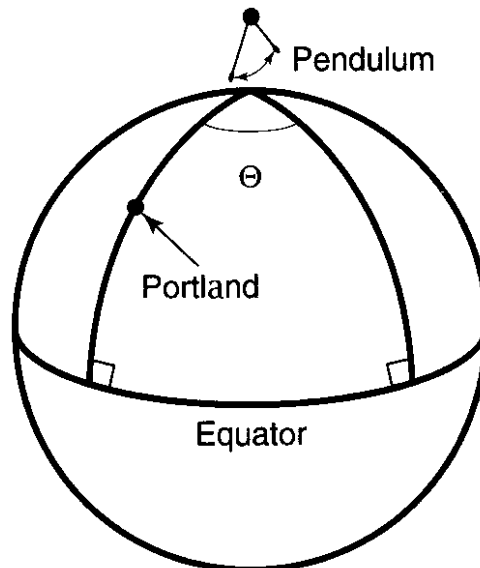


FIG. 18: Itinerary for adiabatic transport of a pendulum on the surface of the earth.

B. Nonholonomic Processes

Let us return to the classical model I used above to develop the notion of an adiabatic process: a perfectly frictionless pendulum whose support is carried around from place to place. I claimed that as long as the motion of the support is very slow, compared to the period of the pendulum (so that the pendulum executes many oscillations before the support has moved appreciably), it will continue to swing in the same plane (or one parallel to it), with the same amplitude (and, of course, the same frequency).

But what if I took this ideal pendulum up to the North Pole, and set it swinging – say, in the direction of Portland (Figure 18). (For the moment, I’ll pretend the earth is not rotating.) Very gently (that is, adiabatically), I carry it down the longitude line passing through Portland, and on beyond, down to the equator. At this stage it is swinging north-south. Now I carry it (still swinging north-south) partway around the equator. And finally, I carry it back up to the North Pole, along the new longitude line. It is clear that the pendulum will no longer be swinging in the same plane as it was when I set out—indeed, the new plane makes an angle Θ with the old one, where Θ is the angle between the southbound and the northbound longitude lines.

As it happens, Θ is equal to the *solid angle* (Ω) subtended (at the center of the earth) by the path around which I carried the pendulum. For this path surrounds a fraction $\Theta/2\pi$ of the northern hemisphere, so its area is $A = (1/2)(\Theta/2\pi)4\pi R^2 = \Theta R^2$ (where R is the radius of the earth), and hence

$$\Theta = A/R^2 \equiv \Omega$$

This is a particularly nice way to express the answer, because it turns out to be independent of the shape of the path (Figure 19).

Incidentally, the Foucault pendulum is an example of precisely this sort of adiabatic transport around a closed loop on a sphere—only this time instead of me carrying the pendulum around, I let the rotation of the earth do the job. For an explanation on why the Foucault pendulum proves the earth rotates, see animations and rollovers by the California Academy of Sciences. <http://www.calacademy.org/products/pendulum/index.html>

The solid angle subtended by a latitude line θ_0 (Figure 20) is

$$\Omega = \int \sin \theta d\theta d\phi = 2\pi (-\cos \theta) \Big|_0^{\theta_0} = 2\pi (1 - \cos \theta_0)$$

Relative to the earth (which has meanwhile turned through an angle of 2π), the daily precession of the Foucault pendulum is $2\pi \cos \theta_0$ – a result that is ordinarily obtained by appeal to Coriolis forces in the rotating reference frame, but is seen in this context to admit a purely geometrical interpretation. For example, at the North Pole, the daily precession is 360° while at San Francisco, it is about 225° .

A system such as this, which does not return to its original state when transported around a closed loop, is said to be nonholonomic. My project here is to study the quantum mechanics of nonholonomic, adiabatic processes. The

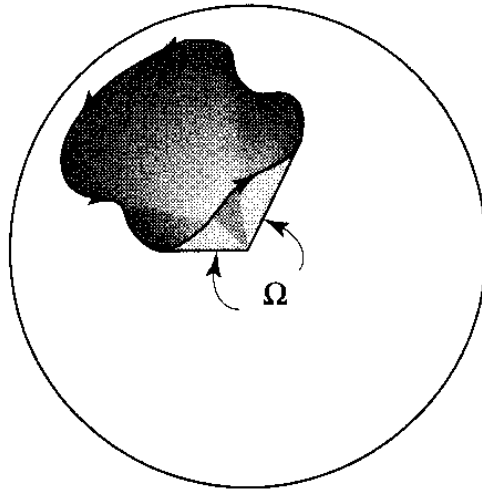


FIG. 19: Arbitrary path on the surface of a sphere, subtending a solid angle Ω .

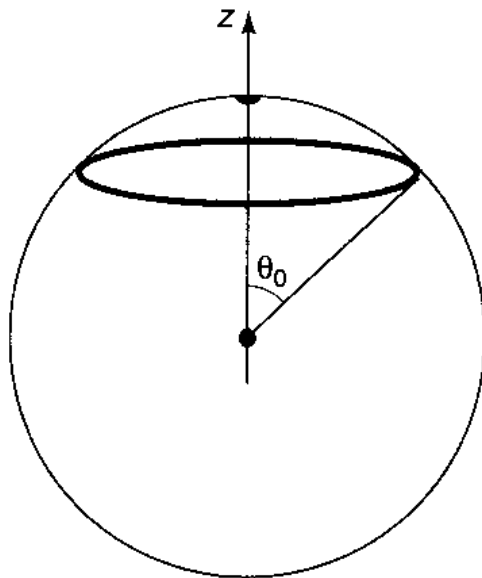


FIG. 20: Path of a Foucault pendulum in the course of one day.

essential question is this: How does the final state differ from the initial state, if the parameters in the Hamiltonian are carried adiabatically around some closed cycle?

C. Berry's Phase

We have showed that a particle which starts out in the n th eigenstate of $\hat{H}(0)$ remains, under adiabatic conditions, in the n th eigenstate of $\hat{H}(t)$, picking up only a time-dependent phase factor. Specifically, its wave function is

$$\Psi_n(t) = e^{i[\theta_n(t) + \gamma_n(t)]} \psi_n(t)$$

where

$$\theta_n(t) = -\frac{1}{\hbar} \int_0^t E_n(t') dt'$$

is the **dynamic phase** (generalizing the “standard” factor $e^{-iE_n t/\hbar}$ to the case where E_n is a function of time) and

$$\gamma_n(t) = i \int_0^t \left\langle \psi_n(t') \left| \frac{\partial}{\partial t'} \psi_n(t') \right. \right\rangle dt'$$

is the so-called **geometric phase**.

Now $\psi_n(t)$ depends on time because there is some parameter $R(t)$ in the Hamiltonian that is changing with time. In the example Figure 17, $R(t)$ would be the width of the expanding square well. Thus

$$\frac{\partial \psi_n}{\partial t} = \frac{\partial \psi_n}{\partial R} \frac{dR}{dt} \quad (45)$$

so

$$\gamma_n(t) = i \int_0^t \left\langle \psi_n(t') \left| \frac{\partial \psi_n}{\partial R} \right. \right\rangle \frac{dR}{dt'} dt' = i \int_{R_i}^{R_f} \left\langle \psi_n \left| \frac{\partial \psi_n}{\partial R} \right. \right\rangle dR$$

where R_i and R_f are the initial and final values of $R(t)$. In particular, if the Hamiltonian returns to its original form after time T , so that $R_i = R_f$, then $\gamma_n(T) = 0$ - nothing very interesting there!

However, I assumed in equation (45) that there is only one parameter in the Hamiltonian that is changing. Suppose there are N of them: $R_1(t), R_2(t), \dots, R_N(t)$; in that case

$$\frac{\partial \psi_n}{\partial t} = \frac{\partial \psi_n}{\partial R_1} \frac{dR_1}{dt} + \frac{\partial \psi_n}{\partial R_2} \frac{dR_2}{dt} + \dots + \frac{\partial \psi_n}{\partial R_N} \frac{dR_N}{dt} = (\nabla_R \psi_n) \cdot \frac{d\mathbf{R}}{dt}$$

where $\mathbf{R} = (R_1, R_2, \dots, R_N)$, and ∇_R is the gradient with respect to these parameters. This time we have

$$\gamma_n(t) = i \int_{\mathbf{R}_i}^{\mathbf{R}_f} \langle \psi_n | \nabla_R \psi_n \rangle \cdot d\mathbf{R}$$

and if the Hamiltonian returns to its original form after a time T , the net geometric phase change is

$$\gamma_n(T) = i \oint \langle \psi_n | \nabla_R \psi_n \rangle \cdot d\mathbf{R} \quad (46)$$

This is a line integral around a closed loop in parameter space, and it is not, in general, zero. Equation (46) was first obtained by M. V. Berry in 1984, and $\gamma_n(T)$ is called **Berry's phase**. Notice that $\gamma_n(T)$ depends only on the path taken, not on how fast that path is traversed (provided, of course, that it is slow enough to validate the adiabatic hypothesis). By contrast, the accumulated dynamic phase,

$$\theta_n(T) = -\frac{1}{\hbar} \int_0^T E_n(t') dt'$$

depends critically on the elapsed time. It is easy to check that Berry's phase is real and can be measured. In fact

$$\nabla_R \langle \psi_n | \psi_n \rangle = 0$$

because ψ_n is assumed normalized. So

$$\begin{aligned} \langle \nabla_R \psi_n | \psi_n \rangle + \langle \psi_n | \nabla_R \psi_n \rangle &= \langle \psi_n | \nabla_R \psi_n \rangle^* + \langle \psi_n | \nabla_R \psi_n \rangle = 0 \\ \langle \psi_n | \nabla_R \psi_n \rangle &\text{ is imaginary} \Rightarrow \gamma_n(t) \text{ is real.} \end{aligned}$$

Another question is that: Is Berry's phase measurable? We are accustomed to thinking that the phase of the wave function is arbitrary – physical quantities involve $|\Psi|^2$, and the phase factor cancels out. But $\gamma_n(T)$ can be measured, if (for example) we take a beam of particles (all in the state Ψ) and split it in two, so that one beam passes through an adiabatically changing potential, while the other does not. When the two beams are recombined, the total wave function has the form

$$\Psi = \frac{1}{2} \Psi_0 + \frac{1}{2} \Psi_0 e^{i\Gamma}$$

where Ψ_0 is the "direct" beam wave function, and Γ is the extra phase (in part dynamic, and in part geometric) acquired by the beam subjected to the varying \hat{H} . In this case

$$\begin{aligned} |\Psi|^2 &= \frac{1}{4} |\Psi_0|^2 (1 + e^{i\Gamma}) (1 + e^{-i\Gamma}) \\ &= \frac{1}{2} |\Psi_0|^2 (1 + \cos \Gamma) = |\Psi_0|^2 \cos^2 (\Gamma/2) \end{aligned}$$

So by looking for points of constructive and destructive interference (where Γ is an even or odd multiple of π , respectively), one can easily measure Γ . Berry, and other early writers, worried that the geometric phase might be swamped by a larger dynamic phase, but it has proved possible to arrange things so as to separate out the two contributions.

VIII. SUMMARY ON PART V

There are quite few examples in Quantum Mechanics which permit exact solutions. We need in most cases the approximation methods. The stationary perturbation theory can be applied to evaluate the corrections to the discrete eigen-energy levels and eigenfunctions if the Hamiltonian \hat{H} can be separated into two parts, \hat{H}_0 , whose solutions are exactly known, and the small perturbation \hat{H}' satisfying $H'_{mn} \ll |E_m^{(0)} - E_n^{(0)}|$.

For non-degenerate case one usually calculate the corrections to energy levels to second order

$$E_n = E_n^{(0)} + H'_{nn} + \sum_{m \neq n} \frac{|H'_{mn}|^2}{E_n^{(0)} - E_m^{(0)}} + \dots$$

and the corrections to eigenfunctions to first order

$$\psi_n = \psi_n^{(0)} + \sum_{m \neq n} \frac{H'_{mn}}{E_n^{(0)} - E_m^{(0)}} \psi_m^{(0)} + \dots$$

For degenerate case, by solving the secular equation

$$\begin{vmatrix} H'_{11} - E_n^{(1)} & H'_{12} & \dots & H'_{1k} \\ H'_{21} & H'_{22} - E_n^{(1)} & \dots & H'_{2k} \\ \dots & \dots & \dots & \dots \\ H'_{k1} & H'_{k2} & \dots & H'_{kk} - E_n^{(1)} \end{vmatrix} = 0$$

one obtains in one stroke both (1) the k possible first order energy shifts $E_n^{(1)}$ (the k roots of the above algebraic equation) and (2) the k possible correct zeroth-order wavefunctions (the corresponding k eigenfunctions $\{c_i^{(0)}\}$, which is nothing but the combination coefficients of ϕ_i - the k degenerate orthonormal eigenstates of \hat{H}_0 corresponding to the same energy $E_n^{(0)}$). One good example of this is the Stark effect of atoms under electric field.

The variational method on the other hand is very useful for estimating the ground state energy E_0 when the exact solutions are not available for \hat{H}_0 . We get closer to the exact result if we use a better trial function. A sophisticated calculation for the ground state of helium atom gives

$$\langle H(\lambda) \rangle_{\min} = -2.85 \frac{e^2}{a_0},$$

a better result than that from perturbation theory.

Time-dependent perturbation theory is responsible for quantum transition probability. To first order the transition probability from initial state k to final state m reads

$$P_{mk}(t) = \left| a_m^{(1)}(t) \right|^2 = \frac{1}{\hbar^2} \left| \int_0^t H'_{mk}(t_1) e^{i\omega_{mk}t_1} dt_1 \right|^2$$

For constant perturbation the transition rate is determined by the Fermi's Golden rule

$$w = \frac{2\pi}{\hbar} |H'_{mk}|^2 \rho(E_m)$$

while for sinusoidal perturbation $\hat{H}'(t) = \hat{F} (e^{i\omega t} + e^{-i\omega t})$, we have

$$w = \frac{2\pi}{\hbar} |F_{mk}|^2 \rho(E_m \pm \hbar\omega)$$

A semiclassical treatment for the absorption and emission of radiation is developed as an application of the time-dependent perturbation theory. One has the Einstein coefficients

$$B_{mk} = B_{km} = \frac{4\pi^2 e^2}{3\hbar^2} |\mathbf{r}_{mk}|^2$$

$$A_{km} = \frac{\hbar\omega_{mk}^3}{\pi^2 c^3} B_{km}$$

and the selection rules for electric dipole moment transition

$$\text{No transition occurs unless } \Delta l = \pm 1, \Delta m = 0, \pm 1$$

Three pictures are equivalent in describing the evolution of the system and the main differences are already summarized in the two tables. We emphasize that the interaction picture is often used in the quantum transition problem. Finally the adiabatic approximation and Berry's Phase are still very hot fields in current research. We mention here only two key formula

$$\Psi_n(t) = e^{i\gamma_n(t)} e^{i\theta_n(t)} \psi_n(t)$$

$$\gamma_n(T) = i \oint \langle \psi_n | \nabla_R \psi_n \rangle \cdot d\mathbf{R}$$

Lecture Notes on Quantum Mechanics - Part VI

Yunbo Zhang

Instituter of Theoretical Physics, Shanxi University

In this part we develop a systematic treatment of angular momentum and related topics. Spin, a special quantum mechanical angular momentum without classical counterpart, will be introduced to distinguish Bosons and Fermions. Permutation of identical particles gives symmetric requirement of the many-particle wavefunctions. Coupling of two angular momenta leads to the famous Clebsch-Gordon coefficients.

Contents

I. Rotations and Angular Momentum	2
A. Commutation relation	2
B. Eigenvalue and eigenvector of angular momentum	5
C. Matrix elements of angular momentum operators	7
II. Spin	9
A. The Uhlenbeck and Goudsmit proposal	10
B. Description of spin state and spin operators	10
C. Pauli matrices	11
D. The complete wave function of an electron	12
III. Addition of of angular momenta	14
A. Complete set of commutation operators in two representations	14
B. Transformation between coupled state and uncoupled state	15
IV. Fine structure of alkali spectra	19
A. Spin-orbit coupling	20
B. Perturbation calculation of spin-orbit coupling energy	21
C. The perturbation energy	22
V. Charged particle in electromagnetic field - Zeeman effect	23
A. Hamiltonian	23
B. Magnetic field	24
C. Zeeman effect	25
1. Normal Zeeman effect	25
2. Anomalous Zeeman effect	26
VI. Spin singlet and triplet	27
VII. Identical Particles	31
A. Bosons and Fermions	31
B. Symmetrization of wavefunctions	33
VIII. Summary of Part VI	37

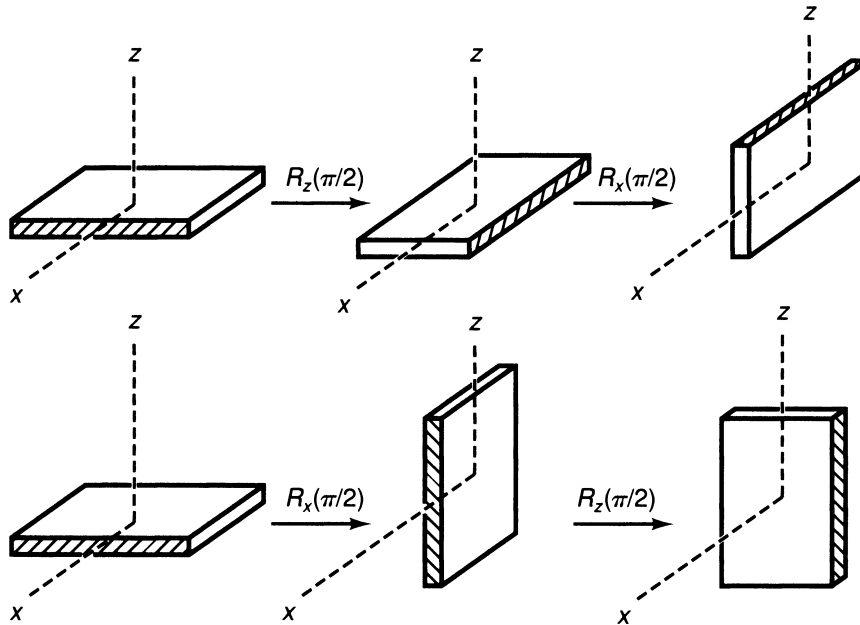


FIG. 1: Example for illustrating the noncommutative property of finite rotations.

I. ROTATIONS AND ANGULAR MOMENTUM

A. Commutation relation

Angular momentum is closely related to the rotation in 3D. Rotations about the same axis commute, whereas rotations about different axes do not. The result of successive rotation depends on the order of rotations. For example, let us consider a 90° rotation about the z -axis, denoted by $R_z(\pi/2)$, followed by a 90° rotation about the x -axis, denoted by $R_x(\pi/2)$; compare this with a 90° rotation about the x -axis, followed by a 90° rotation about the z -axis. The final results are different, as can be seen in Figure 1. So the order of successive finite rotations doesn't commute.

Now we consider an infinitesimal rotation $\delta\alpha$ around axis \mathbf{n} . As in Part IV of this lecture notes, the convention we follow here is that a rotation operation affects the physical system itself as in Figure 1, while the coordinate axes remain unchanged. Had we adopted a different convention, in which a physical system remained fixed but the coordinate axes rotated, the transformation matrix is the transpose of physical system rotated convention. It is obviously important to distinguish the two approaches by using “active rotations” for physical system rotated and “passive rotations” for the coordinate axes rotated. See Ballentine's book Section 7.5 for clarification.

It is important to be aware that when the abstract vector $\Psi(\mathbf{r})$ is represented as a function of space-time coordinates \mathbf{r} , there is an inverse relation between transformations on function space and transformations on coordinates. Transformation invariance means that the value of the new function at the transformed point is the same as the value of the original function at the old point, i.e. $\Psi'(\mathbf{r}') = \Psi(\mathbf{r})$. For a rotation in Euler's angles,

$$R(\alpha, \beta, \gamma)\Psi(\mathbf{r}) = \Psi(R^{-1}(\alpha, \beta, \gamma)\mathbf{r}).$$

So an active rotation will rotate the radius vector \mathbf{r} into

$$R_{\mathbf{n}}(\delta\alpha)\mathbf{r} = \mathbf{r} + \delta\mathbf{r} = \mathbf{r} + \delta\alpha(\mathbf{n} \times \mathbf{r})$$

as we can see from Figure 2 in Part IV.

Two successive infinitesimal rotations, first by $\delta\alpha$ around x -axis, followed by $\delta\beta$ around y -axis, change the vector like

$$\mathbf{r} \rightarrow \mathbf{r}' \rightarrow \mathbf{r}''$$

or

$$\begin{aligned} R_x(\delta\alpha)\mathbf{r} &= \mathbf{r}' = \mathbf{r} + \delta\alpha(\mathbf{x}_0 \times \mathbf{r}) \\ R_y(\delta\beta)\mathbf{r}' &= \mathbf{r}'' = \mathbf{r}' + \delta\beta(\mathbf{y}_0 \times \mathbf{r}') \end{aligned}$$

Notice the vector product formula

$$\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) = (\mathbf{A} \cdot \mathbf{C})\mathbf{B} - (\mathbf{A} \cdot \mathbf{B})\mathbf{C}$$

the net result is

$$\begin{aligned} R_y(\delta\beta)R_x(\delta\alpha)\mathbf{r} &= R_y(\delta\beta)\mathbf{r}' = \mathbf{r}' + \delta\beta(\mathbf{y}_0 \times \mathbf{r}') \\ &= \mathbf{r} + \delta\alpha(\mathbf{x}_0 \times \mathbf{r}) + \delta\beta[\mathbf{y}_0 \times (\mathbf{r} + \delta\alpha(\mathbf{x}_0 \times \mathbf{r}))] \\ &= \mathbf{r} + \delta\alpha(\mathbf{x}_0 \times \mathbf{r}) + \delta\beta(\mathbf{y}_0 \times \mathbf{r}) + \delta\alpha\delta\beta[\mathbf{y}_0 \times (\mathbf{x}_0 \times \mathbf{r})] \\ &= \mathbf{r} + \delta\alpha(\mathbf{x}_0 \times \mathbf{r}) + \delta\beta(\mathbf{y}_0 \times \mathbf{r}) + \delta\alpha\delta\beta[(\mathbf{y}_0 \cdot \mathbf{r})\mathbf{x}_0 - (\mathbf{y}_0 \cdot \mathbf{x}_0)\mathbf{r}] \end{aligned}$$

Let us now change the order of rotation, first by $\delta\beta$ around y -axis, followed by $\delta\alpha$ around x -axis, we have

$$R_x(\delta\alpha)R_y(\delta\beta)\mathbf{r} = \mathbf{r} + \delta\alpha(\mathbf{x}_0 \times \mathbf{r}) + \delta\beta(\mathbf{y}_0 \times \mathbf{r}) + \delta\alpha\delta\beta[(\mathbf{x}_0 \cdot \mathbf{r})\mathbf{y}_0 - (\mathbf{x}_0 \cdot \mathbf{y}_0)\mathbf{r}]$$

It follows that the difference of them is

$$\begin{aligned} \{R_y(\delta\beta)R_x(\delta\alpha) - R_x(\delta\alpha)R_y(\delta\beta)\}\mathbf{r} &= \delta\alpha\delta\beta[(\mathbf{y}_0 \cdot \mathbf{r})\mathbf{x}_0 - (\mathbf{x}_0 \cdot \mathbf{r})\mathbf{y}_0] \\ &= \delta\alpha\delta\beta[(\mathbf{r} \cdot \mathbf{y}_0)\mathbf{x}_0 - (\mathbf{r} \cdot \mathbf{x}_0)\mathbf{y}_0] = \delta\alpha\delta\beta(\mathbf{r} \times (\mathbf{x}_0 \times \mathbf{y}_0)) = \delta\alpha\delta\beta(\mathbf{r} \times \mathbf{z}_0) \\ &= -\delta\alpha\delta\beta(\mathbf{z}_0 \times \mathbf{r}) = \mathbf{r} - R_z(\delta\alpha\delta\beta)\mathbf{r} \\ &= \{R_{\mathbf{n}}(0) - R_z(\delta\alpha\delta\beta)\}\mathbf{r} \end{aligned}$$

where we have used

$$1 = R_{\mathbf{n}}(0)$$

with \mathbf{n} stands for any rotation axis. Thus the final result can be written as

$$R_y(\delta\beta)R_x(\delta\alpha) - R_x(\delta\alpha)R_y(\delta\beta) = R_{\mathbf{n}}(0) - R_z(\delta\alpha\delta\beta) \quad (1)$$

This is an example of the commutation relations between rotation operators about different axes. We use this to deduce the commutation relation of angular momentum in quantum mechanics.

So far we have not used quantum mechanical concepts. The matrix R is just a 3×3 orthogonal matrix acting on a vector. In quantum mechanics, however, wavefunctions live on Hilbert space. We use Dirac's notation to represent a state ket $|\alpha\rangle$. Given a rotation operation R , characterized by a 3×3 orthogonal matrix R , we associate an operator $D(R)$ in the appropriate ket space such that

$$|\alpha\rangle_R = D(R)|\alpha\rangle$$

Note that the 3×3 orthogonal matrix acts on a column matrix made up of the three components of a classical vector, while the operator $D(R)$ acts on state vectors in ket space. The matrix representation of $D(R)$ depends on the dimensionality of the particular ket space in question. For $N = 2$, appropriate for describing a spin-1/2 system with no other degree of freedom, $D(R)$ is represented by a 2×2 matrix; for a spin-1 system, the appropriate representation is a 3×3 unitary matrix, and so on.

To construct the rotation operator $D(R)$, it is again fruitful to examine first its properties under an infinitesimal rotation. The appropriate infinitesimal operator could be written as

$$\hat{Q} = I - i\varepsilon\hat{F} \quad (2)$$

with a Hermitian operator \hat{F} which is called the generator of transformation \hat{Q} . Specifically,

$$\hat{F} \rightarrow \frac{\hat{p}_x}{\hbar}, \varepsilon \rightarrow \delta x$$

for an infinitesimal translation by a displacement δx in the x -direction, and

$$\hat{F} \rightarrow \frac{\hat{H}}{\hbar}, \varepsilon \rightarrow \delta t$$

for an infinitesimal time evolution with time displacement δt . Similarly we define the angular momentum operator \hat{J}_k in such a way that the operator for an infinitesimal rotation around the k -th axis by an angle $\delta\alpha$ can be obtained by letting

$$\hat{F} \rightarrow \frac{\hat{J}_k}{\hbar}, \varepsilon \rightarrow \delta\alpha$$

in equation (2). More generally, we have

$$D_{\mathbf{n}}(\delta\alpha) = 1 - \frac{i}{\hbar}(\hat{\mathbf{J}} \cdot \mathbf{n})\delta\alpha \quad (3)$$

for a rotation about \mathbf{n} by $\delta\alpha$ (Notice that minus signs are missed in eqs. (7.17) (7.20) and (7.21) in Ballentine's book). Specifically

$$\begin{aligned} D_x(\delta\alpha) &= 1 - \frac{i}{\hbar}\delta\alpha\hat{J}_x \\ D_y(\delta\beta) &= 1 - \frac{i}{\hbar}\delta\beta\hat{J}_y \\ D_z(\delta\alpha\delta\beta) &= 1 - \frac{i}{\hbar}\delta\alpha\delta\beta\hat{J}_z \end{aligned}$$

A finite rotation can be obtained by compounding successively infinitesimal rotations about the same axis. For instance, if we are interested in a finite rotation around z -axis by angle ϕ , we consider

$$\begin{aligned} D_z(\phi) &= \lim_{N \rightarrow \infty} \left(1 - i \frac{\hat{J}_z}{\hbar} \frac{\phi}{N} \right)^N \\ &= \exp \left(- \frac{i\hat{J}_z\phi}{\hbar} \right) \end{aligned}$$

We have from the rotation analogy of equation (1)

$$\begin{aligned} \left(1 - \frac{i}{\hbar}\delta\beta\hat{J}_y \right) \left(1 - \frac{i}{\hbar}\delta\alpha\hat{J}_x \right) - \left(1 - \frac{i}{\hbar}\delta\alpha\hat{J}_x \right) \left(1 - \frac{i}{\hbar}\delta\beta\hat{J}_y \right) &= 1 - \left(1 - \frac{i}{\hbar}\delta\alpha\delta\beta\hat{J}_z \right) \\ \left(\frac{i}{\hbar} \right)^2 \delta\alpha\delta\beta \left(\hat{J}_y\hat{J}_x - \hat{J}_x\hat{J}_y \right) &= \frac{i}{\hbar}\delta\alpha\delta\beta\hat{J}_z \end{aligned}$$

Because $\delta\alpha$ and $\delta\beta$ are two arbitrary infinitesimal rotation angles, we thus have the commutation relation

$$\hat{J}_x\hat{J}_y - \hat{J}_y\hat{J}_x = i\hbar\hat{J}_z \quad (4)$$

By the same token

$$\begin{aligned} \hat{J}_y\hat{J}_z - \hat{J}_z\hat{J}_y &= i\hbar\hat{J}_x \\ \hat{J}_z\hat{J}_x - \hat{J}_x\hat{J}_z &= i\hbar\hat{J}_y \end{aligned}$$

which are exactly the same results as we have got in Part IV by defining the angular momentum to be $\hat{\mathbf{r}} \times \hat{\mathbf{p}}$. In general, when the generators of infinitesimal transformation do not commute, the corresponding group of operation is said to be *non-Abelian*. Because of equation (4), the rotation group in 3D is non-Abelian. In contrast, the translation group in 3D is Abelian because \hat{p}_i and \hat{p}_j commute even with $i \neq j$.

We emphasize that in obtaining the commutation relations (4), we have used the following two concepts:

- \hat{J}_k is the generator of rotation about the k -th axis
- Rotations about different axes fail to commute.

It is no exaggeration to say that these commutation relations summarize in a compact manner all the basic properties of rotations in 3D.

B. Eigenvalue and eigenvector of angular momentum

We use a new symbol $\hat{\mathbf{J}}$ to represent the angular momentum because the discussion here applies to angular momentum operators defined more generally by equation (3), which may be orbital, spin, total, even isospin angular momentum. This is important because spin angular momentum, to which our general formalism also applies, has nothing to do with $\hat{\mathbf{r}}$ and $\hat{\mathbf{p}}$. We can construct the Casimir operator

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

which commute each component of $\hat{\mathbf{J}}$. The fundamental commutation relations for general angular momentum reads

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

and

$$[\hat{J}^2, \hat{J}_i] = 0, \quad i = x, y, z$$

We work out the eigenvalues and eigenfunctions of \hat{J}^2 and \hat{J}_z and derive the expression for matrix elements of angular momentum operators, first obtained in a 1926 paper by M. Born, W. Heisenberg and P. Jordan.

Define two non-hermitian operators

$$\hat{J}_{\pm} = \hat{J}_x \pm i\hat{J}_y$$

which are known as *ladder operators*. It follows that

$$\begin{aligned} [\hat{J}^2, \hat{J}_{\pm}] &= 0 \\ [\hat{J}_z, \hat{J}_{\pm}] &= \pm\hbar\hat{J}_{\pm} \\ [\hat{J}_+, \hat{J}_-] &= 2\hbar\hat{J}_z \end{aligned}$$

Indeed,

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

$$[\hat{J}_z, \hat{J}_-] = i\hbar\hat{J}_y + i(i\hbar\hat{J}_x) = -\hbar\hat{J}_x + i\hbar\hat{J}_y = -\hbar\hat{J}_-$$

and

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

We denote the eigenfunctions of \hat{J}^2, \hat{J}_z by $|j, m\rangle$

$$\begin{aligned} \hat{J}^2 |j, m\rangle &= \lambda_j |j, m\rangle \\ \hat{J}_z |j, m\rangle &= m\hbar |j, m\rangle \end{aligned}$$

both eigenvalues λ_j and m are to be determined. What is the physical meaning of \hat{J}_\pm ? To answer this, let us examine how \hat{J}_z act on $\hat{J}_\pm |j, m\rangle$. It is easy to show that $\hat{J}_\pm |j, m\rangle$ are also eigenvectors of \hat{J}_z

$$\begin{aligned}\hat{J}_z(\hat{J}_+ |j, m\rangle) &= (\hat{J}_+ \hat{J}_z + \hbar \hat{J}_+) |j, m\rangle = (m+1)\hbar(\hat{J}_+ |j, m\rangle) \\ \hat{J}_z(\hat{J}_- |j, m\rangle) &= (\hat{J}_- \hat{J}_z - \hbar \hat{J}_-) |j, m\rangle = (m-1)\hbar(\hat{J}_- |j, m\rangle)\end{aligned}$$

except that its eigenvalue is now increased (decreased) by one unit of \hbar becoming $(m \pm 1)\hbar$. So now we see why \hat{J}_\pm , which step one step up (down) on the “ladder” of \hat{J}_z eigenvalues, are known as the ladder operators.

Then start from a simultaneous eigenvector $\hat{J}^2, \hat{J}_z, |j, m\rangle$, we could apply \hat{J}_+ successively, say n times, to obtain another eigenvectors of \hat{J}^2, \hat{J}_z with the \hat{J}_z eigenvalue increased by $n\hbar$, while its \hat{J}^2 eigenvalue is unchanged

$$\begin{aligned}\cdots \leftarrow \hat{J}_-^2 |j, m\rangle \leftarrow \hat{J}_- |j, m\rangle \leftarrow |j, m\rangle \rightarrow \hat{J}_+ |j, m\rangle \rightarrow \hat{J}_+^2 |j, m\rangle \rightarrow \cdots \\ \lambda_j \qquad \lambda_j \qquad \lambda_j \qquad \lambda_j \qquad \lambda_j \\ (m-2)\hbar \quad (m-1)\hbar \quad m\hbar \quad (m+1)\hbar \quad (m+2)\hbar\end{aligned}$$

However, this process can not go on forever: Eventually we’re going to reach a state for which the z -component exceeds the total, and that cannot be. It turns out that there must exist two eigenvectors $|j, m_{\max}\rangle$ and $|j, m_{\min}\rangle$, such that

$$\hat{J}_+ |j, m_{\max}\rangle = 0 \quad (5)$$

$$\hat{J}_- |j, m_{\min}\rangle = 0 \quad (6)$$

Stated another way, the eigenvalue of m can not be increased beyond m_{\max} (decreased below m_{\min}).

Equation (5) also implies

$$\hat{J}_- \hat{J}_+ |j, m_{\max}\rangle = 0$$

But

$$\hat{J}_- \hat{J}_+ = \hat{J}^2 - \hat{J}_z^2 - \hbar \hat{J}_z$$

so

$$\left(\hat{J}^2 - \hat{J}_z^2 - \hbar \hat{J}_z \right) |j, m_{\max}\rangle = 0$$

Because $|j, m_{\max}\rangle$ itself is not a null ket, this relationship is possible only if

$$\lambda_j - m_{\max}^2 \hbar^2 - m_{\max} \hbar^2 = 0$$

or

$$\lambda_j = m_{\max}(m_{\max} + 1)\hbar^2 \quad (7)$$

In a similar manner, we can argue that Equation (6) also implies

$$\hat{J}_+ \hat{J}_- |j, m_{\min}\rangle = 0$$

By writing $\hat{J}_+ \hat{J}_-$ as

$$\hat{J}_+ \hat{J}_- = \hat{J}^2 - \hat{J}_z^2 + \hbar \hat{J}_z$$

and

$$\begin{aligned}\left(\hat{J}^2 - \hat{J}_z^2 + \hbar \hat{J}_z \right) |j, m_{\min}\rangle &= 0 \\ \lambda_j - m_{\min}^2 \hbar^2 + m_{\min} \hbar^2 &= 0\end{aligned}$$

we conclude that

$$\lambda_j = m_{\min}(m_{\min} - 1)\hbar^2 \quad (8)$$

By comparing (7) and (8)

$$m_{\max}(m_{\max} + 1) = m_{\min}(m_{\min} - 1)$$

we infer that

$$m_{\min} = -m_{\max}$$

or

$$m_{\min} = m_{\max} + 1$$

which can not happen. Call $m_{\max} = j$, which is positive, and the allowed values of m lie within

$$-j \leq m \leq j$$

Clearly, we must be able to reach $|j, j\rangle$ by applying \hat{J}_+ successively to $|j, -j\rangle$ a finite number of times. We must therefore have

$$j - (-j) = 2j = \text{positive integer}$$

As a result we get

$$j = \frac{n}{2}$$

i.e., j is either an integer or a half-integer.

$$j : 0, \frac{1}{2}, 1, \frac{3}{2}, \dots$$

while m may take $2j + 1$ different values $j, j - 1, j - 2, \dots - j$. The basic eigenvalue equations now read

$$\begin{aligned} \hat{J}^2 |j, m\rangle &= j(j+1)\hbar^2 |j, m\rangle \\ \hat{J}_z |j, m\rangle &= m\hbar |j, m\rangle \end{aligned} \quad (9)$$

C. Matrix elements of angular momentum operators

When we try to figure out the matrix elements of various angular momentum operators, only eigenvectors of the same j should be taken into account because elements of different j 's all vanish trivially. Assume $|j, m\rangle$ is normalized, we obviously have

$$\begin{aligned} \langle j', m' | \hat{J}^2 |j, m\rangle &= j(j+1)\hbar^2 \delta_{j'j} \delta_{m'm} \\ \langle j', m' | \hat{J}_z |j, m\rangle &= m\hbar \delta_{j'j} \delta_{m'm} \end{aligned}$$

To obtain the matrix elements of \hat{J}_{\pm} , we first consider

$$\begin{aligned} \langle j, m | \hat{J}_+^\dagger \hat{J}_+ |j, m\rangle &= \langle j, m | (\hat{J}^2 - \hat{J}_z^2 - \hbar \hat{J}_z) |j, m\rangle \\ &= \hbar^2 [j(j+1) - m^2 - m] \end{aligned} \quad (10)$$

Now $\hat{J}_+ |j, m\rangle$ must be the same as $|j, m+1\rangle$ (normalized) up to a multiplication constant. Thus

$$\hat{J}_+ |j, m\rangle = c_{jm}^+ |j, m+1\rangle$$

Comparison with (10) leads to

$$|c_{jm}^+|^2 = \hbar^2 [j(j+1) - m(m+1)]$$

So we have determined c_{jm}^+ up to an arbitrary phase factor. It is customary to choose it to be real and positive by convention. So

$$\hat{J}_+ |j, m\rangle = \sqrt{(j-m)(j+m+1)}\hbar |j, m+1\rangle \quad (11)$$

Similarly we can derive

$$\hat{J}_- |j, m\rangle = \sqrt{(j+m)(j-m+1)}\hbar |j, m-1\rangle \quad (12)$$

Finally we determine the matrix elements of \hat{J}_\pm to be

$$\langle j', m' | \hat{J}_\pm | j, m \rangle = \sqrt{(j \mp m)(j \pm m + 1)}\hbar \delta_{j'j} \delta_{m'm \pm 1}$$

Let us turn to some examples. The simplest case is the matrices for $j = \frac{1}{2}$ with two eigenstates

$$|j, m\rangle = \left| \frac{1}{2}, \frac{1}{2} \right\rangle, \left| \frac{1}{2}, -\frac{1}{2} \right\rangle$$

We have

$$\hat{J}^2 |j, m\rangle = \hat{J}^2 \left| \frac{1}{2}, m \right\rangle = \frac{1}{2} \left(\frac{1}{2} + 1 \right) \hbar^2 \left| \frac{1}{2}, m \right\rangle = \frac{3}{4} \hbar^2 \left| \frac{1}{2}, m \right\rangle$$

The matrices of \hat{J}^2 and \hat{J}_z are 2×2 and diagonal

$$\hat{J}^2 = \begin{pmatrix} \frac{3}{4}\hbar^2 & 0 \\ 0 & \frac{3}{4}\hbar^2 \end{pmatrix} = \frac{3}{4}\hbar^2 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$\hat{J}_z = \begin{pmatrix} \frac{1}{2}\hbar & 0 \\ 0 & -\frac{1}{2}\hbar \end{pmatrix} = \frac{1}{2}\hbar \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

while for \hat{J}_\pm , simple calculation

$$\hat{J}_+ \left| \frac{1}{2}, \frac{1}{2} \right\rangle = 0, \quad \hat{J}_+ \left| \frac{1}{2}, -\frac{1}{2} \right\rangle = \hbar \left| \frac{1}{2}, \frac{1}{2} \right\rangle$$

$$\hat{J}_- \left| \frac{1}{2}, \frac{1}{2} \right\rangle = \hbar \left| \frac{1}{2}, -\frac{1}{2} \right\rangle, \quad \hat{J}_- \left| \frac{1}{2}, -\frac{1}{2} \right\rangle = 0$$

gives the matrices

$$\hat{J}_+ = \begin{pmatrix} 0 & \hbar \\ 0 & 0 \end{pmatrix}, \quad \hat{J}_- = \begin{pmatrix} 0 & 0 \\ \hbar & 0 \end{pmatrix}$$

Therefore

$$\hat{J}_x = \frac{1}{2}(\hat{J}_+ + \hat{J}_-) = \frac{\hbar}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \hat{J}_y = \frac{1}{2i}(\hat{J}_+ - \hat{J}_-) = \frac{\hbar}{2} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$$

The next simplest case is $j = 1$ where we choose

$$|\phi_1\rangle = |1, 1\rangle, \quad |\phi_2\rangle = |1, 0\rangle, \quad |\phi_3\rangle = |1, -1\rangle$$

Evidently

$$\hat{J}^2 |1, m\rangle = j(j+1)\hbar^2 |1, m\rangle = 2\hbar^2 |1, m\rangle$$

The matrices are now 3×3

$$\hat{J}^2 = \begin{pmatrix} 2\hbar^2 & 0 & 0 \\ 0 & 2\hbar^2 & 0 \\ 0 & 0 & 2\hbar^2 \end{pmatrix} = 2\hbar^2 \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

$$\hat{J}_z = \begin{pmatrix} \hbar & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -\hbar \end{pmatrix} = \hbar \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

Acting \hat{J}_+ on $|\phi_i\rangle$ gives

$$\hat{J}_+ |1, 1\rangle = 0, \quad \hat{J}_+ |1, 0\rangle = \sqrt{2}\hbar |1, 1\rangle, \quad \hat{J}_+ |1, -1\rangle = \sqrt{2}\hbar |1, 0\rangle$$

i.e.

$$\hat{J}_+ = \begin{pmatrix} 0 & \sqrt{2}\hbar & 0 \\ 0 & 0 & \sqrt{2}\hbar \\ 0 & 0 & 0 \end{pmatrix}$$

Similarly

$$\hat{J}_- |1, 1\rangle = \sqrt{2}\hbar |1, 0\rangle, \quad \hat{J}_- |1, 0\rangle = \sqrt{2}\hbar |1, -1\rangle, \quad \hat{J}_- |1, -1\rangle = 0$$

i.e.

$$\hat{J}_- = \begin{pmatrix} 0 & 0 & 0 \\ \sqrt{2}\hbar & 0 & 0 \\ 0 & \sqrt{2}\hbar & 0 \end{pmatrix}$$

The combination of them gives

$$\hat{J}_x = \frac{1}{2}(\hat{J}_+ + \hat{J}_-) = \begin{pmatrix} 0 & \frac{\sqrt{2}}{2}\hbar & 0 \\ \frac{\sqrt{2}}{2}\hbar & 0 & \frac{\sqrt{2}}{2}\hbar \\ 0 & \frac{\sqrt{2}}{2}\hbar & 0 \end{pmatrix} = \frac{\sqrt{2}}{2}\hbar \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

$$\hat{J}_y = \frac{1}{2i}(\hat{J}_+ - \hat{J}_-) = \begin{pmatrix} 0 & -i\frac{\sqrt{2}}{2}\hbar & 0 \\ i\frac{\sqrt{2}}{2}\hbar & 0 & -i\frac{\sqrt{2}}{2}\hbar \\ 0 & i\frac{\sqrt{2}}{2}\hbar & 0 \end{pmatrix} = \frac{\sqrt{2}}{2}\hbar \begin{pmatrix} 0 & -i & 0 \\ i & 0 & -i \\ 0 & i & 0 \end{pmatrix}$$

Having obtained the matrix elements of \hat{J}_z and \hat{J}_\pm , we can study the rotation operator $D(R)$. If a rotation R is specified by $\hat{\mathbf{n}}$ and ϕ , we can define its matrix elements by

$$D_{m'm}^j(R) = \left\langle j, m' \left| \exp\left(-\frac{i\hat{\mathbf{J}} \cdot \mathbf{n}\phi}{\hbar}\right) \right| j, m \right\rangle$$

which are sometimes called *Wigner functions* after E. P. Wigner, who made pioneering contributions to the group theoretical properties of rotations in quantum mechanics. Notice here the same j -value appears in the ket and bra, which means rotations cannot change the j -value. Often in the literature, the $(2j+1) \times (2j+1)$ matrix formed by $D_{m'm}^j(R)$ is referred to as $(2j+1)$ -dimensional irreducible representation of rotation operator $D(R)$. This leads us to the group theory which is beyond the scope of this course.

II. SPIN

In classical mechanics, a rigid object admits two kinds of angular momentum: orbital ($\mathbf{L} = \mathbf{r} \times \mathbf{p}$), associated with the motion of the center of mass, and spin ($\mathbf{S} = I\omega$) associated with motion about the center of mass. For example, the earth has orbital angular momentum attributable to its annual revolution around the sun, and spin angular momentum coming from its daily rotation about the north-south axis. In the classical context this distinction is largely a matter of convenience, for when you come right down to it, \mathbf{S} is nothing but the sum total of the “orbital”

angular momenta of all the rocks and dirt clouds that go to make up the earth, as they circle around the axis. But an analogous thing happens in quantum mechanics, and here *the distinction is absolutely fundamental*. In addition to orbital angular momentum associated with the motion of the electron around the nucleus, the electron also carries another form of angular momentum, which has nothing to do with motion in space (and which is not, therefore, described by any function of the position variables r, θ, ϕ) but which is somewhat analogous to classical spin (and for which, therefore, we use the same word). But we can not press this analogy too far: The electron (as far as we know) is a structureless point particle, and its spin angular momentum cannot be decomposed into orbital angular momenta of constituent parts. Suffice it to say that elementary particles carry intrinsic angular momentum (\mathbf{S}) in addition to their “extrinsic” angular momentum (\mathbf{L}).

A. The Uhlenbeck and Goudsmit proposal

In 1925, Uhlenbeck and Goudsmit postulated the existence of a new intrinsic property of particles that behaved like an angular momentum as a means of explaining the results of fine structure of spectral lines, anomalous Zeeman effect and Stern-Gerlach experiment. In these experiments one always observe **even** number spectral lines, Zeeman sub-levels or electron beams. This can not be explained by orbital angular momentum because the quantum number l is integer and $(2l + 1)$ must be **odd**. This intrinsic property was later termed spin by Pauli, however, the image of a spinning sphere is not likely an accurate one. This new property needs to be viewed as an intrinsic property like mass and charge that is particular to a given type of particle. Note that, unlike mass and charge, there is no classical analog to spin!

It was Dirac who later showed that spin arises very naturally in a correct relativistic formulation of the quantum theory. This formulation is embodied in the relativistic generalization of the Schrödinger equation called the Dirac equation.

According to the Uhlenbeck and Goudsmit proposal, the spin of a particle should behave like an angular momentum and, therefore, should have an associated magnetic moment

$$\mathbf{M}_s = -\frac{g\mu_B}{\hbar}\hat{\mathbf{S}}$$

where $\hat{\mathbf{S}}$ is the spin angular momentum operator, the projection of which on any axis can only take two possible values

$$\hat{S}_z = \pm\frac{1}{2}\hbar.$$

g is a constant introduced to produce the best fit with experiment and $\mu_B = e\hbar/2m$ is the Bohr magneton. The interaction with a magnetic field is proportional to $-\mathbf{M}_s \cdot \mathbf{B}$, which is the basis of the NMR technique. It is found that good fits to experimental data are obtained when $g = 2$, which means that the *spin gyromagnetic ratio*, defined to be $g\mu_B/\hbar$ is twice as large as the *orbital gyromagnetic ratio* μ_B/\hbar .

B. Description of spin state and spin operators

Two principles should be kept in mind in describing the spin wavefunction and spin operator. First, spin is an intrinsic property of electron. It is not the result of some kind of spatial motion, therefore, it can not be expressed by a wavefunction of spatial coordinate and can not be expressed by operators composed by $\hat{\mathbf{r}}$ and $\hat{\mathbf{p}}$. Secondly, spin is a kind of angular momentum, and obeys all the requirements for an angular momentum. The spin operators representing spin angular momentum has three components $\hat{S}_x, \hat{S}_y, \hat{S}_z$. They obey the general commutation relations of any angular momentum, i.e.

$$\begin{aligned}\hat{S}_x\hat{S}_y - \hat{S}_y\hat{S}_x &= i\hbar\hat{S}_z \\ \hat{S}_y\hat{S}_z - \hat{S}_z\hat{S}_y &= i\hbar\hat{S}_x \\ \hat{S}_z\hat{S}_x - \hat{S}_x\hat{S}_z &= i\hbar\hat{S}_y\end{aligned}$$

One may formally write it in a compact notation

$$\hat{\mathbf{S}} \times \hat{\mathbf{S}} = i\hbar\hat{\mathbf{S}}$$

which means

$$\begin{vmatrix} \mathbf{x}_0 & \mathbf{y}_0 & \mathbf{z}_0 \\ \hat{S}_x & \hat{S}_y & \hat{S}_z \\ \hat{S}_x & \hat{S}_y & \hat{S}_z \end{vmatrix} = i\hbar\hat{\mathbf{S}}.$$

The square of each component

$$\hat{S}_x^2 = \hat{S}_y^2 = \hat{S}_z^2 = \left(\pm\frac{1}{2}\hbar\right)^2 = \frac{\hbar^2}{4}$$

leads to

$$\hat{S}^2 = \frac{3}{4}\hbar^2 = S(S+1)\hbar^2.$$

So the spin quantum number takes the value of $S = \frac{1}{2}$.

In general theory of angular momentum, we have discussed the $j = \frac{1}{2}$ case which is suitable to express the spin angular momentum. Actually this is the spin of the particles that make up ordinary matter (protons, neutrons, and electrons), as well as all quarks and all leptons. Moreover, once you understand spin 1/2, it is a simple matter to work out the formalism for any higher spin. There are just two eigenstates of \hat{J}^2 and \hat{J}_z : $\chi_{\frac{1}{2}} = |\frac{1}{2}, \frac{1}{2}\rangle$, which we call spin up (\uparrow) and $\chi_{-\frac{1}{2}} = |\frac{1}{2}, -\frac{1}{2}\rangle$, which we call spin down (\downarrow). This is a complete set of eigenstates of one electron spin and can be used as basis to span a representation - the spin representation. Using these as basis vectors, the general state of a spin-1/2 particle can be expressed as a two-element column matrix (or spinor)

$$\chi = \begin{pmatrix} a \\ b \end{pmatrix} = a\chi_{\frac{1}{2}} + b\chi_{-\frac{1}{2}}$$

with

$$\chi_{\frac{1}{2}} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \chi_{-\frac{1}{2}} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Meanwhile, the spin operators become 2×2 matrices

$$\begin{aligned} \hat{S}_x &= \frac{\hbar}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, & \hat{S}_y &= \frac{\hbar}{2} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, & \hat{S}_z &= \frac{\hbar}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \\ \hat{S}^2 &= \begin{pmatrix} \frac{3}{4}\hbar^2 & 0 \\ 0 & \frac{3}{4}\hbar^2 \end{pmatrix}, & \hat{S}_+ &= \begin{pmatrix} 0 & \hbar \\ 0 & 0 \end{pmatrix}, & \hat{S}_- &= \begin{pmatrix} 0 & 0 \\ \hbar & 0 \end{pmatrix} \end{aligned}$$

respectively. The ladder operators are defined as before $\hat{S}_{\pm} = \hat{S}_x \pm i\hat{S}_y$.

C. Pauli matrices

It is customary to write $\hat{\mathbf{S}} = \frac{\hbar}{2}\hat{\sigma}$, where $\hat{\sigma} = (\hat{\sigma}_x, \hat{\sigma}_y, \hat{\sigma}_z)$ are the famous **Pauli spin matrices**

$$\hat{\sigma}_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \hat{\sigma}_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \hat{\sigma}_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (13)$$

Pauli spin operators satisfy several important relations. First it is easy to show that

$$\begin{aligned} \hat{\sigma}_j^2 &= 1, & j &= x, y, z \\ \hat{\sigma}_j\hat{\sigma}_k + \hat{\sigma}_k\hat{\sigma}_j &= 0, & \text{for } j &\neq k \end{aligned}$$

These two relations are equivalent to the anti-commutation relation

$$\{\hat{\sigma}_j, \hat{\sigma}_k\} = 2\delta_{jk} \quad (14)$$

We also have the commutation relations

$$[\hat{\sigma}_j, \hat{\sigma}_k] = 2i\epsilon_{jkl}\hat{\sigma}_l \quad (15)$$

Combining (14) and (15) we can obtain

$$\begin{aligned} \hat{\sigma}_x \hat{\sigma}_y &= i\hat{\sigma}_z, \\ \hat{\sigma}_y \hat{\sigma}_z &= i\hat{\sigma}_x, \\ \hat{\sigma}_z \hat{\sigma}_x &= i\hat{\sigma}_y. \end{aligned} \quad (16)$$

Thus we see that, owing to $\hat{\sigma}_j^2 = 1$ (originated from the fact that $\hat{S}_z = \pm \frac{\hbar}{2}$), the spin operators satisfy more rules than other angular momentum operators. Notice also that

$$\hat{\sigma}_j^\dagger = \hat{\sigma}_j, \quad \det(\hat{\sigma}_j) = -1, \quad \text{Tr}(\hat{\sigma}_j) = 0$$

It is possible to establish the Pauli matrices anew from the operation rules. Obviously the operator $\hat{\sigma}_z$ is diagonal in its own representation and is expressed as

$$\hat{\sigma}_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

So what matrix forms will $\hat{\sigma}_x$ and $\hat{\sigma}_y$ be? Suppose $\hat{\sigma}_x$ can be expressed by

$$\hat{\sigma}_x = \begin{pmatrix} a & b \\ b^* & c \end{pmatrix}$$

because it is hermitian. From $\hat{\sigma}_x \hat{\sigma}_z + \hat{\sigma}_z \hat{\sigma}_x = 0$

$$\begin{pmatrix} a & b \\ b^* & c \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} a & b \\ b^* & c \end{pmatrix} = 0$$

one has

$$\begin{pmatrix} a & -b \\ b^* & -c \end{pmatrix} + \begin{pmatrix} a & b \\ -b^* & -c \end{pmatrix} = 0 \Rightarrow a = c = 0.$$

By means of $\hat{\sigma}_x^2 = 1$

$$\begin{pmatrix} 0 & b \\ b^* & 0 \end{pmatrix} \begin{pmatrix} 0 & b \\ b^* & 0 \end{pmatrix} = \begin{pmatrix} bb^* & 0 \\ 0 & bb^* \end{pmatrix} = 1 \Rightarrow bb^* = 1.$$

Choosing $b = b^* = 1$, we have

$$\hat{\sigma}_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$

Finally using $\hat{\sigma}_z \hat{\sigma}_x = i\hat{\sigma}_y$, we obtain

$$\hat{\sigma}_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}.$$

D. The complete wave function of an electron

For an electron with a spatial wave function $\psi(x, y, z)$ and a spin "up", the complete wave function is

$$\Psi_1 = \psi(x, y, z) \chi_{\frac{1}{2}} = \begin{pmatrix} \psi(x, y, z) \\ 0 \end{pmatrix}.$$

For an electron with a spatial wave function $\phi(x, y, z)$ and a spin "down", the complete expression of wave function is

$$\Psi_2 = \phi(x, y, z) \chi_{-\frac{1}{2}} = \begin{pmatrix} 0 \\ \phi(x, y, z) \end{pmatrix}.$$

The superposition of the two above wave functions is

$$\Psi = a\Psi_1 + b\Psi_2 = \begin{pmatrix} a\psi(x, y, z) \\ b\phi(x, y, z) \end{pmatrix}.$$

Example 1 *Measurement of spin: The eigenstates of \hat{S}_z are of course*

$$\chi_{\frac{1}{2}} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \chi_{-\frac{1}{2}} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

with eigenvalues $+\hbar/2$ and $-\hbar/2$, respectively. If you measure \hat{S}_z on a particle in the general state

$$\chi = \begin{pmatrix} a \\ b \end{pmatrix} = a\chi_{\frac{1}{2}} + b\chi_{-\frac{1}{2}},$$

you could get $+\hbar/2$, with probability $|a|^2$, or $-\hbar/2$, with probability $|b|^2$. Since these are the only possibilities, $|a|^2 + |b|^2 = 1$, i.e. the spinor must be normalized. But what if, instead, you chose to measure \hat{S}_x ? What are the possible results, and what are their respective probabilities? According to the generalized statistical interpretation, we need to know the eigenvalues and eigenspinors of \hat{S}_x . The characteristic equation is

$$\begin{vmatrix} -\lambda & \hbar/2 \\ \hbar/2 & -\lambda \end{vmatrix} = 0 \Rightarrow \lambda^2 = (\hbar/2)^2 \Rightarrow \lambda = \pm\hbar/2$$

Not surprisingly, the possible values for \hat{S}_x are the same as those for \hat{S}_z . The eigenspinors are obtained in the usual way:

$$\frac{\hbar}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \pm \frac{\hbar}{2} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \Rightarrow \begin{pmatrix} \beta \\ \alpha \end{pmatrix} = \pm \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$$

so $\beta = \pm\alpha$. Evidently the (normalized) eigenspinors of \hat{S}_x are

$$\chi_{\frac{1}{2}}^{(x)} = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}, \text{ with eigenvalue } +\frac{\hbar}{2}$$

$$\chi_{-\frac{1}{2}}^{(x)} = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix}, \text{ with eigenvalue } -\frac{\hbar}{2}$$

As the eigenvectors of a Hermitian matrix, they span the space; the generic spinor χ can be expressed as a linear combination of them:

$$\chi = \begin{pmatrix} a \\ b \end{pmatrix} = \frac{a+b}{\sqrt{2}}\chi_{\frac{1}{2}}^{(x)} + \frac{a-b}{\sqrt{2}}\chi_{-\frac{1}{2}}^{(x)}$$

If you measure \hat{S}_x , the probability of getting $+\hbar/2$ is $|a+b|^2/2$ and the probability of getting $-\hbar/2$ is $|a-b|^2/2$. It is easy to check that these probabilities add up to 1. Suppose a spin 1/2 particle is in the state

$$\chi = \frac{1}{\sqrt{6}} \begin{pmatrix} 1+i \\ 2 \end{pmatrix}$$

If you measure \hat{S}_z , the probability of getting $+\hbar/2$ is $|(1+i)/\sqrt{6}|^2 = 1/3$, and the probability of getting $-\hbar/2$ is $|2/\sqrt{6}|^2 = 2/3$. If you measure \hat{S}_x , the probability of getting $+\hbar/2$ is $|(3+i)/\sqrt{6}|^2/2 = 5/6$, and the probability of getting $-\hbar/2$ is $|(-1+i)/\sqrt{6}|^2/2 = 1/6$. Evidently the expectation value of \hat{S}_x is

$$\frac{5}{6} \left(+\frac{\hbar}{2} \right) + \frac{1}{6} \left(-\frac{\hbar}{2} \right) = \frac{\hbar}{3}$$

which we could also have obtained more directly:

$$\begin{aligned} \langle S_x \rangle &= \chi^\dagger \hat{S}_x \chi = \begin{pmatrix} \frac{1-i}{\sqrt{6}} & \frac{2}{\sqrt{6}} \end{pmatrix} \frac{\hbar}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \frac{1+i}{\sqrt{6}} \\ \frac{2}{\sqrt{6}} \end{pmatrix} \\ &= \frac{\hbar}{2} \begin{pmatrix} \frac{2}{\sqrt{6}} & \frac{1-i}{\sqrt{6}} \end{pmatrix} \begin{pmatrix} \frac{1+i}{\sqrt{6}} \\ \frac{2}{\sqrt{6}} \end{pmatrix} = \frac{\hbar}{3} \end{aligned}$$

Problem 2 Zeng Jinyan, *Textbook*, p238-239, 1,2,3,4

Problem 3 Show that the Pauli spin matrices satisfy

$$\hat{\sigma}_j \hat{\sigma}_k = \delta_{jk} + i \sum_l \epsilon_{jkl} \hat{\sigma}_l$$

III. ADDITION OF OF ANGULAR MOMENTA

It often happens that there are more than one angular momenta in one system. There may be two possibilities: On the one hand, system contains two particles, each has its own angular momentum. On the other hand, particle may possess more than one angular momenta, for example orbital angular momentum and spin angular momentum.

One essential question is: How these angular momenta behave themselves in a system?

These angular momenta keep their independency. Each angular momentum has its own eigenvalues (\hat{J}^2, \hat{J}_z) and its own eigenstates $(|j, m\rangle)$. The angular momenta do not interfere each other. In this case, these angular momenta are uncoupled.

But there do exist interactions, slight or strong, between (among) these angular momenta. Individual angular momentum thus loses its independency. They add together to form a total angular momentum. In this case, these angular momenta are coupled.

We are interested in the relation between the uncoupled angular momenta and coupled angular momenta. Given two angular momenta, what would the resulted total angular momentum be? From the viewpoint of representation, how would a unitary transformation change the states from uncoupled representation to coupled representation?

A. Complete set of commutation operators in two representations

Consider two angular-momentum operators \mathbf{J}_1 and \mathbf{J}_2 in different subspaces. The components of \mathbf{J}_1 and \mathbf{J}_2 satisfy the usual angular momentum commutation relations

$$\begin{aligned} [\hat{J}_{1j}, \hat{J}_{1k}] &= i\hbar \epsilon_{jkl} \hat{J}_{1l} \\ [\hat{J}_{2j}, \hat{J}_{2k}] &= i\hbar \epsilon_{jkl} \hat{J}_{2l} \end{aligned}$$

However, we have

$$[\hat{J}_{1k}, \hat{J}_{2l}] = 0$$

between any pair of operators from different subspaces. The infinitesimal rotation operator that affects both subspace 1 and subspace 2 is written as

$$\left(1 - \frac{i\hat{\mathbf{J}}_1 \cdot \mathbf{n}\delta\phi}{\hbar}\right) \otimes \left(1 - \frac{i\hat{\mathbf{J}}_2 \cdot \mathbf{n}\delta\phi}{\hbar}\right) = 1 - \frac{i(\hat{\mathbf{J}}_1 \otimes 1 + 1 \otimes \hat{\mathbf{J}}_2) \cdot \mathbf{n}\delta\phi}{\hbar}$$

We denote the total angular momentum by

$$\hat{\mathbf{J}} = \hat{\mathbf{J}}_1 \otimes 1 + 1 \otimes \hat{\mathbf{J}}_2$$

which is more commonly written as

$$\hat{\mathbf{J}} = \hat{\mathbf{J}}_1 + \hat{\mathbf{J}}_2$$

As for the choice of the base kets, we have two options

- Uncoupled representation: Common eigenstates of commutative operators $\hat{J}_1^2, \hat{J}_{1z}, \hat{J}_2^2, \hat{J}_{2z}$ denoted by

$$|j_1, j_2; m_1, m_2\rangle = |j_1, m_1\rangle |j_2, m_2\rangle$$

For given \hat{J}_1, \hat{J}_2 , the $(2j_1 + 1)(2j_2 + 1)$ independent basis eigenstates $|j_1, m_1\rangle |j_2, m_2\rangle$ constitute the uncoupled representation in $(2j_1 + 1)(2j_2 + 1)$ dimensional spaces.

- Coupled representation: Common eigenstates of commutative operators $\hat{J}_1^2, \hat{J}_2^2, \hat{J}^2, \hat{J}_z$ denoted by

$$|j_1, j_2; j, m\rangle \text{ or } |j, m\rangle .$$

It is necessary to show that these operators commute each other. From the definition of the total angular momentum, one obviously has

$$\begin{aligned}\hat{J}_x &= \hat{J}_{1x} + \hat{J}_{2x} \\ \hat{J}_y &= \hat{J}_{1y} + \hat{J}_{2y} \\ \hat{J}_z &= \hat{J}_{1z} + \hat{J}_{2z}\end{aligned}$$

By writing \hat{J}^2 as

$$\hat{J}^2 = (\hat{\mathbf{J}}_1 + \hat{\mathbf{J}}_2)^2 = \hat{J}_1^2 + \hat{J}_2^2 + 2(\hat{J}_{1x}\hat{J}_{2x} + \hat{J}_{1y}\hat{J}_{2y} + \hat{J}_{1z}\hat{J}_{2z})$$

we readily see

$$[\hat{J}^2, \hat{J}_1^2] = 0.$$

Moreover, it is very important to note that even though

$$\begin{aligned}[\hat{J}^2, \hat{J}_{1z}] &= -2i\hbar\hat{J}_{1y}\hat{J}_{2x} + 2i\hbar\hat{J}_{1x}\hat{J}_{2y} \\ [\hat{J}^2, \hat{J}_{2z}] &= 2i\hbar\hat{J}_{1y}\hat{J}_{2x} - 2i\hbar\hat{J}_{1x}\hat{J}_{2y}\end{aligned}$$

we have

$$[\hat{J}^2, \hat{J}_z] = 0$$

that is why we can choose \hat{J}_z in our complete set of commutative operators in coupled representation, but not \hat{J}_{1z} and \hat{J}_{2z} . Likewise, we can not add \hat{J}^2 to the set of uncoupled representation.

B. Transformation between coupled state and uncoupled state

We now consider the unitary transformation that connects the two representations

$$|j, m\rangle = \sum_{m_1, m_2} |j_1, j_2; m_1, m_2\rangle \langle j_1, j_2; m_1, m_2 | j, m\rangle$$

The elements of this transformation matrix $\langle j_1, j_2; m_1, m_2 | j, m\rangle$ are known as **Clebsch-Gordan(CG) coefficients**. There are many important properties of CG coefficients. First, the coefficients vanish unless

$$m = m_1 + m_2.$$

To prove this, first note that

$$(\hat{J}_z - \hat{J}_{1z} - \hat{J}_{2z}) |j, m\rangle = 0$$

Multiplying $\langle j_1, j_2; m_1, m_2 |$ on the left, we have

$$(m - m_1 - m_2) \langle j_1, j_2; m_1, m_2 | j, m\rangle = 0$$

Second, the coefficients vanish unless

$$|j_1 - j_2| \leq j \leq j_1 + j_2$$

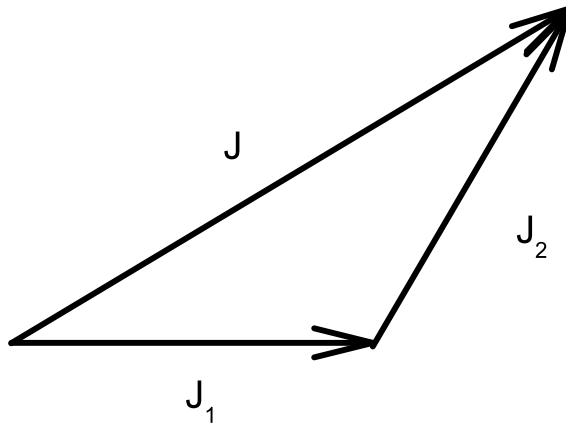


FIG. 2: Triangle rule for addition of angular momenta.

Indeed, given j_1, j_2 we have

$$\begin{aligned} m_{1 \max} &= j_1, & m_{2 \max} &= j_2, \\ m_{\max} &= j_1 + j_2 \Rightarrow j_{\max} &= j_1 + j_2 \end{aligned}$$

But what is j_{\min} ? We could obtain it from the fact that the dimensionality of space should be invariant under the transformation of representations. In the uncoupled representation, we have $(2j_1 + 1)(2j_2 + 1)$ independent basis eigenstates while the dimensionality of the coupled representation is

$$\begin{aligned} \sum_{j=j_{\min}}^{j_{\max}} (2j + 1) &= \frac{1}{2} ((2j_1 + 2j_2 + 1) + (2j_{\min} + 1)) (j_1 + j_2 - j_{\min} + 1) \\ &= (j_1 + j_2 + j_{\min} + 1) (j_1 + j_2 - j_{\min} + 1) \\ &= (2j_1 + 1) (2j_2 + 1) \end{aligned}$$

It follows that

$$\begin{aligned} 4j_1 j_2 + 2j_1 + 2j_2 + 1 &= (j_1 + j_2 + 1)^2 - j_{\min}^2 \\ 0 &= j_1^2 + j_2^2 - 2j_1 j_2 - j_{\min}^2 \\ j_{\min}^2 &= (j_1 - j_2)^2 \end{aligned}$$

i.e.

$$j_{\min} = |j_1 - j_2|$$

so j takes values

$$j = j_1 + j_2, j_1 + j_2 - 1, \dots, |j_1 - j_2| \quad (17)$$

This property appears more clear from the vector model of angular momentum addition. In Figure 2 we visualize \mathbf{J} to be the vectorial sum of \mathbf{J}_1 and \mathbf{J}_2 . Equation (17) means that one side of a triangle is neither larger than the sum of the other two sides, nor smaller than the subtract of them.

CG coefficients are nothing but the unitary transformation matrix elements between two representations, uncoupled and coupled. The third property of CG coefficients is that they form a unitary matrix and by Condon-Shortley convention the matrix elements are taken to be real. An immediate consequence of this is that the inverse coefficient is the same as itself

$$\langle j_1, j_2; m_1, m_2 | j, m \rangle = \langle j, m | j_1, j_2; m_1, m_2 \rangle$$

A real unitary matrix is orthogonal, so we have the orthogonality condition

$$\sum_{jm} \langle j_1, j_2; m_1, m_2 | j, m \rangle \langle j_1, j_2; m'_1, m'_2 | j, m \rangle = \delta_{m_1 m'_1} \delta_{m_2 m'_2}$$

$$\sum_{m_1 m_2} \langle j_1, j_2; m_1, m_2 | j, m \rangle \langle j_1, j_2; m_1, m_2 | j', m' \rangle = \delta_{jj'} \delta_{mm'}$$

Different authors used different notations for this coefficients which we list in the table

Condon-Shortley	$\langle j_1 j_2 m_1 m_2 j m \rangle$
Edmonds	$\langle j_1 j_2 m_1 m_2 j_1 j_2 j m \rangle$
Rose	$C(j_1 j_2 j, m_1 m_2 m)$
Wigner 3-j symbol	$\begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix}$

Related to Condon-Shortley through

$$\begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix} = (-1)^{j_1 - j_2 - m_3} (2j_3 + 1)^{-\frac{1}{2}} \langle j_1 j_2 m_1 m_2 | j_3, -m_3 \rangle,$$

Wigner's 3-j symbol is a more symmetric representation

$$\begin{aligned} \begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix} &= \begin{pmatrix} j_2 & j_3 & j_1 \\ m_2 & m_3 & m_1 \end{pmatrix} = \begin{pmatrix} j_3 & j_1 & j_2 \\ m_3 & m_1 & m_2 \end{pmatrix} \\ &= \begin{pmatrix} j_2 & j_1 & j_3 \\ m_2 & m_1 & m_3 \end{pmatrix} (-1)^{(j_1 + j_2 + j_3)} \\ &= \begin{pmatrix} j_1 & j_2 & j_3 \\ -m_1 & -m_2 & -m_3 \end{pmatrix} (-1)^{(j_1 + j_2 + j_3)} \end{aligned}$$

Example 4 We show by an example how to calculate the CG coefficients. Let $\hat{\mathbf{J}}_1$ be an orbital angular momentum with quantum number $l = 2$ and $\hat{\mathbf{J}}_2$ be a spin angular momentum with quantum number $s = \frac{1}{2}$. The eigenstates of uncoupled representation are

$$\begin{aligned} |j_1, m_1\rangle &: Y_{22}, Y_{21}, Y_{20}, Y_{2,-1}, Y_{2,-2} \\ |j_2, m_2\rangle &: \chi_+ = \left| \frac{1}{2}, \frac{1}{2} \right\rangle, \quad \chi_- = \left| \frac{1}{2}, -\frac{1}{2} \right\rangle \end{aligned}$$

Obviously one has

$$m_{\max} = \frac{5}{2}, \quad j_{\max} = \frac{5}{2}$$

The state in coupled representation is

$$\left| \frac{5}{2}, \frac{5}{2} \right\rangle = Y_{22} \chi_+ \tag{18}$$

from which we apply the operator \hat{J}_- to derive other states with the same j_{\max} . Note that

$$\hat{J}_- = \hat{J}_{1-} + \hat{J}_{2-} = \hat{L}_- + \hat{S}_-$$

For each side of equation (18), by means of

$$\hat{J}_- |j, m\rangle = \sqrt{(j+m)(j-m+1)} \hbar |j, m-1\rangle,$$

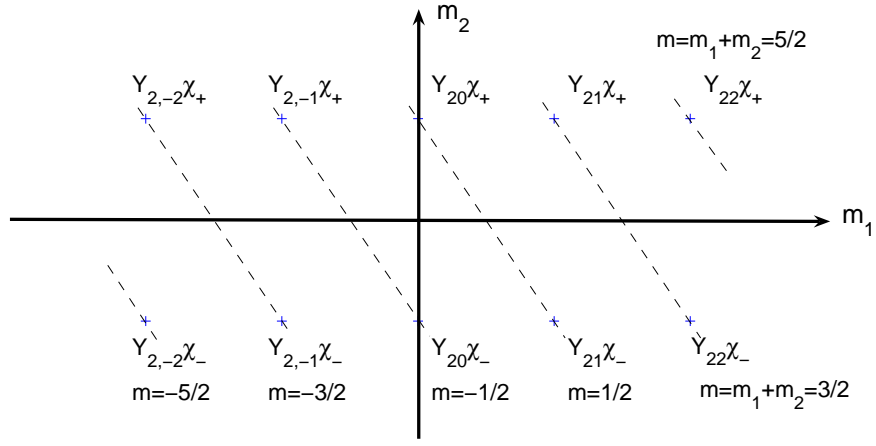


FIG. 3: Example for calculating the CG coefficients.

and

$$\begin{aligned}\hat{L}_- Y_{l,m_l} &= \sqrt{(l+m_l)(l-m_l+1)} \hbar Y_{l,m_l-1} \\ \hat{S}_- \chi_+ &= \hbar \chi_-, \hat{S}_- \chi_- = 0\end{aligned}$$

we obtain

$$\begin{aligned}\hat{J}_- \left| \frac{5}{2}, \frac{5}{2} \right\rangle &= \sqrt{5} \hbar \left| \frac{5}{2}, \frac{3}{2} \right\rangle \\ (\hat{L}_- + \hat{S}_-) Y_{22} \chi_+ &= 2\hbar Y_{21} \chi_+ + \hbar Y_{22} \chi_-\end{aligned}$$

This leads to immediately

$$\left| \frac{5}{2}, \frac{3}{2} \right\rangle = \sqrt{\frac{4}{5}} Y_{21} \chi_+ + \sqrt{\frac{1}{5}} Y_{22} \chi_- \quad (19)$$

Applying the lowering operator once more we get

$$\hat{J}_- \left| \frac{5}{2}, \frac{3}{2} \right\rangle = (\hat{L}_- + \hat{S}_-) \left(\sqrt{\frac{4}{5}} Y_{21} \chi_+ + \sqrt{\frac{1}{5}} Y_{22} \chi_- \right)$$

and

$$\begin{aligned}\left| \frac{5}{2}, \frac{1}{2} \right\rangle &= \sqrt{\frac{3}{5}} Y_{20} \chi_+ + \sqrt{\frac{2}{5}} Y_{21} \chi_- \\ \left| \frac{5}{2}, -\frac{1}{2} \right\rangle &= \sqrt{\frac{2}{5}} Y_{2,-1} \chi_+ + \sqrt{\frac{3}{5}} Y_{20} \chi_- \\ \left| \frac{5}{2}, -\frac{3}{2} \right\rangle &= \sqrt{\frac{1}{5}} Y_{2,-2} \chi_+ + \sqrt{\frac{4}{5}} Y_{2,-1} \chi_- \\ \left| \frac{5}{2}, -\frac{5}{2} \right\rangle &= Y_{2,-2} \chi_-\end{aligned}$$

The relation of these states can be understood from Figure 3. We also see there are two uncoupled states $Y_{21} \chi_+$, $Y_{22} \chi_-$ associated with $m = m_1 + m_2 = \frac{3}{2}$. They can add up to give two independent orthogonal states. Of them, one is the state (19), while the other one belongs to $j = \frac{3}{2}$

$$\left| \frac{3}{2}, \frac{3}{2} \right\rangle = -\sqrt{\frac{1}{5}} Y_{21} \chi_+ + \sqrt{\frac{4}{5}} Y_{22} \chi_-$$

The other $j = \frac{3}{2}$ states are

$$\begin{aligned} \left| \frac{3}{2}, \frac{1}{2} \right\rangle &= -\sqrt{\frac{2}{5}} Y_{20} \chi_+ + \sqrt{\frac{3}{5}} Y_{21} \chi_- \\ \left| \frac{3}{2}, -\frac{1}{2} \right\rangle &= -\sqrt{\frac{3}{5}} Y_{2,-1} \chi_+ + \sqrt{\frac{2}{5}} Y_{20} \chi_- \\ \left| \frac{3}{2}, -\frac{3}{2} \right\rangle &= -\sqrt{\frac{4}{5}} Y_{2,-2} \chi_+ + \sqrt{\frac{1}{5}} Y_{2,-1} \chi_- \end{aligned}$$

These 10 states (6 of them belong to $j = \frac{5}{2}$ and 4 of them belongs to $j = \frac{3}{2}$) are all possible ones in the coupled representation because j can only take values

$$(j_1 + j_2) = 2 + \frac{1}{2} = \frac{5}{2}$$

and

$$(j_1 - j_2) = 2 - \frac{1}{2} = \frac{3}{2}.$$

In summary, in the uncoupled representation we have $(2j_1 + 1)(2j_2 + 1) = 5 \times 2 = 10$ base states, while the dimensionality in coupled representation is the same

$$\begin{aligned} j_A = (j_1 + j_2) &= \frac{5}{2}, & j_B = (j_1 - j_2) &= \frac{3}{2} \\ N = (2j_A + 1) + (2j_B + 1) &= 10 \end{aligned}$$

If, on the other hand, our system is composed by two angular momentum $j_1 = 4$ and $j_2 = 1$, it can be seen that the dimensionality remains the same after the unitary transformation. There are three sets in this case with $j_{A,B,C} = 5, 4, 3$ and the number of bases in uncoupled representation is

$$(2j_1 + 1)(2j_2 + 1) = 9 \times 3 = 27$$

while that for coupled one is

$$(2j_A + 1) + (2j_B + 1) + (2j_C + 1) = 11 + 9 + 7 = 27$$

Problem 5 Zeng Jinyan, Textbook, p239-240, 7

IV. FINE STRUCTURE OF ALKALI SPECTRA

In our study of the hydrogen atom, we took the Hamiltonian to be

$$H = -\frac{\hbar^2}{2m} \nabla^2 - \frac{e^2}{4\pi\epsilon_0 r}$$

(electron kinetic energy plus Coulombic potential energy). But this is not quite the whole story. We have already learned how to correct for the motion of the nucleus: Just replace m by the reduced mass. More significant is the so-called *fine structure*, which is actually due to two distinct mechanisms: a relativistic correction, and spin-orbit coupling. Compared to the Bohr energies, fine structure is a tiny perturbation - smaller by a factor of α^2 , where

$$\alpha = \frac{1}{4\pi\epsilon_0} \frac{e^2}{\hbar c} \simeq \frac{1}{137.036}$$

is the famous fine structure constant. Smaller still (by another factor of α) is the *Lamb shift*, associated with the quantization of the Coulomb field, and smaller by yet another order of magnitude is the *hyperfine structure*, which is due to the magnetic interaction between the dipole moments of the electron and the proton. This hierarchy is summarized in the following table.

Bohr energies:	of order $\alpha^2 mc^2$
Fine structure:	of order $\alpha^4 mc^2$
Lamb shift:	of order $\alpha^5 mc^2$
Hyperfine splitting:	of order $(m/m_p)\alpha^4 mc^2$

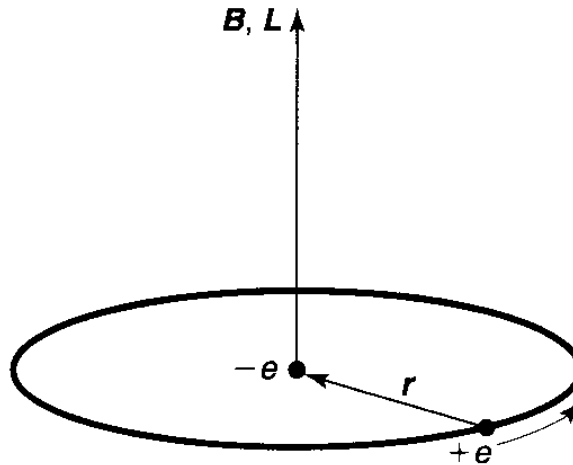


FIG. 4: Hydrogen atom, from the electron's perspective.

A. Spin-orbit coupling

We study the atomic levels of hydrogen-like atoms, that is, atoms with one valence electron outside the closed shell. Alkali atoms such as sodium (Na) and potassium (K) belong to this category. The central potential $V(r)$ appropriate for the valence electron is no longer of the pure Coulomb form. This is because the electrostatic potential $\phi(r)$ that appears in $V(r) = e\phi(r)$, is no longer due just to the nucleus of electric charge $|e|Z$; we must take into account the cloud of negatively charged electrons in the inner shells. The degeneracy characteristic of the pure Coulomb potential are now removed in such a way that the higher l states lie higher for a given n .

Now consider a classical model of spin-orbit coupling as electron moves by a closed orbit around nucleus. The nucleus may also be viewed circulating around electron by a reverse orbit. The circulating nucleus generates a magnetic field B at its center, which exerts a torque on the spinning electron, tending to align its magnetic moment along the direction of the field. If we picture the nucleus from the electron's perspective as a continuous current loop (Figure 4), its magnetic field can be calculated from the Biot-Savart law

$$B = \frac{\mu_0}{4\pi} \frac{2\pi I}{r}$$

with an effective current

$$I = \frac{Zev}{2\pi r}.$$

The magnetic field is

$$B = \frac{1}{4\pi\epsilon_0} \frac{Zev}{c^2 r^2}$$

where we have used $\mu_0\epsilon_0 = 1/c^2$. In order to show the direction of magnetic field, the magnetic field has been written in the vectorial form

$$\mathbf{B} = \frac{1}{4\pi\epsilon_0} \frac{Ze}{c^2 r^2} \left(\frac{\mathbf{r}}{r} \times \mathbf{v} \right)$$

in which $\mathbf{r} \times \mathbf{v}$ is composed of the orbit radius and velocity of nucleus, but apart from the mass of the electron, it is exactly the same as the orbital angular momentum of the electron $\mathbf{L} = \mathbf{r} \times \mathbf{p}$. So

$$\mathbf{B} = \frac{1}{4\pi\epsilon_0} \frac{Ze}{mc^2 r^3} \mathbf{L}. \quad (20)$$

The interaction energy of magnetic field and electron spin magnetic moment is

$$\begin{aligned}\hat{H}' &= -\mathbf{M}_s \cdot \mathbf{B} = -\left(-\frac{g\mu_B}{\hbar}\hat{\mathbf{S}}\right) \cdot \left(\frac{1}{4\pi\epsilon_0}\frac{Ze}{mc^2r^3}\hat{\mathbf{L}}\right) \\ &= -\left(-\frac{e}{m}\hat{\mathbf{S}}\right) \cdot \left(\frac{1}{4\pi\epsilon_0}\frac{Ze}{mc^2r^3}\hat{\mathbf{L}}\right) = \frac{1}{4\pi\epsilon_0}\frac{Ze^2}{m^2c^2r^3}(\hat{\mathbf{L}} \cdot \hat{\mathbf{S}}) \\ &= \frac{1}{m^2c^2}\frac{1}{r}\frac{d}{dr}\left(-\frac{1}{4\pi\epsilon_0}\frac{Ze^2}{r}\right)(\hat{\mathbf{L}} \cdot \hat{\mathbf{S}}) = \frac{1}{m^2c^2}\frac{1}{r}\frac{dV}{dr}(\hat{\mathbf{L}} \cdot \hat{\mathbf{S}})\end{aligned}$$

But there is a serious fraud in this calculation: I did the analysis in the rest frame of the electron, but that's not an inertial system – it accelerates as the electron orbits around the nucleus. You can get away with this if you make an appropriate kinematic correction, known as the **Thomas precession** (after L. H. Thomas). In this context it throws in a factor of 1/2

$$\hat{H}' = \frac{1}{2m^2c^2}\frac{1}{r}\frac{dV}{dr}\hat{\mathbf{L}} \cdot \hat{\mathbf{S}} = \xi(r)\hat{\mathbf{L}} \cdot \hat{\mathbf{S}}$$

This is the spin-orbit interaction; apart from two corrections (the modified gyromagnetic ratio for the electron and the Thomas precession factor – which, coincidentally, exactly cancel one another), it is just what you would expect on the basis of a naive classical model. Physically, it is attributable to the torque exerted on the magnetic dipole moment of the spinning electron, by the magnetic field of the proton, in the electron's instantaneous rest frame. The correct quantum-mechanical explanation of this must await the Dirac's relativistic theory of the electron.

B. Perturbation calculation of spin-orbit coupling energy

The Hamiltonian of the external electron of alkali atom is \hat{H}_0

$$\hat{H}_0 = \frac{1}{2m}\hat{p}^2 + V(r)$$

The eigenfunctions and eigenvalues satisfying the Schrödinger equation

$$\hat{H}_0\psi^{(0)} = E^{(0)}\psi^{(0)}$$

have following features

$$\begin{aligned}\psi^{(0)} &= \psi_{nlm_l}^{(0)} = R_{nl}Y_{lm_l} \\ E^{(0)} &= E_{nl}^{(0)}\end{aligned}$$

Note that the eigenvalues differ from hydrogen atom case in that they also depend on angular quantum number l . We treat the spin orbit coupling as perturbation

$$\hat{H}' = \xi(r)\hat{\mathbf{L}} \cdot \hat{\mathbf{S}}$$

This arises the following questions:

- In considering the zeroth-order wave function, spin freedom of degree must be included.
- The zeroth-order state is degenerate

$$E_{nl}^{(0)} \quad R_{nl}(r)Y_{lm_l}(\theta, \varphi) \quad (2l+1)\text{-fold degenerate}$$

When spin is included, the degeneracy becomes $2(2l+1)$ fold. Thus we are dealing with a degenerate perturbation problem.

- Which of the two representations (uncoupled or coupled) shall we choose as the basis for perturbation? The wavefunctions in uncoupled representation are $R_{nl}Y_{lm_l}\chi_{\pm}$ with $m_l = l, l-1, \dots, -l$ and $\chi = \chi_+, \chi_-$. In the presence of spin-orbit coupling, the Hamiltonian no longer commutes with $\hat{\mathbf{L}}$ and $\hat{\mathbf{S}}$, so the spin and orbital angular momenta are not separately conserved. However, \hat{H}' does commute with \hat{L}^2 , \hat{S}^2 , and the total angular

momentum $\hat{\mathbf{J}} = \hat{\mathbf{L}} + \hat{\mathbf{S}}$. Hence these quantities are conserved. To put it another way, the eigenstates of \hat{L}_z and \hat{S}_z are not “good” states to use in perturbation theory, but the eigenstates of \hat{L}^2 , \hat{S}^2 , \hat{J}^2 and \hat{J}_z are. That means we need to treat the perturbation in the *coupled representation*. The quantum number for total angular momentum takes values $j = l \pm \frac{1}{2}$

$$j_1 = l + \frac{1}{2} : (2j_1 + 1) = 2l + 2$$

$$j_2 = l - \frac{1}{2} : (2j_2 + 1) = 2l$$

with $2(2l + 1)$ degeneracy. The zeroth order wavefunctions for $l = 2, s = \frac{1}{2}$ has already been calculated in the previous section, for example for $j_1 = \frac{5}{2}$

$$\left| \frac{5}{2}, \frac{3}{2} \right\rangle = \sqrt{\frac{4}{5}} Y_{21} \chi_+ + \sqrt{\frac{1}{5}} Y_{22} \chi_-, \text{ etc}$$

Now we have to include the radial wavefunction as well

$$\begin{aligned} |n, l; j, m_j\rangle &= \left| n, 2; \frac{5}{2}, \frac{3}{2} \right\rangle = R_{n2}(r) \sqrt{\frac{4}{5}} Y_{21} \chi_+ + R_{n2}(r) \sqrt{\frac{1}{5}} Y_{22} \chi_- \\ &= \sqrt{\frac{4}{5}} \begin{pmatrix} R_{n2} Y_{21} \\ 0 \end{pmatrix} + \sqrt{\frac{1}{5}} \begin{pmatrix} 0 \\ R_{n2} Y_{22} \end{pmatrix} \\ &= \begin{pmatrix} \sqrt{\frac{4}{5}} R_{n2} Y_{21} \\ \sqrt{\frac{1}{5}} R_{n2} Y_{22} \end{pmatrix} = R_{n2}(r) \left| \frac{5}{2}, \frac{3}{2} \right\rangle \end{aligned}$$

Similarly for $j = \frac{3}{2}$

$$\left| \frac{3}{2}, \frac{3}{2} \right\rangle = -\sqrt{\frac{1}{5}} Y_{21} \chi_+ + \sqrt{\frac{4}{5}} Y_{22} \chi_-$$

and

$$|n, l; j, m_j\rangle = \left| n, 2; \frac{3}{2}, \frac{3}{2} \right\rangle = R_{n2}(r) \left| \frac{3}{2}, \frac{3}{2} \right\rangle$$

C. The perturbation energy

Choosing $|n, l; j, m_j\rangle = R_{n2}(r) |j, m_j\rangle$ as the basis we may readily calculate the energy shift due to the perturbation. Now

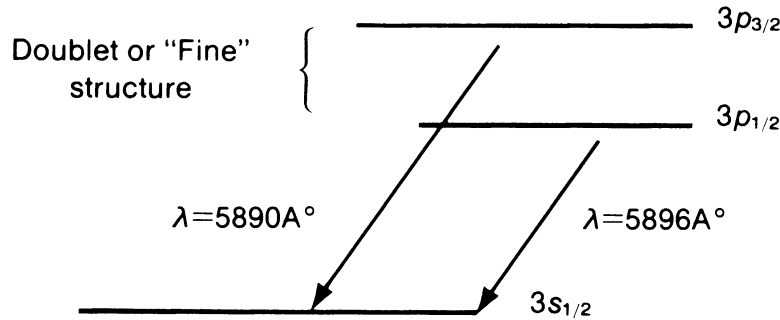
$$\hat{\mathbf{J}}^2 = (\hat{\mathbf{L}} + \hat{\mathbf{S}})^2 = \hat{\mathbf{L}}^2 + \hat{\mathbf{S}}^2 + 2\hat{\mathbf{L}} \cdot \hat{\mathbf{S}}$$

so

$$\hat{\mathbf{L}} \cdot \hat{\mathbf{S}} = \frac{1}{2} (\hat{\mathbf{J}}^2 - \hat{\mathbf{L}}^2 - \hat{\mathbf{S}}^2)$$

In degenerate perturbation theory, if the perturbation is already diagonal in the representation we are using, all we need to do for the first-order energy shift is to take the expectation value

$$\begin{aligned} \langle \hat{H}' \rangle &= \int_0^\infty R_{nl}^* \langle j, m_j | \xi(r) (\hat{\mathbf{L}} \cdot \hat{\mathbf{S}}) R_{nl} | j, m_j \rangle r^2 dr \\ &= \int_0^\infty R_{nl}^* \xi(r) R_{nl} r^2 dr \langle j, m_j | (\hat{\mathbf{L}} \cdot \hat{\mathbf{S}}) | j, m_j \rangle \end{aligned}$$

FIG. 5: Schematic diagram of 3s and 3p and sodium D lines.

where the integral

$$\int_0^\infty R_{nl}^* \xi(r) R_{nl} r^2 dr = c$$

is the same for all elements. And the matrix elements for $\hat{\mathbf{L}} \cdot \hat{\mathbf{S}}$ are

$$\langle j, m_j | (\hat{\mathbf{L}} \cdot \hat{\mathbf{S}}) | j, m_j \rangle = \frac{\hbar^2}{2} (j(j+1) - l(l+1) - 3/4)$$

We conclude that the energy shifts due to spin-orbit interaction are

$$\Delta E = \begin{cases} \frac{c\hbar^2}{2} \left((l + \frac{1}{2})(l + \frac{3}{2}) - l(l+1) - \frac{3}{4} \right) = \frac{c\hbar^2}{2} l & \text{for all } j = l + \frac{1}{2} \text{ states} \\ \frac{c\hbar^2}{2} \left((l - \frac{1}{2})(l + \frac{1}{2}) - l(l+1) - \frac{3}{4} \right) = -\frac{c\hbar^2}{2} (l+1) & \text{for all } j = l - \frac{1}{2} \text{ states} \end{cases}$$

Fine structure breaks the degeneracy in l (that is, for a given n , the different allowed values of l do not all carry the same energy); the energies are determined by n and j . Specifically, consider a sodium atom with ground state configuration $1s^2 2s^2 2p^6 3s^1$. We are interested in the excitation of the valence electron from $3s$ to the nearest level $3p$. Because the central potential is no longer of the pure Coulomb form, $3s$ and $3p$ are now split. The spin-orbit coupling splits the $3p$ level further into two: $3p_{3/2}$ and $3p_{1/2}$ where the subscript stands for $j = l \pm 1/2$. Experimentally, we observe two closely separated yellow lines - known as the sodium D lines - one at 5890\AA , the other at 5896\AA (Figure 5). Here the degeneracy is only partially removed, i.e., the degeneracy for $3p_{3/2}$ and $3p_{1/2}$ are $2 \times 3/2 + 1 = 4$ and $2 \times 1/2 + 1 = 2$, respectively.

V. CHARGED PARTICLE IN ELECTROMAGNETIC FIELD - ZEEMAN EFFECT

A. Hamiltonian

Till now, the interaction we have discussed depends only on the position of the particle and is given by a scalar potential. A quasi-static electromagnetic field described by an electrostatic potential ϕ and a vectorial potential \mathbf{A}

$$\begin{aligned} \mathbf{B} &= \nabla \times \mathbf{A} \\ \mathbf{E} &= -\frac{\partial \mathbf{A}}{\partial t} - \nabla \phi \quad (\text{Gaussian, } \mathbf{E} = -\frac{1}{c} \frac{\partial \mathbf{A}}{\partial t} - \nabla \phi) \end{aligned}$$

interacts with the charged particle with Hamiltonian (equation 55.4 in Classical Mechanics, 2nd edition, by Corben and Stehle)

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

In quantum mechanics, we replace the canonical momentum \mathbf{p} by operator $\hat{\mathbf{p}} = -i\hbar\nabla$, the potential $q\phi$ by V and obtain

$$\begin{aligned}\hat{H} &= \frac{1}{2m}(\hat{\mathbf{p}} - q\mathbf{A})^2 + V \\ &= \frac{1}{2m}(\hat{\mathbf{p}} - q\mathbf{A}) \cdot (\hat{\mathbf{p}} - q\mathbf{A}) + V \\ &= \frac{1}{2m}\hat{\mathbf{p}}^2 + \frac{1}{2m}q^2\mathbf{A}^2 - \frac{1}{2m}q\mathbf{A} \cdot \hat{\mathbf{p}} - \underbrace{\frac{1}{2m}q\hat{\mathbf{p}} \cdot \mathbf{A}} + V.\end{aligned}$$

It follows from

$$\frac{1}{2m}q\hat{\mathbf{p}} \cdot \mathbf{A} = \frac{1}{2m}q(-i\hbar\nabla) \cdot \mathbf{A} = -\frac{i\hbar}{2m}q(\nabla \cdot \mathbf{A}) + \frac{1}{2m}q\mathbf{A} \cdot \hat{\mathbf{p}},$$

and the coulomb gauge

$$\nabla \cdot \mathbf{A} = 0$$

that

$$\hat{H} = \frac{1}{2m}\hat{\mathbf{p}}^2 - \frac{q}{m}\mathbf{A} \cdot \hat{\mathbf{p}} + V + \frac{1}{2m}q^2\mathbf{A}^2.$$

If Gaussian System is used, the final result would be

$$\begin{aligned}\hat{H} &= \frac{1}{2m}\left(\hat{\mathbf{p}} - \frac{q}{c}\mathbf{A}\right)^2 + V \\ &= \frac{1}{2m}\hat{\mathbf{p}}^2 - \frac{q}{mc}\mathbf{A} \cdot \hat{\mathbf{p}} + V + \frac{q^2}{2mc^2}\mathbf{A}^2.\end{aligned}$$

B. Magnetic field

Let the magnetic field be along z -axis,

$$\mathbf{B} = B_0\mathbf{z}_0$$

We choose vectorial potential to be of the form

$$A_x = -\frac{1}{2}B_0y, \quad A_y = \frac{1}{2}B_0x, \quad A_z = 0.$$

and let $q = -e$ (electron).

$$\begin{aligned}\hat{H} &= \frac{1}{2m}\hat{\mathbf{p}}^2 + V + \frac{e}{m}\left(-\frac{1}{2}B_0\hat{y}\hat{p}_x + \frac{1}{2}B_0\hat{x}\hat{p}_y\right) + \frac{e^2}{2m}\frac{B_0^2}{4}(\hat{x}^2 + \hat{y}^2) \\ &= \frac{1}{2m}\hat{\mathbf{p}}^2 + V + \frac{e}{2m}B_0(\hat{x}\hat{p}_y - \hat{y}\hat{p}_x) + \frac{e^2B_0^2}{8m}(\hat{x}^2 + \hat{y}^2) \\ &= \frac{1}{2m}\hat{\mathbf{p}}^2 + V + \frac{e}{2m}B_0\hat{L}_z + \frac{e^2B_0^2}{8m}(\hat{x}^2 + \hat{y}^2).\end{aligned}$$

The third term in the above Hamiltonian

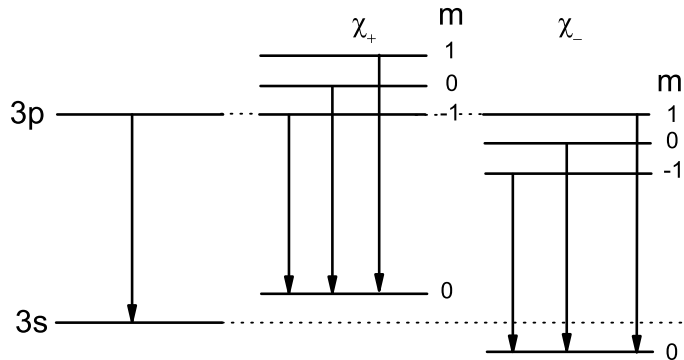
$$\frac{e}{2m}\hat{L}_zB_0 = -\left(-\frac{e}{2m}\hat{\mathbf{L}}\right) \cdot \mathbf{B} = -\boldsymbol{\mu} \cdot \mathbf{B}$$

shows the interaction energy between ‘‘orbital magnetic moment’’ and external magnetic field, while the fourth term $\frac{e^2B_0^2}{8m}(\hat{x}^2 + \hat{y}^2)$ reflects the confinement of magnetic field which corresponds the circular motion of charged particle around magnetic field. It is easy to show that for magnetic field in the laboratory $B < 10\text{Tesla} = 10^5\text{Gauss}$, the B^2 term is much smaller than B term

$$\frac{\frac{e^2B_0^2}{8m}(\hat{x}^2 + \hat{y}^2)}{\frac{e}{2m}B_0\hat{L}_z} \sim \frac{\frac{e^2B_0^2}{8m}a_0^2}{\frac{e\hbar}{2m}B_0} = \frac{eB_0}{4\hbar}a_0^2 = \frac{1.6 \times 10^{-19} \times 10 \times (0.53 \times 10^{-10})^2}{4 \times 1.055 \times 10^{-34}} \sim 1.0 \times 10^{-5}$$

It is not surprising that we have the same result in Gaussian system

$$\frac{\frac{e^2B_0^2}{8mc^2}a_0^2}{\frac{e\hbar}{2mc}B_0} = \frac{eB_0}{4\hbar c}a_0^2 = \frac{4.8 \times 10^{-10} \times 10^5 \times (0.53 \times 10^{-8})^2}{4 \times 3 \times 10^{10} \times 1.055 \times 10^{-27}} \sim 1.0 \times 10^{-5}$$

FIG. 6: Normal Zeeman effect of sodium D lines.

C. Zeeman effect

When an atom is placed in a uniform external magnetic field, the energy levels are shifted. This phenomenon is known as the **Zeeman effect**. While the Zeeman effect in some atoms (e.g., hydrogen) showed the expected equally-spaced triplet, in other atoms the magnetic field split the lines into four, six, or even more lines and some triplets showed wider spacings than expected. These deviations were labeled the “anomalous Zeeman effect” and were very puzzling to early researchers. The explanation of these different patterns of splitting gave additional insight into the effects of electron spin. With the inclusion of electron spin in the total angular momentum, the other types of multiplets formed part of a consistent picture. So what has been historically called the “anomalous” Zeeman effect is really the normal Zeeman effect when electron spin is included.

For a single electron, we should include two more terms in the Hamiltonian: the interaction energy between spin magnetic moment and the external magnetic field, and the spin-orbit coupling term. The total Hamiltonian is now

$$\begin{aligned}\hat{H} &= \frac{1}{2m}\hat{\mathbf{p}}^2 + V(r) + \frac{e}{2m}\hat{L}_z B_0 + \frac{e}{m}\hat{S}_z B_0 + \xi(r)\hat{\mathbf{L}} \cdot \hat{\mathbf{S}} \\ &= \hat{H}_0 + \hat{H}_B + \hat{H}_{LS}.\end{aligned}$$

The nature of the Zeeman splitting depends critically on the strength of the external field in comparison with the internal field (Equation 20) that gives rise to spin-orbit coupling. For if $B_{ext} \ll B_{int}$, then fine structure \hat{H}_{LS} dominates, and \hat{H}_B can be treated as a small perturbation, whereas if $B_{ext} \gg B_{int}$, then the Zeeman effect dominates, and \hat{H}_{LS} becomes the perturbation. In the intermediate zone, where the two fields are comparable, we need the full machinery of degenerate perturbation theory, and it is necessary to diagonalize the relevant portion of the Hamiltonian by hand.

1. Normal Zeeman effect

For strong magnetic field case, external magnetic effect is larger than spin orbit coupling. In this case, the spherical symmetry is completely broken by the strong external field that selects a particular direction in space, the z -axis. We are left with cylindrical symmetry only - that is, invariance under rotation around the z -axis. The spin-orbit interaction is decoupled and $\xi(r)\hat{\mathbf{L}} \cdot \hat{\mathbf{S}}$ can be omitted.

$$\hat{H} = \frac{1}{2m}\hat{\mathbf{p}}^2 + V(r) + \frac{e}{2m}\hat{L}_z B_0 + \frac{e}{m}\hat{S}_z B_0.$$

The good quantum numbers are now n, l, m_l, m_s , or in uncoupled representation, we choose $\psi_{nlm_l\chi_{\pm}} = R_{nl}(r)Y_{lm_l}\chi_{\pm}$, which are nothing but the eigenfunctions of \hat{H} with eigenenergy

$$E_{nl} + \frac{e\hbar}{2m}B_0(m_l + 2m_s) = E_{nl} + \mu_B B_0(m_l \pm 1).$$

For sodium D-line, the frequency shift is the famous *Larmor frequency*

$$\Delta\omega = \frac{1}{\hbar} \frac{e\hbar}{2m} B_0 = \frac{eB_0}{2m}$$

We note that spin only shifts all energy levels up or down and does not contribute to the splitting (Figure 6). Therefore in discussing the splitting of spectral lines in strong magnetic field, we usually ignore the spin effect. Moreover, though there may be many split levels, the selection rules restricted the transition to three kinds. One spectral line thus splits into three (an odd number) with frequencies

$$\omega_0 + \Delta\omega, \omega_0, \omega_0 - \Delta\omega,$$

which is known as *normal Zeeman effect*. The spin-orbit coupling can be considered as first-order perturbation, the energy correction of which is just the expectation value of \hat{H}_{LS} on zeroth-order wavefunctions $R_{nl}(r)Y_{lm_l}\chi_{\pm}$.

2. Anomalous Zeeman effect

In the weak external field case, the spin-orbit coupling \hat{H}_{LS} dominates the interaction energy. Spin and orbit angular momenta couple first into a total angular momentum $\hat{\mathbf{J}}$. It is the magnetic moment of $\hat{\mathbf{J}}$ that interacts with the external field.

$$\hat{H} = \frac{1}{2m} \hat{\mathbf{p}}^2 + V(r) + \xi(r) \hat{\mathbf{L}} \cdot \hat{\mathbf{S}} + \left(\frac{e}{2m} \hat{\mathbf{L}} + \frac{e}{m} \hat{\mathbf{S}} \right) \cdot \mathbf{B}$$

We choose the coupled representation $|n, l; j, m_j\rangle$ and calculate the additional energy due to perturbation of magnetic field

$$\left(\frac{e}{2m} \hat{\mathbf{L}} + \frac{e}{m} \hat{\mathbf{S}} \right) \cdot \mathbf{B} = \frac{e}{2m} (\hat{\mathbf{L}} + 2\hat{\mathbf{S}}) \cdot \mathbf{B} = \frac{e}{2m} (\hat{\mathbf{J}} + \hat{\mathbf{S}}) \cdot \mathbf{B}$$

Note that the degeneracy for j has been removed by spin-orbit coupling. For m_j , owing to the term $\hat{\mathbf{J}} \cdot \mathbf{B} = \hat{J}_z B_0$, different m_j has different perturbation energy and the perturbation is now diagonal in $|j, m_j\rangle$. We could therefore calculate the additional energy by simply taking the expectation value of \hat{H}_B . Note that

$$(\hat{\mathbf{J}} + \hat{\mathbf{S}}) \cdot \mathbf{B} = \left(1 + \frac{\hat{\mathbf{S}}}{\hat{\mathbf{J}}} \right) \hat{\mathbf{J}} \cdot \mathbf{B} = \left(1 + \frac{\hat{\mathbf{J}} \cdot \hat{\mathbf{S}}}{\hat{\mathbf{J}}^2} \right) \hat{J}_z B_0$$

and

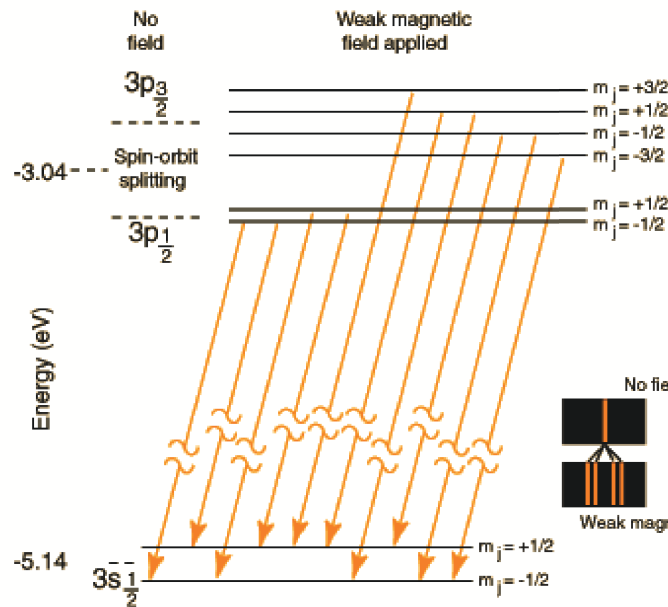
$$\begin{aligned} \hat{\mathbf{J}} &= \hat{\mathbf{L}} + \hat{\mathbf{S}} \rightarrow \hat{\mathbf{L}} = \hat{\mathbf{J}} - \hat{\mathbf{S}} \rightarrow \hat{\mathbf{L}}^2 = \hat{\mathbf{J}}^2 + \hat{\mathbf{S}}^2 - 2\hat{\mathbf{J}} \cdot \hat{\mathbf{S}} \\ \hat{\mathbf{J}} \cdot \hat{\mathbf{S}} &= \frac{1}{2}(\hat{\mathbf{J}}^2 + \hat{\mathbf{S}}^2 - \hat{\mathbf{L}}^2) \end{aligned}$$

so the additional energy

$$\begin{aligned} \Delta E &= \langle \hat{H}_B \rangle = \langle l, j, m_j | \frac{e}{2m} (\hat{\mathbf{J}} + \hat{\mathbf{S}}) \cdot \mathbf{B} | l, j, m_j \rangle \\ &= \langle l, j, m_j | \frac{e}{2m} \left(1 + \frac{\hat{\mathbf{J}}^2 + \hat{\mathbf{S}}^2 - \hat{\mathbf{L}}^2}{2\hat{\mathbf{J}}^2} \right) \hat{J}_z B_0 | l, j, m_j \rangle \\ &= \langle l, j, m_j | \frac{e}{2m} \left(1 + \frac{j(j+1)\hbar^2 + \frac{3}{4}\hbar^2 - l(l+1)\hbar^2}{2j(j+1)\hbar^2} \right) \hat{J}_z B_0 | l, j, m_j \rangle \\ &= \frac{eB_0}{2m} (m_j \hbar) \left(1 + \frac{j(j+1) + \frac{3}{4} - l(l+1)}{2j(j+1)} \right). \end{aligned}$$

Again we take the sodium D-line as an example. The spin-orbit coupling splits the $3p$ level into $3p_{3/2}$ and $3p_{1/2}$ while the $3s$ level is unchanged. For $3p_{3/2}$ level, i.e. $j = 3/2, l = 1$, we have

$$\Delta E = \frac{eB_0}{2m} \left(1 + \frac{\frac{3}{2} \times \frac{5}{2} + \frac{3}{4} - 1 \times 2}{2 \times \frac{3}{2} \times \frac{5}{2}} \right) m_j \hbar = \frac{eB_0}{2m} \frac{4}{3} m_j \hbar,$$

FIG. 7: Anomalous Zeeman effect of sodium D lines.

while m_j takes values of $m_j = \frac{3}{2}, \frac{1}{2}, -\frac{1}{2}, -\frac{3}{2}$. The $3p_{3/2}$ level thus splits into four. For $3p_{1/2}$ level, i.e., $j = 1/2, l = 1$, we have

$$\Delta E = \frac{eB_0}{2m} \left(1 + \frac{\frac{1}{2} \times \frac{3}{2} + \frac{3}{4} - 1 \times 2}{2 \times \frac{1}{2} \times \frac{3}{2}} \right) m_j \hbar = \frac{eB_0}{2m} \frac{2}{3} m_j \hbar,$$

with $m_j = \frac{1}{2}, -\frac{1}{2}$ and the $3p_{1/2}$ level splits into two. Finally, for $3s_{1/2}$ level, we again have doublet

$$\Delta E = \frac{eB_0}{2m} \left(1 + \frac{\frac{1}{2} \times \frac{3}{2} + \frac{3}{4}}{2 \times \frac{1}{2} \times \frac{3}{2}} \right) m_j \hbar = \frac{eB_0}{2m} 2m_j \hbar = \pm \frac{eB_0}{2m} \hbar$$

The spectral line from $3p$ to $3s$ splits into ten as shown in Figure 7, which is known as the *anomalous Zeeman effect*.

Problem 6 Zeng Jinyan, *Textbook*, p240, 9, 10a

VI. SPIN SINGLET AND TRIPLET

Suppose now that we have two electrons (neutral helium, for example. The theory here applies equally to two spin-1/2 particles - for example electron and the proton in the ground state of hydrogen atom): each can have spin up or spin down, so there are four possibilities in all for uncoupled state

$$\chi_+(1)\chi_+(2), \quad \chi_+(1)\chi_-(2), \quad \chi_-(1)\chi_+(2), \quad \chi_-(1)\chi_-(2)$$

or

$$\uparrow\uparrow, \uparrow\downarrow, \downarrow\uparrow, \downarrow\downarrow$$

Question: What is the total angular momentum of the atom?

For one electron, we have the rules

$$\begin{aligned} \hat{S}_x \chi_+ &= \frac{\hbar}{2} \chi_-, & \hat{S}_x \chi_- &= \frac{\hbar}{2} \chi_+ \\ \hat{S}_y \chi_+ &= i \frac{\hbar}{2} \chi_-, & \hat{S}_y \chi_- &= -i \frac{\hbar}{2} \chi_+ \\ \hat{S}_z \chi_+ &= \frac{\hbar}{2} \chi_+, & \hat{S}_z \chi_- &= -\frac{\hbar}{2} \chi_- \end{aligned} \quad (21)$$

For example

$$\hat{S}_x \chi_+ = \begin{pmatrix} 0 & \frac{\hbar}{2} \\ \frac{\hbar}{2} & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \frac{\hbar}{2} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \frac{\hbar}{2} \chi_-, \text{ etc.}$$

We could equivalently choose the uncoupled representation of commutative operators set $(\hat{S}_{1z}, \hat{S}_{2z})$ or the coupled representation of operators set (\hat{S}^2, \hat{S}_z) (We omit \hat{S}_1^2, \hat{S}_2^2 which are constants). Define the total spin operator

$$\hat{\mathbf{S}} = \hat{\mathbf{S}}_1 + \hat{\mathbf{S}}_2$$

we have

$$\hat{S}^2 = \hat{S}_1^2 + \hat{S}_2^2 + 2\hat{\mathbf{S}}_1 \cdot \hat{\mathbf{S}}_2 = \frac{3\hbar^2}{2} + 2(\hat{S}_{1x}\hat{S}_{2x} + \hat{S}_{1y}\hat{S}_{2y} + \hat{S}_{1z}\hat{S}_{2z})$$

Each of the four composite states is an eigenstate of $\hat{S}_z = \hat{S}_{1z} + \hat{S}_{2z}$ – the z -components simply add

$$\begin{aligned} \hat{S}_z \chi_+(1) \chi_+(2) &= \hbar \chi_+(1) \chi_+(2) \\ \hat{S}_z \chi_+(1) \chi_-(2) &= 0, \\ \hat{S}_z \chi_-(1) \chi_+(2) &= 0, \\ \hat{S}_z \chi_-(1) \chi_-(2) &= -\hbar \chi_-(1) \chi_-(2). \end{aligned}$$

In matrix form, \hat{S}_z is diagonal

$$\hat{S}_z = \begin{pmatrix} \hbar & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -\hbar \end{pmatrix}$$

As for \hat{S}^2 we use (21) to obtain

$$\begin{aligned} \hat{S}^2 \chi_+(1) \chi_+(2) &= \frac{3\hbar^2}{2} \chi_+(1) \chi_+(2) + 2 \left(\frac{\hbar}{2}\right)^2 \chi_-(1) \chi_-(2) \\ &\quad + 2 \left(\frac{i\hbar}{2}\right)^2 \chi_-(1) \chi_-(2) + 2 \left(\frac{\hbar}{2}\right)^2 \chi_+(1) \chi_+(2) \\ &= 2\hbar^2 \chi_+(1) \chi_+(2) \\ \hat{S}^2 \chi_+(1) \chi_-(2) &= \frac{3\hbar^2}{2} \chi_+(1) \chi_-(2) + 2 \left(\frac{\hbar}{2}\right) \chi_-(1) \left(\frac{\hbar}{2}\right) \chi_+(2) \\ &\quad + 2 \left(\frac{i\hbar}{2}\right) \chi_-(1) \left(-\frac{i\hbar}{2}\right) \chi_+(2) + 2 \left(\frac{\hbar}{2}\right) \chi_+(1) \left(-\frac{\hbar}{2}\right) \chi_-(2) \\ &= \hbar^2 \chi_+(1) \chi_-(2) + \hbar^2 \chi_-(1) \chi_+(2) \\ \hat{S}^2 \chi_-(1) \chi_+(2) &= \frac{3\hbar^2}{2} \chi_-(1) \chi_+(2) + 2 \left(\frac{\hbar}{2}\right) \chi_+(1) \left(\frac{\hbar}{2}\right) \chi_-(2) \\ &\quad + 2 \left(-\frac{i\hbar}{2}\right) \chi_+(1) \left(\frac{i\hbar}{2}\right) \chi_-(2) + 2 \left(-\frac{\hbar}{2}\right) \chi_-(1) \left(\frac{\hbar}{2}\right) \chi_-(2) \\ &= \hbar^2 \chi_+(1) \chi_-(2) + \hbar^2 \chi_-(1) \chi_+(2) \\ \hat{S}^2 \chi_-(1) \chi_-(2) &= \frac{3\hbar^2}{2} \chi_-(1) \chi_-(2) + 2 \left(\frac{\hbar}{2}\right)^2 \chi_+(1) \chi_+(2) \\ &\quad + 2 \left(-\frac{i\hbar}{2}\right) \chi_+(1) \left(-\frac{i\hbar}{2}\right) \chi_+(2) + 2 \left(-\frac{\hbar}{2}\right) \chi_-(1) \left(-\frac{\hbar}{2}\right) \chi_-(2) \\ &= 2\hbar^2 \chi_-(1) \chi_-(2) \end{aligned}$$

Thus in uncoupled representation the matrix of \hat{S}^2 contains non-diagonal elements

$$\hat{S}^2 = \begin{pmatrix} 2\hbar^2 & 0 & 0 & 0 \\ 0 & \hbar^2 & \hbar^2 & 0 \\ 0 & \hbar^2 & \hbar^2 & 0 \\ 0 & 0 & 0 & 2\hbar^2 \end{pmatrix}$$

the eigenvalues of which are $2\hbar^2, 2\hbar^2, 2\hbar^2, 0$. For eigenvalue $2\hbar^2$, i.e., the total spin quantum number $S = 1$, the corresponding eigenvectors satisfy

$$\begin{pmatrix} 2\hbar^2 & 0 & 0 & 0 \\ 0 & \hbar^2 & \hbar^2 & 0 \\ 0 & \hbar^2 & \hbar^2 & 0 \\ 0 & 0 & 0 & 2\hbar^2 \end{pmatrix} \begin{pmatrix} C_1 \\ C_2 \\ C_3 \\ C_4 \end{pmatrix} = 2\hbar^2 \begin{pmatrix} C_1 \\ C_2 \\ C_3 \\ C_4 \end{pmatrix}$$

where C_1 and C_4 may take any value and we have the restriction that C_2 must be the same as C_3 . There are three independent base kets

$$\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{pmatrix}$$

which are by themselves normalized. For eigenvalue 0, i.e., the total spin quantum number $S = 0$, the corresponding eigenvectors satisfy

$$\begin{pmatrix} 2\hbar^2 & 0 & 0 & 0 \\ 0 & \hbar^2 & \hbar^2 & 0 \\ 0 & \hbar^2 & \hbar^2 & 0 \\ 0 & 0 & 0 & 2\hbar^2 \end{pmatrix} \begin{pmatrix} C_1 \\ C_2 \\ C_3 \\ C_4 \end{pmatrix} = 0$$

in this case we have the eigenvector

$$\begin{pmatrix} 0 \\ \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \\ 0 \end{pmatrix}$$

Thus we have got four common eigenstates of \hat{S}^2, \hat{S}_z . They can be further classified according to eigenvalues of \hat{S}_z

$$\begin{aligned}
 S = 1, M_S = 1, & \quad \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \chi_+(1) \chi_+(2) \\
 S = 1, M_S = -1, & \quad \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} = \chi_-(1) \chi_-(2) \\
 S = 1, M_S = 0, & \quad \begin{pmatrix} 0 \\ \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{pmatrix} = \frac{1}{\sqrt{2}} [\chi_+(1) \chi_-(2) + \chi_-(1) \chi_+(2)] \\
 S = 0, M_S = 0, & \quad \begin{pmatrix} 0 \\ \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \\ 0 \end{pmatrix} = \frac{1}{\sqrt{2}} [\chi_+(1) \chi_-(2) - \chi_-(1) \chi_+(2)]
 \end{aligned}$$

The three states with $S = 1, M_S = 1, 0, -1$ are called spin-triplet states, while the state with $S = 0, M_S = 0$ is called spin-singlet state.

We could, on the other hand, arrive at the same result by means of the addition of angular momenta developed in this chapter. Starting from the highest weight state

$$\chi_+(1) \chi_+(2) = |S = 1, M_s = 1\rangle = |1, 1\rangle$$

we apply the lowering operator on both sides

$$\hat{S}_- |1, 1\rangle = (\hat{S}_{1-} + \hat{S}_{2-}) \chi_+(1) \chi_+(2)$$

and get

$$\sqrt{(1+1)(1-1+1)}\hbar |1, 0\rangle = \hbar \chi_-(1) \chi_+(2) + \hbar \chi_+(1) \chi_-(2)$$

that is

$$|1, 0\rangle = \frac{1}{\sqrt{2}} [\chi_-(1) \chi_+(2) + \chi_+(1) \chi_-(2)]$$

Do it again

$$|1, -1\rangle = \chi_-(1) \chi_-(2).$$

The other possible combination gives

$$|0, 0\rangle = \frac{1}{\sqrt{2}} [\chi_+(1) \chi_-(2) - \chi_-(1) \chi_+(2)]$$

This is illustrated in Figure 8.

It is interesting to note that $\chi_+(1) \chi_-(2)$ and $\chi_-(1) \chi_+(2)$ are not eigenstates of \hat{S}^2

$$\begin{aligned}
 \hat{S}^2 \chi_+(1) \chi_-(2) &= \hbar^2 \chi_+(1) \chi_-(2) + \hbar^2 \chi_-(1) \chi_+(2) \\
 \hat{S}^2 \chi_-(1) \chi_+(2) &= \hbar^2 \chi_-(1) \chi_+(2) + \hbar^2 \chi_+(1) \chi_-(2)
 \end{aligned}$$

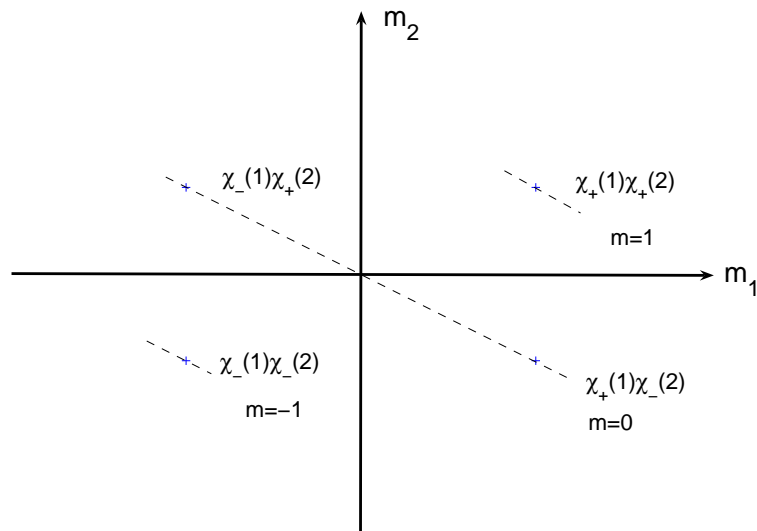


FIG. 8: Spin singlet and triplet.

while suitable combinations of them are

$$\begin{aligned}
 & \hat{S}^2 \left[\frac{1}{\sqrt{2}} \chi_+(1) \chi_-(2) \pm \frac{1}{\sqrt{2}} \chi_-(1) \chi_+(2) \right] \\
 &= \frac{\hbar^2}{\sqrt{2}} [\chi_+(1) \chi_-(2) + \chi_-(1) \chi_+(2)] \pm \frac{\hbar^2}{\sqrt{2}} [\chi_+(1) \chi_-(2) + \chi_-(1) \chi_+(2)] \\
 &= \begin{cases} 2\hbar^2 \left[\frac{1}{\sqrt{2}} \chi_+(1) \chi_-(2) + \frac{1}{\sqrt{2}} \chi_-(1) \chi_+(2) \right] \\ 0 \end{cases}
 \end{aligned}$$

corresponding to eigenvalues $2\hbar^2$ and 0.

We conclude that the combination of two spin-1/2 particles can carry a total spin of 1 or 0, depending on whether they occupy the triplet or the singlet configuration. The triplet states are eigenvectors of \hat{S}^2 with eigenvalue $2\hbar^2$ and the singlet is an eigenvector of \hat{S}^2 with eigenvalue 0.

Problem 7 Zeng Jinyan, *Textbook*, p240, 8

VII. IDENTICAL PARTICLES

A. Bosons and Fermions

Suppose particle 1 is in the (one-particle) state $\psi_a(\mathbf{r})$, and particle 2 is in the state $\psi_b(\mathbf{r})$. In that case the wave function for a two-particle system $\psi(\mathbf{r}_1, \mathbf{r}_2)$ is a simple product $\psi(\mathbf{r}_1, \mathbf{r}_2) = \psi_a(\mathbf{r}_1) \psi_b(\mathbf{r}_2)$. In classical mechanics we can always tell the particles apart – the particles are distinguishable. Quantum mechanically, electrons and all other particles are “utterly identical”, i.e. indistinguishable in principle. We simply construct wavefunctions which are *non-committal* as to which particle is in which state. There are actually two ways to do it:

$$\psi_{\pm}(\mathbf{r}_1, \mathbf{r}_2) = A [\psi_a(\mathbf{r}_1) \psi_b(\mathbf{r}_2) \pm \psi_b(\mathbf{r}_1) \psi_a(\mathbf{r}_2)]$$

Thus the theory admits two kinds of identical particles:

- Bosons, for which we use the plus sign, photons and mesons are bosons with integer spin;
- Fermions, for which we use the minus sign, protons and electrons are fermions with half-integer spin.

This connection between spin and “statistics” (as we shall see, bosons and fermions have quite different statistical properties) can be proved in relativistic quantum mechanics; in the non-relativistic theory it must be taken as an axiom.

There is a more general way to formulate the problem: Let $\psi(q_1, q_2)$ denotes the two-particle wavefunction where q_i contains all the coordinates of particles, spatial and spin. Define the exchange operator \hat{P} which interchanges the two particles

$$\hat{P}\psi(q_1, q_2) = \psi(q_2, q_1)$$

For identical particles, $\psi(q_1, q_2)$ and $\psi(q_2, q_1)$ are the same state, apart from a constant factor λ

$$\hat{P}\psi(q_1, q_2) = \psi(q_2, q_1) = \lambda\psi(q_1, q_2)$$

We do it once more

$$\hat{P}^2\psi(q_1, q_2) = \hat{P}\lambda\psi(q_1, q_2) = \lambda^2\psi(q_1, q_2)$$

and it follows that the eigenvalues of \hat{P} are ± 1

$$\lambda^2 = 1 \rightarrow \lambda = \pm 1$$

i.e.

$$\psi(q_1, q_2) = \pm\psi(q_2, q_1)$$

For the eigenvalue +1

$$\hat{P}\psi(q_1, q_2) = \psi(q_1, q_2)$$

or

$$\psi(q_1, q_2) = \psi(q_2, q_1)$$

This gives a symmetric wavefunction that is suitable to describe Bosons which follow the Bose-Einstein statistics

$$f_{BE}(\varepsilon) = \frac{1}{e^{\frac{\varepsilon}{kT}} - 1}$$

For the eigenvalue -1

$$\hat{P}\psi(q_1, q_2) = -\psi(q_1, q_2)$$

or

$$\psi(q_1, q_2) = -\psi(q_2, q_1)$$

This gives an anti-symmetric wavefunction that is suitable to describe Fermions which follow the Fermi-Dirac statistics

$$f_{FD}(\varepsilon) = \frac{1}{e^{\frac{\varepsilon}{kT}} + 1}$$

Meanwhile classical particles obey Boltzmann-Maxwell statistics

$$f_{BM}(\varepsilon) = e^{-\frac{\varepsilon}{kT}}.$$

This kind of symmetrization requirement, in particular for Fermions, leads to the famous Pauli Exclusion Principle: Two identical Fermions can not occupy the same state. For if $\psi_a = \psi_b$, then

$$\psi_-(q_1, q_2) = \frac{1}{\sqrt{2}} [\psi_a(q_1)\psi_a(q_2) - \psi_a(q_2)\psi_a(q_1)] = 0$$

At ultra-low temperature ($\sim 10^{-9}K$), Fermions go to degeneracy as shown in Figure 9, where particles fill the energy levels two by two (including spin), while Bosons go to Bose-Einstein Condensation, i.e. many bosons occupy the same

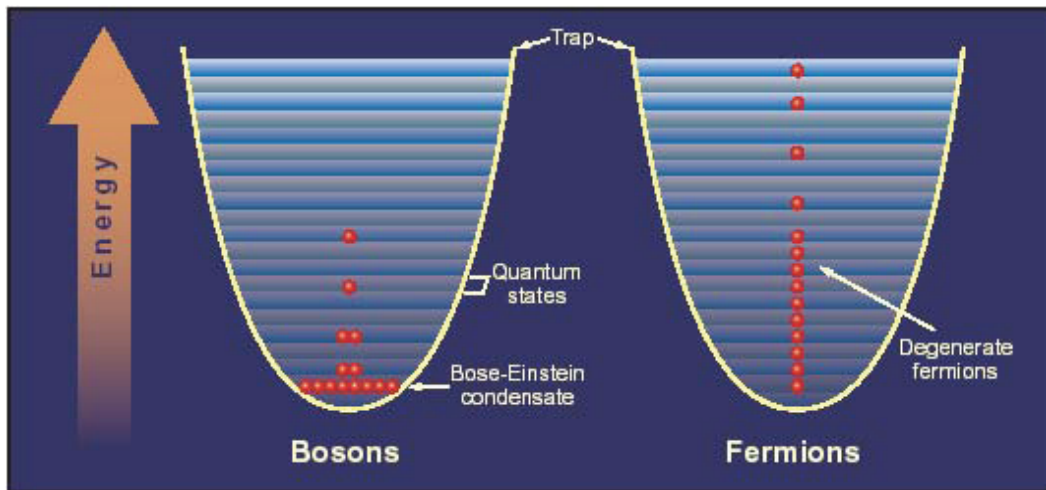


FIG. 9: Fermi degeneracy and Bose-Einstein condensation.

quantum state. The experimental observation for alkali atomic gases won Cornell, Wieman (^{87}Rb), Ketterle (^{23}Na) the 2001 Nobel Prize for Physics.

The normalized wavefunction for N -identical Particles takes the following form: For Fermions, the wavefunction is anti-symmetric

$$\phi_A(q_1, q_2 \dots q_N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \phi_1(q_1) & \phi_1(q_2) & \dots & \phi_1(q_N) \\ \phi_2(q_1) & \phi_2(q_2) & \dots & \phi_2(q_N) \\ \dots & \dots & \dots & \dots \\ \phi_N(q_1) & \phi_N(q_2) & \dots & \phi_N(q_N) \end{vmatrix}$$

which is nothing but the Slater determinant. Upon the exchange of two particles, the wavefunction changes its sign, i.e.

$$\phi_A(\dots q_i, \dots q_j \dots) = -\phi_A(\dots q_j, \dots q_i \dots)$$

Furthermore, if $\phi_i = \phi_j$, we get $\phi_A = 0$, which is Pauli's Principle for many particles. For Bosons, n_i particles may occupy ϕ_i state and

$$\sum_i n_i = N$$

The normalized symmetric wavefunction takes the form

$$\phi_S(q_1, q_2 \dots q_N) = \sqrt{\frac{\prod_i n_i!}{N!}} \sum_p \phi_{p_1}(q_1) \phi_{p_2}(q_2) \dots \phi_{p_N}(q_N)$$

where p stands for any permutation for particles in different states.

B. Symmetrization of wavefunctions

We give some examples on how to symmetrize the wavefunction. First we consider three identical particles of spin-0 Bosons in infinite-depth well. The ground state is simply multiplication of three ground state wavefunctions

$$\left(\sqrt{\frac{2}{a}} \sin \frac{\pi x_1}{a} \right) \left(\sqrt{\frac{2}{a}} \sin \frac{\pi x_2}{a} \right) \left(\sqrt{\frac{2}{a}} \sin \frac{\pi x_3}{a} \right) = \psi_1(x_1) \psi_1(x_2) \psi_1(x_3)$$

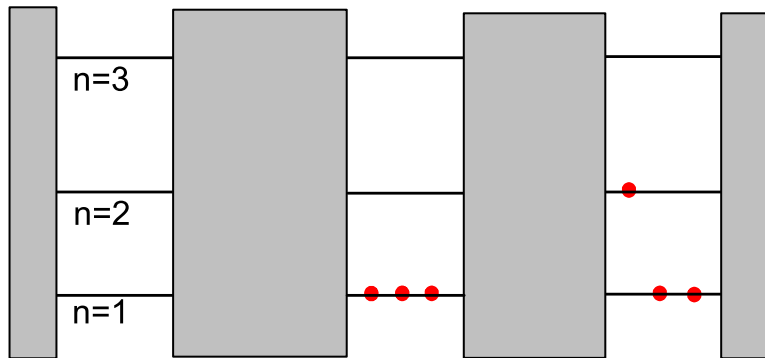


FIG. 10: Symmetrization of wavefunction for three Bosons in infinite potential well.

which is itself symmetric. For the first excited state, one of the Bosons may occupy its first excited state, while the other two still in their ground states (Figure 10)

$$\Psi(1, 2, 3) = \frac{1}{\sqrt{3}} [\psi_1(x_1) \psi_1(x_2) \psi_2(x_3) + \psi_2(x_1) \psi_1(x_2) \psi_1(x_3) + \psi_1(x_1) \psi_2(x_2) \psi_1(x_3)]$$

For arbitrary state the symmetric state is

$$\Psi(1, 2, 3) = \frac{1}{\sqrt{6}} \left\{ \begin{array}{l} \psi_m(x_1) \psi_n(x_2) \psi_p(x_3) + \psi_m(x_1) \psi_p(x_2) \psi_n(x_3) \\ + \psi_n(x_1) \psi_p(x_2) \psi_m(x_3) + \psi_n(x_1) \psi_m(x_2) \psi_p(x_3) \\ + \psi_p(x_1) \psi_m(x_2) \psi_n(x_3) + \psi_p(x_1) \psi_n(x_2) \psi_m(x_3) \end{array} \right\}$$

For two electrons, spin must be considered and we have the Pauli exclusion principle. First we assume no interaction between these two spins. The complete wavefunction is composed of spatial part and spin part

$$\psi_1\alpha, \psi_1\beta, \psi_2\alpha, \psi_2\beta, \psi_3\alpha, \psi_3\beta, \dots$$

with $\alpha = \chi_+$ and $\beta = \chi_-$. The ground state is a combination of electrons with spin up and spin down while keeping them in their spatial ground state ψ_1

$$\begin{aligned} \Psi(1, 2) &= \frac{1}{\sqrt{2}} [\psi_1(x_1) \alpha(1) \psi_1(x_2) \beta(2) - \psi_1(x_1) \beta(1) \psi_1(x_2) \alpha(2)] \\ &= \frac{1}{\sqrt{2}} \begin{vmatrix} \psi_1(x_1) \alpha(1) & \psi_1(x_1) \beta(1) \\ \psi_1(x_2) \alpha(2) & \psi_1(x_2) \beta(2) \end{vmatrix}. \end{aligned}$$

The first excited states, corresponding to one of the electrons is excited to higher level, are degenerate because the two spins can align parallel

$$\Psi(1, 2) = \frac{1}{\sqrt{2}} \begin{vmatrix} \psi_1(x_1) \alpha(1) & \psi_2(x_1) \alpha(1) \\ \psi_1(x_2) \alpha(2) & \psi_2(x_2) \alpha(2) \end{vmatrix}, \Psi(1, 2) = \frac{1}{\sqrt{2}} \begin{vmatrix} \psi_1(x_1) \beta(1) & \psi_2(x_1) \beta(1) \\ \psi_1(x_2) \beta(2) & \psi_2(x_2) \beta(2) \end{vmatrix},$$

or anti-parallel

$$\Psi(1, 2) = \frac{1}{\sqrt{2}} \begin{vmatrix} \psi_1(x_1) \alpha(1) & \psi_2(x_1) \beta(1) \\ \psi_1(x_2) \alpha(2) & \psi_2(x_2) \beta(2) \end{vmatrix}, \Psi(1, 2) = \frac{1}{\sqrt{2}} \begin{vmatrix} \psi_1(x_1) \beta(1) & \psi_2(x_1) \alpha(1) \\ \psi_1(x_2) \beta(2) & \psi_2(x_2) \alpha(2) \end{vmatrix}.$$

as illustrated in Figure 11. All of these wavefunction are antisymmetric because they describe Fermions.

Now we consider the spin-coupling of two-electrons. For single electron the complete wavefunction is

$$\psi(\mathbf{r}_1, \mathbf{S}_1) = \psi(\mathbf{r}_1) \chi(\mathbf{S}_1)$$

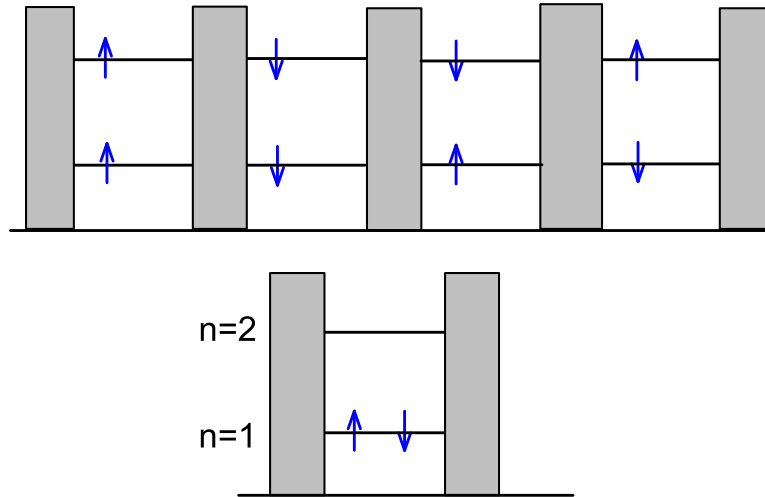


FIG. 11: Symmetrization of wavefunction for two electrons in infinite potential well.

and we denote the two-electron wavefunction as

$$\psi(\mathbf{r}_1, \mathbf{S}_1; \mathbf{r}_2, \mathbf{S}_2) = \psi(\mathbf{r}_1, \mathbf{r}_2) \chi(\mathbf{S}_1, \mathbf{S}_2)$$

Fermions require anti-symmetric complete wavefunctions, which could be realized by choosing:

- spatial wavefunction antisymmetric

$$\psi_A(\mathbf{r}_1, \mathbf{r}_2) = \frac{1}{\sqrt{2}} \{ \psi_m(\mathbf{r}_1) \psi_n(\mathbf{r}_2) - \psi_n(\mathbf{r}_1) \psi_m(\mathbf{r}_2) \}$$

and spin wavefunction symmetric - spin triplet state

$$\chi_S = \begin{cases} \uparrow\uparrow = \alpha(1) \alpha(2) \\ \downarrow\downarrow = \beta(1) \beta(2) \\ \frac{1}{\sqrt{2}} (\uparrow\downarrow + \downarrow\uparrow) = \frac{1}{\sqrt{2}} [\alpha(1) \beta(2) + \beta(1) \alpha(2)] \end{cases}$$

with eigenvalues of $\hat{\mathbf{S}}^2$: $2\hbar^2$, eigenvalues of \hat{S}_z : $\pm\hbar, 0$, respectively.

- or spatial wavefunction symmetric

$$\psi_S(\mathbf{r}_1, \mathbf{r}_2) = \begin{cases} \psi_m(\mathbf{r}_1) \psi_m(\mathbf{r}_2), & m = n \\ \frac{1}{\sqrt{2}} \{ \psi_m(\mathbf{r}_1) \psi_n(\mathbf{r}_2) + \psi_n(\mathbf{r}_1) \psi_m(\mathbf{r}_2) \}, & m \neq n \end{cases}$$

and spin wavefunction anti-symmetric - spin singlet

$$\chi_A = \frac{1}{\sqrt{2}} (\uparrow\downarrow - \downarrow\uparrow) = \frac{1}{\sqrt{2}} [\alpha(1) \beta(2) - \beta(1) \alpha(2)]$$

with eigenvalues of $\hat{\mathbf{S}}^2$: 0, eigenvalue of \hat{S}_z : 0.

Assume that the interaction Hamiltonian between the two electron spins takes the form

$$\hat{H} = -J \hat{\mathbf{S}}_1 \cdot \hat{\mathbf{S}}_2$$

The extension of this to crystal lattice, we may have the Heisenberg exchange model. Now

$$\hat{\mathbf{S}} = \hat{\mathbf{S}}_1 + \hat{\mathbf{S}}_2$$

and

$$\hat{\mathbf{S}}_1 \cdot \hat{\mathbf{S}}_2 = \frac{\hat{\mathbf{S}}^2 - \hat{\mathbf{S}}_1^2 - \hat{\mathbf{S}}_2^2}{2}$$

$J > 0$ means the interaction is ferromagnetic (spins prefer to align in the same direction and the eigenvalue of total spin takes the largest one to minimize the energy, i.e., $\hat{\mathbf{S}}^2 |FM\rangle = 1(1+1)\hbar^2 |FM\rangle$), while for $J < 0$ the interaction is anti-ferromagnetic (spins like to align oppositely and minimization the energy requires the eigenvalue of total spin to be as small as possible, i.e., $\hat{\mathbf{S}}^2 |AFM\rangle = 0$). We have the energy correction due to spin-coupling for spin triplet state

$$\hat{\mathbf{S}}_1 \cdot \hat{\mathbf{S}}_2 = \frac{2\hbar^2 - 2 \times \frac{3}{4}\hbar^2}{2} = \frac{\hbar^2}{4}, \quad \text{energy } E_T = -\frac{J\hbar^2}{4}$$

whereas for spin singlet state

$$\hat{\mathbf{S}}_1 \cdot \hat{\mathbf{S}}_2 = \frac{0 - 2 \times \frac{3}{4}\hbar^2}{2} = -\frac{3}{4}\hbar^2, \quad \text{energy } E_S = \frac{3}{4}J\hbar^2.$$

Example 8 Parahelium and Orthohelium: We have calculated by the variational method the ground state energy of helium atom, the simplest one after hydrogen. In that calculation we did not consider the spin of electrons. As an example of two-electron system (Fermion), the complete wavefunction is antisymmetric. In particular, the ground state would be

$$\psi_S(\mathbf{r}_1, \mathbf{r}_2) = \psi_{100}(\mathbf{r}_1) \psi_{100}(\mathbf{r}_2)$$

if we ignored the repulsive interaction between the two electrons. This is exactly the trial wavefunction what we have used in the variational method. Because it is a symmetric function, the spin state has to be antisymmetric,

$$\chi_A = \frac{1}{\sqrt{2}} [\alpha(1)\beta(2) - \beta(1)\alpha(2)]$$

so the ground state of helium should be a singlet configuration, with the spins oppositely aligned. The first-order energy correction due to the Coulomb repulsion between the two electrons can be calculated by perturbation theory

$$E_{11}^{(1)} = \frac{5Z}{8} \frac{e^2}{a_0}.$$

The excited states of helium consist of one electron in the hydrogen-like ground state and the other in an excited state

$$\psi(\mathbf{r}_1, \mathbf{r}_2) = \psi_{nlm}(\mathbf{r}_1) \psi_{100}(\mathbf{r}_2)$$

We can construct from this both symmetric and antisymmetric combinations, in the usual way

$$\begin{aligned} \psi^S(\mathbf{r}_1, \mathbf{r}_2) &= \frac{1}{\sqrt{2}} [\psi_{nlm}(\mathbf{r}_1) \psi_{100}(\mathbf{r}_2) + \psi_{100}(\mathbf{r}_1) \psi_{nlm}(\mathbf{r}_2)] \\ \psi^A(\mathbf{r}_1, \mathbf{r}_2) &= \frac{1}{\sqrt{2}} [\psi_{nlm}(\mathbf{r}_1) \psi_{100}(\mathbf{r}_2) - \psi_{100}(\mathbf{r}_1) \psi_{nlm}(\mathbf{r}_2)] \end{aligned}$$

the former goes with the antisymmetric spin configuration (the singlet),

$$\psi(\mathbf{r}_1, \mathbf{S}_1; \mathbf{r}_2, \mathbf{S}_2) = \psi^S(\mathbf{r}_1, \mathbf{r}_2) \chi^A(\mathbf{S}_1, \mathbf{S}_2)$$

and they are called parahelium, while the latter requires a symmetric spin configuration (the triplet),

$$\psi(\mathbf{r}_1, \mathbf{S}_1; \mathbf{r}_2, \mathbf{S}_2) = \psi^A(\mathbf{r}_1, \mathbf{r}_2) \chi^S(\mathbf{S}_1, \mathbf{S}_2)$$

and they are known as orthohelium. The ground state is necessarily parahelium; the excited states come in both forms. The first-order energy correction can be calculated for parahelium

$$\langle \psi^S \chi^A | \hat{H}' | \psi^S \chi^A \rangle = K + J, \quad E^S = \epsilon_l + \epsilon_0 + K + J$$

and for orthohelium

$$\langle \psi^A \chi^S | \hat{H}' | \psi^A \chi^S \rangle = K - J, \quad E^A = \epsilon_l + \epsilon_0 + K - J$$

where K and J are direct and exchange integrals

$$K = \frac{e^2}{4\pi\epsilon_0} \int d\mathbf{r}_1 d\mathbf{r}_2 \frac{|\psi_{nlm}(\mathbf{r}_1)|^2 |\psi_{100}(\mathbf{r}_2)|^2}{|\mathbf{r}_1 - \mathbf{r}_2|}$$

$$J = \frac{e^2}{4\pi\epsilon_0} \int d\mathbf{r}_1 d\mathbf{r}_2 \frac{\psi_{nlm}^*(\mathbf{r}_1) \psi_{100}(\mathbf{r}_1) \psi_{100}^*(\mathbf{r}_2) \psi_{nlm}(\mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|}.$$

There is no transitions between singlet and triplet states, thus there are two sets of independent spectra. We summarize the result in the following table

Parahelium $\chi^A(S=0)$	Orthohelium $\chi^S(S=1)$
$\psi_{100}(\mathbf{r}_1) \psi_{100}(\mathbf{r}_2) \chi^A$	Forbidden by Pauli exclusive principle
$[\psi_{nlm}(\mathbf{r}_1) \psi_{100}(\mathbf{r}_2) + \psi_{100}(\mathbf{r}_1) \psi_{nlm}(\mathbf{r}_2)] \chi^A / \sqrt{2}$	$[\psi_{nlm}(\mathbf{r}_1) \psi_{100}(\mathbf{r}_2) - \psi_{100}(\mathbf{r}_1) \psi_{nlm}(\mathbf{r}_2)] \chi^S / \sqrt{2}$

VIII. SUMMARY OF PART VI

In this part we treated the angular momentum in a more systematic way, i.e. instead of defining it as $\mathbf{r} \times \mathbf{p}$, we introduce the angular momentum as the infinitesimal generator of 3D rotation transformation. This definition is more general in the sense that the theory developed in this part applies equally to any angular momentum, orbital, spin or the summation of them. The commutation relation can be written in a compact form

$$[\hat{J}_j, \hat{J}_k] = i\hbar \epsilon_{jkl} \hat{J}_l, \quad j, k, l = x, y, z$$

The Casimir operator \hat{J}^2 commute each one of them, and that is why we choose the common eigenfunction set of \hat{J}^2 and \hat{J}_z . From the basic eigenequations

$$\begin{aligned} \hat{J}^2 |j, m\rangle &= j(j+1)\hbar^2 |j, m\rangle \\ \hat{J}_z |j, m\rangle &= m\hbar |j, m\rangle \\ \hat{J}_+ |j, m\rangle &= \sqrt{(j-m)(j+m+1)}\hbar |j, m+1\rangle \\ \hat{J}_- |j, m\rangle &= \sqrt{(j+m)(j-m+1)}\hbar |j, m-1\rangle \end{aligned}$$

where $\hat{J}_\pm = \hat{J}_x \pm i\hat{J}_y$, we can obtain the matrix form for these operators. For $j = \frac{1}{2}$, one has the three Pauli matrices $\hat{\sigma} = (\hat{\sigma}_x, \hat{\sigma}_y, \hat{\sigma}_z)$

$$\hat{\sigma}_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \hat{\sigma}_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \hat{\sigma}_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

which are related to the spin operator through

$$\hat{\mathbf{S}} = \frac{\hbar}{2} \hat{\sigma}.$$

Spin is a special angular momentum. Besides the usual properties of general angular momentum, Pauli matrices satisfy more rules due to $\hat{\sigma}_j^2 = 1$. For example,

$$\hat{\sigma}_j \hat{\sigma}_k = \delta_{jk} + i \sum_l \epsilon_{jkl} \hat{\sigma}_l$$

We should also append the wavefunction a spin part $\psi(x, y, z) \chi_{\pm\frac{1}{2}}$, with $\chi_{\pm\frac{1}{2}}$ eigenfunctions of \hat{S}_z and $\hat{\mathbf{S}}^2$

$$\chi_{\frac{1}{2}} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \chi_{-\frac{1}{2}} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

The angular momenta can couple to a total angular momentum, therefore we have uncoupled and coupled representations. Uncoupled representation is composed of common eigenstates of commutative operators $\hat{J}_1^2, \hat{J}_{1z}, \hat{J}_2^2, \hat{J}_{2z}$ denoted by $|j_1, j_2; m_1, m_2\rangle$, while in coupled representation we choose $\hat{J}_1^2, \hat{J}_2^2, \hat{J}^2, \hat{J}_z$ as our complete set with eigenkets $|j_1, j_2; j, m\rangle$ or $|j, m\rangle$. They are related by the so-called Clebsch-Gordan(CG) coefficients

$$|j, m\rangle = \sum_{m_1, m_2} |j_1, j_2; m_1, m_2\rangle \langle j_1, j_2; m_1, m_2 | j, m\rangle$$

We have given examples on how to calculate these coefficients and thus the transformation between the two representations. There are basically two types of this kind of coupling, spin-orbital interaction and spin states of two electrons.

- The spin-orbital coupling

$$\hat{H}' = \xi(r) \hat{\mathbf{L}} \cdot \hat{\mathbf{S}}$$

gives rise to the fine structure of alkali atoms where one energy level splits into two corresponding to $j = l \pm 1/2$ respectively. Putting the atom under a uniform external magnetic field, the Hamiltonian is changed into

$$\begin{aligned} \hat{H} &= \frac{1}{2m} \hat{\mathbf{p}}^2 + V(r) + \frac{e}{2m} \hat{L}_z B_0 + \frac{e}{m} \hat{S}_z B_0 + \xi(r) \hat{\mathbf{L}} \cdot \hat{\mathbf{S}} \\ &= \hat{H}_0 + \hat{H}_B + \hat{H}_{LS}. \end{aligned}$$

For strong magnetic field, we have the normal Zeeman effect where the spin-orbital coupling is treated as perturbation and one spectral line splits into an odd number. In a weak field, however, we observed an even number of spectral lines which is known as anomalous Zeeman effect.

- In describing the complete wavefunction of two identical electrons (Fermions, carrying half-integer spin), we need consider the symmetrization requirement of the wavefunctions, i.e. if the spatial wavefunction is antisymmetric, the spin part must be symmetric, which means the spin triplet state

$$\chi_S = \begin{cases} \uparrow\uparrow = \alpha(1)\alpha(2) \\ \downarrow\downarrow = \beta(1)\beta(2) \\ \frac{1}{\sqrt{2}}(\uparrow\downarrow + \downarrow\uparrow) = \frac{1}{\sqrt{2}}[\alpha(1)\beta(2) + \beta(1)\alpha(2)] \end{cases}$$

They are eigenvectors of $\hat{\mathbf{S}}^2$ with eigenvalue $2\hbar^2$ thus carry total spin 1. If the spatial wavefunction is symmetric, the spin part must be antisymmetric, which means the spin singlet state

$$\chi_A = \frac{1}{\sqrt{2}}(\uparrow\downarrow - \downarrow\uparrow) = \frac{1}{\sqrt{2}}[\alpha(1)\beta(2) - \beta(1)\alpha(2)]$$

It is an eigenvector of $\hat{\mathbf{S}}^2$ with eigenvalue 0 thus carries total spin 0. Both of them can be obtained as a result of transformation of uncoupled and coupled representations.

The symmetrization requirement for Fermions gives us the Pauli Exclusion Principle, i.e. two identical Fermions can not occupy the same state, while for Bosons one arrives at Bose-Einstein condensation where many particles are condensed into one single state.

Lecture Notes on Quantum Mechanics - Part VII

Yunbo Zhang

Instituter of Theoretical Physics, Shanxi University

We give a very brief introduction to quantum theory of scattering, and two techniques for calculating the scattering amplitude: partial wave analysis and the Born approximation.

Contents

I. Introduction to Classical Scattering Theory	2
II. Quantum Mechanical Description of Scattering Theory	4
III. Method of Transition Probability	5
IV. Lippman-Schwinger Equation - Green's Function	7
V. Partial Wave Analysis	10
A. l partial wave	10
B. Phase shift	12
VI. Summary on Part VII	14

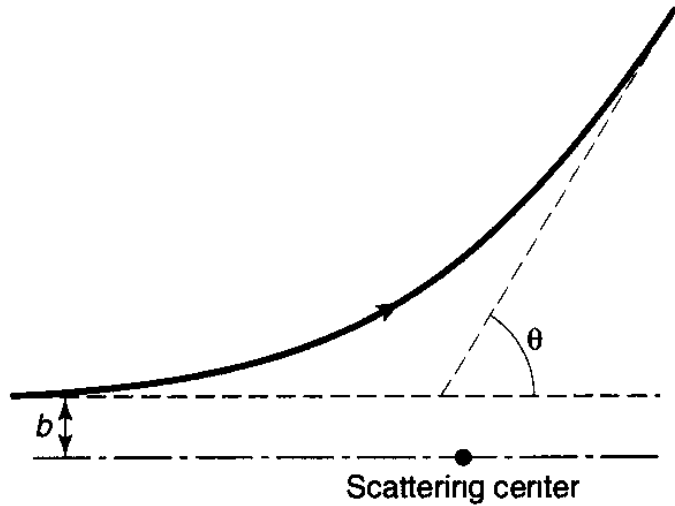


FIG. 1: The classical scattering problem, showing the impact parameter b and the scattering angle θ .

I. INTRODUCTION TO CLASSICAL SCATTERING THEORY

We take as an example the hard-sphere scattering in the classical collision problem. A particle incident on a target will scatter into some direction. There are two critical parameters in this process: the impact parameter b and the scattering angle θ . Assume for simplicity that the target is azimuthally symmetrical, so the trajectory remains in one plane, and that the target is very heavy, so the recoil is negligible. The essential problem of classical scattering theory is this: Given the impact parameter, calculate the scattering angle. Ordinarily, of course, the smaller the impact parameter, the greater the scattering angle (Figure 1).

Suppose the target is a billiard ball, of radius R , and the incident particle is also a billiard ball, which bounces off elastically (Figure 2). In terms of the angle a , the impact parameter is $b = R \sin a$, and the scattering angle is $\theta = \pi - 2a$, so

$$b = R \sin \left(\frac{\pi}{2} - \frac{\theta}{2} \right) = R \cos \frac{\theta}{2}$$

thus

$$\theta = \begin{cases} 2 \cos^{-1} (b/R), & \text{if } b \leq R \\ 0, & \text{if } b \geq R \end{cases}$$

More generally, particles incident within an infinitesimal patch of cross-sectional area $d\sigma$ will scatter into a corresponding infinitesimal solid angle $d\Omega$ (Figure 3). The larger $d\sigma$ is, the bigger $d\Omega$ will be; the proportionality factor $\sigma(\theta) \equiv d\sigma/d\Omega$, is called the *differential scattering cross-section*

$$d\sigma = \sigma(\theta)d\Omega.$$

In terms of the impact parameter and the azimuthal angle ϕ , $d\sigma = b db d\phi$ and $d\Omega = \sin \theta d\theta d\phi$, so

$$\sigma(\theta) = \frac{b}{\sin \theta} \left| \frac{db}{d\theta} \right|$$

(Since θ is typically a decreasing function of b , the derivative is actually negative - hence the absolute value sign.) In the case of hard-sphere scattering,

$$\frac{db}{d\theta} = -\frac{R}{2} \sin \frac{\theta}{2}$$

so

$$\sigma(\theta) = \frac{R \cos(\theta/2)}{\sin \theta} \frac{R}{2} \sin(\theta/2) = \frac{R^2}{4}$$

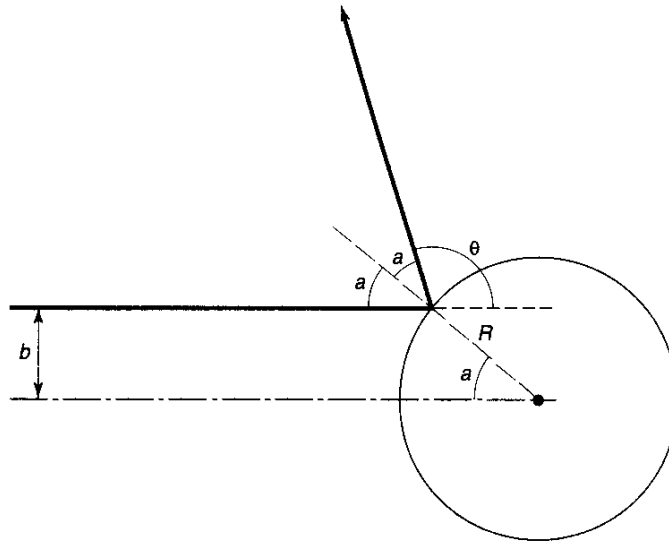
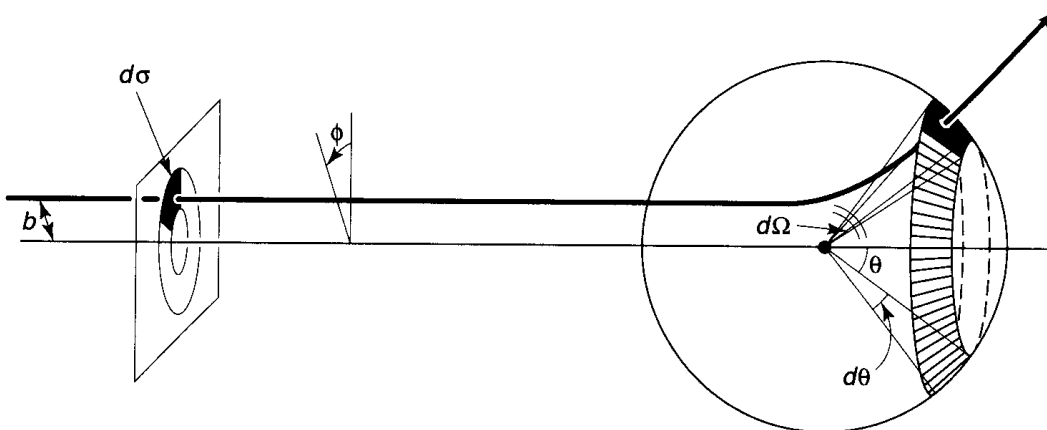


FIG. 2: Elastic hard-sphere scattering.

FIG. 3: Particles incident in the area $d\sigma$ scatter into the solid angle $d\Omega$.

This example is unusual in that the differential cross-section is actually independent of θ . The total cross-section is the integral of $\sigma(\theta)$ over all solid angles:

$$\sigma \equiv \int \sigma(\theta) d\Omega$$

roughly speaking, it is the total area of incident beam that is scattered by the target. For example, in the case of the hard sphere,

$$\sigma = \frac{R^2}{4} \int d\Omega = \pi R^2$$

which is just what we would expect: It's the cross-sectional area of the sphere; billiard balls incident within this area will hit the target, and those farther out will miss it completely. But the virtue of the formalism developed here is that it applies just as well to “soft” targets (such as the Coulomb field of a nucleus) that are not simply “hit or miss”.

Finally, suppose we have a beam of incident particles, with uniform intensity $j_i =$ number of incident particles per unit area, per unit time. The number of particles entering area $d\sigma$ (and hence scattering into solid angle $d\Omega$), per unit time, is $dN = j_i d\sigma = j_i \sigma(\theta) d\Omega$, so

$$\sigma(\theta) = \frac{1}{j_i} \frac{dN}{d\Omega}$$

This is often taken as the definition of the differential cross-section, because it makes reference only to quantities easily measured in the laboratory: If the detector accepts particles scattering into a solid angle $d\Omega$, we simply count the number recorded, per unit time, divide by $d\Omega$, and normalize to the intensity of the incident beam.

Consider the problem of Rutherford scattering: An incident particle of charge q_1 and kinetic energy E scatters off a heavy stationary particle of charge q_2 with potential

$$V(r) = \frac{1}{4\pi\epsilon_0} \frac{q_1 q_2}{r}. \quad (1)$$

From the classical mechanics, we could derive the formula relating the impact parameter to the scattering angle (equation 38.12 in Classical Mechanics by Corben and Stehle)

$$b = \frac{q_1 q_2}{8\pi\epsilon_0 E} \cot\left(\frac{\theta}{2}\right)$$

The differential scattering cross-section for this problem caused by an inverse square law force is

$$\begin{aligned} \sigma(\theta) &= \frac{q_1 q_2}{8\pi\epsilon_0 E} \frac{\cot(\theta/2)}{\sin\theta} \frac{q_1 q_2}{8\pi\epsilon_0 E} \frac{1}{2\sin^2(\theta/2)} \\ &= \frac{1}{2\sin(\theta/2)\cos(\theta/2)} \frac{q_1 q_2 \cos(\theta/2)}{8\pi\epsilon_0 E \sin(\theta/2)} \frac{q_1 q_2}{8\pi\epsilon_0 E} \frac{1}{2\sin^2(\theta/2)} \\ &= \left(\frac{q_1 q_2}{16\pi\epsilon_0 E \sin^2(\theta/2)} \right)^2 \end{aligned}$$

And it is easy to show that the total cross-section for Rutherford scattering is infinite

$$\sigma \equiv \int \sigma(\theta) \sin\theta d\theta d\phi = 2\pi \left(\frac{q_1 q_2}{16\pi\epsilon_0 E} \right)^2 \int \frac{\sin\theta d\theta}{\sin^4(\theta/2)}$$

This integral does not converge, for near $\theta = 0$ (and again near π) we have $\sin\theta \sim \theta$, $\sin(\theta/2) \sim \theta/2$, so the integral goes like

$$16 \times 2\pi \left(\frac{q_1 q_2}{16\pi\epsilon_0 E} \right)^2 \int_0^\epsilon \theta^{-3} d\theta = -8 \times 2\pi \left(\frac{q_1 q_2}{16\pi\epsilon_0 E} \right)^2 \theta^{-2} \Big|_0^\epsilon \rightarrow \infty$$

II. QUANTUM MECHANICAL DESCRIPTION OF SCATTERING THEORY

In the quantum theory of scattering, we imagine an incident plane wave, $\psi_i(z) = e^{ikz}$, traveling in the z -direction, which encounters a scattering potential, producing an outgoing spherical wave (Figure 4). That is, we will look for solutions to the Schrödinger equation of the general form

$$\psi(r, \theta) = e^{ikz} + f(\theta) \frac{e^{ikr}}{r}, \quad \text{for large } r \quad (2)$$

(The spherical wave must carry a factor of $1/r$, because this portion of $|\psi_s|^2$ must go like $1/r^2$ to conserve probability.) The wave number k is related to the energy of the incident particles in the usual way:

$$k = \sqrt{2mE/\hbar^2}$$

(As before, I shall assume the target is azimuthally symmetrical; in the more general case the amplitude f of the outgoing spherical wave could depend on ϕ as well as θ .)

The whole problem is to determine the scattering amplitude $f(\theta)$; it tells you the probability of scattering in a given direction θ , and hence is related to the differential cross-section. Indeed, the probability current density for incident wave, with units of particles per unit area per unit time, is

$$\begin{aligned} j_i &= \frac{i\hbar}{2m} \left[\psi_i \frac{\partial}{\partial z} \psi_i^* - \psi_i^* \frac{\partial}{\partial z} \psi_i \right] \\ &= \frac{i\hbar}{2m} [-ik - ik] = \frac{\hbar k}{m} \end{aligned}$$

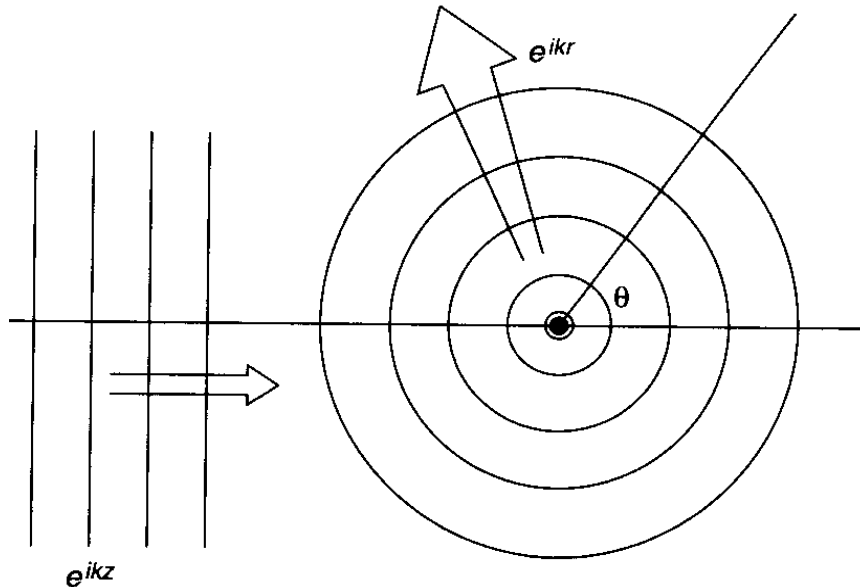


FIG. 4: Scattering of waves; incoming plane wave generates outgoing spherical wave.

We assume the scattering process to be elastic, that is, the internal states of the particles are left unchanged. In the central force field, the conservatives are angular momentum, but not the momentum. The probability current density for scattered wave in radial direction is

$$\begin{aligned}
 j_s &= \frac{i\hbar}{2m} \left[\psi_s \frac{\partial}{\partial r} \psi_s^* - \psi_s^* \frac{\partial}{\partial r} \psi_s \right] \\
 &= \frac{i\hbar}{2m} \left[f(\theta) \frac{e^{ikr}}{r} \frac{\partial}{\partial r} \left(f^*(\theta) \frac{e^{-ikr}}{r} \right) - c.c. \right] \\
 &= \frac{i\hbar}{2m} |f(\theta)|^2 \left[\frac{e^{ikr}}{r} \left(\frac{-ike^{-ikr}}{r} - \frac{e^{-ikr}}{r^2} \right) - c.c. \right] \\
 &= \frac{i\hbar}{2m} |f(\theta)|^2 \left[2 \frac{-ik}{r^2} \right] \\
 &= \frac{\hbar k}{m} \frac{|f(\theta)|^2}{r^2}
 \end{aligned}$$

The number of particles scattered into solid angle $d\Omega$, thus passing area $r^2 d\Omega$, per unit time, is

$$dN = j_s r^2 d\Omega = \frac{\hbar k}{m} |f(\theta)|^2 d\Omega$$

so

$$\sigma(\theta) = \frac{1}{j_i} \frac{dN}{d\Omega} = |f(\theta)|^2$$

Evidently the differential cross-section (which is the quantity of interest to the experimentalist) is equal to the absolute square of the scattering amplitude (which is obtained by solving the Schrödinger equation). In the following we will study three techniques for calculating the scattering amplitude: method of transition probability, the Born approximation and partial wave analysis.

III. METHOD OF TRANSITION PROBABILITY

In Part V, as an application example of Fermi's Golden Rule of the transition probability, we have obtained an expression for the differential cross section

$$\sigma(\theta, \varphi) = \frac{m^2}{4\pi^2 \hbar^4} \left| \iiint V(\mathbf{r}') e^{i(\mathbf{p}_0 - \mathbf{p}) \cdot \mathbf{r}' / \hbar} d^3 \mathbf{r}' \right|^2$$

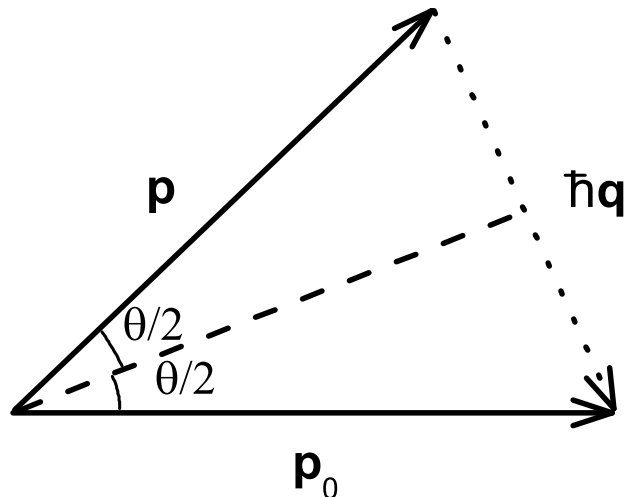


FIG. 5: Two momenta in the Born approximation: \mathbf{p}_0 points in the incident direction, \mathbf{p} in the scattered direction.

For elastic scattering, the momentum is conserved, i.e. the momentum magnitude of incident particle $|\mathbf{p}_0|$ is the same as that of scattered particle $|\mathbf{p}|$. For a spherically symmetrical potential, $V(\mathbf{r}') = V(r')$, the only change is the direction of the momentum and $\sigma(\theta, \varphi) = \sigma(\theta)$. Define

$$\hbar\mathbf{q} = \mathbf{p}_0 - \mathbf{p}$$

and from Figure 5 we see that

$$\hbar q = 2p \sin(\theta/2). \quad (3)$$

Let the polar axis for the \mathbf{r}' integral lie along \mathbf{q} , so that $(\mathbf{p}_0 - \mathbf{p}) \cdot \mathbf{r}' = \hbar q r' \cos \theta'$, the perturbation matrix element can be calculated as (note the different meanings of scattering angle θ and polar angle θ')

$$\begin{aligned} \iiint V(r') e^{i\mathbf{q} \cdot \mathbf{r}'} d^3\mathbf{r}' &= \iiint e^{iqr' \cos \theta'} V(r') r'^2 dr' \sin \theta' d\theta' d\phi' \\ &= 2\pi \iint V(r') r'^2 dr' e^{iqr' \cos \theta'} d(-\cos \theta') = 2\pi \int_0^\infty V(r') r'^2 dr' \left. \frac{e^{iqr' \cos \theta'}}{-iqr'} \right|_0^\pi \\ &= 2\pi \int_0^\infty V(r') r'^2 dr' \frac{e^{iqr'} - e^{-iqr'}}{iqr'} = \frac{2\pi}{iq} \int_0^\infty V(r') 2i \sin(qr') r' dr' \\ &= \frac{4\pi}{q} \int_0^\infty V(r') \sin(qr') r' dr' \end{aligned}$$

and it follows that

$$\sigma(\theta) = \frac{4m^2}{\hbar^4 q^2} \left| \int_0^\infty V(r') \sin(qr') r' dr' \right|^2. \quad (4)$$

Example 1 Yukawa Scattering: The Yukawa potential is a crude model for the binding force in an atomic nucleus. It has the form

$$V(r) = \beta \frac{e^{-\alpha r}}{r}$$

where β and α are constants. The integral in Equation (4) gives

$$\begin{aligned} \int_0^\infty V(r') \sin(qr') r' dr' &= \int_0^\infty \beta \frac{e^{-\alpha r'}}{2ir'} \left(e^{iqr'} - e^{-iqr'} \right) r' dr' \\ &= \frac{\beta}{2i} \int_0^\infty \left(e^{-(\alpha-iq)r'} - e^{-(\alpha+iq)r'} \right) dr' \\ &= \frac{\beta}{2i} \left(\frac{1}{\alpha-iq} - \frac{1}{\alpha+iq} \right) = \frac{\beta q}{\alpha^2 + q^2} \end{aligned}$$

and

$$\sigma(\theta) = \frac{4m^2}{\hbar^4} \frac{\beta^2}{(\alpha^2 + q^2)^2}$$

The total cross-section can be shown to be (Griffiths, prob. 11.12)

$$\sigma = \int \sigma(\theta) \sin\theta d\theta d\phi = \pi \left(\frac{4m\beta}{\alpha\hbar} \right)^2 \frac{1}{(\alpha\hbar)^2 + 8mE}.$$

Example 2 Rutherford scattering: If we put in $\beta = q_1 q_2 / 4\pi\epsilon_0$, $\alpha = 0$, the Yukawa potential reduces to the Coulomb potential (1), describing the electrical interaction of two point charges. Evidently the differential cross-section is

$$\sigma(\theta) = \left(\frac{q_1 q_2}{4\pi\epsilon_0} \frac{2m}{(\hbar q)^2} \right)^2 = \left(\frac{q_1 q_2}{16\pi\epsilon_0 E \sin^2(\theta/2)} \right)^2 \quad (5)$$

where we have used equation (3) and

$$p = \hbar k = \sqrt{2mE}.$$

Equation (5) is precisely the Rutherford formula we have obtained in classical theory. It happens that for the Coulomb potential, classical mechanics, the Born approximation, and quantum field theory all yield the same result.

IV. LIPPMAN-SCHWINGER EQUATION - GREEN'S FUNCTION

The general scattering wavefunctions

$$\psi(\mathbf{r}, \theta) = e^{ikz} + f(\theta) \frac{e^{ikr}}{r}$$

satisfy the time-independent Schrödinger equation,

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

with $V(\mathbf{r})$ the scattering potential. This can be written more succinctly as the superficial form of the Helmholtz equation

$$(\nabla^2 + k^2) \psi = \frac{2m}{\hbar^2} V\psi \quad (6)$$

Note, however, that the “inhomogeneous” term on the right hand itself depends on ψ . Suppose we could find a function $G(\mathbf{r})$ that solves the Helmholtz equation with a delta-function “source”

$$(\nabla^2 + k^2) G(\mathbf{r}, \mathbf{r}') = \delta(\mathbf{r} - \mathbf{r}'), \quad (7)$$

it is easy to show that

$$\psi(\mathbf{r}) = \frac{2m}{\hbar^2} \int d^3\mathbf{r}' G(\mathbf{r}, \mathbf{r}') V(\mathbf{r}') \psi(\mathbf{r}')$$

is a special solution to the inhomogeneous equation (6). Indeed

$$\begin{aligned} (\nabla^2 + k^2) \psi(\mathbf{r}) &= (\nabla^2 + k^2) \frac{2m}{\hbar^2} \int d^3\mathbf{r}' G(\mathbf{r}, \mathbf{r}') V(\mathbf{r}') \psi(\mathbf{r}') \\ &= \frac{2m}{\hbar^2} \int d^3\mathbf{r}' \delta(\mathbf{r} - \mathbf{r}') V(\mathbf{r}') \psi(\mathbf{r}') = \frac{2m}{\hbar^2} V(\mathbf{r}) \psi(\mathbf{r}) \end{aligned}$$

$G(\mathbf{r}, \mathbf{r}')$ is called the *Green's function* for the Helmholtz equation. (In general, the Green's function for a given differential equation represents the “response” to a delta-function source.)

Our first task is to solve Equation (7) for $G(\mathbf{r}, \mathbf{r}')$. This is most easily accomplished by taking the Fourier transform, which turns the differential equation into an algebraic equation. After some heavy analysis, including contour integration we get

$$G(\mathbf{r}, \mathbf{r}') = -\frac{e^{ik|\mathbf{r}-\mathbf{r}'|}}{4\pi|\mathbf{r}-\mathbf{r}'|}$$

(See Zeng Jinyan, textbook, page 356, or Griffiths, page 411, for detail.) The general solution for equation (6) can be expressed as

$$\psi(\mathbf{r}) = \psi_0(\mathbf{r}) + \frac{2m}{\hbar^2} \int d^3\mathbf{r}' G(\mathbf{r}, \mathbf{r}') V(\mathbf{r}') \psi(\mathbf{r}')$$

in which $\psi_0(\mathbf{r})$ satisfy the homogeneous Helmholtz equation

$$(\nabla^2 + k^2) \psi_0(\mathbf{r}) = 0.$$

For the finite range potential, $\psi_0(\mathbf{r})$ can be taken as the incident wavefunction

$$\psi_i(\mathbf{r}) = e^{ikz} = e^{i\mathbf{p}_0 \cdot \mathbf{r}/\hbar}$$

The scattering problem reduces to the following integral form of the Schrödinger equation

$$\begin{aligned} \psi(\mathbf{r}) &= \psi_i(\mathbf{r}) - \frac{m}{2\pi\hbar^2} \int d^3\mathbf{r}' \frac{e^{ik|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|} V(\mathbf{r}') \psi(\mathbf{r}') \\ &= \psi_i(\mathbf{r}) + \psi_s(\mathbf{r}) \end{aligned} \quad (8)$$

which also bears the name of **Lippman-Schwinger equation**. At first glance it looks like an explicit solution to the Schrödinger equation (for any potential) – which is too good to be true. Don't be deceived: There's a ψ under the integral sign on the right-hand side, so we can't do the integral unless we already know the solution! Nevertheless, it is possible to repeat the integral successively

$$\begin{aligned} \psi(\mathbf{r}) &= \psi_0(\mathbf{r}) - \frac{m}{2\pi\hbar^2} \int d^3\mathbf{r}' \frac{e^{ik|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|} V(\mathbf{r}') \left\{ \psi_0(\mathbf{r}') - \frac{m}{2\pi\hbar^2} \int d^3\mathbf{r}'' \frac{e^{ik|\mathbf{r}'-\mathbf{r}''|}}{|\mathbf{r}'-\mathbf{r}''|} V(\mathbf{r}'') \psi_0(\mathbf{r}'') \right\} \\ &= \psi_0(\mathbf{r}) - \frac{m}{2\pi\hbar^2} \int d^3\mathbf{r}' \frac{e^{ik|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|} V(\mathbf{r}') \psi_0(\mathbf{r}') \\ &\quad + \left(\frac{m}{2\pi\hbar^2} \right)^2 \iint d^3\mathbf{r}' d^3\mathbf{r}'' \frac{e^{ik|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|} V(\mathbf{r}') \frac{e^{ik|\mathbf{r}'-\mathbf{r}''|}}{|\mathbf{r}'-\mathbf{r}''|} V(\mathbf{r}'') \psi_0(\mathbf{r}'') + \dots \end{aligned}$$

Here comes the *Born Approximation*: Suppose the incoming plane wave is not substantially altered by the potential; then it makes sense to only keep the first-order, i.e.

$$\psi(\mathbf{r}) \simeq \psi_0(\mathbf{r}) - \frac{m}{2\pi\hbar^2} \int d^3\mathbf{r}' \frac{e^{ik|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|} V(\mathbf{r}') \psi_0(\mathbf{r}')$$

Here \mathbf{r} is the coordinate where one observes the scattered particles, while \mathbf{r}' is in the range of the scattering center. In this sense, the macroscopic coordinate \mathbf{r} is much larger than the microscopic coordinate \mathbf{r}' (Figure 6), i.e. $|\mathbf{r}| \gg |\mathbf{r}'|$ for all points that contribute to the above integral, so

$$|\mathbf{r}-\mathbf{r}'|^2 = r^2 + r'^2 - 2\mathbf{r} \cdot \mathbf{r}' \simeq r^2 \left(1 - \frac{2\mathbf{r} \cdot \mathbf{r}'}{r^2} \right)$$

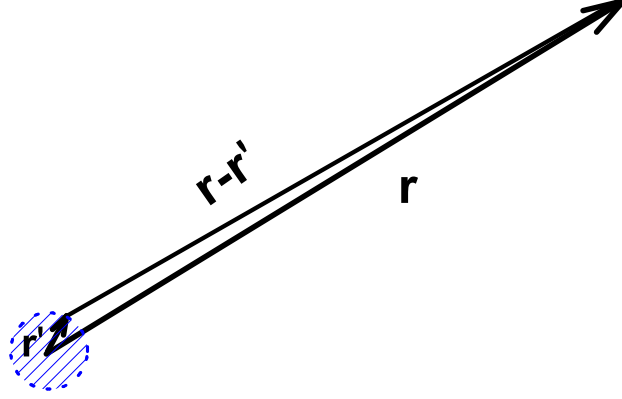


FIG. 6: Macroscopic and microscopic coordinates.

and hence

$$|\mathbf{r} - \mathbf{r}'| \simeq r - \frac{\mathbf{r}}{r} \cdot \mathbf{r}'$$

In the insensitive denominator we can afford to make the more radical approximation $|\mathbf{r} - \mathbf{r}'| \simeq r$; in the sensitive exponent we need to keep the next term. Moreover, we see that

$$k \frac{\mathbf{r}}{r} = \frac{\mathbf{p}}{\hbar}$$

i.e. the momentum of scattered particle. In the case of scattering, we want

$$\psi_0(\mathbf{r}) = e^{ikz} = e^{i\mathbf{p}_0 \cdot \mathbf{r} / \hbar}$$

representing an incident plane wave. For large r , then

$$\begin{aligned} \psi(\mathbf{r}) &\simeq e^{ikz} - \frac{m}{2\pi\hbar^2} \int d^3\mathbf{r}' \frac{e^{ikr} e^{-ik\frac{\mathbf{r}}{r} \cdot \mathbf{r}'}}{r} V(\mathbf{r}') \psi_0(\mathbf{r}') \\ &= e^{ikz} - \frac{m}{2\pi\hbar^2} \frac{e^{ikr}}{r} \int d^3\mathbf{r}' e^{-i\mathbf{p} \cdot \mathbf{r}' / \hbar} V(\mathbf{r}') \psi_0(\mathbf{r}') \end{aligned}$$

This is in the standard form (equation (2)), and we can read off the scattering amplitude

$$\begin{aligned} f(\theta) &= -\frac{m}{2\pi\hbar^2} \int d^3\mathbf{r}' e^{-i\mathbf{p} \cdot \mathbf{r}' / \hbar} V(\mathbf{r}') e^{i\mathbf{p}_0 \cdot \mathbf{r}' / \hbar} \\ &= -\frac{m}{2\pi\hbar^2} \int d^3\mathbf{r}' V(\mathbf{r}') e^{i(\mathbf{p}_0 - \mathbf{p}) \cdot \mathbf{r}' / \hbar} \end{aligned}$$

In Born approximation, the differential cross-section is then

$$\sigma(\theta) = \frac{m^2}{4\pi^2\hbar^4} \left| \int d^3\mathbf{r}' V(\mathbf{r}') e^{i(\mathbf{p}_0 - \mathbf{p}) \cdot \mathbf{r}' / \hbar} \right|^2$$

which is exactly the same as what have obtained from the transition probability method. What we have discussed in last subsection is actually in the first-order Born approximation.

V. PARTIAL WAVE ANALYSIS

Partial wave analysis is a very powerful tool to handle scattering problem in central force field with finite scattering range, especially for low energy. The basic idea is that here the angular momentum and its z -component (\hat{L}^2, \hat{L}_z) are conservatives because of the rotational invariance of the system. We recall that the Schrödinger equation for a spherically symmetrical potential $V(r)$ admits the separable solutions

$$\psi(r, \theta, \phi) = R(r)Y_{lm}(\theta, \phi)$$

where Y_{lm} is a spherical harmonic and $u(r) = rR(r)$ satisfies the radial equation

$$-\frac{\hbar^2}{2m} \frac{d^2 u}{dr^2} + \left(\frac{l(l+1)\hbar^2}{2mr^2} + V(r) \right) u = Eu$$

At very large r the potential goes to zero, and the centrifugal term is negligible, so

$$\frac{d^2 u}{dr^2} \simeq -\frac{2mE}{\hbar^2} u = -k^2 u$$

The general solution is

$$u(r) = Ce^{ikr} + De^{-ikr}$$

the first term represents an outgoing spherical wave, and the second an incoming one – for the scattered wave, we evidently want $D = 0$. At very large r , then,

$$R(r) \approx \frac{e^{ikr}}{r}$$

as we already deduced (on qualitative grounds) in the beginning of this section (Equation 2).

A. l partial wave

Let us now consider the incident particle described by a plane wave traveling along z -axis

$$\psi_i = e^{ikz}$$

which is the eigenstate of momentum and energy with eigenvalues

$$p_x = p_y = 0, \quad p_z = \hbar k, \quad E = \hbar^2 k^2 / 2m$$

The momentum \mathbf{p} , however, is not a conservative because

$$[\mathbf{p}, H] = [\mathbf{p}, V(r)] \neq 0$$

We note the incident wavefunction ψ_i is the eigenstate of \hat{L}_z with eigenvalue 0, but it is not the eigenstate of \hat{L}^2 . Rather, it is a linear combination of eigenstates of \hat{L}^2

$$e^{ikz} = \sum_{l=0}^{\infty} i^l (2l+1) j_l(kr) P^l(\cos \theta) \quad (9)$$

where j_l is a spherical Bessel function, P^l is a Legendre polynomial, and θ is the angle between z and r , i.e., the scattering angle. The expansion of a plane wave in terms of spherical waves is sometimes called *Rayleigh's formula*. It is nothing but the unique solution for the case that the scattering potential were identically zero. *The expansion of incident wave on the eigenstates of conservatives \hat{L}^2 and \hat{L}_z is the key point of partial wave analysis*. Let us write the general scattering wavefunctions with the scattering potential in the form

$$\psi(r, \theta) = \sum_{l=0}^{\infty} i^l (2l+1) A_l R_l(r) P^l(\cos \theta)$$

The reason why we can assume this kind of expansion lies in that:

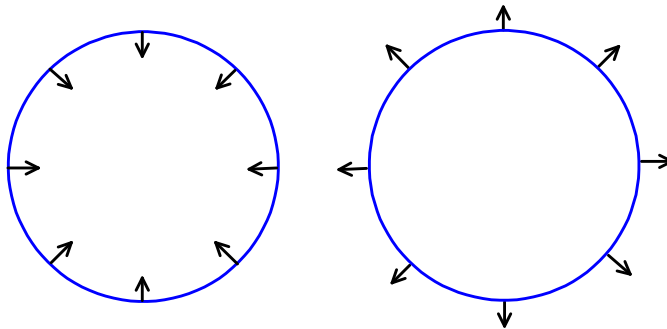


FIG. 7: Incoming and outgoing partial waves.

- The l partial waves in the incident plane wave

$$i^l(2l+1)j_l(kr)P^l(\cos\theta)$$

corresponding to eigenvalues $l(l+1)\hbar^2$ and 0 of \hat{L}^2 and \hat{L}_z , are *scattered separately* because of the conservation of angular momentum, i.e. they won't mix up during the scattering process.

- Moreover, the conservation of \hat{L}_z makes m also a good quantum number and our problem is irrelative of ϕ thus $m = 0$ for all the scattered waves. This simplifies the expansion on spherical harmonics Y_{lm} to that on Legendre polynomials $P^l(\cos\theta)$, which is nothing but Y_{l0} .

The remaining task is to find the detailed form of the radial wavefunction $R_l(r)$ and the corresponding coefficients A_l . Before doing this, let us consider the asymptotic behavior of these partial waves. The asymptotic forms of the Bessel functions in the limit $kr \rightarrow \infty$ are

$$j_l(kr) \rightarrow \frac{1}{kr} \sin\left(kr - \frac{1}{2}l\pi\right)$$

The l partial wave with angular momentum $l(l+1)\hbar^2$

$$\begin{aligned} i^l(2l+1)j_l(kr)P^l(\cos\theta) &\rightarrow \frac{1}{kr}i^l(2l+1)P^l(\cos\theta)\sin\left(kr - \frac{1}{2}l\pi\right) \\ &= \frac{1}{2ikr}i^l(2l+1)P^l(\cos\theta)\left(e^{i(kr-\frac{1}{2}l\pi)} - e^{-i(kr-\frac{1}{2}l\pi)}\right) \\ &= \frac{2l+1}{2kr}i^{l+1}P^l(\cos\theta)e^{-i(kr-\frac{1}{2}l\pi)} - \frac{2l+1}{2kr}i^{l+1}P^l(\cos\theta)e^{i(kr-\frac{1}{2}l\pi)} \end{aligned}$$

becomes essentially summation of an incoming spherical wave and an outgoing spherical wave. Thus

$$\begin{aligned} e^{ikz} &= \sum_l (l \text{ partial wave}) \\ &= \sum_l \left(\begin{array}{l} \text{incoming spherical partial wave with angular momentum } l(l+1)\hbar^2 \\ + \text{outgoing spherical partial wave with angular momentum } l(l+1)\hbar^2 \end{array} \right) \end{aligned}$$

The superposition and interference of those partial waves gives the incident plane wave e^{ikz} (Figure 7).

The physics of scattering is now clear. When the scattering potential is absent, we can analyze the plane wave as the sum of a spherically outgoing wave behaving like e^{-ikr}/r and a spherically incoming wave behaving like e^{ikr}/r for each r . The presence of the scattering center changes only the coefficient of outgoing wave

$$l \text{ outgoing spherical partial wave} \xrightarrow{kr \rightarrow \infty} -s_l \frac{2l+1}{2kr} i^{l+1} P^l(\cos\theta) e^{i(kr-\frac{1}{2}l\pi)}$$

Because of the conservation of angular momentum, the outgoing flux must equal the incoming flux for each partial wave separately. In other words, the coefficient of e^{-ikr}/r must be the same in magnitude as the coefficient of e^{ikr}/r . This means that

$$|s_l| = 1 \quad (10)$$

for all l . The most that can happen is a change in the phase of the outgoing wave and the incoming wave is completely unaffected. Equation (10) is known as the unitarity relation for the l partial wave. Thus superposition of the l partial waves

$$\frac{2l+1}{2kr} i^{l+1} P^l(\cos\theta) e^{-i(kr-\frac{1}{2}l\pi)} - s_l \frac{2l+1}{2kr} i^{l+1} P^l(\cos\theta) e^{i(kr-\frac{1}{2}l\pi)}$$

gives

$$\sum_l \frac{2l+1}{2kr} i^{l+1} P^l(\cos\theta) \left(e^{-i(kr-\frac{1}{2}l\pi)} - s_l e^{i(kr-\frac{1}{2}l\pi)} \right)$$

which should be the same as equation (2). So

$$e^{ikz} + f(\theta) \frac{e^{ikr}}{r} = \sum_l \frac{2l+1}{2kr} i^{l+1} P^l(\cos\theta) \left(e^{-i(kr-\frac{1}{2}l\pi)} - s_l e^{i(kr-\frac{1}{2}l\pi)} \right)$$

i.e.

$$\begin{aligned} & \sum_l \frac{2l+1}{2kr} i^{l+1} P^l(\cos\theta) \left(e^{-i(kr-\frac{1}{2}l\pi)} - e^{i(kr-\frac{1}{2}l\pi)} \right) + f(\theta) \frac{e^{ikr}}{r} \\ &= \sum_l \frac{2l+1}{2kr} i^{l+1} P^l(\cos\theta) \left(e^{-i(kr-\frac{1}{2}l\pi)} - s_l e^{i(kr-\frac{1}{2}l\pi)} \right) \end{aligned}$$

The incoming partial waves on left and right side cancel each other. Multiplication of a factor re^{-ikr} reduces the above equation into

$$-\sum_l \frac{2l+1}{2k} i^{l+1} P^l(\cos\theta) e^{-\frac{i}{2}l\pi} + f(\theta) = -\sum_l \frac{2l+1}{2k} i^{l+1} P^l(\cos\theta) s_l e^{-\frac{i}{2}l\pi}$$

and we get the scattering amplitude

$$\begin{aligned} f(\theta) &= \sum_l \frac{2l+1}{2k} i^{l+1} P^l(\cos\theta) e^{-\frac{i}{2}l\pi} (1 - s_l) \\ &= \sum_l \frac{2l+1}{2k} i P^l(\cos\theta) (1 - s_l) \end{aligned}$$

Evidently the angular distribution of the l partial wave, $P^l(\cos\theta) = Y_{10}(\theta, \phi)$, will not change in the scattering process. The indices of refraction and therefore the wavelengths are different for area $V(r) = 0$ and $V(r) \neq 0$ which gives rise to the change in the phases of outgoing and incoming waves. The combination of them leads to the *phase shift*.

B. Phase shift

The general scattering wavefunctions

$$\psi(r, \theta) = \sum_{l=0}^{\infty} i^l (2l+1) A_l R_l(r) P^l(\cos\theta)$$

is shown to have the form of

$$\sum_l \frac{2l+1}{2kr} i^{l+1} P^l(\cos\theta) \left(e^{-i(kr-\frac{1}{2}l\pi)} - s_l e^{i(kr-\frac{1}{2}l\pi)} \right)$$

We thus see that the only change in the wavefunction at a large distance as a result of scattering is to change the phase of the outgoing wave. Calling this phase $2\delta_l$ (the factor of 2 here is for conventional), we can write

$$s_l = e^{2i\delta_l}$$

with δ_l real. For the full scattering amplitude, we have

$$\begin{aligned} f(\theta) &= \sum_l \frac{2l+1}{2ik} P^l(\cos\theta) (e^{2i\delta_l} - 1) \\ &= \frac{1}{k} \sum_l (2l+1) e^{i\delta_l} \sin\delta_l P^l(\cos\theta) \end{aligned} \quad (11)$$

This expression for $f(\theta)$ rests on the twin principles of rotational invariance and probability conservation. The differential cross section $\sigma(\theta)$ can be obtained by just taking the modulus squared of equation (11). To obtain the total cross section we use the orthogonality of Legendre polynomials and obtain

$$\begin{aligned} \sigma &= \int |f(\theta)|^2 d\Omega = \frac{2\pi}{k^2} \sum_{ll'} (2l+1)(2l'+1) e^{-i\delta_l} e^{i\delta_{l'}} \sin\delta_l \sin\delta_{l'} \\ &\quad \times \int P^l(\cos\theta) P^{l'}(\cos\theta) \sin\theta d\theta \\ &= \frac{2\pi}{k^2} \sum_{ll'} (2l+1)(2l'+1) e^{-i\delta_l} e^{i\delta_{l'}} \sin\delta_l \sin\delta_{l'} \frac{2}{2l+1} \delta_{ll'} \\ &= \frac{4\pi}{k^2} \sum_l (2l+1) \sin^2\delta_l \end{aligned}$$

The application of partial wave analysis relies on finding the phase shift δ_l , i.e., $\delta_0, \delta_1, \delta_2, \dots$. This is valid for the potential which can be taken to be zero beyond some boundary and for low energy of the incident particles. Notice that for larger l , the angular momentum carried by the l partial wave is larger, and the system will less be affected by the central force field because it is farther away from the scattering center. We need only consider the low l partial waves. For very low energy, only s partial wave $l=0$ is taken into account, i.e. $\delta_0 \neq 0, \delta_1 = \delta_2 = \dots = 0$, we have $P^0(\cos\theta) = 1$ and

$$\begin{aligned} \sigma(\theta) &= \frac{1}{k^2} \sin^2\delta_0 \\ \sigma &= \frac{4\pi}{k^2} \sin^2\delta_0 \end{aligned}$$

If, on the other hand, only p wave is considered, $P^1(\cos\theta) = \cos\theta, \delta_1 \neq 0, \delta_0 = \delta_2 = \dots = 0$, then

$$\begin{aligned} \sigma(\theta) &= \frac{1}{k^2} 9 \sin^2\delta_1 \cos^2\theta \\ \sigma &= \frac{4\pi}{k^2} 3 \sin^2\delta_1 \end{aligned}$$

Or, if we keep s and p waves and neglect higher partial waves, $\delta_0 \neq 0, \delta_1 \neq 0, \delta_2 = \delta_3 = \dots = 0$

$$\begin{aligned} \sigma(\theta) &= \frac{1}{k^2} |e^{i\delta_0} \sin\delta_0 + 3e^{i\delta_1} \sin\delta_1 \cos\theta|^2 \\ \sigma &= \frac{4\pi}{k^2} (\sin^2\delta_0 + 3\sin^2\delta_1) \end{aligned}$$

Clearly there is interference between different partial waves.

Finally we give an example on how to calculate the phase shift for s partial wave. The potential model is

$$V(r) = \begin{cases} -V_0, & r < a \\ 0, & r > a \end{cases}$$

We try to find the positive energy solution ($E > 0$) to the radial equation for $l=0$

$$-\frac{\hbar^2}{2m} \frac{d^2u}{dr^2} + \left(\frac{l(l+1)\hbar^2}{2mr^2} + V(r) \right) u = Eu$$

which is quite different with the bound state solution in Part III. The equation reduces to

$$\begin{aligned}\frac{d^2u}{dr^2} &= -\frac{2m(E+V_0)}{\hbar^2}u = -k'^2u, & r < a \\ \frac{d^2u}{dr^2} &= -\frac{2mE}{\hbar^2}u = -k^2u, & r > a\end{aligned}$$

with solutions

$$\begin{aligned}u &= A \sin(k'r + \delta'_0), & r < a \\ u &= B \sin(kr + \delta_0), & r > a\end{aligned}$$

while the solution in $r > a$ is the scattering wave and hence δ_0 is the phase shift. At $r = 0$ the wavefunction $R(r) = u(r)/r$ should be finite, so $\delta'_0 = 0$. At $r = a$ the wavefunction and its derivative should be continuous

$$\begin{aligned}A \sin(k'a) &= B \sin(ka + \delta_0) \\ Ak' \cos(k'a) &= Bk \cos(ka + \delta_0)\end{aligned}$$

so

$$\frac{1}{k'} \tan k'a = \frac{1}{k} \tan(ka + \delta_0)$$

The phase shift is found to be

$$\delta_0 = \tan^{-1}\left(\frac{k}{k'} \tan k'a\right) - ka$$

For

$$V(r) = \begin{cases} +\infty, & r < a \\ 0, & r > a \end{cases}$$

we find

$$\delta_0 = -ka.$$

Inserting this phase shift into the total cross section, we get the result for hard sphere scattering

$$\sigma = \frac{4\pi}{k^2} \sin^2 \delta_0 = 4\pi a^2$$

which is four times larger than the classical result.

VI. SUMMARY ON PART VII

The main task in scattering problem is to calculate the differential cross section: It is defined as the ratio of the number of particles scattered into unit solid angle per unit time $\frac{dN}{d\Omega}$ with the number of incident particles crossing unit area per unit time j_i

$$\sigma(\theta) = \frac{1}{j_i} \frac{dN}{d\Omega}$$

The total cross-section is the integral of $\sigma(\theta)$ over all solid angles

$$\sigma \equiv \int \sigma(\theta) d\Omega$$

In quantum mechanics we look for solutions to the Schrödinger equation of the general form

$$\psi(r, \theta) = e^{ikz} + f(\theta) \frac{e^{ikr}}{r}, \quad \text{for large } r$$

The differential cross-section is equal to the absolute square of the scattering amplitude

$$\sigma(\theta) = |f(\theta)|^2.$$

We provided three methods to calculate this. The Lippman-Schwinger integral equation

$$\psi(\mathbf{r}) = \psi_i(\mathbf{r}) - \frac{m}{2\pi\hbar^2} \int d^3\mathbf{r}' \frac{e^{ik|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|} V(\mathbf{r}')\psi(\mathbf{r}'),$$

in the first order Born approximation, gives exactly the same result as we got through method of transition probability

$$\sigma(\theta) = \frac{4m^2}{\hbar^4 q^2} \left| \int_0^\infty V(r') \sin(qr') r' dr' \right|^2.$$

with

$$\hbar q = 2p \sin(\theta/2)$$

The scattering of a Coulomb potential gives the famous Rutherford formula

$$\sigma(\theta) = \left(\frac{q_1 q_2}{16\pi\epsilon_0 E \sin^2(\theta/2)} \right)^2$$

The Born approximation is more powerful for high energy scattering, while partial wave analysis is often used to calculate the low energy scattering, in which case we only need to take into account the low l partial waves. There is a change in the phase of the outgoing wave and the incoming wave is completely unaffected. The scattering amplitude is a summation for all l partial waves

$$f(\theta) = \frac{1}{k} \sum_l (2l+1) e^{i\delta_l} \sin \delta_l P^l(\cos \theta)$$

and the total cross section

$$\sigma = \frac{4\pi}{k^2} \sum_l (2l+1) \sin^2 \delta_l$$

One thus need to calculate the phase shift δ_l , i.e., $\delta_0, \delta_1, \delta_2, \dots$ for each partial wave. Usually for low energy scattering we only consider the lowest partial waves, s or p waves. For the square spherical potential, we have

$$\delta_0 = \tan^{-1} \left(\frac{k}{k'} \tan k'a \right) - ka$$

while for hard sphere

$$\delta_0 = -ka$$

The quantum result for the total cross section is four times larger than the classical one

$$\sigma = 4\pi a^2.$$