

Quantum Mechanics

Lecture Notes

Dr. Khorshed Kabir

This notes are compiled under supervision of Dr. Khorshed Kabir

UNIVERSITY OF DHAKA

First release, 2019

Thanks to the following individuals

Muhammad Shahnoor Rahman, Department of Physics, University of Dhaka.



Contents

1	sheet-1 :Dirac Delta Function	13
1.1	Other representation of delta function	15
1.2	Properties of the delta function	16
1.2.1	Notes	18
1.3	Derivatives of the delta function	19
1.4	Integration of the delta function	21
1.5	Three - dimensional delta function	21
1.6	Fourier Transformation	23
1.7	Definition of Dirac Delta function and its Representations	24
1.8	Properties of the Delta Function	28
1.9	Derivative of the Delta Function	32
1.10	Integration of the Delta Function	34
1.11	Three dimensional delta function	35
1.12	Fourier transform	37
1.12.1	Parseval Identity	38

2	sheet-2 : Linear Vector Space	17
2.1	Definition	17
2.2	Example of linear vector space	18
2.3	Inner product space or a unitary vector space	20
2.4	Examples of scalar product	21
2.5	Norm of a vector	23
2.6	Metric included by the scalar product	23
2.7	Schwarz's inequality	24
2.8	Analogy of Schwarz inequality with vectors in a three-dimensional Euclidean space \mathbb{R}^3	25
2.9	Orthogonality and linear independence	25
2.9.1	Linear independence	25
2.10	Orthonormality and linear independence	26
2.11	Schmidt orthonormalization method	26
2.12	Dimension of a vector space	27
2.13	Complete vector space	27
2.14	Convergence of a sequence of vectors in a vector space	27
2.14.1	An example of a vector space whose Cauchy sequence does not converge to a vector in the vector space	28
2.15	Hilbert space	29
2.16	Basis vectors in a Hilbert space	29
2.16.1	Finite dimensional space	29
2.16.2	Infinite dimensional vector space	30
2.16.3	Example	31
2.17	Dirac notation	32
2.17.1	Scalar product	32
2.17.2	Dual vector space	32
2.17.3	Bra notation for the vectors of V^*	32
2.17.4	Correspondence between kets and bras	32
2.17.5	The correspondence is anti-linear	33
2.17.6	Dirac notation for the scalar product	33

3	sheet-3 : Linear Vector Space (continued)	35
3.1	Operators in a Hilbert Space	35
3.2	Linear operators	35
3.3	Commutator of two operators	36
3.3.1	Some properties of commutators	36
3.4	Projection operator	37
3.5	Representation of vectors and operators	38
3.6	Matrix representation of ket and bra vectors	39
3.7	Representation of an operator in a basis	40
3.8	Matrix representation of the sum and product of two operators	41
3.9	Operators in a Hilbert space	43
3.9.1	Linear operators	44
3.9.2	Commutator of two operators	45
3.9.3	Some properties of commutators	45
3.10	Projection operator	45
3.11	Representation of vectors and operators	46
3.11.1	Matrix representation of kets	46
3.11.2	Matrix representation of bra vectors	48
3.11.3	Representation of operators	49
3.11.4	Matrix representation of the sum and product of two operators	50
4	sheet-4 : Linear Vector Space (continued)	53
4.1	Adjoint operator	53
4.2	Hermitian or Self-adjoint operator	54
4.3	Inverse operator	55
4.4	Unitary operator	55
4.5	Function of Operators	56
4.6	Adjoint Operator	56
4.6.1	Hermitian or self adjoint operator	57
4.6.2	Inverse operator	58
4.6.3	Unitary operator	58

4.7	Functions of operators	59
4.8	Change of basis	60
4.8.1	Change of representation for kets	61
4.9	Transformation of the matrix elements of an operator due to a change of basis	63
5	sheet-5 : Eigenvalue and Eigenvectors of Operators	65
5.1	Eigenvalues and Eigenvectors of a Hermitian Operator	65
5.2	Determination of eigenvalues and eigenvectors of a Hermitian Operator	67
5.3	Non-degenerate roots	68
5.4	Degenerate roots	69
5.5	Digonalization of a Hermitian Operator	70
5.6	Basis independence of the eigenvalues of an Operator	73
5.7	Infinite dimensional vector space	74
5.8	Completeness condition for the eigenvectors of a Hermitian Operator	75
5.9	Hermitian operator with continuous eigenvalue spectrum	77
5.9.1	Continuous Spectrum	77
5.10	Hermitian operator with continuous and discrete eigenvalue	79
5.11	Problems	79
6	sheet-6 : Coordinate and Momentum Representation	87
7	sheet-7 : Postulates of Quantum Mechanics	89
8	sheet-8 : Compatibility of Observable	91
8.1	Background Materials	91
8.2	Measurement of Two Observables in Quick Succession	93
8.3	Compatible Observables	96
8.3.1	Nondegenerate Case	96
8.3.2	Degenerate Case	98

8.4	Condition for Compatibility of Observables	101
8.5	Labeling of Quantum Mechanical Basis States	106
8.6	Background Materials	108
8.7	Measurement of Two Observables in Quick Succession	110
8.8	Compatible Observables	112
8.8.1	Nondegenerate Case	112
8.8.2	Degenerate Case	114
8.9	Condition for Compatibility of Observables	116
8.10	Labeling of Quantum Mechanical Basis States	122
9	sheet-9 : The Uncertainty Principle	125
10	sheet-10 : Harmonic Oscillator	127
11	sheet-11 : Quantum Dynamics	129
11.1	Quantum Dynamics	129
11.1.1	Schrodinger picture	129
11.1.2	Schrodinger equation	130
12	sheet-12 : The Path Integral Formulation of Quantum Theory	131
12.1	Background Materials	131
12.2	Propagator	134
12.3	Path Integral for the Propagator	135
12.4	Path Integral for a Free Particle	138
12.4.1	check of calculation	142
12.5	Derivation of the propagator for a free particle without using the path integral formula	142
12.6	The Classical Action	143
12.7	Discussion on Path Integral	144
12.7.1	Principle of Least Action	144

12.8	Discussion on The Phase of the Path Integral	146
12.9	Equivalence to the Schrödinger Equation	148
12.10	Potentials of the Form $V = a + bx + cx^2 + d\dot{x} + ex\ddot{x}$	151
12.10.1	Special Cases	152
13	sheet-13 : Potential Wells and Barriers	155
14	sheet-14 : Particle Motion in 3-Dimension	157
15	sheet-15 : The Hydrogen Atom	159
16	sheet-16 : Angular Momentum 1	161
16.1	Commutation Relations of Angular Momentum Operators	162
16.2	Definition of Angular Momentum Operators	164
16.2.1	Square of Angular Momentum	164
16.2.2	Angular Momentum of a System of Particles	165
16.3	Eigenvalue Spectrum of \hat{j}^2 and \hat{j}_z	165
16.4	Matrix Elements of Angular Momentum Operators in the $ jm\rangle$ Basis	169
16.4.1	$j = 1/2$ Representation	171
16.4.2	$j = 1$ Representation	172
17	sheet-17 : Angular Momentum 2: spin 1/2 system	175
17.1	The Spin Operators and Spin States of a Spin-1/2 Particle	175
17.1.1	Properties of Pauli Matrices	178
17.2	Matrix Representations of the Eigenstates of \hat{S}_x , \hat{S}_y and \hat{S}_z	179
17.3	Eigenvalues and Eigenvectors of $\hat{\vec{S}} \cdot \hat{\vec{n}}$	184
17.4	Non-relativistic Description of a Spin-1/2 Particle	186
17.4.1	Basis Vectors	186
17.4.2	State Vector	188
17.5	Electron in an External Magnetic Field. Pauli Equation	190
18	sheet-18 : Rotations and Angular Momentum	193

19	sheet-19 : Addition of Angular Momentum	195
20	sheet-20 : Time Independent Perturbation	197
20.1	Non-degenerate Perturbation Theory	198
20.1.1	First-Order correction to energy: $E_n^{(1)}$	199
20.1.2	First-Order correction to eigenstate	200
20.1.3	Second-Order Correction to Energy: $E_n^{(2)}$	202
21	sheet-21 : Variational Method	205
22	sheet-22 : WKB approximation	207
23	sheet-23 : Time Dependent Perturbation	209
24	sheet-24 : Scattering Theory	211
25	sheet-25 : Integral Equation for Scattering Theory	213
26	sheet-26 : Identical Particles	215
26.1	Indistinguishability of identical particles	215
26.1.1	The wave function of two identical particles	216
26.1.2	Spin Statistics Theorem	217
26.1.3	Statistics of Composite Particles	217
26.1.4	Operators for a system of identical particles	218
26.2	States of Noninteracting Identical Particles	218
26.2.1	Identical Fermions	219
26.2.2	Identical Bosons	220
26.2.3	Exchange Forces	222
26.3	The Helium Atom	225
26.4	Collision Between Identical Particles	230
26.4.1	Scattering of Two Identical Spinless Bosons	230
26.4.2	Partial wave expansion for the scattering amplitude of two spinless bosons ..	233
26.5	Scattering of Two Identical Spin-1/2 Fermions	234
26.5.1	Partial Wave Expansion of the Scattering Amplitude	236

.1	List of Standard Integrals	237
.2	Proofs	237

1. sheet-1 :Dirac Delta Function

Consider the function $D_\varepsilon(x)$ given by

$$D_\varepsilon(x) = \begin{cases} \frac{1}{\varepsilon} & \text{for } -\frac{\varepsilon}{2} \leq x \leq \frac{\varepsilon}{2} \\ 0 & \text{for } |x| > \frac{\varepsilon}{2} \end{cases} \quad (1.1)$$

where ε is a positive parameter. The plot of the function is shown in figure (1.1).

The integral of the function with respect to x is 1, i.e.,

$$\int_{-\infty}^{\infty} D_\varepsilon(x) dx = 1 \quad (1.2)$$

Now imagine making ε smaller. As we decrease ε , the function gets narrower and taller, but the integral of the function(i.e., the area under graph remains constant at the value 1). In the limit $\varepsilon \rightarrow 0$, the function $D_\varepsilon(x)$ collapses to a single point $x = 0$ and gets infinitely tall. So $\lim_{\varepsilon \rightarrow 0} D_\varepsilon(x)$ is not a function at all and the procedure of taking the limit is not justified.

However, we can make the limiting procedure meaningful if multiply $D_\varepsilon(x)$ by some well defined function $f(x)$, integrate over x and then take the limit $\varepsilon \rightarrow 0$. consider the integral

$$\int_{-\infty}^{\infty} D_\varepsilon(x) f(x) dx \quad (1.3)$$

where $f(x)$ is a well-defined function. If ε is significantly small, the variation of $f(x)$ over the effective integration interval $[-\varepsilon/2, \varepsilon/2]$ is negligible and $f(x)$ remains practically equal to $f(0)$, therefore,

$$\int_{-\infty}^{\infty} D_\varepsilon(x) f(x) dx \simeq f(0) \int_{-\infty}^{\infty} D_\varepsilon(x) dx = f(0) \quad (1.4)$$

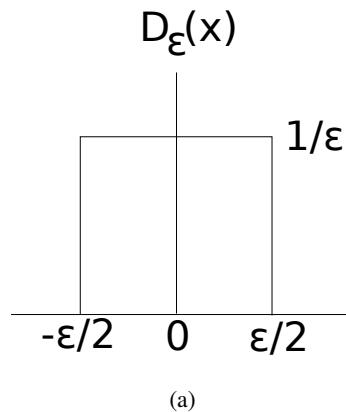


Figure 1.1: Dirac Delta Function

The smaller the value of ε , the better the approximation. In the limit $\varepsilon \rightarrow 0$, the above equation is exact

$$\lim_{\varepsilon \rightarrow 0} \int_{-\infty}^{\infty} D_{\varepsilon}(x) f(x) dx = f(0) \quad (1.5)$$

Now, we define the delta function by the relation

$$\int_{-\infty}^{\infty} \delta(x) f(x) dx \stackrel{def}{=} \lim_{\varepsilon \rightarrow 0} \int_{-\infty}^{\infty} D_{\varepsilon}(x) f(x) dx = f(0) \quad (1.6)$$

This equation is valid for any function $f(x)$ defined at the origin. More generally, $\delta(x - x_0)$ is defined as,

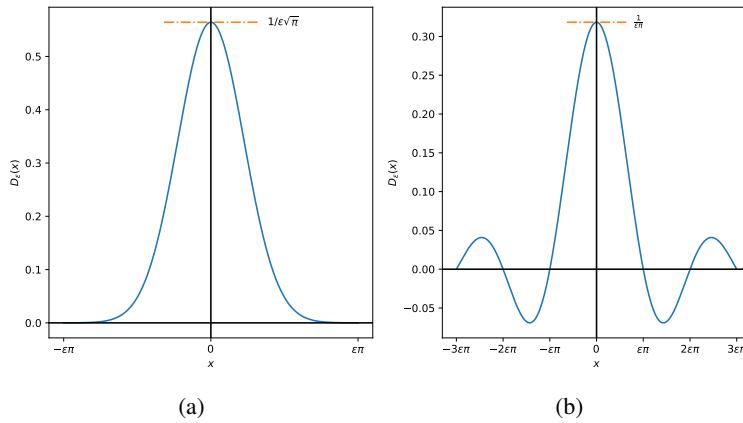
$$\int_{-\infty}^{\infty} \delta(x - x_0) f(x) dx = f(x_0) \quad (1.7)$$

Actually, the integral notation $\int_{-\infty}^{\infty} \delta(x) f(x) dx$ is not justified because $\delta(x)$ is not really a function. Physically, there is no problem since it becomes impossible to distinguish between $D_{\varepsilon}(x)$ and $\delta(x)$ as soon as ε becomes negligible compared to all distances involved in a physical problem. Whenever a mathematical difficulty might arise, all we need to do is to assume that $\delta(x)$ is actually $D_{\varepsilon}(x)$ with ε extremely small but not strictly zero.

Formally, we can express $\delta(x)$ as a limit of a square of proper functions :

$$\lim_{\varepsilon \rightarrow 0} D_{\varepsilon}(x) \equiv \delta(x) \quad (1.8)$$

Here $D_{\varepsilon}(x)$, which is a proper function of x is called the representation of the delta function. One representation is the "square function" given at the beginning. The representation is not unique. There are other functions which approach the delta function when appropriate limits are taken.

Figure 1.2: Plot of function $D_\epsilon(x)$ for (a) equation (1.9) (b) and equation (1.12)

1.1 Other representation of delta function

1. Consider the function

$$D_\epsilon(x) = \frac{1}{\epsilon\sqrt{\pi}} e^{-x^2/\epsilon^2} \quad (\epsilon > 0) \quad (1.9)$$

For each value of the parameter ϵ , this function satisfies

$$\int_{-\infty}^{\infty} D_\epsilon(x) dx = 1 \quad (1.10)$$

When plotted against x , the function has a peak at the origin. The peak has a height of $\frac{1}{\epsilon\sqrt{\pi}}$ and a width of order ϵ (exactly how the width is defined doesn't matter). So if ϵ is allowed to become very small, the peak becomes very tall and very narrow. Outside the peak the function becomes extremely small. Thus we have

$$\delta(x) = \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon\sqrt{\pi}} e^{-x^2/\epsilon^2} \quad (1.11)$$

2. Consider another function

$$D_\epsilon(x) = \frac{1}{\pi} \frac{\sin(x/\epsilon)}{x} \quad (\epsilon > 0) \quad (1.12)$$

For any value of the parameter ϵ we have

$$\int_{-\infty}^{\infty} D_\epsilon(x) dx = 1 \quad (1.13)$$

A plot of the function $D_\epsilon(x)$ shows that it has the value $\frac{1}{\epsilon\pi}$ at $x = 0$ and it oscillates with decreasing amplitude as $|x|$ increases. The width of the central maxima is of the order of ϵ and the period of oscillation with respect to x is $2\pi\epsilon$.

Thus the limit of this function as $\varepsilon \rightarrow 0$ has all the properties of the delta function : it becomes infinitely large at $x = 0$, it has unit integral, and infinitely rapid oscillations as $|x|$ increases means that the entire contribution to an integral containing this function comes from an infinitesimal neighborhood of $x = 0$.

We can therefore write,

$$\delta(x) = \lim_{\varepsilon \rightarrow 0} \frac{1}{\pi} \frac{\sin(x/\varepsilon)}{x} \quad (1.14)$$

3. We can also show that

$$\delta(x) = \lim_{\varepsilon \rightarrow 0} \frac{1}{2\varepsilon} e^{-|x|/\varepsilon} \quad (1.15)$$

$$\delta(x) = \lim_{\varepsilon \rightarrow 0} \frac{1}{\pi} \frac{\varepsilon}{x^2 + \varepsilon^2} \quad (1.16)$$

$$\delta(x) = \lim_{\varepsilon \rightarrow 0} \frac{\varepsilon}{\pi} \frac{\sin^2(x/\varepsilon)}{x^2} \quad (1.17)$$

1.2 Properties of the delta function

It is important to note that, because of its singular @@@@, the δ function cannot be the end result of a calculation, and has meaning only so long as a subsequent integral over its argument is carried out. With this understanding we can write down some relations between delta functions.

Property 1 The delta function is a even function

$$\delta(-x) = \delta(x) \quad (1.18)$$

Property 2

$$x\delta(x) = 0 \quad (1.19)$$

Property 3

$$\delta(ax) = \frac{1}{|a|} \delta(x) \quad (1.20)$$

proof: Consider the integral

$$I = \int_{-\infty}^{\infty} \delta(ax)f(x)dx \quad (1.21)$$

Since the delta function is even in its argument, it doesn't matter if we replace a by $|a|$ in the argument. Thus

$$I = \int_{-\infty}^{\infty} \delta(|a|x)f(x)dx \quad (1.22)$$

Making the change in variable $y = |a|x$ we have,

$$\begin{aligned} I &= \frac{1}{|a|} \int_{-\infty}^{\infty} \delta(y)f(y/|a|)dx \\ &= \frac{1}{|a|} f(0) \end{aligned}$$

or

$$\int_{-\infty}^{\infty} \delta(ax)f(x)dx = \frac{1}{|a|} \int_{-\infty}^{\infty} \delta(x)f(x)dx \quad (1.23)$$

or

$$\delta(ax) = \frac{1}{|a|} \delta(x) \quad (1.24)$$

Property 4 More generally

$$\delta(\phi(x)) = \sum_i \frac{\delta(x - x_i)}{\left| \frac{\partial \phi}{\partial x} \right|_{x_i}} \quad (1.25)$$

where the sum sums over the x_i 's which are simple roots of $\phi(x)$.

proof : let x_1, x_2, \dots, x_N be the simple roots of $\phi(x)$,

In the neighborhood of any one of the simple roots x_i , we can write @@@@ @@@@ @

$$\phi(x) = (x - x_i)\psi(x) \quad (1.26)$$

or ???????

$$\phi(x) = (x - x_i)\psi(x_i) \quad (1.27)$$

where $\psi(x_i) \neq 0$. We have

$$\psi(x_i) = \left| \frac{\partial \phi(x)}{\partial x} \right|_{x=x_i} \quad (1.28)$$

Now, consider the integral

$$\begin{aligned} I &= \int_{-\infty}^{\infty} \delta(\phi(x))f(x)dx \\ &= \sum_{i=1}^N \int_{x_i-\epsilon}^{x_i+\epsilon} \delta[(x - x_i)\psi(x_i)]f(x)dx \\ &= \sum_{i=1}^N \frac{1}{|\psi(x_i)|} \int_{x_i-\epsilon}^{x_i+\epsilon} \delta(x - x_i)f(x)dx \\ &= \sum_{i=1}^N \frac{1}{\left| \frac{\partial \phi(x)}{\partial x} \right|_{x=x_i}} \int_{-\infty}^{\infty} \delta(x - x_i)f(x)dx \\ &= \sum_{i=1}^N \frac{1}{\left| \frac{\partial \phi}{\partial x} \right|_{x=x_i}} f(x_i) \end{aligned}$$

The above result is obtained if we write

$$\delta(\phi(x)) = \sum_{i=1}^N \frac{\delta(x - x_i)}{\left| \frac{\partial \phi}{\partial x} \right|_{x=x_i}} \quad (1.29)$$

Property 5 A frequently used example of the above result is

$$\delta(x^2 - a^2) = \frac{1}{2a} \delta(x - a) + \frac{1}{2a} \delta(x + a) \quad (a > 0) \quad (1.30)$$

Here

$$\phi(x) = x^2 - a^2 = (x-a)(x+a) \quad (1.31)$$

The two simple roots of $\phi(x)$ are at $x=a$ and $x=-a$. Now

$$\left| \frac{\partial \phi}{\partial x} \right|_{x=a} = |2x|_{x=a} = 2a \quad (1.32)$$

$$\left| \frac{\partial \phi}{\partial x} \right|_{x=-a} = |-2x|_{x=-a} = 2a \quad (1.33)$$

\therefore The above result follows.

Property 6

$$f(x)\delta(x-a) = f(a)\delta(x-a) \quad (1.34)$$

Property 7

$$\int \delta(x-y)\delta(y-a)dy = \delta(x-a) \quad (1.35)$$

1.2.1 Notes

Note : 1 We have the identity

$$x\delta(x) = 0 \quad (1.36)$$

The converse is also true and it can be shown that the equation

$$xu(x) = 0 \quad (1.37)$$

has the general solution

$$u(x) = c\delta(x) \quad (1.38)$$

Note : 2 We will now prove an identity which is particularly useful in Quantum Mechanics.

$$\lim_{\epsilon \rightarrow 0^+} \int_{-\infty}^{\infty} \frac{1}{x \pm i\epsilon} f(x) dx = \rho \int_{-\infty}^{\infty} \frac{dx}{x} f(x) \mp i\pi f(0) \quad (1.39)$$

or in short

$$\lim_{\epsilon \rightarrow 0^+} \frac{1}{x \pm i\epsilon} = \rho\left(\frac{1}{x}\right) \mp i\pi\delta(x) \quad (1.40)$$

Where it is understood that the second of these two equations have meaning only within an integral.

The symbol ρ means principle part of an integral where the integral has a single pole. The principle part is defined as

$$\rho \int_{-A}^B \frac{dx}{x} f(x) dx = \lim_{\eta \rightarrow 0^+} \left[\int_{-A}^{-\eta} + \int_{\eta}^B \right] \frac{dx}{x} f(x) \quad (1.41)$$

proof :

$$\frac{1}{x \pm i\epsilon} = \frac{x \mp i\epsilon}{x^2 + \epsilon^2} = \frac{x}{x^2 + \epsilon^2} \mp \frac{i\epsilon}{x^2 + \epsilon^2} \quad (1.42)$$

Now we have

$$\begin{aligned} \lim_{\epsilon \rightarrow 0^+} \frac{1}{\pi} \frac{\epsilon}{x^2 + \epsilon^2} &= \delta(x) \\ \lim_{\epsilon \rightarrow 0^+} (\mp) i \frac{\epsilon}{x^2 + \epsilon^2} &= \mp i\pi\delta(x) \end{aligned} \quad (1.43)$$

Now consider the first term on the right hand side of equation 1.43. We multiply this term by a function $f(x)$ which is regular at the origin and then integrate over x . We get

$$\lim_{\epsilon \rightarrow 0^+} \int_{-\infty}^{\infty} \frac{xf(x)}{x^2 + \epsilon^2} dx = \lim_{\epsilon \rightarrow 0^+} \left[\lim_{\eta \rightarrow 0^+} \left[\int_{-\infty}^{-\eta} \frac{xf(x)}{x^2 + \epsilon^2} + \int_{-\eta}^{\eta} \frac{xf(x)}{x^2 + \epsilon^2} + \int_{\eta}^{\infty} \frac{xf(x)}{x^2 + \epsilon^2} dx \right] \right] \quad (1.44)$$

Note that we take the limit over η first and then we take the limit over ϵ . Consider now the second integral above

$$\lim_{\eta \rightarrow 0^+} \int_{-\eta}^{+\eta} \frac{xf(x)}{x^2 + \epsilon^2} dx = f(0) \lim_{\eta \rightarrow 0^+} \frac{1}{2} [\ln(x^2 + \epsilon^2)]_{x=-\eta}^{x=\eta} = 0 \quad (1.45)$$

If we now reverse the order of the evaluation of limits in equation 1.44, the $\epsilon \rightarrow 0$ limit causes no difficulties in the other two integrals. Thus we have

$$\begin{aligned} &\lim_{\epsilon \rightarrow 0^+} \int_{-\infty}^{\infty} \frac{xdx}{x^2 + \epsilon^2} f(x) \\ &= \lim_{\eta \rightarrow 0^+} \lim_{\epsilon \rightarrow 0^+} \left[\int_{-\infty}^{-\eta} + \int_{\eta}^{\infty} \right] \frac{xdx}{x^2 + \epsilon^2} f(x) \\ &= \lim_{\eta \rightarrow 0^+} \left[\int_{-\infty}^{-\eta} + \int_{\eta}^{\infty} \right] \frac{dx}{x} f(x) \\ &= \rho \int_{-\infty}^{\infty} \frac{1}{x} f(x) dx \end{aligned}$$

This establishes the identity.

1.3 Derivatives of the delta function

One may define the derivative $\delta'(x)$ of the delta function. When ϵ is small, the derivative of $D_\epsilon(x)$ has two peaks close to the origin, one peak is positive and the other is negative as drawn in the figure below

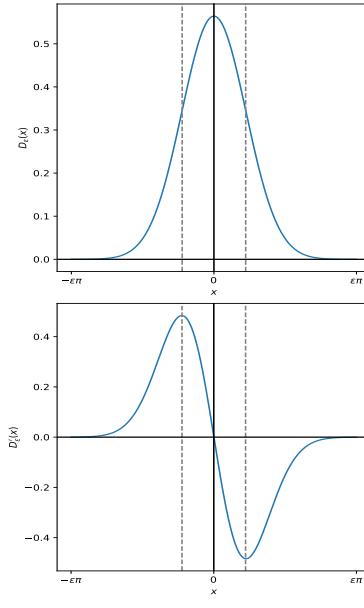
As $\epsilon \rightarrow 0$, each of these peaks becomes very narrow and very tall, and two peaks each approach very close to the origin.

Now the integration by parts gives

$$\int_{-\infty}^{\infty} dx D'_\epsilon(x) f(x) = [D_\epsilon(x) f(x)]_{-\infty}^{\infty} - \int_{-\infty}^{\infty} dx D_\epsilon(x) f'(x) \quad (1.46)$$

Because $D_\epsilon(x)$ tends to zero as $x \rightarrow \pm\infty$, the first term on the right hand side vanishes unless $f(x)$ @@@@ @@@@ violently at infinity. So by letting $\epsilon \rightarrow 0$, we arrive at the definition of $\delta'(x)$

$$\int_{-\infty}^{\infty} \delta'(x) f(x) dx = - \int_{-\infty}^{\infty} \delta(x) f'(x) dx = -f'(0) \quad (1.47)$$



(a)

Figure 1.3: (a) Derivative of Delta Function

From this we immediately get

$$x\delta'(x) = -\delta(x) \quad (1.48)$$

Conversely, it can be shown that the general solution of the equation

$$xu(x) = \delta(x) \quad (1.49)$$

can be written as

$$u(x) = -\delta'(x) + c\delta(x) \quad (1.50)$$

Where the second term arises from the homogeneous equation $x\delta(x) = 0$. From equation 1.47 it also follows that

$$\delta'(-x) = -\delta'(x) \quad (1.51)$$

The n^{th} order derivative of $\delta(x)$ can be defined in the same way. We find

$$\int_{-\infty}^{\infty} \delta^{(n)}(x)f(x)dx = (-1)^n f(0) \quad (1.52)$$

We can prove the following properties:

$$\delta^{(m)}(x) = (-1)^m \delta^{(m)}(-x) \quad (1.53)$$

$$x^{m+1} \delta^{(m)}(x) = 0 \quad (1.54)$$

$$x\delta^{(m)}(x) = -m\delta^{(m-1)}(x) \quad (1.55)$$

1.4 Integration of the delta function

Consider the indefinite integral

$$\Theta_\epsilon(x) = \int_{-\infty}^x D_\epsilon(y) dy \quad (1.56)$$

A graph of $\Theta_\epsilon(x)$ vs x is shown below

As $\epsilon \rightarrow 0$, the step in the function $\Theta_\epsilon(x)$ gets progressively steeper, until, finally, the function changes abruptly from 0 to 1 at $x = 0$. Thus taking the limit $\epsilon \rightarrow 0$ in equation 1.56 we have

$$\Theta(x) = \int_{-\infty}^x \delta(x) dx \quad (1.57)$$

Where

$$\Theta(x) = \begin{cases} 1 & \text{for } x > 0 \\ 0 & \text{for } x < 0 \end{cases} \quad (1.58)$$

If we differentiate equation 1.57 with respect to x ,

we get

$$\frac{d\Theta(x)}{dx} = \delta(x) \quad (1.59)$$

1.5 Three - dimensional delta function

$$\delta(\vec{r}) \stackrel{\text{def}}{=} \delta(x)\delta(y)\delta(z) \quad (1.60)$$

In other words, $\delta(\vec{r})$ is zero if any of the coordinates x, y, z is not equal to zero and $\delta(\vec{r})$ tends to infinity at the origin, i.e., when $x = 0, y = 0, z = 0$, such that

$$\int_{\text{volume}} \delta(\vec{r}) d^3r = 1 \quad (1.61)$$

if the volume of the integration contains the origin. We also have

$$\int \delta(\vec{r}) f(\vec{r}) d^3r = f(0) \quad (1.62)$$

where again the volume of the integration includes the origin. **Note:**

$$\delta(\vec{r} - \vec{r}') = \delta(x - x')\delta(y - y')\delta(z - z') \quad (1.63)$$

$$\int_v \delta(\vec{r} - \vec{r}') d^3r = 1 \quad (1.64)$$

where the volume of integration includes the point \vec{r}' . Otherwise the integral is zero.

$$\int_v \delta(\vec{r} - \vec{r}') f(\vec{r}) d^3r = f(\vec{r}') \quad (1.65)$$

if V includes the point \vec{r}' .

A useful formula Consider the integral

$$\begin{aligned}
 \int_{-\infty}^{\infty} e^{ikx} dx &= \lim_{L \rightarrow \infty} \int_{-L}^L e^{ikx} dx \\
 &= \lim_{L \rightarrow \infty} \frac{1}{ik} (e^{ikL} - e^{-ikL}) \\
 &= \lim_{L \rightarrow \infty} \frac{2}{k} \left(\frac{e^{ikL} - e^{-ikL}}{2i} \right) \\
 &= \lim_{L \rightarrow \infty} \frac{2}{k} \sin(kL) \\
 &= 2\pi \lim_{L \rightarrow \infty} \frac{\sin(kL)}{\pi k} \\
 &= 2\pi \delta(k)
 \end{aligned} \tag{1.66}$$

using

$$\lim_{\epsilon \rightarrow 0} \frac{\sin(x/\epsilon)}{\pi x} = \delta(x) \tag{1.67}$$

Thus

$$\int_{-\infty}^{\infty} e^{ikx} dx = 2\pi \delta(k) \tag{1.68}$$

In equation 1.68 if we integrate with respect to k , we would have $\delta(x)$ on the right hand side,

$$\int_{-\infty}^{\infty} e^{ikx} dk = 2\pi \delta(x) \tag{1.69}$$

Also note that in equation 1.68 we are integrating over its full range of values. Making a change of variable $x \rightarrow -x$ does not change the value of the integral. Hence we also have

$$\int_{-\infty}^{\infty} e^{-ikx} dx = 2\pi \delta(k) \tag{1.70}$$

Similarly, in equation 1.69 making the change $k \rightarrow -k$, doesn't change the value of the integral. So we could also write

$$\int_{-\infty}^{\infty} e^{-ikx} dk = 2\pi \delta(x) \tag{1.71}$$

Thus in summary

$$\int_{\pm\infty}^{\infty} e^{-ikx} dx = 2\pi \delta(k) \int_{\pm\infty}^{\infty} e^{-ikx} dk = 2\pi \delta(x) \tag{1.72}$$

In three dimensions

$$\int_{\text{all space}} e^{\pm i\vec{k} \cdot \vec{r}} d^3 r = (2\pi)^3 \delta(\vec{k}) \quad (1.73)$$

$$\int_{\text{all space}} e^{\pm i(\vec{k} - \vec{k}') \cdot \vec{r}} d^3 r = (2\pi)^3 \delta(\vec{k} - \vec{k}') \quad (1.74)$$

$$\int_{\text{all space}} e^{\pm i\vec{k} \cdot (\vec{r} - \vec{r}')} d^3 r = (2\pi)^3 \delta(\vec{r} - \vec{r}') \quad (1.75)$$

(1.76)

1.6 Fourier Transformation

We can always express a function $f(x)$ in the form

$$f(x) = \int_{-\infty}^{\infty} e^{ikx} \tilde{f}(k) dk \quad (1.77)$$

where $\tilde{f}(k)$ is a function of k , called the fourier transform of $f(x)$. From eqnarray 1.77 we can write

$$\begin{aligned} \int_{-\infty}^{\infty} e^{-ik'x} f(x) dx &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{i(k-k')x} \tilde{f}(k) dk dk' \\ &= 2\pi \int_{-\infty}^{\infty} \delta(k - k') \tilde{f}(k) dk \\ &= 2\pi \tilde{f}(k') \end{aligned} \quad (1.78)$$

Thus

$$\tilde{f}(k) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-ikx} f(x) dx \quad (1.79)$$

Thus functions $f(x)$ and $\tilde{f}(k)$ are Fourier transform of each other. We can write eqnarray 1.77 and 1.79 in a more symmetrical fashion as follows:

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{ikx} \tilde{f}(k) dk \quad (1.80)$$

$$\tilde{f}(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-ikx} f(x) dx \quad (1.81)$$

In three dimension, we can write

$$f(\vec{r}) = \frac{1}{(2\pi)^{3/2}} \int_{\text{all } k\text{-space}} e^{i\vec{k} \cdot \vec{r}} \tilde{f}(\vec{k}) d^3 k \quad (1.82)$$

multiplying eqnarray 1.82 by $e^{-i\vec{k}' \cdot \vec{r}}$ and integrating over \vec{r} , we have

$$\begin{aligned} \int_{\text{all space}} e^{-i\vec{k}' \cdot \vec{r}} f(\vec{r}) d^3 r &= \frac{1}{(2\pi)^{3/2}} \int d^3 r \int d^3 k e^{i(\vec{k} - \vec{k}') \cdot \vec{r}} \tilde{f}(\vec{k}) \\ &= \frac{1}{(2\pi)^{3/2}} \int d^3 k (2\pi)^3 \delta(\vec{k} - \vec{k}') \tilde{f}(\vec{k}) \\ &= \frac{1}{(2\pi)^{-3/2}} \tilde{f}(\vec{k}') \end{aligned} \quad (1.83)$$

Therefore

$$\tilde{f}(\vec{k}) = \frac{1}{(2\pi)^{3/2}} \int e^{-i\vec{k}\cdot\vec{r}} f(\vec{r}) d^3 r \quad (1.84)$$

Thus we can write

$$f(\vec{r}) = \frac{1}{(2\pi)^{3/2}} \int e^{i\vec{k}\cdot\vec{r}} \tilde{f}(\vec{k}) d^3 k \quad (1.85)$$

$$\tilde{f}(\vec{k}) = \frac{1}{(2\pi)^{3/2}} \int e^{-i\vec{k}\cdot\vec{r}} f(\vec{r}) d^3 r \quad (1.86)$$

Parseval's Identity:

We can now prove the important identity

$$\int |f(\vec{r})|^2 d^3 r = \int |\tilde{f}(\vec{k})|^2 d^3 k \quad (1.87)$$

proof:

$$\begin{aligned} \int |f(\vec{r})|^2 d^3 r &= \int f(\vec{r}) f^*(\vec{r}) d^3 r \\ &= \int d^3 r \frac{1}{(2\pi)^{3/2}} \int e^{i\vec{k}\cdot\vec{r}} \tilde{f}(\vec{k}) d^3 k \frac{1}{(2\pi)^{3/2}} \int e^{-i\vec{k}'\cdot\vec{r}} \tilde{f}^*(\vec{k}') d^3 k' \\ &= \frac{1}{(2\pi)^3} \int d^3 k d^3 k' \tilde{f}(\vec{k}) \tilde{f}^*(\vec{k}') \int d^3 r e^{i(\vec{k}-\vec{k}')\cdot\vec{r}} \\ &= \frac{1}{(2\pi)^3} \int d^3 k d^3 k' \tilde{f}(\vec{k}) \tilde{f}^*(\vec{k}') (2\pi)^3 \delta(\vec{k} - \vec{k}') \\ &= \int d^3 k \tilde{f}(\vec{k}) \tilde{f}^*(\vec{k}) \\ &= \int |\tilde{f}(\vec{k})|^2 d^3 k \end{aligned} \quad (1.88)$$

Lecture 1 Dirac Delta Function

1.7 Definition of Dirac Delta function and its Representations

Consider the function $D_\varepsilon(x)$ defined by

$$D_\varepsilon(x) = \begin{cases} 1/\varepsilon & \text{for } -\varepsilon/2 \leq x \leq \varepsilon/2 \\ 0 & \text{for } |x| > \varepsilon/2 \end{cases} \quad (1.89)$$

where ε is a small parameter. The plot of the function is shown below:

The integral of the function with respect to x is unity, i.e.,

$$\int_{-\infty}^{\infty} D_\varepsilon(x) dx = 1. \quad (1.90)$$

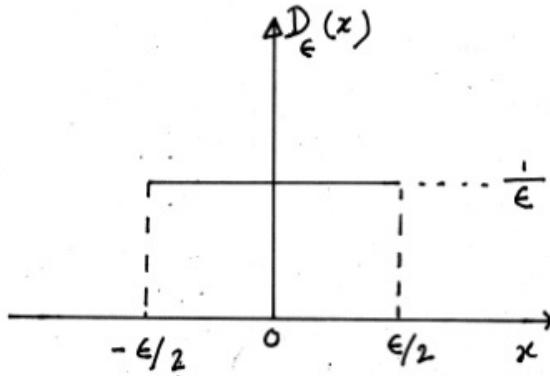


Figure 1.4: Plot of the function defined in Eq. (1.89)

Now, imagine making ϵ smaller. As we decrease ϵ , the function gets narrower and taller, but the integral of the function (i.e., the area under the graph of the function) remains constant at the value 1. In the limit $\epsilon \rightarrow 0$, the function $D_\epsilon(x)$ collapses to a single point, namely $x = 0$, and gets infinitely tall. So, $\lim_{\epsilon \rightarrow 0} D_\epsilon(x)$ is not a function at all and the procedure of taking the limit $\epsilon \rightarrow 0$ is not justified.

However, we can make the limiting procedure meaningful if we multiply $D_\epsilon(x)$ by some well-defined function $f(x)$, integrate over x and then take the limit $\epsilon \rightarrow 0$. Consider the integral

$$\int_{-\infty}^{\infty} D_\epsilon(x) f(x) dx$$

where $f(x)$ is a well-defined function. If ϵ is sufficiently small, the variation of $f(x)$ over the effective integration interval $[-\epsilon/2, \epsilon/2]$ is negligible and $f(x)$ remains practically equal to $f(0)$. Therefore,

$$\int_{-\infty}^{\infty} D_\epsilon(x) f(x) dx \approx f(0) \int_{-\infty}^{\infty} D_\epsilon(x) dx = f(0). \quad (1.91)$$

The smaller the value of ϵ , the better the approximation. In the limit $\epsilon \rightarrow 0$, the above equation is exact:

$$\lim_{\epsilon \rightarrow 0} \int_{-\infty}^{\infty} D_\epsilon(x) f(x) dx = f(0). \quad (1.92)$$

Now, we define the Dirac Delta Function $\delta(x)$ as

$$\int_{-\infty}^{\infty} \delta(x) f(x) dx \stackrel{\text{def}}{=} \lim_{\epsilon \rightarrow 0} \int_{-\infty}^{\infty} D_\epsilon(x) f(x) dx = f(0) \quad (1.93)$$

This equation is valid for any function $f(x)$ defined at the origin. More generally, $\delta(x - x_0)$ is defined as

$$\int_{-\infty}^{\infty} \delta(x - x_0) f(x) dx = f(x_0). \quad (1.94)$$

Actually, the integral notation

$$\int_{-\infty}^{\infty} \delta(x) f(x) dx$$

is not justified because $\delta(x)$ is not really a function. Physically, there is no problem since it becomes impossible to distinguish between $D_\varepsilon(x)$ and $\delta(x)$ as soon as ε becomes negligible compared to all distances involved in a physical problem. Whenever a mathematical difficulty might arise, all we need to do is to assume that $\delta(x)$ is actually $D_\varepsilon(x)$ with ε extremely small but not strictly zero.

Formally, we can express $\delta(x)$ as a limit of a sequence of proper functions:

$$\lim_{\varepsilon \rightarrow 0} D_\varepsilon(x) \equiv \delta(x). \quad (1.95)$$

Here $D_\varepsilon(x)$, which is a proper function of x , is called the representation of the delta function. The representation we have discussed so far is the “square function” given in Eq. (1.89). The representation is not unique. There are other functions which approach the delta function when appropriate limits are taken.

Gaussian Representation

Consider the function

$$D_\varepsilon(x) = \frac{1}{\varepsilon\sqrt{\pi}} e^{-x^2/\varepsilon^2} \quad (\varepsilon > 0). \quad (1.96)$$

For each value of the parameter ε , this function satisfies

$$\int_{-\infty}^{\infty} D_\varepsilon(x) dx = 1. \quad (1.97)$$

This is the normalized Gaussian function whose plot is shown in the figure below.

The Gaussian function has a peak at the origin. The peak has a height $1/\varepsilon\sqrt{\pi}$ and a width of order ε (exactly how the width is defined doesn't matter). So if ε is allowed to be very small, the peak becomes very tall and very narrow. Outside the peak the function becomes extremely small. Thus we have

$$\delta(x) = \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon\sqrt{\pi}} e^{-x^2/\varepsilon^2}. \quad (1.98)$$

Mathematical Notes:

We have the integral

$$\int_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\pi}.$$

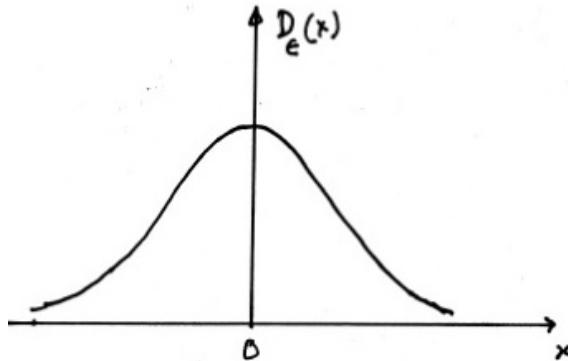


Figure 1.5: Plot of the Gaussian function defined in Eq. (1.96)

Now let us consider the following integral

$$I = \int_{-\infty}^{\infty} e^{-b^2 x^2 + ax} dx.$$

First, write

$$b^2 x^2 + ax = \left(bx + \frac{a}{2b}\right)^2 - \frac{a^2}{4b^2}.$$

Therefore,

$$\begin{aligned} I &= e^{a^2/4b^2} \int_{-\infty}^{\infty} e^{-(bx+a/2b)^2} dx \\ &= e^{a^2/4b^2} (1/b) \int_{-\infty}^{\infty} e^{-z^2} dz \quad (z = bx + \frac{a}{2b}) \\ &= e^{a^2/4b^2} \frac{\sqrt{\pi}}{b}. \end{aligned}$$

Damped Sinusoidal Representation of the Delta Function

Consider another function

$$D_\epsilon(x) = \frac{1}{\pi} \frac{\sin(x/\epsilon)}{x} \quad (\epsilon > 0). \quad (1.99)$$

A plot of the function is shown below. The function $D_\epsilon(x)$ has the value $1/\epsilon\sqrt{\pi}$ at $x = 0$ and it oscillates with decreasing amplitude as $|x|$ increases. The width of the central maximum is of the order of ϵ and the period of oscillations with respect to x is $2\pi\epsilon$.

For any value of ϵ we have

$$\int_{-\infty}^{\infty} D_\epsilon(x) dx = 1. \quad (1.100)$$

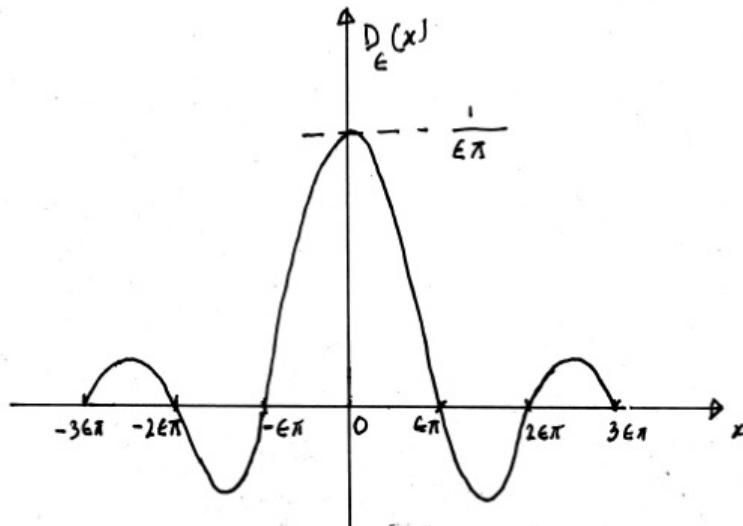


Figure 1.6: Plot of the function defined in Eq. (1.99)

Thus, the limit of the function as $\epsilon \rightarrow 0$ has all the properties of the delta function: it becomes infinitely large at $x = 0$, and infinitely rapid oscillations as $|x|$ increases means that the entire contribution to an integral containing this function comes from an infinitesimal neighborhood of $x = 0$. We can therefore write

$$\delta(x) = \lim_{\epsilon \rightarrow 0} \frac{1}{\pi} \frac{\sin(x/\epsilon)}{x} \quad (1.101)$$

Other representation of the Delta Function

We can also show that

$$\delta(x) = \lim_{\epsilon \rightarrow 0} \frac{1}{2\epsilon} e^{-|x|/\epsilon} \quad (1.102)$$

$$\delta(x) = \lim_{\epsilon \rightarrow 0} \frac{1}{\pi} \frac{\epsilon}{x^2 + \epsilon^2} \quad (1.103)$$

$$\delta(x) = \lim_{\epsilon \rightarrow 0} \frac{\epsilon}{\pi} \frac{\sin^2(x/\epsilon)}{x^2}. \quad (1.104)$$

1.8 Properties of the Delta Function

It is important to note that, because of its singular character, the δ -function cannot be the end result of a calculation, and has meaning only so long as a subsequent integral over its argument is carried out. With this understanding we can write down some relations between delta functions:

1. The delta function is an even function, i.e.,

$$\delta(-x) = \delta(x). \quad (1.105)$$

2.

$$x\delta(x) = 0. \quad (1.106)$$

3.

$$\delta(ax) = \frac{1}{|a|} \delta(x). \quad (1.107)$$

Proof:

Consider the integral

$$I = \int_{-\infty}^{\infty} \delta(ax)f(x)dx.$$

Since the delta function is an even function it doesn't matter if we replace a by $|a|$ in the argument. Thus

$$I = \int_{-\infty}^{\infty} \delta(|a|x)f(x)dx.$$

Making the change of variable

$$y = |a|x$$

we have

$$I = \frac{1}{|a|} \int_{-\infty}^{\infty} \delta(y)f(y/|a|)dy = \frac{1}{|a|} f(0),$$

or,

$$\int_{-\infty}^{\infty} \delta(ax)f(x)dx = \frac{1}{|a|} \int_{-\infty}^{\infty} \delta(x)f(x)dx,$$

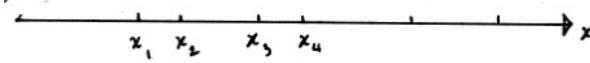
i.e.,

$\delta(ax) = \frac{1}{|a|} \delta(x)$

4. More generally, we have

$$\delta(\phi(x)) = \sum_i \frac{\delta(x - x_i)}{\left| \frac{d\phi}{dx} \right|_{x_i}} \quad (1.108)$$

where the sum runs over the x_i 's which are the simple roots of $\phi(x)$.

Figure 1.7: Simple roots of $\phi(x)$.**Proof:**

Let x_1, x_2, \dots, x_N be the simple roots of $\phi(x)$ (figure below):

In the neighborhood of any of the simple roots x_i , we can write

$$\phi(x) = (x - x_i)\psi(x)$$

where $\psi(x_i) \neq 0$. We have

$$\psi(x_i) = \left. \frac{\partial \phi(x)}{\partial x} \right|_{x=x_i}.$$

Now, consider the integral

$$\begin{aligned} \int_{-\infty}^{\infty} \delta(\phi(x)) f(x) dx &= \sum_{i=1}^N \int_{x_i-\varepsilon}^{x_i+\varepsilon} \delta[(x-x_i)\psi(x_i)] f(x) dx \\ &= \sum_{i=1}^N \frac{1}{|\psi(x_i)|} \int_{x_i-\varepsilon}^{x_i+\varepsilon} \delta(x-x_i) f(x) dx \\ &= \sum_{i=1}^N \frac{1}{\left| \frac{\partial \phi}{\partial x} \right|_{x=x_i}} \int_{-\infty}^{\infty} \delta(x-x_i) f(x) dx \\ &= \sum_{i=1}^N \frac{1}{\left| \frac{\partial \phi}{\partial x} \right|_{x=x_i}} f(x_i) \end{aligned}$$

The above result is obtained if we write

$$\delta(\phi(x)) = \sum_{i=1}^N \frac{\delta(x-x_i)}{\left| \frac{\partial \phi}{\partial x} \right|_{x=x_i}}. \quad \text{Proved.} \quad (1.109)$$

5. A frequently used example of the above result is

$$\delta(x^2 - a^2) = \frac{1}{2a} \delta(x-a) + \frac{1}{2a} \delta(x+a), \quad (a > 0). \quad (1.110)$$

Here

$$\phi(x) = x^2 - a^2 = (x-a)(x+a).$$

The two simple roots of $\phi(x)$ are at $x=a$ and $x=-a$. Now

$$\left| \frac{\partial \phi}{\partial x} \right|_{x=a} = |2x|_{x=a} = 2a$$

and

$$\left| \frac{\partial \phi}{\partial x} \right|_{x=-a} = |2x|_{x=-a} = 2a.$$

Therefore, the above result follows.

6.

$$f(x)\delta(x-a) = f(a)\delta(x-a).$$

7.

$$\int \delta(x-y)\delta(y-a)dy = \delta(x-a).$$

Note 1:

We have the identity

$$x\delta(x) = 0.$$

The converse is also true and it can be shown that the equation

$$xu(x) = 0$$

has the general solution

$$u(x) = c\delta(x).$$

Note 2:

We will now prove an identity which is particularly useful in quantum mechanics:

$$\lim_{\epsilon \rightarrow 0^+} \int_{-\infty}^{\infty} \frac{1}{x \pm i\epsilon} f(x) dx = \mathcal{P} \int_{-\infty}^{\infty} \frac{dx}{x} f(x) \mp i\pi f(0), \quad (1.111)$$

or, in short

$$\lim_{\epsilon \rightarrow 0^+} \frac{1}{x \pm i\epsilon} = \mathcal{P} \left(\frac{1}{x} \right) \mp i\pi\delta(x), \quad (1.112)$$

where it is understood that the second of these two equations have meaning only within an integral. The symbol \mathcal{P} means the the principal part of an integral where the integrand has a simple pole. The principal part is defined as

$$\mathcal{P} \int_{-A}^B \frac{dx}{x} f(x) = \lim_{\eta \rightarrow 0^+} \left[\int_{-A}^{-\eta} + \int_{\eta}^B \right] \frac{dx}{x} f(x). \quad (1.113)$$

Proof:

We can write

$$\frac{1}{x \pm i\epsilon} = \frac{x \mp i\epsilon}{x^2 + \epsilon^2} = \frac{x}{x^2 + \epsilon^2} \mp \frac{i\epsilon}{x^2 + \epsilon^2}. \quad (1.114)$$

Now, we have

$$\lim_{\epsilon \rightarrow 0^+} \frac{1}{\pi} \frac{\epsilon}{x^2 + \epsilon^2} = \delta(x).$$

Therefore,

$$\lim_{\epsilon \rightarrow 0^+} (\mp) i \frac{\epsilon}{x^2 + a^2} = \mp i\pi \delta(x). \quad (1.115)$$

Next, consider the first term on the right hand side of Eq. (1.114). We multiply this term by a function $f(x)$ which is regular at the origin and then integrate over x . We get

$$\begin{aligned} & \lim_{\epsilon \rightarrow 0^+} \int_{-\infty}^{\infty} \frac{xf(x)}{x^2 + a^2} dx \\ &= \lim_{\epsilon \rightarrow 0^+} \left[\lim_{\eta \rightarrow 0^+} \int_{-\infty}^{-\eta} \frac{xf(x)}{x^2 + a^2} dx + \lim_{\eta \rightarrow 0^+} \int_{-\eta}^{\eta} \frac{xf(x)}{x^2 + a^2} dx + \lim_{\eta \rightarrow 0^+} \int_{\eta}^{\infty} \frac{xf(x)}{x^2 + a^2} dx \right] \end{aligned} \quad (1.116)$$

Note that we take the limit over η first and then we take the limit over η . Now consider the second integral of the above equation.

$$\lim_{\eta \rightarrow 0^+} \int_{-\eta}^{\eta} \frac{xf(x)}{x^2 + a^2} dx = f(0) \lim_{\eta \rightarrow 0^+} \frac{1}{2} [\ln(x^2 + \epsilon^2)]_{x=-\eta}^{x=\eta} = 0.$$

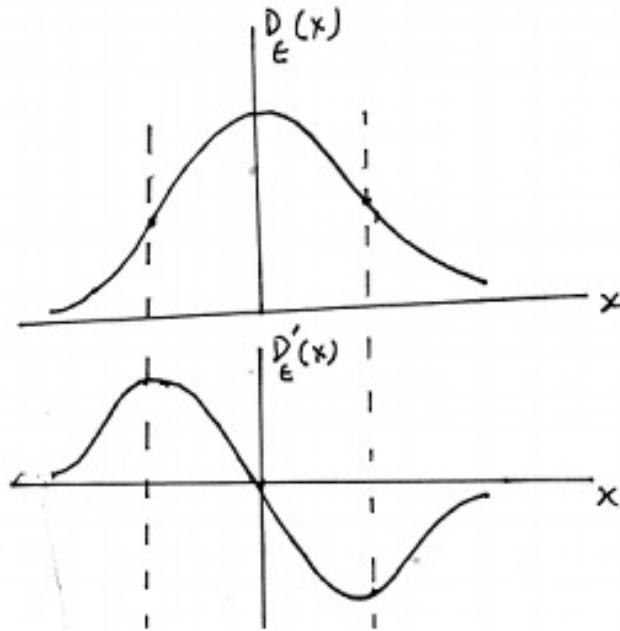
If we now reverse the order of the evaluation of limits in Eq. (1.116), the $\epsilon \rightarrow 0^+$ limit causes no difficulties in the other two integrals. We thus have

$$\begin{aligned} \lim_{\epsilon \rightarrow 0^+} \int_{-\infty}^{\infty} \frac{xdx}{x^2 + a^2} f(x) &= \lim_{\eta \rightarrow 0^+} \lim_{\epsilon \rightarrow 0^+} \left[\int_{-\infty}^{-\eta} + \int_{\eta}^{\infty} \right] \frac{xdx}{x^2 + a^2} f(x) \\ &= \lim_{\eta \rightarrow 0^+} \left[\int_{-\infty}^{-\eta} + \int_{\eta}^{\infty} \right] \frac{dx}{x} f(x) \\ &= \mathcal{P} \int_{-\infty}^{\infty} \frac{1}{x} f(x) dx \end{aligned}$$

This establishes the identity.

1.9 Derivative of the Delta Function

One may define the derivative $\delta'(x)$ of the delta function. When ϵ is small, the derivative of $D_\epsilon(x)$ has two peaks close to the origin, one peak being positive and the other negative as shown in the figure below.

Figure 1.8: Plot of $D_\epsilon(x)$ and $D'_\epsilon(x)$.

As $\epsilon \rightarrow 0$, each of the peaks become very narrow and very tall, and each of the two peaks approach very close to the origin. Now, an integration by parts gives

$$\int_{-\infty}^{\infty} D'_\epsilon(x)f(x)dx = [D_\epsilon(x)f(x)]_{-\infty}^{\infty} - \int_{-\infty}^{\infty} D_\epsilon(x)f'(x)dx. \quad (1.117)$$

Because $D_\epsilon(x)$ tends to zero as $x \rightarrow \pm\infty$, the first term on the right hand side vanishes unless $f(x)$ explodes violently at infinity. So by letting $\epsilon \rightarrow 0$, we arrive at the definition of $\delta'(x)$:

$$\int_{-\infty}^{\infty} \delta'(x)f(x)dx = - \int_{-\infty}^{\infty} \delta(x)f'(x)dx = -f'(0).$$

(1.118)

From this we immediately get

$$x\delta'(x) = -\delta(x).$$

(1.119)

Conversely, it can be shown that the general solution of the equation

$$xu(x) = \delta(x)$$

can be written as

$$u(x) = -\delta'(x) + c\delta(x)$$

where the second term arises from the homogeneous equation

$$x\delta(x) = 0.$$

From the definition (1.118) it also follows that

$$\boxed{\delta'(-x) = -\delta'(x).} \quad (1.120)$$

The n^{th} order derivative of $\delta(x)$ can be defined in the same way. We find

$$\int_{-\infty}^{\infty} \delta^{(n)}(x) f(x) dx = (-1)^n f^{(n)}(0). \quad (1.121)$$

We can also prove the following properties of the derivatives of the delta function:

$$\begin{aligned}\delta^{(m)}(x) &= (-1)^m \delta^{(m)}(-x) \\ x^{m+1} \delta^{(m)}(x) &= 0 \\ x \delta^{(m)}(x) &= -(m-1) \delta^{(m-1)}(x).\end{aligned}$$

1.10 Integration of the Delta Function

Consider the indefinite integral

$$\theta_\epsilon(x) = \int_{-\infty}^x D_\epsilon(y) dy. \quad (1.122)$$

A graph of $\theta_\epsilon(x)$ versus x is shown below.

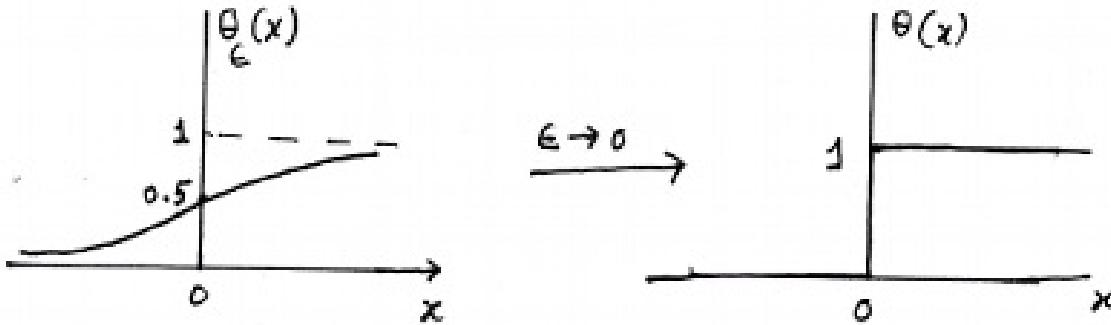


Figure 1.9: The θ -function as an integral of the delta function

As $\epsilon \rightarrow 0$, the step in the function $\theta_\epsilon(x)$ gets progressively steeper, until, finally, the function changes abruptly from 0 to 1 at $x = 0$. Therefore, taking the limit $\epsilon \rightarrow 0$ in Eq. (1.122) we have

$$\theta(x) = \int_{-\infty}^x \delta(y) dy \quad (1.123)$$

where

$$\theta(x) = \begin{cases} 1 & \text{for } x > 0 \\ 0 & \text{for } x < 0. \end{cases} \quad (1.124)$$

If we differentiate Eq. (1.123) with respect to x , we get

$$\frac{d\theta(x)}{dx} = \delta(x). \quad (1.125)$$

1.11 Three dimensional delta function

The three-dimensional delta function $\delta(\vec{r})$ is defined as

$$\delta(\vec{r}) \stackrel{\text{def}}{=} \delta(x)\delta(y)\delta(z). \quad (1.126)$$

In other words, $\delta(\vec{r})$ is zero if any of the coordinates x , y and z is not equal to zero and $\delta(\vec{r})$ tends to infinity at the origin, i.e., when $x = 0$, $y = 0$ and $z = 0$, such that

$$\int_{\text{volume}} \delta(\vec{r}) d^3r = 1 \quad (1.127)$$

if the volume of integration contains the origin. We also have

$$\int \delta(\vec{r}) f(\vec{r}) d^3r = f(0) \quad (1.128)$$

where again the volume of integration contains the origin.

Note:

- $\delta(\vec{r} - \vec{r}') = \delta(x - x')\delta(y - y')\delta(z - z')$
- $\int_V \delta(\vec{r} - \vec{r}') d^3r = 1$
where the volume of integration includes the point \vec{r}' . Otherwise, the integral is zero.
- $\int_V \delta(\vec{r} - \vec{r}') f(\vec{r}) d^3r = f(\vec{r}')$
if V includes the point \vec{r}' .

A useful formula

Consider the integral

$$\begin{aligned} \int_{-\infty}^{\infty} e^{ikx} dx &= \lim_{L \rightarrow \infty} \int_{-L}^L e^{ikx} dx \\ &= \lim_{L \rightarrow \infty} \frac{1}{ik} \left(e^{ikL} - e^{-ikL} \right) \\ &= \lim_{L \rightarrow \infty} \frac{2}{k} \left(\frac{e^{ikL} - e^{-ikL}}{2i} \right) \\ &= \lim_{L \rightarrow \infty} \frac{2}{k} \sin kL \\ &= 2\pi \lim_{L \rightarrow \infty} \frac{\sin kL}{\pi k} \\ &= 2\pi\delta(k), \end{aligned}$$

where we have used

$$\lim_{\epsilon \rightarrow 0} \frac{\sin(x/\epsilon)}{\pi x} = \delta(x). \quad (1.129)$$

Thus, we have the very important formula

$$\boxed{\int_{-\infty}^{\infty} e^{ikx} dx = 2\pi\delta(k)} \quad (1.130)$$

In Eq. (1.130) if we integrate with respect to k , we would have $\delta(x)$ on the right hand side,

$$\boxed{\int_{-\infty}^{\infty} e^{ikx} dk = 2\pi\delta(x)} \quad (1.131)$$

Also note that in Eq. (1.130) we integrate over the full domain of x from $-\infty$ to ∞ . Making a change of variable $x \rightarrow -x$ does not change the value of the integral. Hence we also have

$$\boxed{\int_{-\infty}^{\infty} e^{-ikx} dx = 2\pi\delta(k)} \quad (1.132)$$

Similarly, in Eq. (1.131), making the change of variable $k \rightarrow -k$, doesn't change the value of the integral. So we could also write

$$\boxed{\int_{-\infty}^{\infty} e^{-ikx} dk = 2\pi\delta(x)} \quad (1.133)$$

Thus, in summary

$$\int_{-\infty}^{\infty} e^{\pm ikx} dx = 2\pi\delta(k), \quad (1.134)$$

and

$$\int_{-\infty}^{\infty} e^{\pm ikx} dk = 2\pi\delta(x), \quad (1.135)$$

In three dimensions we have

$$\int_{\text{all space}} e^{\pm i\vec{k}\cdot\vec{r}} d^3r = (2\pi)^3 \delta(\vec{k}). \quad (1.136)$$

$$\int_{\text{all space}} e^{\pm i(\vec{k}-\vec{k}')\cdot\vec{r}} d^3r = (2\pi)^3 \delta(\vec{k}-\vec{k}'). \quad (1.137)$$

$$\int_{\text{all space}} e^{\pm i\vec{k}\cdot(\vec{r}-\vec{r}')} d^3k = (2\pi)^3 \delta(\vec{r}-\vec{r}'). \quad (1.138)$$

1.12 Fourier transform

We can always express a function $f(x)$ in the form

$$f(x) = \int_{-\infty}^{\infty} e^{ikx} \tilde{f}(k) dk \quad (1.139)$$

where $\tilde{f}(k)$ is a function of k , called the Fourier transform of $f(x)$. From Eq. (1.139) we can write

$$\begin{aligned} \int_{-\infty}^{\infty} e^{-ik'x} f(x) dx &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{i(k-k')x} \tilde{f}(k) dk dx \\ &= 2\pi \int_{-\infty}^{\infty} \delta(k - k') \tilde{f}(k) dk \\ &= 2\pi \tilde{f}(k') \end{aligned}$$

Thus,

$$\tilde{f}(k) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-ikx} f(x) dx. \quad (1.140)$$

The functions $f(x)$ and $\tilde{f}(k)$ are Fourier transform of each other. We can write Eqs. (1.139) and (1.140) in a more symmetrical fashion as follows

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{ikx} \tilde{f}(k) dk \quad (1.141)$$

$$\tilde{f}(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-ikx} f(x) dx. \quad (1.142)$$

In three dimensions we can write

$$f(\vec{r}) = \frac{1}{(2\pi)^{3/2}} \int e^{i\vec{k}\cdot\vec{r}} \tilde{f}(\vec{k}) d^3k. \quad (1.143)$$

Multiplying this equation by $e^{-i\vec{k}'\cdot\vec{r}}$ and integrating over \vec{r} , we have

$$\begin{aligned} \int_{\text{all space}} e^{-i\vec{k}'\cdot\vec{r}} f(\vec{r}) d^3r &= \frac{1}{(2\pi)^{3/2}} \int d^3r \int d^3k e^{i(\vec{k}-\vec{k}')\cdot\vec{r}} \tilde{f}(\vec{k}) \\ &= \frac{1}{(2\pi)^{3/2}} \int d^3k (2\pi)^3 \delta(\vec{k} - \vec{k}') \tilde{f}(\vec{k}) \\ &= (2\pi)^{3/2} \tilde{f}(\vec{k}') \end{aligned}$$

Therefore,

$$\tilde{f}(\vec{k}) = \frac{1}{(2\pi)^{3/2}} \int e^{-i\vec{k}\cdot\vec{r}} f(\vec{r}) d^3r. \quad (1.144)$$

In summary,

$$f(\vec{r}) = \frac{1}{(2\pi)^{3/2}} \int e^{i\vec{k}\cdot\vec{r}} \tilde{f}(\vec{k}) d^3k, \quad (1.145)$$

$$\tilde{f}(\vec{k}) = \frac{1}{(2\pi)^{3/2}} \int e^{-i\vec{k}\cdot\vec{r}} f(\vec{r}) d^3r. \quad (1.146)$$

1.12.1 Parseval Identity

We will now prove the important identity

$$\int |f(\vec{r})|^2 d^3r = \int |\tilde{f}(\vec{k})|^2 d^3k. \quad (1.147)$$

Proof:

$$\begin{aligned}
\int |f(\vec{r})|^2 d^3r &= \int f(\vec{r}) f^*(\vec{r}) d^3r \\
&= \int d^3r \frac{1}{(2\pi)^{3/2}} \int e^{i\vec{k}\cdot\vec{r}} \tilde{f}(\vec{k}) d^3k \frac{1}{(2\pi)^{3/2}} \int e^{-i\vec{k}'\cdot\vec{r}} \tilde{f}^*(\vec{k}') d^3k' \\
&= \frac{1}{(2\pi)^3} \int d^3k d^3k' \tilde{f}(\vec{k}) \tilde{f}^*(\vec{k}') \underbrace{\int d^3r e^{i(\vec{k}-\vec{k}')\cdot\vec{r}}}_{(2\pi)^3 \delta(\vec{k}-\vec{k}')} \\
&= \int d^3k d^3k' \tilde{f}(\vec{k}) \tilde{f}^*(\vec{k}') \delta(\vec{k}-\vec{k}') \\
&= \int d^3k \tilde{f}(\vec{k}) \tilde{f}^*(\vec{k}) \\
&= \int d^3k \left| \tilde{f}(\vec{k}) \right|^2.
\end{aligned} \quad (1.148)$$

The proof is now complete.

**DIRAC DELTA FUNCTION
FOR
STUDENTS IN A HURRY**

by

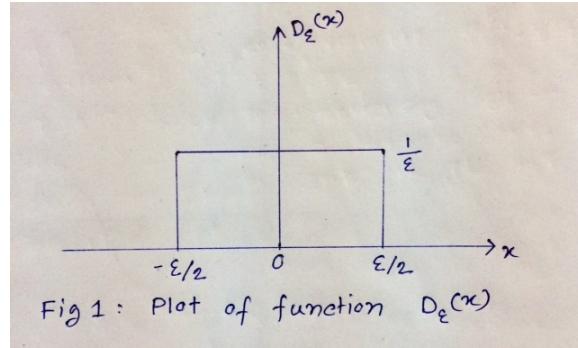
Prof. Dr. Khorshed Ahmed Kabir

Dirac Delta Function

Consider the function $D_\varepsilon(x)$ given by

$$\begin{aligned} D_\varepsilon(x) &= \frac{1}{\varepsilon} ; \text{ for } -\frac{\varepsilon}{2} \leq x \leq \frac{\varepsilon}{2} \\ &= 0 ; \text{ for } |x| > \frac{\varepsilon}{2} \end{aligned}$$

where ε is a positive parameter. The plot of the function is shown below.



The integral of the function with respect to x is 1, i.e,

$$\int_{-\infty}^{\infty} D_\varepsilon(x) dx = 1 \quad (1.149)$$

Now imagine making ε smaller. As we decrease ε , the function gets narrower and taller, but the integral of the function i.e, the area under the graph remain constant at the value 1. In the limit $\varepsilon \rightarrow 0$, the function $D_\varepsilon(x)$ collapses to a single point $x = 0$ and gets infinitely tall. So, $\lim_{\varepsilon \rightarrow 0} D_\varepsilon(x)$ is not a function at all and the procedure of taking the limit $\varepsilon \rightarrow 0$ is not justified.

However, we can make the limiting procedure meaningful if multiply $D_\varepsilon(x)$ by some well defined function $f(x)$, integrate over x and then take the limit $\varepsilon \rightarrow 0$. Consider, the integral

$$\int_{-\infty}^{\infty} D_\varepsilon(x) f(x) dx$$

where $f(x)$ is a well defined function. If ε is significantly small, the variation of $f(x)$ over the effective integration interval $[-\frac{\varepsilon}{2}, \frac{\varepsilon}{2}]$ is negligible and $f(x)$ remain practically equal to $f(0)$. Therefore,

$$\int_{-\infty}^{\infty} D_\varepsilon(x) f(x) dx \simeq f(0) \int_{-\infty}^{\infty} D_\varepsilon(x) dx = f(0) \quad (1.150)$$

The smaller the value of ε , the better the approximation. in the limit $\varepsilon \rightarrow 0$, the above equation is exactly

$$\lim_{\varepsilon \rightarrow 0} \int_{-\infty}^{\infty} D_\varepsilon(x) f(x) dx = f(0) \quad (1.151)$$

Now, we define the delta function by the relation

$$\int_{-\infty}^{\infty} \delta(x)f(x)dx = \lim_{\epsilon \rightarrow 0} \int_{-\infty}^{\infty} D_{\epsilon}(x)f(x)dx = f(0) \quad (1.152)$$

This equation is valid for any function defined at the origin. More generally, $\delta(x - x_0)$ is defined as

$$\int_{-\infty}^{\infty} \delta(x - x_0)f(x)dx = f(x_0) \quad (1.153)$$

Actually, the integral notation $\int_{-\infty}^{\infty} \delta(x)f(x)dx$ is not justified because $\delta(x)$ is not really a function. Physically, there is no problem since it becomes impossible to distinguish between $D_{\epsilon}(x)$ and $\delta(x)$ as soon as ϵ becomes negligible compared to all the distances involved in a physical problem. Whenever a mathematical difficulty arise, all we need to do is to assume that $\delta(x)$ is actually $D_{\epsilon}(x)$ with ϵ extremely small but not strictly zero.

Formally, we can express $\delta(x)$ as a sequence of proper functions

$$\lim_{\epsilon \rightarrow 0} D_{\epsilon}(x) \equiv \delta(x)$$

Here, $D_{\epsilon}(x)$, which is a proper function of x is called the representation of the delta function. One representation is the 'Square Function' given at the beginning. The representation is not unique. There are other functions which approach the delta function when appropriate limits are taken.

Other Representation Of The Delta Function

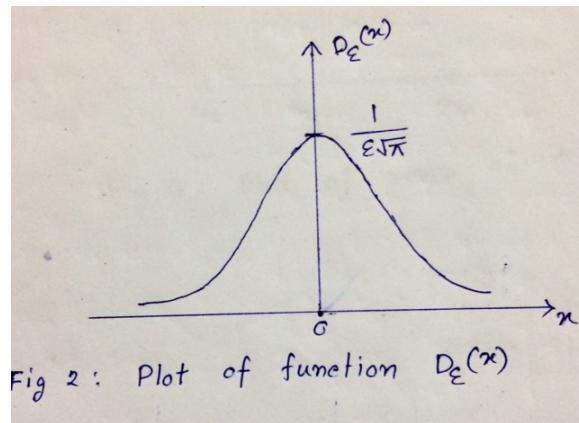
A

Consider the function

$$D_{\epsilon}(x) = \frac{1}{\epsilon\sqrt{\pi}} e^{-x^2/\epsilon^2} \quad \text{with } \epsilon > 0 \quad (1.154)$$

For each value of the parameter ϵ , this function satisfies

$$\int_{-\infty}^{\infty} D_{\epsilon}(x)dx = 1$$



When plotted against x , the function has a peak at the origin. The peak has a height $\frac{1}{\varepsilon\sqrt{\pi}}$ and a width of order ε (exactly how the width is defined doesn't matter). So, if ε is allowed to become very small, the peak becomes very tall and very narrow. Outside the peak, the function becomes extremely small. Thus we have

$$\delta(x) = \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon\sqrt{\pi}} e^{-x^2/\varepsilon^2} \quad (1.155)$$

Note:

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

Let,

$$\begin{aligned} I &= \int_{-\infty}^{\infty} e^{-(b^2x^2+ax)} dx \\ &= e^{a^2/4b^2} \int_{-\infty}^{\infty} e^{-(bx+a/2b)^2} dx \\ &= e^{a^2/4b^2} \frac{1}{b} \int_{-\infty}^{\infty} e^{-z^2} dz \\ &= e^{a^2/4b^2} \frac{\sqrt{\pi}}{b} \end{aligned}$$

By using

$$b^2x^2 + ax = (bx + \frac{a}{2b})^2 - \frac{a^2}{4b^2}$$

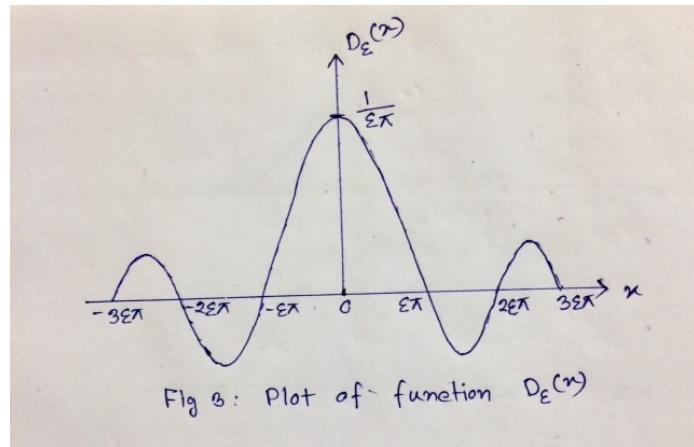
and by letting

$$z = bx + \frac{a}{2b}$$

B

Consider another function

$$D_\varepsilon(x) = \frac{1}{\pi} \frac{\sin(x/\varepsilon)}{x} \quad \text{with } \varepsilon > 0 \quad (1.156)$$



For any value of the parameter ε we have

$$\int_{-\infty}^{\infty} D_\varepsilon(x) dx = 1 \quad (1.157)$$

A plot of the function $D_\varepsilon(x)$ shows that it has the value $\frac{1}{\varepsilon\pi}$ at $x=0$ and it oscillates with decreasing amplitude as $|x|$ increases. The width of the central maxima is of the order of ε and the period of oscillation with respect to x is $2\pi\varepsilon$.

Thus the limit of the function as $\varepsilon \rightarrow 0$ has all the properties of the delta function: it becomes infinitely large at $x=0$, it has unit integral, and infinitely rapid oscillations as $|x|$ increases means that the entire contribution to an integral containing this function comes from an infinitesimal neighbourhood of $x=0$. We can therefore write

$$\delta(x) = \lim_{\varepsilon \rightarrow 0} \frac{1}{\pi} \frac{\sin(x/\varepsilon)}{x} \quad (1.158)$$

C

We can also show that

$$\delta(x) = \lim_{\varepsilon \rightarrow 0} \frac{1}{2\varepsilon} e^{-|x|/\varepsilon}$$

$$\delta(x) = \lim_{\varepsilon \rightarrow 0} \frac{1}{\pi} \frac{\varepsilon}{x^2 + \varepsilon^2}$$

$$\delta(x) = \lim_{\varepsilon \rightarrow 0} \frac{\varepsilon}{\pi} \frac{\sin^2(x/\varepsilon)}{x^2}$$

Properties Of The Delta Function

It is important to note that, because of its singular character, the delta function can not be the end result of a calculation and has meaning only so long as a subsequent integral over its argument is carried out. With this understanding we can write down some relations between delta functions.

1. The delta function is an even function:

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

2.

$$x\delta(x) = 0$$

3.

$$\delta(ax) = \frac{1}{|a|} \delta(x)$$

Proof Of 3

Consider the integral

$$I = \int_{-\infty}^{\infty} \delta(ax)f(x)dx$$

Since the delta function is even in its argument, it doesn't matter if we replace a by $|a|$ in its argument.

Thus

$$I = \int_{-\infty}^{\infty} \delta(|a|x)f(x)dx$$

. Making the change in variable

$$y = |a|x$$

we have

$$I = \frac{1}{|a|} \int_{-\infty}^{\infty} \delta(y)f\left(\frac{y}{|a|}\right)dy = \frac{1}{|a|}f(0)$$

Or,

$$\int_{-\infty}^{\infty} \delta(ax)f(x)dx = \frac{1}{|a|} \int_{-\infty}^{\infty} \delta(x)f(x)dx$$

Or,

$$\delta(ax) = \frac{1}{|a|} \delta(x)$$

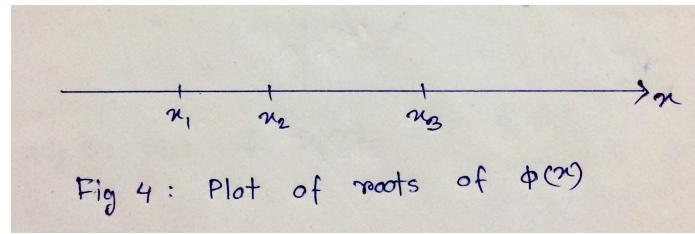
4. More generally,

$$\delta(\phi(x)) = \sum_i \frac{\delta(x - x_i)}{\left| \frac{\partial \phi}{\partial x} \right|_{x_i}}$$

where the sum runs over the x_i 's which are simple roots of $\phi(x)$.

Proof Of 4

Let $x_1, x_2, x_3, \dots, x_N$ be the simple roots of $\phi(x)$



In the neighbourhood of any one of the simple roots x_i , we can write

$$\phi(x) = (x - x_i)\psi(x)$$

where $\psi(x_i) \neq 0$. We have $\psi(x_i) = \left. \frac{\partial \phi(x)}{\partial x} \right|_{x=x_i}$ Now consider the integral

$$\begin{aligned} & \int_{-\infty}^{\infty} \delta(\phi(x)) f(x) dx \\ &= \sum_{i=1}^N \int_{x_i-\epsilon}^{x_i+\epsilon} \delta[(x-x_i)\psi(x_i)] f(x) dx \\ &= \sum_{i=1}^N \frac{1}{|\psi(x_i)|} \int_{x_i-\epsilon}^{x_i+\epsilon} \delta(x-x_i) f(x) dx \\ &= \sum_{i=1}^N \frac{1}{\left| \frac{\partial \phi}{\partial x} \right|_{x=x_i}} \int_{-\infty}^{\infty} \delta(x-x_i) f(x) dx \\ &= \sum_{i=1}^N \frac{1}{\left| \frac{\partial \phi}{\partial x} \right|_{x=x_i}} f(x_i) \end{aligned}$$

The above result is obtained, if we write

$$\delta(\phi(x)) = \sum_{i=1}^N \frac{\delta(x-x_i)}{\left| \frac{\partial \phi}{\partial x} \right|_{x=x_i}}$$

5. A frequently used example of the above result is

$$\delta(x^2 - a^2) = \frac{1}{2a} \delta(x-a) + \frac{1}{2a} \delta(x+a) \quad \text{with } a > 0$$

Here

$$\phi(x) = x^2 - a^2 = (x+a)(x-a)$$

. The two simple roots of $\phi(x)$ are at $x=a$ and $x=-a$. Now,

$$\left| \frac{\partial \phi}{\partial x} \right|_{x=a} = |2x|_{x=a} = 2a$$

And

$$\left| \frac{\partial \phi}{\partial x} \right|_{x=-a} = |2x|_{x=-a} = |-2a| = 2a$$

So, the above result follows.

6.

$$f(x)\delta(x-a) = f(a)\delta(x-a)$$

7.

$$\int_{-\infty}^{\infty} \delta(x-y)\delta(y-a)dy = \delta(x-a)$$

Note 1: We have the identity

$$x\delta(x) = 0$$

. The converse is also true and it can be shown that the equation

$$xu(x) = 0$$

has the general solution

$$u(x) = c\delta(x)$$

Note 2: We will now prove an identity which is particularly useful in quantum mechanics.

$$\lim_{\epsilon \rightarrow 0^+} \int_{-\infty}^{\infty} \frac{1}{x \pm i\epsilon} f(x) dx = P \int_{-\infty}^{\infty} \frac{dx}{x} f(x) \mp i\pi f(0)$$

Or in short

$$\lim_{\epsilon \rightarrow 0^+} \frac{1}{x \pm i\epsilon} = P\left(\frac{1}{x}\right) \mp i\pi\delta(x)$$

where it is understood that the econd of these two equations have meaning only within an integral.

The symbol P means principle part of an integral where the integral has a simple pole. The principle part is defined as

$$P \int_A^B \frac{dx}{x} f(x) = \lim_{n \rightarrow 0^+} \left[\int_A^\eta + \int_\eta^B \right] \frac{dx}{x} f(x)$$

Proof:

$$\frac{1}{x \pm i\epsilon} = \frac{x \mp i\epsilon}{x^2 + \epsilon^2} = \frac{x}{x^2 + \epsilon^2} \mp \frac{i\epsilon}{x^2 + \epsilon^2} \quad (1.159)$$

Now we have

$$\lim_{\epsilon \rightarrow 0^+} \frac{1}{\pi} \frac{\epsilon}{x^2 + \epsilon^2} = \delta(x)$$

So,

$$\lim_{\epsilon \rightarrow 0^+} (\mp) i \frac{\epsilon}{x^2 + \epsilon^2} = \mp i\pi\delta(x) \quad (1.160)$$

Now consider the first term on the right hand side of eqn(11). We multiply this term by a function $f(x)$ which is regular at the origin and then integrate over x . We get,

$$\lim_{\varepsilon \rightarrow 0^+} \int_{-\infty}^{\infty} \frac{xf(x)}{x^2 + \varepsilon^2} dx = \lim_{\varepsilon \rightarrow 0^+} \left[\lim_{\eta \rightarrow 0^+} \int_{-\infty}^{-\eta} \frac{xf(x)dx}{x^2 + \varepsilon^2} + \int_{-\eta}^{\eta} \frac{xf(x)dx}{x^2 + \varepsilon^2} + \int_{\eta}^{\infty} \frac{xf(x)dx}{x^2 + \varepsilon^2} \right] \quad (1.161)$$

Note that we take the limit over η first and then we take the limit over ε . Consider now the second integral above

$$\lim_{\eta \rightarrow 0^+} \int_{-\eta}^{+\eta} \frac{xf(x)dx}{x^2 + \varepsilon^2} = f(0) \lim_{\eta \rightarrow 0^+} \frac{1}{2} [\ln(x^2 + \varepsilon^2)]_{x=-\eta}^{x=\eta} = 0$$

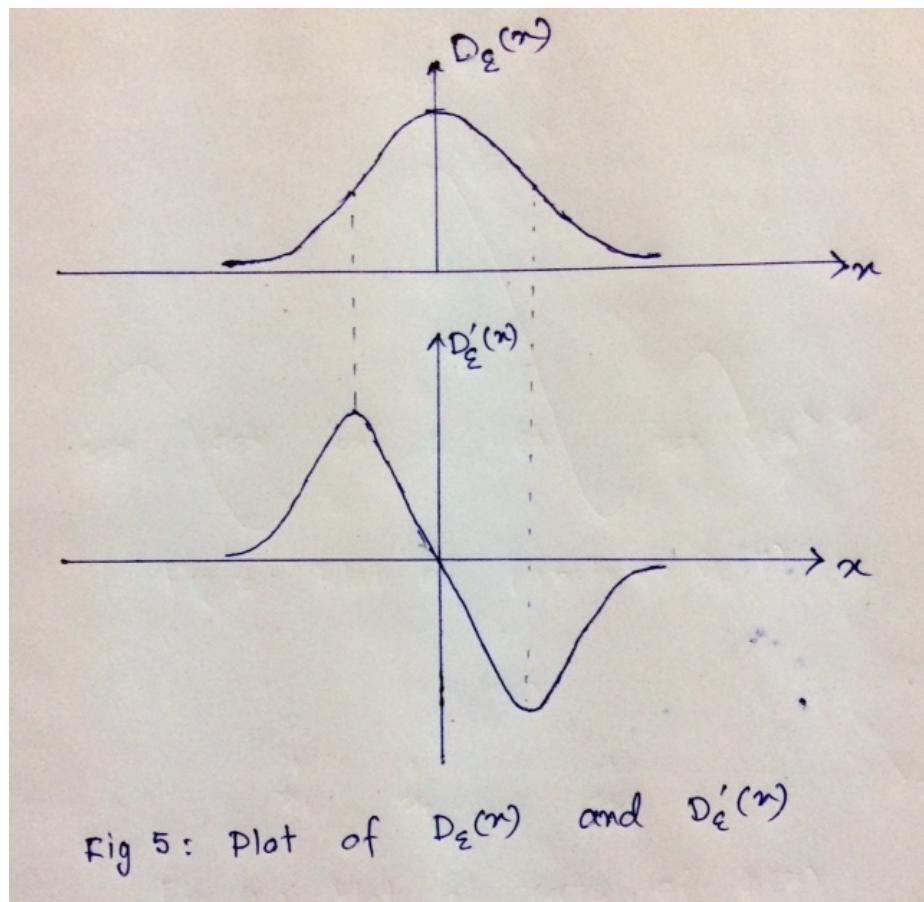
If we now reverse the order of the evaluation of limits in eqn (13), the $\varepsilon \rightarrow 0$ limit causes no difficulties in the other two integrals. We thus have

$$\begin{aligned} \lim_{\varepsilon \rightarrow 0^+} \int_{-\infty}^{\infty} \frac{xf(x)dx}{x^2 + \varepsilon^2} &= \lim_{\eta \rightarrow 0^+} \lim_{\varepsilon \rightarrow 0^+} \left[\int_{-\infty}^{-\eta} + \int_{\eta}^{\infty} \right] \frac{xf(x)dx}{x^2 + \varepsilon^2} \\ &= \lim_{\eta \rightarrow 0^+} \left[\int_{-\infty}^{-\eta} + \int_{\eta}^{\infty} \right] \frac{dx}{x} f(x) \\ &= P \int_{-\infty}^{\infty} \frac{1}{x} f(x) dx \end{aligned}$$

This establishes the identity.

Derivatives Of The Delta Function

One may define the derivative $\delta'(x)$ of the delta function. When ε is small, the derivative of $D_\varepsilon(x)$ has two peaks close to the origin, one peak positive and the other negative as drawn in the figure below.



As $\varepsilon \rightarrow 0$, each of these peaks becomes very narrow and very tall, and two peaks each approach very close to the origin. Now, an integration by parts gives

$$\int_{-\infty}^{\infty} dx D'_\varepsilon(x) f(x) = [D_\varepsilon(x)f(x)]_{-\infty}^{\infty} - \int_{-\infty}^{\infty} dx D_\varepsilon(x) f'(x) \quad (1.162)$$

Because $D_\varepsilon(x)$ tends to zero as $x \rightarrow \pm\infty$, the first term on the right hand side vanishes unless $f(x)$ explodes violently at infinity. So by letting $\varepsilon \rightarrow 0$, we arrive at the definition of $\delta'(x)$:

$$\int_{-\infty}^{\infty} \delta'(x) f(x) dx = - \int_{-\infty}^{\infty} \delta(x) f'(x) dx = -f'(0) \quad (1.163)$$

From this, we immediately get

$$x\delta'(x) = -\delta(x) \quad (1.164)$$

Conversely, it can be shown that the general solution of the equation

$$xu(x) = \delta(x)$$

can be written as

$$u(x) = -\delta'(x) + c\delta(x)$$

where the second term arises from the homogeneous equation

$$x\delta(x) = 0$$

From eqn(15) it also follows that

$$\delta'(-x) = -\delta'(x) \quad (1.165)$$

The n-th order derivative of $\delta(x)$ can be defined in the same way. We find

$$\int_{-\infty}^{\infty} \delta^{(n)}(x)f(x)dx = (-1)^n f(0) \quad (1.166)$$

We can prove following properties:

1.

$$\delta^{(m)}(x) = (-1)^m \delta^{(m)}(-x)$$

2.

$$x^{m+1} \delta^{(m)}(x) = 0$$

3.

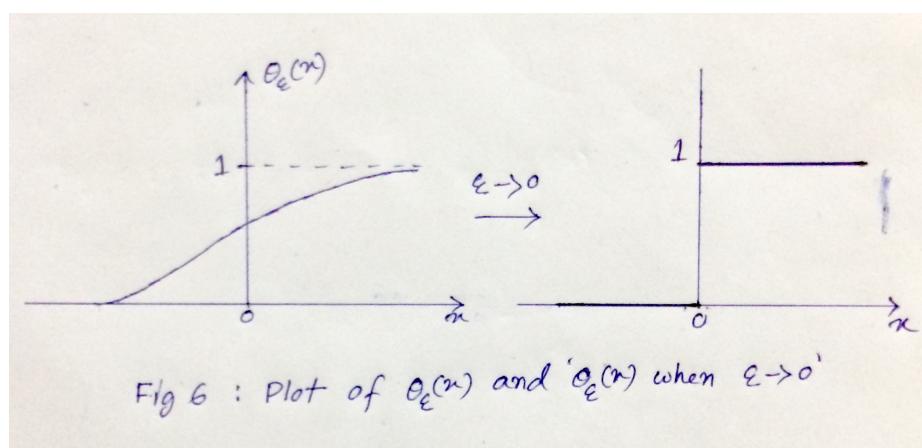
$$x\delta^{(m)}(x) = -m\delta^{(m-1)}(x)$$

Integration Of The Delta Function

Consider the indefinite integral

$$\theta_\epsilon(x) = \int_{-\infty}^x D_\epsilon(y)dy \quad (1.167)$$

A graph of $\theta_\epsilon(x)$ vs x is shown below:



As $\varepsilon \rightarrow 0$, the step in the function $\theta_\varepsilon(x)$ gets progressively steeper, until, finally, the function changes abruptly from 0 to 1 at $x=0$. Thus, taking the limit $\varepsilon \rightarrow 0$ in eqn(19), we have

$$\theta(x) = \int_{-\infty}^x \delta(x) dx \quad (1.168)$$

where

$$\begin{aligned}\theta(x) &= 1 \quad \text{for } x > 0 \\ &= 0 \quad \text{for } x < 0\end{aligned}$$

If we differentiate eqn(20) with respect to x, we get

$$\frac{d\theta(x)}{dx} = \delta(x) \quad (1.169)$$

Three Dimensional Delta Function

We define

$$\delta(\vec{r}) \equiv \delta(x)\delta(y)\delta(z)$$

In other words, $\delta(\vec{r})$ is zero if any of the coordinates x,y and z is not equal to zero and $\delta(\vec{r})$ tends to infinitely at the origin, i.e, when $x=0$, $y=0$ and $z=0$, such that

$$\int_{volume} \delta(\vec{r}) d^3r = 1$$

if the volume of integration contains the origin. We also have

$$\int_{volume} \delta(\vec{r}) f(\vec{r}) d^3r = f(0)$$

where again, the volume of integration contains the origin.

Note:

1.

$$\delta(\vec{r} - \vec{r}') = \delta(x - x')\delta(y - y')\delta(z - z')$$

2.

$$\delta(\vec{r} - \vec{r}') d^3r = 1$$

where the volume of integration includes the point \vec{r}' , otherwise the integral is zero.

3.

$$\int_{volume} \delta(\vec{r} - \vec{r}') f(\vec{r}) d^3r = f(\vec{r}')$$

if volume of integration includes the point \vec{r}'

A Useful Formula

Consider the integral

$$\begin{aligned}
 \int_{-\infty}^{\infty} e^{ikx} dx &= \lim_{L \rightarrow \infty} \int_{-L}^L e^{ikx} dx \\
 &= \lim_{L \rightarrow \infty} \frac{1}{ik} (e^{ikL} - e^{-ikL}) \\
 &= \lim_{L \rightarrow \infty} \frac{2}{k} \left(\frac{e^{ikL} - e^{-ikL}}{2i} \right) \\
 &= \lim_{L \rightarrow \infty} \frac{2}{k} \sin kL \\
 &= 2\pi \lim_{L \rightarrow \infty} \frac{\sin kL}{\pi k} \\
 &= 2\pi\delta(k)
 \end{aligned}$$

Since

$$\lim_{\varepsilon \rightarrow 0} \frac{\sin(x/\varepsilon)}{\pi x} = \delta(x).$$

$$\text{Let, } \varepsilon = \frac{1}{L}$$

So, when, $L \rightarrow \infty$, then $\varepsilon \rightarrow 0$.

$$\text{therefore, } \lim_{L \rightarrow \infty} \frac{\sin kL}{\pi k} = \lim_{\varepsilon \rightarrow 0} \frac{\sin(k/\varepsilon)}{\pi k}$$

Thus

$$\int_{-\infty}^{\infty} e^{ikx} dx = 2\pi\delta(k) \quad (1.170)$$

In eqn(22) if we integrated with respect to k , we would have $\delta(x)$ on the right handside

$$\int_{-\infty}^{\infty} e^{ikx} dk = 2\pi\delta(x) \quad (1.171)$$

Also note that in eqn(22) we are integrating over x over its full range of values. Making a change of variable $x \rightarrow -x$ does not change the value of the integral. Here we also have

$$\int_{-\infty}^{\infty} e^{-ikx} dx = 2\pi\delta(k) \quad (1.172)$$

Similarly in eqn(23), making the change of variable $k \rightarrow -k$, doesn't change the value of the integral. So we could also write

$$\int_{-\infty}^{\infty} e^{-ikx} dk = 2\pi\delta(x) \quad (1.173)$$

Thus in summary,

$$\int_{-\infty}^{\infty} e^{\pm ikx} dx = 2\pi \delta(k)$$

$$\int_{-\infty}^{\infty} e^{\pm ikx} dk = 2\pi \delta(x)$$

In three dimension

$$\int_{all\ space} e^{\pm i\vec{k}\cdot\vec{r}} d^3r = (2\pi)^3 \delta(\vec{k})$$

$$\int_{all\ space} e^{\pm i(\vec{k}-\vec{k}')\cdot\vec{r}} d^3r = (2\pi)^3 \delta(\vec{k}-\vec{k}')$$

$$\int_{all\ space} e^{\pm i\vec{k}\cdot(\vec{r}-\vec{r}')} d^3k = (2\pi)^3 \delta(\vec{r}-\vec{r}')$$

Fourier Transform

We can always express a function $f(x)$ in the form

$$f(x) = \int_{-\infty}^{\infty} e^{ikx} \tilde{f}(k) dk \quad (1.174)$$

where $\tilde{f}(k)$ is a function of k , called the fourier transform of $f(x)$. From eqn(26) we can write

$$\begin{aligned} \int_{-\infty}^{\infty} e^{-ik'x} f(x) dx &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{i(k-k')x} \tilde{f}(k) dk dx \\ &= (2\pi) \int_{-\infty}^{\infty} \delta(k-k') \tilde{f}(k) dk \\ &= (2\pi) \tilde{f}(k') \end{aligned}$$

Thus,

$$\tilde{f}(k) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-ikx} f(x) dx \quad (1.175)$$

The function $f(x)$ and $\tilde{f}(k)$ are Fourier transform of each other. We can write eqn(26) and eqn(27) in a more symmetrical fashion as follows

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{ikx} \tilde{f}(k) dk \quad (1.176)$$

$$\tilde{f}(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-ikx} f(x) dx \quad (1.177)$$

In three dimension, we can write

$$f(\vec{r}) = \frac{1}{(2\pi)^{3/2}} \int_{\text{all } k \text{ space}} e^{i\vec{k} \cdot \vec{r}} \tilde{f}(\vec{k}) d^3 k \quad (1.178)$$

Multiplying this equation by $e^{-i\vec{k}' \cdot \vec{r}}$ and integrating over \vec{r} , we have

$$\begin{aligned} \int_{\text{all space}} e^{-i\vec{k}' \cdot \vec{r}} f(\vec{r}) d^3 r &= \frac{1}{(2\pi)^{3/2}} \int d^3 r \int d^3 k e^{i(\vec{k} - \vec{k}') \cdot \vec{r}} \tilde{f}(\vec{k}) \\ &= \frac{1}{(2\pi)^{3/2}} \int d^3 k (2\pi)^3 \delta(\vec{k} - \vec{k}') \tilde{f}(\vec{k}) \\ &= (2\pi)^{3/2} \tilde{f}(\vec{k}') \end{aligned}$$

Therefore,

$$\tilde{f}(\vec{k}) = \frac{1}{(2\pi)^{3/2}} \int e^{-i\vec{k} \cdot \vec{r}} f(\vec{r}) d^3 r$$

Thus, if we write

$$f(\vec{r}) = \frac{1}{(2\pi)^{3/2}} \int e^{i\vec{k} \cdot \vec{r}} \tilde{f}(\vec{k}) d^3 k$$

then

$$\tilde{f}(\vec{k}) = \frac{1}{(2\pi)^{3/2}} \int e^{-i\vec{k} \cdot \vec{r}} \tilde{f}(\vec{r}) d^3 r$$

Parseval Identity

We can now prove the important identity

$$\int |f(\vec{r})|^2 d^3 r = \int |\tilde{f}(\vec{k})|^2 d^3 k$$

PROOF:

$$\begin{aligned} \int |f(\vec{r})|^2 d^3 r &= \int f(\vec{r}) f^*(\vec{r}) d^3 r \\ &= \int d^3 r \cdot \frac{1}{(2\pi)^{3/2}} \int e^{i\vec{k} \cdot \vec{r}} \tilde{f}(\vec{k}) d^3 k \cdot \frac{1}{(2\pi)^{3/2}} \int e^{-i\vec{k}' \cdot \vec{r}} \tilde{f}^*(\vec{k}') d^3 k' \\ &= \frac{1}{(2\pi)^3} \int d^3 k d^3 k' \tilde{f}(\vec{k}) \tilde{f}^*(\vec{k}') \int d^3 r e^{i(\vec{k} - \vec{k}') \cdot \vec{r}} (2\pi)^3 \delta(\vec{k} - \vec{k}') \\ &= \int d^3 k d^3 k' \tilde{f}(\vec{k}) \tilde{f}^*(\vec{k}') \delta(\vec{k} - \vec{k}') \\ &= \int d^3 k \tilde{f}(\vec{k}) \tilde{f}^*(\vec{k}) \\ &= \int |\tilde{f}(\vec{k})|^2 d^3 k \end{aligned}$$

2. sheet-2 : Linear Vector Space

References

1. Quantum Mechanics - Shanker
2. Quantum Mechanics - Sakurai

2.1 Definition

A linear vector space V is a collection of objects ψ_a, ψ_b, \dots , called vectors, which satisfy the following postulates:

1. If ψ_a and ψ_b are vectors in V , there is a unique vector $\psi_a + \psi_b$ in V , called the sum of ψ_a and ψ_b . In other words, an operation called addition is defined in the vector space such that the space is closed under addition.
2. The vector addition is commutative and associative, i.e.,

$$\psi_a + \psi_b = \psi_b + \psi_a \quad (2.1)$$

$$\psi_a + (\psi_b + \psi_c) = (\psi_a + \psi_b) + \psi_c \quad (2.2)$$

3. There is a vector in V called the null vector and denoted by ϕ satisfying

$$\psi_a + \phi = \phi + \psi_a \quad (2.3)$$

for every ψ_a in V .

4. For every vector ψ_a in V there is another vector ψ'_a in V such that

$$\psi_a + \psi'_a = \phi \quad (2.4)$$

we denote ψ'_a as $-\psi_a$.

(Note : we use the notation $\psi_a - \psi_b$ to mean $\psi_a + (-\psi_b)$).

5. If ψ_a is a vector and λ is an arbitrary number (real or complex), called a scalar, there is a uniquely defined vector $\lambda \psi_a$ in V satisfying.

(a)

$$\lambda(\psi_a + \psi_b) = \lambda\psi_a + \lambda\psi_b \quad (2.5)$$

i.e., multiplication is distributive with respect to vector addition.

(b)

$$(\lambda\mu)\psi_a = \lambda(\mu\psi_a) \quad (2.6)$$

i.e., multiplication by a scalar is associative.

(c)

$$(\lambda + \mu)\psi_a = \lambda\psi_a + \mu\psi_b \quad (2.7)$$

i.e., multiplication is distributive with respect to addition scalars.

- (d) Multiplication by scalars 0 and 1 are defined by

$$0\psi_a = \phi \quad (2.8)$$

$$1\psi_a = \psi_a \quad (2.9)$$

for any ψ_a in V .

2.2 Example of linear vector space

1. Consider all real numbers x in the range $-\infty$ to ∞ , i.e.,

$$-\infty < x < \infty \quad (2.10)$$

$$x \in \text{Re} \quad (\text{Re is the set of all real numbers}) \quad (2.11)$$

Take any two real numbers x_1 and x_2 . If we add two real numbers we get another real number in Re . Thus

$$x_1 + x_2 = \text{Re} \quad (2.12)$$

Next take any real number x . If we multiply x by another real number λ , we get a real number in Re , i.e.,

$$\lambda x \in \text{Re} \quad (2.13)$$

If we take a real number x , then there exists another real number $-x$ such that

$$x + (-x) = 0 \quad (2.14)$$

So the real numbers form a vector space with the real numbers themselves as vectors in the space. The number 0 is the null vector ϕ of the space. The scalars λ by which the vectors are multiplied are also real numbers.

Thus the real numbers form a real linear vector space over a field which are also real numbers. The addition and multiplication are just the normal addition and multiplication of the real numbers.

2. The set of $n - tuples$ of numbers (x_1, x_2, \dots, x_n) . When the addition of vectors and multiplication by a scalar are defined by

$$(x_1, x_2, \dots, x_n) + (y_1, y_2, \dots, y_n) = (x_1 + y_1, x_2 + y_2, \dots, x_n + y_n) \quad (2.15)$$

and

$$\lambda(x_1, x_2, \dots, x_n) = (\lambda x_1, \lambda x_2, \dots, \lambda x_n) \quad (2.16)$$

3. The collection of all square-integrable complex valued functions of a real variable form a vector space. consider all functions

$$f : \text{Re} \rightarrow \mathbb{C} \quad (2.17)$$

here, Re = set of real numbers

\mathbb{C} = set of complex numbers

such that

$$\int_{-\infty}^{\infty} f^*(x)f(x)dx \equiv \int_{-\infty}^{\infty} |f(x)|^2 dx < \infty \text{ (i.e., finite)} \quad (2.18)$$

The sum of two functions and the product of a function by a complex scalar are defined in the usual way.

The reason the square-integrable functions form a (complex) vector space is that the space is closed under addition. The vectors of the space are the square-integrable functions. In other words, it can be shown that if f and g are two vectors, in this case two functions $f(x)$ and $g(x)$ both of which are square integrable, then $f(x) + g(x)$ is also square integrable and have the sum belong to the vector space.

proof:

Let $f(x)$ and $g(x)$ be two square integrable function, i.e.,

$$\int_{-\infty}^{\infty} |f(x)|^2 dx < \infty \quad (2.19)$$

and

$$\int_{-\infty}^{\infty} |g(x)|^2 dx < \infty \quad (2.20)$$

then using the inequality

$$\int_{-\infty}^{\infty} |f+g|^2 dx \leq \left[\sqrt{\int_{-\infty}^{\infty} |f|^2 dx} \sqrt{\int_{-\infty}^{\infty} |g|^2 dx} \right]^2 \quad (2.21)$$

it is obvious that

$$\int_{-\infty}^{\infty} |f+g|^2 dx < \infty \quad (2.22)$$

(i.e., finite).

4. The set of all $n \times n$ matrices with complex elements form a complex linear vector space. For illustration let us take 2×2 matrix A

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad (2.23)$$

where the elements a, b, c and d can, in general, be complex. Then A belongs to a vector (complex) space. We have

$$\phi = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \quad (2.24)$$

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad (2.25)$$

$$-A = \begin{pmatrix} -a & -b \\ -c & -d \end{pmatrix} \quad (2.26)$$

$$\lambda A = \begin{pmatrix} \lambda a & \lambda b \\ \lambda c & \lambda d \end{pmatrix} \quad (2.27)$$

The set of all these matrices form a vector space.

2.3 Inner product space or a unitary vector space

For a general linear vector space, product of vectors (i.e. multiplication of two vectors) need not be defined. However, we will restrict ourselves to spaces in which a scalar product or an inner product is defined.

A linear vector space is called unitary if a scalar product is defined in it. To every pair of vectors ψ_a and ψ_b in V there corresponds a unique scalar (in general complex), called the scalar product. The

scalar product is defined to have the following properties :

$$(\psi_a, \psi_b) = (\psi_b, \psi_a)^* \quad (2.28)$$

$$(\psi_a, \lambda \psi_b) = \lambda (\psi_a, \psi_b) \quad (2.29)$$

$$(\lambda \psi_a, \psi_b) = \lambda^* (\psi_a, \psi_b) \quad (2.30)$$

$$(\psi_a, \psi_b + \psi_c) = (\psi_a, \psi_b) + (\psi_a, \psi_c) \quad (2.31)$$

$$(\psi_a, \psi_a) \geq 0 ; \text{the equality holds only if } \psi_a \text{ is the null vector} \quad (2.32)$$

It follows from the above postulated properties of the scalar product, that the scalar product is linear with respect to post factors, i.e.,

$$(\psi_a, \lambda \psi_b + \mu \psi_c) = \lambda (\psi_a, \psi_b) + \mu (\psi_a, \psi_c) \quad (2.33)$$

and anti-linear with respect to the pre factors, i.e.,

$$(\lambda \psi_a + \mu \psi_b, \psi_c) = \lambda^* (\psi_a, \psi_c) + \mu^* (\psi_b, \psi_c) \quad (2.34)$$

2.4 Examples of scalar product

Example 1 Consider the vector space consisting of all square integrable functions of a real variable in the domain $[a, b]$. This space is denoted by $L^2[a, b]$.

Suppose

$$f \in L^2[a, b] \quad (2.35)$$

i.e.,

$$\int_a^b f^*(x) f(x) dx \equiv \int_a^b |f(x)|^2 dx < \infty \quad (2.36)$$

We can define the scalar product of two vectors f and g as

$$(f, g) \stackrel{\text{def}}{=} \int_a^b f^*(x) g(x) dx = \text{complex number} \quad (2.37)$$

We can show

$$|(f, g)| = \left[\sqrt{\int_a^b |f(x)|^2 dx} \right] \left[\sqrt{\int_a^b |g(x)|^2 dx} \right] \quad (2.38)$$

Since both f and g are square integrable, $|(f, g)|$ is finite, i.e., the scalar product of f and g exists.

The scalar product defined above satisfies all the properties that a scalar product is postulated to have.

Example 2 Now consider the vector space consisting of n -tuples of complex numbers. Such a vector space is denoted as \mathbb{C}^n .

A vector $\psi_a \in \mathbb{C}^n$ may be expressed as

$$\psi_a = \begin{pmatrix} a_1 & a_2 & \dots & a_n \end{pmatrix}^T \quad (2.39)$$

The scalar product may then be defined as

$$(\psi_a, \psi_b) \stackrel{\text{def}}{=} \begin{pmatrix} a_1^* & a_2^* & \dots & a_n^* \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} \quad (2.40)$$

$$= \sum_{i=1}^n a_i^* b_i \quad (2.41)$$

This scalar product also satisfies all the properties of a scalar product.

Example 3 Euclidean 3-space \mathbb{R}^3 . The vectors of \mathbb{R}^3 are 3-tuples of real numbers which could be represented as column vectors. Thus if ψ_a and ψ_b are in \mathbb{R}^3 ,

$$\psi_a = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \quad (2.42)$$

$$\psi_b = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} \quad (2.43)$$

where a_i and b_i are real.

We could define the scalar product of ψ_a and ψ_b as

$$(\psi_a, \psi_b) \stackrel{\text{def}}{=} \begin{pmatrix} a_1 & a_2 & a_3 \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} \quad (2.44)$$

$$= \psi_a^T \psi_b \quad (2.45)$$

$$= \sum_{i=1}^3 a_i b_i \quad (2.46)$$

This scalar product also has all the postulated properties of a scalar product.

In case of the vector space \mathbb{R}^3 , the vectors ψ_a and ψ_b could be represented as directed lines \vec{a} and \vec{b} in a three dimensional coordinate system.

figure

The Scalar product (ψ_a, ψ_b) is the usual dot product

$$\vec{a} \cdot \vec{b} = a_1 b_1 + a_2 b_2 + a_3 b_3 \quad (2.47)$$

$$= |\vec{a}| |\vec{b}| \cos(\theta) \quad (2.48)$$

Where $|\vec{a}|$ and $|\vec{b}|$ are the magnitudes of the vectors \vec{a} and \vec{b} defined as

$$|\vec{a}| \equiv \sqrt{(\psi_a, \psi_a)} = \sqrt{a_1^2 + a_2^2 + a_3^2} \quad (2.49)$$

$$|\vec{b}| \equiv \sqrt{(\psi_b, \psi_b)} = \sqrt{b_1^2 + b_2^2 + b_3^2} \quad (2.50)$$

2.5 Norm of a vector

If a vector space is enclosed with a scalar product, then the scalar product gives us the concept of the "magnitude" or "length" of a vector. In a general vector space the "magnitude" or "length" of a vector is called the norm of the vector. We simply define the norm of a vector ψ_a as

$$\|\psi_a\| \stackrel{\text{def}}{=} \sqrt{(\psi_a, \psi_a)} \quad (2.51)$$

The norm has the following properties:

1.

$$\|\psi_a\| \geq 0 \quad (2.52)$$

the equality holds only if the vector is null.

2.

$$\|\psi_a + \psi_b\| \leq \|\psi_a\| + \|\psi_b\| \quad (2.53)$$

This is called the triangle inequality

3.

$$\|\psi_a - \psi_b\| = \|\psi_b - \psi_a\| \quad (2.54)$$

2.6 Metric included by the scalar product

The norm included by the scalar product allows us to develop the concept of "distance" between vectors in a vector space. We say two vectors ψ_a and ψ_b are 'close' if $\|\psi_a - \psi_b\|$ is small. The metric in a vector space assigns a real number to the vector $\psi_a - \psi_b$. This real number is a measure of how close the two vectors are. We simply define the metric $\mathbf{d}(\psi_a, \psi_b)$ as

$$\mathbf{d}(\psi_a, \psi_b) \stackrel{\text{def}}{=} \|\psi_a - \psi_b\| \quad (2.55)$$

Thus, if there are three vectors ψ_a , ψ_b and ψ_c and if $\mathbf{d}(\psi_a, \psi_b) < \mathbf{d}(\psi_a, \psi_c)$ then we say ψ_a is closer to ψ_b than to ψ_c .

2.7 Schwarz's inequality

We will now prove a very important inequality called Schwarz inequality which states

$$|(\psi_a, \psi_b)| \leq \sqrt{(\psi_a, \psi_a)(\psi_b, \psi_b)} \quad (2.56)$$

or

$$|(\psi_a, \psi_b)| \leq \|\psi_a\| \|\psi_b\| \quad (2.57)$$

proof : Let

$$\psi = \psi_a + \lambda \psi_b \quad (2.58)$$

Then

$$(\psi, \psi) = (\psi_a + \lambda \psi_b, \psi_a + \lambda \psi_b) \quad (2.59)$$

$$= (\psi_a, \psi_a) + \lambda (\psi_a, \psi_b) + \lambda^* (\psi_b, \psi_a) + |\lambda|^2 (\psi_b, \psi_b) \geq 0 \quad (2.60)$$

The best inequality is obtained if λ is chosen so as to minimize the left hand side of the above equation. By differentiation, the value of λ which accomplishes this is found to be

$$\lambda = -\frac{(\psi_b, \psi_a)}{\psi_b, \psi_b} \quad (2.61)$$

Substituting this value of λ in the above equation yields the Schwarz inequality.

We note that the equality sign holds if and only if $(\psi, \psi) = 0$, i.e., ψ is the null vector, i.e., $\psi = \phi$, in other words

$$\psi_a + \lambda \psi_b = \phi(\text{null}) \quad (2.62)$$

i.e.,

$$\psi_a = -\lambda \psi_b + \phi \quad (2.63)$$

$$\psi_a = -\lambda \psi_b \quad (2.64)$$

Hence, the equality holds if ψ_a and ψ_b are multiple of each other, or if ψ_a and ψ_b are "parallel".

It follows from the Schwarz inequality that the scalar product (ψ_a, ψ_b) is finite if the norms of ψ_a and ψ_b are finite.

2.8 Analogy of Schwarz inequality with vectors in a three-dimensional Euclidean space \mathbb{R}^3

In \mathbb{R}^3 , the vectors can be represented by directed lines (i.e. arrows). We have the scalar product ordinary vectors in the form

$$\vec{A} \cdot \vec{B} = |\vec{A}| |\vec{B}| \cos(\theta) \quad (2.65)$$

Since cosine of any angle lies between -1 and $+1$, we have

$$|\vec{A} \cdot \vec{B}| \leq |\vec{A}| |\vec{B}| \quad (2.66)$$

The analogue of this equation for a general vector space is the Schwarz inequality

$$|(\psi_a, \psi_b)| \leq \|\psi_a\| \|\psi_b\| \quad (2.67)$$

2.9 Orthogonality and linear independence

A vector whose norm is unity is called a unit vector. For any given non-null vector, a unit vector can be formed by defining the vector by its norm. Thus

$$u_a = \frac{\psi_a}{\|\psi_a\|} \quad (2.68)$$

is normalized.

Two vectors ψ_a and ψ_b are orthogonal if their inverse product is zero, i.e., if

$$(\psi_a, \psi_b) = 0 \quad (2.69)$$

The unit vectors u_1, u_2, \dots, u_N form an orthogonal set if they are mutually orthonormal , i.e.,

$$(u_i, u_j) = \delta_{ij} \quad (2.70)$$

2.9.1 Linear independence

The set of vectors $\psi_1, \psi_2, \dots, \psi_N$ are linearly independent if none of them can be expressed as a linear combination of the others. Mathematically this means that the equation

$$\sum_{j=1}^N c_j \psi_j = 0 \quad (2.71)$$

cannot be satisfied except by $c_j = 0$ for all j .

2.10 Orthonormality and linear independence

A dot of mutually orthogonal vectors (not necessarily normalized) are necessarily linearly independent. The converse is not true, however. That is, a set of linearly independent vectors may not be mutually orthogonal.

It is always possible to orthonormalize a set of linearly independent vectors. By this we mean that from a given set of N linearly independent vectors, it is possible to form a set of N orthonormal vectors. This procedure is called **Schmidt orthonormalization method**.

2.11 Schmidt orthonormalization method

Suppose $\psi_1, \psi_2, \dots, \psi_N$ is a set of linearly independent vectors. Let

$$u_1 = \frac{\psi_1}{\|\psi_1\|} \quad (2.72)$$

Then $(u_1, u_1) = 1$, i.e., u_1 is normalized. Next construct the vector ψ'_2 as follows:

$$\psi'_2 = \psi_2 - u_1(u_1, \psi_2) \quad (2.73)$$

i.e., to obtain ψ'_2 we have subtracted the 'component' of ψ_2 along the u_1 "direction". Then it follows that

$$(u_1, \psi'_2) = (u_1, \psi_2) - (u_1, u_1)(u_1, \psi_2) \quad (2.74)$$

$$= (u_1, \psi_2) - (u_1, \psi_2) \quad (2.75)$$

$$= 0 \quad (2.76)$$

i.e., ψ'_2 is orthogonal to u_1 . We then normalize ψ'_2 , i.e.,

$$u_2 = \frac{\psi'_2}{\|\psi'_2\|} \quad (2.77)$$

We can continue the process until we exhaust all the vectors. For example, in the next step we can write

$$\psi'_3 = \psi_3 - u_1(u_1, \psi_3) - u_2(u_2, \psi_3) \quad (2.78)$$

We note immediately that ψ'_3 is orthogonal to both u_1 and u_2 , i.e.,

$$(u_1, \psi'_3) = (u_2, \psi'_3) = 0 \quad (2.79)$$

We normalize ψ'_3 to get u_3 , i.e.,

$$u_3 = \frac{\psi'_3}{\|\psi'_3\|} \quad (2.80)$$

Finally, in the $N - th$ step, we write

$$\psi'_N = \psi_N - u_1(u_1, \psi_N) - u_2(u_2, \psi_N) - \dots - u_{N-1}(u_{N-1}, \psi_N) \quad (2.81)$$

ψ'_N is orthogonal to u_1, u_2, \dots, u_{N-1} , i.e.,

$$(u_1, \psi'_N) = (u_2, \psi'_N) = \dots = (u_{N-1}, \psi'_N) = 0 \quad (2.82)$$

Normalizing ψ'_N we get

$$u_N = \frac{\psi'_N}{\|\psi'_N\|} \quad (2.83)$$

Thus, the set $\{u_1, u_2, \dots, u_N\}$ is an orthonormal set of vectors.

2.12 Dimension of a vector space

The vector space V is said to be n-dimensional if there exists n linearly independent vectors, but if $n + 1$ vectors are linearly dependent the dimension may be finite or infinite.

2.13 Complete vector space

Before defining what a complete vector space is we will give some other definitions.

A sequence of vectors $\{\psi_n\}$ in the vector space V is called a **Cauchy sequence** if for every $\varepsilon > 0$ there exists an integer N such that

$$\|\psi_n - \psi_m\| < \varepsilon \quad (2.84)$$

if $n, m > N$. In other words, the vectors in the sequence come 'closer' if the index increases. In particular

$$\|\psi_n - \psi_m\| \rightarrow 0 \text{ as } n, m \rightarrow \infty \quad (2.85)$$

2.14 Convergence of a sequence of vectors in a vector space

A sequence $\{\psi_n\}$, $n = 1, 2, \dots$ in a vector space V converges to a vector ψ in V if for every $\varepsilon > 0$ there exists an integer N such that

$$\|\psi_n - \psi_m\| < \varepsilon \quad (2.86)$$

if $n > N$, that is if

$$\lim_{n \rightarrow \infty} \|\psi - \psi_n\| = 0 \quad (2.87)$$

then

$$\{\psi_n\} \rightarrow \psi \quad (2.88)$$

and the sequence is called a convergent sequence.

Now we can show every convergent sequence is a Cauchy sequence.

proof

Let $\{\psi_n\} \rightarrow \psi$. Then

$$\|\psi_n - \psi_m\| = \|\psi_n - \psi + \psi - \psi_m\| \quad (2.89)$$

$$\leq \|\psi_n - \psi\| + \|\psi - \psi_m\| \quad (2.90)$$

here the triangle inequality is used.

Since $\{\psi_n\}$ is a convergent sequence, each term on the right hand side tends to zero as n and m tends to infinity. Hence $\|\psi_n - \psi_m\| \rightarrow 0$ as $n, m = \infty$, i.e., the sequence $\{\psi_i\}$ with $i = 1, 2, \dots$ is a Cauchy sequence.

The converse of the above statement is not true in general. In other words, *a Cauchy sequence in a vector space may not converge to a vector in the space*. It can be shown that for a finite dimensional vector space the converse is true, i.e., in a finite dimensional vector space a Cauchy sequence is always a convergent sequence. Exceptions may arise in infinite-dimensional vector space.

2.14.1 An example of a vector space whose Cauchy sequence does not converge to a vector in the vector space

Consider the vector space consisting of all continuous function of a single real variable x in the range $[-1, 1]$. In this vector space consider a sequence $\{f_k(x)\}$, $k = 1, 2, \dots$ of the following form:

$$f_k(x) = \begin{cases} 1 & \text{for } \frac{1}{k} \leq x \leq 1 \\ \frac{kx+1}{2} & \text{for } -\frac{1}{k} < x < \frac{1}{k} \\ 0 & \text{for } -1 \leq x \leq -\frac{1}{k} \end{cases} \quad (2.91)$$

$k = 1, 2, 3, \dots$

The graph of the sequence of functions is shown below:

figure

Note that, in this example each $f_k(x)$ is continuous, but their first derivatives are discontinuous.

Let us define the scalar product in this space as

$$(f, g) = \int_{-1}^1 f^*(x)g(x)dx \quad (2.92)$$

So that the metric $\mathbf{d}(f, g)$, i.e., the "distance" between vectors f and g can be defined as

$$\mathbf{d}(f, g) \equiv \|f - g\| \quad (2.93)$$

$$= \sqrt{\int_{-1}^{+1} (f^*(x) - g^*(x))(f(x) - g(x)) dx} \quad (2.94)$$

With this metric we can show that the sequence f_k defined above is indeed a Cauchy sequence. However, looking at the graph above, we see that as k becomes large, f_k approaches the θ function

$$\theta(x) = \begin{cases} 0 & \text{for } x < 0 \\ 1 & \text{for } x > 0 \end{cases} \quad (2.95)$$

Which is a discontinuous function at $x = 0$.

We show that graph of $\theta(x)$ below:

figure

Thus the Cauchy sequence $\{f_k(x)\}$ of continuous function is converging to a discontinuous function which lies outside the vector space V .

If, instead of choosing all continuous function, we had chosen all square integrable function as defining the vector space, then any Cauchy sequence in the vector space would converge to a vector in the space.

Definition : A linear vector space is said to be complete if any Cauchy sequence converges to a vector in the space.

2.15 Hilbert space

A complete linear vector space, finite or infinite dimensional, (and hence @@@@enclosed@@@ with a norm and metric induced by the scalar product), is called a Hilbert space.

A finite-dimensional vector space is always complete. So, a finite dimensional linear vector space in which a scalar product is defined is a Hilbert space.

An infinite dimensional vector space with a scalar product may or may not be complete. Whether or not an infinite dimensional vector space is complete depends upon how exactly the vector space is defined and on the metric.

2.16 Basis vectors in a Hilbert space

2.16.1 Finite dimensional space

In a finite dimensional vector space of dimension n , any set of linearly independent vectors $\psi_1, \psi_2, \dots, \psi_n$ spans the entire space. In other words any vector ψ in the space can be expressed as

linear combination of $\psi_1, \psi_2, \dots, \psi_n$, i.e.,

$$\psi = \sum_{i=1}^n a_i \psi_i \quad (2.96)$$

The vectors $\psi_1, \psi_2, \dots, \psi_n$ form a complete basis for the vector space. The vectors $\psi_1, \psi_2, \dots, \psi_n$, even if linearly independent, may not be orthogonal to each other. It is more convenient to use a set of orthonormal vectors $\phi_1, \phi_2, \dots, \phi_n$ as the basis. Being orthogonal, the vectors $\phi_1, \phi_2, \dots, \phi_n$ are automatically linearly independent. The orthonormal set of basis vectors $\{\phi_i\} i = 1, 2, \dots, n$ can be constructed from the set $\{\psi_i\} i = 1, 2, \dots, n$ by using the schmidt orthonormalization procedure.

Choosing the orthonormal set as the basis, any vector ψ in the vector space can be written as

$$\psi = \sum_{i=1}^n a_i \phi_i \quad (2.97)$$

where

$$(\phi_i, \phi_j) = \delta_{ij} \quad (2.98)$$

using equation 2.98 we have

$$a_i = (\phi_i, \psi) \quad (2.99)$$

2.16.2 Infinite dimensional vector space

In an infinite dimensional vector space the number of basis vectors is infinity. Let $\{\phi_1, \phi_2, \dots\}$ be an infinite set of orthonormal basis vectors spanning the infinite dimensional Hilbert space. This set of basis vectors is said to be complete if any vector ψ in the Hilbert space can be expanded as a linear combination of the basis vectors, i.e.,

$$\psi = \sum_{i=1}^{\infty} a_i \phi_i \quad (2.100)$$

In an infinite dimensional vector space, choosing an infinite number of basis vectors may not ensure that the basis set is complete. It may so happen that there are other linearly independent vectors, may be infinite in numbers, which have been missed in the first choice of the basis vectors.

Whenever we have an infinite sum, as in equation 2.100 is to be understood in the sense that the sequences consisting of the partial sums

$$f_n = \sum_{i=1}^n a_i \phi_i \quad (2.101)$$

$$n = 1, 2, 3, \dots$$

converges to ψ , i.e.,

$$\lim_{n \rightarrow \infty} \|\psi - f_n\| \rightarrow 0 \quad (2.102)$$

Since the vector ψ must have a finite norm, we must have

$$\|\psi\|^2 = (\psi, \psi) = \sum_{i=1}^{\infty} |a_i|^2 < \infty \text{ (finite)} \quad (2.103)$$

If the basis vectors $\{\phi_i\}$ are orthonormal, we have

$$a_i = (\phi_i, \psi) \quad (2.104)$$

So that equation 2.103 can be written as

$$\sum_{i=1}^{\infty} |(\phi_i, \psi)|^2 < \infty \quad (2.105)$$

The scalar a_i can be regarded as the components of ψ in the 'directions' of ϕ_i

2.16.3 Example

Question : Show that the set of all square integrable functions, i.e., set of all functions f such that

$$\int_{-\infty}^{\infty} f^*(x)f(x)dx < \infty \text{ (i.e., finite)} \quad (2.106)$$

belong to a Hilbert space. This Hilbert space is denoted as $L^2(-\infty, \infty)$.

To show this, consider the following

1. If f and g are square integrable functions, so is $f + g$, and hence $f + g$ also belongs to the Hilbert space

$$\|f + g\| \leq \|f\| + \|g\| \quad (2.107)$$

2. We can define the scalar product between f and g as follows:

$$(f, g) \stackrel{def}{=} \int_{-\infty}^{\infty} f^*(x)g(x)dx \quad (2.108)$$

That the scalar product exists follows from the Schwarz inequality

$$|(f, g)| \leq \|f\| \cdot \|g\|, < \infty \quad (2.109)$$

3. It can also be shown that any Cauchy sequence of square integrable functions converges to a limit which is also square integrable. In other words, the space of all square integrable functions is complete.

Hence the linear vector space consisting of all square integrable functions is indeed a Hilbert space.

2.17 Dirac notation

(Cohen-Tannoudji; page 109)

"ket" vectors and "bra" vectors.

Notation

Any element, or vector of a vector space V is called a ket vector, or more simply, a ket. It is represented by the symbol $|\rangle$, inside which is placed a distinctive sign which enables us to distinguish between different kets, for example $|\psi\rangle$.

2.17.1 Scalar product

With each pair of kets $|\phi\rangle$ and $|\psi\rangle$, taken in this order, we associate a complex number, which is their scalar product $(|\phi\rangle, |\psi\rangle)$ and which satisfies various properties discussed earlier (section 2.3).

2.17.2 Dual vector space

Linear functional: A linear functional χ is a linear operation on the kets such that χ operating on a ket $|\psi\rangle$ gives a complex scalar:

$$\chi|\psi\rangle \rightarrow \text{scalar where } |\psi\rangle \in V \quad (2.110)$$

and

$$\chi(\lambda_1|\psi\rangle + \lambda_2|\psi_2\rangle) = \lambda_1\chi|\psi_1\rangle + \lambda_2\chi|\psi_2\rangle \quad (2.111)$$

The set of all linear functions defined in the kets of a vector space V themselves form a linear vector space called the dual space of V and symbolized by V^* .

2.17.3 Bra notation for the vectors of V^*

Any element, or vector, of the space V^* is called a "bra vector", or more simply, a bra. It is symbolized by $\langle|$. For example, the bra $\langle\chi|$ designates the linear functional χ we shall henceforth use the notation $\langle\chi|\psi\rangle$ to denote the number obtained by causing the linear functional $\langle\chi| \in V^*$ to act on the ket $|\psi\rangle \in V$. Thus

$$\chi(|\psi\rangle) = \langle\chi|\psi\rangle \quad (2.112)$$

2.17.4 Correspondence between kets and bras

The existence of the scalar product in V will now enable us to show that we can associate with every ket $|\phi\rangle \in V$ and element of V^* , that is a bra, which will be denoted by $\langle\phi|$.

The ket $|\phi\rangle$ does indeed enable us to define a linear functional, the one which associates with each $|\psi\rangle \in V$ a complex number which is equal to the scalar product $(|\phi\rangle, |\psi\rangle)$. Let $\langle\phi|$ be this linear functional. It is thus defined by the relation

$$\langle\phi|\psi\rangle = (|\phi\rangle, |\psi\rangle) \quad (2.113)$$

2.17.5 The correspondence in anti-linear

Let $\lambda_1 |\phi_1\rangle + \lambda_2 |\phi_2\rangle$ be a ket. Then

$$(\lambda_1 |\phi_1\rangle + \lambda_2 |\phi_2\rangle, \psi) = \lambda_1^* (|\phi_1\rangle, \psi) + \lambda_2^* (|\phi_2\rangle, \psi) \quad (2.114)$$

$$= \lambda_1^* \langle\phi_1|\psi\rangle + \lambda_2^* \langle\phi_2|\psi\rangle \quad (2.115)$$

$$= (\lambda_1^* \langle\phi_1| + \lambda_2^* \langle\phi_2|) |\psi\rangle \quad (2.116)$$

Thus

$$\lambda_1 |\phi_1\rangle + \lambda_2 |\phi_2\rangle \xrightarrow{dc} \lambda_1^* \langle\phi_1| + \lambda_2^* \langle\phi_2| \quad (2.117)$$

Where "dc" is short for dual correspondence.

comment

If λ is a complex number and $|\psi\rangle$ is a ket, then $\lambda |\psi\rangle$ is also a ket. We are sometimes led to write $\lambda |\psi\rangle$ as $|\lambda\psi\rangle$:

$$|\lambda\psi\rangle = \lambda |\psi\rangle \quad (2.118)$$

One must be careful to remember that $\langle\lambda\psi|$ represents the bra associated with the ket $|\lambda\psi\rangle$, since the correspondence between a bra and a ket is anti-linear we have

$$\langle\lambda\psi| = \lambda^* \langle\psi| \quad (2.119)$$

2.17.6 Dirac notation for the scalar product

We now have at our disposal two distinct notations for designating the scalar product of $|\psi\rangle$ by $|\phi\rangle$, namely, $(|\phi\rangle, |\psi\rangle)$ and $\langle\phi|\psi\rangle$, $\langle\phi|$ being the bra associated with the ket $|\phi\rangle$. We shall mostly use the Dirac notation $\langle\phi|\psi\rangle$. In the table below we summarize, in Dirac notation, the properties of the scalar product.

$$\langle\phi|\psi\rangle = \langle\psi|\phi\rangle^* \quad (2.120)$$

$$\langle\phi|\lambda_1\psi_1 + \lambda_2\psi_2\rangle = \lambda_1 \langle\phi|\psi_1\rangle + \lambda_2 \langle\phi|\psi_2\rangle \quad (2.121)$$

$$\langle\lambda_1\phi_1 + \lambda_2\phi_2|\psi\rangle = \lambda_1^* \langle\phi_1|\psi\rangle + \lambda_2^* \langle\phi_2|\psi\rangle \quad (2.122)$$

$$(2.123)$$

$\langle\psi|\psi\rangle$ is real, positive; zero if and only if $|\psi\rangle = \phi$ (null)

3. sheet-3 : Linear Vector Space (continued)

3.1 Operators in a Hilbert Space

An operator is a prescription by which every vector ψ_a in a Hilbert space H is associated with another vector ψ_b in the space:

$$\hat{A} : \psi_1 \rightarrow \psi_b \quad (3.1)$$

for $\psi_a, \psi_b \in H$. We usually employ the notation

$$\psi_b = \hat{A}\psi_a \quad (3.2)$$

In Dirac notation, we write

$$|b\rangle = \hat{A}|a\rangle \quad (3.3)$$

where both $|a\rangle$ and $|b\rangle$ belong to the ket-space. An operator can also act on a bra vector (bra-space is also a Hilbert space; it is dual to the ket space) changing it to another bra-vector. The notation we employ is

$$\langle\psi| = \langle\phi|\hat{A} \quad (3.4)$$

Here the operator \hat{A} acts on the bra-vector $\langle\phi|$ to produce the bra vector $\langle\psi|$. We place the bra-vector on which the operator acts on the left of the operator.

3.2 Linear operators

An operator \hat{A} is said to be a linear operator if it has the following property :

1. For any vector $|a\rangle$ and $|b\rangle$ and any complex number λ_1 and λ_2 , we have

$$\hat{A}(\lambda_1|a\rangle + \lambda_2|b\rangle) = \lambda_1\hat{A}|a\rangle + \lambda_2\hat{A}|b\rangle \quad (3.5)$$

A linear vector operator can act on a bra vector also

$$(\lambda_1\langle a| + \lambda_2\langle b|)\hat{A} = \lambda_1\langle a|\hat{A} + \lambda_2\langle b|\hat{A} \quad (3.6)$$

2. The operator \hat{A} is anti linear if

$$\hat{A}(\lambda_1|a\rangle + \lambda_2|b\rangle) = \lambda_1^*\hat{A}|a\rangle + \lambda_2^*\hat{A}|b\rangle \quad (3.7)$$

3. Two operators \hat{A} and \hat{B} are equal if

$$\hat{A}|\psi\rangle = \hat{B}|\psi\rangle \quad (3.8)$$

for all $|\psi\rangle$ in the vector space.

4. Sum of two operators \hat{A} and \hat{B} is defined as

$$(\hat{A} + \hat{B})|\psi\rangle = \hat{A}|\psi\rangle + \hat{B}|\psi\rangle \quad (3.9)$$

5. Product of two operators \hat{A} and \hat{B} is defined as

$$(\hat{A}\hat{B})|\psi\rangle = \hat{A}(\hat{B}|\psi\rangle) \quad (3.10)$$

This equation says that the operator $\hat{A}\hat{B}$ acting on $|\psi\rangle$ produces the same vector which would be obtained if we first let \hat{B} act on $|\psi\rangle$ and then \hat{A} acts on the resultant of the previous operation. In general $\hat{A}\hat{B} \neq \hat{B}\hat{A}$, although in exceptional cases we may have $\hat{A}\hat{B} = \hat{B}\hat{A}$.

3.3 Commutator of two operators

The commutator of two operators \hat{A} and \hat{B} is defined as

$$[\hat{A}, \hat{B}] \stackrel{\text{def}}{=} \hat{A}\hat{B} - \hat{B}\hat{A} \quad (3.11)$$

In general $[\hat{A}, \hat{B}] \neq 0$ (null operator). If $[\hat{A}, \hat{B}] = 0$, we say that \hat{A} and \hat{B} commute with each other.

3.3.1 Some properties of commutators

$$[\hat{A}, \lambda\hat{B}] = \lambda[\hat{A}, \hat{B}] \quad (3.12)$$

$$[\lambda\hat{A}, \hat{B}] = \lambda[\hat{A}, \hat{B}] \quad (3.13)$$

$$[\hat{A}, \hat{B} + \hat{C}] = [\hat{A}, \hat{B}] + [\hat{A}, \hat{C}] \quad (3.14)$$

$$[\hat{A}, \hat{B}] = -[\hat{B}, \hat{A}] \quad (3.15)$$

$$[\hat{A}, \hat{B}\hat{C}] = \hat{B}[\hat{A}, \hat{C}] + [\hat{A}, \hat{B}]\hat{C} \quad (3.16)$$

$$[\hat{A}\hat{B}, \hat{C}] = \hat{A}[\hat{B}, \hat{C}] + [\hat{A}, \hat{C}]\hat{B} \quad (3.17)$$

3.4 Projection operator

(An important example of a linear operator).

Consider the operator \hat{P}_a defined as

$$\hat{P}_a = |a\rangle\langle a| \quad (3.18)$$

where

$$\langle a|b\rangle = 1 \quad (3.19)$$

Operating by \hat{P}_a on an arbitrary ket $|\psi\rangle$, we have

$$\hat{P}_a |\psi\rangle = |a\rangle\langle a|\psi\rangle \quad (3.20)$$

i.e., \hat{P}_a projects the ket $|\psi\rangle$ along $|a\rangle$. The complex number $\langle a|\psi\rangle$ is the component of $|\psi\rangle$ along $|a\rangle$.

Now, $|\hat{P}_a\rangle$ is a linear operator. To show this consider

$$\hat{P}_a (\lambda_1 |\psi_1\rangle + \lambda_2 |\psi_2\rangle) = |a\rangle\langle a|(\lambda_1 |\psi_1\rangle + \lambda_2 |\psi_2\rangle) \quad (3.21)$$

$$= \lambda_1 |a\rangle\langle a|\psi_1\rangle + \lambda_2 |a\rangle\langle a|\psi_2\rangle \quad (3.22)$$

$$= \lambda_1 \hat{P}_a |\psi_1\rangle + \lambda_2 \hat{P}_a |\psi_2\rangle \quad (3.23)$$

Another important property of the projection operator is

$$\hat{P}_a^2 = \hat{P}_a \quad (3.24)$$

To prove this allow \hat{P}_a^2 to act on a ket.

$$\hat{P}_a^2 |\psi\rangle = \hat{P}_a \hat{P}_a |\psi\rangle \quad (3.25)$$

$$= \hat{P}_a |a\rangle\langle a|\psi\rangle \quad (3.26)$$

$$= |a\rangle\langle a|a\rangle\langle a|\psi\rangle \quad (3.27)$$

$$= |a\rangle\langle a|\psi\rangle \quad (3.28)$$

$$= \hat{P}_a |\psi\rangle \quad (3.29)$$

Exercise Six operator are defined as follows:

$$\hat{A}_1 \psi(x) = [\psi(x)]^2 \quad (3.30)$$

$$\hat{A}_2 \psi(x) = \frac{d}{dx} \psi(x) \quad (3.31)$$

$$\hat{A}_3 \psi(x) = \int_a^x \psi(x') dx' \quad (3.32)$$

$$\hat{A}_4 \psi(x) = x^2 \psi(x) \quad (3.33)$$

$$\hat{A}_5 \psi(x) = \sin(\psi(x)) \quad (3.34)$$

$$\hat{A}_6 \psi(x) = \frac{d^2}{dx^2} \psi(x) \quad (3.35)$$

which of these operators \hat{A}_i are linear operator.

3.5 Representation of vectors and operators

Let ϕ_i be a complete orthonormal basis set in a Hilbert space. Since the basis is orthonormal, we must have

$$(\phi_i, \phi_j) = \delta_{ij} \quad (3.36)$$

An arbitrary vector ψ_a can be written as a linear combination of the basis vectors.

We write

$$\psi_a = \sum_i a_i \phi_i \quad (3.37)$$

Where the scalars a_i are the components of the vector ψ_a along the basis vectors ϕ_i . Using the orthonormality of the basis vectors we immediately have

$$a_i = (\phi_i, \psi_a) \quad (3.38)$$

We can arrange these numbers as a column matrix

$$\begin{pmatrix} a_1 \\ a_2 \\ \vdots \end{pmatrix} \equiv \begin{pmatrix} (\phi_1, \psi_a) \\ (\phi_2, \psi_a) \\ \vdots \end{pmatrix} \quad (3.39)$$

This column matrix is called the representation of the vector ψ_a with respect to the given basis $\{\phi_i\}$. In Dirac notation we represent the vector ψ_a as $|a\rangle$ and the basis vectors ϕ_i are written as $|i\rangle$. We can expand a general ket $|a\rangle$ as a linear combination of the basis kets:

$$|a\rangle = \sum_{i=1}^{\infty} a_i |i\rangle \quad (3.40)$$

Orthonormality of the basis kets can be written as

$$\langle i | j \rangle = \delta_{ij} \quad (3.41)$$

The complex scalar a_i are called the components of the ket $|a\rangle$ along $|i\rangle$. Using the orthonormality condition of the basis vectors (3.40), we have

$$a_i = \langle i | \psi \rangle \quad (3.42)$$

These scalars a_1, a_2, \dots arranged as a column matrix is called the representation of $|a\rangle$ in the basis $\{|i\rangle\}, i = 1, 2, \dots$

Thus

$$|a\rangle \rightarrow \begin{pmatrix} a_1 \\ a_2 \\ \vdots \end{pmatrix} \equiv \begin{pmatrix} \langle 1|a\rangle \\ \langle 2|a\rangle \\ \vdots \end{pmatrix} \quad (3.43)$$

We can write down the representation of any one of the basis vectors in the same basis as

$$|1\rangle \rightarrow \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \end{pmatrix} \quad |2\rangle \rightarrow \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \end{pmatrix} \quad |3\rangle \rightarrow \begin{pmatrix} 0 \\ 0 \\ 1 \\ \vdots \end{pmatrix} \quad (3.44)$$

and so on.

Now, using equation ?? in equation ?? we have

$$|a\rangle = \sum_i a_i |i\rangle \quad (3.45)$$

$$= \sum_i \langle i|a\rangle |i\rangle \quad (3.46)$$

$$= \sum_i |i\rangle \langle i|a\rangle \quad (3.47)$$

$$= \left(\sum_i \hat{P}_i \right) |a\rangle \quad (3.48)$$

where

$$\hat{P}_i = |i\rangle \langle i| \quad (3.49)$$

is the projection operator along $|i\rangle$. Since equation 3.45 is true for all $|a\rangle$ in the vector space (this is because $\{|i\rangle\}$ form a complete set), we must have

$$\sum_i \hat{P}_i = \sum_i |i\rangle \langle i| = \hat{\mathbb{I}} \quad (3.50)$$

where \mathbb{I} is the identity operator. Equation 3.50 is called the **completeness condition** for the basis vectors.

3.6 Matrix representation of ket and bra vectors

The ket vector $|a\rangle$ is represented by a column matrix

$$\begin{pmatrix} a_1 \\ a_2 \\ \vdots \end{pmatrix} \equiv \begin{pmatrix} \langle 1|a\rangle \\ \langle 2|a\rangle \\ \vdots \end{pmatrix} \quad (3.51)$$

in a basis $\{|i\rangle\}$. The dual of ket $|a\rangle$ is the $\langle a|$. But what is the matrix representation of the bra $\langle a|$ in the same basis? To see this we can expand $\langle a|$ as

$$\langle a| = \sum_i \langle a|i\rangle \langle i| \quad (3.52)$$

The $\langle a|$ is represented by a row vector:

$$\langle a| \rightarrow \begin{pmatrix} \langle a|1\rangle & \langle a|2\rangle & \dots \end{pmatrix} = \begin{pmatrix} a_1^* & a_2^* & \dots \end{pmatrix} \quad (3.53)$$

Then the scalar product becomes a number. Thus

$$\langle a|a\rangle = \begin{pmatrix} a_1^* & a_2^* & \dots \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ \vdots \end{pmatrix} \quad (3.54)$$

$$= a_1^*a_1 + a_2^*a_2 + \dots \quad (3.55)$$

$$= \sum_i a_i^*a_i \quad (3.56)$$

$$= \sum_i |a_i|^2 \quad (3.57)$$

Here the quantity $|a_i|^2$ is just a number.

In general

$$\langle b|a\rangle = \sum_i \langle b|i\rangle \langle i|a\rangle \quad (3.58)$$

$$= \begin{pmatrix} \langle b|1\rangle & \langle b|2\rangle & \dots \end{pmatrix} \begin{pmatrix} \langle 1|a\rangle \\ \langle 2|a\rangle \\ \vdots \end{pmatrix} \quad (3.59)$$

$$= \begin{pmatrix} b_1^* & b_2^* & \dots \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ \vdots \end{pmatrix} \quad (3.60)$$

$$= b_1^*a_1 + b_2^*a_2 + \dots \quad (3.61)$$

$$= \sum_i b_i^*a_i \quad (3.62)$$

Here the quantity $b_i^*a_i$ is complex number.

3.7 Representation of an operator in a basis

Consider the equation

$$|b\rangle = \hat{A}|a\rangle \quad (3.63)$$

Let $\{|i\rangle\}$ where $i = 1, 2, \dots$ be a complete set of orthonormal basis states. Taking the component of equation 3.63 along $|i\rangle$, we have

$$\langle i|b\rangle = \langle i|\hat{A}|a\rangle \quad (3.64)$$

$$= \sum_j \langle i|\hat{A}|j\rangle \langle j|a\rangle \quad (3.65)$$

In matrix notation

$$b_i = \sum_j A_{ij} a_j \quad (3.66)$$

where

$$b_i \equiv \langle i|b\rangle \quad (3.67)$$

$$a_j \equiv \langle j|a\rangle \quad (3.68)$$

$$A_{ij} \equiv \langle i|\hat{A}|j\rangle \quad (3.69)$$

Writing in full, equation 3.66 becomes

$$\begin{pmatrix} b_1 \\ b_2 \\ \cdot \\ \cdot \\ \cdot \end{pmatrix} = \begin{pmatrix} A_{11} & A_{12} & \dots \\ A_{21} & A_{22} & \dots \\ \cdot & \cdot & \dots \\ \cdot & \cdot & \dots \\ \cdot & \cdot & \dots \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ \cdot \\ \cdot \\ \cdot \end{pmatrix} \quad (3.70)$$

The matrix $[A]$ with elements $A_{ij} = \langle i|\hat{A}|j\rangle$ is called the matrix representation of the operator \hat{A} with respect to the given basis $\{|i\rangle\}$. Using a basis set, the operator \hat{A} can also be written as

$$\hat{A} = \hat{\mathbb{I}}\hat{A}\hat{\mathbb{I}} \quad (3.71)$$

$$= \left(\sum_i |i\rangle \langle i| \right) \hat{A} \left(\sum_j |j\rangle \langle j| \right) \quad (3.72)$$

$$= \sum_{i,j} |i\rangle \langle i| \hat{A} |j\rangle \langle j| \quad (3.73)$$

$$= \sum_{i,j} |i\rangle A_{ij} \langle j| \quad (3.74)$$

3.8 Matrix representation of the sum and product of two operators

Let

$$\hat{C} = \hat{A} + \hat{B} \quad (3.75)$$

Then

$$C_{ij} = \langle i | \hat{C} | j \rangle \quad (3.76)$$

$$= \langle i | \hat{A} + \hat{B} | j \rangle \quad (3.77)$$

$$= \langle i | \hat{A} | j \rangle + \langle i | \hat{B} | j \rangle \quad (3.78)$$

$$= A_{ij} + B_{ij} \quad (3.79)$$

Next, let \therefore

$$C_{ij} = \langle i | \hat{C} | j \rangle \quad (3.80)$$

$$= \langle i | \hat{A} \hat{B} | j \rangle \quad (3.81)$$

$$= \langle i | \hat{A} \hat{B} | j \rangle \quad (3.82)$$

$$= \sum_k \langle i | \hat{A} | k \rangle \langle k | \hat{B} | j \rangle \quad (3.83)$$

$$= \sum_k A_{ik} B_{kj} \quad (3.84)$$

In full matrix form

$$[C] = [A][B] \quad (3.85)$$

where

$$[A] = \begin{pmatrix} \langle 1 | \hat{A} | 1 \rangle & \langle 1 | \hat{A} | 2 \rangle & \dots \\ \langle 2 | \hat{A} | 1 \rangle & \langle 2 | \hat{A} | 2 \rangle & \dots \\ \vdots & \vdots & \vdots \end{pmatrix} \quad (3.86)$$

and similar expression applies for $[B]$ and $[C]$. This result shows that the matrix of an operator product is equal to the product of the matrices representing the operators, taken in the same order.

Example 1 Using a basis set $\{|i\rangle\}$ write down $\langle b | \hat{A} | a \rangle$ as a matrix product.

Ans

$$\langle b | \hat{A} | a \rangle = \sum_{i,j} \langle b | i \rangle \langle i | \hat{A} | j \rangle \langle j | a \rangle \quad (3.87)$$

$$= \sum_{i,j} b_i^* A_{ij} a_j \quad (3.88)$$

$$= [b]^\dagger [A][a] \quad (3.89)$$

where $[b]^\dagger$ is the matrix representation of $\langle b |$.

$$[b]^\dagger = \begin{pmatrix} b_1^* & b_2^* & \dots \end{pmatrix} \quad (3.90)$$

$[A]$ is the matrix representation of the operator \hat{A}

$$[A] = \begin{bmatrix} A_{11} & A_{12} & \dots \\ A_{21} & A_{22} & \dots \\ \vdots & \vdots & \vdots \end{bmatrix} \quad (3.91)$$

and $[a]$ is the matrix representation of the ket $|a\rangle$.

$$[a] = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \end{pmatrix} \quad (3.92)$$

writing in full

$$\langle b | \hat{A} | a \rangle = \begin{pmatrix} b_1^* & b_2^* & \dots \end{pmatrix} \begin{pmatrix} A_{11} & A_{12} & \dots \\ A_{21} & A_{22} & \dots \\ \vdots & \vdots & \vdots \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ \vdots \end{pmatrix} \quad (3.93)$$

Lecture 3

Linear Vector Space 2

3.9 Operators in a Hilbert space

An operator \hat{A} is a prescription by which every vector ψ_a in a Hilbert space H is associated with another vector ψ_b in the space:

$$\hat{A} : \psi_a \rightarrow \psi_b \quad (3.94)$$

for $\psi_a, \psi_b \in H$. We usually employ the notation

$$\psi_b = \hat{A} \psi_a. \quad (3.95)$$

In Dirac notation, we write

$$|b\rangle = \hat{A} |a\rangle \quad (3.96)$$

where both $|a\rangle$ and $|b\rangle$ belong in the ket space i.e., the Hilbert space H .

An operator can also act on a bra vector (bra-space is also a Hilbert space; it is dual to the ket space) changing it to another bra-vector. The notation we employ is

$$\langle \psi | = \langle \phi | \hat{A}. \quad (3.97)$$

Here the operator \hat{A} acts on the bra-vector $\langle \phi |$ to produce the bra vector $\langle \psi |$. We place the bra-vector on which the operator acts on the left of the operator.

3.9.1 Linear operators

An operator \hat{A} is said to be a linear operator if it has the following property: For any vectors $|a\rangle$ and $|b\rangle$ and any complex number λ_1 and λ_2 , we have

$$\hat{A}(\lambda_1 |a\rangle + \lambda_2 |b\rangle) = \lambda_1 \hat{A}|a\rangle + \lambda_2 \hat{A}|b\rangle. \quad (3.98)$$

A linear operator can act on a bra vector also.

$$(\lambda_1 \langle a| + \lambda_2 \langle b|) \hat{A} = \lambda_1 \langle a| \hat{A} + \lambda_2 \langle b| \hat{A}. \quad (3.99)$$

The operator \hat{A} is antilinear if

$$\hat{A}(\lambda_1 |a\rangle + \lambda_2 |b\rangle) = \lambda_1^* \hat{A}|a\rangle + \lambda_2^* \hat{A}|b\rangle. \quad (3.100)$$

Two operators \hat{A} and \hat{B} are equal if

$$\hat{A}|\psi\rangle = \hat{B}|\psi\rangle \quad (3.101)$$

for all $|\psi\rangle$ in the vector space.

Sum of two operators \hat{A} and \hat{B} is defined as

$$(\hat{A} + \hat{B})|\psi\rangle = \hat{A}|\psi\rangle + \hat{B}|\psi\rangle. \quad (3.102)$$

Product of two operators \hat{A} and \hat{B} is defined as

$$(\hat{A}\hat{B})|\psi\rangle = \hat{A}(\hat{B}|\psi\rangle). \quad (3.103)$$

This equation says that the operator $\hat{A}\hat{B}$ acting on $|\psi\rangle$ produces the same vector which would be obtained if we first let \hat{B} act on $|\psi\rangle$ and then \hat{A} acts on the result of the previous operation. In general $\hat{A}\hat{B} \neq \hat{B}\hat{A}$, although in exceptional cases we may have $\hat{A}\hat{B} = \hat{B}\hat{A}$.

3.9.2 Commutator of two operators

The commutator of two operators \hat{A} and \hat{B} is defined as

$$[\hat{A}, \hat{B}] \stackrel{\text{def}}{=} \hat{A}\hat{B} - \hat{B}\hat{A}. \quad (3.104)$$

In general $[\hat{A}, \hat{B}] \neq 0$ (null operator). If $[\hat{A}, \hat{B}] = 0$, we say that \hat{A} and \hat{B} commute with each other.

3.9.3 Some properties of commutators

We now list some properties of commutators. These properties follow immediately from the definition of the commutator.

$$[\hat{A}, \hat{B}] = -[\hat{B}, \hat{A}]. \quad (3.105)$$

$$[\hat{A}, \lambda\hat{B}] = \lambda[\hat{A}, \hat{B}]. \quad (3.106)$$

$$[\lambda\hat{A}, \hat{B}] = \lambda[\hat{A}, \hat{B}]. \quad (3.107)$$

$$[\hat{A}, \hat{B}\hat{C}] = \hat{B}[\hat{A}, \hat{C}] + [\hat{A}, \hat{B}]\hat{C}. \quad (3.108)$$

$$[\hat{A}\hat{B}, \hat{C}] = \hat{A}[\hat{B}, \hat{C}] + [\hat{A}, \hat{C}]\hat{B} \quad (3.109)$$

3.10 Projection operator

(an important example of a linear operator)

Consider the operator \hat{P}_a defined as

$$\hat{P}_a = |a\rangle\langle a| \quad (3.110)$$

where $\langle a|a\rangle = 1$. Operating by \hat{P}_a on an arbitrary ket $|\psi\rangle$, we have

$$\hat{P}_a |\psi\rangle = |a\rangle\langle a|\psi\rangle \quad (3.111)$$

i.e., \hat{P}_a projects the ket $|\psi\rangle$ along $|a\rangle$. The complex number $\langle a|\psi\rangle$ is called the component of $|\psi\rangle$ along $|a\rangle$. Now, \hat{P}_a is a linear operator. To show this consider

$$\begin{aligned} \hat{P}_a(\lambda_1|\psi_1\rangle + \lambda_2|\psi_2\rangle) &= |a\rangle\langle a|(\lambda_1|\psi_1\rangle + \lambda_2|\psi_2\rangle) \\ &= \lambda_1|a\rangle\langle a|\psi_1\rangle + \lambda_2|a\rangle\langle a|\psi_2\rangle \\ &= \lambda_1\hat{P}_a|\psi_1\rangle + \lambda_2\hat{P}_a|\psi_2\rangle. \end{aligned}$$

Another important property of the projection operator is

$$\hat{P}_a^2 = \hat{P}_a \quad (3.112)$$

To prove this allow \hat{P}_a^2 to act on a ket.

$$\begin{aligned}\hat{P}_a^2 |\psi\rangle &= \hat{P}_a \hat{P}_a |\psi\rangle \\ &= \hat{P}_a |a\rangle \langle a| |\psi\rangle \\ &= |a\rangle \langle a|a\rangle \langle a| |\psi\rangle \\ &= |a\rangle \langle a| |\psi\rangle \\ &= \hat{P}_a |\psi\rangle.\end{aligned}$$

Ex. Six operators are defined as follows:

$$\begin{aligned}\hat{A}_1 \psi(x) &= [\psi(x)]^2 & \hat{A}_4 \psi(x) &= x^2 \psi(x) \\ \hat{A}_2 \psi(x) &= \frac{d}{dx} \psi(x) & \hat{A}_5 \psi(x) &= \sin[\psi(x)] \\ \hat{A}_3 \psi(x) &= \int_a^x \psi(x') dx' & \hat{A}_6 \psi(x) &= \frac{d^2}{dx^2} \psi(x).\end{aligned}$$

Which of the operators \hat{A}_i are linear operators?

3.11 Representation of vectors and operators

3.11.1 Matrix representation of kets

Let $\{\phi_i\}$ be a complete orthonormal basis set in a Hilbert space. Since the basis is orthonormal, we must have

$$(\phi_i, \phi_j) = \delta_{ij}.$$

An arbitrary vector ψ_a can be written as a linear combination of the basis vectors. We write

$$\psi_a = \sum_i a_i \phi_i \quad (3.113)$$

where the scalars a_i are the components of the vector ψ_a along the basis vectors ϕ_i . Using the orthonormality of the basis vectors we immediately have

$$a_i = (\phi_i, \psi_a). \quad (3.114)$$

We can arrange these numbers as a column matrix:

$$\begin{pmatrix} a_1 \\ a_2 \\ \vdots \end{pmatrix} = \begin{pmatrix} (\phi_1, \psi_a) \\ (\phi_2, \psi_a) \\ \vdots \end{pmatrix}.$$

This column matrix is called the representation of the vector ψ_a with respect to the given basis $\{\phi_i\}$. In Dirac notation we represent the vector ψ_a as $|a\rangle$ and the basis vectors ϕ_i are written as $|i\rangle$. We can expand a general ket $|a\rangle$ as a linear combination of the basis kets:

$$|a\rangle = \sum_{i=1}^{\infty} a_i |i\rangle. \quad (3.115)$$

Orthonormality of the basis kets can be written as

$$\langle i|j\rangle = \delta_{ij} \quad (3.116)$$

The complex scalars a_i are called the components of the ket $|a\rangle$ along the basis kets $|i\rangle$. Using the orthonormality condition of the basis vectors (Eq. 3.116), we have

$$a_i = \langle i|a\rangle. \quad (3.117)$$

These scalars a_1, a_2, \dots , arranged as a column matrix is called the representation of $|a\rangle$ in the basis $\{|i\rangle, i = 1, 2, \dots\}$. Thus

$$|a\rangle \rightarrow \begin{pmatrix} a_1 \\ a_2 \\ \vdots \end{pmatrix} \equiv \begin{pmatrix} \langle 1|a\rangle \\ \langle 2|a\rangle \\ \vdots \end{pmatrix}. \quad (3.118)$$

We can write down the representation of any one of the basis vectors in the same basis as

$$|1\rangle \rightarrow \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \end{pmatrix} \quad |2\rangle \rightarrow \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \end{pmatrix} \quad |3\rangle \rightarrow \begin{pmatrix} 0 \\ 0 \\ 1 \\ \vdots \end{pmatrix}, \quad (3.119)$$

and so on.

Now, using Eq. (3.117) in Eq. (3.115) we have

$$\begin{aligned}
 |a\rangle &= \sum_i a_i |i\rangle \\
 &= \sum_i \langle i|a\rangle |i\rangle \\
 &= \sum_i |i\rangle \langle i|a\rangle \\
 &= \left(\sum_i \hat{P}_i \right) |a\rangle
 \end{aligned} \tag{3.120}$$

where

$$\hat{P}_i = |i\rangle \langle i| \tag{3.121}$$

is the projection operator along $|i\rangle$. Since Eq. (8.10) is true for all $|a\rangle$ in the vector space (this is because $\{|i\rangle\}$ for a complete set), we must have

$$\sum_i \hat{P}_i = \sum_i |i\rangle \langle i| = \hat{I} \tag{3.122}$$

where \hat{I} is the identity operator. Eq. (17.64) is called the completeness condition for the basis vectors.

3.11.2 Matrix representation of bra vectors

The ket vector $|a\rangle$ is represented by a column matrix

$$\begin{pmatrix} a_1 \\ a_2 \\ \cdot \\ \cdot \\ \cdot \end{pmatrix} = \begin{pmatrix} \langle 1|a\rangle \\ \langle 2|a\rangle \\ \cdot \\ \cdot \\ \cdot \end{pmatrix}$$

in the basis $\{|i\rangle\}$. The dual of ket $|a\rangle$ is the bra $\langle a|$. What is the matrix representation of the bra $\langle a|$ in the same basis? To see this we can expand $\langle a|$ as

$$\langle a| = \sum_i \langle a|i\rangle \langle i|. \tag{3.123}$$

The bra $\langle a|$ is represented by a row vector:

$$\begin{aligned}
 \langle a| &\rightarrow \left(\langle a|1\rangle \quad \langle a|2\rangle \quad \dots \right) \\
 &= \left(a_1^* \quad a_2^* \quad \dots \right).
 \end{aligned} \tag{3.124}$$

Then the scalar product $\langle a|a\rangle$ becomes a real number:

$$\begin{aligned}
 \langle a|a\rangle &= \begin{pmatrix} a_1^* & a_2^* & \dots \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ \vdots \end{pmatrix} \\
 &= a_1^*a_1 + a_2^*a_2 + \dots \\
 &= \sum_i a_i^*a_i \\
 &= \sum_i |a_i|^2 \quad (\text{real number}). \tag{3.125}
 \end{aligned}$$

In general the scalar product $\langle b|a\rangle$ is

$$\begin{aligned}
 \langle b|a\rangle &= \sum_i \langle b|i\rangle \langle i|a\rangle \\
 &= \begin{pmatrix} \langle b|1\rangle & \langle b|2\rangle & \dots \end{pmatrix} \begin{pmatrix} \langle 1|a\rangle \\ \langle 2|a\rangle \\ \vdots \end{pmatrix} \\
 &= \begin{pmatrix} b_1^* & b_2^* & \dots \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ \vdots \end{pmatrix} \\
 &= \sum_i b_i^*a_i \quad (\text{complex number}). \tag{3.126}
 \end{aligned}$$

3.11.3 Representation of operators

Consider the equation

$$|b\rangle = \hat{A}|a\rangle. \tag{3.127}$$

Let $\{|i\rangle, i = 1, 2, \dots\}$ be a complete set of orthonormal basis states. Taking the component of Eq. (3.127) along $|i\rangle$, we have

$$\begin{aligned}
 \langle i|b\rangle &= \langle i|\hat{A}|a\rangle \\
 &= \sum_j \langle i|\hat{A}|j\rangle \langle j|a\rangle. \tag{3.128}
 \end{aligned}$$

In matrix notation

$$b_i = \sum_j A_{ij}a_j \tag{3.129}$$

where

$$\begin{aligned} b_i &= \langle i | b \rangle \\ a_j &= \langle j | a \rangle \\ A_{ij} &= \langle i | \hat{A} | j \rangle \end{aligned}$$

Writing in full, Eq. (3.129) becomes

$$\begin{pmatrix} b_1 \\ b_2 \\ \cdot \\ \cdot \\ \cdot \end{pmatrix} = \begin{pmatrix} A_{11} & A_{12} & \cdot & \cdot & \cdot \\ A_{21} & A_{22} & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ \cdot \\ \cdot \\ \cdot \end{pmatrix} \quad (3.130)$$

The matrix $[A]$ with elements $A_{ij} = \langle i | \hat{A} | j \rangle$ is called the matrix representation of the operator \hat{A} with respect to the given basis $\{|i\rangle\}$. Using a basis set, the operator \hat{A} can also be written as

$$\begin{aligned} \hat{A} &= \hat{I} \hat{A} \hat{I} \\ &= \left(\sum_i |i\rangle \langle i| \right) \hat{A} \left(\sum_j |j\rangle \langle j| \right) \\ &= \sum_{i,j} |i\rangle \langle i | \hat{A} | j \rangle \langle j| \\ &= \sum_{i,j} |i\rangle A_{ij} \langle j| \end{aligned} \quad (3.131)$$

3.11.4 Matrix representation of the sum and product of two operators

Sum of two operators:

Let

$$\hat{C} = \hat{A} + \hat{B}$$

Then the matrix representation of \hat{C} is

$$\begin{aligned} C_{ij} &= \langle i | \hat{C} | j \rangle \\ &= \langle i | \hat{A} + \hat{B} | j \rangle \\ &= \langle i | \hat{A} | j \rangle + \langle i | \hat{B} | j \rangle \\ &= A_{ij} + B_{ij}. \end{aligned} \quad (3.132)$$

Product of two operators:

Next, suppose

$$\hat{C} = \hat{A} \hat{B}.$$

The matrix representation of \hat{C} is then

$$\begin{aligned}
 C_{ij} &= \langle i | \hat{C} | j \rangle \\
 &= \langle i | \hat{A} \hat{B} | j \rangle \\
 &= \langle i | \hat{A} \hat{I} \hat{B} | j \rangle \\
 &= \sum_k \langle i | \hat{A} | k \rangle \langle k | \hat{B} | j \rangle \\
 &= \sum_k A_{ik} B_{kj}.
 \end{aligned} \tag{3.133}$$

In full matrix form we can write

$$[C] = [A][B] \tag{3.134}$$

where

$$[A] = \begin{pmatrix} \langle 1 | \hat{A} | 1 \rangle & \langle 1 | \hat{A} | 2 \rangle & \dots & \dots & \dots \\ \langle 2 | \hat{A} | 1 \rangle & \langle 2 | \hat{A} | 2 \rangle & \dots & \dots & \dots \\ \vdots & \vdots & \ddots & \ddots & \ddots \\ \vdots & \vdots & \ddots & \ddots & \ddots \\ \vdots & \vdots & \ddots & \ddots & \ddots \end{pmatrix} \tag{3.135}$$

and similarly for $[B]$ and $[C]$. This result shows that the matrix of an operator product is equal to the product of the matrices representing the operators, taken in the same order.

Example

Using a basis set $\{|i\rangle\}$ write down $\langle b | \hat{A} | a \rangle$ as a matrix product.

Ans

$$\begin{aligned}
 \langle b | \hat{A} | a \rangle &= \sum_{i,j} \langle b | i \rangle \langle i | \hat{A} | j \rangle \langle j | a \rangle \\
 &= \sum_{i,j} b_i^* A_{ij} a_j \\
 &= [b]^\dagger [A] [a].
 \end{aligned} \tag{3.136}$$

where $[b]^\dagger$ is the matrix representation of $\langle b |$, i.e.,

$$[b]^\dagger = \begin{pmatrix} b_1^* & b_2^* & \dots \end{pmatrix}, \tag{3.137}$$

$[A]$ is the matrix representation of the operator \hat{A} :

$$[A] = \begin{pmatrix} A_{11} & A_{12} & \dots & \dots & \dots \\ A_{21} & A_{22} & \dots & \dots & \dots \\ \vdots & \vdots & \ddots & \ddots & \ddots \\ \vdots & \vdots & \ddots & \ddots & \ddots \\ \vdots & \vdots & \ddots & \ddots & \ddots \end{pmatrix} \tag{3.138}$$

and $[a]$ is the matrix representation of the ket $|a\rangle$

$$[a] = \begin{pmatrix} a_1 \\ a_2 \\ \cdot \\ \cdot \\ \cdot \end{pmatrix} \quad (3.139)$$

Writing in full, we have

$$\langle b|\hat{A}|a\rangle = \begin{pmatrix} b_1^* & b_2^* & \cdot & \cdot & \cdot \end{pmatrix} \begin{pmatrix} A_{11} & A_{12} & \cdot & \cdot & \cdot \\ A_{21} & A_{22} & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ \cdot \\ \cdot \\ \cdot \end{pmatrix}. \quad (3.140)$$

4. sheet-4 : Linear Vector Space (continued)

4.1 Adjoint operator

Consider the equation

$$|b\rangle = \hat{K}|a\rangle \quad (4.1)$$

The operator \hat{K} carries the ket $|a\rangle$ to the ket $|b\rangle$. The dual of $|a\rangle$ and $|b\rangle$ are the bras $\langle a|$ and $\langle b|$ respectively. Then the operator which carries $\langle a|$ to $\langle b|$ is called the adjoint of \hat{K} and is denoted by \hat{K}^\dagger . Thus in dual space 4.1 is

$$\langle b| = \langle a|\hat{K}^\dagger \quad (4.2)$$

$$\text{ket space } |a\rangle \xrightarrow{\hat{K}} |b\rangle \quad (4.3)$$

$$\text{bra space } \langle a| \xrightarrow{\hat{K}^\dagger} \langle b| \quad (4.4)$$

Here bra space is the dual of ket space and

$$|a\rangle_{\rightarrow}^{dc} \langle a| \quad (4.5)$$

$$|b\rangle_{\rightarrow}^{dc} \langle b| \quad (4.6)$$

dc \equiv dual correspondence.

From equation 4.1 and 4.2 it follows that

$$\langle c|b\rangle = \langle c|\hat{K}|a\rangle \quad (4.7)$$

and

$$\langle b|c\rangle = \langle a|\hat{K}^\dagger|c\rangle \quad (4.8)$$

since

$$\langle b|c\rangle = \langle c|b\rangle^* \quad (4.9)$$

we have

$$\langle c|\hat{K}|a\rangle = \langle a|\hat{K}^\dagger|c\rangle \quad (4.10)$$

equation 4.10 is the defining equation for the adjoint \hat{K}^\dagger of the operator \hat{K} . In scalar product notation

$$\langle c|b\rangle \equiv (\psi_c, \psi_b) \quad (4.11)$$

equation 4.10 can be written as

$$(\psi_c, \hat{K}\psi_a) = (\psi_a, \hat{K}^\dagger\psi_c)^* \quad (4.12)$$

$$= (\hat{K}^\dagger\psi_c, \psi_a) \quad (4.13)$$

In particular, if we take $|c\rangle$ and $|a\rangle$ as the basis states $|i\rangle$ and $|j\rangle$, equation 4.10 becomes

$$\langle i|\hat{K}|j\rangle = \langle j|\hat{K}^\dagger|i\rangle^* \quad (4.14)$$

$$K_{ij} = (\hat{K}^\dagger)_{ji}^* \quad (4.15)$$

$$K_{ji}^* = K_{ij}^* \quad (4.16)$$

$$K_{ij}^* = (K_{ji})^* \quad (4.17)$$

$$[\hat{K}^\dagger] = [K]^\dagger \quad (4.18)$$

i.e., The matrix representation of the adjoint operator is the hermitian conjugate of the matrix representation of \hat{K} .

4.2 Hermitian or Self-adjoint operator

If $\hat{K}^\dagger = \hat{K}$, then \hat{K} is said to be a self-adjoint or a hermitian operator. For a hermitian operator

$$[K] = [K^\dagger] = [K]^\dagger \quad (4.19)$$

or

$$K_{ij} = K_{ji}^* \quad (4.20)$$

i.e., $[K]$ is a hermitian matrix.

Example 1 show that

$$(AB)^\dagger = B^\dagger A^\dagger \quad (4.21)$$

Answer:

$$(\psi_a, \hat{A}\hat{B}\psi_b) = \left((\hat{A}\hat{B})^\dagger \psi_a, \psi_b \right) \quad (4.22)$$

Also

$$(\psi_a, \hat{A}\hat{B}\psi_b) = (\hat{A}^\dagger\psi_a, \hat{B}\psi_b) = (\hat{B}^\dagger\hat{A}^\dagger\psi_a, \psi_b) \quad (4.23)$$

Comparing these equation (equation 4.22 and 4.23) we get,

$$(\hat{A}\hat{B})^\dagger = \hat{B}^\dagger\hat{A}^\dagger \quad (4.24)$$

4.3 Inverse operator

An operator \hat{B} is said to be the inverse of \hat{A} if

$$\hat{A}\hat{B} = \hat{B}\hat{A} = \hat{\mathbb{I}} \quad (4.25)$$

Obviously, if \hat{B} is the inverse of \hat{A} , then \hat{A} is the inverse of \hat{B} . We write

$$\hat{B} \equiv \hat{A}^{-1} \quad (4.26)$$

or

$$\hat{A} = \hat{B}^{-1} \quad (4.27)$$

If equation 4.25 is satisfied,

4.4 Unitary operator

An operator \hat{U} is said to be unitary if

$$\hat{U}\hat{U}^\dagger \equiv \hat{U}^\dagger\hat{U} \equiv \hat{\mathbb{I}} \quad (4.28)$$

i.e., if

$$\hat{U}^\dagger \equiv \hat{U}^{-1} \quad (4.29)$$

Thus, for a unitary operator, its adjoint is @@@@ @@@@ @@@@ @@@@ @@@@ @@@@ @@@@ @@@@

4.5 Function of Operators

Lecture 4 Linear Vector Space 3

4.6 Adjoint Operator

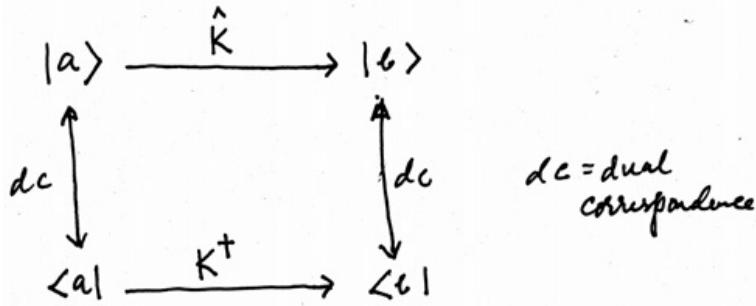
Consider the equation

$$|b\rangle = \hat{K}|a\rangle. \quad (4.30)$$

The operator \hat{K} carries the ket $|a\rangle$ to the ket $|b\rangle$. The dual of $|a\rangle$ and $|b\rangle$ are $\langle a|$ and $\langle b|$, respectively. The operator which carries $\langle a|$ to $\langle b|$ is called the adjoint of \hat{K} and is denoted by \hat{K}^\dagger . Thus, in the dual space, Eq. (4.30) is

$$\langle b| = \langle a|\hat{K}^\dagger. \quad (4.31)$$

We summarize the situation in the figure below.



From Eqs. (4.30) and (4.31) it follows that

$$\langle c|b\rangle = \langle c|\hat{K}|a\rangle \quad (4.32)$$

and

$$\langle b|c\rangle = \langle a|\hat{K}^\dagger|c\rangle. \quad (4.33)$$

Since $\langle b|c\rangle = \langle c|b\rangle^*$, we have

$$\langle c|\hat{K}|a\rangle = \langle a|\hat{K}^\dagger|c\rangle^*. \quad (4.34)$$

Eq. (4.34) is the defining equation for the adjoint \hat{K}^\dagger of the operator \hat{K} . In the scalar product notation (comma notation)

$$\langle c|b\rangle \equiv (\psi_c, \psi_b)$$

Eq. (4.34) can be written as

$$\begin{aligned} (\psi_c, \hat{K}\psi_a) &= (\psi_a, \hat{K}^\dagger\psi_c)^* \\ &= (\hat{K}^\dagger\psi_c, \psi_a). \end{aligned} \quad (4.35)$$

In particular, if we take $|c\rangle$ and $|a\rangle$ as the basis states $|i\rangle$ and $|j\rangle$, Eq. (4.34) becomes

$$\langle i|\hat{K}|j\rangle = \langle j|\hat{K}^\dagger|i\rangle^*$$

or,

$$K_{ij} = (K^\dagger)_{ji}^*$$

or,

$$K_{ji}^\dagger = K_{ij}^*$$

i.e.,

$$K_{ij}^\dagger = (K_{ji})^*.$$

In full matrix notation we can write

$$[\hat{K}^\dagger] = [\hat{K}]^\dagger, \quad (4.36)$$

where $[\hat{K}]$ is the matrix representation of the operator \hat{K} . Similarly for \hat{K}^\dagger . Thus, the matrix representation of the adjoint operator \hat{K}^\dagger is the Hermitian conjugate of the matrix representation of \hat{K} .

4.6.1 Hermitian or self adjoint operator

If $\hat{K}^\dagger = \hat{K}$, then \hat{K} is said to be a self-adjoint or a Hermitian operator. The matrix representing a Hermitian operaor is a Hermitian matrix i.e.,

$$[K] = [K]^\dagger$$

or,

$$K_{ij} = K_{ji}^*.$$

Ex:Show that $(\hat{A}\hat{B})^\dagger = \hat{B}^\dagger\hat{A}^\dagger$.**Ans:**

$$(\psi_a, \hat{A}\hat{B}\psi_b) = ((\hat{A}\hat{B})^\dagger \psi_a, \psi_b). \quad (4.37)$$

Also,

$$(\psi_a, \hat{A}\hat{B}\psi_b) = (\hat{A}^\dagger \psi_a, \hat{B}\psi_b) = (\hat{B}^\dagger \hat{A}^\dagger, \psi_b) \quad (4.38)$$

Comparing Eqs. (4.37) and (4.38), we have

$$(\hat{A}\hat{B})^\dagger = \hat{B}^\dagger\hat{A}^\dagger. \quad (4.39)$$

4.6.2 Inverse operator

An operator \hat{B} is said to be the inverse of the operator \hat{A} if

$$\hat{A}\hat{B} = \hat{B}\hat{A} = \hat{I}. \quad (4.40)$$

Obviously if \hat{B} is the inverse of \hat{A} , then \hat{A} is the inverse of \hat{B} . We write

$$\hat{B} = \hat{A}^{-1}$$

or,

$$\hat{A} = \hat{B}^{-1}$$

if Eq. (4.40) is satisfied.

4.6.3 Unitary operator

An operator \hat{U} is said to be unitary if

$$\hat{U}\hat{U}^\dagger = \hat{U}^\dagger\hat{U} = \hat{I} \quad (4.41)$$

i.e., if

$$\hat{U}^\dagger = \hat{U}^{-1}. \quad (4.42)$$

Thus, for a unitary operator, its adjoint is also its inverse.

4.7 Functions of operators

Consider a real-valued function $f(x)$ of a real variable x . Suppose that the function has a power series expansion

$$f(x) = f_0 + xf_1 + x^2f_2 + \dots . \quad (4.43)$$

If \hat{A} is an operator, we can define the operator $\hat{f}(\hat{A})$ as

$$\hat{f}(\hat{A}) = f_0\hat{I} + \hat{A}f_1 + \hat{A}^2f_2 + \dots . \quad (4.44)$$

As an example of a function of an operator, consider the operator $e^{\lambda\hat{A}}$. This is defined as

$$e^{\lambda\hat{A}} = \hat{I} + \lambda\hat{A} + \frac{\lambda^2}{2!}\hat{A}^2 + \frac{\lambda^3}{3!}\hat{A}^3 + \dots . \quad (4.45)$$

One must be careful in manipulating functions of operators since operators do not commute with each other in general. For example, if $\hat{A}\hat{B} \neq \hat{B}\hat{A}$, then

$$e^{\hat{A}+\hat{B}} \neq e^{\hat{A}}e^{\hat{B}} \neq e^{\hat{B}}e^{\hat{A}}. \quad (4.46)$$

In the special case when $[\hat{A}, \hat{B}]$ is a number times the unit operator, i.e.,

$$[\hat{A}, \hat{B}] = c\hat{I}$$

where c is a number (in general complex), then

$$e^{\hat{A}+\hat{B}} = e^{\hat{A}}e^{\hat{B}}e^{-[\hat{A}, \hat{B}]/2}. \quad (4.47)$$

This result is known as Weyl's identity. For example, in quantum mechanics we have

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

So Weyl's identity would be satisfied by \hat{x} and \hat{p}_x .

Examples

Ex Show that

$$e^{\lambda\hat{A}}\hat{B}e^{-\lambda\hat{A}} = \hat{B} + \lambda[\hat{A}, \hat{B}] + \frac{\lambda^2}{2!}[\hat{A}, [\hat{A}, \hat{B}]] + \frac{\lambda^3}{3!}[\hat{A}, [\hat{A}, [\hat{A}, \hat{B}]]] + \dots . \quad (4.48)$$

Ans:

Let

$$f(\lambda) = e^{\lambda\hat{A}}\hat{B}e^{-\lambda\hat{A}}.$$

Therefore

$$\frac{df(\lambda)}{d\lambda} = e^{\lambda \hat{A}} \hat{A} \hat{B} e^{-\lambda \hat{A}} - e^{\lambda \hat{A}} \hat{B} \hat{A} e^{-\lambda \hat{A}} = e^{\lambda \hat{A}} [\hat{A}, \hat{B}] e^{-\lambda \hat{A}}.$$

Differentiating one more time, we have

$$\frac{d^2 f(\lambda)}{d\lambda^2} = e^{\lambda \hat{A}} [\hat{A}, [\hat{A}, \hat{B}]] e^{-\lambda \hat{A}},$$

and so on. Also

$$f(0) = \hat{B}, \quad \left(\frac{df}{d\lambda} \right)_{\lambda=0} = [\hat{A}, \hat{B}], \quad \left(\frac{d^2 f}{d\lambda^2} \right)_{\lambda=0} = [\hat{A}, [\hat{A}, \hat{B}]], \dots$$

Expanding $f(\lambda)$ in a Taylor series:

$$f(\lambda) = f(0) + \lambda \left(\frac{df}{d\lambda} \right)_{\lambda=0} + \frac{\lambda^2}{2!} \left(\frac{d^2 f}{d\lambda^2} \right)_{\lambda=0} + \dots$$

we have

$$e^{\lambda \hat{A}} \hat{B} e^{-\lambda \hat{A}} = \hat{B} + \lambda [\hat{A}, \hat{B}] + \frac{\lambda^2}{2!} [\hat{A}, [\hat{A}, \hat{B}]] + \dots$$

4.8 Change of basis

Supose we have a set of complete orthonormal basis set $\{|u_i\rangle\}$ in a Hilbert space. The orthonormality and completeness of the vasis set can be expressed as

$$\langle u_i | u_j \rangle = \delta_{ij} \text{ orthonormality} \quad (4.49)$$

and

$$\sum_i |u_i\rangle \langle u_i| = \hat{I} \text{ completeness.} \quad (4.50)$$

In terms of the basis set $\{|u_i\rangle\}$, an arbitrary ket $|\psi\rangle$ of the Hilbert space can be expanded as

$$|\psi\rangle = \sum_i |u_i\rangle \langle u_i| \psi = \sum_i a_i |u_i\rangle, \quad (4.51)$$

where

$$a_i = \langle u_1 | \psi \rangle \quad (4.52)$$

is the component of $|\psi\rangle$ along $|u_i\rangle$. The numbers a_i arranged as a column matrix is called the representation of the ket $|\psi\rangle$ in the basis $\{|u_i\rangle\}$. Thus,

$$|\psi\rangle \longrightarrow \begin{bmatrix} a_1 \\ a_2 \\ \vdots \end{bmatrix} = \begin{bmatrix} \langle u_1 | \psi \rangle \\ \langle u_2 | \psi \rangle \\ \vdots \end{bmatrix}. \quad (4.53)$$

The conjugate of $|\psi\rangle$ in the dual space is written as the bra $\langle\psi|$. The matrix representation of $\langle\psi|$ is the row vector with components $\langle\psi|u_i\rangle$, i.e. $\langle u_i|\psi\rangle^*$. Thus,

$$\langle\psi| \longrightarrow \begin{bmatrix} \langle\psi|u_1\rangle & \langle\psi|u_2\rangle & \cdots \end{bmatrix} = \begin{bmatrix} a_1^* & a_2^* & \cdots \end{bmatrix}. \quad (4.54)$$

Using the basis $\{|u_i\rangle\}$, we can also write down the representation of an operator \hat{A} as a square matrix with elements A_{ij} given by

$$A_{ij} = \langle u_i|\hat{A}|u_j\rangle. \quad (4.55)$$

Writing in full

$$\hat{A} \rightarrow [A] = \begin{bmatrix} A_{11} & A_{12} & A_{13} & \cdots \\ A_{21} & A_{22} & A_{23} & \cdots \\ A_{31} & A_{32} & A_{33} & \cdots \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix} \quad (4.56)$$

Now, we make a change from the basis vectors $\{|u_i\rangle, i = 1, 2, 3, \dots\}$ to a new set of orthonormal basis vectors $\{|u'_i\rangle, i = 1, 2, 3, \dots\}$. The orthonormality and completeness of the new basis states can be expressed as

$$\langle u'_i|u'_j\rangle = \delta_{ij} \quad (4.57)$$

and

$$\sum_i |u'_i\rangle\langle u'_i| = \hat{I}. \quad (4.58)$$

We can also write down the matrix representations of kets and operators in the new basis. We want to find how the components of the ket $|\psi\rangle$ in the new basis relates to the components in the old basis. Similarly, we also want to know how the matrix elements of an operator transform as we make the change of basis.

4.8.1 Change of representation for kets

Let $|\psi\rangle$ be an arbitrary ket in the vector space V . In the new basis $\{|u'_i\rangle\}$, the components a'_i of the ket $|\psi\rangle$ are

$$\begin{aligned} a'_i &= \langle u'_i|\psi\rangle \\ &= \sum_j \langle u'_i|u_j\rangle\langle u_j|\psi\rangle, \end{aligned} \quad (4.59)$$

or,

$$a'_i = \sum_j S_{ij} a_j \quad (4.60)$$

where we have defined

$$S_{ij} = \langle u'_i | u_j \rangle. \quad (4.61)$$

In matrix form Eq. (4.60) is

$$\begin{bmatrix} a'_1 \\ a'_2 \\ \vdots \end{bmatrix} = \begin{bmatrix} \langle u'_1 | u_1 \rangle & \langle u'_1 | u_2 \rangle & \dots \\ \langle u'_2 | u_1 \rangle & \langle u'_2 | u_2 \rangle & \dots \\ \vdots & \vdots & \vdots \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \end{bmatrix}. \quad (4.62)$$

Before proceeding, we shall show that the matrix $[S] = (S_{ij})$ is unitary.

To show that $[S]$ is unitary

We start with the orthonormality of the new basis set, i.e.,

$$\langle u'_i | u'_j \rangle = \delta_{ij}$$

or,

$$\sum_k \langle u'_i | u_k \rangle \langle u_k | u'_j \rangle = \delta_{ij},$$

or,

$$\sum_k S_{ik} S_{jk}^* = \delta_{ij},$$

or,

$$\sum_k S_{ik} S_{kj}^\dagger = \delta_{ij}$$

i.e.,

$$[S][S]^\dagger = [I].$$

Next, we use the orthonormality of the old basis set.

$$\langle u_i | u_j \rangle = \delta_{ij}$$

or,

$$\sum_k \langle u_i | u'_k \rangle \langle u'_k | u_j \rangle = \delta_{ij},$$

4.9 Transformation of the matrix elements of an operator due to a change of basis

or,

$$\sum_k S_{ki}^* S_{kj} = \delta_{ij},$$

or,

$$\sum_k S_{ik}^\dagger S_{kj} = \delta_{ij}$$

i.e.,

$$[S]^\dagger [S] = [I].$$

Thus, we have proved

$$[S][S]^\dagger = [S]^\dagger [S] = [I], \quad (4.63)$$

i.e., $[S]$ is a unitary matrix.

4.9 Transformation of the matrix elements of an operator due to a change of basis

Next, we will discuss how the matrix elements of an operator transform if we make a change of basis. To do so, we proceed as follows.

$$\begin{aligned} A'_{ij} &= \langle u'_i | \hat{A} | u'_j \rangle \\ &= \sum_{k,l} \langle u'_i | u_k \rangle \langle u_k | \hat{A} | u_l \rangle \langle u_l | u'_j \rangle \\ &= \sum_{k,l} \langle u'_i | u_k \rangle A_{kl} \langle u'_j | u_l \rangle^* \\ &= \sum_{k,l} S_{ik} A_{kl} S_{jl}^*. \end{aligned} \quad (4.64)$$

In the full matrix notation we can write

$$[A]' = [S][A][S]^\dagger, \quad (4.65)$$

or, since $[S]$ is unitary,

$$[A]' = [S][A][S]^{-1}. \quad (4.66)$$

Such a transformation of a square matrix is called a similarity transformation.

5. sheet-5 : Eigenvalue and Eigenvectors of Operators

The ket $|\alpha\rangle$ is called the eigenvector or eigenket of the operator A if

$$A|\alpha\rangle = \alpha|\alpha\rangle \quad (5.1)$$

The number α is called the eigenvalue. Thus the effect of \hat{A} on an eigenket of A is merely multiplication by a number.

5.1 Eigenvalues and Eigenvectors of a Hermitian Operator

We now take up the eigenvalue problem of a hermitian operators. Two theorems are of vital important in this content.

Theorem 5.1.1 The eigenvalues of a hermitian operator are real.

Theorem 5.1.2 The eigenvectors of a hermitian operator belonging to different eigenvalues are orthogonal.

Proof. Let A be a hermitian operator and

$$A|\alpha_1\rangle = \alpha_1|\alpha_1\rangle \quad (5.2)$$

$$A|\alpha_2\rangle = \alpha_2|\alpha_2\rangle \quad (5.3)$$

From equation (5.2) we have,

$$\langle\alpha_2|A|\alpha_1\rangle = \alpha_1\langle\alpha_2|\alpha_1\rangle \quad (5.4)$$

Next we take the adjoint of equation (5.3)

$$\langle \alpha_2 | A^\dagger = \alpha_2^* \langle \alpha_2 | \quad (5.5)$$

Since A is hermitian, i.e., $A^\dagger = A$, we get

$$\langle \alpha_2 | A = \alpha_2^* \langle \alpha_2 | \quad (5.6)$$

Hence

$$\langle \alpha_2 | A | \alpha_1 \rangle = \alpha_2^* \langle \alpha_2 | \alpha_1 \rangle \quad (5.7)$$

Combining equation (5.4) and (5.7) we get

$$(\alpha_1 - \alpha_2^*) \langle \alpha_2 | \alpha_1 \rangle = 0 \quad (5.8)$$

If we let $\alpha_2 = \alpha_1$ and recalling that $\langle \alpha_1 | \alpha_1 \rangle \neq 0$, it follows that

$$\alpha_1 - \alpha_1^* = 0 \quad (5.9)$$

i.e., α_1 is real. Since eigenvalues are proved to be real, we can write equation (5.8) as

$$(\alpha_1 - \alpha_2) \langle \alpha_2 | \alpha_1 \rangle = 0 \quad (5.10)$$

If $\alpha_1 \neq \alpha_2$, we must have

$$\langle \alpha_2 | \alpha_1 \rangle = 0 \quad (5.11)$$

i.e., eigenvectors belonging to different eigenvalues are orthogonal. Owing to the linearity of the operators \hat{A} we can normalize the eigenvectors. We shall therefore usually assume that

$$\langle \alpha_1 | \alpha_2 \rangle = \delta_{\alpha_1 \alpha_2} \quad (5.12)$$

Thus, the eigenvectors of a hermitian operator form an orthonormal (and hence linearly independent vectors), i.e.,

$$\langle \alpha_i | \alpha_j \rangle = \delta_{\alpha_i \alpha_j} \quad (5.13)$$

■

5.2 Determination of eigenvalues and eigenvectors of a Hermitian Operator

Let A be a hermitian operator. Consider the eigenvalue equation

$$A |\lambda\rangle = \lambda |\lambda\rangle \quad (5.14)$$

To find the eigenvalue and the corresponding eigenvectors, we have to choose a basis in the vector space and convert the operator equation (5.14) into a matrix equation. For simplicity, we will assume that the vector space is finite dimensional with dimension n .

Now choosing an orthonormal basis set $\{|u_i\rangle\}$, we can cast equation (5.14) as a matrix equation of the following form:

$$\begin{bmatrix} A_{11} & A_{12} & \dots & A_{1n} \\ A_{21} & A_{22} & \dots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & \dots & A_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \lambda \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \quad (5.15)$$

Here x_1, x_2, \dots, x_n are the components of the eigenvector $|\lambda\rangle$ in **directions** $|u_1\rangle, |u_2\rangle, \dots, |u_n\rangle$ respectively, i.e.,

$$x_i \equiv \langle u_i | \lambda \rangle ; i = 1, 2, \dots, n \quad (5.16)$$

Equation (5.15) is a set of linear homogeneous equations which possess non-trivial solutions only if

$$\begin{vmatrix} (A_{11} - \lambda) & A_{12} & \dots & A_{1n} \\ A_{21} & (A_{22} - \lambda) & \dots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & \dots & (A_{nn} - \lambda) \end{vmatrix} = 0 \quad (5.17)$$

or in short

$$\det(A_{ij} - \lambda \delta_{ij}) = 0 \quad (5.18)$$

In matrix notation, we can write

$$|\mathbb{A} - \lambda \mathbb{1}| = 0 \quad (5.19)$$

This equation, which is a polynomial equation of degree n in the unknown λ , is called the secular equation of the matrix \mathbb{A} . Solving this equation we get n roots which we label as

$$\lambda_1, \lambda_2, \dots, \lambda_n$$

Now, we can distinguish two cases. If the n eigenvalues are all distinct, we say that the eigenvalues are *non-degenerate*. However, it may so happen that some of the eigenvalues are repeated. Those eigenvalues which are repeated are called *degenerate* eigenvalues and the number of times an eigenvalue is repeated is called the *order of degeneracy* of that eigenvalue.

5.3 Non-degenerate roots

In this case all the roots λ_i are distinct and there are n of them if the vector space is n dimensional. If A is hermitian, the roots are real. For a non-hermitian operator some or all of the roots may be complex.

Now, for each eigenvalue (root of secular equation) we can solve the eigenvalue equation (5.15) to get n linearly independent eigenvectors $|\lambda_i\rangle$. Since the $|\lambda_i\rangle$'s are linearly independent, they span the n dimensional vector space, i.e., they form a complete set of basis vectors.

If A is hermitian, the eigenvectors are guaranteed to be orthogonal, i.e., $\langle \lambda_i | \lambda_j \rangle = 0$ if $i \neq j$. However, for a non-hermitian operator the eigenvectors may or may not be orthogonal. Using the eigenvectors of A as the basis (This basis is called the eigenbasis of A), the matrix representation of A is

$$A'_{ij} \equiv \langle \lambda_i | A | \lambda_j \rangle = \lambda_j \langle \lambda_i | \lambda_j \rangle \quad (5.20)$$

For hermitian A , we always have $\langle \lambda_i | \lambda_j \rangle = 0$ if $i \neq j$, and, further we can normalize each eigenvector $|\lambda_i\rangle$. Thus, for a hermitian operator, the eigenbasis is an orthogonal set, i.e.,

$$\langle \lambda_i | \lambda_j \rangle = \delta_{ij} \quad (5.21)$$

Therefore, the matrix representation of the operator A in its eigenbasis is diagonal, i.e.,

$$A'_{ij} = \lambda_j \delta_{ij} \quad (5.22)$$

Writing out the matrix (A'_{ij}) in full

$$\mathbb{A}' = \begin{bmatrix} \lambda_1 & 0 & 0 & \dots & 0 \\ 0 & \lambda_2 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \lambda_n \end{bmatrix} \quad (5.23)$$

An operator or a matrix \mathbb{A} is said to be diagonalizable, if we can find a basis in which the matrix becomes diagonal. For a hermitian operator we can always find a basis, the eigenbasis of the operator, in which the matrix representation of the operator is diagonal with the eigenvalues as the

diagonal elements.

For a non-hermitian operator in an n dimensional vector space, there is no guarantee that the matrix representation A'_{ij} in the eigenbasis of the operator is diagonal. This is because, in general, the eigenvectors are not orthogonal, i.e., $\langle \lambda_i | \lambda_j \rangle \neq \delta_{ij}$

5.4 Degenerate roots

The secular equation (5.20) may have roots some or all of which are repeated. So, the number of distinct eigenvalues is now less than the dimension of the vector space.

As an example, suppose we have a six-dimensional vector space ($n = 6$) with three distinct roots $\lambda_1, \lambda_2, \lambda_3$. Suppose λ_1 is repeated three times, λ_2 is repeated two times and λ_3 occurs only once. Thus the six roots of the secular equation are $\lambda_1, \lambda_1, \lambda_1, \lambda_2, \lambda_2, \lambda_3$.

We say λ_1 is three-fold degenerate, λ_2 is two-fold degenerate and λ_3 is non-degenerate. We represent the order of degeneracy of a distinct eigenvalue λ_i by g_{λ_i} . In the present example, $g_{\lambda_1} = 3, g_{\lambda_2} = 2$ and $g_{\lambda_3} = 1$. We have

$$g_{\lambda_1} + g_{\lambda_2} + g_{\lambda_3} = 6 \quad \text{dimension of the vector space} \quad (5.24)$$

Now, it may be shown that, for a hermitian operator if a root λ is g -fold degenerate, there are always g linearly independent eigenvectors corresponding to λ . For a non-hermitian operator, there may not exist as many linearly independent eigenvectors as the order of degeneracy.

In the above example, if λ_1, λ_2 and λ_3 are eigenvalues of a hermitian operator, there are three linearly independent eigenvectors with eigenvalue λ_1 , two linearly independent eigenvectors with eigenvalue λ_2 and one eigenvector with eigenvalue λ_3 . Thus, the total number of linearly independent eigenvectors is six, the same as the dimension of the vector space. Hence these six linearly independent eigenvectors form a *complete basis set of vectors*.

If, however, λ_1, λ_2 and λ_3 are eigenvalues of a non-hermitian operator with the same eigenvalues, there may not exist three linearly independent eigenvectors with eigenvalue λ_1 , or two linearly independent eigenvectors with eigenvalue λ_2 . In such a situation, the number of linearly independent eigenvectors of the non-hermitian operator A is less than the dimension n of the vector space. Hence, these eigenvectors *do not form a basis set* for a n dimensional vector space.

5.5 Digonalization of a Hermitian Operator

Let A be a hermitian operator with distinct eigenvalues $\lambda_1, \lambda_2, \dots$. Some or all of the eigenvalues may be degenerate, with the order or degree of degeneracy of an eigenvalue λ_i being denoted by g_{λ_i} . If $g_{\lambda_j} = 1$ for some λ_j , then λ_j is said to be non-degenerate.

Since A is hermitian there will always be g_{λ_i} linearly independent eigenvectors, each belonging to the same eigenvalue λ_i . We will now require another index, $s^{(i)}$, to distinguish between these linearly independent eigenvectors. We write

$$A |\lambda_i, s^{(i)}\rangle = \lambda_i |\lambda_i, s^{(i)}\rangle \quad (5.25)$$

where $s^{(i)} = 1, 2, \dots, g_{\lambda_i}$. A linear combination of the degenerate eigenvectors is also an eigenvector with the same eigenvalue λ_i . So we have

$$A \left(\sum_{s^{(i)}}^{g_{\lambda_i}} C_{s^{(i)}} |\lambda_i, s^{(i)}\rangle \right) = \lambda_i \left(\sum_{s^{(i)}}^{g_{\lambda_i}} C_{s^{(i)}} |\lambda_i, s^{(i)}\rangle \right) \quad (5.26)$$

Thus, the set of vectors $\{|\lambda_i, s^{(i)}\rangle; \lambda_i \text{ fixed}, s^{(i)} = 1, 2, \dots, g_{\lambda_i}\}$ spans a subspace, called the eigen subspace of λ_i , of the original n dimensional vector space. The eigenvectors belonging to a degenerate eigenvalue need not be orthogonal to each other even if they are linearly independent, as the general theorem of hermitian operators proves the orthogonality of eigenvectors belonging to different eigenvalues.

However, using Schmidt orthonormalization procedure (see section (2.11)), we can get a set of g_{λ_i} orthogonal eigenfunctions of eigenvalue λ_i from a set of g_{λ_i} linearly independent set of eigenfunctions of eigenvalue λ_i .

Thus, all the eigenvectors of the hermitian operator, wherether belonging to same or different eigenvalues can be considered as orthogonal to each other. Further, they are also normalized. Using the set of orthonormal eigenfunctions as the basis, the matrix representation of A is diagonal.

As a concrete example of diagonalization of a hermitian operator, suppose we have a finite seven-dimensional linear vector space. If, all the eigenvectors are non-degenerate, then there are seven distinct eigenvalues $\lambda_1, \dots, \lambda_7$ and corresponding to each eigenvalue there will be one eigenvector $|\lambda_1\rangle, \dots, |\lambda_7\rangle$. These eigenvectors are orthogonal and they are normalized. Using the eigenvectors as

the basis, the matrix representation of A is

$$A = \begin{bmatrix} \lambda_1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \lambda_2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \lambda_3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \lambda_4 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \lambda_5 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \lambda_6 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \lambda_7 \end{bmatrix} \quad (5.27)$$

But if some of the eigenvalues are degenerate, then the number of distinct eigenvalues will be less than seven. Suppose that there are three distinct eigenvalues $\lambda_1, \lambda_2, \lambda_3$. Also suppose that λ_1 is three-fold degenerate and λ_2 and λ_3 both are two-fold degenerate. Thus $g_{\lambda_1} = 3, g_{\lambda_2} = 2, g_{\lambda_3} = 2$ and $g_{\lambda_1} + g_{\lambda_2} + g_{\lambda_3} = 7$ which is the dimension of the vector space.

There are three linearly independent (but not necessarily orthogonal) eigenvectors with eigenvalue λ_1 and two linearly independent eigenvectors for each eigenvalue λ_2 and λ_3 . The eigenvectors with eigenvalue λ_1 can be labeled as $|\lambda_1, s^{(1)}\rangle$ with $s^{(1)} = 1, 2, 3$, i.e., $|\lambda_1, 1\rangle, |\lambda_1, 2\rangle, |\lambda_1, 3\rangle$.

These three eigenvectors span a subspace of the original seven-dimensional vector space H . The subspace is called the eigensubspace of λ_1 and is denoted by H_{λ_1} or simply H_1 . The eigenvectors belonging to λ_2 and λ_3 are labeled similarly. Then two linearly independent eigenvectors with eigenvalue λ_2 span a two-dimensional subspace H_2 and the two linearly independent vectors belonging to λ_3 span the eigensubspace H_3 . These three subspaces make up the full vector space H . We write

$$H = H_1 \bigoplus H_2 \bigoplus H_3 \quad (5.28)$$

The seven linearly independent eigenvectors $\{|\lambda_i, s^{(i)}\rangle, s^{(i)} = 1, 2, \dots, g_{\lambda_i}, i = 1, 2, 3\}$ can now be used as a basis to find the matrix representation of A . If the basis vectors within an eigensubspace are not made orthogonal, the matrix representation of A is block-diagonal as shown below.

$$\begin{array}{c|cc|c|cc} a_{11} & a_{12} & a_{13} & \mathbf{0} & \mathbf{0} \\ \hline a_{21} & a_{22} & a_{23} & \mathbf{0} & \mathbf{0} \\ \hline a_{31} & a_{32} & a_{33} & & & \\ \hline \mathbf{0} & b_{11} & b_{12} & \mathbf{0} & & \\ & b_{21} & b_{22} & & & \\ \hline \mathbf{0} & \mathbf{0} & c_{11} & c_{12} & & \\ & & c_{21} & c_{22} & & \end{array} \quad (5.29)$$

Writing with basis

$$\begin{array}{ccccccc}
 & 1 & 2 & 3 & 4 & 5 & 6 & 7 \\
 & |\lambda_1, 1\rangle & |\lambda_1, 2\rangle & |\lambda_1, 3\rangle & |\lambda_2, 1\rangle & |\lambda_2, 2\rangle & |\lambda_3, 1\rangle & |\lambda_3, 2\rangle \\
 \langle \lambda_1, 1 | & \left[\begin{array}{ccccccc} a_{11} & a_{12} & a_{13} & 0 & 0 & 0 & 0 \\ a_{21} & a_{22} & a_{23} & 0 & 0 & 0 & 0 \\ a_{31} & a_{32} & a_{33} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & b_{11} & b_{12} & 0 & 0 \\ 0 & 0 & 0 & b_{21} & b_{22} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & c_{21} & c_{22} \\ 0 & 0 & 0 & 0 & 0 & c_{21} & c_{22} \end{array} \right] & & & & & (5.30) \\
 \langle \lambda_1, 2 | & & & & & & & \\
 \langle \lambda_1, 3 | & & & & & & & \\
 \langle \lambda_2, 1 | & & & & & & & \\
 \langle \lambda_2, 2 | & & & & & & & \\
 \langle \lambda_3, 1 | & & & & & & & \\
 \langle \lambda_3, 2 | & & & & & & &
 \end{array}$$

Each non-zero block is a square matrix. The first block is a 3×3 matrix, the second one is a 2×2 matrix and the third one is a 2×2 matrix. These blocks themselves are not diagonal if the basis vectors of the three eigensubspaces are not orthogonalized. If we orthogonalize the basis vector in each eigensubspace, then each block will also be diagonal. The matrix representation of A will then be

$$\begin{array}{ccccccc}
 & |\lambda_1, 1\rangle & |\lambda_1, 2\rangle & |\lambda_1, 3\rangle & |\lambda_2, 1\rangle & |\lambda_2, 2\rangle & |\lambda_3, 1\rangle & |\lambda_3, 2\rangle \\
 \langle \lambda_1, 1 | & \left[\begin{array}{ccccccc} \lambda_1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \lambda_1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \lambda_1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \lambda_2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \lambda_2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \lambda_3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \lambda_3 \end{array} \right] & & & & & (5.31) \\
 \langle \lambda_1, 2 | & & & & & & & \\
 \langle \lambda_1, 3 | & & & & & & & \\
 \langle \lambda_2, 1 | & & & & & & & \\
 \langle \lambda_2, 2 | & & & & & & & \\
 \langle \lambda_3, 1 | & & & & & & & \\
 \langle \lambda_3, 2 | & & & & & & &
 \end{array}$$

Thus the matrix representation of a hermitian operator A is diagonalized.

We have proved that a hermitian operator (or a hermitian matrix) is always diagonalizable in a finite dimensional vector space. By diagonalizable we mean that we can always find a basis in which the matrix representation of A is diagonal. This basis is simply the basis consisting of the orthogonalized eigenvectors of A , called eigenbasis of A .

The eigenvectors of a non-hermitian operator may be fewer in numbers than the dimension of the vector space if there is degeneracy. If an eigenvector λ_i is g_i -fold degenerate, then the number of

linearly independent eigenvectors belonging to λ_i may be less than g_i . Therefore, the eigenvectors of a non-hermitian operator cannot form a basis set for the vector space. Therefore, a non-hermitian operator is not diagonalizable.

5.6 Basis independence of the eigenvalues of an Operator

Basis independence refers to representation independence. To find the eigenvalues of a hermitian operator \hat{A} , first we choose an orthonormal basis set $\{|u_i\rangle\}$ and form the matrix representation of the operator. Then we solve the secular equation to find the eigenvalues. Although we have to integrate a basis set to find the eigenvalues, it is easy to verify that the eigenvalues are independent of the choice of the basis.

Indeed, if we choose a new orthogonal set of basis vectors $\{|u'_i\rangle\}$ which are related to the old set according to

$$|u'_i\rangle = \sum_j |u_j\rangle \langle u_j|u'_i\rangle \quad (5.32)$$

Then the new matrix representation of the operator A is related to the old representation by a similarly transformation with a *unitary* matrix. This is easy to see

$$\begin{aligned} A'_{ij} &\equiv \langle u'_i | A | u'_j \rangle \\ &= \sum_{jk} \langle u'_i | u_j \rangle \langle u_j | A | u_k \rangle \langle u_k | u'_j \rangle \\ &= \sum_{jk} S_{ij} A_{jk} S_{jk}^* \\ &= \sum_{jk} S_{ij} A_{jk} S_{jk}^\dagger \end{aligned} \quad (5.33)$$

Where we have defined the matrix S as

$$S_{ij} \equiv \langle u'_i | u_j \rangle \quad (5.34)$$

The matrix S is unitary as shown precisely. In matrix notation, we write equation (5.33) as,

$$A' = SAS^\dagger = SAS^{-1} \quad (5.35)$$

since S is unitary matrix. Then

$$\begin{aligned} \det(\mathbb{A}' - \lambda \mathbb{I}) &= \det(S \mathbb{A} S^{-1} - \lambda S \mathbb{I} S^{-1}) \\ &= \det(S(\mathbb{A} - \lambda \mathbb{I})S^{-1}) \\ &= \det(\mathbb{A} - \lambda \mathbb{I}) \end{aligned} \quad (5.36)$$

Thus, there is no change in the secular equation even if we change the basis set. Since the eigenvalues are the roots of the secular equation, the eigenvalues are representation independent. They are characteristics of the operator \hat{A} itself, and not of any particular representation.

Next, we will show that the determinant and the trace of the matrix representation A are independent of the basis used for the representation.

Since $A' = SAS^{-1}$ we have

$$\begin{aligned}\det(A') &= \det(SAS^{-1}) \\ &= \det(S^{-1}SA) \\ &= \det(A)\end{aligned}\tag{5.37}$$

i.e., the determinant is independent of the representation.

We also have

$$\text{Tr}(A') = \text{Tr}(SAS^{-1}) = \text{Tr}(S^{-1}SA) = \text{Tr}(A)\tag{5.38}$$

i.e., the trace is also independent of the representation. In the above derivations, we have used the identities,

$$\det(AB) = \det(BA)\tag{5.39}$$

$$\text{Tr}(AB) = \text{Tr}(BA)\tag{5.40}$$

where A and B are square matrices. Now, if we use the eigenbasis of the hermitian operator A for the representation, then A is a diagonal matrix with

$$\det A = \prod_i \lambda_i\tag{5.41}$$

$$\text{Tr } A = \sum_i \lambda_i\tag{5.42}$$

If an eigenvalue is g -fold degenerate, then that eigenvalue has to be repeated g times while calculating the determinant and trace of the matrix A .

5.7 Infinite dimensional vector space

We have shown that a linear operator is a finite n -dimensional vector space has n eigenvalues some of which may be repeated. If the operator is hermitian, then the eigenvalues are real and eigenvectors

belonging to different eigenvalues are orthogonal and hence linearly independent.

Further, if an eigenvalue λ of a hermitian operator is g -fold degenerate, then there are g linearly independent eigenvectors corresponding to λ , these degenerate eigenvectors are not necessarily orthogonal even if they are linearly independent. However, we can orthogonalize the degenerate eigenvectors using the Schmidt orthonormalization procedure (see section (2.11)).

Thus in a finite n -dimensional vector space, the eigenvectors of any hermitian operator form a set of orthonormal basis vectors.

In an infinite dimensional vector space, the number of eigenvalues and eigenvectors of a hermitian operators are infinitely many. However, it is possible that the eigenvectors of some hermitian operators do not form a complete set in an infinite dimensional vector space.

Hermitian operators are of vital importance in quantum mechanics because to every observable (e.g., position, linear momentum, angular momentum, spin etc.) we *associate* a corresponding hermitian operator. Of course, there are hermitian operators which are not associated with any observable.

The eigenvectors of a hermitian operator representing a physical observable form a complete set even in an infinite-dimensional Hilbert space. The eigenvectors of a hermitian operator not associated with any observable may not form a complete basis set in an infinite dimensional space.

5.8 Completeness condition for the eigenvectors of a Hermitian Operator

Let us assume that the eigenvalue spectrum of a hermitian operator \hat{A} form a discrete set. In other words, the eigenvalues $a_i, i = 1, 2, \dots$ of the operator are discrete real numbers.

Assume, for the time being, that the eigenvalues are non-degenerate so that there is only one linearly independent eigenvectors $|a_i\rangle$ corresponding to each eigenvalue a_i . The eigenvectors $\{|a_i\rangle, i = 1, 2, \dots\}$ form a complete orthogonal set of basis vectors. Therefore, an arbitrary vector $|\psi\rangle$ of the vector space can be expanded as a linear combination of the vector in the basis set, i.e.,

$$|\psi\rangle = \sum_i a_i |a_i\rangle \quad (5.43)$$

where $c_i = \langle a_i | \psi \rangle$. Therefore, we can write

$$|\psi\rangle = \sum_i \langle a_i | \psi \rangle |a_i\rangle = \sum_i |a_i\rangle \langle a_i | \psi \rangle \quad (5.44)$$

Since $|\psi\rangle$ is arbitrary, we must have

$$\hat{\mathbb{1}} = \sum_i |a_i\rangle \langle a_i| = \sum_i \hat{P}_i \quad (5.45)$$

where

$$\hat{P}_i = |a_i\rangle \langle a_i| \quad (5.46)$$

is the projection along $|a_i\rangle$.

Using the basis $\{|u_i\rangle\}$, any operator \hat{O} can be expressed as

$$\begin{aligned} \hat{O} &= \hat{\mathbb{1}} \hat{O} \hat{\mathbb{1}} = \sum_{i,j} |a_i\rangle \langle a_i| \hat{O} |a_j\rangle \langle a_j| \\ &= \sum_{i,j} |a_i\rangle O_{ij} \langle a_j| \end{aligned} \quad (5.47)$$

where $O_{ij} \equiv \langle a_i | \hat{O} | a_j \rangle$ are the matrix element of \hat{O} in the basis $\{|u_i\rangle\}$. Since basis is the eigenbasis of the operator \hat{A} , the matrix elements of \hat{A} in the basis will be diagonal, i.e.,

$$\hat{A} = \sum_i a_i |a_i\rangle \langle a_i| = \sum_i a_i \hat{P}_i \quad (5.48)$$

Any other operator \hat{B} will in general not be diagonal in the eigenbasis of \hat{A} unless the eigenvectors of \hat{B} and \hat{A} coincide. Later, we will see that two operators \hat{A} and \hat{B} have simultaneous eigenvectors if they commute, i.e., if $[\hat{A}, \hat{B}] = 0$.

Now, we will generalize the notation to include degeneracy, suppose the eigenvalue a_i is a g_i fold degenerate. Then the eigenvectors belonging to the eigenvalue a_i is written as $|a_i, s^{(i)}\rangle$ where $s^{(i)}$ can take values $1, 2, \dots, g_i$. The set of vectors $\{|a_i, s^{(i)}\rangle, s^{(i)} = 1, 2, \dots, g_i; i = 1, 2, \dots\}$ Form a complete orthogonal set. The completeness condition is

$$\sum_{i=1}^{\infty} \sum_{s^i=1}^{g_i} |a_i, s^{(i)}\rangle \langle a_i, s^{(i)}| = \hat{\mathbb{1}} \quad (5.49)$$

and the orthogonality condition is

$$\langle a_i, s^{(i)} | a_j, s^{(j)} \rangle = \delta_{ij} \delta_{s^i s^{(j)}} \quad (5.50)$$

We can rewrite equation (5.49) as (exactly as the non degenerate case)

$$\hat{\mathbb{1}} = \sum_i \hat{P}_i \quad (5.51)$$

where

$$\hat{P}_i = \sum_{s^{(i)}=1}^{g_i} |a_i, s^{(i)}\rangle \langle a_j, s^{(j)}| \quad (5.52)$$

is the projection operator on the eigensubspace of a_i . The operator \hat{A} can then be written in its own eigenbasis as $\hat{A} = \sum_i a_i \hat{P}_i$ with \hat{P}_i given in equation (5.52)

5.9 Hermitian operator with continuous eigenvalue spectrum

In Quantum Mechanics we encounter hermitian operator like position operator, momentum operator whose eigenvalues range over a continuum of real values. Such an eigenvalue spectrum is called continuous. There are also hermitian operators whose eigenvalue spectrum may be both discrete and continuous.

5.9.1 Continuous Spectrum

Let us consider an operator A whose eigenvalues can vary continuously over a certain domain of real numbers

$$A |a\rangle = a |a\rangle \quad (5.53)$$

If there is degeneracy, we will put in a second index s to distinguish between degenerate vectors. Thus we may write $|a\ s\rangle$ to denote a degenerate eigenvector. We assume that there is no degeneracy. In case of degeneracy it is a simple matter to generalize our notations. We assume that the vectors $|a\rangle$ form a complete set. The completeness condition can be written as

$$\int da |a\rangle \langle a| = \hat{1} \quad (5.54)$$

Where the integral extends over the entire domain in which a varies. Usually this domain is $-\infty$ to ∞ .

Two eigenkets $|a\rangle$ and $|a'\rangle$ with $a \neq a'$ are orthogonal because A is hermitian operator, i.e.,

$$\langle a|a'\rangle = 0; \quad a \neq a' \quad (5.55)$$

What will the scalar product be if $a = a'$? Can we take $\langle a|a\rangle = 1$ as in the discrete case where we normalized the eigenkets as $\langle a_i|a_i\rangle = 1$?

The answer is **no**, i.e., in the case where the eigenvalues a vary continuously, the kets $|a\rangle$ cannot be normalized to unity. To see this, expand an arbitrary ket $|f\rangle$ in the eigenbasis $\{|a\rangle\}$ of the operator

\hat{A} . We have

$$|f\rangle = \int da' |a'\rangle \langle a'|f\rangle \quad (5.56)$$

Taking the scalar product of $|f\rangle$ with $|a\rangle$, we get

$$\begin{aligned} \langle a|f\rangle &= \int da' |a'\rangle \langle a'|f\rangle \\ f(a) &= \int da' |a'\rangle f(a') \end{aligned} \quad (5.57)$$

Where we have defined $f(a)$ as $f(a) = \langle a|f\rangle$. In order for equation (5.57) to be valid, we must have

$$\langle a|a'\rangle = \delta(a - a') \quad (5.58)$$

for, with this choice, the right side of equation (5.57) becomes equal to the left side:

$$f(a) = \int da' \delta(a - a') f(a') \quad (5.59)$$

Thus, setting $a' = a$ in equation (5.58) we find

$$\langle a|a\rangle = \delta(0) = \infty \quad (5.60)$$

In other words, the eigenkets $\{|a\rangle\}$ are not normalizable to unity since $\langle a|a\rangle$ is not finite. Therefore, the eigenkets $\langle a|a\rangle$ do not belong to the Hilbert space. However, we can include such eigenkets in the vector space, and the augmented vector space is called the physical Hilbert space.

The kets $\{|a\rangle\}$ are not physically realizable in the sense that no physical state of a system can have a state vector $|\psi\rangle$ which is one of the eigenkets $|a\rangle$. However, the set of eigenkets $\{|a\rangle\}$ can form a basis set because arbitrary ket $|\psi\rangle$ of finite norm can always be expanded in terms of $\{|a\rangle\}$.

As a matter of terminology, we say that the eigenkets belonging to continuously varying eigenvalues of a hermitian operator are "normalizable" to a delta function, i.e. $\langle a|a'\rangle = \delta(a - a')$, even though the kets $|a\rangle$ are not normalizable in the strict mathematical sense, since

$$\| |a\rangle \| = \infty \quad (5.61)$$

In summary, for continuously varying eigenvalues, the orthogonality (5.62) and completeness (5.63) of the eigenvectors of a hermitian operator are written as

$$\langle a|a'\rangle = \delta(a - a') \quad (5.62)$$

$$\hat{\mathbb{1}} = \int da |a\rangle \langle a| \quad (5.63)$$

5.10 Hermitian operator with continuous and discrete eigenvalue

The eigenvalue spectrum of a hermitian operator can be both discrete and continuous, In such a situation we have

$$\hat{A}|a_i\rangle = a_i|a_i\rangle; \quad i = 1, 2, \dots \quad (5.64)$$

for discrete eigenvalues, and

$$\hat{A}|a\rangle = a|a\rangle; \quad a \in D \subset R \quad (5.65)$$

for continuous eigenvalues. The completeness condition is

$$\sum_i |a_i\rangle\langle a_i| + \int da |a\rangle\langle a| = \hat{\mathbb{1}} \quad (5.66)$$

and the orthogonality condition are

$$\langle a_i | a_j \rangle = \delta_{ij} \quad (5.67)$$

$$\langle a | a' \rangle = \delta(a - a') \quad (5.68)$$

$$\langle a_i | a \rangle = 0 \quad (5.69)$$

5.11 Problems

- Find the eigenvalues and the corresponding eigenvectors of the matrix

$$M = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \quad (5.70)$$

can this matrix be diagonalized?

Ans.

The eigenvalue equation is

$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \lambda \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \quad (5.71)$$

The secular equation is then

$$\begin{aligned} \det(M - \lambda \mathbb{1}) &= 0 \\ \begin{vmatrix} 1-\lambda & 1 \\ 0 & 1-\lambda \end{vmatrix} &= 0 \\ (1-\lambda)^2 &= 0 \end{aligned}$$

i.e., $\lambda = 1, 1$ (2 fold degeneracy)

Eigenvector

With $\lambda = 1$ the eigenvalue equation is

$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 & x_2 \end{bmatrix} = 1 \begin{bmatrix} x_1 & x_2 \end{bmatrix}$$

$$\begin{bmatrix} x_1 + x_2 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \quad (5.72)$$

Thus

$$x_1 + x_2 = x_1$$

$$x_2 = 0$$

The element x_1 is arbitrary. Hence

$$|1\rangle = \begin{bmatrix} x_1 \\ 0 \end{bmatrix} \quad (5.73)$$

Normalizing

$$|1\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad (5.74)$$

We have found just one linearly independent eigenvector with $\lambda = 1$. Since M is not hermitian, there is no guarantee that there would be two linearly independent eigenvectors for a two fold degenerate eigenvalue. Here, for the given matrix M , which is non-hermitian, we have only one linearly independent eigenvector corresponding to the two-fold degenerate eigenvalue $\lambda = 1$. So we do not have a complete set of eigenvectors of M to span the two dimensional vector space. Hence M is not diagonalizable by a change of basis, i.e., by a similarity transformation.

2. Find the eigenvalues and the corresponding eigenvectors of the matrix

$$A = \begin{bmatrix} 3 & i \\ -i & 3 \end{bmatrix} \quad (5.75)$$

Ans.

First note that $A^\dagger = A$, i.e., the matrix is hermitian. Hence the eigenvalues would be real and the eigenvectors belonging to distinct eigenvalues would be orthogonal.

The eigenvalue equation is

$$\begin{bmatrix} 3 & i \\ -i & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \lambda \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$\begin{bmatrix} 3 - \lambda & i \\ -i & 3 - \lambda \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \lambda \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad (5.76)$$

The secular equation is

$$\begin{aligned} \begin{vmatrix} 3-\lambda & i \\ -i & 3-\lambda \end{vmatrix} &= 0 \\ (3-\lambda)^2 - (i)(-i) &= 0 \\ (\lambda - 3)^2 - 1 &= 0 \\ (\lambda - 3 - 1)(\lambda - 3 + 1) &= 0 \\ \lambda = 2, 4 & \end{aligned} \tag{5.77}$$

None of the roots are degenerate.

Eigenvector for $\lambda = 2$

$$\begin{bmatrix} 3-2 & i \\ -i & 3-2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \emptyset$$

$$\begin{bmatrix} x_1 + ix_2 \\ -ix_1 + x_2 \end{bmatrix} = \emptyset$$

Thus

$$x_1 + ix_2 = 0 \tag{5.78}$$

$$-ix_1 + x_2 = 0 \tag{5.79}$$

We get the solution $x_1 = -ix_2$. Taking x_2 to be arbitrary

$$|2\rangle = \begin{bmatrix} -ix_2 \\ x_2 \end{bmatrix} \tag{5.80}$$

Normalizing

$$\begin{aligned} \langle 2|2\rangle &= 1 \\ \begin{bmatrix} ix_2^* & x_2^* \end{bmatrix} \begin{bmatrix} -ix_2 \\ x_2 \end{bmatrix} &= 1 \\ 2|x_2|^2 &= 1 \\ |x_2| &= \frac{1}{\sqrt{2}} \end{aligned} \tag{5.81}$$

Take $x_2 \doteq \frac{1}{\sqrt{2}}$.

We could have taken

$$x_2 = -\frac{1}{\sqrt{2}} \tag{5.82}$$

or

$$x_2 = e^{i\phi} \frac{1}{\sqrt{2}} \tag{5.83}$$

In all cases $|x_2| = \frac{1}{\sqrt{2}}$

Normalized eigenvector $|2\rangle$ is

$$|2\rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} -i \\ 1 \end{bmatrix} \quad (5.84)$$

Eigenvector for $\lambda = 4$

$$\begin{bmatrix} 3-4 & i \\ -i & 3-4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \emptyset$$

$$\begin{bmatrix} -x_1 + ix_2 \\ -ix_1 - x_2 \end{bmatrix} = \emptyset$$

Thus

$$-x_1 + ix_2 = 0 \quad (5.85)$$

$$-ix_1 - x_2 = 0 \quad (5.86)$$

We get the solution $x_1 = ix_2$. Taking x_2 to be arbitrary

$$|4\rangle = \begin{bmatrix} ix_2 \\ x_2 \end{bmatrix} \quad (5.87)$$

The value of x_2 has to be found from normalization

$$\begin{aligned} \langle 4 | 4 \rangle &= 1 \\ \begin{bmatrix} -ix_2^* & x_2^* \end{bmatrix} \begin{bmatrix} ix_2 \\ x_2 \end{bmatrix} &= 1 \\ 2|x_2|^2 &= 1 \\ |x_2| &= \frac{1}{\sqrt{2}} \end{aligned} \quad (5.88)$$

Take $x_2 = \frac{1}{\sqrt{2}}$. Normalized eigenvector $|4\rangle$ is

$$|4\rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} i \\ 1 \end{bmatrix} \quad (5.89)$$

Orthogonality of the eigenvectors

$$|2\rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} -i \\ 1 \end{bmatrix} \quad (5.90)$$

$$|4\rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} i \\ 1 \end{bmatrix} \quad (5.91)$$

$$\langle 2 | 4 \rangle = \frac{1}{2} \begin{bmatrix} i & 1 \end{bmatrix} \begin{bmatrix} i \\ 1 \end{bmatrix} = \frac{1}{2} (i^2 + 1) = 0 \quad (5.92)$$

If we take $|2\rangle$ and $|4\rangle$ as the basis the matrix representation of \hat{A} is

$$\hat{A} \rightarrow \begin{bmatrix} \langle 2 | \hat{A} | 2 \rangle & \langle 2 | \hat{A} | 4 \rangle \\ \langle 4 | \hat{A} | 2 \rangle & \langle 4 | \hat{A} | 4 \rangle \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 4 \end{bmatrix} \quad (5.93)$$

Now the similarity transformation that diagonalizes the matrix A . First the matrix S that has the columns as the eigenvectors is

$$S = \begin{bmatrix} |2\rangle & |4\rangle \\ -i/\sqrt{2} & i/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} \quad (5.94)$$

Therefore,

$$A' = \langle 2 | \begin{bmatrix} i/\sqrt{2} & 1/\sqrt{2} \\ -i/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} A \begin{bmatrix} -i/\sqrt{2} & i/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} | 4 \rangle \quad (5.95)$$

$$= \langle 2 | \begin{bmatrix} i/\sqrt{2} & 1/\sqrt{2} \\ -i/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} \begin{bmatrix} 3 & i \\ -i & 3 \end{bmatrix} \begin{bmatrix} -i/\sqrt{2} & i/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} | 4 \rangle \quad (5.96)$$

$$= S^{-1}AS \quad (5.97)$$

3. Find the eigenvalues and the corresponding eigenvectors of the matrix

$$M = \frac{1}{2} \begin{bmatrix} 3 & -1 & 0 \\ -1 & 30 & 0 \\ 0 & 0 & 2 \end{bmatrix} \quad (5.98)$$

Ans.

The matrix M is hermitian. Therefore the eigenvalues are real. The eigenvalues are obtained by solving the secular equation

$$\begin{vmatrix} \frac{3}{2} - \lambda & -\frac{1}{2} & 0 \\ -\frac{1}{2} & \frac{3}{2} - \lambda & 0 \\ 0 & 0 & 1 - \lambda \end{vmatrix} = 0$$

$$(1 - \lambda) \left[\left(\frac{3}{2} - \lambda \right)^2 - \left(-\frac{1}{2} \right) \left(-\frac{1}{2} \right) \right] = 0$$

$$(\lambda - 1)[(\lambda - 3/2 + 1/2)(\lambda - 3/2 - 1/2)] = 0$$

$$(\lambda - 1)(\lambda - 1)(\lambda - 2) = 0$$

Thus the eigenvalues are $\lambda = 1, 1, 2$. The eigenvalue 1 is two fold degenerate and the eigenvalue 2 is non-degenerate, the two distinct eigenvalues are $\lambda_1 = 1$ with $g_1 = 2$ and $\lambda_2 = 2$ with $g_2 = 1$.

Eigenvector for $\lambda = 1$

Since M is hermitian, there will be two linearly independent eigenvectors corresponding to $\lambda = 1$. We will make the two linearly independent eigenvectors orthogonal. The eigenvalue equation is

$$\begin{aligned} \mathbb{A} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} &= 1 \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \\ \begin{bmatrix} \frac{3}{2} - 1 & -\frac{1}{2} & 0 \\ -\frac{1}{2} & \frac{3}{2} - 1 & 0 \\ 0 & 0 & 1 - 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} &= \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \\ \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} & 0 \\ -\frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} &= \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \\ \begin{bmatrix} 1 & -1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} &= \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \end{aligned}$$

Therefore $x_1 = x_2 = x$ (say) with arbitrary x . Also x_3 is arbitrary. Hence

$$|1\rangle = \begin{bmatrix} x \\ x \\ x_3 \end{bmatrix} \quad (5.99)$$

Choose $x = 1$ and $x_3 = 0$ so that

$$|1\rangle^{(1)} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \quad (5.100)$$

Normalizing

$$|1\rangle^{(1)} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = |\lambda = 1, s = 1\rangle \quad (5.101)$$

Next Choose $x = 0$ and $x_3 = 1$

$$|1\rangle^{(2)} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = |\lambda = 1, s = 2\rangle \quad (5.102)$$

These are two orthogonal eigenvectors with eigenvalue $\lambda = 1$.

Eigenvector for $\lambda = 2$

Here $g_{\lambda=2} = 1$. The eigenvalue equation is

$$\begin{bmatrix} \frac{3}{2} - 2 & -\frac{1}{2} & 0 \\ -\frac{1}{2} & \frac{3}{2} - 2 & 0 \\ 0 & 0 & 1 - 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} -\frac{1}{2} & -\frac{1}{2} & 0 \\ -\frac{1}{2} & -\frac{1}{2} & 0 \\ 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

Therefore $x_1 = -x_2 = x$ (say) with arbitrary x . Also $x_3 = 0$. Therefore, eigenvector $|2\rangle$ is of the form

$$|2\rangle = \begin{bmatrix} x \\ -x \\ 0 \end{bmatrix} \quad (5.103)$$

Normalizing

$$|2\rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} = |\lambda = 1, s = 1\rangle \quad (5.104)$$

Similarity Transformation

$$M' = SMS^\dagger \quad (5.105)$$

Now

$$S^\dagger = \begin{bmatrix} |1,1\rangle & |1,2\rangle & |2\rangle \\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} \\ 0 & 1 & 0 \end{bmatrix} \quad (5.106)$$

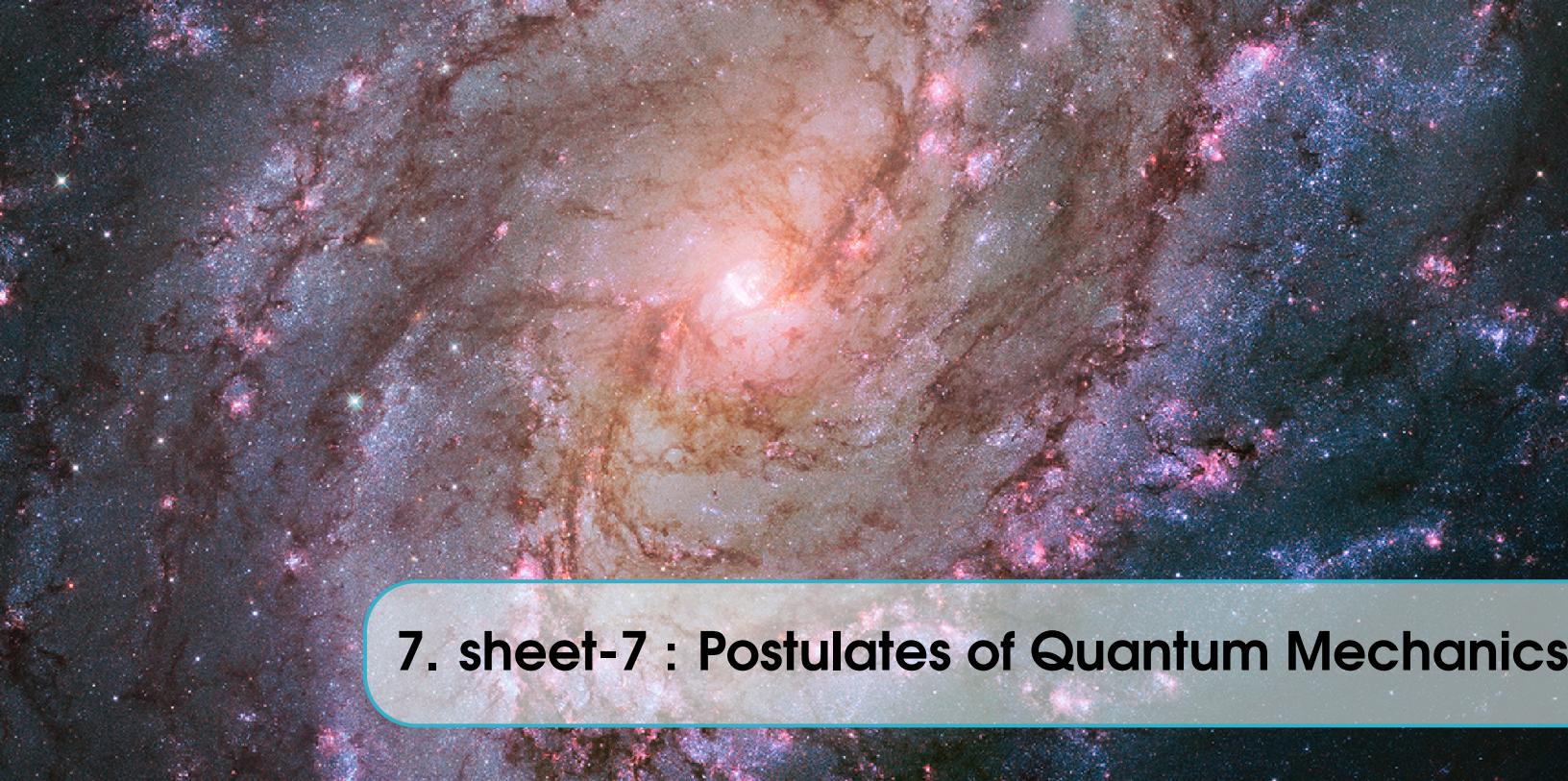
$$S = (S^\dagger)^\dagger = \begin{bmatrix} \langle 1,1| & \langle 1,2| & \langle 2| \\ \langle 1,1| & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ \langle 1,2| & 0 & 0 & 1 \\ \langle 2| & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \end{bmatrix} \quad (5.107)$$

Then the matrix M' is diagonal

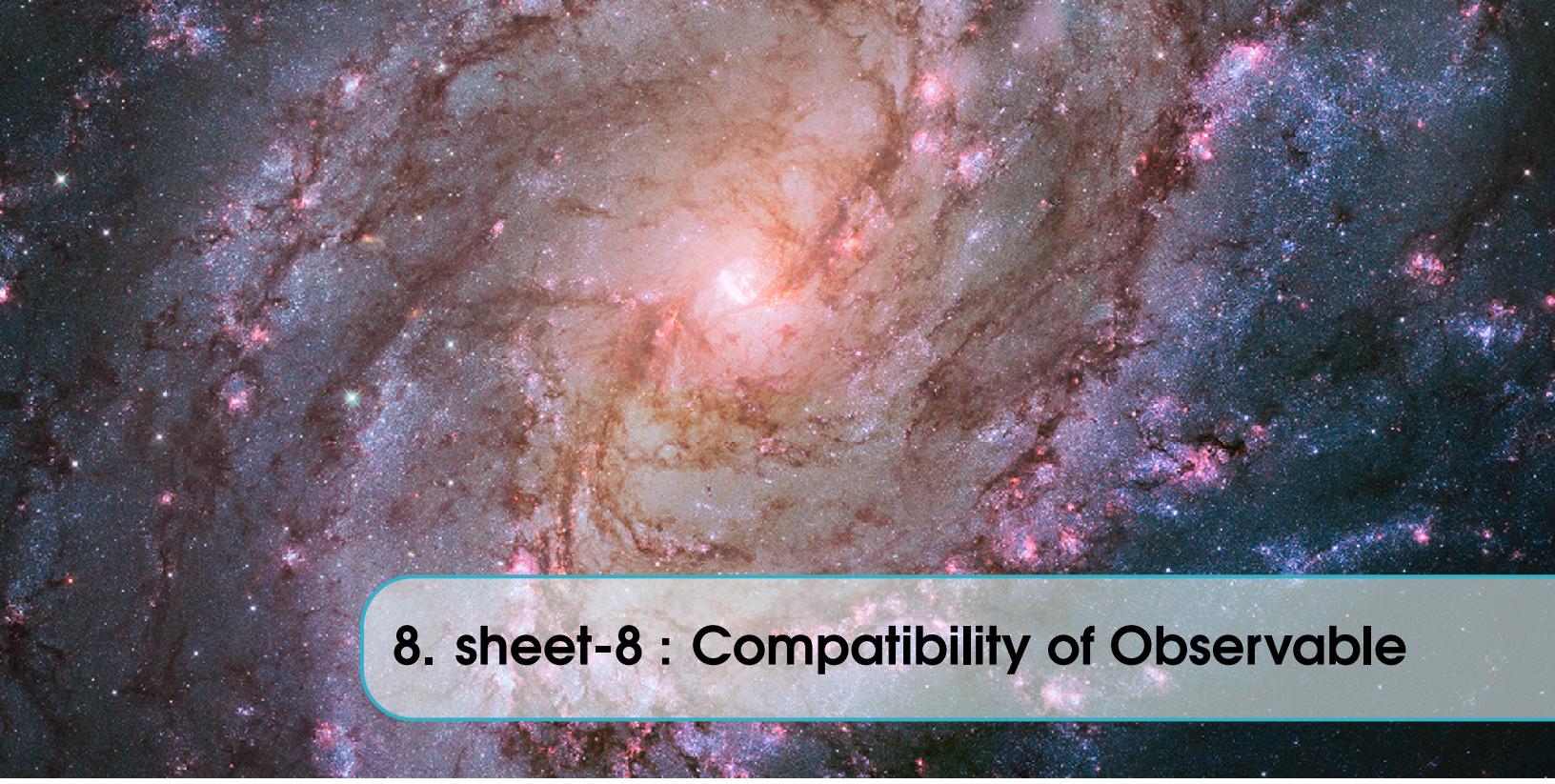
$$M' = S M S^\dagger = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix} \quad (5.108)$$



6. sheet-6 : Coordinate and Momentum Repre



7. sheet-7 : Postulates of Quantum Mechanics



8. sheet-8 : Compatibility of Observable

Lecture 8 Compatibility of Observable : version 1

8.1 Background Materials

Suppose that \hat{A} is the Hermitian operator representing an observable A of a quantum system. The eigenkets of \hat{A} form a complete set of orthonormal basis states. Let $|\psi\rangle$ be the state vector of the system. We can expand $|\psi\rangle$ in the eigenbasis of the operator \hat{A} :

$$|\psi\rangle = \sum_a |a\rangle \langle a| \psi \rangle , \quad (8.1)$$

where a runs over all the eigenvalues of \hat{A} .

In Eq. (8.44), we have assumed that each eigenvalue a is non-degenerate, i.e., for each a there exists only one linearly independent eigenvector $|a\rangle$. Therefore, the eigenvalue itself can be used to label the corresponding eigenket unambiguously.

However, it may so happen that some or all of the eigenvalues of \hat{A} are degenerate, i.e., there may be more than one linearly independent eigenvector corresponding to the same eigenvalue. The number of linearly independent eigenvectors corresponding to a particular eigenvalue a is called the order of degeneracy of the eigenvalue and is denoted by g_a . If the eigenvalue a is degenerate, then just the

eigenvalue itself is not enough to label the eigenstates uniquely. We need another index to distinguish between the g_a linearly independent eigenvectors.

Thus the eigenvectors belonging to a degenerate eigenvalue a may be denoted by $|a, i\rangle$, where the index i can take discrete values $i = 1, 2, 3, \dots, g_a$. The index i may be called the degeneracy index. Later we will see that we can improve the notation by replacing i by the eigenvalues of other Hermitian operators \hat{B}, \hat{C}, \dots , which commute with \hat{A} .

The set of linearly independent eigenvectors $\{|a, i\rangle, i = 1, 2, \dots, g_a\}$, all with the same eigenvalue a , span a g_a -dimensional subspace of the Hilbert space. This subspace is called the eigensubspace of a , and is denoted by H_a . The totality of all the eigensubspaces of the operator \hat{A} constitutes the full Hilbert space.

It is easy to see that any linear combination of the eigenvectors $\{|a, i\rangle, i = 1, 2, \dots, g_a\}$ is also an eigenvector of \hat{A} with the same eigenvalue a . Thus

$$\hat{A} \left(\sum_{i=1}^{g_a} c_i |a, i\rangle \right) = a \left(\sum_{i=1}^{g_a} c_i |a, i\rangle \right), \quad (8.2)$$

where c_i 's are constants. The set of eigenvectors $\{|a, i\rangle, i = 1, 2, \dots, g_a\}$, even though linearly independent, may not be orthogonal to each other because they belong to the same eigenvalue. However, following the Schmidt orthonormalization procedure, we can take linear combinations of the above set of vectors in a special way and obtain a set of g_a orthonormal vectors. The new set of orthonormal (and hence linearly independent) vectors remain eigenvectors of \hat{A} with the same eigenvalue a .

We assume that the Schmidt procedure has been carried out in each eigensubspace. So, in each eigensubspace, we can write

$$\langle a, i | a, j \rangle = \delta_{ij}, \quad (8.3)$$

where $i, j = 1, 2, 3, \dots, g_a$. The orthonormal set of eigenvectors $\{|a, i\rangle, i = 1, 2, \dots, g_a\}$ can be said to span the eigensubspace H_a . The eigenvectors in different eigensubspaces are automatically orthogonal because \hat{A} is Hermitian. Hence, we can assume that **all** eigenvectors of \hat{A} are orthonormal, i.e.,

$$\langle a', i' | a, i \rangle = \delta_{aa'} \delta_{ii'}. \quad (8.4)$$

The full set $\{|a, i\rangle, i = 1, 2, \dots, g_a; a = a_1, a_2, \dots\}$ of eigenvectors of \hat{A} form a complete orthonormal set, i.e., they constitute a basis set for the Hilbert space. The completeness condition of this basis set can be written as

$$\sum_a \sum_{i=1}^{g_a} |a, i\rangle \langle a, i| = \hat{I}. \quad (8.5)$$

Thus the state vector $|\psi\rangle$ of the system can be expanded as

$$\begin{aligned} |\psi\rangle &= \hat{I}|\psi\rangle \\ &= \sum_a \sum_{i=1}^{g_a} |a, i\rangle \langle a, i| \psi \rangle. \end{aligned} \quad (8.6)$$

For simplicity, we will continue to use Eq. (8.44) as the expansion for $|\psi\rangle$ in the eigenbasis of \hat{A} assuming that all the eigenvalues are non-degenerate. In case of degeneracy, notations can be generalized in a straightforward manner, as discussed above.

8.2 Measurement of Two Observables in Quick Succession

Suppose that a quantum system is in the state $|\psi\rangle$. Let A be an observable of the system with corresponding Hermitian operator \hat{A} . We can always expand $|\psi\rangle$ using the eigenbasis of \hat{A} :

$$|\psi\rangle = \sum_a |a\rangle \langle a| \psi \rangle, \quad (8.7)$$

where a runs over the eigenvalue spectrum of \hat{A} . The complex number $\langle a| \psi \rangle$ is the ‘component’ of $|\psi\rangle$ along $|a\rangle$.

We now make a measurement of the observable A on the system in the state $|\psi\rangle$. Since a general state $|\psi\rangle$ may be a superposition of many (perhaps infinitely many) eigenkets of \hat{A} with different eigenvalues, we cannot exactly predict the result of the experiment but can only say that the experiment would yield any of the eigenvalues of \hat{A} for which $\langle a| \psi \rangle \neq 0$.

The outcome is random, i.e., probabilistic, because we cannot predict exactly which eigenvalue will be obtained, but we can assign a probability for obtaining a particular eigenvalue a provided we know the state $|\psi\rangle$ before the measurement. The probability is

$$P_{|\psi\rangle}(a) = |\langle a| \psi \rangle|^2. \quad (8.8)$$

In the course of the measurement, the state of the system collapses to the eigenket $|a\rangle$ if the eigenvalue a is obtained in the measurement. Thus

$$|\psi\rangle \xrightarrow{\text{measurement of A}} |a\rangle. \quad (8.9)$$

A note on the notation is now in order. More generally, the eigenvalue a may be degenerate, and we should say that $|\psi\rangle$ collapses to its normalized projection in the eigensubspace H_a . Thus letting \hat{P}_a denote the projection operator on H_a , we have

$$\hat{P}_a = \sum_{i=1}^{g_a} |a, i\rangle\langle a, i|, \quad (8.10)$$

where i is the degeneracy index. If the eigenvalue a is non-degenerate, the eigensubspace H_a is one-dimensional and the extra index i is not needed, so that the projection operator is simply

$$\hat{P}_a = |a\rangle\langle a| \quad (a \text{ is nondegenerate}). \quad (8.11)$$

Note that \hat{P}_a is Hermitian and satisfies the relation

$$\hat{P}_a^2 = \hat{P}_a. \quad (8.12)$$

After the A-measurement, if the eigenvalue a is obtained, then the collapsed state can be written as

$$|\psi\rangle \longrightarrow \frac{\hat{P}_a|\psi\rangle}{\sqrt{\langle\hat{P}_a|\psi|\hat{P}_a\rangle}} = \frac{\hat{P}_a|\psi\rangle}{\sqrt{\langle\psi|\hat{P}_a|\psi\rangle}} \quad (8.13)$$

Using Eq. (8.10), the collapsed state can be written as

$$\psi \longrightarrow \frac{\sum_{i=1}^{g_a} |a, i\rangle\langle a, i|\psi\rangle}{\sqrt{\sum_{i=1}^{g_a} |\langle a, i|\psi\rangle|^2}} \quad (8.14)$$

If the eigenvalue a is non-degenerate, then the collapsed state is simply

$$\psi \longrightarrow \frac{|a\rangle\langle a|\psi\rangle}{\sqrt{|\langle a|\psi\rangle|^2}} = |a\rangle. \quad (8.15)$$

In any case, whether the eigenvalue a is degenerate or not, the state of the system collapses to the eigenstate of the operator \hat{A} with eigenvalue a . Now the system is in a state with a definite value for the observable A . Further successive measurements of A will yield the same eigenvalue a with 100% certainty.

Next, consider another observable B of the system with corresponding Hermitian operator \hat{B} . If, immediately after the measurement of A , we measure B , are we certain to get a particular eigenvalue b of the operator \hat{B} ? We know that the state of the system after the measurement of A is an eigenstate of \hat{A} , i.e., a state with a definite value a of the observable A . The question we have asked can be restated as follows: in the collapsed state $|a\rangle$, does B have a definite value, i.e., is $|a\rangle$ also an eigenstate of \hat{B} with some eigenvalue b ? The answer to this question is, in general, in the negative, i.e., $|a\rangle$ is not an eigenket of \hat{B} in general. The very important exceptional case where $|a\rangle$ is also an eigenstate of \hat{B} is discussed later.

Since the eigenkets of \hat{B} form a basis set, we can expand $|a\rangle$ as

$$|a\rangle = \sum_b |b\rangle \langle b|a\rangle. \quad (8.16)$$

Therefore, a measurement of B can yield any one of the eigenvalues b for which $\langle b|a\rangle \neq 0$. The probability of obtaining a particular eigenvalue b while the system is in the state $|a\rangle$ is

$$P_{|a\rangle}(b) = |\langle b|a\rangle|^2. \quad (8.17)$$

The state of the system after B is measured collapses from $|a\rangle$ to $|b\rangle$:

$$|a\rangle \xrightarrow{\text{measurement of } B} |b\rangle. \quad (8.18)$$

Thus, the measurement of B generally alters the state of the system because, due to the measurement process, the system is thrown into the eigenstate $|b\rangle$, where b is the eigenvalue of \hat{B} obtained in the measurement. But, since $|b\rangle$ is not, in general, an eigenstate of \hat{A} , the system is no longer in a state with a definite value for the observable A .

If we measure A again immediately after B is measured, we will not get the answer we got the first time, namely a , with certainty. In the second measurement of A there is the possibility of getting any eigenvalue for which $|\langle a|b\rangle| \neq 0$. To repeat, after the first measurement of A , the system is in the eigenstate $|a\rangle$ of \hat{A} but this state is not an eigenstate of the operator \hat{B} except in a special situation discussed below. After the measurement of B , the system is thrown in the eigenstate $|b\rangle$ of \hat{B} which is not an eigenstate of \hat{A} . Therefore, it is not possible, in general, to get the system in a state in which both the observables A and B have definite values.

Now, in successive measurements of A and B , first A then B , on a system initially in the state $|\psi\rangle$, the probability of obtaining the results a and b is

$$P(a, b) = P_{|\psi\rangle}(a)P_{|a\rangle}(b) = |\langle a|\psi\rangle|^2 |\langle b|a\rangle|^2. \quad (8.19)$$

If we reversed the order of measurements, first B and then A , the probability for obtaining the results b for the observable B and a for the observable A would be

$$P(b, a) = P_{|\psi\rangle}(b)P_{|b\rangle}(a) = |\langle b|\psi\rangle|^2 |\langle a|b\rangle|^2. \quad (8.20)$$

We note that, since $|\langle a|\psi\rangle|^2 \neq |\langle b|\psi\rangle|^2$, the two probabilities are not equal in general, i.e., $P(a, b) \neq P(b, a)$.

8.3 Compatible Observables

Now, let us return to the exceptional situation mentioned in the previous section. To recapitulate, let us make two measurements of two observables A and B in quick succession on a system in the state $|\psi\rangle$. First, we measure A , and if the eigenvalue a is obtained, the system's state collapses to an eigenstate of the operator \hat{A} belonging to the eigenvalue a . Then if we measure B immediately afterward, would we be certain to get a particular eigenvalue b of \hat{B} ? Further, if a second measurement of A is made immediately after the measurement of B , what conditions need be fulfilled in order that the result of the first measurement is unaltered, i.e., we will be certain to get the same result a again? We will first consider the case when all eigenvalues of \hat{A} are non-degenerate and then we will consider degeneracy of the eigenvalues.

8.3.1 Nondegenerate Case

Suppose all eigenvalues of \hat{A} are nondegenerate, i.e., for every eigenvalue a , there is only one linearly independent eigenvector $|a\rangle$. Now, as we mentioned earlier, the system's initial state $|\psi\rangle$ collapses to the eigenstate $|a\rangle$ if the eigenvalue a is obtained in the measurement of A . In general, the observable B does not have a definite value in the state $|a\rangle$.

An exception occurs if $|a\rangle$ is also an eigenstate of \hat{B} with some eigenvalue b , i.e.,

$$\hat{B}|a\rangle = b|a\rangle. \quad (8.21)$$

Since $|a\rangle$ is an eigenvector of both \hat{A} and \hat{B} with eigenvalues a and b , respectively, it is more expressive to label the eigenket by both the eigenvalues, i.e.,

$$|a\rangle \equiv |a, b\rangle.$$

Hence

$$\hat{A}|a,b\rangle = a|a,b\rangle \quad (8.22)$$

and

$$\hat{B}|a,b\rangle = b|a,b\rangle. \quad (8.23)$$

Now, the measurement of B immediately after A , is certain to yield b and the state of the system would remain unaltered due to the B -measurement. The change of the state of the system in the two measurements is shown below:

$$|\psi\rangle \xrightarrow{\text{Measure } \hat{A}} |a,b\rangle \xrightarrow{\text{Measure } \hat{B}} |a,b\rangle.$$

There is no change of the state of the system due to the measurement of B because the system was already in an eigenstate of \hat{B} prior to the measurement. Since the state of the system has collapsed to a simultaneous eigenstate of both \hat{A} and \hat{B} after the measurement of A , the system is now in a state where both the observables have definite values, namely, a and b . Further, since the measurement of B does not alter the state, a second measurement of A is certain to yield the previous value a .

If the scenario just described holds in all situations, no matter what is the outcome of the first measurement of A , we say that the observables A and B are compatible. Thus, in summary, if we perform the following sequence of measurements in rapid succession on a system:

1. measure A
2. measure B
3. remeasure A

then, if the result of 3 is certain to be the same as the result of 1, we say that A and B are compatible variables. The condition for compatibility in the nondegenerate case is that every eigenvector of \hat{A} is also an eigenvector of \hat{B} so that the common eigenvectors form a basis of the Hilbert space. Therefore, if observables A and B are compatible, it is always possible to find states of the system, namely the simultaneous eigenstates of \hat{A} and \hat{B} , in which both the observables have definite values.

Note that in our discussions we have assumed the eigenvalues of \hat{A} , the first observable to be measured, are nondegenerate and nothing has been assumed about the degeneracy of the eigenvalues of second observable \hat{B} . The arguments above would be equally valid if we assumed that the eigenvalues of \hat{B} are nondegenerate and no assumptions were made about the degeneracy of the eigenvalues of \hat{A} .

Next, if A and B are compatible, the probability of getting the results a and b in the sequence of measurements: A followed by B , would be

$$P(a, b) = P_{|\psi\rangle}(a)P_{|a,b\rangle}(b) = |\langle a, b | \psi \rangle|^2 \times 1 = |\langle a, b | \psi \rangle|^2. \quad (8.24)$$

If we reverse the order of measurements and assume that the eigenvalues of \hat{B} are also nondegenerate like those of \hat{A} , then $P(b, a)$ would be the same as $P(a, b)$. This result is general for two compatible variables and is true even when the eigenvalues of \hat{A} and \hat{B} are degenerate. This will be discussed in the next section. For non-compatible observables $P(a, b)$ would not be equal to $P(b, a)$.

8.3.2 Degenerate Case

Let us next assume that the eigenvalues of \hat{A} are degenerate. We make no assumption about the degeneracy of the eigenvalues of \hat{B} . As before, we measure A and B in rapid succession, in the order A then B , on a system in the state $|\psi\rangle$. If a particular eigenvalue a of the operator \hat{A} is obtained in the measurement, the state of the system collapses to the normalized projection of $|\psi\rangle$ onto the eigensubspace H_a i.e.,

$$|\psi\rangle \xrightarrow{\text{Measure } A} |\psi'\rangle = \frac{\hat{P}_a |\psi\rangle}{\sqrt{\langle \psi | \hat{P}_a | \psi \rangle}} \in H_a. \quad (8.25)$$

The eigensubspace H_a is g_a -dimensional. The basis vectors of H_a could be chosen as the g_a linearly independent eigenvectors of \hat{A} with the same eigenvalue a , i.e. $\{|a, i\rangle, i = 1, 2, \dots, g_a\}$. Any vector in H_a (infinitely many of them) is an eigenvector of \hat{A} with eigenvalue a , but there are only g_a linearly independent vectors in H_a .

In general, no vector lying wholly in H_a will also be an eigenvector of \hat{B} . So, a subsequent measurement of B yielding some eigenvalue b will throw the system into the eigenstate $|b\rangle$, and this state lying in the eigensubspace H_b , will have components both in H_a and H_a^\perp , where H_a^\perp , called the orthogonal complement of H_a , is the subspace orthogonal to H_a . The subspace H_a^\perp consists of (is the union of) the eigensubspaces of \hat{A} other than H_a . So, a second measurement of A will not yield a , the value obtained in the first measurement, with certainty, i.e., A and B would be incompatible.

However, in exceptional cases it may be possible to find g_a linearly independent eigenvectors of the operator \hat{B} with distinct eigenvalues b_1, b_2, \dots in each eigensubspace H_a and, being linearly independent, the eigenvectors can span the eigensubspace H_a . These eigenvectors of \hat{B} , since they

lie wholly in H_a , are also eigenvectors of \hat{A} with the same eigenvalue a . We may denote these simultaneous eigenvectors of \hat{A} and \hat{B} in H_a as follows:

$$|a, b_i, \alpha_i\rangle, \alpha_i = 1, 2, \dots, g_{b_i}^{(a)},$$

where $g_{b_i}^{(a)}$ represents the order of degeneracy of the eigenvalue b_i in the eigensubspace H_a . The order of degeneracy of b_i in the entire Hilbert space may be greater, for some of the linearly independent eigenvectors of \hat{B} with the same degenerate eigenvalue b_i may lie in eigensubspaces other than H_a .

Thus, in the special case we are considering, we can find simultaneous eigenvectors of both \hat{A} and \hat{B} which span each eigensubspace and so the whole Hilbert space can also be spanned by simultaneous eigenvectors of the two operators. All eigenvectors of \hat{B} , lie wholly in one or the other of the eigensubspaces H_a of the operator \hat{A} , i.e., no eigenvector of \hat{B} has components in more than one eigensubspace H_a .

Now, the state $|\psi'\rangle \in H_a$ immediately after the measurement of A is the normalized projection of the state $|\psi\rangle$ prior to the measurement on to H_a and so the projected vector can be expressed as a linear combination of the simultaneous eigenstates of \hat{A} and \hat{B} which span H_a . Therefore, a subsequent measurement of B would give one of the eigenvalues b_i of the eigenvectors $\{|a, b_i, \alpha_i\rangle\}$ which span H_a . After the B -measurement, the state changes to

$$\begin{aligned} |\psi'\rangle &\rightarrow |\psi''\rangle \\ &= \frac{\hat{P}_{b_i}|\psi'\rangle}{\sqrt{\langle\hat{P}_{b_i}\psi'|\hat{P}_{b_i}\psi'\rangle}} \\ &= \frac{\sum_{\alpha_i=1}^{g_{b_i}^{(a)}} |a, b_i, \alpha_i\rangle\langle a, b_i, \alpha_i|\psi'\rangle}{\sqrt{\langle\psi'|\hat{P}_{b_i}|\psi'\rangle}} \end{aligned} \quad (8.26)$$

In the above formula \hat{P}_{b_i} is the projection operator onto the eigensubspace of b_i i.e., onto H_{b_i} . If b_i is degenerate, then some of the linearly independent eigenvectors with the same eigenvalue b_i may lie in H_a and some may lie in the eigensubspace H_a^\perp orthogonal to H_a and, in the special case under consideration, no eigenvector lies partly in H_a and partly in H_a^\perp . We can therefore write \hat{P}_{b_i} as

$$\begin{aligned} \hat{P}_{b_i} &= \sum_{a', \alpha_i} |a', b_i, \alpha_i\rangle\langle a', b_i, \alpha_i| \\ &= \sum_{\alpha_i} |a, b_i, \alpha_i\rangle\langle a, b_i, \alpha_i| + \sum_{a' \neq a, \alpha_i} |a', b_i, \alpha_i\rangle\langle a', b_i, \alpha_i| \end{aligned} \quad (8.27)$$

where α_i is any additional index required to label the states uniquely. The first term on the right side of the above equation can be considered as the restriction of \hat{P}_{b_i} in H_a and the second term the restriction in H_a^\perp .

Now, the ket the projection operator \hat{P}_{b_i} acts on, i.e., $|\psi'\rangle$, is in H_a and hence cannot be projected onto H_a^\perp . In other words, the second term of Eq. (8.27) acting on $|\psi'\rangle$ gives zero while the first term projects $|\psi'\rangle$ on to that part of H_a which is also a part of H_{b_i} . Therefore, \hat{P}_{b_i} is really the projection of $|\psi'\rangle$ onto the intersection of the eigensubspaces H_a and H_{b_i} . So $|\psi''\rangle$ is in H_a like $|\psi'\rangle$.

We note that the measurement of B has changed the state from $|\psi'\rangle$ to $|\psi''\rangle$, unlike in the nondegenerate case, but the changed state is still in H_a . So, a second measurement of A would certainly give the same eigenvalue a as was obtained in the first measurement of the observable. Therefore, the observables A and B would be compatible.

To summarize, if in every eigensubspace H_a of the eigenvalues of \hat{A} , it is possible to find g_a linearly independent eigenvectors of the operator \hat{B} , then the totality of all the simultaneous eigenvectors in the entire Hilbert space form a complete basis set of vectors. The two observables would then be compatible. We could equally well have said: if in every eigensubspace H_b of the eigenvalues of \hat{B} , we can find g_b linearly independent eigenvectors which are also eigenvectors of \hat{A} , then the two observables A and B are compatible.

This is not to say that any eigenvector of \hat{A} is also an eigenvector of \hat{B} . For example, if we take any vector in H_a , the vector is guaranteed to be an eigenvector of \hat{A} , but not necessarily an eigenvector of \hat{B} . To see this let us consider the vector

$$|\psi\rangle = c_1|a, b_1\rangle + c_2|a, b_2\rangle \in H_a. \quad (8.28)$$

This is a vector in H_a and therefore an eigenvector of \hat{A} with eigenvalue a , but not an eigenvector of \hat{B} . There are infinity of different vectors in H_a but only g_a linearly independent ones which can act as basis for H_a . These linearly independent vectors, though eigenvectors of \hat{A} with eigenvalue a , may not in general be eigenvectors of \hat{B} also. However, by taking appropriate linear combinations of these linearly independent eigenvectors if it is possible to get another set of g_a linearly independent vectors which are also eigenvectors of \hat{B} , then the two observables are compatible. Therefore, for two observables to be compatible, they must have simultaneous eigenvectors which form a complete set of basis states for the Hilbert space.

8.4 Condition for Compatibility of Observables

In the previous section we have stated what we mean when we say two observables are compatible with each other. To recapitulate, if two observables have simultaneous eigenvectors that span the entire Hilbert space, i.e., if the simultaneous eigenvectors form a basis for the Hilbert space, then the observables are said to be compatible.

We will now prove that two observables A and B with corresponding Hermitian operators \hat{A} and \hat{B} , respectively, are compatible if and only if $[\hat{A}, \hat{B}] = 0$. Thus we prove the following theorem:

Theorem:

Two Hermitian operators representing two observables have a complete set of simultaneous eigenvectors if and only if they commute.

An alternative statement of the theorem could be: The necessary and sufficient condition that two observables are compatible is that their operators commute.

Proof:

The proof proceeds in two parts. First, we prove the necessary condition, i.e., we assume that the operators have simultaneous eigenvectors, then we show that the operators commute. Next, we prove the converse (the sufficiency condition), i.e., we assume that the operators commute, then we show they have simultaneous eigenvectors.

(a) Necessary Condition

Let us assume that \hat{A} and \hat{B} have a complete set of simultaneous eigenvectors $|u_n\rangle, n = 1, 2, \dots$. Here each u_n represents a unique set of numbers $\{a, b, i\}$, where a and b are eigenvalues of \hat{A} and \hat{B} , respectively, and i is any other parameter which would be required to label the states uniquely should there be more than one linearly independent eigenvectors with the same values for a and b . Since the complete set of vectors $\{|u_n\rangle, n = 1, 2, \dots\}$ are eigenvectors of both \hat{A} and \hat{B} , we have

$$\hat{A}|u_n\rangle = a_n|u_n\rangle, \quad (8.29)$$

and

$$\hat{B}|u_n\rangle = b_n|u_n\rangle. \quad (8.30)$$

Now, using Eq. (8.73) and (8.74), we can write

$$\begin{aligned} [\hat{A}, \hat{B}]|u_n\rangle &= (\hat{A}\hat{B} - \hat{B}\hat{A})|u_n\rangle \\ &= (a_n b_n - b_n a_n)|u_n\rangle \\ &= 0. \end{aligned} \quad (8.31)$$

Since the simultaneous eigenvectors $\{|u_n\rangle, i = 1, 2, \dots\}$ form a complete set, it follows that

$$[\hat{A}, \hat{B}]|\psi\rangle = 0, \quad (8.32)$$

where $|\psi\rangle$ is an arbitrary vector in the Hilbert space. Hence we have

$$[\hat{A}, \hat{B}] = 0. \quad (8.33)$$

Thus, we have proved that commutivity is a necessary condition for compatibility.

(b) Sufficiency Condition

We now prove the converse, i.e., if \hat{A} and \hat{B} commute, they have simultaneous eigenvectors. First, we will assume that all eigenvalues of \hat{A} are non-degenerate. Next, we will consider the more general situation, namely that, some or all of the eigenvalues of \hat{A} may be degenerate.

Non-degenerate case

If all eigenvalues of \hat{A} are non-degenerate, then it is possible to uniquely label the eigenstates of \hat{A} by the eigenvalues only. Thus let $|a\rangle$ be the eigenstate of the operator \hat{A} belonging to some eigenvalue a . Therefore,

$$\hat{A}|a\rangle = a|a\rangle. \quad (8.34)$$

We start the proof by applying the operator \hat{B} to Eq. (8.78) to get

$$\hat{B}\hat{A}|a\rangle = a\hat{B}|a\rangle. \quad (8.35)$$

Since our assumption is that \hat{A} and \hat{B} commute, we can interchange the order of \hat{A} and \hat{B} on the left hand side of the above equation getting

$$\hat{A}(\hat{B}|a\rangle) = a(\hat{B}|a\rangle). \quad (8.36)$$

From Eq. (8.80) we can conclude that $\hat{B}|a\rangle$ is also an eigenvector of \hat{A} with the same eigenvalue a . Since we have supposed the eigenvalues of \hat{A} are non-degenerate, the vectors $|a\rangle$ and $\hat{B}|a\rangle$ represent the same physical state, i.e., $\hat{B}|a\rangle$ differs from $|a\rangle$ by a constant multiplier which we denote by b . Therefore,

$$\hat{B}|a\rangle = b|a\rangle, \quad (8.37)$$

i.e., all eigenvectors $\{|a\rangle\}$ of \hat{A} are also eigenvectors of \hat{B} . Therefore, we may label these states by both eigenvalues a and b , rather than the single eigenvalue a , i.e., $|a\rangle \equiv |a, b\rangle$.

Degenerate Case

Now, we allow for the more general case that some or all eigenvalues of \hat{A} may be degenerate. In case of degeneracy, the eigenvalue equation for a particular eigenvalue a is written as

$$\hat{A}|a,i\rangle = a|a,i\rangle; i = 1, 2, \dots, g_a, \quad (8.38)$$

where g_a is the order of degeneracy of a . Here a is an element of the set of all the eigenvalues of \hat{A} , i.e., $a \in \{a_1, a_2, \dots\}$. The g_a linearly independent eigenvectors $\{|a,i\rangle; i = 1, 2, \dots, g_a\}$ spanning the eigensubspace H_a are made orthonormal.

If we choose the eigenvectors of \hat{A} belonging to all the eigensubspaces, i.e., the set of eigenvectors $\{|a,i\rangle, i = 1, 2, \dots, g_a; a = a_1, a_2, \dots\}$, as the basis for the Hilbert space, then obviously, the matrix representation of \hat{A} is diagonal. Denoting the matrix representing the operator \hat{A} by \underline{A} , we have

\hat{A}	H_{a_1}	H_{a_2}	H_{a_3}
H_{a_1}	$a_1 \cdot 0$ $0 \cdot a_1$	0	0
H_{a_2}	0	$a_2 \cdot 0$ $0 \cdot a_2$	0
H_{a_3}	0	0	$a_3 \cdot 0$ $0 \cdot a_3$

Figure 8.1: The matrix representation of \hat{A} in the eigenbasis of \hat{A} .

where each diagonal block H_{a_n} is a $g_{a_n} \times g_{a_n}$ dimensional diagonal matrix with the eigenvalue a_n running along the diagonal, other entries being zero.

Next, we ask what would be the matrix representation of \hat{B} in the eigenbasis of \hat{A} . The matrix elements of \hat{B} are written as $\langle a', i' | \hat{B} | a, i \rangle$. Using our assumption that \hat{A} and \hat{B} commute, we have

$$\langle a', i' | [\hat{A}, \hat{B}] | a, i \rangle = 0,$$

i.e.,

$$\langle a', i' | \hat{A}\hat{B} - \hat{B}\hat{A} | a, i \rangle = 0,$$

or,

$$(a' - a) \langle a', i' | \hat{B} | a, i \rangle = 0. \quad (8.39)$$

If $a \neq a'$, then $(a' - a) \neq 0$ and we must have

$$\langle a', i' | \hat{B} | a, i \rangle = 0, \quad \text{if } a' \neq a, \quad (8.40)$$

i.e., \hat{B} does not connect states with different eigenvalues of \hat{A} . In other words, \hat{B} acting on any of the eigenvectors of \hat{A} in the eigensubspace H_a , produces another vector which is also in H_a with no components in other eigensubspaces orthogonal to H_a . The new vector, being in H_a , remains an eigenvector of \hat{A} . Thus, \hat{B} acting on any eigenvector of \hat{A} produces a state which is also an eigenvector of \hat{A} with the same eigenvalue. This is a consequence of the fact that \hat{A} and \hat{B} commute.

From the above arguments, we can conclude that the matrix representation of \hat{B} in the eigenbasis of \hat{A} will be block diagonal as shown below.

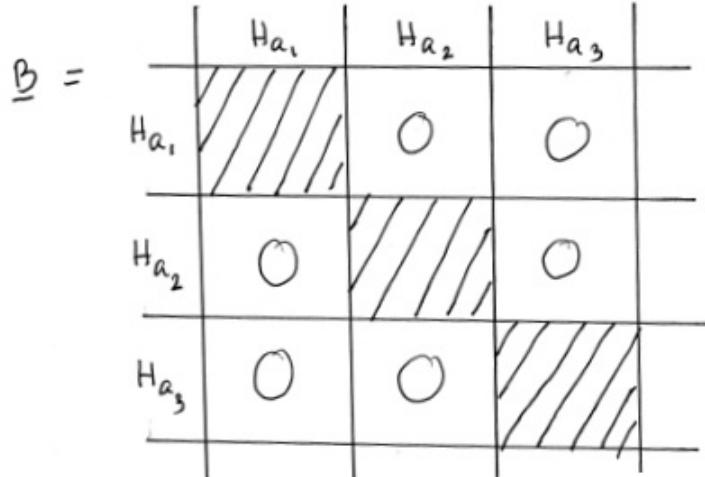


Figure 8.2: The matrix representation of \hat{B} in the eigenbasis of \hat{A} .

In each eigensubspace H_a , the matrix $\underline{B}^{(a)}$ is a $g_a \times g_a$ square matrix, which itself need not be diagonal if we choose an arbitrary basis for H_a . To see this we refer to Eq. (8.84). This equation tells us that, if \hat{A} and \hat{B} commute, the matrix of elements of B are zero if $a \neq a'$, but nothing is concluded about the matrix elements of \hat{B} in any eigensubspace H_a , i.e., when $a = a'$.

However, since \hat{B} is a Hermitian operator, its matrix representation in every eigensubspace H_a is a finite dimensional square Hermitian matrix. We can always diagonalize a finite dimensional Hermitian matrix by a change of basis. The new set of g_a basis vectors in the eigensubspace H_a are linear combinations of the old basis vectors $\{|a, i\rangle, i = 1, 2, \dots, g_a\}$. Therefore, the vectors of the new basis set remain eigenvectors of \hat{A} with the same eigenvalue a . But, the vectors in new basis set, since it diagonalizes \hat{B} , must also be eigenvectors of \hat{B} with eigenvalues b_1, b_2, \dots, b_{g_a} .

Since the vectors in the new basis set are simultaneous eigenvectors of \hat{A} and \hat{B} , we can label them by the eigenvalues of both \hat{A} and \hat{B} . Thus the common eigenvectors, which form the new basis set of H_a , can be written as $\{|a, b_1\rangle, |a, b_2\rangle, \dots, |a, b_{g_a}\rangle\}$. Some of the eigenvalues of \hat{B} may be repeated, in which case we will need another index to distinguish between the linearly independent simultaneous eigenvectors with the same eigenvalues for \hat{A} and \hat{B} . We diagonalize each diagonal block $\underline{B}^{(a_i)}$ of the matrix \underline{B} in the old basis, thereby obtaining simultaneous eigenvectors for both \hat{A} and \hat{B} in the entire

Hilbert space. In the new basis consisting of the simultaneous eigenvectors, the matrix representation of \hat{B} is diagonal as shown in figure (8.6).

\hat{B}	H_{a_1}	H_{a_2}	H_{a_3}	...
H_{a_1}	$b_1^{(1)} \ b_2^{(1)}$ 0 \ddots $b_g^{(1)}$	0	0	
H_{a_2}	0	$b_1^{(2)} \ b_2^{(2)}$ 0 \ddots $b_g^{(2)}$	0	
H_{a_3}	0	0	$b_1^{(3)} \ b_2^{(3)}$ 0 \ddots $b_g^{(3)}$	

Figure 8.3: The matrix representation of \hat{B} in the eigenbasis consisting of simultaneous eigenvectors of \hat{A} and \hat{B} .

The matrix representation of \hat{A} in the new basis is also diagonal and remains the same as in Fig. (8.4).

The diagonal entries $b_i^{(n)}, i = 1, 2, \dots, g_n$ in each eigensubspace H_{a_n} in figure (8.6) are the various eigenvalues of the operator \hat{B} , i.e., $b_i^{(n)} \in \{b_1, b_2, \dots\}$. In a given eigensubspace H_{a_n} , some of the eigenvalues of \hat{B} may be repeated. Further, a particular eigenvalue of \hat{B} may also be repeated in several eigensubspaces H_{a_n} .

Thus, in summary, what we have shown is that, given two Hermitian operators \hat{A} and \hat{B} representing two observables of a quantum system, it is possible to construct a complete set of simultaneous eigenvectors spanning the entire Hilbert space provided $[\hat{A}, \hat{B}] = 0$. Hence, we have proved that $[\hat{A}, \hat{B}] = 0$ is a sufficient condition for two Hermitian operators \hat{A} and \hat{B} to be compatible.

8.5 Labeling of Quantum Mechanical Basis States

Since eigenvectors of a Hermitian operator corresponding to an observable form a complete set, we can label the members of the set by the eigenvalues. Thus, suppose that an observable \hat{A} has

eigenvalues

$$a_1, a_2, a_3, \dots$$

The basis states are then labeled as

$$|a_1\rangle, |a_2\rangle, |a_3\rangle, \dots$$

The labeling would be unambiguous if the eigenvalues were all non-degenerate. However, if some or all eigenvalues are degenerate, we will need another mark of distinction for the eigenvectors.

For example, if an eigenvalue a_n is g_n -fold degenerate, the corresponding eigenvectors may be denoted as

$$|a_n, i\rangle, i = 1, 2, \dots, g_n.$$

An alternative, and better, notation is based on the fact that two compatible observables whose Hermitian operators commute, may be assumed to have the same set of eigenvectors. Hence, if we can find a second observable \hat{B} commuting with the first, such that

$$\hat{B}|a_n, i\rangle = b_i|a_n, i\rangle, i = 1, 2, \dots, g_n \quad (8.41)$$

with eigenvalues b_i all different, then the eigenvalues of \hat{B} may serve to distinguish the eigenvectors. So, we can write

$$|a_n, 1\rangle \equiv |a_n, b_1\rangle$$

$$|a_n, 2\rangle \equiv |a_n, b_2\rangle$$

.

.

.

$$|a_n, g_n\rangle \equiv |a_n, b_{g_n}\rangle.$$

But, if some of the eigenvalues $b_i, i = 1, 2, \dots, g_n$ are equal, we will need a third mark to distinguish the eigenvectors. The third mark may be obtained if we can find a third observable whose operator \hat{C} commutes with both \hat{A} and \hat{B} . Then \hat{A} , \hat{B} and \hat{C} have simultaneous eigenvectors and the eigenvalues of \hat{C} may also be used to label the eigenvectors.

Thus, in general, if we can find a set of mutually commuting Hermitian operators, $\hat{A}, \hat{B}, \hat{C}, \hat{D}, \dots$ whose common eigenvectors can be characterized completely by the eigenvalues a, b, c, d, \dots such that no two eigenvectors have exactly identical set of eigenvalues, then the eigenvalues of these operators can uniquely label the common eigenvectors. Such a set of Hermitian operators is said to be complete. We refer to this set of operators as a complete set of commuting observables (CSCO).

In the new notation, the basis vectors are written as

$$|a, b, c, d, \dots\rangle.$$

The normalization and completeness conditions are then

$$\langle a', b', c', d' \dots | a, b, c, d, \dots \rangle = \delta_{aa'} \delta_{bb'} \dots, \quad (8.42)$$

and

$$\sum_{(a,b,c,d,\dots)} |a, b, c, d, \dots\rangle \langle a, b, c, d, \dots| = \hat{I}. \quad (8.43)$$

Lecture 8 **Compatibility of Observable: version 2**

8.6 Background Materials

Suppose that \hat{A} is the Hermitian operator representing an observable A of a quantum system. The eigenkets of \hat{A} form a complete set of orthonormal basis states. Let $|\psi\rangle$ be the state vector of the system. We can expand $|\psi\rangle$ in the eigenbasis of the operator \hat{A} :

$$|\psi\rangle = \sum_a |a\rangle \langle a| \psi\rangle, \quad (8.44)$$

where a runs over all the eigenvalues of \hat{A} , i.e., $a \in \text{spec}(\hat{A})$, where $\text{spec}(\hat{A}) = (a_1, a_2, a_3, \dots)$ is the set of all eigenvalues of \hat{A} . This set is called the eigenvalue spectrum of \hat{A} .

In Eq. (8.44), we have assumed that each eigenvalue a is non-degenerate, i.e., for each a there exists only one linearly independent eigenvector $|a\rangle$. Therefore, the eigenvalue itself can be used to label the corresponding eigenket unambiguously.

However, it may so happen that some or all of the eigenvalues of \hat{A} are degenerate, i.e., there may be more than one linearly independent eigenvector corresponding to the same eigenvalue. The number of linearly independent eigenvectors corresponding to a particular eigenvalue a is called the order of degeneracy of the eigenvalue and is denoted by g_a . If the eigenvalue a is degenerate, then just the eigenvalue itself is not enough to label the eigenstates uniquely. We need another index to distinguish between the g_a linearly independent eigenvectors all with the same eigenvalue a .

Thus the eigenvectors belonging to a degenerate eigenvalue a may be denoted by $|a, i\rangle$, where the index i can take discrete values $i = 1, 2, 3, \dots, g_a$. The index i may be called the degeneracy index. Later, we will see that we can improve the notation by replacing i by the eigenvalues of other Hermitian operators \hat{B}, \hat{C}, \dots , which commute with \hat{A} .

The set of linearly independent eigenvectors $\{|a, i\rangle, i = 1, 2, \dots, g_a\}$, all with the same eigenvalue a , span a g_a -dimensional subspace of the Hilbert space. This subspace is called the eigensubspace of a , and is denoted by H_a . The union of all the eigensubspaces of the operator \hat{A} constitutes the full Hilbert space.

It is easy to see that any linear combination of the eigenvectors $\{|a, i\rangle, i = 1, 2, \dots, g_a\}$ is also an eigenvector of \hat{A} with the same eigenvalue a . Thus

$$\hat{A} \left(\sum_{i=1}^{g_a} c_i |a, i\rangle \right) = a \left(\sum_{i=1}^{g_a} c_i |a, i\rangle \right), \quad (8.45)$$

where c_i 's are constants. The set of eigenvectors $\{|a, i\rangle, i = 1, 2, \dots, g_a\}$, even though linearly independent, may not be orthogonal to each other because they belong to the same eigenvalue. However, following the Schmidt orthonormalization procedure, we can take linear combinations of the above set of vectors in a special way and obtain a set of g_a orthonormal vectors. The new set of orthonormal (and hence also linearly independent) vectors remain eigenvectors of \hat{A} with the same eigenvalue a .

We assume that the Schmidt procedure has been carried out in each eigensubspace. So, in each eigensubspace, we can write

$$\langle a, i | a, j \rangle = \delta_{ij}, \quad a \in \text{Spec}(\hat{A}) \quad (8.46)$$

where $i, j = 1, 2, 3, \dots, g_a$. The orthonormal set of eigenvectors $\{|a, i\rangle, i = 1, 2, \dots, g_a\}$ can be said to span the eigensubspace H_a . The eigenvectors in different eigensubspaces are automatically orthogonal because \hat{A} is Hermitian. Hence, we can assume that **all** eigenvectors of \hat{A} are orthonormal, i.e.,

$$\langle a', i' | a, i \rangle = \delta_{aa'} \delta_{ii'}. \quad (8.47)$$

The full set $\{|a, i\rangle, i = 1, 2, \dots, g_a; a \in \text{spec}(\hat{A})\}$ of eigenvectors of \hat{A} form a complete orthonormal set, i.e., they constitute a basis set for the Hilbert space. The completeness condition of this basis set

can be written as

$$\sum_{a \in \text{spec}(\hat{A})} \sum_{i=1}^{g_a} |a, i\rangle \langle a, i| = \hat{I}. \quad (8.48)$$

Thus the state vector $|\psi\rangle$ of the system can be expanded as

$$\begin{aligned} |\psi\rangle &= \hat{I}|\psi\rangle \\ &= \sum_{a \in \text{spec}(\hat{A})} \sum_{i=1}^{g_a} |a, i\rangle \langle a, i| \psi\rangle. \end{aligned} \quad (8.49)$$

For simplicity, we will continue to use Eq. (8.44) as the expansion for $|\psi\rangle$ in the eigenbasis of \hat{A} assuming that all the eigenvalues are non-degenerate. In case of degeneracy, notations can be generalized in a straightforward manner, as discussed above.

8.7 Measurement of Two Observables in Quick Succession

Suppose that a quantum system is in the state $|\psi\rangle$. Let A be an observable of the system with corresponding Hermitian operator \hat{A} . We can always expand $|\psi\rangle$ using the eigenbasis of \hat{A} :

$$|\psi\rangle = \sum_a |a\rangle \langle a| \psi\rangle, \quad (8.50)$$

where a runs over the eigenvalue spectrum of \hat{A} . The complex number $\langle a|\psi\rangle$ is the ‘component’ of $|\psi\rangle$ along $|a\rangle$.

We now make a measurement of the observable A on the system in the state $|\psi\rangle$. Since a general state $|\psi\rangle$ may be a superposition of many (perhaps infinitely many) eigenkets of \hat{A} with different eigenvalues, we cannot exactly predict the result of the experiment but can only say that the experiment would yield one of the eigenvalues of \hat{A} for which $\langle a|\psi\rangle \neq 0$.

The outcome is random, i.e., probabilistic, because we cannot predict exactly which eigenvalue will be obtained, but we can assign a probability for obtaining a particular eigenvalue a provided we know the state $|\psi\rangle$ before the measurement. The probability is

$$P_{|\psi\rangle}(a) = |\langle a|\psi\rangle|^2. \quad (8.51)$$

In the course of the measurement, the state of the system collapses to the eigenket $|a\rangle$ if the eigenvalue a is obtained in the measurement. Thus

$$|\psi\rangle \xrightarrow{\text{measurement of } A} |a\rangle. \quad (8.52)$$

Now the system is in a state with a definite value for the observable A . Further successive measurements of A will yield the same eigenvalue a with 100% certainty.

Next, consider another observable B of the system with corresponding Hermitian operator \hat{B} . If, immediately after the measurement of A , we measure B , are we certain to get a particular eigenvalue b of the operator \hat{B} ? We know that the state of the system after the measurement of A is an eigenstate of \hat{A} , i.e., in a state with a definite value a of the observable A . The question we have asked can be restated as follows: in the collapsed state $|a\rangle$, does B have a definite value, i.e., is $|a\rangle$ also an eigenstate of \hat{B} with some eigenvalue b ? The answer to this question is, in general, in the negative, i.e., $|a\rangle$ is not an eigenket of \hat{B} in general. The very important exceptional case where $|a\rangle$ is also an eigenstate of \hat{B} is discussed later.

Since the eigenkets of \hat{B} form a basis set, we can expand $|a\rangle$ as

$$|a\rangle = \sum_b |b\rangle \langle b|a\rangle. \quad (8.53)$$

Therefore, a measurement of B would yield any one of the eigenvalues $b \in \text{spec}(\hat{B})$ for which $\langle b|a\rangle \neq 0$. The probability of obtaining a particular eigenvalue b while the system is in the state $|a\rangle$ is

$$P_{|a\rangle}(b) = |\langle b|a\rangle|^2. \quad (8.54)$$

The state of the system after B is measured collapses from $|a\rangle$ to $|b\rangle$:

$$|a\rangle \xrightarrow{\text{measurement of } B} |b\rangle. \quad (8.55)$$

Thus, the measurement of B generally alters the state of the system just before the measurement because, due to the measurement process, the system is thrown into the eigenstate $|b\rangle$, where b is the eigenvalue of \hat{B} obtained in the measurement. But, since $|b\rangle$ is not, in general, an eigenstate of \hat{A} , the system is no longer in a state with a definite value for the observable A .

If we measure A again immediately after B is measured, we will not get the answer we got the first time, namely a , with certainty. In the second measurement of A there is the possibility of getting any eigenvalue for which $|\langle a|b\rangle| \neq 0$. To repeat, after the first measurement of A , the system is in the eigenstate $|a\rangle$ of \hat{A} but this state is not an eigenstate of the operator \hat{B} except in a special situation discussed below. After the measurement of B , the system is thrown in the eigenstate $|b\rangle$ of \hat{B} which is not an eigenstate of \hat{A} . Therefore, it is not possible, in general, to get the system in a state in which both the observables A and B have definite values.

Now, in successive measurements of A and B , first A then B , on a system initially in the state $|\psi\rangle$, the probability of obtaining the result a and b is

$$P(a, b) = P_{|\psi\rangle}(a)P_{|a\rangle}(b) = |\langle a|\psi\rangle|^2 |\langle b|a\rangle|^2. \quad (8.56)$$

If we reversed the order of measurements, first B and then A , the probability for obtaining the results b for the observable B and a for the observable A would be

$$P(b, a) = P_{|\psi\rangle}(b)P_{|b\rangle}(a) = |\langle b|\psi\rangle|^2 |\langle a|b\rangle|^2. \quad (8.57)$$

We note that, since $|\langle a|\psi\rangle|^2 \neq |\langle b|\psi\rangle|^2$, the two probabilities are not equal in general, i.e., $P(a, b) \neq P(b, a)$.

8.8 Compatible Observables

Now, let us return to the exceptional situation mentioned in the previous section. To recapitulate, let us make two measurements of two observables A and B in quick succession on a system in the state $|\psi\rangle$. First, we measure A , and if the eigenvalue a is obtained, the system's state collapses to an eigenstate of the operator \hat{A} belonging to the eigenvalue a . Then if we measure B immediately afterward, would we be certain to get a particular eigenvalue b of \hat{B} ? Further, if a second measurement of A is made immediately after the measurement of B , what conditions need be fulfilled in order that the result of the first measurement is unaltered, i.e., we will be certain to get the same result a again? We will first consider the case when all eigenvalues of \hat{A} and \hat{B} are non-degenerate and then we will consider degeneracy of the eigenvalues.

8.8.1 Nondegenerate Case

Suppose all eigenvalues of \hat{A} and \hat{B} are nondegenerate. First we measure A on a system which is initially in the state $|\psi\rangle$. The system's initial state $|\psi\rangle$ collapses to the eigenstate $|a\rangle$ if the eigenvalue a is obtained in the measurement of A . In general, the observable B does not have a definite value in the state $|a\rangle$.

An exception occurs if $|a\rangle$ is also an eigenstate of \hat{B} with some eigenvalue b , i.e.,

$$\hat{B}|a\rangle = b|a\rangle. \quad (8.58)$$

Since $|a\rangle$ is an eigenvector of both \hat{A} and \hat{B} with eigenvalues a and b , respectively, it is more expressive to label the eigenket by both the eigenvalues, i.e.,

$$|a\rangle \equiv |a, b\rangle.$$

Hence

$$\hat{A}|a,b\rangle = a|a,b\rangle \quad (8.59)$$

and

$$\hat{B}|a,b\rangle = b|a,b\rangle. \quad (8.60)$$

Now, the measurement of B immediately after A , is certain to yield b and the state of the system would remain unaltered due to the B -measurement. The change of the state of the system in the two measurements is shown below:

$$|\psi\rangle \xrightarrow{\text{Measure } \hat{A}} |a,b\rangle \xrightarrow{\text{Measure } \hat{B}} |a,b\rangle.$$

There is no change of the state of the system due to the measurement of B because the system was already in an eigenstate of \hat{B} prior to the measurement. Since the state of the system has collapsed to a simultaneous eigenstate of both \hat{A} and \hat{B} after the measurement of A , the system is now in a state where both the observables have definite values, namely, a and b . Further, since the measurement of B does not alter the state, a second measurement of A is certain to yield the previous value a .

If the scenario just described holds in all situations, no matter what is the outcome of the first measurement of A , we say that the observables A and B are compatible. Thus, in summary, if we perform the following sequence of measurements in rapid succession on a system:

1. measure A
2. measure B
3. remeasure A

then, if the result of 3 is certain to be the same as the result of 1, we say that A and B are compatible variables. The condition for compatibility in the nondegenerate case is that every eigenvector of \hat{A} is also an eigenvector of \hat{B} so that the common eigenvectors form a basis of the Hilbert space. Therefore, if observables A and B are compatible, it is always possible to find states of the system, namely the simultaneous eigenstates of \hat{A} and \hat{B} , in which both the observables have definite values.

Next, if A and B are compatible, the probability of getting the results a and b in the sequence of measurements: A followed by B , would be

$$P(a,b) = P_{|\psi\rangle}(a)P_{|a,b\rangle}(b) = |\langle a,b|\psi\rangle|^2 \times 1 = |\langle a,b|\psi\rangle|^2. \quad (8.61)$$

If we reverse the order of measurements and assume that the eigenvalues of \hat{B} are also nondegenerate like those of \hat{A} , then $P(b,a)$ would be the same as $P(a,b)$. For non-compatible observables $P(a,b)$ would not be equal to $P(b,a)$.

8.8.2 Degenerate Case

Let us next consider the general case where the eigenvalues of both \hat{A} and \hat{B} may be degenerate. As before, we measure A and B in rapid succession, in the order A then B , on a system in the state $|\psi\rangle$. If a particular eigenvalue a of the operator \hat{A} is obtained in the measurement, the state of the system collapses to the normalized projection of $|\psi\rangle$ onto the eigensubspace H_a i.e.,

$$|\psi\rangle \xrightarrow{\text{Measure } A} |\psi'\rangle = \frac{\hat{P}_a |\psi\rangle}{\sqrt{\langle \psi | \hat{P}_a | \psi \rangle}} \in H_a. \quad (8.62)$$

The eigensubspace H_a is g_a -dimensional. The basis vectors of H_a could be chosen as the g_a linearly independent eigenvectors of \hat{A} with the same eigenvalue a , i.e. $\{|a, i\rangle, i = 1, 2, \dots, g_a\}$. In terms of these basis vectors the projection operator \hat{P}_a onto H_a is written as

$$\hat{P}_a = \sum_{i=1}^{g_a} |a, i\rangle \langle a, i|. \quad (8.63)$$

Any vector in H_a (infinitely many of them) is an eigenvector of \hat{A} with eigenvalue a , but there are only g_a linearly independent vectors in H_a .

A subsequent measurement of B yielding some eigenvalue b will throw the system into the normalized projection of $|\psi'\rangle$ onto the eigensubspace H_b , i.e.,

$$|\psi'\rangle \xrightarrow{\text{Measure } B} |\psi''\rangle \in H_b.$$

The state $|\psi''\rangle$ lying in the eigensubspace H_b , will in general have components both in H_a and H_a^\perp , where H_a^\perp , called the orthogonal complement of H_a , is the subspace orthogonal to H_a . The subspace H_a^\perp consists of (is the union of) the eigensubspaces of \hat{A} other than H_a . So, a second measurement of A will not yield a , the value obtained in the first measurement, with certainty, i.e., A and B would be incompatible.

However, in exceptional cases it may be possible to find g_a linearly independent eigenvectors of the operator \hat{B} with eigenvalues b_1, b_2, \dots (some b_i 's may be repeated) in each eigensubspace H_a and, being linearly independent, the eigenvectors can span the eigensubspace H_a . These eigenvectors of \hat{B} , since they lie wholly in H_a , are also eigenvectors of \hat{A} with the same eigenvalue a . We may denote these simultaneous eigenvectors of \hat{A} and \hat{B} in H_a as follows:

$$|a, b', i_{(ab')}\rangle, i_{(ab')} = 1, 2, \dots, g_{(ab')}, b' = b_1, b_2, \dots. \quad (8.64)$$

Here $g_{(ab')}$ is the order of degeneracy of the eigenvalue b' in the eigensubspace H_a , or, in other words, the order of degeneracy of the eigenvalue pair (ab') in the full Hilbert space H . The order

of degeneracy of b' in the entire Hilbert space may be greater, for some of the linearly independent eigenvectors of \hat{B} with the same degenerate eigenvalue b' may lie in eigensubspaces other than H_a .

Thus, in the special case we are considering, we can find simultaneous eigenvectors of both \hat{A} and \hat{B} which span each eigensubspace H_a and so the whole Hilbert space can also be spanned by simultaneous eigenvectors of the two operators.

In terms of the simultaneous eigenvectors of \hat{A} and \hat{B} , we can write $|\psi'\rangle$ as

$$\begin{aligned} |\psi'\rangle &= N' \hat{P}_a |\psi\rangle \\ &= N' \sum_{b'=b_1, b_2, \dots} \sum_{i_{(ab')}=1}^{g_{(ab')}} |a, b', i_{(ab')}\rangle \langle a, b', i_{(ab')}| \psi \rangle \end{aligned} \quad (8.65)$$

where N' is the normalization constant for $|\psi'\rangle$, i.e.,

$$N' = \frac{1}{\sqrt{\sum_{b'=b_1, b_2, \dots} \sum_{i_{(ab')}=1}^{g_{(ab')}} |\langle a, b', i_{(ab')}| \psi \rangle|^2}} = \frac{1}{\sqrt{\langle \psi | \hat{P}_a | \psi \rangle}}. \quad (8.66)$$

Next, if we measure B immediately after A , we have non zero probability of getting one of the eigenvalues b_i that appears in the basis set $|a, b' \in \{b_1, b_2, \dots\}, i_{ab'}\rangle$ that spans H_a . Suppose we get a particular value b_i . The state of the system then collapses into $|\psi''\rangle$ which is the normalized projection of $|\psi'\rangle$ onto H_{b_i} , i.e.,

$$|\psi'\rangle \xrightarrow{\text{Measure } B} |\psi''\rangle = N'' \hat{P}_{b_i} \hat{P}_a |\psi\rangle \quad (8.67)$$

where N'' is the normalization constant for $|\psi''\rangle$. The projection operator on the eigensubspace H_{b_i} can be written as

$$\hat{P}_{b_i} = \sum_{a'} \sum_{\alpha_{(a'b_i)}=1}^{g_{(a'b_i)}} |a', b_i, \alpha_{(a'b_i)}\rangle \langle a', b_i, \alpha_{(a'b_i)}| \quad (8.68)$$

where $\alpha_{(a'b_i)}$ is the degeneracy index for the eigenvalue pair $(a'b_i)$. Noting the basis vectors are orthonormal, we have

$$\hat{P}_{b_i} \hat{P}_a = \sum_{\alpha_{(ab_i)}=1}^{g_{(ab_i)}} |a, b_i, \alpha_{(ab_i)}\rangle \langle a, b_i, \alpha_{(ab_i)}|. \quad (8.69)$$

The vector $|\psi''\rangle$ immediately after the measurement of B given in Eq. (8.67) can now be written as

$$|\psi''\rangle = N'' \sum_{\alpha_{(ab_i)}=1}^{g_{(ab_i)}} |a, b_i, \alpha_{(ab_i)}\rangle \langle a, b_i, \alpha_{(ab_i)}| \psi \rangle. \quad (8.70)$$

The normalization constant N'' is

$$N'' = \frac{1}{\sqrt{\sum_{\alpha_{(ab_i)}=1}^{g_{(ab_i)}} |\langle a, b_i, \alpha_{(ab_i)} | \psi \rangle|^2}} = \frac{1}{\sqrt{\langle \psi | \hat{P}_{b_i} \hat{P}_a | \psi \rangle}} \quad (8.71)$$

We note that the measurement of B has changed the state from $|\psi'\rangle$ to $|\psi''\rangle$, unlike in the nondegenerate case, but the changed state is still in H_a . So, a second measurement of A would certainly give the same eigenvalue a as was obtained in the first measurement of the observable. Therefore, the observables A and B are compatible. Further, the state $|\psi''\rangle$ is a simultaneous eigenstate of both \hat{A} and \hat{B} . So, in the state $|\psi''\rangle$, both the observables A and B have definite values, namely a and b_i , respectively.

To summarize, if in every eigensubspace H_a of the eigenvalues of \hat{A} , it is possible to find g_a linearly independent of eigenvectors of the operator \hat{B} , then the totality of all the simultaneous eigenvectors in the entire Hilbert space form a complete basis set of vectors. The two observables would then be compatible.

This is not to say that any eigenvector of \hat{A} is also an eigenvector of \hat{B} . For example, if we take any vector in H_a , the vector is guaranteed to be an eigenvector of \hat{A} , but not necessarily an eigenvector of \hat{B} . To see this let us consider the vector

$$|\psi\rangle = c_1|a, b_1\rangle + c_2|a, b_2\rangle \in H_a. \quad (8.72)$$

This is a vector in H_a and therefore an eigenvector of \hat{A} with eigenvalue a , but not an eigenvector of \hat{B} . There are infinity of different vectors in H_a but only g_a linearly independent ones which can act as basis for H_a . These linearly independent vectors, though eigenvectors of \hat{A} with eigenvalue a , may not in general be eigenvectors of \hat{B} also. However, by taking appropriate linear combinations of these linearly independent eigenvectors if it is possible to get another set of g_a linearly independent vectors which are also eigenvectors of \hat{B} , then the two observables are compatible. Therefore, for two observables to be compatible, they must have simultaneous eigenvectors which form a complete set of basis states for the Hilbert space.

8.9 Condition for Compatibility of Observables

In the previous section we have stated what we mean when we say two observables are compatible with each other. To recapitulate, if two observables have simultaneous eigenvectors that span the entire Hilbert space, i.e., if the simultaneous eigenvectors form a basis for the Hilbert space, then the observables are said to be compatible.

We will now prove that two observables A and B with corresponding Hermitian operators \hat{A} and \hat{B} , respectively, are compatible if and only if $[\hat{A}, \hat{B}] = 0$. Thus we prove the following theorem:

Theorem:

Two Hermitian operators representing two observables have a complete set of simultaneous eigenvectors if and only if they commute.

An alternative statement of the theorem could be: The necessary and sufficient condition that two observables are compatible is that their operators commute.

Proof:

The proof proceeds in two parts. First, we prove the necessary condition, i.e., we assume that the operators have simultaneous eigenvectors, then we show that the operators commute. Next, we prove the converse (the sufficiency condition), i.e., we assume that the operators commute, then we show they have simultaneous eigenvectors.

(a) Necessary Condition

Let us assume that \hat{A} and \hat{B} have a complete set of simultaneous eigenvectors $|u_n\rangle, n = 1, 2, \dots$. Here each u_n represents a unique set of numbers $\{a_n, b_n, \alpha_{(a_n b_n)}\}$, where a_n and b_n are eigenvalues of \hat{A} and \hat{B} , respectively, and $\alpha_{(a_n b_n)}$ is any other parameter which would be required to label the states uniquely should there be more than one linearly independent eigenvectors with the same values for the pair $(a_n b_n)$. Since the complete set of vectors $\{|u_n\rangle, n = 1, 2, \dots\}$ are eigenvectors of both \hat{A} and \hat{B} , we have

$$\hat{A}|u_n\rangle = a_n|u_n\rangle, \quad (8.73)$$

and

$$\hat{B}|u_n\rangle = b_n|u_n\rangle. \quad (8.74)$$

Now, using Eq. (8.73) and (8.74), we can write

$$\begin{aligned} [\hat{A}, \hat{B}]|u_n\rangle &= (\hat{A}\hat{B} - \hat{B}\hat{A})|u_n\rangle \\ &= (a_n b_n - b_n a_n)|u_n\rangle \\ &= 0. \end{aligned} \quad (8.75)$$

Since the simultaneous eigenvectors $\{|u_n\rangle, i = 1, 2, \dots\}$ form a complete set, it follows that

$$[\hat{A}, \hat{B}]|\psi\rangle = 0, \quad (8.76)$$

where $|\psi\rangle$ is an arbitrary vector in the Hilbert space. Hence we have

$$[\hat{A}, \hat{B}] = 0. \quad (8.77)$$

Thus, we have proved that commutativity is a necessary condition for compatibility.

(b) Sufficiency Condition

We now prove the converse, i.e., if \hat{A} and \hat{B} commute, they have simultaneous eigenvectors. First, we will assume that all eigenvalues of \hat{A} are non-degenerate. Next, we will consider the more general situation, namely that, some or all of the eigenvalues of \hat{A} may be degenerate.

Non-degenerate case

If all eigenvalues of \hat{A} are non-degenerate, then it is possible to uniquely label the eigenstates of \hat{A} by the eigenvalues only. Thus let $|a\rangle$ be the eigenstate of the operator \hat{A} belonging to some eigenvalue a . Therefore,

$$\hat{A}|a\rangle = a|a\rangle. \quad (8.78)$$

We start the proof by applying the operator \hat{B} to Eq. (8.78) to get

$$\hat{B}\hat{A}|a\rangle = a\hat{B}|a\rangle. \quad (8.79)$$

Since our assumption is that \hat{A} and \hat{B} commute, we can interchange the order of \hat{A} and \hat{B} on the left hand side of the above equation getting

$$\hat{A}(\hat{B}|a\rangle) = a(\hat{B}|a\rangle). \quad (8.80)$$

From Eq. (8.80) we can conclude that $\hat{B}|a\rangle$ is also an eigenvector of \hat{A} with the same eigenvalue a . Since we have supposed the eigenvalues of \hat{A} are non-degenerate, the vectors $|a\rangle$ and $\hat{B}|a\rangle$ represent the same physical state, i.e., $\hat{B}|a\rangle$ differs from $|a\rangle$ by a constant multiplier which we denote by b . Therefore,

$$\hat{B}|a\rangle = b|a\rangle, \quad (8.81)$$

i.e., all eigenvectors $\{|a\rangle\}$ of \hat{A} are also eigenvectors of \hat{B} . Therefore, we may label these states by both eigenvalues a and b , rather than the single eigenvalue a , i.e., $|a\rangle \equiv |a, b\rangle$.

Degenerate Case

Now, we allow for the more general case that some or all eigenvalues of \hat{A} may be degenerate. In case of degeneracy, the eigenvalue equation for a particular eigenvalue a is written as

$$\hat{A}|a, i\rangle = a|a, i\rangle; i = 1, 2, \dots, g_a, \quad (8.82)$$

where g_a is the order of degeneracy of a . Here a is an element of the set of all the eigenvalues of \hat{A} , i.e., $a \in \{a_1, a_2, \dots\}$. The g_a linearly independent eigenvectors $\{|a, i\rangle; i = 1, 2, \dots, g_a\}$ spanning the eigensubspace H_a are made orthonormal.

If we choose the eigenvectors of \hat{A} belonging to all the eigenspaces, i.e., the set of eigenvectors $\{|a, i\rangle, i = 1, 2, \dots, g_a; a = a_1, a_2, \dots\}$, as the basis for the Hilbert space, then obviously, the matrix representation of \hat{A} is diagonal. Denoting the matrix representing the operator \hat{A} by \underline{A} , we have where each diagonal block H_{a_n} is a $g_{a_n} \times g_{a_n}$ dimensional diagonal matrix with the eigenvalue a_n running along the diagonal, other entries being zero.

Next, we ask what would be the matrix representation of \hat{B} in the eigenbasis of \hat{A} . The matrix elements of \hat{B} are written as $\langle a', i' | \hat{B} | a, i \rangle$. Using our assumption that \hat{A} and \hat{B} commute, we have

$$\langle a', i' | [\hat{A}, \hat{B}] | a, i \rangle = 0,$$

\underline{A}	H_{a_1}	H_{a_2}	H_{a_3}
H_{a_1}	$a_1 \quad 0$ $0 \quad a_1$	0	0
H_{a_2}	0	$a_2 \quad 0$ $0 \quad a_2$	0
H_{a_3}	0	0	$a_3 \quad 0$ $0 \quad a_3$

Figure 8.4: The matrix representation of \hat{A} in the eigenbasis of \hat{A} .

i.e.,

$$\langle a', i' | \hat{A} \hat{B} - \hat{B} \hat{A} | a, i \rangle = 0,$$

or,

$$(a' - a) \langle a', i' | \hat{B} | a, i \rangle = 0. \quad (8.83)$$

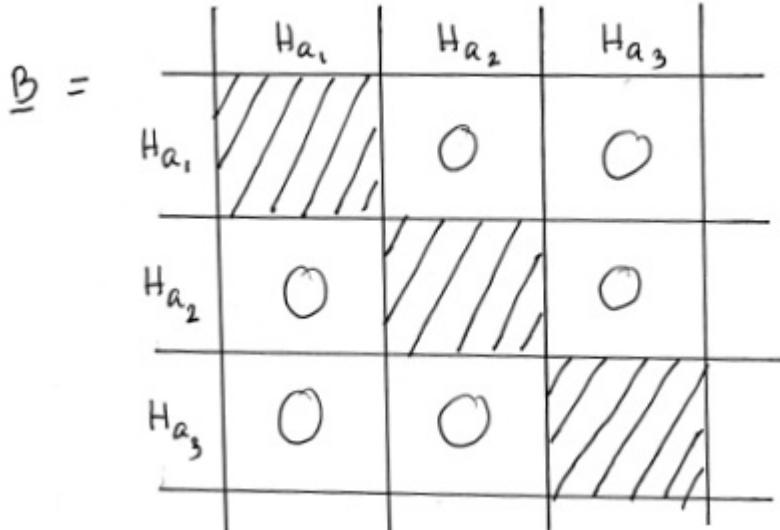
If $a \neq a'$, then $(a' - a) \neq 0$ and we must have

$$\langle a', i' | \hat{B} | a, i \rangle = 0, \quad \text{if } a' \neq a, \quad (8.84)$$

i.e., \hat{B} does not connect states with different eigenvalues of \hat{A} . In other words, \hat{B} acting on any of the eigenvectors of \hat{A} in the eigensubspace H_a , produces another vector which is also in H_a with no components in other eigensubspaces orthogonal to H_a . The new vector, being in H_a , remains an eigenvector of \hat{A} . Thus, \hat{B} acting on any eigenvector of \hat{A} produces a state which is also an eigenvector of \hat{A} with the same eigenvalue. This is a consequence of the fact that \hat{A} and \hat{B} commute.

From the above arguments, we can conclude that the matrix representation of \hat{B} in the eigenbasis of \hat{A} will be block diagonal as shown below.

In each eigensubspace H_a , the matrix $\underline{B}^{(a)}$ is a $g_a \times g_a$ square matrix, which itself need not be diagonal if we choose an arbitrary basis for H_a . To see this we refer to Eq. (8.84). This equation tells us that, if \hat{A} and \hat{B} commute, the matrix elements of B are zero if $a \neq a'$, but nothing is concluded about the matrix elements of \hat{B} in any eigensubspace H_a , i.e., when $a = a'$.

Figure 8.5: The matrix representation of \hat{B} in the eigenbasis of \hat{A} .

However, since \hat{B} is a Hermitian operator, its matrix representation in every eigensubspace H_a , is a finite dimensional square Hermitian matrix. We can always diagonalize a finite dimensional Hermitian matrix by a change of basis. The new set of g_a basis vectors in the eigensubspace H_a are linear combinations of the old basis vectors $\{|a, i\rangle, i = 1, 2, \dots, g_a\}$. Therefore, the vectors of the new basis set remain eigenvectors of \hat{A} with the same eigenvalue a . But, the vectors in new basis set, since it diagonalizes \hat{B} , must also be eigenvectors of \hat{B} with eigenvalues b_1, b_2, b_3, \dots where some of the b'_i 's may be repeated.

Since the vectors in the new basis set are simultaneous eigenvectors of \hat{A} and \hat{B} , we can label them by the eigenvalues of both \hat{A} and \hat{B} . Thus the common eigenvectors, which form the new basis set of H_a , can be written as $\{|a, b_1\rangle, |a, b_2\rangle, \dots, |a, b_{g_a}\rangle\}$ provided the b'_i 's are distinct. Some of the eigenvalues of \hat{B} may be repeated, in which case we will need another index to distinguish between the linearly independent simultaneous eigenvectors with the same values for the eigenvalue pair (ab_i) . We diagonalize each diagonal block $B^{(a_i)}$ of the matrix \underline{B} in the old basis, thereby obtaining simultaneous eigenvectors for both \hat{A} and \hat{B} in each eigensubspace. The set of simultaneous eigenvectors in all eigensubspaces $H_a, a \in \text{spec}(\hat{A})$ can now be used as the basis set for the Hilbert space. In the new basis consisting of the simultaneous eigenvectors, the matrix representation of \hat{B} is diagonal as shown in figure (8.6). The matrix representation of \hat{A} in the new basis is also diagonal and remains the same as in Fig. (8.4).

\hat{B}	H_{a_1}	H_{a_2}	H_{a_3}	.
	$b_1^{(1)}, b_2^{(1)}$ 0	0	0	.
H_{a_1}	0	0	0	$b_{g_1}^{(1)}$
	0	$b_1^{(2)}, b_2^{(2)}$ 0	0	$b_{g_2}^{(2)}$
H_{a_2}	0	0	0	.
	0	0	$b_1^{(3)}, b_2^{(3)}$ 0	$b_{g_3}^{(3)}$
H_{a_3}	0	0	0	.

Figure 8.6: The matrix representation of \hat{B} in the eigenbasis consisting of simultaneous eigenvectors of \hat{A} and \hat{B} .

The diagonal entries $b_i^{(n)}, i = 1, 2, \dots, g_n$ in each eigensubspace H_{a_n} in figure (8.6) are the various eigenvalues of the operator \hat{B} , i.e., $b_i^{(n)} \in \{b_1, b_2, \dots\}$. In a given eigensubspace H_{a_n} , some of the eigenvalues of \hat{B} may be repeated. Further, a particular eigenvalue of \hat{B} may also be repeated in several eigensubspaces H_{a_n} .

Thus, in summary, what we have shown is that, given two Hermitian operators \hat{A} and \hat{B} representing two observables of a quantum system, it is possible to construct a complete set of simultaneous eigenvectors spanning the entire Hilbert space provided $[\hat{A}, \hat{B}] = 0$. Hence, we have proved that $[\hat{A}, \hat{B}] = 0$ is a sufficient condition for two Hermitian operators \hat{A} and \hat{B} to be compatible.

8.10 Labeling of Quantum Mechanical Basis States

Since eigenvectors of a Hermitian operator corresponding to an observable form a complete set, we can label the members of the set by the eigenvalues. Thus, suppose that an observable \hat{A} has eigenvalues

$$a_1, a_2, a_3, \dots$$

The basis states are then labeled as

$$|a_1\rangle, |a_2\rangle, |a_3\rangle, \dots$$

The labeling would be unambiguous if the eigenvalues were all non-degenerate. However, if some or all eigenvalues are degenerate, we will need another mark of distinction for the eigenvectors.

For example, if an eigenvalue a_n is g_n -fold degenerate, the corresponding eigenvectors may be denoted as

$$|a_n, i\rangle, i = 1, 2, \dots, g_n.$$

An alternative, and better, notation is based on the fact that two compatible observables whose Hermitian operators commute, may be assumed to have the same set of eigenvectors. Hence, if we can find a second observable \hat{B} commuting with the first, such that

$$\hat{B}|a_n, i\rangle = b_i|a_n, i\rangle, i = 1, 2, \dots, g_n \quad (8.85)$$

with eigenvalues b_i all different, then the eigenvalues of \hat{B} may serve to distinguish the eigenvectors. So, we can write

$$|a_n, 1\rangle \equiv |a_n, b_1\rangle$$

$$|a_n, 2\rangle \equiv |a_n, b_2\rangle$$

.

.

.

.

$$|a_n, g_n\rangle \equiv |a_n, b_{g_n}\rangle.$$

But, if some of the eigenvalues $b_i, i = 1, 2, \dots, g_n$ are equal, we will need a third mark to distinguish the eigenvectors. The third mark may be obtained if we can find a third observable whose operator \hat{C} commutes with both \hat{A} and \hat{B} . Then \hat{A} , \hat{B} and \hat{C} have simultaneous eigenvectors and the eigenvalues of \hat{C} may also be used to label the eigenvectors.

Thus, in general, if we can find a set of mutually commuting Hermitian operators, $\hat{A}, \hat{B}, \hat{C}, \hat{D}, \dots$ whose common eigenvectors can be characterized completely by the eigenvalues a, b, c, d, \dots such that no two eigenvectors have exactly identical set of eigenvalues, then the eigenvalues of these operators can uniquely label the common eigenvectors. Such a set of Hermitian operators is said to be complete. We refer to this set of operators as a complete set of commuting observables (CSCO).

In the new notation, the basis vectors are written as

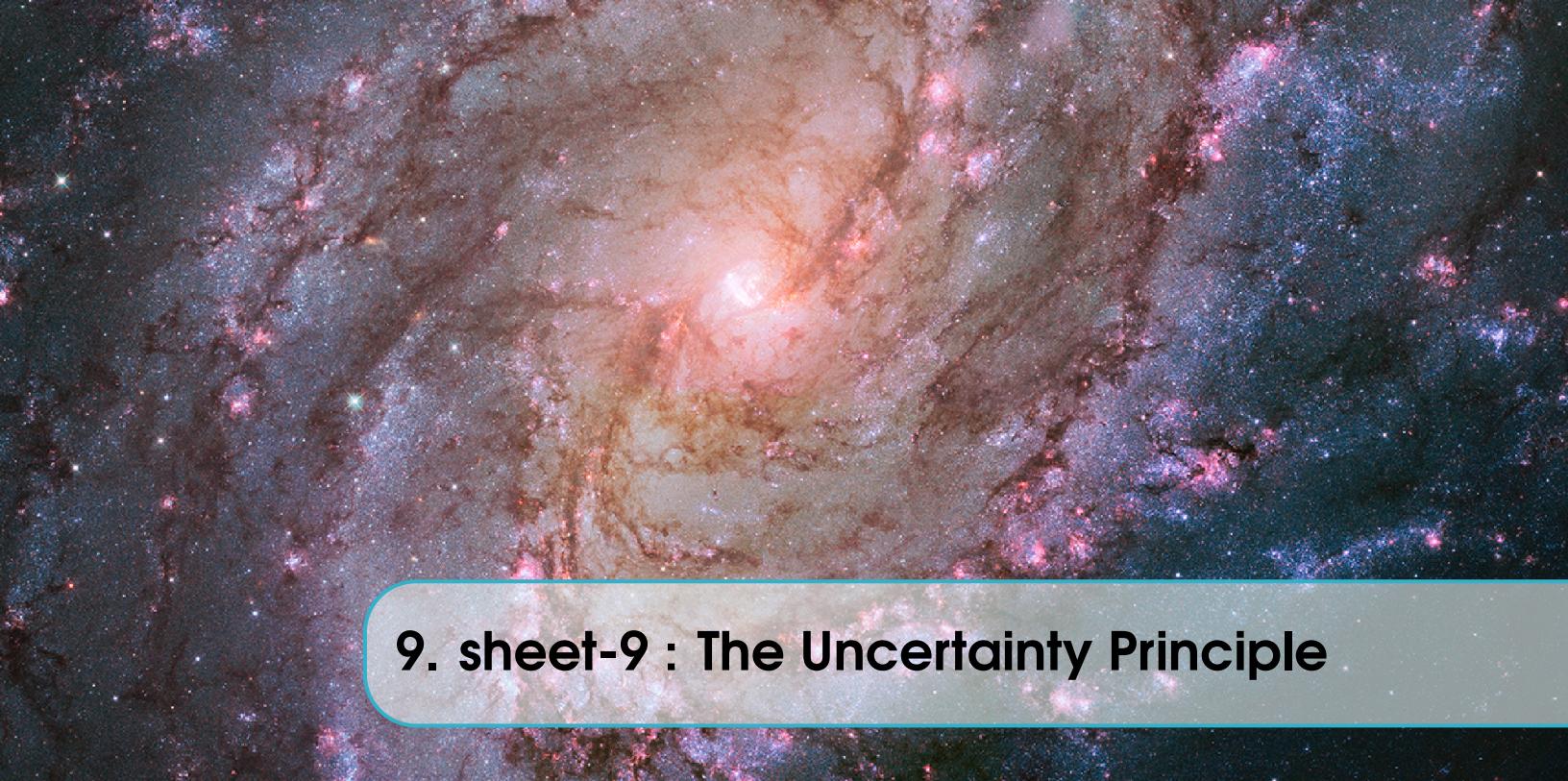
$$|a, b, c, d, \dots\rangle.$$

The normalization and completeness conditions are then

$$\langle a', b', c', d' \dots | a, b, c, d, \dots \rangle = \delta_{aa'} \delta_{bb'} \dots, \quad (8.86)$$

and

$$\sum_{(a,b,c,d,\dots)} |a, b, c, d, \dots\rangle \langle a, b, c, d, \dots| = \hat{I}. \quad (8.87)$$



9. sheet-9 : The Uncertainty Principle



10. sheet-10 : Harmonic Oscillator



11. sheet-11 : Quantum Dynamics

11.1 Quantum Dynamics

11.1.1 Schrodinger picture

The basic question of relativistic quantum dynamics is given an initial state $|\psi(t_0)\rangle$ of the system, how the state at time t , $|t\rangle$, is determined. The assertion that $|\psi(t_0)\rangle$ determines $|\psi(t)\rangle$ is the quantum mechanical form of the principle of causality, and we shall assume it.

In addition, we postulate an extension of the principle of superposition to include the temporal development of states. This states that if $|\psi_1(t_0)\rangle$ and $|\psi_2(t_0)\rangle$ separately evolve into $|\psi_1(t)\rangle$ and $|\psi_2(t)\rangle$, Then a superposition

$$|\psi(t_0)\rangle = \lambda_1 |\psi_1(t_0)\rangle + \lambda_2 |\psi_2(t_0)\rangle$$

develops into

$$|\psi(t)\rangle = \lambda_1 |\psi_1(t)\rangle + \lambda_2 |\psi_2(t)\rangle$$

i.e each component of the state moves independently of each other. This means that $|t\rangle$ can be obtained from an arbitrary initial state by the application of linear operator:

$$|\psi(t)\rangle = T(t, t_0) |\psi(t_0)\rangle$$

The operator T is called the time evolution operator for quantum mechanical state vectors

11.1.2 Schrodinger equation

The exact form of the time evolution operator can be found from the schrodinger equation, which is a postulate of quantum mechanics describing how the state vector changes with time. The schrodinger equation is

$$i\hbar \frac{\partial}{\partial t} |\psi t\rangle = H |\psi t\rangle$$

where H is a linear operator, called the Hamiltonian of the system.

12. sheet-12 : The Path Integral Formulation of

12.1 Background Materials

1. Basis states

$$\hat{Q}_s |q\rangle = q |q\rangle \quad (12.1)$$

$$\hat{P}_s |p\rangle = p |p\rangle \quad (12.2)$$

The states $\{|q\rangle\}$ and $\{|p\rangle\}$ are basis states, i.e., $\mathbb{1}$

$$\int dq |q\rangle \langle q| = \mathbb{1} \quad (12.3)$$

$$\int dp |p\rangle \langle p| = \hat{\mathbb{1}} \quad (12.4)$$

$$(12.5)$$

where the normalization is chosen as

$$\langle q|q'\rangle = \delta(q - q') \quad (12.6)$$

$$\langle p|p'\rangle = \delta(p - p') \quad (12.7)$$

The operators \hat{Q}_s and \hat{P}_s can be expressed in coordinate representation as follows

$$\langle q|\hat{Q}_s = q \langle q| \quad (12.8)$$

$$\langle q|\hat{P}_s = -i\hbar \frac{\partial}{\partial q} \langle q| \quad (12.9)$$

In momentum representaion

$$\langle p | \hat{Q}_s = i\hbar \frac{\partial}{\partial p} \langle p | \quad (12.10)$$

$$\langle p | \hat{P}_s = \langle p | \quad (12.11)$$

The fundamental commutation relation between \hat{Q} and \hat{P} is

$$[\hat{Q}_s, \hat{P}_s] = i\hbar \hat{1} \quad (12.12)$$

For later purposes we will need the momentum eigenstates in coordinate representation, i.e., $\langle q | p \rangle$. To find $\langle q | p \rangle$ we proceed as follows

$$\hat{P}_s | p \rangle = p | p \rangle \quad (12.13)$$

$$\langle q | \hat{P}_s | p \rangle = p \langle q | p \rangle \quad (12.14)$$

$$-i\hbar \frac{\partial}{\partial q} \langle q | p \rangle = p \langle q | p \rangle \quad (12.15)$$

This equation is easy to solve for $\langle q | p \rangle$. We find

$$\langle q | p \rangle = C e^{ipq/\hbar} \quad (12.16)$$

The constant C is chosen such that we have the normalization $\langle p | p' \rangle \delta(p - p')$. Now

$$\begin{aligned} \langle p | p' \rangle &= \int dq \langle p | q \rangle \langle q | p' \rangle \\ &= \int dq C^* e^{-ipq/\hbar} C^{ip'q/\hbar} \\ &= |C|^2 \int_{-\infty}^{\infty} dq e^{-i(p-p')q/\hbar} \\ &= |C|^2 2\pi \delta\left(\frac{p-p'}{\hbar}\right) \\ &= |C|^2 2\pi\hbar \delta(p - p') \end{aligned}$$

Choosing C to be real and positive, we must have

$$C = \frac{1}{\sqrt{2\pi\hbar}} \quad (12.17)$$

Thus

$$\langle q | p \rangle = \frac{1}{\sqrt{2\pi\hbar}} e^{ipq/\hbar} \quad (12.18)$$

2. Quantum Mechanics in Heisenberg picture// The Heisenberg picture of quantum dynamics is obtained from the Schrödinger picture by the following transformation of all kets and all operators

$$|\rangle_H = e^{i\hat{H}t/\hbar} |\rangle_s \quad (12.19)$$

$$\hat{Q}_H(t) = e^{i\hat{H}t/\hbar} \hat{Q}_S e^{-i\hat{H}t/\hbar} \quad (12.20)$$

where we have assumed the system is conservative, i.e., \hat{H} is independent of time.

In the Heisenberg picture, the base kets, for example, the eigenkets of $\hat{Q}_H(t)$ and $\hat{P}_H(t)$ are time dependent. We have

$$\hat{Q}_H(t) |q, t\rangle_H = q |q, t\rangle_H \quad (12.21)$$

$$\hat{P}_H(t) |p, t\rangle_H = p |p, t\rangle_H \quad (12.22)$$

where

$$|q, t\rangle_H = e^{i\hat{H}t/\hbar} |q\rangle \quad (12.23)$$

$$|p, t\rangle_H = e^{i\hat{H}t/\hbar} |p\rangle \quad (12.24)$$

and

$$\hat{Q}_H(t) = e^{i\hat{H}t/\hbar} \hat{Q}_S e^{-i\hat{H}t/\hbar} \quad (12.25)$$

$$\hat{P}_H(t) = e^{i\hat{H}t/\hbar} \hat{P}_S e^{-i\hat{H}t/\hbar} \quad (12.26)$$

The orthogonality and completeness of the Heisenberg picture base kets are

$${}_H \langle q, t | q', t \rangle_H = \langle q | q' \rangle = \delta(q - q') \quad (12.27)$$

$${}_H \langle p, t | p', t \rangle_H = \langle p | p' \rangle = \delta(p - p') \quad (12.28)$$

Note that these are equal time relations. And

$$\hat{\mathbb{1}} = \int dq |q, t\rangle_H {}_H \langle q, t | \quad (12.29)$$

$$\hat{\mathbb{1}} = \int dp |p, t\rangle_H {}_H \langle p, t | \quad (12.30)$$

To show the validity of equation (12.29), for example, we use the transformation of kets and bras from the schrödinger picture to the Heisenberg picture (12.19), i.e.,

$${}_H \langle | = s \langle | e^{-i\hat{H}t/\hbar} \quad (12.31)$$

Thus the right hand side of equation (12.29) can be written as

$$\int dq |q, t\rangle_{HH} \langle q, t| = \int dq e^{i\hat{H}t/\hbar} |q\rangle \langle q| e^{-i\hat{H}t/\hbar} \quad (12.32)$$

$$= e^{i\hat{H}t/\hbar} \left(\int dq |q\rangle \langle q| \right) e^{-i\hat{H}t/\hbar} \quad (12.33)$$

$$= e^{i\hat{H}t/\hbar} \mathbb{1} e^{-i\hat{H}t/\hbar} \quad (12.34)$$

$$= \mathbb{1} \quad (12.35)$$

We note that the state vector in the Heisenberg picture is independent of time, while the state vector in the Schrödinger picture is time-dependent. This is very simply shown as follows:

$$|\psi\rangle_H = e^{i\hat{H}t/\hbar} |\psi(t)\rangle_S \quad (12.36)$$

$$= e^{i\hat{H}t/\hbar} e^{-i\hat{H}t/\hbar} |\psi(0)\rangle_S \quad (12.37)$$

$$= |\psi(0)\rangle_S \quad (12.38)$$

Thus the state ket in the Heisenberg picture is independent of time and is same as the initial state ket in the Schrödinger picture.

12.2 Propagator

The dynamics of a quantum system is completely specified by the "Feynman Kernel", or the propagator or the transition amplitude defined as

$$U(q_2, t_2; q_1, t_1) = \langle q_2, t_2 | q_1, t_1 \rangle_H \quad (12.39)$$

Transforming to the Schrödinger picture base kets, we can write equation (12.39) as

$$\begin{aligned} U(q_2, t_2; q_1, t_1) &= \langle q_2 | e^{-i\hat{H}t_2/\hbar} e^{i\hat{H}t_1/\hbar} | q_1 \rangle \\ &= \langle q_2 | e^{-i\hat{H}(t_2-t_1)/\hbar} | q_1 \rangle \end{aligned} \quad (12.40)$$

We see that the propagator is the matrix element in the coordinate basis of the time-evolution operator in the Schrödinger picture. The physical meaning of the propagator is that it is the probability amplitude of finding the particle at q_2 at time t_2 if the particle was at q_1 at an earlier time t_1 . Knowing the propagator is equivalent to solving the Schrödinger equation, for it allows us to calculate the Schrödinger picture wave function at any moment of time if the wave function is known at an earlier

moment. This is shown below:

$$\psi_S(q, t) = \langle q | \psi_S(t) \rangle \quad (12.41)$$

$$= \langle q | e^{-i\hat{H}t/\hbar} | \psi_S(0) \rangle \quad (12.42)$$

$$= -H \langle q, t | \psi \rangle_H \quad (12.43)$$

$$= \int dq' -H \langle q, t | q', t' \rangle_H -H \langle q', t' | \psi \rangle_H \quad (12.44)$$

$$= \int dq' U(q, t; q', t') \psi_S(q', t') \quad (12.45)$$

The path integral formalism of quantum dynamics provides a means to construct the transition amplitude $-H \langle q, t | q', t' \rangle_H$ from the classical Hamiltonian or Lagrangian alone, without any reference to non commuting operators or Hilbert space vectors.

12.3 Path Integral for the Propagator

We will now calculate

$$U(x, t; x_0, t_0) = -H \langle x, t | x_0, t_0 \rangle \quad (12.46)$$

where $t > t_0$. For this purpose let us devide the time interval (t, t_0) into N equal segments each of duration ϵ . Namely, let

$$\epsilon = \frac{t - t_0}{N} \quad (12.47)$$

In other words, we are discretizing the time interval, and in the end we will take the continuum limit $\epsilon \rightarrow 0$ and $N \rightarrow \infty$. We label the end times t_0 and t and the intermediate times as $t_1, t_2, \dots, t_{N-1}, t_N = t$. Further, we will let $x_N = x$. The intermediate times are

$$t_i = t_0 + i\epsilon, \quad i = 1, 2, \dots, N-1 \quad (12.48)$$

At each intermediate time a complete set of basis states $|x_i, t_i\rangle$ may be inserted:

$$\langle xt | x_0 t_0 \rangle = \int_{-\infty}^{\infty} dx_1 \dots \int_{-\infty}^{\infty} dx_{N-1} \langle xt | x_{N-1} t_{N-1} \rangle \langle x_{N-1} t_{N-1} | x_{N-2} t_{N-2} \rangle \dots \langle x_2 t_2 | x_1 t_1 \rangle \langle x_1 t_1 | x_0 t_0 \rangle \quad (12.49)$$

Here we have omitted the subscript H in the Heisenberg picture basis vectors since there is no scope for confusion. Note that while there are N scalar products in euquation (12.49), there are only $N - 1$ intermediate points so that the number of integrations is $N - 1$. Since $x_N = x$ and $t_N = t$, we can write equation (12.49) as

$$\langle xt | x_0 t_0 \rangle = \int \prod_{i=1}^{N-1} dq_i \prod_{i=0}^{N-1} \langle x_{i+1} t_{i+1} | x_i t_i \rangle \quad (12.50)$$

Figure 12.1: text

Equation (12.50) can be interpreted as follows: A particle that propagates from x_0 at time t_0 to x at time t can take an arbitrary intermediate trajectory, shown in figure (12.1). Such a path is characterized by the coordinate values x_i at intermediate grid @@@@ in the time interval (t_0, t) . One such path is shown in the figure as a zigzag curve. Since each intermediate coordinates x_i ($i = 1, 2, \dots, N - 1$) can vary from $-\infty$ to ∞ , it is essential that all conceivable paths connecting the end points are taken into account. According to the representation principle of Quantum Mechanics they all contribute to the transition amplitude (12.50). Of course, some trajectories may turn out to be more important than others.

We will now calculate the intermediate scalar products which themselves are propagators but over infinitesimal time intervals. An intermediate scalar product has the form $\langle x_{i+1} t_{i+1} | x_i t_i \rangle$. We can calculate this inner product up to first order in ϵ from equation (12.47) as follows

$$\langle x_{i+1} t_{i+1} | x_i t_i \rangle = \langle x_{i+1} | e^{-i\hat{H}t_{i+1}/\hbar} e^{i\hat{H}t_i/\hbar} | x_i \rangle \quad (12.51)$$

$$= \langle x_{i+1} | e^{-i\hat{H}(t_{i+1}-t_i)/\hbar} | x_i \rangle \quad (12.52)$$

$$= \langle x_{i+1} | e^{-i\hat{H}\epsilon/\hbar} | x_i \rangle \quad (12.53)$$

$$= \langle x_{i+1} | \left(\mathbb{1} - \frac{i\epsilon}{\hbar} \hat{H} + \mathcal{O}(\epsilon^2) \right) | x_i \rangle \quad (12.54)$$

We will take \hat{H} to be of the form

$$\hat{H} = \frac{\hat{P}}{2m} + V(\hat{X}) \quad (12.55)$$

Therefore

$$\langle x_{i+1} t_{i+1} | x_i t_i \rangle = \langle x_{i+1} | \left[\mathbb{1} - \frac{i\epsilon}{\hbar} \left(\frac{\hat{P}^2}{2m} + V(\hat{X}) \right) + \mathcal{O}(\epsilon^2) \right] | x_i \rangle \quad (12.56)$$

$$= \int_{-\infty}^{\infty} dp \langle x_{i+1} | p \rangle \langle p | \left[\mathbb{1} - \frac{i\epsilon}{\hbar} \left(\frac{\hat{P}^2}{2m} + V(\hat{X}) \right) + \mathcal{O}(\epsilon^2) \right] | x_i \rangle \quad (12.57)$$

$$= \int_{-\infty}^{\infty} dp \langle x_{i+1} | p \rangle \langle p | x_i \rangle \left[\mathbb{1} - \frac{i\epsilon}{\hbar} \left(\frac{p^2}{2m} + V(x_i) \right) + \mathcal{O}(\epsilon^2) \right] \quad (12.58)$$

$$= \int dp \frac{1}{\sqrt{2\pi\hbar}} e^{ipx_{i+1}/\hbar} \frac{1}{\sqrt{2\pi\hbar}} e^{-ipx_i/\hbar} e^{-i\epsilon/\hbar \left(\frac{p^2}{2m} + V(x_i) \right)} \quad (12.59)$$

$$= \frac{1}{2\pi\hbar} \int_{-\infty}^{\infty} dp e^{ip(x_{i+1}-x_i)/\hbar} e^{-i\epsilon/\hbar \left(\frac{p^2}{2m} + V(x_i) \right)} \quad (12.60)$$

$$= \frac{1}{2\pi\hbar} e^{-i\epsilon V(x_i)/\hbar} \int_{-\infty}^{\infty} e^{ip\epsilon \frac{(x_{i+1}-x_i)}{\hbar}} e^{-i\epsilon/\hbar \left(\frac{p^2}{2m} \right)} dp \quad (12.61)$$

$$= \frac{1}{2\pi\hbar} e^{-i\epsilon V(x_i)/\hbar} \int_{-\infty}^{\infty} e^{ip\epsilon \dot{x}_i/\hbar} e^{-i\epsilon \frac{p^2}{2m\hbar}} dp \quad (12.62)$$

$$= \frac{1}{2\pi\hbar} e^{-i\epsilon V(x_i)/\hbar} \int_{-\infty}^{\infty} e^{-\frac{i\epsilon}{2m\hbar} (p^2 - 2mp\dot{x}_i)} dp \quad (12.63)$$

In the above we have defined

$$\dot{x}_i = \frac{x_{i+1} - x_i}{\epsilon} \quad (12.64)$$

Now

$$p^2 - 2mp\dot{x}_i = (p - m\dot{x}_i)^2 - m^2\dot{x}_i^2 \quad (12.65)$$

We make the change of variable

$$p' = p - m\dot{x}_i \quad (12.66)$$

Therefore equation (12.63) can be written as

$$\langle x_{i+1} t_{i+1} | x_i t_i \rangle = \frac{1}{2\pi\hbar} e^{-i\epsilon V(x_i)/\hbar} e^{-i\epsilon(-m^2\dot{x}_i^2)/2m\hbar} \int_{-\infty}^{\infty} dp' e^{-i\epsilon p'^2/2m\hbar} \quad (12.67)$$

$$= \frac{1}{2\pi\hbar} e^{\frac{i\epsilon}{\hbar} (\frac{1}{2}m\dot{x}_i^2 - V(x_i))} \int_{-\infty}^{\infty} dp' e^{-i\epsilon p'^2/2m\hbar} \quad (12.68)$$

Now we use the standard integral

$$\int_{-\infty}^{\infty} e^{-\alpha x^2} dx = \sqrt{\frac{\pi}{\alpha}} \quad (12.69)$$

to get

$$\int_{-\infty}^{\infty} dp' e^{-i\epsilon p'^2/2m\hbar} = \left(\frac{\pi}{i\epsilon/2m\hbar} \right)^{1/2} = \left(\frac{2\pi\hbar m}{i\epsilon} \right)^{1/2} \quad (12.70)$$

Therefore

$$\langle x_{i+1} t_{i+1} | x_i t_i \rangle = \frac{1}{2\pi\hbar} \left(\frac{2\pi\hbar m}{i\varepsilon} \right)^{1/2} e^{\frac{i\varepsilon}{\hbar} (\frac{1}{2}m\dot{x}_i^2 - V(x_i))} \quad (12.71)$$

$$= \left(\frac{m}{2\pi\hbar i\varepsilon} \right)^{1/2} e^{\frac{i\varepsilon}{\hbar} (\frac{1}{2}m\dot{x}_i^2 - V(x_i))} \quad (12.72)$$

We now substitute equation (12.72) into equation (12.50) to get

$$\langle xt | x_0 t_0 \rangle = \int \prod_{i=1}^{N-1} dx_i \prod_{i=0}^{N-1} \langle x_{i+1} t_{i+1} | x_i t_i \rangle \quad (12.73)$$

$$= \int \prod_{i=1}^{N-1} dx_i \prod_{i=0}^{N-1} \left(\frac{m}{2\pi\hbar i\varepsilon} \right)^{1/2} e^{\frac{i\varepsilon}{\hbar} (\frac{1}{2}m\dot{x}_i^2 - V(x_i))} \quad (12.74)$$

$$= \left(\frac{m}{2\pi\hbar i\varepsilon} \right)^{N/2} \int \prod_{i=1}^{N-1} dx_i e^{\frac{i}{\hbar} \sum_{i=0}^{N-1} \varepsilon (\frac{1}{2}m\dot{x}_i^2 - V(x_i))} \quad (12.75)$$

We now consider a path $x(t')$ connecting the initial and the final space-time point such that the value of $x(t')$ at the intermediate times t_1, t_2, \dots, t_{N-1} are $x(t'_i) = x_i$. Therefore we can write

$$\sum_{i=0}^{N-1} \varepsilon \left(\frac{1}{2}m\dot{x}_i^2 - V(x_i) \right) = \int_{t_0}^t \left[\frac{1}{2}m\dot{x}^2(t') - V(x(t')) \right] dt' \quad (12.76)$$

$$= \int_{t_0}^t L(x(t'), \dot{x}(t')) dt' \quad (12.77)$$

$$= S[x(t')] \quad (12.78)$$

Where $S[x(t')]$ is the action calculated along the particular path. Since we are integrating over x_i ($i = 1, \dots, N-1$), we are effectively summing the exponential in equation (12.75) over all conceivable paths connecting (x_0, t_0) to (x, t) . We define the path integral as

$$\mathcal{D}[x(t')] = \lim_{N \rightarrow \infty} \left(\frac{m}{2\pi\hbar i\varepsilon} \right)^{N/2} \int \prod_{i=1}^{N-1} dx_i \quad (12.79)$$

Therefore, we can write equation (12.75) as

$$\langle x, t | x_0, t_0 \rangle = \int \mathcal{D}[x(t')] e^{\frac{i}{\hbar} S[x(t')]} \quad (12.80)$$

This is the path integral formula for the propagator. We can think of equation (12.80) as a symbolic way of writing equation (12.75) with $N \rightarrow \infty$. In calculating path integrals we use equation (12.75) and set $N \rightarrow \infty$.

12.4 Path Integral for a Free Particle

For a free particle $V = 0$. Therefore the lagrangian is

$$L = T - V = T = \frac{1}{2}m\dot{x}^2(t) \quad (12.81)$$

Figure 12.2: text

the path integral formula for the propagator of a free particle is

$$\langle x, t | x_0, t_0 \rangle = \int \mathcal{D}[x(t')] \exp \left[\frac{i}{\hbar} S[x(t')] \right] \quad (12.82)$$

$$= \lim_{N \rightarrow \infty} \left(\frac{m}{2\pi\hbar i\varepsilon} \right)^{N/2} \int \prod_{i=1}^{N-1} dx_i \exp \left[\frac{i\varepsilon}{\hbar} \sum_{i=0}^{N-1} \frac{1}{2} m \dot{x}_i^2 \right] \quad (12.83)$$

In equation (12.83)

$$\varepsilon = \frac{t - t_0}{N} \quad (12.84)$$

Also \dot{x}_i can be written as

$$\dot{x}_i = \frac{x_{i+1} - x_i}{\varepsilon} \quad (12.85)$$

For notational convenience we let $x_N = x$ where x is the final position. We only integrate over the position the particle may have at intermediate times t_1, t_2, \dots, t_{N-1} .

Using equation (12.85), equation (12.83) can be written as

$$\langle xt | x_0 t_0 \rangle = \lim_{N \rightarrow \infty} \left(\frac{m}{2\pi\hbar i\varepsilon} \right)^{N/2} \int \prod_{i=1}^{N-1} dx_i e^{\frac{i\varepsilon}{\hbar} \sum_{i=0}^{N-1} \frac{m}{2} \left(\frac{x_{i+1} - x_i}{\varepsilon} \right)^2} \quad (12.86)$$

$$= \lim_{N \rightarrow \infty} \left(\frac{m}{2\pi\hbar i\varepsilon} \right)^{N/2} \int \prod_{i=1}^{N-1} dx_i e^{\frac{im}{2\hbar\varepsilon} \sum_{i=0}^{N-1} (x_{i+1} - x_i)^2} \quad (12.87)$$

At this stage it is convenient to make a change of variable

$$y_i = \left(\frac{m}{2\hbar\varepsilon} \right)^{1/2} x_i \quad (12.88)$$

In terms of new variables equation (12.87) is written as

$$\langle xt | x_0 t_0 \rangle = \lim_{N \rightarrow \infty} \left(\frac{m}{2\pi\hbar i\varepsilon} \right)^{N/2} \left(\frac{2\hbar\varepsilon}{m} \right)^{(N-1)/2} \int \prod_{i=1}^{N-1} dy_i e^{-\sum_{i=0}^{N-1} \frac{(y_{i+1} - y_i)^2}{i}} \quad (12.89)$$

We now have to do the gaussian integral over the variables y_1, y_2, \dots, y_{N-1} .

y_1 integral

$$I_1 = \int_{-\infty}^{\infty} dy_1 \exp \left[-\frac{1}{i} [(y_1 - y_0)^2 + (y_2 - y_1)^2] \right] \quad (12.90)$$

consider the exponent

$$(y_1 - y_0)^2 + (y_2 - y_1)^2 = 2y_1^2 - 2(y_0 + y_2)y_1 + (y_0^2 + y_2^2) \quad (12.91)$$

therefore,

$$I_1 = \exp \left[-\frac{1}{i} (y_0^2 + y_2^2) \right] \int_{-\infty}^{\infty} dy_1 \exp \left[-\frac{1}{i} (2y_1^2 - 2(y_0 + y_2)y_1) \right] \quad (12.92)$$

Now we use the standard integral

$$\int_{-\infty}^{\infty} e^{-\alpha x^2 + \beta x} dx = \left(\frac{\pi}{\alpha} \right)^{1/2} \exp \left(\frac{\beta^2}{4\alpha} \right) \quad (12.93)$$

choosing $\alpha = \frac{2}{i}$ and $\beta = \frac{2(y_0 + y_2)}{i}$

$$\begin{aligned} I_1 &= \exp \left[-\frac{1}{i} (y_0^2 + y_2^2) \right] \left(\frac{i\pi}{2} \right)^{1/2} \exp \left[\frac{-4(y_0 + y_2)^2}{4(2/i)} \right] \\ &= \exp \left[-\frac{1}{i} (y_0^2 + y_2^2) \right] \left(\frac{i\pi}{2} \right)^{1/2} \exp \left[\frac{(y_0 + y_2)^2}{2i} \right] \\ &= \left(\frac{i\pi}{2} \right)^{1/2} \exp \left[-\frac{1}{2i} (2(y_0^2 + y_2^2) - (y_0 + y_2)^2) \right] \end{aligned}$$

Thus

$$I_1 = \left(\frac{i\pi}{2} \right)^{1/2} \exp \left[-\frac{1}{2i} (y_2 - y_0)^2 \right] \quad (12.94)$$

Next we do the integral over y_2 . The variable y_2 occurs in the $i = 2$ term in equation (12.89) and also in I_1 in equation (12.94). Therefore, the y_2 integral is

$$\begin{aligned} I_2 &= \int dy_2 \exp \left[-\frac{1}{i} (y_3 - y_2)^2 \right] \left(\frac{i\pi}{2} \right)^{1/2} \exp \left[-\frac{1}{2i} (y_2 - y_0)^2 \right] \\ &= \left(\frac{i\pi}{2} \right)^{1/2} \int dy_2 \exp \left[-\frac{1}{i} (y_3 - y_2)^2 - \frac{1}{2i} (y_2 - y_0)^2 \right] \\ &= \left(\frac{i\pi}{2} \right)^{1/2} \int dy_2 \exp \left[-\frac{1}{i} \{ (y_3 - y_2)^2 + (y_2 - y_0)^2 \} \right] \end{aligned}$$

Consider the term in the curly brackets

$$\begin{aligned} 2(y_3 - y_2)^2 + (y_2 - y_0)^2 &= 2(y_3^2 + y_2^2 - 2y_2y_3) + (y_2^2 + y_0^2 - 2y_0y_2) \\ &= 3y_2^2 - 2y_2(2y_3 + y_0) + (2y_3^2 + y_0^2) \end{aligned}$$

where the first term is quadratic in y_2 , second term is linear in y_2 and the last term is independent of y_2 . We have

$$I_2 = \left(\frac{i\pi}{2} \right)^{1/2} \exp \left[-\frac{1}{2i} (2y_3^2 + y_0^2) \right] \int dy \exp \left[-\frac{1}{2i} (3y_2^2 - 2y_2(2y_3 + y_0)) \right] \quad (12.95)$$

We use standard integral (64) from appendix (26.5.1) and choose

$$\alpha = \frac{3}{2i} \quad (12.96)$$

$$\beta = \frac{(y_0 + 2y_3)}{i} \quad (12.97)$$

$$\begin{aligned} I_2 &= \left(\frac{i\pi}{2} \right)^{1/2} \exp \left[-\frac{1}{2i} (2y_3^2 + y_0^2) \right] \left(\frac{2\pi i}{3} \right)^{1/2} \exp \left[-\frac{(y_0 + 2y_3)^2}{6/i} \right] \\ &= \left(\frac{i\pi}{2} \right)^{1/2} \left(\frac{2\pi i}{3} \right)^{1/2} \exp \left[-\frac{1}{2i} (2y_3^2 + y_0^2) \right] \exp \left[\frac{(y_0 + 2y_3)^2}{6i} \right] \\ &= \left(\frac{i^2\pi^2}{3} \right)^{1/2} \exp \left[-\frac{1}{i} \left\{ \frac{1}{3}y_3^2 + \frac{1}{3}y_0^2 - \frac{2}{3}y_0y_3 \right\} \right] \end{aligned}$$

$$I_2 = \left(\frac{i^2\pi^2}{3} \right)^{1/2} \exp \left[-\frac{(y_3 - y_0)^2}{3i} \right] \quad (12.98)$$

Now the trend is clear. Finally, integrating $(N - 1)$ times we get

$$I_{N-1} = \frac{(i\pi)^{(N-1)/2}}{N^{1/2}} \exp \left[-(y_N - y_0)^2 / Ni \right] \quad (12.99)$$

where $y_N = y$.

Therefore, the path integral formula for the propagator of a free particle is (using the above formula in equation (12.89))

$$\langle xt | x_0 t_0 \rangle_{free} = U_{free}(x, t; x_0, t_0) = \lim_{N \rightarrow \infty} \left(\frac{m}{2\pi\hbar i\varepsilon} \right)^{N/2} \left(\frac{2\hbar\varepsilon}{m} \right)^{(N-1)/2} \frac{(i\pi)^{(N-1)/2}}{N^{1/2}} \exp \left[-(y_N - y_0)^2 / Ni \right] \quad (12.100)$$

previously we defined

$$y = \left(\frac{m}{2\hbar\varepsilon} \right)^{1/2} x \quad (12.101)$$

Therefore

$$\begin{aligned} U(x, t; x_0, t_0) &= \lim_{N \rightarrow \infty} \left(\frac{m}{2\pi\hbar i\varepsilon} \right)^{N/2} \left(\frac{2\hbar\varepsilon}{m} \right)^{(N-1)/2} \frac{(i\pi)^{(N-1)/2}}{N^{1/2}} \exp \left[-\frac{m(x_N - x_0)^2}{2\hbar\varepsilon Ni} \right] \\ &= \lim_{N \rightarrow \infty} \left(\frac{m}{2\pi\hbar i N\varepsilon} \right)^{N/2} \left(\frac{2\pi\hbar Ni\varepsilon}{m} \right)^{(N-1)/2} \exp \left[-\frac{m(x_N - x_0)^2}{2\hbar\varepsilon Ni} \right] \end{aligned}$$

Now $\lim_{N \rightarrow \infty} N\varepsilon = t - t_0 \equiv \Delta t$, also $x_N = x$.

Hence

$$U(x, t; x_0, t_0) = \left(\frac{m}{2\pi\hbar i(t - t_0)} \right)^{1/2} \exp \left[-\frac{m(x - x_0)^2}{2\hbar i(t - t_0)} \right] \quad (12.102)$$

This is the propagator for a free particle obtained using the path integral formula.

12.4.1 check of calculation

We have

$$\lim_{t \rightarrow t_0} \langle xt | x_0 t_0 \rangle = \delta(x - x_0) \quad (12.103)$$

Therefore equation (12.102) must reduce to the delta function $\delta(x - x_0)$ when $t = t_0$. Taking $\Delta = \sqrt{\frac{2\hbar i(t-t_0)}{m}}$ in equation (12.102) we can write

$$U_{free}(x, t; x_0, t_0) = \frac{1}{\pi^{1/2}\Delta} \exp \left[-(x - x_0)^2 / \Delta^2 \right] \quad (12.104)$$

In the limit $t \rightarrow t_0, \Delta \rightarrow 0$. Therefore

$$\lim_{t \rightarrow t_0} U_{free}(x, t; x_0, t_0) = \lim_{\Delta \rightarrow 0} \frac{1}{\pi^{1/2}\Delta} \exp \left[-(x - x_0)^2 / \Delta^2 \right] \quad (12.105)$$

$$= \delta(x - x_0) \quad (12.106)$$

Thus, the free particle propagator (equation (12.102)) has the correct limiting behavior in the limit $t \rightarrow t_0$.

12.5 Derivation of the propagator for a free particle without using the path integral formula

Since for a free particle, the Hamiltonian is simple and its eigenvalues and eigenvectors are known, we can find the propagator $U(x, t; x_0, t_0)$ without using the path integral formula. We now calculate the propagator for a free particle directly without using the path integral formula.

The propagator is

$$\langle xt | x_0 t_0 \rangle \equiv U(x, t; x_0, t_0) \quad (12.107)$$

$$= \langle x | e^{-i\hat{H}(t-t_0)/\hbar} | x_0 \rangle \quad (12.108)$$

$$= \int dp \langle x | e^{-i\hat{H}(t-t_0)/\hbar} | p \rangle \langle p | x_0 \rangle \quad (12.109)$$

$$= \int dp \langle x | e^{-i\frac{p^2(t-t_0)}{2m\hbar}} | p \rangle \langle p | x_0 \rangle \quad (12.110)$$

$$= \int dp e^{-i\frac{p^2(t-t_0)}{2m\hbar}} \langle x | p \rangle \langle p | x_0 \rangle \quad (12.111)$$

$$= \int dp e^{-ip^2\Delta t/2m\hbar} \frac{1}{\sqrt{2\pi\hbar}} e^{ipx/\hbar} \frac{1}{\sqrt{2\pi\hbar}} e^{ipx_0/\hbar} \quad (12.112)$$

$$= \frac{1}{2\pi\hbar} \int_{-\infty}^{\infty} dp \exp[-ip^2\Delta t/2m\hbar + ip(x-x_0)/\hbar] \quad (12.113)$$

Using standard integral from appendix (equation 64) and

$$\alpha = \frac{i\Delta t}{2m\hbar} \quad (12.114)$$

$$\beta = \frac{i(x-x_0)}{\hbar} \quad (12.115)$$

$$\begin{aligned} U(x, t' x_0, t_0) &= \left(\frac{1}{2\pi\hbar} \right) \left(\frac{\pi}{i\Delta t/2m\hbar} \right)^{1/2} e^{-\frac{(x-x_0)^2/\hbar^2}{4(i\Delta t/2m\hbar)}} \\ &= \frac{1}{2\pi\hbar} \left(\frac{2\pi\hbar m}{i\Delta t} \right)^{1/2} e^{-m(x-x_0)^2/2\hbar i\Delta t} \\ &= \left(\frac{m}{2\pi\hbar i\Delta t} \right)^{1/2} e^{-m(x-x_0)^2/2\hbar i\Delta t} \end{aligned}$$

That is,

$$U(x, t; x_0, t_0) = \left(\frac{m}{2\pi\hbar i(t-t_0)} \right)^{1/2} e^{-\frac{m(x-x_0)^2}{2\hbar i(t-t_0)}} \quad (12.116)$$

Which is the same result we obtained earlier by using the path integral formula.

12.6 The Classical Action

Suppose a particle propagates from x_0 at time t_0 to x at a later time t . Of all the conceivable paths from (x_0, t_0) to (x, t) , there is one path for which the action is minimum. This path is called the classical path and the minimum value of the action along the classical is called the classical action.

For a free particle, the classical path is the straight line connecting the points (t_0, x_0) to (t, x) in the space-time diagram.

Figure 12.3: Classical action

The equation for the classical path is then

$$x_{cl}(t') = x_0 + \frac{(x - x_0)}{(t - t_0)}(t' - t_0) \quad (12.117)$$

Here t and t_0 are fixed times and t' 's is the running variable. From the above equation we have

$$\dot{x}_{cl}(t') = \frac{x - x_0}{t - t_0} = constant \quad (12.118)$$

Therefore, the classical action for a free particle is (for a free particle $V = 0$ so we have $L = T - V = T = \frac{1}{2}m\dot{x}^2$)

$$S_{cl} = S[x_{cl}(t')] \quad (12.119)$$

$$= \int_{t_0}^t L(x_{cl}, \dot{x}_{cl}) dt' \quad (12.120)$$

$$= \int_{t_0}^t \frac{1}{2}m\dot{x}_{cl}(t') dt' \quad (12.121)$$

$$= \frac{1}{2}m \left(\frac{x - x_0}{t - t_0} \right)^2 (t - t_0) \quad (12.122)$$

$$= \frac{m(x - x_0)^2}{2(t - t_0)} \quad (12.123)$$

Now, the free propagator is

$$U(x, t; x_0, t_0) = \frac{m}{2\pi\hbar i(t - t_0)} e^{-\frac{m(x-x_0)^2}{2\hbar i(t-t_0)}} \quad (12.124)$$

In terms of classical action we can write

$$U(x, t; x_0, t_0) = \left(\frac{m}{2\pi\hbar i(t - t_0)} \right)^{1/2} e^{\frac{i}{\hbar} S_{cl}} \quad (12.125)$$

12.7 Discussion on Path Integral

Reference: Shankar

12.7.1 Principle of Least Action

If a particle moves from x_0 at time t_0 to a different point x at a later time t , then of all the paths between the points (x_0, t_0) to (x, t) , a classical particle takes the path for which the action is minimum. This is called the principle of least action

Figure 12.4: least action

Suppose a particle follows the path $x(t')$. Then the action for this path is

$$S[x((t'))] = \int_{t_0}^t L(x(t'), \dot{x}(t')) dt' \quad (12.126)$$

Next consider a slightly varied path

$$x(t') + \eta(t') \quad (12.127)$$

where η is very small and

$$\eta(t_0) = \eta(t) = 0 \quad (12.128)$$

Then the action for the varied path is

$$S[x(t') + \eta(t')] = \int_{t_0}^t L(x(t') + \eta(t'), \dot{x}(t') + \dot{\eta}(t')) dt' \quad (12.129)$$

Then upto first order in $\eta(t')$, the variation of the action is

$$\begin{aligned} \delta S[x(t')] &= S[x(t') + \eta(t')] - S[x(t')] \\ &= \int_{t_0}^t \left[\frac{\partial L(x, \dot{x})}{\partial x} \eta(t') + \frac{\partial L(x, \dot{x})}{\partial \dot{x}} \dot{\eta}(t') \right] dt' \\ &= \int_{t_0}^t \left[\frac{\partial L}{\partial x} \eta(t') + \frac{d}{dt'} \left(\frac{\partial L}{\partial \dot{x}} \eta(t') \right) - \frac{d}{dt'} \left(\frac{\partial L}{\partial \dot{x}} \right) \eta(t') \right] dt' \\ &= \int_{t_0}^t \left[\frac{\partial L}{\partial x} - \frac{d}{dt'} \left(\frac{\partial L}{\partial \dot{x}} \right) \right] \eta(t') dt' + \int_{t_0}^t \left[\frac{d}{dt'} \left(\frac{\partial L}{\partial \dot{x}} \eta(t') \right) \right] dt' \end{aligned}$$

The second term on the right hand side of the above equation is zero:

$$\int_{t_0}^t \frac{d}{dt'} \left(\frac{\partial L}{\partial \dot{x}} \eta(t') \right) dt' = \frac{\partial L}{\partial \dot{x}} \eta(t') \Big|_{t'=t_0}^{t'=t} = 0 \quad (12.130)$$

Since $\eta(t) = \eta(t') = 0$. Thus we have

$$\delta S[x(t')] = \int_{t_0}^t \left[\frac{\partial L}{\partial x} - \frac{d}{dt'} \left(\frac{\partial L}{\partial \dot{x}} \right) \right] \eta(t') dt' + \mathcal{O}(\eta^2) \quad (12.131)$$

Now, if the path is the classical path, i.e., $x(t') = x_{cl}(t')$, Then $\delta S = 0$ up to first order in η . Therefore we must have

$$\int_{t_0}^t \left[\frac{\partial L}{\partial x} - \frac{d}{dt'} \left(\frac{\partial L}{\partial \dot{x}} \right) \right]_{cl} \eta(t') dt' = 0 \quad (12.132)$$

Figure 12.5: paths

Since $\eta(t')$ is arbitrary except at the end times t_0 and t , we must have

$$\left[\frac{\partial L}{\partial x} - \frac{d}{dt'} \left(\frac{\partial L}{\partial \dot{x}} \right) \right]_{cl} = 0 \quad (12.133)$$

Thus the variation δS from the classical path is

$$\delta S = \mathcal{O}(\eta^2) \quad (12.134)$$

12.8 Discussion on The Phase of the Path Integral

We derived previously

$$U(x, t; x_0, t_0) = \int \mathcal{D}[x(t')] e^{\frac{i}{\hbar} S[x(t')]} \quad (12.135)$$

Every path contributes a phase factor in the path integral, the phase being $\frac{i}{\hbar} S[x(t')]$ where $x(t')$ is a particular path between the points (t_0, x_0) and (t, x) in the space-time diagram. Historically we can write

$$U = \sum_{\text{all paths}} e^{\frac{i}{\hbar} S[x(t')]} \quad (12.136)$$

The most surprising this about the path integral is that every path, including the classical path $x_{cl}(t')$ gets the same weight, that is to say a complex number of unit modulus.

Of all the paths, there is a special path, called the classical path for which S is minimum or stationary. A slight change in path from the classical one does not change the action, more precisely the change in action is only of second order in the change of path.

Consider a path $x_a(t')$ far away from the classical path. Its contribution to the path integral is

$$Z_a = e^{iS[x_a(t')]/\hbar} \quad (12.137)$$

While doing the path integral if we vary the path from $x_a(t')$ to a neighboring one, there will be slightly change in the action. But, there will be a large change in the phase S/\hbar , since \hbar is small. So, for paths well away from the classical path, contributions cancel because of the large change in phase from one path to the next. However, the situation is different for the classical path and the bundle of paths close to it. Here the action is stationary and so the phase of each of the paths near the classical path is about the same. In other words, the paths in the neighborhood of the classical

Figure 12.6: alternative paths

path contribute constructively to the path integral.

Thus the propagator U is dominated by the paths near the classical path. The classical path is important not because it contributes a lot by itself, but because the paths in the vicinity of the classical path contribute coherently.

How far from the classical path must we deviate before destructive interference sets in? One may say crudely that coherence would be lost once the phase differs from the stationary value $\frac{1}{\hbar}S[x_{cl}(t')]$ by about π , i.e., if the action changes from the classical action by about $\pi\hbar$. For a macroscopic particle this means a very tight constraint on its path since S_{cl} is typically of the order of $1 \text{ erg sec} \approx 10^{27}\hbar$ ^[1]. For a macroscopic particle, a slight change of the path from the classical would change the action by an amount much more than $\pi\hbar$. So, only the classical path contributes to the path integral. Therefore a macroscopic particle has a well defined path, namely the classical path.

For a microscopic particle like an electron, the action is much smaller. Hence for a large variation of the path from the classical one, the change of action remains less than $\pi\hbar$. It follows that a large number of widely varying paths around the classical path contributes coherently to the propagator U . Therefore one cannot say that a microscopic particle follows a definite path as it propagates from one point to another. There is a lot of — in the path that a microscopic particle can choose as it propagates between two fixed space time points.

Consider the following example. A free particle leaves the origin at $t = 0$ and arrives at $x = 1 \text{ cm}$ at $t = 1 \text{ sec}$, the classical path is

$$x_{cl}(t) = at \quad (12.138)$$

where a is a constant with value $a = 1 \text{ cm/sec}$. Choose another path

$$x(t) = bt^2 \quad (12.139)$$

where $b = 1 \text{ cm/sec}^2$.

We will now calculate the change in action for a macroscopic particle of mass $1g$ between these

^[1] $\hbar = 1.0546 \times 10^{-27} \text{ erg sec} = 1.0546 \times 10^{-34} \text{ Js}$

two paths. The action for the classical path is

$$\begin{aligned} S[x_{cl}] &= \int_0^1 \frac{1}{2} m \dot{x}_{cl}^2(t) dt \\ &= \frac{1}{2} m a^2 \times 1 \text{sec} \\ &= \frac{1}{2} \times (1g) \times (1\text{cm/sec})^2 \times (1\text{sec}) = 0.5 \text{erg sec} \end{aligned}$$

While the action for the alternative path is

$$\begin{aligned} S[x(t)] &= \int_0^1 \frac{1}{2} (2bt)^2 dt \\ &= 2b^2 m \int_0^1 t^2 dt \\ &= 2b^2 m \left(\frac{1}{3} \text{sec}^3 \right) \\ &= \frac{2(1\text{cm sec}^{-1})^2 (1g)(1\text{sec}^3)}{3} \\ &= 0.67 \text{ erg sec} \end{aligned}$$

Therefore $\Delta S = 0.17 \text{ erg sec} \approx 1.7 \times 10^{26} \hbar >> \pi \hbar$.

We can therefore ignore non classical paths for the macroscopic particle. On the other hand, for an electron whose mass is $m \approx 10^{-27} \text{g}$, the change in action is $\Delta S \approx \frac{1}{6} \hbar < \pi \hbar$ or the phase difference is $\Delta S/\hbar \approx \frac{1}{6} < \pi$. For the electron the classical path and a wide range of paths around the classical path would contribute to U . It is in such cases assuming that the particle moves in a well defined trajectory $x_{cl}(t)$, leads to conflict with experiment.

12.9 Equivalence to the Schrödinger Equation

In the Schrödinger formalism, the change in the state vector $|\psi\rangle$ over an infinitesimal time is (upto first order in ε)

$$|\psi(\varepsilon)\rangle - |\psi(0)\rangle = -\frac{i\varepsilon}{\hbar} H(t=0) |\psi(0)\rangle \quad (12.140)$$

which in the coordinate basis becomes

$$\psi(x, \varepsilon) - \psi(x, 0) = -\frac{i\varepsilon}{\hbar} \left[-\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} + V(x, 0) \right] \psi(x, 0) \quad (12.141)$$

To compare this result with the path integral prediction to the same order in ε , we begin with

$$\psi(x, \varepsilon) = \int_{-\infty}^{\infty} U(x, \varepsilon; x') \psi(x', 0) dx' \quad (12.142)$$

Figure 12.7: space-time diagram

The calculation of $U(\varepsilon)$ is simplified by the fact that there is no need to integrate over intermediate x 's since there is one slice of time between start to finish.

So,

$$U(x, \varepsilon; x') = \left(\frac{m}{2\pi\hbar i\varepsilon} \right)^{1/2} e^{\frac{i\varepsilon}{\hbar} \left(\frac{1}{2} m \dot{x}'^2 - V(x', 0) \right)} \quad (12.143)$$

Since ε is infinitesimal, we can assume that the path from x' to x in the space-time diagram is linear

So we can write in equation (12.143)

$$\dot{x}' = \frac{x - x'}{\varepsilon} \quad (12.144)$$

In equation (12.143) we can keep $V(x')$ as it is, or replace $V(x')$ by $V\left(\frac{x+x'}{2}\right)$. Replacing $V(x')$ by $V\left(\frac{x+x'}{2}\right)$ doesn't change the result in the first order of ε .

We can now write equation (12.143) as

$$U(x, \varepsilon; x') = \left(\frac{m}{2\pi\hbar i\varepsilon} \right)^{1/2} \exp \left(\frac{i}{\hbar} \left[\frac{m(x-x')^2}{2\varepsilon} - \varepsilon V\left(\frac{x+x'}{2}, 0\right) \right] \right) \quad (12.145)$$

If V is time dependent, we take the time argument of V to be zero since there is already a factor ε before it and any variation of V with time is the interval 0 to ε will produce an effect of second order in ε .

So, substituting equation (12.145) in euqation (12.142) we have

$$\psi(x, \varepsilon) = \left(\frac{m}{2\pi\hbar i\varepsilon} \right)^{1/2} \int_{-\infty}^{\infty} dx' \exp \left[\frac{im(x-x')^2}{2\hbar\varepsilon} \right] \exp \left[-\frac{i\varepsilon}{\hbar} V\left(\frac{x+x'}{2}, 0\right) \right] \psi(x', 0) \quad (12.146)$$

Consider the factor $\exp \left[\frac{im(x-x')^2}{2\hbar\varepsilon} \right]$. It oscillates very rapidly as $(x - x')$ varies since ε is infinitesimal and \hbar is very small. When such a rapidly oscillating function multiplies a smooth function like $\psi(x', 0)$, the integral vanishes for the most part due to the random phase of the exponential. Just as in the case of path integration, the only substantial contribution comes from the region where the phase is stationary. In this case only stationary point is at $x' = x$ where the phase has the minimum value of zero. In terms of $\eta = x' - x$, the region of coherence is, as before,

$$\frac{m\eta^2}{2\varepsilon\hbar} \leq \pi$$

$$|\eta| \leq \left(\frac{2\pi\hbar\varepsilon}{m} \right)^{1/2} \quad (12.147)$$

Consider now

$$\psi(x, \varepsilon) = \left(\frac{m}{2\pi\hbar i\varepsilon} \right)^{1/2} \int_{-\infty}^{\infty} d\eta \exp [im\eta^2/2\hbar\varepsilon] \exp \left[-\frac{i\varepsilon}{\hbar} V(x + \frac{\eta}{2}, 0) \right] \psi(x + \eta, 0) \quad (12.148)$$

We will work to first order in ε and therefore second order in η (see equation (12.147)). We expand

$$\psi(x + \eta, 0) = \psi(x, 0) + \eta \frac{\partial \psi}{\partial x} + \frac{\eta^2}{2} \frac{\partial^2 \psi}{\partial x^2} + \dots \quad (12.149)$$

and

$$\begin{aligned} \exp \left[-\frac{i\varepsilon}{\hbar} V(x + \frac{\eta}{2}, 0) \right] &= 1 - \frac{i\varepsilon}{\hbar} V(x + \frac{\eta}{2}, 0) + \dots \\ &= 1 - \frac{i\varepsilon}{\hbar} V(x, 0) + \dots \end{aligned}$$

Where the terms of the order of $\eta\varepsilon$ are neglected. Equation (12.148) now becomes

$$\psi(x, \varepsilon) = \left(\frac{m}{2\pi\hbar i\varepsilon} \right)^{1/2} \int_{-\infty}^{\infty} \exp \left[\frac{im\eta^2}{2\hbar\varepsilon} \right] \left[1 - \frac{i\varepsilon}{\hbar} V(x, 0) + \dots \right] \left[\psi(x, 0) + \eta \frac{\partial \psi}{\partial x} + \frac{\eta^2}{2} \frac{\partial^2 \psi}{\partial x^2} \right] d\eta \quad (12.150)$$

$$= \left(\frac{m}{2\pi\hbar i\varepsilon} \right)^{1/2} \int_{-\infty}^{\infty} \exp \left[\frac{im\eta^2}{2\hbar\varepsilon} \right] \left[\left(1 - \frac{i\varepsilon}{\hbar} V(x, 0) \right) \psi(x, 0) + \eta \frac{\partial \psi}{\partial x} + \frac{\eta^2}{2} \frac{\partial^2 \psi}{\partial x^2} \right] d\eta \quad (12.151)$$

where terms of the orders of $\varepsilon\eta$ and $\varepsilon\eta^2$ are neglected.

Using standard integrals (from appendix) (65) and (66) we get

$$\begin{aligned} \psi(x, \varepsilon) &= \left(\frac{m}{2\pi\hbar i\varepsilon} \right)^{1/2} \left[\left(\frac{2\pi\hbar i\varepsilon}{m} \right)^{1/2} qty \left(1 - \frac{i\varepsilon}{\hbar} V(x, 0) \right) \psi(x, 0) + \left(\frac{2\pi\hbar i\varepsilon}{m} \right)^{1/2} \frac{i\hbar\varepsilon}{2m} \frac{\partial^2 \psi}{\partial x^2} \right] \\ &= \psi(x, 0) - \frac{i\varepsilon}{\hbar} V(x, 0) \psi(x, 0) - \frac{i\varepsilon}{\hbar} \left(-\frac{\hbar^2}{2m} \right) \frac{\partial^2 \psi}{\partial x^2} \end{aligned}$$

Where we have used $a = \frac{m}{2\hbar i\varepsilon}$

$$\psi(x, \varepsilon) - \psi(x, 0) = -\frac{i\varepsilon}{\hbar} \left[-\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} + V(x, 0) \right] \psi(x, 0) \quad (12.152)$$

Which agrees with the Schrödinger prediction, equation (12.141)

12.10 Potentials of the Form $V = a + bx + cx^2 + d\dot{x} + e\dot{x}\dot{y}$

We wish to compute

$$U(x, t; x') = \int_{x'}^x e^{iS[x(t'')]/\hbar} \mathcal{D}[x(t'')] \quad (12.153)$$

Let us write every path as

$$x(t'') = x_{cl}(t'') + y(t'') \quad (12.154)$$

It follows that

$$\dot{x}(t'') = \dot{x}_{cl}(t'') + \dot{y}(t'') \quad (12.155)$$

Since all paths agree at the end points, $y(0) = y(t) = 0$. When we slice up the time into N parts, we have for intermediate integration variables

$$x_i = x(t''_i) = x_{cl}(t''_i) + y(t''_i) = x_{cl}(t''_i) \equiv x_{cl}(t''_i) + y_i \quad (12.156)$$

Since $x_{cl}(t'')$ is just some constant at t'' ,

$$dx_i = dy_i \quad (12.157)$$

and

$$\int_{x'}^x \mathcal{D}[x(t'')] = \int_0^0 \mathcal{D}[y(t'')] \quad (12.158)$$

Equation (12.153) becomes

$$U(x, t; x') = \int_0^0 \mathcal{D}[y(t'')] \exp \left[\frac{i}{\hbar} S[x_{cl}(t'') + y(t'')] \right] \quad (12.159)$$

The next step is to expand the functional S in a Taylor series about x_{cl}

$$\begin{aligned} S[x_{cl} + y] &= \int_0^t L(x_{cl} + y, \dot{x}_{cl} + \dot{y}) dt' \\ &= \int_0^t \left[L(x_{cl}, \dot{x}_{cl}) + \left(\frac{\partial L}{\partial x} \Big|_{x_{cl}} y + \frac{\partial L}{\partial \dot{x}} \Big|_{x_{cl}} \dot{y} \right) + \frac{1}{2} \left(\frac{\partial^2 L}{\partial x^2} \Big|_{x_{cl}} y^2 + 2 \frac{\partial^2 L}{\partial x \partial \dot{x}} \Big|_{x_{cl}} y \dot{y} + \frac{\partial^2 L}{\partial \dot{x}^2} \Big|_{x_{cl}} \dot{y}^2 \right) \right] dt'' \end{aligned}$$

The series terminates since L is a quadratic polynomial.

The first piece $L(x_{cl}, \dot{x}_{cl})$ integrates to give $S[x_{cl}] \equiv S_{cl}$. The second piece, linear in y and \dot{y} vanishes due to classical equation of motion.

To show this, first recall the classical equation of motion

$$\left[\frac{d}{dt''} \left(\frac{\partial L}{\partial \dot{x}} \right) - \frac{\partial L}{\partial x} \right]_{x_{cl}} = 0 \quad (12.160)$$

Therefore, the linear term in equation (??) can be written as

$$\frac{d}{dt''} \left(\frac{\partial L}{\partial \dot{x}} \right) \Big|_{x_{cl}} y + \frac{\partial L}{\partial \dot{x}} \Big|_{x_{cl}} \dot{y} = \frac{d}{dt''} \left(\frac{\partial L}{\partial \dot{x}} y \right) \quad (12.161)$$

This term integrated over t'' gives zero since

$$y(0) = y(t) = 0 \quad (12.162)$$

To calculate the final piece, note that

$$L = \frac{1}{2} m \dot{x}^2 - a - bx - cx^2 - d \dot{x} - ex \dot{x} \quad (12.163)$$

Hence

$$\frac{1}{2} \frac{\partial^2 L}{\partial x^2} = -c \quad (12.164)$$

$$\frac{\partial^2 L}{\partial x \partial \dot{x}} = -e \quad (12.165)$$

$$\frac{1}{2} \frac{\partial^2 L}{\partial \dot{x}^2} = \frac{m}{2} \quad (12.166)$$

Consequently

$$S[x_{cl}(t'') + y(t'')] = S_{cl} + \int_0^t \left(-cy^2 - ey\dot{y} + \frac{1}{2}m\dot{y}^2 \right) dt'' \quad (12.167)$$

Therefore equation (12.159) becomes

$$U(x, t; x') = \exp \left(\frac{iS_{cl}}{\hbar} \right) \int_0^0 \mathcal{D}[y(t'')] \exp \left[\frac{i}{\hbar} \int_0^t \left(\frac{1}{2}m\dot{y}^2 - cy^2 - ey\dot{y} \right) dt'' \right] \quad (12.168)$$

Since the path integral has no memory of x_{cl} , it can only depend on t . So

$$U(x, t; x') = A(t) \exp \left(\frac{iS_{cl}}{\hbar} \right) \quad (12.169)$$

Where

$$A(t) = \mathcal{D}[y(t'')] \int_0^0 \exp \left[\frac{i}{\hbar} \int_0^t \left(\frac{1}{2}m\dot{y}^2 - cy^2 - ey\dot{y} \right) dt'' \right] \quad (12.170)$$

12.10.1 Special Cases

1. Free Particle

Put $c = e = 0$ in the formula for $A(t)$. In this case we can calculate $A(t)$. We found previously

$$A(t) = \left(\frac{m}{2\pi\hbar i t} \right)^{1/2}$$

2. Harmonic Oscillator

For a harmonic oscillator

$$V(x) = \frac{1}{2}m\omega^2x^2 \quad (12.171)$$

So we set $c = \frac{1}{2}m\omega^2$ and all other coefficient are set to zero. Thus for harmonic oscillator we have

$$A(t) = \int_0^\infty \mathcal{D}[y(t'')] \exp \left[\frac{i}{\hbar} \int_0^t \left(\frac{1}{2}m\dot{y}^2 - \frac{1}{2}m\omega^2 y^2 \right) dt \right] \quad (12.172)$$

The evaluation of this integral is difficult. Note that even if the factor $A(t)$ in $\psi(x,t)$ is not known, we can extract all the probabilistic interpretation at time t .



13. sheet-13 : Potential Wells and Barriers



14. sheet-14 : Particle Motion in 3-Dimension



15. sheet-15 : The Hydrogen Atom



16. sheet-16 : Angular Momentum 1

Lecture 16 Angular Momentum 1

We know the important role played by angular momentum in classical mechanics. The total angular momentum of an isolated physical system is a constant of motion. Furthermore, this is also true in certain cases in which the system is not isolated. For example, if a point particle P, of mass m , is moving in a central potential (one which depends only on the distance between P and a fixed point O), the force to which P is subjected is always directed towards O. The torque about O is consequently zero, and the angular momentum theorem implies that

$$\frac{d\vec{L}}{dt} = 0, \quad (16.1)$$

where \vec{L} is the angular momentum of P with respect to O. This fact has important consequences: the motion of the particle P is confined to a fixed plane passing through the origin O and perpendicular to the angular momentum vector \vec{L} . Moreover, this motion obeys the law of constant aerial velocity (Kepler's second law).

All these properties have their equivalents in quantum mechanics. With the angular momentum \vec{L} of a classical system is associated an observable $\hat{\vec{L}}$, actually a set of three observables \hat{L}_x , \hat{L}_y and \hat{L}_z , which corresponds to the three components of \vec{L} in a Cartesian frame. If the physical system under consideration is a point particle moving under a central potential, then \hat{L}_x , \hat{L}_y and \hat{L}_z are constants of motion in the quantum mechanical sense, that is they commute with the Hamiltonian \hat{H} describing

the particle in the central potential $V(r)$.

We shall denote by the term orbital angular momentum any angular momentum which has a classical analog and $\hat{\vec{L}}$ will denote the corresponding observable. Spin angular momentum will denote the intrinsic angular momentum of an elementary particle. In a complex system, such as an atom, a nucleus or a molecule, the orbital angular momenta $\hat{\vec{L}}_i$ of the various elementary particles of the system combine with each other and with the spin angular momenta $\hat{\vec{S}}_i$ of these same particles to form the total angular momentum $\hat{\vec{J}}$ of the system.

Finally, let us add that $\hat{\vec{J}}$ will also be used to denote an arbitrary angular momentum, when it is not necessary to specify whether we are dealing with an orbital angular momentum, a spin, or a combination of several angular momenta.

16.1 Commutation Relations of Angular Momentum Operators

In classical mechanics, the orbital angular momentum of a particle is defined as

$$\vec{L} = \vec{r} \times \vec{p}. \quad (16.2)$$

The corresponding operator $\hat{\vec{L}}$ in quantum mechanics is obtained by replacing \vec{r} and \vec{p} by their corresponding operators $\hat{\vec{r}}$ and $\hat{\vec{p}}$. Thus

$$\hat{\vec{L}} = \hat{\vec{r}} \times \hat{\vec{p}}, \quad (16.3)$$

i.e.,

$$\begin{aligned}\hat{L}_x &= \hat{y}\hat{p}_z - \hat{z}\hat{p}_y \\ \hat{L}_y &= \hat{z}\hat{p}_x - \hat{x}\hat{p}_z \\ \hat{L}_z &= \hat{x}\hat{p}_y - \hat{y}\hat{p}_x.\end{aligned} \quad (16.4)$$

In the coordinate representation, the position operators are just multiplicative operators and $\vec{p} \rightarrow -i\hbar\vec{\nabla}$. Therefore, in the coordinate representation, we get

$$\begin{aligned}\hat{L}_x &= -i\hbar \left(y \frac{\partial}{\partial z} - z \frac{\partial}{\partial y} \right) \\ \hat{L}_y &= -i\hbar \left(z \frac{\partial}{\partial x} - x \frac{\partial}{\partial z} \right) \\ \hat{L}_z &= -i\hbar \left(x \frac{\partial}{\partial y} - y \frac{\partial}{\partial x} \right).\end{aligned} \quad (16.5)$$

$$(16.6)$$

Equation (16.4) can be compactly written as

$$\hat{L}_i = \epsilon_{ijk} \hat{x}_j \hat{p}_k; i = 1, 2, 3 \quad (16.7)$$

where repeated indices are summed and ϵ_{ijk} is defined as

$$\epsilon_{ijk} = \begin{cases} 1 & \text{if } ijk \text{ is an even permutation of 123} \\ -1 & \text{if } ijk \text{ is an odd permutation of 123} \\ 0 & \text{otherwise.} \end{cases} \quad (16.8)$$

Now, we have the fundamental commutation relations between the position operators and their conjugate momentum operators:

$$[\hat{x}_i, \hat{x}_j] = [\hat{p}_i, \hat{p}_j] = 0, \quad [\hat{x}_i, \hat{p}_j] = i\hbar \delta_{ij}. \quad (16.9)$$

Using the fundamental commutation relations, we can work out the commutation relations between the components of the angular momentum operator. We obtain .

$$[\hat{L}_i, \hat{L}_j] = i\hbar \epsilon_{ijk} \hat{L}_k. \quad (16.10)$$

Equation (16.10) can be written in a more elaborate form as

$$[\hat{L}_x, \hat{L}_y] = i\hbar \hat{L}_z, \quad [\hat{L}_y, \hat{L}_z] = i\hbar \hat{L}_x, \quad [\hat{L}_z, \hat{L}_x] = i\hbar \hat{L}_y \quad (16.11)$$

The commutation relations (16.10) and (16.11) can also be written in the following alternative form:

$$\hat{\vec{L}} \times \hat{\vec{L}} = i\hbar \hat{\vec{L}}. \quad (16.12)$$

Next, we can also easily prove

$$[\hat{L}_i, \hat{x}_j] = i\hbar \epsilon_{ijk} \hat{x}_k \quad (16.13)$$

and

$$[\hat{L}_i, \hat{p}_j] = i\hbar \epsilon_{ijk} \hat{p}_k. \quad (16.14)$$

Now, we define the operator corresponding to the square of the angular momentum as

$$\hat{L}^2 = \hat{L}_x^2 + \hat{L}_y^2 + \hat{L}_z^2. \quad (16.15)$$

It can be shown that \hat{L}^2 commutes with all the components \hat{L}_x , \hat{L}_y and \hat{L}_z , i.e.,

$$[\hat{L}^2, \hat{L}_i] = 0; i = 1, 2, 3 \quad (16.16)$$

i.e.,

$$[\hat{L}^2, \hat{L}] = 0. \quad (16.17)$$

As will be shown later, the commutation relations between the components of the angular momentum operator determine the quantum properties of angular momentum. That is, eigenvalues and eigenvectors of the angular momentum operators are completely determined by the commutation relations and the general properties of the Hilbert space. Therefore, the commutation relations themselves are taken as the definition of angular momentum operators in quantum mechanics.

16.2 Definition of Angular Momentum Operators

A vector operator $\hat{\vec{J}}$ having three components \hat{J}_x , \hat{J}_y and \hat{J}_z is called an angular momentum operator if the components are observables (and hence Hermitian) and obey the commutation relations

$$[\hat{J}_x, \hat{J}_y] = i\hbar\hat{J}_z, \quad [\hat{J}_y, \hat{J}_z] = i\hbar\hat{J}_x, \quad [\hat{J}_z, \hat{J}_x] = i\hbar\hat{J}_y. \quad (16.18)$$

This definition enables us to treat entities which have no classical analog, such as spin, on the same footing as orbital angular momentum.

16.2.1 Square of Angular Momentum

We introduce the operator

$$\hat{J}^2 = \hat{J}_x^2 + \hat{J}_y^2 + \hat{J}_z^2, \quad (16.19)$$

the square of the angular momentum operator $\hat{\vec{J}}$. This operator is Hermitian since \hat{J}_x , \hat{J}_y and \hat{J}_z are Hermitian. Using Eqs. (16.18) we can show that

$$[\hat{J}^2, \hat{J}] = 0 \quad (16.20)$$

i.e., the square of the angular momentum operator commutes with any of the three operators corresponding to the three components.

Angular momentum theory in quantum mechanics is founded entirely on the commutation relations (16.18). Note that, the commutation relations imply that it is impossible to have a state in which the three components of an angular momentum operator have definite values, i.e., they are incompatible observables. However, \hat{J}^2 and any component of $\hat{\vec{J}}$ are compatible.

16.2.2 Angular Momentum of a System of Particles

Let us consider a system of particles. The particles are labeled by Greek indices; $\alpha, \beta = 1, 2, 3, \dots, N$, where N is the number of particles of the system. The angular momentum operators referring to different particles commute, i.e.,

$$[\hat{J}_\alpha, \hat{J}_\beta] = 0, \quad \alpha \neq \beta; \quad (16.21)$$

where α and β refer to particles. The operator corresponding to the total angular momentum of the system of particles is given by

$$\hat{\vec{J}} = \sum_{\alpha=1}^N \hat{J}_\alpha. \quad (16.22)$$

It is easy to verify that $\hat{\vec{J}}$ is an angular momentum operator. We proceed as follows:

$$\begin{aligned} \hat{\vec{J}} \times \hat{\vec{J}} &= \sum_{\alpha} \sum_{\beta} (\hat{J}_\alpha \times \hat{J}_\beta) \\ &= \sum_{\alpha} \sum_{\beta} i\hbar \hat{J}_\alpha \delta_{\alpha\beta} \\ &= \sum_{\alpha} i\hbar \hat{J}_\alpha \\ &= i\hbar \hat{\vec{J}}, \end{aligned} \quad (16.23)$$

i.e., $\hat{\vec{J}}$ is indeed an angular momentum operator.

16.3 Eigenvalue Spectrum of \hat{J}^2 and \hat{J}_z

The components of angular momentum operator do not commute with each other; therefore, they do not have simultaneous eigenvectors. In other words, we cannot find a basis set for which all the components of the angular momentum operator are diagonal. However, since \hat{J}^2 commutes with $\hat{\vec{J}}$, we can find simultaneous eigenvectors for \hat{J}^2 and any one of the components, say \hat{J}_z . Let $|\lambda m\rangle$ be such an eigenvector such that

$$\hat{J}^2 |\lambda m\rangle = \lambda \hbar^2 |\lambda m\rangle \quad (16.24)$$

$$\hat{J}_z |\lambda m\rangle = m\hbar |\lambda m\rangle. \quad (16.25)$$

Here λ labels the eigenvalues of \hat{J}^2 and m those of \hat{J}_z . Now $\hat{J}^2 = \hat{J}_x^2 + \hat{J}_y^2 + \hat{J}_z^2$, being the sum of squares of Hermitian operators, is positive semi-definite, i.e., the eigenvalues λ of \hat{J}^2 are non-negative:

$$\lambda = \frac{\langle \lambda m | \hat{J}^2 | \lambda m \rangle}{\hbar^2} \geq 0. \quad (16.26)$$

Furthermore,

$$\langle \lambda m | \hat{J}^2 | \lambda m \rangle \equiv \langle \hat{J}^2 \rangle = \langle \hat{J}_x^2 \rangle + \langle \hat{J}_y^2 \rangle + \langle \hat{J}_z^2 \rangle \geq \langle \hat{J}_z^2 \rangle \quad (16.27)$$

i.e.,

$$\lambda \geq m^2. \quad (16.28)$$

It is convenient at this stage to introduce two non-Hermitian operators \hat{J}_+ and \hat{J}_- defined by

$$\hat{J}_{\pm} = \hat{J}_x \pm i\hat{J}_y. \quad (16.29)$$

We can now derive the following commutation relations involving \hat{J}_{\pm} :

$$[\hat{J}^2, \hat{J}_{\pm}] = 0 \quad \text{and} \quad [\hat{J}_z, \hat{J}_{\pm}] = \pm \hbar \hat{J}_{\pm}. \quad (16.30)$$

Furthermore, we also have the following useful relations:

$$\hat{J}_- \hat{J}_+ = \hat{J}^2 - \hat{J}_z^2 - \hbar \hat{J}_z \quad (16.31)$$

and

$$\hat{J}_+ \hat{J}_- = \hat{J}^2 - \hat{J}_z^2 + \hbar \hat{J}_z. \quad (16.32)$$

Now, since \hat{J}^2 commutes with both \hat{J}_+ and \hat{J}_- , it follows immediately that

$$\hat{J}^2 (\hat{J}_{\pm} |\lambda m\rangle) = \lambda \hbar^2 (\hat{J}_{\pm} |\lambda m\rangle), \quad (16.33)$$

that is, if $|\lambda m\rangle$ is an eigenket of \hat{J}^2 with eigenvalue $\lambda \hbar^2$, then $\hat{J}_{\pm} |\lambda m\rangle$ are also eigenvectors of \hat{J}^2 with the same eigenvalue.

Next, using the second of the commutation relations in Eq. (16.30), we obtain

$$\begin{aligned} \hat{J}_z (\hat{J}_{\pm} |\lambda m\rangle) &= (\hat{J}_{\pm} \hat{J}_z \pm \hbar \hat{J}_{\pm}) |\lambda m\rangle \\ &= (m \pm 1) \hbar (\hat{J}_{\pm} |\lambda m\rangle). \end{aligned} \quad (16.34)$$

Thus, $\hat{J}_{\pm} |\lambda m\rangle$ are eigenvectors of \hat{J}_z with eigenvalues $(m \pm 1) \hbar$. The operator \hat{J}_+ is therefore called the raising operator, while \hat{J}_- is called the lowering operator.

Now, by successively applying the raising operator \hat{J}_+ on the state $|\lambda m\rangle$ we can increase the value of m in steps of one, keeping λ fixed. However, this process cannot go on indefinitely, for, otherwise, the inequality (16.28), i.e., $\lambda \geq m^2$, will be violated. Thus, there exists a maximum value of m for a given λ . Call this maximum value j . Similarly, by successively applying \hat{J}_- on the state $|\lambda m\rangle$ we can lower the value of m in steps of one. Again this process cannot be continued indefinitely since we cannot violate the inequality (16.28). Therefore, we must have a minimum value of m . Call this minimum value j' .

From the above discussions it is apparent that, starting from the ket $|\lambda j\rangle$, we can reach the ket $|\lambda j'\rangle$ by successive application of the lowering operator \hat{J}_- . Thus, it follows that

$$j - j' = \text{positive integer or zero.} \quad (16.35)$$

Next, we will determine the values of j and j' for a given λ . Since j is the greatest value of m , application of the raising operator \hat{J}_+ to the eigenstate $|\lambda j\rangle$ should not lead to a new eigenket. We must therefore have

$$\hat{J}_+ |\lambda j\rangle = 0. \quad (16.36)$$

Similarly, since j' is the minimum value of m ,

$$\hat{J}_- |\lambda j'\rangle = 0. \quad (16.37)$$

As the next step we apply \hat{J}_- on Eq. (16.36) to get

$$\hat{J}_- \hat{J}_+ |\lambda j\rangle = 0,$$

i.e.,

$$(\hat{J}^2 - \hat{J}_z^2 - \hbar \hat{J}_z) |\lambda j\rangle = 0,$$

or

$$(\lambda - j^2 - j) \hbar^2 |\lambda j\rangle = 0. \quad (16.38)$$

Hence

$$\lambda - j^2 - j = 0,$$

i.e.,

$$\lambda = j(j+1). \quad (16.39)$$

Similarly, by applying \hat{J}_+ on Eq. (16.37), we obtain

$$\hat{J}_+ \hat{J}_- |\lambda j'\rangle = 0,$$

i.e.,

$$(\hat{J}^2 - \hat{J}_z^2 + \hbar \hat{J}_z) |\lambda j'\rangle = 0,$$

or

$$(\lambda - j'^2 + j')\hbar^2|\lambda j'\rangle = 0. \quad (16.40)$$

Hence

$$\lambda - j'^2 + j' = 0,$$

i.e.,

$$\lambda = j'(j' - 1). \quad (16.41)$$

Comparing Eqs. (16.39) and (16.41) we can write

$$j(j+1) = j'(j'-1).$$

Thus, either $j' = -j$ or $j' = j+1$. The second solution is meaningless since j is the greatest value of m . Hence, the only possibility is that

$$j' = -j. \quad (16.42)$$

Using Eq. (16.42) in Eq. (16.35) we obtain

$$2j = \text{positive integer or zero},$$

i.e.,

$$j = 0, \frac{1}{2}, 1, \frac{3}{2}, 2, \dots \quad (16.43)$$

In summary, we have shown that if $|\lambda m\rangle$ is a simultaneous eigenket of \hat{J}^2 and \hat{J}_z with eigenvalues $\lambda\hbar^2$ and $m\hbar$, respectively, then λ can be expressed as $j(j+1)$ and the possible values of j are given as in Eq. (16.43). Furthermore, for a given j , m can assume values from a maximum of j to a minimum of $-j$ in steps of unity, i.e.,

$$m = -j, -j+1, -j+2, \dots, j-1, j. \quad (16.44)$$

It is customary to label the simultaneous eigenkets of \hat{J}^2 and \hat{J}_z by j and m rather than by λ and m . Thus the eigenkets are written as $|jm\rangle$ and the eigenvalue equations are

$$\hat{J}^2|jm\rangle = j(j+1)\hbar^2|jm\rangle, \quad (16.45)$$

and

$$\hat{J}_z|jm\rangle = m\hbar|jm\rangle. \quad (16.46)$$

For a given j , there are $(2j+1)$ eigenkets with m varying from $-j$ to j in steps of one.

16.4 Matrix Elements of Angular Momentum Operators in the $|jm\rangle$ Basis

For a given value of j , we can form $(2j+1)$ linearly independent angular momentum states since m can vary from $-j$ to j in steps of 1, i.e.,

$$m = -j, -j+1, \dots, j-1, j.$$

Thus in the basis $\{|jm\rangle, (j \text{ fixed})\}$ all the operators will be $(2j+1) \times (2j+1)$ matrices.

The matrices for \hat{J}^2 and \hat{J}_z will be diagonal, since the basis kets are simultaneous eigenkets of these two operators. Thus

$$(\hat{J}^2)_{mm'}^{(j)} = \langle jm | \hat{J}^2 | jm' \rangle = j(j+1)\hbar^2 \delta_{mm'} \quad (16.47)$$

and

$$(\hat{J}_z)_{mm'}^{(j)} = \langle jm | \hat{J}_z | jm' \rangle = m\hbar \delta_{mm'} \quad (16.48)$$

To find the matrices for \hat{J}_x and \hat{J}_y , it is convenient to find the matrices for \hat{J}_+ and \hat{J}_- first. To start, we note that

$$\hat{J}_+ |jm\rangle \propto |jm+1\rangle$$

$$\hat{J}_- |jm\rangle \propto |jm-1\rangle$$

Therefore, we write

$$\hat{J}_+ |jm\rangle = c_+ |jm+1\rangle \quad (16.49)$$

and

$$\hat{J}_- |jm\rangle = c_- |jm-1\rangle \quad (16.50)$$

where the constants c_+ and c_- are chosen so that all states $|jm\rangle$, $|jm+1\rangle$ and $|jm-1\rangle$ are normalized.

Now the adjoints of Eqs. (16.49) and (16.50) are written as

$$\langle jm | \hat{J}_- = c_+^* \langle jm+1 | \quad (16.51)$$

and

$$\langle jm | \hat{J}_+ = c_-^* \langle jm-1 |, \quad (16.52)$$

since

$$\hat{J}_+^\dagger = \hat{J}_- \quad \text{and} \quad \hat{J}_-^\dagger = \hat{J}_+.$$

Now, combining Eqs. (16.49) and (16.51) we have

$$\langle jm|\hat{J}_-\hat{J}_+|jm\rangle = |c_+|^2 \langle jm+1|jm+1\rangle,$$

or,

$$\langle jm|\hat{J}^2 - \hat{J}_z^2 - \hbar\hat{J}_z|jm\rangle = |c_+|^2 \langle jm+1|jm+1\rangle,$$

or,

$$\hbar^2 (j(j+1) - m^2 - m) \langle jm|jm\rangle = |c_+|^2 \langle jm+1|jm+1\rangle.$$

Since the states $|jm\rangle$ and $|jm+1\rangle$ are normalized, we have

$$\hbar^2 (j(j+1) - m^2 - m) = |c_+|^2.$$

If c_+ is taken to be real and positive, we have

$$c_+ = \sqrt{j(j+1) - m(m+1)} \hbar$$

or,

$$c_+ = \sqrt{(j-m)(j+m+1)} \hbar. \quad (16.53)$$

Similarly, combining eqs. (16.50) and (16.52) we have

$$\langle jm|\hat{J}_+\hat{J}_-|jm\rangle = |c_-|^2 \langle jm-1|jm-1\rangle,$$

or,

$$\langle jm|\hat{J}^2 - \hat{J}_z^2 + \hbar\hat{J}_z|jm\rangle = |c_-|^2$$

or,

$$|c_-|^2 = (j(j+1) - m(m-1)) \hbar^2.$$

Taking c_- to be real and positive, we have

$$c_- = \sqrt{j(j+1) - m(m-1)} \hbar,$$

or,

$$c_- = \sqrt{(j+m)(j-m+1)} \hbar. \quad (16.54)$$

Summarizing, we have shown

$$\hat{J}_+|jm\rangle = \sqrt{(j-m)(j+m+1)} \hbar |jm+1\rangle \quad (16.55)$$

$$\hat{J}_-|jm\rangle = \sqrt{(j+m)(j-m+1)} \hbar |jm-1\rangle. \quad (16.56)$$

Note that

$$\hat{J}_+|jj\rangle = 0$$

and

$$\hat{J}_-|j-j\rangle = 0.$$

Using Eqs. (16.55) and (16.56) we can find the matrix elements of \hat{J}_+ and \hat{J}_- in the $|jm\rangle$ representation with a fixed value of j . Thus

$$\begin{aligned} (\hat{J}_+)^{(j)}_{mm'} &= \langle jm|\hat{J}_+|jm'\rangle \\ &= \sqrt{(j-m')(j+m'+1)} \hbar \langle jm|jm'+1\rangle \\ &= \sqrt{(j-m')(j+m'+1)} \hbar \delta_{m,m'+1}. \end{aligned} \quad (16.57)$$

Similarly,

$$\begin{aligned} (\hat{J}_-)^{(j)}_{mm'} &= \langle jm|\hat{J}_-|jm'\rangle \\ &= \sqrt{(j+m')(j-m'+1)} \hbar \langle jm|jm'-1\rangle \\ &= \sqrt{(j+m')(j-m'+1)} \hbar \delta_{m,m'-1}. \end{aligned} \quad (16.58)$$

The matrices for \hat{J}_x and \hat{J}_y can be obtained from the relations

$$\hat{J}_x = \frac{1}{2}(\hat{J}_+ + \hat{J}_-), \quad (16.59)$$

and

$$\hat{J}_y = \frac{1}{2i}(\hat{J}_+ - \hat{J}_-). \quad (16.60)$$

Now, let us consider a few examples with definite values of j .

16.4.1 $j = 1/2$ Representation

Suppose $j = 1/2$, then $m = \pm 1/2$ and, therefore, the dimension of the representation is 2×2 . The basis states are $|1/2\ 1/2\rangle$ and $|1/2\ -1/2\rangle$. In the $j = 1/2$ representation we have (with the basis taken in the order $|1/2\ 1/2\rangle$, $|1/2\ -1/2\rangle$):

$$\underline{\hat{J}}^2 = \frac{3}{4} \hbar^2 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad (16.61)$$

and

$$\hat{J}_z = \hbar \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (16.62)$$

We also have

$$\hat{J}_+ = \hbar \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad (16.63)$$

and

$$\hat{J}_- = \hbar \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}. \quad (16.64)$$

Therefore

$$\hat{J}_x = \frac{\hat{J}_+ + \hat{J}_-}{2} = \frac{1}{2} \hbar \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad (16.65)$$

and

$$\hat{J}_y = \frac{\hat{J}_+ - \hat{J}_-}{2i} = \frac{1}{2} \hbar \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}. \quad (16.66)$$

In a more compact notation, we can write

$$\hat{J}_i \doteq \frac{1}{2} \hbar \sigma_i, \quad i = 1, 2, 3 \quad (16.67)$$

where

$$\sigma_1 = \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad (16.68)$$

$$\sigma_2 = \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad (16.69)$$

$$\sigma_3 = \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad (16.70)$$

are the Pauli matrices.

16.4.2 $j = 1$ Representation

In this case $j = 1$ and the basis states are:

$$\{|j=1, m\rangle, m=1, 0, -1\}.$$

Taking the basis states in the order $|11\rangle$, $|10\rangle$, $|1-1\rangle$, the matrices corresponding to \hat{J}^2 and \hat{J}_z are diagonal:

$$\hat{J}^2 \doteq 2\hbar^2 \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad (16.71)$$

$$\hat{J}_z \doteq \hbar \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \quad (16.72)$$

The matrix representation of other operators are

$$\hat{J}_+ \doteq \sqrt{2}\hbar \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \quad (16.73)$$

$$\hat{J}_- \doteq \sqrt{2}\hbar \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \quad (16.74)$$

$$\hat{J}_x = \frac{1}{2}(\hat{J}_+ + \hat{J}_-) \doteq \frac{\hbar}{\sqrt{2}} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad (16.75)$$

$$\hat{J}_y = \frac{1}{2i}(\hat{J}_+ - \hat{J}_-) \doteq \frac{\hbar}{\sqrt{2}} \begin{pmatrix} 0 & -i & 0 \\ i & 0 & -i \\ 0 & i & 0 \end{pmatrix}. \quad (16.76)$$

In this way we can write down the matrix representations of the various angular momentum operators corresponding to any value of j .

END



17. sheet-17 : Angular Momentum 2: spin 1/2

Lecture 17 Angular Momentum 2 Spin-1/2 System

17.1 The Spin Operators and Spin States of a Spin-1/2 Particle

The electron has an intrinsic angular momentum (i.e., spin angular momentum) in addition to orbital angular momentum. Let the operators corresponding to spin angular momentum of the electron be denoted by \hat{S}_x , \hat{S}_y and \hat{S}_z . They satisfy the commutation relations

$$[\hat{S}_i, \hat{S}_j] = i\hbar \epsilon_{ijk} \hat{S}_k. \quad (17.1)$$

We can find the common eigenstates of \hat{S}^2 and any one of the components, \hat{S}_z , say. These eigenstates are denoted by $|sm_s\rangle$ where

$$\hat{S}^2|sm_s\rangle = s(s+1)\hbar^2|sm_s\rangle, \quad (17.2)$$

and

$$\hat{S}_z|sm_s\rangle = m_s\hbar|sm_s\rangle. \quad (17.3)$$

The spin eigenstates $|sm_s\rangle$, $m_s = s, s-1, \dots, -s$, can be taken as the basis states in constructing a general spin state.

The quantity s is called the spin quantum number (or simply the spin) and the quantity m_s is called the projection quantum number, the projection direction being the z -axis. All elementary particles have a definite value for s which is always fixed. For electrons, as for all leptons and quarks, $s = 1/2$. This is why electrons are called a spin-half particles. If $s = 1/2$, the projection quantum number can have only two values, namely $m_s = \pm 1/2$. Therefore, there are only two basis spin states for a spin-1/2 particle. They can be denoted by

$$|1/2\ 1/2\rangle \quad \text{and} \quad |1/2 - 1/2\rangle,$$

or by

$$|S_z +\rangle \quad \text{and} \quad |S_z -\rangle,$$

or simply

$$|+\rangle \quad \text{and} \quad |-\rangle.$$

For a spin-1/2 particle, the eigenvalue equations are:

$$\hat{S}^2|\pm\rangle = \frac{1}{2}\left(\frac{1}{2}+1\right)\hbar^2|\pm\rangle = \frac{3}{4}\hbar^2|\pm\rangle \quad (17.4)$$

and

$$\hat{S}_z|\pm\rangle = \pm\frac{1}{2}\hbar|\pm\rangle. \quad (17.5)$$

Using the states $\{|+\rangle, |-\rangle\}$ as the basis set, any general spin state $|\alpha\rangle$ can be expressed as

$$|\alpha\rangle = |+\rangle\langle +|\alpha\rangle + |-\rangle\langle -|\alpha\rangle, \quad (17.6)$$

or,

$$|\alpha\rangle = c_+|+\rangle + c_-|-\rangle, \quad (17.7)$$

where

$$c_+ = \langle +|\alpha\rangle, \quad \text{and} \quad c_- = \langle -|\alpha\rangle, \quad (17.8)$$

are the “components” of $|\alpha\rangle$ along $|+\rangle$ and $|-\rangle$, respectively. The matrix representation of the general spin state $|\alpha\rangle$ in the basis $\{|+\rangle, |-\rangle\}$ is then

$$|\alpha\rangle \doteq \begin{pmatrix} c_+ \\ c_- \end{pmatrix} \equiv \chi. \quad (17.9)$$

The two-component column matrix (Eq. (17.9)) is called the spinor corresponding to the ket $|\alpha\rangle$ and is denoted by χ . The spinor corresponding to the bra $\langle\alpha|$ is

$$\langle\alpha| \doteq \chi^\dagger = \begin{pmatrix} c_+^* & c_-^* \end{pmatrix}. \quad (17.10)$$

Now we will find the matrix representations of the operators \hat{S}^2 , \hat{S}_x , \hat{S}_y and \hat{S}_z in the basis $\{|+\rangle, |-\rangle\}$. Since the basis set $|+\rangle$ and $|-\rangle$ are eigenstates of \hat{S}^2 and \hat{S}_z , it is obvious that in the given basis \hat{S}^2 and \hat{S}_z are diagonal. Thus

$$\hat{S}^2 \doteq \begin{pmatrix} \langle +|\hat{S}^2|+ \rangle & \langle +|\hat{S}^2|- \rangle \\ \langle -|\hat{S}^2|+ \rangle & \langle -|\hat{S}^2|- \rangle \end{pmatrix}, \quad (17.11)$$

i.e.,

$$\hat{S}^2 \doteq \frac{3}{4}\hbar^2 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \quad (17.12)$$

Similarly,

$$\hat{S}_z \doteq \frac{1}{2}\hbar \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (17.13)$$

To find the matrix representations of \hat{S}_x and \hat{S}_y in the $\{|+\rangle, |-\rangle\}$ basis, we first introduce two non-Hermitian operators, \hat{S}_{\pm} , called raising and lowering operators, and defined by

$$\hat{S}_{\pm} = \hat{S}_x \pm i\hat{S}_y. \quad (17.14)$$

Using Eq. (17.1) we can easily show that

$$[\hat{S}_z, \hat{S}_{\pm}] = \pm\hbar\hat{S}_z. \quad (17.15)$$

In other words, Eq. (17.15) tells us that the effect of \hat{S}_{\pm} when they act on an eigenstate of \hat{S}_z is to give another eigenstate of \hat{S}_z with eigenvalue one unit more (\hat{S}_+) or one unit less (\hat{S}_-). We then have

$$\hat{S}_+|+\rangle = 0, \quad \text{and} \quad \hat{S}_+|-\rangle = \hbar|+\rangle, \quad (17.16)$$

and

$$\hat{S}_-|+\rangle = \hbar|-\rangle, \quad \text{and} \quad \hat{S}_-|-\rangle = 0. \quad (17.17)$$

We can now write down the matrix representation of \hat{S}_+ and \hat{S}_- in the $\{|+\rangle, |-\rangle\}$ basis. We have

$$\hat{S}_+ \doteq \hbar \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad (17.18)$$

$$\hat{S}_- \doteq \hbar \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}. \quad (17.19)$$

Next, using Eq. (17.14) we obtain

$$\hat{S}_x = \frac{1}{2} (\hat{S}_+ + \hat{S}_-) \doteq \frac{1}{2} \hbar \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad (17.20)$$

and

$$\hat{S}_y = \frac{1}{2i} (\hat{S}_+ - \hat{S}_-) \doteq \frac{1}{2} \hbar \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}. \quad (17.21)$$

In summary, we have obtained

$$\hat{S}_x \doteq \frac{1}{2} \hbar \sigma_x, \quad \hat{S}_y \doteq \frac{1}{2} \hbar \sigma_y, \quad \hat{S}_z \doteq \frac{1}{2} \hbar \sigma_z, \quad (17.22)$$

where σ_x , σ_y and σ_z are 2×2 matrices given by

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (17.23)$$

The matrices σ_x , σ_y and σ_z are called Pauli matrices.

17.1.1 Properties of Pauli Matrices

In the following we list some properties of Pauli matrices.

1. $\det(\sigma_j) = -1$, $j = 1, 2, 3$.
2. $\text{Tr}(\sigma_j) = 0$, $j = 1, 2, 3$.
3. $\sigma_x^2 = \sigma_y^2 = \sigma_z^2 = I$ where I is the 2×2 unit matrix.
4. $\sigma_x \sigma_y = -\sigma_y \sigma_x = i \sigma_z$, (x, y, z cyclic)

It follows that $\{\sigma_x, \sigma_y\} = 0$ and $[\sigma_x, \sigma_y] = 2i\sigma_z$, i.e., $[S_x, S_y] = i\hbar S_z$.

5. Property 3 and 4 are sometime condensed in the form

$$\sigma_j \sigma_k = i \epsilon_{jkl} \sigma_l + \delta_{jk} I_{2 \times 2},$$

where $I_{2 \times 2}$ is the two-dimensional unit matrix.

6. We can also easily show

$$\sigma_x \sigma_y \sigma_z = i I_{2 \times 2}.$$

Finally, we will prove an identity which is very useful. If \vec{A} and \vec{B} are two vectors whose components are numbers (or operators which commute with all spin operators), then

$$(\vec{\sigma} \cdot \vec{A})(\vec{\sigma} \cdot \vec{B}) = \vec{A} \cdot \vec{B} I_{2 \times 2} + i \vec{\sigma} \cdot (\vec{A} \times \vec{B}). \quad (17.24)$$

Proof:

$$\begin{aligned}
 (\vec{\sigma} \cdot \vec{A})(\vec{\sigma} \cdot \vec{B}) &= \sum_{j,k} \sigma_j A_j \sigma_k B_k \\
 &= \sum_{j,k} A_j B_k \sigma_j \sigma_k \\
 &= \sum_{j,k} A_j B_k \left[\sum_l i \epsilon_{jkl} \sigma_l + \delta_{jk} I_{2 \times 2} \right] \\
 &= \sum_l i \sigma_l \left(\sum_{j,k} \epsilon_{jkl} A_j B_k \right) + \sum_j A_j B_j I_{2 \times 2} \\
 &= \sum_l i \sigma_l (\vec{A} \times \vec{B})_l + \vec{A} \cdot \vec{B} I_{2 \times 2} \\
 &= i \vec{\sigma} \cdot (\vec{A} \times \vec{B}) + \vec{A} \cdot \vec{B} I_{2 \times 2} \quad (\text{Proved}).
 \end{aligned}$$

17.2 Matrix Representations of the Eigenstates of \hat{S}_x , \hat{S}_y and \hat{S}_z

In the following discussions the eigenstates of $\{\hat{S}^2, \hat{S}_x\}$, $\{\hat{S}^2, \hat{S}_y\}$ and $\{\hat{S}^2, \hat{S}_z\}$ are denoted by

$$|S_x, \pm\rangle, |S_y, \pm\rangle, \text{ and } |S_z, \pm\rangle,$$

respectively. We will use the simultaneous eigenkets of \hat{S}^2 and \hat{S}_z i.e., the kets $|S_z, \pm\rangle$ as the basis. In this basis the matrix representation of the basis kets themselves are very simple. We have

$$|S_z, +\rangle \doteq \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \text{and} \quad |S_z, -\rangle \doteq \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \quad (17.25)$$

Next, let us find the matrix representations of the states $|S_x, \pm\rangle$ in the $\{|S_z, +\rangle, |S_z, -\rangle\}$ basis. First, the eigenvalue equations for the states $|S_x, \pm\rangle$ are:

$$\hat{S}_x |S_x, \pm\rangle = \pm \frac{1}{2} \hbar |S_x, \pm\rangle. \quad (17.26)$$

Let the matrix representation of the state $|S_x, +\rangle$ be

$$|S_x, +\rangle \doteq \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}. \quad (17.27)$$

Therefore, the eigenvalue equation of $|S_x, +\rangle$ in matrix form can be written as

$$\frac{1}{2} \hbar \sigma_x \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \frac{1}{2} \hbar \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \quad (17.28)$$

or,

$$\sigma_x \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix},$$

or,

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

i.e.,

$$\begin{pmatrix} x_2 \\ x_1 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

This equation tells us that $x_1 = x_2$. Therefore, the matrix representation of the state $|S_x, +\rangle$ is of the form

$$|S_x, +\rangle \doteq \begin{pmatrix} x_1 \\ x_1 \end{pmatrix}.$$

Normalizing, we have

$$|S_x, +\rangle \doteq \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}. \quad (17.29)$$

Similarly, we can find the matrix representation for the state $|S_x, -\rangle$:

$$|S_x, -\rangle \doteq \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}. \quad (17.30)$$

Following the same procedure, we find the matrix representations of the states $|S_y, +\rangle$ and $|S_y, -\rangle$ in the $\{|S_z, +\rangle, |S_z, -\rangle\}$ basis. We find

$$|S_y, +\rangle \doteq \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \end{pmatrix}, \quad (17.31)$$

and

$$|S_y, -\rangle \doteq \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -i \end{pmatrix}. \quad (17.32)$$

Summary

In summary, we have obtained the following representations in the $\{|S_z, +\rangle, |S_z, -\rangle\}$ basis:

$$|S_z, +\rangle \doteq \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad |S_z, -\rangle \doteq \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \quad (17.33)$$

Also

$$|S_x, +\rangle \doteq \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad (17.34)$$

and

$$|S_x, -\rangle \doteq \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \end{pmatrix} - \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \quad (17.35)$$

In terms of the abstract vectors we can write

$$|S_x, +\rangle = \frac{1}{\sqrt{2}} |S_z, +\rangle + \frac{1}{\sqrt{2}} |S_z, -\rangle \quad (17.36)$$

$$|S_x, -\rangle = \frac{1}{\sqrt{2}} |S_z, +\rangle - \frac{1}{\sqrt{2}} |S_z, -\rangle \quad (17.37)$$

Next,

$$|S_y, +\rangle \doteq \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \frac{i}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad (17.38)$$

and

$$|S_y, -\rangle \doteq \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -i \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \end{pmatrix} - \frac{i}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \quad (17.39)$$

In terms of the abstract vectors we can write

$$|S_y, +\rangle = \frac{1}{\sqrt{2}} |S_z, +\rangle + \frac{i}{\sqrt{2}} |S_z, -\rangle \quad (17.40)$$

$$|S_y, -\rangle = \frac{1}{\sqrt{2}} |S_z, +\rangle - \frac{i}{\sqrt{2}} |S_z, -\rangle \quad (17.41)$$

Spin Probabilities

Using equations (17.36), (17.37), (17.40) and (17.41) we can work out the probabilities for finding specific spin values in measurements of spin. For example, suppose that the spin state of the electron is $|S_z, +\rangle$. If we measure the x -component of the electron's spin, what is the probability that the measurement will yield the value $\frac{1}{2}\hbar$? Using the postulates of quantum mechanics, we obtain

$$P\left(S_x = \frac{1}{2}\hbar\right) = |\langle S_x, + | S_z, + \rangle|^2 = \left(\frac{1}{\sqrt{2}}\right)^2 = \frac{1}{2}. \quad (17.42)$$

Similarly, the probability that the measurement of S_x gives $-\frac{1}{2}\hbar$ is

$$P\left(S_x = -\frac{1}{2}\hbar\right) = |\langle S_x, -|S_z, +\rangle|^2 = \left(\frac{1}{\sqrt{2}}\right)^2 = \frac{1}{2}. \quad (17.43)$$

This result is dramatically illustrated in an experiment with two Stern-Gerlach arrangements as described below.

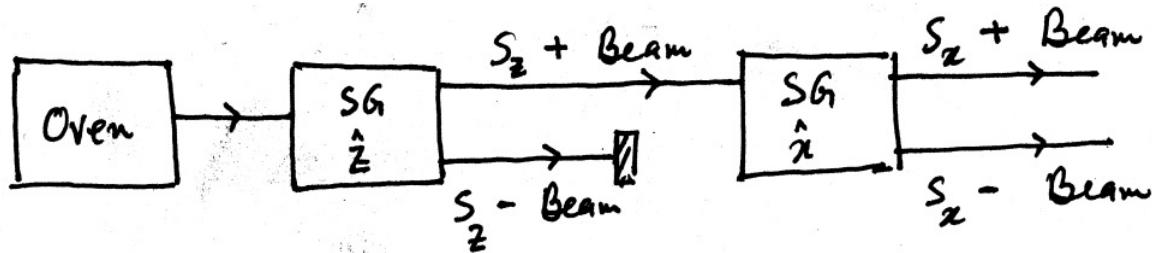
Stern-Gerlach Apparatus for Electron Spin

Figure 17.1: Stern-Gerlach apparatus for electron spin.

Silver atoms from an oven are made to pass through a Stern-Gerlach arrangement in which the inhomogeneity in the magnetic field is along the z -axis ((SG \hat{z})). The beam divides into two components of equal intensity. In one of the beams all the silver atoms are in the $|S_z, +\rangle$ state and in the other beam all of the atoms are in the $|S_z, -\rangle$ state.

Now, suppose that the $|S_z, -\rangle$ beam is blocked. The other beam is allowed to pass through a second Stern-Gerlach apparatus in which the inhomogeneity of the magnetic field is along the x -axis (SG \hat{x}). We find that the $|S_z, +\rangle$ beam which entered the SG \hat{x} apparatus divides into two beams of equal intensities. This shows that if electrons are in the state $|S_z, +\rangle$, then there is an equal probability that a measurement of S_x will give $\frac{1}{2}\hbar$ or $-\frac{1}{2}\hbar$.

17.3 Eigenvalues and Eigenvectors of $\hat{S} \cdot \hat{n}$

So far we have found the eigenvalues and the eigenvectors of the operators \hat{S}_x , \hat{S}_y and \hat{S}_z . We shall now find an expression for the eigenvectors of the operator $\hat{S} \cdot \hat{n}$ in the $|S_z, \pm\rangle$ basis. Here \hat{n} is an arbitrary direction characterized by the polar angle θ and the azimuthal angle ϕ (figure below).

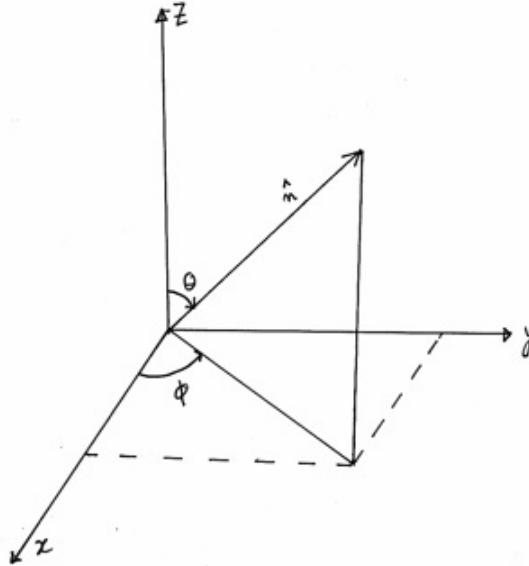


Figure 17.2: A direction \hat{n} in space is characterized by two angles in the spherical coordinate system, namely, the polar angle θ and the azimuthal angle ϕ .

The operator $\hat{S} \cdot \hat{n}$ represents the component of the spin angular momentum along \hat{n} . We can write \hat{n} as

$$\begin{aligned}\hat{n} &= \hat{i}n_x + \hat{j}n_y + \hat{k}n_z \\ &= \hat{i}\sin\theta\cos\phi + \hat{j}\sin\theta\sin\phi + \hat{k}\cos\theta.\end{aligned}\tag{17.44}$$

Now, the matrix representation of the operator $\hat{S} \cdot \hat{n}$ in the $|S_z, \pm\rangle$ basis is $\frac{1}{2}\hbar\vec{\sigma} \cdot \hat{n}$. Therefore,

$$\left(\hat{S} \cdot \hat{n}\right)^2 \doteq \left(\frac{1}{2}\hbar\vec{\sigma} \cdot \hat{n}\right)^2 = \frac{\hbar^2}{4}(\vec{\sigma} \cdot \hat{n})^2 = \frac{\hbar^2}{4}I.\tag{17.45}$$

Hence, the eigenvalues of $\hat{S} \cdot \hat{n}$ are $\pm\hbar/2$.

Let us now find the eigenvectors of $\hat{S} \cdot \hat{n}$ with eigenvalues $+\frac{1}{2}\hbar$. First, we write the eigenvalue equation:

$$\hat{S} \cdot \hat{n} |S_n, +\rangle = \frac{1}{2}\hbar |S_n, +\rangle. \quad (17.46)$$

In the $\{|S_z, +\rangle, |S_z, -\rangle\}$ basis, the matrix representation of the operator $\hat{S} \cdot \hat{n}$ is

$$\begin{aligned} \hat{S} \cdot \hat{n} &\doteq \frac{1}{2}\hbar \vec{\sigma} \cdot \hat{n} \\ &= \frac{1}{2}\hbar \sigma_x \sin \theta \cos \phi + \frac{1}{2}\hbar \sigma_y \sin \theta \sin \phi + \frac{1}{2}\hbar \sigma_z \cos \theta \\ &= \frac{1}{2}\hbar \begin{pmatrix} \cos \theta & \sin \theta e^{-i\phi} \\ \sin \theta e^{i\phi} & -\cos \theta \end{pmatrix}. \end{aligned} \quad (17.47)$$

Let the matrix representation of $|S_n, +\rangle$ in the same basis be

$$|S_n, +\rangle \doteq \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \quad (17.48)$$

i.e.,

$$\langle S_z, + | S_n, + \rangle = x_1, \quad \text{and} \quad \langle S_z, - | S_n, + \rangle = x_2. \quad (17.49)$$

Therefore, the matrix representation of the eigenvalue equation (17.46) is

$$\frac{1}{2}\hbar \begin{pmatrix} \cos \theta & \sin \theta e^{-i\phi} \\ \sin \theta e^{i\phi} & -\cos \theta \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \frac{1}{2}\hbar \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

Canceling the factor $\frac{1}{2}\hbar$ from both sides we get

$$\begin{pmatrix} \cos \theta & \sin \theta e^{-i\phi} \\ \sin \theta e^{i\phi} & -\cos \theta \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix},$$

or,

$$\begin{pmatrix} \cos \theta x_1 + \sin \theta e^{-i\phi} x_2 \\ \sin \theta e^{i\phi} x_1 - \cos \theta x_2 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

Therefore,

$$\cos \theta x_1 + \sin \theta e^{-i\phi} x_2 = x_1,$$

i.e.,

$$(1 - \cos \theta)x_1 = \sin \theta e^{-i\phi} x_2,$$

or,

$$2 \sin^2 \frac{\theta}{2} x_1 = 2 \sin \frac{\theta}{2} \cos \frac{\theta}{2} e^{-i\phi} x_2.$$

From the above equation we get

$$\frac{x_1}{x_2} = \frac{e^{-i\phi/2} \cos \theta/2}{e^{i\phi/2} \sin \theta/2}. \quad (17.50)$$

Hence, the normalized matrix representation for the state $|S_n, +\rangle$ is

$$\begin{aligned} |S_n, +\rangle &\doteq \begin{pmatrix} e^{-i\phi/2} \cos \theta/2 \\ e^{i\phi/2} \sin \theta/2 \end{pmatrix} \\ &= e^{-i\phi/2} \cos \theta/2 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + e^{i\phi/2} \sin \theta/2 \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \end{aligned} \quad (17.51)$$

The above equation can be written in terms of the abstract vectors in the Hilbert space as

$$|S_n, +\rangle = e^{-i\phi/2} \cos \frac{\theta}{2} |S_z, +\rangle + e^{i\phi/2} \sin \frac{\theta}{2} |S_z, -\rangle. \quad (17.52)$$

Proceeding in a similar fashion, we can find the matrix representation of $|S_n, -\rangle$ in the $\{|S_z, +\rangle, |S_z, -\rangle\}$ basis. We get

$$|S_n, -\rangle \doteq \begin{pmatrix} -e^{-i\phi/2} \sin \theta/2 \\ e^{i\phi/2} \cos \theta/2 \end{pmatrix}, \quad (17.53)$$

i.e.,

$$|S_n, -\rangle = -e^{-i\phi/2} \sin \theta/2 |S_z, +\rangle + e^{i\phi/2} \cos \theta/2 |S_z, -\rangle. \quad (17.54)$$

17.4 Non-relativistic Description of a Spin-1/2 Particle

17.4.1 Basis Vectors

So far we have considered quantum mechanical description of a system having spatial degrees of freedom or spin degrees of freedom only. But a system, an electron say, can have both spatial and spin degrees of freedom. How is our formulation to be generalized to include both these degrees of freedom?

We note that to describe a quantum system we first need to establish a basis of states. The basis set could be written in many different ways. To set up a basis, we need a complete set of commuting observables (CSCO) of the system and the simultaneous eigenstates of these observables could serve as a basis.

For a point particle with spin, the CSCO can be chosen in various ways as shown below:

$$\{\hat{R}, \hat{S}^2, \hat{S}_z\}$$

$$\{\hat{P}, \hat{S}^2, \hat{S}_z\}$$

$$\{\hat{H}, \hat{L}^2, \hat{L}_z, \hat{S}^2, \hat{S}_z\}.$$

We shall use the first set of CSCO. Let $|\vec{r}\rangle$ and $|\varepsilon\rangle$ be the eigenkets of \hat{R} and $\{\hat{S}^2, \hat{S}_z\}$, respectively, i.e.,

$$\hat{R}|\vec{r}\rangle = \vec{r}|\vec{r}\rangle \quad (17.55)$$

$$\hat{S}^2|\varepsilon\rangle = s(s+1)\hbar^2|\varepsilon\rangle \quad (17.56)$$

$$\hat{S}_z|\varepsilon\rangle = \frac{1}{2}\hbar\varepsilon|\varepsilon\rangle \quad (\varepsilon = 2m_s). \quad (17.57)$$

For a spin-1/2 particle, $s = 1/2$, therefore, there are only two possible linearly independent basis states with $\varepsilon = 1$ and $\varepsilon = -1$, corresponding to the z -component of spin being equal to $\frac{1}{2}\hbar$ and $-\frac{1}{2}\hbar$, respectively. In other words

$$\hat{S}^2|\pm\rangle = \frac{3}{4}\hbar^2|\pm\rangle, \quad (17.58)$$

$$\hat{S}_z|\pm\rangle = \pm\frac{1}{2}\hbar|\pm\rangle, \quad (17.59)$$

where we have labelled the states as $|+\rangle \equiv |\varepsilon = 1\rangle$ and $|-\rangle \equiv |\varepsilon = -1\rangle$. Now, since the space and spin degrees of freedom are independent, we have

$$[\hat{R}, \hat{S}] = 0. \quad (17.60)$$

The simultaneous eigenkets $|\vec{r}, \varepsilon\rangle$ of the CSCO $\{\hat{R}, \hat{S}^2, \hat{S}_z\}$ can then be written as a product

$$|\vec{r}, \varepsilon\rangle = |\vec{r}\rangle|\varepsilon\rangle. \quad (17.61)$$

That is, the ket space formed by the eigenkets of $\{\hat{R}, \hat{S}^2, \hat{S}_z\}$ is a tensor product of the ket space spanned by the eigenkets $|\vec{r}\rangle$ of \hat{R} and the eigenkets $|\varepsilon\rangle$ of $\{\hat{S}^2, \hat{S}_z\}$. Thus

$$H = H_{\vec{r}} \otimes H_s. \quad (17.62)$$

The orthogonality and completeness conditions of the basis kets $|\vec{r}, \varepsilon\rangle$ are

$$\langle \vec{r}' \varepsilon' | \vec{r} \varepsilon \rangle = \delta(\vec{r} - \vec{r}') \delta_{\varepsilon \varepsilon'} \quad (17.63)$$

and

$$\sum_{\varepsilon} \int d^3r |\vec{r}\varepsilon\rangle \langle \vec{r}\varepsilon| = \hat{I}. \quad (17.64)$$

For spin-1/2 particles, $\varepsilon = \pm$, and the completeness condition, i.e., Eq. (17.64) can be written more explicitly as

$$\int d^3r |\vec{r}+\rangle \langle \vec{r}+| + \int d^3r |\vec{r}-\rangle \langle \vec{r}-| = \hat{I}. \quad (17.65)$$

17.4.2 State Vector

Any general state $|\psi\rangle$ of the particle can be expanded as a linear combination of the basis kets $|\vec{r}\varepsilon\rangle$:

$$\begin{aligned} |\psi\rangle &= \sum_{\varepsilon} \int d^3r |\vec{r}\varepsilon\rangle \langle \vec{r}\varepsilon| \psi \rangle \\ &= \sum_{\varepsilon} \int d^3r |\vec{r}\varepsilon\rangle \psi_{\varepsilon}(\vec{r}) \end{aligned} \quad (17.66)$$

where we have defined

$$\psi_{\varepsilon}(\vec{r}) = \langle \vec{r}\varepsilon | \psi \rangle. \quad (17.67)$$

The numbers $\psi_{\varepsilon}(\vec{r})$ which depend on three continuous indices (i.e., x , y and z) and on one discrete index ε (+ or - for a spin-1/2 particle), can be considered as the ‘coordinates’ of the state $|\psi\rangle$ in the $|\vec{r}\varepsilon\rangle$ basis. Thus, in order to characterize the state of a spin-1/2 particle completely, we have to specify two functions of the space variables x , y and z :

$$\psi_+(\vec{r}) \equiv \langle \vec{r}, + | \psi \rangle, \quad \text{and} \quad \psi_-(\vec{r}) \equiv \langle \vec{r}, - | \psi \rangle. \quad (17.68)$$

These two functions are often written in the form of a two-component column matrix, called a spinor, which we shall denote as $[\psi](\vec{r})$, i.e.,

$$[\psi](\vec{r}) = \begin{pmatrix} \psi_+(\vec{r}) \\ \psi_-(\vec{r}) \end{pmatrix}. \quad (17.69)$$

The bra $\langle \psi |$ associated with the ket $|\psi\rangle$ is given by

$$\begin{aligned} \langle \psi | &= \sum_{\varepsilon} \int d^3r \langle \psi | \vec{r}\varepsilon \rangle \langle \vec{r}\varepsilon | \\ &= \sum_{\varepsilon} \int d^3r \psi_{\varepsilon}^*(\vec{r}) \langle \vec{r}\varepsilon |. \end{aligned} \quad (17.70)$$

The bra $\langle \psi |$ is thus represented by two functions $\psi_+^*(\vec{r})$ and $\psi_-^*(\vec{r})$, written as a row matrix which is the adjoint of $[\psi](\vec{r})$. Thus

$$\langle \psi | \doteq [\psi]^{\dagger}(\vec{r}) = \begin{pmatrix} \psi_+^*(\vec{r}) & \psi_-^*(\vec{r}) \end{pmatrix}. \quad (17.71)$$

With this notation, the scalar product of two vectors $|\psi\rangle$ and $|\phi\rangle$ can be written as

$$\begin{aligned}
 \langle\psi|\phi\rangle &= \sum_{\epsilon} \int d^3r \langle\psi|\vec{\epsilon}\rangle \langle\vec{\epsilon}|\phi\rangle \\
 &= \sum_{\epsilon} \int d^3r \psi_{\epsilon}^*(\vec{r}) \phi_{\epsilon}(\vec{r}) \\
 &= \int d^3r \begin{pmatrix} \psi_+^*(\vec{r}) & \psi_-^*(\vec{r}) \end{pmatrix} \begin{pmatrix} \phi_+(\vec{r}) \\ \phi_-(\vec{r}) \end{pmatrix} \\
 &= \int d^3r [\psi]^\dagger(\vec{r})[\phi](\vec{r}).
 \end{aligned} \tag{17.72}$$

In particular, the normalization of the state vector $|\psi\rangle$ is expressed as

$$\begin{aligned}
 \langle\psi|\psi\rangle &= \int d^3r [\psi]^\dagger(\vec{r})[\psi](\vec{r}) \\
 &= \int d^3r [|\psi_+(\vec{r})|^2 + |\psi_-(\vec{r})|^2] = 1.
 \end{aligned} \tag{17.73}$$

Probabilistic Interpretation of the wavefunctions $\psi_{\pm}(\vec{r})$

We have the following interpretation for the wave functions:

- $|\psi_+(\vec{r})|^2 d^3r$ = probability that the electron is found in a small volume d^3r around \vec{r} with the z -component of the spin $\frac{1}{2}\hbar$. A similar interpretation holds for $|\psi_-(\vec{r})|^2 d^3r$.

If we integrate $|\psi_+(\vec{r})|^2 d^3r$ over \vec{r} we get

- $\int |\psi_+(\vec{r})|^2 d^3r$ = probability that the z -component of the electron's spin is $\frac{1}{2}\hbar$ irrespective of its position. Similarly
- $\int |\psi_-(\vec{r})|^2 d^3r$ = probability that the z -component of the electron's spin is $-\frac{1}{2}\hbar$ irrespective of its position.

As a final comment, we note that the basis vectors for a spin-1/2 particle are tensor products of a ket belonging to $H_{\vec{r}}$ and a ket belonging to H_s (Eq. 17.61). However, the state vector may or may not be a tensor product of a vector in $H_{\vec{r}}$ and a vector in H_s . Only in the special case when the Hamiltonian is spin-independent the state vector $|\psi\rangle$ is of this type, i.e.,

$$|\psi\rangle = |\phi\rangle|\alpha\rangle,$$

where

$$|\phi\rangle \in H_{\vec{r}}, \quad \text{and} \quad |\alpha\rangle \in H_s.$$

In this case we have

$$\begin{aligned}
 \psi_{\epsilon}(\vec{r}) &= \langle\vec{r}\epsilon|\psi\rangle \\
 &= (\langle\vec{r}|\langle\epsilon|)(|\phi\rangle|\alpha\rangle) \\
 &= \langle\vec{r}|\phi\rangle\langle\epsilon|\alpha\rangle \\
 &= \phi(\vec{r})c_{\epsilon},
 \end{aligned} \tag{17.74}$$

where

$$c_\varepsilon = \langle \varepsilon | \alpha \rangle . \quad (17.75)$$

For a spin-1/2 particles $\varepsilon = \pm$, and in the special case of spin-independent Hamiltonian we have

$$\psi_+(\vec{r}) = \phi(\vec{r}) c_+$$

and

$$\psi_-(\vec{r}) = \phi(\vec{r}) c_- .$$

Therefore, the spinor $[\psi](\vec{r})$ is

$$[\psi](\vec{r}) = \begin{pmatrix} \psi_+(\vec{r}) \\ \psi_-(\vec{r}) \end{pmatrix} = \phi(\vec{r}) \begin{pmatrix} c_+ \\ c_- \end{pmatrix} . \quad (17.76)$$

The square of the norm of $|\psi\rangle$ is then given by

$$\langle \psi | \psi \rangle = \langle \phi | \phi \rangle \langle \chi | \chi \rangle = (|c_+|^2 + |c_-|^2) \int d^3 r |\phi(\vec{r})|^2 . \quad (17.77)$$

17.5 Electron in an External Magnetic Field. Pauli Equation

Suppose that an electron is placed in an external time-independent magnetic field \vec{B} . The magnetic moment operator of the electron is

$$\hat{\mu} = \hat{\mu}_{\text{orbit}} + \hat{\mu}_{\text{spin}} , \quad (17.78)$$

where

$$\hat{\mu}_{\text{orbit}} = \frac{q_e}{2m_e} \hat{L} \quad (17.79)$$

and

$$\hat{\mu}_{\text{spin}} = g_s \frac{q_e}{2m_e} \hat{S} . \quad (17.80)$$

Here q_e is the charge of the electron, i.e., $q_e = -e$ with $e = 1.6 \times 10^{-19}$ C, and g_s is the spin gyromagnetic ratio. From Dirac theory we have $g_s = 2$ for electrons. The total magnetic moment of the electron is then

$$\hat{\mu} = \frac{q_e}{2m_e} (\hat{L} + 2\hat{S}) . \quad (17.81)$$

In the spin space of the electron, we can express $\hat{\vec{\mu}}$ as a 2×2 matrix:

$$\hat{\vec{\mu}} = \frac{q_e}{2m_e} (\hat{\vec{L}} \hat{1}_{2 \times 2} + \hbar \vec{\sigma}) \quad (17.82)$$

where we have used the matrix representation of the spin operator of a spin-1/2 particle, i.e.,

$$\hat{\vec{S}} \doteq \frac{1}{2} \hbar \vec{\sigma}, \quad (17.83)$$

and $\hat{1}_{2 \times 2}$ is a unit 2×2 matrix.

Now, the interaction energy of the electron with the external magnetic field is

$$\begin{aligned} V &= -\hat{\vec{\mu}} \cdot \vec{B} \\ &= \mu_B \left(\frac{\hat{\vec{L}}}{\hbar} \hat{1}_{2 \times 2} + \vec{\sigma} \right) \cdot \vec{B}, \end{aligned} \quad (17.84)$$

where

$$\mu_B = \frac{|q_e|\hbar}{2m_e} = \frac{e\hbar}{2m_e} \quad (17.85)$$

is the Bohr magneton. Here e is the magnitude of charge of the electron, hence e is a positive number, $e = 1.6 \times 10^{-19}$ C. The Hamiltonian operator of the electron is then

$$\hat{H} = \left(\frac{\hat{\vec{p}}^2}{2m_e} + V(\hat{\vec{r}}) \right) \hat{1}_{2 \times 2} + \mu_B \left(\frac{\hat{\vec{L}}}{\hbar} \hat{1}_{2 \times 2} + \vec{\sigma} \right) \cdot \vec{B}, \quad (17.86)$$

where $V(\hat{\vec{r}})$ is the potential energy of the electron's interaction with other fields, for example, the Coulomb field of the nucleus if the electron is bound in an atomic orbit. The Schrödinger equation for the electron is then

$$i\hbar \frac{\partial}{\partial t} |\psi\rangle = \hat{H} |\psi\rangle, \quad (17.87)$$

or, in terms of the components of $|\psi\rangle$

$$\begin{aligned} i\hbar \frac{\partial}{\partial t} \begin{pmatrix} \psi_+(\vec{r}, t) \\ \psi_-(\vec{r}, t) \end{pmatrix} &= \\ \left[\left(\frac{-\hbar^2}{2m_e} \nabla^2 + V(\vec{r}) + \frac{\mu_B}{\hbar} \hat{\vec{L}} \cdot \vec{B} \right) \hat{1}_{2 \times 2} + \mu_B \vec{\sigma} \cdot \vec{B} \right] \begin{pmatrix} \psi_+(\vec{r}, t) \\ \psi_-(\vec{r}, t) \end{pmatrix}. \end{aligned} \quad (17.88)$$

This non-relativistic equation for a spin-1/2 particle is known as Pauli equation.



18. sheet-18 : Rotations and Angular Momentum



19. sheet-19 : Addition of Angular Momentum

20. sheet-20 : Time Independent Perturbation

The major task in any practical application of quantum mechanics is to solve the eigenvalue equation of the Hamiltonian H of the system. Considering the bound states, the eigenvalues of H are discrete and corresponding to each eigenvalue there may be one or several linearly independent eigenvectors. The eigenvalue equation, i.e., the time independent Schrödinger equation is

$$H |E_n\rangle = E_n |E_n\rangle \quad (20.1)$$

Except for few special cases the eigenvalue equation cannot be solved exactly. The equation then has to be solved numerically, or approximate methods have to be devised to solve the equation to any desired order of accuracy.

Time dependent perturbation theory applies when H is of the form

$$H = H_0 + V \quad (20.2)$$

where the eigenvalues and eigenvectors of H_0 are completely known and V is an additional time-independent potential called the perturbation.

Let us denote the eigenvalues of H_0 as $E_n^{(0)}$ and the corresponding eigenvectors as $|E_n^{(0)}\rangle$ so that the eigenvalue equation for H_0 is written as

$$H_0 |E_n^{(0)}\rangle = E_n^{(0)} |E_n^{(0)}\rangle \quad (20.3)$$

We assume that all the eigenvalues and the eigenvectors of H_0 are already calculated.

20.1 Non-degenerate Perturbation Theory

We assume that the eigenvalues $E_n^{(0)}$ of H_0 are non-degenerate, i.e., there is only one linearly independent eigenvector $|E_n^{(0)}\rangle$ corresponding to $E_n^{(0)}$. Since the eigenvectors $|E_n^{(0)}\rangle$ form a complete set of vectors, we can express any vector in the Hilbert space as a linear combination of the eigenvectors of H_0 . Further, eigenvectors belonging to different eigenvalues are orthogonal. We also normalize each of the eigenvectors of H_0 . So these eigenvectors form a complete orthonormal set, i.e.,

$$\langle E_n^{(0)} | E_m^{(0)} \rangle = \delta_{nm} \quad (20.4)$$

and

$$\hat{1} = \sum_k |E_k^{(0)}\rangle \langle E_k^{(0)}| \quad (20.5)$$

Now we modify the eigenvalue equation for the full Hamiltonian H (equation (20.1)) as

$$(H_0 + \lambda V) |E_n\rangle_\lambda = E_n \lambda |E_n\rangle_\lambda \quad (20.6)$$

Where we have introduced a real parameter λ whose value lies in the range $(0, 1)$. The eigenvalue $E_n \lambda$ and the eigenvector $|E_n\rangle_\lambda$ in equation (??) are not quite the same as the corresponding quantities in equation (20.1). It is only in the limit the limit $\lambda \rightarrow 1$ would $E_n \lambda$ and $|E_n\rangle_\lambda$ in equation (20.6) coincide with actual values. Furthermore,

$$\lim_{\lambda \rightarrow 0} E_n \lambda = E_n^{(0)} \quad (20.7)$$

$$\lim_{\lambda \rightarrow 0} |E_n\rangle_\lambda = |E_n^{(0)}\rangle \quad (20.8)$$

Thus, as $\lambda \rightarrow 0$, the perturbation is switched off and as $\lambda \rightarrow 1$, the full perturbation V is operative.

We will now set up a perturbative scheme for solving $E_{n,\lambda}$ and $|E_n\rangle_\lambda$ and at the end set $\lambda = 1$. First, we write $E_{n,\lambda}$ and $|E_n\rangle_\lambda$ as power series in λ

$$E_{n,\lambda} = E_n^{(0)} + \lambda E_n^{(1)} + \lambda^2 E_n^{(2)} + \dots \quad (20.9)$$

$$|E_n\rangle_\lambda = |E_n^{(0)}\rangle + \lambda |E_n^{(1)}\rangle + \lambda^2 |E_n^{(2)}\rangle + \dots \quad (20.10)$$

Substituting equation (20.9) and (20.10) in equation (20.6) we have

$$(H_0 + \lambda V) \left(|E_n^{(0)}\rangle + \lambda |E_n^{(1)}\rangle + \lambda^2 |E_n^{(2)}\rangle + \dots \right) = \left(E_n^{(0)} + \lambda E_n^{(1)} + \lambda^2 E_n^{(2)} + \dots \right) \left(|E_n^{(0)}\rangle + \lambda |E_n^{(1)}\rangle + \lambda^2 |E_n^{(2)}\rangle + \dots \right)$$

(20.11)

or

$$\begin{aligned} & H_0 |E_n^{(0)}\rangle + \lambda (H_0 |E_n^{(1)}\rangle + V |E_n^{(0)}\rangle) + \lambda^2 (H_0 |E_n^{(2)}\rangle + V |E_n^{(1)}\rangle) + \dots \\ &= E_n^{(0)} |E_n^{(0)}\rangle + \lambda (E_n^{(1)} |E_n^{(0)}\rangle + E_n^{(0)} |E_n^{(1)}\rangle) + \lambda^2 (E_n^{(2)} |E_n^{(0)}\rangle + E_n^{(1)} |E_n^{(1)}\rangle + E_n^{(0)} |E_n^{(2)}\rangle) + \dots \end{aligned}$$

We will solve this equation order by order in λ . So we equate the coefficient of equal power of λ on both sides of the above equation. We have up to order λ^2

$$\lambda^0 : H_0 |E_n^{(0)}\rangle = E_n^{(0)} |E_n^{(0)}\rangle \quad (20.12)$$

$$\lambda^1 : H_0 |E_n^{(1)}\rangle + V |E_n^{(0)}\rangle = E_n^{(0)} |E_n^{(1)}\rangle + E_n^{(1)} |E_n^{(0)}\rangle \quad (20.13)$$

$$\lambda^2 : H_0 |E_n^{(2)}\rangle + V |E_n^{(1)}\rangle = E_n^{(0)} |E_n^{(2)}\rangle + E_n^{(1)} |E_n^{(1)}\rangle + E_n^{(2)} |E_n^{(0)}\rangle \quad (20.14)$$

Equation (20.12) is considered solved because we have assumed that we know fully the eigenvalues and eigenvectors of H_0 .

20.1.1 First-Order correction to energy: $E_n^{(1)}$

The first order correction to the unperturbed energy of the n -th level is $E_n^{(1)}$. This can be found from equation (20.13). We start by taking the product of equation (20.13) with $\langle E_n^{(0)} |$. We get

$$\langle E_n^{(0)} | H_0 |E_n^{(1)}\rangle + \langle E_n^{(0)} | V |E_n^{(0)}\rangle = E_n^{(0)} \langle E_n^{(0)} | E_n^{(1)}\rangle + E_n^{(1)} \langle E_n^{(0)} | E_n^{(0)}\rangle \quad (20.15)$$

Now

$$H_0 |E_n^{(0)}\rangle = E_n^{(0)} |E_n^{(0)}\rangle \quad (20.16)$$

Since H_0 is hermitian

$$\langle E_n^{(0)} | H_0 = E_n^{(0)} = E_n^{(0)} \langle E_n^{(0)} | \quad (20.17)$$

Also

$$\langle E_n^{(0)} | E_n^{(0)}\rangle = 1 \quad (20.18)$$

Therefore equation (20.15) becomes

$$\begin{aligned} E_n^{(0)} \langle E_n^{(0)} | E_n^{(1)}\rangle + \langle E_n^{(0)} | V |E_n^{(0)}\rangle &= E_n^{(0)} \langle E_n^{(0)} | E_n^{(1)}\rangle + E_n^{(1)} \\ E_n^{(1)} &= \langle E_n^{(0)} | V |E_n^{(0)}\rangle = \equiv V_{nn} \end{aligned} \quad (20.19)$$

This is a fundamental result of time independent perturbation theory of non-degenerate levels. The first order correction to the n -th energy level is the expectation value of the perturbation potential in the unperturbed state.

20.1.2 First-Order correction to eigenstate

The ket $|E_n^{(1)}\rangle$ is the first-order correction to the zeroth-order eigenket $|E_n^{(0)}\rangle$. The ket $|E_n^{(1)}\rangle$ is also found from equation (20.13). First we write $|E_n^{(1)}\rangle$ as a linear combination of $|E_n^{(0)}\rangle$

$$|E_n^{(1)}\rangle = \sum_m |E_m^{(0)}\rangle \langle E_m^{(0)}|E_n^{(1)}\rangle \quad (20.20)$$

$$= \sum_m |E_m^{(0)}\rangle C^{(1)}_{mn} \quad (20.21)$$

Where we have defined

$$C^{(1)}_{mn} = \langle E_m^{(0)}|E_n^{(1)}\rangle \quad (20.22)$$

Using equation (20.21) we write equation (20.13) as

$$\begin{aligned} \sum_m H_0 |E_m^{(0)}\rangle C^{(1)}_{mn} + V |E_n^{(0)}\rangle &= \sum_m E_m |E_m^{(0)}\rangle C^{(1)}_{mn} + E_n^{(1)} |E_n^{(0)}\rangle \\ \text{or, } \sum_m (E_n^{(0)} - E_m^{(0)}) |E_m^{(0)}\rangle C^{(1)}_{mn} &= V |E_n^{(0)}\rangle - E_n^{(1)} |E_n^{(0)}\rangle \end{aligned}$$

Taking the scalar product with $\langle E_k^{(0)}$ we have

$$(E_n^{(0)} - E_k^{(0)}) C^{(1)}_{kn} = \langle E_k^{(0)}|V|E_n^{(0)}\rangle - E_n^{(1)} \delta_{kn} \quad (20.23)$$

If $k = n$, the left side is zero and we recover the result

$$E_n^{(1)} = \langle E_n^{(0)}|V|E_n^{(0)}\rangle, \quad k = n \quad (20.24)$$

Thus, we cannot determine $C_{nn}^{(1)}$ from equation (20.23). This coefficient has to be determined from considerations of normalization of the eigenvectors as discussed later.

Next, if $k \neq n$, then equation (20.23) becomes

$$(E_n^{(0)} - E_k^{(0)}) C_{kn}^{(1)} = \langle E_k^{(0)}|V|E_n^{(0)}\rangle, \quad k \neq n \quad (20.25)$$

or

$$C_{kn}^{(1)} = \frac{\langle E_k^{(0)} | V | E_n^{(0)} \rangle}{E_n^{(0)} - E_k^{(0)}}, \quad k \neq n \quad (20.26)$$

Using equation (20.21) and (20.26) the first order correction to the state is

$$\langle E_n^{(1)} \rangle = \sum_k \langle E_k^{(0)} \rangle C_{kn}^{(1)} \quad (20.27)$$

$$= C_{nn}^{(1)} \langle E_n^{(0)} \rangle + \sum_{k \neq n} \frac{\langle E_k^{(0)} | V | E_n^{(0)} \rangle}{E_n^{(0)} - E_k^{(0)}} \langle E_k^{(0)} \rangle \quad (20.28)$$

$$= C_{nn}^{(1)} \langle E_n^{(0)} \rangle + \sum_{k \neq n} \frac{V_{kn}}{E_n^{(0)} - E_k^{(0)}} \langle E_k^{(0)} \rangle \quad (20.29)$$

Where

$$V_{kn} \equiv \langle E_k^{(0)} | V | E_n^{(0)} \rangle \quad (20.30)$$

Therefo, upto first order in λ , the eigenstate $|E_n\rangle_\lambda$ is (see equation (20.10))

$$|E_n\rangle_\lambda = \langle E_n^{(0)} \rangle + \lambda \langle E_n^{(1)} \rangle + \mathcal{O}(\lambda^2) \quad (20.31)$$

$$= \langle E_n^{(0)} \rangle + \lambda C_{nn}^{(1)} \langle E_n^{(0)} \rangle + \lambda \sum_{k \neq n} \frac{V_{kn}}{E_n^{(0)} - E_k^{(0)}} \langle E_n^{(0)} \rangle + \mathcal{O}(\lambda^2) \quad (20.32)$$

We want to normalize $|E_n\rangle_\lambda$ up to first order

$${}_\lambda \langle E_n | E_n \rangle_\lambda = 1 + \mathcal{O}(\lambda^2) \quad (20.33)$$

using equation (20.32), the normalization condition can be written as (noting λ is real)

$$1 + \lambda C_{nn}^{(1)} + \lambda C_{nn}^{(1)*} + \mathcal{O}(\lambda^2) = 1 + \mathcal{O}(\lambda^2) \quad (20.34)$$

$$or, \quad C_{nn}^{(1)} + C_{nn}^{(1)*} = 0 \quad (20.35)$$

$$i.e., \quad \text{Re}\{C_{nn}^{(1)}\} = 0 \quad (20.36)$$

Thus $C_{nn}^{(1)}$ is a purely imaginary number. We write

$$C_{nn}^{(1)} = i\alpha \quad (\alpha \in \mathfrak{R}) \quad (20.37)$$

Hence equation (20.32) can be written as

$$|E_n\rangle_\lambda = (1 + i\lambda\alpha) |E_n^{(0)}\rangle + \lambda \sum_{k \neq n} \frac{V_{kn}}{E_n^{(0)} - E_k^{(0)}} |E_n^{(0)}\rangle + \mathcal{O}(\lambda^2) \quad (20.38)$$

$$= e^{i\lambda\alpha} |E_n^{(0)}\rangle + \lambda \sum_{k \neq n} \frac{V_{kn}}{E_n^{(0)} - E_k^{(0)}} |E_n^{(0)}\rangle + \mathcal{O}(\lambda^2) \quad (20.39)$$

$$e^{-i\lambda\alpha} |E_n\rangle_\lambda = |E_n^{(0)}\rangle + \lambda \sum_{k \neq n} \frac{V_{kn}}{E_n^{(0)} - E_k^{(0)}} |E_n^{(0)}\rangle + \mathcal{O}(\lambda^2) \quad (20.40)$$

Now $e^{-i\lambda\alpha}$ is an overall phase factor which does not affect the normalization of $|E_n\rangle_\lambda$ up to first order. This factor can be set equal to 1 without loss of generality. So we take $\alpha = 0$, i.e.,

$$\langle E_n^{(0)} | E_n^{(1)} \rangle \equiv C_{nn}^{(1)} = i\alpha = 0 \quad (20.41)$$

i.e., we can choose $|E_n^{(1)}\rangle$ to be orthonormal to $|E_n^{(0)}\rangle$.

Thus, up to first order

$$|E_n\rangle_\lambda = |E_n^{(0)}\rangle + \lambda \sum_{k \neq n} \frac{V_{kn}}{E_n^{(0)} - E_k^{(0)}} |E_n^{(0)}\rangle + \mathcal{O}(\lambda^2) \quad (20.42)$$

Setting $\lambda = 1$ we get the desired eigenket of the full Hamiltonian H up to first order in the perturbing potential, i.e.,

$$|E_n\rangle = |E_n^{(0)}\rangle + |E_n^{(1)}\rangle \quad (20.43)$$

$$= |E_n^{(0)}\rangle + \sum_{k \neq n} \frac{V_{kn}}{E_n^{(0)} - E_k^{(0)}} |E_n^{(0)}\rangle \quad (20.44)$$

20.1.3 Second-Order Correction to Energy: $E_n^{(2)}$

We can find the second order correction to the energy, i.e., $E_n^{(2)}$ from equation (20.14). First, multiply equation (20.14) by $\langle E_n^{(0)} |$

$$\langle E_n^{(0)} | H_0 | E_n^{(2)} \rangle + \langle E_n^{(0)} | V | E_n^{(1)} \rangle = E_n^{(0)} \langle E_n^{(0)} | E_n^{(2)} \rangle + E_n^{(1)} \langle E_n^{(0)} | E_n^{(1)} \rangle + e_n^{(2)} \quad (20.45)$$

Since

$$\langle E_n^{(0)} | H_0 = E_n^{(0)} \langle E_n^{(0)} | \quad (20.46)$$

The first term on the left hand side of equation (20.45) cancels the first term on the right. Therefore, we have

$$\left\langle E_n^{(0)} \middle| V \middle| E_n^{(1)} \right\rangle = E_n^{(1)} \left\langle E_n^{(0)} \middle| E_n^{(1)} \right\rangle + E_n^{(2)} \quad (20.47)$$

$$\text{i.e., } E_n^{(2)} = \left\langle E_n^{(0)} \middle| V \middle| E_n^{(1)} \right\rangle - E_n^{(1)} \left\langle E_n^{(0)} \middle| E_n^{(1)} \right\rangle \quad (20.48)$$

Writing

$$\left\langle E_n^{(1)} \right\rangle = \sum_m \left\langle E_m^{(0)} \right\rangle \left\langle E_m^{(0)} \middle| E_n^{(1)} \right\rangle \quad (20.49)$$

we have

$$E_n^{(2)} = \sum_m \left\langle E_n^{(0)} \middle| V \middle| E_m^{(0)} \right\rangle \left\langle E_m^{(0)} \middle| E_n^{(1)} \right\rangle - E_n^{(1)} \left\langle E_m^{(0)} \middle| E_n^{(1)} \right\rangle \quad (20.50)$$

We now isolate the term with $m = n$ in the summation.

\therefore we have

$$E_n^{(2)} = \left\langle E_n^{(0)} \middle| V \middle| E_m^{(0)} \right\rangle \left\langle E_m^{(0)} \middle| E_n^{(1)} \right\rangle + \sum_{m \neq n} \left\langle E_n^{(0)} \middle| V \middle| E_m^{(0)} \right\rangle \left\langle E_m^{(0)} \middle| E_n^{(1)} \right\rangle - E_n^{(1)} \left\langle E_m^{(0)} \middle| E_n^{(1)} \right\rangle \quad (20.51)$$

But

$$\left\langle E_n^{(0)} \middle| V \middle| E_m^{(0)} \right\rangle = E_n^{(1)} \quad (20.52)$$

So the first term cancels the third term in equation (20.51). We then have

$$E_n^{(2)} = \sum_{m \neq n} \left\langle E_n^{(0)} \middle| V \middle| E_m^{(0)} \right\rangle \left\langle E_m^{(0)} \middle| E_n^{(1)} \right\rangle \quad (20.53)$$

Where the term $m = n$ is excluded from the sum.

Now, we have found previously (20.26)

$$\left\langle E_m^{(0)} \middle| E_n^{(1)} \right\rangle \equiv C_{mn}^{(1)} \quad (20.54)$$



21. sheet-21 : Variational Method



22. sheet-22 : WKB approximation



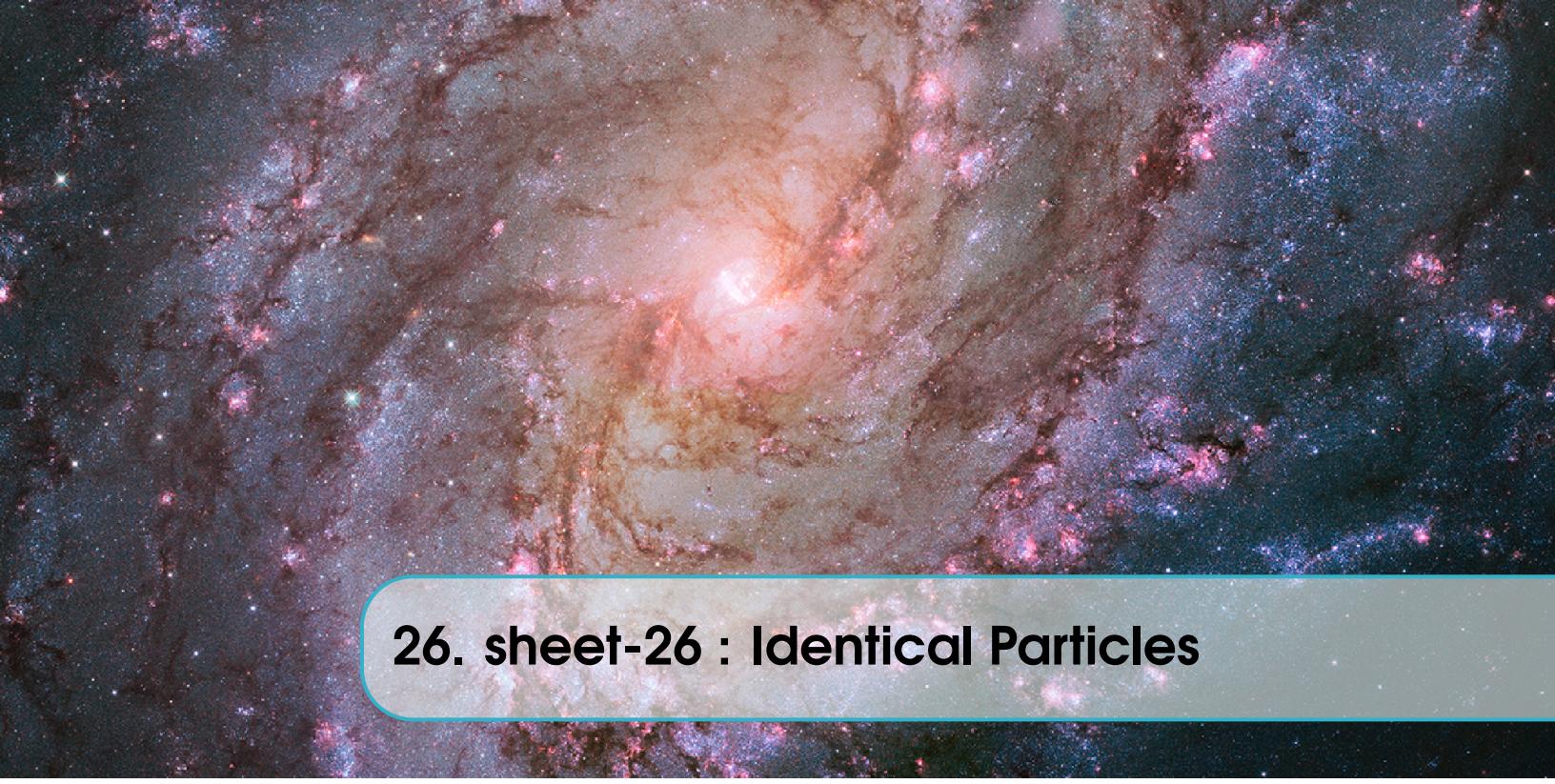
23. sheet-23 : Time Dependent Perturbation



24. sheet-24 : Scattering Theory



25. sheet-25 : Integral Equation for Scattering



26. sheet-26 : Identical Particles

Lecture 26 Identical Particles

26.1 Indistinguishability of identical particles

Two particles are said to be identical when they cannot be distinguished by means of any intrinsic property such as mass, electric charge, spin etc. Thus all the electrons in the universe are identical, as are all the protons, all the hydrogen atoms. But, an electron and a positron are not identical, since, although they have the same mass and same spin, they have different electric charges.

In classical mechanics, identical particles can, in principle, be distinguished from one another by observing their individual paths or by other means like physically labeling them, by different colors, for example. In quantum mechanics the situation is entirely different. By virtue of the uncertainty principle, the concept of the path of an electron (or any microscopic particle) ceases to have any meaning. If the position of an electron is exactly known at a given instant, its coordinates have no definite values even at the next instant. Hence by localizing and numbering the electrons at some instant, we make no progress towards identifying them at subsequent instants. If we localize one of the electrons at a subsequent

instant at some point in space, we cannot say which of the electrons has been detected at this point.

Thus, in quantum mechanics, there is in principle no possibility of separately following each of a number of similar particles and thereby distinguishing them. We may say that, in quantum mechanics, identical particles entirely lose their "individuality", i.e., they are indistinguishable. The principle of indistinguishability of similar particles, play a fundamental part in the quantum theory of systems composed of identical particles.

26.1.1 The wave function of two identical particles

Let us start by considering a system of only two identical particles. Because of the identity of the particles, the states of the system obtained from each other by interchanging the two particles must be completely equivalent physically. This means that, as a result of this interchange, the wave function of the system can change only by an unimportant phase factor. Let $\psi(\xi_1, \xi_2)$ be the wave function of the system, ξ_1 and ξ_2 conventionally denoting the three coordinates and the spin projection for each particle. Then we must have

$$\psi(\xi_1, \xi_2) = e^{i\alpha} \psi(\xi_2, \xi_1), \quad (26.1)$$

where α is some real constant. By repeating the interchange, we return to the original state while the function ψ gets multiplied by $e^{2i\alpha}$. Hence it follows that $e^{2i\alpha} = 1$, or, $e^{i\alpha} = \pm 1$. Thus

$$\psi(\xi_1, \xi_2) = \pm \psi(\xi_2, \xi_1). \quad (26.2)$$

We thus reach the result that there are only two possibilities: the wave function is either *symmetrical* (i.e., it is unchanged when the particles are interchanged) or *antisymmetrical* (i.e., changes sign when this interchange is made). It is obvious that the wave functions of all the states of a given system must have the same symmetry; otherwise the wave function of a state which was a superposition of states of different symmetry would be neither symmetrical nor antisymmetrical.

This result can be immediately generalized to systems consisting of any number of identical particles. For it is clear from the identity of the particles that, if any pair of them has the property of being described by, say, symmetrical wave functions, any other pair of such

particles has the same property. Hence the wave function of identical particles must either be unchanged when any pair of particles are interchanged (and hence when the particles are permuted in any manner), or change sign when any pair are interchanged. In the first case we speak of a symmetrical wave function, and in the second case of an antisymmetrical one.

26.1.2 Spin Statistics Theorem

The property of being described by symmetrical or antisymmetrical wave functions depends on the nature of the particles. Particles described by antisymmetrical functions are said to obey Fermi-Dirac statistics (or to be fermions), while those which are described by symmetrical functions are said to obey Bose-Einstein statistics (or to be bosons)?

According to the laws of relativistic quantum mechanics, the statistics obeyed by particles is uniquely related to their spin: particles with half-integral spin are fermions, and those with integral spin are bosons.

26.1.3 Statistics of Composite Particles

The statistics of complex particles is determined by the number of elementary fermions entering into their composition. For, an interchange of two identical complex particles is equivalent to the simultaneous interchange of several pairs of identical elementary particles. The interchange of bosons does not change the wave function, while the interchange of fermions changes its sign. Hence complex particles containing an odd number of elementary fermions obey Fermi statistics, while those containing an even number obey Bose statistics. This result is, of course, in agreement with the above rule, since a complex particle has an integral or a half-integral spin according as the number of particles with half-integral spin entering into its composition is even or odd.

Thus atomic nuclei of odd mass number (i.e. containing an odd number of neutrons and protons) obey Fermi statistics, and those of even atomic number obey Bose statistics. For atoms, which contain both nuclei and electrons, the statistics is evidently determined by the parity of the sum of the mass number and the atomic number, i.e., $(-1)^{(A+Z)}$.

26.1.4 Operators for a system of identical particles

To say that two particles are identical means that there are no interactions that can distinguish them. Thus any operator corresponding to a physical observable of a collection of N identical particles must treat all the particles on the same footing. Therefore, the operator must be a symmetric function of the coordinates.

Thus, for example, the Hamiltonian $H(\xi_1, \xi_2, \dots, \xi_N)$ must remain unchanged under the interchange of any two particles, and hence under any permutation of coordinates. One calls such an operator symmetric operator.

A system of N identical particles starting out in a completely symmetric, or antisymmetric state, must always remain in such a state. This is because any perturbation $V(\xi_1, \xi_2, \dots, \xi_N)$ that can act to change the state of the system, must be a completely symmetric function of the coordinates of the N particles and therefore commutes with the permutation operator, i.e.,

$$[\hat{P}, \hat{V}] = 0.$$

Therefore, if $\hat{P}|\psi\rangle = \pm|\psi\rangle$, then

$$\hat{P}\hat{V}|\psi\rangle = \hat{V}(\hat{P}|\psi\rangle) = \pm\hat{V}|\psi\rangle,$$

so that $\hat{V}|\psi\rangle$ has the same symmetry as $|\psi\rangle$.

26.2 States of Noninteracting Identical Particles

Suppose we have N noninteracting identical particles each in the same potential well $V(r)$. The Hamiltonian of a single particle, say particle 1, in the well is

$$\hat{H}_0(1) = \frac{\hat{p}_1^2}{2m} + V(\xi_1). \quad (26.3)$$

Let us write the orthonormal energy eigenstates of this Hamiltonian as $\phi_1(\xi_1), \phi_2(\xi_1), \dots$ with corresponding energies $\epsilon_1, \epsilon_2, \dots$. The Hamiltonian for the N particles in the well is just the sum of the individual energy operators of the particles, i.e.,

$$\hat{H} = \hat{H}_0(1) + \hat{H}_0(2) + \dots + \hat{H}_0(N). \quad (26.4)$$

It is easy to write down solutions of the Schrödinger equation

$$\hat{H}\psi(\xi_1, \xi_2, \dots, \xi_N) = E\psi(\xi_1, \xi_2, \dots, \xi_N). \quad (26.5)$$

For example,

$$\psi(\xi_1, \xi_2, \dots, \xi_N) = \phi_{\alpha_1}(\xi_1)\phi_{\alpha_2}(\xi_2)\cdots\phi_{\alpha_N}(\xi_N) \quad (26.6)$$

is a solution which corresponds to the first particle in state α_1 with energy ε_1 , the second particle in state α_2 with energy ε_2 , etc. The total energy of the system is simply

$$E = \varepsilon_{\alpha_1} + \varepsilon_{\alpha_2} + \cdots + \varepsilon_{\alpha_N}. \quad (26.7)$$

In general, Eq. (26.6) is not an admissible solution for N identical particles since it lacks the symmetry under interchange of any two particles. There are many other solutions to Eq. (26.5). Any permutations of $\xi_1, \xi_2, \dots, \xi_n$ on the right hand side of Eq. (26.6) yields, in general, another solution to Eq. (26.5) with the same energy for the system of N identical particles.

To construct admissible eigenstates for N identical particles, we must take linear combination of states of the form Eq. (26.6) that is completely symmetric for bosons or completely antisymmetric for fermions.

26.2.1 Identical Fermions

For two identical fermions, the antisymmetric state is

$$\psi_a(1, 2) = \frac{1}{\sqrt{2}} [\phi_{\alpha_1}(1)\phi_{\alpha_2}(2) - \phi_{\alpha_1}(2)\phi_{\alpha_2}(1)] \quad (26.8)$$

and, generally for N particles

$$\psi_a(1, 2, \dots, N) = \frac{1}{\sqrt{N!}} \sum_p (-1)^P P \phi_{\alpha_1}(1)\phi_{\alpha_2}(2)\cdots\phi_{\alpha_1}(N) \quad (26.9)$$

where the sum is over all the $N!$ permutations of the arguments $1, 2, \dots, N$ of the single particle wave functions. These arguments represent the space and the spin coordinates $\xi_1, \xi_2, \dots, \xi_N$ of the fermions. If the permutation is even, i.e., if the permutation can be obtained by an even number of interchanges of pairs of particles (i.e., even number of

transpositions), then there is a plus sign. If the permutation is odd, there is a minus sign. Eq. (26.9) can be written in the determinant form

$$\psi_a(1, 2, \dots, N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \phi_{\alpha_1}(1) & \phi_{\alpha_1}(2) & \cdots & \phi_{\alpha_1}(N) \\ \phi_{\alpha_2}(1) & \phi_{\alpha_2}(2) & \cdots & \phi_{\alpha_2}(N) \\ \vdots & \vdots & & \vdots \\ \phi_{\alpha_N}(1) & \phi_{\alpha_N}(2) & \cdots & \phi_{\alpha_N}(N) \end{vmatrix}. \quad (26.10)$$

These determinants of one-particle states are called Slater determinants. The antisymmetry of Eq. (26.10) is immediately apparent since interchange of any two columns introduces a factor of -1 . The normalization is $1/\sqrt{N!}$ since Eq. (26.10) consists of $N!$ mutually orthogonal terms.

From the Slater determinant it is also apparent that if two or more sets of individual states are identical, i.e., $\alpha_i = \alpha_j$ etc., the wave function (26.10) vanishes. As a result only one fermion can occupy a given individual quantum state. This statement is known as **Pauli exclusion principle**.

26.2.2 Identical Bosons

Let $\alpha_1, \alpha_2, \dots, \alpha_N$ be the quantum numbers of the single particle states occupied by N identical non-interacting bosons. Some of the quantum numbers may be the same. For a system of bosons, the wave function $\psi(1, 2, \dots, N)$ is given by a sum of products of the form

$$\phi_{\alpha_1}(1)\phi_{\alpha_2}(2)\cdots\phi_{\alpha_N}(N) \quad (26.11)$$

with all possible permutations of the indices $(1, 2, \dots, N)$. This sum clearly possesses the required symmetry property. Thus, for example, for a system of two particles in different states ($\alpha_1 \neq \alpha_2$)

$$\psi(1, 2) = \frac{1}{\sqrt{2}} [\phi_{\alpha_1}(1)\phi_{\alpha_2}(2) + \phi_{\alpha_1}(2)\phi_{\alpha_2}(1)]. \quad (26.12)$$

The factor $1/\sqrt{2}$ is introduced for normalization purposes; all the functions ϕ_1, ϕ_2, \dots are orthogonal and are supposed normalized. In the general case of a system containing an arbitrary number N of particles, the normalized wave function is

$$\psi = \left(\frac{N_1!N_2!\dots}{N!} \right)^{1/2} \sum P \phi_{\alpha_1}(1)\phi_{\alpha_2}(2)\cdots\phi_{\alpha_N}(N), \quad (26.13)$$

where the sum is taken over all distinct permutations of the particles. The numbers N_i represent how many of the suffixes have the same value α_i , i.e., how many identical bosons are accommodated in the same single particle state ϕ_{α_i} . Obviously, $\sum N_i = N$. In the integration of $|\psi|^2$ over the space and spin coordinates $\xi_1, \xi_2, \dots, \xi_N$, all terms vanish except the squared modulus of each term. Since the total number of terms in the sum (26.13) is evidently

$$\frac{N!}{N_1!N_2!\dots},$$

the normalization factor in (26.13) is obtained.

Example

Griffiths page 217

Suppose we have two noninteracting particles, both of mass m in an infinite square well. The one-particle states are:

$$\psi_n(x) = \sqrt{\frac{2}{a}} \sin\left(\frac{n\pi}{a}x\right), \quad E_n = n^2 K$$

where $K = \pi^2 \hbar^2 / 2ma^2$. If the particles are distinguishable, with particle 1 in state n_1 and particle 2 in state n_2 , the composite wave function is the simple product

$$\psi_{n_1 n_2}(x_1, x_2) = \psi_{n_1}(x_1) \psi_{n_2}(x_2), \quad E_{n_1 n_2} = (n_1^2 + n_2^2)K.$$

For example, the ground state is

$$\psi_{11} = \frac{2}{a} \sin(\pi x_1/a) \sin(\pi x_2/a), \quad E_{11} = 2K;$$

the first excited state is doubly degenerate

$$\psi_{12} = \frac{2}{a} \sin(\pi x_1/a) \sin(2\pi x_2/a), \quad E_{12} = 5K,$$

$$\psi_{21} = \frac{2}{a} \sin(2\pi x_1/a) \sin(\pi x_2/a), \quad E_{21} = 5K;$$

and so on.

If the two particles are identical bosons, the ground state is unchanged, but the first excited state is nondegenerate:

$$\frac{\sqrt{2}}{a} [\sin(\pi x_1/a) \sin(2\pi x_2/a) + \sin(2\pi x_1/a) \sin(\pi x_2/a)]$$

still with energy $5K$.

And, if the particles are identical fermions, there is no state with energy $2K$; the ground state is

$$\frac{\sqrt{2}}{a} [\sin(\pi x_1/a) \sin(2\pi x_2/a) - \sin(2\pi x_1/a) \sin(\pi x_2/a)],$$

and its energy is $5K$.

26.2.3 Exchange Forces

From Griffiths

Consider a state consisting of two non interacting particles. Suppose one particle is in the state ψ_a and the other in the state ψ_b . The two states are orthogonal and normalized. If the two particles are distinguishable, and particle number 1 is in the state ψ_a , then the combined wave function is

$$\psi(x_1, x_2) = \psi_a(x_1) \psi_b(x_2). \quad (26.14)$$

If they are identical bosons, the composite wave function is

$$\psi_+(x_1, x_2) = \frac{1}{\sqrt{2}} [\psi_a(x_1) \psi_b(x_2) + \psi_b(x_1) \psi_a(x_2)]; \quad (26.15)$$

and if they are identical fermions, the wave function is

$$\psi_-(x_1, x_2) = \frac{1}{\sqrt{2}} [\psi_a(x_1) \psi_b(x_2) - \psi_b(x_1) \psi_a(x_2)]. \quad (26.16)$$

Let us calculate the expectation value of the square of the separation distance between the two particles:

$$\langle (x_1 - x_2)^2 \rangle = \langle x_1^2 \rangle + \langle x_2^2 \rangle - 2\langle x_1 x_2 \rangle. \quad (26.17)$$

Case 1: Distinguishable Particles.

For the wave function in Equation (26.14), we have

$$\langle x_1^2 \rangle = \int x_1^2 |\psi_a(x_1)|^2 dx_1 \int |\psi_b(x_2)|^2 dx_2 = \langle x^2 \rangle_a$$

i.e., the expectation value of x^2 in the one-particle state ψ_a . We also have

$$\langle x_2^2 \rangle = \int |\psi_a(x_1)|^2 dx_1 \int x_2^2 |\psi_b(x_2)|^2 dx_2 = \langle x^2 \rangle_b,$$

and

$$\langle x_1 x_2 \rangle = \int x_1 |\psi_a(x_1)|^2 dx_1 \int x_2 |\psi_b(x_2)|^2 dx_2 = \langle x \rangle_a \langle x \rangle_a.$$

In this case, then

$$\langle (x_1 - x_2)^2 \rangle_d = \langle x^2 \rangle_a + \langle x^2 \rangle_b - 2\langle x \rangle_a \langle x \rangle_b. \quad (26.18)$$

Incidentally, the answer would, of course, be the same if particle 1 had been in state ψ_b , and particle 2 in state ψ_a .

Case 2: Identical Particles

For the wave functions in Equations (26.15) and (26.16),

$$\begin{aligned} \langle x_1^2 \rangle &= \frac{1}{2} \left[\int x_1^2 |\psi_a(x_1)|^2 dx_1 \int |\psi_b(x_2)|^2 dx_2 \right. \\ &\quad + \int x_1^2 |\psi_b(x_1)|^2 dx_1 \int |\psi_a(x_2)|^2 dx_2 \\ &\quad \pm \int x_1^2 \psi_a(x_1)^* \psi_b(x_1) dx_1 \int \psi_b(x_2)^* \psi_a(x_2) dx_2 \\ &\quad \pm \int x_1^2 \psi_b(x_1)^* \psi_a(x_1) dx_1 \int \psi_a(x_2)^* \psi_b(x_2) dx_2 \left. \right] \\ &= \frac{1}{2} [\langle x^2 \rangle_a + \langle x^2 \rangle_b \pm 0 \pm 0] = \frac{1}{2} (\langle x^2 \rangle_a + \langle x^2 \rangle_b). \end{aligned}$$

Similarly,

$$\langle x_2^2 \rangle = \frac{1}{2} (\langle x^2 \rangle_b + \langle x^2 \rangle_a).$$

(Naturally, $\langle x_1^2 \rangle = \langle x_2^2 \rangle$, since one cannot tell them apart). But

$$\begin{aligned} \langle x_1 x_2 \rangle &= \frac{1}{2} \left[\int x_1 |\psi_a(x_1)|^2 dx_1 \int x_2 |\psi_b(x_2)|^2 dx_2 \right. \\ &\quad + \int x_1 |\psi_b(x_1)|^2 dx_1 \int x_2 |\psi_a(x_2)|^2 dx_2 \\ &\quad \pm \int x_1 \psi_a(x_1)^* \psi_b(x_1) dx_1 \int x_2 \psi_b(x_2)^* \psi_a(x_2) dx_2 \\ &\quad \pm \int x_1 \psi_b(x_1)^* \psi_a(x_1) dx_1 \int x_2 \psi_a(x_2)^* \psi_b(x_2) dx_2 \left. \right] \\ &= \frac{1}{2} [\langle x \rangle_a \langle x \rangle_b + \langle x \rangle_b \langle x \rangle_a \pm \langle x \rangle_{ab} \langle x \rangle_{ba} \pm \langle x \rangle_{ba} \langle x \rangle_{ab}] \\ &= \langle x \rangle_a \langle x \rangle_b \pm |\langle x \rangle_{ab}|^2, \end{aligned}$$

where

$$\langle x \rangle_{ab} = \int x \psi_a(x)^* \psi_b(x) dx. \quad (26.19)$$

Evidently

$$\langle (x_1 - x_2)^2 \rangle_{\pm} = \langle x^2 \rangle_a + \langle x^2 \rangle_b - 2\langle x \rangle_a \langle x \rangle_b \mp 2|\langle x \rangle_{ab}|^2. \quad (26.20)$$

Comparing Eqs. (26.18) and (26.20), we see that the difference lies in the final term:

$$\langle (\Delta x)^2 \rangle_{\pm} = \langle (\Delta x)^2 \rangle_d \mp 2|\langle x \rangle_{ab}|^2. \quad (26.21)$$

Identical bosons (the upper signs) tend to be somewhat closer together, and identical fermions (the lower signs) somewhat further apart, than distinguishable particles in the same two states. Notice that $\langle x \rangle_{ab}$ vanishes unless the two wave functions actually overlap [if $\psi_a(x)$ is zero wherever $\psi_b(x)$ is nonzero, the integral in Eq. (26.19) is itself zero]. So, if ψ_a represents an electron in an atom in Chicago and ψ_b represents an electron in an atom in Seattle, it's not going to make a difference if you antisymmetrize the wave function or not. As a practical matter, therefore it's okay to pretend that electrons with nonoverlapping wave functions are distinguishable.

The interesting case is when there is some overlap of the wave functions. The system behaves as though there were a "force of attraction" between identical bosons pulling them closer together, and a "force of repulsion" between identical fermions pushing them apart. We call it an **exchange force**, although it is not really a force at all — no physical agency is pushing on the particles; rather, it is purely a geometrical consequence of the symmetrization requirement. It is also a strictly quantum mechanical phenomenon, with no classical counterpart. Nevertheless, it has profound consequences.

Consider, for example, the hydrogen molecule (H_2). Roughly speaking, the ground state consists of one electron in the atomic ground state centered on nucleus 1, and one electron in the atomic ground state centered at nucleus 2. If electrons were *bosons*, the symmetrization requirement (or, if you like the "exchange force") would tend to concentrate the electrons toward the middle between the two protons (Figure (26.1 a), and the resulting accumulation of negative charge would attract the protons inward, accounting for the **covalent bond** that

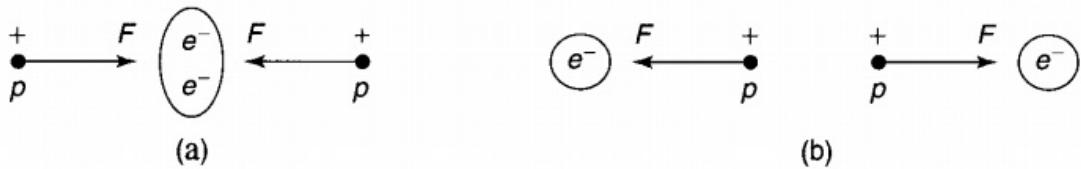


Figure 26.1: Schematic picture of the covalent bond: (a) symmetric configuration produces attractive force; (b) antisymmetric configuration produces repulsive force.

holds the molecule together. Unfortunately, electrons aren't bosons, they are fermions, and this means that the concentration of negative charge should actually be shifted to the wings (Figure 26.1 b), tearing the molecule apart.

But, wait. We have been ignoring *spin*. The complete state of an electron includes not only its position wave function, but also a spinor, describing the orientation of its spin¹:

$$\psi(\vec{r})\chi(s). \quad (26.22)$$

When we put together the two-electron state, it is the *whole works*, not just the spatial part, that has to be antisymmetric with respect to exchange. Now, for a two-particle system where each particle has spin quantum number $s = 1/2$, the singlet combination is antisymmetric (and hence would have to be joined with a *symmetric* spatial function), whereas the three triplet states are all symmetric (and would require an *antisymmetric* spatial function). Evidently, then, the singlet state should lead to bonding, and triplet to antibonding. Sure enough, the chemists tell us that covalent bonding requires the two electrons to occupy the singlet state, with total spin zero.²

26.3 The Helium Atom

The basic Hamiltonian of the helium atom is given by

$$\hat{H} = -\frac{\hbar^2}{2m}\nabla_1^2 - \frac{\hbar^2}{2m}\nabla_2^2 - \frac{1}{4\pi\epsilon_0}\frac{2e^2}{r_1} - \frac{1}{4\pi\epsilon_0}\frac{2e^2}{r_2} + \frac{1}{4\pi\epsilon_0}\frac{e^2}{r_{12}}. \quad (26.23)$$

¹In the absence of coupling between spin and position, we are free to assume that the state is separable in its spin and spatial coordinates. This just says that the probability of getting spin up is independent of the location of the particle. In the presence of coupling, the general state of a particle is a linear combination $\psi_+(\vec{r})\chi_+ + \psi_-(\vec{r})\chi_-$.

²In casual language, it is often said that in the singlet state the two electrons are "oppositely aligned" (one with spin up and the other with spin down). This is something of an over simplification, since the same could be said for the $m = 0$ triplet state. The precise statement is that they are in the singlet configuration.

Suppose the $e^2/4\pi\epsilon_0 r_{12}$ term were absent. Then, with the identity of the electron ignored, the wave function would be just the product of two hydrogenic wave functions with $Z = 1$ changed to $Z = 2$. The total spin is a constant of motion, so that the spin state is either a singlet (antisymmetric) or a triplet (symmetric). The spin states are

$$^1\chi_0 = \frac{1}{\sqrt{2}} [\alpha(1)\beta(2) - \alpha(2)\beta(1)] \quad (\text{antisymmetric}) \quad (26.24)$$

$$^3\chi_1 = \alpha(1)\alpha(2) \quad (\text{symmetric}) \quad (26.25)$$

$$^3\chi_0 = \frac{1}{\sqrt{2}} [\alpha(1)\beta(2) + \alpha(2)\beta(1)] \quad (\text{symmetric}) \quad (26.26)$$

$$^3\chi_{-1} = \beta(1)\beta(2) \quad (\text{symmetric}). \quad (26.27)$$

Suppose that one of the electrons is in the ground state $(nlm) = (100)$, and the other in an excited state (nlm) with $n \neq 1$. Then the spatial wave function of the two electrons would be

$$\phi(\vec{r}_1, \vec{r}_2) = \frac{1}{\sqrt{2}} [\phi_{100}(\vec{r}_1)\phi_{nlm}(\vec{r}_2) \pm \phi_{100}(\vec{r}_2)\phi_{nlm}(\vec{r}_1)] \quad (26.28)$$

where the upper sign is for the spin-singlet state and the lower sign is for the spin-triplet states.

Ground State

First, we consider the ground state. In this case both the electrons are in the orbital state $n = 1, l = 0, m = 0$, i.e., the orbital configuration is $(1s)^2$. In this configuration the space part of the wave function of the two-electron system must necessarily be symmetric. Therefore, the spin state of the system must be the antisymmetric singlet so that the overall wave function which is the product of the spatial wave function and the spin wave function is antisymmetric. So we have

$$\begin{aligned} \psi(\vec{r}_2, \vec{r}_2) &= \psi_{100}(\vec{r}_1)\psi_{100}(\vec{r}_2) ^1\chi_0 \\ &= \frac{Z^3}{\pi a_0^3} e^{-Z(r_1+r_2)/a_0} ^1\chi_0 \quad (Z = 2). \end{aligned} \quad (26.29)$$

This unperturbed wave function gives

$$E = 2 \times 4 \left(-\frac{e^2}{(4\pi\epsilon_0)2a_0} \right) = -8E_{\text{Ry}} = -8 \times 13.6 \text{ eV} = -109 \text{ eV.}$$

for the ground state energy, which is about 30% larger in magnitude than the experimental value of -78.8 eV.

Next, in order to obtain a better value for the ground state energy we take into account the effect of the perturbing potential $e^2/4\pi\varepsilon_0 r_{12}$ in the Hamiltonian. In the first-order perturbation theory,

$$\begin{aligned}
 \Delta E = E^{(1)} &= \langle 1s^2 | e^2 / 4\pi\varepsilon_0 r_{12} | 1s^2 \rangle (\chi_0^\dagger \chi_0) \\
 &= \int \psi_{100}^*(\vec{r}_1) \psi_{100}^*(\vec{r}_2) \frac{e^2}{4\pi\varepsilon_0 r_{12}} \psi_{100}(\vec{r}_1) \psi_{100}(\vec{r}_2) d^3 r_1 d^3 r_2 \\
 &= \frac{Z^6}{\pi^2 a_0^6} \int e^{-2Z(r_1+r_2)/a_0} \frac{e^2}{4\pi\varepsilon_0 r_{12}} d^3 r_1 d^3 r_2 \quad (Z=2) \\
 &= \left(\frac{5}{2}\right) \left(\frac{e^2}{4\pi\varepsilon_0 2a_0}\right) = \left(\frac{5}{2}\right) E_{\text{Ry}}. \tag{26.30}
 \end{aligned}$$

After adding this energy shift, we have

$$E_{(1s)^2} = (-8 + 5/2) E_{\text{Ry}} = -74.8 \text{ eV},$$

which is very close to the experimental value

$$E_{\text{exp}} = -78.8 \text{ eV}.$$

Excited State

Let us briefly consider the excited states. We consider only the case in which one of the electrons is in the single particle ground state (100), i.e., the 1s state, and the other electron in the excited state (nlm) with $n \neq 1$. We write the energy of the state as

$$E = E_{100} + E_{nlm} + \Delta E$$

where

$$E_{100} = E_{1s} = -Z^2 \frac{e^2}{4\pi\varepsilon_0 2a_0}$$

and for hydrogenic atoms

$$E_{nlm} = E_n = -Z^2 \left(\frac{e^2}{4\pi\varepsilon_0 2a_0}\right) \frac{1}{n^2}.$$

The energy correction ΔE is

$$\Delta E = \int \psi^*(\chi; \vec{r}_1 \vec{r}_2) \frac{e^2}{4\pi\varepsilon_0 r_{12}} \psi(\chi; \vec{r}_1 \vec{r}_2) \tag{26.31}$$

where χ represents the spin state of the two electrons. We have

$$\psi(\chi; \vec{r}_1 \vec{r}_2) = \psi(\vec{r}_1 \vec{r}_2) \chi$$

and

$$\psi^*(\chi; \vec{r}_1 \vec{r}_2) = \chi^\dagger \psi^*(\vec{r}_1 \vec{r}_2).$$

Since the operator $e^2/4\pi\epsilon_0 r_{12}$ does not involve spin, it is clear that singlet and triplet states are separated and that the three triplet spin states remain degenerate.

For the singlet (triplet) states we have

$$\begin{aligned} \Delta E_{\text{singlet}} &= \int \frac{1}{\sqrt{2}} [\psi_{100}(\vec{r}_1) \psi_{nlm}(\vec{r}_2) \pm \psi_{100}(\vec{r}_2) \psi_{nlm}(\vec{r}_1)]^* \frac{e^2}{4\pi\epsilon_0 r_{12}} \\ &\quad \times \frac{1}{\sqrt{2}} [\psi_{100}(\vec{r}_1) \psi_{nlm}(\vec{r}_2) \pm \psi_{100}(\vec{r}_2) \psi_{nlm}(\vec{r}_1)] d^3 r_1 d^3 r_2 \end{aligned} \quad (26.32)$$

We can write the above equation as

$$\begin{aligned} \Delta E_{\text{singlet}} &= I \pm J \\ \Delta E_{\text{triplet}} &= I \end{aligned} \quad (26.33)$$

where

$$I = \int d^3 r_1 \int d^3 r_2 |\psi_{100}(\vec{r}_1)|^2 |\psi_{nlm}(\vec{r}_2)|^2 \frac{e^2}{4\pi\epsilon_0 r_{12}} \quad (26.34)$$

and

$$J = \int d^3 r_1 \int d^3 r_2 \psi_{100}^*(\vec{r}_1) \psi_{nlm}^*(\vec{r}_2) \frac{e^2}{4\pi\epsilon_0 r_{12}} \psi_{100}(\vec{r}_2) \psi_{nlm}(\vec{r}_1) \quad (26.35)$$

The upper (lower) sign goes with spin singlet (triplet) states. Obviously, I is positive. We can show that J is also positive. So the net result is such that for the same configuration $(100)^1(nlm)^1$, the spin singlet state lies higher, as shown in the figure below.

The physical interpretation for this is as follows: In the singlet case, the space function is symmetric and the electrons have a tendency to come close to each other. Therefore, the electrostatic repulsion between the electrons is more serious; hence a higher energy results. In the triplet case, the space function is antisymmetric and the electrons tend to avoid each other.

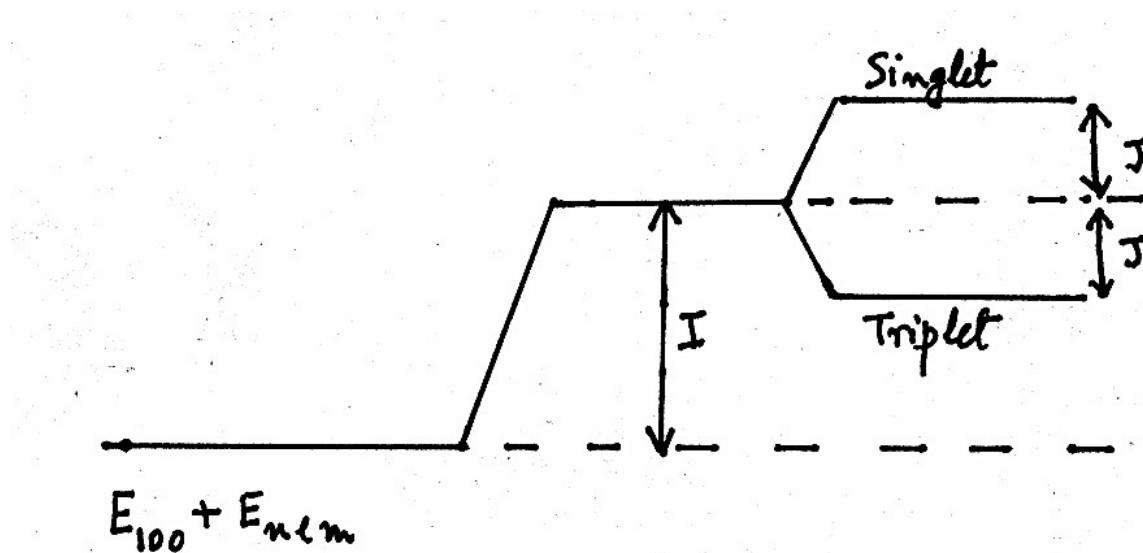


Figure 26.2: Removal of degeneracy between the singlet and triplet states of helium atom in the first order perturbation.

Helium in spin-singlet state is known as parahelium, while helium in the spin-triplet state is known as orthohelium. Each configuration of the two electrons (except for the ground state configuration) splits into the para state and the ortho state, the para state lying higher in energy. For the ground state, only parahelium is possible. The figure below is a schematic energy level diagram for the low-lying configurations of the helium atom.

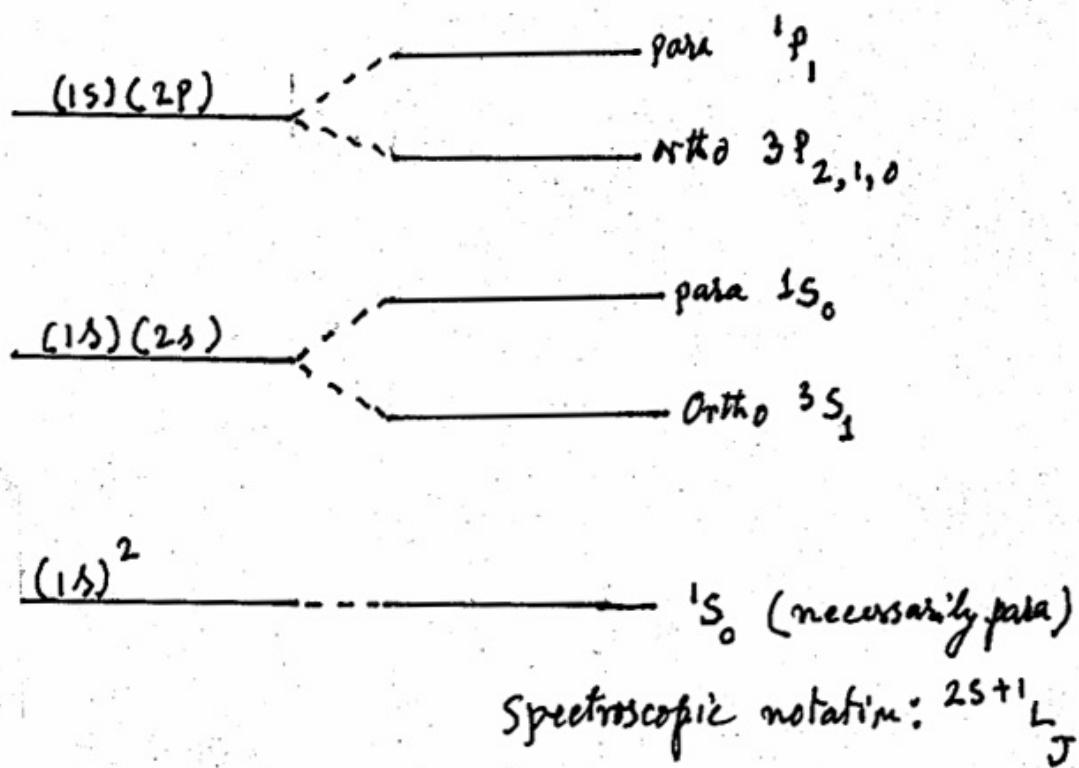


Figure 26.3: Energy levels for the low-lying configurations of the helium atom. The energy levels on the left of the diagram shows the mean of the ortho and para states, i.e., what the energy levels would have been in a given configuration if the electrons were distinguishable particles.

26.4 Collision Between Identical Particles

Read Bransden and Joachain, page 620.

Collisions between identical particles are particularly interesting as a direct illustration of the fundamental differences between classical and quantum mechanics. We shall examine first the elastic scattering of two identical spinless bosons and then analyze elastic collisions between two identical spin-1/2 fermions.

26.4.1 Scattering of Two Identical Spinless Bosons

Let us consider the elastic scattering of two identical bosons of mass m . For simplicity we shall consider only the case of spinless bosons. We work in the centre-of-mass system, in which the time-independent Schrodinger equation is

$$\left[-\frac{\hbar^2}{2\mu} + V(r) \right] \psi(\vec{r}) = E \psi(\vec{r}) \quad (26.36)$$

where $\mu = m/2$ is the reduced mass and $\vec{r} = \vec{r}_1 - \vec{r}_2$ is the relative position vector of the two colliding particles. The situation in the center-of-mass system is illustrated in Fig. (26.4).

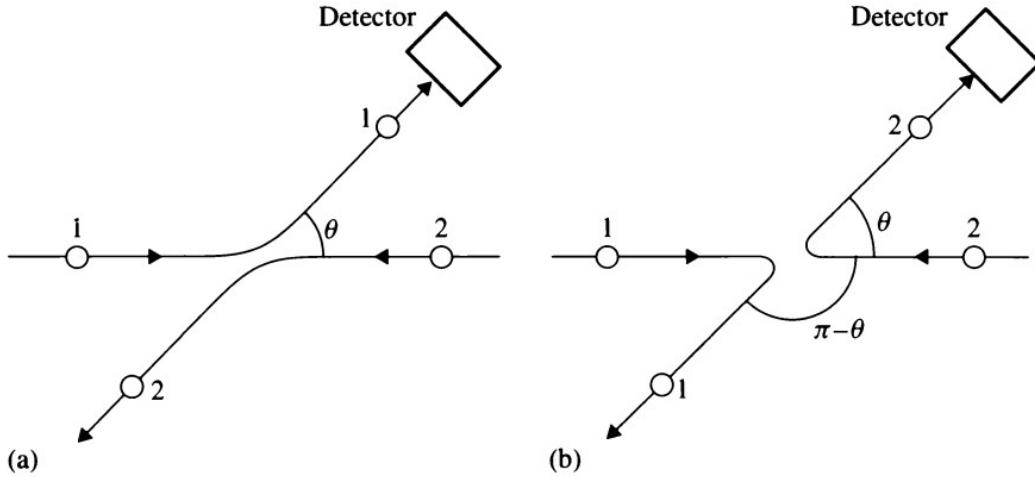


Figure 26.4: The scattering of Identical particles in the centre-of-mass frame.

Two identical particles 1 and 2 approach one another, moving parallel to the z-axis in opposite directions. After an elastic collision the velocity of each particle is changed in direction but remains unchanged in magnitude. A detector counts the particles scattered into the direction characterized by the polar angles (θ, ϕ) . Since the particles 1 and 2 are identical, there is no way of deciding whether a particle recorded by the detector results from a collision event in which particle 1 is scattered in the direction (θ, ϕ) (see Fig. 26.4 (a)), or from a collision process in which particle 2 is scattered in that direction, so that particle 1 is scattered in the opposite direction $(\pi - \theta, \phi + \pi)$ (see Fig. 26.4 (b)).

In classical mechanics the differential cross-section for scattering in the direction (θ, ϕ) would simply be the sum of the differential cross-sections for observation of particle 1 and particle 2 in that direction. If the same were to be true in quantum mechanics, we would obtain for the differential cross-section the ‘classical’ result

$$\frac{d\sigma_{cl}}{d\Omega} = |f(\theta, \phi)|^2 + |f(\pi - \theta, \phi + \pi)|^2 \quad (26.37)$$

where $f(\theta, \phi)$ is the center-of-mass amplitude for scattering in the direction (θ, ϕ) . The amplitude $f(\theta, \phi)$ is related to the asymptotic behavior of the wave function $\psi(\vec{r})$ satisfying

the usual boundary condition

$$\psi(\vec{r}) \sim_{r \rightarrow \infty} e^{ikz} + f(\theta, \phi) \frac{e^{ikr}}{r}. \quad (26.38)$$

However, we shall now show that the expression (26.37) is incorrect.

Indeed, we have seen that wave functions describing systems of identical particles must be properly symmetrized with respect to permutations of the particles. In particular, a wave function describing a system of identical bosons must be completely symmetric. Thus, in the case of two identical spinless bosons, the wave function must be symmetric under the interchange of the spatial coordinates of the two particles. Now the interchange $\vec{r}_1 \leftrightarrow \vec{r}_2$ corresponds to replacing the relative position vector \vec{r} by $-\vec{r}$, which in polar coordinates corresponds to (r, θ, ϕ) being replaced by $(r, \pi - \theta, \phi + \pi)$. The wave function $\psi(\vec{r})$ satisfying the boundary condition (26.38) does not have the required symmetry, but the symmetric combination

$$\psi_+(\vec{r}) = \psi(\vec{r}) + \psi(-\vec{r})$$

is also a solution of the Schrödinger equation and does have the required symmetry, i.e., $\psi_+(-\vec{r}) = \psi_+(\vec{r})$. Using Eq. (26.38), the asymptotic form of $\psi_+(\vec{r})$ is

$$\psi_+(\vec{r}) \sim_{r \rightarrow \infty} \left[e^{ikz} + e^{-ikz} \right] + [f(\theta, \phi) + f(\pi - \theta, \phi + \pi)] \frac{e^{ikr}}{r}. \quad (26.39)$$

The amplitude of the spherically outgoing wave is the symmetric amplitude

$$f_+(\theta, \phi) = f(\theta, \phi) + f(\pi - \theta, \phi + \pi), \quad (26.40)$$

so that the differential cross section is

$$\frac{d\sigma}{d\Omega} = |f(\theta, \phi) + f(\pi - \theta, \phi + \pi)|^2, \quad (26.41)$$

a result which we can write in the form

$$\frac{d\sigma}{d\Omega} = |f(\theta, \phi)|^2 + |f(\pi - \theta, \phi + \pi)|^2 + 2 \operatorname{Re} [f(\theta, \phi)f^*(\pi - \theta, \phi + \pi)]. \quad (26.42)$$

It is important to note that this formula differs from the ‘classical’ result (26.37) by the presence of the third term on the right, which arises from the interference between the amplitudes $f(\theta, \phi)$ and $f(\pi - \theta, \phi + \pi)$.

In the simple case for which the interaction potential is central, the scattering amplitude is independent of the azimuthal angle ϕ . The differential cross-section (26.41) then reduces to

$$\begin{aligned}\frac{d\sigma}{d\Omega} &= |f(\theta) + f(\pi - \theta)|^2 \\ &= |f(\theta)|^2 + |f(\pi - \theta)|^2 + 2\operatorname{Re}[f(\theta)f^*(\pi - \theta)].\end{aligned}\quad (26.43)$$

From Eq. (26.43) we notice that the scattering in the CM frame is symmetric about the angle $\theta = \pi/2$. Moreover, we note that, at $\theta = \pi/2$, the quantum mechanical differential cross section is equal to

$$\frac{d\sigma}{d\Omega}(\theta = \pi/2) = 4|f(\theta = \pi/2)|^2 \quad (26.44)$$

and hence four times as big as if the colliding particles were distinguishable and twice as big as the classical result

$$\frac{d\sigma_{cl}}{d\Omega}(\theta = \pi/2) = 2|f(\theta = \pi/2)|^2. \quad (26.45)$$

26.4.2 Partial wave expansion for the scattering amplitude of two spinless bosons

The amplitude of scattering of two identical spinless bosons is

$$f_+(\theta, \phi) = f(\theta, \phi) + f(\pi - \theta, \phi + \phi). \quad (26.46)$$

For central potentials, we have azimuthal symmetry, i.e., the scattering amplitude is independent of ϕ . Now, the partial wave expansion of the scattering amplitude can be written as

$$\begin{aligned}f_+(\theta) &= f(\theta) + f(\pi - \theta) \\ &= \frac{1}{k} \sum_{l=0}^{\infty} (2l+1)e^{i\delta_l} \sin \delta_l P_l(\cos \theta) + \frac{1}{k} \sum_{l=0}^{\infty} (2l+1)e^{i\delta_l} \sin \delta_l P_l(\cos(\pi - \theta)).\end{aligned}\quad (26.47)$$

Since

$$P_l(\cos(\pi - \theta)) = P_l(-\cos \theta) = (-)^l P_l(\cos \theta),$$

we have

$$\begin{aligned}f_+(\theta) &= f(\theta) + f(\pi - \theta) \\ &= \frac{1}{k} \sum_{l=0}^{\infty} (2l+1)e^{i\delta_l} \sin \delta_l \left[1 + (-1)^l \right] P_l(\cos \theta) \\ &= \frac{2}{k} \sum_{l=\text{even}}^{\infty} (2l+1)e^{i\delta_l} \sin \delta_l P_l(\cos \theta).\end{aligned}\quad (26.48)$$

The partial wave expansion of $f_+(\theta)$ contains only even orbital angular momentum quantum number l , i.e., $l = 0, 2, 4, \dots$. At low energies, the s-wave ($l = 0$) dominates. We then have

$$f_+(\theta) \approx \frac{2}{k} e^{i\delta_0} \sin \delta_0 \quad (\text{low energy}),$$

i.e., the scattering is isotropic in the CM frame. The differential cross section for elastic scattering is

$$\frac{d\sigma}{d\Omega} = |f_+(\theta)|^2 = \frac{4}{k^2} \sin^2 \delta_0. \quad (26.49)$$

The total (i.e., integrated) elastic scattering scattering cross section is found by integrating the elastic differential cross section over all solid angles. The integration is very easy since the differential cross section does not depend upon the scattering angle. We have

$$\sigma_{\text{el}} = \frac{16\pi}{k^2} \sin^2 \delta_0. \quad (26.50)$$

26.5 Scattering of Two Identical Spin-1/2 Fermions

The scattering of identical fermions is more difficult to analyze than that of spinless bosons because of the complications due to the spin. For simplicity, we shall only consider the case of two identical spin-1/2 fermions interacting through central forces. Since the interaction is in general different in the singlet ($S = 0$) and triplet ($S = 1$) spin states of the two fermions, we shall start from two (unsymmetrized) scattering amplitudes $f_s(\theta)$ and $f_t(\theta)$ corresponding respectively to the singlet and triplet cases.

The full wave function describing a system of two identical spin-1/2 fermions must be antisymmetric in the interchange of the two particles, i.e., when all their coordinates (spatial and spin) are interchanged. Now, if the system is in the singlet spin state ($S = 0$), the spin part of the wave function is antisymmetric. Hence the corresponding spatial part of the wave function must be symmetric in the interchange of the position vectors \vec{r}_1 and \vec{r}_2 of the two particles. As a result, the symmetrized singlet scattering amplitude is

$$f_{s+} = f_s(\theta) + f_s(\pi - \theta) \quad (26.51)$$

and the differential cross-section in the singlet spin state is

$$\frac{d\sigma_s}{d\Omega} = |f_s(\theta) + f_s(\pi - \theta)|^2. \quad (26.52)$$

If, on the other hand, the two spin-1/2 fermions are in a triplet spin state ($S = 1$) the corresponding three spin functions are symmetric in the interchange of the spin coordinates of the two particles. The spatial part of the wave function must therefore be antisymmetric in the interchange of the position vectors \vec{r}_1 and \vec{r}_2 , so that the symmetrized triplet scattering amplitude is given by

$$f_{t-} = f_t(\theta) - f_t(\pi - \theta) \quad (26.53)$$

and the differential cross-section in the singlet triplet spin state is

$$\frac{d\sigma_t}{d\Omega} = |f_t(\theta) - f_t(\pi - \theta)|^2. \quad (26.54)$$

If the ‘incident’ and ‘target’ particles are unpolarized (i.e., their spins are randomly orientated), the probability of obtaining triplet states is three times that of singlet states, so that the differential cross-section is given by

$$\begin{aligned} \frac{d\sigma}{d\Omega} &= \frac{1}{4} \frac{d\sigma_s}{d\Omega} + \frac{3}{4} \frac{d\sigma_t}{d\Omega} \\ &= \frac{1}{4} |f_s(\theta) + f_s(\pi - \theta)|^2 + \frac{3}{4} |f_t(\theta) - f_t(\pi - \theta)|^2. \end{aligned} \quad (26.55)$$

For the particular case of *spin-independent* central interactions, where

$$f_s(\theta) = f_t(\theta) = f(\theta) \quad (26.56)$$

we find from (26.55) that

$$\frac{d\sigma}{d\Omega} = |f(\theta)|^2 + |f(\pi - \theta)|^2 - \text{Re}[f(\theta)f^*(\pi - \theta)]. \quad (26.57)$$

We note that this formula differs from the ‘classical’ result by the presence of the third term on the right, which again is an interference term. We also remark that, at $\theta = \pi/2$, the quantum mechanical differential cross-section (26.57) is given by

$$\frac{d\sigma(\theta = \pi/2)}{d\Omega} = |f(\theta = \pi/2)|^2 \quad (26.58)$$

and hence is equal to one-half of the classical result

$$\frac{d\sigma_{cl}(\theta = \pi/2)}{d\Omega} = 2 |f(\theta = \pi/2)|^2.$$

26.5.1 Partial Wave Expansion of the Scattering Amplitude

Consider spin-independent central potential so that the scattering amplitude is the same in the singlet and triplet states of the two particles. Letting the scattering amplitude to be $f(\theta)$ we have,

$$f(\theta) + f(\pi - \theta) = \frac{2}{k} \sum_{l=\text{even}} (2l+1) e^{i\delta_l} \sin \delta_l P_l(\cos \theta) \quad (26.59)$$

and

$$f(\theta) - f(\pi - \theta) = \frac{2}{k} \sum_{l=\text{odd}} (2l+1) e^{i\delta_l} \sin \delta_l P_l(\cos \theta) \quad (26.60)$$

At low energies, the s -wave ($l = 0$) dominates. Therefore

$$\frac{d\sigma}{d\Omega} = \frac{1}{4} \left| \frac{2}{k} e^{i\delta_0} \sin \delta_0 \right|^2 = \frac{1}{k^2} \sin^2 \delta_0, \quad (26.61)$$

and the total (integrated cross section is

$$\sigma_{el}(\text{low energy}) = \int \frac{d\sigma}{d\Omega} d\Omega = \frac{4\pi}{k^2} \sin^2 \delta_0. \quad (26.62)$$

Thus, at low energies, differential cross section and the total elastic cross section for two identical spin-half fermions (under a spin-independent central potential) is four times smaller than the corresponding quantities for two identical spinless bosons.

Note:

We note that

$$\begin{aligned} \left. \frac{d\sigma(\theta = \pi/2)}{d\Omega} \right|_{\text{spinless bosons}} &: \left. \frac{d\sigma(\theta = \pi/2)}{d\Omega} \right|_{\text{classical}} : \left. \frac{d\sigma(\theta = \pi/2)}{d\Omega} \right|_{\text{spin-1/2 fermions}} \\ &= 4 : 2 : 1 \end{aligned}$$

This result can be utilized to determine, experimentally, the spin of a particle like a proton. In this case, the classical cross section is given by Rutherford formula. The fact that the observed cross section for proton-proton scattering at $\theta = \pi/2$ is approximately half of the value given by Rutherford formula, indicates that the proton is a spin-1/2 particle.

Integrals

.1 List of Standard Integrals

$$\int_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\pi} \quad (63)$$

$$\int_{-\infty}^{\infty} e^{-\alpha x^2 + \beta x} dx = \left(\frac{\pi}{2}\right)^{1/2} \exp\left[\frac{\beta^2}{4\alpha}\right] \quad (64)$$

$$\int_{-\infty}^{\infty} e^{-ax^2} dx = \sqrt{\frac{\pi}{a}} \quad (65)$$

$$\int_{-\infty}^{\infty} e^{-ax^2} x^2 dx = \frac{1}{2a} \sqrt{\frac{\pi}{a}} \quad (66)$$

.2 Proofs

$$\int_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\pi} \quad (67)$$

proof:

let,

$$I = \int_{-\infty}^{\infty} e^{-x^2} dx \quad (68)$$

then

$$I^2 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-(x^2+y^2)} dx dy \quad (69)$$

using $r^2 = x^2 + y^2$ since

$$x = r \cos \theta \quad (70)$$

$$y = r \sin \theta \quad (71)$$

we get

$$I^2 = \int_0^{2\pi} \int_0^{\infty} e^{-r^2} r dr d\theta \quad (72)$$

$$= 2\pi \int_0^{\infty} e^{-r^2} r dr \quad (73)$$

let $u = r^2$ therefore $rdr = du/2$

$$I^2 = \pi \int_0^\infty e^{-u} du \quad (74)$$

$$= \pi \left[\frac{e^{-u}}{-1} \right]_0^\infty \quad (75)$$

$$= \pi [0 - (-1)] \quad (76)$$

$$= \pi \quad (77)$$

therefore $I = \sqrt{\pi}$

another

$$I = \int_{-\infty}^{\infty} dx e^{-(b^2 x^2 + ax)} \quad (78)$$

Notice that $(b^2 x^2 + ax) = (bx + \frac{a}{2b})^2 - \frac{a^2}{4b^2}$

$$I = e^{a^2/4b^2} \int_{-\infty}^{\infty} dx e^{-(bx + \frac{a}{2b})^2} \quad (79)$$

let $z = bx + \frac{a}{2b}$ thus $dz = b dx$

$$I = e^{a^2/4b^2} \frac{1}{b} \int_{-\infty}^{\infty} dz e^{-z^2} \quad (80)$$

$$= e^{a^2/4b^2} \frac{\sqrt{\pi}}{b} \quad (81)$$