

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	9
1.1	Other representation of delta function	10
1.2	Properties of the delta function	12
1.2.1	Notes	14
1.3	Derivatives of the delta function	15
1.4	Integration of the delta function	16
1.5	Three - dimensional delta function	17
1.6	Fourier Transformation	18
2	sheet-2 : Linear Vector Space	21
2.1	Definition	21
2.2	Example of linear vector space	22
2.3	Inner product space or a unitary vector space	24
2.4	Examples of scalar product	24
2.5	Norm of a vector	26
2.6	Metric included by the scalar product	26
2.7	Schwarz's inequality	27
2.8	Analogy of Schwarz inequality with vectors in a three-dimensional Euclidean space \mathbb{R}^3	27

2.9	Orthogonality and linear independence	28
2.9.1	Linear independence	28
2.10	Orthonormality and linear independence	28
2.11	Schmidt orthonormalization method	28
2.12	Dimension of a vector space	29
2.13	Complete vector space	29
2.14	Convergence of a sequence of vectors in a vector space	30
2.14.1	An example of a vector space whose Cauchy sequence does not converge to a vector in the vector space	30
2.15	Hilbert space	31
2.16	Basis vectors in a Hilbert space	31
2.16.1	Finite dimensional space	31
2.16.2	Infinite dimensional vector space	32
2.16.3	Example	33
2.17	Dirac notation	33
2.17.1	Scalar product	33
2.17.2	Dual vector space	33
2.17.3	Bra notation for the vectors of V^*	34
2.17.4	Correspondence between kets and bras	34
2.17.5	The correspondence in anti-linear	34
2.17.6	Dirac notation for the scalar product	34
3	sheet-3 : Linear Vector Space (continued)	37
3.1	Operators in a Hilbert Space	37
3.2	Linear operators	37
3.3	Commutator of two operators	38
3.3.1	Some properties of commutators	38
3.4	Projection operator	38
3.5	Representation of vectors and operators	39
3.6	Matrix representation of ket and bra vectors	41
3.7	Representation of an operator in a basis	42
3.8	Matrix representation of the sum and product of two operators	43
4	sheet-4 : Linear Vector Space (continued)	45
4.1	Adjoint operator	45
4.2	Hermitian or Self-adjoint operator	46
4.3	Inverse operator	47
4.4	Unitary operator	47

4.5	Function of Operators	47
5	sheet-5 : Eigenvalue and Eigenvectors of Operators	49
5.1	Eigenvalues and Eigenvectors of a Hermitian Operator	49
5.2	Determination of eigenvalues and eigenvectors of a Hermitian Operator	50
5.3	Non-degenerate roots	51
5.4	Degenerate roots	52
5.5	Digonalization of a Hermitian Operator	53
5.6	Basis independence of the eigenvalues of an Operator	55
5.7	Infinite dimensional vector space	57
5.8	Completeness condition for the eigenvectors of a Hermitian Operator	57
5.9	Hermitian operator with continuous eigenvalue spectrum	59
5.9.1	Continuous Spectrum	59
5.10	Hermitian operator with continuous and discrete eigenvalue	60
5.11	Problems	61
6	sheet-6 : Coordinate and Momentum Representation	67
7	sheet-7 : Postulates of Quantum Mechanics	69
8	sheet-8 : Compatibility of Observable	71
9	sheet-9 : The Uncertainty Principle	73
10	sheet-10 : Harmonic Oscillator	75
11	sheet-11 : Quantum Dynamics	77
11.1	Quantum Dynamics	77
11.1.1	Schrodinger picture	77
11.1.2	Schrodinger equation	77
12	sheet-12 : The Path Integral Formulation of Quantum Theory	79
12.1	Background Materials	79
12.2	Propagator	82
12.3	Path Integral for the Propagator	82
12.4	Path Integral for a Free Particle	85
12.4.1	check of calculation	89

12.5	Derivation of the propagator for a free particle without using the path integral formula	89
12.6	The Classical Action	90
12.7	Discussion on Path Integral	91
12.7.1	Principle of Least Action	91
12.8	Discussion on The Phase of the Path Integral	92
12.9	Equivalence to the Schrodinger Equation	94
12.10	Potentials of the Form $V = a + bx + cx^2 + d\dot{x} + e\ddot{x}$	94
12.10.1	Special Cases	94
13	sheet-13 : Potential Wells and Barriers	95
14	sheet-14 : Particle Motion in 3-Dimension	97
15	sheet-15 : The Hydrogen Atom	99
16	sheet-16 : Angular Momentum 1	101
17	sheet-17 : Angular Momentum 2: spin 1/2 system	103
18	sheet-18 : Rotations and Angular Momentum	105
19	sheet-19 : Addition of Angular Momentum	107
20	sheet-20 : Time Independent Perturbation	109
21	sheet-21 : Variational Method	111
22	sheet-22 : WKB approximation	113
23	sheet-23 : Time Dependent Perturbation	115
24	sheet-24 : Scattering Theory	117
25	sheet-25 : Integral Equation for Scattering Theory	119
26	sheet-26 : Identical Particles	121
.1	List of Standard Integrals	122
.2	Proofs	122

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

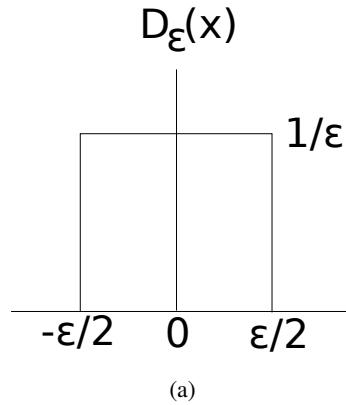


Figure 1.1: Dirac Delta Function

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,

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

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 :

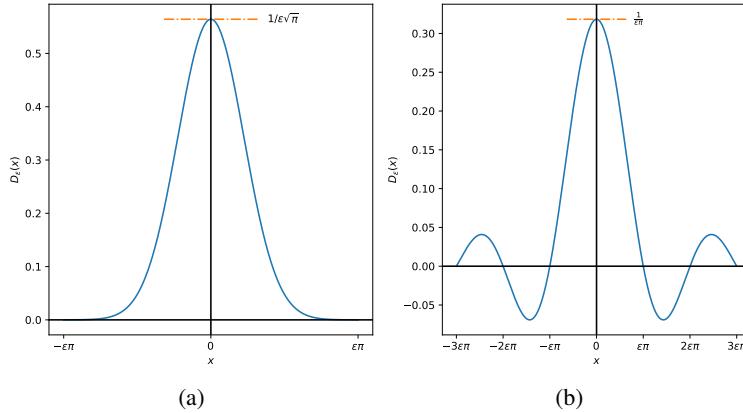
$$\lim_{\varepsilon \rightarrow 0} D_\varepsilon(x) \equiv \delta(x) \quad (1.8)$$

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.

1.1 Other representation of delta function

1. Consider the function

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

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

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 $\epsilon \rightarrow 0$ 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,

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

3. We can also show that

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

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

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

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 an 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,

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

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 :

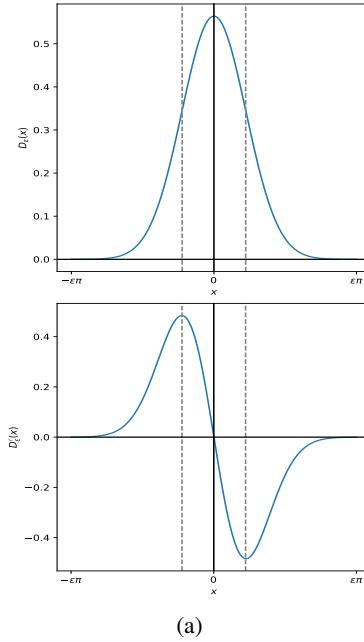
$$\frac{1}{x \pm i\epsilon} = \frac{x \mp i\epsilon}{x^2 + \epsilon^2} = \frac{x}{x^2 + \epsilon^2} \mp \frac{i\epsilon}{x^2 + \epsilon^2} \quad (1.42)$$

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

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



(a)

Figure 1.3: (a) Derivative of Delta Function

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

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

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

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

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 $D_\varepsilon(x)$ has two peaks close to the origin, one peak is positive and the other is 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 the integration by parts gives

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

Because $D_\varepsilon(x)$ 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 $\varepsilon \rightarrow 0$, we arrive at the definition of $\delta'(x)$

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

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

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

A graph of $\Theta_\varepsilon(x)$ 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 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

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

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

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

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

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

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

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

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

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

A useful formula Consider the integral

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

using

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

Thus

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

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

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

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

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

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

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

Thus in summary

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

In three dimensions

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

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

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

$$(1.76)$$

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

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

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

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

Therefore

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

Thus we can write

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

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

Parseval's Identity:

We can now prove the important identity

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

proof:

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

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

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

here, \mathbb{R} = 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

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

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

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

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

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

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

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

$$\psi_a = (a_1 \ a_2 \ \dots \ a_n)^T \quad (2.39)$$

The scalar product may then be defined as

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

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

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

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

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

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

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 $\phi_1, \psi_2, \dots, \phi_n$ as the basis. Being orthogonal, the vectors $\phi_1, \psi_2, \dots, \phi_n$ 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 o 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 cector 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 $\langle \cdot |$. 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

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

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

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

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, $|P_a\rangle$ 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:

$$\langle a| \rightarrow (\langle a|1\rangle \quad \langle a|2\rangle \quad \dots) = (a_1^* \quad a_2^* \quad \dots) \quad (3.53)$$

Then the scalar product becomes a number. Thus

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

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

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

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

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

In general

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

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

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

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

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

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

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

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

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

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

$$= \langle i | \hat{A} \hat{\mathbb{I}} \hat{B} | j \rangle \quad (3.82)$$

$$= \sum_k \langle i | \hat{A} | k \rangle \langle k | \hat{B} | j \rangle \quad (3.83)$$

$$= \sum_k A_{ik} B_{kj} \quad (3.84)$$

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 $\langle b|$.

$$[b]^\dagger = (b_1^* \ b_2^* \ \dots) \quad (3.90)$$

$[A]$ 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

$$\langle b|\hat{A}|a\rangle = (b_1^* \ b_2^* \ \dots) \begin{pmatrix} A_{11} & A_{12} & \dots \\ A_{21} & A_{22} & \dots \\ \vdots & \vdots & \vdots \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ \vdots \end{pmatrix} \quad (3.93)$$

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

$$|a\rangle \xrightarrow{dc} \langle a| \quad (4.5)$$

$$|b\rangle \xrightarrow{dc} \langle b| \quad (4.6)$$

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

$$\langle i|\hat{K}|j\rangle = \langle j|\hat{K}^\dagger|i\rangle^* \quad (4.14)$$

$$K_{ij} = (\hat{K}^\dagger)_{ji}^* \quad (4.15)$$

$$K_{ji}^\dagger = K_{ij}^* \quad (4.16)$$

$$K_{ij}^\dagger = (K_{ji})^* \quad (4.17)$$

$$[\hat{K}^\dagger] = [K]^\dagger \quad (4.18)$$

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:

$$(\psi_a, \hat{A}\hat{B}\psi_b) = ((\hat{A}\hat{B})^\dagger \psi_a, \psi_b) \quad (4.22)$$

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

$$\hat{B} = \hat{A}^{-1} \quad (4.26)$$

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

$$\hat{U}\hat{U}^\dagger = \hat{U}^\dagger\hat{U} = \hat{\mathbb{I}} \quad (4.28)$$

i.e., if

$$\hat{U}^\dagger = \hat{U}^{-1} \quad (4.29)$$

Thus, for a unitary operator, its adjoint is @@@@ @@@@ @@@@ @@@@ @@@@ @@@@

4.5 Function of Operators

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

$$\mathbb{A}' = \begin{bmatrix} \lambda_1 & 0 & 0 & \dots & 0 \\ 0 & \lambda_2 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \lambda_n \end{bmatrix} \quad (5.23)$$

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.

$$\left[\begin{array}{ccc|c|c} a_{11} & a_{12} & a_{13} & \mathbf{0} & \mathbf{0} \\ a_{21} & a_{22} & a_{23} & & \\ a_{31} & a_{32} & a_{33} & & \\ \hline \mathbf{0} & & & b_{11} & b_{12} \\ & & & b_{21} & b_{22} \\ \hline \mathbf{0} & & & \mathbf{0} & \\ & & & c_{11} & c_{12} \\ & & & c_{21} & c_{22} \end{array} \right] \quad (5.29)$$

Writing with basis

$$\begin{array}{ccccccccc} & 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ & |\lambda_1, 1\rangle & |\lambda_1, 2\rangle & |\lambda_1, 3\rangle & |\lambda_2, 1\rangle & |\lambda_2, 2\rangle & |\lambda_3, 1\rangle & |\lambda_3, 2\rangle \\ \langle \lambda_1, 1 | & a_{11} & a_{12} & a_{13} & 0 & 0 & 0 & 0 \\ \langle \lambda_1, 2 | & a_{21} & a_{22} & a_{23} & 0 & 0 & 0 & 0 \\ \langle \lambda_1, 3 | & a_{31} & a_{32} & a_{33} & 0 & 0 & 0 & 0 \\ \langle \lambda_2, 1 | & 0 & 0 & 0 & b_{11} & b_{12} & 0 & 0 \\ \langle \lambda_2, 2 | & 0 & 0 & 0 & b_{21} & b_{22} & 0 & 0 \\ \langle \lambda_3, 1 | & 0 & 0 & 0 & 0 & 0 & c_{21} & c_{22} \\ \langle \lambda_3, 2 | & 0 & 0 & 0 & 0 & 0 & c_{21} & c_{22} \end{array} \quad (5.30)$$

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

$$\begin{array}{c|ccccccc} & |\lambda_1, 1\rangle & |\lambda_1, 2\rangle & |\lambda_1, 3\rangle & |\lambda_2, 1\rangle & |\lambda_2, 2\rangle & |\lambda_3, 1\rangle & |\lambda_3, 2\rangle \\ \langle \lambda_1, 1 | & \lambda_1 & 0 & 0 & 0 & 0 & 0 & 0 \\ \langle \lambda_1, 2 | & 0 & \lambda_1 & 0 & 0 & 0 & 0 & 0 \\ \langle \lambda_1, 3 | & 0 & 0 & \lambda_1 & 0 & 0 & 0 & 0 \\ \langle \lambda_2, 1 | & 0 & 0 & 0 & \lambda_2 & 0 & 0 & 0 \\ \langle \lambda_2, 2 | & 0 & 0 & 0 & 0 & \lambda_2 & 0 & 0 \\ \langle \lambda_3, 1 | & 0 & 0 & 0 & 0 & 0 & \lambda_3 & 0 \\ \langle \lambda_3, 2 | & 0 & 0 & 0 & 0 & 0 & 0 & \lambda_3 \end{array} \quad (5.31)$$

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

similarily transformation with a *unitary* matrix. This is easy to see

$$\begin{aligned}
 A'_{ij} &\equiv \langle u_i | A | u_j' \rangle \\
 &= \sum_{jk} \langle u_i' | u_j \rangle \langle u_j | A | u_k \rangle \langle u_k | u_j' \rangle \\
 &= \sum_{jk} S_{ij} A_{jk} S_{jk}^* \\
 &= \sum_{jk} S_{ij} A_{jk} S_{jk}^\dagger
 \end{aligned} \tag{5.33}$$

Where we have defined the matrix S as

$$S_{ij} \equiv \langle u_i' | u_j \rangle \tag{5.34}$$

The matrix S is unitary as shown precisely. In matrix notation, we write equation (5.33) as,

$$A' = SAS^\dagger = SAS^{-1} \tag{5.35}$$

since S is unitary matrix. Then

$$\begin{aligned}
 \det(A' - \lambda \mathbb{I}) &= \det(SAS^{-1} - \lambda S\mathbb{I}S^{-1}) \\
 &= \det(S(A - \lambda \mathbb{I})S^{-1}) \\
 &= \det(A - \lambda \mathbb{I})
 \end{aligned} \tag{5.36}$$

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 \mathbb{A} and \mathbb{B} are square matrices. Now, if we use the eigenbasis of the hermitian operator A for the representation, then \mathbb{A} is a diagonal matrix with

$$\det \mathbb{A} = \prod_i \lambda_i \quad (5.41)$$

$$\text{Tr } \mathbb{A} = \sum_i \lambda_i \quad (5.42)$$

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 \mathbb{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 these 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

$$\hat{1} = \sum_i |a_i\rangle \langle a_i| = \sum_i \hat{P}_i \quad (5.45)$$

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

$$\begin{aligned} \hat{O} &= \hat{1} \hat{O} \hat{1} = \sum_{i,j} |a_i\rangle \langle a_i| \hat{O} |a_j\rangle \langle a_j| \\ &= \sum_{i,j} |a_i\rangle O_{ij} \langle a_j| \end{aligned} \quad (5.47)$$

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

$$\sum_{i=1}^{\infty} \sum_{s^i=1}^{g_i} |a_i, s^{(i)}\rangle \langle a_i, s^{(i)}| = \hat{1} \quad (5.49)$$

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)

$$\hat{1} = \sum_i \hat{P}_i \quad (5.51)$$

where

$$\hat{P}_i = \sum_{s^{(i)}=1}^{g_i} |a_i, s^{(i)}\rangle \langle a_i, s^{(i)}| \quad (5.52)$$

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

$$\det(M - \lambda \mathbb{1}) = 0$$

$$\begin{vmatrix} 1 - \lambda & 1 \\ 0 & 1 - \lambda \end{vmatrix} = 0$$

$$(1 - \lambda)^2 = 0$$

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

$$\begin{aligned} \begin{bmatrix} 3 & i \\ -i & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} &= \lambda \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \\ \begin{bmatrix} 3-\lambda & i \\ -i & 3-\lambda \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} &= \lambda \begin{bmatrix} 0 \\ 0 \end{bmatrix} \end{aligned} \quad (5.76)$$

The secular equation is

$$\begin{aligned} \begin{vmatrix} 3-\lambda & i \\ -i & 3-\lambda \end{vmatrix} &= 0 \\ (3-\lambda)^2 - (i)(-i) &= 0 \\ (\lambda-3)^2 - 1 &= 0 \\ (\lambda-3-1)(\lambda-3+1) &= 0 \\ \lambda &= 2, 4 \end{aligned} \quad (5.77)$$

None of the roots are degenerate.

Eigenvector for $\lambda = 2$

$$\begin{aligned} \begin{bmatrix} 3-2 & i \\ -i & 3-2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} &= \emptyset \\ \begin{bmatrix} x_1 + ix_2 \\ -ix_1 + x_2 \end{bmatrix} &= \emptyset \end{aligned}$$

Thus

$$x_1 + ix_2 = 0 \quad (5.78)$$

$$-ix_1 + x_2 = 0 \quad (5.79)$$

We get the solution $x_1 = -ix_2$. Taking x_2 to be arbitrary

$$|2\rangle = \begin{bmatrix} -ix_2 \\ x_2 \end{bmatrix} \quad (5.80)$$

Normalizing

$$\begin{aligned} \langle 2 | 2 \rangle &= 1 \\ [ix_2^* & x_2^*] \begin{bmatrix} -ix_2 \\ x_2 \end{bmatrix} = 1 \\ 2|x_2|^2 &= 1 \\ |x_2| &= \frac{1}{\sqrt{2}} \end{aligned} \quad (5.81)$$

Take $x_2 \doteq \frac{1}{\sqrt{2}}$.

We could have taken

$$x_2 = -\frac{1}{\sqrt{2}} \quad (5.82)$$

or

$$x_2 = e^{i\phi} \frac{1}{\sqrt{2}} \quad (5.83)$$

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

$$\langle 4 | 4 \rangle = 1$$

$$\begin{bmatrix} -ix_2^* & x_2^* \end{bmatrix} \begin{bmatrix} ix_2 \\ x_2 \end{bmatrix} = 1$$

$$2|x_2|^2 = 1$$

$$|x_2| = \frac{1}{\sqrt{2}} \quad (5.88)$$

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

$$|2\rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} -i \\ 1 \end{bmatrix} \quad (5.90)$$

$$|4\rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} i \\ 1 \end{bmatrix} \quad (5.91)$$

$$\langle 2 | 4 \rangle = \frac{1}{2} [i \quad 1] \begin{bmatrix} i \\ 1 \end{bmatrix} = \frac{1}{2} (i^2 + 1) = 0 \quad (5.92)$$

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,

$$A' = \begin{bmatrix} \langle 2 | & \langle 4 | \end{bmatrix} \begin{bmatrix} i/\sqrt{2} & 1/\sqrt{2} \\ -i/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} A \begin{bmatrix} |2\rangle & |4\rangle \\ -i/\sqrt{2} & i/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} \quad (5.95)$$

$$= \begin{bmatrix} \langle 2 | & \langle 4 | \end{bmatrix} \begin{bmatrix} i/\sqrt{2} & 1/\sqrt{2} \\ -i/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} \begin{bmatrix} 3 & i \\ -i & 3 \end{bmatrix} \begin{bmatrix} |2\rangle & |4\rangle \\ -i/\sqrt{2} & i/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} \quad (5.96)$$

$$= S^{-1}AS \quad (5.97)$$

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

$$\begin{aligned} & \begin{vmatrix} \frac{3}{2} - \lambda & -\frac{1}{2} & 0 \\ -\frac{1}{2} & \frac{3}{2} - \lambda & 0 \\ 0 & 0 & 1 - \lambda \end{vmatrix} = 0 \\ & (1 - \lambda) \left[\left(\frac{3}{2} - \lambda \right)^2 - \left(-\frac{1}{2} \right) \left(-\frac{1}{2} \right) \right] = 0 \\ & (\lambda - 1) [(\lambda - 3/2 + 1/2)(\lambda - 3/2 - 1/2)] = 0 \\ & (\lambda - 1)(\lambda - 1)(\lambda - 2) = 0 \end{aligned}$$

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

$$\mathbb{A} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 1 \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

$$\begin{bmatrix} \frac{3}{2} - 1 & -\frac{1}{2} & 0 \\ -\frac{1}{2} & \frac{3}{2} - 1 & 0 \\ 0 & 0 & 1 - 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} \frac{1}{2} & -\frac{1}{2} & 0 \\ -\frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} 1 & -1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

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

$$S^\dagger = \begin{bmatrix} |1,1\rangle & |1,2\rangle & |2\rangle \\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} \\ 0 & 1 & 0 \end{bmatrix} \quad (5.106)$$

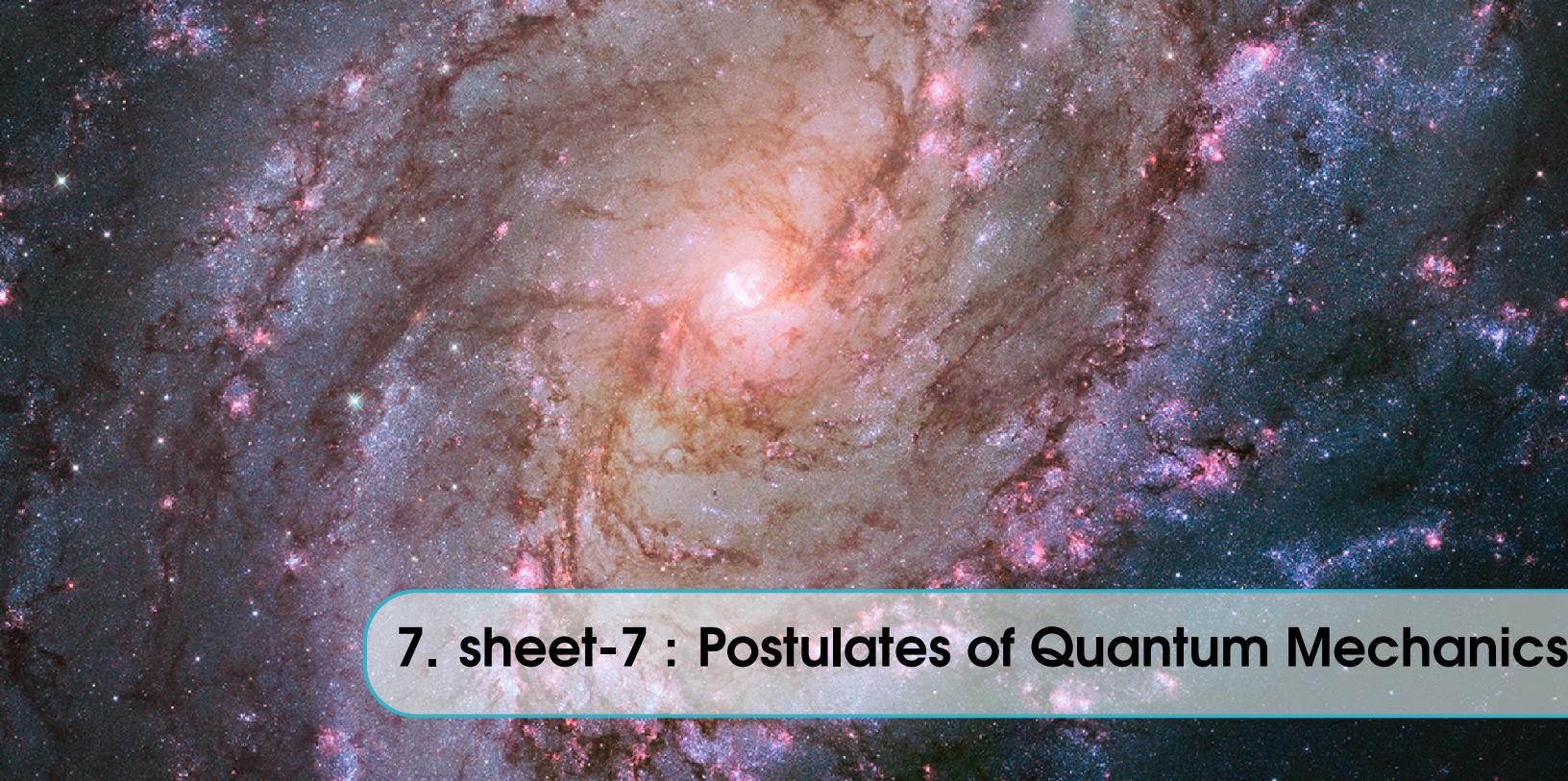
$$S = (S^\dagger)^\dagger = \begin{bmatrix} \langle 1,1 | & \langle 1,2 | & \langle 2 | \\ \langle 1,1 | & \langle 1,2 | & \langle 2 | \\ \langle 1,2 | & \langle 2 | & \langle 2 | \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \end{bmatrix} \quad (5.107)$$

Then the matrix M' is diagonal

$$M' = S M S^\dagger = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix} \quad (5.108)$$



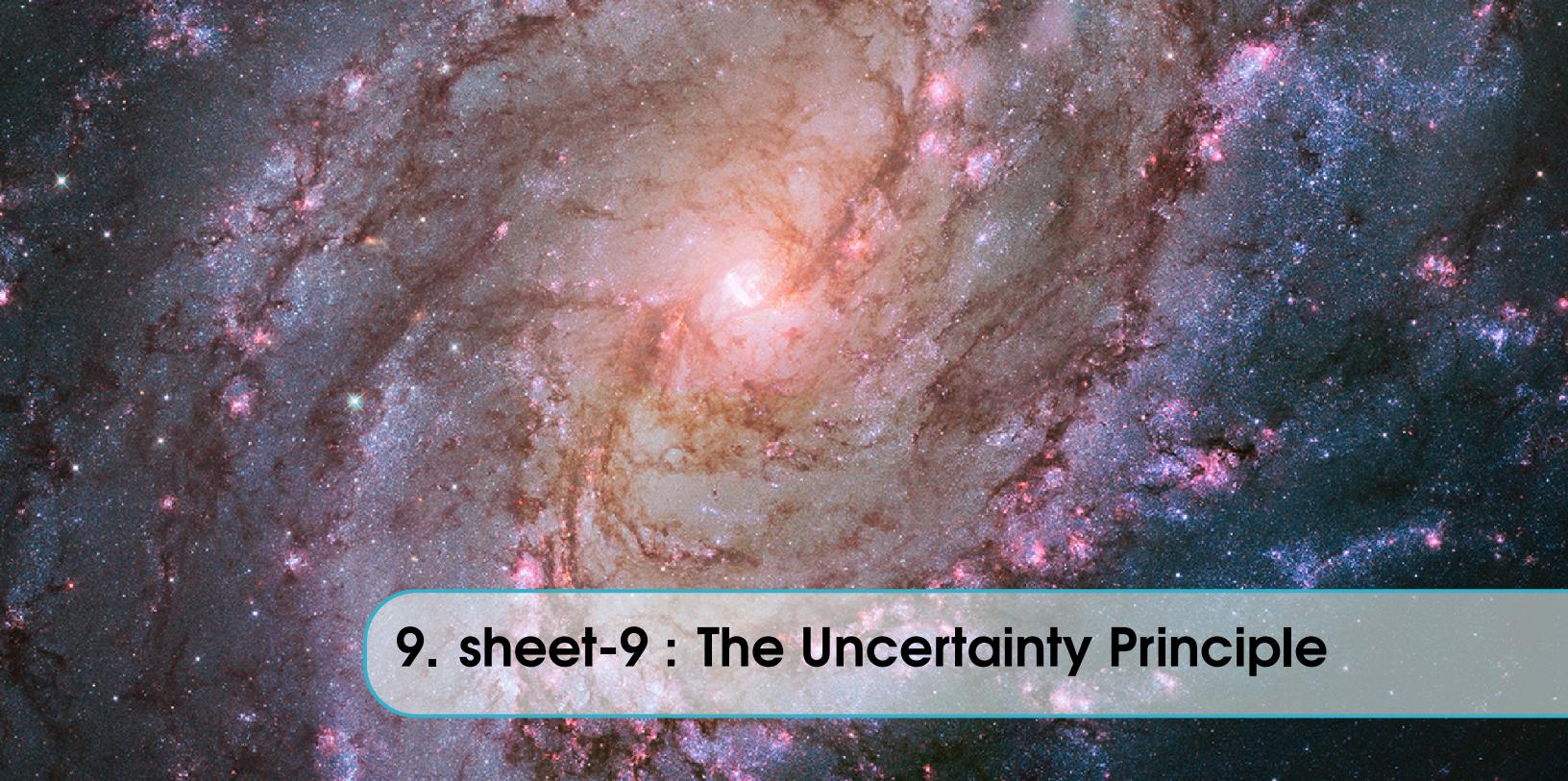
6. sheet-6 : Coordinate and Momentum Repre



7. sheet-7 : Postulates of Quantum Mechanics



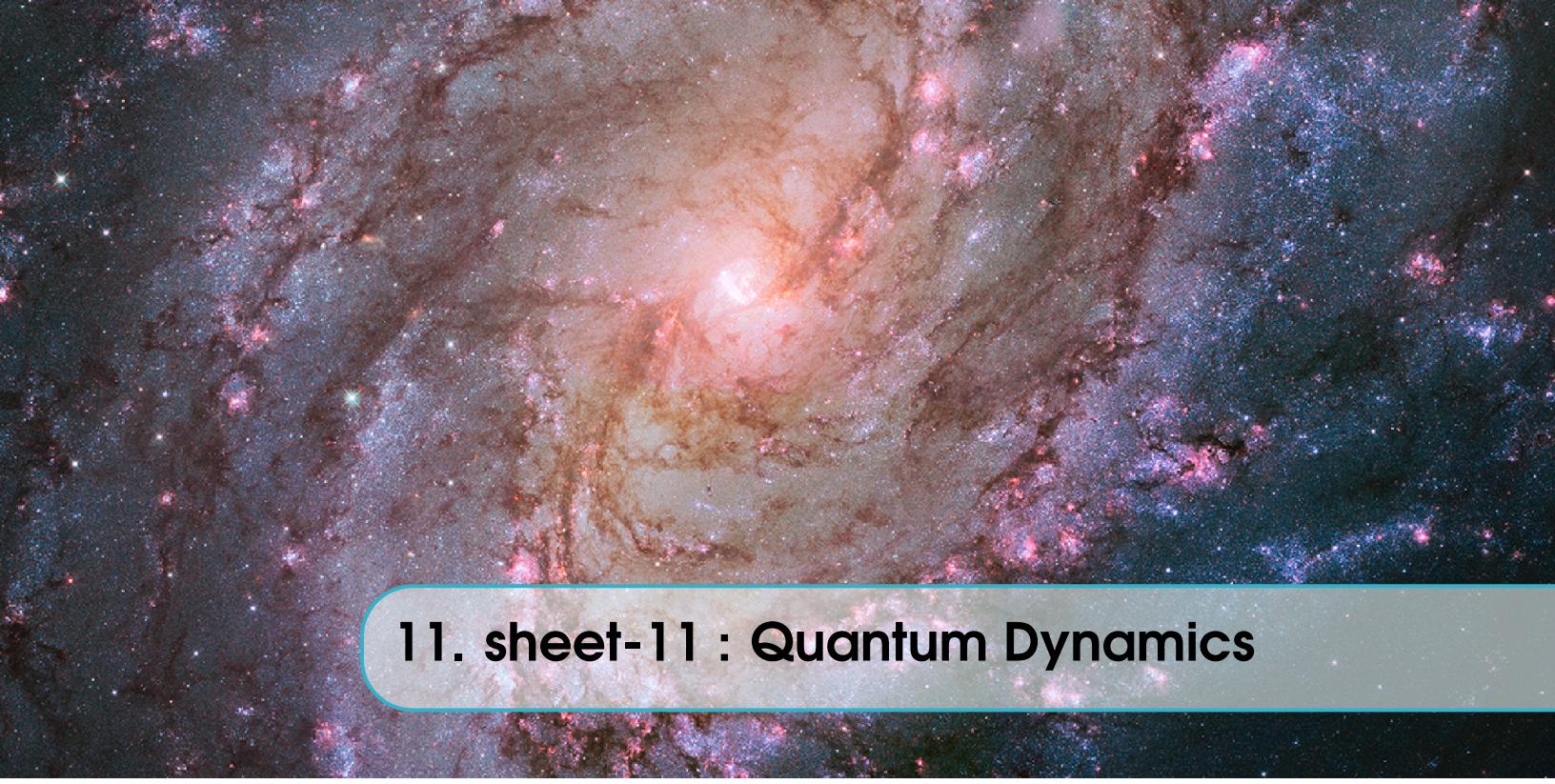
8. sheet-8 : Compatibility of Observable



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 , $|\psi(t)\rangle$, 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 $|\psi(t)\rangle$ 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 representation

$$\langle p|\hat{Q}_s = i\hbar \frac{\partial}{\partial p} \langle p| \quad (12.10)$$

$$\langle p|\hat{P}_s = \langle p| \quad (12.11)$$

The fundamental commutation relation between \hat{Q} and \hat{P} is

$$[\hat{Q}_s, \hat{P}_s] = i\hbar\hat{\mathbb{1}} \quad (12.12)$$

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

$$\langle q|p\rangle = Ce^{ipq/\hbar} \quad (12.16)$$

The constant C is chosen such that we have the normalization $\langle p|p'\rangle \delta(p - p')$. Now

$$\begin{aligned} \langle p|p'\rangle &= \int dq \langle p|q\rangle \langle q|p'\rangle \\ &= \int dq C^* e^{-ipq/\hbar} C e^{ip'q/\hbar} \\ &= |C|^2 \int_{-\infty}^{\infty} dq e^{-i(p-p')q/\hbar} \\ &= |C|^2 2\pi\delta(\frac{p-p'}{\hbar}) \\ &= |C|^2 2\pi\hbar\delta(p-p') \end{aligned}$$

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

$$|\rangle_H = e^{i\hat{H}t/\hbar} |\rangle_s \quad (12.19)$$

$$\hat{\Omega}_H(t) = e^{i\hat{H}t/\hbar} \hat{\Omega}_s e^{-i\hat{H}t/\hbar} \quad (12.20)$$

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

$$\int dq |q, t\rangle_H {}_H \langle q, t | = \int dq - e^{i\hat{H}t/\hbar} |q\rangle \langle q| e^{-i\hat{H}t/\hbar} \quad (12.32)$$

$$= e^{i\hat{H}t/\hbar} \left(\int dq |q\rangle \langle q| \right) e^{-i\hat{H}t/\hbar} \quad (12.33)$$

$$= e^{i\hat{H}t/\hbar} \mathbb{1} e^{-i\hat{H}t/\hbar} \quad (12.34)$$

$$= \mathbb{1} \quad (12.35)$$

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

$$U(q_2, t_2; q_1, t_1) = \langle q_2 | H | q_1, t_1 \rangle_H \quad (12.39)$$

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:

$$\psi_S(q, t) = \langle q | \psi_S(t) \rangle \quad (12.41)$$

$$= \langle q | e^{-i\hat{H}t/\hbar} | \psi_S(0) \rangle \quad (12.42)$$

$$= \langle q | H | \psi \rangle_H \quad (12.43)$$

$$= \int dq' \langle q, t | q', t' \rangle_H \langle q', t' | \psi \rangle_H \quad (12.44)$$

$$= \int dq' U(q, t; q', t') \psi_S(q', t') \quad (12.45)$$

The path integral formalism of quantum dynamics provides a means to construct the transition amplitude $\langle q, t | q', t' \rangle_H$ 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

$$U(x, t; x_0, t_0) = \langle x, t | x_0, t_0 \rangle \quad (12.46)$$

where $t > t_0$. For this purpose let us devide the time integral (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

$$t_i = t_0 + i\epsilon, \quad i = 1, 2, \dots, N-1 \quad (12.48)$$

Figure 12.1: text

At each intermediate time a complete set of basis states $|x_i, t_i\rangle$ may be inserted:

$$\langle xt|x_0t_0\rangle = \int_{-\infty}^{\infty} dx_1 \dots \int_{-\infty}^{\infty} dx_{N-1} \langle xt|x_{N-1}t_{N-1}\rangle \langle x_{N-1}t_{N-1}|x_{N-2}t_{N-2}\rangle \dots \langle x_2t_2|x_1t_1\rangle \langle x_1t_1|x_0t_0\rangle \quad (12.49)$$

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 equation (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

$$\langle xt|x_0t_0\rangle = \int \prod_{i=1}^{N-1} dq_i \prod_{i=0}^{N-1} \langle x_{i+1}t_{i+1}|x_it_i\rangle \quad (12.50)$$

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 $\langle x_{i+1}t_{i+1}|x_it_i\rangle$. We can calculate this inner product up to first order in ϵ from equation (12.47) as follows

$$\langle x_{i+1}t_{i+1}|x_it_i\rangle = \langle x_{i+1}| e^{-i\hat{H}t_{i+1}/\hbar} e^{i\hat{H}t_i/\hbar} |x_i\rangle \quad (12.51)$$

$$= \langle x_{i+1}| e^{-i\hat{H}(t_{i+1}-t_i)/\hbar} |x_i\rangle \quad (12.52)$$

$$= \langle x_{i+1}| e^{-i\hat{H}\epsilon/\hbar} |x_i\rangle \quad (12.53)$$

$$= \langle x_{i+1}| \left(\mathbb{1} - \frac{i\epsilon}{\hbar} \hat{H} + \mathcal{O}(\epsilon^2) \right) |x_i\rangle \quad (12.54)$$

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

$$\langle xt | x_0 t_0 \rangle = \int \mathcal{D}[x(t')] e^{\frac{i}{\hbar} S[x(t')]} \quad (12.80)$$

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

$$\langle x | x_0 t_0 \rangle = \lim_{N \rightarrow \infty} \left(\frac{m}{2\pi\hbar i\varepsilon} \right)^{N/2} \int \prod_{i=1}^{N-1} dx_i e^{\frac{i\varepsilon}{\hbar} \sum_{i=0}^{N-1} \frac{m}{2} \left(\frac{x_{i+1} - x_i}{\varepsilon} \right)^2} \quad (12.86)$$

$$= \lim_{N \rightarrow \infty} \left(\frac{m}{2\pi\hbar i\varepsilon} \right)^{N/2} \int \prod_{i=1}^{N-1} dx_i e^{\frac{im}{2\hbar\varepsilon} \sum_{i=0}^{N-1} (x_{i+1} - x_i)^2} \quad (12.87)$$

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

$$\langle x | x_0 t_0 \rangle = \lim_{N \rightarrow \infty} \left(\frac{m}{2\pi\hbar i\varepsilon} \right)^{N/2} \left(\frac{2\hbar\varepsilon}{m} \right)^{(N-1)/2} \int \prod_{i=1}^{N-1} dy_i e^{-\sum_{i=0}^{N-1} \frac{(y_{i+1} - y_i)^2}{i}} \quad (12.89)$$

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 (2) from appendix (26) and choose

$$\alpha = \frac{3}{2i} \quad (12.96)$$

$$\beta = \frac{(y_0 + 2y_3)}{i} \quad (12.97)$$

$$\begin{aligned} I_2 &= \left(\frac{i\pi}{2}\right)^{1/2} \exp\left[-\frac{1}{2i}(2y_3^2 + y_0^2)\right] \left(\frac{2\pi i}{3}\right)^{1/2} \exp\left[-\frac{(y_0 + 2y_3)^2}{6/i}\right] \\ &= \left(\frac{i\pi}{2}\right)^{1/2} \left(\frac{2\pi i}{3}\right)^{1/2} \exp\left[-\frac{1}{2i}(2y_3^2 + y_0^2)\right] \exp\left[\frac{(y_0 + 2y_3)^2}{6i}\right] \\ &= \left(\frac{i^2\pi^2}{3}\right)^{1/2} \exp\left[-\frac{1}{i}\left\{\frac{1}{3}y_3^2 + \frac{1}{3}y_0^2 - \frac{2}{3}y_0y_3\right\}\right] \\ I_2 &= \left(\frac{i^2\pi^2}{3}\right)^{1/2} \exp\left[-\frac{(y_3 - y_0)^2}{3i}\right] \end{aligned} \quad (12.98)$$

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))

$$\langle xt|x_0t_0\rangle_{free} = U_{free}(x, t; x_0, t_0) = \lim_{N \rightarrow \infty} \left(\frac{m}{2\pi\hbar i\varepsilon}\right)^{N/2} \left(\frac{2\hbar\varepsilon}{m}\right)^{(N-1)/2} \frac{(i\pi)^{(N-1)/2}}{N^{1/2}} \exp\left[-(y_N - y_0)^2 / Ni\right] \quad (12.100)$$

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

$$\langle xt | x_0 t_0 \rangle \equiv U(x, t; x_0, t_0) \quad (12.107)$$

$$= \langle x | e^{-i\hat{H}(t-t_0)/\hbar} | x_0 \rangle \quad (12.108)$$

$$= \int dp \langle x | e^{-i\hat{H}(t-t_0)/\hbar} | p \rangle \langle p | x_0 \rangle \quad (12.109)$$

$$= \int dp \langle x | e^{-i\frac{p^2(t-t_0)}{2m\hbar}} | p \rangle \langle p | x_0 \rangle \quad (12.110)$$

$$= \int dp e^{-i\frac{p^2(t-t_0)}{2m\hbar}} \langle x | p \rangle \langle p | x_0 \rangle \quad (12.111)$$

$$= \int dp e^{-ip^2\Delta t/2m\hbar} \frac{1}{\sqrt{2\pi\hbar}} e^{ipx/\hbar} \frac{1}{\sqrt{2\pi\hbar}} e^{ipx_0/\hbar} \quad (12.112)$$

$$= \frac{1}{2\pi\hbar} \int_{-\infty}^{\infty} dp \exp \left[-ip^2\Delta t/2m\hbar + ip(x - x_0)/\hbar \right] \quad (12.113)$$

Using standard integral from appendix (equation 2) and

$$\alpha = \frac{i\Delta t}{2m\hbar} \quad (12.114)$$

$$\beta = \frac{i(x - x_0)}{\hbar} \quad (12.115)$$

Figure 12.3: Classical action

 \therefore

$$\begin{aligned} U(x, t' | x_0, t_0) &= \left(\frac{1}{2\pi\hbar} \right) \left(\frac{\pi}{i\Delta t / 2m\hbar} \right)^{1/2} e^{-\frac{(x-x_0)^2/\hbar^2}{4(i\Delta t / 2m\hbar)}} \\ &= \frac{1}{2\pi\hbar} \left(\frac{2\pi\hbar m}{i\Delta t} \right)^{1/2} e^{-m(x-x_0)^2/2\hbar i\Delta t} \\ &= \left(\frac{m}{2\pi\hbar i\Delta t} \right)^{1/2} e^{-m(x-x_0)^2/2\hbar i\Delta t} \end{aligned}$$

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.

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 $L = T - V = T = \frac{1}{2}m\dot{x}^2$)

$$S_{cl} = S[x_{cl}(t')] \quad (12.119)$$

$$= \int_{t_0}^t L(x_{cl}, \dot{x}_{cl}) dt' \quad (12.120)$$

$$= \int_{t_0}^t \frac{1}{2}m\dot{x}_{cl}(t') dt' \quad (12.121)$$

$$= \frac{1}{2}m \left(\frac{x-x_0}{t-t_0} \right)^2 (t-t_0) \quad (12.122)$$

$$= \frac{m(x-x_0)^2}{2(t-t_0)} \quad (12.123)$$

Figure 12.4: least action

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

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

$$\begin{aligned} \delta S[x(t')] &= S[x(t') + \eta(t')] - S[x(t')] \\ &= \int_{t_0}^t \left[\frac{\partial L(x, \dot{x})}{\partial x} \eta(t') + \frac{\partial L(x, \dot{x})}{\partial \dot{x}} \dot{\eta}(t') \right] dt' \\ &= \int_{t_0}^t \left[\frac{\partial L}{\partial x} \eta(t') + \frac{d}{dt'} \left(\frac{\partial L}{\partial \dot{x}} \eta(t') \right) - \frac{d}{dt'} \left(\frac{\partial L}{\partial \dot{x}} \right) \eta(t') \right] dt' \\ &= \int_{t_0}^t \left[\frac{\partial L}{\partial x} - \frac{d}{dt'} \left(\frac{\partial L}{\partial \dot{x}} \right) \right] \eta(t') dt' + \int_{t_0}^t \left[\frac{d}{dt'} \left(\frac{\partial L}{\partial \dot{x}} \eta(t') \right) \right] dt' \end{aligned}$$

Figure 12.5: paths

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

$$\int_{t_0}^t \left[\frac{\partial L}{\partial x} - \frac{d}{dt'} \left(\frac{\partial L}{\partial \dot{x}} \right) \right]_{cl} \eta(t') dt' = 0 \quad (12.132)$$

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

$$U = \sum_{all\ paths} e^{\frac{i}{\hbar} S[x(t')]} \quad (12.136)$$

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

$$Z_a = e^{iS[x_a(t')]/\hbar} \quad (12.137)$$

Figure 12.6: alternative paths

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 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 $\Delta S = 0.17 \text{ erg sec} \approx 1.7 \times 10^{26} \hbar \gg \pi \hbar$.

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.

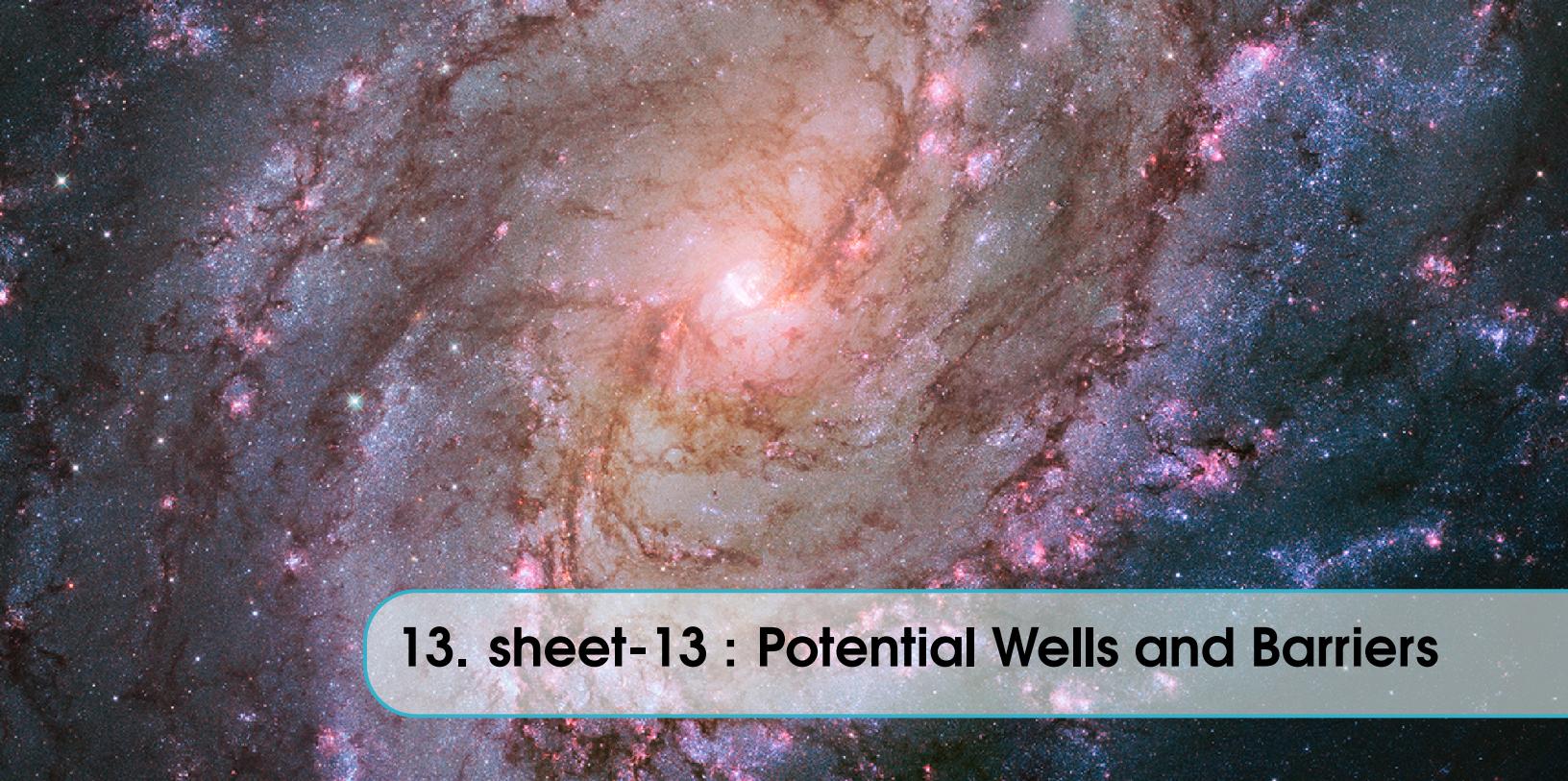
12.9 Equivalence to the Schrodinger Equation

12.10 Potentials of the Form $V = a + bx + cx^2 + dx\dot{x} + ex\ddot{x}$

12.10.1 Special Cases

1. Free Particle

2. Harmonic Oscillator



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



17. sheet-17 : Angular Momentum 2: spin 1/2



18. sheet-18 : Rotations and Angular Momentum



19. sheet-19 : Addition of Angular Momentum



20. sheet-20 : Time Independent Perturbation



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

Integrals

.1 List of Standard Integrals

$$\int_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\pi} \quad (1)$$

$$\int_{-\infty}^{\infty} e^{-\alpha x^2 + \beta x} dx = \left(\frac{\pi}{2}\right)^{1/2} \exp\left[\frac{\beta^2}{4\alpha}\right] \quad (2)$$

.2 Proofs

$$\int_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\pi} \quad (3)$$

proof:

let,

$$I = \int_{-\infty}^{\infty} e^{-x^2} dx \quad (4)$$

then

$$I^2 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-(x^2+y^2)} dx dy \quad (5)$$

using $r^2 = x^2 + y^2$ since

$$x = r \cos \theta \quad (6)$$

$$y = r \sin \theta \quad (7)$$

we get

$$I^2 = \int_0^{2\pi} \int_0^{\infty} e^{-r^2} r dr d\theta \quad (8)$$

$$= 2\pi \int_0^{\infty} e^{-r^2} r dr \quad (9)$$

let $u = r^2$ therefore $r dr = du/2$

$$I^2 = \pi \int_0^{\infty} e^{-u} du \quad (10)$$

$$= \pi \left[\frac{e^{-u}}{-1} \right]_0^\infty \quad (11)$$

$$= \pi [0 - (-1)] \quad (12)$$

$$= \pi \quad (13)$$

therefore $I = \sqrt{\pi}$

another

$$I = \int_{-\infty}^{\infty} dx e^{-(b^2 x^2 + ax)} \quad (14)$$

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

$$I = e^{a^2/4b^2} \int_{-\infty}^{\infty} dx e^{-(bx + \frac{a}{2b})^2} \quad (15)$$

let $z = bx + \frac{a}{2b}$ thus $dz = b dx$

$$I = e^{a^2/4b^2} \frac{1}{b} \int_{-\infty}^{\infty} dz e^{-z^2} \quad (16)$$

$$= e^{a^2/4b^2} \frac{\sqrt{\pi}}{b} \quad (17)$$