

# Notes on The Theoretical Minimum — Quantum Mechanics

Gui-Rong Liang

December 27, 2018

## Contents

<b>1</b>	<b>Mathematical Foundations and Spin systems</b>	<b>2</b>
1.1	Basic concepts and fundamental principles . . . . .	2
1.2	Spin states, spin operators, and Pauli Matrices . . . . .	5
1.3	States and Operators in Continuous basis . . . . .	7
<b>2</b>	<b>Time evolution of quantum states</b>	<b>11</b>
2.1	Time evolution of isolated states and expectation values . . . . .	11

# 1 Mathematical Foundations and Spin systems

## 1.1 Basic concepts and fundamental principles

### 1. Quantum states, operators, Hermitian properties

A quantum state is represented by a ket vector that can be decomposed using the orthogonal basis vectors, in this subsection we mainly focus on discrete basis, thus

$$|A\rangle = \sum_j \alpha_j |j\rangle, \quad \text{with} \quad \alpha_j = \langle j|A\rangle, \quad (1)$$

where the coefficient  $\alpha_j$  is a complex number, and is also called a wave function. Thus

$$|A\rangle = \sum_j |j\rangle \langle j|A\rangle \implies \sum_j |j\rangle \langle j| = 1. \quad (2)$$

The probability of measuring a state corresponding to a certain basis is given by

$$P_j = \alpha_j^* \alpha_j = \langle A|j\rangle \langle j|A\rangle, \quad (3)$$

then the normalization of probability gives to

$$1 = \sum_j P_j = \sum_j \langle A|j\rangle \langle j|A\rangle = \langle A|A\rangle. \quad (4)$$

So the general principle is the state of a system is represented by a unit vector.

Now we introduce the matrix representation of an operator, which turn a vector to another vector,

$$M|A\rangle = |B\rangle. \quad (5)$$

If we decompose A and B into components on basis vectors, we can represent the above equation as,

$$\sum_j M|j\rangle \alpha_j = \sum_j \beta_j |j\rangle, \quad (6)$$

multiplying the above by a basis bra, we have

$$\sum_j \langle k|M|j\rangle \alpha_j = \sum_j \beta_j \langle k|j\rangle = \beta_k, \quad (7)$$

which is denoted by

$$\sum_j M_{kj} \alpha_j = \beta_k, \quad (8)$$

where  $M_{ij} \equiv \langle i|M|j\rangle$  is matrix element of the operator.

We write the above as explicit matrix form (taking 2 dimension as an example) as

$$\begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} = \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix}. \quad (9)$$

Now we take its complex conjugate and make it a equation of row vectors,

$$\begin{pmatrix} \alpha_1^* & \alpha_2^* \end{pmatrix} \begin{pmatrix} M_{11}^* & M_{21}^* \\ M_{12}^* & M_{22}^* \end{pmatrix} = \begin{pmatrix} \beta_1^* & \beta_2^* \end{pmatrix}, \quad (10)$$

which can be represented by the bras notation,

$$\langle A|M^\dagger = \langle B|, \quad (11)$$

where  $M^\dagger$  is called the *Hermitian conjugate* of  $M$ , with the elements of  $M^\dagger$  is given by

$$(M^\dagger)_{jk} = M_{kj}^*. \quad (12)$$

An Hermitian operator is an operator which equals to its own Hermitian conjugate,

$$M = M^\dagger. \quad (13)$$

Physical observable quantities are represented by Hermitian operators, due to the fact that the eigenvalues of Hermitian operators are real. And further, eigenvectors of an Hermitian operator corresponding to different eigenvalues are orthogonal. Next we will prove this.

If

$$L|\lambda\rangle = \lambda|\lambda\rangle, \quad (14)$$

where  $L$  is Hermitian, we have

$$\lambda^* \langle \lambda | \lambda \rangle = \langle \lambda | L^\dagger | \lambda \rangle = \langle \lambda | L | \lambda \rangle = \lambda \langle \lambda | \lambda \rangle \implies \lambda^* = \lambda, \quad (15)$$

Thus  $\lambda$  is real. Then if

$$\begin{cases} L|\lambda_1\rangle = \lambda_1|\lambda_1\rangle \\ L|\lambda_2\rangle = \lambda_2|\lambda_2\rangle, \end{cases} \quad (16)$$

we have

$$\lambda_1 \langle \lambda_1 | \lambda_2 \rangle = \langle \lambda_1 | L | \lambda_2 \rangle = \lambda_2 \langle \lambda_1 | \lambda_2 \rangle \implies (\lambda_1 - \lambda_2) \langle \lambda_1 | \lambda_2 \rangle = 0 \implies \langle \lambda_1 | \lambda_2 \rangle = 0. \quad (17)$$

Even if eigenvectors of the same eigenvalue are not orthogonal, we can make it orthogonal by Gram-Schmidt procedure. Finally we can make all eigenvectors of an Hermitian operator orthogonal, thus they can be used as a basis for the vector space.

Another inference is that if two operators  $L$ ,  $M$  has the same set of eigenvectors, they commute. We'll use the labels  $\lambda_i$  and  $\mu_\alpha$  to denote the eigenvalues of  $L$  and  $M$ , and assume that there is a basis of state-vectors  $|\lambda_i, \mu_\alpha\rangle$  that are simultaneous eigenvectors of both observables.

$$\begin{aligned} L|\lambda_i, \mu_\alpha\rangle &= \lambda_i|\lambda_i, \mu_\alpha\rangle \\ M|\lambda_i, \mu_\alpha\rangle &= \mu_\alpha|\lambda_i, \mu_\alpha\rangle, \end{aligned} \quad (18)$$

the by acting on any of the basis vector, which we omit the subscript later, by the two operators,

$$LM|\lambda, \mu\rangle = \lambda\mu|\lambda, \mu\rangle = ML|\lambda_i, \mu_\alpha\rangle \implies [L, M]|\lambda, \mu\rangle = 0, \quad (19)$$

since any vector is composed of the basis vectors,  $[L, M]$  acting on any vector gives zero,

$$[L, M] = 0, \quad (20)$$

thus they commute. It turns out that the converse of this theorem is also true: if two observables commute, then there is a complete basis of simultaneous eigenvectors of the two observables. Moreover, one may need to specify a larger number of observables to completely label a basis, regardless of the number of observables that are needed, they must all commute, we call this collection a *complete set of commuting observables*.

## 2. The Uncertainty Principle

Suppose the eigenvalues of the observable  $A$  is called  $a$ , then the expectation value of  $A$  is

$$\langle A \rangle = \sum_a a P(a). \quad (21)$$

Roughly speaking this means that  $P(a)$  is centered around the expectation value. What we will mean by "the uncertainty in  $A$  is the so-called *standard deviation*  $\bar{A}$ , which is computed by subtracting from  $A$  its expectation value,

$$\bar{A} = A - \langle A \rangle = A - \langle A \rangle I. \quad (22)$$

The eigenvectors of  $\bar{A}$  are the same as those of  $A$  and the eigenvalues of  $\bar{A}$  is shifted from that of  $A$  by the average value  $\langle A \rangle$ ,

$$\bar{A} = a - \langle A \rangle. \quad (23)$$

The square of the uncertainty of  $A$  is defined by

$$(\Delta A)^2 := \sum_a \bar{A}^2 P(a) = \sum_a (a - \langle A \rangle)^2 P(a) = \langle \Psi | \bar{A}^2 | \Psi \rangle, \quad (24)$$

if the expectation value of  $A$  is zero, then the uncertainty takes the simpler form

$$(\Delta A)^2 = \langle \Psi | A^2 | \Psi \rangle. \quad (25)$$

If  $\vec{X}$  and  $\vec{Y}$  are two vectors, then the *Triangle Inequality* gives

$$|\vec{X}| + |\vec{Y}| \geq |\vec{X} + \vec{Y}|, \quad (26)$$

by squaring and sorting it we have

$$|\vec{X}| |\vec{Y}| \geq \vec{X} \cdot \vec{Y}, \quad (27)$$

and by squaring it again we have

$$|\vec{X}|^2 |\vec{Y}|^2 \geq |\vec{X} \cdot \vec{Y}|^2, \quad (28)$$

which is called the *Cauchy-Schwarz* inequality.

For complex vector spaces, the magnitude of the vectors are given by

$$\begin{aligned} |X| &= \sqrt{\langle X | X \rangle} \\ |X + Y| &= \sqrt{(\langle X | + \langle Y |)(|X\rangle + |Y\rangle)}, \end{aligned} \quad (29)$$

squaring them and using the Triangle Inequality we have

$$2|X||Y| \geq |\langle X | Y \rangle + \langle Y | X \rangle|. \quad (30)$$

Now we take  $|X\rangle$  and  $|Y\rangle$  as follows,

$$\begin{aligned} |X\rangle &= A|\Psi\rangle \\ |Y\rangle &= iB|\Psi\rangle, \end{aligned} \quad (31)$$

where  $A$  and  $B$  are any two observables and  $|\Psi\rangle$  is any ket. Now substitute it into the Triangle Inequality for complex vectors, to get

$$2\sqrt{\langle A^2 \rangle \langle B^2 \rangle} \geq |\langle \Psi | AB | \Psi \rangle - \langle \Psi | BA | \Psi \rangle| = \langle \Psi | [A, B] | \Psi \rangle. \quad (32)$$

Let's suppose for the moment that  $A$  and  $B$  have expectation values of zero, so  $A = \bar{A}$ , and  $\langle A^2 \rangle = (\Delta A)^2$ , thus we rewrite the above inequality as

$$\Delta A \Delta B \geq \frac{1}{2} \langle \Psi | [A, B] | \Psi \rangle. \quad (33)$$

This is the general version of the uncertainty inequality. It says that the product of the uncertainties cannot be smaller than half the magnitude of the expectation value of the commutator. Or to be less quantitative, if the commutator of  $A$  and  $B$  is not zero, then both observables cannot simultaneously be certain. By the way, if the expectation values of  $A$  and  $B$  are not zero, we just use a trick to redefine two new operators in which the expectation values have been subtracted off, then the same relation holds as we expected.

Later we will come to a specific version of the inequality — *Heisenberg's Uncertainty Principle*: The product of the uncertainties of the position and momentum of a particle cannot be less than half of Plank's constant.

## 1.2 Spin states, spin operators, and Pauli Matrices

All possible spin states can be represented in a two dimensional vector space.

$$|A\rangle = \alpha_u|u\rangle + \alpha_d|d\rangle, \quad (34)$$

where the  $u$  and  $d$  represents for pointing up and down when we measure spin along z-axis, and the orthogonality means if we get up we never get down, vice versa.

Now we choose A to be pointing left and right, considering that they're orthonormal, and it is equally likely to be up and down when measured along z, we write them as

$$\begin{cases} |r\rangle = \frac{1}{\sqrt{2}}|u\rangle + \frac{1}{\sqrt{2}}|d\rangle \\ |l\rangle = \frac{1}{\sqrt{2}}|u\rangle - \frac{1}{\sqrt{2}}|d\rangle. \end{cases} \quad (35)$$

And when we represent forward and backward, we must further take into account the equal likelihood of left and right when measured along x-axis, so we take the following form,

$$\begin{cases} |f\rangle = \frac{1}{\sqrt{2}}|u\rangle + \frac{i}{\sqrt{2}}|d\rangle \\ |b\rangle = \frac{1}{\sqrt{2}}|u\rangle - \frac{i}{\sqrt{2}}|d\rangle. \end{cases} \quad (36)$$

Notice that all the above choices have the same freedom by a phase factor  $e^{i\theta}$ .

We will use  $\sigma$  to denote spin operator. Measuring along z-axis gives us

$$\begin{cases} \sigma_z|u\rangle = |u\rangle \\ \sigma_z|d\rangle = -|d\rangle, \end{cases} \quad (37)$$

if we denote  $|u\rangle$  and  $|d\rangle$  as unit vectors as  $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$  and  $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ , we may figure out that

$$\sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (38)$$

Measuring along x-axis gives us

$$\begin{cases} \sigma_x|r\rangle = |r\rangle \\ \sigma_x|l\rangle = -|l\rangle, \end{cases} \quad (39)$$

if we denote  $|r\rangle$  and  $|l\rangle$  as unit vectors as  $\begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}$  and  $\begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix}$ , we may figure out that

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \quad (40)$$

Measuring along y-axis gives us

$$\begin{cases} \sigma_y|f\rangle = |f\rangle \\ \sigma_y|b\rangle = -|b\rangle, \end{cases} \quad (41)$$

if we denote  $|f\rangle$  and  $|b\rangle$  as unit vectors as  $\begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{i}{\sqrt{2}} \end{pmatrix}$  and  $\begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{i}{\sqrt{2}} \end{pmatrix}$ , we may figure out that

$$\sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}. \quad (42)$$

The matrix representations of  $\sigma_x$ ,  $\sigma_y$  and  $\sigma_z$  are called the *Pauli Matrices*. It is easy and useful to verify the commutation relations of components of spins,

$$[\sigma_j, \sigma_k] = 2i\epsilon_{jkl}\sigma_l. \quad (43)$$

Now we may treat  $\sigma_x$ ,  $\sigma_y$  and  $\sigma_z$  as three components of  $\sigma$ , thus  $\sigma$  can be viewed as a general vector, and it can be projected onto a unit vector  $\hat{n}$  point any direction in space, so as to measure spin along that direction.

$$\sigma_n = \vec{\sigma} \cdot \hat{n} = \sigma_x n_x + \sigma_y n_y + \sigma_z n_z = \begin{pmatrix} n_z & n_x - in_y \\ n_x + in_y & -n_z \end{pmatrix}, \quad (44)$$

from the fact that

$$\begin{aligned} \text{tr}(\sigma_n) &= 0 \\ \det(\sigma_n) &= -1, \end{aligned} \quad (45)$$

we could immediately figure out that the two eigenvalues of  $\sigma_n$  are give by  $\pm 1$ . And we could work out its eigenvectors in a certain configuration, if  $\hat{n}$  lies x-z plane, we may write

$$\begin{cases} n_z = \cos \theta \\ n_x = \sin \theta \\ n_y = 0, \end{cases} \quad (46)$$

thus

$$\sigma_n = \begin{pmatrix} \cos \theta & \sin \theta \\ \sin \theta & -\cos \theta \end{pmatrix}, \quad (47)$$

so the eigenvectors correspond to each eigenvalues are given by

$$\begin{cases} \lambda_1 = 1, & |\lambda_1\rangle = \begin{pmatrix} \cos \frac{\theta}{2} \\ \sin \frac{\theta}{2} \end{pmatrix} \\ \lambda_2 = -1, & |\lambda_1\rangle = \begin{pmatrix} -\sin \frac{\theta}{2} \\ \cos \frac{\theta}{2} \end{pmatrix}. \end{cases} \quad (48)$$

Now suppose we prepare a  $|u\rangle$  state, and we measure the probabilities of  $+1$  and  $-1$  along the  $\hat{n}$  direction, we get

$$\begin{cases} P(+1) = \langle u | \lambda_1 \rangle \langle \lambda_1 | u \rangle = \cos^2 \frac{\theta}{2} \\ P(-1) = \langle u | \lambda_2 \rangle \langle \lambda_2 | u \rangle = \sin^2 \frac{\theta}{2} \end{cases} \quad (49)$$

Now we're ready to compute the average value of spin along the  $n$  direction,

$$\langle \sigma_n \rangle = \sum_j \lambda_j P(\lambda_j) = \cos^2 \frac{\theta}{2} - \sin^2 \frac{\theta}{2} = \cos \theta, \quad (50)$$

which agrees perfectly with our expectation.

To proceed, we come to an important theorem,

**The Spin-Polarization Principle:** Any state of a single spin is an eigenvector of some component of the spin.

In other words, given any state

$$|A\rangle = \alpha_u |u\rangle + \alpha_d |d\rangle, \quad (51)$$

there exists some direction  $\hat{n}$ , such that

$$\vec{\sigma} \cdot \vec{n}|A\rangle = |A\rangle. \quad (52)$$

In physics language, we say that the states of a spin are characterized by a *polarization vector*, and along that polarization vector the component of spin is predictably  $+1$ , assuming of course that you know the state-vector. It follows that the expectation value of the spin along the direction  $\hat{n}$  can be expressed as

$$\langle \sigma_n \rangle = \langle \vec{\sigma} \cdot \vec{n} \rangle = 1. \quad (53)$$

In fact, any observable of a spin is represented by a  $2 \times 2$  Hermitian matrix, and has the form

$$\begin{pmatrix} r & w \\ w^* & r' \end{pmatrix}, \quad (54)$$

the implication is that it takes exactly four real parameters to specify this observable. And, there is a neat way to write any spin observable in terms of the Pauli matrices,  $\sigma_x$ ,  $\sigma_y$ ,  $\sigma_z$ , and the unit matrix  $I$ ,

$$S = a\sigma_x + b\sigma_y + c\sigma_z + dI, \quad (55)$$

where  $a$ ,  $b$ ,  $c$  and  $d$  are real numbers.

### 1.3 States and Operators in Continuous basis

#### 1. States and the inner product

A particle at a fixed position  $x_0$  (in one dimension) can be represented by a state vector  $|x_0\rangle$ . In general, a state is a superposition of various position state, since position is continuous variable, we must turn the sum into integral,

$$|\psi\rangle = \int dx \psi(x)|x\rangle, \quad (56)$$

where the coefficient, or wave function  $\psi(x)$  is computed as

$$\psi(x) = \langle x|\psi\rangle, \quad (57)$$

substituting the coefficient into the decomposition we have

$$|\psi\rangle = \int dx |x\rangle\langle x|\psi\rangle, \quad (58)$$

from which we find an identity as

$$I = \int dx |x\rangle\langle x|, \quad (59)$$

which implies summing over all projections  $|x\rangle\langle x|$ , we can recover the original state.

Left multiplying the decomposition of state by a bra vector  $\langle x'|$ , we have

$$\psi(x') = \langle x'|\psi\rangle = \int dx \psi(x) \langle x'|x\rangle, \quad (60)$$

from which we can see the inner product between basis is nothing but the Dirac-delta function,

$$\langle x'|x\rangle = \delta(x' - x), \quad (61)$$

this is the orthonormalization relation of the basis vectors.

Using the above relations we may naturally find the expression of inner product of states in the continuous position basis, just by inserting two different sets of the identity relations of  $x$  and  $x'$ ,

$$\langle \phi | \psi \rangle = \int dx \int dx' \langle \phi | x' \rangle \langle x' | x \rangle \langle x | \psi \rangle = \int dx \int dx' \phi^*(x') \delta(x' - x) \psi(x) = \int dx \phi^*(x) \psi(x), \quad (62)$$

thus the inner product of a state with itself is given by

$$\langle \psi | \psi \rangle = \int dx \psi^*(x) \psi(x), \quad (63)$$

this will remind us of the expression of the probability, and due to the continuous property we will denote our corresponding concept as the "probability density", which is given by

$$P(x) := \psi^*(x) \psi(x), \quad (64)$$

and the probability of state in an interval  $[a, b]$  is given by

$$P([a, b]) = \int_a^b dx P(x) = \int_a^b dx \psi^*(x) \psi(x), \quad (65)$$

and because the total probability equal to 1, our state here is normalized,

$$\langle \psi | \psi \rangle = \int dx \psi^*(x) \psi(x) = \int dx P(x) = 1. \quad (66)$$

## 2. Linear Operators, eigenvectors and eigenvalues

An operator is a machine which can turn a given state into a new state, in terms of the continuous basis, if the coefficients (or the wave function) of the new state is specified, the new state would be determined, thus the effect of the operator is made clear.

Let's set an example. We want to define an operator  $X$  which can turn a state  $|\psi\rangle$  in to a new state  $|\bar{\psi}\rangle$ , now the state  $|\psi\rangle$  is given, so the definition of  $X$  is clear only if we know what the new state  $|\bar{\psi}\rangle$  is. By "the state  $|\psi\rangle$  is given", we mean in terms of any basis, we know exactly what the coefficients of the state  $|\psi\rangle$  are, here we choose the continuous position basis, where we know exactly the wave function  $\psi(x)$ . So the task now shifts from defining  $X$  to giving the expression of  $\bar{\psi}(x)$ , and we define it to be  $\bar{\psi}(x) = x\psi(x)$ , thus we say we complete the definition of the operator  $X$ .

By mathematical expressions, we write the definition of  $X$  as

$$X|\psi\rangle = |\bar{\psi}\rangle \quad \text{s.t.} \quad \bar{\psi}(x) = x\psi(x), \quad (67)$$

or in a more neat form,

$$\langle x | X | \psi \rangle := x \langle x | \psi \rangle. \quad (68)$$

Here we say that  $x$  is a *basis-dependent operator corresponding to  $X$* . Note that an operator only acts on a state in the abstract vector space, while a basis-dependent operator acts on the wave function. Due to the fact that both the abstract state vector and the wave function can be multiplied by a number, sometimes we blur the distinction of the two concept and just write  $X = x$ , but this is not generally true. The clear relation between  $X$  and  $x$  can be seen by inserting an identity relation to the definition, as

$$\int dx' X_{xx'} \psi(x') = \int dx' \langle x | X | x' \rangle \langle x' | \psi \rangle = \langle x | X | \psi \rangle = x \langle x | \psi \rangle = x \psi(x), \quad (69)$$

and by observing we have

$$\begin{cases} X_{xx'} = \langle x | X | x' \rangle = x \delta(x - x') \\ x = \int dx' X_{xx'} = \int dx' \langle x | X | x' \rangle, \end{cases} \quad (70)$$



which clearly shows in position basis  $X$  is a matrix with only non-zero diagonalized elements  $x\delta(x-x')$ , and  $x$  is the sum of the elements of the  $x'$ th column.

Then by bearing the above in mind, we proceed to define a new operator  $D$  as

$$\langle x|D|\psi\rangle := \frac{\partial}{\partial x}\langle x|\psi\rangle, \quad (71)$$

where the partial derivative  $\frac{\partial}{\partial x}$  is the basis-dependent operator corresponding to  $D$ . Notice that now we CANNOT write  $D = \frac{\partial}{\partial x}$ , because  $\frac{\partial}{\partial x}|\psi\rangle$  is meaningless due to the basis-independence of  $|\psi\rangle$ . The relation between  $D$  and  $\frac{\partial}{\partial x}$  can be seen by the similar process as in the case of  $X$  and  $x$ , and here we directly list the results:

$$\begin{cases} D_{xx'} = \langle x|D|x'\rangle = \delta(x-x')\frac{\partial}{\partial x} \\ \frac{\partial}{\partial x} = \int dx' D_{xx'} = \int dx' \langle x|D|x'\rangle. \end{cases} \quad (72)$$

Operators with physical interests are Hermitian operators. A convenient way to judge the hermitian property is by using the inner product. The inner product is number, assuming  $O$  is hermitian, we take the complex conjugate of the inner product,

$$\langle\phi|O|\psi\rangle^* = \langle\psi|O^\dagger|\phi\rangle = \langle\psi|O|\phi\rangle, \quad (73)$$

finding that switching the left and right vector gives the complex conjugate of the original, and this will serve as the criteria for judging the hermitian property in the following.

So are  $X$  and  $D$  hermitian? We first deal with  $X$ ,

$$\begin{aligned} \langle\phi|X|\psi\rangle^* &= \left[ \int dx \langle\phi|x\rangle\langle x|X|\psi\rangle \right]^* = \left[ \int dx \phi^*(x) x\psi(x) \right]^* \\ &= \int dx \psi^*(x) x^*\phi(x) = \int dx \psi^*(x) x\phi(x) \\ &= \langle\psi|X|\phi\rangle, \end{aligned} \quad (74)$$

where we used the fact that the position variable is real. Thus the position operator is indeed Hermitian.

And next the  $D$  operator,

$$\begin{aligned} \langle\phi|D|\psi\rangle^* &= \left[ \int dx \langle\phi|x\rangle\langle x|D|\psi\rangle \right]^* = \left[ \int dx \phi^*(x) \frac{\partial}{\partial x}\psi(x) \right]^* \\ &= \left[ \int_{-\infty}^{\infty} dx \psi^*(x) \phi(x) - \int dx \frac{\partial}{\partial x}\psi^*(x) \phi(x) \right]^* = - \int dx \phi^*(x) \frac{\partial}{\partial x}\psi(x) \\ &= -\langle\psi|D|\phi\rangle \implies D^\dagger = -D, \end{aligned} \quad (75)$$

where we have used the boundary condition that the wave functions vanish at infinity. Thus the  $D$  operator is not Hermitian, but anti-Hermitian. But actually we still want to construct an Hermitian operator out of  $D$ , recall that the easiest way to construct an Hermitian operator out of an anti-Hermitian operator is multiplying it by  $i$  or  $-i$ . So we define a new operator as

$$P := -i\hbar D, \quad (76)$$

and it's easy to check that  $P$  is Hermitian,

$$\langle\phi|P|\psi\rangle^* = [-i\hbar\langle\phi|D|\psi\rangle]^* = i\hbar[-\langle\psi|D|\phi\rangle] = \langle\psi|P|\phi\rangle, \quad (77)$$

and its basis-dependent counterpart is naturally given by  $\hat{p} = -i\hbar\frac{\partial}{\partial x}$ , which can act directly on the wave function,

$$\hat{p}\psi(x) = -i\hbar\frac{\partial}{\partial x}\psi(x). \quad (78)$$

Next we explore the eigenvalues and eigenvectors of  $X$  and  $P$ . Eigenvectors of an operator are, by definitions, vectors that acted by the operator gives the same vector multiplied by a number which is called the eigenvalue. For the  $X$  operator, we are looking for a solution that satisfies,

$$X|\psi\rangle = x_0|\psi\rangle, \quad (79)$$

where  $x_0$  is a specific number. By definition, the LHS in  $x$ -basis is given by

$$\langle x|X|\psi\rangle = x\langle x|\psi\rangle = x\psi(x), \quad (80)$$

and the RHS in the  $x$ -basis would be

$$\langle x|x_0|\psi\rangle = x_0\langle x|\psi\rangle = x_0\psi(x). \quad (81)$$

So the equation we're going to solve is nothing but

$$x\psi(x) = x_0\psi(x), \quad (82)$$

and when obtaining the expression for  $\psi(x)$ , we can easily recover the state vector by the superposition principle,  $|\psi\rangle = \int dx \psi(x)|x\rangle$ .

By rearranging the equation to solve, we get

$$(x - x_0)\psi(x) = 0, \quad (83)$$

which says that  $\psi(x)$  could be non-zero only when  $x = x_0$ , or equivalently, when  $x \neq x_0$ ,  $\psi(x)$  must be zero. We summarize this as follows,

$$\psi(x) = \begin{cases} (\text{non-})\text{zero}, & x = x_0 \\ 0, & x \neq x_0, \end{cases} \quad (84)$$

this reminds us of the Dirac-delta function  $\delta(x - x_0)$ , but we didn't strictly prove that  $\psi(x)$  is indeed  $\delta(x - x_0)$ . Since a straightforward proof is a little troublesome and will lead to some ambiguities (as I tried), we go around that and try to find another clever way. As all we want is a function that satisfies  $x\psi(x) = x_0\psi(x)$ , if we somehow happen to find such a function, we can say that it is the eigenfunction of the operator. The trick here is to use the relation between  $X$  and  $x$ , as follows,

$$x_0\delta(x - x_0) = [x_0\delta(x_0 - x)]^* = \langle x_0|X|x\rangle^* = \langle x|X|x_0\rangle = x\delta(x - x_0), \quad (85)$$

where we have used the Hermitian property of  $X$ , the realness of  $x_0$  and  $\delta(x_0 - x)$ , and the evenness of  $\delta(x_0 - x)$ . Now we see that the Dirac-delta function  $\delta(x - x_0)$  indeed is a solution to the equation, so it is the eigenfunction,  $\psi(x) = \delta(x - x_0)$ , now we recover our eigenvector as

$$|\psi\rangle = \int dx \psi(x)|x\rangle = \int dx \delta(x - x_0)|x\rangle = |x_0\rangle, \quad (86)$$

and finally our eigen-equation will be

$$X|x_0\rangle = x_0|x_0\rangle. \quad (87)$$

Since  $x_0$  can be any value along the  $x$ -axis, this equation can be generalized to

$$X|x\rangle = x|x\rangle, \quad (88)$$

which tells us the simple fact that any state representing a definite position is an eigenvector of the position operator  $X$ , and the eigenvalue is the position coordinate.

Then we play around with  $P$ . Similarly the eigen-equation is given by

$$P|\psi\rangle = p_0|\psi\rangle, \quad (89)$$

and in  $x$ -basis,

$$p_0\psi(x) = p_0\langle x|\psi\rangle = \langle x|p_0|\psi\rangle = \langle x|P|\psi\rangle = -i\hbar\frac{\partial}{\partial x}\langle x|\psi\rangle = -i\hbar\frac{\partial}{\partial x}\psi(x), \quad (90)$$

which is nothing but a simple first-order equation. The solution is given by

$$\psi(x) = e^{\frac{i}{\hbar}p_0x}, \quad (91)$$

and the eigenvector is constructed as

$$|\psi\rangle = \int dx \psi(x)|x\rangle = \int dx e^{\frac{i}{\hbar}p_0x}|x\rangle \equiv |p_0\rangle, \quad (92)$$

which is just the Fourier-transformation of the state  $|x\rangle$ . So the eigen-equation will be recovered as

$$P|p_0\rangle = p_0|p_0\rangle \quad \text{or} \quad P|p\rangle = p|p\rangle, \quad (93)$$

where the physical meaning of  $p$  and the operator  $P$  will be clear later when we take the classical limit. As an aside, here we get a commonly used relation by re-express the solution,

$$\langle x|p\rangle = \langle x|\psi\rangle = \psi(x) = e^{\frac{i}{\hbar}px}, \quad \text{and} \quad \langle p|x\rangle = e^{-\frac{i}{\hbar}px}. \quad (94)$$

## 2 Time evolution of quantum states

### 2.1 Time evolution of isolated states and expectation values

In classical mechanics, if two identical isolated systems start out in different states, they stay in different states. It underlies the fact that information is never lost. The same is true in quantum mechanics, and even much stronger. The direct manifestation is that if two states are initially orthogonal, they remain orthogonal throughout the evolution, or *conservation of orthogonality*.

We denote a state at time  $t$  as  $|\Psi(t)\rangle$ , which originates from the state at time  $t = 0$ ,  $|\Psi(0)\rangle$ , by acting an *time-evolution operator*  $\hat{U}(t)$  on it,

$$|\Psi(t)\rangle = \hat{U}(t)|\Psi(0)\rangle. \quad (95)$$

This equation reflects the fact that quantum evolution of states allows us to compute the probabilities of the outcomes of later experiments, in contrast to the case in classical mechanics, classical determinism allows us to predict the results of experiments.

Mathematically, the conservation of orthogonality tells us

$$\langle\Psi(0)|\Phi(0)\rangle = 0 \implies \langle\Psi(t)|\Phi(t)\rangle = \langle\Psi(0)|\hat{U}^\dagger(t)\hat{U}(t)|\Phi(0)\rangle = 0, \quad (96)$$

if we choose  $|\Psi(0)\rangle$  and  $|\Phi(0)\rangle$  to be any two orthonormal basis of vectors  $|i\rangle$  and  $|j\rangle$ , we could obtain a property of the evolution operator,

$$\langle i|j\rangle = \delta_{ij} = \langle i|\hat{U}^\dagger(t)\hat{U}(t)|j\rangle \implies \hat{U}^\dagger(t)\hat{U}(t) = I, \quad (97)$$

with which the evolution operator satisfies is called *unitary*.

*Principle: The evolution of state-vectors with time is unitary.*

A direct consequence of the unitarity leads to a much stronger version of conservation of orthogonality — *conservation of overlaps*,

$$\langle A(t)|B(t)\rangle = \langle A(0)|\hat{U}^\dagger(t)\hat{U}(t)|B(0)\rangle = \langle A(0)|B(0)\rangle, \quad (98)$$

which says the inner product of  $|A\rangle$  and  $|B\rangle$  does not change with time, or the inner product is conserved. Thus superficially conservation of orthogonality is a special case of conservation of overlap, we have seen that conservation of orthogonality can lead to conservation of overlap, thus they are equivalent.

Time evolution is led by incremental changes, so it would be inspiring to study the evolution operator of an infinitesimal time interval  $\epsilon$ . Another property besides unitarity that an evolution operator should have is *continuity*, which means that the state-vector changes smoothly, or to be more explicitly, when  $\epsilon$  is very small,  $\hat{U}(\epsilon)$  is close to the unit operator, only differing from it by something of order  $\epsilon$ , and when  $\epsilon$  is taken to be zero, the state does not change.

$$U(\epsilon) = I - i\epsilon H, \quad (99)$$

where the  $-i$  and the operator  $H$  is arbitrary and conventional. By plugging its Hermitian conjugate form

$$U^\dagger(\epsilon) = I + i\epsilon H^\dagger \quad (100)$$

into the unitarity equation and expanding it to the first power, we have

$$I = U^\dagger(\epsilon)U(\epsilon) = (I + i\epsilon H^\dagger)(I - i\epsilon H) = I + i\epsilon(H^\dagger - H), \quad (101)$$

or

$$H^\dagger = H. \quad (102)$$

This Hermitian property of  $H$  has the great significance that  $H$  is an observable and has a complete set of orthonormal eigenvalues and eigenvectors.

Here I want to explain the above ansatz in more detail, which is due to my own understanding and I think is more natural.

The most general form of a state obtained by acting the infinitesimal evolution operator on the initial state, considering continuity, takes the form as

$$U(\epsilon)|\Psi\rangle = |\Psi\rangle + \epsilon|\Theta\rangle, \quad (103)$$

where  $|\Theta\rangle$  is some unknown state though, it can be obtained by acting some operator to the original state  $|\Theta\rangle \equiv A|\Psi\rangle$ , so we have

$$U(\epsilon)|\Psi\rangle = |\Psi\rangle + \epsilon A|\Psi\rangle = (I + \epsilon A)|\Psi\rangle, \quad (104)$$

then we have a form of  $U(\epsilon)$  as

$$U(\epsilon) = I + \epsilon A, \quad (105)$$

where the property and meaning of  $A$  is to be explored. Using the unitarity, we have

$$I = U^\dagger(\epsilon)U(\epsilon) = (I + \epsilon A^\dagger)(I + \epsilon A) = I + \epsilon(A^\dagger + A), \quad (106)$$

or

$$A^\dagger = -A, \quad (107)$$

which says  $A$  is anti-Hermitian.

But it is easy to construct an Hermitian operator, which is more desired by us, out of the anti-Hermitian operator just by multiplying the imaginary unit on it, ie,

$$\mathcal{H} \equiv iA = (-i)(-A) = (iA)^\dagger = \mathcal{H}^\dagger, \quad (108)$$

thus we have the form of the infinitesimal evolution operator as

$$U(\epsilon) = I + \epsilon A = I + \epsilon(-i)(iA) = 1 - i\epsilon\mathcal{H}, \quad (109)$$

and due to the Hermitian property of  $\mathcal{H}$ , it should be some kind of observable.

Now we're going to explore more on the evolution of states at an infinitesimal time interval  $\epsilon$ , we write

$$|\Psi(\epsilon)\rangle = \hat{U}(\epsilon)|\Psi(0)\rangle = |\Psi(0)\rangle - i\epsilon H|\Psi(0)\rangle, \quad (110)$$

and rearranging this we have

$$\frac{|\Psi(\epsilon)\rangle - |\Psi(0)\rangle}{\epsilon} = -iH|\Psi(0)\rangle, \quad (111)$$

where the LHS is nothing but the time derivative of the state  $|\Psi\rangle$  at time  $t = 0$ , but since the evolution is assumed to be linear, this should hold at any time  $t$ , so we may write

$$\frac{d|\Psi\rangle}{dt} = -iH|\Psi\rangle, \quad (112)$$

where we wrote partial or total derivative make no difference. This equation tells us how an isolated (not contact with apparatus) state evolves with time, and it is called the *generalized Schrodinger equation*. This is a simple first-order differential equation, and its solution is given by

$$|\Psi(t)\rangle = e^{-iHt}|\Psi(0)\rangle, \quad (113)$$

thus we may extract the explicit form of the evolution operator as  $U(t) = e^{-iHt}$ , and we check it by taking the time to be infinitesimal time interval  $\epsilon$ ,

$$U(t) = e^{-iH\epsilon} \approx 1 - iH\epsilon, \quad (114)$$

which is consistent with our previous reasoning.

As we proceed, we would discover that  $H$  is nothing but the Hamiltonian operator, thus it has the dimension of energy. By checking the dimensionality of the time-dependent Schrodinger equation, we may fix it by a factor of  $\hbar$ , the Planck constant,

$$i\hbar \frac{\partial |\Psi\rangle}{\partial t} = H|\Psi\rangle. \quad (115)$$

This is the standard form of the *time-dependent Schrodinger equation*.

Now that we have derived that  $H$  is a Hermitian operator, we can expand our states in the basis of  $H$ 's eigenvectors, to fully solve the Schrodinger equation,

$$|\Psi\rangle = \sum_j \alpha_j |E_j\rangle, \quad (116)$$

where  $|E_i\rangle$  satisfies

$$H|E_i\rangle = E_i|E_i\rangle, \quad (117)$$

and  $E_i$ s are the eigenvalues of  $H$ , or energy.

Using the time-dependent Schrodinger equation and due the fact that basis vector do not change with time, we have

$$\sum_j \dot{\alpha}_j(t) |E_j\rangle = -\frac{i}{\hbar} H \sum_j \alpha_j(t) |E_j\rangle = -\frac{i}{\hbar} \sum_j E_j \alpha_j(t) |E_j\rangle, \quad (118)$$

or regrouping

$$\sum_j \left\{ \dot{\alpha}_j(t) + \frac{i}{\hbar} E_j \alpha_j(t) \right\} |E_j\rangle = 0, \quad (119)$$

thus every coefficient must be zero, and we have again the simplest differential equation,

$$\frac{d\alpha_j(t)}{dt} = -\frac{i}{\hbar} E_j \alpha_j(t) \quad (120)$$

, and the solution is given by

$$\alpha_j(t) = \alpha_j(0) e^{-\frac{i}{\hbar} E_j t} \quad (121)$$

. This equation contains the underlying assumption that Hamiltonian does not depend explicitly on time. And it's the first time for us to see the deep connection between energy and time. We may extract out coefficients at time zero as

$$\alpha_j(0) = \langle E_j | \Psi(0) \rangle \quad (122)$$

, thus we can write the full solution for the time-dependent Schrodinger equation as

$$|\Psi(t)\rangle = \sum_j \alpha_j(0) e^{-\frac{i}{\hbar} E_j t} |E_j\rangle = \sum_j |E_j\rangle \langle E_j | \Psi(0) \rangle e^{-\frac{i}{\hbar} E_j t}, \quad (123)$$

which emphasizes that we're summing over the basis vectors, and implies that as long as we have the initial state  $|\Psi(0)\rangle$  and the Hamiltonian  $H$  which governs the physical law of evolution, we can have state  $|\Psi(t)\rangle$  at any time as we want.

The average, or expectation value, of an observable is the closest thing in quantum mechanics to a classical value. From a mathematical point of view, an average is defined by the equation

$$\langle \lambda \rangle = \sum_i \lambda_i P(\lambda_i). \quad (124)$$

If we expand a quantum state  $|\Psi\rangle$  in the orthonormal basis of eigenvectors of  $L$ ,

$$|\Psi\rangle = \sum_i \alpha_i |\lambda_i\rangle, \quad (125)$$

we try to compute the following quantity,

$$\langle L \rangle \equiv \langle \Psi | L | \Psi \rangle = \sum_i \sum_j \langle \lambda_j | \alpha_j^* \lambda_i \alpha_i | \lambda_i \rangle = \sum_i \lambda_i (\alpha_i^* \alpha_i) = \sum_i \lambda_i P(\lambda_i) = \langle \lambda \rangle, \quad (126)$$

that's why we use the notation  $\langle L \rangle$  to denote the average. Let's see how this changes with time,

$$\begin{aligned} \frac{d}{dt} \langle L \rangle &= \frac{d}{dt} \langle \Psi(t) | L | \Psi(t) \rangle = \langle \dot{\Psi}(t) | L | \Psi(t) \rangle + \langle \Psi(t) | L | \dot{\Psi}(t) \rangle \\ &= \frac{i}{\hbar} \langle \Psi(t) | (HL - LH) | \Psi(t) \rangle = \frac{i}{\hbar} \langle [H, L] \rangle, \end{aligned} \quad (127)$$

where in the second row we used the Schrodinger equation, and the notation for a commutator. Note that the  $i$  before the commutator is important, since the commutator itself is always imaginary. Equivalently we rewrite it in a shorthand form:

$$\frac{dL}{dt} = -\frac{i}{\hbar} [L, H], \quad (128)$$

where letters denote their average value. This is a very interesting and important equation. It relates the time derivative of the expectation value of an observable  $L$  to the expectation value of another observable  $[L, H]$ . If we assume that the probabilities are nice, narrow, bell-shaped curves, this equation tells us how the peaks of the curves move with time. Equations like this are the closest thing in quantum mechanics to the equations of classical physics.

Recall that in classical mechanics, the time derivative of a physical quantity is given by its Poisson Bracket with the Hamiltonian,

$$\frac{dL}{dt} = \{L, H\}. \quad (129)$$

Due to their similarity, we may identify a correspondence between the commutators and the Poisson Brackets,

$$[L, H] \iff i\hbar \{L, H\}. \quad (130)$$

Actually they have the same mathematical structures (like anti-symmetric, Bianchi identity) called the Lie Algebra, which would be explored later if we have time.

In classical physics, commutators between ordinary observables are zero. But if we want to apply (128) to classical physics, we must assume the commutators are not zero but a very small value of the order of  $\hbar$ . Vice versa, the classical limit is the limit at which  $\hbar$  is negligibly small, thus the  $i\hbar\{L, H\}$  goes back to the vanishing commutator.

Similar to classical mechanics, the equation tells us the expectation value of an observable  $\langle Q \rangle$  does not change with time if it commutes with the Hamiltonian,

$$\frac{d\bar{Q}}{dt} = -\frac{i}{\hbar} \overline{[Q, H]} = 0. \quad (131)$$

Moreover if  $Q$  commutes with the Hamiltonian, the expectation values of all powers of  $Q$ , and even all functions of  $Q$  are conserved.

Just like in classical mechanics, the most obvious conserved quantity is the Hamiltonian itself, since it commutes with itself,

$$\frac{d\bar{H}}{dt} = -\frac{i}{\hbar} \overline{[H, H]} = 0, \quad (132)$$

which says energy is conserved in quantum mechanics.

Now it's more concrete to set the spin as an example. The energy of a spin in a magnetic field is given by

$$H \sim \vec{\sigma} \cdot \vec{B} = \sigma_x B_x + \sigma_y B_y + \sigma_z B_z. \quad (133)$$

If the magnetic field lies along the z-axis, the Hamiltonian is proportional to  $\sigma_z$ , in the following expression we'll absorb all the irrelevant numerical constants into a single constant  $\omega$ , so

$$H = \frac{\hbar\omega}{2} \sigma_z, \quad (134)$$

where we kept  $\frac{\hbar}{2}$  for later convenience.

Now we want to find out how the expectation value of the spin varies with time, we can get

$$\begin{aligned} \langle \dot{\sigma}_x \rangle &= -\frac{i}{\hbar} \langle [\sigma_x, H] \rangle = -\frac{i\omega}{2} \langle [\sigma_x, \sigma_z] \rangle = -\omega \langle \sigma_y \rangle \\ \langle \dot{\sigma}_y \rangle &= -\frac{i}{\hbar} \langle [\sigma_y, H] \rangle = -\frac{i\omega}{2} \langle [\sigma_y, \sigma_z] \rangle = \omega \langle \sigma_x \rangle \\ \langle \dot{\sigma}_z \rangle &= -\frac{i}{\hbar} \langle [\sigma_z, H] \rangle = -\frac{i\omega}{2} \langle [\sigma_z, \sigma_z] \rangle = 0 \end{aligned} \quad (135)$$

This result implies that the 3-vector-operator  $\vec{\sigma}$  precesses like a gyroscope around the direction of the magnetic field. The precession is uniform, with angular velocity  $\omega$ . To be specific, the expectation value for a  $\sigma_z$  measurement does not change with time, but the other two expectation values do change.

All the above formalism in this subsection describes how state evolves with time during two measurements. But something different happens when an observation is made, during an experiment the state of a system jumps unpredictably to an eigenstate of the observable that was measured. This phenomenon is called *the collapse of the wave function*. This implies that to examine the measurement process itself as a quantum mechanical evolution, we must consider the entire experimental setup, including the apparatus, as part of a single quantum system.