

Notes on The Theoretical Minimum — Quantum Mechanics

Gui-Rong Liang

November 5, 2018

Contents

| | | |
|----------|---|----------|
| 1 | Mathematical Foundations and Spins | 2 |
| 1.1 | Quantum states, operators, and Hermitian properties | 2 |
| 1.2 | Time evolution of states and expectation values | 3 |
| 1.3 | Spin states, spin operators, and Pauli Matrices | 6 |

1 Mathematical Foundations and Spins

1.1 Quantum states, operators, and Hermitian properties

A quantum state is represented by a ket vector that can be decomposed using the orthogonal basis vectors,

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

where the coefficient α_j is a complex number. 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 λ 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.

1.2 Time evolution of 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,

$$|\Psi(t)\rangle = \hat{U}(t)|\Psi(0)\rangle. \quad (18)$$

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

Mathematically, the conservation of orthogonality tells us

$$\langle \Psi(0) | \Phi(0) \rangle = 0 \implies \langle \Psi(t) | \Phi(t) \rangle = \langle \Psi(0) | \hat{U}^\dagger(t) \hat{U}(t) | \Phi(0) \rangle = 0, \quad (19)$$

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,

$$\langle i | j \rangle = \delta_{ij} = \langle i | \hat{U}^\dagger(t) \hat{U}(t) | j \rangle \implies \hat{U}^\dagger(t) \hat{U}(t) = I, \quad (20)$$

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*,

$$\langle A(t) | B(t) \rangle = \langle A(0) | \hat{U}^\dagger(t) \hat{U}(t) | B(0) \rangle = \langle A(0) | B(0) \rangle, \quad (21)$$

which says the inner product of $|A\rangle$ and $|B\rangle$ does not change with time, or the inner product is conserved.

Time evolution is led by incremental changes, so it would be inspiring to study the evolution operator of an infinitesimal time interval ϵ . 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 ϵ is very small, $\hat{U}(\epsilon)$ is close to the unit operator, only differing from it by something of order ϵ , and when ϵ is taken to be zero, the state does not change.

$$U(\epsilon) = I - i\epsilon H, \quad (22)$$

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

$$U^\dagger(\epsilon) = I + i\epsilon H^\dagger \quad (23)$$

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

$$I = U^\dagger(\epsilon)U(\epsilon) = (I + i\epsilon H^\dagger)(I - i\epsilon H) = I + i\epsilon(H^\dagger - H), \quad (24)$$

or

$$H^\dagger = H. \quad (25)$$

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

$$U(\epsilon)|\Psi\rangle = |\Psi\rangle + \epsilon|\Theta\rangle, \quad (26)$$

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

$$U(\epsilon)|\Psi\rangle = |\Psi\rangle + \epsilon A|\Psi\rangle = (I + \epsilon A)|\Psi\rangle, \quad (27)$$

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

$$U(\epsilon) = I + \epsilon A, \quad (28)$$

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

$$I = U^\dagger(\epsilon)U(\epsilon) = (I + \epsilon A^\dagger)(I + \epsilon A) = I + \epsilon(A^\dagger + A), \quad (29)$$

or

$$A^\dagger = -A, \quad (30)$$

which says A is anti-Hermitian.

We argue that any operator can be make up of two Hermitian operator F and H , which serve as real and imaginary part individually, as $A = F - iH$, (just take F and H as they have the same eigenvectors as A , and eigenvalues as real and imaginary part of the eigenvalues of A), then it's easy to figure out that an anti-Hermitian operator has no real part, i.e, $A = -iH$, thus we have the form of the infinitesimal evolution operator as

$$U(\epsilon) = I - i\epsilon H, \quad (31)$$

and due to the Hermitian property of 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 ϵ , we write

$$|\Psi(\epsilon)\rangle = \hat{U}(\epsilon)|\Psi(0)\rangle = |\Psi(0)\rangle - i\epsilon|\Psi(0)\rangle, \quad (32)$$

and rearranging this we have

$$\frac{|\Psi(\epsilon)\rangle - |\Psi(0)\rangle}{\epsilon} = -iH|\Psi(0)\rangle, \quad (33)$$

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

$$\frac{\partial|\Psi\rangle}{\partial t} = -iH|\Psi\rangle, \quad (34)$$

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*.

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,

$$i\hbar \frac{\partial|\Psi\rangle}{\partial t} = H|\Psi\rangle. \quad (35)$$

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

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

$$\langle\lambda\rangle = \sum_i \lambda_i P(\lambda_i). \quad (36)$$

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

$$|\Psi\rangle = \sum_i \alpha_i |\lambda_i\rangle, \quad (37)$$

we try to compute the following quantity,

$$\langle L \rangle \equiv \langle \Psi | L | \Psi \rangle = \sum_i \sum_j \langle \lambda_j | \alpha_j^* \lambda_i \alpha_i | \lambda_i \rangle = \sum_i \lambda_i (\alpha_i^* \alpha_i) = \sum_i \lambda_i P(\lambda_i) = \langle \lambda \rangle, \quad (38)$$

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

$$\begin{aligned} \frac{d}{dt} \langle L \rangle &= \frac{d}{dt} \langle \Psi(t) | L | \Psi(t) \rangle = \langle \dot{\Psi}(t) | L | \Psi(t) \rangle + \langle \Psi(t) | L | \dot{\Psi}(t) \rangle \\ &= \frac{i}{\hbar} \langle \Psi(t) | (HL - LH) | \Psi(t) \rangle = \frac{i}{\hbar} \langle [H, L] \rangle, \end{aligned} \quad (39)$$

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 not Hermitian.

Equivalently we rewrite it in a shorthand form:

$$\frac{dL}{dt} = -\frac{i}{\hbar} [L, H], \quad (40)$$

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,

$$\frac{dL}{dt} = \{L, H\}. \quad (41)$$

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

$$[L, H] \iff i\hbar \{L, H\}. \quad (42)$$

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 (40) 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,

$$\frac{d\bar{Q}}{dt} = -\frac{i}{\hbar} \overline{[Q, H]} = 0. \quad (43)$$

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,

$$\frac{d\bar{H}}{dt} = -\frac{i}{\hbar} \overline{[H, H]} = 0, \quad (44)$$

which says energy is conserved in quantum mechanics.

1.3 Spin states, spin operators, and Pauli Matrices

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

$$|A\rangle = \alpha_u |u\rangle + \alpha_d |d\rangle, \quad (45)$$

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

$$\begin{cases} |r\rangle = \frac{1}{\sqrt{2}}|u\rangle + \frac{1}{\sqrt{2}}|d\rangle \\ |l\rangle = \frac{1}{\sqrt{2}}|u\rangle - \frac{1}{\sqrt{2}}|d\rangle. \end{cases} \quad (46)$$

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,

$$\begin{cases} |f\rangle = \frac{1}{\sqrt{2}}|u\rangle + \frac{i}{\sqrt{2}}|d\rangle \\ |b\rangle = \frac{1}{\sqrt{2}}|u\rangle - \frac{i}{\sqrt{2}}|d\rangle. \end{cases} \quad (47)$$

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

We will use σ to denote spin operator. Measuring along z-axis gives us

$$\begin{cases} \sigma_z |u\rangle = |u\rangle \\ \sigma_z |d\rangle = -|d\rangle, \end{cases} \quad (48)$$

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

$$\sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (49)$$

Measuring along x-axis gives us

$$\begin{cases} \sigma_x |r\rangle = |r\rangle \\ \sigma_x |l\rangle = -|l\rangle, \end{cases} \quad (50)$$

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

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \quad (51)$$

Measuring along y-axis gives us

$$\begin{cases} \sigma_y |f\rangle = |f\rangle \\ \sigma_y |b\rangle = -|b\rangle, \end{cases} \quad (52)$$

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

$$\sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}. \quad (53)$$

The matrix representations of σ_x , σ_y and σ_z are called the *Pauli Matrices*.

Now we may treat σ_x , σ_y and σ_z as three components of σ , thus σ 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.

$$\sigma_n = \vec{\sigma} \cdot \hat{n} = \sigma_x n_x + \sigma_y n_y + \sigma_z n_z = \begin{pmatrix} n_z & n_x - in_y \\ n_x + in_y & -n_z \end{pmatrix} \quad (54)$$

If \hat{n} lies x-z plane, we may write

$$\begin{cases} n_z = \cos \theta \\ n_x = \sin \theta \\ n_y = 0, \end{cases} \quad (55)$$

thus

$$\sigma_n = \begin{pmatrix} \cos \theta & \sin \theta \\ \sin \theta & -\cos \theta \end{pmatrix}, \quad (56)$$

its eigenvalues and corresponding eigenvectors are given by

$$\begin{cases} \lambda_1 = 1, & |\lambda_1\rangle = \begin{pmatrix} \cos \frac{\theta}{2} \\ \sin \frac{\theta}{2} \end{pmatrix} \\ \lambda_2 = -1, & |\lambda_2\rangle = \begin{pmatrix} -\sin \frac{\theta}{2} \\ \cos \frac{\theta}{2} \end{pmatrix}. \end{cases} \quad (57)$$

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

$$\begin{cases} P(+1) = \langle u | \lambda_1 \rangle \langle \lambda_1 | u \rangle = \cos^2 \frac{\theta}{2} \\ P(-1) = \langle u | \lambda_2 \rangle \langle \lambda_2 | u \rangle = \sin^2 \frac{\theta}{2} \end{cases} \quad (58)$$

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

$$\langle \sigma_n \rangle = \sum_j \lambda_j P(\lambda_j) = \cos^2 \frac{\theta}{2} - \sin^2 \frac{\theta}{2} = \cos \theta, \quad (59)$$

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

$$|A\rangle = \alpha_u|u\rangle + \alpha_d|d\rangle, \quad (60)$$

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

$$\vec{\sigma} \cdot \vec{n}|A\rangle = |A\rangle. \quad (61)$$

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

$$\langle\sigma_n\rangle = \langle\vec{\sigma} \cdot \vec{n}\rangle = 1. \quad (62)$$