Lecture 1: The superposition principle
Read Baym Chapter 1, Chapter 3; Sakurai Chapter 1

1 Why do we need to learn quantum mechanics (QM)?

Without quantum mechanics, we cannot understand the world.

1. Why can we stand up without collapsing?

2. Why do there exist metal/insulator/semi-conductor, magnetism, superconductivity/superfluidity?

3. What is the origin, past, future, and the fate of the earth, solar system, and even the universe?

2 Questions for further thinking

1. For a microscopic system, should it always be in a state described by a wavefunction?

2. What is the superposition principle?

3. How to normalize eigen-wavefunctions for an operator with continuous eigen-spectra?

4. How to use wavefunctions to represent probability of measurement?

5. What are the constraints for wavefunctions of identical particles?

6. Why to use linear operator to represent observable?

7. What is the precise meaning of quantization?

3 Fundamental principles and notions of QM

1. Correspondence principle

2. Space-time symmetry

3. Statistical interpretation of wavefunctions

4. Superposition principle

5. Fundamental quantization condition

6. Schödinger equation for the pure state and quantum Liouville equation for the mixed state

7. Bose and Fermi statistics

4 How to describe classic systems?

4.1 Classic particles

The state of a classical systems of N-particles can be represented by their coordinates and momenta as $(\vec{r}_1, ..., \vec{r}_N; \vec{p}_1, ..\vec{p}_N)$ which form a complete set of observables. Any other observable F can be expressed as a function of coordinates and momenta as $F(\vec{r}_1, .., \vec{r}_N; \vec{p}_1, ..., \vec{p}_N)$. This is called the "*purely mechanical description*". If at an initial time t_0, $(\vec{r}_1, ..., \vec{r}_N; \vec{p}_1, ..\vec{p}_N, t_0)$ are known, their values at any later time t are also determined by the Hamilton's canonical equation

$$\frac{d}{dt}q_i = \frac{\partial H}{\partial p_i}, \quad \frac{d}{dt}p_i = -\frac{\partial H}{\partial q_i}, \quad (i = 1, ..., n). \tag{1}$$

On the other hand, in many situations we are not able to determine all the values of coordinates and momenta precisely, and thus have to use the "*statistical description*". If coordinates and momenta are repeatedly measured many times, a positive-definite probability distribution function is obtained $\rho(\vec{r}_1, ..., \vec{r}_N; \vec{p}_1, ..., \vec{p}_N)$. The value of the observable F is expressed as

$$\bar{F} = \frac{\int d\vec{r}_1 ... d\vec{r}_N d\vec{p}_1 .. d\vec{p}_N \quad F(\vec{r}_1, .., \vec{r}_N; \vec{p}_1, ..., \vec{p}_N)\rho(\vec{r}_1, ..., \vec{r}_N; \vec{p}_1, ..., \vec{p}_N)}{\int d\vec{r}_1 ... d\vec{r}_N d\vec{p}_1 .. d\vec{p}_N \rho(\vec{r}_1, ..., \vec{r}_N; \vec{p}_1, ..., \vec{p}_N)}. \tag{2}$$

If the system is time-dependent, the distribution function is explicitly dependent on time as $\rho(\vec{r}_1, ..., \vec{r}_N; \vec{p}_1, ..., \vec{p}_N; t)$, and its time evolution follows the Liouville's equation

$$\frac{\partial \rho}{\partial t} + \sum_{i=1}^{N} \left\{ \frac{dq_i}{dt} \frac{\partial \rho}{\partial q_i} + \frac{dp_i}{dt} \frac{\partial \rho}{\partial p_i} \right\} = 0. \tag{3}$$

4.2 Classic waves

Consider an EM wave propagating along the z-direction. Its electric fields can be written as $E_x(r,t) = E_x^0 \cos(kz - \omega t + \alpha_x)$, $E_y(r,t) = E_y^0 \cos(kz - \omega t + \alpha_y)$, $\vec{B} = \hat{z} \times \vec{E}$. In principle, classic physics does *not* need complex number. It is more convenient to represent as

$$E_x(r,t) = E_x e^{i(kz - \omega t)}, \quad E_y(r,t) = E_y e^{i(kz - \omega t)}, \tag{4}$$

where $E_x = E_x^0 e^{i\alpha_x}$ and $E_y = E_y^0 e^{i\alpha_y}$. The polarization of the EM wave is defined as i) If $E_{y(x)} = 0$, then polarization of the EM wave is along $x(y)$-direction. ii) If $E_y = \pm i E_x$, then the EM wave is right(left) circularly polarized. Any general state of this EM wave can be decomposed as linear superposition of either the linear polarized bases i), or the circularly polarized basis.

The energy of the EM wave satisfies the following relation because

$$E_{total} = \frac{|E_x|^2 + |E_y|^2}{8\pi} V = N\hbar\omega, \tag{5}$$

where N is the number of photon. We define the state vector $|\Psi\rangle = (\psi_x, \psi_y)^T$ with $\psi_{x(y)} = \sqrt{\frac{V}{8\pi N\hbar\omega}} E_{x(y)}$, which satisfies the normalization condition $|\psi_x|^2 + |\psi_y|^2 = 1$. We can define a set of basis as

$$|x\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad |y\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \tag{6}$$

which are for polarizations along x and y-directions respectively. Or, we can define We can define the set of basis for circularly polarizations

$$|R\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \end{pmatrix}, \quad |L\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ -i \end{pmatrix}. \tag{7}$$

1. Definition of a row vector $\langle\Psi| = (\psi_x^*, \psi_y^*)$.

2. The scalar product between $|\Psi\rangle$ and $|\Phi\rangle = (\phi_x, \phi_y)$ is defined as $\langle\Psi|\Phi\rangle = \psi_x^*\phi_x + \psi_y^*\phi_y = \langle\Phi|\Psi\rangle^*$.

3. Normalization condition of $|\psi_x|^2 + |\psi_y|^2 = 1$ can be represented as $\langle\Psi|\Psi\rangle = 1$.

4. Any state can be decomposed into linear superpositions of a set of complete bases

$$|\Psi\rangle = \psi_x|x\rangle + \psi_y|y\rangle = \frac{\psi_x - i\psi_y}{\sqrt{2}}|R\rangle + \frac{\psi_x + i\psi_y}{\sqrt{2}}|L\rangle. \tag{8}$$

In other words, we have

$$\begin{aligned} |\Psi\rangle &= |x\rangle\langle x|\Psi\rangle + |y\rangle\langle y|\Psi\rangle = \left\{|x\rangle\langle x| + |y\rangle\langle y|\right\}|\Psi\rangle \\ &= |R\rangle\langle R|\Psi\rangle + |L\rangle\langle L|\Psi\rangle = \left\{|R\rangle\langle R| + |L\rangle\langle L|\right\}|\Psi\rangle. \end{aligned} \tag{9}$$

The function of a polaroid is a projection operator. For example, the x-polaroid and y-polaroid behave as

$$P_x = |x\rangle\langle x|; \quad P_y = |y\rangle\langle y|. \tag{10}$$

And the polaroids of R and L-circularly polarization behave as $P_R = |R\rangle\langle R|$ and $P_L = |L\rangle\langle L|$. For a general projection, which projects out a state $|\Phi\rangle$, its operator is represented as $P_\phi = |\Phi\rangle\langle\Phi|$. For a general state $|\Psi\rangle$, after it passes the polaroid P_ϕ, it becomes $P_\phi|\Psi\rangle = \langle\Phi|\Psi\rangle|\Phi\rangle$, where $\langle\Phi|\Psi\rangle$ is the amplitude passing the polaroid. The EM wave intensity after the projection relative to before is $|P_\phi|\Psi\rangle|^2 = |\langle\Phi|\Psi\rangle|^2$, or, the probability is $|\langle\Phi|\Psi\rangle|^2$.

3

Let us perform the rotation around the z-axis at θ which is denoted as $R(\theta)$. Under this rotation, the bases of $|x\rangle$ and $|y\rangle$ change to $|x'\rangle = R(\theta)|x\rangle$ and $|y'\rangle = R(\theta)|y\rangle$. The rotation operation $R(\theta)$ can be represented as an unitary matrix as

$$(|x'\rangle, |y'\rangle) = (|x\rangle, |y\rangle)R(\theta) \tag{11}$$

where the transformation matrix

$$R_{ij}(\theta) = \langle i|R(\theta)|j\rangle = \langle i|j'\rangle, \tag{12}$$

where i, j equals to x, y, and $j' = x', y'$. More explicitly

$$R(\theta) = \begin{pmatrix} \langle x|x'\rangle & \langle x|y'\rangle \\ \langle y|x'\rangle & \langle y|y'\rangle \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}. \tag{13}$$

Here R is an orthogonal matrix, generally speaking, it can be a unitary matrix, which can be exponentialized as

$$R(\theta) = e^{-\frac{i}{\hbar}S\theta}, \quad \text{with} \quad S = \hbar \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}. \tag{14}$$

For a general state $|\Psi\rangle$, its projections on the two sets of bases satisfy

$$|\Psi\rangle = (|x\rangle, |y\rangle) \begin{pmatrix} \langle x|\Psi\rangle \\ \langle y|\Psi\rangle \end{pmatrix} = (|x'\rangle, |y'\rangle)R^{-1}(\theta) \begin{pmatrix} \langle x|\Psi\rangle \\ \langle y|\Psi\rangle \end{pmatrix}, \tag{15}$$

and thus

$$\begin{pmatrix} \langle x'|\Psi\rangle \\ \langle y'|\Psi\rangle \end{pmatrix} = R^{-1}(\theta) \begin{pmatrix} \langle x|\Psi\rangle \\ \langle y|\Psi\rangle \end{pmatrix} = \begin{pmatrix} \langle x'|x\rangle & \langle x'|y\rangle \\ \langle y'|x\rangle & \langle y'|y\rangle \end{pmatrix} \begin{pmatrix} \langle x|\Psi\rangle \\ \langle y|\Psi\rangle \end{pmatrix}. \tag{16}$$

Define a state $|\Psi'\rangle$ whose coordinates $|x'\rangle$ and $|y'\rangle$ bases are the same as those of $|\Psi\rangle$ in the bases of $|x\rangle$ and $|y\rangle$, and thus

$$|\Psi'\rangle = |x'\rangle\langle x|\Psi\rangle + |y'\rangle\langle y|\Psi\rangle, \tag{17}$$

which shows that

$$\begin{pmatrix} \langle x|\Psi'\rangle \\ \langle y|\Psi'\rangle \end{pmatrix} = \begin{pmatrix} \langle x'|x\rangle & \langle x'|y\rangle \\ \langle y'|x\rangle & \langle y'|y\rangle \end{pmatrix} \begin{pmatrix} \langle x|\Psi\rangle \\ \langle y|\Psi\rangle \end{pmatrix} = R(\theta) \begin{pmatrix} \langle x|\Psi\rangle \\ \langle y|\Psi\rangle \end{pmatrix}. \tag{18}$$

In other words, $|\Psi'\rangle$ is the consequence of applying the rotation around z-axis at θ on the state of $|\Psi\rangle$.

The operator S defined in Eq. 19 is a Hermitian operator, which describe an observable: the spin angular momentum of EM wave along the z-direction. Its eigenvalues are $\pm\hbar$ with

eigenvecotrs $\Psi_{R,L}$

$$S\Psi_{R(L)} = \pm\hbar\Psi_{R(L)}, \quad \Psi_{R(L)} = \begin{pmatrix} 1 \\ \pm i \end{pmatrix}. \tag{19}$$

The right and left circularly polarized light carry spin (projection along the z-axis) $\pm\hbar$.

5   Description of quantum systems

The quantum mechanical systems are similar to the classic waves but with significant difference. For the moment, we only consider states represented by wavefunctions which obey the super-position principle. Such a principle states:

1) For a given system at a given time, the set of all the possible wavefunctions form a linear space on the complex-number field.

2) Consider a state represented by the wavefunction $a\Psi_A + b\Psi_B$ (a, b are complex constants) which is a superposition of wavefunctions Ψ_A and Ψ_B. The measurements of an observable F in the state $a\Psi_A + b\Psi_B$ only yield values from those from measurements in the states Ψ_A or Ψ_B.

3) If in the state Ψ, the measurement of F can yield the value of f, then Ψ can be decomposed as

$$\Psi = \Psi_A + \Psi_B. \tag{20}$$

In the state represented by Ψ_A, the measurement of F only yields the value of f; while such measurements in the state represented by Ψ_B never yield the value of f. Ψ_A is called the eigenstate of F with the eigenvalue f.

4) Under time evolution, the linear relation between wavefunctions does not change. For example, if at t_0 consider a state $\Psi(t_0) = a\Psi_A(t_0) + b\Psi_B(t_0)$, then at time t, we have $\Psi_{(}t) = a\Psi_A(t) + b\Psi_B(t)$.

Examples

1) For a spinless particle, we use $\psi_{\vec{r}'}(\vec{r})$ to represent the state of particle localized at $\vec{r}'$. In such a state, you can only find the particle location at $\vec{r}'$, and later on we can show that $\psi_{\vec{r}'}(\vec{r}) = \delta(\vec{r} - \vec{r}')$. Then in the state $\int_{r' \leq R} \psi_{\vec{r}'}(\vec{r})\psi(\vec{r}')d^3\vec{r}'$, you can never find the particle outside the ball $B_R(0)$. Inside this ball, you can find the particle at the neighborhood $d\vec{r}_1$ around the location of $\vec{r}_1$. The corresponding component is $\psi_{\vec{r}_1}(\vec{r})\psi(\vec{r}_1)d\vec{r}_1$.

2) Suppose a spinless particle is located at $\vec{r}_0$ at time t_0. This wavefunction is denoted as $\psi_{\vec{r}_0}(\vec{r}) = \delta(\vec{r} - \vec{r}_0)$. At a later time t, the amplitude of this particle located at $\vec{r}$ is denoted as

$K(t\vec{r};t_0\vec{r}_0)$, which is called the propagator. Then for an arbitrary wavefunction $\psi(\vec{r},t_0)$, how does it evolve with time?

Solution:

$$\Psi(\vec{r},t_0) = \int d\vec{r}_0 \psi_{\vec{r}_0}(\vec{r})\psi(\vec{r}_0,t_0), \tag{21}$$

in which $\psi_{r_0}(\vec{r})$ is the coordinate eigen-function, and $\psi(\vec{r}_0,t_0)$ is the amplitude. By the superposition law, at time t_1, we have

$$\Psi(\vec{r},t_1) = \int d\vec{r}_0 K(t_1\vec{r};t_0\vec{r}_0)\psi(\vec{r}_0,t_0) = \int d\vec{r}_1 \psi_{\vec{r}_1}(\vec{r}) \int d\vec{r}_0 K(t_1\vec{r}_1;t_0\vec{r}_0)\psi(\vec{r}_0,t_0). \tag{22}$$

At an even later time t_2, we have

$$\Psi(\vec{r},t_2) = \int d\vec{r}_1 K(t_2\vec{r};t_1\vec{r}_1) \int d\vec{r}_0 K(t_1\vec{r}_1;t_0\vec{r}_0)\psi(\vec{r}_0,t_0). \tag{23}$$

Thus the propagator K satisfies the following property

$$K(t_2\vec{r};t_0\vec{r}_0) = \int d\vec{r}_1 K(t_2\vec{r};t_1\vec{r}_1)K(t_1\vec{r}_1;t_0\vec{r}_0), \tag{24}$$

for any $t_2 > t_1 > t_0$.

6   The statistical interpretation of wavefunctions

Different from classical waves, the QM wavefuncitons has the statistical interpretation. For simplicity, we use the case of single particle for example. It can be straightforwardly generalized to many-particle case.

1) A wavefunction $\Psi(\vec{r})$ is an amplitude. $|\Psi(\vec{r})|^2 d\vec{r}$ represent the probability to find the system configuration in the neighborhood $d\vec{r}$ around $\vec{r}$.

2) In the state of Ψ, the probability of the results of measuring an observable F can be represented through bilinear function of Ψ and Ψ^*.

For two eigenstates Ψ_1 and Ψ_2 of F with two different eigenvalues f_1 and f_2, we superpose them for a new state $\Psi_{12} = a_1\psi_1 + a_2\psi_2$ in which we can only find either f_1 or f_2 for measuring F. First of all, Ψ_{12} cannot be an eigenstate of F. Otherwise, say, it has the eigenvalue of f_1, then we can make $\Psi_{12} - a_1\psi_1 = a_2\psi_2$. Since both Ψ and ψ_1 possess the eigenvalue f_1, ψ_2 should be also an eigenstate of F with the same eigenvalue. This contradicts to our initial assumption, thus the measurement in Ψ cannot give a definitive result of F. The question is what is the relative probability of finding f_1 and f_2 denoted by $W_{f_1}(F,\Psi_{12})$ and $W_{f_2}(F,\Psi_{12})$, respectively. As assumed above, they should be expressed by bilinear of

Ψ_{12} and Ψ_{12}^* as

$$
\begin{aligned}
W_{f1}(F,\Psi_{12}) &= A_{11}a_1^*a_1 + A_{12}a_1^*a_2 + A_{12}^*a_2^*a_1 + A_{22}a_2^*a_2, \\
W_{f2}(F,\Psi_{12}) &= B_{11}a_1^*a_1 + B_{12}a_1^*a_2 + B_{12}^*a_2^*a_1 + B_{22}a_2^*a_2.
\end{aligned} \tag{25}
$$

All the coefficients A's and B's are independent of $a_{1,2}$, but are determined by $\Psi_{1,2}$. If $a_1 = 0$, we should have $W_{f_1} = 0$ and thus $A_{22} = 0$. By a similar reasoning, $B_{11} = 0$. Because probability is non-negative, the cross-terms $a_1^*a_2$ and $a_2^*a_1$ vanish which proves $A_{12} = B_{12} = 0$. Also A_{11} and B_{11} should be positive numbers.

The sum of W_{f1} and W_{f2} is the total probability 1, which should be equal to the normalized integral of $|\Psi_{12}|^2$.

$$
\begin{aligned}
A_{11}a_1^*a_1 + B_{22}a_2^*a_2 &= a_1^*a_1 \int \psi_1^*(\vec{r})\psi_1(\vec{r})d\vec{r} + a_2^*a_2 \int \psi_2^*(\vec{r})\psi_2(\vec{r})d\vec{r} \\
&+ a_1^*a_2 \int \psi_1^*(\vec{r})\psi_2(\vec{r})d\vec{r} + a_2^*a_1 \int \psi_2^*(\vec{r})\psi_1(\vec{r})d\vec{r}.
\end{aligned} \tag{26}
$$

Since $a_{1,2}$ are arbitrary, we have

$$
\begin{aligned}
W_{f1}(F,\Psi_{12}) &= \int d\vec{r}\, [a_1\Psi_1(\vec{r})]^*[a_1\Psi_1(\vec{r})], \\
W_{f2}(F,\Psi_{12}) &= \int d\vec{r}\, [a_2\Psi_1(\vec{r})]^*[a_2\Psi_1(\vec{r})], \\
0 &= \int d\vec{r}\Psi_1^*(\vec{r})\Psi_2(\vec{r}).
\end{aligned} \tag{27}
$$

If we choose the normalization of $\Psi_{1,2}$ as $\int |\Psi_{1,2}|^2 dq = 1$, we have $W_{f1} = |a_1|^2$ and $W_{f2} = |a_2|^2$. We have also proved that *two eigenstates of an observable F with different eigenvalues are orthogonal*.

More generally, according to the measurement of an arbitrary observable F on an arbitrary wavefunction Ψ, Ψ can be expressed as the linear superposition of eigenstates F as

$$
\Psi = \sum_{f_0}\Psi_{f_0} + \int df' \Psi_{f'}, \tag{28}
$$

where f_0 represents the discrete eigenvalues of F, and f' represents the continuous eigenvalues. This equation means the completeness of eigenstates of a physical observable. We define the inner product

$$
(\Psi_A, \Psi_B) = \int d\vec{r}\Psi_A^*(\vec{r})\Psi_B(\vec{r}), \tag{29}
$$

then we have (Ψ_{f_0}, Ψ) represents the probability of finding f_0 in the measurement of F. If $f_0 \neq f_0'$, we have $(\Psi_{f_0}, \Psi_{f_0'}) = 0$. For the continuous spectra, $(\int_{\Delta f} df' \Psi_{f}', \Psi)$ represents the probability of finding f' to $f' + \Delta f$ in the measurement of F.

7

7 Expressing observables in terms of operators

Starting from Eq. 28, define an operator $P(f_0, F)\Psi = \Psi_{f_0}$ for discrete eigenvalues, and $P(f', F)\Psi df' = \psi_{f'} df'$ for continuous eigenvalues. $P(f_0, F)$ satisfies $[P(f_0, F)]^2 = P(f_0, F)$ and so does $P(f', F)$, and thus they are projection operators. Define operator $\hat{F}$ as

$$\hat{F} = \sum_{f_0} f_0 P(f_0, F) + \int df' f' P(f', F), \tag{30}$$

then

$$\hat{F}\Psi = \sum_{f_0} f_0 \Psi_{f_0} + \int f' \Psi_{f'} df', \tag{31}$$

The expectation value of F can be expressed as

$$\bar{F} = \frac{\sum_{f_0} f_0 (\psi_{f_0}, \psi_{f_0}) + \sum_{\Delta f} f' \left(\int_{\Delta f} df' \Psi_{f'}, \int_{\Delta f} df' \Psi_{f'} \right)}{(\Psi, \Psi)} = \frac{(\Psi, \hat{F}\Psi)}{(\Psi, \Psi)}. \tag{32}$$

8 Complete set of observables

A complete set of observables $\{F_1, F_2, ...\}$ is defined as follows:
1) All the observables are compatible, *i.e.*, they have common and independent eigenvalues.
2) A set of eigenvalues of this set of observables determine a pure state.

Again we will only use the single particle case for its simplicity. We denote x, y and z and $\vec{r} = (x, y, z)$ for the coordinate operators. The compatibility among x, y and z is assumed. Let us consider x and following Eq. 31 to make the decomposition of the wavefunction $\Psi(\vec{r})$ as

$$\begin{aligned} \Psi(\vec{r}) &= \int dx' \Psi_{x'}(\vec{r}), \\ \hat{x}\Psi &= \int dx' x' \Psi_{x'}(\vec{r}). \end{aligned} \tag{33}$$

where $\Psi_{x'}(\vec{r})$ is the eigenstate of $\hat{r}_x$. This shows that

$$\Psi_{x'}(\vec{r}) = \Psi(\vec{r})\delta(x' - x), \tag{34}$$

and the expression of $\hat{r}_{j,x}$ in the coordinate representation as

$$\hat{x}\Psi(\vec{r}) = x\Psi(\vec{r}). \tag{35}$$

Coordinates as a complete set of observables   We next prove x, y, z are complete set of observable. We define their common eigenstate as $\psi_{x',y',z'}(\vec{r})$, which satisfies

$$\hat{r}_a \psi_{x',y',z'}(\vec{r}) = r'_a \psi_{x',y',z'}(\vec{r}) = r_a \psi_{x',y',z'}(\vec{r}) \tag{36}$$

where $r_a = x, y$ and z, $r'_a = x', y'$ and z'. Then we conclude that $\psi_{x',y',z'}(\vec{r}) \propto \delta(x - x')\delta(y - y')\delta(z - z')$. The set of eigenvalues are $\{x, y, z\}$. It is clear that $\psi_{x',y',z'}(\vec{r})$ form a complete set for wavefunctions $\Psi(\vec{r})$.

Thus we reach the remarkable conclusion that *unlike the case in classical mechanics, coordinates themselves form a complete set of observable without momenta.*

The common eigenvalues of a complete set of observable span a complete basis for wavefunctions

1) Completeness: $\{F_\alpha\}$ with $\{\alpha\} = \{\alpha_1, \alpha_2, ...\}$ is a complete set of observable, and $\{\Psi_{\alpha_1, \alpha_2,....}\}$ represent all the independent common eigenstates.

2) Orthogonal conditions: $(\Psi_\alpha, \Psi_{\alpha'}) = 0$ if $\{\alpha\} \neq \{\alpha'\}$.

3) Normalization. For a set of discretized eigen-value of $\{\alpha\} = \{n\}$, for an arbitrary wavefunction, the probability to find $\{n\}$ is a finite value which is proportional to $(\Psi_{\{n\}}, \Psi_{\{n\}})$, thus it should be normalizable as $(\Psi_{\{n\}}, \Psi_{\{n\}}) = \delta_{\{n\},\{n'\}}$.

For a continuously distributed eigenvalues, for simplicity, we consider a single parameter for eigenvalue $\int_{\Delta\alpha} d\alpha \Psi_\alpha$. The the probability to observe F_α's eigenvalue lying in the interval $\alpha \to \alpha + \Delta\alpha$ should be finite such that $(\int_{\Delta\alpha} d\alpha \Psi_\alpha, \int_{\Delta\alpha} d\alpha \Psi_\alpha)$ is finite. Because $(\psi_{\alpha_1}, \psi_{\alpha_2}) \neq 0$ only at $\alpha_1 = \alpha_2$, $(\Psi_{\alpha_1, \alpha_2})$ has to be proportional to $\delta(\alpha_1 - \alpha_2)$ in order for $(\int_{\Delta\alpha} d\alpha \Psi_\alpha, \int_{\Delta\alpha} d\alpha \Psi_\alpha)$ to have a finite value. Thus for eigenstates with continuous spectra, we need to normalize them using δ-function as

$$(\Psi_\alpha, \Psi_{\alpha'}) = \delta(\alpha - \alpha'). \tag{37}$$

After all the eigenstates are normalized, we can write down the decomposition as

$$\Psi = \sum_{\{n\}} a_{\{n\}} \Psi_{\{n\}} + \int \{d\alpha\} a_{\{\alpha\}} \Psi_{\{\alpha\}}, \tag{38}$$

where $\{m\}$ and $\{n\}$ are the the discrete eigenvalues and $\{\alpha\}$ represents the continuous eigenvalues; $a_{\{n\}} = (\Psi_{\{n\}}, \Psi)$, and $a_{\{m,\alpha\}} = (\Psi_{\{m,\alpha\}}, \Psi)$.

4) Completeness condition: For simplicity, we consider the case of one particle in one dimension. $\psi_{x'}(x)$ is the eigenstate of x with the eigenvalue x'. Let us decompose it in terms of the common eigenstates of a complete set of observable $\{F\}$ as

$$\Psi_{x'}(x) = \sum_n a_n(x') \Psi_n(x) \tag{39}$$

with

$$a_n(x') \quad = \quad (\Psi_n, \Psi_{x'}(x)) = \int \Psi_n^*(x)\delta(x-x')dx = \Psi_n^*(x'). \tag{40}$$

For simplicity, we neglect the continous spectra, which can be added back easily. Thus we have the completeness condition as

$$\sum_n \Psi_n(x)\Psi_n^*(x') = \delta(x-x'). \tag{41}$$

9   Generalization of the concept of wavefunction

Consider the decomposition of wavefunction in terms of the complete set of $F_{\{\alpha\}}$, and for simplicity, we only consider the case that all eigenvalues are discrete.

$$\Psi = \sum_{\{\alpha\}} a_{\{\alpha\}}\Psi_{\{\alpha\}}, \tag{42}$$

where $a_n = (\Psi_n, \Psi)$. a_n is the wavefunction of state Ψ in the representation of α, an $|a_n|^2$ is the probability of finding the eigenstate Ψ_n in the state of Ψ.

Check the expression in the coordinate representation

$$\Psi(\vec{r}) = \int d\vec{r}\Psi(\vec{r})\Psi_{x',y',z'}(\vec{r}), \tag{43}$$

in which $\Psi(\vec{r})$ is the wavefunction in the coordinate representation.