Quantum Mechanics

Lecture Notes

2026 Autumn

Takumi Kuwahara

September 28, 2026

Contents

Cautions

This notes set continues from the companion PDF for Sections 1–2. This notes is based on *Lectures on Quantum Mechanics* by Steven Weinberg. This note is converted from the author's handwritten note to the LaTeX code by the use of AI assistance, therefore there may be errors or inconsistencies. Once you found the errors or inconsistencies, please notify the author.

Notations

The Heaviside-Lorentz unit is used as the convention for the electromagnetic units.

3 General Principles of Quantum Mechanics

3.1 Hilbert space and States

In this subsection, we discuss the Hilbert space and the representation of physical states as vectors within it. One of the postulates in quantum mechanics is that *every physical state can be represented as a vector in a Hilbert space.* A Hilbert space is a kind of a normed complex vector space. A (complex) vector space V is a set of vectors $\Psi, \Psi', \Psi'', \ldots$ with the following properties:

(i). If Ψ, Ψ' and $\Psi'' \in V$, and then, so is $\Psi + \Psi' \in V$.

(ii). Addition operation exists and is associative and commutative.

$$\begin{cases} \Psi + (\Psi' + \Psi'') = (\Psi + \Psi') + \Psi'' \,, \\ \Psi + \Psi' = \Psi' + \Psi \,. \end{cases} \tag{3.1}$$

(iii). Scalar product is defined: when Ψ is vector, so is $\alpha\Psi\,, (\alpha \in \mathbb{C})$ (in the case of real vector space, $\alpha \in \mathbb{R}$). The scalar multiplication is associative and distributive.

$$\begin{cases} \alpha(\alpha'\Psi) = (\alpha\alpha')\Psi \,, \\ \alpha(\Psi + \Psi') = \alpha\Psi + \alpha\Psi' \,, \\ (\alpha + \alpha')\Psi = \alpha\Psi + \alpha'\Psi \,. \end{cases} \tag{3.2}$$

(iv). There is a single zero vector (denoted as $\mathbf{0}$), and

$$\begin{cases} \mathbf{0} + \Psi = \Psi \,. \\ 0 \cdot \Psi = \mathbf{0} \,. \\ \alpha \cdot \mathbf{0} = \mathbf{0} \,. \end{cases} \tag{3.3}$$

In the following, the zero vector is just written as 0 unless it is necessary to distinguish it from the scalar zero.

A normed vector space is a vector space equipped with a norm, which is a map that assigns a non-negative length or size to each vector in the space. To define a norm, we introduce a scalar product (or inner product) $(\cdot, \cdot)$, which is a map that assigns a complex number to each pair of vectors in the space. The inner product must satisfy the following properties:

(i). Linearity: for any vectors Ψ, Ψ', Ψ'' and any scalars α, α', we have

$$(\Psi'', [\alpha\Psi + \alpha'\Psi']) = \alpha(\Psi'', \Psi) + \alpha'(\Psi'', \Psi') \,. \tag{3.4}$$

(ii). Symmetry: associated with the complex conjugate of the inner product.

$$(\Psi', \Psi)^* = (\Psi, \Psi') \,. \tag{3.5}$$

(iii). Positivity: for any vector Ψ, a norm is defined and is positive

$$(\Psi, \Psi) \geq 0\,. \tag{3.6}$$

Equality holds if and only if $\Psi = 0$.

The Dirac's bra-ket notation is widely used in quantum mechanics: for that purpose, we introduce the concept of bra vectors and ket vectors. Ket vectors are defined as vectors in the Hilbert space and are denoted by $|\Psi\rangle$. Bra vectors are defined as the dual vectors, which is linear functional mapping ket vectors in the Hilbert space to complex numbers via the inner product, and are denoted by $\langle\Psi|$. The inner product between two vectors Ψ and Φ can then be written in the bra-ket notation as $\langle\Psi|\Phi\rangle$. In the following, we do not introduce the dual vector space, and the inner product notation is used.

A vector space has a characteristic called *dimensionality*, which is the number of vectors in a basis for the space. The dimensionality is always a non-negative integer, but it does not need to be finite. Indeed, the dimensionality can be infinite in some vector spaces, such as Hilbert spaces used in quantum mechanics. A set of vectors $\Psi_1, \cdots, \Psi_n$ is independent if there is no linear combination of them vanishing, except the trivial combination where all coefficients are zero.

$$\sum_{k=1}^{n} a_k \Psi_k \neq 0\,, \tag{3.7}$$

for any choice of coefficients a_k not all zero. Here, n is the dimensionality of the vector space. For independent vectors $\Psi_1\,, \cdots, \Psi_n$, we can have

$$\sum_{k} a_k \Psi_k = 0\,, \tag{3.8}$$

if and only if all coefficients a_k are zero.

We say that a set of vectors is *orthogonal* when the inner product of each pair of distinct vectors in the set is zero:

$$(\Psi_i, \Psi_j) = 0\,, \tag{3.9}$$

for any $i \neq j$. We can prove that such a set of orthogonal vectors are independent. If the orthogonal set (not need to be independent) satisfies $\sum_i \alpha_i \Psi_i = 0$, by taking the scalar product with Ψ_j, we find

$$\left(\Psi_j, \sum_i \alpha_i \Psi_i\right) = \sum_i \alpha_i (\Psi_j, \Psi_i) = \alpha_j (\Psi_j, \Psi_j) = 0\,, \tag{3.10}$$

so $\alpha_j = 0$ for all j. Meanwhile, the converse does not always hold, namely the vectors of an independent set do not have to be orthogonal. But we can always find independent and orthogonal sets for finite-dimensional vector spaces.

A set of vectors is said to be *complete* if it spans the entire vector space. In other words, any vector Ψ can be expressed as a linear combination of the vectors in the set.

$$\Psi = \alpha_1 \Psi_1 + \cdots + \alpha_n \Psi_n \,. \tag{3.11}$$

The set of vectors is not needed to be independent, but we can always find subset that is both complete and independent. Such a subset is obtained by deleting vectors from the original set until only independent vectors remain. We can call such a subset a *basis* of the vector space. A complete set of orthogonal vectors is also called an *orthogonal basis* for the vector space. Furthermore, the orthogonal basis can be normalized to form an *orthonormal basis*, where each vector has unit norm.

A vector space is said to be *finite-dimensional* if the largest possible number of independent vectors is d. In such a case, any set of d-independent vectors Φ_i is complete because any vector Ψ in the space can be expressed as a linear combination of these d vectors:

$$\Psi = \sum_{i=1}^{d} \alpha_i \Phi_i \,. \tag{3.12}$$

If there is a vector Ψ which is not written by $\sum_{i=1}^{d} \alpha_i \Phi_i$; then Ψ and $\{\Phi_i\}$ form $d+1$ independent vectors, which contradicts the assumption that the maximum number of independent vectors is d. Also, no set of fewer than d vectors $X_i \, (i = 1, \cdots, d-1)$ could be complete. If there exists, each vector Φ_i of d independent vectors could be written as linear combination of the $d - s$ vectors (s is positive integer) X_i:

$$\Phi_i = \sum_{j=1}^{d-s} c_{ij} X_j \,. \tag{3.13}$$

If Φ_i are independent, then any linear combination should be non-zero unless all coefficients are zero.

$$\sum_{k=1}^{d} u_k \Phi_k = \sum_{k\,,j} u_k c_{kj} X_j \neq 0 \,. \tag{3.14}$$

Here, u_k are arbitrary non-zero coefficients. But, we can always take $\sum_{k=1}^{d} u_k c_{kj} = 0$ for all j, leading to a contradiction. It is natural that the d independent vectors cannot be expressed as a linear combination of fewer than d vectors. This is also ensured by the theorem in linear algebra. $(c^T)_{ji}$ is considered as a map from V^d to V^{d-s}, and the existence of a non-zero vector u_k such that $\sum_{k=1}^{d} u_k c_{kj} = 0$ for all j is guaranteed by the rank-nullity theorem:

$$\dim(\ker(c^T)) + \mathrm{rank}(c^T) = d \,. \tag{3.15}$$

Indeed, non-zero $\dim(\ker(c^T))$, namely $\ker(c^T)$ being non-trivial, ensures the existence of a non-zero vector u_k such that $\sum_{k=1}^{d} u_k c_{kj} = 0$ for all j.

Definition of Hilbert space

We have prepared several concepts of normed vector spaces. A Hilbert space is a special type of normed vector space that is complete with respect to the norm induced by an inner product. In the Quantum Mechanics context, Hilbert spaces serve as the mathematical framework for the state space of quantum systems. The dimensionality of the Hilbert spaces is up to the physical system: it can be either finite or infinite. In particular, for infinite-dimension Hilbert spaces, the set is called to be complete in the sense that any vector can be expressed as a linear combination of the infinite numbers of basis vectors.

$$\Psi = \sum_{i=1}^{\infty} \alpha_i \Phi_i \,. \tag{3.16}$$

In other words, there is a sequence of partial sums that converges to any vector Ψ in the Hilbert space: or we can equivalently define the vector $\Omega_N \equiv \Psi - \sum_{i=1}^{N} \alpha_i \Phi_i$ whose norm goes to zero as N approaches infinity, $(\Omega_N, \Omega_N) \to 0$ for $N \to \infty$.

The expansion coefficients are unique for complete orthogonal sets of basis vectors. When a vector Ψ is expanded in terms of a complete orthogonal set of basis vectors $\{\Phi_i\}$, such as Eq. (3.16), the expansion coefficients are given by

$$\alpha_i = \frac{(\Phi_i, \Psi)}{(\Phi_i, \Phi_i)} \,. \tag{3.17}$$

Therefore, any vector in the Hilbert space can be expressed as a linear combination of the complete orthogonal set of basis vectors.

$$\Psi = \sum_{i=1}^{\infty} \frac{(\Phi_i, \Psi)}{(\Phi_i, \Phi_i)} \Phi_i \,. \tag{3.18}$$

By using orthogonality of the basis vectors, we can express the inner product of any two vectors Ψ and Ψ' as

$$(\Psi, \Psi') = \sum_{i} \frac{(\Phi_i, \Psi)^* (\Phi_i, \Psi')}{(\Phi_i, \Phi_i)} \,. \tag{3.19}$$

Probability Interpretation

We can interpret the inner (scalar) products in the Hilbert space as probabilities for measurement outcomes. One can define the transition rate of a state Ψ to a state Φ_i (or measurement outcome corresponding to Φ_i from a state Ψ) as

$$P(\Psi \mapsto \Phi_i) \equiv \frac{|(\Phi_i, \Psi)|^2}{(\Phi_i, \Phi_i)(\Psi, \Psi)} \,. \tag{3.20}$$

From the expression of inner product of a vector with itself, we have

$$(\Psi, \Psi) = \sum_{i} \frac{|(\Phi_i, \Psi)|^2}{(\Phi_i, \Phi_i)} \,. \tag{3.21}$$

Therefore, summing all final states gives

$$\sum_i P(\Psi \to \Phi_i) = \sum_i \frac{|(\Phi_i, \Psi)|^2}{(\Phi_i, \Phi_i)(\Psi, \Psi)} = 1\,. \tag{3.22}$$

This shows that the total sum of all possible measurement outcomes is 1, which allows the probabilistic interpretation of quantum mechanics.

We call $P(\Psi \to \Phi_i)$ the probability of transition from the state Ψ to the state Φ_i. This probability is unchanged under the scaling transformation of the states such that

$$\Psi \to \alpha\Psi, \quad \Phi_i \to \beta_i \Phi_i\,, \tag{3.23}$$

where α and β_i are non-zero complex numbers. Therefore, we can always take any vectors to be normalized such that

$$(\Phi_i, \Phi_i) = 1\,, \qquad (\Psi, \Psi) = 1\,. \tag{3.24}$$

In this orthonormal basis, we can have $P(\Psi \mapsto \Phi_i) = |(\Phi_i, \Psi)|^2$.

There are still phase factors that do not change the probabilities.

$$\Psi \to e^{i\psi}\Psi\,, \quad \Phi_i \to e^{i\varphi_i}\Phi_i \tag{3.25}$$

Two vectors that differ only by a phase factor represent the same physical state: we identify these vectors and we call the equivalence class as *rays*. Indeed, the physical states are represented by rays in the Hilbert space rather than individual vectors.

3.2 Continuum States

Physical states are characterized not only by discrete labels $i = 1, 2, \ldots$ but also by continuum labels (such as position) ξ. Due to continuity of labels, we can define the density of labels $\rho(\xi)$ such that the number of labels in the interval ξ to $\xi + d\xi$ is given by $\rho(\xi)d\xi$. The previous discussions on the discrete labels can be generalized to the case of continuum labels, by taking $\rho(\xi)d\xi$ the approximate number of discrete labels in the interval ξ to $\xi + d\xi$. In other words, the sum over discrete labels i is replaced by an integral over the continuum labels ξ weighted by the density of labels $\rho(\xi)$:

$$\sum_i f_i \to \int f(\xi) \cdot \rho(\xi) d\xi\,. \tag{3.26}$$

The inner product of states with respect to the continuum labels is defined by

$$(\Phi_\xi, \Phi_{\xi'}) = \rho(\xi)\delta_{\xi,\xi'}\,. \tag{3.27}$$

Here, $\delta_{\xi,\xi'}$ is considered as the selection of labels, and $\rho(\xi)$ accounts for the density of labels.

Therefore, any vectors in the Hilbert space can be expanded in terms of the continuum basis $\{\Phi_\xi\}$ as follows:

$$\begin{aligned}\Psi &= \sum_i \frac{(\Phi_i, \Psi)}{(\Phi_i, \Phi_i)}\Phi_i \\ &\to \int \rho(\xi)d\xi \cdot \frac{(\Phi_\xi, \Psi)}{\rho(\xi)}\Phi_\xi = \int (\Phi_\xi, \Psi)\Phi_\xi d\xi \,. \end{aligned} \tag{3.28}$$

The inner product of two vectors Ψ and Ψ' is written in terms of the continuum basis as follows:

$$\begin{aligned}(\Psi, \Psi') &= \sum_i \frac{(\Phi_i, \Psi^*)^*(\Phi_i, \Psi')}{(\Phi_i, \Phi_i)} \\ &\to \int \rho(\xi)d\xi \cdot \frac{(\Phi_\xi, \Psi)^*(\Phi_\xi, \Psi')}{\rho(\xi)} = \int (\Phi_\xi, \Psi)^*(\Phi_\xi, \Psi')d\xi \,. \end{aligned} \tag{3.29}$$

Thus, the normalization condition for the state Ψ, $(\Psi, \Psi) = 1$, is written as

$$\int |(\Phi_\xi, \Psi)|^2 d\xi = 1 \,. \tag{3.30}$$

Compared to the probability interpretation of the measurement outcomes with the discrete basis, the differential probability for the continuum basis is given by

$$dP(\Psi \to \Phi_\xi) = \frac{|(\Phi_\xi, \Psi)|^2}{(\Phi_\xi, \Phi_\xi)} \times \rho(\xi)d\xi = |(\Phi_\xi, \Psi)|^2 \, d\xi \,. \tag{3.31}$$

From the normalization condition, it apparently follows that $\int dP\,(\Psi \to \Phi_\xi) = 1$.

Dirac Delta function

The next question is what is $\rho(\xi)\delta_{\xi',\xi}$ as a "function" of ξ. From the orthonormality condition of the continuum basis, we find the integration property of $\rho(\xi)\delta_{\xi',\xi}$:

$$\int d\xi' \rho(\xi')\delta_{\xi',\xi} = \int d\xi' (\Phi_{\xi'}, \Phi_\xi) = 1 \,, \tag{3.32}$$

as far as ξ is within the domain of integration. Furthermore, analogous to the discrete label, $\sum_{i'} f_{i'}\delta_{ii'} = f_i$, the similar property for the continuum label is imposed,

$$\int d\xi' \rho(\xi')\delta_{\xi',\xi} \cdot f(\xi') = f(\xi) \,. \tag{3.33}$$

Here, ξ is assumed to be within the domain of integration, again. This "function" picks out the special value at ξ. In particular, this function is called as Dirac's delta function and is denoted by $\delta(\xi - \xi')$.

$$\int d\xi' \delta(\xi - \xi') = 1 \,, \qquad \int d\xi' f(\xi')\delta(\xi - \xi') = f(\xi) \,. \tag{3.34}$$

Indeed, the Dirac delta function is not a conventional function in the usual sense, but rather a distribution or generalized function. There are several ways to describe the Dirac delta function as a limit of ordinary functions. The followings are examples of limit expressions.

$$f(x) = \lim_{n\to\infty} \begin{cases} n & \left(-\frac{1}{2n} \le x \le \frac{1}{2n}\right) \\ 0 & (\text{others}) \end{cases} , \tag{3.35}$$

$$f(x) = \lim_{n\to\infty} \sqrt{\frac{n}{\pi}} e^{-nx^2} , \tag{3.36}$$

$$f(x) = \lim_{\epsilon\to 0} \frac{\epsilon}{\pi\left(x^2+\epsilon^2\right)} . \tag{3.37}$$

3.3 Observables

As stated in the first subsection, the physical states are described by the vectors in a Hilbert space. The other important concepts are the observables, which correspond to measurable physical quantities in a quantum system. In quantum mechanics, the observables are represented by *linear Hermitian operators* acting on the Hilbert space of physical states. Let $A\,, B\,, \cdots$ denotes the linear operators and $\Psi, \Phi, \cdots$ denotes the physical states in the Hilbert space. There are several properties of operators acting on the physical states.

(i). The product of the operators is defined as the operator acting on the physical states from right in order, i.e., $(AB)\Psi \equiv A(B\Psi)$.

(ii). The constant scalars of operators are defined as $(\alpha A)\Psi \equiv \alpha(A\Psi)$.

(iii). The addition of operators is defined as $(A+B)\Psi \equiv A\Psi + B\Psi$.

(iv). There is a zero operator $\hat{\mathbb{O}}$ defined as $\hat{\mathbb{O}}\Psi \equiv \mathbf{0}$ for arbitrary Ψ.

Again, in the following, we write the zero operator and the zero vector as 0 for simplicity unless it is necessary to distinguish from the scalar zero. For arbitrary operators and constants, the zero operator satisfies

$$\hat{\mathbb{O}}A = \hat{\mathbb{O}}\,, \quad \hat{\mathbb{O}} + A = A\,, \quad \alpha\hat{\mathbb{O}} = \hat{\mathbb{O}}\alpha = \hat{\mathbb{O}}\,. \tag{3.38}$$

(v). There is a unit operator $\hat{\mathbb{1}}$ defined as $\hat{\mathbb{1}}\Psi \equiv \Psi$ for arbitrary states.

For arbitrary operators, the unit operator satisfies

$$\hat{\mathbb{I}}A = A\hat{\mathbb{1}} = A\,. \tag{3.39}$$

A *linear* operator A acts on states, and satisfies the following linearity conditions:

$$A(\Psi+\Psi') = A\Psi + A\Psi'\,, \quad A(\alpha\Psi) = \alpha A\Psi\,. \tag{3.40}$$

The *adjoint* $A^\dagger$ of any operator A is defined as the operator satisfying the relation

$$(\Psi', A^\dagger \Phi) \equiv (A\Psi', \Psi)\,. \tag{3.41}$$

or equivalently, by using the symmetry of the complex conjugate of the inner product,

$$(\Psi', A^\dagger \Psi) = (\Psi, A\Psi')^*\,. \tag{3.42}$$

It is easy to show that adjoint has the following general properties:

$$(AB)^\dagger = B^\dagger A^\dagger\,, \quad \left(A^\dagger\right)^\dagger = A\,, \quad (\alpha A)^\dagger = \alpha^* A^\dagger\,, \quad (A+B)^\dagger = A^\dagger + B^\dagger\,. \tag{3.43}$$

The adjoint of an operator is also called the Hermitian conjugate. When the Hermitian conjugate of an operator is equal to the operator itself, i.e.,

$$A^\dagger = A\,, \tag{3.44}$$

the operator is called *Hermitian* or *self-adjoint*. The zero operator $\hat{\mathbb{0}}$ and the unit operator $\hat{\mathbb{1}}$ are both Hermitian: $\hat{\mathbb{0}}^\dagger = \hat{\mathbb{0}}$ and $\hat{\mathbb{1}}^\dagger = \hat{\mathbb{1}}$.

Once we introduce a complete orthonormal set $\{\Phi_i\}$, we can define the matrix representation of any linear operator A with respect to this basis.

$$A_{ij} \equiv (\Phi_i, A\Phi_j) \tag{3.45}$$

Since any vectors in Hilbert space are expanded in terms of the complete orthonormal set $\{\Phi_i\}$, the matrix elements of the product of two operators A and B are given by

$$\begin{aligned}(AB)_{ij} &= (\Phi_i, AB\Phi_j) \\ &= \sum_k (\Phi_i, A\Phi_k)(\Phi_k, B\Phi_j) = \sum_k A_{ik}B_{kj}\,.\end{aligned} \tag{3.46}$$

In the intermediate step, we use the expansion of $B\Phi_j = \sum_i (\Phi_i, B\Phi_j)\Phi_i$ in terms of the complete set. We often refer this procedure as *inserting a complete set of states* (in between the operators A and B in this case). In the sense of inserting a complete set, we remind the reader that the inner product can be written as a sum over the complete set:

$$(\Psi, \Psi') = \sum_k (\Psi, \Phi_k)(\Phi_k, \Psi')\,, \tag{3.47}$$

and that the operator matrix elements satisfy $(\Phi_i, AB\Phi_j) = (A^\dagger\Phi_i, B\Phi_j)$. As you can see, the matrix element of the product of the operators has the form of a product of the matrices. The adjoint of an operator in the matrix representation is given by

$$(A^\dagger)_{ij} = A^*_{ji}\,. \tag{3.48}$$

As shown in the previous subsection, we frequently use the continuum label for the states. Let us consider the continuum labels ξ, η, and the corresponding states Φ_ξ and Φ_η, which satisfy the orthonormal condition in the continuum sense:

$$(\Phi_\xi, \Phi_\eta) = \delta(\xi - \eta)\,. \tag{3.49}$$

and we can define the matrix elements of operators in the continuum representation as follows:

$$A_{\xi\eta} \equiv (\Phi_\xi, A\Phi_\eta)\,, \tag{3.50}$$

$$(AB)_{\xi\eta} = \int d\lambda\, A_{\xi\lambda} B_{\lambda\eta}\,. \tag{3.51}$$

We have prepared to discuss the second postulate in quantum mechanics: the observables are represented by linear Hermitian operators acting on the Hilbert space of physical states. A state has a definite value a for an observable, represented by a linear Hermitian operator A, if and only if the state vector Ψ is an *eigenstate* of A.

$$A\Psi = a\Psi\,. \tag{3.52}$$

Here, a is called an *eigenvalue* of A If we also have a different eigenstate Ψ' and its eigenvalue a', namely $A\Psi' = a'\Psi'$, we have

$$\begin{aligned} a(\Psi', \Psi) &= (\Psi', A\Psi) \\ &= (A\Psi', \Psi) = a'^*(\Psi', \Psi)\,. \end{aligned} \tag{3.53}$$

Here, we use the Hermiticity of the operator A. We find the important consequences of eigenvalues and eigenstates of Hermitian operators.

(i). The allowed value of observables are real: $a^* = a$ in the case that $\Psi = \Psi' \neq 0$ and $a' = a$.

(ii). The eigenstates with the different values for any observable are orthogonal: $(\Psi', \Psi) = 0$ for $a' \neq a$

Expectation Values

Let $\{\Psi_r\}$ be a complete set (given by eigenstates for a Hermitian operator A).

$$A\Psi_r = a_r\Psi_r \tag{3.54}$$

The expectation value in a state represented by a normalized vector Ψ is given by

$$\langle A\rangle_\Psi \equiv (\Psi, A\Psi) = \sum_r (\Psi, A\Psi_r)(\Psi_r, \Psi) = \sum_r a_r |(\Psi_r, \Psi)|^2\,. \tag{3.55}$$

In the intermediate step, we insert the completeness relation to expand the state Ψ in terms of the complete set of eigenstates $\{\Psi_r\}$. For any normalized state Ψ, the expectation value of the observable A can be interpreted as the weighted average of its eigenvalues, with the weights given by the probabilities $|(\Psi_r, \Psi)|^2$.

It is easy to see that the n-th power of operator A has the eigenvalues of a_r^n corresponding to the same eigenstates Ψ_r. More generally, we can define functions $f(A)$ of the Hermitian operator A, which act on the eigenstates Ψ_r as

$$f(A)\Psi_r = f(a_r)\Psi_r\,. \tag{3.56}$$

The function of the operator $f(A)$ has the eigenvalues $f(a_r)$ corresponding to the same eigenstates Ψ_r. In other words, the expectation value of a function of the operator $f(A)$ in the eigenstates Ψ_r is simply the function of the corresponding eigenvalue:

$$\langle f(A)\rangle_{\Psi_r} = f(a_r)\,. \tag{3.57}$$

Meanwhile, the expectation value of a function of the operator $f(A)$ in a general state Ψ is not necessarily equal to the function of the expectation value of A in the same state. It is apparent to see

$$\langle f(A)\rangle_{\Psi} = \sum_r f(a_r)\,|(\Psi_r,\Psi)|^2 \neq f\left(\langle A\rangle_{\Psi}\right)\,. \tag{3.58}$$

In particular, for the square of any Hermitian operator A, we have the expectation value

$$\langle A^2\rangle_{\Psi} = \sum_r a_r^2\,|(\Psi_r,\Psi)|^2 \neq \langle A\rangle_{\Psi}^2 = \left(\sum_r a_r\,|(\Psi_r,\Psi)|^2\right)^2\,. \tag{3.59}$$

Indeed, we can show that $\langle A^2\rangle_{\Psi} \geq \langle A\rangle_{\Psi}^2$ and the equality holds when the state Ψ is the eigenstate of A. To see this, we can consider the square of Hermitian operators B^2, which is always a positive semi-definite operator.

$$\left\langle B^2\right\rangle_{\Psi} = \left(\Psi, B^2\Psi\right) = (B\Psi, B\Psi) \geq 0\,. \tag{3.60}$$

The equality holds only if the operator vanishes the state $B\Psi = 0$. Considering the Hermitian operator $B = A - \langle A\rangle_{\Psi}$, we can show that

$$0 \leq \left\langle (A - \langle A\rangle_{\Psi})^2\right\rangle_{\Psi} = \left\langle A^2\right\rangle_{\Psi} - \langle A\rangle_{\Psi}^2\,. \tag{3.61}$$

Then, we can also find that the equality holds when Ψ is an eigenstate of A.

Cauchy-Schwarz inequality

There is a general inequality possessed in any inner product spaces. Applying the statement to the Hilbert space, for any two state vectors Ψ and Ψ', we have

$$|(\Psi',\Psi)|^2 \leq (\Psi',\Psi')(\Psi,\Psi)\,. \tag{3.62}$$

The equality holds when these states are linearly dependent. It is easy to prove this inequality by considering the state

$$\Psi'' \equiv \Psi - \frac{(\Psi',\Psi)}{(\Psi',\Psi')}\Psi'\,, \tag{3.63}$$

and using the positivity of a quantity $(\Psi'',\Psi'')(\Psi',\Psi') \geq 0$. We can also see that the equality holds when the new state vector Ψ'' is zero vector, namely the state vectors Ψ and Ψ' are linearly dependent.

The Cauchy-Schwarz inequality in the Hilbert space gives a generalization of Heisenberg uncertainty principle. We now consider the root mean square deviation of an Hermitian operator A from its expectation value in a state Ψ.

$$\Delta_\Psi A \equiv \left\langle (A - \langle A \rangle_\Psi)^2 \right\rangle_\Psi^{1/2} . \tag{3.64}$$

It is convenient to use a state Ψ_A defined as

$$\Psi_A \equiv \frac{1}{\sqrt{(\Psi, \Psi)}} (A - \langle A \rangle_\Psi) \Psi , \tag{3.65}$$

then we can write $\Delta_\Psi A = (\Psi_A, \Psi_A)^{1/2}$. For any pair of Hermitian operators, A and B, we can define the similar state vectors Ψ_A and Ψ_B. The Cauchy-Schwarz inequality states that

$$\Delta_\Psi A \Delta_\Psi B \geq |(\Psi_A, \Psi_B)| . \tag{3.66}$$

Furthermore, we can rewrite the inner product of the state vectors Ψ_A and Ψ_B as follows.

$$(\Psi_A, \Psi_B) = \frac{1}{(\Psi, \Psi)} (\Psi, [AB - \langle A \rangle_\Psi \langle B \rangle_\Psi] \Psi) . \tag{3.67}$$

Remind that for any complex number c, we have $|c| = \sqrt{(\mathrm{Re}\, c)^2 + (\mathrm{Im}\, c)^2} \geq |\mathrm{Im}\, c|$. The imaginary part of the inner product (Ψ_A, Ψ_B) is given by

$$\mathrm{Im}(\Psi_A, \Psi_B) = \frac{1}{2i} \frac{1}{(\Psi, \Psi)} \left[(\Psi, AB\Psi) - (\Psi, AB\Psi)^* \right] . \tag{3.68}$$

By using the property $(\Psi, AB\Psi)^* = (\Psi, BA\Psi)$ for Hermitian operators, we can derive a generalized uncertainty relation.

$$\Delta_\Psi A \Delta_\Psi B \geq \frac{1}{2} |\langle [A, B] \rangle_\Psi| . \tag{3.69}$$

Here, we define *the commutator of two operators* as

$$[A, B] \equiv AB - BA . \tag{3.70}$$

The definition of $\Delta_\Psi A$ implies that it represents the standard deviation of the observable A in the state Ψ. Therefore, this leads to an important consequence in quantum mechanics: any pair of observables represented by non-commuting Hermitian operators cannot be simultaneously measured with arbitrary precision.

For example, the operators representing position and momentum in quantum mechanics, X and P, satisfy the canonical commutation relation $[X, P] = i\hbar$. Then, in any state Ψ, the uncertainty in position and momentum operators leads to

$$\Delta_\Phi X \Delta_\Phi P \geq \frac{\hbar}{2} . \tag{3.71}$$

This is known as the Heisenberg uncertainty principle.

Meanwhile, when there are two commuting Hermitian operators, A and B, any eigenstate of A can also be chosen to be an eigenstate of B. Such a state is called a *simultaneous eigenstate* of A and B. Indeed, the simultaneous eigenstate labeled by a and b, eigenstates of A and B respectively, satisfies

$$A\Psi_{a,b} = a\Psi_{a,b}\,, \quad B\Psi_{a,b} = b\Psi_{a,b}\,. \tag{3.72}$$

Hence, the operation of the commutator on the simultaneous eigenstate does not lead to any inconsistency, as expected for commuting operators: $[A, B]\Psi_{a,b} = 0$.

Similarly to the trace of matrices, the trace of an operator in a Hilbert space is defined by introducing a complete orthonormal set of state vectors. Let $\{\Psi_i\}$ be a complete orthonormal set of state vectors in the Hilbert space. The trace of an operator A is then defined as

$$\mathrm{Tr}A \equiv \sum_i (\Psi_i, A\Psi_i)\,. \tag{3.73}$$

This definition is quite useful because it allows us to define the trace of an operator independently of the choice of the complete orthonormal set of state vectors, as far as trace exists. Using another orthonormal set $\{\Phi_j\}$, the state vectors Ψ_i can be expressed as a linear combination of Φ_j's. Then, the trace of the operator A can be expressed in terms of the new orthonormal set as follows:

$$\begin{aligned}\mathrm{Tr}A &= \sum_i (\Psi_i, A\Psi_i) \\ &= \sum_{ij} (\Phi_j, A\Psi_i)(\Psi_i, \Phi_j) = \sum_j (\Phi_j, A\Phi_j)\,.\end{aligned} \tag{3.74}$$

In the intermediate step, we use the completeness relation for the orthonormal set $\{\Phi_j\}$. Similarly to the trace of matrices, the trace of an operator has also obvious properties:

(i). Linearity: $\mathrm{Tr}(\alpha A + \beta B) = \alpha \mathrm{Tr}A + \beta \mathrm{Tr}B$.

(ii). Hermitian conjugate: $\mathrm{Tr}A^\dagger = (\mathrm{Tr}A)^*$.

(iii). Cyclic property: $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$.

These properties are derived from the definition of trace in Hilbert space, Eq. (3.73). Meanwhile, the trace is not always defined for all operators, especially in infinite-dimensional Hilbert spaces. For example, the trace of the unit operator $\mathbb{1}$ in a d-dimensional Hilbert space is $\mathrm{Tr}\mathbb{1} = d$. In an infinite-dimensional Hilbert space, this trace is ill-defined.

Related to the uncertainty principle, the trace of the commutator of two operators provides important insights. For instance, consider the position and momentum operators X and P satisfying the canonical commutation relation $[X, P] = i\hbar\mathbb{1}$. In a finite-dimensional Hilbert space, taking the trace of both sides leads to a contradiction:

$$\mathrm{Tr}([X, P]) = \mathrm{Tr}(i\hbar\mathbb{1}) = i\hbar d\,, \tag{3.75}$$

where d is the dimension of the Hilbert space. However, the cyclic property of the trace implies $\mathrm{Tr}([X, P]) = 0$, which contradicts the previous result. Therefore, the uncertainty relation can only be realized in an infinite-dimensional Hilbert space.

3.4 Symmetries

One of important concepts in physics is symmetry. Even in quantum mechanics, symmetries play a crucial role in determining the behavior of physical systems.

Symmetry Principle

When we change our point of view in certain ways, the physics does not change. As an example, the change of the laboratory frame of measurements does not change the physical outcomes. A symmetry transformation is an operation that maps states to different states while preserving the physical properties of the system. Thus, a symmetry transformation can be described by operators mapping rays into different rays. Let Φ be a state vector, and a symmetry transformation is represented by a linear operator U acting on Φ as follows:

$$\Phi \mapsto U\Phi \,. \tag{3.76}$$

The meaning that the physical outcomes do not change is that the probability between any two states remains invariant under the symmetry transformation:

$$P(\Psi \to \Phi) = |(\Phi, \Psi)|^2 \,, \tag{3.77}$$

Here, Φ and Ψ are normalized state vectors. Under the symmetry transformation, the inner product between any two states transforms as follows:

$$(\Phi, \Psi) \mapsto (U\Phi, U\Psi) = (\Phi, U^\dagger U\Psi) \,. \tag{3.78}$$

Here, in the last equality, the adjoint of an operator U is used. One way to make the inner product invariant under the symmetry transformation is to require that U be a *unitary* operator.

$$U^\dagger U = \mathbb{1} \,. \tag{3.79}$$

The existence of an inverse transformation may not be guaranteed for a general operator. For example, rotation and translation have inverses, which undo the original transformation. The inverse of the unitary operator U exists and U^{-1} maps $U\Phi$ back to Φ for any state vector Φ: $U^{-1}U = UU^{-1} = \mathbb{1}$. It is also true that inverse operator U^{-1} also has its inverse $(U^{-1})^{-1}$, which is equal to the original operator U: by multiplying $(U^{-1})^{-1}$ to both sides of $U^{-1}U = \mathbb{1}$, we get $(U^{-1})^{-1} = U$. Applying the inverse transformation U^{-1} to both sides of $U^\dagger U = \mathbb{1}$, we have $U^\dagger = U^{-1}$

There is a question that is this the only way to represent symmetry transformations in quantum mechanics? Indeed, there is an alternative way to represent symmetry

transformations in quantum mechanics, which is given by an operator U being *antiunitary* and *antilinear*.

$$U(\alpha\Psi + \alpha'\Psi') = \alpha^* U\Psi + \alpha'^* U\Psi' \,, \tag{3.80}$$

$$(U\Phi, U\Psi) = (\Phi, \Psi)^* \,, \tag{3.81}$$

respectively. Under the symmetry transformation represented by an antiunitary operator U, the inner product between any two states transforms as the complex conjugate of the original inner product:

$$(\Phi, \Psi) \to (U\Phi, U\Psi) = (\Phi, \Psi)^* \,. \tag{3.82}$$

We remark that the antiunitary operator cannot be a linear operator. If the antiunitary operator were linear, it would contradict the definition of antiunitarity, as shown in the following argument. Let U be antiunitary and linear operator, then

$$\alpha(U\Phi, U\Psi) = (U\Phi, U\alpha\Psi) = (\alpha\Psi, \Phi)^* = \alpha^*(\Psi, \Phi)^* = \alpha^*(U\Phi, U\Psi) \,. \tag{3.83}$$

This is contradict for $\alpha \in \mathbb{C}$. But, most symmetries are represented by linear unitary operators. There is a known theorem that any symmetry transformation in quantum mechanics can be represented either by a unitary or an antiunitary operator: *Wigner's theorem*.

The unitary operators satisfy all the group axioms, and therefore they form a group:

(i). Existence of an identity element: $\mathbb{1}$

(ii). Existence of inverse elements: U^{-1} such that $UU^{-1} = U^{-1}U = \mathbb{1}$

(iii). Closure: If U_1 and U_2 are unitary, then U_1U_2 is also unitary

(iv). Associativity: $(U_1U_2)U_3 = U_1(U_2U_3)$

This fact tells us about the importance of learning *group theory* and *representation theory*. The group theory provides the mathematical framework to study the structure and properties of symmetry groups. Meanwhile, the representation theory becomes important, as it allows us to understand how the symmetry group acts on the Hilbert space of quantum states.

There are two types of symmetries: one is discrete (such as parity or time-reversal), and the other is continuous (such as spatial translations or rotations). As for the continuous symmetries, there is a class of symmetry transformations that can be connected smoothly to the identity transformation. The linear operators arbitrary close to a unit operator $\mathbb{1}$ can be written as $U_\epsilon = \mathbb{1} + i\epsilon T + \mathcal{O}(\epsilon^2)$ with an infinitesimal parameter ϵ. T is called as a *generator* of the continuous symmetry transformation. Such a unitary operator also satisfies the unitarity condition, $U_\epsilon^\dagger U_\epsilon = \mathbb{1}$, and hence the generator T must be Hermitian, $T = T^\dagger$.

So far, we discuss the symmetry transformation of state vectors $\Psi \mapsto U\Psi$. Under the symmetry transformation, the expectation value of an observable A transforms as

$$(\Psi, A\Psi) \mapsto (\Psi, U^{-1}AU\Psi) \,. \tag{3.84}$$

Thus, instead of the symmetry transformation of the state vectors, we can equivalently consider the transformation of the operators.

$$A \to U^{-1}AU\,. \tag{3.85}$$

Transformations of this type are known as a *similarity transformation* in Mathematics. We remark that similarity transformations preserve algebraic properties of operators:

$$U^{-1}AU \times U^{-1}BU = U^{-1}(AB)U\,, \tag{3.86}$$
$$U^{-1}AU + U^{-1}BU = U^{-1}(A+B)U\,, \tag{3.87}$$

As for the infinitesimal unitary transformation, we have

$$A \mapsto A - i\epsilon[T, A] + \mathcal{O}\left(\epsilon^2\right)\,, \tag{3.88}$$

Thus, the infinitesimal unitary transformation of an operator is determined by the commutator with the generator. Since the generator is itself an operator, it can also undergo transformations under the unitary symmetry operations.

There are two equivalent ways to view transformations: one is the active view, and the other is the passive view. In the active view, we consider the transformation as acting directly on the state vectors, while in the passive view, we interpret the transformation as a change of the basis. From the active/passive view of transformation of the basis vectors, the transformation $(\Psi \mapsto U\Psi\,, A \mapsto A)$ is considered to be active, while $(\Psi \mapsto \Psi\,, A \mapsto U^{-1}AU)$ is considered to be passive.

3.5 Space translation

As an example of a symmetry transformation, let us consider space translation: the space translation is a fundamental symmetry in physics, which corresponds to shifting the coordinates of *all* particles by a constant vector. The law of physics should remain invariant even if we change the origin of the coordinate system. Under the space translation by a vector $\boldsymbol{a}$, any particle at position $\boldsymbol{X}_n$ is shifted to $\boldsymbol{X}_n + \boldsymbol{a}$ (n denotes the particle label). In terms of operators acting on the Hilbert space, the symmetry transformation is represented by the unitary operator $U(\boldsymbol{a})$, this unitary transformation gives the similarity transformation of the position operator,

$$U^{-1}(\boldsymbol{a})\boldsymbol{X}_n U(\boldsymbol{a}) = \boldsymbol{X}_n + \boldsymbol{a}\,, \tag{3.89}$$

In the infinitesimal limit, the unitary operator can be expanded to first order in $\boldsymbol{a}$, giving

$$U(\boldsymbol{a}) = 1 - i\frac{\boldsymbol{P}}{\hbar}\cdot\boldsymbol{a} + \mathcal{O}\left(\boldsymbol{a}^2\right)\,. \tag{3.90}$$

Here, $\boldsymbol{P}$ is the generator of space translations. The $\hbar$ in the denominator ensures that the generator $\boldsymbol{P}$ has the dimension of momentum. Then, the commutation relation between the position operator and the generator of translations can be derived.

$$U^{-1}(\boldsymbol{a})\boldsymbol{X}_n U(\boldsymbol{a}) = \boldsymbol{X}_n + \frac{i}{\hbar}\left[\boldsymbol{P}\cdot\boldsymbol{a}, \boldsymbol{X}_n\right] + \mathcal{O}(\boldsymbol{a}^2)\,, \tag{3.91}$$

and this should coincide with the finite translation result $\boldsymbol{X}_n + \boldsymbol{a}$, leading to the commutation relation

$$\frac{i}{\hbar}\left[\boldsymbol{P}\cdot\boldsymbol{a}, \boldsymbol{X}_n\right] = \boldsymbol{a}\,, \tag{3.92}$$

or equivalently,

$$[X_{ni}, P_j] = i\hbar\delta_{ij}\,. \tag{3.93}$$

Here, this commutation relation coincides with the canonical commutation relation between position and momentum operators in quantum mechanics. The question is then: what is the physical meaning of the momentum operator $\boldsymbol{P}$? The space translation makes the coordinates of all particles shift by the same amount, and therefore the generator of translations $\boldsymbol{P}$ corresponds to the total momentum of the system. The total momentum is given by the sum of the individual momenta of all particles, which is represented by the operator $\boldsymbol{P} = \sum_n \boldsymbol{P}_n$, with $\boldsymbol{P}_n$ being the momentum operator of the n-th particle. This ensures that the commutation relations between the total momentum and the position operators of individual particles are consistent with the canonical commutation relations:

$$[X_{ni}, P_{mj}] = i\hbar\delta_{ij}\delta_{nm}\,. \tag{3.94}$$

Here, n, m label the particles, and i, j label the spatial components of the position and momentum operators.

The order of translations does not matter for space translations: a translation by $\boldsymbol{a}$ followed by a translation by $\boldsymbol{b}$ is equivalent to a translation by $\boldsymbol{b}$ followed by a translation by $\boldsymbol{a}$. In terms of the translation operators, this means that the operators corresponding to translations by different vectors commute with each other:

$$U(\boldsymbol{a})U(\boldsymbol{b}) = U(\boldsymbol{b})U(\boldsymbol{a})\,, \tag{3.95}$$

The infinitesimal version of this commutation relation can be derived by expanding the translation operators to the order of $a_i b_j$:

$$U(\boldsymbol{a})U(\boldsymbol{b}) = 1 - \frac{i}{\hbar}\boldsymbol{P}\cdot(\boldsymbol{a}+\boldsymbol{b}) - \frac{1}{\hbar^2}(\boldsymbol{P}\cdot\boldsymbol{a})(\boldsymbol{P}\cdot\boldsymbol{b}) + O(a^2, b^2)\,, \tag{3.96}$$

$$U(\boldsymbol{b})U(\boldsymbol{a}) = 1 - \frac{i}{\hbar}\boldsymbol{P}\cdot(\boldsymbol{a}+\boldsymbol{b}) - \frac{1}{\hbar^2}(\boldsymbol{P}\cdot\boldsymbol{b})(\boldsymbol{P}\cdot\boldsymbol{a}) + O(a^2, b^2)\,. \tag{3.97}$$

It leads to the commutation relation between the components of the momentum operator:

$$[P_i, P_j] = 0\,. \tag{3.98}$$

We can generalize this idea to finite translations by considering the limit of many infinitesimal translations. A translation vector is denoted by $\boldsymbol{r}_N = N\boldsymbol{a}$, where N is a large integer and $\boldsymbol{a}$ is an infinitesimal translation vector. The finite translation vector is a limit of many infinitesimal translations: $\boldsymbol{r} = \lim_{N\to\infty}\boldsymbol{r}_N$. Therefore, the finite

translation operator can be expressed as the limit of many infinitesimal translation operators:

$$U(\boldsymbol{r}) = \lim_{N\to\infty} [U(\boldsymbol{a})]^N = \lim_{N\to\infty} \left[1 - \frac{i}{\hbar}\boldsymbol{P}\cdot\boldsymbol{a}\right]^N = \exp\left(-\frac{i}{\hbar}\boldsymbol{P}\cdot\boldsymbol{r}\right) . \tag{3.99}$$

Here, we can find the s-th power term of the expansion of the N-th power of the infinitesimal translation operator:

$$\left[1 - \frac{i}{\hbar}\boldsymbol{P}\cdot\boldsymbol{a}\right]^N = \sum_{s=0}^{N} \binom{N}{s} \left(-\frac{i}{\hbar}\boldsymbol{P}\cdot\boldsymbol{a}\right)^s , \tag{3.100}$$

and use the definition of exponentiation of operators in the limit $N \to \infty$.

$$e^{iA} = \sum_{n=0}^{\infty} \frac{i^n}{n!} A^n , \tag{3.101}$$

with A denoting an operator.

Let us define Φ_0 as one-particle state with a definite position at the origin. It means that the operation of $\boldsymbol{X}$ on Φ_0 gives zero, $\boldsymbol{X}\Phi_0 = 0$. We can construct a state with a definite position $\boldsymbol{x}$ by applying the translation operator $U(\boldsymbol{x})$ to Φ_0.

$$\Phi_x \equiv U(\boldsymbol{x})\Phi_0 , \quad \boldsymbol{X}\Phi_x = \boldsymbol{x}\Phi_x . \tag{3.102}$$

Now, we can define a momentum-representation state Ψ_p, which is an eigenstate of the momentum operator with eigenvalue $\boldsymbol{p}$: $\boldsymbol{P}\Psi_p = \boldsymbol{p}\Psi_p$. Then, we can consider a scalar product of the momentum-representation state vector Ψ_p and the position-representation state vector Φ_x: (Ψ_p, Φ_x). The scalar product (Ψ_p, Φ_x) can be evaluated using the translation operator.

$$(\Psi_p, \Phi_x) = (\Psi_p, U(\boldsymbol{x})\,\Phi_0) = e^{-i\boldsymbol{p}\cdot\boldsymbol{x}/\hbar}(\Psi_p, \Phi_0) . \tag{3.103}$$

In the last equality, we used the fact that Ψ_p is an eigenstate of the momentum operator with eigenvalue $\boldsymbol{p}$. It is convenient to choose the normalization of the state as

$$(\Psi_p, \Phi_x) \equiv \frac{1}{(2\pi\hbar)^{\frac{3}{2}}} e^{-i\boldsymbol{p}\cdot\boldsymbol{x}/\hbar} . \tag{3.104}$$

The complex conjugate of this scalar product gives the coordinate wave function in the position representation:

$$\psi_p(x) \equiv (\Phi_x, \Psi_p) = \frac{1}{(2\pi\hbar)^{\frac{3}{2}}} e^{i\boldsymbol{p}\cdot\boldsymbol{x}/\hbar} \tag{3.105}$$

This is the plane wave solution representing a particle with definite momentum $\boldsymbol{p}$ in the position representation. The normalization chosen here ensures that the position and momentum eigenstates are properly normalized according to the Dirac delta function

conventions. Here, the scalar product of state vectors in the position representation is given by

$$(\Phi_{x'}, \Phi_x) = \delta^3(\boldsymbol{x}' - \boldsymbol{x})\,, \tag{3.106}$$

which ensures that the position eigenstates are orthonormal according to the Dirac delta function conventions. Then, the scalar product of state vectors in the momentum representation is

$$\begin{aligned}(\Psi_{p'}, \Psi_p) &= \int d^3x (\Psi_{p'}, \Phi_x)(\Phi_x, \Psi_p) \\ &= \frac{1}{(2\pi\hbar)^3} \int d^3x e^{i(\boldsymbol{p}-\boldsymbol{p}')\cdot\boldsymbol{x}/\hbar} = \delta^3(\boldsymbol{p} - \boldsymbol{p}')\,. \end{aligned} \tag{3.107}$$

3.6 Time translation

The second example for the symmetry transformations is the *time translation*. For a state vector $\Psi(t)$ at time t, the time translation operator $U(\tau)$, which induces the time shift $t \to t + \tau$, acts as

$$U(\tau)\Psi(t) = \Psi(t + \tau)\,, \tag{3.108}$$

In a similar way to the previous subsection, we introduce an Hermitian operator H to generate the time translation.

$$U(\tau) \equiv e^{-i\frac{1}{\hbar}H\tau}\,. \tag{3.109}$$

Analogous to the momentum operator generating spatial translations, the Hermitian operator H generates time translations. This can be taken as a definition of the *Hamiltonian* H. Thus, the time dependence of any physical state can be factored out as a unitary evolution generated by the Hamiltonian:

$$\Psi(t) = e^{-i\frac{1}{\hbar}Ht}\Psi(0)\,, \tag{3.110}$$

and this time evolution leaves the scalar product invariant.

$$(\Psi(t), \Psi(t)) = (\Psi(0), \Psi(0))\,, \tag{3.111}$$

Since we have factored out the time dependence as a unitary evolution, the state vector $\Psi(t)$ satisfies the following time evolution equation:

$$i\hbar \cdot \frac{d}{dt}\Psi(t) = H\Psi(t)\,, \tag{3.112}$$

This is nothing but the *time-dependent Schrödinger equation*. One remark is that this is *the time evolution equation for the state vector*, not for the wavefunction in a particular representation.

There are two equivalent formalisms in quantum mechanics: the *Schrödinger picture* and the *Heisenberg picture*. These are related to the choice of whether the time

dependence is assigned to the state vectors (Schrödinger picture: state vectors evolve in time $\Psi(t)$, operators are fixed $A = A(0)$) or to the operators (Heisenberg picture: state vectors are fixed $\Psi = \Psi(0)$, operators evolve in time $A(t)$).

Operators in two different pictures are related by the following unitary transformation (similarity transformation of the operator A):

$$A_H(t) = e^{iHt/\hbar} A_S e^{-iHt/\hbar} \,, \tag{3.113}$$

Here, A_S denotes the operator A in the Schrödinger picture, and $A_H(t)$ denotes the operator A in the Heisenberg picture at time t. The Hamiltonian is also an operator, so we can define the Hamiltonian operators in both the Schrödinger and Heisenberg pictures. However, the Hamiltonian commutes with itself at all times, and the Hamiltonian in two pictures is exactly same.

Symmetry and Conserved Quantity

The time dependence of an operator A in the Heisenberg picture leads to so called *the Heisenberg equation of motion*: the time evolution is governed by the commutator with the Hamiltonian.

$$\frac{d}{dt} A_H(t) = \frac{i}{\hbar} \left[H, A_H(t) \right] \,, \tag{3.114}$$

It is easy to prove the time evolution equation. Since the Hamiltonian H commutes with itself, we can easily compute the derivative of the unitary operator:

$$\frac{d}{dt} \left(e^{-iHt/\hbar} \right) = -\frac{i}{\hbar} H e^{-iHt/\hbar} = -\frac{i}{\hbar} e^{-iHt/\hbar} H \,. \tag{3.115}$$

The Hamiltonian determines the time evolution of the operators in the Heisenberg picture. From the Heisenberg equation of motion, we can see that if an operator commutes with the Hamiltonian, its time derivative vanishes, indicating that $A_H(t)$ is a conserved quantity. Such a case, the operator $A_H(t)$ is time independent, and hence there is no difference between the Schrödinger and Heisenberg pictures for this operator.

Moving on to the discussion of symmetries and conserved quantities in Schrödinger picture. When a symmetry is present in the Schrödinger picture, there is a symmetry transformation connecting the state vectors Ψ and $U\Psi$. Since these state vectors represent the same physical state from different viewpoints (or observers), their time evolutions must be consistent.

$$\Psi(t) = e^{-iHt/\hbar} \Psi(0) \,, \qquad (U\Psi)(t) = e^{-iHt/\hbar} (U\Psi)(0) \,. \tag{3.116}$$

On the other hand, the time evolution of the transformed state is also written as $(U\Psi)(t) = U e^{-iHt/\hbar} \Psi(0)$. Therefore, to be consistent, we must have the relation between the time translation operator and Hamiltonian as follows.

$$e^{-iHt/\hbar} U = U e^{-iHt/\hbar} \,. \tag{3.117}$$

From this relation, we can deduce that the Hamiltonian commutes with the symmetry operator:

$$U^{-1}HU = H\,. \tag{3.118}$$

Once the infinitesimal transformation is considered, we can write $U = e^{ixT}$ for some Hermitian operator T and small parameter x. The commutation relation with the Hamiltonian then becomes

$$[H, T] = 0\,, \tag{3.119}$$

indicating that T is a conserved quantity associated with the symmetry. We can see that the presence of a symmetry leads to a conserved quantity in the system. For example, if the system is invariant under spatial translations, the corresponding conserved quantity is the momentum. Similarly, if the system is invariant under time translations, the conserved quantity is the energy.

So far, we consider the case where the symmetry operator associated with time translation U is unitary and linear. In this case, the Hamiltonian commutes with U, leading to conserved quantities associated with the symmetry.

We may consider the case where the symmetry operator associated with time translation U is antilinear. In this case, the commutation relation with the Hamiltonian may take a different form, and the analysis of conserved quantities needs to be adjusted accordingly. Once we start from Eq. (3.117), we obtain the similarity transformation of the Hamiltonian under the symmetry operator U as $U^{-1}HU = -H$. The Hamiltonian changes sign under the antilinear symmetry transformation, and it implies that a positive energy state is mapped to a negative energy state under this symmetry. If such a symmetry exists, it would imply the existence of negative energy states, which is in conflict with the observed stability of matter. To avoid such a situation, the antilinear transformation associated with time must be accompanied by the time reversal. Once the time reversal is included, the time evolution of the transformed state is

$$(U\Psi)(t) = e^{\frac{i}{\hbar}Ht}(U\Psi)(0)\,, \tag{3.120}$$

and the corresponding commutation relation between the time translation operator and the symmetry operator is modified accordingly.

$$e^{+\frac{i}{\hbar}Ht}U = Ue^{-\frac{i}{\hbar}Ht}\,. \tag{3.121}$$

As a result, we obtain the commutation relation between the Hamiltonian and the symmetry operator, $U^{-1}HU = H$, and no negative energy state is not required after the symmetry transformation.

4 Particle States in Central Potential

It is one of central programs of quantum mechanics to understand the behavior of a state. In this section, we first consider the meaning of Schrödinger equation for a single particle moving in three-dimensional space. In particular, we focus on the case where the particle is subject to a central potential, which depends only on the distance from the origin of coordinates. In the usual textbooks, the treatment often starts from one-dimensional problems. However, for a central potential, the problem can be effectively reduced to a one-dimensional problem along the radial direction. There is no fundamental difference between one-dimensional and three-dimensional treatments in this context.

4.1 Schrödinger Equation for a Single Particle

Let us consider a single particle of mass μ moving in a central potential. The central potential depends only on the distance from the origin, i.e., $V = V(r)$ with $r = \sqrt{x_1^2 + x_2^2 + x_3^2}$. The Hamiltonian for this system is given by

$$H = \frac{1}{2\mu}\boldsymbol{p}^2 + V(r)\,, \tag{4.1}$$

where the first and the second terms represent the kinetic term and the potential term, respectively. The momentum operator is defined as $\boldsymbol{p} = -i\hbar\boldsymbol{\nabla}$.

The time-independent Schrödinger equation for a state with a definite energy E is given by

$$H\psi(\boldsymbol{x}) = \left[-\frac{\hbar^2}{2\mu}\nabla^2 + V(r)\right]\psi(\boldsymbol{x}) = E\psi(\boldsymbol{x})\,, \tag{4.2}$$

$$\nabla^2 \equiv \partial_1^2 + \partial_2^2 + \partial_3^2\,, \quad \partial_i \equiv \frac{\partial}{\partial x_i}\,. \tag{4.3}$$

Here, $\boldsymbol{x} = (x_1, x_2, x_3)$ represents the position vector of the particle in the three-dimensional space, in particular in Cartesian coordinates. For a state with a definite energy E, the time dependence of the wave function is factored out as $e^{-iEt/\hbar}$. To find a solution to the Schrödinger equation under a central potential, it is convenient to switch to spherical coordinates.

$$(x_1, x_2, x_3) \quad \longrightarrow \quad (r, \theta, \varphi)\,. \tag{4.4}$$

Here, r is the radial distance, θ is the polar angle, and φ is the azimuthal angle. In the spherical coordinate system, the Laplacian operator takes the form

$$\nabla^2 = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2}{\partial\varphi^2}\,. \tag{4.5}$$

In an analogy to the classical mechanics, under a spherical potential, the rotational motion of the particle is conserved. It means that the angular momentum of the

particle is conserved. This situation is same even in quantum mechanics, and in the terminology of quantum mechanics, it means that the angular momentum operator commutes with the Hamiltonian. Now, let us define the angular momentum operator in quantum mechanics as

$$\boldsymbol{L} \equiv \boldsymbol{x} \times \boldsymbol{p} = -i\hbar \boldsymbol{x} \times \boldsymbol{\nabla} \,. \tag{4.6}$$

The components of the angular momentum operator in Cartesian coordinates are given by

$$L_i = -i\hbar \sum_{j,k} \epsilon_{ijk} x_j \partial_k \,, \tag{4.7}$$

for i, j, k running over three Cartesian coordinates $1, 2, 3$. ϵ_{ijk} denotes the *Levi-Civita symbol*, which is totally antisymmetric in its indices.

$$\epsilon_{ijk} \equiv \begin{cases} +1 & \text{even permutation} \\ -1 & \text{odd permutation} \\ 0 & \text{otherwise} \end{cases} \,, \tag{4.8}$$

with the convention $\epsilon_{123} = +1$.

To show commutation of the angular momentum operator with Hamiltonian, we first need to establish the commutation relations between the angular momentum operator and the position and momentum operators. Using the commutation relation of the position and momentum operators, $[x_i, \partial_j] = \delta_{ij}$, we can derive the commutation relations between the angular momentum operator and the position and momentum operators as follows.

$$[L_i, x_j] = i\hbar \sum_k \epsilon_{ijk} x_k \,, \qquad [L_i, \partial_j] = i\hbar \sum_k \epsilon_{ijk} \partial_k \,, \tag{4.9}$$

$$[L_i, x^2] = 0 \,, \qquad [L_i, \partial^2] = 0 \,. \tag{4.10}$$

Here, $x^2 \equiv \sum_i x_i^2$ and $\partial^2 \equiv \sum_i \partial_i^2$. The commutation relation with the scalar quantities like x^2 and ∂^2 is zero, as expected for rotationally invariant quantities. For any vectors v_i constructed from x_i and ∂_i, it is possible to show that the commutation relation with the angular momentum operator takes the same form as for the position and momentum operators:

$$[L_i, v_j] = i\hbar \sum_k \epsilon_{ijk} v_k \,. \tag{4.11}$$

It is straightforward to verify the commutation relation with the three-dimensional vectors, which is given by $\boldsymbol{v} = f(x^2, p^2)\boldsymbol{x}$ or $\boldsymbol{v} = f(x^2, p^2)\boldsymbol{p}$. Another possible vector is proportional to the angular momentum operator itself, $\boldsymbol{L} = \boldsymbol{x} \times \boldsymbol{p}$. Indeed, the commutation relation of the angular momentum operator with itself is given by

$$[L_i, L_j] = i\hbar \sum_k \epsilon_{ijk} L_k \,. \tag{4.12}$$

To show the commutation relation of the angular momentum operator with itself, we can use a relation for the Levi-Civita symbol, which is known as the *Jacobi-type identity* for the Levi-Civita symbol.

$$\sum_k (\epsilon_{ijk}\epsilon_{lkm} + \epsilon_{ikl}\epsilon_{jkm} + \epsilon_{ikm}\epsilon_{jlk}) = 0\,. \tag{4.13}$$

For the vector $\boldsymbol{v}$ satisfying Eq. (4.11), the commutation relation of v^2 with the angular momentum operator is given by

$$[L_i, v^2] = \sum_j ([L_i, v_j]v_j + v_j[L_i, v_j]) = 0\,. \tag{4.14}$$

In particular, L^2 commutes with each component of the angular momentum operator, $[L^2, L_i] = 0$. Since the angular momentum operator $\boldsymbol{L}$ commutes with the scalar quantities like x^2 and ∂^2, it commutes with the Hamiltonian for a central potential, and L^2 does as well. It is important to note that while L^2 commutes with each component of the angular momentum operator, the individual components L_i do not commute with each other. Thus, the *simultaneous eigenstates* of the Hamiltonian and the angular momentum operator are typically chosen to be eigenstates of L^2 and one component of $\boldsymbol{L}$, usually L_3. In addition, since the angular momentum operator $\boldsymbol{L}$ commutes with r^2, the operation of $\boldsymbol{L}$ on the wave function actually acts on the angular coordinates (θ, φ).

Now, let us find the expression of the angular momentum operators in terms of spherical coordinates. The Cartesian coordinates are

$$x_1 = r\sin\theta\cos\varphi\,, \quad x_2 = r\sin\theta\sin\varphi\,, \quad x_3 = r\cos\theta\,. \tag{4.15}$$

We can find each component of the angular momentum in the spherical coordinates as follows.

$$L_1 = i\hbar\left(\sin\varphi\frac{\partial}{\partial\theta} + \cot\theta\cos\varphi\frac{\partial}{\partial\varphi}\right)\,, \tag{4.16}$$

$$L_2 = i\hbar\left(-\cos\varphi\frac{\partial}{\partial\theta} + \cot\theta\sin\varphi\frac{\partial}{\partial\varphi}\right)\,, \tag{4.17}$$

$$L_3 = -i\hbar\frac{\partial}{\partial\varphi}\,. \tag{4.18}$$

From the expressions of the angular momentum operators in spherical coordinates, we can find that it is convenient to label the states by the eigenvalue of L_3: the expression of L_3 in the spherical coordinate is rather simple. The square of angular momentum operator L^2 in spherical coordinates is computed by taking the sum of its components squared:

$$L^2 = \sum_{i=1}^{3} L_i^2 = -\hbar^2\left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\varphi^2}\right]\,. \tag{4.19}$$

Then, the Laplacian operator is written in terms of the radial derivative and the square of the angular momentum operator as follows.

$$\nabla^2 = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) - \frac{L^2}{r^2\hbar^2}\,. \tag{4.20}$$

Therefore, the time-independent Schrödinger equation in spherical coordinates is written as follows.

$$E\psi(\boldsymbol{x}) = \left[-\frac{\hbar^2}{2\mu r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + \frac{L^2}{2\mu r^2} + V(r)\right]\psi(\boldsymbol{x})\,. \tag{4.21}$$

Now, let us consider the spectrum of the angular momentum operator L^2. As long as the potential $V(r)$ is not extremely singular at the origin, the wave function behave smoothly near $r \to 0$. In particular, the presence of the angular momentum term in the Hamiltonian implies that the wave function smoothly connected zero near the origin. Thus, introducing an integer ℓ, the wave function behaves as

$$\psi(\boldsymbol{x}) \to r^\ell Y(\theta,\varphi)\,. \tag{4.22}$$

Here, $Y(\theta,\varphi)$ denotes the function depending only on the angular coordinates θ and φ. We assume that the potential $V(r)$ is less singular than $1/r^2$ to allow the potential term in the Schrödinger equation to be negligible compared to the angular momentum term near the origin. Under the assumption, the Schrödinger equation near the origin is dominated by the kinetic term and the angular momentum term, leading to the following equation.

$$0 = -\frac{\hbar^2}{r^2}\frac{\partial}{\partial r}\left[r^2\frac{\partial}{\partial r}\left(r^\ell Y(\theta,\varphi)\right)\right] + \frac{L^2}{r^2}r^\ell Y(\theta,\varphi)\,. \tag{4.23}$$

Therefore, the function $Y(\theta,\varphi)$ must satisfy the following equation for the angular momentum operator L^2 near the origin.

$$L^2 Y(\theta,\varphi) - \hbar^2\ell(\ell+1)Y(\theta,\varphi) = 0\,. \tag{4.24}$$

So far, we consider the behavior of the wave function near the origin. However, once we chose $\psi(\boldsymbol{x})$ as an eigenfunction of L^2 (commute with H), the eigenvalue of L^2 is fixed for the entire wave function.

$$L^2\psi = \hbar^2\ell(\ell+1)\psi(\boldsymbol{x})\,. \tag{4.25}$$

Since the operator L^2 acts only on the angular part of the wave function, the wave function should be proportional to a product of a radial function and an angular function, i.e., we can separate variables as $\psi(\boldsymbol{x}) = R_{n,\ell}(r)Y_\ell^m(\theta,\varphi)$. The functions $Y_\ell^m(\theta,\varphi)$ are labeled by the quantum numbers ℓ and m, corresponding to the eigenvalues of L^2 and L_3, respectively. The functions $Y_\ell^m(\theta,\varphi)$ are called the *spherical harmonics*.

$$L^2 Y_\ell^m(\theta,\varphi) = \hbar^2\ell(\ell+1)Y_\ell^m(\theta,\varphi)\,, \tag{4.26}$$

$$L_3 Y_\ell^m(\theta,\varphi) = \hbar m Y_\ell^m(\theta,\varphi)\,. \tag{4.27}$$

Meanwhile, the functions $R_{n,\ell}(r)$ are labeled by the quantum numbers n and ℓ, corresponding to the energy level and angular momentum quantum numbers, respectively. After separating the variables, we obtain the radial Schrödinger equation,

$$\left[-\frac{\hbar^2}{2\mu r^2}\frac{d}{dr}\left(r^2\frac{d}{dr}\right)+\frac{\hbar^2\ell(\ell+1)}{2\mu r^2}+V(r)\right]R_{n,\ell}(r)=ER_{n,\ell}(r)\,. \tag{4.28}$$

By separating the variables, we reduce the three-dimensional Schrödinger equation to a one-dimensional radial equation for $R_{n,\ell}(r)$, in particular ordinary differential equation (ODE). There is a unitary (normalization) condition for the wave function to be normalizable, which gives us the probability interpretation of the wave function. The integration of the wave function square over the three dimensions should be finite:

$$\int d^3x\,|\psi(\boldsymbol{x})|^2<\infty\,, \tag{4.29}$$

and the finiteness of the radial integral leads to the requirement on the radial wave function $R_{n,\ell}(r)$:

$$\int_0^\infty dr r^2|R_{n,\ell}(r)|^2<\infty\,. \tag{4.30}$$

We may also define the radial wave function as $u(r)\equiv rR(r)$, which gives the radial Schrödinger equation in a form similar to the one-dimensional Schrödinger equation.

$$-\frac{\hbar^2}{2\mu}\frac{d^2u(r)}{dr^2}+\left[V(r)+\frac{\ell(\ell+1)\hbar^2}{2\mu r^2}\right]u(r)=Eu(r)\,, \tag{4.31}$$

and the normalization condition for $u(r)$ becomes

$$\int_0^\infty dr\,|u(r)|^2<\infty\,. \tag{4.32}$$

This is quite similar to the one-dimensional Schrödinger equation, but there are several differences. One of the main differences is the presence of the centrifugal term in the potential term: this term behaves like a repulsive potential for states with nonzero angular momentum, and the wave function approaches zero as $r\to 0$ in such a case. The second difference is the boundary condition for the radial wave function: in one dimension, the wave function is defined over the entire real line, whereas in the radial case, it is defined only for $r\in[0,\infty)$.

4.2 Spherical Harmonics

As discussed in the previous subsection, the angular part of the wave function is described by the spherical harmonics, $Y_\ell^m(\theta,\varphi)$. This subsection is devoted to the properties and mathematical structure of the spherical harmonics. The spherical harmonics are the eigenfunctions of the angular momentum operators L^2 and L_3, satisfying the eigenvalue equations, Eqs. (4.26) and (4.27). We recall that Eq. (4.26) is rewritten as

$\nabla^2\left(r^\ell Y_\ell^m\right) = 0$, and $r^\ell Y_\ell^m$ is the homogeneous polynomial of degree ℓ in $x_\pm \equiv x_1 \pm i x_2$ and x_3. Here, the homogeneous polynomial of degree ℓ contains only terms that are of total degree ℓ in $x_\pm$ and x_3. For example, a homogeneous polynomial of degree one is $\sum_i a_i x_i$, and a homogeneous polynomial of degree two is $\sum_{ij} a_{ij} x_i x_j$. It is apparent because the function $\Phi(r,\theta,\varphi) \equiv r^\ell Y_\ell^m(\theta,\varphi)$ scales as λ^ℓ when r is scaled by λ, which is the defining property of a homogeneous polynomial of degree ℓ.

The eigenvalue m of $L_3 = -i\hbar\partial_\varphi$ just counts the difference between the number of x_+ and x_- factors in the homogeneous polynomial: $m \equiv n_+ - n_-$. Indeed, in the spherical coordinate system, remind that $x_\pm = r\sin\theta e^{\pm i\varphi}$ and $x_3 = r\cos\theta$, and φ dependence of the homogeneous polynomial is entirely contained in the factors of $x_\pm$. For example, $L_3(x_+)^n = \hbar n (x_+)^n$ and $L_3(x_+x_-) = 0$. The function $r^\ell Y_\ell^m$ may $(x_+)^\ell$ or $(x_-)^\ell$, namely $m_{\max} = \ell$ and $m_{\min} = -\ell$. Since m is an integer and m runs from $m = -\ell\,, -\ell+1\,, \cdots \ell-1\,, \ell$, there are $2\ell+1$ possible values of m for a given ℓ.

Let us consider the uniqueness of the spherical function for given labels ℓ and m. Any homogeneous polynomials of order ℓ is written as

$$\sum_{\nu_+,\nu_-} C_{\nu_+\nu_-}\,(x_+)^{\nu_+}\,(x_-)^{\nu_-}\,(x_3)^{\ell-\nu_+-\nu_-}\,. \tag{4.33}$$

The homogeneous polynomial $r^\ell Y_\ell^m$ can contain any terms with $\nu_+ - \nu_- = m$. If there are more than two terms for given m, it is not apparent at first glance that only single combination of these terms remain. Otherwise, there are multiple linearly independent homogeneous polynomials for the same m, and the uniqueness of Y_ℓ^m is not guaranteed. For given ℓ, the powers of orders of x_+ and x_- (and also x_3) is constrained as, $0 \le \nu_+ \le \ell$ and $0 \le \nu_- \le \ell - \nu_+$. Hence, the total number of the homogeneous polynomial of order ℓ is computed as

$$N_\ell = \sum_{\nu_+=0}^{\ell}\sum_{\nu_-=0}^{\ell-\nu_+} 1 = \frac{1}{2}(\ell+1)(\ell+2)\,. \tag{4.34}$$

Meanwhile, the Laplacian of homogeneous polynomial of order ℓ is a homogeneous polynomial of order $\ell-2$: therefore, there are $N_{\ell-2}$ independent constraints on the coefficients of the homogeneous polynomial of order ℓ. Hence, the number of independent homogeneous polynomials of order ℓ that satisfy the Laplace equation is

$$N_\ell - N_{\ell-2} = \frac{1}{2}(\ell+1)(\ell+2) - \frac{1}{2}(\ell-1)\ell = 2\ell+1\,, \tag{4.35}$$

which matches the number of possible values of m for a given ℓ. In other words, the number of independent homogeneous polynomials of order ℓ that satisfy the Laplacian constraints is exactly equal to the number of spherical harmonics Y_ℓ^m for a given ℓ, which is $2\ell+1$: this demonstrates the uniqueness of the spherical harmonics for given ℓ and m.

Because we choose the spherical harmonics Y_ℓ^m to be simultaneous eigenfunctions of L^2 and L_3. The application of L_3 gives the eigenvalue $\hbar m$, and hence the spherical harmonics Y_ℓ^m must have the form $e^{im\varphi}$ in the azimuthal angle φ.

$$Y_\ell^m(\theta,\phi) \propto P_\ell^{|m|}(\cos\theta)e^{im\phi}\,. \tag{4.36}$$

The function $P_\ell^{|m|}(\cos\theta)$ is known as the *associated Legendre function*, and $P_\ell^{|m|}(\cos\theta)$ satisfies the associated Legendre differential equation:

$$\left[-\frac{1}{\sin\theta}\frac{d}{d\theta}\left(\sin\theta\frac{d}{d\theta}\right)+\frac{m^2}{\sin^2\theta}\right]P_\ell^{|m|}(\cos\theta)=\ell(\ell+1)P_\ell^{|m|}(\cos\theta)\,. \tag{4.37}$$

We remark that the differential equation for $P_\ell^{|m|}(\cos\theta)$ only depend on m^2, and hence the function $P_\ell^{|m|}(\cos\theta)$ does not have the dependence on sign of m.

Now, let us see several examples of how to construct Y_ℓ^m. We remind the reader that $x_\pm = r\sin\theta e^{\pm i\varphi}$ and $x_3 = r\cos\theta$. There are several requirements for constructing the spherical harmonics Y_ℓ^m.

(i). $r^\ell Y_\ell^m(\theta,\varphi)$ are homogeneous polynomials.

(ii). $r^\ell Y_\ell^m(\theta,\varphi)$ satisfy the Laplacian constraint, i.e., $\nabla^2(r^\ell Y_\ell^m)=0$.

(iii). The spherical harmonics is normalized, i.e., $\int d\Omega|Y_\ell^m(\theta,\varphi)|^2=1$, where $d\Omega=\sin\theta d\theta d\varphi$.

First, for $\ell=0$ case, there is only one spherical harmonic, which is the homogeneous polynomial of order zero, i.e., a constant function. The corresponding spherical harmonic is

$$Y_0^0(\theta,\varphi)=\frac{1}{\sqrt{4\pi}}\,. \tag{4.38}$$

Next, for $\ell=1$ case, there are three spherical harmonics corresponding to $m=0,\pm1$. There are three different homogeneous polynomials of order one corresponding to the three spherical harmonics. Indeed,

$$Y_1^{\pm1}(\theta,\varphi)\propto\hat{x}_\pm\,,\quad Y_1^0(\theta,\varphi)\propto\hat{x}_3\,. \tag{4.39}$$

Here, $\hat{x}_\pm$ and $\hat{x}_3$ are the normalized Cartesian coordinates on the unit sphere, i.e.,

$$\hat{x}_\pm=\sin\theta e^{\pm i\varphi},\quad \hat{x}_3=\cos\theta\,. \tag{4.40}$$

Thus, once the normalization is fixed, the spherical harmonics for $\ell=1$ are given by

$$Y_1^{\pm1}(\theta,\varphi)=\mp\sqrt{\frac{3}{8\pi}}\sin\theta e^{\pm i\varphi}\,, \tag{4.41}$$

$$Y_1^0(\theta,\varphi)=\sqrt{\frac{3}{4\pi}}\cos\theta\,. \tag{4.42}$$

The sign convention for Y_ℓ^m is chosen $(-1)^m$ for $m>0$, which is known as the Condon-Shortley phase: it is convenient to obtain the all spherical harmonics using the ladder operators $L_\pm=L_1\pm iL_2$.

Finally, we consider the case with $\ell = 2$. There are six homogeneous polynomials, but they are not all independent due to the Laplace equation constraint. Indeed,

$$Y_2^2 \propto (\hat{x}_+)^2 \quad Y_2^{-2} \propto (\hat{x}_-)^2 \,, \tag{4.43}$$

$$Y_2^1 \propto \hat{x}_+\hat{x}_3 \,, \quad Y_2^{-1} \propto \hat{x}_-\hat{x}_3 \,, \tag{4.44}$$

$$Y_2^0 \propto A\hat{x}_+\hat{x}_- + B\hat{x}_3^2 \,. \tag{4.45}$$

As for the spherical harmonics $m = \pm 1\,, \pm_2$, there is single polynomial for each m, but there are two for $m = 0$. A and B are undetermined constants, but there is a constraint from Laplacian. Once we solve $\nabla^2(r^2 Y_2^0) = 0$, we obtain the relation between A and B as $2A + B = 0$. Therefore, we find $Y_2^0 \propto \hat{x}_+\hat{x}_- - 2\hat{x}_3^2$. Once taking a proper normalization into account, we obtain the explicit forms of Y_2^m as follows.

$$Y_2^{\pm 2}(\theta, \varphi) = \sqrt{\frac{15}{32\pi}} \sin^2\theta e^{\pm 2i\varphi} \,, \tag{4.46}$$

$$Y_2^{\pm 1}(\theta, \varphi) = \mp\sqrt{\frac{8}{32\pi}} \sin\theta\cos\theta e^{\pm i\varphi} \,, \tag{4.47}$$

$$Y_2^0(\theta, \varphi) = \sqrt{\frac{5}{16\pi}} \left(3\cos^2\theta - 1\right) \,. \tag{4.48}$$

For later reference, we need to summarize several properties. One is the orthonormal relation of the spherical harmonics:

$$\int d\Omega \left[Y_\ell^m(\theta, \varphi)\right]^* Y_{\ell'}^{m'}(\theta, \varphi) = \delta_{\ell\ell'}\delta_{mm'} \,. \tag{4.49}$$

It is easy to prove the orthonormality of the spherical harmonics with the normalization above. Once we ignore the overall normalization factor, the spherical harmonics can be expressed as $Y_\ell^m(\theta, \varphi) \propto P_\ell^{|m|}(\theta)e^{im\varphi}$. Thus, the integration over φ gives the orthonormal condition for the state vectors with different m values.

$$\int_0^{2\pi} d\varphi e^{-i(m-m')\varphi} = 2\pi\delta_{mm'} \,. \tag{4.50}$$

Then, we can proceed to the integration over θ, which gives the orthonormal condition for the state vectors with different ℓ values, as follows. Since the φ integration gives the orthonormal condition for the m quantum numbers, we only need to consider the θ integration to establish the orthonormality for the ℓ quantum numbers, so let us set $m' = m$ for the θ integration. By considering the following combinations and using the associated Legendre differential equation, we can find

$$\left[\ell(\ell+1) - \ell'(\ell'+1)\right] \int_{-1}^{1} d\cos\theta P_\ell^{|m|}(\cos\theta) P_{\ell'}^{|m|}(\cos\theta) \tag{4.51}$$

$$= -\left[\sin\theta \frac{dP_\ell^{|m|}(\cos\theta)}{d\theta} P_{\ell'}^{|m|}(\cos\theta) - \sin\theta P_\ell^{|m|}(\cos\theta) \frac{dP_{\ell'}^{|m|}(\cos\theta)}{d\theta}\right]_0^\pi = 0 \tag{4.52}$$

In the last equation, the boundary terms vanish because the associated Legendre functions $P_\ell^{|m|}(\cos\theta)$ are regular at the boundaries $\theta = 0$ and $\theta = \pi$ and the sine function $\sin\theta$ goes to zero at these boundaries. Thus, the integral vanishes for $\ell \neq \ell'$, establishing the orthonormality of the associated Legendre functions for different ℓ values.

$$\int_{-1}^{1} d\cos\theta P_\ell^{|m|}(\cos\theta) P_{\ell'}^{|m|}(\cos\theta) = 0 \quad (\ell' \neq \ell)\,, \\ \neq 0 \quad (\ell' = \ell)\,. \tag{4.53}$$

Therefore, taking the proper normalization into account, we can write the orthonormality condition for the associated Legendre functions as

$$\int_{-1}^{1} d\cos\theta P_\ell^{|m|}(\cos\theta) P_{\ell'}^{|m|}(\cos\theta) = \delta_{\ell',\ell}\,. \tag{4.54}$$

In the last of this subsection, we consider the parity property of the spherical harmonics. In the Cartesian coordinates, the parity transformation corresponds to $\hat{\boldsymbol{x}} \to -\hat{\boldsymbol{x}}$. In terms of the spherical coordinates, the parity transformation corresponds to $(r, \theta, \varphi) \to (r, \pi - \theta, \pi + \varphi)$. Since Y_ℓ^m are homogeneous polynomials with rank ℓ of unit vector $\hat{\boldsymbol{x}}$, under the parity transformation $\hat{\boldsymbol{x}} \to -\hat{\boldsymbol{x}}$, they acquire a factor of $(-1)^\ell$.

$$Y_\ell^m(\pi - \theta, \pi + \varphi) = (-1)^\ell Y_\ell^m(\theta, \varphi)\,. \tag{4.55}$$