Chapter 2

Axioms and Tools (Part A)

This chapter covers the mathematical foundations of the fundamental concepts in quantum mechanics. We will introduce and discuss the axioms, as formulated by Paul Dirac and John von Neumann. These axioms establish the mathematical framework of quantum mechanics, which is primarily based on linear algebra. A solid understanding of these concepts is essential for solving quantum mechanical problems throughout the rest of the course.

2.1 Axiom 1: State Vectors

In quantum mechanics, a physical state—for example, the spin of an electron—is represented by a state vector in a complex vector space. The dimensionality of this vector space, in general, is unrelated to the physical dimension of the system under exam (say, the familiar 3-dimensional space of classical mechanics coordinates, for example) but instead is an abstract space. In Stern-Gerlach–type experiments where the only quantum-mechanical degree of freedom is the spin of an atom, the dimensionality is determined by the number of alternative paths the atoms can follow when subjected to a SG apparatus; in the case of the silver atoms of the previous section, the dimensionality is just two, corresponding to the two possible values s_z can assume. Later, we will consider the case of continuous degrees of freedom—for example, the position (coordinate) or momentum of a particle—where the number of alternatives is infinite, in which case the vector space in question is known as a Hilbert space.

Following Dirac, we call a vector in this space a ket and denote it by $|\psi\rangle$. This state ket is postulated to contain complete information about the physical state; everything we are allowed to ask about the state is contained in the ket. For finite-dimensional spaces, quantum states obey familiar linear algebra properties.

We will now introduce the mathematical space in which kets "live", the Hilbert space. A Hilbert space is a vector space over the complex numbers $\mathbb{C}$. Vectors in this space are conventionally called kets and denoted by $|\psi\rangle$. The dual of a ket is called a bra and is denoted by $\langle\psi|$. In the following we denote the dual correspondence with $\leftrightarrow$, and the following relations are postulated

$$|\psi\rangle \leftrightarrow \langle\psi| \ , \tag{2.1.1}$$

thus kets correspond to bras. When a constant c multiplies a ket, then the dual correspondence is the bra times the complex conjugate of the constant c^*

$$c\,|\psi\rangle \leftrightarrow \langle\psi|\,c^*\,. \tag{2.1.2}$$

For finite-dimensional spaces, kets correspond to column vectors ($n \times 1$ matrices), given some representation. Bras correspond to row vectors ($1 \times n$ matrices).

In this Hilbert space, we can define the inner product $\langle\psi|\phi\rangle$ the maps an ordered pair of vectors (in this case to a complex number), and that has the properties:

(a) Positivity: $\langle\psi|\psi\rangle > 0$ for $|\psi\rangle \neq 0$

(b) Linearity: $\langle\phi|\,(a\,|\psi_1\rangle + b\,|\psi_2\rangle) = a(\langle\phi|\psi_1\rangle + b\,\langle\phi|\psi_2\rangle$

(c) Symmetry: $\langle\phi|\psi\rangle = \langle\psi|\phi\rangle^*$, where the star symbol denotes complex conjugation.

Notice again that all these three properties are very natural for complex vectors, once the inner product is identified with the dot product $\langle\psi|\phi\rangle = \vec{\psi}^\dagger \cdot \vec{\phi}$, where the symbol $\dagger$ corresponds to conjugate transpose of the vector.

A quantum state can be consider a ray, which is an equivalence class of vectors that differ by multiplication by a nonzero complex scalar. For any nonzero ray, we can by convention choose a representative of the class, denoted $|\psi\rangle$, that has unit norm:

$$\langle\psi|\psi\rangle = 1 \tag{2.1.3}$$

and all other states obtained multiplying this state by an arbitrary non-zero constant represent the same physical state:

$$c\,|\psi\rangle \equiv |\psi\rangle \tag{2.1.4}$$

Since every ray corresponds to a possible state, given two states $|\phi\rangle\,, |\psi\rangle$, another state can be constructed as the linear superposition of the two:

$$|\psi'\rangle = a\,|\phi\rangle + b\,|\psi\rangle\ . \tag{2.1.5}$$

Notice that the global phase of the state is irrelevant, thus $e^{i\alpha}\,|\psi'\rangle \equiv |\psi'\rangle$. The relative phase in this superposition is however physically significant. For example, the state $a\,|\phi\rangle + b\,|\psi\rangle$ is the same ray as $e^{i\alpha}(a\,|\phi\rangle + b\,|\psi\rangle)$ but it is different from $(a\,|\phi\rangle + e^{i\alpha}b\,|\psi\rangle)$.

2.2 Operators

Operators are the natural companion of state vectors. They are the tool used to do all physically meaningful manipulations of a quantum state. In the following, we denote operators with $\hat{A}$, to distinguish them from scalars. An operator acts on a ket from the left,

$$\hat{A}\,|\psi\rangle = |\psi'\rangle \tag{2.2.1}$$

resulting into another ket $|\psi'\rangle$. In the familiar case of finite-dimensional vector spaces, operators are nothing but matrices acting on vectors in some representation. Thus, the action of a matrix onto a vector results into another vector. In the more general case of Hilbert spaces, observables are linear maps taking vectors to vectors:

$$\hat{A} : |\psi\rangle \mapsto \hat{A}\,|\psi\rangle \tag{2.2.2}$$

In general when an operator acts on a ket, it produces a distinct ket. However, there are special cases in which the application of an operator leads to a constant times the initial ket. Those are known as eigen-kets $|A_n\rangle$ and have the following property

$$\begin{aligned}
\hat{A}\,|A_1\rangle &= a_1\,|A_1\rangle \\
\hat{A}\,|A_2\rangle &= a_2\,|A_2\rangle \\
\hat{A}\,|A_3\rangle &= a_3\,|A_3\rangle \\
&\dots \\
\hat{A}\,|A_n\rangle &= a_n\,|A_n\rangle\ ,
\end{aligned}$$

where the action of the operator is to return the same ket multiplied by a scalar a_n, which are complex-valued in general. Those are eigenvalues of the operator $\hat{A}$, and the corresponding states $|A_1\rangle, |A_2\rangle, |A_3\rangle, \dots, |A_n\rangle$ are the eigen-kets. Again, for finite-dimensional spaces the notion of eigen-ket is strictly equivalent to that of eigen-vectors in linear algebra.

2.2.1 Adjoints and Hermitian Operators

The adjoint of the operator is denoted as $\hat{A}^\dagger$ and is defined by the dual relationship

$$\hat{A}|\psi\rangle \quad \leftrightarrow \quad \langle\psi|\hat{A}^\dagger|, \tag{2.2.3}$$

for all states $|\psi\rangle$. A special class of operators is that of Hermitian operators $\hat{A}$, for which $\hat{A} = \hat{A}^\dagger$. As we will clarify in the following, Hermitian operators in quantum physics play an important role in the measurement process, as they represent observable quantities.

For a Hermitian operator, the following property is satisified:

$$\langle\phi|\hat{A}|\psi\rangle = \langle\psi|\hat{A}|\phi\rangle^\star\,.$$

We show this by using the symmetry property of the inner product $\langle\phi|\hat{A}|\psi\rangle = \langle\phi|\hat{A}\psi\rangle = \langle\psi|\hat{A}|\phi\rangle^\star$. Then, by the definition of adjoint, we have that the dual of $|\hat{A}\psi\rangle \leftrightarrow \langle\psi|\hat{A}^\dagger = \langle\psi|\hat{A}|$, thus $\langle\phi|\hat{A}|\psi\rangle = \langle\psi|\hat{A}^\dagger|\phi\rangle^\star$. If $\hat{A} = \hat{A}^\dagger$ then $\langle\phi|\hat{A}|\psi\rangle = \langle\psi|\hat{A}|\phi\rangle^\star$.

Another important aspect of Hermitian operators is their spectrum. In particular, the eigenvalues of a Hermitian operator $\hat{A}$ are real valued. We also find that the eigen-kets of $\hat{A}$ corresponding to different eigenvalues are orthogonal. We prove these two claims together. Consider the eigensystem of $\hat{A}$

$$\hat{A}|A_i\rangle = a_i|A_i\rangle, \tag{2.2.4}$$

and its dual equivalent

$$\langle A_j|\hat{A}^\dagger = \langle A_j|a_j^\star = \langle A_j|\hat{A}\,. \tag{2.2.5}$$

In the second equality, we used the fact that the operator is Hermitian. Multiplying Eq. (2.2.4) on the left by $\langle A_j|$ and Eq. (2.2.5) on the right by $|A_i\rangle$, and subtract the results we get

$$\langle A_j|A_i\rangle(a_i - a_j^\star) = 0. \tag{2.2.6}$$

When $i = j$ (the two eigenvectors represent the same state), we must have $a_i = a_i^\star$, hence the eigenvalues are real. In the case $i \neq j$, we have that $\langle A_j|A_i\rangle = 0$, under the assumption that the eigenvalues are real and different $(a_i - a_j) \neq 0$. The two eigenvectors are orthogonal.

In the following we will assume that the eigen-kets of operators are taken to be orthonormal, i.e. satisfy the condition

$$\langle A_j|A_i\rangle = \delta_{ij}. \tag{2.2.7}$$

Notice that this condition can always be enforced, because of the orthogonality condition of the previous theorem and because, as we have seen before, the normalization of each $|A_i\rangle$ is arbitrary (we can thus take it to be 1).

2.2.2 Representing state kets with eigen-kets

Choosing an operator $\hat{A}$ and forming its eigen-kets is in general a very important conceptual and practical step needed to represent arbitrary ket states. We will use eigen-kets of operators as base kets to expand arbitrary kets, as much as for an Euclidean space one uses orthogonal unit vectors (coordinates) to represent an arbitrary vector. We provide an example for the latter. Consider an arbitrary vector $\vec{R}$. We may choose to represent it in cartesian coordinates as follows

$$\vec{R} = x\hat{x} + y\hat{y} + z\hat{z} \tag{2.2.8}$$

where $\hat{x}_i$ represent the basis and x, y, z are the projections to this basis. We may choose to reprepsent this vector in an alternative basis, eg the spherical coordinates basis:

$$\vec{R} = R\hat{r} \tag{2.2.9}$$

Both provide valid representations for the vector $\vec{R}$, however typically one of them might be more convenient to use. We will demonstrate this later.

Let us now go back to the arbitrary vector space.A ket $|\psi\rangle$ can be represented as a linear combination of eigen-kets of some operator $\hat{A}$ in this way:

$$|\psi\rangle = \sum_i c_i\,|a_i\rangle\ , \tag{2.2.10}$$

where the complex-valued coefficients c_i are to be determined. Multiplying the above equation with $\langle a_j|$ and using the orthonormality condition $\langle a_i|a_j\rangle = \delta_{ij}$, we obtain:

$$c_i = \langle a_i|\psi\rangle \ . \tag{2.2.11}$$

The expansion coefficients c_i (often referred to as amplitudes of the state on the eigen-kets of $\hat{A}$) are then formally computed as an inner product of a bra $\langle a_j|$ with the state ket $|\psi\rangle$. We also notice that the normalization of the state reads:

$$\langle\psi|\psi\rangle = \sum_{ij} \langle A_j|c_j^* c_i|A_i\rangle = \sum_i |a_i|^2 \tag{2.2.12}$$

thus an equivalent condition for the state to be normalized is that $\langle\psi|\psi\rangle = \sum_i |c_i|^2 = 1$. We can also re-write the expansion of as

$$|\psi\rangle = \sum_i |A_i\rangle \langle A_i|\psi\rangle \tag{2.2.13}$$

From the above result, we deduce a very important relationship called completeness relation or closure and reads

$$\sum_i |A_i\rangle \langle A_i| = \hat{\mathbb{1}} \tag{2.2.14}$$

This relation holds for arbitrary $|\psi\rangle$. Interestingly, the operator $\hat{P}_i = |A_i\rangle \langle A_i|$ is called the projection operator, and it acts on the state ket $|\psi\rangle$. It projects an arbitrary state vector onto an eigenket i of $\hat{A}$.

Exercise 1.1 Show that the projection operator is idempotent, meaning when operated by itself it remains unchanged, i.e. $\hat{P}^2 = \hat{P}$.

2.2.3 Matrix Representation

The projection operator previously introduced is also important to highlight the direct connection between operators and matrices. Specifically, given an arbitrary operator $\hat{B}$, we can insert the completeness relation twice:

$$\hat{B} = \sum_i |A_i\rangle \langle A_i| \hat{B} = \sum_{ij} |A_i\rangle \langle A_i| \hat{B} |A_j\rangle \langle A_j| \tag{2.2.15}$$

and identify $B(i,j) \equiv \left\langle A_i|\hat{B}|A_j \middle| A_i|\hat{B}|A_j \right\rangle$ as the *matrix element* of the operator $\hat{B}$ in the basis of the eigenkets i and j of $\hat{A}$. We can then explicitly write $\hat{B}$ as a $N \times N$ matrix with elements

$$\hat{B} = \begin{pmatrix} \langle A_1|\hat{B}|A_1\rangle & \langle A_1|\hat{B}|A_2\rangle & \cdots & \langle A_1|\hat{B}|A_n\rangle \\ \langle A_2|\hat{B}|A_1\rangle & \langle A_2|\hat{B}|A_2\rangle & \cdots & \langle A_2|\hat{B}|A_n\rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle A_n|\hat{B}|A_1\rangle & \langle A_n|\hat{B}|A_2\rangle & \cdots & \langle A_n|\hat{B}|A_n\rangle \end{pmatrix}, \tag{2.2.16}$$

where we have assumed that the dimensionality of the vector space is finite and equal to n. For example, for our spin $\frac{1}{2}$ case, we have that the dimensionality is $n = 2$.

This explicit matrix representation also allows to better clarify the concept of Hermitian conjugate (adjoint) operator $\hat{A}^\dagger$, since the result of $\langle\phi|\hat{B}|\psi\rangle = \langle\psi|\hat{B}|\phi\rangle^\star$ is true if $\hat{B}$ is an Hermitian operator) directly translates into a condition for the matrix elements

$$B_{ij} = B^\star_{ji}\,, \tag{2.2.17}$$

thus for finite-dimensional Hilbert spaces an Hermitian operator is nothing but an Hermitian matrix.

Further notice that the representation of the operator $\hat{A}$ itself in its eigen-ket basis is nothing but a diagonal matrix, whose elements are the eigenvalues, i.e., $A_{ij} = \delta_{ij} a_i$. It also immediately follows that $\hat{A} = \sum_i a_i |A_i\rangle \langle A_i|$.

To end this section, we will demonstrate with an example how Dirac's notation is a powerful way of writing expressions involving linear operators. Consider the case of an operator $\hat{D} = \hat{B}\hat{C}$. The matrix elements, expressed in the eigenbasis of $\hat{A}$, can be obtained using again the completeness relation:

$$\begin{aligned} D_{ij} =& \langle A_i|\hat{B}\hat{C}|A_j\rangle = \\ =& \langle A_i|\hat{B}\hat{\mathbb{1}}\hat{C}|A_j\rangle = \\ =& \langle A_i|\,\hat{B}\Big(\sum_k |A_k\rangle \langle A_k|\Big)\hat{C}\,|A_j\rangle = \\ =& \sum_k \langle A_i|\,\hat{B}\,|A_k\rangle \langle A_k|\,\hat{C}\,|A_j\rangle = \\ =& \sum_k B_{ik} C_{k,j}\,, \end{aligned} \tag{2.2.18}$$

which correspond to the usual notion of matrix multiplication.

2.2.4 Finding eigen-kets

The explicit representation of operators in terms of matrix elements is also very useful to explicitly find eigen-kets and eigenvalues, given a certain operator. Let us consider again the matrix elements of the operator $\hat{B}$ in the $\hat{A}$ basis, namely: $B_{ij} \equiv \langle A_i|\hat{B}|A_j\rangle$. The eigenvalue equation is

$$\hat{B}\,|B_k\rangle = b_k\,|B_k\rangle\;, \tag{2.2.19}$$

for the unknown b_k and $|B_k\rangle$. We can rewrite this as

$$\begin{aligned} &\hat{B}\sum_j |A_j\rangle \langle A_j|B_k\rangle = b_k\,|B_k\rangle \\ \Rightarrow\; &\langle A_i|\,\hat{B}\sum_j |A_j\rangle \langle A_j|B_k\rangle = b_k \langle A_i|B_k\rangle \\ \Rightarrow\; &\sum_j \langle A_i|\hat{B}|A_j\rangle \langle A_j|B_k\rangle = b_k \langle A_i|B_k\rangle \end{aligned}$$

which in matrix notation is

$$\begin{pmatrix} \langle A_1|\hat{B}|A_1\rangle & \cdots & \langle A_1|\hat{B}|A_n\rangle \\ \vdots & \ddots & \vdots \\ \langle A_n|\hat{B}|A_1\rangle & \cdots & \langle A_n|\hat{B}|A_n\rangle \end{pmatrix} \begin{pmatrix} \langle A_1|B_k\rangle \\ \vdots \\ \langle A_n|B_k\rangle \end{pmatrix} = b_k \begin{pmatrix} \langle A_1|B_k\rangle \\ \vdots \\ \langle A_n|B_k\rangle \end{pmatrix}\,. \tag{2.2.20}$$

The eigenvalues are found, as in standard linear algebra, as solutions of the characteristic equation

$$\det\left(\hat{M} - b_k \hat{I}\right) = 0\,, \tag{2.2.21}$$

and once the b_k are found, we solve the homogeneous linear system

$$(\hat{M} - b_k \hat{I})\vec{v}^k = 0\,, \tag{2.2.22}$$

for the unknown vectors $\vec{v}^k$. This procedure is found in any standard linear algebra book.

Exercise 1.2 Diagonalize the following matrices:

$$B_1 = \begin{pmatrix} 2 & 1 \\ 1 & 4 \end{pmatrix}, \quad B_2 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

Exercise 1.3 Consider the matrix

$$\Omega = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

(a) Is it Hermitian?

(b) Find its eigenvalues and eigenvectors.

(c) Construct the matrix U, where its columns correspond to the eigenvectors you found earlier.

(d) Show that $U^\dagger \Omega U$ is diagonal. What are the diagonal elements? Are they familiar to you?

Exercise 1.4 Consider the matrix

$$\Omega = \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix},$$

where α is a real number.

(a) Show that it is unitary.

(b) Show that the eigenvalues are $e^{\pm i\alpha}$.

(c) Find the corresponding eigenvectors, and show that they are orthogonal.
