1 Basic formalism

We summarize below some of the postulates and definitions basic to our formalism, and present some important results based on these postulates. The formalism is purely mathematical in nature with very little physics input, but it provides the structure within which the physical concepts that will be discussed in the later chapters will be framed.

1.1 State vectors

It is important to realize that the Quantum Theory is a linear theory in which the physical state of a system is described by a vector in a complex, linear vector space. This vector may represent a free particle or a particle bound in an atom or a particle interacting with other particles or with external fields. It is much like a vector in ordinary three-dimensional space, following many of the same rules, except that it describes a very complicated physical system. We will be elaborating further on this in the following.

The mathematical structure of a quantum mechanical system will be presented in terms of the notations developed by Dirac.

A physical state in this notation is described by a "ket" vector, $|\rangle$, designated variously as $|\alpha\rangle$ or $|\psi\rangle$ or a ket with other appropriate symbols depending on the specific problem at hand. The kets can be complex. Their complex conjugates, $|\rangle^*$, are designated by $\langle|$ which are called "bra" vectors. Thus, corresponding to every ket vector there is a bra vector. These vectors are abstract quantities whose physical interpretation is derived through their so-called "representatives" in the coordinate or momentum space or in a space appropriate to the problem under consideration.

The dimensionality of the vector space is left open for the moment. It can be finite, as will be the case when we encounter spin, which has a finite number of components along a preferred direction, or it can be infinite, as is the case of the discrete bound states of the hydrogen atom. Or, the dimensionality could be continuous (indenumerable) infinity, as for a free particle with momentum that takes continuous values. A complex vector space with these properties is called a Hilbert space.

The kets have the same properties as a vector in a linear vector space. Some of the most important of these properties are given below:

(i) $|\alpha\rangle$ and $c\,|\alpha\rangle$, where c is a complex number, describe the same state.
(ii) The bra vector corresponding to $c\,|\alpha\rangle$ will be $c^*\,\langle\alpha|$.

(iii) The kets follow a linear superposition principle

$$a\,|\alpha\rangle + b\,|\beta\rangle = c\,|\gamma\rangle \tag{1.1}$$

where a, b, and c are complex numbers. That is, a linear combination of states in a vector space is also a state in the same space.

(iv) The "scalar product" or "inner product" of two states $|\alpha\rangle$ and $|\beta\rangle$ is defined as

$$\langle\beta|\,\alpha\rangle. \tag{1.2}$$

It is a complex number and not a vector. Its complex conjugate is given by

$$\langle\beta|\,\alpha\rangle^* = \langle\alpha|\,\beta\rangle. \tag{1.3}$$

Hence $\langle\alpha|\,\alpha\rangle$ is a real number.

(v) Two states $|\alpha\rangle$ and $|\beta\rangle$ are orthogonal if

$$\langle\beta|\,\alpha\rangle = 0. \tag{1.4}$$

(vi) It is postulated that $\langle\alpha|\,\alpha\rangle \geq 0$. One calls $\sqrt{\langle\alpha|\,\alpha\rangle}$ the "norm" of the state $|\alpha\rangle$. If a state vector is normalized to unity then

$$\langle\alpha|\,\alpha\rangle = 1. \tag{1.5}$$

If the norm vanishes, then $|\alpha\rangle = 0$, in which case $|\alpha\rangle$ is called a null vector.

(vii) The states $|\alpha_n\rangle$ with $n = 1, 2, \ldots,$ depending on the dimensionality, are called a set of basis kets or basis states if they span the linear vector space. That is, any arbitrary state in this space can be expressed as a linear combination (superposition) of them. The basis states are often taken to be of unit norm and orthogonal, in which case they are called orthonormal states. Hence an arbitrary state $|\phi\rangle$ can be expressed in terms of the basis states $|\alpha_n\rangle$ as

$$|\phi\rangle = \sum_n a_n\,|\alpha_n\rangle \tag{1.6}$$

where, as stated earlier, the values taken by the index n depends on whether the space is finite- or infinite-dimensional or continuous. In the latter case the summation is replaced by an integral. If the $|\alpha_n\rangle$'s are orthonormal then $a_n = \langle\alpha_n\,|\phi\rangle$. It is then postulated that $|a_n|^2$ is the probability that the state $|\phi\rangle$ will be in state $|\alpha_n\rangle$.

(viii) A state vector may depend on time, in which case one writes it as $|\alpha(t)\rangle$, $|\psi(t)\rangle$, etc. In the following, except when necessary, we will suppress the possible dependence on time.

(ix) The product $|\alpha\rangle\,|\beta\rangle$, has no meaning unless it refers to two different vector spaces, e.g., one corresponding to spin, the other to momentum; or, if a state consists of two particles described by $|\alpha\rangle$ and $|\beta\rangle$ respectively.

(x) Since bra vectors are obtained through complex conjugation of the ket vectors, the above properties can be easily carried over to the bra vectors.

1.2 Operators and physical observables

A physical observable, like energy or momentum, is described by a linear operator that has the following properties:

(i) If A is an operator and $|\alpha\rangle$ is a ket vector then

$$A\,|\alpha\rangle = \text{another ket vector.} \tag{1.7}$$

Similarly, for an operator B,

$$\langle\alpha|\,B = \text{another bra vector} \tag{1.8}$$

where B operates to the left

(ii) An operator A is linear if, for example,

$$A\,[\lambda\,|\alpha\rangle + \mu\,|\beta\rangle] = \lambda A\,|\alpha\rangle + \mu A\,|\beta\rangle \tag{1.9}$$

where λ and μ are complex numbers. Typical examples of linear operators are derivatives, matrices, etc. There is one exception to this rule, which we will come across in Chapter 27 which involves the so called time reversal operator where the coefficients on the right-hand side are replaced by their complex conjugates. In this case it is called an antilinear operator.

If an operator acting on a function gives rise to the square of that function, for example, then it is called a nonlinear operator. In this book we will be not be dealing with such operators.

(iii) A is a called a unit operator if, for any $|\alpha\rangle$,

$$A\,|\alpha\rangle = |\alpha\rangle, \tag{1.10}$$

in which case one writes

$$A = \mathbf{1}. \tag{1.11}$$

(iv) A product of two operators is also an operator. In other words, if A and B are operators then AB as well as BA are operators. However, they do not necessarily commute under multiplication, that is,

$$AB \neq BA \tag{1.12}$$

in general. The operators commute under addition, i.e., if A and B are two operators then

$$A + B = B + A. \tag{1.13}$$

They also exhibit associativity, i.e., if A, B, and C are three operators then

$$A + (B + C) = (A + B) + C. \tag{1.14}$$

Similarly $A\,(BC) = (AB)\,C$.

(v) B is called an inverse of the operator A if

$$AB = BA = \mathbf{1}, \tag{1.15}$$

in which case one writes

$$B = A^{-1}. \tag{1.16}$$

(vi) The quantity $|\alpha\rangle\,\langle\beta|$ is called the "outer product" between states $|\alpha\rangle$ and $|\beta\rangle$. By multiplying it with a state $|\gamma\rangle$ one obtains

$$[|\alpha\rangle\,\langle\beta|]\,\gamma\rangle = [\langle\beta|\,\gamma\rangle]\,|\alpha\rangle \tag{1.17}$$

where on the right-hand side the order of the terms is reversed since $\langle\beta|\,\gamma\rangle$ is a number. The above relation implies that when $|\alpha\rangle\,\langle\beta|$ multiplies with a state vector it gives another state vector. A similar result holds for the bra vectors:

$$\langle\gamma\,[|\alpha\rangle\,\langle\beta|] = [\langle\gamma\,|\alpha\rangle]\,\langle\beta|\,. \tag{1.18}$$

Thus $|\alpha\rangle\,\langle\beta|$ acts as an operator.

(vii) The "expectation" value, $\langle A\rangle$, of an operator A in the state $|\alpha\rangle$ is defined as

$$\langle A\rangle = \langle\alpha|\,A\,|\alpha\rangle\,. \tag{1.19}$$

1.3 Eigenstates

(i) If the operation $A\,|\alpha\rangle$ gives rise to the same state vector, i.e., if

$$A\,|\alpha\rangle = (\text{constant}) \times |\alpha\rangle \tag{1.20}$$

then we call $|\alpha\rangle$ an "eigenstate" of the operator A, and the constant is called the "eigenvalue" of A. If $|\alpha\rangle$'s are eigenstates of A with eigenvalues a_n, assumed for convenience to be discrete, then these states are generally designated as $|a_n\rangle$. They satisfy the equation

$$A\,|a_n\rangle = a_n\,|a_n\rangle \tag{1.21}$$

with $n = 1, 2, \ldots$ depending on the dimensionality of the system. In this case one may also call A an eigenoperator.

(ii) If $|\alpha_n\rangle$ is an eigenstate of both operators A and B, such that

$$A\,|\alpha_n\rangle = a_n\,|\alpha_n\rangle\,, \quad \text{and} \quad B\,|\alpha_n\rangle = b_n\,|\alpha_n\rangle \tag{1.22}$$

then we have the results

$$AB\left|\alpha_n\right\rangle = b_n A\left|\alpha_n\right\rangle = b_n a_n\left|\alpha_n\right\rangle, \tag{1.23}$$

$$BA\left|\alpha_n\right\rangle = a_n B\left|\alpha_n\right\rangle = a_n b_n\left|\alpha_n\right\rangle. \tag{1.24}$$

If the above two relations hold for all values of n then

$$AB = BA. \tag{1.25}$$

Thus under the special conditions just outlined the two operators will commute.

1.4 Hermitian conjugation and Hermitian operators

We now define the "Hermitian conjugate" $A^\dagger$, of an operator A and discuss a particular class of operators called "Hermitian" operators which play a central role in quantum mechanics.

(i) In the same manner as we defined the complex conjugate operation for the state vectors, we define $A^\dagger$ through the following complex conjugation

$$[A\left|\alpha\right\rangle]^* = \left\langle\alpha\right| A^\dagger \tag{1.26}$$

and

$$\left\langle\beta\right| A\left|\alpha\right\rangle^* = \left\langle\alpha\right| A^\dagger\left|\beta\right\rangle. \tag{1.27}$$

If on the left-hand side of (1.26), $\left|\alpha\right\rangle$ is replaced by $c\left|\alpha\right\rangle$ where c is a complex constant, then on the right-hand side one must include a factor c^*.

(ii) From (1.26) and (1.27) it follows that if

$$A = \left|\alpha\right\rangle\left\langle\beta\right| \tag{1.28}$$

then

$$A^\dagger = \left|\beta\right\rangle\left\langle\alpha\right|. \tag{1.29}$$

At this stage it is important to emphasize that $\left|\alpha\right\rangle$, $\left\langle\beta\middle|\alpha\right\rangle$, $\left|\alpha\right\rangle\left\langle\beta\right|$, and $\left|\alpha\right\rangle\left|\beta\right\rangle$ are four totally different mathematical quantities which should not be mixed up: the first is a state vector, the second is an ordinary number, the third is an operator, and the fourth describes a product of two states.

(iii) The Hermitian conjugate of the product operator AB is found to be

$$(AB)^\dagger = B^\dagger A^\dagger. \tag{1.30}$$

This can be proved by first noting from (1.27) that for an arbitrary state $\left|\alpha\right\rangle$

$$[(AB)\left|\alpha\right\rangle]^* = \left\langle\alpha\right|(AB)^\dagger. \tag{1.31}$$

If we take

$$B\,|\alpha\rangle = |\beta\rangle \tag{1.32}$$

where $|\beta\rangle$ is another state vector, then the left-hand side of (1.31) can be written as

$$[(AB)\,|\alpha\rangle]^* = [A\,|\beta\rangle]^*\,. \tag{1.33}$$

From the definition given in (1.26) we obtain

$$[A\,|\beta\rangle]^* = \langle\beta|\,A^\dagger = \left[\langle\alpha|\,B^\dagger\right]A^\dagger = \langle\alpha|\,B^\dagger A^\dagger \tag{1.34}$$

where we have used the fact that $\langle\beta| = [B\,|\alpha\rangle]^* = \langle\alpha|\,B^\dagger$. Since $|\alpha\rangle$ is an arbitrary vector, comparing (1.31) and (1.34), we obtain (1.30).

(iv) Finally, if

$$A = A^\dagger \tag{1.35}$$

then the operator A is called "Hermitian."

1.5 Hermitian operators: their eigenstates and eigenvalues

Hermitian operators play a central role in quantum mechanics. We show below that the eigenstates of Hermitian operators are orthogonal and have real eigenvalues.

Consider the eigenstates $|a_n\rangle$ of an operator A,

$$A\,|a_n\rangle = a_n\,|a_n\rangle \tag{1.36}$$

where $|a_n\rangle$'s have a unit norm. By multiplying both sides of (1.36) by $\langle a_n|$ we obtain

$$a_n = \langle a_n|\,A\,|a_n\rangle\,. \tag{1.37}$$

Taking the complex conjugate of both sides we find

$$a_n^* = \langle a_n|\,A\,|a_n\rangle^* = \langle a_n|\,A^\dagger\,|a_n\rangle = \langle a_n|\,A\,|a_n\rangle\,. \tag{1.38}$$

The last equality follows from the fact that A is Hermitian ($A^\dagger = A$). Equating (1.37) and (1.38) we conclude that $a_n^* = a_n$. Therefore, the eigenvalues of a Hermitian operator must be real.

An important postulate based on this result says that since physically observable quantities are expected to be real, the operators representing these observables must be Hermitian.

We now show that the eigenstates are orthogonal. We consider two eigenstates $|a_n\rangle$ and $|a_m\rangle$ of A,

$$A\,|a_n\rangle = a_n\,|a_n\rangle\,, \tag{1.39}$$

$$A\,|a_m\rangle = a_m\,|a_m\rangle\,. \tag{1.40}$$

Taking the complex conjugate of the second equation we have

$$\langle a_m|\,A = a_m\,\langle a_m| \tag{1.41}$$

where we have used the Hermitian property of A, and the fact that the eigenvalue a_m is real. Multiplying (1.39) on the left by $\langle a_m|$ and (1.41) on the right by $|a_n\rangle$ and subtracting, we obtain

$$(a_m - a_n)\,\langle a_m|\,a_n\rangle = 0. \tag{1.42}$$

Thus, if the eigenvalues α_m and α_n are different we have

$$\langle a_m|\,a_n\rangle = 0, \tag{1.43}$$

which shows that the two eigenstates are orthogonal. Using the fact that the ket vectors are normalized, we can write the general orthonormality relation between them as

$$\langle a_m|\,a_n\rangle = \delta_{mn} \tag{1.44}$$

where δ_{mn} is called the Kronecker delta, which has the property

$$\begin{aligned}\delta_{mn} &= 1 \quad \text{for} \quad m = n \\ &= 0 \quad \text{for} \quad m \neq n.\end{aligned} \tag{1.45}$$

For those cases where there is a degeneracy in the eigenvalues, i.e., if two different states have the same eigenvalue, the treatment is slightly different and will be deferred until later chapters.

We note that the operators need not be Hermitian in order to have eigenvalues. However, in these cases none of the above results will hold. For example, the eigenvalues will not necessarily be real. Unless otherwise stated, we will assume the eigenvalues to be real.

1.6 Superposition principle

A basic theorem in quantum mechanics based on linear vector algebra is that an arbitrary vector in a given vector space can be expressed as a linear combination – a superposition – of a complete set of eigenstates of any operator in that space. A complete set is defined to be the set of all possible eigenstates of an operator. Expressing this result for an arbitrary state vector $|\phi\rangle$ in terms of the eigenstates $|a_n\rangle$ of the operator A, we have

$$|\phi\rangle = \sum_n c_n |a_n\rangle \tag{1.46}$$

where the summation index n goes over all the eigenstates with $n = 1, 2, \ldots$. If we multiply (1.46) by $\langle a_m|$ then the orthonormality relation (1.44) between the $|a_n\rangle$'s yields

$$c_m = \langle a_m | \phi\rangle. \tag{1.47}$$

It is then postulated that $|c_m|^2$ is the probability that $|\phi\rangle$ contains $|a_m\rangle$. That is, $|c_m|^2$ is the probability that $|\phi\rangle$ has the eigenvalue a_m. If $|\phi\rangle$ is normalized to unity, $\langle\phi|\phi\rangle = 1$, then

$$\sum_n |c_n|^2 = 1. \tag{1.48}$$

That is, the probability of finding $|\phi\rangle$ in state $|a_n\rangle$, summed over all possible values of n, is one.

Since (1.46) is true for any arbitrary state we can express another state $|\psi\rangle$ as

$$|\psi\rangle = \sum_n c'_n |a_n\rangle . \tag{1.49}$$

The scalar product $\langle\psi|\phi\rangle$ can then be written, using the orthonormality property of the eigenstates, as

$$\langle\psi|\phi\rangle = \sum_m c'^*_m c_m \tag{1.50}$$

with $c'_m = \langle a_m|\psi\rangle$ and $c_m = \langle a_m|\phi\rangle$.

The above relations express the fact that the state vectors can be expanded in terms of the eigenstates of an operator A. The eigenstates $|a_n\rangle$ are then natural candidates to form a set of basis states.

1.7 Completeness relation

We consider now the operators $|a_n\rangle\langle a_n|$, where the $|a_n\rangle$'s are the eigenstates of an operator A, with eigenvalues a_n. A very important result in quantum mechanics involving the sum of the operators $|a_n\rangle\langle a_n|$ over the possibly infinite number of eigenvalues states that

$$\sum_n |a_n\rangle\langle a_n| = \mathbf{1} \tag{1.51}$$

where the **1** on the right-hand side is a unit operator. This is the so called "completeness relation".

To prove this relation we first multiply the sum on the left hand of the above equality by an arbitrary eigenvector $|a_m\rangle$ to obtain

$$\left[\sum_n |a_n\rangle \langle a_n|\right] |a_m\rangle = \sum_n |a_n\rangle \langle a_n| a_m\rangle = \sum_n |a_n\rangle \delta_{nm} = |a_m\rangle \tag{1.52}$$

where we have used the orthonormality of the eigenvectors. Since this relation holds for any arbitrary state $|a_m\rangle$, the operator in the square bracket on the left-hand side acts as a unit operator, thus reproducing the completeness relation.

If we designate

$$P_n = |a_n\rangle \langle a_n| \tag{1.53}$$

then

$$P_n |a_m\rangle = \delta_{nm} |a_m\rangle . \tag{1.54}$$

Thus P_n projects out the state $|a_n\rangle$ whenever it operates on an arbitrary state. For this reason P_n is called the projection operator, in terms of which one can write the completeness relation as

$$\sum_n P_n = \mathbf{1}. \tag{1.55}$$

One can utilize the completeness relation to simplify the scalar product $\langle\psi|\phi\rangle$, where $|\phi\rangle$ and $|\psi\rangle$ are given above, if we write, using (1.51)

$$\langle\psi|\phi\rangle = \langle\psi|\mathbf{1}|\phi\rangle = \langle\psi|\left[\sum_n |a_n\rangle \langle a_n|\right]|\phi\rangle = \sum_n \langle\psi| a_n\rangle \langle a_n|\phi\rangle = \sum_n c_n'^{*} c_n. \tag{1.56}$$

This is the same result as the one we derived previously as (1.50).

1.8 Unitary operators

If two state vectors $|\alpha\rangle$ and $|\alpha'\rangle$ have the same norm then

$$\langle\alpha|\alpha\rangle = \langle\alpha'|\alpha'\rangle. \tag{1.57}$$

Expressing each of these states in terms of a complete set of eigenstates $|a_n\rangle$ we obtain

$$|\alpha\rangle = \sum_n c_n |a_n\rangle \quad \text{and} \quad |\alpha'\rangle = \sum_n c_n' |a_n\rangle . \tag{1.58}$$

The equality in (1.57) leads to the relation

$$\sum_n |c_n|^2 = \sum_n |c_n'|^2 , \tag{1.59}$$

which signifies that, even though c_n may be different from c_n', the total sum of the probabilities remains the same.

Consider now an operator A such that

$$A\,|\alpha\rangle = |\alpha'\rangle. \tag{1.60}$$

If $|\alpha\rangle$ and $|\alpha'\rangle$ have the same norm, then

$$\langle\alpha|\alpha\rangle = \langle\alpha'|\alpha'\rangle = \langle\alpha|\,A^\dagger A\,|\alpha\rangle. \tag{1.61}$$

This implies that

$$A^\dagger A = \mathbf{1}. \tag{1.62}$$

The operator A is then said to be "unitary." From relation (1.60) it is clear that A can change the basis from one set to another. Unitary operators play a fundamental role in quantum mechanics.

1.9 Unitary operators as transformation operators

Let us define the following operator in term of the eigenstates $|a_n\rangle$ of operator A and eigenstates $|b_n\rangle$ of operator B,

$$U = \sum_n |b_n\rangle\langle a_n|. \tag{1.63}$$

This is a classic example of a unitary operator as we show below. We first obtain the Hermitian conjugate of U,

$$U^\dagger = \sum_n |a_n\rangle\langle b_n|. \tag{1.64}$$

Therefore,

$$UU^\dagger = \left[\sum_n |b_n\rangle\langle a_n|\right]\sum_m |a_m\rangle\langle b_m| = \sum\sum |b_n\rangle\langle a_n|a_m\rangle\langle b_m| = \sum |b_n\rangle\langle b_n| = \mathbf{1} \tag{1.65}$$

where we have used the orthonormality relation $\langle a_n|a_m\rangle = \delta_{nm}$ and the completeness relation for the state vectors $|b_n\rangle$ discussed in the previous section. Hence U is unitary.

We note in passing that

$$\sum_n |a_n\rangle\langle a_n| \tag{1.66}$$

is a unit operator which is a special case of a unitary operator when $\langle b_n| = \langle a_n|$.