

# CHAPTER 4

## PRINCIPLES OF QUANTUM MECHANICS

In this Chapter we will continue to develop the mathematical formalism of quantum mechanics, using heuristic arguments as necessary. This will lead to a system of postulates which will be the basis of our subsequent applications of quantum mechanics.

### Hermitian Operators

An important property of operators is suggested by considering the Hamiltonian for the particle in a box:

$$\hat{H} = -\frac{\hbar^2}{2m} \frac{d^2}{dx^2} \quad (1)$$

Let  $f(x)$  and  $g(x)$  be arbitrary functions which obey the same boundary values as the eigenfunctions of  $\hat{H}$ , namely that they vanish at  $x = 0$  and  $x = a$ . Consider the integral

$$\int_0^a f(x) \hat{H} g(x) dx = -\frac{\hbar^2}{2m} \int_0^a f(x) g''(x) dx \quad (2)$$

Now, using integration by parts,

$$\int_0^a f(x) g''(x) dx = -\int_0^a f'(x) g'(x) dx + \left[ f(x) g'(x) \right]_0^a \quad (3)$$

The boundary terms vanish by the assumed conditions on  $f$  and  $g$ . A second integration by parts transforms (3) to

$$+\int_0^a f''(x) g(x) dx - \left[ f'(x) g(x) \right]_0^a$$

It follows therefore that

$$\int_0^a f(x) \hat{H} g(x) dx = \int_0^a g(x) \hat{H} f(x) dx \quad (4)$$

An obvious generalization for complex functions will read

$$\int_0^a f^*(x) \hat{H} g(x) dx = \left( \int_0^a g^*(x) \hat{H} f(x) dx \right)^* \quad (5)$$

In mathematical terminology, an operator  $\hat{A}$  for which

$$\int f^* \hat{A} g d\tau = \left( \int g^* \hat{A} f d\tau \right)^* \quad (6)$$

for all functions  $f$  and  $g$  which obey specified boundary conditions is classified as *hermitian* or *self-adjoint*. Evidently, the Hamiltonian is a hermitian operator. It is postulated that *all* quantum-mechanical operators that represent dynamical variables are hermitian.

### Properties of Eigenvalues and Eigenfunctions

The sets of energies and wavefunctions obtained by solving any quantum-mechanical problem can be summarized symbolically as solutions of the eigenvalue equation

$$\hat{H} \psi_n = E_n \psi_n \quad (7)$$

For another value of the quantum number, we can write

$$\hat{H} \psi_m = E_m \psi_m \quad (8)$$

Let us multiply (7) by  $\psi_m^*$  and the complex conjugate of (8) by  $\psi_n$ . Then we subtract the two expressions and integrate over  $d\tau$ . The result is

$$\int \psi_m^* \hat{H} \psi_n d\tau - \left( \int \psi_n^* \hat{H} \psi_m d\tau \right)^* = (E_n - E_m^*) \int \psi_m^* \psi_n d\tau \quad (9)$$

But by the hermitian property (5), the left-hand side of (9) equals zero. Thus

$$(E_n - E_m^*) \int \psi_m^* \psi_n d\tau = 0 \quad (10)$$

Consider first the case  $m = n$ . The second factor in (10) then becomes the normalization integral  $\int \psi_n^* \psi_n d\tau$ , which equals 1 (or at least a nonzero constant). Therefore the first factor in (10) must equal zero, so that

$$E_n^* = E_n \quad (11)$$

implying that the energy eigenvalues must be real numbers. This is quite reasonable from a physical point of view since eigenvalues represent possible results of measurement. Consider next the case when  $E_m \neq E_n$ . Then it is the second factor in (10) that must vanish and

$$\int \psi_m^* \psi_n d\tau = 0 \quad \text{when} \quad E_m \neq E_n \quad (12)$$

Thus eigenfunctions belonging to different eigenvalues are orthogonal. In the case that  $\psi_m$  and  $\psi_n$  are degenerate eigenfunctions, so  $m \neq n$  but  $E_m = E_n$ , the above proof of orthogonality does not apply. But it is always possible to construct degenerate functions that are mutually orthogonal. A general result is therefore the orthonormalization condition

$$\int \psi_m^* \psi_n d\tau = \delta_{mn} \quad (13)$$

It is easy to prove that a linear combination of degenerate eigenfunctions is itself an eigenfunction of the same energy. Let

$$\hat{H} \psi_{nk} = E_n \psi_{nk}, \quad k = 1, 2 \dots d \quad (14)$$

where the  $\psi_{nk}$  represent a  $d$ -fold degenerate set of eigenfunctions with the same eigenvalue  $E_n$ . Consider now the linear combination

$$\psi = c_1 \psi_{n,1} + c_2 \psi_{n,2} + \dots + c_d \psi_{n,d} \quad (15)$$

Operating on  $\psi$  with the Hamiltonian and using (14), we find

$$\begin{aligned} \hat{H}\psi &= c_1 \hat{H}\psi_{n,1} + c_2 \hat{H}\psi_{n,2} + \dots \\ &= E_n (c_1 \psi_{n,1} + c_2 \psi_{n,2} + \dots) = E_n \psi \end{aligned} \quad (16)$$

which shows that the linear combination  $\psi$  is also an eigenfunction of the same energy. There is evidently a limitless number of possible eigenfunctions for a degenerate eigenvalue. However, only  $d$  of these will be linearly independent.

## Dirac Notation [OPTIONAL]

The term *orthogonal* has been used both for perpendicular vectors and for functions whose product integrates to zero. This actually connotes a deep connection between vectors and functions. Consider two orthogonal vectors  $\mathbf{a}$  and  $\mathbf{b}$ . Then, in terms of their  $x, y, z$  components, labelled by 1, 2, 3, respectively, the scalar product can be written

$$\mathbf{a} \cdot \mathbf{b} = a_1 b_1 + a_2 b_2 + a_3 b_3 = 0 \quad (17)$$

Suppose now that we consider an analogous relationship involving vectors in  $n$ -dimensional space (which you need not visualize!). We could then write

$$a \cdot b = \sum_{k=0}^n a_k b_k = 0 \quad (18)$$

Finally let the dimension of the space become nondenumerably infinite, turning into a continuum. The sum (18) would then be replaced by an integral such as

$$\int a(x) b(x) dx = 0 \quad (19)$$

But this is just the relation for orthogonal functions. A function can therefore be regarded as an abstract vector in a higher-dimensional continuum, known as *Hilbert space*. This is true for eigenfunctions as well. Dirac denoted the vector in Hilbert space corresponding to the eigenfunction  $\psi_n$  by the symbol  $|n\rangle$ . Correspondingly, the complex conjugate  $\psi_m^*$  is denoted by  $\langle m|$ . The integral over the product of the two functions is then analogous to a scalar product of the abstract vectors, written

$$\int \psi_m^* \psi_n d\tau = \langle m| \cdot |n\rangle \equiv \langle m|n\rangle \quad (20)$$

The last quantity is known as a *bracket*, which led Dirac to designate the vectors  $\langle m|$  and  $|n\rangle$  as a “bra” and a “ket,” respectively. The orthonormality conditions (13) can be written

$$\langle m|n\rangle = \delta_{mn} \quad (21)$$

The integral of a “sandwich” containing an operator  $\hat{A}$  can be written very compactly in the form

$$\int \psi_m^* \hat{A} \psi_n d\tau = \langle m|A|n\rangle \quad (22)$$

The hermitian condition on  $\hat{A}$  [cf. Eq (6)] is therefore expressed as

$$\langle m|A|n\rangle = \langle n|A|m\rangle^* \quad (23)$$

### Expectation Values

One of the extraordinary features of quantum mechanics is the possibility for superpositions of states. The state of a system can sometimes exist as a linear combination of other states, for example,

$$\psi = c_1\psi_1 + c_2\psi_2 \quad (24)$$

Assuming that all three functions are normalized and that  $\psi_1$  and  $\psi_2$  are orthogonal, we find

$$\int \psi^* \psi d\tau = |c_1|^2 + |c_2|^2 = 1 \quad (25)$$

We can interpret  $|c_1|^2$  and  $|c_2|^2$  as the probabilities that a system in a state described by  $\psi$  can have the attributes of the states  $\psi_1$  and  $\psi_2$ , respectively. Suppose  $\psi_1$  and  $\psi_2$  represent eigenstates of an observable  $A$ , satisfying the respective eigenvalue equations

$$\hat{A}\psi_1 = a_1\psi_1 \quad \text{and} \quad \hat{A}\psi_2 = a_2\psi_2 \quad (26)$$

Then a large number of measurements of the variable  $A$  in the state  $\psi$  will register the value  $a_1$  with a probability  $|c_1|^2$  and the value  $a_2$  with a probability  $|c_2|^2$ . The average value or *expectation value* of  $A$  will be given by

$$\langle A \rangle = |c_1|^2 a_1 + |c_2|^2 a_2 \quad (27)$$

This can be obtained directly from  $\psi$  by the “sandwich construction”

$$\langle A \rangle = \int \psi^* \hat{A} \psi d\tau \quad (28)$$

or, if  $\psi$  is not normalized,

$$\langle A \rangle = \frac{\int \psi^* \hat{A} \psi d\tau}{\int \psi^* \psi d\tau} \quad (29)$$

Note that the expectation value need not itself be a possible result of a single measurement (like the centroid of a donut, which is located in the hole!). When the operator  $\hat{A}$  is a simple function, not containing differential operators or the like, then (28) reduces to the classical formula for an average value:

$$\langle A \rangle = \int A \rho d\tau \quad (30)$$

### More on Operators

An operator represents a prescription for turning one function into another: in symbols,  $\hat{A}\psi = \phi$ . From a physical point of view, the action of an operator on a wavefunction can be pictured as the process of measuring the observable  $A$  on the state  $\psi$ . The transformed wavefunction  $\phi$  then represents the state of the system *after* the measurement is performed. In general  $\phi$  is different from  $\psi$ , consistent with the fact that the process of measurement on a quantum system produces an irreducible perturbation of its state. Only in the special case that  $\psi$  is an eigenstate of  $A$ , does a measurement preserve the original state. The function  $\phi$  is then equal to an eigenvalue  $a$  times  $\psi$ .

The product of two operators, say  $\hat{A}\hat{B}$ , represents the successive action of the operators, reading from *right to left*—ie., first  $\hat{B}$  then  $\hat{A}$ . In general, the action of two operators in the reversed order, say  $\hat{B}\hat{A}$ , gives a different result, which can be written  $\hat{A}\hat{B} \neq \hat{B}\hat{A}$ . We say that the operators do not *commute*. This can be attributed to the perturbing effect one measurement on a quantum system can have on subsequent measurements. An example of non-commuting operators from everyday life. In our usual routine each morning, we shower and we get dressed. But the result of carrying out these operations in reversed order will be dramatically different!

The *commutator* of two operators is defined by

$$[\hat{A}, \hat{B}] \equiv \hat{A}\hat{B} - \hat{B}\hat{A} \quad (31)$$

When  $[\hat{A}, \hat{B}] = 0$ , the two operators are said to *commute*. This means their combined effect will be the same whatever order they are applied (like brushing your teeth and showering).

The uncertainty principle for simultaneous measurement of two observables  $A$  and  $B$  is closely related to their commutator. The uncertainty  $\Delta a$  in the observable  $A$  is defined in terms of the mean square deviation from the average:

$$(\Delta a)^2 = \langle (\hat{A} - \langle A \rangle)^2 \rangle = \langle A^2 \rangle - \langle A \rangle^2 \quad (32)$$

It corresponds to the *standard deviation* in statistics. The following inequality can be proven for the product of two uncertainties:

$$\Delta a \Delta b \geq \frac{1}{2} |\langle [\hat{A}, \hat{B}] \rangle| \quad (33)$$

The best known application of (33) is to the position and momentum operators, say  $\hat{x}$  and  $\hat{p}_x$ . Their commutator is given by

$$[\hat{x}, \hat{p}_x] = i\hbar \quad (34)$$

so that

$$\Delta x \Delta p \geq \hbar/2 \quad (35)$$

which is known as the *Heisenberg uncertainty principle*. This fundamental consequence of quantum theory implies that the position and momentum of a particle cannot be determined with arbitrary precision—the more accurately one is known, the more uncertain is the other. For example, if the momentum is known exactly, as in a momentum eigenstate, then the position is completely undetermined.

If two operators commute, there is no restriction on the accuracy of their simultaneous measurement. For example, the  $x$  and  $y$  coordinates of a particle can be known at the same time. An important theorem states that two commuting observables can have simultaneous eigenfunctions. To prove this, write the eigenvalue equation for an operator  $\hat{A}$

$$\hat{A} \psi_n = a_n \psi_n \quad (36)$$

then operate with  $\hat{B}$  and use the commutativity of  $\hat{A}$  and  $\hat{B}$  to obtain

$$\hat{B} \hat{A} \psi_n = \hat{A} \hat{B} \psi_n = a_n \hat{B} \psi_n \quad (37)$$

This shows that  $\hat{B} \psi_n$  is also an eigenfunction of  $\hat{A}$  with the same eigenvalue  $a_n$ . This implies that

$$\hat{B} \psi_n = \text{const} \psi_n = b_n \psi_n \quad (38)$$

showing that  $\psi_n$  is a simultaneous eigenfunction of  $\hat{A}$  and  $\hat{B}$  with eigenvalues  $a_n$  and  $b_n$ , respectively. The derivation becomes slightly more complicated in the case of degenerate eigenfunctions, but the same conclusion follows.

After the Hamiltonian, the operators for angular momenta are probably the most important in quantum mechanics. The definition of angular momentum in classical mechanics is  $\mathbf{L} = \mathbf{r} \times \mathbf{p}$ . In terms of its cartesian components,

$$\begin{aligned} L_x &= yp_z - zp_y \\ L_y &= zp_x - xp_z \\ L_z &= xp_y - yp_x \end{aligned} \quad (39)$$

In future, we will write such sets of equation as “ $L_x = yp_z - zp_y$ , *et cyc*,” meaning that we add to one explicitly stated relation, the versions formed by successive cyclic permutation  $x \rightarrow y \rightarrow z \rightarrow x$ . The general prescription for turning a classical dynamical variable into a quantum-mechanical operator was developed in Chap 2. The key relations were the momentum components

$$\hat{p}_x = -i\hbar \frac{\partial}{\partial x}, \quad \hat{p}_y = -i\hbar \frac{\partial}{\partial y}, \quad \hat{p}_z = -i\hbar \frac{\partial}{\partial z} \quad (40)$$

with the coordinates  $x, y, z$  simply carried over into multiplicative operators. Applying (40) to (39), we construct the three angular momentum operators

$$\hat{L}_x = -i\hbar \left( y \frac{\partial}{\partial z} - z \frac{\partial}{\partial y} \right) \quad \textit{et cyc} \quad (41)$$

while the total angular momentum is given by

$$\hat{L}^2 = \hat{L}_x^2 + \hat{L}_y^2 + \hat{L}_z^2 \quad (42)$$

The angular momentum operators obey the following commutation relations:

$$[\hat{L}_x, \hat{L}_y] = i\hbar \hat{L}_z \quad \textit{et cyc} \quad (43)$$



but

$$[\hat{L}^2, \hat{L}_z] = 0 \quad (44)$$

and analogously for  $\hat{L}_x$  and  $\hat{L}_y$ . This is consistent with the existence of simultaneous eigenfunctions of  $\hat{L}^2$  and any one component, conventionally designated  $\hat{L}_z$ . But then these states *cannot* be eigenfunctions of either  $\hat{L}_x$  or  $\hat{L}_y$ .

## Postulates of Quantum Mechanics

Our development of quantum mechanics is now sufficiently complete that we can reduce the theory to a set of postulates.

Postulate 1. The state of a quantum-mechanical system is completely specified by a wavefunction  $\Psi$  that depends on the coordinates and time. The square of this function  $\Psi^*\Psi$  gives the probability density for finding the system with a specified set of coordinate values.

The wavefunction must fulfill certain mathematical requirements because of its physical interpretation. It must be single-valued, finite and continuous. It must also satisfy a normalization condition

$$\int \Psi^*\Psi d\tau = 1 \quad (45)$$

Postulate 2. Every observable in quantum mechanics is represented by a linear, hermitian operator.

The hermitian property was defined in Eq (6). A linear operator is one which satisfies the identity

$$\hat{A}(c_1\psi_1 + c_2\psi_2) = c_1\hat{A}\psi_1 + c_2\hat{A}\psi_2 \quad (46)$$

which is required in order to have a superposition property for quantum states. The form of an operator which has an analog in classical mechanics is derived by the prescriptions

$$\hat{\mathbf{r}} = \mathbf{r}, \quad \hat{\mathbf{p}} = -i\hbar\nabla \quad (47)$$

which we have previously expressed in terms of cartesian components [cf. Eq (40)].

Postulate 3. In any measurement of an observable  $A$ , associated with an operator  $\hat{A}$ , the only possible results are the eigenvalues  $a_n$ , which satisfy an eigenvalue equation

$$\hat{A}\psi_n = a_n \psi_n \quad (48)$$

This postulate captures the essence of quantum mechanics—the quantization of dynamical variables. A continuum of eigenvalues is not forbidden, however, as in the case of an unbound particle.

Every measurement of  $A$  invariably gives one of the eigenvalues. For an arbitrary state (not an eigenstate of  $A$ ), these measurements will be individually unpredictable but follow a definite statistical law, which is the subject of the fourth postulate:

Postulate 4. For a system in a state described by a normalized wave function  $\Psi$ , the average or expectation value of the observable corresponding to  $A$  is given by

$$\langle A \rangle = \int \Psi^* \hat{A} \Psi d\tau \quad (49)$$

Finally,

Postulate 5. The wavefunction of a system evolves in time in accordance with the time-dependent Schrödinger equation

$$i\hbar \frac{\partial \Psi}{\partial t} = \hat{H} \Psi \quad (50)$$

For time-independent problems this reduces to the time-independent Schrödinger equation

$$\hat{H} \psi = E \psi \quad (51)$$

which is the eigenvalue equation for the Hamiltonian operator.

## The Variational Principle

Except for a small number of intensively-studied examples, the Schrödinger equation for most problems of chemical interest *cannot* be solved exactly. The variational principle provides a guide for constructing the best possible approximate solutions of a specified functional form. Suppose that we seek an approximate solution for the ground state of a quantum system described by a Hamiltonian  $\hat{H}$ . We presume that the Schrödinger equation

$$\hat{H} \psi_0 = E_0 \psi_0 \quad (52)$$

is too difficult to solve exactly. Suppose, however, that we have a function  $\tilde{\psi}$  which we think is an approximation to the true ground-state wavefunction. According to the variational principle (or variational theorem), the following formula provides an *upper bound* to the exact ground-state energy  $E_0$ :

$$\tilde{E} \equiv \frac{\int \tilde{\psi}^* \hat{H} \tilde{\psi} d\tau}{\int \tilde{\psi}^* \tilde{\psi} d\tau} \geq E_0 \quad (53)$$

Note that this ratio of integrals has the same form as the expectation value  $\langle H \rangle$  defined by (29). The better the approximation  $\tilde{\psi}$ , the lower will be the computed energy  $\tilde{E}$ , though it will still be greater than the exact value. To prove Eq (53), we suppose that the approximate function can, in concept, be represented as a superposition of the actual eigenstates of the Hamiltonian, analogous to (24),

$$\tilde{\psi} = c_0 \psi_0 + c_1 \psi_1 + \dots \quad (54)$$

This means that  $\tilde{\psi}$ , the approximate ground state, might be close to the actual ground state  $\psi_0$  but is “contaminated” by contributions from excited states  $\psi_1, \dots$ . Of course, none of the states or coefficients on the right-hand side is actually known, otherwise there would no need to worry about approximate computations. By Eq (25), the expectation value of the Hamiltonian in the state (54) is given by

$$\tilde{E} = |c_0|^2 E_0 + |c_1|^2 E_1 + \dots \quad (55)$$

Since all the excited states have *higher* energy than the ground state,  $E_1, E_2 \dots \geq E_0$ , we find

$$\tilde{E} \geq (|c_0|^2 + |c_1|^2 + \dots) E_0 = E_0 \quad (56)$$

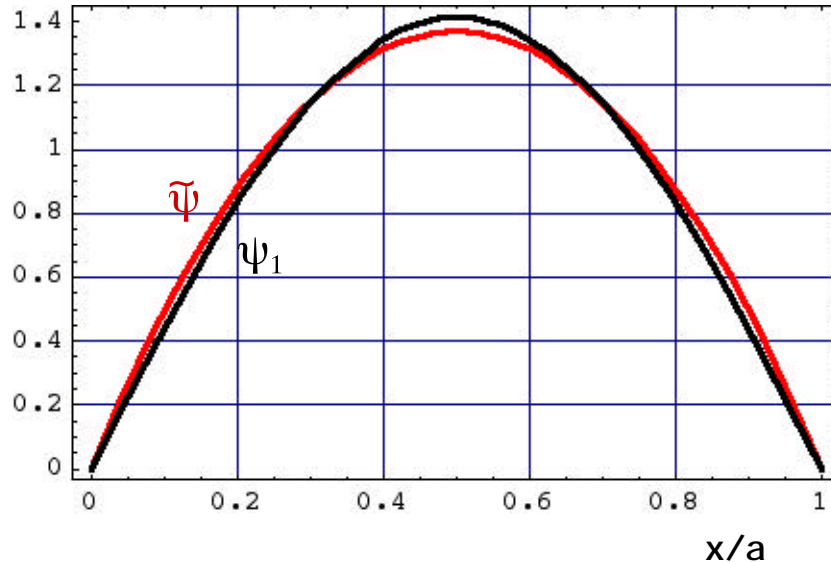
assuming  $\tilde{\psi}$  has been normalized. Thus  $\tilde{E}$  must be greater than the true ground-state energy  $E_0$ , as implied by (53)

As a very simple, although artificial, illustration of the variational principle, consider the ground state of the particle in a box. Suppose we had never studied trigonometry and knew nothing about sines or cosines. Then a reasonable approximation to the ground state might be an inverted parabola such as the normalized function

$$\tilde{\psi}(x) = \left(\frac{30}{a^5}\right)^{1/2} x(a-x) \quad (57)$$

Fig. 1 shows this function along with the exact ground-state eigenfunction

$$\psi_1(x) = \left(\frac{2}{a}\right)^{1/2} \sin \frac{\pi x}{a} \quad (58)$$



**Figure 1.** Variational approximation for particle in a box.

A variational calculation gives

$$\begin{aligned} \tilde{E} &= \int_0^a \tilde{\psi}(x) \left(-\frac{\hbar^2}{2m}\right) \tilde{\psi}''(x) dx = \\ &= \frac{5}{4\pi^2} \frac{h^2}{ma^2} = \frac{10}{\pi^2} E_1 = 1.01321 E_1 \quad (59) \end{aligned}$$

in terms of the exact ground state energy  $E_1 = h^2/8ma^2$ . In accord with the variational theorem,  $\tilde{E} > E_1$ . The computation is in error by about 1%.