

This content has been downloaded from IOPscience. Please scroll down to see the full text.

Download details:

IP Address: 18.116.47.245

This content was downloaded on 05/05/2024 at 16:58

Please note that [terms and conditions apply](#).

You may also like:

[First-Principles Calculations for Cathode, Electrolyte and Anode Battery Materials](#)

[INVERSE PROBLEMS NEWSLETTER](#)

[Concise Short-Channel Effect Model for Inversion-type Ultrathin Silicon-on-Insulator p-Channel Metal-Oxide-Silicon Field-Effect Transistors Based on Partitioning of Thin-Film](#)  
Shiao-Shien Chen Shiao-Shien Chen and James B. Kuo James B. Kuo

[TOPICAL REVIEWS 1994 - 1996](#)

## A Concise Introduction to Quantum Mechanics

Mark S Swanson

## Chapter 4

## Wave mechanics

In 1926 Schrödinger proposed a partial differential equation, now known as the *Schrödinger equation*, that governs the wave function, and formal quantum mechanics was developed from it. Solving the Schrödinger equation and applying the resulting wave function is often referred to as *wave mechanics*, and its basic structure is the subject of this chapter.

#### 4.1 The Schrödinger equation and its general properties

The properties required for the wave function, and therefore the equation that governs it, have been discussed. These are the properties of linear superposition, membership in the space of square integrable functions  $L_2$ , the result (3.15) for a massive particle wave function undergoing a Galilean transformation of the form (1.17), the need to include potential energies, an independence of the scale chosen for the energy, the ability to represent a large variety of spatial boundary conditions through a complete basis set of functions, and an equation that is linear in the time derivative. The most viable candidate for such a governing equation is a *linear partial differential equation of the Sturm–Liouville type*, which typically generates complete sets of functions.

The Schrödinger equation for the wave function can be found by examining the properties of the basis wave functions (3.49) of the particle in a box, given at  $t = 0$  by

$$\Psi_n(x, t = 0) = \sqrt{\frac{1}{L}} \sin\left(\frac{n\pi x}{L}\right) = \sqrt{\frac{1}{L}} \sin\left(\frac{p_n x}{\hbar}\right), \quad (4.1)$$

where  $n$  is a positive integer. Applying the de Broglie wavelength formula (2.18) identifies the magnitude of momentum  $p_n = n\pi\hbar/L$  and the energy  $E_n = p_n^2/2m = n^2\pi^2\hbar^2/2mL^2$  associated with the basis wave function  $\Psi_n$ . The term

associated will now be formalized by noting that the basis wave function of (4.1) satisfies the mathematical relation

$$-\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} \Psi_n(x, t = 0) = \frac{p_n^2}{2m} \Psi_n(x, t = 0) = E_n \Psi_n(x, t = 0). \quad (4.2)$$

Examining (4.2) shows that the quantum mechanical energy is *obtained by applying the differential operator*  $-(\hbar^2/2m) \partial^2/\partial x^2$  to the basis wave function. This differential operator can be written as  $\hat{p}^2/2m$ , where  $\hat{p} = -i\hbar \partial/\partial x$ . Throughout this text operators are designated with a hat, so that  $\hat{p}$  is read as *the operator p*. Since the differential operator  $\hat{p}^2/2m$  gives the energy  $p_n^2/2m$ , the operator  $\hat{p}$  is identified as the *quantum mechanical momentum operator*. In wave mechanics the canonical momentum  $p$  of Newtonian physics becomes the *differential operator*  $\hat{p} = -i\hbar \partial/\partial x$ , and this momentum operator *acts on the wave function*. This is the first example of what is known as the *correspondence principle*, which is the postulate that classical quantities, such as energy, momentum, and angular momentum, correspond to *differential operators* in wave mechanics. It is natural to question why the negative sign is used for the momentum, since either sign works equally well in (4.2). This question will be answered by the result (4.13) for a Galilean transformation on the Schrödinger equation. In (4.2) the operator  $\hat{H} = \hat{p}^2/2m$  corresponds to the classical Hamiltonian  $H = p^2/2m$  with  $p \rightarrow \hat{p}$ , so that

$$\hat{H} \Psi_n(x, t = 0) = \frac{\hat{p}^2}{2m} \Psi_n(x, t = 0) = -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} \Psi_n(x, t = 0) = E_n \Psi_n(x, t = 0). \quad (4.3)$$

The result (4.3) is referred to as an *eigenvalue equation*. The function  $\Psi_n \equiv \Psi_n(x, t = 0)$  is an *eigenfunction* of the differential operator corresponding to the *Hamiltonian*, denoted by  $\hat{H}(\hat{p})$ , and the action of  $\hat{H}$  on the eigenfunction yields the associated *energy eigenvalue*  $E_n = p_n^2/2m$ . The result (4.3) is written  $\hat{H}\Psi_n = E_n\Psi_n$ . Eigenvalue equations such as (4.3) are the predominant mathematical statement of formal quantum mechanics. In the Copenhagen interpretation of quantum mechanics, the eigenvalues provide the *possible experimental values of the observable* associated with the operator in the eigenvalue equation.

The generalization of momentum to higher spatial dimensions uses the gradient,

$$\hat{\mathbf{p}} = -i\hbar\nabla. \quad (4.4)$$

Using (4.4), the complex function  $\psi_p(\mathbf{x}) = \exp(i\mathbf{p} \cdot \mathbf{x}/\hbar)$  and its complex conjugate satisfy

$$\hat{\mathbf{p}} \psi_p(\mathbf{x}) = -i\hbar\nabla\psi_p(\mathbf{x}) = \mathbf{p} \psi_p(\mathbf{x}), \quad \hat{\mathbf{p}} \psi_p^*(\mathbf{x}) = -i\hbar\nabla\psi_p^*(\mathbf{x}) = -\mathbf{p} \psi_p^*(\mathbf{x}). \quad (4.5)$$

The result (4.5) demonstrates that  $\psi_p(\mathbf{x})$  and  $\psi_p^*(\mathbf{x})$  are *eigenfunctions* of the momentum operator,  $\hat{\mathbf{p}} = -i\hbar\nabla$ , corresponding respectively to the *eigenvalues*  $\mathbf{p}$  and  $-\mathbf{p}$ . Insight into these functions is found by looking at the electromagnetic wave (1.40) in terms of its photon content. Using the de Broglie relation,  $\mathbf{k} = \mathbf{p}/\hbar$ , and

the Einstein relation,  $\omega = ck = cp/\hbar = E_p/\hbar$ , gives  $\mathbf{A} = \mathbf{A}_0 \cos((\mathbf{p} \cdot \mathbf{x} - E_p t)/\hbar)$ . This general form for a wave can be applied to the momentum eigenfunctions of (4.5) to create a time-dependent momentum–energy wave function for the Newtonian particle. Normalizing (4.5) with the volume  $V$  of space, the result is a form identical to the generic wave function of (3.19),

$$\Psi_p(\mathbf{x}, t) = \frac{1}{\sqrt{V}} e^{i(\mathbf{p} \cdot \mathbf{x} - E_p t)/\hbar}, \quad (4.6)$$

where  $E_p = p^2/2m$ . In addition to being an eigenfunction of the *momentum operator* (4.4), function (4.6) also satisfies the equation

$$i\hbar \frac{\partial}{\partial t} \Psi_p(\mathbf{x}, t) = E_p \Psi_p(\mathbf{x}, t), \quad (4.7)$$

so that  $\Psi_p$  is also an eigenfunction of the differential operator  $i\hbar \partial/\partial t$  with the eigenvalue  $E_p$ . Although (4.7) was *motivated* by the photon related electromagnetic wave, this general recipe provides an equation that is linear in the time derivative and will now be incorporated into the equation for free Newtonian particles, resulting in the free particle version of the Schrödinger equation.

In order to extend the result (4.7) to the case of a free massive Newtonian particle, the spatial equation (4.2) must allow the eigenvalue  $E_p$  to be identified as the energy of a free massive Newtonian particle,  $E_p = p^2/2m$ . If this identification holds, it then results in a function that satisfies a *linear* partial differential equation, explicitly given by

$$\left( \hat{H}(\hat{\mathbf{p}}) - i\hbar \frac{\partial}{\partial t} \right) \Psi_p(\mathbf{x}, t) = \left( -\frac{\hbar^2}{2m} \nabla^2 - i\hbar \frac{\partial}{\partial t} \right) \Psi_p(\mathbf{x}, t) = 0. \quad (4.8)$$

Using (4.1) shows that equation (4.8) also holds for a particle in a box if the time-dependent wave function takes the form  $\Psi(\mathbf{x}, t) = e^{-iE_p t/\hbar} \Psi_n(\mathbf{x})$ . The final requirement is to show that (4.8) is consistent with Galilean transformations. Using  $\mathbf{p}' = \mathbf{p} - m\mathbf{v}$ ,  $\mathbf{x}' = \mathbf{x} - \mathbf{v}t$ ,  $t' = t$ , and  $E_{p'} = p'^2/2m = (\mathbf{p} - m\mathbf{v})^2/2m = p^2/2m - \mathbf{v} \cdot \mathbf{p} + \frac{1}{2}m\mathbf{v}^2$  in (4.6) gives

$$\Psi_{p'}(\mathbf{x}', t') = e^{i(m\mathbf{v} \cdot \mathbf{x} - \frac{1}{2}m\mathbf{v}^2 t)}/\hbar \frac{1}{\sqrt{V}} e^{i(\mathbf{p} \cdot \mathbf{x} - E_p t)/\hbar} \equiv e^{-i\alpha(\mathbf{v}, \mathbf{x}, t)} \Psi_p(\mathbf{x}, t). \quad (4.9)$$

The phase angle  $\alpha = (m\mathbf{v} \cdot \mathbf{x} - \frac{1}{2}m\mathbf{v}^2 t)/\hbar$  that results from the transformation depends on  $\mathbf{x}$ ,  $t$ , and the *relative velocity*  $\mathbf{v}$ , but *not*  $\mathbf{p}$ . Therefore, result (4.9) is the general behavior (3.15) under a Galilean transformation that is required to make the probability density a scalar function. This phase will be the same *even for linear combinations of*  $\Psi_p$  *with different values of*  $\mathbf{p}$ . This allows mixed momentum states, which are critical to explaining quantum phenomena, to transform according to (3.15).

In addition, the differential equation (4.8) must also be *form invariant* under a Galilean transformation. This is demonstrated by combining the chain rule for partial derivatives with (1.17), which results in

$$\frac{\partial}{\partial x'_i} = \sum_{j=1}^3 \frac{\partial x_j}{\partial x'_i} \frac{\partial}{\partial x_j} + \frac{\partial t}{\partial x'_i} \frac{\partial}{\partial t} = \frac{\partial}{\partial x_i}, \quad (4.10)$$

$$\frac{\partial}{\partial t'} = \sum_{j=1}^3 \frac{\partial x_j}{\partial t'} \frac{\partial}{\partial x_j} + \frac{\partial t}{\partial t'} \frac{\partial}{\partial t} = \sum_{j=1}^3 v_j \frac{\partial}{\partial x_j} + \frac{\partial}{\partial t} = \mathbf{v} \cdot \nabla + \frac{\partial}{\partial t}. \quad (4.11)$$

Using (4.9), (4.10), and (4.11) in (4.8) yields, after some rearrangement,

$$\begin{aligned} \left( -\frac{\hbar^2}{2m} \nabla'^2 - i\hbar \frac{\partial}{\partial t'} \right) \Psi_p(\mathbf{x}', t') &= \left( -\frac{\hbar^2}{2m} \nabla^2 - i\hbar \left( \frac{\partial}{\partial t} + \mathbf{v} \cdot \nabla \right) \right) \Psi_p(\mathbf{x}', t') \\ &= \left( \frac{1}{2m} (-i\hbar \nabla + m\mathbf{v})^2 - \left( i\hbar \frac{\partial}{\partial t} + \frac{1}{2} m\mathbf{v}^2 \right) \right) \\ &\quad \times \left( e^{-i\alpha(\mathbf{v}, \mathbf{x}, t)} \Psi_p(\mathbf{x}, t) \right). \end{aligned} \quad (4.12)$$

The appearance of the combination  $(-i\hbar \nabla + m\mathbf{v})$  in the transformed equation (4.12) shows that Galilean invariance *requires* identifying the momentum operator with the negative sign,  $\hat{\mathbf{p}} = -i\hbar \nabla$ . This follows from the phase angle  $\alpha$  in (4.9), which gives two results,

$$(-i\hbar \nabla + m\mathbf{v}) e^{-i\alpha(\mathbf{v}, \mathbf{x}, t)} = 0, \quad \left( i\hbar \frac{\partial}{\partial t} + \frac{1}{2} m\mathbf{v}^2 \right) e^{-i\alpha(\mathbf{v}, \mathbf{x}, t)} = 0. \quad (4.13)$$

These results show that the transformed momentum  $\hat{\mathbf{p}}' = -i\hbar \nabla'$  and wave function give

$$\hat{\mathbf{p}}' \Psi_p(\mathbf{x}', t') = -i\hbar \nabla' \Psi_p(\mathbf{x}', t') = -i\hbar \nabla \left( e^{-i\alpha} \Psi_p(\mathbf{x}, t) \right) = (\mathbf{p} - m\mathbf{v}) \Psi_p(\mathbf{x}', t'), \quad (4.14)$$

which reproduces the Galilean transformation. Similarly, applying  $i\hbar \partial/\partial t'$  to  $\Psi_p(\mathbf{x}', t')$  yields the the eigenvalue  $(\mathbf{p} - m\mathbf{v})^2/2m$ . In addition, using (4.13) in (4.12) gives

$$\left( -\frac{\hbar^2}{2m} \nabla'^2 - i\hbar \frac{\partial}{\partial t'} \right) \Psi_p(\mathbf{x}', t') = e^{-i\alpha(\mathbf{v}, \mathbf{x}, t)} \left( -\frac{\hbar^2}{2m} \nabla^2 - i\hbar \frac{\partial}{\partial t} \right) \Psi_p(\mathbf{x}, t) = 0, \quad (4.15)$$

showing that (4.8) takes the same form for *all Galilean observers, regardless of relative motion*. Because *both* the solution *and* the equation it obeys satisfy the requirements for a quantum mechanical wave function, (4.8) is the correct quantum mechanical equation for the case of a free Newtonian particle. The linearity in the time derivative is critical to the Galilean invariance of (4.15). In that regard, the d'Alembertian,  $\square = -\nabla^2 + \partial^2/\partial(ct)^2$ , which governs light waves is *not* invariant under Galilean transformations, and therefore is not the correct differential operator for the wave function of a Newtonian particle.

The final step is to generalize (4.8) to the case that a potential energy  $U(\mathbf{x}, t)$  is present. This is done by using the associated classical Hamiltonian,  $H(\mathbf{p}, \mathbf{x}, t)$ , defined by (1.9) so that it includes the potential. The replacement  $H \rightarrow \hat{H}(-i\hbar\nabla, \mathbf{x}, t)$  turns the classical Hamiltonian into a differential operator. The generalization of (4.8) is then known as the *Schrödinger equation*, a linear partial differential equation given by

$$\hat{H}(\hat{\mathbf{p}}, \mathbf{x}, t) \Psi(\mathbf{x}, t) = \hat{H}(-i\hbar\nabla, \mathbf{x}, t) \Psi(\mathbf{x}, t) = i\hbar \frac{\partial}{\partial t} \Psi(\mathbf{x}, t). \quad (4.16)$$

Combining the same transformation (4.9) for the wave function with  $U(\mathbf{x}, t) \rightarrow U(\mathbf{x} - \mathbf{v}t, t)$  leaves (4.16) *form invariant* under a Galilean transformation. Although there is no *a priori* justification for using the same potential present in the classical system, Ehrenfest's theorem, presented later in this chapter, shows that a form of Hamilton's equations (1.13) follow from (4.16). The Hamiltonian operator of (4.16) can therefore be viewed as an example of the *correspondence principle*, which derives all the *operators* of quantum mechanics from their classical antecedents. The best justification for (4.16) is its great success explaining the myriad of quantum phenomena to which it has been applied.

Investigating the general properties of (4.16) begins by noting that the complex conjugate of (4.16) yields a second equation, given by

$$\hat{H}^*(\hat{\mathbf{p}}, \mathbf{x}, t) \Psi^*(\mathbf{x}, t) = -i\hbar \frac{\partial}{\partial t} \Psi^*(\mathbf{x}, t), \quad (4.17)$$

where the process of complex conjugation for  $\hat{H}$  uses  $\hat{\mathbf{p}}^* = i\hbar\nabla$ . Multiplying (4.16) by  $\Psi^*$  and (4.17) by  $\Psi$  and subtracting the latter from the former yields

$$\Psi^*(\hat{H}\Psi) - (\hat{H}^*\Psi^*)\Psi = i\hbar\left(\frac{\partial}{\partial t} \Psi^*\right)\Psi + i\hbar\Psi^*\left(\frac{\partial}{\partial t} \Psi\right) = i\hbar \frac{\partial}{\partial t} (\Psi^* \Psi). \quad (4.18)$$

It is often the case that the classical Hamiltonian has the form  $H = p^2/2m + U(\mathbf{x}, t)$ , where  $U$  is the potential energy. For that case, the quantum mechanical Hamiltonian is

$$\hat{H} = \frac{\hat{p}^2}{2m} + U(\mathbf{x}, t) = -\frac{\hbar^2}{2m}\nabla^2 + U(\mathbf{x}, t). \quad (4.19)$$

If  $U^*(\mathbf{x}, t) = U(\mathbf{x}, t)$ , then the potential energy terms on the left-hand side of (4.18) cancel. The remaining Laplacian terms in (4.18) can be rewritten using the identity

$$(\nabla^2\Psi^*)\Psi - \Psi^*(\nabla^2\Psi) = \nabla \cdot (\nabla\Psi^*\Psi - \Psi^*\nabla\Psi), \quad (4.20)$$

so that, for the case of (4.19) and  $U^*=U$ , (4.18) can be written

$$-\nabla \cdot \mathbf{J}(\mathbf{x}, t) \equiv \nabla \cdot \frac{\hbar}{2im}(\nabla\Psi^*\Psi - \Psi^*\nabla\Psi) = \frac{\partial}{\partial t}(\Psi^*\Psi) = \frac{\partial}{\partial t}\mathcal{P}(\mathbf{x}, t). \quad (4.21)$$

Result (4.21),  $\nabla \cdot \mathbf{J} + \partial\mathcal{P}/\partial t = 0$ , is identical in form to the conservation law for electric charge given by (1.30). Since  $\mathcal{P}(\mathbf{x}, t) = \Psi^*(\mathbf{x}, t)\Psi(\mathbf{x}, t)$  is the probability density and plays the role of the electric charge density, result (4.21) represents the

*conservation of probability.* In three spatial dimensions the vector  $\mathbf{J}$ , given explicitly by

$$\mathbf{J}(\mathbf{x}, t) = i\frac{\hbar}{2m}(\nabla\Psi^*(\mathbf{x}, t)\Psi(\mathbf{x}, t) - \Psi^*(\mathbf{x}, t)\nabla\Psi(\mathbf{x}, t)), \quad (4.22)$$

has the SI units of  $1/(m^2 \text{ s})$ , and can therefore be identified as the *current density for probability*. The result (4.21) satisfies the same relation as (1.31), so that any probability lost or gained in a volume over time is matched by the flow of probability out of or into that volume through its surfaces.

The form of the spatial probability current (4.22) places a continuity requirement on  $\Psi$  and  $\nabla\Psi$ . If  $\mathbf{J}$  is discontinuous at a point, then that point will serve as a *source* or *sink* of probability. This is seen by considering an infinitesimal closed surface  $d\mathbf{S}$  around the point and noting that  $\mathbf{J} \cdot d\mathbf{S} \neq 0$  if  $\mathbf{J}$  is discontinuous at that point. Therefore, in order to avoid creating or destroying probability,  $\mathbf{J}$  must be *spatially continuous*, and this *generally* requires that both  $\Psi$  and  $\nabla\Psi$  must be continuous. If one is discontinuous at a point, then the other must vanish at that point in such a way that  $\mathbf{J}$  remains continuous. The conservation of probability for the very general form (4.19) is critical to interpreting the results of solving the Schrödinger equation. An example for barrier penetration is given in the next chapter.

As an example of probability current, the momentum eigenfunction (4.6) for a particle in a volume  $V$  yields  $\mathbf{J} = \mathbf{p}/(mV)$  and satisfies  $\nabla \cdot \mathbf{J} = 0$ . This is consistent with the right-hand side of (4.21), which vanishes since  $\mathcal{P} = 1/V$  is a constant. The flow of probability through a surface associated with the momentum eigenfunction is therefore in the direction of the classical velocity  $\mathbf{v} = \mathbf{p}/m$  and is equally into and out of *any volume*, including the total volume  $V$  of the system. The total volume can be chosen as a cube of side  $L$ , so that  $V = L^3$ . The flux of probability through a surface  $\mathbf{S}$  of the cube with area  $L^2$  is then given by  $\mathbf{J} \cdot \mathbf{S} = \mathbf{p} \cdot \hat{\mathbf{n}}/mL$ , where  $\hat{\mathbf{n}}$  is normal to the surface. This should be zero if the particle remains in the volume, and this is true only if  $L \rightarrow \infty$ . As a result, (4.6) correctly describes the situation of a particle contained in a volume  $V$  only if the volume is arbitrarily large. This is discussed in the section on the Heisenberg uncertainty principle.

In addition to the potential energy of (4.19) being real, it is often the case that  $U$  has no explicit time-dependence. For such a case, the wave function can undergo *separation of variables*, so that it is written  $\Psi(\mathbf{x}, t) = \psi(\mathbf{x}) T(t)$ . Inserting this in the Schrödinger equation (4.16) and dividing by  $\Psi(\mathbf{x}, t) = \psi(\mathbf{x}) T(t)$  yields

$$\frac{1}{\psi(\mathbf{x})} \left( -\frac{\hbar^2}{2m} \nabla^2 + U(\mathbf{x}) \right) \psi(\mathbf{x}) = \frac{i\hbar}{T(t)} \frac{\partial}{\partial t} T(t). \quad (4.23)$$

For such a case the left-hand side is solely a function of  $\mathbf{x}$ , while the right-hand side is solely a function of  $t$ . Since  $\mathbf{x}$  and  $t$  are independent of each other, both sides of (4.16) must equal the same constant, designated  $E$ . This constant has the same units as  $\hbar/t$ , which is energy. The right-hand side is then very simple to solve, and the function  $T(E, t)$  associated with the choice of  $E$  is given by

$$T(E, t) = \exp\left(-\frac{i}{\hbar}Et\right), \quad (4.24)$$

where (4.24) has been normalized so that  $T^*(E, t) T(E, t) = 1$ . For such a case the wave function has the required property that the energy  $E$  can be modified by an arbitrary amount,  $E \rightarrow E + \delta E$ , since doing so simply creates an overall phase factor,

$$T(E + \delta E, t) = \exp\left(-\frac{i}{\hbar}\delta Et\right) \exp\left(-\frac{i}{\hbar}Et\right) = \exp\left(-\frac{i}{\hbar}\delta Et\right) T(E, t). \quad (4.25)$$

As was the case with the Galilean transformation of (4.9), an overall phase factor has no effect on the quantum mechanical probability density. For a separable wave function the resulting quantum mechanics, like Newtonian mechanics, does not depend on the zero point chosen for the energy. In addition, a separable wave function has the property that the probability density  $\mathcal{P} = \Psi^*(\mathbf{x}, t)\Psi(\mathbf{x}, t) = \psi^*(\mathbf{x})\psi(\mathbf{x})$  is *time-independent* since  $T^*(E, t) T(E, t) = 1$ . For the case that (4.21) holds, the spatial probability current  $\mathbf{J}$  of (4.22) must satisfy  $\nabla \cdot \mathbf{J} = -\partial\mathcal{P}/\partial t = 0$ , and this requires *continuity* for  $\mathbf{J}(\mathbf{x})$ .

For the case of a time-independent potential, the Schrödinger equation (4.23) reduces to an *eigenvalue equation* for  $E$  and the *associated* spatial wave function  $\psi_E(\mathbf{x})$ ,

$$\hat{H}(\hat{\mathbf{p}}, \mathbf{x}) \psi_E(\mathbf{x}) = \left(-\frac{\hbar^2}{2m}\nabla^2 + U(\mathbf{x})\right) \psi_E(\mathbf{x}) = E \psi_E(\mathbf{x}). \quad (4.26)$$

If solvable, the Hamiltonian eigenvalue equation (4.26) generates a set of *constant energy eigenvalues* associated with a set of time-independent eigenfunctions, known as *stationary states*. In the Copenhagen interpretation, solving (4.26) with appropriate boundary conditions determines *all the possible observable energies for the quantum mechanical system* while simultaneously determining the stationary state  $\psi_E(\mathbf{x})$  associated with each energy eigenvalue. It is important to determine the conditions for which the energy eigenvalues are real. This is because an imaginary part for  $E$ , designated  $\epsilon$ , so that  $E = E_R \pm i\epsilon$ , will cause the function (4.24) to exhibit exponential damping or growth over time since  $\exp(-iEt/\hbar) = \exp(-iE_R t/\hbar)\exp(\pm\epsilon t/\hbar)$ . This would correspond to an *unstable solution* since the wave function would not remain normalized and the probabilistic interpretation of the wave function would then be lost. The starting point to examine this possibility is the complex conjugate of (4.26), which yields a second eigenvalue equation involving  $E^*$ ,

$$\hat{H}^* \psi_E^*(\mathbf{x}) = E^* \psi_E^*(\mathbf{x}). \quad (4.27)$$

Multiplying (4.26) by  $\psi^*(\mathbf{x})$  and (4.27) by  $\psi(\mathbf{x})$ , subtracting the latter from the former, and then integrating over the volume of space for which the system is defined yields



$$\int_V dV \left( \psi_E^* (\hat{H} \psi_E) - (\hat{H}^* \psi_E^*) \psi_E \right) = (E - E^*) \int_V dV \psi^*(\mathbf{x}) \psi(\mathbf{x}) = E - E^*, \quad (4.28)$$

since normalization requires that the integral of  $|\psi(\mathbf{x})|^2$  is one. For the case that the Hamiltonian satisfies the condition

$$\int_V dV \psi_E^* (\hat{H} \psi_E) = \int_V dV (\hat{H}^* \psi_E^*) \psi_E, \quad (4.29)$$

the left-hand side of (4.28) will vanish and therefore  $E - E^* = 0$ , insuring that the *energy eigenvalues are real*.

A Hamiltonian with the form (4.19) will now be shown to satisfy (4.29) as long as  $U(\mathbf{x}) = U^*(\mathbf{x})$ . For such a case, the terms involving the potential are the same on both sides of (4.29) and therefore cancel. Suppressing the factor of  $\hbar^2/2m$  and using the identity (4.20), the Laplacian term on the left-hand side of (4.29) can be rewritten as

$$\int_V dV \psi_E^* (\nabla^2 \psi_E) = \int_V dV \nabla \cdot (\psi_E^* \nabla \psi_E - \nabla \psi_E^* \psi_E) + \int_V dV (\nabla^2 \psi_E^*) \psi_E. \quad (4.30)$$

Comparing the first integral on the right-hand side of (4.30) with the spatial probability current (4.22) shows that the argument is proportional to  $\nabla \cdot \mathbf{J}$ . It was shown earlier that  $\nabla \cdot \mathbf{J} = 0$  for a wave function that is separable, as in (4.23). Therefore, only the second integral on the right-hand side of (4.30) remains, which is identical to the Laplacian term on the right-hand side of (4.29). As a result, Hamiltonians of the general form (4.26) with a *real valued* potential  $U(\mathbf{x})$  possess *real energy eigenvalues*. In the next section, a general criterion for real eigenvalues, called *Hermiticity*, will be established.

A Hamiltonian of the form (4.19) also has the property that the spatial eigenfunctions associated with *different* energy eigenvalues are *orthogonal*. If  $\psi_E$  and  $\psi_{E'}$  are two different energy eigenfunctions associated with a Hamiltonian  $\hat{H}$  satisfying, it follows that both  $E$  and  $E'$  are real. The eigenfunctions then satisfy

$$\int_V dV \psi_{E'}^* (\hat{H} \psi_E) = E \int_V dV \psi_{E'}^* \psi_E \quad (4.31)$$

$$\int_V dV (\hat{H}^* \psi_{E'}^*) \psi_E = E' \int_V dV \psi_{E'}^* \psi_E. \quad (4.32)$$

Since the Hamiltonian satisfies (4.29), the left-hand sides of (4.31) and (4.32) are equal. Subtracting (4.32) from (4.31) therefore gives

$$(E - E') \int_V dV \psi_{E'}^* \psi_E = 0. \quad (4.33)$$

Comparing this result to (3.23) shows that the integral appearing in (4.33) is the *inner product* of the two eigenfunctions. This means that spatial eigenfunctions associated with different energy eigenvalues are *orthogonal*, i.e. the integral giving their inner product vanishes since  $E - E' \neq 0$ . The Hamiltonian may also have orthogonal

eigenfunctions associated with the *same* eigenvalue. Such an energy eigenvalue is referred to as *degenerate*.

In general, the Hamiltonians considered in wave mechanics give rise to a *complete set of orthonormal eigenfunctions*. This means that the set of all Hamiltonian eigenfunctions  $\Psi_j(\mathbf{x}, t)$ , where  $j$  is some set of *quantum numbers*, are both orthonormal,

$$\int_V d^n x \Psi_j^*(\mathbf{x}, t) \Psi_k(\mathbf{x}, t) = \delta_{jk}, \quad (4.34)$$

and complete,

$$\sum_j \Psi_j^*(\mathbf{y}, t) \Psi_j(\mathbf{x}, t) = \delta^{(n)}(\mathbf{y} - \mathbf{x}), \quad (4.35)$$

where  $n$  is the spatial dimension of the volume  $V$  system under consideration and the sum in (4.35) is over *all possible quantum numbers*. Such a set of functions typically forms a complete basis for the space of square integrable functions, referred to as  $L_2$ , allowing the *preparation* of the initial wave function so that it can represent an arbitrary experimental situation. The mathematical statement of completeness (4.35) for the eigenfunctions is identical to property (3.40) for the Fourier series. The Dirac delta appearing in (4.35) means that an *arbitrary* initial preparation  $\Psi(\mathbf{x}, 0)$  for the wave function can be *projected* onto this complete set,

$$\begin{aligned} \Psi(\mathbf{x}, 0) &= \int d^n y \Psi(\mathbf{y}, 0) \delta^{(n)}(\mathbf{y} - \mathbf{x}) \\ &= \sum_j \Psi_j(\mathbf{x}, 0) \int d^n y \Psi(\mathbf{y}, 0) \Psi_j^*(\mathbf{y}, 0) = \sum_j a_j \Psi_j(\mathbf{x}, 0), \end{aligned} \quad (4.36)$$

where the coefficient  $a_j$  is the result of the integral in (4.36). Therefore, for the case that the eigenfunctions form a basis of  $L_2$ , an arbitrary preparation of the wave function is possible. The proof of completeness requires the specific details of the eigenfunctions under consideration. However, in all the cases analyzed in this monograph the eigenfunctions will be a complete basis for the function space  $L_2$ .

The Schrödinger equation also has the *extremely important* property that it is linear in the time derivative. This means that the *time evolution* of the wave function is *determined* once its initial form is given or, in the Copenhagen interpretation, *prepared*. In other words, the time evolution of the wave function itself is deterministic and does not need additional initial or final information to be specified, unlike the *light wave equation* (1.39). Assuming the wave function is not singular at the time  $t$  and that the Hamiltonian satisfies  $\partial \hat{H} / \partial t = 0$ , the Schrödinger equation (4.16) allows the Taylor series expansion of the wave function around the time  $t$  to be written

$$\Psi(\mathbf{x}, t + T) = \sum_{n=0}^{\infty} \frac{T^n}{n!} \frac{\partial^n}{\partial t^n} \Psi(\mathbf{x}, t) = e^{-i\hat{H}(\hat{p}, \mathbf{x}) T / \hbar} \Psi(\mathbf{x}, t) \equiv \hat{U}(T) \Psi(\mathbf{x}, t), \quad (4.37)$$

where the time-independence of  $\hat{H}$ , so that  $\hat{H} \partial^n \Psi / \partial t^n = \partial^n (\hat{H} \Psi) / \partial t^n = i\hbar \partial^{n+1} \Psi / \partial t^{n+1}$ , allows the second step. The operator  $\hat{U}(T) = \exp(-i\hat{H} T / \hbar)$  appearing in (4.37) has

the property that  $\hat{U}^*(T) = \hat{U}^{-1}(T)$  as long as  $\hat{H}^* = \hat{H}$ . Examining (4.37) shows that  $\hat{U}(t)$  is responsible for the *time evolution* of the wave function,  $\Psi(\mathbf{x}, t) = \hat{U}(t)\Psi(\mathbf{x}, 0)$ , and for that reason it is referred to as the *evolution operator*. This result is particularly evident if  $\Psi_E(\mathbf{x}, t = 0)$  is an eigenfunction of  $\hat{H}$  with the eigenvalue  $E$ , since (4.24) gives

$$\hat{U}(t)\Psi_E(\mathbf{x}, 0) = e^{-i\hat{H}(\hat{\mathbf{p}}, \mathbf{x})t/\hbar}\Psi_E(\mathbf{x}, 0) = e^{-iEt/\hbar}\Psi_E(\mathbf{x}, 0) = \Psi_E(\mathbf{x}, t). \quad (4.38)$$

The Schrödinger equation therefore embodies the classical concept that the Hamiltonian is responsible for the time evolution of the physical system, an idea first encountered with the classical Poisson bracket formulation (1.15) involving the Hamiltonian. The action of the evolution operator  $\hat{U}(t)$  preserves wave function normalization as long as the Hamiltonian satisfies (4.29), since  $(\hat{U}\Psi)^* = \hat{U}^*\Psi^*$  gives

$$\int_V dV (\hat{U}^*\Psi^*)(\hat{U}\Psi) = \int_V dV \Psi^*\hat{U}^*(\hat{U}\Psi) = \int_V dV \Psi^*\Psi, \quad (4.39)$$

where the second step follows from the assumption that the Hamiltonian satisfies (4.29). To make this step clear, the power series representation of  $\hat{U}$  appearing in (4.37) has terms of the form  $(i\hat{H}T/\hbar)^n$  appearing in it. If  $\hat{H}$  satisfies (4.29), these terms have the property

$$\int_V dV ((i\hat{H}T/\hbar)^n\Psi^*)\hat{U}\Psi = \int_V dV \Psi^*((i\hat{H}T/\hbar)^n\hat{U}\Psi), \quad (4.40)$$

so that the powers of  $i$  do not change sign. An operator that obeys  $\hat{U}^*\hat{U} = 1$  is referred to as a *unitary operator*.

## 4.2 Observables and the wave function

As stressed earlier, in the Copenhagen interpretation of quantum mechanics the wave function encapsulates all the measurable aspects of a physical system. A physical quantity that can be measured or *observed* is referred to as an *observable*, and is given the general expression  $\hat{O}(\hat{\mathbf{p}}, \mathbf{x}, t)$ . In wave mechanics observables are assumed to obey the correspondence principle, where a quantum mechanical observable is given by an *operator* obtained from the classical observable  $O(\mathbf{p}, \mathbf{x}, t)$  by the replacement  $\hat{\mathbf{p}} = -i\hbar\nabla$ , while  $\mathbf{x}$  and  $t$  remain unchanged. This is a basic assumption of quantum mechanics and was used to obtain the Schrödinger equation (4.16) from the classical Hamiltonian. For the moment, the case of an explicitly time-dependent observable will be ignored.

The observables of quantum mechanics often involve combinations of momentum and position. The operator form for these two basic observables has a product that satisfies

$$\hat{p}_k x_j \Psi = -i\hbar \frac{\partial}{\partial x_k}(x_j \Psi) = -i\hbar \delta_{jk} \Psi - i\hbar x_j \frac{\partial}{\partial x_k} \Psi = -i\hbar \delta_{jk} \Psi + x_j \hat{p}_k \Psi, \quad (4.41)$$

which gives an important property of these two observables called their *commutator*,

$$x_j \hat{p}_k - \hat{p}_k x_j \equiv [x_j, \hat{p}_k] = i\hbar \delta_{jk}. \quad (4.42)$$

In general, two observables  $\hat{A}$  and  $\hat{B}$  may not commute, so that  $[\hat{A}, \hat{B}] = \hat{A}\hat{B} - \hat{B}\hat{A} \neq 0$ .

It is possible for the wave function to be prepared at  $t = 0$  as an eigenfunction of a chosen observable  $\hat{O}$ . Examples are the separable wave function (4.26) associated with the Hamiltonian operator  $\hat{H}$  or the eigenfunction (4.6) associated with the momentum operator  $\hat{p}$ . For such a case

$$\hat{O}(\hat{p}, \mathbf{x})\Psi_\lambda(\mathbf{x}, t = 0) = \lambda\Psi_\lambda(\mathbf{x}, t = 0), \quad (4.43)$$

where  $\lambda$  is the eigenvalue of  $\hat{O}$  that is associated with the normalized eigenfunction  $\Psi_\lambda$ . It is critical that the eigenvalues of an observable are *real*, and it is therefore important to establish a general criterion for this to be the case. To that end, the *Hermitian adjoint* of the operator  $\hat{O}$  is designated as  $\hat{O}^\dagger$  and is *defined* by the relation

$$\int_V dV \Psi_1^* (\hat{O}^\dagger \Psi_2) = \int_V dV (\hat{O}^* \Psi_1^*) \Psi_2, \quad (4.44)$$

where  $\hat{O}^*$  is the complex conjugate of the operator  $\hat{O}$ . As mentioned, (4.44) *defines* the Hermitian adjoint of  $\hat{O}$ . If the operator  $\hat{O}$  satisfies  $\hat{O}^\dagger = \hat{O}$ , it is said to be *self-adjoint* or, more commonly, *Hermitian*. This was the case for the Hamiltonian (4.19), since (4.29) is simply the requirement that  $\hat{H}^\dagger = \hat{H}$ . The momentum  $\hat{p} = -i\hbar\nabla$  is also Hermitian, since using  $\hat{p}^* = i\hbar\nabla$  and integrating by parts in (4.44), assuming the endpoints do not contribute, shows that  $\hat{p}^\dagger = \hat{p}$ . If the observable  $\hat{O}$  is *Hermitian* it will have real eigenvalues, which follows from (4.43) and using a normalized eigenfunction  $\Psi_\lambda$ ,

$$\begin{aligned} \lambda^* &= \int_V dV (\Psi_\lambda^* \lambda \Psi_\lambda)^* = \int_V dV (\Psi_\lambda^* [\hat{O} \Psi_\lambda])^* = \int_V dV [\hat{O}^* \Psi_\lambda^*] \Psi_\lambda \\ &= \int_V dV \Psi_\lambda^* [\hat{O}^\dagger \Psi_\lambda] = \int_V dV \Psi_\lambda^* [\hat{O} \Psi_\lambda] = \lambda \int_V dV \Psi_\lambda^* \Psi_\lambda = \lambda. \end{aligned} \quad (4.45)$$

In addition, generalizing result (4.33) for a Hermitian Hamiltonian, the eigenfunctions of a Hermitian operator define an orthonormal set since

$$\lambda' \int_V dV \Psi_{\lambda'}^* \Psi_\lambda = \int_V dV (\hat{O}^* \Psi_{\lambda'}^*) \Psi_\lambda = \int_V dV \Psi_{\lambda'}^* (\hat{O} \Psi_\lambda) = \lambda \int_V dV \Psi_{\lambda'}^* \Psi_\lambda. \quad (4.46)$$

For the case that  $\lambda' - \lambda \neq 0$  the integral vanishes and the eigenfunctions are orthogonal.

However, there is a hidden assumption in the proofs (4.45) and (4.46), which is that  $\Psi_\lambda(\mathbf{x}, t)$  remains an eigenfunction of  $\hat{O}$  if it is an eigenfunction at  $t = 0$ . This holds only if the observable  $\hat{O}$  *commutes* with the Hamiltonian, so that  $[\hat{O}, \hat{H}] = 0$ .

This follows from (4.37), which shows that  $\hat{U}(t) = \exp(-i\hat{H}t/\hbar)$  generates the time evolution of  $\Psi_\lambda(\mathbf{x}, 0)$ . Combining this with the definition of the commutator gives

$$\begin{aligned}\hat{O}\Psi_\lambda(\mathbf{x}, t) &= \hat{O}e^{-i\hat{H}t/\hbar}\Psi_\lambda(\mathbf{x}, 0) = \left[\hat{O}, e^{-i\hat{H}t/\hbar}\right]\Psi_\lambda(\mathbf{x}, 0) + e^{-i\hat{H}t/\hbar}\hat{O}\Psi_\lambda(\mathbf{x}, 0) \\ &= \left[\hat{O}, e^{-i\hat{H}t/\hbar}\right]\Psi_\lambda(\mathbf{x}, 0) + \lambda\Psi_\lambda(\mathbf{x}, t).\end{aligned}\quad (4.47)$$

Result (4.47) shows that an eigenfunction of a Hermitian observable  $\hat{O}$  at  $t = 0$  will be an eigenfunction at a later time *only if the observable commutes with the Hamiltonian*. This follows from the expansion of the evolution operator in powers of  $\hat{H}$ . The related concept of a *complete set of commuting observables* will be discussed later in this section. Furthermore, the eigenfunctions of a particular observable may not be complete, since the eigenfunctions may only span a subspace of the full Hilbert space. An example is the Hermitian operator  $\hat{p}_z$ , which has the eigenfunctions  $e^{ip_z z/\hbar}$ . These eigenfunctions are complete only for arbitrary functions of  $z$ . If  $\hat{p}_x$  and  $\hat{p}_y$  are included, then the eigenfunctions are complete for an arbitrary function  $f(\mathbf{x}) = f(x, y, z)$ .

If the wave function at  $t = 0$  coincides with  $\Psi_\lambda(\mathbf{x}, t = 0)$ , the Copenhagen interpretation states that measuring the observable  $\hat{O}$  *experimentally* at  $t = 0$  will yield  $\lambda$ . Although eigenvalues and eigenfunctions are paired in quantum mechanics, it is important to remember that the wave function  $\Psi$  is not necessarily an eigenfunction of *any* observable since it can always be prepared in the simple mixed state  $\Psi = a\Psi_\lambda + b\Psi_{\lambda'}$ . This is not an eigenfunction of  $\hat{O}$ , since  $\hat{O}\Psi = a\lambda\Psi_\lambda + b\lambda'\Psi_{\lambda'}$ , which, in general, cannot be written  $\lambda\Psi$ . This mixed state can be normalized at  $t = 0$  by using the orthonormality of the eigenfunctions, which yields  $|a|^2 + |b|^2 = 1$  in order for condition (3.1) to be satisfied.

The wave function encapsulates the *probabilities* for all possible outcomes of experimentally measuring  $\hat{O}$ . This can be demonstrated by considering an experiment that prepares the particle at  $t = 0$  with one of *two possible values* for the observable  $\hat{O}$ . The probability that it is prepared with the value  $\lambda$  is  $P_\lambda$  and the probability that it is prepared with the value  $\lambda'$  is  $P_{\lambda'}$ . Since these are the only two possible values, these probabilities must satisfy  $P_\lambda + P_{\lambda'} = 1$ . The statistical average for the measured values of the observable  $\hat{O}$  immediately after preparation will be given by  $\langle\hat{O}\rangle = P_\lambda\lambda + P_{\lambda'}\lambda'$  *if the same preparation and measurement is repeated many times*. As in the statistical mechanics expression (2.3), this is referred to as the *expectation value of  $\hat{O}$* . This same expectation value is obtained from the wave function  $\Psi = a\Psi_\lambda + b\Psi_{\lambda'}$  by evaluating the  $t = 0$  expression

$$\begin{aligned}\langle\hat{O}\rangle &= \int_V dV \Psi^* \hat{O} \Psi = \int_V dV \left(a^* \Psi_\lambda^* + b^* \Psi_{\lambda'}^*\right) (a\lambda\Psi_\lambda + b\lambda'\Psi_{\lambda'}) \\ &= |a|^2\lambda + |b|^2\lambda',\end{aligned}\quad (4.48)$$

where the orthonormality of the two eigenfunctions  $\Psi_\lambda$  and  $\Psi_{\lambda'}$  was used. As discussed in chapter 3, the two probabilities are immediately identified as  $P_\lambda = |a|^2$  and  $P_{\lambda'} = |b|^2$ , the modulus squared of the coefficients of the respective

eigenfunctions in the mixed state. Wave function normalization requires  $|a|^2 + |b|^2 = 1$ , so this gives the required relation  $P_\lambda + P_{\lambda'} = 1$ . This interpretation generalizes to an *arbitrary mixed state*. At any later time, the expectation value of  $\hat{O}$  is still given by (4.48) with the understanding that  $\Psi_\lambda(\mathbf{x}, t) = \hat{U}(t)\Psi_\lambda(\mathbf{x}, 0) = e^{-i\hat{H}t/\hbar}\Psi_\lambda(\mathbf{x}, 0)$  is used in (4.48), yielding

$$\langle \hat{O} \rangle(t) = \int_V dV (e^{i\hat{H}t/\hbar}\Psi^*(\mathbf{x}, 0))\hat{O}(\hat{\mathbf{p}}, \mathbf{x}, t)(e^{-i\hat{H}t/\hbar}\Psi(\mathbf{x}, 0)). \quad (4.49)$$

Result (4.39) shows that the functions  $\Psi_\lambda$  remain orthonormal as long as  $\hat{H}$  is Hermitian. However, unless  $\hat{H}$  and  $\hat{O}$  commute, the expectation value (4.49) will be time-dependent.

These techniques can be demonstrated for the momentum observable,  $\hat{\mathbf{p}} = -i\hbar\nabla$ , which has the eigenfunctions  $\Psi_p(\mathbf{x}, t = 0) = e^{i\mathbf{p}\cdot\mathbf{x}/\hbar}/\sqrt{V}$ . These eigenfunctions correspond to the momentum eigenvalue  $\hat{\mathbf{p}}\Psi_p(\mathbf{x}, t = 0) = \mathbf{p}\Psi_p(\mathbf{x}, t = 0)$ . For the case that the particle is free,  $\hat{H} = \hat{\mathbf{p}}^2/2m$ , these eigenfunctions also correspond to the energy eigenvalue  $\hat{H}\Psi_p(\mathbf{x}, t = 0) = (\mathbf{p}^2/2m)\Psi_p(\mathbf{x}, t = 0) = E_p\Psi_p(\mathbf{x}, t = 0)$ . However, if the Hamiltonian is not free, then these momentum eigenfunctions are *not* eigenfunctions of  $\hat{H}$ . Even if the particle is free, it may be prepared at  $t = 0$  in a mixed state corresponding to a continuum of momentum states. The technique used to derive the continuum version of the Dirac delta (3.41) can be adapted in the limit  $V \rightarrow \infty$  to show that such a wave function can be written at  $t = 0$  in the following form,

$$\Psi(\mathbf{x}, 0) = \lim_{V \rightarrow \infty} \frac{1}{\sqrt{V}} \sum_{\mathbf{n}} a_{\mathbf{n}} e^{i\mathbf{p}_{\mathbf{n}}\cdot\mathbf{x}/\hbar} \rightarrow \int \frac{d^3p}{(2\pi\hbar)^3} \tilde{\Psi}(\mathbf{p}) e^{i\mathbf{p}\cdot\mathbf{x}/\hbar}, \quad (4.50)$$

where  $\mathbf{p}_{\mathbf{n}} = \mathbf{n}\pi\hbar/L$  is the three-dimensional version of  $p_n = n\pi\hbar/L$ , so that  $\mathbf{n} = (n_x, n_y, n_z)$  is a vector with arbitrary positive and negative integer components. It is important to note that (4.50) is *not* an eigenfunction of momentum due to the linear superposition of different momentum eigenfunctions.

Because (4.50) is a linear superposition of momentum eigenfunctions, in physics it is often referred to as a *wave packet*. Result (4.50) is identical to the Fourier transform (3.42) of a function, and this is consistent with the Copenhagen interpretation that the particle may be prepared in an arbitrary state as a superposition of momentum states. Using the  $n$ -dimensional extension of the Dirac delta (3.45) allows (4.50) to be inverted,

$$\begin{aligned} \int_V d^n x \Psi(\mathbf{x}, 0) e^{-i\mathbf{p}\cdot\mathbf{x}/\hbar} &= \int d^n k \int_V \frac{d^n x}{(2\pi\hbar)^n} \tilde{\Psi}(\mathbf{k}) e^{i(\mathbf{k}-\mathbf{p})\cdot\mathbf{x}/\hbar} \\ &= \int d^n k \tilde{\Psi}(\mathbf{k}) \delta^{(n)}(\mathbf{k} - \mathbf{p}) = \tilde{\Psi}(\mathbf{p}). \end{aligned} \quad (4.51)$$

Using the units of  $\Psi(\mathbf{x}, 0)$ ,  $d^n p$ , and  $\hbar^n$  shows that  $\tilde{\Psi}(\mathbf{p})$  has the same units as  $\sqrt{V}$ , where  $V$  is the volume element in  $n$  spatial dimensions. This is consistent with (3.44), which shows that  $\tilde{\Psi}(\mathbf{p}) = \sqrt{V} a_{\mathbf{n}}$ , where  $a_{\mathbf{n}}$  is dimensionless.

In order to understand  $\tilde{\Psi}(\mathbf{p})$ , it is noted that wave function normalization requires

$$\begin{aligned} \int_V d^n x |\Psi(\mathbf{x}, 0)|^2 &= \int \frac{d^n p}{(2\pi\hbar)^n} \frac{d^n k}{(2\pi\hbar)^n} \tilde{\Psi}^*(\mathbf{k}) \tilde{\Psi}(\mathbf{p}) \int_V d^n x e^{i(\mathbf{p}-\mathbf{k})\cdot\mathbf{x}/\hbar} \\ &= \int \frac{d^n p}{(2\pi\hbar)^n} d^n k \tilde{\Psi}^*(\mathbf{k}) \tilde{\Psi}(\mathbf{p}) \delta^{(n)}(\mathbf{k} - \mathbf{p}) \\ &= \int \frac{d^n p}{(2\pi\hbar)^n} |\tilde{\Psi}(\mathbf{p})|^2 = 1. \end{aligned} \quad (4.52)$$

The final integral in (4.52) is the continuum limit of the sum over the coefficients (3.48) that resulted from normalizing an arbitrary wave function for a particle in a box. For the case of a mixed state in (4.48), the modulus squared of the coefficients in the superposition was interpreted as the probability of that *eigenvalue* being present in the prepared state. Using the same quantum mechanical interpretation for the Fourier transform of the wave function, it follows from the wave packet (4.50) and result (4.52) that the quantity

$$\mathcal{P}(\mathbf{p}) = \frac{|\tilde{\Psi}(\mathbf{p})|^2}{(2\pi\hbar)^n} \quad (4.53)$$

can be interpreted as the *probability density in momentum space* at  $t = 0$ . In other words,  $d^n p \mathcal{P}(\mathbf{p})$  is the infinitesimal probability for finding the particle in the infinitesimal momentum volume  $d^n p$  located at the momentum  $\mathbf{p}$  at the time  $t = 0$ , just as  $d^n x |\Psi(\mathbf{x}, 0)|^2$  is the probability of finding the particle in the infinitesimal spatial volume  $d^n x$  located at the position  $\mathbf{x}$  at the time  $t = 0$ . In that regard, using the units of  $\hbar$  and  $\tilde{\Psi}(\mathbf{p})$  shows that  $\mathcal{P}(\mathbf{p})$  does indeed have the units of inverse momentum to the  $n$  power. Therefore, the Copenhagen interpretation of (4.50) is that it represents a *superposition of momentum states*. If the coefficient  $\tilde{\Psi}(\mathbf{p})$  vanishes for some  $\mathbf{p}$ , then the initial wave function has no component corresponding to  $\mathbf{p}$ . For such a case, (4.53) gives zero probability for observing the particle with that momentum at  $t = 0$ .

If  $\partial H/\partial t = 0$  the initial wave function (4.50) evolves to  $\Psi(\mathbf{x}, t) = e^{-i\hat{H}t/\hbar}\Psi(\mathbf{x}, 0)$  at the time  $t$ . The result depends on the specific form of the Hamiltonian  $\hat{H}$ . In the case that the particle is free, so that  $\hat{H} = \hat{\mathbf{p}}^2/2m$ , the momentum eigenfunctions are also eigenfunctions of  $\hat{H}$  with the eigenvalue  $E_p = p^2/2m$ . The wave function at a later time is then given by

$$\Psi(\mathbf{x}, t) = \int \frac{d^3 p}{(2\pi\hbar)^3} \tilde{\Psi}(\mathbf{p}) e^{i(\mathbf{p}\cdot\mathbf{x} - E_p t)/\hbar}, \quad (4.54)$$

where  $\tilde{\Psi}(\mathbf{p})$  is the form (4.51) determined from the wave function preparation. This wave packet will be analyzed in the last section of this chapter for a particular  $\tilde{\Psi}(\mathbf{p})$ .

If the  $\Psi_j(\mathbf{x}, t)$  are an orthonormal set of functions obeying (4.34), then their Fourier transforms (4.51) obey

$$\begin{aligned}
\int \frac{d^n p}{(2\pi\hbar)^n} \tilde{\Psi}_j^*(\mathbf{p}) \tilde{\Psi}_k(\mathbf{p}) &= \int_V d^n x d^n y \Psi_j^*(\mathbf{x}, 0) \Psi_k(\mathbf{y}, 0) \int \frac{d^n p}{(2\pi\hbar)^n} e^{i\mathbf{p}\cdot(\mathbf{x}-\mathbf{y})/\hbar} \\
&= \int_V d^n x d^n y \Psi_j^*(\mathbf{x}, 0) \Psi_k(\mathbf{y}, 0) \delta^{(n)}(\mathbf{x} - \mathbf{y}) \\
&= \int_V d^n x \Psi_j^*(\mathbf{x}, 0) \Psi_k(\mathbf{x}, 0) = \delta_{jk},
\end{aligned} \tag{4.55}$$

so that the Fourier transforms are also orthonormal. Using (4.51) shows that the Fourier transforms of a complete set of eigenfunctions satisfy

$$\begin{aligned}
\frac{1}{(2\pi\hbar)^n} \sum_j \tilde{\Psi}_j^*(\mathbf{p}) \tilde{\Psi}_j(\mathbf{k}) &= \int \frac{d^n x}{(2\pi\hbar)^n} d^n y e^{i(\mathbf{p}\cdot\mathbf{x}-\mathbf{k}\cdot\mathbf{y})/\hbar} \sum_j \Psi_j^*(\mathbf{x}, 0) \Psi_j(\mathbf{y}, 0) \\
&= \int \frac{d^n x}{(2\pi\hbar)^n} d^n y e^{i(\mathbf{p}\cdot\mathbf{x}-\mathbf{k}\cdot\mathbf{y})/\hbar} \delta^{(n)}(\mathbf{x} - \mathbf{y}) \\
&= \int \frac{d^n x}{(2\pi\hbar)^n} e^{i(\mathbf{p}-\mathbf{k})\cdot\mathbf{x}/\hbar} = \delta^{(n)}(\mathbf{p} - \mathbf{k}),
\end{aligned} \tag{4.56}$$

where the completeness of the eigenfunctions, given by (4.35), was used.

The expectation value for the momentum at  $t = 0$  is found using the momentum probability density (4.53),

$$\langle \hat{\mathbf{p}} \rangle = \int \frac{d^n p}{(2\pi\hbar)^n} |\tilde{\Psi}(\mathbf{p}, t)|^2 \mathbf{p} = \int \frac{d^n p}{(2\pi\hbar)^n} \tilde{\Psi}^*(\mathbf{p}) \mathbf{p} \tilde{\Psi}(\mathbf{p}). \tag{4.57}$$

Expression (4.57) can be rewritten as

$$\langle \hat{\mathbf{p}} \rangle = -i\hbar \int_V d^n x \Psi^*(\mathbf{x}, 0) [\nabla \Psi(\mathbf{x}, 0)] = \int_V d^n x \Psi^*(\mathbf{x}, 0) [\hat{\mathbf{p}} \Psi(\mathbf{x}, 0)]. \tag{4.58}$$

The equivalence of (4.58) to (4.57) is demonstrated by inserting (4.50) into (4.58) and using the Dirac delta (3.45) that results from the spatial integration. The result (4.58) for the momentum expectation value at  $t = 0$  is identical to the previous definition (4.48).

To recapitulate, the wave function consistent with the initial state of the system,  $\Psi(\mathbf{x}, t = 0)$ , is prepared as a superposition of a set of *complete basis functions* available for the system under consideration. The coefficients of the superposition are found from the inner product of  $\Psi(\mathbf{x}, t = 0)$  with the basis functions. Typically, the basis set consists of the energy eigenfunctions obtained from the Schrödinger equation, but it can be *any complete set* of orthonormal functions. The future wave function is determined by the evolution operator (4.38). At a later time  $t$ , the expectation value of a *Hermitian observable*  $\hat{O}$  is given by (4.49). In the Copenhagen interpretation, the expectation value (4.49) predicts the statistical average of *many measurements* of the observable at the time  $t$  performed subsequent to an *identical preparation of the wave function*.

The definition (4.49) means that the expectation value of a quantum mechanical observable may depend on time *even if the associated classical observable has no*



*explicit time-dependence*. This is because the wave function appearing in (4.49) evolves in time according to (4.37), which is governed by the Hamiltonian appearing in the Schrödinger equation. This *picture* of quantum mechanics, where the time-dependence of a measurement is contained in the wave function, is known as the *Schrödinger picture*, and the associated quantum mechanical observables derived from the correspondence principle are referred to as *Schrödinger picture observables*, denoted  $\hat{O}_S$ .

There is a second picture of quantum mechanics, known as the *Heisenberg picture*, where the observables themselves embody the time-dependence for their expectation value and the wave function is static. A Heisenberg picture observable, denoted  $\hat{O}_H$ , is obtained from the Schrödinger picture observable by the definition

$$\hat{O}_H(t) = e^{i\hat{H}t/\hbar} \hat{O}_S e^{-i\hat{H}t/\hbar} = U^\dagger(t) \hat{O}_S U(t). \quad (4.59)$$

In the Heisenberg picture the expectation value of the observable (4.59) is given by

$$\langle \hat{O} \rangle(t) = \int dV \Psi^*(\mathbf{x}, t=0) \hat{O}_H(t) \Psi(\mathbf{x}, t=0). \quad (4.60)$$

The Heisenberg picture expression (4.60) is identical to the Schrödinger picture expression (4.49) since (4.37) gives

$$\begin{aligned} \langle \hat{O} \rangle(t) &= \int dV \Psi^*(\mathbf{x}, 0) \hat{O}_H(t) \Psi(\mathbf{x}, 0) \\ &= \int dV \Psi^*(\mathbf{x}, 0) (U^\dagger(t) \hat{O}_S U(t)) \Psi(\mathbf{x}, 0) \\ &= \int dV (U^\dagger(t) \Psi^*(\mathbf{x}, 0)) \hat{O}_S (U(t) \Psi(\mathbf{x}, 0)) \\ &= \int dV \Psi^*(\mathbf{x}, t) \hat{O}_S \Psi(\mathbf{x}, t). \end{aligned} \quad (4.61)$$

Using  $d\Psi/dt = (\hat{H}/i\hbar)\Psi$  and  $d\Psi^*/dt = -(\hat{H}/i\hbar)\Psi^*$ , result (4.61) yields

$$\frac{d}{dt} \langle \hat{O} \rangle = \left\langle \frac{\partial \hat{O}}{\partial t} \right\rangle + \frac{1}{i\hbar} \langle [\hat{O}, \hat{H}] \rangle, \quad (4.62)$$

which holds in either picture as long as the Hamiltonian is Hermitian.

The result (4.62) allows quantum mechanical expectation values to satisfy relationships formally similar to Newtonian mechanics, a result known as *Ehrenfest's theorem*. As an example, since  $[\hat{H}, \hat{H}] = 0$ , result (4.62) shows that

$$\frac{d}{dt} \langle \hat{H} \rangle = \left\langle \frac{\partial \hat{H}}{\partial t} \right\rangle. \quad (4.63)$$

The right-hand side of (4.63) vanishes if  $\hat{H}$  has no explicit time dependence. For such a case, the expectation value of the Hamiltonian is time-independent, and this is the quantum mechanical version of the classical conservation of energy (1.3). Working in one dimension for simplicity, choosing  $\hat{O} = x$  in (4.62), and assuming that the Hamiltonian has the form  $\hat{H} = \hat{p}^2/2m + U(x)$  gives

$$[x, \hat{H}] = \left[ x, \frac{\hat{p}^2}{2m} \right] = \frac{1}{2m}(x\hat{p}^2 - \hat{p}^2x) = \frac{1}{2m}([\hat{p}, x]\hat{p} + \hat{p}[x, \hat{p}]) = i\hbar\frac{\hat{p}}{m}, \quad (4.64)$$

where (4.42) and  $[x, U(x)] = 0$  were used. Inserting this result into (4.62) gives

$$\frac{d}{dt}\langle x \rangle = \frac{\langle \hat{p} \rangle}{m}, \quad (4.65)$$

a form identical to the classical relationship (1.8) for the same Hamiltonian. Similarly, choosing  $\hat{O} = \hat{p}$  and using the same form for the Hamiltonian gives

$$[\hat{p}, \hat{H}] = [\hat{p}, U(x)] = -i\hbar\frac{\partial U}{\partial x}, \quad (4.66)$$

so that (4.62) becomes

$$\frac{d}{dt}\langle \hat{p} \rangle = -\left\langle \frac{\partial U}{\partial x} \right\rangle, \quad (4.67)$$

which is identical in form to Newton's second law (1.11). In order to insure that this holds for all observables, result (4.62) is used to extend the correspondence principle by noting its formal similarity to the Poisson bracket formulation (1.15) of classical mechanics. If  $A$  and  $B$  are two classical observables, then the expectation value of their *quantum mechanical commutator* will be identical to the result obtained from the *classical Poisson bracket* through the substitution

$$\{A, B\} \rightarrow \frac{1}{i\hbar}\langle [\hat{A}, \hat{B}] \rangle. \quad (4.68)$$

The result (4.42) shows that this correspondence holds for the observables  $\mathbf{p}$  and  $\mathbf{x}$  as well as for the general time evolution result (4.62) involving the Hamiltonian.

In this regard, it is important to note that the commutator  $[x_j, \hat{p}_k] = i\hbar\delta_{jk}$  is clearly *not* Hermitian, even though both  $x_j$  and  $\hat{p}_k$  are Hermitian. In general, the product  $\hat{A}\hat{B}$  of two Hermitian operators, satisfying  $\hat{A}^\dagger = \hat{A}$  and  $\hat{B}^\dagger = \hat{B}$ , *may not be Hermitian*. In order to be Hermitian, their product must satisfy (4.45), which means that

$$\int_V d^3x \Psi^*(\mathbf{x}, t)[\hat{A}\hat{B}\Psi(\mathbf{x}, t)] = \int_V d^3x[\hat{A}^*\hat{B}^*\Psi^*(\mathbf{x}, t)]\Psi(\mathbf{x}, t). \quad (4.69)$$

Because the individual operators are Hermitian, it follows that

$$\begin{aligned} \int_V d^3x \Psi^*(\mathbf{x}, t)[\hat{A}\hat{B}\Psi(\mathbf{x}, t)] &= \int_V d^3x \Psi^*(\mathbf{x}, t)[\hat{A}[\hat{B}\Psi(\mathbf{x}, t)]] \\ &= \int_V d^3x[\hat{A}^*\Psi^*(\mathbf{x}, t)][\hat{B}\Psi(\mathbf{x}, t)] \\ &= \int_V d^3x[\hat{B}^*\hat{A}^*\Psi^*(\mathbf{x}, t)]\Psi(\mathbf{x}, t). \end{aligned} \quad (4.70)$$

This demonstrates a general result,  $(\hat{A}\hat{B})^\dagger = \hat{B}^\dagger\hat{A}^\dagger$ . In order for (4.69) and (4.70) to be identical, the product of two Hermitian operators must satisfy  $\hat{A}\hat{B} = (\hat{A}\hat{B})^\dagger = \hat{B}^\dagger\hat{A}^\dagger = \hat{B}\hat{A}$ . In other words, if the two operators *do not commute*, their product is not Hermitian. This is made clear by writing the product as

$$\hat{A}\hat{B} = \frac{1}{2}(\hat{A}\hat{B} - \hat{B}\hat{A}) + \frac{1}{2}(\hat{A}\hat{B} + \hat{B}\hat{A}) \equiv \frac{1}{2}[\hat{A}, \hat{B}] + \frac{1}{2}\{\hat{A}, \hat{B}\}, \quad (4.71)$$

where  $\{\hat{A}, \hat{B}\}$  now stands for the *anticommutator*  $\hat{A}\hat{B} + \hat{B}\hat{A}$ . Using  $(\hat{A}\hat{B})^\dagger = \hat{B}^\dagger\hat{A}^\dagger$ , the Hermitian conjugate of the commutator of two Hermitian operators satisfies  $[\hat{A}, \hat{B}]^\dagger = [\hat{B}^\dagger, \hat{A}^\dagger] = [\hat{B}, \hat{A}] = -[\hat{A}, \hat{B}]$ . This reveals the commutator of two Hermitian operators is *antiHermitian*, changing signs under Hermitian conjugation, whereas the anticommutator is Hermitian since  $\{\hat{A}, \hat{B}\}^\dagger = \{\hat{A}^\dagger, \hat{B}^\dagger\} = \{\hat{A}, \hat{B}\}$ . The expectation value of an antiHermitian operator  $\hat{O}$ , satisfying  $\hat{O}^\dagger = -\hat{O}$ , is given by

$$\begin{aligned} \langle \hat{O} \rangle^* &= \int_V dV [\hat{O}^* \Psi^*(\mathbf{x}, t)] \Psi(\mathbf{x}, t) = \int_V dV \Psi^*(\mathbf{x}, t) [\hat{O}^\dagger \Psi(\mathbf{x}, t)] \\ &= - \int_V dV \Psi^*(\mathbf{x}, t) [\hat{O} \Psi(\mathbf{x}, t)] = -\langle \hat{O} \rangle \end{aligned} \quad (4.72)$$

showing that its expectation value is *pure imaginary*. The anticommutator of two Hermitian operators is Hermitian, so it has a real expectation value. Therefore, (4.71) shows that the product of two Hermitian operators has an expectation value that breaks into a real and an imaginary part. The imaginary part vanishes *only if the two operators commute*.

In the Copenhagen interpretation an observable must yield a real expectation value, and this requires the operator representing the observable to be Hermitian. If the observable chosen for measurement contains products of Hermitian observables, it is necessary that the product involves only Hermitian operators that commute with each other. Otherwise, (4.71) shows that the product breaks into the sum of Hermitian and antiHermitian pieces, yielding a complex expectation value. Therefore, the Copenhagen interpretation of quantum mechanics states that an experiment must choose a *complete set of commuting observables*, hereafter designated CSCO, in order for real expectation values to be obtained for their products. This set of observables contains only Hermitian observables that commute with all other members of the set. Observables that do not commute with any members of the set cannot be measured simultaneously since their products with the members of the set are not Hermitian. Since the fundamental Hermitian observables  $\mathbf{x}$  and  $\hat{\mathbf{p}}$  do not commute, this will almost always require a choice for which observables are to be included in the CSCO. There are additional problems with simultaneously measuring observables that do not commute, and this is discussed in detail in the next section where the Heisenberg uncertainty principle is derived.

An extremely important property of a CSCO is that its members share a *common set of eigenfunctions*. This is readily seen by considering two members of a CSCO,

denoted  $\hat{A}$  and  $\hat{B}$ . If  $\Psi_\lambda$  is an eigenfunction of  $\hat{A}$ , such that  $\hat{A}\Psi_\lambda = \lambda\Psi_\lambda$ , then it follows that

$$\hat{A}(\hat{B}\Psi_\lambda) = \hat{B}(\hat{A}\Psi_\lambda) = \lambda(\hat{B}\Psi_\lambda). \quad (4.73)$$

Result (4.73) shows that  $\hat{B}\Psi_\lambda$  must be proportional to  $\Psi_\lambda$ , so that  $\hat{B}\Psi_\lambda = \lambda'\Psi_\lambda$ , where  $\lambda'$  is the constant of proportionality. This immediately reveals that the constant of proportionality is simply the eigenvalue of  $\hat{B}$  associated with the eigenfunction  $\Psi_\lambda$ . A CSCO therefore shares *a set of eigenfunctions in common*. If these eigenfunctions are complete, then any  $L_2$  wave function may be prepared and arbitrary products of members of the CSCO will have well-defined expectation values. An example of this are the momentum eigenfunctions (4.6). For a free particle, the Hamiltonian  $\hat{H}$  and the momentum  $\hat{p}$  commute and therefore share these eigenfunctions. If the CSCO contains the Hamiltonian, then (4.47) shows that all the observables in the CSCO will have time-independent expectation values.

### 4.3 The Heisenberg uncertainty principle

If two observables do not commute it is still possible to devise an experiment that will attempt to measure both of them simultaneously. For example, passing electrons through a slit of width  $\Delta x$  determines their position up to the slit width  $\Delta x$ . This can be accompanied by a measurement of the range of their momentum *parallel* to the slit, designated  $\Delta p$ . A simple way to find this range of momentum is to measure the angle  $\theta$  at which each electron strikes the screen a distance  $L$  behind the slit. If the velocity perpendicular to the slit is  $v$ , then the time to arrive at the screen is  $T = L/v$ . The distance  $D$  traveled parallel to the slit before striking the screen is  $D \approx L \sin \theta$ , and so the momentum *parallel* to the slit must have been  $\Delta p = m \Delta v \approx m(D/T) = mL \sin \theta/T = mv \sin \theta$ . However, the wave aspect of the electron undergoes diffraction when passing through the slit, spreading the beam so that their impact on the screen occurs within the first diffraction minimum of (1.51),  $\sin \theta = \lambda/\Delta x = h/(mv \Delta x)$ , where the de Broglie wavelength (2.18) was used. Inserting this into the expression for  $\Delta p$  gives the result  $\Delta x \Delta p \approx h$ .

Heisenberg interpreted the slit width as the *uncertainty in the position* of the electron as it passes through the slit, while  $\Delta p$  represents the *uncertainty in the momentum* parallel to the slit *subsequent* to passing through the slit. If  $\Delta x \rightarrow 0$ , reducing the uncertainty in position to zero, it results in  $\Delta p \rightarrow \infty$ , and the uncertainty in the subsequent momentum of the electron parallel to the slit becomes arbitrarily large. The uncertainty in momentum can be reduced to zero,  $\Delta p \rightarrow 0$ , only if  $\Delta x \rightarrow \infty$ , which requires the slit width or uncertainty in position to be arbitrarily large. It is important to note that this occurs even if the source of the electrons is far from the slit, which would classically preclude electrons with any initial momentum parallel to the slit from passing through. As a result, it is the *simultaneous observation of both momentum and position* that is involved in this effect. It will now be shown that the two observables, momentum and position, cannot be measured simultaneously to arbitrary accuracy because they do not commute.

In order to demonstrate this, it is necessary to quantify mathematically the uncertainty, denoted  $\Delta\hat{O}$ , in the measurement of an observable  $\hat{O}$ , and then to derive

a relationship between the uncertainties for observables that involves their commutator. The definition of  $\Delta\hat{O}$  is borrowed from statistics, where deviation from the norm is an important aspect of a statistical correlation. The operator  $\hat{O} - \langle\hat{O}\rangle$  clearly has an expectation value of zero if many measurements are performed after preparing the same wave function since  $\langle\hat{O} - \langle\hat{O}\rangle\rangle = \langle\hat{O}\rangle - \langle\langle\hat{O}\rangle\rangle = \langle\hat{O}\rangle - \langle\hat{O}\rangle = 0$ . However, for an individual measurement it will not necessarily vanish, and the uncertainty  $\Delta\hat{O}$  in measuring  $\hat{O}$  is found by averaging the absolute value of  $\hat{O} - \langle\hat{O}\rangle$  over many measurements. This prevents positive and negative results from cancelling. This is achieved by finding the expectation value of its square and then taking the square root. The result is

$$\Delta\hat{O} = \sqrt{\langle(\hat{O} - \langle\hat{O}\rangle)^2\rangle} = \sqrt{\langle\hat{O}^2\rangle - 2\langle\hat{O}\rangle\langle\hat{O}\rangle + \langle\langle\hat{O}\rangle\rangle^2} = \sqrt{\langle\hat{O}^2\rangle - \langle\hat{O}\rangle^2}. \quad (4.74)$$

This definition of uncertainty vanishes if the wave function is an eigenfunction or *pure state* of the observable, as in (4.43), since  $\langle\hat{O}\rangle = \lambda$  and  $\langle\hat{O}^2\rangle = \lambda^2$ . This is consistent with the idea that a *pure state* of the observable should have no uncertainty in the measurement of that observable. Definition (4.74) is now combined with the Cauchy–Schwarz inequality of (3.24). If  $\hat{A}$  and  $\hat{B}$  are two Hermitian operators, then two new Hermitian operators are defined as  $\hat{\alpha} = \hat{A} - \langle\hat{A}\rangle$  and  $\hat{\beta} = \hat{B} - \langle\hat{B}\rangle$ . The two functions appearing in the Cauchy–Schwarz inequality (3.24) are chosen to be derived from the wave function  $\Psi$  by defining  $f = \hat{\alpha}\Psi$  and  $g = \hat{\beta}\Psi$ . Since  $\langle\hat{\alpha}^2\rangle = (\Delta\hat{A})^2$  and  $\langle\hat{\beta}^2\rangle = (\Delta\hat{B})^2$  this yields

$$\begin{aligned} \int_V dV \hat{\alpha}^*\Psi^*\hat{\alpha}\Psi \int_V dV \hat{\beta}^*\Psi^*\hat{\beta}\Psi &\geq \left| \int_V dV \hat{\alpha}^*\Psi^*\hat{\beta}\Psi \right|^2 \\ \implies \int_V dV \Psi^*\hat{\alpha}^2\Psi \int_V dV \Psi^*\hat{\beta}^2\Psi &\geq \left| \int_V dV \Psi^*\hat{\alpha}\hat{\beta}\Psi \right|^2 \\ \implies (\Delta\hat{A})^2(\Delta\hat{B})^2 &\geq \frac{1}{4} |\langle[\hat{A}, \hat{B}]\rangle + \langle\{\hat{\alpha}, \hat{\beta}\}\rangle|^2, \end{aligned} \quad (4.75)$$

where the Hermitian nature of  $\hat{\alpha}$  and  $\hat{\beta}$  was used in the second step, while the third step used (4.71), definition (4.74), and  $\langle[\hat{\alpha}, \hat{\beta}]\rangle = \langle[\hat{A}, \hat{B}]\rangle$ . Since the expectation value of the commutator is pure imaginary and the expectation value of the anticommutator is pure real, expression (4.75) gives

$$\Delta\hat{A} \Delta\hat{B} \geq \frac{1}{2} \sqrt{|\langle[\hat{A}, \hat{B}]\rangle|^2 + |\langle\{\hat{\alpha}, \hat{\beta}\}\rangle|^2}. \quad (4.76)$$

Result (4.76) is the formal statement of the *Heisenberg uncertainty principle*. If two operators do not commute, then the first term in the square root of (4.76) will not vanish and the product of the uncertainties in their measurements will never be zero. The second term in the square root can be expanded using the definitions of  $\hat{\alpha}$  and  $\hat{\beta}$  to give

$$|\langle\{\hat{\alpha}, \hat{\beta}\}\rangle|^2 = |\langle\{\hat{A} - \langle\hat{A}\rangle, \hat{B} - \langle\hat{B}\rangle\}\rangle|^2 = |\langle\{\hat{A}, \hat{B}\}\rangle - 2\langle\hat{A}\rangle\langle\hat{B}\rangle|^2. \quad (4.77)$$

Even if two operators commute, it is possible that the wave function for the system is such that (4.77) is not zero. If the system is prepared in a state such that (4.77) vanishes, then the Heisenberg uncertainty principle allows the expectation values for commuting observables to be measured simultaneously to arbitrary accuracy. A wave function for which (4.77) vanishes is referred as a *minimum uncertainty wave packet*, and will be discussed again in the next section. For the choice of position and momentum, if (4.77) vanishes then (4.76) gives the *minimum uncertainty relation*

$$\Delta x_j \Delta \hat{p}_k \geq \frac{\hbar}{2} \delta_{jk}. \quad (4.78)$$

The uncertainty principle (4.78) for momentum and position explains why the volume  $V$  in the momentum eigenfunction (4.6) must be arbitrarily large in order to be consistent with a vanishing probability flux into and out of the volume  $V$ . Since the particle is contained in the volume  $V = L^3$ , the length  $L$  determines the uncertainty in the position, so that  $L \approx \Delta x_k$ . However, in the case of a momentum eigenfunction it follows that  $\Delta p_k = 0$  since  $\langle \hat{p}_k^2 \rangle = \langle \hat{p}_k \rangle^2$ . As a result,  $\Delta x_k \rightarrow \infty$ , and this requires an arbitrarily large volume.

It is instructive to evaluate  $\Delta x$  and  $\Delta p$  for the simple case that the wave function represents the ground state of a particle in a one-dimensional box of width  $L$ , given at  $t = 0$  by (3.49) with  $n = 1$ . The uncertainty squared in the particle's position is then

$$\begin{aligned} (\Delta x)^2 &= \langle x^2 \rangle - \langle x \rangle^2 = \frac{2}{L} \int_0^L dx x^2 \sin^2\left(\frac{\pi x}{L}\right) - \left(\frac{2}{L} \int_0^L dx x \sin^2\left(\frac{\pi x}{L}\right)\right)^2 \\ &= \frac{2}{L} \left(\frac{L^3}{6} - \frac{L^3}{4\pi^2}\right) - \left(\frac{2}{L} \frac{L^2}{4}\right)^2 = L^2 \left(\frac{1}{12} - \frac{1}{2\pi^2}\right), \end{aligned} \quad (4.79)$$

so that the uncertainty in position is  $\Delta x = (L/2\pi)\sqrt{(\pi^2 - 6)/3}$ . Similarly, for momentum

$$\begin{aligned} (\Delta \hat{p})^2 &= \langle \hat{p}^2 \rangle - \langle \hat{p} \rangle^2 = \frac{2}{L} \int_0^L dx \frac{\pi^2 \hbar^2}{L^2} \sin^2\left(\frac{\pi x}{L}\right) \\ &\quad - \left(\frac{2}{L} \int_0^L dx \left(-\frac{i\pi\hbar}{L}\right) \sin\left(\frac{\pi x}{L}\right) \cos\left(\frac{\pi x}{L}\right)\right)^2 = \frac{\pi^2 \hbar^2}{L^2}, \end{aligned} \quad (4.80)$$

so that  $\Delta \hat{p} = \pi\hbar/L$ . The product  $\Delta x \Delta \hat{p} = (\hbar/2)\sqrt{(\pi^2 - 6)/3}$  satisfies the inequality (4.78) since the square root is  $\approx 1.14$ , slightly larger than one.

Quantum mechanics was developed to be consistent with Newtonian physics, where the time  $t$  is a parameter rather than an observable. There is *no* correspondence principle that turns time into a quantum mechanical operator  $\hat{T}$ . However, there is an important uncertainty relation involving time for the case that the observable  $\hat{O}$  has no explicit time-dependence,  $\partial \hat{O} / \partial t = 0$ , but does *not* commute with the Hamiltonian  $\hat{H}$ . For such a case, result (4.62) gives  $\langle [\hat{H}, \hat{O}] \rangle = i\hbar d\langle \hat{O} \rangle / dt$ .

Assuming the anticommutator (4.77) vanishes, the uncertainty relationship (4.76) for the energy  $\hat{H}$  and the observable  $\hat{O}$  is given by

$$\Delta E \Delta O \geq \frac{\hbar}{2} \left| \frac{d}{dt} \langle \hat{O} \rangle \right|. \quad (4.81)$$

The more rapidly the expectation value of the observable is changing in time, the greater the uncertainty in simultaneous measurements of the system's energy and the observable will be. The result (4.81) can be restated by defining a time interval  $\Delta t$  from  $d\langle \hat{O} \rangle / dt$ ,

$$\Delta t \equiv \frac{\Delta O}{\left| \frac{d}{dt} \langle \hat{O} \rangle \right|}, \quad (4.82)$$

corresponding to the time interval during which the observable's expectation value significantly changes. This is the case if the expectation value of the chosen observable oscillates with some angular frequency  $\omega$ , since then  $\Delta t \approx 1/\omega$ . Making the identification (4.82) gives a second commonly used uncertainty relation,

$$\Delta E \Delta t \geq \hbar/2. \quad (4.83)$$

It should be remembered that the time interval  $\Delta t$  has an interpretation that depends on the choice for the observable  $\hat{O}$  and the system with which it is associated.

An instructive example of (4.83) is obtained by letting the observable be the position  $x$  of a *free particle* moving in one dimension. Ehrenfest's theorem allows  $d\langle x \rangle / dt$  to be interpreted as the average velocity of the particle, denoted  $v$ . For this case (4.82) gives  $\Delta t = \Delta x / v$ , which is interpreted as the time interval a particle detector must interact with the particle in order to compensate for the uncertainty  $\Delta x$  in the particle's position. In order for  $\Delta t$  to be small, there must be a small uncertainty in the particle's position, and this requires a large uncertainty in the particle's momentum, which in turn creates a large uncertainty in the particle's energy. This interpretation is consistent with the dynamics of a free particle, where the uncertainty in the energy is a function of  $\Delta p$ , since  $\Delta E = \Delta(p^2/2m) = p \Delta p / m$  for the case  $\Delta p \ll p$ . Combining the two expressions gives  $\Delta E \Delta t = p \Delta p \Delta x / mv = \Delta p \Delta x \geq \hbar/2$ , the previous uncertainty relation.

The result (4.83) is commonly extended to all cases involving a finite observation time, where it is interpreted as relating the time interval of observation  $\Delta t$  to the uncertainty  $\Delta E$  in the energy measured for the particle or system. For example, if an unstable particle decays with a lifetime of  $\tau$ , this places an upper limit on the time it can be observed, so that the minimum uncertainty in the particle's energy is  $\Delta E \approx \hbar/2\tau$ . The rho meson has a lifetime of  $\tau \approx 4.5 \times 10^{-24}$  s, corresponding to an uncertainty in its observed energy of  $\approx 1.2 \times 10^{-11}$  J or  $\approx 7 \times 10^7$  eV. This also applies to a quantum state that *decays* to a lower energy state with a lifetime of  $\tau$ . In this case,  $\Delta t = \tau$  creates an uncertainty in the energy associated with the transition, and if a photon is emitted during the transition there will be a range of frequencies,  $\Delta\omega = \Delta E / \hbar \approx 1/2\tau$ , associated with the decay process.

## 4.4 Wave packets

In the Copenhagen interpretation of quantum mechanics the wave function is prepared at some initial time and then evolves in time according to (4.37) until an observation forces the wave function to collapse onto some specific possible values for a complete set of commuting observables. This section discusses consequences of wave function preparation.

Wave function preparation can cause the simultaneous measurements of commuting observables to be accompanied by uncertainties unless their anticommutator (4.77) vanishes, and so it is important to determine the wave functions for which the anticommutator vanishes. This starts by noting that the Cauchy–Schwarz inequality used to derive the uncertainty relation becomes an equality, therefore corresponding to the case of *minimum uncertainty*, when the two functions appearing in (3.24) are proportional to each other. In the derivation of the uncertainty principle this requires that  $\hat{\alpha}\Psi = z\hat{\beta}\Psi$  or  $\hat{\beta}\Psi = \hat{\alpha}\Psi/z$ , where  $z$  is some complex constant. Taking the complex conjugate gives  $\hat{\alpha}^*\Psi^* = z^*\hat{\beta}^*\Psi^*$ . Using  $(\hat{\alpha}\hat{\beta})^\dagger = \hat{\beta}^\dagger\hat{\alpha}^\dagger$  and the Hermitian nature of  $\hat{\beta}$  gives  $(\hat{\beta}^2)^\dagger = (\hat{\beta}^\dagger)^2 = \hat{\beta}^2$ , so that

$$\begin{aligned} \int dV \Psi^*(\hat{\alpha}\hat{\beta}\Psi) &= \int dV (\hat{\beta}^*\hat{\alpha}^*\Psi^*)\Psi = z^* \int dV (\hat{\beta}^{*2}\Psi^*)\Psi \\ &= z^* \int dV \Psi(\hat{\beta}^2\Psi) = \frac{z^*}{z} \int dV \Psi^*(\hat{\beta}\hat{\alpha}\Psi). \end{aligned} \quad (4.84)$$

If  $z$  is pure imaginary, so that  $z = i\lambda$ , then  $z^*/z = -1$  and (4.84) gives the desired result  $\langle\{\hat{\alpha}, \hat{\beta}\}\rangle = 0$ , resulting in *minimum uncertainty*. In the case that momentum and position are chosen for  $\hat{\alpha}$  and  $\hat{\beta}$ , the condition for a *minimum uncertainty* wave function is then

$$\hat{\alpha}\Psi = i\lambda\hat{\beta}\Psi \implies (-i\hbar\nabla - \langle\hat{p}\rangle)\Psi = i\lambda(x - \langle x\rangle)\Psi, \quad (4.85)$$

where  $\lambda$  is an arbitrary *real* constant.

A particularly useful minimum uncertainty wave function is given by preparing the initial wave function as a *Gaussian wave packet*, given in one dimension by

$$\Psi(x, t = 0) = \frac{1}{\sqrt{\gamma\sqrt{\pi}}} e^{-\frac{1}{2}(x-x_o)^2/\gamma^2}, \quad (4.86)$$

where  $\gamma$  is an arbitrary real constant with the units of length. It is straightforward to show (4.86) satisfies the necessary normalization condition (3.1) for the case that the system is defined over the spatial interval  $(-\infty, \infty)$ . The initial expectation values for position and momentum are given by using the probability density  $\mathcal{P}(x, 0) = |\Psi|^2$  to find

$$\begin{aligned} \langle x \rangle &= \frac{1}{\gamma\sqrt{\pi}} \int_{-\infty}^{\infty} dx x e^{-(x-x_o)^2/\gamma^2} = x_o, \\ \langle \hat{p} \rangle &= \frac{i\hbar}{\gamma^3\sqrt{\pi}} \int_{-\infty}^{\infty} dx (x - x_o) e^{-(x-x_o)^2/\gamma^2} = 0. \end{aligned} \quad (4.87)$$



These results show that condition (4.85) is satisfied with  $\lambda = \hbar/\gamma^2$ , verifying that (4.86) is a *minimum uncertainty wave function*. Additional expectation values are given by

$$\begin{aligned}\langle \hat{p}^2 \rangle &= \frac{1}{\gamma\sqrt{\pi}} \int_{-\infty}^{\infty} dx \left( \frac{\hbar^2}{\gamma^2} - \frac{\hbar^2}{\gamma^4} (x - x_o)^2 \right) e^{-(x-x_o)^2/\gamma^2} = \frac{\hbar^2}{2\gamma^2}, \\ \langle x^2 \rangle &= \frac{1}{\gamma\sqrt{\pi}} \int_{-\infty}^{\infty} dx x^2 e^{-(x-x_o)^2/\gamma^2} = \frac{1}{2}\gamma^2 + x_o^2.\end{aligned}\quad (4.88)$$

Combining results (4.87) and (4.88) gives  $\Delta p = \sqrt{\langle \hat{p}^2 \rangle - \langle p \rangle^2} = \hbar/(\sqrt{2}\gamma)$  and  $\Delta x = \sqrt{\langle x^2 \rangle - \langle x \rangle^2} = \gamma/\sqrt{2}$ , corresponding to the minimum uncertainty relation  $\Delta p \Delta x = \hbar/2$ . The Gaussian wave packet (4.86) is initially peaked around the position of maximum probability,  $x_o$ , with the parameter  $\gamma$  determining how sharply it is peaked. The limit  $\gamma \rightarrow 0$  corresponds to  $\Delta x = 0$ , so that the probability density must therefore approach the behavior of the Dirac delta,  $\delta(x - x_o)$ . In fact, it can be shown that an alternative representation of the Dirac delta is given by the  $\gamma \rightarrow 0$  limit of  $|\Psi|^2$  for the case of (4.86). At the other extreme, the limit  $\gamma \rightarrow \infty$  corresponds to  $\Delta p = 0$ , for which the probability density for position approaches a constant over all space, like the momentum eigenfunction (4.6) for the case  $p \rightarrow 0$ .

It is instructive to examine the time evolution of the wave function (4.86) for the case that the particle is free. For that case, the Gaussian is understood as a mixed state of momentum eigenfunctions (4.6) with a momentum distribution  $\tilde{\Psi}(p)$  determined from the initial state (4.86) using (4.51). Setting  $x_o = 0$  and staying in one dimension gives

$$\tilde{\Psi}(p) = \frac{1}{\sqrt{\gamma\sqrt{\pi}}} \int_{-\infty}^{\infty} dx e^{-\frac{1}{2}x^2/\gamma^2} e^{-ipx/\hbar} = \sqrt{2\gamma\sqrt{\pi}} e^{-\frac{1}{2}\gamma^2 p^2/\hbar^2}, \quad (4.89)$$

so that the momentum distribution is also Gaussian. Since the particle is free the time-dependent wave packet that satisfies the Schrödinger equation (4.16) and that coincides with the preparation (4.86) at  $t = 0$  is obtained by writing

$$\Psi(x, t) = \int_{-\infty}^{\infty} \frac{dp}{2\pi\hbar} \tilde{\Psi}(p) e^{i(p x - E_p t)/\hbar}, \quad (4.90)$$

where  $\tilde{\Psi}(p)$  is given by (4.89) and  $E_p = p^2/2m$ . In effect, the wave function is a linear superposition of free particle momentum eigenfunctions, but is not itself an eigenfunction of momentum. Inserting (4.89) into (4.91) gives

$$\Psi(x, t) = \sqrt{2\gamma\sqrt{\pi}} \int_{-\infty}^{\infty} \frac{dp}{2\pi\hbar} e^{-\frac{1}{2}\gamma^2 p^2/\hbar^2} e^{i(p x - E_p t)/\hbar}. \quad (4.91)$$

Since  $E_p = p^2/2m$ , the integral (4.91) is also Gaussian in  $p$  and its evaluation gives

$$\Psi(x, t) = \sqrt{\frac{\gamma}{\sqrt{\pi} \left( \gamma^2 + \frac{i\hbar t}{m} \right)}} e^{-\frac{1}{2}x^2 / \left( \gamma^2 + \frac{i\hbar t}{m} \right)}. \quad (4.92)$$

From (4.92) the time-dependent spatial probability density is given by

$$\mathcal{P}(x, t) = |\Psi(x, t)|^2 = \frac{1}{\sqrt{\pi}} \left( \gamma^2 + \frac{\hbar^2 t^2}{m^2 \gamma^2} \right)^{-1/2} \exp \left( -\frac{x^2}{\left( \gamma^2 + \frac{\hbar^2 t^2}{m^2 \gamma^2} \right)} \right). \quad (4.93)$$

The result (4.93) shows that the probability density remains a normalized Gaussian distribution, but the  $\gamma^2$  in (4.86) is replaced by the time-dependent shape factor  $\Gamma^2(t)$ , where

$$\Gamma^2(t) = \gamma^2 + \frac{\hbar^2 t^2}{m^2 \gamma^2}. \quad (4.94)$$

Even if  $\gamma \approx 0$  initially, so that the initial uncertainty in location is  $\Delta x \approx 0$ , as time passes the probability density (4.93) becomes less sharply peaked. The width of the Gaussian grows, signaling a growing uncertainty in the location determined by the increasing value of  $\Gamma$  in (4.94). After sufficient time has passed, the argument of the exponential tends to zero and the probability density becomes the same everywhere. This is a direct outgrowth of the uncertainty principle. If the particle's initial uncertainty in position is very small, i.e.  $\gamma \approx 0$ , it will possess a large initial uncertainty in momentum,  $\Delta p(0) = \hbar/(\sqrt{2} \gamma)$ , and the uncertainty in its position will therefore grow rapidly since Newtonian dynamics gives  $\Delta x \approx \Delta p(0) t/m$ . This is consistent with (4.94) since at later times  $\Delta x(t) = \Gamma(t) \approx \hbar t/m\gamma \approx \Delta p(0) t/m$  for  $\gamma \approx 0$ . The spreading of the wave packet is also a reflection that the individual momentum eigenfunctions that make up the wave packet have phases that travel at different speeds since  $p x - E_p t = p(x - (p/2m)t)$ . The constituent momentum eigenfunctions therefore separate, resulting in a dispersing wave packet.