3

A First Approach to Quantum Mechanics

In this chapter, we try to understand the main ideas of quantum mechanics. In quantum mechanics, the outcome of a measurement cannot—even in principle—be predicted beforehand; only the *probabilities* for the outcome of the measurement can be predicted. These probabilities are encoded in a *wave function*, which is a function of a *position* variable $\mathbf{x} \in \mathbb{R}^n$. The square of the absolute value of the wave function encodes the probabilities for the position of the particle. Meanwhile, the probabilities for the momentum of the particle are encoded in the *frequency of oscillation* of the wave function. The probabilities can be described using the *position operator* and the *momentum operator*. The time-evolution of the wave function is described by the *Hamiltonian operator*, which is analogous to the Hamiltonian (or energy) function in Hamilton's equations.

3.1 Waves, Particles, and Probabilities

There are two key ingredients to quantum theory, both of which arose from experiments. The first ingredient is *wave–particle duality,* in which objects are observed to have both wavelike and particlelike behavior. Light, for example, was thought to be a wave throughout much of the nineteenth century, but was observed in the early twentieth century to have particle behavior as well. Electrons, meanwhile, were originally thought to be particles, but were then observed to have wave behavior.

B.C. Hall, *Quantum Theory for Mathematicians*, Graduate Texts in Mathematics 267, DOI 10.1007/978-1-4614-7116-5_3,

The second ingredient of quantum theory is its probabilistic behavior. In the two-slit experiment, for example, electrons that are "identically prepared" do not all hit the screen at the same point. Quantum theory postulates that this randomness is fundamental to the way nature behaves. According to quantum mechanics, it is impossible (theoretically, not just in practice) to predict ahead of time what the outcome of an experiment will be. The best that can be done is to predict the *probabilities* for the outcome of an experiment.

These two aspects of quantum theory come together in the wave function. The wave function is a function of a variable $\mathbf{x} \in \mathbb{R}^n$, which we interpret as describing the possible values of the position of a particle, and it evolves in time according to a wavelike equation (the Schrödinger equation). The wave function and its time-evolution account for the wave aspect of quantum theory. The particle aspect of the theory comes from the *interpretation* of the wave function. Although it is tempting to interpret the wave function as a sort of cloud, where we have, say, a little bit of electron-cloud over here, and little bit of electron-cloud over there, this interpretation is not consistent with experiment. Whenever we attempt to measure the position of a *single* electron, we always find the electron at a single point. A single electron in the two-slit experiment is observed at a single point on the screen, not spread out over the screen the way the wave function is. The wave function does not describe something that is directly observable for a single particle; rather, the wave function determines the statistical behavior of a whole sequence of identically prepared particles. See Fig. 1.4 for a dramatic experimental demonstration of this effect.

In the two-slit experiment, for example, it is possible to determine how the wave function behaves as a function of time by solving the (deterministic) Schrödinger equation. Knowledge of the wave function of an individual electron, however, does not determine where that electron will hit the screen. The wave function merely tells us the probability distribution for where the electron might hit the screen, something that is only observable by shooting a whole sequence of electrons at the screen.

It is an oversimplification, but a useful one, to describe the wave–particle aspect of quantum theory in this way: a single electron (or photon, or whatever) acts like a particle, but a large collection of electrons behaves like a wave. A single measurement of a single electron always gives its position as a point, just as we would expect for a particle. This point, however, varies from one electron to the next, even if we shoot each electron toward the screen in precisely the same way. Repeated measurements of identically prepared electrons give a distribution that can, for example, exhibit interference patterns, just as we would expect for a wave. See, again, Fig. 1.4, which should be compared to Figs. 1.1 and 1.2.

It is interesting to note that at the macroscopic scale, where quantum effects are not apparent, light appears to be a wave, whereas electrons appear to be particles. This is the case even though both light and electrons are

really wave–particle hybrids, described in probabilistic terms by a wave function. The difference between the two situations is that photons (the particles of light) have mass zero, whereas electrons have positive mass. This means that photons, unlike electrons, can easily be created and destroyed even at low energies. Thus, the discrete aspect of light—namely, that the energy in light comes only in discrete "quanta," namely the photons—is less evident than the corresponding discrete aspect of electrons.

3.2 A Few Words About Operators and Their Adjoints

In quantum mechanics, physical quantities—such as position, momentum, and energy—are represented by operators on a certain Hilbert space $\mathbf{H}$. These operators are unbounded operators, reflecting that in classical mechanics, these quantities are unbounded functions on the classical phase space. In this section, we look briefly at some technical issues related to unbounded operators and their adjoints. We will delay a full discussion of these technicalities (Chap. 9) until after we have understood the basic ideas of quantum mechanics.

Here and throughout the book, $\mathbf{H}$ will represent a Hilbert space over $\mathbb{C}$, always assumed to be separable. We follow the convention in the physics literature that the inner product be linear in the *second* factor:

$$\langle \phi, \lambda\psi \rangle = \lambda \langle \phi, \psi \rangle ; \quad \langle \lambda\phi, \psi \rangle = \bar{\lambda} \langle \phi, \psi \rangle$$

for all $\phi, \psi \in \mathbf{H}$ and all $\lambda \in \mathbb{C}$.

Recall (Appendix A.3.4) that a linear operator $A : \mathbf{H} \to \mathbf{H}$ is *bounded* if there is a constant C such that $\|A\psi\| \leq C \|\psi\|$ for all $\psi \in \mathbf{H}$. For any bounded operator A, there is a unique bounded operator A^*, called the *adjoint* of A, such that

$$\langle \phi, A\psi \rangle = \langle A^*\phi, \psi \rangle$$

for all $\phi, \psi \in \mathbf{H}$. The existence of A^* follows from the Riesz theorem (Appendix A.4.3), by observing that for each fixed ϕ, the map $\psi \mapsto \langle \phi, A\psi \rangle$ is a bounded linear functional on $\mathbf{H}$. A bounded operator is said to be *self-adjoint* if $A^* = A$.

For various reasons, both physical and mathematical, we want the operators of quantum mechanics operators to be self-adjoint. Once one sees the formulas for these operators, however, one is confronted with a serious technical difficulty: the operators are not bounded.

If A is a linear operator defined on all of $\mathbf{H}$ and having the property that $\langle \phi, A\psi \rangle = \langle A\phi, \psi \rangle$ for all $\phi, \psi \in \mathbf{H}$, then A is automatically bounded. (See Corollary 9.9.) To put this fact the other way around, *an unbounded*

self-adjoint operator cannot be defined on the entire Hilbert space. Thus, to deal with the unbounded operators of quantum mechanics, we must deal with operators that are defined only on a subspace of the relevant Hilbert space, called the **domain** of the operator.

Definition 3.1 *An* ***unbounded operator*** A *on* $\mathbf{H}$ *is a linear map from a dense subspace* $\mathrm{Dom}(A) \subset \mathbf{H}$ *into* $\mathbf{H}$.

More precisely, the operator A is "not necessarily bounded," since nothing in the definition prevents us from having $\mathrm{Dom}(A) = \mathbf{H}$ and having A be bounded.

In defining the adjoint of an unbounded operator, we immediately encounter a difficulty: for a given $\phi \in \mathbf{H}$, the linear functional $\langle \phi, A\cdot \rangle$ may not be bounded, in which case we cannot use the Riesz theorem to define $A^*\phi$. What this means is that the adjoint of A, like A itself, will be defined not on all of $\mathbf{H}$ but only on some subspace thereof.

Definition 3.2 *For an unbounded operator* A *on* $\mathbf{H}$, *the* ***adjoint*** A^* *of* A *is defined as follows. A vector* $\phi \in \mathbf{H}$ *belongs to the domain* $\mathrm{Dom}(A^*)$ *of* A^* *if the linear functional*

$$\langle \phi, A\cdot \rangle \,,$$

defined on $\mathrm{Dom}(A)$, *is bounded. For* $\phi \in \mathrm{Dom}(A^*)$, *let* $A^*\phi$ *be the unique vector* χ *such that*

$$\langle \chi, \psi \rangle = \langle \phi, A\psi \rangle$$

for all $\psi \in \mathrm{Dom}(A)$.

Saying that the linear functional $\langle \phi, A\cdot \rangle$ is bounded means that there is a constant C such that $|\langle \phi, A\psi \rangle| \leq C \, \|\psi\|$ for all $\psi \in \mathrm{Dom}(A)$. If $\langle \phi, A\cdot \rangle$ is bounded, then since $\mathrm{Dom}(A)$ is dense, the BLT theorem (Theorem A.36) tells us that $\langle \phi, A\cdot \rangle$ has a unique bounded extension to all of $\mathbf{H}$. The Riesz theorem then guarantees the existence and uniqueness of χ. The adjoint of an unbounded linear operator is a linear operator on its domain.

We are now ready to define self-adjointness (and some related notions) for unbounded operators.

Definition 3.3 *An unbounded operator* A *on* $\mathbf{H}$ *is* ***symmetric*** *if*

$$\langle \phi, A\psi \rangle = \langle A\phi, \psi \rangle$$

for all $\phi, \psi \in \mathrm{Dom}(A)$. *The operator* A *is* ***self-adjoint*** *if* $\mathrm{Dom}(A^*) = \mathrm{Dom}(A)$ *and* $A^*\phi = A\phi$ *for all* $\phi \in \mathrm{Dom}(A)$. *Finally,* A *is* ***essentially self-adjoint*** *if the closure in* $\mathbf{H} \times \mathbf{H}$ *of the graph of* A *is the graph of a self-adjoint operator.*

That is to say, A is self-adjoint if A^* and A are the *same operator* with the *same domain.* Every self-adjoint or essentially self-adjoint operator is

symmetric, but not every symmetric operator is essentially self-adjoint. For any symmetric operator, $\text{Dom}(A^*) \supset \text{Dom}(A)$ and A^* agrees with A on $\text{Dom}(A)$. The reason a symmetric operator may fail to be self-adjoint is that $\text{Dom}(A^*)$ may be strictly larger than $\text{Dom}(A)$.

Although the condition of being symmetric is certainly easier to understand (and to verify) than the condition of being self-adjoint, self-adjointness is the "right" condition. In particular, the spectral theorem, which is essential to much of quantum mechanics, applies only to operators that are self-adjoint and not to operators that are merely symmetric. If A is essentially self-adjoint, then we can obtain a self-adjoint operator from A simply by taking the closure of the graph of A, and we can then apply the spectral theorem to this self-adjoint operator. Thus, for may purposes, it is enough to have our operators be essentially self-adjoint rather than self-adjoint.

It is generally easy to verify that the operators of quantum mechanics (those representing position, momentum, and so forth) are symmetric on some suitably chosen domain. Proving that these operators are essentially self-adjoint, however, is substantially more difficult. Although establishing essential self-adjointness is a crucial technical issue, it is best not to worry too much about it on a first encounter with quantum mechanics. In this chapter, we will not concern ourselves overly with technical details concerning essential self-adjointness and the precise choice of domain for our operators, depending on Chap. 9 to take care of such matters. For now, we content ourselves with deriving some very elementary properties of symmetric (and thus also self-adjoint) operators.

Proposition 3.4 *Suppose A is a symmetric operator on* $\mathbf{H}$.

1. *For all $\psi \in \text{Dom}(A)$, the quantity $\langle \psi, A\psi \rangle$ is real. More generally, if $\psi, A\psi, \ldots, A^{m-1}\psi$ all belong to* $\text{Dom}(A)$, *then $\langle \psi, A^m\psi \rangle$ is real.*
2. *Suppose λ is an eigenvector for A, meaning that $A\psi = \lambda\psi$ for some nonzero $\psi \in \text{Dom}(A)$. Then $\lambda \in \mathbb{R}$.*

Proof. Since A is symmetric, we have

$$\langle \psi, A\psi \rangle = \langle A\psi, \psi \rangle = \overline{\langle \psi, A\psi \rangle}$$

for all $\psi \in \text{Dom}(A)$. If $\psi, A\psi, \ldots, A^{m-1}\psi$ all belong to the domain of A, we can use the symmetry of A repeatedly to show that

$$\langle \psi, A^m\psi \rangle = \langle A^m\psi, \psi \rangle = \overline{\langle \psi, A^m\psi \rangle}.$$

Meanwhile, if ψ is an eigenvector for A with eigenvalue λ, then

$$\lambda \langle \psi, \psi \rangle = \langle \psi, A\psi \rangle = \langle A\psi, \psi \rangle = \bar{\lambda} \langle \psi, \psi \rangle .$$

Since ψ is assumed to be nonzero, this implies that $\lambda = \bar{\lambda}$. ■

Physically, $\langle \psi, A\psi \rangle$ represents—as we will see later in this chapter—the expectation value for measurements of A in the state ψ, whereas the eigenvalue λ represents one of the possible values for this measurement. On physical grounds, we want both of these numbers to be real. If A is self-adjoint, and not just symmetric, then the spectral theorem will give a canonical way of associating to each $\psi \in \mathbf{H}$ a probability measure on the real line that encodes the probabilities for measurements of A in the state ψ.

3.3 Position and the Position Operator

Let us consider at first a single particle moving on the real line. The wave function for such a particle is a map $\psi : \mathbb{R}^1 \to \mathbb{C}$. Although this map will evolve in time, let us think for now that the time is fixed. The function $|\psi(x)|^2$ is supposed to be the probability density for the position of the particle. This means that the probability that the position of the particle belongs to some set $E \subset \mathbb{R}^1$ is

$$\int_E |\psi(x)|^2 \, dx.$$

For this prescription to make sense, ψ should be normalized so that

$$\int_{\mathbb{R}} |\psi(x)|^2 \, dx = 1. \tag{3.1}$$

That is, ψ should be a unit vector in the Hilbert space $L^2(\mathbb{R})$.

Now, if the function $|\psi(x)|^2$ is the probability density for the position of a particle, then according to the standard definitions of probability theory, the *expectation value* of the position will be

$$E(x) = \int_{\mathbb{R}} x \, |\psi(x)|^2 \, dx, \tag{3.2}$$

provided that the integral is absolutely convergent. More generally, we can compute any *moment* of the position (i.e., the expectation value of some power of the position) as

$$E(x^m) = \int_{\mathbb{R}} x^m \, |\psi(x)|^2 \, dx, \tag{3.3}$$

assuming, again, the convergence of the integral.

A key idea in quantum theory is to express expectation values of various quantities (position, momentum, energy, etc.) in terms of *operators* and the *inner product* on the relevant Hilbert space, in this case, $L^2(\mathbb{R})$. In the case of position, we may introduce the *position operator* X defined by

$$(X\psi)(x) = x\psi(x).$$

That is, X is the "multiplication by x" operator. The point of introducing this operator is that the expectation value of the position [defined in (3.2)] may now be expressed as

$$E(x) = \langle \psi, X\psi \rangle ,$$

where the inner product is the usual one on $L^2(\mathbb{R})$:

$$\langle \phi, \psi \rangle = \int \overline{\phi(x)} \psi(x) \, dx.$$

(Recall that we are following the physics convention of putting the conjugate on the *first* factor in the inner product.)

We use the following notation for the expectation value of the operator X in the state ψ:

$$\langle X \rangle_\psi := \langle \psi, X\psi \rangle .$$

The higher moments of the position, as defined in (3.3), are also computable in terms of the position operator:

$$E(x^m) = \langle \psi, X^m \psi \rangle .$$

At this point, it is not clear that we have gained anything by writing our moments in terms of an operator and the inner product instead of in terms of the integral (3.3). The operator description will, however, motivate a parallel description of moments for the momentum, energy, or angular momentum of a particle in terms of corresponding operators.

It should be noted that, for a given $\psi \in L^2(\mathbb{R})$, $X\psi$ might fail to be in $L^2(\mathbb{R})$. This failure of X to be defined on all of our Hilbert space reflects that X is an unbounded operator, something that we discussed briefly in Sect. 3.2. Even if $X\psi$ is in $L^2(\mathbb{R})$, $X^m\psi$ might fail to be in $L^2(\mathbb{R})$ for some m. Nevertheless, for any unit vector ψ in $L^2(\mathbb{R})$, we have a well-defined probability density on $\mathbb{R}$, given by $|\psi(x)|^2$.

3.4 Momentum and the Momentum Operator

At any fixed time, the wave function $\psi(x)$ of a particle (according to the wave theory postulated by Schrödinger) is a function of a "position" variable x only. Although the wave function ψ directly encodes the probabilities for the position of the particle, through $|\psi(x)|^2$, it is not as clear how information about the particle's momentum is encoded. As it turns out, the momentum is encoded in the *oscillations* of the wave function. A crucial idea in quantum mechanics is the *de Broglie hypothesis*, which we introduced in Sect. 1.2.2 as a way of understanding the allowed energies in the Bohr model of the hydrogen atom. The de Broglie hypothesis proposes a particular relationship between the frequency of oscillation of the wave function—as a function of position at a fixed time—and its momentum.

Proposition 3.5 (de Broglie hypothesis) *If the wave function of a particle has spatial frequency k, then the momentum p of the particle is*

$$p = \hbar k, \tag{3.4}$$

where $\hbar$ is Planck's constant.

The Davisson–Germer electron-diffraction experiments, described in Sect. 1.2.3, strongly support not only the idea that electrons have wavelike behavior, but also the specific relationship (3.4) between the momentum of an electron and the spatial frequency of the associated wave. Of course, Proposition 3.5 is rather vague. To be a bit more precise, Proposition 3.5 is supposed to mean that a wave function of the form $\psi(x) = e^{ikx}$ represents a particle with momentum $p = \hbar k$. [Here, as in Chap. 2, "frequency" is in the angular sense. The cycles-per-unit-distance frequency is $\nu = k/(2\pi)$.]

Now, the function e^{ikx} is obviously not square integrable, so it is not strictly possible for the wave function [which is supposed to satisfy (3.1)] to be e^{ikx}. Let us therefore briefly switch to thinking of a particle on a circle, so that we can avoid certain technicalities. We think of the wave function ψ for a particle on a circle as a 2π-periodic function on $\mathbb{R}$, satisfying the normalization condition

$$\int_0^{2\pi} |\psi(x)|^2 \, dx = 1.$$

For any integer k, it makes sense to say that the normalized wave function $\psi(x) = e^{ikx}/\sqrt{2\pi}$ represents a particle with momentum $p = \hbar k$. In this case, we are supposed to think that the momentum of the particle is definite, that is, nonrandom. If the particle's wave function is $e^{ikx}/\sqrt{2\pi}$, then a measurement of the particle's momentum should (with probability 1) give the value $\hbar k$.

Now, the functions $e^{ikx}/\sqrt{2\pi}$, $k \in \mathbb{Z}$, form an orthonormal basis for the Hilbert space of 2π-periodic, square-integrable functions, which may be identified with $L^2([0, 2\pi])$. Thus, the typical wave function for a particle on a circle is

$$\psi(x) = \sum_{k=-\infty}^{\infty} a_k \frac{e^{ikx}}{\sqrt{2\pi}}, \tag{3.5}$$

where the sum is convergent in $L^2([0, 2\pi])$. If ψ is normalized to be a unit vector, then we have

$$\sum_{k=-\infty}^{\infty} |a_k|^2 = \|\psi\|^2_{L^2([0,2\pi])} = 1. \tag{3.6}$$

For a particle with wave function given by (3.5), the momentum of the particle is no longer definite. Rather, we are supposed to think that a

measurement of the particle's momentum will yield one of the values $\hbar k$, $k \in \mathbb{Z}$, with the probability of getting a particular value $\hbar k$ being $|a_k|^2$. Following elementary probability theory, then, the expectation values for the momentum should be

$$E(p) = \sum_{k=-\infty}^{\infty} \hbar k \, |a_k|^2 , \tag{3.7}$$

and higher moments for the momentum should be

$$E(p^m) = \sum_{k=-\infty}^{\infty} (\hbar k)^m \, |a_k|^2 , \tag{3.8}$$

assuming absolute convergence of the sum.

We would like to encode the moment conditions (3.7) and (3.8) in a *momentum operator* P, which should be defined in such a way that if the particle's wave function ψ is given by (3.5), then $E(p^m) = \langle \psi, P^m \psi \rangle$. We can achieve this relation if P satisfies

$$P e^{ikx} = \hbar k e^{ikx}, \tag{3.9}$$

since then,

$$\langle \psi, P^m \psi \rangle = \sum_{k=-\infty}^{\infty} (\hbar k)^m \, |a_k|^2 = E(p^m). \tag{3.10}$$

The (presumably unique) choice for P satisfying (3.9) is

$$P = -i\hbar \frac{d}{dx}.$$

Returning now to the setting of the real line, it is natural to postulate that the momentum operator P on the line should also be given by $P = -i\hbar \, d/dx$. This operator satisfies the relation

$$P e^{ikx} = (\hbar k) e^{ikx},$$

which is supposed to capture the idea that the wave function e^{ikx} has momentum $\hbar k$. Although the function e^{ikx} is not square-integrable with respect to x, the Fourier transform allows us to build up any square-integrable function as a "superposition" of functions of the form e^{ikx}. (*Superposition* is the term physicists use for a linear combination or the continuous analog thereof, namely an integral.) This means that [by analogy to (3.5)] we have

$$\psi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{ikx} \hat{\psi}(k) \, dk, \tag{3.11}$$

where $\hat{\psi}(k)$ is the Fourier transform of ψ, defined by

$$\hat{\psi}(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-ikx} \psi(x) \, dx. \tag{3.12}$$

(See Appendix A.3.2 for information about the Fourier transform.)

The Plancherel theorem (Theorem A.19) then tells us that the Fourier transform is a unitary map of $L^2(\mathbb{R})$ onto $L^2(\mathbb{R})$. Thus, for any unit vector $\psi \in L^2(\mathbb{R})$,

$$\int_{-\infty}^{\infty} |\psi(x)|^2 \, dx = \int_{-\infty}^{\infty} \left|\hat{\psi}(k)\right|^2 \, dk = 1.$$

In light of what we have in the circle case, it is natural to think that $|\hat{\psi}(k)|^2$ is essentially the probability density for the momentum of the particle. (To be precise, $|\hat{\psi}(k)|^2$ is the probability density for $p/\hbar$.)

We can now express the properties of the momentum operator entirely within the Hilbert space $L^2(\mathbb{R})$, without making explicit mention of the non–square-integrable functions e^{ikx}.

Proposition 3.6 *Define the momentum operator P by*

$$P = -i\hbar \frac{d}{dx}.$$

Then for all sufficiently nice unit vectors ψ in $L^2(\mathbb{R})$, we have

$$\langle \psi, P^m \psi \rangle = \int_{-\infty}^{\infty} (\hbar k)^m \left|\hat{\psi}(k)\right|^2 \, dk \tag{3.13}$$

for all positive integers m. The quantity in (3.13) is interpreted as the expectation value of the mth power of the momentum, $E(p^m)$.

Equation (3.13) should be compared to (3.10) in the case of the circle.

Proof. If ψ is in, say, the Schwartz space (Definition A.15), then, by applying Proposition A.17 m times, we see that the Fourier transform of the nth derivative of ψ is $(ik)^m \hat{\psi}(k)$, and so the Fourier transform of $P^m \psi$ is $(\hbar k)^m \hat{\psi}(k)$. Meanwhile, since the Fourier transform is unitary, we have

$$\langle \psi, P^m \psi \rangle = \int_{-\infty}^{\infty} \overline{\hat{\psi}(k)} (\hbar k)^m \hat{\psi}(k) \, dk,$$

which gives (3.13). (The assumption that ψ be in the Schwartz space is stronger than necessary. The reader is invited to use integration by parts and the definition of the Fourier transform to find weaker assumptions that allow the same conclusion.) ■

3.5 The Position and Momentum Operators

In the following definition, we summarize what we have learned, in the two previous sections, about the position and momentum operators.

Definition 3.7 *For a particle moving in* $\mathbb{R}^1$*, let the quantum Hilbert space be* $L^2(\mathbb{R})$ *and define the* ***position*** *and* ***momentum operators*** X *and* P *by*

$$X\psi(x) = x\psi(x)$$
$$P\psi(x) = -i\hbar\frac{d\psi}{dx}.$$

Neither the position nor the momentum operator is defined as mapping the entire Hilbert space $L^2(\mathbb{R})$ into itself. After all, for $\psi \in L^2(\mathbb{R})$, the function $x\psi(x)$ may fail to be in $L^2(\mathbb{R})$. Similarly, a function ψ in $L^2(\mathbb{R})$ may fail to be differentiable, and even if it is differentiable, the derivative may fail to be in $L^2(\mathbb{R})$. What this means is that X and P are unbounded operators, of the sort discussed briefly in Sect. 3.2. They are defined on suitable dense subspaces $\mathrm{Dom}(X)$ and $\mathrm{Dom}(P)$ of $L^2(\mathbb{R})$. We defer a detailed examination of the domains of these operators until Chap. 9.

A vitally important property of this pair of operators is that they *do not commute*.

Proposition 3.8 *The position and momentum operators* X *and* P *do not commute, but satisfy the relation*

$$XP - PX = i\hbar I, \tag{3.14}$$

This relation is known as the *canonical commutation relation.*

Proof. Using the product rule we calculate that

$$\begin{aligned} PX\psi &= -i\hbar\frac{d}{dx}\left(x\psi(x)\right) \\ &= -i\hbar\psi(x) - i\hbar x\frac{d\psi}{dx} \\ &= -i\hbar\psi(x) + XP\psi, \end{aligned}$$

from which (3.14) follows. ■

There are many important consequences of the relation (3.14), which we will examine at length in Chaps. 11– 14 of the book. For now, we simply note a parallel between (3.14) and the Poisson bracket relationship in classical mechanics: $\{x, p\} = 1$, as follows directly from the definition of the Poisson bracket. This hints at an analogy, which we will explore further in Sect. 3.7, between the *commutator* of two operators A and B on the quantum side (namely, the operator $AB - BA$) and the Poisson bracket of two functions f and g on the classical side.

Proposition 3.9 *For all sufficiently nice functions* ϕ *and* ψ *in* $L^2(\mathbb{R})$*, we have*

$$\langle \phi, X\psi \rangle = \langle X\phi, \psi \rangle$$

and

$$\langle \phi, P\psi \rangle = \langle P\phi, \psi \rangle .$$

Proof. Suppose that ϕ and ψ belong to $L^2(\mathbb{R})$ and that the functions $x\phi(x)$ and $x\psi(x)$ also belong to $L^2(\mathbb{R})$. Then since x is real, we have

$$\int_{-\infty}^{\infty} \overline{\phi(x)} x\psi(x)\, dx = \int_{-\infty}^{\infty} \overline{x\phi(x)}\psi(x)\, dx,$$

where both integrals are convergent because they are both integrals of the product of two L^2 functions.

Meanwhile, for the second claim, let us assume that ϕ and ψ are continuously differentiable and that $\phi(x)$ and $\psi(x)$ tend to zero as x tends to $\pm\infty$. Let us also assume that ϕ, ψ, $d\phi/dx$ and $d\psi/dx$ belong to $L^2(\mathbb{R})$. We note that $d\bar{\phi}/dx$ is the same as $\overline{d\phi/dx}$. Thus, using integration by parts, we obtain

$$-i\hbar \int_{-A}^{A} \overline{\phi(x)} \frac{d\psi}{dx}\, dx = -i\hbar\, \overline{\phi(x)}\psi(x)\Big|_{-A}^{A} + i\hbar \int_{-A}^{A} \overline{\frac{d\phi}{dx}}\psi(x)\, dx.$$

Under our assumptions on ϕ and ψ, as A tends to infinity, the boundary terms will vanish and the remaining integrals will tend (by dominated convergence) to integrals over the whole real line. Thus,

$$\begin{aligned}\int_{-\infty}^{\infty} \overline{\phi(x)} \left(-i\hbar \frac{d\psi}{dx}\right) dx &= i\hbar \int_{-\infty}^{\infty} \overline{\frac{d\phi}{dx}}\psi(x)\, dx \\ &= \int_{-\infty}^{\infty} \overline{\left(-i\hbar \frac{d\phi}{dx}\right)}\psi(x)\, dx,\end{aligned}$$

which is the second claim in the proposition. ■

In the language of Definition 3.3, Proposition 3.9 means that X and P are *symmetric* operators on certain dense subspaces of $L^2(\mathbb{R})$ (the space of functions for which the proposition is proved). It is actually true that X and P are essentially self-adjoint on these domains. The proof of essential self-adjointness, however, will have to wait until Chap. 9.

3.6 Axioms of Quantum Mechanics: Operators and Measurements

In this section we consider the general "axioms" of quantum mechanics. These axioms are not to be understood in the mathematical sense as rules from which all other results are derived in a strictly deductive fashion. Rather, the axioms are the main principles of how quantum mechanics works. Here we look at the "kinematic" axioms, those that apply at one fixed time. There is one additional axiom, governing the time-evolution of the system, which we consider in the next section.

Axiom 1 *The state of the system is represented by a unit vector ψ in an appropriate Hilbert space* $\mathbf{H}$. *If ψ_1 and ψ_2 are two unit vectors in* $\mathbf{H}$ *with $\psi_2 = c\psi_1$ for some constant $c \in \mathbb{C}$, then ψ_1 and ψ_2 represent the same physical state.*

The Hilbert space $\mathbf{H}$ is frequently called the "quantum Hilbert space." This does not, however, mean that $\mathbf{H}$ is some variant of the notion of a Hilbert space, the way a quantum group is a variant of the notion of a group. Rather, "quantum Hilbert space" means simply, "the Hilbert space associated with a given quantum system."

In Axiom 1, it should be noted that unit vectors in $\mathbf{H}$ actually represent only the "pure states" of the theory. There is a more general notion of a "mixed state" (described by a "density matrix") that we will consider in Chap. 19. We will follow the custom in most physics texts of considering at first only pure states.

Axiom 2 *To each real-valued function f on the classical phase space there is associated a self-adjoint operator $\hat{f}$ on the quantum Hilbert space.*

In almost all cases, the operator $\hat{f}$ is unbounded. This unboundedness is unsurprising when we realize that physically relevant functions f on the classical phase space (e.g., position and momentum) are unbounded functions. In the unbounded case, the notion of self-adjointness is rather technical; see Definition 3.3 in Sect. 3.2. In most applications, it is not really necessary to define $\hat{f}$ for *all* functions on the classical phase space, but only for certain basic functions, such as position, momentum, energy, and angular momentum. We will describe the quantizations of these basic functions in this chapter. If one really needs to define $\hat{f}$ for an arbitrary function f (satisfying some regularity assumptions), the standard approach is to use the Weyl quantization scheme, described in Chap. 13.

For a particle moving in $\mathbb{R}^1$, the classical phase space is $\mathbb{R}^2$, which we think of as pairs (x, p) with x being the particle's position and p being its momentum. The quantum Hilbert space in this case is usually taken to be $L^2(\mathbb{R})$ [*not* $L^2(\mathbb{R}^2)$]. In that case, if the function f in Axiom 2 is the position function, $f(x, p) = x$, then the associated operator $\hat{f}$ is the position operator X, given by multiplication by x. If f is the momentum function, $f(x, p) = p$, then $\hat{f}$ is the momentum operator $P = -i\hbar\, d/dx$.

In the physics literature, a function f on the classical phase space is called a *classical observable*, meaning that it is some physical quantity that could be observed by taking a measurement of the system. The corresponding operator $\hat{f}$ is then called a *quantum observable*.

Axiom 3 *If a quantum system is in a state described by a unit vector $\psi \in \mathbf{H}$, the probability distribution for the measurement of some observable f satisfies*

$$E(f^m) = \left\langle \psi, (\hat{f})^m \psi \right\rangle. \tag{3.15}$$

In particular, the expectation value for a measurement of f is given by

$$\left\langle \psi, \hat{f}\psi \right\rangle . \tag{3.16}$$

Note that we have adopted the point of view that even in a quantum mechanical system, what one is measuring is the *classical* observable f. In the quantum case, however, f no longer has a definite value, but only probabilities, which are encoded by the quantum observable $\hat{f}$ and the vector $\psi \in \mathbf{H}$.

If ψ is a nonzero vector in $\mathbf{H}$ but not a unit vector, then (3.16) should be replaced by

$$\frac{\left\langle \psi, \hat{f}\psi \right\rangle}{\langle \psi, \psi \rangle} = \left\langle \tilde{\psi}, \hat{f}\tilde{\psi} \right\rangle ,$$

where $\tilde{\psi} := \psi / \|\psi\|$ is the unit vector associated with ψ. It is convenient to assume that our vectors have been normalized to be unit vectors, simply to avoid having to divide by $\langle \psi, \psi \rangle$ in our expectation values.

Since $\hat{f}$ is assumed to be self-adjoint and every self-adjoint operator is symmetric, Proposition 3.4 tells us that the moments $E(f^m)$, and in particular the expectation value $E(f)$, are real numbers. Since $\hat{f}$ is assumed to be self-adjoint and not just symmetric, the spectral theorem (Chaps. 7 and 10) will give a canonical way of constructing a probability measure $\mu_{A,\psi}$ on $\mathbb{R}$ that may be interpreted as the probability distribution for measurements of A in the state ψ.

Axiom 3 provides motivation for the idea that two unit vectors that differ by a constant represent the same physical state. If $\psi_2 = c\psi_1$ with $|c| = 1$, then for any operator A, we have

$$\langle \psi_2, A\psi_2 \rangle = \langle c\psi_1, Ac\psi_1 \rangle = |c|^2 \langle \psi_1, A\psi_1 \rangle = \langle \psi_1, A\psi_1 \rangle .$$

Thus, the expectation values of all observables are the same in the state ψ_2 as in the state ψ_1.

Notation 3.10 *If A is a self-adjoint operator on $\mathbf{H}$ and $\psi \in \mathbf{H}$ is a unit vector, the expectation value of A in the state ψ is denoted $\langle A \rangle_\psi$ and is defined (in light of Axiom 3) to be*

$$\langle A \rangle_\psi = \langle \psi, A\psi \rangle . \tag{3.17}$$

Proposition 3.11 (Eigenvectors) *If a quantum system is in a state described by a unit vector $\psi \in \mathbf{H}$ and for some quantum observable $\hat{f}$ we have $\hat{f}\psi = \lambda\psi$ for some $\lambda \in \mathbb{R}$, then*

$$E(f^m) = \left\langle (\hat{f})^m \right\rangle_\psi = \lambda^m \tag{3.18}$$

for all positive integers m. The unique probability measure consistent with this condition is the one in which f has the definite value λ, *with probability one.*

What the proposition means is that if ψ is an eigenvector for $\hat{f}$, then measurements of f for a particle in the state ψ are not actually random, but rather always give the answer of λ. If $\hat{f}\psi = \lambda\psi$, then $\left\langle \psi, (\hat{f})^m \psi \right\rangle = \lambda^m \langle \psi, \psi \rangle = \lambda^m$. Thus, by (3.15), we want to find a probability measure μ on $\mathbb{R}$ such that

$$\int_{\mathbb{R}} x^m \, d\mu = \lambda^m, \tag{3.19}$$

for all non-negative integers m. The proposition is claiming that there is one and only one such measure, namely the δ-measure at the point λ.

Because $\hat{f}$ is assumed to be self-adjoint and therefore symmetric, Proposition 3.4 thus tells us that the every eigenvalue for $\hat{f}$ is real.

Proof. The relation (3.18) follows from (3.15) and the fact that $\hat{f}\psi = \lambda\psi$. Meanwhile, if μ is the δ-measure at λ, then certainly (3.19) holds. Meanwhile, since the mth moment grows only exponentially with m, even the most elementary uniqueness results for the moment problem show that the δ-measure is the *only* measure with these moments. (See, e.g., Theorem 8.1 in Chap. 4 of [18].) ■

If, more generally, the state of the system is a *linear combination* of eigenvectors for $\hat{f}$, measurements of f will no longer be deterministic.

Example 3.12 *Suppose $\hat{f}$ has an orthonormal basis $\{e_j\}$ of eigenvectors with distinct (real) eigenvalues λ_j. Suppose also that ψ is a unit vector in* **H** *with the expansion*

$$\psi = \sum_{j=1}^{\infty} a_j e_j. \tag{3.20}$$

Then for a measurement in the state ψ of the observable f, the observed value of f will always be one of the numbers λ_j. Furthermore, the probability of observing the value λ_j is given by

$$\text{Prob}\{f = \lambda_j\} = |a_j|^2 . \tag{3.21}$$

Assuming that ψ is in the domain of $(\hat{f})^m$, it is easy to verify that the probabilities in (3.21) are consistent with the expectation values given in Axiom 3. After all, if ψ is given as in (3.20), then we can readily calculate that $\langle \psi, (\hat{f})^m \psi \rangle$ equals $\sum |a_j|^2 \lambda_j^m$, which is nothing but the mth moment associated with the probability distribution in (3.21). In general, we cannot quite *derive* (3.21) from Axiom 3, since the uniqueness results for the moment problem might not apply. Nevertheless, (3.21) is the most natural candidate for the probabilities, and we will assume that this formula holds.

It is not difficult to extend Example 3.12 to the case where the eigenvalues are not distinct: For *any* sequence $\{\lambda_j\}$ of eigenvalues, the probability of observing some value λ will be the sum of $|a_j|^2$ over all those values of j for which $\lambda_j = \lambda$. For any self-adjoint operator A, the spectral theorem implies that A has *either* an orthonormal basis of eigenvectors *or* some

continuous analog thereof. In particular, given a self-adjoint operator A and a unit vector $\psi \in \mathbf{H}$, the spectral theorem will give us a probability measure μ_ψ^A on $\mathbb{R}$ that we interpret as describing the probabilities for a measurement of A in the state ψ. See Proposition 7.17 in the bounded case and Definition 10.7 in the unbounded case.

Axiom 4 *Suppose a quantum system is initially in a state ψ and that a measurement of an observable f is performed. If the result of the measurement is the number $\lambda \in \mathbb{R}$, then immediately after the measurement, the system will be in a state ψ' that satisfies*

$$\hat{f}\psi' = \lambda\psi'.$$

The passage from ψ to ψ' is called the collapse of the wave function. Here $\hat{f}$ is the self-adjoint operator associated with f by Axiom 2.

Let us assume again that $\hat{f}$ has an orthonormal basis of eigenvectors $\{e_j\}$ with distinct eigenvalues λ_j. Then we can say, more specifically, that if we observe the value λ_j in a measurement of $\hat{f}$ (and we *will* always observe one of the λ_j's) then $\psi' = e_j$. That is, the measurement "collapses" the wave function by throwing away all the components of ψ in the direction of the e_k's, except the one with $k = j$.

This idea of the collapse of the wave function has generated an enormous amount of discussion and controversy. One way to look at the situation is to think that the wave function ψ is not actually the state of the system—although we continue to use the standard physics term, "state." Rather, the wave function is the thing that encodes the *probabilities* for the state of the system. The collapse of the wave function is then something similar to a conditional probability; the probabilities for future measurements of the system should be consistent with the outcome of the measurement we just made. Paul Dirac has described the collapse of the wave function as being not a discontinuous change in the state of the system, but a discontinuous change in our *knowledge of* the state of the system.

In any case, Axiom 4 guarantees the following reasonable principle: If we measure f and then measure f again a very short time later, the result of the second measurement will agree with the result of the first measurement. Thus, immediately after the first measurement, the probabilities for a second measurement of f are not those associated with the vector ψ, but rather those associated with the state ψ'. (Since ψ' is an eigenvector for $\hat{f}$ with eigenvalue λ, Proposition 3.11 tells us that measurements of f in the state ψ' always give the value of λ.)

Note that Axiom 4 only tells us something about the state of the system *immediately* after a measurement. Following the measurement, the state of the system will evolve in time in the usual way (Sect. 3.7). A significant time after the measurement, then, the system will probably no longer be in the state ψ'.

Let us conclude this section by considering an example of how one makes a measurement of a real-world physical system, namely, the hydrogen atom. The Hamiltonian operator $\hat{H}$ for a hydrogen atom has negative eigenvalues of the form

$$-\frac{R}{n^2}, \tag{3.22}$$

where R is the *Rydberg constant* and $n = 1, 2, 3, \ldots$ These energies will be derived in Chap. 18. Negative eigenvalues are of greater interest than positive ones, because negative eigenvalues describes states where the electron is bound to the nucleus. If an electron is placed into a state having energy $-R/n_1^2$, with $n_1 > 1$, it will eventually "decay" into a state with lower energy, say, $-R/n_2^2$, with $n_2 < n_1$. (The most readily observed cases are those with $n_2 = 2$ and $n_2 = 1$.) In the process of decaying, the electron emits a photon, with the energy of the photon being equal to the change in energy of the electron, namely,

$$E_{\text{photon}} = \frac{R}{n_2^2} - \frac{R}{n_1^2}. \tag{3.23}$$

Meanwhile, the frequency of the photon is proportional to its energy. Thus, by observing the frequency of the emitted photon, one can determine the change in energy of the electron and thus determine the values of n_1 and n_2.

A general "bound state" of the hydrogen atom (a state in which the electron is bound to the nucleus), will be a linear combination of eigenvectors for $\hat{H}$ with various different eigenvalues of the form (3.22). To measure the energy of the electron, we simply wait for the electron to decay into a lower-energy state and emit a photon, observe the frequency of the photon, and work backwards to the energy of the electron. If we consider many "identically prepared" electrons, all having the same wave function that is a linear combination of eigenvectors, we will observe many different frequencies for the emitted photons, and thus many different energies for the electron. The *probabilities* for the observed energies of the electron will follow the principle spelled out in Example 3.12.

In basic probability theory, if Y is a random variable then the *variance* σ^2 of Y is computed as

$$\sigma^2 = E\left[(Y - E(Y))^2\right],$$

where E denotes the mean or expectation value of a random variable. The *standard deviation* $\sigma := \sqrt{\sigma^2}$ is a measure of the "typical" deviation from the mean $E(X)$. Observe that the variance may be computed as

$$\begin{aligned} \sigma^2 &= E\left[Y^2 - 2E(Y)Y + E(Y)^2\right] \\ &= E(Y^2) - 2E(Y)^2 + E(Y)^2 \\ &= E(Y^2) - E(Y)^2. \end{aligned} \tag{3.24}$$

Definition 3.13 *If A is a self-adjoint operator on a Hilbert space* $\mathbf{H}$ *and ψ is a unit vector in* $\mathbf{H}$*, let $\Delta_\psi A$ denote the standard deviation associated with measurements of A in the state ψ, which is computed as*

$$\begin{aligned}(\Delta_\psi A)^2 &= \left\langle (A - \langle A\rangle_\psi I)^2 \right\rangle_\psi \\ &= \langle A^2\rangle_\psi - \left(\langle A\rangle_\psi\right)^2.\end{aligned}$$

We refer to $\Delta_\psi A$ as the ***uncertainty*** *of A in the state ψ.*

For any single observable A, it is possible to choose ψ so that $\Delta_\psi A$ is as small as we like. In Chap. 12, however, we will see that when two observables A and B do not commute, then $\Delta_\psi A$ and $\Delta_\psi B$ cannot both be made arbitrarily small for the same ψ. In particular, we will derive there the famous *Heisenberg uncertainty principle*, which states that

$$(\Delta_\psi X)(\Delta_\psi P) \geq \frac{\hbar}{2},$$

for all ψ for which $\Delta_\psi X$ and $\Delta_\psi P$ are defined.

3.7 Time-Evolution in Quantum Theory

3.7.1 The Schrödinger Equation

Up to now, we have been considering the wave function ψ at a fixed time. We now consider the way in which the wave function evolves in time. Recall that in the Hamiltonian formulation of classical mechanics (Sect. 2.5), the time-evolution of the system is governed by the Hamiltonian (energy) function H, through Hamilton's equations. According to Axiom 2, there is a corresponding self-adjoint linear operator $\hat{H}$ on the quantum Hilbert space $\mathbf{H}$, which we call the *Hamiltonian operator* for the system. See Sect. 3.7.4 for an example.

Recall that we motivated the definition of the momentum operator by the *de Broglie hypothesis*, $p = \hbar k$, where k is the spatial frequency of the wave function. We can similarly motivate the time-evolution in quantum mechanics by a similar relation between the energy and the temporal frequency of our wave function:

$$E = \hbar\omega. \tag{3.25}$$

This relationship between energy and temporal frequency is nothing but the relationship proposed by Planck in his model of blackbody radiation (Sect. 1.1.3). Suppose that a wave function ψ_0 has definite energy E, meaning that ψ_0 is an eigenvector for $\hat{H}$ with eigenvalue E. Then (3.25) means that

the time-dependence of the wave function should be purely at frequency $\omega = E/\hbar$. That is to say, if the state of the system at time $t = 0$ is ψ_0, then the state of the system at any other time t should be

$$\psi(t) = e^{-i\omega t}\psi_0 = e^{-iEt/\hbar}\psi_0. \tag{3.26}$$

We can rewrite (3.26) as a differential equation:

$$\frac{d\psi}{dt} = -\frac{iE}{\hbar}\psi = \frac{E}{i\hbar}\psi. \tag{3.27}$$

Note that we are taking "temporal frequency ω" to mean that the time-dependence is of the form $e^{-i\omega t}$, whereas we took "spatial frequency k" to mean that the space-dependence is of the form e^{ikx}, with no minus sign in the exponent. This curious convention is convenient when we look at pure exponential solutions to the free Schrödinger equation (Chap. 4) of the form $\exp[i(kx - \omega t)]$, which describes a solution moving to the right with speed ω/k.

Equation (3.27) tells us the time-evolution for a particle that is initially in a state of definite energy, that is, an eigenvector for the Hamiltonian operator. A natural way to generalize this equation is to recognize that $E\psi$ is nothing but $\hat{H}\psi$, since ψ is just a multiple of ψ_0, which is an eigenvector for $\hat{H}$ with eigenvalue E. Replacing E by $\hat{H}$ in (3.27) leads to the following general prescription for the time-evolution of a quantum system.

Axiom 5 *The time-evolution of the wave function ψ in a quantum system is given by the Schrödinger equation,*

$$\frac{d\psi}{dt} = \frac{1}{i\hbar}\hat{H}\psi. \tag{3.28}$$

Here $\hat{H}$ is the operator corresponding to the classical Hamiltonian H by means of Axiom 2.

Although both Hamilton's equations and the Schrödinger equation involve a Hamiltonian, the two equations otherwise do not seem parallel. Of course, since quantum mechanics is not classical mechanics, we should not expect the two theories to have the same time-evolution. Nevertheless, we might hope to see *some* similarities between the time-evolution of a classical system and that of the corresponding quantum system. Such a similarity can be seen when we consider how the expectation values of observables evolve in quantum mechanics.

Proposition 3.14 *Suppose $\psi(t)$ is a solution of the Schrödinger equation and A is a self-adjoint operator on* **H**. *Assuming certain natural domain conditions hold, we have*

$$\frac{d}{dt}\langle A\rangle_{\psi(t)} = \left\langle \frac{1}{i\hbar}[A, \hat{H}] \right\rangle_{\psi(t)}, \tag{3.29}$$

where $\langle A \rangle_\psi$ *is as in Notation 3.10 and where* $[\cdot, \cdot]$ *denotes the* commutator, *defined as*

$$[A, B] = AB - BA.$$

Equation (3.29) should be compared to the way a function f on the classical phase space evolves in time along a solution of Hamilton's equations: $df/dt = \{f, H\}$. We see, then, that the commutator of operators (divided by $i\hbar$) plays a role in quantum mechanics similar to the role of the Poisson bracket in classical mechanics.

Proof. Let $\psi(t)$ be a solution to the Schrödinger equation and let us compute at first without worrying about domains of the operators involved. If we use the product rule (Exercise 1) for differentiation of the inner product, we obtain

$$\begin{aligned} \frac{d}{dt} \langle \psi(t), A\psi(t) \rangle &= \left\langle \frac{d\psi}{dt}, A\psi \right\rangle + \left\langle \psi, A \frac{d\psi}{dt} \right\rangle \\ &= \frac{i}{\hbar} \left\langle \hat{H}\psi, A\psi \right\rangle - \frac{i}{\hbar} \left\langle \psi, A\hat{H}\psi \right\rangle \\ &= \frac{1}{i\hbar} \left\langle \psi, [A, \hat{H}]\psi \right\rangle, \end{aligned}$$

where in the last step we have used the self-adjointness of $\hat{H}$ to move it to the other side of the inner product. Recall that we are following the convention of putting the complex conjugate on the first factor in the inner product, which accounts for the plus sign in the first term on the second line. Rewriting this using Notation 3.10 gives the desired result.

If A and $\hat{H}$ are (as usual) unbounded operators, then the preceding calculation is not completely rigorous. Since, however, we are deferring a detailed examination of issues of unbounded operators until Chap. 9, let us simply state the conditions needed for the calculation to be valid. For every $t \in \mathbb{R}$, we need to have $\psi(t) \in \mathrm{Dom}(A) \cap \mathrm{Dom}(\hat{H})$, we need $A\psi(t) \in \mathrm{Dom}(\hat{H})$, and we need $\hat{H}\psi(t) \in \mathrm{Dom}(A)$. (These conditions are needed for $[A, \hat{H}]\psi(t)$ to be defined.) In addition, we need $A\psi(t)$ to be a continuous path in $\mathbf{H}$. ■

Note that to see interesting behavior in the time-evolution of a quantum system, there has to be *noncommutativity* present. If all the physically interesting operators A commuted with the Hamiltonian operator $\hat{H}$, then $[\hat{H}, A]$ would be zero and the expectation values of these operators would be constant in time. Noncommutativity of the basic operators is therefore an essential property of quantum mechanics. In the case of a particle in $\mathbb{R}^1$, noncommutativity is built into the commutation relation for X and P, given in Proposition 3.8.

Although it is not reasonable to have *all* physically interesting operators commute with $\hat{H}$, there may be *some* operators with this property. If $[A, \hat{H}] = 0$, then the expectation value of A (and, indeed, all the moments of A) is independent of time along any solution of the Schrödinger equation.

We may therefore call such an operator A a *conserved quantity* (or *constant of motion*). Just as in the classical setting, conserved quantities (when we can find them) are helpful in understanding how to solve the Schrödinger equation.

Proposition 3.14 suggests that the map

$$(A, B) \longmapsto \frac{1}{i\hbar}[A, B],$$

where A and B are self-adjoint operators, plays a role similar to that of the Poisson bracket in classical mechanics. This analogy is supported by the following list of elementary properties of the commutator, which should be compared to the properties of the Poisson bracket listed in Proposition 2.23.

Proposition 3.15 *For any vector space V over $\mathbb{C}$ and linear operators A, B, and C on V, the following relations hold.*

1. $[A, B + \alpha C] = [A, B] + \alpha[A, C]$ *for all* $\alpha \in \mathbb{C}$
2. $[B, A] = -[A, B]$
3. $[A, BC] = [A, B]C + B[A, C]$
4. $[A, [B, C]] = [[A, B], C] + [B, [A, C]]$

Property 4 is equivalent to the *Jacobi identity*,

$$[A, [B, C]] + [B, [C, A]] + [C, [A, B]] = 0, \tag{3.30}$$

as can easily be seen using the skew-symmetry of the commutator.

Proof. The first two properties of the commutator are obvious, and the third is easily verified by writing things out. Property 4 can also be proved by writing things out, but it is slightly messier. Each of the three double commutators on the left-hand side of (3.30) generates four terms, for a total of 12 terms. Each term has the operators A, B, and C multiplied together in some order. It is a straightforward but unenlightening calculation to verify that each of the six possible orderings of A, B, and C occurs twice, with opposite signs. ■

If A and B are bounded *self-adjoint* operators on some Hilbert space, then it is straightforward to check that $(1/(i\hbar))[A, B]$ is again self-adjoint (Exercise 3). If A and B are unbounded self-adjoint operators, then the operator $(1/(i\hbar))[A, B]$ will be self-adjoint under suitable assumptions on the domains of A and B.

Proposition 3.16 *If $\phi(t)$ and $\psi(t)$ are solutions to the Schrödinger equation (3.28), the quantity $\langle \phi(t), \psi(t) \rangle$ is independent of t. In particular, $\|\psi(t)\|$ is independent of t, for any solution $\psi(t)$ of the Schrödinger equation.*

Proof. Using again the product rule, we have

$$\begin{aligned}\frac{d}{dt}\langle\phi(t),\psi(t)\rangle &= \left\langle \frac{1}{i\hbar}\hat{H}\phi(t),\psi(t)\right\rangle + \left\langle \phi(t),\frac{1}{i\hbar}\hat{H}\psi(t)\right\rangle \\ &= -\frac{1}{i\hbar}\left\langle \hat{H}\phi(t),\psi(t)\right\rangle + \frac{1}{i\hbar}\left\langle \phi(t),\hat{H}\psi(t)\right\rangle\end{aligned}$$

Since $\hat{H}$ is self-adjoint, we can move $\hat{H}$ to the other side of the inner product and the derivative is equal to 0. ■

3.7.2 Solving the Schrödinger Equation by Exponentiation

The Schrödinger equation is an example of a equation of the form

$$\frac{dv}{dt} = Av, \tag{3.31}$$

where A is a linear operator on a Hilbert space. (In the Schrödinger case, we have $A = -(i/\hbar)\hat{H}$.) Let us think of (3.31) in the case where the Hilbert space is the finite-dimensional space $\mathbb{C}^n$. In that case, we can think of A as an $n \times n$ matrix, in which case (3.31) is the sort of equation encountered in the elementary theory of ordinary differential equations. The solution of this system (in the finite-dimensional case) can be expressed as

$$v(t) = e^{tA}v_0,$$

where the matrix exponential e^{tA} is defined by a convergent power series and where $v_0 = v(0)$ is the initial condition. If A is diagonalizable, then the exponential can by computed by using a basis of eigenvectors. (See Sect. 16.4 for more information.)

The Schrödinger equation simply replaces $\mathbb{C}^n$ by a Hilbert space $\mathbf{H}$ and the matrix A by the linear operator $-(i/\hbar)\hat{H}$.

Claim 3.17 *Suppose $\hat{H}$ is a self-adjoint operator on $\mathbf{H}$. If a reasonable meaning can be given to the expression $e^{-it\hat{H}/\hbar}$, then the Schrödinger equation can be solved by setting*

$$\psi(t) = e^{-it\hat{H}/\hbar}\psi_0. \tag{3.32}$$

To see why the claim should be true, we expect that we can differentiate the operator-valued expression $e^{-it\hat{H}/\hbar}$ with respect to t as we would in the finite-dimensional case. The differentiation, then, would pull down a factor of $-i\hat{H}/\hbar$, which would indicate that $\psi(t)$ indeed solves the Schrödinger equation. Furthermore, when $t = 0$, $e^{-it\hat{H}/\hbar}$ should be equal to I, so that $\psi(0)$ is indeed ψ_0.

If $\hat{H}$ is a bounded operator (which is rarely the case), then the exponential $e^{-it\hat{H}/\hbar}$ can be defined by a convergent power series, precisely as in the finite-dimensional case. In that case, Claim 3.17 is an easily proved theorem.

In the more typical case where $\hat{H}$ is unbounded, convergence of the series for the exponential is a rather delicate matter, and it is better instead to use the spectral theorem. We leave a general discussion of the spectral theorem to Chaps. 7 and 10, and here consider only the case of a pure point spectrum. A (possibly unbounded) self-adjoint operator $\hat{H}$ is said to have a pure point spectrum if there exists an orthonormal basis $\{e_j\}$ for $\mathbf{H}$ consisting of eigenvectors for $\hat{H}$. If $\hat{H}e_j = E_j e_j$ for some $E_j \in \mathbb{R}$, then the exponential can be defined by requiring that

$$e^{-it\hat{H}/\hbar}e_j = e^{-itE_j/\hbar}e_j. \tag{3.33}$$

The operator $e^{-it\hat{H}/\hbar}$ is unitary and thus bounded; it is the unique bounded operator on $\mathbf{H}$ satisfying (3.33).

It is not precisely true that every self-adjoint operator has an orthonormal basis of eigenvectors, even if the operator is bounded. Nevertheless, given a self-adjoint operator A, the spectral theorem tells us that there is a decomposition of $\mathbf{H}$ into "generalized eigenspaces" for A. It is, however, a bit complicated to state the precise sense of this decomposition, especially in the case of unbounded operators. Still, Claim 3.17 allows us to identify one goal for the spectral theorem: Whatever the spectral theorem says, it ought to allow us to make sense of the expression e^{iaA}, for any self-adjoint operator A and real number a. This goal will indeed be realized, in the bounded case in Chap. 7 and in the unbounded case in Chap. 10.

We should add two points of clarification regarding the expression (3.32). First, in writing (3.32), we have not "really" solved the Schrödinger equation. For this expression to be useful, we need to compute $e^{-it\hat{H}/\hbar}$ in some relatively explicit way. If, for example, we can actually compute an orthonormal basis of eigenvectors for $\hat{H}$, then in light of (3.33), we are on our way to understanding the behavior of the operator $e^{-it\hat{H}/\hbar}$. Second, although $\hat{H}$ is an unbounded operator, which is not defined on all of $\mathbf{H}$ but only on a dense subspace, the operator $e^{-it\hat{H}/\hbar}$ is unitary and defined on all of $\mathbf{H}$. Thus, the right-hand side of (3.32) makes sense for any ψ_0 in $\mathbf{H}$. Nevertheless, we cannot expect that $e^{-it\hat{H}/\hbar}\psi_0$ actually solves the Schrödinger equation (in the natural Hilbert space sense) unless ψ_0 belongs to the domain of $\hat{H}$. (See Lemma 10.17 in Sect. 10.2.)

3.7.3 Eigenvectors and the Time-Independent Schrödinger Equation

As we saw in the preceding section, eigenvectors for the Hamiltonian operator are of great importance in solving the Schrödinger equation. In light of this fact, we make the following definition.

Definition 3.18 *If $\hat{H}$ is the Hamiltonian operator for a quantum system, the eigenvector equation*

$$\hat{H}\psi = E\psi, \quad E \in \mathbb{R}, \tag{3.34}$$

is called the ***time-independent Schrödinger equation****.*

As always in eigenvector equations, we are trying to determine both the numbers E for which (3.34) has a nonzero solution (the eigenvalues) and the corresponding vectors ψ (the eigenvectors). When quantum texts speak of "solving," say, the quantum harmonic oscillator, what they usually mean is finding all of the solutions to the time-independent Schrödinger equation. (See, e.g., Chaps. 5 and 11.) If ψ is a solution to the time-independent Schrödinger equation, then the solution to the time-*dependent* Schrödinger equation with initial condition ψ is simply $\psi(t) = e^{-itE/\hbar}\psi$. Since $\psi(t)$ is just a constant multiple of ψ, we see that $\psi(t)$ represents the same physical state as ψ. Thus, a solution to the time-independent Schrödinger equation is sometimes called a *stationary state*.

3.7.4 The Schrödinger Equation in $\mathbb{R}^1$

Let us now consider the simplest example for the Hamiltonian operator $\hat{H}$. For a particle moving in $\mathbb{R}^1$, recall (Sect. 3.5) that we have identified the position operator X as being multiplication by x and the momentum operator as $P = -i\hbar\, d/dx$. The classical Hamiltonian for such a particle is typically taken to be of the form $H(x,p) = p^2/(2m) + V(x)$, where V is the potential energy function. In that case, we may reasonably take

$$\hat{H} = \frac{P^2}{2m} + V(X).$$

Here the operator $V(X)$ is simply multiplication by the potential energy function $V(x)$. (This operator may also be thought of as the function V applied to the operator X in the sense of the functional calculus coming from the spectral theorem.) We see, then, that

$$\hat{H}\psi(x) = -\frac{\hbar^2}{2m}\frac{d^2\psi}{dx^2} + V(x)\psi(x). \tag{3.35}$$

An operator of the form (3.35), or an analogously defined operator in higher dimensions, is referred to as a *Schrödinger operator*. (The term *Hamiltonian operator* refers more generally to whatever operator governs the time-evolution of a quantum system, regardless of its form.)

If our Hamiltonian is of the form given in (3.35), then the time-dependent Schrödinger equation takes the form

$$\frac{\partial\psi(x,t)}{\partial t} = \frac{i\hbar}{2m}\frac{\partial^2\psi(x,t)}{\partial x^2} - \frac{i}{\hbar}V(x)\psi(x,t), \tag{3.36}$$

which is a linear partial differential equation. By contrast, Newton's equation for a particle in $\mathbb{R}^1$ is a typically nonlinear ordinary differential equation.

For a particle in $\mathbb{R}^1$, the time-independent Schrödinger equation is an ordinary differential equation, one that is linear but that has nonconstant coefficients, unless V happens to be constant. For simple examples of the potential function V, there are relatively standard methods of ordinary differential equations that can be brought to bear on the time-independent Schrödinger equation.

3.7.5 Time-Evolution of the Expected Position and Expected Momentum

Since a quantum particle does not have a fixed position or momentum, it does not make sense to ask whether the particle satisfies Newton's equation. It does, however, make sense to ask whether the *expected values* of the position and momentum satisfy Newton's equation (in the form of Hamilton's equations).

Proposition 3.19 *Suppose $\psi(t)$ is a solution to the Schrödinger equation (3.36) for a sufficiently nice potential V and for a sufficiently nice initial condition $\psi(0) = \psi_0$. Then the expected position and expected momentum in the state $\psi(t)$ satisfy*

$$\frac{d}{dt}\langle X\rangle_{\psi(t)} = \frac{1}{m}\langle P\rangle_{\psi(t)} \tag{3.37}$$

$$\frac{d}{dt}\langle P\rangle_{\psi(t)} = -\langle V'(X)\rangle_{\psi(t)}\,. \tag{3.38}$$

The assumptions in the proposition are there for two reasons: First, to ensure that $\hat{H}$ is actually a self-adjoint operator (see Sect. 9.9) and second, to ensure that the domain assumptions in Proposition 3.14 are satisfied. If we assume, for example, that $V(x)$ is a bounded-below polynomial in x and that ψ_0 belongs to the Schwartz space (A.15), then both of these concerns will be taken care of. Once these technicalities are addressed, the proof of Proposition 3.19 is a straightforward application of Proposition 3.14; see Exercise 4. Note that (3.37) says that in a certain sense, the velocity of a quantum particle is $1/m$ times the momentum, just as in the classical case.

At first glance, it might appear that the pair $(\langle X\rangle_{\psi(t)}, \langle P\rangle_{\psi(t)})$ is a solution to Hamilton's equations, and indeed (3.37) is precisely what Hamilton's equations require. To get a solution to Hamilton's equations, however, we would need the right-hand side of (3.38) to equal $-V'(\langle X\rangle_{\psi(t)})$. But in general,

$$\langle V'(X)\rangle_\psi \neq V'(\langle X\rangle_\psi).$$

Consider, for example, the case $V'(x) = x^3 + x^2$. If ψ is an even function, then $\langle X\rangle_\psi = 0$ and so $V'(\langle X\rangle_\psi) = 0$. But $\langle X^3 + X^2\rangle_\psi$ will not be

zero, because the X^3 term will be zero and the X^2 term will be positive. We conclude, then, that $\langle X \rangle_{\psi(t)}$ and $\langle P \rangle_{\psi(t)}$ usually *do not* evolve along solutions to Hamilton's equations.

There is, however, one case in which $\langle V'(X) \rangle_\psi$ coincides with $V'(\langle X \rangle_\psi)$, and that is the case in which V is quadratic, in which case V' is linear. In that case we have

$$\langle V'(X) \rangle_\psi = \langle aX + bI \rangle_\psi = a \langle X \rangle_\psi + b = V'(\langle X \rangle_\psi).$$

Thus, the expected position and expected momentum *do* follow classical trajectories in the case of a quadratic potential. It is not surprising that this case is special in quantum mechanics, since it is also special in classical mechanics; this is the case in which Newton's law is a linear differential equation.

Although the expected position and expected momentum do not (in general) exactly follow classical trajectories, they will do so approximately under certain conditions. If the wave function $\psi(x)$ is concentrated mostly near a single point $x = x_0$, then $\langle V'(X) \rangle_\psi$ and $V'(\langle X \rangle_\psi)$ will both be approximately equal to $V'(x_0)$. In that case, the expected position and expected momentum of the particle will *approximately* follow a classical trajectory, at least for as long as the wave function remains concentrated near a single point.

3.8 The Heisenberg Picture

The "Heisenberg picture" of quantum mechanics is based on Heisenberg's matrix model of quantum mechanics (Sect. 1.3). In the Heisenberg picture, one thinks of the *operators* (quantum observables) as evolving in time, while the vectors in the Hilbert space (quantum states) remain independent of time. This is to be contrasted with the approach to quantum mechanics we have been using up to now (the "Schrödinger picture"), in which the observables are independent of time and the states evolve in time.

Definition 3.20 *In the Heisenberg picture, each self-adjoint operator A evolves in time according to the operator-valued differential equation*

$$\frac{dA(t)}{dt} = \frac{1}{i\hbar}[A(t), \hat{H}], \tag{3.39}$$

where $\hat{H}$ is the Hamiltonian operator of the system, and where $[\cdot, \cdot]$ is the commutator, given by $[A, B] = AB - BA$.

Note that since $\hat{H}$ commutes with itself, the operator $\hat{H}$ remains constant in time, even in the Heisenberg picture. This observation is the quantum counterpart to the fact that the classical Hamiltonian H remains constant along a solution of Hamilton's equations.

Given the self-adjoint operator $\hat{H}$, the spectral theorem will give us a way to construct a family of unitary operators $e^{-it\hat{H}/\hbar}$, $t \in \mathbb{R}$, and this family of operators computes the time-evolution of states in the Schrödinger picture (Sect. 3.7.2). It is easy to check (at least formally) that the solution to (3.39) can be expressed as

$$A(t) = e^{it\hat{H}/\hbar} A e^{-it\hat{H}/\hbar}. \tag{3.40}$$

Now, if ψ is the state of the system (now considered to be independent of time), then the expectation of $A(t)$ in the state ψ is defined to be $\langle A(t) \rangle_\psi = \langle \psi, A(t)\psi \rangle$. We may then compute that

$$\begin{aligned} \langle A(t) \rangle_\psi &= \left\langle \psi, e^{it\hat{H}/\hbar} A e^{-it\hat{H}/\hbar} \psi \right\rangle \\ &= \left\langle e^{-it\hat{H}/\hbar} \psi, A e^{-it\hat{H}/\hbar} \psi \right\rangle \\ &= \langle \psi(t), A\psi(t) \rangle , \end{aligned}$$

where $\psi(t)$ is time-evolved state of the system in the Schrödinger picture. Here, we have used that the adjoint of $e^{it\hat{H}/\hbar}$ is $e^{-it\hat{H}/\hbar}$, which is formally clear and which is a consequence of the spectral theorem.

Note that in the Schrödinger picture, $\langle \psi(t), A\psi(t) \rangle$ is the expectation value of A in the state $\psi(t)$. We conclude, then, that the Heisenberg picture and the Schrödinger picture give rise to precisely the same expectation values for observables as a function of time, and are therefore physically equivalent. Although we will work primarily with the Schrödinger picture of quantum mechanics, the Heisenberg picture is also important, for example, in quantum field theory.

Proposition 3.21 *Suppose $\hat{H} = P^2/(2m) + V(X)$, where V is a bounded-below polynomial. Then for any $t \in \mathbb{R}$ we have*

$$\hat{H} = \frac{1}{2m} \left(P(t) \right)^2 + V(X(t)). \tag{3.41}$$

Note that since $[\hat{H}, \hat{H}] = 0$, the Hamiltonian $\hat{H}$ is independent of time, even in the Heisenberg picture. Thus, the right-hand side of (3.41) is actually independent of t, even though $P(t)$ and $X(t)$ depend on t. Equation (3.41) holds also for sufficiently nice nonpolynomial functions V, but some limiting argument would be required in the proof. The assumption that V be bounded below is to ensure that $\hat{H}$ is actually an (essentially) self-adjoint operator; compare Sect. 9.10.

Lemma 3.22 *Suppose A is a self-adjoint operator on $\mathbf{H}$ and that $A(\cdot)$ is a solution to (3.39) with $A(0) = A$. Then for any positive integer m, the map*

$$t \mapsto (A(t))^m$$

is also a solution to (3.39).

That is to say, the time-evolution of the mth power of A is the same as the mth power of the time-evolution of A; that is, $A^m(t) = (A(t))^m$.
Proof. If we use (3.40), then the result holds because

$$\begin{aligned} e^{it\hat{H}/\hbar}A^m e^{-it\hat{H}/\hbar} &= e^{it\hat{H}/\hbar}Ae^{-it\hat{H}/\hbar}e^{it\hat{H}/\hbar}Ae^{-it\hat{H}/\hbar}\cdots e^{it\hat{H}/\hbar}Ae^{-it\hat{H}/\hbar} \\ &= \left(e^{it\hat{H}/\hbar}Ae^{-it\hat{H}/\hbar}\right)^m. \end{aligned}$$

It is also easy to check that $A(t)^m$ satisfies the differential equation (3.39). ■

With this lemma in hand, it is easy to prove the proposition.
Proof of Proposition 3.21. On the one hand, since $[\hat{H}, \hat{H}] = 0$, the time-evolved operator $\hat{H}(t)$ is simply equal to $\hat{H}$. On the other hand, if we time-evolve $P^2/(2m) + V(X)$ using Lemma 3.22, we obtain the expression on the right-hand side of (3.41). ■

Proposition 3.23 *Suppose the Hamiltonian of a quantum system is as in Proposition 3.21. Then the operators $X(t)$ and $P(t)$ defined by (3.39) satisfy the following operator-valued differential equation:*

$$\begin{aligned} \frac{dX}{dt} &= \frac{1}{m}P(t) \\ \frac{dP}{dt} &= -V'(X(t)). \end{aligned} \tag{3.42}$$

Proof. See Exercise 7. ■

Proposition 3.23 means that the operator-valued functions $X(t)$ and $P(t)$ satisfy the operator analogs of the classical equations of motion $dx/dt = p(t)/m$ and $dp/dt = -V'(x(t))$. Nevertheless, the *expectation values* of $X(t)$ and $P(t)$ *do not* satisfy the ordinary equations of motion, as we have already seen by calculating in the Schrödinger picture. If we take expectation values in the system (3.42), we get the same answer as in Proposition 3.19, namely,

$$\begin{aligned} \frac{d}{dt}\langle X(t)\rangle_\psi &= \frac{1}{m}\langle P(t)\rangle_\psi \\ \frac{d}{dt}\langle P(t)\rangle_\psi &= -\langle V'(X(t))\rangle_\psi. \end{aligned}$$

These are *not* the classical equations of motion, unless the expectation value of the operator $V'(X(t))$ coincides with V' applied to the expectation value of $X(t)$, which is usually not the case.

3.9 Example: A Particle in a Box

Let us consider quantum mechanics in one space dimension for a particle that is confined to move in a "box," which we describe as the interval $0 \le x \le L$. Our goal is to find all of the eigenvectors and eigenvalues of

the Schrödinger operator, that is, to find solutions of the time-independent Schrödinger equation $\hat{H}\psi = E\psi$. In solving this equation, we may think of the constraint to the box as follows. Imagine a particle moving in $\mathbb{R}^1$ in the presence of a potential V that is 0 for x between 0 and L and takes some very large constant value C on the rest of the real line. Classically, this would mean that the particle has to have very high energy (greater than C) to escape from the box. Quantum mechanically, if we have a solution of the time-independent Schrödinger equation $\hat{H}\psi = E\psi$ for this potential (with $E \ll C$), then we expect ψ to decay rapidly for x outside of the box. (We will see this behavior explicitly in Chap. 5.) In the limit as C tends to infinity, we expect solutions of the time-independent Schrödinger equation to be zero outside the box and to *tend to zero as we approach the ends of the box*.

The upshot of this discussion is that we are looking for smooth functions ψ on $[0, L]$ that satisfy the differential equation

$$-\frac{\hbar^2}{2m}\frac{d^2\psi}{dx^2} = E\psi(x), \quad 0 \leq x \leq L \tag{3.43}$$

and the boundary conditions

$$\psi(0) = \psi(L) = 0. \tag{3.44}$$

For $E > 0$, the solution space to (3.43) will be the span of two complex exponentials, or equivalently a sine and a cosine function:

$$\psi(x) = a \sin\left(\frac{\sqrt{2mE}}{\hbar}x\right) + b\cos\left(\frac{\sqrt{2mE}}{\hbar}x\right). \tag{3.45}$$

If we now impose the boundary condition $\psi(0) = 0$, we get that $b = 0$, leaving only the sine term. If we then impose the condition $\psi(L) = 0$, we will obtain $a = 0$—which would mean that ψ is identically zero—*unless*

$$\sin\left(\frac{\sqrt{2mE}}{\hbar}L\right) = 0. \tag{3.46}$$

Since we are interested in solutions to (3.43) where ψ is not identically zero, we want (3.46) to hold. Thus, the argument of sine function must be an integer multiple of π. This condition imposes a restriction on the value of E, namely that E should be of the form

$$E_j := \frac{j^2\pi^2\hbar^2}{2mL^2}, \tag{3.47}$$

for some positive integer j.

It is a simple exercise (Exercise 8) to verify that for $E \leq 0$, the only solution to (3.43) satisfying the boundary conditions (3.44) is the one with ψ identically zero.

Proposition 3.24 *The following functions are solutions to (3.43) satisfying the boundary conditions (3.44):*

$$\psi_j(x) = \sqrt{\frac{2}{L}} \sin\left(\frac{j\pi x}{L}\right), \quad j = 1, 2, 3, \ldots,$$

and the corresponding eigenvalues E_j are given by (3.47). The functions ψ_j form an orthonormal basis for the Hilbert space $L^2([0, L])$.

Proof. We have already verified the equation and eigenvalue for each ψ_j. It is a simple computation to verify that the ψ_j's are orthonormal, and the elementary theory of Fourier series (Fourier sine series, in this case) shows that the ψ_j's form an orthonormal *basis* for $L^2([0, L])$. ■

The Hamiltonian operator for this problem (in which $V = 0$ inside the box) is given by

$$\hat{H}\psi = -\frac{\hbar^2}{2m}\frac{d^2\psi}{dx^2}.$$

This operator is an unbounded operator and is not defined on the whole Hilbert space $L^2([0, L])$, but only on a dense subspace $\mathrm{Dom}(\hat{H}) \subset L^2([0, L])$. The domain of $\hat{H}$ should be chosen in such a way that $\hat{H}$ is essentially self-adjoint and, thus, symmetric (Sect. 3.2), meaning that

$$\left\langle \phi, \hat{H}\psi \right\rangle = \left\langle \hat{H}\phi, \psi \right\rangle \tag{3.48}$$

for all ϕ, ψ in $\mathrm{Dom}(\hat{H})$. For (3.48) to hold, ϕ and ψ must satisfy appropriate boundary conditions, which will allow the boundary terms in the integration by parts to be zero. (See Exercise 9.)

Mathematically, then, it is necessary to impose some boundary conditions in order for $\hat{H}$ to be an essentially self-adjoint operator. The particular choice of boundary conditions (3.44) is based on the idea of approximating the box by a very large "confining" potential outside the box. See Chap. 9 for an extensive discussion of domain issues for unbounded operator.

3.10 Quantum Mechanics for a Particle in $\mathbb{R}^n$

Up to this point, we have been considering a quantum particle moving in $\mathbb{R}^1$. It is straightforward, however, to generalize to a quantum particle moving in $\mathbb{R}^n$. The Hilbert space for a particle in $\mathbb{R}^n$ is $L^2(\mathbb{R}^n)$, rather than $L^2(\mathbb{R})$. Instead of single position operator, we have n such operators, given by

$$X_j\psi(\mathbf{x}) = x_j\psi(\mathbf{x}), \quad j = 1, \ldots, n.$$

Similarly, we have n momentum operators, given by

$$P_j\psi(\mathbf{x}) = -i\hbar\frac{\partial\psi}{\partial x_j}.$$

As in the $\mathbb{R}^1$ case, X_j does not commute with P_j but satisfies $[X_j, P_j] = i\hbar I$. On the other hand, X_j commutes with X_k and P_j commutes with P_k. Furthermore, X_j commutes with P_k for $j \neq k$. These formulas are referred to as the *canonical commutation relations*.

Proposition 3.25 (Canonical Commutation Relations) *The position and momentum operators satisfy*

$$\begin{aligned} \frac{1}{i\hbar}[X_j, X_k] &= 0 \\ \frac{1}{i\hbar}[P_j, P_k] &= 0 \\ \frac{1}{i\hbar}[X_j, P_k] &= \delta_{jk} I \end{aligned} \tag{3.49}$$

for all $1 \leq j, k \leq n$.

These relations are the quantum counterparts of the Poisson bracket relations among the position and momentum *functions* in classical mechanics. Specifically, the role of the Poisson bracket in Proposition 2.24 is played in Proposition 3.25 by the quantity $(1/(i\hbar))[\cdot, \cdot]$.

If the classical Hamiltonian for a particle in $\mathbb{R}^n$ is of the usual form (kinetic energy plus potential energy), then we may analogously define the Hamiltonian operator to be of the form

$$\hat{H} = \sum_{j=1}^{n} \frac{P_j^2}{2m} + V(\mathbf{X}), \tag{3.50}$$

where $V(\mathbf{X})$ denotes the result of applying the function V to the commuting family of operators $\mathbf{X} = (X_1, \ldots, X_n)$. It it natural to identify $V(\mathbf{X})$ with the operator of multiplication by the function $V(\mathbf{x})$. In that case, we may write $\hat{H}$ more explicitly as

$$\hat{H}\psi(\mathbf{x}) = -\frac{\hbar}{2m}\Delta\psi(\mathbf{x}) + V(\mathbf{x})\psi(\mathbf{x}),$$

where Δ is the Laplacian, given by

$$\Delta = \sum_{j=1}^{n} \frac{\partial^2}{\partial x_j^2}.$$

We refer to an operator of the form (3.50) as a Schrödinger operator.

We may also introduce angular momentum operators defined by analogy to the classical angular momentum functions.

Definition 3.26 *For each pair* (j, k) *with* $1 \leq j, k \leq n$, *define the* ***angular momentum operator*** $\hat{J}_{jk}$ *by the formula*

$$\hat{J}_{jk} = X_j P_k - X_k P_j.$$

As in the classical case, we have $\hat{J}_{jk} = 0$ when $j = k$. When $j \neq k$, X_j and P_k commute, so the order of the factors in the definition of $\hat{J}_{jk}$ is not important. Explicitly, we have

$$\hat{J}_{jk} = -i\hbar \left(x_j \frac{\partial}{\partial x_k} - x_k \frac{\partial}{\partial x_j} \right).$$

The operator in parentheses is the angular derivative $(\partial/\partial\theta)$ in the (x_j, x_k) plane.

When $n = 3$, it is customary to use the quantum counterpart of the classical angular momentum *vector*, namely,

$$\hat{J}_1 := X_2P_3 - X_3P_2; \quad \hat{J}_2 := X_3P_1 - X_1P_3; \quad \hat{J}_3 := X_1P_2 - X_2P_1. \tag{3.51}$$

When $n = 3$, every $\hat{J}_{jk}$ with $j \neq k$ is one of the above three operators or the negative thereof.

3.11 Systems of Multiple Particles

Suppose now we have a system of N quantum particles moving in $\mathbb{R}^n$. If the particles are all of different types (e.g., one electron and one proton), then the Hilbert space for this system is $L^2(\mathbb{R}^{nN})$. That is, the wave function ψ of the system is a function of variables $\mathbf{x}^1, \mathbf{x}^2, \ldots, \mathbf{x}^N$, with each $\mathbf{x}^j$ belonging to $\mathbb{R}^n$. If we normalize ψ to be a unit vector in $L^2(\mathbb{R}^{nN})$, then $|\psi(\mathbf{x}^1, \mathbf{x}^2, \ldots, \mathbf{x}^N)|^2$ is to be interpreted as the joint probability distribution for the positions of the N particles.

We may introduce position operators X_k^j (the kth component of the position of the jth particle) and momentum operators P_k^j in obvious analogy to the definition for a single particle. The typical Hamiltonian operator for such a system is then

$$\hat{H}\psi(\mathbf{x}^1, \ldots, \mathbf{x}^N) = -\sum_{j=1}^{N} \frac{\hbar^2}{2m_j} \Delta_j \psi(\mathbf{x}^1, \ldots, \mathbf{x}^N) + V(\mathbf{x}^1, \ldots, \mathbf{x}^N)\psi(\mathbf{x}),$$

where m_j is the mass of the jth particle. Here Δ_j means the Laplacian with respect to the variable $\mathbf{x}^j \in \mathbb{R}^n$, with the other variables fixed.

As we will see in Chap. 19, the Hilbert space for a composite system, made up of various subsystems, is typically taken to be the (Hilbert) *tensor product* of the individual Hilbert spaces. In the present context, we may think of our system of being made up of N subsystems, each being one of the individual particles. Fortunately, there is a natural isomorphism (Proposition 19.12) between $L^2(\mathbb{R}^{nN})$ and the tensor product of N copies of $\mathbb{R}^n$, so that the approach we are taking here is consistent with the general philosophy.

If the particles in question are identical (say, all electrons), then there is an additional complication to the description of the Hilbert space for the system. In standard quantum theory, we are supposed to believe that "identical particles are indistinguishable." What this means is that the wave function should have the property that if we interchange, say, $\mathbf{x}^1$ with $\mathbf{x}^2$, then the new wave function should represent the same physical state as the original wave function. Recalling that two unit vectors in the quantum Hilbert space represent the same physical state if and only if they differ by a constant of absolute value 1, this means we should have

$$\psi(\mathbf{x}^2, \mathbf{x}^1, \mathbf{x}^3, \ldots, \mathbf{x}^N) = u\psi(\mathbf{x}^1, \mathbf{x}^2, \mathbf{x}^3, \ldots, \mathbf{x}^N),$$

for some constant u with $|u| = 1$. Applying this rule twice gives that ψ is $u^2\psi$, so evidently u must be either 1 or -1.

Particles in quantum mechanics are grouped into two types, according to whether the constant u in the previous paragraph is 1 or -1. Particles with $u = 1$ are called *bosons* and particles with $u = -1$ are called *fermions*. Whether a particle is a boson or a fermion is determined by the *spin* of the particle, a concept that we have not yet introduced. Nevertheless, we can say that particles without spin are bosons. For a collection of N identical spinless particles moving in $\mathbb{R}^3$, the proper Hilbert space is the *symmetric subspace* of $L^2(\mathbb{R}^{3N})$, that is, the space of functions in $L^2(\mathbb{R}^{3N})$ that are invariant under arbitrary permutations of the variables. We will have more to say about spin and systems of identical particles in Chaps. 17 and 19.

3.12 Physics Notation

In quantum mechanics, physicists almost invariably use the *Dirac notation* (or *bra-ket notation*) introduced by Dirac in 1939 [5]. This notation is made up of Notations 3.27–3.29 below. In this section, we explore the Dirac notation along with a few other notational differences between the mathematics and physics literature.

Before proceeding it is important to point out that when using Dirac notation, it is essential that the complex conjugate in the inner product should go on the *first* factor.

Notation 3.27 *A vector ψ in $\mathbf{H}$ is referred to as a **ket** and is denoted $|\psi\rangle$. A continuous linear functional on $\mathbf{H}$ is called a **bra**. For any $\phi \in \mathbf{H}$, let $\langle\phi|$ denote the bra given by*

$$\langle\phi|\,(\psi) = \langle\phi, \psi\rangle\,.$$

That is to say, $\langle\phi|$ is the "inner product with ϕ" functional. The bracket (or bra-ket) of two vectors $\phi, \psi \in \mathbf{H}$ is the result of applying the bra $\langle\phi|$ to the ket $|\psi\rangle$, namely the inner product of the ϕ and ψ, denoted $\langle\phi|\psi\rangle$.

If A is an operator on $\mathbf{H}$ and ϕ is a vector in $\mathbf{H}$, then we can form the linear functional $\langle\phi| A$, i.e., the linear map $\psi \mapsto \langle\phi|A\psi\rangle$. Physicists generally write an expression of this form as

$$\langle\phi\,|A|\,\psi\rangle\,.$$

This notation emphasizes that there are two different ways of thinking of this quantity. We may think of $\langle\phi\,|A|\,\psi\rangle$ either as the linear functional $\langle\phi| A$ applied to the vector $|\psi\rangle$, or as the linear functional $\langle\phi|$ applied to the vector $A\,|\psi\rangle$.

Notation 3.28 *For any ϕ and ψ in $\mathbf{H}$, the expression $|\phi\rangle\langle\psi|$ denotes the linear operator on $\mathbf{H}$ given by*

$$(|\phi\rangle\langle\psi|)\,(\chi) = |\phi\rangle\langle\psi|\chi\rangle = \langle\psi|\chi\rangle\,|\phi\rangle\,.$$

That is, in mathematics notation, $|\phi\rangle\langle\psi|$ is the operator sending χ to $\langle\psi,\chi\rangle\,\phi$.

The operator $|\phi\rangle\langle\psi|$ associates to each (ket) vector $|\chi\rangle$ a new vector in the only way that makes notational sense: We interpret $|\phi\rangle\langle\psi||\chi\rangle$ as the vector $|\phi\rangle$ multiplied by the scalar $\langle\psi|\chi\rangle$.

Notation 3.29 *Given a family of vectors in $\mathbf{H}$ labeled by, say, three indices n, l, and m, rather than denoting these vectors as $|\psi_{n,l,m}\rangle$, a physicist will denote them simply as $|n,l,m\rangle$.*

This notation is not without its pitfalls. If we have two different sets of vectors labeled by the same set of indices, a mathematician can simply label them as $\phi_{n,l,m}$ and $\psi_{n,l,m}$, but the physicist has a problem.

As an example of the Dirac notation, suppose that an operator $\hat{H}$ has an orthonormal basis of eigenvectors ψ_n. A physicist would express the decomposition of a general vector in terms of this basis as

$$I = \sum_n |n\rangle\langle n|\,, \tag{3.52}$$

where ψ_n is represented simply as $|n\rangle$ and where $|n\rangle\langle n|$ is (given that $|n\rangle$ is a unit vector) the orthogonal projection onto the one-dimensional subspace spanned by the vector $|n\rangle$.

Notation 3.30 *In the physics literature, the complex conjugate of a complex number z is denoted as z^*, rather than $\bar{z}$, as in the mathematics literature. What a mathematician calls the adjoint of an operator and denotes by A^*, a physicist calls the Hermitian conjugate of A and denotes by $A^\dagger$. Physicists refer to self-adjoint operators as Hermitian.*

We may express the concept of an adjoint (or Hermitian conjugate) of an operator using Dirac notation, as follows. If A is a bounded operator on $\mathbf{H}$, then $A^\dagger$ is the unique bounded operator such that

$$\langle\psi|\,A = \left\langle A^\dagger\psi\right|.$$

One peculiarity of the physics literature on quantum mechanics is a conspicuous failure of most articles to state what the Hilbert space is. Rather than starting by defining the Hilbert space in which they are working, physicists generally start by writing down the commutation relations that hold among various operators on the space. Thus, for example, a physicist might begin with position and momentum operators X and P, satisfying $[X, P] = i\hbar I$, without ever specifying what space these operators are operating on. The justification for this omission is, presumably, the Stone–von Neumann theorem, which asserts that (provided the operators satisfy the expected "exponentiated" relations) there is, up to unitary equivalence, only one Hilbert space with operators satisfying these relations and on which the operators act irreducibly. (See Chap. 14 for a precise statement of the result.) It is, nevertheless, disconcerting for a mathematician to encounter an entire paper full of computations involving certain operators, without any specification of what space these operators are operating on, let alone *how* the operators act on the space.

This practice among physicists represents something of a role reversal. In the setting of linear algebra, for example, a mathematician might say, "Let V be a n-dimensional vector space over $\mathbb{R}$." If a physicist says, "Oh, so it's $\mathbb{R}^n$," the mathematician will reply, "No, no, you don't have to choose a basis." By contrast, in quantum mechanics, it is the physicist who does not want to choose a particular realization of the space. A physicist will simply write down the commutation relations between, say, X and P. If pressed, the physicist might say that he is working in an irreducible representation of those relations. If a mathematician then says, "Oh, so it's $L^2(\mathbb{R})$," the physicist will reply, "No, no, there is no preferred realization."

Notation 3.31 *Given an irreducible representation of the canonical commutation relations, and given a vector ψ in the corresponding Hilbert space, a physicist will speak of the position wave function $\psi(x)$, defined by*

$$\psi(x) = \langle x|\psi\rangle \,. \tag{3.53}$$

Here, $\langle x|$ is the bra associated with the ket $|x\rangle$, where $|x\rangle$ is supposed to be an eigenvector for the position operator with eigenvalue x.

See, again, Chap. 14 for the precise notion of "irreducible representation of the canonical commutation relations." One may similarly define the *momentum wave function* by taking the inner product of ψ with the eigenvectors of the momentum operator, which are also non-normalizable. See Sect. 6.6 for details.

A mathematician might find Notation 3.31 objectionable on the grounds that the operator X does not actually have any eigenvectors. After all, it is harmless, in view of the Stone–von Neumann theorem, to work in the "Schrödinger representation," in which our Hilbert space is $L^2(\mathbb{R})$ and the position operator X is just multiplication by x. Given a number x_0,

there is no nonzero element ψ of $L^2(\mathbb{R})$ for which $X\psi = x_0\psi$. After all, any ψ satisfying this equation would have to be supported at the point $x = x_0$, in which case ψ would equal zero almost everywhere and would be the zero element of $L^2(\mathbb{R})$. A physicist, on the other hand, would say that the desired eigenfunction is $\psi(x) = \delta(x - x_0)$, where δ is the Dirac delta-"function." The fact that $\delta(x - x_0)$ is not actually *in* the Hilbert space $L^2(\mathbb{R})$ does not concern the physicist; it is simply a "non-normalizable state." The mathematical theory of such non-normalizable states comes under the heading "generalized eigenvectors." See Sect. 6.6 for a discussion of this issue in the case of the eigenvectors of the momentum operator.

A more subtle issue regarding the "position eigenvectors" is that each eigenvector is unique only up to multiplication by a constant. If one wants the momentum operator to act on the position wave function, as defined by (3.53), in the usual way, one must make a consistent choice of normalization of the eigenvectors of the position operators. Specifically, one should choose the constants in such a way that the exponentiated momentum operator $\exp(iaP/\hbar)$ maps $|x\rangle$ to $|x + a\rangle$.

3.13 Exercises

1. Suppose that $\phi(t)$ and $\psi(t)$ are differentiable functions with values in a Hilbert space $\mathbf{H}$, meaning that the limit

$$\frac{d\phi}{dt} := \lim_{h\to 0} \frac{\phi(t+h) - \phi(t)}{h}$$

exists in the norm topology of $\mathbf{H}$ for each t, and similarly for $\psi(t)$. Show that

$$\frac{d}{dt}\langle \phi(t), \psi(t)\rangle = \left\langle \frac{d\phi}{dt}, \psi(t) \right\rangle + \left\langle \phi(t), \frac{d\psi}{dt} \right\rangle.$$

2. Suppose A and B are operators on a *finite-dimensional* Hilbert space and suppose that $AB - BA = cI$ for some constant c. Show that $c = 0$.

   Note: This shows that the commutation relations in (3.8) are a purely infinite-dimensional phenomenon.

3. If A is a bounded operator on a Hilbert space $\mathbf{H}$, then there exists a unique bounded operator A^* on $\mathbf{H}$ satisfying $\langle \phi, A\psi\rangle = \langle A^*\phi, \psi\rangle$ for all ϕ and ψ in $\mathbf{H}$. (Appendix A.4.3.) The operator A^* is called the *adjoint* of A, and A is called *self-adjoint* if $A^* = A$.

   (a) Show that for any bounded operator A and constant $c \in \mathbb{C}$, we have $(cA)^* = \bar{c}A^*$, where $\bar{c}$ is the complex conjugate of c.

(b) Show that if A and B are self-adjoint, then the operator

$$\frac{1}{i\hbar}[A, B]$$

is also self-adjoint.

4. Verify Proposition 3.19 using Proposition 3.14. Note that the operator $V'(X)$ means simply the operator of multiplication by the function $V'(x)$.

5. Suppose that ψ is a unit vector in $L^2(\mathbb{R})$ such that the functions $x\psi(x)$ and $x^2\psi(x)$ also belong to $L^2(\mathbb{R})$. Show that

$$\left\langle X^2 \right\rangle_\psi > \left(\langle X \rangle_\psi \right)^2 .$$

Hint: Consider the integral

$$\int_{-\infty}^{\infty} (x-a)^2 \left|\psi(x)\right|^2 \, dx,$$

where $a = \langle X \rangle_\psi$.

6. Consider the Hamiltonian $\hat{H}$ for a quantum harmonic oscillator, given by

$$\hat{H} = -\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + \frac{k}{2}x^2,$$

where k is the spring constant of the oscillator. Show that the function

$$\psi_0(x) = \exp\left\{ -\frac{\sqrt{km}}{2\hbar}x^2 \right\}$$

is an eigenvector for $\hat{H}$ with eigenvalue $\hbar\omega/2$, where $\omega := \sqrt{k/m}$ is the classical frequency of the oscillator.

Note: We will explore the eigenvectors and eigenvalues of $\hat{H}$ in detail in Chap. 11.

7. Prove Proposition 3.23.

Hint: Show that $[P(t), \hat{H}] = ([P, \hat{H}])(t)$ and $[X(t), \hat{H}] = ([X, \hat{H}])(t)$.

8. (a) Find the general solution to (3.43), where E is a negative real number. Show that the only such solution that satisfies the boundary conditions (3.44) is identically zero.

(b) Establish the same result as in Part (a) for $E = 0$.

9. (a) Suppose ϕ and ψ are smooth functions on $[0, L]$ satisfying the boundary conditions (3.44). Using integration by parts, show that

$$\left\langle \phi, \hat{H}\psi \right\rangle = \left\langle \hat{H}\phi, \psi \right\rangle,$$

where $\hat{H} = -(\hbar^2/2m)\, d^2/dx^2$ and where

$$\langle \phi, \psi \rangle = \int_0^L \overline{\phi(x)}\psi(x)\, dx.$$

(b) Show that the result of Part (a) fails if ϕ and ψ are arbitrary smooth functions (not satisfying the boundary conditions).

10. Let $\hat{J}_1$, $\hat{J}_2$, and $\hat{J}_3$ be the angular momentum operators for a particle moving in $\mathbb{R}^3$. Using the canonical commutation relations (Proposition 3.25), show that these operators satisfy the commutation relations

$$\frac{1}{i\hbar}[\hat{J}_1, \hat{J}_2] = \hat{J}_3; \quad \frac{1}{i\hbar}[\hat{J}_2, \hat{J}_3] = \hat{J}_1; \quad \frac{1}{i\hbar}[\hat{J}_3, \hat{J}_1] = \hat{J}_2.$$

This is the quantum mechanical counterpart to Exercise 19 in the previous chapter.