Chapter 1

Linear algebra and postulates of quantum mechanics

1.1 Introduction

Perhaps the first thing one needs to understand about quantum mechanics is that it has as much to do with mechanics as with, say, electrodynamics, optics, or high energy physics. Rather than describing a particular class of physical phenomena, quantum mechanics provides a *universal theoretical framework*, an "operating system", which can be used in *all* fields of physics — in a fashion similar to mathematics being the universal framework for all natural sciences. The term "quantum mechanics" emerged historically, because first successful applications of this framework were to the motion of electrons in an atom. A better term would be "quantum physics" or "quantum theory".

Quantum treatment of physical phenomena is very different from classical[1]. It is also quite complicated mathematically. Fortunately, however, the predictions of quantum theory are different from classical ones only for relatively simple, microscopic objects. When we increase the complexity of the quantum systems we study, they begin to behave in an increasingly classical fashion, so their quantum description eventually becomes unnecessary. In fact, as we shall see in Chapter 5, the complexity of a physical system is intimately linked to its tendency to lose quantum properties. This is a relief for quantum mechanics students, and also explains why quantum mechanics has not been discovered until the early 20th century: before that time, we (who ourselves are macroscopic entities) have only dealt with macroscopic objects. But as soon as we developed tools to probe the microscopic world, quantum phenomena became manifest.

This is an example of the *correspondence principle*: any new theory should reproduce the results of older well-established theories in those domains where the old theories have been tested. Another example of this principle is the situation with Newtonian mechanics and special relativity. As long as we had to do with objects that move much slower than light, classical mechanics was perfectly sufficient to describe the world around us. But as soon as we became able to observe bodies that move quickly (e.g. the Earth around the sun in the Michelson-Morley experiment), we started to see discrepancies and were forced to develop the theory of relativity. This theory is distinctly different from classical mechanics — yet it is consistent with the latter in the limiting case of low velocities. It would be unwise to use special relativity to describe, for example, a tractor transmission, because the classical approximation is in this case fully sufficient and tremendously simpler. Similarly, using quantum physics to describe macroscopic phenomena would in most cases be overcomplicated and unnecessary.

As discussed above, quantum mechanics is not limited to a particular class of physical systems. However, the quantum treatment gives rise to a certain class of phenomena that are specific to this treatment and are impossible within the framework of classical physics. The most well-known exam-

[1] Under *classical physics* we understand those phenomena whose theoretical description does not contain any quantum features.

ple is perhaps the effect of quantum nonlocality: under certain circumstances, an action performed at a certain place can instantly affect the physical reality at another location that can be very distant from the first one and have no interaction therewith. As a specifically quantum phenomenon, nonlocality is transcendental with respect to subfields of physics: it can be observed with photon polarizations, electron spins, atomic ensembles and many other systems.

Because of quantum mechanics' role as a general framework, we will study it in a fairly rigorous, mathematical fashion. We will introduce definitions and axioms, and then predict phenomena that result from these definitions and axioms, and illustrate these phenomena with examples from different fields of physics. The primary mathematical tool of quantum mechanics is linear algebra. Therefore, in the next three sections we are reviewing its basic concepts. Those students who feel comfortable with their linear algebra skills may wish to skim over these sections; others should perform the exercises to refresh their memory.

1.2 Linear spaces

In this section, we discuss the concept of the linear space. Its definition may appear dry; however, linear spaces are quite easy to visualize as sets of geometric vectors. Like regular numbers, vectors can be added to and subtracted from each other to form new vectors; they can also be multiplied by numbers. Unlike numbers, vectors cannot be multiplied or divided by one another (more exactly, we don't *have to* define multiplication in order to introduce the linear space).

One important peculiarity of the linear algebra used in quantum mechanics is the so-called Dirac notation for vectors[2]. To denote elements of the Hilbert space, instead of writing, for example, $\vec{a}$, we write $|a\rangle$. We shall see later how convenient this notation turns out to be.

Definition 1.1 A *linear (vector) space* $\mathbb{V}$ over a field[3] $\mathbb{F}$ is a set in which the following operations are defined[4]:

1. Addition: $\forall\, |a\rangle, |b\rangle \in \mathbb{V} \qquad \exists$ unique $|a\rangle + |b\rangle \in \mathbb{V}$;
2. Multiplication by a number ("scalar"): $\forall\, |a\rangle \in \mathbb{V}, \forall \lambda \in \mathbb{F} \qquad \exists$ unique $\lambda\, |a\rangle \equiv |a\rangle\, \lambda \in \mathbb{V}$;

These operations must obey the following *axioms*:

1. Commutativity of addition: $|a\rangle + |b\rangle = |b\rangle + |a\rangle$
2. Associativity of addition: $(|a\rangle + |b\rangle) + |c\rangle = |a\rangle + (|b\rangle + |c\rangle)$
3. Existence of the zero element: $\exists\, |\text{zero}\rangle \in \mathbb{V} \qquad \forall\, |a\rangle \in \mathbb{V} \qquad |a\rangle + |\text{zero}\rangle = |a\rangle$
   (**Note**: As an alternative notation for $|\text{zero}\rangle$, we sometimes use "0" but *not* "$|0\rangle$".)
4. Existence of the additive inverse for each element: $\forall\, |a\rangle \;\; \exists\, |b\rangle \;\; |a\rangle + |b\rangle = |\text{zero}\rangle$
   Notation: $|b\rangle \equiv -\, |a\rangle$
5. Distributivity of vector sums: $\lambda(|a\rangle + |b\rangle) = \lambda\, |a\rangle + \lambda\, |b\rangle$;
6. Distributivity of scalar sums: $\qquad (\lambda + \mu)\, |a\rangle = \lambda\, |a\rangle + \mu\, |a\rangle$
7. Associativity of scalar multiplication: $\lambda(\mu\, |a\rangle) = (\lambda\mu)\, |a\rangle$
8. Scalar multiplication identity: for $1 \in \mathbb{F}, \forall\, |a\rangle \in \mathbb{V} \qquad 1 \cdot |a\rangle = |a\rangle$

Definition 1.2 *Subtraction* of vectors in a linear space is defined as follows:

$$|a\rangle - |b\rangle \equiv |a\rangle + (-\, |b\rangle).$$

[2] Paul Dirac (1902-1984) was an English theoretical physicists, one of the founders of quantum mechanics.

[3] *Field* is a term from algebra which means a set of elements that satisfies certain axioms for both addition and multiplication. The sets of rational numbers $\mathbb{Q}$, real numbers $\mathbb{R}$, complex numbers $\mathbb{C}$ are examples of fields. Quantum mechanics usually studies vector spaces over the field of complex numbers.

[4] The symbols $\forall$ and $\exists$ below are called *quantifiers* and mean, respectively "for all" and "there exist".

Exercise 1.1 Which of the following are linear spaces (over the field of complex numbers unless otherwise indicated):

a) $\mathbb{R}$ over $\mathbb{R}$? $\mathbb{R}$ over $\mathbb{C}$? $\mathbb{C}$ over $\mathbb{R}$? $\mathbb{C}$ over $\mathbb{C}$?

b) Polynomial functions? Polynomial functions of degree $\leq n$? $> n$?

c) All functions such that $f(1) = 0$? $f(1) = 1$?

d) All periodic functions of period T?

e) N-dimensional geometric vectors over $\mathbb{R}$?

Exercise 1.2 Prove:

a) there is only one zero in a linear space;

b) if $|a\rangle + |x\rangle = |a\rangle$ for some $|a\rangle \in \mathbb{V}$, then $|x\rangle = |\text{zero}\rangle$;

c) for $0 \in \mathbb{F}$, $\forall |a\rangle \quad 0\,|a\rangle = |\text{zero}\rangle$;

d) $-|a\rangle = (-1)\,|a\rangle$;

e) $-|\text{zero}\rangle = |\text{zero}\rangle$;

f) $\forall |a\rangle,\ -|a\rangle$ is unique.

g) $-(-|a\rangle) = |a\rangle$

h) $|a\rangle = |b\rangle$ if and only if $|a\rangle - |b\rangle = 0$.

1.3 Basis, dimension

Note 1.1 The basis is the smallest subset of a linear space such that all other vectors can be expressed as a linear combination of the basis elements. The term "basis" may suggest that each linear space has only one basis — just like a building can have only one foundation. Actually, as we shall see, in any nontrivial linear space, there are infinitely many infinitely many bases.

Definition 1.3 A set of vectors $|v_i\rangle$ is called *linearly independent* if no nontrivial[5] linear combination $\lambda_1 |v_1\rangle + \ldots + \lambda_N |v_N\rangle$ equals $|\text{zero}\rangle$.

Exercise 1.3 $\{|v_i\rangle\}$ is *not* linearly independent if and only if one of $|v_i\rangle$ can be presented as a linear combination of others.

Exercise 1.4

a) For the linear space of geometric vectors in a plane, show that any two vectors are linearly independent if and only if they are not parallel. Show that any set of three vectors is linearly dependent.

b) For the linear space of geometric vectors in a three-dimensional space, show that any three non-complanar vectors form a linearly independent set.

Definition 1.4 A subset $\{|v_i\rangle\}$ of a vector space $\mathbb{V}$ is called its *spanning set* if any vector in $\mathbb{V}$ can be expressed as a linear combination of $|v_i\rangle$'s.

Exercise 1.5 For the linear space of geometric vectors in a plane, show that any set of at least two vectors, of which at least two are non-parallel, form a spanning set.

Definition 1.5 A *basis* of $\mathbb{V}$ is any linearly independent spanning set. A *decomposition* of a vector into a basis is its expression as a linear combination of the basis elements.

[5] A trivial linear combination is one with all elements being equal to 0.

Definition 1.6 The number of elements in a basis is called the *dimension* of $\mathbb{V}$. Notation: $\dim \mathbb{V}$.

Exercise 1.6* Prove that in a finite-dimensional space, all bases have the same number of elements.

Exercise 1.7 Using the result of Ex. 1.6, show that, in a finite-dimensional space,

a) any linearly independent set of $N = \dim \mathbb{V}$ vectors forms a basis;

b) any spanning set of $N = \dim \mathbb{V}$ vectors forms a basis.

Exercise 1.8 a) For the linear space of geometric vectors in a plane, show that any two non-parallel vectors form a basis.

b) For the linear space of geometric vectors in a three-dimensional space, show that any three non-complanar vectors form a basis.

Note 1.2 In the first half of the course (Chapters 1–2, we will be dealing with linear spaces of finite dimension.

Exercise 1.9 Show that for any element of $\mathbb{V}$, there exists only one decomposition into basis vectors.

Definition 1.7 For a decomposition

$$|a\rangle = \sum_i a_i |v_i\rangle , \tag{1.1}$$

we may use the notation

$$|a\rangle \leftrightarrow \begin{pmatrix} a_1 \\ \vdots \\ a_N \end{pmatrix} . \tag{1.2}$$

This is called the *matrix form* of a vector. The quantities a_i are called the *coefficients* or *amplitudes* of the decomposition.

Note 1.3 When expressing states and operators in the matrix form [such as e.g. in Eq. (1.2)], people frequently use "=" instead of "$\leftrightarrow$". We shall sometimes do the same, although it is not strictly correct because the decomposition depends on the basis while the vector itself does not. A decomposition fully identifies a vector only if we know the decomposition basis.

Exercise 1.10 In the basis $\{|v_i\rangle\}$, one of the elements, $|v_k\rangle$, is chosen. Find the matrix form of the decomposition of this element in this basis.

Exercise 1.11 Consider the linear space of two-dimensional geometric vectors. In geometry, such vectors are usually defined by two numbers (x, y), which correspond to the x- and y-components of the vector, respectively. Does this notation correspond to a decomposition into any basis? If so, which one?

Exercise 1.12 Consider the linear space of two-dimensional geometric vectors. The vectors $\vec{a}$, $\vec{b}$, $\vec{c}$, $\vec{d}$ are oriented with respect to the x axis at angles 0, 45°, 90°, 180° and have lengths $2, 1, 3, 1$, respectively. Do sets $\{\vec{a}, \vec{c}\}$, $\{\vec{b}, \vec{d}\}$, $\{\vec{a}, \vec{d}\}$ form bases? Find decompositions of vector $\vec{b}$ in each of these bases. Express them in the matrix form.

1.4 Inner Product

Although vectors cannot be multiplied by each other in the same way that numbers can, one can define a multiplication operation that maps any pair of vectors onto a number. This operation generalizes the scalar product that is familiar from geometry.

Definition 1.8 $\forall |a\rangle,\ |b\rangle \in \mathbb{V}$ we define an *inner (scalar) product* (or, more informally, in the context of quantum physics, an *overlap*) $\langle a|\ b\rangle \in \mathbb{C}$ such that:

1. $\langle a|\ (|b\rangle + |c\rangle) = \langle a|\ b\rangle + \langle a|\ c\rangle$
2. $\langle a|\ (\lambda\,|b\rangle) = \lambda\,\langle a|\ b\rangle$
3. $\langle a|\ b\rangle = \langle b|\ a\rangle^*$
4. $\langle a|\ a\rangle$ is a real number; $\quad \forall |a\rangle \quad \langle a|\ a\rangle \geq 0; \quad \langle a|\ a\rangle = 0$ if and only if $|a\rangle = 0$

Exercise 1.13 In geometry, the scalar product of two vectors $\vec{a} = (x_a, y_a)$ and $\vec{b} = (x_b, y_b)$ (where all components are real) is defined as $\vec{a} \cdot \vec{b} = x_a x_b + y_a y_b$. Show that this definition has all the properties listed above.

Note 1.4 According to the definition of the inner product, generally $\langle a|\ b\rangle \neq \langle b|\ a\rangle$

Exercise 1.14 For $|x\rangle = \sum_i \lambda_i\,|a_i\rangle$, show that $\langle b|\ x\rangle = \sum_i \lambda_i\,\langle b|\ a_i\rangle$ and $\langle x|\ b\rangle = \sum_i \lambda_i^*\,\langle a_i|\ b\rangle$.

Exercise 1.15 Show that $\forall |a\rangle\ \ \langle \text{zero}|\ a\rangle = \langle a|\ \text{zero}\rangle = 0$.

Definition 1.9 $|a\rangle$ and $|b\rangle$ are called *orthogonal* if $\langle a|\ b\rangle = 0$

Exercise 1.16 A set of mutually orthogonal vectors is linearly independent.

Definition 1.10 $\| |a\rangle \| = \sqrt{\langle a|\ a\rangle}$ is called the *norm (length)* of a vector. Vectors of norm 1 are called *normalized*.

Exercise 1.17 Show that multiplying a state vector by a *phase factor* $e^{i\phi}$, where ϕ is a real number, does not change its norm.

Definition 1.11 A linear space, in which an inner product is defined, is called the *Hilbert space*.

1.5 First Quantum Mechanics Postulate

Quantum Mechanics Postulate I[6]
We will first give s succinct formulation of the Postulate, and then explain its meaning in more detail.

a) A state of a physical system is represented by a vector $|\psi(t)\rangle$ in a Hilbert space of possible states.

b) Incompatible quantum states correspond to orthogonal vectors.

c) All vectors that represent physical quantum states are normalized.

The notions of quantum state and physical system are not clearly defined, but can be understood intuitively. A *physical system* is one or several degrees of freedom associated with a particular physical object that can be studied independently of other degrees of freedom and other objects. For example, if our object is an atom, quantum mechanics can study its motion (one physical system) or its internal state (another physical system). If we wish to study the formation of a molecule out of two atoms, internal and motional states of both atoms affect each other, so we must consider all these degrees of freedom as one physical system. For a molecule itself, quantum mechanics can study its center of mass motion (one physical system), rotational motion (another physical system), vibration of its atoms (third system), quantum states of its electrons (fourth system), etc. On the other hand, any individual atom in a molecule would typically not form a physical system because it cannot be investigated separately from other atoms.

[6] There are no universally accepted postulates of quantum mechanics. Their formulation, as well as numbering, somewhat vary from textbook to textbook.

The notion of a *state* can be defined epistemologically as *what we know* about a physical system, i.e. the information sufficient to predict the future behavior of the system with maximum possible precision. Even if we know the state, a prediction can sometimes only be made in a statistical sense: we can say that with some probability the system will do this, and with a different probability something else. However (if we trust quantum mechanics), it is *fundamentally impossible* to make a better prediction; we can't know more about the future than what we know from its quantum state.

Consider, for example, the following physical system: one motional degree of freedom of a massive particle. One can define its quantum state by saying "the particle's coordinate is exactly $x = 5$ meters". This is a valid definition; we would denote this state as $|x = 5m\rangle$. Another valid state would be $|x = 3m\rangle$. These states are orthogonal ($\langle x = 5m|\ x = 3m\rangle = 0$) because they are "incompatible": if a particle's coordinate is definitely known to be 5 meters, it cannot be detected at the position 3 meters. On the other hand, the particle can be in the state "moving at a speed $v = 4$ meters per second". This is also a valid quantum state; however, we cannot *a priori* say that the particle in this state cannot be detected at $x = 5m$. Hence the inner product $\langle x = 5m|\ v = 4m/s\rangle$ does not have to vanish, and it indeed does not, as we shall see in Chapter 3.

What the first postulate of quantum mechanics says is that if $|x = 5m\rangle$ and $|x = 3m\rangle$ are valid quantum states, then $[|x = 5m\rangle + |x = 3m\rangle]/\sqrt{2}$ [where $1/\sqrt{2}$ is the normalization factor — see Ex. 1.18 for the explanation] is also a valid state. This is called a *superposition* state of the particle being at these two locations. Existence of such states does not make much sense at first: how can a particle be at two places at the same time?

Erwin Schrödinger, one of the founding fathers of quantum physics, gave an even stranger example, talking about a cat in a superposition of being dead and alive. This is the first of the many quantum mysteries and paradoxes we will be encountering in this course. Indeed, the Schrödinger cat is fully compatible with the canons of quantum mechanics, i.e. it is a valid quantum state. However, as we shall see later, this state is extremely fragile and quickly transforms either into the dead state or into the alive state. This is why we don't see too many Schrödinger cats walking around.

Exercise 1.18 What is the normalization factor N of the state of the Schrödinger cat $|\psi\rangle = N\left[2\,|\text{alive}\rangle + i\,|\text{dead}\rangle\right]$ that ensures that $|\psi\rangle$ is a physical state?

Exercise 1.19 What is the dimension of the Hilbert space associated with one motional degree of freedom of a massive particle?
Hint: If you think the answer is obvious, check the solution.

1.6 Polarization of the photon

In order to predict the behavior of a quantum system, we need to know precisely the physical properties of all states in at least one basis of the relevant Hilbert space. Because this requirement is usually difficult to satisfy, quantum mechanics prefers to deal with rather simple systems with few degrees of freedom. We will begin studying quantum mechanics with one of the most simple physical systems: the polarization of the photon[7]. The dimension of its Hilbert space is just two, yet it is quite sufficient to show how amazing the world of quantum mechanics can be.

Suppose we can isolate a single particle of light, photon, from a polarized wave. The photon is a microscopic object and must be treated quantum-mechanically. We begin this treatment by defining the associated Hilbert space. We first notice that the state of the photon obtained from a horizontally polarized wave, whose state we denote as $|H\rangle$, is incompatible with its vertical counterpart, $|V\rangle$: an $|H\rangle$ photon can never be detected in a $|V\rangle$ state. If we prepare a horizontally polarized photon and send it through a polarizing beam splitter, it will always be transmitted and never reflected. This means that states $|H\rangle$ and $|V\rangle$ are orthogonal.

Now we introduce the following rule for other polarization states of photons: any complex linear combination of states $|H\rangle$ and $|V\rangle$

$$|\psi\rangle = A_H e^{i\varphi_H}\,|H\rangle + A_V e^{i\varphi_V}\,|V\rangle\,, \tag{1.3}$$

[7] This is a good place to read the first two sections of Appendix A.

defines the polarization state of the photons that compose the classical wave

$$\vec{E} = \mathrm{Re}[(A_H e^{i\varphi_H}\vec{e}_H + A_V e^{i\varphi_V}\vec{e}_V)e^{ikz-i\omega t}], \tag{1.4}$$

(cf. Eq. (A.1)) with the same amplitudes and phases. For example, if $A_H = A_V$ and $\varphi_H = \varphi_V = 0$, the associated classical wave is $\vec{E} = \mathrm{Re}[A_H(\vec{e}_H + \vec{e}_V)e^{ikz-i\omega t}]$, i.e. linearly polarized at $+45°$. Accordingly, the state $(|H\rangle + |V\rangle)/\sqrt{2}$, where the factor of $\sqrt{2}$ is due to normalization, denotes a single photon with $+45°$ linear polarization. Some further examples are listed in Table 1.1. Note that while the above rule appears intuitive, there is some complex and deep physics behind it, which is beyond the present course.

It follows that states $|H\rangle$ and $|V\rangle$ form an orthonormal basis in the corresponding Hilbert space. First, these states are orthogonal and thus linearly independent (Ex. 1.16). Second, any polarized classical wave can be written in the form (1.4), and thus any polarization state of the photon can be written in accordance with (1.3), i.e. as a linear combination of states $|H\rangle$ and $|V\rangle$. We will call the basis $\{|H\rangle, |V\rangle\}$ the *canonical* basis of our Hilbert space.

We see that the Hilbert space of photon polarization states is two-dimensional. This may be confusing. For linearly polarized photons, there is a continuum of polarization angles — similarly to the continuum of position states in the case of one-dimensional particle motion, discussed in the previous section. So why is one Hilbert space of dimension two and the other of dimension infinity? The reason is that the superposition of linear polarization states with different angles is still a polarization state, e.g.

$$\frac{|0°\rangle + |90°\rangle}{\sqrt{2}} = \frac{|H\rangle + |V\rangle}{\sqrt{2}} = |+45°\rangle\,.$$

On the other hand, a superposition of two position states is not a position state:

$$\frac{|x = 3m\rangle + |x = 5m\rangle}{\sqrt{2}} \neq |x = 4m\rangle\,.$$

Therefore, the latter Hilbert space is much more complex than the former.

For a classical wave, shifting the phases of both horizontal and vertical component by the same amount (i.e. $\varphi_H \to \varphi_H + \varphi_0$, $\varphi_V \to \varphi_V + \varphi_0$, which is equivalent to multiplying the right-hand side by $e^{i\phi}$) does not change the polarization of the wave.

A similar rule applies to quantum states. Multiplying a state vector by $e^{i\phi}$ does not change the physical nature of a state. For example, $|V\rangle$, $i\,|V\rangle$ and $-\,|V\rangle$ represent the same physical object, as well as, say, $(|H\rangle + i\,|V\rangle)/\sqrt{2}$ and $(|V\rangle - i\,|H\rangle)/\sqrt{2}$. This rule turns out to be very general: it works for all states in the entire domain of quantum mechanics.

Table 1.1: Important polarization states.

state	designation	notation
$\cos\theta\,\|H\rangle + \sin\theta\,\|V\rangle$	linear polarization at angle θ to horizontal	$\|\theta\rangle$
$(\|H\rangle + \|V\rangle)/\sqrt{2}$	$+45°$ polarization	$\|+45°\rangle$ or $\|+\rangle$
$(\|H\rangle - \|V\rangle)/\sqrt{2}$	$-45°$ polarization	$\|-45°\rangle$ or $\|-\rangle$
$(\|H\rangle + i\,\|V\rangle)/\sqrt{2}$	Right circular polarization	$\|R\rangle$
$(\|H\rangle - i\,\|V\rangle)/\sqrt{2}$	Left circular polarization	$\|L\rangle$

Note 1.5 The $\pm 45°$ polarization states are also called *diagonal* polarization states.

Note 1.6 There is no common convention associating the handedness of the circular polarization state with the positive or negative sign in the expression $(|H\rangle \pm i\,|V\rangle)/\sqrt{2}$. In this course, we will be using the convention defined in Table 1.1.

Exercise 1.20 Write the matrix form of the decomposition of the diagonal and circular polarization states in the canonical basis.
Answer:

$$|H\rangle \leftrightarrow \begin{pmatrix} 1 \\ 0 \end{pmatrix}; \qquad |V\rangle \leftrightarrow \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{1.5}$$

$$|+\rangle \leftrightarrow \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}; \qquad |-\rangle \leftrightarrow \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \tag{1.6}$$

$$|R\rangle \leftrightarrow \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \end{pmatrix}; \qquad |L\rangle \leftrightarrow \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -i \end{pmatrix} \tag{1.7}$$

1.7 Orthonormal Basis

Definition 1.12 An *orthonormal basis* $\{|v_i\rangle\}$ is a basis whose elements are mutually orthogonal and have norm 1, i.e.

$$\langle v_i|\, v_j\rangle = \delta_{ij}, \tag{1.8}$$

where δ_{ij} is the Kronecker symbol.

Exercise 1.21 Any orthonormal set of N (where $N = \dim \mathbb{V}$) vectors forms a basis. **Hint:** use the result of Ex. 1.7.

Exercise 1.22 Show that, if $\begin{pmatrix} a_1 \\ \vdots \\ a_N \end{pmatrix}$ and $\begin{pmatrix} b_1 \\ \vdots \\ b_N \end{pmatrix}$ are the decompositions of states $|a\rangle$ and $|b\rangle$ in an orthonormal basis, their inner product can be written as

$$\langle a|\, b\rangle = a_1^* b_1 + \ldots + a_N^* b_N. \tag{1.9}$$

Note 1.7 Eq. (1.9) can be expressed in the matrix form using the "row-times-column" rule:

$$\langle a|\, b\rangle = \begin{pmatrix} a_1^* & \ldots & a_N^* \end{pmatrix} \begin{pmatrix} b_1 \\ \vdots \\ b_N \end{pmatrix} \tag{1.10}$$

Note 1.8 Eqs. (1.9) and (1.10) are most frequently used for calculating the inner product. The decompositions of vectors may be different in different bases, and it may appear that the inner product also depends on the basis chosen. This is not the case: according to Defn. 1.8 the inner product depends on the physical states only. It is basis independent.

Note 1.9 $|v_1\rangle = |H\rangle$ and $|v_2\rangle = |V\rangle$ form an orthonormal basis in the Hilbert space of photon polarization states. We will call this basis the *canonical* basis, and decompositions in this basis the *canonical decompositions*.

Exercise 1.23 $\pm 45°$ polarization states form an orthonormal basis. Right and left circular polarization states form an orthonormal basis.

Exercise 1.24 Show that the amplitudes of the decomposition $\begin{pmatrix} a_1 \\ \vdots \\ a_N \end{pmatrix}$ of a vector $|a\rangle$ into an orthonormal basis can be found as follows:

$$a_i = \langle v_i|\, a\rangle\,. \tag{1.11}$$

In other words [see Eq. (1.1)],

$$|a\rangle = \sum_i \langle v_i|\, a\rangle\, |v_i\rangle\,. \tag{1.12}$$

Exercise 1.25 Decompose $|H\rangle$ and $|V\rangle$ into the $\{|+\rangle, |-\rangle\}$ and the $\{|R\rangle, |L\rangle\}$ bases.

Exercise 1.26 If $|a\rangle$ is a normalized state and $\{a_i = \langle v_i|\ a\rangle\}$ is its decomposition in an orthonormal basis $\{|v_i\rangle\}$, then

$$\sum_i |a_i|^2 = 1 \tag{1.13}$$

Exercise 1.27 Decompose $|a\rangle = |+30^\circ\rangle$ and $|b\rangle = |-30^\circ\rangle$ in the $\{|H\rangle, |V\rangle\}$, $\{|+\rangle, |-\rangle\}$, and $\{|R\rangle, |L\rangle\}$ bases. Find the inner product $\langle a|\ b\rangle$ using the result of Ex. 1.22 in all three bases. Do they come out the same?

Exercise 1.28 Suppose $\{|w_i\rangle\}$ is some basis in $\mathbb{V}$. It can be used to find an orthonormal basis $\{|v_i\rangle\}$ by applying the following equation in sequence to each basis element:

$$|v_{k+1}\rangle = \mathcal{N}\left[|w_{k+1}\rangle - \sum_{i=1}^{k} \langle v_i|\ w_{k+1}\rangle\, |v_i\rangle\right], \tag{1.14}$$

where $\mathcal{N}$ is the normalization factor. This is called the *Gram-Schmidt procedure*.

Exercise 1.29 Prove the Cauchy-Schwarz inequality:

$$\forall\, |a\rangle, |b\rangle \qquad |\langle a|\ b\rangle| \leq \| |a\rangle \| \times \| |b\rangle \|. \tag{1.15}$$

Show that the equality is reached if and only if the states $|a\rangle$ and $|b\rangle$ are collinear (i.e. $|a\rangle = \lambda\, |b\rangle$).

Exercise 1.30 Prove the triangle inequality:

$$\forall\, |a\rangle, |b\rangle \qquad \|(|a\rangle + |b\rangle)\| \leq \| |a\rangle \| + \| |b\rangle \| \tag{1.16}$$

1.8 Second Quantum Mechanics Postulate

The second postulate deals with *quantum measurements*, i.e. experiments on obtaining information on the quantum state of the system. In classical, macroscopic physics, the concept of measurement is of technical rather than fundamental nature. This is because we can precisely measure the state and the evolution of the system without disturbing it. For example, a soccer ball will not fly differently dependent on whether the stadium is empty or full of cheering spectators.

In the quantum world, the situation is different: we are big and the things we want to measure are small. Therefore, any measurement will most likely change the quantum state of our system. More importantly, many measurement-like phenomena, in which a quantum state of something microscopic particle affects something macroscopic, occur *inadvertently*, without the experimentalist intention. This can be related, albeit only superficially, to the "butterfly effect"[8], in which a small change in a complex system can result in major consequences some time later. In our world, a huge variety of phenomena, ranging from thermodynamic phase transitions to birth of black holes, and perhaps the existence of the universe itself, are results of quantum fluctuations, and hence can be considered as examples of generalized quantum measurements. But even if such a generalized measurement does not bring about anything dramatic, it will still affect the evolution of a quantum system, and thus needs to be studied.

As said earlier, the laws governing the quantum and classical domains of physics are very different. In order to have a unified picture of the world, we need to have an *interface* between the two, i.e. an understanding when and how the transition between these two "jurisdictions" occurs. A primary element of this interface is provided by the quantum measurement theory.

[8]The term "butterfly effect" originates from a 1972 talk by Edward Lorenz titled "Does the flap of a butterflys wings in Brazil set off a tornado in Texas?".

Before we proceed to formulating the postulate, let us consider an example. Suppose a single photon in state $\alpha|H\rangle + \beta|V\rangle$ hits a polarizing beam splitter (PBS) [Fig. 1.1(a)]. If we were dealing with a classical wave, we would expected it to split: a part would be transmitted through the PBS, and the remainder reflected. But the photon is the smallest energy portion of light, and cannot be divided into parts. So what will happen to it? The experiment shows that the outcome will be random: the photon will go through the PBS with a probability $|\alpha|^2$, and be reflected with a probability $|\beta|^2$.

If a large number N of photons are incident on the PBS (e.g. in the case of a classical wave), on average $|\alpha|^2 N$ of them will be transmitted, and $|\beta|^2 N$ reflected. This means that the total flux of energy in the transmitted and reflected channels will be proportional to $|\alpha|^2$ and $|\beta|^2$, respectively. This is remarkably consistent with the classically expected Eqs. (A.3).

As we know, the part of the classical wave that is transmitted through the PBS will become horizontally polarized. The same happens with photons. After the PBS, the photon state in the transmitted channel will become $|H\rangle$ (and in the reflected channel $|V\rangle$). If we place a series of additional PBS's in the transmitted channel of the first one, the photon will be transmitted through all of these PBS's.

The photon propagating through a PBS gives us an example of the photon polarization state measurement. Into both output channels of the PBS, we can place single-photon detectors — devices that generate a macroscopic electric pulse (*click* in quantum jargon) whenever a photon hits their sensitive areas. Of the two detectors, only one will click — thus providing us with some information about the photon's initial polarization.

We are now ready to formulate our postulate.

Second Quantum Mechanics Postulate An idealized measurement apparatus is associated with some orthonormal basis $\{|v_i\rangle\}$. After the measurement, the apparatus will randomly, with probability

$$\mathrm{pr}_i = |\langle\psi|\, v_i\rangle|^2, \tag{1.17}$$

where $|\psi\rangle$ is the initial state of the system, point to one of the states $|v_i\rangle$. The system (if not destroyed) will then be converted (*projected*, as quantum physicists say) onto the state $|v_i\rangle$.

Definition 1.13 A quantum measurement that proceeds in accordance with the above postulate is called *projective measurement*. The projection of the state measured onto one of the basis elements is also called *collapse* of the quantum state.

Because the initial state $|\psi\rangle$ is normalized, the sum of the probabilities of each measurement outcome is, according to Eq. (1.13), $\sum_i \mathrm{pr}_i = \sum_i |\langle\psi|\, v_i\rangle|^2 = 1$, as one would naturally expect. We also see that the probability of the measurement result does not depend on the overall phase of state $|\psi\rangle$ — in agreement with the fact that this phase has no influence on the physics of a state, as discussed in Sec. 1.6.

Probabilistic behavior of quantum objects has caused a lot of contradiction at the time quantum mechanics was founded. This is because, by the end of the 19th century, the principle of *determinism* was universally accepted: physicists believed that if the initial conditions of a given quantum system is known precisely enough, its future evolution can be predicted arbitrarily well. Quantum physics breached this fundamental belief, and many physicists found it extremely difficult to accept. For example, Albert Einstein made a famous statement that "God does not play dice" and came up with an brilliant *Gedankenexperiment*[9] showing that the postulates of quantum mechanics are in contradiction with the common sense. We will study this Gedankenexperiment in the next two chapters and will see that although quantum mechanics indeed seems to contradict the common sense, the randomness is an intrinsic, experimentally verifiable feature of the world. Einstein was wrong. God does play dice.

Above, we discussed the apparatus for measuring the polarization of the photon in the canonical basis. What if we want to measure it in some other basis? We can take advantage of the optical

[9] "Gedankenexperiment" is the German for "Thought experiment".

element called the waveplate[10] which interconverts polarization states of a photon into one another. Here are some examples.

- The setup shown in Fig. 1.1(a) measures the photon polarization in the canonical ($|H\rangle, |V\rangle$) basis: A polarizing beam splitter sends the horizontal and vertical polarization components to different single-photon detectors. A "click" in one of the detector signifies that a measurement has occurred. This measurement is destructive, because the photon is absorbed by the detector photocathode.
- The setup in Fig. 1.1(b) performs the measurement in the diagonal ($|\pm 45°\rangle$) basis: A $\lambda/2$ waveplate at 22.5° first converts the $+45°$ and $-45°$ components into horizontal and vertical and then a polarizing beam splitter sends these to separate detectors .
- Measurement in the circular polarization ($|R\rangle, |L\rangle$) basis [Fig. 1.1(c)]: a $\lambda/4$ at 0° first converts the circular components into $\pm 45°$ components, then a $\frac{\lambda}{2}$ waveplate, again at 22.5°, converts them into horizontal and vertical components which are then split by a polarizing beam splitter.

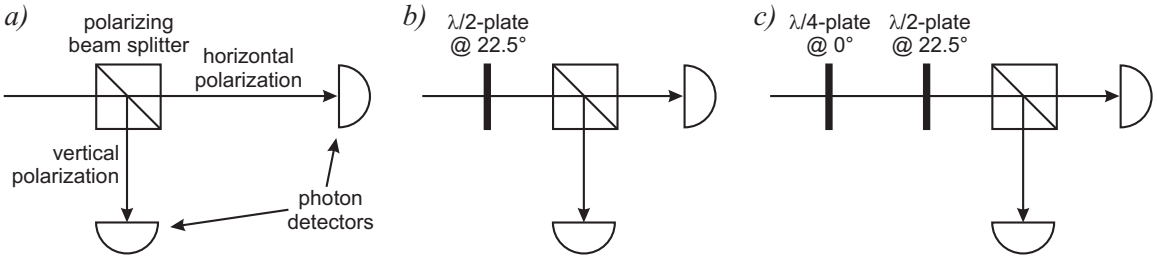


Figure 1.1: Photon polarization measurements in the $\{|H\rangle, |V\rangle\}$ (a), $\{|+\rangle, |-\rangle\}$ (b) and $\{|R\rangle, |L\rangle\}$ (c) bases.

Exercise 1.31 Invent a scheme for the $\{|R\rangle, |L\rangle\}$ basis that would use just one waveplate.

Exercise 1.32 A photon is prepared with a linear polarization 30° to horizontal. Find the probabilities of each outcome if its polarization is measured in the (a) $\{|H\rangle, |V\rangle\}$, (b) $\{|+\rangle, |-\rangle\}$, and (c) $\{|R\rangle, |L\rangle\}$ basis.

Although a single measurement provides us with some information about the initial state of a quantum system, this information is very limited. For example, suppose we have measured a photon in the canonical basis and found that the photon has been transmitted through the PBS. The only thing we learn from this measurement is that the photon was not vertically polarized. But for any other initial state, the result obtained is fully possible.

Suppose now we have performed the same measurement many times. Now we know much more! Since we have the statistics, we can calculate, with some error, $\mathrm{pr}_H = |\langle H|\,\psi\rangle|^2$ and $\mathrm{pr}_V = |\langle V|\,\psi\rangle|^2$, i.e. learn about the absolute values of the state components. But the complex phases of these components are still unknown. For example, if we observe $\mathrm{pr}_H = \mathrm{pr}_V = 1/2$, state $|\psi\rangle$ could be $|H\rangle$ or $|V\rangle$ or $|+\rangle$ or $|-\rangle$ or many other options. What can we do about this?

As we propose the reader to find out independently in the following Exercise, it helps to perform additional sets of measurements in other bases. Then we obtain additional numbers, and it is easier for us to solve the equations for α and β. As it turns out, this approach to measuring quantum states can be generalized to other quantum systems, including those of higher dimension.

Definition 1.14 The procedure of obtaining complete information about the quantum state by performing series of measurements in several different bases on the state's multiple identical copies is called *quantum tomography*.

[10] This is a good place to read the third section of Appendix A.

Exercise 1.33 Suppose multiple polarization measurements of photons identically prepared in the state $|\psi\rangle$ are done in the $\{|H\rangle, |V\rangle\}$, $\{|+\rangle, |-\rangle\}$, $\{|R\rangle, |L\rangle\}$ bases and all six respective probabilities are determined. Show that this information is sufficient to fully determine $|\psi\rangle$ and express its decomposition in the $\{|H\rangle, |V\rangle\}$ basis through p_H, p_+, and p_R. Give an example showing that measuring just in the canonical and diagonal bases is insufficient.

The actual apparatus may be more complicated than a simple projective measurement described by the Second Postulate. For example, a realistic photon detector may fail to register a photon incident on its sensitive area, or produce a "dark count" event in the absence of a photon. A realistic PBS has a finite probability to transmit a vertically polarized photon and reflect a horizontally polarized one. Finally, the photon may simply get lost on its way along the optical path.

However, every measurement apparatus can be presented as some combination of projective measurements and classical, possibly probabilistic, processing of the data. For example, a measurement which contains an imperfect PBS can be viewed as a perfect projective measurement followed by a classical device that is programmed to scramble the measurement result in some fraction of events. An apparatus containing imperfect detectors may be considered to contain a processor that randomly fails to display the result. Such complex measurement apparata have their own mathematical description, which is, however, beyond the scope of this course.[11]

Exercise 1.34 Consider two non-orthogonal states $|a\rangle$ and $|b\rangle$.

a) Show that it is not possible to construct a measurement apparatus that would distinguish these states with certainty.

b)* Show that it is possible to construct a measurement device that would produce, with some probability, results of three types: "definitely not $|a\rangle$", "definitely not $|b\rangle$", and "not sure", and the outputs of the first two types would always be correct.

1.9 Adjoint Space

It is sometimes convenient to think of a scalar product $\langle a|\, b\rangle$ as a product of two objects, $\langle a|$ and $|b\rangle$. This convention is mainly of notational nature, but is used very frequently and is therefore important.

Definition 1.15 For the Hilbert space $\mathbb{V}$, we define the *adjoint space* $\mathbb{V}^\dagger$ (read "V-dagger") which has one-to-one correspondence to $\mathbb{V}$: for each vector $|a\rangle \in \mathbb{V}$ there is one and only one *adjoint* vector $\langle a| \in \mathbb{V}^\dagger$ with the property

$$\text{Adjoint}(\lambda\,|a\rangle + \mu\,|b\rangle) = \lambda^*\,\langle a| + \mu^*\,\langle b|\,. \tag{1.18}$$

"Direct" and adjoint vectors are sometimes called *ket-* and *bra*-vectors, respectively. The rationale behind this terminology, introduced by P. Dirac together with the symbols $\langle|$ and $|\rangle$, is that the bra-ket combination, a "bracket", gives the inner product of the two vectors, i.e. a probability amplitude for a quantum measurement.

Although the adjoint space is a linear space (see an Exercise below), $\mathbb{V}$ and $\mathbb{V}^\dagger$ are different linear spaces. We cannot add a bra-vector and a ket-vector. A good everyday analogy of the adjoint space is the image in a mirror: while there is a one-to-one correspondence between objects and images, and the image perfectly replicates all properties of an object that generated it, it cannot interact with that object.

Exercise 1.35 Show that $\mathbb{V}^\dagger$ is a linear space.

Exercise 1.36 Show that if $\{|v_i\rangle\}$ is a basis in $\mathbb{V}$, $\{\langle v_i|\}$ is a basis in $\mathbb{V}^\dagger$ and if a ket-vector is decomposed in this basis as $|a\rangle = \sum a_i\,|v_i\rangle$, the decomposition of its adjoint is

$$\langle a| = \sum a_i^*\,\langle v_i| \tag{1.19}$$

[11] This is a good place to read Appendix B.

If we want to write a bra-vector in a matrix form, it is convenient to write it as a row rather than column. For example, Eq. (1.19) can be written as

$$\langle a| \leftrightarrow \begin{pmatrix} a_1 \\ \vdots \\ a_N \end{pmatrix}^\dagger = \begin{pmatrix} a_1^* & \dots & a_N^* \end{pmatrix}, \tag{1.20}$$

where the superscript † (in application to a matrix) means transposition and complex conjugation. The inner product (1.10) is then just a matrix product of the representations of the bra- and ket-vectors. Accordingly, we can think of an inner productof two ket-vectors $|a\rangle$ and $|b\rangle$ as a product of a bra-vector and a ket-vector: $(\langle a|)(|b\rangle) \equiv \langle a|\, b\rangle$.

As long as we work in the framework of one particular basis, we can treat bras and kets as rows and columns. Do not forget, though, that all matrix representations change when we choose another basis.

Exercise 1.37 Find matrix forms of states $\langle L|$ and $\langle R|$ in the adjoint canonical basis.

1.10 Linear Operators

Definition 1.16 A *linear operator* $\hat{A}$ is a map[12] of one linear space $\mathbb{V}$ onto another linear space $\mathbb{W}$ such that

a) $\hat{A}(|a\rangle + |b\rangle) = \hat{A}\,|a\rangle + \hat{A}\,|b\rangle$

b) $\hat{A}(\lambda\,|a\rangle) = \lambda\hat{A}\,|a\rangle$

Note 1.10 In this course, we consider only operators that map vector spaces onto themselves ($\mathbb{V} \to \mathbb{V}$).

Definition 1.17 The operator $\hat{\mathbf{1}} : \mathbb{V} \to \mathbb{V}$ which maps every vector onto itself is called the *identity operator*.

Exercise 1.38 Are the following maps linear operators[13]:

a) $\hat{A}\,|a\rangle \equiv 0$

b) $\hat{\mathbf{1}}$

c) $\mathbb{C}^2 \to \mathbb{C}^2 : \hat{A}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ -y \end{pmatrix}$

d) $\mathbb{C}^2 \to \mathbb{C}^2 : \hat{A}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x+y \\ xy \end{pmatrix}$

e) $\mathbb{C}^2 \to \mathbb{C}^2 : \hat{A}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x+1 \\ y+1 \end{pmatrix}$

f) Rotation by angle ϕ in the linear space of two-dimensional geometric vectors (over $\mathbb{R}$)?

Definition 1.18 For any operator $\hat{A}$ and any scalar λ, we can define their product, an operator $\lambda\hat{A}$ which maps vectors according to

$$(\lambda\hat{A})\,|a\rangle \equiv \lambda(\hat{A}\,|a\rangle). \tag{1.21}$$

For any two operators $\hat{A}$ and $\hat{B}$, we can define their sum, an operator $\hat{A} + \hat{B}$ which maps vectors according to

$$(\hat{A} + \hat{B})\,|a\rangle \equiv \hat{A}\,|a\rangle + \hat{B}\,|a\rangle\,. \tag{1.22}$$

[12] A *map* $f : \mathbb{A} \to \mathbb{B}$ from set $\mathbb{A}$ to set $\mathbb{B}$ is a function f such that for every element a in $\mathbb{A}$, there is a unique "image" $f(a)$ in $\mathbb{B}$.

[13] $\mathbb{C}^2$ is the linear space of two-number columns $\begin{pmatrix} x \\ y \end{pmatrix}$.

Note 1.11 When writing products of a scalar with the identity operators, we sometimes omit the symbol $\hat{\mathbf{1}}$ when the context allows no ambiguity. For example, instead of writing $\hat{A} - \lambda\hat{\mathbf{1}}$, we may simply write $\hat{A} - \lambda$.

Exercise 1.39 Show that linear operators form a linear space with the operations of addition and multiplication by a scalar defined as above.

Definition 1.19 *Operator product* $\hat{A}\hat{B}$ is an operator that maps every vector $|a\rangle$ onto $\hat{A}\hat{B}\,|a\rangle \equiv \hat{A}(\hat{B}\,|a\rangle)$. That is, in order to find the action of operator $\hat{A}\hat{B}$ onto a vector, we must first apply $\hat{B}$ to that vector, and then apply $\hat{A}$ to the result.

Note 1.12 It matters in which order the two operators are multiplied, i.e., generally $\hat{A}\hat{B} \neq \hat{B}\hat{A}$. Such operators for which $\hat{A}\hat{B} = \hat{B}\hat{A}$ are said to *commute*. Commutation relations between operators play an important role in quantum mechanics, and will be discussed in detail in Sec. 1.13.

Exercise 1.40 Verify that the operators of counterclockwise rotation by angle $\pi/2$ and reflection about the horizontal axis in the linear space of two-dimensional geometric vectors do not commute.

Exercise 1.41 Show that the multiplication of operators has the property of associativity, i.e. for any three operators one has

$$\hat{A}(\hat{B}\hat{C}) = (\hat{A}\hat{B})\hat{C}. \tag{1.23}$$

It may appear that in order to fully characterize a linear operator, we must tell what it does to every vector. However, this is not the case. In fact, it is enough to tell how the operator maps the elements of some basis $\{|v_1\rangle, \ldots, |v_N\rangle\}$ in $\mathbb{V}$, i.e. it is enough to know the set $\{\hat{A}\,|v_1\rangle, \ldots, \hat{A}\,|v_N\rangle\}$. Then, for any other vector $|a\rangle$, which can be decomposed as

$$|a\rangle = a_1\,|v_1\rangle + \ldots + a_N\,|v_N\rangle\,,$$

we have, thanks to linearity,

$$\hat{A}\,|a\rangle = a_1\hat{A}\,|v_1\rangle + \ldots + a_N\hat{A}\,|v_N\rangle\,. \tag{1.24}$$

How many numerical parameters does one need to completely describe a linear operator? Each image $\hat{A}\,|v_j\rangle$ of a basis element can be decomposed into the same basis:

$$\hat{A}\,|v_j\rangle = \sum_i A_{ij}\,|v_j\rangle\,, \tag{1.25}$$

where, in accordance with Eq. (1.11)

$$A_{ij} = \langle v_i|\left(\hat{A}\,|v_j\rangle\right) \equiv \langle v_i|\,\hat{A}\,|v_j\rangle \tag{1.26}$$

(expressions of the form such as in the right-hand side of Eq. (1.26) are informally called *sandwiches*). For every j, the set of N parameters $A_{1j}, \ldots, A_{Nj}$ fully describes $\hat{A}\,|v_j\rangle$. Accordingly, the set of N^2 parameters A_{ij}, with both i and j varying from 1 to N, contains full information about a linear operator.

Definition 1.20 The *matrix of an operator* in basis $\{|v_i\rangle\}$ is an $N \times N$ square table whose elements are given by Eq. (1.26). The first index of A_{ij} is the number of the row, the second is the number of the column.

Exercise 1.42 Find the matrix of $\hat{\mathbf{1}}$. Show that this matrix does not depend on the choice of basis.

Exercise 1.43 Show that if, in some basis, $|a\rangle \leftrightarrow \begin{pmatrix} a_1 \\ \vdots \\ a_N \end{pmatrix}$, then the state $\hat{A}\,|a\rangle$ is given by the matrix product

$$\hat{A}\,|a\rangle \leftrightarrow \begin{pmatrix} A_{11} & \ldots & A_{1N} \\ \vdots & & \vdots \\ A_{N1} & \ldots & A_{NN} \end{pmatrix} \begin{pmatrix} a_1 \\ \vdots \\ a_N \end{pmatrix} \tag{1.27}$$

In other words, the action of an operator on a vector is equivalent to the multiplication of the corresponding matrices. The matrix of an operator in a known basis fully defines the operator, because using the matrix we can find out what the operator does to every vector (see Ex. 1.43). Operations with operators and vectors are identical to operations with matrices and columns.

Suppose, for example, that you are required to prove that some two operators are equal: $\hat{A} = \hat{B}$. You can do it by choosing some basis and showing the identity for the matrices A_{ij} and B_{ij} of the operators in this basis. Because the matrix fully defines an operator, this would be sufficient. Of course, you will probably want to choose your basis so the matrices A_{ij} and B_{ij} are as easy as possible to calculate.

Caveat: matrices of vectors and operators depend on the basis chosen. On the other hand, operators and states are physical objects and are basis independent. If your calculation cannot be done in one fixed basis, it may be useful to keep it in the Dirac notation to avoid confusion.

Exercise 1.44 Given the matrices of operators $\hat{A}$ and $\hat{B}$ show that

a) $\lambda\hat{A}$

b) $\hat{A} + \hat{B}$

c) $\hat{A}\hat{B}$

are linear operators and find their matrices.

Exercise 1.45 Find the matrices of operators $\hat{A}$ and $\hat{B}$ that correspond to the rotation of the two-dimensional geometric space by angles ϕ and θ, respectively. Find the matrix of $\hat{A}\hat{B}$ using the result of Ex. 1.44 and verify its equivalence to a rotation by $(\phi + \theta)$.

Exercise 1.46 Find the matrix representation of vector $\hat{A}\,|v_k\rangle$, where $|v_k\rangle$ is an element of the basis. k is given, the matrix of $\hat{A}$ is known.

Exercise 1.47 Find, in the canonical basis, the matrix of the linear operator $\hat{A}$ that maps

a) $|H\rangle$ onto $|R\rangle$ and $|V\rangle$ onto $2\,|H\rangle$;

b) $|+\rangle$ onto $|R\rangle$ and $|-\rangle$ onto $|H\rangle$. **Hint:** you first need to find the states $\hat{A}\,|v_1\rangle = \hat{A}\,|H\rangle$ and $\hat{A}\,|v_2\rangle = \hat{A}\,|V\rangle$ and then use Eq. (1.26).

Exercise 1.48 Find the matrix of the operator associated with

a) a $\lambda/2$ plate with its optical axis oriented vertically,

b) a $\lambda/2$ plate with its optical axis oriented at angle 45° to horizontal,

c) a $\lambda/4$ plate with its optical axis oriented vertically,

d)* a $\lambda/4$ plate with its optical axis oriented at angle 30° to horizontal.

Definition 1.21 *Outer products* $|a\rangle\langle b|$ are understood as operators acting as follows:

$$(|a\rangle\langle b|)\,|c\rangle \equiv |a\rangle\,(\langle b|\,c\rangle) = (\langle b|\,c\rangle)\,|a\rangle\,. \tag{1.28}$$

(The second equality comes from the fact that $\langle b|\,c\rangle$ is a number and commutes with everything.)

Exercise 1.49 Show that $|a\rangle\langle b|$ as defined above is a linear operator.

Exercise 1.50 Show that the matrix of the operator $|a\rangle\langle b|$ is a product of the matrix representations of $|a\rangle$ and $\langle b|$:

$$|a\rangle\langle b| \leftrightarrow \begin{pmatrix} a_1 \\ \vdots \\ a_N \end{pmatrix} \begin{pmatrix} b_1^* & \dots & b_N^* \end{pmatrix} = \begin{pmatrix} a_1 b_1^* & \dots & a_1 b_N^* \\ \vdots & & \vdots \\ a_N b_1^* & \dots & a_N b_N^* \end{pmatrix}. \tag{1.29}$$

Exercise 1.51 Find the matrix of the operator $|+\rangle\langle -|$ in the canonical and the $(|R\rangle, |L\rangle)$ bases.

Exercise 1.52 In an orthonormal basis $\{|v_i\rangle\}$,

$$\hat{A} = \sum_{i,j} A_{ij} |v_i\rangle \langle v_j| \tag{1.30}$$

Note 1.13 Eqs. (1.26) and (1.30) are used to switch between the Dirac notation and the matrix notation.

Exercise 1.53 The matrix of operator $\hat{A}$ in the canonical basis is $\begin{pmatrix} 1 & -3i \\ 3i & 4 \end{pmatrix}$. Express this operator in the Dirac notation.

Exercise 1.54 The matrix of operator $\hat{H}$ in the canonical basis is $\frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$.

a) Express this operator in the Dirac notation.

b) Onto which states does $\hat{H}$ map $|H\rangle$ and $|V\rangle$?

c) How can one implement this operator using waveplates?

Definition 1.22 The operator defined in Ex. 1.54 is called the *Hadamard operator*.

Exercise 1.55 Show that for any orthonormal basis $\{|v_i\rangle\}$,

$$\sum_i |v_i\rangle \langle v_i| = \hat{\mathbf{1}} \tag{1.31}$$

Exercise 1.56 Show that $(\langle a| \, b\rangle)(\langle c| \, d\rangle) = \langle a| \, (|b\rangle\langle c|) \, |d\rangle$.

The result of the above exercise means that any sequence of the kind "...-ket-bra-ket-bra-ket-..." can be interpreted both as a sequence of scalar products "bra-ket" and as a sequence of operators "ket-bra". We can place parentheses at will. This property, known as associativity, is very useful in manipulating quantum expressions, for example in the following.

Suppose the matrix of $\hat{A}$ is known in some orthonormal basis $\{|v_i\rangle\}$ and we wish to find its matrix in another orthonormal basis, $\{|w_i\rangle\}$. This can be done by inserting $\hat{\mathbf{1}}$ from Ex. 1.55 as follows:

$$\begin{aligned} (\hat{A}_{ij})_{w\text{-basis}} &= \left\langle w_i \middle| \, \hat{A} \middle| \, w_j \right\rangle = \left\langle w_i \middle| \, \hat{\mathbf{1}}\hat{A}\hat{\mathbf{1}} \middle| \, w_j \right\rangle \\ &= \langle w_i| \left(\sum_k |v_k\rangle \langle v_k| \right) \hat{A} \left(\sum_m |v_m\rangle \langle v_m| \right) |w_j\rangle \\ &= \sum_k \sum_m \langle w_i| \, v_k\rangle \underbrace{\left\langle v_k \middle| \, \hat{A} \middle| \, v_m \right\rangle}_{(A_{km})_{v\text{-basis}}} \langle v_m| \, w_j\rangle \, . \end{aligned} \tag{1.32}$$

The last line can be interpreted as a product of three matrices: the first and the last are the matrices consisting of inner products of elements of the two bases, and the middle one is the matrix of the operator $\hat{A}$ in the basis $\{|v_i\rangle\}$.

Exercise 1.57 Find the matrix of the operator $\hat{A}$ from Ex. 1.53 in the $(|R\rangle, |L\rangle)$ basis

- using the Dirac notation, with the help of the result of Ex. 1.53 and Eq. (1.30);
- according to Eq. (1.32) (i.e. treating the canonical basis as the v-basis, circular as the w-basis, and calculating the matrix $\langle w_i| \, v_k\rangle$).

Check that the results are the same.

Definition 1.23 *Trace* of an operator $\hat{A}$ is the sum of the diagonal elements of its matrix in an orthonormal basis.

Exercise 1.58 Show that the trace of an operator is basis independent.

Definition 1.24 For any basis element $|v_i\rangle$, $\hat{P}_i = |v_i\rangle\langle v_i|$ is called the *projection operator.*

The effect of a measurement on a quantum stated can be expressed in terms of the projection operator associated with a random element of the measurement basis: in a measurement, the state is transformed as $|\psi\rangle \to \hat{P}_i |\psi\rangle = \langle v_i| \psi\rangle |v_i\rangle$ (hence the term "projective measurement"). Note that the length of the state $\hat{P}_i |\psi\rangle$ is not 1 but $\langle v_i| \psi\rangle$. This can be interpreted to reflect the fact that the probability to detect the system in the state $|v_i\rangle$ after the measurement is $|\langle v_i| \psi\rangle|^2$.

Exercise 1.59 Find the matrix of $\hat{\mathbb{P}}_i$ in the basis $\{|v_i\rangle\}$.

1.11 Observable Operators

The Second postulate of quantum physics, as defined in Sec. 1.8, states that a quantum measurement is performed in an orthonormal basis and the measurement result is a random element of that basis. Here we go one step further and associate with each basis element, $|v_i\rangle$, a real number, v_i. Then, instead of saying "the result of the measurement is state $|v_i\rangle$", we say "the result of the measurement is value v_i".

For example, with the measurement of a particle's position, each state with a certain position, e.g. $|x_i\rangle = |x = 3m\rangle$, is naturally associated with a value of the particle's coordinate ($x_i = 3$ m). For other measurements, such as that of a photon polarization, there is no natural connection between basis elements and numbers, but it can be introduced artificially. For example, if we are measuring in the canonical basis, we can associate number 1 with state $|H\rangle$ and -1 with state $|V\rangle$.

The information about the measurement basis and the values associated therewith can be conveniently expressed in the form of an operator,

$$\hat{V} = \sum_i |v_i\rangle \langle v_i| v_i. \tag{1.33}$$

This operator is called the *observable operator*. The elements $|v_i\rangle$ of the associated basis (the observable's *eigenbasis*) are the *eigenstates* or *eigenvectors* of the observable and the corresponding values x_i are its *eigenvalues*. Associating operators with observables may appear unnatural, yet it turns out to be quite useful in calculations and provides a fundamental relation between quantum and classical measurements.

When we wish to measure the observable, we construct a measurement apparatus associated with its eigenbasis. After the measurement, the apparatus will display the eigenvalue corresponding to the eigenvector onto which the state has been projected.

It is important to remember that eigenvalues of an observable operator correspond to physically measurable quantities, and must therefore be real.

Exercise 1.60 Show that the definition of eigenvalues and eigenstates given above is consistent with the traditional one from linear algebra, i.e. that for operator (1.33) and any i, $\hat{V} |v_i\rangle = v_i |v_i\rangle$.

Exercise 1.61 Write the matrix of $\hat{X}$ in the basis $\{|x_i\rangle\}$.

Exercise 1.62 Find the operators associated with the $\{|H\rangle, |V\rangle\}$, $\{|+\rangle, |-\rangle\}$ and $\{|R\rangle, |L\rangle\}$ bases and the eigenvalues ± 1 (respectively) in the Dirac notation. Find the matrices of these operators in the $\{|H\rangle, |V\rangle\}$ basis.

Answer:

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \equiv \hat{\sigma}_z \qquad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \equiv \hat{\sigma}_x \qquad \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \equiv \hat{\sigma}_y \tag{1.34}$$

Definition 1.25 These operators and matrices are called *Pauli operators* and *Pauli matrices*[14].

The following two exercises show that although an observable operator may be possible to implement physically, this physical operation has nothing to do with the operation employed in the measurement.

Exercise 1.63 Propose the implementation of the Pauli operators (up to a phase factor) by means of waveplates. **Hint:** Use Ex. 1.48.

Exercise 1.64 Construct apparata for measuring the Pauli observables of a photon's polarization.

Quantum measurement results are generally uncertain. If we measure the same observable in the same state many times, the result is random, albeit it obeys certain statistics. In classical physics, on the other hand, is we prepare the system in the same state and perform the same measurement, we will observe the same value over and over again. But the correspondence principle demands that quantum behavior becomes classical in the macroscopic limit. This correspondence is established through the expectation value — the weighted average of the values the measured observable takes.

Definition 1.26 Suppose a (not necessarily quantum) experiment on measuring quantity Q can yield any one of N possible outcomes $\{Q_i\}$ $(1 \le i \le N)\}$, with respective probability pr_i. Then the *expectation value* of the outcome is

$$\langle Q \rangle = \sum_{i=1}^{N} \mathrm{pr}_i Q_i. \tag{1.35}$$

Exercise 1.65 Find the expectation of the value displayed on the top side of a fair die.

As an example of how the expectation value of an observable upholds the correspondence principle, let us consider a light wave of $+45^\circ$ polarization entering a polarizing beam splitter. Suppose we repeat this experiment many times and are interested in the difference between the energies of the transmitted and reflected light.

Classically, the amplitudes of the horizontal and vertical polarizations in the wave are equal, so we expect the difference of the two energies to vanish. At the level of single photons, however, we will see random behavior: the photon will be transmitted or reflected with probability $\mathrm{pr}_H = \mathrm{pr}_V = 1/2$. Assigning the energy values $Q_H = \hbar\omega$ and $Q_V = -\hbar\omega$ (where $\hbar\omega$ is the single photon energy) to each of these events, we find that the expectation value for the energy difference is $\mathrm{pr}_H Q_H + \mathrm{pr}_V Q_V = 0$, as in the classical case. A further example of classical behavior of an expectation value is offered by the Ehrenfest theorem (Sec. 3.4).

This example also gives us a good illustration of the quantum-to-classical transition in the macroscopic limit. The more photons we send into the beam splitter, the smaller is the relative difference between the numbers of transmitted and reflected photons. More precisely, according to the laws of statistics, for a total of N photons, the difference will be on the scale of $\sqrt{N}$, so the relative difference scales as $\sqrt{N}/N = \sqrt{N}^{-1}$. For example, if $N = 10000$, we will observe reflection and transmission approximately 5000 times each, with a statistical deviation on a scale of 100. Now because the photon energy is very small (4×10^{-19} Joules for visible light), any experiment involving a macroscopically significant amount of light — even on a scale of nanojoules — will contain an enormous number of photons — and thus the relative difference between the transmitted and reflected energies will be minuscule.

Now let us recall the notion of the observable operator. As we see in the following exercise, this operator offers simple means to calculate the expectation value of a measurement outcome.

Exercise 1.66 We prepare the state $|\psi\rangle$ and measure an observable $\hat{V}$ in this state many times. Show that the expectation value of the observed result is

$$\langle V \rangle = \left\langle \psi \middle| \hat{V} \middle| \psi \right\rangle \tag{1.36}$$

[14]The meaning of subscripts x, y, and z will be clear when we study quantization of angular momenta.

Definition 1.27 The expectation value of an operator in the sense of Eq. (1.36) is also called the *quantum average* or *quantum expectation* of this observable in state $|\psi\rangle$.

In our above example on calculating the expectation values of the difference between the energies of the transmitted and reflected light, the observable is

$$\hat{E} = \hbar\omega \, |H\rangle\langle H| - \hbar\omega \, |V\rangle\langle V| = \hbar\omega\hat{\sigma}_z, \tag{1.37}$$

and its expectation value is

$$\langle E\rangle = \left\langle + \middle| \, \hat{E} \middle| + \right\rangle = 0. \tag{1.38}$$

1.12 Adjoint and self-adjoint operators

So far, we studied linear operators acting on ket-vectors from the left. It turns out that one can also define the action of a linear operator on bra-vectors from the right. As we know (Ex. 1.56), chains of bras and kets have the property of associativity: the product of two inner product $\langle a| \, b\rangle \, \langle c| \, d\rangle$ can be interpreted as an inner product of bra-vector $\langle a|$ and the result of action of operator $|b\rangle\langle c|$ on ket-vector $|d\rangle$. We can, however, also say that

$$\langle a| \, b\rangle \, \langle c| \, d\rangle = \left[\langle a| \, (|b\rangle\langle c|)\right] |d\rangle \, , \tag{1.39}$$

with the operator $|b\rangle\langle c|$ acting on bra-vector $\langle a|$ from the right, generating bra-vector $\langle a| \, b\rangle \, \langle c|$.

This notion is not limited to outer products, but can be applied to any operator. This is because any operator can be written as a sum of inner products, $\hat{A} = \sum_{ij} A_{ij} \, |v_i\rangle\langle v_j|$ (see Ex. 1.52). Using the linearity of the inner product, we find

$$\langle a| \, \hat{A} = \sum_{ij} A_{ij} \, \langle a| \, v_i\rangle \, \langle v_j| \tag{1.40}$$

Note that it is meaningless to write an operator acting on a bra-vector from the left or on a ket-vector from the right.

Exercise 1.67 Show that the matrix form of the vector $\langle a| \, \hat{A}$ is given by the product of the matrix forms of vector $\langle a|$ and operator $\hat{A}$.

Exercise 1.68 Verify that for any operator $\hat{A}$ and vectors $|a\rangle$ and $|c\rangle$,

$$\left(\langle a| \, \hat{A}\right) |c\rangle = \langle a| \left(\hat{A} \, |c\rangle\right). \tag{1.41}$$

Suppose now an that we have an operator $\hat{A}$ that maps ket-vector $|a\rangle$ onto ket-vector $|b\rangle$. What is the operator that maps *bra*-vector $|a\rangle$ onto bra-vector $|b\rangle$? It turns out that this operator is not the same as $\hat{A}$, but has a relatively simple relation thereto.

Definition 1.28 An operator $\hat{A}^\dagger$ ("A-dagger") is called *adjoint* (*Hermitian conjugate*) to $\hat{A}$ if, whenever $|b\rangle = \hat{A} \, |a\rangle$, we also have $\langle b| = \langle a| \, \hat{A}^\dagger$. If $\hat{A} = \hat{A}^\dagger$, the operator is called *Hermitian* or *self-adjoint*.

Unlike bra- and ket-vectors, operators and their adjoints are defined over the same Hilbert space (or, more precisely, in both bra- and Ket- spaces: they act on bra-vectors from the right, and on ket-vectors from the left). This is why self-adjoint operators are possible.

Exercise 1.69 Show that the matrix of $\hat{A}^\dagger$ relates to the matrix of $\hat{A}$ through transposition and complex conjugation.

Exercise 1.70 Show that, for any operator, $(\hat{A}^\dagger)^\dagger = \hat{A}$.

Exercise 1.71 Show that Pauli operators are Hermitian.

Exercise 1.72 Show that $(|c\rangle\langle b|)^\dagger = |b\rangle\langle c|$.

As we see from this exercise, the adjoint of an outer product operator is somehow related to its inverse: if the "direct" operator maps $|b\rangle$ onto $|c\rangle$, its adjoint does the opposite. This is not always the case: for example, the adjoint of $\lambda\hat{\mathbf{1}}$ is $\lambda^*\hat{\mathbf{1}}$, and the product of the two is $|\lambda|^2\hat{\mathbf{1}} \neq \hat{\mathbf{1}}$. However, there is an important class of operators, the so-called unitary operators, for which the inverse is the same as the adjoint. We discuss these operators in detail in Sec. 1.15.

Exercise 1.73 Show that

a)
$$(\hat{A} + \hat{B})^\dagger = \hat{A}^\dagger + \hat{B}^\dagger \tag{1.42}$$

b)
$$(\lambda\hat{A})^\dagger = \lambda^*\hat{A}^\dagger \tag{1.43}$$

c)
$$(\hat{A}\hat{B})^\dagger = \hat{B}^\dagger\hat{A}^\dagger \tag{1.44}$$

If we summarize the results presented in Eqs. (1.18) and (1.44), we obtain a simple rule for finding the adjoint of any expression consisting of vectors and operators:

a) invert the order of all products;

b) conjugate all numbers;

c) replace all kets by bras and vice versa;

d) replace all operators by their adjoints.

$$\text{E.g: Adjoint}\left(\sum_i \lambda_i \hat{A}\hat{B}\,|a_i\rangle\right) = \sum_i \lambda_i^*\,\langle a_i|\,\hat{B}^\dagger\hat{A}^\dagger \tag{1.45}$$

Exercise 1.74 Show that if we apply the above rule to a "sandwich", we obtain a complex conjugate:
$$\langle\phi|\,\hat{A}\,|\psi\rangle = \langle\psi|\,\hat{A}^\dagger\,|\phi\rangle^*\,. \tag{1.46}$$

Exercise 1.75 By a counterexample, show that two operators being Hermitian does not guarantee that their product is also Hermitian.

Definition 1.29 Expression of an operator in the form (1.33) with $\{|x_i\rangle\}$ being an orthonormal basis is called the *spectral decomposition* of the operator.

Definition 1.30 An eigenvalue is called *degenerate* if it corresponds to more than one linearly independent eigenstate.

Exercise 1.76 Show that

a) operators corresponding to physical observables (1.33) are Hermitian;

b)* Any Hermitian operator can be associated with a physical observable, i.e. has a spectral decomposition with real eigenvalues and eigenstates that form an orthonormal basis.

In quantum mechanics, an observable operator is frequently known only in the form of a matrix defined in some basis. In order to understand the physics of this operator, it is sometimes necessary to find the eigenbasis and the eigenvalues of this operator. This is done using the standard procedure of diagonalizing Hermitean operators that is known from linear algebra (and reviewed in the solutions to the exercises below).

As we found in Ex. 1.76, for any Hermitean operator, one can find a set of eigenvalues $\{v_i\}$ and a set of eigenvectors $\{|v_i\rangle\}$ that forms an orthonormal basis. Once this set is found, we can write the operator in the form of Eq. (1.33). The matrix of the operator in basis $\{|v_i\rangle\}$ is then diagonal (Ex. 1.61).

Exercise 1.77 Find the eigenvectors and eigenbases of the Pauli matrices. Verify consistency with the definition given in Ex. 1.62

Exercise 1.78 Same for the rotation of the plane of two-dimensional geometric vectors by angle ϕ. Is this a Hermitian operator?

Exercise 1.79 In a three-dimensional Hilbert space, three operators, in an orthonormal basis $\{|v_1\rangle, |v_2\rangle, |v_3\rangle\}$ have the following matrices:

a) $\hat{L}_x \leftrightarrow \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$,

b) $\hat{L}_y \leftrightarrow \begin{pmatrix} 0 & -i & 0 \\ i & 0 & -i \\ 0 & i & 0 \end{pmatrix}$

c) $\hat{L}_z \leftrightarrow \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}$.

Find their eigenvalues and eigenstates.

Exercise 1.80*

a) Show that if $\left\langle \psi \right| \hat{A} \left| \psi \right\rangle = \left\langle \psi \right| \hat{B} \left| \psi \right\rangle$ for all $|\psi\rangle$, then $\hat{A} = \hat{B}$.

b) Show that if $\left\langle \psi \right| \hat{A} \left| \psi \right\rangle$ is a real number for all $|\psi\rangle$, then $\hat{A}$ is Hermitian.

1.13 Commutator

As we discussed, not all operators commute. The degree of non-commutativity turns out to play an important role in quantum mechanics and is quantified by the operator known as the *commutator*.

Definition 1.31 For any two operators $\hat{A}$ and $\hat{B}$, we define:

$$\textit{Commutator } [\hat{A}, \hat{B}] = \hat{A}\hat{B} - \hat{B}\hat{A}; \tag{1.47}$$

$$\textit{Anticommutator } \{\hat{A}, \hat{B}\} = \hat{A}\hat{B} + \hat{B}\hat{A}. \tag{1.48}$$

Exercise 1.81 Show that:

a)
$$\hat{A}\hat{B} = \frac{1}{2}([\hat{A}, \hat{B}] + \{\hat{A}, \hat{B}\}); \tag{1.49}$$

b)
$$[\hat{A}, \hat{B}] = -[\hat{B}, \hat{A}]; \tag{1.50}$$

c)
$$[\hat{A}, \hat{B}]^\dagger = [\hat{B}^\dagger, \hat{A}^\dagger]; \tag{1.51}$$

d)
$$[\hat{A}, \hat{B}\hat{C}] = [\hat{A}, \hat{B}]\hat{C} + \hat{B}[\hat{A}, \hat{C}]; \tag{1.52}$$

Exercise 1.82 If $\hat{A}$ and $\hat{B}$ are Hermitian, so are

a) $i[\hat{A}, \hat{B}]$

b) $\{\hat{A}, \hat{B}\}$

Exercise 1.83 Find the commutation relations of the Pauli operators. **Hint:** you can do it in the matrix form, in any orthonormal basis.

Exercise 1.84 Consider two Hermitian operators $\hat{A}$ and $\hat{B}$. Show that they are simultaneously diagonalizable (become diagonal in the same orthonormal basis) if and only if $[\hat{A}, \hat{B}] = 0$.

As the Second Postulate of quantum mechanics says, every measurement is associated with some orthonormal basis. We have an option of associating every basis element with a real number, thus defining an observable operator (see Sec. 1.11.

The last exercise shows that the statement that two observable operators commute is equivalent to the measurements defined with these observables being associated with the same orthonormal basis. Such observables are "compatible": a system prepared in an eigenstate $|v_i\rangle$ of observable $\hat{A}$ will remain in this state when observable $\hat{B}$ is measured and the measurement result will be certain: $|v_i\rangle$. If, on the other hand, $\hat{A}$ and $\hat{B}$ don't commute, a system prepared in an eigenstate of observable $\hat{A}$ can give a random result if $\hat{B}$ is measured[15]. The degree of this randomness is quantified by the Heisenberg uncertainty principle, which we study next.

1.14 The uncertainty principle

Before we formulate the uncertainty principle, we first need to define the notion of uncertainty. Let us again consider a (not necessarily quantum) experiment on measuring random variable Q that can take on any one of N possible values $\{Q_i\}$, with respective probabilities pr_i. While the expectation value, $\langle Q \rangle = \sum_{i=1}^{N} \mathrm{pr}_i Q_i$, shows the mean measurement output, the statistical uncertainty shows by how much, on average, a particular measurement result will deviate from the mean (Fig. 1.2).

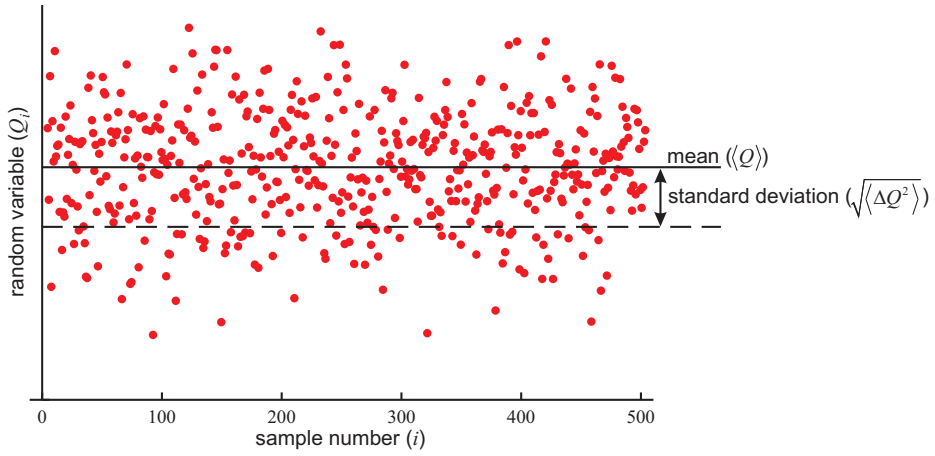


Figure 1.2: Mean and rms standard deviation of a random variable.

Definition 1.32 The *(mean square) variance* of random variable Q is

$$\langle \Delta Q^2 \rangle = \left\langle (Q - \langle Q \rangle)^2 \right\rangle = \sum_i \mathrm{pr}_i \, (Q_i - \langle Q \rangle)^2 \, . \tag{1.53}$$

[15] Even if $\hat{A}$ and $\hat{B}$ don't commute, this does not mean that measuring observable $\hat{B}$ in an eigenstate of $\hat{A}$ will *always* give a random result. For example, suppose the eigenvectors of $\hat{A}$ in a three-dimensional Hilbert space are $|v_1\rangle$, $|v_2\rangle$ and $|v_3\rangle$, and the eigenvectors of $\hat{B}$ are $|v_1\rangle$, $|v_2'\rangle$ and $|v_3'\rangle$ (all eigenvalues are nondegenerate). These sets are different, so $\hat{A}$ and $\hat{B}$ do not commute. However, they have one common eigenstate $|v_1\rangle$, and the system prepared in this state will yield a certain result when either of the two observables are measured.

The *root mean square (rms) standard deviation*, or *uncertainty* of Q is then $\sqrt{\langle \Delta Q^2 \rangle}$.

Exercise 1.85 Show that, for any random variable Q,

$$\langle \Delta Q^2 \rangle = \langle Q^2 \rangle - \langle Q \rangle^2 . \tag{1.54}$$

Exercise 1.86 Calculate the mean square variance of the value displayed on the top side of a fair die.

Exercise 1.87 Show that in the quantum case, the uncertainty associated with measuring observable $\hat{X}$ in state $|\psi\rangle$ is given by the expectation value of the operator

$$\langle \Delta X^2 \rangle = \left\langle \psi \right| \left(\hat{X} - \left\langle \psi \right| \hat{X} \left| \psi \right\rangle \right)^2 \left| \psi \right\rangle \tag{1.55}$$

and that this uncertainty can be calculated according to

$$\left\langle \Delta \hat{X}^2 \right\rangle = \left\langle \psi \right| \hat{X}^2 \left| \psi \right\rangle - \left\langle \psi \right| \hat{X} \left| \psi \right\rangle^2 . \tag{1.56}$$

Exercise 1.88 Show that an observable $\hat{X}$ in a certain quantum state $|\psi\rangle$ has zero uncertainty if an only if $|\psi\rangle$ is an eigenstate of the observable (i.e. $\hat{X}|\psi\rangle = x|\psi\rangle$).

Exercise 1.89 Show that for any Hermitian $\hat{A}$ and $\hat{B}$,

$$\left\langle \{\hat{A}, \hat{B}\} \right\rangle = 2 \operatorname{Re} \left\langle \hat{A}\hat{B} \right\rangle \tag{1.57}$$

$$\left\langle [\hat{A}, \hat{B}] \right\rangle = 2i \operatorname{Im} \left\langle \hat{A}\hat{B} \right\rangle ; \tag{1.58}$$

$$\left| \left\langle [\hat{A}, \hat{B}] \right\rangle \right|^2 \leq 4 \left| \left\langle \hat{A}\hat{B} \right\rangle \right|^2 . \tag{1.59}$$

Exercise 1.90 Show that, for any to Hermitian operators $\hat{A}$, $\hat{B}$, and any state $|\psi\rangle$,

$$\left\langle \psi \right| \hat{A}^2 \left| \psi \right\rangle \left\langle \psi \right| \hat{B}^2 \left| \psi \right\rangle \geq \left| \left\langle \psi \right| \hat{A}\hat{B} \left| \psi \right\rangle \right|^2 . \tag{1.60}$$

Hint: Let $|a\rangle = \hat{A}|\psi\rangle$ and $|b\rangle = \hat{B}|\psi\rangle$ and apply the Cauchy-Schwarz inequality.

Exercise 1.91 Prove the *Heisenberg uncertainty principle*: For Hermitian $\hat{A}$, $\hat{B}$, and any state $|\psi\rangle$

$$\left\langle \Delta \hat{A}^2 \right\rangle \left\langle \Delta \hat{B}^2 \right\rangle \geq \frac{1}{4} \left| \left\langle [\hat{A}, \hat{B}] \right\rangle \right|^2 . \tag{1.61}$$

assuming for simplicity that

$$\left\langle \hat{A} \right\rangle = \left\langle \hat{B} \right\rangle = 0. \tag{1.62}$$

Exercise 1.92 Redo the proof without assuming Eq. (1.62). Would the uncertainty principle (1.61) remain valid if its right-hand side were $\frac{1}{4} \left| \left\langle \{\hat{A}, \hat{B}\} \right\rangle \right|^2$ or $\left| \left\langle \hat{A}\hat{B} \right\rangle \right|^2$?

Exercise 1.93 Show that, if $[\hat{A}\hat{B}] = \epsilon \cdot \hat{\mathbf{1}}$, then the uncertainty product is independent from $|\psi\rangle$:

$$\left\langle \Delta \hat{A}^2 \right\rangle \left\langle \Delta \hat{B}^2 \right\rangle \geq \frac{|\epsilon|^2}{4}. \tag{1.63}$$

Exercise 1.94 Find $\langle \psi | \hat{\sigma}_x | \psi \rangle$,, $\left\langle \psi \right| \Delta \hat{\sigma}_x^2 \left| \psi \right\rangle$, $\langle \psi | \hat{\sigma}_y | \psi \rangle$,, $\left\langle \psi \right| \Delta \hat{\sigma}_y^2 \left| \psi \right\rangle$, as well as $\langle \psi | [\hat{\sigma}_x, \hat{\sigma}_y] | \psi \rangle$ explicitly, in the matrix form, for $|\psi\rangle = |H\rangle$. Verify that the uncertainty principle for $\hat{\sigma}_x$, $\hat{\sigma}_y$ and $|\psi\rangle = |H\rangle$ holds.

The Heisenberg uncertainty principle is one of the most important tenet of quantum physics and one of the primary signatures that distinguishes it from classical. It was also one of the most controversial ideas at the time of quantum mechanics' creation. Similarly to the measurement postulate, the uncertainty principle appeared to be in direct contradiction with the deterministic picture of the world accepted by classical physics. According to the latter, any uncertainty one may have in a measurement is a consequence of an imperfect measurement apparatus, and can be indefinitely reduced by improving that apparatus. In the framework of quantum mechanics, this is not the case: there exist observables that are "incompatible": if one builds an apparatus that is precise at measuring one observable in a particular state of the system, this apparatus is doomed to perform poorly when the other observable is measured, no matter how good it is[16].

Sometimes the uncertainty principle is misconceived as a statement that if one observable is measured with a certain precision, the precision of a *subsequent* measurement of the other observable is limited. While this is correct, it is not what the statement of the uncertainty principle. According to this principle, the uncertainty is an intrinsic property of the state (and the observables), and it does not depend on the sequence in which the measurements are performed.

1.15 Unitary operators

The role that linear operators play in relation to quantum physics is not limited by observables. The very definition of operator as a linear map suggests another important application: an *evolution operator*, $\hat{U}(t)$ can be used to describe evolution of a quantum state $|\psi\rangle$ with time t:

$$|\psi(t)\rangle = \hat{U}(t)\,|\psi(0)\rangle\,. \tag{1.64}$$

Before we define the specific form of the evolution operator, let us first have a more general discussion. What can be said about an evolution operator, without even knowing the physical system in which the evolution occurs? As it turns out, quite a lot. Every evolution operator must map a physical state (i.e. a vector of norm 1) onto another physical state. Let us now discuss some of the consequences of this property.

Definition 1.33 Linear operators that map all vectors of norm 1 onto vectors of norm 1 are called *unitary*.

Exercise 1.95 Show that unitary operators preserve the norm of any vector, i.e. if $|a'\rangle = \hat{U}\,|a\rangle$ then $\langle a|\ a\rangle = \langle a'|\ a'\rangle$.

Exercise 1.96 Show an operator $\hat{U}$ is unitary if and only if it preserves the inner product of any two vectors, i.e. if $|a'\rangle = \hat{U}\,|a\rangle$ and $|b'\rangle = \hat{U}\,|b\rangle$ then $\langle a|\ b\rangle = \langle a'|\ b'\rangle$.

Exercise 1.97 Show that

a) a unitary operator maps any orthonormal basis $\{|w_i\rangle\}$ onto an orthonormal basis. This operator can be written in the form $\hat{U} = \sum_i |v_i\rangle\langle w_i|$, where $|v_i\rangle = \hat{U}\,|w_i\rangle$.

b) conversely, for two orthonormal bases $\{|v_i\rangle\}, \{|w_i\rangle\}$, operator $\hat{U} = \sum_i |v_i\rangle\langle w_i|$ is unitary (in other words, *any* operator that maps an orthonormal basis onto an orthonormal basis is unitary).

Exercise 1.98 Show that an operator $\hat{U}$ is is unitary if and only if $\hat{U}^\dagger\hat{U} = \hat{U}\hat{U}^\dagger = \hat{\mathbf{1}}$

Exercise 1.99 Show that:

[16] The uncertainty principle also contradicts certain primitive interpretations of marxism, which postulates that human ability of cognition is unlimited. In 1948 Soviet Union, this contradiction triggered "debates" between physicists and Communist scholars. The result of these debates was predetermined: all modern physics research would be closed down and physicists would be sent to Gulag. The situation was rescued by I. Kurchatov, the head of the atomic bomb project. He approached Stalin and explained that if the persecution of modern physics continued, the very first project that would have to be closed down would be that on the atomic bomb, which heavily relied on quantum mechanics and the theory of relativity. Stalin was compelled to back off and order the end of all the "debates".

a) If a unitary operator has any eigenvalues, they all have absolute value 1, i.e. can be written as $e^{i\theta},\ \theta \in \mathbb{R}$

b) A diagonalizable operator (i.e. operator whose matrix becomes diagonal in some basis) with eigenvalues of absolute value 1 is unitary.

So we have produced several equivalent definitions of a unitary operator:

- it preserves the norm of a vector;
- it preserves the inner product;
- it maps any orthonormal basis onto an orthonormal basis;
- in some basis, it has a form of a diagonal matrix with diagonal values of absolute value 1;
- its adjoint is its inverse.

If an operator satisfies one of the above definitions, it satisfies all of them. Any operator that describes physical evolution of a quantum state must be unitary.

It is important that all unitary operators are invertible and the inverse of a unitary operator is also a unitary operator. This has quite a deep consequence. If we know the evolution operator and the state resulting from this evolution, we can reconstruct the initial state by applying the inverse evolution operator to the final state. No information is ever lost during a quantum evolution of an isolated quantum system. In the language of statistical physics this means that the entropy of a physical system does not increase during its evolution.

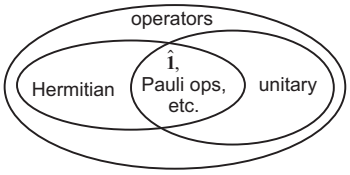


Figure 1.3: Relations among types of operators

The families of Hermitian and unitary operators overlap, but they are not identical (Fig. 1.3). An operator that is both Hermitian and unitary must be inverse to itself.

Exercise 1.100 Verify if the following operators are unitary:

a) Pauli operators;

b) rotation by angle ϕ in the linear space of two-dimensional geometric vectors (over $\mathbb{R}$).

1.16 Functions of operators

The concept of the operator function has many applications in linear algebra and differential equations. As we shall see in the next section, operator functions are also handy in quantum mechanics as they permit easy calculation of evolution operators.

Definition 1.34 Consider a complex function $f(x)$ defined on $\mathbb{C}$. The *operator function* $f(\hat{A})$ of Hermitian operator $\hat{A}$ is the following operator:

$$f(\hat{A}) = \sum_i f(a_i) \left|a_i\right\rangle \left\langle a_i\right|, \tag{1.65}$$

where $\{|a_i\rangle\}$ is an orthonormal basis in which $\hat{A}$ diagonalizes:

$$\hat{A} = \sum_i a_i \left|a_i\right\rangle \left\langle a_i\right|. \tag{1.66}$$

Exercise 1.101 Find the matrix of $\sqrt{\hat{A}}$ and $\ln \hat{A}$ in the orthonormal basis in which

$$\hat{A} \leftrightarrow \begin{pmatrix} 1 & 3 \\ 3 & 1 \end{pmatrix}$$

Exercise 1.102 Find $e^{i\theta\hat{\sigma}_x}$, $e^{i\theta\hat{\sigma}_y}$, $e^{i\theta\hat{\sigma}_z}$ and their matrices in the canonical basis.

Exercise 1.103 Find the matrix of $e^{i\theta\hat{A}}$, where $\hat{A} = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$.

Hint: One of the eigenvalues of $\hat{A}$ is 0, which means that the corresponding eigenstates do not appear in decomposition (1.66). However, the exponential of the corresponding eigenvalue is not zero, and the corresponding eigenstates do show up in the operator function (1.65).

Exercise 1.104 Show that, for any operator $\hat{A}$,

a) $[\hat{A}, f(\hat{A})] = 0$ for any function f;

b) $\hat{A}^{m+n} = \hat{A}^m \hat{A}^n$ for any natural numbers m and n.

Exercise 1.105 Suppose $f(x)$ has a Taylor decomposition $f(x) = f_0 + f_1 x + f_2 x^2 + \ldots$. Show that $f(\hat{A}) = f_0\hat{\mathbf{1}} + f_1\hat{A} + f_2\hat{A}^2 + \ldots$

Exercise 1.106 Show that, if operator $\hat{A}$ is Hermitian, operator $e^{i\hat{A}}$ is unitary and $e^{-i\hat{A}} = \left(e^{-i\hat{A}}\right)^{-1}$.

Exercise 1.107* Let $\vec{v} = (v_x, v_y, v_z)$ be a unit length vector. Show that:

$$e^{i\theta\vec{v}\hat{\vec{\sigma}}} = \cos\theta\,\hat{\mathbf{1}} + i\sin\theta\,\vec{v}\,\hat{\vec{\sigma}}, \tag{1.67}$$

where $\hat{\vec{\sigma}} = (\hat{\sigma}_x, \hat{\sigma}_y, \hat{\sigma}_z)$, $\vec{v}\,\hat{\vec{\sigma}} = v_x\hat{\sigma}_x + v_y\hat{\sigma}_y + v_z\hat{\sigma}_z$. **Hint:** this problem has a simple solution if the basis is chosen cleverly.

Definition 1.35 Suppose state $|\psi(t)\rangle$ depends on certain parameter t. The *derivative* of $|\psi(t)\rangle$ with respect to t is defined as vector

$$\frac{\mathrm{d}\,|\psi\rangle}{\mathrm{d}t} = \lim_{\Delta t\to 0} \frac{|\psi(t+\Delta t)\rangle - |\psi(t)\rangle}{\Delta t}. \tag{1.68}$$

Similarly, the derivative of operator $\hat{Y}(t)$ with respect to t is operator

$$\frac{\mathrm{d}\hat{Y}}{\mathrm{d}t} = \lim_{\Delta t\to 0} \frac{\hat{Y}(t+\Delta t) - \hat{Y}(t)}{\Delta t}. \tag{1.69}$$

Exercise 1.108 Suppose the the matrix form of vector $|\psi(t)\rangle$ is

$$|\psi(t)\rangle = \begin{pmatrix} \psi_1(t) \\ \vdots \\ \psi_N(t) \end{pmatrix}$$

in some basis. Show that

$$\frac{\mathrm{d}\,|\psi\rangle}{\mathrm{d}t} = \begin{pmatrix} \mathrm{d}\psi_1(t)/\mathrm{d}t \\ \vdots \\ \mathrm{d}\psi_N(t)/\mathrm{d}t \end{pmatrix}.$$

Write a similar expression for the matrix form of an operator derivative.

Exercise 1.109 Suppose operator $\hat{A}$ is constant and t is a real parameter. Show that $\frac{\mathrm{d}}{\mathrm{d}t}e^{i\hat{A}t} = i\hat{A}e^{i\hat{A}t}$.

Exercise 1.110* For two operators $\hat{A}$ and $\hat{B}$, suppose that $[\hat{A}, \hat{B}] = ic\hat{\mathbf{1}}$, c being a complex number. Prove the *Baker-Hausdorff-Campbell formula*[17]

$$e^{\hat{A}+\hat{B}} = e^{\hat{A}} e^{\hat{B}} e^{-ic/2} \tag{1.70}$$

in the following steps.

a) Show that

$$[\hat{A}, \hat{B}^n] = nc\hat{B}^{n-1}; \tag{1.71}$$

b) Show that

$$[\hat{A}, e^{\hat{B}}] = ce^{\hat{B}}. \tag{1.72}$$

Hint: use the Taylor series expansion for the exponential.

c) For an arbitrary number λ and operator $\hat{G}(\lambda) = e^{\lambda\hat{A}} e^{\lambda\hat{B}}$, show that

$$\frac{\mathrm{d}\hat{G}(\lambda)}{\mathrm{d}\lambda} = \hat{G}(\lambda)(\hat{A} + \hat{B} + \lambda c) \tag{1.73}$$

d) Solve the differential equation (1.73) to show that

$$\hat{G}(\lambda) = e^{\lambda\hat{A}+\lambda\hat{B}+\lambda^2 c/2}. \tag{1.74}$$

e) Prove the Baker-Hausdorff-Campbell formula using Eq. (1.74).

1.17 Schrödinger equation

Schrödinger equation governs the evolution of quantum states with time. Because of its importance, and because it cannot be derived from what we have learned so far, we can consider it the *Third Quantum Mechanics Postulate*.

In classical physics, the complete set of equations of motion can be obtained from the expression for the *Hamiltonian* (full energy) of the system. That is, the entire information about the predictable behavior of the system, for any initial state, is contained in that Hamiltonian. The same is true for quantum physics. Unlike classical physics, however, the quantum Hamiltonian, being a physical observable quantity, corresponds to a Hermitian operator.

In order to visualize the Hamiltonian operator, let us think of an atom in the framework of the Bohr model. According to this model, there exist certain orbits such that, if the electron moves along one of them, it can stay there for a long time period. Each Bohr's orbit corresponds to a certain constant value of the atomic energy.

Bohr formulated his postulates empirically, based on the experimental data available at that time. But if we review his postulates in the framework of quantum theory, we can see that Bohr's orbits are simply quantum states of the atom that correspond to certain energy values. In other words, they are eigenstates of the Hamiltonian. Denoting these states as $|E_i\rangle$, the atomic Hamiltonian can be written as

$$\hat{H} = \sum_i E_i \left|E_i\right\rangle\left\langle E_i\right|, \tag{1.75}$$

where E_i is the energy associated with each Borh's orbit and the summation is performed over all these orbits.

[17] This is a simplified form of the Baker-Hausdorff-Campbell formula. The full form of this formula is much more complicated and holds for the case when $[\hat{A}, \hat{B}]$ does not commute with $\hat{A}$ or $\hat{B}$.

Now let us formulate the Schrödinger equation. It is a differential equation which relates the current quantum state $|\psi(t)\rangle$ of a physical system, its derivative of with respect to time and the Hamiltonian of the system. Specifically:

$$\frac{d\,|\psi\rangle}{dt} = -\frac{i}{\hbar}\hat{H}\,|\psi\rangle\,, \tag{1.76}$$

where t is time, $\hbar = 1.05457148 \times 10^{-34}$ m^{2}kg/s is the Planck constant, and $\hat{H}$ is the Hamiltonian. Knowing the initial state of the system and its Hamiltonian, we can solve the Schrödinger equation and predict the state at any future moment in time.

Note a special role of time in quantum physics. Unlike all other physical observables, we do not treat it as an operator. There are no eigenstates of time, nor quanta of time, nor does the uncertainty principle hold for measurements of time. Time is simply a continuous variable.

Exercise 1.111 Show that the solution to the Schödringer equation is given by

$$|\psi(t)\rangle = \hat{U}(t)\,|\psi(t=0)\rangle\,, \tag{1.77}$$

where

$$\hat{U}(t) = e^{-i\frac{\hat{H}}{\hbar}t} \tag{1.78}$$

is the evolution operator. Show that this operator is unitary.

Exercise 1.111 shows a way to solve the Schrödinger equation. What we need to do is to calculate the evolution operator, using the tricks learned in the previous section, and then apply it to the initial state. Alternatively, one can write the differential equation (1.76) in the matrix form and solve for each matrix element. Both techniques are illustrated in the following exercise.

Another important result, which follows from the results of Ex. 1.106 and 1.111 is that the inversion of time, i.e. replacing t by $-t$ in Eq. (1.78), inverts the evolution. Of course, this is only a mathematical possibility, as it is not possible to invert time. However, in some systems it may be possible to invert the Hamiltonian, i.e. change the physical conditions surrounding the system in such a way that $\hat{H}$ is replaced by $-\hat{H}$, which would also lead to the inversion of evolution. This phenomenon has many interesting applications, for example, quantum memory for light.

Before we begin to practise solving the Schrödinger equation, a word of caution. The physical system we have used so far, the photon, is a relativistic object, and is evolution follows the laws of quantum electrodynamics rather than non-relativistic quantum mechanics that we are studying now. The energy of the photon equals $\hbar\omega$, where ω is the frequency of the light wave, independently of the polarization. However, for the purpose of practising (before we learned other physical systems to which the physics of the Schrödinger evolution applies to its full extent), let us *pretend* that the energy of the photon can take on different values for different polarization and that we can construct physical settings corresponding to arbitrary Hermitian Hamiltonians.

This assumption is not entirely frivolous. The evolutions that, as we shall see, take place under the Hamiltonians of Ex. 1.112 can occur in birefringent optics with their optical axes oriented at certain angles. This coincidence is however, fortuitous: birefringent effects are of classical optical nature and occur due to different phase velocities associated with different polarizations, rather than different photon energies.

Exercise 1.112 Write the Schrödinger equation for the following Hamiltonians:

a) $\hat{H} = \hbar\omega\hat{\sigma}_z$ (a birefringent material with optical axes oriented along the horizontal and vertical directions);

b) $\hat{H} = \hbar\omega\hat{\sigma}_x$ (a birefringent material with optical axes oriented at 45°);

c) $\hat{H} = \hbar\omega(4\hat{\sigma}_z + 3\hat{\sigma}_x)$ (more complex birefringent properties).

For each case, find the polarization state of the photon at time t if its initial state is either $|\psi(0)\rangle = |H\rangle$ or $|\psi(0)\rangle = |{+45^\circ}\rangle$. Use both techniques: solving the differential equation for the state vector and calculating the evolution operator.

You may have noticed while solving this exercise that the character of the evolution under Schrödinger evolution strongly depends on the initial state. If this state is an eigenstate of the Hamiltonian [e.g. state $|H\rangle$ in part (a) or state $|+\rangle$ in part (b)], the only change imposed by the evolution is a physically unobservable phase factor. Otherwise the state changes its physical nature, evolving from $|+\rangle$ to $|-\rangle$ and back in part (a) and between $|H\rangle$ and $|V\rangle$ in part (b).

The reason for this behavior is easily understood if we look at the expression (1.78) of the evolution operator and recall the definition (1.65) of the operator function. The evolution operator is an exponential of the Hamiltonian, and hence it has the same eigenstates as does the Hamiltonian. Accordingly, the action of this unitary operator on a Hamiltonian eigenstate consists of multiplication by an eigenvalue, which, as we found in Ex. 1.99, has absolute value 1, i.e. is just a phase factor.

Our interpretation of Bohr's orbits as energy eigenstates is thus consistent with Bohr's postulate that electrons are able to remain in these orbits for a long time. In fact, in Chapter 4 we will use this property of Bohr's orbits to calculate their wave functions.

Because of the character of their evolution, the eigenstates of the energy operator are sometimes called *stationary*.