Quantum Mechanics I

Michael Pustilnik

School of Physics
Georgia Institute of Technology
Atlanta, GA 30032

E-mail: pustilnik@gatech.edu

fall 2023 edition

Contents

1 Probabilities and States

1.1 Probabilities

1.1.1 Ensembles, probabilities, and expectation values

Consider a collection of a very large (ideally, infinite) number of identically prepared copies of the system of interest on which measurements are to be made. Borrowing the term from statistical physics, we will call such collection an ***ensemble***. Let A be an observable quantity (henceforth simply an ***observable***) that can be measured on every individual system in the ensemble, and let $\{a\}$ be a set of all possible outcomes of such individual measurements. For simplicity, we will assume this set to be *discrete*.[1]

Imagine now a hypothetical experiment in which we pick at random n systems from ensemble E and measure observable A on each individual system in this *subensemble*. The ***probability*** that a measurement of A yields the outcome a is the limit[2,3]

$$\mathrm{Prob}_E(A=a) = \lim_{n\to\infty} \frac{n_{A=a}}{n}, \tag{1.1}$$

where $n_{A=a}$ is the number of individual measurements resulting in $A=a$. The probabilities (1.1) are, obviously, real dimensionless numbers that satisfy[2]

$$0 \le \mathrm{Prob}(A=a) \le 1, \quad \sum_a \mathrm{Prob}(A=a) = 1. \tag{1.2}$$

The second equation here expresses the obvious property that the probability that a measurement of A results in *some* outcome is 1.

We say that a measurement of A is *certain* to yield $A=a$ if $\mathrm{Prob}(A=a)=1$. Note that in the probabilistic context the words *certain* and *certainty* should be understood as *almost always*. That is, $\mathrm{Prob}(A=a)=1$ does not imply that every individual measurement of A is guaranteed to produce the outcome $A=a$. Instead, it merely means that the outcomes $A\neq a$ are exceedingly rare. Similarly, saying that a measurement of A is certain *not* to yield $A=a$ if $\mathrm{Prob}(A=a)=0$ does not mean that the outcome $A=a$ is never observed.

In practice, it is often easier to design experiments that produce sums of outcomes of individual measurements

$$\sum_{i=1}^{i=n} a_i = \sum_a a n_{A=a}$$

rather than the counts $n_{A=a} = \sum_i \delta_{a_i,a}$ for particular outcomes. Here a_i with $i=1,\ldots,n$ is the outcome of measurement of A on ith individual system in the subensemble of n systems.

[1] In fact, many important observables, most notably positions and momenta of quantum particles, vary continuously. Such continuously-varying observables will be discussed in Chapter 5.

[2] We will often write $\mathrm{Prob}_E(A=a)$ as $\mathrm{Prob}(A=a)$ if the ensemble (E) is obvious from the context.

[3] Eq. (1.1) expresses the so-called *relative frequency* interpretation of probability. The limit $n\to\infty$ in this equation should be understood in the probabilistic sense: it can be shown that the typical (i.e., the most probable) value of the difference between the ratio $n_{A=a}/n$ and its large-n limit $\mathrm{Prob}(A=a)$ decreases with n as $|n_{A=a}/n - \mathrm{Prob}(A=a)| \sim 1/\sqrt{n}$. Rigorous version of this estimate is known in the probability theory as the *law of large numbers*. Typical deviations of the average value $\overline{A}$ from the expectation value $\langle A\rangle$ [see Eq. (1.3)] scale with n the same way: $|\overline{A} - \langle A\rangle| \sim 1/\sqrt{n}$.

If n is large, the ***average value*** of A given by $\overline{A} = n^{-1}\sum_i a_i$ approaches[3] the ***expectation value***

$$\langle A\rangle_E = \lim_{n\to\infty} \overline{A} = \sum_a a \lim_{n\to\infty} \frac{n_{A=a}}{n} = \sum_a a \,\mathrm{Prob}_E(A=a). \tag{1.3}$$

Generalizing Eq. (1.3), we define the expectation value of an arbitrary function of A as

$$\langle f(A)\rangle_E = \sum_a f(a)\,\mathrm{Prob}_E(A=a). \tag{1.4a}$$

It is clear from this definition that taking the expectation value is a linear operation, e.g.,[4]

$$\langle \alpha f(A) + \beta g(A)\rangle = \alpha\langle f(A)\rangle + \beta\langle g(A)\rangle \tag{1.4b}$$

for all functions $f(A)$ and $g(A)$ and all independent of A coefficients α and β.

Note that, in general, $\langle f(A)\rangle \neq f(\langle A\rangle)$. For example,

$$\langle A^2\rangle = \big\langle \big[\langle A\rangle + \big(A - \langle A\rangle\big)\big]^2\big\rangle = \langle A\rangle^2 + \big\langle \big(A-\langle A\rangle\big)^2\big\rangle \geq \langle A\rangle^2, \tag{1.5a}$$

where $\langle A\rangle$ is treated as an independent of A coefficient, so that it coincides with its expectation value, $\langle\langle A\rangle\rangle = \langle A\rangle$. The ***uncertainty*** (also known as the ***standard deviation***) of A is a non-negative quantity[4]

$$\Delta_E A = \sqrt{\langle A^2\rangle - \langle A\rangle^2} \tag{1.5b}$$

that has the same units as A. It is intuitively clear that the uncertainty characterizes a typical spread of the observed values of A about its expectation value.[5] If the uncertainty is very large, $\Delta A \gg |\langle A\rangle|$, then $\pm\Delta A$ (with a random sign) gives a much better estimate of the observed values of A than $\langle A\rangle$ does.

Suppose now that the number of different outcomes N comprising the set $\{a\}$ is finite. The system of N linear equations[2,4]

$$\langle A^m\rangle = \sum_a a^m\,\mathrm{Prob}(A=a), \quad m = 0,\ldots,N-1 \tag{1.6}$$

yields N probabilities $\mathrm{Prob}(A=a)$ in the form of linear combinations of the ***statistical moments*** $\langle A^m\rangle$ with $m = 1,\ldots,N-1$. Eq. (1.4a) then shows that the expectation value of any function of A can also be written as a linear combination of these moments. These observations suggest that the probabilities can be cast in the form of expectation values. Indeed, consider the function $f(A) = \delta_{a,A}$, where a is one of the possible measurement outcomes. Evaluating its expectation value with the help of Eqs. (1.4), we obtain the relation[2,4]

$$\langle \delta_{a,A}\rangle = \sum_{a'} \delta_{a,a'}\,\mathrm{Prob}(A=a') = \mathrm{Prob}(A=a). \tag{1.7}$$

Note that the validity of Eq. (1.7) does not hinge on the number of possible outcomes being finite.

[4] We will often omit explicit references to the ensemble and write the expectation value $\langle f(A)\rangle_E$ and the uncertainty $\Delta_E A$ as $\langle f(A)\rangle$ and ΔA, respectively.

[5] Probability theory puts this intuition on a solid footing in the form of the *Chebyshev's inequality*.

1.1.2 Sequential measurements and compatible observables

Although individual measurements in experiments aimed at revealing probabilities and expectation values discussed in Sec. 1.1.1 may well be and quite often are destructive, they are carried out on *subensembles*, leaving the reminder of the ensemble undisturbed. This protocol allows one to *infer* the results for the entire ensemble without altering it in any way.

One can conceive and sometimes realize other, more sophisticated, measurement schemes. Consider an observable A that can be measured in a non-destructive and reproducible manner. That is, after A has been measured on individual systems, these systems remain available for further measurements, and subsequent measurements of A are certain to reproduce the earlier results. Specifically, imagine an experiment in which one keeps measuring A on individual systems picked from the initial ensemble E_i, systematically discarding them unless the measurements yield the chosen outcome, say, $A = a$. Such ***filtering*** results in a formation of a new, *filtered*, ensemble E_a:

$$\text{initial ensemble } E_i \xrightarrow[\boldsymbol{A=a}]{\textbf{filtering}} \text{filtered ensemble } E_a . \tag{1.8}$$

The filtering-type measurement scheme just described provides conceptually the simplest recipe for *preparing* ensembles endowed with the desired properties. Indeed, by design, a measurement of A on the filtered ensemble E_a is certain to yield $A = a$, i.e., $\text{Prob}_{E_a}(A = a) = 1$. (Of course, the scheme would work only if the probability that individual systems in E_i pass the filter and end up in E_a, i.e., $\text{Prob}_{E_i}(A = a)$, differs from zero.)

Let B be some other observable, with $\{b\}$ being the corresponding set of possible measurement outcomes. Consider a hypothetical experiment consisting of sequential filtering-type measurements of observables of A and B (in this order!),

$$E_i \xrightarrow[\boldsymbol{A=a}]{\textbf{filtering}} E_a \xrightarrow[\boldsymbol{B=b}]{\textbf{filtering}} E_{ab} . \tag{1.9}$$

The probability that individual systems in the initial ensemble E_i successfully pass both filters and end up in the double-filtered ensemble E_{ab} can be written as

$$\text{Prob}_{E_i}(A = a \text{ then } B = b) = \text{Prob}_{E_i}(A = a)\,\text{Prob}_{E_a}(B = b). \tag{1.10}$$

The two factors in the right-hand side of this expression are the probabilities to clear the first and the second filters, respectively.

Observables A and B are called ***compatible*** if for all possible measurement outcomes a and b and for any ensemble E the probability of getting $A = a$ and $B = b$ in a sequence of two filtering-type measurements is independent of the order in which the filters are applied, i.e.,

$$\text{Prob}_E(A = a \text{ then } B = b) = \text{Prob}_E(B = b \text{ then } A = a). \tag{1.11}$$

Obviously, a measurement of B on the double-filtered ensemble E_{ab} [see Eq. (1.9)] is certain to yield $B = b$, i.e., $\text{Prob}_{E_{ab}}(B = b) = 1$. It is easy to show that if observables A and B are compatible, then E_{ab} also inherits from E_a the property that $A = a$ with certainty:

$$\text{Prob}_{E_{ab}}(A = a) = \text{Prob}_{E_a}(A = a) = 1. \tag{1.12}$$

Indeed, for $E = E_a$ the two probabilities in Eq. (1.11) are given by [cf. Eq. (1.10)]

$$\mathrm{Prob}_{E_a}(A = a \text{ then } B = b) = \mathrm{Prob}_{E_a}(A = a)\,\mathrm{Prob}_{E_a}(B = b) = \mathrm{Prob}_{E_a}(B = b), \tag{1.13a}$$

$$\mathrm{Prob}_{E_a}(B = b \text{ then } A = a) = \mathrm{Prob}_{E_a}(B = b)\,\mathrm{Prob}_{E_{ab}}(A = a). \tag{1.13b}$$

(In writing Eq. (1.13a) we took into account that a measurement of A leaves ensemble E_a intact.) For compatible observables [see Eq. (1.11)], the left-hand sides of Eqs. (1.13) coincide. Equating the right-hand sides, we obtain Eq. (1.12). This result applies to all possible outcomes a and b and to any initial ensemble E_i in Eq. (1.9). Accordingly,

compatible observables have certain values for any ensemble prepared by sequential measurements of these observables. (1.14)

We will see in Sec. 2.4.2 below that for the so-called *ideal* measurements [defined in Sec. 1.3.6] the converse statement holds true as well. That is, if two observables have certain values for any ensemble prepared by sequential ideal measurements of these observables, then these observables are compatible in the sense of Eq. (1.11).

The above consideration of sequential filtering-type measurements can be extended to any number of observables, with the same basic conclusion. By measuring compatible observables on every individual system in the initial ensemble and keeping only the systems for which the measurements produce the desired outcomes, one can prepare a new ensemble for which all subsequent measurements of these observables, in any order, are certain to yield the same outcomes. Moreover, the probability that individual systems in the initial ensemble end up in the final, filtered, ensemble, is independent of the order of measurements.

It is clear on physical grounds that it is always possible to choose a ***complete set of compatible observables*** $\{A, B, C, \ldots\}$ such that individual systems in an ensemble for which all these observables have certain values [e.g., $\mathrm{Prob}\,(A = a) = 1$, $\mathrm{Prob}\,(B = b) = 1$, etc.] are truly identical.[6] Such ensembles of identical individual systems are called ***pure***. On the contrary, ensembles consisting of identically prepared (as opposed to identical) individual systems can be viewed as mixtures of pure subensembles and, quite naturally, are referred to as ***mixed***.

Classical physics treats all observables as compatible. Any observable that is not in the chosen[6] complete set of observables specifying a pure ensemble is a function of the observables in the set, and as such also has a certain value. We consider mixed classical ensembles only because of our inability to prepare the pure ones. In other words, in classical physics the need for a probabilistic description arises due to a lack of detailed knowledge of physical characteristics of the system of interest, which in turn reflects the lack of control.

Quantum physics is fundamentally different: whereas some pairs of observables are compatible, others are manifestly not. Thus, for the ensemble E_a defined in Eq. (1.8) the outcome of measurement of an incompatible with A observable B cannot be predicted with certainty even if E_a is a pure ensemble. This leaves us no choice but to settle for the next best thing available, i.e., probabilities of various outcomes of B. It is crucial that the emergence

[6] It should be emphasized that the choice of a complete set of compatible observables is not unique.

of probabilities in this context cannot be attributed to a lack of knowledge or control as it is the case classically. Rather, it is the consequence of the ***incompatibility of observables***, a distinctive feature of quantum physics.

Note that it would be incorrect to say that there is a clear distinction between classical and quantum systems. Instead, one should think of classical and quantum *phenomena*. Although everything in Nature is quantum, quantum effects are often negligible in the sense that a classical description is perfectly adequate. For example, there is nothing wrong with the classical kinetic theory of gases, yet one has to appeal to quantum theory to explain the very existence and stability of atoms and molecules.

1.1.3 Case study: ensemble of spin 1/2 particles

Here we consider an ensemble of identical particles focusing on their spins (intrinsic angular momenta), and ignoring all other characteristics such as positions and momenta. We assume from the outset that ***(i)*** a complete set of compatible observables [see Sec. 1.1.2] is exhausted by a single observable, the component of the spin vector[7] $\mathbf{S}$ along the direction of an arbitrary unit dimensionless vector $\mathbf{n}$, i.e.,

$$S_{\mathbf{n}} = \mathbf{n} \cdot \mathbf{S}. \tag{1.15}$$

Furthermore, we assume that ***(ii)*** measurements of $S_{\mathbf{n}}$ produce only two possible outcomes, and that ***(iii)*** one of these outcomes is given by $S_{\mathbf{n}} = \hbar/2$. We will show now that these assumptions augmented with fairly obvious symmetry arguments lead to a rather complete description of pure ensembles of spin 1/2 particles.

Because space is isotropic, all directions are equivalent, hence the possible outcomes of measurements of $S_{\mathbf{n}'}$ must be the same for all unit vectors $\mathbf{n}'$, including $\mathbf{n}' = -\mathbf{n}$. Therefore, $S_{-\mathbf{n}}$ may take value $\hbar/2$, which in turn implies that $S_{\mathbf{n}} = -S_{-\mathbf{n}} = -\hbar/2$ is the possible value of $S_{\mathbf{n}}$, different from the value $S_{\mathbf{n}} = \hbar/2$ assumed above. Accordingly, the two possible outcomes of measurements of $S_{\mathbf{n}}$ are[8]

$$S_{\mathbf{n}} = \pm\hbar/2. \tag{1.16}$$

We are interested in the expectation value of $\mathbf{a} \cdot \mathbf{S}$, where $\mathbf{a}$ is an arbitrary reference vector (neither unit nor dimensionless) independent of $\mathbf{S}$. Expanding this vector in an orthonormal Cartesian basis $\{\mathbf{x}, \mathbf{y}, \mathbf{z}\}$ as $\mathbf{a} = \mathbf{x}a_{\mathbf{x}} + \mathbf{y}a_{\mathbf{y}} + \mathbf{z}a_{\mathbf{z}}$ and taking advantage of the linearity of the expectation values [recall Eq. (1.4b)], we obtain

$$\langle \mathbf{a} \cdot \mathbf{S} \rangle = a_{\mathbf{x}} \langle S_{\mathbf{x}} \rangle + a_{\mathbf{y}} \langle S_{\mathbf{y}} \rangle + a_{\mathbf{z}} \langle S_{\mathbf{z}} \rangle, \tag{1.17}$$

irrespective of the ensemble. There is nothing specifically quantum about this relation: it simply reflects the vector nature of $\mathbf{a}$ and $\mathbf{S}$. In particular, it shows that the sum in the right-hand side is independent of the choice of the Cartesian basis $\{\mathbf{x}, \mathbf{y}, \mathbf{z}\}$ because these vectors do not appear in the left-hand side. Nevertheless, the quantum character of spin does

[7] Actually, angular momenta, including spin angular momenta, are *pseudo-vectors*. We do not distinguish between the true (or polar) vectors and the pseudo (or axial) vectors in these Notes. Both types of vectors transform the same way under rotations. (Rotations will be discussed in details in Sec. 3.1.1.)

[8] 1/2 in "spin 1/2" refers to $\max\{S_{\mathbf{n}}\} = \hbar/2$ in units of the ***Planck constant*** $\hbar \approx 10^{-27}\,\mathrm{erg}\cdot\mathrm{s}$.

add an unusual flavor to the interpretation of Eq. (1.17). Indeed, our assumption (i) implies that components of vector $\mathbf{S}$ along different directions are mutually incompatible observables. (This, of course, would not be so if $\mathbf{S}$ were a classical angular momentum.) Therefore, the expectation values of $\mathbf{a}\cdot\mathbf{S}$, $S_{\mathbf{x}}$, $S_{\mathbf{y}}$, and $S_{\mathbf{z}}$ cannot be determined by means of non-destructive sequential measurements. Instead, as discussed in the beginning of Sec. 1.1.2, these expectation values are inferred by measuring these observables on four separate subensembles. In this sense, Eq. (1.17) relates the results of four different experiments.

It is often convenient to write Eq. (1.17) as

$$\langle \mathbf{a}\cdot\mathbf{S}\rangle = \mathbf{a}\cdot\langle\mathbf{S}\rangle, \tag{1.18a}$$

where we have introduced vector

$$\langle\mathbf{S}\rangle = \mathbf{x}\langle S_{\mathbf{x}}\rangle + \mathbf{y}\langle S_{\mathbf{y}}\rangle + \mathbf{z}\langle S_{\mathbf{z}}\rangle \tag{1.18b}$$

that depends neither on $\mathbf{a}$ nor on the choice of the Cartesian basis $\{\mathbf{x},\mathbf{y},\mathbf{z}\}$. Similar to the right-hand side of Eq. (1.17), this vector is somewhat unusual: its Cartesian components coincide with the expectation values of the corresponding components of $\mathbf{S}$, which can only be found by measurements on separate subensembles.

Consider now an ensemble of spin 1/2 particles for which measurements of $S_{\mathbf{n}}$ are certain to yield $+\hbar/2$. By assumption (i) it is a pure ensemble uniquely specified by the unit dimensionless vector $\mathbf{n}$, the so-called ***Bloch vector***.[9] We will denote this ensemble by the subscript $\mathbf{n}$ as in the ensemble-defining relation

$$\mathrm{Prob}_{\mathbf{n}}(S_{\mathbf{n}} = \hbar/2) = 1. \tag{1.19}$$

Vector $\langle\mathbf{S}\rangle$ introduced in Eqs. (1.18) depends only on the properties of the ensemble. Since the Bloch vector $\mathbf{n}$ is the only ensemble-characterizing vector at our disposal, $\langle\mathbf{S}\rangle$ for this ensemble must be proportional to it: $\langle\mathbf{S}\rangle_{\mathbf{n}} \propto \mathbf{n}$. The proportionality coefficient can be found by setting $\mathbf{a} = \mathbf{n}$ in Eq. (1.18a) and taking into account the relation $\langle\mathbf{n}\cdot\mathbf{S}\rangle_{\mathbf{n}} = \langle S_{\mathbf{n}}\rangle_{\mathbf{n}} = \hbar/2$ implied by Eq. (1.19). The resulting expression

$$\langle\mathbf{S}\rangle_{\mathbf{n}} = (\hbar/2)\,\mathbf{n} \tag{1.20}$$

applies to all pure ensembles of spin 1/2 particles.[9]

Because components of $\mathbf{S}$ take on only two possible values [see Eq. (1.16)], the corresponding probabilities can be extracted from the expectation values. Eqs. (1.18a) and (1.20) show that

$$\langle S_{\mathbf{n}_1}\rangle_{\mathbf{n}_2} = \langle S_{\mathbf{n}_2}\rangle_{\mathbf{n}_1} = (\hbar/2)(\mathbf{n}_1\cdot\mathbf{n}_2) \tag{1.21}$$

for all unit dimensionless vectors $\mathbf{n}_1$ and $\mathbf{n}_2$. Proceeding as suggested in Sec. 1.1.1 [see Eq. (1.6)], we obtain

$$\mathrm{Prob}_{\mathbf{n}_1}(S_{\mathbf{n}_2} = \hbar/2) = \mathrm{Prob}_{\mathbf{n}_2}(S_{\mathbf{n}_1} = \hbar/2) = \frac{1}{2}(1 + \mathbf{n}_1\cdot\mathbf{n}_2). \tag{1.22}$$

[9] We will see in Sec. 1.3.3 below that for any pure ensemble of spin 1/2 particles there exists a unique Bloch vector $\mathbf{n}$ such that $S_{\mathbf{n}} = \hbar/2$ with probability 1 [cf. Eq. (1.19)]. In other words, the correspondence between the pure ensembles and the Bloch vectors is one-to-one.

The probability of getting, say, $S_{\mathbf{n}_2} = -\hbar/2$ for the ensemble specified by the Bloch vector $\mathbf{n}_1$ is obtained by replacing $\mathbf{n}_2$ in Eq. (1.22) with $-\mathbf{n}_2$. (Indeed, $S_{\mathbf{n}_2} = -\hbar/2$ is equivalent to $S_{-\mathbf{n}_2} = +\hbar/2$.) Notice that, as expected on symmetry grounds, the expectation values (1.21) and thus the probabilities (1.22) depend only on the angle between vectors $\mathbf{n}_1$ and $\mathbf{n}_2$ and are therefore symmetric with respect to the interchange $\mathbf{n}_1 \rightleftarrows \mathbf{n}_2$.

With Eqs. (1.20)-(1.22) at hand, evaluation of various probabilities and expectation values is straightforward. Our derivation of these manifestly non-classical expressions relied on the assumptions made in the beginning of the section. For ensembles of independent spin 1/2 particles, these assumptions can be justified by appealing to the results of Stern-Gerlach experiments. We will see below that all three assumptions emerge naturally from the rules of quantum theory.

1.1.4 Case study: Stern-Gerlach experiment with spin 1/2 particles

Stern-Gerlach-type experiments are performed on *beams* of particles. Of course, a beam is not an ensemble but a single interacting many-particle system. It is plausible, however, that if the density of particles in the beam is sufficiently low, interactions between them are so weak that their effect can be neglected. Such ultra-low-density beam of particles can be viewed as a physical realization of an ensemble of independent particles and treated accordingly. (This turns out to be an excellent approximation under the typical conditions of the Stern-Gerlach-type experiments.)

Passing a beam of particles (henceforth synonymous with an ensemble) through a Stern-Gerlach filter blocking particles with $S_{\mathbf{n}} = -\hbar/2$ produces a filtered beam with $S_{\mathbf{n}} = \hbar/2$. This new beam can be viewed as a pure ensemble, hence all subsequent measurements can be analyzed with the help of Eq. (1.22). However, the initial beam is clearly not pure. It is therefore natural to ask what is the probability that particles in the initial *mixed* beam pass the very first filter, ending up in the filtered (hence "purified") beam.

To get a flavor of the considerations involved, we assume that the initial ensemble E_i is a *mixture* of two pure subensembles $E_\pm$ such that

$$\mathrm{Prob}_{E_\pm}(S_{\mathbf{n}_i} = \pm\hbar/2) = 1, \tag{1.23}$$

where $\mathbf{n}_i$ is an arbitrary unit dimensionless vector. (Note that since $\mathrm{Prob}_{E_\pm}(S_{\mathbf{n}_i} = \mp\hbar/2) = 0$, these subensembles do not overlap: an individual particle cannot simultaneously belong to both E_+ and E_-.) The probability we seek is then a sum of two contributions,

$$\mathrm{Prob}_{E_i}(S_{\mathbf{n}} = \hbar/2) = p_+ \,\mathrm{Prob}_{E_+}(S_{\mathbf{n}} = \hbar/2) + p_- \,\mathrm{Prob}_{E_-}(S_{\mathbf{n}} = \hbar/2). \tag{1.24a}$$

Here $p_\pm$ is the fraction of the total number of particles in E_i that belong to $E_\pm$ and

$$\mathrm{Prob}_{E_\pm}(S_{\mathbf{n}} = \hbar/2) = \frac{1}{2}\big(1 \pm \mathbf{n}\cdot\mathbf{n}_i\big) \tag{1.24b}$$

[see Eq. (1.22)] is the probability that particles in $E_\pm$ pass the filter. Taking into account that $p_+ + p_- = 1$, we obtain

$$\mathrm{Prob}_{E_i}(S_{\mathbf{n}} = \hbar/2) = \frac{1}{2} + \frac{1}{2}(p_+ - p_-)(\mathbf{n}\cdot\mathbf{n}_i). \tag{1.24c}$$

A natural requirement that the probability is independent of the choice of $\mathbf{n}_i$ (recall that $\mathbf{n}_i$ is arbitrary!) leads to $p_+ = p_-$ and

$$\mathrm{Prob}_{E_i}(S_{\mathbf{n}} = \hbar/2) = 1/2, \tag{1.25}$$

as could have been anticipated on symmetry grounds.

There exists a well-developed (and rather elegant) formalism for dealing with mixed ensembles. This formalism does not appear to be as contrived as the calculation outlined above, but the basic idea remains essentially the same: one breaks the mixed ensemble in non-overlapping pure subensembles and requires the results to be independent of the choice of these subensembles. Justification of this approach is well beyond the scope of these Notes.

1.2 Vector spaces

As explained above, the goal of quantum theory is to predict probabilities of outcomes of measurements made on ensembles of identically prepared copies of the system of interest. Like all physical theories, quantum theory consists of a collection of basic rules, and consequences that follow from them. In order to formulate these rules in the most efficient and transparent manner, one needs to be fluent in the proper language (mathematical framework). For quantum theory, this language turns out to be that of linear vector spaces.

1.2.1 Linear vector space

A ***linear vector space*** is a generalization of the familiar three-dimensional space of the ordinary (geometric) vectors. By definition, it is a set of elements called ***ket-vectors*** or simply ***vectors*** (and denoted, following Dirac, by the symbols $|\psi\rangle$) equipped with operations of ***addition*** and ***scalar multiplication***. The addition operation specifies for any pair of vectors in the space a unique third vector, their sum, whereas the scalar multiplication operation associates a complex number (a scalar) and a vector with another vector:[10]

$$|\varphi\rangle, |\psi\rangle \xrightarrow{\textbf{addition}} |\varphi + \psi\rangle, \tag{1.26a}$$

$$c\,, |\psi\rangle \xrightarrow{\textbf{scalar multiplication}} |c\psi\rangle. \tag{1.26b}$$

These operations are postulated to have the following "natural" properties:

$$|\varphi\rangle + |\psi\rangle = |\psi\rangle + |\varphi\rangle \quad \text{(commutativity of addition)}, \tag{1.27a}$$

$$|\varphi + \psi\rangle + |\chi\rangle = |\varphi\rangle + |\psi + \chi\rangle \quad \text{(associativity of addition)}, \tag{1.27b}$$

$$|\psi\rangle + |\text{null}\rangle = |\psi\rangle \quad \text{(existence of the } \textbf{\textit{null}} \text{ vector)}, \tag{1.27c}$$

$$a|\varphi + \psi\rangle = a|\varphi\rangle + a|\psi\rangle \quad \text{(distributivity in vectors)}, \tag{1.27d}$$

$$(a + b)|\psi\rangle = a|\psi\rangle + b|\psi\rangle \quad \text{(distributivity in scalars)}, \tag{1.27e}$$

$$a|b\,\psi\rangle = (ab)|\psi\rangle \quad \text{(associativity of the scalar multiplication)}, \tag{1.27f}$$

$$1|\psi\rangle = |\psi\rangle \quad \text{(how else?)} \tag{1.27g}$$

[10] We will often write $|\varphi + \psi\rangle$ as $|\varphi\rangle + |\psi\rangle$ and $|c\psi\rangle$ as $c|\psi\rangle$ or $|\psi\rangle c$. It is also convenient (and conventional) to write $(-1)|\psi\rangle$ as $|-\psi\rangle$ or $-|\psi\rangle$ as in $|\varphi - \psi\rangle = |\varphi\rangle + |-\psi\rangle = |\varphi\rangle - |\psi\rangle$.

for all vectors $|\varphi\rangle$, $|\psi\rangle$, and $|\chi\rangle$ and all complex numbers a and b. It is easy to show that vector $|\text{null}\rangle$ introduced in Eq. (1.27c) is unique and that for all $|\psi\rangle$ and c

$$0|\psi\rangle = |\text{null}\rangle, \quad c|\text{null}\rangle = |\text{null}\rangle. \tag{1.28}$$

A set of N vectors $\{|\phi_n\rangle; n = 1, \ldots, N\}$ is said to be ***linearly independent*** if the equation

$$\sum_{n=1}^{n=N} c_n|\phi_n\rangle = |\text{null}\rangle \tag{1.29}$$

has no solutions for the coefficients c_n other than that in which all c_n are zero. The largest possible number of vectors a linearly independent set can accommodate, $\mathcal{N} = \max\{N\}$, is the ***dimension*** of the space. This implies that any linearly independent set $\{|\phi_n\rangle\}$ of $\mathcal{N}$ vectors[11] serves as a ***basis*** in the sense that it ***spans*** the whole space: any vector $|\psi\rangle$ in it can be written as a linear combination of the basis vectors,

$$|\psi\rangle = \sum_{n=1}^{n=\mathcal{N}} \psi_n|\phi_n\rangle. \tag{1.30}$$

The coefficients ψ_n in this expansion are ***components*** of vector $|\psi\rangle$ in the basis $\{|\phi_n\rangle\}$. It is easy to see that the components are uniquely determined by the choice of the basis. Moreover, Eqs. (1.27) and (1.30) show that[10]

$$|\varphi + \psi\rangle = \sum_{n=1}^{n=\mathcal{N}} (\varphi_n + \psi_n)|\phi_n\rangle, \quad |c\psi\rangle = \sum_{n=1}^{n=\mathcal{N}} (c\psi_n)|\phi_n\rangle. \tag{1.31}$$

In other words, both the addition and the scalar multiplication operations [recall Eqs. (1.26)] can be carried out component-wise.

1.2.2 Inner product

A linear vector space equipped with an ***inner product*** is called an ***inner-product space***. Similar to the dot product of ordinary vectors, the inner product is a rule assigning a unique *complex* number to any pair of vectors in the space:[12]

$$|\varphi\rangle, |\psi\rangle \xrightarrow{\textbf{inner product}} \text{complex number } \langle\varphi|\psi\rangle. \tag{1.32}$$

It is postulated that

$$\langle\varphi|\psi\rangle = \langle\psi|\varphi\rangle^* \quad (\text{skew-symmetry: } \textbf{the order matters!}), \tag{1.33a}$$

$$\langle\varphi|\psi + \chi\rangle = \langle\varphi|\psi\rangle + \langle\varphi|\chi\rangle \quad (\text{distributivity in kets}), \tag{1.33b}$$

$$\langle\varphi|c\psi\rangle = c\langle\varphi|\psi\rangle \quad (\text{lnearity in kets}), \tag{1.33c}$$

$$\langle\psi|\psi\rangle > 0 \text{ for all } |\psi\rangle \neq |\text{null}\rangle. \tag{1.33d}$$

[11] For the time being, we take the dimension of the space to be finite: $\mathcal{N} < \infty$. Infinite-dimensional spaces are discussed in Sec. 1.2.4 below.

[12] Following Dirac, we say that in $\langle\varphi|\psi\rangle$ the ***ket-vector*** $|\psi\rangle$ is multiplied by the ***bra-vector*** $\langle\varphi|$, indicating the order in which φ and ψ appear in the inner product. Although this language is often used in these notes, the point of view adopted here is that **all vectors are *kets*, there are no stand-alone *bras***. However, it is possible (albeit not necessary for our purposes) to discuss the *bra*-vectors in a mathematically sound manner.

Eqs. (1.33a) and (1.33b) show that $\langle\varphi+\chi|\psi\rangle = \langle\psi|\varphi+\chi\rangle^* = \langle\psi|\varphi\rangle^* + \langle\psi|\chi\rangle^*$, hence

$$\langle\varphi+\chi|\psi\rangle = \langle\varphi|\psi\rangle + \langle\chi|\psi\rangle. \tag{1.34a}$$

Similarly, Eqs. (1.33a) and (1.33c) give

$$\langle a\varphi|b\psi\rangle = a^* b\langle\varphi|\psi\rangle. \tag{1.34b}$$

Thus, the inner product is *linear* in kets and *antilinear* in bras,[12] e.g.,

$$\langle a_1\varphi_1 + a_2\varphi_2|b_1\psi_1 + b_2\psi_2\rangle = a_1^* b_1\langle\varphi_1|\psi_1\rangle + a_1^* b_2\langle\varphi_1|\psi_2\rangle + a_2^* b_1\langle\varphi_2|\psi_1\rangle + a_2^* b_2\langle\varphi_2|\psi_2\rangle.$$

According to Eq. (1.33a), $\langle\psi|\psi\rangle = \langle\psi|\psi\rangle^*$, i.e., $\langle\psi|\psi\rangle$ is a real number. Whereas this number is strictly positive for $|\psi\rangle \neq |\text{null}\rangle$ [see Eq. (1.33d)], it vanishes for $|\psi\rangle = |\text{null}\rangle$:

$$\langle\text{null}|\text{null}\rangle = 0. \tag{1.35}$$

Indeed, setting $c = 0$ in Eq. (1.33c) and taking into account the first equation in (1.28), we find $\langle\varphi|\text{null}\rangle = 0$ for all $|\varphi\rangle$, including $|\varphi\rangle = |\text{null}\rangle$ as in Eq. (1.35).

A very useful property of the inner product is the ***Schwartz inequality***[13]

$$\big|\langle\varphi|\psi\rangle\big|^2 \leq \langle\varphi|\varphi\rangle\langle\psi|\psi\rangle, \tag{1.36}$$

valid for all vectors $|\varphi\rangle$ and $|\psi\rangle$. The inequality is trivially satisfied when one of these vectors is $|\text{null}\rangle$, so that both sides of Eq. (1.36) turn to zero. If both $|\varphi\rangle$ and $|\psi\rangle$ differ from $|\text{null}\rangle$, one can write $|\psi\rangle$ as

$$|\psi\rangle = c|\varphi\rangle + |\psi'\rangle, \tag{1.37a}$$

where vector $|\psi'\rangle$ is orthogonal to $|\varphi\rangle$, i.e., such that $\langle\varphi|\psi'\rangle = 0$. (Two vectors are called ***orthogonal*** if their inner product is zero.) Taking a square, we obtain

$$\langle\psi|\psi\rangle = |c|^2\langle\psi|\psi\rangle + \langle\psi'|\psi'\rangle, \tag{1.37b}$$

which implies the inequality

$$\langle\psi|\psi\rangle \geq |c|^2\langle\psi|\psi\rangle. \tag{1.37c}$$

The coefficient c is found by multiplying[12] both sides of Eq. (1.37a) by $\langle\varphi|$,

$$c = \frac{\langle\varphi|\psi\rangle}{\langle\varphi|\varphi\rangle}. \tag{1.37d}$$

Substitution of this value into Eq. (1.37c) yields the Schwartz inequality (1.36). Obviously, the inequality turns to an equality if and only if $|\psi'\rangle = |\text{null}\rangle$ in Eq. (1.37a), i.e., when vectors $|\varphi\rangle$ and $|\psi\rangle$ are linearly dependent.

[13] The inequalities (1.36) and (1.39c) are analogs of the corresponding relations for the ordinary (geometric) vectors. Indeed, the dot product of vectors $\mathbf{a}$ and $\mathbf{b}$ reads $\mathbf{a}\cdot\mathbf{b} = ab\cos\theta$, where $a = |\mathbf{a}|$, $b = |\mathbf{b}|$,[14] and θ is the angle between $\mathbf{a}$ and $\mathbf{b}$ ($0 \leq \theta \leq \pi$). Because $|\cos\theta| \leq 1$, we have $(\mathbf{a}\cdot\mathbf{b})^2 = a^2b^2\cos^2\theta \leq a^2b^2$ [cf. Eq. (1.36)]. Similarly, $(\mathbf{a}+\mathbf{b})^2 = a^2 + b^2 + 2(\mathbf{a}\cdot\mathbf{b}) \leq a^2 + b^2 + 2ab = (a+b)^2$, hence $|\mathbf{a}+\mathbf{b}| \leq a + b$ [cf. Eq. (1.39c)].

1.2.3 Norm and orthonormal basis

Ordinary vectors have *length*. The analog of length for linear vector spaces is a rule assigning to every vector in the space a unique non-negative real number, the ***norm***,

$$|\psi\rangle \xrightarrow{\textbf{norm}} \text{real number } \|\psi\|. \tag{1.38}$$

The norm is required to satisfy

$$\|\psi\| > 0 \ \text{ for all } \ |\psi\rangle \neq |\text{null}\rangle, \tag{1.39a}$$

$$\|c\psi\| = |c|\,\|\psi\|, \tag{1.39b}$$

$$\|\varphi + \psi\| \leq \|\varphi\| + \|\psi\|. \tag{1.39c}$$

The last relation here is known as the ***triangle inequality***.[13] Setting $c = 0$ in Eq. (1.39b) and taking into account the first equation in (1.28), we obtain $\|\text{null}\| = 0$. Eq. (1.39a) then shows that the only vector with vanishing norm is the null vector, i.e.,

$$\|\psi\| = 0 \ \text{ implies } \ |\psi\rangle = |\text{null}\rangle. \tag{1.40}$$

For an inner-product space [see Sec. 1.2.2] one can introduce the norm by mimicking the corresponding expression for ordinary vectors,[14,15]

$$\|\psi\| = \sqrt{\langle\psi|\psi\rangle}. \tag{1.41}$$

It is easy to verify that properties of the inner product ensure that the norm defined by Eq. (1.41) indeed satisfies the requirements (1.39).

A vector is ***normalized*** if it has a unit norm. A set of mutually orthogonal normalized vectors is called ***orthonormal***. It is easy to see that any such orthonormal set is automatically linearly independent. Therefore, any orthonormal set of $\mathcal{N}$ vectors

$$\{|\phi_n\rangle;\, n = 1, \ldots, \mathcal{N}\}, \quad \langle\phi_m|\phi_n\rangle = \delta_{m,n} \tag{1.42}$$

serves as an ***orthonormal basis*** for the $\mathcal{N}$-dimensional space.[11] It is always possible to construct such basis,[16] and there is an infinite number of orthonormal basis sets at one's disposal.

---

[14] Recall that the length of an ordinary vector $\mathbf{a}$ is given by $a = |\mathbf{a}| = \sqrt{\mathbf{a}\cdot\mathbf{a}}$.

[15] Although the norm can be introduced independently of the inner product, it is the inner-product-based definition that happens to be relevant for quantum theory.

[16] Given an arbitrary linearly independent set $\{|\varphi_n\rangle\}$, we construct a new set $\{|\phi_n\rangle\}$ such that

$$|\phi_n\rangle = \frac{1}{\|\psi_n\|}|\psi_n\rangle \ \text{ with } \begin{cases} |\psi_1\rangle = |\varphi_1\rangle, \\ |\psi_2\rangle = |\varphi_2\rangle - |\phi_1\rangle\langle\phi_1|\varphi_2\rangle, \\ |\psi_3\rangle = |\varphi_3\rangle - |\phi_1\rangle\langle\phi_1|\varphi_3\rangle - |\phi_2\rangle\langle\phi_2|\varphi_3\rangle, \\ \quad\cdots \qquad \cdots \\ |\psi_n\rangle = |\varphi_n\rangle - \sum_{m<n}|\phi_m\rangle\langle\phi_m|\varphi_n\rangle. \end{cases}$$

These vectors obviously satisfy $\langle\phi_m|\phi_n\rangle = \delta_{m,n}$, i.e., the set $\{|\phi_n\rangle\}$ is orthonormal. (This method of constructing an orthonormal set is known as the ***Gramm-Schmidt orthogonalization***.)

Multiplying[12] both sides of $|\psi\rangle = \sum_m \psi_m |\phi_m\rangle$ [cf. Eq. (1.30)] by the basis vector $\langle\phi_n|$ and using the second equation in (1.42), we find

$$\langle\phi_n|\psi\rangle = \sum_{m=1}^{m=\mathcal{N}} \psi_m \underbrace{\langle\phi_n|\phi_m\rangle}_{\delta_{n,m}} = \psi_n .$$

Accordingly, when vector $|\psi\rangle$ is expanded in an orthonormal basis such as (1.42), its components ψ_n are given by the inner products of the basis vectors and $|\psi\rangle$:

$$|\psi\rangle = \sum_{n=1}^{n=\mathcal{N}} \psi_n |\phi_n\rangle, \quad \psi_n = \langle\phi_n|\psi\rangle . \tag{1.43}$$

In addition to the ease of finding components of vectors, working in an orthonormal basis has an advantage that inner products can be written in terms of components as

$$\langle\varphi|\psi\rangle = \sum_{n=1}^{n=\mathcal{N}} \varphi_n^* \psi_n , \tag{1.44a}$$

which gives

$$\|\psi\|^2 = \langle\psi|\psi\rangle = \sum_{n=1}^{n=\mathcal{N}} |\psi_n|^2 \tag{1.44b}$$

for the square of the norm of vector $|\psi\rangle$ [see Eq. (1.41)]. Notice that the sums in the right-hand sides of Eqs. (1.44) consist of only $\mathcal{N}$ terms instead of $\mathcal{N}^2$ as it would be the case for a generic *non*-orthonormal basis.

1.2.4 Hilbert space

Quantum theory deals with vector spaces of special kind, the so-called **Hilbert spaces**. Depending on the problem at hand, the dimension of the pertinent Hilbert space can be either finite or infinite. A finite-dimensional Hilbert space is synonymous with an inner-product space. As discussed in Sec. 1.2.3, it is always possible to introduce an orthonormal basis for such space and expand any vector in the space in this basis [see Eq. (1.43)]. Various manipulations with vectors then conveniently reduce to manipulations with their components [see Eqs. (1.31) and (1.44)].

For $\mathcal{N} = \infty$, the right-hand side of Eq. (1.43) turns to an infinite series, and we face the problem of giving a precise meaning to these series. The Hilbert space concept was introduced to circumvent this difficulty: it turns out that under certain conditions[17] infinite-dimensional inner-product spaces can be dealt with as if they were finite-dimensional. If these conditions are met, expansions of vectors in complete[18] orthonormal sets, such as the

[17] A Hilbert space is formally defined as such inner-product space that every converging sequence[19] of vectors in it has a limit vector[19] that belongs to the space. This definition applies to both finite- and infinite-dimensional spaces.

[18] An orthonormal set of vectors $\{|\phi_n\rangle\}$ is said to be ***complete*** if it is a basis for the space. (Note that in an infinitely-dimensional space having an infinite number of vectors in an orthonormal set is not sufficient to guarantee its completeness.)

infinite-$\mathcal{N}$ versions of the vector series in Eqs. (1.31) and (1.43), are guaranteed to converge,[19] whereas the ordinary (non-vector) series in Eqs. (1.44) converge absolutely.

We will not delve into delicate mathematics of infinite-dimensional Hilbert spaces in these Notes. Instead, we will treat such spaces the physicist's way, i.e., simply by assuming that

all series are convergent unless proven guilty.

There is, however, one property of the Hilbert spaces that is worth mentioning as we will need it later: it can be shown that such spaces include only vectors with a finite norm, i.e., such that $\|\psi\| < \infty$.

1.3 Rules of quantum theory

Here we begin the discussion of the rules upon which the edifice of quantum theory is built. We formulate the rules for the simplest possible setting first, and then modify them to accommodate more complicated situations.

1.3.1 What is a quantum state?

At the heart of quantum theory lies the concept of a ***quantum state***. In classical physics, it makes sense to think and talk about the state of an individual system. Such classical state is uniquely specified by the values of observables $\{A, B, C, \ldots\}$ forming a complete set [see Sec. 1.1.2], e.g., $A = a$, $B = b$, etc. (Recall that in classical physics all observables are compatible.) It is perfectly legitimate and logical to assume that an individual classical system actually *possesses* these values, whereas the role of measurements reduces to merely *revealing* them to an observer.

In quantum theory, the word *state* has a totally different meaning. Indeed, quantum theory does not deal with individual systems. Rather, it makes probabilistic predictions for the outcomes of measurements made on *ensembles* of identically prepared copies of the system of interest. An individual quantum system does not have a state and whether or not it possesses definite values of observables is irrelevant as far as the conventional quantum theory discussed here is concerned. Thus, specifying the state of the system in quantum theory (a *quantum state*) is synonymous with specifying the ensemble to which the system belongs.

A quantum system is said to be in a ***pure state***[20] if the corresponding ensemble is pure [see Sec. 1.1.2]. Conversely, a ***mixed state*** refers to the mixed ensemble. For pure states,[21] the maximal amount of statistical information is at one's disposal in the sense that there

[19] An infinite sequence of vectors $\{|\psi_n\rangle; n = 1, 2, \ldots\}$ ***converges*** if for any $\epsilon > 0$ there exists such M_ϵ that $\|\psi_m - \psi_n\| < \epsilon$ for all $m, n > M_\epsilon$. The sequence converges to the ***limit vector*** $|\psi\rangle$ if $\|\psi - \psi_n\| \xrightarrow{n\to\infty} 0$, i.e., if for any $\epsilon > 0$ there exists such N_ϵ that $\|\psi - \psi_n\| < \epsilon$ for all $n > N_\epsilon$. Similarly, an infinite series $\sum_{n=1}^{n=\infty} |\psi_n\rangle$ converges to the limit vector $|\psi\rangle$ if the sequence of the finite sums $\{|\Psi_n\rangle = \sum_{m=1}^{m=n} |\psi_m\rangle\}$ converges to $|\psi\rangle$.

[20] If you find the phrase "the system is in such-and-such state" confusing, you are not alone. In the early days of quantum theory, no distinction was made between classical and quantum states. As quantum theory evolved, so did the concept of the state, and the word *state* gradually came to mean an *ensemble*. Sadly, as it often happens in physics, the language did not change.

[21] Unless explicitly stated otherwise, hereinafter we consider only pure quantum states.

exists such complete set of mutually compatible observables [see Sec. 1.1.2] that each of these observables has a certain value.[6]

1.3.2 Introducing the rules

The first rule of quantum theory sets in place the mathematical framework. It turns out that every quantum system can be associated with a certain Hilbert space $\mathcal{H}$ (the *space of states*) such that every possible quantum state[21] is represented by a normalized vector in $\mathcal{H}$ and that every such ***state vector*** represents a possible quantum state:[22]

$$\textbf{Rule I:} \quad \begin{array}{c} \textbf{quantum system} \longrightarrow \textbf{Hilbert space } \boldsymbol{\mathcal{H}}; \\ \textbf{quantum state } \boldsymbol{\psi} \rightleftarrows \textbf{normalized vector } \boldsymbol{|\psi\rangle} \textbf{ in } \boldsymbol{\mathcal{H}}. \end{array} \tag{1.45}$$

Built into Rule I is the notion of the ***superposition***: if $|\psi_1\rangle$ and $|\psi_2\rangle$ are two state vectors, then any normalized linear combination of these vectors, i.e., any vector $|\psi\rangle = c_1|\psi_1\rangle + c_2|\psi_2\rangle$ that satisfies $\||\psi\rangle\| = 1$, also represents a quantum state.[21] It is the ease of handling superpositions of states that makes the formalism of Hilbert spaces so well-suited for quantum theory. In fact, Rule I is often referred to as the ***superposition principle***.

Consider now observable A with a discrete set of possible measurement outcomes $\{a\}$. The second rule proclaims that for each a in this set it is possible (at least, in principle) to prepare a quantum state[21] ϕ_a for which a measurement of A is certain to yield $A = a$:

$$\textbf{Rule II:} \quad \textbf{outcome } \boldsymbol{a} \longrightarrow \textbf{state } \boldsymbol{\phi_a} \textbf{ such that } \textbf{Prob}_{\boldsymbol{\phi_a}}(\boldsymbol{A = a}) = \boldsymbol{1}. \tag{1.46a}$$

By Rule I, any such state corresponds to a state vector:

$$\textbf{state } \boldsymbol{\phi_a} \longrightarrow \textbf{state vector } \boldsymbol{|\phi_a\rangle}. \tag{1.46b}$$

For simplicity, we assume in this section that there is only one linearly independent state vector $|\phi_a\rangle$ for each a. In other words, state vectors $|\phi_a\rangle$ with the same a may differ in overall phase factors at most; such outcomes are called ***non-degenerate***. The physical meaning of this restriction is elucidated shortly below, and it will be lifted in Sec. 1.3.4.

Unless the system is in one of the states[20] ϕ_a whose existence is postulated in Rule II, predicting the outcome of measurement of A with certainty is impossible. The third rule (known as the ***Born's rule***) provides a recipe for evaluating probabilities of various measurement outcomes for an arbitrary quantum state ψ:[21]

$$\textbf{Rule III:} \quad \textbf{Prob}_{\boldsymbol{\psi}}(\boldsymbol{A = a}) = \boldsymbol{|\langle \phi_a | \psi \rangle|^2}. \tag{1.47}$$

It is obvious from Eq. (1.47) that state vectors that differ only in overall phase factors result in the same probabilities and thus are physically indistinguishable,[23] i.e.,

$$\textbf{state vectors } \boldsymbol{|\psi\rangle} \textbf{ and } \boldsymbol{e^{i\alpha}|\psi\rangle} \textbf{ represent the same quantum state } \boldsymbol{\psi}. \tag{1.48}$$

[22] It is very common to use "**state $|\psi\rangle$**" as a shorthand for "**state vector $|\psi\rangle$ representing pure quantum state ψ**". One should keep in mind, however, that in view of Eq. (1.48) the "quantum state" is not synonymous with the "state vector".

[23] A mathematician would say that the mapping {states} $\longmapsto$ {state vectors} in Eq. (1.45) is *onto*, but not *one-to-one*. It will be shown in Sec. 2.3.1 that it is possible to formulate Rule I in such a way that this ambiguity does not arise.

The phase factors do not show up in the final results provided that these results are formulated in terms of experimentally relevant quantities such as probabilities and expectation values.

This observation also sheds light upon the physical meaning of the assumption of the absence of degeneracy inherent in Eq. (1.47). Indeed, if state vectors $|\phi_a\rangle$ corresponding to the outcome a [see Eq. (1.46b)] differ in phase factors only, these vectors represent the same state ϕ_a uniquely specified by a. If all possible measurement outcomes in the set $\{a\}$ are non-degenerate, then a single observable A suffices to discriminate between various quantum states, thereby exhausting the complete set of compatible observables [see Sec. 1.1.2].

The probabilities (1.47) must obey Eqs. (1.2). The first equation in (1.2) is satisfied automatically: taking into account that state vectors are normalized [see Eq. (1.45)] and using the Schwartz inequality (1.36), we find $0 \le |\langle\phi_a|\psi\rangle|^2 \le \|\phi_a\|^2\|\psi\|^2 = 1$, hence, as expected, $0 \le \mathrm{Prob}_\psi(A=a) \le 1$. On the contrary, the necessity to satisfy the second equation in (1.2) imposes rather stringent conditions on the state vectors $|\phi_a\rangle$ introduced in Rule II. First of all, Eqs. (1.46) and (1.47) and the obvious relation $\mathrm{Prob}_{\phi_{a'}}(A=a) = \delta_{a,a'}$ imply that state vectors corresponding to different outcomes must be mutually orthogonal,

$$\langle\phi_a|\phi_{a'}\rangle = \delta_{a,a'}. \tag{1.49}$$

Moreover, the orthonormal set $\{|\phi_a\rangle\}$ formed by picking one state vector $|\phi_a\rangle$ for each a turns out to be complete[18] in the sense that it serves as a basis for the Hilbert space $\mathcal{H}$. Indeed, any $|\psi\rangle$ in $\mathcal{H}$ can be written as

$$|\psi\rangle = |\psi_A\rangle + |\psi'\rangle, \quad |\psi_A\rangle = \sum_a \psi_a|\phi_a\rangle, \tag{1.50}$$

where $|\psi'\rangle$ is orthogonal to all vectors in the set $\{|\phi_a\rangle\}$ and thus to their linear combination $|\psi_A\rangle$; it is easy to show[24] that the decomposition $|\psi\rangle = |\psi_A\rangle + |\psi'\rangle$ is unique. If $|\psi'\rangle \ne |\mathrm{null}\rangle$ then vector $|\varphi'\rangle = \|\psi'\|^{-1}|\psi'\rangle$ is normalized and thus, by Rule I, represents a legitimate quantum state φ'. But for this state vector $\langle\phi_a|\varphi'\rangle = 0$, and Rule III gives $\mathrm{Prob}_{\varphi'}(A=a) = 0$ for all a, in contradiction with the second equation in (1.2). Therefore, consistency of our rules requires that $|\psi'\rangle = |\mathrm{null}\rangle$ in Eq. (1.50). Since $|\psi\rangle$ is arbitrary, this implies that the orthonormal set $\{|\phi_a\rangle\}$ spans the entire Hilbert space, i.e., is a basis.

Substituting the expansion of the state vector $|\psi\rangle$ in this basis

$$|\psi\rangle = \sum_a \psi_a|\phi_a\rangle, \quad \psi_a = \langle\phi_a|\psi\rangle \tag{1.51a}$$

[cf. Eq. (1.43)] into Eq. (1.47) shows that the probabilities coincide with the squares of the absolute values of the components (expansion coefficients) ψ_a,

$$\mathrm{Prob}_\psi(A=a) = |\psi_a|^2, \tag{1.51b}$$

[24] Consider two decompositions $|\psi\rangle = |\psi_{A1}\rangle + |\psi'_1\rangle$ and $|\psi\rangle = |\psi_{A2}\rangle + |\psi'_2\rangle$, where vectors $|\psi'_{1,2}\rangle$ are orthogonal to $|\psi_{A1,2}\rangle$. Subtracting the second decomposition from the first, we obtain[10] $|\mathrm{null}\rangle = |\psi_{A1} - \psi_{A2}\rangle + |\psi'_1 - \psi'_2\rangle$. Because the two vectors in the right-hand side here are orthogonal and thus linearly independent, the equation cannot be satisfied unless $|\psi_{A1} - \psi_{A2}\rangle = |\psi'_1 - \psi'_2\rangle = |\mathrm{null}\rangle$, which gives $|\psi_{a1}\rangle = |\psi_{a2}\rangle$ and $|\psi'_1\rangle = |\psi'_2\rangle$. In other words, the decomposition $|\psi\rangle = |\psi_A\rangle + |\psi'\rangle$ is unique.

Using now Eqs. (1.44b) and (1.51b) and taking into account that by Rule I state vectors are normalized, we recover the second equation in (1.2),

$$\sum_a \mathrm{Prob}_\psi(A=a) = \sum_a |\psi_a|^2 = \|\psi\|^2 = 1. \tag{1.51c}$$

The formalism introduced here would be useless unless one is able to find the possible measurement outcomes $\{a\}$ for a given observable of interest A and to construct the corresponding orthonormal set of state vectors $\{|\phi_a\rangle\}$. This shortcoming should be viewed not a fatal flaw, but as a call for further mathematical and physical insight. Indeed, as we will see below, the rules of quantum theory can be also formulated in terms of observables rather than state vectors. Unlike state vectors, observables have direct classical analogs, and their properties can be deduced from symmetry considerations.

1.3.3 Case study: spin 1/2 in light of the rules

In this section we translate the consideration of Sec. 1.1.3 to the language of Hilbert spaces. As in Sec. 1.1.3, we assume here that the complete set of compatible observables consists of a single observable $S_{\mathbf{n}} = \mathbf{n}\cdot\mathbf{S}$, which in turn implies [see Sec. 1.3.2] that the corresponding measurement outcomes $S_{\mathbf{n}} = \pm\hbar/2$ [recall Eq. (1.16)] are non-degenerate. This sets the stage for the application of the rules of quantum theory in the form given in Sec. 1.3.2.

According to Rule II, possible measurement outcomes correspond to state vectors,

$$\text{outcomes } S_{\mathbf{n}} = \pm\hbar/2 \;\longrightarrow\; \text{state vectors } |\pm 1/2;\mathbf{n}\rangle. \tag{1.52}$$

These two state vectors represent the states for which $S_{\mathbf{n}} = \pm\hbar/2$ with probability 1 and, as shown in Sec. 1.3.2, form an orthonormal basis for the two-dimensional Hilbert space of spin 1/2. Components of $\mathbf{S}$ along directions other than $\mathbf{n}$, such as $S_{\mathbf{n}'}$, are incompatible with $S_{\mathbf{n}}$ and give rise to new orthonormal basis sets, such as $|\pm 1/2;\mathbf{n}'\rangle$. Each vector in this basis is a normalized linear combination of vectors $|\pm 1/2;\mathbf{n}\rangle$ and vice versa. This, obviously, is not the case for $\mathbf{n}' = -\mathbf{n}$. Indeed, since $S_{-\mathbf{n}} = -S_{\mathbf{n}}$, the state vectors $|\pm 1/2;\mathbf{n}\rangle$ and $|\mp 1/2;-\mathbf{n}\rangle$ represent the same quantum states, and thus [recall Eq. (1.48)] may differ in overall phase factors at most, i.e.,

$$|\pm 1/2;\mathbf{n}\rangle = |\mp 1/2;-\mathbf{n}\rangle \;\; \text{(up to phase factors)}. \tag{1.53}$$

This observation allows us to simplify the notations and write

$$|\pm 1/2;\mathbf{n}\rangle = |+1/2;\pm\mathbf{n}\rangle = |\pm\mathbf{n}\rangle \;\; \text{(up to phase factors)}. \tag{1.54}$$

In these notations the unit vectors $\pm\mathbf{n}$ in $|\pm\mathbf{n}\rangle$ are the **Bloch vectors** [recall Sec. 1.1.3]. That is, the state vector $|\pm\mathbf{n}\rangle$ represents the state for which $S_{\pm\mathbf{n}} = \hbar/2$ with probability 1.[25] The Born's rule (1.47) and Eq. (1.22) show that inner products of such state vectors satisfy

$$\big|\langle\mathbf{n}|\mathbf{n}'\rangle\big|^2 = \mathrm{Prob}_{\mathbf{n}}(S_{\mathbf{n}'} = \hbar/2) = \mathrm{Prob}_{\mathbf{n}'}(S_{\mathbf{n}} = \hbar/2) = \frac{1}{2}(1+\mathbf{n}\cdot\mathbf{n}'). \tag{1.55}$$

[25] Note that in spite of our general convention,[10] $|-\mathbf{n}\rangle \neq -|\mathbf{n}\rangle$ and $|\mathbf{n}_1\rangle + |\mathbf{n}_2\rangle \neq |\mathbf{n}_1+\mathbf{n}_2\rangle$. Actually, the expression $|\mathbf{n}_1+\mathbf{n}_2\rangle$ is meaningless because $\mathbf{n}_1+\mathbf{n}_2$ is not even a unit vector.

The symmetry of the probabilities with respect to the interchange $\mathbf{n} \rightleftarrows \mathbf{n}'$ noticed in Sec. 1.1.3 is ensured by the property $\langle \mathbf{n}|\mathbf{n}'\rangle = \langle \mathbf{n}'|\mathbf{n}\rangle^*$ of the inner product in the left-hand side.

Consider now the state ψ of spin 1/2 specified by the Bloch vector $\mathbf{n}_\psi$. Expansion of the corresponding state vector $|\mathbf{n}_\psi\rangle$ in the basis $|\pm\mathbf{n}\rangle$ [see Eq. (1.54)] can be parametrized as

$$|\mathbf{n}_\psi(\theta,\phi)\rangle = \cos(\theta/2)|+\mathbf{n}\rangle + e^{i\phi}\sin(\theta/2)|-\mathbf{n}\rangle \quad \text{(up to a phase factor)}, \tag{1.56a}$$

where the angles θ and ϕ belong to the intervals

$$0 \le \theta \le \pi, \quad 0 \le \phi < 2\pi, \tag{1.56b}$$

so that both $\cos(\theta/2)$ and $\sin(\theta/2)$ in Eq. (1.56a) are non-negative. For the Bloch vectors

$$\mathbf{n}_1 = \mathbf{n}_\psi(\pi/2,0), \quad \mathbf{n}_2 = \mathbf{n}_\psi(\pi/2,\pi/2), \quad \mathbf{n}_3 = \mathbf{n}_\psi(0,0) = \mathbf{n} \tag{1.57}$$

Eq. (1.56a) yields $\big|\langle \mathbf{n}_1|\mathbf{n}_2\rangle\big|^2 = \big|\langle \mathbf{n}_2|\mathbf{n}_3\rangle\big|^2 = \big|\langle \mathbf{n}_3|\mathbf{n}_1\rangle\big|^2 = 1/2$. Comparison with Eq. (1.55) then shows that vectors $\mathbf{n}_1$, $\mathbf{n}_2$, and $\mathbf{n}_3$ are perpendicular to each other. Moreover, it turns out that these vectors satisfy $\mathbf{n}_1 \times \mathbf{n}_2 = \mathbf{n}_3$ and thus can be taken as the right-handed Cartesian basis for the Euclidian space of ordinary geometric vectors. We can just as well call these three vectors $\mathbf{x}$, $\mathbf{y}$, and $\mathbf{z}$, respectively, and write

$$|\mathbf{n}_\psi(\theta,\phi)\rangle = \cos(\theta/2)|+\mathbf{z}\rangle + e^{i\phi}\sin(\theta/2)|-\mathbf{z}\rangle \quad \text{(up to an overall phase factor)} \tag{1.58}$$

instead of Eq. (1.56a). This gives, in particular,

$$|\mathbf{x}\rangle = |\mathbf{n}_\psi(\pi/2,0)\rangle = \frac{1}{\sqrt{2}}\big(|\mathbf{z}\rangle + |-\mathbf{z}\rangle\big), \quad |\mathbf{y}\rangle = |\mathbf{n}_\psi(\pi/2,\pi/2)\rangle = \frac{1}{\sqrt{2}}\big(|\mathbf{z}\rangle + i|-\mathbf{z}\rangle\big) \tag{1.59}$$

(again, up to phase factors). Evaluation of the dot products $\mathbf{x}\cdot\mathbf{n}_\psi$, $\mathbf{y}\cdot\mathbf{n}_\psi$, and $\mathbf{z}\cdot\mathbf{n}_\psi$ with the help of Eqs. (1.55), (1.58), and (1.59) is straightforward. The resulting expansion of the Bloch vector $\mathbf{n}_\psi$ in the basis $\{\mathbf{x},\mathbf{y},\mathbf{z}\}$ reads

$$\mathbf{n}_\psi(\theta,\phi) = \sum_{\boldsymbol{\alpha}=\mathbf{x},\mathbf{y},\mathbf{z}} \boldsymbol{\alpha}\,(\boldsymbol{\alpha}\cdot\mathbf{n}_\psi) = \mathbf{x}\sin\theta\cos\phi + \mathbf{y}\sin\theta\sin\phi + \mathbf{z}\cos\theta. \tag{1.60}$$

Eq. (1.60) is, of course, nothing but the standard parametrization of the unit vector $\mathbf{n}_\psi$ in the ***spherical polar coordinates***: the polar angle θ is the angle between $\mathbf{z}$ and $\mathbf{n}_\psi$, whereas the azimuthal angle ϕ is the angle between $\mathbf{x}$ and the projection of $\mathbf{n}_\psi$ onto (x,y)-plane $\mathbf{n}_\psi - \mathbf{z}(\mathbf{z}\cdot\mathbf{n}_\psi)$.[26] Importantly, every pair of angles (θ,ϕ) in the intervals (1.56b) corresponds to a unique Bloch vector $\mathbf{n}_\psi$ and vice versa. On the other hand, it is obvious that any state vector of spin 1/2 can be cast in the form of Eqs. (1.56) or Eq. (1.58) with unique angles θ and ϕ, and thus corresponds to a unique Bloch vector.[27] Accordingly, as mentioned above,[9] pure states of spin 1/2 are in one-to-one correspondence with Bloch vectors:

$$\text{state } \psi \text{ of spin } 1/2 \quad \rightleftarrows \quad \text{Bloch vector } \mathbf{n}_\psi \text{ such that } \mathrm{Prob}_\psi(S_{\mathbf{n}_\psi} = \hbar/2) = 1. \tag{1.61}$$

[26] Vector $\mathbf{n}_\psi$ can be obtained by first rotating vector $\mathbf{z}$ by θ about $\mathbf{y}$ and then rotating the resulting vector by ϕ about $\mathbf{z}$. Rotations will be discussed in details in Sec. 3.1.1 below.

[27] Note that this conclusion stems from Rule I [see Eq. (1.45)] and could not have been reached in Sec. 1.1.3.

Eqs. (1.58) and (1.60) allow us to write any state vector of spin 1/2 (specified by the corresponding Bloch vector) in terms of the basis vectors $|\pm\mathbf{z}\rangle$ [see, e.g., Eqs. (1.59)]. By construction, inner products of any two such state vectors satisfy Eq. (1.55), as one can easily verify by a direct calculation. For example, a change of sign of $\mathbf{n}_\psi$, i.e., the inversion $\mathbf{n}_\psi \to -\mathbf{n}_\psi$, is accomplished by replacing $\theta \to \pi - \theta$ and $\phi \to \pm\pi + \phi$ in Eq. (1.60). Making such replacements in Eq. (1.58), we obtain

$$|-\mathbf{n}_\psi\rangle = \sin(\theta/2)|+\mathbf{z}\rangle - e^{i\phi}\cos(\theta/2)|-\mathbf{z}\rangle \quad \text{(up to a phase factor)}. \tag{1.62}$$

This vector is both normalized and orthogonal to $|\mathbf{n}_\psi\rangle$, hence the two state vectors $|\pm\mathbf{n}_\psi\rangle$ indeed form an orthonormal basis for the two-dimensional Hilbert space of spin 1/2, as they should [recall Eq. (1.54)].

1.3.4 Degeneracy and projections

Imagine now that among all possible *non-degenerate* measurement outcomes of observable A some, say,

$$a_n, \quad n = 1, \ldots, \mathcal{N}_a \tag{1.63}$$

are so close to each other that the difference between them is beyond the resolution of our measurement. That is, as far as the measurement is concerned, these outcomes are indistinguishable: $a_n \approx a$. In this case, the experimentally relevant probability is obtained by summing up the probabilities of getting $A = a_n$ for all $a_n \approx a$,

$$\text{Prob}_\psi(A \approx a) = \sum_{a_n \approx a} \text{Prob}_\psi(A = a_n) = \sum_{n=1}^{n=\mathcal{N}_a} \left|\langle\phi_{an}|\psi\rangle\right|^2, \tag{1.64}$$

where the state vector $|\phi_{an}\rangle$ represents the state ϕ_{an} for which $\text{Prob}_{\phi_{an}}(A = a_n) = 1$ [see Eqs. (1.46) and (1.47)].

Suppose now that all a_n in Eq. (1.63) are identical, $a_n = a$ (as opposed to $a_n \approx a$). This happens when the outcome a is ***$\mathcal{N}_a$-fold degenerate***, i.e., when there exist $\mathcal{N}_a > 1$ mutually orthogonal vectors $|\phi_{an}\rangle$ representing states for which a measurement of A is certain to yield $A = a$.[28] It is natural to expect that in this case the probability of getting $A = a$ is again given by Eq. (1.64),

$$\text{Prob}_\psi(A = a) = \sum_{n=1}^{n=\mathcal{N}_a} \left|\langle\phi_{an}|\psi\rangle\right|^2. \tag{1.65}$$

There is a potential pitfall in this reasoning. Indeed, both Eqs. (1.64) and (1.65) feature the orthonormal set of vectors $\{|\phi_{an}\rangle\}$ spanning $\mathcal{N}_a$-dimensional space $\mathcal{H}_a$, a subspace of the whole Hilbert space $\mathcal{H}$. However, whereas in Eq. (1.64) every $|\phi_{an}\rangle$ corresponds to its own outcome a_n, there is nothing distinguishing the set $\{|\phi_{an}\rangle\}$ in Eq. (1.65) from any other orthonormal basis for the subspace $\mathcal{H}_a$. Fortunately, despite its appearance, Eq. (1.65) turns out to be independent of the choice of the basis, and thus makes a perfect sense.

[28] An obvious albeit trivial example of an observable with this property is $S^2_{\mathbf{n}}$ for spin 1/2.

Indeed, any vector $|\psi\rangle$ in $\mathcal{H}$ can be decomposed as

$$|\psi\rangle = |\psi_a\rangle + |\psi'\rangle, \tag{1.66a}$$

where $|\psi_a\rangle$ belongs to $\mathcal{H}_a$ and $|\psi'\rangle$ is orthogonal to all vectors in $\mathcal{H}_a$. Vector $|\psi_a\rangle$ defined by this relation is the ***projection*** of vector $|\psi\rangle$ onto the subspace $\mathcal{H}_a$; for a given $|\psi\rangle$ the decomposition (1.66a) and thus the projection $|\psi_a\rangle$ is obviously unique.[24] Let $\{|\phi_{an}\rangle\}$ be an arbitrary orthonormal basis for $\mathcal{H}_a$. Since $|\psi'\rangle$ is orthogonal to all $|\phi_{an}\rangle$, we have

$$\langle\phi_{an}|\psi\rangle = \langle\phi_{an}|\psi_a\rangle. \tag{1.66b}$$

Substituting $\langle\phi_{an}|\psi\rangle$ in this form into Eq. (1.65) and taking into account Eq. (1.44b), we obtain

$$\text{Prob}_\psi(A=a) = \sum_{n=1}^{n=\mathcal{N}_a} \left|\langle\phi_{an}|\psi_a\rangle\right|^2 = \|\psi_a\|^2. \tag{1.67}$$

Because the projection $|\psi_a\rangle$ of a given state vector $|\psi\rangle$ onto the subspace $\mathcal{H}_a$ is unique, so is its norm $\|\psi_a\|$, hence the right-hand side of Eq. (1.67) is manifestly basis-independent.

1.3.5 Rules in terms of projections

The discussion in Sec. 1.3.4 suggests that degeneracies can be accommodated by restating Rules II and III [see Eqs. (1.46) and (1.47)] in terms of subspaces and projections. To this end, we note that any vector in $\mathcal{H}_a$ coincides with its projection onto this subspace [see Eq. (1.66a)], hence any state vector $|\psi\rangle$ in $\mathcal{H}_a$ represents a quantum state ψ for which Eq. (1.67) gives $\text{Prob}_\psi(A=a) = \|\psi\|^2 = 1$. This property can be taken as the basis-independent definition of the subspace $\mathcal{H}_a$.

Clearly, instead of state vectors as in Eq. (1.46b), possible measurement outcomes can be associated with subspaces. This observation allows us to replace the original Rule II [see Eqs. (1.46)] with

$$\textbf{Rule II:} \quad \begin{array}{c}\textbf{outcome } \boldsymbol{a} \quad \rightleftarrows \quad \textbf{subspace } \boldsymbol{\mathcal{H}_a} \textbf{ of } \boldsymbol{\mathcal{H}} \textbf{ such that} \\ \textbf{Prob}_{\boldsymbol{\psi}}\boldsymbol{(A=a)=1} \textbf{ for all state vectors } \boldsymbol{|\psi\rangle} \textbf{ in } \boldsymbol{\mathcal{H}_a}.\end{array} \tag{1.68}$$

Note that the correspondence between the outcomes and the subspaces is *one-to-one*, unlike that between the degenerate outcomes and the states. As for the Born's rule (1.47), it can simply be replaced with Eq. (1.67),

$$\textbf{Rule III:} \quad \boldsymbol{\text{Prob}_\psi(A=a) = \|\psi_a\|^2}. \tag{1.69}$$

That is, the probability to find $A=a$ in the state ψ is given by the square of the norm of the projection $|\psi_a\rangle$ of the corresponding state vector $|\psi\rangle$ onto the subspace $\mathcal{H}_a$.

As in Sec. 1.3.2, the rules impose conditions on the subspaces $\mathcal{H}_a$. First of all, to ensure that $\text{Prob}_\psi(A\neq a) = 0$ for all state vectors $|\psi\rangle$ in $\mathcal{H}_a$, any vector in $\mathcal{H}_a$ must be orthogonal to any vector in $\mathcal{H}_{a'}$ with $a'\neq a$ [cf. Eq. (1.49)]. Therefore, the only vector *shared* by the subspaces corresponding to different possible measurement outcomes is $|\text{null}\rangle$. Such subspaces are called ***orthogonal***. Furthermore, the same reasoning as in Sec. 1.3.2 shows that any

vector $|\psi\rangle$ in $\mathcal{H}$ is exhausted by the sum of its projections onto these mutually orthogonal subspaces,

$$|\psi\rangle = \sum_a |\psi_a\rangle, \tag{1.70a}$$

which is another way of saying that the Hilbert space $\mathcal{H}$ is the ***direct sum***[29] of the subspaces $\mathcal{H}_a$,

$$\mathcal{H} = \bigoplus_a \mathcal{H}_a. \tag{1.70b}$$

Indeed, any $|\psi\rangle$ can be written as

$$|\psi\rangle = \textstyle\sum_a |\psi_a\rangle + |\psi'\rangle,$$

where $|\psi_a\rangle$ is the projection of $|\psi\rangle$ onto $\mathcal{H}_a$ and $|\psi'\rangle$ is orthogonal to all vectors in $\oplus_a \mathcal{H}_a$. Similar to Eqs. (1.50) and (1.66a), this decomposition is unique. If $|\psi'\rangle \neq |\text{null}\rangle$ then vector $|\varphi'\rangle = \|\psi'\|^{-1}|\psi'\rangle$ represents a state for which Eq. (1.69) gives $\text{Prob}_\varphi(A=a) = 0$ for all a, in contradiction with the second equation in (1.2). Therefore, $|\psi'\rangle = |\text{null}\rangle$, and the above decomposition reduces to Eq. (1.70a).

1.3.6 Ideal measurement

If the state of the system is known, the rules presented in Sec. 1.3.2 and 1.3.4 allow one to compute the probabilities and expectation values [see Sec. 1.1.1]. In order to describe sequential non-destructive reproducible measurements [see Sec. 1.1.2], we need to understand first how such measurements affect the states and the corresponding state vectors.

As in Sec. 1.1.2, we focus on a filtering-type measurement resulting in a formation of a filtered state (ensemble) for which a measurement of the chosen observable A is certain to yield the desired outcome, say, $A = a$:

$$\text{initial state } \psi_i \xrightarrow[\textbf{yielding } A=a]{\textbf{measurement of } A} \text{filtered state } \psi_a \text{ such that } \text{Prob}_{\psi_a}(A=a) = 1 \tag{1.71}$$

[cf. Eq. (1.8)]. If the outcome a is non-degenerate (i.e., if the corresponding subspace $\mathcal{H}_a$ is one-dimensional), then the filtered state ψ_a coincides with the state ϕ_a introduced in Eqs. (1.46). Eq. (1.71) then translates to the relation between the state vectors

$$|\psi_i\rangle \xrightarrow[\textbf{yielding } A=a]{\textbf{measurement of } A} |\phi_a\rangle \quad \text{(up to phase factors)}. \tag{1.72}$$

However, if the outcome a is $\mathcal{N}_a$-fold degenerate, then the condition $\text{Prob}_{\psi_a}(A=a) = 1$ does not specify the filtered state ψ_a uniquely. All one can say is that the corresponding state vector $|\psi_a\rangle$ belongs to $\mathcal{N}_a$-dimensional subspace $\mathcal{H}_a$ [recall Eq. (1.68)]. To remove the ambiguity, we require the measurement to be ***the least disturbing*** in the sense that it alters

[29] A set that includes all elements from two or more subsets is the ***union*** of these subsets. The standard symbol for such union is $\cup$ as in $\mathcal{H} = \bigcup_a \mathcal{H}_a$. A set that includes only elements that two or more subsets share is the ***intersection*** of these subsets. Intersections are denoted by $\cap$ as in $\bigcap_a \mathcal{H}_a = |\text{null}\rangle$ for orthogonal subspaces. A union of such orthogonal subspaces $\bigcup_a \mathcal{H}_a$ is called a ***direct sum*** and denoted by $\oplus$ as in Eq. (1.70b).

the initial state only to the extent necessary to ensure the desired outcome. In other words, the effect of measurement is to remove from the initial state vector $|\psi_i\rangle$ pieces that do not belong to $\mathcal{H}_a$ [see Eq. (1.66a)], leaving the reminder as is. Taking into account that the filtered state vector must be normalized, we arrive at the definition of an ***ideal measurement*** (also known as the ***projective measurement***):

$$|\psi_i\rangle \xrightarrow[\textbf{yielding } A=a]{\textbf{ideal measurement of } A} \frac{1}{\|\psi_a\|}|\psi_a\rangle \quad \text{(up to phase factors)}, \tag{1.73}$$

where $|\psi_a\rangle$ is a projection of $|\psi_i\rangle$ onto the subspace $\mathcal{H}_a$. We will show in Sec. 2.4.2 that if the initial state ψ_i is characterized by certain values of compatible with A observables [see Sec. 1.1.2], this property is retained in the filtered state ψ_a. It is mainly for this reason the measurement defined by Eq. (1.73) is referred to as *ideal*.

Any real-life measurement is a complicated process during which individual systems interact with the measurement apparatus. Eqs. (1.72) and (1.73) relate the state vector representing the state immediately after the measurement (the filtered state) to that immediately before the measurement (the initial state). Without painstaking analysis, it is impossible to tell in advance whether these relations apply to a given experiment. Fortunately, Eqs. (1.72) and (1.73) happen to describe Stern-Gerlach-type experiments rather well. At the same time, these equations are clearly inadequate if the observable of interest is continuous rather than discrete.[1]

2 Observables and Operators

2.1 Operators on Hilbert spaces

The scalar multiplication operation $|\psi\rangle \longmapsto |c\psi\rangle$ [see Eq. (1.26b)] establishes a correspondence between vectors $|\psi\rangle$ and $|c\psi\rangle$. Generalization of this idea is the concept of an ***operator*** defined as a rule associating each vector $|\psi\rangle$ in $\mathcal{H}$ with another vector $|\phi\rangle$ in the same space,

$$|\psi\rangle \xrightarrow{\textbf{operator } \hat{A}} |\phi\rangle. \tag{2.1a}$$

The latter vector is said to be the result of the ***action***[30] of $\hat{A}$ on $|\psi\rangle$, hence the notation[31]

$$\hat{A}|\psi\rangle = |\phi\rangle. \tag{2.1b}$$

Two operators are said to be ***equal*** if they act the same way:

$$\hat{A} = \hat{B} \text{ if and only if } \hat{A}|\psi\rangle = \hat{B}|\psi\rangle \text{ for all } |\psi\rangle \text{ in } \mathcal{H}. \tag{2.2}$$

2.1.1 Linear operators

We are mostly interested in ***linear operators*** that satisfy[10,31]

$$\hat{A}|\varphi + \psi\rangle = \hat{A}|\varphi\rangle + \hat{A}|\psi\rangle, \quad \hat{A}|c\psi\rangle = c|\hat{A}\psi\rangle \tag{2.3}$$

for all vectors $|\varphi\rangle$ and $|\psi\rangle$ in $\mathcal{H}$ and all complex numbers c. Perhaps the simplest examples of such operators are the ***identity operator*** $\hat{\mathbb{1}}$ that leaves vectors unchanged, and the ***zero operator*** $\hat{\mathbb{0}}$ that turns every vector to $|\text{null}\rangle$,[32]

$$\hat{\mathbb{1}}\,|\psi\rangle = |\psi\rangle, \quad \hat{\mathbb{0}}\,|\psi\rangle = |\text{null}\rangle. \tag{2.4}$$

It can be shown that for linear operators

$$\hat{A} = \hat{B} \text{ if and only if } \langle\psi|\hat{A}|\psi\rangle = \langle\psi|\hat{B}|\psi\rangle \text{ for all } |\psi\rangle \text{ in } \mathcal{H}. \tag{2.5}$$

This criterion is sometimes more convenient than the original definition (2.2). Moreover, in order to prove that two linear operators are equal, it is sufficient to demonstrate that either Eq. (2.2) or Eq. (2.5) holds for all *normalized* vectors in $\mathcal{H}$.

The second equation in (2.3) can be interpreted as the definition of a multiplication of a linear operator by a scalar. That is, $c\hat{A} = \hat{A}c$ is an operator that acts as[31]

$$c\hat{A}|\psi\rangle = c|\hat{A}\psi\rangle. \tag{2.6}$$

In the same spirit, the ***sum*** of operators $\hat{A}$ and $\hat{B}$ is defined as such operator $\hat{A} + \hat{B}$ that

$$(\hat{A} + \hat{B})|\psi\rangle = \hat{A}|\psi\rangle + \hat{B}|\psi\rangle. \tag{2.7}$$

[30] We adopt the convention that operators always act ***left-to-right***, i.e., on *kets*, never on *bras*.[12]

[31] We will often write $\hat{A}|\psi\rangle$ as $|\hat{A}\psi\rangle$.

[32] Note that the equation $\hat{A}|\psi\rangle = |\text{null}\rangle$ does not imply that $\hat{A} = \hat{\mathbb{0}}$ as in the second equation in (2.4) or that $|\psi\rangle = |\text{null}\rangle$ as in the equation $\hat{A}|\text{null}\rangle = |\text{null}\rangle$ obtained by setting $c = 0$ in the second equation in (2.3).

Comparison of Eqs. (2.3), (2.6), and (2.7) with Eqs. (1.27) shows that linear operators form a linear space with respect to the scalar multiplication and addition operations defined by Eqs. (2.6) and (2.7), respectively, and with zero operator $\hat{\mathbb{0}}$ playing the part of the null vector. As expected for a linear space [recall Eqs. (1.28)], we have

$$0\hat{A} = \hat{\mathbb{0}}, \quad c\hat{\mathbb{0}} = \hat{\mathbb{0}}. \tag{2.8}$$

Generalizing Eq. (2.6), we introduce the ***product*** $\hat{A}\hat{B}$ of linear operators $\hat{A}$ and $\hat{B}$ that acts according to

$$\hat{A}\hat{B}|\psi\rangle = \hat{A}|\hat{B}\psi\rangle. \tag{2.9}$$

That is, first $\hat{B}$ acts on $|\psi\rangle$, then $\hat{A}$ acts on $|\hat{B}\psi\rangle = \hat{B}|\psi\rangle$; $\hat{A}\hat{B}$ is obviously linear. This definition is naturally extended to products of three or more operators and integer powers of operators: the rightmost operator always acts first, then the second from the right, etc.

With powers of operators at hand, one can define operator-valued functions of linear operators. Consider an analytic function $f(x)$ whose Taylor expansion about $x = 0$ reads

$$f(x) = f(0)\,x^0 + f'(0)\,x^1 + \frac{1}{2!}\,f''(0)\,x^2 + \dots \tag{2.10a}$$

The corresponding operator function $\hat{f}(\hat{A})$ is obtained by replacing x here with $\hat{A}$, i.e.,[33]

$$\hat{f}(\hat{A}) = f(0)\,\hat{\mathbb{1}} + f'(0)\,\hat{A} + \frac{1}{2!}\,f''(0)\,\hat{A}^2 + \dots, \tag{2.10b}$$

where, by convention, $\hat{A}^0 = \hat{\mathbb{1}}$. For example, the ***exponent*** of $\lambda\hat{A}$ is the series

$$e^{\lambda\hat{A}} = \sum_{n=0}^{n=\infty} \frac{1}{n!}(\lambda\hat{A})^n = \hat{\mathbb{1}} + \lambda\hat{A} + \frac{1}{2!}(\lambda\hat{A})^2 + \frac{1}{3!}(\lambda\hat{A})^3 + \dots \tag{2.11a}$$

This series can be differentiated term-by-term with respect to the *parameter* λ (but not with respect to the operator $\hat{A}$!) with the result

$$\frac{d}{d\lambda}\,e^{\lambda\hat{A}} = \sum_{n=0}^{n=\infty} \frac{1}{n!}\,\lambda^n\hat{A}^{n+1} = \hat{A}\,e^{\lambda\hat{A}} = e^{\lambda\hat{A}}\hat{A}. \tag{2.11b}$$

Thus, $\hat{f}(\lambda) = e^{\lambda\hat{A}}$ is the solution of the first-order differential equation $\frac{d}{d\lambda}\hat{f}(\lambda) = \hat{A}\hat{f}(\lambda)$ subject to the condition $\hat{f}(0) = \hat{\mathbb{1}}$; it can be shown that this solution is unique.

Importantly, the order of operators in the product $\hat{A}\hat{B}$ generally matters. That is, the ***commutator*** of $\hat{A}$ and $\hat{B}$,

$$[\hat{A},\hat{B}] = \hat{A}\hat{B} - \hat{B}\hat{A}, \tag{2.12}$$

may differ from $\hat{\mathbb{0}}$. We say that operators $\hat{A}$ and $\hat{B}$ ***commute*** or that $\hat{A}$ commutes with $\hat{B}$ if $[\hat{A},\hat{B}] = \hat{\mathbb{0}}$. For example, $\hat{A}$ commutes with any power of $\hat{A}$ and thus [see Eq. (2.10b)] with

[33] The operator series such as that in the right-hand side side of Eq. (2.10b) is said to be convergent if vector series $f(0)|\psi\rangle + f'(0)\hat{A}|\psi\rangle + \frac{1}{2!}f''(0)\hat{A}^2|\psi\rangle + \dots$ converges[19] for all $|\psi\rangle$ in $\mathcal{H}$. Vector $\hat{f}(\hat{A})|\psi\rangle$ is the limit vector[19] of this series.

any function of $\hat{A}$, i.e., $[\hat{A}, f(\hat{A})] = \hat{\mathbb{0}}$, and both $\hat{\mathbb{1}}$ and $\hat{\mathbb{0}}$ [see Eq. (2.4)] commute with all operators. It is easy to verify the operator identities

$$[\hat{A}\hat{B}, \hat{C}] = \hat{A}[\hat{B}, \hat{C}] + [\hat{A}, \hat{C}]\hat{B}, \tag{2.13a}$$

$$\big[\hat{A}, [\hat{B}, \hat{C}]\big] + \big[\hat{B}, [\hat{C}, \hat{A}]\big] + \big[\hat{C}, [\hat{A}, \hat{B}]\big] = \hat{\mathbb{0}}, \tag{2.13b}$$

applicable to all linear operators $\hat{A}$, $\hat{B}$, and $\hat{C}$.

When dealing with exponents of linear operators [see Eqs. (2.11)], we will often find useful the expansion

$$e^{\hat{A}}\hat{B}e^{-\hat{A}} = \hat{B} + [\hat{A}, \hat{B}] + \frac{1}{2!}\big[\hat{A}, [\hat{A}, \hat{B}]\big] + \frac{1}{3!}\Big[\hat{A}, \big[\hat{A}, [\hat{A}, \hat{B}]\big]\Big] + \ldots, \tag{2.14}$$

valid for all $\hat{A}$ and $\hat{B}$. In particular, Eq. (2.14) shows that $e^{-\hat{A}}e^{\hat{A}} = e^{-\hat{A}}\hat{\mathbb{1}}e^{\hat{A}} = \hat{\mathbb{1}}$. Inserting the identity operator in this form between every two $\hat{B}$'s in $e^{\hat{A}}\hat{B}^n e^{-\hat{A}}$, we obtain

$$e^{\hat{A}}\hat{B}^n e^{-\hat{A}} = e^{\hat{A}}\hat{B}e^{-\hat{A}}e^{\hat{A}}\hat{B}e^{-\hat{A}}\cdots e^{\hat{A}}\hat{B}e^{-\hat{A}} = \big(e^{\hat{A}}\hat{B}e^{-\hat{A}}\big)^n,$$

so that

$$e^{\hat{A}}\hat{f}(\hat{B})e^{-\hat{A}} = \hat{f}\big(e^{\hat{A}}\hat{B}e^{-\hat{A}}\big) \tag{2.15}$$

for any operator-valued function of $\hat{B}$ defined via Taylor expansion as in Eq. (2.10b).

It is easy to see that for commuting operators $e^{\hat{A}}e^{\hat{B}} = e^{\hat{A}+\hat{B}}$, with $e^{-\hat{A}}e^{\hat{A}} = \hat{\mathbb{1}}$ being the obvious example. However, this relation is inapplicable if $[\hat{A}, \hat{B}] \neq \hat{\mathbb{0}}$. The product of exponents $e^{\hat{A}}e^{\hat{B}}$ simplifies drastically if the operators $\hat{A}$ and $\hat{B}$ commute with their commutator,

$$\big[\hat{A}, [\hat{A}, \hat{B}]\big] = \big[\hat{B}, [\hat{A}, \hat{B}]\big] = \hat{\mathbb{0}}, \tag{2.16a}$$

in which case $e^{\hat{A}}e^{\hat{B}}$ is given by the ***Baker-Hausdorff formula***

$$e^{\hat{A}}e^{\hat{B}} = e^{\hat{A}+\hat{B}}e^{\frac{1}{2}[\hat{A}, \hat{B}]}. \tag{2.16b}$$

Note that because $\hat{A}+\hat{B}$ commutes with $[\hat{A}, \hat{B}]$, the right-hand side of this formula can be written as $\exp\big(\hat{A}+\hat{B}+\frac{1}{2}[\hat{A}, \hat{B}]\big)$.

Hermitian conjugate (synonymous with ***adjoint*** in these Notes) of a linear operator $\hat{A}$ is such linear operator $\hat{A}^\dagger$ that

$$\langle\varphi|\hat{A}|\psi\rangle = \langle\psi|\hat{A}^\dagger|\varphi\rangle^* \text{ for all } |\varphi\rangle \text{ and } |\psi\rangle \text{ in } \mathcal{H}. \tag{2.17a}$$

(Note that the existence of $\hat{A}^\dagger$ for a given $\hat{A}$ is not guaranteed.) It is often convenient to write Eq. (2.17a) in a slightly different but equivalent form[31]

$$\langle\varphi|\hat{A}\psi\rangle = \langle\hat{A}^\dagger\varphi|\psi\rangle, \tag{2.17b}$$

where $\langle\hat{A}^\dagger\varphi|\psi\rangle$ is the inner product of vectors $|\hat{A}^\dagger\varphi\rangle$ and $|\psi\rangle$. Actually, it is sufficient to verify that Eqs. (2.17a) or (2.17b) are satisfied for $|\varphi\rangle = |\psi\rangle$. That is, $\hat{A}^\dagger$ is an Hermitian conjugate of $\hat{A}$ if and only if

$$\langle\psi|\hat{A}|\psi\rangle = \langle\psi|\hat{A}^\dagger|\psi\rangle^* \text{ or } \langle\psi|\hat{A}\psi\rangle = \langle\hat{A}^\dagger\psi|\psi\rangle \text{ for all } |\psi\rangle \text{ in } \mathcal{H}. \tag{2.17c}$$

It is easy to show that

$$(\hat{A}^\dagger)^\dagger = \hat{A}, \quad (\hat{A}+\hat{B})^\dagger = \hat{A}^\dagger + \hat{B}^\dagger, \quad (\hat{A}\hat{B})^\dagger = \hat{B}^\dagger\hat{A}^\dagger, \quad (c\hat{A})^\dagger = c^*\hat{A}^\dagger, \quad [\hat{f}(\hat{A})]^\dagger = \hat{f}(\hat{A}^\dagger). \tag{2.18}$$

To obtain, for example, the third relation here, we apply Eq. (2.17b) twice,

$$\langle\varphi|\hat{A}\hat{B}\psi\rangle = \langle\hat{A}^\dagger\varphi|\hat{B}\psi\rangle = \langle\hat{B}^\dagger\hat{A}^\dagger\varphi|\psi\rangle,$$

and compare this relation with the definition (2.17b).

As mentioned above, linear operators form a linear space. It turns out that

$$\|\hat{A}\| = \max_{\|\psi\|=1} \|\hat{A}\psi\| \tag{2.19}$$

complies with all the requirements for the norm of a vector in a linear space [see Eqs. (1.39)]. A linear operator $\hat{A}$ is called ***bounded*** if $\|\hat{A}\| < \infty$ and ***unbounded*** otherwise. It can be shown that all linear operators in finite-dimensional spaces are bounded. On the contrary, operators in infinite-dimensional spaces may well be unbounded. (As we will see later, this is not the only complication one has to deal with.) Moreover, many operators playing prominent roles in quantum theory happen to be of this type.

Unbounded operators are awkward to work with because these operators cannot even be defined on the whole Hilbert space. Indeed, it is obvious from Eq. (2.19) that if $\hat{A}$ is unbounded then $\mathcal{H}$ necessarily includes such $|\psi\rangle$ that $\hat{A}|\psi\rangle$ has an infinite norm and thus [see Sec. 1.2.4] does not belong to $\mathcal{H}$. A proper treatment of an unbounded operator $\hat{A}$ therefore involves keeping track of its ***domain***, i.e., a subset (not a subspace!) of $\mathcal{H}$ on which $\hat{A}$ is defined as a genuine Hilbert space operator. This implies, in particular, that many definitions given above need to be amended to account for the operator domains.

Unfortunately, dealing with operator domains is rather tricky. The good news is that many results obtained for bounded operators have their unbounded analogs. Therefore, apart from a brief transgression in Sec. 2.5.2, we will discuss in this Chapter only operators acting in finite-dimensional Hilbert spaces (hence bounded, see above). Later on, when we will encounter operators in infinite-dimensional spaces, we will treat these operators in the physicist's manner, i.e., by analogy.

2.1.2 Dirac's notations

In ***Dirac's notations***, linear operators are represented by linear combinations of objects of the type $|\varphi\rangle\langle\psi|$ often referred to as the ***outer products***[34] acting on vector $|\chi\rangle$ in $\mathcal{H}$ as[10]

$$\big(|\varphi\rangle\langle\psi|\big)|\chi\rangle = |\varphi\rangle\langle\psi|\chi\rangle, \tag{2.20}$$

where $\langle\psi|\chi\rangle$ is the inner product of $|\psi\rangle$ and $|\chi\rangle$. It is obvious from this definition that the product of operators $|\varphi_1\rangle\langle\psi_1|$ and $|\varphi_2\rangle\langle\psi_2|$ [see Eq. (2.9)] is given by[35,36]

$$\big(|\varphi_1\rangle\langle\psi_1|\big)\big(|\varphi_2\rangle\langle\psi_2|\big) = |\varphi_1\rangle\langle\psi_1|\varphi_2\rangle\langle\psi_2|. \tag{2.21a}$$

[34] Dirac's outer products are analogs of the so-called *dyadics* in the algebra of geometric vectors.

[35] We will often write $|\alpha\rangle c\langle\beta|$ instead of $c|\alpha\rangle\langle\beta|$.

[36] Eqs. (2.20) and (2.21a) illustrate the general rule: the "flat" sides of the *bra* and *ket* symbols[12] standing next to each other (in this order!) *merge* to form an inner product: $\langle\mathbf{bra}||\mathbf{ket}\rangle \longrightarrow \langle\mathbf{bra}|\mathbf{ket}\rangle$.

Similarly, it follows from Eqs. (1.33a), (1.34b), (2.17), and (2.20) that [35]

$$|a\varphi\rangle\langle b\psi| = |\varphi\rangle a b^* \langle\psi|, \quad \big(|\varphi\rangle c \langle\psi|\big)^\dagger = |\psi\rangle c^* \langle\varphi| \tag{2.21b}$$

for all complex numbers a, b, and c.

Any vector $|\psi\rangle$ in $\mathcal{H}$ can be expanded in an orthonormal basis $\{|\phi_n\rangle\}$ as

$$|\psi\rangle = \sum_n \psi_n |\phi_n\rangle, \quad \psi_n = \langle\phi_n|\psi\rangle. \tag{2.22}$$

[recall Eq. (1.43)]. Acting on both sides of the first equation here with a linear operator $\hat{A}$ and taking into account Eqs. (2.3), we obtain $\hat{A}|\psi\rangle = \sum_n \psi_n \hat{A}|\phi_n\rangle$. Therefore, a linear operator can be defined by specifying its action on the basis vectors. If, for example,

$$\hat{A}|\phi_n\rangle = |\varphi_n\rangle, \tag{2.23a}$$

then, obviously, $\hat{A}$ can be written in Dirac's notations as

$$\hat{A} = \sum_n |\varphi_n\rangle\langle\phi_n|. \tag{2.23b}$$

Note that unlike $|\phi_n\rangle$, vectors $|\varphi_n\rangle$ in Eqs. (2.23) are in general neither normalized nor mutually orthogonal. This, however, is not always the case. For example, for $\hat{A} = \hat{\mathbb{1}}$ Eq. (2.23a) turns to $\hat{\mathbb{1}}|\phi_n\rangle = |\phi_n\rangle$, and Eq. (2.23b) yields the ***completeness relation***

$$\hat{\mathbb{1}} = \sum_n |\phi_n\rangle\langle\phi_n|, \tag{2.24}$$

also known as the ***resolution of identity*** or the ***closure relation***. This relation may be viewed as yet another way of saying that the orthonormal set of vectors $\{|\phi_n\rangle\}$ is the basis for the space, i.e., is complete.[18]

With the help of the completeness relation (2.24), it is easy to write vectors, operators, and their products in one's favorite orthonormal basis. For example, Eq. (2.22) is recovered by writing

$$|\psi\rangle = \hat{\mathbb{1}}|\psi\rangle = \sum_n |\phi_n\rangle \underbrace{\langle\phi_n|\psi\rangle}_{\psi_n} = \sum_n \psi_n |\phi_n\rangle. \tag{2.25a}$$

An inner product of two vectors $\langle\varphi|\psi\rangle$ [see Eq. (1.44a)] is obtained by inserting the identity operator between $\langle\varphi|$ and $|\psi\rangle$,

$$\langle\varphi|\psi\rangle = \langle\varphi|\hat{\mathbb{1}}|\psi\rangle = \sum_n \underbrace{\langle\varphi|\phi_n\rangle}_{\varphi_n^*} \underbrace{\langle\phi_n|\psi\rangle}_{\psi_n} = \sum_n \varphi_n^* \psi_n. \tag{2.25b}$$

To derive the operator relation (2.23b), we write

$$\hat{A} = \hat{A}\hat{\mathbb{1}} = \sum_n \underbrace{\hat{A}|\phi_n\rangle}_{|\varphi_n\rangle} \langle\phi_n| = \sum_n |\varphi_n\rangle\langle\phi_n|. \tag{2.26}$$

As explained above, this relation is not very useful because vectors $|\varphi_n\rangle$ [see Eq. (2.23a)] are neither normalized nor orthogonal. This inconvenience is easily cured with the help of one more identity operator:

$$\hat{A} = \hat{\mathbb{1}}\hat{A}\hat{\mathbb{1}} = \sum_{m,n} |\phi_m\rangle \underbrace{\langle\phi_m|\hat{A}|\phi_n\rangle}_{A_{mn}} \langle\phi_n| = \sum_{m,n} |\phi_m\rangle A_{mn} \langle\phi_n|. \tag{2.27}$$

The coefficients

$$A_{mn} = \langle\phi_m|\hat{A}|\phi_n\rangle \tag{2.28}$$

in this expansion form $\mathcal{N}\times\mathcal{N}$ square matrix, where $\mathcal{N} = \dim\mathcal{H}$ is the dimension of the Hilbert space, and are referred to as the ***matrix elements*** of operator $\hat{A}$ in the orthonormal basis $\{|\phi_n\rangle\}$. The second equation in (2.21b) and Eq. (2.27) show that

$$\hat{A}^\dagger = \sum_{m,n} |\phi_m\rangle A^*_{nm} \langle\phi_n|, \tag{2.29}$$

hence the matrix formed by the matrix elements of $\hat{A}^\dagger$ is the Hermitian conjugate of the corresponding matrix for $\hat{A}$, i.e., $A^\dagger_{mn} = A^*_{nm}$.

It is clear that any expression involving vectors and operators, no matter how complex, can be written in terms of components of vectors and matrix elements of operators. All one needs to do is to insert identity operators in the right places and write these operators in the form of Eq. (2.24). In particular, components of vector $\hat{A}|\psi\rangle$ and matrix elements of the product of two operators $\hat{A}\hat{B}$ [recall Eq. (2.9)] turn to

$$\langle\phi_m|\hat{A}|\psi\rangle = \langle\phi_m|\hat{A}\hat{\mathbb{1}}|\psi\rangle = \sum_n \langle\phi_m|\hat{A}|\phi_n\rangle\langle\phi_n|\psi\rangle = \sum_n A_{mn}\psi_n, \tag{2.30a}$$

$$\langle\phi_m|\hat{A}\hat{B}|\phi_n\rangle = \langle\phi_m|\hat{A}\hat{\mathbb{1}}\hat{B}|\phi_n\rangle = \sum_k \langle\phi_m|\hat{A}|\phi_k\rangle\langle\phi_k|\hat{B}|\phi_n\rangle = \sum_k A_{mk}B_{kn}. \tag{2.30b}$$

We will encounter two quantities characterizing linear operators in finite-dimensional Hilbert spaces that are best defined in terms of their matrix elements (2.28). By definition, the ***trace*** of $\hat{A}$ is the sum of its diagonal matrix elements, i.e.,

$$\operatorname{tr}\hat{A} = \sum_n \langle\phi_n|\hat{A}|\phi_n\rangle. \tag{2.31}$$

The ***determinant*** of $\hat{A}$ is denoted by $\det\hat{A}$ and, as its name suggests, is the determinant of the square matrix formed by the matrix elements of $\hat{A}$. These quantities are useful because both turn out to be independent of the choice of the orthonormal basis.[37] It can be shown that for a product of two operators

$$\operatorname{tr}(\hat{A}\hat{B}) = \operatorname{tr}(\hat{B}\hat{A}), \quad \det(\hat{A}\hat{B}) = \det(\hat{B}\hat{A}) = \det\hat{A}\det\hat{B}. \tag{2.32a}$$

Since, obviously, $\operatorname{tr}(\hat{A}+\hat{B}) = \operatorname{tr}\hat{A} + \operatorname{tr}\hat{B}$, the first relation here can be also written as

$$\operatorname{tr}[\hat{A},\hat{B}] = 0. \tag{2.32b}$$

[37] Note that $\operatorname{tr}\hat{A}$ and $\det\hat{A}$ are well defined objects only for finite-dimensional Hilbert spaces. Indeed, in an infinite-dimensional space both quantities may well be divergent. (Consider, e.g., $\hat{A} = c\hat{\mathbb{1}}$.)

2.1.3 Matrix notations

Clearly, when working in a specific basis, the expansion coefficients is all that is needed. This observation leads to ***matrix notations*** (equivalent to Dirac's) in which *kets* and *bras*[12] are represented by the *column*- and *row*-matrices, respectively,[38]

$$|\psi\rangle \longmapsto \begin{pmatrix} \psi_1 \\ \vdots \\ \psi_{\mathcal{N}} \end{pmatrix}, \quad \langle\varphi| \longmapsto \begin{pmatrix} \varphi_1^* & \cdots & \varphi_{\mathcal{N}}^* \end{pmatrix}, \tag{2.33a}$$

so that the inner product $\langle\varphi|\psi\rangle$ [see Eq. (2.25b)] can be interpreted as a matrix product:

$$\langle\varphi|\psi\rangle = \begin{pmatrix} \varphi_1^* & \cdots & \varphi_{\mathcal{N}}^* \end{pmatrix} \begin{pmatrix} \psi_1 \\ \vdots \\ \psi_{\mathcal{N}} \end{pmatrix} = \sum_n \varphi_n^* \psi_n. \tag{2.33b}$$

Similarly, operators are represented by matrices of their matrix elements [see Eq. (2.28)],

$$\hat{A} \longmapsto \begin{pmatrix} A_{11} & \cdots & A_{1\mathcal{N}} \\ \vdots & \ddots & \vdots \\ A_{\mathcal{N}1} & \cdots & A_{\mathcal{N}\mathcal{N}} \end{pmatrix}. \tag{2.34}$$

Eq. (2.30a) shows that the action of an operator on a vector reduces to a matrix multiplication,

$$\hat{A}|\psi\rangle \longmapsto \begin{pmatrix} A_{11} & \cdots & A_{1\mathcal{N}} \\ \vdots & \ddots & \vdots \\ A_{\mathcal{N}1} & \cdots & A_{\mathcal{N}\mathcal{N}} \end{pmatrix} \begin{pmatrix} \psi_1 \\ \vdots \\ \psi_{\mathcal{N}} \end{pmatrix} = \begin{pmatrix} \sum_n A_{1n}\psi_n \\ \vdots \\ \sum_n A_{\mathcal{N}n}\psi_n \end{pmatrix}, \tag{2.35a}$$

whereas Eq. (2.30b) implies that the matrix representing a product of two operators is obtained by multiplying their matrices,

$$\hat{A}\hat{B} \longmapsto \begin{pmatrix} A_{11} & \cdots & A_{1\mathcal{N}} \\ \vdots & \ddots & \vdots \\ A_{\mathcal{N}1} & \cdots & A_{\mathcal{N}\mathcal{N}} \end{pmatrix} \begin{pmatrix} B_{11} & \cdots & B_{1\mathcal{N}} \\ \vdots & \ddots & \vdots \\ B_{\mathcal{N}1} & \cdots & B_{\mathcal{N}\mathcal{N}} \end{pmatrix}. \tag{2.35b}$$

In practice, choosing between the Dirac's and matrix notations is a matter of taste and convenience. We will use both.

2.2 Hermitian operators

Hermitian (or ***self-adjoint***) operators are linear operators that satisfy

$$\hat{A} = \hat{A}^\dagger. \tag{2.36a}$$

That is, a linear operator $\hat{A}$ is Hermitian if and only if

$$\langle\varphi|\hat{A}|\psi\rangle = \langle\psi|\hat{A}|\varphi\rangle^* \quad \text{or} \quad \langle\varphi|\hat{A}\psi\rangle = \langle\hat{A}\varphi|\psi\rangle \quad \text{for all } |\varphi\rangle \text{ and } |\psi\rangle \text{ in } \mathcal{H} \tag{2.36b}$$

[38] We will often replace the symbol $\longmapsto$ ("maps to") with $=$ ("equals") in equations such as (2.33)–(2.35).

[see Eqs. (2.17a) and (2.17b)], or, equivalently [see Eq. (2.17c)], if and only if

$$\langle\psi|\hat{A}\psi\rangle = \langle\psi|\hat{A}\psi\rangle^* \quad \text{for all } |\psi\rangle \text{ in } \mathcal{H}, \tag{2.36c}$$

i.e., if $\langle\psi|\hat{A}\psi\rangle$ is a real number. Eqs. (2.18) and (2.36a) show that linear combinations of Hermitian operators such as $c_1\hat{A}_1 + c_2\hat{A}_2 + \ldots$ are Hermitian if and only if the coefficients $c_1, c_2, \ldots$ are real, whereas products of Hermitian operators are Hermitian if and only if these operators commute.

2.2.1 Eigenvalue problem

The values of the parameter a for which the equation

$$\hat{A}|\psi\rangle = a|\psi\rangle \tag{2.37}$$

has solutions other than $|\psi\rangle = |\text{null}\rangle$ are the ***eigenvalues*** of operator $\hat{A}$. It follows immediately from Eq. (2.36c) that eigenvalues of Hermitian operators are *real*, i.e.,

$$a = a^*. \tag{2.38}$$

Furthermore, it can be shown[39] that the eigenvalues are solutions of the ***characteristic equation***

$$\det(\hat{A} - a\hat{\mathbb{1}}) = 0. \tag{2.39}$$

The left-hand side of this equation is obviously a polynomial of degree $\mathcal{N} = \dim\mathcal{H}$, the so-called ***characteristic polynomial***. Therefore, the equation has at least one but not more than $\mathcal{N}$ different solutions for a; all these solutions are guaranteed to be real [see Eq. (2.38)].

If a is an eigenvalue of $\hat{A}$ then vectors obeying Eq. (2.37) (including $|\text{null}\rangle$) form a Hilbert space $\mathcal{H}_a$, a subspace of the whole Hilbert space $\mathcal{H}$:

$$\text{eigenvalue } a \text{ of } \hat{A} \;\longrightarrow\; \text{subspace } \mathcal{H}_a \text{ such that } \hat{A}|\varphi_a\rangle = a|\varphi_a\rangle \text{ for all } |\varphi_a\rangle \text{ in } \mathcal{H}_a. \tag{2.40}$$

This subspace is the ***eigenspace*** corresponding to the eigenvalue a. It follows from Eqs. (1.34b), (2.36b), (2.38), and (2.40) that vectors $|\varphi_a\rangle$ and $|\varphi_{a'}\rangle$ taken from different eigenspaces satisfy

$$0 = \langle\hat{A}\varphi_a|\varphi_{a'}\rangle - \langle\varphi_a|\hat{A}\varphi_{a'}\rangle = (a-a')\langle\varphi_a|\varphi_{a'}\rangle,$$

which shows that $\langle\varphi_a|\varphi_{a'}\rangle = 0$ for $a \neq a'$ and thus $\mathcal{H}_a \cap \mathcal{H}_{a'} = |\text{null}\rangle$. Accordingly,

$$\text{eigenspaces corresponding to different eigenvalues are orthogonal.} \tag{2.41}$$

[39] Here we derive Eq. (2.39) for a two-dimensional Hilbert space. In matrix notations, Eq. (2.37) reads

$$\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}\begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix} = a\begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix} \quad\Longrightarrow\quad \begin{cases} (A_{11}-a)\psi_1 + A_{12}\psi_2 = 0, \\ A_{21}\psi_1 + (A_{22}-a)\psi_2 = 0, \end{cases}$$

where $\psi_{1,2}$ are components of $|\psi\rangle \neq |\text{null}\rangle$. If $A_{12} = A_{21} = 0$, the eigenvalues are given by A_{11} and A_{22}, in agreement with Eq. (2.39). In the less obvious case $A_{12} = A_{21}^* \neq 0$ the above equations show that both ψ_1 and ψ_2 differ from zero and give two values for their ratio,

$$\frac{\psi_2}{\psi_1} = -\frac{A_{11}-a}{A_{12}} \quad\text{and}\quad \frac{\psi_2}{\psi_1} = -\frac{A_{21}}{A_{22}-a}.$$

These values coincide (as they must!) if and only if $(A_{11}-a)(A_{22}-a) - A_{12}A_{21} = 0$, which is Eq. (2.39).

Moreover, it turns out that these orthogonal eigenspaces "cover" the entire Hilbert space in the sense that $\mathcal{H}$ is their direct sum,[29]

$$\mathcal{H} = \bigoplus_a \mathcal{H}_a, \tag{2.42}$$

or, equivalently, that any $|\psi\rangle$ in $\mathcal{H}$ is the sum of its projections onto the eigenspaces $\mathcal{H}_a$.

Here is the proof. One can always write $\mathcal{H}$ as

$$\mathcal{H} = \mathcal{H}_a \oplus \mathcal{H}', \tag{2.43}$$

where $\mathcal{H}_a$ is defined in Eq. (2.40) and $\mathcal{H}'$ is the subspace formed by vectors that are orthogonal to every vector in $\mathcal{H}_a$. Eqs. (2.36b) and (2.40) show that for all $|\varphi_a\rangle$ in $\mathcal{H}_a$ and all $|\varphi'\rangle$ in $\mathcal{H}'$

$$\langle\varphi_a|\hat{A}\varphi'\rangle = \langle\hat{A}\varphi_a|\varphi'\rangle = \langle a\varphi_a|\varphi'\rangle = a\langle\varphi_a|\varphi'\rangle = 0,$$

i.e., vector $|\hat{A}\varphi'\rangle$ is orthogonal to all vectors in $\mathcal{H}_a$ and thus, just as $|\varphi'\rangle$, belongs to $\mathcal{H}'$. This observation allows us to consider the eigenvalue problem for $\hat{A}$ on the subspace $\mathcal{H}'$,

$$\hat{A}|\varphi'\rangle = a'|\varphi'\rangle. \tag{2.44}$$

By construction, every eigenvalue a' solving this equation differs from a and the corresponding eigenspace $\mathcal{H}_{a'}$ is orthogonal to $\mathcal{H}_a$, as it should be [recall Eq. (2.41)].

Once one of the eigenvalues of $\hat{A}$ and the corresponding eigenspace have been found, the task of finding the remaining eigenvalues reduces to solving the eigenvalue problem for $\hat{A}$ in a "smaller" space. Repeating these steps, one will eventually determine *all* eigenvalues of $\hat{A}$ along with the corresponding eigenspaces, ending up with

$$\mathcal{H} = \left\{\bigoplus_a \mathcal{H}_a\right\} \oplus \mathcal{H}'' \tag{2.45}$$

for the Hilbert space. If $\mathcal{N} = \dim\mathcal{H}$ is finite, then so is $\mathcal{N}'' = \dim\mathcal{H}''$. Suppose $1 \le \mathcal{N}'' < \mathcal{N}$. The eigenvalue problem for $\hat{A}$ on the subspace $\mathcal{H}''$ [cf. Eq. (2.44)] is then guaranteed to produce at least one new eigenvalue, in contradiction with the assumption that all eigenvalues have been already accounted for in the first term in Eq. (2.45). The only remaining possibility is that $\mathcal{N}'' = 0$. That is, the subspace $\mathcal{H}''$ consists of only one vector, $|\text{null}\rangle$. Because $\oplus_a\mathcal{H}_a$ already includes this vector, Eq. (2.45) reduces to Eq. (2.42). This concludes the proof.

For each eigenspace $\mathcal{H}_a$ one can choose an orthonormal basis[40]

$$\left\{|\phi_{an}\rangle;\, n = 1, \ldots, \mathcal{N}_a = \dim\mathcal{H}_a\right\}, \quad \langle\phi_{am}|\phi_{an}\rangle = \delta_{m,n}. \tag{2.46a}$$

Eq. (2.42) implies that the union[29] of these basis sets $\bigcup_a\{|\phi_{an}\rangle\}$ is an orthonormal basis for the whole Hilbert space $\mathcal{H}$, i.e. [cf. Eq. (2.24)]

$$\hat{\mathbb{1}} = \sum_a \sum_{n=1}^{n=\mathcal{N}_a} |\phi_{an}\rangle\langle\phi_{an}|. \tag{2.46b}$$

[40] We say that the eigenvalue a is ***non-degenerate*** if $\mathcal{N}_a = 1$ and that it is **$\mathcal{N}_a$-fold degenerate** if $\mathcal{N}_a > 1$. In Linear Algebra, $\mathcal{N}_a$ is usually referred to as the ***multiplicity*** of the eigenvalue a.

The basis vectors $\{|\phi_{an}\rangle\}$ are the ***eigenvectors*** of $\hat{A}$. Solving the eigenvalue problem for an Hermitian operator amounts to finding all its eigenvalues and constructing a complete set of eigenvectors, i.e., an orthonormal basis in which every vector satisfies Eq. (2.37). Note that the choice of eigenvectors is not unique.

Substituting the identity operator in the form of Eq. (2.46b) into $\hat{A} = \hat{A}\hat{\mathbb{1}}$ and taking into account that $\hat{A}|\phi_{an}\rangle = a|\phi_{an}\rangle$, we obtain $\hat{A}$ in the basis of its own eigenvectors,

$$\hat{A} = \sum_a \sum_{n=1}^{n=\mathcal{N}_a} |\phi_{an}\rangle\, a\, \langle\phi_{an}|. \tag{2.47a}$$

This expansion is known as the ***spectral decomposition*** of $\hat{A}$. (For our purposes, ***spectrum*** of an operator is synonymous with the set of its eigenvalues.) In the same way, we find spectral decompositions of integer powers of $\hat{A}$,

$$\hat{A}^m = \sum_a \sum_{n=1}^{n=\mathcal{N}_a} |\phi_{an}\rangle\, a^m\, \langle\phi_{an}|. \tag{2.47b}$$

As discussed in Sec. 2.1.1, an operator-valued function of $\hat{A}$ can be defined via a Taylor-like series expansion (2.10b). Even though the operator $\hat{f}(\hat{A})$ is not necessarily Hermitian (consider, e.g., $\hat{f}(\hat{A}) = i\hat{A}$), Eq. (2.47b) implies that any such operator admits a spectral decomposition

$$\hat{f}(\hat{A}) = \sum_a \sum_{n=1}^{n=\mathcal{N}_a} |\phi_{an}\rangle\, f(a)\, \langle\phi_{an}|. \tag{2.48}$$

In fact, Eq. (2.48) can be taken as a definition of $\hat{f}(\hat{A})$ that generalizes and supersedes the original Taylor series-based definition (2.10b). Unlike the series expansion, the new definition (2.48) is not restricted to analytic functions. On the other hand, unlike Eq. (2.48), the series expansion-based definition of $\hat{f}(\hat{A})$ does not rely on the existence of the spectral decomposition (2.47a), and thus is applicable to all linear operators.

It follows immediately from Eqs. (2.47) that

$$\operatorname{tr}\hat{A}^n = \sum_a a^n \mathcal{N}_a, \quad \det\hat{A} = \prod_a a^{\mathcal{N}_a}. \tag{2.49a}$$

Since traces and determinants are independent of the choice of the basis [see Sec. 2.1.2], Eqs. (2.49a) provide an easy way of checking whether the eigenvalues and their degeneracies [40] have been found correctly.[41] It is also obvious from Eq. (2.47a) that the characteristic polynomial [see Eq. (2.39)] is given by

$$\det(\hat{A} - x\hat{\mathbb{1}}) = \prod_a (a - x)^{\mathcal{N}_a}. \tag{2.49b}$$

When x approaches a, this polynomial goes to zero as a power-law: $\det(\hat{A} - x\hat{\mathbb{1}}) \propto (a - x)^{\mathcal{N}_a}$. If the eigenvalue a is known, this property can be utilized to determine its degeneracy $\mathcal{N}_a$.

[41] Eqs. (2.49a) can be also used as an alternative to solving the characteristic equation (2.39). That is, eigenvalues $a_1, a_2, \ldots, a_{\mathcal{N}}$ of $\hat{A}$ (not necessarily all different) can be found by solving the system of linear equations $\sum_i a_i^n = \operatorname{tr}\hat{A}^n$ with $n = 1, \ldots, \mathcal{N}$. (One of these equations can be replaced with $\prod_i a_i = \det\hat{A}$.)

2.2.2 Projectors

Equations such as (2.46b)-(2.48) can be written in a more compact and basis-independent form by introducing the so-called ***projectors***. Projectors are Hermitian operators that satisfy

$$\hat{\mathcal{P}}^2 = \hat{\mathcal{P}}. \tag{2.50}$$

Either by equating the spectral decompositions of $\hat{\mathcal{P}}^2$ and $\hat{\mathcal{P}}$ [recall Eqs. (2.47)] or by acting with $\hat{\mathcal{P}}$ on both sides of the equation $\hat{\mathcal{P}}|\psi\rangle = p|\psi\rangle$, one can show that eigenvalues of $\hat{\mathcal{P}}$ obey the equation $p^2 = p$. This equation has only two solutions, $p = 1$ and $p = 0$. Therefore [recall Eqs. (2.40) and (2.42)], the Hilbert space is a direct sum of the two orthogonal eigenspaces corresponding to these eigenvalues,

$$\mathcal{H} = \mathcal{H}_1 \oplus \mathcal{H}_0\,, \tag{2.51a}$$

and any vector $|\psi\rangle$ in $\mathcal{H}$ is a sum of its projections onto these eigenspaces,

$$|\psi\rangle = |\psi_1\rangle + |\psi_0\rangle, \tag{2.51b}$$

Acting on both sides of this equation with $\hat{\mathcal{P}}$, we find

$$\hat{\mathcal{P}}\,|\psi\rangle = |\psi_1\rangle. \tag{2.51c}$$

Thus, $\hat{\mathcal{P}}$ does what its name (*projector*) suggests: it projects $|\psi\rangle$ onto the eigenspace $\mathcal{H}_1$. Note that $(\hat{\mathbb{1}} - \hat{\mathcal{P}})^2 = \hat{\mathbb{1}} - \hat{\mathcal{P}}$, hence $\hat{\mathbb{1}} - \hat{\mathcal{P}}$ is also a projector. Indeed, $(\hat{\mathbb{1}} - \hat{\mathcal{P}})|\psi\rangle = |\psi_0\rangle$, i.e., this operator projects $|\psi\rangle$ onto $\mathcal{H}_0$.

If $\{|\phi_{an}\rangle\}$ is an orthonormal basis for the $\mathcal{N}_a$-dimensional eigenspace $\mathcal{H}_a$ corresponding to the eigenvalue a [see Eqs. (2.40) and (2.46a)], then the operator

$$\hat{\mathcal{P}}_a = \sum_{n=1}^{n=\mathcal{N}_a} |\phi_{an}\rangle\langle\phi_{an}| \tag{2.52}$$

satisfies $\hat{\mathcal{P}}_a^2 = \hat{\mathcal{P}}_a$ and is therefore a projector. Obviously, $\hat{\mathcal{P}}_a$ projects any vector onto the eigenspace $\mathcal{H}_a$, whereas $(\hat{\mathbb{1}} - \hat{\mathcal{P}}_a)$ projects onto the subspace $\mathcal{H}'$ formed by vectors orthogonal to all vectors in $\mathcal{H}_a$ [see Eq. (2.43)]. Because eigenspaces corresponding to different eigenvalues are orthogonal [see Eq. (2.41)], projectors (2.52) satisfy

$$\hat{\mathcal{P}}_a\,\hat{\mathcal{P}}_{a'} = \hat{\mathcal{P}}_{a'}\hat{\mathcal{P}}_a = \delta_{a,a'}\,\hat{\mathcal{P}}_a\,, \tag{2.53}$$

hence a sum of any number of such ***orthogonal projectors*** is also a projector.[42] For example, $\hat{\mathcal{P}}_a + \hat{\mathcal{P}}_{a'}$ with $a \neq a'$ projects onto $\mathcal{H}_a \oplus \mathcal{H}_{a'}$, whereas $\sum_a \hat{\mathcal{P}}_a$ projects onto the whole Hilbert space $\mathcal{H} = \oplus_a \mathcal{H}_a$ and thus coincides with the identity operator:

$$\hat{\mathbb{1}} = \sum_a \hat{\mathcal{P}}_a\,. \tag{2.54}$$

[42] Note that the right-hand side of Eq. (2.52) also is a sum of projectors $\hat{\mathcal{P}}_{an} = |\phi_{an}\rangle\langle\phi_{an}|$ onto orthogonal one-dimensional subspaces: $\hat{\mathcal{P}}_{an}^2 = \hat{\mathcal{P}}_{an}$, $\hat{\mathcal{P}}_{am}\hat{\mathcal{P}}_{an} = \delta_{m,n}\hat{\mathcal{P}}_{an}$, $\hat{\mathcal{P}}_{an}|\psi\rangle \propto |\phi_{an}\rangle$.

Eq. (2.54) is, of course, nothing but the completeness relation (2.46b) expressed in terms of the orthogonal projectors (2.52) onto the eigenspaces of Hermitian operator $\hat{A}$. Similarly, the spectral decompositions (2.47a) and (2.48) can be written as

$$\hat{A} = \sum_a a\hat{\mathcal{P}}_a, \quad \hat{f}(\hat{A}) = \sum_a f(a)\hat{\mathcal{P}}_a. \tag{2.55}$$

The second equation here can be interpreted as the definition of an operator-valued function of $\hat{A}$ [see the discussion in Sec. 2.2.1]. It is obvious that for any function so defined

$$\hat{f}(\hat{A})\hat{\mathcal{P}}_a = \hat{\mathcal{P}}_a\hat{f}(\hat{A}) = f(a)\hat{\mathcal{P}}_a. \tag{2.56}$$

Consider now operator $\hat{A}$ that has N different eigenvalues a_i, $i = 1, \ldots, N$. Projectors $\hat{\mathcal{P}}_i$ onto the corresponding eigenspaces obey the equations

$$\sum_{i=1}^{i=N} a_i^n \hat{\mathcal{P}}_i = \hat{A}^n, \quad n = 0, 1, \ldots, N-1 \tag{2.57a}$$

This system of N linear equations yields N projectors $\hat{\mathcal{P}}_i$ in the form of polynomials of $\hat{A}$ of degree $N-1$. The second equation in (2.55) then shows that any function of $\hat{A}$ can also be viewed as a polynomial of the same degree, i.e.,

$$\hat{f}(\hat{A}) = \sum_{n=0}^{N-1} c_n \hat{A}^n. \tag{2.57b}$$

The coefficients c_n here can be found by solving the system of N equations

$$\sum_{n=0}^{N-1} c_n a_i^n = f(a_n), \quad i = 1, \ldots, N \tag{2.57c}$$

that can be obtained by multiplying both sides of Eq. (2.57b) by $\hat{\mathcal{P}}_i$ and taking into account Eq. (2.56). Since Hermitian operators on a finite-dimensional Hilbert space are guaranteed to have no more than $\mathcal{N} = \dim \mathcal{H}$ different eigenvalues, functions of these operators can always be written as polynomials of degree $N \leq \mathcal{N} - 1$ rather than as infinite series. This property makes functions of operators easier to handle, and we will use it often.

2.3 Observables as Hermitian operators

Let A be an observable with possible measurement outcomes $\{a\}$. As discussed in Sec. 1.3.5, each outcome a in this set corresponds to a subspace $\mathcal{H}_a$ of $\mathcal{H}$ $\mathcal{H}_a \subseteq \mathcal{H}$ [see Eq. (1.68)]. Furthermore, subspaces corresponding to different outcomes are mutually orthogonal and their direct sum coincides with the whole Hilbert space [see Eqs. (1.70)]. Consider now the operator

$$\hat{A} = \sum_a a\hat{\mathcal{P}}_a,$$

where $\hat{\mathcal{P}}_a$ is a projector onto the subspace $\mathcal{H}_a$ [cf. Eqs. (2.52)-(2.55)]. Because measurement outcomes are real, this operator is Hermitian.

Thus, for any observable one can construct an Hermitian operator with the eigenvalues that coincide with the possible measurement outcomes and with the eigenspaces that coincide with the subspaces introduced in Eq. (1.68). Conversely, it is *plausible* that every Hermitian operator acting on the Hilbert space of the system represents an observable that can (at least, in principle) be measured.

2.3.1 Rules in terms of operators

The above consideration suggests reformulating our rules [see Sec. 1.3.5] in terms of operators. In the new (and final!) version of the second rule the hypothesis that observables correspond to Hermitian operators and vice versa is given the status of a postulate:

$$\textbf{Rule II:}\quad \textbf{observable } \boldsymbol{A} \quad \rightleftarrows \quad \textbf{Hermitian operator } \hat{\boldsymbol{A}}. \tag{2.58}$$

Let $\hat{\mathcal{P}}_a$ be a projector onto the eigenspace $\mathcal{H}_a$ corresponding to the eigenvalue a of operator $\hat{A}$ [see Eq. (2.52)]. As discussed above, the eigenvalue represents one of the possible measurement outcomes of observable A. The probability that a measurement of A in the state ψ represented by the state vector $|\psi\rangle$ yields $A=a$ [see Eq. (1.69)] can then be written as

$$\mathrm{Prob}_\psi(A=a) = \|\psi_a\|^2 = \|\hat{\mathcal{P}}_a\psi\|^2 = \langle\hat{\mathcal{P}}_a\psi|\hat{\mathcal{P}}_a\psi\rangle = \langle\psi|\hat{\mathcal{P}}_a^2|\psi\rangle = \langle\psi|\hat{\mathcal{P}}_a|\psi\rangle, \tag{2.59}$$

where we took into account that $\hat{\mathcal{P}}_a$ is Hermitian and used Eqs. (2.36b), (2.50), and (2.52). These observations are summarized in the operator version of the Born's rule:

$$\textbf{Rule III:}\quad \begin{array}{l}\textbf{measurements of } \boldsymbol{A} \textbf{ yield eigenvalues } \boldsymbol{a} \textbf{ of } \hat{\boldsymbol{A}} \\ \textbf{with probabilities } \mathbf{Prob}_{\boldsymbol{\psi}}(\boldsymbol{A=a}) = \langle\boldsymbol{\psi}|\hat{\boldsymbol{\mathcal{P}}}_{\boldsymbol{a}}|\boldsymbol{\psi}\rangle.\end{array} \tag{2.60}$$

Of course, if all possible measurement outcomes $\{a\}$ of observable A and the corresponding orthogonal subspaces $\mathcal{H}_a$ are already known, there is no apparent benefit in translating the Rules to the operator language. This, however, is not how quantum theory normally works. As we will see later, there is a semi-systematic method of identifying Hermitian operators corresponding to the observables of interest. Having found such operator, one can solve the eigenvalue problem for it, thereby finding all a and $\mathcal{H}_a$.

Actually, solving the eigenvalue problem may not even be necessary. Indeed, with the help of the second equation in (2.55) and Eq. (2.60), the expectation value of $f(A)$ in the state ψ [see Eq. (1.4)] can be written as

$$\langle f(A)\rangle_\psi = \sum_a f(a)\,\mathrm{Prob}_\psi(A=a) = \sum_a f(a)\langle\psi|\hat{\mathcal{P}}_a|\psi\rangle = \langle\psi|\sum_a f(a)\hat{\mathcal{P}}_a|\psi\rangle = \langle\psi|\hat{f}(\hat{A})|\psi\rangle.$$

This result can be viewed as an alternative (and equivalent) form of the Born's rule:

$$\textbf{Rule III (alternative):}\quad \boldsymbol{\langle f(A)\rangle_\psi = \langle\psi|\hat{f}(\hat{A})|\psi\rangle}. \tag{2.61}$$

In particular, the probability of getting $A=a$ in the state ψ can be written as

$$\mathrm{Prob}_\psi(A=a) = \langle\delta_{a,A}\rangle_\psi = \langle\psi|\delta_{a,\hat{A}}|\psi\rangle \tag{2.62a}$$

[see Eqs. (1.7) and (2.61)]. Replacing the operator $\delta_{a,\hat{A}}$ here with its spectral decomposition $\delta_{a,\hat{A}} = \sum_{a'} \delta_{a,a'} \hat{\mathcal{P}}_{a'}$ [see the third equation in (2.55)], we recover Eq. (2.60):

$$\text{Prob}_\psi(A = a) = \sum_{a'} \delta_{a,a'} \langle\psi|\hat{\mathcal{P}}_{a'}|\psi\rangle = \langle\psi|\hat{\mathcal{P}}_a|\psi\rangle. \tag{2.62b}$$

With the help of Rule III in the form of Eq. (2.61), one can evaluate various expectation values without solving first the eigenvalue problem: it is sufficient to know the state vector and the operator representing the observable of interest, in any basis.

The operator formalism also allows us to eliminate the redundancy in the description of quantum states [recall Eq. (1.48)]. Indeed, instead of the state vector $|\psi\rangle$, a pure quantum state ψ can be associated with the projector

$$\hat{\rho}_\psi = |\psi\rangle\langle\psi| \tag{2.63}$$

referred to in this context as the ***state operator*** (also known as the ***density matrix*** or the ***density operator***). Because the phase change $|\psi\rangle \to e^{i\alpha}|\psi\rangle$ does not affect $\hat{\rho}_\psi$ [see the first equation in (2.21b)], the correspondence between the pure quantum states and the density operators is one-to-one. This observation suggests replacing Rule I [see Eq. (1.45)] with

$$\textbf{Rule I:}\quad \textbf{quantum state } \boldsymbol{\psi} \quad \rightleftarrows \quad \textbf{state operator } \hat{\boldsymbol{\rho}}_{\boldsymbol{\psi}} \textbf{ on } \boldsymbol{\mathcal{H}}. \tag{2.64}$$

Let $\{|\phi_n\rangle\}$ be an orthonormal basis for $\mathcal{H}$. The expectation value (2.61) can then be written as

$$\begin{aligned}\langle f(A)\rangle_\psi = \langle\psi|\hat{f}(\hat{A})\hat{\mathbb{1}}|\psi\rangle &= \sum_n \langle\psi|\hat{f}(\hat{A})|\phi_n\rangle\langle\phi_n|\psi\rangle \\ &= \sum_n \langle\phi_n|\psi\rangle\langle\psi|\hat{f}(\hat{A})|\phi_n\rangle = \sum_n \langle\phi_n|\hat{\rho}_\psi\hat{f}(\hat{A})|\phi_n\rangle = \text{tr}\big[\hat{\rho}_\psi\hat{f}(\hat{A})\big],\end{aligned}$$

and is independent of the choice of the basis. As shown above, the recipe for evaluating $\langle f(A)\rangle_\psi$ can be viewed as yet another alternative form of the Born's rule,

$$\textbf{Rule III:}\quad \boldsymbol{\langle f(A)\rangle_\psi = \text{tr}\big[\hat{\rho}_\psi\hat{f}(\hat{A})\big]}. \tag{2.65}$$

Eqs. (2.64) and (2.65) offer merely a glimpse into the state operator formalism. Although we will not pursue this line of reasoning further, it is worth mentioning that there is more to this formalism than its obvious esthetic appeal: it turns out that Rules I and III in the form of Eqs. (2.64) and (2.65) apply to the description of both pure and mixed quantum states [see Sec. 1.3.1].[43]

Eqs. (2.61) and (2.65) show that as far as the evaluation of various expectation values is concerned, neither the solution of the eigenvalue problem nor the construction of the projection operators are required. However, there is no obvious way around it one's goal is to

[43] Density operators for both pure and mixed states satisfy

$$\hat{\rho}_\psi = \hat{\rho}_\psi^\dagger, \quad \text{tr}\,\hat{\rho}_\psi = 1, \quad \langle\varphi|\hat{\rho}_\psi|\varphi\rangle \geq 0 \text{ for all } |\varphi\rangle \text{ in } \mathcal{H}.$$

The difference between the two types of states is that whereas for a pure state $\hat{\rho}_\psi$ is a projector [see Eq. (2.63)], hence $\text{tr}\,\hat{\rho}_\psi^2 = \text{tr}\,\hat{\rho}_\psi = 1$, for a mixed state $\hat{\rho}_\psi^2 \neq \hat{\rho}_\psi$ and $0 < \text{tr}\,\hat{\rho}_\psi^2 < 1$.

describe the sequential measurements [recall Sec. 1.1.2 and 1.3.6]. Indeed, according to our definition [see Eq. (1.73)], an ideal filtering-type measurement *projects* the initial state vector onto an eigenspace of $\hat{A}$,

$$|\psi\rangle \xrightarrow[\text{yielding } A=a]{\text{ideal measurement of } A} \frac{1}{\|\hat{\mathcal{P}}_a\psi\|}\hat{\mathcal{P}}_a|\psi\rangle \text{ (up to phase factors)}. \tag{2.66}$$

Note that $\|\hat{\mathcal{P}}_a\psi\|$ in the right-hand side here can be also written as $[\mathrm{Prob}_\psi(A=a)]^{1/2}$.

2.3.2 Case study: spin 1/2 from the operator perspective

In this section, we will use the results of Sec. 1.1.3 and 1.3.3 to *construct* spin 1/2 operators. We again focus on the observable $S_\mathbf{n} = \mathbf{n}\cdot\mathbf{S}$ with two non-degenerate possible measurement outcomes $S_\mathbf{n} = \pm\hbar/2$. As shown in Sec. 1.3.5 [see Eq. (1.68)] these outcomes correspond to two orthogonal one-dimensional subspaces,

$$S_\mathbf{n} = \pm\hbar/2 \;\longrightarrow\; \text{subspace } \mathcal{H}_\pm \text{ such that } \mathrm{Prob}_\psi(S_\mathbf{n} = \pm\hbar/2) = 1 \text{ for } |\psi\rangle \text{ in } \mathcal{H}_\pm, \tag{2.67}$$

and the Hilbert space is their direct sum, $\mathcal{H} = \mathcal{H}_+ \oplus \mathcal{H}_-$ [see Eq. (1.70b)]. As discussed in Sec. 2.3.1, the operator $\hat{S}_\mathbf{n}$ representing observable $S_\mathbf{n}$ is an Hermitian operator with eigenvalues $\pm\hbar/2$ and eigenspaces $\mathcal{H}_\pm$. The spectral decomposition of $\hat{S}_\mathbf{n}$ [cf. the first equation in (2.55)] therefore reads

$$\hat{S}_\mathbf{n} = (\hbar/2)\,\hat{\mathcal{P}}_+ + (-\hbar/2)\,\hat{\mathcal{P}}_- = \frac{\hbar}{2}(\hat{\mathcal{P}}_+ - \hat{\mathcal{P}}_-), \tag{2.68a}$$

where $\hat{\mathcal{P}}_\pm$ are projectors onto $\mathcal{H}_\pm$. Vectors $|\pm\mathbf{n}\rangle$ introduced in Eqs. (1.52) and (1.54) [see also Eqs. (1.58) and (1.62)] are eigenvectors of operator $\hat{S}_\mathbf{n}$ with eigenvalues $\pm\hbar/2$, i.e., normalized solutions of the equation $\hat{S}_\mathbf{n}|\pm\mathbf{n}\rangle = \pm(\hbar/2)|\pm\mathbf{n}\rangle$. In terms of these vectors, the projectors $\hat{\mathcal{P}}_\pm$ are given by [cf. Eq. (2.52)]

$$\hat{\mathcal{P}}_\pm = |\pm\mathbf{n}\rangle\langle\pm\mathbf{n}|. \tag{2.68b}$$

The vector nature of $\mathbf{n}$ and $\mathbf{S}$ in their dot product $S_\mathbf{n} = \mathbf{n}\cdot\mathbf{S}$ suggests that the corresponding operator $\hat{S}_\mathbf{n}$ can be also written as

$$\hat{S}_\mathbf{n} = \mathbf{n}\cdot\hat{\mathbf{S}} = n_\mathbf{x}\hat{S}_\mathbf{x} + n_\mathbf{y}\hat{S}_\mathbf{y} + n_\mathbf{z}\hat{S}_\mathbf{z}, \tag{2.69a}$$

where $\{\mathbf{x},\mathbf{y},\mathbf{z}\}$ is an arbitrary orthonormal Cartesian basis and

$$\hat{\mathbf{S}} = \mathbf{x}\hat{S}_\mathbf{x} + \mathbf{y}\hat{S}_\mathbf{y} + \mathbf{z}\hat{S}_\mathbf{z} \tag{2.69b}$$

is an Hermitian ***vector operator***[44] representing spin 1/2. Eqs. (2.69) identify $\hat{S}_\mathbf{n}$ with the component of the vector operator $\hat{\mathbf{S}}$ along the direction $\mathbf{n}$. Although quite plausible, these relations may seem suspicious because the operators in their right-hand sides do not have to commute with each other. (In fact, they do not, as we will see shortly below. Moreover, it will be shown in Sec. 2.4 that the non-commutativity of components of $\hat{\mathbf{S}}$ is directly related to the incompatibility of the corresponding observables.) However, it is easy to see that Eqs. (2.69) are in fact consistent with the assumptions made earlier. Indeed, according to Eqs. (1.61),

[44] A vector operator is a geometric vector whose Cartesian components are operators.

(2.5), and (2.61), the operator relations (2.69) hold if and only if the expectation values of the corresponding observables satisfy

$$\langle \mathbf{n}\cdot\mathbf{S}\rangle_{\mathbf{n}_\psi} = \mathbf{n}\cdot\langle\mathbf{S}\rangle_{\mathbf{n}_\psi}, \quad \langle\mathbf{S}\rangle_{\mathbf{n}_\psi} = \mathbf{x}\langle S_\mathbf{x}\rangle_{\mathbf{n}_\psi} + \mathbf{y}\langle S_\mathbf{y}\rangle_{\mathbf{n}_\psi} + \mathbf{z}\langle S_\mathbf{z}\rangle_{\mathbf{n}_\psi} \tag{2.70}$$

for all Bloch vectors $\mathbf{n}_\psi$. Apart from notations, Eqs. (2.70) coincide with Eqs. (1.18). Therefore, the assumptions made in Sec. 1.1.3 and carried over to Sec. 1.3.3 and to the present section inevitably lead to Eqs. (2.70) and thus to Eqs. (2.69).

Vector $\langle\mathbf{S}\rangle_{\mathbf{n}_\psi}$ was found explicitly in Sec. 1.1.3 [see Eq. (1.20)],

$$\langle\mathbf{S}\rangle_{\mathbf{n}_\psi} = \langle\psi|\hat{\mathbf{S}}|\psi\rangle = (\hbar/2)\,\mathbf{n}_\psi. \tag{2.71}$$

This relation ensures that, as expected for vectors, $\langle\mathbf{S}\rangle_{\mathbf{n}_\psi}$ and thus $\hat{\mathbf{S}}$ are independent of the choice of the Cartesian basis $\{\mathbf{x},\mathbf{y},\mathbf{z}\}$. It can be also viewed as the operator-based recipe for finding the Bloch vector for a given pure state of spin 1/2, $\mathbf{n}_\psi = (2/\hbar)\,\langle\psi|\hat{\mathbf{S}}|\psi\rangle$. Replacing vector operator $\hat{\mathbf{S}}$ here with its Cartesian components yields components of $\mathbf{n}_\psi$.

In Eqs. (2.68) $\hat{S}_\mathbf{n}$ is written in the basis of its own eigenvectors $|\pm\mathbf{n}\rangle$. It is not difficult to convert $\hat{S}_\mathbf{n}$ to any other orthonormal basis, such as $|\pm\mathbf{z}\rangle$. Eq. (2.61) shows that the diagonal matrix elements of $\hat{S}_\mathbf{n}$ coincide with the expectation values of $S_\mathbf{n}$ in the states specified by the Bloch vectors $\pm\mathbf{z}$. Using Eqs. (2.70) and (2.71), we find

$$\langle\pm\mathbf{z}|\hat{S}_\mathbf{n}|\pm\mathbf{z}\rangle = \langle S_\mathbf{n}\rangle_{\pm\mathbf{z}} = \pm\,(\hbar/2)\,n_\mathbf{z}\,. \tag{2.72}$$

The off-diagonal matrix elements of the projector $\hat{\mathcal{P}}_+$ [see Eq. (2.68b)] are evaluated with the help of Eqs. (1.58) and (1.60):

$$\langle\pm\mathbf{z}|\hat{\mathcal{P}}_+|\mp\mathbf{z}\rangle = \langle\pm\mathbf{z}|\mathbf{n}\rangle\langle\mathbf{n}|\mp\mathbf{z}\rangle = \langle\pm\mathbf{z}|\mathbf{n}\rangle\langle\mp\mathbf{z}|\mathbf{n}\rangle^* = \frac{1}{2}\,e^{\mp i\phi}\sin\theta = \frac{1}{2}\,(n_\mathbf{x} \mp i n_\mathbf{y}). \tag{2.73a}$$

Changing the sign of $\mathbf{n}$ here, we obtain the matrix elements of $\hat{\mathcal{P}}_-$,

$$\langle\pm\mathbf{z}|\hat{\mathcal{P}}_-|\mp\mathbf{z}\rangle = -\,\langle\pm\mathbf{z}|\hat{\mathcal{P}}_+|\mp\mathbf{z}\rangle. \tag{2.73b}$$

Eqs. (2.68a) and (2.73) then yield the off-diagonal matrix elements of $\hat{S}_\mathbf{n}$,

$$\langle\pm\mathbf{z}|\hat{S}_\mathbf{n}|\mp\mathbf{z}\rangle = (\hbar/2)(n_\mathbf{x} \mp i n_\mathbf{y}). \tag{2.73c}$$

The matrix representing operator $\hat{S}_\mathbf{n}$ in $|\pm\mathbf{z}\rangle$ basis [cf. Eq. (2.34)] therefore reads

$$\hat{S}_\mathbf{n} \longmapsto \begin{pmatrix} \langle+\mathbf{z}|\hat{S}_\mathbf{n}|+\mathbf{z}\rangle & \langle+\mathbf{z}|\hat{S}_\mathbf{n}|-\mathbf{z}\rangle \\ \langle-\mathbf{z}|\hat{S}_\mathbf{n}|+\mathbf{z}\rangle & \langle-\mathbf{z}|\hat{S}_\mathbf{n}|-\mathbf{z}\rangle \end{pmatrix} = \frac{\hbar}{2}\begin{pmatrix} n_\mathbf{z} & n_- \\ n_+ & -n_\mathbf{z} \end{pmatrix}, \quad n_\pm = n_\mathbf{x} \pm i n_\mathbf{y}\,. \tag{2.74}$$

It is straightforward to verify that Eqs. (2.69) are satisfied and that vectors $|\pm\mathbf{n}\rangle$ given by Eqs. (1.60) and (1.62) are indeed eigenvectors of $\hat{S}_\mathbf{n}$ with eigenvalues $\pm\hbar/2$.

Replacing $\mathbf{n}$ in Eq. (2.74) with $\mathbf{x},\mathbf{y}$, and $\mathbf{z}$, we obtain the Cartesian components of $\hat{\mathbf{S}}$,[38]

$$\hat{S}_\mathbf{x} = \frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \hat{S}_\mathbf{y} = \frac{\hbar}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \hat{S}_\mathbf{z} = \frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{2.75}$$

Direct calculation shows that the product of any two of these operators is proportional to the third one:

$$\begin{aligned}
\hat{S}_{\mathbf{x}}\hat{S}_{\mathbf{y}} &= -\hat{S}_{\mathbf{y}}\hat{S}_{\mathbf{x}} = i(\hbar/2)\hat{S}_{\mathbf{z}},\\
\hat{S}_{\mathbf{y}}\hat{S}_{\mathbf{z}} &= -\hat{S}_{\mathbf{z}}\hat{S}_{\mathbf{y}} = i(\hbar/2)\hat{S}_{\mathbf{x}},\\
\hat{S}_{\mathbf{z}}\hat{S}_{\mathbf{x}} &= -\hat{S}_{\mathbf{z}}\hat{S}_{\mathbf{x}} = i(\hbar/2)\hat{S}_{\mathbf{y}}.
\end{aligned} \tag{2.76}$$

Although these relations are specific for spin 1/2, it will be shown in Sec. 3.2.1 below that the commutation relations that follow from Eqs. (2.76) (such as $\bigl[\hat{S}_{\mathbf{x}},\hat{S}_{\mathbf{y}}\bigr] = i\hbar\hat{S}_{\mathbf{z}}$) hold for all angular momentum operators.

Dropping the factors $\hbar/2$ in the right-hand sides of Eqs. (2.75), we get the Cartesian components of the dimensionless vector operator $\hat{\boldsymbol{\sigma}} = (2/\hbar)\hat{\mathbf{S}}$, the so-called ***Pauli matrices***

$$\hat{\sigma}_1 \equiv \hat{\sigma}_{\mathbf{x}} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \hat{\sigma}_2 \equiv \hat{\sigma}_{\mathbf{y}} = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \hat{\sigma}_3 \equiv \hat{\sigma}_{\mathbf{z}} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{2.77a}$$

The utility of these matrices is not limited to the description of spin 1/2 in $|\pm\mathbf{z}\rangle$ basis. Indeed, it turns out that for any linear operator $\hat{A}$ acting in a two-dimensional Hilbert space and for any orthonormal basis the 2×2 matrix representing $\hat{A}$ [see Eq. (2.34)] can be written as a linear combination of the three Pauli matrices (2.77a) and the unit matrix $\hat{\sigma}_0$ representing the identity operator,

$$\hat{A} = \sum_i \alpha_i \hat{\sigma}_i\,. \tag{2.77b}$$

In other words, the four linearly independent Hermitian matrices $\{\hat{\sigma}_i, i = 0,1,2,3\}$ form a basis for the linear space of linear operators [see Sec. 2.1.1] on a two-dimensional Hilbert space. (Notice that the dimension of the operator space is four, not two.) On this space, one can define an inner product [see Sec. 1.2.2] via the relation

$$(\hat{A},\hat{B}) = \frac{1}{2}\,\mathrm{tr}(\hat{A}^\dagger\hat{B}). \tag{2.77c}$$

(To avoid confusion, we will not use the bra/ket notations for this product.) It is easy to verify that the basis $\{\hat{\sigma}_i\}$ is orthonormal with respect to this inner product, i.e., $(\hat{\sigma}_i,\hat{\sigma}_j) = \delta_{i,j}$, and that the coefficients α_i in the expansion (2.77b) are given by the inner products

$$\alpha_i = (\hat{\sigma}_i,\hat{A}) = \frac{1}{2}\,\mathrm{tr}\,(\hat{\sigma}_i^\dagger\hat{A}) = \frac{1}{2}\,\mathrm{tr}\,(\hat{\sigma}_i\hat{A}), \tag{2.77d}$$

as expected for the components of a vector in a linear vector space [cf. Eq. (1.43)]. Note that in the language of spin 1/2 operators, the very existence of the expansion (2.77b) implies that any operator-valued function of components of $\hat{\mathbf{S}}$, no matter how complex, can be written as

$$\hat{f}\bigl(\hat{S}_{\mathbf{x}},\hat{S}_{\mathbf{y}},\hat{S}_{\mathbf{z}}\bigr) = c_0\hat{\mathbb{1}} + c_{\mathbf{x}}\hat{S}_{\mathbf{x}} + c_{\mathbf{y}}\hat{S}_{\mathbf{y}} + c_{\mathbf{z}}\hat{S}_{\mathbf{z}},$$

i.e., it is linear in components of $\hat{\mathbf{S}}$. (The coefficients $c_{\mathbf{x}}$, $c_{\mathbf{y}}$, and $c_{\mathbf{z}}$ here are not necessarily components of a vector.)

2.4 Compatible and incompatible observables

2.4.1 Commuting Hermitian operators

Let $\hat{A}$ and $\hat{B}$ be Hermitian operators with the spectral decompositions [recall the first equation in (2.55)]

$$\hat{A} = \sum_a a\hat{\mathcal{P}}_a, \quad \hat{B} = \sum_b b\hat{\mathcal{P}}_b, \tag{2.78}$$

where a and b are eigenvalues of $\hat{A}$ and $\hat{B}$, respectively, and $\hat{\mathcal{P}}_a$ and $\hat{\mathcal{P}}_b$ are projectors onto the corresponding eigenspaces $\mathcal{H}_a$ and $\mathcal{H}_b$ [see Eq. (2.40)]. It turns out that

$$\boldsymbol{[\hat{A},\hat{B}] = \hat{0}} \textbf{ if and only if } \boldsymbol{[\hat{\mathcal{P}}_a,\hat{\mathcal{P}}_b] = \hat{0}} \textbf{ for all } \boldsymbol{a} \textbf{ and } \boldsymbol{b}. \tag{2.79}$$

Whereas it is obvious that if $[\hat{\mathcal{P}}_a,\hat{\mathcal{P}}_b] = \hat{\mathbb{0}}$ for all a and b then $[\hat{A},\hat{B}] = \hat{\mathbb{0}}$, it is less obvious that $[\hat{A},\hat{B}] = \hat{\mathbb{0}}$ implies $[\hat{\mathcal{P}}_a,\hat{\mathcal{P}}_b] = \hat{\mathbb{0}}$. This assertion can be proven as follows. Using the relation $\hat{A}\hat{\mathcal{P}}_a = a\hat{\mathcal{P}}_a$ [see Eq. (2.56)], we find for $[\hat{A},\hat{B}] = \hat{\mathbb{0}}$ and for all $|\psi\rangle$ in $\mathcal{H}$

$$\hat{A}\big(\hat{B}\hat{\mathcal{P}}_a|\psi\rangle\big) = \hat{B}\hat{A}\hat{\mathcal{P}}_a|\psi\rangle = \hat{B}a\hat{\mathcal{P}}_a|\psi\rangle = a\big(\hat{B}\hat{\mathcal{P}}_a|\psi\rangle\big). \tag{2.80a}$$

Accordingly, vector $\hat{B}\hat{\mathcal{P}}_a|\psi\rangle$ belongs to the eigenspace $\mathcal{H}_a$ [see Eq. (2.40)] of $\hat{A}$ and thus coincides with its projection onto this eigenspace, i.e.,

$$\hat{\mathcal{P}}_a\hat{B}\hat{\mathcal{P}}_a|\psi\rangle = \hat{B}\hat{\mathcal{P}}_a|\psi\rangle. \tag{2.80b}$$

Since $|\psi\rangle$ is arbitrary, Eq. (2.80b) implies [recall Eq. (2.2)] the operator relation

$$\hat{\mathcal{P}}_a\hat{B}\hat{\mathcal{P}}_a = \hat{B}\hat{\mathcal{P}}_a. \tag{2.80c}$$

The left-hand side of Eq. (2.80c) is obviously Hermitian, hence the right-hand side must be Hermitian as well, i.e., $\hat{B}\hat{\mathcal{P}}_a = (\hat{B}\hat{\mathcal{P}}_a)^\dagger = \hat{\mathcal{P}}_a\hat{B}$, which gives $[\hat{B},\hat{\mathcal{P}}_a] = \hat{\mathbb{0}}$. Thus, we have demonstrated that

$$[\hat{A},\hat{B}] = \hat{\mathbb{0}} \text{ with } \hat{A} = \sum_a a\hat{\mathcal{P}}_a \text{ implies } [\hat{B},\hat{\mathcal{P}}_a] = \hat{\mathbb{0}}. \tag{2.81a}$$

This result applies to all Hermitian operators $\hat{A}$ and $\hat{B}$. In particular, replacing $\hat{A}$ in $[\hat{A},\hat{B}]$ here with $\hat{B} = \sum_b b\hat{\mathcal{P}}_b$ and $\hat{B}$ with $\hat{\mathcal{P}}_a$, we obtain

$$[\hat{B},\hat{\mathcal{P}}_a] = \hat{\mathbb{0}} \text{ with } \hat{B} = \sum_b b\hat{\mathcal{P}}_b \text{ implies } [\hat{\mathcal{P}}_a,\hat{\mathcal{P}}_b] = \hat{\mathbb{0}}. \tag{2.81b}$$

Eqs. (2.81) show that $[\hat{\mathcal{P}}_a,\hat{\mathcal{P}}_b] = \hat{\mathbb{0}}$ indeed follows from $[\hat{A},\hat{B}] = \hat{\mathbb{0}}$. This concludes the proof.

If the projectors $\hat{\mathcal{P}}_a$ and $\hat{\mathcal{P}}_b$ commute, then their product

$$\hat{\mathcal{P}}_{ab} = \hat{\mathcal{P}}_a\hat{\mathcal{P}}_b \tag{2.82}$$

satisfies $\hat{\mathcal{P}}_{ab}^2 = \hat{\mathcal{P}}_{ab}$ and is therefore also a projector. The subspace $\mathcal{H}_{ab}$ it projects onto is the intersection[29] of $\mathcal{H}_a$ and $\mathcal{H}_b$, i.e., $\mathcal{H}_{ab} = \mathcal{H}_a \cap \mathcal{H}_b$. (Note that this intersection would not be a Hilbert space if the operators $\hat{A}$ and $\hat{B}$ were not commuting.) By construction, every

vector $|\psi\rangle$ in $\mathcal{H}_{ab}$ simultaneously obeys the equations $\hat{A}|\psi\rangle = a|\psi\rangle$ and $\hat{B}|\psi\rangle = b|\psi\rangle$, i.e., $\mathcal{H}_{ab}$ is an eigenspace of both $\hat{A}$ and $\hat{B}$.

Eqs. (2.53) and (2.82) show that projectors $\hat{\mathcal{P}}_{ab}$ with different indices are orthogonal:

$$\hat{\mathcal{P}}_{ab}\hat{\mathcal{P}}_{a'b'} = \hat{\mathcal{P}}_a\hat{\mathcal{P}}_{a'}\hat{\mathcal{P}}_b\hat{\mathcal{P}}_{b'} = \delta_{a,a'}\delta_{b,b'}\hat{\mathcal{P}}_a\hat{\mathcal{P}}_b = \delta_{a,a'}\delta_{b,b'}\hat{\mathcal{P}}_{ab}\,. \tag{2.83a}$$

Moreover, it follows from Eqs. (2.54), (2.78), and (2.82) that the identity operator can be written as a sum of these orthogonal projectors,

$$\hat{\mathbb{1}} = \hat{\mathbb{1}}^2 = \sum_a \hat{\mathcal{P}}_a \sum_b \hat{\mathcal{P}}_b = \sum_{a,b} \hat{\mathcal{P}}_a\hat{\mathcal{P}}_b = \sum_{a,b} \hat{\mathcal{P}}_{ab}\,. \tag{2.83b}$$

Eq. (2.83a) implies that the eigenspaces $\mathcal{H}_{ab}$ with different indices are orthogonal, whereas the completeness relation Eq. (2.83b) shows that the Hilbert space $\mathcal{H}$ is their direct sum,

$$\mathcal{H} = \bigoplus\nolimits_{a,b} \mathcal{H}_{ab}. \tag{2.84}$$

Therefore, if $\{|\phi_{ab;n}\rangle\}$ is an orthonormal basis for the subspace $\mathcal{H}_{ab}$, then $\bigcup_{a,b}\{|\phi_{ab;n}\rangle\}$ is an orthonormal basis for the whole Hilbert space $\mathcal{H}$. Every vector $|\phi_{ab;n}\rangle$ in this basis is an eigenvector of both $\hat{A}$ and $\hat{B}$ with eigenvalues a and b, respectively. This observation provides an alternative (and equivalent) formulation of theorem (2.79):

$$\textbf{\textit{[}}\hat{\boldsymbol{A}}\textbf{\textit{,}}\hat{\boldsymbol{B}}\textbf{\textit{]}} = \hat{\mathbf{0}} \textbf{ if and only if there exists an orthonormal basis in which every vector is an eigenvector of both } \hat{\boldsymbol{A}} \textbf{ and } \hat{\boldsymbol{B}}. \tag{2.85}$$

This property comes handy when solving eigenvalue problems involving commuting operators.

2.4.2 Compatible observables as commuting operators

It is obvious that any state vector in the simultaneous eigenspace $\mathcal{H}_{ab} = \mathcal{H}_a \cap \mathcal{H}_b$ of the commuting Hermitian operators $\hat{A}$ and $\hat{B}$ [see Sec. 2.4.1] represents a state for which measurements of observables A and B corresponding to these operators are certain to yield $A = a$ and $B = b$, respectively. This suggests that the compatibility of observables has something to do with the commutativity of the corresponding operators. In order to make a definite statement to this effect, we revisit the discussion of sequential measurements [see Sec. 1.1.2].

Consider an experiment that consists of sequential *ideal* filtering-type measurements of observables A and B,

$$\psi_i \xrightarrow[\text{yielding } A=a]{\text{ideal measurement of } A} \psi_a \xrightarrow[\text{yielding } B=b]{\text{ideal measurement of } B} \psi_{ab} \tag{2.86a}$$

[cf. Eq. (1.9)]. As discussed above [see Sec. 1.3.6 and Eq. (2.66)], ideal measurements project state vectors onto the corresponding eigenspaces. That is, the state vectors representing the states ψ_i, ψ_a, and ψ_{ab} in Eq. (2.86a) are related to each other as

$$|\psi_a\rangle \propto \hat{\mathcal{P}}_a|\psi_i\rangle, \quad |\psi_{ab}\rangle \propto \hat{\mathcal{P}}_b|\psi_a\rangle \propto \hat{\mathcal{P}}_b\hat{\mathcal{P}}_a|\psi_i\rangle \tag{2.86b}$$

(up to phase factors and normalization coefficients), where $\hat{\mathcal{P}}_a$ and $\hat{\mathcal{P}}_b$ are projectors onto the eigenspaces of operators $\hat{A}$ and $\hat{B}$, respectively.

The probability that individual systems in the initial state ψ_i pass both filters and end up in the double-filtered state ψ_{ab} can be written as

$$\text{Prob}_{\psi_i}(A = a \text{ then } B = b) = \text{Prob}_{\psi_i}(A = a)\,\text{Prob}_{\psi_a}(B = b)$$

[cf. Eq. (1.10)]. Substituting here

$$\text{Prob}_{\psi_i}(A = a) = \langle\psi_i|\hat{\mathcal{P}}_a|\psi_i\rangle = \|\hat{\mathcal{P}}_a\psi_i\|^2,$$
$$\text{Prob}_{\psi_a}(B = b) = \langle\psi_a|\hat{\mathcal{P}}_b|\psi_a\rangle = \|\hat{\mathcal{P}}_a\psi_i\|^{-2}\,\langle\psi_i|\hat{\mathcal{P}}_a\hat{\mathcal{P}}_b\hat{\mathcal{P}}_a|\psi_i\rangle$$

[see Eqs. (2.60), (2.66), and (2.86b)], we find

$$\text{Prob}_{\psi_i}(A = a \text{ then } B = b) = \langle\psi_i|\hat{\mathcal{P}}_a\hat{\mathcal{P}}_b\hat{\mathcal{P}}_a|\psi_i\rangle. \tag{2.87a}$$

Upon interchanging $A \rightleftarrows B$ and $a \rightleftarrows b$ in this expression, we get the probability that individual systems in ψ_i pass both filters for the case when observable B is measured first,

$$\text{Prob}_{\psi_i}(B = b \text{ then } A = a) = \langle\psi_i|\hat{\mathcal{P}}_b\hat{\mathcal{P}}_a\hat{\mathcal{P}}_b|\psi_i\rangle. \tag{2.87b}$$

According to our definition [see Eq. (1.11)], observables A and B are compatible if for all initial states ψ_i and for all possible measurement outcomes (i.e., eigenvalues) a and b the probabilities in the left-hand sides of Eqs. (2.87) coincide. Equating the right-hand sides of these equations and taking into account Eq. (2.5), we obtain the operator relation

$$\hat{\mathcal{P}}_a\hat{\mathcal{P}}_b\hat{\mathcal{P}}_a = \hat{\mathcal{P}}_b\hat{\mathcal{P}}_a\hat{\mathcal{P}}_b \tag{2.88a}$$

Multiplying both sides of this relation (on the right) by $\hat{\mathcal{P}}_b$ and using $\hat{\mathcal{P}}_b^2 = \hat{\mathcal{P}}_b$, we get

$$(\hat{\mathcal{P}}_a\hat{\mathcal{P}}_b)^2 = \hat{\mathcal{P}}_b\hat{\mathcal{P}}_a\hat{\mathcal{P}}_b. \tag{2.88b}$$

Because the right-hand side of this equation is Hermitian, so is the left-hand side. Taking into account that a square of a linear operator is Hermitian if and only if the operator itself is Hermitian, we obtain $\hat{\mathcal{P}}_a\hat{\mathcal{P}}_b = (\hat{\mathcal{P}}_a\hat{\mathcal{P}}_b)^\dagger = \hat{\mathcal{P}}_b\hat{\mathcal{P}}_a$, i.e., the projectors commute,

$$[\hat{\mathcal{P}}_a, \hat{\mathcal{P}}_b] = \hat{\mathbb{0}} \text{ for all } a \text{ and } b, \tag{2.89a}$$

which in turn implies [recall Eq. (2.79)] that the corresponding operators commute as well,

$$[\hat{A}, \hat{B}] = \hat{\mathbb{0}}. \tag{2.89b}$$

We have just demonstrated that two observables are compatible if and only if the corresponding Hermitian operators commute. In other words, compatible observables are represented by the commuting Hermitian operators and vice versa:

$$\textbf{compatible observables} \quad \rightleftarrows \quad \textbf{commuting Hermitian operators}. \tag{2.90}$$

The concept of a complete set of compatible observables [see Sec. 1.1.2] can now be replaced with the ***complete set of commuting Hermitian operators***. Simultaneous eigenspaces

of these operators are one-dimensional, hence any operator that commutes with all operators in the chosen complete set[6] is a function of the operators in the set.

We have seen in Sec. 1.1.2 [see Eq. (1.12)] that the compatibility of observables A and B implies that for all initial states ψ_i in Eq. (2.86a) and for all possible measurement outcomes a and b the double-filtered state ψ_{ab} inherits from ψ_a the property that $A = a$ with certainty. We will show now that for ideal measurements this is possible if and only if operators $\hat{A}$ and $\hat{B}$ representing the observables of interest commute. Indeed, by Born's rule (2.60), $\mathrm{Prob}_{\psi_{ab}}(A = a) = 1$ if and only if the state vector $|\psi_{ab}\rangle$ satisfies $\hat{\mathcal{P}}_a|\psi_{ab}\rangle = |\psi_{ab}\rangle$. Taking into account the second equation in (2.86b), we obtain

$$\hat{\mathcal{P}}_a\hat{\mathcal{P}}_b\hat{\mathcal{P}}_a|\psi_i\rangle = \hat{\mathcal{P}}_b\hat{\mathcal{P}}_a|\psi_i\rangle. \tag{2.91a}$$

Because vector $|\psi_i\rangle$ here is an arbitrary normalized vector in $\mathcal{H}$, Eq. (2.91a) is equivalent [see Eq. (2.2) and the remark after Eq. (2.5)] to the operator relation

$$\hat{\mathcal{P}}_a\hat{\mathcal{P}}_b\hat{\mathcal{P}}_a = \hat{\mathcal{P}}_b\hat{\mathcal{P}}_a. \tag{2.91b}$$

This relation shows [cf. the derivation of Eq. (2.81a)] that the product $\hat{\mathcal{P}}_b\hat{\mathcal{P}}_a$ is an Hermitian operator, leading again to Eqs. (2.89).

Thus, for ideal sequential measurements (2.86) our definition of the compatibility of observables

$$\mathrm{Prob}_{\psi_i}(A = a \text{ then } B = b) \;=\; \mathrm{Prob}_{\psi_i}(B = b \text{ then } A = a) \tag{2.92a}$$

[see Eq. (1.11)] and the relation

$$\mathrm{Prob}_{\psi_{ab}}(A = a) = \mathrm{Prob}_{\psi_{ab}}(B = b) = 1 \tag{2.92b}$$

that follows from it [see Eq. (1.12)] are both equivalent to $[\hat{A},\hat{B}] = \hat{\mathbb{0}}$. Accordingly, for such measurements Eq. (2.92b) implies Eq. (2.92a), as mentioned in Sec. 1.1.2. In other words, Eq. (2.92b) can be taken as an alternative definition of the compatibility of observables.

Sequential filtering-type measurements [see Eqs. (2.86)] provide a recipe for *preparing* a quantum state ψ_{ab} for which outcomes of measurements of compatible observables A and B can be predicted with certainty. Moreover, since $|\psi_{ab}\rangle \propto \hat{\mathcal{P}}_{ab}|\psi_i\rangle$ [recall (2.86b)], where $\hat{\mathcal{P}}_{ab} = \hat{\mathcal{P}}_a\hat{\mathcal{P}}_b$ [see Eq. (2.82)] is the projector onto the simultaneous eigenspace $\mathcal{H}_{ab}$ of operators $\hat{A}$ and $\hat{B}$, the state vector representing the double-filtered state ψ_{ab} is independent (up to a phase factor) of the order in which the filters are applied.

It is easy to see that the sequential filtering is not the only possible route to ψ_{ab}. Indeed, any function $\hat{F} = \hat{f}(\hat{A},\hat{B})$ of commuting Hermitian operators $\hat{A}$ and $\hat{B}$ admits the spectral decomposition

$$\hat{F} = \hat{f}(\hat{A},\hat{B}) = \sum_{a,b} f(a,b)\,\hat{\mathcal{P}}_{ab}, \tag{2.93}$$

If all coefficients $f(a,b)$ here are real, then operator $\hat{F}$ is Hermitian, hence [recall Eq. (2.58)] it represents an observable F that can in principle be measured. An ideal measurement of this observable yielding $F = f(a,b)$ projects the initial state vector onto $\mathcal{H}_{ab}$, thereby producing the same filtered state ψ_{ab} as above. Moreover, the probability of getting this

outcome $\mathrm{Prob}_{\psi_i}(F=f(a,b))$, i.e., the probability that individual systems in ψ_i end up in ψ_{ab}, coincides with the probabilities (2.92a). In this sense, compatible observables can be measured simultaneously.

2.4.3 Incompatible observables and uncertainty relation

With the help of Eq. (2.61), the square of the uncertainty of observable A in the state ψ [see Eqs. (1.5)] can be written in terms of the corresponding operator $\hat{A}$ as

$$(\Delta_\psi A)^2 = \big\langle \big[A-\langle A\rangle\big]^2\big\rangle = \langle\psi|\big(\hat{A}-\langle A\rangle_\psi \hat{\mathbb{1}}\big)^2|\psi\rangle\,, \quad \langle A\rangle_\psi = \langle\psi|\hat{A}|\psi\rangle. \tag{2.94a}$$

It is easy to see that the uncertainty vanishes if and only if the state vector $|\psi\rangle$ belongs to one of the eigenspaces of $\hat{A}$. Indeed, $\Delta_\psi A$ can be formally written as a norm of a vector,

$$\Delta_\psi A = \|\alpha\|, \quad |\alpha\rangle = \big(\hat{A}-\langle A\rangle_\psi \hat{\mathbb{1}}\big)|\psi\rangle. \tag{2.94b}$$

Therefore [recall Eq. (1.39a)], $\Delta_\psi A = 0$ if and only if $|\alpha\rangle = |\mathrm{null}\rangle$, or, equivalently, if and only if $\hat{A}|\psi\rangle = \langle A\rangle_\psi|\psi\rangle$. Comparison with Eq. (2.40) then shows that in this case the expectation value $\langle A\rangle_\psi$ is an eigenvalue of $\hat{A}$ and that vector $|\psi\rangle$ belongs to the eigenspace corresponding to this eigenvalue.

Consider now a pair of *incompatible* observables A and B. According to Eq. (2.90), the corresponding Hermitian operators $\hat{A}$ and $\hat{B}$ must have a non-vanishing commutator

$$[\hat{A},\hat{B}] = i\hat{C}, \quad \hat{C} = \hat{C}^\dagger \neq \hat{\mathbb{0}}. \tag{2.95}$$

We are interested in a generic state vector $|\psi\rangle$ for which vectors

$$|\alpha\rangle = \big(\hat{A}-\langle A\rangle_\psi \hat{\mathbb{1}}\big)|\psi\rangle, \quad |\beta\rangle = \big(\hat{B}-\langle B\rangle_\psi \hat{\mathbb{1}}\big)|\psi\rangle \tag{2.96a}$$

are different from $|\mathrm{null}\rangle$, so that (see above) both A and B have finite uncertainties

$$\Delta_\psi A = \|\alpha\| > 0\,, \quad \Delta_\psi B = \|\beta\| > 0\,. \tag{2.96b}$$

The Schwartz inequality (1.36) and Eqs. (2.96b) show that

$$(\Delta_\psi A)(\Delta_\psi B) = \|\alpha\|\|\beta\| \geq |\langle\alpha|\beta\rangle| \geq |\mathrm{Im}\langle\alpha|\beta\rangle|\,. \tag{2.97a}$$

Substituting here

$$\mathrm{Im}\langle\alpha|\beta\rangle = \frac{1}{2i}\big[\langle\alpha|\beta\rangle - \langle\beta|\alpha\rangle\big] = \frac{1}{2i}\langle\psi|[\hat{A},\hat{B}]|\psi\rangle = \frac{1}{2}\langle\psi|\hat{C}|\psi\rangle = \frac{1}{2}\langle C\rangle_\psi \tag{2.97b}$$

[see Eqs. (2.95) and (2.96a)], we arrive at the ***uncertainty relation***

$$(\Delta_\psi A)(\Delta_\psi B) \geq \frac{1}{2}\big|\langle C\rangle_\psi\big|. \tag{2.98}$$

This inequality relates the results of three experiments in which mutually incompatible observables A, B, and C are measured on separate subensembles [cf. the remark after Eq. (1.21)].

It is natural to ask for which states ψ the inequality in Eq. (2.98) becomes an equality, i.e., when the product of uncertainties $(\Delta_\psi A)(\Delta_\psi B)$ attains the minimal possible value. For this to happen, both "$\geq$" in Eq. (2.97a) must turn to "$=$". Accordingly, vectors $|\alpha\rangle$ and $|\beta\rangle$ should be linearly dependent [see the remark after Eq. (1.36)], and their inner product $\langle\alpha|\beta\rangle$ should be imaginary. This leads to the condition

$$|\alpha\rangle = ic|\beta\rangle, \quad c = c^*, \tag{2.99}$$

which can be satisfied (if at all) only for some very special state vectors $|\psi\rangle$. We will encounter such minimal uncertainty states later.

2.5 Unitary operators and transformations

2.5.1 Unitary operators

Unitary operators are linear operators that satisfy

$$\hat{U}^\dagger \hat{U} = \hat{\mathbb{1}}, \quad [\hat{U}, \hat{U}^\dagger] = \hat{\mathbb{0}}. \tag{2.100}$$

It is easy to see that products of unitary operators are also unitary, whereas their linear combinations are not.

The first equation in (2.100) and Eqs. (1.41) and (2.17b) show that for all $|\psi\rangle$ in $\mathcal{H}$

$$\|\hat{U}\psi\| = \sqrt{\langle \hat{U}\psi|\hat{U}\psi\rangle} = \sqrt{\langle\psi|\hat{U}^\dagger\hat{U}|\psi\rangle} = \sqrt{\langle\psi|\psi\rangle} = \|\psi\|, \tag{2.101}$$

hence unitary operators have a unit norm,[45] $\|\hat{U}\| = 1$ [recall Eq. (2.19)] and are therefore bounded. From this point of view, the dimension of the Hilbert space is not a serious handicap,[46] and in this section we will allow it to be infinite unless noted otherwise.

Unitary operators are close relatives of the Hermitian operators discussed in Sec. 2.2. Indeed, if $\hat{\Omega}$ is Hermitian, then $\hat{U} = e^{i\hat{\Omega}}$ is a unitary operator with the spectral decomposition [see the last equation in (2.55)]

$$\hat{U} = e^{i\hat{\Omega}} = \sum_\omega e^{i\omega}\,\mathcal{P}_\omega, \tag{2.102}$$

where ω are eigenvalues of $\hat{\Omega}$ and $\hat{\mathcal{P}}_\omega$ are projectors onto the corresponding eigenspaces. We will show now that any unitary operator can be written in this form.

Consider the eigenvalue problem

$$\hat{U}|\psi\rangle = u|\psi\rangle. \tag{2.103a}$$

[45] Unitary operators are not the only ones that have a unit norm. Indeed, any Hermitian operator with eigenvalues satisfying $\max\{|a|\} = 1$ (as it is the case for any projector operator) also has this property.

[46] In particular, convergence[17,19] of the series $\sum_{n=1}^\infty \psi_n|\phi_n\rangle$ in an infinite-dimensional Hilbert space to the limit vector $|\psi\rangle$ implies that the series $\sum_{n=1}^\infty \psi_n \hat{U}|\phi_n\rangle$ converges to $\hat{U}\big(\sum_{n=1}^\infty \psi_n|\phi_n\rangle\big) = \hat{U}|\psi\rangle$. Indeed, taking into account Eq. (2.101), we obtain

$$\big\|\hat{U}|\psi\rangle - \textstyle\sum_{m=1}^{m=n}\psi_m\hat{U}|\phi_m\rangle\big\| = \big\|\hat{U}\big(|\psi\rangle - \sum_{m=1}^{m=n}\psi_m|\phi_m\rangle\big)\big\| = \big\||\psi\rangle - \sum_{m=1}^{m=n}\psi_m|\phi_m\rangle\big\| \xrightarrow{n\to\infty} 0.$$

Comparison of $\|\hat{U}\psi\| = \|u\psi\| = |u|\|\psi\|$ [recall Eq. (1.39b)] with Eq. (2.101) shows that eigenvalues of $\hat{U}$ satisfy $|u| = 1$ and can therefore be parametrized as

$$u = e^{i\omega}, \quad \omega = \omega^*, \quad 0 \le \omega < 2\pi, \tag{2.103b}$$

in agreement with Eq. (2.102). Vectors $|\varphi_\omega\rangle$ satisfying the equation $\hat{U}|\varphi_\omega\rangle = e^{i\omega}|\varphi_\omega\rangle$ form an eigenspace $\mathcal{H}_\omega$, a subspace of $\mathcal{H}$ and a Hilbert space in its own right. Taking into account that $|\varphi_\omega\rangle = e^{-i\omega}\hat{U}|\varphi_\omega\rangle$, we find for all vectors $|\varphi_\omega\rangle$ in $\mathcal{H}_\omega$ and all $|\varphi_{\omega'}\rangle$ in $\mathcal{H}_{\omega'}$

$$\langle\varphi_\omega|\varphi_{\omega'}\rangle = e^{i(\omega-\omega')}\langle\varphi_\omega|\hat{U}^\dagger\hat{U}|\varphi_{\omega'}\rangle = e^{i(\omega-\omega')}\langle\varphi_\omega|\varphi_{\omega'}\rangle,$$

which shows that $\langle\varphi_\omega|\varphi_{\omega'}\rangle = 0$ for $\omega \neq \omega'$. Accordingly, eigenspaces corresponding to different eigenvalues are orthogonal, hence the projectors onto these eigenspaces satisfy

$$\hat{\mathcal{P}}_\omega\hat{\mathcal{P}}_{\omega'} = \delta_{\omega,\omega'}\hat{\mathcal{P}}_\omega\,, \tag{2.104a}$$

as do projectors onto the orthogonal eigenspaces of $\hat{\Omega}$ that appear in Eq. (2.102).

It remains to show that projectors $\hat{\mathcal{P}}_\omega$ sum up to the identity operator, i.e.,

$$\sum_\omega \hat{\mathcal{P}}_\omega = \hat{\mathbb{1}}. \tag{2.104b}$$

To this end, we write $\hat{U}$ as

$$\hat{U} = \hat{V}_1 + i\hat{V}_2, \quad \hat{V}_1 = \frac{1}{2}\big(\hat{U} + \hat{U}^\dagger\big), \quad \hat{V}_2 = \frac{1}{2i}\big(\hat{U} - \hat{U}^\dagger\big). \tag{2.105}$$

The operators $\hat{V}_{1,2}$ are obviously Hermitian. Moreover, the second equation in (2.100) implies that these operators commute, $[\hat{V}_1,\hat{V}_2] = \hat{\mathbb{0}}$. Therefore [see Sec. 2.4.1], these operators admit spectral decompositions

$$\hat{V}_1 = \sum_{v_1,v_2} v_1\hat{\mathcal{P}}_{v_1,v_2}\,, \quad \hat{V}_2 = \sum_{v_1,v_2} v_2\hat{\mathcal{P}}_{v_1,v_2}\,, \tag{2.106a}$$

where projectors $\hat{\mathcal{P}}_{v_1,v_2}$ onto the simultaneous eigenspaces of $\hat{V}_{1,2}$ satisfy

$$\hat{\mathcal{P}}_{v_1,v_2}\hat{\mathcal{P}}_{v_1',v_2'} = \delta_{v_1,v_1'}\delta_{v_2,v_2'}\hat{\mathcal{P}}_{v_1,v_2}\,, \quad \sum_{v_1,v_2}\hat{\mathcal{P}}_{v_1,v_2} = \hat{\mathbb{1}} \tag{2.106b}$$

[see Eqs. (2.83)]. Substitution of Eqs. (2.106a) into the first equation in (2.105) gives

$$\hat{U} = \sum_{v_1,v_2}(v_1 + iv_2)\hat{\mathcal{P}}_{v_1,v_2}\,, \tag{2.107a}$$

hence the simultaneous eigenspaces of operators $\hat{V}_{1,2}$ are also eigenspaces of $\hat{U}$ with eigenvalues $u = v_1 + iv_2$. Comparison with Eq. (2.103b) then shows that the eigenvalues $v_{1,2}$ are in one-to-one correspondence with the parameter ω,

$$v_1(\omega) = \operatorname{Re} u = \cos\omega, \quad v_2(\omega) = \operatorname{Im} u = \sin\omega. \tag{2.107b}$$

This observation allows us to label the projectors by ω instead of $v_{1,2}$:

$$\hat{\mathcal{P}}_{v_1,v_2} = \hat{\mathcal{P}}_{v_1(\omega),v_2(\omega)} \equiv \hat{\mathcal{P}}_\omega. \tag{2.107c}$$

After such relabeling Eqs. (2.106b) turn to Eqs. (2.104). Accordingly, the operator $\hat{U}$ indeed admits a spectral decomposition, and this decomposition has the form of Eq. (2.102)

2.5.2 Unitary transformations

Consider a transformation of the Hilbert space defined by the relation

$$|\psi\rangle \longrightarrow |\psi'\rangle = \hat{U}|\psi\rangle, \tag{2.108}$$

where $\hat{U}$ is a unitary operator. Any such ***unitary transformation*** is, obviously, one-to-one and invertible (indeed, $|\psi\rangle = \hat{\mathbb{1}}|\psi\rangle = \hat{U}^\dagger\hat{U}|\psi\rangle = \hat{U}^\dagger|\psi'\rangle$). Moreover, it leaves the inner product of any two vectors invariant,

$$\langle\varphi'|\psi'\rangle = \langle\hat{U}\varphi|\hat{U}\psi\rangle = \langle\varphi|\hat{U}^\dagger\hat{U}|\psi\rangle = \langle\varphi|\psi\rangle. \tag{2.109}$$

Action of $\hat{U}$ on all vectors in an orthonormal basis $\{|\phi_n\rangle\}$ generates a transformed set of vectors $\{|\phi'_n\rangle\}$ in which

$$|\phi'_n\rangle = \hat{U}|\phi_n\rangle. \tag{2.110a}$$

Eq. (2.109) implies that the transformed set is orthonormal,

$$\langle\phi'_m|\phi'_n\rangle = \langle\phi_m|\phi_n\rangle = \delta_{m,n}. \tag{2.110b}$$

In view of Eq. (2.110a), the operator $\hat{U}$ can be written in Dirac's notations as

$$\hat{U} = \sum_n |\phi'_n\rangle\langle\phi_n| \tag{2.111}$$

[cf. Eqs. (2.23)]. Substituting $\hat{U}$ in this form into $\hat{U}\hat{U}^\dagger$, we obtain

$$\hat{U}\hat{U}^\dagger = \sum_{m,n} |\phi'_m\rangle\underbrace{\langle\phi_m|\phi_n\rangle}_{\delta_{m,n}}\langle\phi'_n| = \sum_n |\phi'_n\rangle\langle\phi'_n|. \tag{2.112}$$

In a finite-dimensional Hilbert space the set $\{|\phi'_n\rangle\}$ contains $\mathcal{N} = \dim\{\mathcal{H}\}$ vectors and thus is guaranteed to be a basis [see Sec. 1.2.1], so that the sum in the right-hand side of Eq. (2.112) coincides with the identity operator. Therefore, for finite $\mathcal{N}$ the relation $\hat{U}^\dagger\hat{U} = \hat{\mathbb{1}}$ that led to Eq. (2.110b) also implies $\hat{U}\hat{U}^\dagger = \hat{\mathbb{1}}$, making the second equation in (2.100) redundant. On the contrary, in an infinite-dimensional Hilbert space merely having an infinite number of vectors in an orthonormal set does not make it complete. However, by virtue of the second equation in (2.100), the left-hand side of Eq. (2.112) equals the identity operator, and Eq. (2.112) yields the completeness relation for the set $\{|\phi'_n\rangle\}$.

We have just shown that a unitary transformation applied to an orthonormal basis generates another orthonormal basis. Conversely, it is easy to see that any linear operator $\hat{U}$ *defined* by Eq. (2.111) with vectors $|\phi_n\rangle$ and $|\phi'_n\rangle$ taken from two orthonormal basis sets satisfies $\hat{U}\hat{U}^\dagger = \hat{U}^\dagger\hat{U} = \hat{\mathbb{1}}$, i.e., is unitary. Accordingly, any change of an orthonormal basis for a Hilbert space is facilitated by a unitary transformation.

Any orthonormal basis can be viewed as a set of eigenvectors of some Hermitian operator:

$$\hat{A} = \sum_n |\phi_n\rangle\, a_n\langle\phi_n|, \quad \hat{A}|\phi_n\rangle = a_n|\phi_n\rangle. \tag{2.113a}$$

(We do not assume the eigenvalues a_n to be all different.) Acting with a unitary operator $\hat{U}$ on both sides of the second equation here and inserting the identity operator $\hat{\mathbb{1}} = \hat{U}^\dagger \hat{U}$ between $\hat{A}$ and $|\phi_n\rangle$ in the left-hand side, we obtain

$$(\hat{U}\hat{A}\hat{U}^\dagger)\hat{U}|\phi_n\rangle = a_n \hat{U}|\phi_n\rangle. \tag{2.113b}$$

Therefore, the ***transformed eigenvectors*** $|\phi_n'\rangle = \hat{U}|\phi_n\rangle$ [cf. Eq. (2.110a)] are eigenvectors of the ***transformed operator***

$$\hat{A}' = \hat{U}\hat{A}\hat{U}^\dagger \tag{2.114a}$$

with the same eigenvalues as the original eigenvectors in Eqs. (2.113a), i.e.,

$$\hat{A}'|\phi_n'\rangle = a_n|\phi_n'\rangle. \tag{2.114b}$$

This observation is often utilized to simplify the solution of the eigenvalue problem: instead of working with $\hat{A}$, one may first look for such $\hat{U}$ that the eigenvalue problem for the transformed operator $\hat{A}'$ is easier to solve.

It is easy to see that the unitary-transformed operators satisfy

$$(\hat{A}^\dagger)' = (\hat{A}')^\dagger, \quad (\hat{A}+\hat{B})' = \hat{A}' + \hat{B}', \quad (\hat{A}\hat{B})' = \hat{A}'\hat{B}', \quad [\hat{f}(\hat{A})]' = \hat{f}(\hat{A}'). \tag{2.115a}$$

Thus, if an operator is Hermitian, then the corresponding unitary-transformed operator is Hermitian as well. If two operators commute, then so do the transformed operators. Furthermore, it follows from Eqs. (2.32a), (2.100), and (2.114a) that[37]

$$\operatorname{tr}\hat{A}' = \operatorname{tr}[\hat{U}\hat{A}\hat{U}^\dagger] = \operatorname{tr}[\hat{U}^\dagger\hat{U}\hat{A}] = \operatorname{tr}\hat{A}, \tag{2.115b}$$

$$\det\hat{A}' = \det[\hat{U}\hat{A}\hat{U}^\dagger] = \det[\hat{U}^\dagger\hat{U}\hat{A}] = \det\hat{A}, \tag{2.115c}$$

i.e., traces and determinants are not affected by unitary transformations, as expected. (Indeed, as discussed above, a unitary transformation amounts to a change of the basis, and traces and determinants are independent of the choice of the basis.) Moreover, the probabilities and expectation values are invariant under unitary transformations provided that *both* the state vector and the operator representing the observable of interest are transformed as prescribed by Eqs. (2.108) and (2.114a), e.g.,

$$\langle\psi'|\hat{f}(\hat{A}')|\psi'\rangle = \langle\psi|\hat{f}(\hat{A})|\psi\rangle \tag{2.116}$$

[see Eq. (2.61) and the last equation in (2.115a)].

2.5.3 Symmetry transformations

Imagine a laboratory completely isolated from the outside world. Because space is uniform and isotropic, no experiment carried out inside such laboratory will be capable of revealing its position or orientation. Similarly, because time is uniform, the results of any such experiment will depend on its duration, but not on the reading on the internal laboratory clock at the start of the experiment.

Consider now the state operator $\hat{\rho}_\psi = |\psi\rangle\langle\psi|$, where $|\psi\rangle$ is the state vector. As discussed in Sec. 2.3.1 [see Eqs. (2.63) and (2.64)], this operator represents a pure quantum state ψ. (Recall that the correspondence $\psi \longleftrightarrow \hat{\rho}_\psi$ is one-to-one.) On the other hand, since the state operator is Hermitian, it also represents an observable that can in principle be measured [recall Eq. (2.58)]. Let $\hat{\rho}_\psi = |\psi\rangle\langle\psi|$ and $\hat{\rho}_\varphi = |\varphi\rangle\langle\varphi|$ be two arbitrary state operators. Then the quantity

$$\mathrm{tr}\,(\hat{\rho}_\psi \hat{\rho}_\varphi) = \langle\psi|\hat{\rho}_\varphi|\psi\rangle = \langle\varphi|\hat{\rho}_\psi|\varphi\rangle = \big|\langle\varphi|\psi\rangle\big|^2 \tag{2.117}$$

[cf. Eqs. (2.61) and (2.65)] can be interpreted either as the expectation value of the observable represented by $\hat{\rho}_\varphi$ in the state represented by $\hat{\rho}_\psi$, or as the expectation value of the observable $\hat{\rho}_\psi$ in the state $\hat{\rho}_\varphi$. Provided that state preparations and measurements are carried out in total isolation from the outside world (e.g., inside an isolated laboratory), the expectation values (2.117) will not depend on the position, orientation, or time of measurement.

This discussion brings about the concept of a ***symmetry transformation***. It is a map of the Hilbert space onto itself that amounts to a one-to-one correspondence between quantum states ψ and ψ' or, equivalently, between the corresponding state operators $\hat{\rho}_\psi$ and $\hat{\rho}_{\psi'}$,

$$\hat{\rho}_\psi = |\psi\rangle\langle\psi| \quad \underset{\text{correspondence}}{\overset{\text{one-to-one}}{\longleftrightarrow}} \quad \hat{\rho}_{\psi'} = |\psi'\rangle\langle\psi'|, \tag{2.118a}$$

that leaves expectation values (2.117) invariant, i.e., such that

$$\big|\langle\varphi|\psi\rangle\big| = \big|\langle\varphi'|\psi'\rangle\big| \text{ for all state vectors } |\varphi\rangle \text{ and } |\psi\rangle. \tag{2.118b}$$

We are interested in transformations that reflect the fundamental symmetries of space and time, i.e., rotations and displacements in space and translations in time. It is clear on physical grounds that these transformations are ***continuous*** in the sense that state operators are continuous functions[47] of some dimensionless parameter ξ. It is always possible to choose this parameter so that the state operator $\hat{\rho}_{\psi'}$ in the right-hand side of Eq. (2.118a) becomes

$$\hat{\rho}_{\psi(\xi)} = |\psi(\xi)\rangle\langle\psi(\xi)|, \tag{2.119}$$

whereas $\hat{\rho}_\psi$ in the left-hand side of Eq. (2.118a) coincides with $\hat{\rho}_{\psi(\xi)}$ at $\xi = 0$.

The continuity of the dependence of $\hat{\rho}_{\psi(\xi)}$ on ξ leads to a drastic simplification: it can be shown rigorously that all such continuous symmetry transformations are unitary,[48]

$$|\psi(\xi)\rangle = \hat{U}(\xi)|\psi(0)\rangle \quad \text{(up to phase factors)}, \tag{2.120}$$

where the unitary operator $\hat{U}(\xi)$ is unique (up to a phase factor) for a given correspondence $\hat{\rho}_{\psi(0)} \longleftrightarrow \hat{\rho}_{\psi(\xi)}$ [see Eqs. (2.118a) and (2.119)]. For the symmetry transformations we consider the parameter ξ and the phase factors of operators $\hat{U}(\xi)$ can be chosen so that these operators have the ***group property***

$$\hat{U}(\xi)\hat{U}(\xi') = \hat{U}(\xi + \xi'). \tag{2.121}$$

[47] The dependence $\hat{A}(\xi)$ is continuous if matrix elements of $\hat{A}(\xi)$ in any basis are continuous functions of ξ.

[48] This observation is one of the consequences of the more general ***Wigner's theorem***. Note that there is no *a priori* reason for the symmetry transformation defined by Eqs. (2.118) to be linear, let alone unitary. It is remarkable that the physics-motivated requirement of continuity (2.119) is sufficient to enforce the unitarity.

This, of course, implies that operators $\hat{U}(\xi)$ at different ξ commute, i.e., $[\hat{U}(\xi),\hat{U}(\xi')]=\hat{\mathbb{0}}$.

For $\xi=\xi'=0$ Eq. (2.121) reduces to $\hat{U}^2(0)=\hat{U}(0)$. Multiplying both sides of this equation by $\hat{U}^\dagger(0)$ and taking into account the unitarity of $\hat{U}(0)$, we obtain

$$\hat{U}(0)=\hat{\mathbb{1}}. \tag{2.122a}$$

Setting $\xi'=-\xi$ in Eq. (2.121) and using Eq. (2.122a), we get $\hat{U}(\xi)\hat{U}(-\xi)=\hat{\mathbb{1}}$. Multiplication of both sides of this equation (on the left) by $U^\dagger(\xi)$ yields

$$U^\dagger(\xi)=\hat{U}(-\xi). \tag{2.122b}$$

For small ξ, the operator $\hat{U}(\xi)$ can be expanded in Taylor series. In view of Eq. (2.122a), the first two terms in this expansion can be written as

$$\hat{U}(\xi)=\hat{\mathbb{1}}-i\xi\hat{K}+\ldots, \tag{2.123a}$$

where $\hat{K}$ is an independent of ξ operator called the ***generator*** of the transformation. The reminder of the series (indicated by the ellipsis) includes contributions of order ξ^2 and higher. Substitution of this expansion into either the unitarity condition $\hat{U}^\dagger(\xi)\hat{U}(\xi)=\hat{\mathbb{1}}$ or into Eq. (2.122b) shows that the generator is Hermitian,

$$\hat{K}=\hat{K}^\dagger. \tag{2.123b}$$

For finite ξ, we write with the help of Eqs. (2.121) and (2.123a)

$$\hat{U}(\xi+\delta\xi)=\hat{U}(\delta\xi)\hat{U}(\xi)=\big[\hat{\mathbb{1}}-i\delta\xi\hat{K}+\ldots\big]\hat{U}(\xi), \tag{2.124a}$$

which can be converted to the first-order differential equation

$$\frac{d}{d\xi}\hat{U}(\xi)=-i\hat{K}\hat{U}(\xi). \tag{2.124b}$$

Note that because $\hat{U}(\delta\xi)$ commutes with $\hat{U}(\xi)$, so does $\hat{K}$: $[\hat{K},\hat{U}(\xi)]=0$. Solution of Eq. (2.124b) subject to the condition (2.122a) is unique and reads [recall Eqs. (2.11)]

$$\hat{U}(\xi)=e^{-i\xi\hat{K}}. \tag{2.125}$$

Note that Eq. (2.120) and the group property (2.121) remain intact if the generator $\hat{K}$ in Eq. (2.125) is replaced with $\hat{K}+k\hat{\mathbb{1}}$, where k is a real number. (Such replacement is equivalent to multiplying $\hat{U}(\xi)$ by the phase factor $e^{-ik\xi}$.) This leaves room for imposing additional constraints on the generator $\hat{K}$.

Because the transformation $\hat{\rho}_{\psi(0)}\longrightarrow\hat{\rho}_{\psi(\xi)}$ amounts to a genuine change of the quantum state, an outside observer will see the probabilities and expectation values that explicitly depend on ξ, e.g.,

$$\langle f(A)\rangle_\xi=\mathrm{tr}\big[\hat{\rho}_{\psi(\xi)}\hat{f}(\hat{A})\big]=\langle\psi(\xi)|\hat{f}(\hat{A})|\psi(\xi)\rangle \tag{2.126a}$$

[cf. Eqs. (2.61) and (2.65)], where $\hat{A}$ is an independent of ξ Hermitian operator representing observable A. With the help of Eq. (2.17b) and the last equation in (2.115a), this expectation value can be rewritten as

$$\langle f(A)\rangle_\xi=\langle\psi(0)|\hat{U}^\dagger(\xi)f(\hat{A})\hat{U}(\xi)|\psi(0)\rangle=\langle\psi(0)|\hat{f}\big[\hat{U}^\dagger(\xi)\hat{A}\hat{U}(\xi)\big]|\psi(0)\rangle. \tag{2.126b}$$

Therefore, instead of transforming the state vectors as in Eqs. (2.120) and (2.126a), one can transform the operators and leave the state vectors as they are,

$$\hat{A} \longrightarrow \hat{A}(\xi) = \hat{U}^\dagger(\xi)\hat{A}\hat{U}(\xi), \quad |\psi(0)\rangle \longrightarrow |\psi(0)\rangle. \tag{2.127a}$$

Differentiating $\hat{A}(\xi)$ with the help of Eq. (2.124b), we obtain the equation

$$\frac{d}{d\xi}\hat{A}(\xi) = i[\hat{K}, \hat{A}(\xi)]. \tag{2.127b}$$

The commutator in the right-hand side here can be also written as $\hat{U}^\dagger(\xi)[\hat{K}, \hat{A}]\hat{U}(\xi)$. Therefore, if operator $\hat{A}$ commutes with $\hat{K}$, then so does $\hat{A}(\xi)$. Eq. (2.127b) then shows that in this case $\hat{A}(\xi)$ [and thus the expectation value $\langle f(A)\rangle_\xi$] is independent of ξ, i.e., observable A is invariant under the transformation considered.

We just described two points of view (sometimes called ***pictures***) on continuous symmetry transformations. In the first picture, it is the state vector that is transformed [see Eqs. (2.120) and (2.126a)]. In the second picture [see Eqs. (2.127a)] one transforms the operator representing the observable of interest. Although the two points of view are equivalent, the second one is more suitable for our purposes. Indeed, whereas state vectors do not have classical analogues, Hermitian operators do - these are the corresponding observables. By comparing equations such as Eq. (2.127b) with their classical counterparts one can associate generators of continuous symmetry transformations with the objects playing the same role in Classical Mechanics. It is in this context the operators representing linear momentum, angular momentum, and energy arise in quantum theory.

Of course, with the unitary operator (2.125) at hand, one can also carry out a unitary transformation of the type discussed in Sec. 2.5.2. In this transformation *both* the operators and the state vectors are transformed according to

$$\hat{A} \longrightarrow \hat{A}(\xi) = \hat{U}(\xi)\hat{A}U^\dagger(\xi), \quad |\psi\rangle \longrightarrow |\psi(\xi)\rangle = \hat{U}(\xi)|\psi\rangle \tag{2.128}$$

[see Eqs. (2.114a) and (2.120)]. As discussed in detail in Sec. 2.5.2, this transformation amounts to merely changing the basis for the Hilbert space and does not affect the probabilities and expectation values [recall Eq. (2.116)]. Transformations of this type are sometimes referred to as the ***passive transformations*** to distinguish them from the ***active transformations*** described by Eqs. (2.126) and (2.127). Notice that in view of Eq. (2.122b) the operators $\hat{A}(\xi)$ defined in Eqs. (2.127a) and (2.128) are related to each other as

$$\hat{A}(\xi)\big|_{\text{active}} = \hat{A}(-\xi)\big|_{\text{passive}}. \tag{2.129}$$

To avoid confusion, henceforth we reserve the notation $\hat{A}(\xi)$ for $\hat{A}(\xi)\big|_{\text{active}} = \hat{U}^\dagger(\xi)\hat{A}U(\xi)$ and write $\hat{A}(\xi)\big|_{\text{passive}}$ either as $\hat{U}(\xi)\hat{A}U^\dagger(\xi)$ or as $\hat{A}(-\xi)$ [see Eqs. (2.127a), (2.128), and (2.129)].

3 Rotations and Angular Momentum

3.1 Rotation transformation

3.1.1 Vectors and rotations

By Euler's rotation theorem, any displacement of a rigid body with a fixed point is equivalent to a rotation about some axis running through that point. In order to understand such rotations, it is sufficient to consider what happens with the *position vector* $\mathbf{r}$ viewed as a directed line segment [49] whose initial point coincides with the fixed point of the rotation.

It is convenient to introduce a unit dimensionless vector $\mathbf{n}$ specifying the direction of the axis of rotation (the axis vector) and to assign a positive (negative) sign to the rotation angle θ according to the usual right (left) screw rule. With this convention, it is obvious that a rotation by θ about $\mathbf{n}$ produces exactly the same result as a rotation by $-\theta$ about $-\mathbf{n}$. This is possible only if the rotated vector $\mathbf{r}_{\theta\mathbf{n}}$ depends on the *angle-axis vector* $\theta\mathbf{n}$ rather than on θ and $\mathbf{n}$ separately.

It is also obvious that two consecutive rotations *about the same axis* (this condition is crucial!) are equivalent to a single rotation by a combined angle, e.g.,

$$\mathbf{r} \xrightarrow[\text{about } \mathbf{n}]{\text{rotation by } \theta} \mathbf{r}_{\theta\mathbf{n}} \xrightarrow[\text{about } \mathbf{n}]{\text{rotation by } \delta\theta} \mathbf{r}_{(\theta+\delta\theta)\mathbf{n}}. \tag{3.1}$$

It is not difficult to show that to first order in $\delta\theta$

$$\mathbf{r}_{(\theta+\delta\theta)\mathbf{n}} = \mathbf{r}_{\theta\mathbf{n}} + \delta\theta\,\mathbf{n}\times\mathbf{r}_{\theta\mathbf{n}} + \dots \tag{3.2a}$$

This relation can be cast in the form of the first-order differential equation

$$\frac{d}{d\theta}\mathbf{r}_{\theta\mathbf{n}} = \mathbf{n}\times\mathbf{r}_{\theta\mathbf{n}}, \tag{3.2b}$$

whose solution subject to the initial condition $\mathbf{r}_{\theta\mathbf{n}}\big|_{\theta=0} = \mathbf{r}$ is unique and reads [50]

$$\mathbf{r}_{\theta\mathbf{n}} = \mathbf{n}(\mathbf{n}\cdot\mathbf{r}) + \big[\mathbf{r} - \mathbf{n}(\mathbf{n}\cdot\mathbf{r})\big]\cos\theta + (\mathbf{n}\times\mathbf{r})\sin\theta. \tag{3.2c}$$

Note that despite its appearance, the right-hand side of this equation is a function of $\boldsymbol{\theta} = \theta\mathbf{n}$; this can be made explicit by writing θ and $\mathbf{n}$ as $|\boldsymbol{\theta}|$ and $\boldsymbol{\theta}/|\boldsymbol{\theta}|$, respectively.

[49] A ***directed line segment*** is a straight line connecting the initial point with the endpoint. It can be visualized as an arrow with the tip at the endpoint. Two such directed line segments are said to be *equivalent* if they have the same length and orientation. A ***vector*** is a geometric object representing equivalent directed line segments. Dealing with vectors (as opposed to directed line segments with fixed initial points) is convenient because all vectors transform under rotations in the same manner, as prescribed by Eqs. (3.2).

Physicists usually think of vectors as of "floating arrows" that have length and direction, but not the initial or end points. Alternatively, one may consider directed line segments with the same initial point (the *origin*) and view each of these line segments as a representative of the entire equivalence class. This point of view is often preferred by mathematicians; in physics, it is traditionally reserved for the position vectors.

[50] The three terms in the right-hand side of Eq. (3.2c) have a transparent meaning: $\mathbf{r}_\parallel = \mathbf{n}(\mathbf{n}\cdot\mathbf{r})$ and $\mathbf{r}_\perp = \mathbf{r} - \mathbf{r}_\parallel = \mathbf{r} - \mathbf{n}(\mathbf{n}\cdot\mathbf{r}) = -\mathbf{n}\times(\mathbf{n}\times\mathbf{r})$ are the parallel and the perpendicular to $\mathbf{n}$ projections of $\mathbf{r}$, respectively, whereas vector $\mathbf{n}\times\mathbf{r} = \mathbf{n}\times\mathbf{r}_\perp$, $|\mathbf{n}\times\mathbf{r}| = |\mathbf{r}_\perp|$, is perpendicular to both $\mathbf{n}$ and $\mathbf{r}$.

Any vector $\mathbf{a}$ can be written as a linear combination of the three unit mutually orthogonal vectors $\{\mathbf{x},\mathbf{y},\mathbf{z}\}$ forming a Cartesian basis for the Euclidian space of geometric vectors,

$$\mathbf{a} = \mathbf{x}\,a_{\mathbf{x}} + \mathbf{y}\,a_{\mathbf{y}} + \mathbf{z}\,a_{\mathbf{z}}. \tag{3.3a}$$

The dot and the cross products of vectors $\mathbf{a}$ and $\mathbf{b}$ can then be expressed in terms of their Cartesian components as

$$\mathbf{a}\cdot\mathbf{b} = a_{\mathbf{x}}b_{\mathbf{x}} + a_{\mathbf{y}}b_{\mathbf{y}} + a_{\mathbf{z}}b_{\mathbf{z}}, \tag{3.3b}$$

$$\mathbf{a}\times\mathbf{b} = \mathbf{x}\,(a_{\mathbf{y}}b_{\mathbf{z}} - a_{\mathbf{z}}b_{\mathbf{y}}) + \mathbf{y}\,(a_{\mathbf{z}}b_{\mathbf{x}} - a_{\mathbf{x}}b_{\mathbf{z}}) + \mathbf{z}\,(a_{\mathbf{x}}b_{\mathbf{y}} - a_{\mathbf{y}}b_{\mathbf{z}}). \tag{3.3c}$$

Although the use of such Cartesian basis expansions is perfectly legitimate and often convenient, it tends to obscure the underlying geometric structure. For this reason, in the following we will rely instead on the well-known vector identities [51]

$$\boldsymbol{\alpha}\cdot\boldsymbol{\beta} = \boldsymbol{\beta}\cdot\boldsymbol{\alpha}, \tag{3.4a}$$

$$\boldsymbol{\alpha}\times\boldsymbol{\beta} = -\,\boldsymbol{\beta}\times\boldsymbol{\alpha}, \tag{3.4b}$$

$$\boldsymbol{\alpha}\cdot(\boldsymbol{\beta}\times\boldsymbol{\gamma}) = \boldsymbol{\beta}\cdot(\boldsymbol{\gamma}\times\boldsymbol{\alpha}) = \boldsymbol{\gamma}\cdot(\boldsymbol{\alpha}\times\boldsymbol{\beta}), \tag{3.4c}$$

$$\boldsymbol{\alpha}\times(\boldsymbol{\beta}\times\boldsymbol{\gamma}) = (\boldsymbol{\alpha}\cdot\boldsymbol{\gamma})\,\boldsymbol{\beta} - (\boldsymbol{\alpha}\cdot\boldsymbol{\beta})\boldsymbol{\gamma}, \tag{3.4d}$$

$$\boldsymbol{\alpha}\times(\boldsymbol{\beta}\times\boldsymbol{\gamma}) + \boldsymbol{\beta}\times(\boldsymbol{\gamma}\times\boldsymbol{\alpha}) + \boldsymbol{\gamma}\times(\boldsymbol{\alpha}\times\boldsymbol{\beta}) = \mathbf{0}, \tag{3.4e}$$

$$(\boldsymbol{\alpha}\times\boldsymbol{\beta})\cdot(\boldsymbol{\gamma}\times\boldsymbol{\delta}) = (\boldsymbol{\alpha}\cdot\boldsymbol{\gamma})(\boldsymbol{\beta}\cdot\boldsymbol{\delta}) - (\boldsymbol{\alpha}\cdot\boldsymbol{\delta})(\boldsymbol{\beta}\cdot\boldsymbol{\gamma}). \tag{3.4f}$$

It is easy to show with the help of these identities that products of rotated vectors satisfy

$$\frac{d}{d\theta}\,(\mathbf{a}_{\theta\mathbf{n}}\times\mathbf{b}_{\theta\mathbf{n}}) = \mathbf{n}\times(\mathbf{a}_{\theta\mathbf{n}}\times\mathbf{b}_{\theta\mathbf{n}}), \quad \frac{d}{d\theta}\,(\mathbf{a}_{\theta\mathbf{n}}\cdot\mathbf{b}_{\theta\mathbf{n}}) = 0. \tag{3.5}$$

Accordingly, the cross product $\mathbf{a}\times\mathbf{b}$ transforms under rotations as a vector, whereas the dot product $\mathbf{a}\cdot\mathbf{b}$ is a geometric ***scalar***, i.e., an object that remains invariant under rotations. The latter observation implies that the length of any vector [14] and the angle between any two vectors are not affected by rotations. In particular, if $\{\mathbf{x},\mathbf{y},\mathbf{z}\}$ is a Cartesian basis, then the corresponding rotated vectors $\{\mathbf{x}_{\theta\mathbf{n}},\mathbf{y}_{\theta\mathbf{n}},\mathbf{z}_{\theta\mathbf{n}}\}$ also have unit length and are mutually orthogonal, thus forming another, rotated, Cartesian basis. Since $\mathbf{x}\cdot(\mathbf{y}\times\mathbf{z})$ is a scalar, rotations preserve the "handedness" of the basis.[52] In fact, by Euler's rotation theorem mentioned above any such handedness-preserving change of the Cartesian basis can be accomplished by an appropriate rotation.

Because the product $\mathbf{a}_{(\theta+\theta')\mathbf{n}}\cdot\mathbf{b}_{\theta'\mathbf{n}}$ is independent of θ', we have

$$\mathbf{a}_{\theta\mathbf{n}}\cdot\mathbf{b} = \mathbf{a}\cdot\mathbf{b}_{-\theta\mathbf{n}}. \tag{3.6}$$

(The left- and right-hand sides here correspond to $\theta'=0$ and $\theta'=-\theta$, respectively.) With the help of this relation, it is easy to verify that in spite of their appearance, Eqs. (3.3) are independent of the choice of the Cartesian basis. Indeed, suppose the basis $\{\mathbf{x},\mathbf{y},\mathbf{z}\}$ is replaced

[51] The Jacobi identity (3.4e) and the Binet-Cauchy identity (3.4f) follow from Eqs. (3.4a)-(3.4d).

[52] For the conventional "right-handed" basis $\mathbf{x}\cdot(\mathbf{y}\times\mathbf{z}) = 1$. The inversion $\mathbf{x}\to-\mathbf{x}$, $\mathbf{y}\to-\mathbf{y}$, $\mathbf{z}\to-\mathbf{z}$ transforms this basis to the "left-handed" one for which $\mathbf{x}\cdot(\mathbf{y}\times\mathbf{z}) = -1$; this cannot be accomplished by a rotation.

with the rotated basis $\{\mathbf{x}_{\theta\mathbf{n}}, \mathbf{y}_{\theta\mathbf{n}}, \mathbf{z}_{\theta\mathbf{n}}\}$. Eq. (3.6) shows that components of an arbitrary vector $\mathbf{a}$ in this basis coincide with the components of the rotated vector $\mathbf{a}_{-\theta\mathbf{n}}$ in the original basis, e.g., $a_{\mathbf{x}_{\theta\mathbf{n}}} = \mathbf{x}_{\theta\mathbf{n}} \cdot \mathbf{a} = \mathbf{x} \cdot \mathbf{a}_{-\theta\mathbf{n}}$. Thus, any such change of basis in the expansion of, say, the dot product of $\mathbf{a}$ and $\mathbf{b}$ [see Eq. (3.3b)] amounts to merely replacing $\mathbf{a}$ and $\mathbf{b}$ with $\boldsymbol{\alpha}_{-\theta\mathbf{n}}$ and $\boldsymbol{\beta}_{-\theta\mathbf{n}}$, respectively, which does not affect the value of the product: $\mathbf{a}_{-\theta\mathbf{n}} \cdot \mathbf{b}_{-\theta\mathbf{n}} = \mathbf{a} \cdot \mathbf{b}$ [recall the second equation in Eq. (3.5)].

3.1.2 Rotation transformation and angular momentum

Consider an observable of vector nature $\mathbf{A}$. If $\mathbf{A}$ were a classical quantity, it would transform under rotations as a vector, i.e., the equation describing the dependence of $\mathbf{A}_{\theta\mathbf{n}}$ on the angle-axis vector $\theta\mathbf{n}$ would have the form of Eq. (3.2b),

$$\frac{d}{d\theta}\mathbf{A}_{\theta\mathbf{n}} = \mathbf{n} \times \mathbf{A}_{\theta\mathbf{n}}. \tag{3.7a}$$

Alternatively, rotation can be viewed as a *continuous canonical transformation*[53]

$$\frac{d}{d\theta}\mathbf{A}_{\theta\mathbf{n}} = \big\{\mathbf{A}_{\theta\mathbf{n}}, \mathbf{n} \cdot \mathbf{J}\big\} \tag{3.7b}$$

generated by the angular momentum $\mathbf{J}$. To ensure quantum-classical correspondence, the Hermitian vector operator[44] $\hat{\mathbf{A}}$ representing observable $\mathbf{A}$ [recall Eq. (2.58)] should transform under rotations in the same fashion. That is, the dependence of the rotated operator $\hat{\mathbf{A}}_{\theta\mathbf{n}}$ on $\theta\mathbf{n}$ is governed by the operator version of Eq. (3.7a),

$$\frac{d}{d\theta}\hat{\mathbf{A}}_{\theta\mathbf{n}} = \mathbf{n} \times \hat{\mathbf{A}}_{\theta\mathbf{n}}, \tag{3.8a}$$

so that the expectation value $\langle \mathbf{A}_{\theta\mathbf{n}} \rangle_\psi \equiv \langle \psi | \hat{\mathbf{A}}_{\theta\mathbf{n}} | \psi \rangle$ obeys the classical equation

$$\frac{d}{d\theta}\langle \mathbf{A}_{\theta\mathbf{n}} \rangle_\psi = \mathbf{n} \times \langle \mathbf{A}_{\theta\mathbf{n}} \rangle_\psi. \tag{3.8b}$$

[53] In the Hamiltonian formulation of Classical Mechanics dynamical quantities are functions of the time-dependent positions and momenta ($\mathbf{r}$ and $\mathbf{p}$ respectively for a single particle), referred to as the *canonical variables*. All equations of the Hamiltonian mechanics can be expressed in terms of the so-called **Poisson brackets**. For a single particle, Poisson bracket of two dynamical quantities $f(\mathbf{r},\mathbf{p})$ and $g(\mathbf{r},\mathbf{p})$ read

$$\{f, g\} = \boldsymbol{\nabla}_\mathbf{r} f \cdot \boldsymbol{\nabla}_\mathbf{p} g - \boldsymbol{\nabla}_\mathbf{p} f \cdot \boldsymbol{\nabla}_\mathbf{r} g,$$

where $\boldsymbol{\nabla}_\mathbf{r}$ and $\boldsymbol{\nabla}_\mathbf{p}$ are gradients with respect to $\mathbf{r}$ and $\mathbf{p}$. Properties of Poisson brackets are very similar to those of quantum commutators [see Eqs. (2.13)],

$$\{fg,h\} = f\{g,h\} + \{f,h\}g, \quad \big\{f,\{g,h\}\big\} + \big\{g,\{h,f\}\big\} + \big\{h,\{f,g\}\big\} = 0.$$

These relations feature scalar dynamical quantities. Generalization to vectors [cf. Eq. (3.7b)] is straightforward, e.g., $\{\mathbf{f}, g\} = \mathbf{x}\{f_\mathbf{x}, g\} + \mathbf{y}\{f_\mathbf{y}, g\} + \mathbf{z}\{f_\mathbf{z}, g\}$.

Canonical transformations are transformations $\mathbf{r} \to \mathbf{r}'(\mathbf{r},\mathbf{p})$, $\mathbf{p} \to \mathbf{p}'(\mathbf{r},\mathbf{p})$ that preserve the values of the Poisson brackets of canonical variables, e.g., $\{r'_\mathbf{x}, p'_\mathbf{x}\} = \{r_\mathbf{x}, p_\mathbf{x}\} = 1$. This ensures that transformed equations of motion retain their form. It is easy to verify that the transformation described by Eq. (3.7b) with $\mathbf{J} = \mathbf{r} \times \mathbf{p}$ is indeed canonical and that the right-hand sides of Eqs. (3.7) coincide.

It is not difficult to verify that the cross and dot products of two such vector operators transform under rotations as vectors and scalars, respectively:

$$\frac{d}{d\theta}\,(\hat{\mathbf{A}}_{\theta\mathbf{n}}\times\hat{\mathbf{B}}_{\theta\mathbf{n}}) = \mathbf{n}\times(\hat{\mathbf{A}}_{\theta\mathbf{n}}\times\hat{\mathbf{B}}_{\theta\mathbf{n}}), \quad \frac{d}{d\theta}\,(\hat{\mathbf{A}}_{\theta\mathbf{n}}\cdot\hat{\mathbf{B}}_{\theta\mathbf{n}}) = \hat{\mathbb{0}} \tag{3.9}$$

[cf. Eqs. (3.5)]. These and other similar equations can be derived using the vector identities (3.4), which can be applied to vector operators provided that their order is kept intact.[54]

On the other hand, the rotation transformation $\hat{\mathbf{A}} \to \hat{\mathbf{A}}_{\theta\mathbf{n}}$ is a continuous symmetry transformation [see Sec. 2.5.3] reflecting the isotropy of space, with the rotation angle θ playing the part of the continuously varying parameter. Therefore,

$$\hat{\mathbf{A}}_{\theta\mathbf{n}} = \hat{R}^\dagger(\theta\mathbf{n})\,\hat{\mathbf{A}}\,\hat{R}(\theta\mathbf{n}) \tag{3.10}$$

[cf. Eq. (2.127a)], where the ***rotation operator*** $\hat{R}(\theta\mathbf{n})$ is unitary. Eq. (3.8a) suggests that the phases of these operators can be chosen so that

$$\hat{R}(\theta\mathbf{n})\,\hat{R}(\theta'\mathbf{n}) = \hat{R}\big[(\theta+\theta')\mathbf{n}\big]. \tag{3.11}$$

Indeed, Eq. (3.8a) is nothing but the operator version of Eq. (3.7a), which in turn is the consequence of the observation [recall (3.1)] that two consecutive rotations about the same axis yield the same result as a single rotation by a combined angle.

As shown in Sec. 2.5.3, the group property (3.11) and the unitarity imply that the rotation operator $\hat{R}(\theta\mathbf{n})$ is an exponent of a scalar operator linear in θ [see Eq. (2.125)]. At the same time, this operator is a function of $\theta\mathbf{n}$, which is a vector. To form a scalar, this vector should be multiplied (in the dot product sense) by some vector, hence

$$\hat{R}(\theta\mathbf{n}) = e^{-i\theta\mathbf{n}\cdot\hat{\mathbf{K}}}, \tag{3.12}$$

where the generator of rotations $\hat{\mathbf{K}}$ is an Hermitian vector operator[44] independent of $\theta\mathbf{n}$. Differentiation of Eq. (3.10) with $\hat{R}(\theta\mathbf{n})$ given by (3.12) yields the quantum counterpart of Eq. (3.7b),

$$\frac{d}{d\theta}\,\hat{\mathbf{A}}_{\theta\mathbf{n}} = \frac{1}{i}\,[\hat{\mathbf{A}}_{\theta\mathbf{n}}, \mathbf{n}\cdot\hat{\mathbf{K}}]. \tag{3.13}$$

[54] For example, the first equation in (3.9) is derived as follows. Differentiating the product $\hat{\mathbf{A}}_{\theta\mathbf{n}}\times\hat{\mathbf{B}}_{\theta\mathbf{n}}$ with the help of Eq. (3.8a) and a similar equation for $\hat{\mathbf{B}}_{\theta\mathbf{n}}$, we obtain

$$\frac{d}{d\theta}\,(\hat{\mathbf{A}}_{\theta\mathbf{n}}\times\hat{\mathbf{B}}_{\theta\mathbf{n}}) = (\mathbf{n}\times\hat{\mathbf{A}}_{\theta\mathbf{n}})\times\hat{\mathbf{B}}_{\theta\mathbf{n}} + \hat{\mathbf{A}}_{\theta\mathbf{n}}\times(\mathbf{n}\times\hat{\mathbf{B}}_{\theta\mathbf{n}}).$$

To simplify the right-hand side of this equation, we note that the Jacobi identity (3.4e) for vectors $\mathbf{n}$, $\mathbf{A}$, and $\mathbf{B}$ (not yet operators!), $\mathbf{n}\times(\mathbf{A}\times\mathbf{B}) + \mathbf{A}\times(\mathbf{B}\times\mathbf{n}) + \mathbf{B}\times(\mathbf{n}\times\mathbf{A}) = \mathbf{0}$, can be also written as

$$(\mathbf{n}\times\mathbf{A})\times\mathbf{B} + \mathbf{A}\times(\mathbf{n}\times\mathbf{B}) = \mathbf{n}\times(\mathbf{A}\times\mathbf{B}).$$

Because the order of vectors $\mathbf{A}$ and $\mathbf{B}$ is maintained throughout this relation, it is safe to replace these vectors with vector operators $\hat{\mathbf{A}}_{\theta\mathbf{n}}$ and $\hat{\mathbf{B}}_{\theta\mathbf{n}}$, respectively. This yields the operator relation

$$(\mathbf{n}\times\hat{\mathbf{A}}_{\theta\mathbf{n}})\times\hat{\mathbf{B}}_{\theta\mathbf{n}} + \hat{\mathbf{A}}_{\theta\mathbf{n}}\times(\mathbf{n}\times\hat{\mathbf{B}}_{\theta\mathbf{n}}) = \mathbf{n}\times(\hat{\mathbf{A}}_{\theta\mathbf{n}}\times\hat{\mathbf{B}}_{\theta\mathbf{n}}),$$

which brings the above equation to the form of the first equation in Eq. (3.9).

The apparent similarity between Eqs. (3.7b) and (3.13) suggests identifying the generator of rotations $\hat{\mathbf{K}}$ with the operator $\hat{\mathbf{J}}$ corresponding to the angular momentum. Note that although the angular momentum is not dimensionless, the right-hand side of Eq. (3.7b) has units of $\mathbf{A}$. Indeed, the Poisson bracket contains products of derivatives with respect to $\mathbf{r}$ and $\mathbf{p}$,[53] which compensates for the length$\times$momentum dimension of $\mathbf{J}$. The role of these units-saving derivatives in the right-hand side of Eq. (3.13) can only be played by the proportionality coefficient in $\hat{\mathbf{K}} \propto \hat{\mathbf{J}}$. This necessitates the introduction of a new fundamental constant absent in Classical Physics. This constant is, of course, nothing but the ***Planck constant*** $\hbar$ that has units of angular momentum. Accordingly, we can write

$$\hat{\mathbf{K}} = \hat{\mathbf{J}}/\hbar. \tag{3.14}$$

Although we have already encountered $\hbar$ before, it is Eq. (3.14) where it makes its first appearance in our formal development of quantum theory. The numerical coefficient in Eq. (3.14) is the matter of convention. We will see in Sec. 3.2.4 below that with the choice (3.14) (as opposed to, say, $\hat{\mathbf{K}} = \hat{\mathbf{J}}/2\hbar$) the eigenvalues of the Cartesian components of spin 1/2 operators are given by $\pm\hbar/2$, just as we have assumed in Sec. 1.1.3, 1.3.3, and 2.3.2.

It should be emphasized that Eqs. (3.8a) and (3.13) have a very different origin. Whereas the former simply tells us what the rotation transformation does, the latter exploits the observation that rotation is a continuous symmetry transformation and as such is unitary. Because unitary transformations do not affect traces of operators on finite-dimensional Hilbert spaces [recall Eq. (2.115b)], Eqs. (3.8a) and (3.13) can be satisfied simultaneously only for vector operators with vanishing traces, i.e., such that $\operatorname{tr}(\mathbf{a}\cdot\hat{\mathbf{A}}) = 0$ for all $\mathbf{a}$, or, equivalently,

$$\operatorname{tr}\hat{\mathbf{A}} \equiv \mathbf{x}\operatorname{tr}\hat{A}_{\mathbf{x}} + \mathbf{y}\operatorname{tr}\hat{A}_{\mathbf{y}} + \mathbf{z}\operatorname{tr}\hat{A}_{\mathbf{z}} = \mathbf{0}. \tag{3.15}$$

Hereinafter, we consider only such traceless operators.[55,56] This applies, in particular, to the generator of rotations $\hat{\mathbf{K}}$ and thus the angular momentum $\hat{\mathbf{J}}$. Indeed, the constraint $\operatorname{tr}\hat{\mathbf{K}} = 0$ can always be enforced by replacing $\hat{\mathbf{K}} \to \hat{\mathbf{K}} + \mathbf{k}\hat{\mathbb{1}}$, where $\mathbf{k}$ is an appropriate ordinary geometric vector.[57]

In ***active rotations*** operators representing observables of vector nature transform according to Eqs. (3.8a) and (3.13), whereas the state vectors remain as they are [see Sec. 2.5.3 and Eq. (2.127)]. One can also consider ***passive rotations*** that do not affect probabilities and expectation values [see Sec. 2.5.2 and Eq. (2.128)]. Let $\mathbf{a}$ be an arbitrary reference vector, such as one of the vectors forming the Cartesian basis, and let $\hat{\mathbf{A}}$ be a vector operator. Provided that $\mathbf{a}$ and $\hat{\mathbf{A}}$ transform under rotations as prescribed by Eqs. (3.2b) and (3.8a), respectively, it is easy to show that

$$\mathbf{a}_{\theta\mathbf{n}}\cdot\hat{\mathbf{A}} = \mathbf{a}\cdot\hat{\mathbf{A}}_{-\theta\mathbf{n}} \tag{3.16a}$$

[55] Importantly, this restriction does not apply to operators on infinite-dimensional spaces, such as the position and momentum operators.

[56] Formally, any vector operator $\hat{\mathbf{A}}'$ with a non-zero trace can be uniquely decomposed as $\hat{\mathbf{A}}' = \hat{\mathbf{A}} + \mathbf{a}\hat{\mathbb{1}}$, where $\hat{\mathbf{A}}$ is traceless [thus subject to both Eqs. (3.8a) and (3.13)] and $\mathbf{a} = \operatorname{tr}\hat{\mathbf{A}}'/\dim\mathcal{H}$. Although nominally an operator, $\mathbf{a}\hat{\mathbb{1}}$ is trivial in the sense that it commutes with all other operators and can be treated as an ordinary geometric vector in products such as $(\mathbf{a}\hat{\mathbb{1}})\cdot\hat{\mathbf{B}} = \mathbf{a}\cdot\hat{\mathbf{B}}$.

[57] The replacement $\hat{\mathbf{K}} \to \hat{\mathbf{K}} + \mathbf{k}\hat{\mathbb{1}}$ has no effect on Eqs. (3.10) and (3.11) [see the the remark after Eq. (2.123)].

[cf. Eq. (3.6)]. Substituting here $\hat{\mathbf{A}}_{-\theta\mathbf{n}} = \hat{R}^\dagger(-\theta\mathbf{n})\,\hat{\mathbf{A}}\,\hat{R}(-\theta\mathbf{n})$ [see Eq. (3.10)] and taking into account that $\hat{R}^\dagger(\theta\mathbf{n}) = \hat{R}(-\theta\mathbf{n})$, we obtain a very useful relation

$$\mathbf{a}_{\theta\mathbf{n}}\cdot\hat{\mathbf{A}} = \hat{R}(\theta\mathbf{n})(\mathbf{a}\cdot\hat{\mathbf{A}})\hat{R}^\dagger(\theta\mathbf{n}). \tag{3.16b}$$

Thus, if $|\phi\rangle$ is an eigenvector of operator $\mathbf{a}\cdot\hat{\mathbf{A}}$ then $\hat{R}(\theta\mathbf{n})|\phi\rangle$ is an eigenvector of $\mathbf{a}_{\theta\mathbf{n}}\cdot\hat{\mathbf{A}}$ with the same eigenvalue [see Eqs. (2.114)]. Eq. (3.16b) shows, in particular, that Cartesian components of a vector operator can be obtained from one another by appropriate rotations, e.g., $\hat{A}_\mathbf{y} = \hat{R}(\pi\mathbf{z}/2)\hat{A}_\mathbf{x}\hat{R}^\dagger(\pi\mathbf{z}/2)$.

3.1.3 Case study: rotations and spin 1/2

In this section we deploy the machinery developed in Sec. 3.1.2 to re-derive the results of Sec. 1.3.3. We again take for granted that the operator $\hat{S}_\mathbf{n} = \mathbf{n}\cdot\hat{\mathbf{S}}$ has two non-degenerate eigenvalues $\pm\hbar/2$ [this ensures that $\operatorname{tr}\hat{S}_\mathbf{n} = 0$]. As in Sec. 1.3.3 and 2.3.2, we use $|\mathbf{n}\rangle$ to denote the eigenvector of $\hat{S}_\mathbf{n}$ with eigenvalue $+\hbar/2$. This vector is a unique (up to a phase factor) normalized solution of the equation

$$\hat{S}_\mathbf{n}|\mathbf{n}\rangle = (\hbar/2)|\mathbf{n}\rangle. \tag{3.17}$$

Consider now a passive rotation transformation [see Eqs. (2.128) and (3.16b)] that replaces $\mathbf{n}$ in $\hat{S}_\mathbf{n}$ with a rotated unit vector $\mathbf{n}_{\theta\mathbf{n}_0}$,

$$\hat{S}_{\mathbf{n}_{\theta\mathbf{n}_0}} = \hat{R}(\theta\mathbf{n}_0)\hat{S}_\mathbf{n}\hat{R}^\dagger(\theta\mathbf{n}_0). \tag{3.18a}$$

The transformed vector $\hat{R}(\theta\mathbf{n}_0)|\mathbf{n}\rangle$ is an eigenvector of $\hat{S}_{\mathbf{n}_{\theta\mathbf{n}_0}}$ with eigenvalue $+\hbar/2$, i.e., in notations of Eq. (3.17),

$$|\mathbf{n}_{\theta\mathbf{n}_0}\rangle = \hat{R}(\theta\mathbf{n}_0)|\mathbf{n}\rangle \quad \text{(up to phase factors)}, \tag{3.18b}$$

where $\mathbf{n}_{\theta\mathbf{n}_0}$ is the rotated Bloch vector.

For any pair of Bloch vectors $\mathbf{n}$ and $\mathbf{n}'$ one can find a rotation transforming $\mathbf{n}$ to $\mathbf{n}_{\theta\mathbf{n}_0} = \mathbf{n}'$. Because the product $\mathbf{n}_{\theta\mathbf{n}_0}\cdot\mathbf{n}_0$ is independent of θ, the axis vector $\mathbf{n}_0$ satisfies $(\mathbf{n}-\mathbf{n}')\cdot\mathbf{n}_0 = 0$. This equation does not define $\mathbf{n}_0$ uniquely, which makes it possible to impose additional constraints on $\theta\mathbf{n}_0$. For our purposes, it is convenient to restrict the rotation angle θ to the interval $0 \le \theta \le \pi$ and, simultaneously, to require the axis vector $\mathbf{n}_0$ to be perpendicular to both $\mathbf{n}$ and $\mathbf{n}'$. For $\mathbf{n}' \neq -\mathbf{n}$ these constraints yield unique θ and $\mathbf{n}_0$,

$$\cos\theta = \mathbf{n}\cdot\mathbf{n}', \quad \mathbf{n}_0 = \frac{\mathbf{n}\times\mathbf{n}'}{|\mathbf{n}\times\mathbf{n}'|}. \tag{3.19}$$

The case $\mathbf{n}' = -\mathbf{n}$ is special: whereas the angle $\theta = \pi$ is still unique, the axis vector $\mathbf{n}_0$ can point in any direction perpendicular to $\mathbf{n}$. Moreover, since $\hat{S}_{\pi\mathbf{n}_0} = \hat{S}_{-\mathbf{n}} = -\hat{S}_\mathbf{n}$, the rotated vector

$$|-\mathbf{n}\rangle = \hat{R}(\pi\mathbf{n}_0)|\mathbf{n}\rangle \quad \text{(up to phase factors)} \tag{3.20a}$$

is an eigenvector of $\hat{S}_\mathbf{n}$ with eigenvalue $-\hbar/2$ and thus [recall Eq. (2.41)] is orthogonal to $|\mathbf{n}\rangle$,

$$\langle\mathbf{n}|-\mathbf{n}\rangle = \langle\mathbf{n}|\hat{R}(\pi\mathbf{n}_0)|\mathbf{n}\rangle = 0 \quad \text{for} \quad \mathbf{n}_0 \perp \mathbf{n}. \tag{3.20b}$$

Accordingly, vectors $|\pm\mathbf{n}\rangle$ (with arbitrary Bloch vector $\mathbf{n}$) form an orthonormal basis for the two-dimensional Hilbert space of spin 1/2.

Substituting the spectral decomposition of the rotation operator

$$\hat{R}(\theta\mathbf{n}_0) = e^{-i\theta\hat{S}_{\mathbf{n}_0}/\hbar} = |\mathbf{n}_0\rangle e^{-i\theta/2}\langle\mathbf{n}_0| + |-\mathbf{n}_0\rangle e^{i\theta/2}\langle-\mathbf{n}_0| \tag{3.21}$$

into Eq. (3.20b), we obtain $|\langle\mathbf{n}|\mathbf{n}_0\rangle| = |\langle\mathbf{n}|-\mathbf{n}_0\rangle|$. On the other hand, $\langle\mathbf{n}|\hat{\mathbb{1}}|\mathbf{n}\rangle = 1$ and the completeness relation $\hat{\mathbb{1}} = |\mathbf{n}_0\rangle\langle\mathbf{n}_0| + |-\mathbf{n}_0\rangle\langle-\mathbf{n}_0|$ yield $|\langle\mathbf{n}|\mathbf{n}_0\rangle|^2 + |\langle\mathbf{n}|-\mathbf{n}_0\rangle|^2 = 1$. These relations show that

$$\big|\langle\mathbf{n}|\mathbf{n}_0\rangle\big|^2 = 1/2 \quad \text{for} \quad \mathbf{n}\perp\mathbf{n}_0\,. \tag{3.22}$$

Using now Eqs. (3.18b), (3.19), (3.21), and (3.22), we get [58]

$$\langle\mathbf{n}|\mathbf{n}'\rangle = \langle\mathbf{n}|\hat{R}(\theta\mathbf{n}_0)|\mathbf{n}\rangle\big|_{\mathbf{n}_0\perp\mathbf{n}} = e^{-i\theta/2}\underbrace{\big|\langle\mathbf{n}|\mathbf{n}_0\rangle\big|^2}_{1/2} + e^{i\theta/2}\underbrace{\big|\langle\mathbf{n}|-\mathbf{n}_0\rangle\big|^2}_{1/2} = \cos(\theta/2)$$

(up to a phase factor), which yields

$$\big|\langle\mathbf{n}|\mathbf{n}'\rangle\big|^2 = \cos^2(\theta/2) = \frac{1}{2}(1+\cos\theta) = \frac{1}{2}\big(1+\mathbf{n}\cdot\mathbf{n}'\big), \tag{3.23}$$

in agreement with Eqs. (1.22) and (1.55).

3.2 Angular momentum

3.2.1 Algebra of angular momentum operators

We return now to the general properties of the quantum rotation transformation introduced in Sec. 3.1.2. Equating the right-hand sides of Eqs. (3.8a) and (3.13) with $\hat{\mathbf{K}}$ given by Eq. (3.14) and setting $\theta = 0$, we obtain

$$\frac{1}{i\hbar}[\hat{\mathbf{A}}, \mathbf{n}\cdot\hat{\mathbf{J}}] = \mathbf{n}\times\hat{\mathbf{A}}.$$

Because this expression is linear in $\mathbf{n}$, it remains valid when $\mathbf{n}$ is replaced with any other vector, say, $\mathbf{b}$. Multiplying both sides of the resulting equation (in the dot product sense) by $i\hbar\mathbf{a}$, we arrive at the commutation relation [59]

$$[\mathbf{a}\cdot\hat{\mathbf{A}}, \mathbf{b}\cdot\hat{\mathbf{J}}] = i\hbar(\mathbf{a}\times\mathbf{b})\cdot\hat{\mathbf{A}}, \tag{3.24}$$

valid for all reference vectors $\mathbf{a}$ and $\mathbf{b}$. Note that the left-hand side of Eq. (3.24)can be also written as $[\mathbf{a}\cdot\hat{\mathbf{J}}, \mathbf{b}\cdot\hat{\mathbf{A}}]$. Elimination of one of the two reference vectors in Eq. (3.24) yields the relations

$$[\hat{\mathbf{A}}, \mathbf{a}\cdot\hat{\mathbf{J}}] = i\hbar(\mathbf{a}\times\hat{\mathbf{A}}), \quad [\hat{\mathbf{J}}, \mathbf{a}\cdot\hat{\mathbf{A}}] = i\hbar(\mathbf{a}\times\hat{\mathbf{A}}), \tag{3.25}$$

[58] Alternatively, one may notice that Eq. (3.22) implies that $\langle\mathbf{n}|\hat{S}_{\mathbf{n}_0}|\mathbf{n}\rangle\big|_{\mathbf{n}_0\perp\mathbf{n}} = 0$. With this relation taken into account, substitution of $\hat{R}(\theta\mathbf{n}_0) = \cos(\theta/2)\hat{\mathbb{1}} - i\sin(\theta/2)(2\hat{S}_{\mathbf{n}_0}/\hbar)$ into $\langle\mathbf{n}|\mathbf{n}'\rangle = \langle\mathbf{n}|\hat{R}(\theta\mathbf{n}_0)|\mathbf{n}\rangle\big|_{\mathbf{n}_0\perp\mathbf{n}}$ immediately yields $\langle\mathbf{n}|\mathbf{n}'\rangle = \cos(\theta/2)$ (up to a phase factor).

[59] Note that in view of Eq. (2.32b), Eq. (3.24) implies Eq. (3.15) provided that the dimension of the Hilbert space is finite.[37]

valid for all $\mathbf{a}$. Either one of these commutation relations is equivalent to Eq. (3.24). Actually, it is possible to get rid of both $\mathbf{a}$ and $\mathbf{b}$ in Eq. (3.24) with the result [60]

$$\hat{\mathbf{A}}\cdot\hat{\mathbf{J}}=\hat{\mathbf{J}}\cdot\hat{\mathbf{A}},\quad \hat{\mathbf{A}}\times\hat{\mathbf{J}}+\hat{\mathbf{J}}\times\hat{\mathbf{A}}=2i\hbar\hat{\mathbf{A}}. \tag{3.26}$$

Note that one needs both equations here to recover Eq. (3.24).

Eq. (3.24) shows that (not quite surprisingly) components of vector operators do not commute with components of $\hat{\mathbf{J}}$. On the contrary (and also not surprisingly), scalar operators remain invariant under rotations and thus commute with $\hat{\mathbf{J}}$. For example, substitution of

$$\hat{\mathbf{A}}_{\theta\mathbf{n}}\cdot\hat{\mathbf{B}}_{\theta\mathbf{n}}=\big[\hat{R}^\dagger(\theta\mathbf{n})\hat{\mathbf{A}}\hat{R}(\theta\mathbf{n})\big]\cdot\big[\hat{R}^\dagger(\theta\mathbf{n})\hat{\mathbf{B}}\hat{R}(\theta\mathbf{n})\big]=\hat{R}^\dagger(\theta\mathbf{n})(\hat{\mathbf{A}}\cdot\hat{\mathbf{B}})\hat{R}(\theta\mathbf{n})$$

into the second equation in (3.9) yields the relation $[\mathbf{n}\cdot\hat{\mathbf{J}},\hat{\mathbf{A}}_{\theta\mathbf{n}}\cdot\hat{\mathbf{B}}_{\theta\mathbf{n}}]=\hat{\mathbb{0}}$, valid for all $\theta\mathbf{n}$. Setting here $\theta=0$ and dropping $\mathbf{n}$, we obtain

$$[\hat{\mathbf{A}}\cdot\hat{\mathbf{B}},\hat{\mathbf{J}}]=\hat{\mathbb{0}}. \tag{3.27}$$

(Direct evaluation of this commutator with the help of Eq. (3.24) yields the same result.)

Eqs. (3.24) and (3.27) apply to all vector operators satisfying Eqs. (3.8a) and (3.13), including $\hat{\mathbf{A}}=\hat{\mathbf{J}}$, in which case the commutation relation (3.24) turns to

$$[\mathbf{a}\cdot\hat{\mathbf{J}},\mathbf{b}\cdot\hat{\mathbf{J}}]=i\hbar(\mathbf{a}\times\mathbf{b})\cdot\hat{\mathbf{J}}. \tag{3.28}$$

This relation can be also written as $[\hat{\mathbf{J}},\mathbf{a}\cdot\hat{\mathbf{J}}]=i\hbar(\mathbf{a}\times\hat{\mathbf{J}})$ or as $\hat{\mathbf{J}}\times\hat{\mathbf{J}}=i\hbar\hat{\mathbf{J}}$ [see Eqs. (3.25) and (3.26)]. Similarly, substituting $\hat{\mathbf{A}}=\hat{\mathbf{B}}=\hat{\mathbf{J}}$ into Eq. (3.27), we obtain

$$[\hat{\mathbf{J}}^2,\hat{\mathbf{J}}]=\hat{\mathbb{0}},\quad \hat{\mathbf{J}}^2\equiv\hat{\mathbf{J}}\cdot\hat{\mathbf{J}}. \tag{3.29}$$

Eqs. (3.28) and (3.29) contain everything we need to construct the angular momentum eigenvectors.

3.2.2 Eigenvalue problem for angular momentum

Eqs. (3.28) and (3.29) show that Cartesian components of the angular momentum operator do not commute with each other,

$$[\hat{J}_{\mathbf{x}},\hat{J}_{\mathbf{y}}]=i\hbar\hat{J}_{\mathbf{z}},\quad [\hat{J}_{\mathbf{y}},\hat{J}_{\mathbf{z}}]=i\hbar\hat{J}_{\mathbf{x}},\quad [\hat{J}_{\mathbf{z}},\hat{J}_{\mathbf{x}}]=i\hbar\hat{J}_{\mathbf{y}}, \tag{3.30}$$

yet they all commute with $\hat{\mathbf{J}}^2=\hat{J}_{\mathbf{x}}^2+\hat{J}_{\mathbf{y}}^2+\hat{J}_{\mathbf{z}}^2$. Therefore [recall Eq. (2.85)], we can consider simultaneous eigenvectors of $\hat{\mathbf{J}}^2$ and, say, $\hat{J}_{\mathbf{z}}$ [61] forming a complete set of commuting operators [see Sec. 2.4.2]. We tentatively denote these eigenvectors by $|\lambda,m\rangle$, where λ and m are real dimensionless numbers parametrizing the eigenvalues of $\hat{\mathbf{J}}^2$ and $\hat{J}_{\mathbf{z}}$, respectively:

$$\hat{\mathbf{J}}^2|\lambda,m\rangle=\hbar^2\lambda|\lambda,m\rangle,\quad \hat{J}_{\mathbf{z}}|\lambda,m\rangle=\hbar m|\lambda,m\rangle,\quad \langle\lambda,m|\lambda',m'\rangle=\delta_{\lambda,\lambda'}\,\delta_{m,m'}\,. \tag{3.31}$$

[60] The first equation in (3.26) follows from $[\mathbf{a}\cdot\hat{\mathbf{A}},\mathbf{a}\cdot\hat{\mathbf{J}}]=\hat{\mathbb{0}}$. The second equation is obtained by applying the identity (3.4f) to $(\mathbf{a}\times\mathbf{b})\cdot(\hat{\mathbf{A}}\times\hat{\mathbf{J}}+\hat{\mathbf{J}}\times\hat{\mathbf{A}})$ and using Eq. (3.24). Alternatively, equations (3.26) can be derived by taking a divergence and a curl with respect to $\mathbf{a}$ of both sides of either one of Eqs. (3.25).

[61] By an appropriate passive rotation transformation [see Eq. (3.16b)] $\mathbf{z}$ in $\hat{J}_{\mathbf{z}}$ can be replaced with any other unit dimensionless vector.

The same reasoning as in Sec. 3.1.3 [see Eqs. (3.18) and (3.20a)] shows that the rotated vector $\hat{R}(\pi\mathbf{n}_0)|\lambda,m\rangle$ with $\mathbf{n}_0 \perp \mathbf{z}$ is an eigenvector of $\hat{J}_\mathbf{z}$ with eigenvalue $-\hbar m$, hence this vector coincides (up to a phase factor) with $|\lambda,-m\rangle$ [see the second equation in (3.31)]. Therefore, eigenvalues of $\hat{J}_\mathbf{z}$ come in pairs: if $|\lambda,m\rangle$ is an eigenvector then there ought to exist also an eigenvector with the opposite value of m, i.e., $|\lambda,-m\rangle$;[62] this ensures that $\mathrm{tr}\,\hat{J}_\mathbf{z} = 0$. On the other hand, Eqs. (3.31) and the Hermiticity of operators $\hat{J}_\mathbf{x}$ and $\hat{J}_\mathbf{y}$ imply[63] that

$$\hbar^2(\lambda - m^2) = \langle\lambda,m|\big(\hat{\mathbf{J}}^2 - \hat{J}_\mathbf{z}^2\big)|\lambda,m\rangle = \langle\lambda,m|\hat{J}_\mathbf{x}^2|\lambda,m\rangle + \langle\lambda,m|\hat{J}_\mathbf{y}^2|\lambda,m\rangle \geq 0\,,$$

which shows that values of m allowed for a given λ satisfy $\lambda \geq m^2 \geq 0$, i.e.,

$$j \equiv \max\{m\} \leq \sqrt{\lambda}\,. \tag{3.32}$$

To make further progress, it is convenient to introduce the non-Hermitian linear combinations of operators $\hat{J}_\mathbf{x}$ and $\hat{J}_\mathbf{y}$ (the so-called ***circular components*** of vector operator $\hat{\mathbf{J}}$)

$$\hat{J}_\pm = \hat{J}_\mathbf{x} \pm i\hat{J}_\mathbf{y}\,, \quad \hat{J}_\pm^\dagger = \hat{J}_\mp\,. \tag{3.33a}$$

It follows from Eqs. (3.29) and (3.30) that these operators obey the commutation relations

$$[\hat{\mathbf{J}}^2, \hat{J}_\pm] = 0\,, \quad [\hat{J}_\mathbf{z}, \hat{J}_\pm] = \pm\hbar\hat{J}_\pm\,, \quad [\hat{J}_+, \hat{J}_-] = 2\hbar\hat{J}_\mathbf{z}\,, \tag{3.33b}$$

so that

$$\begin{aligned}\hat{\mathbf{J}}^2 = \hat{J}_\mathbf{x}^2 + \hat{J}_\mathbf{y}^2 + \hat{J}_\mathbf{z}^2 &= \frac{1}{2}\big(\hat{J}_+\hat{J}_- + \hat{J}_-\hat{J}_+\big) + \hat{J}_\mathbf{z}^2 \\ &= \hat{J}_\mp\hat{J}_\pm \pm \frac{1}{2}[\hat{J}_+, \hat{J}_-] + \hat{J}_\mathbf{z}^2 = \hat{J}_\mathbf{z}^2 \pm \hbar\hat{J}_\mathbf{z} + \hat{J}_\mp\hat{J}_\pm\,.\end{aligned} \tag{3.33c}$$

The second equation in (3.31) and the third equation in (3.33b) give

$$\hat{J}_\mathbf{z}\hat{J}_\pm|\lambda,m\rangle = \big(\hat{J}_\pm\hat{J}_\mathbf{z} + [\hat{J}_\mathbf{z}, \hat{J}_\pm]\big)|\lambda,m\rangle = \big(\hat{J}_\pm\hat{J}_\mathbf{z} \pm \hbar\hat{J}_\pm\big)|\lambda,m\rangle = \hbar(m\pm1)\hat{J}_\pm|\lambda,m\rangle.$$

Comparison with $\hat{J}_\mathbf{z}|\lambda,m\pm1\rangle = \hbar(m\pm1)|\lambda,m\pm1\rangle$ [see Eqs. (3.31)] shows that $\hat{J}_\pm$ acts as the ***raising/lowering*** (or ***ladder***) operator for the eigenvalue of $\hat{J}_\mathbf{z}$,

$$\hat{J}_\pm|\lambda,m\rangle \propto |\lambda,m\pm1\rangle \tag{3.34}$$

with the proportionality coefficient that depends on λ and m. To enforce the bounds (3.32) on the allowed values of m, this coefficient must vanish identically for $m = \pm j$, i.e.,

$$\hat{J}_\pm|\lambda,\pm j\rangle = |\mathrm{null}\rangle\,. \tag{3.35}$$

Accordingly, the last term in $\hat{\mathbf{J}}^2 = \hat{J}_\mathbf{z}^2 + \hbar\hat{J}_\mathbf{z} + \hat{J}_-\hat{J}_+$ [see Eq. (3.33c)] makes no contribution when this operator acts on $|\lambda,j\rangle$:

$$\hat{\mathbf{J}}^2|\lambda,j\rangle = (\hat{J}_\mathbf{z}^2 + \hbar\hat{J}_\mathbf{z})|\lambda,j\rangle = \hbar^2 j(j+1)|\lambda,j\rangle.$$

[62] Accordingly, eigenvalues of $\hat{\mathbf{J}}^2$ are necessarily degenerate. This observation is a manifestation of the general property: if $[\hat{A},\hat{B}] = [\hat{A},\hat{C}] = \hat{\mathbb{0}}$ but $[\hat{B},\hat{C}] \neq \hat{\mathbb{0}}$ then at least one eigenvalue of $\hat{A}$ must be degenerate.

[63] For any $\hat{A} = \hat{A}^\dagger$ and any $|\psi\rangle$ we have $\langle\psi|\hat{A}^2|\psi\rangle = \langle\hat{A}\psi|\hat{A}\psi\rangle \geq 0$, see Eqs. (1.33d), (1.35), and (2.36b).

Comparing this relation with $\hat{\mathbf{J}}^2|\lambda,m\rangle = \hbar^2\lambda|\lambda,m\rangle$ [see Eqs. (3.31)], we obtain

$$\lambda = j(j+1)\,, \tag{3.36}$$

in agreement with Eq. (3.32). Eq. (3.36) shows that the correspondence between the non-negative real numbers λ and j is one-to-one. It is therefore convenient (and customary) to label the eigenvectors by $j = \max\{m\}$ instead of λ, write

$$\hat{\mathbf{J}}^2|j,m\rangle = \hbar^2 j(j+1)|j,m\rangle, \quad \hat{J}_{\mathbf{z}}|j,m\rangle = \hbar m|j,m\rangle, \quad \langle j,m|j',m'\rangle = \delta_{j,j'}\delta_{m,m'} \tag{3.37}$$

instead of Eqs. (3.31), and refer to vectors $|j,m\rangle$ defined by these relations as the ***eigenvectors of angular momentum j***. Hereinafter we follow this convention.

In order to find the dimensionless proportionality coefficient $c_\pm$ in

$$\hat{J}_\pm|j,m\rangle = \hbar c_\pm(j,m)|j,m\pm1\rangle \tag{3.38a}$$

[recall Eq. (3.34)], we take a square of the norm of both sides of this relation. Taking into account the last equation in (3.37), we get

$$\hbar^2|c_\pm(j,m)|^2 = \langle j,m|\hat{J}_\pm^\dagger\hat{J}_\pm|j,m\rangle.$$

Substituting here $\hat{J}_\pm^\dagger\hat{J}_\pm = \hat{J}_\mp\hat{J}_\pm = \hat{\mathbf{J}}^2 - \hat{J}_{\mathbf{z}}^2 \mp \hbar\hat{J}_{\mathbf{z}}$ [see Eq. (3.33c)] and using Eqs. (3.37), we obtain

$$\big|c_\pm(j,m)\big|^2 = j(j+1) - m(m\pm1) = (j\mp m)(j\pm m+1)\,.$$

It is convenient (and conventional) to take the square root of $|c_\pm|^2$ to be real and positive for all $|m|<j$ [this amounts to making a choice of the relative phase of eigenvectors $|j,m\rangle$ and $|j,m\pm1\rangle$ in Eq. (3.38a)], i.e.,

$$c_\pm(j,m) = \sqrt{(j\mp m)(j\pm m+1)}\,. \tag{3.38b}$$

Notice that $c_\pm(j,\pm j) = 0$, in agreement with Eq. (3.35).

Acting on vector $|j,m\rangle$ repeatedly with $\hat{J}_\pm$, one must reach vectors $|j,\pm j\rangle$. (Otherwise, it would be possible to raise/lower m beyond its bounds $\pm j$.) Since $\hat{J}_\pm$ raises/lowers m by 1, this implies that $\max\{m\} - \min\{m\} = 2j$ is an integer number and that m takes on $2j+1$ different values for a given j,

$$m = -j, -j+1, \ldots, j-1, j. \tag{3.39}$$

In other words, eigenvalues of $\hat{\mathbf{J}}^2$ are $(2j+1)$-fold-degenerate.[62]

Because $2j$ is an integer, j can be either an ***integer*** or a ***half-integer***. It turns out that for the angular momentum associated with the particle motion (such as the orbital motion of electrons in atoms) j is always an integer. On the contrary, half-integer j always represents an intrinsic angular momentum of quantum particles that has no classical interpretation, the ***spin***. Each elementary particle is characterized by a certain spin (a particle with an intrinsic angular momentum $j = S$ is referred to as a spin S particle).

In atoms spin and orbital angular momenta of the constituent particles add up, and the total angular momentum of an atom can be both integer or a half-integer. Because atoms often behave as particles, it is very common to think of an atom as of a composite spin S particle, although S now stands for the total angular momentum. For example, silver atoms in the original Stern-Gerlach experiments behave as such composite spin 1/2 particles.

3.2.3 Rotations and angular momentum eigenvectors

In this section we focus on the $(2j+1)$-dimensional eigenspace $\mathcal{H}_j$ of the operator $\hat{\mathbf{J}}^2$ formed by vectors obeying the equation $\hat{\mathbf{J}}^2|\psi\rangle = \hbar^2 j(j+1)|\psi\rangle$, where j is either an integer or a half-integer number. The complete set of commuting operators on this subspace consists of a single operator $\hat{J}_{\mathbf{n}} = \mathbf{n}\cdot\hat{\mathbf{J}}$ with arbitrary unit vector $\mathbf{n}$. Eigenvectors of this operator satisfy

$$\hat{J}_{\mathbf{n}}|m;\mathbf{n}\rangle = \hbar m|m;\mathbf{n}\rangle \tag{3.40}$$

with m given by Eq. (3.39) and form an orthonormal basis for $\mathcal{H}_j$. This equation can be obtained by applying an appropriate passive rotation transformation[61] to the second equation in (3.37). [Vectors $|j,m\rangle$ introduced in Eqs. (3.37) are given by $|m;\mathbf{z}\rangle$ in notations of Eq. (3.40).]

As an illustration, we consider the expectation value of $(J_{\mathbf{n}'})^k$ (here k is a positive integer number) in the state represented by the state vector $|m;\mathbf{n}\rangle$,

$$\langle J^k_{\mathbf{n}'}\rangle_{m,\mathbf{n}} = \langle m;\mathbf{n}|\hat{J}^k_{\mathbf{n}'}|m;\mathbf{n}\rangle. \tag{3.41}$$

Although this expectation value can be evaluated explicitly, here we merely discuss the properties of $\langle J^k_{\mathbf{n}'}\rangle_{m,\mathbf{n}}$ dictated by symmetry.

First of all, it is obvious from Eqs. (3.16b) and (3.40) that vector $\hat{R}(\pi\mathbf{n}_0)|-m;\mathbf{n}\rangle$ with $\mathbf{n}_0\perp\mathbf{n}$ is both an eigenvector of $\hat{J}_{-\mathbf{n}} = \hat{R}(\pi\mathbf{n}_0)\hat{J}_{\mathbf{n}}\hat{R}^\dagger(\pi\mathbf{n}_0)$ with eigenvalue $-\hbar m$, i.e., $|-m;-\mathbf{n}\rangle$ in notations of Eq. (3.40), and an eigenvector of $\hat{J}_{\mathbf{n}} = -\hat{J}_{-\mathbf{n}}$ with eigenvalue $+\hbar m$, i.e., $|m;\mathbf{n}\rangle$. In other words,

$$|m;\mathbf{n}\rangle = |-m;-\mathbf{n}\rangle \quad \text{(up to phase factors)}, \tag{3.42}$$

which generalizes the spin 1/2 result Eq. (1.53). Eq. (3.42) shows that the simultaneous change of the signs of m and $\mathbf{n}$ in the state vector $|m;\mathbf{n}\rangle$ does not affect the quantum state, thus leaving the expectation value (3.41) intact,

$$\langle J^k_{\mathbf{n}'}\rangle_{m,\mathbf{n}} = \langle J^k_{\mathbf{n}'}\rangle_{-m,-\mathbf{n}}\,. \tag{3.43}$$

Next, expectation values do not change if both the operator and the state vector are replaced with the corresponding unitary-transformed objects [see Eq. (2.128)]. In particular, a passive rotation [recall Eq. (3.16b)] amounts to replacing $\mathbf{n}$ and $\mathbf{n}'$ with the corresponding rotated vectors without affecting the value of $\langle J^k_{\mathbf{n}'}\rangle_{m,\mathbf{n}}$. This observation suggests that vectors $\mathbf{n}$ and $\mathbf{n}'$ enter $\langle J^k_{\mathbf{n}'}\rangle_{m,\mathbf{n}}$ via their dot product $\mathbf{n}\cdot\mathbf{n}'$ invariant under rotations. For example, rotation by π about $\mathbf{n}_0\propto\mathbf{n}\times\mathbf{n}'$ (so that $\mathbf{n}_0$ is perpendicular to both $\mathbf{n}$ and $\mathbf{n}'$) is equivalent to the simultaneous change of signs of $\mathbf{n}$ and $\mathbf{n}'$, leading to the relation

$$\langle J^k_{\mathbf{n}'}\rangle_{m,\mathbf{n}} = \langle J^k_{-\mathbf{n}'}\rangle_{m,-\mathbf{n}}\,. \tag{3.44}$$

Combining Eqs. (3.43) and (3.44) and taking into account that $\hat{J}_{-\mathbf{n}'} = -\hat{J}_{\mathbf{n}'}$, we obtain

$$\langle J^k_{\mathbf{n}'}\rangle_{m,\mathbf{n}} = (-1)^k\langle J^k_{\mathbf{n}'}\rangle_{-m,\mathbf{n}}\,, \tag{3.45}$$

which shows that $\langle J^k_{\mathbf{n}'}\rangle_{m,\mathbf{n}}$ is an odd/even function of m for odd/even k.

For $k=1$ the expectation value $\langle J_{\mathbf{n}'}\rangle_{m,\mathbf{n}}$ is obviously linear in $\mathbf{n}'$ hence proportional to the dot product $\mathbf{n}\cdot\mathbf{n}'$, i.e., $\langle J_{\mathbf{n}'}\rangle_{m,\mathbf{n}} \propto (\mathbf{n}\cdot\mathbf{n}')$. The m-dependent coefficient here can be found by setting $\mathbf{n}=\mathbf{n}'$ and taking into account Eq. (3.40), which yields

$$\langle J_{\mathbf{n}'}\rangle_{m,\mathbf{n}} = \hbar m(\mathbf{n}\cdot\mathbf{n}'). \tag{3.46}$$

Interestingly, for $m=0$ this expectation value vanishes identically for all $\mathbf{n}$ and $\mathbf{n}'$. (Of course, states with $m=0$ exist only for integer j.)

Finally, we discuss briefly rotations by 2π. Taking into account Eq. (3.40), we write the spectral decomposition of the rotation operator as

$$\hat{R}(\theta\mathbf{n}) = e^{-i(\theta/\hbar)\hat{J}_{\mathbf{n}}} = \sum_m |m;\mathbf{n}\rangle\, e^{-im\theta}\langle m;\mathbf{n}| \tag{3.47}$$

For $\theta=2\pi$ we have

$$e^{-im\theta} \xrightarrow{\theta=2\pi} e^{-2\pi i m} = e^{-2\pi i(-j+\text{integer})} = (-1)^{2j},$$

and Eq. (3.47) reduces to

$$\hat{R}(2\pi\mathbf{n}) = (-1)^{2j}\hat{\mathbb{1}} = \begin{cases} +\hat{\mathbb{1}}, & j=\text{integer},\\ -\hat{\mathbb{1}}, & j=\text{half-integer}, \end{cases} \tag{3.48}$$

irrespective of $\mathbf{n}$. Using $\hat{R}(\pi\mathbf{n}) = \hat{R}(-\pi\mathbf{n})\hat{R}(2\pi\mathbf{n}) = \hat{R}^\dagger(\pi\mathbf{n})\hat{R}(2\pi\mathbf{n})$ [see Eq. (3.11)] and Eq. (3.48), we obtain another curious relation,

$$\hat{R}(\pi\mathbf{n}) = (-1)^{2j}\hat{R}^\dagger(\pi\mathbf{n}), \tag{3.49}$$

which shows that the operator $\hat{R}(\pi\mathbf{n})$, in addition to being unitary, is also Hermitian for integer j and anti-Hermitian for half-integer j.

The result $\hat{R}(2\pi\mathbf{n}) = -\hat{\mathbb{1}}$ for half-integer j [see Eq. (3.48)] may seem surprising, yet it is neither an artifact, nor a paradox. Indeed, although a rotation by 2π should not affect the quantum state, the corresponding state vector may well acquire a phase factor. Eq. (3.48) tells us that this factor is $(-1)^{2j}$. This observation gives rise to the so-called *superselection rule* for physical observables, which, sadly, is well beyond the scope of these notes.

3.2.4 Case study: spin 1/2 - dotting the *i*'s and crossing the *t*'s

In this section, we return to spin 1/2 particles corresponding to $j=1/2$. To emphasize the spin nature of the intrinsic angular momentum, we denote it by $\hat{\mathbf{S}}$ instead of $\hat{\mathbf{J}}$. As in Sec. 1.3.3 [see Eqs. (1.52) and (1.54)], it is convenient to relabel the simultaneous eigenvectors $|m=\pm 1/2;\mathbf{n}\rangle$ of operators $\hat{\mathbf{S}}^2$ and $\hat{S}_{\mathbf{n}} = \mathbf{n}\cdot\hat{\mathbf{S}}$ [see Eq. (3.40)] as

$$|\pm 1/2;\mathbf{n}\rangle \;\longrightarrow\; |\pm\mathbf{n}\rangle. \tag{3.50}$$

Here we focus on $\mathbf{n}=\mathbf{z}$. It follows from Eqs. (3.37) and (3.38) that the non-vanishing matrix elements of operators $\hat{S}_{\mathbf{z}}$ and $\hat{S}_\pm = \hat{S}_{\mathbf{x}} \pm i\hat{S}_{\mathbf{y}}$ in $|\pm\mathbf{z}\rangle$ basis are

$$\langle\pm\mathbf{z}|\hat{S}_{\mathbf{z}}|\pm\mathbf{z}\rangle = \pm\hbar/2, \quad \langle\pm\mathbf{z}|\hat{S}_\pm|\mp\mathbf{z}\rangle = \hbar. \tag{3.51}$$

Matrix elements of the operator

$$\hat{S}_{\mathbf{n}} = \mathbf{n}\cdot\hat{\mathbf{S}} = n_{\mathbf{z}}\hat{S}_{\mathbf{z}} + \frac{1}{2}\bigl(n_+\hat{S}_- + n_-\hat{S}_+\bigr), \quad n_\pm = n_{\mathbf{x}} \pm i n_{\mathbf{y}} \tag{3.52a}$$

then read

$$\langle \pm\mathbf{z}|\hat{S}_{\mathbf{n}}|\pm\mathbf{z}\rangle = \pm\frac{\hbar}{2}n_{\mathbf{z}}, \quad \langle +\mathbf{z}|\hat{S}_{\mathbf{n}}|-\mathbf{z}\rangle = \langle -\mathbf{z}|\hat{S}_{\mathbf{n}}|+\mathbf{z}\rangle^* = \frac{\hbar}{2}n_- \tag{3.52b}$$

[cf. Eqs. (2.72) and (2.73b)], leading to Eq. (2.74) for $\hat{S}_{\mathbf{n}}$ in matrix notations.

As mentioned in Sec. 2.3.2, direct solution of the eigenvalue problem for $\hat{S}_{\mathbf{n}}$ shows that its eigenvalues are $\pm\hbar/2$ and that the corresponding eigenvectors coincide with vectors $|\pm\mathbf{n}\rangle$ given by Eqs. (1.58) and (1.62). These eigenvectors can be also found by acting on the basis vectors $|\pm\mathbf{z}\rangle$ with appropriate rotation operators [recall Eq. (3.18b)]. Indeed, a unit vector $\mathbf{n}$ specified in the spherical polar coordinates by angles θ and ϕ [see Eq. (1.60)] can be obtained by first rotating vector $\mathbf{z}$ by θ about $\mathbf{y}$, and then rotating the resulting rotated vector by ϕ about $\mathbf{z}$. Accordingly, as one can easily verify,

$$|\pm\mathbf{n}\rangle = \hat{R}(\phi\mathbf{z})\hat{R}(\theta\mathbf{y})|\pm\mathbf{z}\rangle \quad \text{(up to phase factors)}. \tag{3.53}$$

In this section, we applied the general theory of the angular momentum to spin 1/2 and recovered the results obtained earlier by a different method. The coincidence may be viewed as an *a posteriori* justification of the assumptions made in Sec. 1.1.3.

4 Time in Quantum Theory

The notion of time made an uncredited appearance in the discussion of consecutive measurements in Sec. 1.1.2 and 1.3.6. Indeed, in our description, the act of measurement is viewed as an instantaneous and irreversible change of a quantum state. If there were no time, there would be no before, no after, and no change. Of course, this does not mean that quantum states do not evolve with time between measurements, when no one is watching. In fact, they do. This evolution is the subject of the present Chapter.

4.1 Evolution operator and Hamiltonian

4.1.1 Isolated quantum system

The ***time evolution*** (also known as the ***time translation***) is another example of a continuous (thus unitary) symmetry transformation of the type discussed in Sec. 2.5.3. Because time is uniform, results of measurements carried out in an isolated laboratory may depend only on the duration of the experiment. Accordingly, the unitary ***evolution operator*** $\hat{T}$ relating the state vector at time t with that at some later time $t+\tau$ can be taken to be independent of t,

$$|\psi(t+\tau)\rangle = \hat{T}(\tau)|\psi(t)\rangle \quad \text{(up to phase factors)}. \tag{4.1a}$$

Moreover, it should be possible to choose the phases of the evolution operators so that they have the group property

$$\hat{T}(\tau)\hat{T}(\tau') = \hat{T}(\tau+\tau'). \tag{4.1b}$$

The unitarity of $\hat{T}$ and Eq. (4.1b) imply [cf. Eqs. (2.121) and (2.125)] that the evolution operator has the form

$$\hat{T}(\tau) = e^{-i\tau\hat{H}/\hbar}, \tag{4.2}$$

where the generator of time evolution $\hat{H}$ is an independent of time Hermitian operator.

The observable corresponding to this operator can be identified by exploring the similarity with the canonical transformations in Classical Mechanics. Such analysis suggests that $\hat{H}$ represents the ***Hamiltonian*** of the system; for our purposes, the Hamiltonian is synonymous with the ***energy***. Just as it is the case with Eq. (3.14), the role of the Planck constant $\hbar$ in Eq. (4.2) is to render the argument of the exponent $\tau\hat{H}/\hbar$ dimensionless. Importantly, it is the same $\hbar$ that appears in Eq. (3.14); this ensures that the time dependence of various expectation values is governed by the classical equations of motion.

Note that the Hamiltonian and thus its eigenvalues known as ***eigenenergies*** are defined up to an arbitrary constant [see the remark after Eq. (2.125)]. This constant does not enter final results provided that these results are formulated in terms of observable quantities.

4.1.2 Quantum system in a classical environment

The formalism of Sec. 4.1.1 can be adapted for an *approximate* description of evolution of quantum systems interacting with a classical environment. In this approach, it is assumed from the outset that the dependence on time can be described by the equation

$$|\psi(t_1)\rangle = \hat{T}(t_1,t_2)|\psi(t_2)\rangle \tag{4.3a}$$

[cf. Eq. (4.1a)], where the evolution operator $\hat{T}$ is unitary,

$$\hat{T}(t_1,t_2)\hat{T}^\dagger(t_1,t_2) = \hat{T}^\dagger(t_1,t_2)\hat{T}(t_1,t_2) = \hat{\mathbb{1}}, \tag{4.3b}$$

and is endowed with the group property

$$\hat{T}(t_1,t_2)\hat{T}(t_2,t_3) = \hat{T}(t_1,t_3). \tag{4.3c}$$

Unlike for an isolated system [see Eq. (4.1a)], there is no reason for $\hat{T}(t_1,t_2)$ to depend only on the difference of t_1 and t_2. This implies, in particular, that the evolution operators in the left-hand side of Eq. (4.3c) do not have to commute.

In the discussion of consequences of Eqs. (4.3), we will follow closely the derivation in Sec. 2.5.3. For $t_1 = t_2 = t$ Eq. (4.3c) reduces to $\hat{T}^2(t,t) = \hat{T}(t,t)$. Multiplying both sides of this equation by $\hat{T}^\dagger(t,t)$ and taking into account the unitarity property (4.3b), we obtain

$$\hat{T}(t,t) = \hat{\mathbb{1}}. \tag{4.4a}$$

Eqs. (4.3c) and (4.4a) imply that the identity operator can be written as $\hat{\mathbb{1}} = \hat{T}(t_1,t_2)\hat{T}(t_2,t_1)$. Multiplying both sides of this expression by $\hat{T}^\dagger(t_1,t_2)$ on the left and using Eq. (4.3b), we get the relation

$$\hat{T}^\dagger(t_1,t_2) = \hat{T}(t_2,t_1), \tag{4.4b}$$

which shows that $\hat{T}^\dagger$ describes the evolution backward in time.

With Eq. (4.4a) taken into account, the first two terms in the expansion of $\hat{T}(t+\delta t,t)$ in Taylor series in δt can be written as

$$\hat{T}(t+\delta t,t) = \hat{\mathbb{1}} - i(\delta t/\hbar)\hat{H}(t) + \ldots \tag{4.5a}$$

Comparison of this expansion with either Eq. (4.3b) or with Eq. (4.4b)[64] shows that the Hamiltonian $\hat{H}(t)$ is Hermitian,

$$\hat{H}(t) = \hat{H}^\dagger(t). \tag{4.5b}$$

The classical environment enters our description of the time evolution via parameters of the Hamiltonian. Importantly, these parameters and thus the Hamiltonian itself may explicitly depend on time.

It follows from Eqs. (4.3c) and (4.5a) that

$$\hat{T}(t+\delta t,t') = \hat{T}(t+\delta t,t)\hat{T}(t,t') = \big[\hat{\mathbb{1}} - i(\delta t/\hbar)\hat{H}(t) + \ldots\big]\hat{T}(t,t'),$$

which yields the first-order differential equation

$$i\hbar\,\frac{\partial}{\partial t}\,\hat{T}(t,t') = \hat{H}(t)\hat{T}(t,t'). \tag{4.6}$$

It can be shown that Eq. (4.6) subject to the initial condition $\hat{T}(t,t')\big|_{t\to t'} = \hat{\mathbb{1}}$ [see Eq. (4.4a)] yields a unique solution for the evolution operator $\hat{T}(t,t')$. Unfortunately, because in general Hamiltonian operators at different times do not commute, finding this solution is difficult and

[64] Note that $\hat{T}(t,t+\delta t) = \hat{T}\big((t+\delta t)-\delta t, t+\delta t\big) = \hat{\mathbb{1}} + i(\delta t/\hbar)\hat{H}(t+\delta t) + \ldots = \hat{\mathbb{1}} + i(\delta t/\hbar)\hat{H}(t) + \ldots$

quite often impossible. However, there are special cases when Eq. (4.6) can be solved exactly. In particular, Eq. (4.6) simplifies dramatically if $[\hat{H}(t_1), \hat{H}(t_2)] = \hat{\mathbb{0}}$ for all t_1 and t_2 in the interval $t \geq t_{1,2} \geq t'$. It is easy to show that in this case the evolution operator is given by[65]

$$\hat{T}(t,t') = \exp\left[-\frac{i}{\hbar}\int_{t'}^{t} d\tau \hat{H}(\tau)\right]. \tag{4.7a}$$

If the Hamiltonian is independent of time, Eq. (4.7a) further simplifies to

$$\hat{T}(t,t') = \hat{T}(t-t') = e^{-i(t-t')\hat{H}/\hbar}, \tag{4.7b}$$

i.e., in this case the evolution proceeds as if the system were isolated [cf. Eqs. (4.1b) and (4.2)]. As in Classical Mechanics, systems described by time-independent Hamiltonians are often referred to as ***conservative***. (The origin of this term is explained in Sec. 4.1.4 below.)

The evolution operator can also be found exactly if the dependence of the Hamiltonian on time is piece-wise constant. For example, for

$$\hat{H}(t) = \begin{cases} \hat{H}_1, & t < t_0, \\ \hat{H}_2, & t > t_0 \end{cases} \tag{4.8a}$$

Eqs. (4.3c) and (4.7b) give

$$\hat{T}(t,t')\big|_{t>t_0>t'} = \hat{T}(t,t_0)\hat{T}(t_0,t') = e^{-i(t-t_0)\hat{H}_2/\hbar}\, e^{-i(t_0-t')\hat{H}_1/\hbar}. \tag{4.8b}$$

Note that the order of the operators in the right-hand side matters: the two exponents do not commute if $[\hat{H}_1, \hat{H}_2] \neq \hat{\mathbb{0}}$. Note also that although the dependence on time in Eq. (4.8a) is manifestly non-analytic, the evolution operator $\hat{T}(t,t')$ and thus the state vector $|\psi(t)\rangle$ depend on t continuously. Indeed, as it is obvious from Eqs. (4.6), for these dependences to be continuous it is sufficient that matrix elements of the Hamiltonian remain finite at all t.

Finally, it should be emphasized that whereas the consideration of isolated quantum systems in Sec. 4.1.1 is rooted in the fundamental symmetry principles and is therefore exact, describing the interaction of a quantum system with a classical environment as a unitary evolution is at best an approximation. As such, it has obvious and not so obvious limitations. For example, any attempt to derive Eq. (2.66) by treating the interaction of the system of interest with the measurement apparatus in this manner is bound to fail because state vectors immediately after the projective measurement are not related to those immediately before the measurement via a unitary transformation.

[65] Integrals of operators such as that in Eq. (4.7a) are defined in the usual manner. Imagine that the integration range $t' < \tau < t$ is partitioned in n intervals of equal length: $t' = \tau_1 < \tau_2 < \ldots < \tau_n < \tau_{n+1} = t$, where $\tau_{k+1} = \tau_k + \delta\tau$ with $\delta\tau = (\tau_{n+1} - \tau_1)/n = (t-t')/n$. The integral is the limit

$$\lim_{n\to\infty} \delta\tau\big\{\hat{H}(\tau_1) + \hat{H}(\tau_2) + \ldots + \hat{H}(\tau_n)\big\} = \int_{t'}^{t} d\tau \hat{H}(\tau).$$

4.1.3 Survival probability and time-energy uncertainty relation

Suppose at $t = 0$ the system is in the state[20,22] $|\psi(0)\rangle$. By Born's rule, the probability to find it in the same quantum state at later times, the so-called ***survival probability*** (also known as the *dynamical fidelity* or the *Loschmidt echo*), is given by

$$\mathcal{P}(t) = \big|\langle\psi(0)|\psi(t)\rangle\big|^2. \tag{4.9}$$

Assuming that $\hat{H}$ is independent of time and expanding $\hat{T}(t,0) = e^{-it\hat{H}/\hbar}$ [see Eq. (4.7b)] to second order in t, we write the state vector $|\psi(t)\rangle$ at small t as

$$|\psi(t)\rangle = \hat{T}(t,0)|\psi(t)\rangle = \left(\hat{\mathbb{1}} - \frac{it}{\hbar}\hat{H} - \frac{t^2}{2\hbar^2}\hat{H}^2 + \dots\right)|\psi(0)\rangle.$$

Substituting this expansion into Eq. (4.9) and denoting the observable corresponding to $\hat{H}$ (the energy) by E, we obtain

$$\mathcal{P}(t) = \left|1 - \frac{it}{\hbar}\langle E\rangle - \frac{t^2}{2\hbar^2}\langle E^2\rangle + \dots\right|^2 = 1 + \frac{t^2}{\hbar^2}\langle E\rangle^2 - \frac{t^2}{\hbar^2}\langle E^2\rangle + \dots,$$

which gives[66]

$$\mathcal{P}(t) = 1 - \left(\frac{t\Delta E}{\hbar}\right)^2 + \dots, \tag{4.10}$$

where $\Delta E = \sqrt{\langle E^2\rangle - \langle E\rangle^2}$ is the uncertainty of energy. (It will be shown in Sec. 4.1.4 that $\langle E\rangle$, $\langle E^2\rangle$, and thus ΔE are independent of time if $\hat{H} = \text{const}$.)

The result (4.10) is applicable as long as the second term in the right-hand side remains small, i.e., at $t \ll \Delta t$, where

$$\Delta t \sim \hbar/\Delta E. \tag{4.11a}$$

This relation can be regarded as a crude estimate of a ***lifetime*** of a quantum state. Unless the initial state vector $|\psi(0)\rangle$ is an eigenvector of $\hat{H}$, the uncertainty of energy ΔE is strictly positive [see Sec. 2.4.3], resulting in a finite lifetime $\Delta t < \infty$.

One can also write Eq. (4.11a) as

$$\Delta t\,\Delta E \sim \hbar, \tag{4.11b}$$

which can be interpreted as the ***time-energy uncertainty relation***. It should be emphasized, however, that time enters the non-relativistic quantum theory as a parameter as opposed to an observable. (In particular, it is not an accident that there is no time operator.) Therefore, notwithstanding the superficial similarity, Eq. (4.11b) does not have the same status as Eq. (2.98).

[66] Had we neglected the second order in t contribution in the expansion of $\hat{T}(t,0)$, we would have ended up with $\mathcal{P}(t) = 1 + (t\langle E\rangle/\hbar)^2 + \dots$ instead of Eq. (4.10). This expression is clearly wrong, for two reasons. First, the probability cannot exceed 1. Second, probabilities and expectation values cannot depend on $\langle E\rangle$. Indeed, as mentioned in the end of Sec. 4.1.1, the Hamiltonian and thus $\langle E\rangle$ are defined only up to a constant. On the contrary, the uncertainty $\Delta E = \sqrt{\langle (E - \langle E\rangle)^2\rangle}$ is not affected by the replacement $E \to E + E_0$, hence its presence in the result (4.10) is legitimate.

4.1.4 Schrödinger and Heisenberg pictures

Differentiating both sides of the relation $|\psi(t)\rangle = \hat{T}(t,t')|\psi(t')\rangle$ [see Eq. (4.3a)] with the help of Eq. (4.6), we obtain the ***Schrödinger equation***

$$i\hbar \frac{d}{dt}|\psi(t)\rangle = \hat{H}(t)|\psi(t)\rangle. \tag{4.12a}$$

Supplied with an appropriate initial condition such as $|\psi(t)\rangle\big|_{t\to t_0} = |\psi(t_0)\rangle$, this equation yields a unique time-dependent state vector $|\psi(t)\rangle$. Having found this vector, one can evaluate various expectation values at time t, e.g.,

$$\langle f(A)\rangle_t = \langle\psi(t)|\hat{f}(\hat{A})|\psi(t)\rangle. \tag{4.12b}$$

In the right-hand-side of Eq. (4.12b) the state vector $|\psi(t)\rangle$ evolves as prescribed by the Schrödinger equation (4.12a), whereas the operator $\hat{A}$ representing the observable of interest remains as is. This description corresponds to Eqs. (2.120) and (2.126a), and is known as the ***Schrödinger picture*** of time evolution. As discussed in Sec. 2.5.3, an alternative (albeit equivalent) approach to continuous symmetry transformations is to transform the operators, leaving the state vectors as they are [see Eq. (2.127a)]. This observation gives rise to the ***Heisenberg picture*** of time evolution in which the expectation value (4.12b) is written as [cf. Eqs. (2.126b) and (2.127a)]

$$\langle f(A)\rangle_t = \langle\psi(t_0)|\hat{f}(\hat{A}_t)|\psi(t_0)\rangle, \tag{4.13a}$$

where the state vector $|\psi(t_0)\rangle$ is independent of t and

$$\hat{A}_t = \hat{T}^\dagger(t,t_0)\hat{A}\hat{T}(t,t_0) \tag{4.13b}$$

is the ***evolved*** or ***dressed*** operator (as opposed to $\hat{A}$ referred to as ***bare***). Differentiation of this operator with the help of Eq. (4.6) yields the ***Heisenberg equation of motion***

$$\frac{d}{dt}\hat{A}_t = \frac{i}{\hbar}[\hat{H}_t, \hat{A}_t] \tag{4.14a}$$

[cf. Eq. (2.127b)]. Here $\hat{H}_t$ is the dressed Hamiltonian defined similarly to $\hat{A}_t$ [see Eq. (4.13b)],

$$\hat{H}_t = \hat{T}^\dagger(t,t_0)\hat{H}(t)\hat{T}(t,t_0). \tag{4.14b}$$

(Recall that, as discussed in Sec. 4.1.2, the bare Hamiltonian $\hat{H}(t)$ may depend on time.)

The unitarity of the evolution operator implies that commutators of dressed operators such as that in the right-hand side of Eq. (4.14a) coincide with the dressed commutators of the corresponding bare operators:

$$[\hat{H}_t, \hat{A}_t] = \hat{T}^\dagger(t,t_0)[\hat{H}(t), \hat{A}]\hat{T}(t,t_0) = [\hat{H}(t), \hat{A}]_t. \tag{4.14c}$$

If the bare Hamiltonian is independent of time, then $[\hat{H}, \hat{T}(t,t_0)] = \hat{\mathbb{0}}$ [see Eq. (4.7b)], and Eq. (4.14b) gives $\hat{H}_t = \hat{H}$. If the operator $\hat{A}$ representing the observable of interest commutes

with the Hamiltonian, then $[\hat{H}, \hat{A}_t] = [\hat{H}, \hat{A}]_t = \hat{\mathbb{0}}$, and Eq. (4.14a) gives $d\hat{A}_t/dt = \hat{\mathbb{0}}$. Accordingly, in this case $\hat{A}$ remains bare at all times, $\hat{A}_t = \hat{A}$, and the expectation value of any function of A [see Eq. (4.13a)] is constant. Borrowing the term from Classical Mechanics, such observables are referred to as ***integrals of motion***. An obvious example of an integral of motion is the energy itself: for $\hat{A} = \hat{H}$ we have $\langle f(E) \rangle = \text{const}$, i.e., the energy of a system whose evolution is governed by a time-independent Hamiltonian is conserved. It is for this reason such systems are called ***conservative***.

The main advantage of working in the Heisenberg picture is that the equation of motion (4.14a) is somewhat more intuitive than the Schrödinger equation (4.12a). Indeed, as mentioned above, operators, unlike state vectors, have classical analogs. Moreover, as we shall see, Heisenberg equations of motion have the same form as the corresponding classical equations.

4.2 Case study: quantum spin in a classical magnetic field

4.2.1 Hamiltonian

In magnetostatics the energy of a point magnetic moment $\boldsymbol{\mu}$ placed in a magnetic field $\mathbf{B}$ is given by

$$E = -\boldsymbol{\mu} \cdot \mathbf{B}. \tag{4.15a}$$

If the magnetic moment is due to motion of particles with the same charge-to-mass ratio such as electrons in the classical model of an atom, then $\boldsymbol{\mu}$ is proportional to the angular momentum $\mathbf{J}$,

$$\boldsymbol{\mu} = \gamma \mathbf{J}. \tag{4.15b}$$

The coefficient γ here is called the ***gyromagnetic ratio***; this ratio may have either sign. Substituting Eq. (4.15b) into Eq. (4.15a), we write the energy as

$$E = \omega \, \mathbf{n} \cdot \mathbf{J}, \tag{4.15c}$$

where $\omega = |\gamma \mathbf{B}|$ is the so-called ***Larmor frequency*** (its physical meaning will be elucidated in Sec. 4.2.2) and $\mathbf{n}$ is a unit dimensionless vector parallel to $-\gamma \mathbf{B}$.

One can now *guess* the quantum Hamiltonian by promoting E and $\mathbf{J}$ in Eq. (4.15c) to the corresponding Hermitian operators: $E \to \hat{H}$, $\mathbf{J} \to \hat{\mathbf{S}}$. The resulting Hamiltonian describing the interaction of a quantum spin with a classical magnetic field [recall Sec. 4.1.2] reads

$$\hat{H} = \omega \, \mathbf{n} \cdot \hat{\mathbf{S}} = \omega \hat{S}_{\mathbf{n}}. \tag{4.16}$$

Importantly, both ω and $\mathbf{n}$ here may explicitly depend on time. More sophisticated derivations yield essentially the same expression.

4.2.2 Larmor precession

A classical magnetic moment placed in a time-independent magnetic field feels the torque

$$\frac{d}{dt} \mathbf{J} = \boldsymbol{\mu} \times \mathbf{B}. \tag{4.17a}$$

Substituting here $\boldsymbol{\mu}$ from Eq. (4.15b), we obtain the equation of motion

$$\frac{d}{dt}\mathbf{J} = \omega\,\mathbf{n}\times\mathbf{J}, \tag{4.17b}$$

where ω and $\mathbf{n}$ have the same meaning as in Eq. (4.15c). Comparison with Eq. (3.2b) shows that vector $\mathbf{J}$ *precesses* about $\mathbf{n}$ with angular frequency ω, i.e., $\mathbf{J}(t)$ is obtained by rotating vector $\mathbf{J}(0)$ by the angle ωt about $\mathbf{n}$, or, in notations of Sec. 3.1.1,

$$\mathbf{J}(t) = [\mathbf{J}(0)]_{\omega t\mathbf{n}}. \tag{4.17c}$$

Quantum spins exhibit a similar precession. Indeed, Eq. (3.25) with $\hat{\mathbf{A}} = \hat{\mathbf{J}} = \hat{\mathbf{S}}$ and Eq. (4.16) show that

$$[\hat{H}, \hat{\mathbf{S}}] = [\omega\,\mathbf{n}\cdot\hat{\mathbf{S}}, \hat{\mathbf{S}}] = -i\hbar\omega\,\mathbf{n}\times\hat{\mathbf{S}}.$$

The Heisenberg equation of motion for the dressed spin operator $\hat{\mathbf{S}}_t = \hat{T}^\dagger(t,0)\,\hat{\mathbf{S}}\,\hat{T}(t,0)$ [see Eqs. (4.13b), (4.14c), and (4.16)] then reads

$$\frac{d}{dt}\hat{\mathbf{S}}_t = \omega\,\mathbf{n}\times\hat{\mathbf{S}}_t. \tag{4.18}$$

As expected [see the remark in the end of Sec. 4.1.4], this equation has the same form as its classical analog (4.17b). Note also that changing the variable to $\theta = \omega t$ in Eq. (4.18) yields Eq. (3.8a) describing a rotation transformation. Indeed, the evolution operator for a quantum spin placed in a constant magnetic field coincides with the rotation operator,

$$\hat{T}(t,0) = e^{-it\hat{H}/\hbar} = e^{-it\omega\hat{S}_{\mathbf{n}}/\hbar} = \hat{R}(\omega t\mathbf{n}) \tag{4.19}$$

[see Eqs. (3.10), (3.14), (4.7b), and (4.16)], hence the dressed spin operator $\hat{\mathbf{S}}_t$ is nothing but the rotated spin operator: $\hat{\mathbf{S}}_t = [\hat{\mathbf{S}}]_{\omega t\mathbf{n}} = \hat{R}^\dagger(\omega t\mathbf{n})\,\hat{\mathbf{S}}\,\hat{R}(\omega t\mathbf{n})$ [cf. Eq. (3.10)].

So far, we discussed the evolution in the Heisenberg picture. In the Schrödinger picture [see Sec. 4.1.4] the dependence on time is assigned to state vectors rather than operators. This dependence is particularly easy to understand for spin 1/2. Indeed, as explained in Sec. 1.3.3, pure states of spin 1/2 are in one-to-one correspondence with unit dimensionless vectors, the Bloch vectors. Comparison of

$$|\mathbf{n}(t)\rangle = \hat{T}(t,0)|\mathbf{n}(0)\rangle = \hat{R}(\omega t\mathbf{n})|\mathbf{n}(0)\rangle \quad \text{(up to phase factors)} \tag{4.20a}$$

[see Eq. (4.19)] with Eq. (3.18b) shows that the time-dependent Bloch vector $\mathbf{n}(t)$ is obtained by rotating the initial Bloch vector $\mathbf{n}(0)$ by ωt about the direction of the magnetic field $\mathbf{n}$. That is, $\mathbf{n}(t)$ is given by $\mathbf{n}(t) = [\mathbf{n}(0)]_{\omega t\mathbf{n}}$ and obeys the equation

$$\frac{d}{dt}\mathbf{n}(t) = \omega\,\mathbf{n}\times\mathbf{n}(t). \tag{4.20b}$$

It is easy to see that this equation is consistent with the Heisenberg equation of motion (4.18). Indeed, Eq. (4.18) implies that the evolution of the expectation value $\langle\mathbf{S}\rangle_t = \langle\mathbf{n}(0)|\hat{\mathbf{S}}_t|\mathbf{n}(0)\rangle$ is governed by the equation

$$\frac{d}{dt}\langle\mathbf{S}\rangle_t = \omega\,\mathbf{n}\times\langle\mathbf{S}\rangle_t. \tag{4.21}$$

Substituting here $\langle \mathbf{S} \rangle_t = (\hbar/2)\mathbf{n}(t)$ [recall Eq. (2.71)], we recover Eq. (4.20b).

Finally, it should be emphasized that the Hamiltonian (4.16) is an approximation. In real life, a quantum spin placed in a magnetic field will eventually *relax* to the state with the lowest possible energy (the *ground state*); for spin 1/2 this state is $|-\mathbf{n}\rangle$. The dominant mechanism of such relaxation depends on the nature of the spin and of the environment it interacts with. Since the Hamiltonian (4.16) does not include processes responsible for the relaxation, describing the evolution in terms of this Hamiltonian is justified only if the Larmor frequency is much larger than the relaxation rate, i.e., if the magnetic field is sufficiently strong. It is also worth noting that because the Larmor precession has a classical origin, it is not surprising that $\hbar$ does not enter the expression for the Larmor frequency.

4.2.3 Magnetic resonance

Consider the time-dependent Hamiltonian

$$\hat{H}(t) = \omega_0 \hat{S}_{\mathbf{z}} + \omega_1 \big[\hat{S}_{\mathbf{x}} \cos(\omega t) + \hat{S}_{\mathbf{y}} \sin(\omega t)\big] \tag{4.22a}$$

describing a quantum spin placed in a superposition of a constant magnetic field in z-direction and a rotating with angular frequency ω field in (x,y)-plane. (The latter contribution represents, for example, the field of a circularly-polarized wave). It is remarkable that even though $[\hat{H}(t), \hat{H}(t')] \neq \hat{\mathbb{0}}$, the evolution operator can be found exactly. The exact solution is possible because the Hamiltonian (4.22a) can be also written as

$$\hat{H}(t) = \hat{R}(\omega t \mathbf{z}) \big[\omega_0 \hat{S}_{\mathbf{z}} + \omega_1 \hat{S}_{\mathbf{x}}\big] \hat{R}^\dagger(\omega t \mathbf{z}), \tag{4.22b}$$

where $\hat{R}(\omega t \mathbf{z})$ is the rotation operator [see Eq. (3.16b)].

The method we use is rather general and applicable to other problems as well. The idea (due to Dirac) is to seek the solution of the equation

$$i\hbar \frac{d}{dt} \hat{T}(t) = \hat{H}(t) \hat{T}(t), \quad \hat{T}(t) \equiv \hat{T}(t,0) \tag{4.23}$$

[see Eq. (4.6)] in the form

$$\hat{T}(t) = \hat{T}_0(t) \hat{T}_1(t), \tag{4.24a}$$

where $\hat{T}_0(t)$ is the solution of the equation

$$i\hbar \frac{d}{dt} \hat{T}_0(t) = \hat{H}_0(t) \hat{T}_0(t) \tag{4.24b}$$

subject to the initial condition $\hat{T}_0(0) = \hat{\mathbb{1}}$; here $\hat{H}_0(t)$ is Hermitian operator that has units of energy. Just as $\hat{T}(t)$, the evolution operator $\hat{T}_0(t)$ defined by Eq. (4.24b) is unitary, and so is the operator

$$\hat{T}_1(t) = \hat{T}_0^\dagger(t) \hat{T}(t). \tag{4.24c}$$

Differentiating $\hat{T}_1(t)$ with the help of Eqs. (4.23) and (4.24b), we obtain the equation

$$i\hbar \frac{d}{dt} \hat{T}_1(t) = \hat{H}_1(t) \hat{T}_1(t), \quad \hat{H}_1(t) = \hat{T}_0^\dagger(t) \big[\hat{H}(t) - \hat{H}_0(t)\big] \hat{T}_0(t). \tag{4.24d}$$

The problem of finding $\hat{T}(t)$ therefore reduces to choosing such $\hat{H}_0(t)$ that Eqs. (4.24b) and (4.24d) are easier to solve than the original equation (4.23).

For the problem at hand, Eq. (4.22b) suggests the choice $\hat{H}_0 = \omega \hat{S}_{\mathbf{z}}$. Since $\hat{H}_0$ is independent of time, the corresponding evolution operator is given by $\hat{T}_0(t) = \hat{R}(\omega t \mathbf{z})$ [recall Eq. (4.19)]. Moreover, Eqs. (4.22) and the second equation in (4.24d) yield

$$\hat{H}_1 = \omega_1 \hat{S}_{\mathbf{x}} - (\omega - \omega_0)\hat{S}_{\mathbf{z}} = \Omega\, \mathbf{n}_\Omega \cdot \hat{\mathbf{S}} = \Omega \hat{S}_{\mathbf{n}_\Omega}, \tag{4.25a}$$

where the ***Rabi frequency*** Ω and the unit dimensionless vector $\mathbf{n}_\Omega$ are given by

$$\Omega = \sqrt{(\omega - \omega_0)^2 + \omega_1^2}, \quad \mathbf{n}_\Omega = \Omega^{-1}\big[\omega_1 \mathbf{x} - (\omega - \omega_0)\mathbf{z}\big]. \tag{4.25b}$$

This Hamiltonian is also independent of time, hence $\hat{T}_1(t) = \hat{R}(\Omega t \mathbf{n}_\Omega)$. The evolution operator we seek [see Eq. (4.24a)] is then given by the product of two rotation operators,

$$\hat{T}(t) = \hat{R}(\omega t \mathbf{z})\hat{R}(\Omega t \mathbf{n}_\Omega). \tag{4.26}$$

With the exact evolution operator at our disposal, evaluation of various probabilities and expectation values is straightforward. As an example, we focus here on spin 1/2 and ask what is the probability that the spin initially (at $t = 0$) in the state $|-\mathbf{z}\rangle$ will be found in the state $|+\mathbf{z}\rangle$ at later time t, the so-called *spin flip* probability $\mathcal{P}(t)$. By Born's rule [see Eq. (1.47)], this probability is given by

$$\mathcal{P}(t) = \big|\langle +\mathbf{z}|\hat{T}(t)|-\mathbf{z}\rangle\big|^2. \tag{4.27a}$$

Taking into account Eqs. (1.55) and (4.26) and the relation $\hat{R}^\dagger(\omega t \mathbf{z})|+\mathbf{z}\rangle = e^{i\omega t}|+\mathbf{z}\rangle$, we write $\mathcal{P}(t)$ as[67]

$$\mathcal{P}(t) = \big|\langle +\mathbf{z}|\hat{R}(\Omega t \mathbf{n}_\Omega)|-\mathbf{z}\rangle\big|^2 = \big|\langle +\mathbf{z}|\mathbf{n}(t)\rangle\big|^2 = \frac{1}{2}\big[1 + \mathbf{z}\cdot\mathbf{n}(t)\big], \tag{4.27b}$$

where $|\mathbf{n}(t)\rangle = \hat{R}(\Omega t \mathbf{n}_\Omega)|-\mathbf{z}\rangle$. The time-dependent Bloch vector $\mathbf{n}(t)$ here is obtained by rotating vector $-\mathbf{z}$ by Ωt about $\mathbf{n}_\Omega$. Substituting $\mathbf{n} = \mathbf{n}_\Omega$, $\mathbf{r} = -\mathbf{z}$, and $\theta = \Omega t$ into Eq. (3.2c) and taking into account Eq. (4.25b), we arrive at

$$\mathcal{P}(t) = p(\omega)\sin^2(\Omega t/2) \tag{4.28a}$$

with the frequency-dependent coefficient

$$p(\omega) = 1 - (\mathbf{z}\cdot\mathbf{n}_\Omega)^2 = (\omega_1/\Omega)^2 = \frac{\omega_1^2}{(\omega - \omega_0)^2 + \omega_1^2}. \tag{4.28b}$$

This coefficient features a ***resonant peak*** with a characteristic *Lorentzian* shape. The peak is centered at $\omega = \omega_0$, its width is controlled by ω_1, and its height is given by $\max\{p(\omega)\} = 1$.

The exactly solvable model studied here illustrates a very general and important phenomenon: even a very weak periodic perturbation with frequency ω is capable of inducing *resonant transitions* between the unperturbed states with energy difference close to $\hbar\omega$.

[67] Alternatively, one can write the spin flip probability as $\mathcal{P}(t) = 1 - \big|\langle -\mathbf{z}|\hat{R}(\Omega t \mathbf{n}_\Omega)|-\mathbf{z}\rangle\big|^2$, substitute here the rotation operator in the form $\hat{R}(\Omega t \mathbf{n}_\Omega) = \cos(\Omega t/2)\hat{\mathbb{1}} - i\sin(\Omega t/2)(2\hat{S}_{\mathbf{n}_\Omega}/\hbar)$, and use Eq. (1.20), with the same result (4.28).

5 Quantum Mechanics of One-Dimensional Motion

5.1 Mathematics of one-dimensional continuum

5.1.1 Delta-function and Fourier transform

Dirac invented an ingenious method of handling *functionals* of the type $f(x) \longmapsto f(0)$. The idea is to write $f(0)$ formally as an integral

$$\int_{-\infty}^{\infty} dx\, f(x)\,\delta(x) = f(0). \tag{5.1a}$$

As discussed below, the **delta-function** $\delta(x)$ introduced here can often be visualized as an infinitely narrow and infinitely high peak with a unit area.[68] This interpretation suggests that changing the limits of integration in Eq. (5.1a) will not affect the result as long as the point at which the delta-function peaks, i.e., $x = 0$, remains within the integration range. Indeed,

$$\int_a^b dx\, f(x)\delta(x) = \int_{-\infty}^{\infty} dx\,\delta(x) \times \begin{cases} f(x), & a < x < b \\ 0, & \text{otherwise} \end{cases} = \begin{cases} f(0), & a < 0 < b, \\ 0, & \text{otherwise.} \end{cases} \tag{5.1b}$$

Alternatively, $\delta(x)$ can be viewed as a derivative of the ***step function*** $\theta(x)$,

$$\delta(x) = \frac{d}{dx}\theta(x), \quad \theta(x) = \begin{cases} 1, & x > 0, \\ 0, & x < 0, \end{cases} \tag{5.2}$$

which one can easily verify with the help of the integration by parts,

$$\int_a^b dx\, f(x)\,\frac{d}{dx}\theta(x) = f(x)\theta(x)\Big|_a^b - \int_a^b dx\,\theta(x)\,\frac{d}{dx}f(x) = \begin{cases} f(0), & a < 0 < b, \\ 0, & \text{otherwise.} \end{cases}$$

Eq. (5.2) illustrates the general property: delta-functions pop up whenever one differentiates a discontinuous function.

The delta-function can often be dealt with as if it were an ordinary function. For example, the integral $\int dx\, f(x)\,\delta(x - x_0)$ is evaluated by changing the integration variable to $y = x - x_0$:

$$\int_{-\infty}^{\infty} dx\, f(x)\,\delta(x - x_0) = \int_{-\infty}^{\infty} dy\, f(x_0 + y)\,\delta(y) = f(x_0 + y)\big|_{y=0} = f(x_0). \tag{5.3}$$

Consider now a more complicated integral, $\int dx\, f(x)\delta[g(x)]$, where the argument of the delta-function $g(x)$ is an ordinary real function that has one or more isolated zeros x_i,

$$g(x_i) = 0, \quad g'(x_i) = \frac{dg(x)}{dx}\Big|_{x=x_i} \neq 0. \tag{5.4a}$$

Since, obviously, only the immediate vicinities of the points x_i contribute to the integral, it can be written as a sum of contributions from these points,

$$\int_{-\infty}^{\infty} dx\, f(x)\delta[g(x)] = \sum_i \int_{a_i}^{b_i} dx\, f(x)\,\delta[g(x)], \quad a_i < x_i < b_i, \tag{5.4b}$$

[68] Dirac's delta-function is not a function in the usual sense. Rather, it should be understood as an instruction on evaluation of integrals in which it appears. In mathematics, $\delta(x)$ is an example of a ***generalized function***; Dirac called it an *improper function*.

where each interval (a_i, b_i) contains a single zero x_i of $g(x)$. Changing the integration variable to $y = g(x)$, we find

$$\int_{a_i}^{b_i} dx\, f(x)\,\delta[g(x)] = \int_{g(a_i)}^{g(b_i)} dy\, \frac{f[x(y)]}{g'[x(y)]}\,\delta(y) = \frac{f(x_i)}{g'(x_i)}\,\operatorname{sign}\big[g(b_i) - g(a_i)\big] = \frac{f(x_i)}{\big|g'(x_i)\big|}, \quad (5.4c)$$

where we introduced the ***sign function***

$$\operatorname{sign}(x) = \frac{x}{|x|} = 2\theta(x) - 1 = \begin{cases} +1, & x > 0, \\ -1, & x < 0, \end{cases} \quad (5.5)$$

and took into account the obvious relation $\operatorname{sign}\big[g(b_i) - g(a_i)\big] = \operatorname{sign}[g'(x_i)]$. Eq. (5.4c) is equivalent to the identity

$$\delta[g(x)] = \sum_i \frac{1}{\big|g'(x_i)\big|}\,\delta(x - x_i), \quad (5.6)$$

which yields the relations

$$\delta(ax) = \frac{1}{|a|}\,\delta(x), \quad \delta\big(x^2 - a^2\big) = \frac{1}{|2a|}\big[\delta(x + a) + \delta(x - a)\big], \quad (5.7)$$

where $a \neq 0$. The first equation here implies that $\delta(-x) = \delta(x)$. In other words, the delta-function is an even function of its argument.

The ***derivative*** of the delta-function can be defined as the limit

$$\frac{d}{dx}\,\delta(x) = \lim_{a \to +0} \frac{1}{a}\big[\delta(x + a) - \delta(x)\big]. \quad (5.8a)$$

Taking into account Eq. (5.3), we find

$$\int_{-\infty}^{\infty} dx\, f(x)\,\frac{d}{dx}\,\delta(x) = \lim_{a \to +0} \frac{1}{a} \int_{-\infty}^{\infty} dx\, f(x)\big[\delta(x + a) - \delta(x)\big]$$

$$= \lim_{a \to +0} \frac{1}{a}\big[f(-a) - f(0)\big] = -f'(0). \quad (5.8b)$$

Alternatively, one can integrate by parts, with the same result:

$$\int_{-\infty}^{\infty} dx\, f(x)\,\frac{d}{dx}\,\delta(x) = f(x)\delta(x)\Big|_{-\infty}^{\infty} - \int_{-\infty}^{\infty} dx\,\delta(x)\,\frac{d}{dx} f(x) = -f'(0). \quad (5.8c)$$

Yet another alternative is to write nth derivative of $\delta(x)$ as

$$\frac{d^n}{dx^n}\,\delta(x) = \lim_{y \to 0} (-1)^n \frac{\partial^n}{\partial y^n}\,\delta(x - y), \quad (5.9a)$$

which gives

$$\int_{-\infty}^{\infty} dx\, f(x)\,\frac{d^n}{dx^n}\,\delta(x) = \lim_{y \to 0} (-1)^n \frac{\partial^n}{\partial y^n} \int_{-\infty}^{\infty} dx\, f(x)\delta(x - y) = (-1)^n f^{(n)}(0), \quad (5.9b)$$

where $f^{(n)}(x)$ is nth derivative of $f(x)$. This generalizes Eqs. (5.8b) and (5.8c) to arbitrary integer $n > 0$. Replacing $f(x)$ in Eq. (5.9b) with $x^n f(x)$, we obtain the relation

$$x^n\,\frac{d^n}{dx^n}\,\delta(x) = (-1)^n\, n!\,\delta(x). \quad (5.10)$$

[Note that it would be wrong to divide Eq. (5.10) by x^n and write $d^n\delta(x)/dx^n \propto x^{-n}\delta(x)$.]

As mentioned above, the delta-function can often be viewed as an infinitely sharp peak of a unit area. This intuition is formalized as follows: if $\eta(t)$ is such real function that [69,70]

$$\int dt\,\eta(t) = 1, \quad |t\eta(t)| \xrightarrow{|t|\to\infty} 0, \tag{5.11a}$$

then $\delta(x)$ can be written as a limit

$$\delta(x) = \lim_{\alpha\to+0} \delta_\alpha(x), \quad \delta_\alpha(x) = \alpha^{-1}\eta(x/\alpha). \tag{5.11b}$$

Indeed, the first condition in (5.11a) implies that $\int dx\,\delta_\alpha(x) = 1$ for any $\alpha > 0$, including the limit $\alpha \to +0$, whereas the second condition shows that $\lim_{\alpha\to+0}\delta_\alpha(x) = 0$ for any $x \neq 0$.[71]

The second condition in (5.11a) can be relaxed. Consider the function

$$\eta(t) = \frac{\sin t}{\pi t}. \tag{5.12a}$$

For this function, the integral $\int dt\,\eta(t)$ is clearly improper. To give this integral a meaning, we ***regularize*** it (i.e., force it to converge absolutely) as [72,73]

$$\int dt\,\eta(t) = \lim_{\beta\to+0} \int dt\,\eta(t)\,e^{-\beta|t|} = 1. \tag{5.12b}$$

With such regularization, Eqs. (5.11b) and (5.12a) again define a delta-function:

$$\lim_{\alpha\to+0} \int_a^b dx\,\delta_\alpha(x) = \lim_{\alpha\to+0} \int_{a/\alpha}^{b/\alpha} dt\,\eta(t) = \begin{cases} 1, & a < 0 < b, \\ 0, & \text{otherwise} \end{cases}$$

even though $\alpha \to +0$ limit of $\delta_\alpha(x) = (\pi x)^{-1}\sin(x/\alpha)$ at fixed $x \neq 0$ does not exist. This observation provides a counter-example to the naive view that $\delta(x)$ has a definite value for a given x.[68]

One particular representation of the delta-function is especially useful,

$$\delta(x) = \frac{1}{2\pi}\int dk\,e^{ikx}. \tag{5.13a}$$

---

[69] Among the functions $\eta(t)$ satisfying conditions (5.11a) the most popular are [71]

$$\frac{1}{2}\theta\big(1-|t|\big) \text{ (rectangular "box")}, \quad \frac{1}{\pi(1+t^2)} \text{ (Lorentzian)}, \quad \frac{1}{\sqrt{\pi}}\,e^{-t^2} \text{ (Gaussian)}.$$

[70] Hereinafter we will often omit the integration limits in $\int_{-\infty}^{+\infty}$.

[71] One should keep in mind that writing the integral $\int dx f(x)\delta(x)$ as $\lim_{\alpha\to+0}\int dx f(x)\delta_\alpha(x)$ [see Eq. (5.11b)] is legitimate only if the integral $\int dx f(x)\delta_\alpha(x)$ is absolutely convergent.

[72] There are other ways of regularizing improper integrals. For example, one can replace the exponential cutoff $\exp(-\beta|t|)$ in Eqs. (5.12b) and (5.13b) with the Gaussian one, $\exp(-\beta^2 t^2)$, with the same results.[71]

[73] There are many ways to evaluate the integral in Eq. (5.12b) with $\eta(t)$ given by Eq. (5.12a). Perhaps the simplest one is to write $t^{-1}\sin t = \frac{1}{2}\int_{-1}^{1} dz\,e^{izt}$ and change the order of integrations. (Such change is legitimate because in the presence of the exponential cutoff the integral converges absolutely.)

This integral is, of course, also improper. Similar to Eq. (5.12b), we regularize it as [72]

$$\int dk\, e^{ikx} = \lim_{\beta\to+0}\int dk\, e^{ikx-\beta|k|}. \tag{5.13b}$$

Carrying out the integration, we obtain

$$\int_{-\infty}^{\infty} dk\, e^{ikx-\beta|k|} = \int_{-\infty}^{0} dk\, e^{(ix+\beta)k} + \int_{0}^{\infty} dk\, e^{(ix-\beta)k} = \frac{2\beta}{\beta^2+x^2} \xrightarrow{\beta\to+0} 2\pi\delta(x),$$

in agreement with Eq. (5.13a).

With the help of Eq. (5.13a), it is easy to derive the standard ***Fourier transform*** relations

$$f(k) = \int dx\, e^{ikx} f(x), \quad f(x) = \frac{1}{2\pi}\int dk\, e^{-ikx} f(k). \tag{5.14}$$

Indeed, assuming that $f(x)$ is sufficiently well-behaved to allow the change of the order of integrations, we find

$$\int dk\, e^{-ikx} f(k) = \int dk\, e^{-ikx} \underbrace{\int dx' e^{ikx'} f(x')}_{f(k)} = \int dx' f(x') \underbrace{\int dk\, e^{-ik(x-x')}}_{2\pi\delta(x-x')} = 2\pi f(x),$$

which is the second equation in (5.14). Note that since $\delta(k) = \int dx\, e^{ikx}\delta(x) = 1$, Eq. (5.13a) itself can be interpreted as the inverse Fourier transform.

The Fourier transform of a product of two functions is evaluated in the same manner,

$$\begin{aligned}\int dx\, e^{ikx} f(x) g(x) &= \int dx\, e^{ikx} \underbrace{\int \frac{dk'}{2\pi} e^{-ik'x} f(k')}_{f(x)} \underbrace{\int \frac{dk''}{2\pi} e^{-ik''x} g(k'')}_{g(x)} \\ &= \int \frac{dk'}{2\pi}\int \frac{dk''}{2\pi} f(k') g(k'') \underbrace{\int dx\, e^{i(k-k'-k'')x}}_{2\pi\delta(k-k'-k'')} = \int \frac{dq}{2\pi} f(k-q) g(q).\end{aligned} \tag{5.15}$$

Setting here $k = 0$ and taking into account that the Fourier transform of $f^*(x)$ is $f^*(-k)$ [see the first equation in (5.14)], we obtain the ***Parseval identity***

$$\int dx \left|f(x)\right|^2 = \int \frac{dk}{2\pi} \left|f(k)\right|^2. \tag{5.16}$$

This result shows that the Fourier transform $f(k)$ of a square-integrable function $f(x)$ (i.e., such that the integral $\int dx |f(x)|^2$ is converging) is also square-integrable.

A very useful result concerning Fourier transforms is the ***Riemann-Lebesgue lemma***. The lemma applies to absolutely integrable functions (generalized functions such as the delta-function and its derivatives are excluded) and asserts that the Fourier transform of any such function is a continuous function of k satisfying

$$\lim_{|k|\to\infty} f(k) = 0. \tag{5.17}$$

The lemma has nothing to say about how fast $f(k)$ approaches zero. It turns out that the behavior of $f(k)$ at large $|k|$ is determined by discontinuities in $f(x)$ and its derivatives. Suppose, for example, that $f(x)$ has a discontinuity at $x = x_0$, i.e.,

$$f(x_0 + 0) - f(x_0 - 0) = g_0 \neq 0. \tag{5.18a}$$

Integration by parts gives

$$kf(k) = -i\int dx\, f(x)\frac{d}{dx}e^{ikx} = -i\int dx \left\{\frac{d}{dx}\left[f(x)e^{ikx}\right] - e^{ikx}\frac{d}{dx}f(x)\right\} = i\int dx\, e^{ikx}\frac{d}{dx}f(x).$$

Substituting here $df/dx = g_0\delta(x - x_0) + g_1(x)$ [recall Eq. (5.2)], where $g_1(x)$ is a non-singular function subject to the Riemann-Lebesgue lemma (5.17), and taking the limit $|k| \to \infty$, we obtain $\lim_{|k|\to\infty}\big|kf(k)\big| = |g_0| > 0$, so that

$$f(k)\big|_{\text{large }|k|} \propto |k|^{-1}. \tag{5.18b}$$

It is not difficult to show that if nth derivative $d^n f/dx^n$ has a discontinuity but all lower-order derivatives are continuous, then $f(k)$ exhibits a power-law decay $f(k)\big|_{\text{large }|k|} \propto |k|^{-(n+1)}$. Conversely, if *all* derivatives of $f(x)$ are continuous (i.e., if $f(x)$ is an analytic function) then its Fourier transform $f(k)$ decays at large $|k|$ faster than any power (e.g., exponentially).

5.1.2 Operators with continuous spectra

As mentioned above,[1] not all observable quantities are quantized. In our general formalism [see Sec. 2.3.1], such continuously-varying observables should be represented by Hermitian operators with a ***continuous spectrum***. A rigorous treatment of such operators is rather involved. Instead, we proceed heuristically, by exploiting the analogy with operators on finite-dimensional Hilbert spaces [see Sec. 2.2].

We assume from the outset the existence of such basis set $\{|a\rangle\}$ that every vector in it obeys the equation

$$\hat{A}|a\rangle = a|a\rangle \tag{5.19}$$

with *real* eigenvalue a [cf. Eq. (2.38)]. If the eigenvalues vary continuously, the summation over a in the completeness relation [see Eqs. (2.24) and (2.46b)] should be replaced with the integration. That is, instead of $\hat{\mathbb{1}} = \sum_a |a\rangle\langle a|$, we write

$$\hat{\mathbb{1}} = \int da\, |a\rangle\langle a|. \tag{5.20}$$

Acting with the identity operator in this form on the basis vector $|a\rangle$, we obtain

$$\hat{\mathbb{1}}|a\rangle = \int da'\, |a'\rangle\langle a'|a\rangle. \tag{5.21a}$$

On the other hand, it is natural to require that

$$\hat{\mathbb{1}}|a\rangle = |a\rangle. \tag{5.21b}$$

Eqs. (5.21) force us to replace the Kronecker delta in the discrete-spectrum orthonormality condition $\langle a|a'\rangle = \delta_{a,a'}$ [see Eqs. (1.42), (2.41), and (2.46a)] with the delta-function,

$$\langle a|a'\rangle = \delta(a-a'). \tag{5.22}$$

Vectors $|a\rangle$ defined (up to phase factors) by Eqs. (5.19) and (5.22) are referred to as the ***generalized eigenvectors*** of operator $\hat{A}$. Eq. (5.22) shows that similar to the ordinary eigenvectors [see Eq. (2.41)], generalized eigenvectors corresponding to different eigenvalues are orthogonal. However, Eq. (5.22) also shows that these eigenvectors have an infinite norm, and thus do not belong to the Hilbert space $\mathcal{H}$ [see Sec. 1.2.4]. Instead, these vectors live in a larger space for which $\mathcal{H}$ is merely a subspace.

With the help of the completeness relation Eq. (5.20), any vector in $\mathcal{H}$ can be expanded in the basis formed by generalized eigenvectors of $\hat{A}$,

$$|\psi\rangle = \hat{\mathbb{1}}|\psi\rangle = \int da\, |a\rangle\langle a|\psi\rangle = \int da\, |a\rangle \psi(a), \quad \psi(a) = \langle a|\psi\rangle \tag{5.23a}$$

[cf. Eq. (2.25a)]. The ***wave function*** $\psi(a)$ introduced here can be viewed as is the component of $|\psi\rangle$ in this basis. An inner product of two vectors can be written in terms of the corresponding wave functions as

$$\langle\varphi|\psi\rangle = \langle\varphi|\hat{\mathbb{1}}|\psi\rangle = \int da\, \langle\varphi|a\rangle\langle a|\psi\rangle = \int da\, \varphi^*(a)\psi(a) \tag{5.23b}$$

[cf. Eq. (2.25b)]. Because all vectors in $\mathcal{H}$ have a finite norm [see Sec. 1.2.4], the wave functions must be ***square-integrable***:

$$\|\psi\|^2 = \langle\psi|\psi\rangle = \int da\, |\psi(a)|^2 < \infty \text{ for all } |\psi\rangle \text{ in } \mathcal{H}. \tag{5.23c}$$

Any wave function that has delta-function-like singularities would violate this requirement. (For example, $|\psi\rangle \propto |a_0\rangle$ corresponds to $\psi(a) \propto \delta(a-a_0)$, which is not square-integrable.) Accordingly, any vector in $\mathcal{H}$ is necessarily a superposition of a *continuum* of generalized eigenvectors.

Similar to Hermitian operators with discrete spectra [see Eq. (2.47a)], operator $\hat{A}$ is diagonal in the basis of its own eigenvectors, i.e., its spectral decomposition reads

$$\hat{A} = \hat{\mathbb{1}}\hat{A}\hat{\mathbb{1}} = \int da \int da'\, |a\rangle \underbrace{\langle a|\hat{A}|a'\rangle}_{a\,\delta(a-a')} \langle a'| = \int da\, |a\rangle a \langle a|. \tag{5.24a}$$

An operator-valued function of $\hat{A}$ can also be written as a spectral decomposition,

$$\hat{f}(\hat{A}) = \int da\, |a\rangle f(a) \langle a| \tag{5.24b}$$

[cf. Eq. (2.48)]. Substitution of $\hat{f}(\hat{A})$ in this form into Eq. (2.61) yields the expectation value of $f(A)$ in the state represented by the state vector $|\psi\rangle$,

$$\langle f(A)\rangle_\psi = \int da\, f(a)\rho_\psi(a), \tag{5.25}$$

where

$$\rho_\psi(a) = |\psi(a)|^2 \tag{5.26a}$$

is a real function of a that satisfies

$$\rho_\psi(a) \geq 0\,, \quad \int da\, \rho_\psi(a) = 1. \tag{5.26b}$$

Comparison of Eqs. (5.26) with Eqs. (1.2) and (1.4) suggests that this function should be interpreted as the ***probability density*** of observable A. That is,

$$\text{Prob}_\psi(a_1 < A < a_2) = \int_{a_0}^{a_2} da\, \rho_\psi(a) \geq 0 \tag{5.27}$$

is the probability that the outcome of measurement of A lies between a_1 and a_2. Eq. (5.26a) can therefore be viewed as the Born's rule (1.47) adapted for the continuous spectrum. The probability density can be written in the operator form [cf. Eq. (2.62a)] as

$$\rho_\psi(a) = \langle\psi|\,\hat{\delta}(\hat{A} - a\hat{\mathbb{1}})|\psi\rangle\,, \tag{5.28a}$$

where the delta-function of an operator should be understood as the spectral decomposition

$$\hat{\delta}(\hat{A} - a\hat{\mathbb{1}}) = \int da'\, |a'\rangle\, \delta(a' - a) \langle a'|\,. \tag{5.28b}$$

Eqs. (5.28) can be used to evaluate $\rho_\psi(a)$, although it is rarely the most efficient method.

It should be noted that the integration in Eq. (5.20) does not have to extend to the entire real axis. Indeed, it may well happen that the spectrum of $\hat{A}$ consists of two or more disjoint intervals (*bands*) separated by the empty regions (*gaps*) in which $\hat{A}$ has no eigenvalues. Moreover, the continuous spectrum quite often coexists with an ordinary, discrete one. The discrete eigenvalues usually (although not always) lie in the gaps. The corresponding eigenvectors belong to $\mathcal{H}$, form an orthonormal (albeit incomplete) set, and must be included in the completeness relation, so that Eq. (5.20) should be replaced with

$$\hat{\mathbb{1}} = \underbrace{\sum |a\rangle\langle a|}_{\text{discrete}} + \underbrace{\int da\, |a\rangle\langle a|}_{\text{continuous}} \equiv \sum\!\!\!\!\!\!\int |a\rangle\langle a|, \tag{5.29}$$

with the corresponding modifications in Eqs. (5.23)-(5.26). Furthermore, both the discrete and the continuous subsets of the spectrum may also be degenerate. In this case, the eigenvectors $|a\rangle$ should be equipped with additional indexes, and the right-hand side of Eq. (5.29) should include the summation over these indexes [cf. Eq. (2.46b)].

5.2 Position and momentum operators

5.2.1 Displacement transformation

Displacing a classical particle by a amounts to changing its position (coordinate) from x to $x_a = x + a$. The quantum counterpart of such classical displacement is the ***displacement transformation***

$$\hat{x} \xrightarrow{\textbf{displacement by } a} \hat{x}_a = \hat{x} + a\hat{\mathbb{1}}\,, \tag{5.30a}$$

where $\hat{x} = \hat{x}^\dagger$ is the ***position operator***. The displacement transformation is yet another example of a continuous symmetry transformation [see Sec. 2.5.3]. In this case, the symmetry in play is the translational invariance: because space is uniform, an experiment carried out inside an isolated laboratory will not be able reveal its position. As it is the case with all continuous symmetry transformations [recall Eq. (2.127a)], the displaced position operator $\hat{x}_a$ is related to $\hat{x}$ via the unitary transformation

$$\hat{x}_a = \hat{D}^\dagger(a)\,\hat{x}\,\hat{D}(a), \tag{5.30b}$$

where the ***displacement operator*** $\hat{D}(a)$ is a unitary operator that depends continuously on a.

It is obvious from Eq. (5.30a) that two consecutive displacements, by a and b, in any order, yield the same result as a single displacement by $a+b$. This observation implies that it is possible to choose the phases of the displacement operators so that these operators have the group property

$$\hat{D}(a)\hat{D}(b) = \hat{D}(a+b). \tag{5.31}$$

Accordingly [see Eqs. (2.121) and (2.125)], the displacement operator can be written as

$$\hat{D}(a) = e^{-ia\hat{p}/\hbar}. \tag{5.32}$$

Comparison with the canonical transformations in Classical Mechanics suggests identifying the generator of displacements $\hat{p} = \hat{p}^\dagger$ with the ***momentum operator***. As in Eqs. (3.14) and (4.2), the Planck constant $\hbar$ is needed to make the expression in the exponent in Eq. (5.32) dimensionless.[74]

Expanding $\hat{x}_a = e^{ia\hat{p}/\hbar}\,\hat{x}\,e^{-ia\hat{p}/\hbar}$ in $\hat{x}_a = \hat{x} + a\hat{\mathbb{1}}$ [see Eqs. (5.30) and (5.32)] in powers of a with the help of Eq. (2.14), we obtain

$$\hat{x} + [ia\hat{p}/\hbar, \hat{x}] + \frac{1}{2}\big[ia\hat{p}/\hbar, [ia\hat{p}/\hbar, \hat{x}]\big] + \ldots = \hat{x} + a\hat{\mathbb{1}}.$$

This equation must be satisfied for all a, which is possible if and only if operators $\hat{x}$ and $\hat{p}$ obey the commutation relation

$$[\hat{x}, \hat{p}] = i\hbar\hat{\mathbb{1}}. \tag{5.33}$$

This relation can be also written as $[\hat{p}, \hat{x}] = -i\hbar\hat{\mathbb{1}}$, which suggests that various new operator relations can be obtained by simply interchanging x and p and changing the sign of $\hbar$ in the existing ones. Applying this recipe to $e^{ia\hat{p}/\hbar}\,\hat{x}\,e^{-ia\hat{p}/\hbar} = \hat{x} + a\hat{\mathbb{1}}$ [see Eqs. (5.30) and (5.32)] and replacing the parameter a with q that has units of momentum, we obtain the relation

$$e^{-iq\hat{x}/\hbar}\,\hat{p}\,e^{iq\hat{x}/\hbar} = \hat{p} + q\hat{\mathbb{1}}, \tag{5.34}$$

which can be also derived by expanding the left-hand side in powers of q and using Eq. (5.33). Eq. (5.34) shows that the position operator generates displacements of momentum, similar to what $\hat{p}$ does for $\hat{x}$.

[74] Although $c\hbar$ with any $c \neq 0$ would also serve this purpose, the choice $c = 1$ ensures that the cross product of the three-dimensional position and momentum operators $\hat{\mathbf{r}} \times \hat{\mathbf{p}}$ obeys the angular momentum commutation relation (3.28). (We do not discuss three-dimensional motion in these Notes.)

5.2.2 Hilbert space

We arrived at the fundamental commutation relation (5.33) by combining the physical requirement that the position operator should transform under displacements the same way as the observable it represents with the mathematical insight into continuous symmetry transformations. To complete the construction of the position and momentum operators, we need to specify the Hilbert space in which these operators act.

It is easy to see that the commutation relation (5.33) cannot be realized in a finite-dimensional Hilbert space. Indeed, if the dimension of $\mathcal{H}$ were finite, $0 < \dim \mathcal{H} < \infty$, then Eqs. (2.32b) and (5.33) would yield $\dim \mathcal{H} = \operatorname{tr} \hat{\mathbb{1}} = (i\hbar)^{-1} \operatorname{tr}[\hat{x}, \hat{p}] = 0$, contrary to the assumption. The only remaining possibility is that the Hilbert space is infinite-dimensional,

$$\dim \mathcal{H} = \infty. \tag{5.35}$$

Let $|\phi\rangle$ be a vector in $\mathcal{H}$ that satisfies either $\hat{x}|\phi\rangle = x_\phi|\phi\rangle$ or $\hat{p}|\phi\rangle = p_\phi|\phi\rangle$. For this vector Eq. (5.33) gives $\langle\phi|\phi\rangle = \langle\phi|\hat{\mathbb{1}}|\phi\rangle = (i\hbar)^{-1}\langle\phi|[\hat{x},\hat{p}]|\phi\rangle = 0$, hence $|\phi\rangle = |\text{null}\rangle$ [recall Eq. (1.40)]. Thus, the above equations have no solutions in $\mathcal{H}$ other than $|\text{null}\rangle$, i.e., neither $\hat{x}$ nor $\hat{p}$ have genuine eigenvectors.

This discussion suggests that both operators have continuous spectra and all their eigenvectors are generalized [see Sec. 5.1.2]. In particular, the ***position eigenvectors*** $|x\rangle$ are defined (up to phase factors) by the equations

$$\hat{x}|x\rangle = x|x\rangle, \quad \langle x|x'\rangle = \delta(x - x'). \tag{5.36}$$

with real eigenvalues x [cf. (5.19) and (5.22)]. Acting on both sides of the first equation here with the displacement operator $\hat{D}(a) = \hat{D}^\dagger(-a)$, inserting the identity operator $\hat{\mathbb{1}} = \hat{D}^\dagger(a)\hat{D}(a)$ between $\hat{x}$ and $|x\rangle$ in the left-hand side, and using Eqs. (5.30), we obtain

$$\hat{x}\,|x_a\rangle = (x + a)|x_a\rangle, \quad |x_a\rangle = \hat{D}(a)|x\rangle. \tag{5.37}$$

Thus, if x is an eigenvalue of $\hat{x}$ then $x + a$ (with arbitrary a) is also an eigenvalue of $\hat{x}$. In other words, the eigenvalues of the position operator vary continuously between $-\infty$ and ∞.

It can be shown rigorously that the position eigenvectors form a basis, i.e.,

$$\hat{\mathbb{1}} = \int dx\, |x\rangle\langle x| \tag{5.38}$$

[cf. Eq. (5.20)], or, equivalently, that the infinite-dimensional linear vector space formed by complex square-integrable wave functions $\psi(x) = \langle x|\psi\rangle$ with the addition and scalar multiplication operations defined in the usual way and with the inner product

$$\langle\varphi|\psi\rangle = \int dx\, \varphi^*(x)\psi(x) \tag{5.39}$$

[cf. Eq. (5.23b)] is indeed a Hilbert space.[17] This realization of our abstract Hilbert space is often referred to as the ***position space***.

Any operator-valued function of $\hat{x}$ can be written as a spectral decomposition

$$\hat{f}(\hat{x}) = \int dx\, |x\rangle f(x)\langle x| \tag{5.40a}$$

[cf. Eq. (5.24b)]. Taking into account the second equation in (5.36), we obtain

$$\langle x|\hat{f}(\hat{x})|\psi\rangle = \int dx' \langle x|x'\rangle f(x')\langle x'|\psi\rangle = f(x)\langle x|\psi\rangle. \tag{5.40b}$$

Accordingly, the action of $\hat{f}(\hat{x})$ on vector $|\psi\rangle$ in $\mathcal{H}$ amounts to the multiplication of the corresponding position-space wave function $\psi(x) = \langle x|\psi\rangle$ by $f(x)$. This can be written symbolically as

$$\hat{f}(\hat{x}) \xrightarrow[\text{space}]{\text{position}} f(x). \tag{5.40c}$$

These relations show, in particular, that the position operator is *unbounded*. Indeed, the fact that the wave function $\psi(x)$ is square-integrable does not guarantee that $\langle x|\hat{x}|\psi\rangle = x\psi(x)$ is.

It is obvious that any operator-valued function of $\hat{x}$ commutes with $\hat{x}$. Conversely, it turns out that any linear operator that commutes with $\hat{x}$ is necessarily a function of $\hat{x}$. In other words, the position operator $\hat{x}$ exhausts the complete set of commuting operators for the Hilbert space considered. Indeed, it is easy to see that if $[\hat{x},\hat{A}] = \hat{0}$ then the Fourier transform of $\langle x+y|\hat{A}|x\rangle$ with respect to y is independent of k:

$$\frac{\partial}{\partial k}\int dy\, e^{iky}\langle x+y|\hat{A}|x\rangle = i\int dy\, e^{iky} y\langle x+y|\hat{A}|x\rangle = i\int dy\, e^{iky}\langle x+y|[\hat{x},\hat{A}]|x\rangle = 0.$$

Denoting

$$\int dy\, e^{iky}\langle x+y|\hat{A}|x\rangle = \int dy\,\langle x+y|\hat{A}|x\rangle = A(x) \tag{5.41a}$$

and carrying out the inverse Fourier transform [recall Eq. (5.14)], we obtain the relation

$$\langle x+y|\hat{A}|x\rangle = \int \frac{dk}{2\pi}\, e^{-iky} A(x) = A(x)\,\delta(y),$$

which can be also written as

$$\langle x|\hat{A}|x'\rangle = A(x)\,\delta(x-x'). \tag{5.41b}$$

Thus, we have just shown that $[\hat{x},\hat{A}] = \hat{0}$ implies the existence of such function $A(x)$ that

$$\hat{A} = \int dx\,|x\rangle A(x)\langle x|$$

[cf. Eq. (5.40a)], or, equivalently, $\hat{A}|x\rangle = A(x)|x\rangle$. In other words, $\hat{A}$ is a function of $\hat{x}$.

Rather than repeating the same arguments for the momentum operator, we take an obvious shortcut and simply replace x with p in the above relations. The generalized ***momentum eigenvectors*** are defined (up to phase factors) by the equations

$$\hat{p}\,|p\rangle = p|p\rangle, \quad \langle p|p'\rangle = \delta(p-p'), \tag{5.42a}$$

where the eigenvalues p vary continuously between $-\infty$ and ∞ and form a complete set,

$$\hat{\mathbb{1}} = \int dp\,|p\rangle\langle p| \tag{5.42b}$$

[cf. Eqs. (5.36) and (5.38)]. The action of an operator-valued function

$$\hat{f}(\hat{p}) = \int dp\,|p\rangle f(p)\langle p| \tag{5.43a}$$

on vector $|\psi\rangle$ reduces to the multiplication of the momentum-space wave function $\psi(p) = \langle p|\psi\rangle$ by $f(p)$,

$$\hat{f}(\hat{p}) \xrightarrow[\text{space}]{\text{momentum}} f(p) \tag{5.43b}$$

[cf. Eqs. (5.40)]. Moreover, just as the position operator $\hat{x}$, the momentum operator exhausts the complete set of commuting operators: any operator that commutes with $\hat{p}$ is a function of $\hat{p}$ and vice versa. It is then easy to see that any operator that commutes with *both* $\hat{x}$ and $\hat{p}$ is proportional to the identity operator $\hat{\mathbb{1}}$.

5.2.3 Momentum operator in the position space and vice versa

Since the momentum operator $\hat{p}$ does not commute with $\hat{x}$ [see Eq. (1.13)], it cannot be written as a function of $\hat{x}$. This can be demonstrated explicitly.

Comparison of Eqs. (5.36) and (5.37) shows that vector $|x_a\rangle = \hat{D}(a)|x\rangle$ differs from vector $|x+a\rangle$ in an overall phase factor at most, i.e.,

$$\hat{D}(a)|x\rangle = |x+a\rangle \quad \text{(up to phase factors)}, \tag{5.44}$$

It is convenient to choose the phase[75] of the position eigenvector $|x\rangle$ so that

$$|x\rangle = \hat{D}(x)|0\rangle, \quad |0\rangle \equiv |x\rangle\big|_{x=0}. \tag{5.45a}$$

The advantage of this choice is that it eliminates the phase factors in Eq. (5.44):

$$\hat{D}(a)|x\rangle = \hat{D}(a)\hat{D}(x)|0\rangle = \hat{D}(x+a)|0\rangle = |x+a\rangle, \tag{5.45b}$$

where we took into account Eq. (5.31).

Eq. (5.45b) implies that $\langle\psi|\hat{D}(-a)|x\rangle = \langle\psi|x-a\rangle$ for any $|\psi\rangle$. Since $\hat{D}^\dagger(-a) = \hat{D}(a)$ [see Eq. (5.32)], complex conjugation of this expression yields

$$\langle x|\hat{D}(a)|\psi\rangle = \langle x-a|\psi\rangle. \tag{5.46}$$

Differentiating both sides of this equation with respect to a with the help of the obvious relations

$$i\hbar\frac{d}{da}\hat{D}(a) = \hat{p}\hat{D}(a), \quad \frac{\partial}{\partial a}\langle x-a|\psi\rangle = -\frac{\partial}{\partial x}\langle x-a|\psi\rangle$$

and sending a to zero, we obtain

$$\langle x|\hat{p}|\psi\rangle = -i\hbar\frac{d}{dx}\langle x|\psi\rangle. \tag{5.47a}$$

Thus, the momentum operator acts as a *differential operator* in the position space:

$$\hat{p} \xrightarrow[\text{space}]{\text{position}} -i\hbar\frac{d}{dx}. \tag{5.47b}$$

[75] Recall that Eqs. (5.36) and (5.42a) define the eigenvectors only up to phase factors.

It is easy to verify that Eqs. (5.40) and (5.47) are consistent with the commutation relation (5.33). Indeed, Eqs. (5.40b) and (5.47a) show that for any $|\psi\rangle$ in $\mathcal{H}$

$$\langle x|[\hat{x},\hat{p}]|\psi\rangle = \langle x|\hat{x}\hat{p}|\psi\rangle - \langle x|\hat{p}\hat{x}|\psi\rangle$$
$$= x\langle x|\hat{p}|\psi\rangle + i\hbar\frac{d}{dx}\langle x|\hat{x}|\psi\rangle = i\hbar\left[-x\frac{d}{dx} + \frac{d}{dx}x\right]\langle x|\psi\rangle = i\hbar\langle x|\psi\rangle. \quad (5.48a)$$

Using the completeness relation (5.38), we then obtain

$$[\hat{x},\hat{p}]|\psi\rangle = \hat{\mathbb{1}}[\hat{x},\hat{p}]|\psi\rangle = \int dx\,|x\rangle\langle x|[\hat{x},\hat{p}]|\psi\rangle = i\hbar\int dx\,|x\rangle\langle x|\psi\rangle = i\hbar|\psi\rangle. \quad (5.48b)$$

Because $|\psi\rangle$ is arbitrary, Eq. (5.48b) is equivalent to the operator relation (5.33).

Interestingly, Eqs. (5.46)-(5.48) are applicable to generalized eigenvectors as well. In particular, replacing $|\psi\rangle$ in Eq. (5.47a) with the position eigenvector $|x'\rangle$ and taking into account the second equation in (5.36), we obtain matrix elements of the momentum operator in the position basis,

$$\langle x|\hat{p}|x'\rangle = -i\hbar\frac{\partial}{\partial x}\delta(x-x'). \quad (5.49)$$

Notice that although $\langle x|\hat{p}|x'\rangle$ has a singularity at $x \to x'$, it is proportional to the derivative of the delta-function rather than to the delta-function itself. Were it otherwise, the momentum operator $\hat{p}$ would be a function of $\hat{x}$ [cf. Eq. (5.41b)].

It is not difficult to establish a relation between the position and momentum eigenvectors. It follows from Eqs. (5.32) and (5.45a) that

$$\langle p|x\rangle = \langle p|\hat{D}(x)|0\rangle = \langle 0|\hat{D}^\dagger(x)|p\rangle^* = \langle p|0\rangle e^{-ipx/\hbar} \quad (5.50)$$

with vector $|0\rangle$ defined in Eq. (5.45a). Using Eqs. (5.13a), (5.38), and (5.50) and taking into account that the inner product $\langle p|0\rangle$ is independent of x, we obtain

$$\langle p|p'\rangle = \langle p|\hat{\mathbb{1}}|p'\rangle = \int dx\,\langle p|x\rangle\langle x|p'\rangle = \langle p|0\rangle\langle 0|p'\rangle\int dx\,e^{-ix(p-p')/\hbar} = 2\pi\hbar|\langle p|0\rangle|^2\delta(p-p').$$

Comparison with $\langle p|p'\rangle = \delta(p-p')$ [see the second equation in (5.42a)] then gives

$$|\langle p|0\rangle|^2 = (2\pi\hbar)^{-1}. \quad (5.51a)$$

We now choose the phases[75] of the momentum eigenvectors $|p\rangle$ by requiring the inner products $\langle p|0\rangle$ to be real and positive at all p. With this choice,

$$\langle p|0\rangle = |\langle p|0\rangle| = (2\pi\hbar)^{-1/2}, \quad (5.51b)$$

and Eq. (5.50) yields the relation we seek,[76]

$$\langle x|p\rangle = \langle p|x\rangle^* = \frac{e^{ipx/\hbar}}{\sqrt{2\pi\hbar}}. \quad (5.52)$$

[76] Notice that the simultaneous interchange $x \rightleftarrows p$ and $\hbar \rightleftarrows -\hbar$ in $\langle x|p\rangle$ yields $\langle p|x\rangle$. This recipe was guessed in the end of Sec. 5.2.1. It is clear, however, that its applicability to the relations involving eigenvectors hinges on the choices of the phase factors made in Eqs. (5.45a) and (5.51b) and embedded in Eq. (5.52).

This relation and Eqs. (5.38) and (5.42b) imply that when the phases of the position and momentum eigenvectors are chosen according to Eqs. (5.45a) and (5.51b), the position- and the momentum-space wave functions are Fourier transforms of one another:

$$\psi(p) = \langle p|\psi\rangle = \langle p|\hat{\mathbb{1}}|\psi\rangle = \int dx\, \langle p|x\rangle\langle x|\psi\rangle = \int \frac{dx}{\sqrt{2\pi\hbar}}\, e^{-ipx/\hbar}\psi(x), \tag{5.53a}$$

$$\psi(x) = \langle x|\psi\rangle = \langle x|\hat{\mathbb{1}}|\psi\rangle = \int dp\, \langle x|p\rangle\langle p|\psi\rangle = \int \frac{dp}{\sqrt{2\pi\hbar}}\, e^{ipx/\hbar}\psi(p). \tag{5.53b}$$

Because of the similarity between $\hat{x}$ and p, it is natural to expect that the position operator turns to a differential operator in the momentum space. Indeed, taking into account Eqs. (5.40a), (5.52), and Eq. (5.53a), we obtain the analogue of (5.47a),

$$\begin{aligned}\langle p|\hat{x}|\psi\rangle = \int dx\, \langle p|x\rangle x\, \langle x|\psi\rangle &= \int \frac{dx}{\sqrt{2\pi\hbar}}\, e^{-ipx/\hbar} x\, \langle x|\psi\rangle \\ &= i\hbar\frac{d}{dp}\int \frac{dx}{\sqrt{2\pi\hbar}}\, e^{-ipx/\hbar}\langle x|\psi\rangle = i\hbar\frac{d}{dp}\langle p|\psi\rangle,\end{aligned} \tag{5.54a}$$

which can be written symbolically as

$$\hat{x} \xrightarrow[\text{space}]{\text{momentum}} i\hbar\frac{d}{dp}. \tag{5.54b}$$

Eqs. (5.54) can be also obtained by interchanging[76] $x \rightleftarrows p$ and $\hbar \rightleftarrows -\hbar$ in Eqs. (5.47).

5.2.4 Evaluation of expectation values

Because $|\psi(x)|^2$ and $|\psi(p)|^2$ represent the probability densities in the position and momentum spaces, expectation values of functions of x or p are given by

$$\langle f(x)\rangle_\psi = \int dx\, f(x)|\psi(x)|^2, \quad \langle f(p)\rangle_\psi = \int dx\, f(p)|\psi(p)|^2 \tag{5.55}$$

[see Eqs. (5.25)-(5.26)]. It is less obvious what is the expectation value of, say, p^n in terms of the position-space wave function $\psi(x)$. With the help of Eqs. (5.38) and (5.47), this expectation value can be written as

$$\langle p^n\rangle_\psi = \langle\psi|\hat{p}^n|\psi\rangle = \langle\psi|\hat{\mathbb{1}}\hat{p}^n|\psi\rangle = \int dx\, \langle\psi|x\rangle\langle x|\hat{p}^n|\psi\rangle = (-i\hbar)^n\int dx\, \psi^*(x)\frac{d^n}{dx^n}\psi(x). \tag{5.56}$$

We substitute here $\psi(x)$ in the form

$$\psi(x) = \varphi(x)\, e^{i\vartheta(x)}, \quad \int dx\, \varphi^2(x) = 1, \tag{5.57a}$$

where $\varphi(x)$ and $\vartheta(x)$ are continuous real functions. Integrating by parts, we get for $n = 1,2$

$$\langle p\rangle_\psi = \hbar\int dx\, \varphi^2(x)\, \frac{d\vartheta}{dx}, \tag{5.57b}$$

$$\langle p^2\rangle_\psi = \hbar^2\int dx \left|\frac{d\psi}{dx}\right|^2 = \hbar^2\int dx \left\{\left(\frac{d\varphi}{dx}\right)^2 + \varphi^2(x)\left(\frac{d\vartheta}{dx}\right)^2\right\}. \tag{5.57c}$$

Eq. (5.57b) shows that $\langle p\rangle_\psi = 0$ if the phase of the wave function ϑ [see Eq. (5.57a)] is independent of x (e.g., if the wave function is real). Because $d\varphi/dx$ cannot vanish everywhere [this would be incompatible with the second equation in (5.57a)], Eq. (5.57c) yields $\langle p^2\rangle_\psi \geq \hbar^2 \int dx (d\varphi/dx)^2 > 0$, in agreement with the observation [see Sec. 5.1.2] that any state vector is necessarily a superposition of a *continuum* of generalized eigenvectors and thus must have a finite momentum uncertainty $\Delta_\psi p > 0$.

Similar to Eq. (5.56), Eqs. (5.42b) and (5.54) show that

$$\langle x^n\rangle_\psi = \langle\psi|\hat{x}^n|\psi\rangle = (i\hbar)^n \int dp\, \psi^*(p) \frac{d^n}{dp^n}\psi(p). \tag{5.58}$$

Substituting here

$$\psi(p) = \varphi(p)\, e^{-i\vartheta(p)}, \qquad \int dp\, \varphi^2(p) = 1 \tag{5.59a}$$

[notice that the phase ϑ here has the sign opposite to that in Eq. (5.57a)], we find

$$\langle x\rangle_\psi = \hbar \int dp\, \varphi^2(p) \frac{d\vartheta}{dp}, \tag{5.59b}$$

$$\langle x^2\rangle_\psi = \hbar^2 \int dp \left|\frac{d\psi}{dp}\right|^2 = \hbar^2 \int dp \left\{\left(\frac{d\varphi}{dp}\right)^2 + \varphi^2(p)\left(\frac{d\vartheta}{dp}\right)^2\right\}. \tag{5.59c}$$

These expressions can be also obtained by interchanging $x \rightleftarrows p$ in Eqs. (5.57b) and (5.57c).

5.2.5 Minimal uncertainty state

Because the state vector cannot be an eigenvector of $\hat{x}$ or $\hat{p}$, any quantum state is characterized by finite position and momentum uncertainties Δx and Δp. Eqs. (2.98) and (5.33) show that these uncertainties obey the inequality

$$\Delta x \Delta p \geq \hbar/2. \tag{5.60}$$

As discussed in Sec. 2.4.3 [see Eqs. (2.96a) and (2.99)], the product of uncertainties $\Delta x \Delta p$ attains the minimal possible value

$$\Delta x \Delta p = \hbar/2 \tag{5.61a}$$

when vectors $(\hat{x} - \langle x\rangle\hat{\mathbb{1}})|\psi\rangle$ and $(\hat{p} - \langle p\rangle\hat{\mathbb{1}})|\psi\rangle$ are proportional to one another with an imaginary coefficient. This observation yields the equation

$$\frac{i}{\hbar}(\hat{p} - \langle p\rangle\hat{\mathbb{1}})|\psi\rangle = \pm\frac{1}{2a^2}(\hat{x} - \langle x\rangle\hat{\mathbb{1}})|\psi\rangle, \tag{5.61b}$$

where the sign in the right-hand side is to be determined and a is a positive real parameter that has units of length. This parameter has a clear physical meaning. Indeed, it follows from Eqs. (5.61) and the relations $(\Delta x)^2 = \langle\psi|(\hat{x} - \langle x\rangle\hat{\mathbb{1}})^2|\psi\rangle$ and $(\Delta p)^2 = \langle\psi|(\hat{p} - \langle p\rangle\hat{\mathbb{1}})^2|\psi\rangle$ [recall Eq. (2.94b)] that $(\Delta x/a)^2 = 2\Delta x \Delta p/\hbar = 1$, i.e., a is nothing but the position uncertainty,

$$a = \Delta x. \tag{5.61c}$$

It remains to show that the solution of Eq. (5.61b) does in fact exist and to find this solution. To this end, we multiply[12] both sides of Eq. (5.61b) by $\langle x|$ and take into account

Eqs. (5.40) and (5.47). Replacing a with Δx, we obtain the first-order differential equation for the position-space wave function $\psi(x) = \langle x|\psi\rangle$,

$$\frac{d}{dx}\,\psi(x) = \pm\,\frac{x-\langle x\rangle}{2(\Delta x)^2}\,\psi(x) + i\,\frac{\langle p\rangle}{\hbar}\,\psi(x), \tag{5.62a}$$

whose general solution is

$$\ln\psi(x) = \pm\,\frac{(x-\langle x\rangle)^2}{(2\Delta x)^2} + \frac{i\langle p\rangle x}{\hbar} + \text{const}. \tag{5.62b}$$

The normalized solution is obtained by replacing "$\pm$" in Eqs. (5.61a) and (5.62) with "$-$" and reads

$$\psi(x) = \frac{1}{(2\pi)^{1/4}\,(\Delta x)^{1/2}}\,\exp\left[-\left(\frac{x-\langle x\rangle}{2\Delta x}\right)^2 + \frac{i\langle p\rangle x}{\hbar}\right]. \tag{5.63a}$$

This function is a unique (up to a phase factor independent of x) minimal-uncertainty position-space wave function for given $\langle x\rangle$, $\langle p\rangle$, and Δx. Because the Fourier transform of a Gaussian is also a Gaussian, the corresponding momentum-space wave function is given by

$$\psi(p) = \frac{1}{(2\pi)^{1/4}\,(\Delta p)^{1/2}}\,\exp\left[-\left(\frac{p-\langle p\rangle}{2\Delta p}\right)^2 - \frac{i\langle x\rangle p}{\hbar}\right], \tag{5.63b}$$

and is unique for given $\langle x\rangle$, $\langle p\rangle$, and Δp apart from a phase factor independent of p.

In this Section we demonstrated that minimal uncertainty states satisfying Eq. (5.61a) are described by Gaussian wave functions in both the position and the momentum spaces. Conversely, all Gaussian wave functions such as Eqs. (5.63) represent minimal uncertainty states.

5.2.6 Time evolution of a free particle

Classically, the energy of a free particle is the kinetic energy $E = p^2/2m$. Replacing E and p with the corresponding Hermitian operators, we obtain the Hamiltonian

$$\hat{H} = \frac{\hat{p}^2}{2m}, \tag{5.64}$$

where the particle's mass m is not an operator, but a coefficient whose purpose is to take care of the units. (Of course, one can add to this Hamiltonian an arbitrary constant without affecting the results.)

Because $\hat{H}$ commutes with $\hat{p}$, these operators share the eigenvectors [recall Eq. (2.85)]. Since all eigenvectors of $\hat{p}$ are generalized [see Sec. 5.2.2], they do not represent quantum states, only their superpositions do. Thus, free particles cannot be in a state with a certain energy, just like they cannot have a certain momentum. The commutativity of $\hat{H}$ and $\hat{p}$ also implies that the momentum is an integral of motion [see Sec. 4.1.4]. That is, the dressed momentum operator in the Heisenberg picture [see Eqs. (4.13)] coincides with the bare one, $\hat{p}_t = \hat{p}$, so that expectation values of various functions of p are independent of time. Equivalently, one may notice that the action of the evolution operator $\hat{T}(t,0) = e^{-it\hat{H}/\hbar}$

amounts to merely multiplying the momentum-space wave function in the Schrödinger picture $\psi(p,0) = \langle p|\psi(0)\rangle$ by a time-dependent phase factor,

$$\psi(p,t) = \langle p|\psi(t)\rangle = \langle p|\hat{T}(t,0)|\psi(0)\rangle = e^{-ip^2t/2m\hbar}\psi(p,0), \tag{5.65}$$

hence the momentum-space probability density is constant, $|\psi(p,t)|^2 = |\psi(p,0)|^2$.

This does not mean, of course, that the dressed position operator $\hat{x}_t$ or the position-space probability density $|\psi^2(x,t)|$ are also constant. In fact, they are not. Indeed, expanding

$$\hat{x}_t = \hat{T}^\dagger(t,0)\,\hat{x}\,\hat{T}(t,0) = e^{it\hat{H}/\hbar}\,\hat{x}\,e^{-it\hat{H}/\hbar} \tag{5.66a}$$

[see Eq. (4.13b)] in powers of t with the help of Eq. (2.14) and taking into account that

$$[\hat{p}^2,\hat{x}] = -\hat{p}[\hat{x},\hat{p}] - [\hat{x},\hat{p}]\hat{p} = -2i\hbar\hat{p}$$

[see Eqs. (2.13a) and (5.33)], we obtain

$$\hat{x}_t = \hat{x} + \hat{p}t/m. \tag{5.66b}$$

Because the second term in the right-hand side here grows linearly with t, it eventually becomes dominant, so that

$$\langle f(x)\rangle_t = \langle\psi(0)|\hat{f}(\hat{x}_t)|\psi(0)\rangle \xrightarrow{\text{large } t} \langle\psi(0)|\hat{f}(\hat{p}t/m)|\psi(0)\rangle = \int dp\, f(pt/m)|\psi(p,0)|^2 \tag{5.67}$$

for any function of x.

Changing the integration variable in Eq. (5.67) to $x = pt/m$ and comparing the resulting expression with $\langle f(x)\rangle_t = \int dx\, f(x)|\psi(x,t)|^2$ [cf. the first equation in (5.55)], we obtain the relation between the position-space probability density $|\psi(x,t)|^2$ and the momentum-space probability density $|\psi(p,0)|^2$,

$$\big|\psi(x,t)\big|^2 \xrightarrow{\text{large } t} \frac{m}{t}\,\big|\psi(p,0)\big|^2_{p=mx/t}. \tag{5.68}$$

Even if at $t = 0$ the probability density is localized in a small region, Eq. (5.68) shows that $\lim_{t\to\infty}|\psi(x,t)|^2 = 0$ at any fixed x.[77] In other words, at $t\to\infty$ the probability density spreads over the entire space. Such ***probability spreading*** is a generic feature of the free particle evolution.[78]

To better understand this phenomenon, we consider the expectation value of x and its uncertainty Δx. Eq. (5.67) shows that $\langle x^n\rangle_{\text{large } t} = \langle p^n\rangle(t/m)^n$, so that

$$\langle x\rangle_{\text{large } t} = \frac{\langle p\rangle t}{m}, \quad \Delta x_{\text{large } t} = \frac{\Delta p}{m}\,t. \tag{5.69}$$

[77] The result $\lim_{t\to\infty}\psi(x,t) = 0$ can be also obtained by writing $\psi(x,t)$ as the Fourier transform of $\psi(p,t) = e^{-ip^2t/2m\hbar}\psi(p,0)$ [see Eq. (5.53b) and (5.65)] and applying the Riemann-Lebesgue lemma (5.17).

[78] It was Heisenberg's discovery of the spreading of $|\psi(x,t)|^2$ that prompted Max Born to propose the probabilistic interpretation of the wave function and to formulate the eponymous rule.

The first equation here tells us that the "center-of-mass" of the position-space probability density moves with the ***group velocity*** $\langle p\rangle/m$, whereas the second equation shows that its width grows linearly with time. Notice that the ratio $\Delta x/\langle x\rangle$ tends to the time-independent limit $\Delta p/\langle p\rangle$. Thus, if $\Delta p \ll |\langle p\rangle|$, then $\Delta x \ll |\langle x\rangle|$ even at late times. In other words, in this case the probability density broadens with time, but not as fast as it moves.

Interestingly, the Planck constant $\hbar$ does not appear in Eqs. (5.68) and (5.69), which suggests that these results admit a classical interpretation. Indeed, it is obvious that these equations apply also to ensembles of classical particles with randomly distributed momenta.

6 Quantum particle in external potential

6.1 Generalities

6.1.1 The model

In this Chapter, we consider a paradigmatic model of quantum particle of mass m placed in a classical environment described by the time-independent potential $V(x)$,

$$\hat{H} = \frac{\hat{p}^2}{2m} + \hat{V}(\hat{x}), \quad \hat{V}(\hat{x}) = \int dx\, |x\rangle V(x) \langle x|. \tag{6.1}$$

Our main focus is the eigenvalue problem for this operator,

$$\hat{H}|\psi_E\rangle = E|\psi_E\rangle. \tag{6.2a}$$

We are interested in the solutions of Eq. (6.2a) that appear in the completeness relation

$$\hat{\mathbb{1}} = \sum\!\!\!\!\!\!\int |\psi_E\rangle\langle\psi_E| \tag{6.2b}$$

[recall Eq. (5.29)]. (The symbol $\sum\!\!\!\!\!\!\int$ in this expansion stands for the sum/integral over the discrete/continuous subsets of the spectrum and the sum over all possible degeneracies.)

We will work in the position space. Multiplying[12] both sides of Eq. (6.2a) by $\langle x|$ and taking into account Eqs. (5.40) and (5.47), we obtain the differential equation

$$-\frac{\hbar^2}{2m}\frac{d^2}{dx^2}\psi(x) + V(x)\psi(x) = E\psi(x), \quad \psi(x) = \langle x|\psi_E\rangle, \tag{6.3a}$$

often referred to (misleadingly!) as the *time-independent* or *stationary* Schrödinger equation. It is convenient to multiply both sides of this equation by $2m/\hbar^2$ and write it as

$$-\psi'' + v\psi = \varepsilon\psi, \tag{6.3b}$$

where both $v(x) = (2m/\hbar^2)V(x)$ and $\varepsilon = (2m/\hbar^2)E$ have units of $(\text{length})^{-2}$.

6.1.2 Structure of solutions

To be definite, we assume for the time being that the potential $v(x)$ in Eq. (6.3b) is a piece-wise continuous function that remains finite at all finite x, but may diverge at infinity [e.g., $v(x) \propto x^2$ is allowed, but delta-functions and their derivatives are not].[79] It is easy to see that these requirements ensure that both $\psi(x)$ and its derivative $\psi'(x)$ are continuous functions,[80] which in turn implies that $\psi(x)$ and $\psi'(x)$ do not diverge at finite x.

Because Eq. (6.3b) is the second-order ordinary differential equation, it has two linearly independent solutions for any ε (see below). However, not all such solutions are created equal. Indeed, the only solutions of interest are the ones that appear in the expansion

$$\delta(x - x') = \sum\!\!\!\!\!\!\int \psi_\varepsilon(x)\psi^*_\varepsilon(x'). \tag{6.4}$$

[79] Of course, not all potentials of interest are of this type. However, potentials that violate the conditions listed require special care, see, e.g., Sec. 6.2.1 and 6.2.3 below.

[80] Note that the second and higher-order derivatives of $\psi(x)$ do not have to be continuous. For instance, it is obvious from Eq. (6.3b) that a discontinuity in $v(x)$ translates to a discontinuity in $\psi''(x)$.

This expansion is nothing but the completeness relation written in the position space: it can be obtained by multiplying both sides of Eq. (6.2b) by $\langle x|$ and $|x'\rangle$ as in $\langle x|\hat{\mathbb{1}}|x'\rangle$.

To get oriented, we revisit the Hamiltonian of a free particle [see Sec. 5.2.6]. In this case, the potential $v(x)$ vanishes identically and Eq. (6.3b) reduces to

$$-\psi'' = \varepsilon\psi. \tag{6.5a}$$

For any $\varepsilon > 0$ this equation has two linearly independent solutions

$$\psi_{k\pm}(x) = \langle x|\psi_{k\pm}\rangle = \frac{e^{\pm ikx}}{\sqrt{2\pi}}, \quad k = \sqrt{\varepsilon} > 0\,, \tag{6.5b}$$

in which one recognizes the momentum eigenvectors in the position basis (5.52). It is easy to show with the help of Eq. (5.13a) that these solutions satisfy

$$\langle\psi_{k-}|\psi_{k'+}\rangle = \int dx\, \psi^*_{k-}(x)\psi_{k'+}(x) = \int \frac{dx}{2\pi}\, e^{i(k+k')x} = \delta(k+k') = 0$$

(note that both k and k' are positive) and exhaust the completeness relation (6.4),

$$\begin{aligned}\int_0^\infty dk\, \big[\psi_{k+}(x)\psi^*_{k+}(x') + \psi_{k-}(x)\psi^*_{k-}(x')\big] \\ = \int_0^\infty \frac{dk}{2\pi}\, \big[e^{ik(x-x')} + e^{-ik(x-x')}\big] = \int \frac{dk}{2\pi}\, e^{ik(x-x')} = \delta(x-x').\end{aligned} \tag{6.5c}$$

However, Eq. (6.5a) also has two linearly independent solutions for any $\varepsilon < 0$,

$$\psi(x) \propto e^{\pm\kappa x}, \quad \kappa = \sqrt{-\varepsilon} > 0\,. \tag{6.6}$$

These solutions do not enter the completeness relation (6.5c), and, therefore, should be discarded.[81] This can be accomplished by requiring $\psi(x)$ to remain finite at infinity, i.e.,

$$\lim_{|x|\to\infty} |\psi(x)| < \infty. \tag{6.7}$$

It can be shown rigorously that solutions of Eq. (6.3b) (with $v \not\equiv 0$) subject to this condition indeed form a complete set. Such solutions (supplied with appropriate normalization coefficients) will be referred to as the ***eigenfunctions*** of our Hamiltonian. One way to verify that all eigenfunctions have been accounted for is to construct the expansion of the delta-function [cf. Eq. (6.5c)]; this is often quite a daunting task.

Many properties of Eq. (6.3b) can be understood with the help of the ***Wronskian theorem***. In its simplest form, the theorem asserts that the ***Wronskian***

$$W[\psi_1,\psi_2] \equiv \det\begin{pmatrix}\psi_1 & \psi_2\\ \psi_1' & \psi_2'\end{pmatrix} = \psi_1\psi_2' - \psi_2\psi_1' = -\psi_1\psi_2\frac{d}{dx}\ln(\psi_1/\psi_2) \tag{6.8a}$$

[81] Note that Eq. (6.5a) also has two linearly independent solutions for $\varepsilon = 0$. These solutions are $\psi \propto x$ and $\psi = \text{const}$. The first one is eliminated by the condition (6.7). The second one nominally satisfies (6.7), and can be viewed as the value of the integrand in Eq. (6.5c) at the lower border ($k = 0$) of the integration range. There is no need to be concerned about this value because the integral in Eq. (6.5c) does not depend on it.

of any two solutions $\psi_{1,2}$ of Eq. (6.3b) (not necessarily eigenfunctions) is independent of x:

$$\frac{d}{dx}W[\psi_1,\psi_2] = 0. \tag{6.8b}$$

Indeed, since $\psi_{1,2}'' = (v-\varepsilon)\psi_{1,2}$ [see Eq. (6.3b)] we have

$$\frac{d}{dx}W[\psi_1,\psi_2] = \psi_1'\psi_2' + \psi_1\psi_2'' - \psi_2'\psi_1' - \psi_2\psi_1'' = \psi_1(v-\varepsilon)\psi_2 - \psi_2(v-\varepsilon)\psi_1 = 0.$$

Here are some corollaries to this theorem. First of all, it is obvious from Eq. (6.8a) that

$$\text{solutions } \psi_{1,2} \text{ are linearly dependent if and only if } W[\psi_1,\psi_2] = 0. \tag{6.9a}$$

For example, the Wronskian of the two linearly independent free-particle eigenfunctions $\psi_{k\pm}$ [see Eq. (6.5b)] is finite: $W[\psi_{k+},\psi_{k-}] = -ik/\pi \neq 0$, in agreement with Eq. (6.9a).

Next, suppose there exists such x_0 that $\psi_1(x_0) = \psi_2(x_0) = 0$. Then $W[\psi_1,\psi_2] = 0$, hence, by Eq. (6.9a), the functions $\psi_{1,2}$ are linearly dependent. The same is true if both $\psi_{1,2}$ decay to zero at $x \to +\infty$ or at $x \to -\infty$. Thus,

$$\text{solutions that share a zero (including at infinity) are linearly dependent.} \tag{6.9b}$$

Consider now solutions $\psi_{1,2}$ that satisfy $\psi_1(x_0) = \psi_2(x_0)$ and $\psi_1'(x_0) = \psi_2'(x_0)$ at some finite x_0. Since $W[\psi_1,\psi_2] = 0$, we have $\psi_2(x)/\psi_1(x) = \psi_2(x_0)/\psi_1(x_0) = 1$. Accordingly,

$$\text{solution of Eq. (6.3b) with given } \psi(x_0) \text{ and } \psi'(x_0) \text{ is unique.} \tag{6.9c}$$

Finally, let $\phi_{1,2}$ be solutions of Eq. (6.3b) defined by the relations [see Eq. (6.9c)]

$$\phi_1(x_0) = 1, \quad \phi_1'(x_0) = 0, \quad \phi_2(x_0) = 0, \quad \phi_2'(x_0) = 1.$$

Since $W[\phi_1,\phi_2] = 1 \neq 0$, these solutions are linearly independent [recall Eq. (6.9a)]. Any other solution $\psi(x)$ of Eq. (6.3b) is uniquely specified by $\psi(x_0)$ and $\psi'(x_0)$ [recall Eq. (6.9c)], and can be written as a linear combination of $\phi_{1,2}$,

$$\psi(x) = \psi(x_0)\phi_1(x) + \psi'(x_0)\phi_2(x).$$

In other words, as mentioned above,

$$\text{Eq. (6.3b) has only two linearly independent solutions.} \tag{6.9d}$$

We will also need several important results concerning the behavior of solutions of Eq. (6.3b) at large $|x|$. Suppose there exists such x_0 that $\varepsilon < v(x)$ at all $x > x_0$ (or $x < x_0$). It can be shown rigorously that Eq. (6.3b) has solutions that decay to zero at $x \to +\infty$ (or at $x \to -\infty$). Moreover, it can be shown that this decay is at least as fast as exponential. Note that since the decaying solutions share an asymptotic zero (at infinity), they are linearly dependent [see Eq. (6.9b)]. Conversely, if $\varepsilon > v(x)$ at all $x > x_0$ (or $x < x_0$), then Eq. (6.3b) does not have solutions that decay to zero or diverge at $x \to +\infty$ (or at $x \to -\infty$). These

properties are obvious if $v-\varepsilon=\text{const}$ at all $x>x_0$ (or $x<x_0$). For a generic potential $v(x)$ the proofs are elementary, but somewhat lengthy.

We are now ready to discuss various types of eigenfunctions one may encounter when solving Eq. (6.3b). To this end, we consider the potential that tends to finite limits at $x\to\pm\infty$:

$$\lim_{x\to\pm\infty} v(x)=v_\pm, \quad v_+>v_-. \tag{6.10a}$$

For simplicity, we assume that $v(x)$ reaches its limits $v_\pm$ sufficiently fast, so that the asymptotic behavior of eigenfunctions at large $|x|$ can be found by replacing $v(x)$ in Eq. (6.3b) with its limits $v_\pm$. It can be shown that this amounts to the condition

$$\lim_{x\to\pm\infty} x\,[v(x)-v_\pm]=0. \tag{6.10b}$$

We start with $\varepsilon>v_+$. Both linearly independent solutions of Eq. (6.3b) in this case remain finite at $x\to\pm\infty$, thereby conforming with the condition (6.7). Thus, at $\varepsilon>v_+$ the spectrum of our Hamiltonian is *continuous* and *two-fold degenerate*. At large $|x|$ the corresponding eigenfunctions turn to linear combinations of oscillating exponents, $e^{\pm ik_+x}$ at $x\to+\infty$ and $e^{\pm ik_-x}$ at $x\to-\infty$; the "wave vectors" $k_\pm$ here are given by $k_\pm=\sqrt{\varepsilon-v_\pm}$ [cf. Eq. (6.5b)]. Relating the coefficients of these linear combinations is the subject of the ***scattering theory***.

Next, we consider the interval of energies $v_-<\varepsilon<v_+$. At any ε in this range Eq. (6.3b) has a solution that decays at $x\to-\infty$ and oscillates at $x\to+\infty$. All such solutions satisfy the condition (6.7) and are linearly dependent [see Eq. (6.9b)]. Accordingly, this subset of the spectrum is *continuous* and *non-degenerate*. The corresponding eigenfunctions are obviously proportional to e^{κ_-x} with $\kappa_-=\sqrt{v_--\varepsilon}$ at $x\to-\infty$ [cf. Eq. (6.6)], and are given by linear combinations of $e^{\pm ik_+x}$ at $x\to+\infty$.

We turn now to the remaining interval, $\varepsilon<v_-<v_+$. At any ε in this interval Eq. (6.3b) has a solution that decays at $x\to-\infty$ and another solution that decays at $x\to+\infty$. There is no guarantee, however, that these two solutions coincide. For a generic value of ε this is obviously not the case: solution decaying at $x\to-\infty$ diverges at $x\to+\infty$ and vice versa. However, there *may* exist special values of ε for which solutions of Eq. (6.3b) decay to zero at *both* $x\to\pm\infty$. Such ε form a *discrete* subset of the spectrum. There is no degeneracy, the eigenfunctions decay at $x\to\pm\infty$ as $e^{\mp\kappa_\pm x}$ with $\kappa_\pm=\sqrt{v_\pm-\varepsilon}$, and can be normalized to unity. Thus, unlike the generalized eigenfunctions in the continuous part of the spectrum, the discrete-spectrum ones represent genuine quantum states, the so-called ***bound states***.

Of course, a generic potential $v(x)$ does not necessarily supports all three spectral regimes just discussed. In particular, any potential that diverges at $x\to\pm\infty$ [such as $v(x)\propto x^2$] has an infinite number of bound states but leaves no room for the continuous spectrum. Note also that for a generic potential solutions may decay at $x\to\pm\infty$ faster than exponentially.

Finally, we note that whereas a classical particle with energy E can only move within the regions of space where $E>V(x)$ [i.e., $\varepsilon>v(x)$], the above discussion shows that eigenfunctions differ from zero also in the regions where $\varepsilon\le v(x)$. Accordingly, there is a finite probability that a quantum particle will be found in ***classically inaccessible*** regions of space. The ability of quantum particles to penetrate and traverse such classically inaccessible regions is the essence of the ***tunneling effect***.

6.1.3 Bound states

In this section, we discuss properties of bound states in some details. Because the corresponding eigenfunctions (the ***bound state wave functions***) ψ_ε decay at $x \to \pm\infty$ exponentially or faster and do not diverge at finite x, they can be normalized to unity,

$$\int dx\, |\psi_\varepsilon(x)|^2 = 1. \tag{6.11}$$

Since any two such wave functions (with the same ε) are necessarily linearly dependent, they may differ in independent of x overall phase factors at most. Therefore, without loss of generality,

$$\psi_\varepsilon(x) \text{ can be taken to be real.} \tag{6.12}$$

(This observation applies to eigenfunctions in the continuous part of the spectrum as well.[82]) Note that he condition $\operatorname{Im}\psi_\varepsilon = 0$ leaves the sign of $\psi_\varepsilon(x)$ at one's discretion.

Just as eigenvectors of Hermitian operators in finite-dimensional Hilbert spaces [recall Eq. (2.46a)], bound state wave functions form an orthonormal set,

$$\int dx\, \psi_\varepsilon(x)\psi_{\varepsilon'}(x) = \delta_{\varepsilon,\varepsilon'}. \tag{6.13}$$

Indeed, taking into account that $\varepsilon\psi_\varepsilon = -\psi_\varepsilon'' + v\psi_\varepsilon$ [see Eq. (6.3b)], we find for $\varepsilon \neq \varepsilon'$

$$\begin{aligned}(\varepsilon - \varepsilon')\int dx\, \psi_\varepsilon\psi_{\varepsilon'} &= \int dx\, \left[(\varepsilon\psi_\varepsilon)\psi_{\varepsilon'} - \psi_\varepsilon(\varepsilon'\psi_{\varepsilon'})\right] \\ &= \int dx\, \left(\psi_\varepsilon\psi_{\varepsilon'}'' - \psi_{\varepsilon'}\psi_\varepsilon''\right) = \int dx\, \frac{d}{dx}\left(\psi_\varepsilon\psi_{\varepsilon'}' - \psi_{\varepsilon'}\psi_\varepsilon'\right) = 0.\end{aligned}$$

One should keep in mind however that for potentials supporting a continuous spectrum bound state wave functions do not form a complete set [recall Eq. (5.29)].

Because the phase of the bound state wave function $\psi_\varepsilon(x)$ is independent of x, the expectation value of the momentum in this state vanishes identically,[83]

$$\langle p\rangle_\varepsilon = 0 \tag{6.14}$$

[see the remark after Eqs. (5.57)]. At the same time, $\langle p^2\rangle_\varepsilon > 0$ (as opposed to $\langle p^2\rangle_\varepsilon \geq 0$) even in the state with the lowest possible eigenenergy $\varepsilon_0 = \min\{\varepsilon\}$ (i.e., the ***ground state***). Such ***zero-point motion*** is a quantum effect. Indeed, the inequality $\langle p^2\rangle > 0$ follows directly from the uncertainty relation (5.60): in view of Eq. (6.14) we have $\langle p^2\rangle = (\Delta p)^2 \geq (\hbar/2\Delta x)^2 > 0$.

It was shown in Sec. 6.1.2 that bound state energies are bounded from above,

$$\varepsilon < \lim_{x\to\pm\infty} v(x). \tag{6.15a}$$

[82] If Eq. (6.3b) has a solution $\psi(x)$, then its complex conjugate $\psi^*(x)$ is also a solution. Instead of ψ and ψ^*, one can consider *real* functions $\operatorname{Re}\psi = \frac{1}{2}(\psi + \psi^*)$ and $\operatorname{Im}\psi = \frac{1}{2i}(\psi - \psi^*)$. Linearity of Eq. (6.3b) implies that these two functions are solutions. For example, instead of $\psi_{k\pm} \propto e^{\pm ikx}$ for $v \equiv 0$ [see Eq. (6.5b)] we could have chosen linearly independent functions $\psi_{e,k} \propto \cos(kx)$ and $\psi_{o,k} \propto \sin(kx)$. Note that these functions also have definite *parities*: the indices e and o stand for the *even* and *odd* parities, respectively.

[83] Note that if $|\psi\rangle$ is not an eigenvector of $\hat{H}$ then the expectation value $\langle p\rangle_\psi$ does not have to vanish.

It is easy to see that the energies are also bounded from below. Indeed, we have

$$\varepsilon = \frac{2m}{\hbar^2}\langle\psi_\varepsilon|\hat{H}|\psi_\varepsilon\rangle = \frac{1}{\hbar^2}\langle p^2\rangle_\varepsilon + \int dx\, v(x)\psi_\varepsilon^2(x) > \int dx\, v(x)\psi_\varepsilon^2(x) \geq \min\{v(x)\}\int dx\, \psi_\varepsilon^2(x).$$

Taking into account Eq. (6.11), we obtain the lower bound

$$\varepsilon > \min\{v(x)\}. \tag{6.15b}$$

In addition to the ***asymptotic zeros*** at $x \to \pm\infty$, bound state wave functions may also have ***true zeros*** at finite x. Eq. (6.9c) implies that all such zeros are *isolated*. That is, if $\psi_\varepsilon(x_0) = 0$ then $\psi'_\varepsilon(x_0) \neq 0$ (otherwise, $\psi_\varepsilon(x)$ would vanish identically). It turns out that each *classically inaccessible* region $\varepsilon \leq v(x)$ [recall Sec. 6.1.2] contains at most a single zero of $\psi_\varepsilon(x)$, either true or asymptotic.[84] Indeed, suppose the wave function $\psi_\varepsilon(x)$ has two zeros in such region, say, at $x = x_{1,2}$. This implies that $\int_{x_1}^{x_2} dx\, \psi'_\varepsilon(x) = 0$. On the other hand, we have $\psi''_\varepsilon = (v - \varepsilon)\psi_\varepsilon$ [see Eq. (6.3b)] and thus

$$\int_{x_1}^{x_2} dx\, \psi'_\varepsilon(x) = \int_{x_1}^{x_2} dx \int_{x_1}^{x} dy\, \psi''_\varepsilon(y) = \int_{x_1}^{x_2} dx \int_{x_1}^{x} dy \big[v(y) - \varepsilon\big]\psi_\varepsilon(y).$$

Because $\psi_\varepsilon(y) \neq 0$ at $x_1 < y < x_2$, this integral does not vanish unless $v \equiv \varepsilon$, ruling out the existence of two zeros in this case. If $v \equiv \varepsilon$ at all x in the interval $x_1 < x < x_2$ then the unique solution of Eq. (6.3b) vanishing at $x = x_1$ reads $\psi_\varepsilon(x) = (x - x_1)\psi'_\varepsilon(x_1)$, which yields $\psi_\varepsilon(x_2) \neq 0$, again ruling out the existence of the second zero. This concludes the proof.

Remarkably, the total number of the true zeros is known beforehand. Let $\psi_0, \psi_1, \psi_2, \ldots$ be the bound state wave functions labeled in order of increasing eigenenergies, i.e.,

$$\varepsilon_0 < \varepsilon_1 < \varepsilon_2 < \ldots \tag{6.16a}$$

Then it can be shown rigorously that

$$\psi_n(x) \text{ has exactly } n \text{ true zeros}. \tag{6.16b}$$

In particular, the ground state wave function $\psi_0(x)$ has no true zeros at all. Hereinafter, we will adopt this labeling convention. Eqs. (6.16) provide a simple method of verifying that all bound state wave functions have been accounted for. It also allows one to *guess* the shape of the wave function without actually finding it.

One often encounters potentials that are even functions of x, i.e., such that

$$v(x) = v(-x). \tag{6.17a}$$

Changing the variable in Eq. (6.3b) to $-x$ shows that if $\psi(x)$ is a normalized real solution of this equation, then $\psi(-x)$ is a solution as well. Because bound states are not degenerate, the functions $\psi(x)$ and $\psi(-x)$ either coincide or have opposite signs. Accordingly, all bound state wave functions in potentials that satisfy Eq. (6.17a) have a definite ***parity***, either ***even*** or ***odd***. Taking into account Eq. (6.16b), we obtain the relation

$$\psi_n(x) = (-1)^n \psi_n(-x). \tag{6.17b}$$

[84] In particular, if $\varepsilon < v(x)$ at all x then any solution of Eq. (6.3b) that has an asymptotic zero $x \to -\infty$ diverges at $x \to +\infty$ and vice versa. This observation leads to the bound (6.15b).

In other words, ψ_n with an even (odd) index n is an even (odd) function of x.

To close this section, we emphasize that for a generic potential $v(x) = (2m/\hbar^2)V(x)$ that does not diverge at $x \to \pm\infty$ the existence of bound state solutions of Eq. (6.3b) is not guaranteed, nor is it possible to formulate a universal criterion of their existence. It can be shown, however, that for potentials that satisfy

$$\lim_{x\to+\infty} v(x) = \lim_{x\to-\infty} v(x) = v_\infty \tag{6.18a}$$

the sufficient condition for the existence of at least one bound state reads

$$\kappa_0 = \frac{1}{2}\int dx \big[v_\infty - v(x)\big] > 0\,. \tag{6.18b}$$

Note that this criterion does not apply if $\lim_{x\to+\infty} v(x) \neq \lim_{x\to-\infty} v(x)$, i.e., when the potential has uneven "shoulders" as in Eq. (6.10a).

6.2 Case study: bound states in a rectangular potential well

Consider the potential

$$v(x) = \frac{2m}{\hbar^2}V(x) = \begin{cases} u, & |x| > a/2\,, \\ 0, & |x| < a/2\,, \end{cases} \tag{6.19}$$

describing a rectangular potential well of depth u and width a. This potential and its various modifications appear in numerous applications. In particular, it is often used to model electrons in semiconducting heterostructures.

According to the discussion in Sec. 6.1.2, potential (6.19) supports a continuous two-fold degenerate spectrum at energies $\varepsilon > u$. These eigenfunctions will not be discussed here. Instead, we focus on the bound states. Eqs. (6.15) show their energies lie in the range

$$0 < \varepsilon_n < u, \tag{6.20}$$

below the continuous spectrum, and form the hierarchy (6.16a). Although the eigenenergies increase monotonically with the depth of the well,[85]

$$0 < \frac{d}{du}\varepsilon_n(u) < 1, \tag{6.21}$$

it will be shown in Sec. 6.2.1 that they have finite limits at $u \to \infty$,

$$\lim_{u\to\infty} \varepsilon_n(u) = \varepsilon_n(\infty) < \infty. \tag{6.22}$$

[85] Let a_λ be a non-degenerate eigenvalue of Hermitian operator $\hat{A}_\lambda$ that depends on parameter λ and let $|\psi_\lambda\rangle$ be the corresponding eigenvector. It is easy to show that $\partial a_\lambda/\partial\lambda = \langle\psi_\lambda|\partial\hat{A}_\lambda/\partial\lambda|\psi_\lambda\rangle$; this result is known as the ***Hellmann-Feynman theorem***. Treating the depth of the well u in Eq. (6.19) as a parameter and taking into account that $\partial v(x)/\partial u = \theta(|x| - a/2)$, we obtain

$$\frac{d}{du}\varepsilon_n(u) = \int_{|x|>a/2} dx\, \psi_n^2(x)$$

for the problem at hand, from which the inequalities (6.21) follow at once. More generally, if $v_1(x) \geq v_2(x)$ at all x, then the bound state energies corresponding to these potentials satisfy $\varepsilon_{1;n} \geq \varepsilon_{2;n}$.

Moreover, we will see in Sec. 6.2.2 and 6.2.3 that the intervals in which the eigenenergies vary with u do not overlap:

$$\varepsilon_{n-1}(\infty) < \varepsilon_n(u) < \min\bigl\{\varepsilon_n(\infty), u\bigr\}, \tag{6.23}$$

where $\varepsilon_{n-1}(\infty)$ should be replaced with zero for $n=0$. The lower border of this interval,

$$u_n = \begin{cases} 0, & n=0, \\ \varepsilon_{n-1}(\infty), & n>0, \end{cases} \tag{6.24}$$

is the ***threshold depth*** for nth bound state. When the depth of the well u increases, a new bound state splits from the continuous spectrum each time u crosses the corresponding threshold. With further increase of u, the eigenenergy $\varepsilon_n(u)$ moves away from the lower end of the interval (6.23), approaching its upper end at $u\to\infty$. The inequalities (6.21) show that this motion is continuous and smooth. Notice that since $u_0=0$, the ground state is guaranteed to be a bound state for any $u>0$, in agreement with Eqs. (6.18). The total number of bound states N depends on the relation between the depth of the well and its width, or, rather, on the value of the dimensionless parameter $a\sqrt{u}$. It will be shown in Sec. 6.2.2 that $N\propto a\sqrt{u}$ if this parameter large.

All these results follow from Eq. (6.3b). The standard strategy for piecewise-constant potentials such as (6.19) is to find general solutions of Eq. (6.3b) in the regions where the potential is constant, and then match these solutions using the continuity of the wave function and its derivative. For the problem at hand, Eq. (6.3b) reduces to $\psi_n''=-\varepsilon_n\psi_n$ at $|x|<a/2$ and $\psi_n''=(u-\varepsilon_n)\psi_n$ at $|x|>a/2$. Taking into account the parity of the wave functions [recall Eqs. (6.17)] and the inequalities (6.20), we write solutions of these equations vanishing at $x\to\pm\infty$ as

$$\psi_n(x)\Big|_{|x|<a/2} \propto \begin{cases} \cos(k_n x), \text{ even } n, \\ \sin(k_n x), \text{ odd } n, \end{cases} \qquad \psi_n(x)\Big|_{|x|>a/2} \propto \begin{cases} e^{-\kappa_n x}, & x>a/2, \\ e^{\kappa_n x}, & x<-a/2, \end{cases} \tag{6.25a}$$

where

$$k_n(u)=\sqrt{\varepsilon_n(u)}>0, \quad \kappa_n(u)=\sqrt{u-\varepsilon_n(u)}>0. \tag{6.25b}$$

Continuity of $\psi_n(x)$ and $\psi_n'(x)$ at $x=\pm a/2$ implies that their ratio is continuous as well,[86]

$$\frac{\psi_n'(x)}{\psi_n(x)}\Big|_{x\to\pm(a/2-0)} = \frac{\psi_n'(x)}{\psi_n(x)}\Big|_{x\to\pm(a/2+0)}. \tag{6.26}$$

Note that the second equation in (6.25a) shows that $\psi_n(\pm a/2)\neq 0$, hence the limits in Eq. (6.26) are well-defined.

Substitution of the wave functions (6.25) into the matching condition (6.26) yields

$$k_n\tan(k_n a/2)=\kappa_n, \quad k_n^2+\kappa_n^2=u \tag{6.27}$$

[86] The advantage of working with the ratio of ψ_n and ψ_n' is that it allows one not to deal with the proportionality coefficients in Eqs. (6.25).

for even n and similar equations for odd n. Unfortunately, as it is often the case, these equations are transcendental and cannot be solved analytically for arbitrary u. Instead, we focus on the asymptotic behavior of the eigenenergies near the upper and the lower ends of the interval (6.23), i.e., at large u and small $u-u_n$, respectively.

6.2.1 Hard-wall limit

In this section we discuss the bound state energies and the corresponding wave functions in the so-called ***hard-wall limit*** $u\to\infty$. Provided that $\varepsilon_n(\infty)=\lim_{u\to\infty}\varepsilon_n(u)<\infty$ [cf. Eq. (6.22); this assumption will be validated shortly below], we have $\lim_{u\to\infty}\kappa_n(u)=+\infty$. The second equation in (6.25a) then gives

$$\lim_{u\to\infty}\psi_n(x)\big|_{|x|>a/2}=0, \tag{6.28a}$$

which shows that in the limit $u\to\infty$ the potential walls at $x=\pm a/2$ become *impenetrable*, or "*hard*". Because the wave function is continuous at $x=\pm a/2$, Eq. (6.28a) implies that

$$\lim_{u\to\infty}\psi_n(x)\big|_{x=\pm(a/2-0)}=0. \tag{6.28b}$$

Since the wave function and its derivative cannot vanish simultaneously [otherwise, the wave function would vanish identically, see Eq. (6.9c)], we have

$$\lim_{u\to\infty}\psi_n'(x)\big|_{x=\pm(a/2-0)}\neq 0. \tag{6.28c}$$

Accordingly, in the hard-wall limit the derivative of the wave function develops discontinuities at $x=\pm a/2$. This, of course, is not surprising as there is no reason to expect ψ_n' to be continuous if the potential in Eq. (6.3b) is allowed to diverge at finite x.

Instead of solving Eq. (6.3b) and only then sending u to infinity, one can take the hard-wall limit $u\to\infty$ from the outset. Eq. (6.28b) suggests that in this limit the potential (6.19) can be replaced with the ***hard-wall boundary conditions***[87]

$$\psi_n(x)\big|_{x=\pm(a/2-0)}=0. \tag{6.29}$$

This drastically simplifies the problem: it reduces to finding non-vanishing solutions of the equation $-\psi_n''=\varepsilon_n\psi_n$ with $\varepsilon_n\equiv\varepsilon_n(\infty)$ in the finite interval $|x|<a/2$ (i.e., *inside* the potential well).

It is easy to verify that it is impossible to satisfy the boundary conditions if $\varepsilon_n\le 0$, hence $\varepsilon_n=k_n^2>0$, in agreement with Eq. (6.20). The corresponding wave functions are given by the first equation in (6.25a) at $|x|<a/2$ and by $\psi_n(x)=0$ at $|x|>a/2$ [cf. Eq. (6.28a)]. Substituting these wave functions into Eq. (6.29) and taking into account that $k_n>0$, we obtain

$$\frac{k_n a}{2}=\begin{cases}\pi/2+\pi m, & \text{even } n,\\ \pi+\pi m, & \text{odd } n,\end{cases} \tag{6.30a}$$

[87] Note that Eqs. (6.28b) and (6.29) are not identical. In Eq. (6.28b) the limit $x\to\pm(a/2-0)$ is taken first, and only then u is sent to infinity, whereas in Eq. (6.29) the order of limits is reversed. We will see in Sec. 6.2.3 that the boundary condition (6.29) indeed yields correct results.

where $m = 0,1,2,\ldots$ are non-negative integers. For $m = 0$ the wave functions have no zeros at $0 < x < a/2$. Each additional m adds a zero in this interval. Therefore, the total number of the true zeros of $\psi_n(x)$ is

$$n = \begin{cases} 2m, & \text{even } n, \\ 2m+1, & \text{odd } n \end{cases} \tag{6.30b}$$

[recall Eqs. (6.16)]; note that the boundary values (6.29) should be regarded as asymptotic zeros brought from $\pm\infty$ rather than as true zeros. Eqs. (6.30) yield

$$k_n(\infty) = \lim_{u\to\infty} k_n(u) = (\pi/a)(n+1) \tag{6.30c}$$

for both even and odd $n = 0,1,2,\ldots$ The corresponding eigenenergies

$$\varepsilon_n(\infty) = \lim_{u\to\infty} \varepsilon_n(u) = k_n^2(\infty) = (\pi/a)^2(n+1)^2 \tag{6.31}$$

are finite, $0 < \varepsilon_n(\infty) < \infty$, in agreement with the assumption made in the beginning of this section.

The bound state wave functions are easily normalized: we have

$$\int_{-a/2}^{a/2} dx \begin{Bmatrix} \cos^2(k_n x) \\ \sin^2(k_n x) \end{Bmatrix} = \int_{-a/2}^{a/2} dx\, \frac{1}{2}\left[1 \pm \cos(2k_n x)\right] = a/2$$

[the integral of the cosine vanishes because the integration runs over an integer number of its periods, $2k_n a = 2\pi(n+1)$], so that the properly normalized hard-wall-limit wave functions read

$$\psi_n(x) = (2/a)^{1/2}\,\theta\big(a/2 - |x|\big) \times \begin{cases} \cos(k_n x), & \text{even } n, \\ \sin(k_n x), & \text{odd } n, \end{cases} \tag{6.32}$$

where $k_n = k_n(\infty)$. The signs of the wave functions are, of course, arbitrary [see the remark after Eq. (6.12)].

We expect the hard-wall-limit results (6.31) and (6.32) to be a very good approximation for the corresponding finite-u expressions as long as $\varepsilon_n \ll u$, i.e., for low-energy bound states. Of course, for finite u the boundary conditions (6.29) no longer apply, and the wave functions develop exponential "tails" extending to the classically inaccessible regions $|x| > a/2$ [see the second equation in (6.25a)]. However, for states with $\varepsilon_n \ll u$ most of the probability density $\psi_n^2(x)$ is still confined within the interval $|x| < a/2$, and the wave functions in this interval are again given by the oscillating sines and cosines [see the first equation in (6.25a)], albeit with slightly smaller $k_n(u) < k_n(\infty)$.

Irrespective of their origin, the hard-wall-limit eigenfunctions (6.32)] are interesting in their own right. In particular, it is not difficult to show[88] that these functions form a complete set in the sense that

$$\sum_n \psi_n(x)\psi_n(x') = \delta(x - x') \quad \text{for all } |x|, |x'| < a/2, \tag{6.33a}$$

[88] It is straightforward to evaluate the sum in the left-hand side of Eq. (6.33a) explicitly by writing it as a limit $\lim_{\gamma\to+0} \sum_{n=0}^{n=\infty} \psi_n(x)\psi_n(x')\,e^{-\gamma n}$.

so that any function $f(x)$ defined in the interval $|x| < a/2$ can be expanded as

$$f(x) = \sum_n f_n \psi_n(x), \quad f_n = \int_{-a/2}^{a/2} dx\, f(x)\psi_n(x). \tag{6.33b}$$

To illustrate how this expansion works in practice, we discuss the simplest possible example, $f(x) = 1$. Taking into account that in this case $f_n \neq 0$ only for even $n = 2l$, we arrive at the *expansion of unity*

$$1 = \sum_{l=0}^{\infty} f_{2l}\, \psi_{2l}(x), \quad f_{2l} = (-1)^l \frac{\sqrt{8a}}{\pi(2l+1)}. \tag{6.34}$$

Notice that for any finite n, no matter how large, the sum

$$S_n(x) = \sum_{l=0}^{l=n} f_{2l}\, \psi_{2l}(x) \tag{6.35a}$$

is a continuous function of x that satisfies $S_n(x \to \pm a/2) = 0$, in contradiction with Eq. (6.34). This paradox is only apparent and its resolution lies in the non-commutativity of the order of limits $n \to \infty$ and $x \to \pm a/2$. Indeed, it can be shown that

$$\lim_{n\to\infty} S_n(x) = 1 \text{ for any } x \neq \pm a/2. \tag{6.35b}$$

Thus, in order to get unity at $x \to \pm a/2$ as in Eq. (6.34), one must take n to infinity first, and only then send x to $\pm a/2$.

This example shows that expansions such as (6.33b) and (6.34) should be handled with care. In fact, they are valid only in the Hilbert-space sense:[17,19] it is easy to see that the Hilbert-space norm of $1 - S_n$ tends to zero in the limit $n \to \infty$,

$$\big\|1 - S_n\big\|^2 = \int_{-a/2}^{a/2} dx\, [1 - S_n(x)]^2 = a\left[1 - \frac{8}{\pi^2}\sum_{l=0}^{l=n} \frac{1}{(2l+1)^2}\right] \xrightarrow{n\to\infty} 0, \tag{6.36}$$

where we took into account the well-known result $\sum_{n=0}^{\infty}(2n+1)^{-2} = \pi^2/8$.

6.2.2 Bound states on the verge of their disappearance

Eq. (6.20) implies that for each n there exists such u_n that the eigenenergy of nth bound state belongs to the interval

$$u_n < \varepsilon_n(u) < u. \tag{6.37}$$

We interpret u_n as the *threshold depth* of the well for nth bound state. That is, the well can accommodate this state if $u > u_n$, and there is no room for it if $u \leq u_n$. According to Eq. (6.20), the threshold depth satisfies $u_n \geq 0$ for all n. In this section, we show that this inequality is not tight: the exact value of u_n depends on n and is given by Eq. (6.24).

When u decreases, the length of the interval in Eq. (6.37) shrinks to zero, leading to the equation

$$\lim_{u\to u_n+0} \varepsilon_n(u) = u_n. \tag{6.38}$$

An order-of-magnitude estimate of the threshold depth can be obtained by replacing $\varepsilon_n(u)$ in the left-hand side of this equation with the hard-wall-limit result $\varepsilon_n(\infty)$ [see Eq. (6.31)]. This yields

$$u_n \sim (n/a)^2, \quad n \gg 1. \tag{6.39}$$

Substitution of u_n in this form into $u_N \sim u$ then gives the estimate of the total number of bound states $N \sim a\sqrt{u}$ quoted above; this estimate is applicable as long as $N \gg 1$.

We will now evaluate the threshold depth u_n exactly. Consider the bound state wave functions (6.25) in the limit $u \to u_n + 0$. The second equation in (6.25b) and Eq. (6.38) show that $\lim_{u\to u_n+0} \kappa_n(u) = +0$. The second equation in (6.25a) then gives

$$\lim_{u\to u_n+0} \psi_n'(x)\big|_{x=\pm(a/2+0)} = 0. \tag{6.40a}$$

Because the derivative $\psi_n'(x)$ is continuous at $x = \pm a/2$, this leads to the relation

$$\lim_{u\to u_n+0} \psi_n'(x)\big|_{x=\pm(a/2-0)} = 0. \tag{6.40b}$$

Instead of taking the limit $u \to u_n + 0$ in the solution (6.25) of Eq. (6.3b), one can send u to u_n directly in the equation. Ignoring "$+0$", we write

$$-\psi_n'' = u_n \psi_n, \quad |x| < a/2 \tag{6.41a}$$

We are interested in the solutions of this equation subject to the boundary conditions

$$\psi_n'(x)\big|_{x=\pm(a/2-0)} = 0 \tag{6.41b}$$

obtained by changing the order of limits in Eq. (6.40b). [It will be shown in Sec. 6.2.2 that the reduction of the problem to Eqs. (6.41) yields the correct value of u_n.]

It is easy to see that the non-vanishing solutions of Eqs. (6.41) correspond to $u_n = k_n^2 \geq 0$ and have the form of the first equation in (6.25a). Substituting these solutions into the boundary conditions (6.41b) and counting zeros of $\psi_n(x)$ as in Sec. 6.2.1, we obtain

$$\frac{k_n a}{2} = \begin{cases} \pi m, & \text{even } n, \\ \pi/2 + \pi m, & \text{odd } n, \end{cases} \qquad n = \begin{cases} 2m, & \text{even } n, \\ 2m+1, & \text{odd } n, \end{cases} \tag{6.42a}$$

where m are non-negative integers. These equations result in

$$k_n = \pi n/a \tag{6.42b}$$

for both parities. Taking into account Eq. (6.31), we arrive at the exact threshold depth

$$u_n = k_n^2 = (\pi n/a)^2 = \begin{cases} 0, & n = 0, \\ \varepsilon_{n-1}(\infty), & n > 0 \end{cases} \tag{6.43}$$

[cf. Eqs. (6.24) and (6.39)]. With Eq. (6.43) at hand, it is easy to show that the total number of bound states supported by a rectangular potential well of depth u and width a is

$$N = \text{integer part of} \left\{1 + \frac{a\sqrt{u}}{\pi}\right\}. \tag{6.44}$$

As we have anticipated, it scales as $N \propto a\sqrt{u}$ at large $a\sqrt{u} \gg 1$.

6.2.3 Rectangular well of arbitrary depth

In Sec. 6.2.1 and 6.2.2 we discussed the limits $u \to \infty$ and $u \to u_n + 0$. In this section, we go beyond these limits and consider the even-parity bound states[89] in the potential (6.19) with arbitrary $u > 0$.

Using the second equation in Eq. (6.27) to exclude κ_n from the first equation, we obtain

$$\sqrt{u - k_n^2} = k_n \tan(k_n a/2). \tag{6.45a}$$

It is convenient to denote

$$\theta = k_n a/2 = a\sqrt{\varepsilon_n}/2, \quad \nu = a\sqrt{u}/2 \tag{6.45b}$$

and rewrite Eq. (6.45a) in the dimensionless form

$$\sqrt{(\nu/\theta)^2 - 1} = \tan\theta. \tag{6.45c}$$

We are interested in the solutions of Eq. (6.45c) that satisfy

$$0 < \theta < \nu \tag{6.46a}$$

[see Eqs. (6.20) and (6.45b)]. Because at such θ the left-hand side of Eq. (6.45c) is positive, its solutions can only be found in the intervals where $\tan\theta$ in the right-hand side is positive as well, i.e., at

$$\pi m < \theta < \pi m + \pi/2, \tag{6.46b}$$

where $m = 0, 1, 2, \ldots$ are non-negative integers. In each such interval the left-hand side of Eq. (6.45c) monotonically decreases with θ reaching zero at $\theta = \nu$, whereas the right-hand side monotonically increases from $\tan(\pi m + 0) = +0$ to $\tan(\pi m + \pi/2 - 0) = +\infty$. Therefore, the interval (6.46b) contains at most a single solution for θ. Obviously, this solution exists if and only if

$$\nu > \pi m \tag{6.47a}$$

and represents the even-parity bound state with

$$n = 2m. \tag{6.47b}$$

In the original notations, Eqs. (6.46) and (6.47) translate to the inequalities

$$u_n < \varepsilon_n(u) < \min\{\varepsilon_n(\infty), u\}. \tag{6.48}$$

[cf. Eqs. (6.23)], with u_n and $\varepsilon_n(\infty)$ given by Eqs. (6.24) and (6.31), respectively.

We assume now that the condition (6.47a) (i.e., $u > u_n$) is met and consider the dependence of the eigenenergy ε_n on the depth of the well u. Solving Eq. (6.45c) for ν, we obtain

$$\nu(\theta) = \frac{\theta}{\cos(\theta - \pi n/2)}, \quad \pi n/2 < \theta < \pi(n+1)/2. \tag{6.49}$$

[89] The odd-parity bound states can be analyzed in the same manner. Such analysis shows that Eqs. (6.51) and (6.53b) apply to odd n as well.

This function increases monotonically from $\nu(\pi n/2) = \pi n/2$ to $\nu\big(\pi(n+1)/2\big) = +\infty$ and is analytic (i.e., infinitely differentiable) at all θ. Therefore, the corresponding inverse function $\theta(\nu)$ is also monotonically increasing and, by the Lagrange inversion theorem,[90] it is analytic at all $\nu > \pi n/2$. This implies that the function $\theta^2(\nu^2)$ is monotonically increasing and analytic at $\nu^2 > (\pi n/2)^2$, or, equivalently, that the eigenenergy $\varepsilon_n(u)$ grows smoothly with the depth of the well u, in agreement with Eq. (6.21).

Because Eq. (6.45c) is transcendental, the function $\theta(\nu)$ and thus $\varepsilon_n(u)$ cannot be evaluated analytically. It is possible, however, to derive the asymptotes of $\varepsilon_n(u)$ at large and small u. The regime of large u (i.e., $\nu \gg 1$) is easier to understand, and we discuss it first. The derivative of $\ln\theta(\nu)$ with respect to $1/\nu$ is given by

$$\frac{d}{d(1/\nu)} \ln\theta(\nu) = \theta^{-1} \frac{d\nu}{d(1/\nu)} \frac{d\theta}{d\nu} = -\frac{\nu^2}{\theta(d\nu/d\theta)}.$$

Substituting here $\nu(\theta)$ from Eq. (6.49) and integrating the resulting expression, we obtain

$$\ln\left[\frac{\theta(\infty)}{\theta(\nu)}\right] = \int_0^{1/\nu} d(1/\nu') \, \frac{\theta(\nu')}{\cos[\theta(\nu') - \pi n/2] + \theta(\nu')\sin[\theta(\nu') - \pi n/2]}. \tag{6.50a}$$

This equation can be iterated to generate an asymptotic expansion of $\ln\big[\theta(\infty)/\theta(\nu)\big]$ in powers of $1/\nu \ll 1$. The leading term in this expansion is obtained by replacing the integrand in the right-hand side of Eq. (6.50a) with its $\nu' \to \infty$ limit, i.e., unity [recall that $\theta(\infty) = \pi(n+1)/2$]. After such replacement, Eq. (6.50a) immediately yields

$$\ln\left[\frac{\theta(\infty)}{\theta(\nu)}\right] = \frac{1}{\nu}. \tag{6.50b}$$

It is easy to verify that corrections to this result are small, of order $1/\nu^2$. Using now Eqs. (6.45b), we obtain the eigenenergies

$$\frac{\varepsilon_n(u)}{\varepsilon_n(\infty)} = \frac{\theta^2(\nu)}{\theta^2(\infty)}\bigg|_{\nu = a\sqrt{u}/2} = e^{-4/a\sqrt{u}}. \tag{6.51}$$

This equation describes the asymptotic approach of $\varepsilon_n(u)$ with a fixed index n to its hard-wall limit $\varepsilon_n(\infty)$ when the total number of bound states $N \sim a\sqrt{u}$ is taken to infinity. Although this is clearly not the same as the dependence of ε_n on n at fixed N,[91] we expect Eq. (6.51) to be applicable for all $n \ll N$.

[90] Consider a function $x(y)$ that is analytic at $y = y_0$ and satisfies $x(y_0) = x_0$ and $x'(y_0) \neq 0$. The ***Lagrange inversion theorem*** (proved by methods of complex calculus) asserts that the inverse function $y(x)$ is analytic at $x = x_0$. Furthermore, for an arbitrary function $f(y)$ analytic at $y = y_0$ the expansion of $f\big(y(x)\big)$ in Taylor series near $x = x_0$ is given by the ***Lagrange inversion formula***

$$f\big(y(x)\big) = f(y_0) + \sum_{k=1}^{k=\infty} \frac{c_k}{k!}(x - x_0)^k, \quad c_k = \lim_{y \to y_0} \frac{d^{k-1}}{dy^{k-1}} \left[\frac{y - y_0}{x(y) - x_0} \frac{df(y)}{dy}\right]^k.$$

[91] It is easy to show that $\theta(\infty) - \theta(\nu)$ and thus $\varepsilon_n(\infty) - \varepsilon_n(u)$ increase with n.

In practice, one is usually interested in the lowest-order correction to $\varepsilon_n(\infty)$, i.e.,

$$\frac{\varepsilon_n(u)-\varepsilon_n(\infty)}{\varepsilon_n(\infty)} = -\frac{4}{a\sqrt{u}}. \tag{6.52a}$$

To obtain this correction, Eq. (6.50a) is not needed. In fact, it is easy to deduce (6.52a) directly from Eq. (6.45a). Indeed, at large u one can neglect $k_n(u)$ in the left-hand side of Eq. (6.45a) and replace the first $k_n(u)$ in the right-hand side of this equation with $k_n(\infty) = (\pi/a)(n+1)$ [see Eq. (6.30c)]. Because at large u the argument of $\tan\bigl[(a/2)k_n(u)\bigr]$ approaches $\pi(n+1)/2$, we have

$$\tan\bigl[(a/2)k_n(u)\bigr] = \tan\Bigl\{\pi(n+1)/2 - (a/2)[k_n(\infty)-k_n(u)]\Bigr\} \approx \frac{1}{(a/2)[k_n(\infty)-k_n(u)]}.$$

With these approximations, Eq. (6.45a) yields

$$\frac{k_n(u)}{k_n(\infty)} = 1 - \frac{2}{a\sqrt{u}}, \tag{6.52b}$$

from which Eq. (6.52a) follows at once.

We turn now to the regime of small $u-u_n$, i.e., $\nu-\pi n/2 \ll 1$. It is convenient to replace the domain of definition of the function $\nu(\theta)$ [see Eq. (6.49)] with $\pi n/2-\Delta < \theta < \pi n/2+\Delta$. (Of course, the physical meaning can be assigned only to $\nu > \pi n/2$.) If Δ is sufficiently small, $\nu(\theta)$ is monotonically increasing and analytic. Therefore, its inverse $\theta(\nu)$ is uniquely defined, monotonically increasing, and analytic[90] at $\nu(\pi n/2-\Delta) < \nu < \nu(\pi n/2+\Delta)$.

Because for $n=0$ the derivative of $\nu^2(\theta)$ vanishes at $\theta=0$, the expansion of $\theta^2(\nu^2)$ near $\nu^2=0$ cannot be obtained by a direct application of the Lagrange inversion formula.[90] Instead, we use this formula to expand $\theta(\nu)$. Since $\nu(\theta)$ is an odd function, so is its inverse $\theta(\nu)$, hence the expansion of $\theta(\nu)$ near $\nu=0$ includes odd powers of ν only,

$$\theta(\nu) = \nu - \frac{1}{2}\nu^3 + \dots$$

Therefore, the expansion of θ^2 includes only even powers of ν,

$$\theta^2(\nu^2) = \nu^2 - \nu^4 + \dots$$

Taking into account Eqs. (6.45b), we write the ground state energy as

$$\varepsilon_0(u) = u - \frac{1}{4}(au)^2 f_0(a^2u), \tag{6.53a}$$

where $f_0(\xi)$ is an analytic function that satisfies $f_0(0)=1$.

For $n>0$ the function $\nu^2(\theta)$ has a non-zero derivative at $\theta=\pi n/2$. The Lagrange formula[90] then yields the expansion of $\theta^2(\nu^2)$ in powers of $\nu^2-(\pi n/2)^2$,

$$\theta^2(\nu^2) = (\pi n/2)^2 + \bigl[\nu^2-(\pi n/2)^2\bigr] - \frac{1}{4}\bigl[\nu^2-(\pi n/2)^2\bigr]^2 + \dots.$$

Using Eqs. (6.24) and (6.45b), we obtain

$$\varepsilon_n(u) = u - \frac{1}{16}a^2(u-u_n)^2 f_n\left(\frac{u-u_n}{u_n}\right) \tag{6.53b}$$

for $n > 0$; here $f_n(\xi)$ is an analytic function satisfying $f_n(0) = 1$. (Further analysis shows that this function is the same for all $n > 0$.)

When the depth of the well u approaches the threshold value u_n, the eigenenergy $\varepsilon_n(u)$ varies within the interval $u_n < \varepsilon_n(u) < u$ [recall Eq. (6.37)]. Eqs. (6.53) show that $\varepsilon_n(u)$ is in fact much closer to the upper end of this interval:

$$u - \varepsilon_n(u) \ll \varepsilon_n(u) - u_n . \tag{6.54}$$

It is also worth noting that Eq. (6.53a) cannot be obtained by formally setting $n = 0$ in Eq. (6.53b). This observation is not surprising: the ground state is indeed special - it is the only bound state that exists at any $u > 0$.

As it is the case with large u [see Eqs. (6.51) and (6.52)], it is easy to obtain the leading corrections in Eqs. (6.53) directly from Eq. (6.27). Since at small $u - u_n$ the argument of $\tan(k_n a/2)$ is close to $\pi n/2$, we have $\kappa_n/k_n = \tan(k_n a/2) \ll 1$. It is therefore convenient to express all k_n in Eq. (6.27) via κ_n and write this equation as

$$\kappa_n = \sqrt{u - \kappa_n^2}\, \tan\Bigl[(a/2)\sqrt{u - \kappa_n^2}\Bigr]. \tag{6.55a}$$

To obtain $\kappa_n(u)$ in the leading order in small $u - u_n$, we neglect κ_n^2 in the right-hand side and approximate

$$\begin{aligned}
\tan\Bigl[(a/2)\sqrt{u}\Bigr] &= \tan\Bigl\{\pi n/2 + (a/2)\bigl(\sqrt{u} - \sqrt{u_n}\bigr)\Bigr\} \approx (a/2)\bigl(\sqrt{u} - \sqrt{u_n}\bigr) \\
&= (a/2)\sqrt{u}\left\{1 - \left[1 - \frac{u - u_n}{u}\right]^{1/2}\right\} \approx (a/2)\sqrt{u} \times \begin{cases} 1, & n = 0, \\ (u - u_n)/2u, & n > 0. \end{cases}
\end{aligned}$$

With these approximations, Eq. (6.55a) reduces to

$$\kappa_n(u) = \begin{cases} au/2, & n = 0, \\ a(u - u_n)/4, & n > 0. \end{cases}$$

The second equation in Eq. (6.25b) then yields

$$u - \varepsilon_n(u) = \kappa_n^2(u) = \begin{cases} (au/2)^2, & n = 0, \\ \bigl[a(u - u_n)/4\bigr]^2, & n > 0, \end{cases} \tag{6.55b}$$

in agreement with Eqs. (6.53).

Note that the above result for the ground state energy can be also written as

$$\varepsilon_0 - v_\infty = -\kappa_0^2 \tag{6.56}$$

with v_∞ and κ_0 defined in Eqs. (6.18). It turns out that Eq. (6.56) applies to all potentials that satisfy the conditions (6.18). The right-hand side of Eq. (6.56) is the leading term in the asymptotic expansion in powers of a small dimensionless parameter $|a\kappa_0| \ll 1$, where a is now the size of the region where $v_\infty - v(x)$ differs appreciably from zero; its precise value depends on the functional form of $v(x)$.

6.2.4 Narrow- or shallow-well limit: delta-well

Consider a rectangular potential well described by the potential

$$v(x) = \frac{2m}{\hbar^2} V(x) = \begin{cases} 0, & |x| > a/2, \\ -u, & |x| < a/2. \end{cases} \tag{6.57}$$

This potential is obtained by subtracting u from that given by (6.19). Therefore, the consideration of Sec. 6.2.3 remains applicable; the only difference is that one needs to subtract u from the bound state energies found in Sec. 6.2.3.

In this section we focus on the limit

$$a \to +0, \quad u \to +\infty \tag{6.58a}$$

taken under the constraint that au is kept at fixed value

$$au = 2\kappa_0, \tag{6.58b}$$

where the parameter κ_0 is defined as in Eq. (6.18b); it has units of (length)$^{-1}$. In the limit considered $a^2u = 2\kappa_0 a \to +0$, hence [recall Eqs. (6.44)] the well supports only one bound state. Moreover, the same small parameter a^2u controls the higher-order terms in the expansion

$$\varepsilon(u) = -\frac{1}{4}(au)^2 f_0(a^2u) \tag{6.59a}$$

[see Eq. (6.53a)]. In the limit (6.58) we have $f_0(a^2u) \to 1$, and Eq. (6.59a) yields the *exact* bound state energy

$$\varepsilon = -(au/2)^2 \longrightarrow -\kappa_0^2. \tag{6.59b}$$

Eq. (6.59a) was obtained in Sec. 6.2.2 by a rather laborious analysis of the ground state energy at small a^2u. We will show now that it is much easier to derive Eq. (6.59b) by taking the limit (6.58) from the outset. In this limit, the rectangular well potential (6.57) turns to a delta-function,

$$v(x) = -2\kappa_0 \delta(x). \tag{6.60}$$

Since this ***delta-well*** potential does not conform with the conditions listed in the beginning of Sec. 6.1.2, it should be handled with care. It is easy to see that the potential (6.60) is not sufficiently singular to cause a discontinuity, let alone a divergence, in the eigenfunctions. It is however singular enough to cause discontinuities in their derivatives. Indeed, integrating both sides of Eq. (6.3b) over the interval $-b < x < b$, we get

$$\psi'(+b) - \psi'(-b) = \int_{-b}^{+b} dx\, [v(x) - \varepsilon]\psi(x).$$

Substituting here the delta-well potential (6.60) and taking the limit $b \to +0$, we obtain

$$\psi'(+0) - \psi'(-0) = -2\kappa_0 \psi(0). \tag{6.61a}$$

It is convenient to divide both sides of this relation by $\psi(0)$ [we will see shortly below that $\psi(0) \neq 0$] and write

$$\left.\frac{\psi'(x)}{\psi(x)}\right|_{x\to+0} - \left.\frac{\psi'(x)}{\psi(x)}\right|_{x\to-0} = -2\kappa_0. \tag{6.61b}$$

Eq. (6.61b) applies to all eigenfunctions (not only the bound state), and can be interpreted as a boundary condition: one needs to solve Eq. (6.3b) in the potential-free regions $|x| > 0$ and then match these solutions at $x = 0$ using Eq. (6.61b).

Taking into account that the bound state has negative energy $\varepsilon = -\kappa^2 < 0$, we write Eq. (6.3b) at $|x| > 0$ as

$$\psi'' = \kappa^2 \psi. \tag{6.62a}$$

Solution of this equation that vanishes at $x \to \pm\infty$ obviously has the form

$$\psi(x) \propto \begin{cases} e^{\kappa x}, & x < 0, \\ e^{-\kappa x}, & x > 0 \end{cases} \tag{6.62b}$$

[cf. the second equation in (6.25a)], or, with the continuity of $\psi(x)$ at $x = 0$ taken into account,[92]

$$\psi(x) \propto e^{-\kappa|x|}. \tag{6.62c}$$

Substitution of Eq. (6.62b) into the matching condition (6.61b) immediately yields

$$\kappa = \kappa_0, \tag{6.63a}$$

which gives

$$\varepsilon = -\kappa^2 = -\kappa_0^2 \tag{6.63b}$$

for the bound state energy, in agreement with Eq. (6.59b).

This result can be also obtained by substituting Eq. (6.62c) into the equation

$$-\psi'' - 2\kappa_0 \delta(x)\psi(x) = \varepsilon\psi(x). \tag{6.64a}$$

[see Eqs. (6.3b) and (6.60)]. Indeed, we have

$$\frac{d}{dx} e^{-\kappa|x|} = -\kappa \left[\frac{d}{dx}|x|\right] e^{-\kappa|x|} = -\kappa \operatorname{sign}(x) e^{-\kappa|x|},$$

hence [see Eqs. (5.2) and (5.5)]

$$\frac{d^2}{dx^2} e^{-\kappa|x|} = \big[-\kappa \operatorname{sign}(x)\big]^2 e^{-\kappa|x|} - \kappa \left[\frac{d}{dx} \operatorname{sign}(x)\right] e^{-\kappa|x|} = \kappa^2 e^{-\kappa|x|} - 2\kappa\delta(x).$$

Substituting this expression into Eq. (6.64a) and grouping terms of similar functional form, we obtain

$$-(\varepsilon + \kappa^2) e^{-\kappa|x|} + 2(\kappa - \kappa_0)\delta(x) = 0. \tag{6.64b}$$

This equation is satisfied at all x if and only if both terms in the left-hand side vanish identically. This requirement leads to the equations

$$\varepsilon + \kappa^2 = 0, \quad \kappa - \kappa_0 = 0, \tag{6.64c}$$

from which Eq. (6.63b) follows at once.

[92] The normalized bound state wave function (6.62c) reads $\psi(x) = \sqrt{\kappa} e^{-\kappa|x|}$. It is not necessary to keep track of the normalization coefficient if one's goal is merely to find the bound state energy.

Of course, as it is always the case with simple exactly solvable problems, there are numerous other methods of obtaining these results. In particular, instead of solving the differential equation for $\psi(x)$, one can derive an integral equation for the momentum-space wave function $\psi(p)$. For the delta-well potential (6.60) this integral equation turns out to be so simple that its solution does not take much longer than the position-space derivations described above.

6.3 Case study: harmonic oscillator

Here we consider a rather common situation when the potential $V(x)$ in Eq. (6.1) is a smooth function with a minimum at some x, say, at $x=0$. Any such potential can be expanded in Taylor series

$$V(x) = V(0) + \frac{1}{2}V''(0)\,x^2 + \dots \tag{6.65a}$$

with $V''(0) > 0$. To understand the low-energy motion in this potential, it is sufficient to truncate this series at second order and write

$$V(x) \approx \frac{1}{2}V''(0)\,x^2, \tag{6.65b}$$

where we have omitted the constant term $V(0)$.

6.3.1 Classical oscillator

For the potential (6.65b) the classical equation of motion $m\ddot{x} = -V'(x)$ assumes the form

$$m\ddot{x} = -V''(0)\,x. \tag{6.66a}$$

The general solution of this equation can be written as

$$x(t) = x_0\cos(\omega t + \alpha), \tag{6.66b}$$

where

$$\omega = \sqrt{V''(0)/m} \tag{6.66c}$$

is the angular frequency, $x_0 = \max\{x(t)\}$ is the amplitude of oscillations, and α is the phase shift. The approximation (6.65b) that led to Eqs. (6.66) is applicable as long as the amplitude x_0 is so small that at $x = x_0$ the difference between the truncated potential (6.65b) and the exact potential can be safely neglected, i.e., at such x_0 that

$$\bigl|V(x_0) - V(0)\bigr| \ll V''(0)\,x_0^2. \tag{6.67}$$

Because for a classical motion there is no lower bound on x_0, the small amplitude regime when this condition is satisfied is guaranteed to exist.

It is convenient to express $V''(0)$ in the potential (6.65b) via frequency ω [see Eq. (6.66c)] and write the energy of the oscillator as

$$E = \frac{p^2}{2m} + \frac{m\omega^2 x^2}{2}. \tag{6.68}$$

The two terms in the right-hand side here are the kinetic and potential energies, respectively. Taking into account Eq. (6.66b) and the relation $p = m\dot{x}$, we write these energies as

$$\begin{aligned} E_{\rm kin}(t) &= \frac{p^2}{2m} = \frac{m}{2}\,\dot{x}^2 = \frac{1}{2}\,m\omega^2 x_0^2 \sin^2(\omega t + \alpha)\,, \\ E_{\rm pot}(t) &= \frac{m\omega^2 x^2}{2} = \frac{1}{2}\,m\omega^2 x_0^2 \cos^2(\omega t + \alpha)\,. \end{aligned} \tag{6.69a}$$

As the particle moves, the energy sloshes back and forth between these two contributions, but their sum $E = E_{\rm kin}(t) + E_{\rm pot}(t) = m\omega^2 x_0^2/2$ remains constant. Moreover, averaging Eqs. (6.69a) over the period of oscillations[93] yields

$$\overline{E_{\rm kin}} = \overline{E_{\rm pot}} = \frac{1}{2}E. \tag{6.69b}$$

This result is an instance of the classical virial theorem.[94]

6.3.2 Quantum oscillator

Promoting E, p, and x in Eq. (6.68) to the corresponding operators, we obtain the Hamiltonian of a ***quantum harmonic oscillator***

$$\hat{H} = \frac{\hat{p}^2}{2m} + \frac{m\omega^2\hat{x}^2}{2}. \tag{6.70}$$

According to the discussion in Sec. 6.1.2, this Hamiltonian has a purely discrete non-degenerate spectrum, and its eigenvectors form a complete orthonormal set,

$$\hat{H}|\psi_n\rangle = E_n|\psi_n\rangle, \quad \langle\psi_m|\psi_n\rangle = \delta_{m,n}, \quad \sum_{n=0}^{n=\infty} |\psi_n\rangle\langle\psi_n| = \hat{\mathbb{1}}. \tag{6.71}$$

The corresponding position-space wave functions $\psi_n(x) = \langle x|\psi_n\rangle$ are solutions of the differential equation

$$\left\{-\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + \frac{m\omega^2 x^2}{2}\right\}\psi_n(x) = E_n\psi_n(x) \tag{6.72a}$$

---

[93] Averaging over the period is defined as $\overline{F} = \dfrac{1}{T}\displaystyle\int_t^{t+T} dt' F(t')$, where $T = 2\pi/\omega$ is the period of oscillations.

[94] Consider a periodic motion of a classical particle with mass m in the potential $V(x)$. Taking into account the relations $p = m\dot{x}$ and $\dot{p} = -V'(x)$, we obtain

$$\frac{d}{dt}(xp) = \dot{x}p + x\dot{p} = m\dot{x}^2 - xV'(x) = 2E_{\rm kin} - xV'(x).$$

On averaging over the period[93] the left-hand side here vanishes, yielding the ***virial theorem***

$$\overline{E_{\rm kin}} = \frac{1}{2}\,\overline{xV'(x)}.$$

For a harmonic oscillator $V(x) \propto x^2$, hence $\frac{1}{2}xV'(x) = V(x) = E_{\rm pot}$, and the theorem reduces to Eq. (6.69b).

vanishing at $x \to \pm\infty$ [cf. Eqs. (6.3a) and (6.7)]. The ratio of the coefficients $\hbar^2/m$ and $m\omega^2$ of the two terms in the left-hand side of this equation has units of (length)^4, giving rise to the length scale[95]

$$x_\omega = \left(\frac{\hbar^2/m}{m\omega^2}\right)^{1/4} = \sqrt{\hbar/m\omega}. \tag{6.72b}$$

It is therefore convenient to rewrite Eq. (6.72a) as

$$\left\{-x_\omega^2 \frac{d^2}{dx^2} + \frac{x^2}{x_\omega^2}\right\}\psi_n(x) = \epsilon_n \psi_n(x), \tag{6.72c}$$

where

$$\epsilon_n = \frac{2E_n}{\hbar\omega} \tag{6.72d}$$

is the eigenenergy in units of $\hbar\omega/2$ (note that ϵ_n so defined is dimensionless).

For a quadratic potential $V \propto x^2$ working in the position space has no apparent advantage. Indeed, it is easy to show with the help of Eqs. (5.43b) and (5.47b) that the momentum-space wave functions $\psi_n(p) = \langle p|\psi_n\rangle$ obey the equation

$$\left\{-\frac{\hbar^2 m\omega^2}{2}\frac{d^2}{dp^2} + \frac{p^2}{2m}\right\}\psi_n(p) = E_n \psi_n(p). \tag{6.73a}$$

This equation can be also written as

$$\left\{-p_\omega^2 \frac{d^2}{dp^2} + \frac{p^2}{p_\omega^2}\right\}\psi_n(p) = \epsilon_n \psi_n(p), \tag{6.73b}$$

where

$$p_\omega = \sqrt{\hbar m\omega} = \hbar/x_\omega \tag{6.73c}$$

is the momentum scale and ϵ_n is the dimensionless eigenenergy defined in Eq. (6.72d).

Comparison of Eqs (6.72c) and (6.73b) shows that the normalized wave functions $\psi_n(x)$ and $\psi_n(p)$ are given by

$$\psi_n(x) = \frac{1}{\sqrt{x_\omega}}\varphi_n(x/x_\omega), \quad \psi_n(p) = \frac{1}{\sqrt{p_\omega}}\varphi_n(p/p_\omega) \quad \text{(up to phase factors)}, \tag{6.74a}$$

where $\varphi_n(t)$ are real dimensionless functions of a dimensionless variable t that satisfy

$$-\varphi_n'' + t^2\varphi_n = \epsilon_n \varphi_n, \quad \int dt\, \varphi_n^2(t) = 1. \tag{6.74b}$$

Note that the phase factors in Eqs. (6.74a) are not completely arbitrary because with our conventions regarding the phases of the position and momentum eigenvectors [see Eqs. (5.45a)

[95] Note that because the emerging length scale x_ω contains $\hbar$, it has no classical analogue. This length scale can be viewed as an estimate of the amplitude of the zero-point motion [see Sec. 6.1.3]. (Indeed, we will see below that $\Delta x \sim x_\omega$.) We expect the classical results discussed in Sec. 6.3.1 to break down at $x_0 \lesssim x_\omega$. Accordingly, for a generic potential $V(x)$ there is no guarantee that the regime in which the approximation (6.65b) is justified [see (6.67)] exists. [In writing Eq. (6.70) we of course assumed that the potential (6.65b) is exact.]

and (5.51b)], the wave functions $\psi_n(x)$ and $\psi_n(p)$ are Fourier transforms of one another [recall Eqs. (5.53)].

It follows immediately from Eqs. (6.74) that expectation values of functions of x and p in nth eigenstate of $\hat{H}$ are related to each other as

$$\langle f(x/x_\omega)\rangle_n = \langle f(p/p_\omega)\rangle_n, \tag{6.75a}$$

where $f(\xi)$ is an arbitrary function of a dimensionless variable ξ. In particular, for $f(\xi) = \xi^2$ Eq. (6.75a) gives

$$\langle (x/x_\omega)^2\rangle_n = \langle (p/p_\omega)^2\rangle_n. \tag{6.75b}$$

Taking into account the relations

$$\hat{p}^2/m = \hbar\omega(\hat{p}/p_\omega)^2, \quad m\omega^2\hat{x}^2 = \hbar\omega(\hat{x}/x_\omega)^2 \tag{6.76a}$$

and Eqs. (6.70) and (6.73c), we obtain the quantum analog of Eq. (6.69b),

$$\langle E_{\rm kin}\rangle_n = \langle E_{\rm pot}\rangle_n = \frac{1}{2}E_n. \tag{6.76b}$$

Moreover, since the wave functions (6.74a) have definite parities, the expectation values $\langle x\rangle_n$ and $\langle p\rangle_n$ vanish identically, and Eqs. (6.75b) and (6.76) yield

$$\Delta_n x\, \Delta_n p = E_n/\omega \tag{6.76c}$$

for the product of uncertainties of x and p in nth eigenstate of the oscillator.

The first step in solving equations such as Eq. (6.74b) is to understand the behavior of their solutions at large $|t|$. Substituting $\varphi_n(t) \propto e^{\sigma(t)}$ into Eq. (6.74b), we obtain the equation for $\sigma(t) = \ln\varphi_n(t)$,

$$(\sigma')^2 + \sigma'' = t^2 - \epsilon_n. \tag{6.77a}$$

At large $|t| \gg \sqrt{\epsilon_n}$ one can safely neglect ϵ_n in the right-hand side of this equation. As for the left-hand side, the standard trick is to assume that one of the two terms is much larger than the other, solve the equation under this assumption, and then check whether the assumption is correct. If $|\sigma''| \ll (\sigma')^2$ the equation reduces to

$$(\sigma')^2 = t^2, \tag{6.77b}$$

which has two solutions, $\sigma(t) = \pm t^2/2$. (Since $|t|$ is large, there is no need to worry about the integration constant.) Both solutions satisfy $|\sigma''| \ll (\sigma')^2$, hence the assumption we made is correct.[96] The two solutions we found, $\varphi_n(t) \propto e^{\pm t^2/2}$, are the two linearly independent solutions of Eq. (6.74b). The solution we need is the one that vanishes at $t \to \pm\infty$, i.e.,

$$\varphi_n(t)\big|_{\text{large } |t|} \propto e^{-t^2/2}. \tag{6.77c}$$

[96] The assumption $(\sigma')^2 \ll |\sigma''|$ would lead to the equation $\sigma'' = t^2$ instead of Eq. (6.77b). This equation yields $\sigma' \propto t^3$, so that $(\sigma')^2 \propto t^6 \gg \sigma''$, in contradiction with the assumption.

The simplest normalized function behaving as prescribed by Eq. (6.77c) is the Gaussian

$$\varphi_0(t) = \pi^{-1/4} e^{-t^2/2}. \tag{6.78a}$$

Direct substitution shows that this function indeed satisfies Eqs. (6.74b) with the dimensionless eigenenergy

$$\epsilon_0 = 1. \tag{6.78b}$$

Since $\varphi_0(t)$ has no zeros at finite t, it represents the exact ground state ($n = 0$) wave function. Choosing the phase factor of the position-space wave function $\psi_0(x)$ in Eq. (6.74a) so that it is real and positive, we obtain[97]

$$\psi_0(x) = \frac{1}{\sqrt{x_\omega}} \varphi_0(x/x_\omega), \quad \psi_0(p) = \frac{1}{\sqrt{p_\omega}} \varphi_0(p/p_\omega). \tag{6.78c}$$

As expected for a Gaussian, the ground state is a minimal uncertainty state [see Sec. 5.2.5]. Indeed, comparison of Eqs. (6.78c) with Eqs. (5.63) shows that in this state $\Delta_0 x = x_\omega/\sqrt{2}$ and $\Delta_0 p = p_\omega/\sqrt{2}$. Since $x_\omega p_\omega = \hbar$ [see Eq. (6.73c)], this gives $\Delta_0 x \, \Delta_0 p = \hbar/2$, in agreement with Eq. (6.76c).

Evaluation of the remaining wave functions is straightforward, although somewhat tedious. Eqs. (6.77c) and (6.78a) suggest seeking $\varphi_n(t)$ in the form

$$\varphi_n(t) = h_n(t) \varphi_0(t). \tag{6.79}$$

Substitution into the first equation in (6.74b) yields an equation for the function $h_n(t)$. By analyzing this equation, one can determine ϵ_n and $h_n(t)$ for all n [the functions $h_n(t)$ turn out to be polynomials]. Rather than following this route, we discuss below a different, purely algebraic, solution of the eigenvalue problem.

6.3.3 Algebraic solution

It is convenient to write the Hamiltonian (6.70) as

$$\hat{H} = \frac{\hbar\omega}{2} \left[(\hat{x}/x_\omega)^2 + (\hat{p}/p_\omega)^2 \right] \tag{6.80}$$

[see Eqs. (6.76a)], where x_ω and p_ω are the length and momentum scales defined in Eqs. (6.72b) and (6.73c), respectively.

The right-hand side of Eq. (6.80) is a sum of squares of two non-commuting Hermitian operators $\hat{x}/x_\omega$ and $\hat{p}/p_\omega$. We have encountered a similar situation in Sec. 3.2.2, where we have found it to be advantageous to express $\hat{J}_{\mathbf{x}}^2 + \hat{J}_{\mathbf{y}}^2$ via the non-Hermitian circular components $\hat{J}_{\mathbf{x}} \pm i\hat{J}_{\mathbf{y}}$. In the same spirit, we introduce the operators

$$\hat{a} = \frac{1}{\sqrt{2}} \left[(\hat{x}/x_\omega) + i(\hat{p}/p_\omega) \right], \quad \hat{a}^\dagger = \frac{1}{\sqrt{2}} \left[(\hat{x}/x_\omega) - i(\hat{p}/p_\omega) \right] \tag{6.81a}$$

that satisfy

$$[\hat{a}, \hat{a}^\dagger] = \hat{\mathbb{1}}. \tag{6.81b}$$

[97] It follows from Eq. (5.53a) that $\psi_0(p)\big|_{p=0} \propto \int dx \, \psi_0(x)$. For real and positive $\psi_0(x)$, this integral is also real and positive. This observation leads to the second equation in (6.78c).

Substituting

$$\hat{x}/x_\omega = \frac{1}{\sqrt{2}}\big(\hat{a}+\hat{a}^\dagger\big), \quad \hat{p}/p_\omega = -\frac{i}{\sqrt{2}}\big(\hat{a}+\hat{a}^\dagger\big) \tag{6.82}$$

into Eq. (6.80) and using Eq. (6.81b), we obtain

$$\hat{H} = \frac{\hbar\omega}{2}\big(\hat{a}\hat{a}^\dagger+\hat{a}^\dagger\hat{a}\big) = \hbar\omega\big[\hat{n}+(1/2)\hat{\mathbb{1}}\big], \quad \hat{n}=\hat{a}^\dagger\hat{a}. \tag{6.83}$$

The new operator $\hat{n}$ introduced here is a dimensionless Hermitian operator that does not commute with $\hat{a}$ and $\hat{a}^\dagger$,

$$\begin{aligned}
&[\hat{n},\hat{a}] = \hat{a}^\dagger\hat{a}\hat{a} - \hat{a}\hat{a}^\dagger\hat{a} = -[\hat{a},\hat{a}^\dagger]\hat{a} = -\hat{a},\\
&[\hat{n},\hat{a}^\dagger] = \hat{a}^\dagger\hat{a}\hat{a}^\dagger - \hat{a}^\dagger\hat{a}^\dagger\hat{a} = \hat{a}^\dagger[\hat{a},\hat{a}^\dagger] = \hat{a}^\dagger.
\end{aligned} \tag{6.84}$$

Just as $\hat{H}$ [see Eq. (6.71)], this operator has a discrete non-degenerate spectrum,

$$\hat{n}|n\rangle = n|n\rangle, \quad \langle m|n\rangle = \delta_{m,n}, \quad \sum_n |n\rangle\langle n| = \hat{\mathbb{1}}. \tag{6.85a}$$

Eq. (6.83) shows that the eigenvectors $|n\rangle$ introduced here are also eigenvectors of $\hat{H}$,

$$\hat{H}|n\rangle = E_n|n\rangle, \quad E_n = \hbar\omega(n+1/2). \tag{6.85b}$$

Consider now vector

$$|\phi_n\rangle = \hat{a}|n\rangle. \tag{6.86a}$$

Acting with $\hat{n}$ on this vector and using Eq. (6.84) and the first equation in (6.85a), we obtain

$$\hat{n}|\phi_n\rangle = \hat{n}\hat{a}|n\rangle = \big(\hat{a}\hat{n}+[\hat{n},\hat{a}]\big)|n\rangle = \hat{a}(\hat{n}-\hat{\mathbb{1}})|n\rangle = (n-1)\hat{a}|n\rangle = (n-1)|\phi_n\rangle.$$

Comparison with the first equation in (6.85a) then shows that $|\phi_n\rangle \propto |n-1\rangle$, i.e.,

$$\hat{a}|n\rangle = c_n|n-1\rangle. \tag{6.86b}$$

Thus, $\hat{a}$ acts as the *lowering* operator for the eigenvalue n of $\hat{n}$. It follows from the first equation in (6.85a) and Eq. (6.86a) that

$$\langle\phi_n|\phi_n\rangle = \langle n|\hat{a}^\dagger\hat{a}|n\rangle = \langle n|\hat{n}|n\rangle = n. \tag{6.87a}$$

On the other hand, Eq. (6.86b) gives

$$\langle\phi_n|\phi_n\rangle = |c_n|^2\langle n-1|n-1\rangle = |c_n|^2. \tag{6.87b}$$

Eqs. (6.87) show that the coefficient $c_n = \langle n-1|\hat{a}|n\rangle$ in Eq. (6.86b) satisfies $|c_n|^2 = n$. Choosing c_n to be real and non-negative (this amounts to specifying the relative phases of eigenvectors with different eigenvalues), we obtain the non-vanishing matrix elements of $\hat{a}$,

$$\langle n-1|\hat{a}|n\rangle = \sqrt{n}. \tag{6.88a}$$

Replacing here n with $n+1$ and taking a complex conjugate, we get the corresponding relation for the *raising* operator $\hat{a}^\dagger$,

$$\langle n+1|\hat{a}^\dagger|n\rangle = \sqrt{n+1}. \tag{6.88b}$$

Eqs. (1.33d), (1.35), and (6.87a) show that the eigenvalues of $\hat{n}$ are bounded from below,[98]

$$n = \langle n|\hat{n}|n\rangle = \langle \phi_n|\phi_n\rangle \geq 0. \tag{6.89a}$$

It is easy to see that this bound is tight, i.e., the smallest possible eigenvalue of $\hat{n}$ is

$$n_0 = \min\{n\} = 0. \tag{6.89b}$$

Indeed, the lowering operator $\hat{a}$ must annihilate $|n_0\rangle$ (otherwise, it would be possible to get $n < n_0$). Accordingly, $|\phi_{n_0}\rangle = \hat{a}|n_0\rangle = |\text{null}\rangle$. Eq. (6.87a) then gives $n_0 = \langle \phi_{n_0}|\phi_{n_0}\rangle = 0$, in agreement with Eq. (6.89b).

By acting repeatedly with the lowering operator $\hat{a}$ on the eigenvector $|n\rangle$, one must reach the eigenvector with the smallest possible eigenvalue n, i.e., $|0\rangle$. Since each $\hat{a}$ lowers n by 1, this observation implies that eigenvalues of $\hat{n}$ are non-negative integers,

$$n = 0,1,2,3,\ldots \tag{6.90}$$

Conversely, eigenvector $|n\rangle$ can be obtained by acting n times with the raising operator $\hat{a}^\dagger$ on the ground state $|0\rangle$. Taking into account Eq. (6.88b), we obtain the relation

$$|n\rangle = \frac{1}{\sqrt{n!}}(\hat{a}^\dagger)^n|0\rangle. \tag{6.91}$$

The integers (6.90) also serve as the labels for the eigenenergies E_n in Eq. (6.85b); notice that these labels conform with our convention [see Eq. (6.16a)].

The formalism of the raising/lowering operators allows one to evaluate various expectation values in a straightforward manner. For example, using Eqs. (6.82) and (6.88), we find

$$\langle x\rangle_n = \frac{x_\omega}{\sqrt{2}}\langle n|(\hat{a}+\hat{a}^\dagger)|n\rangle = 0, \quad \langle p\rangle_n = -i\frac{p_\omega}{\sqrt{2}}\langle n|(\hat{a}-\hat{a}^\dagger)|n\rangle = 0, \tag{6.92a}$$

as expected [see Sec. 6.3.2]. Similarly, taking into account Eqs. (6.84) and (6.85a), we obtain

$$\langle (x/x_\omega)^2\rangle_n = \langle (p/p_\omega)^2\rangle_n = \frac{1}{2}\langle n|(\hat{a}\hat{a}^\dagger + \hat{a}^\dagger\hat{a})|n\rangle = \frac{1}{2}\langle n|(2\hat{n}+\hat{\mathbb{1}})|n\rangle = n+1/2. \tag{6.92b}$$

These equations yield $\Delta_n x\,\Delta_n p = (n+1/2)\hbar = E_n/\omega$, in agreement with Eq. (6.76c).

6.3.4 Eigenfunctions

As discussed in Sec. 6.3.3, the ground state eigenvector $|0\rangle$ satisfies the equation

$$\hat{a}|0\rangle = |\text{null}\rangle. \tag{6.93a}$$

To find the corresponding position-space wave function $\psi_0(x) = \langle x|0\rangle$, we substitute here $\hat{a}$ from Eq. (6.81a) and multiply both sides of the resulting equation by $\langle x|$. Taking into account Eqs. (5.40c), (5.47b), and (6.73c), we obtain the first-order differential equation

$$\left\{x_\omega\frac{d}{dx} + \frac{x}{x_\omega}\right\}\psi_0(x) = 0. \tag{6.93b}$$

[98] Notice that the bound $n \geq 0$ [see (6.89)] is stronger than $n > -1/2$ implied by the relation $E_n > 0$ and the second equation in Eq. (6.85b).

As in Sec. 6.3.2 [see Eqs. (6.74a) and (6.78c)], it is convenient to write the wave function as

$$\psi_0(x) = \frac{1}{\sqrt{x_\omega}}\,\varphi_0(x/x_\omega), \tag{6.94a}$$

where $\varphi_0(t)$ is a normalized solution of the dimensionless equation

$$\left(\frac{d}{dt} + t\right)\varphi_0(t) = 0. \tag{6.94b}$$

A real and positive solution of this equation coincides with Eq. (6.78a),

$$\varphi_0(t) = \pi^{-1/4}\, e^{-t^2/2}. \tag{6.94c}$$

The position- and the momentum-space wave functions of the excited states can now be found with the help of Eq. (6.91). Taking into account Eq. (6.78c), we obtain [99]

$$\psi_n(x) = \frac{1}{\sqrt{n!}}\,\langle x|\,(\hat a^\dagger)^n|0\rangle = \frac{1}{\sqrt{2^n n!}}\left\{-x_\omega\frac{d}{dx} + \frac{x}{x_\omega}\right\}^n \psi_0(x) = \frac{1}{\sqrt{x_\omega}}\,\varphi_n(x/x_\omega), \tag{6.95a}$$

$$\psi_n(p) = \frac{1}{\sqrt{n!}}\,\langle p|\,(\hat a^\dagger)^n|0\rangle = \frac{(-i)^n}{\sqrt{2^n n!}}\left\{-p_\omega\frac{d}{dp} + \frac{p}{p_\omega}\right\}^n \psi_0(p) = \frac{(-i)^n}{\sqrt{p_\omega}}\,\varphi_n(p/p_\omega) \tag{6.95b}$$

[cf. Eqs. (6.74a)], in agreement with Eqs. (5.53). The functions $\varphi_n(t)$ in these equations are given by

$$\varphi_n(t) = \frac{(-1)^n}{\sqrt{2^n n!}}\left(\frac{d}{dt} - t\right)^n \varphi_0(t). \tag{6.96a}$$

Using the identity

$$\left(\frac{d}{dt} - t\right)^n f(t) = e^{\,t^2/2}\,\frac{d^n}{dt^n}\,e^{-t^2/2} f(t)$$

easily verified by induction, one can write these functions as

$$\varphi_n(t) = \frac{1}{\sqrt{2^n n!}}\,H_n(t)\varphi_0(t). \tag{6.96b}$$

Here $H_n(t)$ are the well-known ***Hermite polynomials*** defined by the relation

$$H_n(t) = (-1)^n\, e^{\,t^2}\,\frac{d^n}{dt^n}\,e^{-t^2}. \tag{6.96c}$$

The first three of these polynomials are

$$H_0(t) = 1, \quad H_1(t) = 2t, \quad H_2(t) = 4t^2 - 2. \tag{6.97a}$$

It is easy to see that

$$H_n(t) = (2t)^n + (\cdots)\,t^{n-1} + \ldots, \tag{6.97b}$$

so that $H_n(t)$ and thus $\varphi_n(t)$ are real and positive at large positive t.[99]

[99] The phases of the wave functions (6.95) are dictated by the choices of the phases of $c_n = \langle n-1|\hat a|n\rangle$ and $\varphi_0(t)$ made in Eqs. (6.88a) and (6.94c), respectively.

It is obvious that Hermite polynomials have definite parities, $H_n(t) = (-1)^n H_n(-t)$. It is not at all obvious (yet true) that all n roots of the equation $H_n(t) = 0$ are real. Using Eq. (6.96c) and the relations

$$e^{t^2}\frac{d}{dt}f(t) = \frac{d}{dt}\left[e^{t^2}f(t)\right] - 2te^{t^2}f(t), \qquad \left(\frac{d^2}{dt^2} + 2t\frac{d}{dt} + 2n\right)\frac{d^{n-1}}{dt^{n-1}}\,e^{-t^2} = 0,$$

it is easy to show that Hermite polynomials satisfy

$$\frac{d}{dt}H_n(t) - 2nH_{n-1}(t) = 0, \tag{6.98a}$$

$$H_{n+1}(t) - 2tH_n(t) + 2nH_{n-1}(t) = 0, \tag{6.98b}$$

$$\frac{d^2}{dt^2}H_n(t) - 2t\frac{d}{dt}H_n(t) + 2nH_n(t) = 0. \tag{6.98c}$$

to be continued...