Chapter 2
The Principles of Quantum Mechanics

This introduction describes the mathematical tools used in the formulation of quantum mechanics and the connections between the physical world and mathematical formalism. Such links constitute the fundamental principles of quantum mechanics.[1] They are valid for every specific realization of these principles. Subsequently, their most immediate consequences are presented. Frequent shortcomings existing in many introductions can thus be avoided.[2]

2.1 Classical Physics

If our vision of a moving object is interrupted by a large billboard and is resumed after the reappearance of the object, we naturally assume that it has traveled all the way behind the billboard (Fig. 2.1). This is implied in the notion of physical reality, one of the postulates in the famous EPR paradox written by Einstein in collaboration with Boris Podolsky and Nathan Rosen [16]: "If, without in any way disturbing a system, we can predict with certainty the value of a physical quantity, then there is an element of physical reality corresponding to this physical quantity."

This classical framework relies on the acceptance of some preconceptions, most notably the existence of the continuous functions of time called trajectories $\mathbf{x}(t)$ [or $\mathbf{x}(t)$ and $\mathbf{p}(t)$, where $\mathbf{p}$ is the momentum of the particle]. A trajectory describes the

---

[1] Presentations of quantum mechanics resting upon few basic principles start with [15], which remains a cornerstone on the subject.

[2] In many presentations it is assumed that the solution of any wave equation for a free particle is the plane wave $\exp[\mathrm{i}(kx - \omega t)]$. Subsequently, the operators corresponding to momentum and energy are manipulated to obtain an equation yielding the plane wave as the solution. This procedure is not very satisfactory because (a) plane waves display some difficulties as wave functions, not being square integrable; (b) quantum mechanics appears to be based on arguments that are only valid within a differential formulation; (c) it leads to the misconception that the position wave function is the only way to describe quantum states.

D. Bes, *Quantum Mechanics*, Graduate Texts in Physics,

DOI 10.1007/978-3-642-20556-9_2,

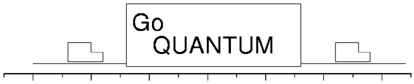

Fig. 2.1 The trajectory of the car behind the billboard as an element of physical reality

motion of a point particle, an abstract concept that does not exist in nature. The concept of trajectory provides an important link between the physical world and its mathematical description. For instance, it allows us to formulate Newton's second law,

$$\mathbf{F}(\mathbf{x}) = \frac{\mathrm{M}\,\mathrm{d}^2\,\mathbf{x}}{\mathrm{d}\,t^2}\,. \tag{2.1}$$

This equation of motion predicts the evolution of the system in a continuous and deterministic way. Since this is a second order equation, the state of a system is specified if the position and the momentum of each particle are known at any one time.

Maxwell's theory of electromagnetism is also part of classical physics. It is characterized in terms of fields, which must be specified at every point in space and time. Unlike particles, fields can be made as small as desired. Electromagnetism is also a deterministic theory.

Essential assumptions in classical physics about both particles and fields are:

- The possibility of non-disturbing measurements
- There is no limit to the accuracy of the values assigned to physical properties

In fact, there is no distinction between physical properties and the numerical values they assume. Schwinger characterizes classical physics as "the idealization of non-disturbing measurements and the corresponding foundations of the mathematical representation, the consequent identification of physical properties with numbers, because nothing stands in the way of the continual assignment of numerical values to these physical properties" [14], p. 11.

Such "obvious" assumptions are no longer valid in quantum mechanics. Therefore, other links have to be created between the physical world and mathematical formalism.

2.2 Mathematical Framework of Quantum Mechanics

According to classical electromagnetism, an inhomogeneous magnetic field **B** directed along an axis (for instance, the z-axis) should bend the trajectory of particles perpendicular to this axis. The amount of bending of these tiny magnets will be proportional to the projection μ_z of their magnetic moment $\boldsymbol{\mu}$. Therefore, if the beam is unpolarized (all values $(-|\boldsymbol{\mu}| \leq \mu_z \leq |\boldsymbol{\mu}|)$ are present), particles are classically expected to impact over a continuous region of a screen. However, it was shown in 1921 by Otto Stern and Walther Gerlach that silver atoms distribute themselves over each of two lines. Since the magnetic moment is proportional to an intrinsic angular momentum called spin ($\boldsymbol{\mu} \propto \mathbf{S}$), it is apparent that only two projections of the spin are allowed by nature [17] (for more details see Sect. 5.2.1).

Note, however, that the fact that a physical quantity may have only two values does not require by itself to abandon classical physics. For instance, your PC is made up from bits, i.e. classical systems that may be in one of two states.[3] As befits classical systems, their state is not altered upon measurement (thus contributing to the stability of classical computers).

A different description is provided by vectors on a plane. While the sum of the two states of a bit does not make sense, the addition of two vectors on a plane is always another vector. Any vector Ψ may be written as a linear combination (Fig. 2.2, top)

$$\Psi = c_x \varphi_x + c_y \varphi_y , \tag{2.2}$$

where c_x, c_y are amplitudes and φ_x,φ_y are two perpendicular vectors of module one. This last property is expressed as $\langle \varphi_i | \varphi_j \rangle = \delta_{ij}$, which is a particular case of the scalar product

$$\langle \Psi | \Psi' \rangle = c_x^* c_x' + c_y^* c_y' . \tag{2.3}$$

In quantum mechanics we allow complex values of the amplitudes. Thus, the definition of normalized vectors becomes

$$\langle \Psi | \Psi \rangle = |c_x|^2 + |c_y|^2 = 1 . \tag{2.4}$$

Another crucial property of the chosen vector space is that the same vector Ψ may be expressed as a combination of other sets of perpendicular vectors χ_x, χ_y along rotated axes (Fig. 2.2, bottom)

$$\begin{aligned} \Psi &= b_x \chi_x + b_y \chi_y , \\ 1 &= |b_x|^2 + |b_y|^2 . \end{aligned} \tag{2.5}$$

This two-dimensional space may be easily generalized to spaces with any number of dimensions, called Hilbert spaces. Here we outline some properties that are specially relevant from the point of view of quantum mechanics. This overview is expanded in Sect. 2.7*.

- As in (2.2), any vector Ψ may be expressed as a linear combination of orthonormal basis states φ_i:

$$\Psi = \sum_i c_i \varphi_i \; ; \qquad c_i = \langle \varphi_i | \Psi \rangle \equiv \langle i | \Psi \rangle , \tag{2.6}$$

[3] Although the bits in your PC function on the basis of quantum processes (for instance, semiconductivity) they are not used in PCs as quantum systems.

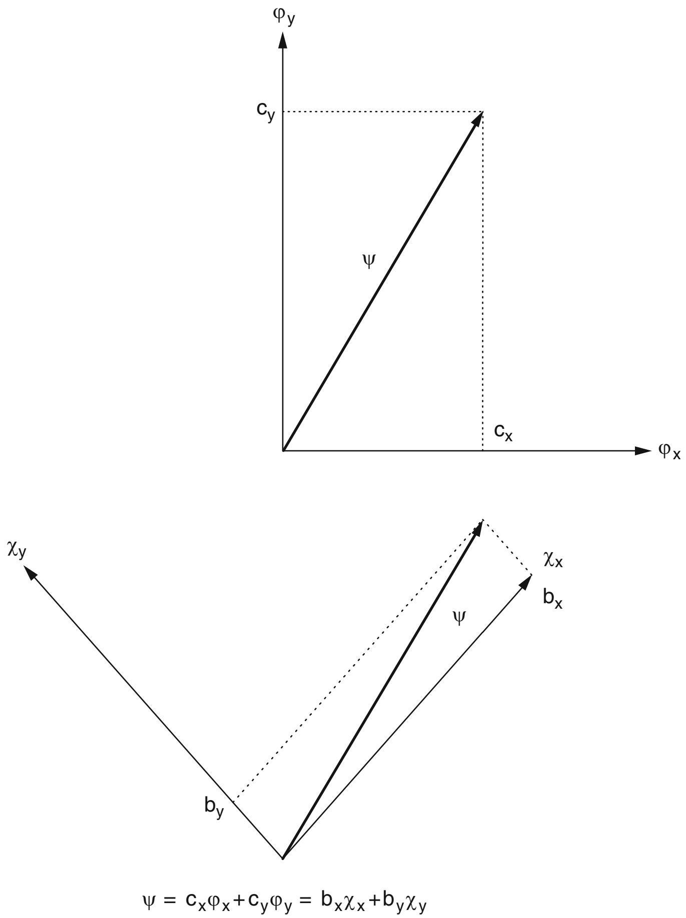

Fig. 2.2 The same vector Ψ can be expressed as the sum of two different systems of basis vectors

$$\delta_{ij} = \langle \varphi_i | \varphi_j \rangle. \tag{2.7}$$

- Linear operators $\hat{Q}$ act on vectors belonging to a Hilbert space, transforming one vector into another: $\hat{Q}\,\Psi = \Phi$. These operators obey a non-commutative algebra, as shown in Sect. 2.7* for the case of rotations in three dimensions. We define the commutation operation through the symbol

$$[\hat{Q}, \hat{R}] \equiv \hat{Q}\,\hat{R} - \hat{R}\,\hat{Q}\,, \tag{2.8}$$

where the order of application of the operators is from right to left ($\hat{Q}\hat{R}\,\Psi = \hat{Q}\,(\hat{R}\Psi)$.

- If the vector $\hat{Q}\varphi_i$ is proportional to φ_i, then φ_i is said to be an eigenvector of the operator $\hat{Q}$. The constant of proportionality q_i is called the eigenvalue

$$\hat{Q}\,\varphi_i = q_i\,\varphi_i\,. \tag{2.9}$$

- The scalar product between a vector $\hat{Q}\Psi_a$ and another vector Ψ_b is called the matrix element of the operator $\hat{Q}$ between the vectors Ψ_a and Ψ_b, and it is symbolically represented[4] as

$$\langle\Psi_b|\hat{Q}\Psi_a\rangle \equiv \langle\Psi_b|Q|\Psi_a\rangle \equiv \langle b|Q|a\rangle\,. \tag{2.10}$$

  The matrix element is said to be diagonal if the same vector appears on both sides of the matrix element ($\langle\Psi|Q|\Psi\rangle$). The matrix elements of the unit operator are the scalar products $\langle\Psi_a|\Psi_b\rangle \equiv \langle a|b\rangle = \langle\Psi_b|\Psi_a\rangle^*$. The norm $\langle\Psi|\Psi\rangle^{1/2}$ is a real, positive number.
- The Hermitian conjugate $\hat{Q}^+$ of an operator $\hat{Q}$ is defined through the relation

$$\langle\varphi_b|Q^+|\varphi_a\rangle = \langle\varphi_a|Q|\varphi_b\rangle^*\,. \tag{2.11}$$

  The operator is said to be Hermitian if

$$\hat{Q}^+ = \hat{Q}\,. \tag{2.12}$$

  The eigenvalues q_i of a Hermitian operator are real numbers and the corresponding eigenvectors φ_i constitute a set of basis vectors.
- The matrix $\mathcal{U}$ with matrix elements $\mathcal{U}_{ab}$ is said to be unitary if the matrix elements of the inverse are given by

$$(\mathcal{U}^{-1})_{ab} = \mathcal{U}^*_{ba}\,. \tag{2.13}$$

  Unitary transformations preserve the norm of the vectors and relate two sets of basis states (see Fig. 2.2)

$$\chi_a = \sum_i \mathcal{U}_{ai}\varphi_i\,. \tag{2.14}$$

These abstract mathematical tools (vectors, Hermitian operators and unitary transformations) may be represented through concrete, well-known mathematical objects, such as column vectors and matrices (Chap. 3), or by means of functions of the coordinate and differential operators (Chap. 4).

2.3 Basic Principles of Quantum Mechanics

In this section we present the quantum mechanical relation between the physical world and the mathematical tools that have been outlined in the previous section. It is formulated through the following quantum principles:

---

[4]The symbols $|a\rangle$ and $\langle a|$ have been called by Dirac ket and bra, respectively [15].

Principle 1. *The state of the system*[5] *is completely described by a vector Ψ – the state vector or state function – belonging to a Hilbert space.*

The state vector Ψ constitutes an unprecedented way of describing nature. It is an abstract entity that carries information about the results of possible measurements. It replaces the classical concepts of position and momentum in the description of physical systems.

The state vector may be multiplied by an arbitrary complex constant and still represent the same physical state. Even if we enforce the requirement of normalization, an arbitrary overall phase is left, which has no physical significance. This is not the case for the relative phase of the terms in the sum $c_a\Psi_a + c_b\Psi_b$, which encodes important physical information.

The fact that the sum of two state vectors is another state vector belonging to the same Hilbert space, i.e. describing another state of the system, is usually called the superposition principle. The sum $c_a\Psi_a + c_b\Psi_b$ must not be interpreted in the sense that we have a conglomerate of systems in which some of them are in the state Ψ_a and some in the state Ψ_b, but rather that the system is simultaneously in both component states. This statement is also valid when the system is reduced to a single particle.

The weirdness of quantum mechanics can be traced back to this superposition. It is fundamentally different from any property of classical particles, which are never found as a linear combination of states associated with different trajectories: a tossed coin may fall as a head or a tail, but not as a superposition of both.

By establishing that the state vector completely describes the state of the system, Principle 1 assumes that there is no way of obtaining information about the system, unless this information is already present in the state vector.

The state vector may only concern some degree(s) of freedom of the physical system, such as position and momentum of a particle and the spin.

The relation between the physical world and states Ψ is more subtle than the classical relation with position and momentum $\mathbf{x}$, $\mathbf{p}$. It relies on principles 2 and 3.

Principle 2. *To every physical quantity there corresponds a single linear operator. In particular, the operators $\hat{x}$ and $\hat{p}$, corresponding to the coordinate and momentum of a particle, fulfil the commutation relation*

$$[\hat{x}, \hat{p}] = \mathrm{i}\hbar\,, \tag{2.15}$$

The commutator is defined in (2.8) and $\hbar$ is Planck's constant h divided by 2π (Table A.1).[6] The constant $\hbar$ provides an estimate of the domain in which quantum

---

[5]This notion of the state of a system is close to the one appearing in classical thermodynamics. It is applied there to many-body systems in which the path of the constituents cannot be traced in practice (for instance, molecules in a gas). However, in quantum mechanics the concept of state is applied even in the case of a single particle.

[6]This commutation relation is related to the classical Poisson bracket (Sect. 2.6.3). It has been derived from relativistic invariance [18], using the fact that spatial translations [generated by $\hat{p}$, see (4.10)] do not commute with Lorentz transformations even in the limit $c \to \infty$.

mechanics becomes relevant. It has the dimensions of classical action (energy × time). Classical physics should be applicable to systems in which the action is much larger than $\hbar$.

This is also fundamentally different from classical physics, for which physical properties are identified with (commuting) numbers (Sect. 2.1).

Since any classical physical quantity may be expressed as a function of coordinate and momentum ($Q = Q(x, p)$), the replacement $x \to \hat{x}$ and $p \to \hat{p}$ in the classical expression $Q(x, p)$ yields the operator $\hat{Q} = Q(\hat{x}, \hat{p})$. The requirement of hermiticity is usually sufficient to account for the non-commutativity of operators [for instance, $xp \leftrightarrow \frac{1}{2}(\hat{x}\hat{p} + \hat{p}\hat{x})$]. Thus, a one-to-one correspondence between physical quantities or observables Q and operators $\hat{Q}$ is established. However, there are also purely quantum operators, such as the spin operators, that cannot be obtained through such substitution.

The operator corresponding to the classical Hamilton function $H(p, x)$ is called the Hamiltonian. For a conservative system

$$\hat{H} = \frac{1}{2M}\hat{p}^2 + V(\hat{x}) , \tag{2.16}$$

where M is the mass of the particle and V is the potential.

Principle 3. *The eigenvalues q_i of an operator $\hat{Q}$ constitute the possible results of the measurements of the physical quantity Q. The probability*[7] *of obtaining the eigenvalue q_i is the modulus squared $|c_i|^2$ of the amplitude of the eigenvector φ_i in the state vector Ψ representing the state of the system.*

Since the results of measurements are real numbers, the operators representing physical observables are restricted to be Hermitian (2.12) and (2.53). In particular, the possible values of the energy E_i are obtained by solving the eigenvalue equation

$$\hat{H}\,\varphi_i = E_i\,\varphi_i . \tag{2.17}$$

As in the case of classical mechanics, quantum mechanics may be applied to very different systems, from single-particle and many-body systems to fields. Thus, quantum mechanics constitutes a framework in which to develop physical theories, rather than a physical theory by itself. A number of simple, typical, well-known problems of a particle moving in a one-dimensional space are discussed in Chaps. 3 and 4.

It is almost as useful to state what the principles do not mean, as to say what they do mean. In the following we list some common misconceptions regarding quantum states [19]:

- "The state vector is similar to other fields present in the physical world." It is fundamentally different from the electric or magnetic fields in electromagnetic

[7] Notions of probability theory are given in Sect. 2.8*.

waves, which carry momentum, energy, etc., and in which any externally caused change propagates at a finite, medium-dependent speed. The state vector does not interact with particles.

- "Energy eigenstates are the only allowed ones." This misconception probably arises from the generalized emphasis on the solution of the eigenvalue equation (2.17) and from its similarity to the correct statement "Energy eigenvalues are the only allowed energies."
- "A state vector describes an ensemble of classical systems." In the standard Copenhagen interpretation, the state vector describes a single system.[8] In none of the acceptable statistical interpretations is the ensemble classical.
- "A state vector describes a single system averaged over some amount of time." The state vector describes a single system at a single instant.

The above three principles are sufficient for the treatment of static situations involving a single particle. Two more principles, concerning many-body systems and the time-evolution of states, are presented in Chaps. 7 and 9, respectively.

2.4 Measurement Process

2.4.1 The Concept of Measurement

In this section we specify some basic concepts involved in the process of measurement.[9]

Two or more systems are in interaction if the presence of one leads to changes in the other, and vice versa. Different initial conditions generally lead to different changes, although this may not always be the case.

A measurement is a process in which a system is put in interaction with a piece of apparatus. The apparatus determines the physical quantity or observable to be measured (length, weight, etc.).

There are two important steps in a measurement. The first is the preparation of the system to be measured, i.e. the determination of the initial state. Bohr's definition of the word "phenomenon" refers to "an observation obtained under specified circumstances, including an account of the whole experimental arrangement" [21], p. 64. This should be contrasted with the EPR definition of physical reality (Sect. 2.1).

The second important step, also crucial in the case of quantum systems, is a (macroscopic) change in the apparatus that should be perceptible by a cognitive system. In many cases this change is produced by a detector at one end of the

[8] See Sect. 12.3.3.

[9] Many definitions included here are extracted from [20].

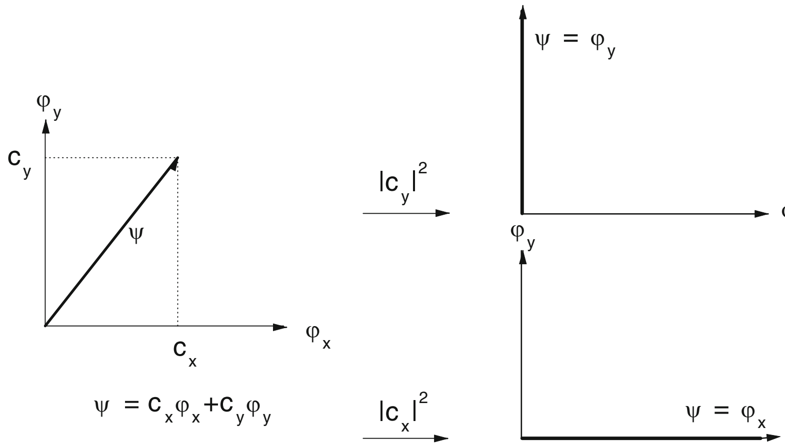

Fig. 2.3 The change in the state vector due to the measurement

apparatus. The magnitude of the physical quantity has a value if the change can be represented in numerical form.

For some systems the interaction with the apparatus does not produce a change in the system or produces a change which is completely determined. This is the case of classical systems (see Sect. 2.1). Our deeply ingrained notion of classical reality has emerged as a consequence of the fact that many independent observers may carry out measurements on the same system without disturbing it (and agree upon the results). On the contrary, this can only be exceptionally achieved in quantum measurements, where a change of the system is generally associated with a change of the apparatus (see Sect. 2.4.2).

2.4.2 Quantum Measurements

The most fundamental difference between a classical and a quantum system is that the latter cannot be measured without being irrevocably altered, no matter how refined the measuring instruments are. This is a consequence of the principles presented in Sect. 2.3.

Assume that the measurement of the physical quantity Q, performed on a system in the state Ψ expanded as in (2.6), yields the result q_j. If the same measurement is repeated immediately afterwards, the same value q_j should be obtained with certainty. Thus, the measurement has changed the previous value of the coefficients $c_i \rightarrow c_i = \delta_{ij}$. In other words, as a result of the measurement, the system "jumps" to an eigenstate of the physical quantity that is being measured (the reduction of the state vector). The only exceptions occur when the initial state is already represented by one of the eigenvectors.

Given an initial state vector Ψ, we do not know in general to which eigenstate the system will jump. Only the probabilities, represented by $|c_i|^2$, are determined (Fig. 2.3). This identification of the probabilities is consistent with the following facts:

- Their value is always positive.
- Their sum is 1 (if the state Ψ is normalized).
- The orthogonality requirement (2.7) ensures that the probability of obtaining any eigenvalue $q_j \neq q_i$ vanishes, if the system is initially represented by an eigenstate φ_i (see Table 2.1).

The fact that, given a state vector Ψ, we can only predict the probability $|c_i|^2$ of obtaining eigenvalues q_i constitutes an indeterminacy inherent to quantum mechanics. Our knowledge about the system cannot be improved, for instance, through a second measurement, since the state Ψ has been transformed into φ_i. The coefficients c_i are also called *probability amplitudes*. The concept of probability implies that we must consider a large number of measurements performed on identical systems, all of them prepared in the same initial state Ψ.

If in the expansion (2.6) there is a subset of basis states φ_k with the same eigenvalue $q_k = q$, the probability of obtaining this eigenvalue is $\sum_k |c_k|^2$. The system is projected after the measurement into the (normalized) state

$$\Psi' = \frac{1}{\sqrt{\sum_k |c_k|^2}} \sum_k c_k \varphi_k. \tag{2.18}$$

So far we have accepted the reduction interpretation of the measurement process without further discussion. This has been, historically, the path followed by most physicists. However, we present one more discussion of the measurement problem in Chap. 14.

The diagonal matrix element is also called the expectation value or mean value of the operator $\hat{Q}$. It is given by the sum of the eigenvalues weighted by the probability of obtaining them

$$\langle\Psi|Q|\Psi\rangle = \sum_i q_i \, |c_i|^2. \tag{2.19}$$

The mean value does not need to be the result q_i of any single measurement, but it is the average value of all the results obtained through the measurement of identical systems.

Measurements of expectation values of non-commuting operators yield the relative phases of the amplitudes (see Sect. 2.6.2).

The uncertainty or standard deviation ΔQ in a given measurement is defined as the square root of the average of the quadratic deviation

$$\begin{aligned} \Delta Q &= \langle\Psi| \left(Q - \langle\Psi|Q|\Psi\rangle\right)^2 |\Psi\rangle^{\frac{1}{2}} \\ &= \left(\langle\Psi|Q^2|\Psi\rangle - \langle\Psi|Q|\Psi\rangle^2\right)^{\frac{1}{2}}, \end{aligned} \tag{2.20}$$

where

$$\langle \Psi | Q^2 | \Psi \rangle = \sum_i q_i^2 \, |c_i|^2. \tag{2.21}$$

We have postulated the existence of links between the physical world and mathematics that are different from those characterizing classical physics. In quantum mechanics, physical quantities are related to (non-commuting) operators; the state vectors are constructed through operations with these mathematical entities; the feedback to the physical world is made by measurements that yield, as possible results, the eigenvalues of the corresponding operators. An example of this two-way connection between formalism and the physical world is the following: assume that the system is constructed in a certain physical state, to which the state vector Ψ is assigned. This assignment is tested by means of various probes, i.e. measurements of observables Q, for which we may know the eigenvector φ_i and, therefore, the probability $|\langle i | \Psi \rangle|^2$ of obtaining the eigenvalues q_i.

This two-way relation between physical world and formalism is not an easy relation [22]:

> "The most difficult part of learning quantum mechanics is to get a good feeling for how the abstract formalism can be applied to actual phenomena in the laboratory. Such applications almost invariably involve formulating oversimplified abstract models of the real phenomena, to which the quantum formalism can effectively be applied. The best physicists have an extraordinary intuition for what features of the actual phenomena are essential and must be represented in the abstract model, and what features are inessential and can be ignored. "

2.5 Some Experimental Consequences of the Basic Principles

2.5.1 Thought Experiments

This section displays some consequences of quantum principles in the form of thought experiments. Alternatively, one may obtain the quantum principles as a generalization of the results of such thought experiments (see [23]).

Let us consider a Hilbert space consisting of only two independent states $\varphi_\pm$. We also assume that these states are eigenstates of an operator $\hat{S}$ corresponding to the eigenvalues ± 1, respectively. Thus the eigenvalue equation $\hat{S}\varphi_\pm = \pm\varphi_\pm$ is satisfied. The scalar products $\langle \varphi_+ | \varphi_+ \rangle = \langle \varphi_- | \varphi_- \rangle = 1$ and $\langle \varphi_+ | \varphi_- \rangle = 0$ are verified. There are many examples of physical observables that may be represented by such operator. For instance, the z-component of the spin[10] is frequently used in these notes (Sects. 2.2, 3.2, 5.2, 9.2, etc.).

[10]Another example is given by the polarization states of the photon (see Sect. 9.8.2†). Most of the two-state experiments are realized by means of such optical devices.

We start by constructing a filter, i.e. an apparatus such that the exiting particles are in a definite eigenstate. In the first part of the apparatus, a beam of particles is split into the two separate $\varphi_\pm$ beams, as in the experiment of Stern and Gerlach (Sect. 5.2.1). In the second part, each beam is pushed towards the original direction. Each separate beam may be masked off at the half-way point. Such an apparatus is sketched in Fig. 2.4a with the φ_- beam masked off. It will be called a φ-filter. It is enclosed within a box drawn with continuous lines.

Any experiment requires first the preparation of the system in some definite initial state. Particles leave the oven in unknown linear combinations Ψ of $\varphi_\pm$ states

$$\Psi = \langle\varphi_+|\Psi\rangle\,\varphi_+ + \langle\varphi_-|\Psi\rangle\,\varphi_- \,. \qquad (2.22)$$

They are collimated and move along the y-axis. In the following cases, we prepare the particles in the filtered state φ_+, by preventing particles in the state φ_- from leaving the first filter (Fig. 2.4b). In the last stage of the experimental set-up we insert another filter as part of the detector, to measure the degree of filtration. The detector includes also a photographic plate which records the arriving particles and is observed by an experimentalist (Fig. 2.4c).

In the first experiment, we place the detector immediately after the first filter (Fig. 2.4d). If the φ_- channel is also blocked in the detector, every particle goes through; if the channel φ_+ is blocked, nothing passes. The amplitudes for these processes are $\langle\varphi_+|\varphi_+\rangle = 1$ and $\langle\varphi_-|\varphi_+\rangle = 0$, respectively. The corresponding probabilities, $|\langle\varphi_+|\varphi_+\rangle|^2$ and $|\langle\varphi_-|\varphi_+\rangle|^2$, also are 1 and 0.

We now consider another set of basis states $\chi_\pm$, also satisfying the orthonormality conditions $\langle\chi_+|\chi_+\rangle = \langle\chi_-|\chi_-\rangle = 1$, $\langle\chi_+|\chi_-\rangle = 0$ (Fig. 2.2, bottom). It is easy to verify that an operator $\hat{R}$, satisfying the eigenvalue equation $\hat{R}\,\chi_\pm = \pm\chi_\pm$, does not commute with $\hat{S}$. Let us perform the necessary modifications of the detector filter so that it can block the particles in either of the states $\chi_\pm$. If $\hat{R}$ corresponds the spin component in the x-direction, the required modification of the detector amounts to a rotation of its filter by an angle $\pi/2$ around the y-axis. Dashed boxes represent filters such that particles exit in the $\chi_\pm$ states (χ-type filters) (Fig. 2.4e).

A particle exiting the first filter in the state φ_+ reorients itself, by chance, within the second filter. This process is expressed within the formalism by expanding the states $\varphi_\pm$ in terms of the new basis set

$$\varphi_\pm = \langle\chi_+|\varphi_\pm\rangle\chi_+ + \langle\chi_-|\varphi_\pm\rangle\chi_- \,. \qquad (2.23)$$

According to Principle 3, a particle emerging from the first filter in the state φ_+ will either emerge from the detector filter in the state χ_+ with probability $|\langle\chi_+|\varphi_+\rangle|^2$ or in the state χ_- with probability $|\langle\chi_-|\varphi_+\rangle|^2$. If the χ_- channel of the second filter is blocked, the particle is either projected into the state χ_+ or is absorbed with probability $1 - |\langle\varphi_+|\chi_+\rangle|^2 = |\langle\varphi_+|\chi_-\rangle|^2$. This result sounds classical: it is the quantum version of the classical Malus law. However, the projection process is probabilistic. Any information about the previous orientation φ_+ is lost.

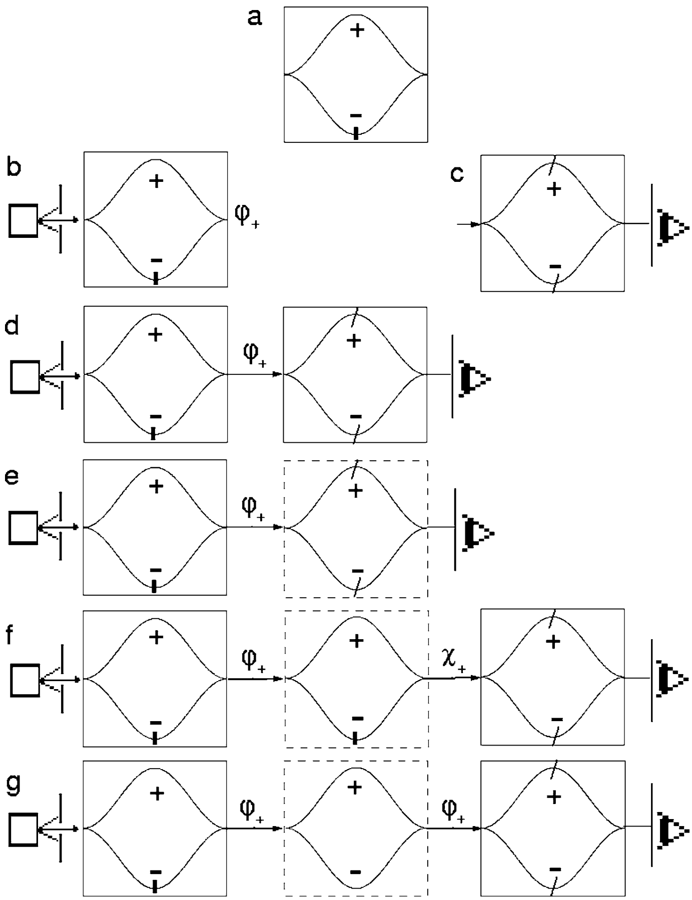

Fig. 2.4 Quantum mechanical thought experiments illustrating the basic principles listed in Sect. 2.3. (**a**) Schematic representation of a filter; (**b**) preparation of the state of a particle; (**c**) detector (filter, photoplate, observer); (**d**)–(**g**) experiments (see text). The *vertical bars* denote fixed path blockings, while the *diagonal bars* indicate paths that can be either opened or closed. For each experiment we perform a measurement in which the upper channel of the detector is open and the down channel blocked, and another measurement with opposite features

We now perform two other experiments which yield results that are spectacularly different from classical expectations. Let us restore the detector filter to the φ-type and introduce a filter of the χ-type between the first filter and the detector (Fig. 2.4f). Thus, particles prepared in the φ_+ state exit the second filter in the χ_+ state. In the spin example, particles leave the first filter with the spin pointing in the direction of the positive z-axis, and the second filter pointing along the positive x-axis. The detector measures the number of particles exiting in one of the $\varphi_\pm$ states (spin pointing up or down in the z-direction). We use now the inverse expansion[11] of (2.23), namely

$$\chi_{\pm} = \langle\varphi_+|\chi_{\pm}\rangle\varphi_+ + \langle\varphi_-|\chi_{\pm}\rangle\varphi_- . \tag{2.24}$$

Thus, the total amplitudes for particles emerging in one of the states $\varphi_{\pm}$ are[12]:

$$\langle\varphi_+|\chi_+\rangle\langle\chi_+|\varphi_+\rangle, \tag{2.25}$$

$$\langle\varphi_-|\chi_+\rangle\langle\chi_+|\varphi_+\rangle . \tag{2.26}$$

Both components $\varphi_{\pm}$ may emerge from the detector filter, in spite of the fact that the fraction of the beam in the φ_- state was annihilated inside the first filter. There is no way in classical physics to explain the reconstruction of the beam φ_-. This example illustrates the quantum rule concerning the impossibility of determining two observables associated with operators which do not commute: a precise determination of R destroys the previous information concerning S.

The result of this experiment is also consistent with Principle 1 in Sect. 2.3, since the state vector χ_+ contains all possible information about the system: its past history is not relevant for what happens to it next. The information is lost because of the blocking mask that has been put inside the second filter.

If we repeat the last experiment, but remove the mask from the second filter (Fig. 2.4g), the total amplitude is given by the sum of the amplitudes associated with the two possible intermediate states

$$\langle\varphi_+|\chi_+\rangle\langle\chi_+|\varphi_+\rangle + \langle\varphi_+|\chi_-\rangle\langle\chi_-|\varphi_+\rangle = \langle\varphi_+|\varphi_+\rangle = 1, \tag{2.27}$$

$$\langle\varphi_-|\chi_+\rangle\langle\chi_+|\varphi_+\rangle + \langle\varphi_-|\chi_-\rangle\langle\chi_-|\varphi_+\rangle = \langle\varphi_-|\varphi_+\rangle = 0 , \tag{2.28}$$

where the closure property has been applied (2.59). All the particles get through in the first case; none in the second case. In going from amplitude (2.26) to (2.28) we get fewer particles, despite the fact that more channels are opened.

It is important that none of the intermediate beams suffers an additional disturbance (for example, the influence of an electric field), which may change the relative phases of the two channels.

The result of this last experiment is equivalent to an interference pattern. Classically, such patterns are associated with waves. However, unlike the case of waves, the particles are always detected as lumps of the same size on a screen placed in front of the exit side of the detector filter. No fractions of a lump are ever detected, as befits the behavior of indivisible particles. Therefore, these experiments display wave–particle duality, which is thus accounted for by Principles 1–3.

---

[11] The amplitudes in (2.23) and (2.24) are related by $\langle\varphi_+|\chi_{\pm}\rangle = \langle\chi_{\pm 1}|\varphi_+\rangle^*$ and $\langle\varphi_-|\chi_{\pm}\rangle = \langle\chi_{\pm 1}|\varphi_-\rangle^*$, according to Table 2.1.

[12] One reads from right to left.

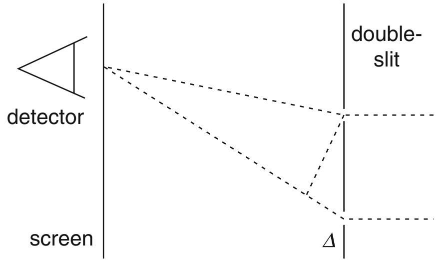

Fig. 2.5 A double-slit experiment. No intensity is detected at points of the screen if the difference Δ between the two paths is an odd multiple of π/k

Feynman has commented this result as follows [23], pp. 1–1: "We choose to examine a phenomenon[13] which is impossible, *absolutely* impossible, to explain in any classical way, and which has in it the heart of quantum mechanics. In reality, it contains the only[14] mystery."

2.5.2 Real Two-Slit Experiments

Thought experiments played a crucial role in the clarification of controversial aspects of quantum mechanics. The discussions between Bohr and Einstein are paradigmatic in this respect (Sect. 15.5.2). However, since the end of the twentieth century, real experiments have replaced thought ones. Not only have earlier views been confirmed, but also more counterintuitive aspects of quantum mechanics have been brought into focus.

Since Thomas Young established the wave nature of light in 1801 (using candles as sources of light), two-slit interference experiments have become crucial to decide between wave and particle behavior (Fig. 2.5).

The usual experimental set-up consists of a source emitting particles and three screens placed on their path. The first screen displays a hole used to collimate the beam. The second one is pierced by two narrow slits placed within the aperture of the beam. The third screen is paved by detectors recording the impact of the particles.

The slits in the intermediate screen define two distinct paths for the particles. Each particle crosses this screen in a linear superposition of two states

$$\Psi = \frac{1}{\sqrt{2}} \left(\varphi_a + \varphi_{a'} \right) . \tag{2.29}$$

The pattern on the detection screen builds up in a pointillist way by accumulation of discrete spots, each one produced by a single particle. It is not possible to predict the final spot for any one particle (only the fringes where no particle will impinge

[13] Matter wave phenomena were experimentally verified for the first time in [5].

[14] Other fundamental issue in quantum physics is entanglement (Chap. 12).

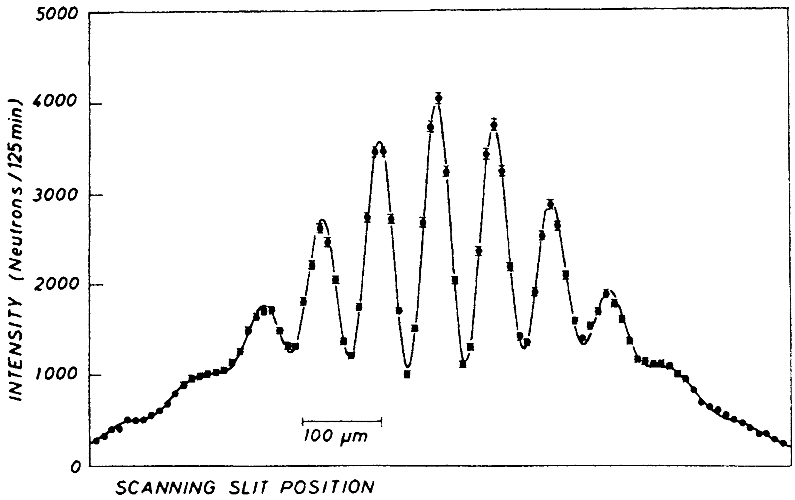

Fig. 2.6 A double-slit diffraction pattern measured with neutrons (Reprinted from [78]. Copyright by the American Physical Society and the Institut für Experimentalphysik, Universität Wien)

can be determined beforehand). After the detection of a large number of particles, a pattern of equidistant fringes with a high density of impacts can be seen at the detection screen.

In a relatively recent version (1988), neutrons of de Broglie wavelength 2 nm (4.34) impinge on two slits 22 μm and 23 μm wide, respectively, separated by a distance of 104 μm. The main results of this experiment are [24]:

- A diffraction pattern indicating the existence of a wave interference phenomenon (Fig. 2.6). The observation plane was located at a distance of 5 m from the slits to insure a resolution of $\approx$100 μm. The solid line represents first-principle predictions from quantum mechanics, including all features of the experimental apparatus.
- The state (2.29) describes a superposition of amplitudes rather than a sum of probabilities, leading to interference terms in the probability $|\Psi|^2$.
- The discreetness of the detection events exhibits the corpuscular nature of the neutrons.
- Neutrons were collected one by one at the observation plane, at a maximum rate of one neutron every 2 s. Therefore, while one neutron was being registered, the next one to arrive was usually still inside the uranium parent nucleus. The particle nature of neutrons is thus also confirmed. Moreover, constraints on the validity of quantum physics to statistical ensembles are ruled out.

Similar interference patterns have been obtained with atoms, molecules and clusters, including fullerenes, a composite molecule of 64 carbon atoms [25].

The superposition giving rise to interference phenomena requires that there is no way to know, even in principle, which path the particle took, a or a'. Interference is destroyed if this information exists, even if it is dispersed in the environment.

2.6 Other Consequences of the Basic Principles

It is shown in Sect. 2.6.1 that the commutation relation between two Hermitian operators $\hat{r}, \hat{s}$ determines the precision with which the values of the corresponding physical quantities may be simultaneously determined. Thus, Heisenberg uncertainty relations between momenta and coordinates become extended to any pair of observables and appear as a consequence of their commutation relations.

We also present the no-cloning theorem (Sect. 2.6.2) and point out the relation between quantum commutators and Poisson brackets (Sect. 2.6.3).

2.6.1 Commutation Relations and the Uncertainty Principle

One assumes two Hermitian operators, $\hat{R}, \hat{S}$, and defines a third (non-Hermitian) operator $\hat{Q}$, such that

$$\hat{Q} \equiv \hat{R} + \mathrm{i}\lambda\hat{S}, \tag{2.30}$$

where λ is a real constant. The minimization with respect to λ of the positively defined norm [see (2.52)]

$$\begin{aligned} 0 \leq \langle\hat{Q}\Psi|\hat{Q}\Psi\rangle &= \langle\Psi|Q^+Q|\Psi\rangle \\ &= \langle\Psi|R^2|\Psi\rangle + \mathrm{i}\lambda\langle\Psi|[R,S]|\Psi\rangle + \lambda^2\langle\Psi|S^2|\Psi\rangle \end{aligned} \tag{2.31}$$

yields the value

$$\begin{aligned} \lambda_{\mathrm{min}} &= -\frac{\mathrm{i}}{2}\langle\Psi|[R,S]|\Psi\rangle/\langle\Psi|S^2|\Psi\rangle \\ &= -\frac{\mathrm{i}}{2}\langle\Psi|[R,S]^+|\Psi\rangle^*/\langle\Psi|S^2|\Psi\rangle \\ &= \frac{\mathrm{i}}{2}\langle\Psi|[R,S]|\Psi\rangle^*/\langle\Psi|S^2|\Psi\rangle. \end{aligned} \tag{2.32}$$

In the second line we have used the definition (2.11) of the Hermitian conjugate. In the last line, the relation $[\hat{R},\hat{S}]^+ = -[\hat{R},\hat{S}]$ stems from the Hermitian character of the operators [see (2.51)]. Substitution of the value λ_{min} in (2.31) yields

$$0 \leq \left(\langle\Psi|R^2|\Psi\rangle - \frac{1}{4}\frac{|\langle\Psi|[R,S]|\Psi\rangle|^2}{\langle\Psi|S^2|\Psi\rangle}\right) \tag{2.33}$$

or

$$\langle\Psi|R^2|\Psi\rangle\,\langle\Psi|S^2|\Psi\rangle \geq \frac{1}{4}\,|\langle\Psi|[R,S]|\Psi\rangle|^2\,. \tag{2.34}$$

The following two operators $\hat{r}, \hat{s}$ have zero expectation value

$$\hat{r} \equiv \hat{R} - \langle\Psi|R|\Psi\rangle\;, \qquad \hat{s} \equiv \hat{S} - \langle\Psi|S|\Psi\rangle, \tag{2.35}$$

and the product of their uncertainties is constrained by [see (2.20)]

$$\Delta r\,\Delta s \geq \frac{1}{2}\,|\langle\Psi|[r,s]|\Psi\rangle|\,. \tag{2.36}$$

Operators corresponding to observables can always be written in the form (2.35). If we prepare a large number of quantum systems in the same state Ψ and then perform some measurements of the observable r in some of the systems, and of s in the others, then the standard deviation Δr of the r-results times the standard deviation Δs of the s-results should satisfy the inequality (2.36). These are intrinsic limitations to the accuracy with which the values of two observables can be determined.

In the case of coordinate and momentum operators, the relation (2.15) yields the Heisenberg uncertainty relation

$$\Delta x\,\Delta p \geq \frac{\hbar}{2}. \tag{2.37}$$

We emphasize the fact that this relation stems directly from basic principles and, in particular, from the commutation relation (2.15). It constitutes an intrinsic limitation upon our knowledge. This limitation cannot be overcome, for instance, by any improvement of the experiment.

If the state of the system is an eigenstate of the operator $\hat{r}$, then a measurement of the observable r yields the corresponding eigenvalue. The value of the observable s associated with a non-commutating operator $\hat{s}$ is undetermined. This is the case of a plane wave describing a particle in free space (Sect. 4.3) for which the momentum may be determined with complete precision, while the particle is spread over all space.

Another consequence of the relation (2.36) is that the state vector Ψ may be simultaneously an eigenstate of $\hat{r}$ and $\hat{s}$ only if these two operators commute, since in this case the product of their uncertainties vanishes. Moreover, if the operators commute and the eigenvalues of $\hat{s}$ are all different within a subset of states, then the matrix elements of $\hat{r}$ are also diagonal within the same subset of states (see Sect. 2.7.1*).

Heisenberg conceived the uncertainty relations to solve the wave-particle paradox. Pure particle behavior requires localization of the particle, while clear wave behavior appears only when the particle has a definite momentum. Heisenberg's interpretation of this was that each of these extreme classical descriptions is satisfied only when the other is completely untenable. Neither picture is valid for

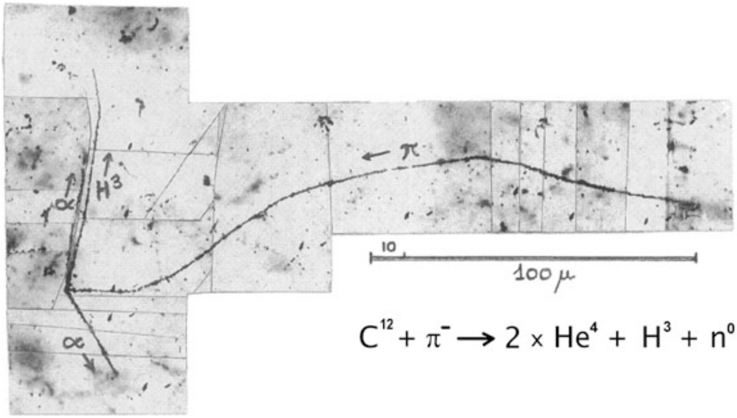

Fig. 2.7 Apparent classical trajectory of a pion. (Reproduced with permission from the authors)

intermediate situations. However, quantum mechanics has to be compatible with the description of the motion of elementary particles (not only with the description of the motion of macroscopic bodies) in terms of trajectories. Heisenberg's answer is that one may construct states Ψ that include a certain amount of localization $\mathbf{p}_0(t)$ and $\mathbf{x}_0(t)$ in both momentum and coordinate. Thus the motion of a particle has some resemblance to classical motion along trajectories. However, there should be a certain spread in the momentum and in the coordinate, such that the amplitudes $\langle \mathbf{p}|\Psi\rangle$ and $\langle \mathbf{x}|\Psi\rangle$, in momentum eigenstates and in position eigenstates, allow uncertainty relations to hold [26].

For an illuminating example, Fig. 2.7 displays the capture of a pion by a carbon nucleus [27]. One can determine the mass, energy and charge of the particles, by measuring the length, the grain density and the scattering direction of their tracks. Let us assume a pion kinetic energy of 10 MeV. Using the pion mass (139 MeV/c^2), one obtains a momentum of p_π=53 MeV/c. The uncertainty in the direction perpendicular to the track may be estimated from the width of the track to be $\approx$1 μm, which yields $\Delta p_\perp \approx 10^{-7}$ MeV/c. The ratio $\Delta p_\perp / p_\pi \approx 10^{-9}$ is too small to produce a visible alteration of the apparent trajectory

2.6.2 No-Cloning Theorem

Let us assume that there are many systems prepared in the state

$$\Psi = a_+ \varphi_+ + a_- \varphi_- \,, \tag{2.38}$$

and that the two amplitudes $a_\pm$ are unknown to us. By measuring two (non-commuting) observables such as

$$\hat{Q} = |\varphi_+\rangle\langle\varphi_+| - |\varphi_-\rangle\langle\varphi_-| \,, \qquad \hat{R} = |\varphi_+\rangle\langle\varphi_-| + |\varphi_-\rangle\langle\varphi_+| \,, \tag{2.39}$$

we obtain the averages

$$\langle\Psi|Q|\Psi\rangle = |a_+|^2 - |a_-|^2 \ , \qquad \langle\Psi|R|\Psi\rangle = a_+^* a_- \, + \, a_-^* a_+ \ , \tag{2.40}$$

from which we can deduce the value of the amplitudes $a_\pm$, including the relative phase.

The situation is different if there is a single copy, since the probability amplitudes are lost after the first measurement. A possible way out would be to produce many copies of the initially single system, but this is prevented by the no-cloning theorem, which also reflects the fragility of quantum states. The theorem says that the state of a particle cannot be copied onto another particle, while the original particle remains in the same state [28]. This is also completely different from what happens in classical mechanics, where we can specify as much as required the state of the system by performing additional measurements without disturbing it.

Suppose the state of two particles is given by

$$\varphi(1)\,\chi(2) \tag{2.41}$$

and that some unitary evolution effects the copying process

$$\varphi(1)\,\varphi(2) = \mathcal{U}\,\varphi(1)\,\chi(2) \ . \tag{2.42}$$

Suppose now that this copying procedure also works for another state

$$\phi(1)\,\phi(2) = \mathcal{U}\,\phi(1)\,\chi(2) \ . \tag{2.43}$$

The scalar product between (2.42) and (2.43) yields

$$\langle\varphi|\phi\rangle^2 = \langle\chi\,\varphi|\mathcal{U}^+\mathcal{U}|\phi\,\chi\rangle = \langle\varphi|\phi\rangle \ . \tag{2.44}$$

Since this equation has two solutions, 0 and 1, either $\varphi = \phi$ or they are mutually orthogonal. Therefore, a general quantum cloning device is impossible.

Even if one allows non-unitary cloning devices, the cloning of non-orthogonal pure states remains impossible unless one is willing to tolerate a finite loss of fidelity in the cloned states.

On the positive side, the impossibility of determining the components of a vector state prevents non-relativistic quantum mechanics to clash with relativity theory (Sects. 12.1 and 13.3).

2.6.3 Commutation Relations and Poisson Brackets

From the commutation relation (2.15), one obtains

$$[\hat{x}^n, \hat{p}] = \mathrm{i}\hbar\, n\, \hat{x}^{n-1} \ , \qquad [\hat{x}, \hat{p}^n] = \mathrm{i}\hbar\, n\, \hat{p}^{n-1} \ , \tag{2.45}$$

which shows that the commutation operation acts as a sort of differential operation. Arbitrary functions of the coordinate and the momenta, $\hat{u} = u(\hat{x}, \hat{p})$, satisfy the following relations:

$$\begin{aligned}
[\hat{u}, \hat{v}] &= -[\hat{v}, \hat{u}] \\
[\hat{u}, c] &= 0 \\
[(\hat{u} + \hat{v}), \hat{w}] &= [\hat{u}, \hat{w}] + [\hat{v}, \hat{w}] \\
[\hat{u}\,\hat{v}, \hat{w}] &= [\hat{u}, \hat{w}]\hat{v} + \hat{u}[\hat{v}, \hat{w}] \\
0 &= [\hat{u}, [\hat{v}, \hat{w}]] + [\hat{v}, [\hat{w}, \hat{u}]] + [\hat{w}, [\hat{u}, \hat{v}]] \,,
\end{aligned} \tag{2.46}$$

as may be verified through power series expansions. These properties are shared by the (classical) Poisson brackets

$$\{u, v\}_{\text{PB}} = \frac{\partial u}{\partial x}\frac{\partial v}{\partial p} - \frac{\partial u}{\partial p}\frac{\partial v}{\partial x} \,. \tag{2.47}$$

Dirac has pointed out that "classical mechanics provides a valid description of dynamical systems under certain conditions. We should thus expect that important concepts in classical mechanics correspond to important concepts in quantum mechanics" [15]. In fact, the quantum mechanical commutator can be obtained from the classical Poisson bracket through the replacement of the coordinate and momentum by the corresponding quantum operators and multiplication by i$\hbar$

$$\mathrm{i}\hbar\{u, v\}_{\text{PB}} \rightarrow [\hat{u}, \hat{v}] \,. \tag{2.48}$$

The commutativity of the classical multiplication rule is obtained in the limiting case $\hbar \rightarrow 0$.

However, the Poisson brackets (2.47) can only be defined with reference to a definite set of coordinates and momenta (although they are invariant under a change of this set). The quantum commutation relation is not limited by this condition and, therefore, has a more fundamental character. A particularly relevant example is the case of the spin components, which do not have classical counterparts (see Sect. 5.2).

2.7* Properties of Hilbert Spaces and Operators

In the following, we briefly review some properties of these mathematical tools.

A Hilbert space is a generalization of the Euclidean, three-dimensional space (see Table 2.1). As in ordinary space, the summation $c_a\Psi_a + c_b\Psi_b$ and the scalar

Table 2.1 Some relevant properties of vectors and operators in Euclidean and Hilbert spaces

	Euclidean space	Hilbert space
Vectors	$\mathbf{r}$	Ψ
Superposition	$\mathbf{r} = c_a \mathbf{r}_a + c_b \mathbf{r}_b$	$\Psi = c_a \Psi_a + c_b \Psi_b$
Scalar product	$\langle \mathbf{r}_a \vert \mathbf{r}_b \rangle = \mathbf{r}_a \cdot \mathbf{r}_b = c_{ab}$	$\langle \Psi_a \vert \Psi_b \rangle = \langle \Psi_b \vert \Psi_a \rangle^* = c_{ab}$
	$c_a, c_b, c_{ab} = \text{real}$	$c_a, c_b, c_{ab} = \text{complex}$
Basis set	$\langle \mathbf{v}_i \vert \mathbf{v}_j \rangle = \delta_{ij}$	$\langle \varphi_i \vert \varphi_j \rangle \equiv \langle i \vert j \rangle = \delta_{ij}$
Dimension ν	3	$2 \leq \nu \leq \infty$
Completeness	$\mathbf{r} = \sum_i x_i \mathbf{v}_i$	$\Psi = \sum_i c_i \varphi_i$
Projection	$x_i = \langle \mathbf{v}_i \vert \mathbf{r} \rangle$	$c_i = \langle \varphi_i \vert \Psi \rangle$
Scalar product	$\langle \mathbf{r}_a \vert \mathbf{r}_b \rangle = \sum_i x_i^{(a)} x_i^{(b)}$	$\langle \Psi_a \vert \Psi_b \rangle = \sum_i \left(c_i^{(a)} \right)^* c_i^{(b)}$
Norm	$\langle \mathbf{r} \vert \mathbf{r} \rangle^{1/2} = \left(\sum_i x_i^2 \right)^{1/2}$	$\langle \Psi \vert \Psi \rangle^{1/2} = \left(\sum_i \vert c_i \vert^2 \right)^{1/2}$
Operators	$\hat{R}_\eta(\theta)\, \mathbf{r}_a = \mathbf{r}_b$	$\hat{Q} \Psi_a = \Psi_b$
Commutators	$[\hat{R}_x(\frac{\pi}{2}), \hat{R}_y(\frac{\pi}{2})] \neq 0$	$[\hat{Q}, \hat{R}]$
Eigenvalues	$\hat{D}_i \mathbf{v}_i = \lambda_i \mathbf{v}_i$	$\hat{Q} \varphi_i = q_i \varphi_i$

product $\langle \Psi_b | \Psi_a \rangle = c_{ab}$ between two vectors are well-defined operations.[15] While the constants c_a, c_b, c_{ab} are real numbers in everyday space, it is essential to allow for complex values in quantum mechanics.

The norm of a vector is defined as the square root of the scalar product of a vector with itself. A vector is said to be normalized if its norm equals 1. Two vectors are orthogonal if their scalar product vanishes. A vector Ψ is linearly independent of a subset of vectors $\Psi_a, \Psi_b, \ldots, \Psi_d$ if it cannot be expressed as a linear combination of them[16] ($\Psi \neq c_a \Psi_a + c_b \Psi_b + \cdots + c_d \Psi_d$).

These last two concepts allow us to define sets of basis vectors φ_i satisfying the requirement of orthonormalization. Moreover, these sets may be complete, in the sense that any vector Ψ may be expressed as a linear combination of them[17] [see (2.6)]. The scalar product $\langle i | \Psi \rangle$ is the projection of Ψ onto φ_i. The scalar product between two vectors Ψ_a, Ψ_b and the square of the norm of the vector Ψ are also given in terms of the amplitudes c_i in Table 2.1.

The number of states in a basis set is the dimension ν of the associated Hilbert space. It has the value 3 in normal space. In this book, we use Hilbert spaces with dimensions ranging from two to infinity.

15 Definition of these fundamental operations is deferred to each realization of Hilbert spaces [(3.2), (3.4) and (4.1), (4.2)]. In the present chapter we use only the fact that they exist and that $\langle a|b \rangle = \langle b|a \rangle^*$.

16 Although the term "linear combination" usually refers only to finite sums, we extend its meaning to also include an infinity of terms.

17 The most familiar case of the expansion of a function in terms of an orthonormal basis set is the Fourier expansion in terms of the exponentials $\exp[ikx]$, which constitute the complete set of eigenfunctions corresponding to the free particle case (see Sect. 4.3).

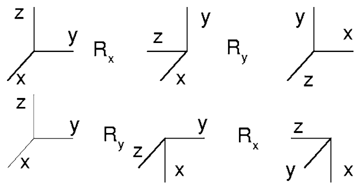

Fig. 2.8 The final orientation of the axes depends on the order of the rotations. R_η here represents a rotation of $\pi/2$ around the η-axis

In ordinary space, vectors are defined by virtue of their transformation properties under rotation operations $\hat{R}_\eta(\theta)$ (η denoting the axis of rotation, and θ the angle). These operations are generally non-commutative, as the reader may easily verify by performing two successive rotations of $\theta = \pi/2$, first around the x-axis and then around the y-axis, and subsequently comparing the result with the one obtained by reversing the order of these rotations (Fig. 2.8). Mathematically, this is expressed through the non-vanishing of the operator (2.8).

In ordinary space, a dilation $\hat{D}$ is an operation yielding the same vector multiplied by a (real) constant. This operation has been generalized in terms of eigenvectors and eigenvalues in (2.9). In general, linear combinations of such eigenvectors do not satisfy the eigenvalue equation.

2.7.1 Some Properties of Hermitian Operators*

The Hermitian conjugate operator $\hat{Q}^+$ is defined through the (2.11). Similarly, we may write

$$\langle \Psi_b | Q | \Psi_a \rangle = \langle \hat{Q} \Psi_a | \Psi_b \rangle^* = \langle \hat{Q}^+ \Psi_b | \Psi_a \rangle \,. \tag{2.49}$$

The following properties are easy to demonstrate

$$\left(\hat{Q} + c\hat{R}\right)^+ = \hat{Q}^+ + c^* \hat{R}^+, \tag{2.50}$$

$$\left(\hat{Q}\hat{R}\right)^+ = \hat{R}^+ \hat{Q}^+. \tag{2.51}$$

According to (2.11), the norm of the state $\hat{Q}\Psi$ is obtained by

$$\langle \hat{Q}\Psi | \hat{Q}\Psi \rangle^{1/2} = \langle \Psi | Q^+ Q | \Psi \rangle^{1/2} \,. \tag{2.52}$$

The norm is a real, positive number.

An operator is said to be Hermitian if it is equal to its own Hermitian conjugate operator

$$\hat{Q}^+ = \hat{Q}. \tag{2.53}$$

Assume now that the state φ_i is an eigenstate of the Hermitian operator $\hat{Q}$ corresponding to the eigenvalue q_i. In this case

$$\langle i|Q|i\rangle = q_i \langle i|i\rangle \,, \qquad \langle i|Q|i\rangle^* = q_i^* \langle i|i\rangle$$
$$\langle i|Q|i\rangle = \langle i|Q|i\rangle^* \;\rightarrow\; q_i = q_i^* . \tag{2.54}$$

Therefore, the eigenvalues of Hermitian operators are real numbers.

Consider now the non-diagonal terms

$$\langle j|Q|i\rangle = q_i \langle j|i\rangle \,, \quad \langle i|Q|j\rangle^* = q_j^* \langle i|j\rangle^* = q_j^* \langle j|i\rangle \,, \tag{2.55}$$

then

$$0 = (q_i - q_j)\,\langle j|i\rangle, \tag{2.56}$$

i.e. two eigenstates, belonging to different eigenvalues, are orthogonal. They may also be orthonormal, upon multiplication by an appropriate normalization constant, which is determined up to a phase.

The eigenvectors of a Hermitian operator constitute a complete set of states for a given system. This means that any state function Ψ, describing any state of the same system, may be expressed as a linear combination of basis states φ_i [see (2.6)].

We define the projection operator (a theoretical filter) $|i\rangle\langle i|$ through the equation

$$|i\rangle\langle i|\,\varphi_j \equiv \langle i|j\rangle\varphi_i = \delta_{ij}\,\varphi_i \,, \tag{2.57}$$

which implies that

$$\sum_i |i\rangle\langle i|\,\Psi = \Psi \tag{2.58}$$

for any Ψ. Thus, unity may be expressed as the operator $\sum_i |i\rangle\langle i|$. From this property stems the closure property, according to which the matrix elements of the product of Hermitian operators may be calculated as the sum over all possible intermediate states of products of the matrix elements corresponding to each separate operator

$$\langle i|QR|j\rangle = \sum_k \langle i|Q|k\rangle\langle k|R|j\rangle. \tag{2.59}$$

2.7.2 Unitary Transformations*

The unitary matrix $(\mathcal{U}_{ai}) = (\langle i|a\rangle)$ in (2.14) transforms the basis set φ_i into the basis set χ_a. Such a matrix does not represent a physical observable and it is not therefore required to be Hermitian.

The inverse transformation to (2.14) is written as

$$\varphi_i = \sum_a \langle a|i\rangle \chi_a \ . \tag{2.60}$$

Therefore, the inverse transformation $\mathcal{U}^{-1}$ is the transposed conjugate:

$$\mathcal{U}^{-1} = (\langle a|i\rangle) = \mathcal{U}^+ \ , \qquad \mathcal{U}^+\mathcal{U} = \mathcal{U}\mathcal{U}^+ = \mathcal{I}, \tag{2.61}$$

where $\mathcal{I}$ is the unit matrix. A matrix satisfying (2.61) is said to be unitary. Equation (2.61) implies that

$$\sum_i \langle a|i\rangle\langle i|b\rangle = \delta_{ab}$$

$$\sum_a \langle i|a\rangle\langle a|j\rangle = \langle i|j\rangle = \delta_{ij} \tag{2.62}$$

If states are transformed according to $\chi = \mathcal{U}\,\varphi$, then the state $\mathcal{U}\,\hat{Q}\,\varphi$ may be written as

$$\mathcal{U}\,\hat{Q}\,\varphi = \mathcal{U}\,\hat{Q}\,\mathcal{U}^+\mathcal{U}\,\varphi = \mathcal{R}\,\chi, \tag{2.63}$$

which yields the rule for the transformation of operators, namely

$$\hat{R} = \mathcal{U}\,\hat{Q}\,\mathcal{U}^+. \tag{2.64}$$

In addition to the norm, unitary transformations preserve the value of the determinant and the trace

$$\det(\langle a|R|b\rangle) = \det(\langle i|Q|j\rangle),$$

$$\mathrm{trace}(Q) \equiv \sum_i \langle i|Q|i\rangle = \mathrm{trace}(R) \equiv \sum_a \langle a|R|a\rangle. \tag{2.65}$$

2.8* Notions on Probability Theory

Probability theory studies the likelihood P_i that the outcome q_i of an event will take place. The limits of P_i are

$$0 \le P_i \le 1 \ . \tag{2.66}$$

If $P_i = 0$, the outcome q_i cannot occur; if $P_i = 1$, it will take place with certainty.

If two events (i, j) are statistically independent, the probability that both i and j take place is given by the product

$$P_{(i \text{ and } j)} = P_i \, P_j \, . \tag{2.67}$$

If two events are mutually exclusive, the probability that one or the other occur is the sum

$$P_{(i \text{ or } j)} = P_i + P_j . \tag{2.68}$$

Probability may be defined as

$$P_i \equiv \lim_{N\to\infty} \frac{n_i}{N} , \tag{2.69}$$

where n_i is the number of outcomes q_i of a total of $N \equiv \sum_i n_i$ outcomes. Since the limit $N \to \infty$ is never attained, in practice N should be made large enough so that the fluctuations become sufficiently small.

The collection of P_is is called the (discrete) probability distribution. The concepts of average $\langle Q\rangle$, root mean square $\langle Q^2\rangle^{1/2}$ and root mean square deviation ΔQ, applied in Sect. 2.4, are given by

$$\begin{aligned}
\langle Q\rangle &= \sum_i q_i\, P_i ,\\
\langle Q^2\rangle^{1/2} &= \left(\sum_i q_i^2\, P_i\right)^{1/2} ,\\
\Delta Q &= \langle (Q-\langle Q\rangle)^2\rangle^{1/2} = \left(\langle Q^2\rangle - \langle Q\rangle^2\right)^{1/2} .
\end{aligned} \tag{2.70}$$

In the case of a continuous distribution, the sums $\sum_i$ are replaced by integrals $\int \mathrm{d}x$. Instead of probabilities P_i one defines probability densities $\rho(x)$ such that

$$\begin{aligned}
1 &= \int_{-\infty}^{\infty} \rho(x)\,\mathrm{d}x ,\\
\langle Q\rangle &= \int_{-\infty}^{\infty} q(x)\,\rho(x)\mathrm{d}x .
\end{aligned} \tag{2.71}$$

Problems

Problem 1. Assume that the state Ψ is given by the linear combination $\Psi = c_1\Psi_1 + c_2\Psi_2$, where the amplitudes c_1, c_2 are arbitrary complex numbers, and both states Ψ_1, Ψ_2 are normalized.

1. Normalize the state Ψ, assuming that $\langle 1|2\rangle = 0$.
2. Find the probability of the system being in the state Ψ_1.

Problem 2. Use the same assumptions as in Problem 2, but $\langle 1|2\rangle = c \neq 0$

1. Find a linear combination $\Psi_3 = \lambda_1\Psi_1 + \lambda_2\Psi_2$, such that it is orthogonal to Ψ_1 and normalized.
2. Express the vector Ψ as a linear combination of Ψ_1 and Ψ_3.

Problem 3. Prove (2.50) and (2.51).
Hint: Apply successively the definition of Hermitian conjugate to the operators $\hat{Q}, \hat{R}$. For instance, start with $\langle\Psi_b|Q\ R|\Psi_a\rangle = \langle\hat{R}\Psi_a|Q^+|\Psi_b\rangle^*= \cdots$.

Problem 4. Show that

$$[\hat{Q}, \hat{R}] = -[\hat{R}, \hat{Q}]$$
$$[\hat{Q}\,\hat{R}, \hat{S}] = [\hat{Q}, \hat{S}]\,\hat{R} + \hat{Q}\,[\hat{R}, \hat{S}]\,.$$

Problem 5. Find the commutation relation between the coordinate operator $\hat{x}$ and the one-particle Hamiltonian (2.16). Discuss the result in terms of the simultaneous determination of energy and position of a particle.

Problem 6. Find the commutation relations

1. $[\hat{p}^n, \hat{x}]$, where n is an integer
2. $[f(\hat{p}), \hat{x}]$. Hint: Expand $f(\hat{p})$ in power series of $\hat{p}$ and apply the previous commutation relation.

Problem 7. Verify that the commutation relation (2.15) is consistent with the fact that the operators $\hat{x}$ and $\hat{p}$ are Hermitian.

Problem 8. Assume the basis set of states φ_i

1. Calculate the effect of the operator $\hat{R} \equiv \sum_i |i\rangle\langle i|$ on an arbitrary state Ψ.
2. Repeat for the operator $\hat{R} \equiv \Pi_i \left(\hat{Q} - q_i\right)$, assuming that the equation $\hat{Q}\varphi_i = q_i\varphi_i$ is satisfied.

Problem 9. Find the relation between the matrix elements of the operators $\hat{p}$ and $\hat{x}$ in the base of eigenvectors of the Hamiltonian (2.16).

Problem 10. Consider the eigenvalue equations

$$\hat{F}\varphi_1 = f_i\varphi_1\,, \quad \hat{F}\varphi_2 = f_2\varphi_2\,, \quad \hat{G}\chi_1 = g_1\chi_1\,, \quad \hat{G}\chi_2 = g_2\chi_2\,, \tag{2.72}$$

and the relations

$$\varphi_1 = \frac{1}{\sqrt{5}}(2\chi_1 + \chi_2)\,, \qquad \varphi_2 = \frac{1}{\sqrt{5}}(\chi_1 - 2\chi_2). \tag{2.73}$$

1. Is it possible to simultaneously measure the observables F and G?

2. Assume that a measurement of F has yielded the eigenvalue f_1. Subsequently G and F are measured (in this order). Which are the possible results and their probabilities?

Problem 11. Consider eigenstates φ_p of the momentum operator. Assume that the system is prepared in the state

$$\Psi = \frac{1}{\sqrt{6}}(\varphi_{2p} + \varphi_p) + \sqrt{\frac{2}{3}}\varphi_{-p}. \tag{2.74}$$

1. What are the possible results of a measurement of the kinetic energy K, and what are their respective probabilities?
2. Calculate the expectation value and the standard deviation of the kinetic energy.
3. What is the vector state after a measurement of the kinetic energy that has yielded the value $k_p = p^2/2M$?

Problem 12. Evaluate, in m.k.s. units, possible values of the precision to which the velocity and the position of a car should be measured to verify the uncertainty relation (2.37).

Problem 13. A 10 MeV proton beam is collimated by means of diaphragms with a 5 mm aperture.

1. Show that the spread in energy ΔE_H, associated with the uncertainty principle, is negligible relative to the total spread, $\Delta E \approx 10^{-3}$ MeV.
2. Calculate the distance x that a proton has to travel to traverse 5 mm in a perpendicular direction, if the perpendicular momentum is due only to the uncertainty principle.

Problem 14. Verify (2.46).