Quantum Mechanics I (SS 2009)

Prof. Dr. G. M. Pastor
Institut für Theoretische Physik
Fachbereich Naturwissenschaften
Universität Kassel

January 16, 2017

Das Skript, das Sie in der Hand haben, ist ein Lernhilfsmittel zur Vorlesung „Quantenmechanik I", wie sie an der Universität Kassel im SS 09 gehalten wurde. Es handelt sich keineswegs um ein vollständiges Werk oder um einen richtigen Beitrag zur der zahlreichen auf dem Gebiet vorhandenen Literatur, deren Bearbeitung den Studenten dringend anempfohlen wird. Darüber hinaus soll die jetzige Fassung nur als ein erster Entwurf betrachtet werden. Dies wird der Leser nicht nur an den „geschickt" verstreuten Druckfehlern und knappen Formulierungen, sondern auch am Fehlen von erläuternden Abbildungen erkennen. Dies ist eine Schwäche, die wir in späteren Fassungen beheben werden.

Kassel, Dezember 2009

Contents

1. Fundamental concepts of classical mechanics

In this chapter we briefly recall the fundamental concepts of classical mechanics starting from Newton's second law and introducing basic notions like linear momentum, energy, Hamiltonian and conservation laws.

1.1. Dynamics of a single particle: Newton's second law

Given a physical system as defined by its constituent particles, their interactions and possible external fields, the purpose of a mechanical theory (either classical, quantum or statistical) is to be able to predict the state of the system at any future time $t > t_0$ on the basis of the sole knowledge of the state of the system at a previous, so-called initial time t_0. The mathematical definition of the state will of course depend on the type of theory one is considering (e.g., classical, quantum or statistical mechanics), but in all cases it should be such that it contains the minimum amount of information needed to predict the outcome of any experiment performed in the system.

In the case of a single particle the state in classical mechanics is defined by the position vector $\vec{r}$ and the linear momentum $\vec{p}$ ($\vec{r}, \vec{p} \in \mathbb{R}^3$). Let $\vec{r}(t)$ be the curve traced by the particle as a function of t, then the linear momentum is defined in terms of the velocity

$$\vec{v} = \frac{d\vec{r}}{dt} = \lim_{\Delta t \to 0} \frac{\vec{r}(t+\Delta t) - \vec{r}(t)}{\Delta t}$$

by

$$\vec{p} = m\,\vec{v} = m\,\dot{\vec{r}}, \tag{1.1}$$

where m is the mass of the particle. Notice that we have implicitly assumed that the position is a continuous (actually differentiable) function of time, i.e., as $\Delta t \to [\vec{r}(t+\Delta t) - \vec{r}(t)] \simeq \vec{v}\,\Delta t$, where $\vec{v}$ is a constant vector. Do you think this is a reasonable assumption?

The fundamental physics involved in the dynamics of a particle is contained in Newton's second law of motion, which in terms of $\vec{r}$ and $\vec{p}$ is given by

$$\vec{F}\big(\vec{r}(t)\big) = \frac{d\vec{p}}{dt} = \dot{\vec{p}}, \tag{1.2}$$

where $\vec{F}$ is the force acting on the particle when it is at the position $\vec{r}$. One says that $\vec{F}(\vec{r})$ is a force vector field ($\vec{F} : \mathbb{R}^3 \to \mathbb{R}^3$). Once more, differentiability of $\vec{p}$ as a function of t is assumed.

If the mass of the particle is independent of time, from (1.1) and (1.2) we obtain the more familiar but less general form

$$\vec{F} = m\,\frac{d\vec{v}}{dt} = m\,\ddot{\vec{r}},$$

where $\ddot{\vec{r}} = \vec{a}$ is the acceleration.

1.2. Constants of motion

In classical mechanics, as well as in quantum mechanics, there are special physical quantities whose value remains constant throughout the time evolution of the system. These quantities have a special physical meaning and provide a fundamental insight on the symmetries underlying the laws of motion.

1.2.1. Linear momentum

The first and most fundamental conservation theorem is the law of *conservation of linear momentum*:

If the force $\vec{F}(\vec{r})$ (or sum of forces) acting on a particle is zero then the momentum $\vec{p}$ is conserved.

This statement (also known as Newton's first law[1]) is an immediate consequence of Eq. (1.2) since

$$\vec{F} = 0 \quad \Rightarrow \quad \dot{\vec{p}} = 0 \quad \Rightarrow \quad \vec{p} = \text{constant}.$$

The linear momentum conservation law is intimately related with translational symmetry. It appears whenever the system is invariant under translations, i.e., when all points in space are equivalent, in the absence of external forces, or when the energy at all points in space is the same. We shall come back to this point in Sec. 1.4.

1.2.2. Angular momentum

The angular momentum $\vec{L}$ of a particle with respect to a given point O (typically the origin of the coordinate system) is defined as

$$\vec{L} = \vec{r} \times \vec{p},$$

where $\vec{r}$ is the position vector going from O to the particle. Clearly, the value of $\vec{L}$ depends on the choice of O. This can be illustrated by considering, for example, a free particle with constant $\vec{p}$ and a point O at a distance d from the trajectory, in which case $L = pd$.

The time dependence of $\vec{L}$ is given by

$$\frac{d\vec{L}}{dt} = \frac{d\vec{r}}{dt} \times \vec{p} + \vec{r} \times \frac{d\vec{p}}{dt}.$$

Since $\vec{p} = m\,\dfrac{d\vec{r}}{dt}$ we have

$$\frac{d\vec{L}}{dt} = \vec{r} \times \vec{F} = \vec{N}, \tag{1.3}$$

---

[1] Actually, the first law is not just a special case of the second one, since it defines the notion of inertial systems, where the second law applies.

where $\vec{F}$ is the force acting on the particle and $\vec{N} = \vec{r} \times \vec{F}$ the torque of the force. Equation (1.3) is a relation between vectors (pseudovectors) which holds irrespectively of the coordinate system, although both $\vec{L}$ and $\vec{N}$ depend on the choice of the origin O.

From Eq. (1.3) the law of *conservation of angular momentum* follows:

If the total torque $\vec{N}$ is zero, the angular momentum $\vec{L}$ is conserved ($\dot{\vec{L}} = 0$).

This conservation law is useful in problems involving central forces, i.e., force fields pointing to a common origin O, which is taken as the origin of the coordinate system. For instance the gravitation field of the sun in the solar system or the electric field of the nucleus $\vec{E} = -Ze\,\vec{r}/r^3$ in the atom.

The conservation of angular momentum is intimately related with the rotational symmetry of the force field. Since Eq. (1.3) is a vector equation it holds for each component independently, i.e., if $N_z = 0$ then L_z is conserved.

1.2.3. Energy conservation: Kinetic and potential energy

Let us consider a particle of constant mass m under the action of an external force $\vec{F}$ and determine the work done by $\vec{F}$ when the particle moves between two points $\vec{r}_1$ and $\vec{r}_2$ along the trajectory. This work is given by the circulation

$$W_{12} = \int_{\vec{r}_1}^{\vec{r}_2} \vec{F} \cdot d\vec{r}.$$

The integral can be calculated by using the parametrization of the trajectory $\vec{r} = \vec{r}(t)$ as a function of time [$\vec{r}_1 = \vec{r}(t_1)$ and $\vec{r}_2 = \vec{r}(t_2)$]. Knowing that $\vec{F} = \dfrac{d\vec{p}}{dt}$ and $\vec{p} = m\,\dfrac{d\vec{r}}{dt}$, or equivalently $d\vec{r} = \dfrac{\vec{p}}{m}\,dt$, we have

$$W_{12} = \int_{t_1}^{t_2} \frac{d\vec{p}}{dt} \cdot \frac{\vec{p}}{m}\,dt.$$

Assuming that the mass of the particle is independent of time and observing that $\dfrac{d}{dt}\,(p^2) = \dfrac{d}{dt}\,(\vec{p}\cdot\vec{p}) = 2\,\vec{p}\cdot\dfrac{d\vec{p}}{dt}$ we obtain

$$W_{12} = \frac{1}{2m}\int_{t_1}^{t_2} \frac{d}{dt}\,(p^2)\,dt = \frac{1}{2m}\,(p_2^2 - p_1^2), \tag{1.4}$$

where $\vec{p}_1 = \vec{p}(t_1)$ and $\vec{p}_2 = \vec{p}(t_2)$ are the linear momenta at the initial and final points of the integration path.

The scalar quantity

$$T = \frac{p^2}{2m} = \frac{p_x^2 + p_y^2 + p_z^2}{2m} \tag{1.5}$$

is known as the kinetic energy of the particle. Eq. (1.4) states that the work of the external force is equal to the change in kinetic energy:

$$W_{12} = T_2 - T_1.$$

It should be noted that the non-relativistic expression (1.5) for T is the simplest one that satisfies the conditions imposed by symmetry: i) Since time is homogeneous, T cannot depend explicitly on t; ii) since space is homogeneous, it cannot depend on $\vec{r}$; iii) since space is isotropic, it must be invariant under rotations, i.e., all directions of the motion must be equivalent; and iv) T must be positive definite and vanish only when the particle is at rest ($\vec{p} = 0$). In relativistic classical mechanics T has the form

$$T = \sqrt{(mc^2)^2 + p^2c^2} - mc^2,$$

which of course satisfies all these general symmetry conditions.

All fundamental forces in nature (gravitational, electrical, etc.) have the property that the circulation of the force field $\vec{F}(\vec{r})$ around any closed path $\mathcal{C}$ is zero, i.e.,

$$\oint_{\mathcal{C}} \vec{F}(\vec{r}) \cdot d\vec{r} = 0 \quad \forall\, \mathcal{C} \text{ closed.}$$

Force fields having this property are said to be conservative. Mathematically this is equivalent to requiring

$$\vec{\nabla} \times \vec{F} = 0 \quad \forall\, \vec{r}$$

or

$$\vec{F}(\vec{r}) = -\vec{\nabla}\, V(\vec{r}) = -\left(\frac{\partial V}{\partial x}, \frac{\partial V}{\partial y}, \frac{\partial V}{\partial z}\right),$$

where $V(\vec{r})$ is a scalar differentiable function called potential energy (V: $\mathbb{R}^3 \to \mathbb{R}$). Well-known examples of potential energy functions are the potential energy of the gravitation at the surface of the earth:

$$V(\vec{r}) = mgz \quad \Rightarrow \quad \vec{F} = -\vec{\nabla}\, V = (0, 0, -mg),$$

or the Coulomb interaction between an electron and the nucleus carrying a charge Ze:

$$V(\vec{r}) = -\frac{Ze^2}{r} \quad \Rightarrow \quad \vec{F}(\vec{r}) = -Ze^2\, \frac{\vec{r}}{r^3}.$$

In the presence of conservative force fields the time dependence of the kinetic energy takes a particularly simple and insightful form. Consider $T = p^2/2m$ along a trajectory of the particle. Differentiating with respect to time one obtains

$$\frac{dT}{dt} = \frac{1}{2m}\, \frac{d}{dt}\, (\vec{p} \cdot \vec{p}) = \frac{1}{m}\, \vec{p} \cdot \frac{d\vec{p}}{dt} = -\vec{\nabla}\, V \cdot \frac{d\vec{r}}{dt} = \frac{-dV}{dt},$$

where we have used that m is independent of time, $\vec{F} = -\vec{\nabla}\, V$ and $\vec{p} = m\, \dfrac{d\vec{r}}{dt}$. Moreover, we know that for any scalar field V: $\mathbb{R}^3 \to \mathbb{R}$, the differential dV is given by

$$dV = V(\vec{r} + d\vec{r}) - V(\vec{r}) = \frac{\partial V}{\partial x}\, dx + \frac{\partial V}{\partial y}\, dy + \frac{\partial V}{\partial z}\, dz = \vec{\nabla}\, V \cdot d\vec{r}.$$

Therefore

$$\frac{dT}{dt} = \frac{-dV}{dt},$$

or equivalently,

$$\frac{d}{dt}(T + V) = 0.$$

In other words the total energy

$$E = T + V = \frac{p^2}{2m} + V(\vec{r}) \tag{1.6}$$

is a constant of motion. This explains why the force fields obtained as the gradient of a potential are called conservative fields.

1.3. Hamiltonian and Hamilton equations for a single particle

The total energy E regarded as a function of the dynamical variables $\vec{r}$ and $\vec{p}$ and eventually time t is known as Hamilton's function or Hamiltonian $H = T + V$. This is the central mathematical object in the powerful Hamiltonian formulation of classical mechanics. Moreover, the corresponding operator in quantum mechanics defines the time evolution of the quantum mechanical state or wave function. It is therefore important to gain some insight into the physical significance of H.

For a single particle and in Cartesian coordinates H is given by

$$H = \frac{p^2}{2m} + V(\vec{r}) = \frac{p_x^2 + p_y^2 + p_z^2}{2m} + V(x, y, z).$$

In order to compact the notation, and to allow for more general sets of coordinates (e.g., spherical or cylindrical) and constraints it is customary to denote the coordinates by q and the conjugated or corresponding momenta by p. Thus, in Cartesian coordinates we have

$$\vec{q} = (q_1, q_2, q_3) = (x, y, z)$$

and the conjugated momenta

$$\vec{p} = (p_1, p_2, p_3) = (p_x, p_y, p_z),$$

which are the usual components of the linear momentum.

Newton's second law and the relation between velocity and momentum can be replaced by a set of first order differential equations known as *Hamiltonian's equations* or, owing to their simple and symmetric form, *canonical equations.* In terms of q_i and p_i they take the form

$$\dot{q}_i = \frac{\partial H}{\partial p_i} \tag{1.7}$$

and

$$\dot{p}_i = -\frac{\partial H}{\partial q_i}. \tag{1.8}$$

It is easy to verify that Eq. (1.7) is equivalent to the definition of linear momentum. For instance for $i = 1$ we have

$$\dot{x} = \dot{q}_1 = \frac{\partial H}{\partial p_1} = \frac{\partial (p^2/2m)}{\partial p_x} = \frac{p_x}{m}.$$

The second set of n equations (n is the dimension of space) are equivalent to Newton's second law. For example for $i = 2$

$$\dot{p}_y = \dot{p}_2 = -\frac{\partial H}{\partial q_2} = -\frac{\partial V}{\partial y} = F_y.$$

The canonical equations are valid in general, including interacting many-particle systems, other curvilinear coordinate systems or in the presence of constraints. Therefore, the Hamiltonian

$$H = H(q_1, \dots q_n, p_1, \dots p_n, t) = T + V + W$$

as a function of q_i, p_i and eventually t (or in Cartesian coordinates as a function of $\vec{p}_i$ and $\vec{r}_i$) univocally defines the time evolution of the system starting from the initial conditions $p_i^0 = p_i(t_0)$ and $q_i^0 = q_i(t_0)$. In particular H contains all the information on possible conservation laws.

For the sake of completeness let us mention that in relativistic mechanics the kinetic plus rest energy of a free particle is related to the momentum $\vec{p}$ by the requirement that the magnitude of the momentum four vector is constant:

$$p_\mu \, p_\mu = p^2 - \frac{T^2}{c^2} = -m^2c^2$$

or equivalently

$$T^2 = c^2p^2 + m^2c^4.$$

Consequently, the Hamiltonian of a relativistic particle under the action of a velocity independent potential V is given by

$$H = T + V = \sqrt{c^2p^2 + m^2c^4} + V.$$

1.4. Symmetry and conservation laws

Let us consider a particle or a set of particles moving along their classical trajectory $q_i(t)$ and $p_i(t)$. The total time derivative of the Hamiltonian along the trajectory is given by

$$\frac{dH}{dt} = \frac{\partial H}{\partial t} + \sum_i \left(\frac{\partial H}{\partial q_i} \dot{q}_i + \frac{\partial H}{\partial p_i} \dot{p}_i \right).$$

Substituting the canonical equations $-\frac{\partial H}{\partial q_i} = \dot{p}_i$ and $\frac{\partial H}{\partial p_i} = \dot{q}_i$ we obtain

$$\frac{dH}{dt} = \frac{\partial H}{\partial t} \quad \forall \, t.$$

This implies that if the system is isolated (i.e., no external potential), or if the external potential is independent of time, H does not depend explicitly on t and the energy is conserved ($\partial H/\partial t = 0$). From the point of view of symmetry, one would say that if all times are equivalent or, in other words, if a translation in time does not change the Hamiltonian, then the energy is conserved:

$$\boxed{\text{Time invariance}} \quad \leftrightarrow \quad \boxed{\text{Energy conservation}}$$

Let us now consider a system whose energy is unchanged upon a translation δx along a given direction x. For example, a rigid body in the gravitational field at the earth surface, which energy remains constant under translations in the x or y horizontal directions. The invariance of H implies

$$0 = \delta H = \frac{\partial H}{\partial x}\,\delta x \quad \Rightarrow \quad \frac{\partial H}{\partial x} = 0 \quad \Rightarrow \quad \dot{p}_x = -\frac{\partial H}{\partial x} = 0 \quad \Rightarrow \quad p_x \text{ is conserved.}$$

The invariance of the Hamiltonian with respect to translations along a given direction implies, actually is equivalent to, the conservation of the corresponding component of the linear momentum:

$$\boxed{\text{Invariance upon translation}} \quad \leftrightarrow \quad \boxed{\text{Linear momentum conservation}}$$

Let us now discuss what happens if the system is invariant upon rotations around a given axis $\hat{n}$. For simplicity we consider a single particle with coordinates $\vec{r} = (x, y, z)$ and momentum $\vec{p} = (p_x, p_y, p_z)$. The changes in $\vec{r}$ and $\vec{p}$ after an infinitesimal rotation with angle $\delta\phi$ around $\hat{n}$ are given by

$$\delta\vec{r} = \hat{n} \times \vec{r}\,\delta\phi$$

and

$$\delta\vec{p} = \hat{n} \times \vec{p}\,\delta\phi.$$

If H is invariant upon rotations around $\hat{n}$ we have

$$\begin{aligned} 0 = \delta H &= \sum_{i=1}^{3} \left(\frac{\partial H}{\partial r_i}\,\delta r_i + \frac{\partial H}{\partial p_i}\,\delta p_i \right) \\ &= -\dot{\vec{p}} \cdot (\hat{n} \times \vec{r})\,\delta\phi + \dot{\vec{r}}_i \cdot (\hat{n} \times \vec{p})\,\delta\phi \\ &= -\delta\phi \left[\hat{n} \cdot (\vec{r} \times \dot{\vec{p}}) + \hat{n} \cdot (\dot{\vec{r}} \times \vec{p}) \right] \\ &= -\delta\phi\,\frac{d}{dt}\,[\hat{n} \cdot (\vec{r} \times \vec{p})] \\ &= -\delta\phi\,\frac{d}{dt}\,(\hat{n} \cdot \vec{L}). \end{aligned}$$

This implies that $\hat{n} \cdot \vec{L}$ is conserved:

$$\boxed{\text{Invariance upon rotation around } \hat{n}} \leftrightarrow \boxed{\text{Conservation of angular momentum } \vec{L} \text{ along } \hat{n}}$$

The fundamental relations between the symmetries of the system or of the Hamiltonian and the conservation laws also hold in quantum mechanics. In fact, the constants of motion are the generators of the infinitesimal transformations which leave the Hamiltonian invariant.

1.5. Many-particle systems

When we consider a system of particles we must distinguish between external forces $\vec{F}_i^{(e)}$, due to the action of external fields applied on each particle i, and internal forces $\vec{F}_{ji}$ due to the interactions between the particles i and j. Newton's second law then takes the form

$$\dot{\vec{p}_i} = \vec{F}_i^{(e)} + \sum_{j \neq i} \vec{F}_{ji},$$

where $\vec{F}_{ji}$ is the force acting on particle i due to particle j. The forces of interest for the following satisfy Newton's third law of action and reaction:

$$\vec{F}_{ij} = -\vec{F}_{ji}$$

and, moreover, the interparticle forces lie along the line joining the particles:

$$\vec{F}_{ij} = (\vec{r}_i - \vec{r}_j)\, f(r_{ij}) \quad \text{where } r_{ij} = \|\vec{r}_i - \vec{r}_j\|.$$

The latter condition is rather strong and does not apply to the electromagnetic forces between moving charges.

1.5.1. Linear momentum

The action-reaction principle (which is of course obeyed by the important electron-electron and electron-nucleus interactions) implies that the time dependence of the total momentum

$$\vec{P} = \sum_i \vec{p}_i$$

is unaffected by the interparticle interactions since

$$\dot{\vec{P}} = \sum_i \dot{\vec{p}_i} = \sum_i \vec{F}_i^{(e)} = \vec{F}^{(e)}. \tag{1.9}$$

The conservation law of the total linear momentum follows: If the total external force is zero the total linear momentum is conserved.

The universal validity of this conservation law can be easily verified over all ranges of length and energy (supernova, firework, nuclear decay, etc.).

We may also use Eq. (1.9) to derive the equation of motion for the center of mass $\vec{R}$ of the many-particle system. Using that

$$\vec{R} = \frac{\sum_i m_i \vec{r}_i}{\sum_i m_i}, \quad M = \sum_i m_i$$

we have

$$\dot{\vec{P}} = \sum_i m_i \ddot{\vec{r}}_i = M\ddot{\vec{R}} = \sum_i \vec{F}_i^{(e)} = \vec{F}^{(e)}.$$

Notice that this equation is not easy to solve when the external force field is inhomogeneous, i.e., $\vec{F}_i = \vec{F}(\vec{r}_i)$.

1.5.2. Angular momentum

The angular momentum of the system is obtained by summing the individual angular momenta of all the particles with respect to the given origin O:

$$\vec{L} = \sum_i \vec{l}_i = \sum_i \vec{r}_i \times \vec{p}_i.$$

Its time dependence is given by

$$\dot{\vec{L}} = \sum_i \left(\dot{\vec{r}}_i \times \vec{p}_i + \vec{r}_i \times \dot{\vec{p}}_i \right) = \sum_i \vec{r}_i \times \dot{\vec{p}}_i = \sum_i \vec{r}_i \times \vec{F}_i.$$

It is meaningful to split the force $\vec{F}_i$ acting on particle i in external and internal contributions:

$$\dot{\vec{L}} = \sum_i \vec{r}_i \times \vec{F}_i^{(e)} + \sum_i \vec{r}_i \times \sum_{j \neq i} \vec{F}_{ji}.$$

Putting together the terms for each pair i and j the second term reads

$$\sum_{i<j} \left(\vec{r}_i \times \vec{F}_{ji} + \vec{r}_j \times \vec{F}_{ij} \right),$$

and using the action-reaction condition on the internal forces $\vec{F}_{ij} = -\vec{F}_{ji}$ we have

$$\sum_{i<j} (\vec{r}_i - \vec{r}_j) \times \vec{F}_{ji}.$$

Finally, for internal forces acting along the line connecting the particles, $\vec{F}_{ji}$ is parallel to $\vec{r}_{ij} = \vec{r}_i - \vec{r}_j$ and thus each term of the sum vanishes ($\vec{F}_{ij} \times \vec{r}_{ij} = 0$). Hence the time dependence of the total angular momentum is determined by the total external torque

$$\vec{N}^{(e)} = \sum_i \vec{r}_i \times \vec{F}_i^{(e)},$$

which results exclusively from the external forces:

$$\dot{\vec{L}} = \vec{N}^{(e)}.$$

The conservation theorem for the angular momentum follows: The total angular momentum of a many-particle system is conserved if the total external torque is zero.

Note that the conservation of angular momentum relies on $\vec{F}_{ji}$ being parallel to $\vec{r}_{ij} = \vec{r}_i - \vec{r}_j$, which does not hold for moving charges. In this case, transfer of angular momentum between the mechanical degrees of freedom and the electromagnetic field is possible, and only the sum of both is conserved.

1.5.3. Energy conservation

In order to introduce the concept of total kinetic energy we consider the work done by all the forces acting on the particles along a path between any two given points in an n-particle trajectory [e.g., $\vec{r}_1(t_1), \dots \vec{r}_n(t_1)$ and $\vec{r}_1(t_2), \dots \vec{r}_n(t_2)$]:

$$\begin{aligned} W_{12} &= \sum_i \int_1^2 \vec{F}_i \cdot d\vec{s}_i = \sum_i \int_1^2 \dot{\vec{p}}_i \cdot \frac{\vec{p}_i}{m_i}\, dt \\ &= \sum_i \int_1^2 \frac{d}{dt}\left(\frac{p_i^2}{2m_i}\right) dt = \int_1^2 \frac{d}{dt}\left(\sum_i \frac{p_i^2}{2m_i}\right) dt \\ &= T_2 - T_1, \end{aligned}$$

where

$$T = \sum_i \frac{p_i^2}{2m_i}.$$

We turn now to the most interesting case of conservative forces. Concerning the external forces the situation is the same as for a single particle, since the force field acts on each particle independently:

$$\vec{F}_i^{(e)}(\vec{r}_i) = -\vec{\nabla}\, v_{\text{ext}}(\vec{r}_i).$$

Consequently, the total potential energy due to the external field is

$$V = \sum_i v_{\text{ext}}(\vec{r}_i).$$

For the interparticle forces we must require $\vec{F}_{ij} = -\vec{F}_{ji}$ and $\vec{F}_{ij} = (\vec{r}_i - \vec{r}_j)\, f(r_{ij})$, where $f(r_{ij})$ depends only on the distance $r_{ij} = |\vec{r}_i - \vec{r}_j|$ between the particles i and j. In order to obtain a scalar function $W(\vec{r}_1, \dots \vec{r}_n)$ such that

$$-\vec{\nabla}_i W = \vec{F}_i^{(\text{int})} = \sum_j \vec{F}_{ji}\ ,$$

we consider first the interaction between two particles i and j and seek for a function $w_{ij}(\vec{r}_i, \vec{r}_j)$ such that

$$-\vec{\nabla}_i w_{ij} = \vec{F}_{ji} = -\vec{F}_{ij} = \vec{\nabla}_j w_{ij}.$$

This can be solved by recalling the calculation of the gradient of a function of the distance $w = w(r)$

$$\frac{\partial w}{\partial x} = \frac{dw}{dr}\,\frac{\partial r}{\partial x} = w'(r)\,\frac{2x}{2\sqrt{x^2+y^2+z^2}} = w'(r)\,\frac{\vec{r}}{r}.$$

Consequently, if we choose

$$w_{ij}(\vec{r}_i, \vec{r}_j) = w_{ij}(r_{ij})$$

with $r_{ij} = |\vec{r}_i - \vec{r}_j|$ we have

$$-\vec{F}_{ji} = \nabla_i w_{ij}(r_{ij}) = w'_{ij}(r_{ij})\,\frac{(\vec{r}_i - \vec{r}_j)}{r_{ij}}$$

and

$$-\vec{F}_{ij} = \nabla_j w_{ij}(r_{ij}) = w'(r_{ij})\,\frac{(\vec{r}_j - \vec{r}_i)}{r_{ij}} = \vec{F}_{ji}.$$

In order to obtain the interaction potential we simply sum over all pairs (i, j):

$$W = \frac{1}{2}\sum_{i\neq j} w_{ij}(r_{ij}) = \sum_{i<j} w_{ij}(r_{ij}).$$

<u>Example:</u> $w_{ij} = \dfrac{q_i\, q_j}{r_{ij}}$ or simply $w(r_{ij}) = \dfrac{e^2}{r_{ij}}$ and $W = \dfrac{1}{2}\displaystyle\sum_{i\neq j}\frac{e^2}{|\vec{r}_i - \vec{r}_j|}$.

Note the factor $\frac{1}{2}$ if the double sum over i and j is used, which results from the fact that each pair of variables ij appears twice, once in $\sum_i$ and once in $\sum_j$. We may explicitly verify that

$$\begin{aligned}
-\frac{\partial W}{\partial x_k} &= -\frac{1}{2}\left(\sum_j \frac{\partial w_{kj}}{\partial x_k} + \sum_i \frac{\partial w_{ik}}{\partial x_k}\right) \\
&= \frac{1}{2}\left(\sum_j \vec{F}_{jk} + \sum_i \vec{F}_{ik}\right) = \sum_j \vec{F}_{jk}.
\end{aligned}$$

Summarizing, the total potential energy

$$V + W = \sum_i v_{\text{ext}}(\vec{r}_i) + \frac{1}{2}\sum_{i\neq j} w_{ij}(|\vec{r}_i - \vec{r}_j|)$$

satisfies

$$-\vec{\nabla}_i(V+W)=\vec{F}_i^{(e)}+\sum_j \vec{F}_{ji}. \tag{1.10}$$

With the help of Eq. (1.10) we can easily determine the time dependence of the total kinetic energy

$$T=\sum_i \frac{p_i^2}{2m_i}$$

as

$$\frac{dT}{dt}=\sum_i \frac{\vec{p}_i}{m}\cdot\dot{\vec{p}}_i=\sum_i -\vec{\nabla}_i(V+W)\cdot\frac{d\vec{r}_i}{dt}=-\frac{d(V+W)}{dt}.$$

Therefore,

$$\frac{d}{dt}(T+V+W)=0,$$

which implies that the total energy, namely, kinetic, plus external potential, plus interaction energy, is a constant of motion. For example, the total energy of a system of n charged particles (e.g., electrons) around the point charge of a nucleus with atomic number Z is given by

$$E=\sum_{i=1}^{n}\frac{p_i^2}{2m}+\sum_i\frac{(-Ze^2)}{|\vec{r}_i|}+\frac{1}{2}\sum_{i\neq j}\frac{e^2}{|\vec{r}_i-\vec{r}_j|}.$$

2. Basic concepts of quantum mechanics

Whenever one attempts to apply classical mechanics to explain atomic or subatomic phenomena, one inevitably comes to very profound contradictions. For instance, in the case of the atom, one would have to conclude that the electrons should fall into the nucleus, since classical charge systems are unstable, and since moving charges in closed trajectories would radiate electromagnetic waves, thereby losing progressively their kinetic energy.

The contradictions are so profound that the formulation of a theory capable of describing microscopic phenomena (i.e., those occuring at very small distances and for particles of very small mass) requires a complete modification of the basic physical concepts and laws. The examples to be discussed below will show that some of the most basic notions of classical mechanics and of our experience with macroscopic phenomena are simply unapplicable to the atomic world. The limitations of classical mechanics are not simple quantitative disagreements but much more fundamental. The main problem is that classical physics does not even provide an appropriate language for describing certain microscopic phenomena, just in qualitative terms.

We will start by discussing a number of experiments that illustrate the concept of wave-particle duality, the complementary principle, the superposition principle, the measurement process in quantum mechanics, and the stochastic nature of the observed events. While the conclusions of these experiments are for the most part negative, in the sense that they show how the classical concepts fail or why they should be abandoned, they also provide some extremely useful insights on the problems that quantum mechanics needs to (and actually does) solve. Moreover, the examples will unravel important clues on how the theory should look like.

2.1. Electromagnetic waves

The equations governing the classical dynamics of electrical and magnetic fields in vacuum are the Maxwell equations

$$\begin{aligned} \vec{\nabla}\cdot\vec{B} &= 0 \qquad & \vec{\nabla}\times\vec{E} + \frac{1}{c}\,\frac{\partial B}{\partial t} &= 0 \\ \vec{\nabla}\cdot\vec{E} &= 0 \qquad & \vec{\nabla}\times\vec{B} - \frac{1}{c}\,\frac{\partial E}{\partial t} &= 0. \end{aligned}$$

With a few simple manipulations one can easily derive the equation for classical electromagnetic waves. Starting from

$$\vec{\nabla}\times\vec{E} + \frac{1}{c}\,\frac{\partial \vec{B}}{\partial t} = 0,$$

take $\vec{\nabla}\times$ to obtain

$$\vec{\nabla}\times(\vec{\nabla}\times\vec{E}) + \frac{1}{c}\,\frac{\partial}{\partial t}\,(\vec{\nabla}\times\vec{B}) = 0.$$

Using a known relation from vector calculus this is written as

$$\vec{\nabla}\,(\vec{\nabla}\cdot\vec{E}) - \nabla^2\vec{E} + \frac{1}{c}\,\frac{\partial}{\partial t}\left(\frac{1}{c}\,\frac{\partial\vec{E}}{\partial t}\right) = 0,$$

which implies

$$\boxed{\nabla^2\vec{E} - \frac{1}{c^2}\,\frac{\partial^2\vec{E}}{\partial t} = 0.}$$

If we focus for simplicity on one component of $\vec{E}$, a scalar electromagnetic field which we denote by ϕ, we have

$$\nabla^2\phi - \frac{1}{c^2}\,\frac{\partial^2\phi}{\partial t} = 0. \tag{2.1}$$

Let us solve this linear homogeneous differential equation with the usual exponential ansatz

$$\phi_{\vec{k}}(\vec{x},t) = e^{i(\vec{k}\cdot\vec{x}-\omega t)},$$

where $\vec{x} = (x_1, x_2, x_3)$ refers to the position vector, $\vec{k}$ to the wave vector, and ω to the angular frequency. Let us see under which conditions $\phi(\vec{x},t) = e^{i(kx-\omega t)}$ satisfies the wave equation:

$$\vec{\nabla}\phi_{\vec{k}} = \vec{\nabla}\left[e^{i(k\vec{x}-\omega t)}\right] = i\vec{k}\,e^{i(k\vec{x}-\omega t)}$$

and thus

$$\vec{\nabla}\cdot(\vec{\nabla}\,\phi_{\vec{k}}) = i\vec{k}\,\vec{\nabla}\left(e^{i(\vec{k}\cdot\vec{x}-\omega t)}\right) = -k^2\,e^{i(kx-\omega t)}.$$

Since $\dfrac{\partial^2}{\partial t^2}\,\phi_{\vec{k}} = -\omega^2 e^{i(\vec{k}\vec{x}-\omega t)}$ we have

$$\nabla^2\phi_{\vec{k}} - \frac{1}{c^2}\,\frac{\partial^2\phi_{\vec{k}}}{\partial t^2} = \left(\frac{\omega^2}{c^2} - k^2\right)e^{i(\vec{k}\vec{x}-\omega t)} = 0,$$

which can only hold if

$$\boxed{\omega = c\,k} \qquad \text{with } k = |\vec{k}|.$$

This is known as the dispersion relation of electromagnetic waves. A dispersion relation gives the energy or frequency of a wave state as a function of the wave vector $\vec{k}$. In the present case the isotropy of space (vacuum) implies that ω depends only on k. Using the relation $\omega = c\,k$ we can write the monochromatic plane wave as

$$\phi_{\vec{k}}(\vec{x},t) = e^{i\,k(\hat{k}\cdot\vec{x}-c\,t)},$$

which explicitly shows that the points of stationary phase propagate with the light velocity $c = \omega/k$ along the direction $\hat{k} = \vec{k}/k$ of the wave vector.

The wave equation is a linear homogeneous equation. This means that if ϕ_1 and ϕ_2 are solutions, then $\alpha\,\phi_1 + \beta\,\phi_2$ is also a solution for any $\alpha, \beta \in \mathbb{C}$. We may then write the most general solution $\phi(\vec{x}, t)$ of the wave equation (2.1) as a linear combination of plane waves $\phi_{\vec{k}} = e^{i(\vec{k}\cdot\vec{x}-\omega t)}$. If $\phi(\vec{x}, t)$ is periodic in space $[\phi(\vec{x}+\vec{a}_i, t) = \phi(\vec{x}, t)$ for $i = 1-3]$ only discrete values of $\vec{k}$ are allowed. For instance, in one dimension $k_n = (2\pi\,n)/a$ with $n \in \mathbb{Z}$ so that $e^{i\,k_n x} = e^{i\,k_n(x+a)}$. However, in the most general case all values of $\vec{k}$ are allowed and the linear combination takes the form of a Fourier integral

$$\phi(\vec{x}, t) = \frac{1}{(\sqrt{2\pi})^3} \int d^3k \; A(\vec{k})\, e^{i(\vec{k}\cdot\vec{x}-\omega t)}. \tag{2.2}$$

The coefficient $A(\vec{k}) \in \mathbb{C}$ represents the weight in amplitude and phase of the plane wave $\phi_{\vec{k}}$ in the electromagnetic wave $\phi(\vec{x}, t)$. Notice that, since $A(\vec{k}) = e^{i\varphi_{\vec{k}}}|A(\vec{k})| \in \mathbb{C}$, any phase shifts $\varphi_{\vec{k}}$ between the plane waves can be taken into account.

The intensity of the radiation at the point $\vec{x}$ and time t is given by the square of the fields $(E^2 + B^2)/8\pi$, which in the present scalar-field approximation corresponds to $|\phi(x, t)|^2$. It is important to notice that, given two waves ϕ_1 and ϕ_2, the superposed solution $\phi_1 + \phi_2$ leads to *interference effects* since the intensity of the superposed waves is not equal to the sum of the intensities:

$$|\phi_1 + \phi_2|^2 \neq |\phi_1|^2 + |\phi_2|^2.$$

We may profit from the general expansion (2.2) to analyze under which conditions a wave packet or light pulse can be localized in some reduced region in space. To this aim we consider a Gaussian wave packet in one dimension:

$$\phi(x, t) = \frac{1}{\sqrt{2\pi\,\Delta k^2}} \int_{-\infty}^{+\infty} dk \;\; \underbrace{e^{-\frac{(k-k_0)^2}{2\Delta k^2}}}_{\substack{A(k)\,\in\,\mathbb{R} \\ \Rightarrow \text{ all plane} \\ \text{waves in phase}}} \; e^{i\,k(x-c\,t)}.$$

We replace for a moment $x - c\,t$ by x and obtain

$$\phi(x, t) = \frac{e^{i\,k_0\,x}}{\sqrt{2\pi\,\Delta k^2}} \int_{-\infty}^{+\infty} dk \;\; e^{-\frac{(k-k_0)^2}{2\Delta k^2}} \; e^{i\,(k-k_0)\,x}$$

and shifting the origin of k to $k - k_0$

$$\phi(x, t) = \frac{e^{i\,k_0\,x}}{\sqrt{2\pi\,\Delta k^2}} \int_{-\infty}^{+\infty} dk \;\; e^{-\frac{k^2}{2\Delta k^2}} \; e^{i\,k\,x}.$$

Taking into account that

$$\frac{k^2}{2\Delta k^2} - i\,k\,x = \left(\frac{k}{\sqrt{2}\Delta k} - i\,\frac{\Delta k\,x}{\sqrt{2}}\right)^2 + \frac{\Delta k^2 x^2}{2}$$

we can write

$$\phi(x,t) = \frac{e^{i\,k_0\,x}}{\sqrt{2\pi\,\Delta k^2}}\, e^{-\frac{\Delta k^2 x^2}{2}} \int_{-\infty}^{+\infty} e^{-\frac{(k-i\,\Delta k^2 x)^2}{2\Delta k^2}} dk.$$

This integral can be solved by changing to the complex variable $z = k - i\,\Delta k^2 x$. Noting that $dz = dk$ we obtain

$$\phi(x,t) = e^{i\,k_0\,x}\, e^{-\frac{\Delta k^2 x^2}{2}}\, \frac{1}{\sqrt{2\pi\,\Delta k^2}} \int e^{-\frac{z^2}{2\Delta k^2}}\, dz = e^{i\,k_0\,x}\, e^{-\frac{x^2\,(\Delta k)^2}{2}}.$$

Finally

$$\phi(x,t) = e^{i\,k_0\,x}\, e^{-\frac{x^2}{2\Delta x^2}},$$

where $\Delta x = 1/\Delta k$. The uncertainty or dispersion Δx in the wave-packet's position is inversely proportional to the uncertainty Δk in the wave vector.

Since the Gaussian packet is an optimal packet in the sense that it reaches the largest localization (smallest Δx) for a given Δk, we conclude that the uncertainty Δk in the value of the wave vector k and in the position of the light pulse (i.e., the photons) are related by

$$\Delta x\,\Delta k \geq 1.$$

For 3 dimensions we can generalize the relation to

$$\Delta x_i\,\Delta k_j \geq \delta_{ij},$$

since the dispersion of k in the direction i has no influence in the localization of $\phi(\vec{x})$ in the directions j with $j \neq i$.

Replacing now x by $x - c\,t$ we obtain

$$\phi(x,t) = e^{i\,k_0\,(x-c\,t)}\, e^{-\frac{(x-c\,t)^2}{2\Delta x^2}}.$$

One observes that the Gaussian wave packet propagates undistorted with the speed c, since the dispersion relation is strictly linear in vacuum: $\omega = c\,k$.

It is interesting to relate the uncertainty or width Δx of the wave packet with the duration Δt of the pulse at a given fixed point in space. For a wave packet with the dispersion Δk the dispersion in frequency is $\Delta\omega = c\,\Delta k$. Moreover, the spatial extension of the packet is

$$\Delta x \geq \frac{1}{\Delta k} = \frac{c}{\Delta\omega}$$

$$\Delta x\,\frac{\Delta\omega}{c} \geq 1$$

$$\frac{\Delta x}{c}\,\Delta\omega \geq 1.$$

The duration of the passage of such a packet at any given point is

$$\Delta t = \frac{\Delta x}{c} \geq \frac{1}{c}\,\frac{c}{\Delta\omega} = \frac{1}{\Delta\omega}.$$

Consequently,

$$\Delta t \, \Delta\omega \geq 1.$$

One concludes that it is impossible to ascribe a well-defined position ($\Delta x = 0$) and a well-defined wave vector ($\Delta k = 0$) or frequency ($\Delta\omega = 0$) to a photon. Moreover, it is not possible to define the frequency ($\Delta\omega = 0$) and the precise point in time of the passage of a pulse ($\Delta t = 0$) at any given point. To be strict one should replace "photon" by "electromagnetic wave" or "ensemble of photons" in the previous statement. These uncertainty relations $\Delta k \, \Delta x \geq 1$ are a typical manifestation of wave properties. They are inherited by quantum mechanical particles with a finite rest mass (e.g., electrons) due to wave–particle duality.

2.2. Non-classical aspects of the electromagnetic field: Photons

2.2.1. The photoelectric effect

The interactions between light and electrons have provided many important clues in the development of quantum theory. The first manifestation of the quantized, particle-like nature of light is the photoelectric effect. The main experimental observations are the following:

i) The rate of electron emission is proportional to the intensity of radiation, i.e., to $|\phi|^2$.

ii) For each metal there is a threshold frequency ω_c such that for $\omega < \omega_c$ no emission at all is observed despite intensity increase (within reasonable limits).

iii) For $\omega > \omega_c$, the largest kinetic energy $E_{\text{kin}}^{\text{max}}$ of the emitted electrons depends linearly on ω, but not on the intensity of the radiation $|\phi|^2$.

iv) The time delay between the start of the incident radiation and electron emission is very short ($< 10^{-9}$ s).

These results are incompatible with the classical theory of electromagnetic radiation. All the observations indicate that the energy is not transferred by the field $|\phi^2|$ but rather that the process occurs in a quantized way.

Einstein's explanation proposes that the emitted electron is scattered by a quantum of light known as photon thereby receiving momentum and energy. The energy E of the photon is related to its frequency by

$$E = \hbar\,\omega$$

and its momentum $\vec{p}$ is given by

$$\vec{p} = \hbar\,\vec{k},$$

where $\hbar = 1.054 \times 10^{-34}$ Joule $\cdot$ s $= 0.66$ eV $\cdot$ fs is *Planck's constant* ($\hbar = h/2\pi$). $\hbar$ has the units of action or of angular momentum.

The intensity of the light corresponds to the number of emitted photons per unit time. The energy of each photon is independent of intensity, it depends only on the frequency ω or the wave length $\lambda = 2\pi/k = 2\pi c/\omega$.

Since $\omega = c\,k \;\Rightarrow\; E = \hbar\,\omega = c\,\hbar\,k = c\,p$. Recalling the relativistic energy-momentum relation $E^2 = p^2c^2 + m^2c^4$ one concludes that photons have vanishing rest mass m. The explanation is as follows:

i) $E_{\text{kin}}^{\text{max}} = \hbar\,\omega - W$, where W is the metal's work function. If $E_{\text{kin}}^{\text{max}} < 0 \;\Rightarrow$ no emission whatever the intensity, since the electrons cannot escape from the metal.

ii) $E_{\text{kin}}^{\text{max}} = \hbar\,\omega - W$ is independent of the number of photons, i.e., of the radiation intensity.

iii) The intensity of the radiation defines the number of photons, which controls the number of emitted electrons.

iv) A single photon can emit an electron. Therefore, no delay in accumulating absorbed energy is involved. According to the classical picture one would have to integrate $|\phi|^2$ during some time before emission can occur, and this time delay would be inversely proportional to the intensity $|\phi|^2$.

2.2.2. The Compton effect

The Compton effect provides a number of additional important clues. Consider a wave packet with small Δk (large spatial extension), i.e., essentially monochromatic. Classically one expects that the electron would gain momentum in the direction $\vec{k}$ of the incident radiation and that the light would be scattered in the form of spherical waves. However, Compton X-ray scattering experiments show the following:

i) The scattered electron often acquires a momentum $\vec{p}$ transversal to the incident $\vec{k}$.

ii) There is no sign of a spherical outgoing wave.

iii) The scattered light is concentrated in a spatially confined packet.

iv) The propagation direction of the scattered light is correlated with the momentum vector $\vec{p}$ of the scattered electron.

v) The wave length and consequently the energy of the scattered light depends on the light's scattering angle θ as

$$\omega' = \omega \left(1 + \frac{2\hbar\,\omega}{mc^2}\, \sin^2 \frac{\theta}{2} \right)^{-1}$$

or equivalently

$$\lambda' - \lambda = 2\lambda_c \sin^2 \frac{\theta}{2},$$

where $\lambda_c = \dfrac{2\pi\,\hbar}{mc}$ is the Compton wave length.

This experiment can be explained by assuming that photons scatter with electrons as individual particles like billard balls, i.e., following energy and momentum conservation according to Einstein's relations $\vec{p} = \hbar\,\vec{k}$ and $E = \hbar\,\omega = c\,p$ for photons, and $E^2 = c^2p^2 + m^2c^4$ for electrons.

In the photoelectronic effect, the need for the quantization of the energy

$$E = \hbar\,\omega$$

of the electromagnetic radiation became clear. This established the relation between energy and frequency of electromagnetic oscillations given by

$$e^{i\,\omega\,t} = e^{i\,\frac{E}{\hbar}\,t}.$$

For the explanation of the Compton effect we need to assume the relation

$$\vec{p} = \hbar\,\vec{k}$$

between the wave vector $\vec{k}$ and momentum $\vec{p}$ of a quantum of radiation. This establishes the relation between momentum and the oscillations of the electromagnetic wave in space given by

$$e^{i\,\vec{k}\cdot\vec{x}} = e^{\frac{i}{\hbar}\,\vec{p}\cdot\vec{x}}.$$

Consequently there is a relation between momentum and wave length:

$$\lambda = \frac{2\pi}{k} = \frac{2\pi\,\hbar}{p}.$$

Since $\omega = c\,k$, we also have $E = c\,p$.

The correlations between the light-scattering direction and its frequency, as well as between the directions of the scattered particles, can be explained by assuming that the interaction occurs between a single photon and an electron, and that the total energy and momentum are conserved in the process. Let us assume that the direction of the wave vector of the incident (scattered) photon is along $\vec{k}$ ($\vec{k}\,'$), and that the electron is initially at rest ($\vec{p} = 0$). Energy and momentum conservation imply

$$\hbar\,\omega + mc^2 = \hbar\omega\,' + \sqrt{m^2c^4 + c^2p'\,^2} \tag{2.3}$$

and

$$\hbar\,\vec{k} = \hbar\,\vec{k}\,' + \vec{p}\,', \tag{2.4}$$

where $\vec{p}\,'$ refers to the momentum of the scattered electron. From Eq. (2.4) we have $\vec{p}\,' = \hbar\,(\vec{k} - \vec{k}\,')$ and therefore

$$\begin{aligned} p'^{\,2} &= \hbar^2\,(\vec{k} - \vec{k}\,')^2 = \hbar^2\,(k^2 + k'^{\,2} - 2\vec{k}\cdot\vec{k}\,') \\ &= \hbar^2\left[(k-k')^2 + 2\,k\,k'\,(1-\cos\theta)\right] \\ &= \hbar^2\left[(k-k')^2 + 4\,k\,k'\,\sin^2\left(\frac{\theta}{2}\right)\right]. \end{aligned} \tag{2.5}$$

Moreover, Eq. (2.3) can be written as

$$\hbar\,c\,(k-k') + mc^2 = \sqrt{m^2c^4 + c^2p'^{\,2}}. \tag{2.6}$$

Taking the square on both sides of Eq. (2.6) and replacing $p'^{\,2}$ by Eq. (2.5) one obtains

$$mc\,(k-k') = 2\hbar\,k\,k'\,\sin^2\frac{\theta}{2}.$$

Finally, introducing the Compton wavelength $\lambda_c = \dfrac{2\pi\,\hbar}{mc}$ we have

$$2\pi\left(\frac{1}{k'} - \frac{1}{k}\right) = 2\,\lambda_c\,\sin^2\frac{\theta}{2}$$

or

$$\lambda' - \lambda = 2\,\lambda_c\,\sin^2\frac{\theta}{2}.$$

A number of nontrivial conclusions can be inferred from the previous partial explanation of the effect:

1) The concept of particles or photons and waves $e^{i\,(\vec{k}\cdot\vec{x} - \omega t)}$ are linked by Einstein's relations
$$E = \hbar\,\omega \quad \text{and} \quad \vec{p} = \hbar\,\vec{k}.$$
In other words, we attach an energy quantum to the electromagnetic field and a momentum $\vec{p}$ to a wave vector. This *wave-particle duality* is a concept that does not exist at all in classical physics. In terms of Bohr's complementary principle one would say that a single classical concept (in this case wave- or particle-like aspects) is not enough to describe atomic phenomena.

2) Another most remarkable feature of the experiment is the stochastic nature of the outcome. It seems that there is no way one could control the momentum of the scattered electron or the wave length of scattered light. Compton incorporates part of the statistical aspect in the interpretation by establishing a relation between the scattered angle and the frequency of the light. But this first analysis is clearly incomplete, since it cannot predict the *probability* for each scattering angle.

3) Classical theory, in contrast, is completely useless, since it would conclude that the initial conditions (which are all the same for all photons) completely define the outcome of the scattering process. The concepts of classical theory are inappropriate since they do not incorporate the statistical aspects of the observed phenomena.

4) Already at this point one may ask oneself a kind of philosophical question: What is the origin of the statistical nature of the experimental outcome? Is it the quantum mechanical state (whatever this is) that has a statistical, somehow non-deterministic nature? Or are the statistical results of the experiment only a consequence of the measurement process, i.e., of the interaction between the quantum mechanical state and the macroscopic world or apparatus that detects scattering angle and wavelength, so that a human can read it? To put it in Einstein's terms, is God playing dices with nature?

2.2.3. Young interference experiment

One of the most remarkable features of quantum physics is the so-called *wave-particle duality* which states that particle and wave aspects of quantum phenomena (in the present case light) are indivisible. In other words, if one attempts to interpret quantum phenomena with classical concepts corresponding to the notions of particles and waves, one realizes that both concepts are needed at the same time. Since these are classically incompatible, it is clear that the classical interpretation and language are inappropriate.

Young's interference experiment for light, and the similar phenomenon of electron diffraction observed for particles having a nonvanishing rest mass, illustrate the problem very clearly. We consider the following arrangement:

(Bild)

The main qualitative observations are the following:

1) If S_2 is closed, one observes a diffraction spot $I_1(x)$.

2) If S_1 is closed and S_2 open, one sees $I_2(x)$.

3) If both sides are open, one sees an interference pattern $I(x) \neq I_1(x) + I_2(x)$.

How to explain this with the particle picture put forward by the explanation of photoelectric effect and Compton's experiment?

1) One could attempt to interpret the single-slit diffraction spot $I_1(x)$ or $I_2(x)$ classically in terms of collisions of the particles (photons) during the passage through the slit. This is actually not quite satisfactory in detail, but anyway this is not so crucial.

2) Let us focus on the interference effect that appears when both slits are open. How can one explain that a particle, which in a classical picture follows a trajectory going, for example, through S_1, knows whether S_2 is open or not?

   One could try to argue that the interference is due to interactions of the photons going through S_1 with photons going through S_2. However, if one reduces the intensity, so that only one or a few photons pass at a time and there is no possible interaction, and one increases proportionally the exposure time, the interference pattern remains. Consequently, interference is not a result of photon-photon interactions.

The observed effect is therefore incompatible with the idea that particles go through one slit or the other. The interference phenomenon is thus *incompatible with the classical notion of path.*

However, we all know that the interference pattern can be easily interpreted by using the notion of waves and the superposition principle, which holds for any linear homogeneous differential equation such as the wave equation

$$\nabla^2\phi - \frac{1}{c^2}\,\frac{\partial^2\phi}{\partial t^2} = 0.$$

The wave picture of light implies that the intensity of the photon beam is proportional to the square of the electric and magnetic fields

$$|E|^2 \quad \text{or} \quad |\phi|^2.$$

The intensity at any point in the screen is given by the square of the sum or superposition of the fields generated by each slit S_1 and S_2 which act as secondary sources:

$$E(x) = E_1(x) + E_2(x) \quad \text{with } E_i = e^{i\vec{k}\cdot(\vec{r}_i - \vec{R}_i)},$$

and

$$\begin{aligned} I(x) &= |E(x)| = |E_1(x) + E_2(x)|^2 \\ &= |E_1(x)|^2 + |E_2(x)|^2 + \underbrace{2\,\mathrm{Re}\{E_1(x)\,E_2^*(x)\}}_{\text{interference term}} \\ &\neq |E_1|^2 + |E_2|^2. \end{aligned}$$

There is no interaction at all between the fields E_1 and E_2, just linear superposition.

You could pragmatically say: “Here the wave explanation works, so maybe we can forget about the particle picture for this experiment (despite Compton’s experiment). Maybe something different happened there.”

But then you decide to look in more detail at the one-photon-at-a-time version of Young’s experiment and you see what’s really going on.

If you reduce the intensity to the limit of one or a few photons at a time and keep the exposure time relatively short, so that only one or a few photons are emitted, you see *no interference pattern* in the photographic plate, but *isolated single-photon dots.* This is completely incompatible with the wave picture which predicts an interference for all intensities, i.e.,

$$I \propto |E(x)|^2.$$

The individual dots follow a random distribution with no correlation between the positions corresponding to successive photons. We recover here Compton's randomness of successively scattered photons.

Only when we increase the exposure time, so that many photons have arrived per unit area and one can no longer distinguish or resolve individual dots, we recover the interference pattern as predicted by the classical wave theory

$$I = |E(x)|^2 = |E_1(x) + E_2(x)|^2.$$

We must conclude that the wave theory, which follows superposition, only predicts the *probability density* for a light quantum or photon to hit the screen at a given point x. In fact $|E(x)|^2$, which is the result of the superposition of $E_1(x)$ and $E_2(x)$, gives the probability density.

While the idea that the classical field gives the probability amplitude is attractive, it doesn't solve the problem that a photon going through S_1 behaves differently if S_2 is open or closed. So we may want to check if the photon does get through one slit or the other. To this aim we put detectors just after the slits. One then observes 50% of the counts on S_1, 50% on S_2, and *never* a count on both. In other words, if we measure the position of the photon we get a well-defined value.

What happens if we put a detector only behind S_2 (that absorbs of course all the photons going through S_2), so that we know that the rest of the photons go certainly through S_1. In this case, as expected, we recover the single-slit result $I_1(x)$ with no interference at all.

At this point we can draw a few important conclusions:

1) If the photons are allowed to traverse different paths, we *either* remain ignorant about which path the quantum particles have traversed and observe an interference pattern, *or* we experimentally determine (i.e., "measure") which was the path that has actually been taken and then lose the interference pattern.

   One often summarizes this by the *complementarity principle* which states that a measurement designed to manifest one classical attribute (e.g., a wave or particle aspect) precludes the possibility of observing the other or at least part of the other classical attribute. *A measurement affects the quantum state of a microscopic particle in an essential way.*

   This is not what happens in classical physics. In classical mechanics the particles have their own dynamical variables (position, momentum, etc.), irrespectively of the measurement, which can in principle be as soft as wished.
   (Example: Photons $\leftrightarrows$ Train.)

2) We must conclude that a photon follows no trajectory. It has *no intrinsic dynamical variables* like position or momentum. These variables appear only as a result of a *measurement.*

3) The photons make random individual spots on the screen, so that there is no possibility of predicting the outcome of a single measurement. The experiment can only attempt to determine the probabilities for individual events. This implies recording a large number of quantum events.

2.3. Non-classical properties of particles having finite rest mass: The uncertainty principle

While the properties of photons and the interaction of light with matter are both interesting and clarifying, a detailed development of the quantum theory of radiation cannot be discussed without a previous background on non-relativistic quantum mechanics. From now on we focus therefore on the properties of electrons, which we will loosely take as a synonym of a quantum object with a non-vanishing rest mass.

We have already pointed out the impossibility of understanding the stability of atoms in a classical framework. For the formulation of quantum mechanics it is also important to realize that electrons show (like photons) a number of features belonging to wave physics. In particular electrons diffract displaying interference patterns when they pass through crystals, in a completely analogous way as electromagnetic waves.

Let us consider an idealized version of the electron diffraction experiment. In the two-slit experiment

(Bild)

one observes the same phenomena found for electromagnetic waves. As already discussed for photons, this result is incompatible with the idea that electrons follow a path. The lack of path or trajectory in quantum dynamics is one of the manifestations of Heisenberg's uncertainty principle, which is probably the most fundamental concept in quantum mechanics.

The fact that electrons lack of a path also implies that they have *no intrinsic dynamical variables* in a classical sense. In particular it means that an electron has no intrinsic value of the position, velocity or momentum. These dynamical variables can only appear as the result of a *measurement.* A measurement is the result of the interaction of an electron with a classical object (typically, but not necessarily a macroscopic object like a photomultiplier detector). These classical objects are called *apparatus.* As a result of a measurement, i.e., as a result of the interaction of the electron and the classical object, the state of the apparatus and the state of the electron change. These changes depend on the initial state of the electron and can thus be used to characterize its state quantitatively. For example, a detector set behind the slits S_1 or S_2 allows one to define the position of the electron.

Notice that the notion of measurement as interaction is completely independent of the presence of a human observer. Moreover, the apparatus need not be macroscopic. It should simply follow classical mechanics to a sufficiently high accuracy. For instance the vapor molecules in a Wilson chamber, which condense to a thin but macroscopic cloud upon the passage of an electron, allow one to determine the electron's path to a low degree of accuracy. As we shall see, momentum or velocity and position can be simultaneously determined provided that both measurements are done with a limited accuracy. Precise simultaneous measurements of both position and momentum are of course not possible, since this would imply that the electron follows a trajectory.

We see that classical mechanics is not only a limiting case of quantum mechanics. It is also needed to formulate the connection between the quantum and the macroscopic worlds, i.e., the relation between the quantum state and the results of measurements. A typical problem in quantum mechanics is to predict the result of a measurement on the basis of the result of a previous measurement. The first measurement (i.e., the interaction with a macroscopic classical object) is often referred to as a “preparation” of the system. An example would be an electron going through a slit with a given kinetic energy. The presence of a slit defines the position of the electron, even if no human is looking at it. Moreover, quantum mechanics must be able to predict the possible values of a given measurement which, as we shall see, are often restricted. For example, part of the energy levels of an atom are discrete.

The measurement is an interaction and therefore affects the state of the electron. This effect becomes increasingly important with increasing precision of the measurement. This might remind you of the analysis of the electromagnetic wave packet, for which a well defined wave length $\Delta k \to 0$ implies an undefined position $\Delta x \to \infty$, and vice versa. For a given accuracy of the measurement the effect of the measurement cannot be made arbitrarily small. If this would be possible, it would mean that this property is an intrinsic property of the electron (e.g., its mass or charge) and not a dynamical variable. The situation is conceptually completely different from what one is used to in classical physics.

Among the different physical observables, the measurement of the coordinates of the particle (position) plays a fundamental role. Successive measurements of the position of an electron do not lie on a straight line. However, as the delay Δt between consecutive measurements is reduced ($\Delta t \to 0$), the distance between the recorded positions tends to vanish. This implies that a position measurement is reproducible. No velocity can be inferred as the limit of $\vec{x}_{i+1} - \vec{x}_i$ for $\Delta t \to 0$. It is not possible to measure position and velocity at the same time, since otherwise the electron would have a trajectory. One concludes that position and velocity cannot be measured simultaneously. However, inexact measurements of both $\vec{x}$ and $\vec{p}$ (or velocity) are possible.

Our previous analysis of a wave packet in the context of classical electromagnetic waves has shown that the uncertainty Δk in the wave vector and the uncertainty Δx in the position of a Gaussian wave packet are related by

$$\Delta k \, \Delta x = 1.$$

Using the de Broglie-Einstein relation between wave vector (or wave length $\lambda = 2\pi/k$) and momentum

$$\vec{p} = \hbar \vec{k}$$

we have for an hypothetical Gaussian electronic wave packet

$$\Delta x \, \Delta p = \hbar.$$

This provides a mathematically more precise statement of Heisenberg's uncertainty principle. It quantifies to what extent simultaneous measurements of position and momentum with a limited degree of accuracy are possible. Notice in particular that $\Delta x \to 0$ implies $\Delta p \to +\infty$ and vice versa.

Later on, as soon as the mathematical formulation of quantum mechanics is available, we shall define

$$\Delta x = \sqrt{\langle\, (x - \langle\, x \,\rangle)^2 \,\rangle}$$

and

$$\Delta p = \sqrt{\langle\, (p - \langle\, p \,\rangle)^2 \,\rangle},$$

and we shall prove that

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

for any quantum mechanical state.

The uncertainty principle is in my views the most important single fundamental concept in quantum mechanics. The reader may wish to stop here and think for himself about the numerous physical implications of the simple relation $\Delta p \, \Delta x \sim \hbar$. Why do electrons diffract through a small opening? Why do electrons not fall to the nucleus, whatever large the nuclear charge is, as a meteorite would fall into the sun? Why do fast electrons look as if they would follow a path in a Wilson chamber? Why is the uncertainty principle irrelevant in the classical macroscopic world, e.g., for cannon shots or for the motion of the moon around the earth? Why is it possible to use electrons to create a moving image in a television set? Is it in principle possible to record a film of a chemical reaction or of a vibrating molecule? If yes, under which conditions? And finally, are two identical quantum mechanical particles (e.g., two electrons) distinguishable or not? Could one devise an experiment to attach a label to an electron to distinguish it from another electron, as one can do with any classical object?

A measurement on a quantum object changes its state. It is therefore not possible in general to predict with certainty the result of a subsequent measurement on the basis of a previous one. We have discussed the example of subsequent measurements of the position. One can only predict the probability for a given result of a measurement, for instance, the probability for an electron to hit the screen at a given point. There are two kinds of measurements in quantum mechanics: those which do not lead with certainty to a given result and those for which the result can be predicted precisely. In the latter case we say that the associated physical quantity has a *definite value*. The measurement of a set of physical quantities, also called *observables*, which have simultaneously definite values and that are such that no other observable can have a definite value at the same time are called *complete set of observables*. A measurement of a complete set of observables defines the quantum state completely, irrespectively of the history of the electron prior to the measurement.

In the following we will focus on complete defined states. Moreover, we consider an elementary particle, i.e., a particle with no internal structure, for which the measurement of the position (x, y, z) in $\mathbb{R}^3$ constitutes a complete set of observables. For all practical purpose this will be an electron ignoring for the moment its intrinsic spin.

3. Fundamental mathematical formalism

The experiments discussed in the previous chapter have illustrated the problems faced by quantum mechanics, the inadequacy of classical concepts such as trajectory and intrinsic dynamical variables, and the concept of measurement in quantum mechanics as an interaction between quantum objects and a classical apparatus. They also provided a number of hints on the probabilistic nature of the outcome of experiments, as well as on the mathematical structure of the theory. For instance, the experiments suggest the notion of interference and superposition, and the relation between momentum and wave length (or wave vector) even for particles having a non-vanishing rest mass (e.g., electrons). We are thus ready to present the mathematical formalism of quantum mechanics. Clearly, a new theory requires new assumptions that cannot be derived from a less general theory. We intend to present these basic assumptions or "postulates" clearly as such. At the same time we shall attempt to provide plausibility arguments based on intuition and the previous experimental discussion.

3.1. The wave function

In the following we shall denote by q the coordinates of the system and by dq the volume element in the coordinate space. Thus, for an electron in one dimension (1D) $q = x$ and $dq = dx$, while in 3D we have $q = \vec{x}$ (or $q = \vec{r}$) and $dq = d^3x$ (or $dq = d^3r$). For a many-particle system $q = (\vec{x}_1, \vec{x}_2, \ldots \vec{x}_n)$ and $dq = d^3x_1, \ldots d^3x_n$.

The fundamental mathematical formulation of the theory relies on the fact that the *state of a quantum mechanical system* at a given time t is completely described by a definite in general complex function

$$\Psi\,(q,t)\,\in \mathbb{C}$$

of the coordinates. The wave function contains the complete information of the physical state. The answer to any question we may ask about the system is contained in $\Psi(q)$. This means that the knowledge of $\Psi(q)$ allows one to predict the probability for each of the various results of the measurement of any other observable (e.g., the momentum $\vec{p}$ of the particle).

The state of the system and thus the wave function $\Psi(q)$ vary in general with time [i.e., $\Psi(q) = \Psi(q,t)$]. According to our first postulate $\Psi(q,t_0)$ describes the quantum state completely at time t_0. Therefore, the knowledge of $\Psi(q,t_0)$ at t_0 must suffice to predict the state of the system and thus $\Psi(q,t)$ at any future time $t > t_0$. Mathematically, this implies that the equations governing the time evolution of $\Psi(q,t)$ must contain at most first-order derivatives with respect to time. If higher derivatives would be involved (like in Newton's equations) one would also need to know $\dot{\Psi}(q,t_0)$ to predict $\Psi(q,t)$ for $t > t_0$. Conversely, if the time-evolution equation involves only first-order time derivatives, the knowledge of $\Psi(q,t_0)$ implies that $\Psi(q,t)$ is univocally defined for all future times. Therefore, once the actual state of the system is defined by $\Psi(q)$, quantum mechanics is absolutely deterministic in the sense that the state $\Psi(q,t)$ is perfectly known to every possible detail at any future time t. As already discussed, it is the outcome of a measurement that is stochastic in nature, but this has nothing to do with time evolution. It is a consequence of the nature of the measurement process and the interaction between the quantum object and a classical apparatus.

3.2. The superposition principle

Once the notion of wave function and quantum state has been introduced we may turn to the chief principle of quantum mechanics: The superposition principle.

Suppose that at time t the system can be in a state 1 described by the wave function $\Psi_1(q)$ and suppose there is another possible state 2 described by $\Psi_2(q)$, we postulate that every function of the form

$$\Psi(q) = c_1\,\Psi_1(q) + c_2\,\Psi_2(q) \tag{3.1}$$

with c_1 and c_2 arbitrary constants in $\mathbb{C}$ describes a possible state of the system. $\Psi(q)$ is said to be a linear combination of Ψ_1 and Ψ_2. Note that c_1 and c_2 are independent of q and t. Moreover, if we know the time dependence of the state of the system when it is in the states 1 and 2 [i.e., $\Psi_1(q,t)$ and $\Psi_2(q,t)$], then the time dependence of the combined state is

$$\Psi(q,t) = c_1\,\Psi_1(q,t) + c_2\,\Psi_2(q,t). \tag{3.2}$$

In other words the linear relation (3.1) holds for all times. Remember that c_1 and c_2 are independent of t. An immediate consequence of the principle of superposition is that all the equations satisfied by any wave function $\Psi(q,t)$ must be linear in Ψ and homogeneous.

Conversely, if the time-evolution equation of the wave function is linear and homogeneous, then the superposition principle holds.

The previous statement can be immediately generalized to n wave functions. Given the possible states $\Psi_1(q), \dots \Psi_n(q)$ of a system, then

$$\Psi(q) = \sum_n a_n\,\Psi_n(q) \tag{3.3}$$

is a possible state for arbitrary $a_n \in \mathbb{C}$. If the set of functions $\Psi_1, \dots \Psi_n$ is such that *any* state of the system can be written in the form (3.3) we say that the set of functions is *complete* or *closed*.

From a mathematical perspective the first part of the superposition principle is equivalent to asserting that the wave functions describing the states of a quantum system form a linear vector space over the field of complex numbers $\mathbb{C}$. The physical principle asserts merely closure, while the other axioms of the definition of vector space follow immediately from the properties of $\mathbb{C}$. In mathematical language the postulate concerning the superposition of time dependences is equivalent to requiring that the time-evolution operator $\hat{U}(t,t_0)$ is linear. The definition of $\hat{U}(t,t_0)$ is

$$\Psi(q,t) = \hat{U}(t,t_0)\,\Psi(q,t_0), \tag{3.4}$$

which simply expresses that $\hat{U}(t, t_0)$ connects or propagates the wave function at two different times t_0 and t. The linearity of $\hat{U}(t, t_0)$ means precisely

$$\hat{U}(c_1\,\Psi_1 + c_2\,\Psi_2) = c_1\,\hat{U}\,\Psi_1 + c_2\,\hat{U}\,\Psi_2. \tag{3.5}$$

Replacing Eq. (3.1) in (3.4) and using (3.5) one obtains indeed (3.2).

At this point it is useful to clarify the abstract notion of *quantum mechanical state* and to provide an appropriate notation for it. From vector algebra we are used to the idea that a vector $\vec{v}$ (for instance the position vector $\vec{r}$), although it can be represented in different orthonormal bases (coordinate systems), always has a well-defined physical meaning as a point in space, which is independent of the basis choice. This imposes precise relations between the coordinates in different bases:

$$\vec{r} = \sum_i x_i\,\hat{e}_i = \sum_i x'_i\,\hat{e}'_i,$$

where x_i (x'_i) are the coordinates in the basis $\{\hat{e}_i\}$ ($\{\hat{e}'_i\}$). Using the properties of the scalar product: $\hat{e}_i \cdot \hat{e}_j = \hat{e}'_i \cdot \hat{e}'_j = \delta_{ij}$ we obtain the coordinates as

$$x_i = \vec{r} \cdot \hat{e}_i \qquad \text{and} \qquad x'_i = \vec{r} \cdot \hat{e}'_i.$$

In quantum mechanics the physical states of the system, which vector properties are guaranteed by the superposition principle, are also independent of the basis, i.e., of the complete set of observables used for defining them. The different choices of complete sets of observables are called representations. Besides the coordinate representation based on the coordinate wave function $\Psi(\vec{x})$ one may consider, for example, the momentum representation based on the momentum wave function $\Phi(\vec{p})$. It is therefore useful to introduce the notion of quantum mechanical state in a representation-independent form which, following Dirac, we denote by the *ket*

$$|\,\Psi\,\rangle.$$

The superposition principle can be stated as

$$|\,\Psi\,\rangle = \sum_n a_n\,|\,n\,\rangle,$$

where $|\,n\,\rangle = |\,\Psi_n\,\rangle$ stands for the *ket* or *vector state* associated to the wave function $\Psi_n(q)$. In this framework the wave function is denoted by

$$\langle\,q\,|\,\Psi\,\rangle = \Psi(q).$$

As a result of the superposition principle it is easy to see that the application

$$\langle\,q\,| : \{\text{vector space}\,|\,\Psi\,\rangle\} \to \mathbb{C}$$

that associates the wave function $\Psi(q)$ to the vector state $|\,\Psi\,\rangle$, namely $|\,\Psi\,\rangle \to \langle\,q\,|\,\Psi\,\rangle = \Psi(q)$, is linear in $|\,\Psi\,\rangle$:

$$\Psi(q) = \langle\,q\,|\,\Psi\,\rangle = \langle\,q\,|\sum_n a_n\,|\,n\,\rangle = \sum_n a_n\,\Psi_n(q) = \sum_n a_n\,\langle\,q\,|\,n\,\rangle.$$

Later on we shall show that $\langle\, q \,|\, \Psi \,\rangle$ has the properties of an inner product in the complex vector space of quantum mechanical states $\{|\, \Psi \,\rangle\}$.

Analogously the momentum-space wave function associated with the state $|\, \Psi \,\rangle$ is given by

$$\langle\, \vec{p} \,|\, \Psi \,\rangle = \Phi(\vec{p}).$$

Of course, $\langle\, \vec{p} \,|\, n \,\rangle = \Phi_n(\vec{p})$ is the momentum-space wave function associated to the state $|\, n \,\rangle$ and

$$\Phi(\vec{p}) = \sum_n a_n \, \Phi_n(\vec{p})$$

is the momentum-space wave function associated to $|\, \Psi \,\rangle = \sum_n a_n \,|\, n \,\rangle$.

3.3. The wave function and the measurements of the coordinates

The connection between the wave function $\Psi(q)$ and the measurement of the coordinates q of the system is given by the third and last fundamental postulate of quantum mechanics: The probability P that a measurement of the coordinates of the system yields values q' in the volume element dq around q is

$$P\{q' \in dq @ q\} = |\Psi(q)|^2 \, dq.$$

One says that $|\Psi(q)|^2$ represents the *probability density* that a measurement of the coordinates yields the value q.

The definition of probability requires that the sum of the probabilities of all possible events be equal to one. This means

$$\int |\Psi(q)|^2 \, dq = 1,$$

where the integration runs over all space. This is known as the *normalization condition* for the wave function. If the integral $\int |\Psi|^2 \, dq$ converges, one can always normalize the wave function by multiplying it by an appropriate constant. However, there are situations where $\int |\Psi|^2 \, dq$ diverges (e.g., for plane waves or for eigenfunctions of continuous spectra). In these cases $|\Psi|^2$ does not represent the probability density, but nevertheless the ratio $|\Psi(q)|^2/|\Psi(q')|^2$ always gives the relative probability for a measurement at any two points q and q'.

We can now use the probability distribution $|\Psi(q)|^2$ in order to compute mean values, also known as *expectation values*, of any function of the coordinates. Assuming that the wave function is properly normalized ($\int |\Psi(\vec{x})|^2 \, d^3x = 1$) the expectation value of the position is given by

$$\langle\, \vec{x} \,\rangle = \int |\Psi(x)|^2 \, \vec{x} \, d^3x$$

and the uncertainty in the position by

$$\Delta x^2 = \langle\, (\vec{x} - \langle\, \vec{x} \,\rangle)^2 \,\rangle = \langle\, x^2 \,\rangle - \langle\, \vec{x} \,\rangle^2,$$

where

$$\langle\, x^2 \,\rangle = \int |\Psi(\vec{x})|^2\, x^2\, d^3x.$$

As usual $\vec{x} = (x_1, x_2, x_3)$ and $x^2 = \sum_i x_i^2$. One can also calculate the average interaction energy between an electron and the nucleus as

$$\begin{aligned} \langle\, V(r) \,\rangle = \Big\langle -\frac{Ze^2}{r} \Big\rangle &= -\int \frac{Ze^2}{r}\, |\Psi(\vec{r})|^2\, d^3r \\ &= -4\pi\, Ze^2 \int_0^{+\infty} r\, |\Psi(r)|^2\, dr \end{aligned}$$

where we have assumed for simplicity that $|\Psi(\vec{r})|^2$ is spherically symmetric (i.e., $|\Psi(\vec{r})|^2 = |\Psi(r)|^2$). In this case the normalization condition reads $4\pi \int_0^{+\infty} r^2\, |\Psi(r)|^2\, dr = 1$.

3.4. Observables

We now turn to the question of the probability of measuring an arbitrary physical property which in quantum mechanics are known as *observables.*

Let us consider a physical quantity f. The values that a physical quantity or observable can take are called *eigenvalues* and the set of possible values is called the *spectrum of eigenvalues* of f. The spectrum of eigenvalues can be *continuous* (e.g., position or momentum eigenvalues) or *discrete* (angular momentum, energy of bound atomic levels, etc.). We consider here for simplicity an observable with a discrete spectrum.

Let f_n denote the n-th eigenvalue ($n = 1, 2, \ldots$) and let $\Psi_n(q)$ be the wave function corresponding to the state where a measurement of the observable f yields with *certainty* the value f_n. The wave function $\Psi_n(q)$ is said to be an eigenfunction or *eigenstate* of the observable f with eigenvalue f_n. We assume that the wave functions Ψ_n are normalized for all n:

$$\int |\Psi_n(q)|^2\, dq = 1.$$

If one considers just one observable, one finds that there are in general several different eigenstates $\Psi_n(q)$ which have the same eigenvalue. For instance, if f is the kinetic energy $p^2/2m$ there are different states with momentum along the x, y or z direction which have the same energy. This would complicate the following discussion, since eigenstates with the same or different f_n need to be treated separately. We therefore assume that f_n refers to a *complete set of observables.* This implies that they characterize the state $|\, \Psi \,\rangle$ and wave function $\Psi_n(q)$ completely and that no other observable can be measured with certainty at the same time. Since the set of values f_n with $n = 1, 2, \ldots$ covers all possible values of f_n, a measurement of the observable f on an arbitrary state $|\, \Psi \,\rangle$ must yield one of the eigenvalues f_n. In accordance with the superposition principle the wave function Ψ must be a linear combination of the eigenstates Ψ_n. We can therefore write

$$\Psi(q) = \sum_n a_n\, \Psi_n(q) \tag{3.6}$$

or, in a representation-independent form,

$$| \Psi \rangle = \sum_n a_n \, | n \rangle$$

with some coefficients a_n. This is a consequence of the fact that f_n characterizes the state completely and that all possible values of f_n are taken into account in the sum. In fact, if there would be a state that cannot be written in this form, one would be able to construct a state which is orthogonal to the subspace spanned by $\{\Psi_1, \Psi_2, \ldots\}$ and for which a measurement of the observable f yields either a known value f_n, which would contradict the assumption that f is a complete set, or a new value, which would contradict the assumption that all possible values of the observable f were included in the expansion (3.6).

We therefore conclude that an arbitrary wave function can always be represented by a linear combination (usually an infinite series) of the eigenstates of a complete set of observables.

Once the state is written in the form (3.6), we postulate that $|a_n|^2$ (i.e., the square modulus of the coefficient of Ψ in the eigenstate Ψ_n) gives the probability P_n that the value f_n is obtained as a result of the measurement of the observable f in the state $\Psi(q)$. This holds provided that $\int |\Psi(q)|^2 \, dq = 1$, otherwise one would divide $|a_n|^2$ by $\int |\Psi(q)|^2 \, dq$.

One may now analyze why this is a reasonable assumption. We know that $\Psi(q)$ defines the state, and since the $\Psi_n(q)$ are fixed, the probability P_n for f_n can only depend on the expansion coefficients a_n. Moreover, we must have $P_n \geq 0$, so that P_n cannot be a linear function of a_n. P_n has to be bilinear (i.e., quadratic) in $\{a_n\}$ and the sum of all probabilities $\sum_n P_n$ must be invariant under any unitary transformation among the eigenstates $| n \rangle$. Moreover, if $\Psi = \Psi_n$ for some n, then $P_m = \delta_{mn} \; \forall \; m$. The only possible choice is then $P_n = |a_n|^2$, which tends to 1 whenever $\Psi \to \Psi_n$.

A few important remarks are due:

i) Notice that if one multiplies the wave function $\Psi(q)$ or the vector state $| \Psi \rangle$ by an arbitrary complex number $e^{i\varphi}$ of modulus 1 ($\varphi \in \mathbb{R}$), none of the results of any possible experiment would change. In other words $\Psi(q)$, $-\Psi(q)$ or $e^{i\varphi} \, \Psi(q)$ represent the same physical state.

   This indeterminacy is irremovable. It has, hovewer, no physical significance since it has no effect on any physical result.

   Some authors say that quantum mechanics is a theory of rays since only the direction of the vector $| \Psi \rangle$ matters. The ray associated to a vector $| \Psi \rangle$ is given by the set $\{c \, | \Psi \rangle$ with $c \in \mathbb{C}$ and $|c| = 1\}$. In fact the restriction $|c| = 1$ is not very important since one can always normalize the probabilities P_n a posteriori.

   Notice that $e^{i\varphi}$ is a constant independent of q and t. Phase factors that depend on position or time do matter [e.g., $e^{i\,(kx - \omega t)}$].

ii) The second important consequence concerns the change in the quantum mechanical state $|\,\Psi\,\rangle$ that is caused by the measurement process. If the result of the measurement of the complete set of observables f is f_n, then the system can only be in the state Ψ_n after the measurement. The measurement acts therefore as a projector

$$\Psi = \sum_n a_n\,\Psi_n \xrightarrow{f \text{ measurement}} |\,\Psi_n\,\rangle$$

with probability $|a_n|^2$. This can be understood as follows. If the result of a measurement is f_n, then a subsequent measurement of f at an instant Δt later must yield the same result as the previous measurement when $\Delta t \to 0$. Taking into account that f_n defines the state completely, then $|\,\Psi\,\rangle = |\,\Psi_n\,\rangle$ after the measurement (besides an irrelevant phase factor). Note that the change in state $|\,\Psi\,\rangle \to |\,\Psi_n\,\rangle$ upon measuring f (and obtaining f_n) occurs even if no human records the result. The change of state occurs because of the *interaction with the apparatus* and has nothing to do with the presence or not of a human observer.

It is clear that the previous history of the quantum state $|\,\Psi\,\rangle$ system is irrelevant just after the measurement. For $t > t_0$ (t_0 time of the measurement) we do not care about the values of all the other a_m with $m \neq n$ if the result of the measurement was f_n.

A measurement in quantum mechanics acts as a filter. It prepares the state in a given eigenstate $|\,\Psi_n\,\rangle$ of the system corresponding to the measured observable f_n.

(Bild)

This would be a selective measurement with only one state filtered. But one can of course filter two or more states $|\,\Psi_n\,\rangle$ and let them interfere, as in our idealized electron diffraction experiment.

(Bild)

iii) If we remove the assumption that f_n is a complete set (which defines $|\,\Psi_n\,\rangle$ univocally besides a phase factor) the probability P_n of measuring a particular eigenvalue f_n would be

$$P_n = \sum_{n'} |a_{n'}|^2,$$

where the sum runs over all the eigenstates n' having the eigenvalue f_n ($\langle\,\Psi_l\,|\,\Psi_m\,\rangle = \delta_{lm}\ \forall\ l,m$). Concerning the change of state resulting from a measurement we have

$$|\,\Psi\,\rangle \xrightarrow{f \text{ measurement}} A\sum_{n'} a_{n'}\,|\,\Psi_{n'}\,\rangle,$$

where n' runs over all the eigenstates $|\,\Psi_{n'}\,\rangle$ having the eigenvalue f_n, which is the result of the measurement, and $A^{-1} = \sum_{n'} |a_{n'}|^2$ is the normalization constant.

We see that the weaker the filtering is, i.e., the larger the number of acceptable results of the measurement, the weaker the change in the quantum state after the measurement, i.e., the weaker the projection.

In this context one often introduces the projection operator

$$\hat{\Lambda}_n = |\,\Psi_n\,\rangle\,\langle\,\Psi_n\,|$$

which describes the result of the measurement of the complete set of eigenvalues f_n on any state $|\,\Psi\,\rangle$:

$$\hat{\Lambda}_n\,|\,\Psi\,\rangle = \hat{\Lambda}_n \sum_m a_m\,|\,\Psi_m\,\rangle = \sum_m a_m\,\hat{\Lambda}_n\,|\,\Psi_m\,\rangle = \sum_m a_m\,|\,\Psi_n\,\rangle\,\underbrace{\langle\,\Psi_n\,|\,\Psi_m\,\rangle}_{\delta_{nm}} = a_n\,|\,\Psi_n\,\rangle.$$

For a simple eigenvalue f_n, which does not define $|\,\Psi_n\,\rangle$ completely, or when the measurement does not discern between different completely defined states $|\,\Psi_{n'}\,\rangle$, the measurement process is described by the operator

$$\hat{\Lambda}'_n = \sum_{n'} |\,\Psi_{n'}\,\rangle\,\langle\,\Psi_{n'}\,|,$$

where the sum runs over the eigenstates having the eigenvalue f_n.

The operator $\hat{\Lambda}_n = |\,\Psi_n\,\rangle\,\langle\,\Psi_n\,|$, which we can simply write $\hat{\Lambda}_n = |\,n\,\rangle\,\langle\,n\,|$, is also known as *outer product* in contrast to the inner product $\langle\,\Psi\,|\,\Phi\,\rangle$ which is a complex number. Notice that $\hat{\Lambda}_n^2 = \hat{\Lambda}_n$ and $\hat{\Lambda}_n'^2 = \hat{\Lambda}'_n$, as for any projector.

iv) We return now to the general case where f_n is a complete set of observables in order to infer some properties about the states $|\,n\,\rangle$ and the corresponding wave functions $\Psi_n(q)$. We know that $|a_n|^2$ is the probability for measuring f_n (provided $\int |\Psi|^2\,dq = 1$) and consequently

$$\sum_n |a_n|^2 = 1. \tag{3.7}$$

If $\Psi(q)$ is not normalized we would have to divide $|a_n|^2$ by $\int |\Psi|^2\,dq$, so that in general we can write

$$\sum_n |a_n|^2 = \int |\Psi|^2\,dq \tag{3.8}$$

and ignore the constraint on the expansion coefficients a_n concerning the normalization of Ψ. From Eq. (3.6) we have

$$\Psi^*(q) = \sum_n a_n^* \, \Psi_n^*(q),$$

so that Eq. (3.8) implies

$$\begin{aligned} \sum_n a_n^* \, a_n &= \int \left[\sum_n a_n^* \, \Psi_n^*(q) \right] \Psi(q) \, dq \\ &= \sum_n a_n^* \int \Psi_n^*(q) \, \Psi(q) \, dq. \end{aligned} \tag{3.9}$$

Since this holds for any values of the complex coefficients a_n, one can derive Eq. (3.9) with respect to a_n^* by considering a_n and a_n^* as independent variables. This yields

$$a_n = \int \Psi_n^*(q) \, \Psi(q) \, dq, \tag{3.10}$$

which is the expression for determining the coefficients a_n in terms of the wave function $\Psi(q)$ and the eigenfunction $\Psi_n(q)$. This is a very important relation between the wave function $\Psi(q)$, which characterizes the state of the system, and the eigenfunction or eigenstate $\Psi_n(q)$ having the definite value f_n of the observable in question. We realize that finding the eigenstates of physical observables will be a central task in solving problems in quantum mechanics.

v) Replacing Eq. (3.6) in Eq. (3.10) we further obtain

$$\begin{aligned} a_n &= \int \Psi_n^*(q) \left(\sum_m a_m \, \Psi_m(q) \right) dq \\ &= \sum_m a_m \int \Psi_n^*(q) \, \Psi_m(q) \, dq \end{aligned}$$

which implies

$$\int \Psi_n^*(q) \, \Psi_m(q) \, dq = \delta_{mn}, \tag{3.11}$$

where δ_{mn} is the Kronecker delta defined by $\delta_{mn} = 1$ for $m = n$ and $\delta_{mn} = 0$ for $m \neq n$. This is known as the *orthogonality condition* between the eigenfunctions Ψ_n. We have demonstrated that the eigenfunctions corresponding to different eigenvalues f_n are orthogonal to each other. Therefore, the set of Ψ_n forms *complete orthonormal basis* of the vector space of all the wave functions of the system.

Notice that Eq. (3.11) is a particular case of Eq. (3.10), in which one sets $\Psi(q)$ equal to $\Psi_m(q)$, since $\Psi(q)$ is equal to $\Psi_m(q)$ precisely when the expansion coefficients $a_n = \delta_{mn}$.

The orthogonality relation (3.11) has been derived assuming that f_n is a complete set of observables which defines Ψ_n univocally. This means that for any two eigenstates Ψ_n and Ψ_m there is some observable whose value in the lists f_n and f_m is different. For simple observables (e.g., the momentum p_x or the kinetic energy $p^2/2m$) there are many states corresponding to the same eigenvalue. However, even in this case the eigenstates are either orthogonal (e.g., because they differ in the eigenvalue of some other observable) or they can be reorthogonalized among the eigenstates having the same eigenvalue. In conclusion one can always assume that all the eigenstates of a physical observable can be written in the form of an orthonormal basis.

3.5. Inner product

In the previous section we have repeatedly found expressions of the form $\int \Psi^*(q)\, \Phi(q)\, dq$, where Ψ and Φ are wave functions. It is therefore useful to analyze the properties of this application $\mathbb{V} \times \mathbb{V} \to \mathbb{C}$ from a more mathematical perspective and in particular to identify it as the inner product in the vector space $\mathbb{V}$ of all quantum mechanical states.

Given two vector states $|\,\Psi\,\rangle$ and $|\,\Phi\,\rangle$ corresponding to the wave functions $\Psi(q)$ and $\Phi(q)$ we define

$$\langle\,\Psi\,|\,\Phi\,\rangle = \int \Psi^*(q)\, \Phi(q)\, dq. \tag{3.12}$$

Denoting the vector space of quantum mechanical states $|\,\Psi\,\rangle$ by $\mathbb{V}$, we may say that $\langle\,\Psi\,|\,\Phi\,\rangle$ is an application from $\mathbb{V} \times \mathbb{V}$ to $\mathbb{C}$. It is easy to see that this application satisfies all the properties of an inner product in $\mathbb{V}$:

i) Positive definiteness:

$$\langle\,\Psi\,|\,\Psi\,\rangle = \int |\Psi|^2\, dq \geq 0 \quad \forall\, |\,\Psi\,\rangle$$

and

$$\langle\,\Psi\,|\,\Psi\,\rangle = 0 \quad \Rightarrow \quad \int |\Psi|^2\, dq = 0 \quad \Rightarrow \quad |\Psi|^2 = 0 \quad \forall\, q \quad \Rightarrow \quad |\,\Psi\,\rangle = 0.$$

ii) Linearity at right:

$$\begin{aligned}\langle\, \Psi \,|\, \alpha\, \Phi_1 + \beta\, \Phi_2 \,\rangle &= \int \Psi^* \,(\alpha\, \Phi_1 + \beta\, \Phi_2)\; dq \\ &= \alpha \int \Psi^*\; \Phi_1\; dq + \beta \int \Psi\; \Phi_2\; dq \\ &= \alpha\, \langle\, \Psi \,|\, \Phi_1 \,\rangle + \beta\, \langle\, \Psi \,|\, \Phi_2 \,\rangle.\end{aligned}$$

iii) Antisymmetry:

$$\langle\, \Phi \,|\, \Psi \,\rangle = \int \Phi^*(q)\; \Psi(q)\; dq = \left[\int \Psi^*(q)\; \Phi(q)\; dq\right]^* = \langle\, \Psi \,|\, \Phi \,\rangle^*.$$

iv) Combining ii) and iii) the so-called antilinearity at left follows:

$$\begin{aligned}\langle\, \alpha\, \Psi_1 + \beta\, \Psi_2 \,|\, \Phi \,\rangle &= \langle\, \Phi \,|\, \alpha\, \Psi_1 + \beta\, \Psi_2 \,\rangle^* = \\ &= [\alpha\, \langle\, \Phi \,|\, \Psi_1 \,\rangle + \beta\, \langle\, \Phi \,|\, \Psi_2 \,\rangle]^* \\ &= \alpha^* \langle\, \Phi \,|\, \Psi_2 \,\rangle^* + \beta^* \langle\, \Phi \,|\, \Psi_2 \,\rangle^* \\ &= \alpha^* \langle\, \Psi_1 \,|\, \Phi \,\rangle + \beta^* \langle\, \Psi_2 \,|\, \Phi \,\rangle.\end{aligned}$$

The coefficients get conjugated when extracted from the left hand side of the inner products.

3.6. Kets and Bras

P. M. Dirac denoted the quantum mechanical vector state $|\, \Psi \,\rangle$ of the system a *ket* and attached to each ket $|\, \Psi \,\rangle$ a *bra* $\langle\, \Psi \,|$ so that $\langle\, \Phi \,|\, \Psi \,\rangle$ is a “bra-ket”. Though this may sound trivial or almost silly at the beginning, the notation in ket and bra form is extremely useful in practice, since it allows to write most relations in quantum mechanics in a representation-independent form. It applies throughout quantum mechanics from the present elementary level all over to relativistic quantum mechanics and complex many-body problems.

For the algebraic manipulations one just needs to know that the bra $\langle\, \Psi \,|$ and the ket $|\, \Psi \,\rangle$ are related one-to-one by

$$\begin{array}{ccc} \text{ket} & \longleftrightarrow & \text{bra} \\ |\, \Psi \,\rangle = \alpha \,|\, \Psi_1 \,\rangle + \beta \,|\, \Psi_2 \,\rangle & & \langle\, \Psi \,| = \alpha^* \langle\, \Psi_1 \,| + \beta^* \langle\, \Psi_2 \,|. \end{array}$$

This follows from ii)–iv).

We may now profit from Dirac’s notation to summarize the previous results obtained using the wave function notation in the more compact and representation-independent ket-bra form. Let $|\,\Psi\,\rangle$ be the ket associated to the wave function $\Psi(q)$, $|\,\Phi\,\rangle$ the one associated to $\Phi(q)$, and $|\,n\,\rangle$ the eigenket corresponding to the eigenvalue f_n of the observable f. Expanding the wave functions $\Psi(q) = \sum_n a_n\,\Psi_n(q)$ and $\Phi(q) = \sum_n b_n\,\Psi_n(q)$ in the eigenfunctions $\Psi_n(q) = \langle\,q\,|\,n\,\rangle$ we have

$$\langle\,\Phi\,|\,\Psi\,\rangle = \sum_{n,m} b_m^*\,a_n \underbrace{\int \Psi_m^*(q)\;\Psi_n(q)\;dq}_{\delta_{mn}} = \sum_n b_n^*\,a_n,$$

which corresponds to the usual inner product in $\mathbb{C}^n$. In particular for the eigenstates $|\,n\,\rangle$ and $|\,m\,\rangle$ we have

$$\langle\,n\,|\,m\,\rangle = \delta_{nm}.$$

We can now write

$$|\,\Psi\,\rangle = \sum_n a_n\,|\,n\,\rangle,$$

where

$$\langle\,n\,|\,\Psi\,\rangle = \langle\,n\,|\sum_m a_m\,|\,m\,\rangle = \sum_m a_m\,\langle\,n\,|\,m\,\rangle = \sum_m a_m\,\delta_{mn} = a_n = \int \Psi_n^*(q)\;\Psi(q)\;dq.$$

The operator $\hat{\Lambda}_n = |\,n\,\rangle\,\langle\,n\,|$, also known as outer product, acts like a projector onto the ket $|\,n\,\rangle$ since for any

$$|\,\Psi\,\rangle = \sum_m a_m\,|\,m\,\rangle$$

it holds

$$\hat{\Lambda}_n\,|\,\Psi\,\rangle = |\,n\,\rangle\,\langle\,n\,|\sum_m a_m\,|\,m\,\rangle = \sum_m a_m\;|\,n\,\rangle\;\underbrace{\langle\,n\,|\,m\,\rangle}_{\delta_{nm}} = a_n\,|\,n\,\rangle.$$

In other words $\hat{\Lambda}_n$ projects $|\,\Psi\,\rangle$ along the state vector $|\,n\,\rangle$. Notice that $\hat{\Lambda}_n^2 = \hat{\Lambda}_n$, as expected for any projector operator.

The completeness of the orthonormal set of eigenstates $|\,n\,\rangle$ is equivalent to stating that any state $|\,\Psi\,\rangle$ can be expressed as a linear combination of them

$$|\,\Psi\,\rangle = \sum_n a_n\,|\,n\,\rangle.$$

Taking into account that $a_n = \langle\,n\,|\,\Psi\,\rangle$ we can write

$$|\,\Psi\,\rangle = \sum_n \langle\,n\,|\,\Psi\,\rangle\,|\,n\,\rangle = \sum_n |\,n\,\rangle\,\langle\,n\,|\,\Psi\,\rangle.$$

Since this holds for any $|\,\Psi\,\rangle$ we must have

$$\sum_n |\,n\,\rangle\,\langle\,n\,| = 1. \tag{3.13}$$

We conclude that an orthonormal basis of kets is complete if and only if Eq. (3.13) holds. This equation is therefore known as the *completeness relation* for the basis of eigenstates. This is an operator identity that expresses a relation between the ensemble of projector operators $\hat{\Lambda}_n = |\, n \,\rangle \langle\, n \,|$.

The completeness relation (3.13) is the most important contribution or innovation of Dirac's ket-bra notation. It is a very powerful tool for deriving algebraic relations between different representations as well as operator relations.

In Dirac's notation one recovers the wave function $\Psi(q)$ by projecting $|\, \Psi \,\rangle$ onto a ket of defined position:

$$\Psi(q) = \langle\, q \,|\, \Psi \,\rangle. \tag{3.14}$$

At this point one may consider Eq. (3.14) as a definition of the bra $\langle\, q \,|$. The definition

$$\langle\, \Psi \,|\, \Phi \,\rangle = \int \Psi^*(q)\; \Phi(q)\; dq$$

can then be written as

$$\begin{aligned} \langle\, \Psi \,|\, \Phi \,\rangle &= \int \langle\, q \,|\, \Psi \,\rangle^* \; \langle\, q \,|\, \Phi \,\rangle \, dq \\ &= \int \langle\, \Psi \,|\, q \,\rangle \; \langle\, q \,|\, \Phi \,\rangle \, dq \\ &= \langle\, \Psi \,\big|\big(\int |\, q \,\rangle \langle\, q \,|\; dq \big)\big|\, \Phi \,\rangle. \end{aligned} \tag{3.15}$$

Noting that Eq. (3.15) holds for any states $|\, \Psi \,\rangle$ and $|\, \Phi \,\rangle$ one concludes that

$$\int |\, q \,\rangle \langle\, q \,|\, dq = 1. \tag{3.16}$$

This relation is known as completeness relation for the coordinate representation.

3.7. Operators

So far we have introduced the wave function $\Psi(q)$ and the more abstract notion of vector state or ket $|\, \Psi \,\rangle$ as the fundamental mathematical objects describing the state of a quantum system and we have established the general relation between $\Psi(q)$ or $|\, \Psi \,\rangle$ and the outcome of any experiment (i.e., the possible values of physical observables or eigenvalues f_n and the probabilities $|a_n|^2$ for observing each eigenvalue). To this aim we have introduced the concepts of eigenvalues and eigenstates, the latter being the quantum states in which a measurement of the observable f yields with certainty the value f_n. It is the purpose of this section to derive the general form of the mathematical equations from which both f_n and Ψ_n are obtained.

Our starting point is the concept of *mean value* or *expectation value* of the observable f in a given state $\Psi(q) = \sum_n a_n \, \Psi_n(q)$ which is given by

$$\langle\, f \,\rangle = \sum_n |a_n|^2 \, f_n. \tag{3.17}$$

This is the usual expression of a mean value known from probability theory (i.e., sum of the possible values of a random variable times the corresponding probabilities). It is equivalent to the result one would obtain by averaging the outcomes of a large number of measurements performed on the same state $|\,\Psi\,\rangle$ or wave function $\Psi(q)$.

It is clear that for each observable f with a spectrum $\{f_n,\, n = 1, 2, \ldots\}$ the expectation value depends only on $\Psi(q)$ since $a_n = \int \Psi_n^*(q)\, \Psi(q)\, dq$. We therefore seek for an expression for $\langle\, f\,\rangle$ that involves $\Psi(q)$ explicitly and not the expansion coefficients a_n. Eq. (3.17) is bilinear in a_n and a_n^* and must therefore be bilinear in Ψ and Ψ^*, since we have the correspondence $\alpha\,\Psi + \beta\,\Psi' \;\leftrightarrow\; \alpha\, a_n + \beta\, a_n'$. The most general form of such a bilinear expression is

$$\langle\, f\,\rangle = \int \Psi^*(q)\, \left(\hat{f}\, \Psi(q)\right) dq, \tag{3.18}$$

where $\hat{f}$ is some linear operator[2] acting on $\Psi(q)$. This operator is simply a linear application $\hat{f} : \mathbb{V} \to \mathbb{V}$, where $\mathbb{V}$ is the vector space of wave functions of the system. The linearity of $\hat{f}$ means that $\hat{f}\,(\alpha\,\Psi + \beta\,\Psi') = \alpha\,(\hat{f}\,\Psi) + \beta\,(\hat{f}\,\Psi')$ for any wave functions Ψ and Ψ' and complex coefficients α and β. $\hat{f}\,\Psi(q)$ is the function obtained by applying the operator $\hat{f}$ to the wave function $\Psi(q)$.

From Eq. (3.17) we have

$$\langle\, f\,\rangle = \sum_n a_n^*\, a_n\, f_n$$

and since

$$a_n = \int \Psi_n^*\, \Psi\, dq \;\Leftrightarrow\; a_n^* = \int \Psi^*\, \Psi_n\, dq$$

we obtain

$$\begin{aligned}\langle\, f\,\rangle &= \sum_n \left(\int \Psi^*\, \Psi_n\, dq\right) a_n\, f_n \\ &= \int \Psi^* \left(\sum_n a_n\, f_n\, \Psi_n\right) dq.\end{aligned} \tag{3.19}$$

Since Eqs. (3.18) and (3.19) hold for any quantum state Ψ we conclude that

$$\hat{f}\, \Psi(q) = \sum_n a_n\, f_n\, \Psi_n(q). \tag{3.20}$$

In Dirac's notation one would write $|\,\Psi\,\rangle = \sum_n a_n\, |\, n\,\rangle$ and $\hat{f}\, |\,\Psi\,\rangle = \sum_n a_n\, f_n\, |\, n\,\rangle$.

One concludes that for any physical quantity or observable f there is a definite linear operator $\hat{f}$ given by Eq. (3.20). The linearity of $\hat{f}$ can be easily verified by noting the linear correspondence between Ψ and the expansion coefficients a_n (i.e., $\alpha\,\Psi + \beta\,\Psi' \;\leftrightarrow\; \alpha\, a_n + \beta\, a_n'$).

---

[2] Throughout these notes we use a circumflex or hat ˆ to distinguish operators from numbers.

A very important consequence of Eq. (3.20) is the so-called *eigenvalue equation* from which both f_n and $\Psi_n(q)$ are derived. Consider Eq. (3.20) for $\Psi(q) = \Psi_n(q)$ for some n, i.e., when $\Psi(q)$ is one of the eigenfunctions of the observable f. In this case $a_n = 1$ and $a_m = 0 \;\; \forall \; m \neq n$ so that

$$\hat{f}\,\Psi_n(q) = f_n\,\Psi_n(q). \tag{3.21}$$

In Dirac's notation Eq. (3.21) reads $\hat{f}\,|\,n\,\rangle = f_n\,|\,n\,\rangle$. The result of applying $\hat{f}$ to $\Psi_n(q)$ is just the multiplication by the eigenvalue f_n. Since this holds for all n, we conclude that the eigenfunctions of the observable f are the solutions of the equation

$$\hat{f}\,\Psi(q) = f\,\Psi(q), \tag{3.22}$$

where f is a constant. The eigenvalues f_n are the values of f for which Eq. (3.22) has a non-trivial solution ($\Psi(q) \neq 0$).

In following sections we shall derive the form of the operators associated to the most important observables based on general physical considerations. A few examples, namely position, momentum and kinetic energy, are given below.

3.7.1. Position, momentum and kinetic energy operators

The operator $\hat{x}$ associated with a measurement of the x coordinate of an electron is simply the multiplication by the x coordinate:

$$\hat{x}\,\Psi(\vec{r}) = x\,\Psi(\vec{r}),$$

where $\vec{r} = (x, y, z)$. This follows from the fact that $|\Psi(\vec{r})|^2$ represents the probability density for a measurement of the position $\vec{r}$ and from the definition of average value

$$\langle\, x\,\rangle = \int \Psi^*(\vec{r})\, x\, \Psi(\vec{r})\, d^3r,$$

which holds for all $\Psi(\vec{r})$. Of course, analogous relations hold for the y and z coordinates, i.e., $\hat{y} = y$ and $\hat{z} = z$.

The momentum operator $\hat{p}_x$ along the x direction is related to the changes in the wave function upon translation along this direction and is therefore proportional to the derivative of the wave function along the x direction. The precise form is

$$\hat{p}_x = -i\,\hbar\,\frac{\partial}{\partial x}, \tag{3.23}$$

which means that

$$\hat{p}_x\,\Psi(\vec{r}) = -i\,\hbar\,\frac{\partial \Psi}{\partial x}.$$

The corresponding expressions for $\hat{p}_y$ and $\hat{p}_z$ follow from symmetry considerations so that

$$\hat{\vec{p}} = -i\,\hbar\,\vec{\nabla} = -i\,\hbar \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right).$$

In order to justify the form (3.23) of the operator $\hat{p}_x$ we recall that according to the de Broglie relation a plane-wave state $e^{i\,\vec{k}\cdot\vec{r}}$ carries a momentum $\vec{p} = \hbar\,\vec{k}$. Thus,

$$\hat{p}_x\, e^{i\,\vec{k}\cdot\vec{r}} = p_x\, e^{i\,\vec{k}\cdot\vec{r}} = \hbar\, k_x\, e^{i\,\vec{k}\cdot\vec{r}} = -i\,\hbar\,\frac{\partial}{\partial x}\, e^{i\,\vec{k}\cdot\vec{r}}.$$

Consider now an arbitrary wave function $\Psi(\vec{r})$ which can be expanded in terms of plane waves as a Fourier transform

$$\Psi(\vec{r}) = \int d^3k\; A(\vec{k})\; e^{i\,\vec{k}\cdot\vec{r}} = \int d^3k\; A(\vec{k})\; e^{i\,\frac{\vec{p}}{\hbar}\cdot\vec{r}}.$$

Applying $\hat{p}_x$ to $\Psi(\vec{r})$ one obtains

$$\begin{aligned}
\hat{p}_x\,\Psi(\vec{r}) &= \int d^3k\; A(\vec{k})\; \hat{p}_x\; e^{i\,\vec{k}\cdot\vec{r}} \\
&= \int d^3k\; A(\vec{k}) \left(-i\,\hbar\,\frac{\partial}{\partial x}\right) e^{i\,\vec{k}\cdot\vec{r}} \\
&= -i\,\hbar\,\frac{\partial}{\partial x}\;\Psi(\vec{r}).
\end{aligned}$$

Therefore $\hat{p}_x = -i\,\hbar\,\dfrac{\partial}{\partial x}$ for any wave function $\Psi(\vec{r})$.

The form of the kinetic energy operator can be derived by using the relation between kinetic energy and momentum known from classical mechanics, which applies to wave packets for which the de Broglie wave length is much shorter than the system's dimensions. In the classical limit we have

$$T = \frac{p^2}{2m}.$$

Consequently, for a classical wave packet we must have

$$\langle\, T\,\rangle = \int \Psi^*(\hat{T}\,\Psi)\, dq = \int \Psi^* \left(\frac{\hat{p}^2}{2m}\right) \Psi\; dq,$$

which implies

$$\hat{T} = \frac{\hat{p}^2}{2m}.$$

Noting that $p_x = -i\,\hbar\,\dfrac{\partial}{\partial x}$ and therefore $p_x^2 = -\hbar^2\,\dfrac{\partial^2}{\partial x^2}$ we obtain

$$\hat{T} = \frac{\hat{p}^2}{2m} = -\frac{\hbar^2}{2m}\,\nabla^2 = -\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right).$$

The previous derivations of the momentum and kinetic energy operators are admittedly not quite rigorous mathematically. Nevertheless, they do provide serious plausibility arguments. Achieving mathematical rigor would imply a too long digression that we cannot afford in this lecture.

3.8. Properties of physical operators

The expectation values

$$\langle\, f\,\rangle = \sum_n f_n\, |a_n|^2 = \int \Psi(q) \left(\hat{f}\,\Psi(q)\right) dq \tag{3.24}$$

and the eigenvalues f_n of the operators associated to physical observables must be real. Note that f_n is a particular case of Eq. (3.24) when $\Psi = \Psi_n$. This important physical condition imposes precise restrictions on the form of the operator $\hat{f}$. Before discussing this problem it is useful to introduce a few mathematical definitions.

For any orthonormal basis $\Psi_n(q)$ (not necessarily a basis of eigenstates) we define the *matrix elements* f_{mn} of an operator $\hat{f}$ in the basis $\{\Psi_n,\, n = 1, 2, \ldots\}$ as

$$f_{mn} = \int \Psi^*_m(q) \left[\hat{f}\,\Psi_n(q)\right] dq = \langle\, m\,|\,\hat{f}\,n\,\rangle = \langle\, m\,|\,\hat{f}\,|\,\hat{n}\,\rangle. \tag{3.25}$$

For the calculation of the matrix elements f_{mn} only the form of the operator $\hat{f}$ in some representation is needed, since knowing how $\hat{f}$ modifies any function we can determine $\hat{f}\,\Psi_n(q)$ or $\hat{f}\,|\,n\,\rangle$ for any basis state $\Psi_n(q)$ or $|\,n\,\rangle$ and compute the inner product with $\Psi^*_m(q)$ or $\langle\, m\,|$.

Conversely, if we know f_{mn} in a complete basis we also know $\hat{f}\,|\,\Psi\,\rangle$ for any $|\,\Psi\,\rangle$, since we can always write

$$\begin{aligned}
\hat{f}\,|\,\Psi\,\rangle &= \left(\sum_m |\,m\,\rangle\,\langle\,m\,|\right) \hat{f}\,|\,\Psi\,\rangle = \sum_m |\,m\,\rangle\,\langle\,m\,|\,\hat{f}\,|\,\Psi\,\rangle = \\
&= \sum_m |\,m\,\rangle\,\langle\,m\,|\,\hat{f}\left(\sum_n |\,n\,\rangle\,\langle\,n\,|\right) |\,\Psi\,\rangle \\
&= \sum_{m,n} |\,m\,\rangle\,\langle\,m\,|\,\hat{f}\,|\,n\,\rangle\,\langle\,n\,|\,\Psi\,\rangle \\
&= \sum_{m,n} |\,m\,\rangle\, f_{mn}\,\langle\,n\,|\,\Psi\,\rangle.
\end{aligned}$$

Notice the repeated use of the completeness relation $\sum_n |\,n\,\rangle\,\langle\,n\,| = 1$. In other words, if we write $|\,\Psi\,\rangle = \sum a_n\,|\,n\,\rangle$ or $\Psi(q) = \sum_n a_n\,\Psi_n(q)$ we obtain $\hat{f}\,|\,\Psi\,\rangle = \sum_m b_m\,|\,m\,\rangle$ or $\hat{f}\,\Psi(q) = \sum_m b_m\,\Psi_m(q)$ with

$$b_m = \sum_n f_{mn}\, a_n.$$

The siutation is analogous to the description of a vector state $|\,\Psi\,\rangle$, which is defined whenever we give its components on a basis

$$|\,\Psi\,\rangle = \sum_n a_n\,|\,n\,\rangle.$$

In the case of operators, giving the value of the matrix elements in a complete basis defines the operator completely. One can actually say that f_{mn} are the components of the operator $\hat{f}$ in the basis $\{|\, n\,\rangle$ with $n = 1, 2, \ldots\}$.

Notice that the definition (3.25) holds for any orthonormal set $\{\Psi_n\}$, which are not necessarily eigenstates of $\hat{f}$. However, if $\hat{f}\,\Psi_n = f_n\,\Psi_n \;\forall\, n$, Eq. (3.25) takes the particularly simple form

$$f_{mn} = f_n\,\langle\, m\,|\, n\,\rangle = f_n\,\delta_{mn}. \tag{3.26}$$

In this case, one says that the matrix f_{mn} is diagonal. Of course, the converse is also true. If Eq. (3.26) holds for all m and n in a complete orthonormal basis, then the states $\{\Psi_n,\, n = 1, 2, \ldots\}$ are the eigenstates of $\hat{f}$.

Given an operator $\hat{f}$ we define the *conjugate operator* $\hat{f}^*$ by the relation

$$(\hat{f}\,\Psi)^* = \hat{f}^*\,\Psi^*.$$

In other words, if $\hat{f}\,\Psi = \Phi \;\Rightarrow\; \hat{f}^*\,\Psi^* = \Phi^*$. It is easy to show that $\hat{f}^*$ is linear, since $\hat{f}$ is linear. Note that $(\hat{f}^*)^* = \hat{f}$.

For an arbitrary operator $\hat{f}$ we define the *transpose operator* $\hat{f}^t$ by the relation

$$\int \Phi\,(\hat{f}\,\Psi)\,dq = \int \Psi\,(\hat{f}^t\,\Phi)\,dq.$$

Notice that we simply interchange the functions Φ and Ψ. In addition we have $(\hat{f}^t)^t = \hat{f}$. The *adjoint* or *complex-conjugate-transpose* operator $\hat{f}^\dagger$ of an operator $\hat{f}$ is defined by

$$\hat{f}^\dagger = (\hat{f}^*)^t = (\hat{f}^t)^*.$$

It is easy to see that the order of the transposition and conjugation does not matter since

$$\int \Phi^*\,(\hat{f}^*)^t\,\Psi\,dq = \int \Psi\,\hat{f}^*\,\Phi^*\,dq = \left(\int \Psi^*\,\hat{f}\,\Phi\,dq\right)^* = \left(\int \Phi\,\hat{f}^t\,\Psi^*\,dq\right)^* = \int \Phi^*\,(\hat{f}^t)^*\,\Psi\,dq$$

for all $\Phi(q)$ and $\Psi(q)$.

An operator $\hat{f}$ and its adjoint $\hat{f}^\dagger$ satisfy the following important relation

$$\langle\, \Phi\,|\,\hat{f}\,\Psi\,\rangle = \langle\,\hat{f}^\dagger\,\Phi\,|\,\Psi\,\rangle \tag{3.27}$$

or equivalently

$$\int \Phi^*(q)\,\Big(\hat{f}\,\Psi(q)\Big)\,dq = \int \Big[\hat{f}^\dagger\,\Phi(q)\Big]^*\;\Psi(q)\,dq. \tag{3.28}$$

Notice that in the left-hand side $\hat{f}$ acts on $\Psi(q)$ while in the right-hand side $\hat{f}^\dagger$ acts on $\Phi(q)$. This can be proven straightforwardly:

$$\int \Phi^*\,(\hat{f}\,\Psi)\,dq = \int \Psi\,(\hat{f}^t\,\Phi^*)\,dq = \int \Psi\,(\hat{f}^\dagger\,\Phi)^*\,dq.$$

The matrix elements of $\hat{f}$ and $\hat{f}^\dagger$ are related to each other by

$$(\hat{f}^\dagger)_{mn} = (\hat{f}_{nm})^*. \tag{3.29}$$

As in the usual adjoint of matrices in $\mathbb{C}^{n\times n}$, the indices are transposed and the result conjugated. Eq. (3.29) can be demonstrated by using Eq. (3.28) since

$$(\hat{f}^\dagger)_{mn} = \langle\, m \,|\, \hat{f}^\dagger\, n \,\rangle = \langle\, \hat{f}\, m \,|\, n \,\rangle = \langle\, n \,|\, \hat{f}\, m \,\rangle^* = (f_{nm})^*. \tag{3.30}$$

Using this relation one can show that $(\hat{f}^*)_{mn} = [(\hat{f}^t)_{nm}]^*$. To this aim one replaces $\hat{f}$ by $\hat{f}^t$ in Eq. (3.30) and uses that $(\hat{f}^t)^\dagger = [(\hat{f}^t)^t]^* = \hat{f}^*$.

Conversely, let us assume that $\hat{f}^\dagger$ is defined by means of its matrix elements on a complete basis as $(\hat{f}^\dagger)_{mn} = (f_{nm})^*$. This is equivalent to

$$\langle\, m \,|\, \hat{f}^\dagger\, n \,\rangle = \langle\, n \,|\, \hat{f}\, m \,\rangle^* = \langle\, \hat{f}\, m \,|\, n \,\rangle.$$

For arbitrary states $|\,\Psi\,\rangle = \sum_n a_n\, |\, n\,\rangle$ and $|\,\Phi\,\rangle = \sum_n b_n\, |\, n\,\rangle$ we have

$$\begin{aligned}\langle\, \Phi \,|\, \hat{f}^\dagger\, \Psi \,\rangle &= \sum_{m,n} b_m^*\, a_n \,\langle\, m \,|\, \hat{f}^\dagger\, n \,\rangle = \sum_{m,n} b_m^*\, a_n \,\langle\, \hat{f}\, m \,|\, n \,\rangle \\ &= \langle\, \hat{f}\, \Phi \,|\, \Psi \,\rangle,\end{aligned}$$

which means that the matrix-element relation (3.30) is equivalent to the definition of adjoint operator as given by Eq. (3.27) or (3.28).

After this mathematical digression we may return to the properties of operators associated to physical observables. Since $\langle\, f \,\rangle$ is real for all quantum states Ψ, all its eigenvalues f_n are real. Let Ψ_n be a complete set of eigenstates then

$$\hat{f}_{mn} = \delta_{mn}\, f_n \in \mathbb{R} \quad \Rightarrow \quad f_{mn} = (f_{nm})^* \quad \Rightarrow \quad \hat{f}_{mn} = (\hat{f}^\dagger)_{mn} \;\; \forall\; m, n.$$

This implies

$$\hat{f} = \hat{f}^\dagger.$$

One concludes that all operators $\hat{f}$ of real physical quantities are equal to their adjoint. These operators are called *self-adjoint* or *hermitian.*

It is easy to see that the converse is also true. If an operator is hermitian, i.e., $\hat{f}^\dagger = \hat{f}$, then all its eigenvalues f are real. In fact

$$\hat{f}\,|\,\Psi\,\rangle = f\,|\,\Psi\,\rangle \quad \Rightarrow \quad f = \langle\, \Psi \,|\, \hat{f} \,|\, \Psi \,\rangle = \langle\, \hat{f}\, \Psi \,|\, \Psi \,\rangle^* = \langle\, \Psi \,|\, \hat{f}^\dagger \,|\, \Psi \,\rangle^* = \langle\, \Psi \,|\, \hat{f} \,|\, \Psi \,\rangle^* = f^*.$$

This does not imply, of course, that any hermitian operator corresponds to a meaningful physical quantity.

Knowing that all operators of physical observables are hermitian (i.e., $\hat{f}^\dagger = \hat{f}$) we can prove the orthogonality of eigenstates corresponding to different eigenvalues directly. Let us consider two eigenstates

$$\hat{f}\,|\, n\,\rangle = f_n\,|\, n\,\rangle \qquad \text{and} \qquad \hat{f}\,|\, m\,\rangle = f_m\,|\, m\,\rangle.$$

It follows that

$$\langle\, m\,|\,\hat{f}\,n\,\rangle = \langle\, m\,|\,f_n\,n\,\rangle = f_n\,\langle\, m\,|\,n\,\rangle$$

and at the same time

$$\begin{aligned}\langle\, m\,|\,\hat{f}\,n\,\rangle &= \langle\, \hat{f}^\dagger\,m\,|\,n\,\rangle = \langle\, \hat{f}\,m\,|\,n\,\rangle && (\hat{f}^\dagger = \hat{f})\\ &= \langle\, f_m\,m\,|\,n\,\rangle = f_m^*\,\langle\, m\,|\,n\,\rangle = f_m\,\langle\, m\,|\,n\,\rangle && (f_m \in \mathbb{R}).\end{aligned}$$

Subtracting the two identities one obtains

$$0 = (f_m - f_n)\,\langle\, m\,|\,n\,\rangle,$$

which implies that $\langle\, m\,|\,n\,\rangle = 0$, whenever $f_m \neq f_n$.

In general f refers to a complete system of physical quantities f, g, h, etc. In this case we have operators $\hat{f}$, $\hat{g}$, $\hat{h}$, etc. and eigenvalues f_n, g_n, h_n with $n = 1, 2, \ldots$. The eigenstates Ψ_n must therefore be the solutions of a system of equations $\hat{f}\,\Psi = f\,\Psi$, $\hat{g}\,\Psi = g\,\Psi$, etc.

3.9. Sums, products and commutators of operators

Operators can be added and multiplied by constants. Since

$$\langle\, \alpha\,\hat{f} + \beta\,\hat{g}\,\rangle = \alpha\,\langle\, \hat{f}\,\rangle + \beta\,\langle\, \hat{g}\,\rangle$$

we must have

$$(\alpha\,\hat{f} + \beta\,\hat{g})\,\Psi = \alpha\,\hat{f}\,\Psi + \beta\,\hat{g}\,\Psi.$$

If the quantities f and g have a common complete set of eigenvalues f_n and g_n, i.e., if

$$\hat{f}\,\Psi_n = f_n\,\Psi_n \qquad \text{and} \qquad \hat{g}\,\Psi_n = g_n\,\Psi_n$$

for the same set of quantum states Ψ_n, we have

$$(\alpha\,\hat{f} + \beta\,\hat{g})\,\Psi_n = (\alpha\,f_n + \beta\,g_n)\,\Psi_n.$$

However, in the general case where $\hat{f}$ and $\hat{g}$ cannot be determined simultaneously with certainty, the eigenvalues of $\hat{f} + \hat{g}$ bear no relation to the eigenvalues of $\hat{f}$ and $\hat{g}$.

For any α and $\beta \in \mathbb{R}$ and hermitic operators $\hat{f} = \hat{f}^\dagger$ and $\hat{g} = \hat{g}^\dagger$ it is clear that

$$\begin{aligned}(\alpha\,\hat{f} + \beta\,\hat{g})^\dagger &= \alpha^*\,\hat{f}^\dagger + \beta^*\,\hat{g}^\dagger\\ &= \alpha\,\hat{f} + \beta\,\hat{g}.\end{aligned}$$

This implies that a linear combination of hermitic operators $\alpha\,\hat{f} + \beta\,\hat{g}$ with real coefficients α and β is also hermitic.

The multiplication of operators is defined like the composition of functions

$$\Psi \xrightarrow{\hat{f}} \hat{f}\,\Psi = \Phi \xrightarrow{\hat{g}} \chi = \hat{g}\,\Phi = \hat{g}\,(\hat{f}\,\Psi).$$

In general two operators $\hat{f}$ and $\hat{g}$ corresponding to two observables need not commute. In order to quantify this difference one actually defines the commutator

$$[\hat{f},\hat{g}] = \hat{f}\,\hat{g} - \hat{g}\,\hat{f},$$

which is in general a new operator. It is easy to see that

$$[\hat{f},\hat{g}] = -[\hat{g},\hat{f}]$$

and that

$$[\hat{f}\,\hat{g},\hat{h}] = \hat{f}\,[\hat{g},\hat{h}] + [\hat{f},\hat{h}]\,\hat{g}.$$

Notice that if $[\hat{f},\hat{h}] = 0$ and $[\hat{g},\hat{h}] = 0$, i.e., if $\hat{f}$ and $\hat{g}$ commute with $\hat{h}$, then $\hat{f}\,\hat{g}$ also commutes with $\hat{h}$. However, this does not imply that $\hat{f}$ and $\hat{g}$ commute.

The commutator of two operators corresponding to the observables f and g is very important, since it determines whether these two quantities can be measured simultaneously with certainty, i.e., whether they can take simultaneously definite values.

In order to show this, let us first consider the case where two observables f and g have definite values in the complete set of eigenfunctions Ψ_n, i.e.,

$$\hat{f}\,\Psi_n = f_n\,\Psi_n \qquad \text{and} \qquad \hat{g}\,\Psi_n = g_n\,\Psi_n.$$

It then follows that

$$(\hat{f}\,\hat{g})\,\Psi_n = \hat{f}\,(\hat{g}\,\Psi_n) = g_n\,(\hat{f}\,\Psi_n) = g_n\,f_n\,\Psi_n \tag{3.31}$$

and analogously

$$(\hat{g}\,\hat{f})\,\Psi_n = \hat{g}\,(\hat{f}\,\Psi_n) = f_n\,(\hat{g}\,\Psi_n) = f_n\,g_n\,\Psi_n. \tag{3.32}$$

Comparing Eqs. (3.31) and (3.32) we conclude that

$$(\hat{g}\,\hat{f})\,\Psi_n = (\hat{f}\,\hat{g})\,\Psi_n$$

for all states of a complete basis. Thus,

$$\hat{g}\,\hat{f} = \hat{f}\,\hat{g}$$

and the operators commute.

The proof of the converse is more subtle. In this case we consider a complete basis of eigenstates of $\hat{f}$, which exists since $\hat{f} = \hat{f}^\dagger$, and look for a basis of eigenstates of $\hat{g}$. Let us assume that $[\hat{f},\hat{g}] = \hat{f}\,\hat{g} - \hat{g}\,\hat{f} = 0$ and consider the matrix elements

$$\langle\, m\,|\,\hat{f}\,\hat{g}\,|\,n\,\rangle = \langle\, m\,|\,\hat{g}\,\hat{f}\,|\,n\,\rangle, \tag{3.33}$$

where $|\,m\,\rangle$ and $|\,n\,\rangle$ are eigenstates of $\hat{f}$. Using that $\hat{f}^\dagger = \hat{f}$ and that $f_n \in \mathbb{R}$ we have

$$\langle\, m\,|\,\hat{f}\,\hat{g}\,|\,n\,\rangle = \langle\,\hat{f}^\dagger\,m\,|\,\hat{g}\,|\,n\,\rangle = \langle\,\hat{f}\,m\,|\,\hat{g}\,|\,n\,\rangle = \langle\, f_m\,m\,|\,\hat{g}\,|\,n\,\rangle = f_m^*\langle\, m\,|\,\hat{g}\,|\,n\,\rangle = f_m\,\langle\, m\,|\,\hat{g}\,|\,n\,\rangle$$

and

$$\langle\, m \,|\, \hat{g}\, \hat{f} \,|\, n \,\rangle = \langle\, m \,|\, \hat{g}\, f_n \,|\, n \,\rangle = f_n \,\langle\, m \,|\, \hat{g} \,|\, n \,\rangle.$$

Using Eq. (3.33) we obtain

$$(f_m - f_n)\,\langle\, m \,|\, \hat{g} \,|\, n \,\rangle = 0.$$

Consequently, $\langle\, m \,|\, \hat{g} \,|\, n \,\rangle = g_{mn}$ vanishes for every pair of eigenstates of $\hat{f}$ with different eigenvalues $f_m \neq f_n$. Moreover, since $\hat{g}^\dagger = \hat{g}$, linear combinations of the eigenstates of $\hat{f}$ with the same eigenvalues f_n allow one to diagonalize $\hat{g}$ within these subspaces so that one can achieve $g_{mn} = 0 \;\; \forall\; m \neq n$. This implies that $\hat{g}$ and $\hat{f}$ have a common complete set of eigenstates.

Examples:

Consider the different components $(\hat{x}, \hat{y}, \hat{z})$ of the position operator $\hat{\vec{r}}$. It is clear that

$$\hat{x}\, \hat{y}\, \Psi(\vec{r}) = x\, y\, \Psi(\vec{r}) = y\, x\, \Psi(\vec{r}) = \hat{y}\, \hat{x}\, \Psi(\vec{r}).$$

Therefore, all the components x_i of the position of the electron, or even of a system of many electrons, can be determined simultaneously with arbitrary accuracy. The same holds for all the components of the momentum since

$$\begin{aligned}\hat{p}_x\, \hat{p}_y\, \Psi(\vec{r}) &= \left(-i\,\hbar\, \frac{\partial}{\partial x}\right)\left(-i\,\hbar\, \frac{\partial \Psi}{\partial y}\right) = \\ &= -\hbar^2\, \frac{\partial^2\, \Psi}{\partial x\, \partial y} = -\hbar^2\, \frac{\partial^2\, \Psi}{\partial y\, \partial x} = \hat{p}_y\, \hat{p}_x\, \Psi(\vec{r}).\end{aligned}$$

However, the same components of the position and momentum operators $\hat{x}$ and $\hat{p}_x$ do not commute:

$$\begin{aligned}[\hat{x}, \hat{p}_x]\, \Psi = (\hat{x}\, \hat{p}_x - \hat{p}_x\, \hat{x})\, \Psi &= x\left(-i\,\hbar\, \frac{\partial \Psi}{\partial x}\right) - \left(-i\,\hbar\, \frac{\partial (x\, \Psi)}{\partial x}\right) \\ &= -i\,\hbar\, x\, \frac{\partial \Psi}{\partial x} + i\,\hbar\, \Psi + i\,\hbar\, x\, \frac{\partial \Psi}{\partial x} = i\,\hbar\, \Psi.\end{aligned}$$

Consequently,

$$[\hat{x}, \hat{p}_x] = i\,\hbar.$$

One concludes that position and momentum cannot be determined simultaneously with arbitrary accuracy, in agreement with the Heisenberg principle. Of course, different components of position and momentum commute (e.g, $[\hat{x}, \hat{p}_y] = 0$), so that in general we have

$$[\hat{x}_k, \hat{p}_l] = i\,\hbar\, \delta_{kl}.$$

Summarizing, we have achieved the following important result: Two observables f and g can be measured *simultaneously* to an *arbitrary accuracy*, i.e., they can have *simultaneously definite values*, if and only if the corresponding operators $\hat{f}$ and $\hat{g}$ commute. Therefore, if the operators commute one says that the observables are *compatible.*

The simplest example of a product of operators is an operator raised to some integer power $\hat{f}^\nu$ ($\nu \in \mathbb{N}$). It is clear that the eigenvalues of $\hat{f}^\nu$ are the ν-th power of the eigenvalues f_n of $\hat{f}$. For any function $\Phi(f)$ admitting a Taylor expansion $\Phi(f) = \sum_\nu \alpha_\nu f^\nu$ the operator $\Phi(\hat{f})$ has the same eigenstates Ψ_n as $\hat{f}$ and the eigenvalues $\Phi(f_n)$. It is easy to see that if $\hat{f}$ is hermitic ($\hat{f}^\dagger = \hat{f}$) then $[\Phi(\hat{f})]^\dagger = \Phi(\hat{f})$. To show this it is enough to prove the hermiticity of $\hat{f}^\nu$ by induction: $\langle\, m \,|\, \hat{f}^{\nu+1} \,|\, n \,\rangle = \langle\, m \,|\, \hat{f}\, \hat{f}^\nu \,|\, n \,\rangle = \langle\, \hat{f}^\dagger m \,|\, \hat{f}^\nu \,|\, n \,\rangle = \langle\, \hat{f}\, m \,|\, \hat{f}^\nu\, n \,\rangle = \langle\, (\hat{f}^\nu)^\dagger\, \hat{f}\, m \,|\, n \,\rangle = \langle\, \hat{f}^{\nu+1}\, m \,|\, n \,\rangle = \langle\, n \,|\, \hat{f}^{\nu+1} \,|\, m \,\rangle^*$, where we have used the induction hypothesis $(\hat{f}^\nu)^\dagger = \hat{f}^\nu$.

It is important to note that the product $\hat{f}\,\hat{g}$ of two hermitian operators $\hat{f} = \hat{f}^\dagger$ and $\hat{g} = \hat{g}^\dagger$ is *not necessarily* hermitian, unless the two operators commute, since $(\hat{f}\,\hat{g})^\dagger = \hat{g}^\dagger \hat{f}^\dagger$. In the latter case the eigenstate basis of the product operator is the eigenstate basis common to $\hat{f}$ and $\hat{g}$ and has $f_n\, g_n$ as eigenvalues. This is due to the particular form of the transpose of a product of operators

$$\begin{aligned}
\int \Phi\,(\hat{f}\,\hat{g}\,\Psi)\, dq &= \int \Phi\, \Big[\hat{f}\,(\hat{g}\,\Psi)\Big]\, dq = \int (\hat{g}\,\Psi)\;(\hat{f}^t\,\Phi)\, dq \\
&= \int (\hat{f}^t\,\Phi)\;(\hat{g}\,\Psi)\, dq = \int \Psi\, \Big[\hat{g}^t\,(\hat{f}^t\,\Phi)\Big]\, dq \\
&= \int \Psi\, \hat{g}^t\, \hat{f}^t\, \Phi\, dq.
\end{aligned}$$

This implies that $(\hat{f}\,\hat{g})^t = \hat{g}^t\,\hat{f}^t$. Since $(\hat{f}\,\hat{g}\,\Psi)^* = [\hat{f}\,(\hat{g}\,\Psi)]^* = \hat{f}^*\,(\hat{g}\,\Psi)^* = \hat{f}^*\,\hat{g}^*\,\Psi$ we have

$$(\hat{f}\,\hat{g})^\dagger = \hat{g}^\dagger\,\hat{f}^\dagger.$$

For hermitic operators $\hat{f} = \hat{f}^\dagger$ and $\hat{g} = \hat{g}^\dagger$ this implies

$$(\hat{f}\,\hat{g})^\dagger = \hat{g}\,\hat{f}.$$

It is easy to verify that for any $\hat{f}$ and $\hat{g}$, $(\hat{f}\,\hat{g} + \hat{g}\,\hat{f})$ and $i[\hat{f}, \hat{g}] = i(\hat{f}\,\hat{g} - \hat{g}\,\hat{f})$ are always hermitian.

Exercise: Using these properties show that the orbital angular momentum operator $\hat{\vec{l}} = \hat{\vec{r}} \times \hat{\vec{p}}$ is hermitian.

3.10. Dirac's delta function $\delta(x)$

Before discussing the properties of eigenfunctions having a continuous spectrum of eigenvalues it is useful to recall the main properties and some important representations of Dirac's delta function.

Dirac's delta function $\delta(x)$ is defined by the following conditions:

$$\delta(x) = 0 \quad \forall\; x \neq 0 \tag{3.34}$$

and

$$\int_{-\infty}^{+\infty} \delta(x)\, dx = 1. \tag{3.35}$$

Clearly, $\delta(x)$ is not well defined as a function since (3.34) and (3.35) imply $\delta(x) \xrightarrow{x\to 0} +\infty$. Moreover, from (3.34) and (3.35) we have

$$\int_{-\varepsilon}^{\varepsilon} \delta(x)\, dx = 1 \quad \forall\, \varepsilon > 0. \tag{3.36}$$

It follows that the integral

$$\int_{-\infty}^{x} \delta(x')\, dx' = \theta(x) = \begin{cases} 0 & x < 0 \\ 1/2 & x = 0 \\ 1 & x > 0 \end{cases}$$

is the Heaviside step function. The value for $x = 0$ follows from the fact that $\delta(x)$ is even, i.e., $\delta(x) = \delta(-x)\ \ \forall\, x$, and thus $\displaystyle\int_{-\infty}^{0} \delta(x')\, dx' = \int_{0}^{+\infty} \delta(x')\, dx' = 1/2$.

From the fundamental theorem of analysis follows the differential representation of the delta function:

$$\delta(x) = \frac{d\,\theta}{dx}\,(x).$$

The delta function is not a function in the usual sense, but a *generalized function* or *distribution function* which are defined as the *regular sequence* of particularly well-behaved functions. To be more explicit mathematically, a particularly well-behaved function has derivatives of all orders at all points and falls at least as fast as $|x|^{-n}$ for $x \to \infty$ for all n. A sequence of particularly well-behaved functions $f_\tau(x)$ with respect to a parameter τ ($\tau > 0$) is said to be a regular sequence when $\displaystyle\lim_{\tau\to 0} \int_{-\infty}^{+\infty} f_\tau(x)\ g(x)\, dx$ exists for any particularly well-behaved function $g(x)$.

The central property of the delta function is that for any function that is continuous at $x = 0$ it holds

$$\int_{-\infty}^{+\infty} f(x)\ \delta(x)\, dx = f(0). \tag{3.37}$$

This follows from $\delta(x) = 0$ for $x \neq 0$, or equivalently $\displaystyle\int_{-\varepsilon}^{\varepsilon} \delta(x)\, dx = 1\ \ \forall\, \varepsilon > 0$, and the continuity of $f(x)$. One can immediately generalize Eq. (3.37) as

$$\int f(x)\ \delta(x-a)\, dx = f(a) \tag{3.38}$$

for all functions f and points a, provided that $f(x)$ is continuous at $x = a$ and that the integration range includes the point $x = a$.

There are many useful representations of $\delta(x)$, as a derivative, regular sequence of functions, integral or series:

i) $\delta(x) = \dfrac{d\theta}{dx}$, where $\theta(x)$ is the Heaviside step function.

ii) $\delta^{(3)}(\vec{r}) = \delta(x)\,\delta(y)\,\delta(z) = -\dfrac{1}{4\pi}\,\nabla^2\left(\dfrac{1}{r}\right)$.

This can be easily proved by applying Gauss theorem to the vector field $\vec{E} = -\vec{\nabla}\left(\dfrac{1}{r}\right) = \dfrac{\vec{r}}{r^3}$.

iii) $\delta(x) = \lim\limits_{\varepsilon\to 0} \delta_\varepsilon(x)$, where $\delta_\varepsilon(x) = \begin{cases} 1/\varepsilon & \text{for } |x| < \varepsilon/2 \\ 0 & \text{otherwise} \end{cases}$ is a square function.

iv) $\delta(x) = \lim\limits_{\sigma\to 0} \delta_\sigma(\varepsilon)$, where $\delta_\sigma(\varepsilon) = \dfrac{1}{\sqrt{2\pi\sigma^2}}\, e^{-\frac{x^2}{2\sigma^2}}$ is a Gaussian function.

v) $\delta(x) = \lim\limits_{\lambda\to 0} \delta_\lambda(\varepsilon)$, where $\delta_\lambda(\varepsilon) = \dfrac{1}{\pi}\,\dfrac{\lambda}{\lambda^2 + x^2}$ is a Lorentzian function.

vi) $\delta(x) = \lim\limits_{K\to +\infty} \delta_K(\varepsilon)$, where $\delta_K(\varepsilon) = \dfrac{1}{2\pi}\displaystyle\int_{-K}^{K} e^{i\,k\,x} dk = \dfrac{1}{\pi}\,\dfrac{\sin(Kx)}{x}$.

Of course, these limits do not exist in the sense of the usual definition. They have to be interpreted in the sense of a regular sequence, i.e., inside an integral. In fact, the limit $\lim\limits_{\varepsilon\to 0}\displaystyle\int \delta_\varepsilon(x)\; f(x)\, dx$ always exists. It is in this sense that the integral

$$\int \delta(x)\; f(x)\, dx = \lim_{\varepsilon\to 0}\int \delta_\varepsilon(x)\; f(x)\, dx$$

is to be interpreted.

Among other useful properties of the δ-function we have

$$\delta(ax) = \frac{\delta(x)}{|a|}$$

or the more general expression

$$\delta(f(x)) = \sum_i \frac{\delta(x - x_i)}{|f'(x_i)|},$$

where x_i are the roots of $f(x_i) = 0$ and $f'(x_i) \neq 0$. These relations can be proved straightforwardly:

$$\int_{-\infty}^{+\infty} \delta(ax)\; f(x)\, dx = \frac{1}{|a|}\int_{-\infty}^{+\infty} \delta(x')\; f\left(\frac{x'}{|a|}\right)\, dx' = \frac{1}{a}\, f(0).$$

3.10.1. Fourier integral representation of $\delta(x)$

The Fourier integral theorem states that we can express any function $f(x)$ in the form

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} F(k)\, e^{i\,k\,x}\, dk, \tag{3.39}$$

where the Fourier transform is

$$F(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f(x)\, e^{-i\,k\,x}\, dx. \tag{3.40}$$

Replacing (3.40) in (3.39) we obtain

$$\begin{aligned} f(x) &= \frac{1}{2\pi} \int_{-\infty}^{+\infty} \left(\int_{-\infty}^{+\infty} f(x')\, e^{-i\,k\,x'}\, dx' \right) e^{i\,k\,x}\, dk \\ &= \int_{-\infty}^{+\infty} \left(\frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{i\,k\,(x-x')}\, dk \right) f(x')\, dx'. \end{aligned}$$

Consequently,

$$\delta(x - x') = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{i\,k\,(x-x')}\, dk \tag{3.41}$$

is an integral representation of the δ-function. As in any representation Eq. (3.41) is used in connection with an integral of the form $\int \delta(x)\, f(x)\, dx$.

3.10.2. Representation of $\delta(x)$ by a complete basis of orthonormal functions

Consider a complete orthonormal basis in a Hilbert space $\{\Psi_n(x),\, n = 1, 2, \ldots\}$. Completeness means that an arbitrary function can be written in the form

$$\Psi(x) = \sum_n a_n\, \Psi_n(x). \tag{3.42}$$

Using the orthonormality of the basis we have

$$a_n = \int \Psi_n^*(x)\, \Psi(x)\, dx, \tag{3.43}$$

and substituting (3.43) in (3.42) we obtain

$$\Psi(x) = \sum_n \left(\int \Psi_n^*(x')\, \Psi(x')\, dx' \right) \Psi_n(x)$$

or

$$\Psi(x) = \int dx' \left[\sum_n \Psi_n^*(x')\, \Psi_n(x) \right] \Psi(x').$$

Therefore

$$\delta(x' - x) = \sum_n \Psi_n^*(x')\, \Psi_n(x). \tag{3.44}$$

In general $\Psi_n(x)$ is either real or $\Psi_n^*(x)$ is another element of the basis for a different n so that $\delta(x - x')$ as given by Eq. (3.44) can be shown to be real and symmetric [i.e., $\delta(x - x') = \delta(x' - x)$].

3.11. The continuum spectrum

The properties of eigenvalues and eigenvectors discussed so far for observables having a discrete spectrum can be generalized with a few but significant changes to observables having a continuous spectrum. Let f be an observable (actually a complete set of observables) having a continuous spectrum. Assuming that all possible values of f are taken into account we can expand any wave function in the form of an integral

$$\Psi(q) = \int a_f\, \Psi_f(q)\, df, \tag{3.45}$$

where a_f are complex coefficients that depend on the continuous variable f. The wave functions $\Psi_f(q)$ are eigenstates of the operator $\hat{f}$ corresponding to the observable f: $\hat{f}\, \Psi_{f'} = f'\, \Psi_{f'}$ or $\hat{f} |\, f' \,\rangle = f' |\, f' \,\rangle$. In the case of a continuous spectrum the normalization of the eigenstates $\Psi_f(q)$ is more complex and we cannot require that the norm is equal to 1. We look however for a normalization in which the probability P that a measurement of the physical quantity f in a system with wave function $\Psi(q)$ yields a value f' in the range $(f, f + df)$ satisfies the proportionality relation

$$P[f' \in (f, f + df)] \propto |a_f|^2\, df. \tag{3.46}$$

In other words, the square modulus of the expansion coefficient $|a_f|^2$ is proportional to the probability density of measuring the value f, if the system is in the state $\Psi(q)$ [see Eq. (3.45)]. If the wave function is normalized ($\int |\Psi|^2\, dq = 1$) we must have

$$\int |a_f|^2\, df = 1.$$

Otherwise we must divide $|a_f|^2$ by $\int |\Psi|^2\, dq$. In general we have

$$\int |a_f|^2\, df = \int |\Psi(q)|^2\, dq. \tag{3.47}$$

From Eq. (3.45) we have

$$\Psi^*(q) = \int a_f^*\, \Psi_f^*(q)\, dq$$

and thus

$$\begin{aligned}\int a_f^*\, a_f\, df &= \int \left(\int a_f^*\, \Psi_f^*(q)\, df \right) \Psi(q)\, dq \\ &= \int a_f^* \left(\int \Psi_f^*(q)\, \Psi(q)\, dq \right) df. \qquad (3.48)\end{aligned}$$

Since a_f is an arbitrary complex function of f, we may consider a_f and a_f^* as independent variables. Taking $\partial/\partial a_f^*$ in Eq. (3.48), keeping a_f and thus $\Psi(q)$ constant, we obtain

$$a_f = \int \Psi_f^*(q)\, \Psi(q)\, dq = \langle\, f \,|\, \Psi \,\rangle. \qquad (3.49)$$

This expression gives expansion coefficients of $\Psi(q)$ in terms of the eigenstates $\Psi_f(q)$ of the continuum spectrum. It is analogous to the corresponding expression $a_n = \int \Psi_n^*\, \Psi\, dq = \langle\, n \,|\, \Psi \,\rangle$ for the discrete spectrum.

Replacing Eq. (3.45) in Eq. (3.49) we obtain

$$\begin{aligned}a_f &= \int \Psi_f^*(q) \left(\int a_{f'}\, \Psi_{f'}(q)\, df' \right) dq \\ &= \int a_{f'} \left(\int \Psi_f^*(q)\, \Psi_{f'}(q)\, dq \right) df'.\end{aligned}$$

Since this holds for any function a_f of f we must have

$$\int \Psi_f^*(q)\, \Psi_{f'}(q)\, dq = \delta(f' - f) = \langle\, f \,|\, f' \,\rangle. \qquad (3.50)$$

This is the *orthonormalization condition* for the eigenfunctions of the continuum spectrum. Notice that this follows simply from Eq. (3.47) and the expansion (3.45). It replaces the orthonormalization rule

$$\int \Psi_m^*(q)\, \Psi_n(q)\, dq = \delta_{mn} = \langle\, n \,|\, m \,\rangle$$

for a discrete spectrum, basically by replacing Kronecker's delta by Dirac's delta function.

We may now derive the *completeness relation* in the continuum spectrum. For this we replace Eq. (3.49) in Eq. (3.45), which gives

$$\begin{aligned}\Psi(q) &= \int \left(\int \Psi_f^*(q')\, \Psi(q')\, dq' \right) \Psi_f(q)\, df \\ &= \int \left(\int \Psi_f^*(q')\, \Psi_f(q)\, df \right) \Psi(q')\, dq'.\end{aligned}$$

This implies

$$\int \Psi_f^*(q')\, \Psi_f(q)\, df = \delta(q' - q). \qquad (3.51)$$

The analogous relation for a discrete spectrum is

$$\sum_n \Psi_n^*(q')\,\Psi_n(q) = \delta(q'-q), \tag{3.52}$$

which was demonstrated in section 3.10.2 [Eq. (3.44)].

Using Dirac's notation we can write

$$|\,\Psi\,\rangle = \int a_f\,|\,f\,\rangle\,df \tag{3.53}$$

and

$$\langle\,f\,|\,f'\,\rangle = \int \Psi_f^*(q)\,\Psi_{f'}(q)\,dq = \delta(f'-f), \tag{3.54}$$

while in the discrete spectrum we have

$$|\,\Psi\,\rangle = \sum_n a_n\,|\,n\,\rangle$$

and

$$\langle\,m\,|\,n\,\rangle = \int \Psi_m^*(q)\,\Psi_n(q)\,dq = \delta_{mn}.$$

From Eqs. (3.53) and (3.54) we obtain

$$\langle\,f\,|\,\Psi\,\rangle = \int a_{f'}\,\langle\,f\,|\,f'\,\rangle\,df' = \int a_{f'}\,\delta(f'-f)\,df' = a_f \tag{3.55}$$

which coincides with Eq. (3.49) noting that

$$\langle\,f\,|\,\Psi\,\rangle = \int \Psi_f^*(q)\,\Psi(q)\,dq. \tag{3.56}$$

Finally, replacing Eq. (3.55) in Eq. (3.53) we have

$$|\,\Psi\,\rangle = \int |\,f\,\rangle\,\langle\,f\,|\,\Psi\,\rangle\,df$$

which implies

$$\int df\,|\,f\,\rangle\,\langle\,f\,| = \mathbb{1}. \tag{3.57}$$

This is the completeness relation for the continuum spectrum in analogy with the known expression

$$\sum_n |\,n\,\rangle\,\langle\,n\,| = \mathbb{1}$$

for a discrete spectrum.

In particular the position is a complete set of eigenvalues. Therefore,

$$\int dq \,|\, q \,\rangle \,\langle\, q \,| = \mathbb{1} \quad \text{or} \quad \int d^3 r \,|\, \vec{r} \rangle \,\langle\, \vec{r} \,| = \mathbb{1}.$$

We look now for the *eigenstates of the position operator* $\hat{q} = q$, which we know is simply the multiplication by q ($q = x, y, z$). Let $\Psi_{q'}(q)$ be the eigenfunction with defined position q'. Then we must have

$$\hat{q}\, \Psi_{q'}(q) = q\, \Psi_{q'}(q), \tag{3.58}$$

since the operator $\hat{q}$ is simply multiplication by q, and

$$\hat{q}\, \Psi_{q'}(q) = q'\, \Psi_{q'}(q) \tag{3.59}$$

for all q. Notice that Eq. (3.58) follows from the definition of the position operator $\hat{q}$ which holds for any $\Psi(q)$, while Eq. (3.59) is the condition for $\Psi_{q'}(q)$ to be an eigenfunction of $\hat{q}$ with the eigenvalue q'. This implies

$$(q - q')\, \Psi_{q'}(q) = 0 \quad \forall\, q.$$

Consequently,

$$\Psi_{q'}(q) = 0 \quad \forall\, q \neq q'.$$

The *normalization condition* (3.50) reads

$$\int \Psi^*_{q'}(q)\, \Psi_{q''}(q)\, dq = \delta(q'' - q'),$$

which is satisfied if we set

$$\Psi_{q'}(q) = \delta(q - q').$$

The *expansion coefficient* $a_{q'}$ of the wave function $\Psi(q)$ is then given by

$$a_{q'} = \int \Psi^*_{q'}(q)\, \Psi(q)\, dq = \int \delta(q - q')\, \Psi(q)\, dq = \Psi(q').$$

The probability of measuring a position $q' \in (q, q + dq)$ is $|a_q|^2\, dq = |\Psi(q)|^2\, dq$, as it should.

In Dirac's ket-bra form we denote the ket with defined position q' by $|\, q' \,\rangle$. A general vector state $|\, \Psi \,\rangle$ would be expanded like

$$|\, \Psi \,\rangle = \int a_{q'} \,|\, q' \,\rangle\, dq' = \int \Psi(q') \,|\, q' \,\rangle\, dq'.$$

Thus

$$\langle\, q \,|\, \Psi \,\rangle = a_q = \Psi(q).$$

The wave function is the projection of the state $|\, \Psi \,\rangle$ in the eigenstate with defined position q. Moreover, if we want to change the representation, i.e., the complete set of observables used to describe the state of the system, we can write

$$\begin{aligned}
\Psi(q) = \langle\, q \,|\, \Psi \,\rangle &= \langle\, q \,|\left(\int |\, f \,\rangle \langle\, f \,|\, df\right)|\, \Psi \,\rangle \\
&= \int \underbrace{\langle\, q \,|\, f \,\rangle}_{\Psi_f(q)}\, \underbrace{\langle\, f \,|\, \Psi \,\rangle}_{a_f}\, df \\
&= \int a_f\, \Psi_f(q)\, df.
\end{aligned}$$

All relations and transformations between the various representations can be obtained by using Dirac's ket-bra notation, together with the orthonormalization and completeness relations.

The projection of the ket $|\, \Psi \,\rangle$ in the eigenstates of $\hat{f}$, namely

$$\langle\, f \,|\, \Psi \,\rangle = a_f,$$

is analogous to the wave function $\Psi(q) = a_q$ with the only difference that in the case of a_f the observable f has been considered to represent the state and not the position q. Actually $a_f = \langle\, f \,|\, \Psi \,\rangle$ defines the state $|\, \Psi \,\rangle$ as much as $a_q = \langle\, q \,|\, \Psi \,\rangle = \Psi(q)$ defines it. Therefore $a_f = \langle\, f \,|\, \Psi \,\rangle$ is known as the wave function in the f representation.

A particularly relevant example of a representation in the continuum is the momentum representation

$$\Phi(\vec{p}) = \langle\, \vec{p} \,|\, \Psi \,\rangle.$$

Let us consider the eigenstates

$$\hat{p}_x \,|\, p'_x \,\rangle = p'_x \,|\, p'_x \,\rangle.$$

Knowing that in coordinate representation $\hat{p}_x = -i\,\hbar\,\dfrac{\partial}{\partial x}$ we have

$$\langle\, x \,|\, \hat{p}_x \,|\, p'_x \,\rangle = -i\,\hbar\,\frac{\partial}{\partial x}\,\langle\, x \,|\, p'_x \,\rangle = p'_x\,\langle\, x \,|\, p'_x \,\rangle.$$

The solution to this differential equation with respect to the variable x has the form

$$\langle\, x \,|\, p'_x \,\rangle = A\, e^{\frac{i}{\hbar}\, p'_x\, x},$$

where A is a constant to be determined from the normalization condition

$$\begin{aligned}\langle\, p_x'' \,|\, p_x' \,\rangle &= \int \langle\, p_x'' \,|\, x \,\rangle \, \langle\, x \,|\, p_x' \,\rangle \, dx = |A|^2 \int e^{\frac{i}{\hbar}\,(p_x' - p_x'')} \, dx \\ &= |A|^2 \, \hbar \int \frac{dx}{\hbar} \, e^{\frac{i}{\hbar}\,(p_x' - p_x'')\,x} = 2\,\pi\,\hbar\,|A|^2\,\delta(p_x' - p_x'').\end{aligned}$$

Consequently $|A|^2 = (2\,\pi\,\hbar)^{-1}$ and

$$\langle\, x \,|\, p_x \,\rangle = \frac{1}{\sqrt{2\,\pi\,\hbar}} \, e^{\frac{i}{\hbar}\,p_x\,x}.$$

Finally, let us mention that there are observables with partly discrete and continuous spectra (e.g., the energy of atoms). In this case we have

$$|\,\Psi\,\rangle = \sum_n a_n \,|\, n \,\rangle + \int a_f \,|\, f \,\rangle \, df,$$

where $|\, n \,\rangle$ and $|\, f \,\rangle$ refer, respectively, to the eigenstates of the discrete and continuous spectra.

4. Quantum dynamics

In this section we focus on the time dependence of the wave function, the corresponding kets, and observable quantities.

4.1. The Hamilton operator

From the superposition principle we know that the equations governing the time dependence of the wave function must be linear and homogeneous. Moreover, since the knowledge of $\Psi(q)$ at any given time t_0 defines the state at any future time t, they must involve only first-order derivatives with respect to time. In the most general form we can write

$$i\,\hbar\,\frac{\partial\Psi}{\partial t} = \hat{H}\,\Psi, \tag{4.1}$$

where $\hat{H}$ is a linear operator known as the *Hamiltonian* or *Hamilton operator*. The factor $i\,\hbar$ is introduced for convenience. We would like to demonstrate some general properties of $\hat{H}$.

If a wave function is normalized at a given time

$$\int |\Psi(q,t_0)|^2\,dq = 1,$$

then $|\Psi(q,t)|^2\,dq$ represents the probability for an electron to have a coordinate $q' \in (q, q+dq)$ and must remain normalized at any future time. We must therefore have

$$0 = \frac{\partial}{\partial t}\int |\Psi(q,t)|^2\,dq = \int \frac{\partial\Psi^*}{\partial t}(q)\,\Psi(q)\,dq + \int \Psi^*(q)\,\frac{\partial\Psi}{\partial t}(q)\,dq$$

for all t. Replacing Eq. (4.1) and its complex conjugate $-i\,\hbar\,\dfrac{\partial\Psi^*}{\partial t} = (H\,\Psi)^* = \hat{H}^*\,\Psi^*$ we have

$$\begin{aligned} 0 &= \int \Big[-(\hat{H}^*\,\Psi^*)\,\Psi\,dq + \Psi^*(\hat{H}\,\Psi)\Big]\,dq \\ &= \int (\Psi^*\,H\,\Psi - \Psi^*\,H^\dagger\,\Psi)\,dq \\ &= \int \Psi^*(H - H^\dagger)\,\Psi\,dq. \end{aligned}$$

Since this holds for any state Ψ, we must have

$$\hat{H} = \hat{H}^\dagger.$$

The Hamiltonian is therefore hermitian.

As in classical mechanics the Hamiltonian defines the dynamics of the system. If a system is isolated, i.e., it is not subject to any varying external fields, the Hamiltonian cannot depend on time. In fact, in the absence of varying external field all times are equivalent for the system and therefore the laws for the time evolution must be the same at all times.

4.2. Stationary states

Since the Hamiltonian is hermitic and time-independent (isolated system), its eigenstates and eigenvalues given by

$$\hat{H}\,\Psi_n = E_n\,\Psi_n$$

are also time-independent. As in classical mechanics the quantity associated to the Hamiltonian and which is conserved in isolated systems is the energy of the system. The eigenvalues E_n of $\hat{H}$ are known as *eigenenergies.*

The time dependence of the eigenstates of $\hat{H}$ is very easy to solve. It is given simply by a phase factor, since

$$i\hbar\,\frac{\partial\Psi_n}{\partial t} = \hat{H}\,\Psi_n = E_n\,\Psi_n$$

implies

$$\Psi_n(q,t) = e^{-\frac{i}{\hbar}E_n(t-t_0)}\,\Psi(q,t_0). \tag{4.2}$$

The probability amplitude a_f for any time-independent observable $\hat{f}$ ($\partial\hat{f}/\partial t = 0$) in the state $\Psi_n(q,t)$ is given by

$$a_f(t) = \int \Psi_f^*(q)\,\Psi_n(q,t)\,dq = e^{-\frac{i}{\hbar}E_n(t-t_0)}\int \Psi_f^*(q)\,\Psi_n(q,t_0)\,dq.$$

Consequently, the probability $|a_f|^2$ of measuring the value f and the expectation value $\langle\,\Psi_n\,|\,\hat{f}\,|\,\Psi_n\,\rangle$ are independent of t. Notice that we have just used that the operator $\hat{f}$ and thus its eigenstates $\Psi_f(q)$ do not depend on time. In particular the probability density $|\Psi_n(q,t)|^2$ for any value of the position is time-independent in an eigenstate of $\hat{H}$ [see also Eq. (4.2)]. For these reasons the eigenstates of $\hat{H}$ are called *stationary states.*

The stationary state with the lowest energy is called the *ground-state* or *normal state* of the system and the associated energy eigenvalue E_0 is the *ground-state energy.* The stationary states with higher energy eigenvalues $E_n > E_0$ are called *excited states.*

An arbitrary wave function at a given time t_0 can be expanded in terms of the wave functions $\Psi_n(q)$ of stationary states as

$$\Psi(q,t_0) = \sum_n a_n\,\Psi_n(q),$$

where the $\Psi_n(q)$ are the solutions of the eigenvalue equation $H\,\Psi_n = E_n\,\Psi_n$. As usual the square modulus $|a_n|^2$ represents the probability of finding the system in the state Ψ_n. The time dependence of Ψ is then given by

$$\Psi(q,t) = \sum_n a_n\,\Psi_n(q)\;e^{-\frac{i}{\hbar}E_n(t-t_0)}.$$

This is the result of the superposition of the time dependences of the different stationary states composing $\Psi(q,t_0)$. Notice that as soon as more than one value of E_n is involved in the expansion of $\Psi(q)$, the probablity density $|\Psi(q)|^2$ oscillates in time and is therefore not stationary.

4.3. Conserved quantities

We have seen that two physical observables f and g can be measured simultaneously if and only if the corresponding operators commute ($[\hat{f},\hat{g}]=0$). A quantity or observable f that can be measured simultaneously with the energy, and whose operator $\hat{f}$ is independent of time, is called a *conserved quantity.* The reason for this name is that under these conditions the average value

$$\langle\, f\,\rangle = \langle\,\Psi\,|\,\hat{f}\,|\,\Psi\,\rangle$$

in any state $|\,\Psi\,\rangle$ is independent of time. In fact, also the probability distribution $|a_n|^2$ of the outcome of the measurement of the observable $\hat{f}$ is independent of t. The proof of this important property is actually very simple. Since f and E can be measured simultaneously, one can find a complete set of eigenstates of $\hat{H}$ that are also eigenstates of $\hat{f}$:

$$\hat{H}\,|\,n\,\rangle = E_n\,|\,n\,\rangle$$

and

$$\hat{f}\,|\,n\,\rangle = f_n\,|\,n\,\rangle,$$

where E_n, f_n and $|\,n\,\rangle$ are independent of time. For any state

$$|\,\Psi(t)\,\rangle = \sum_n a_n\, e^{-\frac{i}{\hbar}\,E_n\,t}\,|\,n\,\rangle$$

we have

$$\begin{aligned}\langle\,\hat{f}\,\rangle = \langle\,\Psi\,|\,\hat{f}\,|\,\Psi\,\rangle &= \sum_{nm} a_m^*\,a_n\,e^{\frac{i}{\hbar}\,(E_m-E_n)\,t}\,\underbrace{\langle\,m\,|\,\hat{f}\,|\,n\,\rangle}_{\delta_{mn}\,f_n}\\ &= \sum_n |a_n|^2\,f_n.\end{aligned}$$

Since the expansion coefficient at time t is $a_n(t) = a_n(0)\,e^{-\frac{i}{\hbar}E_n t}$, it is clear that the probability $|a_n|^2$ of finding the system in state $|\,n\,\rangle$, or measuring the value f_n, is independent of t for all $|\,n\,\rangle$. One concludes that in quantum mechanics a conserved property is characterized by the condition $[\hat{f},\hat{H}]=0$ and $\partial\hat{f}/\partial t = 0$. Let us recall that in classical mechanics conserved quantities are characterized by having a vanishing Poisson bracket

$$\{f,H\} = \sum_i \left(\frac{\partial f}{\partial q_i}\frac{\partial H}{\partial p_i} - \frac{\partial f}{\partial p_i}\frac{\partial H}{\partial q_i}\right) = 0.$$

4.4. Degenerate energy levels

One often finds that there are various different eigenstates that correspond to the same energy. In these cases one says that the corresponding energy level is *degenerate.* Physically this means that the energy alone does not form a complete set of observables.

An example of a system having a degenerate energy spectrum is the free particle whose energy is given by $E = p^2/2m$. Except for $\vec{p} = 0$ there are infinitely many eigenstates (two eigenstates in 1D) having the same energy.

It is interesting to see that if $[\hat{f}, \hat{H}] = 0$ and $|\, n \,\rangle$ is an eigenstate of $\hat{H}$ with energy E_n, then $|\, n' \,\rangle = \hat{f}\, |\, n \,\rangle$ is also a stationary state with the same energy:

$$\hat{H}\, |\, n' \,\rangle = \hat{H}\, \hat{f}\, |\, n \,\rangle = \hat{f}\, \hat{H}\, |\, n \,\rangle = E_n\, \hat{f}\, |\, n \,\rangle = E_n\, |\, n' \,\rangle.$$

In fact, $\hat{f}$ can be time-dependent ($\partial \hat{f}/\partial t \neq 0$) provided it commutes with $\hat{H}$ at all times. Notice that $|\, n \,\rangle$ need not be an eigenstate of $\hat{f}$. Consequently, as soon as $|\, n \,\rangle$ is not an eigenstate of $\hat{f}$, there are degenerate levels. This situation is found when there are two operators $\hat{f}$ and $\hat{g}$ that commute with $\hat{H}$, but not between themselves. Consider

$$[\hat{f}, \hat{H}] = 0 \quad \text{and} \quad [\hat{g}, \hat{H}] = 0, \quad \text{but} \quad [\hat{f}, \hat{g}] \neq 0$$

and let $\{|\, n \,\rangle$ with $n = 1, 2, \ldots\}$ be a set of eigenstates of $\hat{f}$ and $\hat{H}$ with energy E_n. Then there is at least one state $|\, n \,\rangle$ for which

$$\hat{g}\, |\, n \,\rangle \neq g\, |\, n \,\rangle,$$

i.e., $|\, n \,\rangle$ is not an eigenstate of $\hat{g}$, since otherwise $\hat{f}$ and $\hat{g}$ would commute. Taking into account that $\hat{g}\, |\, n \,\rangle$ has the same energy E_n as $|\, n \,\rangle$ and that $\hat{g}\, |\, n \,\rangle \neq |\, n \,\rangle$ it is clear that the energy level E_n is degenerate. An important example of non-commuting conserved quantities are the components $\hat{L}_x$, $\hat{L}_y$ and $\hat{L}_z$ of the angular momentum in a central symmetric field.

Exercise: Consider a system composed of two non-interacting parts whose Hamiltonian can be written as $\hat{H} = \hat{H}_1(q_1) + \hat{H}_2(q_2)$. Let $\Psi_1(q_1)$ and $\Psi_2(q_2)$ denote the eigenstates of $\hat{H}_1$ and $\hat{H}_2$ respectively (i.e., $\hat{H}_1\, \Psi_1 = E_1\, \Psi_1$ and $\hat{H}_2\, \Psi_2 = E_2\, \Psi_2$). Show that the composite state $\Psi(q_1, q_2) = \Psi_1(q_1)\, \Psi_2(q_2)$ is an eigenstate of $\hat{H}$ with energy $E = E_1 + E_2$.

4.5. Finite vs. infinite motion

As we shall see the stationary states can have a discrete or continuous energy spectrum. One can show that stationary states of the discrete spectrum necessarily correspond to a finite motion, which means that neither the system nor any part of it moves to infinity at any time. This can be easily understood since the eigenfunctions of the discrete spectrum are square integrable, i.e.,

$$\int |\Psi_n|^2\, dq = 1.$$

Therefore, $|\Psi_n|^2 \to 0$ for $q \to \infty$ faster than q^{-D}, where D refers to the dimensions of the coordinate space.

In contrast the stationary states of a continuous energy spectrum have non-square-integrable wave functions satisfying the orthonormality condition

$$\int \Psi^*_E(q)\, \Psi_{E'}(q)\, dq = \delta(E - E').$$

In particular the norm of $\Psi_E(q)$ diverges. The stationary states of a continuum spectrum correspond therefore to an infinite motion. Even if at $t = 0$ the wave function $\Psi(q, t_0)$ is localized, one can show that the probability of finding the particle at $q \to \infty$ becomes finite for sufficiently long t. In fact the probability $\int_V |\Psi_E(q, t)|^2\, dq$ of finding the particle in any finite volume V vanishes for sufficiently long time t.

4.6. The Schrödinger equation

We know that the momentum $\vec{p}$ of a particle defines its state completely, very much like the position $\vec{r}$ does. Moreover, for a free particle, $\vec{p}$ is conserved and well defined. Therefore $\vec{p}$ and the energy can be measured simultaneously. The Hamiltonian of a free particle can thus be expressed in terms of the operators $\hat{p}_x$, $\hat{p}_y$ and $\hat{p}_z$, since one can find a complete set of eigenstates with defined values of $\vec{p}$ and E. The isotropy of space requires that $\hat{H}$ can only depend on $\hat{p}^2 = \hat{p}_x^2 + \hat{p}_y^2 + \hat{p}_z^2$. Moreover, the homogenity of space and time requires that for a free particle $\hat{H}$ cannot depend on t or $\hat{\vec{r}} = (\hat{x}, \hat{y}, \hat{z})$. The precise form of $\hat{H}$ must satisfy the correspondence principle, i.e., the limit of classical mechanics. Therefore for a free particle we have

$$\hat{H} = \frac{\hat{p}^2}{2m}, \tag{4.3}$$

where m is the particle's mass. It should be noted that Eq. (4.3) can be derived without any reference to classical mechanics by requiring simply the Galileo invariance of the equations of motion for $\Psi(q, t)$.

In the presence of an external potential $U(x, y, z)$ the Hamiltonian is given as in classical mechanics by

$$\begin{aligned} \hat{H} &= \frac{\hat{p}^2}{2m} + U(\hat{x}, \hat{y}, \hat{z}) \\ &= -\frac{\hbar^2}{2m}\, \nabla^2 + U(x, y, z). \end{aligned}$$

The stationary states are therefore given by the time-independent *Schrödinger equation*

$$\hat{H}\, \Psi = E\, \Psi$$

which takes the form

$$-\frac{\hbar^2}{2m}\, \nabla^2\, \Psi + U(x, y, z)\, \Psi = E\, \Psi.$$

For a *free particle* the equation reads

$$-\frac{\hbar^2}{2m}\nabla^2\,\Psi = E\,\Psi.$$

This is solved by the eigenstates

$$\Phi_{\vec{p}}(\vec{r}) = \frac{1}{(2\pi\hbar)^{3/2}}\;e^{\frac{i}{\hbar}\,\vec{p}\cdot\vec{r}}$$

of the momentum operator $\hat{\vec{p}}$, where $\vec{p} = (p_x, p_y, p_z)$ is an arbitrary vector in $\mathbb{R}^3$. Here we have chosen the normalization condition $\langle\,\vec{p}\,|\,\vec{p}\,'\,\rangle = \delta(\vec{p}-\vec{p}\,')$. The corresponding energy is

$$E = \frac{p^2}{2m} = \frac{p_x^2 + p_y^2 + p_z^2}{2m}.$$

Therefore the energy spectrum $E \geq 0$ is continuous, unbound and infinitely degenerate except for $E = 0$ (twofold degenerate in 1D). The complete time-dependent stationary states are given by

$$\Phi_{\vec{p}}(\vec{r},t) = \frac{1}{(2\pi\hbar)^{3/2}}\;e^{-\frac{i}{\hbar}\,E\,t + \frac{i}{\hbar}\,\vec{p}\cdot\vec{r}}.$$

4.7. The variational principle

Mathematically it is always possible to reformulate a differential equation in a variational form by finding an appropriate functional[3] such that its minimization with respect to the function yields the original differential equation. In the case of the Schrödinger equation the variational principle takes a very simple and physically transparent form. One can in fact show that the stationary states are the extremes of the expectation value of the energy subject to the normalization constraint. Let

$$E[\Psi] = \langle\,\Psi\,|\,\hat{H}\,|\,\Psi\,\rangle = \int \Psi^*(q)\,\hat{H}\,\Psi(q)\,dq \tag{4.4}$$

be the average energy in the wave function $\Psi(q)$. $E[\Psi]$ is a functional of $\Psi(q)$. The norm of Ψ, given by

$$\langle\,\Psi\,|\,\Psi\,\rangle = \int \Psi^*(q)\,\Psi(q)\,dq, \tag{4.5}$$

is also a functional of Ψ. Actually the precise value of the norm $\int |\Psi|^2\,dq$ is not important, *provided that it remains constant* when Ψ is varied. Therefore, the following also applies to states of the continuum spectrum. To be explicit we require $\langle\,\Psi\,|\,\Psi\,\rangle = 1$. The extrema

---

[3] A functional is an application that assigns to each function a number in $\mathbb{R}$ or $\mathbb{C}$.

of Eq. (4.4) subject to Eq. (4.5) are obtained by introducing a Lagrange multiplier E, so that we actually seek for the extrema of the auxiliary functional

$$F[\Psi] = \int \Psi^*(q)\, \hat{H}\, \Psi(q)\, dq - E \left(\int \Psi^*(q)\, \Psi(q)\, dq - 1 \right). \tag{4.6}$$

Since $\Psi(q) \in \mathbb{C}$ we may consider the variations $\delta\Psi(q)$ and $\delta\Psi^*(q)$ as independent variables, instead of taking the variations of the real and imaginary parts. At the extrema the variation δF of F with respect to $\Psi^*(q)$ and $\Psi(q)$ must vanish. Varying Ψ^* keeping Ψ constant, we have

$$\delta F = \int \delta\Psi^*(q) \left[\hat{H}\, \Psi(q) - E\, \Psi(q) \right] dq.$$

Since $\delta F = 0$ for any function $\delta\Psi^*(q)$ we must have

$$\hat{H}\, \Psi(q) = E\, \Psi(q) \tag{4.7}$$

at all q, which is the Schrödinger equation.[4] Taking the variations of F given by Eq. (4.6) with respect to $\delta\Psi(q)$ keeping $\Psi^*(q)$ constant yields

$$\delta F = \int \left[\Psi^*(q)\, \hat{H} - E\, \Psi^*(q) \right] \delta\Psi(q)\, dq = \int \left\{ [H\, \Psi(q)]^* - E\, \Psi^*(q) \right\} \delta\Psi(q)\, dq.$$

Requiring $\delta F = 0$ for all $\delta\Psi(q)$ also implies Eq. (4.7).

The absolute minimum of $E[\Psi]$ subject to the normalization condition $\langle\, \Psi \,|\, \Psi \,\rangle = 1$ is the lowest eigenvalue of $\hat{H}$, i.e., the ground state energy E_0. We may therefore write

$$E_0 = \min_{\Psi} \frac{\langle\, \Psi \,|\, \hat{H} \,|\, \Psi \,\rangle}{\langle\, \Psi \,|\, \Psi \,\rangle}. \tag{4.8}$$

This result can be obtained more directly by computing $E[\Psi]$ in the form

$$E[\Psi] = \sum_n |a_n|^2\, E_n, \tag{4.9}$$

where $|\, \Psi \,\rangle = \sum_n a_n \,|\, n \,\rangle$ is expanded in the stationary states $|\, n \,\rangle$ of $\hat{H}$ ($H \,|\, n \,\rangle = E_n \,|\, n \,\rangle$). Here we assume for simplicity that $\sum_n |a_n|^2 = 1$. From Eq. (4.9) it is clear that the absolute minimum of $E[\Psi]$ is the lowest eigenvalue of $\hat{H}$, i.e., the ground-state energy E_0.

Eq. (4.9) also gives a useful hint for obtaining the excited states from a variational procedure. Suppose that one has determined the ground state $|\, \Psi_0 \,\rangle$ or that one knows its symmetry. One can then minimize $E[\Psi] = \langle\, \Psi \,|\, \hat{H} \,|\, \Psi \,\rangle$ by requiring not only $\langle\, \Psi \,|\, \Psi \,\rangle = 1$ but also $\langle\, \Psi_0 \,|\, \Psi \,\rangle = 0$. This means minimizing $E[\Psi]$ keeping $a_0 = 0$ and $\sum_n |a_n|^2 = 1$, which corresponds to $a_1 = 1$ and $a_m = 0 \;\; \forall\, m \neq 1$. In this way the first excited state Ψ_1

---

[4] Here we have used the fundamental theorem of the calculus of variations.

and its energy E_1 are obtained. The procedure can be iterated in principle at will. The n-th excited state is the minimum of $E[\Psi] = \langle \Psi \,|\, \hat{H} \,|\, \Psi \rangle$ with the additional conditions

$$\langle \Psi \,|\, \Psi \rangle = 1 \quad \text{and} \quad \langle \Psi_m \,|\, \Psi \rangle = 0 \quad \forall \, m < n.$$

Using the variational principle one can show that the ground-state wave function for a single-particle system has no nodes. Consequently the excited states necessarily have nodes, since otherwise $\int \Psi_n^* \, \Psi_0 \; dq \neq 0$. Moreover, this implies that the ground state is *nondegenerate*, since otherwise nodes could be introduced by forming linear combinations of two or more states having the ground-state energy.

If the motion occurs in a restricted region, Ψ_0 vanishes at the boundaries and has no nodes inside. If the size of the restricted region increases (e.g., if the size of a nanoparticle increases), the energy of all levels decreases since the number of accessible functions for the minimization of $E[\Psi]$ increases. We recover Heisenberg's uncertainty principle implying lower Δp^2 and $\langle \frac{p^2}{2m} \rangle$ for larger Δx^2.

4.8. Properties of the Schrödinger equation and stationary wave functions

The single-particle Schrödinger equation reads

$$\hat{H} \, \Psi = E \, \Psi$$

with

$$\hat{H} = \frac{\hat{p}^2}{2m} + U(x,y,z) = -\frac{\hbar^2}{2m} \nabla^2 + U(x,y,z)$$

and can be written as

$$-\frac{\hbar^2}{2m} \nabla^2 \, \Psi + (U(x,y,z) - E) \, \Psi = 0. \tag{4.10}$$

For simplicity we often focus on the one-dimensional case for which Eq. (4.10) reads

$$-\frac{\hbar^2}{2m} \frac{d^2\Psi}{dx^2} + [U(x) - E] \, \Psi = 0. \tag{4.11}$$

Local continuity and boundary conditions on $\Psi(\vec{r})$

In order to identify the conditions that physical solutions of Eqs. (4.10) and (4.11) must satisfy, several cases should be considered:

i) If $U(\vec{r})$ is continuous at $\vec{r}$, then $\Psi(\vec{r})$ has continuous second-order (partial) derivatives at this point. Obviously, the first-order derivates $\vec{\nabla}\Psi$ or $\frac{d\Psi}{dx}$ and the wave function $\Psi(\vec{r})$ are also continuous at $\vec{r}$.

ii) If $\vec{r}$ belongs to a surface of finite discontinuities of $U(\vec{r})$, or if $U(x)$ is discontinuous at x in 1D, the second-order derivative $\dfrac{d^2\Psi}{dx^2} = \dfrac{2m}{\hbar^2}\,[U(x) - E]\Psi$ also shows a finite discontinuity at this point. However, the integral of $\dfrac{d^2\Psi}{dx^2}$, i.e., $\dfrac{d\Psi}{dx}$, remains continuous at this point:

$$\frac{d\Psi}{dx}\,(x+\eta) - \frac{d\Psi}{dx}\,(x) = \int_x^{x+\eta} \frac{d^2\Psi}{dx^2}\,(x')\,dx' \xrightarrow{\eta\to 0} 0,$$

where we have used that $\dfrac{d^2\Psi}{dx^2} = \dfrac{2m}{\hbar^2}\,(U - E)\Psi$ is everywhere finite. Since $\dfrac{d\Psi}{dx}$ is continuous the wave function itself is also continuous even in the presence of finite discontinuities of the potential $U(\vec{r})$.

iii) Finally, we consider a potential $U(\vec{r})$ that becomes infinitely repulsive in a region of finite measure (i.e., a finite volume in 3D or a segment of finite length in 1D, not an isolated point). If $U(\vec{r}) \to \infty$ in a finite (non-infinitesimal) region the wave function must strictly vanish throughout this region, since otherwise the energy E would diverge: $\Psi(\vec{r}) \to 0$ for $U(\vec{r}) \to +\infty$. At the boundary of the domain where $U(\vec{r}) \to +\infty$, $\Psi(x)$ is continuous, and thus vanishes, but $\dfrac{d\Psi}{dx}$ has a finite discontinuity. To show this we can write

$$\frac{d\Psi}{dx}\,(x_0+\eta) - \frac{d\Psi}{dx}\,(x_0-\eta) = \int_{x_0-\eta}^{x_0+\eta} \frac{2m}{\hbar^2}\,(U-E)\,\Psi\,dx' = \frac{2m}{\hbar^2}\,\langle\,(U-E)\,\Psi\,\rangle\,2\eta,$$

where x_0 refers to a point at the boundary of the infinitely repulsive domain and $\langle\ldots\rangle$ indicates the average in the interval $[x_0-\eta,\, x_0+\eta]$. We may now take the limit $U \to +\infty$ keeping $\langle\,(U-E)\,\Psi\,\rangle\,\eta$ constant. Therefore the discontinuity of $\dfrac{d\Psi}{dx}$ remains finite even at points where the discontinuity in $U(\vec{r})$ diverges. A finite discontinuity in $\dfrac{d\Psi}{dx}$ implies of course that $\Psi(x)$ is continuous.

The previous arguments also apply to the limit of arbitrarily small width of the divergent region. If the potential has a divergency only at one point and the divergency is integrable of the form $U(x) = \alpha\,\delta(x)$, then $\Psi(x)$ remains finite and continuous but $\dfrac{d\Psi}{dx}$ is discontinuous. To illustrate the origin of the discontinuity of $\dfrac{d\Psi}{dx}$ we can write

$$\begin{aligned}\frac{d\Psi}{dx}\,(\eta) - \frac{d\Psi}{dx}\,(-\eta) &= \int_{-\eta}^{\eta} \frac{d^2\Psi}{dx^2}\,(x')\,dx' \\ &= \frac{2m}{\hbar^2}\int_{-\eta}^{\eta} \big(U(x') - E\big)\,\Psi(x')\,dx' = \frac{2m}{\hbar^2}\int_{-\eta}^{\eta} \big[\alpha\,\delta(x') - E\big]\,\Psi(x')\,dx'\end{aligned}$$

$$= \frac{2m}{\hbar^2} \left(\alpha\, \Psi(0) - 2\eta E \langle \Psi \rangle\right) \xrightarrow{\eta \to 0} \frac{2m}{\hbar^2}\, \alpha\, \Psi(0).$$

Since the discontinuity of $\frac{d\Psi}{dx}$ at $x = 0$ is finite, $\Psi(x)$ remains continuous even in the presence of locally divergent potentials.

Eigenenergy lower bound

Concerning the eigenenergies E it is easy to see that they are necessarily larger than the minimum value $U_{\min}$ of the potential. In fact $U(\vec{r}) \geq U_{\min} \;\; \forall\, \vec{r}$ implies $\langle\, U \,\rangle \geq U_{\min}$. Moreover, the kinetic energy $T = \langle \frac{\hat{p}^2}{2m} \rangle$ is always positive. Consequently,

$$E = \langle\, \Psi \,|\, \hat{T} + \hat{U} \,|\, \Psi \,\rangle \geq \langle\, \Psi \,|\, \hat{U} \,|\, \Psi \,\rangle \geq U_{\min}.$$

The nature of the energy spectrum, either discrete or continuous, is conditioned by the long-distance behavior of the potential energy. Let us consider a potential $U(\vec{r})$ that vanishes at infinite or, more precisely, such that $\lim_{r\to+\infty} U(r, \theta, \varphi) = 0 \quad \forall\; \theta, \varphi$. In this case all possible negative-energy eigenstates have a wave function that vanishes for $r \to \infty$. The motion is thus localized and the spectrum is discrete. Conversely, if the energy is positive, the motion is not bounded and the spectrum is continuous. In the case where $U > 0\ \forall\ \vec{r}$ and $U(\vec{r}) \xrightarrow{r\to\infty} 0$ there is no discrete spectrum, since $E \geq U_{\min} = 0$ and $E > 0$ implies unbounded motion.

Let us finally recall that the time-independent Schrödinger equation in the absence of magnetic fields

$$-\frac{\hbar^2}{2m}\, \nabla^2\, \Psi + U\Psi = E\, \Psi$$

is real. The same holds for the above-mentioned continuity boundary conditions on $\Psi(\vec{r})$ and $\vec{\nabla}\Psi$. Therefore, in the absence of magnetic field the time-independent stationary states $\Psi_n(\vec{r})$ can be chosen to be real. If the level is nondegenerate $\Psi_n(x)$ is necessarily real. In any case $\Psi^*(x)$ is always an eigenstate of energy E, whenever $\Psi(x)$ is an eigenstate with energy E. Consequently, it is always possible to choose $\Psi(x)$ to be real by considering $\Psi + \Psi^*$ and $i\,(\Psi - \Psi^*)$.

4.9. Motion in one dimension: The square potential well

If $U(\vec{r})$ depends only on one coordinate or, more generally, if $U(\vec{r}) = U_1(x) + U_2(y) + U_3(z)$, the Hamiltonian can be splitted in the sum of three terms each of which involves the variables x, y and z separately. The stationary states are then given by the product $\Psi(\vec{r}) = \Psi_1(x)\, \Psi_2(y)\, \Psi_3(z)$, where Ψ_i are the solutions of one-dimensional equations of the form

$$-\frac{\hbar^2}{2m}\, \frac{d^2\Psi}{dx^2} + U(x)\Psi = E\, \Psi.$$

In order to illustrate some of the properties of $\Psi(x)$ we consider a square potential well given by

$$U(x) = \begin{cases} U_0 > 0 & \text{for } |x| > a/2 \\ 0 & \text{for } |x| \le a/2. \end{cases}$$

In the region I, where $|x| < a/2$, the Schrödinger equation reads

$$-\frac{\hbar^2}{2m}\frac{d^2\Psi}{dx^2} = E\,\Psi$$

with $E > 0$. The solutions have therefore the form

$$\Psi(x) = A\,e^{i\,k\,x} + B\,e^{-i\,k\,x}$$

with

$$E = \frac{\hbar^2 k^2}{2m}$$

and $k > 0$ ($k \in \mathbb{R}$). For x outside the well, in regions II ($x < -a/2$) and III ($x > a/2$), we have

$$-\frac{\hbar^2}{2m}\frac{d^2\Psi}{dx^2} = (E - U_0)\Psi \quad \text{with} \quad (E - U_0) < 0.$$

The solutions for $E < U_0$ have the form

$$\Psi_\pm(x) = e^{\pm\kappa\,x}$$

with

$$U_0 - E = \frac{\hbar^2\kappa^2}{2m} \quad \text{or} \quad \kappa = \sqrt{\frac{2m}{\hbar^2}\,(U_0 - E)} > 0.$$

The spectrum for $E < U_0$ is discrete and therefore the eigenfunctions must be integrable. Thus we have

$$\Psi(x) = C\,e^{\kappa\,x} \quad \text{for} \quad x < -a/2 \quad \text{(Region I)}$$

and

$$\Psi(x) = D\,e^{-\kappa\,x} \quad \text{for} \quad x > a/2 \quad \text{(Region III)},$$

where C and D are constants ($\kappa > 0$).

Notice that the particle has a finite probability to be found outside the well (U_0 finite). However, the penetration length $1/\kappa$ vanishes for $U_0 \to +\infty$, as already discussed in the previous section. The discrete eigenvalues $E < U_0$ and the constants A, B, C and D are determined by imposing the continuity of $\Psi(x)$ and $\Psi'(x)$ at the well boundaries $x = \pm\,a/2$. This is discussed in some detail below.

For $E > U_0$ the motion is infinite and the spectrum is therefore continuum. In the regions I and III the solutions have the general plane-wave form

$$\Psi(x) = A'\, e^{i\,k'\,x} + B'\, e^{-i\,k'\,x}$$

with $k = \sqrt{\frac{2m}{\hbar^2}\,(E - U_0)} \geq 0 \quad (|x| > a/2)$. However, notice that the constants A' and B' may be different in regions I and III.

As a first example of a system having a completely discrete energy spectrum we consider the case $U_0 \to +\infty$, for which the motion is confined to region II. $\Psi(x)$ vanishes in regions I and III and the continuity of $\Psi(x)$ imposes the constraints $\Psi(\pm\, a/2) = 0$ at the well boundaries. We thus have

$$\Psi(a/2) = A\, e^{i\,k\,a/2} + B\, e^{-i\,k\,a/2} = 0 \tag{4.12}$$

and

$$\Psi(-a/2) = A\, e^{-i\,k\,a/2} + B\, e^{i\,k\,a/2} = 0. \tag{4.13}$$

Multiplying Eq. (4.12) by $e^{-i\,k\,a/2}$ ($e^{i\,k\,a/2}$) and Eq. (4.13) by $e^{i\,k\,a/2}$ ($e^{-i\,k\,a/2}$) and subtracting one obtains

$$B\left(e^{-i\,k\,a} - e^{i\,k\,a}\right) = 0 \quad \Rightarrow \quad B\,\sin k\,a = 0$$

and

$$A\left(e^{i\,k\,a} - e^{-i\,k\,a}\right) = 0 \quad \Rightarrow \quad A\,\sin k\,a = 0.$$

Since A and B cannot be simultaneously zero, we have $\sin k\,a = 0$, which implies $k = \frac{n\,\pi}{a}$ with $n = 1, 2, \ldots$. Note that $n = 0$ is excluded since it yields a constant $\Psi(x) \equiv 0$. The corresponding eigenenergies take thus the values

$$E_n = \frac{\pi^2\,\hbar^2}{2m\,a^2}\,n^2 \quad \text{with} \quad n = 1, 2, \ldots\ .$$

The condition $\Psi(-a/2) = 0$ implies

$$A + B\,e^{i\,k\,a} = A + B\,e^{i\,n\,\pi} = A + (-1)^n\,B = 0,$$

which is equivalent to the relation

$$A + B\,e^{-i\,k\,a} = A + B\,e^{-i\,n\,\pi} = A + (-1)^n\,B = 0$$

derived from the constraint $\Psi(a/2) = 0$. Consequently, $A = B = \sqrt{\frac{2}{a}}$ and

$$\Psi_n(x) = \sqrt{\frac{2}{a}}\,\cos\left(\frac{n\,\pi\,x}{a}\right) \quad \text{for } n \text{ odd.}$$

Similarly $A = -B = \sqrt{\frac{2}{a}}$ and

$$\Psi_n(x) = \sqrt{\frac{2}{a}}\,\sin\left(\frac{n\,\pi\,x}{a}\right) \quad \text{for } n \text{ even.}$$

Notice that $\Psi_n(x)$ is either even or odd $[\Psi_n(-x) = (-1)^{n+1}\,\Psi_n(x)]$ and that the ground state $\Psi_1(x)$ is even and nodeless. These are general properties of even potentials $[U(-x) = U(x)]$.

4.10. Summary of quantum dynamics

4.10.1. The time-dependent Schrödinger equation

The superposition principle implies that the dynamics (time evolution) of the wave function $\Psi(q,t)$ is given by a first-order linear differential equation of the general form

$$i\,\hbar\,\frac{\partial\Psi}{\partial t} = \hat{H}\,\Psi,$$

where $\hat{H}$ is the Hamilton operator or Hamiltonian of the system. For a single particle in an external potential $U(\vec{r})$ it is given by

$$\hat{H} = \frac{\hat{p}^2}{2m} + U(\vec{r}).$$

In general we know that the dynamics must preserve the normalization of the wave function, i.e.,

$$\int |\Psi(q,t)|^2\,dq \quad \text{is independent of } t.$$

This requires that $\hat{H}^\dagger = \hat{H}$ is an hermitic operator, which means just that for all $\Psi_1(\vec{r})$ and $\Psi_2(\vec{r})$

$$\int \Psi_1^*(q)\;[\hat{H}\,\Psi_2(q)]\,dq = \int [\hat{H}\,\Psi_1(q)]^*\;\Psi_2(q)\,dq.$$

4.10.2. Stationary states

Since $\hat{H}^\dagger = \hat{H}$ we can find a complete set of eigenstates $\Psi_n(q)$, i.e.,

$$\hat{H}\,\Psi_n(q) = E_n\,\Psi_n(q),$$

where $E_n \in \mathbb{R}$ are the eigenenergies. $\Psi_n(q)$ are independent of time since $\hat{H}$ does not depend explicitly on time. The eigenstates of $\hat{H}$ have a very simple time dependence:

$$\Psi_n(q,t) = \Psi_n(q)\;e^{-\frac{i}{\hbar}E_n(t-t_0)},$$

where we have set $\Psi_n(q,t_0) = \Psi_n(q)$. The eigenstates $\Psi_n(q)$ of $\hat{H}$ are called stationary states since for any observable $\hat{f}$ that does not depend explicitly on time we have that

$$|a_f|^2 = \Big| \int \Psi_f^*(q)\; \Psi_n(q,t)\, dq \Big|^2$$

is independent of time ($\hat{f}\,\Psi_f = f\,\Psi_f$).

The time dependence of an arbitrary state is obtained straightforwardly if the complete set of stationary states $\Psi_n(q)$ is known. One then expands

$$\Psi(q,t_0) = \sum_n a_n\, \Psi_n(q)$$

in terms of Ψ_n, and the superposition implies

$$\Psi(q,t) = \sum_n a_n\, \Psi_n(q)\; e^{-\frac{i}{\hbar} E_n (t-t_0)}$$

for all $t \geq t_0$.

4.10.3. The energy spectrum

The spectrum of eigenenergies can be discrete or continuous. A discrete spectrum, actually a stationary state Ψ_n belonging to a discrete energy spectrum, corresponds to a finite motion that is restricted in space, since the corresponding wave function is square-integrable:

$$\int dq\; |\Psi_n(q)|^2 = 1.$$

In contrast an energy eigenstate Ψ_E of the continuous spectrum corresponds to an infinite motion that is unbounded. These are known as scattering states. Even if localized states can be constructed with stationary states of the continuous spectrum, one can show that the particle always moves to infinity for large enough time. In this case

$$\int_V |\Psi_E(q,t)|^2\, dq \xrightarrow{t\to\infty} 0 \qquad \text{for any finite volume } V.$$

A given eigenvalue E_n of $\hat{H}$ can correspond to a single stationary state, to several states, or to infinitely many states. In the first case we say that the energy level is nondegenerate, otherwise it is said to be degenerate. In most cases the energy levels are degenerate, which means that the energy alone does not suffice to define the quantum state univocally. As examples one can mention the p, d, and f orbitals in atoms, or a free particle with $\vec{p} \neq 0$.

4.10.4. The variational principle

The stationary state with the lowest energy is known as the ground state, while the other states are called excited states. The ground state energy E_0 and the corresponding wave function can be derived from a variational procedure

$$E_0 = \min_{\Psi} \frac{\langle\, \Psi \,|\, \hat{H} \,|\, \Psi \,\rangle}{\langle\, \Psi \,|\, \Psi \,\rangle}.$$

The excited states $\Psi_n(q)$ can be derived by minimizing

$$E[\Psi] = \langle\, \Psi \,|\, \hat{H} \,|\, \Psi \,\rangle$$

under the constraints $\langle\, \Psi \,|\, \Psi \,\rangle = 1$ and $\langle\, \Psi \,|\, \Psi_m \,\rangle = 0 \;\; \forall\, m < n$. For the ground state Ψ_0 only the constraint $\langle\, \Psi_0 \,|\, \Psi_0 \,\rangle = 1$ applies.

4.10.5. Properties of the stationary Schrödinger equation

We consider the time-independent Schrödinger equation

$$\hat{H}\, \Psi = E\, \Psi \tag{4.14}$$

or, making the form of $\hat{H}$ explicit,

$$-\frac{\hbar^2}{2m}\, \nabla^2\, \Psi + U(x,y,z)\Psi = E\, \Psi \tag{4.15}$$

in 3D and

$$-\frac{\hbar^2}{2m}\, \frac{d^2\Psi}{dx^2} + U(x)\Psi = E\, \Psi \tag{4.16}$$

in 1D. Since the Schrödinger equation is a second-order differential equation, it is the second derivative of $\Psi(\vec{r})$ that shows discontinuities at the points where the potential is discontinuous. Consequently,

i) $\Psi(x)$ and $\overrightarrow{\nabla\Psi}$ are continuous provided that $U(\vec{r})$ has only finite discontinuities.

ii) If $U(\vec{r}) \to +\infty$ in a finite volume in space (or in a finite segment in 1D), then $\Psi(\vec{r}) = 0$ in this volume. Otherwise E would diverge.

iii) At the boundary of regions excluded to the particles $[\Psi(\vec{x}) = 0]$ the wave function remains continuous but $\overrightarrow{\nabla\Psi}$ is discontinuous.

If the potential has a lower bound $U_{\text{min}} = \min\{U(\vec{r}) \;\text{ for }\; \vec{r} \in \mathbb{R}^3\}$, the ground state energy E_0 satisfies $E_0 \geq U_{\text{min}}$. If $U \to -\infty$ for a finite number of points (e.g., at the nuclei of atoms), the ground state remains finite and the particle does not fall into the center of attraction provided that the potential does not decrease (diverge) very rapidly with distance, actually as long as

$$U(\vec{r}) \sim \frac{1}{r^\alpha} \quad \text{with} \quad \alpha < 2.$$

The reason is the positive kinetic energy contribution which diverges as Δr^{-2} when the wave function is localized in a region Δr.

In the absence of magnetic field $\hat{H}$ is real. Moreover, the boundary conditions on $\Psi(\vec{r})$ are also real. Therefore, if $\Psi(\vec{r})$ is a solution of Eq. (4.14), then $\Psi^*(\vec{r})$ is also a solution with the same eigenenergy E. Thus, the stationary state wave functions can be taken to be real [proportional to $\Psi + \Psi^*$ and $(\Psi - \Psi^*)/i$]. Nondegenerate states have necessarily real $\Psi(\vec{r})$. Otherwise one can search the energy eigenstates assuming a real $\Psi(\vec{r})$.

4.11. The square potential well

In this section we consider the problem of a one-dimensional potential of the form

$$U(x) = \begin{cases} U_0 > 0 & \text{for } \ x < 0 & \text{I} \\ 0 & \text{for } \ 0 < x < a & \text{II} \\ U_0 & \text{for } \ x > a & \text{III.} \end{cases}$$

The potential being piecewise constant, we determine the energy eigenfunctions by solving Eq. (4.16) in the three regions I–III and by matching the solutions at the boundaries.

In regions I and III the Schrödinger equation reads

$$-\frac{\hbar^2}{2m}\frac{d^2\Psi}{dx^2} + U_0\,\Psi = E\,\Psi \quad \Leftrightarrow \quad \frac{2m}{\hbar^2}(U_0 - E)\Psi = \frac{d^2\Psi}{dx^2}.$$

For the discrete spectrum of bound states having $0 \le E \le U_0$ we obtain

$$\Psi(x) = e^{\pm\kappa\, x}$$

with $\kappa^2 = \frac{2m}{\hbar^2}(U_0 - E) > 0 \ \Rightarrow \ \kappa = \frac{\sqrt{2m(U_0 - E)}}{\hbar^2}$. Since $\Psi(x) \to 0$ for $|x| \to +\infty$ (bound state), we must have

$$\Psi_1(x) = A_1\, e^{\kappa\, x} \qquad \text{for } x < 0 \quad \text{(region I)}$$

and

$$\Psi_3(x) = A_3\, e^{-\kappa\, x} \qquad \text{for } x > a \quad \text{(region III)}$$

with $\kappa > 0$ and A_1, $A_3 \in \mathbb{R}$. Let us recall that the discrete spectrum is non-degenerate and therefore the eigenfunctions are real.

In region II we have

$$\frac{d^2\Psi}{dx^2} = -\frac{2m\,E}{\hbar^2}\,\Psi$$

with $E > 0$. This implies

$$\Psi(x) = e^{\pm i\,k\,x}$$

with $k^2 = \dfrac{2m\,E}{\hbar^2} > 0 \;\Rightarrow\; k = \dfrac{\sqrt{2m\,E}}{\hbar}$. The most general solution in region II has the form

$$\Psi_2(x) = A\, e^{i\,k\,x} + B\, e^{-i\,k\,x}$$

with A and $B \in \mathbb{C}$ $(k > 0)$. Since we are looking for real solutions we consider

$$\Psi_2 \pm \Psi_2^* = (A \pm B^*)\, e^{i\,k\,x} + (B \pm A^*)\, e^{-i\,k\,x}.$$

Without loss of generality we can write

$$\Psi_2(x) = A_2\, \sin(k\,x + \varphi) \qquad \text{for } 0 < x < a$$

with $k > 0$, A_2 and $\varphi \in \mathbb{R}$.

The continuity constraints on $\Psi(x)$ and $\Psi' = \dfrac{d\Psi}{dx}$ can be replaced by the requirement that $\Psi(x)$ and $\Psi'(x)/\Psi(x)$ are continuous. Considering the logarithmic derivative Ψ'/Ψ is more practical since the continuity conditions define Ψ up to a multiplicative constant.

At $x = 0$ we must have $\Psi_1 = \Psi_2$, i.e.,

$$A_1 = A_2\, \sin\delta.$$

Noting that $\Psi_1'(x) = A_1\,\kappa\, e^{\kappa\,x}$ and $\Psi_2'(x) = A_2\,k\,\cos(k\,x + \delta)$, $\Psi_1'/\Psi_1 = \Psi_2'/\Psi_2$ implies

$$\begin{aligned}
\frac{\kappa\,A_1}{A_1} &= \frac{k\,A_2\,\cos\delta}{A_2\,\sin\delta} \\
k &= \tan\delta\,\kappa = \tan\delta\,\sqrt{\frac{2m}{\hbar^2}\,(U_0 - E)} \\
k &= \tan\delta\,\sqrt{\frac{2m}{\hbar^2}\left(U_0 - \frac{\hbar^2\,k^2}{2m}\right)} \\
k &= \tan\delta\,\sqrt{\frac{2m\,U_0}{\hbar^2} - k^2} \qquad (k > 0).
\end{aligned} \tag{4.17}$$

Note that $\tan\delta > 0$. At $x = a$ we must have

$$A_2\,\sin(k\,a + \delta) = A_3$$

and

$$\begin{aligned}
\frac{\Psi'}{\Psi} &= \frac{-\kappa\, e^{-\kappa\,a}}{e^{-\kappa\,a}} = \frac{k\,\cos(k\,a + \delta)}{\sin(k\,a + \delta)} \\
k &= -\tan(k\,a + \delta)\,\kappa = -\tan(k\,a + \delta)\,\sqrt{\frac{2m\,U_0}{\hbar^2} - k^2}.
\end{aligned} \tag{4.18}$$

Note that $\tan(k\,a+\delta) < 0$. We can obtain k and δ from Eqs. (4.17) and (4.18) as follows. From Eq. (4.17) we have

$$k^2 = \frac{sin^2\delta}{\cos^2\delta}\left(\frac{2m\,U_0}{\hbar^2} - k^2\right)$$

$$k^2(1-\sin^2\delta) = \sin^2\delta\left(\frac{2m\,U_0}{\hbar^2} - k^2\right)$$

$$k^2 = \sin^2\delta\,\frac{2m\,U_0}{\hbar^2}$$

$$\sin\delta = \pm\frac{\hbar\,k}{\sqrt{2m\,U_0}}.$$

In this way we obtain

$$\delta = \sin^{-1}\frac{\hbar\,k}{\sqrt{2m\,U_0}} + n\,\pi \tag{4.19}$$

with $k>0$, $\tan\delta > 0$ and $n \in \mathbb{Z}$. In the following we choose $0 \le \sin^{-1}x \le \pi/2$ to ensure that $\tan\delta > 0$. Analogously from Eq. (4.18) we have

$$\sin(k\,a+\delta) = \mp\frac{\hbar\,k}{\sqrt{2m\,U_0}} \tag{4.20}$$

with $k>0$, $\tan(k\,a+\delta) < 0$ and $n \in \mathbb{Z}$. Notice that the signs in Eqs. (4.19) and (4.20) are opposite since $\dfrac{\tan\delta}{\tan(k\,a+\delta)} = -1$. Thus we obtain

$$k\,a+\delta = -\sin^{-1}\frac{\hbar\,k}{\sqrt{2m\,U_0}} + n'\,\pi$$

with $0 \le \sin^{-1}x \le \pi/2$. The minus sign ensures that $\tan(k\,a+\delta) < 0$. Finally, k is obtained from the transcendental equation

$$n'\,\pi - \sin^{-1}\left(\frac{\hbar\,k}{\sqrt{2m\,U_0}}\right) = k\,a + \sin^{-1}\left(\frac{\hbar\,k}{\sqrt{2m\,U_0}}\right) + n\,\pi$$

with n and $n' \in \mathbb{Z}$. We can then write $m = n' - n$

$$k\,a = m\,\pi - 2\,\sin^{-1}\left(\frac{\hbar\,k}{\sqrt{2m\,U_0}}\right)$$

$$\frac{1}{2}\,(m\,\pi - k\,a) = \sin^{-1}\left(\frac{\hbar\,k}{\sqrt{2m\,U_0}}\right)$$

$$\sin\left[\frac{1}{2}\,(m\,\pi - k\,a)\right] = \frac{\hbar}{\sqrt{2m\,U_0}}\,k,$$

where $0 \le (m\,\pi - k\,a)/2 \le \pi/2$. It follows that

$$\sin\left(\frac{m\,\pi}{2}\right)\cos\left(\frac{k\,a}{2}\right) + \cos\left(\frac{m\,\pi}{2}\right)\sin\left(\frac{k\,a}{2}\right) = \frac{\hbar}{\sqrt{2m\,U_0}}\,k \tag{4.21}$$

and

$$0 \le \frac{m\,\pi}{2} - \frac{k\,a}{2} \le \frac{\pi}{2} \qquad \Leftrightarrow \qquad \frac{(m-1)\,\pi}{2} \le \frac{k\,a}{2} \le \frac{m\,\pi}{2}$$

$$\Leftrightarrow \qquad \begin{cases} \tan\left(\frac{k\,a}{2}\right) > 0 & \text{for } \ m \text{ odd} \\ \tan\left(\frac{k\,a}{2}\right) < 0 & \text{for } \ m \text{ even.} \end{cases}$$

For m odd we have $\sin(m\,\pi/2) = (-1)^{(m-1)/2}$. Thus, Eq. (4.21) implies

$$(-1)^{(m-1)/2}\ \cos\left(\frac{k\,a}{2}\right) = \frac{\hbar\,\sqrt{2}}{a\,\sqrt{m\,U_0}}\,\frac{k\,a}{2} \qquad \text{and} \qquad \tan\left(\frac{k\,a}{2}\right) > 0. \tag{4.22}$$

For m even we have $\cos(m\,\pi/2) = (-1)^{m/2}$. Thus, Eq. (4.21) implies

$$(-1)^{m/2}\ \sin\left(\frac{k\,a}{2}\right) = \frac{\hbar}{a}\,\frac{\sqrt{2}}{\sqrt{m\,U_0}}\,\frac{k\,a}{2} \qquad \text{and} \qquad \tan\left(\frac{k\,a}{2}\right) < 0. \tag{4.23}$$

Once k is known from Eq. (4.22) or (4.23) the eigenenergies are given by

$$E_n = \frac{\hbar^2\,k^2}{2m} = \frac{2\,\hbar^2}{m\,a^2}\left(\frac{k\,a}{2}\right)^2.$$

4.12. General properties of one-dimensional systems

A system is said to be one-dimensional when it is governed by the one-dimensional Schrödinger equation

$$-\frac{\hbar^2}{2m}\,\frac{d^2\Psi(x)}{dx^2} + U(x)\,\Psi(x) = E\,\Psi(x). \tag{4.24}$$

This is the case when $U(\vec{r})$ depends only on x, but also when $U(\vec{r}) = U_1(x) + U_2(y) + U_3(z)$ so that the eigenfunctions can be written in the form $\Psi(\vec{r}) = \phi_1(x)\,\phi_2(y)\,\phi_3(z)$. Some general properties of the motion in one dimension are summarized below.

Degeneracy of the energy spectrum: The energy levels of the discrete spectrum, i.e., when the motion is finite, are nondegenerate. The same holds when the particle cannot move in one direction, i.e., when the spectrum is continuous but $\Psi(x) \xrightarrow{x\to+\infty} 0$ or $\Psi(x) \xrightarrow{x\to-\infty} 0$.

To prove this assertion consider that there were two solutions Ψ_1 and Ψ_2 of Eq. (4.24). In this case we would have

$$\frac{\Psi_1''}{\Psi_1} = \frac{2m\,(U-E)}{\hbar} = \frac{\Psi_2''}{\Psi_2} \qquad \forall\ x,$$

which implies

$$\Psi_1''\,\Psi_2 - \Psi_2''\,\Psi_1 = 0 \quad \text{ or } \quad \Psi_1'\,\Psi_2 - \Psi_2'\,\Psi_1 = \text{ constant}.$$

Taking the limit of $x \to \pm\infty$ in the direction for which $\Psi_{1,2}(x) \to 0$, we find that the constant is zero:

$$\Psi_1'\,\Psi_2 - \Psi_2'\,\Psi_1 = 0.$$

Integrating $\dfrac{\Psi_1'}{\Psi_1} = \dfrac{\Psi_2'}{\Psi_2}$ we obtain

$$\ln\Psi_1 = \ln\Psi_2 + \text{ constant} \quad \forall\, x$$

or

$$\Psi_1 = c\,\Psi_2 \quad \forall\, x.$$

The two functions can only differ in an irrelevant constant (phase) factor.

Nodes of the wave function: The eigenfunction of the discrete spectrum corresponding to the n-th eigenvalue has n nodes, i.e., it vanishes n times. The ground state is nodeless and the n-th excited state has n nodes. This is known as oscillation theorem. If the motion is restricted to a finite region (e.g., $|\Psi(x)| = 0$ for $|x| > a$ due to $U(x) = +\infty$ for $|x| > a$), then the oscillation theorem applies to the finite region where $\Psi(x) \neq 0$.

Symmetric potentials: Let us consider a potential that is even, i.e., $U(x) = U(-x)$, and let $\Psi(x)$ be a stationary state of energy E. Taking into account that $\dfrac{d^2\Psi(-x)}{dx^2} = \dfrac{d^2\Psi(x)}{dx^2}$ it is clear that the function $\Psi(-x)$ is also an eigenstate with the same energy E. In fact, starting from Eq. (4.24) and replacing x by $-x$ we have

$$\frac{d^2\Psi(-x)}{dx^2} + U(x)\,\Psi(-x) = E\,\Psi(-x),$$

where we have used that $U(x) = U(-x)$. There are then two possibilities. If the energy level is nondegenerate (e.g., discrete spectrum), then $\Psi(x)$ and $\Psi(-x)$ can only differ by a constant:

$$\Psi(x) = c\,\Psi(-x).$$

Repeating the change of sign again we obtain

$$\Psi(x) = c\,\Psi(-x) = c^2\,\Psi(x),$$

which implies $c = \pm 1$. Consequently,

$$\Psi(x) = \pm\Psi(-x)$$

is either even or odd. In this case one says that Ψ has defined parity. Taking into account that the n-th eigenstate has n nodes we conclude that the ground state Ψ_0 is necessarily even, that the first excited state is odd, and that in general the n-th eigenstate Ψ_n has the same parity as n.

If the stationary is doubly degenerate, the parity of the wave function is not univocally defined by its energy. However we can always construct the states

$$\Psi_+(x) = \Psi(x) + \Psi(-x) \tag{4.25}$$

and

$$\Psi_-(x) = \Psi(x) - \Psi(-x) \tag{4.26}$$

which are, respectively, even and odd. For instance, if we have

$$\Psi(x) = e^{i\,k\,x},$$

the corresponding even and odd states are

$$\Psi_+ = e^{i\,k\,x} + e^{-i\,k\,x} = 2\,\cos kx$$

and

$$\Psi_- = e^{i\,k\,x} - e^{-i\,k\,x} = 2\,i\,\sin kx.$$

Notice, however, that Ψ_+ and Ψ_- as given by Eqs. (4.25) and (4.26) are in general not normalized and can eventually be zero, if the original function $\Psi(x)$ has a defined parity. In conclusion, for even potentials $U(x) = U(-x)$ one can always assume that the stationary states have defined parity, i.e., that they are even or odd.

Scattering states: We consider now a potential $U(x)$ which tends to finite limiting values for large distances. Let $U(x) \xrightarrow{x\to-\infty} 0$ and $U(x) \xrightarrow{x\to+\infty} U_0 > 0$. Since the discrete spectrum corresponds to a finite motion we can only find discrete energy levels for $U_{\min} \leq E < 0$. In this case the wave function decreases exponentially in both directions as soon as $E < U(x)$.

For $0 < E < U_0$ the motion is infinite in the direction $x \to -\infty$ and therefore the spectrum is continuous. As already discussed the eigenvalues are nondegenerate since $\Psi(x) \to 0$ for $x \to +\infty$ $(E < U_0)$. For $x \to -\infty$, $U(x) \to 0$ and we can neglect the potential in order to infer the limiting value of the wave function. For $x \to -\infty$ we therefore have a stationary plane way of the form

$$\Psi(x) = A\,\sin(k\,x + \delta),$$

where $k = \sqrt{2m\,E}/\hbar$. In the other limit $(x \to +\infty)$ we can approximate $U(x) \simeq U_0$ and the wave function takes the limiting form

$$\Psi(x) = B\,e^{-\kappa\,x},$$

where $\kappa = \sqrt{2m\,(U_0 - E)}/\hbar$. This follows from $\dfrac{d^2\Psi(x)}{dx^2} \cong \dfrac{2m}{\hbar}\,(U_0 - E)$ for $x \to +\infty$. As expected $\Psi(x)$ decreases exponentially in the classically forbidden region where $E < U(x)$.

Finally for $E > U_0$ the spectrum is continuous and doubly degenerate. In this case both solutions of the second order differential equation are physically acceptable (no divergent behavior). The asymptotic forms of the stationary state are

$$\Psi(x) = A_1\, e^{i\, k_1\, x} + B_1\, e^{-i\, k_1\, x}$$

with $k_1 = \sqrt{2m\,(E - U_0)}/\hbar$ for $x \to +\infty$, and

$$\Psi(x) = A_2\, e^{i\, k_2\, x} + B_2\, e^{-i\, k_2\, x}$$

with $k_2 = \sqrt{2m\,E}/\hbar$ for $x \to -\infty$. Notice that the form of $\Psi(x)$ is similar for both limits but that wave vectors or momenta $p = \pm\hbar\, k$ are different. As we will discuss in more detail below the term $e^{i\, k\, x}$ ($k > 0$) corresponds to a particle having a positive velocity and current density, i.e., moving to the right, while $e^{-i\, k\, x}$ corresponds to a particle moving to the left.

4.13. Transmission and reflection coefficients

We consider a potential as in the previous section such that $U(x) \to 0$ for $x \to -\infty$ and $U(x) \to U_0 > 0$ for $x \to +\infty$. For simplicity we can assume that $U(x)$ increases monotonously with x, but this is not really necessary. Suppose we are dealing with classical particles that are moving from left to right under the action of such a force field $\vec{F} = -\vec{\nabla}\, U$. If the energy of the incident particles is $E < U_0$, a classical particle would move from left to right with a velocity v that decreases with increasing x until it reaches the return point x_c given by $U(x_c) = E$. After that, the particle is always reflected by the potential wall and continues its motion back towards $x \to -\infty$ with reversed v. If the incident energy is $E > U_0$ the classical particle overcomes the barrier and continues its motion to $x \to +\infty$, albeit with a reduced velocity $v = \sqrt{2\,(E - U_0)/m}$. In quantum mechanics, however, two new phenomena appear. First, there is a finite probability of finding the particle beyond the classical return point x_c (i.e., $|\Psi(x)|^2 > 0$ for $x > x_c$). And second, there is a finite (non-vanishing) probability that the particle is reflected even for an incident kinetic energy $E > U_0$.

We consider an incident particle from the left with $E > U_0$, which can eventually be reflected in the direction $x \to -\infty$, or transmitted to $x \to +\infty$. The wave function has thus the form

$$\Psi(x) = A_2\; e^{i\, k_2\, x} \quad \text{where} \quad k_2 = \sqrt{2m\,(E - U_0)}/\hbar$$

for $x \to +\infty$, and

$$\Psi(x) = A_1\; e^{i\, k_1\, x} + B_1\; e^{-i\, k_1\, x} \quad \text{where} \quad k_1 = \sqrt{2m\,E}/\hbar$$

for $x \to -\infty$. In order to introduce the notions of transmission and reflection coefficients we consider the *probability current density* j, which measures the flux of probability across a surface (or a given point in 1D). For the transmitted wave ($x \to +\infty$), j is given by

$j_t = \dfrac{\hbar\, k_2}{m}\, |A_2|^2 = \dfrac{p_2}{m}\, |A_2|^2$, while for the incident wave it is given by $j_i = \dfrac{\hbar\, k_1}{m}\, |A_1|^2$. The concept of probability current density is discussed in detail in Sec. 4.16.

We define the transmission coefficient T as the ratio between the transmitted and the incident probability current densities:

$$T = \frac{j_t}{j_i} = \frac{k_2\, |A_2|^2}{k_1\, |A_1|^2}.$$

Analogously, the probability current density of the reflected wave is $j_r = -\dfrac{\hbar\, k_1}{m}\, |B_1|^2$ (i.e., pointing to the negative x direction). The reflection coefficient R is then defined by the ratio

$$R = \frac{|j_r|}{|j_i|} = \frac{|B_1|^2}{|A_1|^2}$$

between the reflected and the incident probability current densities.

In any stationary state the probability of finding an electron in any finite volume is independent of time. Therefore the total flux of the probability current density over the surface of any volume must vanish. This means that the incident probability flux must be equal to the sum of the reflected and transmitted flux:

$$\frac{\hbar\, k_1}{m}\, |A_1|^2 = \frac{\hbar\, k_2}{m}\, |A_2|^2 + \frac{\hbar\, k_1}{m}\, |B_1|^2.$$

Consequently, we have

$$T + R = 1.$$

It is interesting to point out that the reflection and transmission coefficients are the same for positive and negative directions of the incoming particle. To show this we write the asymptotic values of a general solution of the Schrödinger equation for a given energy E:

$$\Psi(x) = \begin{cases} A_1\, e^{i\, k_1\, x} + B_1\, e^{-i\, k_1\, x} & \text{for} \;\; x \to -\infty \\ A_2\, e^{i\, k_2\, x} + B_2\, e^{-i\, k_2\, x} & \text{for} \;\; x \to +\infty. \end{cases} \tag{4.27}$$

We consider first the particular case $A_1 = 1$ and $B_1 = 0$ and denote the corresponding complex coefficients of the asymptotic values for $x \to +\infty$ as α and α':

$$\Psi_1(x) = \begin{cases} e^{i\, k_1\, x} & \text{for} \;\; x \to -\infty \\ \alpha\, e^{i\, k_2\, x} + \alpha'\, e^{-i\, k_2\, x} & \text{for} \;\; x \to +\infty. \end{cases}$$

Second, we consider the case $A_1 = 0$ and $B_1 = 1$ and we write

$$\Psi_2(x) = \begin{cases} e^{-i\, k_1\, x} & \text{for} \;\; x \to -\infty \\ \beta\, e^{i\, k_2\, x} + \beta'\, e^{-i\, k_2\, x} & \text{for} \;\; x \to +\infty. \end{cases}$$

Since the Schrödinger equation is linear we can construct the general solution $\Psi(x)$ given by Eq. (4.27) as a linear combination $\Psi(x) = A_1\,\Psi_1 + B_1\,\Psi_2$. Comparing the asymptotic behaviors for $x \to +\infty$ we infer

$$A_2 = \alpha\, A_1 + \beta\, B_1 \tag{4.28}$$

and

$$B_2 = \alpha' A_1 + \beta' B_1 \tag{4.29}$$

for all A_1 and $B_1 \in \mathbb{C}$. Since the Hamiltonian is real, Ψ^* is a stationary state whenever $\Psi(x)$ is a stationary state. We can therefore construct the function $\Psi(x)$ given by Eq. (4.27) by combining $\Psi_1^*(x)$ and $\Psi_2^*(x)$ with the coefficients B_1 and A_1, respectively. Indeed

$$\begin{aligned}\Psi(x) &= B_1\,\Psi_1^*(x) + A_1\,\Psi_2^*(x)\\ &= \begin{cases} A_1\, e^{i\,k_1\,x} + B_1\, e^{-i\,k_1\,x} & \text{for } x \to -\infty \\ (\alpha'^* B_1 + \beta'^* A_1)\, e^{i\,k_2\,x} + (\alpha^* B_1 + \beta^* A_1)\, e^{-i\,k_2\,x} & \text{for } x \to +\infty \end{cases}\end{aligned}$$

has the same asymptotic behavior for $x \to -\infty$ as the stationary state given by Eq. (4.27) and must therefore coincide with it for all x. Comparing with Eq. (4.27) we conclude that

$$B_2 = \alpha^* B_1 + \beta^* A_1. \tag{4.30}$$

The reflection coefficient for a particle coming from left to right [$B_2 = 0$ in Eq. (4.27)] is given by

$$R_l = \left|\frac{B_1}{A_1}\right|^2.$$

Since $B_2 = 0$ we have, using Eq. (4.30), that $\alpha^* B_1 + \beta^* A_1 = 0$. This implies

$$R_l = \left|\frac{\beta^*}{\alpha^*}\right|^2.$$

The reflection coefficient for a particle coming from the right [$A_1 = 0$ in Eq. (4.27)] is given by

$$R_r = \left|\frac{A_2}{B_2}\right|^2.$$

Since $A_1 = 0$ we have, using Eq. (4.30), $B_2 = \alpha^* B_1$ and, using Eq. (4.28), $A_2 = \beta\, B_1$. Thus,

$$R_r = \left|\frac{\beta}{\alpha^*}\right|^2 = R_l.$$

Since $T + R = 1$, the transmission coefficients T_l and T_r for left and right incidence are also the same. Notice that in classical mechanics we have $R_r = R_l = 0$ and $T_r = T_l = 1$ $(E > U_0)$. Therefore, the left-right symmetry also holds in the classical case.

If $E < U_0$, then $E - U(x) < 0$ for $x \to -\infty$ and k_2 becomes purely imaginary. The wave function $\Psi_2(x) \to 0$ for $x \to +\infty$ and the transmitted current density vanishes. In this case $T = 0$ and $R = 1$. However, notice that there is a finite probability of finding the particle beyond the classical turning point x_c. This probability decreases exponentially for $x \to +\infty$.

A few simple examples can be worked out analytically, for instance, the square potential step. For $E < E_0$, $R = 1$ and $T = 0$. For $E \geq U_0$ we have $R \to 1$ for $E \to U_0$ and $R \to 0$ for $E \to +\infty$.

Another interesting example is the rectangular potential barrier. In this case we see that the transmission coefficient does not vanish even for $E < U_0$. Since there is always a reflected wave, we must have $R \geq 0$ and $T \leq 1$. The possibility that the particles traverse a potential energy barrier is known as tunnel effect. It is a well-known specific feature of quantum mechanics. The phenomenon appears in a number of experiments in nanophysics, condensed matter, atomic physics, chemistry, etc.

4.14. Time dependence of observables and operators

In quantum mechanics the derivative of observables with respect to time cannot be defined as in classical physics. Consider for example the position of the particle as observable and its derivative, the velocity. In classical mechanics the position is an intrinsic dynamical variable of the particle and one can measure it with arbitrary accuracy at any time. The velocity $v(t)$ is then defined as

$$v(t) = \lim_{\Delta t \to 0} \frac{x(t + \Delta t) - x(t)}{\Delta t},$$

where $x(t)$ refers to the position at time t. However, in quantum mechanics the particles do not have an intrinsic value of the position. The position appears only as the result of a *measurement*, and the measurement alters the quantum state in an essential way.[5] A particle having a definite position at time t will not have a definite position at a subsequent time $t + \Delta t$. Therefore, the result of the differences $\Delta x = x(t + \Delta t) - x(t)$ can never be predicted with certainty. Only the wave functions $\Psi(x, t)$ and the probability distribution of the position $|\Psi(x, t)|^2$ are well-defined at all times. Therefore, only the average value of the position $\langle\, x \,\rangle$ and any other higher moments $\langle\, x^n \,\rangle$ of $|\Psi(x, t)|^2$ have well-defined values at all times.

The previous arguments apply to any observable. Except for the conserved quantities of a specific physical situation (i.e., for a specific Hamiltonian), an observable having a definite value at an instant t will not have a definite value at time $t + \Delta t$. Only the probability distribution $|a_f|^2$ is well-defined. It is therefore natural to define time

---

[5] As already discussed, the measurement of an observable f projects the quantum state in an eigenstate $\Psi_{f'}$ of the observable, where f' is the result of the measurement.

derivatives of observables in terms of mean values. We introduce the derivative of the physical quantity $\hat{f}$ by means of the operator $\dot{\hat{f}} = d\hat{f}/dt$, which is defined by the relation

$$\langle\, \Psi \,|\, \frac{d\hat{f}}{dt} \,|\, \Psi \,\rangle = \frac{d}{dt} \langle\, \Psi \,|\, \hat{f} \,|\, \Psi \,\rangle. \tag{4.31}$$

In other words, the average value of the quantity $\dot{\hat{f}}$ is equal to the time derivative of the average value of $\hat{f}$ for any state $|\, \Psi \,\rangle$. Notice that in this way $\dot{\hat{f}}$ is univocally defined, since two operators having the same average on any state $|\, \Psi \,\rangle$ are necessarily equal. Moreover, $\hat{f}^\dagger = \hat{f}$ implies $\dot{\hat{f}}^\dagger = \dot{\hat{f}}$. Consequently, if $\hat{f}$ is a physical observable, $\dot{\hat{f}}$ also corresponds to an observable.

An explicit expression for $\dot{\hat{f}}$ can be readily obtained from Eq. (4.31) by calculating the derivative of $\langle\, \Psi \,|\, \hat{f} \,|\, \Psi \,\rangle = \int \Psi^* \hat{f}\, \Psi \, dq$ taking into account the time dependence of $\Psi = \Psi(q,t)$, as well as a possible time dependence of $\hat{f} = \hat{f}(t)$. A straightforward differentiation yields

$$\frac{d}{dt} \langle\, \Psi \,|\, \hat{f} \,|\, \Psi \,\rangle = \int \left(\frac{\partial \Psi^*}{\partial t}\, \hat{f}\, \Psi + \Psi^* \, \frac{\partial \hat{f}}{\partial t}\, \Psi + \Psi^* \hat{f}\, \frac{\partial \Psi}{\partial t} \right) dq.$$

From the Schrödinger equation $i\,\hbar\, \frac{\partial \Psi}{\partial t} = \hat{H}\, \Psi$ we have $-i\,\hbar\, \frac{\partial \Psi^*}{\partial t} = (\hat{H}\, \Psi)^*$, and therefore

$$\int \frac{\partial \Psi^*}{\partial t}\, \hat{f}\, \Psi \, dq = \frac{i}{\hbar} \int (\hat{H}\, \Psi)^* \hat{f}\, \Psi \, dq = \int \Psi^* \, \hat{H}\, \hat{f}\, \Psi \, dq,$$

where we have used the hermiticity of $\hat{H}$ ($\hat{H}^\dagger = \hat{H}$). It follows that

$$\frac{d}{dt} \langle\, \hat{f} \,\rangle = \int \left(\Psi^* \, \frac{\partial \hat{f}}{\partial t}\, \Psi + \frac{i}{\hbar}\, \Psi^* \, \hat{H}\, \hat{f}\, \Psi - \frac{i}{\hbar}\, \Psi^* \, \hat{f}\, \hat{H}\, \Psi \right) dq.$$

Applying the definition (4.31) of $\dot{\hat{f}}$ we have

$$\int \Psi^* \dot{\hat{f}}\, \Psi \, dq = \int \Psi^* \left[\frac{\partial \hat{f}}{\partial t} + \frac{i}{\hbar} \left(H\hat{f} - \hat{f}\hat{H} \right) \right] \Psi \, dq.$$

Since this holds for all Ψ, we conclude that

$$\dot{\hat{f}} = \frac{\partial \hat{f}}{\partial t} + \frac{i}{\hbar}\, [\hat{H}, \hat{f}]. \tag{4.32}$$

At the present stage Eq. (4.32) should be regarded as the definition of the operator $\dot{\hat{f}}$ associated to a new observable, namely, the derivative of the observable $\hat{f}$. Notice that $\dot{\hat{f}}$ depends on the Hamiltonian $\hat{H}$, which contrasts with the usual definition of time derivative in classical mechanics.

In many cases of interest (e.g., $\vec{r}$, $\vec{p}$, $\vec{l}=\vec{r}\times\vec{p}$, etc.) the operator $\hat{f}$ does not depend explicitly on time ($\partial\hat{f}/\partial t=0$), so that

$$\dot{\hat{f}}=\frac{i}{\hbar}\left[\hat{H},\hat{f}\right]. \tag{4.33}$$

Assuming that $\hat{H}$ does not depend on time (isolated system) we conclude that $\dot{\hat{f}}$ does not depend explicitly on time, either. From this perspective only the wave functions $\Psi(q,t)$ or the representation-independent kets $|\,\Psi(t)\,\rangle$ depend on time. The operators associated to the observables are time-independent. This is the so-called *Schrödinger picture* in which the quantum dynamics is governed by the Schrödinger equation. An alternative approach to quantum dynamics is to consider Eq. (4.32) or (4.33) as a differential equation for the operator $\hat{f}$, which is now regarded as being time-dependent. This is the so-called *Heisenberg picture.* In the present section and in the following sections (4.15–4.17) we stick to the Schrödinger picture in which the operators are independent of time (except for an eventual explicit dependence) and the wave functions depend on t following the Schrödinger equation. The Heisenberg picture will be discussed in Sec. 4.18.

An important class of observables are the conserved quantities which mean value does not depend on time for any state $\Psi(q,t)$. In previous sections we have shown that the operators $\hat{f}$ of conserved quantities do not depend explicitly on time and commute with $\hat{H}$. Eq. (4.32) shows that this is equivalent to requiring $\dot{\hat{f}}=0$. In other words, the operator $\dot{\hat{f}}$ corresponding to the time derivative of the observable $\hat{f}$ (in short, the time derivative $\dot{\hat{f}}$ of the observable $\hat{f}$) vanishes if and only if the observable is conserved. Let us recall that if the particle is in a state $\Psi_{f'}$ with a defined value f' of a conserved observable $\hat{f}$ at a time t ($\hat{f}\,\Psi_{f'}=f'\,\Psi_{f'}$), it preserves the same defined value f' of $\hat{f}$ at all subsequent times.

Finally, one may notice that Eq. (4.32) implies

$$\left(\dot{\hat{f}}\right)^{\dagger}=\frac{\partial\hat{f}^{\dagger}}{\partial t}-\frac{i}{\hbar}\left[\hat{f}^{\dagger},\hat{H}\right]=\frac{\partial\hat{f}^{\dagger}}{\partial t}+\frac{i}{\hbar}\left[\hat{H},\hat{f}^{\dagger}\right].$$

In particular for hermitic operators we have $\left(\dot{\hat{f}}\right)^{\dagger}=\dot{\hat{f}}$. In other words, the time derivative of a physical observable is also a physical observable.

4.15. Velocity and acceleration: Ehrenfest theorem

As a first application of the derivative of an operator we determine the velocity operator

$$\hat{\vec{v}}=\frac{d}{dt}\,\hat{\vec{r}}.$$

The x component of the velocity is given by

$$\hat{v}_x=\frac{i}{\hbar}\,[\hat{H},\hat{x}],$$

since $\hat{x}$ does not depend explicitly on time. Using that $\hat{H} = \hat{p}^2/2m + U(\vec{r})$ we have

$$[\hat{H}, \hat{x}] = \frac{1}{2m}\,[\hat{p}_x^2, x] = \frac{1}{2m}\,(\hat{p}_x\,[\hat{p}_x, \hat{x}] + [\hat{p}_x, \hat{x}]\,\hat{p}_x) = -\frac{i\,\hbar}{m}\,\hat{p}_x.$$

Consequently,

$$\hat{v}_x = \frac{\hat{p}_x}{m}.$$

Proceeding analogously for the other components we obtain

$$\hat{\vec{v}} = \frac{\hat{\vec{p}}}{m}.$$

This relation is formally the same as in classical mechanics. It implies that in the classical limit, when the de Broglie wavelength is much smaller than the system dimensions, we recover the classical value of the velocity.

The eigenstates of the velocity and momentum are the same. A state having a defined velocity (or momentum) cannot have a defined position and vice versa ($[\hat{x}_i, \hat{p}_j] = i\,\hbar\,\delta_{ij}$). This means that after a measurement of the position at time t, the velocity is undefined and therefore the position at an infinitesimally close time $t + \Delta t$ is undefined. The theory reproduces the experimental observations which demonstrate the breakdown of the concept of trajectory in the microscopic world.

Let us determine the derivative of the momentum $\dfrac{d\hat{\vec{p}}}{dt}$ which is related to the acceleration operator $\hat{\vec{a}} = \dfrac{d\hat{\vec{v}}}{dt}$ by $\dfrac{d\hat{\vec{p}}}{dt} = m\,\hat{\vec{a}}$. From Eq. (4.32) we have

$$\frac{d\hat{\vec{p}}}{dt} = \frac{i}{\hbar}\,[\hat{H}, \hat{\vec{p}}],$$

since $\hat{\vec{p}}$ does not depend explicitly on time ($\partial\hat{p}/\partial t = 0$). Using that $\hat{H} = \hat{p}^2/2m + U(\vec{r})$ we have, for the x component,

$$\begin{aligned} \dot{\hat{p}}_x &= \frac{i}{\hbar}\,[U(\vec{r}), \hat{p}_x] = \frac{i}{\hbar}\,(U\hat{p}_x - \hat{p}_x\,U) \\ &= U\,\frac{\partial}{\partial x} - \frac{\partial}{\partial x}\,U = U\,\frac{\partial}{\partial x} - \frac{\partial U}{\partial x} - U\,\frac{\partial}{\partial x} \\ &= -\frac{\partial U}{\partial x}. \end{aligned}$$

Analogous expressions hold for the other components and therefore

$$\frac{d\hat{\vec{p}}}{dt} = -\vec{\nabla}\,U(\vec{r}) = \hat{\vec{F}}(\vec{r}).$$

This relation, also known as Ehrenfest theorem, is formally the same as Newton's equation in classical mechanics. As before it ensures that the classical limit is properly obtained. Moreover, for any quantum state $\Psi(\vec{r},t)$ the time derivative of the average momentum

$$\frac{d}{dt}\langle\,\hat{\vec{p}}\,\rangle = \frac{d}{dt}\langle\,\Psi\,|\,\hat{\vec{p}}\,|\,\Psi\,\rangle = \left\langle\,\Psi\,\middle|\,\frac{d\hat{\vec{p}}}{dt}\,\middle|\,\Psi\,\right\rangle = \langle\,\hat{\vec{F}}\,\rangle$$

follows the average force $\langle\,\hat{\vec{F}}\,\rangle$. In other words $\langle\,\hat{\vec{p}}\,\rangle$ follows a classical equation of motion at any instant t, provided that the quantum mechanical average of the force $\langle\vec{F}\rangle$ is calculated using the wave function $|\,\Psi(t)\,\rangle$. Notice that $\langle\vec{F}\rangle$ is not a simple function of $\langle\vec{r}\rangle$ and $\langle\vec{p}\rangle$, except for free particles and the harmonic oscillator $\left[U(x) = k\,x^2/2\right]$. One should therefore not exaggerate the meaning of this exact relation. While the Ehrenfest theorem is seldom used for electrons, it is often applied to compute the dynamics of nuclei in molecules and clusters in the framework of the Born-Oppenheimer approximation or other semiclassical approximations. Here the nuclear motion is treated classically under the action of the average force derived from the electronic ground-state wave function corresponding to the given nuclear coordinates.

4.16. The probability current density

The probability density $|\Psi(q,t)|^2$ gives the probability per unit volume of finding the electron in a small volume dq around point q at time t. It represents the particle density at point q and time t. It is therefore very interesting to investigate the time dependence of $|\Psi|^2$ and to relate it to the flux of probability density across the surface of the small volume dq:

$$\frac{\partial|\Psi|^2}{\partial t} = \frac{\partial\Psi^*}{\partial t}\,\Psi + \Psi^*\,\frac{\partial\Psi}{\partial t} = \frac{i}{\hbar}\,(\hat{H}\,\Psi)^*\,\Psi - \frac{i}{\hbar}\,\Psi^*\,(\hat{H}\,\Psi).$$

Using that

$$\hat{H} = \hat{H}^* = -\frac{\hbar^2}{2m}\,\nabla^2 + U(\vec{r})$$

we have

$$\frac{\partial|\Psi|^2}{\partial t} = -\frac{i\hbar}{2m}\,\left(\Psi\,\nabla^2\Psi^* - \Psi^*\,\nabla^2\Psi\right). \tag{4.34}$$

We can now rewrite the right-hand side with the help of known identities from vector calculus:

$$\vec{\nabla}\cdot(\varphi\,\vec{A}) = \vec{\nabla}\,\varphi\cdot\vec{A} + \varphi\,\vec{\nabla}\cdot\vec{A}$$

implies

$$\vec{\nabla}\cdot(\Psi\;\vec{\nabla}\Psi^*) = \vec{\nabla}\Psi\cdot\vec{\nabla}\Psi^* + \Psi\;\nabla^2\Psi^* \tag{4.35}$$

and

$$\vec{\nabla} \cdot (\Psi^* \, \vec{\nabla}\Psi) = \vec{\nabla}\Psi^* \cdot \vec{\nabla}\Psi + \Psi^* \, \nabla^2\Psi. \tag{4.36}$$

Subtracting Eqs. (4.35) and (4.36) we obtain the identity

$$\vec{\nabla} \cdot (\Psi \, \vec{\nabla}\Psi^* - \Psi^* \, \vec{\nabla}\Psi) = \Psi \, \nabla^2\Psi^* - \Psi^* \, \nabla^2\Psi, \tag{4.37}$$

which can be substituted in Eq. (4.34) to yield

$$\frac{\partial|\Psi|^2}{\partial t} = -\frac{i\,\hbar}{2m} \, \vec{\nabla} \cdot (\Psi \, \vec{\nabla}\Psi^* - \Psi^* \, \vec{\nabla}\Psi).$$

This has the form of a continuity equation

$$\frac{\partial \varrho}{\partial t} + \vec{\nabla} \cdot \vec{\jmath} = 0$$

provided that we identify the density $\varrho = |\Psi|^2$ with the probability density and

$$\vec{\jmath} = \frac{i\,\hbar}{2m} \, (\Psi \, \vec{\nabla}\Psi^* - \Psi^* \, \vec{\nabla}\Psi) \tag{4.38}$$

with the probability current density. Eq. (4.38) can also be written in the form

$$\vec{\jmath} = \frac{1}{2m} \left[\Psi \, (\hat{\vec{p}} \, \Psi)^* + \Psi^* (\hat{\vec{p}} \, \Psi) \right] = \frac{1}{m} \, \mathbb{R}\mathrm{e} \, \{ \Psi^* (\hat{\vec{p}} \, \Psi) \}$$

or

$$\vec{\jmath} = \frac{1}{2} \left[\Psi \, (\hat{\vec{v}} \, \Psi)^* + \Psi^* (\hat{\vec{v}} \, \Psi) \right] = \mathbb{R}\mathrm{e} \, \{ \Psi^* (\hat{\vec{v}} \, \Psi) \}.$$

Notice that $\vec{\jmath}$ vanishes for real $\Psi(x)$ since $\hat{\vec{p}} = -i\,\hbar\,\vec{\nabla}$ and $\hat{\vec{v}} = \hat{\vec{p}}/m$ are purely imaginary. Therefore in the absence of magnetic field (real $\hat{H}$) $\vec{\jmath}$ vanishes for all nondegenerate stationary states.

In order to clarify the physical interpretation of $\vec{\jmath}$ we can integrate $\dfrac{\partial|\Psi|^2}{\partial t}$ over a finite volume V. The rate of change in the probability $\int_V |\Psi|^2 \, dq$ of finding the particle in the volume V is given by

$$\frac{d}{dt} \int_V |\Psi|^2 \, dV = \int_V \frac{\partial|\Psi|^2}{\partial t} \, dV = - \int_V \vec{\nabla} \cdot \vec{\jmath} \, dV.$$

Using Gauss theorem we can write the last integral as the flux of $\vec{\jmath}$ through the closed surface S surrounding the volume V, and obtain

$$\frac{d}{dt} \int_V |\Psi|^2 \, dV = - \int_S \vec{\jmath} \cdot d\vec{s}.$$

Let us recall that the surface-element vector $d\vec{s} = ds \, \hat{n}$ is given by the surface differential ds multiplied by the outwards pointing normal to the surface $\hat{n}$. If $\vec{\jmath} \cdot d\vec{s} > 0$ the flux of probability current density across ds goes outside the volume V, and the contribution to the change in the probability of finding the particle inside V is negative. The total flux of $\vec{\jmath}$ across the surface S enclosing V gives the probability that the particle abandons the volume V per unit time.

As an example consider $\Psi(\vec{r}) = A\, e^{i\vec{k}\cdot\vec{r}}$. In this case $\hat{\vec{p}}\,\Psi = -i\,\hbar\,\vec{\nabla}\Psi = A\,\hbar\,\vec{k}\,e^{i\,k\cdot\vec{r}}$ and $\Psi^*\,\hat{\vec{p}}\,\Psi = |A|^2\,\hbar\,\vec{k}$ so that

$$\vec{\jmath} = |A|^2\,\frac{\hbar\,\vec{k}}{m} = |A|^2\,\frac{\vec{p}}{m}.$$

The current density $\vec{\jmath}$ is equal to the probability $|A|^2$ of finding the particle at $\vec{r}$ times the velocity $\vec{v} = \dfrac{\vec{p}}{m}$. The analogy with a classical fluid is clear. Notice that for real $\Psi(x)$, for instance, $\Psi(x) = \sin(\vec{k}\cdot\vec{r})$, we have $\hat{\vec{p}}\,\Psi = -i\,\hbar\,\vec{k}\,\cos(\vec{k}\cdot\vec{r})$, $\Psi^*\,\hat{\vec{p}}\,\Psi = -i\,\hbar\,\vec{k}\,\sin(\vec{k}\cdot\vec{r})\,\cos(\vec{k}\cdot\vec{r})$ and $\vec{\jmath} = 0 \ \ \forall\ \vec{r}$.

4.17. The time-evolution operator

So far we have implicitly worked in the so-called Schrödinger picture in which the quantum dynamics is derived exclusively from the time dependence of the wave function $\Psi(q,t)$ or of the corresponding ket $|\,\Psi(t)\,\rangle$. In this case the operators associated to the various observables are independent of time, except for a possible explicit time dependence found in particular cases. As we shall see this is not the only possibility of describing the quantum dynamics. In order to discuss these alternative approaches it is useful to introduce the concept of time-evolution operator.

Our starting point is the time evolution of an arbitrary ket

$$|\,\Psi(t)\,\rangle = \sum_n a_n\, e^{-\frac{i}{\hbar}\,E_n\,t}\,|\,\Psi_n(0)\,\rangle, \tag{4.39}$$

where $|\,\Psi(0)\,\rangle = \sum_n a_n\,|\,\Psi_n(0)\,\rangle$ and $|\,\Psi_n(0)\,\rangle$ are the stationary states given by $\hat{H}\,|\,\Psi_n(0)\,\rangle = E_n\,|\,\Psi_n(0)\,\rangle$. Using that

$$e^{-\frac{i}{\hbar}\,E_n\,t}\,|\,\Psi_n(0)\,\rangle = e^{-\frac{i}{\hbar}\,\hat{H}\,t}\,|\,\Psi_n(0)\,\rangle$$

we can rewrite Eq. (4.39) as

$$|\,\Psi(t)\,\rangle = \sum_n a_n\, e^{-\frac{i}{\hbar}\,\hat{H}\,t}\,|\,\Psi_n(0)\,\rangle = e^{-\frac{i}{\hbar}\,\hat{H}\,t}\,|\,\Psi(0)\,\rangle. \tag{4.40}$$

Note that if $\hat{A}$ is a linear operator, $\hat{A}^n$ is also linear $\forall\ n \in \mathbb{Z}$ and therefore $F(\hat{A})$ is also linear for any function F. Eq. (4.40) can be written in a more general form

$$|\,\Psi(t)\,\rangle = \hat{U}(t,t_0)\,|\,\Psi(t_0)\,\rangle, \tag{4.41}$$

where $\hat{U}(t,t_0)$ is the *time-evolution operator*. For closed systems, for which $\hat{H}$ is independent of time, we have

$$\hat{U}(t,t_0) = e^{-\frac{i}{\hbar}\,\hat{H}\,(t-t_0)}. \tag{4.42}$$

$\hat{U}(t, t_0)$ propagates the state $|\,\Psi\,\rangle$ from time t_0 to time t. The superposition principle implies that Eq. (4.41) is valid in general with $\hat{U}(t, t_0)$ being a linear operator (even in cases where $\hat{H}$ depends on time). The conservation of the norm, i.e.,

$$\langle\,\Psi(t)\,|\,\Psi(t)\,\rangle = \langle\,\Psi(t_0)\,|\,\Psi(t_0)\,\rangle$$

for all $|\,\Psi\,\rangle$, t and t_0, requires that

$$\hat{U}(t, t_0)^\dagger\,\hat{U}(t, t_0) = \mathbb{1}$$

or

$$\hat{U}(t, t_0)^{-1} = \hat{U}(t, t_0)^\dagger = \hat{U}(t_0, t), \tag{4.43}$$

since

$$\langle\,\Psi(t)\,|\,\Psi(t)\,\rangle = \langle\,\Psi(t_0)\,|\,\hat{U}^\dagger(t, t_0)\,\hat{U}(t, t_0)\,|\,\Psi(t_0)\,\rangle.$$

In other words, $\hat{U}$ is a unitary operator. Note that Eq. (4.43) holds if and only if $\hat{H}$ is hermitic ($\hat{H} = \hat{H}^\dagger$). One can easily verify Eq. (4.43) in the case of a closed system, where $\hat{H}$ is independent of time and $\hat{U}$ is given by Eq. (4.42).

Notice that Eqs. (4.41) and (4.42) contain the same physical information as Eq. (4.39) or as the Schrödinger equation. They have, however, the advantage of being written in a representation-independent form.

4.18. The Schrödinger and Heisenberg pictures

The discussion of quantum dynamics has been based on a description in which the wave functions $\Psi(q, t)$ or kets $|\,\Psi(t)\,\rangle$ depend on time and the operators associated to the observables stay fixed. This approach was first introduced by Schrödinger and is therefore known as *Schrödinger picture*. However, this is not the only possibility of describing quantum dynamics. Another very important approach consists in attaching all the time dependence to the operators keeping the kets $|\,\Psi\,\rangle$ independent of time. This approach, first introduced by Heisenberg, is known as the *Heisenberg picture*.

In order to discuss the relation between the two pictures and to avoid confusions we attach the subscripts S and H to the corresponding kets, bras and operators. Thus, the Schrödinger operator $\hat{f}_S$ associated with the observable f is independent of time[6] and the Schrödinger kets $|\,\Psi(t)\,\rangle_S$ are given by

$$|\,\Psi(t)\,\rangle_S = e^{-\frac{i}{\hbar}\hat{H}\,t}\,|\,\Psi(0)\,\rangle_S. \tag{4.44}$$

---

[6] For simplicity we focus on observables and operators that do not depend explicitly on time ($\frac{\partial \hat{f}}{\partial t} = 0$). In particular we assume that the Hamiltonian $\hat{H}$ corresponds to an isolated system and is therefore time-independent.

In the Heisenberg approach the quantum states $|\,\Psi\,\rangle_H$ are assumed to be independent of time. We take $t=0$ as reference time and define the $|\,\Psi\,\rangle_H$ in terms of the Schrödinger ket as

$$|\,\Psi\,\rangle_H = |\,\Psi(0)\,\rangle_S. \tag{4.45}$$

We seek now for the form of the Heisenberg operator $\hat{f}_H$ which must satisfy

$$_H\langle\,\Psi\,|\,\hat{f}_H(t)\,|\,\Psi\,\rangle_H = \ _S\langle\,\Psi(t)\,|\,\hat{f}_S\,|\,\Psi(t)\,\rangle_S \tag{4.46}$$

for all states $|\,\Psi\,\rangle$ and times t. Using that

$$|\,\Psi(t)\,\rangle_S = e^{-\frac{i}{\hbar}\hat{H}\,t}\,|\,\Psi(0)\,\rangle_S = e^{-\frac{i}{\hbar}\hat{H}\,t}\,|\,\Psi\,\rangle_H$$

and the corresponding relation for the bras,

$$_S\langle\,\Psi(t)\,| = \ _H\langle\,\Psi\,|\,e^{\frac{i}{\hbar}\hat{H}\,t},$$

we obtain

$$\hat{f}_H(t) = e^{\frac{i}{\hbar}\hat{H}\,t}\,\hat{f}_S\,e^{-\frac{i}{\hbar}\hat{H}\,t} \tag{4.47}$$

with the initial condition $\hat{f}_H(0) = \hat{f}_S$. The latter is a consequence of choosing $t=0$ as reference time [see Eq. (4.45)], as can be easily verified by setting $t=0$ in Eq. (4.46) and by noting that $\langle\,\Psi(0)\,|\,\hat{f}_H(0)\,|\,\Psi(0)\,\rangle = \langle\,\Psi(0)\,|\,\hat{f}_S\,|\,\Psi(0)\,\rangle$ for all possible initial states $|\,\Psi(0)\,\rangle$. The two approaches are thus completely equivalent since for any states $|\,\alpha\,\rangle$ and $|\,\beta\,\rangle$ we have

$$_S\langle\,\alpha(t)\,|\,\hat{f}_S\,|\,\beta(t)\,\rangle_S = \ _H\langle\,\alpha\,|\,\hat{f}_H(t)\,|\,\beta\,\rangle_H.$$

In the Schrödinger picture the kets or wave functions evolve in time following the Schrödinger equation or equivalently Eq. (4.44). In the Heisenberg picture the kets $|\,\Psi\,\rangle_H$ are independent of time and coincide with the initial state $|\,\Psi(0)\,\rangle_S$ of the Schrödinger approach. Thus,

$$|\,\Psi(t)\,\rangle_S = e^{-\frac{i}{\hbar}\hat{H}\,t}\,|\,\Psi\,\rangle_H \tag{4.48}$$

and

$$|\,\Psi\,\rangle_H = e^{\frac{i}{\hbar}\hat{H}\,t}\,|\,\Psi\,\rangle_S. \tag{4.49}$$

It is important to note that the reference time at which $|\,\Psi(0)\,\rangle_S = |\,\Psi\,\rangle_H$ (here $t=0$) is the same for all states. For example, if one considers a fixed state $|\,\alpha\,\rangle$ at different times t in the S-picture, this implies that the initial state $|\,\alpha(0)\,\rangle_S = e^{\frac{i}{\hbar}\hat{H}\,t}\,|\,\alpha\,\rangle$ leading to the same $|\,\alpha\,\rangle = e^{-\frac{i}{\hbar}\hat{H}\,t}\,|\,\alpha(0)\,\rangle_S$ at different t must depend on time. In this case, the corresponding Heisenberg ket $|\,\alpha\,\rangle_H = e^{\frac{i}{\hbar}\hat{H}\,t}\,|\,\alpha\,\rangle$ is different for different t. It has to evolve with time in the opposite sense (as the initial Schrödinger ket) in order that it corresponds to the same Schrödinger state $|\,\alpha\,\rangle$ at all t.

Let us consider the S-picture and calculate the probability amplitude ${}_S\langle\,\alpha\,|\,\Psi(t)\,\rangle_S$ of finding a state $|\,\Psi(t)\,\rangle_S$ in a fixed state $|\,\alpha\,\rangle_S$. From Eq. (4.48) we have

$$ {}_S\langle\,\alpha\,|\,\Psi(t)\,\rangle_S = {}_S\langle\,\alpha\,|\,e^{-\frac{i}{\hbar}\hat{H}\,t}\,|\,\Psi\,\rangle_H. $$

Using Eq. (4.49) we can write

$$ |\,\alpha\,\rangle_H = e^{\frac{i}{\hbar}\hat{H}\,t}\,|\,\alpha\,\rangle_S \quad \Rightarrow \quad {}_H\langle\,\alpha\,| = {}_S\langle\,\alpha\,|\,e^{-\frac{i}{\hbar}\hat{H}\,t} $$

and therefore

$$ {}_S\langle\,\alpha\,|\,\Psi(t)\,\rangle_S = {}_H\langle\,\alpha\,|\,\Psi\,\rangle_H. $$

The probability amplitude $\langle\,\alpha\,|\,\Psi\,\rangle$ and the transition probability $|\langle\,\alpha\,|\,\Psi\,\rangle|^2$ do not depend on the picture used for evaluating them provided that the relation $|\,\Psi\,\rangle_H = |\,\Psi(0)\,\rangle_S$ is fulfilled by all states.

In the S-picture the operators are usually independent of time ($\partial\hat{f}/\partial t = 0$), while in the H-picture they are functions of t. Although the time dependence is explicit in Eq. (4.47) [pretty much like the one of $|\,\Psi(t)\,\rangle_S$ in Eq. (4.44)], this equation is in general difficult to solve due to the exponential dependence on $\hat{H}$. One would like to derive a simpler first-order differential equation from which $\hat{f}_H(t)$ can be inferred. Differentiating Eq. (4.47) with respect to t, one obtains

$$ \frac{d\hat{f}_H}{dt} = \frac{i}{\hbar}\,\hat{H}\,e^{\frac{i}{\hbar}\hat{H}\,t}\,\hat{f}_H(0)\,e^{-\frac{i}{\hbar}\hat{H}\,t} - \frac{i}{\hbar}\,e^{\frac{i}{\hbar}\hat{H}\,t}\,\hat{f}_H(0)\,e^{-\frac{i}{\hbar}\hat{H}\,t}\,\hat{H} $$

or

$$ \frac{d\hat{f}_H}{dt} = \frac{i}{\hbar}\left[\hat{H}, \hat{f}_H(t)\right], \tag{4.50} $$

where we have used that the operator $\hat{f}$ in the Schrödinger picture is independent of time (i.e., $\partial\hat{f}_s/\partial t = 0$). This is the *Heisenberg equation of motion* which takes the role of the Schrödinger equation as the fundamental equation governing the quantum dynamics in the H-picture. Notice that Eq. (4.50) is a first-order homogeneous linear differential equation in $\hat{f}_H(t)$ which is completely defined by the Hamiltonian $\hat{H}$ of the system. The superposition principle holds therefore for Heisenberg operators. Eq. (4.50) and the initial condition $\hat{f}(0) = \hat{f}_S$ define $\hat{f}(t)$ univocally.

If the operator $\hat{f}$ depends explicitly on time (i.e., $\partial\hat{f}_S/\partial t \neq 0$), the condition

$$ {}_H\langle\,\Psi\,|\,\hat{f}_H(t)\,|\,\Psi\,\rangle_H = {}_S\langle\,\Psi(t)\,|\,\hat{f}_S(t)\,|\,\Psi(t)\,\rangle_S, \tag{4.51} $$

which defines $\hat{f}_H(t)$, leads to

$$ \hat{f}_H(t) = e^{\frac{i}{\hbar}\hat{H}\,t}\,\hat{f}_S(t)\,e^{-\frac{i}{\hbar}\hat{H}\,t}. $$

The Heisenberg equation of motion then reads

$$\frac{d\hat{f}_H(t)}{dt} = \frac{\partial \hat{f}_H(t)}{\partial t} + \frac{i}{\hbar}\left[\hat{H}, \hat{f}_H(t)\right], \tag{4.52}$$

where, as for any operator,

$$\frac{\partial \hat{f}_H}{\partial t} = e^{\frac{i}{\hbar}\hat{H}\,t}\ \frac{\partial \hat{f}_S(t)}{\partial t}\ e^{-\frac{i}{\hbar}\hat{H}\,t}.$$

The formalism can be generalized to the case where the Hamiltonian depends on time. This situation is found in systems under the action of external fields. In this case the equations are somewhat different, since the time evolution operator $\hat{U}(t)$ does not take the simple form (4.42). Replacing Eq. (4.41) in Eq. (4.51) one easily obtains that the proper general definition of the Heisenberg operator is

$$\hat{f}_H(t) = \hat{U}^\dagger(t)\ \hat{f}_S(t)\ \hat{U}(t). \tag{4.53}$$

Differentiating Eq. (4.41) and comparing with the Schrödinger equation one finds that

$$i\,\hbar\,\frac{\partial \hat{U}(t)}{\partial t} = \hat{H}_S(t)\ \hat{U}(t).$$

Differentiation of Eq. (4.53) finally yields

$$\frac{d\hat{f}_H(t)}{dt} = \frac{\partial \hat{f}_H(t)}{\partial t} + \frac{i}{\hbar}\left[\hat{H}_H(t), \hat{f}_H(t)\right],$$

which is formally very similar to Eq. (4.52), except for the fact that $\hat{H}_H(t)$ also needs to be propagated in time.

Notice that Eq. (4.50) for $t = 0$ coincides with the definition of the derivative of an observable with respect to time in the Schrödinger picture that we discussed in Secs. 4.14 and 4.15 [see Eq. (4.32)]. Thus $\frac{d\hat{f}_H}{dt}$ has the same physical significance as the time derivative of the observable $\hat{f}$ at time t. In particular

$$\frac{d\hat{x}_H(t)}{dt} = \frac{\hat{p}_H(t)}{m}$$

and

$$m\,\frac{d^2\hat{x}_H(t)}{dt^2} = -\vec{\nabla}\,V_H(x)$$

at all times.

As an example in order to illustrate the physical meaning of the uncertainty relations and of Heisenberg operators we would like to investigate the possibility of measuring the position of a free electron at two different times $t_0 = 0$ and t. For this purpose we calculate the commutator $[\hat{x}(0), \hat{x}(t)] = [x, \hat{x}(t)]$ where we have replaced $\hat{x}(0) = \hat{x}_S = x$ by the time-independent Schrödinger operator. Recalling that

$$\hat{x}(t) = e^{\frac{i}{\hbar}\hat{H}\,t}\,\hat{x}(0)\,e^{-\frac{i}{\hbar}\hat{H}\,t} = e^{\frac{i}{\hbar}\hat{H}\,t}\,x\,e^{-\frac{i}{\hbar}\hat{H}\,t}$$

we have

$$\begin{aligned}
[\hat{x}(0), \hat{x}(t)] &= \left[x, e^{\frac{i}{\hbar}\hat{H}\,t}\,x\,e^{-\frac{i}{\hbar}\hat{H}\,t}\right] \\
&= e^{\frac{i}{\hbar}\hat{H}\,t}\left[x, x\,e^{-\frac{i}{\hbar}\hat{H}\,t}\right] + \left[x, e^{\frac{i}{\hbar}\hat{H}\,t}\right]x\,e^{-\frac{i}{\hbar}\hat{H}\,t} \\
&= e^{\frac{i}{\hbar}\hat{H}\,t}\,x\left[x, e^{-\frac{i}{\hbar}\hat{H}\,t}\right] + \left[x, e^{\frac{i}{\hbar}\hat{H}\,t}\right]x\,e^{-\frac{i}{\hbar}\hat{H}\,t}.
\end{aligned}$$

Since for free electrons $\hat{H} = \hat{p}^2/2m$, and using that

$$[x, F(\hat{p})] = i\,\hbar\,\frac{\partial F}{\partial p}\,(\hat{p}),$$

we have

$$\left[x, e^{\pm\frac{i}{\hbar}\frac{\hat{p}^2}{2m}\,t}\right] = \pm\,i\,\hbar\left(\frac{i\,t}{\hbar}\right)\frac{\hat{p}}{m}\,e^{\pm\frac{i}{\hbar}\frac{\hat{p}^2}{2m}\,t} = \mp\,\frac{\hat{p}\,t}{m}\,e^{\pm\frac{i}{\hbar}\frac{\hat{p}^2}{2m}\,t} = \mp\,\frac{\hat{p}\,t}{m}\,e^{\pm\frac{i}{\hbar}\hat{H}\,t}.$$

Consequently,

$$\begin{aligned}
[\hat{x}(0), \hat{x}(t)] &= e^{\frac{i}{\hbar}\hat{H}\,t}\,x\,\frac{\hat{p}\,t}{m}\,e^{-\frac{i}{\hbar}\hat{H}\,t} - \frac{\hat{p}\,t}{m}\,e^{\frac{i}{\hbar}\hat{H}\,t}\,x\,e^{-\frac{i}{\hbar}\hat{H}\,t} \\
&= \frac{t}{m}\,e^{\frac{i}{\hbar}\hat{H}\,t}\,(x\,\hat{p} - \hat{p}\,x)\,e^{-\frac{i}{\hbar}\hat{H}\,t} = \frac{i\,\hbar\,t}{m},
\end{aligned}$$

where we have used that $[\hat{p}, \hat{H}] = 0$ and $[\hat{x}, \hat{p}] = i\,\hbar$. Notice that the commutator vanishes only for $t = 0$ (equal times). Therefore, it is not possible to measure the position at two subsequent times $t = 0$ and $t > 0$ with arbitrary accuracy! As t increases, $|\,[\langle\, x(0), x(t)\,\rangle]\,| = \hbar\,t/m$ increases. Consequently, for a given uncertainty $\Delta x(0)$ in the position at $t = 0$, the uncertainty $\Delta x(t)$ in the position at time t increases with t.

Besides the previous straightforward calculation there is a more elegant and simpler way to compute the commutator $[\hat{x}(0), \hat{x}(t)]$. We consider the equation of motion

$$\frac{d\hat{x}}{dt} = \frac{i}{\hbar}\left[\hat{H}, \hat{x}(t)\right] = \frac{-i}{\hbar}\left[x, \hat{H}\right] = \left(\frac{-i}{\hbar}\right)i\,\hbar\,\frac{\partial\hat{H}}{\partial p} = \frac{\hat{p}}{m} = \frac{\hat{p}(0)}{m}, \tag{4.54}$$

where we have used that

$$[x, \hat{F}(p)] = i\,\hbar\,\frac{\partial\hat{F}}{\partial p}$$

and that $[\hat{p}, \hat{H}] = 0$ for free particles, which implies that $\hat{p}(t) = \hat{p}(0)$ is independent of t.

Eq. (4.54) for $\dfrac{d\hat{x}}{dt}$ can be integrated straightforwardly to give

$$\hat{x}(t) = \hat{x}(0) + \frac{\hat{p}(0)}{m}\, t.$$

Consequently,

$$[\hat{x}(0), \hat{x}(t)] = \frac{t}{m}\,[\hat{x}, \hat{p}] = \frac{i\hbar t}{m}. \tag{4.55}$$

In order to analyze the physical meaning of Eq. (4.55) we consider the uncertainty inequality

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

where $\Delta\hat{A} = \hat{A} - \langle\, \hat{A}\,\rangle$ and $\Delta\hat{B} = \hat{B} - \langle\, \hat{B}\,\rangle$. Applying this relation to the dispersion or uncertainty $\langle\, (\Delta\hat{x})^2\,\rangle_t = \langle\, [\,\hat{x}(t) - \langle\, \hat{x}(t)\,\rangle\,]^2\,\rangle$ of the position at time t and $\langle\, (\Delta\hat{x})^2\,\rangle_0 = \langle\, [\,\hat{x}(0) - \langle\, \hat{x}(0)\,\rangle\,]^2\,\rangle$ at time $t = 0$, one obtains

$$\sqrt{\langle\, (\Delta\hat{x})^2\,\rangle_t}\, \sqrt{\langle\, (\Delta\hat{x})^2\,\rangle_0} \geq \frac{\hbar t}{2m}$$

for $t \geq 0$. As physically expected, the product of uncertainties increases with t and vanishes only for $t = 0$. For any given $t > 0$ (whatever small) $\sqrt{\langle\, (\Delta\hat{x})^2\,\rangle_t}$ diverges as the measurement of the position at $t = 0$ becomes arbitrarily sharp $\left[\,\langle\, (\Delta\hat{x})^2\,\rangle_0 \to 0\,\right]$.

The Heisenberg formulation allows us to characterize conserved quantities in a very simple and transparent way. In the H-picture a quantity is conserved if the time derivative of the corresponding operator vanishes. From Eq. (4.52) we conclude that this is the case when the operator does not depend explicitly on time (i.e., $\partial\hat{f}_S/\partial t = 0$) and $[\hat{H}, \hat{f}_H] = 0$. This is equivalent to $\partial f_S/\partial t = 0$ and $[\hat{H}, \hat{f}_S] = 0$, since $[\hat{H}, e^{\frac{i}{\hbar} H t}] = 0$. The condition characterizing conserved quantities is of course the same as the one derived in previous sections within the S-picture.

4.19. The uncertainty relation for arbitrary observables

The uncertainty relation between position and momentum has been discussed in different contexts. The impossibility of measuring the position and momentum of a particle at the same time demonstrates most clearly that quantum particles do not follow a classical trajectory. This fact has changed profoundly our views of the physical world and thus led to the new concepts inherent to quantum theory. Furthermore, we have shown that two physical observables A and B can be measured simultaneously with arbitrary accuracy in all states $|\,\Psi\,\rangle$, if and only if the corresponding operators commute (i.e., $[\hat{A}, \hat{B}] = 0$). It is the purpose of this section to quantify the relation between the commutator $[\hat{A}, \hat{B}]$ and the minimum possible value of the mean square deviations of the outcomes of the measurement of A and B in an arbitrary state $|\,\Psi\,\rangle$.

Let us recall that, given the operator $\hat{A}$ of a physical observable and a complete set of eigenstates

$$\hat{A}\,|\,a\,\rangle = a\,|\,a\,\rangle,$$

the probability of measuring the value a of A in the state $|\,\Psi\,\rangle$ is

$$P_\Psi(a) = |\langle\,a\,|\,\Psi\,\rangle|^2.$$

The average value of A according to the probability distribution $P_\Psi(a)$ is therefore

$$\langle\,A\,\rangle = \sum_a a\,P_\Psi(a) = \langle\,\Psi\,|\,\hat{A}\,|\,\Psi\,\rangle.$$

The variance ΔA^2 of the probability distribution $P_\Psi(a)$ is given by

$$\Delta A^2 = \sum_a \left(a - \langle\,A\,\rangle\right)^2\, P_\Psi(a) = \langle\,\Psi\,|\,(\hat{A} - \langle\,A\,\rangle)^2\,|\,\Psi\,\rangle. \tag{4.56}$$

Developing $(\hat{A} - \langle\,A\,\rangle)^2$ in Eq. (4.56) we recover the usual relation

$$\Delta A^2 = \langle\,\Psi\,|\,(\hat{A} - \langle\,A\,\rangle)^2\,|\,\Psi\,\rangle = \langle\,\Psi\,|\,\hat{A}^2\,|\,\Psi\,\rangle - \langle\,\Psi\,|\,\hat{A}\,|\,\Psi\,\rangle^2 \tag{4.57}$$

known from statistics. ΔA^2 gives a measure of the dispersion of the outcome of a measurement of $\hat{A}$ in the state $|\,\Psi\,\rangle$. The *uncertainty of the observable* A in the state $|\,\Psi\,\rangle$ is defined as the mean square deviation

$$\Delta A = \sqrt{\langle\,\Psi\,|\,(\hat{A} - \langle\,A\,\rangle)^2\,|\,\Psi\,\rangle}. \tag{4.58}$$

Consider for example an eigenstate $|\,a'\,\rangle$ of $\hat{A}$. In this case one says that the observable A is sharp since $\hat{A}\,|\,a'\,\rangle = a'\,|\,a'\,\rangle$ implies $\langle\,a'\,|\,\hat{A}\,|\,a'\,\rangle = a'$ and $\Delta A^2 = \langle\,a'\,|\,(\hat{A} - a')^2\,|\,a'\,\rangle = 0$.

In this context it is important to recall that for any two observables A and B, the probability distribution of the measured values $P_\Psi(a) = |\langle\,a\,|\,\Psi\,\rangle|^2$ and $P_\Psi(b) = |\langle\,b\,|\,\Psi\,\rangle|^2$ are not independent of each other, since they both derive from the same state $|\,\Psi\,\rangle$. Our goal here is to quantify this relation in terms of the uncertainties ΔA and ΔB.

Using the Schwarz inequality

$$|\langle\,\Psi\,|\,\varphi\,\rangle|^2 \leq \langle\,\Psi\,|\,\Psi\,\rangle\,\langle\,\varphi\,|\,\varphi\,\rangle,$$

which is valid for any state vectors $|\,\Psi\,\rangle$ and $|\,\varphi\,\rangle$, and a few algebraic manipulations the following important generalization of Heisenberg's uncertainty principle can be shown.

For any hermitic operators $\hat{A} = \hat{A}^\dagger$ and $\hat{B} = \hat{B}^\dagger$ and any arbitrary state $|\,\Psi\,\rangle$ it holds that

$$\sqrt{\langle\,\Psi\,|\,(\hat{A} - \langle\,A\,\rangle)^2\,|\,\Psi\,\rangle}\,\sqrt{\langle\,\Psi\,|\,(\hat{B} - \langle\,B\,\rangle)^2\,|\,\Psi\,\rangle} \geq \frac{1}{2}\left|\langle\,\Psi\,|\,[\hat{A}, \hat{B}]\,|\,\Psi\,\rangle\right|. \tag{4.59}$$

Taking advantage of the definition of uncertainty or mean square deviation ΔA [see Eqs. (4.56)–(4.58)] we can write the inequality (4.59) in a more compact form as

$$\Delta A\,\Delta B \geq \frac{1}{2}\,|\langle\,[\hat{A},\hat{B}]\,\rangle|. \tag{4.60}$$

The average of the commutator $[\hat{A},\hat{B}]$ gives a measure of the degree of incompatibility of a simultaneous measurement of A and B in the quantum state $|\,\Psi\,\rangle$. In other words, $|\langle\,[\hat{A},\hat{B}]\,\rangle|$ measures the minimum degree of uncertainty ΔB to be expected in a measurement of B if the observable A is measured at the same time with an accuracy ΔA. Moreover, the uncertainty relation (4.60) provides a means of quantifying to what extent a measurement of A alters the outcome of a measurement of B. In fact we see that unless $[\hat{A},\hat{B}] = 0$, the more accurate the measurement of A is ($\Delta A \to 0$), the more unpredictable the outcome of a measurement of B becomes ($\Delta B \to \infty$).

For incompatible observables ($[\hat{A},\hat{B}] \neq 0$) having an unbounded eigenvalue spectrum (e.g., $\hat{x}$ and $\hat{p}$) one usually finds $\Delta B \to \infty$ when $\Delta A \to 0$ since for most states $|\,\Psi\,\rangle$ the average $\langle\,\Psi\,|\,[\hat{A},\hat{B}]\,|\,\Psi\,\rangle \neq 0$. This is of course always the case when $[\hat{A},\hat{B}]$ is a constant like for $\hat{x}$ and $\hat{p}$. However, for operators with bounded eigenvalues (e.g., the components $\hat{L}_x$, $\hat{L}_y$ and $\hat{L}_z$ of the angular momentum $\hat{\vec{L}}$), the uncertainty or mean square deviation cannot diverge. Thus, ΔB remains finite even if $\Delta A = 0$. This is possible because $\langle\,\Psi\,|\,[\hat{A},\hat{B}]\,|\,\Psi\,\rangle = 0$ when $|\,\Psi\,\rangle$ is an eigenstate of $\hat{A}$. For example, for the angular momentum operators we have $[\hat{L}_x,\hat{L}_y] = i\,\hbar\,\hat{L}_z \;\Rightarrow\; \Delta L_x^2\;\Delta L_y^2 \geq \dfrac{\hbar^2}{4}\,|\langle\,L_z\,\rangle|^2$. If we consider an eigenstate of $\hat{L}_x$, i.e., $\hat{L}_x\,|\,\Psi\,\rangle = l_x\,|\,\Psi\,\rangle$, we have $\Delta L_x = 0$ and ΔL_y finite, since $\langle\,\Psi\,|\,\hat{L}_z\,|\,\Psi\,\rangle = 0$ when $|\,\Psi\,\rangle$ is an eigenstate of $\hat{L}_x$ or $\hat{L}_y$.

A few examples are worth mentioning in order to assess the implications of the uncertainty relation. In the case of compatible observables (i.e., $[\hat{A},\hat{B}] = 0$) the lower bound is zero and there is no limitation for measuring A and B with arbitrary accuracy at the same time. The states satisfying $\Delta A = 0$ and $\Delta B = 0$ are the simultaneous eigenstates of $\hat{A}$ and $\hat{B}$. Of course nothing precludes considering states $|\,\Psi\,\rangle$ where either A, B or both are not sharp (i.e., $\Delta A > 0$ and $\Delta B > 0$). We can apply Eqs. (4.59) and (4.60) to the fundamental commutation relation

$$[x,p] = i\,\hbar$$

in order to obtain

$$\Delta x\,\Delta p \geq \frac{\hbar}{2}, \tag{4.61}$$

which physical consequences have already been discussed in length.

Motivated by the inequality (4.61), we may investigate the mathematical conditions for having optimal-uncertainty kets $|\,\Psi\,\rangle$, such that the equality sign in (4.60) holds, and determine the corresponding wave functions $\Psi(x) = \langle\,x\,|\,\Psi\,\rangle$. One can show that

$$\Delta A\,\Delta B = \frac{1}{2}\,\langle\,[\hat{A},\hat{B}]\,\rangle$$

when $\widehat{\Delta A}\,|\,\Psi\,\rangle$ and $\widehat{\Delta B}\,|\,\Psi\,\rangle$ are collinear, i.e.,

$$\widehat{\Delta A}\,|\,\Psi\,\rangle = \lambda\,\widehat{\Delta B}\,|\,\Psi\,\rangle \tag{4.62}$$

with $\lambda \in \mathbb{C}$, where $\widehat{\Delta A} = \hat{A} - \langle\, A\,\rangle$ and $\widehat{\Delta B} = \hat{B} - \langle\, B\,\rangle$ are the zero-average deviation operators. This follows from the Schwarz inequality in which the equal sign holds only when the corresponding kets are a multiple of each other ($|\,\alpha\,\rangle = \lambda\,|\,\beta\,\rangle$). A further condition needs to be imposed to $|\,\Psi\,\rangle$, namely

$$\langle\,\Psi\,|\,\{\widehat{\Delta A},\widehat{\Delta B}\}\,|\,\Psi\,\rangle = \langle\,\Psi\,|\,(\widehat{\Delta A}\,\widehat{\Delta B} + \widehat{\Delta B}\,\widehat{\Delta A})\,|\,\Psi\,\rangle = 0.$$

Using Eq. (4.62) one can show that this implies that λ must be purely imaginary ($\lambda^* = -\lambda$).

To illustrate these optimal-uncertainty kets let us consider $\hat{x}$ and $\hat{p}$. The condition (4.62) reads

$$x\,\Psi(x) = \lambda\left[-i\,\hbar\,\frac{\partial}{\partial x}\,\Psi(x)\right], \tag{4.63}$$

where we have assumed for simplicity $\langle\,\Psi\,|\,\hat{x}\,\Psi\,\rangle = 0$ and $\langle\,\Psi\,|\,\hat{p}\,|\,\Psi\,\rangle = 0$. Note that $\langle\,x\,\rangle$ can be shifted at will by redefining the origin of the coordinates. Moreover, the average momentum can be shifted by considering $\Psi'(x) = \Psi(x)\,e^{\frac{i}{\hbar}\,p'\,x}$ since

$$\hat{p}\left(\Psi(x)\,e^{i\,\frac{p'}{\hbar}\,x}\right) = e^{i\,\frac{p'}{\hbar}\,x}\,\hat{p}\,\Psi(x) + \;p'\,e^{i\,\frac{p'}{\hbar}\,x}\,\Psi(x) \;\Rightarrow\; \langle\,\Psi'\,|\,\hat{p}\,|\,\Psi'\,\rangle = \langle\,\Psi\,|\,\hat{p}\,|\,\Psi\,\rangle + p'.$$

Eq. (4.63) can be written as

$$\frac{i}{\lambda\,\hbar}\,x\,dx = \frac{d\Psi}{\Psi}$$

and integrated to give

$$\frac{i}{2\,\lambda\,\hbar}\,x^2 = \ln\Psi + c \quad\Rightarrow\quad \Psi(x) = a\,e^{\frac{i}{2\,\lambda\,\hbar}\,x^2} = a\,e^{-\frac{x^2}{2\,\hbar\,|\lambda|}},$$

where in the last step we have set $\lambda = -i\,|\lambda|$ in order that $\Psi(x)$ is normalizable ($i/\lambda < 0$). This shows that the Gaussians are the optimal wave packets in the sense that they correspond to the best possible compromise between uncertainty in position and momentum. This example should also help to emphasize that $\frac{1}{2}\,|\langle\,[\hat{A},\hat{B}]\,\rangle|$ should always be regarded as a lower bound of $\Delta A\,\Delta B$ and that only very special kets actually show this lowest value for a given pair of observables. Furthermore, we can also use $|\langle\,[\hat{A},\hat{B}]\,\rangle|$ as a less rigorous order-of-magnitude estimate of $\Delta A\,\Delta B$, since the lower bound can be attained.

A particularly interesting case of the uncertainty relation is found when one of the observables is the energy, i.e., $\hat{B} = \hat{H}$ and $\Delta B = \Delta E$. In this case we have

$$\Delta A\,\Delta E \geq \frac{|\langle\,[\hat{H},\hat{A}]\,\rangle|}{2} = \frac{\hbar}{2}\left|\left\langle\,\frac{d\,\hat{A}}{dt}\,\right\rangle\right| = \frac{\hbar}{2}\left|\frac{d\langle\,A\,\rangle}{dt}\right|, \tag{4.64}$$

since $\dfrac{d\hat{A}}{dt} = \dfrac{i}{\hbar}\,[\hat{H}, \hat{A}]$. This implies that conserved quantities can be measured with arbitrary accuracy at the same time as the energy ($\Delta E\,\Delta A \geq 0$ for $d\hat{A}/dt = 0$). In the case of non-conserved quantities, Eq. (4.64) allows us to estimate the time Δt after which the quantity A (more precisely, the probability distribution for A) has changed significantly in the given state $|\,\Psi\,\rangle$. As we shall see in the following section, this characteristic time is essentially independent of the particular property considered and simply given by the uncertainty in the energy ΔE of the state $|\,\Psi\,\rangle$.

4.20. The energy-time uncertainty relation

We consider a state $|\,\Psi\,\rangle$ at $t = 0$ and an observable A and would like to estimate the time Δt required for the ket $|\,\Psi\,\rangle$ to change significantly from the perspective of this observable. In other words, how much time Δt do we need to wait in order to see a significant change in the probability distribution $P_\Psi(a)$ of the observable $\hat{A}$ in the state $\Psi(t)$? At time t the quantity A has a given probability distribution $P_\Psi(a) = |\langle\, a\,|\,\Psi\,\rangle|^2$ with an average $\langle\, A\,\rangle$ and an uncertainty or mean-square deviation ΔA. The time Δt required for observing a significant change in the probability distribution $P_\Psi(a)$ can be estimated by requiring that after this time the change in the average of $\langle\, A\,\rangle$ should be comparable with the width ΔA of the original probability distribution $P_\Psi(a)$. Mathematically this is expressed by the condition

$$|\Delta\langle\, A\,\rangle| = \left|\frac{d\,\langle\, A\,\rangle}{dt}\right| \Delta t \simeq \Delta A.$$

Using the uncertainty relation

$$\Delta E\,\Delta A \geq \frac{\hbar}{2}\left|\frac{d\langle\, A\,\rangle}{dt}\right|$$

we obtain

$$\Delta E\,\Delta t \geq \frac{\hbar}{2}. \tag{4.65}$$

Remarkably, this inequality relating Δt with the energy uncertainty ΔE in the state Ψ is *independent* of the considered observable $\hat{A}$. Eq. (4.65) is usually referred to as the *energy-time uncertainty relation.* It should be however noted that the inequality (4.65) is conceptually very different from the uncertainty relation (4.60) between two incompatible observables. In the latter case we compare the uncertainty of two quantities that are measured simultaneously in a given state $|\,\Psi\,\rangle$. In contrast in the relation (4.65) we compare the uncertainty in the energy ΔE with the time Δt after which the quantum state $|\,\Psi\,\rangle$ changes significantly. Only the energy is a physical magnitude, the time being simply a parameter in quantum mechanics. In particular there is no operator associated to time.

Physically, the energy-time uncertainty relation states that the time evolution of a state occurs faster when the energy of the system is known less precisely. In particular an energy eigenstate has $\Delta E = 0$. It is therefore a stationary state and the time required for any physical property to change is infinite. More precisely, the probability distribution $P_\Psi(a)$ for any physical property remains unchanged at all times.

Notice that the relation (4.65) is the same for all observables. Its universal validity allows us to illustrate it in a variety of physical situations. Let us consider for example a wave packet with a given spatial extension Δx. The uncertainty in the momentum can be estimated by

$$\Delta p \simeq \frac{\hbar}{\Delta x}$$

and the corresponding uncertainty in the energy is given by

$$\Delta E = \frac{dE}{dp}\,\Delta p = \frac{p}{m}\,\Delta p = v_G\,\Delta p, \tag{4.66}$$

where v_G is known as the group velocity of the wave packet. The time required for the wave packet to move a distance comparable with its spatial extension is

$$\Delta t = \frac{\Delta x}{v_G}. \tag{4.67}$$

Combining (4.66) and (4.67) we have

$$\Delta E \cdot \Delta t = v_G\,\Delta p \cdot \frac{\Delta x}{v_G} = \Delta p\,\Delta x \geq \frac{\hbar}{2}.$$

The broader the energy uncertainty of the wave packet is, the shorter is the time Δt for it to pass, i.e., the faster the time evolution of $|\,\Psi\,\rangle$.

As a further example let us consider a state given by the superposition of two stationary states $|\,\Psi_1\,\rangle$ and $|\,\Psi_2\,\rangle$ with energies E_1 and E_2 ($E_1 \neq E_2 \Rightarrow \langle\,\Psi_1\,|\,\Psi_2\,\rangle = 0$). At time t we have

$$|\,\Psi(t)\,\rangle = a_1\,e^{-\frac{i}{\hbar}\,E_1\,t}\,|\,\Psi_1\,\rangle + a_2\,e^{-\frac{i}{\hbar}\,E_2\,t}\,|\,\Psi_2\,\rangle,$$

where a_1 and a_2 are complex numbers defining the initial state $|\,\Psi(0)\,\rangle$. A measure of the change in $|\,\Psi(t)\,\rangle$ as a function of time is given by the modulus of the correlation amplitude

$$c(t) = \langle\,\Psi(0)\,|\,\Psi(t)\,\rangle = |a_1|^2\,e^{-\frac{i}{\hbar}\,E_1\,t} + |a_2|^2\,e^{-\frac{i}{\hbar}\,E_2\,t},$$

where we have used that $\langle\,\Psi_1\,|\,\Psi_2\,\rangle = 0$. Thus,

$$|c(t)|^2 = \left||a_1|^2 + |a_2|^2\,e^{-\frac{i}{\hbar}\,\Delta E\,t}\right|^2,$$

where $\Delta E = E_2 - E_1$. We see that $|c(t)|^2$ oscillates with a period

$$T = \frac{2\,\pi\,\hbar}{\Delta E}.$$

Therefore, the characteristic time of the dynamics of the system satisfies

$$\Delta t \, \Delta E \simeq 2\,\pi\,\hbar.$$

The energy-time uncertainty also applies to quasi-stationary states $|\,\Psi\,\rangle$ having a finite lifetime τ. In this context τ represents the typical time required for the system to decay into other states, i.e., for $|\,\Psi\,\rangle$ to change significantly. The relation (4.65) implies that quasi-stationary states have a finite energy width $\Delta E = \Gamma$ of the order of $\hbar/\tau$. In spectroscopy Γ is often referred to as the line width of the corresponding excited energy level.

5. The harmonic oscillator

The harmonic oscillator is a very important problem in quantum mechanics. On the one side it is simple enough to be solvable analytically using the fundamental concepts of quantum theory. Therefore it constitutes a very useful example of quantum mechanical methodology. On the other side the harmonic oscillator appears in a number of different physical situations including the vibrations of the atoms in molecules and solids, as well as the quantum theory of radiation. Understanding its properties in detail is of the utmost importance.

The basic assumption or formulation of the problem is that the potential energy takes the form

$$U(x) = \frac{1}{2}\, k\, x^2,$$

where k is a constant. It is important to remark that any potential can be approximated by this form close to minimum. This actually holds in n dimensions once the coordinate system is taken along the normal modes. Close to an n-dimensional minimum one can always write $U = \frac{1}{2} \sum_{i=1}^{n} k_i\, q_i^2$, where q_i refers to the normal coordinates. Since U is the sum of potentials of the form $U = \sum_{i=1}^{n} U_i\,(q_i)$, the n-dimensional problem can be separated in n one-dimensional ones, each of them corresponding to an harmonic oscillator. In the following we will therefore focus on the 1D case.

5.1. The classical oscillator

Let us recall the main properties of the classical solution. Since $U(x) = \frac{1}{2}\, k\, x^2$, the force is given by $F_x = -\frac{\partial U}{\partial x} = -k\, x$. The equation of motion reads

$$m\, \frac{dx^2}{dt^2} = -k\, x. \tag{5.1}$$

The restoring force is proportional to the displacement from the equilibrium position $x = 0$. We can easily solve Eq. (5.1) with an exponential ansatz of the form

$$x(t) = A\, e^{i\,\omega\, t}.$$

Substituting in Eq. (5.1) we obtain

$$-m\,\omega^2\, e^{i\,\omega\, t} = -k\, e^{i\,\omega\, t},$$

which implies

$$m\,\omega^2 = k. \tag{5.2}$$

Taking $\omega = \sqrt{k/m}$, the most general solution has the form

$$x(t) = A\, e^{i\,\omega\, t} + B\, e^{-i\,\omega\, t}, \tag{5.3}$$

or equivalently

$$x(t) = x_M \cos(\omega t + \varphi),$$

where x_M is the largest possible displacement $[\,|\cos(\omega t + \varphi)| \leq 1]$. The velocity

$$v(t) = \frac{dx}{dt} = -x_M\, \omega \sin(\omega t + \varphi)$$

vanishes at the turning points $\pm x_M$. The total energy

$$E = \frac{p^2}{2m} + \frac{1}{2}\, k\, x^2$$

is of course conserved. Since $p = m\, v = 0$ at the turning points $\pm x_M$ we have

$$E = \frac{1}{2}\, k\, x_M^2$$

or $x_M = \sqrt{2\,E/k}$. The classical motion is thus constrained to a finite region given by $|x| \leq x_M = \sqrt{2\,E/k}$.

In order to make the dependence on the vibrational frequency apparent it is customary to replace k by using Eq. (5.2) and to write the Hamiltonian in the form

$$H = \frac{p^2}{2m} + \frac{m\,\omega^2}{2}\, x^2.$$

For the sake of future comparison it is useful to express the classical solution (5.3) in terms of the initial position $x(0)$ and momentum $p(0)$. From (5.3) we have

$$x(t) = A\, e^{i\,\omega\, t} + B\, e^{-i\,\omega\, t} = (A + B) \cos\omega\, t + i\,(A - B) \sin\omega\, t$$

and

$$\begin{aligned} p(t) &= m\, \dot{x}(t) = i\, m\, \omega\, A\, e^{i\,\omega\, t} - i\, m\, \omega\, B\, e^{-i\,\omega\, t} \\ &= i\, m\, \omega\,(A - B) \cos\omega\, t - m\, \omega\,(A + B) \sin\omega\, t \end{aligned}$$

with $x(0) = A + B$ and $p(0) = i\, m\, \omega\,(A - B)$. Thus we can write

$$p(t) = p(0) \cos\omega\, t - m\, \omega\, x(0) \sin\omega\, t \tag{5.4}$$

and

$$x(t) = x(0) \cos\omega\, t + \frac{p(0)}{m\, \omega} \sin\omega\, t. \tag{5.5}$$

5.2. The quantum oscillator

In order to solve the quantum mechanical problem we consider first the Heisenberg picture and determine the time dependence of the operators $\hat{x}(t)$ and $\hat{p}(t)$. The equations of motion are

$$i\hbar\,\frac{d\hat{p}}{dt} = \left[\hat{p},\hat{H}\right] = \left[\hat{p},\frac{m\,\omega^2\,\hat{x}^2}{2}\right] = -i\,\hbar\,\frac{\partial}{\partial x}\left(\frac{m\,\omega^2\,\hat{x}^2}{2}\right) = -i\,\hbar\,m\,\omega^2\,\hat{x}$$

$$\Rightarrow \quad \frac{d\hat{p}}{dt} = -m\,\omega^2\,\hat{x} \tag{5.6}$$

and

$$i\,\hbar\,\frac{d\hat{x}}{dt} = \left[\hat{x},\hat{H}\right] = \left[\hat{x},\frac{\hat{p}^2}{2m}\right] = i\,\hbar\,\frac{\partial}{\partial p}\left(\frac{\hat{p}^2}{2m}\right) = i\,\hbar\,\frac{\hat{p}}{m}$$

$$\Rightarrow \quad \frac{d\hat{x}}{dt} = \frac{\hat{p}}{m}. \tag{5.7}$$

Multiplying Eq. (5.7) by $i\,m\,\omega$ and subtracting to Eq. (5.6) we obtain

$$\begin{aligned}\frac{d}{dt}\,(\hat{p} - i\,m\,\omega\,\hat{x}) &= -i\,m\,\omega\,\frac{\hat{p}}{m} - m\,\omega^2\,\hat{x} \\ &= -i\,\omega\,(\hat{p} - i\,m\,\omega\,\hat{x})\,.\end{aligned}$$

The non-hermitian operator[7]

$$\hat{a} = \frac{\hat{p} - i\,m\,\omega\,\hat{x}}{\sqrt{2\,m\,\hbar\,\omega}} \tag{5.8}$$

satisfies the simple equation

$$\frac{d\hat{a}}{dt} + i\,\omega\,\hat{a} = 0.$$

This implies

$$\hat{a}(t) = \hat{a}(0)\;e^{-i\,\omega\,t},$$

where $\hat{a}(0) = \dfrac{\hat{p}(0) - i\,m\,\omega\,\hat{x}(0)}{\sqrt{2\,m\,\hbar\,\omega}}$ refers to the Schrödinger operator. Let us recall that at $t = 0$ the operators for the H- and S-pictures are identical. Introducing the hermitic conjugate operator

$$\hat{a}^\dagger = \frac{\hat{p} + i\,m\,\omega\,\hat{x}}{\sqrt{2\,m\,\hbar\,\omega}}, \tag{5.9}$$

---

[7] At this point the normalizing factor $(2\,m\,\hbar\,\omega)^{-1/2}$ is not relevant. However, it will prove very useful later on in order to simplify the commutation rules for $\hat{a}$ and $\hat{a}^\dagger$. Note that $\hat{a}$ is dimensionless and that $m\,\omega\,x$ has the units of momentum.

which time dependence is given by $\hat{a}^\dagger(t) = \hat{a}^\dagger(0)\, e^{i\,\omega\,t}$, we can solve for $\hat{p}$ and $\hat{x}$ as

$$\begin{aligned}\hat{p}(t) &= \left(\frac{m\,\hbar\,\omega}{2}\right)^{1/2} \left[\hat{a}(t) + \hat{a}^\dagger(t)\right] \qquad (5.10)\\ &= \hat{p}(0)\,\cos\omega\,t - m\,\omega\,\hat{x}(0)\,\sin\omega\,t\end{aligned}$$

and

$$\begin{aligned}\hat{x}(t) &= i\left(\frac{\hbar}{2\,m\,\omega}\right)^{1/2} \left[\hat{a}(t) - \hat{a}^\dagger(t)\right] \qquad (5.11)\\ &= \hat{x}(0)\,\cos\omega\,t + \frac{\hat{p}(0)}{m\,\omega}\,\sin\omega\,t. \qquad (5.12)\end{aligned}$$

These equations are formally the same as the classical solution (5.4) and (5.5). In particular the average values of the position and momentum follow the classical trajectory as predicted by the Ehrenfest theorem.

From Eq. (5.12) it is interesting to calculate the commutator

$$[\hat{x}(0), \hat{x}(t)] = \frac{\sin\omega\,t}{m\,\omega}\,[\hat{x}(0), \hat{p}(0)] = \frac{i\,\hbar}{m\,\omega}\,\sin\omega\,t$$

between the position operator at times $t = 0$ and t. This shows that the observables $\hat{x}(t)$ and $\hat{x}(0)$ are incompatible except at the same time $t = 0$ and at integer multiples of the half period $T/2 = \pi/\omega$. The uncertainty relation implies

$$\Delta x(0)\,\Delta x(t) \geq \frac{\hbar}{2\,m\,\omega}\,|\sin\omega\,t|$$

and in the limit of short times $t \ll 1/\omega$

$$\Delta x(0)\,\Delta x(t) \geq \frac{\hbar}{2\,m}\,t.$$

For small t, $\Delta x(0)\,\Delta x(t)$ increases as expected with increasing t. Moreover, for all finite $t \neq 2\,\pi\,n/\omega$ with $n \in \mathbb{N}$, $\Delta x(t) \to +\infty$ for $\Delta x(0) \to 0$. Notice that for short times the details of the potential play no role since $[\hat{x}(0), \hat{x}(t)]$ is independent of ω for $\omega\,t \ll 1$.

Since $\hat{a}(t)$ and $\hat{a}^\dagger(t)$ have a simple oscillatory time dependence, it is interesting to express $\hat{H}$ in terms of them by replacing $\hat{p}$ and $\hat{x}$ using Eqs. (5.10) and (5.11). From Eq. (5.10) we have

$$\frac{\hat{p}^2}{2m} = \frac{1}{2m}\,\frac{m\,\hbar\,\omega}{2}\,\left(\hat{a}^2 + \hat{a}^{\dagger\,2} + \hat{a}\,\hat{a}^\dagger + \hat{a}^\dagger\,\hat{a}\right)$$

and from Eq. (5.11)

$$\frac{m\,\omega^2}{2}\,\hat{x}^2 = \frac{m\,\omega^2}{2}\left(-\frac{\hbar}{2\,m\,\omega}\right)\left(\hat{a}^2 + \hat{a}^{\dagger\,2} - \hat{a}\,\hat{a}^\dagger - \hat{a}^\dagger\,\hat{a}\right).$$

Summing the kinetic and potential energy contributions we obtain

$$\hat{H} = \frac{\hbar\,\omega}{2}\,\left(\hat{a}\,\hat{a}^\dagger + \hat{a}^\dagger\,\hat{a}\right). \tag{5.13}$$

Notice that, since $\hat{H}$ is hermitic and time-independent, we could have already expected that $\hat{a}$ and $\hat{a}^\dagger$ appear in bilinear expressions of the form $\hat{a}\,\hat{a}^\dagger + \hat{a}^\dagger\,\hat{a}$.

In order to determine the eigenenergies of $\hat{H}$ it is useful to compute the commutator

$$\begin{aligned}\left[\hat{a}, \hat{a}^\dagger\right] &= \frac{1}{2\,m\,\hbar\,\omega}\,\left[\hat{p} - i\,m\,\omega\,\hat{x},\ \hat{p} + i\,m\,\omega\,\hat{x}\right] \\ &= \frac{1}{2\,m\,\hbar\,\omega}\,\left[i\,m\,\omega\,(-i\,\hbar) - i\,m\,\omega\,(i\,\hbar)\right] \\ &= \frac{1}{2\,m\,\hbar\,\omega}\,2\,m\,\omega\,\hbar = 1.\end{aligned}$$

The commutation rule

$$\left[\hat{a}, \hat{a}^\dagger\right] = 1 \tag{5.14}$$

is of central importance to the problem. Using it we can write $\hat{H}$ in the form

$$\hat{H} = \hbar\,\omega\,\left(\hat{a}^\dagger\,\hat{a} + \frac{1}{2}\right).$$

The hermitian operator

$$\hat{n} = \hat{a}^\dagger\,\hat{a}$$

is positive definite since for all $|\,\Psi\,\rangle$

$$\langle\,\Psi\,|\,\hat{a}^\dagger\,\hat{a}\,|\,\Psi\,\rangle = \langle\,\hat{a}\,\Psi\,|\,\hat{a}\,\Psi\,\rangle \geq 0. \tag{5.15}$$

Therefore the eigenenergies have $\hbar\,\omega/2$ as a lower bound. The presence of a finite (non-zero) lower bound for the ground state energy, that is proportional to $\hbar$, is clearly a consequence of Heisenberg's uncertainty principle. The phenomenon is sometimes referred to as the zero point motion of the oscillator.

The stationary states are clearly the eigenstates of $\hat{n}$. Let $|\,n\,\rangle$ denote the normalized eigenstate with eigenvalue n:

$$\hat{n}\,|\,n\,\rangle = n\,|\,n\,\rangle \quad \text{with} \quad \langle\,n\,|\,m\,\rangle = \delta_{nm}.$$

Notice that Eq. (5.15) implies that $n \geq 0$. In other words, the eigenvalues n have zero as a lower bound. In terms of $|\,n\,\rangle$ we have

$$\hat{H}\,|\,n\,\rangle = E_n\,|\,n\,\rangle \quad \text{with} \quad E_n = \left(n + \frac{1}{2}\right)\hbar\,\omega.$$

It is easy to see that $\hat{a}\,|\,n\,\rangle$ is either zero or an eigenstate of $\hat{n}$ with eigenvalue $n-1$. To show this we calculate the effect of $\hat{n}$ on the state $\hat{a}\,|\,n\,\rangle$:

$$\begin{aligned}\hat{n}\,\hat{a}\,|\,n\,\rangle &= \hat{a}^\dagger\,\hat{a}\,\hat{a}\,|\,n\,\rangle = (\hat{a}\,\hat{a}^\dagger - 1)\,\hat{a}\,|\,n\,\rangle = \hat{a}\,\hat{a}^\dagger\,\hat{a}\,|\,n\,\rangle - \hat{a}\,|\,n\,\rangle \\ &= (n-1)\,\hat{a}\,|\,n\,\rangle.\end{aligned}$$

Thus $\hat{a}\,|\,n\,\rangle$ is either zero or it is proportional to an eigenstate $|\,n'\,\rangle$ of $\hat{n}$ which eigenvalue is reduced by one ($n' = n-1$). Let us compute the square norm of this state:

$$\langle\,\hat{a}\,n\,|\,\hat{a}\,n\,\rangle = \langle\,n\,|\,\hat{a}^\dagger\,\hat{a}\,|\,n\,\rangle = n \geq 0.$$

This implies that n must be a positive integer or zero, since for any non-integer value of n one would obtain an eigenstate of $\hat{n}$ with a negative eigenvalue by successive application of $\hat{a}$. Notice that $\langle\,\hat{a}\,n\,|\,\hat{a}\,n\,\rangle$ vanishes only for $n = 0$. Thus we must have that for all $n > 0$, a positive integer exists such that $n - k = 0$. Otherwise we would violate the lower bound existing on the eigenvalues of $\hat{n}$.

Moreover, it is easy to see that all non-negative integers are actually eigenvalues of $\hat{n}$, since for any $|\,n\rangle$ ($\hat{n}\,|\,n\,\rangle = n\,|\,n\,\rangle$), $\hat{a}^\dagger\,|\,n\,\rangle$ is also an eigenstate of $\hat{n}$ with eigenvalue $n+1$:

$$\hat{n}\,\hat{a}^\dagger\,|\,n\,\rangle = \hat{a}^\dagger\,\hat{a}\,\hat{a}^\dagger\,|\,n\,\rangle = \hat{a}^\dagger\,(\hat{a}^\dagger\,\hat{a} + 1)\,|\,n\,\rangle = (n+1)\,\hat{a}^\dagger\,|\,n\,\rangle.$$

The square norm of $\hat{a}^\dagger\,|\,n\,\rangle$ is given by

$$\langle\,\hat{a}^\dagger\,n\,|\,\hat{a}^\dagger\,n\,\rangle = \langle\,n\,|\,\hat{a}\,\hat{a}^\dagger\,|\,n\,\rangle = \langle\,n\,|\,(1+\hat{n})\,|\,n\,\rangle = n+1.$$

Finally, choosing the relative phases between $|\,n\,\rangle$, $\hat{a}\,|\,n\,\rangle$ and $\hat{a}^\dagger\,|\,n\,\rangle$ to be zero, we have

$$\hat{a}\,|\,n\,\rangle = \sqrt{n}\,|\,n-1\,\rangle$$

and

$$\hat{a}^\dagger\,|\,n\,\rangle = \sqrt{n+1}\,|\,n+1\,\rangle,$$

where we have used the orthonormality condition $\langle\,n\,|\,m\,\rangle = \delta_{nm}$. The eigenenergies are then given by

$$E_n = \left(n + \frac{1}{2}\right)\hbar\,\omega$$

with n a non-negative integer ($n \geq 0$). The operator $\hat{a}^\dagger$ ($\hat{a}$) increases (decreases) the level of excitation by one and is therefore known as the *creation (annihilation) operator*. The operator $\hat{n} = \hat{a}^\dagger\,\hat{a}$ counts the level of excitation and is therefore known as the *number operator*. Starting from an eigenstate of $\hat{n}$, applying $\hat{a}$ causes a quantum of energy $\hbar\,\omega$ to disappear ($E_{n-1} = E_n - \hbar\,\omega$), while $\hat{a}^\dagger$ adds a quantum of energy to the system ($E_{n+1} = E_n + \hbar\,\omega$).

5.3. Some properties of the stationary kets

Before we determine the wave functions $\langle\, x \,|\, n \,\rangle = \Psi_n(x)$ of the harmonic oscillator it is interesting to derive some average properties of the stationary eigenkets $|\, n \,\rangle$ in the so-called occupation number representation. For this purpose we use the commutation rules for the creation, annihilation and number operators and the properties of $|\, n \,\rangle$. Starting from

$$\hat{x} = i \left(\frac{\hbar}{2\, m\, \omega} \right)^{1/2} (\hat{a} - \hat{a}^\dagger)$$

and

$$\hat{p} = \left(\frac{m\, \hbar\, \omega}{2} \right)^{1/2} (\hat{a} + \hat{a}^\dagger)$$

we have $\langle\, n \,|\, \hat{x} \,|\, n \,\rangle = 0$ and $\langle\, n \,|\, \hat{p} \,|\, n \,\rangle = 0$, since both $\hat{a}$ and $\hat{a}^\dagger$ change the occupation number n by one and since $\langle\, n \,|\, m \,\rangle = \delta_{nm}$. The average position and momentum of all stationary states is zero. This is consistent with the fact that the eigenfunctions of an even potential have defined parity (even or odd).

In order to compute the uncertainty in position and momentum we calculate

$$\begin{aligned}
\langle\, n \,|\, \hat{p}^2 \,|\, n \,\rangle &= \left(\frac{m\, \hbar\, \omega}{2} \right) \langle\, n \,|\, (\hat{a} + \hat{a}^\dagger)^2 \,|\, n \,\rangle \\
&= \left(\frac{m\, \hbar\, \omega}{2} \right) \langle\, n \,|\, (\hat{a}^2 + \hat{a}\, \hat{a}^\dagger + \hat{a}^\dagger\, \hat{a} + \hat{a}^{\dagger\, 2}) \,|\, n \,\rangle \\
&= \left(\frac{m\, \hbar\, \omega}{2} \right) \langle\, n \,|\, (1 + 2\, \hat{a}^\dagger\, \hat{a}) \,|\, n \,\rangle = m\, \hbar\, \omega \left(n + \frac{1}{2} \right) = m\, E_n
\end{aligned}$$

and

$$\begin{aligned}
\langle\, n \,|\, \hat{x}^2 \,|\, n \,\rangle &= - \left(\frac{\hbar}{2\, m\, \omega} \right) \langle\, n \,|\, (\hat{a} - \hat{a}^\dagger)^2 \,|\, n \,\rangle \\
&= \frac{\hbar}{2\, m\, \omega} \langle\, n \,|\, a\, a^\dagger + a^\dagger\, a \,|\, n \,\rangle = \frac{\hbar}{m\, \omega} \left(n + \frac{1}{2} \right) = \frac{E_n}{m\, \omega^2}.
\end{aligned}$$

As expected, both $\langle\, \hat{x}^2 \,\rangle$ and $\langle\, \hat{p}^2 \,\rangle$ increase with increasing excitation level. Notice that the average kinetic and potential energies are both equal to $E_n/2$. For the uncertainty product one obtains

$$\Delta x\, \Delta p = \left(n + \frac{1}{2} \right) \hbar.$$

Thus the ground state corresponds to the lowest possible value allowed by the uncertainty relation (4.60). From our previous discussion of uncertainty-optimal wave functions in 1D we may already conclude that the ground-state wave function $\langle\, x \,|\, 0 \,\rangle$ is a Gaussian.

5.4. The eigenfunctions

In order to determine the wave functions $\Psi_n(x) = \langle\, x \,|\, n \,\rangle$ we start with the condition defining the ground state $|\, 0 \,\rangle$, namely, $\hat{a}\,|\, 0 \,\rangle = 0$. Using that $\hat{a} = \dfrac{\hat{p} - i\, m\, \omega\, \hat{x}}{\sqrt{2\, m\, \hbar\, \omega}}$ we can write

$$\langle\, x \,|\, (\hat{p} - i\, m\, \omega\, \hat{x})\,|\, 0 \,\rangle = 0$$

or equivalently

$$-i\,\hbar\,\frac{\partial}{\partial x}\,\langle\, x \,|\, 0 \,\rangle - i\, m\, \omega\, x\,\langle\, x \,|\, 0 \,\rangle = 0, \tag{5.16}$$

where we have used that $\langle\, x \,|\, \hat{p} \,|\, \alpha \,\rangle = -i\,\hbar\,\dfrac{\partial}{\partial x}\,\langle\, x \,|\, \alpha \,\rangle$ for any ket $|\, \alpha \,\rangle$. Let us recall that $\langle\, x \,|\, \hat{p} \,|\, \alpha \,\rangle$ represents the wave function obtained by applying the operator $\hat{p}$ to the state $|\, \alpha \,\rangle$. In wave-function language (position representation) $\langle\, x \,|\, \hat{p}\, \alpha \,\rangle$ corresponds to $\hat{p}\,\Psi_\alpha(x) = -i\,\hbar\,\dfrac{\partial\, \Psi_\alpha(x)}{\partial x} = -i\,\hbar\,\dfrac{\partial\,\langle\, x \,|\, \alpha \,\rangle}{\partial x}$, since by definition $\Psi_\alpha(x) = \langle\, x \,|\, \alpha \,\rangle$.

Once the connection between the stationary kets $|\, n \,\rangle$ and the wave function $\Psi_n(x)$ has been established it is helpful to change to the more familiar notation

$$\Psi_n(x) = \langle\, x \,|\, n \,\rangle.$$

The equation (5.16) defining the ground-state wave function then reads

$$\frac{\partial}{\partial x}\,\Psi_0(x) + \frac{m\,\omega}{\hbar}\,x\,\Psi_0(x) = 0. \tag{5.17}$$

At this point it is useful to introduce the characteristic length

$$\lambda = \left(\frac{\hbar}{m\,\omega}\right)^{1/2}$$

of the quantum oscillator. Physically, λ represents the distance at which the potential energy equals the ground-state energy $\hbar\,\omega/2$ or, in other words, the return point of a classical oscillator with energy $E = \hbar\,\omega/2$ (i.e., $m\,\omega^2\,\lambda^2/2 = \hbar\,\omega/2$). From Eq. (5.17) we obtain

$$d\Psi_0 = -\frac{x}{\lambda^2}\,\Psi_0\,dx$$

$$\frac{1}{\Psi_0}\,d\Psi_0 = -\frac{x}{\lambda^2}\,dx$$

$$\ln \Psi_0 = -\frac{x^2}{2\,\lambda^2} + c$$

$$\Psi_0(x) = A\,e^{-\frac{x^2}{2\,\lambda^2}},$$

where c and $A = e^c$ are constants. The normalization condition implies

$$\Psi_0(x) = \left(\frac{1}{\pi\,\lambda^2}\right)^{1/4}\ e^{-\frac{x^2}{2\,\lambda^2}} = \left(\frac{m\,\omega}{\pi\,\hbar}\right)^{1/4}\ e^{-\frac{x^2}{2\,\lambda^2}}.$$

The wave function of the excited states can be generated by applying the creation operator:

$$\begin{aligned}
\hat{a}^\dagger\,|\,0\,\rangle &= |\,1\,\rangle \\
(\hat{a}^\dagger)^2\,|\,0\,\rangle &= \hat{a}^\dagger\,|\,1\,\rangle = \sqrt{2}\,|\,2\,\rangle \\
(\hat{a}^\dagger)^3\,|\,0\,\rangle &= \sqrt{2}\ \ \hat{a}^\dagger\,|\,2\,\rangle = \sqrt{6}\,|\,3\,\rangle \\
(\hat{a}^\dagger)^n\,|\,0\,\rangle &= \sqrt{(n-1)!}\ \ \hat{a}^\dagger\,|\,n-1\,\rangle = \sqrt{n!}\,|\,n\,\rangle.
\end{aligned}$$

Consequently,

$$|\,n\,\rangle = \frac{1}{\sqrt{n!}}\ (\hat{a}^\dagger)^n\,|\,0\,\rangle,$$

$$\Psi_n(x) = \frac{1}{\sqrt{n!}}\ \langle\,x\,|\,(\hat{a}^\dagger)^n\,|\,0\,\rangle = \frac{1}{\sqrt{n!}}\ \left\langle\,x\,\middle|\,\left[\frac{(\hat{p}+i\,m\,\omega\,\hat{x})}{\sqrt{2\,m\,\hbar\,\omega}}\right]^n\,\middle|\,0\right\rangle$$

and

$$\Psi_n(x) = (n!\;2\,m\,\hbar\,\omega)^{-1/2}\ \left(-i\,\hbar\,\frac{\partial}{\partial x} + i\,m\,\omega\,x\right)^n\ \Psi_0(x).$$

Since $\Psi_0(x) \propto e^{-\frac{1}{2}\left(\frac{x}{\lambda}\right)^2}$, the derivatives and multiplication by x yield a polynomial of degree n times the same Gauss function. Finally, the eigenfunctions take the form

$$\Psi_n(x) = \left(\frac{1}{\pi\,\lambda^2}\right)^{1/4}\ \frac{1}{(\sqrt{2})^n\ \sqrt{n!}}\ e^{-\frac{1}{2}\left(\frac{x}{\lambda}\right)^2}\ H_n\left(\frac{x}{\lambda}\right),$$

where $\lambda = \sqrt{\hbar/m\,\omega}$ and $H_n(\xi)$ are the Hermite polynomials:

$$\begin{aligned}
H_0(\xi) &= 1 \\
H_1(\xi) &= 2\,\xi \\
H_2(\xi) &= 4\,\xi^2 - 2 \\
H_3(\xi) &= 8\,\xi^3 - 12\,\xi, \quad \text{etc.}
\end{aligned}$$

Notice that $\Psi_n(-x) = (-1)^n\,\Psi_n(x)$. As expected, the eigenfunctions have alternating parity, the ground state being nodeless and therefore even. The n-th wave function has n nodes located at the roots of $H_n(\xi)$, which are all real.

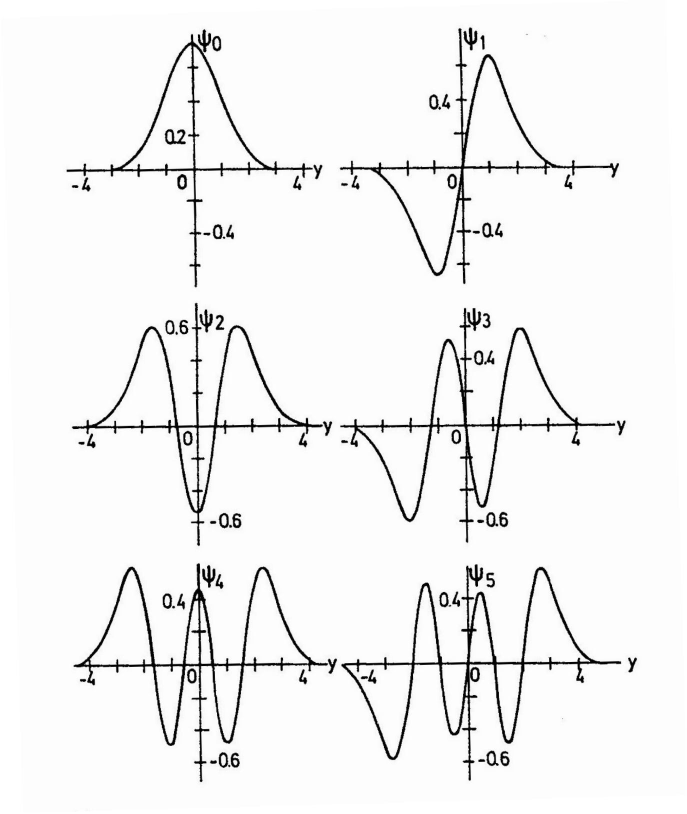


Figure 1: The stationary state wave functions $\Psi_n(y)$ of the harmonic oscillator for $n \leq 5$, as a function of $y = x/\lambda$ $(\lambda = \sqrt{\hbar/m\,\omega})$ [5]. Notice how far the particle penetrates in the classically forbidden region $|x/\lambda| > \sqrt{2\,n+1}$. It is easy to see that the classical return point x_M^n for the n-th excited state with energy $E_n = \hbar\,\omega\,(n + 1/2)$ is given by $(x_M^n/\lambda)^2 = 2\,n + 1$.

6. Angular momentum

In the absence of external fields or external potential $V(x)$ the space is homogeneous and the linear momentum $\hat{p} = -i\,\hbar\,\vec{\nabla}$ is a conserved quantity. The correlation between homogeneity of space (i.e., all points in space are equivalent or $\partial\hat{H}/\partial x = \partial\hat{H}/\partial y = \partial\hat{H}/\partial z = 0$) and momentum conservation is well known from classical mechanics ($\dot{\vec{p}} = -\vec{\nabla}H$). In addition we know that in the absence of external fields the space is also isotropic, i.e., invariant with respect to rotations. This implies that all directions in space are equivalent, and we have actually used this fact in order to infer that the free-particle Hamiltonian $H \propto p^2$. In classical mechanics the isotropy of space implies the conservation of angular momentum, which is an important observable even in the presence of an external field, particularly when it is centrally symmetric. The purpose of this chapter is to introduce the concept of angular momentum in quantum mechanics and to investigate its most fundamental properties such as its eigenvalues and eigenfunctions.

6.1. Definition and commutation rules

Following the correspondence principle the angular momentum operator $\hat{\vec{l}}$ is defined as in classical mechanics by the vector relation

$$\hbar\,\hat{\vec{l}} = \hat{\vec{r}} \times \hat{\vec{p}}, \tag{6.1}$$

where $\hat{\vec{r}} = (x, y, z) = (x_1, x_2, x_3)$ and $\hat{\vec{p}} = -i\,\hbar\,\vec{\nabla}$. The factor $\hbar$ is introduced in order to simplify the equations. Thus, $\vec{l}$ is measured in units of $\hbar$ and is therefore dimensionless. For the individual components we have

$$\hbar\,\hat{l}_x = y\,\hat{p}_z - z\,\hat{p}_y, \tag{6.2}$$

$$\hbar\,\hat{l}_y = z\,\hat{p}_x - x\,\hat{p}_z, \tag{6.3}$$

$$\hbar\,\hat{l}_z = x\,\hat{p}_y - y\,\hat{p}_x. \tag{6.4}$$

Notice that a cyclic permutation of the coordinates (i.e., $x \to y$, $y \to z$ and $z \to x$) allows one to obtain (6.3) from (6.2), (6.4) from (6.3) and (6.2) from (6.4). This symmetry will be used to avoid repeating calculations which results can be inferred by cyclic permutations. Moreover, one observes that the order of $\hat{x}_i$ and $\hat{p}_j$ in Eqs. (6.2)–(6.4) is irrelevant, since the same components are never multiplied together and $[\hat{x}_i, \hat{p}_j] = 0$ for $i \neq j$.

The fundamental commutation rule between the components of $\vec{l}$ can be calculated straightforwardly:

$$\begin{aligned}
\hbar^2\left[\hat{l}_x, \hat{l}_y\right] &= [y\,\hat{p}_z - z\,\hat{p}_y, z\,\hat{p}_x - x\,\hat{p}_z] \\
&= [y\,\hat{p}_z, z\,\hat{p}_x] + [z\,\hat{p}_y, x\,\hat{p}_z] \\
&= y\,[\hat{p}_z, z]\,\hat{p}_x + x\,[z, \hat{p}_z]\,\hat{p}_y
\end{aligned}$$

$$= -i\,\hbar\,y\,\hat{p}_x + i\,\hbar\,x\,\hat{p}_y$$

$$= i\,\hbar^2\,\hat{l}_z = i\,\hbar\,\hat{L}_z.$$

From now on we measure $\hat{\vec{l}}$ in units of $\hbar$, i.e., we replace $\hat{\vec{l}}$ by $\hat{\vec{l}}/\hbar$, so that $\hat{\vec{l}}$ becomes dimensionless. Thus,

$$\left[\hat{l}_x, \hat{l}_y\right] = i\,\hat{l}_z.$$

Cyclic permutation of the variables implies

$$\left[\hat{l}_m, \hat{l}_j\right] = i\,\hat{l}_k, \tag{6.5}$$

where (m, j, k) is a cyclic permutation of $(1, 2, 3)$. This is often written in the compact form

$$\left[\hat{l}_m, \hat{l}_j\right] = i\,\varepsilon_{mjk}\,\hat{l}_k, \tag{6.6}$$

where m, j and k refer to the components $(m, j, k = 1\text{–}3)$ and ε_{mjk} is the Levy-Civita symbol (or antisymmetric unit tensor of rank 3) satisfying $\varepsilon_{mjk} = 1$ for (m, j, k) a cyclic permutation of $(1, 2, 3)$, $\varepsilon_{mjk} = -1$ for (m, j, k) a cyclic permutation of $(2, 1, 3)$ and $\varepsilon_{mjk} = 0$ in all other cases, i.e., when two indices are equal. In Eq. (6.6) a sum over the repeated index k is implied $\left(\sum_{k=1}^{3}\right)$ in which only one term is non-vanishing.

Eq. (6.6) is the fundamental *angular momentum commutation rule.* Notice that the different components do not commute with each other. Therefore, any pair of them cannot have definite values simultaneously, except when all the components are zero $(\Delta l_x\,\Delta l_y \geq |\langle\,[\hat{l}_x, \hat{l}_y]\,\rangle|/2 = |\langle\,\hat{l}_z\,\rangle|/2)$. The angular momentum (a pseudovector) is in this respect fundamentally different from other vectors like position $\vec{r}$ or linear momentum $\vec{p}$, which components can all be measured simultaneously with arbitrary accuracy.

The commutation rules between $\vec{l}$, $\vec{r}$ and $\vec{p}$ can be easily found. It is straightforward to show that

$$\left[\hat{l}_x, x\right] = \left[\hat{l}_y, y\right] = \left[\hat{l}_z, z\right] = 0$$

and

$$\left[\hat{l}_x, y\right] = \frac{1}{\hbar}\,[y\,p_z - z\,p_y, y] = \frac{-z}{\hbar}\,[p_y, y] = i\,z.$$

Cyclic permutation yields

$$\left[\hat{l}_m, x_j\right] = i\,\varepsilon_{mjk}\,x_k, \tag{6.7}$$

where a sum over repeated indices is implied. Analogously for the linear momentum we have

$$\left[\hat{l}_x, \hat{p}_x\right] = \left[\hat{l}_y, \hat{p}_y\right] = \left[\hat{l}_z, \hat{p}_z\right] = 0$$

and

$$\left[\hat{l}_x, \hat{p}_y\right] = \frac{1}{\hbar}\,[y\,\hat{p}_z - z\,\hat{p}_y, \hat{p}_y] = \frac{1}{\hbar}\,[\hat{y}, \hat{p}_y]\,\hat{p}_z = i\,\hat{p}_z,$$

which can be written in a more compact form as

$$\left[\hat{l}_m, \hat{p}_j\right] = i\,\varepsilon_{mjk}\,\hat{p}_k. \tag{6.8}$$

A set of three operators $\hat{\vec{v}} = (\hat{v}_1, \hat{v}_2, \hat{v}_3)$ satisfying the commutation rule $[\hat{l}_m, \hat{v}_j] = i\,\varepsilon_{mjk}\,\hat{v}_k$ (like $\hat{\vec{p}}$ and $\hat{\vec{r}}$) are known as *vector operators.* Their components transform like usual vectors under rotations.[8]

With the components of $\hat{\vec{l}}$ we can construct, as for any vector, the operator of the *square modulus of* $\hat{\vec{l}}$ given by

$$\hat{l}^{\,2} = \hat{l}_x^2 + \hat{l}_y^2 + \hat{l}_z^2.$$

It is easy to see that $\hat{l}^{\,2}$ commutes with all the components of $\hat{\vec{l}}$:

$$\begin{aligned}
\left[\hat{l}^{\,2}, \hat{l}_x\right] &= \left[\hat{l}_y^2, \hat{l}_x\right] + \left[\hat{l}_z^2, \hat{l}_x\right] \\
&= \hat{l}_y\left[\hat{l}_y, \hat{l}_x\right] + \left[\hat{l}_y, \hat{l}_x\right]\hat{l}_y + \hat{l}_z\left[\hat{l}_z, \hat{l}_x\right] + \left[\hat{l}_z, \hat{l}_x\right]\hat{l}_z \\
&= i(-\hat{l}_y\,\hat{l}_z - \hat{l}_z\,\hat{l}_y + \hat{l}_z\,\hat{l}_y + \hat{l}_y\,\hat{l}_z) = 0.
\end{aligned}$$

Since the same holds for $\hat{l}_y$ and $\hat{l}_z$ (cyclic permutation leaves $\hat{l}^{\,2}$ unchanged), we have

$$\left[\hat{l}^{\,2}, \hat{l}_k\right] = 0 \qquad \forall\; k = 1, 2, 3. \tag{6.9}$$

One concludes that the square modulus $\hat{l}^{\,2}$ and any one of the components (e.g., $\hat{l}_z$) can have definite values at the same time, although the different components of $\vec{l}$ cannot have definite values simultaneously. Notice that Eq. (6.9) has been derived using exclusively the fundamental angular momentum commutation rule (6.6). Therefore, it also holds for the intrinsic angular momentum $\hat{\vec{S}}$ of elementary particles (e.g., the spin of the electron) which fulfills Eq. (6.6) but which can not be written in terms of position and momentum operators, as the orbital angular momentum in Eq. (6.1). In fact, $\hat{\vec{l}}$ commutes with the square modulus $\hat{v}^2 = \hat{v}_x^2 + v_y^2 + v_z^2$ of any vector operator $\hat{\vec{v}}$ such as $\vec{r}$ or $\vec{p}$. Physically, this corresponds to the fact that rotations do not alter the moduli of vectors.

---

[8] The commutation rules (6.7) and (6.8) can be derived more elegantly by using the known expression $(\vec{x} \times \vec{p})_k = \varepsilon_{kmn}\,x_m\,p_n$ for the components of the vector product (sum over repeated indices is implied). Noting that $\hbar\,\hat{l}_k = \varepsilon_{kmn}\,x_m\,p_n$ we have $\hbar\,[\hat{l}_k, x_j] = \varepsilon_{kmn}\,[x_m\,p_n, x_j] = \varepsilon_{kmn}\,x_m\,[p_n, x_j] = -i\,\hbar\,\varepsilon_{kmn}\,x_m\,\delta_{nj} = -i\,\hbar\,\varepsilon_{kmj}\,x_m = i\,\hbar\,\varepsilon_{kjm}\,x_m$ and $\hbar\,[\hat{l}_k, \hat{p}_j] = \varepsilon_{kmn}\,[x_m\,p_n, p_j] = \varepsilon_{kmn}\,p_n\,\delta_{mj}\,(i\,\hbar) = i\,\hbar\,\varepsilon_{kjn}\,p_n$.

6.2. Translations and rotations

The purpose of this section is to provide with a physical or geometrical interpretation of the vector operators $\hat{\vec{p}}$ and $\hat{\vec{l}}$ of the linear and angular momentum. We would like to calculate how a wave function $\Psi(\vec{r})$ changes when we translate it by an infinitesimal displacement $\delta\vec{a}$. Let $\tilde{\Psi}$ be the translated wave function which is given by

$$\hat{T}(\delta\vec{a})\,\Psi(\vec{r}) = \tilde{\Psi}(\vec{r}) = \Psi(\vec{r}-\delta\vec{a}) = \Psi(\vec{r}) - \delta\vec{a}\cdot\vec{\nabla}\Psi(\vec{r}) + O(\delta a^2).$$

Using the definition of the momentum operator $\hat{\vec{p}} = -i\,\hbar\,\vec{\nabla}$ we can write

$$\tilde{\Psi}(\vec{r}) = \left[1 - \frac{i}{\hbar}\,(-i\,\hbar\,\delta\vec{a}\cdot\vec{\nabla})\right]\Psi(\vec{r}) + O(\delta a^2) = \left(1 - \frac{i}{\hbar}\,\delta\vec{a}\cdot\hat{\vec{p}}\right)\Psi(\vec{r}) + O(\delta a^2).$$

We conclude that $\hat{\vec{p}}$ is the operator describing an infinitesimal translation of the wave function. We actually say that $\hat{\vec{p}}$ is the *generator of infinitesimal translations.* The translated wave function $\tilde{\Psi}$ is usually written in terms of the linear operator $\hat{T}(\delta\vec{a})$ for infinitesimal translations as

$$\tilde{\Psi}(\vec{r}) = \hat{T}(\delta\vec{a})\;\Psi(\vec{r})$$

with

$$\hat{T}(\delta\vec{a}) = 1 - \frac{i}{\hbar}\,\delta\vec{a}\cdot\hat{\vec{p}} + O(\delta a^2).$$

Notice that $\delta\hat{U} = \hat{T}(\delta\vec{a})$ is unitary to first order in $\delta\vec{a}$, since $\delta\hat{U}\,\delta\hat{U}^\dagger = 1 + O(\delta a^2)$, where we have used that $\hat{\vec{p}} = \hat{\vec{p}}^\dagger$ is hermitic. The fact that $\delta\hat{U}$ is unitary was of course expected, since a translation (infinitesimal or finite) does not change the norm of $\Psi(\vec{r})$.

We say that an observable $\hat{A}$ is *invariant* under a given transformation, in the present case a translation, if the result of applying the operator to the translated function, $\hat{A}\,\tilde{\Psi}$, is the same as the translation of the function $\hat{A}\,\Psi$. Mathematically, invariance upon translation means

$$\hat{A}\,\tilde{\Psi} = \hat{A}\,\hat{T}(\delta\vec{a})\,\Psi = \hat{T}(\delta\vec{a})\,\hat{A}\,\Psi$$

for all Ψ, which is equivalent to

$$\hat{A}\,\hat{T}(\delta\vec{a}) = \hat{T}(\delta\vec{a})\,\hat{A}$$

or

$$\hat{A}\,\hat{\vec{p}} = \vec{p}\,\hat{A} \quad \Leftrightarrow \quad \left[\vec{p}, \hat{A}\right] = 0. \tag{6.10}$$

We conclude that the linear momentum operator commutes with translational invariant observables and vice versa. Moreover, in the case where the Hamiltonian $\hat{H}$ is translational invariant

$$\frac{d\hat{\vec{p}}}{dt} = \frac{i}{\hbar}\left[\hat{H}, \hat{\vec{p}}\right] = 0 \tag{6.11}$$

and the linear momentum is a conserved quantity.

An alternative, maybe more explicit proof of the previous statements is obtained by calculating the change in an operator, for example the Hamiltonian $\hat{H}$, resulting from a translation $\delta\vec{a}$ of the coordinates:

$$\begin{aligned} \delta\hat{H} &= \hat{H}\left(\vec{r}+\delta\vec{a},\hat{\vec{p}}\right) - \hat{H}\left(\vec{r},\hat{\vec{p}}\right) = \delta\vec{a}\cdot\vec{\nabla}_{\vec{r}}\,\hat{H} = \sum_{i=1}^{3}\delta a_i\,\frac{\partial\hat{H}}{\partial x_i} = \frac{i}{\hbar}\sum_{i=1}^{3}\delta a_i\left[\hat{p}_i,\hat{H}\right] \\ &= \frac{i}{\hbar}\,\delta\vec{a}\cdot\left[\hat{\vec{p}},\hat{H}\right], \end{aligned} \tag{6.12}$$

where we have used that

$$\left[\hat{p}_i,\hat{F}\right] = -i\,\hbar\,\frac{\partial\hat{F}}{\partial x_i} \tag{6.13}$$

for any function $\hat{F}(\vec{r})$ of the position operator $\hat{\vec{r}}$ $\left[\hat{\vec{r}}=(x_1,x_2,x_3)\text{ and }\hat{\vec{p}}=(\hat{p}_1,\hat{p}_2,\hat{p}_3)\right]$.

Let us recall that a similar relation

$$\left[\hat{x}_i,\hat{G}\right] = i\,\hbar\,\frac{\partial\hat{G}}{\partial p_i} \tag{6.14}$$

holds for the commutator of the position operator $\hat{\vec{r}}=(x_1,x_2,x_3)$ and any function $G(\vec{p})$ of the linear momentum.[9]

We now turn to the somewhat more complicated case of rotations. We would like to calculate how the wave function $\Psi(\vec{r})$ changes when we rotate the system by a small angle $\delta\varphi$ around an arbitrary axis $\hat{n}$. We know that the change $\delta\vec{r}$ in the position vector $\vec{r}$ is given by

$$\delta\vec{r} = \delta\vec{\varphi}\times\vec{r},$$

where $\delta\vec{\varphi}=\delta\varphi\,\hat{n}$. The rotated wave function $\tilde{\Psi}$ is given by

$$\begin{aligned} \tilde{\Psi}(\vec{r}) &= \Psi(\vec{r}-\delta\vec{r}) = \Psi(\vec{r}) - \delta\vec{r}\cdot\vec{\nabla}\,\Psi(\vec{r}) + O\left(\delta r^2\right) \\ &= \left[1-\frac{i}{\hbar}\left(\delta\vec{\varphi}\times\vec{r}\right)\cdot\hat{\vec{p}}\right]\Psi(\vec{r}) + O\left(\delta r^2\right). \end{aligned}$$

Using the relation $(\vec{a}\times\vec{b})\cdot\vec{c}=\vec{a}\cdot(\vec{b}\times\vec{c})$, which holds also for non-commuting operators, we have

$$\begin{aligned} \tilde{\Psi}(\vec{r}) &= \left[1-\frac{i}{\hbar}\,\delta\vec{\varphi}\cdot\left(\vec{r}\times\hat{\vec{p}}\right)\right]\Psi(\vec{r}) + \delta\left(\delta\varphi^2\right) \\ &= \left(1-i\,\delta\vec{\varphi}\cdot\vec{l}\right)\Psi(\vec{r}) + O\left(\delta\varphi^2\right). \end{aligned}$$

---

[9] The proof of Eqs. (6.13) and (6.14) is straightforward by induction once $F(\vec{r})$ are expanded in a Taylor series. In the case of Eq. (6.13) we have $[p,x^n]=x\left[p,x^{n-1}\right]+[p,x]\,x^{n-1}=-i\,\hbar\,x\,(n-1)\,x^{n-2}-i\,\hbar\,x^{n-1}=-i\,\hbar\,n\,x^{n-1}$. The proof of Eq. (6.14) is analogous.

We conclude that $\hbar\hat{\vec{l}} = \vec{r} \times \hat{\vec{p}}$ is the operator describing an infinitesimal rotation of the wave function. One therefore says that $\hat{\vec{l}}$ is the *generator of infinitesimal rotations.* As in the case of translations it is useful to write

$$\tilde{\Psi}(\vec{r}) = \hat{D}(\delta\vec{\varphi})\,\Psi(r),$$

where the operator of an infinitesimal rotation $\hat{D}(\delta\vec{\varphi})$ is given by

$$\hat{D}(\delta\vec{\varphi}) = 1 - i\,\delta\vec{\varphi}\cdot\hat{\vec{l}} + O(\delta\varphi^2).$$

As in the case of translations, $\hat{D}(\delta\vec{\varphi})$ is unitary up to first order in $\delta\varphi$ $\Big[\hat{D}(\delta\vec{\varphi})\,\hat{D}(\delta\vec{\varphi})^\dagger = 1$ $+O\left(\delta\varphi^2\right)\Big]$ since $\hat{\vec{l}} = \hat{\vec{l}}^\dagger$ is hermitic.

The arguments used to derive Eqs. (6.10) and (6.11) apply to any symmetry transformation $\delta\hat{U}$. We can therefore conclude that the orbital angular momentum $\hat{\vec{l}}$ commutes with any rotational invariant observable. Moreover, if the Hamiltonian is rotational invariant

$$\frac{d\hat{\vec{l}}}{dt} = \frac{i}{\hbar}\left[\hat{H}, \hat{\vec{l}}\right] = 0$$

and the angular momentum is conserved.

The calculation of the change in the Hamiltonian (or any other operator) under an infinitesimal rotation is somewhat more complicated, since a rotation implies a change in the position $\delta\vec{r} = \delta\vec{\varphi} \times \vec{r}$ as well as in the momentum $\delta\vec{p} = \delta\vec{\varphi} \times \vec{p}$. We consider a single-particle Hamiltonian $\hat{H} = \hat{H}(\vec{r}, \vec{p})$ given by the sum of the kinetic energy operator $\hat{T}(\vec{p}) = \hat{p}^2/2\,m$ and the potential energy $V(\vec{r})$.[10] The change in the Hamiltonian following an infinitesimal rotation is

$$\begin{aligned}
\delta\hat{H} &= \hat{H}(\vec{r} + \delta\vec{\varphi} \times \vec{r}, \vec{p} + \delta\vec{\varphi} \times \vec{p}) - \hat{H}(\vec{r}, \vec{p}) \\
&= (\delta\vec{\varphi} \times \vec{r}) \cdot \vec{\nabla}_{\vec{r}}\hat{H} + (\delta\vec{\varphi} \times \vec{p}) \cdot \vec{\nabla}_{\vec{p}}\hat{H} \\
&= \delta\vec{\varphi} \cdot (\vec{r} \times \vec{\nabla}_{\vec{r}}\hat{H}) + \delta\vec{\varphi} \cdot (\vec{p} \times \vec{\nabla}_{\vec{p}}\hat{H}) \\
&= \delta\vec{\varphi} \cdot (\vec{r} \times \vec{\nabla}_{\vec{r}}V) + \delta\vec{\varphi} \cdot (\vec{p} \times \vec{\nabla}_{\vec{p}}\hat{T}).
\end{aligned}$$

Using the relations

$$i\left[\hat{\vec{l}}, \hat{F}(\vec{r})\right] = \vec{r} \times \vec{\nabla}_r\,\hat{F}$$

and

$$i\left[\hat{\vec{l}}, \hat{G}(\vec{p})\right] = \vec{p} \times \vec{\nabla}_{\vec{p}}\,\hat{G}$$

---

[10] The extension to a many-particle Hamiltonian $\hat{H}(\vec{r}_1, \ldots \vec{r}_N, \vec{p}_1, \ldots \vec{p}_N)$ including interactions is rather straightforward.

we can write

$$\begin{aligned}\delta\hat{H} &= i\,\delta\vec{\varphi}\cdot\left[\hat{\vec{l}},\hat{V}\right] + i\,\delta\vec{\varphi}\cdot\left[\hat{\vec{l}},\hat{T}\right] \\ &= i\,\delta\vec{\varphi}\cdot\left[\hat{\vec{l}},\hat{H}\right].\end{aligned}$$

Consequently, the angular momentum operator $\hat{\vec{l}}$ commutes with rotational invariant observables. Moreover, if the Hamiltonian is invariant under rotations, the angular momentum is conserved. The most important and fundamental examples of rotational invariant Hamiltonians are the atoms.

6.3. Angular momentum operator in spherical coordinates

Taking into account that $\hat{\vec{l}}$ is conserved when the external potential $V = V(r)$ is spherically symmetric, it is important to know how to express $\hat{\vec{l}}$ and $\hat{l}^2$ in spherical coordinates. Let us recall the usual relation between spherical coordinates (r, φ, θ) and the Cartesian coordinates (x, y, z):

$$\begin{aligned} x &= r\,\sin\theta\,\cos\varphi \\ y &= r\,\sin\theta\,\sin\varphi \\ z &= r\,\cos\theta. \end{aligned} \tag{6.15}$$

For any function $\Psi(\vec{r})$ we have

$$\frac{\partial\Psi}{\partial\varphi} = \frac{\partial\Psi}{\partial x}\,\frac{\partial x}{\partial\varphi} + \frac{\partial\Psi}{\partial y}\,\frac{\partial y}{\partial\varphi},$$

since z is independent of φ. Using Eqs. (6.15) we have

$$\begin{aligned}\frac{\partial\Psi}{\partial\varphi} &= -r\,\sin\theta\,\sin\varphi\,\frac{\partial\Psi}{\partial x} + r\,\sin\theta\,\cos\varphi\,\frac{\partial\Psi}{\partial y} \\ &= -y\,\frac{\partial\Psi}{\partial x} + x\,\frac{\partial\Psi}{\partial y} \\ &= \frac{i}{\hbar}\left(y\,i\hbar\,\frac{\partial\Psi}{\partial x} - x\,i\hbar\,\frac{\partial\Psi}{\partial y}\right) = \frac{i}{\hbar}\left(x\,\hat{p}_y - y\,\hat{p}_x\right)\Psi.\end{aligned}$$

Recalling that $\hbar\,\vec{l} = \vec{r}\times\vec{p}$, this implies

$$\hat{l}_z = -i\,\frac{\partial}{\partial\varphi} \quad\Leftrightarrow\quad \hat{L}_z = -i\hbar\,\frac{\partial}{\partial\varphi}.$$

As expected the z component of the angular momentum operator represents the effect of an infinitesimal rotation of the wave function around the z axis. It is actually analogous to a Cartesian component of the linear momentum operator, for example, $\hat{p}_x = -i\,\hbar\,\dfrac{\partial}{\partial x}$.

The eigenfunctions of $\hat{l}_z$ satisfy the equation

$$-i\,\frac{\partial\,\Psi(r,\theta,\varphi)}{\partial\varphi} = m\,\Psi(r,\theta,\varphi),$$

where $m \in \mathbb{R}$ denotes the eigenvalue. This can be easily integrated to yield

$$\Psi_m(r,\varphi,\theta) = f(r,\theta)\,e^{i\,m\,\varphi}. \tag{6.16}$$

If we focus on single-valued functions, Ψ_m must be periodic in φ with a period $2\,\pi$, which implies that the eigenvalues of $\hat{l}_z$ must be integers $m = 0, \pm 1, \pm 2, \dots$.

For the orbital angular momentum $\hbar\,\vec{l} = \vec{r} \times \vec{p}$, integer values of m and single-valued wave functions $\Psi_m(r,\varphi,\theta)$ are the only possibility. However, from a more general perspective one could also expect m to be a half integer $m = \pm 1/2, \pm 3/2, \dots$. To see this, one should recall that the wave function Ψ cannot be directly measured in experiment. Instead, the experimentally accessible quantity is the probability distribution given by $|\Psi|^2$. Therefore, it is actually $|\Psi|^2$ that always has to be single-valued, i.e., periodic in φ with period $2\,\pi$. Of course, applying this condition to the eigenfunctions of $\hat{l}_z$ given by Eq. (6.16) gives no information on the possible values of m, since $|e^{i\,m\,\varphi}|^2 = 1$ for all $m \in \mathbb{R}$. The point is that the single-valued condition on $|\Psi|^2$ must hold for any physical state and therefore for any linear combination of eigenstates of $\hat{l}_z$. Let

$$\Psi = \alpha\;e^{i\,m\,\varphi} + \beta\;e^{i\,m'\,\varphi},$$

then

$$\begin{aligned}|\Psi|^2 &= |\alpha|^2 + |\beta|^2 + \alpha\,\beta^*\;e^{i\,(m-m')\,\varphi} + \alpha^*\beta\;e^{-i\,(m-m')\,\varphi}\\ &= |\alpha|^2 + |\beta|^2 + 2\;\mathrm{Re}\,\{\alpha\,\beta^*\;e^{i\,(m-m')\,\varphi}\}.\end{aligned}$$

The $2\,\pi$ periodicity of $|\Psi|^2$ implies that $m - m'$ must be an integer. Taking $m' = -m$ one concludes that $2\,m$ must be an integer, and thus the eigenvalues of angular momentum operators must be *integer* or *half-integer*. The physical consequences of half-integer values of the angular momentum, as found for instance in the case of the electron spin ($S = 1/2$), will be discussed later on.

6.4. Eigenvalues and eigenstates of $\hat{l}_z$ and $\hat{l}^2$

We have shown that $\left[\hat{l}_z, \hat{l}^{\,2}\right] = 0$ and therefore it is possible to find a complete basis of simultaneous eigenstates of $\hat{l}^{\,2}$ and $\hat{l}_z$. We denote these eigenstates by $|\,\lambda, m\,\rangle$ so that

$$\hat{l}^{\,2}\,|\,\lambda, m\,\rangle = \lambda\,|\,\lambda, m\,\rangle$$

with $\lambda \geq 0$ and

$$\hat{l}_z\,|\,\lambda, m\,\rangle = m\,|\,\lambda, m\,\rangle.$$

The derivation of the eigenvalues of $\hat{l}^2$ and $\hat{l}_z$ will be done on the sole basis of the commutation rule

$$\left[\hat{l}_x, \hat{l}_z\right] = i\,\hat{l}_z$$

and its cyclic permutations. It is therefore valid for any type of angular momentum: orbital or intrinsic, integer or half-integer. We have already shown that

$$\left[\hat{l}^2, \hat{l}_z\right] = 0.$$

Since

$$\hat{l}^2 = \hat{l}_x^2 + \hat{l}_y^2 + \hat{l}_z^2$$

it is clear that

$$\langle\, \alpha \,|\, \hat{l}^2 - \hat{l}_z^2 \,|\, \alpha \,\rangle = \langle\, \alpha \,|\, \hat{l}_x^2 + \hat{l}_y^2 \,|\, \alpha \,\rangle \geq 0$$

for all states $|\, \alpha \,\rangle$. In particular for the eigenstates $|\, \lambda, m \,\rangle$ we have

$$\lambda \geq m^2$$

or

$$\sqrt{\lambda} \geq |m|.$$

This means that for any given eigenvalue of $\hat{l}^2$ the eigenvalues of $\hat{l}_z$ are bounded. We shall denote by $l = l_z^{\max}$ the largest possible value of m for the given λ. Note that l can be smaller than $\sqrt{\lambda}$.

The eigenstates having different m but the same λ of $\hat{l}^2$ can be related to each other with the help of the so-called ladder operators

$$\hat{l}_+ = \hat{l}_x + i\,\hat{l}_y$$

and

$$\hat{l}_- = \hat{l}_x - i\,\hat{l}_y,$$

in a similar way as the creation and annihilation operators relate different eigenstates of the number operator in the harmonic oscillator. One first notices the important commutation relations

$$\left[\hat{l}_z, \hat{l}_\pm\right] = \left[\hat{l}_z, \hat{l}_x\right] \pm i \left[\hat{l}_z, \hat{l}_y\right] = i\,\hat{l}_y \pm \hat{l}_x = \pm \hat{l}_\pm .$$

This implies

$$\hat{l}_z\,\hat{l}_+ - \hat{l}_+\,\hat{l}_z = \hat{l}_+ \quad \Leftrightarrow \quad \hat{l}_z\,\hat{l}_+ = \hat{l}_+ \left(\hat{l}_z + 1\right).$$

Consider now an eigenstate $|\,\lambda, m\,\rangle$, then we have

$$\hat{l}_z\left(\hat{l}_+\,|\,\lambda, m\,\rangle\right) = \hat{l}_+\left(\hat{l}_z + 1\right)|\,\lambda, m\,\rangle = (m+1)\,\hat{l}_+\,|\,\lambda, m\,\rangle.$$

Therefore, $\hat{l}_+\,|\,\lambda, m\rangle$ is either zero or it is an eigenstate of $\hat{l}_z$ with the eigenvalue l_z increased by one. We can then write

$$\hat{l}_+\,|\,\lambda, m\,\rangle = c_+(m)\,|\,\lambda, m+1\,\rangle,$$

where $c_+(m)$ is a constant (eventually equal to zero) that depends on m. Moreover, we know that the eigenvalues of $\hat{l}_z$ for a given λ are bounded by the relation $l_z \leq l = l_z^{\max}$, where $l \leq \lambda$. We must therefore have $c_+(l) = 0$ or equivalently

$$\hat{l}_+\,|\,\lambda, l\,\rangle = 0, \tag{6.17}$$

since otherwise, by applying $\hat{l}_+$, one would obtain an eigenstate of $\hat{l}_z$ with eigenvalue larger than l.

The equation derived in the following is important in order to relate λ and l:

$$\begin{aligned}
\hat{l}_-\,\hat{l}_+ &= (\hat{l}_x - i\,\hat{l}_y)\,(\hat{l}_x + i\,\hat{l}_y) \\
&= \hat{l}_x^2 + \hat{l}_y^2 + i\,(\hat{l}_x\,\hat{l}_y - \hat{l}_y\,\hat{l}_x) \\
&= \hat{l}_x^2 + \hat{l}_y^2 - \hat{l}_z \\
&= \hat{l}^2 - \hat{l}_z^2 - \hat{l}_z.
\end{aligned} \tag{6.18}$$

Applying $\hat{l}_-\,\hat{l}_+$ to $|\,\lambda, l\,\rangle$ and using Eqs. (6.17) and (6.18) we have

$$0 = \hat{l}_-\,\hat{l}_+\,|\,\lambda, l\,\rangle = \left(\hat{l}^2 - \hat{l}_z^2 - \hat{l}_z\right)|\,\lambda, l\,\rangle = \left(\lambda - l^2 - l\right)|\,\lambda, l\,\rangle.$$

One concludes that

$$\lambda = l\,(l+1).$$

Since λ is a simple function of $l = l_z^{\max}$, it is meaningful to label the eigenstates of $\hat{l}^2$ in terms of l as $|\,l, m\,\rangle$ instead of using the actual eigenvalue λ of $\hat{l}^2$ $[\lambda = l\,(l+1)]$. Thus, summarizing our results in the new notation, we have

$$\begin{aligned}
&|m| \leq l, \\
&\hat{l}^2\,|\,l, m\,\rangle = l\,(l+1)\,|\,l, m\,\rangle, \\
&\hat{l}_z\,|\,l, m\,\rangle = m\,|\,l, m\,\rangle,
\end{aligned}$$

and

$$\hat{l}^+\,|\,l, m\,\rangle = c^+(m)\,|\,l, m+1\,\rangle$$

with $c^+(l) = 0$.

Proceeding as in Eq. (6.18) it is easy to see that

$$\hat{l}_\mp \hat{l}_\pm = \hat{l}^2 - \hat{l}_z^2 \mp \hat{l}_z. \tag{6.19}$$

We can now repeat the previous reasoning for $\hat{l}_-$. In this case we have

$$\left[\hat{l}_z, \hat{l}_-\right] = -\hat{l}_-.$$

Therefore,

$$\hat{l}_z \left(\hat{l}_- \,|\, l\, m\,\rangle\right) = \left(-\hat{l}_- + \hat{l}_- \hat{l}_z\right) |\, l\, m\,\rangle = (m-1)\, \hat{l}_- \,|\, l\, m\,\rangle,$$

which implies

$$\hat{l}_- \,|\, l\, m\,\rangle = c_-(m)\,|\, l, m-1\,\rangle.$$

In other words, $\hat{l}_- \,|\, l\, m\,\rangle$ is an eigenstate of $\hat{l}_z$ with its eigenvalue reduced by 1. For the minimum value of $m = l_z^{\min}$ we must have

$$\hat{l}_- \,|\, l, l_z^{\min}\,\rangle = 0. \tag{6.20}$$

Using Eqs. (6.19) and (6.20) one obtains

$$0 = \hat{l}_+ \hat{l}_- \,|\, l, l_z^{\min}\,\rangle = \left(\hat{l}^2 - \hat{l}_z^2 + \hat{l}_z\right) |\, l, l_z^{\min}\,\rangle = \left[(l^2 + l) - (l_z^{\min})^2 + l_z^{\min}\right] |\, l, l_z^{\min}\,\rangle,$$

which implies

$$l_z^{\min^2} - l_z^{\min} = l^2 + l.$$

One concludes that

$$l_z^{\min} = -l$$

since the other root $l_z = l+1$ makes no sense. Consequently,

$$\boxed{-l \le m \le l.}$$

We may now compute the normalization coefficients $c^+(m)$ and $c^-(m)$ relating $|\, l, m\,\rangle$ and $|\, l, m \pm 1\,\rangle$ as

$$\hat{l}_\pm \,|\, l, m\,\rangle = c_\pm(m)\,|\, l, m \pm 1\,\rangle$$

$$\Rightarrow \quad \langle\, l, m\,|\, \hat{l}_\mp = \langle\, l, m \pm 1\,|\, c_\pm(m)^*.$$

It follows that

$$\begin{aligned} |c_\pm(m)|^2 &= \langle\, l, m\,|\, \hat{l}_\mp \hat{l}_\pm \,|\, l, m\,\rangle \\ &= \langle\, l, m\,|\, \hat{l}^2 - \hat{l}_z^2 \mp \hat{l}_z \,|\, l\, m\,\rangle \\ &= l(l+1) - m(m \pm 1) \ge 0, \end{aligned}$$

and thus

$$c_\pm(m) = \sqrt{l(l+1) - m(m\pm 1)} = \sqrt{(l \mp m)\,(l \pm m + 1)},$$

where we have chosen the arbitrary phase to be equal to 1. Finally,

$$\hat{l}_+ \,|\, l, m \,\rangle = \sqrt{l(l+1) - m(m+1)} \,|\, l, m+1 \,\rangle$$

and

$$\hat{l}_- \,|\, l, m \,\rangle = \sqrt{l(l+1) - m(m-1)} \,|\, l, m-1 \,\rangle.$$

Note that

$$l_+ \,|\, l, m \,\rangle \neq 0 \quad \forall\, m < l,$$

and

$$l_- \,|\, l, m \,\rangle \neq 0 \quad \forall\, m > -l.$$

One may start from $|\, l, -l \,\rangle$ and apply repeatedly $\hat{l}_+$ to obtain

$$\begin{gathered} |\, l, -l \,\rangle \\ |\, l, -l+1 \,\rangle \\ |\, l, -l+2 \,\rangle \\ \vdots \\ |\, l, -l+k = l \,\rangle. \end{gathered}$$

It is clear that $k = 2\,l \in \mathbb{N}$, since otherwise we would find $|c_+(m)|^2 < 0$. Consequently, $2\,l \in \mathbb{N}$, which implies that l is *integer* or *half-integer*. Moreover, notice that the values of m differ always by 1. For instance, for $l = 1$ we have

$$|\, l, m \,\rangle = \{|\, 1, -1 \,\rangle,\ |\, 1, 0 \,\rangle \text{ and } |\, 1, 1 \,\rangle\},$$

and for $l = 3/2$

$$|\, l, m \,\rangle = \{|\, 3/2, -3/2 \,\rangle,\ |\, 3/2, -1/2 \,\rangle,\ |\, 3/2, 1/2 \,\rangle \text{ and } |\, 3/2, 3/2 \,\rangle\}.$$

We can now summarize the expression for the matrix elements of the operators $\hat{J}^2$, $\hat{J}_z$, $\hat{J}_+$ and $\hat{J}_-$ in the J^2, J_z representation, also known as $|\, j, m \,\rangle$ representation:

$$\begin{aligned} \hat{J}^2 \,|\, j, m \,\rangle &= j\,(j+1)\,\hbar^2 \,|\, j, m \,\rangle \\ \langle\, j\, m' \,|\, \hat{J}^2 \,|\, j, m \,\rangle &= j\,(j+1)\,\hbar^2\, \delta_{mm'} \\ \\ \hat{J}_z \,|\, j, m \,\rangle &= m\,\hbar \,|\, j, m \,\rangle \end{aligned}$$

$$\langle\, j, m' \,|\, \hat{J}_z \,|\, j, m \,\rangle = m\, \hbar\, \delta_{mm'}$$

$$\begin{aligned}\langle\, j, m' \,|\, \hat{J}_+ \,|\, j, m \,\rangle &= \sqrt{j\,(j+1) - m\,(m+1)}\; \hbar\, \delta_{m',\,m+1} \\ &= \sqrt{(j-m)\,(j+m+1)}\; \hbar\, \delta_{m',\,m+1}\end{aligned}$$

$$\begin{aligned}\langle\, j, m' \,|\, \hat{J}_- \,|\, j, m \,\rangle &= \sqrt{j\,(j+1) - m\,(m-1)}\; \hbar\, \delta_{m',\,m-1} \\ &= \sqrt{(j+m)\,(j-m+1)}\; \hbar\, \delta_{m',\,m-1}.\end{aligned}$$

The $2j+1$ states $\{|\, j, m \,\rangle\} = \{|\, j, -j \,\rangle, |\, j, -j+1 \,\rangle, \ldots, |\, j, j-1 \,\rangle, |\, j, j \,\rangle\}$ span a subspace $\mathbb{S}$ which is *invariant* upon rotations, i.e., any state within $\mathbb{S}$ remains within $\mathbb{S}$ after any rotation in SO(3). This subspace, or the set of $2j+1$ states $|\, j, m \,\rangle$ that span it, are said to be *irreducible*, since no linear subspace included in $\mathbb{S}$ with a lower dimension than $\mathbb{S}$ can be found, which is invariant upon all rotations in SO(3). The latter can be clearly seen by noting that all basis states $|\, j, m \,\rangle$ can be obtained starting from any of them and applying $\hat{J}_+$ and $\hat{J}_-$, which represent linear combinations of infinitesimal rotations around the axis x and y.

Examples:

1) The case $j = 1/2$ is relevant for the electronic spin and therefore deserves special attention. We consider the basis $\{|\, 1/2, 1/2 \,\rangle = |\, z, \uparrow \,\rangle = |\, z, + \,\rangle = \binom{1}{0}$ and $|\, 1/2, -1/2 \,\rangle = |\, z, \downarrow \,\rangle = |\, z, - \,\rangle = \binom{0}{1}\}$.

$$\hat{J}^2 = \frac{3}{4}\,\hbar^2 = j\,(j+1)\,\hbar \qquad \hat{J}_z = \frac{\hbar}{2}\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}$$

$$\hat{J}_+ = \hbar\begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix} \qquad \hat{J}_- = \hbar\begin{pmatrix}0 & 0\\ 1 & 0\end{pmatrix}$$

$$\hat{J}_x = \frac{1}{2}\,(\hat{J}_+ + \hat{J}_-) = \frac{\hbar}{2}\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix} \qquad \hat{J}_y = \frac{-i}{2}\,(\hat{J}_+ - \hat{J}_-) = \frac{\hbar}{2}\begin{pmatrix}0 & -i\\ i & 0\end{pmatrix}$$

The 2×2 matrices

$$\sigma_x = \begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix},\ \sigma_y = \begin{pmatrix}0 & -i\\ i & 0\end{pmatrix} \text{ and } \sigma_z = \begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}$$

are known as Pauli matrices. They have a few quite interesting properties:

$$\sigma_x^2 = \sigma_y^2 = \sigma_z^2 = 1,$$

and

$$\sigma_i\, \sigma_j = i\, \varepsilon_{ijk}\, \sigma_k.$$

Writing the spin operator in vector form as $\hat{\vec{S}} = \left(\hat{S}_x, \hat{S}_y, \hat{S}_z\right)$ and introducing $\vec{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$ we have

$$\hat{\vec{S}} = \frac{\hbar}{2}\,\vec{\sigma}.$$

2) The spin-1 case: For $J = 1$ we have $|\,j, m\,\rangle = \{|\,1, 1\,\rangle,\ |\,1, 0\,\rangle,\ |\,1, -1\,\rangle\}$ and $J^2 = j\,(j+1)\,\hbar^2 = 2\,\hbar^2$. The explicit form of the operators is the following:

$$\hat{J}^2 = 2\,\hbar^2 \begin{pmatrix} 1 & & \\ & 1 & \\ & & 1 \end{pmatrix} \qquad \hat{J}_z = \hbar \begin{pmatrix} 1 & & \\ & 0 & \\ & & -1 \end{pmatrix}$$

$$\hat{J}_+ = \hbar \begin{pmatrix} 0 & \sqrt{2} & 0 \\ 0 & 0 & \sqrt{2} \\ 0 & 0 & 0 \end{pmatrix} \qquad \hat{J}_- = \hbar \begin{pmatrix} 0 & 0 & 0 \\ \sqrt{2} & 0 & 0 \\ 0 & \sqrt{2} & 0 \end{pmatrix}$$

$$\hat{J}_x = \frac{J_+ + J_-}{2} = \frac{\hbar}{2} \begin{pmatrix} 0 & \sqrt{2} & \\ \sqrt{2} & 0 & \sqrt{2} \\ & \sqrt{2} & 0 \end{pmatrix} \qquad \hat{J}_y = \frac{\hbar}{2\,i} \begin{pmatrix} 0 & \sqrt{2} & \\ -\sqrt{2} & 0 & \sqrt{2} \\ & -\sqrt{2} & 0 \end{pmatrix}.$$

6.5. Angular momentum operator in spherical coordinates

In order to calculate the orbital angular momentum operator in spherical coordinates we need to find the corresponding expression for the nabla operator. Consider a function $\Psi(\vec{r})$ and its gradient

$$\vec{\nabla}\Psi = (\nabla\Psi)_r\,\hat{e}_r + (\nabla\Psi)_\theta\,\hat{e}_\theta + (\nabla\Psi)_\varphi\,\hat{e}_\varphi \tag{6.21}$$

expressed in terms of its spherical components. For any displacement vector

$$d\vec{r} = dr\ \hat{e}_r + r\,d\theta\ \hat{e}_\theta + r\ \sin\theta\ d\varphi\ \hat{e}_\varphi$$

the change in $\Psi(\vec{r})$ is given by

$$d\Psi = \vec{\nabla}\Psi \cdot d\vec{r} = (\nabla\Psi)_r\ dr + (\nabla\Psi)_\theta\ r\,d\theta + (\nabla\Psi)_\varphi\ r\ \sin\theta\ d\varphi. \tag{6.22}$$

On the other side the change $\vec{\nabla}\Psi \cdot d\vec{r}$ in the function $\Psi(\vec{r}) = \Psi(r, \theta, \varphi)$ can be expressed in terms of its partial derivatives as

$$d\Psi = \vec{\nabla}\Psi \cdot d\vec{r} = \frac{\partial\Psi}{\partial r}\ dr + \frac{\partial\Psi}{\partial\theta}\ d\theta + \frac{\partial\Psi}{\partial\varphi}\ d\varphi. \tag{6.23}$$

Comparing Eqs. (6.22) and (6.23) and replacing in Eq. (6.21) it follows that

$$\vec{\nabla} = \frac{\partial}{\partial r}\ \hat{e}_r + \frac{1}{r}\ \frac{\partial}{\partial\theta}\ \hat{e}_\theta + \frac{1}{r\ \sin\theta}\ \frac{\partial}{\partial\varphi}\ \hat{e}_\varphi.$$

The orbital angular momentum operator can then be obtained as

$$\hat{\vec{L}} = -i\,\hbar\,(\vec{r}\times\vec{\nabla}) = -i\,\hbar \begin{vmatrix} \hat{e}_r & \hat{e}_\theta & \hat{e}_\varphi \\ r & 0 & 0 \\ \frac{\partial}{\partial r} & \frac{1}{r}\,\frac{\partial}{\partial\theta} & \frac{1}{r\,\sin\theta}\,\frac{\partial}{\partial\varphi} \end{vmatrix}$$

$$= -i\,\hbar \left[\hat{e}_\varphi\,\frac{\partial}{\partial\theta} - \hat{e}_\theta\,\frac{1}{\sin\theta}\,\frac{\partial}{\partial\varphi} \right].$$

Notice that $\hat{\vec{L}}$ has no component along the radial direction $\hat{e}_r$. Moreover, $\hat{\vec{L}}$ acts only on the angular variables θ and ϕ. In order to obtain the expressions for $\hat{L}_x$, $\hat{L}_y$ and $\hat{L}_z$ we express $\hat{e}_\varphi$ and $\hat{e}_\theta$ in Carthesian coordinates as

$$\hat{e}_\varphi = (-\sin\varphi, \cos\varphi, 0)$$

and

$$\hat{e}_\theta = (\cos\theta\,\cos\varphi, \cos\theta\,\sin\varphi, -\sin\theta)$$

to obtain

$$\hat{L}_z = -i\,\hbar\,\frac{\partial}{\partial\varphi}$$

and

$$\hat{L}_\pm = \hbar\,e^{\pm i\,\varphi} \left(\pm\frac{\partial}{\partial\theta} + i\,\cot\theta\,\frac{\partial}{\partial\varphi} \right).$$

A straightforward calculation yields

$$\hat{L}_+\,\hat{L}_- = \hbar^2 \left[-\frac{\partial}{\partial\theta^2} - i\,\frac{\partial}{\partial\varphi} - \cot\theta\,\frac{\partial}{\partial\theta} - \cot^2\theta\,\frac{\partial^2}{\partial\varphi^2} \right].$$

Using that $\hat{L}^2 = \hat{L}_+\,\hat{L}_- - \hbar\,\hat{L}_z + \hat{L}_z^2$ one obtains

$$\hat{L}^{\,2} = -\frac{\hbar^2}{\sin^2\theta} \left[\sin\theta\,\frac{\partial}{\partial\theta}\left(\sin\theta\,\frac{\partial}{\partial\theta}\right) + \frac{\partial^2}{\partial\varphi^2} \right].$$

Notice that $\hat{L}^{\,2}$ depends only on the angular variables θ and φ, i.e., it does not act on the radial variable r. In this context it is useful to relate $\hat{L}^{\,2}$ with the Laplace operator:

$$\nabla^2 = \frac{1}{r^2}\,\frac{\partial}{\partial r}\left(r^2\,\frac{\partial}{\partial r}\right) - \frac{\hat{L}^{\,2}}{r^2\,\hbar^2}.$$

The kinetic energy operator $\hat{T}$ can then be written as

$$\hat{T} = \frac{\hat{p}^2}{2\,m} = \underbrace{-\frac{\hbar^2}{2\,m}\,\frac{1}{r^2}\,\frac{\partial}{\partial r}\left(r^2\,\frac{\partial}{\partial r}\right)}_{\hat{T}_r} + \underbrace{\frac{\hat{L}^{\,2}}{2\,m\,r^2}}_{\hat{T}_{\theta,\varphi}},$$

where $\hat{T}_r$ depends only on r and $\hat{T}_{\theta,\varphi}$ depends only on θ and φ. Taking into account that $\hat{L}^2$ and $\hat{T}_{\theta,\varphi}$ act only on θ and φ, and that the radial part of the kinetic energy $\hat{T}_r$ acts only on r, we can seek for the eigenfunction of $\hat{T}$, and even of the Hamiltonian $\hat{H} = \hat{T} + V(r)$ in the presence of a spherically symmetric potential $V(r)$, by separating radial and angular variables as

$$\Psi_{lm}(\vec{r}) = R(r)\, Y_{lm}(\theta,\varphi),$$

where $Y_{lm}(\theta,\varphi)$ is an eigenfunction of $\hat{L}^2$ and $\hat{L}_z$.

The angular dependent functions $Y_{lm}(\theta,\varphi)$ are known as *spherical harmonics.* They are eigenfunctions of $\hat{L}_z$, and therefore

$$-i\,\hbar\,\frac{\partial}{\partial\varphi}\,Y_{lm}(\theta,\varphi) = m\,\hbar\,Y_{lm}(\theta,\varphi) \quad\Rightarrow\quad Y_{lm}(\theta,\varphi) = f(\theta)\,e^{i\,m\,\varphi},$$

where $f(\theta)$ is a function of θ alone. Replacing $Y_{lm}(\theta,\varphi)$ in the equation

$$\hat{L}^2\,Y_{lm}(\theta,\varphi) = l\,(l+1)\,\hbar\,Y_{lm}(\theta,\varphi)$$

one obtains a differential equation for f of the form

$$\left[\frac{d}{dz}\,(1-z^2)\,\frac{d}{dz} + \left(l\,(l+1) - \frac{m^2}{1-z^2}\right)\right] f(z) = 0,$$

where $z = \cos\theta$. The spherical harmonics thus take the form

$$Y_{lm}(\theta,\varphi) = \sqrt{\frac{2\,l+1}{4\,\pi}\,\frac{(l-m)!}{(l+m)!}}\;P_l^m(\cos\theta)\,e^{i\,m\,\varphi}$$

for $m \geq 0$, where $P_l^m(\cos\theta)$ are the associated Legendre polynomials. Besides an arbitrary choice of phase factor, we have used the normalization condition

$$\int_0^{2\,\pi}\int_{-1}^{1} d\Omega\; Y^*_{l'm'}(\theta,\varphi)\,Y_{lm}(\theta,\varphi) = \delta_{ll'}\,\delta_{mm'}.$$

For $m = 0$, $P_l^m(\cos\theta)$ reduces to the usual Legendre polynomials $P_l(\cos\theta)$. Thus,

$$Y_{l\,0}(\theta,\varphi) = \sqrt{\frac{2\,l+1}{4\,\pi}}\;P_l(\cos\theta).$$

For $m < 0$ we have

$$Y_{l,\,-|m|}(\theta,\varphi) = (-1)^m\,Y^*_{l,\,|m|}(\theta,\varphi).$$

The spherical harmonics satisfy the following important completeness relation:

$$\sum_{l=0}^{+\infty}\sum_{m=-l}^{l} Y^*_{lm}(\theta',\varphi')\,Y_{lm}(\theta,\varphi) = \delta(\varphi-\varphi')\,\delta(\cos\theta - \cos\theta').$$

Consequently, an arbitrary function $\Psi(r, \theta, \varphi)$ can be written in the form

$$\Psi(r, \theta, \varphi) = \sum_{l=0}^{+\infty} \sum_{m=-l}^{l} R_{lm}(r)\, Y_{lm}(\theta, \varphi).$$

Examples:

i) $l = 0$

$$Y_{00}(\theta, \varphi) = \frac{1}{\sqrt{4\,\pi}} \qquad \leftarrow \text{independent of } \theta \text{ and } \varphi.$$

ii) $l = 1$

$$Y_{10}(\theta, \varphi) = \sqrt{\frac{3}{4\,\pi}}\, \cos\theta \qquad \leftarrow \text{independent of } \varphi.$$

$$\begin{aligned} Y_{1;\pm1}(\theta, \varphi) &= \mp\sqrt{\frac{3}{8\,\pi}}\, \sin\theta\, e^{\pm i\,\varphi} \\ &= \mp\sqrt{\frac{3}{8\,\pi}}\, \sin\theta\,(\cos\varphi \pm i\, \sin\varphi) \end{aligned}$$

Notice that $Y_{10} \propto \cos\theta \propto z$ corresponds to the atomic orbital p_z and that $Y_{1;\pm1} \propto (x \pm i\,y)$ corresponds to the atomic orbital $p_x \pm i\,p_y$. The eigenfunctions of $\hat{L}^2$ and $\hat{L}_z$ are classified according to the value of l as

l	0	1	2	3	
	s	p	d	f	orbitals.

The spherical harmonics have well-defined inversion symmetry, i.e., symmetry with respect to the transformation $\vec{r} \to -\vec{r}$, or equivalently $(r, \theta, \varphi) \to (r, \pi - \theta, \varphi + \pi)$. In fact, upon inversion we have

$$Y_{lm}(\pi - \theta, \varphi + \pi) = (-1)^l\, Y_{lm}(\theta, \varphi).$$

This implies even (odd) parity for l even (odd).

An alternative derivation of the eigenfunctions of L^2 and L_z would be to apply our knowledge of the operator algebra. Starting from

$$L_+\, |\, l\, l\, \rangle = 0,$$

and using that

$$\hat{L}_+ = \hbar\, e^{i\,\varphi} \left(\frac{\partial}{\partial\theta} + i\, \cot\theta\, \frac{\partial}{\partial\varphi} \right)$$

as well as

$$\langle\, \vec{r}\, |\, l\, l\, \rangle = Y_{l\,l}(\theta, \varphi) = e^{i\,l\,\varphi}\, f(\theta),$$

we have

$$0 = e^{i\,l\,\varphi}\,\frac{\partial f}{\partial \theta} + i\,\frac{\cos\theta}{\sin\theta}\,(i\,l)\;e^{i\,l\,\varphi}\,f(\theta),$$

which implies

$$\frac{\partial f}{\partial \theta} - l\,\frac{\cos\theta}{\sin\theta}\,f(\theta) = 0.$$

This equation is satisfied by $f(\theta) = c\,\sin^l(\theta)$ so that

$$Y_{l\,l}(\theta,\varphi) = c_l\,\sin^l(\theta)\;e^{i\,l\,\varphi}.$$

Repeated application of $\hat{L}_-$ allows one to obtain all the other $Y_{lm}(\theta,\varphi)$ with $m < l$. The normalization constant is determined by the condition

$$\int d\Omega\;|Y_{l\,l}(\theta,\varphi)|^2 = 1.$$

With an appropriate choice of the phase factor, so that $(\hat{L}^-)^l\;Y_{l\,l}(\theta,\varphi)$ has the same sign as $P_l(\cos\theta)$, one obtains

$$c_l = \frac{(-1)^l}{2^l\,l!}\,\sqrt{\frac{(2\,l+1)\,(2l)!}{4\,\pi}}.$$

Taking into account that

$$\hat{L}^- = \hbar\,e^{-i\,\varphi}\left(-\frac{\partial}{\partial\theta} + i\,\cot\theta\,\frac{\partial}{\partial\varphi}\right)$$

and

$$\hat{L}^-\,|\,l,m\,\rangle = \hbar\,\sqrt{(l+m)\,(l-m+1)}\,|\,l,m-1\,\rangle$$

we finally arrive to the recursive relation

$$Y_{l\,m-1}(\theta,\varphi) = \frac{1}{\sqrt{(l+m)\,(l-m+1)}}\;e^{-i\,\varphi}\left(-\frac{\partial}{\partial\theta} + i\,\cot\theta\,\frac{\partial}{\partial\varphi}\right)Y_{lm}(\theta,\varphi).$$

7. Spin

Consider a system consisting of several parts such as a nucleus. We would like to describe the quantum mechanical state of the nucleus as a whole for a given fixed internal energy E. It is clear that the position (or momentum) of the nucleus is not enough to describe its QM state, since we must take into account that the nucleus can have different internal states. Taking into account the rotational invariance of isolated systems we can attempt to describe the state of the nucleus by its total angular momentum L. It often happens that all the states of a given energy E have the same total angular momentum L. However, even in this case, we must also take into account the possible orientations of $\vec{L}$ in space (unless $L = 0$). These have all the same E and are characterized by the $2\,L + 1$ possible projections of $\vec{L}$ along a given axis, say, the z axis. Therefore, in order to describe the state of the nucleus, we need to consider not only its coordinates $\vec{r}$ (or momentum $\vec{p}$), but also an additional discrete variable, namely the projection of the angular momentum along a given direction in space. The state of a composite particle can then be described by the wave function

$$\Psi_{nL} = \Psi_{nL}(\vec{r}, L_z).$$

There is no reason why elementary particles, that is, particles that appear to be structureless in the considered energy range of the interactions,[11] should not have such an internal degree of freedom. In fact there are many examples in nature where a particle “at rest” or in a state bearing no orbital angular momentum (e.g., a spherically symmetric state $\Psi(\vec{r}) = \Psi(r)$ or in a plane wave state $\Psi(\vec{r}) = \frac{1}{\sqrt{V}}\, e^{i\,\vec{p}\,\vec{r}}$) shows a finite (non-vanishing) *intrinsic angular momentum* $j_0 \neq 0$. Moreover, the value of j_0 is very often half-integer.

The intrinsic angular momentum of an elementary particle is called *spin* and the vector operator associated to this observable is denoted by $\hat{\vec{s}} = (\hat{s}_x, \hat{s}_y, \hat{s}_z)$. The commutation rules for $\hat{s}_x$, $\hat{s}_y$ and $\hat{s}_z$ follow from the commutation rules of the rotations (matrices) in $\mathbb{R}^3$ and are the same as those of the orbital angular momentum $\vec{l} = \vec{r} \times \vec{p}$, since $\hat{s}_x$, $\hat{s}_y$ and $\hat{s}_z$ are the generators of infinitesimal rotations (around the axis x, y and z) acting on the spin variables. Therefore

$$[\hat{s}_x, \hat{s}_y] = i\,\hbar\,\hat{s}_z, \qquad [\hat{s}_y, \hat{s}_z] = i\,\hbar\,\hat{s}_x, \quad \text{and} \quad [\hat{s}_z, \hat{s}_x] = i\,\hbar\,\hat{s}_y.$$

All the properties derived for the angular momentum $\vec{j}$ hold of course for $\vec{s}$. In particular $\hat{s}^2$ and $\hat{s}_z$ form a complete set of compatible observables. The corresponding eigenvalues are $s(s+1)$ and $m_s = -s, -s+1, \ldots, s-1, s$. Since the length of the spin is fixed by the nature of the particle considered (e.g., $s = 1/2$ for an electron), the angular momentum $\hbar\,\vec{s}$ vanishes in the limit $\hbar \to 0$. Thus, the spin has no classical equivalent.

For the description of the state of a particle we not only need to consider the usual classical variables such as position $\vec{r}$ or momentum $\vec{p}$, but also the spin variable m_s, also denoted by σ, which characterizes the intrinsic state. The length s of the spin is fixed, like the mass or the charge of the given particle, and therefore does not depend on the

---

[11] The fact that the notion of *elementary particle* is *energy dependent* shall be discussed below.

particle state. In three dimensions four compatible observables are needed to describe the state of a particle having $s \neq 0$. These can be written as $|\,\vec{r}, m_s\,\rangle$ or $|\,\vec{p}, m_s\,\rangle$, for example. One may also define the state by the observables $\hat{H}$, $\hat{l}^{\,2}$, $\hat{l}_z$ and $\hat{s}_z$ or $\hat{H}$, $\hat{l}^{\,2}$, $\hat{\jmath}^{\,2}$ and $\hat{\jmath}_z$, where $\vec{\jmath} = \vec{l} + \vec{s}$ is the total angular momentum. In this case one denotes the states as $|\,E, l, m_l, m_s\,\rangle$ and $|\,E, l, j, j_z\,\rangle$, respectively. The wave function

$$\langle\,\vec{r}, \sigma\,|\,\Psi\,\rangle = \Psi(\vec{r}, \sigma) = \Psi_\sigma(\vec{r})$$

is thus a function of four variables (3 continuous and 1 discrete). One often attaches an index to the wave function corresponding to the different values of $m_s = \sigma$. These are known as the spin components of the wave function. Notice that the spin variables are completely independent of the usual classical variables $\vec{r}$ or $\vec{p}$. One can therefore regard the kets of a particle with spin as the elements of a Hilbert space, which is the direct product of the usual Hilbert space associated to the classical variables and $(2\,s + 1)$-dimensional space associated to the spin variable. For example,

$$|\,\vec{p}, m_s\,\rangle = |\,\vec{p}\,\rangle \cdot |\,m_s\,\rangle$$

or

$$|\,E, l, m_l, m_s\,\rangle = |\,E, l, m_l\,\rangle\;|\,m_s\,\rangle.$$

The spin operators act on the spin variable only and therefore commute with the operators acting on the coordinates such as $\vec{r}$, $\vec{p}$, $\vec{l}$, etc. If one rotates a particle with spin, one must take into account the effect of the rotation on both the coordinates and the spin variable. The operator describing an infinitesimal rotation of the orbital part of the ket is given by

$$\hat{U}_{\text{orb}}(\delta\phi) = 1 - \frac{i}{\hbar}\,\hat{n} \cdot \hat{\vec{l}}\,\delta\phi + O(\delta\phi^2),$$

where $\delta\phi$ denotes the angle and $\hat{n}$ the axis of rotation. The rotation of the spin variables is realized by the operator

$$\hat{U}_{\text{spin}}(\delta\phi) = 1 - \frac{i}{\hbar}\,\hat{n} \cdot \hat{\vec{s}}\,\delta\phi + O(\delta\phi^2),$$

where $\hat{\vec{s}}$ is the spin operator. The operator corresponding to the rotation of the complete ket is therefore given by the product

$$\begin{aligned}\hat{U} &= \hat{U}_{\text{spin}}\,\hat{U}_{\text{orb}} = \hat{U}_{\text{orb}}\,\hat{U}_{\text{spin}}\\ &= 1 - \frac{i}{\hbar}\,\hat{n} \cdot (\hat{\vec{l}} + \hat{\vec{s}})\,\delta\phi + O(\delta\phi^2).\end{aligned}$$

Consequently, the total angular momentum

$$\vec{\jmath} = \vec{l} + \vec{s}$$

is the generator of infinitesimal rotations. It is easy to verify that

$$[\hat{\jmath}_i, \hat{\jmath}_k] = i\,\varepsilon_{ikl}\,\hat{\jmath}_l,$$

from which it follows that $\hat{\jmath}^2$ and $\hat{\jmath}_z$ form a complete set of observables with eigenvalues $j(j+1)$ and $m = -j, -j+1, \ldots, j-1, j$. It is therefore important to determine which are the eigenvalues of $\hat{\jmath}^2$, i.e., the possible values of j, that can be obtained for given values of l and s.

For a system of particles the total angular momentum is given by the sum of the operators acting on each variable:

$$\vec{J} = \sum_i \vec{\jmath}_i \quad \text{with} \quad \vec{\jmath}_i = \vec{l}_i + \vec{s}_i,$$

which we can put in the form

$$\vec{J} = \vec{L} + \vec{S}$$

with $\vec{L} = \sum_i \vec{l}_i$ and $\vec{S} = \sum_i \vec{s}_i$.

Example: The electron spin and spinor states

Electrons are particles having $s = 1/2$. The dimension of the spin space is therefore $2\,s + 1 = 2$. We can write the wave function as $\Psi_\sigma(\vec{r})$ with $\sigma = 1/2$ and $-1/2$ or as a two-component or *spin-1/2 spinor*:

$$|\,\Psi\,\rangle = \begin{pmatrix} \Psi_{1/2}(\vec{r}) \\ \Psi_{-1/2}(\vec{r}) \end{pmatrix} = \begin{pmatrix} \Psi_{\uparrow}(\vec{r}) \\ \Psi_{\downarrow}(\vec{r}) \end{pmatrix} = \begin{pmatrix} \Psi_{+}(\vec{r}) \\ \Psi_{-}(\vec{r}) \end{pmatrix}.$$

For the discussion of the spin dependence one can drop the dependence on the classical variables $\vec{r}$ or $\vec{p}$ (e.g., one may consider a given position $\vec{r}$ or imagine that the electron has a defined momentum $\vec{p}$) and simply write

$$|\,\Psi\,\rangle = \begin{pmatrix} c_+ \\ c_- \end{pmatrix} \quad \text{with} \quad |c_+|^2 + |c_-|^2 = 1.$$

The eigenstates of

$$\hat{s}_x = \frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

are given by

$$|\,x, +\,\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \text{and} \quad |\,x, -\,\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix},$$

which correspond to the eigenvalues $s_x = \pm\,\hbar/2$. The ket $|\,x, +\,\rangle$ ($|\,x, -\,\rangle$) represents a state in which the spin of the electron points along the positive (negative) x direction. In terms of the eigenstates of $\hat{s}_z$, namely

$$|\,z, +\,\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{and} \quad |\,z, -\,\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

we can write

$$|\,x,\pm\,\rangle = \frac{1}{\sqrt{2}}\left(|\,z,+\,\rangle \pm |\,z,-\,\rangle\right).$$

And the corresponding components of the wave functions are

$$\Psi^x_\pm(\vec{r}) = \langle\,\vec{r}\,|\,x,\pm\,\rangle = \frac{1}{\sqrt{2}}\left(\langle\,\vec{r}\,|\,z,+\,\rangle \pm \langle\,\vec{r}\,|\,z,-\,\rangle\right) = \frac{1}{\sqrt{2}}\left[\Psi^z_{1/2}(\vec{r}) \pm \Psi^z_{-1/2}(\vec{r})\right].$$

Let us finally notice that if one prepares an electron in a state $|\,x,+\,\rangle$ having its spin pointing along the positive x direction (e.g., as a result of a Stern-Gerlach filtering experiment with the magnetic field gradient along the x direction) and one measures the projection of the spin along a perpendicular direction (e.g., by means of a second Stern-Gerlach experiment with the magnetic field gradient along the z direction) the probability of measuring the value $s_z = +\,\hbar/2$ or $s_z = -\,\hbar/2$ is $|\langle\,z,\pm\,|\,x,+\,\rangle|^2 = 1/2$.

The addition of spin and orbital angular momenta is discussed in the following section.

7.1. Addition of angular momentum

We consider two commuting angular momentum operators $\hat{\vec{J}}_1$ and $\hat{\vec{J}}_2$ acting on different subspaces $\mathcal{S}_1$ and $\mathcal{S}_2$. These can be the spin $\hat{\vec{S}}$ and orbital momentum $\hat{\vec{L}}$ of an electron, for which we have

$$[\hat{\vec{L}},\hat{\vec{S}}\,] = 0,$$

since they act on different variables, or the total angular momenta $\hat{\vec{\jmath}}_1$ and $\hat{\vec{\jmath}}_2$ of two electrons, which satisfy

$$[\hat{\vec{\jmath}}_1,\hat{\vec{\jmath}}_2] = 0$$

for the same reason. The usual commutation relations in component form read

$$[\hat{J}_{1\,k},\hat{J}_{1\,l}] = i\,\hbar\,\varepsilon_{klm}\,\hat{J}_{1\,m}$$

and

$$[\hat{J}_{2\,k},\hat{J}_{2\,l}] = i\,\hbar\,\varepsilon_{klm}\,\hat{J}_{2\,m}$$

with

$$[\hat{J}_{1\,k},\hat{J}_{2\,l}] = 0$$

for all $k,l,m = 1$–3. Since $\hat{\vec{J}}_1$ and $\hat{\vec{J}}_2$ are angular momentum operators, and choosing as usual z as the quantization direction, we have

$$[\hat{J}_{1\,z},\hat{J}_1^2] = 0$$

and

$$[\hat{J}_{2\,z}, \hat{J}_2^2] = 0.$$

Recalling that $[\hat{\vec{J}}_1, \hat{\vec{J}}_2] = 0$, it is clear that $\hat{J}_{1\,z}$, $\hat{J}_1^2$, $\hat{J}_{2\,z}$ and $\hat{J}_2^2$ are compatible observables. Therefore, we can characterize our kets by the corresponding quantum numbers as

$$|\, j_1, j_2;\, m_1, m_2\,\rangle \equiv \underbrace{|\, j_1\, m_1\,\rangle}_{\phi \text{ orbital}}\ \underbrace{|\, j_2\, m_2\,\rangle}_{\chi \text{ spin}}.$$

As usual $\hat{J}_{\alpha\,z}\,|\, j_\alpha\, m_\alpha\,\rangle = m_\alpha\,|\, j_\alpha\, m_\alpha\,\rangle$ and $\hat{J}_\alpha^2\,|\, j_\alpha\, m_\alpha\,\rangle = j_\alpha\,(j_\alpha + 1)\,|\, j_\alpha\, m_\alpha\,\rangle$ for $\alpha = 1$ and 2. If we now perform a rotation R on the system we must rotate or *transform* the orbital ket and the spin ket, or the coordinates of both electrons, if we are dealing with a two-electron system. The corresponding operators are given by

$$\hat{D}_1(R) = e^{-\frac{i}{\hbar}\,\hat{\vec{J}}_1\cdot\hat{n}\,\phi}$$

and

$$\hat{D}_2(R) = e^{-\frac{i}{\hbar}\,\hat{\vec{J}}_2\cdot\hat{n}\,\phi}.$$

The operator transforming the complete ket is

$$\hat{D}(R) = \hat{D}_1(R)\;\hat{D}_2(R) = \hat{D}_2(R)\;\hat{D}_1(R),$$

where it is understood that D_1 acts on the variables of subspace $\mathcal{S}_1$ and D_2 on the variables of $\mathcal{S}_2$. In terms of $\hat{\vec{J}}_1$ and $\hat{\vec{J}}_2$ we have

$$\hat{D}(R) = e^{-\frac{i}{\hbar}\,\hat{\vec{J}}_1\cdot\hat{n}\,\phi}\; e^{-\frac{i}{\hbar}\,\hat{\vec{J}}_2\cdot\hat{n}\,\phi} = e^{-\frac{i}{\hbar}\,(\hat{\vec{J}}_1+\hat{\vec{J}}_2)\cdot\hat{n}\,\phi}.$$

The generator of rotations of the composite system $\mathcal{S}_1 \oplus \mathcal{S}_2$ is therefore

$$\hat{\vec{J}} = \hat{\vec{J}}_1 + \hat{\vec{J}}_2,$$

i.e., the vector sum of the individual operators. The fundamental property of $\hat{\vec{J}} = \hat{\vec{J}}_1 + \hat{\vec{J}}_2$ reads

$$[\hat{J}_k, \hat{J}_l] = i\,\hbar\,\varepsilon_{klm}\,\hat{J}_m, \tag{7.1}$$

i.e., $\hat{\vec{J}}$ is an angular momentum operator. The usual properties of angular momenta follow:

i) $[\hat{J}^2, \hat{J}_z] = [\hat{J}^2, \hat{J}_y] = [\hat{J}^2, \hat{J}_x] = 0.$

ii) $[\hat{J}_z, \hat{J}^{\pm}] = \pm\hbar\,\hat{J}^{\pm}$.

iii) One finds the same kind of eigenstates $|\,j\,m\,\rangle$ with $-j \leq m \leq j$.

iv) The values of j are either integer or half-integer, and the values of m are $m = -j, -j+1, \ldots, j-1, j$.

v) The ladder operators $\hat{J}^{\pm} = \hat{J}_x \pm i\,\hat{J}_y$ satisfy $\hat{J}^{\pm}\,|\,j\,m\,\rangle = \sqrt{j\,(j+1) - m\,(m\pm 1)}\,|\,j, m\pm 1\,\rangle$.

It is therefore important to construct the states $|\,j\,m\,\rangle$ with defined $\hat{J}^2\,|\,j\,m\,\rangle = j(j+1)\,|\,j\,m\,\rangle$ and $\hat{J}_z\,|\,j\,m\,\rangle = m\,|\,j\,m\,\rangle$ in terms of the states $|\,j_1, j_2;\, m_1, m_2\,\rangle = |\,j_1, m_1\,\rangle\,|\,j_2, m_2\,\rangle$ of $\mathcal{S}_1 \oplus \mathcal{S}_2$.

The proof of Eq. (7.1) is straightforward:

$$
\begin{aligned}
[\hat{J}_k, \hat{J}_l] &= [\hat{J}_{1k} + \hat{J}_{2k},\ \hat{J}_{1l} + \hat{J}_{2l}] \\
&= [\hat{J}_{1k}, \hat{J}_{1l}] + [\hat{J}_{2k}, \hat{J}_{2l}] \\
&= i\,\hbar\,\varepsilon_{klm}\,\hat{J}_{1m} + i\,\hbar\,\varepsilon_{klm}\,\hat{J}_{2m} \\
&= i\,\hbar\,\varepsilon_{klm}\,\hat{J}_m.
\end{aligned}
$$

In addition we have

$$
[\hat{J}_k, \hat{J}_{1l}] = [\hat{J}_{1k}, \hat{J}_{2l}] = i\,\hbar\,\varepsilon_{klm}\,\hat{J}_{1m}.
$$

This implies that $\hat{\vec{J}}_1$ and $\hat{\vec{J}}_2$ are *vector operators*. Therefore $[\hat{J}_1^2, \hat{\vec{J}}] = [\hat{J}_2^2, \hat{\vec{J}}] = 0$. We can prove this statement for vector operators in general:

$$
\begin{aligned}
[\hat{V}^2, \hat{J}_z] &= [\hat{V}_x^2 + \hat{V}_y^2 + \hat{V}_z^2, \hat{J}_z] \\
&= \hat{V}_x\,[\hat{V}_x, \hat{J}_z] + [\hat{V}_x, \hat{J}_z]\,\hat{V}_x + \hat{V}_y\,[\hat{V}_y, \hat{J}_z] + [\hat{V}_y, \hat{J}_z]\,\hat{V}_y \\
&= \hat{V}_x\,(-i\,\hbar\,\hat{V}_y) + (-i\,\hbar\,\hat{V}_y)\,\hat{V}_x + \hat{V}_y\,(i\,\hbar\,\hat{V}_x) + (i\,\hbar\,\hat{V}_x)\,\hat{V}_y \\
&= 0.
\end{aligned}
$$

In conclusion, besides $\hat{J}_{1z}$, $\hat{J}_1^2$, $\hat{J}_{2z}$ and $\hat{J}_2^2$, we also have the following set of compatible operators $\hat{J}_1^2$, $\hat{J}_2^2$, $\hat{J}^2$ and $\hat{J}_z$. The states of the composite system $\mathcal{S}_1 \oplus \mathcal{S}_2$ can be characterized by the corresponding quantum numbers as $|\,j_1, j_2;\, j, m\,\rangle$, where

$$
\begin{aligned}
\hat{J}_1^2\,|\,j_1, j_2;\, j, m\,\rangle &= j_1\,(j_1+1)\,|\,j_1, j_2;\, j, m\,\rangle \\
\hat{J}_2^2\,|\,j_1, j_2;\, j, m\,\rangle &= j_2\,(j_2+1)\,|\,j_1, j_2;\, j, m\,\rangle
\end{aligned}
$$

$$\hat{J}^2 \,|\, j_1, j_2;\, j, m \,\rangle = j\,(j+1)\,|\, j_1, j_2;\, j, m \,\rangle$$

$$\hat{J}_z \,|\, j_1, j_2;\, j, m \,\rangle = m \,|\, j_1, j_2;\, j, m \,\rangle.$$

What are the possible values of j in terms of j_1 and j_2? How are the states $|\, j_1, j_2;\, j, m \,\rangle$, which are eigenstates of $\hat{J}_1^2$, $\hat{J}_2^2$, $\hat{J}^2$ and $\hat{J}_z$, written in terms of the states $|\, j_1, j_2;\, m_1, m_2 \,\rangle$, which are eigenstates of $\hat{J}_1^2$, $\hat{J}_2^2$, $\hat{J}_{1z}$ and $\hat{J}_{2z}$? In general we can have

$$|\, j_1, j_2;\, j, m \,\rangle = \sum_{m_1=-j_1}^{j_1} \sum_{m_2=-j_2}^{j_2} |\, j_1, j_2;\, m_1, m_2 \,\rangle \,\langle\, j_1, j_2;\, m_1, m_2 \,|\, j_1, j_2;\, j, m \,\rangle.$$

$\langle\, j_1, j_2;\, m_1, m_2 \,|\, j_1, j_2;\, j, m \,\rangle$ are the so-called *Clebsch-Gordan* or *vector addition coefficients*. An important property is the conservation of the total z component of the angular momentum:

$$\langle\, j_1, j_2;\, m_1, m_2 \,|\, j_1, j_2;\, j, m \,\rangle = 0$$

unless

$$\boxed{m_1 + m_2 = m.} \tag{7.2}$$

To prove Eq. (7.2) we compute

$$\langle\, j_1, j_2 \,|\, m_1, m_2 \,|\, \hat{J}_z - \hat{J}_{1z} - \hat{J}_{2z} \,|\, j_1, j_2;\, j, m \,\rangle = 0$$

(since $\hat{J}_z = \hat{J}_{1z} + \hat{J}_{2z}$) by acting with $\hat{J}_z$ on the ket and with $\hat{J}_{1z}$ and $\hat{J}_{2z}$ on the bra. This implies

$$(m - m_1 - m_2)\,\langle\, j_1, j_2;\, m_1, m_2 \,|\, j_1, j_2;\, j, m \,\rangle = 0,$$

which proves the statement.

The second important property is

$$\langle\, j_1, j_2;\, m_1, m_2 \,|\, j_1, j_2;\, j, m \,\rangle = 0$$

unless

$$\boxed{|j_1 - j_2| \le j \le j_1 + j_2.} \tag{7.3}$$

This can be demonstrated as follows: Since $\hat{J}_z \,|\, j_1, j_2;\, m_1, m_2 \,\rangle = (m_1 + m_2)\,|\, j_1, j_2;\, m_1, m_2 \,\rangle$, the largest possible value of J_z is $j_1 + j_2$. There is only one state with $J_z = j_1 + j_2$, namely,

$$|\, j_1, j_2,\, j = j_1 + j_2,\, m = j_1 + j_2 \,\rangle = |\, j_1, j_2;\, m_1 = j_1, m_2 = j_2 \,\rangle.$$

Applying $\hat{J}_+ = \hat{J}_{1+} + \hat{J}_{2+}$ we obtain

$$\hat{J}_+ \,|\, j_1, j_2;\; m_1 = j_1, m_2 = j_2 \,\rangle = 0$$

$$\Rightarrow \quad |\, j_1, j_2;\; m_1 = j_1, m_2 = j_2 \,\rangle = |\, j_1, j_2;\; j = j_1 + j_2,\; m = j_1 + j_2 \,\rangle.$$

There are 2 kets for $m = j_1 + j_2 - 1$. One is proportional to $\hat{J}_- \,|\, j = j_1 + j_2,\; m = j_1 + j_2 \,\rangle$, and the other has $j = j_1 + j_2 - 1$. There are 3 kets for $m = j_1 + j_2 - 2$, etc. In general, one finds $k + 1$ kets with $m = j_1 + j_2 - k$ until $j_2 - k = \min\{j_1, j_2\} = j_2$, where we have chosen $j_2 \leq j_1$ by convention. This proves the inequality (7.3).

One can easily verify that the sum of the numbers $2j+1$ of orthogonal states $|\, j_1, j_2;\; j, m \,\rangle$ corresponding to each value of j gives $(2j_1 + 1)\,(2j_2 + 1)$, which is the number of states $|\, j_1, j_2;\; m_1, m_2 \,\rangle = |\, j_1, m_1 \,\rangle \,|\, j_2, m_2 \,\rangle$ that can be formed by combining the two angular momenta (direct product):

$$\sum_{j=|j_1-j_2|}^{j_1+j_2} (2j + 1) = (2j_1 + 1)\,(2j_2 + 1).$$

This means that the dimension of the subspace is not modified by the change of representation (unitary transformation).

The spectrum of $\hat{J}^2$ is then given by

$$j = j_1 + j_2,\; j_1 + j_2 - 1,\; \ldots\; |j_1 - j_2|.$$

Notice that the difference between successive values of j is always 1, even if j can be half integer (e.g., $j = 5/2$, $3/2$ and $1/2$ for $j_1 = 3/2$ and $j_2 = 1$). Moreover, there is only one sequence of $2j + 1$ states having $m = -j,\; -j + 1,\; \ldots\; j$ for each value of j. This is usually summarized by the triangular relation

$$|j_1 - j_2| \leq j \leq j_1 + j_2.$$

Let us finally recall that the actual eigenvalues of $\hat{J}^2$ are as usual $j(j + 1)$, since

$$\hat{J}^2 \,|\, j, m \,\rangle = j(j + 1) \,|\, j, m \,\rangle.$$

Example: $L = 2$ and $S = 1/2$

In order to ilustrate how the different eigenstates of $\hat{J}^2$ are constructed we would like to determine the states with defined total angular momentum $\vec{J} = \vec{L} + \vec{S}$ for an electron ($S = 1/2$) in an atomic orbital ($L = 2$). We denote the eigenstates with defined m_L and m_S by $|\, L, S;\; m_L, m_S \,\rangle$ ($\hat{L}_z \,|\, L, S;\; m_L, m_S \,\rangle = m_L \,|\, L, S;\; m_L, m_S \,\rangle$ with $m_L = \pm 2$, ± 1 and 0, and $\hat{S}_z \,|\, L, S;\; m_L, m_S \,\rangle = m_S \,|\, L, S;\; m_L, m_S \,\rangle$ with $m_S = \pm 1/2$) and the eigenstates of $\hat{J}^2$ and $\hat{J}_z$ by $|\, L, S;\; J, m_J \,\rangle$. There is only one possible state with maximal $m_J = L + S = 5/2$. It is given by

$$|\, 2,\; 1/2;\; J = 5/2,\; m_J = 5/2 \,\rangle = |\, 2,\; 1/2;\; m_L = 2,\; m_S = 1/2 \,\rangle.$$

There are two states with $m_J = 3/2$, one with $m_L = 1$ and $m_S = 1/2$ and the other with $m_L = 2$ and $m_S = -1/2$. Appropriate linear combinations of these two states yield the following eigenstates of $\hat{J}^2$ and $\hat{J}_z$:

$$|\,2,\,1/2;\; J = 5/2,\; m_J = 3/2\,\rangle$$

and

$$|\,2,\,1/2;\; J = 3/2,\; m_J = 3/2\,\rangle.$$

How many states can we find with defined m_L and m_S and $m_L + m_S = m_J = 1/2$? These are

$$|\,2,\,1/2;\; m_L = 0,\; m_S = 1/2\,\rangle$$

and

$$|\,2,\,1/2;\; m_L = -1,\; m_S = 1/2\,\rangle.$$

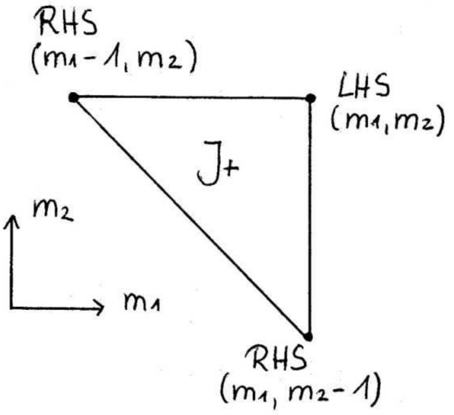
Since application of $\hat{J}_-$ on the eigenstates having $J = 5/2$ and $J = 3/2$ with $m_J = 3/2$ yields two orthogonal states with $m_J = 1/2$, it is clear that no new values of J are obtained for $m_J = 1/2$. The states $|\,2,\,1/2;\; J = 5/2,\; m_J = 1/2\,\rangle$ and $|\,2,\,1/2;\; J = 3/2,\; m_J = 1/2\,\rangle$ are obtained by appropriate linear combination of the states $|\,2,\,1/2;\; m_L = 0,\; m_S = -1/2\,\rangle$ and $|\,2,\,1/2;\; m_L = -1,\; m_S = 1/2\,\rangle$. The same reasoning applies to negative values of $m_J = -1/2$, $-3/2$ and $-5/2$. Therefore, the values of J resulting from the sum of $L = 2$ and $S = 1/2$ are simply $J = 5/2$ and $3/2$, in agreement with the inequality (7.3).

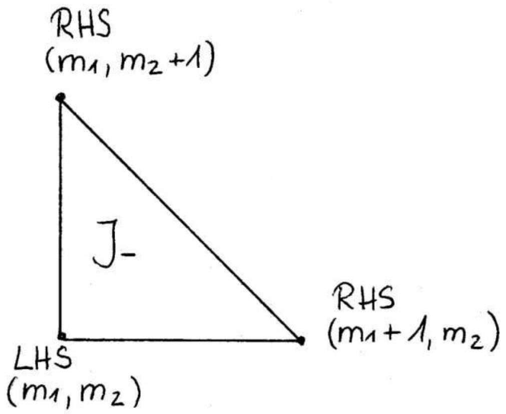
Example: $j_1 = 2$ and $j_2 = 2$ Consider now the sum of two angular momenta $j_1 = 2$ and $j_2 = 2$. The largest value of $J_z = m_J = m_1 + m_2$ is obtained for $m_1 = m_2 = 2$, which corresponds to the state

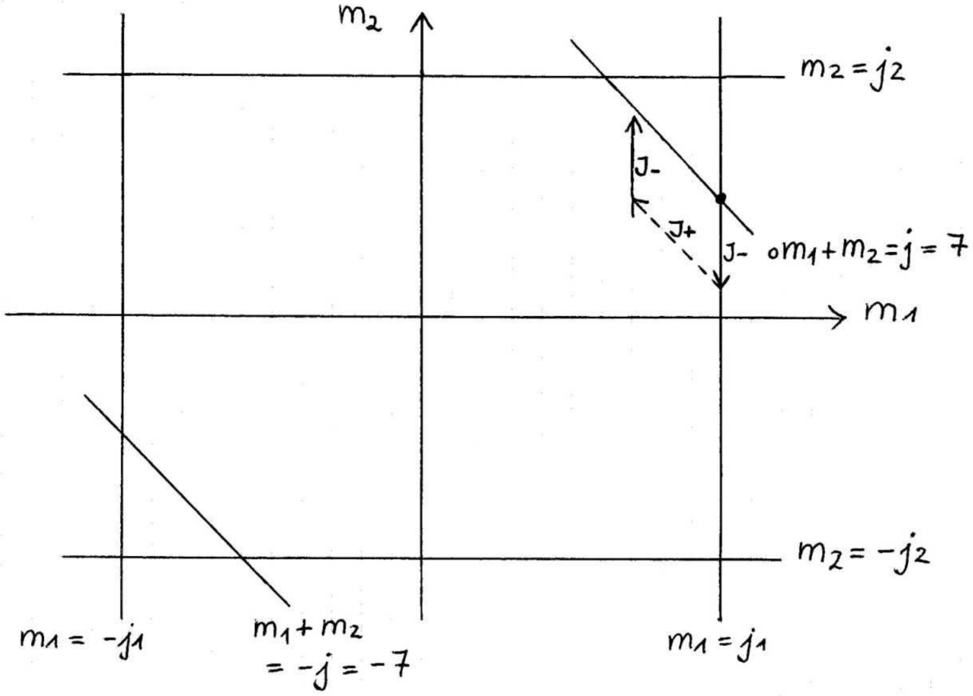
$$|\,2,\,2;\; J = 4,\; m_J = 4\,\rangle = |\,2,\,2;\; m_1 = 2, m_2 = 2\,\rangle.$$

For $J_z = 3$ we find two states $[(m_1, m_2) = (1,\,2)$ and $(m_1, m_2) = (2,\,1)]$, for $J_z = 2$ we find three states $[(m_1, m_2) = (0,\,2)$, $(1,\,1)$ and $(2,\,0)]$, for $J_z = 1$ there are four $[(m_1, m_2) = (-1,\,2)$, $(0,\,1)$, $(1,\,0)$ and $(2,\,-1)]$ and for $J_z = 0$ there are five $[(m_1, m_2) = (-2,\,2)$, $(-1,\,1)$, $(0,\,0)$, $(1,\,-1)$ and $(2,\,-2)]$. In other words, we always find that the dimension of the subspace with a given J_z increases by one each time we decrease J_z $(4 \geq J_z \geq 0)$. Consequently, the values of J obtained from the sum of $j_1 = 2$ and $j_2 = 2$ are $J = 4$, 3, 2, 1 and 0, in agreement with the inequality (7.3).

7.2. The recursion relation for the Clebsch-Gordan coefficients







The purpose of this section is to determine the expansion coefficients of $|\, j_1, j_2;\, j, m\,\rangle$ in terms of $|\, j_1, j_2;\, m_1, m_2\,\rangle$. We compute $\langle\, j_1, j_2;\, m_1, m_2\,|\, \hat{J}_\pm \,|\, j_1, j_2;\, j, m\,\rangle$ in two equivalent ways. First, we operate with $(\hat{J}_\pm)^\dagger = \hat{J}_\mp$ on the bra or left-hand side (LHS) and, second, we operate with $\hat{J}_\pm$ on the ket or right-hand side (RHS):

$$\langle\, j_1, j_2;\, m_1, m_2\,|\, \hat{J}_\pm \,|\, j_1, j_2;\, j, m\,\rangle =$$

$$\begin{aligned}
&\text{LHS} \quad \longrightarrow \quad \sqrt{(j\mp m)\,(j\pm m+1)}\;\langle\, j_1,j_2;\; m_1,m_2\,|\, j_1,j_2;\; j,m{\pm}1\,\rangle =\\
&\text{RHS} \quad \longrightarrow \quad = \sqrt{(j_1\pm m_1)\,(j_1\mp m_1+1)}\;\langle\, j_1,j_2;\; m_1{\mp}1,m_2\,|\, j_1,j_2;\; j,m\,\rangle +\\
&\qquad\qquad\qquad + \sqrt{(j_2\pm m_2)\,(j_2\mp m_2+1)}\;\langle\, j_1,j_2;\; m_1,m_2{\mp}1\,|\, j_1,j_2;\; j,m\,\rangle. \qquad (7.4)
\end{aligned}$$

Notice that m is defined entirely by m_1 and m_2, since we must always have $m_1+m_2 = m\pm 1$. Equation (7.4) defines the Clebsch-Gordan (CG) *recursion relations.* They form a set of triangular relations which can be illustrated by the following figure.

Given one coefficient, for example, for $m_1 = j_1$ and $m_2 + m_1 = j$, we can determine all others without any ambiguity. The value of the first coefficient is fixed by the normalization condition

$$\sum_{\substack{m_1\\ m_2}} |\langle\, j_1 j_2;\; m_1 m_2\,|\, j_1 j_2; jm\,\rangle|^2 = 1$$

and by a sign or phase convention. One can also have complex CG coefficients or, to be more precise, the solutions of the recursion relations (7.4) can be complex. However, in this case the complex conjugate is also always a solution. The CG coefficients are therefore usually chosen to be real.

8. The central symmetric field

8.1. General formulation

We consider the motion of two different interacting particles, for example, an electron moving around a nucleus or an electron-positron pair. Taking into account the isotropy of space, the interaction energy $U(r)$ between the particles can only depend on the distance $r = |\vec{r}_2 - \vec{r}_1|$ between the particles. Denoting by m_1 and m_2 the particle masses, the Hamilton operator of the two-particle system is given by

$$\hat{H} = -\frac{\hbar^2}{2m_1}\nabla_1^2 - \frac{\hbar^2}{2m_2}\nabla_2^2 + U(r).$$

As in classical mechanics it is convenient to change the coordinates by using the relative coordinate

$$\vec{r} = \vec{r}_2 - \vec{r}_1$$

and the center of mass coordinate

$$\vec{R} = \frac{m_1\vec{r}_1 + m_2\vec{r}_2}{m_1 + m_2} \quad \Leftrightarrow \quad M\vec{R} = m_1\vec{r}_1 + m_2\vec{r}_2,$$

where $M = m_1 + m_2$ is the total mass. Since

$$\vec{r}_1 = \vec{R} - \frac{m_2}{M}\vec{r}$$

and

$$\vec{r}_2 = \vec{R} + \frac{m_1}{M}\vec{r},$$

and denoting the vector components as $\vec{r} = (x, y, z)$, $\vec{R} = (X, Y, Z)$, $\vec{r}_1 = (x_1, y_1, z_1)$ and $\vec{r}_2 = (x_2, y_2, z_2)$ we have

$$\begin{aligned}\frac{\partial\Psi}{\partial x_{1,2}} &= \frac{\partial\Psi}{\partial x}\cdot\frac{\partial x}{\partial x_{1,2}} + \frac{\partial\Psi}{\partial X}\cdot\frac{\partial X}{\partial x_{1,2}} \\ &= \mp\frac{\partial\Psi}{\partial x} + \frac{m_{1,2}}{M}\frac{\partial\Psi}{\partial X},\end{aligned}$$

where $-$ $(+)$ corresponds to 1 (2). From this it follows that

$$\frac{\partial^2\Psi}{\partial x_{1,2}} = \mp\left(\mp\frac{\partial^2\Psi}{\partial x^2} + \frac{\partial^2\Psi}{\partial X\,\partial x}\frac{m_{1,2}}{M}\right) + \frac{m_{1,2}}{M}\left(\mp\frac{\partial^2\Psi}{\partial x\,\partial X} + \frac{\partial^2\Psi}{\partial X^2}\frac{m_{1,2}}{M}\right)$$

and therefore

$$\frac{1}{m_1}\frac{\partial^2\Psi}{\partial x_1^2} + \frac{1}{m_2}\frac{\partial^2\Psi}{\partial x_2^2} = \left(\frac{1}{m_1} + \frac{1}{m_2}\right)\frac{\partial^2\Psi}{\partial x^2} + \frac{m_1 + m_2}{M^2}\frac{\partial^2\Psi}{\partial X^2} + \frac{2}{M}\frac{\partial^2\Psi}{\partial X\,\partial x} - \frac{2}{M}\frac{\partial^2\Psi}{\partial x\,\partial X}.$$

The Hamilton operator can then be written in the form

$$\hat{H} = -\frac{\hbar^2}{2M}\,\nabla_R^2 - \frac{\hbar^2}{2\mu}\,\nabla_r^2 + U(r),$$

where ∇_R^2 refers to the Laplacian with respect to the center of mass coordinate $\vec{R}$, ∇_r^2 to the Laplacian with respect to the relative coordinate $\vec{r}$, and $\mu = \dfrac{m_1\, m_2}{m_1 + m_2}$ to the reduced mass. Since the mass of the nucleus m_2 is much larger than the mass of the electron $m_1 = m_e$ we may set $\mu \simeq m_e$.

The separation of variables $\vec{r}$ and $\vec{R}$ in $\hat{H}$ allows us to write its eigenfunctions $\Psi(\vec{r}_1, \vec{r}_2)$ as product states of the form

$$\Psi(\vec{r}, \vec{R}) = \phi(\vec{R})\,\Psi(\vec{r}),$$

where $\phi(\vec{R})$ describes the free-particle motion of the center of mass and $\Psi(\vec{r})$ the relative motion of the particles under the action of the interparticle potential $U(r)$.

We focus on the relative motion and write the Schrödinger equation as

$$-\frac{\hbar^2}{2\mu}\,\nabla^2\,\Psi + U(r)\,\Psi = E\,\Psi.$$

Using the expression for the Laplace operator in spherical coordinates

$$\nabla^2 = \frac{1}{r^2}\,\frac{\partial}{\partial r}\left(r^2\,\frac{\partial}{\partial r}\right) + \frac{1}{r^2}\,\frac{1}{\sin^2\theta}\left[\sin\theta\,\frac{\partial}{\partial\theta}\left(\sin\theta\,\frac{\partial}{\partial\theta}\right) + \frac{\partial^2}{\partial\phi^2}\right]$$

and the expression for the square of the angular momentum (in units of $\hbar$)

$$\hat{l}^2 = -\frac{1}{\sin^2\theta}\left[\sin\theta\,\frac{\partial}{\partial\theta}\left(\sin\theta\,\frac{\partial}{\partial\theta} + \frac{\partial^2}{\partial\phi^2}\right)\right]$$

we can write the Hamiltonian as

$$\hat{H} = -\frac{\hbar^2}{2\mu}\,\frac{1}{r^2}\,\frac{\partial}{\partial r}\left(r^2\,\frac{\partial}{\partial r}\right) + \frac{\hbar^2}{2\mu}\,\frac{\hat{l}^2}{r^2} + U(r). \tag{8.1}$$

If one introduces the radial component of the linear momentum

$$\hat{p}_r\,\Psi = -i\,\hbar\,\frac{1}{r}\,\frac{\partial}{\partial r}\,(r\,\Psi) = -i\,\hbar\left(\frac{\partial\Psi}{\partial r} + \frac{\Psi}{r}\right),$$

one can bring $\hat{H}$ in the form known from classical mechanics:

$$\hat{H} = \frac{\hat{p}_r^2}{2\mu} + \frac{1}{2\mu}\,\frac{\hat{L}^2}{r^2} + U(r), \tag{8.2}$$

where $\vec{L} = \hbar\,\vec{l} = \hat{\vec{r}} \times \hat{p}$.

Recalling that $\hat{\vec{r}}$ and $\hat{\vec{p}}$ are vector operators and that the corresponding squares $\hat{r}^2$ and $\hat{p}^2$ are invariant upon rotation $\left(\text{i.e., } \left[\hat{r}^2, \hat{\vec{l}}\right] = \left[\hat{p}^2, \hat{\vec{l}}\right] = 0\right)$, it is easy to see that the Hamiltonian given by Eq. (8.1) or (8.2) commutes with $\vec{l}$, since the interparticle potential U depends only on r. We may therefore search for the common eigenstates of $\hat{H}$, $\hat{l}^{\,2}$ and $\hat{l}_z$. The eigenfunctions of $\hat{l}^{\,2}$ and $\hat{l}_z$ have the form

$$\Psi = R(r)\, Y_{lm}(\theta, \varphi), \tag{8.3}$$

where $R(r)$ is the *radial wave function* and $Y_{lm}(\theta, \varphi)$ are spherical harmonics satisfying

$$\hat{l}^{\,2}\, Y_{lm} = l(l+1)\, Y_{lm}$$

and

$$\hat{l}_z\, Y_{lm} = m\, Y_{lm}.$$

Replacing Eq. (8.3) in Eq. (8.1) and using that $\vec{l}$ acts only on the angular variables θ and φ we obtain that the radial part of the stationary states is given by

$$-\frac{\hbar^2}{2\mu}\,\frac{1}{r^2}\,\frac{d}{dr}\left(r^2\,\frac{dR}{dr}\right) + \frac{\hbar^2}{2\mu}\,\frac{l(l+1)}{r^2}\,R + U(r)\,R = E\,R.$$

This is the Schrödinger equation for the radial wave function. Notice that R depends on E and l but not on the z component of the angular momentum $l_z = m$. Consequently, all stationary states are $2l+1$ degenerate. This is of course an expected consequence of $\left[\hat{H}, \hat{\vec{l}}\right] = 0$ and in particular $\left[\hat{H}, \hat{l}_\pm\right] = 0$. In the case of a discrete spectrum one usually denotes the radial part by $R_{\nu l}(r)$, where $\nu \geq 0$ is an integer, known as *radial quantum number*, which gives the number of nodes in $R_{\nu l}(r)$ for $r > 0$.

In order to investigate the properties of $R(r)$ it is useful to make the substitution

$$R(r) = \frac{\chi(r)}{r}.$$

Using that

$$\begin{aligned}
\frac{1}{r^2}\,\frac{d}{dr}\left(r^2\,\frac{dR}{dr}\right) &= \frac{1}{r^2}\,\frac{d}{dr}\left[r^2\left(\frac{1}{r}\,\frac{d\chi}{dr} - \frac{1}{r^2}\,\chi\right)\right] = \frac{1}{r^2}\,\frac{d}{dr}\left(r\,\frac{d\chi}{dr} - \chi\right) \\
&= \frac{1}{r^2}\,\frac{d\chi}{dr} + \frac{1}{r}\,\frac{d^2\chi}{dr^2} - \frac{1}{r^2}\,\frac{d\chi}{dr} \\
&= \frac{1}{r}\,\frac{d^2\chi}{dr^2}
\end{aligned}$$

one obtains that the equation for χ reads

$$-\frac{\hbar^2}{2\mu}\,\frac{d^2\chi}{dr^2} + \frac{\hbar^2}{2\mu}\,\frac{l(l+1)}{r^2}\,\chi + U\,\chi = E\,\chi, \tag{8.4}$$

which looks exactly as a 1D Schrödinger equation with the effective potential $V(r) = U(r) + \hbar^2 l(l+1)/(2\mu r^2)$. Since $R(r)$ must be finite everywhere, it follows that

$$\chi(0) = 0.$$

This condition, which is rather obvious if $U(r)$ is finite everywhere, also holds when $U(r)$ diverges at the origin ($r \to 0$) as $U(r) \sim r^s$ with $s > -2$ [2]. This will be discussed in more detail below. The normalization condition for bound states requires

$$1 = \int_0^{+\infty} |R(r)|^2\, r^2\, dr = \int_0^{+\infty} |\chi(r)|^2\, dr.$$

Equation (8.4) is equivalent to a one-dimensional (1D) Schrödinger equation for $\chi(r)$ with the boundary condition $r \geq 0$ and $\chi(0) = 0$.

8.2. Properties of the motion in a spherical potential

Based on our knowledge of the motion in one dimension (see Sec. 4.12) a few general conclusions can be derived concerning the motion in a spherically symmetric field:

i) Since the motion is bounded on one side [$\chi(0) = 0$], there is only one eigenstate of Eq. (8.4) for each energy E. This implies that *for a given l* there is only one radial solution χ_{El} for each value of E. This does not preclude, however, that different values of l might lead to different χ_{lE} having the same eigenenergy. This is actually the case in the Coulomb field $U(r) = -\alpha/r$ $(\alpha > 0)$. Such degeneracies are said to be *accidental*, since they are not a consequence of rotational symmetry.

ii) Taking into account that the angular part is completely defined by the values of l and m, we conclude that E, l and m form a complete set of compatible observables according to which the eigenstates

$$\Psi_{Elm}(\vec{r}) = R_{El}(r)\, Y_{lm}(\theta, \varphi)$$

can be classified. Alternatively, one may consider ν, l and m as complete set of quantum numbers.

iii) Further insight in the structure of the energy spectrum can be obtained by applying the oscillator theorem, which states that the eigenfunction corresponding to the n-th eigenvalue of a discrete spectrum in 1D has $n-1$ nodes (i.e., the lowest-energy state has no nodes, the following 1 node, etc.). Notice that the effective potential

$$U_l(r) = U(r) + \frac{\hbar^2}{2m}\,\frac{l(l+1)}{r^2}$$

of the 1D Schrödinger equation for $\chi_{\nu l}(r)$ depends on l. We therefore fix the value of l and conclude that the lowest-energy eigenstate χ_{0l} for each l has no nodes. The number of nodes in $\chi_{\nu l}$ increases by one as we increase the level of excitation ν for any given l. Assigning $\nu = 0$ to the lowest-energy state and ordering the states by increasing energy, it follows that ν defines the number of nodes in $\chi(r)$ and in $R(r) = \chi/r$ for all finite values of r [excluding $r = 0$ where $\chi(0) = 0$]. The number ν is known as *radial quantum number*, l as *azimuthal* or *orbital quantum number* and m as *magnetic quantum number.*

iv) Since the centrifugal potential $\dfrac{\hbar^2}{2m}\,\dfrac{l(l+1)}{r^2}$ is strictly positive, it is clear that the ground state of a central symmetric problem always has $l = 0$. This is easily seen by applying the variational principle. In fact, using the radial part χ of a hypothetical ground state for $l \geq 1$ as a variational ansatz for the $l = 0$ equation, would always yield a strictly lower energy. Consequently, the ground state is non-degenerate ($2l + 1 = 1$ for $l = 0$).

v) The previous variational argument also implies that the lowest possible energy for a given l increases monotonously with l. While E increases with increasing l for a given ν, and also E increases with ν for a given l, there is no simple rule concerning different ν and l. For instance, in the Coulomb field $U(r) = -\alpha/r$, the eigenstate having $l = 0$ and $\nu = 1$ (i.e., the $2s$-orbital) is degenerate with the eigenstate having $l = 1$ and $\nu = 0$ (i.e., the $2p$-orbital).

vi) Finally, we can determine the behavior of $R(r)$ near the origin. We restrict ourselves to potentials that do not diverge very rapidly at the origin, satisfying

$$\lim_{r\to 0} r^2\, U(r) = 0, \tag{8.5}$$

which includes all cases of interest including the unscreened Coulomb field. The condition (8.5) implies that the particle does not fall into the center [2]. The Schrödinger equation reads

$$-\frac{\hbar^2}{2\mu}\,\frac{1}{r^2}\,\frac{d}{dr}\left(r^2\,\frac{dR}{dr}\right) + \frac{\hbar^2}{2\mu}\,\frac{l(l+1)}{r^2}\,R + UR = ER. \tag{8.6}$$

Multiplying Eq. (8.6) by r^2 and taking the limit $r \to 0$ we conclude by using (8.5) that for small r the radial wave function satisfies the equation

$$\frac{d}{dr}\left(r^2\,\frac{dR}{dr}\right) = l(l+1)\,R.$$

We seek $R(r)$ for $r \to 0$ in the form of a power law

$$R(r) = A\,r^s + O(r^{s+1})$$

and focus on the leading term. Since

$$\frac{d}{dr}\left[r^2\,\frac{d}{dr}\,(r^s)\right] = s\,\frac{d}{dr}\left(r^{s+1}\right) = s(s+1)\,r^s,$$

the condition on s reads

$$s(s+1) = l(l+1),$$

whose solutions are $s = l \geq 0$ and $s = -(l+1) \leq -1$. The latter is of course unphysical, since R would diverge at the origin.[12] Thus,

$$R_{nl}(r) \sim r^l.$$

The larger l is, the stronger is the centrifugal potential. Therefore, the probability of finding the particle close to the origin decreases faster as l increases.

8.3. The Coulomb field

A particularly important example of central potential is the Coulomb field of the form

$$U(r) = -\frac{\alpha}{r}$$

with $\alpha > 0$, which applies to Hydrogen-like atoms ($\alpha = Z\,e^2$). In this potential the eigenenergies take the form

$$E_n = -\frac{\mu\,\alpha^2}{2\,\hbar^2}\,\frac{1}{n^2},$$

where $n = 1, 2, 3, \ldots$ is called the *principal quantum number*. The integer n is strictly positive ($n \geq 1$) and must satisfy the condition

$$n \geq l + 1.$$

Notice that the energy is independent of l, except for the constraint

$$l \leq n - 1.$$

The ground state corresponds to $n = 1$ and $l = 0$. As expected, it is non-degenerate. The radial quantum number ν, giving the number of nodes of the radial wave function, is related to n and l by

$$\nu = n - l - 1.$$

---

[12] In fact, $s = -(l+1)$ with $l \geq 1$ would imply that $R(r) \propto 1/r^{l+1}$ is not integrable at the origin. The remaining case $l = 0$ implies $R(r) \propto 1/r$, which is integrable. However, $\nabla^2\,(1/r) = -4\pi\delta(\vec{r})$ would require that $U(r) \propto \delta(\vec{r})$ near the origin, which is also unphysical.

For a given n, l takes the values

$$l = 0, 1, \ldots, n-1.$$

This additional degeneracy (beyond the $2l+1$ degeneracy associated with the magnetic quantum number m) is an *accidental degeneracy*, which is specific of the Coulomb field. The principal quantum number n and the azimuthal quantum number l are used to label the orbitals in atomic physics:

	principal	orbital	radial	magnetic
orbitals	n	l	ν	$2l+1$
$1s$	1	0	0	1
$2s$	2	0	1	1
$2p$	2	1	0	3
$3s$	3	0	2	1
$3p$	3	1	1	3
$3d$	3	2	0	5

The lowest energy radial functions $R_{nl}(r)$ are

$$R_{1\,0} = 2\, e^{-r}$$

$$R_{2\,0} = \frac{1}{\sqrt{2}}\, e^{-r/2} \left(1 - \frac{r}{2}\right)$$

$$R_{2\,1} = \frac{1}{2\sqrt{6}}\, e^{-r/2}\, r$$

$$R_{3\,0} = \frac{2}{3\sqrt{3}}\, e^{-r/3} \left(1 - \frac{2}{3}\, r + \frac{2}{27}\, r^2\right)$$

$$R_{3\,1} = \frac{8}{27\sqrt{6}}\, e^{-r/3}\, r \left(1 - \frac{1}{6}\, r\right)$$

$$R_{3\,2} = \frac{4}{81\sqrt{30}}\, e^{-r/3}\, r^2.$$

The reader may wish to verify that the general properties discussed in the previous section are fulfilled.

9. Perturbation theory

The large majority of problems in quantum mechanics cannot be solved in an analytical way. As in any other branch of physics, approximation methods are of central importance. Besides the practical motivation for providing a quantitative or qualitative insight on otherwise inaccessible phenomena, the formal development of approximate methods often provides a valuable new perspective in the structure of the theory, which is rather universal and thus helps to develop a quantum mechanical intuition. This is particularly true in the case of perturbation theory.

The general problem can be formulated as follows. We consider a Hamiltonian

$$\hat{H} = \hat{H}_0 + \hat{V},$$

which can be splitted in two terms. The first one is the unperturbed Hamiltonian $\hat{H}_0$, whose eigenstates $|\, n^0 \,\rangle$ and eigenenergies E_n^0 are exactly known. The second term is the perturbation, which is assumed to be small, in some sense to be quantified below. In the following we consider first the stationary case, in which both $\hat{H}_0$ and $\hat{V}$ are independent of time. Later on we shall discuss perturbations $\hat{V}$ introducing a time dependence.

9.1. Symmetry and perturbation

Splitting $\hat{H}$ in the form $\hat{H} = \hat{H}_0 + \hat{V}$ is in some cases not obvious, nor is the best choice of a complete set of compatible observables spanning the Hilbert space in which $\hat{H}_0$ and $\hat{H}$ operate. In any case we assume that the properties of $\hat{H}_0$ are completely understood. By this we mean that we know a complete set of eigenstates $|\, n^0, \alpha \,\rangle$ with energies E_n^0, where α denotes the eigenvalues of the observables which might be needed in order to define the states completely:

$$\hat{H}_0 |\, n^0, \alpha \,\rangle = E_n^0 \,|\, n^0, \alpha \,\rangle.$$

In order to illustrate the various possible choices of $|\, n^0, \alpha \,\rangle$ and the role of symmetry in $\hat{H}_0$ and $\hat{V}$, let us consider the three dimensional isotropic harmonic oscillator

$$\begin{aligned} \hat{H}_0 &= \frac{\hat{p}^2}{2m} + \frac{1}{2} m\omega^2 \hat{r}^2 \\ &= \frac{1}{2m}\left(\hat{p}_x^2 + \hat{p}_y^2 + \hat{p}_z^2\right) + \frac{1}{2} m\omega^2 \left(x^2 + y^2 + z^2\right), \end{aligned}$$

whose eigenvalues are

$$E_N^0 = \left(N + \frac{3}{2}\right)\hbar\omega = \left(n_x + n_y + n_z + \frac{3}{2}\right)\hbar\omega = E^0_{n_x\, n_y\, n_z}.$$

The spectrum is highly degenerate since the eigenvalues depend only on

$$N = n_x + n_y + n_z.$$

The degeneracy is actually

$$g_N = (N+1)(N+2)/2 = \sum_{n_x=0}^{N} (N - n_x + 1).$$

The degeneracy of the spectrum of $\hat{H}_0$ opens various possibilities for the quantum numbers α classifying the states $|\, n,\, \alpha\,\rangle$ within each degenerate manifold. This choice reflects different symmetries of the underlying states, which becomes non-trivial when perturbations are introduced.

Besides the obvious Cartesian representation $|\, n_x,\, n_y,\, n_z\,\rangle$ with product wave functions $\langle\, x,\, y,\, z\,|\, n_x,\, n_y,\, n_z\,\rangle = \Psi_{n_x}(x)\Psi_{n_y}(y)\Psi_{n_z}(z)$, one may consider a spherical representation $|\, N,\, l,\, m\,\rangle$ in which the wave functions $\langle\, r,\, \theta,\, \varphi\,|\, N,\, l,\, m\,\rangle = R_{Nl}(r)\, Y_{lm}(\theta,\, \varphi)$ are the product of spherical harmonics and a solution of the radial Schrödinger equation with angular momentum l. In addition one may also consider polar coordinates, in which case the states $|\, N,\, n_z,\, m\,\rangle$ are characterized by the polar, z and magnetic quantum numbers N, n_z and m. Of course, unitary transformations within each degenerate subspace allow us to obtain any of these representations as linear combinations from each other.

Consider now the following perturbations:

$$\text{a)}\ \ V_a(\vec{r}) = \lambda\, v(r),$$

$$\text{b)}\ \ V_b(\vec{r}) = \lambda_1\, v_1(x) + \lambda_2\, v_2(y) + \lambda_3\, v_3(z) \text{ and}$$

$$\text{c)}\ \ V_c(\vec{r}) = \lambda\, v_{12}(x^2 + y^2) + \lambda'\, v_3(z).$$

It is clear that the difficulty in finding the eigenstates of $\hat{H} = \hat{H}_0 + \hat{V}$ will be very different, depending on the choice of the unperturbed states. For example, if we consider the $V_a = \lambda\, v(r)$ in Cartesian representation $|\, n_x,\, n_y,\, n_z\,\rangle$, we will find matrix elements among all the states having the same energy ($n_x + n_y + n_z = N$). Therefore, the changes in the eigenstates will be very significant even if the strength of the perturbation tends to zero (i.e., $V_a(\vec{r}) = \lambda\, v(r)$ with $\lambda \to 0$). Thus, Cartesian states $|\, n_x,\, n_y,\, n_z\,\rangle$ do not provide a meaningful starting point for a perturbative treatment of V_a. In contrast, the spherical basis $|\, N,\, l,\, m\,\rangle$, in which V_a is already diagonal within each degenerate subspace and the matrix elements are independent of m ($\langle\, N',\, l', m'\,|\, V_a\,|\, N,\, l,\, m\,\rangle = \delta_{ll'}\, \delta_{mm'} \langle\, N'\,|\, V_a\,|\, N\,\rangle$) is perfectly suited, since the shape of $|\, N,\, l,\, m\,\rangle$ does not change in the limit $\lambda \to 0$. Thus, we may expect that the corrections resulting from V_a can be introduced in an iterative way. Furthermore, it is not necessary to solve the eigenvalue problem for different m, since they are independent of m. Similar considerations immediately show that $|\, n_x,\, n_y,\, n_z\,\rangle$ and $|\, N,\, n_z,\, m\,\rangle$ are the appropriate choices for V_b and V_c, respectively.

A few general conclusions may be drawn:

i) Finding the stationary states of $\hat{H} = \hat{H}_0 + \hat{V}$ is considerably simplified by symmetry considerations, particularly when some of the constants of motion under $\hat{H}_0$ remain constants of motion under the action of the perturbation $\hat{V}$.

ii) Not all choices of stationary states $|\, n^0 \,\rangle$ of $\hat{H}_0$ are good starting points for finding the eigenstates $|\, n \,\rangle$ of $\hat{H}$. Particularly in the presence of degeneracies, the changes introduced by $\hat{V}$ may be very significant even in the limit of $\hat{V} \to 0$ (i.e., $|\langle\, n^0 \,|\, n \,\rangle| \nrightarrow 1$ for $\hat{V} \to 0$). In order that a perturbation expansion remains meaningful we must ensure that $|\, n \,\rangle \xrightarrow{V \to 0} |\, n^0 \,\rangle$.

iii) If symmetry of the perturbation $\hat{V}$ is lower than the symmetry of $\hat{H}_0$, the degeneracy of the spectrum is in general reduced. For example, the perturbation V_b will in general remove the degeneracy of $\hat{H}_0$, since it removes isotropy, while the perturbation V_a will remove the degeneracy between states having different l (but the same N). The $2l+1$ degeneracy among the different m remains of course unaltered under the action of V_a.

9.2. The Rayleigh-Schrödinger perturbation expansion

We intend to obtain a general iterative solution of the eigenvalues problem

$$\hat{H}|\, n \,\rangle = E_n|\, n \,\rangle, \tag{9.1}$$

where

$$\hat{H} = \hat{H}_0 + \hat{V} \tag{9.2}$$

and $\hat{V}$ can be regarded as "small" in a sense to be defined with some mathematical rigor below. The idea is to seek for an iterative solution of the form

$$E_n = E_n^{(0)} + E_n^{(1)} + E_n^{(2)} + \cdots$$

or equivalently,

$$\Delta_n = E_n - E_n^{(0)} = \Delta_n^{(1)} + \Delta_n^{(2)} + \cdots$$

and

$$|\, n \,\rangle = |\, n_0 \,\rangle + |\, n_1 \,\rangle + |\, n_2 \,\rangle + \cdots ,$$

where the k-order corrections to the eigenenergies $\Delta_n^{(k)}$ and eigenstates $|\, n_k \,\rangle$ are of the order of V_{lm}^k, where $V_{lm} = \langle\, l_0 \,|\, \hat{V} \,|\, m_0 \,\rangle$ are the matrix elements of the perturbation $\hat{V}$. In order to allow for a clear identification of the contributions of the different orders in $\hat{V}$, it is useful to write

$$\hat{H} = \hat{H}_0 + \lambda\, \hat{V}, \tag{9.3}$$

$$\Delta_n = \lambda\, \Delta_n^{(1)} + \lambda^2\, \Delta_n^{(2)} + \cdots , \tag{9.4}$$

$$|\, n \,\rangle = |\, n_0 \,\rangle + \lambda\, |\, n_1 \,\rangle + \lambda^2\, |\, n_2 \,\rangle + \cdots , \tag{9.5}$$

and to regard $\hat{H}$, $\Delta_n(\lambda)$ and $|\, n \,\rangle_\lambda$ as functions of the parameter $\lambda \in [0,\, 1]$. In this way $\lambda = 0$ corresponds to the unperturbed case $\hat{H}_0$. Finally, at the end of the calculations we may set $\lambda = 1$ and recover Eqs. (9.1) and (9.2).

9.3. The non-degenerate case

We consider first an unperturbed Hamiltonian $\hat{H}_0$ having a non-degenerate spectrum $E_n^{(0)}$. As we shall see, this hypothesis is only required for the unperturbed eigenstate $|\, n_0 \,\rangle$ to which the perturbed eigenstate $|\, n \,\rangle$ tends when $\lambda \to 0$. We aim to calculate each $|\, n \,\rangle$ in terms of $|\, n_0 \,\rangle$ knowing that $\langle n_0 | n \rangle \neq 0$ for all λ and that $|\, n \,\rangle \to |\, n_0 \,\rangle$ for $\lambda \to 0$. The degeneracy of the other unperturbed levels $E_m^{(0)}$ with $E_m^{(0)} \neq E_n^{(0)}$ is not relevant for calculating $|\, n \,\rangle$. Later on we will remove the restriction on the non degeneracy of $|\, n_0 \,\rangle$. We further assume that the spectrum is discrete. Again this applies only to the state $|\, n_0 \,\rangle$ to which the perturbation is applied, i.e., which is actually being calculated. $\hat{H}_0$ itself may also have a continuous spectrum, at higher energies, for example. Extending the formulae to states belonging to the continuum spectrum is in principle straightforward by replacing the sums over the discrete spectrum by the corresponding integrals over the continuum spectrum.

Our aim is to solve the eigenvalue problem

$$\hat{H}|\, n \,\rangle = E_n|\, n \,\rangle,$$
$$\left(\hat{H}_0 + \lambda \hat{V}\right)|\, n \,\rangle = E_n|\, n \,\rangle,$$

which we rewrite in the form

$$\left(\hat{H}_0 - E_n^{(0)}\right)|\, n \,\rangle = \left(\Delta_n - \lambda \hat{V}\right)|\, n \,\rangle. \tag{9.6}$$

This form is quite appealing because it groups on the left side known operators and constants of order zero, keeping on the right side all the terms of order λ or higher. Still we need to solve for $|\, n \,\rangle$ in order to obtain an expression which is useful for iterative approximations.

In order to solve for $|\, n \,\rangle$ one may regard Eq. (9.6) as an inhomogeneous linear equation, where

$$\left(\hat{H}_0 - E_n^{(0)}\right)|\, n \,\rangle = 0 \tag{9.7}$$

is the homogeneous part and

$$\left(\Delta_n - \lambda\hat{V}\right)|\, n \,\rangle$$

the inhomogeneity. The solution of the homogeneous part (9.7) is obviously proportional to $|\, n_0 \,\rangle$ and can be written as $|\, n_0 \,\rangle\langle\, n_0 \,|\, n \,\rangle$. However, notice that Eq. (9.6) provides no information on the projection $\langle\, n_0 \,|\, n \,\rangle$. In fact, $\left(\hat{H}_0 - E_n^{(0)}\right)$ projects out any component of $|\, n \,\rangle$ along $|\, n_0 \,\rangle$, since the latter has the unperturbed energy $E_n^{(0)}$. As for any eigenvector, one of the non-vanishing components of $|\, n \,\rangle$, in the present case $\langle\, n_0 \,|\, n \,\rangle$, can always be chosen freely, provided that the normalization of $|\, n \,\rangle$ is addressed *a posteriori.* Since

$$|\, n \,\rangle \to |\, n_0 \,\rangle \quad \text{for} \quad \lambda \to 0$$

we assume

$$\boxed{\langle\, n_0 \,|\, n \,\rangle = 1} \tag{9.8}$$

for all λ. Consequently,

$$\begin{aligned}\langle\, n \,|\, n \,\rangle &= \langle\, n \,| \left(\sum_m |\, m_0 \,\rangle\langle\, m_0 \,| \right) |\, n \,\rangle \\ &= 1 + \sum_{m \neq n} |\langle\, m_0 \,|\, n \,\rangle|^2\end{aligned}$$

necessarily exceeds unity. The renormalization of $|\, n \,\rangle$ is not important to lowest order in λ, as it will be discussed below.

In order to find a particular solution of Eq. (9.6) including the inhomogeneity, one is tempted to multiply Eq. (9.6) from the left by the Green's operator

$$\hat{G}_0 = \left(\hat{H}_0 - E_n^{(0)} \right)^{-1} .$$

But of course $\left(\hat{H}_0 - E_n^{(0)} \right)$ is invertible only in the subspace orthogonal to $|\, n_0 \,\rangle$. To circumvent this problem we introduce the projection operators

$$\hat{P}_n = |\, n_0 \,\rangle\langle\, n_0 \,|$$

and its complement

$$\hat{Q}_n = 1 - \hat{P}_n = \sum_{m \neq n} |\, m_0 \,\rangle\langle\, m_0 \,|.$$

These allow us to split

$$|\, n \,\rangle = \left(\hat{P}_n + \hat{Q}_n \right) |\, n \,\rangle = |\, n_0 \,\rangle\langle\, n_0 \,|\, n \,\rangle + \hat{Q}_n |\, n \,\rangle \tag{9.9}$$

in the part belonging to the kernel of $\left(\hat{H}_0 - E_n^{(0)} \right)$ and the part where $\left(\hat{H}_0 - E_n^{(0)} \right)$ is regular. Notice that $[\hat{Q}_n, \hat{H}_0] = 0$ because they are diagonal on the same basis. We consider the operator

$$\frac{\hat{Q}_n}{\hat{H}_0 - E_n^{(0)}} = \sum_{m \neq n} \frac{|\, m_0 \,\rangle\langle\, m_0 \,|}{E_m^{(0)} - E_n^{(0)}},$$

which is well defined since $E_n^{(0)}$ is nondegenerate, and multiply with it Eq. (9.6) from the left to obtain

$$\begin{aligned}\frac{\hat{Q}_n}{\hat{H}_0 - E_n^{(0)}} \left(\hat{H}_0 - E_n^{(0)} \right) |\, n \,\rangle &= \sum_{m \neq n} \frac{|\, m_0 \,\rangle\langle\, m_0 \,|}{E_m^{(0)} - E_n^{(0)}} \left(\hat{H}_0 - E_n^{(0)} \right) |\, n \,\rangle \\ &= \sum_{m \neq n} |\, m_0 \,\rangle\langle\, m_0 \,|\, n \,\rangle \\ &= \hat{Q}_n |\, n \,\rangle\end{aligned}$$

This implies

$$\hat{Q}_n|\, n\,\rangle = \frac{\hat{Q}_n}{\hat{H}_0 - E_n^{(0)}} \left(\Delta_n - \lambda\, \hat{V}\right) |\, n\,\rangle. \tag{9.10}$$

This is the particular solution of the inhomogeneous equation. Combining (9.8) and (9.10) in (9.9) we have

$$\boxed{|\, n\,\rangle = |\, n_0\,\rangle + \frac{\hat{Q}_n}{\hat{H}_0 - E_n^{(0)}} \left(\Delta_n - \lambda\, \hat{V}\right) |\, n\,\rangle}. \tag{9.11}$$

This exact equation is most useful for obtaining $|\, n\,\rangle$ by successive iterations, each one involving a higher-order correction in λ. For example, one may start with $|\, n_0\,\rangle$ (the limit of $|\, n\,\rangle$ for $\lambda \to 0$) insert it on the right-hand side in order to obtain the correction of $|\, n\,\rangle$ to first order in λ. Replacing then this approximation in the RHS of Eq. (9.11) gives $|\, n\,\rangle$ to second order in λ and so on. Formally, one would replace Eq. (9.11) in itself iteratively *ad infinitum* and thus obtain a series expansion of $|\, n\,\rangle$ as a function of λ.

But before doing that we need an iterative equation for the correction of the energy levels Δ_n, which also enter Eq. (9.11). This we obtain from the relation

$$\langle\, n_0\,|\left(\Delta_n - \lambda\, \hat{V}\right) |\, n\,\rangle = 0\,,$$

which is easily inferred from Eq. (9.6) by recalling that $\hat{H}_0|\, n_0\,\rangle = E_n^{(0)}|\, n_0\,\rangle$. Using the normalization condition $\langle\, n_0\,|\, n\,\rangle = 1$ we have

$$\boxed{\Delta_n = \lambda\langle\, n_0\,|\,\hat{V}\,|\, n\,\rangle\,,} \tag{9.12}$$

which shows that the knowledge of $|\, n\,\rangle$ to the order k in λ yields the correction Δ_n of the nth energy level to the order $k+1$. This can then be inserted in Eq. (9.11) to infer the correction of $|\, n\,\rangle$ to the order $k+1$.

The basic assumptions of the Rayleigh-Schrödinger perturbation expansion are that $\hat{V}$ is small in some sense and that Eqs. (9.11) and (9.12) can be solved by iterations. In the following we develop such a solution, derive the corresponding perturbation expansion of Δ_n and $|\, n\,\rangle$, and finally analyze the convergence conditions.

We expand Δ_n and $|\, n\,\rangle$ in powers of λ in order to facilitate keeping track of the contributions with different orders in $\hat{V}$:

$$\Delta_n = \sum_{k=1}^{\infty} \lambda^k\, \Delta_n^{(k)} \tag{9.13}$$

and

$$|\, n\,\rangle = \sum_{k=1}^{\infty} \lambda^k\,|\, n_k\,\rangle. \tag{9.14}$$

Replacing this expansions in Eq. (9.12) we obtain

$$\sum_{k=1}^{\infty} \lambda^k \, \Delta_n^{(k)} = \lambda \, \langle \, n_0 \, | \, \hat{V} \sum_{k=0}^{\infty} \lambda^k \, | \, n_k \, \rangle,$$

which implies

$$\boxed{\Delta_n^{(k+1)} = \langle \, n_0 \, | \, \hat{V} \, | \, n_k \, \rangle.} \tag{9.15}$$

The $(k+1)$-order corrections to Δ_n is obtained straightforwardly from the k-order correction to the eigenstate. Replacing the expansions (9.13) and (9.14) in Eq. (9.11) we obtain

$$\sum_{k=0}^{\infty} \lambda^k \, | \, n_k \, \rangle = | \, n_0 \, \rangle + \frac{\hat{Q}_n}{\hat{H}_0 - E_n^{(0)}} \left[\sum_{k=1}^{\infty} \lambda^k \, \Delta_n^{(k)} - \lambda \, \hat{V} \right] \sum_{k'=0}^{\infty} \lambda^{k'} | \, n_{k'} \, \rangle \tag{9.16}$$

The second term on the right-hand side can be rearranged as follows:

$$\begin{aligned} & \frac{\hat{Q}_n}{\hat{H}_0 - E_n^{(0)}} \left[\sum_{k=1}^{\infty} \sum_{k'=0}^{\infty} \lambda^{k+k'} \Delta_n^{(k)} | \, n_{k'} \, \rangle - \sum_{k=0}^{\infty} \lambda^{k+1} \hat{V} | \, n_k \, \rangle \right] \\ & \qquad = \frac{\hat{Q}_n}{\hat{H}_0 - E_n^{(0)}} \left[\sum_{m=1}^{\infty} \lambda^m \sum_{k=1}^{m} \Delta_n^{(k)} | \, n_{m-k} \, \rangle - \sum_{k=1}^{\infty} \lambda^k \hat{V} | \, n_{k-1} \, \rangle \right] \end{aligned} \tag{9.17}$$

Comparing (9.17) with the left-hand side of (9.16) and equating the kets corresponding to the same powers of λ, we obtain for $k \geq 1$

$$\boxed{| \, n_k \, \rangle = \sum_{k'=1}^{k} \Delta_n^{(k')} \frac{\hat{Q}_n}{\hat{H}_0 - E_n^{(0)}} | \, n_{k-k'} \, \rangle - \frac{\hat{Q}_n}{\hat{H}_0 - E_n^{(0)}} \, \hat{V} \, | \, n_{k-1} \, \rangle.} \tag{9.18}$$

Again, as in the case of $\Delta_n^{(k)}$, the k-order correction to the eigenvector can be obtained by operating on the states $| \, n_{k'} \, \rangle$ with $k' < k$.

We may now make these expressions explicit for the most important lowest orders. The first-order correction of the energy level reads

$$\Delta_n^{(1)} = \langle \, n_0 \, | \, \hat{V} \, | \, n_0 \, \rangle, \tag{9.19}$$

which represents the average of the perturbation in the corresponding unperturbed state $| \, n_0 \, \rangle$, a transparent result worth keeping in mind.

For the first-order correction of $| \, n \, \rangle$ we have

$$| \, n_1 \, \rangle = \Delta_n^{(1)} \frac{\hat{Q}_n}{\hat{H}_0 - E_n^{(0)}} | \, n_0 \, \rangle - \frac{\hat{Q}_n}{\hat{H}_0 - E_n^{(0)}} \, \hat{V} \, | \, n_0 \, \rangle.$$

Knowing that $\hat{Q}_n|\, n_0 \,\rangle = 0$, we obtain

$$
\begin{aligned}
|\, n_1 \,\rangle &= -\frac{\hat{Q}_n}{\hat{H}_0 - E_n^{(0)}}\, \hat{V}|\, n_0 \,\rangle \\
&= \sum_{m\neq n} \frac{|\, m_0 \,\rangle\langle\, m_0 \,|}{E_n^{(0)} - E_m^{(0)}}\, \hat{V}|\, n_0 \,\rangle .
\end{aligned}
$$

It is useful to denote the matrix elements of the perturbation between the unperturbed eigenstates in a more compact way as

$$
V_{mn} = \langle\, m_0 \,|\, \hat{V} \,|\, n_0 \,\rangle .
$$

Thus, we can write

$$
|\, n_1 \,\rangle = \sum_{m\neq n} \frac{V_{mn}}{E_n^{(0)} - E_m^{(0)}} |\, m_0 \,\rangle . \tag{9.20}
$$

To first order in $\hat{V}$, the amplitude of the state $|\, m_0 \,\rangle$ in the eigenstate $|\, n \,\rangle$ is directly proportional to the matrix element V_{mn} between $|\, n_0 \,\rangle$ and $|\, m_0 \,\rangle$, and inversely proportional to the energy difference between these states. This is a very fundamental and widely applicable trend in quantum mechanics. This explains why the same matrix element V_{mn} connecting two states $|\, m_0 \,\rangle$ and $|\, n_0 \,\rangle$ can have a huge effect on the eigenstates and eigenenergies if the unperturbed levels $E_m^{(0)}$ and $E_n^{(0)}$ are close ($|E_m^{(0)} - E_n^{(0)}| \sim V_{mn}$), while the effects can be negligible if the unperturbed energies are far appart ($|E_m^{(0)} - E_n^{(0)}| \gg |V_{mn}|$).

The second-order correction to Δ_n is

$$
\Delta_n^{(2)} = \langle\, n_0 \,|\, \hat{V}\, \frac{\hat{Q}_n}{\hat{H}_0 - E_n^{(0)}}\, \hat{V} \,|\, n_0 \,\rangle = \sum_{m\neq n} \frac{|V_{nm}|^2}{E_n^{(0)} - E_m^{(0)}} . \tag{9.21}
$$

This second-order correction is very important in practice, since very often the first-order correction vanishes ($V_{nn} = 0$) for example due to symmetry reasons. The trends implied by (9.21), being independent of explicit form of $\hat{V}$, have a universal validity in quantum mechanics. For example, in the case of the ground state we have $E_n^{(0)} - E_m^{(0)} < 0\ \forall m$, which implies that *any perturbation* tends to stabilize the ground state. This can be interpreted on the basis of the variational principle. Indeed, if one regards the zeroth-order approximation of the ground state wave-function $|\, n_0 \,\rangle$ as a variational Ansatz for the exact ground state $|\, n \,\rangle$, it is clear that

$$
E_n = \langle\, n \,|\, \hat{H}_0 + \hat{V} \,|\, n \,\rangle \leq \langle\, n_0 \,|\, \hat{H}_0 + \hat{V} \,|\, n_0 \,\rangle = E_n^{(0)} + \Delta_n^{(1)} .
$$

The situation is of course different for the excited states. Eq. (9.21) tells us that any two levels $E_k^{(0)}$ and $E_l^{(0)}$ connected by a matrix element V_{kl} tend to repel each other: If

$E_k^{(0)} < E_l^{(0)}$ the second-order correction of the level k is negative, while the correction on the level l due to the same coupling is positive.

Finally, Eqs. (9.20) and (9.21), but also the general expansions (9.15) and (9.18), give us a crude criterion for the convergence of the Rayleigh-Schrödinger perturbation expansion, namely,

$$|V_{mn}| = |\langle\, m_0 \,|\, \hat{V} \,|\, n_0 \,\rangle| \ll |E_m^{(0)} - E_n^{(0)}|. \tag{9.22}$$

Notice that this is only a necessary condition. Moreover, its validity usually depends strongly on the particular state $|\, n \,\rangle$ that one is trying to approximate. A well isolated state $|\, n_0 \,\rangle$, with large gaps $|E_n^{(0)} - E_m^{(0)}|$ to all other unperturbed levels may satisfy (9.22) very well and already the first-order correction (9.20) may be a very good approximation to $|\, n \,\rangle$. However, another state in the same system belonging to a group of close by non-degenerate levels may be much harder to approximate accurately. Rigorous mathematical conditions are not available, at least not in a form having practical relevance. Every problem must be analyzed individually from a physical perspective, in order to conclude if a given level of perturbation approximation is satisfactory, or if the perturbation expansion is sound at all.

A very nice example of the limitations of the simple criterion (9.22), but also of the ways to take advantage of perturbation theory, is discussed in the book by Gottfried [1]. Consider a one dimensional harmonic oscillator $\hat{H}_0 = m\omega^2 x^2/2$ with a perturbation

$$V(x) = \gamma\, \hbar\, \omega \left(\frac{x}{x_0}\right)^k$$

that is cubic or quartic in x (i.e., $k = 3$ or 4). $x_0 = \sqrt{\hbar/m\,\omega}$ stands for the length scale of the oscillator. It is clear that taking γ sufficiently small one can always satisfy (9.22), i.e., $|V_{mn}| \ll \hbar\omega$ for an arbitrary number of states above the ground state. Nevertheless, the consequences of $\hat{V}$ on the spectrum are dramatic for $k = 3$ and for $k = 4$ with $\gamma < 0$, in which cases the spectrum becomes continuous, even for arbitrarily small $|\gamma|$. Only $k = 4$ and $\gamma > 0$ (in general k even and $\gamma > 0$) allows a safe perturbation expansion. Still, even for $k = 3$ or for $k = 4$ with $\gamma < 0$, the states derived from the unperturbed ground state and low-lying excited states correspond to strong resonances in the continuum of the perturbed oscillator (γ small). Their position in the spectrum and their lifetime can then be calculated by using a perturbation expansion.

9.4. Wave function normalization

Before closing the discussion of non-degenerate perturbation theory we would like to calculate the norm of the approximation to $|\, n \,\rangle$. This is certainly larger than 1, since we assumed $\langle\, n \,|\, n_0 \,\rangle = 1$ for all λ in order to derive the perturbation expansion [see Eq. (9.8)]. We consider the first order approximation, which according to Eq. (9.20) is given by

$$|\, n \,\rangle = |\, n_0 \,\rangle + \sum_{m \neq n} \frac{V_{mn}}{E_n^{(0)} - E_m^{(0)}} |\, m_0 \,\rangle + O(V^2). \tag{9.23}$$

This implies

$$\langle\, n\,|\, n\,\rangle = 1 + \sum_m \frac{|V_{mn}|^2}{\left(E_n^{(0)} - E_m^{(0)}\right)^2} + O\left(V^3\right).$$

The normalization factor is thus

$$\langle\, n\,|\, n\,\rangle^{-1/2} = 1 - \frac{1}{2}\sum_m \frac{|V_{mn}|^2}{\left(E_n^{(0)} - E_m^{(0)}\right)^2} + O(V^3),$$

which implies that the wave function $|\, n\,\rangle$ as given by (9.23) is properly normalized to first order in V. In other words, although $\langle\, n\,|\, n\,\rangle$ depends on V or λ, the probability of finding the system in a state other than $|\, n_0\,\rangle$ (after the action of the perturbation) is of second order in V.

9.5. The degenerate case

We consider the case in which the unperturbed states, whose energy and eigenstates we want to compute under the action of the perturbation V, belong to a level having a degeneracy g. Let us denote these states by $|\, n_0, \alpha\,\rangle$, where

$$\hat{H}_0|\, n_0, \alpha\,\rangle = E_n^{(0)}|\, n_0, \alpha\,\rangle$$

with $\alpha = 1, \cdots, g$. The choice of these states is a priori arbitrary, since we can combine them linearly at will. However, the choice of the unperturbed states ceases to be arbitrary if one requires that the changes that they experience under the action of a very small perturbation remain small. To be explicit, we denote the eigenstates of

$$\hat{H} = \hat{H}_0 + \lambda\,\hat{V}$$

by $|\, n, \alpha\,\rangle$, so that

$$\hat{H}|\, n, \alpha\,\rangle = E_{n\alpha}|\, n, \alpha\,\rangle.$$

The basic condition in order that the development

$$|\, n, \alpha\,\rangle = |\, n_0, \alpha\,\rangle + \lambda\,|\, n_1, \alpha\,\rangle + \lambda^2|\, n_2, \alpha\,\rangle + \cdots$$

in a power series of λ remains meaningful is that

$$|\, n, \alpha\,\rangle \xrightarrow{\lambda\to 0} |\, n_0, \alpha\,\rangle. \tag{9.24}$$

This means that the unperturbed states must be adapted to the perturbation $\hat{V}$, so that the changes in $|\, n, \alpha\,\rangle$ are arbitrarily small when $\lambda \to 0$.

Let us illustrate the problem with a simple two-state Hamiltonian

$$\hat{H}_0 = \begin{pmatrix} \varepsilon_0 & 0 \\ 0 & \varepsilon_0 \end{pmatrix}$$

whose energy levels are degenerate. Let us denote the orthonormal kets in the unperturbed basis by $|\,1\,\rangle$ and $|\,2\,\rangle$. For instance, $|\,1\,\rangle$ and $|\,2\,\rangle$ may describe two identical orbitals on atoms that are far apart. If we introduce the perturbation

$$\hat{V} = \begin{pmatrix} \Delta & 0 \\ 0 & -\Delta \end{pmatrix}$$

we obtain two distinct energies $\varepsilon_\pm = \varepsilon_0 \pm \Delta$. The perturbation removes the degeneracy, but the states $|\,1\,\rangle$ and $|\,2\,\rangle$ remain the eigenstates of the perturbed Hamiltonian $\hat{H} = \hat{H}_0 + \hat{V}$. The reason for this is, of course, that the perturbation is diagonal in the unperturbed basis $|\,1\,\rangle$ and $|\,2\,\rangle$. Moreover, if we denote the eigenstates of $\hat{H}$ by $|\,+\,\rangle$ and $|\,-\,\rangle$, with $\hat{H}|\,+\,\rangle = \varepsilon_+|\,+\,\rangle$ and $\hat{H}|\,-\,\rangle = \varepsilon_-|\,-\,\rangle$, it is clear that the condition (9.24) is trivialy satisfied, since $|\,+\,\rangle$ ($|\,-\,\rangle$) coincides with $|\,1\,\rangle$ ($|\,2\,\rangle$).

However, if the perturbation is non-diagonal, i.e.,

$$\hat{V} = \begin{pmatrix} 0 & t \\ t & 0 \end{pmatrix},$$

the changes introduced by $\hat{V}$ are dramatic. We know that in this case the eigenstates are $|\,+\,\rangle = \frac{1}{\sqrt{2}}(|\,1\,\rangle + |\,2\,\rangle)$ and $|\,-\,\rangle = \frac{1}{\sqrt{2}}(|\,1\,\rangle - |\,2\,\rangle)$, even in the limit of $t \to 0$. The angle between $|\,+\,\rangle$ and $|\,1\,\rangle$ or $|\,2\,\rangle$ is $\pi/4$, the largest value one can obtain between pairs of orthogonal vectors in a plane! The only way to avoid such a discontinuous change in $|\,n,\alpha\,\rangle$ upon switching on an arbitrary small V ($t \to 0$) is to choose the unperturbed basis such that there are no off-diagonal matrix elements of the perturbation between degenerate unperturbed states. In other words, before starting with any perturbative expansion of $|\,n,\alpha\,\rangle$ and $E_{n\alpha}$ in powers of λ, we need to adapt the unperturbed states to the perturbation under study, by solving the so-called *secular equation*

$$\det\left[V_{\alpha\beta} - \varepsilon\,\delta_{\alpha\beta}\right] = 0\,, \tag{9.25}$$

where

$$V_{\alpha\beta} = \langle\, n_0, \alpha \,|\, \hat{V} \,|\, n_0, \beta \,\rangle \tag{9.26}$$

are the matrix elements of $\hat{V}$ between the eigenstates of $\hat{H}_0$ in the degenerate subspace under consideration ($\alpha, \beta = 1, \ldots, g$). This represents a finite $g \times g$ eigenvalue problem.

Solving Eq. (9.25) implies building the linear combinations

$$|\,\tilde{n}_0, \gamma\,\rangle = \sum_{\alpha=1}^{g} c_\alpha^\gamma\,|\,n_0, \alpha\,\rangle$$

with $\gamma = 1, \ldots, g$, so that

$$\sum_\beta V_{\alpha\beta}\, c_\beta^\gamma = \Delta_\gamma^{(1)}\, c_\alpha^\gamma\,.$$

In this way one obtains the new unperturbed states $|\,\tilde{n}_0,\gamma\,\rangle$ and the corresponding first order corrections to the energy levels $\Delta_\gamma^{(1)}$, which are the eigenvectors and eigenvalues of the matrix $V_{\alpha\beta}$ [see Eqs. (9.25) and (9.26)]. Notice that the actually important property of $|\,\tilde{n}_0,\gamma\,\rangle$ is that $\hat{V}$ is diagonal in the g-dimensional subspace spanned by them. The fact that the eigenvalues $\Delta_\gamma^{(1)}$ give the first-order corrections to the energy levels coincides with (9.19), since $\Delta_\gamma^{(1)} = \langle\,\tilde{n}_0,\gamma\,|\,\hat{V}\,|\,\tilde{n}_0,\gamma\,\rangle$.

Once that appropriate zeroth-order states $|\,\tilde{n}_0,\gamma\,\rangle$ are obtained, the corrections due to the coupling to other states $|\,m_0,\beta\,\rangle$ having $E_m^{(0)} \neq E_n^{(0)}$ can be calculated just as in the non-degenerate case. This holds, however, only if at first order [Eq. (9.25)] the perturbation removes the degeneracy completely, i.e., if all the $\Delta_\gamma^{(1)}$ are different, and if no further degeneracies with other states are accidentally introduced by the first-order corrections $\Delta_\gamma^{(1)}$. The latter can be safely assumed in the limit of small $V_{\alpha\beta}$.

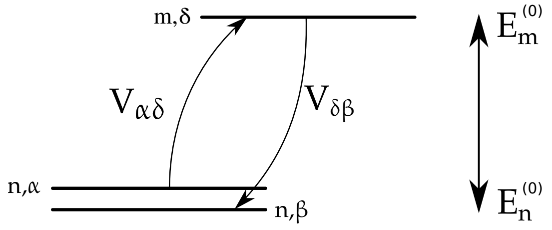
It often happens, however, that the perturbation has no matrix elements between the unperturbed orbitals $|\,n_0,\alpha\,\rangle$, or that some degeneracies still remain after the diagonalization of V. In this case the second-order approximation needs to be calculated. Although the equations tend to look more complicated, the basic idea remains, as before, to find the unperturbed states $|\,\tilde{n}_0,\gamma\,\rangle$ which do not change in the limit of $V \to 0$. The only difference is that now all the processes up to second order in V need to be taken into account. The secular equation to be solved reads

$$\det\left[\langle\, n_0,\alpha\,|\,\hat{V}\,|\,n_0,\beta\,\rangle + \sum_{\substack{m\neq n\\ \delta}} \frac{\langle\, n_0,\alpha\,|\,\hat{V}\,|\,m_0,\delta\,\rangle\langle\, m_0,\delta\,|\,\hat{V}\,|\,n_0,\beta\,\rangle}{E_n^{(0)}-E_m^{(0)}} - \varepsilon\,\delta_{\alpha\beta}\right] = 0$$

where α and β belong to the g-fold degenerate subspace $(\alpha,\beta = 1,\dots,g)$ and $|\,m_0,\delta\,\rangle$ refers to the unperturbed states outside the degenerate subspace (i.e., $E_m^{(0)} \neq E_n^{(0)}$). This secular equation has the same form as Eq. (9.25) with the effective coupling

$$V_{\alpha\beta}^{eff} = \langle\, n_0,\alpha\,|\,\hat{V}\,|\,n_0,\beta\,\rangle + \sum_{\substack{m\neq n\\ \delta}} \frac{\langle\, n_0,\alpha\,|\,\hat{V}\,|\,m_0,\delta\,\rangle\,\langle\, m_0,\delta\,|\,\hat{V}\,|\,n_0,\beta\,\rangle}{E_n^{(0)}-E_m^{(0)}}\,.$$

between the states $|\,n_0,\alpha\,\rangle$ and $|\,n_0,\beta\,\rangle$. The following picture illustrates how the effective second-order coupling arises:



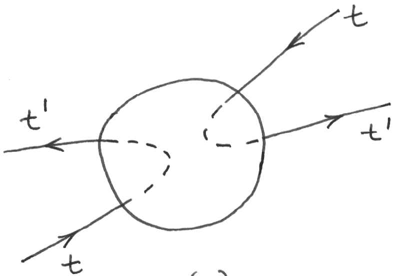
10. Identity of particles

The physical world and in particular atoms, molecules and solids are made of many interacting particles. In principle the N body wave function $\Psi(x_1 \dots x_N)$ and the Schrödinger equation $i\hbar\frac{\partial\Psi}{\partial t} = \hat{H}\Psi$ contain all the information for describing the physical states of a given system of N identical particles and their time evolution. However, neither the Schrödinger equation is easy to solve in general nor the wave function $\Psi(x_1 \dots x_N)$ appears to be the most practical representation of many-particle states. The representation in terms of *occupation numbers* usually known as *second quantization* provides a much more appropriate framework. Most of the theoretical developments on the many-body properties of matter rely on this approach.

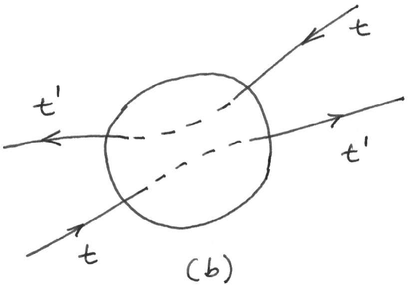
Second quantization is a formalism in which the occupation numbers of an arbitrary complete set of single-particle states play the role of independent variables, instead of the coordinates x_i of the individual particles, as in the usual wave function $\Psi(x_1 \dots x_N)$. The transitions between different many-particle states can be visualized as changes in the occupations of simpler single-particle orbitals. This is particularly useful in order to formulate, visualize and understand the fundamental physical processes defining the electronic and magnetic properties of matter. For instance, the relevant orbitals responsible for chemical bonding, optical transitions, conductivity, magnetism etc. can be focused by an appropriate choice of the single-particle basis. The most important interactions and energy scales can then be readily identified. In the following we discuss the principle of indistinguishability of identical particles and its consequences on the symmetry of many-body wave functions. A detailed discussion of second quantization may be found in the books by Landau & Lifshitz or Fetter & Wallecka [2, 6] or in the lecture notes on Quantum Mechanics II (`http://www.physik.uni-kassel.de/pastor`).

10.1. The principle of indistinguishability of identical particles

In classical physics the particles preserve their individuality despite having the same physical properties. A "numbering" of the particles is possible with which one can follow the trajectory of each particle individually along its path. This applies, whatever the number of particles is, and in particular for particles that are exactly identical in all respects. In quantum mechanics the situation is entirely different, since the notion of deterministic path ceases to have any meaning as a consequence of Heisenberg's uncertainty principle. Even if a numbering of the particles were possible at some time, for example if we measure the position of each particle in the system at time t, there is no possibility of tracking the positions of the particles at any future (or past) time $t' > t$ ($t' < t$), since the coordinates have no definite values even at times arbitrarily close to t. If we then localize (or measure) an electron at given instant $t' > t$, it is impossible to say which electron (among the N previously localized ones) has arrived at this point. The lack of a single deterministic path can be illustrated in the following scattering picture [7]:



(a)



(b)

In quantum mechanics identical particles entirely lose their individuality and become completely indistinguishable. *No experimental measurement can ever remove this indistinguishability.* This is the principle of indistinguishability of identical particles, which has many far-reaching consequences.

Consider two observers O and O' who prepare two physically identical quantum mechanical states of a system consisting of two identical particles, but adopt different conventions for labelling the electronic coordinates of the two particles. For example, O and O' measure the position of the particles at x_1 and x_2 or they prepare the scattering of two identical wave packets. Let $|\,\Psi\,\rangle$ be the state considered by O and $\Psi(x_1,x_2) = \langle\, x_1,x_2\,|\,\Psi\,\rangle$ the coordinate wave function. And let $|\,\Psi'\,\rangle$ be the state considered by O' with the coordinate wave function $\Psi'(x_1,x_2) = \langle\, x_1,x_2\,|\,\Psi'\,\rangle$. As stated above, the only difference between $|\,\Psi\,\rangle$ and $|\,\Psi'\,\rangle$ is the way in which the particles are labelled, i.e., $\Psi'(x_1,x_2) = \Psi(x_2,x_1)$.

The principle of indistinguishability of identical particles states that $|\,\Psi\,\rangle$ and $|\,\Psi'\,\rangle$ are equivalent representations of the same physical states with completely equivalent physical properties. Therefore, for any quantum mechanical state $|\,\beta\,\rangle$, the probability of finding the system described by $|\,\Psi\,\rangle$ or by $|\,\Psi'\,\rangle$ in the state $|\,\beta\,\rangle$ must be the same. Mathematically, this means that

$$|\langle\,\beta\,|\,\Psi\,\rangle|^2 = |\langle\,\beta\,|\,\Psi'\,\rangle|^2 \quad \forall\,|\,\beta\,\rangle.$$

In particular for $|\,\beta\,\rangle = |\,\Psi\,\rangle$ we have

$$|\langle\,\Psi\,|\,\Psi\,\rangle|^2 = 1 = |\langle\,\Psi\,|\,\Psi'\,\rangle|^2.$$

Consequently, taking into account that $\langle\,\Psi'\,|\,\Psi'\,\rangle = 1$, we must have

$$|\,\Psi'\,\rangle = e^{i\alpha}|\,\Psi\,\rangle.$$

In fact, two normalized states which overlap has the absolute value 1 can only differ by a multiplicative phase factor.[13] Then we have

$$\langle\, x_1,x_2\,|\,\Psi'\,\rangle = e^{i\alpha}\langle\, x_1,x_2\,|\,\Psi\,\rangle$$

---

[13] To prove this one may write $|\,\Psi'\,\rangle = a\,|\,\Psi\,\rangle + b\,|\,\Psi\,\rangle$ with $\langle\,\Psi\,|\,\Psi\,\rangle = 0$. This implies $|\langle\,\Psi\,|\,\Psi'\,\rangle|^2 = |a|^2 = 1$ and $\langle\,\Psi'\,|\,\Psi'\,\rangle = |a|^2 + |b|^2 = 1$. It follows that $|\,\Psi'\,\rangle = a\,|\,\Psi\,\rangle$ with $a = e^{i\,\alpha}$.

or

$$\Psi'(x_1, x_2) = e^{i\alpha}\, \Psi(x_1, x_2),$$

and thus

$$\Psi(x_2, x_1) = e^{i\alpha}\, \Psi(x_1, x_2).$$

By repeating the interchange we obtain

$$\Psi(x_1, x_2) = e^{i\alpha}\, \Psi(x_2, x_1) = e^{2i\alpha}\, \Psi(x_1, x_2),$$

which implies that $e^{2i\alpha} = 1$ or $e^{i\alpha} = \pm 1$. One concludes that

$$\boxed{\Psi(x_1, x_2) = \pm\Psi(x_2, x_1).}$$

In other words, the wave function of two identical particles is either *symmetric* or *antisymmetric* with respect to the interchange of the two coordinates.

The previous arguments can immediately be generalized to any two particles x_i, x_j in an N-particle system. Therefore,

$$\Psi(x_1, \ldots x_i, \ldots x_j, \ldots x_N) = \pm\Psi(x_1, \ldots x_j, \ldots x_i, \ldots x_N)$$

for all i and j. Since the particles are indistinguishable, the same sign holds for any two particles in the system. The wave function $\Psi(x_1, \ldots x_N)$ is either fully symmetrical or fully antisymmetrical with respect to the interchange of variables. The superposition of states with different symmetry is not possible, since the resulting wave function would neither be symmetrical nor antisymmetrical.

The particles in nature are thus divided in two disjoint groups. The particles having symmetrical wave functions are called *bosons* and are said to obey Bose-Einstein statistics. The particles with antisymmetrical wave functions are called *fermions* and are said to obey Fermi-Dirac statistics. The property of being a boson or a fermion is of course a fundamental property that depends on the nature of the particle. Experiment shows that there is a one-to-one correspondance between the fermionic or bosonic character and the intrinsic spin of the particles: Bosons are particles with integer spin, while Fermions are particles with half-integer spin. Most elementary particles are fermions (e^-, e^+, p, n). However, photons and a number of elementary excitations in condensed matter (phonons, magnons, etc.) are bosons.

In the case of complex particles (e.g., an α particle) the interchange of two particles can be regarded as the simultaneous interchange of its constituents. Therefore, the statistics of complex particles is fermionic if the number of elementary fermions in the particle is odd, or bosonic if the number of fermions is even. Thus, an α particle composed by two protons and two neutrons is a boson. ^{3}He atoms are fermions while ^{4}He atoms are bosons. This has crucial consequences on the low-temperature properties of these two isotopes. Notice that the integer/half-integer rule holds also for complex particles, since an even (odd) number of half-integer elementary particles corresponds to a total spin which is integer (half-integer).

To conclude this section, let us consider the effect of an arbitrary permutation on the coordinates. Let $P : [1, N] \to [1, N]$ be a permutation in the natural interval $[1, N]$ which we denote by

$$P = \begin{pmatrix} 1 & 2 & \dots & N \\ P(1) & P(2) & \dots & P(N) \end{pmatrix}$$

or simply $P = [P(1), P(2), \dots P(N)]$. The *order of the permutation* $O(P) = p$ is defined as the number of transpositions required to bring the sequence $[P(1), P(2), \dots P(N)]$ into the normal ordering $[1, 2, \dots N]$. For example,

$$P = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} \text{ has } p = 1, \quad \text{while} \quad P = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} \text{ has } p = 2.$$

It is then easy to see that for bosons

$$\Psi\left(x_{P(1)}, x_{P(2)}, \dots x_{P(N)}\right) = \Psi(x_1, x_2, \dots x_N), \tag{10.1}$$

while for fermions

$$\Psi\left(x_{P(1)}, x_{P(2)}, \dots x_{P(N)}\right) = (-1)^p \, \Psi(x_1, x_2, \dots x_N). \tag{10.2}$$

This fundamental property of the many-particle wave function [Eq. (10.1) or (10.2)] has many far-reaching consequences for the properties of matter. In the case of Fermions it leads to the Pauli exclusion principle with which the structure of Mendeleev's periodic table of the elements can be explained.

A. Systems of units

The cgs Gaussian system

Throughout this lecture and the associated exercises we will try to stick to the cgs (centimeter-gram-second) system in the Gaussian version for electromagnetic units. The Maxwell equations thus take the form

$$\vec{\nabla} \cdot \vec{D} = 4\pi\rho \qquad \vec{\nabla} \cdot \vec{B} = 0$$

$$\vec{\nabla} \times \vec{H} = \frac{4\pi}{c}\,\vec{J} + \frac{1}{c}\,\frac{\partial \vec{D}}{\partial t} \qquad \vec{\nabla} \times \vec{E} + \frac{1}{c}\,\frac{\partial \vec{B}}{\partial t} = 0.$$

The Lorentz force per unit charge is given by $\dfrac{\vec{F}}{q} = \vec{E} + \dfrac{\vec{v}}{c} \times \vec{B}$.

The following are the units corresponding to some important physical magnitudes:

[Energy] = ergs = erg
[Charge] = electrostatic unit = esu = statcoulomb = statC
1 Coulomb = 1 C = 3 × 10^9 statC
[Potential] = statV = 1 erg/esu
[Current] = statA = esu/s
[Current density] = statA/cm^2

Some physical constants in Gaussian units:

Bohr radius $a_0 = \dfrac{\hbar^2}{me^2} = 0.529$ Å $= 5.29 \times 10^{-9}$ cm
Electron mass $= m_e = 9.109 \times 10^{-28}$ g
Electron charge $= e = 4.803 \times 10^{-10}$ esu
Planck constant $\hbar = \dfrac{h}{2\pi} = 1.054 \times 10^{-27}$ erg × s
$c = 2.998 \times 10^{10}$ cm/s
The dimensionless fine structure constant is $\alpha = \dfrac{e^2}{\hbar c} \cong \dfrac{1}{137}$.

Atomic units

A very practical unit system for the calculations, particularly if one focuses on the properties of electrons, is the system of atomic units denoted by au. In this system the following fundamental physical constants are set equal to 1:

Electron charge $e = 1$ au
Electron mass $m_e = 1$ au
Planck's constant $\hbar = \dfrac{h}{2\pi} = 1$ au

The Bohr radius $a_0 = \dfrac{\hbar^2}{me^2} = 1$ au $= 5.29 \times 10^{-9}$ cm $= 0.529$ Å becomes the unit of length.

The Hartree, i.e., twice the ionization energy of the H atom, is given by

$$E_h = \frac{me^4}{\hbar^2} = \frac{e^2}{a_0} = \frac{\hbar^2}{m\,{a_0}^2} = 1 \text{ au} = 27,2 \text{ eV} = 4.36 \times 10^{-11} \text{ erg} = 2 \text{ Ry}$$

Finally, 1 eV $= 1.6 \times 10^{-19}$ J $= 1.6 \times 10^{-12}$ erg.

References

[1] K. Gottfried, Quantum Mechanics Volume I: Fundamentals (W. A. Benjamin Inc., New York, Amsterdam, 1966)

[2] L.D. Landau and E.M. Lifshitz, Quantum Mechanics (Non-relativistic Theory) (Butterworth-Heinemann, Oxford, 1958)

[3] C. Cohen-Tannoudji, B. Diu and F. Laloë, Quantenmechanik, Band 1 (De Gruyter, Berlin, 2008)

[4] G. B. Arfken and H. J. Weber, Mathematical Methods for Physicists (Elsevier, 2005)

[5] F. Schwabl, Quantenmechanik (QM I), (Springer, Berlin/Heidelberg/New York, 2005)

[6] A. L. Fetter and J. D. Wallecka, Quantum Theory of Many-particle systems (Mc Graw Hill, San Farncisco, 1971); Reprinted by Dover, Mineola, 2003.

[7] J.J. Sakurai, Modern Quantum Mechanics (Addison Wesley, 1985)

Acknowledgement
Die Vorbereitung dieses Skriptes ist nur durch die sorgfältige und interessierte Mitarbeit von Frau Andrea Wecker und Herr Sergej Riemer möglich gewesen, wofür der Autor und hoffentlich auch der Leser sehr dankbar sind.