QUANTUM PHYSICS

November 24, 2021

Preliminary version (not for circulation) of the notes
for the course "Quantum Physics"

V. Penna, Dipartimento di Fisica, Politecnico di Torino

COURSE PROGRAM

1 Review of results and methods of Newton's Mechanics

To better understand the novel character of the Lagrangian formulation of Classical Mechanics it is useful to review both the principles and the main results of the Newton approach. The latter allows one to predict the dynamical evolution of systems formed by many independent particles once the force fields acting on the system are known. Rigid bodies, that can be viewed as a special subclass of many-particle systems, and the study of their motions represent one of the most significant application of Newton's formulation. The latter consists of a series of general, fundamental theorems of classical Mechanics which, in addition to be valid for any many-particle system, have allowed to achieve a complete characterization of the dynamics of rigid bodies.

1.1 Newton's Laws and fundamental Theorems

The Newtonian formulation of classical Mechanics is based on the three, famous Newton Laws. In addition to the first law (the principle of Inertia: "The velocity of a body remains constant in the absence of external forces"), the **second Newton Law** states that

$$m\,\vec{a} = \vec{F}(\vec{r})$$

where

$$\vec{r} = x\vec{u}_x + y\vec{u}_y + z\vec{u}_z\,, \qquad \vec{a} = \frac{d^2\vec{r}}{dt^2} = \vec{u}_x\frac{d^2x}{dt^2} + \vec{u}_y\frac{d^2y}{dt^2} + \vec{u}_z\frac{d^2z}{dt^2}\,,$$

$\vec{u}_x$, $\vec{u}_y$, $\vec{u}_z$ being the unit vectors of a Cartesian reference frame. The remarkable **operational value** of the Second Newton Law becomes evident by observing how this law provides the differential equations governing the dynamics of a particle of mass m under the action of a given force field $\vec{F}(\vec{r})$

$$m\,\frac{d^2x}{dt^2} = F_x(x,y,z)\,, \quad m\,\frac{d^2y}{dt^2} = F_y(x,y,z)\,, \quad m\,\frac{d^2z}{dt^2} = F_z(x,y,z)\,.$$

The solutions of such second-order differential equations together with the knowledge of initial conditions $\vec{r}(0)$ (position vector at $t = 0$) and $\vec{v}(0)$ (velocity at $t = 0$) fully determine the behavior of the particle (its trajectory $\vec{r}(t)$ and thus its dynamical state) with the passing of time.

Third Newton's Law: the action of a mass m_1 on a mass m_2, represented by the force $\vec{F}_{12}$, is always equal and opposite the action of mass m_2 on mass m_1, the reaction represented by the force $\vec{F}_{21}$, namely $\vec{F}_{21} = -\vec{F}_{12}$.
The third Law allows one to obtain the equations determining the evolution of many-particle systems where each particle undergoes the action of the other particle and, possibly, of external force fields $\vec{F}^E(\vec{r})$

$$m_i\,\frac{d^2\vec{r}_i}{dt^2} = \vec{F}_i(\vec{r}_i)\,, \quad \vec{F}_i(\vec{r}_i) = \vec{F}_i^E(\vec{r}_i) + \sum\nolimits_{i\neq j}\vec{F}_{ji}$$

where internal forces usually have the form $\vec{F}_{ji} = \vec{F}(\vec{r}_i - \vec{r}_j)$. Vector $\vec{r}_i$ represents the position of the i-th particle. Complex systems such as planetary systems (e. g., the solar system), gas of molecules/atoms immersed in external magnetic and/or electric fields, and classical fluids are well described by such equations.

The main results of the Newton formulation of Classical Mechanics are contained and summarized in the following theorems

1) Equation of motion of the Center of Mass (CM): $\quad M\,\vec{A}_{cm} \,=\, \vec{F}^{E}$.

In the previous equation

$$\vec{F}^{E} = \sum_i \vec{F}_i^{E}(\vec{r}_i)\,, \quad \vec{A}_{cm} \,=\, \frac{d^2\vec{R}_{cm}}{dt^2} = \frac{1}{M}\sum_i m_i\,\vec{a}_i\,, \qquad \text{with} \quad \vec{R}_{cm} = \frac{1}{M}\sum_i m_i\,\vec{r}_i\,,$$

represent the external forces and the CM acceleration, respectively. The motion of the Center of Mass, described by $\vec{R}_{cm}(t)$, takes place without any influence of internal forces $\vec{F}_{ji}$. This result offers a considerable amount of information about the Center-of-Mass dynamics which appears to be totally independent from the dramatically complex motions of single particles governed by internal interactions. This is the first important result in the task of determining the ordered behaviors that characterize many-particle systems.

2) Equation of motion for the Angular Momentum of many-particle systems.

$$\frac{d\vec{L}}{dt} \,=\, \vec{M}^{E} - M\,\vec{V}_P \wedge \vec{V}_{cm}\,,$$

where $\vec{V}_P$ is the velocity of the moving pole P, $\vec{V}_{cm}$ the Center-of-Mass velocity

$$\vec{L} = \sum_i m_i\,\vec{r}_i \wedge \vec{v}_i\,, \quad \vec{M}^{E} = \sum_i \vec{M}_i^{E}\,.$$

The latter term $\vec{M}^{E}$ represents the momenta of external forces where $\vec{M}_i^{E} = \vec{r}_i \wedge \vec{F}_i^{E}$ and $\vec{r}_i$ is the position vector of mass m_i with respect to pole P. One should remind that velocities $\vec{v}_i$ are referred to the origin of the inertial reference frame which, in general, does not coincide with P. The main result of this theorem is that the dynamics of any rigid body is effectively described in terms of angular momentum by the previous equation. As regards a generic many-particle systems, similar to the Center-of-Mass theorem, the previous equation provides general information about the rotational activity of the system through $\vec{L}$.

3) Work integral and kinetic energy of a many-particle system. Theorem:

$$W_{AB} \equiv K_B - K_A\,.$$

The work from the space configuration A to configuration B and the kinetic energy are given by

$$W_{AB} = \sum_i W_{AB}(i) \,= \sum_i \int_{A\ \Gamma_i}^{B} \vec{F}_i \cdot d\vec{r}_i\,, \qquad K = \sum_i K(i) = \frac{1}{2}\sum_i m_i \vec{v}_i^{\,2}$$

respectively. Forces $\vec{F}_i$ include both external and internal forces. Forces are not necessarily conservative. The presence of Γ_i in the integrals reminds us that such works are, in general, dependent on path Γ_i. For each particle

$$\vec{F}_i = \vec{F}_i^E + \vec{F}_i^d + \sum_{j \neq i} \vec{F}_{ij}$$

where forces $\vec{F}_i^d$ are dissipative. Assuming that internal forces are **conservative** and that $\vec{F}_i^E$ represent the **conservative** contribution of external forces (part of the latter are incorporated in the dissipative terms $\vec{F}_i^d$), one has

$$\vec{F}_{ij} = \vec{F}(\vec{r}_j - \vec{r}_i) = -\nabla_{\vec{r}_j} U_{ij}\,, \qquad \vec{F}_i^E = -\nabla_{\vec{r}_i} U_i^E\,,$$

where $U_{ij} = U(\vec{r}_j - \vec{r}_i)$ and $U_i^E = U^E(\vec{r}_i)$ represent the corresponding potential energies.

The preceding three theorems lead to three fundamental results representing well-known **conservation laws** which, in addition to be extraordinarily useful tools both in the theoretic study of physical systems and in applications, have paved the way to further developments of the theory of dynamical systems. The most prominent example is the energy conservation of conservative systems which has provided the basis for developing the **Hamiltonian formulation of classical mechanics**.

1.2 Conservation Laws

Conservation of Momentum.

$$\vec{F}^E = 0 \;\Rightarrow\; M\,\vec{A}_{cm} = 0 \;\Rightarrow\; \frac{d}{dt} M\,\vec{V}_{cm} = 0 \;\Rightarrow\; M\,\vec{V}_{cm} \equiv \sum_i m_i\,\vec{v}_i = const$$

The quantity $\vec{P}_{tot} = \sum_i m_i\,\vec{v}_i$ represents the **total momentum** of the system which, in the absence of external forces, appears to be **conserved**. Despite the time evolution of single-particle velocities $\vec{v}_i$ (and thus of microscopic momenta $\vec{p}_i = m_i\vec{v}_i$) can be arbitrarily complex their summation is constant in time.
Another implication is that **velocity** $\vec{V}_{cm}$ is **constant**. As an example, one should remember that thanks to this theorem the description of solar system becomes extremely simple: the latter can be seen as an inertial system whose center, by means of further approximations, identifies with the sun. More ingeneral, the Conservation of Momentum represents the basis for studying any scattering process, elastic and inelastic collisions.

Conservation of Angular Momentum. Assume that pole P does not change its position in time ($\vec{V}_P = 0$). If the momentum of external forces $\vec{M}^E = 0$ then

$$\frac{d\vec{L}}{dt} = \vec{M}^E \;\Rightarrow\; \vec{L} = \sum_i m_i\,\vec{r}_i \wedge \vec{v}_i = const\,.$$

The second Kepler law determining the areal velocity of planets follows form this conservation law. Once more, we note how, Owing to $\vec{M}^E = 0$, in many-particle systems, the summation of single-particle contributions $\vec{L}_i = m_i\,\vec{r}_i \wedge \vec{v}_i$ is constrained to be constant even if each $\vec{L}_i$ features complex time evolutions governed by internal forces. This conservation law explains as well the properties of gyroscope, a mechanical device able to preserve its orientation independently form the motion of the platform on which is mounted.

Conservation of mechanical energy. In the absence of dissipative forces

$$W_{AB} = \sum_i W_{AB}(i) \,= \sum_i \int_{A\ \ \Gamma_i}^{B} \vec{F}_i \cdot d\vec{r}_i = -\Big(U_B - U_A \Big)$$

with the total potential energy

$$U \,= U^E + U^I = \sum_i U_i^E + \frac{1}{2} \sum_i \sum_{j \neq i} U_{ij}\,.$$

Then

$$K_B - K_A = -\Big(U_B - \, U_A \Big) \;\; \Rightarrow \quad K_B + U_B \,=\, K_A + U_A\,,$$

showing how the summation of the kinetic and potential energies corresponding to the initial configuration A (at time $t = t_A$) is equal to the summation of the kinetic and potential energies corresponding to the final configuration B (at time $t = t_B$). Then, at any time,

$$E = \, K + U \,=\, \frac{1}{2} \sum_i m_i v_i^2 \,+\, \sum_i U_i^E + \frac{1}{2} \sum_i \sum_{j \neq i} U_{ij} = \, const\,.$$

This probably represents one of the most important results of Classical Mechanics. The derivation of the Hamilton formulation of Classical Mechanics (which, in turn, is a key element in the foundation of quantum Mechanics) is rooted in this energy-conservation law.

2 The Lagrangian formulation of Classical Mechanics

2.1 Functional derivative and variational calculus

The (first) derivative of a function $f(x)$ represents the first contribution in the Taylor expansion

$$f(x+\Delta x) = f(x) + \frac{df}{dx}\,\Delta x + \frac{1}{2}\,\frac{d^2 f}{dx^2}\,\Delta x^2 + \ ...$$

where Δx is the deviation from the reference value x. The condition $df/dx = 0$ allows one to identifies the value of x correspondingto stationary points (minima, maxima, and saddles). The definition of function and, correspondingly, of derivative can be generalized to include functions depending on objects exhibiting a structure (such as curves or surfaces) certainly more complex than the simple scalar variable x occurring in $f(x)$ describing the points of real axis. As a simple example consider the work integral of a two-dimensional force field defined in the plane xy

$$W = \int_{A\ \ \Gamma}^{B} \vec{F}(\vec{r}) \cdot d\vec{r} = \int_{t_1}^{t_2} \vec{F}[\,\vec{r}(t)\,] \cdot \vec{v}(t)\,dt = \int_{t_1}^{t_2} \Big(F_x[x(t),y(t)]\,\dot{x}(t) \ + \ F_y[x(t),y(t)]\,\dot{y}(t) \Big)\,dt$$

where $\vec{r}(t) = x(t)\vec{u}_x + y(t)\vec{u}_y$ describes the curve Γ with $\vec{r}(t_1) = A$ and $\vec{r}(t_2) = B$. Work W **explicitly depends on** Γ through the parametric formulas $x(t)$ and $y(t)$ occurring in the position vector $\vec{r}(t)$. Information about Γ is encoded in the analytic form of $x(t)$ and $y(t)$. In this sense W is a functional depending both on $\vec{r}(t)$ and on $\vec{v}(t)$.

Consider a generic functional

$$F[\Gamma] = \int_{t_1}^{t_2} f(\,x(t), y(t), \dot{x}(t), \dot{y}(t))\,dt\,, \qquad \Gamma\ : \vec{r}_\Gamma(t) = \ x(t)\vec{u}_x + y(t)\vec{u}_y$$

with $\vec{r}_\Gamma(t_1) = A$ and $\vec{r}_\Gamma(t_2) = B$ are the initial and the final point, respectively, of a portion of Γ. A **small variation** of Γ is represented by a curve γ

$$\vec{r}_\gamma(t) = \ X(t)\,\vec{u}_x + Y(t)\,\vec{u}_y \ = \ \Big(x(t) + \delta x(t)\Big)\vec{u}_x + \ \Big(y(t) + \delta y(t)\Big)\vec{u}_y = \ \vec{r}_\Gamma(t) \ + \ \delta\vec{r}(t)$$

vector $\delta\vec{r} = \delta x\,\vec{u}_x + \ \delta y\,\vec{u}_y$ describes the deviations of γ from Γ. Of course, $\delta x = \epsilon a(t)$ and $\delta y = \epsilon b(t)$ where ϵ is small enough to ensure the two conditions $|\delta x| << |x|$ and $|\delta y| << |y|$ (γ is a small deformation of Γ). Deviations δx and δy plays the same role of the structureless deviation Δx of x in $f(x+\Delta x)$.

An important assumption is that deviations are zero in the two extreme points A (at $t = t_1$) and B (at $t = t_2$) of the arc AB of Γ. Explicitly,

$$\delta x(t_1) = \delta x(t_2) = 0\,, \qquad \delta y(t_1) = \delta y(t_2) = 0\,. \tag{1}$$

$$F[\gamma] = \int_{t_1}^{t_2} f(\,X(t), Y(t), \dot{X}(t), \dot{Y}(t))\, dt = \int_{t_1}^{t_2} f\Big(\,x(t) + \delta x,\, y(t) + \delta y,\, \dot{x}(t) + \delta\dot{x},\, \dot{y}(t) + \delta\dot{y}\Big)\, dt$$

Since δx, δy, $\delta\dot{x}$ and $\delta\dot{y}$ are small quantities

$$\begin{aligned} F[\gamma] &\simeq \int_{t_1}^{t_2} \left[f\Big(\,x(t),\, y(t),\, \dot{x}(t),\, \dot{y}(t)\Big) + \frac{\partial f}{\partial x}\,\delta x + \frac{\partial f}{\partial y}\,\delta y + \frac{\partial f}{\partial \dot{x}}\,\delta\dot{x} + \frac{\partial f}{\partial \dot{y}}\,\delta\dot{y} \right] dt \\ &= F[\Gamma] \,+ \int_{t_1}^{t_2} \left[\frac{\partial f}{\partial x}\,\delta x + \frac{\partial f}{\partial y}\,\delta y + \frac{\partial f}{\partial \dot{x}}\,\delta\dot{x} + \frac{\partial f}{\partial \dot{y}}\,\delta\dot{y} \right] dt \end{aligned}$$

To simplify the second term, whose counterpart for a simple function $f(x)$ is given by

$$\frac{df}{dx}\,\Delta x \quad \text{in} \quad f(x + \Delta x) = f(x) + \frac{df}{dx}\,\Delta x\,, \tag{2}$$

we observe that

$$\begin{aligned} \frac{\partial f}{\partial \dot{x}}\,\delta\dot{x} &= \frac{\partial f}{\partial \dot{x}}\,\delta\frac{d}{dt}x = \frac{\partial f}{\partial \dot{x}}\,\frac{d}{dt}\,\delta x = \frac{d}{dt}\left[\frac{\partial f}{\partial \dot{x}}\,\delta x\right] - \left[\frac{d}{dt}\frac{\partial f}{\partial \dot{x}}\right]\,\delta x\,, \\ \frac{\partial f}{\partial \dot{y}}\,\delta\dot{y} &= \,...\, = \frac{d}{dt}\left[\frac{\partial f}{\partial \dot{y}}\,\delta y\right] - \left[\frac{d}{dt}\frac{\partial f}{\partial \dot{y}}\right]\,\delta y\,. \qquad\qquad , \end{aligned}$$

The effect of the integral on such terms implies that

$$\begin{aligned} \int_{t_1}^{t_2} \left[\frac{\partial f}{\partial \dot{x}}\,\delta\dot{x} + \frac{\partial f}{\partial \dot{y}}\,\delta\dot{y}\right] dt &= \int_{t_1}^{t_2} \left[\frac{d}{dt}\left[\frac{\partial f}{\partial \dot{x}}\delta x\right] - \left[\frac{d}{dt}\frac{\partial f}{\partial \dot{x}}\right]\,\delta x + \frac{d}{dt}\left[\frac{\partial f}{\partial \dot{y}}\,\delta y\right] - \left[\frac{d}{dt}\frac{\partial f}{\partial \dot{y}}\right]\,\delta y\right] dt \\ &= \left[\frac{\partial f}{\partial \dot{x}}\,\delta x\right]_{t_1}^{t_2} + \left[\frac{\partial f}{\partial \dot{y}}\,\delta y\right]_{t_1}^{t_2} - \int_{t_1}^{t_2} \left[\left[\frac{d}{dt}\frac{\partial f}{\partial \dot{x}}\right]\,\delta x + \left[\frac{d}{dt}\frac{\partial f}{\partial \dot{y}}\right]\,\delta y\right] dt \\ &= -\int_{t_1}^{t_2} \left[\left[\frac{d}{dt}\frac{\partial f}{\partial \dot{x}}\right]\,\delta x + \left[\frac{d}{dt}\frac{\partial f}{\partial \dot{y}}\right]\,\delta y\right] dt \end{aligned}$$

owing to the boundary conditions (1). The final form of $F[\gamma]$ will be

$$F[\gamma] = F[\Gamma] + \int_{t_1}^{t_2} \left[\frac{\partial f}{\partial x}\,\delta x + \frac{\partial f}{\partial y}\,\delta y - \left[\frac{d}{dt}\frac{\partial f}{\partial \dot{x}}\right]\,\delta x - \left[\frac{d}{dt}\frac{\partial f}{\partial \dot{y}}\right]\,\delta y\right] dt$$

entailing the first-order variations

$$F[\gamma] - F[\Gamma] = \int_{t_1}^{t_2} \left[\left[\frac{\partial f}{\partial x} - \left(\frac{d}{dt}\frac{\partial f}{\partial \dot{x}}\right)\right]\,\delta x + \left[\frac{\partial f}{\partial y}\,\delta y - \left(\frac{d}{dt}\frac{\partial f}{\partial \dot{y}}\right)\right]\,\delta y\right] dt$$

By comparing such an expression with formula (2), where $\frac{df}{dx}$ is the coefficient of δx in $f(x + \Delta x)$, we find that

$$\frac{\delta F}{\delta x} = \frac{\partial f}{\partial x} - \left(\frac{d}{dt}\frac{\partial f}{\partial \dot{x}}\right)\,, \qquad \frac{\delta F}{\delta y} = \frac{\partial f}{\partial y} - \left(\frac{d}{dt}\frac{\partial f}{\partial \dot{y}}\right) \tag{3}$$

represent the **functional derivatives** of $F[\gamma]$ these corresponding to the coefficients of elementary deviations δx and δy. Such formulas reproduce the standard definition of functional derivative.

Example 1. Find the path of minimum length connecting two points A and B on the cartesian plane. Of course, we expect that such a path corresponds to a straight line. Suppose that $\vec{r} = x(t)\,\vec{u}_x + y(t)\,\vec{u}_y$ describes a planar curve γ parametrized by t and connecting A and B. Based on the definition of arc length, one has

$$L(A,B) = \int_{A\ \gamma}^{B} ds \;=\; \int_{t_A\ \gamma}^{t_B} v(t)\,dt \;=\; \int_{t_A\ \gamma}^{t_B} \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2}\,dt$$

$R(\dot{x};\dot{y}) = \sqrt{\dot{x}^2+\dot{y}^2} = \sqrt{\vec{v}^{\,2}} = \; v(t)$

$$0 = \frac{\partial R}{\partial x} - \left(\frac{d}{dt}\frac{\partial R}{\partial \dot{x}}\right) = -\left(\frac{d}{dt}\frac{\partial R}{\partial \dot{x}}\right) = -\frac{d}{dt}\frac{\dot{x}}{R} \;\;\Rightarrow\;\; \frac{\dot{x}}{R} = c_1 = const$$

$$0 = \frac{\partial R}{\partial y} - \left(\frac{d}{dt}\frac{\partial R}{\partial \dot{y}}\right) = -\left(\frac{d}{dt}\frac{\partial R}{\partial \dot{y}}\right) = -\frac{d}{dt}\frac{\dot{y}}{R} \;\;\Rightarrow\;\; \frac{\dot{y}}{R} = c_2 = const$$

Since

$$c_1^2 + c_2^2 = \frac{\dot{x}^2}{R^2} \;+\; \frac{\dot{y}^2}{R^2} = 1\,,$$

then

$$c_1 = \cos\alpha\,, \quad c_2 = \sin\alpha\,.$$

Integrating in time

$$\dot{x} = \; R(t)\,c_1 = \; R(t)\;\cos\alpha\,, \quad \dot{y} = \; R(t)\,c_2 = \; R(t)\;\sin\alpha\,,$$

gives the solutions

$$x(t) = x_0 + f(t)\;\cos\alpha\,, \quad y(t) = y_0 + f(t)\;\sin\alpha\,, \qquad \text{with } \dot{f} = R(t)$$

which describes a straight line able to join $A = (x_0, y_0)$ and B when a suitable choice of α is made. To select a straight line of this type, one must introduce the further conditions that $f(t)$ is an increasing function of t, $f(t_1) = 0$ (entailing that $x(t_1) = x_0$, $y(t_1) = y_0$) and $f(t_2)$ such that $x(t_2) = x_0 + f(t_2)\;\cos\alpha \equiv x_B$, $y(t_2) = y_0 + f(t_2)\;\cos\alpha \equiv= y_B$.

Example 2. A constant function is such that $df(x)/dx = 0$. The work integral of a conservative force is the functional counterpart of a constant function in that, by definition, the work integral must be independent on the path γ. It is interesting to discover what conditions one finds by imposing

$$\frac{\delta W}{\delta x} = 0\,, \quad \frac{\delta W}{\delta y} = 0\,, \quad \frac{\delta W}{\delta z} = 0\,,$$

on the functional

$$W = \int_{A\ \ \Gamma}^{B} \vec{F}(\vec{r}) \cdot d\vec{r} = \int_{t_1}^{t_2} \Big(F_x(x,y,z)\,\dot{x}(t) \ + \ F_y(x,y,z)\,\dot{y}(t) + F_y(x,y,z)\,\dot{z}(t) \Big)\, dt\,.$$

$$\frac{\delta W}{\delta x} = \frac{\partial}{\partial x}\Big(F_x\dot{x} \ + \ F_y\dot{y} + F_z\dot{z} \Big) - \frac{d}{dt}\frac{\partial}{\partial \dot{x}}\Big(F_x\dot{x} \ + \ F_y\dot{y} + F_z\dot{z} \Big)$$

$$= \dot{x}\frac{\partial F_x}{\partial x} + \ \dot{y}\frac{\partial F_y}{\partial x} + \dot{z}\frac{\partial F_z}{\partial x} - \frac{d}{dt}F_x = \dot{x}\frac{\partial F_x}{\partial x} + \ \dot{y}\frac{\partial F_y}{\partial x} + \dot{z}\frac{\partial F_z}{\partial x} - \dot{x}\frac{\partial F_x}{\partial x} - \dot{y}\frac{\partial F_x}{\partial y} - \dot{z}\frac{\partial F_x}{\partial z}$$

$$= \dot{y}\frac{\partial F_y}{\partial x} + \dot{z}\frac{\partial F_z}{\partial x} - \dot{y}\frac{\partial F_x}{\partial y} - \dot{z}\frac{\partial F_x}{\partial z} = \dot{y}\left(\frac{\partial F_y}{\partial x} - \frac{\partial F_x}{\partial y}\right) + \dot{z}\left(\frac{\partial F_z}{\partial x} - \frac{\partial F_x}{\partial z}\right)$$

One has

$$\frac{\delta W}{\delta x} = \dot{y}\left(\frac{\partial F_y}{\partial x} - \frac{\partial F_x}{\partial y}\right) + \dot{z}\left(\frac{\partial F_z}{\partial x} - \frac{\partial F_x}{\partial z}\right) = 0$$

Likewise, the variations with respect to y and z give

$$\frac{\delta W}{\delta y} = \ ... \ = \dot{z}\left(\frac{\partial F_z}{\partial y} - \frac{\partial F_y}{\partial z}\right) + \dot{x}\left(\frac{\partial F_x}{\partial y} - \frac{\partial F_y}{\partial x}\right) = 0$$

$$\frac{\delta W}{\delta z} = \ ... \ = \dot{x}\left(\frac{\partial F_x}{\partial z} - \frac{\partial F_z}{\partial y}\right) + \dot{y}\left(\frac{\partial F_y}{\partial z} - \frac{\partial F_z}{\partial y}\right) = 0$$

Since velocity components $\dot{x}$, $\dot{y}$, and $\dot{z}$ are arbitrary quantities the only way to satisfy the previous conditions is to impose

$$\frac{\partial F_x}{\partial z} - \frac{\partial F_z}{\partial y} = 0\,, \qquad \frac{\partial F_y}{\partial z} - \frac{\partial F_z}{\partial y} = 0\,, \qquad \frac{\partial F_y}{\partial x} - \frac{\partial F_x}{\partial y} = 0$$

These equations are the well-known conditions within vector analysis: they are equivalently expressed by means of operator curl through the formula $\mathrm{curl}\vec{F} = 0$. Such a **condition** implies that field $\vec{F}$ is **conservative** namely $\vec{F} = -\nabla U$. If this is the case, in fact, one has $\mathrm{curl}\vec{F} = \mathrm{curl}(-\nabla U) = 0$ regardless the scalar field U generating $\vec{F}$.

2.2 Generalized coordinates and Configuration Space

In the Newtonian formulation of Classical Mechanics the **dynamical state** of a particle at time t is determined once the position vector and the velocity

$$\vec{r}(t) = x(t)\,\vec{u}_x + y(t)\vec{u}_y + z(t)\vec{u}_z\,, \qquad \vec{v}(t) = \frac{d\,\vec{r}}{dt} = \dot{x}(t)\,\vec{u}_x + \dot{y}(t)\vec{u}_y + \dot{z}(t)\vec{u}_z\,,$$

respectively, are known at time t. Momentum $\vec{p}(t) = m\,\vec{v}(t)$ can be also used instead of $\vec{v}(t)$. Then, in the presence of N independent particles, the dynamical state is given by

$$\vec{r}_1(t),\, \vec{r}_2(t),\, ...\,, \vec{r}_N(t), \qquad \vec{p}_1(t),\, \vec{p}_2(t),\, ...\,, \vec{p}_N(t),$$

involving a set of $3N + 3N$ independent variables. The study of specific dynamical problems, however, shows that, in many cases, such a (cartesian) vector description is neither effective nor simple. The number of variables actually necessary to ensure a correct description of a dynamical system can be much smaller than $6N$ even if the systems involves N particles.

In general, this is due the presence of **internal constraints** that embody the geometric intrinsic properties of the system. Significant examples can be found by considering rigid bodies in which the distance between any pair of particles (viewed as elementary masses constituting the body) is fixed. The new coordinates of such a restricted set are named **generalized coordinates**.

Example 1. Consider a (homogeneous) thin rod with mass M and length R. The rod is totally free. Its motion, in general, will be the combination of a free translation in the 3D ambient space (orbital motion) and of a rotation of the rod around its center of mass. In principle, the position of the rod is specified through the position vectors of the two extremes

$$\vec{r}_1 = x_1\vec{u}_x + y_1\vec{u}_y + z_1\vec{u}_z\,, \qquad \vec{r}_2 = x_2\vec{u}_x + y_2\vec{u}_y + z_2\vec{u}_z\,.$$

Altogether, **six cartesian coordinates** are involved, **three** for each extreme. Since the rod is a rigid body the **number of variables actually necessary** to determine the rod position **are only five** its length

$$(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2 \equiv R^2 = \textit{constant}\,.$$

being a constant of motion. These are three cartesian coordinates X, Y, and Z, determining the center-of-mass position $\vec{R} = X\,\vec{u}_x + Y\,\vec{u}_y + Z\,\vec{u}_z$, and two angles ϕ and θ, determining the rod orientation in the center-of-mass reference. This can be easily proven by defining the unit vector $\vec{u}$ integral with the rod motion and parallel to the rod. The position

$$\vec{r}_i = r_i\,\vec{u}(\phi,\theta)\,, \qquad \vec{u}(\phi,\theta) = \vec{u}_x\cos\phi\,\sin\theta + \vec{u}_y\sin\phi\,\sin\theta + \vec{u}_z\cos\theta$$

of each particle of mass m_i forming the rod (imagine to divide the rod in many point-like elementary portions labeled by i) is fully determined by ϕ and θ. Then

$$X\,, \quad Y\,, \quad Z\,, \quad \phi\,, \quad \theta\,,$$

are the five **generalized coordinates** sufficient to fix the space configuration of the rod.

Example 2. Consider a generic rigid body (RB) with mass M. To characterize its motion consider two distinct points $\vec{r}_1$ and $\vec{r}_2$ such that the center of mass is at some point $\vec{R}$ along the straight line connecting $\vec{r}_1$ and $\vec{r}_2$. Note that the portion of the body corresponding to the elementary masses placed on this straight line can be seen as an ideal thin rod, perfectly equivalent to the one of Example 1.
The RB motion is, in general, the combination of a free translation of the center of mass in the 3D ambient space (orbital motion) and of a rotation of our "ideal" rod around the RB center of mass. A third angle ψ, however, is necessary to describe the motion of the remaining masses of the RB placed around the ideal rod. Then

$$X\,, \quad Y\,, \quad Z\,, \quad \phi\,, \quad \theta\,, \quad \psi\,,$$

are the six **generalized coordinates** sufficient to fix the space configuration of the RB.

Comment. The number of **generalized coordinates** characterizing a system can undergo further restriction when further constraints are involved in the dynamics.

Example 3. Suppose that the thin rod of Example 1 is no longer free to translate because a pivot has been placed in the point corresponding to the center of mass. This amounts to imposing the **three constraints** $X = Y = Z = 0$. The rod motion is thus described by **two generalized coordinates**, the angles ϕ and θ.

Example 4. Consider a spherical physical pendulum represented by a thin homogeneous rod with mass M, length L and a pivot in one of its extremes. In this case one of the two positions $\vec{r}_1$ and $\vec{r}_2$ of the rod extremes (see Example 1) is fixed. Assume for simplicity that the first extreme is placed at the origing ($\vec{r}_1 = 0$) of a given cartesian reference. Then the position of the center of mass

$$\vec{R} = X\,\vec{u}_x + Y\,\vec{u}_y + Z\,\vec{u}_z$$

is sufficient to fix the rod position in the space. In this sense X, Y and Z represent generalized coordinates with respect to the six coordinates involved by $\vec{r}_1$ and $\vec{r}_2$. An equivalent and more advantageous set is that of spherical coordinates R, ϕ, α:

$$\vec{R} = R\,\vec{u}_R(\phi,\alpha)\,, \quad R = |\vec{R}|\,, \quad \vec{u}_R(\phi,\alpha) = \vec{u}_x \cos\phi\, \sin\alpha + \vec{u}_y \sin\phi\, \sin\alpha - \vec{u}_z \cos\alpha$$

where unit vector $\vec{u}_R$ is attached and parallel to the rod and ϕ describes the rod position around the vertical axis Z. Angle α is the angle between the rod and Z: for $\alpha = 0$ the free extreme reachs the lowest position (α differs from $\theta = \pi - \alpha$ of Examples 1-3). Since the rod is rigid one must take into account the **further constraint**

$$R := \sqrt{X^2 + Y^2 + Z^2} \equiv \frac{L}{2} \qquad (L \,\text{is constant in time})$$

thus reducing the number of generalized coordinates involved in the dynamics. Coordinates ϕ and α surviving such a reduction are named **degrees of freedom**.

Example 5. Consider a particle confined in a thin toroidal domain of radius R in the plane xy. Its motion takes place on a circle. In this case the generalized coordinates are the cylindrical coordinates ρ, ϕ and z such that $x = \rho\cos\phi$ and $y = \rho\sin\phi$. Since the particle is subject to two **constraints**

$$z = 0\,, \qquad \rho = \sqrt{x^2 + y^2} = R$$

the three coordinates **reduce to a unique coordinate**, the angle ϕ, and

$$\vec{r}(\phi) = \rho\,(\cos\phi\,\vec{u}_x + \sin\phi\,\vec{u}_y)\,.$$

Example 6. The two-body system. This example is very interesting because makes it evident how the presence of **conserved quantitites** (the so-called constants of motion) can simplify the description of the system by **reducing the number of degrees of freedom** actively partecipating in the dynamics.

Two vectors $\vec{r}_i = x_i\,\vec{u}_x + y_i\vec{u}_y + z_i\vec{u}_z$ with $i = 1, 2$ indicate the particle positions then involving 6 cartesian coordinates. Generalized coordinates are found when considering the relative-position and center-of-mass vectors

$$\vec{r}_1 = \vec{R} + \frac{m_2}{M}\,\vec{r}\,, \quad \vec{r}_2 = \vec{R} - \frac{m_1}{M}\,\vec{r} \quad \Leftrightarrow \quad \vec{r} = \vec{r}_1 - \vec{r}_2\,, \quad \vec{R} = \frac{m_1\vec{r}_1 + m_2\vec{r}_2}{M}\,.$$

The new coordinates are

$$X\,, \quad Y\,, \quad Z\,, \quad (\text{relevant to}\,\tilde{\mathrm{R}}) \qquad\qquad \mathrm{x}\,, \quad \mathrm{y}\,, \quad \mathrm{z}\,, \quad (\text{relevant to}\,\tilde{\mathrm{r}})\,.$$

The force fields acting on m_1 and m_2 (such masses could represent a planet and the Sun) reduce to the mutual gravitational interaction

$$\vec{F}_{12} = -\gamma m_1 m_2 \frac{\vec{r}_2 - \vec{r}_1}{|\vec{r}_1 - \vec{r}_2|^3} = +\gamma m_1 m_2 \frac{\vec{r}}{r^3}\,, \qquad \vec{F}_{21} = -\gamma m_1 m_2 \frac{\vec{r}_1 - \vec{r}_2}{|\vec{r}_1 - \vec{r}_2|^3} = -\gamma m_1 m_2 \frac{\vec{r}}{r^3}$$

which, being internal forces, satisfy the condition $\vec{F}_{12} = -\vec{F}_{21}$. Then, owing to the theorem on the center-of-mass motion (and to its corollary on total momentum), equation $\vec{F}_{12} + \vec{F}_{21} = 0$ entails that $\vec{P} = \vec{p}_1 + \vec{p}_2 = M\,\vec{V}$ is a constant of motion

$$P_X = const\,, \quad P_Y = const\,, \quad P_Z = const \quad \Leftrightarrow \quad V_X = const\,, \quad V_Y = const\,, \quad V_Z = const\,.$$

Due to the **absence of external forces**, the center of mass displays a uniform motion with a constant $\vec{V}$ and $\vec{R}(t) = \vec{R}_0 + \vec{V}\,t$ showing how the **two-body system is inertial**. Hence, generalized coordinates X, Y and Z, the components of $\vec{R}$, can be eliminated by assuming a reference frame attached to the center of mass.

From this new standpoint, the generalized coordinates of the two-body problem reduce to x, y and z, the components of relative-position vector $\vec{r}$. Another fundamental conservation law further reduces the coordinate number. In the absence of external forces (if the internal ones are central forces, as in the present case) the **total angular momentum is conserved**. Then the motion of the two masses must take place on a common plane. One can choose a reference frame that coincides with plane xy. This choice is equivalent to setting $z \equiv 0$. **Only two generalized coordinates** survives at this point: the polar coordinates r, ϕ of plane xy.

Example 7. The double pendulum. This mechanical device can be constructed by means two thin rods. An extreme of the first rod is attached to a fixed pivot, while the second extreme is jointed to an extreme of the second rod. The system moves under the action of gravity. Suppose that the motion takes place in the plane xy. In this case two generalized coordinates, the angle θ describing the position of the first rod (as for a simple pendulum) and an angle β describing the slope of the second rod with respect to the firts one. One easily shows that such two angles result from a reduction process of four (cartesian) coordinates: two of these are relevant to the position vector of the first-rod free extreme, and the other two relevant to the vector describing the second-rod center of mass.

Conclusions. Generalized coordinates, conventionally denoted by

$$q_1\,,\ q_2\,,\ \cdots,\ q_M\,,$$

provide an alternative, more effective description to the one in terms of Cartesian coordinates. Their number M can be much smaller than $3N$, the number of Cartesian coordinates involved by a system of N particles. In this perspective, the (Cartesian) position vectors and velocities appear to be quantities depending on q_i and $\dot{q}_i$

$$\vec{r}_i = \vec{r}_i(q_1, q_2, \ldots q_M)\,, \qquad \vec{v}_i = \vec{v}_i(q_1, q_2, \ldots q_M, \dot{q}_1, \dot{q}_2, \ldots \dot{q}_M)\,.$$

Of course, if the system is formed by N **free** particles, totally independent the one from the other, one has $M \equiv 3N$. Conversely, 6 coordinates are sufficient to describe the configurations of a N-particle rigid body where $3N >> M = 6$. The choice of such new coordinates allows one to take into account the geometric features of the systems leading to a simpler description of motion. Rigid bodies well represent this situation.
Likewise, generalized coordinates appear to be extremely useful in the presence of **constraints**. The latter are typically represented by a set of equations

$$f_r(q_1, q_2, \ldots q_M) = 0 \qquad \text{with } r \in [1, N_c]$$

In this case the constraints are said to be **holonomic**. Generalized coordinates often give the possibility to embody in a simple way the constraints in the dynamics of the system. To

this end one must solve equations $f_r = 0$. This amounts to rewriting them in the form of N_c relations $q_r = g_r(\{q_\alpha\})$ where $\{q_\alpha\}$ indicates a restricted set of S coordinates $q_1, q_2, ...q_s$.

The number S of generalized coordinates which **can vary independently** in the presence of constraints is called the **number degrees of freedom** of the system.

Definition. A system is said to be **holonomic** if the constraint equations $f_r = 0$ can be solved. In this case the number S of surviving coordinates is equal to the number of degrees of freedom namely $S = M - N_c$ where N_c is the number of constrains.

Example. In example 1, the **thin free rod** in the 3-dimensional space can described by 3+3 Cartesian coordinates which are constrained by the rod length R. A simpler description is found by considering six generalized coordinates (X, Y, Z) (center-of-mass coordinates) and (R, ϕ, θ) (spherical coordinates describing the rod in with respect to the center of mass). However, $5 = 6 - 1$ generalized coordinates (X, Y, Z, ϕ, θ) are sufficient due to the constraint $f(R, \phi, \theta) = R^2 - L^2/4 = 0$.

In example 3, the three constraints

$$f_1(X, Y, Z, \phi, \theta) := X = 0\,, \quad f_2(X, Y, Z, \phi, \theta) := Y = 0\,, \quad f_3(X, Y, Z, \phi, \theta) := Z = 0\,,$$

have been imposed (in addition to the rod-length $R^2 - L^2/4 = 0$) which eliminate the motion of the center of mass. 6 generalized coordinates for the free rod have been reduced to $6 - 3 - 1 = 2$ coordinates (**the effective degrees of freedom**) owing to 4 constraints.

Example. In example 4, the center of mass of the **spherical pendulum** is described by X, Y and Z which represent a possible choice of generalized coordinates. These are sufficient to determine the rod position in the space. Since the rod is a rigid body the further constraint can be included

$$f(X, Y, Z) = X^2 + Y^2 + Z^2 - L^2/4 = 0 \quad \Rightarrow \quad Z = \pm\sqrt{L^2/4 - (X^2 + Y^2)}$$

where $|\vec{R}| = L/2$ and L the rod length. The second eqaution shows that the constraint equation can be **solved** and Z turns out to be depending on X and Y. The number of degrees of freedom is thus $3 - 1 = 2$ corresponding to X and Y. Another possible choice is to identify **generalized coordinates** with **spherical coordinates**. Then the rod position can be described by means of spherical coordinates R, ϕ and α through

$$\vec{R} = R\,\vec{u}_R(\phi, \alpha) \;\; \rightarrow \quad X = R\cos\phi\,\sin\alpha\,, \;\; Y = R\sin\phi\,\sin\alpha\,, \;\; Z = -R\,\cos\alpha\,.$$

In this case the previous constraint assumes the simple form $f(R, \phi, \alpha) = R^2 - L^2/4 = 0$ showing that the number of **degrees of freedom** is $3 - 1 = 2$, the angles ϕ, α.

All the systems investigated in the sequel belong to the class of **holonomic systems**.

Note that a rigid body formed by N particles represent an intrinsically constrained system in that the equations fixing the distance d_{in} between particles $|\vec{r}_i - \vec{r}_n|^2 - d_{in}^2 = 0$ are nothing but constraints. Of course, the number of constraints is of the order of the Avogadro number. Such **internal** constraints determine the geometry and the mass distribution of the body.

Configuration space. This space, denoted by $\mathcal{C}$, describes any possible configuration of a given system through the set of the generalized coordinates. The structure of the configuration space thus reflects, through its coordinates, the geometric features of the body (a simple particle, a rod, a generic 3D rigid body, a body with symmetric shape, and so on) and the presence of possible constraints. In the following list the configuration spaces of the preceding seven examples are explicitly defined. The manifold including $\mathcal{C}$, written on the right, represent the initial enlaraged space describing the system configurations before imposing the constraints.

Case 1 $(X, Y, Z, \phi, \theta) = (X, Y, Z) \times (\phi, \theta) \in \mathcal{C}_1 = \mathbb{R}^3 \times \mathbb{S}^2 \subset \mathbb{R}^3 \times \mathbb{R}^3 = \mathbb{R}^6$

Case 2 $(X, Y, Z, \phi, \theta, \psi) = (X, Y, Z) \times (\phi, \theta) \times \psi \in \mathcal{C}_2 = \mathbb{R}^3 \times \mathbb{S}^2 \times \mathbb{S} \subset \mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}^2 = \mathbb{R}^8$

Case 3 $(\phi, \theta) \in \mathcal{C}_3 = \mathbb{S}^2 \subset \mathbb{R}^3$

Case 4 $(\phi, \alpha) \in \mathcal{C}_4 = \mathbb{S}^2 \subset \mathbb{R}^3$

Case 5 $(\phi) \in \mathcal{C}_5 = \mathbb{S}^1 \subset \mathbb{R}^3$

Case 6 $(r, \phi) \in \mathcal{C}_6 = \mathbb{R}^2 \subset \mathbb{R}^3 = \{(x, y, z)\}$

Case 7 $(\phi, \beta) \in \mathcal{C}_7 = \mathbb{S}^1 \times \mathbb{S}^1 = \mathbb{T}^2 \subset \mathbb{R}^2 \times \mathbb{R}^2 = \mathbb{R}^4$

2.3 The Lagrangian formulation

The use of the so-called **generalized coordinates** denoted by q_1, q_2, ... q_N and of the relevant **generalized velocities** $\dot{q}_1$, $\dot{q}_2$, ..., $\dot{q}_N$ in place of the usual cartesian position vectors $\vec{r}_i$ and velocities $\vec{v}_i$ (with $i \in [1, N]$ allows one to take into account in an explicit way the intrinsic constraints possibly conditioning the evolution of physical systems. This representation of system configurations extremely simplifies the derivation of dynamical equations.

Coordinates q_i minimize the amount of information necessary to describe the system in that they reflect the effective geometry in which the dynamics takes place. The recognition of the effective degrees of freedom of a system is thus equivalent to incorporating the its constraints. Let us consider some examples based on those discussed in the previous subsection.

1) Single particle of mass m confined in a very thin toroidal domain. This is the system relevant to **Example 5**. The effective configuration space reduces to a **ring** of radius R representing the **restricted portion of the 3D space** where m stands. The position vector and the velocity are given by

$$\vec{r}(\theta) = x(\theta)\,\vec{u}_x + y(\theta)\vec{u}_y\,, \qquad \vec{v}(\theta,\dot{\theta}) = \frac{d\,\vec{r}}{dt} = \dot{\theta}\,(\,-y(\theta)\,\vec{u}_x + x(\theta)\vec{u}_y\,)$$

respectively, where $x = R\cos\theta$ and $y = R\sin\theta$. To describe the position and velocity of m it is sufficient to know, at each time, $\phi(t)$ and $\dot{\phi}(t)$. Such variables allow one to describe any physical quantity of interest

$$E = \frac{m\vec{v}^{\,2}}{2} + U(\vec{r}) = \frac{mR^2}{2}\,\dot{\theta}^2 + W(\theta)\,, \quad \vec{L} = m\vec{r}\wedge\vec{v} = mR^2\dot{\theta}\,\vec{u}_z\,,$$

where $W(\theta) = U\big(\vec{r}(\theta)\big)$ is the energy potential along the ring.

2) Physical pendulum represented by a homogeneous rod with mass M, length L and the pivot placed in one of its extremes. This system is obtained from **Example 4** by setting $\phi \equiv \pi/2$. Then oscillations take place in the vertical plane yz. Let m_i be the elementary masses forming the rod. The kinetic and potential energies depend in a simple way on a single coordinate (the angle α between the rod and the vertical axis z on which the pivot stands) and on the relevant generalized velocity $\dot{\alpha}$

$$\phi \equiv \pi/2 \;\;\Rightarrow\;\; \vec{r}_i = r_i\,\vec{u}_R(\pi/2,\alpha) = r_i\,(\vec{u}_y\sin\alpha\,-\vec{u}_z\cos\alpha)\,, \quad \vec{v}_i = r_i\,\dot{\alpha}\,(\,\vec{u}_y\cos\alpha\,+\vec{u}_z\sin\alpha\,)$$

$$K = \sum_i \frac{m_i}{2}\vec{v}_i^{\,2} = \frac{I}{2}\,\dot{\alpha}^2\,, \quad I = \sum_i m_i\,r_i^2\,, \qquad U(\vec{R}) = -M\vec{g}\cdot\vec{R} = M\,g\,Z = -M\,g\,\frac{L}{2}\,\cos\alpha\,.$$

Note that in U vector $\vec{R} = (X, Y, Z)$, $R = |\vec{R}| = L/2$ is the position of the center of mass (corresponding the rod center), $Z = -R\cos\alpha$ and $\vec{g} = -g\,\vec{u}_z$ is the gravitational field .

3) Free rod in the gravitational field. This is the system discussed in **Example 1**. Its motion, in general, is the combination of the free translation in the 3D ambient space (orbital motion) and of rotations around its center of mass described by (X, Y, Z, ϕ, θ). For each point of the rod

$$\vec{r_i} = r_i\,\vec{u}(\phi,\theta)\,, \qquad \vec{u}(\phi,\theta) = \vec{u}_x\cos\phi\,\sin\theta + \vec{u}_y\sin\phi\,\sin\theta + \vec{u}_z\cos\theta$$

(where $\vec{u}$ is the unit vector attached and parallel to the rod). The rotational kinetic energy reads

$$K_{rot} = \sum_i \frac{m_i}{2}\vec{v}_i{}^2 = \sum_i \frac{m_i}{2} r_i^2 \left(\frac{d\vec{u}}{dt}\right)^2 = \ldots = \frac{I}{2}\,(\,\dot\theta^2 + \dot\phi^2\sin^2\theta)$$

while the contribution due to the translation motion of the center of mass, placed at $\vec{R} = (X, Y, Z)$, is given by

$$K_{tr} = \frac{M}{2}\vec{V}^2 = \frac{M}{2}\left(\dot X^2 + \dot Y^2 + \dot Z^2\right).$$

The potential energy is simply $U(\vec{R}) = -M\vec{g}\cdot\vec{R} = M\,g\,Z$.

4) Two-body problem. This is the system discussed in **Example 6**. Let us derive the angular momentum and the potential and kinetic energies expressed in terms of generalized coordinates and velocities. By observing that

$$\vec{r}_1 = \vec{R} + \frac{m_2}{M}\,\vec{r}, \quad \vec{r}_2 = \vec{R} - \frac{m_1}{M}\,\vec{r}, \quad \Rightarrow \vec{v}_1 = \vec{V} + \frac{m_2}{M}\,\vec{v}, \quad \vec{v}_2 = \vec{V} - \frac{m_1}{M}\,\vec{v}\,,$$

where $\vec{V} = d\vec{R}/dt$ is the center-of-mass velocity and $\vec{v}$ is the relative velocity, the kinetic energy can be rewritten as

$$K = \frac{m_1}{2}\,\vec{v}_1^2 + \frac{m_2}{2}\,\vec{v}_2^2 = \frac{m_1+m_2}{2}\,\vec{V}^2 + \frac{\mu}{2}\,\vec{v}^2 = \frac{m_1+m_2}{2}\,\vec{V}^2 + \frac{\mu}{2}\,(\dot r^2 + r^2\dot\phi^2)$$

(the two terms of the latter formula, as expected, obeys the first König theorem) thanks to the polar-coordinate picture of planar position vector $\vec{r}$

$$\vec{r} = r\,\vec{u}_r(\phi)\,, \;\; \vec{u}_r(\phi) = \vec{u}_x\cos\phi + \vec{u}_y\sin\phi\,, \qquad \vec{v} = \vec{u}_r\,\dot r + \vec{u}_\phi\,r\,\dot\phi\,.$$

The total energy is

$$E = \frac{m_1+m_2}{2}\,\vec{V}^2 + \frac{\mu}{2}\,(\dot r^2 + r^2\dot\phi^2) - \frac{\gamma\,m_1\,m_2}{r}\,,$$

while the angular momentum reads

$$\vec{L} = m_1\vec{r}_1\wedge\vec{v}_1 + m_2\vec{r}_2\wedge\vec{v}_2 = \vec{R}\wedge(m_1\vec{v}_1 + m_2\vec{v}_2) + \vec{r}\wedge\left(m_1\frac{m_2}{M}\vec{v}_1 - m_2\frac{m_1}{M}\vec{v}_2\right)$$

$$= \vec{R}\wedge\,M\,\vec{V} + \mu\,\vec{r}\wedge\vec{v} = \vec{R}\wedge\,M\,\vec{V} + \mu\,r^2\dot\phi\,\vec{u}_z\,, \qquad \mu = \frac{m_2 m_1}{m_1+m_2} \quad \text{(reduced mass)}\,,$$

(the two terms in the final formula, as expected, obeys the second König theorem). All such quantities depends on X, Y, Z r and ϕ and on the relevant generalized velocities $\dot X$, $\dot Y$, $\dot Z$ $\dot r$ and $\dot\phi$.

Such examples show that it is possible to express physical quantities such as the kineti energy, the potential energy, the angular momentum (and any other significant physical quantities) in terms of generalized coordinates and velocities.

Definition of Lagrangian. The Lagrangian of a system described by generalized coordinates and velocities q_i and $\dot{q}_i$, respectively, with $i = 1, 2, ... M$ is defined by

$$L = L(q_i, \dot{q}_i) = K(q_i, \dot{q}_i) - U(q_i)$$

where the notation $(q_i, \dot{q}_i)$ and (q_i) used in L has the meaning $U(q_i) = U(q_1, q_2, ..., q_M)$ and $K(q_i, \dot{q}_i) = K(q_1, q_2, ..., q_M, \dot{q}_1, \dot{q}_2 .... \dot{q}_M)$.

Definition of Action. The Action (or Action Integral) of a system with Lagrangian L, relevant to its motion from time t_1 to time t_2, is given by

$$A[\gamma] = \int_{t_1\ \gamma}^{t_2} dt\, L(q_i, \dot{q}_i)$$

where γ is the path in the configuration space describing the evolution from the initial configuration $\gamma(t_1) = (\, q_i(t_1), \dot{q}_i(t_1)\,)$ to the final one $\gamma(t_2) = (\, q_i(t_2), \dot{q}_i(t_2)\,)$. The state of the system at each time in the interval $[t_1, t_2]$ is represented by $(\, q_i(t), \dot{q}_i(t)\,)$.

The Hamilton Principle. The motion of the system from time t_1 to time t_2 is such that the Action $A[\gamma]$ has a stationary value for the **actual path** describing the motion, namely, the **variation of the Action** is such that $\delta A[\gamma] = 0$.

The meaning of such a statement is that, even if infinitely many different paths can be used to connect points $\gamma(t_1)$ and $\gamma(t_2)$, only those for which A is stationary provide the correct description of the evolution of the systems.

Theorem. The Lagrange equations describing the motion of the system follow from the Hamilton Principle.

Let us derive Lagrange's equation from the condition $\delta A[\gamma] = 0$. The latter can be cast into a more explicit form by means of variational derivatives. Derivatives are performed with respect to path γ. Then

$$\delta A[\gamma] = 0 \quad \Rightarrow \qquad 0 = \frac{\delta A}{\delta q_i} = \frac{\partial L}{\partial q_i} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i}$$

where the general formula (3) derived in subsection 2.1 has been used. For illustrative purposes, let us **review the main steps** of the calculation giving the functional derivative of A

$$\delta A[\gamma] = \int_{t_1\ \gamma}^{t_2} dt\, \delta L(q_i, \dot{q}_i) = \int_{t_1\ \gamma}^{t_2} dt\, \left[\sum_i \frac{\partial L}{\partial q_i}\, \delta q_i + \sum_i \frac{\partial L}{\partial \dot{q}_i}\, \delta \dot{q}_i \right]$$

$$= \int_{t_1\ \gamma}^{t_2} dt \sum_i \left[\frac{\partial L}{\partial q_i}\, \delta q_i + \frac{\partial L}{\partial \dot{q}_i}\, \frac{d}{dt} \delta q_i \right] = \int_{t_1\ \gamma}^{t_2} dt \sum_i \left[\frac{\partial L}{\partial q_i}\, \delta q_i + \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_i}\, \delta q_i \right) - \left(\frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} \right)\, \delta q_i \right]$$

$$= \left(\frac{\partial L}{\partial \dot{q}_i} \, \delta q_i \right)_{t_1}^{t_2} + \int_{t_1 \ \gamma}^{t_2} dt \sum_i \left[\frac{\partial L}{\partial q_i} - \left(\frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} \right) \right] \delta q_i = 0 \quad \Rightarrow \quad \frac{\partial L}{\partial q_i} - \left(\frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} \right) = 0$$

The Lagrange equations are thus given by

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} - \frac{\partial L}{\partial q_i} = 0 \,, \qquad p_i := \frac{\partial L}{\partial \dot{q}_i} \,, \tag{4}$$

where terms p_i are named **generalized momenta**. Note that momenta p_i are in one-to-one correspondence with generalized velocities $\dot{q}_i$. This derivation can be shown to be valid also for time-dependent Lagrangians $L(q_i, \, \dot{q}_i, \, t)$.

Comment. The deep meaning of the Hamilton principle and the implications thereof are extremely important: as a matter of fact, this principle provides a new basis on which to develop Classical Mechanics and represents an alternative to Newton's principles. If the equations of motion of a given system are derived from the three Newton laws (or, in case of many-particle systems, by applying the theorems of Newton's approach), within the Lagrangian scenario such equations are provided by formulas (4). The equations obtained within such two schemes appear to be perfectly equivalent.

The Lagrangian formulation is 1) general (it can be applied to, essentially, any classical system or model), 2) complete (it contains the same information embodied in the Newton approach), 3) unifying (it offers a common scheme for describing the dynamics of models and systems that are very different) and 4) autonomous (its variational character represents an approach independent from Newton's formulation).

The advantages offered by the new scheme should be evident: once the generalized coordinates of the system have been identified, the knowledge of Lagrangian L leads in a quite direct way to the equations characterizing the dynamics of the system. The equivalence between these two formualtions can be proven in a rigorous way. In the sequel, we verify such an equivalence by means of several examples.

2.4 Examples

Example 1. One-dimensional Harmonic oscillator. Generalized coordinates and velocities: x, $\dot{x}$ (no difference with the usual Cartesian quantities)

$$L = K - U \,, \qquad K = \frac{m}{2} \, \dot{x}^2 \,, \qquad U = + \frac{k}{2} \, (x - L)^2 \,.$$

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{x}} - \frac{\partial L}{\partial x} = 0 \quad \Rightarrow \quad \frac{d}{dt} \frac{\partial K}{\partial \dot{x}} + \frac{\partial U}{\partial x} = 0 \quad \Rightarrow \quad \frac{d(m \, \dot{x})}{dt} + k \, (x - L) = 0$$

The resulting equation is the well-known harmonic-oscillator equation

$$m \, \ddot{x} = -k \, (x - L) \qquad \Leftrightarrow \qquad m \, a = F(x) \,.$$

Example 2. Physical pendulum (thin rod of length ℓ with an extreme attached to a pivot placed in the origin). Oscillations take place in the plane zy around axix x. Generalized coordinates and velocities: α, $\dot{\alpha}$.

$$L = K - U\,, \qquad K = \frac{I}{2}\,\dot{\alpha}^2\,, \qquad U = +M\,g\,Z = -M\,g\,\frac{\ell}{2}\,\cos\alpha\,.$$

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{\alpha}} - \frac{\partial L}{\partial \alpha} = 0 \quad \Rightarrow \quad \frac{d}{dt}\frac{\partial K}{\partial \dot{\alpha}} + \frac{\partial U}{\partial \alpha} = 0 \quad \Rightarrow \quad \frac{d(I\,\dot{\alpha})}{dt} + Mg\frac{\ell}{2}\,\frac{\partial}{\partial \alpha}(-\cos\alpha) = 0$$

The resulting equation is the well-known pendulum equation

$$I\,\ddot{\alpha} + Mg\frac{\ell}{2}\,\sin\alpha\, = 0 \qquad \Leftrightarrow \qquad \frac{d\vec{L}}{dt} = \vec{M} \quad \text{with } \vec{M} \equiv \vec{R} \wedge M\,\vec{g}\,,\ |\vec{R}| = \ell/2\,.$$

Example 3. Spherical pendulum. This is a physical pendulum whose oscillations are not confined to plane zy. Generalized coordinates and velocities: ϕ, α, $\dot{\phi}$, $\dot{\alpha}$.

$$L = K - U\,, \qquad K = \frac{I}{2}\,(\,\dot{\alpha}^2 + \,\dot{\phi}^2 \sin^2\alpha\,)\,, \qquad U = M\,g\,Z = -M\,g\,\frac{L}{2}\,\cos\alpha\,.$$

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{\phi}} - \frac{\partial L}{\partial \phi} = 0 \quad \Rightarrow \quad \frac{d}{dt}\frac{\partial K}{\partial \dot{\phi}} = \frac{d}{dt}(I\,\dot{\phi}\sin^2\alpha)\, = 0 \quad \Rightarrow \quad I\,\dot{\phi}\,\sin^2\alpha = const$$

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{\alpha}} - \frac{\partial L}{\partial \alpha} = 0 \quad \Rightarrow \quad \frac{d}{dt}\frac{\partial K}{\partial \dot{\alpha}} - \frac{\partial K}{\partial \alpha} + \frac{\partial U}{\partial \alpha} = 0 \quad \Rightarrow \quad I\,\ddot{\alpha} - I\dot{\phi}^2 \sin\alpha\,\cos\alpha + Mg\frac{L}{2}\,\sin\alpha\, = 0\,.$$

Of course, for $\phi = 0$, the simple-pendulum equation of the previous example is recovered. An interesting aspect of spherical-pendulum equations is the presence of the constant of motion $I\,\dot{\phi}\,\sin^2\alpha$ which can be identified with a component L_z of angular momentum $\vec{L}$.

Proof. Recall that $\vec{r}_i = r_i\,\vec{u}(\phi,\alpha)$ and $\vec{u}(\phi,\alpha) = \,\vec{u}_x\cos\phi\,\sin\alpha\, + \vec{u}_y \sin\phi\,\sin\alpha\, - \vec{u}_z\cos\alpha$

$$\vec{L} = \sum_i m_i\,\vec{r}_i \wedge \vec{v}_i = \sum_i m_i\,r_i^2\vec{u} \wedge \frac{d\vec{u}}{dt} = I\,\vec{u} \wedge \frac{d\vec{u}}{dt}$$

$$= I\,(\cos\phi\,\sin\alpha,\,\sin\phi\,\sin\alpha,\,-\cos\alpha) \wedge \Big[\dot{\phi}\sin\alpha(-\sin\phi,\,\cos\phi\,,\,0)$$

$$+\dot{\alpha}(\cos\phi\,\cos\alpha,\,\sin\phi\,\cos\alpha,\,\sin\alpha)\Big]$$

$$= I\,\Big(\vec{u}_x(\,\dot{\alpha}\sin\phi + \dot{\phi}\cos\alpha\sin\alpha\,\cos\phi\,) + \vec{u}_y(-\dot{\alpha}\cos\phi + \dot{\phi}\cos\alpha\sin\alpha\,\sin\phi\,) + \vec{u}_z\,\dot{\phi}\,\sin^2\alpha\Big)$$

An appropriate use of this constant of motion simplifies the dynamical problem which is reduced to a unique equation for coordinate α

$$L_z = I\,\dot{\phi}\,\sin^2\alpha \;\;\Rightarrow\;\;\; \dot{\phi} = \frac{L_z}{I\,\sin^2\alpha} \;\;\Rightarrow\;\;\; I\,\ddot{\alpha} - I\left(\frac{L_z}{I\,\sin^2\alpha}\right)^2 \sin\alpha\,\cos\alpha + Mg\frac{L}{2}\,\sin\alpha\, = 0$$

$$I\,\ddot{\alpha} - \frac{L_z^2\,\cos\alpha}{I\,\sin^3\alpha} + Mg\frac{L}{2}\,\sin\alpha\, = 0$$

Comment. It is worth noting how the Lagrangian formalism provides in a direct way the constant of motion L_z as a byproduct of Lagrange-equation derivation. The explicit calculation of $\vec{L}$ and the proof that $dL_z/dt = 0$ are bypassed because this information is already incorporated in the Lagrangian.

Example 4. Two-body system in the presence of mutual gravitational forces. The generalized coordinates are X, Y, Z, (concerning the center of mass) and r and ϕ (concerning the planar orbital motion).

$$L = K - U\,, \qquad K = \frac{M}{2}\,(\dot{X}^2 + \dot{Y}^2 + \dot{Z}^2) + \frac{\mu}{2}\,(\,\dot{r}^2 + r^2\dot{\phi}^2\,)\,, \qquad U = -\frac{\gamma m_1 m_2}{r}\,.$$

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{X}} - \frac{\partial L}{\partial X} = 0 \;\;\Rightarrow\;\;\; \frac{d}{dt}\frac{\partial K}{\partial \dot{X}} = \frac{d}{dt}M\dot{X} = 0 \;\;\Rightarrow\;\;\; P_X = M\dot{Y} = const$$

Likewise

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{Y}} - \frac{\partial L}{\partial Y} = 0\,,\;\; \frac{d}{dt}\frac{\partial L}{\partial \dot{Z}} - \frac{\partial L}{\partial Z} = 0 \;\;\Rightarrow\;\;\; P_Y = M\dot{Y} = const\,,\;\; P_Z = M\dot{Z} = const$$

The three components P_X, P_Y, P_Z of total momentum are **constants of motion**. This confirms the total-momentum conservation ensuing from the center-of-mass theorem found in the Newtonian scheme.

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{r}} - \frac{\partial L}{\partial r} = 0 \;\;\Rightarrow\;\;\; \frac{d}{dt}\frac{\partial K}{\partial \dot{r}} - \frac{\partial K}{\partial r} + \frac{\partial U}{\partial r} = 0 \;\;\Rightarrow\;\;\; \mu\ddot{r} - \mu\,r\dot{\phi}^2 + \frac{\gamma m_1 m_2}{r^2} = 0$$

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{\phi}} - \frac{\partial L}{\partial \phi} = 0 \;\;\Rightarrow\;\;\; \frac{d}{dt}\frac{\partial K}{\partial \dot{r}} = 0 \;\;\Rightarrow\;\;\; \mu\,r^2\dot{\phi} = const$$

A **fourth constant of motion** appears which represents the third component of $\vec{L}$, $L_z = \mu\,r^2\dot{\phi}$. Again the Lagrange equations confirm the laws of the Newton formulation (in this

specific case the second Kepler law). By introducing the information that $\dot\phi = L_z/(\mu\, r^2)$ in the equation for r the dynamics reduces to a **unique equation**

$$\mu\ddot r - \frac{L_z^2}{\mu\, r^3} + \frac{\gamma m_1 m_2}{r^2} = 0 \qquad \Leftrightarrow \qquad \mu\,\frac{d^2\vec r}{dt^2} = -\gamma m_1 m_2 \frac{\vec r}{r^3}\,,$$

which is perfectly equivalent to the two-body Newton equation (shown on the right).

Example 5. Two masses subject to internal and external harmonic forces. Masses m_1 and m_2 can move in the interval $0 \le x \le L$ of the axis x. The generalized coordinates are simply x_1, x_2. The mass m_1 is attracted by a spring (with spring constant k_1 and length ℓ_1) towards $x = 0$, while the mass m_2 is attracted by a second spring (with spring constant k_2 and length ℓ_2) towards $x = L$. A third spring (with spring constant k_3 and length ℓ_3) connect the two masses. Assume that $L > \ell_1 + \ell_2 + \ell_3$.

$$K = \frac{m_1}{2}\,\dot x_1^2 + \frac{m_2}{2}\,\dot x_2^2\,, \qquad U = \frac{k_1}{2}\,(x_1-\ell_1)^2 \,+\, \frac{k_2}{2}\,(L-x_2-\ell_2)^2 \,+\, \frac{k_3}{2}\,(x_2-x_1-\ell_3)^2$$

$$\frac{d}{dt}\frac{\partial L}{\partial \dot x_1} - \frac{\partial L}{\partial x_1} = 0 \;\Rightarrow\; \frac{d}{dt}\frac{\partial K}{\partial \dot x_1} + \frac{\partial U}{\partial x_1} = 0 \;\Rightarrow\; m_1\,\ddot x_1 + k_1\,(x_1-\ell_1) - k_3\,(x_2-x_1-\ell_3) = 0$$

$$\frac{d}{dt}\frac{\partial L}{\partial \dot x_2} - \frac{\partial L}{\partial x_2} = 0 \;\Rightarrow\; \frac{d}{dt}\frac{\partial K}{\partial \dot x_2} + \frac{\partial U}{\partial x_2} = 0 \;\Rightarrow\; m_2\,\ddot x_2 - k_2\,(L-x_2-\ell_2) + k_3\,(x_2-x_1-\ell_3) = 0$$

Such equations reproduce the Newton equations

$$m_1\,\ddot x_1 = -\,k_1\,(x_1-\ell_1) + k_3\,(x_2-x_1-\ell_3)\,, \qquad m_2\,\ddot x_2 = +\,k_2\,(L-x_2-\ell_2) - k_3\,(x_2-x_1-\ell_3)$$

Example 6. Free rod in the gravitational field of strength g.

$$X\,,\; Y\,,\; Z\,,\; (\text{center of mass}) \qquad \phi\,,\; \theta\,,\; (\text{rod rotations})$$

$$L = K - U\,, \qquad K = \frac{M}{2}\,(\dot X^2 + \dot Y^2 + \dot Z^2) + \frac{I}{2}\,(\,\dot\theta^2 + \,\dot\phi^2 \sin^2\theta\,)\,, \qquad U = +M\,g\,Z\,.$$

$$\frac{d}{dt}\frac{\partial L}{\partial \dot X} - \frac{\partial L}{\partial X} = 0 \quad\Rightarrow\quad \frac{d}{dt}\frac{\partial K}{\partial \dot X} = \frac{d}{dt}M\dot X = 0 \quad\Rightarrow\quad M\dot X = const$$

$$\frac{d}{dt}\frac{\partial L}{\partial \dot Y} - \frac{\partial L}{\partial Y} = 0 \quad\Rightarrow\quad \frac{d}{dt}\frac{\partial K}{\partial \dot Y} = \frac{d}{dt}M\dot Y = 0 \quad\Rightarrow\quad M\dot Y = const$$

$$\frac{d}{dt}\frac{\partial L}{\partial \dot Z} - \frac{\partial L}{\partial Z} = 0 \quad\Rightarrow\quad \frac{d}{dt}\frac{\partial K}{\partial \dot Z} = -\frac{\partial U}{\partial Z} \quad\Rightarrow\quad M\ddot Z = -M\,g$$

The components $P_X = M\dot X$ and $P_Y = M\dot Y$ of total momentum are conserved. In addition one must consider the Lagrange equations for ϕ and θ

$$\frac{d}{dt}\frac{\partial L}{\partial \dot\phi} - \frac{\partial L}{\partial \phi} = 0 \quad\Rightarrow\quad \frac{d}{dt}\frac{\partial K}{\partial \dot\phi} = \frac{d}{dt}(I\,\sin^2\theta\dot\phi) \,= 0 \quad\Rightarrow\quad I\,\dot\phi\,\sin^2\theta = const$$

$$\frac{d}{dt}\frac{\partial L}{\partial \dot\theta} - \frac{\partial L}{\partial \theta} = 0 \quad\Rightarrow\quad \frac{d}{dt}\frac{\partial K}{\partial \dot\theta} - \frac{\partial K}{\partial \theta} = 0 \quad\Rightarrow\quad I\,\ddot\theta - I\dot\phi^2 \sin\theta\,\cos\theta \,= 0\,.$$

3 The Hamiltonian formulation of Classical Mechanics

Within the Lagrangian formulation the dynamical state of a system is specified by the generalized coordinates q_1, q_2, ..., q_M (describing the configuration space) and by the relevant generalized velocities $\dot{q}_1$, $\dot{q}_2$, ..., $\dot{q}_M$.

An **alternative description** of the dynamics of a system can be found by exploiting the **generalized momenta**

$$p_j = \frac{\partial L}{\partial \dot{q}_j} = h_j(q_i, \dot{q}_i)\,, \quad j \in [1, M]$$

defined in the previous chapter. Functions $h_i(q_i, \dot{q}_i)$ evidence the fact that generalized momenta depend, in general, on generalized coordinates and velocities. The list of variables

$$q_1\,,\ q_2\,,\ ...\, q_M\,, \qquad p_1 = \frac{\partial L}{\partial \dot{q}_1}\,,\ p_2 = \frac{\partial L}{\partial \dot{q}_2}\,, \ ...\, p_M = \frac{\partial L}{\partial \dot{q}_M}$$

contains the necessary information to specify the dynamical state of a system. The **equivalence of these two different representations** of dynamics is proven if velocities $\dot{q}_i$ can be expressed as functions of q_i and p_i. If functions h_j can be inverted giving

$$p_j = h_j(q_i, \dot{q}_i) \qquad \Rightarrow \qquad \dot{q}_j = f_j(q_i, p_i)$$

the equivalence is proved. To this end, consider the kinetic energy

$$K = \frac{1}{2}\sum_n \sum_k A_{nk}(q_i)\,\dot{q}_n \dot{q}_k \quad \Rightarrow \quad p_j = \frac{\partial L}{\partial \dot{q}_j} = \frac{\partial K}{\partial \dot{q}_j} = \frac{1}{2}\sum_n \sum_k A_{nk}(q_i)\,\left[\dot{q}_n \frac{\partial \dot{q}_k}{\partial \dot{q}_j} + \dot{q}_k \frac{\partial \dot{q}_n}{\partial \dot{q}_j}\right]$$

$$p_j = \frac{1}{2}\sum_n \sum_k A_{nk}(q_i)\,\left[\,\dot{q}_n \delta_{kj} + \,\dot{q}_k \delta_{nj}\right] = \sum_n A_{nj}(q_i)\,\dot{q}_n \qquad (\,A_{nk} = A_{kn}\,)$$

In general, matrix $A = ||A_{nk}||$ can be shown to be **invertible**, meaning that there exists a matrix B such that $B \circ A = A \circ B = \mathbb{I}$. Then

$$p_j = \sum_n A_{nj}(q_i)\,\dot{q}_n \qquad \Rightarrow \qquad \dot{q}_n = \sum_k B_{kn}(q_i)\,p_k\,, \quad B_{kn}A_{nj} = \delta_{nj}$$

This proof is based on the assumption that K is a quadratic function of velocities, a property that characterizes the most part of physical systems. Exceptions to this rule are represented by systems with a singular Lagrangian (for example, a gas of planar point-like vortices).

Comment. The spectral theorem of Matrix Theory shows that any finite-dimensional symmetric matrix A with real elements can be diagonalized by means of an orthogonal matrices O ($O^{-1} = O^t$). If A is a positive-definite matrix its eigenvalues are $\lambda_j > 0$. This entails that A is invertible.

3.1 Hamilton function. Poisson Brackets.

The Hamiltonian (or Hamilton function) of a system is defined by the equation

$$H(q_n, p_i) = \sum\nolimits_j p_j \dot{q}_j - L(q_n, \dot{q}_i) \qquad \text{where} \quad \dot{q}_j = f_j(q_i, p_i)\,.$$

Based on this definition it is possible to derive the equations of motion governing the evolution of variables q_n, p_i. To perform the following calculations one should recall that $\dot{q}_j$ must be viewed as functions of q_i, and p_i

$$\frac{\partial H}{\partial p_k} = \frac{\partial}{\partial p_k}\sum_j p_j \dot{q}_j - \frac{\partial L}{\partial p_k} = \dot{q}_k + \sum_j p_j \frac{\partial \dot{q}_j}{\partial p_k} - \sum_n \frac{\partial \dot{q}_n}{\partial p_k}\frac{\partial L}{\partial \dot{q}_n}$$

$$= \dot{q}_k + \sum_j p_j \frac{\partial \dot{q}_j}{\partial p_k} - \sum_n \frac{\partial \dot{q}_n}{\partial p_k} p_n = \dot{q}_k$$

$$\frac{\partial H}{\partial q_k} = \sum_j p_j \frac{\partial \dot{q}_j}{\partial q_k} - \frac{\partial L}{\partial q_k} - \sum_n \frac{\partial \dot{q}_n}{\partial q_k}\frac{\partial L}{\partial \dot{q}_n}$$

$$= \sum_j p_j \frac{\partial \dot{q}_j}{\partial q_k} - \frac{\partial L}{\partial q_k} - \sum_n \frac{\partial \dot{q}_n}{\partial q_k} p_n = -\frac{\partial L}{\partial q_k} = -\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_k} = -\dot{p}_k$$

In conclusion, one finds the so-called **Hamilton equations**

$$\dot{q}_k = \frac{\partial H}{\partial p_k}\,, \qquad \dot{p}_k = -\frac{\partial H}{\partial q_k}\,,$$

whose solutions completely determine $q_k(t)$ and $p_k(t)$. The 2M-dimensional space described by $(q_1, ..., q_M, p_1, ..., p_M)$ is called **phase space**. While the Lagrange equations represent a system of M independent, second-order equations, the Hamilton equations form a system of $2M$ **independent, first-order equations**.

It can be shown that the **Hamiltonian is a constant of motion**. Proof:

$$\frac{dH}{dt} = \sum_k \left[\dot{q}_k \frac{\partial H}{\partial q_k} + \dot{p}_k \frac{\partial H}{\partial p_k}\right] = \sum_k \left[\frac{\partial H}{\partial p_k}\frac{\partial H}{\partial q_k} + \left(-\frac{\partial H}{\partial q_k}\right)\frac{\partial H}{\partial p_k}\right] = 0$$

The physical meaning of the Hamiltonian is particularly important: H can be shown to represent the **total energy of the system**. As a consequence the total energy of the system is a **conserved quantity**. Based on the relation $\dot{q}_n = B_{kn}\, p_k$ and on the inverse relation $p_j = A_{ij}\, \dot{q}_i$ (summations on k and j is implied) one has

$$H(q_n, p_i) = \sum\nolimits_j p_j \dot{q}_j - L \;= \sum\nolimits_j A_{ij}\, \dot{q}_i \dot{q}_j \; - (K - U) = \; 2K \; - K \; + U \; = \; K \; + \; U$$

showing that H is the energy of the system. On the other hand,

$$H(q_n, p_i) = K + U = \frac{1}{2}\sum_i \sum_j A_{ij}\,\dot{q}_i \dot{q}_j + U = \frac{1}{2}\sum_i \sum_j A_{ij} \sum_k B_{ki}\,p_k \sum_r B_{rj}\,p_r + U$$

$$= \frac{1}{2}\sum_j \sum_k \delta_{jk}\,p_k \sum_r B_{rj}\,p_r + U = \frac{1}{2}\sum_k \sum_r B_{rk}\,p_k p_r + U$$

This result shows that H is **quadratic function** of generalized momenta.

A basic tool of Hamiltonian formulation is the definition of Poisson Brackets. These are evinced from the calculation of time derivatives of physical quantities. Let A a function of Hamiltonian coordinates q_k and p_k, also called **canonical variables**

$$\frac{dA}{dt} = \sum_k \left[\dot{q}_k \frac{\partial A}{\partial q_k} + \dot{p}_k \frac{\partial A}{\partial p_k}\right] = \sum_k \left[\frac{\partial H}{\partial p_k}\frac{\partial A}{\partial q_k} + \left(-\frac{\partial H}{\partial q_k}\right)\frac{\partial A}{\partial p_k}\right]$$

Such an expression, obtained by exploiting the Hamilton equations, implicitly defines the so-called (canonical) **Poisson Brackets** of two any functions A and B of variables q_k and p_k

$$\{A, B\} = \sum_k \left[\frac{\partial A}{\partial q_k}\frac{\partial B}{\partial p_k} - \frac{\partial B}{\partial p_k}\frac{\partial A}{\partial q_k}\right] \tag{5}$$

Properties (A, B, F are functions of q_j, p_k)

$$\{A, B\} = -\{B, A\}\,, \quad \{d\,A, c\,B\} = d\,c\,\{A, B\}\,,\ d, c \in \mathbb{R}$$

$$\{A, B + C\} = \{A, B\} + \{A, C\}\,, \quad \{A + D, B\} = \{A, B\} + \{D, B\}$$

$$\{A\,B, F\} = A\,\{B, F\} + B\,\{A, F\} \ \rightarrow \quad \{F, A\,B\} = A\,\{F, B\} + B\,\{F, A\}$$

$$\left\{A, \sum\nolimits_i c_i B_i\right\} = \sum_i c_i\,\{A, B_i\}\,, \quad \left\{\sum\nolimits_i c_i A_i, B\right\} = \sum_i c_i \{A_i, B\}\,,\ c_i \in \mathbb{R}\,.$$

$$\{A, \{B, F\}\} + \{B, \{F, A\}\} + \{F, \{A, B\}\} = 0$$

At the operational level, the Poisson Brackets are extremely useful because they allow one to discover **if** a function A, describing some property of the system relevant to H, **represents a constant of motion**. Remarkably, this result is achieved without solving the dynamical equations in an explicit way in that if $\{A, H\} = 0$ then

$$\frac{dA}{dt} = \{A, H\} = 0 \qquad \Rightarrow \quad A = const \ \Rightarrow \ \text{conserved quantity}\,.$$

This result represents an considerably useful tool for discovering the symmetry properties of (Hamiltonian) dynamical systems. Thanks to the **Poisson Brackets** the Hamilton equations can expressed in the alternative, simple form

$$\dot{q}_k = \{q_k, H\}\,, \qquad \dot{p}_k = \{p_k, H\}\,.$$

As for the Lagrange equations, let us consider some significant applications.

3.2 Examples

Example 1. One-dimensional Harmonic oscillator. Canonical variables: x, p.

$$K = \frac{m}{2}\,\dot{x}^2 \quad \Rightarrow \quad p = \frac{\partial L}{\partial \dot{x}} = \frac{\partial K}{\partial \dot{x}} = m\dot{x}\,, \qquad U = +\frac{k}{2}\,(x-L)^2\,.$$

$$H = K + U = \frac{m}{2}\,\dot{x}^2 + \frac{k}{2}\,(x-L)^2 = \frac{p^2}{2m} + \frac{k}{2}\,(x-L)^2$$

$$\text{Hamilton equations}: \quad \dot{x} = \frac{\partial H}{\partial p} = \frac{p}{m}\,, \qquad \dot{p} = -\frac{\partial H}{\partial x} = -k(x-L)\,.$$

By combining such equation one recovers the well-known harmonic-oscillator equation

$$\ddot{x} = \frac{1}{m}\dot{p}\,,\ \ \dot{p} = -k(x-L) \qquad \Rightarrow \qquad m\,\ddot{x} = -k\,(x-L)$$

The Hamiltonian formulations is therefore equivalent to the preceding formulations.

Example 2. Physical pendulum. Canonical variables: α, p_α.

$$K = \frac{I}{2}\,\dot{\alpha}^2 \quad \Rightarrow \quad p_\alpha = \frac{\partial L}{\partial \dot{\alpha}} = \frac{\partial K}{\partial \dot{\alpha}} = I\dot{\alpha} \equiv L_x\,, \qquad U = -M\,g\,\frac{L}{2}\,\cos\alpha\,.$$

$$H = K + U = \frac{I}{2}\,\dot{\alpha}^2 - M\,g\,\frac{L}{2}\,\cos\alpha = \frac{p_\alpha^2}{2I} - M\,g\,\frac{L}{2}\,\cos\alpha$$

$$\dot{\alpha} = \frac{\partial H}{\partial p_\alpha} = \frac{p_\alpha}{I} \qquad \dot{p}_\alpha = -\frac{\partial H}{\partial \alpha} = -M\,g\,\frac{L}{2}\,\sin\alpha$$

By combining such equation one recovers the well-known harmonic-oscillator equation

$$\ddot{\alpha} = \frac{\dot{p}_\alpha}{I}\,,\ \ \dot{p}_\alpha = -M\,g\,\frac{L}{2}\,\sin\alpha \qquad \Rightarrow \qquad I\,\ddot{\alpha} = -\,M\,g\,\frac{L}{2}\,\sin\alpha$$

Example 3. Spherical pendulum (oscillations are not confined to plane zy). Canonical variables: ϕ, α, p_ϕ, p_α.

$$K = \frac{I}{2}\,(\,\dot{\alpha}^2 + \,\dot{\phi}^2\sin^2\alpha\,) \quad \Rightarrow \quad p_\alpha = \frac{\partial L}{\partial \dot{\alpha}} = \frac{\partial K}{\partial \dot{\alpha}} = I\dot{\alpha}\,, \quad p_\phi = \frac{\partial L}{\partial \dot{\phi}} = \frac{\partial K}{\partial \dot{\phi}} = I\,\dot{\phi}\sin^2\alpha\,,$$

$$U = -M\,g\,\frac{L}{2}\,\cos\alpha\,.$$

$$H = K + U = \frac{I}{2}\,(\,\dot\alpha^2 + \,\dot\phi^2\sin^2\alpha\,) - M\,g\,\frac{L}{2}\,\cos\alpha = \frac{p_\alpha^2}{2I} \,+\, \frac{p_\phi^2}{2I\sin^2\alpha} \,-\, M\,g\,\frac{L}{2}\,\cos\alpha$$

$$\dot\alpha = \frac{\partial H}{\partial p_\alpha} = \frac{p_\alpha}{I}\,, \qquad \dot p_\alpha = -\frac{\partial H}{\partial \alpha} = \frac{p_\phi^2\cos\alpha}{I\sin^3\alpha} \,-M\,g\,\frac{L}{2}\,\sin\alpha\,,$$

$$\dot\phi = \frac{\partial H}{\partial p_\phi} = \frac{p_\alpha}{I\sin^2\alpha}\,, \qquad \dot p_\phi = -\frac{\partial H}{\partial \phi} = 0\,.$$

The fourth equation discloses the presence of the conserved quantity $p_\phi = const$. Similar to the Lagrangian case, p_ϕ can be exploited to simplify the dynamics. The first two equations reduce to a single one

$$\ddot\alpha = \frac{\dot p_\alpha}{I}\,, \quad \dot p_\alpha = \frac{p_\phi^2\cos\alpha}{I\sin^3\alpha} \,- M\,g\,\frac{L}{2}\,\sin\alpha \qquad \Rightarrow \qquad I\ddot\alpha = \frac{p_\phi^2\cos\alpha}{I\sin^3\alpha} \,- M\,g\,\frac{L}{2}\,\sin\alpha$$

Since p_ϕ is a constant, the latter equation appears to be completely decoupled from the remaining canonical variables. As expected, it perfectly reproduces the equation found in the Lagrangian approach. The third equation has an auxiliary role: once $\alpha(t)$ has been calculated $\phi(t)$ is obtained by a simple time integration.

Comment. It is worth noting how the Hamiltonian formalism provides in a direct way the constant of motion p_ϕ as a byproduct of the derivation of Hamilton equations. A circumstance that occurred also within the Lagrangian approach to the same system.

Example 4. Two-body system in the presence of mutual gravitational forces. The canonical variables are X, Y, Z, P_X, P_Y P_Z (concerning the center of mass) and r, ϕ, p_r and p_ϕ (concerning the planar orbital motion).

$$K = \frac{M}{2}\,(\dot X^2 + \dot Y^2 + \dot Z^2) + \frac{\mu}{2}\,(\,\dot r^2 + \,r^2\dot\phi^2\,)\,, \;\; U = -\frac{\gamma m_1 m_2}{r} \quad \Rightarrow$$

$$P_X = \frac{\partial L}{\partial \dot X} = \frac{\partial K}{\partial \dot X} = M\dot X\,, \quad P_Y = \frac{\partial L}{\partial \dot X} = \frac{\partial K}{\partial \dot Y} = M\dot Y\,, \quad P_Z = \frac{\partial L}{\partial \dot Z} = \frac{\partial K}{\partial \dot Z} = M\dot Z\,,$$

$$p_r = \frac{\partial L}{\partial \dot r} = \frac{\partial K}{\partial \dot r} = \mu\dot r\,, \quad p_\phi = \frac{\partial L}{\partial \dot\phi} = \frac{\partial K}{\partial \dot\phi} = \mu r^2\dot\phi\,,$$

$$H = K + U = \frac{P_X^2 + P_Y^2 + P_Z^2}{2M} \,+ \frac{p_r^2}{2\mu} + \frac{p_\phi^2}{2\mu r^2} - \frac{\gamma m_1 m_2}{r} \quad \Rightarrow$$

$$\dot X = \frac{\partial H}{\partial P_X} = \frac{P_X}{M}\,, \qquad \dot Y = \frac{\partial H}{\partial P_Y} = \frac{P_Y}{M}\,, \qquad \dot Z = \frac{\partial H}{\partial P_Z} = \frac{P_Z}{M}\,,$$

$$\dot{P}_Y = -\frac{\partial H}{\partial Y} = 0\,, \qquad \dot{P}_X = -\frac{\partial H}{\partial X} = 0\,, \qquad \dot{P}_Z = -\frac{\partial H}{\partial Z} = 0\,,$$

the ensuing three constant of motions, P_X, P_Y, and P_Z describe the total-momentum conservation following from the center-of-mass theorem. The remaining equations are

$$\dot{r} = \frac{\partial H}{\partial p_r} = \frac{p_r}{\mu}\,, \qquad \dot{p}_r = -\frac{\partial H}{\partial r} = \frac{p_\phi^2}{\mu r^3} - \frac{\gamma m_1 m_2}{r^2}$$
$$\dot{\phi} = \frac{\partial H}{\partial p_\phi} = \frac{p_\phi}{\mu r^2}\,, \qquad \dot{p}_\phi = -\frac{\partial H}{\partial \phi} = 0\,.$$

The last equation evidences the presence of a well-known constant of motion: $p_\phi = \mu r^2 \dot{\phi} = const$ representing the third component of the angular momentum relevant to the planar motion of the two bodies. Once more the results of both the Lagrangian and the Newtonian approaches are fully confirmed. The equations for r and p_r can be reduced to a single one where p_ϕ is a fixed quantity

$$\mu\,\ddot{r} = \frac{p_\phi^2}{\mu r^3} - \frac{\gamma m_1 m_2}{r^2}\,.$$

Example 5. Two masses subject to internal and external harmonic forces. Masses m_1 and m_2 can move in the interval $0 \le x \le L$ of the axis x (recall that $L > \ell_1 + \ell_2 + \ell_3$). The canonical variables are x_1, x_2, p_1, p_2.

$$K = \frac{m_1}{2}\,\dot{x}_1^2 + \frac{m_2}{2}\,\dot{x}_2^2 \;\Rightarrow \qquad p_1 = \frac{\partial L}{\partial \dot{x}_1} = m_1\,\dot{x}_1\,,\; p_1 = \frac{\partial L}{\partial \dot{x}_2} = m_2\,\dot{x}_2\,,$$

$$U = \frac{k_1}{2}\,(x_1 - \ell_1)^2 + \frac{k_2}{2}\,(L - x_2 - \ell_2)^2 + \frac{k_3}{2}\,(x_2 - x_1 - \ell_3)^2$$

$$H = \frac{p_1^2}{2m_1} + \frac{p_2^2}{2m_2} + \frac{k_1}{2}\left(x_1 - \ell_1\right)^2 + \frac{k_2}{2}\left(L - x_2 - \ell_2\right)^2 + \frac{k_3}{2}\left(x_2 - x_1 - \ell_3\right)^2$$

$$\dot{x}_1 = \frac{\partial H}{\partial p_1} = \frac{p_1}{m_1}\,, \quad \dot{p}_1 = -\frac{\partial H}{\partial \dot{x}_1} = -k_1\left(x_1 - \ell_1\right) + k_3\left(x_2 - x_1 - \ell_3\right),$$
$$\dot{x}_2 = \frac{\partial H}{\partial p_2} = \frac{p_2}{m_2}\,, \quad \dot{p}_2 = -\frac{\partial H}{\partial \dot{x}_2} = +k_2\left(L - x_2 - \ell_2\right) - k_3\left(x_2 - x_1 - \ell_3\right)$$

The first and the third equation give

$$\ddot{x}_1 = \frac{\dot{p}_1}{m_1}\,, \quad \ddot{x}_2 = \frac{\dot{p}_2}{m_2}$$

which, combined with the second and fourth equations, reproduce the Newton equations

$$m_1\ddot{x}_1 = -\,k_1\Big(x_1 - \ell_1\Big) + \,k_3\Big(x_2 - x_1 - \ell_3\Big)\,,\;\; m_2\ddot{x}_2 = +\,k_2\Big(L - x_2 - \ell_2\Big) - \,k_3\Big(x_2 - x_1 - \ell_3\Big)\,.$$

4 The Schrödinger equation

At the beginning of the last century a series of fundamental experiments made more and more evident the inadequacy of classical physics in providing a satisfactory explanation of an increasing number of phenomena at the microscopic and atomic scales. Contradictory results were found when adopting a classical interpretation of processes involving high-frequency radiation, conductivity properties of solids, the interaction of radiation and matter, the apparent wavelike behavior of electrons and atoms, and many other aspects of atomic and condensed-matter physics. The need of a new, more effective theoretical basis able to provide a consistent interpretation of this rich phenomenology led to the formulation of Quantum Mechanics.

Among many crucial experiments, one should recall, in particular,

1) the measure of the energy (per unit of volume and frequency) of Black-Body Radiation showing that the (classical) Rayleigh-Jeans law fails to describe the high-frequency radiation while the (quantum) Planck radiation law perfectly matches experimental data (Planck, 1900),
2) the Photoelectric Effect leading to the assumption that light consists of photons representing quanta of energy $E = \hbar\, w$ (Einstein, 1905),
3) the Davisson-Germer experiment (1927) where an electron beam striking a grid (a crystal lattice of suitably small lattice constant) undergoes a diffraction effect. This gives rise to interference patterns typical of visible light or X rays,
4) the Thomson experiment (1928) where an electron beam going through a thin foil of gold determines well-known diffraction patterns (represented by a series of concentric circles) on a photographic plate.
5) the existence of discrete atomic emission and absorption spectra,
6) the Franck-Hertz experiment (1913) evidencing the discrete energy levels of atoms.

4.1 Derivation of the Schrödinger equation

Many experiments showed that, similar to the classical light, electron beams exhibit a wavelike behavior. For example, when the beam strikes a crystal lattice (electrons elastically collide with atoms that act as scattering centres) they determine diffractions patterns owing to the effect of coherent interference.
Abandoning a deterministic description of particles in terms of position and velocity, electrons were thus assumed to be represented by **elementary waves** of the form

$$\psi(x,t) = Ae^{i(kx-wt)}\,, \qquad k = \frac{2\pi}{\lambda} \quad \text{with} \quad \lambda = 2\pi\frac{\hbar}{p}$$

where λ is the wavelength of the electron (postulated in 1924 by De Broglie), p the electron momentum and $\hbar = 1.054 \times 10^{-34}\, J\, s$ the Planck constant. As in the case of electromagnetic waves, the energy was assumed to be $E = \hbar w$. Such a representation, by the way, allowed to combine plane waves originated by different "sources" to explain interference effects.
The possibility to **superimpose different waves** suggested that the equation governing the

time evolution $\psi(x,t)$ had a **linear character** as the Maxwell equations describing electromagnetic waves. Then, if one considers, for example, an equation similar to the D'Alambert equation for electromagnetic fields, the substitution of $\psi(x,t) = Ae^{i(kx-wt)}$ in

$$\frac{\partial^2}{\partial t^2}\psi - \sigma\,\frac{\partial^2}{\partial x^2}\psi = 0 \tag{6}$$

entails

$$(w^2 - \sigma\,k^2)\psi = 0 \quad \Rightarrow \quad w^2 = \sigma\,k^2.$$

The resulting dispersion relation $w^2 = \sigma\,k^2$ is able to reproduce the characteristic energy-momentum relation for electromagnetic waves $E = \hbar w = c\hbar k = c\,p$ if one assumes $\sigma = c^2$ with $c =$ light speed. However, it **does not succeed** in reproducing the correct relation between energy and momentum $E = p^2/(2m)$ in which E features a **quadratic dependence** on k instead of a linear one:

$$w^2 = \sigma\,k^2 \quad \Rightarrow \quad \hbar^2 w^2 = \sigma\,\hbar^2 k^2 \quad \Rightarrow \quad E^2 = \sigma\,p^2 \qquad \text{if } E = \hbar w\,,\ p = \frac{2\pi\,\hbar}{\lambda} \equiv \hbar k\,.$$

Since $\psi(x,t)$ is a wave function now describing **matter particles**, equation (6) must be rejected and a new wave equation must be postulated for $\psi(x,t)$ which is consistent with $E = p^2/(2m)$. The correct form for the equation describing matter particles was found to be

$$i\frac{\partial\psi}{\partial t} + \sigma\,\frac{\partial^2\psi}{\partial x^2} = 0 \ \Rightarrow \tag{7}$$
$$\left(w - \sigma\,k^2\right)\,\psi = 0 \ \Rightarrow \quad w = \sigma\,k^2 \ \Rightarrow \quad E = \hbar w = \hbar\sigma k^2\,.$$

If one assumes $\sigma \equiv \hbar/(2m)$ then

$$E = \hbar w = \hbar\sigma k^2 = \hbar\,\frac{\hbar k^2}{2m} = \frac{p^2}{2m}$$

leading to the **expected expression for the energy of a free particle**. It is worth noting how the different forms of equations (6) and (7) well match the famous energy-momentum equation

$$E = \sqrt{p^2c^2 + m^2c^4}$$

of the Einstein relativistic theory stating that electromagnetic waves and matter wave functions must feature distinct energy-momentum relations. In the light case, whose elementary constituents are photons with zero mass, one recovers $E = \sqrt{p^2c^2 + m^2c^4} = cp$ while for non-relativistic particles ($p << mc$) one has

$$E = \sqrt{p^2c^2 + m^2c^4} = mc^2\sqrt{1 + p^2/m^2c^2} \simeq mc^2\left(1 + \frac{p^2}{2m^2c^2}\right) = mc^2 + \frac{p^2}{2m}$$

which, up to a constant term, provides the standard formula for the kinetic energy.

Recalling that the motion of particles takes place in the 3D ambient space, the final form of the **Schrödinger equation** for a free particle was

$$i\hbar\frac{\partial\psi}{\partial t} = -\frac{\hbar^2}{2m}\nabla^2\psi \tag{8}$$

with $\nabla^2 = \Delta = \partial^2_{x_1} + \partial^2_{x_2} + \partial^2_{x_3}$, whose elementary solutions are

$$\psi(\vec{x},t) = Ae^{i\left(\vec{k}\cdot\vec{x}-\frac{Et}{\hbar}\right)}\,, \qquad k = |\vec{k}| = \frac{2\pi}{\lambda} \qquad \Rightarrow \quad E = \frac{\hbar^2\vec{k}^2}{2m} = \frac{\vec{p}^2}{2m}\,. \tag{9}$$

Two important informations emerge from equation (8). The fact that

$$-\frac{\hbar^2}{2m}\nabla^2\psi \;=\; E\,\psi\,, \qquad\qquad -i\,\hbar\,\nabla\psi \;= \hbar\vec{k}\,\psi = \vec{p}\,\psi$$

suggests that, within Quantum Mechanics, **energy and momentum are represented by operators**

$$E \;\Leftrightarrow\; H = -\frac{\hbar^2}{2m}\nabla^2\,, \qquad \vec{p} \;\Leftrightarrow\; \hat{p} = -i\,\hbar\,\nabla$$

where ∇ is the gradient operator and **momentum operator** $\hat{p}$, formed by three vector components

$$\hat{p}_1 = -i\hbar\frac{\partial}{\partial x_1} = -i\hbar\partial_{x_1}\,,\;\; \hat{p}_2 = -i\hbar\frac{\partial}{\partial x_2} = -i\hbar\partial_{x_2}\,,\;\; \hat{p}_3 = -i\hbar\frac{\partial}{\partial x_3} = -i\hbar\partial_{x_3}\,, \quad \hat{p} = \sum_{j=1}^{3}\hat{p}_j\vec{u}_j$$

should not be confused with the classical momentum $\vec{p} = \vec{u}_x p_x + \vec{u}_y p_y + \vec{u}_z p_z$. Symbol $\hat{H}$ is called **Hamilton operator** or simply **Hamiltonian**. The definition of the **momentum operator** allows to write the Hamiltonian as a function of momentum opertor components $\hat{p}_1$, $\hat{p}_2$ and $\hat{p}_3$

$$\hat{H} = -\frac{\hbar^2}{2m}\nabla^2 = -\frac{\hbar^2}{2m}\left(\,\partial^2_{x_1} + \partial^2_{x_2} + \partial^2_{x_3}\,\right) = \frac{1}{2m}\left[\,(-i\hbar\partial_{x_1})^2 + (-i\hbar\partial_{x_2})^2 + (-i\hbar\partial_{x_3})^2\,\right]$$

$$= \frac{1}{2m}\left[\,\hat{p}_1^2 + \hat{p}_2^2 + \hat{p}_3^2\,\right] = \frac{\hat{p}^2}{2m}$$

showing a complete similarity to the classical formula.
The Schrödinger equation for a **general problem** where a quantum particle **undergoes a potential** $U(\vec{r})$ can be evinced from the apparent analogy between classical and quantum formulas. More specifically, the Hamilton operator in the presence of U is deduced from the classical energy

$$\text{if} \quad E = \frac{\vec{p}^2}{2m} \;\Leftrightarrow\; \hat{H} = \frac{1}{2m}\hat{p}^2 \quad \text{then} \qquad E = \frac{\vec{p}^2}{2m} + U(\vec{r}) \;\Longrightarrow\; \hat{H} = \frac{1}{2m}\hat{p}^2 + U(\vec{r})$$

Then the **Schrödinger equation** for a problem including a potential $U(\vec{r})$ is simply given by

$$i\hbar\frac{\partial\psi}{\partial t} = \hat{H}\psi \qquad \text{with} \quad \hat{H} = \frac{\hat{p}^2}{2m} + U(\vec{r})\,. \tag{10}$$

Superposition Principle. If $\psi_1(\vec{r},t)$ and $\psi_2(\vec{r},t)$ represent two solutions of a Schrödinger problem $i\hbar\partial_t\psi_k = \hat{H}\psi_k$, with $k = 1, 2$ then any linear combination $\psi = c_1\psi_1 + c_2\psi_2$ with c_1, $c_2 \in \mathbb{C}$ is a solution of the same problem

$$i\hbar\frac{\partial\psi}{\partial t} = c_1 i\hbar\frac{\partial\psi_1}{\partial t} + c_2 i\hbar\frac{\partial\psi_1}{\partial t} = c_1\hat{H}\psi_1 + c_2\hat{H}\psi_2 = \hat{H}\left(c_1\psi_1 + c_2\psi_2\right) = \hat{H}\psi$$

Thus a general solution can be formed by considering an arbitrarily complex superposition of solutions $\Phi = \sum_k c_k\psi_k$. This property entails that the whole set of solutions of a given Schrödinger problem is a **linear vector space**. This property is discussed in the sequel.

Schrödinger equation for many-particle systems.

$$i\hbar\partial_t\Psi(\vec{r}_1,\vec{r}_2,...\vec{r}_N,t) = \hat{H}\Psi(\vec{r}_1,\vec{r}_2,...\vec{r}_N,t)\,, \qquad \hat{H} = \sum_{i=1}^{N}\frac{\hat{p}_i^2}{2m_i} + U(\vec{r}_1,\vec{r}_2,...\vec{r}_N)\,. \tag{11}$$

where

$$\hat{p}_i^2 = -\hbar^2\nabla_i^2 = -\hbar^2\Big(\partial_{x_i}^2 + \partial_{y_i}^2 + \partial_{z_i}^2\Big)\,.$$

4.2 Density of Probability and Continuity Equation

A basic assumption in the development of Quantum Theory concerns the information about the state of the system embodied in the wave function $\psi(\vec{r},t)$. The quantities

$$\rho(\vec{r},t) = |\psi(\vec{r},t)|^2\,, \qquad dP = \rho(\vec{r},t)d^3\vec{r}\,,$$

represent the **probability density** to find an electron at time t at position $\vec{r}$ and the probability that the electron (and, more in general, any quantum particle) occupies at time t volume $d^3\vec{r}$, respectively. As a consequence, if V is the volume of the ambient space of a given the system

$$P(V_0) = \int_{V_0} \rho(\vec{r},t)d^3r \,<\, 1\,, \quad P(V) = \int_V \rho(\vec{r},t)\, d^3r = \,1\,,$$

are the probability that a particle occupies at time t volume $V_0 \subset V$ and the total probability (necessarily equal to 1), respectively. Note that if the ambient space V is the whole 3D space in the preceding integral $V \equiv \mathbb{R}^3$. This probabilistic interpretation was initially suggested and supported by the experiments involving the electron diffraction.

Continuity equation. The time evolution of $\rho(\vec{r},t)$ can be shown to obey a well-known hydrodynamic equation leading to define the probability-density current.

$$i\hbar\partial_t\psi = \left(-\frac{\hbar^2}{2m}\nabla^2 + U(\vec{r})\right)\psi \quad \Rightarrow \quad -i\hbar\partial_t\psi^* = \left(-\frac{\hbar^2}{2m}\nabla^2 + U(\vec{r})\right)\psi^*$$

($\hat{H}$ has been written in an explicit way)

$$i\hbar\partial_t|\psi|^2 = i\hbar\psi^*\partial_t\psi + i\hbar\psi\partial_t\psi^* = \psi^*\left(-\frac{\hbar^2}{2m}\nabla^2 + U(\vec{r})\right)\psi - \psi\left(-\frac{\hbar^2}{2m}\nabla^2 + U(\vec{r})\right)\psi^*$$

$$= -\frac{\hbar^2}{2m}\psi^*\nabla^2\psi + \frac{\hbar^2}{2m}\psi\nabla^2\psi^* = -\frac{\hbar^2}{2m}\Big[\psi^*\nabla^2\psi - \psi\nabla^2\psi^*\Big] = -\frac{\hbar^2}{2m}\Big[\nabla\Big(\psi^*\nabla\psi - \psi\nabla\psi^*\Big)\Big]$$

Then

$$i\hbar\partial_t|\psi|^2 = -i\,\hbar\,\nabla\vec{j} \text{ with } \quad \vec{j} = \frac{\hbar}{2m\,i}\Big[\psi^*\nabla\psi - \psi\nabla\psi^*\Big] \quad \Rightarrow \quad \partial_t\rho + \,\mathrm{div}\vec{j} = 0$$

where the latter formula has the form of a **continuity equation**. Since ρ was assumed to represent the (probability) density, $\vec{j}$ is the **probability-density current**. Recall that $\mathrm{div}\vec{j} = \nabla\vec{j} = \partial_x j_x + \partial_y j_y + \partial_z j_z$.

Current $\vec{j}$ can be also expressed in the standard form $\vec{j} = \rho\vec{v}$. Substituting $\psi = |\psi|e^{i\theta}$ in $\vec{j}$ one has

$$\vec{j} = \frac{\hbar}{2mi}\Big[\psi^*\nabla\psi - \psi\nabla\psi^*\Big] = \frac{\hbar}{2mi}\Big[\psi^*\left(e^{i\theta}\nabla|\psi| + i|\psi|e^{i\theta}\nabla\theta\right) - C.C.\Big]$$

$$= \frac{\hbar}{2mi}\Big[\big(|\psi|\nabla|\psi| + i|\psi|^2\nabla\theta\big) - C.C.\Big] = \frac{\hbar}{m}|\psi|^2\nabla\theta \;\; \rightarrow \;\; \vec{v} = \frac{\hbar}{m}\,\nabla\theta\,.$$

By integrating the two terms of the continuity equation on volume V one finds the integral equation

$$\int_V \partial_t\rho\, d^3r + \int_V \mathrm{div}\vec{j}\, d^3r = 0\,,$$

which, thanks to the Gauss-Green Theorem (or divergence Theorem)

$$\int_V \mathrm{div}\vec{j}\, d^3r = \int_S \vec{j}\cdot\vec{u}\, dS\,,$$

can be recast into a new form

$$\partial_t P(V) = \int_V \partial_t\rho\, d^3r = -\int_S \vec{j}\cdot\vec{u}\, dS\,. \tag{12}$$

where surface S is the boundary of volume V. The latter equation represents the integral version of the continuity equation. This result states that the change in time of the probability $P(V)$ relevant to any volume V is determined by the total flux of the probability-density current across S. If the system involves a **finite ambient space** V, then $\vec{j}$ **vanishes on** S since, by definition, $\psi = 0$ on S for any wave functions $\psi \in L^2(V)$. Then $\partial_t P(V) = 0$ entails that the total probability $P(V)$ is constant in time.

If $V = \mathbb{R}^3$ (meaning that the ambient space of a particle described by ψ is the whole 3D space) then the distance R of points on S from the origin becomes infinitely large. If ψ is a normalizable wave function then integral

$$\int_{\mathbb{R}^3} |\psi(\vec{r},t)|^2\, d^3r = 1 < \infty\,,$$

where ψ must be rapidly vanishing for $r \to \infty$. Hence $|\vec{j}|\, dS \to 0$ for $R \to \infty$ and

$$\int_{\mathbb{R}^3} \partial_t\rho\, d^3r = \partial_t \int_{\mathbb{R}^3} \rho\, d^3r = 0 \quad \Rightarrow \quad P(V) = \int_{\mathbb{R}^3} \rho\, d^3r = constant \tag{13}$$

This proof shows that the normalization integral must have the same value at any time namely it is a time-independent quantity. Quantum mechanically, the total probability is a conserved quantity.

Comment. One can determine qualitatively the dependence of ψ on r for $r \to \infty$ which ensures the vanishing of the surface integral. Wave function ψ can be written as $\psi \simeq f(r)/r^k$ where $|f(r)| < C$ for large r. Then

$$\psi \simeq \frac{f(r)}{r^k} \;\; \Rightarrow \;\; \vec{j} \propto i(\psi^*\nabla\psi - \psi\nabla\psi^*) = i\left[\psi^*\left(\frac{1}{r^k}\nabla f - \frac{kf\vec{r}}{r^{k+2}}\right) - C.C.\right]$$

$$= i\left[\left(\frac{f^*}{r^{2k}}\nabla f - \frac{k|f|^2\vec{r}}{r^{2k+2}}\right) - C.C.\right] = i\frac{1}{r^{2k}}\left(f^*\nabla f - C.C.\right)$$

and

$$\vec{j}\cdot d\vec{S} = \vec{j}\cdot\vec{u}_r\, r^2 d\Omega \propto\; i\left[(f^*\nabla f - C.C.)\cdot\vec{u}_r\right]\frac{1}{r^{2k-2}}\, d\Omega \;\to\; 0$$

provided $k > 1$. The inequality $k > 1$ is thus **sufficient to guartantee** that integral (12) vanishes and thus supports the interpretation of $P(V)$ as the total probability for $V = \mathbb{R}^3$ since integral (13) is **finite and constant**. A more careful inspection of integral (13)

$$\int_{\mathbb{R}^3} |\psi(\vec{r})|^2\, d^3r = \int_0^\infty \int_0^\pi \int_0^{2\pi} \frac{|f|^2}{r^{2k}}\, r^2 dr \sin\theta d\theta d\phi = 1$$

shows (a posteriori) how $k > 3/2$ is required to avoid a diverging behavior of the integral. This condition is acceptable because it simply reinforces the previous inequality $k > 1$. Note that $|(f^*\nabla f - C.C.)\cdot\vec{u}_r|$ is finite for $r \to \infty$. For example, with a spherical wave, $f = A\, e^{ikr}$

$$(f^*\nabla f - C.C.)\cdot\vec{u}_r = (|f|^2 ik\nabla r - C.C.)\cdot\vec{u}_r = 2A^2\, ik\,\frac{\vec{r}}{r}\cdot\vec{u}_r = \; 2A^2 ik\,.$$

4.3 Physical operators and commutation relations

The derivation of the Schrödinger equation has shown how classical variables are represented, within Quantum Mechanics, by operators. Operators are of two types. **The first class** is formed by the operators corresponding to the position vector $\vec{x} = \vec{u}_1 x_1 + \vec{u}_2 x_2 + \vec{u}_3 x_3$ its components x_1, x_2, and x_3 and, in general, any function $f(x_1, x_2, x_3)$ of coordinates

$$\hat{x} \leftrightarrow \vec{x}\,, \qquad \hat{x}_i \leftrightarrow x_i\,, \qquad \hat{f} = f(\hat{x}_1, \hat{x}_2, \hat{x}_3) \leftrightarrow f = f(x_1, x_2, x_3)\,.$$

Their action on a generic wave function ψ is particularly simple (they are said multiplication operators)

$$\hat{x}\,\psi(\vec{x}) \equiv \vec{x}\psi(\vec{x}) \qquad \hat{x}_i\,\psi(\vec{x}) \equiv x_i\psi(\vec{x}) \qquad \hat{f}\;\psi(\vec{x}) \equiv f(x_1, x_2, x_3)\,\psi(\vec{x})$$

and entails that multiplication operators can be identified with their classical counterparts. **The second class** is formed by operators involving derivatives. These are, for example,

$$\hat{p} = -i\hbar\nabla \leftrightarrow \vec{p}, \qquad \hat{p}_j = -i\hbar\partial_{x_j} \leftrightarrow p_j\,, \qquad \hat{T} = \frac{\hat{p}^2}{2m} = -\frac{\hbar^2}{2m}\nabla^2 \leftrightarrow T = \frac{\vec{p}^2}{2m}\,,$$

associated to the classical momentum vector, its components and the kinetic energy. Their action on state ψ involves the use of derivatives

$$\hat{p}_j\psi = -i\hbar\frac{\partial\psi}{\partial x_j}\,, \quad \hat{T}\psi = -\frac{\hbar^2}{2m}\left(\frac{\partial^2\psi}{\partial x_1^2} + \frac{\partial^2\psi}{\partial x_2^2} + \frac{\partial^2\psi}{\partial x_3^2}\right)$$

The angular-momentum operator and its component provide a third example showing that operators of such classes can be mixed. Since, classically, $\vec{L} = \vec{x} \wedge \vec{p}$ with $L_j = \epsilon_{jmn} x_m p_n$

$$\hat{L} = \hat{x} \wedge \hat{p} = -i\hbar\vec{x} \wedge \nabla = -i\hbar\epsilon_{jmn} x_m \frac{\partial}{\partial x_n}\,, \tag{14}$$

$$\hat{L}_1 = -i\hbar(x_2\partial_{x_3} - x_3\partial_{x_2})\,, \qquad \hat{L}_2 = -i\hbar(x_3\partial_{x_1} - x_1\partial_{x_3})\,, \qquad \hat{L}_3 = -i\hbar(x_1\partial_{x_2} - x_2\partial_{x_1})\,.$$

General properties of linear operators. Consider a generic operator A. The action of A is such that $A\,\psi(\vec{x}) = \phi(\vec{x}) \in L^2(V)$ where $L^2(V)$ represents the **space of square integrable functions** and the domain V is the space of the physical system under consideration. This equation states that the effect of the action of A is to generate another function ϕ still belonging to $L^2(V)$. Other properties

i) Physical operators are linear: $A(c_1\psi_1 + c_2\psi_2) = c_1 A\psi_1 + c_2 A\psi_2 = c_1\phi_1 + c_2\phi_2$

ii) The sum of two operators: $(A + B)\psi = A\psi + B\psi$

iii) Product of two operator: $AB\,\psi = A\,(B\,\psi)$

iv) Unit operator (or identity operator): $\mathbb{I}\,A = \mathbb{I}\,A = A$, $\mathbb{I}\,\psi = \psi$

v) Zero operator: $0\,A = 0\,A = 0$, $0\,\psi = 0$

Definition of commutator. The commutator of two physical operators A and B is defined by

$$[A,B] = AB - BA \quad \Rightarrow \quad [A,B]\psi = A\,(B\,\psi) - B\,(A\,\psi)\,.$$

A commutator is antisymmetric $[A,B] = -[B,A]$ and bilinear

$$[A, c_1B_1 + c_2B_2] = c_1[A,B_1] + c_2[A,B_2]\,, \quad [c_1A_1 + c_2A_2, B] = c_1[A_1,B] + c_2[A_2,B]\,,$$

where A, B, A_i, and B_i are operators and $c_1, c_2 \in \mathbb{C}$. Let us consider some significant examples involving operators relevant to canonical variables x_n and p_k

$$[x_i,x_j] = 0 \;\Leftrightarrow\; x_ix_j\psi = x_jx_i\psi \qquad [\hat{p}_k,\hat{p}_j] = 0 \;\Leftrightarrow\; -\hbar^2\frac{\partial}{\partial x_k}\frac{\partial\psi}{\partial x_j} = -\hbar^2\frac{\partial}{\partial x_j}\frac{\partial\psi}{\partial x_k}$$

The first commutator is trivial (namely it is zero) because scalar variables can be exchanged. The second commutator is also trivial because the derivation order can be exchanged. A third more interesting commutators is

$$[x_k,\hat{p}_j] = \,i\,\hbar\,\delta_{kj} \;\Leftrightarrow\; \left(-i\hbar x_k\frac{\partial}{\partial x_j} + i\hbar\frac{\partial}{\partial x_j}\,x_k\right)\psi = -i\hbar x_k\frac{\partial\psi}{\partial x_j} + i\hbar\frac{\partial}{\partial x_j}\,(x_k\psi)$$

$$= -i\hbar x_k\frac{\partial\psi}{\partial x_j} + i\hbar\frac{\partial\psi}{\partial x_j}\,x_k + i\hbar\frac{\partial x_k}{\partial x_j}\,\psi = \,i\,\hbar\,\delta_{kj}\,\psi$$

showing that operators relevant to **canonically conjugate** variables have a non vanishing commutators

$$[x_1,\hat{p}_1] = i\hbar\,, \qquad [x_2,\hat{p}_2] = i\hbar\,, \qquad [x_3,\hat{p}_3] = i\hbar\,.$$

Note that, at a formal level, the previous commutators should be written as $[x_n,\hat{p}_k] = i\hbar\,\delta_{kj}\mathbb{I}$ where the identity operator is included. Since its action is trivial, $\mathbb{I}$ it is usually understood.

Commutators of $\hat{p}_k$ with a generic function $f(\hat{x}) = f(x_1,x_2,x_3)$ of position operator $\hat{x}$

$$[\hat{p}_k, f]\psi = \hat{p}_k(f\psi) - f\,\hat{p}_k\psi = -i\hbar\partial_{x_k}(f\psi) - f\,\hat{p}_k\psi$$

$$= -i\hbar\psi(\partial_{x_k}f) - i\hbar\,f(\partial_{x_k}\psi) + i\hbar\,f(\partial_{x_j}\psi) = -i\hbar(\partial_{x_k}f)\psi \qquad \Rightarrow \quad [\hat{p}_k,f] = -i\hbar\,\frac{\partial f}{\partial x_k}\,.$$

Operator version of Leibnitz's rule

$$[A,BC\,] = \,B\,[A,C]\, + [A,B]\,C \tag{15}$$

Proof:

$$[A, BC\,] = A\,(BC) - (BC)\,A = ABC - BCA = ABC + (-BAC + BAC) - BCA$$

$$= (ABC - BAC) + (BAC - BCA) = (AB - BA)C + B(AC - CA) = [A, B]C\, + B[A, C]\,.$$

Properties. The following commutators can be derived rather easily

$$[\,\hat{L}_j, \hat{p}_n\,] = +i\hbar\,\epsilon_{jnk}\hat{p}_k\,, \qquad [\,\hat{L}_j, \hat{x}_n\,] = +i\hbar\,\epsilon_{jnk}\hat{x}_k\,. \tag{16}$$

$$[\,\hat{L}_j, \hat{p}^2\,] = 0\,, \qquad [\,\hat{L}_j, \hat{x}^2\,] = 0\,. \tag{17}$$

where $\hat{p}^2 = \hat{p}_1^2 + \hat{p}_2^2 + \hat{p}_3^2$ and $\hat{x}^2 = \hat{x}_1^2 + \hat{x}_2^2 + \hat{x}_3^2$.

1) $[\,\hat{L}_j, \hat{p}_n\,] = [\,\epsilon_{jrs}x_r\hat{p}_s, \hat{p}_n\,] = \epsilon_{jrs}[\,x_r, \hat{p}_n\,]\hat{p}_s = \epsilon_{jrs}i\hbar\delta_{rn}\hat{p}_s = i\hbar\,\epsilon_{jnk}\hat{p}_k$

2) $[\,\hat{L}_j, \hat{x}_n\,] = [\,\epsilon_{jrs}x_r\hat{p}_s, x_n\,] = \epsilon_{jrs}x_r[\,\hat{p}_s, x_n\,] = \epsilon_{jrs}x_r(-i\hbar\delta_{sn}) = i\hbar\,\epsilon_{jnk}x_k$

3) $[\,\hat{L}_j, \hat{p}^2\,] = \sum_k [\,\hat{L}_j, \hat{p}_k^2\,] = \hat{p}_k[\,\hat{L}_j, \hat{p}_k\,] + [\,\hat{L}_j, \hat{p}_k\,]\hat{p}_k$

$$= \hat{p}_k\, i\hbar\,\epsilon_{jks}\hat{p}_s + i\hbar\,\epsilon_{jks}\hat{p}_s\hat{p}_k = \, i\hbar\,\epsilon_{jks}\hat{p}_k\hat{p}_s + i\hbar\,\epsilon_{jsk}\hat{p}_k\hat{p}_s = 0$$

since $\epsilon_{jks} = -\epsilon_{jsk}$. Note that, in addition, $\sum_k \sum_s \epsilon_{jks}F_{ks} = 0$ for any symmetric symbol $F_{ks} = F_{sk}$ and thus the two terms vanish also separately.

4) $[\,\hat{L}_j, \hat{x}^2\,] = \sum_k [\,\hat{L}_j, x_k^2\,] = x_k[\,\hat{L}_j, x_k\,] + [\,\hat{L}_j, x_k\,]x_k$

$$= x_k\, i\hbar\,\epsilon_{jks}x_s + i\hbar\,\epsilon_{jks}x_s x_k = \, i\hbar\,\epsilon_{jks}x_k x_s + i\hbar\,\epsilon_{jsk}x_k x_s = 0$$

Canonical quantization.

Note that the **same results** obtained from the previous commutators by using operators $\hat{x}_j$, $\hat{p}_j$, $\hat{x}^2$, $\hat{p}^2$ and $\hat{L}_j$ can be found for their classical counterparts x_j, p_j, $\vec{x}^2$, $\vec{p}^2$ and L_j when using the Poisson brackets in place of commutators and setting $i\hbar \equiv 1$. So, for example, one finds $\{L_j, p_n\} = \epsilon_{jnk}x_k$. Commutators and Poisson brackets thus feature an evident similarity. In particular, this allows one to **establish a formal connection** with Hamiltonian Mechanics and the definition of Poisson brackets

$$\{A, B\} = \sum_n \left(\frac{\partial A}{\partial x_n}\frac{\partial B}{\partial p_n} \, - \, \frac{\partial B}{\partial x_n}\frac{\partial A}{\partial p_n} \right) \text{ entailing } \{x_k, p_j\} = \, \delta_{kj}$$

when A and B are coordinates and momenta, and to state a new important procedure.

This is the Heisenberg **canonical quantization** procedure which provides a scheme to effect the quantization of classical systems based on such a formal similarity. It states that, in the quantization process, canonical variables x_j, p_k are replaced by operators according to the rule

$$x_j,\, p_k \;\; \Rightarrow Q(x_j) = \hat{x}_j,\, Q(p_k) = \hat{p}_k$$

while **canonical Poisson Brackets are replaced by commutators**

$$\{x_k, p_j\} \;\; \Rightarrow Q(\{x_k, p_j\}) = [x_k, p_j] := \, i\hbar \, \times \, \{x_k, p_j\}\,.$$

The relevance of this, apparently simple, statement stands in its **generality**. Whenever a classical Hamiltonian system must be quantized one knows that, based on the previous rule, canonical variables q_j and p_k must be replaced with operators $\hat{q}_j$ and $\hat{p}_k$ whose commutators $[\hat{q}_j, \hat{p}_n]$ are perfectly known: $[\hat{q}_j, \hat{p}_k] = i\hbar\,\{q_j, p_k\} = i\hbar\delta_{jk}$. If $\hat{q}_j$ are multiplicative operators then the explicit form of operator $\hat{p}_k$ is easily evinced from their commutators (see the example with the Cartesian coordinate x_n and the relavant momentum operators $\hat{p}_k = -i\hbar\partial_k$). Also, in principle, one is able to construct the operator $\hat{A}$ corresponding to any classical function $A(q_j, p_k)$ and the commutator of any pair of operator $\hat{A}$ and $\hat{B}$. Actually, however, the definition of a generic operator $\hat{A}$ is affected by the so-called (unsolved) **ordering problem**. Elementary example: function $x_1^2 p_1$ can be written, quantum-mechanically, in three ways

$$1)\; x_1^2\hat{p}_1\,, \qquad 2)\; x_1\hat{p}_1 x_1 = x_1^2\hat{p}_1 - i\hbar x_1\,, \qquad 3)\; \hat{p}_1 x_1^2 = x_1\hat{p}_1 x_1 - i\hbar x_1 = x_1^2\hat{p}_1 - 2i\hbar x_1\,.$$

Writing simmetrized expressions (to get operators with a Hermitian character) such as

$$x_1\hat{p}_1 x_1\,, \qquad \frac{1}{2}(x_1^2\hat{p}_1 + \hat{p}_1 x_1^2)\,, \qquad \frac{1}{3}(x_1^2\hat{p}_1 + x_1\hat{p}_1 x_1 + \hat{p}_1 x_1^2)$$

does not solve the problem since many simmetrized equivalent forms are available.

Definition of scalar product of two wave functions. This definition is particularly important. Its role will be discussed in more detail in the section devoted to the Hilbert spaces of physical states. Wave functions representing such physical states can be viewed as vectors and $L^2(V)$ exhibits the structure of **vector space**. The scalar product in $L^2(V)$ is given by

$$(\phi, \psi) := \int d^3x\, \phi^*(\vec{x})\, \psi(\vec{x}) = (\psi, \phi)^*\,, \qquad (\psi, \psi) := \int d^3x\, |\psi(\vec{x})|^2 \; > \; 0\,.$$

While the usual scalar product $(\vec{a}, \vec{b}\,) := \vec{a}\cdot\vec{b} = \sum_k a_k b_k$ of real vector spaces is bilinear

$$(\lambda_1\vec{a}_1 + \lambda_2\vec{a}_2)\cdot\vec{b} = \lambda_1\,\vec{a}_1\cdot\vec{b} + \lambda_2\,\vec{a}_2\cdot\vec{b}\,, \quad \vec{a}\cdot(\gamma_1\vec{b}_1 + \gamma_2\vec{b}_2) = \gamma_1\,\vec{a}\cdot\vec{b}_1 + \gamma_2\,\vec{a}\cdot\vec{b}_2$$

in the present case the form (ϕ, ψ) is **sesquilinear** namely

$$(\,c_1\phi_1 + c_2\phi_2,\, \psi) = c_1^*(\phi_1,\, \psi) + c_2^*(\phi_2,\, \psi)\,, \quad (\phi,\, c_1\psi_1 + c_2\psi_2) = c_1(\phi,\, \psi_1) + c_2(\phi,\, \psi_2)$$

Similar to matrices M of some N-dimensional vector space acting on (ordinary) vectors $\vec{a}$, $\vec{b}$ for which one has

$$\vec{c} = M\vec{a}\,,\;\; c_j = \sum_k M_{jk} a_k \;\;\Rightarrow\;\; \vec{b}\cdot\vec{c} = \sum_j b_j c_j = \sum_j \sum_k b_j M_{jk} a_k = \vec{b}\cdot M\vec{a}\,,$$

given an operator A and a wave function $\xi(\vec{x}) = A\,\psi(\vec{x})$ one can obtain a similar formula

$$(\phi, \xi) := \int d^3x\, \phi^*(\vec{x})\, \xi(\vec{x}) = \int d^3x\, \phi^*(\vec{x})\, A\psi(\vec{x}) = (\phi, A\psi)\,.$$

Note the formal analogy with the scalar product $\vec{b}\cdot M\vec{a} = \Sigma_j b_j (M\vec{a})_j$: the role of the summation is played by integral while function $\phi(\vec{x})$ (as well as $\xi(\vec{x})$) represents the vector component labeled by the (vector) "index" $\vec{x}$.

This in turn allows one to define the **adjoint operator**: operator A^+ is said to be the adjoint operator to A if the following equality holds

$$(\phi, A\psi) = (A^+\phi, \psi)$$

Explicitly,

$$(\phi, A\psi) := \int d^3x\, \phi^*(\vec{x})\, A\psi(\vec{x}) \equiv (A^+\phi, \psi) := \int d^3x\, \left(A^+\phi(\vec{x})\right)^*\, \psi(\vec{x})$$

Since $(\psi_1, \psi_2)^* = (\psi_2, \psi_1)$ the latter definition can assume the equivalent form

$$(\psi, A^+\phi) = (\phi, A\psi)^*\,. \tag{18}$$

Property. One can show that $(AB)^+ = B^+A^+$. To prove such a formula, consider that, owing to definition (18), one has

$$\big(\psi, (AB)^+\phi\big) = \big(\phi, (AB)\psi\big)^* = \big((AB)\psi, \phi\big)$$

On the other hand, from definition (18) $(\psi, A^+\phi) = (\phi, A\psi)^* = (A\psi, \phi)$, so that

$$\big((AB)\psi, \phi\big) = \big(A(B\psi), \phi\big) = \big(B\psi, A^+\phi\big) = \big(\psi, B^+A^+\phi\big) \;\;\Rightarrow \big(\psi, (AB)^+\phi\big) = \big(\psi, B^+A^+\phi\big)$$

so that $(AB)^+ = B^+A^+$. Using instead the first definition one has

$$\big((AB)^+\psi, \phi\big) = \big(\psi, (AB)\phi\big)$$

and

$$\big(\psi, (AB)\phi\big) = \big(\psi, A(B\phi)\big) = \big(A^+\psi, B\phi\big) = \big(B^+(A^+)\psi, \phi\big)$$

leading to the same conclusion.

An operator is **hermitian** (or self-adjoint) if $A = A^+$. Then

$$(\psi, A^+\phi) := (\phi, A\psi)^* \equiv (\psi, A\phi)$$

This property will be discussed in detail in the subsequent sections.

4.4 Time-evolution of wave functions and stationary states

If Hamiltonian H is time independent the Schrödinger problem (SP)

$$i\hbar\partial_t\psi(\vec{x},t) = H\,\psi(\vec{x},t)\,, \qquad H = \frac{\hat{p}^2}{2m} + U(\vec{x})\,,\ \ \hat{p}^2 = -\hbar^2\nabla^2$$

can be solved by assuming that $\psi(\vec{x},t) = f(t)\psi(\vec{x})$. In this case the previous SP becomes

$$i\hbar\psi(\vec{x})\partial_t f(t) = f(t)\,H\,\psi(\vec{x}) \quad \Rightarrow \quad i\hbar\frac{1}{f(t)}\partial_t f(t) = \frac{1}{\psi(\vec{x})}\,H\,\psi(\vec{x})$$

Since the left and right sides depend on the independent variables t and $\vec{x}$, respectively, then the only posssibilty to solve this equation is that

$$i\hbar\frac{1}{f(t)}\partial_t f(t) = E = \frac{1}{\psi(\vec{x})}\,H\,\psi(\vec{x})$$

where E is a constant. In this case

$$f(t) = e^{-iE\,t/\hbar}\,, \qquad H\,\psi(\vec{x}) = E\,\psi(\vec{x}) \tag{19}$$

where the latter equation is called the **time-independent Schrödinger equation** and solutions

$$\psi(\vec{x},t) = e^{-iE\,t/\hbar}\psi(\vec{x})$$

are known as the **stationary states** or stationary solutions. Similar to matrix problems where one looks for the eigenvectors (and the relevant eigenvalues) of some matrix, the time-independent Schrödinger equation represents an **eigenvalue equation**

$$\left(-\frac{\hbar^2}{2m}\nabla^2 + U(\vec{x})\right)\psi_E(\vec{x}) = E\,\psi_E(\vec{x})$$

where one looks for wave functions $\psi_E(\vec{x})$ such that the action $H\,\psi_E$ of operator H simply gives $E\,\psi_E$. Stationary states ψ have been labeled with the subscript index E to remind us that each **eigenfunction** ψ_E is associated to a specific eigenvalue E of H (these are called energy eigenvalues). In the subsequent sections this similarity will be evidenced when interpreting operator H (and any other physical operator) as infinite-dimensional matrix (Dirac's picture of Quantum Mechanics). It is important to emphasize how solving $H\,\psi_E = E\,\psi_E$ amounts to **solving a differential equation** due to the presence in H of the kinetic-energy operator.

The **most general solution** of the Schrödinger problem is a linear superposition of different energy eigenfunctions of the form

$$\psi(\vec{x},t) = \sum_n C(E_n)\,e^{-iE_n\,t/\hbar}\psi_{E_n}(\vec{x})\,, \qquad \psi(\vec{x},t) = \int_D dE\,C(E)\,e^{-iE\,t/\hbar}\psi_E(\vec{x}) \tag{20}$$

The first formula refers to Hamiltonians with a discrete spectrum of eigenvalues E_0, E_1, ... E_n ... while the second one refers to Hamiltonians with a continuous spectrum where $E \in D \subseteq \mathbb{R}$. Situations involving a mixed spectrum are also rather common.

Formal solution of the SP. Once the initial wave function $\psi(\vec{x},0)$ describing the state of the system at $t=0$ is given, the Schrödinger equation $i\hbar\partial_t\psi(\vec{x},t) = H\,\psi(\vec{x},t)$ is formally solved by

$$\psi(\vec{x},t) = e^{-it\,H/\hbar}\,\psi(\vec{x},0)\,. \tag{21}$$

An important application of this formula is the derivation (see section 11) of the Heisenberg representation of quantum systems (in which physical operators are time dependent and physical states do not evolve in time) from the Schrödinger representation (in which physical states are time dependent). This formula deserves to be mentioned for another important reason: it represents the very initial step for developing an alternative formulation of Quantum Mechanics, the Path-Integral formulation due to Richard Feynman (1948).

1) Consider the simple case of stationary states $\psi(\vec{x},t) = e^{-iE\,t/\hbar}\psi_E(\vec{x})$. After observing that the initial state is such that $\psi(\vec{x},0) \equiv \psi_E(\vec{x})$, one has

$$\psi(\vec{x},t) = e^{-it\,H/\hbar}\,\psi(\vec{x},0) = e^{-it\,H/\hbar}\,\psi_E(\vec{x}) = \left[\mathbb{I} - it\,\frac{H}{\hbar} + \frac{1}{2}(-it)^2\left(\frac{H}{\hbar}\right)^2 + \,...\right]\,\psi_E(\vec{x})$$

where $e^{-it\,H/\hbar}$ has been expressed via the Taylor expansion of the exponential function. One should recall that, similar to matrices, functions $f(A)$ of a physical operator A can be expressed, with the due care, by means of the Taylor expansion. If $A = H$

$$H\psi_E(\vec{x}) = E\,\psi_E(\vec{x}) \;\Rightarrow\; H^s\psi_E(\vec{x}) = E^s\,\psi_E(\vec{x}) \;\Rightarrow$$

$$F(H)\,\psi_E(\vec{x}) = \sum_s F_s\,H^s\,\psi_E(\vec{x}) = \sum_s F_s\,E^s\,\psi_E(\vec{x}) = F(E)\,\psi_E(\vec{x})$$

Then

$$\psi(\vec{x},t) = \sum_{s=0}^{\infty}\frac{(-it)^s}{s!}\left(\frac{H}{\hbar}\right)^s\,\psi_E(\vec{x}) = \sum_{s=0}^{\infty}\frac{(-it)^s}{s!}\left(\frac{E}{\hbar}\right)^s\,\psi_E(\vec{x}) = e^{-it\,E/\hbar}\,\psi_E(\vec{x})$$

reproducing the simple solution corresponding to time-dependent stationary states.
2) In the **general case**, $\psi(\vec{x},0)$ is not a stationary state but a generic initial state. The following calculation confirms the validity of the formal solution and offers, in parallel, a nice example where the time derivative of an operator is effected

$$i\hbar\partial_t\psi(\vec{x},t) = i\hbar\partial_t\left(e^{-it\,H/\hbar}\,\psi(\vec{x},0)\right) = i\hbar\partial_t\left(e^{-it\,H/\hbar}\right)\,\psi(\vec{x},0)$$

$$= i\hbar\partial_t\sum_{s=0}^{\infty}\frac{(-i)^s t^s}{\hbar^s\,s!}\,H^s\,\psi(\vec{x},0) = i\hbar\sum_{s=0}^{\infty}\frac{s\,(-i)^s t^{s-1}}{\hbar^s\,s!}\,H^s\,\psi(\vec{x},0) = i\hbar\sum_{s=1}^{\infty}\frac{s\,(-i)^s t^{s-1}}{\hbar^s\,s!}\,H^s\,\psi(\vec{x},0)$$

$$= i\hbar\sum_{n=0}^{\infty}\frac{(n+1)\,(-i)^{n+1}t^n}{\hbar^{n+1}\,(n+1)!}\,H^{n+1}\,\psi(\vec{x},0) = \,H\sum_{n=0}^{\infty}\frac{(-i)^n t^n}{\hbar^n\,n!}\,H^n\,\psi(\vec{x},0) = \,H\,\psi(\vec{x},t)\,.$$

Formula

$$\psi(\vec{x},t) = \sum_{s=0}^{\infty}\frac{(-it)^s}{s!}\left(\frac{H}{\hbar}\right)^s\,\psi(\vec{x},0) \tag{22}$$

can be truncated, for times t small enough, thus providing an approximation to state $\psi(\vec{x},t)$. Formula (22) can be utilized as well within the Theory of Coherent States. Thanks to the properties of coherent states (and, particularly, due to the properties of the dynamical algebra associated to H), each term $H^s\psi(\vec{x},0)$, for each s, can be calculated iteratively and the summation on all the terms can be effected producing a well-defined analytic expression of $\psi(\vec{x},t)$.

Remark. Formula (22) clearly shows how the time evolution is a **unitary process** which preserves the norm of the initial state. After defining $U_t := e^{-it\,H/\hbar}$, one has

$$(\psi(\vec{x},t),\psi(\vec{x},t)) = (U_t\psi(\vec{x},0),U_t\psi(\vec{x},0)) = (\psi(\vec{x},0),U_t^+U_t\psi(\vec{x},0)) = (\psi(\vec{x},0),\psi(\vec{x},0) = 1$$

where one easily checks that $U_t^+ = (e^{-it\,H/\hbar})^+ = e^{+it\,H^+/\hbar} = e^{+it\,H/\hbar} = U_t^{-1}$ entailing that $U_t^+U_t = \mathbb{I}$ or equivalently $U_t^{-1} = U_t^+$, the equation defining unitary operators. This result crucially depends on the **hermitian character** of the Hamiltonian: $H = H^+$. The latter property is discussed in Section 4.6, within the derivation of the Ehrenfest Theorem.

In conclusion, one can check whether formula (21) contains the same amount of information embodied in the general solution (20). By considering the initial state as a superposition $\psi(\vec{x},0) = \sum_n C(E_n)\psi_{E_n}(\vec{x})$ of energy eigenfunctions one finds

$$\psi(\vec{x},t) = e^{-it\,H/\hbar}\,\psi(\vec{x},0) = \sum_n C(E_n)e^{-it\,H/\hbar}\,\psi_{E_n}(\vec{x}) = \sum_n C(E_n)\,e^{-iE_n\,t/\hbar}\psi_{E_n}(\vec{x}) \qquad (23)$$

showing how (21) and the general solution (20) are **perfectly equivalent**.

Time-dependent Hamiltonians. A formal solution of the Schrödinger equation can be found even for a time-dependent Hamiltonian $H(t)$. By time integrating the Schrödinger equation one obtains the formula

$$\psi(\vec{x},t) = \psi(\vec{x},0) + \int_0^t dt_1 \frac{H(t_1)}{i\hbar}\psi(\vec{x},t_1)\,,$$

which can be exploited to obtain the formal solution we are interested in through a simple iterative scheme (see also the simple example at the end of this section). By expressing, in turn, $\psi(\vec{x},t_1)$ in terms of this integral formula one finds

$$\psi(\vec{x},t) = \psi(\vec{x},0) + \int_0^t dt_1 \frac{H(t_1)}{i\hbar}\left[\psi(\vec{x},0) + \int_0^{t_1} dt_2 \frac{H(t_2)}{i\hbar}\psi(\vec{x},t_2)\right]$$

$$= \psi(\vec{x},0) + \int_0^t dt_1 \frac{H(t_1)}{i\hbar}\psi(\vec{x},0) + \int_0^t dt_1 \frac{H(t_1)}{i\hbar}\int_0^{t_1} dt_2 \frac{H(t_2)}{i\hbar}\psi(\vec{x},t_2).$$

Effecting once more this substitution for $\psi(\vec{x},t_2)$ gives

$$\psi(\vec{x},t) = \psi(\vec{x},0) + \int_0^t dt_1 \frac{H(t_1)}{i\hbar}\psi(\vec{x},0) + \int_0^t dt_1 \frac{H(t_1)}{i\hbar}\int_0^{t_1} dt_2 \frac{H(t_2)}{i\hbar}\psi(\vec{x},0) + R_3(H)$$

with

$$R_3(H) = \int_0^t dt_1 \frac{H(t_1)}{i\hbar} \int_0^{t_1} dt_2 \frac{H(t_2)}{i\hbar} \int_0^{t_2} dt_3 \frac{H(t_3)}{i\hbar} \psi(\vec{x}, t_3).$$

After the Nth iteration one has

$$\begin{aligned}\psi(\vec{x},t) = \Big[\mathbb{I} &+ \int_0^t dt_1 \frac{H(t_1)}{i\hbar} + \int_0^t dt_1 \frac{H(t_1)}{i\hbar} \int_0^{t_1} dt_2 \frac{H(t_2)}{i\hbar} \\ &+ \int_0^t dt_1 \frac{H(t_1)}{i\hbar} \int_0^{t_1} dt_2 \frac{H(t_2)}{i\hbar} \int_0^{t_2} dt_3 \frac{H(t_3)}{i\hbar} ... \\ &+ \int_0^t dt_1 \frac{H(t_1)}{i\hbar} \int_0^{t_1} dt_2 \frac{H(t_2)}{i\hbar} ... \int_0^{t_{N-1}} dt_N \frac{H(t_N)}{i\hbar} \Big] \psi(\vec{x},0) + R_{N+1}(H)\end{aligned}$$

namely

$$\psi(\vec{x},t) = \left[\mathbb{I} + \sum_{k=1}^{N} \int_0^t dt_1 \frac{H(t_1)}{i\hbar} \int_0^{t_1} dt_2 \frac{H(t_2)}{i\hbar} \int_0^{t_2} dt_3 \frac{H(t_3)}{i\hbar} ... \int_0^{t_{k-1}} dt_k \frac{H(t_k)}{i\hbar} \right] \psi(\vec{x},0)$$

$$+ R_{N+1}(H) \tag{24}$$

with

$$R_{N+1}(H) = \int_0^t dt_1 \frac{H(t_1)}{i\hbar} \int_0^{t_1} dt_2 \frac{H(t_2)}{i\hbar} ... \int_0^{t_N} dt_{N+1} \frac{H(t_{N+1})}{i\hbar} \psi(\vec{x}, t_{N+1}).$$

The expression (24) giving $\psi(\vec{x},t)$ is called the **Neumann series**. For $N \to \infty$ the residual state $R_{N+1}(H)$ can be shown to tend to zero so that

$$\psi(\vec{x},t) = \left[\mathbb{I} + \sum_{k=1}^{\infty} \int_0^t dt_1 \frac{H(t_1)}{i\hbar} \int_0^{t_1} dt_2 \frac{H(t_2)}{i\hbar} \int_0^{t_2} dt_3 \frac{H(t_3)}{i\hbar} ... \int_0^{t_{k-1}} dt_k \frac{H(t_k)}{i\hbar} \right] \psi(\vec{x},0)$$

The series contained in the square brackets represents the generalization of the exponential operator $e^{-iHt/\hbar}$ when H is time dependent. It is interesting to note how, if one assumes that H is time independent, the kth term of the previous summation reduces to

$$\int_0^t dt_1 \int_0^{t_1} dt_2 \int_0^{t_2} dt_3 ... \int_0^{t_{k-1}} dt_k \frac{H^k}{(i\hbar)^k} = \frac{H^k}{(i\hbar)^k} \frac{t^k}{k!}$$

thus reproducing the series associated to the exponential $e^{-iHt/\hbar}$ of the formal solution (21) (see also (22)).

Comment. The essence of the previuos iterative scheme can be exemplified by the following problem. The geometric series $S(x) = 1 + x + x^2 + x^3 + ... = \sum_{k=0}^{\infty} x^k$ features the nice property

$$S(x) = 1 + xS(x) \;\; \text{implying that} \;\; S(x) = \frac{1}{1-x}\,.$$

This provides an elementary scheme to find the analytic formula resulting from the summation of infinitely many terms. Consider the "inverse" problem: we do not know function $S(x)$ but we know the equation $S(x) = 1 + xS(x)$ it satisfies. By utilizing this formula in an iterative way, from

$$S(x) = 1 + xS(x) = S(x) = 1 + x(1 + xS(x)) = 1 + x + x^2 S(x) = 1 + x + x^2(1 + xS(x)) =$$

$$= 1 + x + x^2 + x^3 S(x) = ... = 1 + x + x^2 + ...x^k + x^{k+1} S(x)$$

we recover, for $k \to \infty$ and $|x| < 1$, the explicit formula $S(x) = \sum_{k=0}^{\infty} x^k$ describing $S(x)$ as a series. In the same spirit, the iterative scheme described above shows how the time-integral form of the Schrödinger equation can be used to find the corresponding formal solution as an infinite series of repeated time integrals.

4.5 Free-particle Schrödinger problem and localized states

The Schrödinger problem for a free particle in $V = \mathbb{R}^3$ is easily solved

$$i\hbar\partial_t\phi_k(\vec{x},t) = -\frac{\hbar^2}{2m}\,\nabla^2\phi_k(\vec{x},t)\,, \qquad \phi_k(\vec{x},t) = \frac{e^{i(\vec{k}\cdot\vec{x}-wt)}}{\sqrt{(2\pi)^3}}\,, \qquad w = \frac{\hbar k^2}{2m} = \frac{E}{\hbar}\,.$$

Physically, the stationary solutions $\phi_k(\vec{x},t)$ are simple plane waves which, at $t = 0$, reduce to $\phi_k(\vec{x})$. The latter, in turn, are the solutions of the **eigenvalue equation** (19) for Hamiltonian H representing the **time-independent Schrödinger equation**. The general solution has the form

$$\phi(\vec{x},t) = \int d^3k\, a(\vec{k})\,\phi_k(\vec{x},t) = \int d^3k\, a(\vec{k})\,\frac{1}{\sqrt{(2\pi)^3}}\,\exp\left[i\left(\vec{k}\cdot\vec{x} - \frac{\hbar k^2}{2m}\,t\right)\right]$$

Proof: the left side of the preceding Schrödinger equation

$$i\hbar\partial_t\phi(\vec{x},t) = \int d^3k\, a(\vec{k})\,(i\hbar\partial_t)\phi_k(\vec{x},t) = \int d^3k\, a(\vec{k})\,\hbar\omega\phi_k(\vec{x},t) = \int d^3k\, a(\vec{k})\,\frac{\hbar^2k^2}{2m}\,\phi_k(\vec{x},t)$$

clearly coincides with the right side of the same equation

$$-\frac{\hbar^2}{2m}\,\nabla^2\phi(\vec{x},t) = \int d^3k\, a(\vec{k})\,\left(-\frac{\hbar^2}{2m}\,\nabla^2\right)\,\phi_k(\vec{x},t) = \int d^3k\, a(\vec{k})\,\left(\frac{\hbar^2k^2}{2m}\right)\,\phi_k(\vec{x},t)$$

Then the two quantities coincide. Functions $\phi_k(\vec{x})$ will be shown to form a basis in the section devoted to the Hilbert space of physical states.

Localized states. Assume that

$$a(\vec{k}) = A\,e^{-(\vec{k}-\vec{q})^2\,d^2}$$

$$\psi(\vec{x},t) = \frac{A}{\sqrt{(2\pi)^3}}\int d^3k\, e^{-(\vec{k}-\vec{q})^2\,d^2}\,\exp\left[i\left(\vec{k}\cdot\vec{x} - s\,k^2\right)\right] \qquad \leftarrow \quad s = s(t) = \frac{\hbar\,t}{2m}$$

By observing that

$$(\vec{k}-\vec{q})^2 = (k_1-q_1)^2 + (k_2-q_2)^2 + (k_3-q_3)^2\,, \quad \vec{k}\cdot\vec{x} = k_1x_1 + k_2x_2 + k_3x_3$$

the integral can be rewritten as

$$\psi(\vec{x},t) = \frac{A}{\sqrt{(2\pi)^3}}\int_{\mathbb{R}}dk_1\int_{\mathbb{R}}dk_2\int_{\mathbb{R}}dk_3\,\prod_j e^{-(k_j-q_j)^2\,d^2}\,\exp\left[i\left(k_jx_j - s\,k_j^2\right)\right]$$

$$= \prod_{j=1}^{3}\frac{A_i}{\sqrt{(2\pi)}}\int_{\mathbb{R}}dk_j\,e^{-(k_j-q_j)^2\,d^2}\,\exp\left[i\left(k_jx_j - s\,k_j^2\right)\right] = \prod_{j=1}^{3}\phi_i(x_i,t)$$

The structure of the integral is such that it can be decomposed in three independent factors each one corresponding to the wave function of a one-dimensional problem relevant to one of the three axis x_1, x_2 and x_3

$$I = \int_{\mathbb{R}} dk\, e^{-(k-q)^2\, d^2}\, \exp\left[i\left(kx - s\,k^2\right)\right] = \int_{\mathbb{R}} dk\, \exp\left[-(k^2+q^2-2qk)\, d^2 + \, i(\, kx - s\,k^2\,)\right]$$

$$= \, e^{-q^2\, d^2} \int_{\mathbb{R}} dk\, \exp\left[-(\, d^2 + is\,)\, k^2 \, + (2q\, d^2 \, + ix)\, k\right]$$

$$= \, e^{-q^2\, d^2} \int_{\mathbb{R}} dk\, e^{-\nu\, k^2 \,+\, \eta\, k} \qquad \leftarrow \;\; \nu = d^2 + is\,, \;\; \eta = \, 2q\, d^2 + ix$$

$$= \, e^{-q^2\, d^2} \int_{\mathbb{R}} dk\, \exp\left[-\nu\, \left(k - \frac{\eta}{2\nu}\right)^2 \, + \frac{\eta^2}{4\nu}\right] = \, e^{-q^2\, d^2 \,+\, \frac{\eta^2}{4\nu}} \int_{\mathbb{R}} dk\, e^{-\nu\, \left(k - \frac{\eta}{2\nu}\right)^2}$$

$$= \, e^{-q^2\, d^2 \,+\, \frac{\eta^2}{4\nu}} \int_{\mathbb{R}} dK\, e^{-\nu\, K^2} \, = \, e^{-q^2\, d^2 \,+\, \frac{\eta^2}{4\nu}} \sqrt{\frac{\pi}{\nu}}$$

$$I = \, \exp\left[-q^2\, d^2 \, + \, \frac{(\, 2q\, d^2 + ix\,)^2}{4\nu}\right] \sqrt{\frac{\pi}{\nu}}\,, \qquad \nu = d^2 + is\,, \;\; s = \frac{\hbar\, t}{2m}$$

The resulting wave function is

$$\psi(\vec{x}, t) = \prod_j \phi_j(x_j, t)\,, \qquad \phi_j(x_j, t) = \frac{A_j}{\sqrt{2\pi}} \sqrt{\frac{\pi}{\nu}} \, \exp\left[-q_j^2\, d^2 \, + \, \frac{(\, 2q_j\, d^2 + ix_j\,)^2}{4\nu}\right]. \qquad (25)$$

One easily shows that this wavefunction describes a **localized state**. The probability density is given by $|\psi|^2 = |\phi_1|^2 |\phi_2|^2 |\phi_3|^2$. To find $|\phi_j|^2$ with $j = 1, 2, 3$ one must calculate the real part of the Gaussian-like exponent. Since $q_j^2\, d^2$ is real let us consider the remaining part of the exponent

$$Re\, \frac{(\, 2q_j\, d^2 + ix_j\,)^2}{4\nu} = Re\, \left[\frac{\nu^*}{4|\nu|^2} \left(\, 4q_j^2 d^4 \, - x_j^2 \, + 4i\, x_j q_j d^2\, \right)\right]$$

$$= Re\, \left[\frac{d^2 \, - \, is}{4(d^4 + s^2)} \left(\, 4q_j^2 d^4 \, - x_j^2 \, + 4i\, x_j q_j d^2\, \right)\right]$$

$$= Re\, \left[\frac{d^2}{4|\nu|^2} \left(\, 4q_j^2 d^4 \, - x_j^2 \, \right) + \, \frac{s}{4|\nu|^2} \left(+4\, x_j q_j d^2\, \right) \, + i\, (\, ...\,)\right]$$

$$= d^2\, \frac{(\, 4q_j^2 d^4 \, - x_j^2\,) + 4s\, x_j q_j}{4|\nu|^2} = \, d^2\, \frac{4q_j^2 d^4 \, + 4s^2 q_j^2 - (x_j - 2s\, q_j)^2}{4|\nu|^2}$$

$$= \, 4q_j^2 d^2\, \frac{d^4 \, + s^2}{4|\nu|^2} - \, d^2\, \frac{(x_j - 2s\, q_j)^2}{4|\nu|^2} = \, q_j^2 d^2 \, - \, d^2\, \frac{(x_j - 2s\, q_j)^2}{4(d^4 + s^2)}$$

The resulting probability density has the form

$$|\phi_j(x_j,t)|^2 = \frac{|A_j|^2}{2|\nu|}\, e^{-2q_j^2\, d^2}\, \exp\left[2\, Re\, \left(\frac{(\,2q_j\, d^2 + i x_j\,)^2}{4\nu}\right)\right]$$
$$= \frac{|A_j|^2}{2|\nu|}\, e^{-2q_j^2\, d^2}\, \exp\left[2\, q_j^2 d^2\, -\, d^2\, \frac{(x_j - 2s\, q_j)^2}{2(d^4+s^2)}\right]$$

$$|\phi_j(x_j,t)|^2 = \frac{|A_j|^2}{2\sqrt{d^4+s^2}}\, \exp\left[-\, d^2\, \frac{(x_j - 2s\, q_j)^2}{2(d^4+s^2)}\right] \qquad 2s\, q_j = 2\frac{\hbar\, t}{2m}\, q_j = \frac{\hbar\, q_j}{m}\, t = \, v_j\, t$$

The probability density $|\phi_j(x_j,t)|^2$ is a Gaussian distribution. Then state $\phi_j(x_j,t)$ describes a **localized state** (Gaussian wave packet) along axis x_j while state $\psi(\vec{x},t)$ describes a **localized state** in the 3-dimensional space. The localization maximum is at $t = 0$ when the Gaussian size is given by d.
The Gaussian maximum corresponds to $x_j = v_j t$ which shows that the velocity of the Gaussian wave packet is v_j. For this reason, at least initially, the Gaussian wave packet mimicks the behavior of classical particles whose momentum and position are always well defined. In the 3-dimensional space, for $t \simeq 0$, the particle is localized in a finite portion of the space and its velocity essentially corresponds to that of the Gaussian peak

$$|\vec{x}| \leq \, \sqrt{2}\, d\,, \qquad \vec{v} = \frac{\hbar\, \vec{q}}{m}\,.$$

Since the denominator of the exponent contains the term $s^2 \propto t^2$, the Gaussian size represented by $\sqrt{2\,(d^4+s^2)}$ gets larger and larger as time goes on. As a consequence, the particle position is less and less defined and, for $t \to \infty$, the Gaussian distribution tends to a flat distribution (complete delocalization of the quantum particle).

Exercise. Prove that, in wavefunction $\psi(\vec{x},t) = \prod_j \phi_j(x_j,t)$ described by equation (25)

$$\phi_j(x_j,t) = \frac{A_j}{\sqrt{2\pi}}\, \sqrt{\frac{\pi}{\nu}}\, \exp\left[-q_j^2\, d^2\, +\, \frac{(\,2q_j\, d^2 + i x_j\,)^2}{4\nu}\,\right].$$
$$= \frac{A_j}{\sqrt{2\pi}}\, \sqrt{\frac{\pi}{\nu}}\, e^{-\frac{d^2}{4|\nu|^2}(x_j - 2sq_j)^2}\, e^{\frac{i}{4|\nu|^2}\left(4q_j x_j d^4 - 4sq_j^2 d^4 + s x_j^2\right)}$$
$$= \frac{A_j}{\sqrt{2\pi}}\, \sqrt{\frac{\pi}{\nu}}\, e^{-\frac{d^2}{4|\nu|^2}\left(x_j - 2sq_j\right)^2}\, e^{i q_j x_j}\, e^{-isq_j^2 + \frac{is}{4|\nu|^2}\left(x_j - 2sq_j\right)^2}$$

in which the second exponential factor represents a **time-independent plane wave**. The latter becomes the dominating term for $t \to \infty$.

4.6 The Ehrenfest Theorem

The average value of a physical operator A relevant to the state ψ of the system is given by

$$\langle A\rangle = (\psi, A\psi) = \int_V d^3x \psi^*(\vec{x},t) A\psi(\vec{x},t)\,.$$

This new quantity, also called the expectation value of operator A relevant to the state ψ, will be shown to possess an intrinsic probabilistic meaning based on the notion of complete system of physical states. This aspect will be discussed in the next sections.
Average values of physical operators exhibit a time evolution with interesting properties. To understand this point it is necessary to show that the Hamiltonian operator satisfies the condition

$$(\phi, H^+\psi) := (\psi, H\phi)^* \equiv (\phi, H\psi)\,,$$

$\forall\ \psi, \phi \in L^2(V)$, where the first equality follows from the adjoint-operator definition. This means that operator H **must be Hermitian**.

$$(\psi, H\phi)^* = \int_V d^3x \left(\psi^* H\,\phi\right)^* = \int_V d^3x\,\psi\left(\frac{-\hbar^2}{2m}\nabla^2 + U\right)\phi^*$$

$$= \frac{-\hbar^2}{2m}\int_V d^3x\,\psi\nabla^2\phi^* + \int_V d^3x\,\psi U\phi^* = (\phi, U\psi) - \frac{\hbar^2}{2m}\int_V d^3x\,\psi\nabla^2\phi^*$$

The last integral can be recast into the following form through a repeated integration by part

$$\int_V d^3x\,\psi\nabla^2\phi^* = \int_V d^3x \left(\nabla(\psi\nabla\phi^*) - (\nabla\psi)(\nabla\phi^*)\right)$$

$$= \int_{S_V} \psi\nabla\phi^*\cdot\vec{u}\,dS - \int_V d^3x\,(\nabla\psi)(\nabla\phi^*) = -\int_V d^3x\,(\nabla\psi)(\nabla\phi^*)$$

$$= -\int_V d^3x \left(\nabla(\phi^*\nabla\psi) - \phi^*(\nabla^2\psi)\right) = \ldots = \int_V d^3x\,\phi^*(\nabla^2\psi)$$

Then

$$(\psi, H\phi)^* = (\phi, U\psi) - \frac{\hbar^2}{2m}\int_V d^3x\,\phi^*(\nabla^2\psi) = \int_V d^3x\,\phi^*\left(\frac{-\hbar^2}{2m}\nabla^2 + U\right)\psi = (\phi, H\psi)\,. \quad (26)$$

Let us consider the time evolution of expectation values. In view of the fact that

$$i\hbar\partial_t\psi = H\,\psi\,, \qquad -i\hbar\partial_t\psi^* = H\,\psi^* \qquad \text{with } H = \frac{-\hbar^2\,\nabla^2}{2m} + U(\vec{x})\,,$$

the time derivative of $\langle A\rangle$

$$\frac{\partial}{\partial t}\langle A\rangle = \int_V d^3x \Big((\partial_t\psi^*)A\psi + \psi^*(\partial_t A)\psi + \psi^* A(\partial_t\psi)\Big) = (\dot{\psi}, A\psi) + (\psi, A\dot{\psi})$$

(we have assumed that $\partial_t A = 0$) gives

$$\frac{\partial}{\partial t}\langle A\rangle = \left(\frac{1}{i\hbar}H\psi, A\psi\right) + (\psi, A\,\frac{1}{i\hbar}H\psi) = -\frac{1}{i\hbar}\,(H\psi, A\psi) + \frac{1}{i\hbar}(\psi, A\,H\psi)$$

Owing to the Hermitian character of Hamiltonian $(\psi, H^+\phi) := (H\psi, \phi) \equiv (\psi, H\phi)$ one has $(H\psi, A\psi) = (\psi, H^+A\psi) = (\psi, HA\psi)$. Then

$$\frac{\partial}{\partial t}\langle A\rangle = -\frac{1}{i\hbar}\,(\psi, H\,A\psi) + \frac{1}{i\hbar}(\psi, A\,H\psi) = \frac{1}{i\hbar}\Big(\psi,\,(A\,H - H\,A)\,\psi\Big) = \frac{1}{i\hbar}(\psi,\,[A\,H]\,\psi)\,.$$

It is worth noting the profound similarity of this formula with that describing the time derivative of (classical) time-independent observables A in terms of Poisson Brackets

$$\frac{dA}{dt} = \{A, H\}\,.$$

As an example, consider the expectation values of the momentum and position operators

$$\frac{\partial}{\partial t}\langle \hat{x}_j\rangle = \frac{1}{i\hbar}(\psi,\,[\hat{x}_j\,H]\,\psi)\,, \qquad \frac{\partial}{\partial t}\langle \hat{p}_j\rangle = \frac{1}{i\hbar}(\psi,\,[\hat{p}_j\,H]\,\psi)$$

$$[\hat{x}_j\,H] = \frac{1}{2m}\,[\hat{x}_j\,\hat{p}^2] = \frac{1}{2m}\,[\hat{x}_j\,\hat{p}_j^2] = \frac{1}{2m}\,\Big(\hat{p}_j[\hat{x}_j\,\hat{p}_j] + [\hat{x}_j\,\hat{p}_j]\hat{p}_j\Big) = \frac{1}{2m}\,2\,i\hbar\,\hat{p}_j = \frac{i\hbar}{m}\,\hat{p}_j\,.$$

$$[\hat{p}_j\,H] = [\hat{p}_j\,U] = -i\hbar\frac{\partial U}{\partial x_j}$$

The previous equations become

$$\frac{\partial}{\partial t}\langle \hat{x}_j\rangle = \frac{1}{i\hbar}\left(\psi,\,\frac{i\hbar}{m}\,\hat{p}_j\,\psi\right) = \frac{1}{m}\,(\psi,\,\hat{p}_j\psi) = \frac{1}{m}\,\langle \hat{p}_j\rangle\,,$$

$$\frac{\partial}{\partial t}\langle \hat{p}_j\rangle = \frac{1}{i\hbar}\left(\psi,\,-i\hbar\frac{\partial U}{\partial x_j}\,\psi\right) = -\left(\psi,\,\frac{\partial U}{\partial x_j}\psi\right) = -\langle\frac{\partial U}{\partial x_j}\rangle\,.$$

These reproduce the well-known classical equations

$$\frac{dx_j}{dt} = \{x_j, H\} = \frac{1}{2m}\{x_j, p_j^2\} = \frac{p_j}{m}\,, \qquad \frac{dp_j}{dt} = \{p_j, U\} = -\frac{\partial U}{\partial x_j}\,,$$

showing that the expectation values of physical operators are governed by the same classical equations of the corresponding classical variables. This result represents the **Ehrenfest Theorem**. Comment: the correspondence is complete only if U contains, at most, linear and quadratic terms of coordinates x_j namely $U = \sum_{ik} U_{ik}x_ix_k + \sum_i c_ix_i$. In this case, one finds

$$W(x_n) = \frac{\partial U}{\partial x_n} = \sum_k (U_{nk} + U_{kn})x_k + c_n \quad \Rightarrow \quad \langle W(x_n)\rangle \equiv W\big(\,\langle x_i\rangle\,\big)$$

This result is in agreement with the Bohr **Correspondence Principle** stating that for large quantum numbers quantum-mechanical laws reduce to classical laws.

5 Hilbert space of physical states

Let us review the main definitions and properties characterizing a linear vector space. A suitable generalization of the latter leads to the definition of **Hilbert space**. This space is equipped with the formal structure apt to represent the space of physical states $L^2(V)$ formed by wave functions.

A linear vector space L over a field F (usually either $F = \mathbb{R}$ or $F = \mathbb{C}$) is a set of elements (called vectors)

$$|a\rangle\,,\ \ |b\rangle\,,\ \ |c\rangle\,,\ ...$$

equipped with two operations: the **vector addition** and the **scalar multiplication**

$$\text{for any } |x\rangle, |y\rangle \in L \ \rightarrow \quad |x\rangle + |y\rangle = |z\rangle \in L\,, \qquad \text{for any } \lambda \in F\,,\ |x\rangle \in L \ \rightarrow\ \lambda|x\rangle = |v\rangle \in L\,.$$

Such operations must satisfy the following axioms

1) $|x\rangle + |y\rangle = |y\rangle + |x\rangle$ (commutativity of addition),

2) $(|x\rangle + |y\rangle) + |z\rangle = |x\rangle + (|y\rangle + |z\rangle) = |x\rangle + |y\rangle + |z\rangle$ (associativity of addition),

3) $(\lambda\,\gamma)\,|x\rangle = \lambda\,(\gamma\,|x\rangle) = \lambda\,\gamma\,|x\rangle$ (associativity of multiplication),

4) $(\lambda + \gamma)\,|x\rangle = \lambda|x\rangle + \gamma|x\rangle\,,\ \ \lambda(|x\rangle) + |y\rangle) = \lambda|x\rangle + \lambda|y\rangle$ (distributivity of multiplication)

5) $|0\rangle + |x\rangle = |x\rangle$ (zero element of addition),

6) $|x'\rangle + |x\rangle = |0\rangle = 0 \ \rightarrow\ |x'\rangle = -|x\rangle$ (inverse element of addition).

Basis of a vector space. A set $\mathcal{B}(L) = \{|x_k\rangle,\, k \in [1, N]\}$ of N linearly independent vectors forms a basis if any vector $|v\rangle \in L$ can be written as a linear combination $|v\rangle = \sum_k v_k|x_k\rangle$ of vectors $|x_k\rangle$. Equivalently, a set of linearly independent vectors is said to be **complete** if it forms a basis. Recall that N vectors are linearly independent when the relation $\sum_k C_k|x_k\rangle = 0$ is satisfied only if all $C_k \in F$ are zero.

Scalar product or inner product. This is an operation (x, y) that associates a number of field F to any pair of vectors $|x\rangle$ and $|y\rangle$. In view of future application assume that $F = \mathbb{C}$. The properties of the scalar product are

$$(x, c_1y_1 + c_2y_2) = c_1(x, y_1) + c_2(x, y_2)\,, \quad (c_1x_1 + c_2x_2, y) = c_1^*(x, y_1) + c_2^*(x, y_2)$$

Due to the second property (the product is antilinear) such a scalar product is said to be **sesquilinear**. Note that if field $F = \mathbb{R}$ then the scalar product is simply "bilinear" because $c_i = c_i^*$. The definition of dual vectors

$$|v\rangle = \sum_k v_k|x_k\rangle \ \Rightarrow\ \langle v| = \sum_k v_k^*\langle x_k|$$

(and thus of dual basis) allows one to get a more explict formula for the scalar product

$$(a,b) = \langle a|b\rangle := \sum_j \sum_k a_j^* b_k \, \langle x_j|x_k\rangle$$

where the explicit value of any product $\langle a|b\rangle$ is thus entrusted to the (knowledge of the) products $\langle x_j|x_k\rangle$ between basis vectors. This definition of scalar product leads to define the **norm** of a vector

$$|a| = \sqrt{\langle a|a\rangle}\,, \qquad \langle a|a\rangle = \sum\nolimits_{k=1}^{N} \sum\nolimits_{j=1}^{N} a_j^* a_k \, \langle x_j|x_k\rangle\,.$$

Orthonormal Basis. To emphasize the fact that we consider an orthonormal basis, we denote $|x_k\rangle$ with $|k\rangle$. If $\mathcal{B}(L) = \{|k\rangle,\, k \in [\,N]\}$ is such that $\langle j|k\rangle = 0$ whenever $j \neq k$ and $\langle k|k\rangle = 1$ (the basis vectors have norm equal to 1) the basis is orthonormal. Then the scalar product simplifies

$$\langle x_j|x_k\rangle = \langle j|k\rangle = \delta_{jk} \;\Rightarrow\quad (a,b) = \langle a|b\rangle := \sum\nolimits_{k=1}^{N} a_k^* b_k\,, \quad \langle a|a\rangle := \sum\nolimits_{k=1}^{N} |a_k|^2\,.$$

The **norm of a vector** in the orthonormal basis thus has the form $|a| = \sqrt{\sum_{k=1}^{N} |a_k|^2}$.

Comment. To better emphasize and clarify the difference between vectors $|a\rangle = \sum_k a_k\,|k\rangle$ and $\langle b| = \sum_j b_j^*\,\langle j|$ (these are called ket vectors and bra vectors, respectively, within the Dirac picture of Quantum Mechanics) one can visualize them by means of the standard matrix representation of such objects. Each ket vector can be seen as a column vector (single-column matrix) while each bra vector corresponds to a row vector (single-row matrix) whose scalar product $\langle a|b\rangle := \sum_{k=1}^{N} b_k^* a_k$ reproduces the standard rule of matrix multiplication.

Euclidean vector space. To conclude this part, let us review the familiar case of Euclidean vector spaces and, in particular, the case of 3D space. In this case the field is that of real numbers: $F = \mathbb{R}$.

$$|v\rangle = \sum_k v_k |k\rangle \;\Rightarrow\; \vec{v} = \sum_k v_k \vec{u}_k$$

where $|k\rangle$ identify with unit vectors $\vec{u}_k$ of the Cartesian frame. The distinction between $|v\rangle$ and $\langle v|$ is only formal and dual vector $\langle v| = \sum_k \langle k|v_k^*$ identifies with $|v\rangle$ since $v_k^* \equiv v_k$. The scalar product is

$$\langle a|b\rangle := \sum\nolimits_{k=1}^{3} b_k a_k = \vec{a}\cdot\vec{b}\,, \qquad \langle a|a\rangle := \sum\nolimits_{k=1}^{3} a_k^2 = |a|^2\,,$$

since the (Cartesian) unit vectors satisfy the relation $\vec{u}_j \cdot \vec{u}_k = \delta_{jk}$.

Hilbert space. An infinite-dimensional (linear vector) space over the field $F = \mathbb{C}$ whose vectors has finite norm

$$|v\rangle = \sum_{k=1}^{\infty} v_k |k\rangle \,, \qquad \langle v|v\rangle = \sum_{k=1}^{\infty} |v_k|^2 \,<\, \infty$$

is called Hilbert space. The fact that $|v| = \sqrt{\langle v|v\rangle}$ is finite is a consequence of the converging character of series representing $\langle v|v\rangle$ which is made of infinitely many contributions. The use of the Cauchy-Schwarz inequality is of crucial importance in ensuring a correct definition of scalar product in the Hilbert space

$$|\langle a|b\rangle|^2 \leq \langle a|a\rangle \, \langle b|b\rangle \;\; \Rightarrow \left| \sum_{k=1}^{\infty} a_k^* b_k \right|^2 \,\leq\, |a|^2 |b|^2 \,<\, \infty$$

Since, by definition, norms $|a|$ and $|b|$ are finite then $|\langle a|b\rangle|$ is finite.

5.1 Vector space of wave functions

The quantum dynamics of a particle contained in the ambient space $V \subseteq \mathbb{R}^3$ and subject to potential $U(\vec{x})$ is governed by the Schrödinger equation

$$i\hbar \, \partial_t \psi = \hat{H} \, \psi \,, \qquad H = \frac{\hat{p}^2}{2m} + U(\vec{x}) \,.$$

Its solutions are time-dependent wave functions $\psi = \psi(\vec{x}, t)$ satisfying, at each time, the normalization condition. For this reason they belong to the space $L^2(V)$ of **square integrable functions**, well-known within the theory of Vector Spaces, for which

$$\int_V d^3x \, |\psi(\vec{x}, t)|^2 = const < \infty \,.$$

The discussion relevant to the continuity equations has shown that **the normalization condition is satisfied at any time**. In this sense it can be seen as a constant of motion. Then, any (normalized) functions of $L^2(V)$ can be viewed as a wave function describing 1) the initial state of the system or 2) the physical state possibly involved in the time evolution of the system at some time t. The fundamental property of space $L^2(V)$ is thus to contain all the wave functions representing the physical states of the system.

The Hilbert-space structure of $L^2(V)$

Axioms (1)-(6) listed in the previous section are satisfied by wave functions. Hence, space $L^2(V)$ has the structure of **vector space**. The principal ingredient in this representation of physical states is that the linear combination of two wave functions of $L^2(V)$ still belongs to $L^2(V)$. Then the additivity property of vectors is the consequence of the already mentioned **superposition principle**. The vector-space structure provides an interesting interpretation of physical states in terms of vectors and involves the fundamental notion of **basis**.

Assume that $L^2(V)$ is equipped with the **scalar product** of two functions. This was defined in the previous section by

$$(\psi, \phi) = \int_V d^3x\, \psi^*(\vec{x})\phi(\vec{x})\,.$$

Then a system of orthonormal functions $\{\phi_k(\vec{x}), k = 1, 2, ...\infty\}$ is defined by

$$(\phi_n, \phi_i) = \int_V d^3x\, \phi_n^*(\vec{x})\phi_i(\vec{x}) = \delta_{ni}\,. \tag{27}$$

Orthonormal Basis. Such a system is said to be **complete** (and in this case is called orthonormal **basis**) if any function ψ of $L^2(V)$ can be represented in terms of its **Fourier series** namely

$$\psi(\vec{x}) \equiv \sum_k c_k\, \phi_k(\vec{x}) \quad \text{with} \quad c_k = (\phi_k, \psi)\,.$$

Thanks to equation (27) the formula for c_k is easily proven

$$(\phi_k, \psi) = \int_V d^3x\, \phi_k^*(\vec{x})\psi(\vec{x}) = \int_V d^3x\, \phi_k^*(\vec{x}) \sum_n c_n\phi_n(\vec{x}) = \sum_n c_n\, (\phi_k, \phi_n) = \sum_n c_n\, \delta_{kn} = c_k\,.$$

The implication of the previous formulas is noticeable: each wave function appears to be identified by an (infinte) sequence of numbers c_1, c_2, ...c_n, ... which can be viewed as the components of vector ψ. This **suggests** that there exists a one-to-one correspondence between functions ψ and the vectors of an infinite-dimensional vector space (Hilbert space)

$$\psi(\vec{x}) \qquad \Leftrightarrow \qquad |c\rangle = \sum_{k=1}^{\infty} c_k\, |k\rangle\,. \tag{28}$$

The formula defining the **norm** of a wave function reinforces the vector interpretation of ψ. This is easily expressed in terms of scalar quantities c_k

$$(\psi, \psi) = \int_V d^3x\, \psi^*(\vec{x})\psi(\vec{x}) = \sum_n \sum_k c_n^* c_k \int_V d^3x\, \phi_n^*(\vec{x})\phi_k(\vec{x}) = \sum_k \sum_n c_n^* c_k\, (\phi_k, \phi_n)$$

$$= \sum_k \sum_n c_n^* c_k\, \delta_{ni} = \sum_k |c_k|^2 \quad \Rightarrow \quad (\psi, \psi) = \sum_k |c_k|^2$$

The latter is the well-known **equation of Parseval**.

Comment. For **generic** orthonormal systems the (formal) Fourier series $\sum_k f_k\phi_k$ of some function f can be shown to satisfy the inequality $\sum_k |f_k|^2 \leq (f, f)$. This reflects the fact that $f \sim \sum_k f_k\phi_k$ meaning that, despite the series is the best approximation of f, it does not necessarily converge to f. If the orthonormal systems $\mathcal{B} = \{\phi_k\}$ is complete then $\sum_k f_k\phi_k$ converges "in mean" to $f(x)$ for each point of V. The previous inequality becomes an equation, the Parseval equation, when $\mathcal{B}$ is complete.

Theorem of Fisher-Riesz. We mention, in conclusion, the fundamental theorem of Fisher-Riesz: Given a space $L^2(V)$ equipped with an orthonormal basis $\mathcal{B}(V) = \{\phi_k, k \in [1, \infty]\}$, the inequality

$$\sum_{k=0}^{\infty} |c_k|^2 < \infty \,,$$

where c_k are numbers **choosen arbitrarily**, is a condition necessary and sufficient to ensure that there exists a function $\psi \in L^2(V)$ such that its Fourier series is given by $\psi = \sum_k c_k \phi_k$.

While our preceding discussion showed that ψ is easily associated to an infinite-dimensional vector $|c\rangle$, this theorem fully specifies the relation between functions and vectors evidencing the existence of a one-to-one correspondence between the space $L^2(V)$ and the Hilbert space of vectors. The condition $\sum_k |c_k|^2 < \infty$, valid for some vector $|v\rangle = \sum_k c_k |k\rangle$ of the Hilbert space, is sufficient to state that there exists a square integrable function $\psi = \sum_k c_k \phi_k$ in $L^2(V)$. The space of vectors $|v\rangle$ thus induces the Hilbert-space structure on $L^2(V)$.

Completeness relation. The knowledge of a basis $\{|k\rangle, k \in [1, N]\}$ of a finite-dimensional vector space allows one to find a special way to represent the **identity operator** in terms of the basis vectors and their dual counterpart. Since $\langle i|v\rangle = \sum_k v_k \langle i|k\rangle = \sum_k v_k \delta_{ik} = v_i$, one has

$$|v\rangle = \sum_k v_k |k\rangle = \sum_k |k\rangle\langle k|v\rangle = \left(\sum_k |k\rangle\langle k| \right) |v\rangle \;\; \Rightarrow \;\; \mathbb{I} \equiv \sum_k |k\rangle\langle k|$$

The latter equation is called the **completeness relation**. The use of this formula thus produces the explicit representation of a vector in terms of its components. In the case of (vector) space $L^2(V)$ one has

$$\psi(\vec{x}) = \sum_k c_k \phi_k(\vec{x}) = \sum_k \left(\int_V d^3y \phi_k^*(\vec{y}) \psi(\vec{y}) \right) \phi_k(\vec{x}) = \int_V d^3y \left(\sum_k \phi_k^*(\vec{y}) \phi_k(\vec{x}) \right) \psi(\vec{y})$$

The only function with this property is the Dirac delta function $\delta(y - x)$ whose action in a 1D domain V is, by definition,

$$\int_V dy \, \delta(y - x) f(y) = f(x) \,, \qquad x, y \in V$$

while in a 3D domain $V = X \times Y \times Z$

$$\int_V d^3y \, \delta^3(\vec{y} - \vec{x}) f(\vec{y}) = \int_X dy_1 \int_Y dy_2 \int_Z dy_3 \, \delta(y_1 - x_1)\delta(y_2 - x_2)\delta(y_3 - x_3) f(y_1, y_2, y_3)$$

$$= \int_X dy_1 \int_Y dy_2 \, \delta(y_1 - x_1)\delta(y_2 - x_2) \int_Z dy_3 \, \delta(y_3 - x_3) f(y_1, y_2, y_3)$$

$$= \int_X dy_1 \int_Y dy_2 \, \delta(y_1 - x_1)\delta(y_2 - x_2) f(y_1, y_2, x_3) = \int_X dy_1 \, \delta(y_1 - x_1) \int_Y dy_2 \, \delta(y_2 - x_2) f(y_1, y_2, x_3)$$

$$= \int_X dy_1\, \delta(y_1 - x_1) f(y_1, x_2, x_3) = f(\vec{x})$$

Therefore, in $L^2(V)$, the **completeness relation** reads

$$\delta^3(\vec{y} - \vec{x}) = \sum_{k=1}^{\infty} \phi_k^*(\vec{y}) \phi_k(\vec{x})\,, \tag{29}$$

The latter allows one to represent the action of the identity operator in the space $L^2(V)$

$$\int_V d^3y\, \delta^3(\vec{y} - \vec{x}) \Big(\ldots \Big) = \mathbb{I} \quad \Leftrightarrow \qquad \mathbb{I} \equiv \sum_k |k\rangle\langle k|$$

The formula describing the completeness relation contains an important information: since there are, in principle, infinitely many independent basis in vector space $L^2(V)$, then there exist infinitely many ways of representing the Dirac delta functions.

Formal properties of function $\delta(x)$.

$$\delta(x) = \delta(-x)\,, \qquad \delta(ax) = \frac{1}{|a|}\,\delta(x)\,, \quad a \in \mathbb{R}\,,$$
$$f(x)\,\delta(x - x_0) = f(x_0)\,\delta(x - x_0)\,, \qquad x\,\delta(x) = 0\,.$$

Some definitions of the 1-dimensional Dirac delta function $\delta(x)$

$$\delta(x) := \lim_{n\to\infty} \frac{1}{\pi}\,\frac{n}{1 + n^2x^2}\,, \qquad \delta(x) := \lim_{n\to\infty} \frac{1}{\pi}\,\frac{\sin^2(nx)}{nx^2}\,, \qquad \delta(x) := \lim_{n\to\infty} \frac{n\, e^{-n^2x^2}}{\sqrt{\pi}}$$

These definitions must satisfy the condition

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

Each definition is valid for a class of functions with specific properties. For example, the first definitions works with functions such that $f(x)/|x| \to 0$ for $|x| \to \infty$. Conversely, if one considers $f(x) = e^{cx}$ only the third definition can be used. Another extremely common definition, which is valid for the any function of $L^2(\mathbb{R})$, is given by the integral formula

$$\delta(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} dk\, e^{ik\,x}\,.$$

5.2 Plane-wave basis in a box and Fourier transforms

The system of **plane waves in a box** (we consider plane waves with $t = 0$)

$$\phi_k(\vec{x}) = \frac{e^{i\vec{k}\cdot\vec{x}}}{\sqrt{L^3}}\,, \qquad \vec{k} = \frac{2\pi}{L}\,(n_1, n_2, n_3)\,,\ \ n_i \in \mathbb{Z}\,,$$

supplies a particularly clear example of an orthonormal basis. The latter is defined in the (vector) space $L^2(V)$ where $V = \{\vec{x} : |x_1|, |x_2|, |x_3| \leq L/2\}$. In this case the ambient space is a simple cubic box. This artifice is currently used in many applications to take into account the fact that physical systems often feature a finite dimensional character. Functions ϕ_k incorporate the so-called **periodic boundary conditions** ($\phi_k(\vec{x}) = \phi_k(x_1, x_2, x_3)$) being

$$\phi_k(x_1 + L, x_2, x_3) = \phi_k(x_1, x_2 + L, x_3) = \phi_k(x_1, x_2, x_3 + L) = \phi_k(x_1, x_2, x_3)\,.$$

The orthonormality condition is easily proven

$$(\phi_k, \phi_q) = \int_V d^3x\, \frac{e^{-i\vec{k}\cdot\vec{x}}}{\sqrt{L^3}}\, \frac{e^{i\vec{q}\cdot\vec{x}}}{\sqrt{L^3}} = \int_V d^3x\, \frac{e^{i(\vec{q}-\vec{k})\cdot\vec{x}}}{L^3}$$

$$= \int_{-L/2}^{L/2} dx_1 \int_{-L/2}^{L/2} dx_2 \int_{-L/2}^{L/2} dx_3\, \frac{e^{i(q_1-k_1)x_1} e^{i(q_2-k_2)x_2} e^{i(q_3-k_3)x_3}}{L^3} = \prod_j \left(\int_{-L/2}^{L/2} dx_j \frac{e^{i(q_j-k_j)x_j}}{L} \right)$$

$$= \prod_j \left(\frac{e^{i(q_j-k_j)L/2} - e^{-i(q_j-k_j)L/2}}{iL(q_j - k_j)} \right) = \prod_j \left(\frac{e^{\pi i(m_j-n_j)} - e^{-\pi i(m_j-n_j)}}{iL(q_j - k_j)} \right) = 0$$

where $\vec{q} = \frac{2\pi}{L}\,(m_1, m_2, m_3)$ and $\vec{k} = \frac{2\pi}{L}\,(n_1, n_2, n_3)$ with $m_i, n_i \in \mathbb{Z}$. If $\vec{q} = \vec{k}$, and thus $m_j = n_j$ for each j, one finds that $(\phi_k, \phi_k) = 1$ so that, in general, one has

$$(\phi_k, \phi_q) = \delta_{\vec{k},\vec{q}} = \delta_{n_1,m_1}\, \delta_{n_2,m_2} \delta_{n_3,m_3}$$

The following, well-known representation of the Dirac delta function

$$\delta^3(\vec{y} - \vec{x}) = \sum_{\vec{k}} \frac{e^{i\vec{k}\cdot(\vec{x}-\vec{y})}}{L^3} = \sum_{\vec{k}} \frac{e^{i\vec{k}\cdot\vec{x}}}{\sqrt{L^3}}\, \frac{e^{-i\vec{k}\cdot\vec{y}}}{\sqrt{L^3}} \quad \Rightarrow \qquad \delta^3(\vec{y} - \vec{x}) = \sum_{\vec{k}} \phi_k^*(\vec{y})\phi_k(\vec{x})$$

provides the basic ingredient of this example in that the completeness condition (29) appears to be satisfied. This confirms the basic result of Fourier Analysis (and, more in general, of Functional Analysis) that the **plane-wave system** is **complete** in $L^2(V)$. In this basis any function of $L^2(V)$ can be expressed by means of its **Fourier series**

$$\psi(\vec{x}) = \sum\nolimits_{\vec{k}} a_{\vec{k}}\, \phi_k(\vec{x}) \qquad \text{where} \quad a_{\vec{k}} = (\phi_k, \psi) = \int_V d^3y\, \phi_k^*(\vec{y})\psi(\vec{y})\,.$$

The Fourier series describing $\psi(\vec{x})$ in the actual basis is called the **inverse Fourier transform** while $a_{\vec{k}}$ is the **Fourier transform** associated to $\psi(\vec{x})$.

The use of the completeness relation confirms, as expected, this property

$$\psi(\vec{x}) = \int_V d^3y\, \delta^3(\vec{y}-\vec{x})\psi(\vec{y}) = \int_V d^3y \sum_{\vec{k}} \phi_k^*(\vec{y})\phi_k(\vec{x})\psi(\vec{y}) = \sum_{\vec{k}} \phi_k(\vec{x})(\phi_k,\psi) = \sum_{\vec{k}} a_{\vec{k}}\,\phi_k\,.$$

The Parseval equation is expressed by

$$(\psi,\psi) = \int_V d^3x\,|\psi(\vec{x})|^2 = \int_V d^3x\,\psi^*(\vec{x})\,\sum_{\vec{k}} a_{\vec{k}}\,\phi_k(\vec{x}) = \sum_{\vec{k}} a_{\vec{k}} \int_V d^3x\,\psi^*(\vec{x})\phi_k(\vec{x})$$

$$= \sum_{\vec{k}} a_{\vec{k}}\,(\phi_k,\psi)^* = \sum_{\vec{k}} |a_{\vec{k}}|^2$$

Physically, the functions $\phi_k(\vec{x})$ of the basis are simple plane waves representing solutions of the **energy eigenvalue equation**, namely, of the stationary Schrödinger equation **for a free particle in a "periodic" box**

$$\hat{H}\,\phi_k(\vec{x}) = E_k\,\phi_k(\vec{x})\,, \qquad \hat{H} = -\frac{\hbar^2}{2m}\nabla^2\,, \quad E_k = \frac{\hbar^2k^2}{2m} = \hbar\,w_k\,.$$

For a generic t, such a wave functions provide the solutions to time-dependent Schrödinger problem in the box V (with periodic boundary conditions)

$$i\hbar\partial_t\phi_k(\vec{x},t) = \hat{H}\,\phi_k(\vec{x},t)\,, \qquad \phi_k(\vec{x},t) = e^{-i\,E_kt/\hbar}\phi_k(\vec{x}) \;=\; \frac{e^{i(\vec{k}\cdot\vec{x}-E_kt/\hbar)}}{\sqrt{L^3}}\,.$$

For each component of momentum operator $\hat{p}$ they satisfy the equations

$$\hat{p}_i\phi_k(\vec{x}) = \hbar k_i\,\phi_k(\vec{x})\,, \quad k_i = \frac{2\pi n_i}{L}\,, \quad n_i = 0,\,\pm1,\,\pm2\,... \tag{30}$$

showing that the action of $\hat{p}_i$ associates the discretized momentum $\hbar k_i$ to $\phi_k(\vec{x})$. The fact that the particle is assumed to be confined in volume V is responsible for the **discreteness of momentum vectors** $\hbar\vec{k}$.

Comment: At the formal level, this plane-wave basis can be used to represent any wave-funtions ψ defined in V through its Fourier series. Hence, for example, square integrable functions defined in the whole 3D space (and thus in V) are included in this class. This means that the Fourier series of $\psi(\vec{x})$ is well defined for each $\vec{x}\in V$ even if ψ is defined in the whole 3D space.

In general, however, we shall see that the study of a specific system described by some Hamiltonian $H = \hat{p}^2/(2m) + U$ can suggest a **different choice of the basis** owing to the possible **symmetries** or, more in general, to the **geometry** naturally associated with potential U. A careful choice of the basis often leads to a **simpler and more effective description** of physical states ψ relevant to H. A frequent use of the present plane-wave basis is made for systems where potential U exhibits space periodicity (e. g., an electron in a crystal lattice).

This circumstance leads to consider the more general plane-wave basis involving Fourier transforms defined in whole 3D ambient space.

5.3 Comments and examples

Example 1. Consider the 1D box $V = \{\vec{x} : |x| \leq L/2\}$. Any function of $L^2(V)$ can be represented as a **Fourier series** by means of its **Fourier transform**

$$\psi(x) = \sum_k a(k)\,\phi_k(x) \quad \text{with } a(k) = \int_V dy\,\phi_k^*(y)\psi(y)$$

$$\psi(x) = \int_V dy\,\delta(y-x)\psi(y) = \int_V dy \sum_k \phi_k^*(y)\phi_k(x)\psi(y) = \sum_k \phi_k(x) \int_V dy\,\phi_k^*(y)\psi(y)$$

Consider the **simplest function** $f(x) = x$ defined on the whole real axis $\mathbb{R}$. Since

$$\int_{-\infty}^{\infty} dx\,|f(x)|^2 = \int_{-\infty}^{\infty} dx\,x^2 \to \infty\,, \qquad \int_{-L/2}^{L/2} dx\,|f(x)|^2 = \int_{-L/2}^{L/2} dx\,x^2 = \frac{L^3}{12} < \infty$$

function x does not belong to $L^2(\mathbb{R})$. It is instead a square integrable function of $L^2(V)$.

In view of the fact that $k = 2\pi n/L$ is, up to a constant factor, the integer $n = 0, \pm 1, \pm 2, ...$, one can adopt the notation

$$\phi_k(x) = \frac{e^{ikx}}{\sqrt{L}} = \frac{e^{i\,K_n\,x}}{\sqrt{L}} = \phi_n(x)\,, \qquad K_n = \frac{2\pi\,n}{L}$$

Let us calculate the Fourier transform (for short, set $K_n = 2\pi\,n/L$)

$$\begin{aligned} a_n &= \int_V dy\,\phi_n^*(y) f(y) = \int_{-L/2}^{L/2} dy\,\frac{e^{-i\,K_n\,y}}{\sqrt{L}}\,y = -i\int_{-L/2}^{L/2} \frac{dy}{\sqrt{L}}\,y\,\sin(K_n\,y) \\ &= -i\int_{-L/2}^{L/2} \frac{dy}{\sqrt{L}}\left[\frac{1}{K_n}\frac{d}{dy}\left(-y\,\cos(K_n\,y)\right) + \frac{1}{K_n}\cos(K_n\,y)\right] \\ &= -\frac{(-i)}{K_n\sqrt{L}}\Big[y\,\cos(K_n\,y)\Big]_{-L/2}^{L/2} + (-i)\int_{-L/2}^{L/2}\frac{dy}{K_n\sqrt{L}}\,\cos(K_n\,y) \\ &= \frac{i}{K_n\sqrt{L}}\Big[y\,\cos(K_n\,y)\Big]_{-L/2}^{L/2} = \frac{i}{K_n\sqrt{L}}\Big[\frac{L}{2}\,\cos(\pi\,n) - \frac{(-L)}{2}\,\cos(-\pi\,n)\Big] \\ &= \frac{i}{K_n\sqrt{L}}\,L\,\cos(\pi\,n) = (-i)\,\frac{L^{3/2}(-1)^{n+1}}{2\pi\,n} \end{aligned}$$

Then, after observing that $a_0 = 0$

$$x = \sum_{n=-\infty}^{\infty} a_n\,\phi_n(x) = \frac{-i\,L^{3/2}}{2\pi}\sum_{n=-\infty}^{\infty}\frac{(-1)^{n+1}}{n}\,\frac{e^{i\,K_n\,x}}{\sqrt{L}} = \frac{-iL}{2\pi}\sum_{n=-\infty}^{\infty}\frac{(-1)^{n+1}}{n}\,i\,\sin(K_n\,x)$$

The final form of the **Fourier representation of** $f(x) = x$ **reads**

$$x = +\frac{L}{\pi}\sum_{n=1}^{\infty}\frac{(-1)^{n+1}}{n}\,\sin(K_n\,x)$$

This representation is valid in the interval $-L/2 < x < L/2$. The extreme points are excluded because the series exhibits a discontinuity.

Example 2. The action of Dirac's delta function. Consider the one-dimensional Dirac delta function $\delta(y-x)$ and a square integrable function $f(x)$ defined in the interval $V=\{-\infty\leq x\leq\infty\}\equiv\mathbb{R}$. Since the size of this interval is infinitley large this case is similar to the one of plane waves in $\mathbb{R}^3$.

$$\int_V dy\,\delta(y-x)f(y)=f(x)\,,\qquad x,y\in V$$

$$\delta(y-x)=\int_{-\infty}^{\infty}dk\,\frac{e^{ik\,(x-y)}}{2\pi}$$

One can easily **check of the action** of $\delta(y-x)$ in a specific case. After observing that

$$\int_V dy\,\delta(y-x)f(y)=\int_V dy\,\int_{-\infty}^{\infty}dk\,\frac{e^{ik\,(x-y)}}{2\pi}f(y)=\int_{-\infty}^{\infty}dk\,\frac{e^{ik\,x}}{2\pi}\int_V dy\,e^{-iky}f(y)$$

let us consider the function $f(y)=e^{-by^2}$ (this is certainly square integrable). One has

$$\int_V dy\,e^{-iky}f(y)=\int_{-\infty}^{\infty}dy\,e^{-iky}\,e^{-by^2}=\int_{-\infty}^{\infty}dy\,\exp\left[-\frac{k^2}{4b}-b\left(y+\frac{ik}{2b}\right)^2\right]$$

$$=e^{-k^2/(4b)}\int_{-\infty}^{\infty}dy\,\exp\left[-b\left(y+\frac{ik}{2b}\right)^2\right]=e^{-k^2/(4b)}\int_{-\infty}^{\infty}dz\,\exp\left[-bz^2\right]=e^{-k^2/(4b)}\sqrt{\frac{\pi}{b}}$$

in which the use of the new complex variable $z=y+\frac{ik}{2b}$ is equivalent to deforming the integration path. Such a deformation can be shown to be insignificant. Then

$$\int_V dy\,\delta(y-x)f(y)=\,....\,=\int_{-\infty}^{\infty}dk\,\frac{e^{ik\,x}}{2\pi}\,e^{-k^2/(4b)}\sqrt{\frac{\pi}{b}}=\frac{1}{2\sqrt{\pi b}}\int_{-\infty}^{\infty}dk\,e^{ik\,x}\,e^{-k^2/(4b)}$$

This expression should be equal to $f(x)=e^{-bx^2}$. Check

$$\frac{1}{2\sqrt{\pi b}}\int_{-\infty}^{\infty}dk\,e^{ik\,x}\,e^{-k^2/(4b)}=e^{-bx^2}\frac{1}{2\sqrt{\pi b}}\int_{-\infty}^{\infty}dk\,\exp\left[-\frac{1}{4b}\,(k-2bix)^2\right]$$

$$=e^{-bx^2}\frac{1}{2\sqrt{\pi b}}\int_{-\infty}^{\infty}dq\,\exp\left[-\frac{q^2}{4b}\right]=e^{-bx^2}\,\frac{1}{2\sqrt{\pi b}}\times\sqrt{\frac{\pi}{1/4b}}=e^{-bx^2}\,.$$

Again the use of the new complex variable $q=k-2bix$ does not alter the result of the Gaussian integral. Note that this example implicitly involves the use of the Fourier transform $a(k)$ of $f(x)$ and, in the final part, the calculation of the inverse Fourier transform (which is expected to give $f(x)$)

$$a(k)=\int_V dy\,e^{-iky}f(y)=e^{-k^2/(4b)}\sqrt{\frac{\pi}{b}}\,,\qquad\frac{1}{2\sqrt{\pi b}}\int_{-\infty}^{\infty}dk\,e^{ik\,x}\,e^{-k^2/(4b)}=e^{-bx^2}\,.$$

Note that this exercise is equivalent to calculating the Forurier transform of e^{-bx^2} and then to reproducing the initial function by means of $a(k)$ through its inverse Fourier transform.

5.4 Plane-wave basis and quantum states in the 3D space

The **basis** constituted by the system of orthogonal functions (plane waves with $t = 0$)

$$\phi_k(\vec{x}) = \frac{e^{i\vec{k}\cdot\vec{x}}}{(2\pi)^{3/2}}\,, \qquad \vec{k} \in \mathbb{R}^3$$

is another well-known example. Such basis allows to represent the functions of (vector) space $L^2(V)$ in the case when $V = \mathbb{R}^3$ is the whole ambient space.

Completeness relation. The new basis involves the following, crucial definition of the Dirac delta function in $V = \mathbb{R}^3$

$$\delta^3(\vec{y}-\vec{x}) = \int d^3k\, \frac{e^{i\vec{k}\cdot(\vec{x}-\vec{y})}}{(2\pi)^3} = \int d^3k\, \frac{e^{i\vec{k}\cdot\vec{x}}}{(2\pi)^{3/2}}\, \frac{e^{-i\vec{k}\cdot\vec{y}}}{(2\pi)^{3/2}} \;\Rightarrow\; \delta^3(\vec{y}-\vec{x}) = \int d^3k\, \phi_k^*(\vec{y})\phi_k(\vec{x}) \quad (31)$$

where the latter formula represents the **completeness** relation for the present case. The definition of $\delta^3(\vec{y}-\vec{x})$ given in the preceding subsection shows that summation $\sum_{\vec{k}}$ has been replaced with integration $\int d^3k$.

Note that, to simplify the calculations involved in the subsequent discussion, we imply the symbol $\mathbb{R}^3$ for the integration domain both in the space integrals $\int d^3x$ and in the k-dependent integrals $\int d^3k$.

Orthonormality condition: This condition, characterizing a basis of $L^2(V)$ where V is a finite volume (as for the plane-wave basis in a box), in the actual case assumes a simpler, **more general form** due to the fact that $\vec{k}$ is a continuous vector parameter

$$(\phi_k, \phi_q) = \int d^3x\, \frac{e^{-i\vec{k}\cdot\vec{x}}}{(2\pi)^{3/2}}\, \frac{e^{i\vec{q}\cdot\vec{x}}}{(2\pi)^{3/2}} = \int d^3x\, \frac{e^{i(\vec{q}-\vec{k})\cdot\vec{x}}}{(2\pi)^3} = \delta^3(\vec{k}-\vec{q})$$

The last equality is justified by comparing the last integral with definition of the Dirac delta function given in formula (31): one easily sees that in this pair of formulas the role of space vectors $\vec{x}$, $\vec{y}$ and that of the so-called wave vectors $\vec{k}$, $\vec{q}$ have been exchanged. Likewise, integration involves $\vec{x}$ instead of $\vec{k}$. Apart from this, the latter formula is, in turn, a Dirac delta function depending on $\vec{k}-\vec{q}$. The unusual feature of this basis is that, while states with $\vec{k} \neq \vec{q}$ give $(\phi_k, \phi_q) = 0$ and thus are **orthogonal**, for $\vec{k} = \vec{q}$ one has

$$(\phi_k, \phi_k) = \int d^3x\, \frac{1}{(2\pi)^3} \to \infty$$

meaning that **plane waves cannot be normalized** in $\mathbb{R}^3$.

This pathology is not surprising in that $|\phi_k|^2 = 1/(2\pi)^3$ (the probability density relevant to plane waves is constant everywhere). Then, in $\mathbb{R}^3$, the normalization of plane waves necessarily diverges. This problem does not represent an inherent contradiction of the theory discussed so far. It should be viewed as an intrinsic, natural consequence of the fact that one consider a space with $V \to \infty$. What is important is that this "anomaly" in no way affects the representation of wave functions of $L^2\left(\mathbb{R}^3\right)$.

Probabilistic interpretation: This interpretation of the density function $|\psi(\vec{x})|^2$ is still valid. Consider the following integrals

$$P(D) = \int_D d^3x\, |\psi(\vec{x})|^2\,,\;\; P(V_1) = \int_{V_1} d^3x\, |\phi_k|^2 = \frac{V_1}{(2\pi)^3}\,,\;\; P(V_2) = \int_{V_2} d^3x\, |\phi_k|^2 = \frac{V_2}{(2\pi)^3}\,.$$

The first integral, where ψ is some normalizable function, gives the usual information in terms of probability $P(D)$ to find a particle in the domain D. The other two integrals, where ϕ_k is not normalizable, again provide information in terms of probability: the quantities $P(V_1)$ and $P(V_2)$ are proportional to the probability to find a particle in domain V_1 and V_2, respectively. Then $P(V_1)/P(V_2) = V_1/V_2$ can be seen as a relative probability.

Fourier series. Any function of $L^2\left(\mathbb{R}^3\right)$ can be represented as

$$\psi(\vec{x}) = \int d^3k\, a(\vec{k})\, \phi_k(\vec{x}) = \int d^3k\, a(\vec{k})\, \frac{e^{i\vec{k}\cdot\vec{x}}}{(2\pi)^{3/2}} \qquad \text{where} \quad a(\vec{k}) = \int d^3y\, \phi_k^*(\vec{y})\psi(\vec{y})\,,$$

which is equivalent to the completeness of the wave-function basis described by equation (31)

$$\psi(\vec{x}) = \int d^3y\, \delta^3(\vec{y}-\vec{x})\psi(\vec{y}) = \int d^3y \int d^3k\, \phi_k^*(\vec{y})\phi_k(\vec{x})\psi(\vec{y}) = \int d^3k\; \phi_k(\vec{x}) \int d^3y\, \phi_k^*(\vec{y})\psi(\vec{y})$$

While the **Fourier transform** $a(\vec{k})$ associated to $\psi(\vec{x})$ has essentially the same form as the case where V is finite, the **inverse Fourier transform** describing $\psi(\vec{x})$ and previously expressed by a (Fourier) **series** becomes an integral in the actual basis.

Parseval equation. The continuous version in the current basis of the Parseval equation is easily derived

$$(\psi,\psi) = \int d^3x\, |\psi(\vec{x})|^2 = \int d^3x \int d^3k\, a^*(\vec{k})\, \phi_k^*(\vec{x}) \int d^3q\, a(\vec{q})\, \phi_q(\vec{x})$$

$$= \int d^3k \int d^3q\, a^*(\vec{k})\, a(\vec{q}) \int d^3x\, \phi_k^*(\vec{x})\phi_q(\vec{x}) = \int d^3k \int d^3q\, a^*(\vec{k})\, a(\vec{q})\, \delta^3(\vec{k}-\vec{q}) = \int d^3k\, |a(\vec{k})|^2$$

Note that the interpretation of $|\psi(\vec{x})|^2$ as probability density in ambient space V is implicitly extended to $|a(\vec{k})|^2$ which can be seen as the **probability density** in the momentum-vector space $\mathbb{R}^3$ described (up to the constant $\hbar$) by $\vec{k}$.

Time-dependent wavefunctions. If a wavefunction depends on time (this case is very common because any solution of some Schrödinger problem is a time-dependent function) then the defining equation of Fourier transform shows that the latter inherits this time dependence

$$\psi(\vec{x},t) = \int d^3k\, a(\vec{k},t)\, \phi_k(\vec{x}) = \int d^3k\, a(\vec{k},t)\, \frac{e^{i\vec{k}\cdot\vec{x}}}{(2\pi)^{3/2}} \qquad \text{with} \quad a(\vec{k},t) = \int d^3y\, \phi_k^*(\vec{y})\psi(\vec{y},t)\,.$$

The Parseval equation becomes

$$\int d^3x\, |\psi(\vec{x},t)|^2 = \int d^3k\, |a(\vec{k},t)|^2\,.$$

5.5 Expectation values of momentum and position operators

Similar to the plane waves of $L^2(V)$ in a finite periodic box, functions $\phi_k(\vec{x})$ of $L^2(\mathbb{R}^3)$ are **solutions** of the eigenvalue problem for the free-particle Hamiltonian $H = \hat{p}^2/(2m)$ (see equation 19)

$$\hat{H}\phi_k(\vec{x}) \;=\; E(k)\,\phi_k(\vec{x})\,, \quad E(k) = \hbar\, w = \frac{\hbar^2 k^2}{2m}$$

where $E(k)$ is the **eigenvalue** associated to **eigenfunctions** $\phi_k(\vec{x})$ through the latter eigenvalue equation. For this reason plane waves $\phi_k(\vec{x})$ can be used as well to construct the **stationary solutions**, at a generic time $t > 0$, to the free-particle Schrödinger problem in $V = \mathbb{R}^3$

$$i\hbar\partial_t\phi_k(\vec{x},t) = \hat{H}\,\phi_k(\vec{x},t)\,, \qquad \phi_k(\vec{x},t) = \frac{e^{i(\vec{k}\cdot\vec{x}-Et/\hbar)}}{(2\pi)^{3/2}}\,, \qquad E(k) = \frac{\hbar^2 k^2}{2m}\,. \tag{32}$$

Moreover, functions $\phi_k(\vec{x})$ feature the properties similar to those described by equation (30) for the periodic-box plane waves

$$\hat{p}_j\,\phi_k(\vec{x}) = -i\hbar\partial_{x_j}\frac{e^{i\vec{k}\cdot\vec{x}}}{(2\pi)^{3/2}} = \hbar k_j\,\phi_k(\vec{x}) \;\;\Leftrightarrow\;\; \hat{p}\,\phi_k(\vec{x}) = \hbar\vec{k}\,\phi_k(\vec{x})\,. \tag{33}$$

involving operator $\hat{p}$ and components $\hat{p}_j$ thereof. These make it evident that vector $\vec{p} = \hbar\vec{k}$ is associated to the action $\hat{p}$ and, more specifically, that $p_1 = \hbar k_1$, $p_2 = \hbar k_2$ and $p_3 = \hbar k_3$ are the eigenvalues of $\hat{p}_i\,\phi_k = \hbar k_i\,\phi_k$, the eigenvalue equations of the vector components of $\hat{p}$.

All such results will be discussed in the subsequent section within the unifying context of properties and theorems of Hermitian opertors.

This basis is quite general in that it allows one to represent any wave functions $\psi(\vec{x})$ describing the states of any physical system in $V = \mathbb{R}^3$ included the time-dependent states, $\psi(\vec{x},t)$

$$\psi(\vec{x},t) = \int d^3k\, a(\vec{k},t)\,\phi_k(\vec{x}) \quad\Leftrightarrow\quad a(\vec{k},t) = \big(\phi_k(\vec{x}),\,\psi(\vec{x},t)\big)\,. \tag{34}$$

In the **special case** of the free-particle Schrödinger problem (32) the most general solutions is written as

$$\psi(\vec{x},t) = \int d^3k\, a(\vec{k})\,\phi_k(\vec{x},t) = \int d^3k\, a(\vec{k})\,\frac{1}{(2\pi)^{3/2}}\,\exp\left[i\left(\vec{k}\cdot\vec{x} - \frac{\hbar k^2}{2m}\,t\right)\right]$$

$$= \int d^3k\, a(\vec{k},t)\,\frac{1}{(2\pi)^{3/2}}\,e^{i\,\vec{k}\cdot\vec{x}}\,, \qquad a(\vec{k},t) = a(\vec{k})\,\exp\left(-i\frac{\hbar k^2}{2m}\,t\right)\,. \tag{35}$$

showing that the time-dependent exponential factor can be included in the Fourier transform $a(\vec{k})$. An application of this formula has been made for constructing **spatially localized states** (25) characterized by a gaussian distribution of momenta $\vec{p} = \hbar\vec{k}$ around a given value $\vec{k} \equiv \vec{q}$ (see subsection (4.5)).

Let us consider the **general case** where ψ is generic function of $L^2(V)$ not necessarily representing a solution of the free-particle Schrödinger problem. The subsequent discussion shows that the plane-wave basis is particularly useful to understand the physical meaning of the expectation value of momentum operator $\hat{p}$.

Expectation value of position operator $\hat{x} = \vec{x}$. At time t, it is given by

$$(\psi, \vec{x}\,\psi) = \int d^3x\, \vec{x}\, |\psi(\vec{x}, t)|^2$$

showing how the contribution of each position $\vec{x}$ is weighted by the probability density $\rho(\vec{x}, t) = |\psi(\vec{x}, t)|^2$. The meaning of this formula is quite intuitive: if $\rho(\vec{x}, t) = |\psi(\vec{x}, t)|^2$ described the matter density (instead of the probability density) the previous formula would be the position of the center of mass of the system.

Expectation value of momentum operator $\hat{p} = -i\hbar\nabla$. This is very expressive as well. At time t it is given by

$$(\psi, \vec{p}\,\psi) = \int d^3x\, \psi^*(\vec{x}, t)\, (-i\hbar\nabla)\, \psi(\vec{x}, t) = \int d^3x\, \psi^*(\vec{x}, t)\, (-i\hbar\nabla) \int d^3k\, \frac{a(\vec{k}, t)}{(2\pi)^{3/2}}\, e^{i\,\vec{k}\cdot\vec{x}}$$

$$= \int \frac{d^3x}{(2\pi)^{3/2}}\, \psi^*(\vec{x}, t) \int d^3k\, a(\vec{k}, t)\, \hbar\vec{k}\, e^{i\,\vec{k}\cdot\vec{x}} = \int d^3k\, a(\vec{k}, t)\, \hbar\vec{k} \int \frac{d^3x}{(2\pi)^{3/2}}\, \psi^*(\vec{x}, t)\, e^{i\,\vec{k}\cdot\vec{x}}$$

$$= \int d^3k\, a(\vec{k}, t)\, \hbar\vec{k} \int \frac{d^3x}{(2\pi)^{3/2}}\, e^{i\,\vec{k}\cdot\vec{x}} \int d^3q\, a^*(\vec{q}, t)\, \frac{e^{-i\,\vec{q}\cdot\vec{x}}}{(2\pi)^{3/2}}$$

$$= \int d^3k \int d^3q\, a(\vec{k}, t) a^*(\vec{q}, t)\, \hbar\vec{k} \int \frac{d^3x}{(2\pi)^3}\, e^{i\,\vec{k}\cdot\vec{x}}\, e^{-i\,\vec{q}\cdot\vec{x}}$$

$$= \int d^3k \int d^3q\, a(\vec{k}, t) a^*(\vec{q}, t)\, \hbar\vec{k}\, \delta^3(\vec{k} - \vec{q}) = \int d^3k\, a(\vec{k}, t) a^*(\vec{k}, t)\, \hbar\vec{k}$$

Concluding, one has

$$(\psi, \vec{p}\,\psi) = \int d^3k\, |a(\vec{k}, t)|^2\, \hbar\vec{k}\,. \tag{36}$$

Probabilistic meaning of $|a(\vec{k}, t)|^2$. Equations (34) and (36) show that the role of Fourier-transform factor $a(\vec{k}, t)$ in $\psi(\vec{x}, t)$ is to weight the momentum eigenstate $\phi_k(\vec{x})$ corresponding to momentum $\vec{p} = \hbar\,\vec{k}$, as confirmed by the Parseval equation

$$\int d^3x\, |\psi(\vec{x}, t)|^2 = \int d^3k\, |a(\vec{k}, t)|^2 = 1\,.$$

For the special class of free-particle solutions (35), formula (36) assumes the simple form

$$(\psi, \vec{p}\,\psi) = \int d^3k\, |a(\vec{k}, t)|^2\, \hbar\vec{k} = \int d^3k\, |a(\vec{k})|^2\, \hbar\vec{k}\,.$$

Similar to $\rho(\vec{x},t) = |\psi(\vec{x},t)|^2$ describing the probability to find a particle at some position $\vec{x}$, the term $|a(\vec{k},t)|^2$ represents **the probability that the system has the momentum corresponding to eigenstate** $\phi_k(\vec{x})$ in the 3-dimensional momentum space. In other words, $|a(\vec{k},t)|^2$ describes the **distribution of momentum** $\hbar\vec{k}$ characterizing a given state $\psi(\vec{x},t)$.

Formula (36) quite naturally confirms such an interpretation. It evidences the fact that the expectation value of $\vec{p}$ is the average value of vector $\hbar\vec{k}$ relevant to distribution $|a(\vec{k},t)|^2$ in the 3-dimensional momentum space $\{\,\vec{p} = \hbar\vec{k} \in \mathbb{R}^3\}$. If $\psi(\vec{x},t)$ is a free-particle solution (35) such a distribution is time independent being $|a(\vec{k},t)|^2 \equiv |a(\vec{k})|^2$.

Expectation value of the kinetic-energy operator $K = \hat{p}^2/(2m) = -(\hbar^2/2m)\,\nabla^2$

$$(\psi, K\,\psi) = \int d^3x\,\psi^*(\vec{x},t)\,\left(-\frac{\hbar^2}{2m}\nabla^2\right)\,\psi(\vec{x},t) = \frac{-\hbar^2}{2m}\int d^3x\,\psi^*(\vec{x},t)\,\nabla^2\,\int d^3k\,\frac{a(\vec{k},t)}{(2\pi)^{3/2}}\,e^{i\,\vec{k}\cdot\vec{x}}$$

$$= \frac{-\hbar^2}{2m}\int d^3x\,\psi^*(\vec{x},t)\int d^3k\,\frac{a(\vec{k},t)}{(2\pi)^{3/2}}\,(-k^2)e^{i\,\vec{k}\cdot\vec{x}} = \frac{-\hbar^2}{2m}\int d^3x\,\psi^*(\vec{x},t)\int d^3k\,a(\vec{k},t)\,(-k^2)\psi_k(\vec{x})$$

$$= \frac{-\hbar^2}{2m}\int d^3k\,a(\vec{k},t)\,(-k^2)\int d^3x\,\psi^*(\vec{x},t)\psi_k(\vec{x}) = \frac{-\hbar^2}{2m}\int d^3k\,a(\vec{k},t)\,(-k^2)\,(\psi_k,\psi(t))^*$$

Since $(\psi_k,\psi(t)) = a(\vec{k},t)$

$$(\psi, K\,\psi) = \frac{-\hbar^2}{2m}\int d^3k\,a(\vec{k},t)\,(-k^2)\,a^*(\vec{k},t) = \int d^3k\,\left(\frac{\hbar^2\,k^2}{2m}\right)\,|a(\vec{k},t)|^2\,,$$

showing that the expectation value $\langle K\rangle = (\psi, K\psi)$ is the average of the (kinetic) energy eigenvalue $E(k) = \hbar^2\,k^2/(2m)$ weighted by probabilities $|a(\vec{k},t)|^2$. Once more, if ψ represents the general solution of the free-particle Schrödinger problem, then $|a(\vec{k},t)|^2 = |a(\vec{k})|^2$.

Comment. The calculations leading to the previous expectation values $(\psi,\vec{p}\,\psi)$ and $(\psi, K\,\psi)$ can be performed in an alternative way by exploiting the properties of the plane-wave basis. Consider, for example,

$$(\psi,\vec{p}\,\psi) = \int d^3x\,\psi^*(\vec{x},t)\,(-i\hbar\nabla)\,\psi(\vec{x},t) = \int d^3x\,\psi^*(\vec{x},t)\,(-i\hbar\nabla)\,\int d^3k\,a(\vec{k},t)\,\phi_k(\vec{x})$$

$$= \int d^3x\,\psi^*(\vec{x},t)\int d^3k\,\hbar\vec{k}\,a(\vec{k},t)\,\phi_k(\vec{x}) = \int d^3k\,\hbar\vec{k}\,a(\vec{k},t)\,(\psi,\phi_k)$$

$$= \int d^3k\;\hbar\vec{k}\;a(\vec{k},t)\,(\phi_k,\psi)^* = \int d^3k\;\hbar\vec{k}\;|a(\vec{k},t)|^2\,, \qquad a^*(\vec{k},t) = (\phi_k,\psi)^*\,.$$

Expectation value of energy operator H. If $\psi_{E_n}(\vec{x})$ represents the energy eigenfunction such that $H\psi_{E_n}(\vec{x}) = E_n\psi(\vec{x})$, the expectation value of the Hamiltonian H associated to the **most general solution** of the Schrödinger problem can be found by resorting to equation (20). One finds the (quantum) **average energy**

$$E = (\psi, H\psi) = \int_{\mathbb{R}^3} d^3x\ \psi^*(\vec{x},t)H\psi(\vec{x},t)$$

$$= \sum_m \sum_n C^*(E_m)C(E_n)\ e^{i(E_m-E_n)\ t/\hbar} \int_{\mathbb{R}^3} d^3x\ \psi^*_{E_m}(\vec{x})H\psi_{E_n}(\vec{x})$$

$$= \sum_m \sum_n C^*(E_m)C(E_n)\ e^{i(E_m-E_n)\ t/\hbar} E_n\Big(\psi^*_{E_m}, \psi_{E_n}\Big) = \sum_n |C(E_n)|^2\ E_n$$

where the orthogonality condition $(\psi_{E_m}, \psi_{E_n}) = \delta_{m,n}$ has been utilized. The use of the latter is justified by the fact that the eigenfunction set $\{\psi_{E_n}(\vec{x}), n \in [0,\infty]\}$ can be shown to be **complete** (namely, it forms a basis). This property is discussed in the next section. The important point is that, once more, the final form of $E = (\psi, H\psi)$ exhibits the expected meanigful form describing the average of some quantity. In the current case, it combines all the possible energy values, each one weighted by $|C(E_n)|^2$ representing the probability to find in ψ the component ψ_{E_n} and thus the corresponding energy E_n. Interestingly, the resulting averge energy is **time independent**, a property that mimicks the time independence of classical Hamiltonians (and attests the validity of the correspondence principle).

6 Hermitian Operators and Eigenvalue Equations

This class of operators plays a fundamental role in Quantum Mechanics since any physical observable must be represented by a hermitian operator. The property of being hermitian is crucial in that it entails that the expectation value of an (hermitian) operator must be **real**. This circumstance, in turn, is essential because, in view of the operator character of physical observable within Quantum Mechanics, the only measurable quantities are the expectation values of an operator. Since the results of any experimental measurement is necessarily expressed in terms of real numbers the fact that physical operators are hermitian appears to be indispensable. It is thus particularly important to understand the properties of hermitian operators.

Definition of the **adjoint** operator A^+ of a given operator A

$$(\psi_2, A^+\,\psi_1) = (\psi_1, A\,\psi_2)^* \quad \Rightarrow \quad \int d^3x\,\psi_2^*\,A^+\,\psi_1 = \left(\int d^3x\,\psi_1^*\,A\,\psi_2\right)^*$$

Since $(\psi, \phi) = (\phi, \psi)^*$, then an equivalent definition is

$$\Big((\psi_2, A^+\,\psi_1)\Big)^* = \Big((\psi_1, A\,\psi_2)^*\Big)^* \quad \Rightarrow \quad (A^+\psi_1, \psi_2) = (\psi_1, A\,\psi_2)$$

namely

$$\int d^3x\,(A^+\,\psi_1)^*\psi_2 = \int d^3x\,\psi_1^*\,A\,\psi_2$$

Definition: an operator is **hermitian** if $A = A^+$. In this case

$$(\psi_2, A\,\psi_1) = (\psi_1, A\,\psi_2)^* \quad \Rightarrow \quad \int d^3x\,\psi_2^*\,A\,\psi_1 = \left(\int d^3x\,\psi_1^*\,A\,\psi_2\right)^*$$

The following, well-known, physical operators are hermitian

$$x_i\,, \quad f(x_1, x_2, x_3)\,, \quad \hat{p}_i = -i\hbar\frac{\partial}{\partial x_i}\,, \quad \hat{p}^2 = \hat{p}_1^2 + \hat{p}_2^2 + \hat{p}_3^2\,, \quad \hat{L}_j = \hat{x} \wedge \hat{p} = \epsilon_{jnm} x_n \hat{p}_m$$

It is interesting to check the validity of this statement in some cases

$$(\psi_2, x_i\,\psi_1) = \int d^3x\,\psi_2^*\,x_i\,\psi_1 = \left(\int d^3x\,\psi_2\,x_i^*\,\psi_1^*\right)^* = \left(\int d^3x\,\psi_1^*\,x_i\,\psi_2\right)^* = (\psi_1, x_i\,\psi_2)^*$$

$$(\psi_2, f\psi_1) = \int d^3x\,\psi_2^*\,f(\vec{x})\,\psi_1 = \left(\int d^3x\,\psi_2\,f^*(\vec{x})\,\psi_1^*\right)^* = \left(\int d^3x\,\psi_1^*\,f(\vec{x})\,\psi_2\right)^* = (\psi_1, f\psi_2)^*$$

Note that the latter formula has already been proved when we showed that $H = H^+$ (see equation (26)). In that case, f was represented by the potential energy $U(\vec{x})$.

$$(\psi_2, \hat{p}_3\,\psi_1) = \int d^3x\,\psi_2^*\,(-i\hbar\partial_{x_3})\,\psi_1 = -i\hbar\int d^3x\,\psi_2^*\,\partial_{x_3}\psi_1 = -i\hbar\int dx^3\,(\partial_{x_3}(\psi_2^*\psi_1) - \psi_1\partial_{x_3}\psi_2^*)$$

$$= -i\hbar\int_{\mathbb{R}} dx_1\int_{\mathbb{R}} dx_2\int_{\mathbb{R}} dx_3\left(\partial_{x_3}(\psi_2^*\psi_1) - \psi_1\partial_{x_3}\psi_2^*\right)$$

$$= -i\hbar\int_{\mathbb{R}} dx_1\int_{\mathbb{R}} dx_2\left(\psi_2^*\psi_1\right)_{-\infty}^{\infty} + i\hbar\int dx^3\,\psi_1\partial_{x_3}\psi_2^* = i\hbar\int dx^3\,\psi_1\partial_{x_3}\psi_2^*$$

$$= \left(\int dx^3\,\psi_1^*(-i\hbar\partial_{x_3})\psi_2\right)^* = (\psi_1, \hat{p}_3\,\psi_2)^*$$

where the fact that $\psi_1(\vec{x})$, $\psi_2(\vec{x}) \to 0$ for $|\vec{x}| \to \infty$ (and thus when $x_3 \to \infty$) has been used to eliminate $(\psi_2^*\psi_1)_{-\infty}^{\infty}$. Similar calculations, involving repeated integrations by part, allows one to show the hermitian character of angular-momentum components $\hat{L}_j$, of Laplacian ∇^2 and thus of the kinetic energy.

Eigenvalue equation. In the preceding section we have encounterd two eigenvalue equations relevant to operator $\vec{p}$ and its components $\vec{p}_j$ (see equation (33))

$$\hat{p}\,\phi_k(\vec{x}) = \hbar\vec{k}\,\phi_k(\vec{x})\,, \qquad \hat{p}_j\,\phi_k(\vec{x}) = \hbar k_j\,\phi_k(\vec{x})\,,$$

where the eigenfuntions $\phi_k(\vec{x})$ corresponds to vectors of the plane-wave basis. Another significant example is the time-independent Schrödinger equation (19) corresponding to the eigenvalue equation $H\,\psi(\vec{x}) = E\,\psi(\vec{x})$ for the energy opertor H.

In general, for any operator A it is possible to write down its eigenvalue equation

$$A\,\psi_m(\vec{x}) = \alpha_m\,\psi_m(\vec{x})\,, \qquad \alpha_m = \textbf{eigenvalue}$$

where the subscript m of α_m implies that the eigenfunction $\psi_m(\vec{x})$ is associated to a specific eigenvalue (see, e. g., the case of momentum operators). Eigenvalues α_m can be either discrete –an integer label enumerates them, see e. g. formula (30) for the momentum-opertor components of a particle in a box– or continuous, as for the momentum-opertor components of a particle in the 3D space. From now on we assume that operator A is hermitian.

Theorem 1. The eigenvalues of hermitian operators are real. Proof:

$$A\,\psi_m = \alpha_m\,\psi_m \;\;\Rightarrow\;\; (\psi_m, A\,\psi_m) = (\psi_m, \alpha_m\,\psi_m) = \alpha_m\,(\psi_m, \psi_m)$$

$$(\psi_m, A\,\psi_m)^* = \alpha_m^*\,(\psi_m, \psi_m)^*$$

On the other hand, if A is hermitian $(\Phi_2, A\,\Phi_1) \equiv (\Phi_2, A^+\Phi_1) = (\Phi_1, A\,\Phi_2)^*$ for any pair of functions Φ_1, $\Phi_2 \in L^2(V)$. Hence $(\psi_m, A\,\psi_m)^* \equiv (\psi_m, A\,\psi_m)$.

Then $(\psi_m, A\,\psi_m)^* \equiv (\psi_m, A\,\psi_m) = \alpha_m^*\,(\psi_m, \psi_m)$, implying that

$$0 = (\psi_m, A\,\psi_m) - (\psi_m, A\,\psi_m)^* = (\alpha_m - \alpha_m^*)\,(\psi_m, \psi_m) \;\;\Rightarrow\;\; \alpha_m = \alpha_m^*\,.$$

Theorem 2. The eigenfunctions of a hermitian operator belonging to different eigenvalues are necessarily orthogonal. Proof:

$$A\,\psi_m = \alpha_m\,\psi_m\,, \quad A\,\psi_n = \alpha_n\,\psi_n \;\;\Rightarrow$$

$$(\psi_n, A\,\psi_m) = (\psi_n, \alpha_m\,\psi_m) = \alpha_m\,(\psi_n, \psi_m)\,, \quad (\psi_m, A\,\psi_n) = (\psi_m, \alpha_n\,\psi_n) = \alpha_n\,(\psi_m, \psi_n)$$

Once more, since A is hermitian $(\Phi_2, A\,\Phi_1) = (\Phi_1, A\,\Phi_2)^*$. Then $(\psi_n, A\,\psi_m) \equiv (\psi_m, A\,\psi_n)^*$

$$\Rightarrow\;\; \alpha_m\,(\psi_n, \psi_m) \equiv \Big(\alpha_n\,(\psi_m, \psi_n)\Big)^* \;\;\Rightarrow\;\; \alpha_m\,(\psi_n, \psi_m) \equiv \alpha_n^*\,(\psi_m, \psi_n)^*\,.$$

By observing that $\alpha_n^* = \alpha_n$ and $(\psi_m, \psi_n)^* = (\psi_n, \psi_m)$, one finds

$$(\alpha_m\, - \alpha_n)\,(\psi_n, \psi_m) \;=\; 0$$

entailing that, if $\alpha_m \neq \alpha_n$, eigenstates ψ_n and ψ_m **are necessarily orthogonal**. This result is extremely important. Suppose that the spectrum (the set of eigenvalues) of A is such that all the eigenvalues are different. In this case the spectrum is said to be **non degenerate**. Then the set of eigenvectors forms an **orthonormal system**.

Degenerate eigenvalues. If the **spectrum is degenerate** then there exists at least a group of $M > 1$ eigenstates $\{\psi_n : n \in [1, M]\}$ having the same eigenvalue. To simplify the notation we have assumed that these are the first M elements of the whole eigenstate set. In this case since $(\psi_n, \psi_m) = 0$ is no longer valid one has

$$(\psi_n, \psi_m) = C_{nm} \;\text{ with }\quad C_{nm} = (\psi_n, \psi_m) = (\psi_m, \psi_n)^* = C_{mn}^*\,,$$

where the latter equation follows from the properties of scalar product of $L^2(V)$. Within Matrix Theory, a matrix $C = ||C_{nm}||$ with this property corresponds to a Hermitian matrix $C = C^+$ where $C^+ = ||(C^+)_{nm}|| = ||C_{mn}^*||$.
Comment. One easily verifies that such a definition perfectly matches the one used for physical operators A. Consider the basis $\{|n\rangle : n \in [1, M]\}$ such that $\langle n|C|i\rangle = C_{ni}$. Condition $C^+ \equiv C$ entails that

$$\langle n|C^+|i\rangle \equiv \langle n|C|i\rangle \;\;\Rightarrow\;\; C_{ni}^+ \equiv C_{ni} \;\;\Rightarrow\;\; C_{in}^* = (C_{in})^* = C_{ni} \;\;\Rightarrow\;\; \big(\langle i|C|n\rangle\big)^* = \langle n|C|i\rangle$$

Note that, the latter definition essentially coincides with that defining Hermitian operators $(\phi, A\psi) = (\psi, A\phi)^*$.

The important point is that, based on the theorems of Linear Algebra, it is possible to reduce a Hermitian matrix C to its **diagonal form** Λ by means of some **unitary matrix** U

$$U^+CU = \Lambda \;\Rightarrow\quad (U^+)_{in}C_{nm}U_{mj} = \lambda_i\delta_{ij}$$

(we use the convention that a repeated index implies summation). Unitary matrices are defined by the property $U^+ = U^{-1}$ and thus satisfy the identity $U^+U = U\,U^+ = \mathbb{I}$.

The **new set of eigenfunctions** $\Phi_n = U_{jn}\psi_j$, representing a linear combination of the initial eigenfunctions ψ_j, form an **orthonormal set**. To see this property one must rewrite $(U^+)_{in}C_{nm}U_{mj}$ as follows

$$(U^+)_{in}C_{nm}U_{mj} = U^*_{ni}\,(\psi_n,\psi_m)\,U_{mj} = U_{mj}\,U^*_{ni}\,(\psi_n,\psi_m) = U_{mj}\left(U_{ni}\psi_n,\psi_m\right)$$

$$= \left(U_{ni}\psi_n, U_{mj}\psi_m\right) = (\Phi_i,\Phi_j)$$

Then

$$(U^+)_{in}C_{nm}U_{mj} = \lambda_i\delta_{ij} \;\Rightarrow\quad (\Phi_i,\Phi_j) = \lambda_i\delta_{ij}$$

where the factor λ_i can be absorbed in Φ_i and Φ_j. The set $\{\psi_n : n \in [1,M]\}$ will be replaced by the new orthogonal set $\{\Phi_n : n \in [1,M]\}$.

Concluding, given a Hermitian operator A, independently from the fact that its spectrum is degenerate or not, the set of eigenvectors can be reduced to an **orthonormal system**.

Property. In general, the set of linearly independent eigenfunctions associated to a hermitian operator can be shown to generate a **basis** in a space $L^2(V)$ (for a suitable choice of V), namely, it represents a **system of orthonormal functions which is complete**. This result (the so-called Spectral Theorem of functional analysis) is the generalization of the theorem showing that the eigenvectors of a finite-dimensional hermitian matrix form a basis.

This property entails that, given a Hermitian operator A acting on the functions of $L^2(V)$, its eigenfunctions can be used to represent any function of $L^2(V)$. Assume that the spectrum of A is **discrete**

$$A\,\psi_n = \alpha_n\,\psi_n \;\Rightarrow\quad \psi = \sum_n f_n\,\psi_n\,,\;\; \forall\,\psi \in L^2(V) \quad \text{with}\;\; f_n = (\psi_n,\psi)\,.$$

Moreover

$$(\psi,\psi) \;=\; \sum_n |f_n|^2 = 1$$

where $|f_n|^2$ **represents the probability that the system is in the state** ψ_n and thus the probability that effecting a measurement of A the result found is α_n. The expectation value embodies this information

$$\langle A\rangle = (\psi, A\psi) = \sum_n f^*_n \sum_m f_m(\psi_n, A\psi_m) = \sum_n\sum_m f^*_n f_m\alpha_m\,(\psi_n,\psi_m) = \sum_n\sum_m f^*_n f_m\alpha_m\,\delta_{nm}$$

giving the **fundamental result**

$$\langle A \rangle = \sum_m \alpha_m \, |f_m|^2$$

showing how the expectation values, the quantum averges of physical operators, necessarily are real quantities. This is an important feature in that expectation values represent the quantities measured experimentally. Two examples have been discussed in the previous section, where the expectation values of momentum components $\hat{p}_j$ (and thus, more in general, the momentum operator $\hat{p}$) and of the kinetic-energy operator have been shown to fulfill this property (see formula (36)).

The **completeness relation** has the usual form

$$\delta^s(\vec{y} - \vec{x}) \, = \, \sum_n \, \psi_n^*(\vec{y}) \, \psi_n(\vec{x}) \, .$$

If Hermitian operator A acting on the functions of $L^2(V)$ features a **continuous** spectrum the previous formulas take the form

$$A \, \psi_\lambda = \, \lambda \, \psi_\lambda \;\; \Rightarrow \quad \psi = \int_D d\lambda f(\lambda) \, \psi_\lambda \, , \;\; \forall \, \psi \in L^2(V) \quad \text{with} \;\; f(\lambda) = (\psi_\lambda, \psi) \, .$$

Depending on the specific operator one deals with, eigenvalue $\lambda \in \mathbb{R}$ or on a subset D of $\mathbb{R}$. The next subsection supplies some examples. For example, it is shown that while for momentum components $-\infty \leq \lambda \leq \infty$, the kinetic-energy operator features $\lambda \geq 0$. The orthogonality of the basis and Parseval equation are described by the two formulas

$$(\psi_\eta, \, \psi_\lambda) \, = \, \delta \, (\, \eta - \, \lambda \,) \, , \qquad (\psi, \psi) \, = \, \int_D d\lambda \, |f(\lambda)|^2 = 1 \, ,$$

respectively, where $|f_\lambda|^2$ represents the probability relevant to state ψ_λ. The expectation value embodies this information

$$\langle A \rangle = (\psi, A\psi) = \int_D d\lambda f^*(\lambda) \int_D d\eta \, f(\eta) \, (\psi_\lambda, A\psi_\eta) = \int_D d\lambda \int_D d\eta \, f^*(\lambda) f(\eta) \, \eta \, (\psi_\lambda, \psi_\eta)$$

$$= \int_D d\lambda \int_D d\eta \, f^*(\lambda) f(\eta) \, \eta \, \delta(\lambda - \eta) = \int_D d\lambda \, \lambda \, |f(\lambda)|^2 \, .$$

The completeness relation has the form

$$\delta^s(\vec{y} - \vec{x}) \, = \, \int_D d\lambda \, \psi_\lambda^*(\vec{y}) \, \psi_\lambda(\vec{x}) \, .$$

6.1 Examples and comments

Example 1. Consider the momentum operator of a particle confined in 1-dimensional box $V = \{-L/2 < x < L/2\}$ with periodic boundary conditions. The eigenvalue equation for momentum operator $\hat{p}$ is

$$\hat{p} = -i\hbar\partial_x \;\Rightarrow\quad \hat{p}\,\phi_k(x) = \hbar\,k\,\phi_k(x)$$

is solved by

$$\phi_k(x) = \frac{e^{ik\,x}}{\sqrt{L}}\,, \qquad k\,\in\mathbb{R}\,.$$

Note that, apparently, $\hbar k$ is **continuous**. This is the result **if one ignores the periodic boundary conditions**. Conversely, the inclusion of this information amounts to imposing that $\phi_k(x+L) = \phi_k(x)$, namely,

$$\frac{e^{ik\,(x+L)}}{\sqrt{L}} = \frac{e^{ik\,x}}{\sqrt{L}} \quad\Leftrightarrow\quad k_n = \frac{2\pi}{L}\,n\,,\;\; n\in\mathbb{Z}$$

leading to a discretized set of eigenvalues. The final set of eigenstates exactly coincides with the plane-wave basis

$$\phi_n(\vec{x}) = \frac{e^{ik_n\,x}}{\sqrt{L}}\,, \qquad k_n = \frac{2\pi}{L}\,n\,,\;\; n\in\mathbb{Z}\,.$$

This example elucidates the crucial role of periodic boundary conditions.

Example 2: **degeneracy**. The same set of eigenstates (the plane waves) satisfies the eigenvalue equation for the energy operator $H = \hat{p}^2/(2m)$

$$H = \frac{\hat{p}^2}{2m} = -\frac{\hbar^2}{2m}\partial_x^2 \;\Rightarrow\quad \frac{\hat{p}^2}{2m}\,\phi_n(\vec{x}) = E_n\,\phi_n(\vec{x})\,,\quad E_n = \frac{\hbar^2 k_n^2}{2m}$$

The spectrum features an infinite **degeneracy**: for each n both $\phi_{+n}(\vec{x})$ and $\phi_{-n}(\vec{x})$ belong to eigenvalue E_n. Such two eigenstates, however, are orthogonal for each n, consistent with the fact that plane waves form a basis.

Example 3. In a 3-dimensional box momentum operator $\hat{p}$ has three components $\hat{p}_1$, $\hat{p}_2$, and $\hat{p}_3$. The corresponding three eigenvalue equations are easily solved

$$\hat{p}_j = -i\hbar\partial_{x_j}\,,\;\; j=1,\,2,\,3 \;\Rightarrow\quad \hat{p}_j\,\phi_{n_j}(x_j) = \hbar k_j\,\phi_{n_j}(x_j)$$

where

$$\phi_{n_j}(x_j) = \frac{e^{ik_j\,x_j}}{\sqrt{L}}\,, \qquad k_j = \frac{2\pi}{L}\,n_j\,,\;\; n_j\in\mathbb{Z}$$

If one considers a single momentum components, e. g. $\hat{p}_1$, eigenstates should be written as $\psi_{n_1}(x_1,x_2,x_3) = \phi_{n_1}(x_1)\,f(x_2,x_3)$ which satisfies the eigenvalue equation

$$\hat{p}_1\,\psi_{n_1} = \hat{p}_1\,\phi_{n_1}(x_1)\,f(x_2,x_3) = f(x_2,x_3)\,\hat{p}_1\,\phi_{n_1}(x_1) = \hbar k_1\,\psi_{n_1}\,f(x_2,x_3) = \hbar k_1\,\psi_{n_1}$$

In this case, the eigenvalue equation does not provide any information about $f(x_2, x_3)$ which remains totally undetermined. This situation can be seen as the case of **maximum degeneracy** due to the complete arbitrariness of function $f(x_2, x_3)$. **To remove the degeneracy** one needs supplementary information which, for example, can be achieved by requiring that the eigenvalue equations relevant to $\hat{p}_2$ and $\hat{p}_3$ are satisfied. By imposing such a condition one finds that

$$\psi(x_1, x_2, x_3) = \phi_{n_1}(x_1)\,\phi_{n_2}(x_2)\,\phi_{n_3}(x_3) = \frac{e^{ik_1\,x_1}}{\sqrt{L}}\,\frac{e^{ik_2\,x_2}}{\sqrt{L}}\,\frac{e^{ik_3\,x_3}}{\sqrt{L}}\,,\;\; k_j = \frac{2\pi}{L}\,n_j\,,\;\; n_j \in \mathbb{Z}$$

simultaneously diagonalize operators $\hat{p}_1$, $\hat{p}_2$ and $\hat{p}_3$. On the other hand, this is nothing but that the wave function of the plane-wave basis in a box which can be better expressed as

$$\phi_k(\vec{x}) = \frac{e^{i\vec{k}\cdot\vec{x}}}{\sqrt{L^3}}\,, \qquad \vec{k} = \frac{2\pi}{L}\,(n_1, n_2, n_3)\,,\;\; n_i \in \mathbb{Z}$$

Comment 1. The **special circumstance** that the momentum components $\hat{p}_j$ are commuting quantities $[\hat{p}_m, \hat{p}_n] = 0$ for each m and n has enabled us to solve the problem by diagonalizing each component separately. This property is described by a fundamental theorem discussed in the sequel.

Example 4. The previous problem is important when one must derive the eigenvalues and eigenvectors of the free-particle energy operator

$$H = \frac{\hat{p}_1^2 + \hat{p}_2^2 + \hat{p}_3^2}{2m} = -\frac{\hbar^2}{2m}\Big(\partial^2_{x_1} + \partial^2_{x_2} + \partial^2_{x_3}\Big) = -\frac{\hbar^2}{2m}\,\Delta$$

The request to diagonalize in parallel $\hat{p}_1$, $\hat{p}_2$, and $\hat{p}_3$ is implicitly contained in the fact that one must solve the eigenvalue equations for Δ. Clearly, the solution has the form $e^{i\vec{k}\cdot\vec{x}}$. As in example 1, the **periodic boundary conditions** will restrict the range of k_1, k_2 and k_3. The eigenvalue equation is satisfied by plane waves

$$H\,\phi_k(\vec{x}) \;=\; E_k\,\phi_k(\vec{x})\,, \qquad E_k = \frac{\hbar^2}{2m}(k_1^2 + k_2^2 + k_3^2)\,, \quad k_j = \frac{2\pi\,n_j}{L}\,,\; n_j \in \mathbb{Z}$$

This spectrum is **degenerate** in that the same value of $n_1^2 + n_2^2 + n_3^2$ can be realized by different triplets (n_1, n_2, n_3). Plane waves $\phi_k(\vec{x})$ form an orthonormal basis.

Comment 3. The general solution (20) of the Schrödinger problem can be reinterpreted by considering the properties of Hermitian operators. The energy operator is Hermitian. The eigenstates of the relevant eigenvalue equation

$$H\,\psi_n(\vec{x}) = E_n\psi_n(\vec{x})\,,\;\; n = 0,\, 1,\, 2,$$

are the **stationary states** of the Schrödinger problem $i\hbar\partial_t\psi = H\psi$. We have assumed to consider a discrete spectrum where index n labels the eigenvalues (this is case, e. g., when a

particle is trapped in a potential well) Eigenstates $\psi_n(\vec{x})$ form an **orthonormal basis**. The **general solution**

$$\psi(\vec{x},t) = \sum_n C(E_n) e^{-itE_n/\hbar} \psi_n(\vec{x})$$

not only provides the most general way of representing solutions to some Schrödinger problem, but represents the **characteristic expansion within the energy-eigenvalue basis** for **any time-dependent states**.

Comment 4. Even if the Theorems concerning Hermitian operators have implicitly assumed that such operators had a discrete spectrum, the results supplied by such theorems can be extended to **operators with a continuous spectrum**.

Example 4. Consider the momentum operator $\hat{p}$ and its three components $\hat{p}_1$, $\hat{p}_2$, and $\hat{p}_3$ for a particle in the space $\mathbb{R}^3$. The corresponding three eigenvalue equations are easily solved

$$\hat{p}_j = -i\hbar\partial_{x_j}\,,\ j = 1,\,2,\,3 \ \ \Rightarrow \quad \hat{p}_j\,\phi_{k_j}(x_j) = \hbar k_j\,\phi_{k_j}(x_j)$$

where the eigenvalues of

$$\phi_{k_j}(x_j) = \frac{e^{ik_j\,x_j}}{\sqrt{L}}\,, \qquad k_j \in \mathbb{R}$$

are continuous (necessarily real) quantities. In this case, **no boundary condition** is involved which restricts the permitted values of eigenvalues. The resulting basis is that of plane waves that, in spite of their diverging normalization integral, provides the fundamental representation of (normalizable) functions by means of the **inverse Fourier integral**. The **eigenvalue equation** for operator $\hat{p}^2$ and thus for the free-particle hamiltonian $H = \hat{p}^2/(2m)$ is solved as well

$$\hat{p}^2\,\phi_{\vec{k}}(\vec{x}) = \hbar^2(k_1^2 + k_2^2 + k_3^2)\,\phi_{\vec{k}}(\vec{x})\,, \qquad \phi_{\vec{k}}(\vec{x}) = \phi_{k_1}(x_1)\,\phi_{k_2}(x_2)\,\phi_{k_3}(x_3)\,.$$

In this case the eigenvalue degeneracy is due to any vector $\vec{k}$ satisfying $k_1^2 + k_2^2 + k_3^2 = \text{const}$.

6.2 The Axioms of Quantum Theory

I. The physical states of a system are described by wave functions $\Psi(\vec{x}) \in \mathbb{C}$.
For any state, the quantity $|\Psi(\vec{x})|^2$ is the probability density of finding a particle at time t at the position $\vec{x}$.

II. The physical observables are represented by Hermitian operators.

III. The expectation value of the physical observable rapresented by operator A is given by

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

IV. Time evolution of physical states $\Psi(\vec{x},t)$ is governed by the Schrödinger equation

V. A measurement of operator A selecting the value λ (within the eigenvalue set of the relevant spectrum) entails that the wavefunction of the system changes to the corresponding eigenfunction Ψ_λ.

7 The Harmonic oscillator

This system represents a paradigmatic example within the Quantum Theory. It allows to demonstrate the properties of a quantum system in a particularly transparent way.

Hamiltonian and eigenvalue equation

$$H = \frac{p^2}{2m} + \frac{m\omega^2}{2}\,x^2 \quad \Rightarrow \quad \left(\frac{-\hbar^2}{2m}\frac{d^2}{dx^2} + \frac{m\omega^2}{2}\,x^2\right)\psi(x) \,=\, E\,\psi(x) \tag{37}$$

The spectrum-generating-algebra method. This method allows one to perform the diagonalization of H by exploiting the algebraic properties of the so-called lowering and raising operators (or destruction/construction operators)

$$a = \sqrt{\frac{m\omega}{2\hbar}}\left(x + \frac{i\,p}{m\omega}\right), \qquad a^+ = \sqrt{\frac{m\omega}{2\hbar}}\left(x - \frac{i\,p}{m\omega}\right) \tag{38}$$

also available in the form

$$a = \frac{1}{\sqrt{2}}\left(\frac{x}{\lambda} + \frac{i\,\lambda\,p}{\hbar}\right), \quad a^+ = \frac{1}{\sqrt{2}}\left(\frac{x}{\lambda} - \frac{i\,\lambda\,p}{\hbar}\right), \qquad \lambda = \sqrt{\frac{\hbar}{m\omega}}$$

Note that

$$[\,a\,,\,a^+\,] = \left[\sqrt{\frac{m\omega}{2\hbar}}\left(x + \frac{i\,p}{m\omega}\right), \sqrt{\frac{m\omega}{2\hbar}}\left(x - \frac{i\,p}{m\omega}\right)\right] = \frac{m\omega}{2\hbar}\left(\left[x, \frac{-i\,p}{m\omega}\right] + \left[\frac{i\,p}{m\omega}, x\right]\right)$$

$$= -2i\,\frac{m\omega}{2\hbar}\,\frac{1}{m\omega}\,[\,x,\,p\,] = -2i\,\frac{m\omega}{2\hbar}\,\frac{1}{m\omega}\,i\hbar = 1 \quad \Rightarrow \qquad [\,a\,,\,a^+\,] = 1\,.$$

By inverting the previous formulas one has

$$x = \sqrt{\frac{\hbar}{2m\omega}}\,\left(a \,+\, a^+\right), \qquad p = \,-i\sqrt{\frac{\hbar m\omega}{2}}\,\left(a \,-\, a^+\right).$$

Hamiltonian

$$H = \frac{p^2}{2m} + \frac{m\omega^2}{2}\,x^2 = -\frac{1}{2m}\,\frac{\hbar m\omega}{2}\,\left(a \,-\, a^+\right)^2 + \frac{m\omega^2}{2}\,\frac{\hbar}{2m\omega}\,\left(a \,+\, a^+\right)^2$$

$$= -\,\frac{\hbar\omega}{4}\,\left(a^2 \,+ (a^+)^2 \,- a^+a \,- a\,a^+\right) + \frac{\hbar\omega}{4}\,\left(a^2 \,+ (a^+)^2 \,+ a^+a \,+ a\,a^+\right)$$

$$= -\,\frac{\hbar\omega}{4}\,\left(-a^+a \,- a\,a^+\right) + \frac{\hbar\omega}{4}\,\left(a^+a \,+ a\,a^+\right) = \frac{\hbar\omega}{2}\,\left(a^+a \,+ a\,a^+\right) = \frac{\hbar\omega}{2}\,\left(2a^+a \,+ 1\right)$$

The identity $a\,a^+ = 1 + a^+a$ owing to $[a, a^+] = 1$ has been used. Then

$$H = \hbar\omega \left(a^+a + \frac{1}{2} \right)$$

The problem of diagonalizing H is thus equivalent to diagonalizing operator $\hat{n} = a^+a$ called **number operator**. To this end, assume that the spectrum of $\hat{n}$ is known

$$\hat{n}\,\psi_n = n\,\psi_n$$

where n is a non negative real number in that

$$(\psi_n,\, \hat{n}\,\psi_n) = (\psi_n,\, a^+a\,\psi_n) = (\psi_n,\, a^+\,(a\,\psi_n)) = (a\,\psi_n,\, a\,\psi_n) = (\phi_n, \phi_n) \geq 0\,.$$

Since $(\psi_n,\, \hat{n}\,\psi_n) = (\psi_n,\, n\,\psi_n) = n\,(\psi_n,\, \psi_n) = n$ then n is a real, non negative number.

Operators 1, a, a^+ and $\hat{n}$ are the **generators** of the so-called **Weyl-Heisenberg algebra**. The defining commutators are

$$[a, a^+] = 1\,, \quad [a, \hat{n}] = a\,, \quad [a^+, \hat{n}] = -a^+\,, \tag{39}$$

in addition to the trivial ones $[1, a] = [1, a^+] = [1, \hat{n}] = 0$. The algebra is the set of elements formed by any (Hermitian) linear combination $\xi = ca^+ + c^*a + r\hat{n} + s\,1$ of the generators where $c \in \mathbb{C}$ and $r,\, s \in \mathbb{R}$. The distinctive feature of a set of elements equipped with the algebra structure is that any pair of elements ξ and ξ' has a commutator $[\,\xi, \xi'\,]$ which still represents an element of the algebra.

Lemma 1. Wave function $a^+\psi_n$ is an eigenfunction of $\hat{n}$ with eigenvalue $n + 1$

$$\hat{n}\,a^+\psi_n = \Big(\hat{n}\,a^+ - a^+\hat{n} + a^+\hat{n}\Big)\,\psi_n = \Big([\hat{n}\,a^+] + a^+\hat{n}\Big)\,\psi_n$$

$$= \Big(a^+ + a^+\hat{n}\Big)\,\psi_n = a^+\Big(1 + \hat{n}\Big)\,\psi_n = a^+\big(1 + n\big)\,\psi_n = (1 + n)\,a^+\psi_n$$

Then

$$a^+\psi_n = C_n\psi_{n+1} \quad \Rightarrow \quad \psi_{n+1} = \frac{1}{C_n}\,a^+\psi_n\,, \qquad C_n \in \mathbb{C}\,.$$

Normalization

$$1 = (\psi_{n+1},\, \psi_{n+1}) = \left(\frac{1}{C_n}\,a^+\psi_n,\, \frac{1}{C_n}\,a^+\psi_n\right) = \frac{1}{|C_n|^2}\,\big(\,a^+\psi_n,\, a^+\psi_n\big) = \frac{1}{|C_n|^2}\,\big(\,\psi_n,\, a\,a^+\psi_n\big)$$

$$= \frac{1}{|C_n|^2}\,(\,\psi_n,\, (\hat{n} + 1)\,\psi_n) = \frac{1}{|C_n|^2}\,(\,\psi_n,\, (n + 1)\,\psi_n) = \frac{n+1}{|C_n|^2}\,(\,\psi_n,\, \psi_n) = \frac{n+1}{|C_n|^2}$$

Then

$$a^+\psi_n \,=\, \sqrt{n+1}\psi_{n+1} \quad \Leftrightarrow \quad \psi_{n+1} = \frac{1}{\sqrt{n+1}}\, a^+\psi_n \tag{40}$$

Lemma 2. Wave function $a\,\psi_n$ is an eigenfunction of $\hat{n}$ with eiegnvalue $n-1$

$$\hat{n}\, a\,\psi_n = \Big(\hat{n}\, a - a\,\hat{n}\, + a\,\hat{n}\,\Big)\,\psi_n = \Big([\hat{n}\, a\,] + a\,\hat{n}\,\Big)\,\psi_n$$

$$= \Big(-a\, + a\,\hat{n}\,\Big)\,\psi_n = a\,\Big(-1+\hat{n}\,\Big)\,\psi_n = a\,\big(\,n-1\,\big)\,\psi_n = \,(\,n-1\,)\,a\,\psi_n$$

Then

$$a\,\psi_n \,=\, D_n\psi_{n-1} \quad \Rightarrow \qquad \psi_{n-1} = \frac{1}{D_n}\, a\,\psi_n \qquad D_n \in \mathbb{C}\,.$$

Normalization

$$1 = (\psi_{n-1},\, \psi_{n-1}) = \left(\frac{1}{D_n}\, a\,\psi_n,\, \frac{1}{D_n}\, a\,\psi_n\right) = \frac{1}{|D_n|^2}\,(\,a\,\psi_n,\, a\,\psi_n) = \frac{1}{|D_n|^2}\,\big(\,\psi_n,\, a^+ a\,\psi_n\big)$$

$$= \frac{1}{|D_n|^2}\,(\,\psi_n,\, \hat{n}\,\psi_n) = \frac{1}{|D_n|^2}\,(\,\psi_n,\, n\,\psi_n) = \frac{n}{|D_n|^2}\,(\,\psi_n,\, \psi_n) = \frac{n}{|D_n|^2}$$

Then

$$a\,\psi_n \,=\, \sqrt{n}\,\psi_{n-1} \quad \Leftrightarrow \quad \psi_{n-1} = \frac{1}{\sqrt{n}}\, a\,\psi_n \tag{41}$$

Equation $1 = n/|D_n|^2$ entails that $n \geq 0$. Then the lowest value of n is zero. Then, starting from the **ground state** ψ_0 such that $a\,\psi_0 = 0$, one can reconstruct the whole hierarchy of eigenstates. To this end formula (40) plays a crucial role

$$\psi_n = \frac{1}{\sqrt{n}}\, a^+\psi_{n-1} = \frac{1}{\sqrt{n}}\, a^+\left(\frac{1}{\sqrt{n-1}}\, a^+\psi_{n-2}\right) = \frac{1}{\sqrt{n(n-1)}}\,(a^+)^2\psi_{n-2}$$

$$= \,...\, = \frac{1}{\sqrt{n(n-1)...(n-s+1)}}\,(a^+)^s\psi_{n-s} = \frac{1}{\sqrt{n!}}\,(a^+)^n\psi_0$$

The resulting formula

$$\psi_n = \frac{1}{\sqrt{n!}}\,(a^+)^n\psi_0 \tag{42}$$

is particularly important because, if ψ_0 is known, one can reconstruct the entire set of eigenfunctions. In this sense **the algebra is generating the spectrum**.

7.1 Eigenvalues and eigenstates

The HO ground state. Wave function ψ_0 must solve

$$a\,\psi_0 = 0 \;\;\Rightarrow\;\; \frac{1}{\sqrt{2}}\left(\frac{x}{\lambda} + \frac{i\,\lambda\,p}{\hbar}\right)\psi_0 = 0 \;\;\Rightarrow\;\; \left(\frac{x}{\lambda} + \lambda\,\frac{d}{dx}\right)\psi_0 = 0$$

whose solution is almost immediate

$$\psi_0(x) = C\,e^{-x^2/(2\lambda^2)}\,, \qquad C = (\lambda\sqrt{\pi})^{-1/2}$$

Thanks to formula (42)

$$\begin{aligned}\psi_n = \frac{1}{\sqrt{n!}}\,(a^+)^n\psi_0 &= \frac{C}{\sqrt{n!}}\left(\frac{1}{\sqrt{2}}\left(\frac{x}{\lambda} - \frac{i\,\lambda\,p}{\hbar}\right)\right)^n e^{-x^2/(2\lambda^2)} \\ &= \frac{C}{\sqrt{2^n\,n!}}\left(\frac{x}{\lambda} - \lambda\,\frac{d}{dx}\right)^n e^{-x^2/(2\lambda^2)} \end{aligned} \tag{43}$$

Definition. The Hermite polynomials are defined by

$$H_n(y) = (-1)^n\,e^{y^2}\,\frac{d^n}{dy^n}\,e^{-y^2}$$

$$H_0 = 1\,, \quad H_1 = 2y\,, \quad H_2 = 4y^2 - 2\,, \quad H_3 = 8y^3 - 12y\,, \quad H_4 = 16y^4 - 48y^2 + 12$$

An equivalent definition is based on the differential equation

$$\left[\,\frac{d^2}{dy^2} - 2y\,\frac{d}{dy} + 2n\,\right]\,H_n(y) = 0$$

Lemma

$$\psi_n = \frac{1}{\sqrt{n!}}\,(a^+)^n\psi_0 = \frac{C}{\sqrt{2^n\,n!}}\;e^{-x^2/(2\lambda^2)}\,H_n(x/\lambda) \tag{44}$$

where $H_n(y)$ is the n-th Hermite polynomial. **Proof**: owing to formula (43)

$$\psi_n = \frac{C}{\sqrt{2^n\,n!}}\left(\frac{x}{\lambda} - \lambda\,\frac{d}{dx}\right)^n e^{-x^2/(2\lambda^2)} = \frac{C}{\sqrt{2^n\,n!}}\left(y - \frac{d}{dy}\right)^n e^{-y^2/2}$$

where $x = \lambda\,y$. Consider the following factorization

$$\left(y - \frac{d}{dy}\right)^n e^{-y^2/2} = \left(y - \frac{d}{dy}\right)\left(y - \frac{d}{dy}\right)\,...\,\left(y - \frac{d}{dy}\right)\,e^{+y^2/2}\,e^{-y^2}$$

$$= e^{+y^2/2}\,e^{-y^2/2}\left(y - \frac{d}{dy}\right)\,e^{+y^2/2}\,e^{-y^2/2}\left(y - \frac{d}{dy}\right)\,e^{+y^2/2}\,...\,e^{-y^2/2}\left(y - \frac{d}{dy}\right)\,e^{+y^2/2}e^{-y^2}$$

$$e^{-y^2/2}\left(y-\frac{d}{dy}\right)\,e^{+y^2/2}F(y)=e^{-y^2/2}\left(y\,e^{+y^2/2}F-\,y\,e^{+y^2/2}F-e^{+y^2/2}\frac{dF}{dy}\right)=(-1)\,\frac{dF}{dy}$$

Then

$$\psi_n=\frac{C}{\sqrt{2^n\,n!}}\left(y-\frac{d}{dy}\right)^n\,e^{-y^2/2}=\frac{C}{\sqrt{2^n\,n!}}\,e^{+y^2/2}(-1)^n\,\frac{d^n}{dy^n}\,e^{-y^2}$$

$$=\frac{C}{\sqrt{2^n\,n!}}\,e^{-y^2/2}\left((-1)^n e^{+y^2}\,\frac{d^n}{dy^n}\,e^{-y^2}\right)=\frac{C}{\sqrt{2^n\,n!}}\,e^{-y^2/2}\,H_n(y)\qquad QED$$

Energy spectrum. Energy eigenfunctions and the corresponding eigenvalues are given by

$$\psi_n=\frac{1}{\sqrt{2^n\lambda\,n!\sqrt{\pi}}}\,e^{-x^2/(2\lambda^2)}\,H_n(x/\lambda)\;\;\Leftrightarrow\;\;\;H\,\psi_n\,=E_n\,\psi_n\,,\quad E_n=\hbar\omega\,(n+1/2)\,.$$

Harmonic-oscillator basis. Eigenstates ψ_n of the Hermitian operator H form a complete system of orthonormal eigenstates

$$(\psi_m,\psi_n)=\int_{-\infty}^{\infty}dx\,\psi_m(x)\,\psi_n(x)\,=\int dx\,\frac{e^{-x^2/\lambda^2}}{\lambda\sqrt{2^n 2^m n!m!\pi}}\,H_m(x/\lambda)\,H_n(x/\lambda)$$

$$=\int dy\,\frac{e^{-y^2}}{\sqrt{2^n 2^m n!m!\pi}}\,H_m(y)\,H_n(y)=\frac{1}{\sqrt{2^n 2^m n!m!\pi}}\,\int dy\,e^{-y^2}\,H_m(y)\,H_n(y)=\delta_{mn}$$

The value of the latter integral is obtained by means of well-known calculations within the Theory of Orthogonal Polynomials. The **completeness** entails that

$$\sum_{n=0}^{\infty}\psi_n(y)\,\psi_n(y')\,=\,\delta(y-y')\,.$$

Proof of the orthogonality relation $(\psi_m,\psi_n)=\delta_{mn}$. Assume that $m<n$. In case $m>n$ one exchanges ψ_m with ψ_n being $(\psi_m,\psi_n)=(\psi_n,\psi_m)$. It is possible to show that

$$\left(e^{-x^2/2\lambda^2}x^s,\psi_n\right)=0\quad\text{for each}\;\;s<n\,.$$

If this statement is correct, by observing that ψ_m is a linear combination of powers x^s with $s\le m$, then

$$(\psi_m,\psi_n)=\left(\,e^{-x^2/2\lambda^2}\sum_s c_s x^s,\,\psi_n\,\right)=\sum_s c_s\,\left(\,e^{-x^2/2\lambda^2}x^s,\,\psi_n\,\right)=0$$

whenever $m < n$. In the case when $m > n$ one gets the same conclusion because ψ_m can be exchanged with ψ_n. Then, by substituting $\psi_n = N_n\, e^{-x^2/\lambda^2}\, H_n$ in the latter formula, one has

$$\Big(e^{-x^2/2\lambda^2} x^s, \psi_n\Big) = \int_{-\infty}^{\infty} dx\, e^{-x^2/\lambda^2}\, x^s\, N_n\, H_n(x/\lambda) = N_n\, \lambda^{s+1} \int_{-\infty}^{\infty} dy\, e^{-y^2}\, y^s\, H_n(y)$$

$$= N_n\, \lambda^{s+1} \int dy\, e^{-y^2}\, y^s \left((-1)^n e^{y^2} \frac{d^n}{dy^n} e^{-y^2} \right) = N_n\, \lambda^{s+1} (-1)^n \int dy\, E^{-1}\, y^s\, E D^n E^{-1}$$

$$= N_n\, \lambda^{s+1} (-1)^n \int dy\, y^s\, D^n E^{-1}$$

The latter integral can be shown to be zero by performing repeated integrations by part

$$\int dy\, y^s\, D^n E^{-1} = \int dy\, \big[\, D(y^s\, D^{n-1} E^{-1}) - s\, y^{s-1}\, D^{n-1} E^{-1} \,\big]$$

$$= \Big(y^s\, D^{n-1} E^{-1} \Big)_{-\infty}^{\infty} - s \int dy\, \big(\, y^{s-1}\, D^{n-1} E^{-1} \,\big) = -s \int dy\, y^{s-1}\, D^{n-1} E^{-1}$$

$$= -s \int dy\, \big[\, D(y^{s-1}\, D^{n-2} E^{-1}) - (s-1)\, y^{s-2}\, D^{n-2} E^{-1} \,\big]$$

$$= (-)^2\, s(s-1) \int dy\, \big[\, y^{s-2}\, D^{n-2} E^{-1} \,\big] = (-)^r\, s(s-1)\, ...(s-r+1) \int dy\, \big[\, y^{s-r}\, D^{n-r} E^{-1} \,\big]$$

$$= (-)^s\, s! \int dy\, \big[\, D^{n-s} E^{-1} \,\big] = (-)^s\, s! \, \big[\, D^{n-s-1} E^{-1} \,\big]_{-\infty}^{\infty} = 0 \qquad QED\,.$$

Probability density and expectation values.

Recall that

$$x = \sqrt{\frac{\hbar}{2m\omega}}\, \big(\, a \,+\, a^+ \,\big)\,, \qquad p = \,-i \sqrt{\frac{\hbar m\omega}{2}}\, \big(\, a \,-\, a^+ \,\big)\,.$$

$$\langle x \rangle = (\psi_n, x\, \psi_n) = \sqrt{\frac{\hbar}{2m\omega}}\, \big(\psi_n, (a \,+\, a^+)\, \psi_n\big) = \sqrt{\frac{\hbar}{2m\omega}}\, \Big(\big(\psi_n, a^+ \psi_n\big) + \big(\psi_n, a\, \psi_n\big) \Big) = 0\,.$$

$$\langle p \rangle = (\psi_n, p\, \psi_n) = -i \sqrt{\frac{\hbar m\omega}{2}}\, \big(\psi_n, (a - a^+)\, \psi_n\big) = -i \sqrt{\frac{\hbar m\omega}{2}} \Big(\big(\psi_n, a^+ \psi_n\big) - \big(\psi_n, a\, \psi_n\big) \Big) = 0\,.$$

Concerning $\langle x^2 \rangle$ and $\langle p^2 \rangle$ one finds

$$\langle x^2 \rangle = (\psi_n, x^2\, \psi_n) = \frac{\hbar}{2m\omega}\, \big(\psi_n, (a \,+\, a^+)^2\, \psi_n\big) = \frac{2n+1}{2}\, \lambda^2\,,$$

$$\langle p^2 \rangle = (\psi_n, p^2\, \psi_n) = -\frac{\hbar m\omega}{2} \big(\psi_n, (a \,-\, a^+)^2\, \psi_n\big) = \frac{2n+1}{2\,\lambda^2}\, \hbar^2$$

with $\lambda^2 = \hbar/(mw)$

$$\Delta x = \sqrt{\langle x^2\rangle - \langle x\rangle^2} = \sqrt{\langle x^2\rangle} = \sqrt{\frac{2n+1}{2}}\,\lambda\,,$$

$$\Delta p = \sqrt{\langle p^2\rangle - \langle p\rangle^2} = \sqrt{\langle p^2\rangle} = \sqrt{\frac{2n+1}{2}}\,\frac{\hbar}{\lambda}\,.$$

This reproduces the well-known Heisenberg uncertainty relation

$$\Delta x\,\Delta p \;=\; \sqrt{\frac{2n+1}{2}}\,\lambda\,\sqrt{\frac{2n+1}{2}}\,\frac{\hbar}{\lambda} \;=\; \frac{2n+1}{2}\,\hbar \;\geq\; \frac{1}{2}\,\hbar\,,$$

that will be discussed in the next section.

Probability density. From the classical formula $x(t) = A\,\sin(wt+\phi)$ describing the harmonic-oscillator motion, one can derive the classical probability of finding the particle in a given interval (x_1, x_2) in $[-A, +A]$

$$dx = Aw\,\cos(wt+\phi)\,dt \;\;\Rightarrow\;\; dx = \frac{2\pi}{T}\,\sqrt{A^2-x^2}dt \;\;\Rightarrow\;\; \frac{dx}{\pi\,\sqrt{A^2-x^2}} = \frac{dt}{T/2}$$

$$P(x_1,x_2) = \frac{t_2-t_1}{T/2} = \int_{x_1}^{x_2}\frac{dx}{\pi\,\sqrt{A^2-x^2}}$$

The latter can be shown to coincide for $n >> 1$ (semiclassical limit) with the corresponding integral of the probability density

$$\int_{x_1}^{x_2}|\psi_n(x)|^2dx \;\;\rightarrow\;\; P(x_1,x_2) = \int_{x_1}^{x_2}\frac{dx}{\pi\,\sqrt{A^2-x^2}}$$

Note that the Hermite polynomial $H_n(y)$ embodied in ψ_n features n **zeros distributed in the interval** $-A \leq x \leq +A$ with $A \simeq \lambda\sqrt{2n}$. For $|x| > A$ the harmonic-oscillator wavefunctions are such that $\psi_n(x) \approx x^n e^{-x^2/(2\lambda^2)}$ since $H_n(x/\lambda) \approx x^n$ and the exponential factor rapidly dominates x^n making ψ_n vanishingly small. The size of the zero interval $[-A, A]$ is confirmed by the comparison of the classical energy with the corresponding energy eigenvalue E_n for large n:

$$E = \frac{mw^2}{2}\,A^2 \;\Leftrightarrow\; E_n = \hbar w(n+1/2) \;\Rightarrow\; A^2 \simeq \frac{2}{mw^2}\,\hbar w(n+1/2) \;\Rightarrow\; A \simeq \sqrt{\frac{2n\,\hbar}{mw}} = \lambda\sqrt{2n}\,,$$

showing that the **region where oscillations of mass** m **take place** corresponds to the region where $|\psi_n|^2$ is non zero, namely, to the region containing the zeros of polynomial H_n. Classical inversion points $x = \pm A$ coincide with the values of x for which the Gaussian factor of $|\psi_n|^2$ causes a rapid decreasing of the **probability density**.

8 The Heisenberg Uncertainty Relation

Consider two Hermitian operators A_1 and A_2 with non vanishing commutator $[A_1, A_2] = iG$. The new operators involving the expectation values of A_1 and A_2

$$D_1 = A_1 - \langle A_1 \rangle\,, \;\; D_2 = A_2 - \langle A_2 \rangle\,, \qquad \langle A_r \rangle := (\psi, A_r \psi)\,,\; r = 1,\, 2$$

where $|\psi\rangle$ is a generic physical state, feature the commutator

$$\begin{aligned}[D_1, D_2] = [\, A_1 - \langle A_1 \rangle,\, A_2 - \langle A_2 \rangle\,] &= [A_1,\, A_2 - \langle A_2 \rangle\,] - [\langle A_1 \rangle,\, A_2 - \langle A_2 \rangle\,] \\ &= [A_1,\, A_2 - \langle A_2 \rangle\,] = [A_1,\, A_2]\,,\end{aligned}$$

since any operator commutes with quantities such as $\langle A_1 \rangle$ and $\langle A_2 \rangle$ (one should recall that the identity operator is implied $\langle A_r \rangle \equiv \langle A_r \rangle\, \mathbb{I}$).

Derivation of the Uncertainty Relation.

Step 1. To begin with, consider

$$|\langle G \rangle|^2 = |\langle [D_1, D_2] \rangle|^2 = |\, \langle D_1 D_2 \rangle - \langle D_2 D_1 \rangle\, |^2$$

Since $(\psi_1, A\psi_2)^* = (\psi_2, A^+ \psi_1)$ and, by definition, $D_r = D_r^+$

$$\langle D_1 D_2 \rangle^* = (\psi, (D_1 D_2)\psi)^* = (\psi, (D_1 D_2)^+ \psi) = (\psi, (D_2^+ D_1^+)\psi) = (\psi, D_2 D_1 \psi) = \langle D_2 D_1 \rangle$$

Then, by setting $\langle D_1 D_2 \rangle = \alpha + i\beta \in \mathbb{C}$, one has

$$|\langle G \rangle|^2 = |\, \langle D_1 D_2 \rangle - \langle D_1 D_2 \rangle^*\, |^2 = 4\beta^2 \leq\, 4|\alpha + i\beta|^2 \;\; \Rightarrow \;\; |\langle G \rangle|^2 \leq\, 4\, |\langle D_1 D_2 \rangle|^2$$

Step 2. The quantity $\langle D_1 D_2 \rangle$ can be rewritten as

$$\langle D_1 D_2 \rangle = (\psi, D_1 D_2 \psi) = (\psi, D_1 (D_2 \psi)) = (D_1^+ \psi, D_2 \psi) = (D_1 \psi, D_2 \psi)$$

where the adjoint-operator definition $(\psi, A\phi) = (A^+ \psi, \phi)$ and the fact that $A = D_1 = D_1^+$ have been used. Then

$$|\langle G \rangle|^2 \leq\, 4\, |\langle D_1 D_2 \rangle|^2 =\, 4|(D_1 \psi, D_2 \psi)|^2$$

Step 3. Thanks to the Schwarz inequality $|(\psi_1, \psi_2)|^2 \leq (\psi_1, \psi_1)\, (\psi_2, \psi_2)$, by setting $D_1 \psi = \psi_1$ and $D_2 \psi = \psi_2$, one has

$$\begin{aligned}|\langle G \rangle|^2 \leq\, 4|(D_1 \psi, D_2 \psi)|^2 \;\leq\; 4(\psi_1, \psi_1)\, (\psi_2, \psi_2) &= 4\, (D_1 \psi, D_1 \psi)(D_2 \psi, D_2 \psi) \\ &= 4\, (\psi, D_1^2 \psi)(\psi, D_2^2 \psi) = 4\, \langle D_1^2 \rangle \langle D_2^2 \rangle\end{aligned}$$

The quantities $\langle D_r^2 \rangle$ are recognized to be the root-mean-square deviations (denoted by ΔA_r) from the mean value $\langle A_r \rangle$

$$\langle D_r^2 \rangle = \langle (A_r - \langle A_r \rangle)^2 \rangle = \langle (A_r^2 \, - 2 A_r \, \langle A_r \rangle \, + \langle A_r \rangle^2) \rangle = \langle A_r^2 \rangle \, - 2 \langle A_r \rangle \, \langle A_r \rangle \, + \langle A_r \rangle^2$$

$$= \langle A_r^2 \rangle \, - \langle A_r \rangle^2 \;\; \Rightarrow \qquad (\Delta A_r)^2 = \langle A_r^2 \rangle \, - \langle A_r \rangle^2$$

The final formula is

$$|\langle G \rangle|^2 \leq 4 \, (\Delta A_1)^2 \, (\Delta A_2)^2 \;\; \Rightarrow \qquad (\Delta A_1)^2 \, (\Delta A_2)^2 \, \geq \, \frac{1}{4} \, |\langle [A_1, A_2] \rangle|^2 \, .$$

This states that the information one has about the expectation values of two non commuting operators A_1 and A_2 is intrinsically affected by uncertainties ΔA_1 and ΔA_2, respectively, that cannot be reduced in an arbitrary way. Then, the information one gets from a measurement is

$$A_1 \;\; \Rightarrow \;\; \langle A_1 \rangle \pm \Delta A_1 \qquad A_2 \;\; \Rightarrow \;\; \langle A_2 \rangle \pm \Delta A_2$$

where reducing ΔA_1 necessarily entails the increase of ΔA_2 (and viceversa). Then improving our information on some physical observable corresponds to losing information about an other one.

The latter inequality represents the **Minimum Uncertainty product** or relation (Heisenberg, 1930). For canonically conjugate variables such as position and momentum operators the Minimum Uncertainty relation gives

$$[x_j, p_n] = i\hbar \, \delta_{nj} \;\; \Rightarrow \quad (\Delta x_j)^2 \, (\Delta p_n)^2 \, \geq \, \frac{\hbar^2}{4} \, \delta_{nj} \;\; \Rightarrow \quad \Delta x_j \, \Delta p_j \, \geq \, \frac{\hbar}{2} \, ,$$

This leads to dramatic consequences that modify our conception of physical reality. In particular, the deterministic character of Classical Mechanics, where the dynamical state of a particle is totally defined at each time if the relevant equations of motion have been solved, is completely lost. A precise measurement of position (Δx_j small) entails large deviation Δp_n so that the (expectation value of) momentum totally undetermined. Conversely, knowing the velocity of a moving particle with arbitrary precision induces an intrinsic, complete delocalization of the particle whose position is affected by an arbitarily large deviation. The idea of "trajectory" described by a particle becomes meaningless as well as the idea of "curve of motion" in the phase space of the Hamiltonian formulation.

8.1 Commuting operators

Theorem 2. If two Hermitian operators are such that $[A, B] = 0$ then they have a common set of eigenfunctions.

Non degenerate case. Assume that the spectrum of A is non degenerate and consider an eigenstate ψ with eigenvalue α. Then $A\,\psi = \alpha\,\psi$ and

$$AB\,\psi = BA\,\psi = B\,\alpha\,\psi = \alpha\,B\,\psi\,,$$

showing that state $\psi_B = B\,\psi$ is an eigenstate of A. On the other hand, since ψ is the **only** eigenstate associated to α, one necessarily finds that $\psi_B \propto \psi$, namely,

$$\psi_B = B\,\psi = \beta\psi\,, \qquad \beta \in \mathbb{R}$$

showing that ψ is an eigenstate of B. Since ψ is a generic eigenstates of A the previous proof is valid for any eigenstate of A.

Degenerate case. Let α be the degenerate eigenvalue. Then there exists an eigenstate subset $B_M = \{\psi_i, i \in [1, M] : A\psi_i = \alpha\psi_i\}$. Consider the action of A on $B\psi_i$

$$A(B\,\psi_i) = AB\,\psi_i = BA\,\psi_i = \alpha\,(B\psi_i)$$

which shows how $(B\,\psi_i)$ is, in turn, an eigenstate associated to α. As a consequence $B\psi_i$ can be expressed as a linear superposition of vectors of B_M

$$B\psi_i = C_{in}\psi_n \quad \text{with} \quad (\psi_k, B\psi_i) = C_{ik}\,.$$

Fortunately, it is possible to reorganize eigenstates ψ_i so as to get a new set ϕ_k with the property that $B\phi_i = \beta_i\phi_i$.
To this end, it is useful to observe that matrix C is Hermitian, namely, $C^+_{ik} := C^*_{ki} \equiv C_{ik}$. Proof:

$$C^*_{ki} = (C_{ki})^* = (\psi_i, B\psi_k)^* = (\psi_k, B^+\psi_i) = (\psi_k, B\psi_i) = C_{ik}$$

Then C can be diagonalized by means of a unitary transformation U (recall that unitary matrices are such that $U^{-1} = U^+$)

$$U^+C\,U \equiv D \quad \Leftrightarrow \quad U^+_{ij}C_{jn}\,U_{nm} = U^*_{ji}C_{jn}\,U_{nm} \equiv D_{im} = d_i\delta_{im}$$

The desired set is found to be constituted by vectors $\phi_i = U^*_{ji}\psi_j$

$$B\phi_i = BU^*_{ji}\psi_j = U^*_{ji}\,(B\psi_j) = U^*_{ji}\,(C_{jn}\psi_n) = U^*_{ji}\,C_{jn}\,\delta_{nm}\,\psi_m$$

$$= U^*_{ji}\, C_{jn}\, U_{nk} U^*_{mk}\, \psi_m = \, (U^*_{ji}\, C_{jn}\, U_{nk})\, U^*_{mk}\, \psi_m = \, d_i \delta_{ik}\, \phi_k = \, d_i\, \phi_i \, ,$$

where the property of U

$$U\, U^+ = \mathbb{I} \quad \Leftrightarrow \quad U_{nk} U^+_{km} = U_{nk} U^*_{mk} = \delta_{nm}$$

representing its unitary character has been used. This proves that the choice of the new basis of vectors ϕ_i allows one to diagonalize both B and A (note that, obviously, $A\phi_i = \alpha\phi_i$).

Comment. The commutativity condition $[A, B] = 0$ of two hermition operators A and B is both physically significant and often advantageous in the solution of quantum problems. This clealry emerges from the following observations.

1) Suppose that $A = H$, the Hamilton operator. Since, within the canonical quantization rule, $[H, B] = 0$ is associated to $\{B, H\} = 0$ then operator B represents the quantum counterpart of a (classical) constant of motion B which in the Hamiltonian formalism satisfies the equation

$$\frac{dB}{dt} = \{B, H\} = 0 \quad \rightarrow \quad B = \text{time independent quantity.}$$

2) A more evident link between $\{B, H\} = 0$ and its quantum counterpart $[H, B] = 0$ is supplied by the Ehrenfest theorem. This states that

$$i\hbar \frac{d}{dt} \langle B \rangle_t = \langle [B, H] \rangle = 0 \; \rightarrow \quad \langle B \rangle_t = \text{time independent quantity.}$$

The expectation value of B behaves as a constant of motion thus confirming that it reproduces the behavior of its classical counterpart, the function B such that $\{B, H\} = 0$.

3) Quantum-mechanically, finding an operator B such that $[H, B] = 0$ entails that B and H feature a (common) **basis** of eigenfunctions which diagonalize these operators, namely, they satisfy $H\psi_m = E_m \psi_m$ and $B\psi_m = \beta_m \psi_m$. If the state is described by a generic solution $\psi(x, t)$ of the Schrödinger problem (a superposition of stationary states) then one easily shows that indeed $\langle B \rangle_t =$ constant since B is digonalized by the same eigenfunctions of H. Explicitly,

$$\langle B \rangle_t = (\psi(x,t), B\psi(x,t)) = ... = \sum_m |C_m|^2 \beta_m \quad \text{with } \psi(x,t) = \sum_m C_m e^{-itE_m/\hbar} \psi_m(x)$$

4) If two hermitian operators are such that $[A, B] = 0$ then the Heisenberg inequality takes the form $(\Delta A)^2 (\Delta B)^2 \geq |\langle [A, B] \rangle|^2 / 4 = 0$. As a consequence no mathematical constraint prevents the possibility to arbitrarily reduce ΔA, ΔB to zero. In this special case the measurement of the expectation values $\langle A \rangle \pm \Delta A$ and $\langle B \rangle \pm \Delta B$ is not affected by quantum deviations being $\Delta A, \Delta B = 0$. As an example, consider the case when the state of the system is described by one of the eigenfunctions ψ_m of the **common basis** of eigenfunctions satisfying the eigenvalue equations $A\psi_m = \alpha_m \psi_m$ and $B\psi_m = \beta_m \psi_m$. Then one easily discovers that $\Delta A, \Delta B = 0$. Check: For example, one has $(\Delta A)^2 = \langle A^2 \rangle - \langle A \rangle^2 = 0$ since $\langle A^2 \rangle = \alpha_m^2$ and $\langle A \rangle = \alpha_m$.

8.2 Complete set of operators

Definition. A set A, B ... P of Hermitian operators is called a **complete set of operators** if it contains the largest possible number of operators that commute the one with the other. The common set of eigenstates must be non degenerate. A **complete** set of eigenfunctions of the operator A is called a **basis of** A.

Examples

1) 1-dimensional systems with $H = \frac{\hat{p}^2}{2m} + U(x)$. Three choices are possible

$$i)\ \hat{x} = x\,, \qquad ii)\ \hat{p} = -i\hbar\partial_x\,, \qquad iii)\ H$$

2) 3D systems with $H = (p_1^2 + p_2^2 + p_3^2)/(2m)$ but no potential. Classically, the number of degrees of freedom is 6 corresponding to the six canonical variables x_i and p_i. The maximum number of observables corresponding to mutually commuting operators is three. In view of fundamental commutators

$$[L_j, L_m] = i\hbar\epsilon_{jmn}L_n\,, \quad [L_j, x_m] = i\hbar\epsilon_{jmn}x_n\,, \quad [L_j, p_m] = i\hbar\epsilon_{jmn}p_n\,,$$
$$[L_j, x^2] = 0\,, \quad [L_j, p^2] = 0\,, \quad [L_j, L^2] = 0\,,$$

with $x^2 = \sum_i x_i^2$ and $p^2 = \sum_i p_i^2$, there are many possibilities

$$i)\ x_1,\, x_2,\, x_3\,, \qquad ii)\ p_1,\, p_2,\, p_3\,, \qquad iii)\ x_1,\, x_2,\, p_3\,, \qquad iv)\ L_3,\, L^2, H\,....$$
$$v)\ p_i,\, L_i,\, H \propto p^2\,, \qquad vi)\ \ x_k,\, L_k,\, x^2\,,\, ...$$

The choice of angular-momentum triplet L_1, L_2 and L_3 is excluded since $[L_n, L_m] = i\epsilon_{nmk}L_k$. The fact that H commutes with various operators represents an important information. These conserved quantities are

$$p_1,\, p_2,\, p_3, \quad L_i\,,\ i = 1, 2, 3\,, \quad L^2 = L_1^2 + L_2^2 + L_3^2$$

in addition to H. Considering such constants of motion, the more advantageous choices should be cases (ii) and (iv) because the three observables correspond to constants of motion in addition to be mutually commuting operators.

3) 3D systems with $H = (p_1^2 + p_2^2 + p_3^2)/2m + U(\vec{x})$. The presence of U restricts the number of possibilities since $[p_i, H] \neq 0$, $[L_i, H] \neq 0$,

$$i)\ x_1,\, x_2,\, x_3\,, \qquad ii)\ p_1,\, p_2,\, p_3\,, \qquad iii)\ x_1,\, x_2, p_3\,,....$$

The absence of conserved quantities drastically reduces the available choices for the triplet of observable quantities.

4) 3D systems with $H = (p_1^2 + p_2^2 + p_3^2)/2m + U(|\vec{x}|)$. In this case $[L_i, H] = 0$ and thus $[L^2, H] = 0$ since U depends on $x = |\vec{x}|$. The choices are

$$i)\ x_1,\ x_2,\ x_3\,, \qquad ii)\ p_1,\ p_2,\ p_3\,, \qquad iii)\ x_1,\ x_2,\ p_3\,, \qquad iv)\ L_3,\ L^2, H\ ...$$

$$v)\ x^2,\ L_3,\ L^2 \qquad vi)\ p^2,\ L_3, L^2\ ....$$

The two preferred choices are (ii) (this involves the description of quantum states in the plane-wave basis) and (iv) which involves three conserved quantities.

Comment. These examples make it evident that, while many choices are possible which give a **complete set** of observables, only a restricted number of these lead to the optimal situation where the complete set is formed by conserved quantities.

Theorem. An operator which commutes with the operators forming a complete set can be represented as a function of such operators.

8.3 Coherent states

Coherent states $\psi_z(x)$ form that **special class of quantum states which minimizes the product of uncertainties** of the Heisenberg inequality, namely

$$\Delta x\,\Delta p \,\equiv\, \frac{\hbar}{2}\,, \qquad (\Delta A)^2 = (\psi_z, A^2\psi_z) - (\psi_z, A\psi_z)^2\,, \quad A = x, p\,.$$

Such states can be derived in several independent ways. Here, we consider the **annihilation operator coherent states** defined as the eigenstates of the annihilation operator a by the equation

$$a\,\psi_z(x) \,=\, z\,\psi_z(x) \tag{45}$$

Note that owing to

$$a = \sqrt{\frac{m\omega}{2\hbar}}\left(x + i\frac{p}{m\omega}\right)\,, \qquad a^+ = \sqrt{\frac{m\omega}{2\hbar}}\left(x - i\frac{p}{m\omega}\right)$$

ψ_z identify with the solutions of $a\,\psi_z = z\,\psi_z$ viewed as a differential problem

$$\sqrt{\frac{m\omega}{2\hbar}}\left(x + \frac{\hbar}{m\omega}\,\frac{d}{dx}\right)\,\psi_z = z\,\psi_z\,.$$

The solution has the form

$$\psi_z(x) = C\,e^{-\sigma\,x^2 + i\eta\,x} \;\;\Rightarrow\;\; \sqrt{\frac{m\omega}{2\hbar}}\left(x + \frac{\hbar}{m\omega}\,(-2\sigma x + i\eta)\right)\,\psi_z = z\,\psi_z$$

$$\sqrt{\frac{m\omega}{2\hbar}}\left(x - \frac{2\hbar}{m\omega}\,\sigma x + i\,\frac{\hbar}{m\omega}\,\eta\right)\,\psi_z = z\,\psi_z \;\;\Rightarrow$$

$$\frac{2\hbar}{m\omega}\,\sigma = 1\,, \quad \sqrt{\frac{m\omega}{2\hbar}}\,i\,\frac{\hbar}{m\omega}\,\eta = z \quad\Rightarrow\quad \sigma = \frac{m\omega}{2\hbar} = \frac{1}{2\lambda^2}\,, \quad i\eta = \sqrt{\frac{2m\omega}{\hbar}}\,z = \frac{\sqrt{2}}{\lambda}\,z$$

where $\lambda = \sqrt{\hbar/(m\omega)}$ is the harmonic-oscillator characteristic length. Then

$$\psi_z(x) = C\,e^{-\sigma\,x^2 + i\eta\,x} = \psi_z(x) = C\,\exp\left[-\frac{x^2}{2\lambda^2} + \frac{\sqrt{2}\,z}{\lambda}\,x\right]\,. \tag{46}$$

Normalization of coherent state $\psi_z(x)$.

$$(\psi_z, \psi_z) = 1 \;\;\Rightarrow\quad 1 = C^2\int dx\,\exp\left[-\frac{x^2}{\lambda^2} + \frac{\sqrt{2}\,(z + z^*)}{\lambda}\,x\right]\,, \quad \mathrm{Re}z = v$$

$$= C^2 \int dx\, \exp\left[-\frac{x^2}{\lambda^2} + \frac{2\sqrt{2}\, v}{\lambda}\, x \right] = C^2 \int dx\, \exp\left[-\frac{(x - \sqrt{2}\,\lambda v)^2}{\lambda^2} + 2v^2 \right]$$

$$= C^2\, e^{2v^2} \int dy\, \exp\left[-\frac{y^2}{\lambda^2} \right] = C^2\, e^{2v^2}\, \lambda\, \sqrt{\pi} \;\; \rightarrow \;\; C = \frac{e^{-v^2}}{\sqrt{\lambda\, \sqrt{\pi}}}$$

The **physical meaning of eigenvalue** z can be easily evidenced by considering the projection of equation $a\, \psi_z = z\, \psi_z$ on ψ_z

$$\sqrt{\frac{m\omega}{2\hbar}} \left(x + i\, \frac{p}{m\omega} \right) \psi_z = z\, \psi_z \;\; \Rightarrow \;\; \sqrt{\frac{m\omega}{2\hbar}} \left((\psi_z, x\psi_z) + \frac{i}{m\omega}\, (\psi_z, p\psi_z) \right) = z\, (\psi_z, \psi_z)$$

$$z = \sqrt{\frac{m\omega}{2\hbar}} \left(\langle x \rangle + \frac{i}{m\omega}\, \langle p \rangle \right) \;\; \Rightarrow \;\; \mathrm{Re}z = \sqrt{\frac{m\omega}{2\hbar}}\, \langle x \rangle = \frac{\langle x \rangle}{\lambda\sqrt{2}}, \;\; \mathrm{Im}z = \frac{\langle p \rangle}{\sqrt{2\hbar m\omega}} = \frac{\lambda\, \langle p \rangle}{\hbar\sqrt{2}}.$$

This allows one to recast ψ_z into a particularly expressive form

$$\psi_z(x) = C\, \exp\left[-\frac{x^2}{2\lambda^2} + \frac{\sqrt{2}\, z}{\lambda}\, x \right] = C\, \exp\left[-\frac{x^2}{2\lambda^2} + \frac{\sqrt{2}}{\lambda}\, \frac{\langle x \rangle}{\lambda\sqrt{2}}\, x + \frac{i\,\sqrt{2}}{\lambda}\, \frac{\lambda\, \langle p \rangle}{\hbar\sqrt{2}}\, x \right]$$

$$= C\, \exp\left[i\, \frac{\langle p \rangle}{\hbar}\, x - \frac{x^2 - 2x\, \langle x \rangle}{2\lambda^2} \right] = C\, e^{-\frac{\langle x \rangle^2 + i\langle p \rangle/\hbar\, x}{2\lambda^2}}\, \exp\left[-\frac{(x - \langle x \rangle)^2}{2\lambda^2} \right] = \frac{e^{i\, x\, \frac{\langle p \rangle}{\hbar}}}{\sqrt{\lambda\sqrt{\pi}}}\, e^{-\frac{(x - \langle x \rangle)^2}{2\lambda^2}}$$

The latter shows that coherent states are the combination of a localized Gaussian-like state and of a plane wave with wave vector $\langle p \rangle / \hbar$.

Heisenberg uncertainty relation. It is now relatively easy to evaluate the uncertainty relation of x and p over eigenstates (46) by exploiting once more $a\, \psi_z = z\, \psi_z$ and the fact that

$$(\psi_z, a\psi_z) = z\, (\psi_z, \psi_z) = z \;\; \Rightarrow \;\; (\psi_z, a^+\psi_z) = (\psi_z, a\psi_z)^* = z^*$$

Since

$$x = \sqrt{\frac{\hbar}{2m\omega}} \Big(a + a^+ \Big), \qquad p = -i \sqrt{\frac{\hbar m\omega}{2}} (a - a^+),$$

it follows that

$$\langle x \rangle = (\psi_z, x\, \psi_z) = \sqrt{\frac{\hbar}{2m\omega}} \Big((\psi_z, a\, \psi_z) + (\psi_z, a^+\psi_z) \Big) = \sqrt{\frac{\hbar}{2m\omega}} \Big(z + z^* \Big), \tag{47}$$

$$\langle p\rangle = (\psi_z, p\,\psi_z) = -i\sqrt{\frac{\hbar m\omega}{2}}\Big((\psi_z, a\,\psi_z) - (\psi_z, a^+\psi_z)\Big) = -i\sqrt{\frac{\hbar m\omega}{2}}(z - z^*)\,, \tag{48}$$

while the expectation values of x^2 and p^2 read

$$\langle x^2\rangle = (\psi_z, x^2\psi_z) = \frac{\hbar}{2m\omega}\,\langle a^2 + (a^+)^2 + 2a^+a + \mathbb{I}\rangle$$

$$= \frac{\hbar}{2m\omega}\Big((\psi_z, (a^+)^2\psi_z) + (\psi_z, a^2\psi_z) + 2(\psi_z, a^+a\psi_z) + (\psi_z, \psi_z)\Big) = \frac{\hbar}{2m\omega}[(z + z^*)^2 + 1]\,,$$

$$\langle p^2\rangle = (\psi_z, p\,\psi_z) = -\frac{\hbar m\omega}{2}\langle a^2 + (a^+)^2 - 2a^+a - \mathbb{I}\rangle$$

$$= \frac{-\hbar}{2m\omega}\Big((\psi_z, (a^+)^2\psi_z) + (\psi_z, a^2\psi_z) - 2(\psi_z, a^+a\psi_z) - (\psi_z, \psi_z)\Big) = \frac{\hbar m\omega}{2}[1 - (z - z^*)^2]\,,$$

giving

$$\Delta_x^2 = \langle x^2\rangle - \langle x\rangle^2 = \frac{\hbar}{2m\omega}[(z + z^*)^2 + 1] - \left(\sqrt{\frac{\hbar}{2m\omega}}\Big(z + z^*\Big)\right)^2 = \frac{\hbar}{2m\omega}\,, \tag{49}$$

$$\Delta_p^2 = \langle p^2\rangle - \langle p\rangle^2 = \frac{\hbar m\omega}{2m\omega}[1 - (z - z^*)^2] - \left(-i\sqrt{\frac{\hbar m\omega}{2}}\Big(z - z^*\Big)\right)^2 = \frac{\hbar m\omega}{2}\,, \tag{50}$$

and

$$\Delta_x^2\,\Delta_p^2 = \frac{\hbar}{2m\omega}\,\frac{\hbar m\omega}{2} = \frac{\hbar^2}{4}\,. \tag{51}$$

This proves the distinctive feature of the annihilation-operator coherent states ψ_z, namely, the fact that such states are characterized by **minimum uncertainties** Δ_x and Δ_p.

Alternative derivation of $\psi_z(x)$. By expanding states ψ_z in the basis of the number-operator eigenfunctions

$$\psi_z(x) = \sum_{n=0}^{\infty} f_n\,\psi_n(x) \qquad (\,a^+\psi_n = \sqrt{n+1}\psi_{n+1},\ a\psi_n = \sqrt{n}\psi_{n-1}\,)$$

one has

$$a\psi_z(x) = a\sum_{n=0}^{\infty} f_n\psi_n(x) = \sum_{n=0}^{\infty} f_n\,\sqrt{n}\,\psi_{n-1}(x) = \sum_{n=1}^{\infty} f_n\,\sqrt{n}\,\psi_{n-1}(x) = \sum_{n=0}^{\infty}\sqrt{n+1}f_{n+1}\,\psi_n(x)$$

Then

$$a\,\psi_z(x) = z\,\psi_z(x) \quad \Rightarrow \quad \sum_{n=0}^{\infty}\sqrt{n+1}f_{n+1}\,\psi_n(x) = z\sum_{n=0}^{\infty} f_n\,\psi_n(x)$$

Hence, by comparing the left and the right side of the equation, one finds that the coefficients f_n of the expansion are related by the simple recursive formula

$$\sqrt{n}f_n = zf_{n-1}$$

giving

$$f_n = \frac{z}{\sqrt{n}}f_{n-1} = \frac{z}{\sqrt{n}}\left(\frac{z}{\sqrt{n-1}}f_{n-2}\right) = \frac{z^2}{\sqrt{n}\sqrt{n-1}}f_{n-2} = \ldots = \frac{z^n}{\sqrt{n!}}f_0$$

Therefore

$$\psi_z(x) = f_0\sum_{n=0}^{\infty}\frac{z^n}{\sqrt{n!}}\psi_n(x)\,.$$

Parameter f_0 is determined by the normalization condition

$$1 = (\psi_z,\psi_z) = |f_0|^2\sum_{m,n}\frac{(z^*)^m}{\sqrt{m!}}\frac{z^n}{\sqrt{n!}}(\psi_m,\psi_n) = |f_0|^2\sum_{n=0}^{\infty}\frac{|z|^{2n}}{n!} = |f_0|^2\,e^{|z|^2}\,. \tag{52}$$

The final form reads

$$\psi_z(x) = e^{-\frac{|z|^2}{2}}\sum_{n=0}^{\infty}\frac{z^n}{\sqrt{n!}}\,\psi_n(x)\,. \tag{53}$$

where

$$\psi_n = \frac{1}{\sqrt{2^n\lambda\,n!\sqrt{\pi}}}\;e^{-x^2/(2\lambda^2)}\,H_n(x/\lambda)$$

Comment 1. We thus have found two independent but equivalent representations (46) and (53) of coherent states ψ_z. Combining such formulas one finds

$$\psi_z(x) = e^{-\frac{|z|^2}{2}}\sum_{n=0}^{\infty}\frac{z^n}{\sqrt{n!}}\,\psi_n(x) \equiv C\,\exp\left[-\frac{x^2}{2\lambda^2} + \frac{\sqrt{2}\,z}{\lambda}\,x\right]$$

$$e^{-\frac{|z|^2}{2}}\sum_{n=0}^{\infty}\frac{z^n}{\sqrt{n!}}\,\frac{e^{-x^2/(2\lambda^2)}}{\sqrt{2^n\lambda\,n!\sqrt{\pi}}}\,H_n(x/\lambda) \equiv C\,\exp\left[-\frac{x^2}{2\lambda^2} + \frac{\sqrt{2}\,z}{\lambda}\,x\right]$$

$C = e^{-v^2}/\sqrt{\lambda\sqrt{\pi}}$, $v = Re\,z$. This can be shown to reduce to the formula for the **generating function** of Hermite polynomials defined within the orthogonal polynomial Theory. If one remembers that function ψ_n form a basis, then

$$\psi_z(x) \,=\sum_n C_n(z)\,\psi_n(x)\,, \qquad (\psi_n,\psi_z)=C_n(z)=\,e^{-\frac{|z|^2}{2}}\,\frac{z^n}{\sqrt{n!}}\,.$$

Formula (53) is therefore the Fourier series of $\psi_z(x)$ in the harmonic-oscillator basis.

Comment 2. The set $\{\psi_z(x), z\in\mathbb{C}\}$ does not form an orthogonal basis in that

$$(\psi_z,\psi_\xi)=e^{-\frac{|z|^2+|\xi|^2}{2}}\,\sum_{m=0}^{\infty}\sum_{n=0}^{\infty}\frac{(z^*)^m\,\xi^n}{\sqrt{m!\,n!}}\,\big(\psi_m(x),\,\psi_n(x)\big)$$

$$=e^{-\frac{|z|^2+|\xi|^2}{2}}\,\sum_{n=0}^{\infty}\frac{(z^*\,\xi)^n}{n!}\,=e^{-\frac{|z|^2+|\xi|^2}{2}}e^{z^*\xi}\neq 0\,.$$

The latter, however, shows that $(\psi_z,\psi_\xi)=1$ for $z=\xi$. Then coherent state exhibit the correct normaliziation.

Time evolution of coherent states. The search of a general solution to the Schrödinger problem of the harmonic oscillator provides a very nice application of formula (21) describing the time propagation of a given initial state. If one assume that the initial state $\psi(x,0)$ is a coherent state $\psi_{z_0}(x)$ then an entire class of solutions to the Schrödinger problem is easily derived. To this end consider

$$i\hbar\partial_t\psi(x,t) = H\,\psi(x,t) \qquad \text{with } H = \hbar w(\hat{n}+1/2)\,, \tag{54}$$

and the formal solution (21) with $\psi(x,0)$ given by formula (53)

$$\psi(x,t) = e^{-itH/\hbar}\psi_{z_0}(x) = e^{-itH/\hbar}\,e^{-\frac{|z_0|^2}{2}}\sum_{n=0}^{\infty}\frac{z_0^n}{\sqrt{n!}}\,\psi_n(x) \;=\; e^{-\frac{|z_0|^2}{2}}\sum_{n=0}^{\infty}\frac{z_0^n}{\sqrt{n!}}\,e^{-itH/\hbar}\psi_n(x)$$

$$=\; e^{-\frac{|z_0|^2}{2}}\sum_{n=0}^{\infty}\frac{z_0^n}{\sqrt{n!}}\,e^{-itw(\hat{n}+1/2)}\psi_n(x) \;=\; e^{-\frac{|z_0|^2}{2}}e^{-\frac{i}{2}tw}\sum_{n=0}^{\infty}\frac{z_0^n}{\sqrt{n!}}\,e^{-itw\hat{n}}\psi_n(x)$$

$$=\; e^{-\frac{|z_0|^2}{2}}e^{-\frac{i}{2}tw}\sum_{n=0}^{\infty}\frac{z_0^n}{\sqrt{n!}}\,e^{-itwn}\psi_n(x) \;=\; e^{-\frac{|z_0|^2}{2}}e^{-\frac{i}{2}tw}\sum_{n=0}^{\infty}\frac{1}{\sqrt{n!}}\,\left(z_0e^{-itw}\right)^n\psi_n(x)$$

$$=\; e^{-\frac{i}{2}tw}\psi_{z(t)}(x)\,, \qquad z(t) = z_0e^{-itw}.$$

This results shows how, if the initial state is a coherent state, then the **time-propagation formula** (21) easily provides a solution to equation (54). A remarkable properties is that $\psi(x,t)$, up to the phase factor $e^{-itw/2}$, **preserves the coherent-state character** of the initial state $\psi_{z_0}(x)$. Also, one should note that **any choice** of the coherent-state parameter z_0 provides a different time-dependent solution. Let us denote the new solution with

$$\Phi_{z(t)}(x) = \;e^{-\frac{i}{2}tw}\psi_{z(t)}(x)\,, \qquad z(t) = z_0e^{-itw}.$$

As a consequence, by exploiting equations (47) and (48), one finds

$$\bar{x}(t) = \langle\hat{x}\rangle = (\Phi_z, x\,\Phi_z) = (\psi_z, x\,\psi_z) = \sqrt{\frac{\hbar}{2m\omega}}\big(z(t)+z^*(t)\big) = \sqrt{\frac{2\hbar}{m\omega}}\,|z_0|\,\cos(wt-\phi_0)$$

$$\bar{p}(t) = \langle\hat{p}\rangle = (\Phi_z, \hat{p}\,\Phi_z) = (\psi_z, \hat{p}\,\psi_z) = -i\sqrt{\frac{\hbar m\omega}{2}}\big(z(t)-z^*(t)\big) = -\sqrt{2\hbar m\omega}\,|z_0|\sin(wt-\phi_0)\,.$$

Comparison of quantum and classical dynamics. The latter formulas show that the time-dependent expectation values for $\hat{x}$ and $\hat{p}$ perfectly reproduce the time evolution of the position $x(t)$ and momentum $p(t)$ obtained from the harmonic-oscillator (classical) equation

$$\ddot{x} + w^2\,x \;=\; 0 \;\;\Rightarrow\;\;\; x(t) = A\cos(wt-\phi_0)\,,\;\; p(t) = m\dot{x} = -m\,Aw\,\sin(wt-\phi_0)\,.$$

One has $\bar{x}(t) \equiv x(t)$ and $\bar{p}(t) \equiv p(t)$ provided $A = |z_0|\sqrt{2\hbar/(mw)}$. The property that, classically, the total energy is constant at each time

$$H = \frac{p^2(t)}{2m} + \frac{m\omega^2}{2}\, x^2(t) = \frac{m\omega^2}{2}\, A^2 = E$$

must be compared with the expectation value of operator energy H

$$\langle H \rangle = (\Phi_z, H\,\Phi_z) = (\psi_z, H\,\psi_z) = \Big(\psi_z,\, \hbar w\,(a^+ a + 1/2)\,\psi_z\Big)$$

$$= \hbar w \Big(\psi_z,\, a^+ a\,\psi_z\Big) + \frac{\hbar w}{2}\Big(\psi_z,\, \psi_z\Big) = \hbar w \Big(a\psi_z,\, a\,\psi_z\Big) + \frac{\hbar w}{2} = \hbar w \Big(z\psi_z,\, z\,\psi_z\Big) + \frac{\hbar w}{2}$$

$$= \hbar w\,|z|^2 + \frac{\hbar w}{2} = \hbar w\,|z_0|^2 + \frac{\hbar w}{2} = E + \frac{\hbar w}{2}\,.$$

It results that, consistent with the classical energy E, $\langle H \rangle$ is constant. However, the quantum energy $\langle H \rangle$ is intrinsically different from the classical one E due to $\hbar w/2$: the latter is the so-called **zero-point energy** of the (quantum) harmonic oscillator and represents a typical quantum effect. This term, contained by construction in the harmonic-oscillator eigenvalues $E_n = \hbar w\,(n + 1/2)$, is the consequence of the fact that $[a, a^+] = 1 \neq 0$, namely, that, quantum-mechanically, $aa^+ \neq a^+ a$ (see the derivation of the HO Hamiltonian in terms of operator $\hat{n}$). Even if, classically, energy $E = mw^2 A^2$ vanishes when oscillation amplitude $A \to 0$, quantum-mechanically the the lowest energy is $\hbar w/2$ even if $|z_0| \to 0$. In this case the lowest energy state is given by

$$\Phi_z(x) = e^{-i\,\hbar w\, t/2}\,\psi_z(x) \;\to\; \Phi_0(x) = e^{-i\,\hbar w\, t/2}\,\psi_0(x)\,,$$

namely the coherent state, up to a phase factor, collapses into the HO ground state.

The average values $\bar{x}(t)$ and $\bar{p}(t)$ of operators $\hat{x}$ and $\hat{p}$ derived from **time-dependent coherent states** solution of the (**quantum**) Schrödinger problem behave like their classical counterpart. This demonstrates how the coherent-state picture of quantum systems provides the description of quantum phenomena that better approaches the classical one. Also, the perfect match between the trajectories described by $x(t)$ and $p(t)$ and those described by the average values $\bar{x}(t)$ and $\bar{p}(t)$ confirms the predictions of the Eherenfest theorem on the semiclassical behavior of expectation values.

The "quantum trajectory" $(\bar{x}(t), \bar{p}(t))$ in the phase space coincides with $(x(t), p(t))$ even if one should recall that, while $(x(t), p(t))$ describes a precise point in the phase space, state $(\bar{x}(t), \bar{p}(t))$ is **unavoidably affected by the quantum uncertainties** Δx and Δp. This destroys the deterministic character of classical Physics based on the assumption that it is possible to possess, at each time, a complete information on the state of the system.

Harmonic-oscillator eigenstates and Heisenberg uncertainty relation. Let us compare the uncertainties obtained by using the Harmonic-oscillator eigenstates with those obtained in the coherent-state picture. Recall that

$$\langle \hat{x} \rangle = (\psi_n, x\,\psi_n) = 0\,, \qquad \langle \hat{x}^2 \rangle = (\psi_n, x^2\,\psi_n) = \frac{\hbar}{2m\omega}(2n+1)$$

and

$$\langle \hat{p} \rangle = (\psi_n, \hat{p}\,\psi_n) = 0\,, \qquad \langle \hat{p}^2 \rangle = (\psi_n, \hat{p}\,\psi_n) = \frac{\hbar m\omega}{2}(2n+1)$$

giving

$$(\Delta p)^2(\Delta x)^2 = \langle \hat{x}^2 \rangle\,\langle \hat{p}^2 \rangle = \frac{\hbar}{2m\omega}(2n+1)\,\frac{\hbar m\omega}{2}(2n+1) = \frac{\hbar^2}{4}(2n+1)^2\,.$$

The comparison with the Heisenberg uncertainty relation in terms of coherent states

$$\Delta p\,\Delta x = \frac{\hbar}{2}(2n+1) \qquad \Leftrightarrow \quad \Delta p\,\Delta x \equiv \frac{\hbar}{2} \quad (\text{coherent states})\,,$$

shows that $\Delta p\,\Delta x$ rapidly increases for large n. The ground state $n = 0$ is the state for which the uncertainty relation gets its lowest possible value and coincides with the one obtained in the coherent-state picture.

Alternative derivation of coherent-state solutions to problem (54) Suppose that the solution has the form $\psi(x,t) = e^{i\alpha}\psi_z(x)$ where $z = z(t)$ and $\alpha = \alpha(t)$ and $\psi_z(x)$ is given by formula (53). By effecting a suitable choice of $z(t)$ and $\alpha(t)$ one can find a solutions of the Schrödinger problem for the harmonic-oscillator Hamiltonian. From

$$i\hbar\partial_t\Big(e^{i\alpha}\psi_z\Big) = e^{i\alpha}H\,\psi_z \quad \Rightarrow$$

one gets

$$-\hbar\dot{\alpha}\,e^{i\alpha}\psi_z \,+\, e^{i\alpha}e^{-\frac{|z|^2}{2}}\,\sum_{n=0}^{\infty}\Big(i\hbar\partial_t z^n\Big)\frac{1}{\sqrt{n!}}\,\psi_n(x) \,=e^{i\alpha}\,e^{-\frac{|z|^2}{2}}\,\sum_{n=0}^{\infty}\frac{z^n}{\sqrt{n!}}\,H\,\psi_n(x)$$

where factor $e^{i\alpha}$ can be eliminated everywhere and the time derivative $\partial_t z^n$ entails that

$$-\hbar\dot{\alpha}\,\psi_z + e^{-\frac{|z|^2}{2}}\,\sum_{n=0}^{\infty}\Big(i\hbar n\,\frac{\dot{z}}{z}\Big)\frac{z^n}{\sqrt{n!}}\,\psi_n(x) \,=e^{-\frac{|z|^2}{2}}\,\sum_{n=0}^{\infty}\frac{z^n}{\sqrt{n!}}\,\Big(\,\hbar w(n+1/2)\Big)\,\psi_n(x)$$

$$-\hbar\dot{\alpha}\,\psi_z + e^{-\frac{|z|^2}{2}}\,\sum_{n=0}^{\infty}\Big(i\hbar n\,\frac{\dot{z}}{z}\Big)\frac{z^n}{\sqrt{n!}}\,\psi_n(x) \,=e^{-\frac{|z|^2}{2}}\,\sum_{n=0}^{\infty}\frac{z^n}{\sqrt{n!}}\,\Big(\,\hbar wn\Big)\,\psi_n(x) + \frac{\hbar w}{2}\,\psi_z(x)$$

If $z(t) = z_0\,e^{-iwt}$

$$-\hbar\dot{\alpha}\,\psi_z + e^{-\frac{|z|^2}{2}}\sum_{n=0}^{\infty}\Big(i\hbar n\,(-iw)\Big)\frac{z^n}{\sqrt{n!}}\,\psi_n(x) = e^{-\frac{|z|^2}{2}}\sum_{n=0}^{\infty}\frac{z^n}{\sqrt{n!}}\Big(\hbar w n\Big)\,\psi_n(x) + \frac{\hbar w}{2}\,\psi_z(x)$$

The latter reduces to

$$-\hbar\dot{\alpha}\,\psi_z = \frac{\hbar w}{2}\,\psi_z(x) \quad \Rightarrow \quad -\hbar\dot{\alpha} = \frac{\hbar w}{2} \quad \Rightarrow \quad \alpha = -\frac{\hbar w}{2}\,t\,.$$

Parameter $\alpha(t)$ is then determined and the final form of $\psi(x,t)$ is

$$\psi(x,t) = e^{-i\,\hbar w\,t/2}\,\psi_z(x) = e^{-i\,\hbar w\,t/2}e^{-\frac{|z|^2}{2}}\sum_{n=0}^{\infty}\frac{z^n(t)}{\sqrt{n!}}\,\psi_n(x)\,, \quad z(t) = z_0\,e^{-iwt} = |z_0|\,e^{-iwt+i\phi_0}$$

which coincides with the solution $\Phi_{z(t)}(x)$ of the Schrödinger problem obtained by means of the propagator formula.

9 Angular momentum

The definition and the properties of angular-momentum operator $\hat{L}$ have been given in section 4.3. According to formula (14)

$$\hat{L} = \vec{x} \wedge \hat{p} = -i\hbar \vec{x} \wedge \nabla = \vec{e}_j \, \epsilon_{jmn} x_m p_n \quad \Rightarrow \quad L_j = -i\hbar \epsilon_{jmn} x_m \frac{\partial}{\partial x_n} \,,$$

namely,

$$\hat{L}_1 = -i\hbar(x_2 \partial_{x_3} - x_3 \partial_{x_2}) \,, \qquad \hat{L}_2 = -i\hbar(x_3 \partial_{x_1} - x_1 \partial_{x_3}) \,, \qquad \hat{L}_3 = -i\hbar(x_1 \partial_{x_2} - x_2 \partial_{x_1}) \,.$$

Angular-momentum components were shown to feature the properties (16) and (17)

$$[\, \hat{L}_j, \hat{p}_n \,] = +i\hbar \, \epsilon_{jnk} \hat{p}_k \,, \qquad [\, \hat{L}_j, \hat{x}_n \,] = +i\hbar \, \epsilon_{jnk} \hat{x}_k \,.$$

$$[\, \hat{L}_j, \hat{p}^2 \,] = 0 \,, \qquad [\, \hat{L}_j, \hat{x}^2 \,] = 0 \,,$$

where $\hat{p}^2 = \hat{p}_1^2 + \hat{p}_2^2 + \hat{p}_3^2$ and $\hat{x}^2 = \hat{x}_1^2 + \hat{x}_2^2 + \hat{x}_3^2$.

Comment. Since in the sequel vector operators $\hat{x}$, $\hat{p}$ and $\hat{L}$ cannot be confused with their classical counterpart the symbol $\hat{}$ is eliminated.

The previous properties are particularly important in the study of systems whose Hamiltonian depends on angular-momentum components, namely, $H = H(L_1, L_2, L_3, ...)$. In several applications (the most significant is the case of the two-body Hamiltonian with rotational symmetry closely related to the Hydrogen-atom Hamiltonian) the Hamiltonian commutes with both L^2 and L_j

$$[L^2, H] = 0 \,, \quad \text{while } [L_j, H] = 0 \,.$$

which thus represent conserved quantities (good quantum numbers). Therefore, one is interested in finding a basis able to describe the eigenfunctions of L^2 and the relevant eigenvalues. Unlike operators p^2 and x^2 whose components p_1, p_2, p_3 and x_1, x_2, x_3, respectively, commute $[p_i, p_j] = 0$, $[x_i, x_j] = 0$ and thus can be diagonlized simultaneously, components L_j of $\hat{L}$ **do not commute** the one with the other. Then, at most, one is allowed to find a common basis for L^2 and one of components L_j since $[L_j, L^2] = 0$ for each j (see subsection 8.1). A standard choice is the one corresponding to the pair L^2 and L_3.

9.1 Spectrum of angular-momentum operators

To determine the common set of eigenfunctions it is useful to introduce the new operators $L_+ = L_1 + iL_2$ and $L_- = L_1 - iL_2$ such that

$$(L_\pm)^+ = (L_1 \pm iL_2)^+ = L_1 \pm (-iL_2) = L_\mp \,,$$

$$[L_3, L_\pm] = \, [L_3, L_1 \pm iL_2] = [L_3, L_1] \pm i[L_3, L_2] = i\hbar L_2 \pm i(-i\hbar L_1) = \pm \, \hbar \, L_\pm \,,$$

$$[L_+, L_-] = [L_1 + iL_2, L_1 - iL_2] = [L_1 + iL_2, L_1] - i[L_1 + iL_2, L_2] = i[L_2, L_1] - i[L_1, L_2]$$
$$= -2i[L_1, L_2] = -2i(i\hbar L_3) = 2\hbar\, L_3\,.$$

Since L_1, L_2 and L_3 form an algebra (as well as L_+, L_- and L_3), similar to the scheme applied in the case of harmonic-oscillator Hamiltonian, one can implement a procedure based on raising and lowering operators. The role of $\hat{n}$, a and a^+ is played by L_3, L_- and L_+, respectively (spectrum-generating-algebra method).

Assume that ϕ_m is an eigenfunction of L_3. Then

$$L_3\phi_m = \hbar m\, \phi_m\,.$$

Since $[L^2, L_3] = 0$, in addition to L_3, function ϕ_m must diagonalize L^2. For future convenience, assume that the eigenvalue of operator $L^2 = L_1^2 + L_2^2 + L_3^2$ has the form $\hbar^2\ell(\ell+1)$ so that

$$L^2\,\phi_m \;=\; \hbar^2\ell(\ell+1)\,\phi_m\,, \quad \forall\, m\,. \tag{55}$$

At present no information is available concerning index ℓ except that $\ell(\ell+1) \geq 0$ and hence that $\ell \geq 0$. Note that one should write $\phi_m = Y_{\ell m}$ to make it evident that **a second quantum number is involved**. In the sequel, however, index ℓ is implied and symbol ϕ_m is used to simplify calculations as much as possible.

Based on the fact that $L_3L_+ - L_+L_3 = +\hbar L_+$ and $L_3L_- - L_-L_3 = -\hbar L_-$ one finds

$$L_3L_+\phi_m = (L_+L_3 + \hbar L_+)\,\phi_m = L_+(L_3\phi_m) + \hbar L_+\phi_m = \hbar m\, L_+\phi_m + \hbar L_+\phi_m = \hbar(m+1)\, L_+\phi_m$$

and

$$L_3L_-\phi_m = (L_-L_3 - \hbar L_-)\,\phi_m = L_-(L_3\phi_m) - \hbar L_-\phi_m = \hbar m\, L_-\phi_m - \hbar L_-\phi_m = \hbar(m-1)\, L_-\phi_m$$

Summarizing

$$L_3L_\pm\phi_m = \hbar(m \pm 1)\, L_\pm\phi_m \quad \Rightarrow \qquad L_\pm\phi_m = \lambda_\pm\phi_{m\pm1}$$

This formula makes evident the reason why $L_\pm$ are called **raising/lowering operators**.

Derivation of the norm $\lambda_\pm$ of $L_+\phi_m$ and of $L_-\phi_m$.

1) $(L_+\phi_m, L_+\phi_m) = (\phi_m, (L_+)^+L_+\phi_m) = (\phi_m, L_-L_+\phi_m) = \Big(\phi_m, (L^2 - L_3^2 - \hbar L_3)\phi_m\Big)$

$$= \Big(\phi_m, (\hbar^2\ell(\ell+1) - \hbar^2m^2 - \hbar^2m)\phi_m\Big) = \hbar^2(\ell(\ell+1) - m(m+1))$$

where

$$L_-L_+ = (L_1 - iL_2)(L_1 + iL_2) = L_1^2 + L_2^2 - iL_2L_1 + iL_1L_2 = L_1^2 + L_2^2 + i[L_1, L_2]$$

$$= L^2 - L_3^2 + i\,(i\hbar L_3) = L^2 - L_3^2 - \hbar L_3$$

2) $\quad (L_-\phi_m, L_-\phi_m) = (\phi_m, (L_-)^+ L_-\phi_m) = (\phi_m, L_+L_-\phi_m) = \Big(\phi_m, (L^2 - L_3^2 + \hbar L_3)\phi_m\Big)$

$$= \Big(\phi_m, (\hbar^2\ell(\ell+1) - \hbar^2 m^2 + \hbar^2 m)\phi_m\Big) = \hbar^2(\ell(\ell+1) - m(m-1))$$

where

$$L_+L_- = (L_1 + iL_2)(L_1 - iL_2) = L_1^2 + L_2^2 + iL_2L_1 - iL_1L_2 = L_1^2 + L_2^2 - i[L_1, L_2]$$

$$= L^2 - L_3^2 - i\,(i\hbar L_3) = L^2 - L_3^2 + \hbar L_3$$

Then

$$\begin{cases} L_+ Y_{\ell,\,m} = \hbar\sqrt{\ell(\ell+1) - m(m+1)}\, Y_{\ell,\,m+1} \\ \\ L_- Y_{\ell,\,m} = \hbar\sqrt{\ell(\ell+1) - m(m-1)}\, Y_{\ell,\,m-1} \end{cases} \tag{56}$$

where the ℓ-dependence of $\phi_m \equiv Y_{\ell,m}$ has been made evident.

Range of eigenvalue m. Since $(L_\pm\phi_m, L_\pm\phi_m) = \hbar^2(\ell(\ell+1) - m(m\pm 1))$ is non negative then the following inequalities

$$\ell(\ell+1) - m(m+1) \geq 0\,, \qquad \ell(\ell+1) - m(m-1) \geq 0\,,$$

must be satisfied simultaneously. Since $m = \ell+1$ violates the first inequality while $m = -\ell-1$ violates the second one, then **the range of** m resulting from the latter inequalities is given by $-\ell \leq m \leq \ell$. Another important information is that

$$L_+\phi_{+\ell} = \hbar\sqrt{\ell(\ell+1) - \ell(\ell+1)}\,\phi_{\ell+1} = 0\,, \quad L_-\phi_{-\ell} = \hbar\sqrt{\ell(\ell+1) + \ell(-\ell-1)}\,\phi_{-\ell-1} = 0\,,$$

showing that the raising/lowering operators cannot generate new eigenstates with m outside the range

$$-\ell \leq m \leq \ell\,.$$

The whole set of eigenstates ϕ_m can be derived from the repeated action of L_-

$$L_-\phi_\ell = \hbar\sqrt{\ell(\ell+1) - \ell(\ell-1)}\,\phi_{\ell-1} = \hbar\sqrt{2\ell}\,\phi_{\ell-1}\,,$$

$$L_-^2\phi_\ell = L_-(\,\hbar\sqrt{2\ell}\,\phi_{\ell-1}\,) = \hbar\sqrt{2\ell}\,L_-\phi_{\ell-1} = \hbar\sqrt{2\ell}\,\hbar\sqrt{\ell(\ell+1) - (\ell-1)(\ell-2)}\,\phi_{\ell-2}$$

$$= \hbar^2\sqrt{2\ell}\,\sqrt{2(2\ell-1)}\,\phi_{\ell-2}\,.$$

In general

$$L_-^s \phi_\ell = \hbar^s \sqrt{s!} \sqrt{\frac{(2\ell)!}{(2l-s)!}} \phi_{\ell-s}$$

showing that $L_-^s \phi_\ell \propto \phi_{\ell-s} \equiv \phi_{-\ell}$ **provided index** s assumes tha value $s = 2\ell$. After observing that s **is a positive integer**, one can conclude that necessarily $\ell = s/2$ is either integral or a half-integral

$$\ell = 0, 1, 2, \ldots \infty\,, \qquad \ell = 1/2, 3/2, 5/2, \ldots \infty\,.$$

In the next section ℓ will be shown to assume integral values. It is worth noting how the preceding derivation of the spectrum of L_3 and L^2 is totally based on the algebraic properties of the angular-momentum algebra. It confirms the utility of the **spectrum-generating-algebra method** successfully employed in the case of the harmonic-oscillator spectrum.

In conclusion, we mention two useful formulas. Consider the state $\psi_0 = Y_{\ell,0}$

$$(L_+)^s Y_{\ell,0} = \hbar^s \sqrt{\frac{(\ell+s)!}{(\ell-s)!}} Y_{\ell,s}\,, \quad (L_-)^s Y_{\ell,0} = \hbar^s \sqrt{\frac{(\ell+s)!}{(\ell-s)!}} Y_{\ell,-s} \tag{57}$$

These can be used (see the next subsection) to derive the whole set of eigenfunctions of L_3 once the wavefunction $Y_{\ell,0}$ corresponding to $m = 0$ has been given.

9.2 Angular momentum and eigenfunctions in spherical coordinates

Spherical coordinates r, φ and θ are implicitly defined by

$$x_1 = r\sin\theta\,\cos\varphi\,, \quad x_2 = r\sin\theta\,\sin\varphi\,, \quad x_3 = r\cos\theta\,, \qquad x_1^2 + x_2^2 + x_3^2 = r^2\,.$$

The use of these new coordinates is particularly advantageous for finding the eigenfunctions of L_3 and L^2. In general, the operators relevant to the angular-momentum operator can be shown to take the form

$$\begin{cases} L_1 = x_2 p_3 - x_3 p_2 = -i\hbar\left(x_2\frac{\partial}{\partial x_3} - x_3\frac{\partial}{\partial x_2}\right) = +i\hbar\left(\sin\varphi\,\frac{\partial}{\partial\theta} + \cos\varphi\,\cot\theta\frac{\partial}{\partial\varphi}\right)\,, \\ L_2 = x_3 p_1 - x_1 p_3 = -i\hbar\left(x_3\frac{\partial}{\partial x_1} - x_1\frac{\partial}{\partial x_3}\right) = -i\hbar\left(\cos\varphi\,\frac{\partial}{\partial\theta} - \sin\varphi\,\cot\theta\frac{\partial}{\partial\varphi}\right) \\ L_3 = x_1 p_2 - x_2 p_1 = -i\hbar\left(x_1\frac{\partial}{\partial x_2} - x_2\frac{\partial}{\partial x_1}\right) = -i\hbar\,\frac{\partial}{\partial\varphi}\,. \end{cases} \tag{58}$$

$$L_\pm = \hbar e^{\pm i\varphi}\left[\pm\frac{\partial}{\partial\theta} + i\,\frac{\cos\theta}{\sin\theta}\,\frac{\partial}{\partial\varphi}\right], \qquad L^2 = -\hbar^2\left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\,\frac{\partial^2}{\partial\varphi^2}\right].$$

Eigenvalue equations $L^2 Y_{\ell,m} = \hbar^2\ell(\ell+1)\,Y_{\ell,m}$ and $L_3 Y_{\ell,m} = \hbar m\,Y_{\ell,m}$ thus read

$$-\left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\,\frac{\partial^2}{\partial\varphi^2}\right] Y_{\ell,m} = \ell(\ell+1)\,Y_{\ell,m}\,, \quad -i\,\frac{\partial}{\partial\varphi}\,Y_{\ell,m} = m\,Y_{\ell,m}\,, \tag{59}$$

respectively. By assuming that $Y_{\ell,m} = \Phi_m(\varphi)\,F_{\ell m}(\theta)$, the second equation gives

$$F_{\ell m}(\theta)(-i)\,\frac{\partial}{\partial\varphi}\,\Phi_m(\varphi) = m\,\Phi_m(\varphi)\,F_{\ell m}(\theta) \;\Rightarrow\; (-i)\,\frac{\partial}{\partial\varphi}\,\Phi_m(\varphi) = m\,\Phi_m(\varphi)$$

$$\Rightarrow \quad \Phi_m(\varphi) = \frac{1}{\sqrt{2\pi}}\,e^{im\,\varphi}\,,$$

while owing to $\partial_\varphi^2\Phi_m(\varphi) = -m^2\Phi_m(\varphi)$ the first one simplies as follows

$$-\left[\frac{\partial^2}{\partial\theta^2} + \frac{\cos\theta}{\sin\theta}\,\frac{\partial}{\partial\theta} - \frac{m^2}{\sin^2\theta}\right] F_{\ell m} = \ell(\ell+1)\,F_{\ell m}\,. \tag{60}$$

The solutions to the latter are provided within mathematical Physics by the theory of Orthogonal Polynomials. They correspond to the **associated Legendre functions**

$$F_{\ell m} = P_{\ell m}(\cos\theta)\,,$$

whose definition for $m \geq 0$ is given by

$$\xi = \cos\theta\,, \qquad P_{\ell m}(\xi) = (1-\xi^2)^{m/2}\frac{d^m}{d\xi^m}\,P_\ell(\xi)\,, \quad P_\ell(\xi) = \frac{1}{2^\ell\,\ell!}\,\frac{d^\ell}{d\xi^\ell}(\xi^2-1)^\ell$$

$$P_0(\xi) = 1\,, \quad P_1(\xi) = \xi\,, \quad P_2(\xi) = \frac{1}{2}\,(3\xi^2-1)\,, \quad P_3(\xi) = \frac{1}{2}\,(5\xi^3-3\xi)\,...$$

The functions $P_\ell(\xi)$ represent the so-called **Legendre polynomials**. They are defined in the interval $-1 \leq \xi \leq +1$. In view of the properties demonstrated in the previous section one should recall that

$$-\ell \leq\ m\ \leq +\ell \quad \text{with} \quad \ell = 0, 1, 2,\ ....\infty\,.$$

Note that the case when ℓ is a half-integer ($\ell = 1/2,\ 3/2,\ 5/2\,....$) must be excluded. In this case index m, the eigenvalue of L_3, would assume half-integral values $m = r/2$ (r integer) thus involving **unphysical** eigenfunctions $\Phi_m(\varphi)$

$$\Phi_m(\varphi+2\pi) = \frac{e^{im\,(\varphi+2\pi)}}{\sqrt{2\pi}} = \frac{e^{im\,\varphi}}{\sqrt{2\pi}}\,e^{i2m\,\pi} = \,(-1)^r\,\frac{e^{im\,\varphi}}{\sqrt{2\pi}} = (-1)^r\,\Phi_m(\varphi) \neq \Phi_m(\varphi)\,.$$

Wavefunction Φ_m is no longer a single valued function, namely, it assumes a different value when going around the origin of the plane xy. Physical states cannot feature this pathological behavior. The undesired extra-factor $(-1)^r$ is removed as soon as ℓ assumes integral values.

Summary. Representing L^2 in terms of spherical coordinates allows one to reduce eigenvalue equation $L^2 Y_{\ell,m} = \hbar^2\ell(\ell{+}1)\,Y_{\ell,m}$ to differential equation (60). Its solution, found by assuming the factorized form $Y_{\ell,m} = \Phi_m(\varphi)F_{\ell,m}(\theta)$, leads, in parallel, to identify eigenfunction $\Phi_m(\varphi)$ for the (second) eigenvalue equation $L_3\Phi_m = \hbar m\,\Phi_m$. Then the orthogonal-polynomials theory allows to identify $F_{\ell,m}(\theta)$ in equation (59) with the associated Legendre functions $P_{\ell,m}(\cos\theta)$. The **spectrum-generating-algebra** method (giving information on m, ℓ and the L_3-spectrum structure) provides an alternative way to find $Y_{\ell,m}$ from the Legendre polynomials $Y_{\ell,0}(\cos\theta) = P_\ell(\cos\theta)/\sqrt{2\pi}$ where $m = 0$. This is shown at the end of this section.

Property 1. For negative values of m the **associated Legendre functions** are given by

$$P_{\ell m}(\xi) = (-1)^m\,\frac{(\ell+m)!}{(\ell-m)!}\,P_{\ell|m|}(\xi)\,. \tag{61}$$

Property 2. The following orthogonality condition can be shown to be valid for any $m \leq \ell$, h (with m positive)

$$\int_{-1}^{+1} d\xi\,P_{\ell m}(\xi)P_{hm}(\xi) = \frac{2}{2\ell+1}\,\frac{(\ell+m)!}{(\ell-m)!}\,\delta_{\ell h}\,, \quad (\,m>0\,) \tag{62}$$

Property (61) allows one to extend this equation to negative values of m.

Property 3. Since functions

$$Y_{\ell,m}(\theta,\varphi) = (-1)^m N_{\ell m}\, \Phi_m(\varphi) P_{\ell m}(\cos\theta)$$

diagonalize two Hermitian operators **they form a basis**. This is a complete orthonormal system in the space of functions $L^2(\mathbb{S}^2)$ defined on the sphere $\mathbb{S}^2$ and thus described by φ, θ.

$$(Y_{h,n},\, Y_{\ell,m}) = \int_0^{\pi}\int_0^{2\pi} \sin\theta\, d\theta\, d\varphi\, Y^*_{h,n}(\theta,\varphi)\, Y_{\ell,m}(\theta,\varphi) = \delta_{\ell h}\,\delta_{mn}\,.$$

Proof.

$$\int_0^{\pi}\int_0^{2\pi} \sin\theta\, d\theta\, d\varphi\, Y^*_{h,n}(\theta,\varphi)\, Y_{\ell,m}(\theta,\varphi)$$

$$= N_{hn}N_{\ell m}\int_0^{\pi}\sin\theta\, d\theta\, P_{hn}(\cos\theta)\, P_{\ell m}(\cos\theta)\int_0^{2\pi} d\varphi\, \Phi^*_n(\varphi)\,\Phi_m(\varphi)$$

$$= N_{hn}N_{\ell m}\int_0^{\pi}\sin\theta\, d\theta\, P_{hn}(\cos\theta)\, P_{\ell m}(\cos\theta)\,\delta_{mn} = N_{hm}N_{\ell m}\delta_{mn}\int_{+1}^{-1} d\xi\, P_{hm}(\xi)\, P_{\ell m}(\xi)$$

$$= N^2_{\ell m}\delta_{mn}\delta_{\ell h}\,\frac{2}{2\ell+1}\,\frac{(\ell+m)!}{(\ell-m)!} \quad \Rightarrow \quad N_{\ell m} = \sqrt{\frac{2\ell+1}{2}\,\frac{(\ell-m)!}{(\ell+m)!}}\,,$$

where, after introducing the integration variable $\xi = \cos\theta$ with $d\xi = -\sin\theta\, d\theta$, property (62) has been used.

Property 4. Completeness relation

$$\sum_{\ell=0}^{\infty}\sum_{m=-\ell}^{\ell} Y_{\ell,m}(\theta,\varphi)\, Y^*_{\ell,m}(\theta',\varphi') = \frac{1}{\sin\theta}\,\delta(\theta-\theta')\,\delta(\varphi-\varphi')\,.$$

Note that, for any function $f(\theta,\varphi)$ defined on the sphere one has

$$\int_{4\pi} d\Omega\, f(\theta,\varphi)\,\frac{1}{\sin\theta}\,\delta(\theta-\theta')\,\delta(\varphi-\varphi') = \int_0^{\pi}\int_0^{2\pi} d\theta d\varphi\, f(\theta,\varphi)\,\delta(\theta-\theta')\,\delta(\varphi-\varphi') = f(\varphi',\theta')\,.$$

The **probability density** relevant to the eigenfunctions of L^2 and L_3 is given by

$$|Y_{\ell,m}|^2 = \frac{1}{2\pi}\, N^2_{\ell,m}\,|P_{\ell,m}(\theta)|^2$$

being $|\Phi_m(\varphi)|^2 = 1/(2\pi)$. This shows that the probability density is independent from variable φ and is **symmetric** with respect to rotations around the z axis.

Orbitals. States with $\ell = 0$, $\ell = 1$, $\ell = 2$ and $\ell = 3$ are called **s-orbitals**, **p-orbitals**, **d-orbitals** and **f-orbitals**, respectively.

Examples. Eigenfunctions of L^2 and L_z

$$Y_{\ell,m}(\theta,\varphi) = N_{\ell,m}\, P_{\ell,m}(\theta)\, \Phi_m(\varphi)\,, \qquad -\ell \leq\, m\, \leq +\ell\,, \quad \ell = 0, 1, 2, \ldots \infty$$

$\Longrightarrow \ell = 0,\, m = 0$

$$Y_{0,0}(\theta,\varphi) = N_{0,0}\, P_{0,0}(\theta)\, \Phi_0(\varphi) = \frac{1}{\sqrt{4\pi}}$$

$\Longrightarrow \ell = 1,\, m = -1, 0. + 1$

$$Y_{1,-1}(\theta,\varphi) = -\sqrt{\frac{3}{8\pi}}\, \sin\theta\, e^{-i\varphi}\,, \quad Y_{1,0}(\theta,\varphi) = \sqrt{\frac{3}{4\pi}}\, \cos\theta\,, \quad Y_{1,+1}(\theta,\varphi) = -\sqrt{\frac{3}{8\pi}}\, \sin\theta\, e^{+i\varphi}\,,$$

$\Longrightarrow \ell = 2,\, m = -2, -1, 0. + 1, +2$

$$Y_{2,-2}(\theta,\varphi) = \sqrt{\frac{15}{32\pi}}\, \sin^2\theta\, e^{-2i\varphi}\,, \quad Y_{2,-1}(\theta,\varphi) = -\sqrt{\frac{15}{8\pi}}\, \sin\theta\cos\theta\, e^{-i\varphi}\,,$$

$$Y_{2,0}(\theta,\varphi) = \sqrt{\frac{5}{16\pi}}\, (3\cos^2\theta - 1)$$

$$Y_{2,+1}(\theta,\varphi) = -\sqrt{\frac{15}{8\pi}}\, \sin\theta\cos\theta\, e^{+i\varphi}\,, \quad Y_{2,+2}(\theta,\varphi) = \sqrt{\frac{15}{32\pi}}\, \sin^2\theta\, e^{+2i\varphi}\,.$$

Biatomic molecule. This system features a Hamiltonian of the form $H = L^2/(2I)$. The same type of Hamiltonian characterizes a thin rod of fixed length. The latter case, where the rod can be seen as an ensemble of point-like masses distributed along a straight line, includes the former one (only two point-like masses, the atoms, are present). The position of mass m_i forming the rod is

$$\vec{r}_i = r_i\, \vec{u}(\phi,\theta)\,, \quad \vec{u}(\phi,\theta) = (\cos\phi\, \sin\theta, \sin\phi\, \sin\theta, \cos\theta)\,, \tag{63}$$

$$\frac{d\vec{u}}{dt} := \vec{u}_t = \dot{\phi}\, \sin\theta\, (-\sin\phi, \cos\phi, 0) + \dot{\theta}\, (\cos\phi\, \cos\theta, \sin\phi\, \cos\theta, -\sin\theta)\,. \tag{64}$$

Then

$$\vec{v}_i^{\,2} = r_i^2\, (\vec{u}_t)^2 = r_i^2 \left(\dot{\phi}^2\, \sin^2\theta + \dot{\theta}^2\right) \;\Rightarrow\; H = \frac{1}{2}\sum_i m_i \vec{v}_i^{\,2} = \frac{I}{2}\left(\dot{\phi}^2\, \sin^2\theta + \dot{\theta}^2\right)$$

with $I = \sum_i m_i r_i^2$. Note that the scalar product $\vec{u}\cdot\vec{u}_t = 0$ is vanishing. Since the angular momentum is

$$\vec{L} = \sum\nolimits_i m_i \vec{r}_i \wedge \vec{v}_i = \left(\sum\nolimits_i m_i r_i^2\right) \vec{u} \wedge \frac{d\vec{u}}{dt} = I\, \vec{u} \wedge \vec{u}_t \;\Rightarrow\quad H = \frac{L^2}{2I}\,.$$

Proof:

$$\vec{L}^2 = I^2\ (\vec{u}\wedge\vec{u}_t)^2 = I^2\ (\vec{u}\wedge\vec{u}_t)\cdot(\vec{u}\wedge\vec{u}_t) = I^2\,\vec{u}\cdot[\vec{u}_t\wedge(\vec{u}\wedge\vec{u}_t)]$$

$$= I^2\,\vec{u}\cdot[\vec{u}\,(\vec{u}_t\cdot\vec{u}_t) - \vec{u}_t\,(\vec{u}\cdot\vec{u}_t)] = I^2\,\left[\vec{u}^2\,\vec{u}_t^2 - (\vec{u}\cdot\vec{u}_t)^2\right] = I^2\,\vec{u}_t^2 = I^2\left(\dot{\phi}^2\,\sin^2\theta + \dot{\theta}^2\right). \quad (65)$$

Quantum-mechanically, the energy operator $\hat{H} = \hat{L}^2/(2I)$ exhibits a spectrum characterized by the quantum number ℓ

$$\hat{H}Y_{\ell,m}(\theta,\phi) = E_\ell\,Y_{\ell,m}(\theta,\phi)\,, \quad E_\ell = \frac{\hbar^2}{2I}\,\ell(\ell+1)\,.$$

This does not depend on the second quantum number m in that opertor L_3 is not involved in a direct way in Hamiltonian $\hat{H}$. For this reason any state representing a generic linear combination of $Y_{\ell,m}$ with a given ℓ has the same energy eigenvalue

$$\psi_\ell(\theta,\phi) = \sum_{m=-\ell}^{\ell} C_m\,Y_{\ell,m}(\theta,\phi) \quad \rightarrow \quad \hat{H}\psi_\ell = \sum_{m=-\ell}^{\ell} C_m\,\hat{H}Y_{\ell,m} = \sum_{m=-\ell}^{\ell} C_m\,E_\ell Y_{\ell,m} = E_\ell\psi_\ell\,.$$

The **degeneracy** of energy eigenvalues is due to the fact that since $[\hat{H},\hat{L}_3] = 0$ for each ℓ-dependent energy eigenvalue E_ℓ there exist $2\ell+1$ states $Y_{\ell,m}$ with $-\ell \le m \le \ell$ associated to the same eigenvalue E_ℓ.

To get a complete **quantum** description of the biatomic molecule and of the corresponding **energy eigenstates** and eigenvalues, one should remind that the center of mass of the biatomic molecule can translate in the ambient space. Then the total energy operator must include the kinetic energy of the center of mass $\hat{P}^2/(2M)$

$$\hat{H}_{tot} = \frac{\hat{P}^2}{2M} + \frac{\hat{L}^2}{2I}\,, \quad \hat{P} = -i\hbar\nabla_X\,,$$

where $\hat{X} = \vec{X} = \vec{u}_1X_1 + \vec{u}_2X_2 + \vec{u}_3X_3$ is the position operator of the center of mass and $\hat{P} = \vec{u}_1P_1 + \vec{u}_2P_2 + \vec{u}_3P_3$ is the momentum operator whose component $P_n = -i\hbar\partial_{X_n}$ satisfy the canonical commutators $[X_k, P_n] = i\hbar\delta_{kn}$. The eigenvalue problem for the total Hamiltonian $\hat{H}_{tot}\Psi = E\Psi$ can be solved by assuming that Ψ has the factorized form $\Psi = \Phi_{\mathbf{K}}(X)\,Y_{\ell,m}(\varphi,\theta)$ in which $\Phi_{\mathbf{K}}$ is a plane wave. One has

$$\left(\frac{\hat{P}^2}{2M} + \frac{\hat{L}^2}{2I}\right)\Phi_{\mathbf{K}}(X)\,Y_{\ell,m}(\varphi,\theta) = E\,\Phi_{\mathbf{K}}(X)\,Y_{\ell,m}(\varphi,\theta) \ \text{ with } \quad \Phi_{\mathbf{K}}(X) = \frac{1}{(2\pi)^{3/2}}\,e^{i\vec{K}\vec{X}}\,,$$

whose left-hand side gives

$$Y_{\ell,m}\,\frac{\hat{P}^2}{2M}\Phi_{\mathbf{K}} + \Phi_{\mathbf{K}}\frac{\hat{L}^2}{2I}Y_{\ell,m} = \frac{\hbar^2K^2}{2M}\,\Phi_{\mathbf{K}}\,Y_{\ell,m} + \frac{\hbar^2\ell(\ell+1)}{2I}\,\Phi_{\mathbf{K}}\,Y_{\ell,m} = E\,\Phi_{\mathbf{K}}\,Y_{\ell,m}\,.$$

Then, the energy eigenvalues of the total Hamiltonian are

$$E = E(\vec{K},\ell) = \frac{\hbar^2K^2}{2M} + \frac{\hbar^2\ell(\ell+1)}{2I} \ \text{ associated to } \quad \Phi_{\mathbf{K}}(X)\,Y_{\ell,m}(\varphi,\theta)\,.$$

Application of the SGA method.
One can show that $L_+ Y_{\ell,m} = \sqrt{\ell(\ell+1) - m^2 - m}\, Y_{\ell,m+1}$, namely, $Y_{\ell,m+1}$ is generated by the action of L_+ on $Y_{\ell,m}$. A similar calculation shows that $Y_{\ell,m-1}$ is generated by the action of L_- on $Y_{\ell,m}$. Since $\xi = \cos\theta$ operator L_+ becomes

$$L_+ = \hbar e^{i\varphi} \left[\frac{\partial}{\partial\theta} + i \frac{\cos\theta}{\sin\theta} \frac{\partial}{\partial\varphi} \right] = \hbar e^{i\varphi} \left[-\sqrt{1-\xi^2} \frac{\partial}{\partial\xi} + i \frac{\xi}{\sqrt{1-\xi^2}} \frac{\partial}{\partial\varphi} \right].$$

Let us consider

$$L_+ Y_{\ell,m} = L_+ (-1)^m N_{\ell m} \frac{e^{im\varphi}}{\sqrt{2\pi}} P_{\ell,m}(\xi) = (-1)^m \frac{N_{\ell m}}{\sqrt{2\pi}} L_+ e^{im\varphi} P_{\ell,m}(\xi).$$

Due to the action of φ-derivative the term $L_+ e^{im\varphi} P_{\ell,m}(\xi)$ reduces to

$$L_+ e^{im\varphi} P_{\ell,m}(\xi) = \hbar e^{i\varphi} \left[-\sqrt{1-\xi^2} \frac{\partial}{\partial\xi} + i \frac{\xi}{\sqrt{1-\xi^2}} \frac{\partial}{\partial\varphi} \right] e^{im\varphi} P_{\ell,m}(\xi)$$

$$= \hbar e^{i(m+1)\varphi} \left[-\sqrt{1-\xi^2}\, \partial_\xi - \frac{m\,\xi}{\sqrt{1-\xi^2}} \right] P_{\ell,m}(\xi)$$

By observing that, in the latter formula,

$$\sqrt{1-\xi^2}\, \partial_\xi P_{\ell,m}(\xi) = \sqrt{1-\xi^2} \partial_\xi \Big[(1-\xi^2)^{m/2} \partial_\xi^m P_\ell(\xi) \Big] =$$

$$= \sqrt{1-\xi^2} \Big[(1-\xi^2)^{m/2} \partial_\xi^{m+1} P_\ell(\xi) - 2\xi \frac{m}{2} (1-\xi^2)^{(m-2)/2} \partial_\xi^m P_\ell(\xi) \Big]$$

$$= (1-\xi^2)^{(m+1)/2} \partial_\xi^{m+1} P_\ell(\xi) - \frac{\xi m}{\sqrt{1-\xi^2}} (1-\xi^2)^{m/2} \partial_\xi^m P_\ell(\xi) = P_{\ell,m+1}(\xi) - \frac{\xi m\, P_{\ell,m}(\xi)}{\sqrt{1-\xi^2}}$$

one obtains

$$L_+ e^{im\varphi} P_{\ell,m}(\xi) = \hbar\, e^{i(m+1)\varphi} \left[-P_{\ell,m+1}(\xi) + \frac{\xi m P_{\ell,m}(\xi)}{\sqrt{1-\xi^2}} - \frac{m\,\xi P_{\ell,m}(\xi)}{\sqrt{1-\xi^2}} \right]$$

namely

$$L_+ Y_{\ell,m}(\varphi,\theta) = L_+ (-1)^m N_{\ell m} \frac{e^{im\varphi}}{\sqrt{2\pi}} P_{\ell,m}(\xi) = -\hbar\, (-1)^m N_{\ell m} \frac{e^{i(m+1)\varphi}}{\sqrt{2\pi}} P_{\ell,m+1}(\xi)$$

$$= \hbar \sqrt{\ell(\ell+1) - m(m+1)}\, Y_{\ell,m+1}(\varphi,\theta) \qquad\qquad QED.$$

being $N_{\ell m} := \sqrt{(\ell-m)!/(\ell+m)!} = \sqrt{(\ell-m)(\ell+m+1)}\, N_{\ell\, m+1}$

Example. The derivation of L_1 in terms of spherical coordinates is discussed to exemplify the analytic calculations necessary to perform the coordinate change. Note that

$$r = \sqrt{x_1^2 + x_2^2 + x_3^2}\,, \quad \varphi = \operatorname{arctg}\frac{x_2}{x_1}\,, \quad \theta = \operatorname{arctg}\frac{\sqrt{x_1^2 + x_2^2}}{x_3}\,.$$

$$\frac{\partial f}{\partial x_3} = \frac{\partial r}{\partial x_3}\frac{\partial f}{\partial r} + \frac{\partial \theta}{\partial x_3}\frac{\partial f}{\partial \theta} = \frac{x_3}{r}\frac{\partial f}{\partial r} - \frac{\sqrt{x_1^2 + x_2^2}}{r^2}\frac{\partial f}{\partial \theta} = \cos\theta\frac{\partial f}{\partial r} - \frac{\sin\theta}{r}\frac{\partial f}{\partial \theta}$$

$$\begin{aligned}\frac{\partial f}{\partial x_2} &= \frac{\partial r}{\partial x_2}\frac{\partial f}{\partial r} + \frac{\partial \theta}{\partial x_2}\frac{\partial f}{\partial \theta} + \frac{\partial \phi}{\partial x_2}\frac{\partial f}{\partial \varphi} = \frac{x_2}{r}\frac{\partial f}{\partial r} + \frac{x_3}{r^2}\left(\partial_2\sqrt{x_1^2 + x_2^2}\right)\frac{\partial f}{\partial \theta} + \frac{x_1}{x_1^2 + x_2^2}\frac{\partial f}{\partial \varphi} \\ &= \sin\theta\sin\varphi\frac{\partial f}{\partial r} + \frac{\cos\theta}{r}\frac{x_2}{\sqrt{x_1^2 + x_2^2}}\frac{\partial f}{\partial \theta} + \frac{\cos\varphi}{r\sin\theta}\frac{\partial f}{\partial \varphi} \\ &= \sin\theta\sin\varphi\frac{\partial f}{\partial r} + \frac{\cos\theta}{r}\frac{\sin\varphi\sin\theta}{\sin\theta}\frac{\partial f}{\partial \theta} + \frac{\cos\varphi}{r\sin\theta}\frac{\partial f}{\partial \varphi} \\ &= \sin\theta\sin\varphi\frac{\partial f}{\partial r} + \frac{\cos\theta\sin\varphi}{r}\frac{\partial f}{\partial \theta} + \frac{\cos\varphi}{r\sin\theta}\frac{\partial f}{\partial \varphi}\end{aligned}$$

Then

$$\begin{aligned}x_2\frac{\partial f}{\partial x_3} - x_3\frac{\partial f}{\partial x_2} &= r\sin\theta\sin\varphi\left(\cos\theta\frac{\partial f}{\partial r} - \frac{\sin\theta}{r}\frac{\partial f}{\partial \theta}\right) \\ &\quad -r\cos\theta\left(\sin\theta\sin\varphi\frac{\partial f}{\partial r} + \frac{\cos\theta\sin\varphi}{r}\frac{\partial f}{\partial \theta} + \frac{\cos\varphi}{r\sin\theta}\frac{\partial f}{\partial \varphi}\right) \\ = -\sin^2\theta\sin\varphi\frac{\partial f}{\partial \theta} - \cos^2\theta\sin\varphi\frac{\partial f}{\partial \theta} - \cos\varphi\frac{\cos\theta}{\sin\theta}\frac{\partial f}{\partial \varphi} &= -\sin\varphi\frac{\partial f}{\partial \theta} - \cos\varphi\frac{\cos\theta}{\sin\theta}\frac{\partial f}{\partial \varphi}\end{aligned}$$

Similar calculations allows one to derive the expressions of L_3 and L_2 in spherical coordinates.

10 The central-potential Hamiltonian. The hydrogen atom.

The classical Hamiltonian of a mass undergoing the force field generated by a central (energy) potential $V(r)$ (note that V depends on $r = |\vec{x}|$) can be recast into the new form

$$H = \frac{\vec{p}^2}{2m} + V(r) \equiv \frac{L^2}{2mr^2} + \frac{p_r^2}{2m} + V(r)\,, \tag{66}$$

where radial momentum p_r and r are canonically conjugate, namely, $\{r, p_r\} = 1$ and $L = |\vec{L}|$. By exploiting these two independent expressions of H one can show that

$$\{L^2, H\} = 0\,, \quad \{L_k, H\} = 0\,.$$

As a consequence L^2 is a **constant of motion** as well as components L_k of angular momentum $\vec{L}$. Since, in addition, one has $\{L_k, L^2\} = 0$, the study of the dynamics relevant to H can take advantage from three constant of motions: H, L^2 and L_3 (the latter represents the standard choice among the three components L_k). The corresponding quantum problem will feature the same conserved quantitites

$$[H, L^2] = 0\,, \quad [H, L_z] = 0\,, \quad [L^2, L_z] = 0$$

which, owing to their commutativity, can be simultaneously observed and thus provide three good quantum numbers.

10.1 Spherical coordinates, conserved quantities and quantization

(1) Based on the first expression of H one easily proves that $\{L_k, H\} = 0$. To this end one must show that $\{L_k, p^2\} = 0$ and $\{L_k, V(r)\} = 0$. One should recall that

$$r^2 = x_1^2 + x_2^2 + x_3^2\,, \quad p^2 = p_1^2 + p_2^2 + p_3^2\,, \quad L_k = \epsilon_{kmn} x_m p_n\,,$$

(in the latter repeated indexes m and n imply summation) and that Poisson Brackets (5) becomes, in Cartesian coordinates x_i and momenta p_i,

$$\{A, B\} = \sum\nolimits_{i=1}^{3} \left[\frac{\partial A}{\partial x_i}\frac{\partial B}{\partial p_i} - \frac{\partial B}{\partial x_i}\frac{\partial A}{\partial p_i} \right].$$

Proof.

$$\{L_k, p^2\} = \sum\nolimits_{i=1}^{3} \left[\frac{\partial L_k}{\partial x_i}\frac{\partial p^2}{\partial p_i} - \frac{\partial p^2}{\partial x_i}\frac{\partial L_k}{\partial p_i} \right] = \sum\nolimits_{i=1}^{3} \frac{\partial}{\partial x_i}\Big(\epsilon_{kmn} x_m p_n\Big)\, 2p_i$$

$$= 2\sum\nolimits_{i=1}^{3} \epsilon_{kmn} p_n \frac{\partial x_m}{\partial x_i}\, p_i = 2\sum\nolimits_{i=1}^{3} \epsilon_{kmn} p_n \delta_{mi}\, p_i = 2\,\epsilon_{kmn} p_n p_m = 0$$

$$\{L_k, r^2\} = \sum_{i=1}^{3} \left[\frac{\partial L_k}{\partial x_i}\frac{\partial r^2}{\partial p_i} - \frac{\partial r^2}{\partial x_i}\frac{\partial L_k}{\partial p_i} \right] = \sum_i \left[-\frac{\partial r^2}{\partial x_i}\frac{\partial L_k}{\partial p_i} \right] = -\sum_{i=1}^{3} 2x_i \frac{\partial}{\partial p_i}\Big(\epsilon_{kmn} x_m p_n\Big)$$

$$= -2\sum_{i=1}^{3} x_i\, \epsilon_{kmn} x_m \frac{\partial p_n}{\partial p_i} = 2\sum_{i=1}^{3} x_i\, \epsilon_{kmn} x_m \delta_{ni} = 2\,\epsilon_{kmn} x_n x_m = 0\,.$$

(**Comment**. In passing, we note that an alternative derivation of $\{L_k, r^2\}$ and $\{L_k, p^2\}$ is obtained by exploiting the equality $\{AB, C\} = A\{B, C\} + B\{A, C\}$ (see the properties listed below equation (5))

$$\{L_k, p^2\} = \{\epsilon_{kmn} x_m p_n, p^2\} = \epsilon_{kmn} p_n \left\{x_m, \Sigma_i p_i^2\right\} = \epsilon_{kmn} p_n \{x_m, p_m^2\} = 2\epsilon_{kmn} p_n p_m = 0\,,$$

$$\{L_k, r^2\} = \{\epsilon_{kmn} x_m p_n, r^2\} = \epsilon_{kmn} x_m \left\{p_n, \Sigma_i x_i^2\right\} = \epsilon_{kmn} x_m \{p_n, x_n^2\} = 2\epsilon_{kmn} x_n x_m = 0\,.$$

Note that the quantum version of such calculations was performed in subsection (4.3) to show that $[L_k, r^2] = 0 = [L_k, p^2]$)

To complete the proof, we show that thanks to $\{L_k, r^2\} = 0$ then $\{L_k, V(r)\} = 0$

$$\{L_k, V(r)\} = \sum_{i=1}^{3} \left[\frac{\partial L_k}{\partial x_i}\frac{\partial V}{\partial p_i} - \frac{\partial V}{\partial x_i}\frac{\partial L_k}{\partial p_i} \right] = \sum_{i=1}^{3} \left[-\frac{\partial V}{\partial x_i}\frac{\partial L_k}{\partial p_i} \right] = -\sum_{i=1}^{3} \frac{\partial V}{\partial r}\frac{\partial r}{\partial x_i}\frac{\partial L_k}{\partial p_i}$$

$$= \frac{1}{2r}\frac{\partial V}{\partial r}\sum_{i=1}^{3} \left[-\frac{\partial r^2}{\partial x_i}\frac{\partial L_k}{\partial p_i} \right] = \frac{1}{2r}\frac{\partial V}{\partial r}\{L_k, r^2\} = 0\,,$$

$$\text{and then} \qquad \{L_k, H\} = \frac{1}{2m}\{L_k, p^2\} + \{L_k, V\} = 0\,. \qquad QED\,.$$

As a consequence $\{L^2, H\} = 0$. **Proof**.

$$\{L^2, H\} = \{\Sigma_k L_k^2, H\} = \Sigma_k \{L_k^2, H\} = \Sigma_k\, 2L_k\, \{L_k, H\} = 0\,, \qquad QED\,.$$

(2) Quantum-mechanically the second expression of H is particularly useful to obtain the energy operator $\hat{H}$ and to perform its diagonalization. One easily derives **classical Hamiltonian** (66) as a function of spherical coordinates r, θ, and φ and of the relevant momenta p_r, p_θ, and p_φ from Lagrangian $\mathcal{L} = K - V(r)$. If $\vec{r} = r\vec{u}_r$ is the position of m, one has

$$K = \frac{m}{2}\vec{v}^2 = \frac{m}{2}\left(\frac{d}{dt}\, r\vec{u}_r\right)^2 = \frac{m}{2}\left(\dot{r}\vec{u}_r + r\,\frac{d\vec{u}_r}{dt}\right)^2 = \frac{m}{2}\left(\dot{r}^2\vec{u}_r^2 + r^2\left(\frac{d\vec{u}_r}{dt}\right)^2\right)$$

$$= \frac{m}{2}\left(\dot{r}^2 + r^2\left(\dot{\theta}^2 + \dot{\varphi}^2\sin^2\theta\right)\right) = \frac{m}{2}\,\dot{r}^2 + \frac{L^2}{2mr^2}\,, \quad \left(\vec{u}_r = \vec{u}(\varphi, \theta), \text{see eqs.(63), (64)}\right)$$

In the latter equation L^2 is found by using formula (65) in which $I = \Sigma_i m_i r_i^2$ (for a rod formed by many masses m_i) reduces to $I = mr^2 + mr^2$ in the case of **biatomic molecules**. Then, with a single mass, $I = mr^2$, giving the squared angular momentum

$$L^2 = m^2 r^4 \left(\dot{\theta}^2 + \dot{\varphi}^2 \sin^2 \theta \right) .$$

Then the generalized momenta

$$p_\varphi = \frac{\partial \mathcal{L}}{\partial \dot{\varphi}} = \frac{\partial K}{\partial \dot{\varphi}} = mr^2 \sin^2 \theta \, \dot{\varphi} \,, \;\; p_\theta = \frac{\partial \mathcal{L}}{\partial \dot{\theta}} = \frac{\partial K}{\partial \dot{\theta}} = mr^2 \, \dot{\theta} \,, \;\; p_r = \frac{\partial \mathcal{L}}{\partial \dot{r}} = \frac{\partial K}{\partial \dot{r}} = m\dot{r} \,. \quad (67)$$

give $\dot{r} = p_r/m$, $\dot{\varphi} = p_\varphi/(mr^2 \sin^2 \theta)$ and $\dot{\theta} = p_\theta/(mr^2)$, which, in turn, lead to

$$K = \frac{p_r^2}{2m} + \frac{p_\theta^2}{2mr^2} + \frac{p_\varphi^2}{2mr^2 \sin^2 \theta} = \frac{p_r^2}{2m} + \frac{L^2}{2mr^2} \qquad \text{with } \; L^2 = p_\theta^2 + \frac{p_\varphi^2}{\sin^2 \theta} \,.$$

This demonstrates formula (66)

$$H = K + V(r) = \frac{L^2}{2mr^2} + \frac{p_r^2}{2m} + V(r) \,, \qquad QED \,.$$

Comment. The new canonical variables feature the canonical Poisson brackets

$$\{r, p_r\} = 1 \,, \quad \{\varphi, p_\varphi\} = 1 \,, \quad \{\theta, p_\theta\} = 1 \,,$$

$$\{r, p_\varphi\} = \{r, p_\theta\} = 0 \,, \quad \{\theta, p_r\} = \{\theta, p_\varphi\} = 0 \,, \quad \{\varphi, p_r\} = \{\varphi, p_\theta\} = 0 \,.$$

Consistent with the previous calculations, one can check that $\{H, L^2\} = 0$ within the new spherical-coordinate picture

$$\{H, L^2\} = \left\{ \frac{L^2}{2mr^2} + \frac{p_r^2}{2m} + V(r), \, L^2 \right\} = \frac{L^2}{2m} \left\{ \tfrac{1}{r^2}, L^2 \right\} + \tfrac{1}{2m} \left\{ p_r^2, L^2 \right\} + \left\{ V(r), L^2 \right\} = 0$$

in that r and p_r have zero Poisson brackets with any function of p_θ, p_φ, θ and φ. This is the case for L^2 which depends on p_θ, p_φ, $\sin\theta$. Likewise, after noting that with the new coordinates $L_3 = p_\varphi$, one finds

$$\{H, L_3\} = \left\{ \frac{L^2}{2mr^2} + \frac{p_r^2}{2m} + V(r), \, p_\varphi \right\} = \frac{1}{2mr^2} \left\{ L^2, \, p_\varphi \right\} = 0 \,.$$

Comment. It is important to note that, **thanks to the presence of constant of motion** L^2, H reduces to the Hamiltonian of a simple 1-dimensional **radial problem** involving two canonical variables r, p_r and potential $V(r)$. This preliminary analysis, at the classical level, **paves the way to the quantization process** concerning H.

Quantization. While the **angular-momentum operator** is known (its formula has been derived in Section (9)) the quantum counterpart of p_r has not been identified so far. The **canonical quantization rule** based on classical Poisson brackets suggests that

$$[r, p_r] = i\hbar \;\; \Rightarrow \;\; p_r = -i\hbar \frac{\partial}{\partial r} .$$

The latter formula is not correct even if it involves the correct commutator. In particular, p_r turns out to be non Hermitian. The solution to this problem is found by observing that any operator of the form $p_r = -i\hbar\partial_r + f(r)$ satisfies $[r, p_r] = i\hbar$. This arbitrariness allows one to find the correct definition for operator p_r. This is given by

$$\vec{p}^{\,2} = \frac{L^2}{r^2} + p_r^2 \qquad \text{if} \quad p_r := -i\hbar \left(\frac{\partial}{\partial r} + \frac{1}{r} \right) \equiv -i\hbar \frac{1}{r} \frac{\partial}{\partial r} r .$$

Proof. The explicit derivation of L^2 based on classical calculations is **not equivalent** to derivation based on momentum and position operators. **Classically**, by exploiting the identity $\vec{a} \cdot (\vec{b} \wedge \vec{c}) = \vec{b} \cdot (\vec{c} \wedge \vec{a})$

$$L^2 = (\vec{r} \wedge \vec{p})^2 = (\vec{r} \wedge \vec{p}) \cdot (\vec{r} \wedge \vec{p}) = \vec{r} \cdot \Big(\vec{p} \wedge (\vec{r} \wedge \vec{p}) \Big)$$

$$= \vec{r} \cdot \Big[\vec{r}(\vec{p} \cdot \vec{p}) - \vec{p}(\vec{p} \cdot \vec{r}) \Big] = \vec{r}^{\,2}\, \vec{p}^{\,2} - (\vec{r} \cdot \vec{p})^2 = \vec{r}^{\,2}\, \vec{p}^{\,2} - \vec{r}^{\,2} p_r^2 \quad \Rightarrow \quad \vec{p}^{\,2} = p_r^2 + \frac{L^2}{r^2} .$$

Definition (67) of momentum $p_r = m\dot{r} = \vec{u}_r \cdot m\vec{v} = (\vec{r} \cdot \vec{p})/r$ has been used which shows that p_r is the **radial component** of momentum $\vec{p}$.

Quantum-Mechanically, the identity $\vec{a} \cdot (\vec{b} \wedge \vec{c}) = \vec{b} \cdot (\vec{c} \wedge \vec{a})$ is not valid any longer because vector $\vec{a}$ does not commute with $\vec{b}$ and $\vec{c}$. Thus the order of vectors cannot be altered. The calculation of L^2 is based on $\vec{L} = \vec{u}_k \epsilon_{kmn} x_m p_n$ where x_m and p_m are operators (for short, we use the same classical vector symbols $\vec{r}$ and $\vec{p}$)

$$L^2 = \vec{L} \cdot \vec{L} = \vec{u}_k \cdot \vec{u}_h \, \epsilon_{kmn} x_m p_n \epsilon_{hrs} x_r p_s = \delta_{hk} \, \epsilon_{kmn} x_m p_n \epsilon_{hrs} x_r p_s = \epsilon_{kmn} \epsilon_{krs} \, x_m p_n x_r p_s$$

$$= (\delta_{mr}\delta_{ns} - \delta_{ms}\delta_{nr}) \, x_m p_n x_r p_s = x_m p_n x_m p_n - x_m p_n x_n p_m = x_m (x_m p_n - i\hbar\delta_{mn}) p_n$$

$$-x_m (x_n p_n - i\hbar\delta_{nn}) p_m = r^2 p^2 - i\hbar x_n p_n - x_m x_n p_n p_m + 3i\hbar x_m p_m = \; ... \; = \vec{r}^{\,2}\, \vec{p}^{\,2} - (\vec{r} \cdot \vec{p})^2 + i\hbar \vec{r} \cdot \vec{p}$$

with $\vec{r} \cdot \vec{p} = -i\hbar \vec{r} \cdot \nabla = -i\hbar r \partial_r$ (note the presence of the extra term $i\hbar \vec{r} \cdot \vec{p}$ absent in the classical version of this formula) entailing

$$\vec{p}^{\,2} = \frac{L^2}{r^2} + \frac{1}{r^2} (\vec{r} \cdot \vec{p})^2 - \frac{i\hbar}{r^2} \vec{r} \cdot \vec{p} = \frac{L^2}{r^2} - \frac{\hbar^2}{r^2} \left(r \frac{\partial}{\partial r} r \frac{\partial}{\partial r} \right) - \frac{\hbar^2}{r} \frac{\partial}{\partial r}$$

$$= \frac{L^2}{r^2} - \hbar^2 \left(\frac{1}{r} \frac{\partial}{\partial r} r \frac{\partial}{\partial r} + \frac{1}{r} \frac{\partial}{\partial r} \right) = \frac{L^2}{r^2} - \hbar^2 \left(\frac{\partial^2}{\partial r^2} + \frac{2}{r} \frac{\partial}{\partial r} \right) .$$

Then, quantum-mechanically

$$\vec{p}^{\,2} = \frac{L^2}{r^2} + p_r^2 \qquad \text{if} \quad p_r := -i\hbar \left(\frac{\partial}{\partial r} + \frac{1}{r} \right) \equiv -i\hbar \frac{1}{r} \frac{\partial}{\partial r} r \qquad QED .$$

The new definition of p_r ensures the **hermiticity** of operator p_r. To this end one must prove that $(\psi_1, p_r^+\psi_2) := (\psi_2, p_r\psi_1)^* \equiv (\psi_1, p_r\psi_2)$. **Proof**:

$$(\psi_2, p_r\psi_1)^* = \left(\int d^3x\psi_2^* \left(-i\hbar\frac{\partial}{\partial r} - \frac{i\hbar}{r}\right)\psi_1\right)^* = \int d^3x\psi_2 \left(+i\hbar\frac{\partial}{\partial r} + \frac{i\hbar}{r}\right)\psi_1^* = \int d^3x\frac{i\hbar}{r}\psi_2\psi_1^*$$

$$+i\hbar\int d\Omega\int r^2dr\psi_2\frac{\partial\psi_1^*}{\partial r} = \int d^3x\psi_2\left(\frac{i\hbar}{r}\right)\psi_1^* + i\hbar\int d\Omega\int dr\left(\frac{\partial}{\partial r}\Big(r^2\psi_2\psi_1^*\Big) - \psi_1^*\frac{\partial}{\partial r}\Big(r^2\psi_2\Big)\right)$$

$$= \int d^3x\psi_2\left(\frac{i\hbar}{r}\right)\psi_1^* + i\hbar\int d\Omega\left(r^2\psi_2\psi_1^*\right)_0^\infty + i\hbar\int d\Omega\int dr\left(-2r\psi_1^*\psi_2 - r^2\psi_1^*\frac{\partial\psi_2}{\partial r}\right)$$

$$= \int d^3x\frac{i\hbar\psi_2\psi_1^*}{r} + i\hbar\int d^3x\left(-\frac{2\psi_1^*\psi_2}{r} - \psi_1^*\frac{\partial\psi_2}{\partial r}\right) = i\hbar\int d^3x\psi_1^*\left(-\frac{1}{r} - \frac{\partial}{\partial r}\right)\psi_2 = (\psi_1, p_r\psi_2)\,.$$

10.2 Diagonalization of the central-potential problem

The time-independent Schrödinger equation for Hamiltonian

$$H = \frac{\hat{p}^2}{2m} + V(r) = \frac{\hat{L}^2}{2mr^2} + \frac{\hat{p}_r^2}{2m} + V(r)\,, \tag{68}$$

reads

$$\left[\frac{L^2}{2mr^2} - \frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial r^2} + \frac{2}{r}\frac{\partial}{\partial r}\right) + V(r)\right]\psi(r,\theta,\varphi) = E\,\psi(r,\theta,\varphi)$$

The separation ansatz $\psi(r,\theta,\varphi) = R(r)Y_{\ell,m}(\theta,\varphi)$ gives

$$\left[\frac{\hbar^2\ell(\ell+1)}{2mr^2} - \frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial r^2} + \frac{2}{r}\frac{\partial}{\partial r}\right) + V(r)\right]R(r) = E\,R(r)$$

where the eigenvalue equation $L^2Y_{\ell,m} = \hbar^2\ell(\ell+1)\,Y_{\ell,m}$ (see formula (59)) has been used.

The two labels ℓ and m relevant to L^2 and L_3 are good quantum numbers in that in addition to $[L^2, L_3] = 0$ one also has

$$[L^2, H] = [L_3, H] = 0\,.$$

Recall that $[L_3, H] = 0$ follows from $[L_k, p^2] = 0$. This means that operators H, L^2 and L_3 are **simultaneously diagonalizable**.

The equation for $R(r)$ can be reduced to a 1-dimensional Schrödinger equation by setting $R(r) = u(r)/r$

$$\left(\frac{\partial^2}{\partial r^2} + \frac{2}{r}\frac{\partial}{\partial r}\right)\frac{u}{r} = \frac{\partial}{\partial r}\left(\frac{1}{r}\frac{\partial u}{\partial r} - \frac{u}{r^2}\right) + \frac{2}{r^2}\frac{\partial u}{\partial r} - \frac{2u}{r^3}$$

$$= \frac{1}{r}\frac{\partial^2 u}{\partial r^2} - \frac{1}{r^2}\frac{\partial u}{\partial r} - \frac{1}{r^2}\frac{\partial u}{\partial r} + \frac{2u}{r^3} + \frac{2}{r^2}\frac{\partial u}{\partial r} - \frac{2u}{r^3} = \frac{1}{r}\frac{\partial^2 u}{\partial r^2} - \frac{2}{r^2}\frac{\partial u}{\partial r} + \frac{2}{r^2}\frac{\partial u}{\partial r} = \frac{1}{r}\frac{\partial^2 u}{\partial r^2}$$

$$\frac{-\hbar^2}{2mr}\frac{\partial^2 u}{\partial r^2} + \frac{\hbar^2 \ell(\ell+1)}{2mr^2}u + V(r)\,u \;=\; E\,\frac{u}{r} \quad \Rightarrow$$

$$\left[-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial r^2} + \frac{\hbar^2\ell(\ell+1)}{2mr^2} + V(r)\right] u = E\,u\,. \tag{69}$$

The latter shows that, similar to the two-body classical problem (see the effective potential obtained within the Kepler problem), the mass of the radial problem undergoes an effective potential including the repulsive centrifugal term $\hbar^2\ell(\ell+1)/(2mr^2)$. Various conditions must be considered to determine the class of solutions physically acceptable. In general, we expect that ψ can be normalized. This amounts to imposing

$$\int d^3x|\psi|^2 = \int_{4\pi} d\Omega |Y_{\ell,m}|^2 \int_0^\infty dr\, r^2 \frac{|u|^2}{r^2} = \int_0^\infty dr\,|u|^2 < \infty$$

which implies that $u(r) \simeq 1/r^{\epsilon+1/2}$ with $\epsilon > 0$ for large r (for $\epsilon = 0$ the integral exhibits a logarithmic divergence). Then we consider how the analytic form of the effective potential affects the behavior of $u(r)$ for both $r \to 0$ and $r \to \infty$. Note that the divergence of the effective potential for $r \to 0$ requires that the condition $u(0) = 0$ must be imposed.

Limit $r \to 0$. In case $V(r) = -e^2Z/r$ (Coulomb potential) the dominating term is the centrifugal one. The previous equation, rewritten as

$$\frac{1}{r^2}\left[-r^2\frac{\hbar^2}{2m}\frac{\partial^2}{\partial r^2} + \frac{\hbar^2\ell(\ell+1)}{2m} + r^2\left(V(r) - E\right)\right] u = 0$$

for $r \simeq 0$ reduces to

$$\left[-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial r^2} + \frac{\hbar^2\ell(\ell+1)}{2mr^2}\right] u = 0 \quad \Rightarrow \quad u(r) = Ar^{\ell+1} + Br^{-\ell}\,,$$

since $r^2\left(V(r) - E\right) \simeq 0$. Function $u(r)$ is the general solution of this equation. Owing to $u(0) = 0$ this class of solutions is restricted to the case $B = 0$. Then $u(r) = r^{\ell+1}\sum_n a_n r^n$.

Limit $r \to \infty$. In this case the radial equation for $u(r)$ reduces to

$$-\frac{\hbar^2}{2m}\frac{\partial^2 u}{\partial r^2} = E\,u \quad \Rightarrow \quad u = De^{kr} + Ce^{-kr}\,.$$

In order to ensure a correct normalization of $u(r)/r$ one must assume $D = 0$.

From the effective potential $V(r) + \hbar^2\ell(\ell+1)/(2mr^2)$ it follows that **bound states** correspond to negative energies E

$$-\frac{\hbar^2\,k^2}{2m} = E \quad \Rightarrow \quad k = \sqrt{\frac{2m|E|}{\hbar^2}}$$

The latter shows that, thanks to $E < 0$, parameter k is real (instead of imaginary). This involves an exponentially decreasing $u(r)$ and thus a normalizable function.

Solution of the radial problem. In Hydrogen-like atoms a single electron of charge $-e$ interacts with a positive nucleus having charge $+Ze$ due to its Z protons. Then

$$\left[-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial r^2} + \frac{\hbar^2\ell(\ell+1)}{2mr^2} - \frac{Ze^2}{r} \right] u = E\,u \;\rightarrow\; \left[\frac{\partial^2}{\partial r^2} - \frac{\ell(\ell+1)}{r^2} + \frac{2mZe^2}{\hbar^2\,r} \right] u = \frac{2m|E|}{\hbar^2}\,u$$

$$r = r_0\,x \;\rightarrow\; \frac{1}{r_0^2}\left[\frac{\partial^2}{\partial x^2} - \frac{\ell(\ell+1)}{x^2} + \frac{2mZe^2r_0}{\hbar^2\,x} \right] u = \frac{2m|E|}{\hbar^2}\,u$$

$$r_0^2 = \frac{\hbar^2}{2m|E|} = \frac{1}{k^2}\,, \quad x_0 = \frac{2mZe^2r_0}{\hbar^2} = \frac{Ze^2}{\hbar}\sqrt{\frac{2m}{|E|}} \qquad \Rightarrow \tag{70}$$

$$\left[\frac{\partial^2}{\partial x^2} - \frac{\ell(\ell+1)}{x^2} + \frac{x_0}{x} - 1 \right] u = 0\,, \tag{71}$$

Step 1. Substitute $u(r) = x^{\ell+1}e^{-x}W(x)$ where $x = r/r_0 = r\,k$. In this way the behaviors for large and small r is incorporated in the solution. Since

$$\frac{\partial^2 u}{\partial x^2} = \frac{\partial}{\partial x}\left[W\,\frac{\partial}{\partial x}\Big(x^{\ell+1}e^{-x}\Big) + x^{\ell+1}e^{-x}\frac{\partial W}{\partial x} \right]$$

$$= W\,\frac{\partial^2}{\partial x^2}\Big(x^{\ell+1}e^{-x}\Big) + 2\,\frac{\partial W}{\partial x}\,\frac{\partial}{\partial x}\Big(x^{\ell+1}e^{-x}\Big) + x^{\ell+1}e^{-x}\frac{\partial^2 W}{\partial x^2}$$

$$= W\,\frac{\partial}{\partial x}\Big((\ell+1)x^\ell e^{-x} - x^{\ell+1}e^{-x}\Big) + 2\,\frac{\partial W}{\partial x}\,\Big((\ell+1)x^\ell e^{-x} - x^{\ell+1}e^{-x}\Big) + x^{\ell+1}e^{-x}\frac{\partial^2 W}{\partial x^2}$$

$$= W\,\Big((\ell+1)\ell\,x^{\ell-1}e^{-x} - (\ell+1)x^\ell e^{-x} - (\ell+1)x^\ell e^{-x} + x^{\ell+1}e^{-x} \Big)$$

$$+2\,\frac{\partial W}{\partial x}\,\Big((\ell+1)x^\ell e^{-x} - x^{\ell+1}e^{-x}\Big) + x^{\ell+1}e^{-x}\frac{\partial^2 W}{\partial x^2}$$

$$= W\,\Big((\ell+1)\ell - 2(\ell+1)x + x^2 \Big)x^{\ell-1}e^{-x} + 2\,\frac{\partial W}{\partial x}\,\Big((\ell+1) - x\Big)x^\ell e^{-x} + \frac{\partial^2 W}{\partial x^2}x^{\ell+1}e^{-x}$$

Equation (71) becomes

$$\Big(\,(\ell+1)\ell\,-2(\ell+1)x+x^2\,\Big)x^{\ell-1}e^{-x}\,W+2\,\Big((\ell+1)-x\Big)x^{\ell}e^{-x}\,\frac{\partial W}{\partial x}\,+\,x^{\ell+1}e^{-x}\,\frac{\partial^2 W}{\partial x^2}$$
$$+\left(-\frac{\ell(\ell+1)}{x^2}+\frac{x_0}{x}-\,1\,\right)\,x^{\ell+1}e^{-x}W\,=\,0\,,$$

$$\Big(-2(\ell+1)+x\,\Big)x^{\ell}\,W+2\,\Big((\ell+1)-x\Big)x^{\ell}\frac{\partial W}{\partial x}\,+\,\frac{\partial^2 W}{\partial x^2}x^{\ell+1}+(x_0-\,x\,)\,x^{\ell}W\,=\,0\,,$$

$$x\,\frac{\partial^2 W}{\partial x^2}\,+\,2\,\Big((\ell+1)-x\Big)\frac{\partial W}{\partial x}+\big(x_0-2(\ell+1)\,\big)\,W\,=\,0\,, \tag{72}$$

Step 2. Consider the Taylor expansion $W=\sum_{k=0}^{\infty}\,a_k\,x^k$. Substituting such an expression in the latter equation gives

$$\sum_{k=0}^{\infty}a_k\,\Big[\,k(k-1)\,x^{k-1}\,+\,2(\ell+1-x)\,k\,x^{k-1}+(x_0-2(\ell+1))\,x^k\Big]\,=\,0$$

$$\sum_{k=0}^{\infty}\,\Big[\,\Big(k(k-1)\,+\,2(\ell+1)\,k\Big)a_k\,x^{k-1}+(x_0-2(\ell+1)-2k\,)a_k\,x^k\Big]\,=\,0$$

$$\sum_{k=0}^{\infty}\,\Big[\,\Big(k(k+1)\,+\,2(\ell+1)\,(k+1)\Big)a_{k+1}\,x^k+(x_0-2(\ell+1)-2k\,)a_k\,x^k\Big]\,=\,0$$

$$\Big(k(k+1)\,+\,2(\ell+1)\,(k+1)\Big)a_{k+1}\,+(x_0-2(\ell+1)-2k\,)a_k\,=0\quad\Rightarrow$$

$$a_{k+1}\,=-\frac{x_0-2(\ell+1+k)}{(k+1)\,(\,k+2\ell+2)}a_k$$

The latter recursion relation determines any a_k if a_0 is given. In order to ensure a non diverging behavior of $W(x)$ for $x\to\infty$ one must consider polynomial-like solutions such that

$$x_0-2(\ell+1+k)=0\,,\quad k\equiv N=0,1,2,3\,...$$

so that, in $W(x)$, $a_{N+1}=a_{N+2}=a_{N+3}=\,...\,=0$. As a consequence

$$x_0=2(\ell+1+N)\quad\Rightarrow\quad\frac{Ze^2}{\hbar}\sqrt{\frac{2m}{|E|}}=2(\ell+1+N)\quad\Rightarrow\quad\frac{Z^2e^4}{\hbar^2}\frac{2m}{|E|}=4(\ell+1+N)^2$$

$$\Rightarrow \quad E = -\frac{mZ^2e^4}{2\hbar^2(\ell+1+N)^2} = E_N$$

where definition (70) has been used. N is called the **radial quantum number**. The **remarkable effect** is that in order to ensure a correct physical behavior of W (by terminating the series representing W) **the energy E must be quantized**. The conventional description of **energy eigenvalues** involves the integer $n = N + \ell + 1$, called the **principal quantum number**, giving

$$E_n = -\frac{mZ^2e^4}{2\hbar^2\, n^2}\,, \qquad n = 1, 2, \ldots \infty. \tag{73}$$

Then for a given n

$$n = \text{fixed} \quad \Rightarrow \quad \ell = 0, 1, 2, \ldots n-1\,.$$

The **degeneracy of the energy eigenvalues** is complicated by the fact that, for each ℓ, the index m of the eigenfunctions $Y_{\ell,m}$ of L_3 and L^2 is such that $-\ell \le m \le \ell$. Then for a given n

$$\sum_{\ell=0}^{n-1}(2\ell+1) = 2\frac{n(n-1)}{2} + n = \boldsymbol{n}^2$$

represents the number of energy eigenfunctions corresponding to the same E_n.

Step 3. After determining the eigenvalues, the remaining problem is to characterize the polynomial-like solutions W in order to obtain a complete definition of the **energy eigenfunctions** corresponding to each E_n. By means of substitution $y = 2x$ and of the quantization condition $x_0 = 2(\ell + 1 + N) = 2n$ equation (72)

$$x\,\frac{\partial^2 W}{\partial x^2} + 2\left((\ell+1) - x\right)\frac{\partial W}{\partial x} + \big(x_0 - 2(\ell+1)\big)\,W = 0\,,$$

can be rewritten as

$$y\,\frac{\partial^2 W}{\partial y^2} + \left(s+1-y\right)\frac{\partial W}{\partial y} + N\,W = 0\,, \quad s = 2\ell+1\,,\ \ N = n-\ell-1\,. \tag{74}$$

This is the differential equation defining the **associated Laguerre polynomials**. The explicit form of its solutions, up to an arbitrary costant factor, is given by

$$W(x) = L_N^s(y) = \sum_{k=0}^{N}\frac{(-1)^k\,(N+s)!}{k!\,(k+s)!\,(N-k)!}\,y^k\,.$$

Polynomials $L_N^s(y)$ exhibit various interesting properties that are proved within the Theory of classical orthogonal Polynomials. The previous formula for such polynomials can be derived by means of

$$L_N^s(y) = (-1)^s\frac{d^s}{dy^s}\,L_{N+s}(y)\,, \quad L_r(y) = e^y\,\frac{d^r}{dy^r}\,e^{-y}\,y^r$$

where $L_r(y)$ represent the **Laguerre polynomials**. A fundamental property of such polynomials is the identity

$$\int_0^\infty dy\, e^{-y}\, y^s\, L_r^s(y) L_q^s(y) = \delta_{rq}\, \frac{(r+s)!}{r!}$$

based on which the system of associated Laguerre polynomials can be shown to form a basis. More interestingly, the previous equation, combined with the Laguerre-polynomial recursion relation

$$(N+1)L_{N+1}^s - (2N+s+1-y)L_N^s + (s+N)L_{N-1}^s = 0\,,$$

can be exploited to prove

$$\int_0^\infty dy\, e^{-y}\, y^{s+1} \left(L_N^s(y)\right)^2 = (2N+s+1)\, \frac{(N+s)!}{N!}\,. \tag{75}$$

To this end it is sufficient to rewrite $yL_N^s(y)$ in the latter integral as a linear combination of L_{N-1}^s, L_N^s and L_{N+1}^s by resorting to the previous recursion relation.

This will be involved in the calculation of the normalization factor of ψ. Then the energy eigenfunctions turn out to have the form

$$\psi = C\, Y_{\ell,m}(\theta,\varphi)\, R(r) = C\, Y_{\ell,m}(\theta,\varphi)\, \frac{u(r)}{r} = C\, Y_{\ell,m}(\theta,\varphi)\, \frac{1}{r}\, x^{\ell+1}\, e^{-x}\, W(x)$$

$$= C\, Y_{\ell,m}(\theta,\varphi)\, \frac{1}{r}\, x^{\ell+1}\, e^{-x}\, L_N^s(2x)$$

In view of definitions

$$x = \frac{r}{r_0} = k\, r\,, \quad k = \frac{\sqrt{2m_e|E|}}{\hbar}\,, \quad y = 2x\,, \quad s = 2\ell+1\,, \quad N = n-\ell-1$$

(electron mass m has been denoted with m_e to avoid confusion with the L_3 quantum number m) one finds

$$\psi_{n\ell m} = C\, Y_{\ell,m}(\theta,\varphi)\, (2k\, r)^\ell\, 2^{-\ell} k\, e^{-k\, r}\, L_{n-\ell-1}^s(2kr)$$

$$= D\, Y_{\ell,m}(\theta,\varphi)\, (2k\, r)^\ell\, e^{-k\, r}\, L_{n-\ell-1}^{2\ell+1}(2kr)\,, \quad D = 2^{-\ell}\, k\, C\,.$$

Note that parameter k is dependent on quantum number n being $E = E_n$

$$k = k_n = \frac{\sqrt{2m_e|E|}}{\hbar} = \frac{\sqrt{2m_e}}{\hbar}\, \sqrt{|E_n|} = \frac{\sqrt{2m_e}}{\hbar} \sqrt{\frac{m_e Z^2 e^4}{2\hbar^2\, n^2}} = \frac{Z e^2 m_e}{\hbar^2\, n} = \frac{Z}{a\, n}$$

where $a = \hbar^2/(e^2 m_e) = 0.529\, 10^{-10}\, m$ is the **Bohr radius**. The normalization constant D follows from the scalar product

$$\Big(\psi_{n\,\ell\,m},\,\psi_{n'\ell'm'}\Big) = \delta_{\ell'\ell}\,\delta_{m'm}\,\delta_{n'n}\,\frac{D^2\,2n\,(n+\ell)!}{(2k)^3(n-\ell-1)!} \;=\; 1$$

$$\Rightarrow \quad D_{n\ell} = \sqrt{\frac{(2k)^3(n-\ell-1)!}{2n\,(n+\ell)!}}\,.$$

Summarizing, the **energy eigenfunctions** and the relevant **eigenvalues** are

$$\psi_{n\ell m} = R_{n,\ell}(r)\,Y_{\ell,m}(\theta,\varphi) = D_{n\ell}\,(2k_n\,r)^{\ell}\,e^{-k_n\,r}\,L^{2\ell+1}_{n-\ell-1}(2k_n r)\,Y_{\ell,m}(\theta,\varphi)\,, \tag{76}$$

with $\ell = 0, 1, 2, \ldots n-1$, $|m| \le \ell$, and

$$E_n = -\,\frac{m_e Z^2 e^4}{2\hbar^2\,n^2} \;=\; -\,\frac{(Ze)^2}{2a\,n^2} \;=\; -\,\frac{m_e c^2}{2}\,\alpha^2\,\frac{Z^2}{n^2}\,, \qquad n = 1, 2, 3, \ldots \infty \tag{77}$$

where

$$\alpha = \frac{e^2}{\hbar\,c} = \frac{1}{137}\,, \quad e = 1.6\times 10^{-19}\,C\,, \quad c = 2.998\times 10^{8}\,\frac{m}{s}\,, \quad h = 6.626\times 10^{-34}\,J\,s$$

$$m_e = 9.11\times 10^{-31}\,Kg\,, \quad 1\,eV = 1.6\times 10^{-19}\,J\,, \quad m_e c^2 = 8.187\times 10^{-15}\,J \;= 0.51098\,MeV\,.$$

In the third expression of equation (77), α is called **fine-structure constant**.

In the hydrogen atom the firts-level energy ($Z = n = 1$) is $E_1 = -m_e c^2\alpha^2/2 = -13.6\,eV$. Then the ionization energy (to get the electron at an infinite distance) is $|E_1|$.

Radial probability density. The probability that an electron in the state $\psi_{n\ell m}$ is in the position $(r, \theta\varphi)$ is given by $|\psi_{n\ell m}(r, \theta\varphi, t)|^2\, r^2 dr\, d\Omega$ with $d\Omega = \sin\theta d\theta d\varphi$. Angular integration gives the radial probability density

$$P_{n\ell}(r)dr = \int_{4\pi} d\Omega\, dr\, r^2\, |\psi_{n\ell m}(r, \theta\varphi)|^2$$

$$= D_{n\ell}^2\, (2k_n\, r)^{2\ell}\, e^{-2k_n\, r} \left(L_{n-\ell-1}^{2\ell+1}(2k_n r)\right)^2 r^2 dr \;\times \int_{4\pi} d\Omega |Y_{\ell,m}(\theta, \varphi)|^2 \;= R_{n\ell}^2(r)\, r^2\, dr \;\rightarrow$$

$$P_{n\ell}(r) = r^2\, R_{n,\ell}^2(r) \;= r^2 \times\; D_{n\ell}^2\, (2k_n\, r)^{2\ell}\, e^{-2k_n\, r} \left(L_{n-\ell-1}^{2\ell+1}(2k_n r)\right)^2 .$$

Quantum number m is clearly unimportant. The associated Laguerre polynomial $L_{n-\ell-1}^{s}(2k_n r)$ can be shown to have $N = n - \ell - 1$ distinct zeros that therefore are inherited by $P_{n,\ell}(r)$. Eigenstates with $\ell = 0$ are nonzero at $r = 0$ while any other eigenfunction with $\ell > 0$ vanishes at the origin. Note that, even if $R_{n,0}^2(r) \neq 0$ at $r = 0$ the radial probability density $P_{n,0}(r)$ is zero at $r = 0$. Then the probability to find the electron close to $r = 0$ tends to zero.

Atomic spectra. The energy eigenvalues of the hydrogen atom allow to predict the atomic spectra due to the transition between different energy levels

$$\hbar w_{mn} = E_m - E_n = \frac{m_e c^2}{2}\, \alpha \left(\frac{1}{n^2} - \frac{1}{m^2}\right)$$

This formula perfectly matches the experimental observations.

Radial-position expectation values. By using the recursion relation of associated Laguerre polynomials one can show

$$\langle r\rangle_{n\ell} = \int d^3x\, |\psi_{n\ell m}(r, \theta, \varphi)|^2\, r = \frac{a}{2Z}(3n^2 - \ell(\ell + 1))\,,$$

and for $\ell = n - 1$ (the maximum allowed value) $\langle r\rangle_{n,n-1} = (a/Z)n(n + 1/2)$. Moreover

$$\langle r^2\rangle_{n,n-1} = \int d^3x\, |\psi_{n\ell m}(r, \theta, \varphi)|^2\, r^2 = \frac{a^2}{Z^2} n^2(n + 1)(n + 1/2)\,.$$

Then

$$\Delta r = \sqrt{\langle r^2\rangle_{n,n-1} - \langle r\rangle_{n,n-1}^2} = \frac{a\, n}{Z}\sqrt{\frac{1}{2}(n + 1/2)} \quad \Rightarrow \quad \frac{\Delta r}{\langle r\rangle_{n,n-1}} = \frac{1}{\sqrt{2n + 1}}\,,$$

showing that the relevance of the radial uncertainty becomes more and more negligible for large n. Hence, for large energies, the **orbitals** more and more appear to be well defined spherical shells on which electrons move. This well reproduces the **classical behavior** expected for large energies (semiclassical limit). The radial probability density

$$P_{n,n-1}(r) = \; C\, r^{2n}\, e^{-2k_n\, r} \quad \Rightarrow \quad \frac{d}{dr} P_{n,n-1} = C\, r^{2n}\, e^{-2k_n\, r} \left(\frac{2n}{r} - 2k_n\right) = 0$$

reachs its maximum at $r = n/k_n = an^2/Z$ which identifies with $\langle r\rangle_{n,n-1}$ for large n.

Classical orbits of Kepler's problem. The quantum state with $\ell = 0$ and thus $m = 0$ corresponds to the classical case $\vec{L} = \vec{r} \wedge \vec{p} = 0$ meaning that the motion takes place along a straight line. For a given n, the opposite situation corresponds to states with $\ell = n - 1$. The condition $\ell = n - 1$ is equivalent to setting $p_r = m\dot{r} = 0$ (circular orbit) in (the classical formula)

$$E = \frac{p_r^2}{2m} + \frac{L^2}{2mr^2} + U(r) \equiv \frac{L^2}{2mr^2} + U(r)\,, \quad r = constant$$

at a fixed energy E which should be equal to E_n. Indeed, in such an extreme case, if one substitutes $L^2 \equiv \hbar^2\ell(\ell+1) = \hbar^2 n(n-1)$ and $r \equiv \langle r\rangle_{n,n-1} = an(n+1/2)/Z$ in the previous formula then the quantized energy $E = E_n$ is recovered for sufficiently large n.

For $\ell = n - 1$ the motion takes place on a circular orbit contained in a plane orthogonal to angular momentum $\vec{L}$ which, in general, is not parallel to the z axis of the Cartesian reference frame. In this case one would have $L_2, L_1 \neq 0$. Quantum mechanically, however, the latter are not observable. This is the reason why the degeneracy of energy eigenstates involves the quantum number $\hbar\, m$: there is **no preferred position** of $\vec{L}$ and thus no preferred projection L_3. The classical scenario is recovered if $m = +\ell = n - 1$ involving the maximum value of $\langle L_3\rangle$. In this case the condition $L_2, L_1 \equiv 0 \;\Leftrightarrow\; \vec{L} = L_3\vec{u}_3$ is fulfilled and the plane of the circular orbit coincides with the xy plane.

Calculation of normalization constant D. To determine the normalization constant D one has to calculate

$$\Big(\psi_{n\,\ell\,m},\, \psi_{n'\ell'm'}\Big) = \int_{4\pi} d\Omega \int_0^\infty dr\, r^2\, \psi^*_{n\,\ell\,m}\psi_{n'\ell'm'}$$

$$= D^2 \int_{4\pi} d\Omega Y^*_{\ell,m}(\theta,\varphi) Y_{\ell',m'}(\theta,\varphi) \int_0^\infty dr\, r^2 (2k'r)^{\ell'}\, e^{-k'\,r}\, L^{2\ell'+1}_{n'-\ell'-1}(2k'r)\, (2kr)^{\ell}\, e^{-k\,r}\, L^{2\ell+1}_{n-\ell-1}(2kr)$$

where $k' = k_{n'} = Z/(an')$ and $k = k_n = Z/(an)$. The second integral can be shown to be zero for $n' \neq n$ so that

$$\Big(\psi_{n\,\ell\,m},\, \psi_{n'\ell'm'}\Big) = D^2\, \delta_{\ell'\ell}\, \delta_{m'm}\, \delta_{n'n} \int_0^\infty dr\, r^2\, (2kr)^{2\ell}\, e^{-2k\,r} \left(L^{2\ell+1}_{n-\ell-1}(2kr)\right)^2$$

$$= \ldots = \delta_{\ell'\ell}\, \delta_{m'm}\, \delta_{n'n}\, \frac{D^2}{(2k)^3} \int_0^\infty dy\, y^{2\ell+2}\, e^{-y} \left(L^{2\ell+1}_{n-\ell-1}(y)\right)^2$$

$$= \delta_{\ell'\ell}\, \delta_{m'm}\, \delta_{n'n}\, \frac{D^2}{(2k)^3}\, [2(n-\ell-1) + (2\ell+1) + 1\,]\, \frac{[\,(n-\ell-1) + (2\ell+1)\,]!}{(n-\ell-1)!}$$

$$= \delta_{\ell'\ell}\, \delta_{m'm}\, \delta_{n'n}\, \frac{D^2}{(2k)^3}\, 2n\, \frac{(n+\ell)!}{(n-\ell-1)!} \quad \Rightarrow \quad D_{n\ell} = \sqrt{\frac{(2k)^3\,(n-\ell-1)!}{2n\,(n+\ell)!}}\,.$$

where equation (75) has been used.

10.3 Two-body problem

Similar to a planet around the sun, the motion of an electron around an atomic nucleus is a two-body problem which can be reduced to a one-body problem involving the so-called **reduced mass**. This problem has been already investigated in the introduction of Section (2.3) and in subsection (3.2) (see example 4). Classically, within a generic inertial frame, one has the Hamiltonian

$$H = \frac{\vec{p}_1^{\,2}}{2m_1} + \frac{\vec{p}_2^{\,2}}{2m_2} + U(\vec{x}_1 - \vec{x}_2)\,.$$

Owing to

$$\vec{x}_1 = \vec{R} + \frac{m_2}{M}\,\vec{x}\,, \quad \vec{x}_2 = \vec{R} - \frac{m_1}{M}\,\vec{x} \quad \Rightarrow \quad \vec{v}_1 = \vec{V} + \frac{m_2}{M}\,\vec{v}\,, \quad \vec{v}_2 = \vec{V} - \frac{m_1}{M}\,\vec{v}\,,$$

where $\vec{R} = (X_1, X_2, X_3)$ is the center-of-mass position, $\vec{V}$ is the center-of-mass velocity, $\vec{x} = \vec{x}_1 - \vec{x}_2$ the relative-position vector (this corresponds to vector $\vec{r}$ used in the study of the central-potential problem) and $\vec{v} = \vec{v}_1 - \vec{v}_2$ the relevant velocity. Momentum vectors

$$\vec{P} = m_1\vec{v}_1 + m_2\vec{v}_2\,, \quad \vec{p} = \mu\,\vec{v}\,, \quad \mu = \frac{m_1\,m_2}{m_1 + m_2}\,.$$

describe the **total momentum** (a conserved quantity) and the momentum associated to $\vec{v}$, respectively. A simple calculation shows that

$$\frac{\vec{p}_1^{\,2}}{2m_1} + \frac{\vec{p}_2^{\,2}}{2m_2} = \frac{\vec{P}^2}{2M} + \frac{\vec{p}^{\,2}}{2\mu} \quad \Rightarrow \quad H = \frac{\vec{P}^2}{2M} + \frac{\vec{p}^{\,2}}{2\mu} + U(\vec{x})\,.$$

The Poisson brackets of the new canonical variables are easily derived from those of the old variables

$$\{x_{\alpha k}, p_{\beta j}\} = \delta_{\alpha\beta}\delta_{kj} \;\; \alpha, \beta = 1, 2 \quad \Rightarrow \quad \{X_k, P_j\} = \delta_{kj}\,, \;\; \{x_k, p_j\} = \delta_{kj}\,.$$

By implementing the canonical quantization process one finds

$$[X_i, P_j] = i\hbar\,\delta_{kj}\,, \;\; [x_k, p_j] = i\hbar\,\delta_{kj}\,,$$

$$\hat{P}_k = -i\hbar\frac{\partial}{\partial X_k}\,, \quad \hat{P} = -i\hbar\,\nabla_X\,, \quad \hat{p}_n = -i\hbar\frac{\partial}{\partial x_n}\,, \quad \hat{p} = -i\hbar\,\nabla_x\,.$$

Note that, similar to the classical case,

$$[\hat{P}_j, \hat{H}] = 0\,, \qquad \hat{H} = \frac{\hat{P}^2}{2M} + \frac{\hat{p}^2}{2\mu} + U(\vec{x})\,.$$

The three components of $\hat{P}$ correspond to **observable quantities** and thus are related to good quantum numbers. The eigenfunction of momentum operators $\hat{P}_j$ and thus of $\hat{P}^2$ are simple plane waves

$$\phi_K(\vec{X}) = \frac{1}{(2\pi)^{3/2}} e^{i\vec{K}\cdot\vec{X}} \;\; \text{such that} \quad \hat{P}^2\,\phi_K(\vec{X}) = -\hbar^2\,\nabla_X^2\phi_K(\vec{X}) = \hbar^2 K^2\,\phi_K(\vec{X})\,.$$

In view of this property one can implement the separation ansatz leading to the wave function

$$\Psi(\vec{X},\vec{x}) = e^{i\vec{K}\cdot\vec{X}}\,\psi(\vec{x})\,.$$

The stationary Schrödinger problem for $\hat{H}$ reduces to the eigenvalue equation for the central-potential problem

$$\hat{H}\Psi(\vec{X},\vec{x}) = E_{tot}\,\Psi(\vec{X},\vec{x}) \;\Rightarrow\quad \left(\frac{\hat{P}^2}{2M} + \frac{\hat{p}^2}{2\mu} + V(r)\right)\phi_K\,\psi = E\,\phi_K\,\psi$$

$$\Rightarrow\quad \psi\frac{\hat{P}^2}{2M}\,\phi_K + \phi_K\left(\frac{\hat{p}^2}{2\mu} + V(r)\right)\psi = E\,\phi_K\,\psi\,,\Rightarrow\qquad E_{tot} = \frac{\hbar^2K^2}{2M} + E\,,$$

where the dependence on $x = |\vec{x}| = r$ of $U(\vec{x}) = V(r)$ has been evidenced, the eigenstate ψ corresponds to $\psi_{n,\ell,m}$, and eigenvalue E is the eigenvalue E_n of $\psi_{n,\ell,m}$ obtained in the previous section from the solution of the eigenvalue problem for Hamiltonian (68). In fact, the reduced Hamiltonian operator $H_r = \hat{p}^2/(2\mu) + V(r)$ found above is completely equivalent to Hamiltonian operator (68) for the central-potential problem if the condition

$$m_2 >> m_1 \quad\Rightarrow\quad \mu = \frac{m_1\,m_2}{m_1+m_2} \simeq m_1$$

is used, and $m_1 = m_e$ and m_2 are assumed to represent the electron mass and the atomic-nucleus mass, respectively.

11 The Dirac picture of physical states and operators

In chapter (5) and, more specifically, in section (5.1) the space of physical states $L^2(V)$ has been shown to possess the structure of Hilbert space. This property was established through the theorem of Fisher-Riesz. The latter states that wave functions $\psi(\vec{x})$ and infinite-dimensional vectors $|v\rangle = \sum_k v_k |k\rangle$ are in the one-to-one correspondence described by formula (28)

$$\psi = \psi_v(\vec{x}) = \sum_i v_i \, \phi_i(\vec{x}) \quad \Leftrightarrow \qquad |v\rangle = \sum_i v_i \, |i\rangle \, .$$

The subscript v has been introduced to make evident such a correspondence. One should recall that, given ψ_v, the corresponding vector $|v\rangle$ is completely determined its components being given by $v_i = (\phi_i, \psi_v)$. The equivalence between these different representations of physical states, is also supported by the equivalence of the relevant scalar products

$$(\psi_u, \psi_v) = \sum_i u_i^* \, v_i \, = \langle u|v\rangle \, .$$

Such results pave the way to the more abstract but, at the same time, most general formulation of quantum theory, the so-called Dirac picture or representation of Quantum Mechanics, in which physical states are represented by bra and ket vectors, and physical operators correspond to matrices.

Within the vector formalism, physical operators correspond to matrices. Consider a physical operator A and an orthonormal complete system $\mathcal{B} = \{\phi_j(x) : \, j \in [0, \infty]\}$. Note that, as in the case of the harmonic oscillator, the most part of physical systems involve basis with infinitely many vectors. For the sake of simplicity, we have considered a basis with a discrete index j even if, in many situations, **the index can be continuous**. The matrix elements of a given operator A is defined as follows

$$A \qquad \Rightarrow \qquad A_{nm} = (\phi_n, \, A \, \phi_m) = \int_V dx \, \phi_n^*(x) A \, \phi_m(x)$$

This circumstance entails that, similar to wave function $\psi_v(x)$ associated to the vector $|v\rangle$, **operator A can be associated to a matrix**

$$\tilde{A} = \sum_i \sum_j |i\rangle A_{ij} \langle j| \quad \Rightarrow \tag{78}$$

$$\langle n|\tilde{A}|m\rangle = \langle n| \Big(\sum_i \sum_j |i\rangle A_{ij} \langle j| \Big) |m\rangle = \sum_i \sum_j \langle n|i\rangle A_{ij} \langle j|m\rangle = \sum_i \sum_j \delta_{ni} \, A_{ij} \, \delta_{jm} = A_{nm}$$

If $A = A^+$ (physical operators are Hermitian) the corresponding matrix is Hermitian. This means that $(A^+)_{nm} = A_{nm}$. Proof:

$$(A^+)_{nm} := A^*_{mn} = (\phi_m, \, A \, \phi_n)^* = (\phi_n, \, A^+ \phi_m) = (\phi_n, \, A \, \phi_m) = A_{nm} \quad QED \, .$$

If $\mathcal{B}$ is the basis associated to operator A then $A_{mn} = a_n \, \delta_{mn}$ where a_n are the eigenvalues of A, namely, $A\phi_n = a_n \phi_n$. The relevant matrix is diagonal $\tilde{A} = \Sigma_n \Sigma_j |n\rangle A_{nj} \langle j| = \Sigma_n \, a_n |n\rangle\langle n|$.

11.1 Change of basis

Consider the new basis $\mathcal{B}' = \{\phi'_k(x) : k \in [0,\infty]\}$. Then the matrix elements of A and ψ_v are expressed as

$$\psi_v(\vec{x}) = \sum_k v'_k\, \phi'_k(\vec{x}) \;\text{ and } \qquad A'_{nm} = (\phi'_n,\, A\,\phi'_m)\,. \tag{79}$$

Lemma. Two independent basis such as $\mathcal{B}$ and $\mathcal{B}'$ are connected by a **unitary transformation**. Proof: any element of $\mathcal{B}'$ can be represented in $\mathcal{B}$

$$\phi'_n = \sum_m S_{mn}\,\phi_m\,, \quad S_{mn} = (\phi_m,\,\phi'_n\,) = \int_V dx\,\phi^*_m(x)\,\phi'_n(x) \tag{80}$$

A matrix S is **unitary** if

$$S\,S^+ = S^+S\, = \mathbb{I} \quad \Leftrightarrow \quad \sum_i S_{mi}S^+_{in} = \sum_i S^+_{mi}S_{in} = \delta_{mn}\,, \quad S^+_{in} = S^*_{ni}\,,\; S^+_{mi} = S^*_{im}\,.$$

Then, being $S_{mn} = (\phi_m,\,\phi'_n\,)$, one has

$$\sum_i S_{mi}S^+_{in} = \sum_i S_{mi}S^*_{ni} = \sum_i (\phi_m,\,\phi'_i\,)(\phi_n,\,\phi'_i\,)^*$$

$$= \sum_i \int_V dx\,\phi^*_m(x)\,\phi'_i(x)\left(\int_V dy\,\phi^*_n(y)\,\phi'_i(y)\right)^*$$

$$= \sum_i \int_V dx\,\phi^*_m(x)\,\phi'_i(x)\int_V dy\,\phi_n(y)\,(\phi'_i(y))^* = \int_V dx\int_V dy\,\phi^*_m(x)\,\phi_n(y)\,\left(\sum_i (\phi'_i(y))^*\phi'_i(x)\right)$$

$$= \int_V dx\int_V dy\,\phi^*_m(x)\,\phi_n(y)\,\delta(x-y) = (\phi_m,\phi_n) = \delta_{mn}\,. \qquad QED$$

Likewise, one can prove that $\sum_i S^+_{mi}S_{in} = \delta_{mn}$.

Corollary 1. Equation (80) can be easily inverted thus giving ϕ_i in terms of ϕ'_n

$$\phi_i = \sum_n S^+_{ni}\,\phi'_n\,, \quad S^+_{ni} = S^*_{in}\,. \tag{81}$$

Proof.

$$\sum_n S^+_{ni}\,\phi'_n = \sum_n S^+_{ni}\sum_m S_{mn}\,\phi_m = \sum_m \left(\sum_n S_{mn}S^+_{ni}\right)\,\phi_m = \sum_m \delta_{im}\,\phi_m\, = \phi_i\,.$$

Corollary 2. Thanks to the preceding Lemma, one can show that vector components v_m and v'_n are related by the two transformations

$$v'_n = \sum_m S^+_{nm}\, v_m, \qquad v_m = \sum_i S_{mi}\, v'_i\,. \tag{82}$$

Proof: owing to formula (80), one finds that

$$v'_n = (\phi'_n, \psi_v) = \left(\sum_m S_{mn}\phi_m,\, \psi_v\right) = \sum_m S^*_{mn}\,(\phi_m,\, \psi_v) = \sum_m S^*_{mn}\, v_m = \sum_m S^+_{nm}\, v_m\,.$$

The second formula is found by exploiting the fact that $\sum_i S_{mi}S^+_{in} = \delta_{mn}$

$$\sum_i S_{mi}v'_i = \sum_i S_{mi}\, \sum_k S^+_{ik}\, v_k = \sum_k \delta_{mk}\, v_k = \, v_m\,. \qquad QED$$

Corollary 3. Unitary transformations **preserve the correct normalization** of physical states. Proof.

$$(\psi_v, \psi_v) \equiv \sum_m |v_m|^2 \, = \sum_m \left(\sum_i S_{mi}\, v'_i\right)^* \sum_j S_{mj}\, v'_j = \sum_i \sum_j \sum_m S^+_{im} S_{mj}\, (v'_i)^*\, v'_j$$

$$= \sum_i \sum_j \delta_{ij}\, (v'_i)^*\, v'_j = \sum_j |v'_j|^2\,.$$

This property is fundamental in that it ensures that **the probabilistic interpretation** inherent in the definition of quantum-mechanical states **is maintained even if the basis is changed**.

Corollary 4. The transformation of operator A is realized by exploiting once more formula (80)

$$A'_{nm} = (\phi'_n,\, A\,\phi'_m) = \left(\sum_i S_{in}\phi_i,\, A\, \sum_j S_{jm}\phi_j\right) = \sum_i S^*_{in} \sum_j S_{jm}\,(\phi_i,\, A\,\phi_j) \quad \Rightarrow$$

$$A'_{nm} = \sum_i \sum_j S^+_{ni} S_{jm}\, A_{ij}$$

Corollary 5. Covariant character of vector representation. The previous Lemma and transformations (82) imply that the description of vector $|v\rangle$ changes as follows

$$|v\rangle = \sum_n v_n |n\rangle = \sum_n \left(\sum_i S_{ni}\, v'_i\right) |n\rangle = \sum_i v'_i\, |i'\rangle\,, \qquad |i'\rangle = \sum_n S_{ni} |n\rangle \tag{83}$$

where the latter definition is consistent with equation (80). Likewise, in the new basis, the (matrix) operator $\tilde{A}$ is represented by

$$\tilde{A} = \sum_n \sum_m |n'\rangle A'_{nm} \langle m'| \tag{84}$$

Proof

$$|i'\rangle = \sum_n S_{ni}|n\rangle \quad \Rightarrow \quad \sum_i S^+_{im}|i'\rangle = \sum_i \sum_n S^+_{im} S_{ni}|n\rangle \quad \Rightarrow \quad \sum_i S^+_{im}|i'\rangle = |m\rangle$$

$$\langle m| = \Big(|m\rangle\Big)^+ = \Big(\sum_i S^+_{im}|i'\rangle\Big)^+ = \Big(\sum_i S^*_{mi}|i'\rangle\Big)^+ = \sum_i \big(S^*_{mi}\big)^* \langle i'| = \sum_i S_{mi}\langle i'|$$

$$\tilde{A} = \sum_n \sum_m A'_{nm}|n'\rangle\,\langle m'| = \sum_n \sum_m |n'\rangle \Big(\sum_i \sum_j S^+_{ni} S_{jm}\, A_{ij}\Big) \langle m'|$$

$$= \sum_i \sum_j A_{ij} \Big(\sum_n S^+_{ni}\,|n'\rangle\Big)\Big(\sum_m S_{jm}\,\langle m'|\Big) = \sum_i \sum_j A_{ij}\,|i\rangle\,\langle j|\,. \qquad QED$$

11.2 The Dirac picture

The discussion of the preceding section has shown that 1) representations of physical states relevant to two different basis must be connected by a unitary transformation (unitarity ensures that the norm of physical states is preserved), and 2) that physical states and operators can be expressed in terms of abstract vectors and matrices, respectively.

Concerning the first point it is useful to supply some example to clarify the relation between **representation** and **basis**. For example, in a 1D system, where the ambient space is $V = \mathbb{R}$, there are at least three well-known representation of physical states:

1) Harmonic-oscillator eigenstates: this basis is the set of states diagonalizing the number operator $\hat{n} = a^+ a$ and thus the harmonic-oscillator Hamiltonian $H\psi_n = \hbar w(n + 1/2)\psi_n$.

$$\psi_n(x) = \frac{1}{\sqrt{2^n \lambda\, n! \sqrt{\pi}}}\; e^{-x^2/(2\lambda^2)}\, H_n(x/\lambda)\,, \qquad (\psi_n, \psi_m) = \delta_{nm}\,, \;\; n,\, m \,=\, 0,\, 1,\, 2,\, 3,\, ...$$

2) Plane waves: this basis is the set of states satisfying the eigenvalue equation $\hat{p}\phi_k = \hbar k \phi_k$, $p = \hbar k \in \mathbb{R}$ for the momentum operator $\hat{p} = -i\hbar\partial_x$.

$$\phi_k(x) = \frac{1}{\sqrt{2\pi}}\; e^{+i\,k\,x}\;, \qquad (\phi_k, \phi_q) = \delta(k - q)\,.$$

3) position-operator eigenstates: this basis is the set of states satisfying the eigenvalue equation $\hat{x}\phi_\xi(x) = \xi\,\phi_\xi(x)$ with $\xi \in \mathbb{R}$ of position operator $\hat{x} = x$

$$\phi_\xi(x) = \delta(x - \xi)\,, \qquad (\phi_\xi, \phi_\eta) = \int_{\mathbb{R}} dx\, \delta(x - \xi)\delta(x - \eta) = \delta(\xi - \eta)\,.$$

A generic state ψ can be represented in the three following equivalent ways

$$\psi(x) = \sum_n c_n \, \psi_n(x) = \int_{\mathbb{R}} dk \, c(k) \phi_k(x) = \int_{\mathbb{R}} d\xi \, c(\xi) \phi_\xi(x)$$

each one associated to one of the three basis, where

$$c_n = (\psi_n, \psi) = \int_{\mathbb{R}} dx \, \frac{H_n(x/\lambda)}{\sqrt{2^n \lambda \, n! \sqrt{\pi}}} \, e^{-x^2/(2\lambda^2)} \, \psi(x)$$

$$c(k) = (\phi_k, \psi) = \int_{\mathbb{R}} dx \, \phi_k^*(x) \, \psi(x) = \int_{\mathbb{R}} dx \, \frac{e^{-i\,kx}}{\sqrt{2\pi}} \, \psi(x)$$

$$c(\xi) = (\phi_\xi, \psi) = \int_{\mathbb{R}} dx \, \delta(x - \xi) \, \psi(x) = \, \psi(\xi) \, .$$

Note that many other basis can be used for representing a physical state $\psi(x)$: in principle, any Hamiltonian $H = \hat{p}^2/(2m) + U(x)$ with a potential defined on the whole real axis is able to provide its own set of eigenstates and thus a new orthonormal basis.

Such examples show the common feature that **states and eigenstates are represented in terms of wave functions** while, as a necessary consequence, operators are either multiplication or differential operators. The Dirac picture abandons the formalism based on the wave-function representation of physical state to adopt the vector formalism suggested by the Fisher-Riesz theorem

$$\psi(x) \quad \Rightarrow \quad |\psi\rangle \, .$$

From now on, instead of the previous notation where $\psi_v(x)$ is associated to some vector $|v\rangle$, we use the more expressive notation in which vector $|\psi\rangle$ is associated to a given wave function $\psi(x)$ and is denoted with the same symbol ψ. If a state describes a specific physical property because it is the eigenstate of some Hermitian operator and thus it is associated to a specific eigenvalue

$$A\phi_\alpha(x) = \alpha \, \phi_\alpha(x) \quad \Rightarrow \qquad \phi_\alpha(x) \; \rightarrow \; |\alpha\rangle \, ,$$

namely, wave function ϕ_α is replaced by the ket vector $|\alpha\rangle$. Then

$$\begin{aligned} \hat{n} \, \psi_n(x) = n \, \psi_n(x) \quad &\Rightarrow \qquad \psi_n(x) \; \rightarrow \; |n\rangle \\ \hat{p} \, \phi_k(x) = \, \hbar k \, \phi_k(x) \quad &\Rightarrow \qquad \phi_k(x) \; \rightarrow \; |k\rangle \\ \hat{x} \, \phi_\xi(x) = \xi \, \phi_\xi(x) \quad &\Rightarrow \qquad \phi_\xi(x) \; \rightarrow \; |\xi\rangle \end{aligned}$$

Hence

$$\psi(x) = \sum_n c_n \, \psi_n(x) = \int_{\mathbb{R}} dk \, c(k) \phi_k(x) = \int_{\mathbb{R}} d\xi \, c(\xi) \phi_\xi(x)$$

(the latter is almost trivial in that $c(\xi) = \psi(\xi)$) will take the form

$$|\psi\rangle = \sum_n c_n \, |n\rangle = \int_{\mathbb{R}} dk \, c(k) \, |k\rangle = \int_{\mathbb{R}} d\xi \, c(\xi) |\xi\rangle = \, ... \, = \sum_\alpha C_\alpha \, |\alpha\rangle \tag{85}$$

where the latter representation reminds us that the basis of other Hermitian operators could be adopted to describe states and operators. This vector-like representation of states is certainly **more abstract**. In fact, a wave function, **independently from the information that possibly carries** (as in the case of eigenstates of some Hermitian operator), always provides information about the probability to find a particle in some point of the space since $|\psi|^2$ is the probability density.

Basis and unitary transformations. Based on the discussion of the preceding section, the three basis $\{|n\rangle : n \in [0,\infty]\}$, $\{|k\rangle : k \in \mathbb{R}\}$ and $\{|\xi\rangle : \xi \in \mathbb{R}\}$ are equivalent, namely, they are connected by **unitary transformations**. Since such basis are complete we can exploit the identity representations

$$\mathbb{I} = \sum\nolimits_n |n\rangle\langle n| = \int_{\mathbb{R}} dk\, |k\rangle\langle k| = \int_{\mathbb{R}} d\xi\, |\xi\rangle\langle\xi| \tag{86}$$

The three following cases well clarify how such unitary transformations are defined.

Case 1. By using the **coordinate representation** of the identity operator $\mathbb{I}$ in the **energy representation** of $|\psi\rangle$, one finds

$$|\psi\rangle = \sum\nolimits_n c_n\, |n\rangle = \int_{\mathbb{R}} dx\, |x\rangle\langle x| \sum\nolimits_n c_n\, |n\rangle = \int_{\mathbb{R}} dx \sum\nolimits_n c_n\, \langle x|n\rangle |x\rangle\,. \tag{87}$$

The comparison of the latter with the representation of $|\psi\rangle$ in the position-state basis $\{|x\rangle\}$

$$|\psi\rangle \equiv \int_{\mathbb{R}} dx\, \psi(x)\, |x\rangle \quad \text{implies that} \quad \psi(x) = c(x) = \sum\nolimits_n c_n\, \langle x|n\rangle$$

(recall that $c(x) = \psi(x)$). Since we know that $\psi(x) = \Sigma_n\, c_n\, \psi_n(x)$ is the Fourier series of ψ in the basis $\{\psi_n\}$, then the meaning of $\langle x|n\rangle$ is shown by the identity

$$\langle x|n\rangle = \psi_n(x)\,.$$

This calculation describes a **change of basis** analogous to the one performed in formula (83). One should recall that formula (83) described a generic change of basis

$$|v\rangle = \sum\nolimits_n v_n |n\rangle = \sum\nolimits_i v'_i\, |i'\rangle\,,$$

with

$$v_n = \sum\nolimits_i S_{ni}\, v'_i\,, \quad v'_j = \sum\nolimits_n S^+_{jn} v_n\,, \qquad |i'\rangle = \sum\nolimits_n S_{ni} |n\rangle\,, \quad |m\rangle = \sum\nolimits_n S^+_{im} |i'\rangle$$

entailing

$$\langle m|i'\rangle = S_{mi}\,, \quad \langle i'|m\rangle = \langle m|i'\rangle^* = S^*_{mi} = S^+_{im}\,.$$

If the role of basis $\{|i\rangle'\}$ is played by $\{|x\rangle : x \in \mathbb{R}\}$, in view of the following analogies

$$v_n \;\leftrightarrow\; c_n\,, \qquad v'_i \;\leftrightarrow\; \psi(x)\,, \qquad |i\rangle' \;\leftrightarrow\; |x\rangle\,, \qquad |n\rangle \;\leftrightarrow\; |n\rangle\,,$$

one finds

$$\langle n|x\rangle = S_{nx} = \psi_n^*(x)\,(\equiv \psi_n(x))\,, \quad \langle x|n\rangle = S_{nx}^* = S_{xn}^+ = \psi_n(x)\,,$$

and

$$|x\rangle = \sum_n S_{nx}|n\rangle = \sum_n \psi_n^*(x)|n\rangle\,, \quad |n\rangle = \int dx\, S_{xn}^+|x\rangle = \int dx\, \psi_n(x)|x\rangle\,.$$

Case 2. Likewise, by using the the **momentum representation** of the identity operator in the **coordinate representation** of $|\psi\rangle$, one finds

$$|\psi\rangle = \int_{\mathbb{R}} dx\, \psi(x)\,|x\rangle = \int_{\mathbb{R}} dk\,|k\rangle\langle k| \int_{\mathbb{R}} dx\, \psi(x)\,|x\rangle = \int_{\mathbb{R}} dk \int_{\mathbb{R}} dx\, \psi(x)\,\langle k|x\rangle|k\rangle \tag{88}$$

$$\text{that must be compared with} \qquad |\psi\rangle = \int_{\mathbb{R}} dk\, c(k)\,|k\rangle\,,$$

where the meaning of $c(k)$ is defined by the Fourier integral $\psi(x) = \int_{\mathbb{R}} dk\; c(k)\phi_k(x)$ (plane-wave basis). The comparison implies that

$$c(k) = \int_{\mathbb{R}} dx\, \psi(x)\,\langle k|x\rangle \;\; \Rightarrow \;\; \langle k|x\rangle = \phi_k^*(x) = \frac{e^{-ikx}}{\sqrt{2\pi}} \equiv S_{xk}^* = S_{kx}^+\,, \quad \langle x|k\rangle = S_{xk} = \phi_k(x)\,.$$

Therefore the **unitary transformation** connecting the x basis and the k basis identifies with the Fourier transform of function $\psi(x)$

$$|x\rangle = \int_{\mathbb{R}} dk\,\langle k|x\rangle\,|k\rangle = \int_{\mathbb{R}} dk\,\frac{e^{ikx}}{\sqrt{2\pi}}\,|k\rangle\,, \quad |k\rangle = \int dx\,\langle x|k\rangle\,|x\rangle = \int_{\mathbb{R}} dx\,\frac{e^{-ikx}}{\sqrt{2\pi}}\,|x\rangle\,.$$

Case 3. Finally, by using the the **momentum representation** of the identity opertor in the **energy representation** of $|\psi\rangle$, one finds

$$|\psi\rangle = \sum_n c_n\,|n\rangle = \int_{\mathbb{R}} dk\,|k\rangle\langle k| \sum_n c_n\,|n\rangle = \int_{\mathbb{R}} dk \sum_n c_n\,\langle k|n\rangle|k\rangle \equiv \int_{\mathbb{R}} dk\, c(k)\,|k\rangle \tag{89}$$

implying that

$$c(k) = \sum_n c_n\,\langle k|n\rangle \quad \Rightarrow \quad \langle k|n\rangle = \psi_n(k)$$

To make more evident the meaning of the latter one can exploit once more $\int_{\mathbb{R}} dx\,|x\rangle\langle x|$

$$\langle k|n\rangle = \psi_n(k) = \langle k|\left(\int_{\mathbb{R}} dx\,|x\rangle\langle x|\right)|n\rangle = \int_{\mathbb{R}} dx\,\langle k|x\rangle\langle x|n\rangle = \int_{\mathbb{R}} dx\,\frac{e^{ikx}}{\sqrt{2\pi}}\,\psi_n^*(x) = (\phi_k, \psi_n)$$

showing that $\langle k|n\rangle = \psi_n(k)$ is the (complex conjugate of the) Fourier transform of the energy eigenstates.

$$\langle n|k\rangle = S_{kn} = \psi_n^*(k)\,, \quad \langle k|n\rangle = S_{nk}^+ = \psi_n(k)\,.$$

11.3 Applications

Thanks to the Dirac (vector) notation for physical states one has

$$|\psi\rangle = \sum_m c_m |m\rangle\,,\;\; c_m = (\phi_m, \psi)\,, \qquad \tilde{A} = \sum_m \sum_j |m\rangle\langle j| A_{mj}\,,\;\; A_{mj} = (\phi_m, A\phi_j)$$

where $\psi(x)$ and A represent a generic state and a generic operator, respectively, in the coordinate picture.

Reformulation of the Schrödinger equation in terms of vector components

$$i\hbar\partial_t|\psi\rangle = H\,|\psi\rangle \quad \Rightarrow \quad i\hbar\sum_m(\partial_t c_m)|m\rangle = \sum_m c_m\,H\,|m\rangle \quad \Rightarrow$$

Vector $|\psi\rangle$ is time dependent, namely, its components $c_m = c_m(t)$ are time dependent.

$$i\hbar\langle n|\left(\sum_m \dot{c}_m|m\rangle\right) = \langle n|\sum_m c_m\,H\,|m\rangle \quad \Rightarrow \quad i\hbar\sum_m \dot{c}_m\langle n|m\rangle = \sum_m c_m\,\langle n|H\,|m\rangle$$

$$\Rightarrow \quad i\hbar\dot{c}_n = \sum_m c_m\,H_{nm}$$

This demonstrates that any Schrödinger problem is always expressible as a linear system of equations. If basis $\{|m\rangle\}$ is the **basis of energy eigenstates** then $H|m\rangle = E_m|m\rangle$ then

$$i\hbar\dot{c}_n = \sum_m c_m\,H_{nm}\,,\;\; H_{nm} = \langle n|H|m\rangle = \delta_{nm}\,E_m \quad \Rightarrow \quad i\hbar\dot{c}_n = \,E_n\,c_n$$

entailing that $c_m = c_m(0)e^{-itE_m/\hbar}$

$$|\psi\rangle = \sum_m c_m(t)\,|m\rangle = \sum_m c_m(0)\,e^{-itE_m/\hbar}\,|m\rangle\,.$$

Coordinate representation. Position and momentum operators (symbols x, ξ and η represents coordinates)

$$\tilde{x} = \int d\eta|\eta\rangle\langle\eta|\tilde{x}\int d\xi\,|\xi\rangle\langle\xi| = \int d\xi\int d\eta|\eta\rangle x_{\eta\xi}\langle\xi|\,,$$

$$x_{\eta\xi} = \langle\xi|\tilde{x}|\eta\rangle = (\phi_\xi(x), \hat{x}\,\phi_\eta(x)) = \eta\,\delta(\eta-\xi)\,.$$

An equivalent, simpler way to derive the latter equation is $\langle\xi|\tilde{x}|\eta\rangle = \eta\langle\xi|\eta\rangle = \eta\,\delta(\eta-\xi)$ owing to the fact that $|\eta\rangle$ is an eigenstate of $\tilde{x}$ and therefore $\tilde{x}|\eta\rangle = \eta|\eta\rangle$.

$$\tilde{p} = \int d\eta\,|\eta\rangle\langle\eta|\,\tilde{p}\int d\xi\,|\eta\rangle\langle\xi| = \int d\xi\int d\eta|\eta\rangle p_{\eta\xi}\langle\xi|\,,$$

$$p_{\eta\xi} = \langle\xi|\tilde{p}|\eta\rangle = (\phi_\xi(x), \hat{p}\,\phi_\eta(x)) = \int dx \phi^*_\xi(x)(-i\hbar\partial_x)\phi_\eta(x)$$

$$= \int dx \delta(x-\xi)\big[\,(-i\hbar\partial_x)\delta(x-\eta)\,\big] = (-i\hbar\partial_\xi)\,\delta(\xi-\eta)$$

Derivation of the Schrödinger equation in the coordinate representation

$$i\hbar\partial_t|\psi\rangle = H\,|\psi\rangle \quad \Rightarrow \quad i\hbar\partial_t \int d\xi\psi(\xi)|\xi\rangle = H\int d\xi\,\psi(\xi)|\xi\rangle \quad \Rightarrow$$

$$i\hbar\int d\xi\dot{\psi}(\xi)\,|\xi\rangle = \int d\xi\,\psi(\xi)H|\xi\rangle \quad \Rightarrow \quad i\hbar\,\langle x|\int d\xi\,\dot{\psi}(\xi)|\xi\rangle = \langle x|\int d\xi\psi(\xi)H|\xi\rangle$$

The equation has been projected on state $\langle x|$. Then

$$i\hbar\int dx\,\dot{\psi}(\xi)\langle x|\xi\rangle = \langle x|\int d\xi\,\psi(\xi)H|\xi\rangle \;\Rightarrow\; i\hbar\int dx\,\dot{\psi}(\xi)\,\delta(x-\xi) = \int d\xi\,\psi(\xi)\,\langle x|H|\xi\rangle$$

$$i\hbar\,\dot{\psi}(x) = \int d\xi\psi(\xi)\langle x|H|\xi\rangle\,, \qquad \text{with } \langle x|H|\xi\rangle = \frac{1}{2m}\langle x|\tilde{p}^2|\xi\rangle + \langle x|V(\tilde{x})|\xi\rangle$$

$$\langle x|\tilde{p}^2|\xi\rangle = (\phi_x, \hat{p}^2\phi_\xi) = \int d\eta\,\phi^*_x(\eta)\Big(-\hbar^2\partial^2_\eta\Big)\phi^*_\xi(\eta) = -\hbar^2\partial^2_x\,\delta(x-\xi)\,,$$

$$\langle x|V(\tilde{x})|\xi\rangle = \langle x|V(\xi)|\xi\rangle = \langle x|\xi\rangle V(\xi) = V(x)\delta(x-\xi)\,.$$

It follows that

$$i\hbar\,\dot{\psi}(x) = \int d\xi\psi(\xi)\left[\frac{-\hbar^2}{2m}\,\partial^2_x\,\delta(x-\xi)\;+\;V(x)\delta(x-\xi)\right]$$

$$i\hbar\,\dot{\psi}(x) = \frac{-\hbar^2}{2m}\,\partial^2_x\int d\xi\psi(\xi)\,\delta(x-\xi) + V(x)\int d\xi\psi(\xi)\,\delta(x-\xi)$$

giving (recall that $\psi(x) = \psi(x,t)$: time dependence has been implied to simplify the previous calculations)

$$i\hbar\,\dot{\psi}(x,t) = \frac{-\hbar^2}{2m}\,\partial^2_x\psi(x,t)\;+V(x)\psi(x,t) = \Big[\,\frac{\hat{p}^2}{2m}\;+V(x)\,\Big]\,\psi(x,t)\,. \qquad QED$$

This shows that the Dirac formalism is **totally equivalent** to the coordinate representation of Quantum Mechanics based on the wave-functions picture of quantum states.

Momentum representation.

$$\tilde{x} = \int q|q\rangle\langle q|\tilde{x} \int dk\, |k\rangle\langle k| = \int dq \int dk |q\rangle x_{qk} \langle k|\,,$$

$$x_{qk} = \langle q|\tilde{x}|k\rangle = (\phi_q(x), \hat{x}\,\phi_k(x)) = \int dx\, \frac{e^{ix(k-q)}}{2\pi}\, x = i\partial_q\, \delta(q-k)$$

$$\tilde{p} = \int q|q\rangle\langle q|\tilde{p} \int dk\, |k\rangle\langle k| = \int dq \int dk |q\rangle p_{qk} \langle k|\,,$$

$$p_{qk} = \langle q|\tilde{p}|k\rangle = (\phi_q(x), \hat{p}\,\phi_k(x)) = \int dx \phi_q^*(x)(-i\hbar\partial_x)\phi_k(x)$$

$$= \hbar k \int dx \phi_q^*(x)(-i\hbar\partial_x)\phi_k(x) = \hbar k\, \delta(q-k)$$

Of course, on can consider the simpler derivation based on $\tilde{p}|k\rangle = k|k\rangle$.

Reformulation of the Schrödinger equation in the momentum representation. Exercise: show that

$$i\hbar\partial_t|\psi\rangle = H\,|\psi\rangle \quad \Leftrightarrow \quad i\hbar\,\dot{\Phi}(k,t) = \Big[\,\frac{-\hbar^2\,k^2}{2m} + V(\hat{x})\Big]\,\Phi(k,t)\,.$$

where

$$\langle k|\psi\rangle = c_k(t) = \Phi(k,t)\,, \qquad \hat{x} = +i\hbar\partial_k\,.$$

11.4 The Schrödinger and Heisenberg Representations

These two representations describe two independent and, in a sense, complementary approaches for investigating the time evolution of quantum-mechanical systems.

Schrödinger representation. Within this representation the time evolution of a system is encoded in quantum states which obey the Schrödinger equation

$$i\hbar\partial_t\psi(\vec{x},t) \;=\; H\,\psi(\vec{x},t)\,, \qquad \psi(\vec{x},t) = \;e^{-itH/\hbar}\,\psi(\vec{x},0)$$

the latter formula representing the so-called formal solution discussed in subsection 4.4 within the wave-function picture of Quantum Mechanics. In this representation **physical operators are time independent** (unless they feature time dependence by definition) whereas the correponding expectation values are time dependent since

$$\langle A\rangle_t = (\psi, A\psi) \equiv \sum_r c_r^*(t)(\phi_r, A\sum_s c_s(t)\phi_s) = \sum_r\sum_s c_r^*(t)c_s(t)(\phi_r, A\phi_s)$$

if $\psi(\vec{x},t)=\sum_s c_s(t)\phi_s(\vec{x})$ in some basis $\mathcal{B}=\{\phi_s(\vec{x}), s=0,1,2,...\,\infty\}$. If the basis coincides with the energy basis, namely, if $\phi_s(\vec{x})$ satisfies the eigenvalue equation $H\phi_s=E_s\phi_s$, then $\langle A\rangle_t$ takes the form

$$\langle A\rangle_t=\sum_r\sum_s c_r^*(t)c_s(t)(\phi_r,A\phi_s)=\sum_r\sum_s c_r^*(0)c_s(0)e^{it(E_r-E_s)/\hbar}A_{r,s},\qquad A_{r,s}:=(\phi_r,A\phi_s)$$

where $A_{r,s}$ is a generic matrix element in the matrix-like representation of operators within the Dirac picture. It is important to note how ϕ_s and ϕ_r are the (space-dependent part of the) stationary states solving the Schrödinger problem

$$\phi_s(\vec{x},t)=e^{-itE_s/\hbar}\phi_s(\vec{x})$$

and $c_s(t)=c_s(0)e^{-itE_s/\hbar}$. For large t, the terms in $\langle A\rangle_t$ with $s\neq r$ are fast oscillating and thus can be neglected so that $\langle A\rangle_t$ becomes

$$\langle A\rangle_t\simeq\sum_s|c_s(0)|^2A_{s,s}\ .$$

Comment 1. In the presence of many-body systems the probability $|c_s(0)|^2$ can be assumed to describe the distribution of particle within the canonical ensemble: by setting $|c_s(0)|^2=e^{-\beta E_s}/Z$ (with Z representing the partion function and $\beta=1/(kT)$, T= temperature) one recovers the well-known formula of quantum Statistical Physics

$$\langle A\rangle_t\simeq\sum_s\frac{e^{-\beta E_s}}{Z}A_{s,s}=\sum_s\frac{e^{-\beta E_s}}{Z}(\phi_s,A\phi_s)=\sum_s(\rho\phi_s,A\phi_s)=Tr(\rho A)$$

where $\rho=e^{-\beta H}/Z$ is the density operator and $Tr(\rho\ A)=\sum_s(\phi_s,\rho A\phi_s)$ is the trace operator.

Comment 2. As shown in the previous subsection, the time evolution of a state $\psi(\vec{x},t)=\sum_s c_s(t)\phi_s(\vec{x})$ (represented in a generic basis –for example the plane-wave basis– different from the energy basis) is equivalent to solving a linear system. By projecting the Schrödinger equation $i\hbar\partial_t\psi(\vec{x},t)=H\psi(\vec{x},t)$ on a generic state $\phi_k(\vec{x},t)$, namely, by writing

$$(\phi_k,i\hbar\partial_t\psi)=(\phi_k,H\psi)$$

and observing that $\partial_t\psi=\sum_s\dot{c}_s(t)\phi_s$ one finds

$$i\hbar\sum_s\dot{c}_s(t)(\phi_k,\phi_s)=\sum_s\dot{c}_s(t)(\phi_k,H\phi_s)$$

giving in turn

$$i\hbar\dot{c}_k(t)=\sum_s\dot{c}_s(t)H_{k,s},\qquad H_{k,s}=(\phi_k,H\phi_s)$$

which represents a "simple" system of first-order linear equations whose solution can be achieved, in principle, by standard techniques. This is the Schrödinger equation expressed in terms of vector components already derived in subsection 11.3.

Heisenberg representation. Based on the equation for the expectation value $\langle A\rangle_t$ in the Schorödinger picture, one can define the Heisenberg representation A_t for the evolution of a given operator A as follows

$$\langle A\rangle_t = (\psi(\vec{x},t), A\psi(\vec{x},t)\rangle = (\psi_0 e^{-itH/\hbar}, Ae^{-itH/\hbar}\psi_0) = (\psi_0, e^{+itH/\hbar}Ae^{-itH/\hbar}\psi_0) = (\psi_0, A_t\psi_0)$$

where $\psi_0 = \psi(\vec{x},0)$ and the new operator

$$A_t = e^{+itH/\hbar}Ae^{-itH/\hbar}$$

has been introduced which shows how the time dependence relevant to the evolution of the system is now incorporated in the operator A_t. Operator A_t satisfies the so-called **Heisenberg equation**

$$\frac{d}{dt}A_t \;=\; \frac{1}{i\hbar}\,[A_t, H] + \frac{\partial}{\partial t}A_t$$

where the presence of the partial derivative take into account the fact that, in general, operator $A_t = A_t(t) = e^{+itH/\hbar}A(t)e^{-itH/\hbar}$ might feature an explict dependence on time. **Proof**:

$$\frac{d}{dt}A_t = \frac{d}{dt}e^{+itH/\hbar}A(t)e^{-itH/\hbar} = \left(\frac{d}{dt}e^{+itH/\hbar}\right)\,A(t)e^{-itH/\hbar} + e^{+itH/\hbar}\left(\frac{d}{dt}A(t)\right)e^{-itH/\hbar}$$

$$+e^{+itH/\hbar}A(t)\left(\frac{d}{dt}e^{-itH/\hbar}\right) = \frac{iH}{\hbar}\,e^{+itH/\hbar}A(t)e^{-itH/\hbar} + \frac{\partial}{\partial t}A_t - \,e^{+itH/\hbar}A(t)\,\frac{iH}{\hbar}e^{-itH/\hbar}$$

$$= -\,\frac{H}{i\hbar}\,A_t \,+\, A_t\,\frac{H}{i\hbar} \,+ \frac{\partial}{\partial t}A_t = \frac{1}{i\hbar}\,[\,A_t, H\,] + \frac{\partial}{\partial t}A_t\,. \qquad QED$$

Except for the extra factor $1/(i\hbar)$ this formula shows a quite evident similarity to the formula describing the time derivative of a function of canonical variables in terms of Poisson Brackets $\dot{f} = \{f, H\} + \partial_t f$ where once more $\partial_t f$ includes a possible explicit time dependence of f.

Suppose that $H = H(x,p)$, namely, Hamiltonian H is a function of canonical variables (for short we consider a 1-dimensional system with a single coordinate). Then

$$x_t = e^{+itH/\hbar}xe^{-itH/\hbar}\,, \quad p_t = e^{+itH/\hbar}pe^{-itH/\hbar} \quad H_t = e^{+itH/\hbar}H(x,p)e^{-itH/\hbar} = H(x_t,p_t)\,.$$

Nevertheless, owing to $[H,H]=0$,

$$H_t = e^{+itH/\hbar}He^{-itH/\hbar} = H\,e^{+itH/\hbar}e^{-itH/\hbar} = H$$

One can conclude that Hamiltonian H is unaffected by the time evolution of its constituent operators x_t and p_t, consistent with the fact that H must be a constant of motion, not only within Classical Mechanics, but also within the quantum formalism. More in general, one can conclude that, **conservation laws** characterizing Classical Mechanics can be extended to Quantum Mechanics because the definition of **constant of motion** can be applied to any time-dependent operator A_t (with no explicit time dependence $\partial_t A_t = 0$) such that

$$\frac{d}{dt}A_t = \frac{1}{i\hbar}\,[\,A_t, H\,] + \frac{\partial}{\partial t}A_t = 0\,, \quad \Leftrightarrow \quad [\,A_t, H\,] = \,0\,.$$

Note that $[\,A_t, H\,] = 0 \;\Leftrightarrow\; [\,A, H\,] = \,0$ in that $[\,A_t, H\,] = [U_t A\,U_t^+, H] = \,... = U_t[A,H]U_t^+ = 0$ where $U_t = \exp(-itH/\hbar)$.

11.5 The action of physical operators on quantum states

In general any Hermitian operator A can be used to generate a unitary transformation whose action on physical states has a definte, physically meanigfull, interpretation. A transformation of this form is obtained through the exponentiation of A

$$T(\phi) = e^{i\phi\, A} = \sum_{s=0}^{\infty} \frac{(i\phi)^s}{s!}\, A^s\,, \qquad T^+(\phi) = e^{(i\phi\, A)^+} = e^{-i\phi\, A^+} = e^{-i\phi\, A} = T(-\phi) = T^{-1}(\phi)$$

where the first formula represents the defining eqaution for the exponential operator generated by A depending on a real parameter ϕ and the second series of equalities makes evident the unitary character of $T(\phi)$. The latter is well visible due to the fact the corresponding inverse operator $T^{-1}(\phi)$ is its hermitian conjugate $T^+(\phi)$.

Translation operators. The simplest possible example is that involving the momentum operator $\hat{p} = -i\hbar\partial_x$

$$T(\lambda) = e^{i\lambda\, p/\hbar} = \sum_{s=0}^{\infty} \frac{(i\lambda)^s}{s!} \left(\frac{p}{\hbar}\right)^s = \sum_{s=0}^{\infty} \frac{(i\lambda)^s}{s!} \left(-i\frac{\partial}{\partial x}\right)^s = \sum_{s=0}^{\infty} \frac{\lambda^s}{s!}\, \frac{\partial^s}{\partial x^s}$$

Its application to some wave function $\psi(x)$ gives

$$T(\lambda)\psi(x) = e^{i\phi\, p/\hbar}\psi(x) = \sum_{s=0}^{\infty} \frac{\lambda^s}{s!}\, \frac{\partial^s \psi}{\partial x^s} = \psi(x+\lambda)\,.$$

The last equality directly follows from the Taylor-expansion formula representing a function at $x' = x + \lambda$

$$\psi(x') = \psi(x+\lambda) = \sum_{s=0}^{\infty} \frac{\lambda^s}{s!}\, \frac{\partial^s \psi}{\partial x^s}$$

where λ is a sufficiently small deviations within a neighbourhood of the base point x whose size depends on the analytical properties of function ψ. Then $T(\lambda)$ is a **translation operator** (or displacement operator) whose action is given by $T(\lambda)\psi(x) = \psi(x+\lambda)$. Its generaliztion to the 3-dimensional space is easily obtained

$$T\big(\vec{\lambda}\big)\, \psi(x_1, x_2, x_3) = e^{i\,\vec{\lambda}\cdot\vec{p}/\hbar}\psi(x_1, x_2, x_3) = e^{i\,\lambda_1 p_1/\hbar}\, e^{i\,\lambda_2 p_2/\hbar}\, e^{i\,\lambda_3 p_3/\hbar}\psi(x_1, x_2, x_3)$$

$$= e^{\lambda_1 \frac{\partial}{\partial x_1}} e^{\lambda_2 \frac{\partial}{\partial x_2}} e^{\lambda_3 \frac{\partial}{\partial x_3}}\, \psi(x_1, x_2, x_3) = e^{\lambda_2 \frac{\partial}{\partial x_2}} e^{\lambda_3 \frac{\partial}{\partial x_3}}\, \psi(x_1+\lambda_1, x_2, x_3)$$
$$= \,...\, = \psi\big(x_1+\lambda_1,\, x_2+\lambda_2,\, x_3+\lambda_3\big)$$

Rotation operators. Another interesting example is that involving one of the three component of the angular-momentum operator $\vec{L} = \vec{x} \wedge \vec{p}$. Let us consider the component L_3

$$L_3 = -i\hbar\big(x_1 \frac{\partial}{\partial x_2} - x_2 \frac{\partial}{\partial x_1}\big) \quad \Rightarrow \quad R(\theta) = e^{i\,\theta\, L_3/\hbar} = e^{\theta\, \left(x_1 \frac{\partial}{\partial x_2} - x_2 \frac{\partial}{\partial x_1}\right)}$$

To understand the effect of the action of $R(\theta)$ it is sufficient to consider an infinitesimal rotation $(0 < \theta << \pi)$

$$R(\theta)\,\psi(x_1,x_2,x_3) = e^{i\,\theta\,L_3/\hbar}\,\psi(x_1,x_2,x_3) = \left(\mathbb{I} \,+\, i\,\theta\,\frac{L_3}{\hbar} \,+\, ...\right)\psi(x_1,x_2,x_3)$$

$$\simeq \left(\mathbb{I} \,+\, \theta\,\left(x_1\frac{\partial}{\partial x_2} \,-\, x_2\frac{\partial}{\partial x_1}\right)\right)\psi(x_1,x_2,x_3) = \psi(x_1,x_2,x_3) + \theta\,\left(x_1\frac{\partial\psi}{\partial x_2} \,-\, x_2\frac{\partial\psi}{\partial x_1}\right).$$

The formula defining a coordinate change due to a rotation about the axis x_3 is the planar vector transformation

$$\vec{x}' \;=\; \left[\begin{array}{cc} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{array}\right]\,\vec{x}$$

which, written in components, reads

$$x_1' = x_1\,\cos\theta \,-\, x_2\,\sin\theta\,, \qquad x_2' = x_1\,\sin\theta \,+\, x_2\,\cos\theta\,.$$

and, if θ is "small", reduces to

$$x_1' \simeq x_1 \,-\, x_2\,\theta\,, \qquad x_2' \simeq x_2 \,+\, x_1\,\theta\,, \qquad \cos\theta \simeq 1,\ \sin\theta \simeq \theta\,.$$

A function ψ depending on the rotated vector $\vec{x}'$ (namely, on new coordinates x_1', x_2') can written as

$$\psi\big(\vec{x}'\big) = \psi\big(x_1',\,x_2',\,x_3\big) = \psi\big(x_1\,\cos\theta \,-\, x_2\,\sin\theta,\,x_1\,\sin\theta \,+\, x_2\,\cos\theta,\,x_3\big)\,.$$

If θ is infinitesimal, one gets

$$\psi\big(x_1',\,x_2',\,x_3\big) = \psi\big(x_1 - \,\theta\,x_2,\,x_2 + \,\theta\,x_1,\,x_3\big) = \psi\big(x_1,\,x_2,\,x_3\big) + (-\,\theta\,x_2)\,\frac{\partial\psi}{\partial x_1} + (\theta\,x_1)\,\frac{\partial\psi}{\partial x_2}$$

where the latter expression is the Taylor expansion with respect to the infinitesimal displacements $-\,\theta\,x_2$ and $+\theta\,x_1$. The comparison of the latter formula with the previous formula for $R(\theta)\,\psi(x_1,x_2,x_3)$ shows that

$$R(\theta)\,\psi(x_1,x_2,x_3) \equiv\; \psi\big(x_1 - \,\theta\,x_2,\,x_2 + \,\theta\,x_1,\,x_3\big)\,.$$

In the case when θ is a finite angle, the latter becomes

$$R(\theta)\,\psi(x_1,x_2,x_3) \equiv\; \psi\big(x_1\,\cos\theta \,-\, x_2\,\sin\theta,\,x_1\,\sin\theta \,+\, x_2\,\cos\theta,\,x_3\big)\,.$$

The action of $R(\theta)$ generated by L_3 is thus **completely defined** and amounts to effecting a coordinate transformation in the plane (x_1,x_2) representing to a **rotation**. No calculation is necessary to obtain the formulas for the rotation $R(\alpha)$ (around axis x_1) and $R(\beta)$ (around axis x_2) generated by L_1 and L_2, respectively. These can be derived in a direct way from the latter formula. For example

$$R(\beta)\,\psi(x_1,x_2,x_3) = e^{i\,\beta\,L_2/\hbar}\,\psi(x_1,x_2,x_3) \equiv\; \psi\big(x_1\,\cos\beta \,+\, x_3\,\sin\beta,\,x_2\,,\,x_3\,\cos\beta - \,x_1\,\sin\beta\,\big)\,.$$

The most general rotation can be show to be generated by the combined action of L_1, L_2 and L_3 through the formula

$$R(\alpha,\beta,\theta) = \exp\left[\, i\,(\,\alpha\, L_1\hbar \,+\, \beta\, L_2\hbar \,+\, \theta\, L_3\hbar\,)\,\right] .$$

Time-propagation operator. The formal solution of a Schrödinger problem described by equation (21) provides another significant example

$$\psi(\vec{x},t) = U(t)\,\psi(\vec{x},0) = e^{-it\,H/\hbar}\,\psi(\vec{x},0)\,.$$

Here the unitary operator $U(t)$ is generated by the (hermitian) Hamilton operator H. It describes the time propagation of the initial state $\psi(\vec{x},0)$. If t is small

$$\psi(\vec{x},t) \simeq \left(\mathbb{I} \,-\, i\,t\,\frac{H}{\hbar} \,+\, ...\right)\psi(\vec{x},0)$$

Differently from the case of rotations where the action of L_k on wave functions can be easily calculated, the action of H on $\psi(\vec{x},0)$ is in general complicated because of the complex (at least quadratic) structure of H in terms of operators x_j and p_n. This scheme can be easily extended to any time $t' = t + \Delta t$. At time $t + \Delta t$ the propagated state $\psi(\vec{x}, t + \Delta t)$ can be expressed both by means of the propagation operator

$$\psi(\vec{x},t+\Delta t) = U(t+\Delta t)\psi(\vec{x},0) = e^{-i(t+\Delta t)\,H/\hbar}\psi(\vec{x},0) = e^{-i\Delta t\;H/\hbar}\left(e^{-it)\;H/\hbar}\psi(\vec{x},0)\right)$$

$$\simeq \left(\mathbb{I} \,-\, i\Delta t\,\frac{H}{\hbar} \,+\, ...\right)\psi(\vec{x},t)$$

and by means of the series expansion

$$\psi(\vec{x},t+\Delta t) = \psi(\vec{x},t) + \frac{\partial\psi(\vec{x},t)}{\partial t}\;\Delta t + \;...\;.$$

Then, by comparing such expressions, one has

$$\psi(\vec{x},t) - i\Delta t\,\frac{H}{\hbar}\;\psi(\vec{x},t) + \;... = \psi(\vec{x},t) + \frac{\partial\psi}{\partial t}\;\Delta t + \;...\;,$$

namely, the Schorödinger equation

$$-\frac{i}{\hbar}H\psi(\vec{x},t) = \frac{\partial\psi(\vec{x},t)}{\partial t}$$

Exercise. Determine the action of the propagator generated by the free-particle Hamiltonian

$$H = \frac{p^2}{2m} = -\frac{\hbar^2}{2m}\partial_x^2 \quad \Rightarrow \quad \psi(\vec{x},t) = U(t)\,\psi(\vec{x},0) = \exp\left(-i\,\frac{\hbar\,t}{2m}\,\frac{\partial^2}{\partial x^2}\right)\,\psi(\vec{x},0)\,.$$

12 The spin operator

The Stern-Gerlach experiment (1922) showed that a beam of atoms having **zero total angular momentum** is splitted into two beams due to the presence of a **non uniform magnetic field** $\vec{B}$. Silver atoms have this property in that, except for a single external electron, the electronic cloud features spherical symmetry and thus $\vec{L}_{tot} = 0$. Since the external electron is in the state 5s ($\ell = 0$) also its contribution to the total angular momentum is zero. To see this consider the effective Hamiltonian describing the silver atom

$$H = \frac{\vec{p}^{\,2}}{2\,m} - \vec{\mu} \cdot \vec{B}(\vec{x})\,, \quad \vec{\mu} = -\frac{e}{2m_e c}\,\vec{L}$$

where $\vec{\mu}$ is the atom magnetic moment of a hydrogen-like atom containing a single electron (its derivation is discussed in section 13). Silver atoms are well represented by this simple model owing to their electron-cloud structure. The **splitting effect** was somewhat surprising in that the force due to the coupling of $\vec{\mu}$ with a space-dependent field $\vec{B}(\vec{x})$

$$\vec{F} = -\nabla\big(-\vec{\mu} \cdot \vec{B}(\vec{x})\big) = \mu_3\,\frac{\partial B_3}{\partial x_3}\,\vec{u}_3\,,$$

was expected to be zero. In fact, quantum-mechanically, the condition that $\ell = 0$ implies that

$$\mu_3 = -\frac{e}{2m_e c}\,L_3 \;\propto\; \hbar\, m = 0\,,$$

$m = 0$ being the unique eigenvalue that operator L_3 can assume for states corresponding to type-s orbitals of hydrogen-like atoms.

This circumstance allowed to understand that, in addition to the well-known magnetic moment relevant to angular momentum $\vec{L}$, the electron is equipped with an **intrinsic angular momentum**, the spin $\vec{S}$, relevant to the electron rotation about its own axis. Spin operator $\vec{S}$ must be **added** to $\vec{L}$ in $\vec{\mu}$. Since the spin operator features two physical states (corresponding to two eigenvalues $\pm\hbar/2$ of S_3, the third component of $\vec{S}$) this was sufficient to supply a convincing explanation of the two-beam splitting effect detected experimentally. Many other experiments involving the observation of atomic transition (such as the anomalous Zeeman effect or the study of He atom structure) confirmed that the inclusion of spin degrees of freedom is an essential ingredient to construct realistic theoretical models.

The study of atomic systems and, more in general, of elementary particles led to formulate the **spin-statistic theorem** involving the fact that elementary particles are of two types: fermions (possessing half-integer spin such as, for example, electrons, protons and neutrons) and bosons (possessing integer/zero spin such as, e. g., photons or composite particles formed by an even number of fermions). This aspect will be discussed in a separate section.

12.1 Matrix representation of spin operators

From a formal point of view, spin operators essentially possess the same vector and algebraic properties of angular momentum $\vec{L}$. The main difference concerns the way their operator character is realized. This difference will disappear within the Dirac picture of Quantum Mechanics. While the vector components of $\vec{L}$ are represented by differential-like operators $L_k = -i\hbar\epsilon_{kmn}\, x_m\, \partial_n$ (see section 9.2), the vector components of $\vec{S}$ are matrices

$$S_3 = \frac{\hbar}{2}\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} = \frac{\hbar}{2}\,\sigma_3\,, \quad S_1 = \frac{\hbar}{2}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \frac{\hbar}{2}\,\sigma_1\,, \quad S_2 = \frac{\hbar}{2i}\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} = \frac{\hbar}{2}\,\sigma_2\,. \tag{90}$$

Matrices σ_1, σ_2 and σ_3, implicitly defined by the previous equations, are called Pauli spin matrices. It is a simple exercise to prove that spin components S_k fulfill the following commutators

$$[S_1, S_2] = i\hbar S_3\,, \quad [S_2, S_3] = i\hbar S_1\,, \quad [S_3, S_1] = i\hbar S_2\,,$$

while

$$S_- := S_1 - iS_2 = \hbar\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \quad S_+ := S_1 + iS_2 = \hbar\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix},$$

entailing

$$S_1 = \frac{S_+ + S_-}{2}\,, \qquad S_2 = \frac{S_+ - S_-}{2i}\,,$$

fulfill commutators

$$[S_3, S_+] = +\hbar S_+\,, \quad [S_3, S_-] = -\hbar S_-\,.$$

Apart from the substitution of symbol L with S, such formulas exactly reproduce the commutation relations of angular momentum $[L_m, L_r] = i\hbar\epsilon_{mrq}L_q$ and $[L_3, L_\pm] = \pm\hbar\, L_\pm$. Operator $\vec{S}^2$ exhibits the same properties of $\vec{L}^2$

$$\vec{S}^{\,2} = S^2 = S_1^2 + S_2^2 + S_3^2\,, \qquad [S^2, S_1] = [S^2, S_2] = [S^2, S_3] = 0$$

showing that S^2 and S_m with $m = 1, 2, 3$ can be observed simultaneously. The **basis of the eigenstates** of S_3 is the standard choice for representing the vector space of **physical states**. Such eigenstates satisfy the eigenvalue equation

$$S_3\,|s, m_s\rangle = \hbar m_s\,|s, m_s\rangle\,, \quad m_s = \pm 1/2\,, \qquad |s, +1/2\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad |s, -1/2\rangle = \begin{bmatrix} 0 \\ 1 \end{bmatrix}. \tag{91}$$

Here index s plays the same role of label ℓ relevant to $\vec{L}$

$$\vec{S}^{\,2}\,|s, m_s\rangle = \hbar^2\, s(s+1)\,|s, m_s\rangle$$

Proof.

$$S_1^2 = S_2^2 = S_3^2 = \frac{\hbar^2}{4}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \frac{\hbar^2}{4}\,\mathbb{I} \quad \Rightarrow \quad \vec{S}^{\,2} = S_1^2 + S_2^2 + S_3^2 = \frac{3}{4}\,\hbar^2\,\mathbb{I}$$

$$\vec{S}^{\,2}\,|s,m_s\rangle = \frac{3}{4}\,\hbar^2\,|s,m_s\rangle \quad \Rightarrow \quad s(s+1) = \frac{3}{4} \quad \Rightarrow \quad s = \frac{1}{2}\,.$$

Then, for any vector $|s,m_s\rangle$, the (representation) index s **takes the value** $s = \frac{1}{2}$. For this reason spin $\vec{S}$ is called **one-half spin** or $s = 1/2$ spin.

Action of $S_\pm$ on the basis eigenstates. For short, we set $|s,\pm 1/2\rangle = |\pm 1/2\rangle$ since s assumes a single value.

$$S_+|+1/2\rangle = 0\,, \quad S_+|-1/2\rangle = \hbar\,|+1/2\rangle\,, \tag{92}$$

$$S_-|+1/2\rangle = \hbar\,|-1/2\rangle\,, \quad S_-|-1/2\rangle = 0\,. \tag{93}$$

This property makes evident the reason why $S_\pm$ are called **raising/lowering operators**.

Hamiltonians including spin-dependent terms feature physical states called **spinor states**

$$|\Psi\rangle = \sum_{m_s=\pm 1/2} \psi_{m_s}(\vec{x})\,|m_s\rangle = \psi_-(\vec{x})\,|\,{}^{-1}\!/_2\rangle + \psi_+(\vec{x})\,|\,{}^{+1}\!/_2\rangle$$

$$= \psi_-(\vec{x}) \begin{bmatrix} 0 \\ 1 \end{bmatrix} + \psi_+(\vec{x}) \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \psi_+(\vec{x}) \\ \psi_-(\vec{x}) \end{bmatrix} \tag{94}$$

Given, for example, the Hamiltonian

$$H = H_0 + H_\alpha = \frac{\vec{p}^2}{2m} + V(\vec{x}) + \vec{\alpha}\cdot(\vec{L} + g\vec{S})\,,$$

its Schrödinger equation takes the form

$$i\hbar\,\frac{\partial}{\partial t}|\Psi\rangle = H\,|\Psi\rangle \;\rightarrow\; i\hbar\,\frac{\partial}{\partial t}\begin{bmatrix} \psi_+ \\ \psi_- \end{bmatrix} = H \begin{bmatrix} \psi_+ \\ \psi_- \end{bmatrix}$$

which becames

$$i\hbar\,\frac{\partial}{\partial t}\begin{bmatrix} \psi_+ \\ \psi_- \end{bmatrix} = \begin{bmatrix} (H_0 + \vec{\alpha}\cdot\vec{L})\,\psi_+ \\ (H_0 + \vec{\alpha}\cdot\vec{L})\,\psi_- \end{bmatrix} + \frac{g\,\hbar}{2}\left[\begin{bmatrix} 0 & \alpha_1 - i\alpha_2 \\ \alpha_1 + i\alpha_2 & 0 \end{bmatrix} + \begin{bmatrix} \alpha_3 & 0 \\ 0 & -\alpha_3 \end{bmatrix}\right]\begin{bmatrix} \psi_+ \\ \psi_- \end{bmatrix}$$

giving in turn **two coupled Schrödinger equations** (the presence of $\vec{S}$ causes the coupling)

$$i\hbar\,\frac{\partial}{\partial t}\psi_+ = \left(H_0 + \vec{\alpha}\cdot\vec{L} + \frac{g\,\hbar}{2}\alpha_3\right)\psi_+ + \frac{g\,\hbar}{2}(\alpha_1 - i\alpha_2)\,\psi_-\,,$$

$$i\hbar\,\frac{\partial}{\partial t}\psi_- = \left(H_0 + \vec{\alpha}\cdot\vec{L} - \frac{g\,\hbar}{2}\alpha_3\right)\psi_- + \frac{g\,\hbar}{2}(\alpha_1 + i\alpha_2)\,\psi_+\,.$$

With $\vec{\alpha} = e\vec{B}/2m_e c$, where $\vec{B}$ is the magnetic field, and $g = 2$ one finds the well-known **Pauli equation** representing the nonrelativistic limit of **Dirac's equation**

$$i\hbar \frac{\partial}{\partial t} \begin{bmatrix} \psi_+ \\ \psi_- \end{bmatrix} = \left(-\frac{\hbar^2}{2m_e}\nabla^2 + V(\vec{x}) + \frac{e}{2m_e c}\,\vec{B}\cdot\vec{L} + \frac{e}{m_e c}\vec{B}\cdot\vec{S} \right) \begin{bmatrix} \psi_+ \\ \psi_- \end{bmatrix} .$$

This describes a free electron under the action of a constant magnetic field. Its derivation, in the absence of spin $\vec{S}$, is discussed in the subsequent section.

The Jaynes-Cummings (JC) model. The spin formalism is often used to model systems whose spectrum features a small number of energy levels. In this case, the spin does not represent a specific property of particles belonging to the system. For esample, matrices such as S_3 and $S_\pm$ can be exploited to represent microscopic processes characterized by two quantum states and by the corresponding energy eigenvalues.

The JC model describes the interaction between matter and radiation. The latter consists of an electromagnetic wave with a definite wave vector $\vec{k}$ whose energy is typically represented by the term $H_{em} = \hbar w a^+ a$ ($w = c|\vec{k}|$ is the light frequency) that can be derived from the general formula for the electromagnetic-field energy. The matter is represented by atoms in which an electron can be excited from a level with energy $E_1 = E_0 - \hbar w/2$ to a level with a larger energy $E_2 = E_0 + \hbar w/2$. This system simply involves two energy levels and its Hamiltonian is then represented by the 2x2 matrix $H_{at} = E_0 \mathbb{I}_2 + w\, S_3$. The total Hamiltonian reads

$$H = H_{em} + H_{at} + H_I = \hbar w a^+ a + E_0 \mathbb{I}_2 + w\, S_3 + \gamma\Big(a^+ S_- + S_+ a\Big)$$

in which the γ-dependent term describe the interaction and the constant term E_0 can be ignored. The physical states for this systems are $|n, m\rangle = \psi_n(x)|\frac{1}{2}, m\rangle$ which diagonalize both H_{em} and H_{at}

$$H_{em}|n,m\rangle = \big(\hbar w a^+ a\ \psi_n(x)\big)\ |\tfrac{1}{2},m\rangle = \hbar w n\ \psi_n(x)\ |\tfrac{1}{2},m\rangle = \hbar w n\ |n,m\rangle\,,$$

$$H_{at}|n,m\rangle = \psi_n(x)\ \big(w S_3|\tfrac{1}{2},m\rangle\big) = \hbar w \frac{m}{2}\ \psi_n(x)\ |\tfrac{1}{2},m\rangle = \hbar w \frac{m}{2}\ |n,m\rangle\,.$$

One should recall, in this model, that quantum number n describes the number of photons of the field. The **interaction term** contains terms $a^+ S_-$ and $S_+ a$ describing the processes

$$\begin{aligned} a^+ S_- |n, +\tfrac{1}{2}\rangle = a^+ S_- \psi_n(x)|+\tfrac{1}{2}\rangle = a^+ \psi_n(x) \times S_-|+\tfrac{1}{2}\rangle = \hbar\sqrt{n+1}\psi_{n+1}(x)|-\tfrac{1}{2}\rangle \\ = \hbar\sqrt{n+1}|n+1,-\tfrac{1}{2}\rangle\,, \\ a\ S_+|n+1,-\tfrac{1}{2}\rangle = a S_+ \psi_{n+1}(x)|+\tfrac{1}{2}\rangle = a\psi_{n+1}(x) \times S_+|+\tfrac{1}{2}\rangle = \hbar\sqrt{n+1}\psi_n(x)|-\tfrac{1}{2}\rangle \\ = \hbar\sqrt{n+1}|n,+\tfrac{1}{2}\rangle\,. \end{aligned}$$

The first represents the transition $E_2 \to E_1$ from the excited to the unexcited atomic state ($|n, +\frac{1}{2}\rangle \to |n+1, -\frac{1}{2}\rangle$) involving the creation of a photon $n \to n+1$. The second represents the opposite process with the destruction of a photon ($n+1 \to n$) in parallel with the transition

$E_1 \to E_2$ from the unexcited to the excited atomic state ($|n+1,-\frac{1}{2}\rangle \to |n,+\frac{1}{2}\rangle$). It is a simple exercise to show that $\hbar w a^+ a + w S_3$ commutes with H and that states

$$|E_n(\pm 1)\rangle = \frac{1}{\sqrt{2}}\big(|n+1,-\tfrac{1}{2}\rangle \pm |n,+\tfrac{1}{2}\rangle\big)$$

are the eigenstates of H for each value $n = 0, 1, 2, ...$
Proof. In view of the previous formulas one has

$$(a^+ S_- + a S_+)|E_n(\pm 1)\rangle = (a^+ S_- + a S_+)\frac{1}{\sqrt{2}}\big(|n+1,-\tfrac{1}{2}\rangle \pm |n,+\tfrac{1}{2}\rangle\big)$$

$$= \frac{1}{\sqrt{2}}\Big(a S_+|n+1,-\tfrac{1}{2}\rangle \pm a^+ S_-|n,+\tfrac{1}{2}\rangle\Big) = \hbar\sqrt{n+1}\,\frac{1}{\sqrt{2}}\Big(|n,+\tfrac{1}{2}\rangle \pm |n+1,-\tfrac{1}{2}\rangle\Big)$$

$$= \pm\hbar\sqrt{n+1}\,\frac{1}{\sqrt{2}}\Big(|n+1,-\tfrac{1}{2}\rangle \pm |n,+\tfrac{1}{2}\rangle\Big) = \pm\hbar\sqrt{n+1}|E_n(\pm 1)\rangle$$

showing how $|E_n(\pm 1)\rangle$ is an eigenstate of $a^+ S_- + a S_+$. On the other hand, the same property holds for $\hbar w a^+ a + w\ S_3$

$$(\hbar w a^+ a + w\ S_3)|E_n(\pm 1)\rangle = \frac{1}{\sqrt{2}}\Big((\hbar w a^+ a + w\ S_3)|n+1,-\tfrac{1}{2}\rangle \pm (\hbar w a^+ a + w\ S_3)|n,+\tfrac{1}{2}\rangle\Big)$$

$$= \frac{1}{\sqrt{2}}\Big((\hbar w(n+1) - \hbar w/2)|n+1,-\tfrac{1}{2}\rangle \pm (\hbar w n + \hbar w/2)|n,+\tfrac{1}{2}\rangle\Big)$$

$$= \Big(\hbar w n + \hbar w/2\Big)|E_n(\pm 1)\rangle$$

Then, by combining these results, one finds

$$H|E_n(\pm 1)\rangle = \Big[\hbar w a^+ a + w\ S_3 + \gamma(a^+ S_- + a S_+)\Big]|E_n(\pm 1)\rangle = E_n(\pm 1)|E_n(\pm 1)\rangle$$

providing the explicit form of the energy eigenvalues

$$E_n(\pm 1) = \hbar w(n + 1/2) \pm \gamma\sqrt{n+1}\ .$$

12.2 Addition of angular momenta

The subsequent discussion provides a simple, intuitive derivation of the vector basis associated to the **angular momentum obtained by adding vectors** such as $\vec{L}$ and $\vec{S}$. This is constructed by means of the two basis

$$\{|\ell, m\rangle : |m| \leq \ell\} \quad \text{and} \quad \{|s, m\rangle : |m| \leq s = 1/2\}\,.$$

We consider the two cases $\vec{J} = \vec{L} + \vec{S}$ and $\vec{J} = \vec{S}_1 + \vec{S}_2$ involving two one-half spins. In the first case, concerning operator $\vec{J} = \vec{L} + \vec{S}$, one easily checks that

$$[J_n, J_m] = [S_n + L_n, S_m + L_m] = [L_n, L_m] + [S_n, S_m] = \, i\hbar\,\epsilon_{mnk} L_k + \, i\hbar\,\epsilon_{mnk} S_k = \, i\hbar\,\epsilon_{mnk} J_k\,.$$

It is thus evident that $\vec{J}$, $\vec{L}$ and $\vec{S}$ feature the **same algebraic structure**. As a consequence

$$[J^2, J_k] = 0 \tag{95}$$

because components J_k satisfy the same commutators of S_k or L_k. In particular, $[J^2, J_3] = 0$ entails that j and m_j relevant to J^2 and J_3 provide a pair of good quantum numbers.

Range of the eigenvalues relevant to J^2 and J_3. We determine the range of j, the quantum numbers associated to J^2, in a heuristic way. Operators $\hat{S}^2$ and $\hat{L}^2$ are associated to the eigenvalues

$$\hbar^2 s(s+1) \leftrightarrow \hat{S}^2 = S_1^2 + S_2^2 + S_3^2\,, \qquad \hbar^2 \ell(\ell+1) \leftrightarrow \hat{L}^2 = L_1^2 + L_2^2 + L_3^2\,.$$

In the **semiclassical limit**, namely for $\ell, s >> 1$, one can identify $\hat{S}$ and $\hat{L}$ with vectors $\vec{S}$ and $\vec{L}$ so that

$$\hbar^2 s^2(1 + 1/s) \simeq \hbar^2 s^2 = \vec{S}^2\,, \qquad \hbar^2 \ell^2(1 + 1/\ell) \simeq \hbar^2 \ell^2 = \vec{L}^2.$$

Then, classically, if $\vec{L}$ and $\vec{S}$ have length $|\vec{L}| = \hbar\ell$ and $|\vec{S}| = \hbar s$ (with $s < \ell$) the new vector $\vec{J} = \vec{L} + \vec{S}$ has length ranging in the interval $\hbar(\ell - s) \leq |\vec{J}| \leq \hbar(\ell + s)$. By analogy, the **quantum counterpart** of $\vec{J}^2$ is characterized by the eigenvalue $\hbar^2 j(j+1)$ where j (corresponding to $|\vec{J}|$ in the semiclassical limit) will range in

$$(\ell - s) \leq j \leq (\ell + s).$$

This heuristic reasoning can be confirmed by exact theoretical calculations. Because of the properties of angular momentum, the third component $J_3 = S_3 + L_3$ of the new operator $\hat{J}$ features an eigenvalue m_j such that

$$-j \leq m_j \leq +j \qquad \text{with } \ell - s \leq j \leq \ell + s\,.$$

By considering the trial states $|\ell, m\rangle|s, m_s\rangle$, the explicit form of m_j can be evinced from

$$J_3\,|\ell, m\rangle|s, m_s\rangle = L_3\,|\ell, m\rangle|s, m_s\rangle + |\ell, m\rangle S_3|s, m_s\rangle = \hbar m\,|\ell, m\rangle|s, m_s\rangle + \hbar m_s|\ell, m\rangle|s, m_s\rangle$$

$$= \hbar m_j |\ell, m\rangle |s, m_s\rangle \,, \qquad \text{where } m_j = m + m_s \,.$$

and one should recall that

$$-s \leq m_s \leq +s \,, \qquad -\ell \leq m \leq +\ell \,.$$

After observing that tha highest value of m_j is j, one easily recovers the two extreme cases discussed above: the largest value $j = \ell + s$ corresponds to the maximal eigenvalue of L_3 when (classically) $\vec{L}$ and $\vec{S}$ in $\vec{J}$ are **parallel** vectors. Likewise, the lowest value $j = \ell - s$ corresponds to the maximal eigenvalue of L_3 when $\vec{L}$ and $\vec{S}$ are **antiparallel** vectors.

Commuting operators. In the initial scheme, operators L^2, S^2, L_3 and S_3 provide good quantum numbers in that they commute the one with the other. In the new scheme based on operator $\vec{J}$ one easily checks that in addition to $[J^2, J_3] = 0$ involving the pair of good quantum numbers j and m_j, owing to

$$J^2 = (\vec{S} + \vec{L})^2 = \vec{S}^2 + \vec{L}^2 + 2\vec{S} \cdot \vec{L}$$

one has

$$[J^2, S_3] = [\vec{S}^2 + \vec{L}^2 + 2\vec{S} \cdot \vec{L}, S_3] = [\vec{S}^2 + 2\vec{S} \cdot \vec{L}, S_3] = 2 \sum_j L_j [S_j, S_3] \neq 0 \,.$$

Likewise

$$[J^2, L_3] = \, ... \, = 2 \sum_j S_j [L_j, L_3] \neq 0 \,.$$

By combining such nonzero commutators, however, one obtains $[J^2, L_3] + [J^2, S_3] = \, ... \, = 0$ confirming the result $[J^2, J_3] = 0$ contained in formula (95).

Conversely,

$$[J^2, L^2] = [L^2 + S^2 + 2S \cdot L, L^2] = \sum_k S_k [L_k, L^2] = 0 \,, \quad [J^2, S^2] = ... \, = 0$$

and

$$[L^2, J_3] = [L^2, L_3 + S_3] = [L^2, L_3] = 0 \,, \quad [S^2, J_3] = [S^2, L_3 + S_3] = [S^2, S_3] = 0 \,.$$

Then, the commuting operators are

$$S^2 \,, \quad L^2 \,, \quad J^2 \,, \quad J_3 \,, \quad \Leftrightarrow \quad s, \; \ell, \; j, \; m_j$$

and the relevant eigenvalues form the set of good quantum numbers of the new scheme.

New basis $\{ \, |j, m_j, \ell, s\rangle \, \}$. An important observation concerning states $|\ell, m\rangle \, |s, m_s\rangle$ is that the same value m_j can be obtained both with $m + 1$, $m_s = -1/2$ and with m, $m_s = +1/2$. This degeneracy entails that the states of the basis relevant to J^2 and J_3 are represented by

$$|j, m_j, \ell, s\rangle = \sum\nolimits^*_{m, m_s} C(m, m_s) |\ell, m\rangle \, |s, m_s\rangle \,, \quad j = \ell \pm 1/2 \,, \quad s = 1/2 \,,$$

where the superscript * implies the constraint $m + m_s =$ constant. Note that such a state satisfies, by construction,

$$L^2|j, m_j, \ell, s\rangle = \hbar^2\ell(\ell+1)\,|j, m_j, \ell, s\rangle\,, \qquad S^2|j, m_j, \ell, s\rangle = \hbar^2 s(s+1)\,|j, m_j, \ell, s\rangle\,,$$

and

$$J_3|j, m_j, \ell, s\rangle = \hbar(m + m_s)\,|j, m_j, \ell, s\rangle\,.$$

The latter can be easily demonstrated thanks to the condition $m + m_s =$ constant

$$J_3|j, m_j, \ell, s\rangle = (S_3 + L_3)\,|j, m_j, \ell, s\rangle =$$

$$= \sum\nolimits^*_{m,m_s} C(m, m_s)(S_3 + L_3)|\ell, m\rangle\,|s, m_s\rangle = (m_s + m)\sum\nolimits^*_{m,m_s} C(m, m_s)|\ell, m\rangle\,|s, m_s\rangle\,.$$

The double series of states relevant to $j = \ell - 1/2$ and $j = \ell + 1/2$ can be shown to have the form

$$\begin{cases} |\ell - \frac{1}{2}, m_j, \ell, s\rangle = \alpha_-|\ell, m_j - \frac{1}{2}\rangle\,|s, +\frac{1}{2}\rangle \;+\; \beta_-|\ell, m_j + \frac{1}{2}\rangle\,|s, -\frac{1}{2}\rangle \\ \\ |\ell + \frac{1}{2}, m_j, \ell, s\rangle = \alpha_+|\ell, m_j - \frac{1}{2}\rangle\,|s, +\frac{1}{2}\rangle + \beta_+|\ell, m_j + \frac{1}{2}\rangle\,|s, -\frac{1}{2}\rangle \end{cases} \tag{96}$$

with

$$-(\ell - 1/2) \le m_j \le \ell - 1/2\,, \qquad -(\ell + 1/2) \le m_j \le \ell + 1/2\,,$$

respectively, and

$$\alpha_\pm = \pm\sqrt{\frac{\ell \pm m_j + 1/2}{2\ell + 1}} = \pm\beta_\mp\,.$$

Note that both $|\alpha_-|^2 + |\beta_-|^2 = 1$ and $|\alpha_+|^2 + |\beta_+|^2 = 1$ are fulfilled.

Comment. In order **to determine coefficients** $C(m, m_s)$ one can exploit the spectrum-generating algebra method already applied for the harmonic oscillator and the definition of the angular-momentum basis $\{|\ell, m\rangle : |m| \le \ell\}$. With $j = \ell + \frac{1}{2}$, for example, one can identify the "highest" state (with respect to the action of raising operator J_+) through the condition

$$J_+|X\rangle = (L_+ + S_+)|X\rangle \equiv 0 \;\Rightarrow\; |X\rangle = |\ell, \ell\rangle\,\left|s, +\tfrac{1}{2}\right\rangle\,, \quad \text{where } L_+|\ell, \ell\rangle = 0\,,\;\; S_+\left|s, +\tfrac{1}{2}\right\rangle = 0\,.$$

Since $\max(m_j) = j = \ell + 1/2$ with $j = \ell + s$ for a given ℓ and $s = 1/2$, then

$$|j, j, \ell, s\rangle \;\equiv\; |X\rangle = |\ell, \ell\rangle\,\left|s, +\tfrac{1}{2}\right\rangle\,.$$

Likewise, from $J_-|Y\rangle = 0$ one identifies the "lowest" state $|Y\rangle = |j, j, -\ell, s\rangle \;\equiv\; |\ell, -\ell\rangle\,\left|s, -\tfrac{1}{2}\right\rangle$. The repeated action of J_- generates the whole series of states forming the basis

$$(J_-)^k|j, j, \ell, s\rangle = \lambda_k\,|j, j - k, \ell, s\rangle\,.$$

This terminates for $k = 2j+1$ namely $(J_-)^{2j+1}|j,j,\ell,s\rangle = 0$. Parameters λ_k are found by imposing the normalization condition on the states generated by this procedure.

General case. $\vec{J} = \vec{J}_a + \vec{J}_b$

$$J_{a3}|j_a,m_a\rangle = \hbar\, m_a\, |j_a,m_a\rangle\,, \;\; |m_a| \le j_a\,, \qquad J_a^2|j_a,m_a\rangle = \hbar^2\, j_a(j_a+1)\, |j_a,m_a\rangle$$

$$J_{b3}|j_b,m_b\rangle = \hbar\, m_b\, |j_b,m_b\rangle\,, \;\; |m_b| \le j_b\,, \qquad J_b^2|j_b,m_b\rangle = \hbar^2\, j_b(j_b+1)\, |j_b,m_b\rangle$$

The new states have the form

$$|j,m_j,j_a,j_b\rangle = \sum\nolimits^*_{m_a,m_b} C(\,m_a,m_b\,)|j_a,m_a\rangle\, |j_b,m_b\rangle$$

where the summation on m_a, m_b undergoes the constraint $m_a + m_b = m_j = \text{constant}$. Rather obviously, such states diagonalize J_a^2, J_b^2 as well as J_3

$$J_3|j,m_j,j_a,j_b\rangle = (m_a+m_b) \sum\nolimits^*_{m_a,m_b} C(\,m_a,m_b\,)|j_a,m_a\rangle\, |j_b,m_b\rangle\,.$$

By assuming, for example, $j_a \ge j_b$

$$j_a - j_b \;\le\; j \;\le\; j_a + j_b\,, \qquad -j \;\le\; m_j = m_a + m_b \;\le\; -j\,.$$

For given j_a, j_b index m_a assumes $2j_a+1$ values while m_b assumes $2j_b+1$ values. The total number of independent product states is $(2j_a+1)(2j_b+1)$ which is necessarily equal to the number of new states $|j,m_j,j_a,j_b\rangle$.

12.3 Exercise: addition of two spins

$\vec{J} = \vec{J}_a + \vec{J}_b = \vec{S}_a + \vec{S}_b$. Since $j_a = j_b = 1/2$ then $m_a, m_b = \pm 1/2$

$$S_{a3}\,|\tfrac{1}{2},m_a\rangle = \hbar\, m_a\; |\tfrac{1}{2},m_a\rangle\;, \qquad S_a^2\,|\tfrac{1}{2},m_a\rangle = \hbar^2\,\frac{3}{4}\;|\tfrac{1}{2},m_a\rangle\;,$$

$$S_{b3}\,|\tfrac{1}{2},m_b\rangle = \hbar\, m_b\; |\tfrac{1}{2},m_b\rangle\;, \qquad S_b^2\,|\tfrac{1}{2},m_b\rangle = \hbar^2\,\frac{3}{4}\;|\tfrac{1}{2},m_b\rangle\;.$$

First case. Since $j_a \equiv j_b = 1/2$ one has $j = j_a - j_b = 0$ and $m_j \equiv 0$. The **single state** with $m_j = 0 = m_a + m_b$ is obtained by combining states with $m_a = -m_b = \pm 1/2$

$$|j,m_j,j_a,j_b\rangle \equiv |0,0,1/2,1/2\rangle = \frac{1}{\sqrt{2}}\,(|j_a,{-\tfrac{1}{2}}\rangle\;|j_b,{+\tfrac{1}{2}}\rangle + \sigma\,|j_a,{+\tfrac{1}{2}}\rangle\;|j_b,{-\tfrac{1}{2}}\rangle)$$

with $j_a \equiv j_b = 1/2$. After observing that $\sigma = \pm 1$ ensures in both cases the correct normalization, the choice $\sigma = -1$ is **necessary** to have $J^2|0,0,1/2,1/2\rangle = 0$, consistent with $j = 0$.
Proof

$$J^2 = J_3^2 + J_1^2 + J_2^2 = J_3^2 + \frac{1}{2}\Big(J_+J_- + J_-J_+\Big) = J_3^2 + \frac{1}{2}\Big(2J_+J_- - 2\hbar\, J_3\Big)$$

Then $J^2|0,0,\frac{1}{2},\frac{1}{2}\rangle = J_+J_-|0,0,\frac{1}{2},\frac{1}{2}\rangle$, entailing that $J_+J_-|0,0,\frac{1}{2},\frac{1}{2}\rangle = 0$ must be satisfied. By exploiting formulas (92) and (93), one finds that

$$J_+J_-|0,0,\tfrac{1}{2},\tfrac{1}{2}\rangle = (S_{a+} + S_{b+})(S_{a-} + S_{b-})\frac{1}{\sqrt{2}}\Big(|-\tfrac{1}{2}\rangle\,|+\tfrac{1}{2}\rangle + \sigma\,|+\tfrac{1}{2}\rangle\,|-\tfrac{1}{2}\rangle\Big)$$

$$= \big(S_{a+} + S_{b+}\big)\frac{\hbar}{\sqrt{2}}\Big(|-\tfrac{1}{2}\rangle\,|-\tfrac{1}{2}\rangle + \sigma\,|-\tfrac{1}{2}\rangle\,|-\tfrac{1}{2}\rangle\Big) = \hbar\,\frac{\sigma+1}{\sqrt{2}}\,\big(S_{a+} + S_{b+}\big)\,|-\tfrac{1}{2}\rangle\,|-\tfrac{1}{2}\rangle = 0\,.$$

provided $\sigma = -1$. State $|0,0,\frac{1}{2},\frac{1}{2}\rangle$ is called **singlet state**.

Second case. In this case $j = j_a + j_b = 1$ and $m_j = 0, \pm 1$. The new states have the form (indexes $j_a = j_b = 1/2$ are used in the following discussion to distinguish state of spin $\vec{S}_a$ from states of spin $\vec{S}_b$)

$$\begin{cases} |1,1,j_a,j_b\rangle = \big|j_a,+\frac{1}{2}\big\rangle\,\big|j_b,+\frac{1}{2}\big\rangle = |1,+1\rangle \\ \\ |1,0,j_a,j_b\rangle = \frac{1}{\sqrt{2}}\Big(\big|j_a,+\frac{1}{2}\big\rangle\,\big|j_b,-\frac{1}{2}\big\rangle + \nu\,\big|j_a,-\frac{1}{2}\big\rangle\,\big|j_b,+\frac{1}{2}\big\rangle\Big) = |1,0\rangle \\ \\ |1,-1,j_a,j_b\rangle = \big|j_a,-\frac{1}{2}\big\rangle\,\big|j_b,-\frac{1}{2}\big\rangle = |1,-1\rangle \end{cases} \tag{97}$$

Such states, called **triplet states**, satisfy $J_3|1,m_j,\frac{1}{2},\frac{1}{2}\rangle = \hbar m_j\,|1,m_j,\frac{1}{2},\frac{1}{2}\rangle$ and the normalization condition. The undetermined parameter ν is such that $\nu = +1$ in order **to ensure the orthogonality** of states $|1,m_j,j_a,j_b\rangle$ with respect to state $|0,0,j_a,j_b\rangle$ (their index j correspond to different values).

It is interesting to check that the action of $J_\pm$ indeed reproduces that of raising and lowering operators. For example

$$J_+|1,-1,j_a,j_b\rangle = (S_{a+} + S_{b+})\;|j_a,-\tfrac{1}{2}\rangle\,|j_b,-\tfrac{1}{2}\rangle$$

$$= \Big(S_{a+}\,|j_a,-\tfrac{1}{2}\rangle\Big)\,|j_b,-\tfrac{1}{2}\rangle + |j_a,-\tfrac{1}{2}\rangle\;S_{b+}\,|j_b,-\tfrac{1}{2}\rangle = |j_a,+\tfrac{1}{2}\rangle\,|j_b,-\tfrac{1}{2}\rangle + |j_a,-\tfrac{1}{2}\rangle\,|j_b,+\tfrac{1}{2}\rangle$$

The latter coincides with state $|1,0,j_a,j_b\rangle$ up to the normalization factor $1/\sqrt{2}$. Another example is

$$J_+|1,0,j_a,j_b\rangle = (S_{a+} + S_{b+})\;\frac{1}{\sqrt{2}}\Big(|j_a,+\tfrac{1}{2}\rangle\,|j_b,-\tfrac{1}{2}\rangle + |j_a,-\tfrac{1}{2}\rangle\,|j_b,+\tfrac{1}{2}\rangle\Big)$$

$$= (S_{a+} + S_{b+})\;\frac{1}{\sqrt{2}}\,|j_a,+\tfrac{1}{2}\rangle\,S_{b+}\,|j_b,-\tfrac{1}{2}\rangle + \frac{1}{\sqrt{2}}\Big(S_{a+}\,|j_a,-\tfrac{1}{2}\rangle\Big)\,|j_b,+\tfrac{1}{2}\rangle = \sqrt{2}\;|j_a,+\tfrac{1}{2}\rangle\;|j_b,+\tfrac{1}{2}\rangle$$

which, apart from the extra factor $\sqrt{2}$, once more generates the expected vector $|1,1,j_a,j_b\rangle$.

An interesting result emerging from the previous analysis is that the **addition process** so far discussed allows one to derive the **matrix realizations** of the angular-momentum and spin operators in many independent ways together with the relevant vector basis. While angular momenta are characterized by integral j, spin operators correspond to j assuming half-integral values.

One should recall that, while the method for adding angular momenta and spins provides a constructive, almost iterative, procedure to generate angular momenta with an (in principle) arbitrarily large representation index j, the spectrum-generating algebra method supplies the **basic procedure** for determining the matrix realization of angular momenta and spins whose essence is to be entirely based on the commutation-relation properties. This procedure was discussed in subsection 9.1: its application to the angular momentum $\vec{L}$ led to determine the standard basis through the action of operators $L_{\pm}$.

In general, index j relevant to $\vec{J}$ (as well as indexes ℓ or s relevant to $\vec{L}$ and $\vec{S}$, respectively) is called the **representation index**. It shows that the matrix dimension of a given representation of spin/angular momentum operators is $(2j+1)\times(2j+1)$ since the number of the standard-basis vectors is $2j+1$. As a consequence, index j plays a double role: in addition to describe the formal properties of $\vec{J}$, it determines the eigenvalue of $J^2|j,m_j\rangle = \hbar^2 j(j+1)|j,m_j\rangle$ which provides, at the quantum-mechanical level, information about the magnitude of operator J^2.

Angular momentum and Dirac's notation. Similar to components S_3, $S_{\pm}$ of spin $\vec{S}$ represented by 2×2 matrices, angular momentum $\vec{J}$ (together with the relevant physical states) can be represented for any j **by means of the Dirac notation** in the more abstract form involving matrices (matrix formulation of quantum mechanics) in place of the differential-operator form (58) used for the hydrogen-atom Schrödinger problem. This formulation is interesting in that one finds realizations of angular-momentum operators that are much simpler than those based on the wave-function representation of physical operators. The construction of such realization is based on the knowledge of the spectrum of J_3 and the use of the **spectrum-generating algebra method** discussed in subsection 9.1.

The following example completes the discussion of subsection 9.1 showing that, in the spirit of Dirac's picture, the spectrum-generating algebra method allows one to determine (we consider the simple case $j=1$) the matrix realization of angular momentum and the corresponding vector basis. Since the basis is formed by three elements, operators J_3, $J_{\pm}$ are necessarily 3×3 matrices. The (natural) vector form of $|1,m_j\rangle$ is

$$|1,+1\rangle = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad |1,0\rangle = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad |1,-1\rangle = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

which replace the set of functions associated to the Legendre polynomials $Y_{j,m_j}(\theta,\phi)$ with $j=1$ and $m_j=-1,\,0,+1$. Recalling that J_3 is diagonal, by definition, one finds

$$J_3 = \hbar \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad \Leftrightarrow \quad J_3|1,m_j\rangle = \hbar m_j\,|1,m_j\rangle$$

The action of angular-momentum raising/lowering operators $J_\pm$ is defined by formula (56) which, in the present case, takes the form

$$\begin{cases} J_+ \,|j, m_j\rangle = \hbar\sqrt{j(j+1) - m_j(m_j+1)}\,|j, m_j+1\rangle \\ \\ J_- \,|j, m_j\rangle = \hbar\sqrt{j(j+1) - m_j(m_j-1)}\,|j, m_j-1\rangle \end{cases} \tag{98}$$

This nice property follows from the fact that $J_\pm$ satisfy commutators $[J_3, J_\pm] = \pm J_\pm$ and $[J_+, J_-] = 2J_3$. One should recall that the derivation of formulas (56) was totally based on the commutators, namely, on the **algebraic structure** associated to such physical operators. The application of formulas (98) gives

$$\begin{aligned} J_-|1,+1\rangle &= \hbar\sqrt{j(j+1) - 1(1-1)}\,|1, 1-1\rangle = \hbar\sqrt{2}\,|1,0\rangle\,, \\ J_-\,|1,\,0\rangle \;&=\; \hbar\sqrt{1(1+1) - 0(0-1)}\,|1,-1\rangle = \hbar\sqrt{2}\,|1,-1\rangle\,. \end{aligned}$$

Assume that the matrix elements of J_- are unknown parameters a, b, c ... r. Within the matrix formalism the latter equations become

$$J_-|1,+1\rangle = \hbar \begin{bmatrix} a & b & c \\ d & e & f \\ p & q & r \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \hbar\sqrt{2} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \Rightarrow a = p = 0\,,\; d = \sqrt{2}\,,$$

$$J_-|1,\,0\rangle \;=\; \hbar \begin{bmatrix} 0 & b & c \\ \sqrt{2} & e & f \\ 0 & q & r \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \hbar\sqrt{2} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \Rightarrow\; b = e = 0\,,\; q = \sqrt{2}\,.$$

Finally, from $J_-\,|1,\,-1\rangle = 0$ one has $c = f = r = 0$ giving

$$J_- = \hbar\sqrt{2} \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \quad \Rightarrow \quad J_+ = \hbar\sqrt{2} \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

The latter satisfy commutators $[J_3, J_\pm] = \pm\hbar\, J_\pm$ and $[J_+, J_-] = 2\hbar\, J_3$ thus **confirming the validity** of the matrix representation of J_3 and $J_\pm$. for example

$$J_+J_- - J_-J_+ = 2\hbar^2 \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} - 2\hbar^2 \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

$$= 2\hbar^2 \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} - 2\hbar^2 \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = 2\hbar^2 \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix} = 2\hbar\, J_3\,.$$

13 Motion in the electromagnetic field

Uniform magnetic field. Classically, the motion of a mass m with charge q is described by the equation involving the Lorentz force

$$m\,\ddot{\vec{x}} \,=\, q\left(\vec{E}+\frac{\vec{v}}{c}\wedge\vec{B}\right) \tag{99}$$

where $\vec{E}=-\nabla\Phi$ and the **magnetic field** $\vec{B}$ (note that in the Gauss system of units $\vec{B}=\vec{H}$ with 1 gauss $=10^{-4}$ tesla $=10^{-4}$weber/m^2) is assumed to be **constant**. In this case the vector potential is given by

$$\vec{A}=\frac{1}{2}\vec{B}\wedge\vec{x}=\frac{B}{2}\left(-x_2\vec{u}_1+x_1\vec{u}_2\right),\quad \vec{B}=B\vec{u}_3\;\;\Rightarrow\;\;\;\mathrm{curl}\vec{A}=\mathrm{curl}\left(\frac{1}{2}\vec{B}\wedge\vec{x}\right)=\vec{B}\,.$$

The Lagrangian has the form

$$L=K-U=\frac{m}{2}\dot{\vec{x}}^{\,2}\,+\,\frac{q}{c}\,\dot{\vec{x}}\cdot\vec{A}\,-\,q\,\Phi=\frac{m}{2}\dot{\vec{x}}^{\,2}\,+\frac{qB}{2c}\left(x_1\dot{x}_2-x_2\dot{x}_1\right)\,-\,q\,\Phi\,.$$

To check the validity of the latter we derive the relevant Lagrangian equations

$$\frac{d}{dt}\frac{\partial L}{\partial\dot{x}_1}-\frac{\partial L}{\partial x_1}=\frac{d}{dt}\left(m\dot{x}_1-\frac{qB}{2c}\,x_2\right)-\left(\frac{qB}{2c}\,\dot{x}_2-q\,\frac{\partial\Phi}{\partial x_1}\right)=m\ddot{x}_1-\frac{qB}{c}\,\dot{x}_2+q\,\frac{\partial\Phi}{\partial x_1}=0$$

$$\frac{d}{dt}\frac{\partial L}{\partial\dot{x}_2}-\frac{\partial L}{\partial x_2}=\frac{d}{dt}\left(m\dot{x}_2+\frac{qB}{2c}\,x_1\right)-\left(-\frac{qB}{2c}\,\dot{x}_1-q\,\frac{\partial\Phi}{\partial x_2}\right)=m\ddot{x}_2+\frac{qB}{c}\,\dot{x}_1+q\,\frac{\partial\Phi}{\partial x_1}=0$$

The resulting equations

$$m\ddot{x}_1=+\frac{qB}{c}\,\dot{x}_2-q\,\frac{\partial\Phi}{\partial x_1}\,,\quad m\ddot{x}_2=-\frac{qB}{c}\,\dot{x}_1-q\,\frac{\partial\Phi}{\partial x_2}\,,$$

coincide with the components of the Lorentz equations where $E_1=-\partial_1\Phi$ and $E_2=-\partial_2\Phi$. The ensuing **Hamiltonian** is found by calculating the generalized momenta

$$p_1=\frac{\partial L}{\partial\dot{x}_1}=m\dot{x}_1-\frac{qB}{2c}\,x_2\,,\;\;p_2=\frac{\partial L}{\partial\dot{x}_2}=m\dot{x}_2+\frac{qB}{2c}\,x_1\,,\;\;p_3=\frac{\partial L}{\partial\dot{x}_3}=m\dot{x}_3\;\;\Rightarrow$$

$$H=\vec{p}\cdot\dot{\vec{x}}-L=\left(m\dot{\vec{x}}+\frac{q}{c}\,\vec{A}\right)\cdot\dot{\vec{x}}-\frac{m}{2}\dot{\vec{x}}^{\,2}\,+\,q\,\Phi-\,\frac{q}{c}\,\dot{\vec{x}}\cdot\vec{A}=+\frac{m}{2}\left(\dot{x}_1^2+\dot{x}_2^2+\dot{x}_2^3\right)\,+\,q\,\Phi$$

$$\Rightarrow\quad H=\frac{p_3^2}{2m}+\frac{1}{2m}\left(p_1+\frac{qB}{2c}\,x_2\right)^2+\frac{1}{2m}\left(p_2-\frac{qB}{2c}\,x_1\right)^2\,+\,q\,\Phi\,.$$

The relevant Poisson brackets read

$$\{x_k,p_n\}=\delta_{kn}\,,\quad\{p_k,p_n\}=\{x_k,x_n\}=0\,.$$

Comment. It is worth noting that in this model the velocity components $\dot{x}_1$ and $\dot{x}_2$ orthogonal to $\vec{B}$ have non zero Poisson brackets. To see this, let us define the **generalized momenta** P_1 and P_2

$$P_k\,:=m\dot{x}_k=\,p_k+\,\frac{q}{c}\,A_k\;\;\Rightarrow\quad\{P_1,P_2\}=\{p_1+\,\frac{q}{c}\,A_1,p_2+\,\frac{q}{c}\,A_2\}$$

$$= \frac{q}{c}\{p_1,\, A_2\} + \frac{q}{c}\{\, A_2, p_1\} = \frac{q}{c}\left(\frac{\partial A_2}{\partial x_1} - \frac{\partial A_1}{\partial x_2}\right) = \frac{q}{c}\left(\frac{B}{2} + \frac{B}{2}\right) = \frac{qB}{c}$$

This rather counterintuitive fact is due to the presence of magnetic field B. As a consequence generalized momenta P_1 and P_2 will be, quantum-mechanically, non commuting operators. Based on the **canonical quantization rule**, one has, in addition to $[p_k, p_n] = [x_k, x_n] = 0$

$$[x_k, p_n] = \, i\hbar\, \delta_{kn}\,, \quad [P_1, P_2] = \, i\hbar\, \frac{qB}{c}\,.$$

General case, $\vec{B}$ non uniform. The Lorentz equation has the usual form

$$m\,\ddot{\vec{x}} \,=\, q\left(\vec{E} + \frac{\vec{v}}{c} \wedge \vec{B}\right)$$

where, based on the Maxwell-equation theory, the electric field $\vec{E}$ and the magnetic field $\vec{H}$ can be expressed in terms of the vector potential $\vec{A}$ and the scalar potential Φ

$$\vec{E} = -\nabla\Phi - \frac{1}{c}\partial_t \vec{A}\,, \qquad \vec{B} = \,\mathrm{curl}\,\vec{A}\,.$$

The **Lagrangian** that reproduces such an equation is given by

$$L = K - U = \frac{m}{2}\dot{\vec{x}}^2 \,-\, q\,\Phi + \,\frac{q}{c}\,\dot{\vec{x}}\cdot\vec{A} \qquad \text{where } U\big(\vec{x},\dot{\vec{x}}\big) = \,+q\,\Phi(\vec{x}) - \,\frac{q}{c}\,\dot{\vec{x}}\cdot\vec{A}\,.$$

Proof.

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{x}_i} - \frac{\partial L}{\partial x_i} = 0 \;\;\Rightarrow\;\; \frac{d}{dt}\left(m\dot{x}_i + \frac{q}{c}\,A_i\right) - \left(-q\frac{\partial\Phi}{\partial x_i} + \frac{q}{c}\,\dot{x}_k\frac{\partial A_k}{\partial x_i}\right) = 0 \;\Rightarrow$$

$$m\ddot{x}_i + \frac{q}{c}\dot{x}_k\frac{\partial A_i}{\partial x_k} \,-\, \left(-q\frac{\partial\Phi}{\partial x_i} + \frac{q}{c}\,\dot{x}_k\frac{\partial A_k}{\partial x_i}\right) = 0 \;\Rightarrow\;\; m\ddot{x}_i = \,-q\frac{\partial\Phi}{\partial x_i} \,+\, \frac{q}{c}\,\dot{x}_k\left(\frac{\partial A_k}{\partial x_i} - \frac{\partial A_i}{\partial x_k}\right)$$

Then

$$m\ddot{\vec{x}} = \,-q\,\nabla\Phi + \,\frac{q}{c}\,\left(\dot{x}_k\,\nabla A_k - (\dot{\vec{x}}\cdot\nabla)\vec{A}\right) \quad \text{gives} \quad m\ddot{\vec{x}} = \,q\,\vec{E} + \,\frac{q}{c}\,\dot{\vec{x}}\wedge\vec{B}$$

being

$$\dot{\vec{x}}\wedge\vec{B} = \dot{\vec{x}}\wedge\mathrm{curl}\vec{A} = \vec{u}_i\dot{x}_j\epsilon_{ijm}\epsilon_{mkn}\frac{\partial A_n}{\partial x_k} = \vec{u}_i\dot{x}_j\Big(\delta_{ik}\delta_{jn} - \delta_{in}\delta_{jk}\Big)\frac{\partial A_n}{\partial x_k}$$

$$= \vec{u}_k\dot{x}_n\frac{\partial A_n}{\partial x_k} - \vec{u}_n\dot{x}_k\frac{\partial A_n}{\partial x_k} = \dot{x}_n\nabla A_n - (\dot{\vec{x}}\cdot\nabla)\vec{A} \qquad QED\,.$$

The ensuing **Hamiltonian** reads

$$p_k = \frac{\partial L}{\partial\dot{x}_k} = \,m\dot{x}_k + \,\frac{q}{c}\,A_k \;\;\Rightarrow\;\; H = \vec{p}\cdot\dot{\vec{x}} - L = \left(m\dot{\vec{x}} + \frac{q}{c}\,\vec{A}\right)\cdot\dot{\vec{x}} - \frac{m}{2}\dot{\vec{x}}^2 \,+\, q\,\Phi - \,\frac{q}{c}\,\dot{\vec{x}}\cdot\vec{A}$$

$$\Rightarrow\;\; H = +\frac{m}{2}\dot{\vec{x}}^2 \,+\, q\,\Phi = \frac{1}{2m}\left(\vec{p} - \frac{q}{mc}\,\vec{A}\right)^2 + \,q\,\Phi\,.$$

This concludes the derivation of the magnetic Hamiltonian in the presence of an arbitrary $\vec{B}$.

13.1 Derivation of the quantum Hamiltonian. Normal Zeeman effect.

The motion of a single electron (with charge $q = -e$ and mass $m = m_e$) undergoing the action of the atomic nucleus is given by (the operator symbol $\hat{}$ is omitted)

$$H_B = \frac{1}{2m_e}\left(\vec{p} + \frac{e}{c}\,\vec{A}\right)^2 - e\Phi(x) = \frac{1}{2m_e}\left[p_3^2 + \left(p_1 + \frac{e}{c}\,A_1\right)^2 + \left(p_2 + \frac{e}{c}\,A_2\right)^2\right] - \frac{Ze^2}{x}$$

$$= \frac{1}{2m_e}\left[p_3^2 + \left(p_1 - \frac{eB}{2c}\,x_2\right)^2 + \left(p_2 + \frac{eB}{2c}\,x_1\right)^2\right] - \frac{Ze^2}{x}$$

$$= \frac{1}{2m_e}\left(p_3^2 + p_1^2 + p_2^2\right) + \frac{eB}{2m_e c}\,(-x_2 p_1 + x_1 p_2) + \frac{e^2B^2}{8m_e c^2}\,\left(x_2^2 + x_1^2\right) - \frac{Ze^2}{x}$$

$$H_B = \frac{\vec{p}^2}{2m_e} + \frac{eB}{2m_e c}\,L_3 + \frac{e^2B^2}{8m_e c^2}\,\left(x_2^2 + x_1^2\right) - \frac{Ze^2}{x} \tag{100}$$

In the presence of **weak magnetic fields** the B^2-dependent term can be neglected while the B-dependent term can viewed as a perturbation. Then

$$H_B \simeq \frac{\vec{p}^2}{2m_e} - \frac{Ze^2}{x} + \frac{eB}{2m_e c}\,L_3 = H_0 + \frac{eB}{2m_e c}\,L_3\,. \tag{101}$$

Note that, in H_B, the term

$$H_\mu = \frac{eB}{2m_e c}\,L_3 = \frac{e}{2m_e c}\,\vec{B}\cdot\vec{L} = -\vec{\mu}\cdot\vec{B}\,, \quad \vec{\mu} = -\frac{e}{2m_e c}\,\vec{L}$$

exhibits the usual form of the potential energy characterizing a magnetic dipole immersed in a field $\vec{B} = B\vec{u}_3$. Vector operator $\vec{\mu}$ is the **magnetic moment** classically predicted by the Ampere law. The (classical) current i of a closed circuit of section S determines a magnetic moment

$$\vec{\mu} = \frac{-e}{2m_e c}\,\vec{L} = \frac{-e}{2m_e c}\,m_e R^2\,\dot{\phi}\,\vec{u}_3 = -\frac{e}{c}\,R^2\,\frac{\pi}{T}\,\vec{u} = \frac{1}{c}\,S\,i\,\vec{u}\,,$$

the circuit current due to the electron motion corresponding to $i = -e/T$ and the circuit section being $S = \pi R^2$. The radius R is of the same order as the atomic radius. This explains the meaning of the energy term H_μ which describes the coupling of the magnetic moment of the electron, due to its orbital motion, with field $\vec{B}$.

Energy eigenvalues and eigenstates. Since, by construction, $\vec{B}$ is parallel to the axis z then $\vec{B}\cdot\vec{L} = BL_3$. The energy eigenvalue problem reads

$$H_B\psi = E_{nm}\,\psi \;\Rightarrow\; \left(H_0 + H_\mu\right)\psi_{n\ell m} = E_{nm}\,\psi_{n\ell m}\,, \quad H_\mu = \frac{eB}{2m_e c}\,L_3\,,$$

where we expect that E_{nm} depends on m due to the presence of operator L_3 in H_B. Thanks to the fact that $[\,H_0,\,L_3\,] = 0$ the term H_μ commutes with H_0 and therefore $[\,H_B,\,H_\mu\,] = 0$.

States ψ can be identified with eigenfunctions $\psi_{n\ell m}$, the **hydrogen-atom eigenstates**. One finds

$$H_0\,\psi_{n\ell m} = E_n\,\psi_{n\ell m}\,, \qquad H_\mu\,\psi_{n\ell m} = \frac{eB}{2m_e c}\,\hbar m\,\psi_{n\ell m}$$

The presence of H_μ **eliminates the degeneracy** characterizing quantum number m of states $\psi_{n\ell m}$ giving the energy eigenvalues

$$E_{nm} = -\frac{(Ze)^2}{2a\,n^2} + \frac{eB}{2m_e c}\,\hbar m = -\frac{(Ze)^2}{2a\,n^2} + \mu_B B\,m\,, \qquad \mu_B = \frac{e\hbar}{2m_e c}\,, \quad a = \frac{\hbar^2}{m_e\,e^2}$$

The quantity $\mu_B = 0,927\,10^{-20}$erg/gauss ($1\,\text{erg} = 10^{-7}\,\text{J}$) is called the **Bohr magneton** and describes the **magnetic-moment quanta**. In the presence of weak magnetic fields (for example, $B \leq 10^5\,$gauss) the magnetic-like energy $\mu_B B \simeq 10^{-22} J$ can be compared with the hydrogen-atom electron energy $|E_1| = -13.6\,eV \simeq -2\times 10^{-18}\,J$. With this B, the third term in (100) becomes

$$\frac{e^2B^2}{8m_e c^2} \times a^2 = \frac{(\mu_B B)^2 m_e a^2}{2\hbar^2} \approx 10^{-27}\ J$$

when observing that $x_2 \approx x_1 \approx a$. This justifies its suppression.

The main aspect characterizing the new spectrum is that a single level associated to E_n for $B = 0$ undergoes a **splitting effect** generating $2\,n - 1$ new energy levels. More specifically,

$$-\ell \leq m \leq \ell\,, \quad \ell = 0, 1, 2, \ldots n-1 \ \Rightarrow \qquad -(n-1) \leq m \leq n-1$$

Examples. Apart from the simple case $n = 1 \to \ell = 0 \to m = 0$ one has

$$n = 2 \ \to \ m = -1,\,0,\,+1 \quad \begin{cases} m = -1 \to \quad \ell = 1 \\ m = 0 \to \quad \ell = 0, 1 \\ m = +1 \to \quad \ell = 1 \end{cases}$$

$$n = 3 \ \to \ m = -2,-1,\,0,\,+1,\,+2 \quad \begin{cases} m = -2 \to \quad \ell = 2 \\ m = -1 \to \quad \ell = 1, 2 \\ m = 0 \to \quad \ell = 0, 1, 2 \\ m = +1 \to \quad \ell = 1, 2 \\ m = +2 \to \quad \ell = 2 \end{cases}$$

Then for each pair of quantum numbers (n, m) the range of ℓ is $|m| \leq \ell \leq n-1$ and the relevant degeneracy is $n - |m|$.

The emergence of new levels (recall that $\mu_B = e\hbar/(2m_e c^2)$ is the Bohr magneton)

$$n = 2, |m| \leq 1 \quad \begin{cases} E_{2,-1} = E_2 + \mu_B B(-1) \\ E_{2,0} = E_2 \\ E_{2,+1} = E_2 + \mu_B B(+1) \end{cases}, \quad n = 3, |m| \leq 2 \quad \begin{cases} E_{3,-2} = E_3 + \mu_B B(-2) \\ E_{3,-1} = E_3 + \mu_B B(-1) \\ E_{3,0} = E_3 \\ E_{3,+1} = E_3 + \mu_B B(+1) \\ E_{3,+2} = E_3 + \mu_B B(+2) \end{cases}$$

determines additional transitions whose effective number is restricted by the **selection rule** $\Delta m = 0, \pm 1$. Each transition of the case $B = 0$ (for example from $n = 3$ to $n = 1$ or from $n = 2$ to $n = 1$) is now replaced by a group of transitions absent when $B = 0$. Such an effect represents the so-called **normal Zeeman effect**. The number of the predicted (new) transitions reproduces, however, a limited number of experimental observations. In general, a **larger number** of new energy levels appears to be necessary to reach a satisfactory agreement with experimental data.

This leads to investigate a more complex Hamiltonian where the effects (so far ignored) of the **electron spin** are taken into account. Within the improved theoretical framework including spin operators the splitting of the $B = 0$ energy levels in a **larger number of energy levels** is called the **anomalous Zeeman effect**.

Comment on selection rules and electric-dipole transitions. In principle, in a hydrogen-like atom any transition $|n, \ell, m\rangle \to |n', \ell', m'\rangle$ is allowed. This apparent freedom is limited by the fact that any energy change $E_n - E_{n'}$ $(n \neq n')$ causes the emission of electromagnetic radiation characterized by frequency $w_{nn'} = (E_n - E_{n'})/\hbar$. The classical formula for the intensity of the emitted radiation suggests the quantum formula whereby selection rules are found

$$I = \frac{2e^2}{3c^3}\left(\frac{d^2\mathbf{r}}{dt^2}\right)^2 \quad \Rightarrow \quad I_{\nu',\nu} = \frac{2e^2}{3c^3}\left(\frac{d^2}{dt^2}\langle t, \nu'|\mathbf{r}|\nu, t\rangle\right)^2 \simeq \frac{2e^2}{3c^3}\, w_{nn'}^4 \, \left|\langle t, \nu'|\mathbf{r}|\nu, t\rangle\right|^2 ,$$

where $\nu' = (n', \ell', m')$, $\nu = (n, \ell, m)$, $\mathbf{r}(t)$ describes the (classical) position of the oscillating dipole charge, and the expectation value $\langle t, m', \ell', n'|\mathbf{r}|\, n, \ell, m, t\rangle$ represents its quantum counterpart $\mathbf{r}$ being the position operator. Stationary states $|\nu, t\rangle = e^{-itE_n/\hbar}|\nu, 0\rangle$ have been used. **Selection rules** describe the cases when $\langle t, \nu'|\mathbf{r}|\nu, t\rangle$ is **non zero** and thus the emission of radiation takes place.

13.2 Inclusion of spin $\vec{S}$ and of relativistic effects

The Stern-Gerlach experiment (1922) showed that an atomic beam with atoms having **zero total angular momentum** $(\ell = 0)$ is splitted into two beams due to the presence of a magnetic field $B \neq 0$. Such an effect has demonstrated how electrons are equipped with a $1/2$ spin $\vec{S}$ whose magnetic moment must be added to the moment of angular momentum $\vec{L}$.

For illustrative purposes it is interesting to write the extended Hamiltonian which includes both spin-dependent terms and relativistic terms

$$H = H_0 + H_\mu + H_{SL} + H_R\,, \qquad H_0 = \frac{\vec{p}^2}{2m} - \frac{Ze^2}{r}\,.$$

Term

$$H_\mu = -(\vec{\mu} + \vec{\mu}_s)\cdot\vec{B} = \frac{e}{2m_e c}\left(\vec{L} + g\,\vec{S}\right)\cdot\vec{B} = \frac{e}{2m_e c}\left(L_3 + g\,S_3\right)B\,,$$

includes the coupling of spin magnetic moment with $\vec{B}$. Parameter g (gyromagnetic factor) can be proved to assume the value $g = 2$ (within the theory of angular momentum and, in particular, according to the scheme for the angular-momenta addition).

Terms H_{SL} and H_R are derived from the relativistic theory of electron based on the Dirac equation in the limit $v/c << 1$. In particular,

$$H_{SL} = \frac{1}{2m_e^2c^2}\,\frac{1}{r}\,\frac{dU}{dr}\,\vec{S}\cdot\vec{L} \,=\, \frac{Ze^2}{2m_e^2c^2r^3}\,\vec{S}\cdot\vec{L}\,, \qquad U(r) = q\,\Phi(r) = \,-\frac{Ze^2}{r}\,,$$

Another term of relativistic origin is

$$H_R = \,-\frac{(\vec{p}^{\,2})^2}{8m_e^3c^2} \,+\, ...$$

which follows, for $v^2/c^2 << 1$, from the relativistic energy-momentum equation

$$E = \sqrt{c^2\vec{p}^{\,2} + m_e^2c^4} = m_ec^2\sqrt{1 + \frac{\vec{p}^{\,2}}{m_e^2c^2}} \simeq\, m_ec^2 + \,\frac{\vec{p}^{\,2}}{2m_e} \,-\, \frac{(\vec{p}^{\,2})^2}{8m_e^3c^2} \,+\, ...$$

While kinetic term $\vec{p}^{\,2}/(2m_e)$ is contained in H_0, constant term $m_e^2c^4$ (describing the electron rest energy) can be included in turn in H_0 and involves a simple shift of the energy levels.

Comment. Recall that

$$\vec{S}\cdot\vec{L} \approx \hbar^2\,, \quad r \simeq a = \frac{\hbar^2}{m_e\,e^2}\,; \qquad \alpha = \frac{e^2}{\hbar\,c} \simeq \frac{1}{137}\,, \quad \alpha\,a \,=\, \frac{\hbar}{m_ec}\,.$$

Then the magnitude of energies E_{SL} and $|E_R|$ relevant to H_{SL} and H_R, respectively, is given by

$$E_{SL} = \frac{Ze^2}{2m_e^2c^2r^3}\,\vec{S}\cdot\vec{L} \approx \frac{Ze^2}{2m_e^2c^2a^3}\,\hbar^2 = \frac{Ze^2\alpha^2\,a^2}{2a^3} = \,Z\alpha^2\frac{e^2}{2a} \,=\, Z\alpha^2|E_1|\,,$$

$$(p \approx \hbar k = \hbar/a \;\rightarrow) \qquad |E_R| = \frac{p^4}{8m_e^3c^2} \approx \frac{\hbar^4}{8m_e^3c^2a^4} = \frac{\hbar^2}{8m_e}\frac{(\alpha\,a)^2}{a^4} = \frac{\hbar^2\alpha^2}{4m_ee^2}\,\frac{|E_1|}{a} = \frac{\alpha^2}{4}\,|E_1|\,,$$

where $|E_n| = (Ze)^2/(2an^2)$ with $n = 1$, $Z = 1$ is the first level of the hydrogen atom, confirming that such terms represent small perturbations of H_0. Note that $\alpha^2 \simeq 5 \times 10^{-5}$.

Anomalous Zeeman effect and Paschen-Back effect. Due to the perturbative character of H_R and H_{SL}, let us consider

$$H \simeq H_0 + H_\mu = \frac{\vec{p}^{\,2}}{2m} - \frac{Ze^2}{r} \,+ \frac{e}{2m_ec}\Big(L_3 + g\,S_3\Big)\,B\,, \tag{102}$$

thus neglecting relativistic effects. The presence of the magnetic moment of $\vec{S}$ seems to be sufficient to cause the occurrence of a larger number of energy levels. The eigenstates of such a Hamiltonian are readily found to be the spinors

$$\Psi_{n,\ell,m,m_s}(r,\theta,\varphi) = \psi_{n,\ell,m}(r,\theta,\varphi)\,|s,m_s\rangle \tag{103}$$

since

$$\left(H_0 + \frac{e\,B}{2m_e c}\,L_3\right)\,\psi_{n,\ell,m} \;=\; \left(E_n + \frac{e\,B}{2m_e c}\,\hbar m\right)\,\psi_{n,\ell,m}\,, \qquad S_3\,|s,m_s\rangle \;=\; \hbar m_s\,|s,m_s\rangle\,. \quad (104)$$

Then (note that H_0 and L_3 act on $\psi_{n,\ell,m}$ while S_3 acts on $|s,m_s\rangle$)

$$\left(H_0 + \frac{e\,B}{2m_e c}\,(L_3 + gS_3)\right)\,\psi_{n,\ell,m}\,|s,m_s\rangle = \left[\left(H_0 + \frac{e\,B}{2m_e c}\,L_3\right)\,\psi_{n,\ell,m}\right]\,|s,m_s\rangle +$$

$$+\psi_{n,\ell,m}\,\left(\frac{egB}{2m_e c}\,S_3\,|s,m_s\rangle\right) = \left(E_n + \frac{e\,B}{2m_e c}\,\hbar m\right)\,\psi_{n,\ell,m}\,|s,m_s\rangle + \psi_{n,\ell,m}\,\left(\frac{egB}{2m_e c}\,\hbar m_s\,|s,m_s\rangle\right)$$

$$= \left(E_n + \frac{e\,B}{2m_e c}\,\hbar\big(m + \,g m_s\,\big)\right)\,\psi_{n,\ell,m}\,|s,m_s\rangle \quad (105)$$

(recalll that $s = 1/2$, $m_s = \pm 1/2$). The latter shows that energy eigenvalues are given by

$$E_n(m,m_s) = \,E_n + \frac{e\hbar\,B}{2m_e c}\,\big(\,m + \,g\,m_s\,\big)\,, \qquad g = 2\,.$$

Since $m_s = \pm 1/2$, the new spectrum features a two-level splitting for each energy level of the Hamiltonian (101) where no spin term has been included. For **strong magnetic fields** the previous formula well reproduces the energy levels of hydrogen atoms (**Paschen-Back effect**) observed experimentally. For $n = 2$ and $n = 3$ these are given by

$$n = 2,\ |m| \le 1,\ m_s = \pm 1/2 \quad \begin{cases} E_2(-1,\pm) = \,E_2 + \mu_B B(-1 \pm 1) \\ E_2(0,\pm) = \,E_2 \pm \mu_B B \\ E_2(+1,\pm) = \,E_2 + \mu_B B(+1 \pm 1) \end{cases}$$

$$n = 3,\ |m| \le 2,\ m_s = \pm 1/2 \quad \begin{cases} E_3(-2,\pm) = \,E_3 + \mu_B B(-2 \pm 1) \\ E_3(-1,\pm) = \,E_3 + \mu_B B(-1 \pm 1) \\ E_3(0,\pm) = \,E_3 \pm \mu_B B \\ E_3(+1,\pm) = \,E_3 + \mu_B B(+1 \pm 1) \\ E_3(+2,\pm) = \,E_3 + \mu_B B(+2 \pm 1) \end{cases}$$

Part of such levels coincide. For example, one easily check that $E_n(m,-1/2) = E_n(m-2,+1/2)$. Due to selection rules $\Delta m_s = 0$ and $\Delta m = 0, \pm 1$, however, the resulting spectrum exactly reproduces the transitions characterizing the normal Zeeman effect. Apparently, the inclusion of the new spin term does not contribute to interpret the experimental observations relevant to the Zeeman effect.

The **correct approach to this problem** consists in considering the quantum numbers of operators

$$J^2\,, \quad J_3\,, \quad L^2\,, \quad S^2\,,$$

in place of those relevant to L^2, L_3, S^2, S_3 used so far. This suggestion comes from the model Hamiltonian where the **relativistic term H_{SL} is included**. If H_{SL} is **not neglected** the choice of observables J^2, J_3, L^2 and S^2 becomes **almost unavoidable** since the following commutators do not vanish

$$[\,\vec{S}\cdot\vec{L},\,S_3]\neq 0\,,\quad [\,\vec{S}\cdot\vec{L},\,L_3]\neq 0\,.$$

With the new choice, in addition to $[\,\vec{S}\cdot\vec{L},\,S^2]=0$ and $[\,\vec{S}\cdot\vec{L},\,L^2]=0$, one easily shows that

$$[\,\vec{S}\cdot\vec{L},\,J^2]=0\,,\quad [\,\vec{S}\cdot\vec{L},\,J_3]=0\,,\qquad \vec{J}=\vec{S}+\vec{L}\,,$$

where $\vec{J}^2=(\vec{S}+\vec{L})^2=\vec{S}^2+\vec{L}^2+2\vec{S}\cdot\vec{L}$ has been used. As a consequence $[H,\,J_3]=0$ and $[H,\,J^2]=0$. Within this alternative scheme (recall that $g=2$) one has

$$H=H_0+H_\mu+H_{SL}\simeq H_0+H_\mu=\frac{\vec{p}^2}{2m}-\frac{Ze^2}{r}+\frac{e}{2m_ec}\Big(J_3+\,S_3\Big)\,B\,, \tag{106}$$

where the B-dependent term can be seen as a perturbative term for B sufficiently weak whereas relativistic terms such as H_{LS} are neglected in the present calculation because they represent higher-order corrections to energy levels. The good quantum numbers are, in addition to n,

$$\ell\,,\quad s=1/2\,,\quad \ell-s\leq j\leq \ell+s\,,\quad -j\leq m_j\leq +j\,.$$

In the first approach where $H_\mu\,\propto\,L_3+2S_3$ we used eigenstates (103) represented by the direct product of the hydrogen-problem wave functions $\psi_{n,\ell,m}$ with spinors $|s,m_s\rangle$. Within the new approach, apart from term containing S_3, Hamiltonian (106) is diagonal in the basis of eigenstates

$$|n,j,m_j,\ell,s\rangle$$

where we have adopted the **Dirac notation** in order to simplify as much as possible the formalism. The radial part $R_{n\ell}(r)$ and the angular-momentum contribution $Y_{\ell,m}(\theta,\varphi)$ of eigenstate (76) is now represented by $|n,\ell,m\rangle$. The latter has been incorporated in $|j,m_j,\ell,s\rangle$ as well as spin states $|s,m_s\rangle$ according to the rules for the angular-momentum addition. The introduction of the new basis (see formula (96)) well illustrates this process where index m_j replaces both m and m_s

$$|n,j,m_j,\ell,s\rangle=\sum\nolimits^{*}_{m+m_s}C(m,m_s)|n,\ell,m\rangle\,|s,m_s\rangle\,,\quad j=\,\ell\pm 1/2\,,\quad s=\,1/2\,.$$

Note that state $|\ell,m\rangle$ has been replaced, within the Dirac picture of quantum states, with $|n,\ell,m\rangle$ including the principal quantum number n. Written in terms of wavefunctions, state $|n,\ell,m\rangle$ corresponds to the Hydrogen-atom eigenfunction (76) given by $\psi_{n,\ell,m}(r,\theta,\phi)=R_{n,\ell}(r)Y_{\ell,m}(\theta,\phi)$. In the following calculations, index n is not involved in a direct way. Thus its presence will be **implied in the new states**, namely, we set $|n,j,m_j,\ell,s\rangle\equiv|j,m_j,\ell,s\rangle$. For $j=\ell\pm 1/2$ one has

$$\Big|\ell-\tfrac{1}{2},m_j,\ell,s\Big\rangle=\alpha_-\,\Big|\ell,m_j-\tfrac{1}{2}\Big\rangle\,\Big|s,+\tfrac{1}{2}\Big\rangle\,+\,\beta_-\,\Big|\ell,m_j+\tfrac{1}{2}\Big\rangle\,\Big|s,-\tfrac{1}{2}\Big\rangle$$

$$\left|\ell+\tfrac{1}{2},m_j,\ell,s\right\rangle = \alpha_+\left|\ell,m_j-\tfrac{1}{2}\right\rangle\left|s,+\tfrac{1}{2}\right\rangle + \beta_+\left|\ell,m_j+\tfrac{1}{2}\right\rangle\left|s,+\tfrac{1}{2}\right\rangle$$

with

$$-(\ell-1/2)\le m_j\le \ell-1/2\,,\qquad -(\ell+1/2)\le m_j\le \ell+1/2\,,$$

respectively, and

$$\alpha_\pm = \pm\sqrt{(\ell\pm m_j+1/2)/(2\ell+1)} = \pm\beta_\mp\,.$$

The range of j and m_j follows from the theory of **angular-momentum addition** discussed in subsection 12.2.

The perturbation theory allows one to calculate the first-order corrections to the energy eigenvalues of the unperturbed Hamiltonian. In general, the perturbed Hamiltonian has the form $H = H_{un} + \epsilon H_p$ where H_{un} and H_p represent the unperturbed Hamiltonian and the perturbation term, respectively. Parameter ϵ allows one to connect the Hamiltonian $H = H_0 + H_p$ (fully including the perturbation for $\epsilon \to 1$) with the unperturbed one at $\epsilon = 0$. Perturbation H_p is assumed to be "small" with respect to H_{un} in the sense that matrix elements of H_p, calculated in the energy-state basis $\{|E_n^0\rangle\}$ of H_{un}, are small with respect to the eigenvalues of H_{un}. The eigenvalues **including the first-order correction** are given by

$$E_n^1 = \langle E_n^0|H|E_n^0\rangle = E_n^0 + \epsilon\,\langle E_n^0|H_p|E_n^0\rangle\,,\quad \epsilon = 1\,,$$

where E_n^0 are the eigenvalues of the unperturbed case $H_{un}|E_n^0\rangle = E_n^0|E_n^0\rangle$. In the present case

$$H_{un} = \frac{\hat{p}^2}{2m} - \frac{Ze^2}{r} + \frac{e\,B}{2m_ec}J_3,\quad H_p = \frac{eB}{2m_ec}\,S_3 \equiv \mu_B B\,\frac{S_3}{\hbar},$$
$$|E_n^0\rangle \;\to\; \left|n,\ell\pm\tfrac{1}{2},m_j,\ell,s\right\rangle\,,$$

can be seen as the unperturbed Hamiltonian, the perturbation term, and the eigenstates $|E_n^0\rangle$ of the unperturbed case, respectively. Indeed, eigenvalues $E_n^0 \approx 10\,eV$ (recall that, in the hydrogen atom, $|E_1^0| = 13.5\,eV$ with $Z = 1$) are significantly larger than $\mu_B\,B$ with $\mu_B = 0.927\times10^{-20}\text{erg}/G \simeq 0.58\times10^{-8}eV/G$ if B is not larger than $B \approx 10^5\,G$. Note that $\mu_B\,B$ determines the magnitude of the matrix elements of H_p. The perturbation theory is briefly discussed in section 14.

In the first case $j = \ell - \frac{1}{2}$ one has

$$\langle H\rangle \;=\; E_n + \frac{eB}{2m_ec}\left\langle s,\ell,m_j,\ell-\tfrac{1}{2}\right|\Big(J_3+\;S_3\Big)\left|\ell-\tfrac{1}{2},m_j,\ell,s\right\rangle$$
$$=\; E_n + \frac{eB}{2m_ec}\,\hbar\,m_j \;+\; \frac{eB}{2m_ec}\left\langle s,\ell,m_j,\ell-\tfrac{1}{2}\right|S_3\left|\ell-\tfrac{1}{2},m_j,\ell,s\right\rangle$$

Expectation value of S_3

$$\left\langle s,\ell,m_j,\ell-\tfrac{1}{2}\right|S_3\left|\ell-\tfrac{1}{2},m_j,\ell,s\right\rangle = \Big(\left\langle -\tfrac{1}{2},s\right|\left\langle \ell,m_j+\tfrac{1}{2}\right|\beta_-^* + \left\langle +\tfrac{1}{2},s\right|\left\langle \ell,m_j-\tfrac{1}{2}\right|\alpha_-^*\Big)\times$$
$$S_3\Big(\alpha_-\left|\ell,m_j-\tfrac{1}{2}\right\rangle\left|s,+\tfrac{1}{2}\right\rangle \;+\; \beta_-\left|\ell,m_j+\tfrac{1}{2}\right\rangle\left|s,-\tfrac{1}{2}\right\rangle\Big)$$

$$= |\alpha_-|^2 \left\langle m_j - \tfrac{1}{2}, \ell \,|\, \ell, m_j - \tfrac{1}{2} \right\rangle \left\langle +\tfrac{1}{2}, s \right| S_3 \left| s, +\tfrac{1}{2} \right\rangle$$
$$+ |\beta_-|^2 \left\langle m_j + \tfrac{1}{2}, \ell \,|\, \ell, m_j + \tfrac{1}{2} \right\rangle \left\langle -\tfrac{1}{2}, s \right| S_3 \left| s, -\tfrac{1}{2} \right\rangle$$
$$= |\alpha_-|^2 \left\langle +\tfrac{1}{2}, s \right| S_3 \left| s, +\tfrac{1}{2} \right\rangle + |\beta_-|^2 \left\langle -\tfrac{1}{2}, s \right| S_3 \left| s, -\tfrac{1}{2} \right\rangle = |\alpha_-|^2 \frac{\hbar}{2} + |\beta_-|^2 \left(-\frac{\hbar}{2} \right)$$
$$= \frac{\hbar}{2} \left(\frac{\ell - m_j + 1/2}{2\ell + 1} - \frac{\ell + m_j + 1/2}{2\ell + 1} \right) = -\frac{\hbar\, m_j}{2\ell + 1}\,.$$

The resulting expectation value of the energy reads

$$E_n\big(\ell - \tfrac{1}{2},\, m_j\big) = \langle H \rangle \;=\; E_n + \frac{eB\hbar}{2m_e c}\, m_j \,\left(1 \,-\, \frac{1}{2\ell + 1} \right) =\; E_n + \mu_B\, B\, m_j\, \frac{2\ell}{2\ell + 1}\,.$$

Likewise, **in the second case** $j = \ell + \frac{1}{2}$,

$$\langle H \rangle \;=\; E_n + \frac{eB}{2m_e c}\, \langle s, \ell, m_j, \ell + \tfrac{1}{2} | \Big(J_3 + \, S_3 \Big) | \ell + \tfrac{1}{2}, m_j, \ell, s \rangle$$
$$=\; E_n + \frac{eB}{2m_e c}\, \hbar\, m_j \,+ \frac{eB}{2m_e c}\, \langle s, \ell, m_j, \ell + \tfrac{1}{2} | S_3 | \ell + \tfrac{1}{2}, m_j, \ell, s \rangle$$

$$\langle s, \ell, m_j, \ell + \tfrac{1}{2} | S_3 | \ell + \tfrac{1}{2}, m_j, \ell, s \rangle =$$

$$= \langle s, \ell, m_j, \ell - \tfrac{1}{2} | S_3 \Big(\alpha_+ | \ell, m_j - \tfrac{1}{2} \rangle \, | s, +\tfrac{1}{2} \rangle \,+\, \beta_+ | \ell, m_j + \tfrac{1}{2} \rangle \, | s, -\tfrac{1}{2} \rangle \Big)$$

$$= |\alpha_+|^2 \langle +\tfrac{1}{2}, s | S_3 | s, +\tfrac{1}{2} \rangle \,+\, |\beta_+|^2 \langle -\tfrac{1}{2}, s | S_3 | s, -\tfrac{1}{2} \rangle = |\alpha_+|^2 \frac{\hbar}{2} \,+\, |\beta_+|^2 \left(-\frac{\hbar}{2} \right)$$
$$= \frac{\hbar}{2} \left(\frac{\ell + m_j + 1/2}{2\ell + 1} - \frac{\ell - m_j + 1/2}{2\ell + 1} \right) = +\frac{\hbar\, m_j}{2\ell + 1}\,.$$

giving

$$E_n\big(\ell + \tfrac{1}{2},\, m_j\big) = \langle H \rangle \;=\; E_n + \frac{eB\hbar}{2m_e c}\, m_j \,\left(1 \,+\, \frac{1}{2\ell + 1} \right) =\; E_n + \mu_B\, B\, m_j\, \frac{2\ell + 2}{2\ell + 1}\,.$$

Comparison with the energy levels of the scheme based on quantum numbers ℓ, m, m_s and $s = 1/2$. In the "old" scheme

$$E_n(m, m_s) = \; E_n + \frac{e\hbar}{2m_e c} \left(m + \, g\, m_s \right) B\,, \qquad g = 2\,.$$

$$n=1,\ |m|\le 0,\ m_s=\pm 1/2\,,\qquad E_1\left(0,\pm\tfrac{1}{2}\right)=\ E_1+\mu_B\,B\left(\,0\,\pm 1\,\right)$$

$$n=2,\,|m|\le 1,\,m_s=\pm 1/2\ \begin{cases} E_2\left(-1,\pm\tfrac{1}{2}\right)=\ E_2+\mu_B B(-1\pm 1) \\ E_2\left(0,\pm\tfrac{1}{2}\right)=\ E_2\pm\mu_B B \\ E_2\left(+1,\pm\tfrac{1}{2}\right)=\ E_2+\mu_B B(+1\pm 1) \end{cases}$$

Within the new scheme, the spectrum of the anomalous Zeeman effect is

$$E_n\big(\ell\pm\tfrac{1}{2},\,m_j\big)=\ E_n+\mu_B\,B\,m_j\,\frac{2\ell+1\pm 1}{2\ell+1}$$

$$n=1,\,\ell=0\ \rightarrow\ \ |m_j|\le 1/2,\,j=0+1/2\,,\qquad E_1\big(+\tfrac{1}{2},\,\pm\tfrac{1}{2}\big)=\ E_1+\mu_B\,B\,(\pm 1)\ ,$$

(note that $j=-1/2$ is excluded). In this case (in the absence of the spin-orbit term H_{SL}) the first (splitted) energy level is the same as in the normal Zeeman effect. With $n=2$, in addition to

$$n=2,\,\ell=0\ \rightarrow\ \ m_j=\pm 1/2,\,j=1/2\,,\qquad E_2\big(0+\tfrac{1}{2},\,\pm\tfrac{1}{2}\big)=\ E_2+\mu_B\,B\,(\pm 1)\ ,$$

(once more $j=-1/2$ is excluded) one finds

$$n=2,\,\ell=1\ \rightarrow\ \ |m_j|\le j,\ j=1\pm 1/2\ \begin{cases} E_2\left(1-\tfrac{1}{2},-\tfrac{1}{2}\right)=\ E_2+\mu_B B(-\tfrac{1}{2})\,\tfrac{2}{3} \\ E_2\left(1-\tfrac{1}{2},+\tfrac{1}{2}\right)=\ E_2+\mu_B B(+\tfrac{1}{2})\,\tfrac{2}{3} \\ \\ E_2\left(1+\tfrac{1}{2},-\tfrac{3}{2}\right)=\ E_2+\mu_B B(-\tfrac{3}{2})\,\tfrac{4}{3} \\ E_2\left(1+\tfrac{1}{2},-\tfrac{1}{2}\right)=\ E_2+\mu_B B(-\tfrac{1}{2})\,\tfrac{4}{3} \\ E_2\left(1+\tfrac{1}{2},+\tfrac{1}{2}\right)=\ E_2+\mu_B B(+\tfrac{1}{2})\,\tfrac{4}{3} \\ E_2\left(1+\tfrac{1}{2},+\tfrac{3}{2}\right)=\ E_2+\mu_B B(+\tfrac{3}{2})\,\tfrac{4}{3} \end{cases}$$

This second case well evidences the different energy-level structure of the new scheme. By taking into account the selection rules $\Delta\ell=\pm 1$ and $\Delta m_j=0,\pm 1$ (following from $\Delta m=0,\pm 1$, $\Delta m_s=0$) the transitions predicted by the new spectrum perfectly match those observed experimentally.

14 Perturbation theory

Real (physical) systems often feature a Hamiltonian whose complex structure does not allow one to perform the diagonalization process in a simple way. Usually, this is due to the coupling with external fields, the inclusion of spin degrees of freedom or of relativistic effects. In many cases, however, the derivation of the energy spectrum and of the relevant energy eigenstates can be treated in an approximate way by exploiting the fact that a component of the Hamiltonian represents a perturbation. In general, this situation is represented by

$$H = H_0 + \tau\, V\,, \quad H_0|n\rangle = E_n^0\,|n\rangle$$

where τ is a parameter that can assume arbitrarily small values, and E_n^0 and $|n\rangle$ are the eigenvalues and the eigenstates, respectively, of the **unperturbed problem**. Parameter τ allows one to connect the unperturbed Hamiltonian at $\tau = 0$ to the Hamiltonian $H = H_0 + V$, fully including the perturbation for $\tau \to 1$. Perturbation V is assumed to be "small" with respect to H_0 in the sense that matrix elements $\langle m|V|n\rangle$ of V, calculated in the energy-state basis of H_0, are small with respect to the eigenvalues of H_0.

The basic assumption consists in expanding both the eigenvalues and the eigenstates of H in the form of power series

$$|E_n\rangle = |a_n\rangle + \tau|b_n\rangle + \tau^2|c_n\rangle + \tau^3|d_n\rangle + \ ...\ ,$$
$$E_n = E_n^0 + \tau\, E_n^1 + \tau^2\, E_n^2 + \tau^3\, E_n^3 + \ ...\ ,$$

where $H|E_n\rangle = E_n|E_n\rangle =$ can be written as

$$[H_0 + \tau\, V]|E_n\rangle = \left[\sum_{k\geq 0} \tau^k E_n^k\right]\ |E_n\rangle\,.$$

If states $|n\rangle$ are **non degenerate** ($E_n^0 \neq E_m^0$ if $|n\rangle \neq |m\rangle$) then $|a_n\rangle = |n\rangle$.

The comparison of

$$H|E_n\rangle = H_0|n\rangle + \tau H_0|b_n\rangle + \tau^2 H_0|c_n\rangle + \tau^3 H_0|d_n\rangle + \ ...+$$
$$+\tau V|n\rangle + \tau^2 V|b_n\rangle + \tau^3 V|c_n\rangle + \tau^4 V|d_n\rangle\ +\ ...\ =$$
$$= H_0|n\rangle + \tau\Big(H_0|b_n\rangle + V|n\rangle\Big) + \tau^2\Big(H_0|c_n\rangle + V|b_n\rangle\Big) + \tau^3\Big(H_0|d_n\rangle + V|c_n\rangle\Big)\ +\ ...$$

with

$$E_n|E_n\rangle = E_n^0|n\rangle + \tau\Big(E_n^0|b_n\rangle + E_n^1|n\rangle\Big) + \tau^2\Big(E_n^0|c_n\rangle + E_n^1|b_n\rangle + E_n^2|n\rangle\Big)$$
$$+\tau^3\Big(E_n^0|d_n\rangle + E_n^1|c_n\rangle + E_n^2|b_n\rangle + E_n^3|n\rangle\Big) + \ ...$$

gives

$$H_0|n\rangle = E_n^0|n\rangle\,,$$
$$H_0|b_n\rangle + V|n\rangle = E_n^0|b_n\rangle + E_n^1|n\rangle\,,$$
$$H_0|c_n\rangle + V|b_n\rangle = E_n^0|c_n\rangle + E_n^1|b_n\rangle + E_n^2|n\rangle\,,$$
$$H_0|d_n\rangle + V|c_n\rangle = E_n^0|d_n\rangle + E_n^1|c_n\rangle + E_n^2|b_n\rangle + E_n^3|n\rangle\,,$$

when considering the first, second and third order contributions. These can be rewritten as

$$\left(H_0 - E_n^0\right)|n\rangle = 0\,,$$
$$\left(H_0 - E_n^0\right)|b_n\rangle = -V|n\rangle + E_n^1|n\rangle\,,$$
$$\left(H_0 - E_n^0\right)|c_n\rangle = -V|b_n\rangle + E_n^1|b_n\rangle + E_n^2|n\rangle\,,$$
$$\left(H_0 - E_n^0\right)|d_n\rangle = -V|c_n\rangle + E_n^1|c_n\rangle + E_n^2|b_n\rangle + E_n^3|n\rangle\,,$$

Assume that $\langle k|b_n\rangle = 0$, $\langle k|c_n\rangle = 0$, ... for $k = n$, namely

$$|b_n\rangle = \sum_{i\neq n} B_i(n)|i\rangle\,, \quad |c_n\rangle = \sum_{i\neq n} C_i(n)|i\rangle\,, \quad \ldots\,.$$

This assumption is justified by the fact that, in the present expansion, the contribution of the state $|n\rangle$ to $|E_n\rangle$ is already taken into account thorugh $|a_n\rangle \equiv |n\rangle$. Then the previous equations state that

$$0 = -\langle n|V|n\rangle + E_n^1\,, \quad 0 = -\langle n|V|b_n\rangle + E_n^2\,, \quad 0 = -\langle n|V|c_n\rangle + E_n^3\,, \ldots$$

while multiplying by a generic $\langle k| \neq \langle n|$ gives

$$\left(E_k^0 - E_n^0\right)\langle k|b_n\rangle = -\langle k|V|n\rangle\,,$$
$$\left(E_k^0 - E_n^0\right)\langle k|c_n\rangle = -\langle k|V|b_n\rangle + E_n^1\langle k|b_n\rangle\,,$$
$$\left(E_k^0 - E_n^0\right)\langle k|d_n\rangle = -\langle k|V|c_n\rangle + E_n^1\langle k|c_n\rangle + E_n^2\langle k|b_n\rangle\,,$$

where $\langle k|b_n\rangle = B_k(n)$, $\langle k|c_n\rangle = C_k(n)$, $\langle k|d_n\rangle = D_k(n)$..., whose explicit form turns out to be

$$B_k(n) = -\frac{\langle k|V|n\rangle}{E_k^0 - E_n^0}\,, \qquad C_k(n) = -\frac{\langle k|V|b_n\rangle}{E_k^0 - E_n^0} + E_n^1\,\frac{\langle k|b_n\rangle}{E_k^0 - E_n^0}\,,$$
$$D_k(n) = -\frac{\langle k|V|c_n\rangle}{E_k^0 - E_n^0} + E_n^1\,\frac{\langle k|c_n\rangle}{E_k^0 - E_n^0} + E_n^2\,\frac{\langle k|b_n\rangle}{E_k^0 - E_n^0}\,, \ldots$$

In particular, $C_k(n)$ have the form

$$C_k(n) = \frac{1}{E_k^0 - E_n^0} \sum_{q \neq n} B_q(n) \left[-\langle k|V|q\rangle + E_n^1 \langle k|q\rangle \right]$$

$$C_k(n) = \frac{1}{E_n^0 - E_n^0} \left[\frac{\langle n|V|n\rangle\langle k|V|n\rangle}{(E_k^0 - E_n^0)} - \sum_{q \neq n} \frac{\langle q|V|n\rangle\langle k|V|q\rangle}{(E_q^0 - E_n^0)} \right] ,$$

Up to the second order, the energy reads

$$E_n = E_n^0 + \tau\, E_n^1 + \tau^2\, E_n^2 = E_0(m) + \tau\, \langle n|V|n\rangle + \tau^2\, \langle n|V|b_n\rangle + \ ...$$

$$= E_n^0 + \tau\, \langle n|V|n\rangle - \tau^2 \sum_{k \neq n} \frac{\langle n|V|k\rangle\langle k|V|n\rangle}{E_k^0 - E_n^0}$$

while the modified eigenstate is

$$|E_n\rangle = |n\rangle - \tau \sum_{k \neq n} \frac{\langle k|V|n\rangle}{E_k^0 - E_n^0} |k\rangle + \tau^2 \sum_{k \neq n} C_k(n)\, |k\rangle + \ ... \, .$$

Example. The harmonic oscillator is characterized by the spring constant $k_0 = m w_0^2$

$$H_0 = \frac{p^2}{2m} + \frac{k_0}{2} x^2 = \hbar w_0 \left(a^+ a + 1/2 \right) \ \Rightarrow \quad H = \frac{p^2}{2m} + \frac{k_0 + \epsilon k_0}{2} x^2 = H_0 + \epsilon\, \frac{k_0}{2} x^2$$

where $V = \epsilon\, k_0 x^2/2$ represents the perturbation for $\epsilon << 1$. Note that the role of τ can be confused with that of ϵ in the sense that instead of considering $\tau \in [0, 1]$ one can assume $\tau \in [0, \epsilon]$ with the new perturbation $V' = k_0 x^2/2$. This circumstance is rather frequent in physical applications. It allows to regard τ as a small quantity controlling in a direct way the magnitude of perturbation V. We exploit the fact that the spectrum of the new oscillator with $k = k_0 + \epsilon k_0$ is completely known to make a **comparison with the spectrum obtained in the perturbation scheme**. The spectrum of the new oscillator is

$$E_n(\tau) = \hbar w_0 \sqrt{(1+\tau)} \Big(n + 1/2 \Big)$$

being

$$k = (1+\tau) k_0 = m w^2(\tau) \ \rightarrow \ w(\tau) = \sqrt{\frac{k_0}{m}(1+\tau)} = w_0 \sqrt{(1+\tau)} \, .$$

$$x = \sqrt{\frac{\hbar}{2 m w_0}} \left(a^+ + a \right) \quad p = -i \sqrt{\frac{\hbar m w_0}{2}} \left(a - a^+ \right)$$

$$E_n^1 = \langle n|V'|n\rangle = \frac{k_0}{2} \langle n|x^2|n\rangle = \frac{k_0}{2}\, \frac{\hbar}{2 m w_0} \langle n| \Big((a^+)^2 + a^2 + 2n + 1 \Big) |n\rangle = \frac{\hbar\, w_0}{2} \, (\, n + 1/2) \, .$$

$$E_n^2 = -\sum_{k\neq n} \frac{|\langle k|V'|n\rangle|^2}{E_k^0 - E_n^0} = -\frac{|\langle n+2|V'|n\rangle|^2}{E_{n+2}^0 - E_n^0} - \frac{|\langle n-2|V'|n\rangle|^2}{E_{n-2}^0 - E_n^0}$$

$$= -\frac{\hbar^2 w_0^2}{16}\left[\frac{|\langle n+2|((a^+)^2 + a^2 + 2n+1)|n\rangle|^2}{2\hbar w_0} + \frac{|\langle n-2|((a^+)^2 + a^2 + 2n+1)|n\rangle|^2}{-2\hbar w_0}\right]$$

$$= -\frac{\hbar w_0}{32}\left|\sqrt{(n+1)(n+2)}\right|^2 + \frac{\hbar w_0}{32}\left|\sqrt{n(n-1)}\right|^2 = \frac{\hbar w_0}{32}(-4n-2) = -\frac{\hbar w_0}{8}(n+1/2)$$

$$E_n(\tau) = E_n^0 + \tau E_n^1 + \tau^2 E_n^2 + \ldots = \hbar w_0\left(n+1/2\right) + \tau\,\frac{\hbar\, w_0}{2}\,(\,n+1/2) - \tau^2\frac{\hbar w_0}{8}(n+1/2) + \ldots.$$

$$= \hbar\, w_0\left(n+1/2\right)\left(1 + \frac{\tau}{2} - \frac{\tau^2}{8} + \ldots.\right)$$

The first three terms of this expansion reproduce the Taylor expansion of the exact formula $E_n(\tau) = \hbar w_0\sqrt{(1+\tau)}\,(n+1/2)$.

15 Identical particles and Spin-Statistic Theorem

A wide class of physical systems is formed by many **identical particles**, namely, by particles that feature the same intrinsic properties (such as mass, charge and spin) and thus cannot be distinguished the one from the other. For example, many (identical) electrons aggregate to form atoms. Likewise, atomic nuclei appear to be formed by many protons while, at a larger space scale, the so-called Bose-Einstein condensates are typically formed by $\approx 10^6$ atoms of the same atomic species. More in general, macroscopic physical systems typically involving particle numbers of the order of the Avogadro number (such as gases, liquids, and solids but also superconductors, superfluids or magnetic systems) are formed by identical particles whose collective behavior and properties are investigated within classical and quantum Statistical Mechanics.

Classically, identical particles can be distinguished by means of their positions and, more in general, due to their different trajectories. Conversely, in quantum-mechanical systems, the fact that the aggregation of many identical particles takes place in microscopic regions and involves complex many-body interactions makes it impossible to identify individual particles. The circumstance that particles are indistinguishable has led to discover **new fundamental effects that have no classical counterpart**.

Symmetric many-particle Hamiltonians. The Hamiltonian of many (identical) particles necessarily exhibits a **symmetric form** well represented by

$$H = H(\mathbf{x}_1, \mathbf{x}_2, \, ... \, \mathbf{x}_N) = H_0 + H_U = \sum_{j=1}^{N} \left[\frac{\mathbf{p}_j^2}{2m} + V(\mathbf{x}_j) \right] + \frac{1}{2} \sum_{j=1}^{N} \sum_{\ell=1 \, (\neq j)}^{N} U(|\mathbf{x}_j - \mathbf{x}_\ell|) \, .$$

Two-body interactions are described by $U(|\mathbf{x}_j - \mathbf{x}_\ell|)$ in H_U. The exchange of particle i and n (consisting in the position-vector exchange $\mathbf{x}_i \leftrightarrow \mathbf{x}_n$ and entailing the momentum-vector exchange $\mathbf{p}_i \leftrightarrow \mathbf{p}_n$) leaves H unchanged. In general, Hamiltonian H can include terms depending on spin operators. This is the case, for example, when atomic electrons undergo the action of an external magnetic field $\mathbf{B}$. Spin-orbit interactions (describing the coupling of spin magnetic moments of electrons with the magnetic field caused by the atomic-nucleus electric current) also involve the spin operators relevant to each electron. A many-electron Hamiltonian describing such interactions is

$$H_{so}(B) = H \, + H_{so} \, + H_B = H_0 \, + H_U \, + H_{so} \, + H_B \, ,$$

with

$$H_B = -\frac{e\mathbf{B}}{2m_e c} \cdot \sum_i (\, \mathbf{L}_i + 2\mathbf{S}_i \,) \, , \qquad H_{so} = + \sum_i \frac{1}{2m_e^2 c^2 r_i} \frac{dV(r_i)}{dr_i} \, (\mathbf{L}_i \cdot \mathbf{S}_i) \, ,$$

where $V(\mathbf{x}_i) = V(r_i)$ $(|\mathbf{x}_i| = r_i)$ in H_0 is the atomic-nucleus energy potential, $\mathbf{L}_i = \mathbf{x}_i \wedge \mathbf{p}_i$ is the angular-momentum operator, and $\mathbf{S}_i$ is the spin operator for i-th electron. An even more

general model which also includes spin-spin interactions could be represented by

$$H_{so+ss}(B) = H_{so}(B) + \frac{1}{2}\sum_n \sum_{n\neq i} \eta_{in}\,(\mathbf{S}_i \cdot \mathbf{S}_n)\,, \qquad r_i = |\mathbf{x}|\,,$$

where $\eta_{ij} = \eta_{ji}$. The exchange of two electrons in $H_{so+ss}(B)$ thus requires the operator exchange $\mathbf{S}_i \leftrightarrow \mathbf{S}_n$ in addition to $\mathbf{x}_i \leftrightarrow \mathbf{x}_n$ and $\mathbf{p}_i \leftrightarrow \mathbf{p}_n$ which entails as well $\mathbf{L}_i \leftrightarrow \mathbf{L}_n$. Condition $\eta_{ij} = \eta_{ji}$ once more ensures that $H_{so+ss}(B)$ is symmetric.

15.1 Symmetrized states

Construction of symmetrized wave functions. Consider a generic quantum state of N identical particle. The exchange of two particles, represented by

$$\psi = \Psi(\mathbf{x}_1, ...\,\mathbf{x}_j, ...\,\mathbf{x}_i, ...\,\mathbf{x}_N) \;\;\Rightarrow\;\; \psi' = P_{ji}\,\psi(\mathbf{x}_1 ...\,\mathbf{x}_j, ...\,\mathbf{x}_i, ...) = \psi(\mathbf{x}_1, ...\,\mathbf{x}_i, ...\,\mathbf{x}_j, ...\,\mathbf{x}_N)\,,$$

is the basic process that allows one to construct wavefunctions exhibiting a defined symmetry character. The **permutation operator** P_{ji} effects the exchange of particles j and i by permuting the relevant positions $\mathbf{x}_j \leftrightarrow \mathbf{x}_i$.

Note that in case ψ has the factorized form $\psi = \psi_{\alpha_1}(\mathbf{x}_1)\psi_{\alpha_2}(\mathbf{x}_2)\,...\,\psi_{\alpha_N}(\mathbf{x}_N)$ where each single-particle wave function $\psi_{\alpha_i}(\mathbf{x}_i)$ is labeled by quantum numbers α_i associated to the particle with position $\mathbf{x}_i$ then, in the new state ψ', labels α_j and α_i will be associated to particles with position $\mathbf{x}_i$ and $\mathbf{x}_j$, respectively.

The repeated action of permutation operators P_{ji} allows one to construct both **symmetric and antisymmetric wave functions**. For example, with $N = 3$ particles

$$\Phi_s(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) = \frac{1}{\sqrt{3!}}\Big(1 + P_{13}P_{12} + P_{23}P_{12} + P_{12} + P_{23} + P_{13}\Big)\,\psi(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)$$

$$= \frac{\psi(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) + \psi(\mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_1) + \psi(\mathbf{x}_3, \mathbf{x}_1, \mathbf{x}_2) + \psi(\mathbf{x}_2, \mathbf{x}_1, \mathbf{x}_3) + \psi(\mathbf{x}_1, \mathbf{x}_3, \mathbf{x}_2) + \psi(\mathbf{x}_3, \mathbf{x}_2, \mathbf{x}_1)}{\sqrt{3!}}$$

Likewise

$$\Phi_a(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) = \frac{1}{\sqrt{3!}}\Big(1 + P_{13}P_{12} + P_{23}P_{12} - P_{12} - P_{23} - P_{13}\Big)\,\psi(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)$$

$$= \frac{\psi(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) + \psi(\mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_1) + \psi(\mathbf{x}_3, \mathbf{x}_1, \mathbf{x}_2) - \psi(\mathbf{x}_2, \mathbf{x}_1, \mathbf{x}_3) - \psi(\mathbf{x}_1, \mathbf{x}_3, \mathbf{x}_2) - \psi(\mathbf{x}_3, \mathbf{x}_2, \mathbf{x}_1)}{\sqrt{3!}}$$

where each term involving an odd number of permutations now involves an extra factor (-1). One easily checks that whenever two any position vectors $\mathbf{x}_i$ and $\mathbf{x}_j$ are exchanged in $\Phi_s(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)$ then state Φ_s is unchanged. In general, one has

$$\Phi_s(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) = \Phi_s(\mathbf{x}_i, \mathbf{x}_j, \mathbf{x}_k)$$

where $i \neq j \neq k \neq i$. With Φ_a, instead, one has

$$\left\{ \begin{array}{rcl} \Phi_a(\mathbf{x}_2, \mathbf{x}_1, \mathbf{x}_3) & = & -\Phi_a(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) \\ \Phi_a(\mathbf{x}_1, \mathbf{x}_3, \mathbf{x}_2) & = & -\Phi_a(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) \\ \Phi_a(\mathbf{x}_3, \mathbf{x}_2, \mathbf{x}_1) & = & -\Phi_a(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) \end{array} \right. .$$

The generalization of such formulas to the N particle case gives the completely symmetric/antisymmetric wavefunctions Φ_s and Φ_a defined as

$$\Phi_s(\mathbf{x}_1, \mathbf{x}_2, ... \mathbf{x}_N) = \frac{1}{\sqrt{N!}} \sum_P P\, \psi(\mathbf{x}_1, \mathbf{x}_2, ... \mathbf{x}_N) \tag{107}$$

$$\Phi_a(\mathbf{x}_1, \mathbf{x}_2, ... \mathbf{x}_N) = \frac{1}{\sqrt{N!}} \sum_P (-1)^{\eta_P} P\, \psi(\mathbf{x}_1, \mathbf{x}_2, ... \mathbf{x}_N) \tag{108}$$

where $P : 1, 2, ... N \Rightarrow P(1), P(2), ... P(N)$ represents one of the $N!$ permutations that exchanges N different symbols. Any permutation P is realized by combining the action $P_{ij}\, \psi(\mathbf{x}_1, ... \mathbf{x}_j, ... \mathbf{x}_i, ...) := \psi(\mathbf{x}_1, ... \mathbf{x}_i, ... \mathbf{x}_j, ...)$ of different exchange operators P_{ij} as shown by the two preceding examples involving three particles. Formally, this property amounts to stating that the set of all P forms a discrete group (the so-called symmetric group). The P-dependent index η_P occurring in Φ_a is such that $\eta_P = 0$ or $\eta_P = 1$ if the number of two-symbol interchanges generating a permutation P is **even** or **odd**, respectively. Concluding, any quantum state of N identical particles can be recast into one of the two forms Φ_s and Φ_a. One easily checks that the symmetrized wave functions defined above are such that

$$\begin{aligned} \Phi_s(\mathbf{x}_1, ... \mathbf{x}_j, ... \mathbf{x}_i, ... \mathbf{x}_N) &= +\Phi_s(\mathbf{x}_1, ... \mathbf{x}_i, ... \mathbf{x}_j, ... \mathbf{x}_N)\,, \\ \Phi_a(\mathbf{x}_1, ... \mathbf{x}_j, ... \mathbf{x}_i, ... \mathbf{x}_N) &= -\Phi_a(\mathbf{x}_1, ... \mathbf{x}_i, ... \mathbf{x}_j, ... \mathbf{x}_N)\,, \end{aligned}$$

when two any particles are exchanged in the symmetric and antisymmetric case, respectively.

Antisymmetric states thus feature the **remarkable property** that $\Phi_a = 0$ whenever two coordinates $\mathbf{x}_i$ and $\mathbf{x}_j$ coincide. Physically, this means that the many-particle probability density $\rho = |\Phi_a(\mathbf{x}_1, ..., \mathbf{x}_N)|^2$ is zero as soon as one considers spatial configurations where two (or more) particles are placed at the same spatial position. Antisymmetric states then represent a class of particles that systematically tend to avoid the one with the other. Such a macroscopic **repulsive effect** is totally absent in the case of particles described by symmetric states. The probability density $|\Phi_s(\mathbf{x}_1, ..., \mathbf{x}_N)|^2$ never vanishes when two coordinates $\mathbf{x}_i$ and $\mathbf{x}_j$ coincide. Conversely, in the case when Φ_s describes the ground state, one can show that $|\Phi_s|^2$ reaches a maximum when all the particles are concentrated at the same point. In the previous example with three particles, if $x_1 \simeq x_2 \simeq x_3$ then $\psi(\mathbf{x}_i, \mathbf{x}_j, \mathbf{x}_k) \simeq \psi(\mathbf{x}_1, \mathbf{x}_1, \mathbf{x}_1)$. So no destructive intereference effects due to the possibly different phases of complex quantities $\psi(\mathbf{x}_i, \mathbf{x}_j, \mathbf{x}_k)$ are present and one obtains the maximum value of $|\Phi_s|^2$

$$|\Phi_s(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)|^2 \simeq 6|\psi(\mathbf{x}_1, \mathbf{x}_1, \mathbf{x}_1)|^2.$$

Symmetrization of energy eigenstates. Based on such a symmetrization scheme one easily constructs symmetric and antisymmetric states **associated to the same energy eigenvalue** E. To this end it is sufficient to apply all the possible permutations P to the energy eigenstates ψ_E of the eigenvalue problem

$$H\,\psi_E(\mathbf{x}_1, \mathbf{x}_2, \ldots \mathbf{x}_N) = E\,\psi_E(\mathbf{x}_1, \mathbf{x}_2, \ldots \mathbf{x}_N)\,.$$

For any permutation P one finds

$$PH\,\psi_E = HP\,\psi_E\,, \quad PH\,\psi_E = E\,P\psi_E \;\Rightarrow\; HP\,\psi_E = E\,P\psi_E$$

since H has a **symmetric form**. Owing to the linear character of definitions (107) and (108) stating that wavefunctions Φ_s and Φ_a are linear superpositions of states $P\,\psi(\mathbf{x}_1, \mathbf{x}_2, \ldots \mathbf{x}_N)$, it follows that

$$H\,\Phi_{sE} = E\,\Phi_{sE} \quad \textbf{and} \quad H\,\Phi_{aE} = E\,\Phi_{aE}\,.$$

Then, given a generic stationary solution Ψ_t of the time-dependent Schrödinger problem

$$i\hbar\partial_t\Psi_t = H\,\Psi_t\,, \qquad \Psi_t = e^{-iEt/\hbar}\,\psi_E(\mathbf{x_1}, \mathbf{x_2}, \ldots \mathbf{x_N})\,,$$

one easily obtains the **symmetrized stationary solutions**

$$\Psi_{sE}(t) = e^{-iEt/\hbar}\,\Phi_{sE}(\mathbf{x_1}, \mathbf{x_2}, \ldots \mathbf{x_N})\,, \quad \Psi_{aE}(t) = e^{-iEt/\hbar}\,\Phi_{aE}(\mathbf{x_1}, \mathbf{x_2}, \ldots \mathbf{x_N})\,.$$

Notice that, if the unsymmetrized wavefunction ψ, in addition to energy E, is labeled by other quantum numbers λ_1, λ_2 ... then Φ_a and Φ_s will feature the same list of quantum numbers.

Time propagation and symmetrization. We have showed that stationary solutions obtained from the energy-eigenvalue equation are easily symmetrized. In this case, it is rather obvious that states $\Phi_{aE}(t) = e^{-iEt/\hbar}\Phi_{aE}$ and $\Phi_{sE}(t) = e^{-iEt/\hbar}\Phi_{sE}$, solutions of $i\hbar\partial_t\psi_t = H\,\psi_t$, maintain a definite symmetry character at any time t. It is more interesting to consider the time evolution of states Φ_a or Φ_s, endowed with a definite symmetry character, corresponding an arbitrary state $\psi(\mathbf{x}_1, \ldots \mathbf{x}_N)$ which does not diagonalize the Hamiltonian. One easily proves that the time propagation of a symmetric (antisymmetric) state generates a symmetric (antisymmetric) state. Let $\Phi_a(\mathbf{r}_1, \ldots \mathbf{r}_N, t)$ be the N-particle state at time t. Then

$$\Phi_a(\mathbf{r}_1, \ldots \mathbf{r}_N, t+\epsilon) = e^{-i\epsilon H/\hbar}\,\Phi_a(\mathbf{r}_1, \ldots \mathbf{r}_N, t) = \Phi_a(\mathbf{r}_1, \ldots \mathbf{r}_N, t) - i\frac{\epsilon}{\hbar}\,H\Phi_a(\mathbf{r}_1, \ldots \mathbf{r}_N, t)$$

is the N-particle state at time $t+\epsilon$ with $\epsilon << t$. The new state $H\Phi_a(\mathbf{r}_1, \ldots \mathbf{r}_N, t)$ is still antisymmetric because H is a symmetric operator. Then $\Phi_a(\mathbf{r}_1, \ldots \mathbf{r}_N, t+\epsilon)$ is antisymmetric. If one considers instead a symmetric state $\Phi_s(\mathbf{r}_1, \ldots \mathbf{r}_N, t)$ the same scheme leads to observe that $\Phi_s(\mathbf{r}_1, \ldots \mathbf{r}_N, t+\epsilon)$ will be symmetric because $H\,\Phi_s(\mathbf{r}_1, \ldots \mathbf{r}_N, t)$ turns out to be symmetric.

15.2 Bosons, fermions and exclusion principle

Separable Hamiltonians. Let us consider the rather common case when H can be separated in N independent Hamiltonians H_i. To this end one must ignore both the two-body potential U in H_U and the spin-spin interactions depending on $\vec{S}_i \cdot \vec{S}_j$ contained in the general model $H_{so+ss}(B)$. Then, based on the previous examples, Hamiltonian H could have the form

$$H = \sum_i H_i\,, \qquad H_i = H_0(\mathbf{x}_i) + \gamma\,\vec{B}\cdot(\vec{L}_i + 2\vec{S}_i) + c_i\,\vec{L}_i\cdot\vec{S}_i \tag{109}$$

which includes the terms relevant to the spin-orbit coupling and to the spin-B coupling. Owing to their perturbation character, the contributions of the terms that one neglects can be included a posteriori by applying the stationary perturbation theory to the unperturbed energy states of H_0 and to the relevant energy spectrum. States ψ of H_0 are **completely factorized**

$$\Psi = \Psi(\mathbf{x}_1, \mathbf{x}_2, \ldots \mathbf{x}_N) = \prod_{j=1}^{N} \Psi_{\alpha_j}(\mathbf{x}_j)\,, \qquad H_i\,\Psi_{\alpha_i}(\mathbf{x}_i) = E(\alpha_i)\,\Psi_{\alpha_i}(\mathbf{x}_i)\,. \tag{110}$$

Symbol Ψ has been used in place of ψ to emphasize the fact that such states are **spinor states** (see equation (94)) due to the presence of $\vec{S}_i$-dependent terms in the single-particle Hamiltonian H_i. State Ψ_{α_j} (where $\alpha_j = (v_j, \sigma_j)$) has the form

$$\Psi_{\alpha_j}(\mathbf{x}_j) = \Psi_{v_j\sigma_j}(\mathbf{x}_j) = \sum_{\mu_j=\pm 1/2} C_{\mu_j}(\sigma_j)\,\psi_{v_j}^{\mu_j}(\mathbf{x}_j)\;|\mu_j\rangle$$

$$= C_{+1/2}(\sigma_j)\,\psi_{v_j}^{+1/2}(\mathbf{x}_j)\left|+\tfrac{1}{2}\right\rangle_j + C_{-1/2}(\sigma_j)\,\psi_{v_j}^{-1/2}(\mathbf{x}_j)\left|-\tfrac{1}{2}\right\rangle_j$$

$$= C_{+1/2}(\sigma_j)\,\psi_{v_j}^{+}(\mathbf{x}_j)\begin{bmatrix}1\\0\end{bmatrix}_j + C_{-1/2}(\sigma_j)\,\psi_{v_j}^{-}(\mathbf{x}_j)\begin{bmatrix}0\\1\end{bmatrix}_j = \begin{bmatrix} f_{\alpha_j}^{+}\\ f_{\alpha_j}^{-}\end{bmatrix}$$

Note that the spin state $|s, m_s\rangle$ defined by formulas (91) has been replaced with $|\sigma_j\rangle$ (where $s = 1/2$ is implied) in which index j labeling $\sigma_j = \pm 1/2$ allows one to associate state $|\sigma_j\rangle$ to a specific spin operator $\vec{S}_j$. Within such a notation the third-component eigenvalue equation takes the form $S_{j3}\,|\sigma_j\rangle = \hbar\sigma_j\,|\sigma_j\rangle$.

Index $\alpha_j = (v_j, \sigma_j)$ includes the information relevant to the spinor character of Ψ_{α_j} and the fact that it diagonalizes spin-dependent terms. In the presence of a radial potential $V(\mathbf{x}_i) = V(|\mathbf{x}_i|)$, index $v_i = (n_i, \ell_i, m_i)$, occurring in single-particle states Ψ_{α_j}, represents the quantum numbers relevant to the contribution of the radial part of Hamiltonian $H_0(\mathbf{x}_i)$, to operator $\mathbf{L}_i^2$, and to the third component of angular momentum $\mathbf{L}_i$, respectively.

Comment. Spinor states involve a more general definition of scalar product which takes into account the presence of spin components and the fact that $\langle\mu_j|\sigma_j\rangle = \delta_{\mu_j\sigma_j}$. More explicitly

$$\left\langle +\tfrac{1}{2}\middle|-\tfrac{1}{2}\right\rangle_j = [1,0]\cdot\begin{bmatrix}0\\1\end{bmatrix} = 0\,, \quad \left\langle -\tfrac{1}{2}\middle|+\tfrac{1}{2}\right\rangle_j = [0,1]\cdot\begin{bmatrix}1\\0\end{bmatrix} = 0\,, \quad \left\langle \pm\tfrac{1}{2}\middle|\pm\tfrac{1}{2}\right\rangle_j = 1$$

Then, the normalization condition $(\Psi_{\alpha_j}, \Psi_{\alpha_j}) = 1$ is achieved by defining the scalar product of spinor states

$$\Big(\Psi_{\alpha_j}, \Psi_{\alpha_j}\Big) = \left(\sum_{\nu_j} C_{\nu_j}(\sigma_j)\, \psi^{\nu_j}_{\alpha_j}(\mathbf{x}_j)\, |\nu_j\rangle\,,\, \sum_{\mu_j} C_{\mu_j}(\sigma_j)\, \psi^{\mu_j}_{\alpha_j}(\mathbf{x}_j)\, |\mu_j\rangle\right)$$

$$= \sum_{\nu_j}\sum_{\mu_j} C^*_{\nu_j}(\sigma_j) C_{\mu_j}(\sigma_j)\, \langle\nu_j|\mu_j\rangle\, \big(\psi^{\nu_j}_{\alpha_j}(\mathbf{x}_j),\, \psi^{\mu_j}_{\alpha_j}(\mathbf{x}_j)\big) = \sum_{\mu_j} |C_{\mu_j}(\sigma_j)|^2\, \big(\psi^{\mu_j}_{\alpha_j}(\mathbf{x}_j),\, \psi^{\mu_j}_{\alpha_j}(\mathbf{x}_j)\big)$$

$$= \sum_{\mu_j} |C_{\mu_j}(\sigma_j)|^2 \int d^3x\, |\psi^{\mu_j}_{\alpha_j}(\mathbf{x}_j)|^2 = |C_{-1/2}(\sigma_j)|^2 + |C_{+1/2}(\sigma_j)|^2\,.$$

Example 1. Consider the case when **spin-orbit interactions** are **neglected** in the single-electron hamiltonian (109) ($c_i = 0$). Then H_i is formed by the two commuting parts $H_0(\mathbf{x}_i) + \gamma\vec{B}\cdot\vec{L}_i$ (depending on coordinates) and $\gamma\vec{B}\cdot\vec{S}_i = B_1 S_{1i} + B_2 S_{2i} + B_3 S_{3i}$ (depending on spin). In this case the relevant eigensates have the **factorized form**

$$\Psi_{\alpha_i}(\mathbf{x}_i) = \psi_{v_i}(\mathbf{x}_i)\, |f(\sigma_i)\rangle$$

where the spin state is independent from the coordinate-dependent part ψ_{v_i}. The spin part $|f(\sigma_i)\rangle$ is the superposition

$$|f(\sigma_i)\rangle = C_{-\frac{1}{2}}(\sigma_i)\left|-\tfrac{1}{2}\right\rangle_i + C_{+\frac{1}{2}}(\sigma_i)\left|+\tfrac{1}{2}\right\rangle_i\,.$$

Example 2. If $\vec{B}\cdot\vec{S} \equiv B_3\, S_{3i}$ ($\vec{B}$ is parallel to axis z), the single-electron hamiltonian H_i essentially reduces to that described by Hamiltonian (102). Then $C_{\mu_i}(\sigma_i) = \delta_{\mu_i\sigma_i}$ giving

$$\Psi_{\alpha_i}(\mathbf{x}_i) = \psi_{v_i}(\mathbf{x}_i)\, |\sigma_i\rangle\,, \qquad \sigma_i = \pm\frac{1}{2}\,,\;\; S_{3i}\, |\sigma_i\rangle = \hbar\sigma_i\, |\sigma_i\rangle\,.$$

Similar to state (103) (see also equations (104) and (105)), state $\Psi_{\alpha_j}(\mathbf{x}_j)$ diagonalizes Hamiltonian H_j containing $B_3\, S_3$.

Example 3. As a final example, assume that H_i contains the non diagonal term $B_1\, S_{1i}$ (with $B_2 = B_3 = 0$). Then $C_{\mu_i}(\sigma_i) \neq \delta_{\mu_i\sigma_i}$. One has ($\sigma_i = \pm 1/2$)

$$\Psi_{\alpha_i}(\mathbf{x}_i) = \psi_{v_i}(\mathbf{x}_i)\, |f(\sigma_i)\rangle\,, \qquad |f(\sigma_i)\rangle = \frac{1}{\sqrt{2}}\Big(\left|-\tfrac{1}{2}\right\rangle + 2\sigma_i \left|+\tfrac{1}{2}\right\rangle\Big)\,,\;\; S_{1\,i}\, |f(\sigma_i)\rangle = \hbar\sigma_i\, |f(\sigma_i)\rangle\,.$$

Symmetric wavefunctions and bosons. Based on formula (107), the symmetric wavefunction corresponding to a factorized state $\Psi = \Psi_{\alpha_1}(\mathbf{x}_1)\Psi_{\alpha_2}(\mathbf{x}_2)\,...\,\Psi_{\alpha_N}(\mathbf{x}_N)$ is given by

$$\Phi_s(\mathbf{x}_1,\, ...\,\mathbf{x}_j,\, ...\,\mathbf{x}_N) = \frac{1}{\sqrt{N!}}\sum_P P\, [\Psi_{\alpha_1}(\mathbf{x}_1)\Psi_{\alpha_2}(\mathbf{x}_2)\,...\,\Psi_{\alpha_N}(\mathbf{x}_N)] \tag{111}$$

if the N particles of Ψ **are associated with** N **different quantum numbers** α_j. Conversely, if $\alpha_1 = \alpha_2 = ... = \alpha_N$, state (111) reduces to the trivial state

$$\Phi_s(\mathbf{x}_1,\, \mathbf{x}_2,\, ...\, \mathbf{x}_N) = \Psi_{\alpha_1}(\mathbf{x}_1)\, \Psi_{\alpha_1}(\mathbf{x}_2) ... \Psi_{\alpha_1}(\mathbf{x}_N)\,,$$

which is symmetric by construction. Intermediate cases where N_1 particles are in state $\Psi_{\alpha_1}(\mathbf{x})$, N_2 particles are in state $\Psi_{\alpha_2}(\mathbf{x})$, ... and N_r particles are in state $\Psi_{\alpha_r}(\mathbf{x})$, and thus the number of different quantum numbers α_j is smaller than N, are still represented by state (111). Note that $\sum_{s=1}^{r} N_s = N$. In such cases, after introducing the information that N_1 symbols α_j are equal to α_1, N_2 symbols α_j are equal to α_2, ... and N_r symbols α_j are equal to α_r, state (111) reduces to

$$\Phi_s(\mathbf{x}_1,\, ...\, \mathbf{x}_j,\, ...\, \mathbf{x}_N) = \frac{\sqrt{N_1!N_2!..N_r!}}{\sqrt{N!}} \times$$

$$\sum_P{}^{*}\, P \left[\prod_{p_1=1}^{N_1} \Psi_{\alpha_1}(\mathbf{x}_{p_1}) \prod_{p_2=N_1+1}^{N_1+N_2} \Psi_{\alpha_2}(\mathbf{x}_{p_2}) ... \prod_{p_r=N_1+...N_{r-1}+1}^{N_1+N_2+...N_r} \Psi_{\alpha_r}(\mathbf{x}_{p_r}) \right] \tag{112}$$

where the symbol $*$ in $\sum_P^*$ means that permutations P exchanging position $\mathbf{x}_j$ of the group of wavefunctions $\Psi_{\alpha_j}(\mathbf{x}_j)$ with the same quantum number α_j must be ignored.

Example. Consider the simple 3-boson case $\Psi = \Psi_{\alpha_1}(\mathbf{x}_1)\Psi_{\alpha_2}(\mathbf{x}_2)\Psi_{\alpha_3}(\mathbf{x}_3)$ with $\alpha_1 = \alpha_2 \neq \alpha_3$. Then, state Φ_s given by

$$\frac{\Psi(\mathbf{x}_1,\mathbf{x}_2,\mathbf{x}_3) + \Psi(\mathbf{x}_2,\mathbf{x}_3,\mathbf{x}_1) + \Psi(\mathbf{x}_3,\mathbf{x}_1,\mathbf{x}_2) + \Psi(\mathbf{x}_2,\mathbf{x}_1,\mathbf{x}_3) + \Psi(\mathbf{x}_1,\mathbf{x}_3,\mathbf{x}_2) + \Psi(\mathbf{x}_3,\mathbf{x}_2,\mathbf{x}_1)}{\sqrt{3!}}$$

is no longer formed by six different states being

$$\Psi(\mathbf{x}_1,\mathbf{x}_2,\mathbf{x}_3) = \Psi_{\alpha_1}(\mathbf{x}_1)\Psi_{\alpha_1}(\mathbf{x}_2)\Psi_{\alpha_3}(\mathbf{x}_3) = \Psi(\mathbf{x}_2,\mathbf{x}_1,\mathbf{x}_3)\,,$$

$$\Psi(\mathbf{x}_2,\mathbf{x}_3,\mathbf{x}_1) = \Psi_{\alpha_1}(\mathbf{x}_2)\Psi_{\alpha_1}(\mathbf{x}_3)\Psi_{\alpha_3}(\mathbf{x}_1) = \Psi(\mathbf{x}_3,\mathbf{x}_2,\mathbf{x}_1)\,,$$

$$\Psi(\mathbf{x}_3,\mathbf{x}_1,\mathbf{x}_2) = \Psi_{\alpha_1}(\mathbf{x}_3)\Psi_{\alpha_1}(\mathbf{x}_1)\Psi_{\alpha_3}(\mathbf{x}_2) = \Psi(\mathbf{x}_1,\mathbf{x}_3,\mathbf{x}_2)\,.$$

Formula (112) confirms this result giving the wavefunction

$$\Phi_s = \frac{\sqrt{2!}}{\sqrt{3!}} \left[\Psi(\mathbf{x}_1,\mathbf{x}_2,\mathbf{x}_3)+\Psi(\mathbf{x}_2,\mathbf{x}_3,\mathbf{x}_1)+\Psi(\mathbf{x}_3,\mathbf{x}_1,\mathbf{x}_2)\right] = \sum_{k=1}^{3} \sum_{n,m\neq k} \frac{\Psi_{\alpha_1}(\mathbf{x}_n)\Psi_{\alpha_1}(\mathbf{x}_m)\Psi_{\alpha_3}(\mathbf{x}_k)}{2\sqrt{3}}$$

Identical particles described by a symmetric wavefunction $\Phi_s(\mathbf{x}_1, \mathbf{x}_2,\, ...\, \mathbf{x}_N)$ are called **bosons** and obey the **Bose-Einstein statistic**. The **distinctive feature** of aggregates of bosons is

evidenced by formula (112), in which r groups of N_s particles, with $\sum_{s=1}^{r} N_s = N$, feature the same quantum number α_j and thus are in the same physical state $\Psi_{\alpha_j}(\mathbf{x}_j)$.

Factor $\sqrt{N_1!N_2!...N_r!}/\sqrt{N!}$, in addition to ensure the correct normalization of Φ_s, embodies this information. More specifically, it evidences the fact that, while in state (111) all symbols α_j are different and thus $N!$ different contributions $P[\Psi_{\alpha_1}(\mathbf{x}_1)\Psi_{\alpha_2}(\mathbf{x}_2)...\Psi_{\alpha_N}(\mathbf{x}_N)]$ are superimposed, in state (112) a **smaller number** (represented by $N!/(N_1!...N_r!)$) of such contributions is involved owing to presence of r groups of bosons with the same quantum number α_j. Hence the number of bosons that can occupy a given state $\Psi_{\alpha_j}(\mathbf{x}_j)$ is in principle **unlimited**.

Owing to this crucial property a gas of bosons exhibits, at zero temperature, the so-called **Bose-Einstein condensation**. This phenomenon is characterized by the fact that a macroscopic number of particles ($\approx 10^6$) collapses in a **single quantum state**, the one corresponding to the single-particle wavefunction with zero momentum.

Antisymmetric wavefunctions and fermions. This is obtained by applying formula (108) to the unsymmetrized state $\Psi = \Psi_{\alpha_1}(\mathbf{x}_1)\Psi_{\alpha_2}(\mathbf{x}_2)...\Psi_{\alpha_N}(\mathbf{x}_N)$. The results is the well-known formula

$$\Phi_a(\mathbf{x}_1, ...\mathbf{x}_j, ...\mathbf{x}_N) = \frac{1}{\sqrt{N!}} det \begin{bmatrix} \Psi_{\alpha_1}(\mathbf{x}_1) & \Psi_{\alpha_1}(\mathbf{x}_2) & ... & ... & \Psi_{\alpha_1}(\mathbf{x}_N) \\ \Psi_{\alpha_2}(\mathbf{x}_1) & \Psi_{\alpha_2}(\mathbf{x}_2) & ... & ... & \Psi_{\alpha_2}(\mathbf{x}_N) \\ ... & ... & ... & ... & ... \\ ... & ... & ... & ... & ... \\ \Psi_{\alpha_N}(\mathbf{x}_1) & \Psi_{\alpha_N}(\mathbf{x}_2) & ... & ... & \Psi_{\alpha_N}(\mathbf{x}_N) \end{bmatrix} \tag{113}$$

entailing that $\Phi_a = 0$ **whenever two particles have** $\alpha_i = \alpha_j$, namely, **they exhibit the same quantum numbers**.

Wavefunction $\Phi_a(\mathbf{x}_1, ...\mathbf{x}_j, ...\mathbf{x}_N)$, called Slater's determinat, describes systems of identical particles called **fermions** that obey the **Fermi-Dirac statistic**. Aggregates of identical fermions are thus characterized by the property that two any fermions cannot have the same quantum numbers α_j, namely, each fermion always pertains to a different quantum state $\Psi_{\alpha_j}(\mathbf{x}_j)$. Hence the number of fermions that can occupy a given state $\Psi_{\alpha_j}(\mathbf{x}_j)$ is limited to 0 or 1. This result is known as the **Pauli exclusion principle**.

The experimental observations confirm that all known particles belong to one of such two classes and thus particle aggregates are always described either by completely symmetric or by completely antisymmetric wavefunctions. The spin-statistics theorem completes this scenerio stating that bosons exhibit an integer spin while fermions feature half-integer spin.

Spin-Statistic theorem [3], [9]. A system of identical particles with **integer** spin is characterized by a wavefunction which is **symmetric** under the exchange $\mathbf{x}_j \leftrightarrow \mathbf{x}_i$ of two any particles $\Psi(\mathbf{x}_1, ...\mathbf{x}_j, ...\mathbf{x}_i, ...\mathbf{x}_N) = +\Psi(\mathbf{x}_1, ...\mathbf{x}_i, ...\mathbf{x}_j, ...\mathbf{x}_N)$. These particles are said to obey **Bose-Einstein statistics** and are called **bosons**. A system of identical particles with **half-integer** spin is characterized by a wavefunction which is **antisymmetric** under the exchange $\mathbf{x}_j \leftrightarrow \mathbf{x}_i$ of two any particles $\Psi(\mathbf{x}_1, ...\mathbf{x}_j, ...\mathbf{x}_i, ...\mathbf{x}_N) = -\Psi(\mathbf{x}_1, ...\mathbf{x}_i, ...\mathbf{x}_j, ...\mathbf{x}_N)$. These particles are said to obey **Fermi-Dirac statistics** and are called **fermions**.

Comment 1. Within nonrelativistic quantum mechanics this Theorem has been introduced as an empirical law (symmetrization principle [3]) widely supported by experimental observations. Its rigorous proof is given within relativistic Quantum Field Theory (see, for example, reference [9]).

Comment 2. Aggregates of particles exhibiting a stable character because their are sufficiently tightly bound (namely the interactions among particles are strong) can be seen as **composite particles** with a definite statistical character. Atoms are stable aggregates composed by electrons, protons and neutrons (all having spin 1/2). Atoms such as ^{87}Rb and ^{41}K are isotopes whose statistical character depends on the number of neutrons. Since the Rb atomic number is 37 and the ^{87}Rb nucleons are 87 then 50 neutrons are contained in its nucleus. Likewise, the K atomic number is 19 so that ^{41}K has 22 neutrons. The resulting total spin is thus integer. These isotopes have been used very recently to realize Bose-Einstein condensates. The K isotope ^{40}K instead exhibits a fermionic character. This isotope has been used to realize boson-fermion mixtures together with ^{87}Rb. Other well-known examples of bosonic and fermionic particles in Condensed Matter physics are ^{4}He and ^{3}He atoms, respectively.

Symmetrized spin wavefunctions. Two-particle states gives the possibility to construct in a simple way states with definite symmetry by means of spin wavefunctions. With two spin-1/2 fermions (for example, two electrons) 1 and 2 one finds

$$\left|+\tfrac{1}{2}\right\rangle_1\left|+\tfrac{1}{2}\right\rangle_2\,,\quad \frac{1}{\sqrt{2}}\left(\left|+\tfrac{1}{2}\right\rangle_1\left|-\tfrac{1}{2}\right\rangle_2+\left|-\tfrac{1}{2}\right\rangle_1\left|+\tfrac{1}{2}\right\rangle_2\right)\,,\quad \left|-\tfrac{1}{2}\right\rangle_1\left|-\tfrac{1}{2}\right\rangle_2\,,$$

$$\frac{1}{\sqrt{2}}\left(\left|+\tfrac{1}{2}\right\rangle_1\left|-\tfrac{1}{2}\right\rangle_2-\left|-\tfrac{1}{2}\right\rangle_1\left|+\tfrac{1}{2}\right\rangle_2\right)$$

where the first three states (triplet states) are symmetric and the fourth (singlet state) is antisymmetric when the exchange $1 \leftrightarrow 2$ is effected. These two set of vectors are the two basis found by implemeting rules for the addition of two spins $\mathbf{J}=\mathbf{S}_1+\mathbf{S}_2$. By denoting the previous four states by $\xi_{m_j,\,j}(1,2)$ one finds for $j=1$

$$\begin{cases} \chi_{+1,1}(1,2) = \left|+\frac{1}{2}\right\rangle_1\left|+\frac{1}{2}\right\rangle_2 \leftrightarrow \;\; m_1=m_2=+1/2\,,\; m_j=1 \\ \\ \chi_{0,1}(1,2) = \frac{1}{\sqrt{2}}\left(\left|+\frac{1}{2}\right\rangle_1\left|-\frac{1}{2}\right\rangle_2+\left|-\frac{1}{2}\right\rangle_1\left|+\frac{1}{2}\right\rangle_2\right) \;\leftrightarrow\;\; m_1=-m_2=\pm 1/2\,,\; m_j=0 \\ \\ \chi_{-1,1}(1,2) = \left|-\frac{1}{2}\right\rangle_1\left|-\frac{1}{2}\right\rangle_2 \;\leftrightarrow\;\; m_1=m_2=-1/2\,,\; j=1\,,\; m_j=-1 \end{cases}$$

This symmetric basis corresponds to states (97) labelled by index $|m_j|=|m_1+m_2|\leq j=1$. For $j=0$ one has

$$\chi_{0,0}(1,2)=\frac{1}{\sqrt{2}}\left(\left|+\tfrac{1}{2}\right\rangle_1\left|-\tfrac{1}{2}\right\rangle_2-\left|-\tfrac{1}{2}\right\rangle_1\left|+\tfrac{1}{2}\right\rangle_2\right)\;\leftrightarrow\;\; m_1=-m_2=\pm 1/2\,,\; m_j=0$$

Such states fulfil the properties

$$\chi_{m_j,1}(1,2) = +\,\chi_{m_j,1}(2,1) \quad \text{and} \quad \chi_{0,0}(1,2) = -\,\chi_{0,0}(2,1)\,.$$

Therefore, one obtains three symmetric states with $j = 1$ (triplet states) and an antisymmetric state with $j = 0$ (singlet state).

15.3 Helium atom

The (simplified version of the) He-atom problem offers an interesting example where two-electron symmetrized states include the contributions both of spatial wavefunctions and of the spin wavefunctions χ_{j,m_e} just defined. Solutions of $H\Phi(\mathbf{r_1},\mathbf{r_2}) = E\Phi(\mathbf{r_1},\mathbf{r_2})$ where

$$H = H_0 + U = \sum_{j=1}^{2}\left[\frac{\mathbf{p}_j^2}{2m_e} + V(\mathbf{r}_j)\right] + U(\mathbf{r}_1,\mathbf{r}_2)\,, \quad V(\mathbf{r}_j) = -\frac{Ze^2}{|\mathbf{r}_j|}\,,\ U = \frac{e^2}{|\mathbf{r}_1-\mathbf{r}_2|}\,,$$

have, in general, the form $\Phi(\mathbf{r_1},\mathbf{r_2}) = \chi(1,2)\,\psi(\mathbf{r_1},\mathbf{r_2})$.

Even if H does not contain spin-dependent terms one has to include the spin wavefunction $\chi(1,2)$ in ψ in order to take into account the **statistical properties** of electrons that, having half-integer spin, are fermions. Note that solving $H\,\Phi = E\,\Phi$ amounts to finding the solutions of $H\,\psi = E\,\psi$ since, due to the spin independence of the Hamiltonian, $H\,\chi(1,2)\,\psi = \chi(1,2)\,H\,\psi$. If $U(\mathbf{r}_1,\mathbf{r}_2)$ is viewed as a perturbative term, the **unperturbed problem** exhibits two classes of solutions, the first symmetric and the other antisymmetric

$$\psi_s = \frac{1}{\sqrt{2}}\left[u(\mathbf{r}_1)\,v(\mathbf{r}_2) + u(\mathbf{r}_2)\,v(\mathbf{r}_1)\right]\,, \quad \psi_a = \frac{1}{\sqrt{2}}\left[u(\mathbf{r}_1)\,v(\mathbf{r}_2) - u(\mathbf{r}_2)\,v(\mathbf{r}_1)\right]\,.$$

Single-electron wavefunctions u and v of the unperturbed problem are the hydrogen-atom wavefunctions (with $Z = 2$)

$$u(\mathbf{r}_i) = \psi_{n,\ell,m}(\mathbf{r}_i)\,, \quad v(\mathbf{r}_i) = \psi_{n',\ell',m'}(\mathbf{r}_i)$$

which satisfy $[\mathbf{p}_i^2/2m_e + V(\mathbf{r}_i)]\,\psi(\mathbf{r}_i) = E_n\,\psi(\mathbf{r}_i)$. The inclusion of the spin wavefunctions gives the symmetrized states

$$\Phi_1(\mathbf{r_1},\mathbf{r_2}) = \chi_{1,\mu}(1,2)\,\psi_a(\mathbf{r}_1,\mathbf{r}_2)\,,\ \mu = 0,\pm1\,, \qquad \Phi_2(\mathbf{r_1},\mathbf{r_2}) = \chi_{0,0}(1,2)\,\psi_s(\mathbf{r}_1,\mathbf{r}_2)\,,$$

$$\Phi_3(\mathbf{r_1},\mathbf{r_2}) = \chi_{1,\mu}(1,2)\,\psi_s(\mathbf{r}_1,\mathbf{r}_2)\,,\ \mu = 0,\pm1\,, \qquad \Phi_4(\mathbf{r_1},\mathbf{r_2}) = \chi_{0,0}(1,2)\,\psi_a(\mathbf{r}_1,\mathbf{r}_2)\,.$$

where, owing to $\chi_{\mu,1}(1,2) = +\,\chi_{\mu,1}(2,1)$ and $\chi_{0,0}(1,2) = -\,\chi_{0,0}(2,1)$, with $\mu = -1,\,0,+1$,

$$\Phi_1(\mathbf{r_2},\mathbf{r_1}) = \chi_{1,\mu}(2,1)\,\psi_a(\mathbf{r}_2,\mathbf{r}_1) = \chi_{1,\mu}(1,2)\,(-1)\psi_a(\mathbf{r}_1,\mathbf{r}_2) \;=\; -\Phi_1(\mathbf{r_1},\mathbf{r_2})$$

$$\Phi_2(\mathbf{r_2},\mathbf{r_1}) = \chi_{0,0}(2,1)\,\psi_s(\mathbf{r}_2,\mathbf{r}_1) = (-1)\chi_{0,0}(1,2)\,\psi_s(\mathbf{r}_1,\mathbf{r}_2) \;=\; -\Phi_2(\mathbf{r_1},\mathbf{r_2})$$

while

$$\Phi_3(\mathbf{r_2},\mathbf{r_1}) = +\Phi_3(\mathbf{r_1},\mathbf{r_2})\,, \quad \Phi_4(\mathbf{r_2},\mathbf{r_1}) = +\Phi_4(\mathbf{r_1},\mathbf{r_2})\,.$$

Then Φ_1 and Φ_2 feature the requested **antisymmetric form** whereas states Φ_3 and Φ_4, displaying a symmetric character, must be excluded because electrons are particles with half-integer spin (fermions) and thus their wavefunctions must be antisymmetric (Spin-Statistic theorem).

Note that, in Φ_1, antisymmetric wavefunction $\psi_a(\mathbf{r}_1,\mathbf{r}_2)$ is associated with symmetric spin state $\chi_{\mu,1}(1,2)$ having **total spin** $j = 1$, whereas, in Φ_2, symmetric wavefunctions $\psi_s(\mathbf{r}_1,\mathbf{r}_2)$ is associated with antisymmetric spin state $\chi_{0,0}(1,2)$ having **total spin** $j = 0$.

Such symmetrized states allows one to get a qualitatively good approximation of the He-atom spectrum when electron-electron term $U(\mathbf{r}_1, \mathbf{r}_2)$ is included by means of the standard perturbation theory.

Spectrum and energy eigenstates of the unperturbed case. In the absence of U, if

$$H_0(\mathbf{r}_i)\, u(\mathbf{r}_i) = E_n\, u(\mathbf{r}_i)\,, \quad H_0(\mathbf{r}_i)\, v(\mathbf{r}_i) = E_{n'}\, v(\mathbf{r}_i)$$

one finds

$$H_0\, \Phi_1 = H_0 \chi_{1,\mu}\, \psi_a = \chi_{1,\mu}\, \big(H_0(\mathbf{r}_1) + H_0(\mathbf{r}_2)\big) \frac{1}{\sqrt{2}} \Big(u(\mathbf{r}_1)\, v(\mathbf{r}_2) - u(\mathbf{r}_2)\, v(\mathbf{r}_1)\Big)$$

$$= \chi_{1,\mu}\, \big(E_u + E_v\big)\, \frac{u(\mathbf{r}_1)\, v(\mathbf{r}_2) - u(\mathbf{r}_2)\, v(\mathbf{r}_1)}{\sqrt{2}} = \big(E_n + E_{n'}\big)\, \Phi_1\,.$$

Likewise

$$H_0\, \Phi_2 = H_0\, \chi_{0,0}\, \psi_s = \chi_{0,0} \big(H_0(\mathbf{r}_1) + H_0(\mathbf{r}_2)\big) \frac{u(\mathbf{r}_1)\, v(\mathbf{r}_2) + u(\mathbf{r}_2)\, v(\mathbf{r}_1)}{\sqrt{2}} = \ ... = \ \big(E_n + E_{n'}\big)\, \Phi_2\,.$$

Then Φ_1 and Φ_2 are degenerate states because feature the same energy eigenvalue. According to the standard perturbation theory, whenever the unperturbed problem exhibit **degenerate states**, one must verify that the perturbation term U is such that its matrix elements $(\Phi_r, U\Phi_s) \propto \delta_{rs}$. The validity of the latter allows one to use the formulas derived within the non degenerate case in the case with degeneracy. If the previous condition is not fulfilled states Φ_r must be replaced with a new group of degenerate state $\xi_k = \Sigma_r C_{kr} \Phi_r$ such that $(\xi_r, U\xi_s) \propto \delta_{rs}$. Fortunately, in the present case,

$$(\Phi_1, U\, \Phi_2) = \big(\chi_{1,\mu}\, \psi_a, U\, \chi_{0,0}\, \psi_s\big) = \big(\chi_{1,\mu}, \chi_{0,0}\big)\, \big(\chi_{1,\mu}\, \psi_a, U\, \chi_{0,0}\, \psi_s\big) = 0\,.$$

Note that $(\Phi_1, U\, \Phi_2) = (\chi_{1,\mu}, \chi_{0,0})\, (\psi_a,\, \psi_s)$: the scalar product factorizes into a spinor part multiplied by the standard wavefuction part.

Energy spectrum with electron-electron interaction U. The energy levels including the first-order corrections to the energy spectrum are represented by a double series

$$E_{oh} = E_{n,n'} + E_1 = \big(E_n + E_{n'}\big) + (\Phi_1, U\, \Phi_1)$$

$$E_{ph} = E_{n,n'} + E_1 = \big(E_n + E_{n'}\big) + (\Phi_2, U\, \Phi_2)$$

relevant to the so-called **orthohelium** levels and **parahelium** levels, respectively, associated, in turn, to antisymmetric spatial wavefunctions ψ_a and to symmetric spatial wavefunctions ψ_s. One has

$$(\Phi_1, U\, \Phi_1) = (\chi_{1,\mu},\, \chi_{1,\mu})\, (\psi_a,\, U\psi_a) = (\psi_a,\, U\psi_a) = \int d^3x_1\, d^3x_2 \psi_a^*(\mathbf{r}_1, \mathbf{r}_2)\, U(\mathbf{r}_1, \mathbf{r}_2)\, \psi_a(\mathbf{r}_1, \mathbf{r}_2)$$

$$= \int d^3x_1\, d^3x_2\, \frac{u^*(\mathbf{r}_1)\, v^*(\mathbf{r}_2) - u^*(\mathbf{r}_2)\, v^*(\mathbf{r}_1)}{\sqrt{2}}\, U(\mathbf{r}_1, \mathbf{r}_2)\, \frac{u(\mathbf{r}_1)\, v(\mathbf{r}_2) - u(\mathbf{r}_2)\, v(\mathbf{r}_1)}{\sqrt{2}}$$

$$= \int d^3x_1 d^3x_2 \frac{U(1,2)}{2}\Big(|u(1)|^2|v(2)|^2 - u^*(1)\,v^*(2)u(2)\,v(1) - u^*(2)\,v^*(1)u(1)\,v(2) + |u(2)|^2|v(1)|^2\Big)$$

$$= \int d^3x_1\, d^3x_2\, \frac{U(1,2)}{2}\Big(|u(1)|^2|v(2)|^2 + |u(2)|^2|v(1)|^2\Big)$$

$$-\Big(u^*(1)\,v^*(2)u(2)\,v(1) + u^*(2)\,v^*(1)u(1)\,v(2)\Big)$$

$$= \int d^3x_1\, d^3x_2\, U(1,2)\,|u(1)|^2|v(2)|^2 - \int d^3x_1\, d^3x_2\, U(1,2)\,\Big(u^*(1)\,v^*(2)u(2)\,v(1)\Big).$$

Summarizing $(\Phi_1, U\,\Phi_1) = I_C - I_Q$ where

$$I_C = \int d^3x_1\, d^3x_2\, U(\mathbf{r}_1,\mathbf{r}_2)\,|u(\mathbf{r}_1)|^2|v(\mathbf{r}_2)|^2 = \int d^3x_1\, d^3x_2\, \frac{e^2\,|u(\mathbf{r}_1)|^2|v(\mathbf{r}_2)|^2}{|\mathbf{r}_1 - \mathbf{r}_2|}$$

$$= \int d^3x_1\,(-e\,|u(\mathbf{r}_1)|^2|)\,V(\mathbf{r}_1)\,,$$

$$I_Q = \int d^3x_1\, d^3x_2\, U(\mathbf{r}_1,\mathbf{r}_2)\,u(\mathbf{r}_1)\,v(\mathbf{r}_2)\,u^*(\mathbf{r}_2)\,v^*(\mathbf{r}_1)\,.$$

Likewise,

$$(\Phi_2, U\,\Phi_2) = (\chi_{0,0},\,\chi_{0,0})\,(\psi_s,\,U\psi_s) = (\psi_s,\,U\psi_s) = \int d^3x_1\, d^3x_2 \psi_a^*(\mathbf{r}_1,\mathbf{r}_2)\,U(\mathbf{r}_1,\mathbf{r}_2)\,\psi_a(\mathbf{r}_1,\mathbf{r}_2)$$

$$= \int d^3x_1\, d^3x_2\, \frac{u^*(\mathbf{r}_1)\,v^*(\mathbf{r}_2) + u^*(\mathbf{r}_2)\,v^*(\mathbf{r}_1)}{\sqrt{2}}\,U(\mathbf{r}_1,\mathbf{r}_2)\,\frac{u(\mathbf{r}_1)\,v(\mathbf{r}_2) + u(\mathbf{r}_2)\,v(\mathbf{r}_1)}{\sqrt{2}}$$

Compared with $(\Phi_1, U\,\Phi_1)$, one easily note that the only difference is the substitution $- \to +$ in the wavefunction ψ_a which becomes ψ_s. Then

$$(\Phi_2, U\,\Phi_2) = I_C + I_Q\,.$$

Integral I_C can be interpreted, at a semiclassical level, as the electrostatic energy of the charge density $-e|u(\mathbf{r})|^2$ in the potential field $V(\mathbf{r}_1)$ generated by the charge density $-e|v(\mathbf{r})|^2$. The unconventional (purely **quantum**) term is so-called **exchange term** I_Q that has no classical interpretation. This is consistent with the fact that it issues from the electron exchange $u(\mathbf{r}_1)\,v(\mathbf{r}_2) \to u(\mathbf{r}_2)\,v(\mathbf{r}_1)$ generating the second term in ψ_a.

Orthohelium and parahelium energy levels are thus given by

$$\Phi_1 \leftrightarrow E_{oh} = \big(E_n + E_{n'}\big) + I_C - I_Q\,,$$

$$\Phi_2 \leftrightarrow E_{ph} = \big(E_n + E_{n'}\big) + I_C + I_Q\,,$$

respectively, showing that, for a given pair of states $u\,v$, one always has $E_{oh} < E_{ph}$. Whenever the two electrons are in the same spatial state (meaning that $u = v$) state $\Phi_1 = 0$ being $\psi_a = 0$, consistent with the **Pauli exclusion principle**. The orthohelium series is thus characterized by eigenstates where electrons never share the same spatial state.

Hence the helium **ground state** identifies with the lowest state of the parahelium series, where the two electrons are in the same spatial state $n = n' = 1$, $\ell = \ell' = 0$ and $m' = m = 0$ ($\leftrightarrow u = v$) but electrons exhibit opposite spin in the antisymmetric spin state $\chi_{0,0}$. The relevant eigenvalue $2E_1 + I_C + I_Q$ (with $n = n' = 1$) appears to be smaller than the lowest eigenvalue of the orthohelium series $E_1 + E_2 + I_C - I_Q$ (with $n = 1$, $n' = 2$) even if the exchange-term contribution $-I_Q$ is negative. In general, the only eigenstates where the two helium electrons can occupy the same spatial state ($\leftrightarrow u = v$) are those belonging to the parahelium series.

The counterintuitive aspect of this analysis is that, even if the Hamiltonian does not explicitly include spin-dependent terms, one has to take into account the spinor character of Helium-atom energy eigenstates in order to obtain a realistic description of the energy spectrum. Electron-electron interaction does not depends only on U but also, in a less evident but certainly important way, through the coupling of the relevant spins owing to the Pauli principle. Opposite spin are requested in order to accomodate two electrons in the same spatial state. A nice interpretation of this fact is that electrons with parallel spins in the same spatial states undergo a (sort of) infinitely large repulsion. The following effective model well represents this situations

$$H \equiv \sum_{j=1}^{2} \left[\frac{\mathbf{p}_j^2}{2m_e} + V(\mathbf{r}_j) \right] + I_C - \frac{I_Q}{2} \left[1 + \frac{4}{\hbar^2} \mathbf{S}_1 \cdot \mathbf{S}_2 \right] . \tag{114}$$

It provides an equivalent model which reproduces the same energy levels obtained by including the perturbation U in $H = H_0$. Model (114) evidences the role of spin-spin interactions of the two electrons.

16 Quantization of the electromagnetic field

The dynamics of the electromagnetic (EM) fields is described by the Maxwell equations (we use the simplified vector notation $\vec{A} = \mathbf{A}$)

$$\text{div}\ \mathbf{B} = 0, \qquad \text{div}\ \mathbf{D} = \rho$$

$$\text{curl}\ \mathbf{H} - \partial_t \mathbf{D} = \mathbf{j}, \qquad \text{curl}\ \mathbf{E} + \partial_t \mathbf{B} = 0,$$

where $\mathbf{E}$ is the electric field, $\mathbf{D}$ is the electric displacement field, $\mathbf{B}$ the magnetic field and $\mathbf{H}$ the magnetic-field intensity. In the presence of charged matter the density ρ describes the charge distribution while $\mathbf{j}$ represents possible electric currents. In the absence of a medium which is electrically or magnetically active the relations $\mathbf{B} = \mu_0\mu_r\mathbf{H}$ and $\mathbf{D} = \epsilon_0\epsilon_r\mathbf{E}$ reduce to $\mathbf{B} = \mu_0\mathbf{H}$ and $\mathbf{D} = \epsilon_0\mathbf{E}$ being $\mu_r,\ \epsilon_r = 1$.

The first and fourth equations allow one to introduce a simplified representation where only two fields (the vector potential $\mathbf{A}$ and the scalar potential $\nabla\Phi$) are involved. Based on the theorems of differential geometry, the equation $\text{div}\ \mathbf{B} = 0$ implies that there exists a field $\mathbf{A}$ such that $\mathbf{B} = \text{curl}\ \mathbf{A}$. The fourth equation that can be written as

$$0 = \text{curl}\ \mathbf{E} + \partial_t \text{curl}\ \mathbf{A} = \text{curl}\Big(\mathbf{E} + \partial_t \mathbf{A}\Big), \qquad \text{entailing} \qquad \mathbf{E} + \partial_t \mathbf{A} + \nabla\Phi = 0\ .$$

Then the identity $\mathbf{E} = -\partial_t\mathbf{A} - \nabla\Phi$ follows from the fact that an irrotational field is defined up to an arbitrary gradient-like term $\nabla\Phi$.

For $\rho = 0$ and $\mathbf{j} = 0$ (no charge or current sources) the Maxwell equations become $\text{div}\ \mathbf{B} = 0$, $\text{div}\ \mathbf{D} = 0$ and

$$\text{curl}\ \mathbf{B} - \mu_0\epsilon_0\ \partial_t\mathbf{E} = 0, \qquad \text{curl}\ \mathbf{E} + \partial_t\mathbf{B} = 0, \qquad \mu_0\epsilon_0 = \frac{1}{c^2}.$$

which describes the **electromagnetic waves**. By exploiting the symmetries of Maxwell equations we assume that $\Phi = 0$ and $\text{div}\ \mathbf{A} = 0$ (technically speaking this choice is called the Coulomb gauge). Then EM waves are described by

$$\mathbf{E} = -\partial_t\mathbf{A}, \qquad \mathbf{B} = \text{curl}\ \mathbf{A}.$$

It is a simple exercise to show that with this representation of $\mathbf{E}$ and $\mathbf{B}$ the fourth Maxwell equation reduces to an identity

$$\text{curl}\ \Big(-\partial_t\mathbf{A}\Big) + \partial_t \text{curl}\mathbf{A} = 0 \qquad \rightarrow \quad 0\ =\ 0,$$

while the third equation becomes

$$-\Delta\mathbf{A}\ + \frac{1}{c^2}\partial_t^2\mathbf{A} = 0,$$

by means of the vector identity $\text{curl}\big(\text{curl}\mathbf{A}\big) = \nabla\text{div}\mathrm{A} - \Delta\mathrm{A}$. The D'Alambert equation for $\mathbf{A}$ thus describes the EM field.

16.1 Reduction of EM-field Hamiltonian to an ensemble of oscillators.

The most **general solution** of the D'Alambert equation for $\mathbf{A}$ is written as a superposition of plane waves

$$\mathbf{A}(t,\mathbf{x}) = \sum\nolimits_{\mathbf{k}} a_{\mathbf{k}}(t)\; e^{i\mathbf{k}\cdot\mathbf{x}}\; \mathbf{u}_{\mathbf{k}}.$$

Since

$$\mathbf{E} = -\partial_t \mathbf{A} = \sum_{\mathbf{k}} \dot{a}_{\mathbf{k}}(t) e^{i\mathbf{k}\cdot\mathbf{x}} \mathbf{u}_{\mathbf{k}} \equiv \sum_{\mathbf{k}} \mathbf{E}_{\mathbf{k}}, \quad \mathbf{B} = \mathrm{curl}\mathbf{A} = \sum_{\mathbf{k}} a_{\mathbf{k}}(t) e^{i\mathbf{k}\cdot\mathbf{x}} \Big(i\mathbf{k} \wedge \mathbf{u}_{\mathbf{k}}\Big) \equiv \sum_{\mathbf{k}} \mathbf{B}_{\mathbf{k}},$$

then, for each $\mathbf{k}$, the component $\mathbf{E}_{\mathbf{k}}$ of field $\mathbf{E}$ is orthogonal to the corresponding component $\mathbf{B}_{\mathbf{k}}$ of $\mathbf{B}$. Also, due to the Coulomb-gauge condition $\mathrm{div}\mathbf{A} = 0$ entailing $\mathbf{k}\cdot\mathbf{u}_{\mathbf{k}} = 0$ both $\mathbf{E}_{\mathbf{k}}$ and $\mathbf{B}_{\mathbf{k}}$ lie in the plane orthogonal to wavevector $\mathbf{k}$. Substituting the general solution $\mathbf{A}(t,\mathbf{x})$ in the D'Alambert equation we find

$$\begin{aligned} 0 = -\Delta\mathbf{A} + \frac{1}{c^2}\partial_t^2\mathbf{A} &= -\sum_{\mathbf{k}} a_{\mathbf{k}}(t)(-\mathbf{k}^2)\; e^{i\mathbf{k}\cdot\mathbf{x}}\mathbf{u}_{\mathbf{k}} - \frac{1}{c^2}\sum_{\mathbf{k}} \ddot{a}_{\mathbf{k}}(t) e^{i\mathbf{k}\cdot\mathbf{x}}\mathbf{u}_{\mathbf{k}} \\ &= \sum_{\mathbf{k}} \Big(\mathbf{k}^2\; a_{\mathbf{k}}(t) + \frac{1}{c^2}\ddot{a}_{\mathbf{k}}\Big) e^{i\mathbf{k}\cdot\mathbf{x}}\mathbf{u}_{\mathbf{k}} \end{aligned} \tag{115}$$

where

$$\Delta\mathbf{A} = \sum_{\mathbf{k}} a_{\mathbf{k}}(t)(-\mathbf{k}^2)\; e^{i\mathbf{k}\cdot\mathbf{x}}\mathbf{u}_{\mathbf{k}}, \qquad \partial_t^2\mathbf{A} = \sum_{\mathbf{k}} \ddot{a}_{\mathbf{k}}(t) e^{i\mathbf{k}\cdot\mathbf{x}}\mathbf{u}_{\mathbf{k}}$$

have been used. Since the planewaves form a basis then equation (115) entails that, for each $\mathbf{k}$, the term in the brackets must be zero. Then, the most general solution reads

$$\ddot{a}_{\mathbf{k}} + c^2\mathbf{k}^2\; a_{\mathbf{k}}(t) = 0 \qquad \rightarrow \quad a_{\mathbf{k}}(t) = B_{\mathbf{k}}(0)\; e^{+iw_{\mathbf{k}}t} + C_{\mathbf{k}}(0)\; e^{-iw_{\mathbf{k}}t} \tag{116}$$

with the **well-know dispersion relation** $w_{\mathbf{k}}^2 = c^2\mathbf{k}^2$ of EM waves. The important information emerging from equation (116) is that such an equation describes a **harmonic oscillator** and all the solutions $a_{\mathbf{k}}(t)$ simply describe **elementary oscillations**. This is a crucial information in order to perform the **EM field quantization**.

A better characterization of the general solution is achieved by observing that vectors $\mathbf{u}_{\mathbf{k}}$ lie in the plane orthogonal to $\mathbf{k}$. Such condition, described by $\mathbf{k}\cdot\mathbf{u}_{\mathbf{k}} = 0$, follows from the assumption that $\mathrm{div}\mathbf{A} = 0$. In the light of this **two polarizations** of the oscillating fields are allowed which will be described by two orthogonal unit vectors $\mathbf{u}_{\mathbf{k}\sigma}$ with $\sigma = 1, 2$. Then

$$\mathbf{A} = \sum_{\sigma}\sum_{\mathbf{k}} a_{\mathbf{k}\sigma}(t)\; e^{i\mathbf{k}\cdot\mathbf{x}}\mathbf{u}_{\mathbf{k}\sigma} = \sum_{\sigma}\sum_{\mathbf{k}} \Big(B_{\mathbf{k}\sigma}\; e^{+iw_{\mathbf{k}}t} + C_{\mathbf{k}\sigma}\; e^{-iw_{\mathbf{k}}t}\Big)\; e^{i\mathbf{k}\cdot\mathbf{x}}\mathbf{u}_{\mathbf{k}\sigma}$$

To obtain real valued vector fields $\mathbf{E}$ and $\mathbf{B}$ we must consider a less general solution containing the information that $\mathbf{A} = \mathbf{A}^*$. One finds

$$\sum_{\sigma}\sum_{\mathbf{k}} a_{\sigma\mathbf{k}}(t)\; e^{i\mathbf{k}\cdot\mathbf{x}}\mathbf{u}_{\sigma\mathbf{k}} = \sum_{\sigma}\sum_{\mathbf{k}} a^*_{\sigma\mathbf{k}}(t)\; e^{-i\mathbf{k}\cdot\mathbf{x}}\mathbf{u}_{\sigma\mathbf{k}}$$

which, after the substitution $\mathbf{k} \to -\mathbf{k}$ in the left-hand side, entails

$$a_{+\mathbf{k}\sigma} = a^*_{-\mathbf{k}\sigma} \quad \to \quad C_{+\mathbf{k}\sigma} = B^*_{-\mathbf{k}\sigma} \, .$$

Note that vector $\mathbf{u}_{-\mathbf{k}\sigma}$, occuring with the substitution $\mathbf{k} \to -\mathbf{k}$, quite reasonably can be assumed to satisfy the condition $\mathbf{u}_{\mathbf{k}\sigma} = \mathbf{u}_{-\mathbf{k}\sigma}$. The vector potential then takes the final form

$$\mathbf{A} = \sum_\sigma \sum_{\mathbf{k}} \Big(B_{\mathbf{k}\sigma}(t) + B^*_{-\mathbf{k}\sigma}(t) \Big) \, e^{i\mathbf{k}\cdot\mathbf{x}} \mathbf{u}_{\mathbf{k}\sigma} = \sum_\sigma \sum_{\mathbf{k}} \Big(B_{\mathbf{k}\sigma} \, e^{+iw_{\mathbf{k}}t} + B^*_{-\mathbf{k}\sigma} \, e^{-iw_{\mathbf{k}}t} \Big) \, e^{i\mathbf{k}\cdot\mathbf{x}} \mathbf{u}_{\mathbf{k}\sigma}$$

The EM-field Hamiltonian is given by

$$H = \int_V d^3x \left(\frac{\epsilon_0}{2} \, \mathbf{E}^2 + \frac{1}{2\mu_0} \, \mathbf{B}^2 \right) \qquad \text{where } \mathbf{E} = -\partial_t \mathbf{A}, \;\; \mathbf{B} = \text{curl}\mathbf{A}.$$

By substituting the explicit form of $\mathbf{A}$ in terms of plane waves and exploiting the property that plane waves form a basis (the detailed calculation is discussed in the final comment of this section; a more extended discussion can be found in W. Greiner, *Quantum Mechanics* (Special Chapters)) one reduces H to the form

$$\begin{aligned} H &= \sum_\sigma \sum\nolimits_{\mathbf{k}} \left(\frac{\epsilon_0}{2} w^2_{\mathbf{k}} + \frac{1}{2\mu_0} \mathbf{k}^2 \right) 2L^3 \; B^*_{\mathbf{k}\sigma}(t) B_{\mathbf{k}\sigma}(t) \\ &= \frac{1}{2\mu_0} \sum_\sigma \sum\nolimits_{\mathbf{k}} \left(\frac{w^2_{\mathbf{k}}}{c^2} + \mathbf{k}^2 \right) 2L^3 \; B^*_{\mathbf{k}\sigma}(t) B_{\mathbf{k}\sigma}(t) = \sum_\sigma \sum\nolimits_{\mathbf{k}} 2\epsilon_0 w^2_{\mathbf{k}} L^3 \; |B_{\mathbf{k}\sigma}(t)|^2 \end{aligned}$$

where $\mathbf{k}^2 c^2 = w^2_{\mathbf{k}}$ and $\mu_0 \epsilon_0 = 1/c^2$ have been used. By defining the new complex variables

$$b_{\mathbf{k}\sigma}(t) = \gamma_k \; B_{\mathbf{k}\sigma}(t), \qquad \gamma_k = \left(\frac{2L^3 \epsilon_0 w_k}{\hbar} \right)^{1/2} ,$$

the Hamiltonian becomes

$$H = \sum_\sigma \sum\nolimits_{\mathbf{k}} \hbar w_{\mathbf{k}} \; |b_{\mathbf{k}\sigma}(t)|^2 \, . \tag{117}$$

We know that modes $B_{\mathbf{k}\sigma}(t)$ (and thus $b_{\mathbf{k}\sigma}(t)$ which identify with $B_{\mathbf{k}\sigma}(t)$ up to the constant γ_k) satisfy the harmonic-oscillator equation (116). Then the latter can be rewritten as

$$\ddot{b}_{\mathbf{k}\sigma} + w^2_{\mathbf{k}} \; b_{\mathbf{k}\sigma} = 0, \qquad w^2_{\mathbf{k}} = c^2 \mathbf{k}^2. \tag{118}$$

These equations can be derived by exploiyìting the Hamiltonian formalism **once suitable Poisson brackets have been defined**. The latter are given by

$$\{F, G\} = \sum_\sigma \sum_{\mathbf{k}} \frac{1}{i\hbar} \left(\frac{\partial F}{\partial b_{\mathbf{k}\sigma}} \frac{\partial G}{\partial b^*_{\mathbf{k}\sigma}} - \frac{\partial G}{\partial b_{\mathbf{k}\sigma}} \frac{\partial F}{\partial b^*_{\mathbf{k}\sigma}} \right) ,$$

whose most elementary case is given by $F = b_{\mathbf{k}\sigma}$, $G = b^*_{\mathbf{n}\gamma}$

$$\{b_{\mathbf{k}\sigma}, b^*_{\mathbf{n}\gamma}\} = \frac{1}{i\hbar} \delta_{\sigma\gamma} \delta_{\mathbf{kn}}.$$

Such brackets simply represent the reformulation in terms of $b_{\mathbf{k}\sigma}$, $b^*_{\mathbf{n}\gamma}$ of the usual Poisson brackets $\{q_{\mathbf{k}\sigma}, p_{\mathbf{n}\gamma}\} = \delta_{\sigma\gamma}\delta_{\mathbf{kn}}$ for the canonical coordinates $q_{\mathbf{k}\sigma}$ and momenta $p_{\mathbf{n}\gamma}$ (see the final comment below). What is important is that **they reproduce the expected dynamical equations** for variables $b_{\mathbf{q}}$ (for brevity we set $\mathbf{q} = (\mathbf{k}, \sigma)$). To see this let us apply the standard Hamilton procedure for deriving the Hamilton equations

$$\dot{b}_{\mathbf{q}} = \{b_{\mathbf{q}}, H\} = \frac{1}{i\hbar}\frac{\partial H}{\partial b^*_{\mathbf{q}}} = \frac{1}{i\hbar}\frac{\partial}{\partial b^*_{\mathbf{q}}}\sum_{\sigma}\sum\nolimits_{\mathbf{k}} \hbar w_{\mathbf{k}}\,|b_{\sigma\mathbf{k}}|^2 = -iw_{\mathbf{q}}\,b_{\mathbf{q}}$$

$$\dot{b}^*_{\mathbf{q}} = \{b_{\mathbf{q}}, H\} = -\frac{1}{i\hbar}\frac{\partial H}{\partial b_{\mathbf{q}}} = ... \;= +iw_{\mathbf{q}}\,b_{\mathbf{q}}$$

By performing a second time derivation on $\dot{b}_{\mathbf{q}}$ we find

$$\ddot{b}_{\mathbf{q}} = \frac{d}{dt}(-iw_{\mathbf{q}}\,b_{\mathbf{q}}) = -iw_{\mathbf{q}}\,\dot{b}_{\mathbf{q}} = -iw_{\mathbf{q}}(-iw_{\mathbf{q}}\,b_{\mathbf{q}}) = -w^2_{\mathbf{q}}\,\dot{b}_{\mathbf{q}} \qquad \rightarrow \;\; \ddot{b}_{\mathbf{q}} + w^2_{\mathbf{q}}\,\dot{b}_{\mathbf{q}} = 0$$

perfectly matching equations (118). This result proves that the new Poisson brackets are correct in that they reproduce the expected dynamical equations.

16.2 Field quantization, Fock states

The main reason why the previous Poisson brackets represent an extremely precious tool is that they allows us to apply to H the **canonical quantization procedure**. By implementing the latter (remember that it essentailly consists in the rule $\{F, G\} \rightarrow [F, G] = i\hbar\,\{F, G\}$) we get, in the current case,

$$\{b_{\mathbf{k}\sigma}, b^*_{\mathbf{n}\gamma}\} = \frac{1}{i\hbar}\delta_{\sigma\gamma}\delta_{\mathbf{kn}} \qquad \rightarrow \qquad [b_{\mathbf{k}\sigma}, b^+_{\mathbf{n}\gamma}] = \delta_{\sigma\gamma}\delta_{\mathbf{kn}},$$

together with

$$b^*_{\mathbf{k}\sigma}b_{\mathbf{k}\sigma} \qquad \rightarrow \qquad \frac{1}{2}\Big(b^+_{\mathbf{k}\sigma}b_{\mathbf{k}\sigma} + b_{\mathbf{k}\sigma}b^+_{\mathbf{k}\sigma}\Big) = \Big(b^+_{\mathbf{k}\sigma}b_{\mathbf{k}\sigma} + 1/2\Big).$$

The symmetrized quantum formula is justified by the fact both $b^+_{\mathbf{k}\sigma}b_{\mathbf{k}\sigma}$ and $b_{\mathbf{k}\sigma}b^+_{\mathbf{k}\sigma}$ can be, in principle, associated to the classical quantity $|b_{\mathbf{k}\sigma}|^2$. Then the Hamiltonian assumes the quantum form

$$H = \sum_{\sigma}\sum\nolimits_{\mathbf{k}} \hbar w_{\mathbf{k}}\,\Big(b^+_{\mathbf{k}\sigma}b_{\mathbf{k}\sigma} + 1/2\Big) \tag{119}$$

where $1/2$ represents the zero energy contribution of ground state. The entire procedure discussed in this subsection aims to justify the quantum form (119) of the EM Hamiltonian.

One immediately recognizes that the quantity $b^+_{s,k}b_{s,k}$ is the number operator $\hat{n}_{s,k} = b^+_{s,k}, b_{s,k}$ encountered in the study of the quantum harmonic oscillator. Each mode (s, k) is thus associated to a specific number operator $\hat{n}_{s,k}$. The eigenstates of the harmonic oscillator $|n_{s,k}\rangle$ can be used to construct eigenstates of H_{EM}. These have the factorized form

$$|\vec{n}\rangle = \prod_{s}\prod_{k}|n_{s,k}\rangle = |..., n_{1,p}, n_{2,p}, ...n_{1,q}, n_{2,q}, ...\rangle$$

in which $k \equiv \vec{k}$ assumes all the vector values characterizing the wave-vector domain, and each $n_{s,k}$ counts the number of field excitations associated to a given mode $b_{s,k}$. Then

$$\hat{n}_{s,k}|\vec{n}\rangle = n_{s,k}|\vec{n}\rangle$$

The $n_{s,k}$ EM-field excitations represent the photons constituting the fraction of the total photon population associated with the frequency w_k of mode $b_{s,k}$. One finds

$$H_{EM}|\vec{n}\rangle = \Big(E_0 + \sum_{s=1,2}\sum_{k} \hbar w_k b^+_{s,k} b_{s,k}\Big)|\vec{n}\rangle = E(\{n_{s,k}\})|\vec{n}\rangle$$

with the energy eiegenvalue

$$E(\{n_{s,k}\}) = E_0 + \sum_{s=1,2}\sum_{k} \hbar w_k n_{s,k}\,.$$

The latter clearly shows how the total energy is the sum of the contributions $\hbar w_k n_{s,k}$ weighted by the number of photons $n_{s,k}$ having the mode energy $\hbar w_k$.

Comment 1. The Poisson brackets

$$\{b_{\mathbf{k}\sigma}, b^*_{\mathbf{n}\gamma}\} = \frac{1}{i\hbar}\delta_{\sigma\gamma}\delta_{\mathbf{kn}}.$$

represent the reformulation of the usual Poisson brackets in terms of coordinates $q_{\mathbf{k}}$ and momenta $p_{\mathbf{k}}$. The Hamiltonian of the harmonic oscillator (for the sake of simplicity we drop the polarization index)

$$H_{\mathbf{k}} = \frac{p^2_{\mathbf{k}}}{2m} + \frac{m w^2_{\mathbf{k}}}{2} q^2_{\mathbf{k}}$$

involves coordinates $q_{\mathbf{k}}$ and momenta $p_{\mathbf{k}}$. To simplfy the Hamiltonian one typically introduces complex variables $b_{\mathbf{k}}$ and $b^*_{\mathbf{k}}$ by means of formulas

$$q_k = \sqrt{\frac{\hbar}{2mw_k}}(b_{\mathbf{k}} + b^*_{\mathbf{k}}), \qquad p_k = i\sqrt{\frac{\hbar m w_k}{2}}(b^*_{\mathbf{k}} - b_{\mathbf{k}}),$$

leading to the form $H_k = \hbar w_{\mathbf{k}} |b_{\mathbf{k}}|^2$. By means of the inverse formulas

$$b_{\mathbf{k}} = \sqrt{\frac{mw_k}{2\hbar}}\left(q_k + \frac{ip_k}{mw_k}.\right), \qquad b^*_{\mathbf{k}} = \sqrt{\frac{mw_k}{2\hbar}}\left(q_k - \frac{ip_k}{mw_k},\right),$$

one easily checks that new Poisson brackets $\{b_{\mathbf{n}}, b^*_{\mathbf{k}}\} = \delta_{\mathbf{nk}}/i\hbar$ can be derived from the canonical ones $\{q_{\mathbf{n}}, p_{\mathbf{k}}\} = \delta_{\mathbf{nk}}$. In this sense, they appear to be equivalent. One has

$$\{b_{\mathbf{k}}, b^*_{\mathbf{n}}\} = \frac{mw_{\mathbf{k}}}{2\hbar}\left\{\left(q_{\mathbf{k}} + \frac{ip_{\mathbf{k}}}{mw_{\mathbf{k}}}\right), \left(q_{\mathbf{n}} - \frac{ip_{\mathbf{n}}}{mw_{\mathbf{n}}}\right)\right\} = -\frac{mw_{\mathbf{k}}}{2\hbar}\left\{q_{\mathbf{k}}, \frac{ip_{\mathbf{n}}}{mw_{\mathbf{n}}}\right\} + \frac{mw_{\mathbf{n}}}{2\hbar}\left\{q_{\mathbf{n}}, \frac{ip_{\mathbf{k}}}{mw_{\mathbf{k}}}\right\}$$

$$= -\frac{iw_{\mathbf{k}}}{2\hbar w_{\mathbf{n}}}\ \{q_{\mathbf{k}}, p_{\mathbf{n}}\} + \frac{iw_{\mathbf{n}}}{2\hbar w_{\mathbf{k}}}\ \{q_{\mathbf{n}}, p_{\mathbf{k}}\} = \frac{1}{i\hbar}\delta_{\mathbf{kn}}. \qquad QED$$

Comment 2. Derivation of the momentum-mode picture of the EM Hamiltonian.

$$\mathbf{A} = \sum_{\sigma} \sum_{\mathbf{k}} \Big(B_{\mathbf{k}\sigma}\, e^{+iw_{\mathbf{k}}t} + B^{*}_{-\mathbf{k}\sigma}\, e^{-iw_{\mathbf{k}}t} \Big)\, e^{i\mathbf{k}\cdot\mathbf{x}} \mathbf{u}_{\mathbf{k}\sigma}$$

By using the time derivative

$$\partial_t \mathbf{A} = \sum_{\sigma} \sum_{\mathbf{k}} iw_{\mathbf{k}} \Big(B_{\mathbf{k}\sigma}\, e^{+iw_{\mathbf{k}}t} - B^{*}_{-\mathbf{k}\sigma}\, e^{-iw_{\mathbf{k}}t} \Big)\, e^{i\mathbf{k}\cdot\mathbf{x}} \mathbf{u}_{\mathbf{k}\sigma}$$

we find

$$\int_V d^3x\, \mathbf{E}^2 = \int_V d^3x\, (\partial_t \mathbf{A})^2 = \sum_{\sigma} \sum_{\mathbf{k}} \sum_{\gamma} \sum_{\mathbf{q}} \int_V d^3x\, \mathbf{u}_{\mathbf{k}\sigma} \cdot \mathbf{u}_{\mathbf{q}\gamma} \times$$

$$iw_{\mathbf{k}} \Big(B_{\mathbf{k}\sigma}\, e^{+iw_{\mathbf{k}}t} - B^{*}_{-\mathbf{k}\sigma}\, e^{-iw_{\mathbf{k}}t} \Big)\, iw_{\mathbf{q}} \Big(B_{\mathbf{q}\gamma}\, e^{+iw_{\mathbf{q}}t} - B^{*}_{-\mathbf{q}\gamma}\, e^{-iw_{\mathbf{q}}t} \Big)\, e^{i(\mathbf{q}+\mathbf{k})\cdot\mathbf{x}}$$

$$= -\sum_{\sigma} \sum_{\mathbf{k}} \sum_{\gamma} \sum_{\mathbf{q}} \mathbf{u}_{\mathbf{k}\sigma} \cdot \mathbf{u}_{\mathbf{q}\gamma}\, w_{\mathbf{k}} w_{\mathbf{q}} \times$$

$$\Big(B_{\mathbf{k}\sigma}\, e^{+iw_{\mathbf{k}}t} - B^{*}_{-\mathbf{k}\sigma}\, e^{-iw_{\mathbf{k}}t} \Big) \Big(B_{\mathbf{q}\gamma}\, e^{+iw_{\mathbf{q}}t} - B^{*}_{-\mathbf{q}\gamma}\, e^{-iw_{\mathbf{q}}t} \Big)\, L^3\, \delta_{\mathbf{q}+\mathbf{k}}$$

$$= -\sum_{\sigma} \sum_{\mathbf{k}} \sum_{\gamma} L^3 \mathbf{u}_{\mathbf{k}\sigma} \cdot \mathbf{u}_{\mathbf{k}\gamma}\, w^2_{\mathbf{k}} \Big(B_{\mathbf{k}\sigma}\, e^{+iw_{\mathbf{k}}t} - B^{*}_{-\mathbf{k}\sigma}\, e^{-iw_{\mathbf{k}}t} \Big) \Big(B_{-\mathbf{k}\gamma}\, e^{+iw_{\mathbf{k}}t} - B^{*}_{\mathbf{k}\gamma}\, e^{-iw_{\mathbf{k}}t} \Big)$$

$$= \sum_{\sigma} \sum_{\mathbf{k}} L^3 w^2_{\mathbf{k}} \Big(|B_{\mathbf{k}\sigma}|^2 + |B_{-\mathbf{k}\sigma}|^2 - B_{\mathbf{k}\sigma} B_{-\mathbf{k}\sigma}\, e^{+2iw_{\mathbf{k}}t} - B^{*}_{-\mathbf{k}\sigma} B^{*}_{\mathbf{k}\sigma}\, e^{-2iw_{\mathbf{k}}t} \Big)$$

$$= \sum_{\sigma} \sum_{\mathbf{k}} L^3 w^2_{\mathbf{k}} \Big(2|B_{\mathbf{k}\sigma}|^2 - B_{\mathbf{k}\sigma} B_{-\mathbf{k}\sigma}\, e^{+2iw_{\mathbf{k}}t} - B^{*}_{-\mathbf{k}\sigma} B^{*}_{\mathbf{k}\sigma}\, e^{-2iw_{\mathbf{k}}t} \Big)\, .$$

The identity $\mathbf{u}_{\mathbf{k}\sigma} \cdot \mathbf{u}_{\mathbf{k}\gamma} = \delta_{\gamma\sigma}$ has been used. Likewise, by exploiting the expression

$$\mathrm{curl}\mathbf{A} = \sum_{\sigma} \sum_{\mathbf{k}} \Big(B_{\mathbf{k}\sigma}\, e^{+iw_{\mathbf{k}}t} + B^{*}_{-\mathbf{k}\sigma}\, e^{-iw_{\mathbf{k}}t} \Big)\, e^{i\mathbf{k}\cdot\mathbf{x}} i\mathbf{k} \wedge \mathbf{u}_{\mathbf{k}\sigma}$$

we find

$$\int_V d^3x\, \mathbf{B}^2 = \int_V d^3x\, (\mathrm{curl}\mathbf{A})^2 = \sum_{\sigma} \sum_{\mathbf{k}} \sum_{\gamma} \sum_{\mathbf{q}} \int_V d^3x\, (i\mathbf{k} \wedge \mathbf{u}_{\mathbf{k}\sigma}) \cdot (i\mathbf{q} \wedge \mathbf{u}_{\mathbf{q}\gamma}) \times$$

$$\Big(B_{\mathbf{k}\sigma}\, e^{+iw_{\mathbf{k}}t} + B^{*}_{-\mathbf{k}\sigma}\, e^{-iw_{\mathbf{k}}t} \Big) \Big(B_{\mathbf{q}\gamma}\, e^{+iw_{\mathbf{q}}t} + B^{*}_{-\mathbf{q}\gamma}\, e^{-iw_{\mathbf{q}}t} \Big)\, e^{i(\mathbf{q}+\mathbf{k})\cdot\mathbf{x}}$$

$$= -\sum_{\sigma} \sum_{\mathbf{k}} \sum_{\gamma} \sum_{\mathbf{q}} (\mathbf{k} \wedge \mathbf{u}_{\mathbf{k}\sigma}) \cdot (\mathbf{q} \wedge \mathbf{u}_{\mathbf{q}\gamma}) \times$$

$$\Big(B_{\mathbf{k}\sigma}\, e^{+iw_{\mathbf{k}}t} + B^*_{-\mathbf{k}\sigma}\, e^{-iw_{\mathbf{k}}t}\Big)\, \Big(B_{\mathbf{q}\gamma}\, e^{+iw_{\mathbf{q}}t} + B^*_{-\mathbf{q}\gamma}\, e^{-iw_{\mathbf{q}}t}\Big)\, L^3\delta_{\mathbf{q}+\mathbf{k}}$$

$$= \sum_{\sigma}\sum_{\mathbf{k}}\sum_{\gamma}(\mathbf{k}\wedge\mathbf{u}_{\mathbf{k}\sigma})\cdot(\mathbf{k}\wedge\mathbf{u}_{\mathbf{k}\gamma})L^3\Big(B_{\mathbf{k}\sigma}e^{iw_{\mathbf{k}}t} + B^*_{-\mathbf{k}\sigma}e^{-iw_{\mathbf{k}}t}\Big)\, \Big(B_{-\mathbf{k}\gamma}e^{iw_{\mathbf{k}}t} + B^*_{\mathbf{k}\gamma}e^{-iw_{\mathbf{k}}t}\Big)$$

Since $(\mathbf{k}\wedge\mathbf{u}_{\mathbf{k}\sigma})\cdot(\mathbf{k}\wedge\mathbf{u}_{\mathbf{k}\gamma}), = k^2\ \mathbf{u}_{\mathbf{k}\sigma}\cdot\mathbf{u}_{\mathbf{k}\gamma} = k^2\ \delta_{\gamma\sigma}$ then

$$\int_V d^3x\ \mathbf{B}^2 = \sum_{\sigma}\sum_{\mathbf{k}}\sum_{\gamma}\ L^3k^2\ \Big(B_{\mathbf{k}\sigma}B_{-\mathbf{k}\sigma}\ e^{2iw_{\mathbf{k}}t} + |B_{\mathbf{k}\sigma}|^2 + |B_{-\mathbf{k}\sigma}|^2 + B^*_{-\mathbf{k}\sigma}B^*_{\mathbf{k}\sigma}\ e^{-2iw_{\mathbf{k}}t}\Big)$$

$$= \sum_{\sigma}\sum_{\mathbf{k}}\ L^3k^2\ \Big(2|B_{\sigma\,\mathbf{k}}|^2 + B_{\mathbf{k}\sigma}B_{-\mathbf{k}\sigma}\ e^{2iw_{\mathbf{k}}t} + B^*_{-\mathbf{k}\sigma}B^*_{\mathbf{k}\sigma}\ e^{-2iw_{\mathbf{k}}t}\Big)$$

By substituting these quantities in the EM-field Hamiltonian we find

$$H = \frac{\epsilon_0}{2}\int_V d^3x\ (\ \mathbf{E}^2 + c^2\ \mathbf{B}^2)$$

$$= \frac{\epsilon_0}{2}\Big[\sum_{\sigma}\sum_{\mathbf{k}} L^3w_{\mathbf{k}}^2\ \Big(2|B_{\mathbf{k}\sigma}|^2 - B_{\mathbf{k}\sigma}B_{-\mathbf{k}\sigma}\ e^{2iw_{\mathbf{k}}t} - B^*_{-\mathbf{k}\sigma}B^*_{\mathbf{k}\sigma}\ e^{-2iw_{\mathbf{k}}t}\Big)$$

$$+c^2\sum_{\sigma}\sum_{\mathbf{k}}\ L^3k^2\ \Big(2|B_{\mathbf{k}\sigma}|^2 + B_{\mathbf{k}\sigma}B_{-\mathbf{k}\sigma}\ e^{2iw_{\mathbf{k}}t} + B^*_{-\mathbf{k}\sigma}B^*_{\mathbf{k}\sigma}\ e^{-2iw_{\mathbf{k}}t}\Big)\Big]$$

Thanks to the dispersion relation $w_{\mathbf{k}}^2 = c^2k^2$ we obtain

$$H = \sum_{\sigma}\sum_{\mathbf{k}} 2\epsilon_0L^3w_{\mathbf{k}}^2\ |B_{\mathbf{k}\sigma}|^2 \qquad\qquad QED$$

A Classical vs quantum configurations with N particles in $r < N$ states

Let us compare classical and quantum configurations formed by N particles distributed in $r < N$ physical states. The simplest example is represented by the completely symmetric wavefunction of N bosons obtained by using α_1, α_2, ... α_N different quantum numbers

$$\psi_S(\mathbf{x}_1, ...\mathbf{x}_j, ...\mathbf{x}_N) = \frac{1}{\sqrt{N!}} \sum_P P\ [\psi_{\alpha_1}(\mathbf{x}_1)\psi_{\alpha_2}(\mathbf{x}_2) ... \psi_{\alpha_N}(\mathbf{x}_N)]$$

To simplify the notation one can assume $\psi_{\alpha_k}(\mathbf{x}_i) \equiv \psi_k(i)$. Then, for example, with $N = 3$

$$\psi_S(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) = \frac{1}{\sqrt{3!}}\Big[\psi_1(1)\psi_2(2)\psi_3(3) + \psi_1(2)\psi_2(3)\psi_3(1) + \psi_1(3)\psi_2(1)\psi_3(2)$$

$$+\psi_1(2)\psi_2(1)\psi_3(3) + \psi_1(1)\psi_2(3)\psi_3(2) + \psi_1(3)\psi_2(2)\psi_3(1)\Big]$$

If the symmetrized wavefunction for $N = N_1 + N_2 + ... N_r$ bosons is obtained by using α_1, α_2, ... α_r different quantum numbers then

$$\Phi_s(\mathbf{x}_1, ...\mathbf{x}_j, ...\mathbf{x}_N) = \frac{\sqrt{N_1!N_2!..N_r!}}{\sqrt{N!}} \times$$

$$\sum_P{}^* P \left[\prod_{p_1=1}^{N_1} \psi_{\alpha_1}(\mathbf{x}_{p_1}) \prod_{p_2=N_1+1}^{N_1+N_2} \psi_{\alpha_2}(\mathbf{x}_{p_2}) ... \prod_{p_r=N_1+...N_{r-1}+1}^{N_1+N_2+...N_r} \psi_{\alpha_r}(\mathbf{x}_{p_r})\right]$$

Consider the case when the symmetric state with $N = 4$ bosons is such that single-particle states $\psi_{\alpha_1}(\mathbf{x})$, $\psi_{\alpha_2}(\mathbf{x})$ and $\psi_{\alpha_3}(\mathbf{x})$ are occupied by 2, 1 and 1 bosons, respectively.

$$\Phi_s(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4) = \frac{\sqrt{N_1!N_2!N_3!}}{\sqrt{N!}} \sum_P{}^* P\ [\psi_{\alpha_1}(\mathbf{x}_1)\psi_{\alpha_1}(\mathbf{x}_2) \times \psi_{\alpha_2}(\mathbf{x}_3) \times \psi_{\alpha_3}(\mathbf{x}_4)]$$

$$= \frac{\sqrt{2!\,1!\,1!}}{\sqrt{4!}} \Big[\ \psi_1(1)\psi_1(2)\,\psi_2(3)\psi_3(4) + \psi_1(1)\psi_1(3)\,\psi_2(2)\psi_3(4) + \psi_1(1)\psi_1(4)\,\psi_2(2)\psi_3(3)$$

$$+\psi_1(2)\psi_1(3)\,\psi_2(1)\psi_3(4) + \psi_1(2)\psi_1(4)\,\psi_2(1)\psi_3(3) + \psi_1(3)\psi_1(4)\,\psi_2(1)\psi_3(2)$$

$$+\psi_1(1)\psi_1(2)\,\psi_2(4)\psi_3(3) + \psi_1(1)\psi_1(3)\,\psi_2(4)\psi_3(2) + \psi_1(1)\psi_1(4)\,\psi_2(3)\psi_3(2)$$

$$+\psi_1(2)\psi_1(3)\,\psi_2(4)\psi_3(1) + \psi_1(2)\psi_1(4)\,\psi_2(3)\psi_3(1) + \psi_1(3)\psi_1(4)\,\psi_2(2)\psi_3(1)\Big]$$

The 12 states forming this symmetric wave function are, quantum-mechanically, perfectly equivalent the one to the other. Owing to the **boson indistiguishability**, each states of this superposition, for example $\psi_1(2)\psi_1(3)\psi_2(1)\psi_3(4)$, embodies the same physical information (suppose that ψ_i corresponds to energy E_i)

$$2\,\text{boson with energy}\,E_1\,,\quad 1\,\text{boson with energy}\,E_2\,,\quad 1\,\text{boson with energy}\,E_3\,,$$

of the remaining 11 states

$$\psi_1(1)\psi_1(2)\psi_2(3)\psi_3(4)\,,\quad \psi_1(1)\psi_1(2)\psi_2(4)\psi_3(3)\,,\,....\,.$$

This is the origin of the substantial difference between the symmetric state $\Phi_s(\mathbf{x}_1,\mathbf{x}_2,\mathbf{x}_3,\mathbf{x}_4)$ –a superposition of states in which the 12 states have the same relevance– and the corresponding **classical configurations** in which the 12 states involved, collectively exhibiting the same physical properties, appear to be different in that the 4 particles involved are **distinguishable particles**. In particular, within a fully classical description the different trajectories of each particle allow one to **distinguish**, **at any time**, two any configurations where particles have been exchanged. For example

$$\text{bosons}\,1+2\,\text{with energy}\,E_1\,,\quad \text{boson}\,3\,\text{with energy}\,E_2\,,\quad \text{boson}\,4\,\text{with energy}\,E_3\,,$$

$$\text{bosons}\,1+3\,\text{with energy}\,E_1\,,\quad \text{boson}\,4\,\text{with energy}\,E_2\,,\quad \text{boson}\,2\,\text{with energy}\,E_3\,,$$

which are completely indistinguishable and thus equivalent at the quantum-mechanical level. Symmetric state such as Φ_s thus represent the dramatic **information collapse** occurring in any ensemble of identical bosons.

In case the particles were spinless **fermions** no quantum state could be formed under the present conditions (4 fermions distributed among 3 quantum states) because **two fermions cannot have the same quantum number** and then no antisymmetric wavefunction can be constructed.

However, if spins are included among the properties of the 4 particles forming the system, instead of the 3 single-particle states ψ_1, ψ_2, and ψ_3 one has 3×2 different quantum states ψ_{k,σ_k} with $\sigma_i = \pm \equiv \pm 1/2$ (spin up/down). As in the case of bosons, we assume that index k in ψ_{k,σ_k} describes the energy eigenvalue E_k. Such states can be used to construct several 4-fermion states in which fermions occupy different single-particle quantum states

$$\Psi_{1+}(1)\Psi_{1-}(2)\Psi_{3\,\sigma_3}(3)\Psi_{4\,\sigma_4}(4)\,,\quad \sigma_3,\sigma_4 = \pm = \pm 1/2\,.$$

For given values of spin σ_3 and σ_4 one obtains 24 different states by effecting the permutations of particle indexes 1, 2, 3, and 4. Note that states $\Psi_{1\pm}$ necessarily exhibit opposite spin

corresponding to the same energy E_1 whereas states $\Psi_{3\,\sigma_3}$, $\Psi_{4\,\sigma_4}$ are such that both $\sigma_3 = \sigma_4$ and $\sigma_3 \neq \sigma_4$ are allowed due to the diversity of energies E_3, E_4.

A natural way to start the **antisymmetrization process** leading to get fully antisymmetrized states is to consider states of the form

$$\Big(\Psi_{1+}(1)\Psi_{1-}(2) - \Psi_{1+}(2)\Psi_{1-}(1)\Big)\,\Psi_{3\,\sigma_3}(3)\Psi_{4\,\sigma_4}(4)\,,$$

$$\Big(\Psi_{1+}(2)\Psi_{1-}(3) - \Psi_{1+}(3)\Psi_{1-}(2)\Big)\,\Psi_{3\,\sigma_3}(4)\Psi_{4\,\sigma_4}(1)\,, \quad \ldots\ .$$

One obtains 24 states which, however, are partially antisymmetric. In the first case, e. g., the antisymmetry property is restricted to the exchange $1 \leftrightarrow 2$. In order to extend the same property to the exchange $3 \leftrightarrow 4$ one must consider a "more symmetric" state such as

$$\Big(\Psi_{1+}(1)\Psi_{1-}(2) - \Psi_{1+}(2)\Psi_{1-}(1)\Big)\,\Big(\Psi_{3\,\sigma}(3)\Psi_{4\,\sigma'}(4) - \Psi_{3\,\sigma}(4)\Psi_{4\,\sigma'}(3)\Big)\,.$$

One can construct 6 different states of this form by effecting the permutations of particle indexes 1, 2, 3, and 4. This, however, is not sufficient to ensure that the new states are antisymmetric for exchanges such as, for example, $1 \leftrightarrow 3$, and $2 \leftrightarrow 3$. It becomes clear that only considering all the permutations of particle indexes 1, 2, 3, and 4 one achieves the completely antisymmetric form summarized by the Slater-determinant formula (113)

$$\Phi_a(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4) = \frac{1}{\sqrt{4!}}\, det \begin{bmatrix} \Psi_{1-}(1) & \Psi_{1+}(1) & \Psi_{3\sigma}(1) & \Psi_{4\sigma'}(1) \\ \Psi_{1-}(2) & \Psi_{1+}(2) & \Psi_{3\sigma}(2) & \Psi_{4\sigma'}(2) \\ \Psi_{1-}(3) & \Psi_{1+}(3) & \Psi_{3\sigma}(3) & \Psi_{4\sigma'}(3) \\ \Psi_{1-}(4) & \Psi_{1+}(4) & \Psi_{3\sigma}(4) & \Psi_{4\sigma'}(4) \end{bmatrix} \tag{120}$$

The determinant-like form of this superposition is necessary to get state Φ_a with the expected property. The previous antisymmetrization process once more represents the **information collapse** where the distinctive characters of 24 different states progressively disppear to reach the full antisymmetry. In a more classical sense, 24 configurations that should be different if particles were distinguishable merge in a unique one. This parallel, however, is not permitted in that spin has no classical counterpart. This collapse is similar to that observed in the case with 4 bosons where 12 quantum states are superimposed to form a unique symmetric states.

B Time-dependent Variational Principle

Solving the Schrödinger problem of systems involving some complex potential or interaction term due to external fields is a common situation when dealing with real systems. In addition, similar to atomic systems where many electrons undergo the nucleus action but also interact the one with the other, real systems often exhibit a many-body character which constitutes a further source of complexity. A simple example providing a clear description of the procedure based on the Time-dependent Variational Principle is the following problem proposed by Dirac. Consider a one-dimensional potential problem where $H = (p^2/2m) + U(x)$ and $U(x)$ represents some complex potential. Assume that the solution of the Schrödinger problem has the form

$$\psi(x) = \Lambda\, e^{i\phi} e^{i\frac{\alpha x}{\hbar}} e^{-\frac{(x-\beta)^2}{2\sigma^2}}\,,$$

where the α-dependent term represents a plane wave while the Gaussian term describes the particle localization. All the parameters are functions of time whose form is unknown. Evaluating the expectation values of x, p, $(x-\beta)^2$ allows one to interpret the role played by such parameters. Concerning p one finds

$$\langle\psi|p|\psi\rangle = \int_{\mathbb{R}} dx\, \psi^*(x)(-i\hbar\partial_x)\psi(x) = \int_{\mathbb{R}} dx\, \left[\,\alpha + i\frac{(x-\beta)}{\sigma^2}\,\right] \Lambda^2 e^{-\frac{(x-\beta)^2}{\sigma^2}} = \alpha \int_{\mathbb{R}} dx\, |\psi(x)|^2 = \alpha$$

since

$$\int_{\mathbb{R}} dx\, |\psi(x)|^2 = 1\,, \qquad \Lambda^2 \int_{\mathbb{R}} dx\, (x-\beta)\, e^{-\frac{(x-\beta)^2}{\sigma^2}} = 0\,.$$

Notice that Λ can be fixed so as to ensure the correct normalization of $\psi(x)$

$$1 = \int_{\mathbb{R}} dx\, |\psi(x)|^2 = \Lambda^2 \int_{\mathbb{R}} dx\, e^{-\frac{(x-\beta)^2}{\sigma^2}} = \Lambda^2 \int_{\mathbb{R}} dy\, \sigma e^{-y^2} = \Lambda^2\, \sigma\sqrt{\pi} \Longrightarrow \Lambda = \frac{1}{\sqrt{\sigma\sqrt{\pi}}}\,.$$

Concerning x and $(x-\beta)^2$ one has

$$\langle\psi|x|\psi\rangle = \int_{\mathbb{R}} dx\, x|\psi(x)|^2 = \int_{\mathbb{R}} dx\, [\,(x-\beta)+\beta\,]|\psi(x)|^2 = \beta \int_{\mathbb{R}} dx|\psi(x)|^2 = \beta\,.$$

and

$$\langle\psi|(x-\beta)^2|\psi\rangle = \int_{\mathbb{R}} dx\, (x-\beta)^2|\psi(x)|^2 = \Lambda^2 \int_{\mathbb{R}} dy\, y^2 e^{-\frac{y^2}{\sigma^2}} =$$

$$= -\Lambda^2 \left[\partial_s \int_{\mathbb{R}} dy\, e^{-sy^2}\right]_{s=1/\sigma^2} = -\Lambda^2 \left[\partial_s \sqrt{\pi/s}\right]_{s=1/\sigma^2} = \frac{1}{2}\sigma^2\,,$$

showing, in turn, the physical meaning of β and of σ, respectively.

When solving some Schrödinger problem is particularly difficult, a standard strategy consists in assuming/imposing state $|\psi\rangle$ to satisfy the weaker version of the Schrödinger problem

$\langle\psi|[i\hbar\partial_t - H]|\psi\rangle = 0$. Of course, the state $|\psi\rangle$ obtained within this approach represents, in general, an approximation to the exact solution. By calculating

$$i\hbar\langle\psi|\partial_t|\psi\rangle = \int_{\mathbb{R}} dx \left[-\hbar\dot{\phi} - \dot{\alpha}x + i\hbar\dot{\beta}\frac{(x-\beta)}{\sigma^2}\right]|\psi(x)|^2 = -\left[\hbar\dot{\phi} + \dot{\alpha}\beta\right],$$

$$\langle\psi|H|\psi\rangle = \langle\psi|\left[\frac{p^2}{2m} + U\right]|\psi\rangle = \int_{\mathbb{R}} dx \left[-\frac{\hbar^2}{2m}\psi^*(x)(\partial_x^2)\psi(x) + |\psi(x)|^2 U(x)\right] = \frac{\alpha^2}{2m} + \mathcal{U}(\beta)$$

one take $\langle\psi|[i\hbar\partial_t - H]|\psi\rangle = 0$ into the form

$$-\hbar\dot{\phi} = \dot{\alpha}\beta + \frac{\alpha^2}{2m} + \mathcal{U}(\beta)\,,$$

giving in turn

$$\hbar\phi = \int_0^t d\tau \left[-\dot{\alpha}\beta - \frac{\alpha^2}{2m} - \mathcal{U}(\beta)\right] = C + \int_0^t d\tau \left[\dot{\beta}\alpha - \frac{\alpha^2}{2m} - \mathcal{U}(\beta)\right],$$

with $C = -[\beta\alpha]_0^t$. It is natural to interpret

$$\mathcal{L} = \dot{\beta}\alpha - \frac{\alpha^2}{2m} - \mathcal{U}(\beta)\,, \qquad \mathcal{H} = \frac{\alpha^2}{2m} + \mathcal{U}(\beta)$$

as an effective Lagrangian and an effective Hamiltonian, respectively. This follows from the observation that $\alpha^2/2m$ is a kinetic-energy term (α is the expectation value of p) and $\mathcal{U}(\beta)$ is an effective potential (β is the expectation value of x). One should recall that, α and β are undetermined functions of time. Based on this classical analogy it is natural to regard $\hbar\phi$ as an **effective action**. Then, imposing the condition that $\hbar\phi$ is **stationary** generates the relevant Lagrange equations

$$\frac{d}{dt}\frac{\partial\mathcal{L}}{\partial\dot{\alpha}} - \frac{\partial\mathcal{L}}{\partial\alpha} = 0\,, \quad \frac{d}{dt}\frac{\partial\mathcal{L}}{\partial\dot{\beta}} - \frac{\partial\mathcal{L}}{\partial\beta} = 0\,.$$

Substituting the explicit form $\mathcal{L} = \dot{\beta}\alpha - \frac{\alpha^2}{2m} - \mathcal{U}(\beta)$ of the Lagrangian we obtain the equations

$$\dot{\alpha} = -\frac{\partial\mathcal{H}}{\partial\beta}\,, \qquad \dot{\beta} = +\frac{\partial\mathcal{H}}{\partial\alpha}\,,$$

that exhibit a well recognizable Hamiltonian form. We conclude that the original, purely quantum, Hamiltonian problem can be reformulated as a classical problem described by an effective Hamiltonian $\mathcal{H}$ whose equations are standardly given by $\dot{\alpha} = \{\alpha, \mathcal{H}\}$, and $\dot{\beta} = \{\beta, \mathcal{H}\}$, once the Poisson brackets $\{\beta, \alpha\} =$ have been defined. This procedure can be generalized, in principle, to arbitrarily complex quantum systems.

References

[1] W. Greiner, *Quantum Mechanics (Special Chapters)* (Springer-Verlag, Berlin, Heidelberg 1998)

[2] W. Greiner and J. Reinhardt *Field Quantization* (Springer-Verlag, Berlin, Heidelberg 1996)

[3] J. J. Sakurai, *Modern Quantum Mechanics* (Addison-Wesley Publishing Company, Reading, Massachusetts 1994)

[4] L. Pitaevskii and S. Stringari, *Bose-Einstein Condensation* (Clarendon Press, Oxford 2003)

[5] Ph. W. Courteille, V. S. Bagnato, and V. I. Yukalov, *Bose-Einstein Condensation of Trapped Atomic Gases*, Laser Physics 11 (6) (2001) 659-800

[6] Zhang W M, Feng D H, and Gilmore R 1990 *Rev. Mod. Phys.* **62**, 867

[7] V. I. Yukalov, *Cold Bosons in Optical Lattices*, Laser Physics 19 (1) (2009) 1-110

[8] L. D. Landau and E. M. Lifshitz, *Quantum Mechanics*, Vol. 3 of Course of Theoretical Physics (Pergamon Press, Oxford 1991)

[9] V. B. Berestetskii, E. M. Lifshitz and L. Pitaevskii, *Quantum Electrodynamics*, Vol. 4 of Course of Theoretical Physics (Pergamon Press, Oxford 1980).