Francisco CHINESTA, Emmanuelle ABISSET-CHAVANNE - The Schrödinger Equation - 2018

Chapter 1
The Schrödinger equation

Chinesta, Francisco; Abisset-Chavanne Emmanuelle

1.1 The history of quantum mechanics in small bits

Niels Bohr, one of the father of quantum mechanics, claimed that ‘if quantum mechanics doesn’t profoundly shocked you, you haven’t understood it yet”, and in the same time, he confessed that he never understood it completely. There are numerous books devoted to the history and main concepts of quantum mechanics [53, 42, 30, 24, 45, 11, 41, 32]. In such a history it is very difficult to define the starting point. In the particular case of the atom history, coming back to the Democritus ideas (four centuries before Christ) is, perhaps, a bit excessive!

The first relevant concepts were the Newton ideas about the light that he supposed composed by particles. However at the same period Huygens claimed that in fact light is a wave. This last theory was widely accepted after the experiments of Thomas Young (at the beginning of the nineteen century) and Augustin Fresnel a little bit later. This theory was definitively consolidated from the Maxwell works on electromagnetism.

The idea of the atom was firstly addressed by Robert Boyle (17th century), by Newton in his works on physics and optics, by Lavoisier (18th century) in his studies on combustion, by Dalton (at the beginning of the 19th century) and by Gay-Lussac and Avogadro in the 19th century.

Atoms were considered in physics by Maxwell and Boltzmann, in their theory of gases, the basis of statistical mechanics. But these ideas were not accepted by the scientific community, and some recognized scientist attacked this theory, not confirmed at that time experimentally, where the issue of reversibility was the major protagonist, motivating, apparently, Boltzmann suicide in 1906. The reality of the atoms was established in one of the three capital works published by Einstein in 1905. One of them concerned the relativity theory, the second one the photoelectric effect (from which he received the Nobel Prize in 1921) and the third one concerned the description of the Brownian motion based on the existence of atoms in continuous movement.

At the end of the 19^{th} century there was a certain controversy about the essence of the radiation produced by a metallic filament driving an electrical current in a vacuum environment. J.J. Thompson in the Cavendish laboratory (founded by Maxwell in Cambridge, U.K.) conducted some key experiments leading to the electron discovery, that was the first elementary particle to be identified. For this discovery, J.J. Thompson received the Nobel Prize. Thompson, in the purest British tradition imagined the atom like a pudding or even better, like a watermelon where the negative electrons were uniformly distributed into a positive matrix. Research was very active at that time in some topics related to the nature of the X-rays or the radioactivity, with some of the most salient scientists involved in it: Henry Becquerel, Marie and Pierre Curie or Ernest Rutherford. The last one identified the alpha and beta radiations (the gamma one was discovered later) composed by positive and negative particles moving at high speed. The gamma rays seemed similar to the X-rays but showing lower wavelengths.

At the beginning of the 20^{th} century Rutherford (working at the University of Manchester) conducted some key experiments evidencing the alpha beam scattering when it impacted gold foils. To explain the observed results, incompatible with the Thompson atom model, Rutherford proposed a model for the atom consisting of electrons rotating around a densely positive nucleus, and he received in turn the Nobel Prize. However the Rutherford model had an important weakness because an accelerated charged particle emits radiation. The related energy loss should motivate its collapse, in contrast with the experimental evidence on the atoms stability.

At that time, another important research topic concerned the radiation coming from a black body. Classical models predicted an infinite energy. The main contribution in this field was provided by Planck who, using the statistical theory of thermodynamics developed by Boltzmann, concluded that the energy is quantified and related to the frequency ν from the expression: $E = \mathrm{h}\nu$ where h is known as the Planck constant and its value in the metric system of units is $\mathrm{h} = 6.626 \times 10^{-34} (\mathrm{J} \cdot \mathrm{s})$.

Planck works were considered in the model elaborated by Einstein (that deserved his Nobel Prize) concerning the photoelectric effect, that was explained from the introduction of the photon (light quanta). Thus, some centuries later, the Newton's idea about the discrete essence of the light (composed of particles) was renewed.

A new proposal of the atom model was given by Niels Bohr who suggested that electrons are located in orbitals having a permitted energy according to the energy quantification. To move from one orbit to another it must receive or loss an amount of energy corresponding to an integer number of quanta. This simple model allowed to describe successfully the chemical structure of the matter.

However, if we accept that light has a double character: particles and waves, why not extend this assertion to any other particle? This was the idea of Louis de Broglie who, starting from the Planck and Einstein results, established in his PhD thesis that momentum, p, and frequency, ν, are related from $pc = \mathrm{h}\nu$, where c represents the speed of light. This relation has deep consequences, because in its left-hand side appears the momentum (characteristic of particles) while the right-hand side involves frequency (characteristic of waves), stating a subtle duality between particles and waves.

Obviously its experimental corroboration was not simple because in order to produce wave diffraction one needs to proceed with a diaphragm whose dimension must be of the same order than the wavelength λ. From the de Broglie's model it results that $\lambda p = $ h. Thus, in order to obtain large enough wavelengths, momentum p must be small enough, as it is the case when considering lightweight particles. De Broglie during his PhD defense, answering a question of a sceptic member of his evaluation committee, suggested that his hypothesis could be verified by diffracting electrons in a crystal, scenario that makes compatible the wavelength (atomic distance) and the particles momentum.

The crucial experiment suggested by de Broglie during his PhD dissertation was finally conducted simultaneously by Davisson (in USA) and George Thompson (J.J. Thompson's son) in U.K. Both of them received the Nobel Prize. Thus, ironically, J.J. Thompson received the Nobel Prize for proving that electrons are particles and his son, George Thompson, received the same prize but now for proving that electrons are "also" waves!

It is important to note that interference, observed in these experiments, and easily understood within the wave framework, can be also explained from the Feynman path-integral, where each particle follows 'simultaneously" any possible trajectory while interacting with itself.

Another eminent and key scientist was W. Pauli, who introduced the fourth quantum number (the one related to the spin) as well as the exclusion principle which states that it is impossible to find two electrons in the same quantum state, and allowed a better understanding of the Bohr's atom model. At that time, two statistics were introduced, the one of Bose-Einstein (that applies on particles with integer spin that are not concerned by the exclusion principle) and the one due to Fermi-Dirac (that applies on particles with semi-integer spin and for which the exclusion principle applies). These statistics are of capital importance for describing the structure of matter as well as for explaining some exotic behaviors as the ones related to superconductivity and superfluidity.

The end of our brief overview concerns three attempts to describe the quantum mechanics reality. The first approach was introduced in Gottingen by the team led by Max Born, who, in collaboration with Heisenberg, proposed the matrix mechanics to describe quantum dynamics. The second approach was initiated in Cambridge by Paul Dirac who, starting from Heisenberg's works proposed a new algebra (the quantum algebra) which incorporates the matrix mechanics as a particular case. Finally, Erwin Schrödinger, inspired by de Broglie's works, proposed the introduction of a wavefunction describing the distribution of the particles in the whole space as well as the partial differential equation governing its evolution, known as wave-equation. The last involves a continuous and unbounded medium (only the energy is discrete), where the evolution of the scalar and complex unknown field (the wavefunction) is governed by a partial differential equation, at first glance so simple! Even if this last approach can be considered within the general formalism of the quantum algebra, it seems conceptually simpler and more natural because of its connections with the well established physics of waves.

Obviously the three conceptual schemes, the first two more particle-oriented and the last one clearly wave-oriented, are equivalent as proved by Paul Dirac. However, today the Schrödinger formalism continues to be more popular, being preferred by a number of scientists working in the field of solid state physics and computational chemistry. Thus, this approach constitutes a powerful tool for describing the structure and mechanics of matter.

The wave-particle duality, the Heisenberg uncertainty principle (stating that a particle cannot "have" simultaneously a position and a momentum) and the reduction of the wavefunction (typically during an observation that produces the 'instantaneous" reduction of the wavefunction in favor of an apparent material-particle), among many others, are at the origin of a number of paradoxes (some of them still open) that continue to passionate scientists and non scientists! The interested reader can refer to some of the numerous available books previously referred.

1.2 Planck versus the ultraviolet catastrophe

In this section we revisit the classical theory of black-body radiation (Rayleigh-Jeans theory) and the so-called ultraviolet catastrophe. In order to solve this major issue, Planck introduced the energy quantization, considered later by Einstein to explain the photoelectric effect.

1.2.1 Rayleigh-Jeans theory

A stationary electric field in 1D writes

$$E = E_0 \sin\left(2\pi \frac{x}{\lambda}\right) \sin(2\pi \nu t), \tag{1.1}$$

where E_0 is the amplitude and λ and ν respectively the wavelength and the frequency, with $c = \lambda \nu$ (c being the speed of light). If the stationary wave is constrained to exist in an 1D cavity of length L with perfectly conductive boundaries, implying $E(x=0) = E(x=L) = 0$, the following relation must be verified

$$\frac{2L}{\lambda} = n, \quad n = 1,2,3,\cdots, \tag{1.2}$$

or

$$\nu = \frac{cn}{2L}, \quad n = 1,2,3,\cdots. \tag{1.3}$$

Thus, the number of allowed frequencies within the interval $[\nu, \nu + d\nu]$ becomes $\frac{2L}{c} d\nu$, in fact the double $\frac{4L}{c} d\nu$ because we have two different polarizations.

When considering a cubic cavity $[0,L]^3$, and considering for a given direction expressed by the angles (α, β, γ) with respect to the cartesian axes

$$\begin{cases} \lambda = \lambda_x \cos\alpha \\ \lambda = \lambda_y \cos\beta \\ \lambda = \lambda_z \cos\gamma \end{cases}, \tag{1.4}$$

the compatibility conditions read

$$\begin{cases} \frac{2L}{\lambda_x} = n_x \\ \frac{2L}{\lambda_y} = n_y \\ \frac{2L}{\lambda_z} = n_z \end{cases}, \tag{1.5}$$

or

$$\begin{cases} \frac{2L}{\lambda} \cos\alpha = n_x \\ \frac{2L}{\lambda} \cos\beta = n_y \\ \frac{2L}{\lambda} \cos\gamma = n_z \end{cases}. \tag{1.6}$$

Summing the square of all them (with $\cos^2\alpha + \cos^2\beta + \cos^2\gamma = 1$) it results

$$\left(\frac{2L}{\lambda}\right)^2 = n_x^2 + n_y^2 + n_z^2 \rightarrow \frac{2L}{\lambda} = \sqrt{n_x^2 + n_y^2 + n_z^2}, \tag{1.7}$$

or with $r = \sqrt{n_x^2 + n_y^2 + n_z^2}$ to

$$r = \frac{2L}{c}\nu. \tag{1.8}$$

With the volume related to r being the one of the spherical shell, $4\pi r^2 dr$ divided by 8 for considering a single octant, and using the relation between r and the frequency ν, it results that the number of allowed frequencies in $[\nu, \nu + d\nu]$ is

$$\frac{\pi}{2}\left(\frac{2L}{c}\right)^3 \nu^2 d\nu, \tag{1.9}$$

in fact the double because of the two possible polarizations.

The energy density $\rho(\nu)$ (the total energy divided by the cavity volume L^3) can be computed using the energy equipartition theorem, which associates a kinetic energy (E_c) of $K_b T/2$ per degree of freedom:

$$\rho(\nu)d\nu = \frac{8\pi\nu^2 K_b T}{c^3} d\nu, \tag{1.10}$$

whose integral

$$\int_0^\infty \rho(\nu)d\nu = \int_0^\infty \frac{8\pi\nu^2 K_b T}{c^3} d\nu \tag{1.11}$$

diverges, i.e.

$$\int_0^\infty \rho(\nu)d\nu = \infty, \tag{1.12}$$

that constitutes the so-called ultraviolet catastrophe.

1.2.2 Planck theory

To avoid the divergence of the integral, Planck proposed a discrete energy distribution

$$E(\nu) = nh\nu, \quad n = 0, 1, 2, 3, \cdots, \tag{1.13}$$

which, using the Boltzmann theory, provides at equilibrium the following probability distribution:

$$p(n) = \frac{e^{-\frac{E_n}{K_b T}}}{\sum_{i=0}^{\infty} e^{-\frac{E_i}{K_b T}}}, \tag{1.14}$$

that is the probability, at equilibrium, to have n photons related to the frequency ν.

Now, the energy average $\overline{E}_\nu$ reads

$$\overline{E}_\nu = \sum_{n=0}^{\infty} E_n p(n) = \frac{\sum_{n=0}^{\infty} nh\nu e^{-\frac{E_n}{K_b T}}}{\sum_{n=0}^{\infty} e^{-\frac{E_n}{K_b T}}}, \tag{1.15}$$

that using the change of variable $x = e^{-\frac{h\nu}{K_b T}}$, becomes

$$\overline{E}_\nu = h\nu \frac{\sum_{n=0}^{\infty} nx^n}{\sum_{n=0}^{\infty} x^n} = h\nu \frac{x + 2x^2 + 3x^3 + \cdots}{1 + x + x^2 + \cdots} = h\nu x \frac{1 + 2x + 3x^2 + \cdots}{1 + x + x^2 + \cdots}. \tag{1.16}$$

Taking into account the relations

$$\begin{cases} \frac{1}{1-x} = 1 + x + x^2 + \cdots \\ \frac{1}{(1-x)^2} = 1 + 2x + 3x^2 + \cdots \end{cases}, \tag{1.17}$$

the energy average reads

$$\overline{E}_\nu = \frac{h\nu x}{1-x} = \frac{h\nu}{e^{\frac{h\nu}{K_b T}} - 1}. \tag{1.18}$$

The energy density then reads

$$\rho(\nu)d\nu = \frac{8\pi\nu^2 \overline{E}_\nu}{c^3} d\nu = \frac{8\pi\nu^2}{c^3} \frac{h\nu}{e^{\frac{h\nu}{K_b T}} - 1} d\nu, \tag{1.19}$$

whose integral does not diverge anymore

$$\rho = \int_0^\infty \rho(\nu) d\nu = \frac{8\pi^5}{15h} \left(\frac{K_b T}{h}\right)^4, \tag{1.20}$$

from which the Stefan-Boltzmann law $\rho \propto T^4$ can be noticed.

1.3 An intuitive approach to the Schrödinger equation

Even if the Schrödinger equation could be assumed as a first principe (avoiding the necessity of deriving it), in this section we prefer deriving it following a non-rigorous procedure, but with the merit of being almost intuitive.

We consider as starting point some key results:

- The de Broglie equation establishing the relation between momentum p and wavelength λ: $\lambda p = h$;
- The Einstein equation establishing the relation between energy and frequency (according to Planck): $E = h\nu$;
- The general expression of a sinusoidal traveling wave:
$$\Psi(x,t) = \sin 2\pi \left(\frac{x}{\lambda} - \nu t \right); \tag{1.21}$$
- The non-relativistic expression of energy:
$$E = \frac{p^2}{2m} + V, \tag{1.22}$$
with V the potential. Because the fact of using the non-relativistic expression of the energy, the resulting Schrödinger equation becomes non-relativistic. Dirac considered its relativistic counterpart and derived the relativistic counterpart of the Schrödinger equation (the so-called Dirac equation).

Now, we assume a linear physics and focus on the simplest scenario that concerns a particle evolving in the space with constant potential V_0. In absence of interactions (forces results from the potential gradient, here vanishing) the momentum and the energy should remain constant, according to the de Broglie and Einstein expressions $p\lambda = h$ and $E = h\nu$. Thus, the energy expression (1.22) becomes

$$\frac{h^2}{2m\lambda^2} + V_0 = \nu h, \tag{1.23}$$

or by introducing $k = \frac{2\pi}{\lambda}$, $\omega = 2\pi\nu$ and $\hbar = \frac{h}{2\pi}$, the energy expression (1.23) can be rewritten as

$$\frac{\hbar^2 k^2}{2m} + V_0 = \omega\hbar, \tag{1.24}$$

and the sinusoidal wave (1.21) reads

$$\Psi(x,t) = \sin(kx - \omega t). \tag{1.25}$$

Now, in absence of interactions, the evolution of the wavefunction $\Psi(x,t)$ is expected to be given by a travelling wave like (1.25). Its space and time derivatives read

$$\begin{cases} \frac{\partial\Psi}{\partial x} = k\cos(kx-\omega t) \\ \frac{\partial^2\Psi}{\partial x^2} = -k^2\sin(kx-\omega t) \\ \frac{\partial\Psi}{\partial t} = -\omega\cos(kx-\omega t) \end{cases}, \tag{1.26}$$

whose comparison with the energy expression (1.24) suggest

$$\alpha\frac{\partial^2\Psi}{\partial x^2} + V\Psi = \beta\frac{\partial\Psi}{\partial t}. \tag{1.27}$$

To compute the expression of α and β, the wavefunction (1.25) is substituted into (1.27) with $V = V_0$ leading to

$$-\alpha k^2\sin(kx-\omega t) + \sin(kx-\omega t)V_0 = -\beta\omega\cos(kx-\omega t), \tag{1.28}$$

that has no solution for any choice of α and β.

To increase our chances we consider a more general form of the traveling wave

$$\Psi(x,t) = \cos(kx-\omega t) + \gamma\sin(kx-\omega t). \tag{1.29}$$

Introducing this expression into (1.27) yields

$$(-\alpha k^2 + V_0 + \beta\omega\gamma)\cos(kx-\omega t) + (-\alpha k^2\gamma + V_0\gamma - \beta\omega)\sin(kx-\omega t) = 0, \tag{1.30}$$

that implies

$$\begin{cases} -\alpha k^2 + V_0 = -\beta\omega\gamma \\ -\alpha k^2 + V_0 = \frac{\beta\omega}{\gamma} \end{cases}. \tag{1.31}$$

By subtracting both equations (1.31) it results

$$-\gamma - \frac{1}{\gamma} = 0, \tag{1.32}$$

that implies

$$\gamma = -\frac{1}{\gamma}, \tag{1.33}$$

that is, $\gamma^2 = -1$, or $\gamma = \pm i$. Substituting this value in the first equation in (1.31) results

$$-\alpha k^2 + V_0 = \pm i\beta\omega, \tag{1.34}$$

whose comparison with the expression of the energy (1.24) leads to

$$\begin{cases} \alpha = -\frac{\hbar^2}{2m} \\ \beta = \pm i\hbar \end{cases}. \tag{1.35}$$

Without loss of generality, we choose β to be positive. We can now write the Schrödinger equation:

$$-\frac{\hbar^2}{2m}\frac{\partial^2\Psi}{\partial x^2}+V\Psi=i\hbar\frac{\partial\Psi}{\partial t}. \tag{1.36}$$

It is important to note that as the Heisenberg uncertainty principle, discussed in the next chapter, establishes a huge difference between classical and quantum particles (the last cannot have a defined position and momentum simultaneously), in the wave description just described, the wavefunction seems to be also a quite strange wave, with a part of it defined in the complex dimension.

1.4 The Feynman approach

Feynman considered that for a free particle the wavefunction evolves from its initial state $\Psi(y,t=0)$ to the final one $\Psi(x,t)$

$$\Psi(y,t=0)\rightarrow\Psi(x,t). \tag{1.37}$$

The assumption of linearity implies that given two evolutions

$$\begin{cases}\Psi_I(y,t=0)\rightarrow\Psi_I(x,t)\\ \Psi_{II}(y,t=0)\rightarrow\Psi_{II}(x,t)\end{cases}, \tag{1.38}$$

it results

$$\alpha\Psi_I(y,t=0)+\beta\Psi_{II}(y,t=0)\rightarrow\alpha\Psi_I(x,t)+\beta\Psi_{II}(x,t). \tag{1.39}$$

From the Schwartz kernel theorem we can write

$$\Psi(x,t)=\int_y\mathscr{G}(x,y;t)\Psi(y,t=0)dy, \tag{1.40}$$

with the kernel function $\mathscr{G}(\bullet)$ independent on the origin of the coordinate system, i.e. $\mathscr{G}(x,y;t)=\mathscr{G}(x-y;t)$, and $\mathscr{G}(x,y;t)=\mathscr{G}\left((x-y)^2;t\right)$ in the case of assuming isotropy. A dimensional analysis allows writing

$$\mathscr{G}(x,y;t)=\mathscr{G}\left(\frac{(x-y)^2m}{\hbar t}\right). \tag{1.41}$$

Now, we consider a final state attained from two different paths

$$\Psi(z;t+t')=\int_y\mathscr{G}(z,y;t+t')\Psi(y,t=0)dy, \tag{1.42}$$

and

$$\Psi(z;t+t')=\int_x\mathscr{G}(z,x;t')\Psi(x,t)dx, \tag{1.43}$$

that imply the composition rule

$$\mathscr{G}(z,y;t+t') = \int_x \mathscr{G}(z,x;t')\mathscr{G}(x,y;t)dx, \tag{1.44}$$

illustrated in Fig. 1.1.

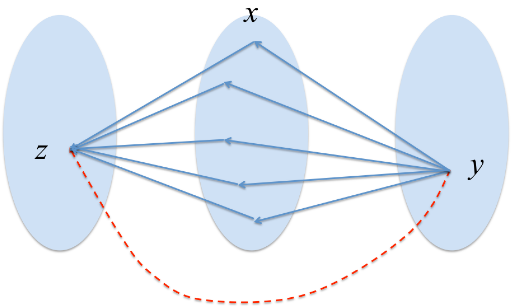

Fig. 1.1 Evolution from $\Psi(y;t=0)$ to $\Psi(z;t+t')$.

Using the previous results (1.41) and (1.44) with the normality condition

$$\int_x \overline{\Psi}(x,t)\Psi(x,t)dx = 1, \tag{1.45}$$

with $\overline{\Psi}$ the conjugate of Ψ, the expression of the kernel can be derived

$$\mathscr{G}(x,y;t) = \frac{i}{t}\frac{m}{2\pi\hbar}e^{\frac{i}{t}\frac{(x-y)^2 m}{\hbar}}, \tag{1.46}$$

where reversibility is ensured because $\frac{i}{t} = \frac{-i}{-t}$.

Thus, if the path $y \rightarrow x$ implies a phase change

$$\varphi = \frac{(x-y)^2 m}{\hbar t}, \tag{1.47}$$

$x \rightarrow z$ increases the phase change

$$\varphi' = \varphi + \frac{(z-x)^2 m}{\hbar t'}, \tag{1.48}$$

but because the integral is performed at the intermediate state according to Eq. (1.44) an interference appears. Thus, if for George Tompson the electron interference experience was the proof of its wave nature, for Feynman the interference was simply the signature of a particle that is taking simultaneously all the possible trajectories interacting with itself (interaction is created by the integral in (1.44)).

Now, from the kernel identified, the differential equation governing the evolution of the wavefunction can be obtained, and as expected it coincides with the Schrödinger equation.

1.5 The Schrödinger equation

In classical mechanics the evolution of a particle can be determined from the knowledge of its position $\mathbf{r}$ and velocity $\mathbf{v}$ (or its momentum $\mathbf{p} = m\mathbf{v}$, with m the particle mass) at a point of its trajectory. However, in quantum mechanics position and momentum can not be associated simultaneously to a particle, according to the Heisenberg uncertainty principle. Thus, more than speaking about position or momentum we should speak about probability distribution related to position $P_r = |\Psi|^2$ and momentum $P_p = |\Psi_p|^2$. The necessity of considering the square of Ψ or Ψ_p results from the fact that both functions are complex and the resulting probabilities P_r and P_p must be necessarily real numbers.

Precisely because of the uncertainty principle, these two wavefunctions Ψ and Ψ_p are not independent, the first one being the Fourier's transform of the second one. The properties of the Fourier transform guarantee that the higher the localization in space, the lower its localization in the momentum variable and vice-versa [51]. Of course, from the physical point of view this fact induces a conceptual difficulty because in classical mechanics two informations are needed to determine the evolution of the system (i.e. position and velocity) and the quantum system seems requiring only one Ψ (or Ψ_p, because both are related from the Fourier transform). This paradox is only apparent because both wavefunctions are complex, and then they have a real and an imaginary part. Thus, the knowledge of a quantum system requires knowing the evolution of the wavefunction Ψ (or Ψ_p) introduced by Erwin Schrödinger in the twenties of the past century, that only needs the solution of the equation governing its distribution or its evolution in the transient case.

For the sake of simplicity we are first introducing the Schrödinger equation considering neither the relativistic effects nor the spin. If we assume a system composed of N_P particles, the evolution of the joint wavefunction $\Psi = \Psi(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_{N_P}, t)$ is governed by the Schrödinger equation whose dimensionless form results [22]

$$i\frac{\partial \Psi}{\partial t} = -\sum_{p=1}^{N_p} \frac{1}{2m_p} \nabla_p^2 \Psi + \sum_{p=1}^{N_p} \sum_{k=p+1}^{N_p} V_{pk} \Psi, \tag{1.49}$$

where each particle p is distributed in the whole physical space $\mathbf{r}_p = (x_p, y_p, z_p) \in \mathbb{R}^3$ and $i=\sqrt{-1}$. The differential operator ∇_p^2 is defined in the conformational space of each particle, i.e.

$$\nabla_p^2 = \frac{\partial^2}{\partial x_p^2} + \frac{\partial^2}{\partial y_p^2} + \frac{\partial^2}{\partial z_p^2}. \tag{1.50}$$

The Coulomb's potential describing the inter-particle interactions writes

$$V_{pk} = \frac{q_p \cdot q_k}{||\mathbf{r}_p - \mathbf{r}_k||}, \tag{1.51}$$

where the masses m_p are unity for electrons, charges q_j are -1 for electrons and $+Z_j$ (atomic numbers) for nuclei.

Equation (1.49) can be rewritten in the compact form

$$(\mathscr{H} - \mathscr{T})\,\Psi = 0, \tag{1.52}$$

where $\mathscr{H}$ denotes the Hamiltonian operator and $\mathscr{T} = -i\frac{\partial}{\partial t}$. This equation was proposed, not deduced, and it is today considered as a first principle, like the Newton equation in classical mechanics.

It is important to note that Eq. (1.49) can be integrated from an initial condition $\Psi(\mathbf{x}, t=0)$, leading to $\Psi(\mathbf{x}, t)$, with $t \in (0, T]$. If we consider $0 < \tau < T$, and we do not observe the state of the system at $t = \tau$, then for $t > \tau$, $\Psi(\mathbf{x}, t)$ corresponds to the solution of Eq. (1.52) with the initial condition $\Psi(\mathbf{x}, t=0)$. However, if we observe the system at time $t = \tau$, the observation process modifies irreversibly the wavefunction, that is, $\Psi(\mathbf{x}, \tau - \varepsilon) \neq \Psi(\mathbf{x}, \tau + \varepsilon)$ (with $\varepsilon \to 0$). Then, after the observation the system is described by the solution of Eq. (1.52) with the initial condition $\Psi(\mathbf{x}, \tau + \varepsilon)$ that differs from the solution $\Psi(\mathbf{x}, t)$ computed from $\Psi(\mathbf{x}, t=0)$.

By assuming the separation of variables the wavefunction can be decomposed according to

$$\Psi = \Psi_r\left(\mathbf{r}_1, \ldots, \mathbf{r}_{N_p}\right) \cdot \Psi_t(t). \tag{1.53}$$

Introducing this expression into the Schrödinger equation (1.52) and dividing by $\Psi_r \cdot \Psi_t$ yields

$$\frac{\Psi_t \cdot \mathscr{H}(\Psi_r)}{\Psi_r \cdot \Psi_t} - \frac{\Psi_r \cdot \mathscr{T}(\Psi_t)}{\Psi_r \cdot \Psi_t} = 0, \tag{1.54}$$

or

$$\frac{\mathscr{H}(\Psi_{\mathbf{r}})}{\Psi_r} - \frac{\mathscr{T}(\Psi_t)}{\Psi_t} = 0. \tag{1.55}$$

The first term depends on the space variables and the second one on the time, implying that both terms are equal to a constant E (that as discussed later represents the energy), that is

$$\begin{cases} \frac{\mathscr{T}(\Psi_t)}{\Psi_t} = E \\ \\ \frac{\mathscr{H}(\Psi_r)}{\Psi_r} = E \end{cases}. \tag{1.56}$$

The integration of the first relation in (1.56) implies

$$\Psi_t(t) = A\ \mathrm{e}^{-iEt}, \tag{1.57}$$

and the second one leads to the following eigenproblem

$$\mathscr{H}(\Psi_r(\mathbf{r}_1, \ldots, \mathbf{r}_{Np})) = E\Psi_r(\mathbf{r}_1, \ldots, \mathbf{r}_{Np}). \tag{1.58}$$

The eigenfunctions $\Psi_r^{(n)}(\mathbf{r}_1,\ldots,\mathbf{r}_{Np})$ related to the eigenvalues E_n define an orthogonal basis. Obviously, each function

$$\Psi^{(n)} = A_n\, e^{-iE_n t} \cdot \Psi_r^{(n)}(\mathbf{r}_1,\ldots,\mathbf{r}_{Np}), \tag{1.59}$$

is time-independent because the physical meaning of the solution is done by the complex norm of the wavefunction, that is

$$|\Psi^{(n)}| = \overline{\Psi^{(n)}} \cdot \Psi^{(n)} = A_n^2\, e^{-iE_n t} \cdot e^{iE_n t} \cdot \overline{\Psi_r^{(n)}} \cdot \Psi_r^{(n)} = A_n^2\, \overline{\Psi_r^{(n)}} \cdot \Psi_r^{(n)}, \tag{1.60}$$

where $\overline{a}$ denotes the conjugate of a.

Of course, the general transient solution can be written as

$$\Psi(\mathbf{r}_1,\ldots,\mathbf{r}_{N_p},t) = \sum_{n=1}^{\infty} A_n\, e^{-iE_n t} \cdot \Psi_r^{(n)}(\mathbf{r}_1,\ldots,\mathbf{r}_{N_p}), \tag{1.61}$$

that obviously is not time-independent anymore.

In what follows, we assume a system composed of N_e electrons and N_n nuclei, and a time-independent solution of the Schrödinger equation. The Born-Oppenheimer model writes the total wavefunction of the combined system (electrons and nuclei) as a sum of products of wavefunctions related to the electrons and to the nuclei. Born and Oppenheimer showed that the total wavefunction can be reasonably approximated from a single product of a nuclei wavefunction $\Psi_n = \Psi_n(\mathbf{R}_1,\ldots,\mathbf{R}_{N_n})$ and an electronic wavefunction $\Psi_e = \Psi_e(\mathbf{r}_1,\ldots,\mathbf{r}_{N_e};\mathbf{R}_1,\ldots,\mathbf{R}_{N_n})$ that depends parametrically on the nuclei coordinates. Introducing $\Psi = \Psi_n \cdot \Psi_e$ into the Schrödinger's eigenproblem $\mathscr{H}(\Psi) = E\Psi$ it results

$$\Psi_n \cdot \mathscr{H}_e(\Psi_e) + \Psi_e \cdot \mathscr{H}_n(\Psi_n) = E\ \Psi_e \cdot \Psi_n, \tag{1.62}$$

or

$$\frac{\mathscr{H}_e(\Psi_e)}{\Psi_e} + \frac{\mathscr{H}_n(\Psi_n)}{\Psi_n} = E, \tag{1.63}$$

where the first term depends on the electron coordinates and the second one on the nuclei coordinates. Thus we can write

$$\frac{\mathscr{H}_e(\Psi_e)}{\Psi_e} = E_e, \tag{1.64}$$

implying

$$\mathscr{H}_e(\Psi_e) = E_e\ \Psi_e, \tag{1.65}$$

and

$$\frac{\mathscr{H}_n(\Psi_n)}{\Psi_n} = E - E_e, \tag{1.66}$$

that results in

$$(\mathscr{H}_n + E_e)\Psi_n = E\ \Psi_n. \tag{1.67}$$

The difference between the dynamics of electron and nuclei is derived in the following way: (i) the nuclei are typically tens of thousands times heavier that the electrons; (ii) the particles constituting the molecules are in equilibrium so, in average, they have similar kinetic energies (equipartition theorem); (iii) thus the ratio of the square of their velocities (electrons and nuclei) will be roughly the inverse ratio of their masses (velocity ratio of the order of hundred); (iv) then the wavelength associated with the nuclei is a hundred of times lower that the one corresponding to the electrons. The last conclusion comes from the fact that according to de Broglie $\lambda_n^{-1} \sim p_n = m_n v_n \approx 10^4 m_e 10^{-2} v_e = 10^2 p_e \sim 10^2 \lambda_e^{-1}$ (where $\bullet_n$ and $\bullet_e$ refer respectively to the nuclei and electrons properties).

Thus, electrons see stationary nuclei and the nuclei see electrons distributed (in a sort of mean). From now on we assume the nuclei fixed in the physical space and write

$$\mathscr{H}(\Psi(\mathbf{r}_1,\ldots,\mathbf{r}_{N_e};\mathbf{R}_1,\ldots,\mathbf{R}_{N_n})) = E\ \Psi(\mathbf{r}_1,\ldots,\mathbf{r}_{N_e};\mathbf{R}_1,\ldots,\mathbf{R}_{N_n}), \tag{1.68}$$

where the parametrical dependence of the electronic distribution on the nuclei position has been emphasized. The Hamiltonian can be written as

$$\mathscr{H}(\Psi) = -\sum_{e=1}^{N_e} \frac{1}{2}\nabla_e^2 \Psi + \sum_{e=1}^{N_e}\sum_{n=1}^{N_n} V_{en}\Psi + \sum_{e=1}^{N_e}\sum_{e'=e+1}^{N_e} V_{ee'}\Psi, \tag{1.69}$$

where

$$V_{ee'} = \frac{1}{||\mathbf{r}_e - \mathbf{r}_{e'}||}, \tag{1.70}$$

and

$$V_{en} = -\frac{Z}{||\mathbf{r}_e - \mathbf{R}_n||}. \tag{1.71}$$

1.6 Relations between position and momentum wavefunctions

As previously indicated $\Psi(\mathbf{x},t)$ and $\Psi_p(\mathbf{p},t)$ are Fourier transforms of each other in order to fulfill the Heisenberg uncertainty principle addressed later. Thus, proceeding in the 1D case for the sake of simplicity we have

$$\Psi_p(p,t) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} \Psi(x,t) e^{-i\frac{px}{\hbar}}\, dx = \mathscr{F}(\Psi(x,t)), \tag{1.72}$$

where $\mathscr{F}(\bullet)$ refers the Fourier transform and the use of the constant $\hbar$ ($\hbar = \frac{\mathrm{h}}{2\pi}$) is introduced into the Fourier transform definition for addressing later the Heisenberg uncertainty principle. The inverse transformation writes

$$\Psi(x,t) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} \Psi_p(p,t) e^{i\frac{px}{\hbar}} dx. \tag{1.73}$$

This definition of the Fourier transform is not practical when using standard convolution expressions, but this is not an issue here and moreover the convolution expression can be modified accordingly.

Using the Fourier transform properties we can write

$$p^n \Psi_p = \mathscr{F}\left(\left(\frac{\hbar}{i}\right)^n \frac{\partial^n \Psi}{\partial x^n}\right). \tag{1.74}$$

In what follows we will also make use of the Parseval's theorem that for $\Psi_p = \mathscr{F}(\Psi)$ and $\Phi_p = \mathscr{F}(\Phi)$ states

$$\int_{-\infty}^{\infty} \overline{\Psi}_p \cdot \Phi_p \, dp = \int_{-\infty}^{\infty} \overline{\Psi} \cdot \Phi \, dx. \tag{1.75}$$

1.6.1 Calculating expectations

As soon as the electronic distribution Ψ is known one could compute the one related to the electronic momentum Ψ_p and from both the different expected values of position, momentum, energy, etc. Imagine a simple system consisting of a single electron in the one-dimensional space, i.e. $\mathbf{r} = x$. Thus, writing: $P_x = |\Psi|^2$ and $P_p = |\Psi_p|^2$ the expected value for the position reads

$$\langle x \rangle = \int_{\mathbb{R}} x P_x \, dx = \int_{\mathbb{R}} x \, \overline{\Psi}(x) \cdot \Psi(x) \, dx, \tag{1.76}$$

with $\mathbb{R} = (-\infty, \infty)$.

The analogue expression for the momentum results

$$\langle p \rangle = \int_{\mathbb{R}} p P_p \, dx = \int_{\mathbb{R}} p \, \overline{\Psi}_p(p) \cdot \Psi_p(p) \, dp, \tag{1.77}$$

that using Eq. (1.74) and the Parseval theorem results

$$\langle p \rangle = \int_{\mathbb{R}} p P_p \, dp = \int_{\mathbb{R}} \overline{\Psi}(x) \frac{\hbar}{i} \frac{\partial \Psi(x)}{\partial x} \, dx. \tag{1.78}$$

Analogously, we can compute $\langle p^2 \rangle$

$$\langle p^2 \rangle = \int\limits_{\mathbb{R}} \overline{\Psi}(x) \left(\frac{\hbar}{i}\right)^2 \frac{\partial^2 \Psi(x)}{\partial x^2}\ dx. \tag{1.79}$$

Thus, for a polynomial $g(p)$ it results

$$\langle g \rangle = \int\limits_{\mathbb{R}} \overline{\Psi}(x)\ g\left(\frac{\hbar}{i}\frac{\partial}{\partial x}\right) \Psi(x)\ dx, \tag{1.80}$$

where $g\left(\frac{\hbar}{i}\frac{\partial}{\partial x}\right)$ refers the replacement of p in the polynomial $g(p)$ by the differential operator $\frac{\hbar}{i}\frac{\partial}{\partial x}$.

By considering $f(x)$ and $g(p)$ and using the linearity it results

$$\langle f + g \rangle = \int\limits_{\mathbb{R}} \overline{\Psi}(x)\ \left(f + g\left(\frac{\hbar}{i}\frac{\partial}{\partial x}\right)\right) \Psi(x)\ dx. \tag{1.81}$$

When applied to the Hamiltonian $\mathscr{H}(x,p)$

$$\mathscr{H} = \frac{p^2}{2m} + V(x), \tag{1.82}$$

it results

$$\langle \mathscr{H} \rangle = \int\limits_{\mathbb{R}} \overline{\Psi}(x) \mathscr{H} \Psi(x)\ dx = \int\limits_{\mathbb{R}} \overline{\Psi}(x)\ \left(V(x) - \frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}\right) \Psi(x)\ dx. \tag{1.83}$$

All the previous results can be easily extended to 3D scenarios. Thus, as soon as the wavefunction $\Psi(\mathbf{x},t)$ is known, the different expectations related to position, momentum and/or energy can be easily computed.

1.7 Heisenberg uncertainty principle

If we assume for a while (and without loss of generality) $\langle x \rangle = 0$ and $\langle p \rangle = 0$,

$$(\Delta x)^2 = \langle (x - \langle x \rangle)^2 \rangle = \langle x^2 \rangle, \tag{1.84}$$

and similarly

$$(\Delta p)^2 = \langle p^2 \rangle. \tag{1.85}$$

If we define the integral

$$\mathscr{I}(\alpha) = \int_{\mathbb{R}} \left| \left(\alpha\hbar\frac{\partial}{\partial x} - x\right)\Psi \right|^2\ dx \geq 0, \tag{1.86}$$

its expansion reads

$$\mathscr{I}(\alpha) = \int_{\mathbb{R}} \left(\alpha\hbar \frac{\partial}{\partial x} - x \right) \overline{\Psi} \cdot \left(\alpha\hbar \frac{\partial}{\partial x} - x \right) \Psi \, dx =$$

$$\alpha^2\hbar^2 \int_{\mathbb{R}} \frac{\partial \overline{\Psi}}{\partial x} \frac{\partial \Psi}{\partial x} \, dx + \int_{\mathbb{R}} x^2 \overline{\Psi} \Psi \, dx - \alpha\hbar \left(\int_{\mathbb{R}} \frac{\partial \overline{\Psi}}{\partial x} x \Psi \, dx + \int_{\mathbb{R}} x \overline{\Psi} \frac{\partial \Psi}{\partial x} \right). \quad (1.87)$$

Now, by integrating by parts the first term in the right-hand side and taking into account that the wavefunction vanishes when $\mathbf{x} \to \infty$, it results

$$\hbar^2 \int_{\mathbb{R}} \frac{\partial \overline{\Psi}}{\partial x} \frac{\partial \Psi}{\partial x} \, dx = -\hbar^2 \int_{\mathbb{R}} \overline{\Psi} \frac{\partial^2 \Psi}{\partial x^2} \, dx = (\Delta p)^2. \quad (1.88)$$

On the other hand, integrating by parts the third term yields

$$\int_{\mathbb{R}} \frac{\partial \overline{\Psi}}{\partial x} x \Psi \, dx = - \int_{\mathbb{R}} \overline{\Psi} \left(\Psi + x \frac{\partial \Psi}{\partial x} \right) \, dx = -1 - \int_{\mathbb{R}} \overline{\Psi} x \frac{\partial \Psi}{\partial x} \, dx, \quad (1.89)$$

leading to the polynomial expression

$$\mathscr{I}(\alpha) = \alpha^2 (\Delta p)^2 + \alpha\hbar + (\Delta x)^2 \geq 0. \quad (1.90)$$

Since this quantity is nonnegative, the discriminant of the quadratic polynomial

$$\hbar^2 - 4(\Delta p)^2 (\Delta x)^2 \leq 0 \quad (1.91)$$

is negative, that leads to the usual expression of the Heisenberg uncertainty principle

$$(\Delta p)^2 (\Delta x)^2 \geq \frac{\hbar^2}{4}. \quad (1.92)$$

When $\langle x \rangle \neq 0$ and/or $\langle p \rangle \neq 0$ we obtain the same result as proved in [51]. All the previous results can be easily extended to 3D scenarios.

The main consequences are that quantum particles are quite different to their classical counterpart, because in quantum theory positon and velocity cannot exist simultaneous, if position is perfectly defined (it exists in a classical sense) the ‘quantum particle” has a totally undetermined velocity and vice-versa.

It exists another uncertainty relationship relating energy and time (even if the last is not an operator). This expression gives an approximate value for the length of time taken for the energy to change by one standard deviation. In order to get an accurate energy measurement it should change slowly. This result is of major interest for explaining the quantum tunnel effect.

1.8 Observable and its time evolution

The expected value of any observable F (Hermitian operator on a complex vector space, like position, energy, ...),

$$\langle F \rangle = \int_{\mathbb{R}} \overline{\Psi} F \Psi \, dx, \tag{1.93}$$

can be written in a more compact form as

$$\langle F \rangle = (\Psi, F\Psi), \tag{1.94}$$

where

$$(\Psi, \Phi) = \int_{\mathbb{R}} \overline{\Psi} \cdot \Phi \, dx. \tag{1.95}$$

When considering the time-independent Hamiltonian observable $\langle \mathscr{H} \rangle = (\Psi, \mathscr{H}\Psi)$, the time evolution of its expectation is given by

$$\frac{d\langle \mathscr{H} \rangle}{dt} = \left(\frac{\partial \Psi}{\partial t}, \mathscr{H}\Psi \right) + \left(\Psi, \mathscr{H} \frac{\partial \Psi}{\partial t} \right), \tag{1.96}$$

where the time-independent nature of $\mathscr{H}$ implies $\frac{\partial \mathscr{H}}{\partial t} = 0$.

By integrating by parts and taking into account both the Hermitian character of observables (the Hamiltonian in the present case) and the fact that Ψ vanishes when $x \to \infty$ gives

$$\frac{d\langle \mathscr{H} \rangle}{dt} = \left(\frac{\partial \Psi}{\partial t}, \mathscr{H}\Psi \right) + \left(\mathscr{H}\Psi, \frac{\partial \Psi}{\partial t} \right), \tag{1.97}$$

that using the Schrödinger equation $\mathscr{H}\Psi = i\hbar \frac{\partial \Psi}{\partial t}$ leads to

$$\frac{d\langle \mathscr{H} \rangle}{dt} = \left(\frac{\partial \Psi}{\partial t}, i\hbar \frac{\partial \Psi}{\partial t} \right) + \left(i\hbar \frac{\partial \Psi}{\partial t}, \frac{\partial \Psi}{\partial t} \right) = 0, \tag{1.98}$$

that states the energy conservation.

For any time-independent observable $\mathscr{O}$ ($\mathscr{O}$ does not depend on t), it results that

$$\langle \mathscr{O} \rangle = \int_{\mathbb{R}} \overline{\Psi} \mathscr{O} \Psi \, dx, \tag{1.99}$$

or using the notation previously introduced

$$\langle \mathscr{O} \rangle = (\Psi, \mathscr{O}\Psi). \tag{1.100}$$

The time evolution of observable $\mathscr{O}$ reads

$$\frac{d\langle \mathscr{O} \rangle}{dt} = \left(\frac{\partial \Psi}{\partial t}, \mathscr{O}\Psi \right) + \left(\Psi, \mathscr{O} \frac{\partial \Psi}{\partial t} \right), \tag{1.101}$$

that using again the Schrödinger equation gives

$$\frac{d\langle\mathscr{O}\rangle}{dt} = \left(-\frac{i}{\hbar}\mathscr{H}\Psi, \mathscr{O}\Psi\right) + \left(\Psi, -\frac{i}{\hbar}\mathscr{O}\mathscr{H}\Psi\right), \tag{1.102}$$

that can be rewritten as

$$\frac{d\langle\mathscr{O}\rangle}{dt} = \frac{i}{\hbar}\left((\mathscr{H}\Psi, \mathscr{O}\Psi) - (\Psi, \mathscr{O}\mathscr{H}\Psi)\right), \tag{1.103}$$

or

$$\frac{d\langle\mathscr{O}\rangle}{dt} = \frac{i}{\hbar}\left((\Psi, \mathscr{H}\mathscr{O}\Psi) - (\Psi, \mathscr{O}\mathscr{H}\Psi)\right), \tag{1.104}$$

that finally leads to

$$\frac{d\langle\mathscr{O}\rangle}{dt} = \frac{i}{\hbar}\left(\Psi, (\mathscr{H}\mathscr{O} - \mathscr{O}\mathscr{H})\Psi\right), \tag{1.105}$$

or its compact counterpart

$$\frac{d\langle\mathscr{O}\rangle}{dt} = \frac{i}{\hbar}\langle[\mathscr{H}, \mathscr{O}]\rangle, \tag{1.106}$$

where the so-called commutator [•] (Poisson bracket) is defined as

$$[\mathscr{H}, \mathscr{O}] = \mathscr{H}\mathscr{O} - \mathscr{O}\mathscr{H}. \tag{1.107}$$

When the observable depends explicitly on time it results that

$$\frac{d\langle\mathscr{O}\rangle}{dt} = \left\langle\frac{d\mathscr{O}}{dt}\right\rangle + \frac{i}{\hbar}\langle[\mathscr{H}, \mathscr{O}]\rangle. \tag{1.108}$$

1.8.1 The Ehrenfest theorem

From these results, it is very easy to prove the Ehrenfest's theorem that states

$$\begin{cases} \langle F\rangle = \frac{d\langle p\rangle}{dt} \\ \langle p\rangle = m\frac{d\langle x\rangle}{dt} \end{cases}, \tag{1.109}$$

establishing a perfect analogy between classical and quantum mechanics, in the sense that quantum expectation fulfills classical Newton mechanics.

To prove the previous relations we consider both of them

$$\frac{d\langle p\rangle}{dt} = \frac{i}{\hbar}\left\langle\left[\frac{p^2}{2m} + V(x), p\right]\right\rangle, \tag{1.110}$$

and

$$\frac{d\langle x\rangle}{dt} = \frac{i}{\hbar}\left\langle\left[\frac{p^2}{2m} + V(x), x\right]\right\rangle, \tag{1.111}$$

respectively, where the fact that position and momentum operators do not have time dependences was taken into account. By developing the first relation, taking into account that the momentum commutes with itself and substituting the momentum by $p \to -i\hbar \frac{\partial}{\partial x}$, it results that

$$\frac{d\langle p\rangle}{dt} = \int_{\mathbb{R}} \overline{\Psi} V(x) \frac{\partial \Psi}{\partial x}\, dx - \int_{\mathbb{R}} \overline{\Psi} \frac{\partial}{\partial x} \left(V(x)\Psi\right)\, dx. \tag{1.112}$$

Developing the second integral in the previous equation, it finally results

$$\frac{d\langle p\rangle}{dt} = -\int_{\mathbb{R}} \overline{\Psi} \frac{\partial V(x)}{\partial x} \Psi\, dx, \tag{1.113}$$

or

$$\frac{d\langle p\rangle}{dt} = -\left\langle \frac{\partial V(x)}{\partial x} \right\rangle = \langle F\rangle, \tag{1.114}$$

that generalizes Newton's second law.

Now, coming back to the second expression in Eq. (1.109) and by using the commutation relation (that again reflects the Heisenberg uncertainty principle)

$$[p,x] = -i\hbar, \tag{1.115}$$

it results

$$\frac{d\langle x\rangle}{dt} = \frac{i}{\hbar} \left\langle \left[\frac{p^2}{2m} + V(x), x \right] \right\rangle = \frac{i}{\hbar} \left\langle \left[\frac{p^2}{2m}, x \right] \right\rangle = \frac{i}{2m\hbar} \langle p\,[p,x] + [p,x]p\rangle = \frac{1}{m}\langle p\rangle. \tag{1.116}$$

Again, all the previous results can be easily extended to the 3D case.

1.9 Charge density and interatomic potentials: The Hellmann-Feynman theorem

As soon as the electronic distribution is known, the electron density $\rho_e(\mathbf{r})$ can be easily computed by applying

$$\rho_e(\mathbf{r}) = \int\limits_{\mathbb{R}^{3(N_e-1)}} \overline{\Psi} \cdot \Psi\, d\mathbf{r}_1 \cdots d\mathbf{r}_{e-1} d\mathbf{r}_{e+1} \cdots d\mathbf{r}_{N_e}, \tag{1.117}$$

allowing us to compute the electronic density at each point of the space by adding all the electron contributions according to

$$\rho(\mathbf{r}) = \sum_{e=1}^{N_e} \rho_e(\mathbf{r}). \tag{1.118}$$

Now, the Hellmann-Feynman theorem [51] leads to a direct interpretation of the electronic effects on the nuclei in terms of the classical electrostatics. The applied force on a nucleus k results

$$\mathbf{f}_k = -\nabla_k(\widetilde{V}_{nk} + \widetilde{V}_{ek}), \tag{1.119}$$

where $\widetilde{V}_{nk}$ is the electrostatic contribution inter-nuclei

$$\widetilde{V}_{nk} = \sum_{n=1, n\neq k}^{N_n} \frac{Z^2}{\|\mathbf{R}_n - \mathbf{R}_k\|}, \tag{1.120}$$

and $\widetilde{V}_{ek}$ is the one associated with the electronic distribution:

$$\widetilde{V}_{ek} = -Z\left(\int_{\mathbb{R}^3} \frac{\rho(\mathbf{r})}{\|\mathbf{r} - \mathbf{R}_k\|} d\mathbf{r}\right). \tag{1.121}$$

Thus, it seems clear that knowing the electronic distribution, the electronic density can be computed and from it the effects that the electronic distribution has on the nuclei. The solution of the Schrödinger equation allows computing the exact inter-atomic potentials that could be then used to perform accurate molecular dynamics simulations. However, the multi-dimensional character of the Schrödinger equation makes very difficult such solution even for moderate populations of electrons. Two possibilities exist: (i) the one that consists in introducing some simplifications in the treatment of that equation; or (ii) the one related to the proposal of empirical (or quantum inspired) inter-atomic potentials to be used in the molecular dynamics framework as discussed in Chapter 3. *Ab initio* simulations are performed using quantum-based potential whereas molecular dynamics uses empirical potentials.

Before analyzing the state of the art on the computational treatment of the Schrödinger equation we are introducing a last key concept, the one related to the Pauli exclusion principle.

1.10 The Pauli exclusion principle

The wavefunction of many-electrons must reflect the fact that electrons are indistinguishable. Electrons do not know that we have labeled them. If we use $\mathbf{r}_1$ to describe the coordinates of one electron, and $\mathbf{r}_2$ for another, then

$$|\Psi(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \ldots, \mathbf{r}_{Ne})|^2 = |\Psi(\mathbf{r}_2, \mathbf{r}_1, \mathbf{r}_3, \ldots, \mathbf{r}_{N_e})|^2, \tag{1.122}$$

that is, the probability distribution of electrons cannot depend on the way that we have labeled them. Thus, if Π is any permutation of the N_e electronic coordinates, then

$$\Pi|\Psi(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \ldots, \mathbf{r}_{Ne})|^2 = |\Psi(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \ldots, \mathbf{r}_{N_e})|^2, \tag{1.123}$$

that implies just two possibilities

$$\Pi\Psi(\mathbf{r}_1,\mathbf{r}_2,\mathbf{r}_3,\ldots,\mathbf{r}_{Ne}) = \Psi(\mathbf{r}_1,\mathbf{r}_2,\mathbf{r}_3,\ldots,\mathbf{r}_{N_e}), \tag{1.124}$$

or

$$\Pi\Psi(\mathbf{r}_1,\mathbf{r}_2,\mathbf{r}_3,\ldots,\mathbf{r}_{Ne}) = -\Psi(\mathbf{r}_1,\mathbf{r}_2,\mathbf{r}_3,\ldots,\mathbf{r}_{N_e}), \tag{1.125}$$

that can be written as

$$\Pi\Psi(\mathbf{r}_1,\mathbf{r}_2,\mathbf{r}_3,\ldots,\mathbf{r}_{Ne}) = a\Psi(\mathbf{r}_1,\mathbf{r}_2,\mathbf{r}_3,\ldots,\mathbf{r}_{N_e}), \tag{1.126}$$

where the exclusion Pauli principe states that for electrons $a = -\mathbf{1}$.

Thus, the most general statement of the Pauli principle for electrons establishes that an acceptable wavefunction for many electrons must be antisymmetric with respect to the exchange of the coordinates of any two electrons.

If we imagine a system composed of two non-interacting electrons ($V_{e_1e_2} = 0$) the resulting wavefunction can be written in the one-term separated form

$$\Psi(\mathbf{r}_1,\mathbf{r}_2) = \Psi_1(\mathbf{r}_1)\cdot\Psi_2(\mathbf{r}_2), \tag{1.127}$$

but because of the Pauli exclusion constraint this function must be adjusted to ensure its antisymmetry ($a = -1$ in Eq. (1.126)) by considering the Slater determinant

$$\Psi(\mathbf{r}_1,\mathbf{r}_2) = \Psi_1(\mathbf{r}_1)\cdot\Psi_2(\mathbf{r}_2) - \Psi_1(\mathbf{r}_2)\cdot\Psi_2(\mathbf{r}_1) = \begin{vmatrix} \Psi_1(\mathbf{r}_1) & \Psi_2(\mathbf{r}_1) \\ \Psi_1(\mathbf{r}_2) & \Psi_2(\mathbf{r}_2) \end{vmatrix}, \tag{1.128}$$

where the use of the determinant ensures that changing the electron's label corresponds to exchanging the two associated columns in the determinant, and the associated change in the wavefunction sign. However, this strategy introduces some technical difficulties by increasing the number of terms involved in the wavefunction expression.

Until now we have not included anything in the description to take into account the electron spin. One could think that as there is no mention of the spin in the Hamiltonian, the spin should have little relevance to the energy of the system. However, the Pauli principle applies to the exchange of all the coordinates (space and spin) so in order to comply with it the spin must be introduced into the particle coordinates.

If we do not introduce the spin into the description of a system composed of two electrons, the solution must verify: $\Psi(\mathbf{r}_1,\mathbf{r}_2) = -\Psi(\mathbf{r}_2,\mathbf{r}_1)$, but if the spin is considered, the expression $\Psi(\mathbf{r}_1,\mathbf{r}_2) = \Psi(\mathbf{r}_2,\mathbf{r}_1)$ holds if the two electrons have opposite spin.

We can give a simple solution to this issue by including the spin into the particles coordinates by defining the extended coordinates $\mathbf{x}_i$ from

$$\mathbf{x}_i = (\mathbf{r}_i, s_i) = (x_i, y_i, z_i, s_i), \tag{1.129}$$

where s_i is the spin coordinate of electron i, that can take two values that we denote by α and β (in fact $-1/2$ and $1/2$ in the case of electrons). Thus, two electrons can be allocated in the same orbital as soon as they have different spin. Pauli principle is of major importance for explaining the electronic structure of matter and also atomic bonding.

Remark. When one is solving the transient Schrödinger equation, if the initial condition verifies the Pauli exclusion principle, the computed transient solution will verify the antisymmetry without considering any particular treatment. However, when the time-independent solution is searched the antisymmetry must be enforced, and the most common way to enforce it is the use of the Slater's determinant. It is important to note that the computation of transient solutions needs very efficient integration schemes able to ensure the energy conservation while guaranteeing the stability.

1.11 On the numerical solution of the Schrödinger equation

Due to the high dimensionality of the Schrödinger equation its solution is only possible for very reduced populations of electrons. For this reason different approximated methods have been proposed and extensively used. We are describing in the next chapter the two most widely used: the Hartree-Fock (HF) method and the Density Functional Theory (DFT), the first one can be applied only for a moderate number of electrons and it has been extensively used in the context of quantum chemistry to analyze the structure and behavior of molecules. However, in the case of crystals, the number of electrons becomes too large to make possible their simulation by the direct solution of the Schrödinger equation or by using the Hartree-Fock technique. In this context the most successful technique was and continues to be the DFT. Both techniques (HF and DFT) are un practice approximated techniques that work in some cases and in other cases could become crude approximations.

The three techniques (direct Schrödinger solution, HF and DFT) can be applied for solving both the time-independent and the transient Schrödinger equations. Transient solutions are very common in the context of quantum gas dynamics (physics of plasma) but they are more scarce in material science when the structure and properties of molecules or crystals are concerned. For this reason, in that follows we are focusing on the solution of the time-independent Schrödinger equation which leads to the solution of its associated multidimensional eigenproblem.

Before considering in the next chapter the two approximated techniques (HF and DFT) we consider here the most natural and accurate one, the one consisting in the direct solution of the Schrödinger equation. In the time-independent case, usually only the so-called ground distribution (the one related to the minimum energy) is searched. In any case, the high dimensional spaces in which the Schrödinger equation is defined leads to the curse of dimensionality. Nowadays it is widely accepted that classical discretization methods, based on the use of a grid (or a mesh), fail when the space dimension approaches few tens. Some attempts to address multidimensional equations have been proposed: the simplest choice consists in using stochastic techniques (described in Chapter 3.2) because in this case the computation complexity does not scale exponentially with the space dimension. However, these simulations are expensive and introduce statistical noise. Moreover, only the moments of the distribution can be computed accurately, because an accurate description of the distribution itself requires too many trajectories of the stochastic process. A second possibility consists in using sparse grid techniques but nowadays the use of these techniques only allows solving models in spaces rarely exceeding few tens. A third possibility consists in employing separated representations (at the origin of the so-called Proper Generalized Decomposition) and that allowed an accurate and fast solution of models involving many particles [6]. Even if the use of separated representations in computational mechanics is quite unusual, its use has been extensively considered in the context of quantum chemistry, in particular it is at the basis of the Hartree-Fock method described in the next chapter. Within the context of separated representations the main issue when addressing the solution of the Schrödinger equation is related to the antisymmetry of the wavefunction enforced by using the Slater determinant whose number of terms explodes with the space dimension (number of particles involved in the quantum system).

In order to illustrate the limits of classical discretization techniques for addressing the direct solution of the Schrödinger equation we consider two quotations extracted from [35]:

> *"For the solution of the time-dependent problems, different levels of approximation have been used, which range between the direct discretization of the TDSE – time dependent Schrödinger equation – as the most precise but computationally most expensive choice, and time-dependent density functional theory (TDDFT), which is appealing from a computational point of view but turned out to be too crude an approximation to capture important features of the problem, ..."*

> *"The present state-of-the-art of numerically solving the time-dependent Schrödinger equation directly for realistic laser pulses is limited to two electron systems. The most successful calculations involve the largest massively parallel computers available. It is clear that the direct solution of the linear time-dependent Schrödinger equation has reached its computational limits"*

or the analysis addressed in [38]:

> *' ... the theory of Everything is not even remotely a theory of every thing. We know this equation [the Schrödinger equation] is correct because it has been solved accurately for small number of particles -isolated atoms and small molecules - and found to agree in minute detail with experiments. However it cannot be solved accurately when the number of particles exceeds about 10. No computer existing, or that will ever exist, can break this barrier because it is a catastrophe of dimension. If the amount of computer memory required to represent the quantum wavefunction of one particle is N then the amount required to represent the wavefunction of k particles is N^k. It is possible to perform approximate calculations ... but the schemes for approximating are not first-principles deductions but rather art keyed to experiment ... "*

One could think that this pessimistic perspective is based on the fact that the complexity scaling with N^k is due to the fact of using a grid instead of using separated representations. In some cases, the use of a separated representation circumvents the curse of dimensionality, leading to a complexity scaling with $N \cdot k$ instead of N^k, however some quantum systems explore the whole configurational space whose dimension corresponds to N^k (for example the spin glass problem considered in [10] consisting of a set of k quantum spins interacting by random Heisenberg exchanges where the configuration space dimension becomes 2^k and obviously the system can explore the whole configuration space). In these cases no solution exists! Thus, we must assume the existence of numerically tractable and intractable quantum systems. In what follows we are considering only the first ones and addressing their solutions by using approximated techniques.