Electronic structure of matter

(a document to the course; there is the table of contents at the end)

Martin Konôpka

Department of Physics, INPE, FEIIT SUT in Bratislava, `martin.konopka@stuba.sk`

last update: 28th December 2022

very few small pieces of the text around the end, especially in the appendices, are still in the Slovak language

Almost all commonly perceived manifestations of matter are somehow related to electrons in atoms, molecules or crystals:

- **We see objects around us**. For this, light is needed. A photon (a quantum of energy of light) is emitted from an atom (or from a moelcule or cystal). It carries some energy. This energy had to be released somehow by the atom. In the early 20^{th} century, physisists discovered that an atom is made up of a small and heavy positively charged nucleus and light and negatively charged electrons moving around the nucleus. Electrons moving in the field of a nucleus have some energy. This energy can be decreased by emiting a photon. So **how we visually perceive or measure objects with optical instruments is related to the electronic structure of atoms, molecules and solids**.
- **Mechanical properties of matter** (elasticity, hardness, viscosity, etc). They may, of course, depend also on atomic masses. Above all, however, they are influenced by the nature of the interactions and forces acting between atoms and molecules. These interactions are related to the distribution of electrons in matter. The nuclei of atoms usually enter the description of these properties effectively only as point (classical) particles with certain masses, charges and possibly also with internal angular momenta (spins of the nuclei).
- **Chemical reactions**. They are a unique example in that very rich manifestations of electronic structure of atoms, ions, molecules as well as radicals can be observed. Some substances react together, others do not. It depends on their ability to form a chemical bond, and this ability depends on the specific properties of the electron shells of the atoms and molecules in the substance.
- **Electrical conductivity**. Metals are typical conductors. Their electrical conductivity is caused by the presence of conduction electrons, which can move under the influence of the loaded electric field (i.e voltage). Whether a material contains conductive electrons or other charge carriers (e.g., holes) also depends on the type of atoms (and the number of electrons in them), i.e., what electronic structure they form.

Since the physics of the world of electrons, atoms, molecules and crystals is quantum physics, this lecture will make extensive use of it.

1 Reminder of Basis Postulates of Quantum Mechanics

Although some of you have already completed the course of Quantum Mechanics (QM), it will first be useful to recall in some condensed form and summarize what we will need from it. In this introductory lecture, we will briefly go through the postulates of QM. They are postulates of wave QM, i.e. QM written in the formalism of wave functions, and for the sake of simplicity only for a one-particle system. We will approach by exposing selected and condensed topics mostly according to [1].

1.1 The First Postulate of Quantum Mechanics

Electrons in matter, as well as other objects of the microworld with relatively small energies, are generally not well described by classical concepts of position and momentum. Instead, we use wavefunctions and a certain statistical apparatus.

The probability density of finding a particle at certain point $\vec{r}$ is defined by formula

$$\rho(\vec{r}, t) = \frac{\mathrm{d}P}{\mathrm{d}V}$$

And here is the postulate:

To every state of a particle, a complex function $\psi(\vec{r}, t)$ is assigned, such that it perfectly characterises the state (completely describes it). Square of the absolute value of the function is equal to the probability density of finding the particle at the point $\vec{r}$ at time t:

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

So we have

$$\mathrm{d}P(\vec{r}, t) = \rho(\vec{r}, t)\,\mathrm{d}V$$

Implication 1: For a finite spatial domain Ω, the following formula applies:

$$P(\Omega, t) = \int_\Omega \rho(\vec{r}, t)\,\mathrm{d}^3r$$

Implication 2: Because the particle must be somewhere, we have

$$\int \psi^*(\vec{r}, t)\psi(\vec{r}, t)\,\mathrm{d}^3r = 1$$

From this we obtain two conditions for ψ:

(1) In order for this intergal to exist, it must hold that

$$\lim_{|\vec{r}| \to \infty} \psi = 0$$

(2) ψ has to be such that the whole integral does not depend on time (although $\psi^*\psi$ may depend on time).

1.2 The Second Postulate of Quantum Mechanics

This postulate concerns the determination of the mean values of physical quantities. Before we express it, we must introduce certain concepts and learn certain knowledge about operators in QM.

1.2.1 Mean Values of the Coordinates. Pure Quantum Ensemble

Measured quantities in quantum-mechanical systems usually acquire random values. Therefore, it makes sense to ask about their average (mean) values. In this introduction, let us consider an example from elementary statistics [1]: we have a bag with coins of different values:
N_1 coins with a value h_1, N_2 coins with a value h_2, ..., N_n coins with a value h_n.
The total number of coins is

$$N_1 + N_2 + \cdots + N_n = N$$

The total value of coins is

$$N_1h_1 + N_2h_2 + \cdots + N_nh_n = h$$

The average value of one coin is

$$\bar{h} \;=\; \frac{h}{N} \;=\; \frac{N_1}{N}h_1 + \frac{N_2}{N}h_2 + \cdots + \frac{N_n}{N}h_n$$

Suppose that all the coins have same size and weight. We shake the bag well and take out one randomly selected coin. What is the likehood that it has the value h_1? Obviously,

$$P_1 = \frac{N_1}{N}\ ; \quad \text{a obdobne ďalšie:} \quad P_2 = \frac{N_2}{N}\,, \ldots,\ P_n = \frac{N_n}{N}$$

Thus, the average value of one coin can be expressed by formula

$$\bar{h} = \sum_{i=1}^{n} P_i h_i \tag{1}$$

Let us now focus our attention on the particle moving, for simplicity, along a segment of a straight line. Imagine that we know the probability density $\rho(x)$ of finding the particle at any point x. [It need not be a quantum-mechanical (QM) particle; we just need such a one motion of which has a random nature.] What is the mean (average) coordinate of such a particle? Countrary to the example with coins, here we have a continguous set of possible values, that is also and infinite set. Anyway, we find the following result:

$$\bar{x} = \int_a^b \mathrm{d}x\, \rho(x)\, x$$

And now let us go to QM. We know that the probability density of finding a QM particle is $\rho(x,t) = \psi^*(x,t)\psi(x,t)$. Therefore

$$\bar{x}(t) = \int_a^b \psi^*(x,t)\, x\, \psi(x,t)\, \mathrm{d}x$$

It might have been written more compactly but we will see that is has a sense to expand it as we have done. Compared to the previous examples, we have also the time here for ρ may vary with time. But the time stands here as a trivial parameter only. (There is no integration over the time and the same time is on both sides of the equality.)

What actually do we mean by a mean value (e.g. of a coordinate) in QM? For instance, we would like to measure the average position of an electron in a hydrogen atom being in a certain quantum state (characterised by a certain wavefunction). Or, what might be easier done, the average position of an electron in a potential well, in the simplest case one-dimensional.[1] In general, however, it is a three-dimensional problem and in such a case we, therefore, would aim to determine not only $\bar{x}$, but also $\bar{y}$ and $\bar{z}$. We would basically proceed as follows: We would prepare a large number of hydrogen atoms in an identical manner, so that all of them would be described by the same wavefunction $\psi(\vec{r},t)$. Such an ensemble (a set of systems with the same wavefunction) is called a ***pure quantum ensemble***. We would placed a tiny measuring device at each of the atoms such that it would be able to record the position of the electron on demand at time t. Using the measures position vectors from the individual atoms, we woudl then calculated their average value. We would declare this to be the quantum-mechanical ***expectation value*** (a mean value).

Knowing what to understand by the expectation value of a position vector in QM, let us express it with the aid of the wavefunction. When considering the three dimensions, we have to write, for instance,

$$\mathrm{d}P = \rho(\vec{r},t)\, \mathrm{d}x\, \mathrm{d}y\, \mathrm{d}z = \rho(\vec{r},t)\, \mathrm{d}^3 r$$

and other notations are also being used for the volume element $\mathrm{d}^3 r$. Otherwise, however, the formula will have the same structure:

$$\boxed{\bar{x}(t) = \int \psi^*(\vec{r},t)\, x\, \psi(\vec{r},t)\, \mathrm{d}^3 r} \tag{2}$$

and the integration goes over the entire (infinite) space; in such a case (multidimensional integrals) we usually write a single integration symbol only and we omit writing the integration bounds. Of course, formulae for the y and z coordinates would be written quite analogously.

---

[1] In a real experiment, it would mean that one of the dimensions – the length – of such a structure would be much larger than the remaining two dimensions. Such a groove might technologically be formed on a surface of a solid material or at an interface of surfaces. Even a two-dimensional structure, i.e. a two-dimensional potential well, could be formed on a surface.

1.2.2 Mean Value of the x-component of the Momentum

We will not derive this; we will only recall what you should already know:

$$\boxed{\bar{p}_x(t) = \int \psi^*(\vec{r},t)\, \hat{p}_x\, \psi(\vec{r},t)\, \mathrm{d}^3 r} \tag{3}$$

where

$$\hat{p}_x = -\mathrm{i}\hbar \frac{\partial}{\partial x} \tag{4}$$

and we would analogously write it for $\hat{p}_y$ a $\hat{p}_z$.

1.2.3 Position Operator, Momentum Operator and Other Operators

We can now indicate what the 2nd postulate will be about: in QM, every physical quantity F has a corresponding operator, which we denote by the symbol $\hat{F}$, such that

$$\bar{F}(t) = \int \psi^*(\vec{r},t)\, \hat{F}\, \psi(\vec{r},t)\, \mathrm{d}^3 r$$

We will make this postulate more accurate later. To be able to do so, we first have to examine propeties of operators being used in QM. We now define several further QM oparators on the basis of the ***correspondence with classical mechanics***.

$$\hat{\vec{p}} = \vec{e}_x\hat{p}_x + \vec{e}_y\hat{p}_y + \vec{e}_z\hat{p}_z = -\mathrm{i}\hbar\left(\vec{e}_x\frac{\partial}{\partial x} + \vec{e}_y\frac{\partial}{\partial y} + \vec{e}_z\frac{\partial}{\partial z}\right) \tag{5}$$

Therefore

$$\boxed{\hat{\vec{p}} = -\mathrm{i}\hbar\vec{\nabla}} \tag{6}$$

The ***angular momentum operator***:

$$\boxed{\hat{\vec{L}} \;=\; \hat{\vec{r}} \times \hat{\vec{p}} \;=\; -\mathrm{i}\hbar\, \hat{\vec{r}} \times \vec{\nabla}} \tag{7}$$

The ***kinetic energy operator***:

$$\boxed{\hat{T} = \frac{\hat{\vec{p}}\,.\,\hat{\vec{p}}}{2m}} \;=\; -\frac{\hbar^2}{2m}\vec{\nabla}\,.\,\vec{\nabla} \;\equiv\; -\frac{\hbar^2}{2m}\Delta \tag{8}$$

The ***operator of the potential energy*** of a point charge q in an external electric field with the intensity $\vec{E}(\vec{r}) = -\operatorname{grad} U(\vec{r})$:

$$\boxed{\hat{V}(\vec{r}) = q\, U(\vec{r})} \tag{9}$$

It is a simple operator in the sense that it is expressed by a usual number; e.g., there is no derivative int its expression. For instance, assuming a coulombic field generated by a (fixed in space) charge Q, we have $U(\vec{r}) = Q/(4\pi\varepsilon_0 r)$.

1.2.4 Properties of the Operators Used in Quantum Mechanics [1, 2, 3, 4]

Definition 1: *Let D_1 and D_2 be two sets of functions (not necessarily different). By the operator $\hat{A}$ is called a prescription (rule) that assigns a function $g \in D_2$ to each function $f \in D_1$, which we write symbolically $g = \hat{A}f$.*

Definition 2: *An operator $\hat{A}$ defined on a set D is called linear if*

$$\hat{A}(c_1 f_1 + c_2 f_2) = c_1\hat{A}f_1 + c_2\hat{A}f_2\ , \quad \forall f_1, f_2 \in D\ , \quad \forall c_1, c_2 \in \mathbb{C} \tag{10}$$

Definition 3: *Given is an $\hat{A}$. If there exists such an operator $\hat{A}^\dagger$ that the equality*

$$\int f^* \hat{A}\, f\, \mathrm{d}\tau = \int (\hat{A}^\dagger f)^* f\, \mathrm{d}\tau \quad \text{pre } \forall f \in D \tag{11}$$

holds, then the operator $\hat{A}^\dagger$ is called Hermitian-conjugate to the operator $\hat{A}$.

Definition 4: *If $\hat{A}^\dagger = \hat{A}$, then we say that $\hat{A}$ is a hermitian operator.*[2]

Example 1: The operator of multiplication by a complex constant: $\hat{A} = c$.

$$\int_{-\infty}^{\infty} f^* c\, f \,\mathrm{d}x = \int_{-\infty}^{\infty} (c^* f)^* f \,\mathrm{d}x$$

From this we see that $c^\dagger = c^* \neq c$. It means that multiplication by a complex constant is not a hermitian operator in general. It might be only in the case of zero imaginary part of c.

Example 2: The operator of the derivative with respect to a coordinate: $\hat{A} = \dfrac{\partial}{\partial x}$. We assume that it operates on a set of functions such that the integrals written bellow exist and the function f goes to zero for $x \to \pm\infty$. Subsequently, with the aid of integration by parts, we prove that

$$\int_{-\infty}^{\infty} f^* \frac{\partial}{\partial x} f \,\mathrm{d}x = \int_{-\infty}^{\infty} \left(-\frac{\partial f}{\partial x}\right)^* f \,\mathrm{d}x$$

Therefore

$$\left(\frac{\partial}{\partial x}\right)^\dagger = -\frac{\partial}{\partial x}$$

meaning that the derivative operator is not a hermitian one too.

Example 3: Operator $\hat{x} = x$ (i.e., multiplication by a coordinate, which is a real quantity). Using the precedure as in Example 1 we see that at last we have a hermitian operator:

$$\hat{x}^\dagger = \hat{x}$$

Example 4: Operator $\hat{p}_x = -\mathrm{i}\hbar\dfrac{\partial}{\partial x}$. Using a procedure similar as in Example 2 we find out that

$$\hat{p}_x^\dagger = \hat{p}_x$$

that is, the operator of the x-component of the momentum is hermitian (although it contains the derivative; it, however, includes also i). We will see that this is not an accidence and it has a deeper and wider meaning.

Theorem 1: *A physical quantity F has in each (quantum) state a real mean value if and only if it is evaluated as an expectation value from some hermitian operator (which we denote as $\hat{F}$).*

Remark: Even if in this theorem we mention a "physical quantity" and "quantum state", it is a purely mathematical theorem (with its consequences for physics).

Proof:

(A) Let $\bar{F} = (\bar{F})^*$. Using this assumption, we should prove that $\hat{F} = \hat{F}^\dagger$. Of course, we have to suppose that the operator $\hat{F}^\dagger$ exists; otherwise there would be nothing to prove. We have

$$\bar{F} = \int \psi^* \hat{F} \psi \,\mathrm{d}\tau = \int (\hat{F}^\dagger \psi)^* \psi \,\mathrm{d}\tau$$

as well as

$$\bar{F}^* = \left(\int \psi^* \hat{F} \psi \,\mathrm{d}\tau\right)^* = \int (\hat{F}\psi)^* \psi \,\mathrm{d}\tau$$

Because $\bar{F} = (\bar{F})^*$, the equality

$$\int (\hat{F}^\dagger \psi)^* \psi \,\mathrm{d}\tau = \int (\hat{F}\psi)^* \psi \,\mathrm{d}\tau$$

---

[2] In English-written literature, both *hermitian* and *hermitean* adjectives can be found. The first of them is much more frequent but in the book [3] the second variant is used.

must hold for any function ψ from the set under consideration. The last equality can only be fulfilled if $\hat{F} = \hat{F}^\dagger$ what was to be proved.

(B) It now remains to conduct the proof in the reverse direction: startin from the assumption $\hat{F} = \hat{F}^\dagger$, we have to prove that $\bar{F} = (\bar{F})^*$. It is similarly simple and you can try it as a homework.

Since any physical quantity can acquire real values only, the theorem just proven implies that the respective operator in QM has to be a hermitian operator.

Theorem 2: *Let there be a Hermitian-conjugate operator $\hat{A}^\dagger$ to the operator $\hat{A}$. Let A be a linear operator. Then for all $f_1, f_2 \in D$ the following equality holds:*

$$\int f_1^* \hat{A} f_2 \, \mathrm{d}\tau = \int (\hat{A}^\dagger f_1)^* f_2 \, \mathrm{d}\tau \tag{12}$$

This statement also applies in the opposite direction.

Proof: It can be looked at in books [2, 3, 4] `(but you do not need to know it on the exam).`
The equation in the last theorem can be written more transparently

$$\left(\int f_1^* \hat{A} f_2 \, \mathrm{d}\tau \right)^* = \int f_2^* \hat{A}^\dagger f_1 \, \mathrm{d}\tau \tag{13}$$

We will go through further definitions and theorems briefly; more detailed wording and the proof can be found in the literature [2, 3, 4].

Definition 5: (Product of operators.) By the product of two operators we understand the operator $\hat{C} = \hat{B}\hat{A}$ such that

$$\hat{C} f = \hat{B}(\hat{A} f)$$

Theorem 3: *If $\hat{C} = \hat{B}\hat{A}$ and if hermitian-conjugate operators to $\hat{A}$ and $\hat{B}$ exist, then*

$$\hat{C}^\dagger = \hat{A}^\dagger \hat{B}^\dagger$$

Proof: We have done it as an exercise using Theorem 2.

Definition 6: *The expression*

$$[\hat{A}, \hat{B}] \equiv \hat{A}\hat{B} - \hat{B}\hat{A}$$

is called the commutator of the operators $\hat{A}$, $\hat{B}$.
In the subject *Quantum Mechanics*, you have learnt that, for instance,

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

1.2.5 Eigenfunctions and Eigenvalues of Operators

Assume we have an operator $\hat{A}$ (it need not be hermitian). If its action on a function f generates the same function, or at most multiplied by a constant, such a function is called an ***eigenfunction*** of the operator $\hat{A}$:

$$\hat{A} f_a = a f_a$$

a is the respective ***eigenvalue*** of the operator; we attached the index a to the function to highlight its association with the eigenvalue a.

The set of all eigenvalues of an operator is called the ***spectrum*** of the operator. An operator can have many eigenfunctions and eigenvalues. If the set of the eigenvalues is countable, the operator a *discrete* spectrum. If it is not countable, the operator has a *continuous* spectrum. For instance, the equation

$$\hat{p}_x \exp\left(\frac{\mathrm{i}}{\hbar}ax\right) = a\,\exp\left(\frac{\mathrm{i}}{\hbar}ax\right)$$

hold for the operator $\hat{p}_x$ for any real a. Therefore, the eigenvalues of the operator $\hat{p}_x$ are all real numbers, i.e. it has a *continuous* spectrum.

The operator of the x-coordinate also has a continuous spectrum:

$$\hat{x}f = xf$$

On the other hand, many operators have discrete spectra, for example the operator $\hat{L}_z$ and others. And there are also operators that have mixed spectra: a part is discrete, other part of is continuous. This is the case of the hydrogen atom Hamiltonian and also of many other very important operators.

To keep the notation and mathematical apparatus simple, we will formulate some parts of the quantum theory apparatus only for operators having discrete spectra. In cases of continuos and mixed spectra, procedures would in some cases be similar; it is sufficient to imagine the associated "index" acquires values from a continuous set or from a set that is continuous by parts. Continuous spectra, however, bring also non-trivial complications, for example impossibility to normalise eigenfunctions; for instance, an attempt to determine a norm of an eigenfunction of the momentum operator will fail:

$$\int_{-\infty}^{\infty} \left|\exp\left(\frac{\mathrm{i}}{\hbar}ax\right)\right|^2 \,\mathrm{d}x \to \infty$$

that is, the integral does not exist.

Theorem 4: *Eigevalues of a hermitian operator that correspond to normalisable eigenfunctions are real numbers [2].*

Proof: Directly follows from Theorem 1. It can also be found in the cited books.

Often there are several eigenfunctions for one eigenvalue:

$$\hat{A}f_{j,\alpha} = \Lambda_j f_{j,\alpha}\,, \quad \alpha \in \{1, 2, \ldots, g\}$$

Then we say that the operator has a *degenerate* spectrum or that the eigenvalue is degenerate (if $g = 2$, then doubly, if $g = 3$, then triply, etc).

Theorem 5: *Eigenfunctions corresponding to different eigenvalues of a hermitian operator are mutually orthogonal [2, 3, 4].*

Remark 1: That means that $\int f_m^* f_n \,\mathrm{d}\tau \propto \delta_{mn}$ hold. If the eigenfunctions are normalisable, i.e. if the corresponding integrations converge, then we usually normalise the eigenfunctions so that $\int f_m^* f_n \,\mathrm{d}\tau = \delta_{mn}$.
Remark 2: The theorem does not imply orthogonality of linearly-independent eigenfunctions $f_{n,j}$ corresponding to the same eigenvalue (i.e. to a degenerate one).
Proof: Can be found in the cited books.

1.2.6 Formulation of the Second Postulate of Quantum Mechanics

In quantum mechanics, every physical quantity F has a linear hermitian operator $\hat{F}$ associated to it such that the mean value (the expectation value) of F (at time t) in a state described by a wavefunction $\psi(\vec{r}, t)$ is given by

$$\bar{F}(t) = \int \psi^*(\vec{r}, t)\hat{F}\,\psi(\vec{r}, t)\,\mathrm{d}^3 r \tag{14}$$

The physical quantity F can only take values that are eigenvalues of the operator $\hat{F}$. The relations

$$\hat{x} = x\,, \quad \hat{p}_x = -\mathrm{i}\hbar\,\frac{\partial}{\partial x}$$

are assumed (postulated) as well as analogous ones for y and z components. Operators of remaining physical quantities, that have classical analogues, are determined according to the expressions in classical physics (here mechanics) using the components of coordinates and momenta.

Remark 1: The property that F can only acquire the values that are eigenvalues of the respective operator, must be understood in the sense that each individual measurement of the quantity on a particular quantum-mechanical system can only give some of the eigenvalues of $\hat{F}$ as a result. The average value can, of course, be different. It goes in the same way as in the case of the bag with the coins (sec.1.2.1): Each time we reach into a bag for a random one coin, we will necessarily select only the monetary value that is minted on the coins, e.g. 50 cents (which in our comparison corresponds to one of the eigenvalues). But when we make an average of many pulls, the average financial value can be e.g. 73 cents, a value that no single coin has. Additional related supporting argumentation for the 2^{nd} postulate is provided in Appendix A.1. The argumentation can be found in the literature as well, for instance in [2].

Remark 2: In a case of a many-particle system, there would be integrations over coordinates of the additional particles of the system in (14).

1.3 The Third Postulate of Quantum Mechanics

If a state described by a wave function $\psi_1(\vec{r},t)$ exists as well as a state described by $\psi_2(\vec{r},t)$, then the state

$$\psi(\vec{r},t) = c_1\psi_1(\vec{r},t) + c_2\psi_2(\vec{r},t)\,, \qquad c_1, c_2 \in \mathbb{C} \tag{15}$$

is also possible in priciple. This postulate is called the **superposition principle**. Whether this or that superposition state is experimentally achievable, however, is a different matter. The postulate should be understood that in a theory we can work with any such superposition states.

1.4 The Fourth Postulate of Quantum Mechanics

According to the 2^{nd} postulate of QM, we know how to construct the operator of the energy of a particle in an external field. We call it Hamilton's operator, in short Hamiltonian. If the particle is in an external field with the potential energy $V(\vec{r},t)$, then its Hamiltonian is

$$\hat{H} = -\frac{\hbar^2}{2m}\Delta + V(\vec{r},t) \tag{16}$$

Thus, the potential energy may be time-dependent. A typical situation of this kind is an atom placed in the field of an electrmagnetic (EM) wave. As some other example, we have had the time-independent potential energy $V(\vec{r}) = qU(\vec{r})$, pozri (9). The fourth postulate of QM says:

The equation of motion for the wave function of a state is the Schrödinger equation.

$$\boxed{\mathrm{i}\hbar\frac{\partial}{\partial t}\psi(\vec{r},t) = \hat{H}\psi(\vec{r},t)} \tag{17}$$

where $\hat{H}$ is the Hamiltonian of the given quantum-mechanical system.

It means that if we know a state wave function $\psi(\vec{r},t_0)$ (at certain time t_0), then, by solving the Schrödinger equation (SchE), we are able to determine the state wave function $\psi(\vec{r},t)$ at any later time t.

It is easy to convince that the SchE is consistent with the superpositon principle: If wave functions $\psi_1(\vec{r},t)$ and $\psi_2(\vec{r},t)$ are solutions to the SchE, then also any linear combination of them is a solution [see also (15)].

2 Stationary states (a brief reminder)

Consider a physical system with a time-independent Hamiltonian $\hat{H}$. Let $u_n(\vec{r})$ be some of its eigenfunctions and E_n its respective eigenvalue, that is the eigenenergy:

$$\hat{H}u_n(\vec{r}) = E_n u_n(\vec{r}) \tag{18}$$

Suppose we prepare the QM system at time $t = 0$ in this very state:

$$\psi(\vec{r}, 0) = u_n(\vec{r})$$

How the state of the system will evolve in time? By substituting into the time-dependent SchE (17), we can easily make sure that the function

$$\psi(\vec{r}, t) = u_n(\vec{r}) \exp\left(-\frac{\mathrm{i}}{\hbar}E_n t\right) \tag{19}$$

is solution of the equation. It means that up to a periodically oscillating phase factor $\exp\left(-\frac{\mathrm{i}}{\hbar}E_n t\right)$ (the absolute value of which does not change) is the state always the same: $u_n(\vec{r})$. Therefore, eigenstates of a Hamiltonian of given physical system are called *statioanry states.* Their time evolution just described is called *free* time evolution.

Equation (18) is often called the *time-independent Schrödinger equation* or the *stationary* SchE.

3 (Non)-commuting operators and the uncertainty relation

We have learnt that in QM, operators are assigned to classical quantities such as the position vector or momentum, (the correspondence principle). Numerical values of the quantities in QM are obtained using averaging, for instance

$$\bar{x} = \int_{-\infty}^{\infty} \psi^*(x)\,\hat{x}\,\psi(x)\,\mathrm{d}x\ , \qquad \bar{p}_x = \int_{-\infty}^{\infty} \psi^*(x)\,\hat{p}_x\,\psi(x)\,\mathrm{d}x \tag{20}$$

While the numbers commute among themselves when multiplied, it may not be the case with operators:

$$\hat{x}\hat{p}_x - \hat{p}_x\hat{x} = \mathrm{i}\hbar$$

We know the notion of a *commutator*: e.g. $[\hat{x}, \hat{p}_x] = \mathrm{i}\hbar$. We have also learnt that eigenfunctions and eigenvalues of the operators (those that correspond to physical quantities) are important in QM because experimental measurements of some quantity can only yield one of the eigenvalues of the corrsponding operator as a result. For example, we have the operator and equation

$$\hat{H}\psi_n(x) = E_n\psi_n(x)$$

for eigenenergies of some given system. In the following subsection, we will discuss commuting operators. In another subsection, we will briefly discuss non-commuting operators.

3.1 Common Eigenfunctions of Commuting Operators

The goal of this section is to prove (at least partially) a very important statement saying that commuting operators have common eigenfunctions (not eigenvalues). We will use this property, for example, in our search for the eigenfunctions of the hydrogen-atom and like ions Hamilton operator. We will demonstrate the proof of the statement at least for the case of the simple situation of non-degenarate spectra of the operators. at least

The Case of Non-degenerate Spectra

Theorem 6: *Assume that operators $\hat{A}$ and $\hat{B}$ commute and both have non-degenerate spectra. Then any eigenfunction of A is also an eigenfunction to (of) B and vice versa (i.e. any eigenfunction of B is also an eigenfunction to A).*

Proof: Let $Af = af$. Act with the operator B from the left side:

$$B(Af) = B(af)\,;\ \text{komutujú}\ \ \Rightarrow\ \ A(Bf) = a(Bf)$$

It means that the function

$$Bf \equiv f' \tag{21}$$

is also an eigenfunctions of the operator A,

$$Af' = af'$$

and it even corresponds to the same eigenvalue as the function f. Because the operator A has, according to the assumption, a non-degenerate spectrum , f' can differ from f at most trivially, i.e. by an unessential constant multiple:

$$f' = cf \tag{22}$$

Thus, f and f' are essentally the same eigenfunctions of the operator A. By combining equations (21) and (22), we obtain

$$Bf = cf$$

meaning that the function f is an eigenfunction of the operator B as well, what was to be proved.

Remark: Thus, commuting operators have common *eigenfunctions*, not eigenvalues.

The Case of Degenerate Spectra

We should now prove an analogue of the above theorem for the general case, i.e. the case when $\hat{A}$ and $\hat{B}$ may have degenerate spectra.

Theorem 7: *If operators $\hat{A}$ and $\hat{B}$ commute, then it is possible to construct a complete set of their eigenfunctions such that they are common to both these operators. There is also the statement valid in the reverse direction: It operators $\hat{A}$ and $\hat{B}$ have common eigenfunctions and if these eigenfunctions form a complete system, then the operators $\hat{A}$ and $\hat{B}$ commute.* [2, 3].

The proof is not difficult to understand. It is only more lenghty and it uses some knowledge from algebra. Essentially, it is done by constructings the common eigenfunctions. We will not do it but we will keep in mind the content of the theorem.

3.2 The Uncertainty Relation

We just recall what you should know from the course of *Quantum Mechanics*.

Let F and G be physical quantities with operators $\hat{F}$ and $\hat{G}$. In case that the commutator of $\hat{F}$ and $\hat{G}$ is non-zero, the two quantities are incompatible, i.e. we cannot measure them at the same time with arbitrary precisions.

Briefly, if $[\hat{F}, \hat{G}] \neq 0$ then F and G cannot both be determined accurately. There will be some uncertainty in F or in G or (most likely) in both.

For example,

$$\Delta x\, \Delta p_x \geq \frac{\hbar}{2} \tag{23}$$

This is a principal equality which can be derived with the help of considerations about wave packets, or using a more formal and more general approach [2, 3]. Therefore, the uncertainty relation (23) cannot be understood in terms of some imperfection of experimental apparatus, but as a fundamental property of particles. This property cannot, of course, be noticed on some relatively heavy particles of matter, such as e.g. a grain of sand, but for very light particles of the microworld, such as e.g. electron, this uncertainty is significant.

If two operators switch, such as $\hat{p}_x$ and $\hat{p}_y$, commute, then in principle it is possible to measure the corresponding quantities with arbitrary precision, i.e. with zero uncertainties (if we disregard the imperfections of measuring instruments). Also formally, the inequality

$$\Delta p_x \, \Delta p_y \geq 0 \tag{24}$$

can be derived for such (commuting) operators [2, 3]. Obviosly, it does not put any bound on the (always non-negative, of course) uncertainties Δp_x, Δp_y.

4 Angular Momentum in Quantum Mechanics

Angular momentum (AM) is an important quantity already in classical mechanics. It is such because it belongs to the integrals of motion, which are the quantities not changing their value (while the system is a closed, i.e. an isolated one). Integrals, i.e. constans of motion (in classical mechanics) are the total mechanical energy of a system, its total momentum and the total angular momentum. As we will see, in QM the AM is even a more important quantity, so it is worth dealing with it. Knowledge about the AM will be a springboard to the study of motion of a particle in a spherically symmetric field – hydrogen atom for example. In the next chapter we will wish to determine eigenfunctions of Hamiltonian of such a particle. For the sake of motivation, let us go a little ahead of the exposition and say that the Hamiltonian for a spherically symmetric field commutes with the operators describing the angular momentum. Therefore, if we find proper eigenfunctions for, e.g. the operator of the quadrat of the AM, then they might also be eigenfunctions for the particle in the spherically symmetri field.

4.1 Definitions and Basic Commutation Relations

Angular momentum of a single classical particle is

$$\vec{L} = \vec{r} \times \vec{p} \tag{25}$$

According to the correspondence principle, we then construct the corresponding quantum-mechanical operator

$$\hat{\vec{L}} = \hat{\vec{r}} \times \hat{\vec{p}} \tag{26}$$

Its cartesian components are (we will write the z one in detail)

$$\hat{L}_x = \hat{y}\hat{p}_z - \hat{z}\hat{p}_y \,, \qquad \hat{L}_y = \hat{z}\hat{p}_x - \hat{x}\hat{p}_z \,, \qquad \hat{L}_z = \hat{x}\hat{p}_y - \hat{y}\hat{p}_x = -\mathrm{i}\hbar \left(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x} \right) \tag{27}$$

They are all hermitian operators, which is easy to ascertain based on Theorem 3 and the relations like $[\hat{y}, \hat{p}_z] = 0$. As an exercise we show that $\hat{L}_x$, $\hat{L}_y$, $\hat{L}_z$ do not commute and that their commutators are

$$\boxed{[\hat{L}_x, \hat{L}_y] = \mathrm{i}\hbar\hat{L}_z \,, \qquad [\hat{L}_y, \hat{L}_z] = \mathrm{i}\hbar\hat{L}_x \,, \qquad [\hat{L}_z, \hat{L}_x] = \mathrm{i}\hbar\hat{L}_y} \tag{28}$$

A physical consequence, according to the uncertainty principle, then is impossibility to know values of two cartesian components of the AM precisely at once. A mathematical consequence according to Theorem 7 formulated above is impossibility to find a complete system of common eigenfunctions for a pair of operators, say for $\hat{L}_x$ and $\hat{L}_y$; such a system does not exist. Regarding the AM, its magnitude is, however, of interest. In classical mechanics we would calculate a value $L \equiv |\vec{L}|$. In QM it turn out to be more practicle to calculate the square of the magnite of the AM. Let us thus introduce the operator of the square of the angular momentum:

$$\hat{L}^2 = \hat{L}_x^2 + \hat{L}_y^2 + \hat{L}_z^2 \tag{29}$$

And let us explore its commutation properties. As an exercise we prove that

$$\boxed{[\hat{L}^2, \hat{L}_x] = [\hat{L}^2, \hat{L}_y] = [\hat{L}^2, \hat{L}_z] = 0} \tag{30}$$

For the pair of operators $\hat{L}^2, \hat{L}_x$ a complete system of common eigenfunctions thus exists. And so for the pairs $\hat{L}^2, \hat{L}_y$ a $\hat{L}^2, \hat{L}_z$. We will be seeking them; it will be sufficient to do for $\hat{L}^2, \hat{L}_z$. Later, they will prove very important in the search of stationary states (eigenstates) of the hydrogen atom and so on. We will also find the corresponding eigenvalues which will tell us what can be possible values of, for instance, $\bar{L}_z$ or $\bar{L}^2$.

4.2 Spherical Coordinates

We will see that it is often easier to work with the AM operators and corresponding eigenfunctions if they are expresses using spherical coordinates ϑ, φ, r. Let us recall the transformations formulas:

$$x = r \sin\vartheta \cos\varphi \tag{31a}$$
$$y = r \sin\vartheta \sin\varphi \tag{31b}$$
$$z = r \cos\vartheta \tag{31c}$$

$$\vartheta = \arccos\left(\frac{z}{\sqrt{x^2+y^2+z^2}}\right) \tag{32a}$$
$$\varphi = \operatorname{atan2}(y, x) \tag{32b}$$
$$r = \sqrt{x^2+y^2+z^2} \tag{32c}$$

4.3 Eigenfunctions and Eigenvalues of the $\hat{L}_z$ Operator

Let us take any differentiable function $f = f(\vec{r})$ and try to express its derivative with respect to φ:

$$\frac{\partial f}{\partial \varphi} = \frac{\partial f}{\partial x}\underbrace{\frac{\partial x}{\partial \varphi}}_{-y} + \frac{\partial f}{\partial y}\underbrace{\frac{\partial y}{\partial \varphi}}_{x} + \frac{\partial f}{\partial z}\underbrace{\frac{\partial z}{\partial \varphi}}_{0}$$

According to (27) we have

$$\boxed{\hat{L}_z = -\mathrm{i}\hbar\frac{\partial}{\partial \varphi}} \tag{33}$$

The equation for the eigenfunctions and eigenvalues (eigensystem) of the operator $\hat{L}_z$ is written as

$$\hat{L}_z \Phi_a(\varphi) = a \Phi_a(\varphi) \tag{34}$$

where according to Theorem 4 a has to be a real number. By substituting for $\hat{L}_z$, the differential equation (DE)

$$\frac{\mathrm{d}\Phi_a}{\mathrm{d}\varphi} + \frac{a}{\mathrm{i}\hbar}\Phi_a = 0 \tag{35}$$

is formed. The solution of this equation (found in the exercise) is

$$\Phi_a(\varphi) = c \exp\left(\frac{\mathrm{i}}{\hbar} a \varphi\right) \tag{36}$$

where c is an arbitrary complex constant. We can easily be convinced of this solution by substitution. Given the geometric meaning of the φ angle, we will require the condition

$$\Phi_a(\varphi + 2\pi) = \Phi_a(\varphi) \tag{37}$$

By using it we get

$$\exp\left(\frac{\mathrm{i}}{\hbar} a 2\pi\right) = 1$$

From this $a = m\hbar$, where $m \in \mathbb{Z}$. So we see that the z component of the AM can only take on values that are integer multiples of $\hbar$. The same is of course true also for the x and y components. (These axes are mutually equivalent to the z one by their nature; it is sufficient to rename the axes and we get x from z, for example.) As it is seen as well, the eigenvalues of the $\hat{L}_x$, $\hat{L}_y$, and $\hat{L}_z$ operators are non-degenerate.

We usually choose the value of the constant c by convention so that the normalisation

$$\int_0^{2\pi} \Phi^* \Phi \, \mathrm{d}\varphi = 1 \tag{38}$$

is satisfied, which gives (if we want c to be a real and positive constant) $c = 1/\sqrt{2\pi}$. Let us summarise:

$$\boxed{\hat{L}_z \Phi_m(\varphi) = m\hbar\, \Phi_m(\varphi)}\,, \quad \text{where} \quad \boxed{\Phi_m(\varphi) = \frac{1}{\sqrt{2\pi}}\, e^{\mathrm{i}m\varphi}}\,, \quad m \in \mathbb{Z} \tag{39}$$

The eigenvalues of the operator $\hat{L}_z$ are thus non-degenerate. We can also easily see that the corresponding functions are – as it should be also according to Theorem 5 – orthogonal to each other (verified in the exercise):

$$\int_0^{2\pi} \Phi_m^*(\varphi)\, \Phi_{m'}(\varphi)\, \mathrm{d}\varphi = \delta_{mm'} \tag{40}$$

Approaching the end of this section, it is suitable to say that if we already have found expression (39) for the eigenfunctions of $\hat{L}_z$, then it is easy, with the help of the Euler formula and the relations (31a), (31b), (31c) between sperical and cartesian coordnates, to reexpress the expression in cartesian coordinates (derivation done in the exercise):

$$\Phi_m(\varphi) \equiv f_m^{(z)}(x,y) = \frac{1}{\sqrt{2\pi}} \left(\frac{x + \mathrm{i}y}{x - \mathrm{i}y}\right)^{m/2} \tag{41}$$

The expressions for the eigenfunctions of $\hat{L}_x$ a $\hat{L}_y$ in terms of cartesian coordinates can now be immediatelly obtained by the cyclic permutation of the coordinates in (41). $\hat{L}_x$

There is no such symmetry or equivalence within the three spherical coordinates as it is within the cartesian ones. Therefore the eigenfunctions of the operators $\hat{L}_x$ and $\hat{L}_y$ in *spherical* coordinates will not have expressions of the same form as in the case of the operator $\hat{L}_z$ It would however be easy to find expressions for the operators $\hat{L}_x$ and $\hat{L}_y$ in spherical coordinates since we know, from the exercise, their expressions of the kind (41) in cartesian coordinates.

4.4 Common Eigenfunctions of Operators $\hat{L}_z$ a $\hat{L}^2$ (Part 1). Separation of Variables in Spherical Coordinates

Finding eigenfunctions is related to finding of eigenvalues and therefore it will be helpful for us to determine what can in QM be magnitude of the angular momentum (or its square). In this section, however, we will only suggest one procedure, a basic one, based on solving a partial differential equation, a bit "heavy-footed" compared to another method and at the same time less general.

We will use the commutation

$$[\hat{L}^2, \hat{L}_z] = 0 \tag{42}$$

from which, according to either Theorem 6 or Theorem 7 implies that the two operators have common eigenfunctions. So first let us see if the functions $\exp(\mathrm{i}m\varphi)$, which are the eigenfunctions of $\hat{L}_z$, happen to be eigenfunctions for $\hat{L}^2$ as well. The expression

$$\hat{L}^2 = -\hbar^2 \left[\frac{1}{\sin\vartheta}\frac{\partial}{\partial\vartheta}\left(\sin\vartheta \frac{\partial}{\partial\vartheta}\right) + \frac{1}{\sin^2\vartheta}\frac{\partial^2}{\partial\varphi^2}\right] = -\hbar^2 \nabla^2_{\vartheta,\varphi} \tag{43}$$

for the operator $\hat{L}^2$ in spherical coordinates (will be explained in exercises) will serve us in this. Let us note that there is no r in this expression. $\hat{L}^2$, however, depends also on the spherical angle ϑ and so its eigenfunctions will in general depend on this angle as well. Therfore the simple functions $\exp(\mathrm{i}m\varphi)$ (depending on φ only) will not be eigenfunctions of $\hat{L}^2$. But how to understand this knowing the fact that the operators $\hat{L}^2$ and $\hat{L}_z$ commute? The only possible explanations (with respect to Theorem 6) is that ***$\hat{L}^2$ has*** a ***degenerate spectrum*** (and we will convince ourselves in detail that it is so). But, on the other hand, we have proved that any eigenfunction of $\hat{L}_z$ *must* contain the factor $\exp(\mathrm{i}m\varphi)$. So the eigenfunctions of $\hat{L}_z$ such that they are *at the same time* also eigenfunctions of $\hat{L}^2$, must take the form

$$K(r,\vartheta)\, e^{\mathrm{i}m\varphi} \tag{44}$$

where $K(r,\vartheta)$ is a function of its two variables. Let us notice that the expression (44) is an eigenfunction of the operator $\hat{L}_z$ for arbitrary chosen function $K(r,\vartheta)$. We can easily be convinced of this by substituting (44) into (34).

And also note that the dependence on r in (44) is not necessary, because r is not found neither in the expression for $\hat{L}_z$ nor in $\hat{L}^2$. The angle ϑ must, however, in general be there, since it enters the operator $\hat{L}^2$. The equation

$$\hat{L}^2 Y_l^m(\vartheta,\varphi) = \lambda_{lm} Y_l^m(\vartheta,\varphi) \tag{45}$$

must therefore be solved, together with the already partially solved equation

$$\hat{L}_z Y_l^m(\vartheta,\varphi) = \hbar m Y_l^m(\vartheta,\varphi) \tag{46}$$

for the (at this point) unknown common eigenfunctions for which we however know, using formula (44), that it must be possible to express them in the separated form[3]

$$\boxed{Y_l^m(\vartheta,\varphi) = \Theta_l^m(\vartheta)\,\Phi_m(\varphi)} \tag{47}$$

To prevent a misinterpretation: m in Y_l^m and in Θ_l^m is not and exponent but a superscript. λ_{lm} are still completely unknown eigevalues; we only know that thay must be real. We understand the index m; it is associated with the variable φ. We do not specify the index l yet, but since we have the new variable ϑ, an index associated with it will apparently be needed. So let the l is the index and is has to be written at the function Θ. We have written also the index m to it; we do not know yet if it will be needed, i.e. if Θ will depend on m (it will). Even if it was not needed, we would not do anything wrong by writing it there.

As can be easily convinced, and we have already indicated it in between (44) and (45), the form (47) could be made more complicated by multiplying it by an arbitrary r-dependent function and such a form would still present a common eigenfunction of the operators $\hat{L}^2$ and $\hat{L}_z$ (because the variable r is just a constant for these operators; they contain neither derivatives with respect to r nor r itself). Such a multiplication by an arbitrary r-dependent function would be just an unnecessary complication for now; we are now trying to find the simplest possible common eigenfunctions of the operators $\hat{L}^2$ and $\hat{L}_z$. Therefore, have omitted the function $K(r,\vartheta)$, which depends also on the variable r, and replaced it by the function $\Theta_l^m(\vartheta)$, which depends on the only necessary spatial variable ϑ.

The solutions (eigenfunctions) of the form (47) are products of two functions, one of which only depends on the variable ϑ and the other on φ only. It is therefore a ***factorised*** form of the solution. The variables are separated in it. Therefore, we talk about ***separation*** of the variables.

After substituting the form (47) into eq. (45), we get a differential equation (DE)

$$\text{we do not even write it} \tag{48}$$

We should now solve the equation to get Θ_l^m and then the eigenfunctions of the operator $\hat{L}^2$ we are looking for. Such a straightforward approach is especially suitable when meeting QM for the first time; it employs solving the DE with the aid of the mathematical analysis aparatus and is rather lengthy. In this course we will better demonstrate an elegant and much more general akgebraic approach how to determine the eigenvalues λ_{lm} of the operator $\hat{L}^2$. In doing this we will also see that they will not depend on the index m which means they will be degenerate [because the operator $\hat{L}^2$ will in general have more eigenfunctions $Y_l^m(\vartheta,\varphi)$ for a single eigenvalue λ_l]. In the above consideration using Theorems 6 and 7, we have already proved the fact that the eigenvalues will be in *some* way degenerate. In next section we will in a more general way demonstrate, how precisely the are degenerate.

4.5 Eigenvalues of the Angular Momenta Operators

It turns out that commutation relations of the form (28) apply not on for the orbital angular momentum,[4] but also for all other angular momenta, for example spin. Therefore, in QM we distinguish what kind of AM it is: for example the operator (26) is the ***orbital*** AM operator. Spin of a particle is associated with an (unspecified) operator

---

[3] In some literature, for instance in [2], a notation Y_{lm} is being used for these functions. However, most of today's sources and literature use the notation Y_l^m, while the symbol Y_{lm} is reserved for closely related so-called real spherical harmonics, which in their essence are the real and imaginary components of the functions Y_l^m.

[4] The orbital AM is the usual AM, which is in classical physics calculated as $\vec{r} \times \vec{p}$ and in quantum physics with the use of the oparator $\hat{\vec{r}} \times \hat{\vec{p}}$.

of ***spin*** AM. Sum of the orbital and spin AM is the *total* AM and it has its operator as well. We are not going to deal with this now. However, we will perform the following analysis by *assuming* the validity of commutation relations of the kind (28) and we will derive consequences almost purely from this assumption. Instead of symbols such as $\hat{L}_x$, we will use notations $\hat{J}_x$, $\hat{J}_y$, $\hat{J}_z$ to emphasize that it does not have to be just an orbital AM. These three operators will therefore form a vector operator $\hat{\vec{J}}$. For simplicity, we will consider it dimensionless. We make this by choosing $\hbar = 1$ (or by omitting $\hbar$). In many literature, if possible, the Planck constant $\hbar$ is also set equal to 1 and is not even written.

Commutation Relations. Let us therefore consider linear hermitian operators for which commutation relations [2, 3]

$$[\hat{J}_x, \hat{J}_y] = \mathrm{i}\hat{J}_z\,, \qquad [\hat{J}_y, \hat{J}_z] = \mathrm{i}\hat{J}_x\,, \qquad [\hat{J}_z, \hat{J}_x] = \mathrm{i}\hat{J}_y \tag{49}$$

apply. These relationships are often taken as the definition of what we consider in QM to be the angular momentum. Let us define an operator

$$\hat{J}^2 = \hat{J}_x^2 + \hat{J}_y^2 + \hat{J}_z^2 \tag{50}$$

Exactly as we had above [see (30)], it can be proved here purely on the basis of relations (49) that (done as an exercise)

$$[\hat{J}^2, \hat{J}_x] = [\hat{J}^2, \hat{J}_y] = [\hat{J}^2, \hat{J}_z] = 0 \tag{51}$$

This, according to Theorem 7, means that the eigenfunctions of the operators $\hat{J}_z$, $\hat{J}^2$ can be constructed to be common to both of these operators. However, our task now will be to find eigenvalues of $\hat{J}_z$ and $\hat{J}^2$ in particular and to learn something about the mentioned eigenfunctions. However, instead of the eigenfunctions labeled as Y_l^m, we will use a more generally applicable notation $|j, m\rangle$ and we will call this abstract symbol an eigenvector. For now let us just take it as a notation [2, 3]. So we have to solve the double task [compare to the pair of equations (45) and (46)]

$$\boxed{\begin{aligned} \hat{J}^2|j, m\rangle &= \eta_j|j, m\rangle \qquad (52)\\ \hat{J}_z|j, m\rangle &= m|j, m\rangle \qquad (53)\end{aligned}}$$

We do not assume anything about the value of m yet, we just know that it should be real because of the hermicity assumed above. The same for η_j. We take the index j for a more general analogue of the index l from section 4.4 about the orbital angular momentum. By the notation introduced in equations (52) and (53) we already anticipate (insipred by the knowledge acquired earlier) that the eigenvalues η_j of the operator $\hat{J}^2$ will not depend on m. However, we will verify this guessed property by a derivation; therefore, it will not stay guessed. The independence of the eigenvalues η_j of the operator $\hat{J}^2$ on m is on a more general level the same property as is the independence of the eigenvalues λ_{lm} of the operator $\hat{L}^2$ on m. Although without a proof but still with a sufficient motivation, we have already mentioned the latter independence at the end of section 4.4. The independence of η_j on m means that the eigenvalues of the operator $\hat{J}^2$ (hence also of $\hat{L}^2$) will be degenerate. In solving given task we will really prove it and we opt to use the simple symbol η_j (i.e. without m) already from the start because it is a practical notation and there will be no need to change it later.

Raising and Lowering Operators. Let us now define the pair of operators (for the sake of simple notation we start to omit the hats)

$$J_+ = J_x + \mathrm{i}J_y\,, \qquad J_- = J_x - \mathrm{i}J_y \tag{54}$$

Obviously, they are not hermitian operators but it is not a problem. It is useful to know their commuation properties, for example immediately found relations

$$[J^2, J_+] = [J^2, J_-] = 0 \tag{55}$$

as well as further ones which we know from exercise:[5]

$$[J_z, J_+] = J_+\,, \qquad [J_z, J_-] = -J_-\,, \qquad [J_+, J_-] = 2J_z \tag{56}$$

---

[5] It is easy to show using (54) and (49). For instance, $[J_z, J_+] = [J_z, J_x + \mathrm{i}J_y] = \mathrm{i}J_y + \mathrm{i}(-\mathrm{i})J_x = J_x + \mathrm{i}J_y = J_+$.

As we will see, the followng additional identities will also be needed (proved in exercise)[6]

$$J_+J_- = J^2 - J_z^2 + J_z \tag{57a}$$
$$J_-J_+ = J^2 - J_z^2 - J_z \tag{57b}$$

Let us now explore the action of the operators $J_\pm$ on the unknown eigenvectors $|j,m\rangle$.

$$J_+|j,m\rangle \overset{\text{def}}{=} |j,m\rangle_+\,, \qquad\qquad J_-|j,m\rangle \overset{\text{def}}{=} |j,m\rangle_- \tag{58}$$

Let us at first try (and calculate) this:

$$J^2|j,m\rangle_+ \;=\; J^2J_+|j,m\rangle \;=\; J_+J^2|j,m\rangle \;=\; J_+\eta_j|j,m\rangle \;=\; \eta_j|j,m\rangle_+$$

And analogously

$$J^2|j,m\rangle_- \;= \qquad\qquad\qquad\qquad\qquad\qquad\qquad = \;\eta_j|j,m\rangle_-$$

Thus, the vectors $|j,m\rangle_+$ and $|j,m\rangle_-$ are also eigenvectors of the operators J^2, with the eigenvalue in both cases being the same η_j as is for the eigenvecor $|j,m\rangle$. (59)

Examine also the operation

$$\begin{aligned} J_z|j,m\rangle_+ \;&=\; J_zJ_+|j,m\rangle \;=\; (J_+ + J_+J_z)|j,m\rangle \;=\; |j,m\rangle_+ + J_+m|j,m\rangle \;= \\ &= (1+m)|j,m\rangle_+ \end{aligned}$$

and analogously also

$$J_z|j,m\rangle_- \;= (-1+m)|j,m\rangle_-$$

Thus, the vectors $|j,m\rangle_\pm$ are also eigenvectors of the operator J_z, with eigenvalues being $m \pm 1$. (60)

We already see from the above said, that the operator $\boldsymbol{J^2}$ has a ***degenerate spectrum***, because, e.g., the three different eigenvectors $|j,m\rangle_-$, $|j,m\rangle$ a $|j,m\rangle_+$ correspond to the same eigenvalue η_j. What regards to J_z, let us assume non-degeneracy of its spectrum. We have a good motivation for this as its particular example, L_z, has the non-degenerate spectrum. Later we will convince ourselves about the correctness of this assumption also for the general AM. Based on the two results framed above, it can then be said that $|j,m\rangle_+$ is (up to a constant) equal to $|j,m+1\rangle$ and analogously for $|j,m\rangle_-$. We write it as follows:

$$|j,m\rangle_+ \equiv J_+|j,m\rangle \;=\; C^{(+)}_{jm}\,|j,m+1\rangle \tag{61a}$$
$$|j,m\rangle_- \equiv J_-|j,m\rangle \;=\; C^{(-)}_{jm}\,|j,m-1\rangle \tag{61b}$$

where $C^{(+)}_{jm}$ a $C^{(-)}_{jm}$ are as yet unspecified constants. In the notation as for usual functions, we will express this by equations

$$\begin{aligned} Y^{(+)}_{j,m} \equiv J_+Y_j^m \;&=\; C^{(+)}_{jm}\,Y_j^{m+1} \\ Y^{(-)}_{j,m} \equiv J_-Y_j^m \;&=\; C^{(-)}_{jm}\,Y_j^{m-1} \end{aligned}$$

The operators $J_\pm$ modify the eigenstates $|j,m\rangle$ by increasing or decreasing the eigenvalue m: for example, by acting with J_+ on state $|j,m\rangle$, we create state $|j,m{+}1\rangle$ (up to a less significant constant). Therefore, J_+ is called the ***raising*** operator and J_- the ***lowering*** operator.

---

[6] This also can be done easily: $J_+J_- = (J_x + \mathrm{i}J_y)(J_x - \mathrm{i}J_y) = J_x^2 + \mathrm{i}J_yJ_x - \mathrm{i}J_xJ_y + J_y^2 = J_x^2 + J_y^2 + J_z$.

Determination of the Constants $C^{(\pm)}_{jm}$. Let us try to express the square of the norm of the functions $Y^{(+)}_{j,m}$, i.e. the scalar product[7] $\left\langle Y^{(+)}_{j,m} \middle| Y^{(+)}_{j,m} \right\rangle$:

$$\int \left[Y^{(+)}_{j,m}\right]^* Y^{(+)}_{j,m}\,\mathrm{d}\tau \;=\; \int \left[C^{(+)}_{jm}\, Y^{m+1}_{j}\right]^* C^{(+)}_{jm}\, Y^{m+1}_{j}\,\mathrm{d}\tau \;=\; \left|C^{(+)}_{jm}\right|^2$$

We assume that the functions Y^m_j (also the vectors $|j,m\rangle$) are normalised to 1. The scalar product can also be written as follows:

$$\int \left[Y^{(+)}_{j,m}\right]^* Y^{(+)}_{j,m}\,\mathrm{d}\tau \;=\; \int \left[J_+ Y^m_j\right]^* J_+ Y^m_j\,\mathrm{d}\tau \;=\; \int \left(Y^m_j\right)^* J_+^\dagger J_+ Y^m_j\,\mathrm{d}\tau \;=\; \int \left(Y^m_j\right)^* J_- J_+ Y^m_j\,\mathrm{d}\tau$$

We used the fact the the operators J_+ and J_- are hermitian conjugate to each other (mutually hermitian conjugate) – it is known from exercises. Preto platí

$$\left|C^{(+)}_{jm}\right|^2 = \int \left(Y^m_j\right)^* J_- J_+ Y^m_j\,\mathrm{d}\tau$$

In the abstract formalism of Dirac bra and ket vectors using the vectors $|j,m\rangle$ we write it as follows (and let us write it also for the $C^{(-)}_{jm}$, what we would derive quite analogously):

$$\left|C^{(+)}_{jm}\right|^2 = \langle j,m|J_-J_+|j,m\rangle\,, \qquad \left|C^{(-)}_{jm}\right|^2 = \langle j,m|J_+J_-|j,m\rangle \tag{62}$$

From the exercise we know the operator identities (57a) and (57b). Using them we get

$$\begin{aligned}
\left|C^{(+)}_{jm}\right|^2 &= \langle j,m|J^2|j,m\rangle - \langle j,m|J_z^2|j,m\rangle - \langle j,m|J_z|j,m\rangle &&= \eta_j - m^2 - m \\
\left|C^{(-)}_{jm}\right|^2 &= &&= \eta_j - m^2 + m
\end{aligned}$$

Thus, we find that relations

$$\left|C^{(+)}_{jm}\right|^2 = \eta_j - m(m+1) \tag{63a}$$

$$\left|C^{(-)}_{jm}\right|^2 = \eta_j - m(m-1) \tag{63b}$$

apply. Their left-hand sides clearly say that the expressions must be non-negative. Thus ($\wedge$ – *and*, i.e. the conjunct)

$$\boxed{\eta_j \geq m(m+1)} \quad \wedge \quad \boxed{\eta_j \geq m(m-1)}\,, \qquad \forall m \tag{64}$$

This implies limitations on m *at given* η_j; the values of m must be bounded from above and below.

Upper Limit. Generically, however, equation

$$J_+|j,m\rangle = C^{(+)}_{jm}\,|j,m+1\rangle$$

applies [see (61a)], which would “want” to increment m without termination. To terminate this climb, $C^{(+)}_{j,m_{\max}}$ has to be 0 for certain $m_{\max}$. Eq. (63a) then implies that $\eta_j = m_{\max}(m_{\max}+1)$. Note that now we are using the symbol j only in the meaning of an index, both in the eigenvalues eta_j and in the eigenfunctions Y^m_j (and in the case of the abstract notation also in the eigenvectors $|j,m\rangle$). However, we have not yet assigned any value to the j index; the index itself alone does not enter any formula, that is it is an unused symbol. So let’s start using it instead of the lengthy $m_{\max}$:

$$j \equiv m_{\max} \tag{65}$$

The consequence of eq. (63a) is then written shorter:

$$\boxed{\eta_j = j(j+1)} \tag{66}$$

From the above exposition, it is reaaly possible to observe [see eq. (63a)] that, if, by subsequent incrementing of the number m, we were proceeding higher and higher (from the vector $|m\rangle$ going to $|m+1\rangle$ etc), then this climb would be stopped since $C^{(+)}_{jj} = 0$. at the value of $m = m_{\max} \equiv j$

---

[7] The scalar product S, called also the dot product, of functions f and g is defined by the expression $S = \int f^* g\ \mathrm{d}\tau$, in which the integration goes over the whole range of the variable τ, in which the functions f and g are defined. In the Diract formalism, the scalar product is written as $S = \langle f|g\rangle$. The symbol $\langle f|$ alone is called *bra* and $|g\rangle$ is called emphket (from the word *brackets*).

Lower Limit. Generically, relation

$$J_-|j,m\rangle = C^{(-)}_{jm}\,|j,m-1\rangle$$

applies as well [see (61b)], which, on the other hand, would “want” to decrement m without termination. To terminate this steping down, $C^{(-)}_{j,m_{\min}}$ has to be 0 for certain $m_{\min}$. From the second of the inequalities (64) if follows that

$$\eta_j = m_{\min}(m_{\min}-1) \tag{67}$$

that is

$$m^2_{\min} - m_{\min} - j(j+1) = 0$$

The solutions of this quadratic equation are the roots

$$m_{\min} \in \{j+1, -j\}$$

The second of them is obviosly the one we need. Thus

$$m_{\min} = -j$$

In the end, let us recall: just as the vanishing $C^{(+)}_{j,j}$ stopped us as we went up the “ladder”, so now (when we move down) we have $C^{(-)}_{j,-j} = 0$ instead.

Eigenvalues. Thus, for given η_j (equivalently for given j), the following values of m are possible:

$$m \in \{-j, -j+1, \ldots, j\}\,, \qquad\qquad \text{i.e. } 2j+1 \text{ values}$$

The *number* of values must definitely be a non-negative integral number (integer), and in this particular case the number is at least 1. Therefore, $\boxed{2j \text{ is a non-negative integer}}$; hence $\boxed{j \text{ is a non-negative integer or half-integer.}}$ and $\boxed{m \text{ are integral numbers or half-integers}}$.

So let us summarize what we have found:

$$\boxed{\begin{array}{rcr} J^2|j,m\rangle &=& j(j+1)\,|j,m\rangle \\ J_z|j,m\rangle &=& m\,|j,m\rangle \end{array}} \tag{68}$$

$$j \in \{0,1,2,\ldots\}, \ \text{ or } j \in \left\{\frac{1}{2}, \frac{3}{2}, \frac{5}{2}, \ldots\right\}, \qquad\qquad m \in \{-j, -j+1, \ldots, j\}$$

So, we have determined the eigenvalues of the operators J^2 and J_z. We see that the eigenvalues of the operator J^2 are degenerate, because for one j we have $2j+1$ of different m values, and thus so many different eigenvectors.

Thus, by a purely algebraic procedure, we obtained the eigenvalues of the corresponding operators from the postulated (but at least in the case of the orbital angular momentum justified) commutation relations. While for the orbital angular momentum we got the integral numbers m as the eigenvalues of L_z, we have derived at least the mathematical possibility of half-integral ms from relations (49). The question is whether such a mathematical possibility is realised somewhere in nature. Experiments answer *yes.* This possibility is realized in the case of the spin AM of an electron and other fermions. Spin is the internal AM of a particle. Different angular momenta of a system (in the simplest case of a single particle) add up. Therefore also the resulting (composite, summed up) AM of an electron has a half-integer z component (and thus also a projection of the AM to any axis since we can choose the z direction arbitrarily). Let us remind again that each (also a composite) AM in quantum mechanics must obey the commutation relations of the kind (49). That this is the case is shown in particular by the agreement of the experiments and the theory built on this assumption.

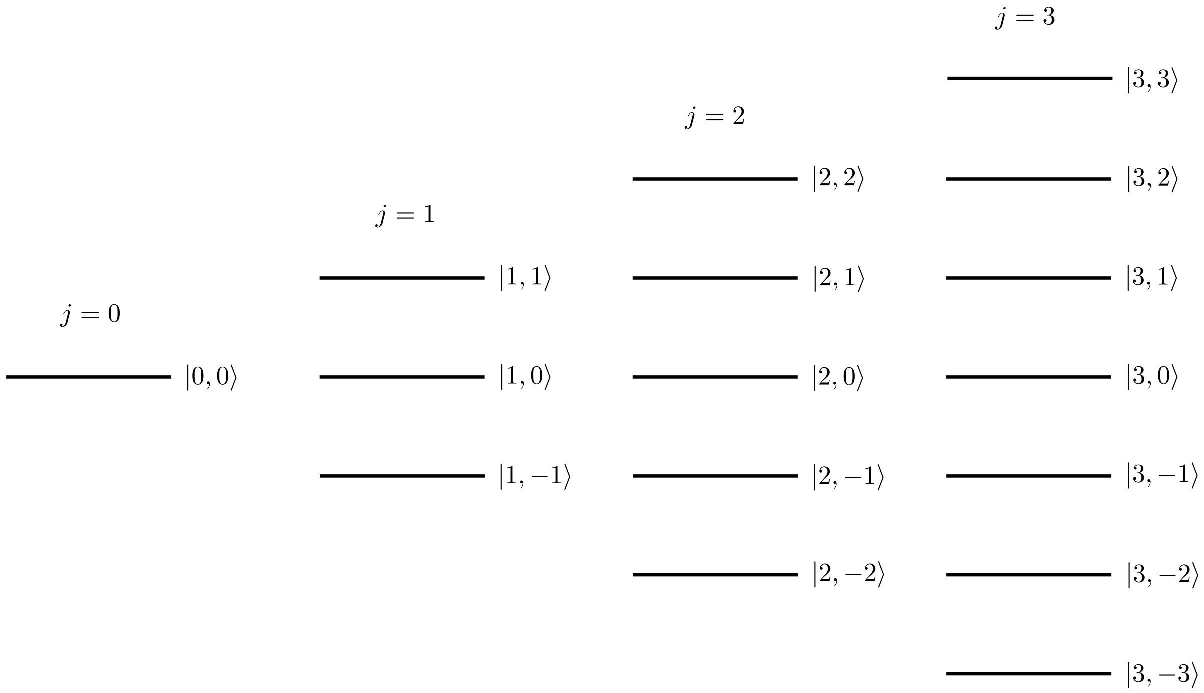

Figure 1: A scheme of the common eigenvectors of the operators J^2 and J_z for the lowest integer j's.

Phase Convention. We have not yet found any specific expressions for the eigenvectors $|j,m\rangle$, nor unambiguous values of the constants $C^{(\pm)}_{jm}$, only the squares of their absolute values. The phase (argument) of these constants cannot be unambiguously determined; it just needs to be defined. Most often we choose it to have the constants real and positive. The we get the following formulae expressing the action of the rasing and lowering operators [see also (61) and (66)]:

$$\boxed{\begin{aligned} J_+|j,m\rangle &= \sqrt{j(j+1)-m(m+1)}\;|j,m+1\rangle \equiv \sqrt{(j-m)(j+m+1)}\;|j,m+1\rangle \\ J_-|j,m\rangle &= \sqrt{j(j+1)-m(m-1)}\;|j,m-1\rangle \equiv \sqrt{(j+m)(j-m+1)}\;|j,m-1\rangle \end{aligned}} \qquad \begin{aligned}&(69\text{a})\\&(69\text{b})\end{aligned}$$

This phase convention is called the ***Condon-Shortley convention***.

A schematic representation of the spectrum of common eigenvectors of the operators J^2 and J_z for some of the lowest *integral* values of j is shown in Fig. 1. Within each of the "ladders" we can "move" using the action of the rasing and lowering operators $J\pm$.

4.6 Common Eigenfunctions of Operators $\hat{L}_z$ and $\hat{L}^2$ (Part 2)

It is worth remembering that in the general discussion of the AM in the previous section, the numbers j and m could either be integer or half-integers. The half-interal ones, as it in QM turns out, are important for the spin AM. Now, however, we resume the interrupted analysis of the orbital AM, which we in the traditional way started in section 4.4. We would complete the ***traditional approach*** by ***solving the partial DE*** (48) ***in spherical coordinates*** ϑ, φ. Although we only started the traditional approach, we learnt some new things (knowledge). Armed by the knowledge of the rasing and lowering oparators, we will now complete the analysis of the orbital AM in a different way.

Instead of the index j, we will be using l for the orbital AM. And what is essential, we have already found out that we have the integral ms in the case of the orbital AM; see (39). Hence l will also be integers. Instead of the mathematically demanding and leghty procedure indicated at the end of section (4.4), we are now going to look for the common eigenfunction of $\hat{L}_z$ and $\hat{L}^2$ ***by exploring the effect of these operators on polynomials in cartesian coordinates***. We know from the exercise that the operator $\hat{L}^2$ can in cartesian coordinates be reexpressed in the form

$$\hat{L}^2 = (y^2+z^2)\hat{p}_x^2 + (z^2+x^2)\hat{p}_y^2 + (x^2+y^2)\hat{p}_z^2 - 2(xy\hat{p}_x\hat{p}_y + yz\hat{p}_y\hat{p}_z + zx\hat{p}_z\hat{p}_z) + 2\mathrm{i}\hbar(x\hat{p}_x + y\hat{p}_y + z\hat{p}_z) \tag{70}$$

from which we easily get the final form

$$\boxed{\hat{L}^2 = \hbar^2\left[-(y^2+z^2)\nabla_x^2 - (z^2+x^2)\nabla_y^2 - (x^2+y^2)\nabla_z^2 + 2(yz\nabla_{yz}^2 + zx\nabla_{zx}^2 + xy\nabla_{xy}^2) + 2(x\nabla_x + y\nabla_y + z\nabla_z)\right]} \tag{71}$$

We can see that it contains products of polynomials and partial derivatives and if the derivative is n^{th}, the multiplying polynomial is of the degree n. This is also the case with the $\hat{L}^2$ operator:

$$\boxed{\hat{L}_z = -\mathrm{i}\hbar\,(x\nabla_y - y\nabla_x)} \tag{72}$$

For brevity we will set $\hbar = 1$ in calculations of this section. We will do calculations of sections 4.6.1 and 4.6.2 as an exercise.

4.6.1 Action of $\hat{L}_z$ and $\hat{L}^2$ on Polynomials $f = ax + by$

It is a homogeneous polynomial of the 1st degree (sometime perhaps called a *linear form*).

Action of $\hat{L}_z$. Let us thus first try to explore what we get by acting of $\hat{L}_z$ on such a polynomial; x and y are cartesian coordinates, i.e. real quantities. As yet a and b are arbitrary constants.

$$\hat{L}_z f = -\mathrm{i}(x\nabla_y - y\nabla_x)(ax+by) = -\mathrm{i}(-ay+bx) = -\mathrm{i}bx + \mathrm{i}ay$$

We got a polynomial somehow resembnling the original f. We have to solve (among other things) the problem

$$\hat{L}_z f = \lambda f \tag{73}$$

with λ being an eigenvalue. We have already solved this task in section 4.3, but the functions found there were eigenfunctions of $\hat{L}_z$ only, not of $\hat{L}^2$. Therefore we are now solving the problem again and other way. Upon substituting of what we got few lines above into equation (73), we get

$$-\mathrm{i}bx + \mathrm{i}ay = \lambda ax + \lambda by$$

This equation has to apply for aby point in space, that is also for any real x and y. So we obtain equalities

$$\begin{aligned} -\mathrm{i}b &= \lambda a\,, & \mathrm{i}a &= \lambda b \\ \Rightarrow \qquad b &= \mathrm{i}a\lambda\,, & \mathrm{i}a &= \mathrm{i}a\lambda^2 \end{aligned} \tag{74}$$

Therefore

$$\text{if } a \neq 0 \text{ then} \qquad \lambda^2 = 1 \tag{75}$$

For $a \neq 0$ we have thus found eigenvalues $\lambda \in \{+1, -1\}$ of the operator $\hat{L}_z$ and eigenfunctions (in the form of polynomials)

$$f = ax + \mathrm{i}\lambda ay = a(x + \mathrm{i}\lambda y) = ar\sin\vartheta\,(\cos\varphi + \mathrm{i}\lambda\sin\varphi) = ar\sin\vartheta\,e^{\mathrm{i}\lambda\varphi} \tag{76}$$

The eigenvalues ±1 really belong to the set of those that we have found for the operator $\hat{L}_z$ in section 4.3. We have there determined the form $K\exp(\mathrm{i}\lambda\varphi)$ for the eigenfunctions, with K being anything independent on the angle φ. And it is really so: the multiplicating factor $ar\sin\vartheta$ does not include φ. What if we had $a = 0$? Then we would have $b = 0$ and would obtain a trivial function identically equal to zero. It is an uninteresting solution as it does not bring any information.

Action of $\hat{L}^2$. Let us now act on the given polynomial by the operator of the square od the AM, expressed in the form (71). We easily obtain the results

$$\hat{L}^2 f = 2(x\nabla_x + y\nabla_y + z\nabla_z)(ax+by) = 2ax + 2by = 2f \tag{77}$$

We see that this polynomial is an eigenfunction for $\hat{L}^2$ even for arbitrarily chosen a, b. The corresponding eigenvalue is the number 2. (Writing also $\hbar$, it would be $2\hbar^2$.)

Preliminary Summary for the Homogeneous Polynomials of the 1st Degree. Of course, if we want the polynomial f to be an eigenfunction for both the operators, we have to restrict a and b by the condition (74): $b = \mathrm{i}a\lambda$. And let us note that the eigenvalue 2 is just $l(l+1)$ for $l = 1$. This l and the eigenvalues $\lambda = \pm 1$ determined few lines above agree with the result (68) (in which j and m were used instead of l and λ). We can write the equations solved here and the solutions found now as follows:

$$\hat{L}^2 f_{1,+1} = 2\, f_{1,+1}\,, \qquad \hat{L}^2 f_{1,-1} = 2\, f_{1,-1}\,, \qquad \hat{L}_z f_{1,+1} = +1\, f_{1,+1}\,, \qquad \hat{L}_z f_{1,-1} = -1\, f_{1,-1}$$

or, more briefly,

$$\boxed{\hat{L}^2 f_{1,\pm 1} = 2\, f_{1,\pm 1}}\,, \qquad \boxed{\hat{L}_z f_{1,\pm 1} = \pm 1\, f_{1,\pm 1}} \tag{78}$$

where

$$\boxed{f_{1,\pm 1} = a(x \pm \mathrm{i}y) = ar \sin\vartheta\, e^{\pm \mathrm{i}\varphi}} \tag{79}$$

Instead of f_{lm}, we almost could use the already introduced general notation Y_l^m according to (47). We will not do it because the functions f_{lm} do not have the proper normalisation constants which are being used for Y_l^m; f_{lm} even include the dependence on the spherical variable r and we know that this one does not have to be there; aiming to determine as simple common eigenfunctions (of the operators $\hat{L}^2$ and $\hat{L}_z$) as possible, we will remove the r later on. In doing the derivation, it was more practical to keep r there.

Determination of the Eigenfunction $f_{1,0}$. We will complete the *triplet* of the functions $f_{1,m}$ only after $f_{1,0}$ is determined as well. We have learnt a little bit above [formulae (69)] that once one of the functions f_{lm} is known (i.e. for one particular m), then, using the effect of the rasing and/or lowering operator, we can determine f_{lm} for all other ms. So, let us use the second of those formulae. We get

$$f_{1,0} = \frac{1}{\sqrt{2}}\, \hat{L}_- f_{1,+1} \tag{80}$$

where $f_{1,+1}$ is given by the expresion (79). To obtain the results, we need to calculate the effect of $\hat{L}_-$ on $f_{1,+1}$.

$$\begin{aligned}\hat{L}_- f_{1,+1} &= (L_x - \mathrm{i}L_y)[a(x+\mathrm{i}y)] = a\{[-\mathrm{i}(y\nabla_z - z\nabla_y)] - \mathrm{i}[-\mathrm{i}(z\nabla_x - x\nabla_z)]\}(x+\mathrm{i}y) = \\ &= a(-\mathrm{i}y\nabla_z + \mathrm{i}z\nabla_y - z\nabla_x + x\nabla_z)(x+\mathrm{i}y) = a(-z-z) = -2az\end{aligned}$$

Upon substituting into (80), we obtain

$$\boxed{f_{1,0} = -\sqrt{2}\, az = -\sqrt{2}\, ar\cos\vartheta} \tag{81}$$

The Constant a. If one just requires $f_{1,+1}$, $f_{1,0}$, $f_{1,-1}$ to be some common eigenfunctions for L^2 a L_z, then the ***constant a*** can be an arbitrary complex number and it ***does not have to be the same for neither $f_{1,-1}$, nor $f_{1,+1}$***, neither $f_{1,0}$. Each of the three functions can have its own constant: if any eigenfunction is multiplied by any constant, the function remains to be the eigenfunction. The choice of these constants is a matter a practical convenience and normalisation. In our study of the general AM, we have introduced the Condon-Shortley convention [see (69)]. In the case of the orbital AM which we are studying now, we will choose the constants to be consistent with the Condon-Shortley convention. For example, using a procedure similar to the above one, i.e. by utilising the action of the lowering opeerator, we easily find out that

$$f_{1,-1} = \frac{1}{\sqrt{2}} L_- f_{1,0} = -a(x - \mathrm{i}y) = -ar\sin\vartheta e^{-\mathrm{i}\varphi} \tag{82}$$

We see that the $f_{1,-1}$ determined in this way (i.e. in accordance with the Condon-Shortley convention) has an opposite sign than $f_{1,-1}$ determined above [eq. (79)].

4.6.2 Action of $\hat{L}_z$ and $\hat{L}^2$ on the Polynomials $f = ax^2 + by^2 + cxy$

In a way analogous to the one in section 4.6.1, let us now examine the effect of the operators of the AM on homogeneous polynomial of the 2nd degree (i.e. on quadratic forms, although the word *form* may rarely be used in English). We again choose the forms independent on z, because otherwise they could not be eigenfunctions for $\hat{L}_z$, as it is to convince.

Action of $\hat{L}_z$.

$$\hat{L}_z f = -\mathrm{i}(x\nabla_y - y\nabla_x)(ax^2 + by^2 + cxy) = -\mathrm{i}(-2axy + 2bxy + cx^2 - cy^2)$$

$$\begin{aligned} \hat{L}_z f = \lambda f \quad \Longrightarrow \quad 2\mathrm{i}(a-b)xy - \mathrm{i}cx^2 + \mathrm{i}cy^2 = \lambda a x^2 + \lambda b y^2 + \lambda c xy \quad \Longrightarrow \\ -\mathrm{i}c = \lambda a \\ \mathrm{i}c = \lambda b \\ 2\mathrm{i}(a-b) = \lambda c \end{aligned}$$

Let us first focus on the solutions with $\lambda \neq 0$. They are found to be as follows:

$$b = -a\,, \qquad c = \mathrm{i}\lambda a\,, \qquad \lambda \in \{+2, -2\}$$

This gives eigenfunctions of $\hat{L}_z$ for $\lambda = \pm 2$ as follows:

$$f_{\pm} = a(x^2 - y^2) \pm \mathrm{i}2axy = a(x \pm \mathrm{i}y)^2 \tag{83}$$

Action of $\hat{L}^2$. Are the functions just written above eigenfunctions also for $\hat{L}^2$? [Look at the expression (71).] Let us examine it and do it for general a, b, c to have it more interesting.

$$\begin{aligned} \hat{L}^2 f = {} & [-(y^2+z^2)\nabla_x^2 - (z^2+x^2)\nabla_y^2 - (x^2+y^2)\nabla_z^2]\,(ax^2+by^2+cxy) + \\ & + 2\,(yz\nabla_{yz}^2 + zx\nabla_{zx}^2 + xy\nabla_{xy}^2)\,(ax^2+by^2+cxy) + \\ & + 2\,(x\nabla_x + y\nabla_y + z\nabla_z)\,(ax^2+by^2+cxy) = \\ & = -2a\,(y^2+z^2) - 2b\,(z^2+x^2) + \\ & + 2\,cxy + \\ & + 2\,(2ax^2 + 2by^2 + cxy + cxy) = \\ & = (4a-2b)x^2 + (4b-2a)y^2 - (2a+2b)z^2 + 6cxy \end{aligned}$$

As we see, not every homogeneous polynomial of the 2nd degree form the title 4.6.2 is an eigenfunction of $\hat{L}^2$. But if we take those of them which are eigenfunctions of $\hat{L}_z$, i.e. those obeying $b = -a$ while c can be arbitrary in this case, we obtain

$$\hat{L}^2 f = 6[a(x^2 - y^2) + cxy] = 6f \tag{84}$$

For now, we have learnt enough about the polynomials of the 2nd degree. We focused on the solutions with $\lambda \neq 0$ only. The case of zero λ is not a solution now as it can easily be found: then we would have $c = 0$, $a = b$ and such a polynomial would not be an eigenfunction for $\hat{L}^2$.

Preliminary Summary for Homogeneous Polynomials of the 2nd Degree

$$\boxed{\hat{L}^2 f_{2,\pm 2} = 6\, f_{2,\pm 2}}\,, \qquad \boxed{\hat{L}_z f_{2,\pm 2} = \pm 2\, f_{2,\pm 2}} \tag{85}$$

where

$$\boxed{f_{2,\pm 2} = a(x \pm \mathrm{i}y)^2 = ar^2 \sin^2\vartheta\, e^{\pm \mathrm{i}2\varphi}} \tag{86}$$

and let us note that $6 = l(l+1)$ for $l = 2$.

Determination of the Eigenfunction $f_{2,+1}$. We already know how to do it: according to (69), we obtain

$$f_{2,+1} = \frac{1}{\sqrt{4}}\, \hat{L}_- f_{2,+2} \tag{87}$$

Hence, we need to calculate

$$
\begin{aligned}
&\hat{L}_- f_{2,+2} = (L_x - \mathrm{i}L_y)[a(x+\mathrm{i}y)^2] = a[(L_x - \mathrm{i}L_y)\,(x^2 - y^2 + 2\mathrm{i}xy] = \\
&= aL_x\,(-y^2 + 2\mathrm{i}xy) - \mathrm{i}aL_y(x^2 + 2\mathrm{i}xy) = \\
&= a[-\mathrm{i}z2y + 2\mathrm{i}\,\mathrm{i}zx - \mathrm{i}(-\mathrm{i})z2x - 2\mathrm{i}zy] = 2a(-2\mathrm{i}yz - 2xz) = -4az(x+\mathrm{i}y)
\end{aligned}
$$

Thus[8]

$$\boxed{f_{2,+1} = -2az(x+\mathrm{i}y) = -2ar^2\cos\vartheta\sin\vartheta e^{\mathrm{i}\varphi}} \tag{88}$$

Determination of the Eigenfunction $f_{2,0}$. The derivation goes through analogous steps (using the lowering operator). The result is

$$\boxed{f_{2,0} = -\frac{2}{\sqrt{6}}\,a\,(x^2 + y^2 - 2z^2) = \frac{2}{\sqrt{6}}\,ar^2\,(3\cos^2\vartheta - 1)} \tag{89}$$

4.6.3 Overall Summary

The eigenfunctions Y_l^m sought for are called ***spherical harmonic functions***, or, more briefly, spherical harmonics. Except for the insignificant factors, they are proportional to the functions f_{lm} found above. The eigenfunctions Y_l^l, as we have seen, can be found by examining the effect of the operators L_z and L^2 and L2 on polynomials. Y_l^m for $m < l$ can be determined by application of L_-. Thus, we have determined the solutions – the eigenvalues and common eigenfunctions (at least some of them) – of the equations

$$\hat{L}^2 Y_l^m = \hbar^2 l(l+1) Y_l^m \tag{90a}$$

$$\hat{L}_z Y_l^m = \hbar m Y_l^m \tag{90b}$$

$$l \in \{0, 1, 2, \dots\}, \quad m \in \{-l, -l+1, \dots, l\} \tag{90c}$$

in which we explicitely displayed the the eigenvalues found. Although we have not calculated it for a general l, we see that obviously (and it is indeed so) the following will apply:

- The lowest harmonic: a constant (Y_0^0; is can be seen even without calculating).
- We do not put the coefficient r^l in the expression Y_l^m, because the operators of the AM do not depend on r.
- The spherical functions, as we see, can be written in the form

  $$\boxed{Y_l^m(\vartheta,\varphi) = \Theta_l^m(\vartheta)\Phi_m(\varphi)} \tag{91}$$

  which we have found above [(47)], but we did not know yet, what indices would be at Θ. The functions $\Phi_m(\varphi)$ have form (39), which satisfies the standard normalisation in accordance with (40). We will not write the functions $\Theta_l^m(\vartheta)$ alone. Instead, we directly write the function $Y_l^m(\vartheta,\varphi)$ (a little below)

The spherical harmonic functions satisfy the Laplace equation; hence their name “harmonic”.

Parity. The spherical harmonics with an even l do not change their sign upon the $\vec{r} \to -\vec{r}$ inversion while those with an odd l do:

$$Y_l^m(-\vec{r}) = (-1)^l\, Y_l^m(\vec{r}) \tag{92}$$

Thus, we numerically quantify the parity by values $(-1)^l$.

---

[8] Some intermediate calculations and results needed for these and similar calculations will be convenient to write:

$$L_x(x^2) = 0\,,\quad L_x(y^2) = 2\mathrm{i}yz\,,\quad L_x(z^2) = -2\mathrm{i}yz\,,\quad L_y(x^2) = -2\mathrm{i}xz\,,\quad L_y(z^2) = 2\mathrm{i}xz$$

$$L_x(zx) = -\mathrm{i}xy\,,\quad L_x(yz) = \mathrm{i}(z^2 - y^2)\,,\quad L_y(xy) = -\mathrm{i}yz\,,\quad L_y(zx) = \mathrm{i}(x^2 - z^2)\,,\quad L_y(yz) = \mathrm{i}xy$$

There is no need to calculate explicitely all of the, for several of them can be derived by the cyclic permutations of the coordinates.

Orthogonality and Normalisation.

$$\int Y_l^{m*}(\vartheta,\varphi) Y_{l'}^{m'}(\vartheta,\varphi)\,\mathrm{d}\Omega = \delta_{ll'}\delta_{mm'} \tag{93}$$

where $\mathrm{d}\Omega = \sin\vartheta\,\mathrm{d}\vartheta\,\mathrm{d}\varphi$ is an element of the spatial angle and the integration goes over the entire range.

Completeness. The spherical harmonics form a complete set of functions on a unit sphere. In other words, any function of the variables ϑ, φ can be expressed as a linear combination of the sphererical harmonics.

$$f(\vartheta,\varphi) = \sum_{l=0}^{\infty}\sum_{m=-l}^{l} c_{lm} Y_l^m(\vartheta,\varphi) \tag{94}$$

Listing of the Lowest Spherical Harmonic Functions.

$$Y_0^0(\vartheta,\varphi) = \sqrt{\frac{1}{4\pi}} \tag{95a}$$

$$Y_1^1(\vartheta,\varphi) = -\sqrt{\frac{3}{8\pi}}\,\sin\vartheta\,e^{\mathrm{i}\varphi} \tag{95b}$$

$$Y_1^0(\vartheta,\varphi) = \sqrt{\frac{3}{4\pi}}\,\cos\vartheta \tag{95c}$$

$$Y_1^{-1}(\vartheta,\varphi) = \sqrt{\frac{3}{8\pi}}\,\sin\vartheta\,e^{-\mathrm{i}\varphi} \tag{95d}$$

$$Y_2^2(\vartheta,\varphi) = \sqrt{\frac{15}{32\pi}}\,\sin^2\vartheta\,e^{\mathrm{i}2\varphi} \tag{95e}$$

$$Y_2^1(\vartheta,\varphi) = -\sqrt{\frac{15}{8\pi}}\,\sin\vartheta\cos\vartheta\,e^{\mathrm{i}\varphi} \tag{95f}$$

$$Y_2^0(\vartheta,\varphi) = \sqrt{\frac{5}{4\pi}}\left(\frac{3}{2}\cos^2\vartheta - \frac{1}{2}\right) \tag{95g}$$

$$Y_2^{-1}(\vartheta,\varphi) = \sqrt{\frac{15}{8\pi}}\,\sin\vartheta\cos\vartheta\,e^{-\mathrm{i}\varphi} \tag{95h}$$

$$Y_2^{-2}(\vartheta,\varphi) = \sqrt{\frac{15}{32\pi}}\,\sin^2\vartheta\,e^{-\mathrm{i}2\varphi} \tag{95i}$$

The following sections of the current paragraph (4.6.3) need to be ***read carefully*** and known about, but you will not have to know how to derive the formulae that appear here, neither memorise them.

General Formula for the Spherical Harmonic Functions in the Condon-Shortley Phase Convention [6, 2, 3, 5].

$$Y_l^m(\vartheta,\varphi) = (-1)^m\sqrt{\frac{2l+1}{4\pi}\frac{(l-m)!}{(l+m)!}}\,P_l^m(\cos\vartheta)\,e^{\mathrm{i}m\varphi} \qquad \text{for } m\in\{0,1,\ldots,l\} \tag{96}$$

Evaluations of the spherical functions at negative indices m can be obtained with the aid of the relation (which we will not prove)

$$Y_l^{-m}(\vartheta,\varphi) = (-1)^m\left[Y_l^m(\vartheta,\varphi)\right]^* \qquad \text{for } m\in\{-l,-l+1,\ldots,l\} \tag{97}$$

Hence, in the Condon-Shortley phase convention, the spherical functions with positive ms will oscilate with m because of the factor $(-1)^m$. There will be no such oscillations in the case of the negative-m spherical function. $P_l^m(x)$ is the associated Legendre function of the degree l and order m. It is defined using the Legendre polynomials $P_l(x)$:

$$P_l^m(x) = (1-x^2)^{m/2}\frac{\mathrm{d}^m}{\mathrm{d}x^m}P_l(x) \qquad \text{for } m \geq 0 \tag{98}$$

The Legendre polynomials can be expressed using the relation

$$P_l(x) = \frac{1}{2^l l!}\frac{\mathrm{d}^l}{\mathrm{d}x^l}[(x^2-1)^l] \tag{99}$$

which is the Rodrigues formula.

The Condon-Shortley convention is ***usual in quantum mechanics***. Recall that we have introduced it in relation with the general angular momentum [equations (69)]. Indeed, in calculations such as those we did using the lowering operator in sections 4.6.1 and 4.6.2, the same signs of the spherical harmonic functions are obtained as we have written in the listing (95a)–(95i). To make the convention for the spherical harmonics completely described, they are defined so that:

- the functions Y_l^l for even l are positive, that is, for example, we will have $a > 0$ in the formula $f_{2,+2} = a(x+\mathrm{i}y)^2 = ar^2 \sin^2\vartheta\, e^{\mathrm{i}2\varphi}$ [see (86)],
- the functions will be negative for odd l, see for instance Y_1^{-1} above.

In literature, one can often see spherical harmonics typed with both their indices as subscripts (Y_{lm}); this is most frequently being used for the *real* spherical harmonics, which are the functions expressed from the real and imaginary parts of the usual (the complex) spherical harmonics (96). In the book [2], however, this notation is used for the complex spherical harmonics (96). We can also find formula (96) used for negative indices m. This can be done if we, e.g., define the Legendre polynomials for the negative values of m too, which would not be difficult [6, 5].

5 Particle in a Spherically Symmetric Force Field

An example is the hygrogen atom in which the electron moves in a coulombic field.

5.1 General Spherically Symmetric Field

At least at the beginning, we will solve a more general problem, a one with an almost arbitrary sperically symmetric potential energy. The task is to determine the eigenenergies and corresponding wavefunctions of such a Hamiltonian [1, 2, 3], i.e. to find solutions to the equation

$$\boxed{H\psi = E\psi} \tag{100}$$

where the Hamiltonian is

$$H = -\frac{\hbar^2}{2m}\Delta + V(r) \tag{101}$$

$V(r)$ is the spherically symmetric potential energy. It is often being called a *potential*, but its dimension (units) is that of energy.[9] We will show as an exercise that[10]

$$[H, L_x] \;=\; [H, L_y] \;=\; [H, L_z] \;=\; [H, L^2] \;=\; 0 \tag{103}$$

---

[9] $V(r)$ differs from a potential by just a constant multiplicator; this can easily be seen if we express explicitly that it is a potential energy of a point charge q in an electrostatic potential $U(r)$: $V(r) = qU(r)$. Here $U(r)$ is really a potential, i.e. also by its dimension.

[10] Using (43), (105) and (106), it can be shown that the Hamiltonian (101) can be expressed in the form

$$\boxed{H = H_r + \frac{1}{2mr^2}L^2} \quad \text{where} \quad H_r = -\frac{\hbar^2}{2mr^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + V(r) \tag{102}$$

The AM operators do not depend on r. Then, using also commutation relations of the type $[L^2, L_x] = 0$, is is easily seen that formulae (103) really apply.

Therefore, it is possible to find commnon eigenfunctions of chosen three operators, for instance L_z, L^2, H. (We cannot add a fourth one since the cartesian components of $\vec{L}$ do not commute.) We will use the selected commutating operators to find solutions to the problem (100). We have already found the common eigenfunctions for the two operators, L_z and L^2: the spherical harmonic functions $Y_l^m(\vartheta, \varphi)$. In order for some eigenfunction of the Hamiltonian H to be also an eigenfunction of the operators L_z and L^2, the angular dependence of the eigenfunction must remain the same as it is in the case of the spherical harmonic functions. Therefore, the solutions to the problem (100) must be sought in the (factorised) form

$$\psi(\vec{r}) = R(r) Y_l^m(\vartheta, \varphi) \tag{104}$$

We need to find the unknown functions $R(r)$ and of course also the eigenvalue (energies) E. For the Laplace operator appearing in (101), it is advantageous to use the expression using spherical coordinates. (which we will learn about in the exercise):

$$\Delta \equiv \vec{\nabla}^2 = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2 \frac{\partial}{\partial r}\right) + \frac{1}{r^2}\nabla^2_{\vartheta,\varphi} \tag{105}$$

where

$$\nabla^2_{\vartheta,\varphi} = \frac{1}{\sin\vartheta}\frac{\partial}{\partial\vartheta}\left(\sin\vartheta \frac{\partial}{\partial\vartheta}\right) + \frac{1}{\sin^2\vartheta}\frac{\partial^2}{\partial\varphi^2} \tag{106}$$

The Laplace operator and the derivatives in it act on functions in the followig sense:

$$\vec{\nabla}^2 f = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2 \frac{\partial f}{\partial r}\right) + \text{and similarly the remaining terms}$$

We then write our Hamiltonian as follows:

$$H = -\frac{\hbar^2}{2m}\left[\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2 \frac{\partial}{\partial r}\right) + \frac{1}{r^2}\nabla^2_{\vartheta,\varphi}\right] + V(r) \tag{107}$$

Substitute the Hamiltonian (107) and the proposed form (104) of the solution into the stationary SchE (100). We will get the equation (omitting indices l and m and the function variables)

$$-\frac{\hbar^2}{2m}\left\{\frac{1}{r^2}\frac{\partial}{\partial r}\left[r^2 \frac{\partial(RY)}{\partial r}\right] + \frac{1}{r^2}\nabla^2_{\vartheta,\varphi}(RY)\right\} + VRY = E\,RY \tag{108}$$

We will now use the standard procedure for solving separable differential equations (DEs): we do the derivatives in this way:

$$\frac{\partial(RY)}{\partial r} = Y\frac{\mathrm{d}R}{\mathrm{d}r}$$

and divide the whole equation by the expression RY. We regroup the terms and get the equation

$$\frac{1}{R}\frac{\mathrm{d}}{\mathrm{d}r}\left(r^2\frac{\mathrm{d}R}{\mathrm{d}r}\right) + \frac{2m}{\hbar^2}r^2[E - V(r)] = -\frac{1}{Y}\nabla^2_{\vartheta,\varphi}Y \tag{109}$$

This is already a differential equation in a *separate* form, because one group of terms depends on only one of the variables (r), while the other group depends only on the remaining variables ϑ, φ. Therefore, no matter how we change e.g. the variable r, the right side of the equation will certainly not change, and therefore neither its left side. Thus

$$-\frac{1}{Y}\nabla^2_{\vartheta,\varphi}Y = \text{konšt} \overset{\text{def}}{=} \lambda \tag{110}$$

Knowing that [see (43)]

$$L^2 = -\hbar^2\nabla^2_{\vartheta,\varphi}$$

we obtain

$$L^2\,Y = \hbar^2\lambda Y \tag{111}$$

which is the problem that we already dealt with and solved [results (90a), (90b) and further in that section]. This time we have “peeled off” the sub-task [equation (111)] from the larger task – from the problem of the particle in a spherically symmetric field. After simple rearangements and substiting $\lambda = l(l+1)$, differential equation (109) then takes the form

$$\frac{1}{r^2}\frac{\mathrm{d}}{\mathrm{d}r}\left(r^2\frac{\mathrm{d}R}{\mathrm{d}r}\right) + \left\{\frac{2m}{\hbar^2}[E - V(r)] - \frac{l(l+1)}{r^2}\right\}R = 0 \tag{112}$$

This is sometime called the *radial* Schrödinger equation. $R(r)$ is the radial wavefunction. If we do the outer derivative with respect to r in radial SchE (112), we obtain its other form:

$$\boxed{\frac{\mathrm{d}^2R}{\mathrm{d}r^2}+\frac{2}{r}\frac{\mathrm{d}R}{\mathrm{d}r}+\left\{\frac{2m}{\hbar^2}[E-V(r)]-\frac{l(l+1)}{r^2}\right\}R=0} \tag{113}$$

This is an ordinary dofferential equation of the 2^{nd} order.

Physical Meaning of the Term with l.

$$V_l=V(r)+\frac{\hbar^2}{2m}\frac{l(l+1)}{r^2} \tag{114}$$

This is an effective potential energy of given particle with inclusion of the contribution of the centrifugal force. Details can be found in the book [2] and other.

Simplification by a Substitution.

$$\chi(r)=rR(r) \tag{115}$$

Using this substitution [1], the radial SchE (113) is simplified to[11]

$$\boxed{\frac{\mathrm{d}^2\chi}{\mathrm{d}r^2}+\left\{\frac{2m}{\hbar^2}[E-V(r)]-\frac{l(l+1)}{r^2}\right\}\chi=0} \tag{116}$$

This ordinary differential equation (DE) belong to the class of singular differencial equations. The total wavefunction for the spherical problem is then, according to (104), expressed as

$$\psi(\vec{r})=\frac{\chi(r)}{r}\,Y_l^m(\vartheta,\varphi) \tag{117}$$

Restriction to Bound States. It is known that problem (100) of a particle in a central (i.e. a spherically symmetric) field has two kinds of solutions:

- *Bound states* – these are states, a wavefunction of which is localised around the force centre which mean that it vanishes are large r: $\lim_{|\vec{r}|\to\infty}|\psi(\vec{r})|=0$. Physically, this means that there is a high probability of finding the particle near the centre and tiny (practically zero) find it somewhere far from it.
- *Scattering states* – these are states, a wavefunction of which is delocalised, that is the above-written limit is non-zero. Physically, this means that the particle can with a non-negligible probability be found even at large distances from the force centre

We will discuss further properties of this two kinds of states a little later. However, we state already here that *we will only search for bound states.* Their wave functions, as we have written above, acquire non-negligible values in a certain restricted spatial domain. Therefore, we require the normalisation condition[12]

$$\boxed{\int|\psi(\vec{r})|^2\mathrm{d}^3r=1} \tag{118}$$

to apply. We express the integration element using the spherical coordinates:

$$\mathrm{d}^3r=r^2\sin\vartheta\,\mathrm{d}r\,\mathrm{d}\vartheta\,\mathrm{d}\varphi=r^2\mathrm{d}r\,\mathrm{d}\Omega \tag{119}$$

From the normalisation condition (118), and from the conventional normalisation of the spherical functions, $\int|Y_l^m(\vartheta,\varphi)$ 1 [more generally expressed by the orthonormality (93)], the following condition for the normalisation of the auxiliary function $\chi(r)$ is obtained:

$$\int_0^\infty|\chi(r)|^2\mathrm{d}r=1 \tag{120}$$

For such an integral to exists, $\chi(r)$ has to converge to zero sufficiently rapidly (for $r\to\infty$).

---

[11] To obtain it, we need to derived the relations $\frac{\mathrm{d}R}{\mathrm{d}r}=-\frac{\chi}{r^2}+\frac{1}{r}\frac{\mathrm{d}\chi}{\mathrm{d}r}$ and also $\frac{\mathrm{d}^2R}{\mathrm{d}r^2}=\frac{2}{r^3}\chi-\frac{2}{r^2}\frac{\mathrm{d}\chi}{\mathrm{d}r}+\frac{1}{r}\frac{\mathrm{d}^2\chi}{\mathrm{d}r^2}$.

[12] In the case of scattering states, the wave function could not be normalized to a finite number because it is not limited to a finite region of space. By convention, we would "normalise" it to the δ function.

Solving Equation (116) in a Neighborhood of the Singular Point $r = 0$. Assume that the potential of the given problem fulfills the condition

$$\lim_{r\to 0}\left[r^2 V(r)\right] = 0 \tag{121}$$

Although this somewhat reduces the generality of the class of potentials being studied, let's realize that this is fulfilled for any potential that doesn't diverge at the origin, and even for many that diverge. It is especially important that the condition (121) is fulfilled also by the Coulomb potential for which the potential energy is $V(r) \propto 1/r$ (i.e. which diverges at the origin). So, we can neglect some terms in singular DE (116) for $r \to 0$ and obtain the equation

$$\frac{\mathrm{d}^2\chi}{\mathrm{d}r^2} - \frac{l(l+1)}{r^2}\chi = 0 \tag{122}$$

We try to find a solution of the equation in the neighbourhood of the singular point $r = 0$ in a form of the power series

$$\chi(r) = r^{\alpha}\sum_{k=0}^{\infty} d_k\, r^k$$

(The factor r^{α} stands there for a case if the series started from the term r^{α} and it is practical to be prepared to that in advance.) It we limit our treatment to really very small r, then it will be sufficient to consider the lowest term of the series: $\chi = d_0 r^{\alpha}$. We substitute it to equation (122) and obtain

$$\alpha(\alpha-1)\, d_0\, r^{\alpha-2} - \frac{l(l+1)}{r^2}\, d_0\, r^{\alpha} = 0 \tag{123}$$

and subsequently

$$\alpha(\alpha-1) = l(l+1)$$

which is a quadratic equation for α (at a given quantum number l). Its solutions are

$$\alpha_1 = l+1 \quad\Rightarrow\quad \chi \propto r^{l+1}\,, \qquad\qquad \alpha_2 = -l \quad\Rightarrow\quad \chi \propto \frac{1}{r^l}$$

We reject the second of these solutions as unphysical, because such a radial wave function $\chi(r)/r$ would certainly be unbounded at $r \to 0$. Therefore

$$\chi = c_1\, r^{l+1} + c_2\,\frac{1}{r^l}\,, \qquad \text{kde } c_2 = 0$$

Solving the Equation (116) for Large r. Assume that the potential energy V of the particle under study fulfills

$$\lim_{r\to\infty} V(r) = V_{\infty} = \text{const} \tag{124}$$

Then we can neglect the term with $1/r^2$ in DE (116) and the equation thus take the form

$$\frac{\mathrm{d}^2\chi}{\mathrm{d}r^2} + \frac{2m}{\hbar^2}(E - V_{\infty})\,\chi = 0 \tag{125}$$

This is a linear ordinary DE of the 2^{nd} order with constant coefficients, thus very simple in its form. During your studies, you probably encountered the equation of this form many times, for the first time in the study of the linear harmonic oscillator (where the independent variable was the time t, not the distance r). Thus we know that it is easy to solve analytically. However, we must distinguish the cases in this equation:

- $E - V_{\infty} > 0$; The equation is then really the same, from a mathematical point of view, as the one for a harmonic oscillator, with oscillating solutions, i.e. functions of the type $\cos kr$, $\sin kr$, or, equivalently, $e^{\pm \mathrm{i}kr}$. k is a real number: $k = \sqrt{2m(E-V_{\infty})/\hbar^2}$. These are functions that definitely do not exhibit localisation to some finite spatial domain. Therefore, they correspond to the above mentioned scattering states [see above eq. (118)]. Thus we have just found out the ***stattering states have energies greater than V_{∞}***. Although the scattering states cannot be normalised, they are not sheerly unphysical and they are of great importance for physics. In this course, however, we will not address them in details, as we have stated it above eq. (118).

- $E - V_\infty < 0$; equation (125) then has solutions in a form of non-oscillating exponential functions. We are going to address these solutions and as we will soon see, one of them is physical and corresponds to a bound state.

Denote

$$\boxed{\varkappa^2 = \frac{2m}{\hbar^2}(V_\infty - E)} > 0 \tag{126}$$

Then

$$\frac{\mathrm{d}^2\chi}{\mathrm{d}r^2} - \varkappa^2\chi = 0 \quad \Longrightarrow \quad \chi = B_1 e^{-\varkappa r} + B_2 \underbrace{e^{\varkappa r}}_{\text{diverguje}} \quad \Longrightarrow \quad B_2 = 0 \quad \Longrightarrow \quad \chi = B_1 e^{-\varkappa r}$$

The solution of the form $e^{\varkappa r}$, where $\varkappa > 0$, is unphysical, because for large r it would yield infinite values of the wave function. Hence we dropped it. But the solution $e^{-\varkappa r}$ gives a bound wave function which is even localised in the vicinity of the potential centre. Therefore, this solution corresponds to a bound state. So, with the help of (126) we found out that ***bound states have energies smaller***[13] ***than*** $\mathbf{V_\infty}$.

A classical analogue to our system under study is a body in the gravitational field of a mass centre: for instance, a planet or a comet in the gravitational field of the Sun, as an (almost) motionless mass centre. If the mechanical (kinetic plus potential) energy the moving body is negative, the body does not leave the space of the solar system and it orbits an ellipse around the Sun. This is the case of planets as well as the periodically returning cometes. If, however, the mechanical energy of the body is positive, it keeps moving away from the solar system – will not return to it. This is the case of, e.g., space probes Pioneer 10 and Pioneer 11 to which the so-called escape velocity was given, sufficient to untie them from the Sun. This is the classical analogue to a scattering state.

Summary. The radial function $R = \chi(r)/r$ of a bound state in the problem of a spherically symmetric potential has the asymptotic behaviour *approximately* as follows:

$$\boxed{R_l(r) \propto \begin{cases} r^l\,, & r \to 0 \\ \dfrac{1}{r} e^{-\varkappa r}\,, & r \to \infty \end{cases}} \tag{127}$$

where $\varkappa$ is given by formula (126). In deriving these results, we used the assumptions that the potential energy $V(r)$ around the origin fulfills the condition $\lim_{r\to 0}\left[r^2 V(r)\right] = 0$ and that at large distances from the centre $\lim_{r\to\infty} V(r) = V_\infty = \text{const.}$ $l \in \{0, 1, 2, \dots\}$ is the quantum number following from solving the angular part of the whole problem formulated by equation (100). We added the quantum number l as an index to the radial wave function because the function depends on it, as it follows from the above said. To determine the radial functions completely, we must, of course, specify a particular form of the potential energy $V(r)$. We will then discover that, apart from the quantum number l, the radial function will depend on another quantum number.

It is highlighted above that the asymptotics (127) is only approximate, in a kind of framework sense. We will see this later in the case of the hydrogen atom, and we may be able to wonder why the above considerations about asymptotics at singular points were not entirely consistent.

The whole wave function (104), since it is also to have a probabilistic interpretation, must be normalized to 1, which is expressed by equation (118). And since the spherical harmonic functions are also normalized to 1 [see (93) for $l = l'$, $m = m'$], the radial wave functions must also be normalized to 1:

$$\int_0^\infty R_l^2(r)\, r^2\, \mathrm{d}r = 1 \tag{128}$$

We have already expressed this equivalently by condition (120) for the $chi(r)$ function.

---

[13] The usual case is $V_\infty = 0$. For example, the Coulombic potential energy satisfies $\lim_{r\to\infty} q_1 q_2/(4\pi\varepsilon_0\, r) = 0$. Therefore we often say that the bound states have negative energies and the scattering states positive energies. Let us, however, work with a general V_∞ in the case of a general spherically symmetric potential. V_∞.

5.2 The Hydrogen Atom and Like Ions

This is a particular example of a spherically symmetric force field or a potential; hence we will discuss this whole section as an **exercise**.

We now make the potential energy $V(r)$ specific:

$$V(r) = -\frac{1}{4\pi\varepsilon_0}\frac{Ze^2}{r} \tag{129}$$

Thus, we take the value of Ze as the nucleus charge; the case of $Z = 1$ would correspond to the hydrogen atom, the case of $Z = 2$ to the kation He$^+$ of the helium atom, to the kation Li^{2+}, etc. To save writing, let's introduce the notation [2]

$$e'^2 = \frac{e^2}{4\pi\varepsilon_0} \tag{130}$$

5.2.1 The Eigenenergies and Wavefunctions

The task to be solved is equation (116) for negative energies, for those correspond to bound states. For the actual potential energy, the constant $\varkappa$ [see (126)] will be expressed as

$$\varkappa = \sqrt{\frac{2m}{\hbar^2}(-E)} \tag{131}$$

as $V_\infty = 0$ [see (124)]. The constant $\varkappa$ is an inverse length, as for its physical dimension. Therefore, it will be useful for us to introduce a dimensionless distance from the nucleus:

$$\boxed{\rho = 2\varkappa r} = \sqrt{-\frac{8m}{\hbar^2}E}\, r \tag{132}$$

We will now rewrite the DE (116) using the dimensionless distance ρ. Instead of $\chi = \chi(r)$, it then becomes convenient to use some other, differently denoted function;[14] let it be $\tilde{\chi}(\rho)$:

$$\tilde{\chi}(\rho) = \chi(r) \tag{133}$$

We express

$$\frac{\mathrm{d}\chi}{\mathrm{d}r} = \frac{\mathrm{d}\tilde{\chi}}{\mathrm{d}\rho}\frac{\mathrm{d}\rho}{\mathrm{d}r} = 2\varkappa\frac{\mathrm{d}\tilde{\chi}}{\mathrm{d}\rho}\,, \qquad \frac{\mathrm{d}^2\chi}{\mathrm{d}r^2} = 4\varkappa^2\frac{\mathrm{d}^2\tilde{\chi}}{\mathrm{d}\rho^2}$$

and obtain the DE (after dividing by the constant $4\varkappa^2$)

$$\frac{\mathrm{d}^2\tilde{\chi}}{\mathrm{d}\rho^2} + \left[\frac{2mE}{\hbar^2 4\varkappa^2} + \frac{2m}{\hbar^2}\frac{Ze'^2}{2\varkappa\rho} - \frac{l(l+1)}{\rho^2}\right]\tilde{\chi} = 0$$

We will use the expression for $\varkappa$ according to (131) to cast the equation in the form

$$\frac{\mathrm{d}^2\tilde{\chi}}{\mathrm{d}\rho^2} + \left[-\frac{1}{4} - \frac{\varkappa}{E}\frac{Ze'^2}{2\rho} - \frac{l(l+1)}{\rho^2}\right]\tilde{\chi} = 0$$

This motivates us to introduce also the constant

$$\beta = -\frac{Ze'^2\varkappa}{2E} \tag{134}$$

the value of which is positive (as $E < 0$ and which can be, using (131), expressed also in the form

$$\beta = \sqrt{-\frac{m\,Z^2\,e'^4}{2E\hbar^2}} \tag{135}$$

---

[14] In similar cases, a different notation is usually omitted. Strictly, however, a function expressed using another variable should be labelled differently. So at least now we're doing it here, though not elsewhere. This depends on the circumstances when it is appropriate and, conversely, when a different notation would be a just an unnecessary complication.

(See also [2], where this constant is introduced for $Z = 1$.) So, we finally arrive at a compact DE

$$\frac{\mathrm{d}^2\tilde{\chi}}{\mathrm{d}\rho^2} + \left[-\frac{1}{4} + \frac{\beta}{\rho} - \frac{l(l+1)}{\rho^2}\right]\tilde{\chi} = 0 \tag{136}$$

According to the general[15] result (127) and using (115), we already know how its solutions behave at the singular points:

$$\chi(r) = \tilde{\chi}(\rho) \propto \begin{cases} r^{l+1}\,, & r \to 0 \\ e^{-\varkappa r}\,, & r \to \infty \end{cases} \propto \begin{cases} \rho^{l+1}\,, & \rho \to 0 \\ e^{-\rho/2}\,, & \rho \to \infty \end{cases} \tag{137}$$

Of course, we want to find the form of the solution also elsewhere, not only in these extreme points. So we can look for it in the form

$$\boxed{\tilde{\chi}(\rho) = \rho^{l+1}e^{-\rho/2}\,v(\rho)} \tag{138}$$

where $v(\rho)$ is some unknown function to be determined. Substitute this Ansatz into DE (136). After rather lengthy but simple manipulations, we obtain a differential equation for $v(\rho)$:

$$\rho\,\frac{\mathrm{d}^2 v}{\mathrm{d}\rho^2} + [2(l+1) - \rho]\,\frac{\mathrm{d}v}{\mathrm{d}\rho} + (\beta - l - 1)v = 0 \tag{139}$$

which is no more a singular DE. We look for the solution to it in a form of a power series

$$\boxed{v(\rho) = \sum_{k=0}^{\infty} c_k\rho^k} \tag{140}$$

This yields expressions for the derivatives:

$$v'(\rho) = \sum_{k=1}^{\infty} k\,c_k\rho^{k-1}\,, \quad v''(\rho) = \sum_{k=2}^{\infty} k(k-1)\,c_k\rho^{k-2}$$

When we substiture all this into the last DE, we get

$$\sum_{k=2}^{\infty} k(k-1)c_k\rho^{k-1} + [2(l+1) - \rho]\sum_{k=1}^{\infty} kc_k\rho^{k-1} + (\beta - l - 1)\sum_{k=0}^{\infty} c_k\rho^k = 0$$

We now reexpress this equation to the form

$$\sum_{k=0}^{\infty}[\ \ \text{something}_k\ \ ]\,\rho^k = 0$$

This will be possible if we properly shift the summation indices. The resulting equation will be

$$\sum_{k=0}^{\infty}\left\{c_{k+1}[k(k+1) + 2(l+1)(k+1)] - c_k(k + l + 1 - \beta)\right\}\rho^k = 0 \tag{141}$$

For this must hold for any (non-negative) ρ, the coefficient multiplying the expression ρ^k has to be zero for each index k. From this finding, we obtain the recurrent formula

$$\boxed{c_{k+1} = \frac{k + l + 1 - \beta}{(k+1)(2l + 2 + k)}\,c_k} \tag{142}$$

So, we succeeded in finding a way how to complete the solution of the DE (139), and by this also of the whole problem of the hydrogen atom. But will this solution automatically be a physical one? The coefficient c_k determined are to

---

[15] although that one in some cases is not fully accurate, but it will not matter

be substituted to series (140). Let us explore if the radial wave function with the $v(\rho)$ determined in the above way tends to zero for $\rho \to \infty$ as it ought to. For this purpose, let us first see how the series (140) converges.

$$\frac{c_{k+1}}{c_k} = [k \to \infty] = \frac{1}{k}$$

We know that the same ratio is found also for the exponential function e^ρ:

$$e^\rho = \sum_{k=0}^{\infty} \underbrace{\frac{1}{k!}}_{b_k} \rho^k \quad \Rightarrow \quad \frac{b_{k+1}}{b_k} = [k \to \infty] = \frac{1}{k}$$

It means that the function $v(\rho)$ will behave as an exponential function for large ρ:

$$v(\rho) \propto e^\rho \quad \text{pre } \rho \to \infty$$

Now look at Ansatz (138). We see that for the $v(\rho)$ just examined

$$\tilde{\chi}(\rho) \propto \rho^{l+1} e^{\rho/2} \quad \text{for } \rho \to \infty$$

i.e. it diverges. Thus, the determined solution is unphysical. However, the hope of finding a physical solution will come to life when we realise this: if the coefficient c_k is zero at certain k, then, according to the recurrent formula (142), also all higher coefficients vanish. By this, the infinite series (140) becomes a polynomial and consequently its values will be quantifiable for arbitrarily large ρ (it will be convergent, in other words). Denote its degress as n_r. Let us examine its asymptotics:

$$\lim_{\rho\to\infty} \tilde{\chi}(\rho) = \lim_{\rho\to\infty} \rho^{l+1} e^{-\rho/2} v(\rho) \propto \lim_{\rho\to\infty} \rho^{l+1} e^{-\rho/2} \rho^{n_r} = 0$$

Thus, such a solution is physical. So, we want

$$c_{n_r+1} = 0\,, \quad c_{n_r} \neq 0 \tag{143}$$

Recurrent formula (142) applied to $k = n_r$ then yields that

$$\beta = n_r + l + 1 \tag{144}$$

Looking at the definction of the constant β [formula (134)], we already see that the ***eigenenergy*** E will depend on the indices, that is, it will be *quantised*. We will make this more specifice later. As we also see, we got β as a natural number (a positive integer), since l is a non-negative integer and the degree of the polynomial is

$$n_r \in \{0, 1, 2, \dots\}$$

Therefore, we have a good motivation to change the notation of β to n:

$$\beta = n \ \in \mathbb{N} \tag{145}$$

Already at this point we can, using (135) and (130), express the eigenenergies of the hydrogen atom or its like ion[16]:

$$\boxed{\boxed{E_n = -\frac{m}{\hbar^2}\left(\frac{e^2}{4\pi\varepsilon_0}\right)^2 \frac{1}{2}\frac{Z^2}{n^2}}}\,, \quad n = 1, 2, 3, \dots \tag{147}$$

---

[16] In Hartree atomic units, $e = m = \hbar = \dfrac{1}{4\pi\varepsilon_0} = 1$. Thus, in these units $e'^2 = 1$ and

$$\beta = \sqrt{-\frac{Z^2}{2E}}\,, \qquad \boxed{\boxed{E_n = -\frac{1}{2}\frac{Z^2}{n^2}}}\,, \quad n = 1, 2, 3, \dots\,, \qquad \varkappa_n = \frac{Z}{n}\,, \qquad \boxed{\rho = 2\,\frac{Z}{n}\,r} \tag{146}$$

which are often more practical and easier to remember expression; *cf.* also (151).

We have not discussed this expression in connection with the contributions of the kinetic and potential energy, but you can read something about it in Appendix B, for example.

The analysis done above [see (144) and (145)] implies that a certain value of the quantum number n can be obtained from several combinations of the numbers n_r and l. Different n_r values imply different degrees of the polynomials, so also different eigenfunctions. Similarly, different values of l imply different spherical functions, thus also different eigenfunction too. In other words, the eigenenergy E will be *degenerate* (except from the case of $n = 1$, as we will see). So, the indices are coupled according to

$$n = n_r + l + 1\,, \qquad n_r,\, l \in \{0, 1, 2, \dots\} \tag{148}$$

that is (and for completeness we add also the quantum number m which should not be confused with the mass denoted by the same symbol)

$$\boxed{\boxed{n \in \{1, 2, 3, \dots\}\,, \qquad l \in \{0, 1, \dots, n-1\}\,, \qquad m \in \{-l,\, -l+1,\, \dots,\, l\}}}\,, \qquad \boxed{n_r = n - l - 1} \tag{149}$$

We now summarise what we have learnt up till now about the eigenfunctions and let us try to find expression for them. We have found out that the radial function will depend on the quantum number l; see (127) in the section on the general spherically symmetric potential energy. Then, in the present section, we discovered that it will depend also on the quantum number n. Therefore, we will write $R(r) = R_{nl}(r)$. In order not to be lost in the sequence of the various substitutions and steps, we first recall and run through the subsequent expressions (104), (115), (132), (133), (138), (140), (142), (148) which lead us to the finding a mathematical expression for the wavefunction. By putting together the appropriate formulae, we get

$$\boxed{\boxed{\psi_{nlm}(r, \vartheta, \varphi) = R_{nl}(r) Y_l^m(\vartheta, \varphi)}} \tag{150}$$

The index n is called the ***principal*** quantum number; it determines the energy of the eigenstate. The index l is called the ***orbital*** quantum number; it determines the magnitude of the angular momentum. The index m is called the ***magnetic*** quantum number; it determines the projection of the angular momentum on the z axis (i.e. the z-component). The scaling constant $\varkappa$ defined by formula (131) depends on the eigenenergy; this dependence can be converted on the dependence on the principal quantum number:

$$\varkappa = \boxed{\varkappa_n = \frac{me'^2}{\hbar^2}\,\frac{Z}{n}}\,, \qquad \boxed{\rho = 2\varkappa_n\, r} \tag{151}$$

The radial function will then be expressed as

$$R_{nl}(r) \;=\; \frac{\chi(r)}{r} \;=\; 2\varkappa_n\,\frac{\tilde{\chi}(\rho)}{\rho} \;=\; 2\varkappa_n\,\frac{1}{\rho}\;\rho^{l+1}e^{-\rho/2}\,v(\rho) \;=\; 2\varkappa_n\,\rho^l e^{-\rho/2}\sum_{k=0}^{n_r} c_k\rho^k$$

t. j.

$$\boxed{R_{nl}(r) = 2\varkappa_n\, e^{-\rho/2}\sum_{k=0}^{n_r} c_k\rho^{l+k}} \tag{152}$$

The above exposition implies that the radial functions $R_{nl}(r)$ and, of course, the auxiliary radial functions $\chi(r)$ a $\tilde{\chi}(\rho)$, can be determined to be ***real***.

Degeneracy of the Levels. Using formulae (147), (149) and (150) we can see that the level $n = 2$ is 4 times degenerate[17] Generally, the n^{th} level is n^2 times degenerate, i.e. there are n^2 mutually linearly independent eigenfunctions corresponding to the energy level E_n. (The number n^2 is not difficult to calculate. you just have to think about it and know how to add an arithmetic sequence.) We also see, whence the degeneracy comes from: a part of it comes originates in the sperical harmonic functions Y_l^m, in which for each l we have $2l + 1$ different values of m. This degeneracy is related to the spherical symetry of the problem under study [2, 3]. The other part originates in the radial component of the solution, as for one principal quantum number n, we have n different magnitudes of the angular momentum, thus n different indices l.

---

[17] In these considerations, we do not take into account the spin degeneracy. We only consider the eigenvalues and eigenfunctions of Hamiltonian (101) with the potential energy (129).

5.2.2 Calculations of the Radial Wavefunctions

In principle, we have already determined the radial wavefunctions by the expression (152) and by the recurrent formula (142). There are several reasons, why this is not a final form yet. For example, we have not yet determined the coefficients c_k; we only know that we will have to apply the recurrent formula and surely also the normalisation condition. We are now going to do these calculations as an exercise. For this purpose, we first write in one place all needed formulae and quantities in a convenient form.

The auxiliary constant $\varkappa_n$ [see (151)] has a dimension of inverse length. Therefore, it can be very conveniently expressed in terms of the first Bohr radius a_{B} of the hydrogen atom,

$$a_{\mathrm{B}} = \frac{\hbar^2 4\pi\varepsilon_0}{me^2} = 0{,}5291772083\,.\,10^{-10}\,\mathrm{m} \tag{153}$$

or, here even more conveniently, with the help of the generalised first Bohr radius a_Z, which characterises an hydrogen-like ion with Z protons:

$$a_Z = \frac{a_{\mathrm{B}}}{Z} \tag{154}$$

The constant $\varkappa_n$ can then be expressed as

$$\varkappa_n = \varkappa_n(Z) = \frac{1}{na_Z} \tag{155}$$

where we displayed its dependence on the proton number Z. Further formulae needed are

$$\rho = 2\varkappa_n r \tag{156}$$

$$\tilde{\chi}_{nl}(\rho) \equiv \chi_{nl}(r) = rR_{nl}(r) \tag{157}$$

We are going to search the radial functions on the space of real functions; we have found out that it is possible to do so. We will express the normalisation condition, for instance specifically for the function $\tilde{\chi}_{nl}(\rho)$, according to (120) by the equation

$$\boxed{\frac{1}{2\varkappa_n}\int_0^\infty \tilde{\chi}_{nl}^2(\rho)\,\mathrm{d}\rho = 1} \tag{158}$$

which is perhaps a most practical one to calculate the coefficients using the normalisation condition. We will also need the recurrent formula (142), which we now rewrite using n instead of β:

$$c_{k+1} = \frac{k+l+1-n}{(k+1)(2l+2+k)}\,c_k \tag{159}$$

We will do our following calculations of the coefficients c_k with the use of the functions $\tilde{\chi}_{nl}(\rho)$. Thus, according to (157), (156) and (152) we express

$$\tilde{\chi}_{nl}(\rho) = e^{-\rho/2}\sum_{k=0}^{n_r} c_k \rho^{l+k+1} \tag{160}$$

Wavefunction for $n = 1$. According to (149), $l = n_r = 0$. Using this we obtain $\tilde{\chi} = \rho e^{-\rho/2}c_0$. Using the normalisation condition, we get $c_0 = \sqrt{\varkappa_1}$. We note in passing that it is worth to calculate integrals of the form

$$I_n \equiv \int_0^\infty x^n e^{-x}\,\mathrm{d}x = n! \tag{161}$$

We arrive at the results

$$R_{10}(r) = \left(\frac{1}{a_Z}\right)^{3/2} 2\,e^{-r/a_Z} \tag{162}$$

Wavefunction for $n = 2, l = 0$. In this case $n_r = 1$; thus

$$\psi_{2,0,0}(r,\vartheta,\varphi) = Y_0^0(\vartheta,\varphi)\;2\varkappa_2\,e^{-\rho/2}\,(c_0 + c_1\rho) \tag{163}$$

whereas according to eq. (142) $c_1 = -c_0/2$ etc. (An exercise.)

Wavefunctions for $n = 2, l = 1$. In this case $n_r = 0$; thus

$$\psi_{2,1,m}(r,\vartheta,\varphi) = Y_1^m(\vartheta,\varphi)\, 2\varkappa_2\, e^{-\rho/2}\, c_0\, \rho \tag{164}$$

atď. (An exercise.)

Listing of the Lowest Radial Functions for a Hydrogen-like Ion with the proton number Z. (This includes also the hydrogen atom itself as the particular case of $Z = 1$.)

$$R_{10}(r) = \left(\frac{1}{a_Z}\right)^{3/2} 2\, e^{-r/a_Z} \tag{165a}$$

$$R_{20}(r) = \left(\frac{1}{2a_Z}\right)^{3/2} \left(2 - \frac{r}{a_Z}\right) e^{-r/(2a_Z)} \tag{165b}$$

$$R_{21}(r) = \left(\frac{1}{2a_Z}\right)^{3/2} \left(\frac{r}{a_Z\sqrt{3}}\right) e^{-r/(2a_Z)} \tag{165c}$$

$$R_{30}(r) = \left(\frac{1}{3a_Z}\right)^{3/2} 2 \left[1 - \frac{2}{3}\frac{r}{a_Z} + \frac{2}{27}\left(\frac{r}{a_Z}\right)^2\right] e^{-r/(3a_Z)} \tag{165d}$$

$$R_{31}(r) = \left(\frac{1}{3a_Z}\right)^{3/2} \frac{4}{9}\sqrt{2} \left(1 - \frac{1}{6}\frac{r}{a_Z}\right) \frac{r}{a_Z}\, e^{-r/(3a_Z)} \tag{165e}$$

$$R_{32}(r) = \left(\frac{1}{3a_Z}\right)^{3/2} \frac{2}{27}\sqrt{\frac{2}{5}} \left(\frac{r}{a_Z}\right)^2 e^{-r/(3a_Z)} \tag{165f}$$

6 Approximate Methods of Solving the Stationary Schrödinger Equation for Bound States

Thus, the task is to solve the problem [1]

$$\hat{H} u_n(\vec{r}) = E_n u_n(\vec{r}) \tag{166}$$

where $\hat{H}$ is a time-independent Hamiltonian of a system. Now we use symbols $u_n(\vec{r})$ for its eigenfunctions. To keep the notation simple, we limit our treatment to one-particle systems; the principle of the method can, however, be directly generalised to many-particle systems.

6.1 The Variational Method

Definition 7: *A function $f(\vec{r})$ is called quadratically integrable if*

$$\int |f(\vec{r})|^2\, \mathrm{d}^3 r < \infty$$

(The integration is taken over the entire space.) In other words, f is a normalisable function.

Theorem 8 (the variational principle): *Assume that $\hat{H}$ is a hermitian operator with a discrete spectrum with the lowest eigenvalue being E_0. Let the eigenfunctions of the operator $\hat{H}$ form a complete orthonormal system (set). The following inequality then holds for arbitrary quadratically integrable function $f(\vec{r})$*

$$\frac{\displaystyle\int f^*(\vec{r})\hat{H} f(\vec{r})\, \mathrm{d}^3 r}{\displaystyle\int f^*(\vec{r}) f(\vec{r})\, \mathrm{d}^3 r} \geq E_0 \tag{167}$$

Proof: In accordance with (166), we denote the eigenfunctions and eigenvalues of the operator $\hat{H}$ as $u_n(\vec{r})$ and E_n. The assumed *completeness* of the set of the functions $u_n(\vec{r})$ in other words mean that any function can be expressed as their linear combination. Exactly in this way we do with the function $f(\vec{r})$: rozvinieme ju podľa sústavy vlastných funkcií operátora $\hat{H}$. we will expand it in a series By convention, we will be using the indexing of the eigenfunctions and eigenvalues so that they would start from $n = 0$ (the lowest eigenvalue) and $E_0 \leq E_1 \leq E_2 \leq \ldots$. Individual eigenvalues may also be degenerate, i.e. some of them may be equal each other. The expansion under consideration then reads

$$f(\vec{r}) = \sum_{n=0}^{\infty} c_n u_n(\vec{r}) \tag{168}$$

The ***orthonormality*** of the eigenfunctions of the operator $\hat{H}$ assumed by the theorem is expressed as

$$\int u_m^*(\vec{r}) u_n(\vec{r})\, \mathrm{d}^3 r = \delta_{mn} \tag{169}$$

We can similarly expand

$$f^*(\vec{r}) = \sum_{n=0}^{\infty} c_n^* u_n^*(\vec{r})$$

Substitute now these expressions into the left-hand side of the variational principle (167). Calculate

$$\int f^*(\vec{r}) \hat{H} f(\vec{r})\, \mathrm{d}^3 r = \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} c_m^* c_n \int \mathrm{d}r^3\, u_m^*(\vec{r}) \hat{H} u_n(\vec{r}) = [\text{see } (166)] = \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} c_m^* c_n \int \mathrm{d}r^3\, u_m^*(\vec{r}) E_n u_n(\vec{r}) =$$

$$= [\text{the orthonormality}] = \sum_{n=0}^{\infty} |c_n|^2 E_n$$

Now we can immediatelly write also the results for the denominator of the variational principle (167):

$$\int f^*(\vec{r}) f(\vec{r})\, \mathrm{d}^3 r = \sum_{n=0}^{\infty} |c_n|^2$$

It is obvious that

$$\begin{gathered} E_0 \leq E_0 \\ E_0 \leq E_1 \\ \ldots\ldots \\ E_0 \leq E_n \\ \ldots\ldots \end{gathered}$$

We multiply each of these inequalities by its respective $|c_n|^2$. We then add the inequalities together. In this way we get

$$E_0 \sum_{n=0}^{\infty} |c_n|^2 \leq \sum_{n=0}^{\infty} E_n |c_n|^2$$

i.e.

$$E_0 \leq \frac{\displaystyle\int f^*(\vec{r}) \hat{H} f(\vec{r})\, \mathrm{d}^3 r}{\displaystyle\int f^*(\vec{r}) f(\vec{r})\, \mathrm{d}^3 r}$$

what needed to be proved.

Remark: If we were interested how to calculate (at least formally) the coefficients c_n in the proof just done, we would take equation (168), multiplied it by the function $u_m^*(\vec{r})$ and integrate over the integration domain. Using the orthonormality (169), we would obtain the expression

$$c_n = \int u_n^*(\vec{r}) f(\vec{r})\, \mathrm{d}^3 r$$

Coefficients are being expressed in this way very often in quantum physics and especially in studies of electronic structure,

How to determine the eigenfunctions and eigenvalues using the variational principle (167)? The principle itself does not provide any result for the functions nd values. We can, however, construct a trial function $f(\vec{r}, \alpha_1, \alpha_2, \ldots, \alpha_p) \equiv f(\vec{r}, \alpha)$, in which $\alpha \equiv (\alpha_1,\, \alpha_2,\, \ldots,\, \alpha_p)$ are some parameters. We substitute the function into the left-hand side of the variational principle (167). By doing this, we get a function F which depends on the chosen parameters:

$$\frac{\displaystyle\int f^*(\vec{r};\alpha)\hat{H} f(\vec{r};\alpha)\,\mathrm{d}^3 r}{\displaystyle\int f^*(\vec{r};\alpha) f(\vec{r};\alpha)\,\mathrm{d}^3 r} = F(\alpha_1, \ldots, \alpha_p) \tag{170}$$

We determine such values of the parameters α_j, at which the function f has its global minimum. Let us denote them by $\tilde{\alpha}_1, \ldots, \tilde{\alpha}_p$. In this way, we managed to approach the unknown exact lowest eigenenergy E_0 as much as possible for the proposed form of the trial function. We will declare that the lowest of the values F we were able to determine be an approximation to the ground-state energy. It is an upper estimate for the exact energy. We will declare that the trial function (the one with the optimal values of the parameters) is an approximation to the ground-state wavefunction.

Although the above method may seem rather crude, it can give excellent results with a suitably chosen trial function and a sufficient number of parameters. Since we usually cannot find exact solutions to problem (166) (except for a few cases, such as the hydrogen atom), we must use approximate methods of solution. It is the variation method that is used very often, which we will talk about later.

The variation method can also be used to find some of the lowest excited states of the system (i.e. those that have their eigenenergies greater than E_0). We will show this for the case when the ground state is non-degenerate (which is a common situation).

Theorem 9: *Assume that $\hat{H}$ is an operator as in Theorem 8 and that, in addition, its lowest eigenvalue E_0 (the ground state) is non-degenarate and its respective eigenfunction is $u_0(\vec{r})$. Let $g(\vec{r}$ be a quadratically integrable function orthogonal to $_0(\vec{r})$. The inequality*

$$\frac{\displaystyle\int g^*(\vec{r})\hat{H} g(\vec{r})\,\mathrm{d}^3 r}{\displaystyle\int g^*(\vec{r}) g(\vec{r})\,\mathrm{d}^3 r} \geq E_1 \tag{171}$$

then holds, in which E_1 is the energy of the first excited state (and it can also be degenerate)

Proof: Similarly as in the previous theorem, we expand the function under consideration into a series of the eigenfunctions of the operator $\hat{H}$: $\hat{H}$:

$$g(\vec{r}) = \sum_{n=1}^{\infty} c_n u_n(\vec{r})$$

This time, however, we have omitted the eigenfunction $u_0(\vec{r})$ of the lowest level from the linear combination, because according to the assumption, $g(\vec{r})$ should be orthogonal to $u_0(\vec{r})$. (We can easily be convinced explicitely that if we include the term with $c_0 u_0(\vec{r})$ into the summation, the orthogonality

$$\int g^*(\vec{r})\, u_0(\vec{r})\,\mathrm{d}^3 r = 0$$

would not apply.) We proceed similarly to the previous proof, but with the difference that we write a sequence of non-strict inequalities.

$$E_1 \leq E_n \quad \text{pre } n \geq 1$$

and again, we apply the summation only from the index 1. In this way we arrive at the inequality (171) of Theorem 9, what needed to be proved.

Theorem 9 applied to the Hamiltonian operator thus allows us in principle to find an upper estimate of the eigenenergy of the first excited level and the approximate wave unction of this level (which can also be degenerate). We need to optimise the trial function for this calculation, similarly as in the case of Theorem 8. According to Theorem 9, we construct the function to be orthogonal to the ground state. After an optimisation, the trial function will become an approximate eigen-wavefunction of the first excited state. Thus, before calculating the excited state, we need – at least in principle – to have a determined wave function of the ground state (using a variational caclulation vased on Theorem 8 for example). Since the wave function of the ground state is usually only approximately known, the calculation of the excited state in this way on the basis of the knowledge of the approximate eigenstate will generally be less accurate than the calculation of the ground state. And it can be seen that similarly we could look for higher and higher levels (but increasingly less accurate).

However, the specific implementation of the variational method is such that we calculate the ground state and also a certain number of excited states at once, by searching for the whole system of eigenvectors and values of the matrix (we will talk about this in section 6.1.1). Then we can achieve that the accuracy of determining the excited states will be similar for a given Hamiltonian as the accuracy of determining the ground state (and at the same time such as the numerical diagonalization method used is able to achieve). However, as we will learn later, in the search for excited states of many-electron atoms or molecules, we also encounter principally different fundamental difficulties, and these in particular will impair the accuracy of the results.

Even ***degeneracy*** of any of the levels ***is not an obstacle*** to the application of the variational principle. We have formulated Theorem 9 for the case of non-degenerate ground state for the sake of brevity and greater clarity only.

In Theorems 8 and 9, we assumed that the spectrum of the operator $\hat{H}$ is discrete and that the corresponding eigenfunctions form a complete set of functions. However, even this (absolute) *completeness is not necessary*; if we want to find the wave functions and energies corresponding to the discrete part of the spectrum by a variational method (that is bound states, i.e. spatially localized and bounded eigenfunctions), then it is sufficient that the set of functions $u_n(\vec{r})$ allows to expand in series (168) any spatially localized function $f(\vec{r})$, because we don't even want to find another one.

Theorems 8 and 9 show us that stationary SchE can be understood as equivalent to the variational principle. Not only is this an important theoretical knowledge, but it gives us certainty in practical calculations that we cannot "shoot under (the target)" when calculating the energy.

6.1.1 An often used Version of the Variational Method

The trial function is most frequently (especially in numerical calculations on computers) being seached in the form

$$f(\vec{r};\alpha) = \sum_{i=1}^{p} \alpha_i f_i(\vec{r}) \tag{172}$$

where $f_i(\vec{r})$ are some known functions. We used a shortcut notation for the set of the variational parameters on the left-hand side (LHS): $\alpha \equiv (\alpha_1, \alpha_2, \ldots, \alpha_p)$. If some (usually unknown) function is expanded in a series of known functions, the sequence of the functions [in this case $f_i(\vec{r})$] is called a basis set, or briefly a *basis*. We should now to minimise the function $F(\alpha)$ defined by fraction (170), that is, to determine the optimal parameters $\tilde{\alpha}$, for which the function $F(\alpha)$ acquires its minimum value. Substituting for $f(\vec{r};\alpha)$ gives

$$F(\alpha) = \frac{\displaystyle\sum_i \sum_j \alpha_i^* \, H_{ij} \, \alpha_j}{\displaystyle\sum_i \sum_j \alpha_i^* \, S_{ij} \, \alpha_j} \tag{173}$$

where

$$H_{ij} = \int f_i^*(\vec{r}) \, \hat{H} \, f_j(\vec{r}) \, \mathrm{d}^3 r \tag{174}$$

$$S_{ij} = \int f_i^*(\vec{r}) f_j(\vec{r}) \, \mathrm{d}^3 r \tag{175}$$

H_{ij} are the matrix elements of the Hamiltonian in the given basis. S_{ij} are the overlap integrals, i.e. scalar products. Hence, the basis functions f_i neeed not be normalised to 1. They even need not be (mutually) orthogonal. (If they are, then calculations with them are simpler, od course.) Since we are in quantum mechanics, the function $f(\vec{r}, \alpha)$ should be (once the opetimal parameters have been determined) a wave function of the electron, possibly of some other particle. For this reason, we may already at this point require its normalisation to unity:

$$\int f^*(\vec{r}; \alpha) f(\vec{r}; \alpha)\, \mathrm{d}^3 r = 1 \tag{176}$$

Therefore, the denominator of fraction (173) has to be 1 for a correct function $f(\vec{r}; \alpha)$:

$$\sum_i \sum_j \alpha_i^*\, S_{ij}\, \alpha_j = 1 \tag{177}$$

However, in doing a minimisation of the function $F(\alpha)$, the trial function $f(\vec{r}, \alpha)$ is being varied in a rather arbitrary (e.g. also random) way. These variations would in general lead to violation of the condition (176). Thus, if we aimed to straightforwardly minimise $F(\alpha)$ [given by formula (173)], we would relly have to consider and to write also its denominator.[18] However, we better want to avoid this as it woult not be practical. We prefer to minimise the simplified function

$$F_{\text{čitateľ}}(\alpha) = \sum_{i=1}^{p} \sum_{j=1}^{p} \alpha_i^*\, H_{ij}\, \alpha_j \tag{178}$$

while maintaining condition (177). This mathematical task can be solved using the *Lagrange multiplier* (LM) method:[19] we define the new function

$$F(\alpha, \lambda) = F_{\text{čitateľ}}(\alpha) + \lambda \left(1 - \sum_{i,j=1}^{p} \alpha_i^*\, S_{ij}\, \alpha_j \right) \tag{179}$$

The the yet unknown constant λ is called the Lagrange multiplier. The extrema of the function $F(\alpha, \lambda)$ are determined by calculating its partial derivatives with respect to the particular arguments, which are the (in general) complex parameters α_j. It is a set of p complex variables, i.e. $2p$ real variables. We decompose each of the α_j to its real and imaginary parts: $\alpha_j = X_j + \mathrm{i} Y_j$. By doing this, we obtain the following equations (the necessary conditions for the extremum):

$$\frac{\partial F}{\partial X_1} = 0\,, \quad \frac{\partial F}{\partial Y_1} = 0\,, \quad \frac{\partial F}{\partial X_2} = 0\,, \quad \frac{\partial F}{\partial Y_2} = 0\,, \quad \ldots\ldots\ldots\,, \quad \frac{\partial F}{\partial X_p} = 0\,, \quad \frac{\partial F}{\partial Y_p} = 0 \tag{180}$$

It is shown in the short Appendix C that the set of $2p$ equations is equivalent the set of $2p$ equations

$$\frac{\partial F}{\partial \alpha_1} = 0\,, \qquad \frac{\partial F}{\partial \alpha_2} = 0\,, \qquad \ldots\ldots\ldots\,, \qquad \frac{\partial F}{\partial \alpha_p} = 0 \tag{181a}$$

$$\frac{\partial F}{\partial \alpha_1^*} = 0\,, \qquad \frac{\partial F}{\partial \alpha_2^*} = 0\,, \qquad \ldots\ldots\ldots\,, \qquad \frac{\partial F}{\partial \alpha_p^*} = 0 \tag{181b}$$

in which the partial derivatives with respect to the complex variables α_j a α_j^* are used. We express the function to be minimised by the formula

$$F(\alpha, \lambda) = \lambda + \sum_{i,j=1}^{p} \alpha_i^*\, (H_{ij} - \lambda S_{ij})\, \alpha_j \tag{182}$$

We could take λ for a variational parameter too. By doing this, we would obtain one additional equation with a partial derivative, but this equation would not yield any new information. Therefore, the first term (the LM alone)

---

[18] The denominator corrects for the possible improper normalisation of the function $f(\vec{r}, \alpha)$ and guarantees that the energy is evaluated as if the function $f(\vec{r}, \alpha)$ was normalised properly.

[19] Its principle is nicely geometrically explained form instance in [2].

is omitted in the literature [2]. By calculating the derivatives and markinh that they shall be zero, we obtain

$$\frac{\partial F}{\partial \alpha_k} = \sum_{i=1}^{p} \alpha_i^* (H_{ik} - \lambda S_{ik}) \overset{\text{let be}}{=} 0 \tag{183a}$$

$$\frac{\partial F}{\partial \alpha_k^*} = \sum_{j=1}^{p} (H_{kj} - \lambda S_{kj})\, \alpha_j \overset{\text{let be}}{=} 0 \tag{183b}$$

$$k \in \{1, 2, \dots p\}$$

The $\overset{\text{let be}}{=} 0$ equations are actually two sets, each of which contains p algebraic linear homogeneous equations with p unknowns (α_i^* alebo α_j). It is convenient to express these equations in explicit matrix forms and to put the terms with λ to their right-hand sides. For example, the second part of system (183) can be written as follows:

$$\begin{pmatrix} H_{11} & H_{12} & \dots\dots & H_{1p} \\ H_{21} & H_{22} & \dots\dots & H_{2p} \\ \dots & \dots & \dots\dots & \dots \\ H_{p1} & H_{p2} & \dots\dots & H_{pp} \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \dots \\ \alpha_p \end{pmatrix} = \lambda \begin{pmatrix} S_{11} & S_{12} & \dots\dots & S_{1p} \\ S_{21} & S_{22} & \dots\dots & S_{2p} \\ \dots & \dots & \dots\dots & \dots \\ S_{p1} & S_{p2} & \dots\dots & S_{pp} \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \dots \\ \alpha_p \end{pmatrix} \tag{184}$$

It is recommended to write also the first part and in this way to find out that by its complex conjugation [20] we obtain an equation almost the same as (184); they would differ only by λ^* in the second equation instead of λ. But this implies that (if it is difficult to understand, you really should write both the systems) $\lambda^* = \lambda$; thus the LM λ is real.[21]

Equation (184) has a form of the stationary SchE in a matrix notation using the considered basis functions $f_i(\vec{r})$; see (172). From the point of view of the linear algebra, it is a generalised equation[22] for the eigenvalues λ and the eigenvectors written as columns using the quantities α_i. (It would be the usual, i.e. not the generalised problem, if the matrix S was not there, or, equivalently, if it would be the identity matrix.) After the optimisation, the ***Lagrange multiplier λ*** acquires the meaning of the ***approximate eigenenergy***.[23] It would be exact if we used a complete basis, which would typically mean an infinit number of the basis functions, i.e. a numerically non-tractable problem. Practically, a finite number p is often sufficient to obtain highly accurate results.

The matrix H on the LHS of equation (184) is a matrix form of the Hamiltonian $\hat{H}$ in the considered basis; we often use also the notion of *matrix representation* of the Hamiltonian (in given basis). Calculation of eigenvalues and eigenvectors is called a ***diagonalisation*** of a matrix.[24] The matrix S is called the ***overlap matrix*** for it expresses how much (if at all) the basis functions overlap each other.

By moving all the terms in (184) to LHS, i.e. to arrange them as they were in (183b), we arrive at wan equation of the form *matrix* times *vector* = *zero vektor*. A non-zero solution of the corresponding linear set can only exist, if the determinant of the corresponding matrix vanishes:

$$\begin{vmatrix} H_{11} - \lambda S_{11} & H_{12} - \lambda S_{12} & \dots\dots & H_{1p} - \lambda S_{1p} \\ H_{21} - \lambda S_{21} & H_{22} - \lambda S_{22} & \dots\dots & H_{2p} - \lambda S_{2p} \\ \dots & \dots & \dots\dots & \dots \\ H_{p1} - \lambda S_{p1} & H_{p2} - \lambda S_{p2} & \dots\dots & H_{pp} - \lambda S_{pp} \end{vmatrix} = 0 \tag{185}$$

---

[20] To do this, we need to know that $H_{ik}^* = H_{ki}$ and analogously pre S_{ik}. It follows from Theorem 2 [see (12) and (174)].

[21] Note also, that even $F(\alpha)$ itself is real according to (170) for example, which is not a surprising finding, as the numerator is an expectation value of a hermitian operator. And it should be clear now that $F(\alpha, \lambda)$ is real too.

[22] We learnt the generalised eigenvalue and eigenvector problem in the subject *Počítačová fyzika*. We formulated it by equation $A \cdot x = \lambda B \cdot x$. If the matrix B is not a singular one, then this problem can be transformed to $(B^{-1} \cdot A) \cdot x = \lambda x$, that is the usual task to find the eigenveectors and eigenvalues of matrix $B^{-1} \cdot A$.

[23] It should be fully obvious in the case of an orthonormal basis; then $S_{ij} = \delta_{ij}$. However, the LM (after the optimisation) has the meaning of the approximate eigenenergy even in cases when the matix S differs from the identity matrix; to understand this, realise that we could obtain equation (184) as follows, without even using the variational principle: **(i)** We start from the SchE $\hat{H}\psi = \lambda\psi$, by which we define that the parameter λ is the eigenenergy. **(ii)** We express ψ as a linear combination of the basis functions $f_j(\vec{r})$ and substitute to the SchE. **(iii)** We multiply both sides of the equation by f_k^* from the left and integrate. **(iv)** We write the system obtained in this way in a matrix form. The system will have the form (184), i.e. the same as would be obtained from the variational principle. Hence, λ will have the meaning of the approximate eigenenergy also in the equations following from the variational principle.

[24] This calculation can mathematically be expressed as a matrix operation that results in a transformation of the Hamiltonian into a diagonal matrix $\text{diag}(\lambda_1, \lambda_2, \dots, \lambda_p)$.

It is a pth order algebraic equation, also called the *charakteristic* equation (of the respective square matrix), or the *secular* equation. In general, it has p different roots λ,

$$\lambda_1, \lambda_2, \ldots, \lambda_p \tag{186}$$

(some of then may be equal). For each λ, we then determine the set of parameters α. By doing this we obtain the approximate solutions of the given problem (166). The lowest of the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_p$ is an approximate ground-state eigenenergy of the given Hamiltonian. The higher eigenvalues correspond to the approximate eigenenergies of the excited states.

We could obtain equation (184) even without using the variational theorem as it was explained in footnote 23. So, what good was the variational principle for us? Apart other things, it assures us that in our search for the eigenvalues we can not underestimate the exact ground state energy; see also the comment above the title of this section (6.1.1).

6.2 The Perturbation Method

Due to lack of time, we will not deal with the perturbation method. In quantum physics, this is generally an extremely important method. It is used e.g. to describe atoms and molecules embedded in an external electric or magnetic field, provided that the external field represents a weak effect compared to the effect of the internal field of the atom or molecule. If you are interested, you can get acquainted with the perturbation method, e.g. in the book [2] or [3]. More advanced formulations of the perturbation method (e.g. the perturbation method MP2 [7]) are used e.g. also in quantum chemistry to calculate important corrections to eigenenergies and functions obtained by other methods (typically by the Hartree-Fock method, which is one of the variational methods).

7 Internal Angular Momentum and Internal Magnetic Dipole Moment of Electron

In this section we will be talking about spin of an electron, which is a short term for the intrinsic internal angular momentum of an electron.[25] It will not be a detailed discussion as you should already know something about spin from the course *Quantum Mechanics*. It will be a similar dense summary with an emphasis on some systematics or order of steps, as we did in Chapter 1 on the postulates of wave quantum mechanics. In wave WM, only one quantum mechanical particle is discussed using a wave function, and its spin is ignored. In essence, the content of this chapter can be characterized by adding to the previous four postulates a fifth, which postulates the spin electron and the method of its description.

7.1 Experimental Facts Confirming Existence of Spin in Quantum Mechanics

We first recall the usual angular momentum and the usual magnetic moment known from the basic physics course: If a charged particle moves along a circle (or even along a more complicated loop), it has a corresponding *mechanical* (orbital) angular momentum (AM) $\vec{\ell} = \vec{r} \times \vec{p}$. This movement also creates the corresponding magnetic dipole moment (briefly magnetic moment, MM) of the magnitude $\mu = IS$, where S is the loop area and I is the current flowing through this thought circuit. If the charge of the particle is q and its mass m, then the relation between its MM and its orbital AM is

$$\boxed{\vec{\mu} = \frac{q}{2m}\,\vec{\ell}} \tag{187}$$

which can easily be derived [2].

Experiments show that the electron also has its *intrinsic* internal angular momentum and the corresponding intrinsic or internal magnetic dipole moment. These internal moments exist regardless of whether and how the electron moves in the space of the usual coordinates (in the “orbital” space). The internal angular momentum

---

[25] Other particles may also have their spins, however.

of a particle is called *spin*. The mentioned experiments or phenomena indicating the spin of the electron can be summarized in several groups [1], which we will only briefly list here:

1. fine structure of spectral lines
2. the Stern-Gerlach experiment
3. gyromagnetic phenomena: experiments of Einstein and de Haas
4. magnetooptic phenomena: the anomalous Zeeman effect

The existence of a quantum number associated with the internal state of an electron was postulated in 1925 by W. Pauli based on the study of atomic spectra. However, it was not he who introduced the term spin. In the same year, a little later, Kronig, Uhlenbeck, and Goudsmit interpreted this new quantum number as a manifestation of the momentum of the electron and introduced the term spin. According to the KUG hypothesis, a projection of this AM to a chosen axis can acquire two values only: $\pm\hbar/2$. In experiments of the type SG or EdH, it was also possible to measure the magnetic moment $\vec{\mu}_s$ corresponding to the spin and the ratio between $\vec{\mu}_s$ and the spin AM $\vec{s}$. It was found that [compare to (187)]

$$\boxed{\vec{\mu}_s = \frac{q}{m_\mathrm{e}}\,\vec{s}}\,, \qquad \text{where } q = -e, \quad m_\mathrm{e} = \text{the electron mass} \tag{188}$$

So, this (gyromagnetic) ratio is twice of the gyromagnetic ratio for the orbital AM. (It turns out that the latter ratio is only approximatelly twice of the former one.) By the "chosen axis" we usually consider the z axis. Therefore, we write that the projection of electron spin on the z axis is

$$s_z \in \left\{ +\frac{\hbar}{2}, -\frac{\hbar}{2} \right\} \tag{189}$$

As in QM we many times use the atomic units in which $\hbar = 1$, we then say that the electron spin acquires half-integer values. If we say that electron spin is $1/2$, we mean the *magnitude* of the spin projection in the units of $\hbar$.

The relation between the MM and AM is often expressed with the aid of the *Bohr magneton* μ_B:

$$\mu_\mathrm{B} = \frac{e\hbar}{2m_\mathrm{e}} \tag{190}$$

The mechanical (orbital) moment then becomes

$$\vec{\mu} = -g_L\,\mu_\mathrm{B}\,\frac{1}{\hbar}\,\vec{\ell} \tag{191}$$

where $g_L = 1$ is the so-called orbital g-factor. Analogously for the spin AM, we express

$$\vec{\mu}_s = -g_S\,\mu_\mathrm{B}\,\frac{1}{\hbar}\,\vec{s} \tag{192}$$

where $g_S \approx 2$ is the spin g-factor.

The theory and a better understanding of spin in the framework of QM was develped by Pauli in 1927. While spin had to be introduced into the standard QM by the postulating based on experimental facts (and Pauli significantly contributed to this – the Pauli equation from 1927 is especially known), in relativistic QM developed by Dirac in 1928, spin follows directly from theory. Spin cannot be satisfactorily interpreted as the rotation of a particle about its own axis.

7.2 Eigenvalues and Eigenvectors of Spin Operators

In Section (4.5), we talked about the theoretical description of AM and postulated that any angular momentum in QM will be defined so that the relevant operators must satisfy the commutation relations (49). Using this, we obtained the eigenvalues of the AM operators according to equations (68) and in addition the knowledge that the values of the quantum number j can only be non-negative integers or half-integers. We also found out that if, for some QM system, the highest AM projection on the z axis is $j\hbar$, then the number of the different projections is $2j + 1$. The whole procedure of that section could have been understood purely mathematically. But even in this early stage of our study we saw its physical realisation: the mechanical (i.e. orbital) angular momentum in QM. That time we did not know any physical realisation of a half-integral j. Now, knowing that the spin AM exists and acquires the two values of the projection, and that they are half-integral, we begin to see that the mathematical theory of section (4.5) will perfectly suit us to describe electron spin. It is sufficient to take $j = 1/2 = s$. Thus, once equations (68) are made specific for the case of electron spin (and with the explicit writing out of the constant $\hbar$), they become

$$\boxed{\begin{aligned} \hat{s}^2 \left|\frac{1}{2}, m_s\right\rangle &= \frac{3}{4}\hbar^2 \left|\frac{1}{2}, m_s\right\rangle \\ \hat{s}_z \left|\frac{1}{2}, m_s\right\rangle &= m_s\hbar \left|\frac{1}{2}, m_s\right\rangle \end{aligned}} \tag{193}$$

$$m_s \in \left\{-\frac{1}{2}, +\frac{1}{2}\right\}$$

These equations also define our notation that will be used below. Instead of the more generally used symbol j, we will be using $s = 1/2$ to denote the spin of one electron. We say that:

$$\begin{aligned} m_s &= +\frac{1}{2} \quad \ldots\ldots \quad \text{spin hore} \\ m_s &= -\frac{1}{2} \quad \ldots\ldots \quad \text{spin dole} \end{aligned}$$

Thus, only two linearly independent common eigenvectors of the operators $\hat{s}^2$ a $\hat{s}_z$ exist: the vectors

$$\left|\frac{1}{2}, +\frac{1}{2}\right\rangle, \quad \left|\frac{1}{2}, -\frac{1}{2}\right\rangle \tag{194}$$

denoting “spin up” and “spin down”.

7.3 Pauli Matrices

In quantum mechanics, it often proves practical to use matrix formalism. In it, instead of the notation (194), we use two-component column vectors:

$$\left|\frac{1}{2}, +\frac{1}{2}\right\rangle \to \tilde{\chi}_+ = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \tag{195a}$$

$$\left|\frac{1}{2}, -\frac{1}{2}\right\rangle \to \tilde{\chi}_- = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{195b}$$

It is a certain particular representation of vectors (194). Then it is natural to expect that the operators $\hat{s}^2$ and $\hat{s}_z$, which enter equations (193), will also have some specific matrix representation (expression), using 2×2 matrices. And of course, a matrix representation will be associated to the operators $\hat{s}_x$ and $\hat{s}_y$ as well. We denote the matrix expressions for the operators $\hat{s}_x$, $\hat{s}_y$ and $\hat{s}_z$ by symbols s_x, s_y, s_z. If we write these matrix expressions (at least using the symbols), then it becomes easy to construct the matrix expression of the operator $\hat{s}^2$:

$$s^2 = s_x^2 + s_y^2 + s_z^2 \tag{196}$$

We have the folowong requirements for these matrices:

1. They must be 2×2 (which we already said).
2. They must be hermitian (because they represent hermitian operators).
3. The (common) eigenvectors of the matrices s^2 a s_z must be column vectors $\tilde{\chi}_+$ a $\tilde{\chi}_-$ defined by formulae (195a), (195b).
4. The have to satisfy commutation relations (49) which are generally valid for AM operators.

Doing an explicit calculation, one can be convinced that these requirements are fulfilled by the followong matrices:

$$s_x = \frac{\hbar}{2}\sigma_x\,, \qquad s_y = \frac{\hbar}{2}\sigma_y\,, \qquad s_z = \frac{\hbar}{2}\sigma_z \tag{197}$$

where σ_x, σ_y, σ_z are Pauli matrices:

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma_y = \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix}, \qquad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{198}$$

The distribution of their matrix elements depends partiall on our definition of the ordering of the basis vectors. The ordering is by convention such as in (194), i.e. the first vector is the one with the positive spin, the second with the negative spin. The Pauli matrices themselves satisfy the commutation relations

$$[\sigma_x, \sigma_y] = 2\mathrm{i}\sigma_z \quad \text{a cyklicky ďalšie.} \tag{199}$$

The algebra of Pauli matrices can be read in more detail e.g. in the books [2, 3]. For example, the following identity holds:

$$\sigma_x^2 = \sigma_y^2 = \sigma_z^2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{200}$$

Then we obtain

$$s^2 = \frac{3}{4}\hbar^2 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{201}$$

[See also the upper equation in (193).]

7.4 Wavefunction of a Spin 1/2 Particle

In the postule no. 1 of the wave quantum mechanics, we introduced the wave function. Now we have to generalise it to be able to describe also the electron spin [1]. We consider one particle whose spin value can be generally indefinite, i.e. a superposition of the spin-up ($+1/2$) and spin-down ($-1/2$) state. We shortly call it a state with uncertain projection of the spin (on a chosen axis which we denote as z). However, the whole thing is more complicated, because the particle has not only the spin degree of freedom, but also the translational degrees of freedom (the usual spatial coordinates x, y, z, i.e. the "orbital" degrees of freedom); hence the wave function must depend not only on the spin, but also on the spatial coordinates. It should also be noted that even in the case of an indefinite spin, when measuring the projection, we acquire only the value $+1/2$ or $-1/2$.[26] The uncertainty lies in the fact that the result of an *individual measurement* on a state with indefinite spin is random, it cannot be predicted.[27] If we found out by the detection device that the particle has spin $+1/2$, then the space-dependent wave function would be some, let's denote it $\varphi_+(\vec{r}, t)$. If we found spin $-1/2$, then the spatially dependent wave function would be $\varphi_-(\vec{r}, t)$. There is no reason to assume that the spatial part should necessarily be the same in the two cases (although it is quite often the case). The corresponding probability densities are:

$$\rho_+(\vec{r}, t) = |\varphi_+(\vec{r}, t)|^2\,, \qquad \rho_-(\vec{r}, t) = |\varphi_-(\vec{r}, t)|^2 \tag{202}$$

The total probability density to find the particle at time t at point $\vec{r}$ is

$$\rho(\vec{r}, t) = \rho_+(\vec{r}, t) + \rho_-(\vec{r}, t) \tag{203}$$

---

[26] This was discussed in the 2^{nd} postulate of QM, although not specifically in relation to spin.

[27] However, we can know in advance the *probability* of measuring a particular value; this is if we know the wave function of the measured state.

and the equality

$$\int \rho(\vec{r}, t)\, \mathrm{d}^3 r = 1 \tag{204}$$

must hold. To describe the whole wave function, we introduce the 2×1 matrix

$$\varphi(\vec{r}, t) = \begin{pmatrix} \varphi_+(\vec{r}, t) \\ \varphi_-(\vec{r}, t) \end{pmatrix} \tag{205}$$

which is called *spinor* and it can also be expressed by

$$\varphi(\vec{r}, t) = \tilde{\chi}_+ \varphi_+(\vec{r}, t) + \tilde{\chi}_- \varphi_-(\vec{r}, t) \tag{206}$$

where $\tilde{\chi}_\pm$ are column vectors (195). It can be easily verified that this matrix satisfies

$$\varphi^\dagger(\vec{r}, t)\varphi(\vec{r}, t) = \rho_+ + \rho_- = \rho \tag{207}$$

By this paragraph, we have actually generalised or adapted the first postulate of QM so that we can also describe a particle with spin $1/2$.

However, we will not deal with the description of spin using spinors further, because extending this formalism to the description of many-electron systems would be impractical, at least for our purpose. Instead of spinors, we can use a wave function, that will have a *spin coordinate* as an argument, which we will talk about in the next sections.

8 Systems of Many Electrons

In this section we will consider mainly atoms and molecules with many electrons. We have reviewed the postulates of QM in chapter 1. For simplicity and brevity, we introduced them for one-particle systems. Therefore, it is now necessary to extend and complete them so that we can use QM for systems with many electrons and correctly take into account the spin of electrons.

8.1 Generalisation of the 1st Postulate: Many-Particle Wave Function[28]

$$\Psi(\vec{r}_1, \sigma_1, \vec{r}_2, \sigma_2, \ldots, \vec{r}_N, \sigma_N; t) \tag{208}$$

is the wave function of the N-particle system.

$$|\Psi(\vec{r}_1, \sigma_1, \vec{r}_2, \sigma_2, \ldots, \vec{r}_N, \sigma_N; t)|^2\, \mathrm{d}^3 r_1 \mathrm{d}^3 r_2 \ldots \mathrm{d}^3 r_N \tag{209}$$

is the *probability* that at time t, particle 1 is found around the point $\vec{r}_1$ in the volume element $\mathrm{d}^3 r_1$ and having the spin σ_1 and simultaneously particle 2 is found around the point $\vec{r}_2$ in the volume element $\mathrm{d}^3 r_2$ and having the spin σ_2 etc. The expression $|\Psi|$ (without the volume elements) is the corresponding *probability density.* It is found practical to define the meaning of the many-particle wave-function as it stated above even if we have a system of identical particles for which we do not really know which of them is the 1st one, which is the 2nd etc.

Normalisation:

$$\begin{aligned} \sum_{\sigma_1=\pm 1} \sum_{\sigma_2=\pm 1} \cdots \sum_{\sigma_N=\pm 1} \int \mathrm{d}^3 r_1 \mathrm{d}^3 r_2 \ldots \mathrm{d}^3 r_N\, |\Psi(\vec{r}_1, \sigma_1, \vec{r}_2, \sigma_2, \ldots, \vec{r}_N, \sigma_N; t)|^2 \equiv \\ \equiv \int \mathrm{d}x_1 \mathrm{d}x_2 \ldots \mathrm{d}x_N\, |\Psi(x_1, x_2, \ldots, x_N)|^2 = 1 \end{aligned} \tag{210}$$

x_i is a short notation for the spatial and spin coordinates together, i.e. $(x_i) \equiv (\vec{r}_i, \sigma_i)$. For even more brevity, we have introduced the formal integration sysmbol over x_i. It needs to be understood as including summation over the spin coordinte σ_i; it follows directly from the introduced notation. The particles assumed in this section need not necessarily be electrons, neither have to be identical. But the formulae with the summations over spins assume spin $1/2$. We will now introduce examples of particular forms of the wave function.

---

[28] In English-written literature, the term *many-body* wave function is sometime being used.

8.1.1 One Particle ($N = 1$)

An elementary form of one-particle wave function dependent on both spatial and spin coordinate can be expressed as a product of spatial and spin function:

$$\boxed{\Psi(\vec{r},\sigma) = \varphi(\vec{r})\,\chi(\sigma)} \tag{211}$$

where $\varphi(\vec{r})$ is some spatially dependent function and $\chi(\sigma)$ a spin-dependent one. If this is really to be a quite simple form of wave function, then the spin-dependent part should either represent spin up or spin down state, and not a linear combination of spins. Therefore, the elementary (and most important) examples of the function $\chi(\sigma)$ are the following particular functions:

$$\boxed{\chi_+(\sigma) = \begin{cases} 1, & \sigma = +1 \\ 0, & \sigma = -1 \end{cases}}, \qquad \boxed{\chi_-(\sigma) = \begin{cases} 0, & \sigma = +1 \\ 1, & \sigma = -1 \end{cases}} \tag{212}$$

$\chi_+(\sigma)$ is the eigenfunction of operator $\hat{s}_z$ for the eigenvalue $+\hbar/2$. Analogously, $\chi_-(\sigma)$ corresponds to the eigenvalue $-\hbar/2$. Thus, $\chi_+(\sigma)$ a $\chi_-(\sigma)$ are one-particle spin wave functions denoting states with spin up and down, respectively. We were using similar symbols ($\tilde{\chi}_+$, $\tilde{\chi}_-$) in the matrix (spinor) formalism of section 7.3, but here we are not using the formlism. To distinguish the notation, we use the letters χ without the tildes here, although they physically represent the same as do expressions (195). Note that the functions (212) involve both the spin coordinate (σ) and the spin index ($+$ or $-$), that is the spin quantum number. This is how it should be when expressing spin-dependent eigenfunctions: [29] both the spin coordinate and spin quantum number should appear in their notation. However, if a spin-dependent function were not an eigenfunction of the operator $\hat{s}_z$, then a particle in such a state would not have a sharp (i.e., definite) value of its spin projection on the z-axis, and then the spin index $+$ or $-$ could not be ascribed to such a function.

8.1.2 Two Particles ($N = 2$)

Now we are able to write elementary examples of two-particle wave functions as follows:

$$\boxed{\Psi_{n_1,\nu_1;n_2,\nu_2}(\vec{r}_1,\sigma_1,\vec{r}_2,\sigma_2) = \varphi_{n_1}(\vec{r}_1)\,\chi_{\nu_1}(\sigma_1)\,\varphi_{n_2}(\vec{r}_2)\,\chi_{\nu_2}(\sigma_2)} \tag{213}$$

where ν_1, ν_2 are the spin indices (quantum numbers) $+$ or $-$. The indices n_1, n_2 are quantum numbers (or sets of quantum numbers) of the orbital, i.e. of the spatially dependent functions $\varphi_{n_1}(\vec{r}_1)$, $\varphi_{n_2}(\vec{r}_2)$. Although the wave function just written cannot yet represent a physical state of a pair of electrons, because it does not meet the antisymmetry requirement, which we will learn later. However, by a linear combination of at least two wave functions of type (213), we will easily be able to construct an antisymmetric wave function. However, the form (213) could represent e.g. the wave function of the proton-electron system (if we consider the proton as a QM particle). [30]

8.2 The 4th Postulate: Schrödinger Equation for the Many-Particle Wave Function

$$\mathrm{i}\hbar\frac{\partial\Psi}{\partial t} = \hat{H}\Psi \tag{214}$$

Thus, this equation has the same form as it would have for one particle. Therefore, the time-independent SchE will also the form as for a single particle.

8.3 System of Identical Particles

Quantum-mechanical particles of the same kind are indistinguishable.

---

[29] Eigenfunction of the operator $\hat{s}_z$, possible of some other.

[30] We will approximate protons by classical particles in this course. Very often this is a completely sufficient approximation.

This is also a postulate, consequencies of which are confirmed by experiments. We will learn about the consequences of this postulate later and will work with it in this subsection as well. You can see a more detailed discussion of indistinguishability, e.g. in [2], beginning of chap. 15 or in [3], beginning of chap. XIV.

Consider a general many-body wave function [1]

$$\Psi(x_1, x_2, \ldots, x_N; t) \tag{215}$$

The particles even need not be electrons at this stage of our exposition. We only assume them to be of the same kind, that is identical particles. As we have stated above, in quantum physics we believe that they are indistinguishable. It means that if we make an interchange of any two of the particles (in the mathematical formalism, this is expressed by ***interchange of the corresponding coordinates***), nothing must change physcially. Therefore, the wave function may at most change its phase upon such interchange of the particles. We therefore demand the following:

$$\Psi(x_1, \ldots, x_j, \ldots, x_i, \ldots, x_N; t) = e^{\mathrm{i}\alpha}\, \Psi(x_1, \ldots, x_i, \ldots, x_j, \ldots, x_N; t) \tag{216}$$

where α is some real constant; we will soon see what specific values it can acquire.

It turns our convenient to introduce an operator for formal description of the interchange of the particles i and j. We define ot by the equation

$$\hat{P}_{ij}\, F(x_i, x_j) = F(x_j, x_i) \tag{217}$$

in which we started to use the more concise notation expressing only the coordinates of the electrons to be exchanges. F can be arbitrary function of the considered arguments (possibly also of time); in this definition, it need not yet be the wave function Ψ. Even a higher degree of brevity is achieved by denoting the arguments using numbers only, i.e.

$$F(1, 2, \ldots, N) \tag{218}$$

and the action of the operator is

$$\hat{P}_{ij}\, F(i, j) = F(j, i) \tag{219}$$

From the point of view of combinatorics, this operator make a permutation Using the operator, we can now rewrite equation (216) as follows:

$$\underbrace{\hat{P}_{ij}\Psi(i, j; t)}_{\Psi(j, i; t)} = e^{\mathrm{i}\alpha}\, \Psi(i, j; t) \tag{220}$$

We see that a *physically correct* (or realistic [2]) *wave function of a system of identical particles must be an eigenfunction of the particle interchange operator.*

8.4 Eigenvalues and Eigenfunctions of the $\hat{P}_{ij}$ operator

$$\hat{P}_{ij} f(i, j) = \lambda f(i, j) \tag{221}$$

Let us act by the operator $\hat{P}_{ij}$ on this equation from the left. We will obtain two occurrencies of the operator on the LHS of the equation which will return the indices i and j to their original order:

$$f(i, j) = \lambda \hat{P}_{ij} f(i, j) = \lambda^2 f(i, j)$$

Therefore

$$\boxed{\lambda = \pm 1} \tag{222}$$

Eigenfunctions for $\lambda = +1$ are those that obey the property $f(j, i) = f(i, j)$. These are called ***symmetric*** functions. Eigenfunctions for $\lambda = -1$ are those that obey the property $f(j, i) = -f(i, j)$. These are called ***antisymmetric*** functions.

The particle exchange operator is linear. It is also a hermitian one since it has real eigenvalues. And, importantly, we are now already able to formulate the findings at the end of the last section more specifically: ***A physically correct wave function of a system of indentical particles has to be either symmetric or antisymmetric upon exchange of the coordinates.***

8.5 Commutation $[\hat{H}, \hat{P}_{ij}] = 0$, Bosons, Fermions, Permutation Symmetry of a Wave Function

The Hamilton operator corresponds to a physical quantity; therefore it will not change upon interchange of two indetical particles:

$$\hat{P}_{ij}\hat{H} = \hat{H} \tag{223}$$

This is a consequence of the postulated indistinguishability of the particles. The last equation written represents the effect of the operator $\hat{P}_{ij}$ on another operator ($\hat{H}$). Now act by the operator $\hat{P}_{ij}$ on the *function* $\hat{H}F(i,j)$, where $F(i,j)$ is any function of multiple spatial and spin coordinates. Again, we write explicitly only those of them that are interchanged by the given operator.

$$\hat{P}_{ij}\{\hat{H}F(i,j)\} = \hat{H}\{\hat{P}_{ij}F(i,j)\}$$

because the particle interchange operator will simply "cross" over the Hamiltonian and will not do anything with it. For this holds for arbitrary function F, it implies that

$$\boxed{\boxed{[\hat{H}, \hat{P}_{ij}] = 0}} \tag{224}$$

We have learnt in secion 3.1 what are consequences of commutation of two operators: a complete set of eigenfunctions for each of the operators can be constructed such that all these functions are eigenfunctions of both the operators (Theorem 7). So *by meeting the requirement that the wave function be an eigenfunction of the particle interchange operator, we do not lose the ability to construct this function so that it is also an eigenfunction of the Hamiltonian of the system.*

Now a question arises: is the many-particle wave function symmetric or antisymmetric upon interchange of some of its two coordinates? Available data and analysis show that [2, 3, 4]:

A system of identical bosons (particles of an integer spin) is always described by a symmetric wave function. A system of identical fermions (particles of a half-integer spin) is always described by an antisymmetric wave function.

This statement must be considered to be an independent postulate in non-relativistic QM (in addition to the postulates we have formulated in section 1). It applies also for time-dependent function, not only for stationary ones. These properties of wave functions can be proved theoretically in the quantum field theory. (I.e., they are just some derived facts, not postulates, in the quantum field theory.) Using an explicit formula, the antisymmetry is expressed by the relation The antisymmetry is

$$\Psi(j,i;t) = -\Psi(i,j;t) \tag{225}$$

8.6 The Pauli Principle

Let us now write equation (221) for the eigensystem of the operator in more detail and by writing the most general wave function possible, i.e. also time-dependent, to see that the Pauli principle applies very generally, not only to stationary states.

$$\hat{P}_{ij}\Psi(x_i,x_j;t) = \lambda\Psi(x_i,x_j;t) \tag{226}$$

It means that in this section we consider a wave function either symmetric or antisymmetric; only these two kinds can be eigenfunctions of $\hat{P}_{ij}$. See section 8.4. And it also means (according to the postulate at the end of the previous section) that we consider a wave function of a system of identical particles, either bosons or fermions. We now expand the N-particle wave function $\Psi(x_i,x_j;t)$ in some complete set of orthogonal functions as follows (see Appendix A.2):[31]

$$\Psi(x_i,x_j;t) = \sum_{n_i}\sum_{n_j} C_{n_i n_j}(\bar{x};t)\,\phi_{n_i}(x_i)\,\phi_{n_j}(x_j) \tag{227}$$

---

[31] It was explained in the lecture too: we first imagine the function $\Psi(x_i,x_j;t)$ as dependent on single variable only, the other variables having some fixed values. Etc.

where $\bar{x}$ is the set of coordinates $x_1, \ldots, x_N$ excluding the coordinates x_i, x_j. n_i, n_j are summations indices which also represent quantum numbers. Typically, they are composite indices.

For example, if we were going to express a state of electrons in an atom by the wave function $\Psi(x_i, x_j; t)$, the single-particle basis functions $\phi_{n_i}(x_i)$ would then be eigenfunctions of the hydrogen atom or similar functions; thus, the simple symbol n_i would in fact mean

$$n_i \to (n_i, l_i, m_i, \nu_i)$$

where $n_i \in \mathbb{N}$ on the RHS is the principal quantum number, $l_i \in \{0, 1, \ldots, n_i - 1\}$ is the orbital quantum number, $m_i \in \{-l_i, -l_i{+}1, \ldots, l_i\}$ je magnetické kvantové číslo and ν_i is the spin quantum number (by convention, its values can be $\pm 1/2$ or ± 1, or it is only denoted by the symbols $+$, $-$). We have discussed the spin quantum numbers in sections 8.1.1 and 8.1.2.

We can interpret expression (227) as follows: assuming that the particles with the exception of the i^{th} and j^{th} one are at positions denoted by the multi-coordinate $\bar{x}$, $C_{n_i n_j}(\bar{x}; t)$ is the probability amplitude[32] to find the particle i in the state ϕ_{n_i} and at the same time to find the particle j in the state ϕ_{n_j}. A more detailed and elementary explanation of the meaning of expansion (227) is explained in Appendix A.2. Now, substitute expansion (227) into(226); we can omit the arguments in the coefficients $C_{n_i n_j}(\bar{x}; t)$ for the sake of brevity.

$$\hat{P}_{ij} \sum_{n_i} \sum_{n_j} C_{n_i n_j}\, \phi_{n_i}(x_i)\, \phi_{n_j}(x_j) = \lambda \sum_{n_i} \sum_{n_j} C_{n_i n_j}\, \phi_{n_i}(x_i)\, \phi_{n_j}(x_j)$$

We apply the permutation operator in the LHS.

$$\text{LHS} = \sum_{n_i} \sum_{n_j} C_{n_i n_j}\, \phi_{n_i}(x_j)\, \phi_{n_j}(x_i)$$

In next step, we rename the summation indices: $n_i \leftrightarrow n_j$ (which is a trivial operation for the symbol used as a summation index can be arbitrary).

$$\text{LHS} = \sum_{n_j} \sum_{n_i} C_{n_j n_i}\, \phi_{n_j}(x_j)\, \phi_{n_i}(x_i)$$

Let us now equate the LHS with the RHS. In doing so, we write the summation symbols $\sum$ in the same order as they are on the RHS; the result does not depend on the order.

$$\sum_{n_i} \sum_{n_j} C_{n_j n_i}\, \phi_{n_j}(x_j)\, \phi_{n_i}(x_i) = \lambda \sum_{n_i} \sum_{n_j} C_{n_i n_j}\, \phi_{n_i}(x_i)\, \phi_{n_j}(x_j)$$

From this, we obtain the equation

$$\sum_{n_i} \sum_{n_j} \left(C_{n_j n_i} - \lambda C_{n_i n_j} \right)\, \phi_{n_i}(x_i)\, \phi_{n_j}(x_j) = 0\,, \qquad\qquad \forall x_i, x_j$$

We multiple this equation by functions $\phi^*_{m_i}(x_i)$ a $\phi^*_{m_j}(x_j)$ from the left side and subsequently we do integration over the spatial coordinates and summation over the spin coordinates. Employing the orthogonality

$$\int \phi^*_m(x)\, \phi_n(x)\, \mathrm{d}x \propto \delta_{m,n} \tag{228}$$

(in fact, there is the summation over the spin coordinate but we denote it all using just the integration symbol for brevity) we obtain the set of algebraic equations[33]

$$C_{n_j n_i} = \lambda\, C_{n_i n_j} \tag{230}$$

---

[32] If we had two particles only in the system, then $C_{n_1 n_2}(\bar{x}; t)$ would be a true probability amplitude (amplitude of probability). But since we aim to consider an N-particle system, we introduced a little of technical and interpretive complication to our proof of the Pauli principle, in particular, we had to say "assuming that the particles with the exception of ...".

[33] From the mentioned orthogonality of the one-particle functions, the orthogonality for the two-particle functions also follows:

$$\int \psi^*_M(x_i, x_j)\, \psi_N(x_i, x_j)\, \mathrm{d}x_i\, \mathrm{d}x_j \propto \delta_{MN} \tag{229}$$

with $\psi_M(x_i, x_j) = \phi_{m_i}(x_i)\, \phi_{m_j}(x_j)$, and analogously $\psi_N(x_i, x_j)$. $M = (m_i\; m_j)$, $N = (n_i\; n_j)$ are composite (double)indices. Ordering of their components is unimportant.

It means that the expansion coefficients $C_{n_i n_j}$ are either symmetric or antisymmetric with respect to the exchange od their indices. This could already be deduced directly from expansion (227).

What is the probability that at time t both particle i and particle j will be in the same state ϕ_n? To explore it, let us look at the coefficients with the same indices in expansion (227), that is $n_i = n_j$. According to (230), the case of $n_i = n_j = n$ implies

$$C_{nn} = \lambda\, C_{nn} \tag{231}$$

For $\lambda = 1$ (bosons), it is an identity. For $\lambda = -1$ (fermions) we obtain $C_{nn} = 0$. It means that

probability to find two fermions to occupy the same single-particle state $\phi_n(x)$ is vanishing, which is the statement of the Pauli exclusion principle.

In other words, *two fermions cannot occupy the same individual quantum state* [3]. For example, if an electron in an orbital around an atomic nucleus is characterised by the quantum numbers n, l, m and, in addition, by the projection of its spin on the z axis (which can be understood as a spin quantum number and denoted as m_s or ν), then any state of any other electron (in the same atom) must differ in at least one of the four quantum numbers. (The spin quantum numbers have been explained in sections 8.1.1 and 8.1.2.)

In what follows, we will mostly discuss not a general wave function (which can also be time-dependent) but stationary states only, i.e. eigenfunctions of the Hamiltonian.[34]

8.7 Wave Function of a Two Electron System

The best known and most frequently met representants of two-electron systems are the helium atom and the hydrogen molecule H_2 [8]. We will work out the helium atom in an exercise and will calculate its ground-state energy using a simple version of the variational method with one parameter (see Appendix D). We obtained or will obtain the result – especially the energy of its ground state – which has a remarkable quantitative accuracy given the simple analytically manageable method we use for it. In this section, we will focus on better understanding the *wave function* of two-electron systems, and not only the wave function of the ground state, but also the excited states.

It is worth saying in advance that *the exact eigenfunctions of any two- and many-electron* ***interacting*** *system cannot be expressed analytically.* The obstacle is caused by the Coulomb interaction of electrons with each other, which is a difficult problem when examining the electronic structure. In other words, we cannot solve the corresponding Schrödinger equation (neither stationary nor time-dependent) exactly for such a system. Numerically, however, at least the ground state of the two-electron system could be found with virtually any accuracy.

To understand the electronic structure, it is necessary to know at least qualitatively a correct analytical form of the wave function. We propose the wave function of helium (also of a similar ion, but briefly speaking only helium) expressed in the form of the product $\psi(\vec{r}_1, \vec{r}_2) = \varphi(\vec{r}_1)\varphi(\vec{r}_2)$ of the hydrogen orbitals. We have done it so in the exercise; see (D.4). Let us now examine particular forms of two-electron wave functions in the context of the general requirements that a correct wave function has to meet.

8.7.1 Independent Electrons

At first, we have to realise that the factorised form $\varphi(\vec{r}_1)\varphi(\vec{r}_2)$ [eq. (D.4)] cannot be an exact eigenfunction of the helium Hamiltonian (D.1). To make sure of this, consider a two-electron Hamiltonian that can be written as a sum of commuting single-particle operators as follows:

$$\hat{H}^{\mathrm{ind}} = \hat{h}_1(\vec{r}_1) + \hat{h}_2(\vec{r}_2) \tag{232}$$

It is therefore the sum of mutually *independent* operators, because the individual variables (coordinates) are *separated.* Assume we know the exact eigenfunctions of these one-particle Hamiltonians:

$$\hat{h}_1(\vec{r})\varphi_1(\vec{r}) = \mathcal{E}_1\varphi_1(\vec{r})\,, \qquad \hat{h}_2(\vec{r})\varphi_2(\vec{r}) = \mathcal{E}_2\varphi_2(\vec{r}) \tag{233}$$

---

[34] Determination of eigenenergies of a Hamiltonian is the principal task in the electronic structure theory and also in this course.

Let us now construct the factorised function

$$\psi(\vec{r}_1, \vec{r}_2) = \varphi_1(\vec{r}_1)\varphi_2(\vec{r}_2) \tag{234}$$

and calculate what the Hamiltonian $\hat{H}^{\text{ind}}$ will do with it:

$$\hat{H}^{\text{ind}}\psi(\vec{r}_1, \vec{r}_2) = \left[\hat{h}_1(\vec{r}_1) + \hat{h}_2(\vec{r}_2)\right] \varphi_1(\vec{r}_1)\varphi_2(\vec{r}_2) = \mathcal{E}_1\varphi_1(\vec{r}_1)\varphi_2(\vec{r}_2) + \mathcal{E}_2\varphi_1(\vec{r}_1)\varphi_2(\vec{r}_2)$$

thus

$$\hat{H}^{\text{ind}}\psi(\vec{r}_1, \vec{r}_2) = (\mathcal{E}_1 + \mathcal{E}_2)\,\psi(\vec{r}_1, \vec{r}_2) \tag{235}$$

The above procedure can be reversed in its sequence (albeit in a little more complicated way) and thus it can be shown that if some product function $\varphi_1(\vec{r}_1)\varphi_2(\vec{r}_2)$ is an (exact) eigenfunction of some Hamiltonian dependent on the variables $\vec{r}_1$ and $\vec{r}_2$, then the Hamiltonian can be expressed as the sum of type $\hat{h}_1(\vec{r}_1) + \hat{h}_2(\vec{r}_2)$. So we state:

The product (factorised) wave function (234) *is the (exact) eigenfunction of Hamiltonian* (232) *having separated variables. The energy of such a two-particle system is the sum of the energies of the individual particles* (here electrons).

Two electrons described by the Hamiltonian of the type $\hat{h}_1(\vec{r}_1) + \hat{h}_2(\vec{r}_2)$ are mutually *independent*, they do not affect one another in any way. However, the product wave function cannot be an exact eigenfunction of the helium Hamiltonian, because this cannot be written as a sum of the type $\hat{h}_1(\vec{r}_1) + \hat{h}_2(\vec{r}_2)$. The interaction term $\hat{W} = 1/|\vec{r}_1 - \vec{r}_2|$, which cannot be decomposed into a separated form, prevents this. Thus, in our exercise, we actually found (or will only find) an exact ground state of some effective Hamiltonian of the form

$$H^{\text{eff}} = -\frac{1}{2}\nabla_1^2 - \frac{1}{2}\nabla_2^2 - \frac{Z_{\text{eff}}}{r_1} - \frac{Z_{\text{eff}}}{r_2} \tag{236}$$

only (expressed in the atomic units), where Z_{eff} expressed by (D.44) is the effective nuclear charge calculated by the optimisation. As we can see, H^{eff}, is just a particular example of the separated form $\hat{H}^{\text{ind}}$. Thanks to especially the mentioned optimisation, we were able to consider the factorised wave function of type $\varphi_1(\vec{r}_1)\varphi_2(\vec{r}_2)$ as at least approximately a good eigenfunction of the helium Hamiltonian. If we did not optimise anything, then the electron shell of such a helium would only be an addition of two electron shells of the helium cations on each other, without any manifestation of the electron interaction. Since we have done the optimisation, the electrons in such a helium model are not considered to be completely independent of each other from a physical point of view. Their interaction is included in such a way that the charge of the nucleus in such a model is (effectively) smaller than the actual charge of the helium nucleus. It is an effect of the partial screening of the nucleus field by the other electron. Mathematically, however, the form of the wave function remains at a simple level of independent electrons, which is of course very practical. This is used in mean-field methods, where an electron moves as if without interacting with the other electrons in an average field generated by the nuclei and the other electrons. In the context of electronic structure, these are the Hartree method and also the Hartree-Fock method, which we will talk about later.

8.7.2 Antisymmetry of Wave Function

We have learnt in section (8.5) that a wave function for a system of identical fermions must be antisymetric upon interchange of the coordinates of any two fermions, i.e. it has to switch its sign. The product wave function we used for the helium atom is obviously not of this kind. We used $\varphi_1 = \varphi_2 = \varphi$ for the ground state of helium (thus both electrons occupying the same atomic orbital). Such a function is symmetric; therefore we now denote it as $\psi_{\text{S}}(\vec{r}_1, \vec{r}_2)$:

$$\psi_{\text{S}}(\vec{r}_1, \vec{r}_2) = \varphi(\vec{r}_1)\varphi(\vec{r}_2) \tag{237}$$

How do we deal with the problem that it is not antisymmetric? We realise that we have not considered any spin in whis wave function. A complete form of a wave function has to include spin too; we use the spin coordinates for this. Only such a complete (total) wave function must be antisymmetric. In addition to this property, we want it to express that two electrons (e.g., in a helium atom or in an H_2 molecule) occupy the same spatial orbitals, but differ in their spins.

If we express the total wave function in the form

$$\Psi(\vec{r}_1, \sigma_1, \vec{r}_2, \sigma_2) = \psi_{\text{S}}(\vec{r}_1, \vec{r}_2)\,\chi_{\text{A}}(\sigma_1, \sigma_2) \tag{238}$$

where $\chi_\mathrm{A}(\sigma_1,\sigma_2)$ is some antisymmetric function of the spin coordinates, then also the product (238) will be an antisymmetric function. Even without a *deeper reasoning* (we will do it later), it is so far acceptable that the antisymmetric spin wave function can be expressed as follows:

$$\chi_\mathrm{A}(\sigma_1,\sigma_2)=\frac{1}{\sqrt{2}}\left[\chi_+(\sigma_1)\chi_-(\sigma_2)-\chi_-(\sigma_1)\chi_+(\sigma_2)\right] \tag{239}$$

where $\chi_+(\sigma)$ a $\chi_-(\sigma)$ are one-particle spin wave functions denoting states with spin "up" and "down" [see (212)]:

$$\chi_+(\sigma)=\begin{cases}1, & \sigma=+1\\ 0, & \sigma=-1\end{cases} \qquad \chi_-(\sigma)=\begin{cases}0, & \sigma=+1\\ 1, & \sigma=-1\end{cases} \tag{240}$$

The multiplication constant $1/\sqrt{2}$ in the wave function is to ensure the proper normalisation:

$$\langle\chi_\mathrm{A}|\chi_\mathrm{A}\rangle=\sum_{\sigma_1=\pm1}\sum_{\sigma_2=\pm1}\chi_\mathrm{A}^*(\sigma_1,\sigma_2)\,\chi_\mathrm{A}(\sigma_1,\sigma_2)=1$$

In this way we have solved the problem that the purely spatial (i.e. orbital) wave function of helium, as we had proposed it by formula (237), did not obey the anisymmetry requirement. The total wave function (238) cannot be expressed as a product of two functions, one of which would depend on the coordinates of the first electron only, the other on the coordinates of the second. Thanks to the form (237) we can say that in state (238), both electrons occupy identical spatial orbital, but differ in the projections of their spins on the z axis. Consequently, the Pauli principle, which says that two electrons cannot occupy the same one-particle state, is satisfied as well. Ours differ in their spins which suffices to satisfy the Pauli principle.

It is necessary to realise that functions not obeying the antisymmetry principle can also be eigenfunctions of Hamiltonian $\hat{H}^\mathrm{ind}$ [eq. (232)] (and also of other Hamiltonian, see below). For instance, a mathematically correct eigenfunction is also the (purely spatial-dependent) function (234), which is neither symmetric nor antisymmetric. If we multiplied it by any spin-dependent function, such total wave function would still be a mathematically correct eigenfunction of Hamiltonian (232). However, we have learnt in section 8.5 that a correct total (both spatial and spin-dependent) wave function for a system of electrons must be antisymmetric. Of the many mathematically correct eigenfunctions of the Hamiltonian, only those that satisfy the antisymmetry requirementm, are physically significant (correct).

8.7.3 Clasification of States in the Helium Atom

At the the end of the last section, we made the statement in the sense that whatever is the spin part of the total wave function, this wave function will still be a mathematically correct eigenfunction of the Hamiltonian $\hat{H}^\mathrm{ind}$ [eq. (232)], assuming of course that the spatial part $\psi(\vec{r}_1,\vec{r}_2)$ is mathematically correct. This holds not only for $\hat{H}^\mathrm{ind}$; consider the exact non-relativistic Hamiltonian of the helium atom or like ion (shortly just helium):

$$\hat{H}=-\frac{\hbar^2}{2m}\nabla_1^2-\frac{\hbar^2}{2m}\nabla_2^2-\frac{1}{4\pi\varepsilon_0}\frac{Ze^2}{r_1}-\frac{1}{4\pi\varepsilon_0}\frac{Ze^2}{r_2}+\frac{1}{4\pi\varepsilon_0}\frac{e^2}{|\vec{r}_1-\vec{r}_2|} \tag{241}$$

Let $\psi(\vec{r}_1,\vec{r}_2)$ be one of its purely spatial eigenfunctions:

$$\hat{H}\psi(\vec{r}_1,\vec{r}_2)=E\psi(\vec{r}_1,\vec{r}_2)$$

(Of course, this is not a factorised function.) Write the total wave function as follows:

$$\Psi(\vec{r}_1,\sigma_1,\vec{r}_2,\sigma_2)=\psi(\vec{r}_1,\vec{r}_2)\,\chi(\sigma_1,\sigma_2) \tag{242}$$

Ψ is an eigenfunction of $\hat{H}$ with *arbitrary* $\chi(\sigma_1,\sigma_2)$, for the non-relativistic helium Hamiltonian [eq. (241)] does not dependent on spin variables. Such a Hamiltonián (and also any other that is independent of spin variables) in no way effects on the purely spin function χ. To be convinced explicitely about it, calculate as follows:

$$\hat{H}\Psi(\vec{r}_1,\sigma_1,\vec{r}_2,\sigma_2)=\hat{H}[\psi(\vec{r}_1,\vec{r}_2)\,\chi(\sigma_1,\sigma_2)]=\chi(\sigma_1,\sigma_2)\hat{H}\psi(\vec{r}_1,\vec{r}_2)=E\Psi(\vec{r}_1,\sigma_1,\vec{r}_2,\sigma_2)$$

that is, $\Psi(\vec{r}_1, \sigma_1, \vec{r}_2, \sigma_2)$ is really the eigenfunction.

And if $\psi(\vec{r}_1, \vec{r}_2)$ were an only approximate purely spatial eigenfunction of $\hat{H}$, then $\Psi(\vec{r}_1, \sigma_1, \vec{r}_2, \sigma_2)$ would be an approximate total eigenfunction.

However, the independence of a Hamiltonian of spin coordinates also means that such a Hamiltonian commutes with any spin operator. Realise that this holds not only for the simplified Hamiltonian $\hat{H}^{\mathrm{ind}}$ [eq. (232)], but also, for example, for the exact non-relativistic helium Hamiltonian $\hat{H}$ [eq. (241)] containing the electron-electron interaction contribution. For example, the identities

$$[\hat{H}, \hat{S}_z] = [\hat{H}, \hat{S}^2] = 0 \tag{243}$$

hold, in which

$$\hat{\vec{S}} = \hat{\vec{s}}_1 + \hat{\vec{s}}_2 \tag{244}$$

is the total spin angular momentum (AM) operator of a two-electron system and

$$\hat{S}^2 = (\hat{S}_x)^2 + (\hat{S}_y)^2 + (\hat{S}_z)^2 \tag{245}$$

is the definition of the operator of the square of the total spin AM of the system. By a simple explicit calculation, it can be convinced that this operator has also the from

$$\hat{S}^2 = (\hat{s}_1 + \hat{s}_2)^2 \tag{246}$$

which we intuitively expect too. As we have seen in previous lectures, the pair of AM operators describing the projection on the z-axis and the square of AM magnitude are especially important. Therefore, in accordance with (243), it is convenient and practical to construct the eigenfunctions of the spin-independent Hamiltonian so that they are also eigenfunctions of the operators $\hat{S}_z$ and $\hat{S}^2$. [35]

Let us, therefore, verify whether the wave function $\Psi = \psi_{\mathrm{S}}\chi_{\mathrm{A}}$ [eq. (238)] including the anisymmetric spin function (239) is an eigenfunction of the operators $\hat{S}_z$ and $\hat{S}^2$. (Now begins the "deeper reasoning" that we have mentioned in sec. 8.7.2, but we skipped it.) The orbital part of the wave function will not do anything wrong in this verification, because it does not contain spin degrees of freedom (spin coordinates) and we simply "swap" the spin operator over it, as we will see below. First, calculate

$$\hat{S}_z\Psi(\vec{r}_1, \sigma_1, \vec{r}_2, \sigma_2) \;\equiv\; \hat{S}_z\Psi(1,2) \;=\; \hat{S}_z\left[\psi_{\mathrm{S}}(1,2)\chi_{\mathrm{A}}(1,2)\right] = [\text{prehodíme}] = \psi_{\mathrm{S}}(1,2)\hat{S}_z\chi_{\mathrm{A}}(1,2)$$

If the equality (not verified yet) $\hat{S}_z\chi_{\mathrm{A}}(1,2) = \lambda\chi_{\mathrm{A}}(1,2)$, would be true, then we would get $\hat{S}_z\Psi(1,2) = \lambda\Psi(1,2)$, that is the function Ψ would be an eigenfunction of the operator $\hat{S}_z$ *regardless of the form of the purely orbital function* ψ_{S}. Hence, it is sufficient to check whether the spin function $\chi_{\mathrm{A}}(1,2)$ is an eigenfunction of the operator $\hat{S}_z$. If this is so, then the total wave function $\Psi(1,2)$ is an eigenfunction too. So, do the verification:

$$\begin{aligned}\hat{S}_z\chi_{\mathrm{A}}(1,2) &= \frac{1}{\sqrt{2}}\left(\hat{s}_{1z} + \hat{s}_{2z}\right)\left[\chi_+(1)\chi_-(2) - \chi_-(1)\chi_+(2)\right] = \\ &= \frac{1}{\sqrt{2}}\left\{\left[\frac{\hbar}{2}\chi_+(1)\chi_-(2) - \frac{\hbar}{2}\chi_+(1)\chi_-(2)\right] - \left[-\frac{\hbar}{2}\chi_-(1)\chi_+(2) + \frac{\hbar}{2}\chi_-(1)\chi_+(2)\right]\right\} = 0\end{aligned} \tag{247}$$

---

[35] And this is not all, because proper eigenfunctions must also be eigenfunctions of the particle exchange operator, see (217) and the end of the relevant paragraph. In addition, the spin-independent Hamiltonian of the atom also commutes with the operators of the total *orbital* angular momentum, which for helium are the Cartesian components of the operator

$$\hat{\vec{L}} = \hat{\vec{\ell}}_1 + \hat{\vec{\ell}}_2$$

and also the important operator $\hat{L}^2$. Therefore, the correct and practically expressed eigenfunction of the spin-independent Hamiltonian should also be an eigenfunction of the operators $\hat{L}_z$ and $\hat{L}^2$. Fortunately, we are already familiar with this, at least in the case of the hydrogen atom, but we see that things get complicated. We would have an even more complex task if we included the spin-orbital interaction in the Hamiltonian, i.e. that the Hamiltonian would also depend on the spin coordinates. There, in general, the Hamiltonian of an atom would commute with the total angular momentum $\hat{J} = \hat{L} + \hat{S}$ and with its square $\hat{J}^2$, but not with the orbital or spin moment separately. We return to the commutation with $\hat{J}_z$ and $\hat{J}^2$ because this applies to an atom in general. These problems are easier to study by examining the invariance of the Hamiltonian under different rotations in orbital and spin space [2]. The angular momentum operators are also operators expressing rotations in orbital or spin space. However, we do not have time to deal with this.

Thus, the two-electron spin function $\chi_{\mathrm{A}}(1,2)$ is really an eigenfunction of the spin-sum operator and the corresponding eigenvalue is 0, i.e. the corresponding two-particle state demonstrates zero projection of the total spin on the z axis. However, it is necessary to verify, whether this spin function is also an eigenfunction of the operator $\hat{S}^2$. To do this, we express the operator in the form

$$\hat{S}^2 = \hat{S}_+\hat{S}_- + (\hat{S}_z)^2 - \hbar\hat{S}_z \tag{248}$$

which we have derived for the general angular momentum in the exercise; see formula (57a). We will employ also one of the equalities (69) expressing the effect of the raising and lowering operators on the angular momentum eigenstates. So, we can calculate:

$$\hat{S}^2\chi_{\mathrm{A}}(1,2) = \hat{S}_+\hat{S}_-\chi_{\mathrm{A}} + \hat{S}_z\underbrace{\hat{S}_z\chi_{\mathrm{A}}}_{0} - \hbar\underbrace{\hat{S}_z\chi_{\mathrm{A}}}_{0}$$

Thus, we need to evaluate

$$\hat{S}_-\chi_{\mathrm{A}} = (\hat{s}_{1-} + \hat{s}_{2-})\,\frac{1}{\sqrt{2}}\,[\chi_+(1)\chi_-(2) - \chi_-(1)\chi_+(2)]$$

Realise that $\hat{s}_{1-}$ acts on functions with the spin coodinate of the electron 1 [i.e. on $\chi_+(1), \chi_-(1)$] and analogously, $\hat{s}_{2-}$ on the electron 2. We do the multiplication of the expressions in the parentheses, use the mentioned formula (69b) for $j = 1/2$ and m either $+1/2$ or $-1/2$. After a few lines we get

$$\hat{S}_-\chi_{\mathrm{A}} = 0$$

Therefore, the equality

$$\hat{S}^2\chi_{\mathrm{A}}(1,2) = 0 \tag{249}$$

holds. It means that the spin wave function $\chi_{\mathrm{A}}(1,2)$ [and consequently also the total wave function $\Psi(1,2)$] is also an eigenfunction of the operator $\hat{S}^2$ and the corresponding eigenvalues is zero again.

Thus, the antisymmetric two-electron spin wave function [eq. (239)] represents a pair of electrons with zero projection of the sum of their spins on the z axis (and also on other arbitrary axis) and also with zero total spin (i.e. with zero eigenvalue of the operator $\hat{S}^2$).

Štandardne sa pre spinové vlnové funkcie používa značenie

$$\chi^S_{S_z} \tag{250}$$

is being used for the spin wave functions. The upper index marks the value of S in the equation

$$\hat{S}^2\chi^S_{S_z} = \hbar^2 S(S+1)\chi^S_{S_z} \tag{251}$$

and the lower index the value of S_z in the equation

$$\hat{S}_z\chi^S_{S_z} = \hbar S_z\chi^S_{S_z} \tag{252}$$

[See, for instance, equations (68), in which we were omitting $\hbar$ for the sake of brevity.] From the indices S and S_z we know what eigenvalues of the operators $\hat{S}^2$ and $\hat{S}_z$ their common eigenfunction corresponds to, so we know everything about the total spin of the system.

Singlet: $\boldsymbol{S = 0}$, $\boldsymbol{S_z = 0}$. Thus, we write the spin function (239) as follows:

$$\chi_{\mathrm{A}}(\sigma_1,\sigma_2) = \boxed{\chi^0_0 = \frac{1}{\sqrt{2}}\,[\chi_+(\sigma_1)\chi_-(\sigma_2) - \chi_-(\sigma_1)\chi_+(\sigma_2)]} \tag{253}$$

It is called *singlet*. A system in the singlet state (the He atom, for instance) has anti-parallel spins of its electrons and the total spin 0.

Triplet: $S = 1$, $S_z \in \{-1, 0, +1\}$. From the theory in section 4.5 we know, that spin functions with non-zero spin should obviosly to exist. (We still consider a two-electron system.) These are the functions [2, 3]

$$\boxed{\begin{aligned}
\chi^1_1 &= \chi_+(\sigma_1)\chi_+(\sigma_2) && (254a)\\
\chi^1_0 &= \frac{1}{\sqrt{2}}\left[\chi_+(\sigma_1)\chi_-(\sigma_2) + \chi_-(\sigma_1)\chi_+(\sigma_2)\right] && (254b)\\
\chi^1_{-1} &= \chi_-(\sigma_1)\chi_-(\sigma_2) && (254c)
\end{aligned}}$$

and together they are called a *triplet*. This name originates from the property that if an atom is inserted in an external magnetic field, then the corresponding energy level (which would otherwise be degenerate if we neglect other, very weak effects) splits into three sublevels. If the atom in in the singlet state, then there is no such splitting. In this document, we do not present calculations verifying that (254) are really the eigenfuntions of the operators $\hat{S}_z$ and $\hat{S}^2$, but we have written the hints in this course and we have been doing calculations of this kind couple of times and you ought to be able to accomplish them. The procedure for χ^1_1 is straightforward and similar to that above for $\chi^0_0 = \chi_\mathrm{A}$ and also shorter. We can then calculate the vectros χ^1_0 a χ^1_{-1} using the lowering operator $\hat{S}_-$, i.e. with the aid of formula (69b). Or, we can directly verify them as we did with χ_A. Recall that in the Dirac ket-vectors notation, the components of the triplet are written

$$|1, 1\rangle, \quad |1, 0\rangle, \quad |1, -1\rangle \tag{255}$$

Note the the three functions of the triplet are symmetric. Consequently, if we aim to form a total wave function out of them, its orbital part must be antisymmetric.

A system in any triplet state has the magnitude of the total spin described by the quantum number $S = 1$. If the state is χ^1_1, then also $S_z = 1$, which means that both electrons have their spin projections on the z axis equal to $+1/2$. We say that the spins are parallel and equally oriented (i.e. they point to the same direction), which is an inaccurate informal wording. The spins are of this character also in the state χ^1_{-1}, but then $s_z = -1/2$. (Lower-case letters are used for symbols related to one particle, upper-case for the whole system.) The spin projections in the state χ^1_0 are mutually opposite but the total spin magnitude is still determined by the value $S = 1$; hence the square if the total spin magnitude is $\hbar^2 S(S + 1) = 2\hbar^2$.

An illustrative picture is that the arrow of the spin AM has equal length for each state of the triplet $\chi^1_{S_z}$ [its square is $1(1 + 1)\hbar^2 = 2\hbar^2$] given by the number $S = 1$. In the case of $S_z = 1$, this vector is oriented along the positive z-axis direction. In the case of $S_z = -1$, it is oriented to the negative direction, and in the case of $S_z = 0$, it is perpendicular to the z axis. If we, for example, looked at the state χ^1_1 from a viewpoint of a differently oriented coordinate system, we would have to write it as, e.g. ${\chi'}^1_{-1}$ (this would only be in the case if the new z axis was oriented exactly oppositely), but in general, we would have to express it as a linear combination of all the three components, i.e. ${\chi'}^1 = \sum_{S_z=-1}^{1} \chi^1_{S_z}$. Of course, it would be a triplet state also in this other coordinate system, but in general, if viewed from the different point of view, it would not have a sharp (definite) value of S_z We also see the the three vectors of a triplet are mutually of the same importance; we also say that they transform among themselves upon rotations of the coordinate system.

And why the "arrow length" of the spin AM in a triplet state in not equal to the value of $1/2 + 1/2 = 1$ (in $\hbar = 1$ units)? After all, it is about two electrons and if they point their spin AM to the same directions, we should obtain $1/2 + 1/2 = 1$. It is because of the quantum-mechanical uncertainty. It is true that the "arrow length", more practically its square, is determined by the operator $\hat{S}^2 = \hat{S}^2_x + \hat{S}^2_y + \hat{S}^2_z$. It is the QM expectation value of this operator in the given state; shortly $\langle\hat{S}^2\rangle$. But because the cartesian components $\hat{S}_x$, $\hat{S}_y$, $\hat{S}_z$ do not commute among themselves, it is impossible to determine sharp (definite) values S_x, S_y, S_z. Consequently, it is not possible to determine the square of the "arrow length" in the classical way as $S^2_x + S^2_y + S^2_z$. We have $S_z = 1$ for the state χ^1_1 and we might think that we just need to take $S_x = S_y = 0$ and in this way determine the square of the arrow length; we would get 1 which is a wrong value because the correct one for the triplet state is $S(S + 1) = 2$. The value of 2 is also the QM expectation value of this operator in any triplet state: $\langle\chi^1_{S_z}|\hat{S}^2|\chi^1_{S_z}\rangle = 2$ (in the units of $\hbar^2$).

There are no higher multiplets for the helium atom because two electrons cannot yield a total spin greater than 1. For this, we would need a system of more electrons, e.g. the lithium atom at least.

Any function of a triplet is orthogonal to the singlet function:

$$\langle\chi^1_{S_z}|\chi^0_0\rangle = \sum_{\sigma_1=\pm1}\sum_{\sigma_2=\pm1}\chi^{1*}_{S_z}(\sigma_1,\sigma_2)\,\chi^0_0(\sigma_1,\sigma_2) = 0 \tag{256}$$

It can easily be verified and it generally follows from the fact that they are eigenfunctions corresponding to different eigenvalues of a hermitian operator, specifically of $\hat{S}^2$. The individual components of the triplet are mutually orthogonal too:

$$\langle\chi^1_{S_z}|\chi^1_{S'_z}\rangle = \delta_{S_z,S'_z} \tag{257}$$

Other multiplets. Hélium aj iné dvojelektrónové sústavy teda umožňujú len existenciu singletných a tripletných spinových stavov elektrónového obalu, teda stavov s celkovými elektrónovými spinmi S buď 0 alebo 1. Jednoelektrónová sústava, ako napr. atóm vodíka, má zasa celkový elektrónový spin S rovný vždy $1/2$, čo umožňuje dva rôzne priemety spinu na os z (dve rôzne hodnoty S_z), a túto dvojicu spinových stavov nazývame v angličtine aj v iných jazykoch ***doublet*** (čítaj dablet). Pre lítium ($N = 3$) by sme mali aj stavy s $S = 3/2$, čo dáva počet rôznych priemetov na os z $2S + 1 = 4$, a takáto štvorica spinových stavov sa v odbornej literatúre nazýva ***quartet***. Pre vyššie multiplety sa pri štúdiu elektrónovej štruktúry (ale aj u NMR) často stretávame aj s názvami ***quintet*** a ***sextet***, popr. aj ďalšími.

9 The Hartree-Fock Approximation

For an N-electron system, the complete non-relativistic Hamiltonian with added spin contributions takes the form[36]

$$\hat{H} = \sum_{i=1}^{N}\left[-\frac{\hbar^2}{2m}\nabla_i^2 + \hat{v}_{\text{ext}}(\vec{r}_i)\right] + \frac{1}{2}\sum_{i,j=1}^{N}{}' \frac{e^2}{4\pi\varepsilon_0}\frac{1}{|\vec{r}_i-\vec{r}_j|} + \hat{H}^{\text{spin}} \tag{258}$$

where the prime (the comma) at the summation over i, j indicates that cases with $i = j$ are not accounted for. The term with the operators $\hat{v}_{\text{ext}}(\vec{r}_i)$ represents the electrostatic interaction of the i-th electron with the given (*external*) potential, which itself is a sum of the Coulomb potentials due to the nuclei and, in addition, it may include contributions from possible other sources (e.g. a capacitor field, into which the molecule may be inserted:[37]

$$\hat{v}_{\text{ext}}(\vec{r}) = -\sum_I \frac{1}{4\pi\varepsilon_0}\frac{Z_I e^2}{|\vec{r}-\vec{R}_I|} + \hat{v}_{\text{other}}(\vec{r}) \tag{259}$$

where I labels the individual nuclei of the molecule or a crystal, which we assume to be fixed, i.e. motionless (which often is a good approximation) and are positioned at the points $\vec{R}_I$. The charges of the nuclei are $Z_I e$. Even if the nuclei were moving slowly, our description using the stationary SchE would usually be satisfactory, because with slow motion, the electronic structure manages to adapt to the instantaneous positions of the nuclei. Neglect of the kinetic energy of the nuclei is called the *Born-Oppenheimer approximation*. We did not include the Coulomb energy of the nuclei into the Hamiltonian (258), i.e. the term

$$\frac{1}{2}\sum_{\substack{I,J\\ I\neq J}} \frac{1}{4\pi\varepsilon_0}\frac{Z_I Z_J\, e^2}{|\vec{R}_I-\vec{R}_J|} \tag{260}$$

With constant positions of the nuclei, this term represents only a constant shift in the total energy of the system and its inclusion would therefore be trivial. Hence, we will omit it. However, it would be necessary to include it at least in cases where we would study e.g. dependence of the energy of the ground state of the molecule on the positions of the nuclei, vibrational frequencies, etc. An external magnetic field is not included in Hamiltonian (258).

---

[36] The existence of spin follows from the relativistic theory. We do not study a relativistic theory, but since we need to include spin into the description, we added it into the theory "by hand" (we postulated it in Chapter 7 based on known experimental facts and also on what we have learnt about the angular momentum earlier in section 4.5).

[37] The external field in which an electron is located is therefore, in this context, understood to be the electrostatic field generated by the nuclei and by any other external sources.

Our task is to find the solution of the problem

$$\hat{H}\Psi = E\Psi \tag{261}$$

at least for the ground state. For Hamiltonian (258), we cannot find an exact solution of this equation mainly due to the term with electron-electron repulsion. The Hartree method and especially the Hartree-Fock one makes it possible to find at least an approximate solution.

9.1 Motivation for Further Steps

The task to finding the eigenstates of Hamiltonian (258) would be greatly simplified if we could replace this Hamiltonian with an effective Hamiltonian of the form that is a sum of one-particle terms:

$$\hat{H}^{\text{eff}} = \sum_{i=1}^{N} \hat{h}_i^{\text{eff}}(\vec{r}_i, \sigma_i) \tag{262}$$

This form implies as if the individual electrons were not interacting with each other, and therefore we often briefly call such a Hamiltonian a ”non-interacting Hamiltonian.” In fact, such or a formally similar Hamiltonian may involve at least an indirect interaction between electrons, as we will see later. For the sake of generality, we now consider possible dependence of the Hamiltonian on spin, although we will abandon it later; in traditional expositions of the Hartree method (HM) and the Hartree-Fock method (HFM), spin usually does not enter the Hamiltonian. However, its inclusion would not present a difficulty. We will label the spatial and spin coordinates in different ways, depending on what is convenient for particular purpose: $(\vec{r}_i, \sigma_i) \equiv (x_i) \equiv (i)$.

Assume that we know the solutions of each of the effective one-particle Hamiltonians:

$$\hat{h}_i^{\text{eff}}\phi_i(i) = \mathcal{E}_i\phi_i(i) \tag{263}$$

(For this is a one-particle problem, solutions to it would not be, hopefully, difficult to determine at least numerically.) It can then be easily shown that the solution for the total Hamiltonian (262) is the product function

$$\Psi_{\text{HP}}(1, 2, \ldots, N) \;=\; \phi_1(1)\,\phi_2(2)\;\ldots\;\phi_N(N) \tag{264}$$

that is

$$H^{\text{eff}}\,\Psi_{\text{HP}} = E\,\Psi_{\text{HP}} \tag{265}$$

with the eigenenergy being the sum of the energies of the individual electrons:[38]

$$E = \sum_{i=1}^{N} \mathcal{E}_i \tag{266}$$

A wave function of the form (264) is called ***Hartree product***.

9.2 Antisymmetrisation of the Wave Function

Although the wave function of the form (264) may also be an exact solution for the effective Hamiltonian (262), it is not physically satisfactory because it is not antisymmetric. It is not even symmetric, so it does not respect the indistinguishability of electrons at all. (However, it would be symmetric at least if we chose functions of the same form for ϕ_1, ..., ϕ_N.) We will show how to make an antisymmetric function from it.

$\boldsymbol{N = 2}$.

$$\phi_1(1)\,\phi_2(2) \;\longrightarrow\; \frac{1}{\sqrt{2}}[\phi_1(1)\,\phi_2(2) - \phi_1(2)\,\phi_2(1)] \;=\; \frac{1}{\sqrt{2}}\begin{vmatrix} \phi_1(1) & \phi_2(1) \\ \phi_1(2) & \phi_2(2) \end{vmatrix} \;=\; \Psi(1,2) \tag{267}$$

---

[38] It will be somehow more difficult in the HF method.

$N = 3$. This is more complicated here, so we will first show how to make a symmetric function from the Hartree product and then an antisymmetric one. From the trio $1, 2, 3$ representing the coordinate of electrons, we create all possible permutations (physically representing mutual exchanges of electrons):

$$123 \quad 132 \quad 231 \quad 213 \quad 312 \quad 321 \tag{268}$$

We write down the symmetrised function as follows:

$$\begin{aligned} \Psi_{\text{sym}} = \mathcal{N} \big[& \phi_1(\mathbf{1})\, \phi_2(\mathbf{2})\, \phi_3(\mathbf{3}) + \phi_1(\mathbf{1})\, \phi_2(\mathbf{3})\, \phi_3(\mathbf{2}) + \\ & + \phi_1(\mathbf{2})\, \phi_2(\mathbf{3})\, \phi_3(\mathbf{1}) + \phi_1(\mathbf{2})\, \phi_2(\mathbf{1})\, \phi_3(\mathbf{3}) + \\ & + \phi_1(\mathbf{3})\, \phi_2(\mathbf{1})\, \phi_3(\mathbf{2}) + \phi_1(\mathbf{3})\, \phi_2(\mathbf{2})\, \phi_3(\mathbf{1}) \big] \end{aligned} \tag{269}$$

where $\mathcal{N}$ is a normalisation constant such that $\langle \Psi | \Psi \rangle = 1$. It is easy to convince that the antisymmetric function will be similar to the last one, but in such a way that at each odd permutation the sign of the respective term will be changed: [39]

$$\begin{aligned} \Psi(1,2,3) \ = \ \frac{1}{\sqrt{6}} \big[& \phi_1(\mathbf{1})\, \phi_2(\mathbf{2})\, \phi_3(\mathbf{3}) - \phi_1(\mathbf{1})\, \phi_2(\mathbf{3})\, \phi_3(\mathbf{2}) + \\ & + \phi_1(\mathbf{2})\, \phi_2(\mathbf{3})\, \phi_3(\mathbf{1}) - \phi_1(\mathbf{2})\, \phi_2(\mathbf{1})\, \phi_3(\mathbf{3}) + \\ & + \phi_1(\mathbf{3})\, \phi_2(\mathbf{1})\, \phi_3(\mathbf{2}) - \phi_1(\mathbf{3})\, \phi_2(\mathbf{2})\, \phi_3(\mathbf{1}) \big] \end{aligned} \tag{270}$$

For instance, the 123 permutation is 0th, that is even. The 321 permutation is formed by one interchange of the order of the electrons 1 and 3 compared to the originally defined order (123), so it is odd. We get the 312 permutation from the original order by two consecutive elementary swaps, so it is an even permutation. Possibly by four swaps, but it's also an even number. Again, it is easy, albeit a little longer, to be convinced that we can also write antisymmetric three-electron function (270) as a determinant:

$$\Psi(1,2,3) \ = \ \frac{1}{\sqrt{6}} \begin{vmatrix} \phi_1(1) & \phi_2(1) & \phi_3(1) \\ \phi_1(2) & \phi_2(2) & \phi_3(2) \\ \phi_1(3) & \phi_2(3) & \phi_3(3) \end{vmatrix} \tag{271}$$

A General N. There is $N!$ permutations of N numbers. We produce the antisymmetric function from the Hartree product (264) as a linear combination of $N!$ terms, which we write as follows:

$$\Psi(1,2,\ldots,N) \ = \ \frac{1}{\sqrt{N!}} \sum_{p=0}^{N!-1} (-1)^p \ \phi_1(1p)\ \phi_2(2p)\ \ldots \phi_N(Np) \tag{272}$$

The symbols $1p$, ..., Np are numbers obtained by p^{th} permutation of the original order 1, ..., N. The zeroth permutations represents the original order. It does not matter what the order of the other permutations is. Only the parity of a permutation is important (whether it is even or odd), and the parity does not depend on the chosen order. For illustration, e.g. for $N = 4$ in brief symbolic notation,

$$\begin{aligned} \Psi(1,2,3,4) = \frac{1}{\sqrt{4!}} \Big(\ & 1234 - 1243 + 1342 - 1324 + 1423 - 1432 + \\ & + 2143 - 2134 + 2314 - 2341 + 2431 - 2413 + \\ & + 3124 - 3142 + 3241 - 3214 + 3412 - 3421 + \\ & + 4132 - 4123 + 4213 - 4231 + 4321 - 4312 \quad \Big) \end{aligned}$$

By ***permutation*** in these notes we mean either the whole ordered N-tuple of numbers, or just an elementary swapping of a pair of numbers to get a given N-tuple; the specific meaning needs to be understood from the context.

---

[39] An equivalent alternative would be to change the sign at each even permutation.

However, we will get the same resulting antisymmetric wave function as (272) also in the case if we interchange just the indices of the wave functions and will keep the order of the coordinates untouched (original):

$$\Psi(1,2,\ldots,N) \;=\; \frac{1}{\sqrt{N!}} \sum_{p=0}^{N!-1} (-1)^p \; \phi_{1p}(1)\; \phi_{2p}(2)\; \ldots \phi_{Np}(N) \tag{273}$$

The wave function (273) can again be written as a determinants, as it is taught in algebra:

$$\Psi(1,2,\ldots,N) \;=\; \frac{1}{\sqrt{N!}} \begin{vmatrix} \phi_1(1) & \phi_2(1) & \ldots & \phi_N(1) \\ \phi_1(2) & \phi_2(2) & \ldots & \phi_N(2) \\ \ldots & \ldots & \ldots & \ldots \\ \phi_1(N) & \phi_2(N) & \ldots & \phi_N(N) \end{vmatrix} \tag{274}$$

A wave function expressed in this way is called a ***Slater determinant***. It is a well-known property of determinants from algebra that if we interchange two rows, the sign of the determinant changes.[40] An interchange of two rows, as can be seen, physically corresponds to an interchange of two electrons. Therefore, the Slater determinant really ensures that the wave function changes its sign to the opposite upon interchange of any two electrons.

In order not to have to write long formulas of the type (273) or (274) every time, we sometimes use the expression

$$\Psi(1,2,\ldots,N) \;=\; \hat{A}\,[\phi_1(1)\;\phi_2(2)\;\ldots\phi_N(N)] \tag{276}$$

where $\hat{A}$ is the anisymmetrisation operator. The way how its acts on a function (on a Hartree product), is obvious from the above explanation.

In section 8.6 we learnt that two electrons cannot occupy the same one-particle state. Therefore, for instance, the state

$$\Psi(1,2,3) = \hat{A}\,[\phi(1)\;\phi(2)\;\phi_3(3)]$$

(in which the first two spinorbitals are equal, i.e. $\phi_1 = \phi_2 = \phi$), should have zero probability of its realisation. If we look at expression (270), we see that in such a case Ψ is really zero. Again, this property of determinants is know from algebra: if any two functions of the Hartree product are the same, the corresponding Slater determinant is identically zero. Thus, the choice of the wave function in the form of a Slater determinant will also ensure the fulfilment of the Pauli principle and, as can be seen, it is closely related to the fact that such a wave function is antisymmetric.

At the end let us note that if Hartree product (264) is an eigenstate of the effective Hamiltonian (262),

$$H^{\text{eff}}\,\Psi_{\text{HP}} = E\,\Psi_{\text{HP}}$$

then also any state created by some permutation of the Hartree product will be an eigenfunction corresponding to the same eigenenergy. Then we come to the conclusion that the corresponding antisymmetric wave function $\Psi = \hat{A}\Psi_{\text{HP}}$ is also an eigenfunction of this Hamiltonian and corresponds to the same energy E . And let us summarise our findings of this section: we have learnt how to produce a wave function of the correct permutation symmetry from the Hartree product (264).

---

[40] Also if we interchange two columns. However, if we are talking about the exchange of *coordinates* of electrons, then for the form (274), the exchange of rows is important. However, the columnar Slater determinant is often used, i.e.

$$\Psi(1,2,\ldots,N) \;=\; \frac{1}{\sqrt{N!}} \begin{vmatrix} \phi_1(1) & \phi_1(2) & \ldots & \phi_1(N) \\ \phi_2(1) & \phi_2(2) & \ldots & \phi_2(N) \\ \ldots & \ldots & \ldots & \ldots \\ \phi_N(1) & \phi_N(2) & \ldots & \phi_N(N) \end{vmatrix} \tag{275}$$

It is identically equal to determinant (274). For the determinant (275), an interchange of two electrons is mathematically represented by the interchange of two columns. It is just a matter of preference whether we use the form (274) or (275).

9.3 The Hartree-Fock Self-Consistent Field Method

The Task To Be Solved. We have to find an approximate ground state of the N-electron system described by the Hamiltoian

$$\boxed{\hat{H} = \sum_{i=1}^{N} \hat{h}(i) + \frac{1}{2} \sum_{\substack{i,j=1 \\ i\neq j}}^{N} \frac{1}{|\vec{r}_i - \vec{r}_j|}} \tag{277}$$

where

$$\boxed{\hat{h}(i) \equiv \hat{h}(\vec{r}_i) = -\frac{\nabla_i^2}{2} + \hat{v}_{\text{ext}}(\vec{r}_i)} \tag{278}$$

Thus, we will be using the atomic units in this section. The Hamiltonian defines the basis physical parameters (positions and charges of the nuclei and the number of electrons) of the task to be solved. As we can see, the task is a little bit simpler comparing the operator (258) since we neglect the spin-dependent terms. This is, however, often an excellent or very good approximation, especially for lighter atoms.[41] This time we will denote the many-particle eigenfunction to be determined as Φ, not Ψ. Hence, we have to solve the task $\hat{H}\Phi = E\Phi$ [rov. (261)].

The Proposed Form of The Solution. In the Hartree-Fock method (HFM), we will search the wave function Φ *in a form of the Slater determinant* (274), that is, the function will obey the antisymmetry requirement.[42] Writing its arguments in detail, it is

$$\Phi(\vec{r}_1, \sigma_1, \vec{r}_2, \sigma_2, \ldots, \vec{r}_N, \sigma_N).$$

For practical reasons, this determinant is often written not as usually but using the expansion (273):

$$\boxed{\Phi(1, 2, \ldots, N) \;=\; \frac{1}{\sqrt{N!}} \sum_{p=0}^{N!-1} (-1)^p \; \phi_{1p}(1) \; \phi_{2p}(2) \; \ldots \phi_{Np}(N)} \tag{279}$$

The spinorbitals $\phi_1(\vec{r}, \sigma), \ldots, \phi_N(\vec{r}, \sigma)$ are the basic building blocks in out construction of the many-body wave function Φ. They are unknown functions to be determined.

The Functional Representing the Energy for a Chosen Wave Function (279). HFM is a particular realisation of the variational method; see section 6.1. In the sense of this method, we will consider the function (279) as a trial function on which the variational method will be applied. The role of the variational parameters will be played by the one-particle spinorbitals[43] ϕ_i. If we ensured the value of the denominator of the fraction (167) equal to 1, we could determine the total energy of the system by minimisation of the expression

$$\mathcal{G} = \int \Phi^*(1, \ldots, N) \; \hat{H} \; \Phi(1, \ldots, N) \, \mathrm{d}\tau \geq E_0 \tag{280}$$

where $\mathrm{d}\tau \equiv \mathrm{d}x_1 \mathrm{d}x_2 \ldots \mathrm{d}x_N$ means integration over the spatial coordinates of all the electrons and also summation over their spin coordinates. It is because at the correct normalisation

$$\int \Phi^*(1, \ldots, N) \, \Phi(1, \ldots, N) \, \mathrm{d}\tau = 1 \tag{281}$$

$\mathcal{G}$ would represent the quantum-mechanical expectation value of the energy of the system being in the state Φ. Because the expression $\mathcal{G}$ depends on *functions* (by which we mean ϕ_i), it is called a ***functional***. We will ensure the

---

[41] By not considering the spin degrees of freedom *in the Hamiltonian*, we consider a purely non-relativistic Hamiltonian. It is because the presence of spin follows from the relativistic quantum electrodynamics. This still does not prevent us from considering the spin degrees of freedom in wave functions.

[42] We just note that there is also the ***Hartree method*** (HM). In it, the wave function is sought in the form of the Hartree product. This is not sufficient for electronic structure, but the HM would be suitable as an intermediate step to the HF method; after studying it, the HFM could be mastered more easily and better understood. Due to lack of time, we cannot focus on the HM, but you can see it in the Appendix E. Since it assumes a simpler form of the wave function, it is a simpler method than the HFM.

[43] In reality, even those will later be expressed as linear combinations of some known functions. Thus, the corresponding coefficients will become the variational parameters, which is practical.

normalisation of Φ by demanding orthonormality of the spinorbitals:

$$\boxed{\int \phi_i^*(\vec{r},\sigma)\,\phi_j(\vec{r},\sigma)\,\mathrm{d}x \;\equiv\; \langle\phi_i|\phi_j\rangle \;=\; \delta_{ij}}, \qquad \forall i,j \tag{282}$$

These conditions are a common element in the construction of the HF method [7, 8] and facilitate its derivation. The conditions of the orthonormality will be satisfied using Lagrange multipliers. Therefore, we define an augmented functional

$$E_{\text{HF}} = \mathcal{G} - \sum_{i,j=1}^{N} \lambda_{ij}\left[\langle\phi_i|\phi_j\rangle - \delta_{ij}\right] \tag{283}$$

in which λ_{ij} are the mentioned multipliers. Thus, instead of the simpler functional (280), we will minimise E_{HF}. The one-particle functions ϕ_i standing in (279) are unknown and our task is to determine them so that the value of E_{HF} can be as low as possible. Hence, the spinorbitals ϕ_i have, at least formally, a role of variational parameters. It is practical to split the whole functional (283) to several terms and then simplify as follows:

$$\boxed{E_{\text{HF}} \;=\; \mathcal{G} + \mathcal{L} \;=\; \mathcal{G}^{(1)} + \mathcal{G}^{(2)} + \mathcal{L}} \tag{284}$$

where

$$\mathcal{G}^{(1)} \equiv \int \Phi^*(1,\ldots,N)\left[\sum_{i=1}^{N}\hat{h}(i)\right]\Phi(1,\ldots,N)\,\mathrm{d}x_1\ldots\mathrm{d}x_N \tag{285}$$

is the contribution from the single-particle terms of the Hamiltonian,

$$\mathcal{G}^{(2)} \equiv \int \Phi^*(1,\ldots,N)\left[\frac{1}{2}{\sum_{i,j=1}^{N}}{}'\,\frac{1}{r_{ij}}\right]\Phi(1,\ldots,N)\,\mathrm{d}x_1\ldots\mathrm{d}x_N \tag{286}$$

is the contribution from the two-particle terms of the Hamiltonian. The prime at the summation over i, j indicates that the cases of $i = j$ are omitted. The third component of the functional E_{HF} is the term responsible for the orthonormality of the spinorbitals:

$$\mathcal{L} = -\sum_{i,j=1}^{N} \lambda_{ij}\left[\langle\phi_i|\phi_j\rangle - \delta_{ij}\right] \tag{287}$$

We now explicitly substitute the Slater determinant (279) into these three components of the Hartree-Fock functional E_{HF}. We start with $\mathcal{G}^{(1)}$ and obtain

$$\mathcal{G}^{(1)} \;=\; \sum_{i=1}^{N}\int \phi_i^*(1)\,\hat{h}(1)\,\phi_i(1)\,\mathrm{d}x_1 \;\equiv\; \sum_{i=1}^{N}\langle\phi_i|\hat{h}|\phi_i\rangle \tag{288}$$

In order to work from definition (285) toward the final expression, eq. (288), it was necessary in particular:

- Explicitly use the antisymmetric function (279).
- Utilise orthonormality of the spinorbitals, thanks to which only the term with $p' = p$ of the sum over the permutations p' remains non-vanishing.
- Later, it was necessary to realise in a certain step of the manipulations that in the resulting sum over the permutations (for any chosen i), each term (each value) is repeated $(N-1)!$ times. Therefore, it was possible to express the summation as follows:

$$\sum_{p=0}^{N!-1}\text{something} \;=\; (N-1)!\sum_{j=1}^{N}\text{something similar, depending on the index } j,$$
$$\text{independent of } i.$$

- There is still the sum over i there that was to the left of the sum over the permutations. The sum over i can be calculated trivially for nothing depends on the index i there.
- At the end, only the sum over j will be left there. For the elegance of the final result, we denote this summation index to i.

Although we do not take over the Hartree method (HM), we will at least mention that we would also come to a term of the form (288) there, and much easier, because in the HM, the wave function has only one term.

In a similar, although a more complex way, we now begin to express the two-particle contribution $\mathcal{G}^{(2)}$ to the total functional E_{HF}. The derivation is even longer and really more difficult, but with proper notation and consistency in the individual steps of the calculation, it is not extremely difficult. Here are at least a few key points.

- Again, it is necessary to use the orthonormality of the spinorbitals, thanks to which only two non-zero terms remain from the sum over the permutations p' (for given i and j):
  - the one with $p' = p$ [as it also was in the derivation of $\mathcal{G}^{(1)}$]
  - also the term created by such a permutation p' that has only the indices i and j interchanged compared to the ordering in the permutation p; We can denoted it as p'_0. Thus, if the permutation p is even, then the p'_0 is odd and *vice versa*. Consequently, $p + p'_0$ in the second non-vanishing term is an odd number.
- Similarly to the derivation of $\mathcal{G}^{(1)}$, here it also is necessary to realise in a certain step that, although for any chosen i, j there are $N!$ contributions in the summation over the permutations p, not all of them are different from each other. Different from each other, let us call them unique, are $N(N-1)/2$ contributions for any given pair of indices i, j. Thus, each of the unique contributions is repeated $N!/[N(N-1)/2]$ times.

We arrive at

$$\mathcal{G}^{(2)} = \frac{1}{2} \sum_{i,j=1}^{N}{}' \left[\int \phi_i^*(1)\, \phi_j^*(2)\, \frac{1}{r_{12}}\, \phi_i(1)\, \phi_j(2)\, \mathrm{d}x_1\, \mathrm{d}x_2 - \int \phi_i^*(1)\, \phi_j^*(2)\, \frac{1}{r_{12}}\, \phi_j(1)\, \phi_i(2)\, \mathrm{d}x_1\, \mathrm{d}x_2 \right] \tag{289}$$

We got the two terms (apart from the fact that there are also the summations there) because of the antisymmetry of the wave function. Therefore, the second term has exchanged electron coordinates (or the wave function indices) and has the opposite sign to the first term. In the HM, we would get only the first term; you may check (E.13). Using the Dirac bra and ket vector notation, the writing will be shorter:

$$\mathcal{G}^{(2)} = \frac{1}{2} \sum_{i,j=1}^{N}{}' \left[\left\langle \phi_i \phi_j \left| r_{12}^{-1} \right| \phi_i \phi_j \right\rangle - \left\langle \phi_i \phi_j \left| r_{12}^{-1} \right| \phi_j \phi_i \right\rangle \right] \tag{290}$$

In such a notation, however, it is necessary to remember on which variables the individual functions depend and to keep the introduced order.

Thus, we have constructed the whole functional $E_{\text{HF}} = \mathcal{G}^{(1)} + \mathcal{G}^{(2)} + \mathcal{L}$ to calculate the ground-state energy:

$$\boxed{E_{\text{HF}}[\phi] \;=\; \sum_{i=1}^{N} \langle \phi_i | \hat{h} | \phi_i \rangle + \frac{1}{2} \sum_{i,j=1}^{N}{}' \left\langle \phi_i \phi_j \left| r_{12}^{-1} \right| \phi_i \phi_j \right\rangle - \frac{1}{2} \sum_{i,j=1}^{N}{}' \left\langle \phi_i \phi_j \left| r_{12}^{-1} \right| \phi_j \phi_i \right\rangle - \sum_{i,j=1}^{N} \lambda_{ij} \left[\langle \phi_i | \phi_j \rangle - \delta_{ij} \right]} \;\geq\; E_0 \tag{291}$$

The One-Particle Contribution. The first term represents the sum of one-particle energies, which are the kinetic energies of the electrons plus their potential energies in the given external field, i.e. in the potential $\hat{v}_{\text{ext}}(\vec{r})$.

The Two-Particle Contribution $+\frac{1}{2}$ .... To know the physical meaning of the second term, let's write it down in more detail, using spatial integrals and summations over the spin coordinates. In doing this, let us return for a while to the SI units. We express the spinorbitals more explicitly employing the usual factorised form

$$\phi_i(x) = \varphi_i(\vec{r})\chi_i(\sigma) \tag{292}$$

in which the (usual spatial) orbitals are orthonormal and the spin functions too,[44] see (212). Then, placing the integrals, sums, and other symbols as practically as possible, we get

$$\begin{aligned} 2^{\text{nd}}\text{ term of (291)} &= \frac{1}{2}\sum_{i,j=1}^{N}{}' \int \mathrm{d}^3r_1\mathrm{d}^3r_2\ \varphi_i^*(\vec{r}_1)\varphi_j^*(\vec{r}_2)\ \frac{1}{4\pi\varepsilon_0}\frac{(-e)^2}{|\vec{r}_1-\vec{r}_2|}\ \varphi_i(\vec{r}_1)\varphi_j(\vec{r}_2)\ \times \\ &\quad \times \underbrace{\sum_{\sigma_1=\pm1}\sum_{\sigma_2=\pm1} \chi_i^*(\sigma_1)\chi_j^*(\sigma_2)\chi_i(\sigma_1)\chi_j(\sigma_2)}_{1} \end{aligned} \tag{293}$$

The double sum over the spins written in the second row is equal to 1 due to the *normalisation* of the spin functions; for, e.g., an electron with the coordinate σ_1, we have $\sum_{\sigma_1=\pm1}\chi_i^*(\sigma_1)\chi_i(\sigma_1)=1$ for each index i. The expression

$$-e\varphi_i^*(\vec{r}_1)\varphi_i(\vec{r}_1)=\rho_i(\vec{r}_1) \tag{294}$$

can be, in a quantum-mechanical sense[45] understood as electric charge density generated by the electron 1 in the orbital φ_i. Quite analogously we also see the density $\rho_j(\vec{r}_2)$ in (293). Therefore, the 2nd term of formula (291) can be expressed as follows:

$$\boxed{\frac{1}{4\pi\varepsilon_0}\,\frac{1}{2}\sum_{i,j=1}^{N}{}'\left\langle\phi_i\phi_j\left|r_{12}^{-1}\right|\phi_i\phi_j\right\rangle = \frac{1}{2}\sum_{i,j=1}^{N}{}'\int \mathrm{d}^3r_1\mathrm{d}^3r_2\,\frac{1}{4\pi\varepsilon_0}\,\frac{\rho_i(\vec{r}_1)\rho_j(\vec{r}_2)}{|\vec{r}_1-\vec{r}_2|}} \tag{295}$$

We see that the integral standing there is nothing but the electrostatic Coulomb energy of the interaction of two charge distributions with the spatial densities $\rho_i(\vec{r}_1)$ and $\rho_j(\vec{r}_2)$. That is, the purely classical contribution, well known from the basic physics course too! For these reasons the expression (integral)

$$\boxed{J_{ij}\ \equiv\ \left\langle\phi_i\phi_j\left|r_{12}^{-1}\right|\phi_i\phi_j\right\rangle} = \int \mathrm{d}^3r_1\mathrm{d}^3r_2\ \varphi_i^*(\vec{r}_1)\varphi_j^*(\vec{r}_2)\,\frac{1}{|\vec{r}_1-\vec{r}_2|}\ \varphi_i(\vec{r}_1)\varphi_j(\vec{r}_2)\ >\ 0 \tag{296}$$

is called the ***Coulomb integral*** [8, 11] (also Coulombic). We would obtain it also in the HM [you may check (E.13)]. Looking at its form which employs the integral in (295) we see, that, because of the equal signs of the interacting charges, it is really positive. Let us conclude our discussion of expression (295) using other words: it represents the electron-electron electrostatic Coulomb repulsion (an energy) of the given N-electron system.

The Two-Particle Contribution $-\frac{1}{2}$ .... In a similar way, let us try to find a meaning of the third term of formula (291). Doing so, we arrive at the result

$$\boxed{-\frac{1}{2}\sum_{i,j=1}^{N}{}'\left\langle\phi_i\phi_j\left|r_{12}^{-1}\right|\phi_j\phi_i\right\rangle = -\frac{1}{2}\sum_{i,j=1}^{N}{}'\,\delta_{\chi_i\chi_j}\int \mathrm{d}^3r_1\mathrm{d}^3r_2\,\frac{\varphi_i^*(\vec{r}_1)\varphi_j^*(\vec{r}_2)\varphi_j(\vec{r}_1)\varphi_i(\vec{r}_2)}{|\vec{r}_1-\vec{r}_2|}} \tag{297}$$

The factor $\delta_{\chi_i\chi_j}$ emerged there from the contribution of the spin functions thanks to their *orthogonality and normalisation.*[46] This time it is not easy to get densities from the orbitals. Therefore, the third term of formula (291) does

---

[44] Although it might seem that there may be N different spin functions χ_1, χ_2, ..., χ_N there, in fact each of them is just either χ_+ or χ_-, that is chi_{+1} or χ_{-1} in the different notation.

[45] $\varphi_i^*(\vec{r})\varphi_i(\vec{r})$ is not a classical density, but a probability density in the sense of the 1st postulate of QM. In an analogous sense, $-e|\varphi_i(\vec{r}|^2$ is then a charge density.

[46] It is not complicated and not hard to get at all, one just have to proceed carefully:

$$\sum_{\sigma_1=\pm1}\sum_{\sigma_2=\pm1}\chi_i^*(\sigma_1)\chi_j^*(\sigma_2)\chi_j(\sigma_1)\chi_i(\sigma_2) = \sum_{\sigma_1=\pm1}\chi_i^*(\sigma_1)\chi_j(\sigma_1)\underbrace{\sum_{\sigma_2=\pm1}\chi_j^*(\sigma_2)\chi_i(\sigma_2)}_{\delta_{\chi_j\chi_i}} = \delta_{\chi_i\chi_j}\delta_{\chi_j\chi_i} = \delta_{\chi_i\chi_j}$$

Thus, the sum marked by the curled bracket is nonzero only if the spin function for the spinorbital ϕ_j is the same as the spin function for the spinorbital ϕ_i. And the value of this summation, of course, no longer depends on any spin coordinate; it is just a number (0 or 1). In the same way we will evaluate the remaining summation, i.e. the one over σ_1.

not have any simple classical interpretation. It emerged from the description as a consequence of the antisymmetry of the wave function of identical fermions. Hence, it is a purely non-classical contribution. The integral

$$\boxed{K_{ij} \;\equiv\; \left\langle \phi_i\phi_j \left| r_{12}^{-1} \right| \phi_j\phi_i \right\rangle} \;=\; \delta_{\chi_i\chi_j} \int \mathrm{d}^3r_1\,\mathrm{d}^3r_2\; \varphi_i^*(\vec{r}_1)\varphi_j^*(\vec{r}_2)\,\frac{1}{|\vec{r}_1-\vec{r}_2|}\,\varphi_j(\vec{r}_1)\varphi_i(\vec{r}_2) \;\geq\; 0 \tag{298}$$

is called the ***exchange integral*** [8, 11]. Without proof, we stated that if it is nonzero, it is positive [7], similarly as the Coulomb integral is always positive. Comparing to the Coulomb integral, *the indices* of the spinorbitals i, j on the right of $1/r_{12}$ *are exchanged* in the exchange integral.

Thus, the inclusion of antisymmetry leads to the exchange integral which *lowers* the total energy, and therefore the Hartree-Fock method gives a lower, or better, ground state energy than the Hartree method, in which the exchange integrals do not appear. The term *exchange interaction* is also being used. However, it must not be misunderstood as some physical interaction, but only as a way or concept to represent the corresponding contribution to the total energy of the system described by the single-determinant wave function [7].

The factor $\delta_{\chi_i\chi_j}$ in (297) causes that ***only spinorbitals with the same spins*** contribute to the exchange interaction. It does not even make sense to consider the exchange interaction and the exchange integral for electrons or spinorbitals with different spins.

Intuitively, we can partially understand the effect of the energy lowering as a consequence of the *Pauli repulsion*: two electrons with equally oriented spins must differ in something else, otherwise they would violate the Pauli exclusion principle. Therfore, they differ in their spatial wave functions which are such (different from one another, even mutually orthogonal) that keep the electrons apart. By this, they reduce their Coulomb electrostatic energy and consequently also the total energy of the system. However, a thorough understanding of the energy lowering would also have to take into account the electrons-nuclei energies. Therefore, the above interpretation is not accurate enough.

Finally, we remind that although the exchange interaction is non-classical in nature, it is still derived from the classical electrostatic interaction, because the exchange integrals include the factor $1/|\vec{r}_1 - \vec{r}_2|$. Realise: the corresponding energy is just a certain part of the quantum-mechanical exspectation value of the potential energy operator of the electrons.

The Last Contribution to the Functional E_{HF} – it is only present to guarantee the orthonormality of the spinorbitals. If this is achieved, the term vanishes.

The conditions $j \neq i$ in the above double-summations is no longer necessary. This can be seen from the fact that the expressions in the functional E_{HF} $\left\langle \phi_i\phi_j \left| r_{12}^{-1} \right| \phi_i\phi_j \right\rangle$ and $\left\langle \phi_i\phi_j \left| r_{12}^{-1} \right| \phi_j\phi_i \right\rangle$ are subtracted from one another at $j = i$. So, their non-physicality does not matter, and if it suits us, we will consider summations without the condition $j \neq i$. In this context, let us see how the Coulomb sum (295) will change upon inclusion of the terms with $j = i$. For this purpose, it is advantageous to move the summations over i and j on the right of the integrals and calculate (now already without the prime at $\sum$)

$$\sum_{i,j=1}^{N} \rho_i(\vec{r}_1)\rho_j(\vec{r}_2) = \rho(\vec{r}_1)\rho(\vec{r}_2) \tag{299}$$

where

$$\rho(\vec{r}) = \sum_{i=1}^{N} \rho_i(\vec{r}) \tag{300}$$

is the *total* charge density from the electrons. Thus, the Coulomb contribution to E_{HF} augmented in this way can be expressed as follows:

$$\frac{1}{4\pi\varepsilon_0}\,\frac{1}{2}\sum_{i,j=1}^{N} \left\langle \phi_i\phi_j \left| r_{12}^{-1} \right| \phi_i\phi_j \right\rangle \;=\; \frac{1}{2}\int \mathrm{d}^3r_1\,\mathrm{d}^3r_2\,\frac{1}{4\pi\varepsilon_0}\,\frac{\rho(\vec{r}_1)\rho(\vec{r}_2)}{|\vec{r}_1-\vec{r}_2|} \tag{301}$$

In the context of electronic-structure methods, this quantity is called *Hartree energy*. It is again an electrostatic interaction of the charge density, but this time of the total one, i.e. it also contains the non-physical contribution of the interaction of the electron with itself (the so-called *self-interaction* or *self-energy*).

Minimisation of the Functional (and of the Energy). We want to find out at what function ϕ_i the functional E_{HF} is minimal. This is something analogous to finding the minimum of a function, when we calculate the derivative of the function. The total differential of the function vanishes at the minimum of the function. Here, however, we have to look for a minimum of the functional. Therefore, we will make variations of the functional E_{HF}, which means that we will examine how it changes (varies) upon small variations of the functions ϕ_i which it depends on. Consider the following variation of the functions ϕ_i:

$$\boxed{\phi_i \longrightarrow \phi_i + \delta\phi_i} \tag{302}$$

The functional then varies as follows:

$$\boxed{E_{\text{HF}}[\phi] \quad \longrightarrow \quad E_{\text{HF}}[\phi + \delta\phi] = E_{\text{HF}}[\phi] + \delta E_{\text{HF}}} \tag{303}$$

and it could analogously be written also for its individual components $\mathcal{G}^{(1)}$, $\mathcal{G}^{(2)}$, and $\mathcal{L}$. Hence, the sum of the one-particle integrals in (288) gets

$$\begin{aligned} \mathcal{G}^{(1)}[\phi] \quad \longrightarrow \quad \mathcal{G}^{(1)}[\phi + \delta\phi] &= \sum_{i=1}^{N} \int (\phi_i + \delta\phi_i)^* \, \hat{h}(i) \, (\phi_i + \delta\phi_i) \, \mathrm{d}x \; = \\ &= \mathcal{G}^{(1)}[\phi] + \underbrace{\sum_{i=1}^{N} \int \delta\phi_i^* \, \hat{h}(i) \, \phi_i \, \mathrm{d}x + \sum_{i=1}^{N} \int [\hat{h}(i) \, \phi_i]^* \delta\phi_i \, \mathrm{d}x \; +}_{\delta\mathcal{G}^{(1)}} \\ &\quad + \text{terms in } 2^{\text{nd}} \text{ order in } \delta\phi_k \,, \text{ which are negligible} \end{aligned} \tag{304}$$

Note that the second term in $\delta\mathcal{G}^{(1)}$ is complex conjugate (c.c.) to the first one. So, we obtain

$$\delta\mathcal{G}^{(1)} = \sum_{i=1}^{N} \langle \delta\phi_i | \hat{h} | \phi_i \rangle + \text{c.c.} \tag{305}$$

In calculations with two-particle integrals too, we will omit writing the arguments of the functions ϕ in some places for the sake of brevity. We again remind ourselves that it is necessary to remember which function depends on $x_i \equiv \vec{r}_i, \sigma_i$ (briefly i) and which on x_j. For the variation of the sum of the two-particle integrals (289) we get, using a procedure in a manner similar to the above one, only more complex, the following result [using that $r_{ij} = r_{ji}$ and that we can arbitrarily rename and interchange summation indices ($i \leftrightarrow j$) with one another]:

$$\begin{aligned} \delta\mathcal{G}^{(2)} = & \sum_{i,j=1}^{N}{}' \int \delta\phi_i^* \, \phi_j^* \, \frac{1}{r_{12}} \, \phi_i \, \phi_j \, \mathrm{d}x_1 \, \mathrm{d}x_2 + \sum_{i,j=1}^{N}{}' \int \phi_i^* \, \phi_j^* \, \frac{1}{r_{12}} \, \delta\phi_i \, \phi_j \, \mathrm{d}x_1 \, \mathrm{d}x_2 \\ & - \sum_{i,j=1}^{N}{}' \int \delta\phi_i^* \, \phi_j^* \, \frac{1}{r_{12}} \, \phi_j \, \phi_i \, \mathrm{d}x_1 \, \mathrm{d}x_2 - \sum_{i,j=1}^{N}{}' \int \phi_i^* \, \phi_j^* \, \frac{1}{r_{12}} \, \delta\phi_j \, \phi_i \, \mathrm{d}x_1 \, \mathrm{d}x_2 \end{aligned} \tag{306}$$

Note that the second terms in each row are complex conjugate to the firts terms. In the compact notation, we write down this variation as follows:

$$\delta\mathcal{G}^{(2)} = \left[\sum_{i,j=1}^{N}{}' \, \langle \delta\phi_i \, \phi_j | r_{12}^{-1} | \phi_i \, \phi_j \rangle + \text{k.z.} \right] - \left[\sum_{i,j=1}^{N}{}' \, \langle \delta\phi_i \, \phi_j | r_{12}^{-1} | \phi_j \, \phi_i \rangle + \text{k.z.} \right] \tag{307}$$

To evaluate also the variation of the functional E_{HF}, not only of $\mathcal{G}$, it remains to evaluate the variation of the term with the Largrange multiplier, see (287). This can be calculated easily, resulting to

$$\delta\mathcal{L} = - \sum_{i,j=1}^{N} \lambda_{ij} \langle \delta\phi_i | \phi_j \rangle \; + \text{k.z.} \tag{308}$$

Now we are able to write what the variation of the whole functional E_{HF}, that is the value of

$$\delta E_{\text{HF}} = E_{\text{HF}}[\phi + \delta\phi] - E_{\text{HF}}[\phi] = \delta\mathcal{G}^{(1)} + \delta\mathcal{G}^{(2)} + \delta\mathcal{L} \tag{309}$$

is equal to. To accomplish this, we need to collect the results (304), (306) and (308). We arrive at

$$\delta E_{\mathrm{HF}} = \sum_{i=1}^{N} \langle \delta\phi_i | \hat{h} | \phi_i \rangle + \sum_{i,j=1}^{N}{}' \langle \delta\phi_i \, \phi_j | r_{12}^{-1} | \phi_i \, \phi_j \rangle - \sum_{i,j=1}^{N}{}' \langle \delta\phi_i \, \phi_j | r_{12}^{-1} | \phi_j \, \phi_i \rangle - \sum_{i,j=1}^{N} \lambda_{ij} \langle \delta\phi_i | \phi_j \rangle \; + \text{c.c.} \tag{310}$$

and we better rewrite this result using integrals and at the same time we pull the common parts to the left:

$$\begin{aligned} \delta E_{\mathrm{HF}} \; &= \sum_{i=1}^{N} \int \mathrm{d}x_1 \, \delta\phi_i^*(1) \Big[\\ & \hat{h}(1)\phi_i(1) + \sum_{j=1}^{N}{}' \int \mathrm{d}x_2 \, \phi_j^*(2) \, r_{12}^{-1} \, \phi_i(1) \, \phi_j(2) - \sum_{j=1}^{N}{}' \int \mathrm{d}x_2 \, \phi_j^*(2) \, r_{12}^{-1} \, \phi_j(1) \, \phi_i(2) - \sum_{j=1}^{N} \lambda_{ij} \, \phi_j(1) \\ & \Big] \; + \text{c.c.} \end{aligned} \tag{311}$$

where c.c. denotes the terms complex conjugate to the former ones. We keep writing the primes at the sums, that is, we are omitting summations over $j = i$ although this is no longer necessary.

As we have said before, we are trying to find out for what functions ϕ_i is the functional E_{HF} is minimal. Just as a function has zero first derivative around its extremal point, ie zero change in the first order, so a functionial around its extremum has vanishing variation. Therefore, in order to find the minimising spinorbitals ϕ_i, we require

$$\delta E_{\mathrm{HF}} = 0 \tag{312}$$

To satisfy this for any small variations $\delta\phi_i$,

the expression in the square brackets (311) *must vanish.* (313)

Before we write down zeroness of the expression in the square brackets, we introduce two important terms and their notation, so that we can properly write the integrals over x_2 in the square brackets in (311).

The Coulomb Operator. The expression

$$\boxed{\hat{\mathcal{J}}_j(1) = \int \mathrm{d}x_2 \, \phi_j^*(2) \, r_{12}^{-1} \, \phi_j(2)} = \int \mathrm{d}^3 r_2 \; \varphi_j^*(\vec{r}_2) \, \frac{1}{|\vec{r}_1 - \vec{r}_2|} \, \varphi_j(\vec{r}_2) \tag{314}$$

is called the *Coulomb operátor.* Its name comes from the fact that it expresses (in the atomic units) the Coulomb electrostatic energy of an electron located at point $\vec{r}_1$ with an electron cloud that is created by the wave function ϕ_j. It can be seen from formula (314) that it is an operator expressed by *real* numerical values; it is a real function of the variable $\vec{r}_1$. Hence, the Coulombov operátor is *hermitian.*

The Exchange Operator. The second term in the square brackets of (311) is more complicated than the first one. We will define the corresponding operator indirectly only, through its effect on the spinorbital. The form of this operator will look rather artificially and is motivated by the equations being easy to write. The square brackets in the following definition (it is the framed part) are not necessary but they are convenient for clarity.

$$\boxed{\hat{\mathcal{K}}_j(1)\phi_i(1) = \left[\int \mathrm{d}x_2 \, \phi_j^*(2) \, r_{12}^{-1} \, \phi_i(2) \right] \phi_j(1)} = \left[\delta_{\chi_j \chi_i} \int \mathrm{d}^3 r_2 \; \varphi_j^*(\vec{r}_2) \, \frac{1}{|\vec{r}_1 - \vec{r}_2|} \, \varphi_i(\vec{r}_2) \right] \varphi_j(\vec{r}_1) \chi_j(\sigma_1) \tag{315}$$

The operator $\hat{\mathcal{K}}_j$ is called the *exchange operator.* Its name comes from the fact that it is located in the term of the functional that was created by exchanging the coordinates of the two electrons. The presence of this operator stems from the indistinguishability of electrons and from the requirement of antisymmetry of the wave function, and is therefore of a non-classical nature. At the same time, however, it is a manifestation of electrostatic interaction, since the expression r_{12}^{-1} is found in this operator. Although it may seem that the $\hat{\mathcal{K}}_j$ operator should also carry the index

i of the function it acts on, it is not the case.[47] $\hat{\mathcal{K}}_j(1)$ processes any function it acts on by integrating it over x_2 and also produces the function $\phi_j(1)$. It is therefore an operator that *depends on* ϕ_j, *processes* any given function (for instance ϕ_i) in a defined way and there is no reason to put the symbol of the function on that it is currently acting, on the operator notation. Directly from definitions 3 and 4 in section 1.2.4, we can easily prove that the exchange operator is also *hermitian.*

The Hartree-Fock Equations. Equation (313) can now be written as follows:

$$\boxed{\left[\hat{h}(1) + {\sum_{j=1}^{N}}' \hat{\mathcal{J}}_j(1) - {\sum_{j=1}^{N}}' \hat{\mathcal{K}}_j(1)\right] \phi_i(1) = \sum_{j=1}^{N} \lambda_{ij}\, \phi_j(1)} \tag{320}$$

This system of equations is called the *Hartree-Fock equations* (HFE) [7]. It will be possible to simplify it even more to the so-called canonical form in which $\lambda_{ij} \propto \delta_{ij}$; therefore the summation on the right-hand side will disappear. The term in the square brackets is called the ***Fock operator***:

$$\boxed{\hat{f}(1) = \hat{h}(1) + \sum_{j=1}^{N} \left[\hat{\mathcal{J}}_j(1) - \hat{\mathcal{K}}_j(1)\right]} \tag{321}$$

As mentioned above, it is not necessary to omit the values with $i = j$ in the summations. This is and important simplifications because, thanks to it, there is only one and the same Fock operator in the HFE[48] that acts on each spinorbital ϕ_i and the HFE can then be written down briefly:

$$\boxed{\boxed{\hat{f}(1)\phi_i(1) = \sum_{j=1}^{N} \lambda_{ij}\, \phi_j(1)}} \tag{322}$$

Finally, realise that the Fock operator is *hermitian.* This comes as a consequence of $\hat{h}(1)$, $\hat{\mathcal{J}}_j(1)$ and $\hat{\mathcal{K}}_j(1)$ being hermitian operators.

---

[47] The formal expression of the symbol $\hat{\mathcal{K}}_j$ depends on i of that spinorbital and therefore it might seem that the exchange operator should also get the index i:

$$\hat{\mathcal{K}}_j^{(i)}(1) = \left[\int \mathrm{d}x_2\, \phi_j^*(2)\, r_{12}^{-1}\, \phi_i(2)\right] \frac{\phi_j(1)}{\phi_i(1)} \tag{316}$$

For a better understanding of the exchange operator, realise that it can act on any function, not just the spinorbitals ϕ_i. Let us first consider any function $\phi(x)$ expressible as a linear combination of our spinorbitals:

$$\phi(x) = \sum_{k=1}^{N} c_k \phi_k(x) \tag{317}$$

Then we express the effect of the exchange operator on $\phi(x)$ due to its linearity as follows:

$$\hat{\mathcal{K}}_j(1)\phi(1) = \sum_{k=1}^{N} c_k \hat{\mathcal{K}}_j(1)\phi_k(1) = \sum_{k=1}^{N} c_k \left[\int \mathrm{d}x_2\, \phi_j^*(2)\, r_{12}^{-1}\, \phi_k(2)\right] \phi_j(1) = \left[\int \mathrm{d}x_2\, \phi_j^*(2)\, r_{12}^{-1} \sum_{k=1}^{N} c_k \phi_k(2)\right] \phi_j(1) \tag{318}$$

i.e.

$$\boxed{\boxed{\hat{\mathcal{K}}_j(1)\phi(1) = \left[\int \mathrm{d}x_2\, \phi_j^*(2)\, r_{12}^{-1}\, \phi(2)\right] \phi_j(1)}} \tag{319}$$

It can be seen that nothing prevents us from extending the definition of the exchange operator so that it can act on any function of the variable x_1, not only on the above-mentioned linear combination (317). The action of the exchange operator $\hat{\mathcal{K}}_j(1)$ must therefore be understood as creating the spinorbital $\phi_j(1)$ and multiplying it by the number $\int \phi_j^*(2)\, r_{12}^{-1}\, \phi(2)\, \mathrm{d}x_2$. This number therefore depends on the function on that $\hat{\mathcal{K}}_j(1)$ acts.

[48] This is an essential difference and perhaps also a surprising simplification in comparison to Hartree equations (E.29).

Solution of the HFE. The Self-Consistent Field. The HFE is a system of N integro-differential equation for the unknow functions ϕ_i. Thus, by solving these equations we find the functions that minimise the functional E_{HF}. In such case, the value of this functional is an approximate energy of the ground state of the system. We do not have enough time in this course to properly study the ways how to solve the HFE, so we will describe them only in a roughest sketch.

In principle, the HFE are solved using subsequent iterations: in the beginning, we choose some guess functions

$$\phi_1^{(0)}, \phi_2^{(0)}, \ldots, \phi_N^{(0)} \tag{323}$$

For instance, if we solve the HFE for an atom, we can take the exactly know spinorbitals of a hydrogen-like ion as the initial guess $\phi_i^{(0)}$. We then determine the initial Fock operator using the guess functions (the operator is certainly not good yet). Using the Fock operator, we then determine (in a manner not specified here) more accurate (although still very rough) spinorbotals

$$\phi_1^{(1)}, \phi_2^{(1)}, \ldots, \phi_N^{(1)} \tag{324}$$

Using these, we again construct the Fock operator; it should now be more closer to the accurate one. And so on, and so on, we do the iterations and once we finish. We can make the decision to finish, for example, if the difference between two outputs of the subsequent iterations becomes negligible. Then the spinorbitals ϕ_i will be consistent with the Fock operator. The resulting elecrostatic field from the considered electrons is called *self-consistent field* (SCF). Thus, each electron moves in this SCF. It is actually a field created by QM averaging of the field from the other electrons and, in addition, the field from the nuclei is also accounted for. The term *mean field*[49] is also being used, although in also a more general sense.

An alternative, and nowadays more and more frequently used way to solve the HFE, is the direct minimisation of the HF functional. The advantage of this method is its higher robustness, as it is a direct application of the variation principle.

The Energy of the Ground State. This energy is given by the minimum of the functional (291). Because ϕ_i are orthogonal, we obtain (assuming that they are the minimising spinorbitals)

$$E_{\mathrm{HF}} = \sum_{i=1}^{N} \langle \phi_i | \hat{h} | \phi_i \rangle + \frac{1}{2} \sum_{i,j=1}^{N} \left\langle \phi_i \phi_j \left| r_{12}^{-1} \right| \phi_i \phi_j \right\rangle - \frac{1}{2} \sum_{i,j=1}^{N} \left\langle \phi_i \phi_j \left| r_{12}^{-1} \right| \phi_j \phi_i \right\rangle \tag{325}$$

Using the expressions (314) and (315) for the Coulombov and exchange operators, we can also write the HF energy as follows:

$$\boxed{E_{\mathrm{HF}} = \sum_{i=1}^{N} \int \mathrm{d}x \; \phi_i^*(x) \left\{ \hat{h}(x) + \frac{1}{2} \sum_{j=1}^{N} \left[\hat{\mathcal{J}}_j(x) - \hat{\mathcal{K}}_j(x) \right] \right\} \phi_i(x)} \tag{326}$$

The difference between the exact non-relativistic energy and the Hartree-Fock energy is called the ***correlation energy***:

$$E_{\mathrm{corr}} = E_0 - E_{\mathrm{HF}} < 0 \tag{327}$$

(Here we consider a ground state only.) Its magnitude is small in comparison with the magnitude of the total energy E_0, of the order of e.g. one percent of the total energy. Nevertheless, the effects associated with electron correlation use to be important. Determining the ground state energy and wave function beyond the HF method is still a major and fundamental challenge in the study of electronic structure. In cases of large systems (large molecules, clusters or crystals), it often happens that even current methods and computational resources do not make it possible to calculate satisfactorily ground-state energy and wave function of a given system.

---

[49] also in statistical physics

9.4 A Mathematical Excursion: Matrix Representations, Unitary Transformations

We have not taken too much from this section (9.4), only those parts that are written in the larger font. By the way, these parts are just a simple generalization of what has been taken over in the ***Computational Physics*** course about diagonalisation of a symmetric matrix and expressing this diagonalization by an orthogonal transformation. The generalisation of a symmetric matrix is **hermitian** matrix. The generalisation of an orthogonal matrix is a **unitary** matrix.

Matrix representations. Majme sadu navzájom *lineárne nezávislých* funkcií φ_n tvoriacich *úplnú sústavu funkcií*; úplnú aspoň v tom zmysle, že pomocou ich lineárnej kombinácie vieme s požadovanou presnosťou vyjadriť ľubovoľnú funkciu, ktorú pri štúdiu daného problému vyjadriť potrebujeme. Tú sústavu (sadu, postupnosť) funkcií potom nazývame *báza.* Samotné funkcie voláme bázové funkcie. Okrem lineárnej nezávislosti predpokladajme aj ich vzájomnú ortogonálnosť a normovanosť na 1:

$$\int \varphi_m^* \varphi_n \, \mathrm{d}\tau = \langle \varphi_m | \varphi_n \rangle = \delta_{mn} \tag{328}$$

pričom integrujeme cez bližšie nešpecifikované premenné, od ktorých tie funkcie závisia. Tých integračných premenných môže byť veľa a nemusia byť len spojité (ako napr. x, y, z), ale môžu byť medzi nimi aj spinové súradnice, v prípade ktorých sa sumuje, nie integruje. Symboly ako napr. φ_n nazývame funkcie. Závisia od konkrétnych premenných, napr. x, y, z. Abstraktné zápisy ako $|\varphi_n\rangle$ nazývajme vektory.

Uvažujme ľubovoľný lineárny operátor $\hat{A}$. Pôsobme ním na ľubovoľnú funkciu f takú, ktorá sa dá vyjadriť ako lineárna kombinácia bázových funkcií φ_n. Výsledkom bude nejaká iná funkcia. Označíme si ju g:

$$\boxed{\hat{A} f = g}\,, \quad \text{t. j.} \quad \hat{A} |f\rangle = |g\rangle \tag{329}$$

Keďže $\{\varphi_n\}$ je úplná sústava, aj g sa musí dať vyjadriť ako ich lineárna kombinácia. Zapíšme to pre obe tie funkcie takto:

$$f = \sum_n f_n \varphi_n\,, \qquad g = \sum_n g_n \varphi_n \tag{330}$$

kde f_n a g_n sú koeficienty v tých lineárnych kombináciách (rozvojové koeficienty).[50] Dosaďme tieto rozvoje do (329). Dostaneme

$$\hat{A} \sum_n f_n \varphi_n = \sum_n g_n \varphi_n$$

Lineárny operátor prejde cez rozvojové koeficienty (nejaké komplexné čísla) triviálne podľa Definície 2 (10). Tak dostaneme

$$\sum_n f_n \hat{A} \varphi_n = \sum_n g_n \varphi_n \tag{331}$$

Pôsobením $\hat{A}$ na bázovú funkciu φ_n vznikne funkcia $\hat{A}\varphi_n$, ktorá sa tiež dá zapísať v danej báze:

$$\hat{A} \varphi_n = \sum_k A_{kn}\, \varphi_k \tag{332}$$

Príslušné rozvojové koeficienty sme teda označili A_{kn}. Index n je tam potrebný, lebo ide o rozvoj funkcie φ_n. Pri pôsobení $\hat{A}$ na inú bázovú funkciu, napr. na $\varphi_{n'}$, by sme totiž dostali iné rozvojové koeficienty ($A_{kn'}$), a preto ich treba označovať aj indexom bázovej funkcie, na ktorú pôsobia. Dosaďme rozvoj (332) do (331):

$$\sum_n f_n \sum_k A_{kn} \varphi_k = \sum_n g_n \varphi_n$$

Prenásobme túto rovnicu zľava funkciou φ_m^* a preintegrujme:

$$\sum_n \sum_k f_n A_{kn} \int \varphi_m^* \varphi_k \, \mathrm{d}\tau = \sum_n g_n \int \varphi_m^* \varphi_n \, \mathrm{d}\tau$$

Využijeme ortonormalitu (328) bázových funkcií, čím dostaneme

$$\boxed{\boxed{\sum_n A_{mn} f_n = g_m}} \tag{333}$$

To je sústava algebraických rovníc, ktoré dávajú do vzťahu tri rôzne sady rozvojových koeficientov. Sady koeficientov f_n, g_n definované rovnicami (330) sa dajú zapísať aj ako stĺpcové vektory. Koeficienty A_{mn} zasa tvoria štvorcovú maticu. Ak

---

[50] Netreba si ich teda popliesť s nejakými funkciami. Inde sme totiž symboly typu f_n používali ako bázové funkcie, pričom rozvojové koeficienty sme značili inak; pozri napr. stať 6.1.1 o variačnej metóde.

si počet bázových funkcií označíme N a indexovať začneme od 1, tak posledne napísaná sústava sa dá zapísať aj maticovo-vektorovo:

$$\begin{pmatrix} A_{11} & \dots & A_{1N} \\ \vdots & & \vdots \\ A_{N1} & \dots & A_{NN} \end{pmatrix} \begin{pmatrix} f_1 \\ \vdots \\ f_N \end{pmatrix} = \begin{pmatrix} g_1 \\ \vdots \\ g_N \end{pmatrix} \tag{334}$$

Aj (333) je maticovo-vektorový zápis, ale v zložkách. Aby sme plne pochopili aj význam koeficientov A_{mn}, vráťme sa ku rovnici (332), ktorou boli definované. Prenásobme ju zľava funkciou φ_m^* a preintegrujme:

$$\int \varphi_m^* \hat{A} \varphi_n \, \mathrm{d}\tau = \sum_k A_{kn} \int \varphi_m^* \varphi_k \, \mathrm{d}\tau \quad \Longrightarrow \quad \boxed{A_{mn} = \int \varphi_m^* \, \hat{A} \, \varphi_n \, \mathrm{d}\tau} = \langle \varphi_m | \hat{A} | \varphi_n \rangle \tag{335}$$

Keď teraz porovnáme abstraktnú rovnosť (329) so zápisom (333) alebo (334), môžeme skonštatovať, že tie maticovo-vektorové zápisy predstavujú istú *reprezentáciu* abstraktného zápisu (329). Matica A s prvkami (335) je maticovou reprezentáciou operátora $\hat{A}$ v nami použitej báze $\{\varphi_n\}$. Vektor s prvkami f_n je reprezentáciou funkcie f v danej báze. Obdobne vektor prvkov g_n.

Ak chceme explicitne vedieť, ako určíme koeficienty f_n, g_n, zoberieme definičné rovnice (330), prenásobíme ich zľava φ_m^*, preintegrujeme a dostaneme (po premenovaní indexu)

$$\boxed{f_n = \int \varphi_n^* \, f \, \mathrm{d}\tau} = \langle \varphi_n | f \rangle \, , \qquad \boxed{g_n = \int \varphi_n^* \, g \, \mathrm{d}\tau} = \langle \varphi_n | g \rangle \tag{336}$$

Ešte sa pozrime na maticovú reprezentáciu súčinu dvoch operátorov: $\hat{C} = \hat{A}\hat{B}$. Nech $\hat{C} = \hat{A}\hat{B}$ pôsobí na nejakú funkciu f, následkom čoho vznikne funkcia h: $\hat{C}f = h$. V reprezentácií pomocou danej bázy to zapíšeme

$$\sum_n C_{mn} f_n = h_m \tag{337}$$

Ako vyjadríme C_{mn} pomocou maticových prvkov A_{mn} a B_{mn}? Pomocou vyjadrenia $C_{mn} = \int \varphi_m^* \hat{A}\hat{B} \varphi_n \, \mathrm{d}\tau$ nie je ťažké sa dopracovať k výsledku

$$C_{mn} = \sum_k A_{mk} B_{kn} \tag{338}$$

t. j. ide o bežné násobenie matíc, $C = AB$, čo je aj intuitívne očakávaný výsledok.

Ešte doplňme, že podľa Vety 2 – pozri (12) – pre ľubovoľný lineárny operátor $\hat{A}$, ku ktorému existuje hermitovsky združený, platí

$$\int \varphi_m^* \hat{A} \varphi_n \, \mathrm{d}\tau = \int (\hat{A}^\dagger \varphi_m)^* \varphi_n \, \mathrm{d}\tau \quad \Longrightarrow \quad \boxed{A_{mn}^* = (A^\dagger)_{nm}} \tag{339}$$

Maticovou reprezentáciou hermitovsky združeného operátora teda je matica $A^\dagger$, ktorá je transponovaná a navyše komplexne združená ku matici A. V špeciálnom prípade, ak by sme mali nejaký hermitovský operátor $\hat{H}$, t. j. platilo by

$$\hat{H}^\dagger = \hat{H} \tag{340}$$

tak pre ich maticové reprezentácie by sme dostali

$$H_{mn}^* = H_{nm} \tag{341}$$

z čoho vyplýva, že diagonálne prvky sú reálne. To je známa a dôležitá vlastnosť hermitovských matíc a úzko súvisí s tým, že vlastné hodnoty hermitovského operátora sú reálne čísla.

Unitary Transformations.

Definícia 7: *Lineárny spojitý [2] operátor $\hat{U}$ sa nazýva unitárny, ak k nemu existuje inverzný operátor $\hat{U}^{-1}$ a platí*

$$\hat{U}^\dagger \hat{U} = \hat{1} \tag{342}$$

kde $\hat{1}$ je jednotkový operátor.

Čiže $\hat{U}^{-1} = \hat{U}^\dagger$ a následne dostávame, že platí aj $\hat{U}\hat{U}^\dagger = \hat{1}$. Pre maticové reprezentácie v nejakej zvolenej ortonormovanej báze potom musí platiť

$$U^\dagger U = U U^\dagger = I \tag{343}$$

kde I je štvorcová matica, ktorá má na diagonále jednotky a všade inde nuly (jednotková matica). Významnou matematickou vlastnosťou unitárnych operátorov je, že zachovávajú skalárny súčin. Nech $|f\rangle$ a $|g\rangle$ sú nejaké vektory. Unitárne transformované vektory sú

$$|f'\rangle = \hat{U}|f\rangle \, , \qquad |g'\rangle = \hat{U}|g\rangle \tag{344}$$

Potom

$$\langle f'|g'\rangle = \int f'^* g' \,\mathrm{d}\tau = \int \left(\hat{U}f\right)^* \left(\hat{U}g\right) \,\mathrm{d}\tau = \int f^* \hat{U}^\dagger \left(\hat{U}g\right) \,\mathrm{d}\tau = \int f^* g \,\mathrm{d}\tau$$

teda

$$\langle f'|g'\rangle = \langle f|g\rangle \tag{345}$$

Aj maticový element nejakého operátora $\hat{A}$, teda číslo

$$\langle f|\hat{A}|g\rangle$$

je skalárny súčin, konkrétne súčin vektorov $|f\rangle$ a $\hat{A}|g\rangle$. Preto sa maticové elementy musia unitárnou transformáciou zachovávať. Počítajme, čo z toho vyplynie.

$$\langle f|\hat{A}|g\rangle = \int f^* \hat{A}\, g \,\mathrm{d}\tau = \int f^* \hat{U}^\dagger \underbrace{\hat{U}\hat{A}\hat{U}^\dagger}_{\hat{A}'} \underbrace{\hat{U}g}_{g'} \,\mathrm{d}\tau = \int f^* \hat{U}^\dagger \hat{A}' g' \,\mathrm{d}\tau = \int \left(\hat{U}f\right)^* \hat{A}' g' \,\mathrm{d}\tau = \int f'^* \hat{A}' g' \,\mathrm{d}\tau$$

teda

$$\boxed{\langle f'|\hat{A}'|g'\rangle = \langle f|\hat{A}|g\rangle} \tag{346}$$

Zavedené označenie

$$\hat{A}' = \hat{U}\hat{A}\hat{U}^\dagger \tag{347}$$

teda treba rozumieť ako unitárne transformovaný operátor $\hat{A}$. Operátory sa teda unitárne transformujú tak, že ich treba násobiť zľava aj sprava unitárnymi operátormi $\hat{U}$ a $\hat{U}^\dagger$. A tak isto to bude aj s maticovými reprezentáciami operátorov: ak A je matica reprezentujúca operátor $\hat{A}$ a U, $U^\dagger$ sú matice reprezentujúce unitárne operátory $\hat{U}$, $\hat{U}^\dagger$, tak unitárne transformovaná matica bude $A' = UAU^\dagger$.

An extremely important feature of hermitian matrices is that they can be diagonalised by unitary transformations. It should be understood as follows: Let H be a hermitian $N \times N$ matrix. Then is has N eigenvalues, in general different from each other. Denote them by $\mathcal{E}_1, ..., \mathcal{E}_N$, According to Theorem 4 of section 1.2.5, they are real. The corresponding eigenvectors are columnar. Suppose that j-th eigenvector has its components denoted by X_{ij}, $i \in \{1, \ldots, N\}$, and, of course, also $j \in \{1, \ldots, N\}$.

$$\begin{pmatrix} H_{11} & \ldots & H_{1N} \\ \vdots & & \vdots \\ H_{N1} & \ldots & H_{NN} \end{pmatrix} \begin{pmatrix} X_{1j} \\ \vdots \\ X_{Nj} \end{pmatrix} = \mathcal{E}_j \begin{pmatrix} X_{1j} \\ \vdots \\ X_{Nj} \end{pmatrix} , \quad j \in \{1, \ldots, N\} \tag{348}$$

According to Theorem 5, in section 1.2.5, the individual vectors are orthogonal to each other and we will normalise them to 1, so they will be orthonormalised. When we stack all the N eigenvectors next to each other, we get the square matrix

$$X = \begin{pmatrix} X_{11} & \ldots & X_{1N} \\ \vdots & & \vdots \\ X_{N1} & \ldots & X_{NN} \end{pmatrix} \tag{349}$$

We express the mentioned orthonormality of the eigenvectors, specifically of the i-th one to the j-th one as follows:

$$(X^*_{1i}, \ldots, X^*_{Ni}) \begin{pmatrix} X_{1j} \\ \vdots \\ X_{Nj} \end{pmatrix} = \delta_{ij} , \qquad \text{i.e.} \quad \sum_{k=1}^{N} X^*_{ki} X_{kj} = \delta_{ij} , \qquad i, j \in \{1, 2, \ldots, N\} \tag{350}$$

The last written system of equations for the orthogonality can be rewritten in the form

$$\sum_{k=1}^{N} (X^\dagger)_{ik} X_{kj} = \delta_{ij} , \qquad i, j = 1, 2, \ldots, N \tag{351}$$

using the compact notation $X^\dagger X = I$, which corresponds to the definition of the unitary operator [*cf.* (342)]; specifically, we now have the operator represented by the matrix X. I is the identity matrix[51] $N \times N$. Thus, we have proved this important finding:

---

[51] not a *unit* or *unity* matrix! A unit matrix is such that has units (1) everywhere.

Eigenvectors of a hermitian matrix form a unitary matrix.

Using the X, (348), which is a system of systems of equations, can be written as follows:

$$\begin{pmatrix} H_{11} & \dots & H_{1N} \\ \vdots & & \vdots \\ H_{N1} & \dots & H_{NN} \end{pmatrix} \begin{pmatrix} X_{11} & \dots & X_{1N} \\ \vdots & & \vdots \\ X_{N1} & \dots & X_{NN} \end{pmatrix} = \begin{pmatrix} X_{11} & \dots & X_{1N} \\ \vdots & & \vdots \\ X_{N1} & \dots & X_{NN} \end{pmatrix} \begin{pmatrix} \mathcal{E}_1 & \dots & 0 \\ \vdots & & \vdots \\ 0 & \dots & \mathcal{E}_N \end{pmatrix} \tag{352}$$

or compactly

$$\boxed{HX = X\mathrm{diag}(\mathcal{E})} \tag{353}$$

where $\mathrm{diag}(\mathcal{E})$ is a matrix that has the values $\mathcal{E}_1$, ..., $\mathcal{E}_N$ on its diagonal and zeros everywhere else. Let us now multiple the last equation by the matrix $X^\dagger$ from the left. Because it is unitary, we get

$$\boxed{\boxed{\mathrm{diag}(\mathcal{E}) = X^\dagger H X}} \tag{354}$$

Thus, the diagonal matrix of the eigenvalues of the hermitian matrix H is obtained by a certain unitary transformation of the matrix H. Specifically, by such a unitary transformation, whose matrix is equal to (349). Therefore, calculation of eigenvalues of a matrix is often being called ***diagonalisation*** of the matrix. In doing this, we can also calculate the eigenvectors. Usually we can calculate both the eigenvalues and the eigenvectors with one iterative algorithm [9].

A special case of a hermitian matrix is the *symmetric matrix*. This is when H is real. And analogously, a special case of a unitary matrix is the *orthogonal matrix*. This one is also real. Thus, a symmetric matrix can be diagonalised by an orthogonal transformation.

9.5 The Canonical Form of the Hartree-Fock Equations

`Again, in this part you only need to look at what is typeset in the larger font.`

Imagine that we have solved the HF equations, (322). Hence we know the single-particle spinorbitals ϕ_i as well as the matrix λ. Recall, however, that we were concerned with determining the wave function and the eigenenergy of the ground state of a *many-particle* system defined by Hamiltonian (277), i.e. we were trying to find an (approximate) solution of the problem $\hat{H}\Phi = E\Phi$ for the ground state. We were searching that many-particle function in the form of a Slater determinant (274) [or (275)] and the approximate ground-state energy is given by the minimum of the functional (291); it is sufficient to substitute the determined optimal spinorbitals into it.

Let us try to investigate what happens to the determinantal wave function Φ if we apply some unitary transformation to the spinorbitals ϕ_i [7].

$$\begin{pmatrix} U_{11} & \dots & U_{1N} \\ \vdots & & \vdots \\ U_{N1} & \dots & U_{NN} \end{pmatrix} \begin{pmatrix} \phi_1 \\ \vdots \\ \phi_N \end{pmatrix} = \begin{pmatrix} \phi'_1 \\ \vdots \\ \phi'_N \end{pmatrix} \tag{355}$$

We get some new spinorbitals which we denoted by ϕ'_i. Briefly, we accomplished the transformation

$$\{\phi_i\} \longrightarrow \{\phi'_i\}\,, \qquad \text{kde} \quad \phi'_i = \sum_{j=1}^{N} U_{ij}\phi_j \tag{356}$$

What will be the value of the determinant Φ' composed of the transformed spinorbitals?

Najprv si zapíšme maticu, ktorú používame pri skladaní Slaterovho determinantu Φ podľa (275). Označme ju M:

$$M = \begin{pmatrix} \phi_1(1) & \dots & \phi_1(N) \\ \vdots & & \vdots \\ \phi_N(1) & \dots & \phi_N(N) \end{pmatrix} \tag{357}$$

Všimnime si, že maticu transformovaných spinorbitálov ϕ'_i teraz dostaneme maticovým násobením:

$$\begin{pmatrix} U_{11} & \dots & U_{1N} \\ \vdots & & \vdots \\ U_{N1} & \dots & U_{NN} \end{pmatrix} \begin{pmatrix} \phi_1(1) & \dots & \phi_1(N) \\ \vdots & & \vdots \\ \phi_N(1) & \dots & \phi_N(N) \end{pmatrix} = \begin{pmatrix} \phi'_1(1) & \dots & \phi'_1(N) \\ \vdots & & \vdots \\ \phi'_N(1) & \dots & \phi'_N(N) \end{pmatrix} \tag{358}$$

alebo stručne a s prehodeným poradím strán takto:

$$M' = UM \tag{359}$$

Z algebry je známe, že pre (štvorcové) matice platí

$$\det(AB) = \det(A)\det(B) \tag{360}$$

Preto

$$\underbrace{\det(M')}_{\Phi'} = \det(U)\underbrace{\det(M)}_{\Phi} \tag{361}$$

Pre unitárne matice z definície platí

$$U^\dagger U = I$$

Preto

$$1 = \det(U^\dagger)\det(U) = [\det(U)]^* \det(U) = |\det(U)|^2$$

čiže sa dá písať

$$\det(U) = e^{i\varrho} \tag{362}$$

kde ϱ je nejaké reálne číslo (ktorého hodnotu sme nezistili, ale ani nie je dôležitá).

According to (361), the transformed Slater determinant will be

$$\boxed{\Phi' = \Phi e^{i\varrho}} \tag{363}$$

that is, it differs from the original one by a phase factor only. However, it is known from QM that multiplying the wave function of a system by any constant (of magnitude 1 for correct normalisation) does not change anything on the physical properties of the system.

Naozaj: ak nejaká $\Psi(t)$ je riešením SchR

$$\mathrm{i}\hbar\frac{\partial\Psi(t)}{\partial t} = \hat{H}(t)\Psi(t) \tag{364}$$

tak vynásobením tejto rovnice ľubovoľnou konštantou napr. tvaru $e^{i\varrho}$ vidíme, že aj $\Psi'(t) = \Psi(t)e^{i\varrho}$ je riešením SchR pre *ten istý* hamiltonián. A úplne obdobne to platí aj pre stacionárnu SchR a náš jednodeterminantový stav Φ.

Therefore, if we find some spinorbitals $\{\phi_i\}_1^N$ using the HF method and then we calculate some new spinorbitals from them by an arbitrary unitary transformation, then these new ones will in principle be as good a solution of the HF equations as the original spinorbitals. True physical significance has the many-particle wave function, not the single-particle spinorbitals. These are just auxiliary "building blocks" to express the many-particle wave function. Since energy is a physical quantity, it must be conserved in unitary transformations of spinorbitals; is *invariant* under them. We would easily be convinced of this by explicit calculation by evaluating the functional $E_{\mathrm{HF}}[\phi]$ (291) for ϕ'. We would get $E_{\mathrm{HF}}[\phi'] = E_{\mathrm{HF}}[\phi]$.

It turns out the the ***Fock operator*** (321) is invariant ***under unitary transformations of the spinorbitals*** (it does not change if we replace ϕ_i with ϕ'_i).

Ideme to dokázať. Najprv sa pozrime na člen $\hat{h}(1)$ vo Fockovom operátore. Tento člen od spinorbitálov nezávisí, a teda je, celkom triviálne, invariantný voči ich zmenám. Ďalšou skupinou členov je súčet Coulombových operátorov. Počítajme, aký bude, keď ho vyjadríme pomocou transformovaných spinorbitálov (356).

$$\sum_{j=1}^N \hat{\mathcal{J}}'_j(1) = \sum_{j=1}^N \int \mathrm{d}x_2\, \phi'^*_j(2)\, r_{12}^{-1}\, \phi'_j(2) = \sum_{j=1}^N \int \mathrm{d}x_2 \sum_{k,l=1}^N U^*_{jk}U_{jl}\, \phi^*_k(2)\, r_{12}^{-1}\, \phi_l(2) =$$
$$= \int \mathrm{d}x_2 \sum_{k,l=1}^N \phi^*_k(2)\, r_{12}^{-1}\, \phi_l(2) \underbrace{\sum_{j=1}^N U^*_{jk}U_{jl}}_{\delta_{kl}} = \int \mathrm{d}x_2 \sum_{k=1}^N \phi^*_k(2)\, r_{12}^{-1}\, \phi_k(2) = \sum_{j=1}^N \hat{\mathcal{J}}_j(1)$$

> Súčet Coulombových operátorov je teda invariantný voči ľubovoľnej unitárnej transformácii spinorbitálov. Ešte overme invariantnosť súčtu výmenných operátorov. To sa spraví trochu náročnejšie, lebo výmenný operátor má zložitejšie štruktúrovanú formu; pozri poznámku 47 pod čiarou a rovnice (315), (316), (317), (318) a (319). Počítajme teda pôsobenie $\hat{\mathcal{K}}'_j(1)$ na ľubovoľnú funkciu ϕ:
>
> $$\sum_{j=1}^{N}\hat{\mathcal{K}}'_j(1)\phi(1) = \sum_{j=1}^{N}\int \mathrm{d}x_2\, \phi'^*_j(2)\, r_{12}^{-1}\, \phi(2)\phi'_j(1) = \sum_{j=1}^{N}\int \mathrm{d}x_2 \left[\sum_{m=1}^{N} U^*_{jm}\phi^*_m(2)\right] r_{12}^{-1}\, \phi(2) \left[\sum_{n=1}^{N} U_{jn}\phi_n(1)\right]$$
>
> Tak isto ako pri dokazovaní invariantnosti Coulombovho operátora, aj tu sa suma cez j dá vypočítať, teda
>
> $$\sum_{j=1}^{N} U^*_{jm}U_{jn} = \delta_{mn}$$
>
> Tak dostávame
>
> $$\sum_{j=1}^{N}\hat{\mathcal{K}}'_j(1)\phi(1) = \sum_{j=1}^{N}\hat{\mathcal{K}}_j(1)\phi(1)$$
>
> pre ľubovoľnú funkciu ϕ. Aby toto mohlo byť splnené, musí platiť aj
>
> $$\sum_{j=1}^{N}\hat{\mathcal{K}}'_j(1) = \sum_{j=1}^{N}\hat{\mathcal{K}}_j(1) \tag{365}$$
>
> čiže aj súčet výmenných operátorov je invariantný voči U. Tak zisťujeme, že Fockov operátor (321) je tiež invariantný, čo bolo treba dokázať.

Now take the HF equations in their compact form (322) and write them down in the matrix-vector notation:

$$\hat{f}(1)\begin{pmatrix}\phi_1(1)\\ \vdots \\ \phi_N(1)\end{pmatrix} = \begin{pmatrix}\lambda_{11} & \dots & \lambda_{1N}\\ \vdots & & \vdots \\ \lambda_{N1} & \dots & \lambda_{NN}\end{pmatrix}\begin{pmatrix}\phi_1(1)\\ \vdots \\ \phi_N(1)\end{pmatrix} \tag{366}$$

or compactly

$$\hat{f}(1)\vec{\phi}(1) = \lambda\,\vec{\phi}(1) \tag{367}$$

We multiply the last written equation from the left by the unitary matrix and insert the identity matrix in the form $U^\dagger U$ into the equation (by this, nothing is changed but the identity matrix is suitable to be inserted there):

$$\hat{f}(1)\,U\phi(1) = U\,\lambda\,U^\dagger U\,\phi(1) \tag{368}$$

On the left-hand side, it was correct to interchange the matrix U, contaning numerical quantities only, with the operator $\hat{f}$ [$\hat{f}(1)$ is just a single-component operator, not a matrix or vector]. The matrix λ is hermitian; it must be, otherwise the functional E_{HF} would not be real, but it is defined as real. We have learn above that a hermitian matrix can be diagonalised by a unitary transformation. Hence, also for our matrix λ there exists a unitary transformation that accomplishes its diagonalisation. So we get

$$\hat{f}(1)\,\phi'(1) = \mathrm{diag}(\mathcal{E})\;\phi'(1) \tag{369}$$

where $\mathrm{diag}(\mathcal{E})$ is a matrix that has the eigenvalues of the matrix λ on the diagonal and zeros everywhere else. If we write it down in components and without the primes for simplicity, we arrive at

$$\boxed{\boxed{\hat{f}(x)\phi_i(x) = \mathcal{E}_i\,\phi_i(x)}} \tag{370}$$

This is the ***canonical*** form of the HF equations. We obtained it using a unitary transformation of the originaly found spinorbitals. Thus, the spinorbitals are not the same as in (320), but for simplicity we use the same notation for the canonical orbitals. However, the operator $\hat{f}$ is the same because ot its invariance. As we have said above, the original spinorbitals yield the same energy E_{HF} as the new, transformed ones. Thus, it cannot be said that some are more physical than the other. But some (usually the diagonalising ones) may be more practical than others.

Especially in quantum chemistry, the HFE are a basis or starting calculation for several more accurate (but computationally more demanding) ***ab initio*** methods.[52]

---

[52] Clasically [ab iˈnitió], in later times [ab iˈnicio/iˈnício]; this is a term (from Latin) being used in scientific literature. In English, they alternatively use the term *first principles*; these are thus methods that calculate the electronic structure only from the basic laws of physics, i.e. from the SchE, without a use of empirical parameters or formulae. An *ab initio* method is primarily the Hartree-Fock one, which is, however, relatively inaccurate because it does not include the correlation energy.

9.6 Interpretation of the Solutions of the HF Equations

Also in this part you only need to look at what is typeset in the larger font.

The solutions of HF equations (370) are the spinorbitals ϕ_i and the eigenenergies $\mathcal{E}_i$. An important following result is also the approximate energy E_{HF} [see (325) of the ground state of the system under study. The many-body wave function Φ is rarely evaluated; it is not necessary, it is sufficient to work with the single-particle spinorbitals. Using them, as we have seen, the energy E_{HF} can be calculated as well as other physical quantities. As we have found out and stated beneath equation (364), the spinorbitals alone do not have an unambiguos physical significance. The canonical HF spinorbitals obtained by solving (370) are no more physical than any other related to them by a unitary transformation. Nevertheless, either the canonical spinorbitals (or only the corresponding orbitals) or some convenient unitarily transfomed ones exhibiting certain localisation in accordance with our intuition, can often be useful for better understanding chemical bonds or at least to give valuable intuitive insights. Such considerations, however, sometimes do not have a sufficient rational basis.

9.6.1 The Orbital Energies and the Total Energy

(We should talk about spinorbital energies, because the $\mathcal{E}_i$ that we found, correspond to spinorbitals. But for brevity, let us call them orbital energies.)

Z kanonických HF rovníc (370), v ktorých vystupuje Fockov operátor (321), dostávame

$$\mathcal{E}_i = \langle\phi_i|\hat{f}|\phi_i\rangle = \langle\phi_i|\hat{h}|\phi_i\rangle + \sum_{j=1}^{N}\Big[\langle\phi_i|\hat{\mathcal{J}}_j|\phi_i\rangle - \langle\phi_i|\hat{\mathcal{K}}_j|\phi_i\rangle\Big] \tag{371}$$

Orbitálnu energiu $\mathcal{E}_i$ môžme interpretovať ako energiu elektrónu obsadzujúceho stav ϕ_i nachádzajúceho sa v elektrostatickom vonkajšom poli plus v spriemerovanom poli ostatných $N-1$ elektrónov a mínus výmenná energia. Za Coulombov a výmenný operátor dosadíme podľa ich definícií (314) a (315). Zároveň si aj zavedieme štandardné kompaktné značenie pre jedno a dvojčasticové maticové elementy, ktoré veľmi zostručňuje zápis a používa sa najmä v HF teórii. Pomerne kompaktné značenie sme zaviedli už rovnicou (290), najmä v porovnaní s (289). Štandardne sa však v učebniciach kvantovej chémie používa ešte stručnejší zápis [7], a to buď „fyzikálny“ alebo „chemický“. My si uvedieme ten fyzikálny:

$$\langle i|h|j\rangle = \langle\phi_i|h|\phi_j\rangle = \int \mathrm{d}x\ \phi_i^*(x)\ \hat{h}\ \phi_j^*(x) \tag{372}$$

$$\langle ij|kl\rangle = \langle\phi_i\phi_j|\phi_k\phi_l\rangle = \int \mathrm{d}x_1\,\mathrm{d}x_2\ \phi_i^*(x_1)\phi_j^*(x_2)\ r_{12}^{-1}\ \phi_k(x_1)\phi_l(x_2) \tag{373}$$

$$\langle ij||kl\rangle = \langle ij|kl\rangle - \langle ij|lk\rangle = \int \mathrm{d}x_1\,\mathrm{d}x_2\ \phi_i^*(x_1)\phi_j^*(x_2)\ r_{12}^{-1}\ [\phi_k(x_1)\phi_l(x_2) - \phi_l(x_1)\phi_k(x_2)] \tag{374}$$

Orbitálne energie teraz môžeme zapísať takto:

$$\boxed{\mathcal{E}_i} = \langle i|h|i\rangle + \sum_{j=1}^{N}(\ \langle ij|ij\rangle - \langle ij|ji\rangle\ ) = \boxed{\langle i|h|i\rangle + \sum_{j=1}^{N}\langle ij||ij\rangle} \tag{375}$$

Pomocou stručného značenia zapíšeme aj HF energiu (325):

$$E_{\text{HF}} = \sum_{i=1}^{N}\langle i|h|i\rangle + \frac{1}{2}\sum_{i,j=1}^{N}\langle ij||ij\rangle \tag{376}$$

Let us compare the Hartree-Fock energy (325) of the ground state with the sum of all the orbital energies. We see that

$$\boxed{E_{\text{HF}} \neq \sum_{i=1}^{N}\mathcal{E}_i} \tag{377}$$

Thus, the ground-state energy is not the sum of the energies of the individual electrons.

$$\sum_{i=1}^{N}\mathcal{E}_i = \sum_{i=1}^{N}\langle i|h|i\rangle + \sum_{i,j=1}^{N}\langle ij||ij\rangle \tag{378}$$

This is because each interaction enters the above sum of the orbital energies twice; the factor of $1/2$ is missing there.

9.6.2 The Orbital Energies and the Koopmans Theorem

`Also in this section you only need to look at what is typeset in the larger font.`

We are going to examine the physical meaning of the orbital energies. However, we first realise that the Fock operator, (321), has an infinite number of eigenfunctions. By solving the HF equations, we get the sequences of the lowest N functions and the corresponding energies that we have studied so far: $\{\phi_i,\ \mathcal{E}_i\}_{i=1}^{N}$. However, by solving the HF equations, we also find the Fock operator, and for this we can then calculate the rest of the eigensystem: $\{\phi_i,\ \mathcal{E}_i\}_{i=N+1}^{\infty}$; see equations (370). According to the literature [7], we will agree on the following notation for the indices: The indices $a, b, \ldots$ will be used for $i \in \{1, 2, \ldots, N\}$, thus for the lowest N single-electron states. These states describe electrons forming the Slater many-particle function; therefore, we call them *occupied* states or spinorbitals. The indices $r, s, \ldots$ will be used for $i > N$, that is for the unoccupied spinorbitals,o which are usually called empty or (especially in quantum-chemistry literature) *virtual*. Really: there are only N electrons in the system, and therefore, if the system is in the ground state (Φ), then only N of the lowest spinorbitals can be occupied, each by one electron. Possible excited states of the system can then be at least roughly described by creating a new determinant in which one of the occupied spinorbitals ϕ_a is replaced by one of the empty spinorbitals ϕ_r.

It should be noted that statements of the type "electron occupies the level i" (and has energy $\mathcal{E}_i$ and is described by the wave function ϕ_i) are inaccurate. In fact, the system is described by some many-particle function. However, such statemens or expressions correspond to the one-particle picture provided by the HF method, are also intuitively understandable, and are therefore it is convenient to use them.

S využitím (375) a identity [podľa (374)]

$$\langle ii||ii\rangle = 0 \tag{379}$$

hneď vieme napísať

$$\mathcal{E}_a = \langle a|h|a\rangle + \sum_{\substack{b=1\\ \neq a}}^{N} \langle ab||ba\rangle \tag{380}$$

$$\mathcal{E}_r = \langle r|h|r\rangle + \sum_{b=1}^{N} \langle rb||rb\rangle \tag{381}$$

V rovnici pre obsadený stav sme teda mohli zo sumovania vypustiť člen, ktorý by bol aj tak nulový.

Uvažujme teraz tri fyzikálne sústavy:

1. Tú N-elektrónovú, pre ktorú sme počítali základný stav HF metódou. Označme príslušný determinant ket vektorom aj s indexom N a dolným indexom 0 (prízvukujúcim, že ide o základný stav). Zapíšme pritom aj príslušne označenú celkovú energiu sústavy (376).

$$\left|{}^{N}\Phi_0\right\rangle , \qquad {}^{N}E_0 \;=\; \left\langle {}^{N}\Phi_0 \left|\hat{H}\right| {}^{N}\Phi_0\right\rangle \;=\; \sum_{a=1}^{N} \langle a|h|a\rangle + \frac{1}{2} \sum_{a,b=1}^{N} \langle ab||ab\rangle \tag{382}$$

2. $(N-1)$-elektrónovú sústavu takú, ktorú získame vytrhnutím elektrónu z (obsadenej) hladiny c základného stavu. Ostatné elektróny pritom necháme v takých spinorbitáloch, a akých sú. Príslušný determinant a energiu zapíšme obdobne ako v predošlom bode:

$$\left|{}^{N-1}\Phi_c\right\rangle , \qquad {}^{N-1}E_c \;=\; \left\langle {}^{N-1}\Phi_c \left|\hat{H}\right| {}^{N-1}\Phi_c\right\rangle \;=\; \sum_{\substack{a=1\\ \neq c}}^{N} \langle a|h|a\rangle + \frac{1}{2} \sum_{\substack{a=1\\ \neq c}}^{N} \sum_{\substack{b=1\\ \neq c}}^{N} \langle ab||ab\rangle \tag{383}$$

   Je zrejmé, že vybraním jedného elektrónu zmeníme elektrické pole vo vnútri sústavy a stav $\left|{}^{N-1}\Phi_c\right\rangle$ nebude základným stavom $N-1$ elektrónovej sústavy ani v priblížení HF metódy. HF základný stav by sme dostali tak, že by sme spinorbitály existujúcich $N-1$ elektrónov prispôsobili novému poľu, teda by sme vlastne museli spraviť nový HF výpočet. Stav $\left|{}^{N-1}\Phi_c\right\rangle$ však často môžeme pokladať aspoň za približný $(N-1)$-elektrónový základný stav.

3. $(N+1)$-elektrónovú sústavu takú, ktorú získame dodaním elektrónu do (prázdnej) hladiny r základného stavu. Ostatné elektróny pritom opäť necháme v takých spinorbitáloch, a akých sú. Príslušný determinant a energiu označíme podobne ako v predošlých bodoch, ale energiu vypočítame len univerzálnou formulou platnou pre ľubovoľný jednodeterminantový stav:

$$\left|{}^{N+1}\Phi^r\right\rangle , \qquad {}^{N+1}E^r \;=\; \left\langle {}^{N+1}\Phi^r \left|\hat{H}\right| {}^{N+1}\Phi^r\right\rangle \;=\; \sum_{i}^{\text{occ}} \langle i|h|i\rangle + \frac{1}{2} \sum_{i,j}^{\text{occ}} \langle ij||ij\rangle \tag{384}$$

Rozsah occ znamená, že sa sumuje cez všetky obsadené spinorbitály, nech sú to ktorékoľvek. Konkretizácia pre $N+1$ elektrónov v hladinách $1, 2, \ldots, N, (\ldots), r$ je jednoduchá, ale vyžaduje viac písania.

The ***ionisation energy*** (sometime not very properly called ionisation potential) of an atom, molecule or ion is the amount of energy needed to release an electron from the electron shell of given system (the atom, ...). Thus, within the one-determinant approximation (which is used in the HF method), we could consider as the approximate ionisation energy the energy needed to release an electron from the highest occupied (the N-th) level. However, for the purpose of interpreting the orbital energies, consider the release of an electron from any of the first N levels. The ionisation energy defined in such a more general sense is:

$$I = {}^{N-1}E_c - {}^{N}E_0 \tag{385}$$

Jeho výpočet na základe vyjadrení (382), (383) a (379) nie je ťažký ani zdĺhavý. Využíva aj identity

$$\langle ij|kl\rangle = \langle ji|lk\rangle\,, \qquad \langle ij||kl\rangle = \langle ji||lk\rangle \tag{386}$$

vyplývajúce z definícií (373) a (374). Výsledkom je

$$\boxed{I = -\mathcal{E}_c} \tag{387}$$

The ***electron affinity*** is defined in an analogous way: when we add another electron to the molecule, the energy of the molecule usually decreases. The more it decreases, the greater the electron affinity, which is then a positive value.

Thus, its definition in the framework of the HF theory is

$$A = {}^{N}E_0 - {}^{N+1}E^r \tag{388}$$

The calculation is again not difficult and its result is

$$\boxed{A = -\mathcal{E}_r} \tag{389}$$

Expression (387) and (389) for the ionisation energy and the electron affinity form the content of the Koopmans theorem. For reasons that we will not be explaining here, the ionisation energies according to (387) are a relatively good approximation to experimental values, while the electron affinities according to (389) are usually quite wrong. They are even often negative, although they should be positive [7].

An important quantity is also the *electronegativity* EN of an atom or a molecule. Its definition according to Mulliken is

$$\mathrm{EN} = \frac{I + A}{2} \tag{390}$$

9.7 The Spin-Restricted and Unrestricted Variants of the HF Method

In solving the HFE, the unknown spinorbitals are expanded in a chosen basis of known functions and the coefficients of this expansion are then the unknowns to be determined. As we have seen, the HFE can be transfomed to form (263) from motivation section 9.1. [It is even a simpler form than (E.29) at the HE, since the effective Hamiltonian in (263) is the same for all spinorbitals – there is no i index at it.] This leads to the idea or concept of the energy levels $\mathcal{E}_i$ and also to the concept that an electron with wave function ϕ_i occupies the energy level $\mathcal{E}_i$. It is often the case that these levels are doubly degenerate, that is

$$\mathcal{E}_1 = \mathcal{E}_2\,, \qquad \mathcal{E}_3 = \mathcal{E}_4\,, \qquad \ldots\ldots\ldots\,, \qquad \mathcal{E}_{N-1} = \mathcal{E}_N$$

and the spinorbitals ϕ_1, ϕ_2 then differ in their spin component only; similarly the spinorbitals ϕ_3, ϕ_4 have identical spatial parts and differ in their spin parts only. And so on for all other pairs of the spinorbitals (if we properly label them). In other words, each energy level is occupied by two electrons which differ in their spins only. One of them has its spin "up", the other "down". We then say that the system (an atom, molecule or crystal) exhibits ***spin degeneracy***. In such case, only $N/2$ unknow spatial orbitals are to be determined. If for this or any other reason, we force the spatial components of the pairs of the spinorbitals to be identical, we say about the ***spin-restricted Hartree-Fock*** theory (or just restricted HF, in brief). It can well be used especially for systems with an even number of electrons (but not for all such). The restricted HFE expressed in a particular non-orthogonal basis (e.g. in the basis of the gaussian functions) are called the ***Roothan equations***. If we do not require the identical pairs of the spatial orbitals, we say about the ***unrestricted*** HF method or theory.

10 The Homogeneous Electron Gas

For chemical reactions and electrical conductivity, especially the valence electrons of atoms or molecules, including atoms in extended systems,[53] are important. Among such materials, conductive crystals, i.e. common metals (which do not usually consist of a one monocrystal, but often considering a monocrystal of given material is sufficient to study many of the material's properties), have an important place. The simplest metals are the lighter elements of Group 1 of the Mendeleev Periodic Table: alkali metals Li, Na, K. Their atoms have only one valence (outer) electron. Thus, their highest electron shell is similar to the shell of the hydrogen atom, which is $1s^1$. E.g. lithium has shells $1s^2$, $2s^1$, so its outer shell is only half full. The inner one is completely filled (i.e. closed) and therefore relatively stable, inert. Alkali metals are also called *simple*.

If there are many atoms in the system, the valence electrons are only weakly bound in them, they easily jump from atom to atom and thus move across the crystal lattice. It cannot be said that any valence electron would be bound to a particular atom. In the case of a simple metal, all the valence electrons are thus weakly bound and their energies are from a certain continuous energy interval (the conduction band). The conduction electrons in the ground state (ideally at a temperature of $T \approx 0\,\mathrm{K}$) occupy only the lower half of the energy levels of the conduction band.

A metal as a whole is electrically neutral under normal conditions. Due to the mobility of the conduction electrons, their spatial distribution is such that electrical neutrality is also ensured locally. A significant deviation from local neutrality under normal conditions in a metal is not even possible because there is not enough energy for it. Thus, there is a compensating positive charge in a close neighbourhood of each electron. Therefore, the Coulomb interaction between electrons is significantly screened. *In the roughest approximation, we can even look at electrons as non-interacting.* Thus, to understand some of the basic properties of a metal, it is sufficient to consider conduction electrons as a gas of non-interacting (independent) electrons.

10.1 Non-Interacting Electrons

Consider a homogeneous gas of non-interacting $1/2$-spin fermions occupying a macroscopic (sufficiently large) volume. We want to calcuate their total energy. By its nature, it will certainly be a kinetic energy, because non-interacting particles cannot have a potential energy. Their mutual independence means that at the beginning, it will be sufficient to examine properties of one such electron. Obviously, if the space where those electrons are located is large enough, it doesn't matter if the space has a shape of a cuboid, a sphere, a cube, and so on. Across a sufficiently large part of the considered space, the wave function of such an electron – a free particle – can be described as a plane wave:

$$\psi_{\vec{k}}(\vec{r}) = A\,e^{\mathrm{i}\vec{k}.\vec{r}} = A\,e^{\mathrm{i}(k_x x + k_y y + k_z z)} \tag{391}$$

and the corresponding eigenenergy is

$$\mathcal{E}_{\vec{k}} = \frac{\hbar^2 k^2}{2m} \tag{392}$$

A is an unspecified normalisation constant. Since we now assume truly independent particles, and not those that would interact at least through a mean field in the HF method, the total energy of the gas will be the sum of the energies of the individual particles [*cf.* (377)]. And the many-particle wave function will be a Slater determinant build of one-particle functions. However, there is still one condition that otherwise independent electrons must respect: the Pauli principle. The principle implies for a given system that the wave function $\psi_{\vec{k}}(\vec{r})$ can describe no more than two electrons; if they are two, they must differ in the values of their spins. In consequence, the energy level $\mathcal{E}_{\vec{k}}$ can be occupied by no more than two electrins differing in their spins. Since electrons in matter tend to occupy the lowest possible energy levels at low temperatures, the individual levels will be occupied starting from the lowest one up to a certain higher level. Note that if there were no Pauli principle, all (non-interacting) electrons would occupy the lowest possible energy level. Thus, although in the Hamiltonian

$$\hat{H} = \sum_{i=1}^{N} -\frac{\hbar^2}{2m}\vec{\nabla}_i^2 \tag{393}$$

---

[53] They can be e.g. crystals, which are condensed substances with periodic repetition of a certain group of atoms. But they can also be non-periodic solids, which can often be described as amorphous.

there is no interaction, the Pauli principle still causes that even the independent fermions at least "know" that their neighbours already occupied some level.

There are no restrictions of the values of the wave vector $\vec{k}$. This may lead to technical difficulties, for example how to calculate summations over all occupied states. Also, how to normalise a plane wave in the infinite space. This is usually solved by introducing *periodic boundary conditions* (PBC).[54] Instead of the infinite space, we imagine only a certain section it having the shape of (in the simplest case) a cube with the side L and we impose the PBC on the wave functions:

$$\begin{aligned} \psi_{\vec{k}}(x+L,\ y,\ z) &= \psi_{\vec{k}}(x,y,z) \\ \psi_{\vec{k}}(x,\ y+L,\ z) &= \psi_{\vec{k}}(x,y,z) \qquad\qquad \forall\ x,y,z \in \mathbb{R} \\ \psi_{\vec{k}}(x,\ y,\ z+L) &= \psi_{\vec{k}}(x,y,z) \end{aligned} \tag{394}$$

In order for the first of the equations (394) to be satisfied, the following must hold:

$$\exp\left\{\mathrm{i}[k_x(x+L)+k_y y+k_z z]\right\} = \exp\left\{\mathrm{i}[k_x x+k_y y+k_z z]\right\} \quad\Longrightarrow\quad \exp\left(\mathrm{i}k_x L\right)=1 \quad\Longrightarrow\quad k_x = \nu_x \frac{2\pi}{L}$$

ν_x can be any integer. It apply analogously to the two other boundary conditions. Thus we get restrictions on the real wave vectors which were otherwise arbitrary up to now. The values allowed by the defined boundary conditions are

$$\vec{k}_{\vec{\nu}} = \frac{2\pi}{L}\left(\nu_x,\nu_y,\nu_z\right)\ , \qquad \nu_x,\nu_y,\nu_z \in \mathbb{Z} \tag{395}$$

The form of the wave function, of course, remains as we found it at the beginning: (391). This time, however, we can also find the normalisation constant A, because we will require that

$$\int_{(\Omega)} \left|\psi_{\vec{k}}(\vec{r})\right|^2 \ \mathrm{d}^3 r = 1 \tag{396}$$

where it is indicated that the integration goes over the volume $\Omega = L^3$. Hencem, the normalised one-electron wave functions are

$$\boxed{\psi_{\vec{\nu}}(\vec{r}) = \frac{1}{\sqrt{\Omega}}\ e^{\mathrm{i}\vec{k}_{\vec{\nu}}.\vec{r}}} \tag{397}$$

The corresponding eigenenergies are calculated using formula (392). We can consider the numbers ν_x, ν_y, ν_z as quantum numbers.

Of course, the introduction of the PBC is an artifical procedure. We can therefore ask whether the results we get under its assumption will be correct. We first realize that we have chosen a very large, periodically replicated volume Ω. The artificial forcing that after a huge distance L the wave function must begin to repeat its values has a negligible effect on local physical quantities somewhere within the considered volume. In addition, we will see that this hand-inserted artificial length and volume scale falls out of the results due to a fraction simplification. It is therefore practical in our considerations, but its specific value is not important. Finally, when we think better about PBC later, we will realise that by PBC we have only chosen a certain (regular) sampling of k-space.

Now let us calculate the total energy of the ground state of the gas in the volume Ω; gas is also everywhere else, but we calculate the energy per volume. According to what we said under formula (392), the following applies:

$$E = \sum_{i=1}^{N} \mathcal{E}_{\vec{k}_i} \tag{398}$$

In this, we assume that N is the number of electrons per volume Ω. The electrons occupy the states $\psi_{\vec{\nu}}$. We can imagine that we store the N given electrons in quantum states (labeled by) $\vec{k}_{\vec{\nu}}$ gradually, from the lowest state ($\nu_x = \nu_y = \nu_z = 0$) up to the higher ones. We place two electrons in each state, differing only in their spins. Since we build the basic many-particle state of the system in this way, in occupying the states we cannot omit any of the one-particle states, because the resulting N-electron state would not have the lowest possible energy. So, the total energy E can be expressed by summation over wave vectors (395),

$$E = \sum_{\vec{k}} g_{\vec{k}}\ \mathcal{E}_{\vec{k}} = \sum_{\nu_x=-\infty}^{\infty}\ \sum_{\nu_y=-\infty}^{\infty}\ \sum_{\nu_z=-\infty}^{\infty} g_{\vec{k}_{\vec{\nu}}}\ \mathcal{E}_{\vec{k}_{\vec{\nu}}} \tag{399}$$

---

[54] – a very common acronym in English texts on electronic structure.

where $g_{\vec{k}}$ are the *occupation numbers*: they are equal to 2 for the occupied states and 0 for the unoccupied ones. So, we can formally sum over the all wave vectors. The mesh of the allowed k-vectors (395) is quite dense for large L; neighboring vectors differ by $2\pi/L$ only. Therefore, it will be correct if we replace the summation in (399) by integration:

$$E = \sum_{\vec{k}} g_{\vec{k}}\, \mathcal{E}_{\vec{k}} \longrightarrow \int \mathrm{d}^3k\, \rho_{\vec{k}}\, g_{\vec{k}}\, \mathcal{E}_{\vec{k}} \tag{400}$$

where

$$\rho_{\vec{k}} = \frac{1}{\left(\dfrac{2\pi}{L}\right)^3} = \frac{\Omega}{8\pi^3} \tag{401}$$

is the *density of states* in the 3-dimensional k-space and as it can be seen, it is constant, that is, independent of the position $\vec{k}$ in the reciprocal space. Thus, by its physical dimension and also meaning, $\rho_{\vec{k}}$ is the number of k-vectors (i.e. also the number opf states $\psi_{\vec{k}}$) per unit of volume of the reciprocal space:[55]

$$\rho_{\vec{k}} = \frac{\mathrm{d}\mathcal{N}}{\mathrm{d}^3k} \tag{402}$$

The total energy E can then be calculated by integrating according to (400), while the occupation numbers $g_{\vec{k}}$ are nonzero only for wave vectors with sizes up to a certain maximum value, which we denote by k_F. As can be seen with consideration of (392), the integrated function will depend only on the size of the wave vector, not on its direction. Therefore, it will be advantageous to integrate using spherical coordinates in reciprocal space:

$$\begin{aligned} E &= \frac{\Omega}{8\pi^3} \int \mathrm{d}^3k\, g_{\vec{k}}\, \mathcal{E}_{\vec{k}} = \frac{\Omega}{8\pi^3} \int_0^{k_\mathrm{F}} \mathrm{d}k\, k^2 \int_0^{\pi} \mathrm{d}\vartheta \sin\vartheta \int_0^{2\pi} \mathrm{d}\varphi\, 2\, \frac{\hbar^2 k^2}{2m} = \\ &= \frac{\Omega}{8\pi^3}\, 8\pi\, \frac{\hbar^2}{2m} \int_0^{k_\mathrm{F}} \mathrm{d}k\, k^4 = \frac{\Omega}{10\pi^2} \frac{\hbar^2}{m} k_\mathrm{F}^5 \end{aligned} \tag{403}$$

teda

$$E = \frac{1}{10\pi^2} \frac{\hbar^2}{m} k_\mathrm{F}^5\, \Omega \tag{404}$$

We realise that we have integrated over the volume of a sphere with a centre at the origin and a radius k_F. The volume of this sphere in reciprocal space is:

$$\Omega_\mathrm{F} = \frac{4}{3}\pi k_\mathrm{F}^3$$

and it is the volume of the reciprocal space that is fully occupied by the given N electrons in the ground state. In this consideration, the already mentioned volume of $(2\pi/L)^3$ per two electrons (differing in their spins) is accounted for. Therefore, the total number of electrons can be expressed as

$$N = 2\, \frac{\frac{4}{3}\pi k_\mathrm{F}^3}{\left(\dfrac{2\pi}{L}\right)^3} = \frac{1}{3\pi^2} k_\mathrm{F}^3\, \Omega$$

Let us introduce the notation

$$n = \frac{N}{\Omega} \tag{405}$$

It is density of electrons in the usual (they also say direct) space, that is, the basic parameter of the system under study. Because we are studying a homogeneous electron gas, it is a constant. Thus, using the density, we can express the above-introduced parameter

$$\boxed{k_\mathrm{F} = \left(3\pi^2 n\right)^{1/3}} \tag{406}$$

which is called the *Fermi wave number* and it is therefore the size of those wave vectors that are related to the highest occupied states. It is noteworthy that the expression (406) remains valid for interacting electrons with density n

---

[55] Where necessary, differently defined densities of states are being used in literature.

too [12] (but we will not prove it). The energy of the energetically highest electrons of the gas at the temperature of 0 K is called the ***Fermi energy***:

$$\mathcal{E}_{\mathrm{F}} = \frac{\hbar^2 k_{\mathrm{F}}^2}{2m} \tag{407}$$

The average total energy E per one electron of the system can now be expressed using formulae (404), (405), (406) and (407) as follows:

$$\boxed{\frac{E}{N} = \frac{3}{5}\,\mathcal{E}_{\mathrm{F}}} \tag{408}$$

After a simple derivation, the result for the total energy expressed using the electron density becomes

$$\boxed{E = \frac{3}{10}\,\left(3\pi^2\right)^{2/3}\,\frac{\hbar^2}{m}\;n^{5/3}\;\Omega} \tag{409}$$

It is ***kinetic*** energy, because we obtained it for the gas of non-interacting electrons, which cannot have potential energy. It is the sum of the kinetic energies of the individual electrons.

11 The Thomas-Fermi Model

Thus, formula (409) gives the energy of the non-interacting electron gas in the volume Ω. The model of Thomas and Fermi (TF) extrapolates the validity of this relation to relatively general situations with inhomogeneous electron densities in systems, such as atoms and possibly molecules, ions and extended systems. According to this model, we imagine the volume of the system divided into small volumes $\Delta\Omega$. In each of these volume elements, we assume the validity of formula (409), taking the average density in this element for the density and the value $\Delta\Omega$ for its volume. It is certainly no longer a non-interacting gas, but Thomas and Fermi nevertheless make the bold assumption described. The total kinetic energy is thus obtained by summation over all volume elements $\Delta\Omega$, i.e. by integrating over the entire volume of the system. In the atomic units ($\hbar = m = 1$) this energy gets

$$T_{\mathrm{TF}}[n] = C_{\mathrm{F}} \int n^{5/3}(\vec{r})\;\mathrm{d}^3 r\,, \qquad C_{\mathrm{F}} = \frac{3}{10}\,\left(3\pi^2\right)^{2/3} \approx 2.871 \tag{410}$$

T_{TF} is called the *Thomas-Fermi functional* for kinetic energy. Thus, kinetic energy of the system of electrons is calculated ***from the density only*** in the model. In order to be able to calculate the energy of the ground state of the system, it is necessary to add other components to the energy. In the basic TF model [11, 12], the following is added:

- The elektrostatic interaction of the charge density $-en(\vec{r})$ with the external potential $u_{\mathrm{ext}}(\vec{r})$, which in the TF theory usually is the potential coming from the atomic nucleus; however, if we study a more complex system, it is the total electrostatic potential from the nuclei plus from other external sources:

  $$u_{\mathrm{ext}}(\vec{r}) = \sum_I \frac{Z_I}{|\vec{r} - \vec{R}_I|} \;+\; u_{\mathrm{other}}(\vec{r}) \tag{411}$$

  It is an external potential defined exactly as we used it in the HF theory [*cf.* (259)], but there we multiplied it by the charge of electron, thus getting the potential energy operator $v_{\mathrm{ext}}(\vec{r})$ from the potential. Thus, the following holds in the SI system:

  $$v_{\mathrm{ext}}(\vec{r}) \equiv -e u_{\mathrm{ext}}(\vec{r}) \tag{412}$$

  In this section, however, we are otherwise using the Hartree atomic units in which $v_{\mathrm{ext}}(\vec{r}) \equiv -u_{\mathrm{ext}}(\vec{r})$. In addition, in TF theory we do not work with operators, but only with classical concepts, so we do not even need the operator symbol $\hat{v}_{\mathrm{ext}}$ here.

- The classical Hartree energy describing the electrostatic interaction of electrons with each other in terms of their total density; there is such a contribution even in the HF theory, although it is primarily expressed there through the orbitals; In the HF theory, the Hartree energy is in the Coulomb electrostatic integral.

In this way we obtain the energy of the system as a functional of the density:

$$\boxed{E_{\mathrm{TF}}[n] = C_{\mathrm{F}} \int n^{5/3}(\vec{r})\,\mathrm{d}^3r - \int u_{\mathrm{ext}}(\vec{r})n(\vec{r})\,\mathrm{d}^3r + \frac{1}{2}\int \frac{n(\vec{r})n(\vec{r}\,')}{|\vec{r}-\vec{r}\,'|}\,\mathrm{d}^3r\,\mathrm{d}^3r'} \equiv T_{\mathrm{TF}} + E_{\mathrm{ext}} + E_{\mathrm{Hartree}} \tag{413}$$

This is the functional of the Thomas-Fermi theory of atoms (and it could be evaluated for molecules and crystals too). Its first term (kinetic energy) is non-classical, the second and third ones are the classical electrostatic interactions of a continuous charge distribution with an external field and with itself, as we know them from the basic physics course.

If the TF model is correct, then by finding the minimum of the functional E_{TF} we determine the energy of the ground state of the considered atom or other system. Thus, the task will be to find the minimising density $n(\vec{r})$. That one must, however, integrate to the required number of electrons of the atom under study:

$$\int n(\vec{r})\,\mathrm{d}^3r = N \tag{414}$$

We impose this using a Lagrange multiplier as an equality constraint. Therefore, the augmented functional

$$\boxed{\Omega_{\mathrm{TF}}[n] = E_{\mathrm{TF}}[n] - \mu\left[\int n(\vec{r})\,\mathrm{d}^3r - N\right]} \tag{415}$$

is to be minimised. μ is the Lagrange multiplier. To find the minimum of this functional, we must first calculate its variation and then set it equal to zero. We find the variation of the functional similarly as when we talked about the Hartree-Fock equations: we will calculate how the functional changes when the density $n(\vec{r})$ changes by a small amount $\delta n(\vec{r})$:

$$\delta\Omega_{\mathrm{TF}}[n] = \Omega_{\mathrm{TF}}[n+\delta n] - \Omega_{\mathrm{TF}}[n] \tag{416}$$

In order not to have to write too long formulae, we will write the whole functional as a sum of several terms:

$$\Omega_{\mathrm{TF}} = T_{\mathrm{TF}} + E_{\mathrm{ext}} + E_{\mathrm{Hartree}} - \mu\left[\int n(\vec{r})\,\mathrm{d}^3r - N\right] \tag{417}$$

First, calculate

$$\delta T_{\mathrm{TF}}[n] = T_{\mathrm{TF}}[n+\delta n] - T_{\mathrm{TF}}[n]$$

$$T_{\mathrm{TF}}[n+\delta n] = C_{\mathrm{F}}\int (n+\delta n)^{5/3}\,\mathrm{d}^3r = C_{\mathrm{F}}\int n^{5/3}\left(1+\frac{\delta n}{n}\right)^{5/3}\mathrm{d}^3r$$

The density variations in the search for the functional minimum are very small, such that $|\delta n|/n \ll 1$. Then we can apply Taylor's expansion

$$(1+x)^p \approx 1+px\,, \qquad |x| \ll 1 \tag{418}$$

to a part of the subintegral expression and obtain

$$T_{\mathrm{TF}}[n+\delta n] = C_{\mathrm{F}}\int \left(n^{5/3} + \frac{5}{3}n^{2/3}\delta n\right)\mathrm{d}^3r$$

Therefore

$$\boxed{\delta T_{\mathrm{TF}}[n] = \frac{5}{3}C_{\mathrm{F}}\int n^{2/3}\,\delta n\;\mathrm{d}^3r} \tag{419}$$

Let us now calculate

$$\begin{aligned}
E_{\mathrm{Hartree}}[n+\delta n] &= \frac{1}{2}\int \frac{\left[n(\vec{r})+\delta n(\vec{r})\right]\left[n(\vec{r}\,')+\delta n(\vec{r}\,')\right]}{|\vec{r}-\vec{r}\,'|}\,\mathrm{d}^3r\,\mathrm{d}^3r' = \\
&= E_{\mathrm{Hartree}}[n] + \frac{1}{2}\int \frac{n(\vec{r})\delta n(\vec{r}\,')}{|\vec{r}-\vec{r}\,'|}\,\mathrm{d}^3r\,\mathrm{d}^3r' + \frac{1}{2}\int \frac{\delta n(\vec{r})n(\vec{r}\,')}{|\vec{r}-\vec{r}\,'|}\,\mathrm{d}^3r\,\mathrm{d}^3r' + \\
&+ \frac{1}{2}\int \frac{\delta n(\vec{r})\delta n(\vec{r}\,')}{|\vec{r}-\vec{r}\,'|}\,\mathrm{d}^3r\,\mathrm{d}^3r'
\end{aligned}$$

The last term – the one with the product of variations – is infinitely smaller than the terms linear in the variation δn. The two linear terms of the last formula are identical because the notations of the integration variables do not matter. So we get

$$\boxed{\delta E_{\text{Hartree}}[n] = \int \frac{n(\vec{r})\delta n(\vec{r}\,')}{|\vec{r}-\vec{r}\,'|}\ \mathrm{d}^3r\,\mathrm{d}^3r'} \tag{420}$$

Variations of the remaining two contributions to the functional Ω_{TF} can be calculated analogously (and very simply). Bu summing up we arrive at expression

$$\delta\Omega_{\text{TF}}[n] = \frac{5}{3}C_{\text{F}}\int n^{2/3}\,\delta n\ \mathrm{d}^3r - \int u_{\text{ext}}(\vec{r})\delta n(\vec{r})\ \mathrm{d}^3r + \int \frac{n(\vec{r}\,')\delta n(\vec{r})}{|\vec{r}-\vec{r}\,'|}\ \mathrm{d}^3r\,\mathrm{d}^3r' - \mu\int \delta n(\vec{r})\ \mathrm{d}^3r \tag{421}$$

and if we put it into one integral and take out of the brackets what can be taken, we obtain the following expression:

$$\delta\Omega_{\text{TF}}[n] = \int \mathrm{d}^3r\left[\frac{5}{3}\,C_{\text{F}}\,n^{2/3}(\vec{r}) - u_{\text{ext}}(\vec{r}) + \int \mathrm{d}^3r'\,\frac{n(\vec{r}\,')}{|\vec{r}-\vec{r}\,'|} - \mu\right]\delta n(\vec{r}) + \\ + \text{terms proportional to higher powers in } \delta n \tag{422}$$

It follows from the condition of the minimum of the functional that we must require

$$\delta\Omega_{\text{TF}} = 0 \tag{423}$$

for any small variations of the electron density. For $\delta\Omega_{\text{TF}}$ to be zero under these conditions, the following must apply:

$$\boxed{\boxed{\frac{5}{3}\,C_{\text{F}}\,n^{2/3}(\vec{r}) - u_{\text{ext}}(\vec{r}) + \int \frac{n(\vec{r}\,')}{|\vec{r}-\vec{r}\,'|}\ \mathrm{d}^3r' = \mu}} \tag{424}$$

By solving this integral equation for an unknown density (such that it is non-negative and integrates to N), we can finally evaluate also the ground state energy of the TF model. The Lagrange multiplier μ has a meaning of the Fermi energy according to the TF model (which we will not justify).

Thus, the TF theory (1927) provides an incredibly simple description of the complex many-particle problem. Instead of a wave function which has $3N$ spatial variables, the density with only 3 variables is sufficient. The tax for this simplicity is the inaccuracy of the TF model, even after improvements, which have been many in history (eg the inclusion of exchange energy). For example, this method is unable even qualitatively correctly describe molecular bonds (nuclei of a molecule would not be kept together according to the TF model). However, the TF theory is an excellent motivation for a rigorous approach to the description of the ground state using the density. This rigorous approach – density functional theory – began to develop in 1964, when Hohenberg and Kohn published their innovative work.

12 Density Functional Theory

We will be using abbreviation DFT according to the English term *Density-Functional Theory*. The fundamentals of this theory were laid by Pierre Hohenberg a Walter Kohn in their work [13]. We will introduce the two basic theorems of DFT according to this original work.

12.1 Electron Density

Since DFT theorems work with the concept of electron density, it will be useful to first clarify how this density can be determined from a wave function. Let a wave function of an N-electron system is $\Psi(\vec{r}_1,\sigma_1,\ldots,\vec{r}_N,\sigma_N)$. At this stage of our the exposition, it may not even be a wave function of the ground state, it may even depend on time, although we do not explicitly indicate it. However, we do require the standard normalisation:

$$\sum_{\sigma_1}\cdots\sum_{\sigma_N}\int \mathrm{d}^3r_1\cdots\int \mathrm{d}^3r_N\,|\Psi(\vec{r}_1,\sigma_1,\ldots,\vec{r}_N,\sigma_N)|^2 = 1 \tag{425}$$

We interpret the expression $|\Psi(\vec{r}_1, \sigma_1, \dots, \vec{r}_N, \sigma_N)|^2 \ \mathrm{d}^3 r_1 \dots \mathrm{d}^3 r_N$ as the probability to find:

electron 1 in a state with spin σ_1 within the volume $\mathrm{d}^3 r_1$ in the neighbourhood of point $\vec{r}_1$
and at the same time (in the sense of the logic)
electron 2 in a state with spin σ_2 within the volume $\mathrm{d}^3 r_2$ in the neighbourhood of point $\vec{r}_2$
and at the same time
....
electron N a state with spin σ_N within the volume $\mathrm{d}^3 r_N$.

In these considerations, we must not forget that the electrons are in fact indistinguishable and the wave function is antisymmetric. However, the wave function is conventionally being normalised according to (425). According to this normalisation, the unity is obtained by summing up the probabilities of the occurrence of the N electrons here and there in the way as if they were distinguishable. Other normalisation would bring substantial complications.

Calculation of $n(\vec{r})$ Using the Probability Density If we omit summation and integration over one of the coordinates, e.g. over the first one, we obtain the probability density of finding *any* electron in the spin state σ at the point $\vec{r}$:

$$\mathcal{P}_\sigma(\vec{r}) = \sum_{\sigma_2} \cdots \sum_{\sigma_N} \int \mathrm{d}^3 r_2 \cdots \int \mathrm{d}^3 r_N \, |\Psi(\vec{r}, \sigma, \vec{r}_2, \sigma_2, \dots, \vec{r}_N, \sigma_N)|^2 \tag{426}$$

We highlighted *any*, because due to the indistinguishability of electrons and the antisymmetry of the wave function $[\Psi(j, i; t) = -\Psi(i, j; t)$, see the end of section 8.5], it does not matter which orbital-spin coordinate $(\vec{r}_i, \sigma_i)$ we would omit from the summation and integration; from a given wave function Ψ, we would always get the same function $\mathcal{P}_\sigma(\vec{r})$. Of course, thanks to normalisation (425) it holds

$$\sum_\sigma \int \mathrm{d}^3 r \, \mathcal{P}_\sigma(\vec{r}) = 1 \tag{427}$$

that is, if we accomplish the search over the entire orbital-spin space, we find the searched electron with the 100% probability (although we can not say if it is the first one or the second etc). If we are not interested in what spin an electron has, then we get the probability density of finding some electron (with any spin) at point $\vec{r}$ by summing up over the both spins:

$$\mathcal{P}(\vec{r}) = \sum_\sigma \mathcal{P}_\sigma(\vec{r}) \tag{428}$$

and it is clear that

$$\int \mathcal{P}(\vec{r}) \, \mathrm{d}^3 r = 1 \tag{429}$$

It is now clear that the *electron density* $n(\vec{r})$ is proportional to the function $\mathcal{P}(\vec{r})$:

$$n(\vec{r}) = N \, \mathcal{P}(\vec{r}) \tag{430}$$

The proportionality constant must be the total number of electrons N, because the integral of the electron density over the whole space must give the total number of the electrons:

$$\int n(\vec{r}) \, \mathrm{d}^3 r = N \tag{431}$$

Recall that we have already explicitly encountered the electron density n of the N-electron system, or (which is an equivalent quantity), the charge density ρ generated by such a system: For the first time, it was in our study of the HF method [eq. (300)]. For the second time, we encountered density when studying the homogeneous electron gas, for which from the very definition of the problem, naturally, the density did not depend on the position [eq. (405)]. Then we worked with it, already spatially dependent, in the theory of Thomas and Fermi; see for instance eq. (413). *Now we have finally related the electron density to the wave function of the N-electron system.* Based on formulae (430), (428) and (426) we now write this relation explicitly:

$$\boxed{n(\vec{r}) = N \sum_\sigma \sum_{\sigma_2} \cdots \sum_{\sigma_N} \int \mathrm{d}^3 r_2 \cdots \int \mathrm{d}^3 r_N \, |\Psi(\vec{r}, \sigma, \vec{r}_2, \sigma_2, \dots, \vec{r}_N, \sigma_N)|^2} \tag{432}$$

In the following sections, we will need to be able to express the potential energy of an electron in a given external field using the electron density generated by that electron.

Calculation of $n(\vec{r})$ Using the Operator of the Density of Electrons.

This part -- it is written in a smaller font and narrower text -- does not need to be known, but it is recommended to at least look at it for interest.

Hustota elektrónov je fyzikálna veličina a preto sa dá definovať aj jej operátor; samotnú hustotu potom bude možné počítať ako kvantovomechanickú strednú hodnotu tohoto operátora v danom ľubovoľnom stave Ψ:

$$n(\vec{r}) = \langle\Psi|\hat{n}(\vec{r})|\Psi\rangle \tag{433}$$

Aké je vyjadrenie operátora hustoty elektrónov? Treba ho skonštruovať v súlade s princípom korešpondencie, teda na základe 2. postulátu kvantovej mechaniky preberanom v odseku 1.2.6. Vychádzať preto treba z formuly pre hustotu klasických elektrónov; tá sa potom dá prehlásiť za operátor. Klasické elektróny sú bodové častice, a preto v miestach, kde je niektorý z elektrónov, je hustota nekonečná, a v ostatných miestach priestoru je nulová. Teda

$$\boxed{\hat{n}(\vec{r}) = \sum_{i=1}^{N} \delta(\vec{r} - \hat{\vec{r}}_i)} \tag{434}$$

kde $\delta(\vec{r})$ je 3-rozmerná Diracova delta funkcia spĺňajúca podmienky [2]

$$\delta(\vec{r}) = \delta(-\vec{r})\,, \qquad \int \delta(\vec{r})\,\mathrm{d}^3r = 1 \tag{435}$$

a aj všeobecnejšiu podmienku

$$\int f(\vec{r})\,\delta(\vec{r} - \vec{r}_0)\,\mathrm{d}^3r = f(\vec{r}_0) \tag{436}$$

pre ľubovoľnú „slušne" sa správajúcu funkciu $f(\vec{r})$ a pre ľubovoľný bod $\vec{r}_0$ v priestore. Všimnime si, že pre vyššie zavedenú klasickú hustotu (a zároveň QM operátor) $\hat{n}(\vec{r})$ platí

$$\int \hat{n}(\vec{r})\,\mathrm{d}^3r = N \tag{437}$$

čo je v súlade s tým, čo od správne zavedenej hustoty očakávame. Teraz môžeme vyjadriť samotnú QM strednú hodnotu hustoty podľa (433):

$$\begin{aligned}
n(\vec{r}) &= \langle\Psi|\, \sum_{i=1}^{N} \delta(\vec{r} - \hat{\vec{r}}_i)\, |\Psi\rangle = \\
&= \sum_{i=1}^{N} \sum_{\sigma_1} \cdots \sum_{\sigma_N} \int \mathrm{d}^3r_1 \ldots \mathrm{d}^3r_N\ \Psi^*(\vec{r}_1, \sigma_1, \ldots, \vec{r}_N, \sigma_N)\ \delta(\vec{r} - \vec{r}_i)\ \Psi(\vec{r}_1, \sigma_1, \ldots, \vec{r}_N, \sigma_N) = \\
&= \sum_{i=1}^{N} \sum_{\sigma_1} \cdots \sum_{\sigma_N} \int \left(\prod_{\substack{j=1 \\ j \neq i}}^{N} \mathrm{d}^3r_j \right) \ \Psi^*(\vec{r}_1, \sigma_1, \ldots, \vec{r}_i \to \vec{r}, \sigma_i, \ldots, \vec{r}_N, \sigma_N)\ \Psi(\vec{r}_1, \sigma_1, \ldots, \vec{r}_i \to \vec{r}, \sigma_i, \ldots, \vec{r}_N, \sigma_N)
\end{aligned}$$

Vďaka antisymetrii vlnovej funkcie je jedno, aké je i – pre každé i dostaneme taký istý príspevok. Preto (ak použijeme $i = 1$)

$$n(\vec{r}) = N \sum_{\sigma_1} \cdots \sum_{\sigma_N} \int \left(\prod_{j=2}^{N} \mathrm{d}^3r_j \right) |\Psi(\vec{r}, \sigma_1, \ldots, \vec{r}_N, \sigma_N)|^2 \tag{438}$$

čo je vyjadrenie identické s výsledkom (432), ktorý sme dostali výpočtom pomocou hustoty pravdepodobnosti. Výpočet pomocou operátora $\hat{n}(\vec{r})$ možno považovať za elegantnejší.

12.2 Definition of the Problem under Study

Consider a system of N electrons. Suppose their motion is influenced by the external electrostatic potential $\hat{v}_{\text{ext}}(\vec{r})$ and by the mutual Coulomb interactions of the electrons. The external potential includes both the Coulomb field of the atomic nuclei and possible other electrostatic field, exactly as we assumed in the theory of Hartree and Fock. We do not consider spin interactions or other relativistic effects, so there are no spin-dependent terms in the Hamiltonian. Therefore, the total Hamiltonian will be

$$\hat{H} = \hat{T} + \hat{V}_{\text{ext}} + \hat{W} \tag{439}$$

where

$$\hat{T} = \sum_{i=1}^{N} \left(-\frac{\hbar^2}{2m} \right) \vec{\nabla}_i^2 \tag{440}$$

is the kinetic-energy operator of the whole system,

$$\hat{V}_{\text{ext}} = \sum_{i=1}^{N} \hat{v}_{\text{ext}}(\vec{r}_i) \tag{441}$$

is the potential-energy operator of the whole system and [see possibly also (411) and (412)]

$$\hat{v}_{\text{ext}}(\vec{r}) = -\sum_{I} \frac{1}{4\pi\varepsilon_0} \frac{Z_I e^2}{|\vec{r} - \vec{R}_I|} \;+\; \hat{v}_{\text{other}}(\vec{r}) \tag{442}$$

Thus, $\hat{V}_{\text{ext}}$ is the sum of the potential energies of the individual electrons. Every electron moves in the same external potential[56] $\hat{u}_{\text{ext}}(\vec{r}) = \hat{v}_{\text{ext}}(\vec{r})/(-e)$. This must be from the very definition of the problem, because all the N electrons belong to the same system (e.g. a molecule or a crystal) and every electron thus feels a field of the same nuclei plus a possible additional ("other") field. $\hat{v}_{\text{ext}}$ and $\hat{u}_{\text{ext}}$ are the same functions as v_{ext} and u_{ext} in the theory of Thomas and Fermi, but here we also add hats to them to emphasize that DFT is, unlike the TF model, a fully quantum theory. We were writing hat also in the HF theory [eq. (259), (278)]. Finally

$$\hat{W} = \frac{1}{2} \sum_{i,j=1}^{N}{}' \frac{e^2}{4\pi\varepsilon_0} \frac{1}{|\vec{r}_i - \vec{r}_j|} \tag{443}$$

is the operator of the electron-electron interaction energy.

The main task in the DFT is *to determine the ground-state electron density and energy.* The fundamental spatial variable is the density $n(\vec{r})$. Wave functions (or rather a set of single-particle orbitals) have, as we shall see, only an auxiliary role in DFT.

12.3 Potential Energy of Electrons in the External Field

The total energy of the system in any normalized state Ψ can be found as the quantum-mechanical expectation value of the energy operator:

$$E = \langle \Psi | \hat{H} | \Psi \rangle \;=\; \langle \Psi | \hat{T} | \Psi \rangle + \langle \Psi | \hat{V}_{\text{ext}} | \Psi \rangle + \langle \Psi | \hat{W} | \Psi \rangle \tag{444}$$

Let us now consider the contribution from the external field to show that it can be easily expressed in terms of density $n(\vec{r})$.

$$\begin{aligned} E_{\text{ext}} &= \langle \Psi | \hat{V}_{\text{ext}} | \Psi \rangle = \qquad (445) \\ &= \sum_{\sigma_1} \cdots \sum_{\sigma_N} \int \mathrm{d}^3 r_1 \ldots \mathrm{d}^3 r_N \; \Psi^*(\vec{r}_1, \sigma_1, \ldots, \vec{r}_N, \sigma_N) \left[\sum_{i=1}^{N} \hat{v}_{\text{ext}}(\vec{r}_i) \right] \Psi(\vec{r}_1, \sigma_1, \ldots, \vec{r}_N, \sigma_N) = \\ &= \sum_{i=1}^{N} \sum_{\sigma_1} \cdots \sum_{\sigma_N} \int \mathrm{d}^3 r_1 \ldots \mathrm{d}^3 r_N \; |\Psi(\vec{r}_1, \sigma_1, \ldots, \vec{r}_N, \sigma_N)|^2 \; \hat{v}_{\text{ext}}(\vec{r}_i) \end{aligned}$$

Due to the antisymmetry of the wave function, each of the N terms of this sum is the same. (No matter what i we take, the contribution for each i is the same.) Therefore, taking into account (432) we get

$$\boxed{E_{\text{ext}} = \int n(\vec{r}) \hat{v}_{\text{ext}}(\vec{r}) \mathrm{d}^3 r = \int \underbrace{(-e) n(\vec{r})}_{\rho(\vec{r})} \; \hat{u}_{\text{ext}}(\vec{r}) \mathrm{d}^3 r} \tag{446}$$

---

[56] In a strict sense, $u_{\text{ext}}(\vec{r})$ is called to be a potential (usually coming from the nuclei, i.e. positive), while $v_{\text{ext}}(\vec{r}) = -e u_{\text{ext}}(\vec{r})$ is the potential energy of a point particle with charge $(-e)$ in the field $u_{\text{ext}}(\vec{r})$. For the sake of brevity, we will also sometimes call $\hat{v}_{\text{ext}}(\vec{r}) = v_{\text{ext}}(\vec{r})$ potential. We use this quantity here in the sense of an operator; in QM we cannot consider an electron as a classical point particle. The actual potential energy of the electrons in the external field must be calculated from the wave function according to (445) or from the density according to (446).

Thus, even in quantum mechanics, the quantum mechanical expectation value of the potential energy of electrons in a given external field is calculated by the formula known from classical physics: as if the electrons were forming a continuous charge distribution with the charge density $\rho(\vec{r})$ and this charge distribution in the potential $\hat{u}_{\text{ext}}(\vec{r})$ has the electrostatic energy E_{ext}. This is exactly the way how we (in the atomic units) calculated this energy in the model of Thomas and Fermi, see the second term in (413).

12.4 Density as Basic Variable

While in sections 12.1 and 12.3 Ψ could have been any normalised wave function of the N- electrons system (even a time-dependent one), here by Ψ we will understand the ground-state wave function. We will assume the ground state of the system under study to be non-degenerate. The ground-state wave function is then unambiguously determined the the number N and by the external potential $\hat{v}_{\text{ext}}(\vec{r})$. To understand this, realise that if we know the number of electrons of a system and the course of the external potential, then we can unambiguously determine its Hamiltonian (439); whatever the system is, the contributions $\hat{T}$ and $\hat{W}$ always have the forms (440) and (443) (and for different systems, these formulae can only differ in different numbers N). And if we know the Hamiltonian, then at least in principle we can solve the stationary SchE and in this way to determine the wave function and energy of the ground state. Note also that if we have two different systems (for example two different molecules), but with the same numbers of electrons, then their Hamiltonians differ only in the external potentials. So once again – the ground state wave function Ψ is uniquely determined by the number of N and by the external potential $\hat{v}_{\text{ext}}(\vec{r})$. We do not write the hats on v_{ext} further for brevity.

The 1st theorem of Hohenberg and Kohn: reports that a less obvious statement – in reverse direction – also applies [13]:

v_{ext} *is a unique functional of the ground-state density* $n(\vec{r})$*, apart from a trivial additive constant.*

Alebo, trochu inými slovami [11], *the external potential* $v_{\text{ext}}(\vec{r})$ *is determined, within a trivial additive constant, by the electron density* $n(\vec{r})$*.*

A comment to elucidate the constant: We know that if we add any constant to potential energy or potential, nothing changes in physics of the system: the force on a particle is $\vec{F} = -\vec{\nabla} v_{\text{ext}}$ and the additive constant has no influence on the result of the derivative. Also in QM, only energy differences are important.

We also notice that if we know the electron density, the number of electrons in the system is clearly determined by it:

$$N = \int n(\vec{r})\,\mathrm{d}^3 r \tag{447}$$

A consequence: Since, according to the 1st HK theorem, the density of the ground state uniquely determines the potential, and thus also the Hamiltonian, the wave function of the ground state and of all other states are then uniquely determined (apart from trivial multiplication constants). *Therefore, all properties of the system all fully determined if just the ground-state density is given* [12].

Proof of the 1st HK theorem: it proceeds by *reductio ad absurdum*[57] and is strikingly simple:

Máme danú hustotu $n(\vec{r})$ istého základného stavu. Predpokladajme, že by 1. HK teoréma neplatila, teda že by existovali (aspoň) dva netriviálne odlišné vonkajšie potenciály, $v_{\text{ext}}(\vec{r})$ a $v'_{\text{ext}}(\vec{r})$, také, že by oba dávali (vyriešením SchR) tú istú hustotu základného stavu $n(\vec{r})$. Tie dva rôzne potenciály by nutne viedli ku dvom rôznym hamiltoniánom $\hat{H}$ a $\hat{H}'$ a tým aj ku rôznym vlnovým funkciám Ψ a Ψ' základných stavov. Predpokladajme normovanosť týchto vlnových funkcií. Energie základných stavov Ψ a Ψ' si označme E a E'. Potom platí

$$E = \langle\Psi|\hat{H}|\Psi\rangle\,, \qquad E' = \langle\Psi'|\hat{H}'|\Psi'\rangle$$

---

[57] clasically [re'duktió ad ab'surdum], in later times [re'dukcio ad ab'surdum] (from Latin)

Podľa všeobecne platného variačného princípu QM (167) platí[58]

$$E' = \langle \Psi' | \hat{H}' | \Psi' \rangle < \langle \Psi | \hat{H}' | \Psi \rangle$$

Použili sme ostrú nerovnosť, lebo podľa predpokladu uvažujeme sústavu s nedegenerovaným základným stavom. Ak teda Ψ' dáva minimum, tak Ψ musí dať vyššiu energiu. Ďalej to upravujme:

$$E' < \langle \Psi | \hat{H}' | \Psi \rangle = \langle \Psi | \hat{H} | \Psi \rangle + \langle \Psi | \hat{H}' - \hat{H} | \Psi \rangle = E + \langle \Psi | \hat{H}' - \hat{H} | \Psi \rangle$$

Druhý člen na pravej strane sa dá upravovať s prihliadnutím ku (439) (445) a (446) takto:

$$\langle \Psi | \hat{H}' - \hat{H} | \Psi \rangle = \langle \Psi | \hat{V}'_{\text{ext}} - \hat{V}_{\text{ext}} | \Psi \rangle = \int n(\vec{r}) \left[v'_{\text{ext}}(\vec{r}) - v_{\text{ext}}(\vec{r}) \right] \mathrm{d}^3 r$$

Hustota $n(\vec{r})$ je totiž podľa predpokladu len jedna pre oba potenciály. Takže dostávame

$$E' < E + \int n(\vec{r}) \left[v'_{\text{ext}}(\vec{r}) - v_{\text{ext}}(\vec{r}) \right] \mathrm{d}^3 r \tag{448}$$

Obdobný výpočet môžeme spraviť tak, že navzájom zameníme čiarkované a nečiarkované veličiny:

$$E < \langle \Psi' | \hat{H} | \Psi' \rangle = \langle \Psi' | \hat{H}' | \Psi' \rangle + \langle \Psi' | \hat{H} - \hat{H}' | \Psi' \rangle = E' + \langle \Psi' | \hat{H} - \hat{H}' | \Psi' \rangle \quad \Rightarrow$$

$$E < E' + \int n(\vec{r}) \left[v_{\text{ext}}(\vec{r}) - v'_{\text{ext}}(\vec{r}) \right] \mathrm{d}^3 r \tag{449}$$

Sčítaním nerovností (448) a (449) dostávame

$$E' + E < E + E'$$

which is an obvious non-sense. Predpoklad, že by (v prípade sústav s nedegenerovaným základným stavom) pre danú hustotu $n(\vec{r})$ existoval viac ako jeden vonkajší potenciál v_{ext}, teda nie je správny, lebo vedie ku nezmyselnému dôsledku.[59]

12.5 The Variational Principle

The 2^{nd} theorem of Hohenberg and Kohn:

There exists a universal [i.e. independent of $v_{\text{ext}}(\vec{r})$] functional of the density, $F_{\text{HK}}[n]$ such that the expression

$$E_v[n] \equiv \int v_{\text{ext}}(\vec{r})\, n(\vec{r})\, \mathrm{d}^3 r \;+\; F_{\text{HK}}[n] \tag{450}$$

has as its minimum value the exact ground-state energy of the N-electron system for any given potential $v_{\text{ext}}(\vec{r})$. The density that minimises the functional $E_v[n]$ is the exact ground-state density [13].

Mathematically expressed, the variational principle of the 2^{nd} Hohenberg and Kohn theorems is that

$$E_0 \leq E_v[n] \tag{451}$$

for any trial density that satisfies the conditions

$$n(\vec{r}) \geq 0\,, \qquad \int n(\vec{r})\, \mathrm{d}^3 r = N \tag{452}$$

---

[58] Ten princíp samozrejme platí aj pre mnohočasticové sústavy. My sme si ho zapisovali len pomocou jednočasticových vlnových funkcií, lebo vtedy sme poznali len také, ale všimnime si, že na jeho platnosti a dôkaze sa nič nezmení, keď namiesto $f(\vec{r})$ a $\mathrm{d}^3 r$ budeme uvažovať mnohočasticové funkcie, dokonca aj so spinovými premennými.

[59] Nešlo teda o typický dôkaz sporom, lebo v takom by sa prišlo do sporu s nejakým vopred vysloveným predpokladom. My sme len prišli ku nezmyslu. Ale aj dôkazy sporom sa radia medzi logicko-matematické postupy s názvom *reductio ad absurdum*.

Proof of the 2^{nd} HK theorem:

We already know how to express the energy of a given system in any normalised N-electron state Ψ – we can do it using formula (444). For the purpose of our proof, we would now need to rewrite this formula so that the energy expressed by it depends on the density of the electrons, and not on the wave function. Is it possible? Yes; realise that the 1st HK theorem has the consequence, mentioned below it, that the ground-state wave function itself is fully determined by the ground state (GS) density, and is therefore also a functional of the density. And since everything can be calculated from the wave function, e.g. the kinetic energy (in the sense of the QM expectation value), the electron-electron interaction energy, also the total energy, we can also consider these as density functionals. Therefore, we can really write for the energy (444) corresponding to state Ψ and also for the individual components of the energy that they can be expressed as density functionals (and in this expression there will no longer be a dependence on wave functions):

$$E_v[n] = T[n] + W[n] + \int n(\vec{r}) v_{\text{ext}}(\vec{r}) \mathrm{d}^3 r \tag{453}$$

where

$$T[n] = \langle \Psi | \hat{T} | \Psi \rangle \,, \qquad W[n] = \langle \Psi | \hat{W} | \Psi \rangle \tag{454}$$

We now define the functional

$$F_{\text{HK}}[n] \equiv T[n] + W[n] \tag{455}$$

The total energy can then be expressed by the formula of the form (450). By this we have completed the part of the proof of the 2nd HK theorem, i.e. we have proved that *there exists* some universal functional $F_{\text{HK}}[n]$, using that we can also express the energy of the GS.

Above, we added the index v to the total energy, highlighting that it depends on the external potential $v_{\text{ext}}(r)$, which, as we already know, is also a functional of the density, so it would not be necessary to talk about the dependence on $v_{\text{ext}}(r)$. However, for formal reasons, if we need and want to, we can also understand $E_v[n]$ as a value that depends on $v_{\text{ext}}(r)$ and on $n(\vec{r})$ as on two independent functions. For a given chosen potential $v_{\text{ext}}(\vec{r})$, we can then test different (trial) densities $n(\vec{r})$ and find out which of them gives the lowest value of the functional $E_v[n]$. If we took a density other than the GS density, then this density would correspond to a wave function other than the GS function. We remind that with these considered variations of the density, we keep the external potential $v_{\text{ext}}(r)$ *in formula* (453) fixed in its originally given form. Based on the variational principle of QM taken over in section 6.1, it must then be true that if the density and thus the wave function deviates from the wave function of the ground state, the value of $E_v[n]$ must necessarily increase. By this we have completed the proof of the 2^{nd} HK theorem. We also note that although the variational principle taken over in section 6.1 has been written for the case of a single particle, its generalization to N particles would be trivial.

An alternative formulation of the potential-and-density relation also exists. This formulation also eliminates the problem of possible degenerate ground state, the solution of which we have omitted in our exposition. The authors of this formulation are Levy and Lieb [11, 12]. In their formulation, they provide an alternative proof of the 2^{nd} theorem of Hohenberg and Kohn.

We now see that the heuristic approach of Thomas and Fermi, in which the minimum of the functional (415) is sought with respect to density, has been given a rigorous basis by the theorems of Hohenberg and Kohn (as well as by the works of Levy and Lieb). However, the problem of the TF model is in the qualitatively inaccurate expression of the functionals $T_{\text{TF}}[n]$ and $W[n]$. For example, $W[n]$ in the TF theory includes only the classical part of the interaction energy; the exchange energy is completely missing, as well as the correlation energy. Even the work of HK [13] itself, although it proves that there exists some universal functional, it does not find its exact explicit form, only approximations for certain limiting cases. Even later, this functional could never be found exactly. However, a suitable method was proposed in 1965 by Kohn and Sham, which we will talk about at least a little in the next section.

12.6 Kohn-Sham Ansatz

The difficult-to-solve system of N interacting electrons was replaced by Kohn and Sham (KS) in their approach to the DFT by an auxiliary system of N independent, i.e. non-interacting electrons [14]. *This auxiliary system must be* *such as to provide* ***the same electron density*** *of the ground state as the original problem of interacting particles.* This is the KS ansatz (an "educated guess" how to solve a given problem [Wikipedia]). In English-written literature, the term *Kohn-Sham mapping* is sometimes used to indicate that ***Kohn and Sham mapped the system of N interacting electrons to a system of N non-interacting ones, but so as the electron density to be the same.*** Due to the assumption of the independent electrons, their kinetic energy is easily determined. Of course, independent electrons would only have the kinetic energy plus the energy E_{ext} of the interaction with a given external potential. This energy would be completely wrong and we want the correct ground state energy. Therefore, we must add the missing contributions $E_{\text{Hartree}}[n]$ and $E_{\text{xc}}[n]$ to the following expression for the total energy:

$$E_{\text{KS}}[n] = T_s[n] + \int n(\vec{r}) v_{\text{ext}}(\vec{r}) \mathrm{d}^3 r + E_{\text{Hartree}}[n] + E_{\text{xc}}[n] \tag{456}$$

where T_s is the mentioned kinetic energy of the *neinteracting* electrons, E_{Hartree} is the Hartree energy and E_{xc} is the ***exchange-correlation energy*** (i.e. the exchange one and the correlation one added together). For understanding, (456) must be compared with (453). It can be seen that the Hartree energy is a (quantitatively significant) part of the total interaction energy W. But the total W in addition contains the QM contributions: the exchange-correlation energy, so we had to add these to (456) in the form of the E_{xc} term. It is obvious that the exchange-correlation energy defined in this way also contains a part of the kinetic energy, because we have included into T_s only such a kinetic energy that corresponds to the independent electrons.

The system of independent electrons can usually be described by a one-determinant wave function, as it is in the Hartree-Fock theory. So in DFT we introduce auxiliary (spin) orbitals, which the Slater determinant is composed of (but it doesn't really need to be built and evaluated). These orbitals, when calculated correctly, must generate the correct ground state density. The kinetic energy T_s is then determined from these auxiliary orbitals as in the HF method, i.e. it is (in principle) the *exact* kinetic energy of the *non-interacting* electrons. Thus, not as in the Thomas-Fermi model, in which the kinetic energy is determined very inaccurately directly from the density according to (410). Realise that the TF kinetic energy is really only correct for homogeneous gas of non-interacting electrons (but in the TF model it is also used for inhomogeneous one, which then has its undesirable consequences.) In DFT, we have transformed the problem to the problem of non-interacting particles, but they generally form an *inhomogeneous* gas and therefore it is correct to calculate their kinetic energy from the orbitals. Thus, in practical use of DFT, we do not completely get rid of wave functions; the one-particle ones are needed.

A special feature of DFT is the mentioned exchange-correlation energy, for which various approximations are being proposed. The basic model is called the ***local-density approximation*** [14] (LDA):[60]

$$E_{\text{xc}}^{\text{LDA}}[n] = \int n(\vec{r})\, \epsilon_{\text{xc}}(n(\vec{r}))\, \mathrm{d}^3 r \tag{457}$$

where $\epsilon_{\text{xc}}(n)$ is the exchange-correlation energy of *homogeneous* electron gas of density n per one electron. The dependence of ϵ_{xc} on n for a homogeneous electron gas can be calculated with sufficient accuracy by specialised methods. Using formula (457), we can then approximaltely determine the exchange-correlation energy $E_{\text{xc}}[n]$, which makes it possible to practically use the density-functional theory. The LDA model appears to be a rough approximation because in atoms, molecules and crystals, density exhibits strong inhomogeneities. It is a similar difficulty as in the Thomas-Fermi model. Rather surprisingly, however, DFT with the LDA functional is a very good approximation, qualitatively better than the TF model. During the development of DFT, of course, various improvements were developed for the $E_{\text{xc}}[n]$ function, e.g. gradient corrections, under which ϵ_{xc} depends not only on the density at a given location, but also on its gradient [see, for instance, the contribution of authors J.P. Perdew, S. Kurth in [15], formula (1.230)]. In this respect, e.g. the **GGA** *(Generalised Gradient Approximation)* model, and in particular its implementation PBE (acronym according to the authors' names) is successful [15, 12]. However, no functional proposed so far is accurate (and it can hardly be expected that such a one, being also practically usable, will ever be found). Nevertheless, DFT in conjunction with the KS mapping (of the interacting to the non-interacting problem) is the most widely used *ab initio* method.[61] This is because it is a good compromise between accuracy and computational demands. Walter Kohn is the winner of the 1998 Nobel Prize in Chemistry for his key work on DFT.

---

[60] a well-know abbreviation in electronic structure theory

[61] The term *ab initio* is, as already mentioned, used primarily for methods of quantum chemistry which employ only basic physical laws and mathematical-numerical procedures to solve the relevant equations. In a broader sense, *ab initio* methods include DFT methods, although quantum chemists do not usually call them that, because the designs of the $E_{\text{xc}}[n]$ functionals tend to be constructed in ways that do not guarantee uniform quality of results for different systems. E.g. for some molecules we get highly accurate results with DFT,

Finally, we only verbally mention that the KS ansatz, after accomplishing variation (minimum search) similar to that in the TF method, leads to the ***Kohn-Sham equations*** for the above-mentioned auxiliary one-particle orbitals and the corresponding one-particle eigenenergies [14, 12, 11]. The KS equations by their form resemble the HF equations a little.

---

for many others (even solids) also very good, but in some cases DFT (with a specific functional, e.g. LDA) fails. This can often be helped by designing another functional, but this can be worse for some other structures, or for determining some other parameters of a given structure.

A Expansion of a wave function in a complete set of functions

A.1 Functions of one variable. Expectation values of quantities

Assume we have a complete set of functions

$$\{u_n(x)\} \tag{A.1}$$

We omit writing the range of the values of the index n. It usually is infinite with the indexing starting either from 0 or from 1 or from $-\infty$, possibly from some other values. This does not matter now and a concrete choice depends on a concrete task to be solved. The essential property is that the functions form a complete set which means that any function can be written as a linear combination of the functions $u_n(x)$:

$$\boxed{\psi(x) = \sum_n c_n u_n(x)} \tag{A.2}$$

In addition to the completeness, let us assume also *orthogonality* of of the basis functions $u_n(x)$:

$$\int_{-\infty}^{\infty} u_m^*(x) u_n(x)\,\mathrm{d}x = 0 \qquad \text{ak } m \neq n \tag{A.3}$$

Such functions are definitely linearly independent: no one of them can be expressed as a linear combination of the *remaining* functions.[62] In addition, there are practical (and in quantum mechanics also physical) reasons to assume

$$\int_{-\infty}^{\infty} |u_n(x)|\,\mathrm{d}x = 1 \qquad \text{pre každé } n \tag{A.5}$$

that is the normalisation to unity. The last two properties (orthogonality plus normalisation) are called *orthonormality* in a single word. We express it

$$\boxed{\int_{-\infty}^{\infty} u_m^*(x) u_n(x)\,\mathrm{d}x = \delta_{mn}} \tag{A.6}$$

The above formulae and procedure can also be understood purely as some mathematical formalism. Since we want them to use for quantum physics, let us suppose that $\{u_n(x)\}$ is a set of eigenfunctions fo some hermitian operator and that $\psi(x)$ is a wave function.[63] It has to be normalised to unity:

$$\int_{-\infty}^{\infty} |\psi(x)|^2\,\mathrm{d}x = 1 \tag{A.7}$$

We now subtitute expansion (A.2) into this condition. Using orthonormality (A.6), we arrive at

$$\boxed{\sum_n |c_n|^2 = 1} \tag{A.8}$$

---

[62] In order to demonstrate the above mentioned linear independency, let us try, for examples, to express the function $u_k(x)$ as s linear combination of the basis functions $u_n(x)$:

$$u_k(x) = \sum_n c_n u_n(x) \tag{A.4}$$

Multiple the equation from the left by the function $u_m^*(x)$ (with m being any of the possible values of the index) and integrate. We obtain

$$\int_{-\infty}^{\infty} u_m^*(x) u_k(x)\,\mathrm{d}x = \sum_n c_n \int_{-\infty}^{\infty} u_m^*(x) u_n(x)\,\mathrm{d}x$$

i.e., employing the orthonormality (A.6),

$$\delta_{mk} = \sum_n c_n \delta_{mn} = c_m, \qquad \forall m$$

which in other sybols reads $c_n = \delta_{nk}$. Thus, we obtain the only non-vanishing term in the expansion (A.4):

$$u_k(x) = \sum_n \delta_{nk} u_n(x) = u_k(x)$$

[63] It may depend also on time but we need not stress this possible dependence now.

This result suggests that *the number* $|c_n|^2 = P_n$ *should be interpreted as a probability that the particle is in a state described by the wave function* $u_n(x)$. As the probabilities sum up to 1, it means that one of the possibilities will definitly happen. Since the particle exists, it has to be in some state. And since our linear combination (A.2) includes all possibilites, one of them must occur. For example, if we roll the dice, the particular probabilities are $1/6$ and they sum up to 1. The quantities c_n are called probability amplitudes. We can find an expression for them as follows: we multiply expnasion (A.2) from the left side by the function $u_m^*(x)$ and integrate (we have been doing manipulations of this kind). Using the orthonormality, we arrive at

$$\boxed{c_n = \int_{-\infty}^{\infty} u_n^*(x)\,\psi(x)\,\mathrm{d}x} \tag{A.9}$$

We provide further support for the above interpretation of the numbers $|c_n|^2$ as follows: We have said that the functions $u_n(x)$ are eigenfunctions of some hermitian operator, which we now denote as $\hat{F}$:

$$\hat{F}u_n(x) = F_n u_n(x) \tag{A.10}$$

where F_n is (a real) eigenvalues corresponding to the eigenfunction $u_n(x)$. Assume the particle be in a state described by a wave function $\psi(x)$; what is then the expectation value of the quantity F in this state? According to the 2^{nd} postulate of quantum mechanics, this expectation value can be calculated as follows:

$$\bar{F} = \int_{-\infty}^{\infty} \psi^*(x)\,\hat{F}\,\psi(x)\,\mathrm{d}x \tag{A.11}$$

By substituting expansion (A.2), using (A.10) and orthonormality (A.6), we obtain

$$\boxed{\bar{F} = \sum_n |c_n|^2\,F_n} \tag{A.12}$$

Let us now compare this result with formula (1) in our discussion of the bag with the coins (section 1.2.1). We see that the numbers $|c_n|^2$ have a probabilistic interpretation: $|c_n|^2$ is the probability of finding the value F_n in a measurement of the quantity F. According to (A.10) we can alternatively formulate this finding as follows: $|c_n|^2$ is the probability to find the system in the state u_n. The complex number c_n itself is the probability *amplitute* of the result F_n, i.e. of the system being in the state u_n.

Therefore, if some c_n in expansion (A.2) vanishes, the probability of finding the particle in the corresponding state u_n is zero. The content of this section complements the argumentation to the 2^{nd} postulate of QM (section 1.2).

Although we have done all the argumentation for a function of the single variable x, a generalisation to three variables $(x, y, z) = \vec{r}$ would be trivial and coul be done just by a simple renaming of x to $\vec{r}$ and by using triple integrals d^3r instead of simple ones.

A.2 Generalisation of the Argumentation to Many Particle Wave Function

Ak teraz máme vlnovú funkciu popisujúcu dve častice, je to funkcia dvoch vektorových premenných $\Psi(\vec{r}_1, \vec{r}_2)$. Chceme ju rozvinúť do nejakého úplného systému funkcií. Tento úplný systém musí tiež byť tvorený funkciami dvoch premenných:

$$\Psi(\vec{r}_1, \vec{r}_2) = \sum_n C_n\, w_n(\vec{r}_1, \vec{r}_2) \tag{A.13}$$

Typicky používaným spôsobom konštrukcie úplnej sústavy funkcií dvoch premenných je vyrobiť ich z jednočasticových bázových funkcií:

$$w_{n_1,n_2}(\vec{r}_1, \vec{r}_2) = u_{n_1}(\vec{r}_1)\, u_{n_2}(\vec{r}_2) \tag{A.14}$$

Že sa to tak dá, sme sa presvedčili na prednáške o Pauliho princípe (časť 8.6); pozri poznámka pod čiarou ku formule (227). Index n v (A.13) teda môže byť nejaký kompozitný index: $n \equiv (n_1, n_2)$, ale to je len technická záležitosť. Funkcie w_{n_1,n_2} tvoria úplnú ortonormovanú sústavu funkcií dvoch premenných:

$$\int w^*_{m_1,m_2}(\vec{r}_1, \vec{r}_2)\, w_{n_1,n_2}(\vec{r}_1, \vec{r}_2)\, \mathrm{d}^3r_1\, \mathrm{d}^3r_2 = \delta_{m_1,m_2}\delta_{n_1,n_2} \qquad \text{(ortonormovanosť)}$$

ako sa dá ľahko presvedčiť.[64] Rozvoj (A.13) podrobnejšie napíšeme

$$\Psi(\vec{r}_1, \vec{r}_2) = \sum_{n_1} \sum_{n_2} C_{n_1,n_2}\, u_{n_1}(\vec{r}_1)\, u_{n_2}(\vec{r}_2) \tag{A.15}$$

Interpretácia je taká, že C_{n_1,n_2} je amplitúda pravdepodobnosti toho, že časticu 1 nájdeme v stave u_{n_1} a zároveň časticu 2 v stave u_{n_2}.

Platí teda, že ak je niektoré C_{n_1,n_2} v rozvoji (A.15) nulové, tak to znamená, že je nulová pravdepodobnosť nájsť časticu 1 v stave u_{n_1} a zároveň časticu 2 v stave u_{n_2}. Toto sa priamo využíva v dôkaze Pauliho princípu v časti 8.6; pozri vyjadrenie (231), kde pre fermióny vychádza $C_{n,n} = 0$, čiže nulová pravdepodobnosť nájsť dve častice v tom istom stave.

B Partition of the Eigenenergy of the Hydrogen Atom of a Similar Ion into the Kinetic and Potential Energies

We are going to split energy (147) into its kinetic and potential parts. Before doing so, we note that the ways how it can be expressed may seem to be very different each other (although they are equivalent):

$$E_n \;=\; -\frac{m}{\hbar^2}\left(\frac{e^2}{4\pi\varepsilon_0}\right)^2 \frac{1}{2}\frac{Z^2}{n^2} \;=\; \frac{1}{2}\frac{1}{4\pi\varepsilon_0}\frac{(Ze)(-e)}{a_Z}\frac{1}{n^2} \;=\; -\frac{1}{2}\frac{\hbar^2}{2m}\frac{1}{a_Z^2}\frac{1}{n^2} \tag{B.1}$$

We have introduced the first of the expressions in our derivation of the eigenenergies. To find the second one, which can easily be done, we used an expression including the 1st Bohr radius of a hydrogen-like ion:

$$a_Z = \frac{a_{\mathrm{B}}}{Z} \tag{B.2}$$

where a_{B} is the standard 1st Bohr radius, i.e. the one for the hydrogen atom:

$$a_{\mathrm{B}} = \frac{\hbar^2 4\pi\varepsilon_0}{me^2} = 0{,}5291772083\,.\,10^{-10}\,\mathrm{m} \tag{B.3}$$

The second expression reminds us the coulombic potential energy between a nucleus of the charge Ze and an electron with the charge $-e$. But it is just half of such energy. The third expression for the eigenenergy in (B.1) could, of course, be easily derived and we have chosen its form to remind us the kinetic energy $p^2/(2m)$, where the momentum $p = \hbar/a_Z$ and there is also the multiplicator $-1/2$ there.

If we did a calculation according the the Bohr model (which can be done easily and we will not give the derivation here), we would obtain the kinetic and potential energies as follows:

$$\boxed{E_n^{\mathrm{kin}} = -E_n} = \frac{\hbar^2}{2m}\frac{1}{a_Z^2}\frac{1}{n^2}\,, \qquad \boxed{E_n^{\mathrm{pot}} = 2E_n} = \frac{1}{4\pi\varepsilon_0}\frac{(Ze)(-e)}{a_Z}\frac{1}{n^2} \tag{B.4}$$

The particular portions of the energy really have the classical (although not physically accurate) interpretation as we have written above.

The calculations of the portions of the total energy E_n can, of course, be done correctly in the full quantum-mechanical way. A straightforward way is as follows: The quantum-mechanical expectation value of the kinetic energy of an electron in a hydrogen atom or like ion in the state $\psi_{nlm}(\vec{r})$ is

$$\bar{T} = \int \psi^*_{nlm}(\vec{r}) \left(-\frac{\hbar^2}{2m}\Delta\right) \psi_{nlm}(\vec{r})\, \mathrm{d}^3 r \tag{B.5}$$

Analogously, the quantum-mechanical expectation value of the potential energy of the electron in this state in the field of a nucleus of charge Ze is

$$\bar{V} = \int \psi^*_{nlm}(\vec{r}) \left(-\frac{1}{4\pi\varepsilon_0}\frac{Ze^2}{r}\right) \psi_{nlm}(\vec{r})\, \mathrm{d}^3 r \tag{B.6}$$

---

[64] Presvedčiť sa o ortonormovanosti je ľahké. Presvedčiť sa o úplnosti býva zvyčajne ťažšie, ale nemusíme si prednášku matematicky a technicky príliš komplikovať.

where $\psi_{nlm}(\vec{r})$ is the respective eigenstate (150). Accomplishing relatively lengthy calculations, we would arrive at

$$\bar{T} = E_n^{\text{kin}}\,, \qquad \bar{V} = E_n^{\text{pot}} \tag{B.7}$$

that is the same result as from the Bohr model. It could be elegantely derived using the virial theorem which is a quite general one, applying to many situations in both classical and quantum physics and it says that $\bar{V} = -2\bar{T}$. But we have not derived the theorem.

Finally, it is worth recalling that the total energy of the hydrogen atom (or a like ion) has a sharp (define, zero-uncerytainty) value ***in the state*** $\boldsymbol{\psi_{nlm}(\vec{r})}$. It follows from the fact that it is an eigenenergy of the operator of the total (kinetic plus potential) energy. We would easily calculate in mathematically: we would find the mean quadratic deviation (the second power of the uncertainty) vanishing. ***However, neither the kinetic energy alone, nor the potential energy, has a sharp value in this state.*** It is because the energy E_n^{kin} is not an eigenenergy of the kinetic-energy operator $\hat{T}$. Similarly, E_n^{pot} is not an eigenenergy of the potential-energy operator $\hat{V}$ of the system under study. As for these particular energies, we can only undestand them as mean values. In this respect, it is useful to realise that $[\hat{T}, \hat{V}] \neq 0$.

C Search for a Local Extremum of a Functions of Many Complex Variables

Let f be a complex function of the complex variable $z = x + \mathrm{i}y$. Therefore, we also can imagine it as a function of two real variables x, y. Assume that the partial derivatves of f with respect to the real variables x, y exist. How do we calculate its *partial* derivative withg respect to the complex variable z? As follows:

$$\left.\frac{\partial f}{\partial z}\right|_{z^*} = \frac{\partial f}{\partial x}\left.\frac{\partial x}{\partial z}\right|_{z^*} + \frac{\partial f}{\partial y}\left.\frac{\partial y}{\partial z}\right|_{z^*} \tag{C.1}$$

since x, y are mutually independent real variables. In the procedure we are talking about, we will also formally consider z and z^* as mutually independent variables, although this may sound strange. The vertical lines $|_{z^*}$ emphasize that in the partial derivatives with respect to z we consider z^* as a constant. We express the complex variables z, z^* using x, y by the equations

$$z = x + \mathrm{i}\,y\,, \qquad z^* = x + \mathrm{i}\,y \tag{C.2}$$

and the corresponding inverse relations are

$$x = \frac{1}{2}(z + z^*)\,, \qquad y = \frac{1}{2\mathrm{i}}(z - z^*) \tag{C.3}$$

Then we can easily calculate the partial derivatives with respect to z and analogously also with respect to z^*; we obtain

$$\left.\frac{\partial f}{\partial z}\right|_{z^*} = \frac{1}{2}\left(\frac{\partial f}{\partial x} - \mathrm{i}\,\frac{\partial f}{\partial y}\right), \qquad \left.\frac{\partial f}{\partial z^*}\right|_{z} = \frac{1}{2}\left(\frac{\partial f}{\partial x} + \mathrm{i}\,\frac{\partial f}{\partial y}\right) \tag{C.4}$$

They are called ***Wirtinger derivatives***. We can immediately invert the two equations (C.4) as follows:

$$\frac{\partial f}{\partial x} = \left.\frac{\partial f}{\partial z}\right|_{z^*} + \left.\frac{\partial f}{\partial z^*}\right|_{z}, \qquad \frac{\partial f}{\partial y} = \mathrm{i}\left(\left.\frac{\partial f}{\partial z}\right|_{z^*} - \left.\frac{\partial f}{\partial z^*}\right|_{z}\right) \tag{C.5}$$

Now we should look for some extremal point of the function f in the variables x, y. Therefore, we write down the necessary conditions for the extremum: $\partial f/\partial x = 0$, $\partial f/\partial y = 0$. Using the above written equations, we fid out that this pair of conditions is equivalent to the pair of equations $\partial f/\partial z = 0$, $\partial f/\partial z^* = 0$. That is, the following equivalence holds:

$$\boxed{\left(\frac{\partial f}{\partial x} = 0 \quad \wedge \quad \frac{\partial f}{\partial y} = 0\right) \qquad \Longleftrightarrow \qquad \left(\frac{\partial f}{\partial z} = 0 \quad \wedge \quad \frac{\partial f}{\partial z^*} = 0\right)} \tag{C.6}$$

The formal constancy of the complex variable no derivative is taken with respect to is no longer emphasized here for brevity. The generalisation of this procedure for more complex variables z_1, z_2, ..., z_p is straightforward. Finally, we emphasize the need to strictly distinguish partial derivatives with respect to z from the total derivatives with respect to z. In the situation we are studying, a total derivation does not even exist, because the function (182) does not satisfy the Cauchy-Riemann conditions.

D The Helium Atom and Like Ions by the Variational Method

We solve this problem as an exercise.

D.1 Formulation of the Task and the Proposed Form of Its Solution

The system under study can be H^-, He, Li^+, Be^{2+}, B^{3+}, C^{4+}, ... The Hamiltonian of each of such two-electron systems has the form

$$H = -\frac{\hbar^2}{2m}\nabla_1^2 - \frac{\hbar^2}{2m}\nabla_2^2 - \frac{1}{4\pi\varepsilon_0}\frac{Ze^2}{r_1} - \frac{1}{4\pi\varepsilon_0}\frac{Ze^2}{r_2} + \frac{1}{4\pi\varepsilon_0}\frac{e^2}{|\vec{r}_1 - \vec{r}_2|} \tag{D.1}$$

Z is the number of protons in the nucleus which can be 1, 2, 3, ... The numbers of electrons in the cloud for the system under study is always 2. The problem to be solved is

$$H\psi(\vec{r}_1, \vec{r}_2) = E\psi(\vec{r}_1, \vec{r}_2) \tag{D.2}$$

and we want to determine its lowest-energy solution, i.e. the ground state. We will use the variational principle of quantum mechanics:

$$\frac{\langle\psi|H|\psi\rangle}{\langle\psi|\psi\rangle} \geq E_0 \tag{D.3}$$

We choose

$$\psi(\vec{r}_1, \vec{r}_2) = \phi(\vec{r}_1)\phi(\vec{r}_2) \tag{D.4}$$

to be our trial wave function. In this form, $\phi(\vec{r})$ is the ground state of the hydrogen atom or a like ion and we will write its form in next section. The two-particle wave function (D.4) has a very simple product form meaning that the electrons are considered as if were mutually independent, i.e. non-interacting with each other and moving in the field of the helium atom (or of a similar ion). We will include their interaction later in an indirect approximate way. We will do it by chosing the auxiliary single-particle function ψ such that it will depend on a certain parameter; the value of the parameter will be set as if the function did not “feel” the complete field of the nucleus, but only a field partially screened (or weakened) by the effect of the other electron. Although the interaction of the electron with the other one will not be included explicitly, it will at least be included indirectly. Below we specify further steps of our procedure.

D.2 The Ground State of a Hydrogen-Like Ion

This is an auxiliary section in which we will “derive” how the wave function of a hydrogen-like ion (i.e. of a single-electron system, countrary to our main task, which is a two-electron problem) looks like. Although we have derived the wave function, we perhaps may not have it on hand right now and looking into literature we may perhaps find just the hydrogen-atom wavefunction only (i.e. for the proton number 1) and not the wave functions for the hydrogen-like ions. For the ions, we get the wave function by simply modifying the hydrogen wave function as described in the following lines.

We know the the hydrogen atom groud state has the $1s$-type wave function, that is

$$\phi_{100}(\vec{r}) = \frac{1}{(\pi a_\mathrm{B}^3)^{1/2}}\, e^{-r/a_\mathrm{B}} \tag{D.5}$$

where

$$a_\mathrm{B} = \frac{\hbar^2 4\pi\varepsilon_0}{me^2} \tag{D.6}$$

is the 1st Bohr radius. We aim to find a generalisatin of wave function (D.5) valid for a hydrogen-like ion with a nucleus of charge βe. It is the dependence of the single-particle wave function of type (D.5) on the charge of the nucleus that will now be important to us. It is obvious that one of the two es in the product ee in expression (D.6) originates from the nucleus, the other from the electron. Therefore, we obtain the wave function of the hydrogen-like ion from (D.5) by the substitution

$$e^2 \longrightarrow \beta e^2$$

from which the substitution

$$a_{\mathrm{B}} \longrightarrow a_\beta = \frac{a_{\mathrm{B}}}{\beta}$$

follows. βe is the nuclear charge of the hydrogen-like ion under consideration. We stress that this β is just an auxiliary parameter introduced independently of the true nuclear charge value Z of the two-electron ion. In what follows, for brevity, we will denote the standard 1st Bohr radius simply a:

$$a \equiv a_{\mathrm{B}}$$

For the wave function of a hydrogen-like ion, we get the form

$$\boxed{\phi(\vec{r}) = \frac{1}{(\pi a_\beta^3)^{1/2}}\, e^{-r/a_\beta} = \left(\frac{\beta^3}{\pi a^3}\right)^{1/2} e^{-\beta r/a}} \tag{D.7}$$

Thus, it is a wave function of the *one-electron* ion. It would correspond to the hydrogen atom wave function in the special case of $\beta = 1$. In other special case, $\beta = 2$, it would corresponds to the wave function of the He$^+$ ion.

Our goal pursued in the following paragraphs is to find a solution to the problem (D.2). The corresponding wave function ψ is thus understood as the wave function of a two-electron atom or ion. Wave function (D.7) will serve as an auxiliary mathematical object to achieve the stated goal.

D.3 The Energy for the Chosen Wave Function ψ

Thus, the trial helium atom (or a like ion) wave function will be [see (D.4)]

$$\boxed{\psi(\vec{r}_1, \vec{r}_2) = \frac{\beta^3}{\pi a^3}\, e^{-\beta(r_1+r_2)/a}} \tag{D.8}$$

As we see, it depends on the parameter β. We are not going to determine the value of this parameter immediately, but we reveal in advance that we will later consider it a variation parameter. The proposed wave function is normalised to unity ($\int \psi^*\psi\, \mathrm{d}^3r_1\, \mathrm{d}^3r_2 = 1$, which can easily be verified since $\int \phi^*\phi\, \mathrm{d}^3r = 1$). Therefore, in our application of the variational principle (D.3), it is not necessary to write the denominator and we write the energy for the proposed state ψ as follows:

$$\boxed{E = E(\beta) = \int \psi^*(\vec{r}_1, \vec{r}_2)\, H\, \psi(\vec{r}_1, \vec{r}_2)\, \mathrm{d}^3r_1\, \mathrm{d}^3r_2} \tag{D.9}$$

We emphasized that the value of this energy depends on the auxiliary parameter β. On the RHS of this equation, the parameter β is found only in the wave function. The Hamiltonian H is independent of the parameter. Substitute the Hamiltonian (D.1) into the integral (D.9) and we can write the resulting expression in the form of the sum

$$\boxed{E = T + V + W} \tag{D.10}$$

where

$$T = \int \psi^*(\vec{r}_1, \vec{r}_2) \left(-\frac{\hbar^2}{2m}\right) \nabla_1^2\, \psi(\vec{r}_1, \vec{r}_2)\, \mathrm{d}^3r_1\, \mathrm{d}^3r_2 + \int \psi^*(\vec{r}_1, \vec{r}_2) \left(-\frac{\hbar^2}{2m}\right) \nabla_2^2\, \psi(\vec{r}_1, \vec{r}_2)\, \mathrm{d}^3r_1\, \mathrm{d}^3r_2 \tag{D.11}$$

$$V = \int \psi^*(\vec{r}_1, \vec{r}_2) \left(-\frac{Ze^2}{4\pi\varepsilon_0\, r_1}\right) \psi(\vec{r}_1, \vec{r}_2)\, \mathrm{d}^3r_1\, \mathrm{d}^3r_2 + \int \psi^*(\vec{r}_1, \vec{r}_2) \left(-\frac{Ze^2}{4\pi\varepsilon_0\, r_2}\right) \psi(\vec{r}_1, \vec{r}_2)\, \mathrm{d}^3r_1\, \mathrm{d}^3r_2 \tag{D.12}$$

$$W = \int \psi^*(\vec{r}_1, \vec{r}_2)\, \frac{e^2}{4\pi\varepsilon_0\, |\vec{r}_1 - \vec{r}_2|}\, \psi(\vec{r}_1, \vec{r}_2)\, \mathrm{d}^3r_1\, \mathrm{d}^3r_2 \tag{D.13}$$

D.3.1 Calculation of the Kinetic Energy (*T*)

$$
\begin{aligned}
T &= -\frac{\hbar^2}{2m}\int \phi^*(\vec{r}_1)\phi^*(\vec{r}_2)\,\nabla_1^2\,\phi(\vec{r}_1)\phi(\vec{r}_2)\;\mathrm{d}^3r_1\,\mathrm{d}^3r_2 - \frac{\hbar^2}{2m}\int \phi^*(\vec{r}_1)\phi^*(\vec{r}_2)\,\nabla_2^2\,\phi(\vec{r}_1)\phi(\vec{r}_2)\;\mathrm{d}^3r_1\,\mathrm{d}^3r_2 = \\
&= -\frac{\hbar^2}{2m}\int \mathrm{d}^3r_1\,\phi^*(\vec{r}_1)\nabla_1^2\,\phi(\vec{r}_1)\int \mathrm{d}^3r_2\;\phi^*(\vec{r}_2)\phi(\vec{r}_2) - \frac{\hbar^2}{2m}\int \mathrm{d}^3r_2\,\phi^*(\vec{r}_2)\nabla_2^2\,\phi(\vec{r}_2)\int \mathrm{d}^3r_1\;\phi^*(\vec{r}_1)\phi(\vec{r}_1)
\end{aligned}
$$

Since the orbitals $\phi(\vec{r})$ of the hydrogen-like ion are normalized to 1, i.e.

$$\int \phi^*(\vec{r})\phi(\vec{r})\;\mathrm{d}^3r = 1$$

the corresponding integrals do not need to be written further. The choice of particular symbols for integration variables does not matter, and therefore each of the two addends in the expression for T will be the same. So we get

$$T = \left(-\frac{\hbar^2}{2m}\right)2\int \phi^*(\vec{r})\,\nabla^2\,\phi(\vec{r})\,\mathrm{d}^3r \tag{D.14}$$

We now substitute for ϕ according to expression (D.7) and obtain

$$T = -\frac{\hbar^2}{2m}\,2\,\frac{\beta^3}{\pi a^3}\underbrace{\int e^{-\beta r/a}\,\nabla^2\,e^{-\beta r/a}\;\mathrm{d}^3r}_{\mathcal{I}_1} \equiv -\frac{\hbar^2}{2m}\,2\,\frac{\beta^3}{\pi a^3}\,\mathcal{I}_1 \tag{D.15}$$

We denoted the integral in this expression by $\mathcal{I}_1$. Calculate it in spherical coordionates. Recall that the Laplace operator in these coordinates acquires the form

$$\vec{\nabla}^2 = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + \frac{1}{r^2}\nabla^2_{\vartheta,\varphi} \tag{D.16}$$

where

$$\nabla^2_{\vartheta,\varphi} = \frac{1}{\sin\vartheta}\,\frac{\partial}{\partial\vartheta}\left(\sin\vartheta\frac{\partial}{\partial\vartheta}\right) + \frac{1}{\sin^2\vartheta}\,\frac{\partial^2}{\partial\varphi^2} \tag{D.17}$$

So, calculate

$$
\begin{aligned}
\mathcal{I}_1 &\equiv \int e^{-\beta r/a}\,\nabla^2\,e^{-\beta r/a}\;\mathrm{d}^3r \;=\; \int_0^\infty \mathrm{d}r\;r^2\int \mathrm{d}\Omega\;e^{-\beta r/a}\,\nabla^2\,e^{-\beta r/a} \;= \\
&= \int_0^\infty \mathrm{d}r\;r^2\int \mathrm{d}\Omega\;e^{-\beta r/a}\left[\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + \frac{1}{r^2}\nabla^2_{\vartheta,\varphi}\right]e^{-\beta r/a}
\end{aligned}
$$

The following holds:

$$\nabla^2_{\vartheta,\varphi}\,e^{-\beta r/a} = 0$$

The integration over the spatial angle Ω will be simple because nothing depends on the angles in the function to be integrated. We have

$$\int \mathrm{d}\Omega \equiv \int_0^\pi \mathrm{d}\vartheta\,\sin\vartheta\int_0^{2\pi}\mathrm{d}\varphi = 4\pi$$

and so, we can write

$$\mathcal{I}_1 = 4\pi\int_0^\infty \mathrm{d}r\;r^2 e^{-\beta r/a}\,\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\,\frac{\partial e^{-\beta r/a}}{\partial r}\right) = 4\pi\int_0^\infty \mathrm{d}r\;e^{-\beta r/a}\,\frac{\partial}{\partial r}\left(r^2\,\frac{\partial e^{-\beta r/a}}{\partial r}\right)$$

We will do this integral easily using the *per partes* (by parts) method. The results is

$$\mathcal{I}_1 = -\frac{\pi a}{\beta} \tag{D.18}$$

Therefore, the quantum-mechanical expectation value of the kinetic energy of the two electrons under study in the state ψ will be

$$\boxed{T = 2\,\frac{\hbar^2}{2m}\left(\frac{\beta}{a}\right)^2} = \beta^2 \quad \text{(in a. u.)} \tag{D.19}$$

We explicitly highlighted the factor of 2 to remind ourselves that the kinetic energy is a sum of the kinetic energies of the two electrons in the atom or ion under study. We see that Hartree atomic units [in which $m = e = \hbar = 1$, $\varepsilon_0 = 1/(4\pi)$] make it extremely easy to write some formulae. If in a specific problem the use of atomic units would significantly facilitate us, e.g. a derivation procedure, the atomic units ought to be used. We will soon get to such an opportunity.

D.3.2 Calculation of the Electrons-Nucleus Potential Energy (V)

$$\boxed{V = -2\,\frac{e^2}{4\pi\varepsilon_0}\,Z\,\frac{\beta}{a}} = -2Z\beta \quad \text{(in a. u.)} \tag{D.20}$$

D.3.3 Calculation of the Electron-Electron Potential Energy (W)

The most interesting contribution to the calculated energy E is given by expression (D.13). It expresses the energy of the electron-electron interaction. We substitute formula (D.8) for ψ and rewrite the value of W as follows:

$$W = \frac{e^2}{4\pi\varepsilon_0}\,\frac{1}{\pi^2}\left(\frac{\beta}{a}\right)^6 \int \exp\left[-\frac{2\beta(r_1+r_2)}{a}\right]\frac{1}{r_{12}}\,\mathrm{d}^3r_1\,\mathrm{d}^3r_2 \;\equiv\; \frac{e^2}{4\pi\varepsilon_0}\,\frac{1}{\pi^2}\left(\frac{\beta}{a}\right)^6 \mathcal{I}_3 \tag{D.21}$$

Thus, in this expression we have denoted the integral alone by

$$\mathcal{I}_3 = \int \exp\left[-\frac{2\beta(r_1+r_2)}{a}\right]\frac{1}{r_{12}}\,\mathrm{d}^3r_1\,\mathrm{d}^3r_2 \tag{D.22}$$

and we calculate it using spherical coordinates as follows:

$$\mathcal{I}_3 = \int_0^\infty \mathrm{d}r_1\, r_1^2 \exp\left(-\frac{2\beta}{a}r_1\right)\int \mathrm{d}\Omega_1 \int_0^\infty \mathrm{d}r_2\, r_2^2 \exp\left(-\frac{2\beta}{a}r_2\right)\int \frac{\mathrm{d}\Omega_2}{r_{12}} \tag{D.23}$$

Integration over the spatial angle Ω_2 can be elegantly managed using an analogy with electrostatics (**the trick of Prof. Peter Lichard, in which a uniformly charged spherical surface with radius r_2 is considered and the electrostatic potential at point r_1, which can be at any point in space, is calculated from it**). We obtain

$$\int \frac{\mathrm{d}\Omega_2}{r_{12}} = \begin{cases} \dfrac{4\pi}{r_1} & \text{pre } r_1 > r_2 \\[2ex] \dfrac{4\pi}{r_2} & \text{pre } r_1 \le r_2 \end{cases} \tag{D.24}$$

If we introduce the so-called Heaviside step function

$$\Theta(x) = \begin{cases} 1 & \text{for } x > 0 \\ 1/2 & \text{for } x = 0 \\ 0 & \text{for } x < 0 \end{cases} \tag{D.25}$$

then the result of the integration over the spatial angle Ω_2 can also be written as follows:

$$\int \frac{\mathrm{d}\Omega_2}{r_{12}} = \frac{4\pi}{r_1}\,\Theta(r_1 - r_2) + \frac{4\pi}{r_2}\,\Theta(r_2 - r_1) \tag{D.26}$$

We substitute this expression into formula (D.23) and consequently the integral break into the sum

$$\begin{aligned}\mathcal{I}_3 = &\int_0^\infty \mathrm{d}r_1\, r_1^2 \exp\left(-\frac{2\beta}{a} r_1\right) \int \mathrm{d}\Omega_1 \int_0^\infty \mathrm{d}r_2\, r_2^2 \exp\left(-\frac{2\beta}{a} r_2\right) \frac{4\pi}{r_1}\, \Theta(r_1 - r_2) \\ &+ \int_0^\infty \mathrm{d}r_1\, r_1^2 \exp\left(-\frac{2\beta}{a} r_1\right) \int \mathrm{d}\Omega_1 \int_0^\infty \mathrm{d}r_2\, r_2^2 \exp\left(-\frac{2\beta}{a} r_2\right) \frac{4\pi}{r_2}\, \Theta(r_2 - r_1)\end{aligned} \tag{D.27}$$

It turns out that both lines of this expression are the same. To see it, it is sufficient to interchange the ordering of the integrals in the second line. (This is possible because so far the integration bounds are the constant values or the infinities, not variables.)

$$\begin{aligned}\mathcal{I}_3 = &\int_0^\infty \mathrm{d}r_1\, r_1^2 \exp\left(-\frac{2\beta}{a} r_1\right) \int \mathrm{d}\Omega_1 \int_0^\infty \mathrm{d}r_2\, r_2^2 \exp\left(-\frac{2\beta}{a} r_2\right) \frac{4\pi}{r_1}\, \Theta(r_1 - r_2) \\ &+ \int_0^\infty \mathrm{d}r_2\, r_2^2 \exp\left(-\frac{2\beta}{a} r_2\right) \int \mathrm{d}\Omega_1 \int_0^\infty \mathrm{d}r_1\, r_1^2 \exp\left(-\frac{2\beta}{a} r_1\right) \frac{4\pi}{r_2}\, \Theta(r_2 - r_1) = \\ &= 4\pi \int_0^\infty \mathrm{d}r_1\, r_1^2 \exp\left(-\frac{2\beta}{a} r_1\right) \int_0^{r_1} \mathrm{d}r_2\, r_2^2 \exp\left(-\frac{2\beta}{a} r_2\right) \frac{4\pi}{r_1} \\ &+ 4\pi \int_0^\infty \mathrm{d}r_2\, r_2^2 \exp\left(-\frac{2\beta}{a} r_2\right) \int_0^{r_2} \mathrm{d}r_1\, r_1^2 \exp\left(-\frac{2\beta}{a} r_1\right) \frac{4\pi}{r_2}\end{aligned} \tag{D.28}$$

(In both lines, the integration over Ω_1 yields the value of 4π.) When we look at these two added expressions, we see that they differ only in that the indices 1 and 2 are interchanged. However, the notation used for integration variables cannot matter. Therefore, *both those lines are the same* and we can write

$$\begin{aligned}\mathcal{I}_3 &= 2\,.\,4\pi \int_0^\infty \mathrm{d}r_1\, r_1^2 \exp\left(-\frac{2\beta}{a} r_1\right) \int_0^{r_1} \mathrm{d}r_2\, r_2^2 \exp\left(-\frac{2\beta}{a} r_2\right) \frac{4\pi}{r_1} = \\ &= 2(4\pi)^2 \int_0^\infty \mathrm{d}r_1\, r_1 \exp\left(-\frac{2\beta}{a} r_1\right) \int_0^{r_1} \mathrm{d}r_2\, r_2^2 \exp\left(-\frac{2\beta}{a} r_2\right)\end{aligned} \tag{D.29}$$

The integrals that appear there can already be calculated by basic methods. First, the integral over r_2 must be calculated, because it has the value r_1 as its upper bound. The overall result is

$$\mathcal{I}_3 = 2(4\pi)^2\, \frac{5}{8} \left(\frac{a}{2\beta}\right)^5 \tag{D.30}$$

According to (D.21), the integrál W, which has the physical dimension of energy, is then equal to

$$\boxed{W = \frac{e^2}{4\pi\varepsilon_0} \frac{5}{8} \frac{\beta}{a}} = \frac{5}{8}\beta \quad \text{(in a. u.)} \tag{D.31}$$

D.4 The Energy for the Chosen Wave Function ψ (continuation)

The total energy E, as written down by formulae (D.9)) and (D.10), can now be expressed by adding contributions (D.19), (D.20) and (D.31) as follows:

$$E = 2\,\frac{\hbar^2}{2m} \left(\frac{\beta}{a}\right)^2 - 2\,\frac{e^2}{4\pi\varepsilon_0}\, Z\, \frac{\beta}{a} + \frac{e^2}{4\pi\varepsilon_0} \frac{5}{8} \frac{\beta}{a} \tag{D.32}$$

To get rid of a number of now irrelevant constants, we express it in Hartree atomic units $[m = e = \hbar = 1,$ $\varepsilon_0 = 1/(4\pi)\,]$. In these, the Bohr radius a gets equal to 1 [see (D.6)]. We obtain a much simpler expression

$$\boxed{E = \beta^2 - 2Z\beta + \frac{5}{8}\beta = E(\beta)} \tag{D.33}$$

The unit of energy in Hartree atomic units is 1 Hartree, which is

$$1\,\mathrm{Ha} \;=\; \frac{e^2}{4\pi\varepsilon_0}\,\frac{1}{a} \;\equiv\; \frac{e'^2}{a} \;=\; 27.2113834\,\mathrm{eV} = 4.35974380\,.\,10^{-18}\,\mathrm{J} \tag{D.34}$$

as can also be deduced from the energy formulae written above. Thus, an energy in atomic units can also be understood as a (dimensionless) value of the energy expressed relative to the value of 1 Ha. As already mentioned, the unit of distance in the atomic units is

$$1\,\mathrm{Bohr} = a \;\equiv a_{\mathrm{B}} = \frac{\hbar^2 4\pi\varepsilon_0}{me^2} = 0.5291772083\,.\,10^{-10}\,\mathrm{m} \tag{D.35}$$

D.5 Minimisation of the Energy by the Variational Method

The energy E expressed by formula (D.33) is an energy of the two electrons described by wave function (D.4) [see also (D.7) and (D.8)] and moving in the central field of the nucleus with the charge Ze. The wave function (D.4) itself corresponds in its form to two mutually independent non-interacting electrons, each of them would move in a central field of the nucleus with charge βe. Hence, the function ψ will certainly not be correct for the problem under study (helium ot its like ion), especially because it completely ignores the mutual repulsive interaction of the electrons and, in addition, it includes the parameter β, the value of which we did not even specify. Nevertheless, we can improve it as much as possible, or rather set it as best we can, i.e. optimise it. The only way to do it is to find the best possible parameter β. The criterion of optimality of the parameter β will be the energy (D.33), which we will try to get as low as possible, in accordance with the variational principle of quantum mechanics (Theorem 8). Mathematically, this means finding the minimum of the function $E(\beta)$. So, we employ derivatives:

$$\frac{\partial E}{\partial \beta} = 2\beta - 2Z + \frac{5}{8}$$

and set the condition

$$\frac{\partial E}{\partial \beta} = 0$$

Using it, we obtain the result for the optimal value of the parameter β:

$$\beta_{\mathrm{opt}} = Z - \frac{5}{16} \tag{D.36}$$

Thus, the mininal energy sought will be

$$\boxed{E_{\mathrm{min}} = E(\beta_{\mathrm{opt}}) = -\left(Z - \frac{5}{16}\right)^2} = -\beta_{\mathrm{opt}}^2 \tag{D.37}$$

This is our approximate result for the energy of the ground state of the helium atom ot its like ion. Let us now look at the cases of the individual proton numbers Z (table 1).

D.6 The Ionisation Energy of Helium and of the Like Ions

The first ionisation energy of a (neutral) helium atom is the minimum energy needed to pull one electron out of it. It is assumed that both helium and the resulting He$^+$ ion are in their ground states. An electron removed by the supply of ionisation energy moves away from the atom and stops (or has negligible kinetic energy). Since it is far from the nucleus after being torn out, it also has negligible potential energy. The first ionisation energy of the helium atom can therefore be calculated as follows:

$$E_1^{\mathrm{ion}}(\mathrm{He}) = E(\mathrm{He}^+) - E(\mathrm{He}) \tag{D.38}$$

He$^+$ is a hydrogen-like ion; thus we know its ground-state energy exactly: We will get it from formula (147) with $n = 1$, $Z = 2$. As for the helium atom itself, we have just determined its energy approximately by the variational method: formula (D.37) with $Z = 2$.

Table 1: The energies of the helium atom and its like ions obtained by the one-parameter variational method.

Z	značka	E_{min} (Ha)	E_{min} (eV)
1	H^-	-0.4727	-12.86
2	He	-2.8477	-77.49
3	Li^+	-7.2227	-196.53
4	Be^{2+}	-13.5977	-370.01
5	B^{3+}	-21.9727	-597.91
6	C^{4+}	-32.3477	-880.22

Similarly we define the ionisation energies of the helium/like ions H^-, Li^+, etc. These, however, will not be the 1st ionisation energies of neutral atoms, but of the ions. To však už nebudú 1. ionizačné energie neutrálnych atómov, ale iónov. For instance, using the formula

$$E^{\text{ion}}(\text{Li}^+) = E(\text{Li}^{2+}) - E(\text{Li}^+) \tag{D.39}$$

we calculate this energy for the lithium kation. Table 2 lists and compares these values also with experimental ones taken from [1]. All the energies in this table are in Hartrees. The negative value in the case of the anion H^- is non-physical. We determined it by the equation

Table 2: The ionisation energies of helium and its like ions, obtained by the one-parameter variational method. All the energies in this table are in the atomic units (Ha). The experimental energies have been taken from [1].

Z	symbol	experiment	the variational method
1	H^-	0.055	-0.0273
2	He	0.90331(4)	0.8477
3	Li^+	2.7798(5)	2.7227
4	Be^{2+}	–––	5.5977
5	B^{3+}	–––	9.4727
6	C^{4+}	14.407(4)	14.3477

$$E^{\text{ion}}(\text{H}^-) = E(\text{H}) - E(\text{H}^-) \tag{D.40}$$

The negative value indicates that the variational method that we used is completely incapable to calculate the ionisation energy of this anion. This is because the energy of the anion H^- is too high by this method (−0.4727 Ha). For the other ions (as well as helium itself), however, the simple variational method explained here gives surprisingly good results (given how simple it is). If we used a variational method with more parameters, we would get results closer to the experimental ones and the value of the ionisation energy for H^- would be positive. Physically, a negative ionization energy would mean that the ion H^- would be unstable. In fact, this ion exists and is extremely important for the opacity of the atmosphere of the Sun and similar stars (D. Chalonge, 1946). The H^- ion can exist in a stable way in its ground state only. (Its bound excited state does not exist.) One proton is able to keep at two electrons only in their ground state.

D.7 The Effect of Screening

For the ground state energy of the helium atom and similar ions, we derived the expression

$$E = -\frac{e'^2}{a}\left(Z - \frac{5}{16}\right)^2 \tag{D.41}$$

This expression can be broken down into the sum of two identical ones:

$$E = -\frac{1}{2}\frac{e'^2}{a}\left(Z - \frac{5}{16}\right)^2 - \frac{1}{2}\frac{e'^2}{a}\left(Z - \frac{5}{16}\right)^2 \tag{D.42}$$

Each of these two terms has a form such as the ground state energy of a hydrogen-like ion (i.e., the energy of a single-electron system); see expression (147), which can indeed be easily adapted to the form

$$E_1 = -\frac{1}{2}\frac{e'^2}{a} Z^2 \tag{D.43}$$

However, in the above expression (D.42) for the energy of the helium atom or a similar ion, the value of $Z - 5/16$ acts as if it were a proton number, not the value of Z. Therefore, we can introduce an effective proton number

$$Z_{\text{eff}} = Z - \frac{5}{16} \tag{D.44}$$

For helium, its value is $Z_{\text{eff}}(\text{He}) = 27/16$. The electrons thus shield or *screen* each other from the nucleus, and each of them moves like in a spherically symmetric field of one and the same effective nucleus (in the field that is the sum of the field of the real nucleus and the other electron). Such an interpretation is correct thanks to the approximation we used, namely that we have written the wave function in a factorised form separating the variables $\vec{r}_1$ and $\vec{r}_2$.

E The Hartree Method

The Task to Be Solved. Máme riešiť problém (261). V Hartreeho metóde budeme neznámu vlnovú funkciu Ψ hľadať v tvare Hartreeho súčinu. Pri výklade Hartreeho metódy sa obvykle ignoruje spin [1], čo spravíme aj my. Budeme teda riešiť týmto spôsobom zjednodušenú verziu úlohy (261). Vlnovú funkciu označíme a vyjadríme výrazom

$$\psi(\vec{r}_1, \ldots, \vec{r}_N) = \varphi_1(\vec{r}_1) \ldots \varphi_N(\vec{r}_N) \tag{E.1}$$

Je to tiež Hartreeho súčin, tentoraz závislý iba od priestorových súradníc. Hartreeho metóda je istou realizáciou variačnej metódy; pozri odsek 6.1. V zmysle tejto metódy potom funkciu (E.1) budeme považovať za pokusnú funkciu, na ktorú aplikujeme variačnú metódu. Ak by sme zabezpečili, že menovateľ zlomku (167) vo variačnej metóde by bol rovný 1, celkovú energiu sústavy by sme mohli hľadať minimalizáciou výrazu

$$\mathcal{G} = \int \psi^*(\vec{r}_1, \ldots, \vec{r}_N)\ \hat{H}\ \psi(\vec{r}_1, \ldots, \vec{r}_N)\, \mathrm{d}\tau \geq E_0 \tag{E.2}$$

kde $\mathrm{d}\tau \equiv \mathrm{d}^3 r_1 \ldots \mathrm{d}^3 r_N$. $\mathcal{G}$ vtedy predstavuje kvantovomechanickú strednú hodnotu energie sústavy nachádzajúcej sa v stave ψ. Jednotkovosť menovateľa v (167) znamená, že mnohočasticová funkcia ψ je normovaná na 1:

$$\int \psi^*(\vec{r}_1, \ldots, \vec{r}_N)\, \psi(\vec{r}_1, \ldots, \vec{r}_N)\, \mathrm{d}\tau = 1 \tag{E.3}$$

Keďže výraz $\mathcal{G}$ závisí od *funkcií* (máme na mysli tie φ_i), nazývame ho *funkcionál.* Normovanie ψ na 1 dosiahneme tým, že aj pre jednočasticové funkcie budeme požadovať, aby platilo

$$\int \varphi_i^*(\vec{r})\, \varphi_i(\vec{r})\, \mathrm{d}^3 r = 1\,, \qquad \forall i \tag{E.4}$$

Splnenie týchto normovacích podmienok zabezpečíme použitím Lagrangeových multiplikátorov. Preto definujeme rozšírený funkcionál

$$G = \mathcal{G} - \sum_{i=1}^{N} \lambda_i \left(\int \varphi_i^*(\vec{r})\, \varphi_i(\vec{r})\, \mathrm{d}^3 r - 1 \right) \tag{E.5}$$

kde λ_i sú spomínané Lagrangeove multiplikátory. Namiesto jednoduchšieho funkcionálu (E.2) teda budeme minimalizovať G. Jednočasticové funkcie φ_i vystupujúce v (E.1) sú neznáme a našou úlohou je nájsť ich tak, aby bola hodnota G čo najmenšia. Funkcie φ_i teda majú úlohu variačných parametrov.

Fyzikálne parametre problému, ktorý treba riešiť, sú definované Hamiltoniánom. Ten zoberme podobný ako (258), ale bez spinovej časti, čo je často veľmi dobré priblíženie. Hamiltonián teda teraz zapíšeme

$$\boxed{\hat{H} = \sum_{i=1}^{N} \hat{h}(i) + \frac{1}{2} \sum_{i,j=1}^{N} \hat{w}(i,j)} \tag{E.6}$$

kde

$$\boxed{\hat{h}(i) \equiv \hat{h}(\vec{r}_i) = -\frac{\hbar^2}{2m} \vec{\nabla}_i^2 + \hat{v}_{\text{ext}}(\vec{r}_i)} \tag{E.7}$$

a

$$\boxed{\hat{w}(i,j) = \begin{cases} \dfrac{e^2}{4\pi\varepsilon_0} \dfrac{1}{|\vec{r}_i - \vec{r}_j|}\,, & i \neq j \\ 0\,, & i = j \end{cases}} \tag{E.8}$$

The Functional Representing Energy for Given Wave Function (E.1). Celý funkcionál (E.5) je praktické rozpísať si a následne zjednodušiť takto:

$$G \;=\; \mathcal{G} + \mathcal{L} \;=\; \mathcal{G}^{(1)} + \mathcal{G}^{(2)} + \mathcal{L} \tag{E.9}$$

kde

$$\mathcal{G}^{(1)} \equiv \int \psi^*(\vec{r}_1, \ldots, \vec{r}_N) \left[\sum_{i=1}^{N} \hat{h}(i) \right] \psi(\vec{r}_1, \ldots, \vec{r}_N)\, \mathrm{d}\tau = \sum_{i=1}^{N} \int \varphi_i^*\, \hat{h}(i)\, \varphi_i\, \mathrm{d}^3 r_i = \tag{E.10}$$

$$= \big(\text{na označení integ. prem. nezáleží}\big) = \sum_{i=1}^{N} \int \varphi_i^*(\vec{r})\, \hat{h}(\vec{r})\, \varphi_i(\vec{r})\, \mathrm{d}^3 r \tag{E.11}$$

$$\mathcal{G}^{(2)} \equiv \int \psi^*(\vec{r}_1, \ldots, \vec{r}_N) \left[\frac{1}{2} \sum_{i,j=1}^{N} \hat{w}(i,j) \right] \psi(\vec{r}_1, \ldots, \vec{r}_N)\, \mathrm{d}\tau = \tag{E.12}$$

$$= \frac{1}{2} \sum_{i,j=1}^{N} \int \varphi_i^*(\vec{r}_i)\, \varphi_j^*(\vec{r}_j)\, \hat{w}(i,j)\, \varphi_i(\vec{r}_i)\, \varphi_j(\vec{r}_j)\; \mathrm{d}^3 r_i\, \mathrm{d}^3 r_j \tag{E.13}$$

$$\mathcal{L} \equiv -\sum_{i=1}^{N} \lambda_i \left[\int \varphi_i^*(\vec{r})\, \varphi_i(\vec{r})\, \mathrm{d}^3 r - 1 \right] \tag{E.14}$$

Minimisation of the Functional (and of the Energy). Chceme zistiť, pri akých funkciách φ_i bude funkcionál G minimálny. Ide o niečo analogické ku hľadaniu minima funkcie, kedy sa funkcia derivuje. Tu však máme hľadať minimum funkcionálu. Namiesto jednoduchého derivovania budeme funkcionál G varírovať, čo znamená, že skúsime, ako sa zmení pri malej zmene funkcií φ_i, od ktorých závisí. Uvažujme teda takúto variáciu funkcií φ_i:

$$\boxed{\varphi_i \longrightarrow \varphi_i + \delta\varphi_i} \tag{E.15}$$

Potom sa funkcionál zmení takto:

$$\boxed{G[\varphi] \quad \longrightarrow \quad G[\varphi + \delta\varphi] = G[\varphi] + \delta G} \tag{E.16}$$

a obdobne sa to dá písať aj pre jeho jednotlivé zložky $\mathcal{G}^{(1)}$, $\mathcal{G}^{(2)}$ a $\mathcal{L}$. Pre súčet jednočasticových integrálov v sume vyššie teda máme

$$\begin{aligned}\mathcal{G}^{(1)}[\varphi] \quad \longrightarrow \quad \mathcal{G}^{(1)}[\varphi+\delta\varphi] &= \sum_{i=1}^{N} \int (\varphi_i+\delta\varphi_i)^* \, \hat{h}(i) \, (\varphi_i+\delta\varphi_i) \, \mathrm{d}^3 r_i \; = \\ &= \mathcal{G}^{(1)}[\varphi] + \underbrace{\sum_{i=1}^{N} \int \delta\varphi_i^* \, \hat{h}(i) \, \varphi_i \, \mathrm{d}^3 r_i + \sum_{i=1}^{N} \int [\hat{h}(i) \, \varphi_i]^* \delta\varphi_i \, \mathrm{d}^3 r_i}_{\delta\mathcal{G}^{(1)}} \; + \\ &+ \text{členy 2. rádu v } \delta\varphi_k \, , \text{ ktoré sú zanedbateľné}\end{aligned} \tag{E.17}$$

Všimnime si, že druhý člen v $\delta\mathcal{G}^{(1)}$ je komplexne združený k prvému.

Aj pri počítaní s dvojčasticovými integrálmi budeme miestami kvôli stručnosti vynechávať písanie argumentov funkcií φ; ak sú vynechané, tak platí $\varphi_i \equiv \varphi_i(i) \equiv \varphi_i(\vec{r}_i)$, $\varphi_j \equiv \varphi_j(j) \equiv \varphi_j(\vec{r}_j)$. Pre variáciu sumy dvojčasticových integrálov (E.13) dostávame postupom podobným než vyššie, len zložitejším, toto [pričom využijeme, že $\hat{w}(i,j) = \hat{w}(j,i)$ a že sumačné indexy môžeme ľubovoľne premenovať, aj vymeniť $(i \leftrightarrow j)$ medzi sebou]:

$$\delta\mathcal{G}^{(2)} = \sum_{i,j=1}^{N} \int \delta\varphi_i^* \, \varphi_j^* \, \hat{w}(i,j) \, \varphi_i \, \varphi_j \, \mathrm{d}^3 r_i \, \mathrm{d}^3 r_j + \sum_{i,j=1}^{N} \int \varphi_i^* \, \varphi_j^* \, \hat{w}(i,j) \, \delta\varphi_i \, \varphi_j \, \mathrm{d}^3 r_i \, \mathrm{d}^3 r_j \tag{E.18}$$

Aj tu je druhý člen komplexne združený k prvému. Aby sme spočítali aj variáciu funkcionálu G, nielen $\mathcal{G}$, zostáva ešte spočítať variáciu člena s Lagrangeovými multiplikátormi, pozri (E.14). Tá sa počíta ľahko a je

$$\delta\mathcal{L} = -\sum_{i=1}^{N} \lambda_i \left(\int \delta\varphi_i^* \, \varphi_i \, \mathrm{d}^3 r + \int \varphi_i^* \, \delta\varphi_i \, \mathrm{d}^3 r \right) \tag{E.19}$$

Teraz už vieme napísať, čomu sa rovná variácia celého funkcionálu G, pozri (E.5), (E.9). Potrebujeme na to zozbierať výsledky (E.17), (E.18) a (E.19). Dostávame

$$\delta G \equiv G[\varphi+\delta\varphi] - G[\varphi] = \delta\mathcal{G}^{(1)} + \delta\mathcal{G}^{(2)} + \delta\mathcal{L} \tag{E.20}$$

a teda

$$\delta G = \sum_{i=1}^{N} \int \mathrm{d}^3 r_i \, \delta\varphi_i^* \left(\hat{h}(i) + \sum_{j=1}^{N} \int \mathrm{d}^3 r_j \, \varphi_j^* \, \hat{w}(i,j) \, \varphi_j \; - \lambda_i \right) \varphi_i \; + \; \text{k.z.} \tag{E.21}$$

kde k.z. označuje členy komplexne združené s predošlými.

Ako sme povedali už skôr, snažíme sa hľadať, pri akých funkciách φ_i je funkcionál G minimálny. Tak ako pri funkcii je v okolí jej extrému nulová prvá derivácia, čiže v prvom ráde nulová zmena, tak pri funkcionáli je v okolí jeho extrému nulová variácia. Preto kvôli nájdeniu minimalizujúcich funkcií φ_i požadujeme

$$\delta G = 0 \tag{E.22}$$

Aby toto bolo splnené pre ľubovoľné variácie $\delta\varphi_i$, musí byť výraz $(\ldots\ldots)\varphi_i$ vo formule (E.21) nulový.[65] Musia teda platiť rovnice

$$\left[\hat{h}(i) + \sum_{j=1}^{N} \int \mathrm{d}^3 r_j \, \varphi_j^*(j) \, \hat{w}(i,j) \, \varphi_j(j) \right] \varphi_i(i) = \lambda_i \varphi_i(i) \tag{E.23}$$

Pripomeňme, že $\hat{w}(i,i) \equiv 0$, a teda členy s $j = i$ v týchto rovniciach vypadnú. „k.z.“ v rovnici (E.21) nám dá len komplexne združenú rovnicu ku práve napísanej, teda žiadnu novú rovnicu. Na označení integračnej premennej $\vec{r}_j$ v rovniciach (E.23) nemôže záležať. Ani písanie indexu i vo vonkajšej premennej $\vec{r}_i$ teraz už nie je nutné. Rovnice (E.23) sa preto dajú písať (trochu podrobnejšie) aj takto:

$$\boxed{\left[\hat{h}(\vec{r}) + \sum_{\substack{j=1 \\ j\neq i}}^{N} \frac{e^2}{4\pi\varepsilon_0} \int \mathrm{d}^3 r' \, \varphi_j^*(\vec{r}\,') \, \frac{1}{|\vec{r}-\vec{r}\,'|} \, \varphi_j(\vec{r}\,') \right] \varphi_i(\vec{r}) = \lambda_i \varphi_i(\vec{r})} \qquad i \in \{1, 2, \ldots, N\} \tag{E.24}$$

---

[65] Samotný obsah tých zátvoriek nemá číselnú hodnotu, je to len operátor. Preto v tom nulovom výraze musí byť sprava aj φ_i.

Rovnice (E.24) predstavujú sústavu N integrálno-diferenciálnych rovníc pre neznáme funkcie φ_i. Vyriešením týchto rovníc teda nájdeme funkcie, ktoré extremalizujú (zvyčajne minimalizujú) funkcionál G. Hodnota tohoto funkcionálu v takom prípade je približnou vlastnou energiou základného stavu sústavy. Sústava rovníc (E.24) [tak isto aj (E.23)] sa nazýva ***Hartreeho rovnice*** (HR). Jednotlivé rovnice tejto sústavy svojou formou pripomínajú bezčasovú Schrödingerovu rovnicu.

The physical meaning of the sum in the Hartree equations. Na jeho pochopenie stačí najprv v i-tej HR skúmať jeden člen (t. j. pre jedno j rôzne od i). Hodnota výrazu $\varphi_i^*(\vec r\,')\varphi_i(\vec r\,')$ je hustota pravdepodobnosti výskytu elektrónu s vlnovou funkciou $\varphi_j(\vec r\,')$ v mieste $\vec r\,'$. Výraz $(-e)|\varphi_j(\vec r\,')|^2$ je potom hustota elektrického náboja vytváraná takým elektrónom; je to hustota v kvantovo-mechanickom zmysle priemernej hodnoty v danom bode priestoru. Výraz

$$\int \frac{1}{4\pi\varepsilon_0}\,\frac{(-e)|\varphi_j(\vec r\,')|^2}{|\vec r-\vec r\,'|}\,\mathrm{d}^3r' \;\overset{\text{def}}{=}\; u_j(\vec r) \tag{E.25}$$

je priemerný elektrostatický potenciál v mieste $\vec r$ vytváraný elektrónom s vlnovou funkciou $\varphi_j(\vec r\,')$. Preto

$$\sum_{\substack{j=1\\ j\neq i}}^{N} u_j(\vec r) \;\overset{\text{def}}{=}\; U_i^{\text{Hartree}}(\vec r) \tag{E.26}$$

je ustrednený elektrostatický potenciál, ktorý vytvárajú v mieste $\vec r$ všetky elektróny okrem i-teho (t. j. toho, ktorý obsadzuje orbitál φ_i). Voláme ho aj *Hartreeho potenciál.* Celá suma cez j v HR teda je

$$\boxed{\sum_{\substack{j=1\\ j\neq i}}^{N} \frac{e^2}{4\pi\varepsilon_0}\int \mathrm{d}^3r'\;\varphi_j^*(\vec r\,')\,\frac{1}{|\vec r-\vec r\,'|}\,\varphi_j(\vec r\,') \;=\; -e\,U_i^{\text{Hartree}}(\vec r) \;=\; V_i^{\text{Hartree}}(\vec r)} \tag{E.27}$$

a má význam potenciálnej energie i-teho elektrónu v ustrednenom poli všetkých ostatných elektrónov. Je *priestorovo závislá* (preto sa nazýva potenciálna) a často ju preto tiež nazývajú Hartreeho potenciál [ale treba mať na pamäti, že v SI sústave majú potenciál a potenciálna energia odlišné jednotky a teda nie sú to totožné veličiny, i keď sa líšia len o triviálny násobok $(-e)$].

V tejto súvislosti si ešte všimnime sumu $\mathcal{G}^{(2)}$ elektrón-elektrónových odpudivých energií vyjadrenú príspevkom (E.13 Na základe vyššie zavedenej Hartreeho potenciálnej energie sa dá zapísať

$$\mathcal{G}^{(2)} = \frac{1}{2}\sum_{i=1}^{N}\int \varphi_i^*(\vec r)\,V_i^{\text{Hartree}}(\vec r)\,\varphi_i(\vec r)\,\mathrm{d}^3r \tag{E.28}$$

Orbitals in the Hartree Equations. The Effective Orbitals-Dependent Hamiltonian. HR (E.24) sa teda dajú kompaktne zapísať

$$\underbrace{\left[\hat h(\vec r) + V_i^{\text{Hartree}}(\vec r)\right]}_{\hat h_i^{\text{eff}}}\varphi_i(\vec r) = \lambda_i\varphi_i(\vec r) \tag{E.29}$$

Veľmi pripomínajú sústavu navzájom nezávislých rovníc (263) z motivačnej časti. Je tu však jedna komplikácia: efektívny jednočasticový potenciál v (E.29) závisí od orbitálov φ_j, čo sú neznáme funkcie. Tieto neznáme funkcie v HR vystupujú v kubickej forme. HR sú preto nelineárne a nie sú až tak jednoduché, ako sme si to na začiatku predstavovali v motivačnej časti.

Jednotlivé orbitály φ_i sú vlastnými funkciami navzájom odlišných efektívnych hamiltoniánov (ktoré sa preto tiež musia indexovať). Vyriešením HR tak dostaneme orbitály, ktoré nie sú navzájom ortogonálne. Podmienku ortogonálnosti sme ani nikde nepoužili. Naozaj: pri minimalizácii sme len naložili podmienku (E.4), že orbitály majú byť normované na 1.

Solution of the Hartree equations. HR predstavujú sústavu N integrálno-diferenciálnych rovníc. Rieši sa metódou postupných iterácií: na začiatku si zvolíme nejaké štartovacie funkcie

$$\varphi_1^{(0)}, \varphi_2^{(0)}, \dots, \varphi_N^{(0)} \tag{E.30}$$

Napr. ak riešime HR pre atóm, tak za $\varphi_i^{(0)}$ môžeme zvoliť presne známe vlastné funkcie pre vodíku podobný ión. Z týchto štartovacích funkcií určíme začiatočnú hodnotu Hartreeho potenciálu $U_i^{(0)}(\vec{r})$ (ktorá je určite ešte veľmi nesprávna). Z $U_i^{(0)}(\vec{r})$ potom riešením HR dostaneme už spresnené (ale stále veľmi hrubé) jednočasticové funkcie

$$\varphi_1^{(1)}, \varphi_2^{(1)}, \dots, \varphi_N^{(1)} \tag{E.31}$$

a aj prvý odhad Lagrangeových multiplikátorov $\lambda_i^{(1)}$. Tým máme ukončenú prvú iteráciu. A tak ďalej iterujeme, až raz skončíme, a to napr. vtedy, keď rozdiel medzi výstupmi po sebe idúcich iterácií bude zanedbateľný. Vtedy už budú orbitály φ_i konzistentné s Hartreeho potenciálom. Výsledné elektrostatické pole od uvažovaných elektrónov nazývame ***self-konzistentné pole***; pojem samosúhlasné pole sa používa menej často.

O význame vlastných hodnôt a jednočasticových funkcií sme si povedali pri štúdiu Hartreeho-Fockovej metódy. Samotná Hartreeho metóda sa v praxi používa zriedka, lebo nerešpektuje antisymetriu vlnovej funkcie.

F Functionals

F.1 An Intuitive Explanation of Functionals

Taylor expansion of a function of one variable:

$$f(x+\Delta x) = f(x) + \frac{1}{1!}\left.\frac{\mathrm{d}f}{\mathrm{d}x}\right|_x \Delta x + \frac{1}{2!}\left.\frac{\mathrm{d}^2 f}{\mathrm{d}x^2}\right|_x (\Delta x)^2 + \frac{1}{3!}\left.\frac{\mathrm{d}^3 f}{\mathrm{d}x^3}\right|_x (\Delta x)^3 + \dots \tag{F.1}$$

Taylor expansion of a function of m variables:

$$\begin{aligned} & f(x_1+\Delta x_1,\ x_2+\Delta x_2,\ \dots,\ x_n+\Delta x_m) = \\ & f(x_1, x_2, \dots, x_m) + \frac{1}{1!}\sum_{i=1}^{m}\left.\frac{\partial f}{\partial x_i}\right|_x \Delta x_i + \frac{1}{2!}\sum_{i=1}^{m}\sum_{j=1}^{m}\left.\frac{\partial^2 f}{\partial x_i \partial x_j}\right|_x (\Delta x_i)\,(\Delta x_j) + \dots \end{aligned} \tag{F.2}$$

The vertical lines $|_x$ says us: do the derivative and then evaluate it at x. For the functions on n variables, we used the shortcut

$$x \equiv x_1, x_2, \dots, x_m \tag{F.3}$$

A Functional. It is a mathematical form that depends on some *function*, i.e. not on an elementary variable or variables like x_1, ..., x_m. A nice examples is the Thomas-Fermi (TF) kinetic-energy functional (410):

$$T_{\mathrm{TF}}[n] = C_{\mathrm{F}} \int n^{5/3}(\vec{r})\ \mathrm{d}^3 r \tag{F.4}$$

It depends on the electron density $n(\vec{r})$, which itself is a function. Let us further use the notation $F[n]$ for a functional depending of a function $n(\vec{r})$ (which need not necessarily be a density). The function $n(\vec{r})$ is defined on some spatial domain $\mathcal{D}$ (which can be either finite such as a cubic box or a sphere) or infinite. In any such case, the domain contain an infinite number of spatial points that typically form a continuum and we will consider this type of damains. We can, however, *discretise* the spatial domain, i.e. to divide it between some finite number of grid points. Such a procedure is often being done for a purpose like numerical integration (quadrature) and, certainly, we would find several other examples from numerical mathematic. Here, however, we use the discretisation of space for the purpose of a theoretical analysis. Let the grid points be

$$\vec{r}_1, \vec{r}_2, \dots, \vec{r}_m \tag{F.5}$$

In such a discrete representation, the functional $F[n]$ becomes a function of m spatial variables:

$$F(n(\vec{r}_1), n(\vec{r}_2), \ldots, n(\vec{r}_m)) \tag{F.6}$$

To make notation more compact we introduce

$$n_i \equiv n(\vec{r}_i) \tag{F.7}$$

and then the liist of the variables on which the function F depends, is $n_1, n_2, \ldots, n_m$:

$$F = F(n_1, n_2, \ldots, n_m) \tag{F.8}$$

Let us now consider that the function $n(\vec{r})$ is slightly modified by an amount $\delta n(\vec{r})$:

$$n(\vec{r}) \longrightarrow n(\vec{r}) + \delta n(\vec{r}) \tag{F.9}$$

How the functional $F[n]$ changes upon such variation of $n(\vec{r})$? Formally, we express the variation $\delta F[n]$ of the functional $F[n]$ as follows:

$$F[n] \longrightarrow F[n + \delta n] = F[n] + \delta F[n] \tag{F.10}$$

The last equation defines what is meant by a *variation* of a functional, $\delta F[n]$; the term variation is used for functionals, not for functions. Since we have discretised the functional [converted it to the function (F.8)] and since the quantities $\delta n(\vec{r})$ are small enough, we can now take advantage of the Taylor expansion (F.2) to answer the above question by calculating $F[n + \delta n]$ in the discrete representation:

$$\begin{aligned} F(n_1 + \delta n_1, \ldots, n_m + \delta n_m) &= F(n_1, \ldots, n_m) + \\ + \sum_{i=1}^{m} \left.\frac{\partial F}{\partial n_i}\right|_n \delta n_i &+ \frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{m} \left.\frac{\partial^2 F}{\partial n_i \partial n_j}\right|_n \delta n_i \, \delta n_j + \ldots \end{aligned} \tag{F.11}$$

Because the grid can be very dense, we can replace the summations over the grid points by integrals over the spatial domain $\mathcal{D}$:

$$\begin{aligned} F(n_1 + \delta n_1, \ldots, n_m + \delta n_m) &= F(n_1, \ldots, n_m) + \\ &+ \int_{\mathcal{D}} \left.\frac{\partial F}{\partial n(\vec{r})}\right|_n \delta n(\vec{r}) \, \mathrm{d}^3 r + \\ &+ \frac{1}{2} \int_{\mathcal{D}} \mathrm{d}^3 r \int_{\mathcal{D}} \mathrm{d}^3 r' \left.\frac{\partial^2 F}{\partial n(\vec{r}) \partial n(\vec{r}\,')}\right|_n \delta n(\vec{r}) \, \delta n(\vec{r}\,') \; + \; \ldots \end{aligned} \tag{F.12}$$

Thus, after the short trip to common functions, we are back to the world of functionals; the last formula in the brief language of functionals is rewritten as

$$\begin{aligned} F[n + \delta n] &= F[n] + \int_{\mathcal{D}} \mathrm{d}^3 r \left.\frac{\delta F}{\delta n(\vec{r})}\right|_n \delta n(\vec{r}) + \\ &+ \frac{1}{2} \int_{\mathcal{D}} \mathrm{d}^3 r \int_{\mathcal{D}} \mathrm{d}^3 r' \left.\frac{\delta^2 F}{\delta n(\vec{r}) \delta n(\vec{r}\,')}\right|_n \delta n(\vec{r}) \, \delta n(\vec{r}\,') \; + \; \ldots \end{aligned} \tag{F.13}$$

The expressions

$$\frac{\delta F}{\delta n(\vec{r})}, \quad \frac{\delta^2 F}{\delta n(\vec{r}) \delta n(\vec{r}\,')}$$

are the first and the second functional derivatives of the given functional, $F[n]$, respectively. In the above exansions, they are evaluate at the density $n(\vec{r})$ [not at $n(\vec{r}) + \delta n(\vec{r})$]. By comparing (F.13) to (F.12) we immediatelly see the principal meaning of the concept of a functional derivative.

By definition, a variation of the functional is defined by [see (F.10)]

$$\delta F[n] = F[n + \delta n] - F[n] \tag{F.14}$$

Now, if we are searching for a local extremum (a minimum or maximum) of the functional, the necessary condition expressed in the discretised form is

$$\left.\frac{\partial F}{\partial n_i}\right|_{\text{ext}} = 0\,, \quad \forall i \in \{1, 2, \ldots, m\} \tag{F.15}$$

We see that the functional form of this condition is

$$\boxed{\left.\frac{\delta F}{\delta n(\vec{r})}\right|_{n_0} = 0} \tag{F.16}$$

(the necessary condition of a local extremum of the functional). Very often, this equation is sufficient to use do determine a minimum or maximum. Its solution is the extermising function $n_0(\vec{r})$, i.e. the one at which the functional takes its minimum or maximum (may not be a global one).

If δn is very small in (F.13), the first two terms then become sufficient to keep and we can then express variation (F.14) around some arbitrary chosen function $n(\vec{r})$ as

$$\delta F[n] = \int_{\mathcal{D}} \mathrm{d}^3 r \; \left.\frac{\delta F}{\delta n(\vec{r})}\right|_{n} \delta n(\vec{r})\,, \quad \text{for } \delta n(\vec{r}) \to 0 \tag{F.17}$$

If the chosen function $n(\vec{r})$ is $n_0(\vec{r})$, then the variation $\delta F[n]$ vanishes. Hence, we derived the frequently used formula of the *variational calculus*:

$$\boxed{\delta F[n] = 0} \quad \text{at a local minimum or maximum} \tag{F.18}$$

It is just as when differential of a common function like $f(x)$ of eq. (F.1) vanishes at any extremal point of the function.

F.2 On the Kohn-Sham Mapping (or Ansatz)

The Hohenberg-Kohn total-energy functional is given by (453):

$$E_v[n] = T[n] + W[n] + \int n(\vec{r}) v_{\text{ext}}(\vec{r})\; \mathrm{d}^3 r \tag{F.19}$$

In this expression, the forms of $T[n]$ and $W[n]$ are unknown; hence such a functional can not have any direct practical use. According to Kohn and Sham, we can, however, introduce the auxiliary (or reference) system of non-interacting electrons such as to provide the same ground-state density. A wave function of such a system has the form of a Slater determinant, exactly as in the Hartree-Fock theory. (So, both in the HF theory and in the DFT, total wave functions are some Slater determinants.)

Denote the non-interacting kinetic energy by symbol $T_s[n]$ [see (456)]. We now can add and substract $T_s[n]$ in (F.19) and regroup the terms:

$$E_v[n] = T_s[n] + \int n(\vec{r}) v_{\text{ext}}(\vec{r})\; \mathrm{d}^3 r + W[n] + T[n] - T_s[n] \tag{F.20}$$

$T[n] - T_s[n]$ is some difficult part of the kinetic energy. We also know that the electron-electron interation energy, $W[n]$, contains an “easy” component (that is, easy to express using the density). It is the Hartree energy, obviosly a relatively big component of $W[n]$; see (301) for its first occurence in our course, then (413) of the Thomas-Fermi theory, and finally (456) of the Kohn-Sham theory. So, decompose also $W[n]$:

$$W[n] = E_{\text{Hartree}}[n] + W_{\text{rest}} \tag{F.21}$$

Thus, the total-energy functional (F.20) can be expressed as

$$\begin{aligned} E_v[n] &= T_s[n] + \int n(\vec{r}) v_{\text{ext}}(\vec{r})\; \mathrm{d}^3 r + E_{\text{Hartree}}[n] + \underbrace{T[n] - T_s[n] + W_{\text{rest}}}_{E_{\text{xc}}[n]} = \\ &= E_{\text{KS}}[n] \end{aligned} \tag{F.22}$$

(It is the same expression as (456). And it has to be understood as the definition of $E_{\text{xc}}[n]$.) So, the difficult-to-express but, fortunatelly, relatively small part was denoted as $E_{\text{xc}}[n]$. It is called ***exchange-correlation energy***. We know from the HF theory that there must be the exchange-energy contribution to the total energy. Since this contributions is neither in $T_s[n]$ nor in the interaction energy with the external field, it then obviously must be a part of $T[n] - T_s[n] + W_{\text{rest}} \equiv E_{\text{xc}}[n]$. A similar consideration holds for the correlation energy.

Now, if we want to find a local minimum of the Kohn-Sham functional, (F.22), we have to keep the correct number of the electrons, N, in the minimisation. We do in using a Lagrange multiplier, exactly as we have done it in the TF theory, see (415). Therefore, again an augmented functional,

$$\Omega_{\text{KS}}[n] = E_{\text{KS}}[n] - \mu \left[\int n(\vec{r})\, \mathrm{d}^3 r - N \right] \tag{F.23}$$

is to be minimised. We do the minimisation by requesting that at the minimising density $n_0(\vec{r})$,

$$\delta\Omega_{\text{KS}}[n] = 0 \tag{F.24}$$

for very small variations $\delta n(\vec{r})$ around the minimising density. Alternatively, we can set the condition (see section F.1)

$$\left. \frac{\delta\Omega_{\text{KS}}}{\delta n(\vec{r})} \right|_{n_0} = 0 \tag{F.25}$$

This yields the Kohn-Sham equation for the ground-state density n_0 which is the density that minimises the functional $\Omega_{\text{KS}}[n]$, also (and importantly) the Kohn-Sham functional (F.22) [and also the orginal functional (F.19) of Hohenberg and Kohn] as that one is equal to the KS functional by definition. Note also, that people often omit the term with N in (F.23) for it is a constant only and has no effect in the serach for the minimum.

In practical calcuations, we are not able to find the exact functional $E_{\text{xc}}[n]$. Therefore, although the exact equality of the HK and KS functionals holds in theory, we do not meet it in practice.

References

[1] Peter Lichard, lectures from the acad. year 1991/1992, Faculty of Mathematics and Physics of the Comenius University in Bratislava. Poossibly also scripts *Kvantová mechanika* (Univerzita Komenského, Bratislava 1984).

[2] Ján Pišút, Ladislav Gomolčák, Vladimír Černý, *Úvod do kvantovej mechaniky*, 2^{nd} edition (Alfa, Bratislava 1983).

[3] Albert Messiah, *Quantum Mechanics*, Two Volumes Bound as One (Dover Publications, inc., Mineola 1999).

[4] Claude Cohen-Tannoudji, Bernard Diu, Franck Laloë, *Quantum Mechanics, Vol. 1.*, 2^{nd} edition (Wiley-VCH, Weinheim 2019).

[5] `https://en.wikipedia.org/wiki/Spherical_harmonics` *Spherical harmonics.*

[6] George B. Arfken, *Mathematical methods for physists*, 3^{rd} edition (Academic press, inc., Orlando, 1985).

[7] Atilla Szabo, Neil S. Ostlund, *Modern quantum chemistry* (Dover publications, inc., Mineola, 1996).

[8] Frank L. Pilar, *Elementary quantum chemistry*, 2^{nd} edition (McGraw-Hill Pub. 1990, New York; Dover 2001, Mineola).

[9] William H. Press, Saul A. Teukolsky, William T. Vetterling, Brian P. Flannery, *Numerical recipes*, 3^{rd} edition, (Cambridge University Press, Cambridge 2007).

This is the latest version of the „Numerical recipes". It makes effor to promote object-oriented programming in `C++` and it includes two new chapters. The topics we deal with are presented in most cases quite similarly in this book as in the older edition *Numerical recipes in C* [10].

[10] William H. Press, Saul A. Teukolsky, William T. Vetterling, Brian P. Flannery, *Numerical recipes in C, Second Edition*, (Cambridge University Press, Cambridge 1992);
the full text is freely available at `http://s3.amazonaws.com/nrbook.com/book_C210.html` .

There are older and newer editions of the „Numerical recipes"; see `http://numerical.recipes` .

At some topics it is, however, better to read the theory from the most up-to-date edition [9].

[11] Robert G. Parr, Weitao Yang, *Density Functional Theory of Atoms and Molecules*, (Oxford University Press, New York 1989).

[12] Richard M. Martin, *Electronic Structure*, (Cambridge University Press, Cambridge 2004).

[13] P. Hohenberg, W. Kohn, *Inhomogeneous Electron Gas*, Physical Review **136**, B864 (1964).

[14] W. Kohn, L.J. Sham, *Self-Consistent Equations Including Exchange and Correlation Effects*, Physical Review **140**, A1133 (1965).

[15] C. Fiolhais, F. Nogueira, M. Marques (editori), LECTURE NOTES IN PHYSICS, *A Primer in Density Functional Theory*, (Springer-Verlag Berlin Heidelberg 2003).

Contents

---

[66] In English-written literature, the term *many-body* wave function is sometime being used.