
Remarks on Wave Function Theory and Methods

4

Dariusz Kędziera and Anna Kaczmarek-Kedziera

Contents

Introduction: What and Why?	124
Quantum Mechanics for Dummies	125
On the Way to Quantum Chemistry	133
Variational Principle: An Indicator	136
Perturbation Calculus: The Art of Estimation	137
One-Electron Approximation: Describe One and Say Something About All	142
Hartree–Fock Method: It Is Not That Sophisticated	147
Møller–Plesset Perturbation Theory: HF Is Just the Beginning	152
Beyond the HF Wave Function	159
Coupled Cluster Approximation: The Operator Strikes Back	165
Conclusions	170
Bibliography	171

Abstract

Methods of computational chemistry seem to often be simply a melange of undecipherable acronyms. Frequently, the ability to characterize methods with respect to their quality and applied approximations or to ascribe the proper methodology to the physicochemical property of interest is sufficient to perform research. However, it is worth knowing the fundamental ideas underlying the computational techniques so that one may exploit the approximations intention-

D. Kędziera (✉)

Department of Chemistry and Photochemistry of Polymers, Faculty of Chemistry, Nicolaus Copernicus University, Toruń, Poland
e-mail: teodar@chem.umk.pl

A. Kaczmarek-Kedziera

Faculty of Chemistry, Nicolaus Copernicus University, Toruń, Poland
e-mail: teoadk@chem.umk.pl

ally and efficiently. This chapter is an introduction to quantum chemistry methods based on the wave function search in one-electron approximation.

Introduction: What and Why?

Quantum chemistry is a branch of science originating from quantum mechanics that focuses on investigations of chemical systems. The mathematical roots of quantum chemistry allow it to be treated as a methodology for solving an eigenvalue equation for operators or – even simpler – finding solutions for some differential equations. We cannot totally escape this way of thinking, since this is how things really are. However, a chemist will comprehend quantum chemistry more as a helping tool in experimental work, supporting description of chemical reactions, a tool that plays a role similar to a spectrophotometer or a chromatographic column, a tool that can provide information about the system under consideration, and a powerful tool, the popularization of which was achieved thanks to the fast progress of computer power and the hard work of people who made the transformation from pure theory to computer programs possible. Their efforts were appreciated – in 1998, the Nobel Prize was awarded to Walter Kohn “for his development of the density functional theory” and John Pople “for his development of computational methods.”

From the point of view of the experimentalist, the apparatus of quantum chemistry can be perceived similarly as the NMR spectrometer. One knows that the quality of the obtained NMR spectrum depends not only on the magnetic field of the magnet but also on the signal-processing capabilities. To successfully use NMR spectroscopy in experimental work, detailed knowledge about the technology of production and preparation of the magnets and electronic equipment is not a requisite. It is enough to keep in mind that with the given frequency one gets corresponding accuracy and information. All the rest is simply skill in sample preparation and expertise in spectrum interpretation. For effective usage of computational techniques of quantum chemistry, one must be aware of applied approximation to tune the accuracy of calculations and possess knowledge of the physicochemical phenomenon one wants to describe.

The aim of the present chapter is to provide a gentle introduction to basic quantum chemistry methods – the methods of solving the electronic Schrödinger equation. The chapter is intended for people starting their adventure with computational chemistry and wanting it to become the tool, not the aim itself.

When discussing quantum chemistry, we cannot totally avoid quantum mechanics. However, let us use another comparison: Traveling abroad it is good to know some basic expressions in the local language of the country you go to. It makes life easier and gives pleasure in interpersonal contacts. Still, no one expects a tourist to speak a language as fluently as a native speaker. Therefore, to efficiently apply computational techniques in experimental research, one has to learn some basic quantum mechanical terms that will help during the journey through the remainder of this chapter.

Quantum Mechanics for Dummies

We will begin with the basic terms of quantum mechanics. In this section, they will be introduced in an intuitive manner, to enable understanding of the next sections' content, even by beginners.

We consider a system of N electrons in the field produced by the potentials arising from nuclei (the nuclei are not treated as particles consisting of nucleons but just as a point source of the electrostatic potential). We are interested only in one particular case:

- The probability of finding the electrons of the considered system on the infinite distance from the nuclei is equal to zero. In other words, we want our system to constitute the whole entirety, not breaking into separate and independent parts (this would be the case of the interaction of two electrons with no attraction – two negative charges would repel each other to infinity).
- The energies of this system constitute the discrete spectrum.
- We want to know only the lowest value of energy (the wider approach can be found in the next volume of the present book).

With such limitations, we do not need to consider all of the different general cases and can simply concentrate on the bound-state chemistry.

A central notion in quantum chemistry is a wave function. This is a function characterizing a state of the system. Therefore, it depends on the variables that are adequate for the given system. This means that the wave function has to depend, at least, on spatial coordinates describing motions of the particles in the investigated system. Moreover, the wave function depends on so-called spin variables (spin is an additional degree of freedom included a posteriori in nonrelativistic quantum mechanics). This spin dependency can be built into the wave function by introducing a spin function. For instance, for the electron with a label 1, its wave function depends on the spatial coordinates x_1, y_1 and z_1 and is multiplied by the spin function $\alpha(\sigma_1)$ or $\beta(\sigma_1)$, where σ_1 is a spin variable. The spin functions must fulfill the following requirements:

$$\int \alpha^*(\sigma_1) \alpha(\sigma_1) d\sigma_1 = \int \beta^*(\sigma_1) \beta(\sigma_1) d\sigma_1 = 1 \quad (1)$$

$$\int \alpha^*(\sigma_1) \beta(\sigma_1) d\sigma_1 = \int \beta^*(\sigma_1) \alpha(\sigma_1) d\sigma_1 = 0, \quad (2)$$

where the integration is carried out over the spin variable, which can be treated as the integration variable only. Such a construction may seem to be somehow unnatural; however, it is a convenient way of ascribing spins to the electrons without dealing with its origins.

In general, the wave function must depend on time to reproduce information about the time evolution of the system. However, since we are interested only in the ground state of the system, we can neglect the time dependency. Considering the bound state is equivalent to imposing the condition of the square integrability of the wave function. The integral over all variables in the full range must exist:

$$\iint \dots \int f^* f d\tau = q, \quad (3)$$

where q is a finite real number. An asterisk under the integral denotes the complex conjugate; it comes from the fact that the wave function can be complex in general. The square-integrability condition ensures that the wave function vanishes for the infinite values of all spatial variables and, therefore, that our molecule is kept together. In the above expression, the integration intervals and the integration variables are not stated explicitly. For the investigated N -electron system, the wave function depends on the $3N$ spatial variables (for each particle i , we have x_i , y_i , and z_i coordinates) and additionally N spin variables (σ_i for the particle i):

$$f = f(x_1, y_1, z_1, \sigma_1, x_2, y_2, z_2, \sigma_2, \dots, x_N, y_N, z_N, \sigma_N). \quad (4)$$

The volume element in this $4N$ -dimensional space is

$$d\tau = dV \cdot d\sigma, \quad (5)$$

where the spatial part can be written as

$$dV = dx_1 dy_1 dz_1 dx_2 dy_2 dz_2 \dots dx_N dy_N dz_N, \quad (6)$$

and the spin part is

$$d\sigma = d\sigma_1 d\sigma_2 \dots d\sigma_N. \quad (7)$$

The spatial variables change from $-\infty$ to ∞ and spin variables can take allowed values. One can see that writing all of the integrals, variables, and volume elements explicitly takes time and a lot of paper, even for relatively small systems. Therefore, one usually keeps them in mind, not writing them down.

The wave function contains all of the information about the state of the system. In order to extract it, operators are applied. An operator can be understood by an analogy to a function. The function ascribes a number to a number and the operator ascribes a function to another function. In other words, the operator is a recipe for how to obtain one function from another:

$$\widehat{A}f = g. \quad (8)$$

We will denote operators by hats above the symbol to distinguish them from functions and numbers. One of the particularly interesting cases is when the function g is proportional to the function f ,

$$g = af, \quad (9)$$

where a is a number. Then Eq. 8 takes the form

$$\widehat{A}f = af. \quad (10)$$

Equation 10 is called an eigenvalue equation of the operator \widehat{A} . The function f fulfilling this equation is called an eigenfunction and a is an eigenvalue of the operator \widehat{A} . The Schrödinger equation

$$\widehat{H}\Psi = E\Psi \quad (11)$$

is a typical eigenvalue equation in which the Hamilton operator \widehat{H} extracts the information about the energy E of the system from the wave function Ψ .

Like in the case of the wave function, we will not consider operators in general. Let us concentrate on the Hamilton operator and its properties to simplify our discussion. We need our operators ascribed to observables (Hamiltonian among others) to satisfy the following requirements:

- Linearity – the operators must fulfill the condition

$$\widehat{A}(\alpha f + \beta g) = \alpha\widehat{A}f + \beta\widehat{A}g, \quad (12)$$

where now α and β are numbers. This seems simple and obvious; however, it is not the property of all operators. For instance, the square root is not a linear operator, since the square root of the sum is not equal to the sum of the square roots.

- Real eigenvalues – only real values can be measured in a laboratory.

For these reasons, we will be interested in so-called Hermitian operators that can be defined by the relation

$$\iint \dots \int f_1^* (\widehat{A}f_2) d\tau = \iint \dots \int f_2 (\widehat{A}f_1)^* d\tau. \quad (13)$$

All the functions, variables, and integration intervals remain the same as in Eq. 3. Writing of all these things in the expressions was already troublesome enough, and things become even more complicated when operators appear. In order to make life easier, Dirac notation can be applied. In this notation Eq. 13 has the form

$$\langle f_1 | \widehat{A}f_2 \rangle = \langle \widehat{A}f_1 | f_2 \rangle, \quad (14)$$

where the left-hand side can be equivalently written as $\langle f_1 | \widehat{A} | f_2 \rangle$, and the integral of Eq. 3 becomes simply

$$\langle f | f \rangle = q. \quad (15)$$

In this very convenient notation, it is also assumed that the integration intervals and variables flow from the context.

Let us look at Hermitian operators more carefully, considering them in the example of Hamiltonian. It has already been mentioned that such operators have real eigenvalues. Furthermore, the eigenfunctions of the Hermitian operator that correspond to different eigenvalues are orthogonal. In other words, for

$$\widehat{H} f_1 = E_1 f_1 \quad \text{and} \quad \widehat{H} f_2 = E_2 f_2, \quad (16)$$

where $(E_1 \neq E_2)$, one has

$$\langle f_1 | f_2 \rangle = \langle f_2 | f_1 \rangle = 0. \quad (17)$$

This will be a very useful property, since it will cause various terms in complicated expressions to vanish. In the case of degeneration, or, in other words, when one of the eigenvalues corresponds to two or more eigenfunctions, the eigenfunctions f_1 and f_2 can be orthogonalized.

It is worth considering the integral

$$\langle f_1 | \widehat{H} | f_1 \rangle. \quad (18)$$

Since f_1 is the eigenfunction of \widehat{H} with the eigenvalue E_1 , it is obvious that

$$\langle f_1 | \widehat{H} | f_1 \rangle = \langle f_1 | E_1 f_1 \rangle = E_1 \langle f_1 | f_1 \rangle. \quad (19)$$

It would be certainly more convenient if the result was a single number – the eigenvalue E_1 . This would be the case if $\langle f_1 | f_1 \rangle = 1$ or if we say the function f_1 was normalized to unity. It would be consistent with the interpretation of the integral $\langle f_1 | f_1 \rangle$ as the probability of finding the system in the whole space – it should be surely equal 1. This is a very handy requirement. Any function that does not possess this property can be normalized by multiplying by the normalization factor $\mathcal{N} = 1/\sqrt{\langle f_1 | f_1 \rangle}$. Then, the new function \tilde{f}_1 will be given as

$$\tilde{f}_1 = \mathcal{N} f_1. \quad (20)$$

This new function \tilde{f}_1 is also an eigenfunction of the Hamiltonian, since f_1 was only divided by the number $\sqrt{\langle f_1 | f_1 \rangle}$ and Hamiltonian is linear:

$$\widehat{H} \tilde{f}_1 = \widehat{H} \mathcal{N} f_1 = \mathcal{N} \widehat{H} f_1 = \mathcal{N} E f_1 = E \mathcal{N} f_1 = E \tilde{f}_1. \quad (21)$$

In the case of unnormalized functions, the expression for the eigenvalue E_1 can be obtained from Eq. 19:

$$E_1 = \frac{\langle f_1 | \widehat{H} | f_1 \rangle}{\langle f_1 | f_1 \rangle}. \quad (22)$$

However, many of the applied functions are not the eigenfunctions of the Hamiltonian. Therefore, let us investigate another interesting integral,

$$\langle g | \widehat{H} | g \rangle, \quad (23)$$

where g is not an eigenfunction of \widehat{H} . In order to calculate this integral, an alternate important property of the Hermitian operators needs to be exploited: the fact that their eigenfunctions constitute a complete basis set. Each function depending on the same variables as the eigenfunctions can be expressed as the linear combination of the basis functions. Now, this concept seems to be hard-core mathematics; however, anybody using computational techniques knows well that the two things one must input to the ab initio program are the method and the basis set. Hence, let us make a break from the general considerations of operators and abstract space functions and concentrate for a while on the basis set concept in the example of simple trigonometric functions.

In a calculus course, one learns how to express a function using a set of other functions. For example, consider $\sin x$ function and expand it in the Taylor series around 0:

$$\sin x = \sum_{i=1}^{\infty} \frac{(-1)^{i-1}}{(2i-1)!} x^{2i-1} = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \dots \quad (24)$$

In Eq. 24, $\sin x$ function is expressed in the basis set of monomials:

$$\sin x = \sum_{k=1}^{\infty} c_k x^k, \quad (25)$$

where c_k are the expansion coefficients that need to be determined. In our case it is simple, since c_k result directly from the Taylor expansion and are equal:

Table 1 Taylor expansion of x function and standard deviation for various expansion lengths

n	F_n	σ_n
1	x	3.93×10^{-1}
3	$x - \frac{1}{6}x^3$	4.05×10^{-2}
5	$x - \frac{1}{6}x^3 + \frac{1}{120}x^5$	2.08×10^{-3}
7	$x - \frac{1}{6}x^3 + \frac{1}{120}x^5 - \frac{1}{5,040}x^7$	6.39×10^{-5}

$$c_k = \begin{cases} \frac{(-1)^{(k-1)/2}}{k!} & \text{for odd } k, \\ 0 & \text{for even } k. \end{cases} \quad (26)$$

The summation in Eqs. 24 and 25 goes from 1 to ∞ . In practice, finite and possibly short expansions are applied:

$$F_n(x) = \sum_{i=1}^n c_i x^i. \quad (27)$$

This truncation of the series introduces an approximation to our function.

Let us analyze the x function in the range $x \in \langle -\frac{\pi}{2}; \frac{\pi}{2} \rangle$. The standard deviation works well as the accuracy measure:

$$\sigma_n \sqrt{\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} (\sin x - F_n(x))^2 dx}. \quad (28)$$

Table 1 summarizes data for small n values. Increasing the number of expansion terms causes a decrease of the standard deviation values and more accurate representation of the original x function. Given the required accuracy of the calculation, the necessary value length of the expansion, n , can be found.

The following question arises: Why use Taylor expansion instead of x function itself, if one needs to worry about the expansion accuracy? The answer is straightforward: simplifications and savings. It is much easier to operate on the polynomials than on the trigonometric functions (for instance, the integral $\int (x^i)^2 dx$ is much easier to handle than $\int \sin^2 x dx$). Moreover, the required accuracy can often be obtained with a relatively short expansion.

Let us now make the considerations more general. As was stated before, the set of eigenfunctions of the Hermitian operator \hat{H} is complete and orthonormal – functions are orthogonal and normalized:

$$\forall_{i,j} \left\langle f_i \left| f_j \right. \right\rangle = \delta_{ij}, \quad (29)$$

where δ_{ij} is a Kronecker symbol that takes value 1 for $i = j$ and 0 otherwise. The basis set completeness means that each function depending on the same set of variables can be expressed by the basis functions:

$$g = \sum_{i=1}^{\infty} c_i f_i, \quad (30)$$

where coefficients c_i need to be found. Knowing the normalized g function makes this task simple, because of the orthonormality of the $\{f_i\}$ set, the coefficients will be equal:

$$c_i = \langle f_i | g \rangle, \quad (31)$$

since

$$\langle f_i | g \rangle = \sum_{j=1}^{\infty} c_j \langle f_i | f_j \rangle = \sum_{j=1}^{\infty} c_j \delta_{ij} = c_i. \quad (32)$$

(Only one term for $i = j$, c_i , remains; all other vanish for the Kronecker delta equals zero if $i \neq j$.) Similarly,

$$\langle g | g \rangle = \sum_{j=1}^{\infty} c_j^* c_j. \quad (33)$$

However, things are not that easy, since we usually apply the expansion (Eq. 30) when we do not know the g function. Thus, the integrals (Eqs. 31 and 33) should be perceived rather as the interpretation of the c_i coefficients than the direct recipe for the calculations. From Eq. 30, the g function can be treated as the linear combination of the f_i functions. Moreover (see Eq. 33), the probability that a system described by the function g is in the state f_j is given by $c_j^* c_j$.

Now let us consider the expression

$$\langle g | \hat{H} | g \rangle \equiv \langle \hat{H} \rangle_g, \quad (34)$$

when g is not the Hamiltonian eigenfunction. Using the expansion Eq. 30 and the fact that f_i are the Hamiltonian eigenfunctions (Eq. 16), one obtains

$$\langle g | \hat{H} | g \rangle = \sum_{j=1}^{\infty} c_j^* c_j E_j. \quad (35)$$

The above integral is called an average (expectation) value, and Eq. 35 for Hamiltonian carries the information about the average energy of the system in the state described by the g function. Looking closer at Eq. 35 shows that this average energy is simply a weighted average of all possible E_j energies of the system. The weights are determined by the $c_j^* c_j$ products – the probability of finding the system in the f_j state. It should be noticed that we used the linearity of the Hamiltonian operator to achieve this result.

Conclusion? Very optimistic: We can say something about the sought energy value not knowing the eigenfunctions of the operator of interest, since for calculations of Eq. 35, we do not need f_j functions. Strange? Not at all, if we recall some linear algebra: Using three basis vectors, we can describe each and every point in the 3D space. Likewise, the wave function can be perceived as a vector, the Hermitian operator as the symmetric transformation matrix, an integral $\langle f | g \rangle$ as the dot product, orthogonality of functions as orthogonality of vectors, and normalization as dividing the vector components by its length.

Now, when the “linear algebra” term has already appeared, let us see how it is applied for solving the eigenequation. Almost all calculations are performed by applying the basis functions. This means that the unknown function Ψ describing the investigated system is expressed in the basis of known functions χ_i (see Eq. 30):

$$\Psi \approx \sum_{i=1}^n c_i \chi_i = \Phi, \quad (36)$$

where Ψ is the eigenfunction of the Hamiltonian corresponding to a given eigenvalue E (Eq. 11). The task is to find such c_i coefficients that the function Φ would be the best approximation to Ψ . Since Φ is an approximation to the wave function, the corresponding energy will also be only approximated. Let us call the approximation E_Φ . Basis functions χ_i are not the Hamiltonian eigenfunctions; therefore, to estimate the energy, an average value must be calculated. Therefore, substituting Eq. 36 for Eq. 11 and multiplying by $\Psi^* = \langle \sum_i c_i \chi_i |$ on both sides gives¹

$$\text{LHS} = \left\langle \sum_i c_i \chi_i \left| \hat{H} \left| \sum_j c_j \chi_j \right. \right. \right\rangle = \sum_i \sum_j c_i^* c_j \langle \chi_i | \hat{H} | \chi_j \rangle, \quad (37)$$

$$\text{RHS} = E_\Phi \sum_i \sum_j c_j^* c_j \langle \chi_i | \chi_j \rangle. \quad (38)$$

¹Here, different subscripts appear on the both sides of the integral. The sum does not depend on the name of the summation index; thus, any subscript can be applied. However, one should not apply the same index on both sides of the integral, since it can cause the erroneous omission of the off-diagonal terms. Compare the overlap integral

$$\begin{aligned} \langle \Psi | \Psi \rangle &= \left\langle \sum_i c_i \chi_i \left| \sum_j c_j \chi_j \right. \right\rangle = \langle c_1 \chi_1 + c_2 \chi_2 + c_3 \chi_3 + \dots | c_1 \chi_1 + c_2 \chi_2 + c_3 \chi_3 + \dots \rangle \\ &= c_1^* c_1 \langle \chi_1 | \chi_1 \rangle + c_1^* c_2 \langle \chi_1 | \chi_2 \rangle + c_1^* c_3 \langle \chi_1 | \chi_3 \rangle + \dots \end{aligned}$$

The explicit writing of all terms shows that not only the integrals $\langle \chi_i | \chi_i \rangle$ with the same function on both sides are present, but also the contributions $\langle \chi_i | \chi_j \rangle$ with $i \neq j$. Therefore, the diversification of the subscripts prevents mistakes.

In order to further simplify the notation, let us denote $\langle \chi_i | \hat{H} | \chi_j \rangle$ by H_{ij} and $\langle \chi_i | \chi_j \rangle$ by S_{ij} . Then,

$$\sum_i \sum_j c_j^* c_j H_{ij} = E_\Phi \sum_i \sum_j c_i^* c_j S_{ij}. \quad (39)$$

Equivalently, in the matrix form

$$\mathbf{H}\mathbf{c} = \mathbf{S}\mathbf{c}E_\Phi, \quad (40)$$

where \mathbf{H} is the Hamiltonian matrix with the elements H_{ij} , \mathbf{S} is called the overlap matrix and is built of the overlap integrals S_{ij} , and \mathbf{c} denotes the vector of the c_i coefficients. The basis sets applied in practice are usually non-orthogonal, which causes the off-diagonal terms in the \mathbf{S} matrix to not vanish.

Such a method of finding approximate eigenvalues and eigenvectors of the Hamiltonian is known as the Ritz method and is frequently applied in quantum chemistry.

This simple introduction of basic terms of quantum mechanics is obviously far from complete. One can notice the lack of further discussion of the degeneration, the continuum spectrum, and many other topics. For these we encourage the reader to dive into the following excellent books on quantum mechanics and chemistry: Atkins and Friedman (2005), Griffiths (2004), Levine (2008), Lowe and Peterson (2005), McQuarrie and Simon (1997), Pielka (2007), Ratner and Schatz (2000), and Szabo and Ostlund (1996).

On the Way to Quantum Chemistry

For the sake of simplification, we assume that the energy of the ground state of our system differs from other energy values. This allows one to avoid embroilment in technical details that are unnecessary at this point. Our system is described by the wave function Ψ fulfilling the Schrödinger Eq. 11. It is important to notice that this eigenvalue equation can be solved exactly only for hydrogen atoms. Any more complicated system requires approximate techniques. In order to explain this complication, let us look into the Hamilton operator. For the system of N electrons and M nuclei, the full Hamiltonian is a sum of the following terms:

- Kinetic energy of electrons, \hat{T}_e
- Kinetic energy of nuclei, \hat{T}_n
- Energy of interactions between electrons, \hat{V}_{ee}
- Energy of interactions between nuclei, \hat{V}_{nn}
- Energy of interactions between a nucleus and an electron, \hat{V}_{ne}

In the atomic units, these terms have the following form:

$$\hat{T}_e = -\frac{1}{2} \sum_{i=1}^N \nabla_{\mathbf{r}_i}^2 \quad (41)$$

$$\hat{T}_n = -\sum_{i=1}^M \frac{1}{2m_i} \nabla_{\mathbf{R}_i}^2 \quad (42)$$

$$\hat{V}_{ee} = \sum_{i=1}^N \sum_{j>i}^N \frac{1}{r_{ij}} \quad (43)$$

$$\hat{V}_{nn} = \sum_{i=1}^M \sum_{j>i}^M \frac{Z_i Z_j}{R_{ij}} \quad (44)$$

$$\hat{V}_{ne} = -\sum_{i=1}^N \sum_{j=1}^M \frac{Z_j}{|\mathbf{r}_i - \mathbf{R}_j|} \quad (45)$$

where m_i is the mass of the nucleus i , Z_i stands for the nuclear charge, and \mathbf{r}_{ij} denotes the distance between the electrons i and j , $r_{ij} = |\mathbf{r}_i - \mathbf{r}_j|$. Likewise, R_{ij} refers to the internuclear distance, $R_{ij} = |\mathbf{R}_i - \mathbf{R}_j|$. The presence of the mutual distances between the particles causes a serious problem when solving the Schrodinger equation; it does not allow one to decouple the equations.

Fortunately, from the chemist's point of view, such a Hamiltonian is not very useful. The chemist is not interested in each and every bit of information one can get about any N -electron M -nuclei system; however, she or he is focused on the given molecule, its conformations, interactions with the environment, and properties (spectroscopic, magnetic, electric, and so on). What makes quantum mechanics a valuable tool for chemists is the Born–Oppenheimer approximation, discussed in detail in the previous chapter of this volume. Let us briefly summarize it to maintain consistent notation throughout the chapter.

The chemist is concerned with the relative positions of the nuclei in the molecule and with the internal energy, but not with the motions of the molecule as a whole. This motion can be excluded from our consideration, for example, by elimination of the center-of-mass translation. Moreover, the intermolecular (electrostatic) forces acting on electrons and nuclei would be similar. This would cause much slower internal motion of the heavy nuclei in comparison to light electrons. For this reason, the approximate description of electron motion with parametric dependence on the static positions of nuclei is justified. Such reasoning leads to adiabatic approximation and finally to Born–Oppenheimer approximation.

According to this approximation, the Hamiltonian can be written as

$$\hat{H} = \hat{T}_n + \hat{H}_e + \hat{V}_{nn}, \quad (46)$$

where \widehat{T}_n now has a meaning of nuclear kinetic energy of the molecule for which the center of mass is stopped (however, there are still oscillations and rotations), and

$$\widehat{H}_e = \widehat{T}_e + \widehat{V}_{ee} + \widehat{V}_{en} \quad (47)$$

is called an electronic Hamiltonian, and it represents the energy of the system after omitting the nuclear kinetic energy term and nuclear repulsion. One can now focus on the solution of the equation of the form

$$\left[\widehat{T}_n(\mathbf{R}) + \widehat{H}_e(\mathbf{r}; \mathbf{R}) + \widehat{V}_{nn}(\mathbf{R}) \right] \Psi(\mathbf{r}, \mathbf{R}) = E \Psi(\mathbf{r}, \mathbf{R}). \quad (48)$$

Here, the dependence on the electronic spatial variables $\mathbf{r} = (x_1, y_1, z_1, \dots, x_N, y_N, z_N)$ and the nuclear spatial variables $\mathbf{R} = (X_1, Y_1, Z_1, \dots, X_M, Y_M, Z_M)$ is written explicitly. The semicolon sign in the \widehat{H}_e term denotes the parametric dependence – for various \mathbf{R} the various electronic equations are obtained.

With such a Hamiltonian, it seems reliable to distinguish also the nuclear $f(\mathbf{R})$ and electronic $\Psi_e(\mathbf{r}; \mathbf{R})$ part in the wave function

$$\Psi(\mathbf{r}, \mathbf{R}) \approx \Psi_e(\mathbf{r}; \mathbf{R}) f(\mathbf{R}), \quad (49)$$

which leads to a significant reduction of the problem.

Now Eq. 48 can be separated into three equations:

$$\widehat{H}_e(\mathbf{r}; \mathbf{R}) \Psi_e(\mathbf{r}; \mathbf{R}) = E_e(\mathbf{R}) \Psi_e(\mathbf{r}; \mathbf{R}) \quad (50)$$

$$\left(\widehat{H}_e(\mathbf{r}; \mathbf{R}) + \widehat{V}_{nn}(\mathbf{R}) \right) \Psi_e(\mathbf{r}; \mathbf{R}) = U(\mathbf{R}) \Psi_e(\mathbf{r}; \mathbf{R}) \quad (51)$$

$$\left(\widehat{T}_n(\mathbf{R}) + \widehat{U}(\mathbf{R}) \right) f(\mathbf{R}) = E f(\mathbf{R}) \quad (52)$$

The first two describe electronic motion for a given position of nuclei. The difference between $E_e(\mathbf{R})$ and $U(\mathbf{R})$ is that in $U(\mathbf{R})$ nuclear repulsion energy is taken into account. These equations are milestones in our considerations for two reasons. First, since we are now talking about “fixed positions of the nuclei,” finally we have got molecules instead of an unspecified system containing some electrons and some nuclei. The second is hidden in Eq. 52: Electronic energy and nuclear repulsion energy constitute the potential, in which nuclei are moving. That is why the proper description of electronic movement in a molecule is so important: The electrons glue the whole molecule together.

Our attention in the rest of the chapter will be focused only on Eq. 50; hence, to simplify notation, all the subscripts denoting the electronic case will be omitted:

$$\widehat{H}_e \rightarrow \widehat{H} \quad (53)$$

$$\Psi_e \rightarrow \Psi \quad (54)$$

$$\widehat{H}_e \Psi_e (r; R) = E_e \Psi_e (r; R) \rightarrow \widehat{H} \Psi = E \Psi \quad (55)$$

The electronic wave function Ψ satisfies all the requirements discussed in the previous sections, depends on the coordinates of N electrons, and additionally must be antisymmetric with respect to the exchange of coordinates of two electrons.²

It should be noted that the analytic solution of Eq. 50 is not known even for the smallest molecule, i.e., H_2 . Therefore, the approximate techniques must be applied to extract the necessary information about molecules of interest. Quantum mechanics provides two tools:

- Variational principle
- Perturbation theory

Variational Principle: An Indicator

The variational principle allows one to judge the quality of the obtained solutions. It can be formulated as follows: For the arbitrary trial function χ that is square-integrable, differentiable, and antisymmetric and depends on the same set of variables as a sought ground-state Ψ_0 function, we have

$$E_0 \leq \frac{\langle \chi | \widehat{H} | \chi \rangle}{\langle \chi | \chi \rangle}, \quad (56)$$

where E_0 is the ground-state energy corresponding to Ψ_0 (Eq. 50). The important consequence of the variational principle is that to estimate the energy of the system, one does not need to solve the eigenequation (this we already know; see Eq. 35), and moreover – what is crucial – the estimated energy value will always not be lower than the exact eigenvalue E_0 .

The proof of the inequality (Eq. 56) is straightforward and can be derived from Eqs. 30 and 35. The function χ that satisfies the above requirements can be expanded on the basis of the Hamiltonian eigenfunctions:

$$\chi = \sum_{i=0}^{\infty} c_i f_i, \quad (57)$$

²In order to explain the antisymmetry requirement, we have to refer again to theory that is beyond the scope of the present chapter. Let us simply state here that wave functions must be antisymmetric without belaboring the point. This will mean that the exchange of the coordinates of the two electrons causes the wave function to change the sign: $\Psi(\tau_1, \tau_2) = -\Psi(\tau_2, \tau_1)$.

where f_i fulfill the eigenproblem $\widehat{H} f_i = E_i f_i$. Thus,

$$\frac{\langle \chi | \widehat{H} | \chi \rangle}{\langle \chi | \chi \rangle} = \frac{\sum_{i=0}^{\infty} c_i^* c_i E_i}{\sum_{i=0}^{\infty} c_i^* c_i} \geq \frac{\sum_{i=0}^{\infty} c_i^* c_i E_0}{\sum_{i=0}^{\infty} c_i^* c_i} = E_0, \quad (58)$$

with the assumption that E_0 is the lowest of all Hamiltonian eigenvalues (Atkins and Friedman 2005; Levine 2008; Lowe and Peterson 2005; McQuarrie and Simon 1997; Piela 2007; Szabo and Ostlund 1996).

Perturbation Calculus: The Art of Estimation

Due to the variational principle that is satisfied for the electronic Hamiltonian, the group of methods of searching for parameters optimizing the energy value can be constructed. The way of verification of the given wave function is the corresponding energy value: The lower, the better. Besides this “quality control,” the variational principle does not give the prescription for the choice of the trial wave functions. Here comes the perturbation calculus – the method frequently applied in physics for the estimation of the functions or values on the basis of partial knowledge about the solutions of the investigated problem. We will consider here the Rayleigh–Schrödinger variant of the perturbation calculus (Atkins and Friedman 2005; Levine 2008; Lowe and Peterson 2005; McQuarrie and Simon 1997; Piela 2007; Ratner and Schatz 2000).

Let us assume that the total electronic Hamiltonian of the investigated system can be divided into

$$\widehat{H} = \widehat{H}^0 + \widehat{H}^1, \quad (59)$$

in such a fashion that we know the exact solutions of

$$\widehat{H}^0 \Psi_k^{(0)} = E_k^{(0)} \Psi_k^{(0)}, \quad (60)$$

where the subscript k enumerates the eigenvalues of the \widehat{H}^0 operator in such a way that $E_0^{(0)}$ is the lowest energy. Now, one can say that the operator \widehat{H} describes the system for which \widehat{H}^0 is an unperturbed operator and \widehat{H}^1 denotes a perturbation. We can assume that if the change in the system represented by \widehat{H}^1 is minor, then the functions $\Psi_k^{(0)}$ would be a good approximation to Ψ_k . Considering \widehat{H}^0 , one postulates its Hermiticity and that its eigenvalues are not degenerate (in our case, for the ground state at least $E_0^{(0)}$ must not be equal to any other eigenvalue). This condition will become clear in a moment.

Knowing only the unperturbed solutions (Eq. 60), we would like to say something more about the ground-state energy of the investigated system. Nothing is easier – we can calculate the average value of the full electronic Hamiltonian with

the $\Psi_0^{(0)}$ function. The variational principle states that the resulting energy will be not lower than the exact energy:

$$E_0 \leq \left\langle \Psi_0^{(0)} \left| \widehat{H}^0 + \widehat{H}^1 \right| \Psi_0^{(0)} \right\rangle = E_0^{(0)} + E_0^{(1)}. \quad (61)$$

The term modifying $E_0^{(0)}$ is simply

$$E_0^{(1)} = \left\langle \Psi_0^{(0)} \left| \widehat{H}^1 \right| \Psi_0^{(0)} \right\rangle. \quad (62)$$

So far, the only new thing is the manner of partitioning the total energy into the energy of the unperturbed system and the corrections (where $E_0^{(1)}$ is not the only term):

$$E_0 = E_0^{(0)} + E_0^{(1)} + E_0^{(2)} + \dots \quad (63)$$

Likewise, the wave function can be written as

$$\Psi_0 = \Psi_0^{(0)} + \Psi_0^{(1)} + \Psi_0^{(2)} + \dots, \quad (64)$$

where $\Psi_0^{(1)}, \Psi_0^{(2)}$ and so forth are the corrections to the wave function of the unperturbed system $\Psi_0^{(0)}$. Now, the electronic Schrödinger Eq. 50 becomes

$$\begin{aligned} & \left(\widehat{H}^0 + \widehat{H}^1 \right) \left(\Psi_0^{(0)} + \Psi_0^{(1)} + \Psi_0^{(2)} + \dots \right) \\ & = \left(E_0^{(0)} + E_0^{(1)} + E_0^{(2)} + \dots \right) \left(\Psi_0^{(0)} + \Psi_0^{(1)} + \Psi_0^{(2)} + \dots \right). \end{aligned} \quad (65)$$

Introducing the expansions (Eqs. 63 and 64) does not increase our knowledge about the energy or the wave function; it is only a different way of expressing the unknowns by other unknowns. However, now we have a starting point for further investigations.

The comparison of the terms on the left- and right-hand side of the above expression is instructive. Let us regard as similar the terms with the same sum of the superscripts (so-called perturbation order, by analogy to the multiplication and ordering of polynomials). Simple multiplication in Eq. 65 and directing the terms of the same order to separate equations gives

$$\widehat{H}^0 \Psi_0^{(0)} = E_0^{(0)} \Psi_0^{(0)}, \quad (66)$$

$$\widehat{H}^0 \Psi_0^{(1)} + \widehat{H}^1 \Psi_0^{(0)} = E_0^{(0)} \Psi_0^{(1)} + E_0^{(1)} \Psi_0^{(0)}, \quad (67)$$

$$\begin{aligned} \widehat{H}^0 \Psi_0^{(2)} + \widehat{H}^1 \Psi_0^{(1)} &= E_0^{(0)} \Psi_0^{(2)} + E_0^{(1)} \Psi_0^{(1)} + E_0^{(2)} \Psi_0^{(0)}. \\ &\vdots \end{aligned} \quad (68)$$

These equations link the corrections to the wave function and to the energy. Before detailed investigation of the subsequent corrections, one more thing should be underlined. Up to now, the function Ψ_0 is not normalized; only $\Psi_k^{(0)}$ are normalized. Until the corrections to Ψ_0 were found, we would not be able to normalize it. We can only write the normalization constant as $\mathcal{N} = \frac{1}{\sqrt{\langle \Psi_0 | \Psi_0 \rangle}}$. However, it is not necessary at this moment. Now the intermediate normalization condition is more useful:

$$\langle \Psi_0^{(0)} | \Psi_0 \rangle = 1. \quad (69)$$

Such a concept is based on the information that the eigenfunctions of \widehat{H}^0 form an orthonormal complete set of functions (that is one of the reasons why the Hermiticity of \widehat{H}^0 was required) and they can be applied to express any other function, for instance, Ψ_0 as

$$\Psi_0 = \sum_{k=0}^{\infty} c_k \Psi_k^{(0)} + \sum_{k \neq 0}^{\infty} c_k \Psi_k^{(0)}. \quad (70)$$

In this linear combination, the function $\Psi_0^{(0)}$ has a distinguished meaning ($c_0 = 1$); this is the approximation of the wave function of the considered system. Therefore, one can require that $\Psi_0^{(0)}$ does not have a contribution to the higher corrections: $\Psi_0^{(1)}$, $\Psi_0^{(2)}$, and so on:

$$\sum_{k \neq 0}^{\infty} c_k \Psi_k^{(0)} = \Psi_0^{(1)} + \Psi_0^{(2)} + \dots \quad (71)$$

Here, the benefits from the intermediate normalization are obvious: The function $\Psi_0^{(0)}$ is orthogonal to each of the corrections (or, in other words, the corrections are defined in such a way that they are orthogonal to $\Psi_0^{(0)}$).

Therefore, there is an additional set of equations to be satisfied:

$$\langle \Psi_0^{(0)} | \Psi_0^{(n)} \rangle = \delta_{0n}, \quad (72)$$

where the superscript n denotes the n th-order correction to the ground-state wave function Ψ_0 . Now we can go back to Eqs. 66, 67, and 68 and extract the corrections to energy. For this purpose, each of the equations must be multiplied from the left-hand side by $\Psi_0^{(0)}$ and integrated:

$$E_0^{(0)} = \langle \Psi_0^{(0)} | \widehat{H}^0 | \Psi_0^{(0)} \rangle, \quad (73)$$

$$E_0^{(1)} = \langle \Psi_0^{(0)} | \widehat{H}^1 | \Psi_0^{(0)} \rangle, \quad (74)$$

$$E_0^{(2)} = \langle \Psi_0^{(0)} | \widehat{H} | \Psi_0^{(1)} \rangle, \quad \vdots \quad (75)$$

The integrals $\langle \Psi_0^{(0)} | \widehat{H}^0 | \Psi_0^{(n)} \rangle$ vanish, since

$$\langle \Psi_0^{(0)} | \widehat{H}^0 | \Psi_0^{(n)} \rangle = \langle \widehat{H}^0 \Psi_0^{(0)} | \Psi_0^{(n)} \rangle = E_0^{(0)} \langle \Psi_0^{(0)} | \Psi_0^{(n)} \rangle = 0. \quad (76)$$

Thus, obtaining the energy corrections of any order is straightforward. The general expression for the n th-order correction can be written as

$$E_0^{(n)} = \langle \Psi_0^{(0)} | \widehat{H}^1 | \Psi_0^{(n-1)} \rangle \quad \text{for } n > 1. \quad (77)$$

The problem is that to obtain the corrections to the energy in the second or higher orders, the corrections to the wave function are necessary. Then, let us try to find $\Psi_0^{(1)}$. This function can be expressed as the linear combination of the functions from the orthonormal set $\{\Psi_k^{(0)}\}$ for $k \neq 0$:

$$\Psi_0^{(1)} = \sum_{k \neq 0}^{\infty} c_k^{(1)} \Psi_k^{(0)}, \quad (78)$$

where $c_k^{(1)}$ are the expansion coefficients in the first-order correction. Again, the whole thing reduces to finding the coefficients c_k . Substituting Eq. 78 to Eq. 67 gives

$$\left(\widehat{H}^0 - E_0^{(0)} \right) \sum_{k \neq 0}^{\infty} c_k^{(1)} \Psi_k^{(0)} = \left(E_0^{(1)} - \widehat{H}^1 \right) \Psi_0^{(0)}. \quad (79)$$

Integrating this equation with the $\Psi_l^{(0)}$ function leads to

$$\begin{aligned} \text{LHS} &= \left\langle \Psi_l^{(0)} \left| \widehat{H}^0 - E_0^{(0)} \right| \sum_{k \neq 0}^{\infty} c_k \Psi_k^{(0)} \right\rangle = \sum_{k \neq 0}^{\infty} c_k \left\langle \Psi_l^{(0)} \left| \widehat{H}^0 - E_0^{(0)} \right| \Psi_k^{(0)} \right\rangle \\ &= \sum_{k \neq 0}^{\infty} c_k \left(E_l^{(0)} - E_0^{(0)} \right) \left\langle \Psi_l^{(0)} \left| \Psi_k^{(0)} \right\rangle = \sum_{k \neq 0}^{\infty} c_k \left(E_l^{(0)} - E_0^{(0)} \right) \delta_{lk} \\ &= c_l \left(E_l^{(0)} - E_0^{(0)} \right) \end{aligned} \quad (80)$$

and

$$\begin{aligned} \text{RHS} &= \left\langle \Psi_l^{(0)} \left| E_0^{(1)} - \widehat{H}^1 \right| \Psi_0^{(0)} \right\rangle = E_0^{(1)} \left\langle \Psi_l^{(0)} \left| \Psi_0^{(0)} \right\rangle - \left\langle \Psi_l^{(0)} \left| \widehat{H}^1 \right| \Psi_0^{(0)} \right\rangle \\ &= - \left\langle \Psi_l^{(0)} \left| \widehat{H}^1 \right| \Psi_0^{(0)} \right\rangle. \end{aligned} \quad (81)$$

Altogether, these allow one to write the coefficients of the expansion (Eq. 78) as

$$c_l = \frac{\langle \Psi_l^{(0)} | \widehat{H}^1 | \Psi_0^{(0)} \rangle}{E_0^{(0)} - E_l^{(0)}}. \quad (82)$$

Hence, the first correction to the wave function is already known:

$$\Psi_0^{(1)} = \sum_{k \neq 0}^{\infty} \frac{\langle \Psi_k^{(0)} | \widehat{H}^1 | \Psi_0^{(0)} \rangle}{E_0^{(0)} - E_k^{(0)}} \Psi_k^{(0)}, \quad (83)$$

and, thereby, the second-order correction to the energy can be calculated as

$$E_0^{(2)} = \sum_{k \neq 0}^{\infty} \frac{\langle \Psi_k^{(0)} | \widehat{H}^1 | \Psi_0^{(0)} \rangle}{E_0^{(0)} - E_k^{(0)}} \langle \Psi_0^{(0)} | \widehat{H}^1 | \Psi_k^{(0)} \rangle. \quad (84)$$

This is also equivalently written as

$$E_0^{(2)} = \sum_{k \neq 0}^{\infty} \frac{\left| \langle \Psi_k^{(0)} | \widehat{H}^1 | \Psi_0^{(0)} \rangle \right|^2}{E_0^{(0)} - E_k^{(0)}}. \quad (85)$$

The energy difference in the denominator of the above expression cannot be equal to zero, and for this reason the non-degenerated ground state was assumed. The higher-order corrections are sought in a similar manner, which just requires more operations.

One of the interesting issues is the problem of variationality of the perturbation calculus built upon the variational Hamiltonian. This is, however, a sophisticated problem for advanced readers and will not be discussed here. It should be added, in summary, that the manner of partitioning of the Hamilton operator was arbitrary. The only prerequisites were the Hermitian character of the operators (for the eigenfunctions to form the orthonormal set) and the non-degenerated ground-state eigenenergy and nothing more. One should also remember that, in practice, even solving the unperturbed problem cannot be performed exactly and the approximations must be applied. The consequence could be the loss of accuracy for higher-order corrections. Moreover, good convergence of the perturbation expansion can be expected when the consecutive corrections are small in comparison to the total estimated value. However, in such a case, the low orders of the series would reproduce the sought value with relatively good accuracy. Thus, application of the low orders of perturbation calculus is highly recommended.

One-Electron Approximation: Describe One and Say Something About All

Equipped with general knowledge about the tools for the Schrödinger equation solution, one can move to many-electron systems.

The electronic Hamiltonian for any many-electron system in atomic units has the following form (compare Eqs. 41, 42, 43, 44, and 45):

$$\hat{H} = -\frac{1}{2} \sum_{i=1}^N \Delta_{r_i} + \sum_{i=1}^N \sum_{j=1}^M \frac{Z_j}{|\mathbf{r}_i - \mathbf{R}_j|} + \sum_{i=1}^N \sum_{j=i+1}^N \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}. \quad (86)$$

Let us look more closely. In the first term, we sum up over the number of electrons N ; in the second term, the summations run over the number of electrons N and number of nuclei M ; and the third term contains the double sum over the number of electrons N . Thus, one can simplify the notation of the first two terms:

$$-\frac{1}{2} \sum_{i=1}^N \Delta_{r_i} + \sum_{i=1}^N \sum_{j=1}^M \frac{Z_j}{|\mathbf{r}_i - \mathbf{R}_j|} = \sum_{i=1}^N \left(-\frac{1}{2} \Delta_{r_i} + \sum_{j=1}^M \frac{Z_j}{|\mathbf{r}_i - \mathbf{R}_j|} \right). \quad (87)$$

Now, denoting the term in parenthesis by $\hat{h}(i)$,

$$\hat{h}(i) = -\frac{1}{2} \Delta_{r_i} + \sum_{j=1}^M \frac{Z_j}{|\mathbf{r}_i - \mathbf{R}_j|}, \quad (88)$$

we get the part of the Hamiltonian depending only on the coordinates of one electron i (and nuclear coordinates, but it does not bother us). Hence, the total electronic Hamiltonian (Eq. 86) can be rewritten as the sum of one-electron and two-electron contributions:

$$\hat{H} = \sum_{i=1}^N \hat{h}(i) + \sum_{i=1}^N \sum_{j=i+1}^N \hat{g}(i, j), \quad (89)$$

where we introduced a symbol:

$$\hat{g}(i, j) = \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}. \quad (90)$$

It should be noticed that each of the one-electron Hamiltonians $\hat{h}(i)$ describes a single electron in the field of some potentials. Therefore, the exact solutions of the eigenvalue problem for these one-electron operators are available. The problem lies

in the $\hat{g}(i, j)$ operator that couples two electrons together: It is not possible to separate their coordinates exactly.

All electronic Hamiltonians have the same general form; they differ only in the number of electrons N and the nuclear potential hidden in $\hat{h}(i)$. One can choose any possible chemical compounds and try to write the corresponding equations; however, one would soon note the similarity of all of them. Therefore, we will not invest time in describing the procedure for a polypeptide or a nanotube, but for simplicity we will start from the two-electron helium atom (the simplest many-electron system) and later try to generalize the considerations. For the helium atom:

- $N = 2$ – two electrons
- $M = 1$ – one nucleus

The generalization of the helium discussion into the larger (N -electron) systems should be straightforward:

$$\hat{h}(1) + \hat{h}(2) = \sum_{i=1}^2 \hat{h}(i) \rightarrow \sum_{i=1}^N \hat{h}(i), \quad (91)$$

$$\hat{g}(1, 2) = \sum_{i=1}^2 \sum_{j=i+1}^2 \hat{g}(i, j) \rightarrow \sum_{i=1}^N \sum_{j=i+1}^N \hat{g}(i, j), \quad (92)$$

and finally

$$\hat{h}(i) = -\frac{1}{2}\Delta_{r_i} + \frac{Z_1}{|\mathbf{r}_i - \mathbf{R}_1|} \rightarrow -\frac{1}{2}\Delta_{r_i} + \sum_{j=1}^M \frac{Z_j}{|\mathbf{r}_i - \mathbf{R}_j|}. \quad (93)$$

Recall that the exact solutions of the one-electron problem are known, and the task is to solve the full problem. The ideas of the perturbation theory were explained in the previous section. Now it is the time to apply the knowledge. The one-electron part can be treated as the unperturbed Hamiltonian and the rest as the perturbation:

$$\hat{H}^0 = \hat{h}(1) + \hat{h}(2), \quad \hat{H}^1 = \hat{g}(1, 2), \quad (94)$$

where the normalized solutions for the one-electron part are known:

$$\hat{h}(1)\phi_i(1) = \epsilon_i\phi_i(1), \quad (95)$$

$$\hat{h}(2)\phi_j(2) = \epsilon_j\phi_j(2). \quad (96)$$

These functions require more attention. Although the electronic Hamiltonian – and thereby the one-electron operators – do not act on the spin variables, the wave

functions ϕ_k must carry on the spin dependence. Therefore, the function $\phi_k(l)$ is a product of the spatial part depending on the three spatial coordinates of the electron l and on the spin part and is called a spin-orbital. For simplicity, the set of coordinates τ_l is written as a label of the electron, i.e., l . Such convention will be applied from now on, with an exception where the τ_l labeling is really needed.

If an operator can be written as a sum of contributions acting on different variables, its eigenfunction takes a form of the product of the eigenfunctions of the subsequent operators in the summation. In the case of the helium atom, where \widehat{H}^0 is a sum of $\widehat{h}(1)$ and $\widehat{h}(2)$, the wave function $\Psi(1, 2)$ can be denoted as the product of one-electron functions:

$$\Psi(1, 2) = \phi_i(1)\phi_j(2). \quad (97)$$

The eigenproblem for such a function gives the eigenvalue that is simply the sum of the one-electron eigenvalues:

$$\begin{aligned} [\widehat{h}(1) + \widehat{h}(2)] \phi_i(1)\phi_j(2) &= \widehat{h}(1)\phi_i(1)\phi_j(2) + \widehat{h}(2)\phi_i(1)\phi_j(2) \\ &= [\widehat{h}(1)\phi_i(1)] \phi_j(2) + [\widehat{h}(2)\phi_j(2)] \phi_i(1) \\ &= [\epsilon_i \phi_i(1)] \phi_j(2) + [\epsilon_j \phi_j(2)] \phi_i(1) \\ &= [\epsilon_i + \epsilon_j] \phi_i(1)\phi_j(2). \end{aligned} \quad (98)$$

However, here the antisymmetry requirement should also be taken into account. The many-electron function must change the sign with respect to the exchange of labels of any two electrons:

$$\Psi(1, 2) = -\Psi(2, 1). \quad (99)$$

The product (Eq. 97) is not antisymmetric; the interchange of the electron labels leads to

$$\Psi(2, 1) = \phi_i(2)\phi_j(1) \neq -\Psi(1, 2), \quad (100)$$

and the result is a function different from the original $\Psi(1, 2)$. But another function,

$$\Psi(1, 2) \sim [\phi_i(1)\phi_j(2) - \phi_i(2)\phi_j(1)], \quad (101)$$

satisfies the antisymmetry condition.

Moreover, the wave function $\Psi(1, 2)$ has to be normalized. This can be achieved by calculating the following (overlap) integral:

$$\begin{aligned}
\langle \Psi(1, 2) | \Psi(1, 2) \rangle &= \langle \mathcal{N}(\phi_i(1)\phi_j(2) - \phi_j(1)\phi_i(2)) | \mathcal{N}(\phi_i(1)\phi_j(2) - \phi_j(1)\phi_i(2)) \rangle \\
&= \mathcal{N}^2 \left(\langle \phi_i(1)\phi_j(2) | \phi_i(1)\phi_j(2) \rangle - \langle \phi_i(1)\phi_j(2) | \phi_j(1)\phi_i(2) \rangle \right. \\
&\quad \left. - \langle \phi_j(1)\phi_i(2) | \phi_i(1)\phi_j(2) \rangle + \langle \phi_j(1)\phi_i(2) | \phi_j(1)\phi_i(2) \rangle \right).
\end{aligned} \tag{102}$$

The spin-orbitals of electrons 1 and 2 are mutually independent; thus,

$$\langle \phi_i(1)\phi_j(2) | \phi_k(1)\phi_l(2) \rangle = \langle \phi_i(1) | \phi_k(1) \rangle \langle \phi_j(2) | \phi_l(2) \rangle = \delta_{ik}\delta_{jl}, \tag{103}$$

where $\langle \phi_i(1) | \phi_k(1) \rangle = \delta_{ik}$ arises from the fact that the eigenfunctions of the Hermitian operator $\widehat{h}(1)$ form the orthonormal set. Hence, in Eq. 102 only the first and last integral will be nonvanishing, and, finally,

$$\langle \Psi(1, 2) | \Psi(1, 2) \rangle = 2\mathcal{N}^2 = 1, \tag{104}$$

and, therefore, the normalization constant \mathcal{N} must equal $1/\sqrt{2}$. In order to fulfill the normalization and the antisymmetry request, the trial wave function can be written as

$$\Psi(1, 2) = \frac{1}{\sqrt{2}} (\phi_i(1)\phi_j(2) - \phi_j(1)\phi_i(2)) = \frac{1}{\sqrt{2}} \begin{vmatrix} \phi_i(1) & \phi_i(2) \\ \phi_j(1) & \phi_j(2) \end{vmatrix}. \tag{105}$$

Thereafter, the expectation value of the Hamiltonian can be calculated using the same tricks as in Eq. 102³:

$$\begin{aligned}
\langle \widehat{H}(1, 2) \rangle_{\Psi(1,2)} &= \langle \Psi(1, 2) | \widehat{H}(1, 2) | \Psi(1, 2) \rangle \\
&= \langle \phi_i(1) | \widehat{h}(1) | \phi_i(1) \rangle + \langle \phi_j(2) | \widehat{h}(2) | \phi_j(2) \rangle \\
&\quad + \langle \phi_i(1)\phi_j(2) | \frac{1}{r_{12}} | \phi_i(1)\phi_j(2) \rangle \\
&\quad - \langle \phi_i(1)\phi_j(2) | \frac{1}{r_{12}} | \phi_j(1)\phi_i(2) \rangle.
\end{aligned} \tag{106}$$

³Since the value of the integral does not depend on the name of the variable, the obvious equality $\int_a^b f(x)dx = \int_a^b f(y)dy$ in the above case takes the form

$$\langle \phi_i(1)\phi_j(2) | \frac{1}{r_{12}} | \phi_j(1)\phi_i(2) \rangle = \langle \phi_i(2)\phi_j(1) | \frac{1}{r_{12}} | \phi_j(2)\phi_i(1) \rangle.$$

The only difference with respect to the overlap integral is that the integrals containing $1/r_{12}$ cannot be separated.

In the spirit of perturbation calculus, the expectation value (Eq. 106) can be treated as the energy corrected up to the first order of perturbation expansion. According to the variational principle, this integral is an upper bound of the exact energy of the two-electron system under consideration. The recipe for correcting the trial function is given by the perturbation theory: Apply functions corresponding to the remaining states of the unperturbed system in the expansion. In other words, in order to improve the wave function, the expansion built from products of the remaining states of the systems should be used.

Let us summarize. The many-electron function of the system can be approximately written as the antisymmetrized product of the one-electron functions being the solutions for the one-electron eigenvalue problem. This is the idea of the popular one-electron approximation (Atkins and Friedman 2005; Lowe and Peterson 2005; McQuarrie and Simon 1997; Ratner and Schatz 2000). In the N -electron case, one obtains the trial function as the antisymmetrized product of N one-electron functions that can be written in the form of the Slater determinant:

$$\Psi(1, 2, \dots, N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \phi_1(1) & \phi_1(2) & \dots & \phi_1(N) \\ \phi_2(1) & \phi_2(2) & \dots & \phi_2(N) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_N(1) & \phi_N(2) & \dots & \phi_N(N) \end{vmatrix}. \quad (107)$$

$1/\sqrt{N!}$ factor ensures the normalization of the wave function. Often, instead of writing the whole determinant, only the diagonal is written down:

$$\begin{vmatrix} \phi_1(1) & \phi_1(2) & \dots & \phi_1(N) \\ \phi_1(2) & \phi_2(2) & \dots & \phi_2(N) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_N(1) & \phi_N(2) & \dots & \phi_N(N) \end{vmatrix} = |\phi_1(1)\phi_2(2) \dots \phi_N(N)|. \quad (108)$$

It is important to realize that the many-electron function in the form of the Slater determinant is only an approximation. Intuition tells us that describing the many-electron system using only one-electron functions cannot be exact. One needs to be aware of the fact that such an approach causes the loss of some information included in the sought wave function. In particular, the one-electron function cannot “see” another electron; therefore, the terms coupling the mutual electron positions are missing in the one-electron approximation.

For example, consider the two-electron function (Piela 2007)

$$F(1, 2) \sim (e^{-ar_1-br_2-cr_{12}} - e^{-ar_2-br_1-cr_{12}}), \quad (109)$$

where r_1, r_2 are the electron–nucleus distances, r_{12} denotes the distance between two electrons, and a, b , and c stand for the coefficients. $F(1, 2)$ contains the factor c correlating the electron motions. In the one-electron approach, such a function would be approximated by the antisymmetrized function

$$f(1, 2) \sim (e^{-ar_1-br_2} - e^{-ar_2-br_1}), \quad (110)$$

with total neglect of this correlation. From the practical point of view, this means that the trial function written as the single determinant does not allow reproduction of the exact electronic energy, even when using the best possible one-electron functions for its construction.

The question arises: Why use such an approach, knowing from the very beginning that it is bad? The answer is simple. First, including the electron correlation in the wave function is very expensive. For two electrons, one additional term appears; for three, three terms; for four, six. In general, for N electrons there are $N(N - 1)/2$ terms (the triangle of the $N \times N$ matrix without the diagonal elements). Therefore, the number of coefficients describing the electron correlation is much bigger than the number of one-electron terms. Calculating even the one-electron coefficients is very time-consuming, and moreover, calculating the overlap integrals and Hamiltonian matrix elements becomes prohibitively complicated with the correlated functions. Second, perturbation theory provides ways of improving the results. Expansion built on a higher number of determinants will lead to better energy. Third, a chemist does not usually need exact data but only an appropriate accuracy (furnishing a house does not require caliper measurements, just a quick glance to estimate the size of the door and the furniture).

Therefore, let us stick to the one-electron approximation. The next section will explain how to find the best possible spin-orbitals.

Hartree–Fock Method: It Is Not That Sophisticated

Now we are prepared to concentrate on the methods for solving the electron equation. As you will see, they are only an extension of the already-discussed techniques. We start with the fundamental Hartree–Fock method.

The main goal of this approximation is to find the spin-orbitals applied for construction of the Slater determinant that will best reproduce the exact wave function. We again begin our considerations with the two-electron system. The problem is that the operator $\hat{h}(i)$ does not contain the part arising from the potential of the second electron, and an operator responsible for this missing part must be found. What we do know is that such an interaction is included in the two-electron part of Eq. 106:

$$\left\langle \frac{1}{r_{12}} \right\rangle_{\Psi(1,2)} = \left\langle \phi_1(1)\phi_j(2) \left| \frac{1}{r_{12}} \right| \phi_i(1)\phi_j(2) \right\rangle - \left\langle \phi_i(1)\phi_j(2) \left| \frac{1}{r_{12}} \right| \phi_j(1)\phi_i(2) \right\rangle. \quad (111)$$

Unfortunately, Eq. 111 contains the integrals that cannot be exactly separated into a product of the simpler integrals depending only on the coordinates of one electron. What we can propose is rewriting this equation in the following form:

$$\left\langle \frac{1}{r_{12}} \right\rangle_{\Psi(1,2)} = \langle \phi_i(1) | \widehat{v}(1) | \phi_i(1) \rangle + \langle \phi_j(2) | \widehat{v}(2) | \phi_j(2) \rangle + \text{the rest}, \quad (112)$$

where $\widehat{v}(1)$ and $\widehat{v}(2)$ are introduced to extract only the terms depending on the coordinates of a single electron from Eq. 111 and “the rest” is what remains from $\left\langle \frac{1}{r_{12}} \right\rangle_{\Psi(1,2)}$ after this extraction.

If such extraction were possible, the operators $\widehat{v}(1)$ and $\widehat{v}(2)$ could be applied to improve our one-electron operators, which leads to the following equations:

$$\left[\widehat{h}(1) + \widehat{v}(1) \right] \phi_i(1) = \epsilon_i \phi_i(1), \quad (113)$$

$$\left[\widehat{h}(2) + \widehat{v}(2) \right] \phi_j(2) = \epsilon_j \phi_j(2). \quad (114)$$

The solutions of such equations (spin-orbitals ϕ_i, ϕ_j) can be used to build up the Slater determinant. They ensure better approximation than Eqs. 95 and 96, since they somehow provide for the influence of the second electron.

Therefore, our goal now is to utilize Eq. 106 (treating ϕ_i and ϕ_j) as known functions) to find the best form of the operators $\widehat{v}(1)$ and $\widehat{v}(2)$. Adding zero, written as

$$0 = \left\langle \frac{1}{r_{12}} \right\rangle_{\Psi(1,2)} - \left\langle \frac{1}{r_{12}} \right\rangle_{\Psi(1,2)}, \quad (115)$$

to Eq. 111 allows one to ascribe, for instance,

$$\begin{aligned} \langle \phi_i(1) | \widehat{v}(1) | \phi_i(1) \rangle &= \left\langle \phi_i(1) \phi_j(2) \left| \frac{1}{r_{12}} \right| \phi_i(1) \phi_j(2) - \phi_j(1) \phi_i(2) \right\rangle, \\ \langle \phi_j(2) | \widehat{v}(2) | \phi_j(2) \rangle &= \left\langle \phi_i(1) \phi_j(2) \left| \frac{1}{r_{12}} \right| \phi_i(1) \phi_j(2) - \phi_j(1) \phi_i(2) \right\rangle, \\ \text{the rest} &= - \left\langle \phi_i(1) \phi_j(2) \left| \frac{1}{r_{12}} \right| \phi_i(1) \phi_j(2) - \phi_j(1) \phi_i(2) \right\rangle. \end{aligned}$$

Consider in more detail the integral containing $\widehat{v}(1)$. We want it to be expressed in such a way that only the coordinates of the electron labeled by 1 are explicitly written under the integral:

$$\begin{aligned} \langle \phi_i(1) | \widehat{v}(1) | \phi_i(1) \rangle &= \left\langle \phi_i(1) \phi_j(2) \left| \frac{1}{r_{12}} \right| \phi_i(1) \phi_j(2) - \phi_j(1) \phi_i(2) \right\rangle \\ &= \int \int \phi_i^*(\tau_1) \phi_j^*(\tau_2) \frac{1}{r_{12}} \phi_i(\tau_1) \phi_j(\tau_2) d\tau_1 d\tau_2 \\ &\quad - \int \int \phi_i^*(\tau_1) \phi_j^*(\tau_2) \frac{1}{r_{12}} \phi_j(\tau_1) \phi_i(\tau_2) d\tau_1 d\tau_2 \\ &= \int \phi_i^*(\tau_1) \left(\int \phi_j^*(\tau_2) \frac{1}{r_{12}} \phi_j(\tau_2) d\tau_2 \right) \phi_i(\tau_1) d\tau_1 \end{aligned}$$

$$\begin{aligned}
& - \int \phi_i^* (\tau_1) \left(\int \phi_j^* (\tau_2) \frac{1}{r_{12}} \phi_i (\tau_2) d\tau_2 \right) \phi_j (\tau_1) d\tau_1 \\
& = \int \phi_i^* (\tau_1) \left(\widehat{J}_j(1) \phi_i (\tau_1) \right) d\tau_1 \\
& \quad - \int \phi_j^* (\tau_1) \left(\widehat{K}_j(1) \phi_i (\tau_1) \right) d\tau_1 \tag{116} \\
& = \left\langle \phi_i(1) \left| \widehat{J}_j(1) \right| \phi_i(1) \right\rangle - \left\langle \phi_i(1) \left| \widehat{K}_j(1) \right| \phi_i(1) \right\rangle \\
& = \left\langle \phi_i(1) \left| \widehat{J}_j(1) - \widehat{K}_j(1) \right| \phi_i(1) \right\rangle,
\end{aligned}$$

where two operators were defined:

$$\widehat{J}_j(1) \phi_i(1) = \left(\int \phi_j^* (\tau_2) \frac{1}{r_{12}} \phi_j (\tau_2) d\tau_2 \right) \phi_i(1), \tag{117}$$

$$\widehat{K}_j(1) \phi_i(1) = \left(\int \phi_j^* (\tau_2) \frac{1}{r_{12}} \phi_i (\tau_2) d\tau_2 \right) \phi_j(1). \tag{118}$$

Despite a slightly different way of defining, they are still ordinary operators. An operator is a function of a function. The operators $\widehat{J}_j(1)$ and $\widehat{K}_j(1)$ act on the function $\phi_i(1)$, producing another function, as was written in Eq. 8. The operator $\widehat{J}_j(1)$ acting on $\phi_i(1)$ transforms it into the same function:

$$\phi_i(1) \xrightarrow{\widehat{J}_j(1)} \left(\int \phi_j^* (\tau_2) \frac{1}{r_{12}} \phi_j (\tau_2) d\tau_2 \right) \phi_i(1), \tag{119}$$

and the operator $\widehat{K}_j(1)$ produces $\phi_j(1)$:

$$\phi_i(1) \xrightarrow{\widehat{K}_j(1)} \left(\int \phi_j^* (\tau_2) \frac{1}{r_{12}} \phi_i (\tau_2) d\tau_2 \right) \phi_j(1). \tag{120}$$

The expressions on the right-hand side of the arrows are some functions of a dependent variable τ_1 (the integration eliminates the dependence on τ_2 ; however, its result is not a number but a function depending on τ_1). The interpretation of the $\widehat{J}_j(1)$ and $\widehat{K}_j(1)$ operators by ascribing them to observables is not straightforward. These operators appear in the equations when we try to write the interaction between two electrons as the average value of the one-electron operator calculated with the one-electron function. However, in fact, they do appear as the difference:

$$\widehat{v}_i^{\text{HF}}(1) = \widehat{J}_j(1) - \widehat{K}_j(1). \tag{121}$$

And the physical sense should be sought in this difference. Here, $\widehat{v}_i^{\text{HF}}(1)$ is an operator of the average interaction of the electron labeled as 1, described by a spin-orbital ϕ_i with the second electron characterized by $\phi_j(2)$. It should be noticed that the potential $\widehat{v}_i^{\text{HF}}(1)$ depends on $\phi_j(2)$, since this spin-orbital is necessary to define $\widehat{J}_j(1)$ and $\widehat{K}_j(1)$ operators.

Similarly,

$$\widehat{v}_j^{\text{HF}}(2) = \widehat{J}_i(2) - \widehat{K}_i(2), \quad (122)$$

where the action of operators $\widehat{J}_j(1)$ and $\widehat{K}_j(1)$ on the function $\phi_j(2)$ is defined by

$$\widehat{J}_i(2)\phi_j(2) = \left(\int \phi_i^*(\tau_1) \frac{1}{r_{12}} \phi_i(\tau_1) d\tau_1 \right) \phi_j(2) \quad (123)$$

$$\widehat{K}_i(2)\phi_j(2) = \left(\int \phi_i^*(\tau_1) \frac{1}{r_{12}} \phi_j(\tau_1) d\tau_1 \right) \phi_i(2). \quad (124)$$

Here, one more important circumstance should be mentioned. The electronic Hamiltonian does not depend on spin. However, the electronic wave function is spin dependent. During the construction of the HF equations, this dependence is introduced to the operators \widehat{K} since they depend on two different spin-orbitals.

Let us briefly review. We want to have a one-electron operator that includes the interaction between the electrons in some averaged way. Such an operator must depend on the function describing the motion of the second, adjacent electron. Therefore, for the two-electron case, two coupled equations must be solved (Eqs. 113 and 114). The word “coupled” that distinguishes this set of equations from Eqs. 95 and 96 is crucial. Denoting

$$\widehat{f}_i(1) = \widehat{h}(1) + \widehat{v}_i^{\text{HF}}(1), \quad (125)$$

one can rewrite the above equations as

$$\widehat{f}_k(1)\phi_k(1) = \epsilon_k\phi_k(1), \quad \text{for } k = i, j. \quad (126)$$

$\widehat{f}_i(1)$ is called the Fock operator and Eq. 114 gives the Hartree–Fock equations.

It should be noted that the label of electrons determines only the name of the integration variables, and the result of the integration does not depend on the name of the variable. Therefore, what is really important is the label of the spin-orbital. It will be even more pronounced in the N -electron case, when the electron labels are applied only to show that the operator acts on one or two electron coordinates. The form of the N -electron wave function depends only on the spin-orbitals and not on the electron labels; they are just the integration variables.

The most popular way of solving Eq. 126 is the iterative procedure. It starts from the guessed or chosen spin-orbitals $\phi_i(1)$ and $\phi_j(2)$ applied to construct the potentials $\widehat{v}_i^{\text{HF}}(1)$ and $\widehat{v}_j^{\text{HF}}(2)$. Next, the obtained potentials are substituted to Eq. 125 and solutions of Eq. 126 give the improved form of the $\phi_i(1)$ and $\phi_j(2)$ spin-orbitals. These, on the other hand, are treated as the starting point again, and the whole procedure is repeated until the starting and final orbitals of the given iteration

do not differ much. This technique is called self-consistent field (SCF). Often the abbreviations HF for Hartree–Fock method and SCF are used interchangeably. They can also be joined together as SCF–HF, denoting the self-consistent way of solving Hartree–Fock equations.

Let us assume that the spin-orbitals are already known. Concentrate on the calculations of the average value of the Hamiltonian with the determinant build of these spin-orbitals using the operators defined previously. Writing down the two-electron part:

$$\begin{aligned} \left\langle \frac{1}{r_{12}} \right\rangle_{\Psi(1,2)} &= \left\langle \phi_i(1) \left| \widehat{J}_j(1) - \widehat{K}_j(1) \right| \phi_i(1) \right\rangle \\ &= \left\langle \phi_j(2) \left| \widehat{J}_i(2) - \widehat{K}_i(2) \right| \phi_j(2) \right\rangle, \end{aligned}$$

one obtains the integrals that can be denoted by

$$\begin{aligned} J_{ij} &= \left\langle \phi_i(1) \left| \widehat{J}_j(1) \right| \phi_i(1) \right\rangle, \\ K_{ij} &= \left\langle \phi_i(1) \left| \widehat{K}_j(1) \right| \phi_i(1) \right\rangle. \end{aligned}$$

With this notation, the expression (Eq. 106) takes the following form:

$$\left\langle \widehat{H}(1, 2) \right\rangle_{\Psi(1,2)} = \left\langle \phi_i(1) \left| \widehat{f}(1) \right| \phi_i(1) \right\rangle + \left\langle \phi_j(2) \left| \widehat{f}(2) \right| \phi_j(2) \right\rangle - (J_{ij} - K_{ij}), \quad (127)$$

or if one wants to apply the spin-orbital energies ϵ calculated earlier:

$$\left\langle \widehat{H}(1, 2) \right\rangle_{\Psi(1,2)} = \epsilon_i + \epsilon_j - (J_{ij} - K_{ij}). \quad (128)$$

Now it is time to generalize these considerations into the N -electron case. The recipe for this transformation was given previously (Eqs. 91, 92, and 93). For the system of N electrons, the set of N -coupled equations of the form

$$\widehat{f}_i(1)\phi_i(1) = \epsilon_i\phi_i(1), \quad \text{for } i = 1, \dots, N \quad (129)$$

must be solved. Here,

$$f_i(1) = h(1) + \sum_{j \neq i}^N \left(\widehat{J}_j(1) - \widehat{K}_j(1) \right). \quad (130)$$

The summation in the expression for the one-electron Fock operator arises from the fact that now the given electron described by the spin-orbital i interacts with $(N - 1)$ electrons in the states determined by the remaining spin-orbitals.

In searching for the ground state energy, one is interested in the lowest possible energy. Therefore, the functions of choice are the spin-orbitals corresponding to the lowest values of ϵ . Such a set of spin-orbitals, called occupied, is opposite to any other solutions of the Fock equations corresponding to higher energies. These are known as virtual (unoccupied) orbitals.

Finally, the average value of the Hamiltonian can be written as

$$\left\langle \widehat{H}(1, 2, \dots, N) \right\rangle_{\Psi^{(0)}(N)} = \sum_{i=1}^N \epsilon_i - \sum_{i=1}^N \sum_{j>i}^N (J_{ij} - K_{ij}) = E_0^{\text{HF}}$$

It should be emphasized that the energy estimated in this manner is not the simple sum of the orbital energies. If the terms $\sum_{i=1}^N \sum_{j>i}^N (J_{ij} - K_{ij})$ are neglected, the double summation of the electron–electron interaction would take place (Atkins and Friedman 2005; Cramer 2004; Jensen 2006; Levine 2008; Lowe and Peterson 2005; McQuarrie and Simon 1997; Piela 2007; Ratner and Schatz 2000; Roos and Widmark 2002; Szabo and Ostlund 1996).

Møller–Plesset Perturbation Theory: HF Is Just the Beginning

From here forward, we will treat the Hartree–Fock function as the basis for the further investigations and denote it as $\Psi_0^{(0)}$, where the subscript 0 indicates the ground state and the superscript (0) is the reference function. We will also omit the explicit writing of the dependence of the Hamiltonian and the wave function on the coordinates of N electrons. As a consequence,

$$\left\langle \Psi_0^{(0)} \left| \widehat{H} \right| \Psi_0^{(0)} \right\rangle = E_0^{\text{HF}}. \quad (131)$$

One of the possible ways of improving the HF results is the application of the perturbation theory (Atkins and Friedman 2005; Cramer 2004; Jensen 2006; Levine 2008; McQuarrie and Simon 1997; Piela 2007; Ratner and Schatz 2000; Szabo and Ostlund 1996). The Hamiltonian (Eq. 89) can be partitioned into the unperturbed part and the perturbation in the following manner:

$$\widehat{H}^0 = \sum_{i=1}^N \widehat{f}(i), \quad \widehat{H}^1 = \sum_{i=1}^N \left(\sum_{j>i}^N \frac{1}{r_{ij}} - \widehat{v}^{\text{HF}}(i) \right) \quad (132)$$

(compare Eq. 125). Up to the first order in the perturbation theory, the energy is

$$E_0^{\text{HF}} = E_0^{(0)} + E_0^{(1)} = \left\langle \Psi_0^{(0)} \left| \widehat{H}^0 + \widehat{H}^1 \right| \Psi_0^{(0)} \right\rangle = \left\langle \widehat{H} \right\rangle_{\Psi_0^{(0)}}. \quad (133)$$

The correction to the HF energy appears in the second order:

$$E_0^{(2)} = \sum_{k \neq 0}^{\infty} \frac{\left| \langle \Psi_0^{(0)} | \widehat{H}^1 | \Psi_k^{(0)} \rangle \right|^2}{E_0^{(0)} - E_k^{(0)}}. \quad (134)$$

During calculations, the expansion of the spin-orbitals in the finite basis set is applied. This allows identification of only the finite number of spin-orbitals. But the number of all possible $\Psi_k^{(0)}$ functions can still be horrifyingly large. In practice, this is equivalent to the finite but long expansion in the above summation. In quantum chemistry we are interested in the best quality results with moderate expenses and are continuously searching for more economical methods.

Careful examination of the second-order energy correction $E_0^{(2)}$ shows that a significant number of its terms do not contribute to the final result. Some work invested in the manipulation of expressions allows one not only to learn the basic computational apparatus but also to save a lot of effort.

In order to proceed comfortably, the notation should again be simplified. Let us drop the superscript (0) denoting the zeroth-order functions (other functions will not appear in our considerations). Additionally, we omit the subscript k from the determinants containing the virtual spin-orbitals. In return, we explicitly specify the pattern of spin-orbital exchange. Using Eq. 108, the Slater determinant can be written as

$$\left| \Psi_0 \right\rangle = |\phi_1(1)\phi_2(2)\phi_3(3)\phi_4(4) \dots \phi_N(N)\rangle. \quad (135)$$

The numbers of the occupied spin-orbitals vary from 1 to N . Thus, the virtual spin-orbitals will be labeled starting from $N + 1$ onward. Now consider the example of the determinant in which the occupied spin-orbital ϕ_3 is exchanged for the virtual one, ϕ_{N+8} . The new determinant can be written as

$$\left| \Psi_3^{N+8} \right\rangle = \left| \phi_1(1)\phi_2(2)\phi_{N+8}(3)\phi_4(4) \dots \phi_N(N) \right\rangle, \quad (136)$$

where the subscript in Ψ_3^{N+8} denotes the occupied orbital that is exchanged and the superscript denotes the virtual one that takes its place. To generalize, one could denote the occupied orbitals building the Ψ_0 function by first alphabet letters a, b, c, d, \dots and the virtuals by p, q, r, s, \dots . Therefore, Ψ_3^{N+8} can be written as Ψ_a^p , where $a = 3$ and $p = N + 8$. Likewise, any determinant can be represented.

Thereafter, the influence of the function choice on the values of the integrals appearing in the electronic energy calculations can be investigated. Similarly, as in the case of Hamiltonian, the integrals can also be divided into two groups with one-electron operator

$$\widehat{o}_1 = \sum_{i=1}^N \widehat{o}(i), \quad (137)$$

where in place of $\hat{o}(i)$ operators $\hat{h}, (i)$ or $\hat{f}(i)$ will be used, and with two-electron operator,

$$\hat{O}_2 = \sum_i \sum_{j>i} \left(\frac{1}{r_{ij}} \right). \quad (138)$$

To get a full picture, we should analyze the following types of integrals:

$$\langle \Psi_0 | \hat{o}_1 | \Psi_a^p \rangle \quad \text{and} \quad \langle \Psi_0 | \hat{O}_2 | \Psi_a^p \rangle, \quad (139)$$

$$\langle \Psi_0 | \hat{o}_1 | \Psi_{ab}^{pq} \rangle \quad \text{and} \quad \langle \Psi_0 | \hat{O}_2 | \Psi_{ab}^{pq} \rangle, \quad (140)$$

$$\langle \Psi_0 | \hat{o}_1 | \Psi_{abc}^{pqr} \rangle \quad \text{and} \quad \langle \Psi_0 | \hat{O}_2 | \Psi_{abc}^{pqr} \rangle. \quad (141)$$

Any other will be equal to zero (Levine 2008; Szabo and Ostlund 1996).

Recall that the determinant of the $N \times N$ matrix can be represented as the sum of the $N!$ products of the matrix elements. According to Laplace's formula, a determinant can be expanded along a row or a column. Thus, calculation of the $\langle \Psi_0 | \hat{o}_1 | \Psi_a^p \rangle$ integrals, when the occupied orbital a in Ψ_0 has been exchanged with the virtual orbital p in Ψ_a^p , can be performed by expanding the determinant Ψ_0 along the a -th row and the determinant Ψ_a^p along the p -th row:

$$\Psi_0 = \frac{1}{\sqrt{N!}} \sum_{i=1}^N \phi_a(i) C_{ai}, \quad (142)$$

$$\Psi_a^p = \frac{1}{\sqrt{N!}} \sum_{i=1}^N \phi_p(i) C_{pi}. \quad (143)$$

The cofactors C can be perceived as the $(N - 1)$ -electron determinants that have been obtained from Ψ_0 and Ψ_a^p via the elimination of the a -th and p -th spin-orbitals, respectively. Thus, after elimination of what is different in the two determinants, we get both cofactors equal to each other. Hence,

$$\begin{aligned} \langle \Psi_0 | \hat{o}_1 | \Psi_a^p \rangle &= \frac{1}{N!} \left\langle \sum_{i=1}^N \phi_a(i) C_{ai} \left| \sum_{j=1}^N \hat{o}(j) \right| \sum_{k=1}^N \phi_p(k) C_{pi} \right\rangle \\ &= \frac{1}{N!} \sum_{i=1}^N \langle \phi_a(i) | \hat{o}(i) | \phi_p(i) \rangle \langle C_{ai} | C_{pi} \rangle \\ &= \frac{1}{N} \sum_{i=1}^N \langle \phi_p(i) | \hat{o}(i) | \phi_p(i) \rangle \\ &= \langle \phi_a(1) | \hat{o}(1) | \phi_p(1) \rangle, \end{aligned} \quad (144)$$

where

$$\langle C_{ai} | C_{pi} \rangle = (N - 1)!, \quad (145)$$

$$\sum_{i=1}^N \langle \phi_a(i) | \widehat{\partial}(i) | \phi_a(i) \rangle = N \langle \phi_a(1) | \widehat{\partial}(1) | \phi_a(1) \rangle \quad (146)$$

was applied. If the cofactors are not the same – this would be the case when the determinants differ from two or more spin-orbitals – the corresponding overlap matrix is equal to zero according to the orthogonality condition. Thus,

$$\begin{aligned} \langle \Psi_0 | \widehat{\partial}_1 | \Psi_a^p \rangle &= \langle \phi_a(1) | \widehat{\partial}(1) | \phi_p(1) \rangle, \\ \langle \Psi_0 | \widehat{\partial}_1 | \Psi_{ab}^{pq} \rangle &= 0, \\ \langle \Psi_0 | \widehat{\partial}_1 | \Psi_{abc}^{pqr} \rangle &= 0. \end{aligned} \quad (147)$$

Similar (but a little more time-consuming) considerations for the two-electron operators lead to the following expressions:

$$\begin{aligned} \langle \Psi_0 | \widehat{\mathcal{O}}_2 | \Psi_a^p \rangle &= \sum_{i=1}^N \left(\left\langle \phi_a(1) \phi_i(2) \left| \frac{1}{r_{12}} \right| \phi_p(1) \phi_i(2) \right\rangle \right. \\ &\quad \left. - \left\langle \phi_a(1) \phi_i(2) \left| \frac{1}{r_{12}} \right| \phi_i(1) \phi_p(2) \right\rangle \right) \\ \langle \Psi_0 | \widehat{\mathcal{O}}_2 | \Psi_{ab}^{pq} \rangle &= \left\langle \phi_a(1) \phi_b(2) \left| \frac{1}{r_{12}} \right| \phi_p(1) \phi_q(2) \right\rangle \\ &\quad - \left\langle \phi_a(1) \phi_b(2) \left| \frac{1}{r_{12}} \right| \phi_q(1) \phi_p(2) \right\rangle \\ \langle \Psi_0 | \widehat{\mathcal{O}}_2 | \Psi_{abc}^{pqr} \rangle &= 0. \end{aligned} \quad (148)$$

In the first contact with these equations, one can have the feeling that something is lost here. We start from the integrals with the N -electron functions and we finish with the integral of only the electrons labeled by 1 and 2. What happened to the rest? Again, it should be emphasized here that the electron labels symbolize only the integration variables. What matters is the functions of these variables, namely, spin-orbitals. Thus, in the integral with the one-electron operator, the electron label 1 means that the integration is performed only over the variables of one electron. Likewise, the two-electron operator integral depends on the variables of two electrons, which is symbolized by two labels: 1 and 2.

In the integration of one-electron expressions, the electron label can be omitted without any harm: $\langle \phi_a(1) | \widehat{\partial}_1(1) | \phi_a(1) \rangle$. Likewise, in the two-electron case, we can declare that the spin-orbitals are written in the given order. So,

$$\left\langle \phi_x(1) \phi_y(2) \left| \frac{1}{r_{12}} \right| \phi_v(1) \phi_z(2) \right\rangle = \left\langle \phi_x \phi_y \left| \frac{1}{r_{12}} \right| \phi_v \phi_z \right\rangle. \quad (149)$$

Moreover, now there is no reason to explicitly write the ϕ symbol. Thereby, the next simplification of the notation is obvious:

$$\begin{aligned}\langle \phi_a | \widehat{o}_1 | \phi_a \rangle &= \langle a | \widehat{o}_1 | a \rangle, \\ \langle \phi_x \phi_y | \frac{1}{r_{12}} | \phi_v \phi_z \rangle &= \langle xy | \frac{1}{r_{12}} | vz \rangle,\end{aligned}\tag{150}$$

and finally, since the combination of the integrals $\langle xy | 1/r_{12} | vz \rangle - \langle xy | 1/r_{12} | zv \rangle$ appears frequently, the following symbol is introduced:

$$\langle xy || zv \rangle = \langle xy | \frac{1}{r_{12}} | vz \rangle - \langle xy | \frac{1}{r_{12}} | zv \rangle.\tag{151}$$

Now Eqs. 147 and 148 can be rewritten as

$$\begin{aligned}\langle \Psi_0 | \widehat{o}_1 | \Psi_a^p \rangle &= \langle a | \widehat{o} | p \rangle, \\ \langle \Psi_0 | \widehat{o}_1 | \Psi_{ab}^{pq} \rangle &= 0, \\ \langle \Psi_0 | \widehat{o}_1 | \Psi_{abc}^{pqr} \rangle &= 0, \\ \langle \Psi_0 | \widehat{O}_2 | \Psi_a^p \rangle &= \sum_{i=1}^N \langle ai || pi \rangle, \\ \langle \Psi_0 | \widehat{O}_2 | \Psi_{ab}^{pq} \rangle &= \langle ab || pq \rangle, \\ \langle \Psi_0 | \widehat{O}_2 | \Psi_{abc}^{pqr} \rangle &= 0.\end{aligned}\tag{152}$$

These simple equations, known as the Slater rules, allow for the following general remark: If the one-electron operator is integrated with functions that differ by more than one spin-orbital, the corresponding integral vanishes. Similarly, the result is zero for the integration of two-electron operators with functions differing by more than two spin-orbitals. Hitherto, only the integrals with Ψ_0 were considered. However, it is easy to notice that functions Ψ_{ab}^{pq} and Ψ_{abc}^{pqr} differ by only one exchange (spin-orbital $\phi_c \rightarrow \phi_r$), etc. Therefore, the above considerations can be also applied in any other cases.

In this abundance of equations, our main goal cannot be lost: All these derivations were necessary to limit the types of the Ψ_k functions present in the MP2 energy expression. Now, with a recognition of the above Slater rules, one can safely neglect the integrals with the pairs of the Ψ_k functions differing with more than two spin-orbital exchanges. But this is not everything.

Let us consider the integral

$$\langle \Psi_0 | \widehat{H} | \Psi_a^p \rangle = \langle a | \widehat{h} | p \rangle + \sum_j \left(\langle aj | \frac{1}{r_{ij}} | pj \rangle - \langle aj | \frac{1}{r_{ij}} | jp \rangle \right).\tag{153}$$

Since

$$\left\langle aj \left| \frac{1}{r_{ij}} \right| pj \right\rangle = \left\langle a \left| \widehat{J}_j \right| p \right\rangle, \quad (154)$$

and

$$\left\langle aj \left| \frac{1}{r_{ij}} \right| jp \right\rangle = \left\langle a \left| \widehat{K}_j \right| p \right\rangle, \quad (155)$$

one gets

$$\left\langle \Psi_0 \left| \widehat{H} \right| \Psi_a^p \right\rangle = \left\langle a \left| \widehat{H}^0 \right| p \right\rangle, \quad (156)$$

which has to vanish, because

$$\left\langle a \left| \widehat{H}^0 \right| p \right\rangle = \left\langle a \left| \sum_{i=1}^N f(i) \right| p \right\rangle = \langle a | f | p \rangle = \epsilon_p \delta_{ap} = 0, \quad (157)$$

for it was assumed that $a \neq p$.

The obtained result,

$$\left\langle \Psi_0 \left| \widehat{H} \right| \Psi_a^p \right\rangle = 0, \quad (158)$$

is known as the Brillouin theorem. Applying it allows one also to show that

$$\left\langle \Psi_0 \left| \widehat{H}^1 \right| \Psi_a^p \right\rangle = 0. \quad (159)$$

Indeed, using

$$\widehat{H}^1 = \widehat{H} - \widehat{H}^0, \quad (160)$$

one obtains

$$\begin{aligned} \left\langle \Psi_0 \left| \widehat{H}^1 \right| \Psi_a^p \right\rangle &= \left\langle \Psi_0 \left| \widehat{H} \right| \Psi_a^p \right\rangle - \left\langle \Psi_0 \left| \widehat{H}^0 \right| \Psi_a^p \right\rangle \\ &= 0 - \left(\sum_{i=1}^N \epsilon_i \right) \left\langle \Psi_0 \left| \Psi_a^p \right\rangle \right. \\ &= 0. \end{aligned} \quad (161)$$

Now is the time for an important conclusion: In order to calculate the correction to the energy in MP2, the functions arising from Ψ_0 only by the exchange of precisely two spin-orbitals need to be applied, not more, not less. Life becomes easier when not calculating zero contributions in a complicated way.

Let us exploit the above knowledge to transform Eq. 134. The integral on the right-hand side will be calculated with the functions Ψ_{ab}^{pq} . Assuming $b > a$ and $q > p$ allows avoidance of the integration with the same functions. The upper limit for

the summations over a and b will be equal to N . For the remaining spin-orbitals, it should be ∞ ; however, in practice, the finite basis is applied and the upper limit will be determined by the basis set size. Let us look into the numerator of Eq. 134 carefully. Denoting the total Fock operator as $\widehat{F} = \sum_i \widehat{f}(i)$, one gets

$$\begin{aligned} \langle \Psi_0 | \widehat{H}^1 | \Psi_{ab}^{pq} \rangle &= \langle \Psi_0 | \widehat{H} | \Psi_{ab}^{pq} \rangle - \langle \Psi_0 | \widehat{F} | \Psi_{ab}^{pq} \rangle \\ &= \langle ab | \frac{1}{r_{ij}} | pq \rangle - \langle ab | \frac{1}{r_{ij}} | pq \rangle, \end{aligned} \quad (162)$$

since

$$\langle \Psi_0 | \widehat{F} | \Psi_{ab}^{pq} \rangle = 0 \quad (163)$$

(integration of the one-electron operator with the functions differing by two spin-orbitals; see Eq. 152).

Now consider the denominator of Eq. 134. The energy of the ground state is simply a sum of N lowest spin-orbital energies:

$$E_0^{(0)} = \sum_{i=1}^N \epsilon_i. \quad (164)$$

The energy of the zeroth order for the function Ψ_{ab}^{pq} is a similar sum with ϵ_a and ϵ_b replaced by ϵ_p and ϵ_q :

$$E_{(pq)}^{(0)} = E_0^{(0)} - \epsilon_a - \epsilon_b + \epsilon_p + \epsilon_q. \quad (165)$$

Thus, the final form of Eq. 134 is

$$E_0^{(2)} = \sum_{b>a} \sum_{q>p} \frac{|\langle ab || pq \rangle|^2}{\epsilon_p + \epsilon_q - \epsilon_a - \epsilon_b}. \quad (166)$$

Why was it worth our hard work? Not only for satisfaction. These derivations are necessary to understand the mechanisms employed in computational methods of quantum chemistry. With the Slater rules, it is straightforward to recognize the vanishing integrals. The double exchanges appear to be most important, since these are the only exchanges that give rise to the energy corrections in the MP2 method.

According to the perturbation calculus, the wave function corrected in the first order can be written as

$$\Psi \approx \Psi_0 + \sum_{k \neq 0} c_k \Psi_k, \quad (167)$$

where c_k is given by

$$c_k = \frac{\langle \Psi_0 | \widehat{H}^1 | \Psi_k \rangle}{E_0 - E_k}. \quad (168)$$

From the Slater rules, it is easy to estimate that the only nonvanishing terms will arise from the double spin-orbital exchange in the wave function. The intuitive statement is that the low-order corrections should have the larger impact and the above considerations lead to the most important accomplishment of this section: The largest contribution to the corrections of the Hartree–Fock function arises from the functions with the doubly exchanged spin-orbitals.

Beyond the HF Wave Function

Having done all this hard work, one can now sit comfortably in an armchair and think. The main goal of the quantum chemistry is to find the best possible description of the state of the system (the best possible wave function). The exact solutions of the Hamiltonian eigenproblem are unavailable and all we have are approximations. We have become used to approximations in everyday life. The important thing is to realize that the Hartree–Fock solutions can be improved. At the beginning of this chapter, various properties of operators were discussed. Among others, it was stated that the eigenfunctions of the Hermitian operators constitute the complete sets and any other function of the same variables can be represented by applying them. The one-electron spin-orbitals that are the eigenfunctions of the Fock operator are accessible. They form the complete set, but only for the one-electron functions. However, they can be applied to build up the N -electron determinants. The set of all possible determinants is also complete and, therefore, can be applied to express any N -electron function (Cramer 2004; Jensen 2006; Levine 2008; Lowe and Peterson 2005; Piela 2007; Ratner and Schatz 2000; Roos and Widmark 2002; Szabo and Ostlund 1996):

$$\Psi = c_0 \Psi_0 + \sum_{a,p} c_a^p \Psi_a^p + \sum_{a,b,p,q} c_{ab}^{pq} \Psi_{ab}^{pq} + \sum_{a,b,c,p,q,r} c_{abc}^{pqr} \Psi_{abc}^{pqr} + \dots \quad (169)$$

For this purpose, only the coefficients $c_0, c_a^p, c_{ab}^{pq}, c_{abc}^{pqr}, \dots$ need to be found. In the ideal case, all the summations would be infinite and the problem must be reduced. Still, instead of using the infinite expansions, the finite and relatively small number of terms can be sufficient. Moreover, solving the Hartree–Fock equations in the finite basis, one possesses only the finite number of orbitals that can be exchanged.

Anyway, it is instructive to see how large the number of terms in Eq. 169 can be. Consider the methane molecule CH_4 . Calculations with the minimal basis set (each orbital described by a single one-electron function; for carbon single functions for each of the orbitals, 1 s , 2 s , $2p_x$, $2p_y$, $2p_z$, and for hydrogen a single function for 1 s orbital) require 9 orbitals/18 spin-orbitals for the 10-electron system. From the probability theory, the number of combinations (K) of k elements from the n -element set can be calculated as

$$K = \binom{n}{k} = \frac{n!}{(n-k)!k!}. \quad (170)$$

In the case of methane, 10 electrons can be placed in 18 spin-orbitals on $\binom{18}{10} = 43,758$ ways. This is equivalent to the 43,758 terms in the expansion (Eq. 169). Impressive. And one needs to remember that the minimal basis set gives relatively bad Hartree–Fock solutions and is not recommended in *ab initio* calculations. However, increasing the basis set size causes the number of expansion terms to grow dramatically. For instance, in the case of the so-called double- ζ basis set (two functions per each orbital) for methane, one has 36 spin-orbitals, which makes 254,186,856 combinations! And double- ζ is still not much . . .

Therefore, it is necessary to find some way to reduce the size of the problem. The symmetry of the molecules can be applied here, and the fact that the chosen determinants (or their linear combinations) must be the given functions of the spin operators can be beneficial. Moreover, one would like to eliminate from the expansion the determinants that are not crucial for the quality of the wave function, and their neglect does not cause the deterioration of the description of the system (or causes only slight deterioration). In other words, only the determinants that have the significant contribution to the total energy must be chosen for the wave function construction. Let us begin with the classification of the determinants, taking into account the number of the spin-orbitals exchanged with respect to Ψ_0 . For this purpose, the averaged value of Hamiltonian calculated with Ψ will be useful. We can write the wave function expansion as

$$\Psi = c_0\Psi_0 + \mathbf{S}\mathbf{c}_S + \mathbf{D}\mathbf{c}_D + \mathbf{T}\mathbf{c}_T + \mathbf{Q}\mathbf{c}_Q + \dots \quad (171)$$

The symbols' meaning can be clearly deciphered by comparison with Eq. 169. \mathbf{S} denotes a vector build of the determinants constructed from Ψ_0 by single exchanges, and \mathbf{c}_S is a vector of coefficients corresponding to the functions in \mathbf{S} :

$$\mathbf{S}\mathbf{c}_S = \sum_a \sum_q c_a^q \Psi_a^q. \quad (172)$$

In other words, \mathbf{S} contains all the functions with the single exchanged spin-orbital. Similarly, \mathbf{D} would be the combination of the functions with double exchanges, \mathbf{T} with triple exchanges, and so forth. With such a notation, the function Ψ can be treated as the scalar product of the basis vectors $\Psi_0, \mathbf{S}, \mathbf{D}, \dots$, and the coefficient vector

$$|\Psi\rangle = [|\Psi_0\rangle, |\mathbf{S}\rangle, |\mathbf{D}\rangle, |\mathbf{T}\rangle, |\mathbf{Q}\rangle, \dots] \cdot \begin{bmatrix} c_0 \\ \mathbf{c}_S \\ \mathbf{c}_D \\ \mathbf{c}_T \\ \mathbf{c}_Q \\ \vdots \end{bmatrix}. \quad (173)$$

Using this notation, the Hamiltonian \widehat{H} of the system can be linked in an elegant way to a matrix \mathbf{H} . Let us apply this form of the wave function for the calculation of the Hamiltonian average value. To simplify the expressions, let us limit ourselves to the truncated expansion:

$$\Psi_{\text{SD}} = [\Psi_0 \mathbf{S} \mathbf{D}] \begin{bmatrix} c_0 \\ \mathbf{c}_\mathbf{S} \\ \mathbf{c}_\mathbf{D} \end{bmatrix} = c_0 \Psi_0 + \mathbf{S} \mathbf{c}_\mathbf{S} + \mathbf{D} \mathbf{c}_\mathbf{D}. \quad (174)$$

The average value of the Hamiltonian can now be written as

$$\langle H \rangle_{\Psi_{\text{SD}}} = \langle \Psi_{\text{SD}} | \widehat{H} | \Psi_{\text{SD}} \rangle = [c_0 \mathbf{c}_\mathbf{S} \mathbf{c}_\mathbf{D}]^\dagger \begin{bmatrix} \langle \Psi_0 | \\ \langle \mathbf{S} | \\ \langle \mathbf{D} | \end{bmatrix} \widehat{H} \begin{bmatrix} | \Psi_0 \rangle, | \mathbf{S} \rangle, | \mathbf{D} \rangle \end{bmatrix} \begin{bmatrix} c_0 \\ \mathbf{c}_\mathbf{S} \\ \mathbf{c}_\mathbf{D} \end{bmatrix}. \quad (175)$$

The vector multiplication leads to the following expression:

$$\begin{aligned} \langle H \rangle_{\Psi_{\text{SD}}} &= c_0^* \langle \Psi_0 | \widehat{H} | \Psi_0 \rangle c_0 + c_0^* \langle \Psi_0 | \widehat{H} | \mathbf{S} \rangle \mathbf{c}_\mathbf{S} + c_0^* \langle \Psi_0 | \widehat{H} | \mathbf{D} \rangle \mathbf{c}_\mathbf{D} \\ &\quad + \mathbf{c}_\mathbf{S}^\dagger \langle \mathbf{S} | \widehat{H} | \Psi_0 \rangle c_0 + \mathbf{c}_\mathbf{S}^\dagger \langle \mathbf{S} | \widehat{H} | \mathbf{S} \rangle \mathbf{c}_\mathbf{S} + \mathbf{c}_\mathbf{S}^\dagger \langle \mathbf{S} | \widehat{H} | \mathbf{D} \rangle \mathbf{c}_\mathbf{D} \\ &\quad + \mathbf{c}_\mathbf{D}^\dagger \langle \mathbf{D} | \widehat{H} | \Psi_0 \rangle c_0 + \mathbf{c}_\mathbf{D}^\dagger \langle \mathbf{D} | \widehat{H} | \mathbf{S} \rangle \mathbf{c}_\mathbf{S} + \mathbf{c}_\mathbf{D}^\dagger \langle \mathbf{D} | \widehat{H} | \mathbf{D} \rangle \mathbf{c}_\mathbf{D}. \end{aligned} \quad (176)$$

Such an equation is not very useful, since we still do not know the c_0 , $\mathbf{c}_\mathbf{S}$ and $\mathbf{c}_\mathbf{D}$ coefficients determining the Ψ_{SD} function. The only thing that can be said about them so far comes from the normalization requirement for Ψ_{SD} :

$$1 = c_0^* c_0 + \mathbf{c}_\mathbf{S}^\dagger \mathbf{c}_\mathbf{S} + \mathbf{c}_\mathbf{D}^\dagger \mathbf{c}_\mathbf{D}. \quad (177)$$

This is not enough to uniquely determine the wave function. However, going back to Eq. 175 and multiplying only the inside vectors, one obtains

$$\mathbf{H}_{\text{SD}} = \begin{bmatrix} \langle \Psi_0 | \\ \langle \mathbf{S} | \\ \langle \mathbf{D} | \end{bmatrix} \widehat{H} \begin{bmatrix} | \Psi_0 \rangle, | \mathbf{S} \rangle, | \mathbf{D} \rangle \end{bmatrix} = \begin{bmatrix} \langle \Psi_0 | \widehat{H} | \Psi_0 \rangle & \langle \Psi_0 | \widehat{H} | \mathbf{S} \rangle & \langle \Psi_0 | \widehat{H} | \mathbf{D} \rangle \\ \langle \mathbf{S} | \widehat{H} | \Psi_0 \rangle & \langle \mathbf{S} | \widehat{H} | \mathbf{S} \rangle & \langle \mathbf{S} | \widehat{H} | \mathbf{D} \rangle \\ \langle \mathbf{D} | \widehat{H} | \Psi_0 \rangle & \langle \mathbf{D} | \widehat{H} | \mathbf{S} \rangle & \langle \mathbf{D} | \widehat{H} | \mathbf{D} \rangle \end{bmatrix}. \quad (178)$$

We can then associate finding the approximate Hamiltonian eigenvalues with its matrix in the Ψ_{SD} basis:

$$\langle \mathbf{H} \rangle_{\Psi_{\text{SD}}} = [c_0^* \mathbf{c}_S^\dagger \mathbf{c}_D^\dagger] = \begin{bmatrix} \langle \Psi_0 | \widehat{H} | \Psi_0 \rangle & \langle \Psi_0 | \widehat{H} | \mathbf{S} \rangle & \langle \Psi_0 | \widehat{H} | \mathbf{D} \rangle \\ \langle \mathbf{S} | \widehat{H} | \Psi_0 \rangle & \langle \mathbf{S} | \widehat{H} | \mathbf{S} \rangle & \langle \mathbf{S} | \widehat{H} | \mathbf{D} \rangle \\ \langle \mathbf{D} | \widehat{H} | \Psi_0 \rangle & \langle \mathbf{D} | \widehat{H} | \mathbf{S} \rangle & \langle \mathbf{D} | \widehat{H} | \mathbf{D} \rangle \end{bmatrix} \begin{bmatrix} c_0 \\ \mathbf{c}_S \\ \mathbf{c}_D \end{bmatrix}. \quad (179)$$

Because of the Hermitian character of the Hamilton operator, the \mathbf{H}_{SD} matrix is symmetric and real. Its diagonalization provides the set of the eigenvalues corresponding to its eigenvectors. We are interested in the ground-state energy and thus, we need only the lowest eigenvalue of the \mathbf{H}_{SD} matrix and the respective normalized eigenvector Ψ_{SD} .

Knowing the procedure for the finite basis (only single- and double-orbital exchanges), we can see how it looks for the full (Eq. 171) expansion. The matrix notation leads to the average value of the Hamiltonian, written as

$$\langle H \rangle_\Psi = [c_0, \mathbf{c}_S, \mathbf{c}_D, \mathbf{c}_T, \mathbf{c}_Q] \begin{bmatrix} \langle \Psi_0 | \\ \langle \mathbf{S} | \\ \langle \mathbf{D} | \\ \langle \mathbf{T} | \\ \langle \mathbf{Q} | \\ \vdots \end{bmatrix} \widehat{H} [\Psi_0 | \mathbf{S} \rangle, | \mathbf{D} \rangle, | \mathbf{T} \rangle, | \mathbf{Q} \rangle, \dots] \begin{bmatrix} c_0 \\ \mathbf{c}_S \\ \mathbf{c}_D \\ \mathbf{c}_T \\ \mathbf{c}_Q \\ \vdots \end{bmatrix}. \quad (180)$$

The vector multiplication permits one to perceive the average value as the eigenproblem of the Hamiltonian matrix:

$$\mathbf{H} = \begin{bmatrix} \langle \Psi_0 | \widehat{H} | \Psi_0 \rangle & 0 & \langle \Psi_0 | \widehat{H} | \mathbf{D} \rangle & 0 & 0 & \dots \\ 0 & \langle \mathbf{S} | \widehat{H} | \mathbf{S} \rangle & \langle \mathbf{S} | \widehat{H} | \mathbf{D} \rangle & \langle \mathbf{S} | \widehat{H} | \mathbf{T} \rangle & 0 & \dots \\ \langle \mathbf{D} | \widehat{H} | \Psi_0 \rangle & \langle \mathbf{D} | \widehat{H} | \mathbf{S} \rangle & \langle \mathbf{D} | \widehat{H} | \mathbf{D} \rangle & \langle \mathbf{D} | \widehat{H} | \mathbf{T} \rangle & \langle \mathbf{D} | \widehat{H} | \mathbf{Q} \rangle & \dots \\ 0 & \langle \mathbf{T} | \widehat{H} | \mathbf{S} \rangle & \langle \mathbf{T} | \widehat{H} | \mathbf{D} \rangle & \langle \mathbf{T} | \widehat{H} | \mathbf{T} \rangle & \langle \mathbf{T} | \widehat{H} | \mathbf{Q} \rangle & \dots \\ 0 & 0 & \langle \mathbf{Q} | \widehat{H} | \mathbf{D} \rangle & \langle \mathbf{Q} | \widehat{H} | \mathbf{T} \rangle & \langle \mathbf{Q} | \widehat{H} | \mathbf{Q} \rangle & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}. \quad (181)$$

It can be clearly seen that some blocks in this matrix are equal to zero. This happens in two cases:

- The integrals between Ψ_0 and functions of the \mathbf{S} type (single exchange of spin-orbitals) vanish due to the Brillouin theorem, as was shown in the previous section.

- The integrals between the functions that differ by more than two exchanges, for instance, **S** and **Q** type, vanish due to the Slater rules.

Even not knowing combinatorics, one can expect that the number of functions in a block will grow drastically with and increase in the number of exchanges (block **S** will contain less functions than **D** etc.). A bit of thinking in the beginning would help to save a lot of time by not calculating zero integrals. Let's see: Only in the case of $\langle \Psi_0 | \widehat{H} | \mathbf{D} \rangle$ and $\langle \mathbf{S} | \widehat{H} | \mathbf{S} \rangle$ blocks should all the elements be calculated. The remaining matrices are sparse. For example, the $\langle \mathbf{D} | \widehat{H} | \mathbf{D} \rangle$ block contains the integrals of the following types: $\langle \Psi_{ab}^{pq} | \widehat{H} | \Psi_{ab}^{pq} \rangle$ (the same function on both sides), $\langle \Psi_{ab}^{pq} | \widehat{H} | \Psi_{ab}^{pr} \rangle$ (differing by one exchange), $\langle \Psi_{ab}^{pq} | \widehat{H} | \Psi_{ab}^{rs} \rangle$ (differing by two exchanges), $\langle \Psi_{ab}^{pq} | \widehat{H} | \Psi_{ac}^{pq} \rangle$ (differing by three exchanges), and $\langle \Psi_{ab}^{pq} | \widehat{H} | \Psi_{dc}^{pq} \rangle$ (differing by four exchanges). Obviously, the two latter cases produce zeros.

With the large number of exchanges, the size of the blocks grows abruptly, but most of the elements would be equal to zero. The simplification in this case would be the limitation of the **H** matrix size by the elimination of the functions including more than a given number of exchanges from the expansion. Let us leave only the single exchange block. Thus, the Hamiltonian matrix has the form

$$\mathbf{H}_s = \begin{bmatrix} \langle \Psi_0 | \widehat{H} | \Psi_0 \rangle & 0 \\ 0 & \langle \mathbf{S} | \widehat{H} | \mathbf{S} \rangle \end{bmatrix}. \quad (182)$$

This is the block diagonal matrix. One of the properties of such matrices is that their eigenvalue set is a sum of the eigenvalues of the diagonal blocks. This means that the lowest possible eigenvalue is $E_0 = \langle \Psi_0 | \widehat{H} | \Psi_0 \rangle$, and in consequence, there is no improvement in the ground-state energy with respect to the Hartree–Fock theory when taking only the single orbital exchanges.

Hence, let us also include the functions of the **D** type. Now the Hamilton matrix can be written as

$$\mathbf{H}_{SD} = \begin{bmatrix} \langle \Psi_0 | \widehat{H} | \Psi_0 \rangle & 0 & \langle \Psi_0 | \widehat{H} | \mathbf{D} \rangle \\ 0 & \langle \mathbf{S} | \widehat{H} | \mathbf{S} \rangle & \langle \mathbf{S} | \widehat{H} | \mathbf{D} \rangle \\ \langle \mathbf{D} | \widehat{H} | \Psi_0 \rangle & \langle \mathbf{D} | \widehat{H} | \mathbf{S} \rangle & \langle \mathbf{D} | \widehat{H} | \mathbf{D} \rangle \end{bmatrix}. \quad (183)$$

It is no longer a block diagonal matrix – all blocks contribute to its eigenvalues and one can count on some improvement. An interesting observation, however, is that here the functions with the single spin-orbital exchange also have influence on the energy via the $\langle \mathbf{S} | \widehat{H} | \mathbf{D} \rangle$ and $\langle \mathbf{D} | \widehat{H} | \mathbf{S} \rangle$ blocks.

Next, subsequent groups of functions can be applied containing more than two spin-orbital exchanges. However, the calculations become prohibitively expensive, even for moderate size of the systems, and the consecutive corrections are smaller and smaller. The distinguished character of the double spin-orbital exchange was already discussed within the MP2 method. Now one can also expect that including double exchanges produces reasonable results with the moderate computational cost. Then, why not save more and diagonalize only

$$\mathbf{H}_D = \begin{bmatrix} \langle \Psi_0 | \hat{H} | \Psi_0 \rangle & \langle \Psi_0 | \hat{H} | \mathbf{D} \rangle \\ \langle \mathbf{D} | \hat{H} | \Psi_0 \rangle & \langle \mathbf{D} | \hat{H} | \mathbf{D} \rangle \end{bmatrix} \quad (184)$$

instead of \mathbf{H}_{SD} ? This can be done; however, savings are not that great, since the number of \mathbf{S} functions is significantly smaller than the number of \mathbf{D} functions. Thus, if one can afford \mathbf{H}_D diagonalization, \mathbf{H}_{SD} diagonalization is probably also within easy reach.

The above reasoning has led to the sequence of quantum chemistry methods. The best results can be obtained within full CI (FCI) by applying the full expansion (Eq. 171) within the given basis set. This is certainly the most expensive variant. Cheaper – but also worse – are, respectively, CISD based on the \mathbf{H}_{SD} matrix and CID neglecting single exchanges (Cramer 2004; Jensen 2006; Levine 2008; Lowe and Peterson 2005; Piela 2007; Ratner and Schatz 2000; Szabo and Ostlund 1996).

So far, the reference function has been a single determinant. Such an approach is very limited. For instance, it does not allow one to describe a dissociation process. Correct description of dissociation requires at least one determinant for each subsystem. And, even in the cases when multi-determinant reference state description is not obligatory, such an elastic wave function will provide an improved description of the system of interest (Cramer 2004; Jensen 2006; Levine 2008; Piela 2007; Roos and Widmark 2002).

The multi-determinant wave function Ψ depends both on the expansion coefficients and on the spin-orbitals building up the determinants. Both these sets of variables can be optimized simultaneously. The particular case of this procedure, when taking only the first expansion term, is the Hartree–Fock approximation (SCF–HF). Therefore, the optimization of the multi-determinant wave function is called the multiconfiguration (MC) SCF method. Even without a detailed study of the MC–SCF equations, an improvement in the results with respect to the HF energy can be expected. However, this approach is much more expensive, since the spin-orbitals are optimized several times. Again, a time saving is desired. Therefore, let us search for the spin-orbitals with the highest influence on the total energy value. It has been observed that not all doubly exchanged functions provide the same contribution to the energy. Some improve the result more and others less. This is due to the spin-orbital energy differences. The exchange of the spin-orbitals of significantly different energies does not contribute much to the total energy improvement. Therefore, it can be requested that the exchange is included in calculations only if the energy difference between the involved spin-orbitals is

smaller than some given value. Hence, only some groups of spin-orbitals can be exchanged.

Up to now, the spin-orbital notion was used. However, let us switch to the orbital language that is frequently used for MC–SCF considerations.

For the N -electron system, the orbitals can be divided into three groups:

- Core orbitals, which are not varied, since they have too low orbital energies, but are applied in the wave function expansion (doubly occupied orbitals)
- Active orbitals, which are exchanged in the expansion (partially occupied orbitals)
- Virtual orbitals, which are not varied and not applied in the expansion (unoccupied orbitals)

Instead of optimization of all the orbitals, only the active orbitals will be varied within the complete active space self-consistent field (CASSCF) approximation. In the acronym of this method, the number of active orbitals and active electrons is also provided for the given system. For instance, CASSCF (6,4) denotes the calculations with the expansion including all the possible exchanges of the four electrons within the six active orbitals. The CASSCF approach leads to all possible exchanges in the given active space, and for a moderately sized system, the size of the active spaces can quickly exceed the computational resources. In such a case, the solution can be the restricted active space self-consistent field (RASSCF) method, which supplies a way of limiting the size of the active space.

Additionally, one needs to remember that for a powerful tool such as perturbation theory, there is no obstacle to applying the multi-determinant reference function as the unperturbed function in perturbation calculus. Thus, similar to the SCF–HF and MP2 approaches, CASPT2 would be the second-order perturbation theory complete active space method – the perturbationally corrected CASSCF.

Coupled Cluster Approximation: The Operator Strikes Back

It would seem that all the straightforward ways to improve wave function in the one-electron approximation have been exploited. However, we now next discuss one of the most accurate (and simultaneously most expensive) methods applied in quantum chemistry.

The idea is simple. Consider again the expansion (Eq. 171). Introducing an operator

$$\widehat{C} = \widehat{C}_0 + \widehat{C}_1 + \widehat{C}_2 + \widehat{C}_3 + \widehat{C}_4 + \dots, \quad (185)$$

defined as

$$\widehat{C}_0 |\Psi_0\rangle = c_0 |\Psi_0\rangle, \quad (186)$$

$$\widehat{C}_1 |\Psi_0\rangle = \mathbf{S}c_s, \quad (187)$$

$$\widehat{C}_2 |\Psi_0\rangle = \mathbf{Dc}_D, \quad (188)$$

$$\widehat{C}_3 |\Psi_0\rangle = \mathbf{Tc}_T, \quad (189)$$

$$\widehat{C}_4 |\Psi_0\rangle = \mathbf{Qc}_Q, \quad (190)$$

⋮

allows one to write Eq. 171 in a very compact form:

$$\Psi = \widehat{C} \Psi_0. \quad (191)$$

Now the problem of finding the appropriate expansion can be replaced by the problem of finding the adequate operator. This is the essence of the coupled cluster (CC) method. Here the assumption is made that the wave function can be expressed by

$$\Psi = e^{\widehat{T}} \Psi_0, \quad (192)$$

where Ψ_0 is a reference function (depending on the approach, this can be the one-determinant HF function or the multi-determinant function arising from MC-SCF) and \widehat{T} is a sought operator. Applying the expansion of the exponential function, it can be written that

$$e^{\widehat{T}} = \widehat{1} + \widehat{T} + \frac{1}{2!} \widehat{T}^2 + \frac{1}{3!} \widehat{T}^3 + \dots \quad (193)$$

Such an expanded form makes the interpretation of the \widehat{T} operators easier. Putting

$$\widehat{T} = \widehat{T}_1 + \widehat{T}_2 + \widehat{T}_3 + \widehat{T}_4 + \dots, \quad (194)$$

one can identify the subsequent \widehat{T}_i operators as corresponding to i -tuple exchanges of the spin-orbitals in the reference function:

$$\widehat{T}_1 \Psi_0 = \sum_{a,p} t_a^p \Psi_a^p, \quad (195)$$

$$\widehat{T}_2 \Psi_0 = \sum_{a,b,p,q} t_{ab}^{pq} \Psi_{ab}^{pq}, \quad (196)$$

and so forth. The coefficients t (called “amplitudes”) are in general not equivalent to the c coefficients in the CI expansion (see Eq. 171). In order to find their mutual relation, let us consider the approximate operator:

$$\widehat{T} \approx \widehat{T}_1 + \widehat{T}_2 + \widehat{T}_3 + \widehat{T}_4. \quad (197)$$

The operator (Eq. 193) takes the form

$$\begin{aligned}
 e^{\widehat{T}_1 + \widehat{T}_2 + \widehat{T}_3 + \widehat{T}_4} &= \widehat{1} \\
 &+ \widehat{T}_1 + \widehat{T}_2 + \widehat{T}_3 + \widehat{T}_4 \\
 &+ \frac{1}{2!} \left(\widehat{T}_1 + \widehat{T}_2 + \widehat{T}_3 + \widehat{T}_4 \right)^2 \\
 &+ \frac{1}{3!} \left(\widehat{T}_1 + \widehat{T}_2 + \widehat{T}_3 + \widehat{T}_4 \right)^3 \\
 &+ \frac{1}{4!} \left(\widehat{T}_1 + \widehat{T}_2 + \widehat{T}_3 + \widehat{T}_4 \right)^4.
 \end{aligned} \tag{198}$$

Limiting ourselves to the terms corresponding to not more than four spin-orbital exchanges and writing it in the ordered way according to the number of exchanges, one gets

$$\begin{aligned}
 e^{\widehat{T}_1 + \widehat{T}_2 + \widehat{T}_3 + \widehat{T}_4} &= \widehat{1} \\
 &+ \widehat{T}_1 \\
 &+ \widehat{T}_2 + \frac{1}{2} \widehat{T}_1^2 \\
 &+ \widehat{T}_3 + \widehat{T}_1 \widehat{T}_2 + \frac{1}{3} \widehat{T}_1^3 \\
 &+ \widehat{T}_4 + \widehat{T}_1 \widehat{T}_3 + \frac{1}{2} \widehat{T}_2^2 + \frac{1}{2} \widehat{T}_1^2 \widehat{T}_2 + \widehat{T}_1^4.
 \end{aligned} \tag{199}$$

Now the direct comparison can be made:

$$\widehat{C}_1 = \widehat{T}_1, \tag{200}$$

$$\widehat{C}_2 = \widehat{T}_2 + \frac{1}{2} \widehat{T}_1^2, \tag{201}$$

$$\widehat{C}_3 = \widehat{T}_3 + \widehat{T}_1 \widehat{T}_2 + \frac{1}{3} \widehat{T}_1^3, \tag{202}$$

$$\widehat{C}_4 = \widehat{T}_4 + \widehat{T}_1 \widehat{T}_3 + \frac{1}{2} \widehat{T}_2^2 + \frac{1}{2} \widehat{T}_1^2 \widehat{T}_2 + \widehat{T}_1^4. \tag{203}$$

We have the relation between the \widehat{C}_i and \widehat{T}_i operators, but still neither \widehat{C}_i nor \widehat{T}_i are known. Recall from the earlier sections that the double exchanges have a significant influence on the energy improvement with respect to the Hartree–Fock results. Taking double exchanges into account within the coupled cluster formalism means that the operators \widehat{T}_1 and \widehat{T}_2 need to be determined. However, as a side effect, they also allow inclusion of some not negligible contributions arising from the triple and higher exchanges. In the above comparison, the \widehat{T}_1 and \widehat{T}_2 operators recover two out of three terms in \widehat{C}_3 and three out of five terms in \widehat{C}_4 . This is the power of the CC method.

Unfortunately, the strength of this method does not go together with ease of calculations. Obtaining the expressions for the operators \widehat{T} is ransomed with

compromises. Not only is the operator expansion (Eq. 194) truncated, but the basis set is finite. Moreover, the variational character of the method is sacrificed.

In order to realize the complications, let us consider step-by-step the energy calculation within the CC formalism. We begin, as usual, with the electron Schrödinger Eq. 50. Substituting Eq. 192 gives

$$\widehat{H}e^{\widehat{T}}\Psi_0 = Ee^{\widehat{T}}\Psi_0. \quad (204)$$

Taking into account that, due to (Eq. 193),

$$\langle \Psi_0 | \Psi \rangle = \langle \Psi_0 | e^{\widehat{T}} | \Psi_0 \rangle = \langle \Psi_0 | \Psi_0 \rangle = 1, \quad (205)$$

the energy can be calculated as

$$E = \langle \Psi_0 | \widehat{H}e^{\widehat{T}} | \Psi_0 \rangle = \langle \Psi_0 | \widehat{H} | \Psi \rangle. \quad (206)$$

This is not the Hamiltonian average value expression. Additionally, the operator inside the bracket is not Hermitian. But, until we assume that (Eq. 192) is true, such an approach works. We can also construct an integral:

$$\langle \Psi_{ab}^{pq} | \widehat{H}e^{\widehat{T}} | \Psi_0 \rangle = E \langle \Psi_{ab}^{pq} | e^{\widehat{T}} | \Psi_0 \rangle, \quad (207)$$

which is the consequence of Eq. 204 and will be applied in the near future.

We should now concentrate on the way of determining the form of amplitudes. To simplify the considerations, we can assume

$$\widehat{T} \approx \widehat{T}_2, \quad (208)$$

which is equivalent to the CCD variant. We are interested in finding the amplitudes t_{ab}^{pq} . The final result of the calculations will be the approximate energy:

$$E_{\text{CCD}} = \langle \Psi_0 | \widehat{H}e^{\widehat{T}_2} | \Psi_0 \rangle. \quad (209)$$

The information about the amplitude t_{ab}^{pq} can be extracted from the integral

$$t_{ab}^{pq} = \langle \Psi_{ab}^{pq} | \widehat{T}_2 | \Psi_0 \rangle \quad (210)$$

(see Eq. 196). However, the amplitudes are still not known, since we do not know the \widehat{T}_2 operator. Therefore, one more equation is necessary to elicit the sought information. Let us begin with the approximated expression (Eq. 207):

$$\begin{aligned} \langle \Psi_{ab}^{pq} | \widehat{H} e^{\widehat{T}_2} | \Psi_0 \rangle &= E_{\text{CCD}} \langle \Psi_{ab}^{pq} | e^{\widehat{T}_2} | \Psi_0 \rangle \\ &= \langle \Psi_0 | \widehat{H} e^{\widehat{T}_2} | \Psi_0 \rangle \langle \Psi_{ab}^{pq} | e^{\widehat{T}_2} | \Psi_0 \rangle. \end{aligned} \quad (211)$$

The expansion (Eq. 193) tailored to the present case,

$$e^{\widehat{T}_2} = \widehat{1} + \widehat{T}_2 + \frac{1}{2} \widehat{T}_2^2 + \dots, \quad (212)$$

and substituted to the left-hand side of Eq. 211 gives

$$\langle \Psi_{ab}^{pq} | \widehat{H} e^{\widehat{T}} | \Psi_0 \rangle = \langle \Psi_{ab}^{pq} | \widehat{H} \left(\widehat{1} + \widehat{T}_2 + \frac{1}{2} \widehat{T}_2^2 \right) | \Psi_0 \rangle. \quad (213)$$

Further terms are not necessary; in such a case, the functions on both sides of the integral would differ with four and more spin-orbital exchanges (and Hamiltonian is still a sum of one- and two-electron operators). Similarly, the expansion in the integral $\langle \Psi_0 | \widehat{H} e^{\widehat{T}_2} | \Psi_0 \rangle$ will also be truncated on the second term:

$$\langle \Psi_0 | \widehat{H} e^{\widehat{T}_2} | \Psi_0 \rangle = \langle \Psi_0 | \widehat{H} \left(\widehat{1} + \widehat{T}_2 \right) | \Psi_0 \rangle. \quad (214)$$

Remembering that

$$E_0^{(0)} = \langle \Psi_0 | \widehat{H} | \Psi_0 \rangle, \quad (215)$$

one gets

$$\langle \Psi_0 | \widehat{H} e^{\widehat{T}_2} | \Psi_0 \rangle = E_0^{(0)} + \langle \Psi_0 | \widehat{H} \widehat{T}_2 | \Psi_0 \rangle. \quad (216)$$

The last integral on the right-hand side of Eq. 211, $\langle \Psi_{ab}^{pq} | e^{\widehat{T}} | \Psi_0 \rangle$, can be nonvanishing only if the functions on the right and left side are the same. This is possible for

$$\langle \Psi_{ab}^{pq} | e^{\widehat{T}_2} | \Psi_0 \rangle = \langle \Psi_{ab}^{pq} | \widehat{T}_2 | \Psi_0 \rangle. \quad (217)$$

This is the integral that can provide information about the desired amplitudes of Eq. 210. Putting all these together, one gets

$$\langle \Psi_{ab}^{pq} | \widehat{H} \left(\widehat{1} + \widehat{T}_2 + \frac{1}{2} \widehat{T}_2^2 \right) | \Psi_0 \rangle = \left(E_0 + \langle \Psi_0 | \widehat{H} \widehat{T}_2 | \Psi_0 \rangle \right) \langle \Psi_{ab}^{pq} | \widehat{T}_2 | \Psi_0 \rangle. \quad (218)$$

Therefore, the amplitude $t_{ab}^{pq} = \langle \Psi_{ab}^{pq} | \hat{T}_2 | \Psi_0 \rangle$ can be expressed as

$$t_{ab}^{pq} = \frac{\langle \Psi_{ab}^{pq} | \hat{H} (\hat{1} + \hat{T}_2 + \frac{1}{2} \hat{T}_2^2) | \Psi_0 \rangle}{E_0 + \langle \Psi_0 | \hat{H} \hat{T}_2 | \Psi_0 \rangle}. \quad (219)$$

Unluckily, this does not mean that the amplitudes are known. Still, the above expression also contains the t_{ab}^{pq} amplitudes on the right-hand side in the \hat{T}_2 operators. Moreover, all other amplitudes are also present on the right-hand side. The consequence of this aggravation is that the CC equations cannot be solved separately, one by one. All together, the complicated set of nonlinear equations must be handled. The number of equations is equal to the number of sought amplitudes. This is the main reason for the huge computational cost of the CC calculations, even though the variationality of the method was abandoned (Atkins and Friedman 2005; Cramer 2004; Jensen 2006; Levine 2008; Piela 2007; Roos and Widmark 2002).

It can be seen that solving the CC equations is quite complicated, even in the simplified case of the CCD approach. If one wanted to use the variational Hamiltonian and apply its average value, the following integrals would appear:

$$\langle \Psi | \hat{H} | \Psi \rangle = \langle \Psi_0 | e^{\hat{T}} \hat{H} e^{\hat{T}} | \Psi_0 \rangle = \langle e^{\hat{T}} \Psi_0 | \hat{H} | e^{\hat{T}} \Psi_0 \rangle. \quad (220)$$

In order to calculate them, one needs to know the form of all the \hat{T}_i operators, since not only the function on the right-hand side of the above integral will contain the exchanged spin-orbitals but also the function on the left-hand side. Therefore, one needs to calculate terms like $\langle \hat{T}_3 \Psi_0 | \hat{H} | \hat{T}_2 \Psi_0 \rangle$ and many others. This causes the significant increase of the computational costs of the CC method.

Like in the MP n case, the CC method is worth using for the short expansion of the \hat{T} operator. Thus, relatively good accuracy is obtained with a moderate price.

Conclusions

We have finally reached the end of the zeroth iteration in the process of learning quantum chemistry methods. The beginner may feel saturated or even overwhelmed; however, we hope that this chapter arouses interest. Our aim was to show that simple ideas underlie quantum chemistry methods. The purpose is to put complicated things in a simpler and more convenient form. One of the most popular rules in computational chemistry is as follows: "If you cannot calculate something, divide it into parts in such a way that you can calculate some contribution while the other is too difficult." For instance, nonrelativistic energy can be divided into HF energy and correlation energy. Correlation energy accounts for the contribution that we cannot calculate in practice, but methods such as MP n , CC, and CI allow one to find some

part of it. It may happen (and it often does!) that what we can calculate will be enough.

This chapter should be treated as the introduction to more advanced handbooks or as a guide through the symbols and concepts applied in the later parts of this book. Thus, some of the concepts are just touched upon, and many are omitted. If the reader noticed this and wants to know more, it means that this chapter has met its goal.

Acknowledgments Authors are grateful to Dr Krzysztof Strasburger, Dr Andrej Antušek, Agnieszka Zawada, and Łukasz Mentel for helpful comments.

Bibliography

- Atkins, P., & Friedman, R. (2005). *Molecular quantum mechanics*. Oxford: Oxford University Press.
- Cramer, C. J. (2004). *Essentials of computational chemistry: Theories and models*. Chichester: Wiley.
- Griffiths, D. J. (2004). *Introduction to quantum mechanics* (2nd ed.). San Francisco: Benjamin Cummings.
- Jensen, F. (2006). *Introduction to computational chemistry*. Chichester: Wiley.
- Levine, I. N. (2008). *Quantum chemistry* (6th ed.). Englewood Cliffs: Prentice Hall.
- Lowe, J. P., & Peterson, K. (2005). *Quantum chemistry* (3rd ed.). Boston: Academic.
- McQuarrie, D. A., & Simon, J. D. (1997). *Physical chemistry: A molecular approach*. Sausalito: University Science.
- Piela, L. (2007). *Ideas of quantum chemistry*. Amsterdam: Elsevier.
- Ratner, M. A., & Schatz, G. C. (2000). *Introduction to quantum mechanics in chemistry*. Upper Saddle River: Prentice Hall.
- Roos, B. O., & Widmark, P.-O. (Eds.). (2002). *European summerschool in quantum chemistry*. Lund: Lund University.
- Szabo, A., & Ostlund, N. S. (1996). *Modern quantum chemistry: Introduction to advanced electronic structure theory*. Mineola: Dover.