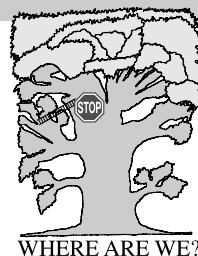


Chapter 8

ELECTRONIC MOTION IN THE MEAN FIELD: ATOMS AND MOLECULES



Where are we?

We are in the upper part of the main trunk of the TREE.

An example

What is the electronic structure of atoms? How do atoms interact in a molecule? Two *neutral* moieties (say, hydrogen atoms) attract each other with a force *of a similar order of magnitude* to the Coulombic forces between two ions. This is quite surprising. What pulls these neutral objects to one another? These questions are at the foundations of chemistry.

What is it all about

Hartree–Fock method – a bird’s eye view (▲)

p. 329

- Spinorbitals
- Variables
- Slater determinants
- What is the Hartree–Fock method all about?

The Fock equation for optimal spinorbitals (Δ)

p. 334

- Dirac and Coulomb notations
- Energy functional
- The search for the conditional extremum
- A Slater determinant and a unitary transformation
- Invariance of the \hat{J} and \hat{K} operators
- Diagonalization of the Lagrange multipliers matrix
- The Fock equation for optimal spinorbitals (General Hartree–Fock method – GHF)
- The closed-shell systems and the Restricted Hartree–Fock (RHF) method
- Iterative procedure for computing molecular orbitals: the Self-Consistent Field method

Total energy in the Hartree–Fock method (Δ)

p. 351

Computational technique: atomic orbitals as building blocks of the molecular wave function (▲)

p. 354

- Centring of the atomic orbital
- Slater-type orbitals (STO)

- Gaussian-type orbitals (GTO)
- Linear Combination of Atomic Orbitals (LCAO) Method
- Basis sets of Atomic Orbitals
- The Hartree–Fock–Roothaan method (SCF LCAO MO)
- Practical problems in the SCF LCAO MO method
- Back to foundations (♦) p. 369**
 - When does the RHF method fail?
 - Fukutome classes
- RESULTS OF THE HARTREE–FOCK METHOD p. 379**
- Mendeleev Periodic Table of Chemical Elements (▲) p. 379**
 - Similar to the hydrogen atom – the orbital model of atom
 - Yet there are differences...
- The nature of the chemical bond (Δ) p. 383**
 - H_2^+ in the MO picture
 - Can we see a chemical bond?
- Excitation energy, ionization potential, and electron affinity (RHF approach) (Δ) p. 389**
 - Approximate energies of electronic states
 - Singlet or triplet excitation?
 - Hund's rule
 - Ionization potential and electron affinity (Koopmans rule)
- Localization of molecular orbitals within the RHF method (Δ) p. 396**
 - The external localization methods
 - The internal localization methods
 - Examples of localization
 - Computational technique
 - The σ , π , δ bonds
 - Electron pair dimensions and the foundations of chemistry
 - Hybridization
- A minimal model of a molecule (▲) p. 417**
 - Valence Shell Electron Pair Repulsion (VSEPR)

The Born–Oppenheimer (or adiabatic) approximation is the central point of this book (note its position in the TREE). Thanks to the approximation, we can consider *separately* two *coupled* problems concerning molecules:

- the motion of the electrons at fixed positions of the nuclei (to obtain the electronic energy),
- the motion of nuclei in the potential representing the electronic energy of the molecule (see Chapter 7).

From now on we will concentrate on the *motion of the electrons at fixed positions of the nuclei* (the Born–Oppenheimer approximation, p. 229).

To solve the corresponding eq. (6.18), we have at our disposal the variational and the perturbation methods. The latter one should have a reasonable starting point (i.e. an unperturbed system). This is not the case in the problem we want to consider at the moment. Thus, only the variational method remains. If so, a class of trial functions should be proposed. In this chapter the trial wave function will have a very specific form, bearing significant importance for the theory. We mean here what is called Slater determinant, which is

composed of molecular orbitals. At a certain level of approximation, each molecular orbital is a “home” for two electrons. We will now learn on how to get the optimum molecular orbitals (Hartree–Fock method). Despite some quite complex formulas, which will appear below, the main idea behind them is extremely simple. It can be expressed in the following way.

Let us consider the road traffic, the cars (electrons) move at fixed positions of buildings (nuclei). The motion of the cars proves to be very complex (as it does for the electrons) and therefore the problem is extremely difficult. How can such a motion be described in an approximate way? To describe such a complex motion one may use the so called mean field approximation (paying the price of poor quality). In the mean field approximation method we focus on the motion of *one car only* considering its motion in such way that the car avoids *those streets that are usually most jammed*. In this chapter we will treat the electrons in a similar manner (leaving the difficulties of considering the correlation of the motions of the electrons to Chapter 10). Now, the electrons will not feel the true electric field of the other electrons (as it should be in a precise approach), but rather their *mean* electric field, i.e. averaged over their motions.

Translating it into quantum mechanical language, the underlying assumptions of the mean field method for the N identical particles (here: electrons) are as follows:

- there is a certain “effective” *one-particle* operator $\hat{F}(i)$ of an identical mathematical form for all particles $i = 1, 2, \dots, N$, which has the eigenfunctions φ_k , i.e. $\hat{F}\varphi_k = \varepsilon_k\varphi_k$, such that
- $\langle \Psi | \hat{H} \Psi \rangle \approx \langle \tilde{\Psi} | \hat{H}^{ef} \tilde{\Psi} \rangle$, where $\tilde{\Psi}$ is an approximate wave function (to the exact wave function Ψ , both functions normalized) for the total system, \hat{H} is the electronic Hamiltonian (in the clamped nuclei approximation, Chapter 6), and $\hat{H}^{ef} = \sum_{i=1}^N \hat{F}(i)$. In such a case the eigenvalue equation $\hat{H}^{ef} \prod_{i=1}^N \varphi_i(i) = E_0 \prod_{i=1}^N \varphi_i(i)$ holds, and the approximate total energy is equal to $E_0 = \sum_{i=1}^N \varepsilon_k$, as if the particles were independent.

Any mean field method needs to solve two problems:

- How should $\tilde{\Psi}$ be constructed using N eigenfunctions φ_k ?
- What is the form of the one-particle effective operator \hat{F} ?

These questions will be answered in the present chapter.

Such effectively independent, yet interacting particles, are called quasiparticles or – as we sometimes use to say – bare particles dressed up by the interaction with others.

It is worth remembering that the mean field method bears several different names in chemistry:

- one-determinant approximation,
- one-electron approximation,
- one-particle approximation,
- molecular orbital method,
- independent-particle approximation,
- mean field approximation,
- Hartree–Fock method,
- self-consistent field method (as regards practical solutions).

It will be shown how the mean field method implies that mile-stone of chemistry: the periodic table of chemical elements.

Next, we will endeavour to understand *why* two atoms create a chemical bond, and also what affects the ionization energy and the electron affinity of a molecule.

Then, still within the molecular orbital scheme, we will show how we can reach a localized description of a molecule, with chemical bonds between *some* atoms, with the inner electronic shells, and the lone electronic pairs. The last terms are elements of a rich and very useful language commonly used by chemists.

Why this is important?

Contemporary quantum chemistry uses better methods than the mean field, described in this chapter. We will get to know them in Chapters 10 and 11. Yet all these methods *start from the mean field approximation* and in most cases they only perform cosmetic changes in energy and in electron distribution. For example, the methods described here yield about 99% of the total energy of a system.¹ There is one more reason why this chapter is important. Methods beyond the one-electron approximation are – computationally – very time-consuming (hence they may be applied only to small systems), while the molecular orbital approach is the “daily bread” of quantum chemistry. It is a sort of standard method, and the standards have to be learnt.

What is needed?

- Postulates of quantum chemistry (Chapter 1, necessary).
- Operator algebra, Hermitian operators (Appendix B, p. 895, necessary).
- Complete set of functions (Chapter 1, necessary).
- Hilbert space (Appendix B, p. 895, recommended).
- Determinants (Appendix A, p. 889, absolutely necessary).
- Slater–Condon rules (Appendix M, p. 986, only the results are needed).
- Lagrange multipliers (Appendix N, p. 997, necessary).
- Mulliken population analysis (Appendix S, p. 1015, occasionally used).

Classical works

This chapter deals with the basic theory explaining electronic structure of atoms and molecules. This is why we begin by Dimitrii Ivanovich Mendeleev who discovered in 1865, when writing his book “*Osnovy Khimii*” (“*Principles of Chemistry*”), St Petersburg, Tovarishestvo Obshchestvennaya Polza, 1869–71, his famous periodic table of elements – one of the greatest human achievements. ★ Gilbert Newton Lewis in the paper “*The Atom and the Molecule*” published in the *Journal of the American Chemical Society*, 38 (1916) 762 and Walter Kossel in an article “*Über die Molekülbildung als Frage des Atombaus*” published in *Annalen der Physik*, 49 (1916) 229, introduced such important theoretical tools as the octet rule and stressed the importance of the noble gas electronic configurations. ★ As soon as quantum mechanics was formulated in 1926, Douglas R. Hartree published several papers in the *Proceedings of the Cambridge Philosophical Society*, 24 (1927) 89, 24 (1927) 111, 26 (1928) 89, entitled “*The Wave Mechanics of an Atom with a Non-Coulomb Central Field*”, containing the computations for atoms such large as Rb and Cl. These were *self-consistent ab initio*² computations. . . , and the wave function was assumed to be the *product* of spinor-

¹In physics and chemistry we are seldom interested in the total energy. The energy differences of various states are of importance. Sometimes such precision is not enough, but the result speaks for itself.

²That is, derived from the first principles of (non-relativistic) quantum mechanics! Note, that these young people worked incredibly fast (no e-mail, no PCs).

bitals. ★ The LCAO approximation (for the solid state) was introduced by Felix Bloch in his PhD thesis “*Über die Quantenmechanik der Elektronen in Kristallgittern*”, University of Leipzig, 1928, and three years later Erich Hückel used this method to describe the first molecule (benzene) in a publication “*Quantentheoretische Beiträge zum Benzolproblem. I. Die Elektronenkonfiguration des Benzols*”, which appeared in *Zeitschrift für Physik*, 70 (1931) 203. ★ Vladimir Fock introduced the antisymmetrization of the spinorbital product in his publication “*Näherungsmethode zur Lösung des quantenmechanischen Mehrkörperproblems*” in *Zeitschrift für Physik*, 61 (1930) 126 and *ibid.* 62 (1930) 795. ★ John Slater proposed the idea of the multi-configurational wave function (“*Cohesion in Monovalent Metals*”, *Physical Review*, 35 (1930) 509). ★ The Hartree–Fock method in the LCAO approximation was formulated by Clemens C.J. Roothaan in his work “*New Developments in Molecular Orbital Theory*” published in the *Reviews of Modern Physics*, 23 (1951) 69, and, independently, by George G. Hall in a paper “*The Molecular Orbital Theory of Chemical Valency*” in *Proceedings of the Royal Society (London)*, A205 (1951) 541. ★ The physical interpretation of the orbital energies in the Hartree–Fock method was given by Tjalling C. Koopmans in his only quantum chemical paper “*On the Assignment of Wave Functions and Eigenvalues to the Individual Electron of an Atom*” published in *Physica*, 1 (1933/1934) 104. ★ The first localized orbitals (for the methane molecule) were computed by Charles A. Coulson despite the difficulties of war time (*Transactions of the Faraday Society*, 38 (1942) 433). ★ Hideo Fukutome, first in *Progress in Theoretical Physics*, 40 (1968) 998, and then in several following papers, analyzed general solutions for the Hartree–Fock equations from the symmetry viewpoint, and showed exactly eight classes of such solutions.

In the previous chapter the motion of the nuclei was considered. In the Born–Oppenheimer approximation (Chapter 6) the motion of the nuclei takes place in the potential, which is the electronic energy of a system (being a function of the nuclei position, \mathbf{R} , in the configurational space). The electronic energy $E_k^0(\mathbf{R})$ is an eigenvalue given in eq. (6.8) (adapted to the polyatomic case, hence $R \rightarrow \mathbf{R}$): $\hat{H}_0\psi_k(\mathbf{r}; \mathbf{R}) = E_k^0(\mathbf{R})\psi_k(\mathbf{r}; \mathbf{R})$. We will now deal exclusively with this equation, i.e. we will consider the electronic motion at fixed positions of the nuclei (clamped nuclei). Thus, our goal is two-fold: we are interested in what the electronic structure looks like and in how the electronic energy depends on the positions of the nuclei.³

Any theoretical method applicable to molecules may be also used for atoms, albeit very accurate wave functions, even for simple atoms, are not easy to calculate.⁴ In fact for atoms we know the solutions quite well only in the mean field approximation, i.e. the atomic orbitals. Such orbitals play an important role as building blocks of many-electron wave functions.

³In the previous chapter the ground-state electronic energy $E_0^0(\mathbf{R})$ was denoted as $V(\mathbf{R})$.

⁴If an atom is considered in the Born–Oppenheimer approximation, the problem is even simpler, the electronic equation also holds; we can then take, e.g., $\mathbf{R} = \mathbf{0}$. People still try to compute correlated wave functions (i.e. beyond the mean field approximation, see Chapter 10) for heavier atoms. Besides, relativistic effects (see Chapter 3) play increasingly important roles for such atoms. Starting with magnesium, they are larger than the correlation corrections. Fortunately, the relativistic corrections for atoms are largest for the inner electronic shells, which are the least important for chemists.

8.1 HARTREE–FOCK METHOD – A BIRD’S EYE VIEW

Douglas R. Hartree (1897–1958) was born and died in Cambridge. He was a British mathematician and physicist, professor at Manchester University, and then professor of mathematical physics at Cambridge. Until 1921 his interest was in the development of numerical methods for anti-aircraft artillery (he had some experience from the 1st World War), but a lecture by Niels Bohr has completely changed his career. Hartree immediately started investigating atoms. He used the atomic wave function in the form of the spinorbital *product*. Hartree learnt to use machines to solve differential equations while in Boston, and then he built one for himself at Cambridge. The machine was invented by Lord Kelvin, and constructed by Vannevar Bush in the USA. The machine integrated equations using a circle which rolled on a rotating disc. Later the first electronic computer, ENIAC, was used, and Hartree was asked to come and help to compute missile trajectories. An excerpt from *“Solid State and Molecular Theory”*, Wiley, London, 1975 by



John C. Slater: *“Douglas Hartree was very distinctly of the matter-of-fact habit of thought that I found most congenial. The hand-waving magical type of scientist regarded him as a ‘mere computer’. Yet he made a much greater contribution to our knowledge of the behaviour of real atoms than most of them did. And while he limited himself to atoms, his demonstration of the power of the self-consistent field for atoms is what has led to the development of that method for molecules and solids as well”.*

Before introducing the detailed formalism of the Hartree–Fock method, let us first look at its principal features. It will help us to understand our mathematical goal.

First of all, the positions of the nuclei are frozen (Born–Oppenheimer approximation) and then we focus on the wave function of N electrons. Once we want to move nuclei, we need to repeat the procedure from the beginning (for the new position of the nuclei).

Vladimir A. Fock (1898–1974), Russian physicist, professor at the Leningrad University (Sankt Petersburg), led investigations on quantum mechanics, gravity theory, general relativity theory, and in 1930, while explaining atomic spectra, invented the *antisymmetrization* of the spinorbitals product.



8.1.1 SPINORBITALS

Although this comparison is not precise, the electronic wave function for a molecule is built of segments, as a house is constructed from bricks.

The electronic wave function of a molecule containing N electrons depends on $3N$ Cartesian coordinates of the electrons and on their N spin coordinates (for each electron, its $\sigma = \frac{1}{2}$ or $-\frac{1}{2}$). Thus, it is a function of position in $4N$ -dimensional space. This function will be created out of simple “bricks”, i.e. *molecular spinorbitals*. Each of those will be a function of the coordinates of *one* electron only:

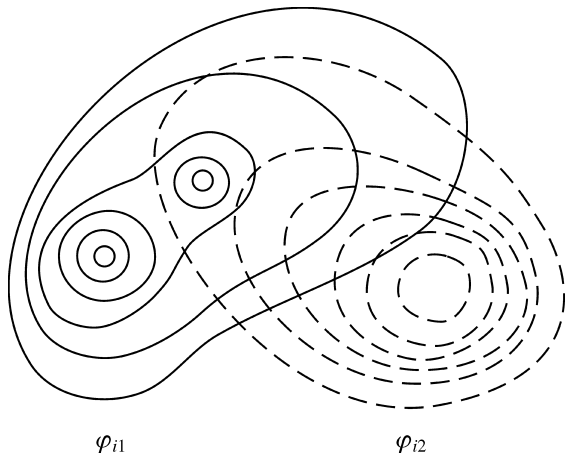


Fig. 8.1. According to eq. (8.1) a spinorbital is a mixture of α and β orbital components: $\varphi_{i1}(\mathbf{r})$ and $\varphi_{i2}(\mathbf{r})$, respectively. Figure shows two sections of such a spinorbital (z denotes the Cartesian axis perpendicular to the plane of the page): section $z = 0, \sigma = \frac{1}{2}$ (solid isolines) and section $z = 0, \sigma = -\frac{1}{2}$ (dashed isolines). In practical applications most often a restricted form of spinorbitals is used: either $\varphi_{i1} = 0$ or $\varphi_{i2} = 0$, i.e. a spinorbital is taken as an orbital part times spin function α or β .

three Cartesian coordinates and one spin coordinate (cf. Chapter 1). A spinorbital is therefore a function of the coordinates in the 4D space,⁵ and in the most general case a normalized spinorbital reads as (Fig. 8.1)

$$\phi_i(\mathbf{r}, \sigma) = \varphi_{i1}(\mathbf{r})\alpha(\sigma) + \varphi_{i2}(\mathbf{r})\beta(\sigma), \tag{8.1}$$

where the orbital components φ_{i1} and φ_{i2} (square-integrable functions) that depend on the position \mathbf{r} of the electron can adopt *complex* values, while the spin functions α and β , which depend on the spin coordinate σ , are defined in Chapter 1, p. 28. In the vast majority of quantum mechanical calculations the spinorbital ϕ_i is a *real* function, and φ_{i1} and φ_{i2} are such that either $\varphi_{i1} = 0$ or $\varphi_{i2} = 0$. *Yet for the time being we do not introduce any significant⁶ restrictions for the spinorbitals.* Spinorbital ϕ_i will adopt different complex values for various spatial coordinates as well as for a given value⁷ of the spin coordinate σ .

8.1.2 VARIABLES

Thus the variables, on which the wave function depends, are as follows:

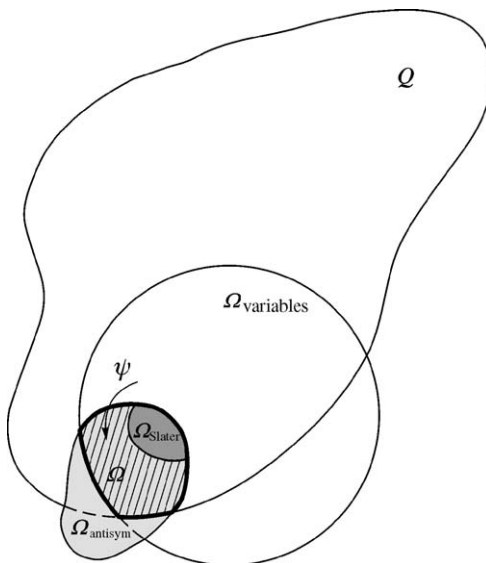
- x_1, y_1, z_1, σ_1 or briefly 1,
- x_2, y_2, z_2, σ_2 or briefly 2,
-
- x_N, y_N, z_N, σ_N or briefly N ,

where x_i, y_i, z_i are the Cartesian coordinates and σ_i is the spin coordinate of electron i .

The true wave function ψ (i.e. the eigenfunction of the Hamiltonian \hat{H}) belongs (see Fig. 8.2) to the set Ω which is the common part of the following sets:

⁵The analogy of a house and bricks fails here, because both the house and the bricks come from the same 3D space.
⁶The normalization condition does not reduce the generality of the approach.
⁷That is, we put $\sigma = \frac{1}{2}$ or $\sigma = -\frac{1}{2}$.

Fig. 8.2. Diagram of the sets, among which the solution ψ of the Schrödinger equation is sought. The Q set is the one of all square-integrable functions, $\Omega_{\text{variables}}$ is the set of the functions with variables as those of the solution of Schrödinger equation, ψ , and Ω_{antisym} is the set of the functions which are antisymmetric with respect to the exchange of coordinates of any two electrons. The solutions of the Schrödinger equation, ψ , will be sought in the common part of these three sets: $\psi \in \Omega = Q \cap \Omega_{\text{variables}} \cap \Omega_{\text{antisym}}$. The Ω_{Slater} represents the set of single Slater determinants built of normalizable spinorbitals. The exact wave function ψ always belongs to $\Omega - \Omega_{\text{Slater}}$.



- set Q of all square-integrable functions,
- set $\Omega_{\text{variables}}$ of all the functions dependent on the above mentioned variables,
- set Ω_{antisym} of all the functions which are antisymmetric with respect to the mutual exchange of the coordinates of any two electrons (p. 33).

$$\psi \in \Omega - \Omega_{\text{Slater}} \quad \text{with} \quad \Omega = Q \cap \Omega_{\text{variables}} \cap \Omega_{\text{antisym}}.$$

John C. Slater (1901–1976), American physicist, for 30 years a professor and dean at the Physics Department of the Massachusetts Institute of Technology, then at the University of Florida Gainesville and the Quantum Theory Project at this school. His youth was in the stormy period of the intense development of quantum mechanics, and he participated vividly in it. For example, in 1926–1932 he published articles on the ground state of the helium atom, on the screening constants (Slater orbitals), on the antisymmetrization of the wave function (Slater determinant), and on the algorithm for calculating the integrals (the Slater–Condon rules). In this period he made the acquaintance of John Van Vleck, Robert Mulliken, Arthur Compton, Edward Condon and Linus Pauling. In Europe (Zurich and Leipzig) he exchanged ideas with Werner Heisenberg,



Friedrich Hund, Peter Debye, Felix Bloch, Douglas Hartree, Eugene Wigner, Albert Einstein, Erich Hückel, Edward Teller, Nevil Mott, and John Lennard-Jones. The frequency of the appearance of his name in this book is the best testament to his great contribution to quantum chemistry.

8.1.3 SLATER DETERMINANTS

There should be something in the theory which assures us that, if we renumber the electrons, no theoretical prediction will change. The postulate of the *antisymmetric* character of the wave function with respect to the *exchange* of the coordinates of *any* two electrons, certainly ensures this (Chapter 1, p. 28). The solution of the Schrödinger equation for a given stationary state of interest should be sought amongst *such* functions.

A *Slater determinant* is a function of the coordinates of N electrons, which *automatically* belongs to Ω :

$$\psi = \frac{1}{\sqrt{N!}} \begin{vmatrix} \phi_1(1) & \phi_1(2) & \dots & \phi_1(N) \\ \phi_2(1) & \phi_2(2) & \dots & \phi_2(N) \\ \dots & \dots & \dots & \dots \\ \phi_N(1) & \phi_N(2) & \dots & \phi_N(N) \end{vmatrix},$$

where ϕ_i are the orthonormal⁸ one-electron⁹ functions, i.e. molecular spinorbitals. The Slater determinants form a subset $\Omega_{\text{Slater}} \subset \Omega$.

A Slater determinant carries two important attributes of the exact wave function:

- Suppose we want to calculate the probability density that two electrons with the same spin coordinate σ are in the same place, i.e. such that two electrons have *all* their coordinates (spatial and spin ones) identical. If so, then the two columns of the above mentioned determinant are identical. And this means that the determinant becomes equal to zero.¹⁰ From this and from the continuity of the wave function we may conclude that:

electrons of the same spin cannot approach each other.

- Let us now imagine two electrons with opposite values of their spin coordinate σ . If these two electrons take the same position in space, the Slater determinant will not vanish, because in the general case there is nothing that forces $\phi_i(1)$ to be equal to $\phi_i(2)$, when $1 \equiv (r_1, \sigma = \frac{1}{2})$ and $2 \equiv (r_1, \sigma = -\frac{1}{2})$ for

⁸It is *most often* so, and then the factor standing in front of the determinant ensures normalization. The spinorbitals could be *non-normalized* (but, if they are to describe a stationary state, they should be square-integrable). They also do not need to be mutually orthogonal, but certainly they *need to be linearly independent*. Any attempt to insert the linearly-dependent functions in the determinant will have a “tragic outcome” – we will get 0. It comes from the properties of the determinant (if a row is a linear combination of the others, the determinant is zero). It also follows that if we have a set of non-orthogonal spinorbitals in a Slater determinant, we could orthogonalize them by making the appropriate linear combinations. This would multiply the original Slater determinant by an irrelevant constant. This is why it is *no loss of generality to require the spinorbitals to be orthonormal*.

⁹In the theory of the atomic nucleus, the determinant wave function for the nucleons (fermions) is also used.

¹⁰Indeed, this is why we exist. Two objects built out of fermions (e.g., electrons) cannot occupy the same position in space. If it were not so, our bodies would sink in the ground.

$i = 1, 2, \dots$. From this, and from the continuity of the wave function, we conclude that:

electrons of opposite spins can approach each other.

8.1.4 WHAT IS THE HARTREE–FOCK METHOD ALL ABOUT?

The Hartree–Fock method is a *variational* one (p. 196) and uses the variational wave function in the form of a *single Slater determinant*.

In other words we seek (among the Ω_{Slater} set of trial functions) the determinant (ψ_{HF}), which results in the lowest mean value of the Hamiltonian.

In this case the mathematical form of the spinorbitals undergoes variation – change $\varphi_{i1}(\mathbf{r})$ as well as $\varphi_{i2}(\mathbf{r})$ in eq. (8.1) (however you want) to try to lower the mean value of the Hamiltonian as much as possible. The output determinant which provides the *minimum* mean value of the Hamiltonian is called the Hartree–Fock function. The Hartree–Fock function is an *approximation* of the true wave function (which satisfies the Schrödinger equation $H\psi = E\psi$), because the former is indeed the optimal solution, but *only among single Slater determinants*. The Slater determinant is an antisymmetric function, but an antisymmetric function does not necessarily need to take the shape of a Slater determinant.

Taking the variational wave function in the form of one determinant means an automatic limitation to the subset Ω_{Slater} for searching for the optimum wave function. In fact, we should search the optimum wave function in the set $\Omega - \Omega_{\text{Slater}}$. Thus it is an *approximation* for the solution of the Schrödinger equation, with no chance of representing the exact result.

The true solution of the Schrödinger equation is never a single determinant.

Why are Slater determinants used so willingly? There are two reasons for this:

- a determinant is a kind of “template”.¹¹ Whatever you put inside, the result (if not zero) is antisymmetric *by definition*, i.e. it automatically satisfies one of the postulates of quantum mechanics.
- it is constructed out of simple “bricks” – the one-electron functions (spinorbitals).

The Slater determinants built out of the complete set of spinorbitals do form the complete set.

¹¹ An interesting analogy to the history of algebra appears here. The matrix (lat. *matrix*) took its name from the printing stamp, because the latter indeed served the inventor of matrix algebra, James Joseph Sylvester (1814–1897), for automatically “cutting out” the determinants.

Because of this, the true wave function can take the form of a linear combination of the determinants (we will discuss this later in Chapter 10).

8.2 THE FOCK EQUATION FOR OPTIMAL SPINORBITALS

8.2.1 DIRAC AND COULOMB NOTATIONS

The spatial *and* spin coordinate integrals (ϕ are the spinorbitals, φ are the orbitals) in the *Dirac notation* will be denoted with angle brackets $\langle \rangle$ (\hat{h} denotes a one-electron operator and r_{12} – the distance between electrons 1 and 2, $dV_1 = dx_1 dy_1 dz_1$, $dV_2 = dx_2 dy_2 dz_2$), for the one-electron integrals:

$$\langle i|\hat{h}|j\rangle \equiv \sum_{\sigma_1} \int dV_1 \phi_i^*(1) \hat{h} \phi_j(1) \equiv \int d\tau_1 \phi_i^*(1) \hat{h} \phi_j(1), \quad (8.2)$$

and for the two-electron integrals:

$$\begin{aligned} \langle ij|kl\rangle &\equiv \sum_{\sigma_1} \sum_{\sigma_2} \int dV_1 \int dV_2 \phi_i^*(1) \phi_j^*(2) \frac{1}{r_{12}} \phi_k(1) \phi_l(2) \\ &\equiv \int d\tau_1 d\tau_2 \phi_i^*(1) \phi_j^*(2) \frac{1}{r_{12}} \phi_k(1) \phi_l(2). \end{aligned} \quad (8.3)$$

The integrals over the spatial (only) coordinates will be denoted by round brackets $\langle \rangle$, for the one-electron integrals:

$$\langle i|\hat{h}|j\rangle \equiv \int dV_1 \varphi_i^*(1) \hat{h}(1) \varphi_j(1), \quad (8.4)$$

and for the two-electron integrals:

$$\langle ij|kl\rangle \equiv \int dV_1 \int dV_2 \varphi_i^*(1) \varphi_j^*(2) \frac{1}{r_{12}} \varphi_k(1) \varphi_l(2). \quad (8.5)$$

This is called Dirac notation (of the integrals).¹²

8.2.2 ENERGY FUNCTIONAL

Applying the first Slater–Condon rule¹³ we get the following equation for the *mean value of Hamiltonian* (without nuclear repulsion) calculated using the normalized

¹²Sometimes one uses *Coulomb notation* $\langle ij|kl\rangle_{\text{Dirac}} \equiv \langle ik|jl\rangle_{\text{Coulomb}}$, also $\langle ij|kl\rangle_{\text{Dirac}} \equiv \langle ik|jl\rangle_{\text{Coulomb}}$. Coulomb notation emphasizes the physical interpretation of the two electron integral, as the energy of the Coulombic interaction of two charge distributions $\varphi_i^*(1)\varphi_k(1)$ for electron 1 and $\varphi_j^*(2)\varphi_l(2)$ for electron 2. Dirac notation for the two-electron integrals emphasizes the two-electron functions “bra” and “ket” from the general Dirac notation (p. 19). *In the present book we will consequently use Dirac notation* (both for integrals using spinorbitals, and for those using orbitals, the difference being emphasized by the type of bracket). Sometimes the self-explaining notation $\langle i|\hat{h}|j\rangle \equiv \langle \phi_i|\hat{h}|\phi_j\rangle$, etc. will be used.

¹³Appendix M, p. 986; please take a look at this rule (you may leave out its derivation).

Slater one-determinant function ψ , i.e. the energy functional $E[\psi]$

$$E[\psi] = \langle \psi | \hat{H} | \psi \rangle = \sum_{i=1}^N \langle i | \hat{h} | i \rangle + \frac{1}{2} \sum_{i,j=1}^N (\langle ij | ij \rangle - \langle ij | ji \rangle), \quad (8.6)$$

where the indices symbolize the spinorbitals, and the symbol \hat{h}

$$\hat{h}(1) = -\frac{1}{2}\Delta_1 - \sum_{a=1}^M \frac{\mathcal{Z}_a}{r_{a1}} \quad (8.7)$$

is the one-electron operator (in atomic units) of the kinetic energy of the electron plus the operator of the nucleus–electron attraction (there are M nuclei).

8.2.3 THE SEARCH FOR THE CONDITIONAL EXTREMUM

We would like to find such spinorbitals (“the best ones”), that *any* change in them leads to an *increase* in energy $E[\psi]$. But the changes of the spinorbitals need to be such that the above formula still holds, and it would hold only by assuming the *orthonormality* of the spinorbitals. This means that there are some constraints for the changed spinorbitals:

$$\langle i | j \rangle - \delta_{ij} = 0 \quad \text{for } i, j = 1, 2, \dots, N. \quad (8.8)$$

Thus we seek the *conditional* minimum. We will find it using the Lagrange multipliers method (Appendix N, p. 997). In this method the equations of the constraints multiplied by the Lagrange multipliers are added to the original function which is to be minimized. Then we minimize the function as if the constraints did not exist.

conditional
minimum

We do the same for the functionals. The necessary condition for the minimum is that the variation¹⁴ of

stationary points

$$E - \sum_{ij} L_{ij} (\langle i | j \rangle - \delta_{ij})$$

is equal zero.

The variation of a functional is defined as the *linear* part of the functional change coming from a change in the function which is its argument.

¹⁴However, this is not a sufficient condition, because the vanishing of the differential for certain values of independent variables happens not only for minima, but also for maxima and saddle points (*stationary points*).

Variation is an analogue of the differential (the differential is just the linear part of the function's change). Thus we calculate the linear part of a change (variation):

$$\delta \left(E - \sum_{ij} L_{ij} \langle i|j \rangle \right) = 0 \quad (8.9)$$

using the (yet) undetermined Lagrange multipliers L_{ij} and we set the variation equal to zero.¹⁵

The stationarity condition for the energy functional

It is sufficient to vary *only the functions complex conjugate to the spinorbitals* or only the spinorbitals (cf. p. 197), yet the result is always the same. We decide the first.

Substituting $\phi_i^* \rightarrow \phi_i^* + \delta\phi_i^*$ in (8.6) (and retaining only linear terms in $\delta\phi_i^*$ to be inserted into (8.9)) the variation takes the form (the symbols δi^* and δj^* mean $\delta\phi_i^*$ and $\delta\phi_j^*$)

$$\sum_{i=1}^N \left(\langle \delta i | \hat{h} | i \rangle + \frac{1}{2} \sum_{ij} (\langle \delta i, j | ij \rangle + \langle i, \delta j | ij \rangle - \langle \delta i, j | ji \rangle - \langle i, \delta j | ji \rangle - 2L_{ij} \langle \delta i | j \rangle) \right) = 0. \quad (8.10)$$

Now we will try to express this in the form:

$$\sum_{i=1}^N \langle \delta i | \dots \rangle = 0.$$

Since the δi^* may be arbitrary, the equation $|\dots\rangle = 0$ (called the Euler equation in variational calculus), results. This will be our next goal.

Noticing that the sum indices and the numbering of electrons in the integrals are arbitrary we have the following equalities

$$\begin{aligned} \sum_{ij} \langle i, \delta j | ij \rangle &= \sum_{ij} \langle j, \delta i | ji \rangle = \sum_{ij} \langle \delta i, j | ij \rangle, \\ \sum_{ij} \langle i, \delta j | ji \rangle &= \sum_{ij} \langle j, \delta i | ij \rangle = \sum_{ij} \langle \delta i, j | ji \rangle, \end{aligned}$$

and after substitution in the expression for the variation, we get

$$\sum_i \left(\langle \delta i | \hat{h} | i \rangle + \frac{1}{2} \sum_j (\langle \delta i, j | ij \rangle + \langle \delta i, j | ij \rangle - \langle \delta i, j | ji \rangle - \langle \delta i, j | ji \rangle - 2L_{ij} \langle \delta i | j \rangle) \right) = 0. \quad (8.11)$$

¹⁵Note that $\delta(\delta_{ij}) = 0$.

Let us rewrite this equation in the following manner:

$$\sum_i \left\langle \delta i \left| \left(\hat{h} \phi_i(1) + \sum_j \left(\int d\tau_2 \frac{1}{r_{12}} \phi_j^*(2) \phi_j(2) \phi_i(1) - \int d\tau_2 \frac{1}{r_{12}} \phi_j^*(2) \phi_i(2) \phi_j(1) - L_{ij} \phi_j(1) \right) \right) \right| \right\rangle_1 = 0, \quad (8.12)$$

where $\langle \delta i | \dots \rangle_1$ means integration over coordinates of electron 1 and $d\tau_2$ refers to *the spatial coordinate integration and spin coordinate summing for electron 2*. The above must be true for *any* $\delta i^* \equiv \delta \phi_i^*$, which means that *each individual term* in parentheses needs to be equal to zero:

$$\begin{aligned} \hat{h} \phi_i(1) + \sum_j \left(\int d\tau_2 \frac{1}{r_{12}} \phi_j^*(2) \phi_j(2) \cdot \phi_i(1) - \int d\tau_2 \frac{1}{r_{12}} \phi_j^*(2) \phi_i(2) \cdot \phi_j(1) \right) \\ = \sum_j L_{ij} \phi_j(1). \end{aligned} \quad (8.13)$$

The Coulombic and exchange operators

Let us introduce the following linear operators:

- a) *two Coulombic operators: the total operator $\hat{J}(1)$ and the orbital operator $\hat{J}_j(1)$, defined via their action on an arbitrary function $u(1)$ of the coordinates of electron 1*

Coulombic and exchange operators

$$\hat{J}(1)u(1) = \sum_j \hat{J}_j(1)u(1) \quad (8.14)$$

$$\hat{J}_j(1)u(1) = \int d\tau_2 \frac{1}{r_{12}} \phi_j^*(2) \phi_j(2) u(1) \quad (8.15)$$

- b) *and similarly, two exchange operators: the total operator $\hat{K}(1)$ and the orbital operator $\hat{K}_j(1)$*

$$\hat{K}(1)u(1) = \sum_j \hat{K}_j(1)u(1) \quad (8.16)$$

$$\hat{K}_j(1)u(1) = \int d\tau_2 \frac{1}{r_{12}} \phi_j^*(2) u(2) \phi_j(1). \quad (8.17)$$

Then eq. (8.13) takes the form

$$(\hat{h}(1) + \hat{J}(1) - \hat{K}(1))\phi_i(1) = \sum_j L_{ij} \phi_j(1). \quad (8.18)$$

The equation is nice and concise except for one thing. It would be even nicer if the right-hand side were proportional to $\phi_i(1)$ instead of being a linear combination of all the spinorbitals. In such a case the equation would be similar to the eigenvalue problem and we would like it a lot. It would be similar but not identical, since the operators \hat{J} and \hat{K} include the *sought* spinorbitals ϕ_i . Because of this, the equation would be called the pseudo-eigenvalue problem.

8.2.4 A SLATER DETERMINANT AND A UNITARY TRANSFORMATION

How can we help? Let us notice that we do not care too much about the spinorbitals themselves, because these are *by-products* of the method which is to give the optimum mean value of the Hamiltonian, and the corresponding N -electron wave function. We can choose some other spinorbitals, such that the mean value of the Hamiltonian does not change and the Lagrange multipliers matrix is diagonal. Is this at all possible? Let us see.

Let us imagine the linear transformation of spinorbitals ϕ_i , i.e. in matrix notation:

$$\phi' = A\phi, \quad (8.19)$$

where ϕ and ϕ' are vertical vectors containing components ϕ_i . A vertical vector is uncomfortable for typography, in contrast to its transposition (a horizontal vector), and it is easier to write the transposed vector: $\phi'^T = [\phi'_1, \phi'_2, \dots, \phi'_N]$ and $\phi^T = [\phi_1, \phi_2, \dots, \phi_N]$. If we construct the determinant built of spinorbitals ϕ' and not of ϕ , an interesting chain of transformations will result:

$$\begin{aligned} & \frac{1}{\sqrt{N!}} \begin{vmatrix} \phi'_1(1) & \phi'_1(2) & \dots & \phi'_1(N) \\ \phi'_2(1) & \phi'_2(2) & \dots & \phi'_2(N) \\ \dots & \dots & \dots & \dots \\ \phi'_N(1) & \phi'_N(2) & \dots & \phi'_N(N) \end{vmatrix} \\ &= \frac{1}{\sqrt{N!}} \begin{vmatrix} \sum_i A_{1i}\phi_i(1) & \dots & \sum_i A_{1i}\phi_i(N) \\ \sum_i A_{2i}\phi_i(1) & \dots & \sum_i A_{2i}\phi_i(N) \\ \dots & \dots & \dots \\ \sum_i A_{Ni}\phi_i(1) & \dots & \sum_i A_{Ni}\phi_i(N) \end{vmatrix} \end{aligned} \quad (8.20)$$

$$\begin{aligned} &= \det \left\{ A \frac{1}{\sqrt{N!}} \begin{bmatrix} \phi_1(1) & \phi_1(2) & \dots & \phi_1(N) \\ \phi_2(1) & \phi_2(2) & \dots & \phi_2(N) \\ \dots & \dots & \dots & \dots \\ \phi_N(1) & \phi_N(2) & \dots & \phi_N(N) \end{bmatrix} \right\} \\ &= \det A \cdot \frac{1}{\sqrt{N!}} \begin{vmatrix} \phi_1(1) & \phi_1(2) & \dots & \phi_1(N) \\ \phi_2(1) & \phi_2(2) & \dots & \phi_2(N) \\ \dots & \dots & \dots & \dots \\ \phi_N(1) & \phi_N(2) & \dots & \phi_N(N) \end{vmatrix}. \end{aligned} \quad (8.21)$$

We have therefore obtained our initial Slater determinant multiplied by a *number*: $\det \mathbf{A}$. Thus, provided that $\det \mathbf{A}$ is not zero,¹⁶

the new wave function would provide the same mean value of the Hamiltonian.

The only problem from such a transformation is loss of the normalization of the wave function. Yet we may even preserve the normalization. Let us choose such a matrix \mathbf{A} , that $|\det \mathbf{A}| = 1$. This condition will hold if $\mathbf{A} = \mathbf{U}$, where \mathbf{U} is a *unitary* matrix.¹⁷ This means that

if a unitary transformation \mathbf{U} is performed on the orthonormal spinorbitals (when \mathbf{U} is real, we call \mathbf{U} an *orthogonal* transformation), then the new spinorbitals ϕ' are also orthonormal.

This is why a unitary transformation is said to represent a rotation in the Hilbert space: the mutually orthogonal and perpendicular vectors do not lose these features upon rotation.¹⁸ This can be verified by a direct calculation:

$$\begin{aligned} \langle \phi'_i(1) | \phi'_j(1) \rangle &= \left\langle \sum_r U_{ir} \phi_r(1) \left| \sum_s U_{js} \phi_s(1) \right. \right\rangle \\ &= \sum_{rs} U_{ir}^* U_{js} \langle \phi_r(1) | \phi_s(1) \rangle = \sum_{rs} U_{ir}^* U_{js} \delta_{rs} \\ &= \sum_r U_{ir}^* U_{jr} = \delta_{ij}. \end{aligned}$$

Thus, in the case of a unitary transformation even the normalization of the total one-determinant wave function is preserved; at worst the phase χ of this function will change (while $\exp(i\chi) = \det \mathbf{U}$), and this factor does not change either $|\psi|^2$ or the mean value of the operators.

8.2.5 INVARIANCE OF THE \hat{J} AND \hat{K} OPERATORS

How does the Coulombic operator change upon a unitary transformation of the spinorbitals? Let us see,

$$\hat{J}(1)' \chi(1) = \int d\tau_2 \frac{1}{r_{12}} \sum_j \phi_j'^*(2) \phi_j'(2) \chi(1)$$

¹⁶The \mathbf{A} transformation thus cannot be singular (see Appendix A, p. 889).

¹⁷For a unitary transformation $\mathbf{U}\mathbf{U}^\dagger = \mathbf{U}^\dagger\mathbf{U} = \mathbf{1}$. The matrix \mathbf{U}^\dagger arises from \mathbf{U} via the exchange of rows and columns (this does not influence the value of the determinant), and via the complex conjugation of all elements (and this gives $\det \mathbf{U}^\dagger = (\det \mathbf{U})^*$). Finally, since $(\det \mathbf{U})(\det \mathbf{U}^\dagger) = 1$ we have $|\det \mathbf{U}| = 1$.

¹⁸Just as three fingers held at right angles do not cease to be of the same length (normalization) after rotation of your palm and continue to be orthogonal.

$$\begin{aligned}
&= \int d\tau_2 \frac{1}{r_{12}} \sum_j \sum_r U_{jr}^* \phi_r^*(2) \sum_s U_{js} \phi_s(2) \chi(1) \\
&= \int d\tau_2 \frac{1}{r_{12}} \sum_{r,s} \left(\sum_j U_{js} U_{jr}^* \right) \phi_r^*(2) \phi_s(2) \chi(1) \\
&= \int d\tau_2 \frac{1}{r_{12}} \sum_{r,s} \left(\sum_j U_{rj}^\dagger U_{js} \right) \phi_r^*(2) \phi_s(2) \chi(1) \\
&= \int d\tau_2 \frac{1}{r_{12}} \sum_{r,s} \delta_{sr} \phi_r^*(2) \phi_s(2) \chi(1) \\
&= \int d\tau_2 \frac{1}{r_{12}} \sum_r \phi_r^*(2) \phi_r(2) \chi(1) = \hat{J}(1) \chi(1).
\end{aligned}$$

The operator $\hat{J}(1)'$ proves to be *identical* with the operator $\hat{J}(1)$. Similarly we may prove the *invariance of the operator K* .

The operators \hat{J} and \hat{K} are invariant with respect to any unitary transformation of the spinorbitals.

In conclusion, while deriving the new spinorbitals from a unitary transformation of the old ones, we do not need to worry about \hat{J} and \hat{K} since they remain the same.

8.2.6 DIAGONALIZATION OF THE LAGRANGE MULTIPLIERS MATRIX

Eq. (8.18) may be written in matrix form:

$$[\hat{h}(1) + \hat{J}(1) - \hat{K}(1)]\phi(1) = L\phi(1), \quad (8.22)$$

where ϕ is a column of spinorbitals. Transforming $\phi = U\phi'$ and multiplying the Fock equation by U^\dagger (where U is a unitary matrix), we obtain

$$U^\dagger[\hat{h}(1) + \hat{J}(1) - \hat{K}(1)]U\phi(1)' = U^\dagger L U \phi(1)', \quad (8.23)$$

because \hat{J} and \hat{K} did not change upon the transformation.

The U matrix *may* be chosen such that $U^\dagger L U$ is the *diagonal* matrix.

Its diagonal elements¹⁹ will now be denoted as ε_i . Because $\hat{h}(1) + \hat{J}(1) - \hat{K}(1)$ is a linear operator we get equation

¹⁹Such diagonalization is possible because L is a Hermitian matrix (i.e. $L^\dagger = L$), and each Hermitian matrix may be diagonalized *via* the transformation $U^\dagger L U$ with the unitary matrix U . Matrix L is indeed Hermitian. It is clear when we write the complex conjugate of the variation $\delta(E - \sum_{ij} L_{ij} \langle i|j \rangle) = 0$. This gives $\delta(E - \sum_{ij} L_{ij}^* \langle j|i \rangle) = 0$, because E is real, and after the change of the summation indices $\delta(E - \sum_{ij} L_{ji}^* \langle i|j \rangle) = 0$. Thus, $L_{ij} = L_{ji}^*$, i.e. $L = L^\dagger$.

$$U^\dagger U(\hat{h}(1) + \hat{J}(1) - \hat{K}(1))\phi(1)' = U^\dagger LU\phi(1)' \quad (8.24)$$

or alternatively

$$(\hat{h}(1) + \hat{J}(1) - \hat{K}(1))\phi(1)' = \varepsilon\phi(1)', \quad (8.25)$$

where $\varepsilon_{ij} = \varepsilon_i\delta_{ij}$.

8.2.7 THE FOCK EQUATION FOR OPTIMAL SPINORBITALS (GENERAL HARTREE-FOCK METHOD – GHF)

We leave out the “prime” to simplify the notation²⁰ and write *the Fock equation* for a single spinorbital:

THE FOCK EQUATION IN THE GENERAL HARTREE-FOCK METHOD (GHF)

$$\hat{F}(1)\phi_i(1) = \varepsilon_i\phi_i(1), \quad (8.26)$$

where the Fock operator \hat{F} is

$$\hat{F}(1) = \hat{h}(1) + \hat{J}(1) - \hat{K}(1). \quad (8.27)$$

These ϕ_i are called *canonical spinorbitals*, and are the solution of the Fock equation, ε_i is the *orbital energy* corresponding to the spinorbital ϕ_i . It is indicated in brackets that both the Fock operator and the molecular spinorbital depend on the coordinates of one electron only (exemplified as electron 1).²¹

Fock operator
canonical
spin-orbitals
orbital energy

²⁰This means that we finally forget about ϕ' (we pretend that they have never appeared), and we will deal only with such ϕ as correspond to the *diagonal* matrix of the Lagrange multipliers.

²¹The above derivation looks more complex than it really is. The essence of the whole machinery will now be shown as exemplified by two coupled (bosonic) harmonic oscillators, with the Hamiltonian $\hat{H} = \hat{T} + \hat{V}$ where $\hat{T} = -\frac{\hbar^2}{2m_1}\frac{\partial^2}{\partial x_1^2} - \frac{\hbar^2}{2m_2}\frac{\partial^2}{\partial x_2^2}$ and $V = \frac{1}{2}kx_1^2 + \frac{1}{2}kx_2^2 + \lambda x_1^4 x_2^4$, with $\lambda x_1^4 x_2^4$ as the coupling term. Considering the bosonic nature of the particles (the wave function is symmetric, see Chapter 1), we will use $\psi = \phi(1)\phi(2)$ as a variational function, where ϕ is a normalized spinorbital. The expression for the mean value of the Hamiltonian takes the form

$$\begin{aligned} E[\phi] &= \langle \psi | \hat{H} | \psi \rangle = \langle \phi(1)\phi(2) | (\hat{h}(1) + \hat{h}(2))\phi(1)\phi(2) \rangle + \lambda \langle \phi(1)\phi(2) | x_1^4 x_2^4 \phi(1)\phi(2) \rangle \\ &= \langle \phi(1)\phi(2) | \hat{h}(1)\phi(1)\phi(2) \rangle + \langle \phi(1)\phi(2) | \hat{h}(2)\phi(1)\phi(2) \rangle + \lambda \langle \phi(1) | x_1^4 \phi(1) \rangle \langle \phi(2) | x_2^4 \phi(2) \rangle \\ &= \langle \phi(1) | \hat{h}(1)\phi(1) \rangle + \langle \phi(2) | \hat{h}(2)\phi(2) \rangle + \lambda \langle \phi(1) | x_1^4 \phi(1) \rangle \langle \phi(2) | x_2^4 \phi(2) \rangle \\ &= 2\langle \phi | \hat{h} \phi \rangle + \lambda \langle \phi | x^4 \phi \rangle^2, \end{aligned}$$

where one-particle operator $\hat{h}(i) = -\frac{\hbar^2}{2m_i}\frac{\partial^2}{\partial x_i^2} + \frac{1}{2}kx_i^2$.

The change of E , because of the variation $\delta\phi^*$, is $E[\phi + \delta\phi] - E[\phi] = 2\langle \phi + \delta\phi | \hat{h} \phi \rangle + \lambda \langle \phi + \delta\phi | x^4 \phi \rangle^2 - [2\langle \phi | \hat{h} \phi \rangle + \lambda \langle \phi | x^4 \phi \rangle^2] = 2\langle \phi | \hat{h} \phi \rangle + 2\langle \delta\phi | \hat{h} \phi \rangle + \lambda \langle \phi | x^4 \phi \rangle^2 + 2\lambda \langle \delta\phi | x^4 \phi \rangle \langle \phi | x^4 \phi \rangle + \lambda \langle \delta\phi | x^4 \phi \rangle^2 - [2\langle \phi | \hat{h} \phi \rangle + \lambda \langle \phi | x^4 \phi \rangle^2]$.

Unrestricted Hartree–Fock method (UHF)

The GHF method derived here is usually presented in textbooks as the *unrestricted Hartree–Fock method* (UHF). Despite its name, UHF is not a fully unrestricted method (as the GHF is). In the UHF we assume (cf. eq. (8.1)):

- orbital components φ_{i1} and φ_{i2} are *real* and
- *there is no mixing of the spin functions* α and β , i.e. either $\varphi_{i1} = 0$ and $\varphi_{i2} \neq 0$ or $\varphi_{i1} \neq 0$ and $\varphi_{i2} = 0$.

8.2.8 THE CLOSED-SHELL SYSTEMS AND THE RESTRICTED HARTREE–FOCK (RHF) METHOD

Double occupation of the orbitals and the Pauli exclusion principle

When the number of electrons is even, the spinorbitals are usually formed out of orbitals in a very easy (and simplified with respect to eq. (8.1)) manner, by multiplication of each orbital by the spin functions²² α or β :

$$\phi_{2i-1}(\mathbf{r}, \sigma) = \varphi_i(\mathbf{r})\alpha(\sigma) \quad (8.28)$$

$$\phi_{2i}(\mathbf{r}, \sigma) = \varphi_i(\mathbf{r})\beta(\sigma), \quad i = 1, 2, \dots, \frac{N}{2}, \quad (8.29)$$

where – as it can be clearly seen – there are twice as few occupied orbitals φ as occupied spinorbitals ϕ (*occupation means that a given spinorbital appears in the Slater determinant*²³) (see Fig. 8.3). Thus we introduce an artificial *restriction* for spinorbitals (some of the consequences will be described on p. 369). This is why the method is called the Restricted Hartree–Fock.

There are as many spinorbitals as electrons, and therefore there can be a maximum of two electrons per orbital.

If we wished to occupy a given orbital with more than two electrons, we would need once again to use the spin function α or β when constructing the spinorbitals,

Its linear part, i.e. the variation, is $\delta E = 2\langle\delta\phi|\hat{h}\phi\rangle + 2\lambda\langle\delta\phi|x^4\phi\rangle\langle\phi|x^4\phi\rangle$. The variation $\delta\phi^*$ has, however, to ensure the normalization of ϕ , i.e. $\langle\phi|\phi\rangle = 1$. After multiplying by the Lagrange multiplier 2ε , we get the extremum condition $\delta(E - 2\varepsilon\langle\phi|\phi\rangle) = 0$, i.e. $2\langle\delta\phi|\hat{h}\phi\rangle + 2\lambda\langle\delta\phi|x^4\phi\rangle\langle\phi|x^4\phi\rangle - 2\varepsilon\langle\delta\phi|\phi\rangle = 0$. This may be rewritten as $2\langle\delta\phi|[\hat{h} + \lambda\bar{x}^4x^4 - \varepsilon]\phi\rangle = 0$, where $\bar{x}^4 = \langle\phi|x^4\phi\rangle$, which gives ($\delta\phi^*$ is arbitrary) the Euler equation $[\hat{h} + \lambda\bar{x}^4x^4 - \varepsilon]\phi = 0$, i.e. the analogue of the Fock equation (8.27): $\hat{F}\phi = \varepsilon\phi$ with the operator $\hat{F} = [\hat{h} + \lambda\bar{x}^4x^4]$. Let us emphasize that the operator \hat{F} is a one-particle operator, *via* the notation $\hat{F}(1)\phi(1) = \varepsilon\phi(1)$, while $\hat{F}(1) = [\hat{h}(1) + \lambda\bar{x}^4x_1^4]$.

It is now clear what the mean field approximation is: the two-particle problem is reduced to a *single-particle* one (denoted as number 1), and the influence of the second particle is *averaged over its positions* ($\bar{x}^4 = \langle\phi|x^4\phi\rangle = \langle\phi(2)|x_2^4\phi(2)\rangle$).

²²It is not necessary, but quite comfortable. This means: $\phi_1 = \varphi_1\alpha$, $\phi_2 = \varphi_1\beta$, etc.

²³And only this. When the Slater determinant is written, the electrons lose their identity – they are not anymore distinguishable.

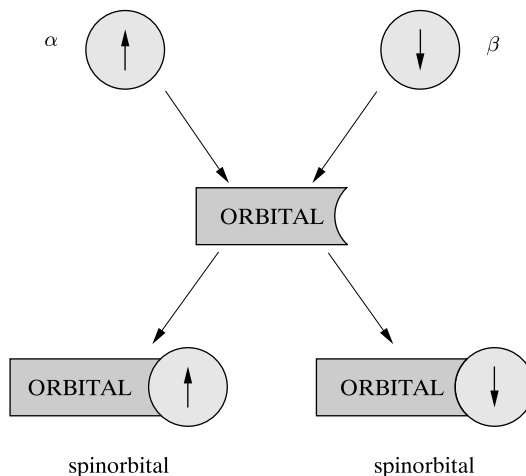


Fig. 8.3. Construction of a spinorbital in the RHF method (i.e. a function x, y, z, σ) as a *product* of an orbital (a function of x, y, z) and one of the two spin functions $\alpha(\sigma)$ or $\beta(\sigma)$.

i.e. repeating a spinorbital. This would imply two identical rows in the Slater determinant, and the wave function would equal zero. This cannot be accepted. The above rule of maximum double occupation is called the Pauli exclusion principle.²⁴ Such a formulation of the Pauli exclusion principle requires two concepts: the postulate of the antisymmetrization of the electronic wave function, p. 28, and double orbital occupancy. The first of these is of fundamental importance, the second is of a technical nature.²⁵

We often assume the double occupancy of orbitals within what is called the *closed shell*. The latter term has an approximate character (Fig. 8.4). It means that for the studied system, there is a large energy difference between HOMO and LUMO orbital energies.

closed shell

HOMO is the Highest Occupied Molecular Orbital, and LUMO is the Lowest Unoccupied Molecular Orbital. The unoccupied molecular orbitals are called *virtual orbitals*.

²⁴From “*Solid State and Molecular Theory*”, Wiley, London, 1975 by John Slater: “... I had a seminar about the work which I was doing over there – the only lecture of mine which happened to be in German. It has appeared that not only Heisenberg, Hund, Debye and young Hungarian PhD student Edward Teller were present, but also Wigner, Pauli, Rudolph Peierls and Fritz London, all of them on their way to winter holidays. Pauli, of course, behaved in agreement with the common opinion about him, and disturbed my lecture saying that “*he had not understood a single word out of it*”, but Heisenberg has helped me to explain the problem. (...) Pauli was extremely bound to his own way of thinking, similar to Bohr, who did not believe in the existence of photons. Pauli was a warriorlike man, a kind of dictator...”

²⁵The concept of orbitals, occupied by electron pairs, exists only in the mean field method. We will leave this idea in the future, and the Pauli exclusion principle should survive as a postulate of the antisymmetry of the electronic wave function (more generally speaking, of the wave function of fermions).

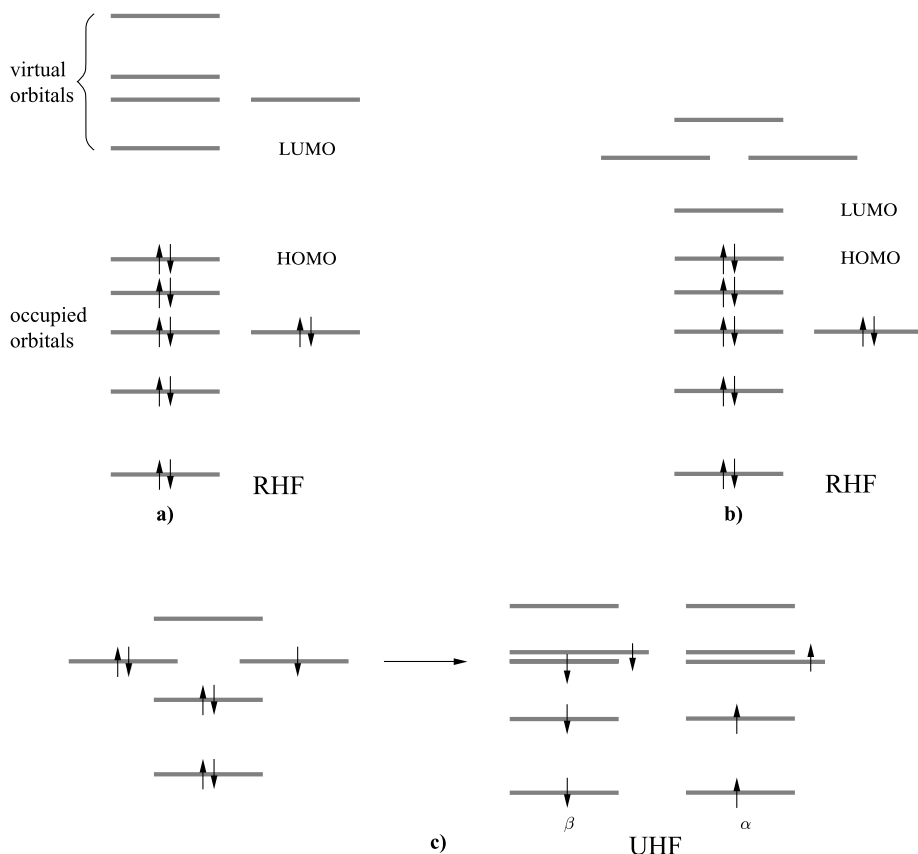


Fig. 8.4. The closed (a), poorly closed (b) and open (c) shell. The figure shows the occupancy of the molecular orbitals together with the corresponding spin functions (spin up and spin down for α and β functions): in case (a) and (b) the double occupancy of the lowest lying orbitals (on the energy scale) has been assumed; in the case (c) there is also an attempt to doubly occupy the orbitals (left-hand side), but a dilemma appears about which spinorbitals should be occupied. For example, in Fig. (c) we have decided to occupy the β spinorbital ("spin down"), but there is also a configuration with the α spinorbital ("spin up") of the same energy. This means that we need to use a scheme which allows different orbitals for different spins, e.g., UHF. The UHF procedure gives different orbitals energies for the α and β spins. One possibility is shown on the right-hand side of Fig. (c).

A CLOSED SHELL

A closed shell means that the HOMO is doubly occupied as are all the orbitals which are equal or lower in energy. The occupancy is such that the mathematical form of the Slater determinant does not depend on the spatial orientation of the x, y, z axis. Using group theory nomenclature (Appendix C), this function transforms according to *fully symmetric irreducible representation* of the symmetry group of the electronic Hamiltonian.

If a shell is not closed, it is called “open”.²⁶ We assume that there is a unique assignment for which molecular spinorbitals²⁷ within a closed shell are occupied in the ground state. The concept of the closed shell is approximate because it is not clear what it means when we say that the HOMO–LUMO energy distance²⁸ is large or small.²⁹

We need to notice that HOMO and LUMO have somewhat different meanings. As will be shown on p. 393, $-\varepsilon_{\text{HOMO}}$ represents an approximate ionization energy, i.e. binding energy of an electron interacting with the $(N - 1)$ -electron system, while $-\varepsilon_{\text{LUMO}}$ is an approximate electron affinity energy, i.e. energy of an electron interacting with the N -electron system.

The Fock equations for a closed shell (RHF method) can be derived in a very similar way as in the GHF method. This means the following steps:

- we write down the expression for the mean value of the Hamiltonian as a functional of the *orbitals* (the summation extends over all the *occupied orbitals*, there are $N/2$ of them, as will be recalled by the upper limit denoted by MO):³⁰

$$E = 2 \sum_i^{\text{MO}} (i|\hat{h}|i) + \sum_{i,j}^{\text{MO}} [2(ij|ij) - (ij|ji)];$$
- we seek the conditional minimum of this functional (Lagrange multipliers method) allowing for the variation of the *orbitals* which takes their orthonormality into account $\delta E = 2 \sum_i^{\text{MO}} (\delta i|\hat{h}|i) + \sum_{i,j}^{\text{MO}} [2(\delta ij|ij) - (\delta ij|ji) + 2(i\delta j|ij) - (i\delta j|ji)] - \sum_{i,j}^{\text{MO}} L'_{ij}(\delta i|j) = 0;$

²⁶Sometimes we use the term semi-closed shell, if it is half-occupied by the electrons and we are interested in the state bearing maximum spin. In this case the Slater determinant is a good approximation. The reasons for this is, of course, the uniqueness of electron assignment to various spinorbitals. *If there is no uniqueness (as in the carbon atom), then the single-determinant approximation cannot be accepted.*

²⁷The adjective “molecular” is suggested even for calculations for an atom. In a correct theory of electronic structure, the number of nuclei present in the system should not play any role. Thus, from the point of view of the computational machinery, an atom is just a molecule with one nucleus.

²⁸The decision to occupy only the lowest energy MOs (so called *Aufbau Prinzip*; a name left over from the German origins of quantum mechanics) is accepted under the *assumption* that the total energy differences are sufficiently well approximated by the differences in the orbital energies.

²⁹Unless the distance is zero. The helium atom, with the two electrons occupying the $1s$ orbital (HOMO), is a $1s^2$ shell of impressive “closure”, because the HOMO–LUMO energy difference calculated in a good quality basis set (6-31G**, see p. 364) of atomic orbitals is of the order of 62 eV. On the other hand, the HOMO–LUMO distance is zero for the carbon atom, because in the ground state 6 electrons occupy the $1s$, $2s$, $2p_x$, $2p_y$ and $2p_z$ orbitals. There is room for 10 electrons, and we only have six. Hence, the occupation (configuration) in the ground state is $1s^2 2s^2 2p^2$. Thus, both HOMO and LUMO are the $2p$ orbitals, with zero energy difference. If we asked for a single sentence describing why carbon compounds play a prominent role in Nature, it should be emphasized that, for carbon atoms, the HOMO–LUMO distance is equal to zero and that the orbital levels ε_{2s} and ε_{2p} are close in energy.

On the other hand, the beryllium atom is an example of a closed shell, which is not very tightly closed. Four electrons are in the lowest lying configuration $1s^2 2s^2$, but the orbital level $2p$ (LUMO) is relatively close to $2s$ (HOMO) (10 eV for the 6-31G** basis set is not a small gap, yet it amounts much less than that of the helium atom).

³⁰And not spinorbitals; see eqs. (M.17) and (M.18).

- we derive the Euler equation for this problem from $(\delta i | \dots) = 0$. In fact it is the Fock equation expressed in *orbitals*³¹

$$\hat{F}(1)\varphi_i(1) = \varepsilon_i\varphi_i(1), \quad (8.30)$$

closed-shell
Fock operator

where φ are the *orbitals*. The *Fock operator* is defined for the closed shell, as

$$\hat{F}(1) = \hat{h}(1) + 2\hat{J}(1) - \hat{K}(1), \quad (8.31)$$

where the first term (see eq. (8.7)) is the sum of the kinetic energy operator of electron 1 and the operator of the interaction of this electron with the nuclei in the molecule, the next two terms, i.e. Coulombic \hat{J} and exchange \hat{K} operators, are connected with the potential energy of the interaction of electron 1 with all electrons in the system, and they are defined (slightly differently than before for \hat{J} and \hat{K} operators³²) *via* the action on any function (χ) of the position of electron 1:

$$\begin{aligned} 2\hat{J}(1)\chi(1) &= \sum_{i=1}^{\text{MO}} 2\hat{J}_i(1)\chi(1) = \sum_{i=1}^{\text{MO}} 2 \int dV_2 \frac{1}{r_{12}} \varphi_i^*(2)\varphi_i(2)\chi(1) \\ &\equiv 2 \sum_i^{\text{MO}} \int dV_2 \frac{1}{r_{12}} |\varphi_i(2)|^2 \chi(1), \end{aligned} \quad (8.32)$$

$$\hat{K}(1)\chi(1) = \sum_{i=1}^{\text{MO}} \hat{K}_i(1)\chi(1) = \sum_{i=1}^{\text{MO}} \int dV_2 \frac{1}{r_{12}} \varphi_i^*(2)\chi(2)\varphi_i(1), \quad (8.33)$$

where integration is now exclusively over the spatial coordinates³³ of electron 2. Factor 2 multiplying the Coulombic operator results (as the reader presumably guessed) from the *double* occupation of the orbitals.

Interpretation of the Coulombic operator

The Coulombic operator is nothing else but a calculation of the Coulombic potential (with the opposite sign as created by all the *electrons*, Fig. 8.5) at the position of

³¹After a suitable unitary transformation of orbitals, analogous to what we have done in GHF case.

³²Because we have orbitals here, and not spinorbitals.

³³Simply, the summation over the spin coordinates has already been done when deriving the equation for the mean value of the Hamiltonian.

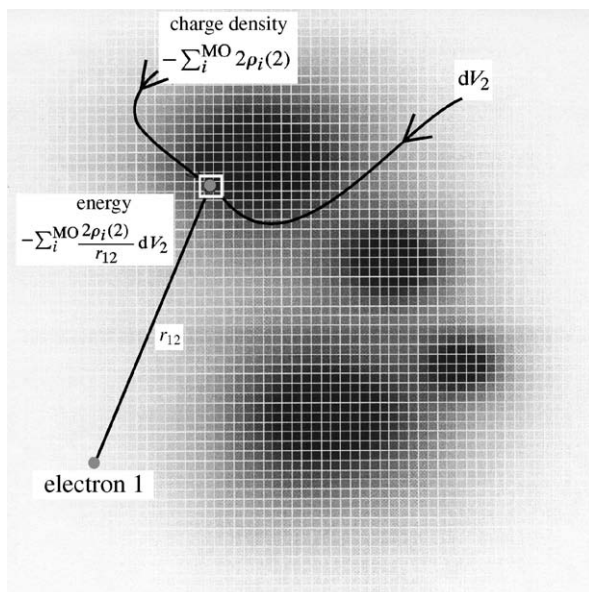


Fig. 8.5. Point-like electron 1 interacts with the total electron density (shown as electron cloud with density $\sum_i^{\text{MO}} 2\rho_i(2)$). To compute the interaction energy the total electron density is chopped into small cubes. The interaction energy of electron 1 with one of such cubes of volume dV_2 containing charge $-\sum_i^{\text{MO}} 2\rho_i(2) dV_2$ is calculated according to the Coulomb law: charge \times charge divided by their distance: $\frac{-1 \times (-1) \sum_i^{\text{MO}} 2\rho_i(2) dV_2}{r_{12}}$ or, alternatively, as charge -1 times electric potential produced by a single cube at electron 1. The summation over all cubes gives $\int \frac{\sum_i^{\text{MO}} 2\rho_i(2)}{r_{12}} dV_2 = 2\hat{J}$.

electron 1. Indeed, such a potential coming from an electron occupying molecular orbital φ_i is equal to

$$\int \frac{\rho_i(2)}{r_{12}} dV_2, \quad (8.34)$$

where $\rho_i(2) = \varphi_i(2)^* \varphi_i(2)$ is the probability density of finding electron 2 described by orbital φ_i . If we take into account that the orbital φ_i is occupied by *two* electrons, and that the number of the doubly occupied molecular orbitals is $N/2$, then the electrostatic potential calculated at the position of the electron 1 is

$$\int \frac{\sum_i^{\text{MO}} 2\rho_i(2)}{r_{12}} dV_2 = \sum_i^{\text{MO}} 2\hat{J}_i = 2\hat{J}(1).$$

The same expression also means an interaction of two *elementary* charges 1 and 2, one of each represented by a diffused cloud with a given charge density distribution $\rho(2) = \sum_i^{\text{MO}} 2\rho_i(2)$.

Integration in the formula for the operator \hat{J} is a consequence of the approximation of independent particles.

This approximation means that, in the Hartree–Fock method, the electrons do not move in the electric field of the other point-like electrons, but in the *mean static field of all the electrons* represented by electron cloud ρ . It is as if a driver (one of the electrons) in Paris did not use the position of other cars, but a map showing only the traffic intensity *via* the probability density cloud. The driver would then have a diffuse image of other vehicles,³⁴ and could not satisfactorily optimize the position towards other cars (it means higher energy for the molecule under study).

THE MEAN FIELD

This is typical for all the mean field methods. In these methods, *instead of watching the motion of other objects in detail, we average* these motions, and the problem simplifies (obviously, we pay the price of lower quality).

mean field

However, this trick is ingenious and worth remembering.³⁵

Coulombic self-interaction

There is a problem with this. From what we have said, it follows that the electron 1 uses the “maps” of *total* electron density, i.e. including its own contribution to the density.³⁶ This looks strange though. Let us take a closer look, maybe something has been missed in our reasoning. Note first of all that the repulsion of electron 1 (occupying, say, orbital k) with the electrons, which is visible in the Fock operator, reads as $(\varphi_k | (2\hat{J} - \hat{K}) \varphi_k)$ and not as $(\varphi_k | (2\hat{J}) \varphi_k)$. Let us write it down in more details:

$$\begin{aligned} (\varphi_k | (2\hat{J} - \hat{K}) \varphi_k) &= \int dV_1 |\varphi_k(1)|^2 \sum_{i=1}^{\text{MO}} 2 \int dV_2 \frac{1}{r_{12}} \varphi_i^*(2) \varphi_i(2) \\ &\quad - \sum_{i=1}^{\text{MO}} \int dV_1 \varphi_k(1)^* \varphi_i(1) \int dV_2 \frac{1}{r_{12}} \varphi_i^*(2) \varphi_k(2) \\ &= \int dV_1 |\varphi_k(1)|^2 \sum_{i=1}^{\text{MO}} 2 \int dV_2 \frac{1}{r_{12}} \varphi_i^*(2) \varphi_i(2) \end{aligned}$$

³⁴ An effect similar to the action of fog or alcohol. Both lead to miserable consequences.

³⁵ We use it every day, although we do not call it a mean field approach. Indeed, if we say: “*I will visit my aunt at noon, because it is easier to travel out of rush hours*”, or “*I avoid driving through the centre of town, because of the traffic jams*”, in practice we are using the mean field method. We average the motions of all citizens (including ourselves!) and we get a “map” (temporal or spatial), which allows us to optimize *our own* motion. The motion of our fellow-citizens *disappears*, and we obtain a *one-body* problem.

³⁶ Exactly as happens with real city traffic maps.

$$\begin{aligned}
& - \int dV_1 \varphi_k(1)^* \varphi_k(1) \int dV_2 \frac{1}{r_{12}} \varphi_k^*(2) \varphi_k(2) \\
& - \sum_{i(\neq k)}^{\text{MO}} \int dV_1 \varphi_k(1)^* \varphi_i(1) \int dV_2 \frac{1}{r_{12}} \varphi_i^*(2) \varphi_k(2) \\
& = \iint dV_1 dV_2 \frac{1}{r_{12}} \rho_k(1) [\rho(2) - \rho_k(2)] - \sum_{i(\neq k)}^{\text{MO}} (ki|ik),
\end{aligned}$$

where $\rho_k = |\varphi_k(1)|^2$, i.e. the distribution of electron 1 interacts electrostatically with all the *other* electrons,³⁷ i.e. with the distribution $[\rho(2) - \rho_k(2)]$ with ρ denoting the total electron density $\rho = \sum_{i=1}^{\text{MO}} 2|\varphi_i|^2$ and $-\rho_k$ excluding from it the self-interaction energy of the electron in question. Thus, the Coulombic and exchange operators together ensure that an electron interacts electrostatically with other electrons, not with itself.

Electrons with parallel spins repel less

There is also an exchange remainder $-\sum_{i(\neq k)}^{\text{MO}} (ki|ik)$, which is just a by-product of the antisymmetrization of the wave function (i.e. the Slater determinant), which tells us that in the Hartree–Fock picture electrons of the same spin functions³⁸ repel less. What??? As shown at the beginning of the present chapter, two electrons of the same spin cannot occupy the same point in space, and therefore (from the continuity of the wave function) they avoid each other. It is as if they repelled each other, because of the Pauli exclusion principle, *in addition* to their Coulombic repulsion. Is there something wrong in our result then? No, everything is OK. The necessary antisymmetric character of the wave function says simply that the same spins should keep apart. However, when the electrons described by the same spin functions keep apart, this obviously means *their Coulombic repulsion is weaker than that of electrons of opposite spins*.³⁹ This is what the term $-\sum_{i(\neq k)}^{\text{MO}} (ki|ik)$ really means.

Hartree method

The exchange operator represents a (non-intuitive) result of the antisymmetrization postulate for the total wave function (Chapter 1) and it has no classical inter-

³⁷The fact that the integration variables pertain to electron 2 is meaningless, it is just a definite integration and the name of the variable does not count at all.

³⁸When deriving the total energy expression (Appendix M), only those exchange terms survived, which correspond to the parallel spins of the interacting electrons. Note also, that for real orbitals (as in the RHF method), every exchange contribution $-(ki|ik) \equiv -\int dV_1 \varphi_k(1) \varphi_i(1) \int dV_2 \frac{1}{r_{12}} \varphi_i(2) \varphi_k(2)$ means a repulsion, because this is a self-interaction of the cloud $\varphi_k \varphi_i$.

³⁹Note that in the Hamiltonian, the Coulombic repulsion of the electrons is spin-independent. This suggests that when trying to improve the description (by going beyond the Hartree–Fock approximation), we have to worry more about correlation of electrons with the opposite spin functions (e.g., those occupying the same orbital).

pretation. If the variational wave function were the *product* of the spinorbitals⁴⁰ (Douglas Hartree did this in the beginning of the quantum chemistry)

$$\phi_1(1)\phi_2(2)\phi_3(3)\cdots\phi_N(N),$$

then we would get the corresponding Euler equation, which in this case is called the Hartree equation

$$\begin{aligned}\hat{F}_{\text{Hartree}}(1)\phi_i(1) &= \varepsilon_i\phi_i(1), \\ \hat{F}_{\text{Hartree}}(1) &= \hat{h}(1) + \sum_{j(\neq i)}^N \hat{J}_j(1),\end{aligned}$$

where \hat{F}_{Hartree} corresponds to the Fock operator. Note that there is no self-interaction there.

8.2.9 ITERATIVE PROCEDURE FOR COMPUTING MOLECULAR ORBITALS: THE SELF-CONSISTENT FIELD METHOD

The following is a typical technique for solving the Fock equation.

First, we meet the difficulty that in order to solve the Fock equation we should first ... know its solution. Indeed, the Fock equation is not an eigenvalue problem, but a pseudo-eigenvalue problem, because the Fock operator depends on the *solutions* (obviously, unknown). Regardless of how strange it might seem, we deal with this situation quite easily using an iterative approach. This is called the self-consistent field method (SCF). In this method (Fig. 8.6) we

SCF iterations

- assume at the beginning (zero-th iteration) a certain shape of molecular orbitals;⁴¹
- introduce these orbitals to the Fock operator, thus obtaining a sort of “caricature” of it (the zero-order Fock operator);
- solve the eigenvalue problem using the above “Fock operator” and get the molecular orbitals of the first iteration;
- repeat the process until the shape of the orbitals does not change in the next iteration, i.e. until the Fock equations are solved.⁴²

⁴⁰Such a function is not legal – it does not fulfil the antisymmetrization postulate. This illegal character (caused by a lack of the Pauli exclusion principle) would sometimes give unfortunate consequences: e.g., more than two electrons would occupy the 1s orbital, etc.

⁴¹These are usually the any-sort “orbitals”, although recently, because of the *direct SCF* idea (we calculate the integrals whenever they are needed, i.e. at each iteration), an effort is made to save computational time per iteration and therefore to provide as good-quality a starting function as possible. We may obtain it *via* an initial calculation with some two-electron integrals neglected.

⁴²Using our driver analogy, we may say that at the beginning the driver has *false* maps of the probability density (thus the system energy is high – in our analogy the car repair costs are large). The next iterations (repair costs effectively teach all the drivers) improve the map, the final energy decreases, and at the very end we get the best map possible. The mean energy is the lowest possible (within the mean field method). A further energy lowering is only possible beyond the Hartree–Fock approximation, i.e. outside of the mean field method, which for the drivers means not using maps. A suspicious

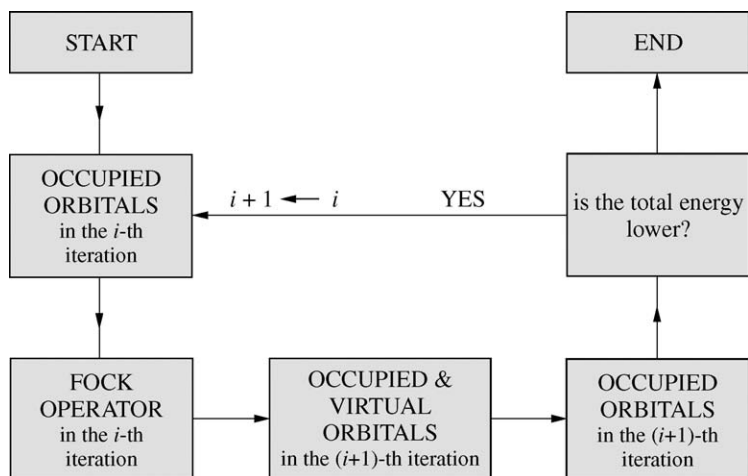


Fig. 8.6. Iterative solution of the Fock equation (the self-consistent field method, SCF). We:
 – start from any set of occupied orbitals (zeroth iteration),
 – insert them to the Fock operator,
 – solve the Fock equation,
 – obtain the molecular orbitals of the first approximation,
 – choose those of the lowest energy as the occupied ones and if your criterion of the total energy is not satisfied, repeat the procedure.

8.3 TOTAL ENERGY IN THE HARTREE–FOCK METHOD

In Appendix M, p. 986, we derived the following expressions for the mean value of the Hamiltonian using the normalized determinant (without a constant additive term for the nuclear repulsion energy V_{nn} , SMO means summation over the spinorbitals $i = 1, \dots, N$; in the RHF method, the MO summation limit means summation over the orbitals $i = 1, \dots, N/2$)

$$\begin{aligned}
 E'_{\text{HF}} &= \sum_i^{\text{SMO}} \langle i | \hat{h} | i \rangle + \frac{1}{2} \sum_{i,j=1}^{\text{SMO}} [\langle ij | ij \rangle - \langle ij | ji \rangle] \\
 &\equiv \sum_i^{\text{SMO}} h_{ii} + \frac{1}{2} \sum_{i,j=1}^{\text{SMO}} [J_{ij} - K_{ij}].
 \end{aligned} \tag{8.35}$$

person (scientist) should be careful, because our solution *may depend on the starting point used*, i.e. from the initial, completely arbitrary orbitals. Besides, the iteration process does not necessarily need to be convergent. But it appears that the solutions in the Hartree–Fock method are usually independent on the zero-th order MOs, and convergence problems are very rare. This is surprising. This situation is much worse for better-quality computations, where the AOs of small exponents are included (*diffuse orbitals*). Then we truly meet the problem already described (p. 292) of searching for the global energy minimum among a multitude of local ones.

If double occupancy is assumed (i.e. the flexibility of the variational wave function is *restricted*) we may transform this expression in the following way

$$\begin{aligned}
 E'_{\text{RHF}}(\text{double occupancy}) &= \sum_i^{\text{MO}} (\langle i\alpha | \hat{h} | i\alpha \rangle + \langle i\beta | \hat{h} | i\beta \rangle) \\
 &\quad + \frac{1}{2} \sum_i^{\text{MO}} \sum_j^{\text{SMO}} [\langle i\alpha, j | i\alpha, j \rangle - \langle i\alpha, j | j, i\alpha \rangle + \langle i\beta, j | i\beta, j \rangle - \langle i\beta, j | j, i\beta \rangle] \\
 &= 2 \sum_i^{\text{MO}} (i | \hat{h} | i) + \frac{1}{2} \sum_i^{\text{MO}} \sum_j^{\text{MO}} [\langle i\alpha, j\alpha | i\alpha, j\alpha \rangle + \langle i\alpha, j\beta | i\alpha, j\beta \rangle - \langle i\alpha, j\alpha | j\alpha, i\alpha \rangle \\
 &\quad - \langle i\alpha, j\beta | j\beta, i\alpha \rangle + \langle i\beta, j\alpha | i\beta, j\alpha \rangle + \langle i\beta, j\beta | i\beta, j\beta \rangle \\
 &\quad - \langle i\beta, j\alpha | j\alpha, i\beta \rangle - \langle i\beta, j\beta | j\beta, i\beta \rangle] \\
 &= 2 \sum_i^{\text{MO}} (i | \hat{h} | i) + \frac{1}{2} \sum_i^{\text{MO}} \sum_j^{\text{MO}} [4(ij | ij) - 2(ij | ji)] \\
 &= 2 \sum_i^{\text{MO}} (i | \hat{h} | i) + \sum_i^{\text{MO}} \sum_j^{\text{MO}} [2(ij | ij) - (ij | ji)].
 \end{aligned}$$

This finally gives

$$E'_{\text{RHF}} = 2 \sum_i^{\text{MO}} (i | \hat{h} | i) + \sum_{i,j}^{\text{MO}} [2(ij | ij) - (ij | ji)] \equiv 2 \sum_i^{\text{MO}} h_{ii} + \sum_{i,j}^{\text{MO}} [2\mathcal{J}_{ij} - \mathcal{K}_{ij}]. \quad (8.36)$$

Given the equality $\langle i | \hat{h} | i \rangle = (i | \hat{h} | i)$, these integrals have been written here as h_{ii} . The *Coulombic and exchange* integrals expressed in spinorbitals are denoted J_{ij} and K_{ij} and expressed in orbitals as \mathcal{J}_{ij} and \mathcal{K}_{ij} .

Both formulae (8.35) and (8.36) may give different results, because in the first, no double occupancy is assumed (we will discuss this further on p. 372).

The additive constant corresponding to the internuclear repulsion (it is constant, since the nuclei positions are frozen)

$$V_{nn} = \sum_{a < b} \frac{Z_a Z_b}{R_{ab}}, \quad (8.37)$$

has not yet been introduced and thus the full Hartree–Fock energy is

$$E_{\text{RHF}} = E'_{\text{RHF}} + V_{nn}. \quad (8.38)$$

Note that

$$\sum_{i,j}^{\text{MO}} [2\mathcal{J}_{ij} - \mathcal{K}_{ij}] \equiv \sum_{i,j}^{\text{MO}} [2(ij|ij) - (ij|ji)] = \left\langle \psi_{\text{HF}} \left| \sum_{i < j} \frac{1}{r_{ij}} \right| \psi_{\text{HF}} \right\rangle = V_{ee}. \quad (8.39)$$

Hence, V_{ee} is the mean electronic repulsion energy in our system.⁴³

It is desirable (interpretation purposes) to include the orbital energies in the formulae derived. Let us recollect that the orbital energy ε_i is the mean value of the Fock operator for orbital i , i.e. the energy of an effective electron described by this orbital. Based on formulae (8.31)–(8.33), this can be expressed as (i stands for the molecular orbital)

$$\varepsilon_i = h_{ii} + \sum_j^{\text{MO}} [2\mathcal{J}_{ij} - \mathcal{K}_{ij}], \quad (8.40)$$

and this in turn gives an elegant expression for the Hartree–Fock electronic energy

$$E'_{\text{RHF}} = \sum_i^{\text{MO}} [h_{ii} + \varepsilon_i]. \quad (8.41)$$

From eqs. (8.36), (8.40) and (8.39), the total electronic energy may be expressed as

$$E'_{\text{RHF}} = \sum_{i=1}^{\text{MO}} 2\varepsilon_i - V_{ee}. \quad (8.42)$$

It can be seen that the total electronic energy (i.e. E'_{RHF}) is *not* the sum of the orbital energies of electrons $\sum_i 2\varepsilon_i$.

And we would already expect full additivity, since the electrons in the Hartree–Fock method are treated as independent. Yet “independent” does not mean “non-interacting”. The reason for the non-additivity is that for *each* electron we need to calculate its effective interaction with *all* the electrons, hence we would get too much repulsion.⁴⁴ Of course, the total energy, and not the sum of the orbital energies, is the most valuable. Yet in many quantum chemical problems we interpret orbital energy lowering as energetically profitable. And it turns out that such an interpretation has an approximate justification. Works by Fraga, Politzer and Rue-

⁴³Please recall that $\langle \psi_{\text{HF}} | \hat{H} | \psi_{\text{HF}} \rangle = E_{\text{RHF}}$ and V_{ee} is, therefore, the Coulombic interaction of electrons.

⁴⁴For example, the interaction of electron 5 and electron 7 is calculated *twice*: as the interaction $\frac{1}{r_{57}}$ and $\frac{1}{r_{75}}$

denberg⁴⁵ show that at the equilibrium geometry of a molecule, the formula

$$E_{\text{RHF}} = E'_{\text{RHF}} + V_{nn} \approx \frac{3}{2} \sum_{i=1}^{\text{MO}} 2\epsilon_i, \quad (8.43)$$

works with 2%–4% precision, and even better results may be obtained by taking a factor of ... 1.55 instead of $\frac{3}{2}$.

8.4 COMPUTATIONAL TECHNIQUE: ATOMIC ORBITALS AS BUILDING BLOCKS OF THE MOLECULAR WAVE FUNCTION

atomic orbital

We have to be careful because the term “atomic orbital” is used in quantum chemistry with a double meaning. These are: (i) orbitals of the mean field for a particular atom, or (ii) functions localized in the space about a given centre. We nearly always use what are known as *exponential basis sets*.⁴⁶

exponential
basis sets

$$g(\mathbf{r}) = f(x, y, z) \exp(-\zeta r^n),$$

where $f(x, y, z)$ is a polynomial. Such an atomic orbital is localized (centred) around $(0, 0, 0)$. The larger the exponent ζ , the more effective is this centring.

For $n = 1$, we have what is called the STO – *Slater-Type Orbitals*, and for $n = 2$ the GTO – *Gaussian Type Orbitals*.

8.4.1 CENTRING OF THE ATOMIC ORBITAL

Atomic orbital $g(\mathbf{r})$ may be shifted by a vector \mathbf{A} in space [translation operation $\hat{U}(\mathbf{A})$, see Chapter 2] to result in the new function $\hat{U}(\mathbf{A})g(\mathbf{r}) = g(\hat{T}^{-1}(\mathbf{A})\mathbf{r}) = g(\hat{T}(-\mathbf{A})\mathbf{r}) = g(\mathbf{r} - \mathbf{A})$, because $\hat{T}^{-1}(\mathbf{A}) = \hat{T}(-\mathbf{A})$. Hence the orbital centred at a given point (indicated by a vector \mathbf{A}) is (Fig. 8.7):

$$g(\mathbf{r} - \mathbf{A}) = f(x - A_x, y - A_y, z - A_z) \exp[-\zeta |\mathbf{r} - \mathbf{A}|^n]. \quad (8.44)$$

Different centring of the atomic orbitals is used in practice, although if the complete set of the orbitals were at our disposal, then it might be centred in a *single* point.

⁴⁵S. Fraga, *Theor. Chim. Acta* 2 (1964) 406; P. Politzer, *J. Chem. Phys.* 64 (1976) 4239; K. Ruedenberg, *J. Chem. Phys.* 66 (1977) 375.

⁴⁶Atomic orbitals (the first meaning) are usually expressed as linear combinations of atomic orbitals in the second meaning. There may be, but does not need to be, a nucleus in the centre. If we set an atomic nucleus at this point, we emphasize the important fact that an electron will reside close to the nucleus.

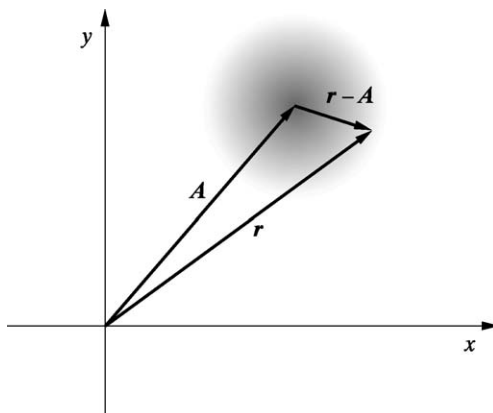


Fig. 8.7. The AO $g(\mathbf{r})$ centred at the point shown by vector A , means the creation of the orbital $g(\mathbf{r} - A)$. A linear combination of such orbitals can describe any smooth function of the position in space, of any degree of complexity.

It is more economic, however, to allow using the incomplete set and the possibility of AO centred in various points of space.

We could construct a molecular orbital of any complexity *exclusively* using the orbitals $g(\mathbf{r}) = \exp(-\zeta|\mathbf{r} - A|^n)$, i.e. the $f(x, y, z) = \text{const}$, colloquially known as the $1s$ orbitals. It is clear that we could do it in any “hole-repairing” (plastering-like) procedure.⁴⁷ But why do we not do it like this? The reason is simple: the number of such atomic orbitals that we would have to include in the calculations would be too large. Instead, chemists allow for higher-order polynomials $f(x, y, z)$. This makes for more efficient “plastering”, because, instead of spherically symmetric objects ($1s$), we can use orbitals $g(\mathbf{r})$ of virtually any shape (*via* an admixture of the p, d, f, \dots functions). For example, how a rugby-ball shaped orbital can be achieved is shown in Fig. 8.8.

8.4.2 SLATER-TYPE ORBITALS (STO)

The Slater-type orbitals⁴⁸ differ from the atomic orbitals of the hydrogen atom (see p. 178). The first difference is that the radial part is simplified in the STOs.

⁴⁷Frost even derived the method of FSGO – *Floating Spherical Gaussian Orbitals*, A.A. Frost, *J. Chem. Phys.* 47 (1967) 3707, i.e. Gaussian type orbitals of variationally-chosen positions. Their number is truly minimal – equal to the number of occupied MOs.

⁴⁸We will distinguish two similar terms here: Slater-type orbitals and Slater orbitals. The latter is reserved for special Slater-type orbitals, in which the exponent is easily computed by considering the effect of the screening of nucleus by the internal electronic shells. The screening coefficient is calculated according to the Slater rules, see below. Slater formulated these rules by watching the orbitals delivered by his young coworkers from the computing room. As he writes: “when the boys were computing”, he noticed that we can quite easily *predict* the orbital shape without any calculation. It is enough to introduce screening, which could easily be predicted from numerical experience.

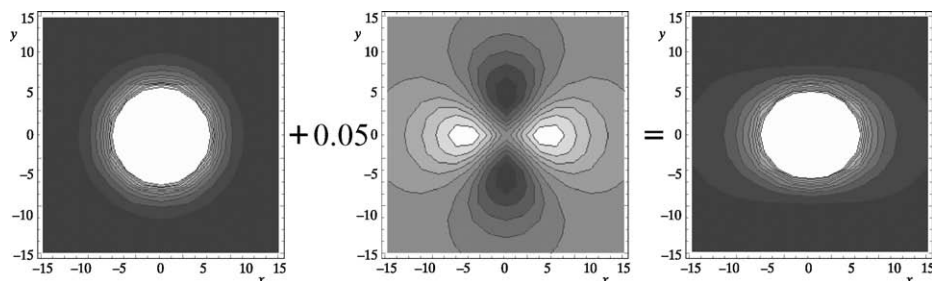


Fig. 8.8. An example of function (xy section, in a.u.) modelling by a linear combination of AOs. If a tiny admixture of the $3d_{x^2-y^2}$ function is added to the spherically symmetric $1s$ orbital (a football ball, both orbitals with 0.5 orbital exponent). We will get shrinking in one direction, and elongation in the other (the dimension in the third direction is unchanged), i.e. a flattened rugby ball. In our case the tiny admixture means 0.05. If the admixture were of the $2p$ type, the ball would look more like an egg. As we see, nearly everything can be simulated like this. This is essence of the LCAO method.

We put r^k , where k is a natural number or zero, instead of Laguerre polynomials⁴⁹ given on p. 179. The second difference is in the orbital exponent, which has no constraint except that it has to be positive.⁵⁰

The STOs have a great advantage: they decay with distance from the centre in a similar way to the “true” orbitals – let us recall the exponential vanishing of the hydrogen atom orbitals (see Chapter 4).⁵¹ STOs would be fine, but finally we have

⁴⁹This means that the radial part of a STO has no nodes. Because of this, STOs of the same angular dependence, in contrast to the hydrogen-like atom orbitals, are not orthogonal.

⁵⁰Otherwise the orbital would not be square-integrable. To get a rough idea of how the atomic orbitals for a particular atom look, Slater orbitals have been proposed: $1s, 2s, 2p, \dots$. They are Slater-type orbitals with $\zeta = \frac{Z-\sigma}{n}$, where Z stands for the nuclear charge, σ tells us how other electrons screen (i.e. effectively diminish) the charge of the nucleus ($\sigma = 0$ for an atom with a single electron), and n is the principal quantum number. The key quantity σ , is calculated for each orbital of an atom using simple rules of thumb (designed by Prof. Slater after examining his students’ computer outputs). We focus on the electron occupying the orbital in question, and we try to see what it sees. The electron sees that the nucleus charge is screened by its fellow electrons. The Slater rules are as follows:

- write down the electronic configuration of an atom grouping the orbitals in the following way: $[1s][2s2p][3s3p][3d]\dots$,
- electrons to the right give zero contribution,
- other electrons in the same group contribute 0.35 each, except $[1s]$ which contributes 0.30,
- for an electron in an $[nsnp]$ group each electron in the $n-1$ group contributes 0.85, for lower groups each contributes 1.0 and for the $[nd]$ or $[nf]$ groups, all electrons in groups to the left contribute 1.0.

Example: The carbon atom. Configuration in groups: $[1s^2][2s^22p^2]$. There will be two σ ’s: $\sigma_{1s} = 0.30$, $\sigma_{2s} = \sigma_{2p} = 3 \cdot 0.35 + 2 \cdot 0.85 = 2.75$. Hence, $\zeta_{1s} = \frac{6-0.30}{1} = 5.70$, $\zeta_{2s} = \zeta_{2p} = \frac{6-2.75}{2} = 1.625$. Hence, $1s_C = N_{1s} \exp(-5.70r)$, $2s_C = N_{2s} \exp(-1.625r)$, $2p_{x,C} = N_{2p} x \exp(-1.625r)$, $2p_{y,C} = N_{2p} y \exp(-1.625r)$, $2p_{z,C} = N_{2p} z \exp(-1.625r)$.

⁵¹It has been proved that each of the Hartree–Fock orbitals has the same asymptotic dependence on the distance from the molecule (N.C. Handy, M.T. Marron, H.J. Silverstone, *Phys. Rev.* 180 (1969) 45), i.e. $\text{const} \cdot \exp(-\sqrt{-2\varepsilon_{\text{max}}}r)$, where ε_{max} is the orbital energy of HOMO. Earlier, people thought the orbitals decay as $\exp(-\sqrt{-2\varepsilon_i}r)$, where ε_i is the orbital energy expressed in atomic units. The last formula, as is easy to prove, holds for the atomic orbitals of hydrogen atoms (see p. 178). R. Ahlrichs,

to compute a large number of the integrals needed.⁵² And here is a real problem. Since the Hamiltonian contains the electron–electron interactions, integrals appear with, in general, four atomic orbitals (of different centres). These integrals are difficult to calculate, and are therefore excessively computer time-consuming.

8.4.3 GAUSSIAN-TYPE ORBITALS (GTO)

If the exponent in eq. (8.44) is equal to $n = 2$, we are dealing with Gaussian Type Orbitals (GTO).

The most important among them are $1s$ -type orbitals:

$$\chi_p \equiv G_p(\mathbf{r}; \alpha_p, \mathbf{R}_p) = \left(\frac{2\alpha_p}{\pi} \right)^{\frac{3}{4}} \exp(-\alpha_p |\mathbf{r} - \mathbf{R}_p|^2), \quad (8.45)$$

where α_p is the orbital exponent, \mathbf{R}_p is the vector indicating the centre of the orbital, and the factor standing before the expression is the normalization constant. Why are $1s$ -type orbitals so important? Because we may construct “everything” (even s , p , d -like orbitals) out of them using proper linear combinations. For example, the difference of two $1s$ orbitals, centred at $(a, 0, 0)$ and $(-a, 0, 0)$, is similar to the $2p_x$ orbital (Fig. 8.9).

The most important reason for the great progress of quantum chemistry in recent years is replacing the Slater-type orbitals, formerly used, by Gaussian-type orbitals as the expansion functions.

Orbital size

Each orbital extends to infinity and it is impossible to measure its extension using a ruler. Still, the α_p coefficient may allow comparison of the sizes of various orbitals. And the quantity

$$\rho_p = (\alpha_p)^{-\frac{1}{2}} \quad (8.46)$$

may be called (which is certainly an exaggeration) the *orbital radius* of the orbital χ_p , because⁵³

orbital radius

$$\int_0^{\rho_p} \int_0^\pi \int_0^{2\pi} \chi_p^2 d\tau = 4\pi \int_0^{\rho_p} \chi_p^2 r^2 dr = 0.74, \quad (8.47)$$

M. Hoffmann-Ostenhoff, T. Hoffmann-Ostenhoff, J.D. Morgan III, *Phys. Rev. A* 23 (1981) 2106 have shown that at a long distance r from an atom or a molecule, the square root of the *ideal* electron density satisfies the inequality: $\sqrt{\rho} \leq C(1+r)^{\frac{(Z-N+1)}{\sqrt{2\varepsilon}}-1} \exp[-(2\varepsilon)]$, where ε is the first *ionization potential*, Z is the sum of the nuclear charges, N is the number of electrons, and C is a constant.

⁵²The number of necessary integrals may reach billions.

⁵³See, e.g., I.S. Gradshteyn, J.M. Ryzhik, “*Table of Integrals, Series, and Products*”, Academic Press, Orlando, 1980, formula 3.381.

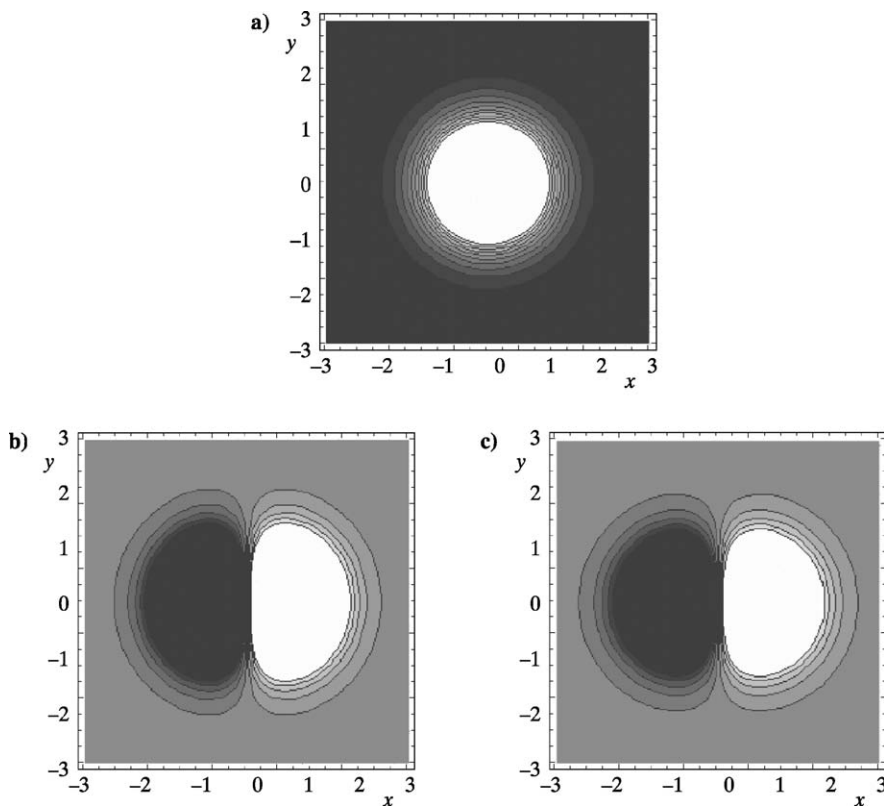


Fig. 8.9. Two spherically symmetric Gaussian-type orbitals (xy section, in a.u.) of the “1s-type” $G(\mathbf{r}; 1, \mathbf{0})$ (a) are used to form the difference orbital (b): $G(\mathbf{r}; 1, -0.5\mathbf{i}) - G(\mathbf{r}; 1, +0.5\mathbf{i})$, where \mathbf{i} is the unity vector along the x axis. For comparison (c) the Gaussian-type p_x orbital is shown: $xG(\mathbf{r}; 1, \mathbf{0})$. It can be seen that the spherical orbitals may indeed simulate the $2p$ ones. Similarly, they can model the spatial functions of arbitrary complexity.

where the integration over r goes through the inside of a sphere of radius ρ_p . This gives us an idea about the part of space in which the orbital has an important amplitude. For example, the $1s$ hydrogen atom orbital can be approximated as a linear combination of three $1s$ GTOs (here centred on the origin of the coordinate system; such a popular approximation is abbreviated to STO-3G):⁵⁴

$$1s \approx 0.64767G_1(\mathbf{r}; 0.151374, \mathbf{0}) + 0.40789G_2(\mathbf{r}; 0.681277, \mathbf{0}) + 0.07048G_3(\mathbf{r}; 4.50038, \mathbf{0}) \quad (8.48)$$

which corresponds to the following radii ρ of the three GTOs: 2.57, 1.21 and 0.47 a.u.

⁵⁴S. Huzinaga, *J. Chem. Phys.* 42 (1965) 1293.

Product of GTOs

The product of two Gaussian-type 1s orbitals (even if they have different centres) is a single Gaussian-type 1s orbital.⁵⁵

The case of GTOs other than 1s does not give any trouble, but the result is slightly different. The product of the exponential factors is, of course, the 1s-type GTO, shown above. The polynomials of x, y, z standing in both GTOs multiplied by each other (recall the dependence of the polynomial on the orbital centring,

⁵⁵ It is the most important feature of GTOs (along with the square dependence in the exponent).

Let us take two (not normalized) GTOs 1s: $\exp(-a(\mathbf{r}-\mathbf{A})^2)$ and $\exp(-b(\mathbf{r}-\mathbf{B})^2)$, the first centred on the point shown by vector \mathbf{A} , the second – by vector \mathbf{B} . It will be shown that their product is the Gaussian-type orbital

$$\exp(-a(\mathbf{r}-\mathbf{A})^2) \exp(-b(\mathbf{r}-\mathbf{B})^2) = N \exp(-c(\mathbf{r}-\mathbf{C})^2),$$

with parameters $c = a + b$, $\mathbf{C} = (a\mathbf{A} + b\mathbf{B})/(a + b)$, $N = \exp[-\frac{ab}{a+b}(\mathbf{A}-\mathbf{B})^2]$.

Vector \mathbf{C} shows the centre of the new Gaussian-type orbital. It is identical to the centre of mass position, where the role of mass is played by the orbital exponents a and b .

Here is the proof:

$$\begin{aligned} \text{Left side} &= \exp(-ar^2 + 2ar\mathbf{A} - aA^2 - br^2 + 2br\mathbf{B} - bB^2) \\ &= \exp(-(a+b)r^2 + 2r(a\mathbf{A} + b\mathbf{B})) \exp[-(aA^2 + bB^2)] \\ &= \exp(-cr^2 + 2c\mathbf{C}r) \exp[-(aA^2 + bB^2)], \\ \text{Right side} &= N \exp(-c(\mathbf{r}-\mathbf{C})^2) = N \exp[-c(r^2 - 2\mathbf{C}r + C^2)] = \text{Left side}, \end{aligned}$$

if $N = \exp(cC^2 - aA^2 - bB^2)$. It is instructive to transform the expression for N , which is a kind of amplitude of the Gaussian-type orbital originating from the multiplication of two GTOs. So,

$$\begin{aligned} N &= \exp[(a+b)C^2 - aA^2 - bB^2] = \exp\left(\frac{(a^2A^2 + b^2B^2 + 2ab\mathbf{A}\mathbf{B})}{(a+b)} - aA^2 - bB^2\right) \\ &= \exp\left(\frac{1}{a+b}(a^2A^2 + b^2B^2 + 2ab\mathbf{A}\mathbf{B} - a^2A^2 - abA^2 - b^2B^2 - abB^2)\right) \\ &= \exp\left(\frac{1}{a+b}(2ab\mathbf{A}\mathbf{B} - abA^2 - abB^2)\right) \\ &= \exp\left(\frac{ab}{a+b}(2\mathbf{A}\mathbf{B} - A^2 - B^2)\right) = \exp\left(\frac{-ab}{a+b}(\mathbf{A}-\mathbf{B})^2\right). \end{aligned}$$

This is what we wanted to show.

It is seen that if $\mathbf{A} = \mathbf{B}$, then amplitude N is equal to 1 and the GTO with the $a + b$ exponent results (as it should). The amplitude N strongly depends on the distance $|\mathbf{A} - \mathbf{B}|$ between two centres. If the distance is large, the N is very small, which gives the product of two distant GTOs as practically zero (in agreement with common sense). It is also clear that if we multiply two strongly contracted GTOs ($a, b \gg 1$) of different centres, the “GTO-product” is again small. Indeed, let us take, e.g., $a = b$. We get $N = \exp[-a/2|\mathbf{A} - \mathbf{B}|^2]$.

formula (8.44)), can always be presented as a certain polynomial of x', y', z' taken versus the new centre C . Hence, in the general case,

the product of any two Gaussian-type orbitals is a linear combination of Gaussian-type orbitals.

Integrals

If somebody wanted to perform alone⁵⁶ quantum chemical calculations, they would immediately face integrals to compute, the simplest among them being the $1s$ -type. Expressions for these integrals are given in Appendix P on p. 1004.

8.4.4 LINEAR COMBINATION OF ATOMIC ORBITALS (LCAO) METHOD

Algebraic approximation

LCAO MO

Usually we apply the self-consistent field approach with the LCAO method; this is then the SCF LCAO MO.⁵⁷ In the SCF LCAO MO method, each molecular orbital is presented as a linear combination of atomic orbitals χ_s

$$\varphi_i(1) = \sum_s^M c_{si} \chi_s(1) \quad (8.49)$$

where the symbol (1) emphasizes that each of the atomic orbitals, and the resulting molecular orbital, depend on the spatial coordinates of *one* electron only (*say*, electron 1). The coefficients c_{si} are called the LCAO coefficients.

The approximation, in which the molecular orbitals are expressed as linear combinations of the atomic orbitals, is also called the *algebraic approximation*.⁵⁸

⁵⁶That is, independent of existing commercial programs, which only require the knowledge of how to push a few buttons.

⁵⁷*Linear Combination of Atomic Orbitals – Molecular Orbitals*. This English abbreviation helped Polish quantum chemists in totalitarian times (as specialists in “MO methods”, MO standing for the mighty “citizen police” which included the secret police). It was independently used by Professors Wiktor Kemula (University of Warsaw) and Kazimierz Gumiński (Jagiellonian University). A young coworker of Prof. Gumiński complained, that despite much effort he still could not get the official registered address in Cracow, required for employment at the university. The Professor wrote a letter to the officials, and asked his coworker to deliver it in person. The reaction was immediate: “*Why didn’t you show this to us earlier?!?*”.

⁵⁸It was introduced in solid state theory by Felix Bloch (his biography is on p. 435), and used in chemistry for the first time by Hückel.

Erich Hückel (1896–1980), German physicist, professor at the universities in Stuttgart and Marburg, student of Bohr and Debye. Erich Hückel, presumably inspired by his brother Walter, an eminent organic chemist, created a simplified version of the Hartree–Fock method, which played a major role in linking the quantum theory with chemistry. Even today, although this tool is extremely simplistic and has been superseded by numerous and much better computational schemes, Hückel theory is valued as an initial insight into the electronic structure of some categories of molecules and solids.

Curiosity: these people liked to amuse themselves with little rhymes. Felix Bloch has translated a poem by Walter Hückel from German to English. It does not look like a



great poetry, but deals with the famous Erwin (Schrödinger) and his mysterious function ψ :

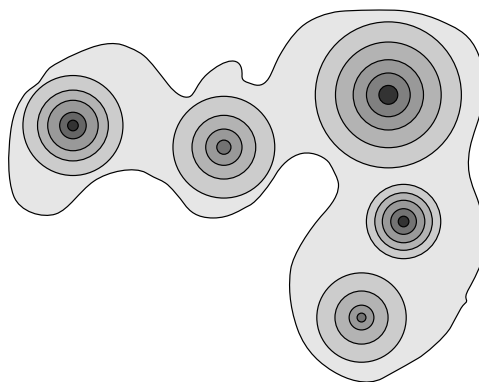
*“Erwin with his ψ can do
Calculations quite a few.
But one thing has not been seen,
Just what does ψ really mean”.*

Why is it so useful? Imagine we do not have such a tool at our disposal. Then we are confronted with defining a function that depends on the position in 3D space and has a quite complex shape (Fig. 8.10). If we want to do it accurately, we should provide the function values at many points in space, say for a large grid with a huge number of nodes, and the memory of our PC will not stand it. Also, in such an approach one would not make use of the fact that the function is smooth. We find our way through by using atomic orbitals. For example, even if we wrote that a molecular orbital is in fact a *single* atomic orbital (we can determine the latter by giving *only four* numbers: three coordinates of the orbital centre and the orbital exponent), although very primitive, this would carry a lot of physical intuition (truth...): (i) the spatial distribution of the probability of finding the electron is concentrated in some small region of space, (ii) the function decays exponentially when we go away from this region, etc.

“Blocks” of molecular orbitals φ_i are constructed out of “primary building blocks” – the one-electron functions χ_s (in the jargon called *atomic orbitals*), which

atomic orbitals
(AO)

Fig. 8.10. The concept of a molecular orbital (MO) as a linear combination of atomic orbitals (LCAO), a section view. From the point of view of mathematics, it is an expansion in a series of a complete set of functions. From the viewpoint of physics, it is just recognizing that when an electron is close to nucleus a , it should behave in a similar way as that required by the atomic orbital of atom a . From the point of view of a bricklayer, it represents the construction of a large building from soft and mutually interpenetrating bricks.



are required to fill two basic conditions:

- they need to be square-integrable,
- they need to form the complete set, i.e. “everything” can be constructed from this set (any smooth square-integrable function of x, y, z),

and several *practical* conditions:

- they should be *effective*, i.e. each single function should include a part of the physics of the problem (position in space, decay rate while going to ∞ , etc.),
- should be “flexible”, i.e. their parameters should influence their shape to a large extent,
- the resulting *integrals should be easily computable* (numerically and/or analytically), see p. 360.

In *computational practice*, unfortunately, we fulfil the second set of conditions only to some extent: the set of orbitals taken into calculations (i.e. *the basis set*) is always limited, because computing time means money, etc. In some calculations for crystals, we also remove the first set of conditions (e.g., we often use *plane waves*: $\exp(ik \cdot r)$, and these are not square-integrable).

Interpretation of LCAO. If in Fig. 8.10, we take the linear combination of five atomic orbitals and provide a reasonable choice of their centres, the exponents and the weights of the functions, we will get quite a good approximation of the ideal orbital. We account for the advantages as follows: instead of providing a huge number of function values at the grid nodes, we master the function using only $5 \times 5 = 25$ numbers.⁵⁹

The idea of LCAO MO is motivated by the fact that the molecular orbital should consist of spatial sections (atomic orbitals), because in a molecule in the vicinity of a given atom, an electron should be described by an atomic orbital of *this* atom. The essence of the LCAO approach is just the connection (unification) of such sections. But only some AOs are important in practice. This means that the main effort of constructing MOs is connected to precise shaping and polishing, by inclusion of more and more of the necessary AOs.⁶⁰

Effectiveness of AOs mixing

When could we expect that two normalized AOs will have comparable LCAO coefficients in a low-energy MO? Two rules hold (both can be deduced from eq. (D.1)) for the *mixing effectiveness* of the AOs, obtained from numerical experience:

EFFECTIVENESS OF AO MIXING

- AOs must correspond to comparable energies (in the meaning of the mean value of the Fock operator),
- AOs must have large overlap integral.

⁵⁹Three coordinates of the centre, the exponent and the coefficient c_{si} standing at AO altogether give five parameters per one AO.

⁶⁰Which plays the role of the filling mass, because we aim for a beautiful shape (i.e. ideal from the point of view of the variational method) for the MOs.

plane waves

mixing effectiveness
AO

Let us see what we obtain as the orbital energies⁶¹ (in a.u.) for several important atoms:

	1s	2s	2p	3s	3p
H	-0.5	—	—	—	—
C	-11.34	-0.71	-0.41	—	—
N	-15.67	-0.96	-0.51	—	—
O	-20.68	-1.25	-0.62	—	—
F	-26.38	-1.57	-0.73	—	—
Cl	-104.88	-10.61	-8.07	-1.07	-0.51

Now, which orbitals will mix effectively when forming methane? The hydrogen atom offers the 1s orbital with energy -0.5 . As we can see from the table, there is no possibility of effectively mixing with the carbon 1s orbital, while the 2s and 2p are very good candidates. Note that

the orbital energies of all the outer-most (the so called valence) orbitals are similar for all the elements (highlighted as bold in the table), and therefore they are able to mix effectively, i.e. to lower energy by forming chemical bonds.

This is why chemistry is mainly the science of outer shell orbitals.

The mathematical meaning of LCAO. From mathematical point of view, formula (8.49) represents an expansion of an unknown function φ_i in a series of the known functions χ_s , which belong to a certain complete set, thus M should be equal ∞ . In real life, we need to truncate this series, i.e. use some limited M .

AO basis set

8.4.5 BASIS SETS OF ATOMIC ORBITALS

BASIS SET

The set of the AOs $\{\chi_s\}$ used in the LCAO expansion is called a basis set.

The choice of the basis set functions χ (the incomplete set) is one of the most important *practical* (numerical) problems of quantum chemistry. Yet, because it is of a technical character, we will just limit ourselves to a few remarks.

Although atomic functions do not need to be *atomic* orbitals (e.g., they may be placed *in-between* nuclei), in most cases they are centred directly on the nuclei⁶² of the atoms belonging to the molecule under consideration. If M is small (in the less precise calculations), the Slater atomic orbitals discussed above are often used as the expansion functions χ_s ; for larger M (in more accurate calculations), the

⁶¹J.B. Mann, "Atomic Structure Calculations. I. Hartree-Fock Energy Results for the Elements H through Lr", Report LA-3690 (Los Alamos National Laboratory, 1967).

⁶²It is about the choice of the local coordinate system at the nucleus.

relation between χ_s and the orbitals of the isolated atoms is lost, and χ_s are chosen based on the numerical experience gathered from the literature.⁶³

8.4.6 THE HARTREE–FOCK–ROOTHAAN METHOD (SCF LCAO MO)

Clemens C.J. Roothaan (b. 1916), American physicist, professor at the University of Chicago. He became interested in this topic, after recognizing that in the literature people write about the effective one-electron operator, but he could not find its mathematical expression.



George G. Hall (b. 1925), Irish physicist, professor of Mathematics at the University of Nottingham. His scientific achievements are connected to localized orbitals, ionization potentials, perturbation theory, solvation and chemical reactions.



The Hartree–Fock (HF) equations are nonlinear differential-integral equations, which can be solved by appropriate numerical methods. For example, in the case of atoms and diatomics the orbitals may be obtained in a numerical form.⁶⁴ High accuracy at long distances from the nuclei is their great advantage. However, the method is very difficult to apply for larger systems.

A solution is the use of the LCAO MO method (algebraization of the Fock equations). It leads to simplification of the computational scheme of the Hartree–Fock method.⁶⁵ If the LCAO expansion is introduced to the expression for the total energy, then formula (8.41) (together with $\varepsilon_i = (i|\hat{F}|i)$) gives:

$$\begin{aligned}
 E'_{\text{HF}} &= \sum_i [h_{ii} + (i|\hat{F}|i)] = \sum_{i=1}^{\text{MO}} \sum_{rs} c_{ri}^* c_{si} [(r|\hat{h}|s) + (r|\hat{F}|s)] \\
 &\equiv \frac{1}{2} \sum_{rs} P_{sr} [h_{rs} + F_{rs}],
 \end{aligned} \tag{8.50}$$

bond-order
matrix

where \mathbf{P} in the RHF method is called the *bond-order matrix*,

⁶³For those who love such problems, we may recommend the article by S. Wilson “*Basis Sets*” in the book “*Ab initio Methods in Quantum Chemistry*”, ed. by K.P. Lawley, 1987, p. 439. In fact this knowledge is a little magic. Certain notations describing the quality of basis sets are in common use. For example, the symbol 6-31G* means that the basis set uses GTOs (G), the hyphen divides two electronic shells (here K and L, see p. 381). The K shell is described by a single atomic orbital, which is a certain linear combination (a “contracted orbital”) of six GTOs of the 1s type, and the two digits, 31, pertain to the L shell and denote two contracted orbitals for each valence orbital (2s, 2p_x, 2p_y, 2p_z), one of these contains three GTOs, the other one GTO (the latter is called “contracted”, with a bit of exaggeration). The starlet corresponds to d functions used additionally in the description of the L shell (called polarization functions).

⁶⁴J. Kobus, *Adv. Quantum Chem.* 28 (1997) 1.

⁶⁵The LCAO approximation was introduced to the Hartree–Fock method, independently, by C.C.J. Roothaan, *Rev. Modern Phys.* 23 (1951) 69 and G.G. Hall, *Proc. Royal Soc.* A205 (1951) 541.

$$P_{sr} = 2 \sum_i^{\text{MO}} c_{ri}^* c_{si},$$

and the summation goes over all the occupied MOs. The symbols h_{rs} and F_{rs} , introduced here, are the matrix elements of the corresponding operators. In consequence, a useful expression for the total energy in the HF method may be written as

$$E_{\text{HF}} = \frac{1}{2} \sum_{rs} P_{sr} (h_{rs} + F_{rs}) + \sum_{a < b} \frac{Z_a Z_b}{R_{ab}}, \quad (8.51)$$

where the first summation goes over the atomic orbitals (AO). For completeness, we also give the expression for F_{rs}

$$F_{rs} = (r | \hat{h} + 2\hat{J} - \hat{K} | s) = h_{rs} + \sum_i^{\text{MO}} [2(r | i | s i) - (r | i | i s)], \quad (8.52)$$

where i is the index of a MO, and r and s denote the AOs.

Expressing everything in AOs we obtain:

$$\begin{aligned} F_{rs} &= h_{rs} + \sum_i^{\text{MO}} \sum_{pq}^{\text{AO}} c_{pi}^* c_{qi} [2(rp | sq) - (rp | qs)] \\ &= h_{rs} + \sum_{pq}^{\text{AO}} P_{qp} \left[(rp | sq) - \frac{1}{2} (rp | qs) \right], \end{aligned} \quad (8.53)$$

where the summation goes over the AOs. We will use these formulae in the future.

In the SCF LCAO MO method, the Fock equations (complicated differential-integral equations) are solved in a very simple way. From (8.49) and (8.30) we have

$$\hat{F} \sum_s c_{si} \chi_s = \varepsilon_i \sum_s c_{si} \chi_s. \quad (8.54)$$

Making the scalar product with χ_r for $r = 1, 2, \dots, M$ we obtain

$$\sum_s (F_{rs} - \varepsilon_i S_{rs}) c_{si} = 0. \quad (8.55)$$

This is equivalent to the *Roothaan matrix equation*:⁶⁶

$$\mathbf{F}\mathbf{c} = \mathbf{S}\mathbf{c}\boldsymbol{\varepsilon} \quad (8.56)$$

Roothaan
matrix equation

where \mathbf{S} is the matrix of the overlap integrals $\langle \chi_r | \chi_s \rangle$ involving the AOs, $\boldsymbol{\varepsilon}$ is the diagonal matrix of the orbital energies⁶⁷ ε_i , and \mathbf{F} is the Fock operator matrix.

⁶⁶Left-hand side: $\sum_s F_{rs} c_{si}$, right-hand side: $\sum_{s,l} S_{rs} c_{sl} \varepsilon_{li} = \sum_{s,l} S_{rs} c_{sl} \delta_{li} \varepsilon_i = \sum_s S_{rs} c_{si} \varepsilon_i$. Comparison of both sides of the equation gives the desired result.

⁶⁷In fact some approximations to them. Their values approach the orbital energies, when the basis set of AOs gets closer to the complete basis set.

Each of these matrices is square (of the rank M). F depends on c (and this is why it is a *pseudo*-eigenvalue equation).

The Hartree–Fock–Roothaan matrix equation is solved iteratively:

- a) we assume an initial c matrix (i.e. also an initial P matrix; often in the zero-th iteration we put $P = 0$, as if there were no electron repulsion),
- b) we find the F matrix using matrix P ,
- c) we solve the Hartree–Fock–Roothaan equation (see Appendix L, p. 984) and obtain the M MOs, we choose the $N/2$ occupied orbitals (those of lowest energy),
- d) we obtain a new c matrix, and then a new P , etc.,
- e) we go back to a).

The iterations are terminated when the total HF energy (more liberal approach) or the coefficients c (less liberal one) change less than the assumed threshold values. Both these criteria (ideally fulfilled) may be considered as a sign that the output orbitals are already *self-consistent*. Practically, these are never the exact solutions of the Fock equations, because a *limited* number of AOs was used, while expansion to the complete set requires the use of an infinite number of AOs (the total energy in such a case would be called the *Hartree–Fock limit energy*).

Hartree–Fock
limit

After finding the MOs (hence, also the HF function) in the SCF LCAO MO approximation, we may calculate the total energy of the molecule as the mean value of its Hamiltonian. We need only the *occupied* orbitals, and not the *virtual* ones for this calculation.

The Hartree–Fock method only takes care of the total energy *and completely ignores the virtual orbitals*, which may be considered as a kind of by-product.

8.4.7 PRACTICAL PROBLEMS IN THE SCF LCAO MO METHOD

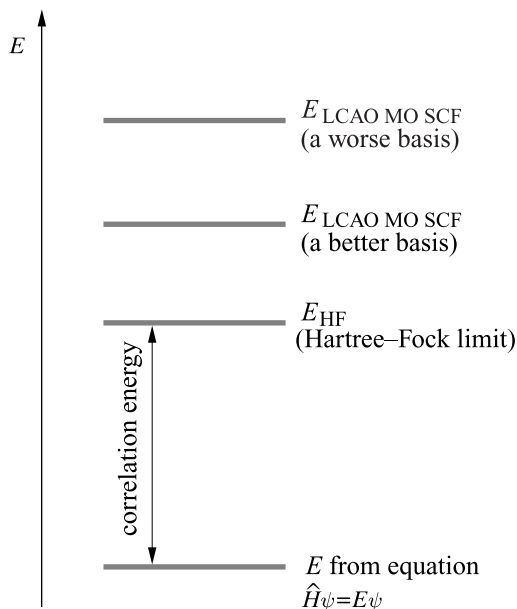
Size of the AO basis set

NUMBER OF MOs

The number of MOs obtained from the SCF procedure is always equal to the number of the AOs used. Each MO consists of various contributions of the same basis set of AOs (the apparent exception is when, *due to symmetry*, the coefficients at some AOs are equal to zero).

For double occupancy, M needs to be larger or equal to $N/2$. Typically we are forced to use large basis sets ($M \gg N/2$), and then along the occupied orbitals we get $M - N/2$ unoccupied orbitals, which are also called *virtual orbitals*. Of course, we should aim at the best *quality* MOs (i.e. how close they are to the solutions of the Fock equations), and avoiding large M (computational effort is proportional to M^4), but in practice a better basis set often means a larger M . The variational

Fig. 8.11. The Hartree–Fock method is variational. The better the wave function, the lower the mean value of the Hamiltonian. An extension of the AO basis set (i.e. adding new AOs) *has to* lower the energy, and the ideal solution of the Fock equations gives the “Hartree–Fock limit”. The ground-state eigenvalue of the Hamiltonian is thus *always* lower than the limit.



principle implies the ordering of the total energy values obtained in different approximations (Fig. 8.11).

It is required that as large a basis set as possible is used (mathematics: we approach the complete set), but we may also ask if a basis set dimension may be *decreased* freely (economy!). Of course, the answer is no! The absolute limit M is equal to half the number of the electrons, because only then can we create M spinorbitals and write the Slater determinant. However, in quantum chemistry rather misleadingly, we call the minimal basis set the basis set resulting from inner shell and valence orbitals in the corresponding atoms. For example, the minimum basis set for a water molecule is: $1s$, $2s$ and three $2p$ orbitals of oxygen and two $1s$ orbitals of hydrogen atoms, seven AOs in total (while the truly minimal basis would contain only $10/2 = 5$ AOs).

minimal basis
set

“Flip-flop”

The M MOs result from each iteration. We order them using the increasing orbital energy ε criterion, and then we use the $N/2$ orbitals of the lowest orbital energy in the Slater determinant – we call it the occupation of MOs by electrons. We might ask why we make the lowest lying MOs occupied? *The variational principle does not hold for orbital energies.* And yet we do so (not trying all possible occupations), and only very rarely we get into trouble. The most frequent trouble is that the criterion of orbital energy leads to the occupation of one set of MOs in odd iterations, and another set of MOs in even ones (typically both sets differ by including/excluding one of the two MOs that are neighbours on the energy scale) and the energy re-

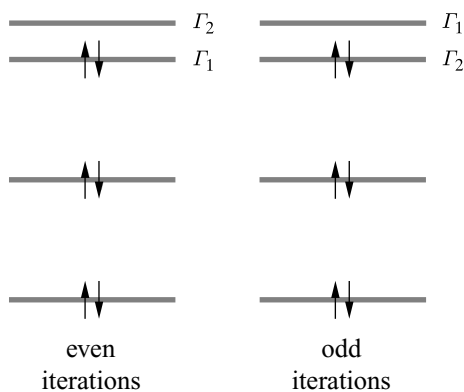


Fig. 8.12. A difficult case for the SCF method (“flip-flop”). We are sure that the orbitals exchange in subsequent iterations, because they differ in symmetry (Γ_1, Γ_2).

sulting from the odd iterations is different from that of the even ones.⁶⁸ Such behaviour of the Hartree–Fock method is indeed annoying⁶⁹ (Fig. 8.12).

Dilemmas of the AOs centring

Returning to the total energy issue, we should recall that in order to decrease the total energy, we may move the nuclei (so far frozen during the HF procedure). This is called the geometry optimization. Practically all calculations should be repeated for each nuclear geometry during such optimization.⁷⁰ And there is one more subtlety. As was said before, the AOs are most often centred on the nuclei. When the nuclei are moved, the question arises whether a nucleus should pull its AOs to a new place, or not.⁷¹ If *not*, then this “slipping off” the nuclei will significantly increase the energy (independent of, whether the geometry is improved or not). If *yes*, then in fact we use different basis sets for each geometry, hence in each case we search for the solution in a slightly different space (because it is spanned by other basis sets). People use the second approach. It is worth notifying that the problem would disappear if the basis set of AOs were complete.

The problem of AO centring is a bit shameful in quantum chemistry. Let us consider the LCAO approximation and a real molecule such as Na_2CO_3 . As mentioned above, the LCAO functions have to form a complete set. But which functions? Since they have to form a complete set, they may be chosen as the eigenfunc-

⁶⁸“Flip-flop” is the common name for this sort of behaviour.

⁶⁹There are methods for mastering this rodeo by using the matrix \mathbf{P} in the k -th iteration, not taken from the previous iteration (as usual), but as a certain linear combination of \mathbf{P} from the $k-1$ and $k-2$ iterations. When the contribution of \mathbf{P} from the $k-2$ iteration is large, in comparison with that from the $k-1$ iteration, it corresponds to a gentle attempt at quietening the nervous stallion.

⁷⁰Let us take an example of CH_4 . First, we set any starting geometry, say, a square-like planar. Now, we try to change the configuration to make it out-of-plane (the energy goes down). Taking the HCH angles as all equal (tetrahedral configuration) once more lowers the total energy computed. Putting all the CH bonds of equal length gives even lower energy. Finally, by trying different CH bond lengths we arrive at the optimum geometry (for a given AO basis set). In practice, such geometry changes are made automatically by computing the gradient of total energy. The geometry optimization is over.

⁷¹Even if the AOs were off the nuclei, we would have the same dilemma.

tions of a certain Hermitian operator (e.g., the energy operator for the harmonic oscillator or the energy operator for the hydrogen atom or the uranium atom). We decide, and we are free to choose. In addition to this freedom, we add another freedom, that of the centring. Where should the eigenfunctions (of the oscillator, hydrogen or uranium atom) of the complete set be centred, i.e. positioned in space? Since it is the complete set, *each way of centring is OK by definition*. It really looks like this if we hold to principles.

But in practical calculations, we never have the complete set at our disposal. We always need to limit it to a certain finite number of functions, and it does not represent any complete set. Depending on our computational resources, we limit the number of functions. We usually try to squeeze the best results from our time and money. How do we do it? We apply our physical intuition to the problem, believing that it will pay off. First of all, intuition suggests the use of functions for some atom which is present in the molecule, and not those of the harmonic oscillator, or the hydrogen or uranium atom, which are absent from our molecule. And here we meet another problem. *Which* atom, because we have Na, C and O in Na_2CO_3 . It appears that

the solution close to optimum is to take as a basis set *the beginnings* of several complete sets – each of them centred on one of the atoms.

So, we could centre the $1s$, $2s$, $2p$, $3s$ orbitals on both Na atoms, and the $1s$, $2s$, $2p$ set on the C and O atoms.⁷²

8.5 BACK TO FOUNDATIONS...

8.5.1 WHEN DOES THE RHF METHOD FAIL?

The reason for any Hartree–Fock method failure can be only one thing: the wave function is approximated as a *single* Slater determinant. All possible catastrophes come from this. And we might even deduce when the Hartree–Fock method is not appropriate for description of a particular real system. First, let us ask when a single determinant would be OK? Well, if out of all determinants which could be constructed from a certain spinorbital basis set, only its energy (i.e. the mean value of Hamiltonian for this determinant) were close to the true energy of the molecule. In such a case, only this determinant would matter in the linear combination of

⁷²This is nearly everything, except for a small paradox, that if we are moderately poor (reasonable but not extensive basis sets), then our results will be good, but if we became rich (and we perform high-quality computations using very large basis sets for each atom) then we would get into trouble. This would come from the fact that our basis set starts to look like six distinct complete sets. Well, that looks too good, doesn't it? We have an *overcomplete* set, and trouble must come. The overcompleteness means that any orbital from one set is already representable as a linear combination of another complete set. You would see strange things when trying to diagonalize the Fock matrix. No way! Be sure that you would be begging to be less rich.

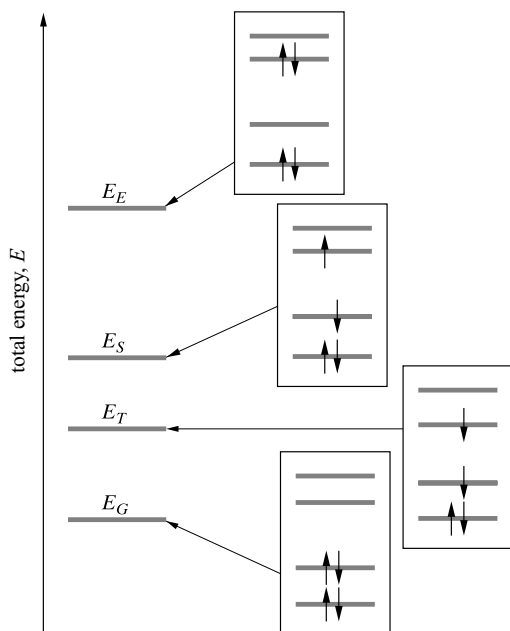


Fig. 8.13. In exact theory there is no such a thing as molecular orbitals. In such a theory we would only deal with the many-electron states and the corresponding energies of the molecule. If, nevertheless, we decided to *stick to the one-electron approximation*, we would have the MOs and the corresponding orbital energies. These one-electron energy levels can be occupied by electrons (0, 1 or 2) in various ways (the meaning of the occupation is given on p. 342), and a many-electron wave function (a Slater determinant) corresponds to each occupation. This function gives a certain mean value of the Hamiltonian, i.e. the total energy of the molecule. *In this way one value of the total energy of the molecule corresponds to a diagram of orbital occupation.* The case of the S and T states is somewhat more complex than the one shown here, and we will come back to it on p. 390.

determinants,⁷³ and the others would have negligible coefficients. It could be so,⁷⁴ if the energies of the occupied orbitals were much lower than those of the virtual ones (*“Aufbau Prinzip”*, p. 380). Indeed, various electronic states of different total energies may be approximately formed while the orbitals scheme is occupied by electrons (Fig. 8.13), and if the virtual levels are at high energies, the total energy calculated from the “excited determinant” (replacement: occupied spinorbital → virtual spinorbital) would also be high.

In other words, the danger for the RHF method is when the energy difference between HOMO and LUMO is small. For example, RHF completely fails to describe metals properly⁷⁵ Always, when the HOMO–LUMO gap is small, expect bad results.

Incorrect description of dissociation by the RHF method

An example is provided by the H_2 molecule at long internuclear distances.

In the simplest LCAO MO approach, two electrons are described by the *bonding orbital* (χ_a and χ_b are 1s orbitals centred on the H nuclei, a and b , respectively)

⁷³The Slater determinants form the complete set, p. 334. In the configuration interaction method (which will be described in Chapter 10) the electronic wave function is expanded using Slater determinants.

⁷⁴We shift here from the total energy to the one-electron energy, i.e. to the orbital picture.

⁷⁵It shows up as strange behaviour of the total energy per metal atom, which exhibits poorly-decaying oscillations with an increasing number of atoms. In addition, the exchange interactions, notorious for fast (exponential) decay as calculated by the Hartree–Fock method, are of a long-range character (see Chapter 9).

$$\varphi_{\text{bond}} = \frac{1}{\sqrt{2(1+S)}}(\chi_a + \chi_b), \quad (8.57)$$

but there is another orbital, an *antibonding* one

antibonding
orbital

$$\varphi_{\text{antibond}} = \frac{1}{\sqrt{2(1-S)}}(\chi_a - \chi_b). \quad (8.58)$$

These names stem from the respective energies. For the bonding orbital:

$$E_{\text{bond}} = \frac{H_{aa} + H_{ab}}{1+S} < H_{aa},$$

and for the antibonding orbital

$$E_{\text{antibond}} = \frac{H_{aa} - H_{ab}}{1-S} > H_{aa}.$$

These approximate formulae are obtained if we accept that the molecular orbital satisfies a sort of “Schrödinger equation” using an effective Hamiltonian (say, an analogue of the Fock operator): $\hat{H}_{\text{ef}}\varphi = E\varphi$ and after introducing notation: the overlap integral $S = (\chi_a|\chi_b)$, $H_{aa} = (\chi_a|\hat{H}_{\text{ef}}\chi_a)$, the *resonance integral*⁷⁶ $H_{ab} = H_{ba} = (\chi_a|\hat{H}_{\text{ef}}\chi_b) < 0$. The resonance integral H_{ab} , and the overlap integral S , decay exponentially when the internuclear distance R increases.

resonance
integral

INCORRECT DISSOCIATION LIMIT OF THE HYDROGEN MOLECULE

Thus we have obtained the *quasi-degeneracy* (a near degeneracy of two orbitals) for long distances, while we need to occupy *only one* of these orbitals (bonding one) in the HF method. The antibonding orbital is treated as virtual, and as such, is *completely ignored*. However, as a matter of fact, for long distances R , it corresponds to the same energy as the bonding energy.

We have to pay for such a huge drawback. And the RHF method pays, for its result significantly deviates (Fig. 8.14) from the exact energy for large R values (tending to the energy of the two isolated hydrogen atoms). This effect is known as an “incorrect dissociation of a molecule” in the RHF method (here exemplified by the hydrogen molecule). The failure may be explained in several ways and we have presented one point of view above.

incorrect
dissociation

If one bond is broken and another is formed in a molecule, the HF method does not need to fail. It appears that RHF performs quite well in such a situation, because two errors of similar magnitude (Chapter 10) cancel each other.⁷⁷

⁷⁶This integral is negative. It is its sign which decides the energy effect of the chemical bond formation (because H_{aa} is nearly equal to the energy of an electron in the H atom, i.e. $-\frac{1}{2}$ a.u.).

⁷⁷Yet the description of the transition state (see Chapter 14) is then of lower quality.

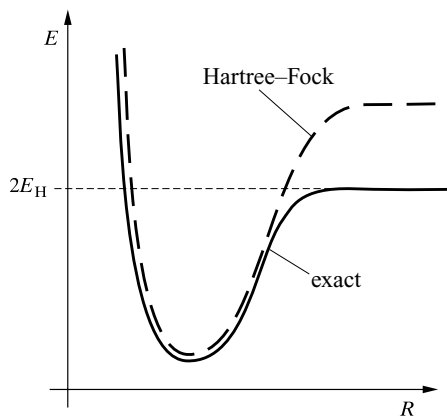


Fig. 8.14. Incorrect dissociation of H_2 in the molecular orbital (i.e. HF) method. The wave function in the form of one Slater determinant leads to dissociation products, which are neither atoms, nor ions (they should be two ground-state hydrogen atoms with energy $2E_H = -1$ a.u.).

8.5.2 FUKUTOME CLASSES

Symmetry dilemmas and the Fock operator

We have derived the general Hartree–Fock method (GHF, p. 341) providing completely free variations for the spinorbitals taken from formula (8.1). As a result, the Fock equation of the form (8.26) was derived.

We then decided to limit the spinorbital variations *via our own* condition of the double occupancy of the molecular orbitals as the real functions. This has led to the RHF method and to the Fock equation in the form (8.30).

The Hartree–Fock method is a complex (nonlinear) procedure. Do the HF solutions have any symmetry features as compared to the Hamiltonian ones? This question may be asked both for the GHF method, and also for any *spinorbital constraints* (e.g., the RHF constraints). The following problems may be addressed:

- Do the output orbitals belong to the irreducible representations of the symmetry group (Appendix C on p. 903) of the Hamiltonian? Or, if we set the nuclei in the configuration corresponding to symmetry group G , will the canonical orbitals transform according to some irreducible representations of the G group? Or, still in other words, does the Fock operator exhibits the symmetry typical of the G group?
- Does the same apply to electron density?
- Is the Hartree–Fock determinant an eigenfunction of the \hat{S}^2 operator?⁷⁸
- Is the probability density of finding a $\sigma = \frac{1}{2}$ electron equal to the probability density of finding a $\sigma = -\frac{1}{2}$ electron at any point of space?

Instabilities in the Hartree–Fock method

The above questions are connected to the *stability of the solutions*. The HF solution is stable if any change of the spinorbitals leads to a *higher* energy than the one

stability of
solutions

⁷⁸For \hat{S}_z it is always an eigenfunction.

found before. We may put certain conditions for spinorbital changes. *Relaxing the condition of double occupancy may take various forms*, e.g., the paired orbitals may be equal but *complex*, or all orbitals may be *different* real functions, or we may admit them as different complex functions, etc. Could the energy increase along with this gradual orbital constraints removal? No, an energy increase is, of course, impossible, because of the variational principle, the energy might, however, remain constant or decrease.

The general answer to this question (the character of the energy change) cannot be given since it depends on various things, such as the molecule under study, interatomic distances, the AOs basis set, etc. However, as shown by Fukutome⁷⁹ using a group theory analysis, there are exactly eight situations which *may* occur. Each of these leads to a characteristic shape of the set of occupied orbitals, which is given in Table 8.1. We may pass the borders between these eight classes of GHF method solutions while changing various parameters.⁸⁰

The Fukutome classes may be characterized according to total spin as a function of position in space:

- The first two classes RHF (TICS) and CCW correspond to identical electron spin densities for α and β electrons at any point of space (total spin density equal to zero). This implies double orbital occupancy (the orbitals are real in RHF, and complex in CCW).
- The further three classes ASCW, ASDW and ASW are characterized by the non-vanishing spin density keeping a certain direction (hence A = *axial*). The popular ASDW, i.e. UHF (no. 4) class is worth mentioning.⁸¹ We will return to the UHF function in a moment.
- The last three classes TSCW, TSDW, TSW correspond to spin density with a total non-zero spin, where direction in space varies in a complex manner.⁸²

UHF method

The Fukutome classes allow some of the posed questions to be answered:

- The resulting RHF MOs *may* belong (and most often do) to the irreducible symmetry representations (Appendix C in p. 903) of the Hamiltonian. But this is not necessarily the case.
- In the majority of calculations, the RHF electron density shows (at molecular geometry close to the equilibrium) spatial symmetry identical with the point symmetry group (the nuclear configuration) of the Hamiltonian. But the RHF method may also lead to *broken symmetry solutions*. For example, a system composed of the equidistant H atoms uniformly distributed on a circle shows bond alternation, i.e. symmetry breaking of the BOAS type.⁸³

broken
symmetry
BOAS

⁷⁹A series of papers by H. Fukutome starts with the article in *Prog. Theor. Phys.* 40 (1968) 998 and the review article *Int. J. Quantum Chem.* 20 (1981) 955. I recommend a beautiful paper by J.-L. Calais, *Adv. Quantum Chem.* 17 (1985) 225.

⁸⁰In the space of the parameters it is something like a phase diagram for a phase transition.

⁸¹UHF, i.e. *Unrestricted Hartree-Fock*.

⁸²See J.-L. Calais, *Adv. Quantum Chem.* 17 (1985) 225.

⁸³BOAS stands for the *Bond-Order Alternating Solution*. It has been shown, that the translational symmetry is broken and that the symmetry of the electron density distribution in polymers exhibits a unit

Table 8.1. Fukutome classes (for φ_{i1} and φ_{i2} see eq. (8.1))

Class	Orbital components $\begin{bmatrix} \varphi_{11} & \varphi_{21} & \dots & \varphi_{N1} \\ \varphi_{12} & \varphi_{22} & \dots & \varphi_{N2} \end{bmatrix}$	Remarks	Name
1	$\begin{bmatrix} \varphi_1 & 0 & \varphi_2 & 0 & \dots & \varphi_{N/2} & 0 \\ 0 & \varphi_1 & 0 & \varphi_2 & \dots & 0 & \varphi_{N/2} \end{bmatrix}$	φ_i real	RHF \equiv TICS ¹
2	$\begin{bmatrix} \varphi_1 & 0 & \varphi_2 & 0 & \dots & \varphi_{N/2} & 0 \\ 0 & \varphi_1 & 0 & \varphi_2 & \dots & 0 & \varphi_{N/2} \end{bmatrix}$	φ_i complex	CCW ²
3	$\begin{bmatrix} \varphi_1 & 0 & \varphi_2 & 0 & \dots & \varphi_{N/2} & 0 \\ 0 & \varphi_1^* & 0 & \varphi_2^* & \dots & 0 & \varphi_{N/2}^* \end{bmatrix}$	φ_i complex	ASCW ³
4	$\begin{bmatrix} \varphi_1 & 0 & \varphi_2 & 0 & \dots & \varphi_{N/2} & 0 \\ 0 & \chi_1 & 0 & \chi_2 & \dots & 0 & \chi_{N/2} \end{bmatrix}$	φ, χ real	UHF \equiv ASDW ⁴
5	$\begin{bmatrix} \varphi_1 & 0 & \varphi_2 & 0 & \dots & \varphi_{N/2} & 0 \\ 0 & \chi_1 & 0 & \chi_2 & \dots & 0 & \chi_{N/2} \end{bmatrix}$	φ, χ complex	ASW ⁵
6	$\begin{bmatrix} \varphi_1 & \chi_1 & \varphi_2 & \chi_2 & \dots & \varphi_{N/2} & \chi_{N/2} \\ -\chi_1^* & \varphi_1^* & -\chi_2^* & \varphi_2^* & \dots & -\chi_{N/2}^* & \varphi_{N/2}^* \end{bmatrix}$	φ, χ complex	TSCW ⁶
7	$\begin{bmatrix} \varphi_1 & \chi_1 & \varphi_2 & \chi_2 & \dots & \varphi_{N/2} & \chi_{N/2} \\ \tau_1 & \kappa_1 & \tau_2 & \kappa_2 & \dots & \tau_{N/2} & \kappa_{N/2} \end{bmatrix}$	$\varphi, \chi, \tau, \kappa$ real	TSDW ⁷
8	$\begin{bmatrix} \varphi_1 & \chi_1 & \varphi_2 & \chi_2 & \dots & \varphi_{N/2} & \chi_{N/2} \\ \tau_1 & \kappa_1 & \tau_2 & \kappa_2 & \dots & \tau_{N/2} & \kappa_{N/2} \end{bmatrix}$	$\varphi, \chi, \tau, \kappa$ complex	TSW ⁸

¹Also, according to Fukutome, TICS, i.e. *Time-reversal-Invariant Closed Shells*.

²*Charge Current Waves*.

³*Axial Spin Current Waves*.

⁴*Axial Spin Density Waves*.

⁵*Axial Spin Waves*.

⁶*Torsional Spin Current Waves*.

⁷*Torsional Spin Density Waves*.

⁸*Torsional Spin Waves*.

triplet instability

- The RHF function is always an eigenfunction of the \hat{S}^2 operator (and, of course, of the \hat{S}_z). This is no longer true, when extending beyond the RHF method (*triplet instability*).
- The probability densities of finding the $\sigma = \frac{1}{2}$ and $\sigma = -\frac{1}{2}$ electron coordinate are different for the majority of Fukutome classes (“*spin waves*”).

Example: Triplet instability

The wave function in the form of a Slater determinant is always an eigenfunction of the \hat{S}_z operator, and if in addition double occupancy is assumed (RHF) then it is also an eigenfunction of the \hat{S}^2 operator, as exemplified by the hydrogen molecule in Appendix Q on p. 1006.

cell twice as long as that of the nuclear pattern [J. Paldus, J. Čížek, *J. Polym. Sci., Part C* 29 (1970) 199, also J.-M. André, J. Delhalle, J.G. Fripiat, G. Hennico, J.-L. Calais, L. Piela, *J. Mol. Struct. (Theochem)* 179 (1988) 393]. The BOAS represents a feature related to the Jahn–Teller effect in molecules and to the Peierls effect in the solid state (see Chapter 9).

And what about the UHF method?

Let us study the two electron system, where the RHF function (the TICS Fukutome class) is:

$$\psi_{\text{RHF}} = \frac{1}{\sqrt{2}} \begin{vmatrix} \phi_1(1) & \phi_1(2) \\ \phi_2(1) & \phi_2(2) \end{vmatrix},$$

and both spinorbitals have a *common real orbital part* φ : $\phi_1 = \varphi\alpha$, $\phi_2 = \varphi\beta$.

Now we allow for a *diversification* of the orbital part (keeping the functions *real*, i.e. staying within the ASDW Fukutome class, usually called UHF in quantum chemistry) for both spinorbitals. We proceed slowly from the closed-shell situation, using as the orthonormal spinorbitals:

$$\phi'_1 = N_-(\varphi - \delta)\alpha, \quad \phi'_2 = N_+(\varphi + \delta)\beta,$$

where δ is a small real correction to the φ function, and N_+ and N_- are the normalization factors.⁸⁴ The electrons hate each other (Coulomb law) and may thank us for giving them separate apartments: $\varphi + \delta$ and $\varphi - \delta$. We will worry about the particular mathematical shape of δ in a minute. For the time being let us see what happens to the UHF function:

$$\begin{aligned} \psi_{\text{UHF}} &= \frac{1}{\sqrt{2}} \begin{vmatrix} \phi'_1(1) & \phi'_1(2) \\ \phi'_2(1) & \phi'_2(2) \end{vmatrix} \\ &= \frac{1}{\sqrt{2}} N_+ N_- \begin{vmatrix} [\varphi(1) - \delta(1)]\alpha(1) & [\varphi(2) - \delta(2)]\alpha(2) \\ \phi'_2(1) & \phi'_2(2) \end{vmatrix} \\ &= \frac{1}{\sqrt{2}} N_+ N_- \left\{ \begin{vmatrix} \varphi(1)\alpha(1) & \varphi(2)\alpha(2) \\ \phi'_2(1) & \phi'_2(2) \end{vmatrix} - \begin{vmatrix} \delta(1)\alpha(1) & \delta(2)\alpha(2) \\ \phi'_2(1) & \phi'_2(2) \end{vmatrix} \right\} \\ &= \frac{1}{\sqrt{2}} N_+ N_- \left\{ \begin{vmatrix} \varphi(1)\alpha(1) & \varphi(2)\alpha(2) \\ [\varphi(1) + \delta(1)]\beta(1) & [\varphi(2) + \delta(2)]\beta(2) \end{vmatrix} - \begin{vmatrix} \delta(1)\alpha(1) & \delta(2)\alpha(2) \\ [\varphi(1) + \delta(1)]\beta(1) & [\varphi(2) + \delta(2)]\beta(2) \end{vmatrix} \right\} \\ &= \frac{1}{\sqrt{2}} N_+ N_- \left\{ \begin{vmatrix} \varphi(1)\alpha(1) & \varphi(2)\alpha(2) \\ \varphi(1)\beta(1) & \varphi(2)\beta(2) \end{vmatrix} + \begin{vmatrix} \varphi(1)\alpha(1) & \varphi(2)\alpha(2) \\ \delta(1)\beta(1) & \delta(2)\beta(2) \end{vmatrix} - \begin{vmatrix} \delta(1)\alpha(1) & \delta(2)\alpha(2) \\ \varphi(1)\beta(1) & \varphi(2)\beta(2) \end{vmatrix} - \begin{vmatrix} \delta(1)\alpha(1) & \delta(2)\alpha(2) \\ \delta(1)\beta(1) & \delta(2)\beta(2) \end{vmatrix} \right\} \\ &= N_+ N_- \psi_{\text{RHF}} \\ &\quad + \frac{1}{\sqrt{2}} N_+ N_- \{ [\varphi(1)\delta(2) - \varphi(2)\delta(1)][\alpha(1)\beta(2) + \alpha(1)\beta(2)] \} \\ &\quad - \frac{1}{\sqrt{2}} N_+ N_- \begin{vmatrix} \delta(1)\alpha(1) & \delta(2)\alpha(2) \\ \delta(1)\beta(1) & \delta(2)\beta(2) \end{vmatrix}. \end{aligned}$$

⁸⁴Such a form is not fully equivalent to the UHF method, in which a general form of real orbitals is allowed.

The first and last functions are singlets ($S_z = 0, S = 0$), while the second function represents a *triplet state* ($S_z = 0, S = 1$), Appendix Q on p. 1006. Thus a small diversification of the orbital functions leads to some *triplet (second term) and singlet (third term) admixtures* to the original singlet function $N_+N_- \psi_{\text{RHF}}$ (called *triplet contamination*). The former is proportional to δ and the latter to δ^2 . Now the total wave function is no longer an eigenfunction of the \hat{S}^2 operator. How is this possible? If one electron has a spin coordinate of $\frac{1}{2}$ and the second one of $-\frac{1}{2}$, aren't they paired? Well, not necessarily, because one of the triplet functions (which describes the *parallel* configuration of both spins⁸⁵) is $[\alpha(1)\beta(2) + \alpha(1)\beta(2)]$.

ASDW

Is the resulting UHF energy (calculated for such a function) lower than the corresponding RHF energy (calculated for ψ_{RHF}), i.e. is the RHF solution unstable towards ASDW-type spinorbitals changes (no. 4 in the Table of Fukutome classes)?

It depends on a particular situation. A while before, we promised to consider what the δ function should look like for the hydrogen molecule. In the RHF method, both electrons occupy the same molecular orbital φ . If we assume within the UHF method that whenever one electron is close to the a nucleus, the second one prefers to be closer to b , this would happily be accepted by the electrons, since they repel each other (the mean value of the Hamiltonian would decrease, this is welcome). Taking the $\delta = \varepsilon \tilde{\varphi}$ function (where $\tilde{\varphi}$ is the antibonding orbital, and $\varepsilon > 0$ is a small coefficient) would have such consequences. Indeed, the sum $\varphi + \delta = \varphi + \varepsilon \tilde{\varphi}$ takes larger absolute value preferentially at one of the nuclei⁸⁶ (Fig. 8.15). Since both orbitals correspond to electrons with opposite spins, there will be some net spin on each of the nuclei. This nicely justifies the name of Axial Spin Density Wave (ASDW) Fukutome gave to the UHF method.

AMO method

A similar reasoning pertaining function $\varphi - \delta = \varphi - \varepsilon \tilde{\varphi}$ results in opposite preferences for the nuclei. Such a particular UHF method, which uses virtual orbitals $\tilde{\varphi}$ to change RHF orbitals, carries the friendly name of the AMO approach.⁸⁷

Now,

$$\begin{aligned}\psi_{\text{UHF}} &= N_+N_- \psi_{\text{RHF}} \\ &+ \frac{1}{\sqrt{2}} N_+N_- \varepsilon \{ [\varphi(1)\tilde{\varphi}(2) - \varphi(2)\tilde{\varphi}(1)] [\alpha(1)\beta(2) + \alpha(1)\beta(2)] \} \\ &- \frac{1}{\sqrt{2}} N_+N_- \varepsilon^2 \begin{vmatrix} \tilde{\varphi}(1)\alpha(1) & \tilde{\varphi}(2)\alpha(2) \\ \tilde{\varphi}(1)\beta(1) & \tilde{\varphi}(2)\beta(2) \end{vmatrix} \\ &= N_+N_- [\psi_{\text{RHF}} + \varepsilon\sqrt{2}\psi_{T''} - \varepsilon^2\psi_E],\end{aligned}$$

⁸⁵To call them parallel is an exaggeration, since they form an angle 70.5° (see Chapter 1, p. 28), but this is customary in physics and chemistry.

⁸⁶In our example, the approximate bonding orbital is $\varphi = \frac{1}{\sqrt{2}}(1s_a + 1s_b)$, and $\tilde{\varphi} = \frac{1}{\sqrt{2}}(1s_a - 1s_b)$, hence $\varphi + \varepsilon\tilde{\varphi} = \frac{1}{\sqrt{2}}[(1 + \varepsilon)1s_a + (1 - \varepsilon)1s_b]$, while $\varphi - \varepsilon\tilde{\varphi} = \frac{1}{\sqrt{2}}[(1 - \varepsilon)1s_a + (1 + \varepsilon)1s_b]$. Thus one of the new orbitals has a larger amplitude at nucleus a , while the other one has it at nucleus b (as we had initially planned).

⁸⁷*Alternant Molecular Orbitals*; P.-O. Löwdin is its author, *Symp. Mol. Phys.*, Nikko (Tokyo Maruzen), 1954, p. 13, also R. Pauncz, "Alternant Molecular Orbitals", Saunders, Philadelphia, 1967.

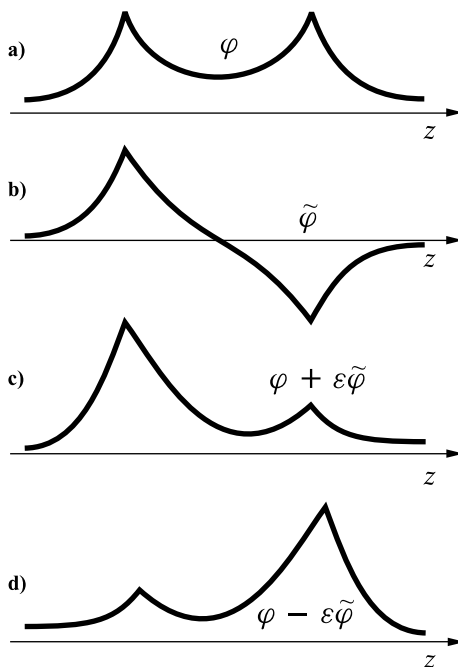


Fig. 8.15. The effect of mixing the bonding orbital φ (Fig. a) with the antibonding orbital $\tilde{\varphi}$ (Fig. b). A small admixture (c) of $\tilde{\varphi}$ to the orbital φ leads to an increase of the probability amplitude of the resulting orbital at the left nucleus, while a subtraction of $\tilde{\varphi}$ (d) leads to a larger probability amplitude of the resulting orbital at the right nucleus. Thus it results in partial separation of the spins $\frac{1}{2}$ and $-\frac{1}{2}$.

where the following notation is used for normalized functions: ψ_{RHF} for the ground state of the energy E_{RHF} , $\psi_{T''}$ for the triplet state of the energy E_T , and ψ_E for the singlet state with a doubly occupied antibonding orbital that corresponds to the energy E_E .

Let us calculate the mean value of the Hamiltonian using the ψ_{UHF} function. Because of the orthogonality of the spin functions (remember that the Hamiltonian is independent of spin) we have

$\langle \psi_{\text{RHF}} | \hat{H} \psi_{T''} \rangle = \langle \psi_{\text{RHF}} | \psi_{T''} \rangle = 0$ and obtain (with accuracy up to ε^2 terms)

$$\begin{aligned} \bar{E}_{\text{UHF}} &\approx \frac{\langle \psi_{\text{RHF}} | \hat{H} \psi_{\text{RHF}} \rangle + 2\varepsilon^2 \langle \psi_{T''} | \hat{H} \psi_{T''} \rangle - 2\varepsilon^2 \langle \psi_{\text{RHF}} | \hat{H} \psi_E \rangle}{\langle \psi_{\text{RHF}} | \psi_{\text{RHF}} \rangle + 2\varepsilon^2 \langle \psi_{T''} | \psi_{T''} \rangle} \\ &= \frac{E_{\text{RHF}} + 2\varepsilon^2 E_T - 2\varepsilon^2 (\varphi\varphi|\tilde{\varphi}\tilde{\varphi})}{1 + 2\varepsilon^2} \approx E_{\text{RHF}} + 2\varepsilon^2 [(E_T - E_{\text{RHF}}) - (\varphi\varphi|\tilde{\varphi}\tilde{\varphi})], \end{aligned}$$

where the Taylor expansion and the III Slater–Condon rule have been used (p. 986): $\langle \psi_{\text{RHF}} | \hat{H} \psi_E \rangle = (\varphi\varphi|\tilde{\varphi}\tilde{\varphi}) > 0$. The last integral is greater than zero, be-

Per-Olov Löwdin (1916–2000), Swedish chemist and physicist, student of Pauli, professor at the University of Uppsala (Sweden), founder and professor of the Quantum Theory Project at Gainesville University (Florida, USA), very active in organizing the scientific life of the international quantum chemistry community



cause it corresponds to the Coulombic self-repulsion of a certain charge distribution.

It is now clear that everything depends on the sign of the square bracket. If $E_T \gg E_{\text{RHF}}$, then the spatial diversification of the opposite spin electrons (connected with the stabilization of $-2e^2(\varphi\varphi|\tilde{\varphi}\tilde{\varphi})$) will not pay because in such a case $\bar{E}_{\text{UHF}} \gg E_{\text{RHF}}$. However, if the E_T is close to the ground state energy, then the *total energy*

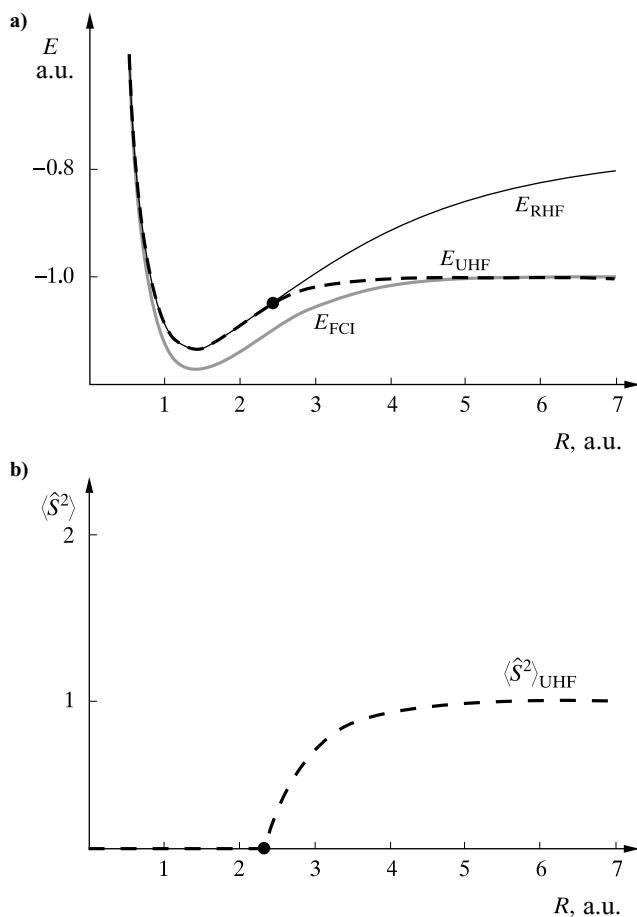


Fig. 8.16. (a) The mean value of Hamiltonian (E) calculated by the RHF and UHF methods. The lowest curve (E_{FCI}) corresponds to the accurate result (called the full configuration interaction method, see Chapter 10). (b) The mean value of the \hat{S}^2 operator calculated by the RHF and UHF methods. The energies $E_{\text{RHF}}(R)$ and $E_{\text{UHF}}(R)$ are identical for internuclear distances $R < 2.30$ a.u. For larger R values the two curves separate, and the RHF method gives an incorrect description of the dissociation limit, while the UHF method still gives a correct dissociation limit. For $R < 2.30$ a.u., the RHF and UHF wave functions are identical, and they correspond to a singlet, while for $R > 2.30$ the UHF wave function has a triplet contamination. T. Helgaker, P. Jørgensen, J. Olsen, “*Molecular Electronic Structure Theory*”, Wiley, Chichester, © 2000, reproduced with permission of John Wiley and Sons Ltd.

will decrease upon the addition of the triplet state, i.e. the RHF solutions will be unstable towards the AMO-type change of the orbitals.

This is the picture one obtains in numerical calculations for the hydrogen molecule (Fig. 8.16). At short distances between the atoms (up to 2.30 a.u.) the interaction is strong and the triplet state is of high energy. Then the variational principle does not allow the triplet state to contribute to the ground state and the UHF and the RHF give the same result. But beyond the 2.30 a.u. internuclear distance, the triplet admixture results in a small stabilization of the ground state and the UHF energy is lower than the RHF. For very long distances (when the energy difference between the singlet and triplet states is very small), the energy gain associated with the triplet component is very large.

We can see from Fig. 8.16.b the drama occurring at $R = 2.30$ a.u. for the mean value of the \hat{S}^2 operator. For $R < 2.30$ a.u. the wave function preserves the singlet character, for larger R the triplet addition increases fast, and at $R = \infty$ the mean value of the square of the total spin \hat{S}^2 is equal to 1, i.e. half-way between the $S(S + 1) = 0$ result for the singlet ($S = 0$) and the $S(S + 1) = 2$ result for the triplet ($S = 1$), since the UHF determinant is exactly 50% : 50% singlet and triplet mixture. Thus, *one determinant (UHF) is able to describe properly the dissociation of the hydrogen molecule in its ground state (singlet), but at the expense of a large spin contamination (triplet admixture).*

RESULTS OF THE HARTREE–FOCK METHOD

8.6 MENDELEEV PERIODIC TABLE OF CHEMICAL ELEMENTS

8.6.1 SIMILAR TO THE HYDROGEN ATOM – THE ORBITAL MODEL OF ATOM

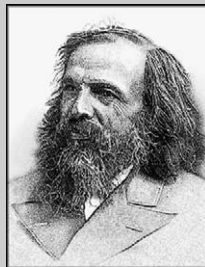
The Hartree–Fock method gives an approximate wave function for the atom of any chemical element from the Mendeleev periodic table (*orbital picture*). The Hartree–Fock method stands behind the *orbital model of atoms*. The model says essentially that a single Slater determinant can describe the atom to an accuracy that in most cases satisfies chemists. To tell the truth, the orbital model is in principle false,⁸⁸ but it is remarkable that nevertheless the conclusions drawn from it agree with experiment, at least qualitatively. It is quite exciting that

the electronic structure of all elements can be generated to a reasonable accuracy using the *Aufbau Prinzip*, i.e. a certain scheme of filling the atomic orbitals of the hydrogen atom.

⁸⁸Because the contributions of other Slater determinants (configurations) is not negligible (see Chapter 10).

Dimitrii Ivanovich Mendeleev (1834–1907), Russian chemist, professor at the University in Petersburg, and later controller of the Russian Standards Bureau of Weights and Measures (after he was expelled from the University by the tsarist powers for supporting a student protest). He was born in Tobolsk, as the youngest of fourteen children of a headmaster. In 1859 young Mendeleev – thanks to a tsarist scholarship – went to Paris and Heidelberg, where he worked with Robert Bunsen and Gustav Kirchhoff. After getting his Ph.D. in 1865, he became at 32 professor of Chemistry at the University in Sankt Petersburg. Since he had no good textbook, he started to write his own (“Principles of chemistry”). This is when he discovered one of the major human generalizations (1869): the periodicity law of chemical elements.

In 1905 he was nominated for the Nobel Prize, but lost by one vote to Henri Moissan, the discoverer of fluorine. The Swedish Royal Academy thus lost *its* chance, because in a year or so Mendeleev died. Many scientists have had similar intuition as had Mendeleev,



but it was Mendeleev who completed the project, who organized the known elements in the Table, and who predicted the existence of unknown elements. The following example shows how difficult it was for science to accept the Periodic Table. In 1864 John Newlands presented to The Royal Society in London his work showing similarities of the light elements, occurring for each eighth element with increasing atomic mass. The President of the meeting, quite amused by these considerations, suggested: “*haven’t you tried to organize them according to the alphabetic order of their names?*”.

Thus, the simple and robust orbital model serves chemistry as a “work horse”. Let us take some examples. All the atoms are build on a similar principle. A nodeless spherically symmetric atomic orbital (called $1s$) of the lowest orbital energy, next, the second lowest (and also the spherically symmetric, one radial node) is called $2s$, etc. Therefore, when filling orbital energy states by electrons some electronic shells are formed: K ($1s^2$), L ($2s^2 2p^6$), ..., where the maximum for shell orbital occupation by electrons is shown.

The very foundations of a richness around us (its basic building blocks being atoms in the Mendeleev Periodic Table) result from a very simple idea, that the proton and electron form a stable system called the hydrogen atom.

8.6.2 YET THERE ARE DIFFERENCES...

The larger the atomic number, the more complex the electronic structure. For neutral atoms the following occupation scheme applies.

Aufbau Prinzip

The Aufbau Prinzip relies on a scheme of orbital energies, Fig. 8.17. We cannot however expect that all nuances of atomic stabilities and of the ions corresponding to them might be deduced from a single simple rule like the Aufbau Prinzip, and not from the hard work of solving the Schrödinger equation (plus also the relativistic effects, Chapter 3) individually for each particular system.

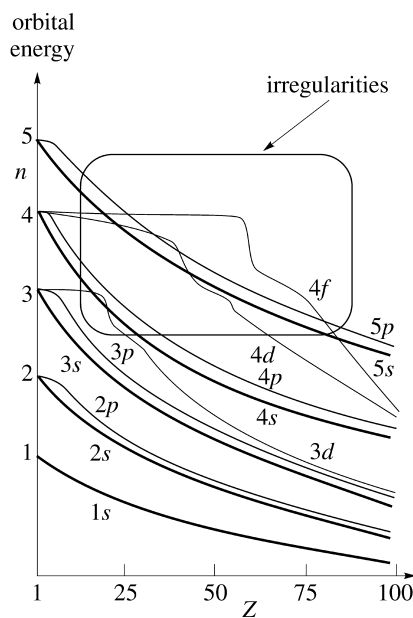


Fig. 8.17. A diagram of the order (in an energy scale) of the orbital energies as functions of the atomic number Z . This diagram, together with the Aufbau Prinzip, allows to write down the electronic configurations of atoms and explains the physical and chemical properties of chemical elements (adapted from P. Atkins, “Physical Chemistry”, sixth ed., Oxford University Press, Oxford, 1998).

From Fig. 8.17 can see that:

- the orbital energy depends not only on the principal quantum number n , but also on the angular quantum number⁸⁹ l , and the larger the l , the higher the energy,
- since for large n the *Aufbau Prinzip* is not always valid, the levels of a given n overlap in the energy scale with the $n' = n + 1$ levels.

Even so, the consecutive occupation of the electronic shells by electrons leads to a quasi-periodicity (sometimes called the periodicity) of the electronic configurations, and in consequence a quasi-periodicity of all chemical and physical properties of the elements.

Example 1. Noble gases. The atoms He, Ne, Ar, Kr, Xe, Rn have a remarkable feature, that all the subshells below and including $ns\ np$ subshell are fully occupied.

	configuration	number of electrons
He:	$1s^2$	2
Ne:	$1s^2 2s^2 2p^6$	$10 = 2 + 8$
Ar:	$1s^2 2s^2 2p^6 3s^2 3p^6$	$18 = 2 + 8 + 8$
Kr:	$1s^2 2s^2 2p^6 3s^2 3p^6 3d^{10} 4s^2 4p^6$	$36 = 2 + 8 + 8 + 18$
Xe:	$1s^2 2s^2 2p^6 3s^2 3p^6 3d^{10} 4s^2 4p^6 4d^{10} 5s^2 5p^6$	$54 = 2 + 8 + 8 + 18 + 18$
Rn:	$1s^2 2s^2 2p^6 3s^2 3p^6 3d^{10} 4s^2 4p^6 4d^{10} 5s^2 5p^6 4f^{14} 5d^{10} 6s^2 6p^6$	$86 = 2 + 8 + 8 + 18 + 18 + 32$

⁸⁹If the nucleus were large, then orbitals of different l would have different orbital energies. This explains the energy differences for the s, p, d, \dots levels, because the outer shell electrons move in the field of the nucleus shielded by the inner shell electrons (thus, in a field of something that can be seen as a large pseudo-nucleus).

According to the discussion on p. 363, what chemistry is all about is the outermost occupied orbitals that participate in forming chemical bonds. The noble gases stand out from other elements by completing their electronic shells, no wonder then that they are distinguished by very special chemical properties. The noble gases do not form chemical bonds.⁹⁰

Example 2. Alkali metals. The atoms Li, Na, K, Rb, Cs, Fr have the following dominant electronic configurations (the inner shells have been abbreviated by reporting the corresponding noble gas atom configuration):

	inner shells	valence configuration
Li	[He]	$2s^1$
Na	[Ne]	$3s^1$
K	[Ar]	$4s^1$
Rb	[Kr]	$5s^1$
Cs	[Xe]	$6s^1$
Fr	[Rn]	$7s^1$

No wonder that the elements Li, Na, K, Rb, Cs, Fr exhibit similar chemical and physical properties. Let us take any property we want, e.g., what will we get if the element is thrown into water. Lithium is a metal that reacts slowly with water, producing a colourless basic solution and hydrogen gas. Sodium is a metallic substance, and with water is a very dangerous spectacle (wild dancing flames). It reacts rapidly with water to form a colourless basic solution and hydrogen gas. The other alkali metals are even more dangerous. Potassium is a metal as well, and reacts very rapidly with water giving a colourless basic solution and hydrogen gas. Rubidium is a metal which reacts very rapidly with water producing a colourless basic solution and hydrogen gas. Cesium metal reacts rapidly with water. The result is a colourless solution and hydrogen gas. Francium is very scarce and expensive, and probably no one has tried its reaction with water. *We may however expect, with very high probability, that if the reaction were made, it would be faster than that with cesium and that a basic solution would be produced.*

However maybe all elements react rapidly with water to form a colourless basic solution and hydrogen gas? Well, this is not true. The noble gases do not. Helium does not react with water. Instead it dissolves slightly in it to the extent of about $8.61 \text{ cm}^3/\text{kg}$ at 293 K. Also neon does not react with water, but it does dissolve in it – just about $10.5 \text{ cm}^3/\text{kg}$ at 293 K. Argon, krypton, xenon and radon also do not react with water. They dissolve in it to the extent of 33.6, 59.4, 108.1 and $230 \text{ cm}^3/\text{kg}$ at 293 K, respectively. It is clear that these elements form a family that does not react with water at 293 K, but instead dissolves (slightly) in water.⁹¹ The reason is that all these elements have closed (i.e. fully occupied) shells, whereas a chemical reaction needs the opening of closed shells (see Chapter 14).

⁹⁰We have to add though, that the closed shells of the noble gases can be opened either in extreme physical conditions or by using aggressive compounds. Then, they may form chemical bonds.

⁹¹Note, that the concentration increases monotonically.

Example 3. Halogens. Let us see whether there are other families. Let us concentrate on atoms which have p^5 as the outer-most configuration. Using our scheme of orbital energies we produce the following configurations with this property: $[\text{He}]2s^22p^5$ with 9 electrons, i.e. F, $[\text{Ne}]3s^23p^5$ with 17 electrons, i.e. Cl, $[\text{Ar}]3d^{10}4s^24p^5$ with 35 which corresponds to Br, $[\text{Kr}]4d^{10}5s^25p^5$ with 53 electrons which is iodine, $[\text{Xe}]4f^{14}5d^{10}6s^26p^5$ means 85 electrons, i.e. astatine, or At. Are these elements similar? What happens to halogens in contact with water? Maybe they react very rapidly with water producing a colourless basic solution and hydrogen gas like the alkali metals, or do they just dissolve in water like the noble gases? Let us see.

Fluorine reacts with water to produce oxygen, O_2 , and ozone O_3 . This is strange in comparison with alkali metals. Next, chlorine reacts with water to produce hypochlorite, OCl^- . Bromine and iodine do a similar thing producing hypobromite OBr^- and hypoiodite OI^- . Nothing is known about the reaction of astatine with water. Apart from the exceptional behaviour of fluorine,⁹² there is no doubt we have a family of elements. This family is different from the noble gases and from the alkali metals.

Thus, the families show evidence that elements differ widely among families, but much less within a family, with rather small (and often monotonic) changes within it. This is what (quasi) periodicity is all about. The families are called *groups* (usually columns) in the Mendeleev Table.

The Mendeleev Periodic Table represents a kind of compass in chemistry. Instead of having a sort of wilderness, where all the elements exhibit their unique physical and chemical properties as *deus ex machina*, we obtain *understanding* that the animals are in a zoo, and are not unrelated, that there are some families, which follow from similar structure and occupancy of the outer electronic shells. Moreover, it became clear that there are cages in the zoo waiting for animals yet to be discovered. The animals could have been described in detail *before they were actually found by experiment*. This periodicity pertains not only to the chemical and physical properties of elements, but also to all parameters that appear in theory and are related to atoms, molecules and crystals.

8.7 THE NATURE OF THE CHEMICAL BOND

As shown on p. 371, the MO method explains the nature of the chemical bond *via* the argument that the orbital energy in the molecule is lower than that in the isolated atom. But why is this so? Which interactions decide bond formation? Do they have their origin in quantum or in classical mechanics?

To answer these questions, we will analyze the simplest case: chemical bonding in a molecular ion H_2^+ . It seems that quantum mechanics is not required here: we

⁹²For light elements the details of the electronic configuration play a more important role. For example, hydrogen may also be treated as an alkali metal, but its properties differ widely from the properties of the other members of this family.

deal with one repulsion and two attractions. No wonder there is bonding, since the net effect is one attraction. But the same applies, however, to the dissociated system (the hydrogen atom and the proton). Thus, the story is becoming more subtle.

8.7.1 H_2^+ IN THE MO PICTURE

Let us analyze chemical bonding as viewed by the poor version of the MO method (only two $1s$ hydrogen atom orbitals are used in the LCAO expansion, see Appendix R on p. 1009). Much can be seen thanks to such a poor version. The mean kinetic energy of the (only) electron of H_2^+ , residing on the bonding MO $\varphi = [2(1+S)]^{-1/2}(a+b)$, is given as (a and b denote the atomic $1s$ orbitals centred, respectively, on the a and b nuclei)

$$\bar{T} \equiv (\varphi|\hat{T}|\varphi) = \frac{T_{aa} + T_{ab}}{1+S}, \quad (8.59)$$

where S is the overlap integral $S = (a|b)$, and

$$T_{aa} = \left(a \left| -\frac{1}{2}\Delta \right| a \right) = T_{bb},$$

$$T_{ab} = \left(a \left| -\frac{1}{2}\Delta \right| b \right) = T_{ba}.$$

The non-interacting hydrogen atom and the proton have the mean kinetic energy of the electron equal to T_{aa} . The kinetic energy change is thus

$$\Delta T = \bar{T} - T_{aa} = \frac{T_{ab} - ST_{aa}}{1+S}. \quad (8.60)$$

The denominator is always positive, and the numerator (as known from computational experience) is negative for any internuclear distance. This means that the kinetic energy of the electron decreases upon molecule formation.⁹³ Hence,

kinetic energy stabilizes the molecule but not the atom.

Let us note (please recall the a and b functions are the eigenfunctions of the hydrogen atom Hamiltonian), that $T_{ab} = E_H S - V_{ab,b}$ and $T_{aa} = E_H - V_{aa,a}$, where

⁹³This agrees with intuition, which suggests that an electron now has more space for penetration ("larger box", see p. 145), and the energy levels in the box (potential energy is zero in the box, therefore we mean kinetic energy here) decrease, when the box dimension increases. This example shows that some abstract problems which can be solved exactly (here the particle in the box), serve as a beacon for more complex problems.

E_H is the ground state energy of the H atom,⁹⁴ and

$$V_{ab,b} = V_{ab,a} = -\left(a \left| \frac{1}{r_b} \right| b\right),$$

$$V_{aa,a} = -\left(a \left| \frac{1}{r_a} \right| a\right).$$

Now, ΔT can be presented as

$$\Delta T = -\frac{V_{ab,a} - S V_{aa,a}}{1 + S}, \quad (8.61)$$

because the terms with E_H cancel each other. In this way the change in kinetic energy of the electron when a molecule is formed may be *formally* presented as the integrals describing the potential energy.

Now let us calculate the change in the mean potential energy. The mean potential energy of the electron (the nucleus–nucleus interaction will be added later) equals to

$$\bar{V} = (\varphi | V | \varphi) = \left(\varphi \left| -\frac{1}{r_a} - \frac{1}{r_b} \right| \varphi \right) = \frac{(V_{aa,a} + V_{aa,b} + 2V_{ab,a})}{1 + S} \quad (8.62)$$

while in the hydrogen atom it was equal to $V_{aa,a}$. The difference, ΔV , is

$$\Delta V = \frac{(-S V_{aa,a} + 2V_{ab,a} + V_{aa,b})}{1 + S}. \quad (8.63)$$

We can see that when the change in *total* electronic energy $\Delta E_{\text{el}} = \Delta T + \Delta V$ is calculated, some kinetic energy terms will cancel the corresponding potential energy terms, and *potential energy will dominate during bond formation*:

$$\Delta E_{\text{el}} = \frac{V_{ab,a} + V_{aa,b}}{1 + S}. \quad (8.64)$$

To obtain the change, ΔE , in the total energy of the system during bond formation, we have to add the term $1/R$ describing the nuclear repulsion

$$\Delta E = \frac{V_{ab,a}}{1 + S} + \frac{V_{aa,b}}{1 + S} + \frac{1}{R}. \quad (8.65)$$

This formula is identical (because $V_{ab,a} = V_{ab,b}$) to the difference in orbital energies in the molecule H_2^+ and in the hydrogen atom, as given in Appendix R on p. 1009.

⁹⁴For example, $T_{ab} = (a | -\frac{1}{2}\Delta | b) = (a | -\frac{1}{2}\Delta - \frac{1}{r_b} + \frac{1}{r_b} | b) = E_H S + (a | \frac{1}{r_b} | b) = E_H S - V_{ab,b}$.

The formula can be easily interpreted. Let us first consider the electron density described by the φ orbital: $\varphi^2 = [2(1+S)]^{-1}(a^2 + b^2 + 2ab)$. Let us note that the density can be divided into the part close to nucleus a , that close to nucleus b , and that concentrated in the bonding region⁹⁵

$$\varphi^2 = \rho_a + \rho_b + \rho_{ab}, \quad (8.66)$$

where $\rho_a = [2(1+S)]^{-1}a^2$, $\rho_b = [2(1+S)]^{-1}b^2$, $\rho_{ab} = [(1+S)]^{-1}ab$. It can be seen,⁹⁶ that the charge associated with ρ_a is $[-2(1+S)]^{-1}$, the charge connected with the nucleus b is the same, and the overlap charge ρ_{ab} is $-S/(1+S)$. Their sum gives $-2/[2(1+S)] - 2S/[2(1+S)] = -1$ (the unit electronic charge). The formula for ΔE may also be written as (we use symmetry: the nuclei are identical, and the a and b orbitals differ only in their centres):

$$\Delta E = \frac{V_{ab,a}}{[2(1+S)]} + \frac{V_{ab,b}}{[2(1+S)]} + \frac{V_{aa,b}}{[2(1+S)]} + \frac{V_{bb,a}}{[2(1+S)]} + \frac{1}{R}. \quad (8.67)$$

Now it is clear that this formula exactly describes the Coulombic interaction (Fig. 8.18.a,b):

- of the electron cloud from the a atom (with density $\frac{1}{2}\rho_{ab}$) with the b nucleus, and vice versa (the first two terms of the expression),
- of the electron cloud of density ρ_a with the b nucleus (third term),
- of the electron cloud of density ρ_b with the a nucleus (fourth term),
- of the a and b nuclei (fifth term).

If we consider *classically* a proton approaching a hydrogen atom, the only terms for the total interaction energy are (Fig. 8.18.c):

$$\Delta E_{\text{class}} = V_{aa,b} + \frac{1}{R}. \quad (8.68)$$

The difference between ΔE and ΔE_{class} only originates from the difference in electron density, calculated quantum mechanically and classically, cf. Fig. 8.18.b,c. The ΔE_{class} is a *weak* interaction (especially for long distances), and tends to $+\infty$ for small R , because⁹⁷ of the $1/R$ term. This can be understood because ΔE_{class} is the difference between two Coulombic interactions: of a point charge with a spherical charge cloud, and of the respective two point charges (called *penetration energy*). ΔE contains two more terms in comparison with ΔE_{class} : $V_{ab,a}/[2(1+S)]$ and $V_{ab,b}/[2(1+S)]$, and both decrease exponentially to $V_{aa,a} = -1$ a.u., when R decreases to zero. Thus these terms are not important for long distances, stabilize the molecule for intermediate distances (and provide the main contribution to the chemical bond energy), and are dominated by the $1/R$ repulsion for small distances.

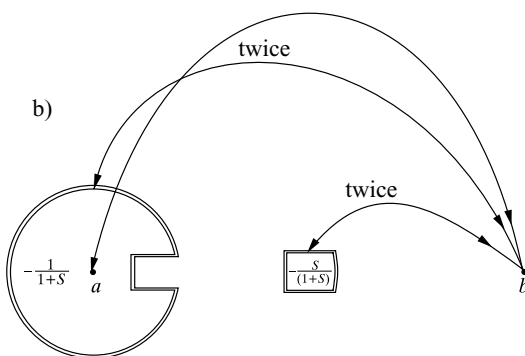
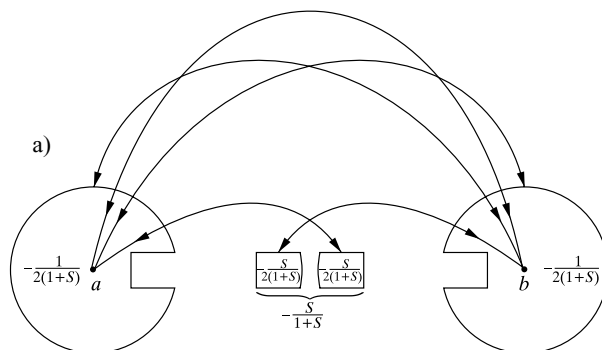
⁹⁵Function $a(1)b(1)$ has the highest value in the middle of the bond.

⁹⁶After integrating of ρ_a .

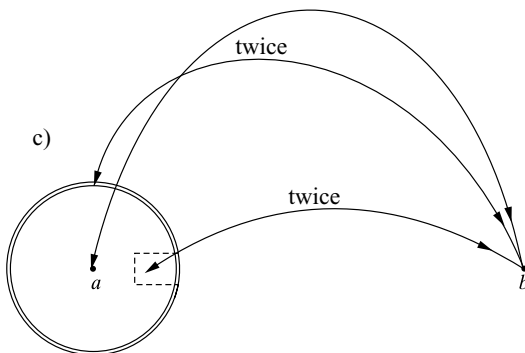
⁹⁷ $V_{aa,a}$ is finite.

Fig. 8.18. The nature of the chemical bond in the H_2^+ molecule (schematic interpretation):

— (a) *The quantum picture of the interaction.* The total electron density $\varphi^2 = \rho_a + \rho_b + \rho_{ab}$, consists of three electronic clouds $\rho_a = [2(1+S)]^{-1}a^2$ bearing the $-\frac{1}{2(1+S)}$ charge concentrated close to the a nucleus, a similar cloud $\rho_b = [2(1+S)]^{-1}b^2$ concentrated close to the b nucleus and the rest (the total charge is -1) $\rho_{ab} = [(1+S)]^{-1}ab$ bearing the charge of $-2\frac{S}{2(1+S)}$, concentrated in the middle of the bond. The losses of the charge on the a and b atoms have been shown schematically, since the charge in the middle of the bond originates from these losses. The interactions have been denoted by arrows: there are all kinds of interactions of the fragments of one atom with the fragments of the second one.



— (b) *The quantum picture – summary* (we will need it in just a moment). This scheme is similar to (a), but it has been emphasized that the attraction of ρ_a by nucleus b is the same as the attraction of ρ_b by nucleus a , hence they were both presented as one interaction of nucleus b with charge of $-2\rho_a$ at a (hence the double contour line in the figure). In this way two of the interaction arrows have disappeared as compared to (a).



— (c) *The classical picture of the interaction between the hydrogen atom and a proton.* The proton (nucleus b) interacts with the electron of the a atom, bearing the charge of $-1 = -2\frac{1}{2(1+S)} - 2\frac{S}{2(1+S)}$ and with nucleus a . Such division of the electronic charge indicates that it consists of two fragments ρ_a [as in (b)] and of two $-\frac{S}{2(1+S)}$ charges [i.e. similar to (b), but *centred in another way*]. The only difference as compared to (b) is, that in the classical picture nucleus b interacts with two *quite distant* electronic charges (put in the vicinity of nucleus a), while in the quantum picture [schemes (a) and (b)] the same charges attract themselves at short distance.

In the quantum case, for the electron charge cloud connected with the a nucleus, a^2 is decreased by a charge of $S/(1+S)$, which shifts to the halfway point towards nucleus b . In the classical case, there is no charge shift – the whole charge is close to a . In both cases there is the nucleus–nucleus and the nucleus–electron interaction. The first is identical, but the latter is completely different in both cases. Yet even in the latter interaction, there is something in common: the interaction of the nucleus with the major part of the electron cloud, with charge $-[1 - S/(1+S)] = -1/(1+S)$. The difference in the cases is the interaction with the remaining part of the electron cloud,⁹⁸ the charge $-S/(1+S)$.

In the classical view this cloud is located close to *distant* nucleus a , in the quantum view it is in the *middle of the bond*. The latter is much better for bonding. This interaction, of the (negative) electron cloud ρ_{ab} in the middle of the bond with the positive nuclei, stabilizes the chemical bond.

8.7.2 CAN WE SEE A CHEMICAL BOND?

If a substance forms crystals, it may be subjected to X-ray analysis. Such an analysis is quite exceptional, since it is one of very few techniques (which include neutronography and nuclear magnetic resonance spectroscopy), which can show atomic positions in space. More precisely, the X-ray analysis shows electronic density maps, because the radiation sees electrons, not nuclei. The inverse is true in neutronography. If we have the results of X-ray and neutron scattering, we can subtract the electron density of atoms (positions shown by neutron scattering) from the electron density of the molecular crystal (shown by X-ray scattering). This difference would be a consequence of the chemical bonding (and to a smaller extent of the intermolecular interactions). This method is called X–N or X–Ray minus Neutron Diffraction.⁹⁹ Hence differential maps of the crystal are possible, where we can see the shape of the “additional” electron density at the chemical bond, or the shape of the electron deficit (negative density) in places where the interaction is antibonding.¹⁰⁰

⁹⁸This simple interpretation gets more complex when further effects are considered, such as contributions to the energy due to the polarization of the spherically symmetric atomic orbitals or the exponent dependence of the $1s$ orbitals (i.e. the dimensions of these orbitals) on the internuclear distance. When there are several factors at play (some positive, some negative) and when the final result is of the order of a single component, then we decide which component carries responsibility for the outcome. The situation is similar to that in Parliament, when two MPs from a small party are blamed for the result of a vote (the party may be called the balancing party) while perhaps 200 others who also voted in a similar manner are left in peace.

⁹⁹There is also a pure X-ray version of this method. It uses the fact that the X-ray reflections obtained at large scattering angles see only the spherically symmetric part of the atomic electron density, similarly to that which we obtain from neutron scattering.

¹⁰⁰R. Boese, *Chemie in unserer Zeit* 23 (1989) 77; D. Cremer, E. Kraka, *Angew. Chem.* 96 (1984) 612.

From the differential maps we can estimate (by comparison with standard substances):

- 1) the strength of a chemical bond *via* the value of the positive electron density at the bond,

2) the deviation of the bond electron density (perpendicular intersection) from the cylindrical symmetry, which gives information on the π character of the chemical bond,

3) the shift of the maximum electron density towards one of the atoms which indicates the *polarization* of the bond,

4) the shift of the maximum electron density away from the straight line connecting the two nuclei, which indicates bent (banana-like) bonding.

bond
polarization

This opens up new possibilities for comparing theoretical calculations with experimental data.

8.8

EXCITATION ENERGY, IONIZATION POTENTIAL, AND ELECTRON AFFINITY (RHF APPROACH)

8.8.1

APPROXIMATE ENERGIES OF ELECTRONIC STATES

Let us consider (within the RHF scheme) the simplest closed-shell system with both electrons occupying the same orbital φ_1 . The Slater determinant, called ψ_G (G from the *ground state*) is built from two spinorbitals $\phi_1 = \varphi_1\alpha$ and $\phi_2 = \varphi_1\beta$. We also have the virtual orbital φ_2 , corresponding to orbital energy ε_2 , and we may form two other spinorbitals from it. We are now interested in the energies of *all the possible excited states* which can be formed from this pair of orbitals. These states will be represented as Slater determinants, built from φ_1 and φ_2 orbitals with the appropriate electron occupancy. We will also assume that excitations do not deform the φ orbitals (which is, of course, only partially true). Now all possible states may be listed by occupation of the ε_1 and ε_2 orbital levels, see Table 8.2.

Table 8.2. All possible occupations of levels ε_1 and ε_2

level	function	ψ_G	ψ_T	$\psi_{T'}$	ψ_1	ψ_2	ψ_E
ε_2		–	α	β	β	α	$\alpha\beta$
ε_1		$\alpha\beta$	α	β	α	β	–

E is a doubly excited electronic state, T and T' are two of three possible triplet states of the same energy. If we require that any state should be an eigenfunction of the \hat{S}^2 operator (it also needs to be an eigenfunction of \hat{S}_z , but this condition is fortunately fulfilled by all the functions listed above), it appears that only ψ_1 and

ψ_2 are illegal. However, their combinations:

$$\psi_S = \frac{1}{\sqrt{2}}(\psi_1 - \psi_2) \quad (8.69)$$

$$\psi_{T''} = \frac{1}{\sqrt{2}}(\psi_1 + \psi_2). \quad (8.70)$$

are legal. The first describes the singlet state, and the second the triplet state (the third function missing from the complete triplet set).¹⁰¹ This may be easily checked by inserting the spinorbitals into the determinants, then expanding the determinants, and separating the spin part. For ψ_S , the spin part is typical for the singlet, $\alpha(1)\beta(2) - \alpha(2)\beta(1)$, for T, T' and T'' the spin parts are, respectively, $\alpha(1)\alpha(2)$, $\beta(1)\beta(2)$ and $\alpha(1)\beta(2) + \alpha(2)\beta(1)$. This is expected for triplet functions with components of total spin equal to 1, -1, 0, respectively (Appendix Q).

Now let us calculate the mean values of the Hamiltonian using the states mentioned above. Here we will use the Slater–Condon rules (p. 986), which soon¹⁰² produce in the MO representation:

$$E_G = 2h_{11} + \mathcal{J}_{11}, \quad (8.71)$$

$$E_T = h_{11} + h_{22} + \mathcal{J}_{12} - \mathcal{K}_{12}, \quad (8.72)$$

(for all three components of the triplet)

$$E_S = h_{11} + h_{22} + \mathcal{J}_{12} + \mathcal{K}_{12}, \quad (8.73)$$

$$E_E = 2h_{22} + \mathcal{J}_{22}, \quad (8.74)$$

where $h_{ii} = (\varphi_i|\hat{h}|\varphi_i)$, and \hat{h} is a one-electron operator, the same as that appearing in the Slater–Condon rules, and *explicitly* shown on p. 335, \mathcal{J}_{ij} and \mathcal{K}_{ij} are two two-electron integrals (Coulombic and exchange): $\mathcal{J}_{ij} = (ij|ij)$ and $\mathcal{K}_{ij} = (ij|ji)$.

The orbital energies of a molecule (calculated for the state with the doubly occupied φ_1 orbital) are:

$$\varepsilon_i = (\varphi_i|\hat{F}|\varphi_i) = (\varphi_i|\hat{h} + 2\hat{\mathcal{J}} - \hat{\mathcal{K}}|\varphi_i), \quad (8.75)$$

where

$$\hat{\mathcal{J}}(1)\chi(1) = \int dV_2 \varphi_1^*(2)\varphi_1(2) \frac{1}{r_{12}}\chi(1), \quad (8.76)$$

¹⁰¹Let us make a convention, that in the Slater determinant $\frac{1}{\sqrt{2}}\det|\phi_1(1)\phi_2(2)|$, the spinorbitals are organized according to increasing orbital energy. This is important because only then are the signs in formulae (8.69) and (8.70) valid.

¹⁰²For E_G the derivation of the final formula is given on p. 352 (E'_{RHF}). The other derivations are simpler.

$$\hat{K}(1)\chi(1) = \int dV_2 \varphi_1^*(2)\chi(2)\frac{1}{r_{12}}\varphi_1(1). \quad (8.77)$$

Thus, we get:

$$\varepsilon_1 = h_{11} + \mathcal{J}_{11}, \quad (8.78)$$

$$\varepsilon_2 = h_{22} + 2\mathcal{J}_{12} - \mathcal{K}_{12}. \quad (8.79)$$

Now, the energies of the electronic states can be expressed in terms of orbital energies:

$$E_G = 2\varepsilon_1 - \mathcal{J}_{11}, \quad (8.80)$$

$$E_T = \varepsilon_1 + \varepsilon_2 - \mathcal{J}_{11} - \mathcal{J}_{12} \quad (8.81)$$

(for the ground singlet state and for the three triplet components of the common energy E_T). The distinguished role of φ_1 (in E_T) may be surprising (since the electrons reside on φ_1 and φ_2), but φ_1 is indeed distinguished, because the ε_i values are derived from the Hartree–Fock problem with the *only* occupied orbital φ_1 . So we get:

$$E_S = \varepsilon_1 + \varepsilon_2 - \mathcal{J}_{11} - \mathcal{J}_{12} + 2\mathcal{K}_{12}, \quad (8.82)$$

$$E_E = 2\varepsilon_2 + \mathcal{J}_{22} - 4\mathcal{J}_{12} + 2\mathcal{K}_{12}. \quad (8.83)$$

Now it is time for conclusions.

8.8.2 SINGLET OR TRIPLET EXCITATION?

The Jabłoński diagram plays an important role in molecular spectroscopy (Fig. 8.19). It shows three energy levels: the ground state (G), the first excited singlet state (S), and the metastable in-between state. Later on researchers identified this metastable state as the lowest triplet (T).¹⁰³

Let us compute the energy difference between the singlet and triplet states:

$$E_T - E_S = -2\mathcal{K}_{12} < 0 \quad (8.84)$$

This equation says that

a molecule always has lower energy in the excited triplet state than in the excited singlet state (both states resulting from the use of the same orbitals),

Aleksander Jabłoński (1898–1980), Polish theoretical physicist, professor at the John Casimirus University in Vilnius, then at the Nicolaus Copernicus University in Toruń, studied photoluminescence problems.



¹⁰³A. Jabłoński, *Nature* 131 (1933) 839; G.N. Lewis, M. Kasha, *J. Am. Chem. Soc.* 66 (1944) 2100.

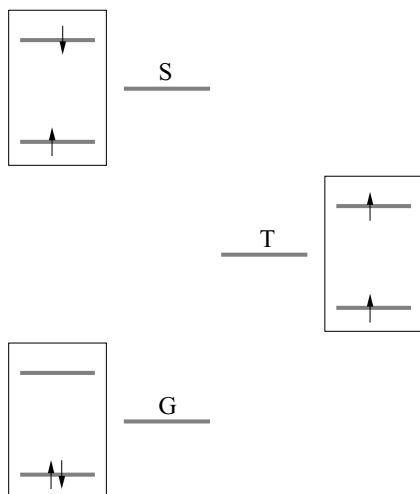


Fig. 8.19. The Jabłoński diagram. The ground state is G. The energy of the singlet excited state (S) is *higher* than the energy of the corresponding triplet state (T; that resulting from use of the same orbitals).

because $\mathcal{K}_{12} = (\varphi_1(1)\varphi_2(2)|\frac{1}{r_{12}}|\varphi_2(1)\varphi_1(2))$ is always positive being the interaction of two identical charge distributions (interpretation of an integral, real functions assumed). This rule holds firmly for the energy of the two lowest (singlet and triplet) states.

8.8.3 HUND'S RULE

The difference between the energies of the ground and triplet states is:

$$E_T - E_G = (\varepsilon_2 - \varepsilon_1) - \mathcal{J}_{12}. \quad (8.85)$$

This result has a simple interpretation. The excitation of a single electron (to the triplet state) costs some energy ($\varepsilon_2 - \varepsilon_1$), but (since $\mathcal{J}_{12} > 0$) there is also an energy gain ($-\mathcal{J}_{12}$) connected with the removal of the (mutually repulsing) electrons from the “common apartment” (orbital φ_1) to the two separate “apartments” (φ_1 and φ_2). Apartment φ_2 is admittedly on a higher floor ($\varepsilon_2 > \varepsilon_1$), but if $\varepsilon_2 - \varepsilon_1$ is small, then it may still pay to move.

In the limiting case, if $\varepsilon_2 - \varepsilon_1 = 0$, the system prefers to put electrons in separate orbitals and with the same spins (according to the empirical Hund rule, Fig. 8.20).

Friedrich Hermann Hund (1896–1997), professor of theoretical physics at the Universities in Rostock, in Leipzig (1929–1946), Jena, Frankfurt am Main, and finally Göttingen, where in his youth he had worked with Born and Franck. He applied quantum theory to atoms, ions and molecules and discovered his famous empirical rule in 1925 (biography in



German: *Intern. J. Quantum Chem.* S11 (1977) 6).

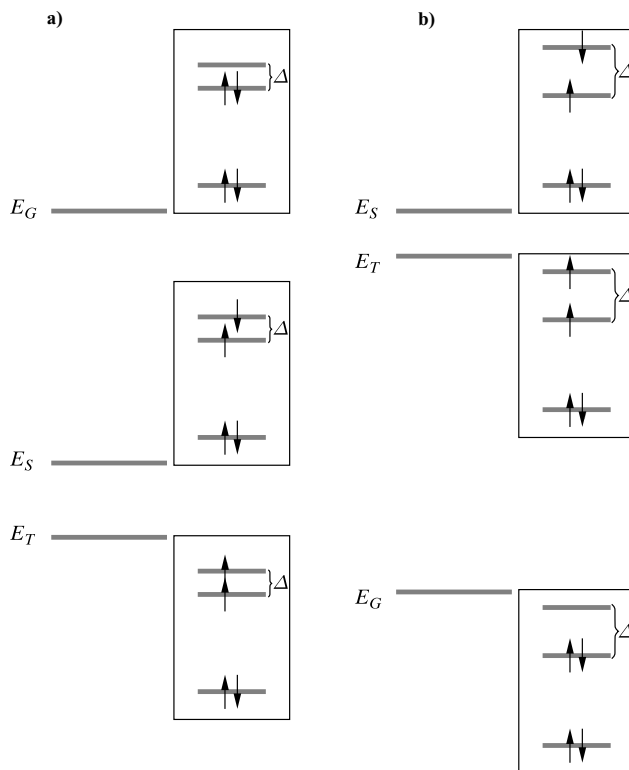


Fig. 8.20. Energy of each configuration (E_G , E_T , E_S ; left side of the pictures (a) and (b)) corresponds to an electron occupation of the orbital energy levels (shown in boxes). Two electrons of the HOMO face a dilemma:

— is it better for one of them (fortunately, they are not distinguishable...) to make a sacrifice and move to the upper-floor apartment (then they can avoid each other), Fig. (a);
 — or is it better to occupy a common apartment on the lower floor (...but electrons do not like each other), Fig. (b).

If the upper floor is not too high in the energy scale (small Δ , Fig. (a)), then each of the electrons occupies a separate apartment and they feel best having their spins parallel (triplet state). But when the upper floor energy is very high (large Δ , Fig. b), then both electrons are forced to live in the same apartment, and in that case they have to have antiparallel spins (this ensures lower energy).

The Hund's rule pertains to case (a) in its extreme form ($\Delta = 0$). When there are several orbitals of the same energy and there are many possibilities for their occupation, then the state with the lowest energy is such that the electrons each go to a separate orbital, and the alignment of their spins is "parallel" (see p. 32).

8.8.4 IONIZATION POTENTIAL AND ELECTRON AFFINITY (Koopmans RULE)

The ionization potential of the molecule M is defined as the minimum energy needed for an electron to detach from the molecule. The electron affinity energy of the molecule M is defined as the minimum energy for an electron detachment from

M^- . Let us assume again naively, that *during these operations the molecular orbitals and the orbital energies do not undergo any changes*. In fact, of course, everything changes, and the computations should be repeated for each system separately (the same applies in the previous section for excitations).

In our two-electron system, which is a model of any closed-shell molecule,¹⁰⁴ the electron removal leaves the molecule with one electron only, and its energy has to be

$$E_+ = h_{11}. \quad (8.86)$$

However,

$$h_{11} = \varepsilon_1 - \mathcal{J}_{11}. \quad (8.87)$$

This formula looks like trouble! After the ionization there is only a single electron in the molecule, while here some electron–electron repulsion (integral \mathcal{J}) appears! But everything is fine, because we still use the two-electron problem as a reference, and ε_1 relates to the two-electron problem, in which $\varepsilon_1 = h_{11} + \mathcal{J}_{11}$. Hence,

IONIZATION ENERGY

The ionization energy is equal to the negative of the orbital energy of an electron:

$$E_+ - E_G = -\varepsilon_1. \quad (8.88)$$

To calculate the electron affinity energy we need to consider a determinant as large as 3×3 , but this proves easy if the useful Slater–Condon rules (Appendix M) are applied. Rule number I gives (we write everything using the spinorbitals, then note that the three spinorbitals are derived from two orbitals, and then sum over the spin variables):

$$E_- = 2h_{11} + h_{22} + \mathcal{J}_{11} + 2\mathcal{J}_{12} - \mathcal{K}_{12}, \quad (8.89)$$

and introducing the orbital energies we get

$$E_- = 2\varepsilon_1 + \varepsilon_2 - \mathcal{J}_{11}, \quad (8.90)$$

which gives

$$E_- - E_G = \varepsilon_2 \quad (8.91)$$

Hence,

ELECTRON AFFINITY

The electron affinity is the difference of the energies of the system without an electron and one representing an anion, $E_G - E_- = -\varepsilon_2$. It is equal approximately to the negative energy of the virtual orbital on which the electron lands.

¹⁰⁴Koopmans theorem applies for this case.

A comment on Koopmans theorem

The MO approximation is, of course, a rough approximation to reality. So is Koopmans theorem, which proves to be poorly satisfied for most molecules. But these approximations are often used for practical purposes. This is illustrated by a certain quantitative relationship, derived by Grochala et al.¹⁰⁵

The authors noted, that a very simple relationship holds surprisingly well for the equilibrium *bond lengths* R of four objects: the ground state M_0 of the closed shell molecule, its excited triplet state M_T , its radical-cation $M^{+\cdot}$, and radical-anion $M^{-\cdot}$:

$$R(M_T) = R(M^{-\cdot}) + R(M^{+\cdot}) - R(M_0).$$

The above relationship is similar to that pertaining to the corresponding energies

$$E(M_T) = E(M^{-\cdot}) + E(M^{+\cdot}) - E(M_0),$$

which may be deduced, basing on certain approximations, from Koopmans theorem,¹⁰⁶ or from the Schrödinger equation while neglecting the two-electron operators (i.e. Coulomb and exchange). The difference between these two expressions is, however, fundamental: the latter holds for the four species at *the same* nuclear geometry, while the former describes the geometry *changes* for the “relaxed” species.¹⁰⁷ The first equation proved to be satisfied for a variety of molecules: ethylene, cyclobutadiene, divinylbenzene, diphenylacetylene, *trans*-N₂H₂, CO, CN⁻, N₂, and NO⁺. It also inspired Andreas Albrecht to derive general inequalities, holding for any one-electron property. The first equation, inspired by Koopmans theorem, was analyzed in detail within density functional theory¹⁰⁸ (described in Chapter 11). It is not yet clear, if it would hold beyond the one-electron approximation, or for experimental bond lengths (these are usually missing, especially for polyatomic molecules).

Tjalling Charles Koopmans (1910–1985), American econometrist of Dutch origin, professor at Yale University (USA), introduced mathematical procedures of linear programming to economics, and received the Nobel Prize for in 1975 “for work on the theory of optimum allocation of resources”.



¹⁰⁵W. Grochala, A.C. Albrecht, R. Hoffmann, *J. Phys. Chem. A* 104 (2000) 2195.

¹⁰⁶Let us check it using the formulae derived by us: $E(M_T) = \varepsilon_1 + \varepsilon_2 - \mathcal{J}_{11} - \mathcal{J}_{12}$, and $E(M^{-\cdot}) + E(M^{+\cdot}) - E(M_0) = [2\varepsilon_1 + \varepsilon_2 - \mathcal{J}_{11}] + [\varepsilon_1 - \mathcal{J}_{11}] - [2\varepsilon_1 - \mathcal{J}_{11}] = \varepsilon_2 + \varepsilon_1 - \mathcal{J}_{11}$. The equality is obtained after neglecting \mathcal{J}_{12} as compared to \mathcal{J}_{11} .

¹⁰⁷If we assume that a geometry change in these states induces an energy increase proportional to the square of the change, and that the curvature of all these parabolas is identical, then the above relationship would be easily proved. The problem is that these states have significantly different force constants, and the curvature of parabolas strongly varies among them.

¹⁰⁸P.W. Ayers, R.G. Parr, *J. Phys. Chem. A* 104 (2000) 2211.

8.9 LOCALIZATION OF MOLECULAR ORBITALS WITHIN THE RHF METHOD

The canonical MOs derived from the RHF method are usually delocalized over the whole molecule, i.e. their amplitudes are significant for all atoms in the molecule. This applies, however, mainly to high energy MOs, which exhibit a similar AO amplitude for most atoms. Yet the canonical MOs of the inner shells are usually very well localized. The canonical MOs are occupied, as usual, by putting two electrons on each low lying orbital (the Pauli exclusion principle).

The picture obtained is in contrast to chemical intuition, which indicates that the electron pairs are localized within the chemical bonds, free electron pairs and inner atomic shells. The picture which agrees with intuition may be obtained after the localization of the MOs.

The localization is based on making new orbitals as linear combinations of the canonical MOs, a fully legal procedure (see p. 338). Then, the determinantal wave function, as shown on p. 338, expressed in the new spinorbitals, takes the form $\psi' = (\det A)\psi$. For obvious reasons, the total energy will not change in this case. If linear transformation applied is an orthogonal transformation, i.e. $A^T A = \mathbf{1}$, or a unitary one, $A^\dagger A = \mathbf{1}$, then the new MOs preserve orthonormality (like the canonical ones) as shown on p. 339. We emphasize that we can make *any non-singular*¹⁰⁹ linear transformation A , not only orthogonal or unitary ones. This means something important, namely

the solution in the Hartree–Fock method depends on the *space* spanned by the occupied orbitals (i.e. on the set of all linear combinations which can be formed from the occupied MOs), and not on the orbitals only.

The new orbitals do not satisfy the Fock equations (8.30), these are satisfied by canonical orbitals only.

The localized orbitals (being some other orthonormal basis set in the space spanned by the canonical orbitals) satisfy the Fock equation (8.18) with the off-diagonal Lagrange multipliers.

Can a chemical bond be defined in a polyatomic molecule?

Unfortunately, the view to which chemists get used, i.e. the chemical bonds between pairs of atoms, lone electron pairs, inner shells, can be derived in an infinite

¹⁰⁹For any singular matrix $\det A = 0$, and this should not be allowed (p. 339).

number of ways (because of the arbitrariness of transformation A), and in each case the effects of localization ... vary. Hence,

we cannot uniquely define the chemical bond in a polyatomic molecule.

It is not a tragedy, however, because what really matters is the probability density, i.e. the square of the complex modulus of the *total* many-electron wave function. The concept of the (localized or delocalized) molecular orbitals represents simply an attempt to divide this total density into various spatially separated although overlapping parts, each belonging to a single MO. It is similar to dividing an apple into N parts. The freedom of such a division is unlimited. For example, we could envisage that each part would have the dimension of the apple (“delocalized orbitals”), or an apple would be simply cut axially, horizontally, concentrically etc. into N equal parts, forming an analogue of the localized orbitals. Yet each time the full apple could be reconstructed from these parts.

As we will soon convince ourselves, the problem of defining a chemical bond in a polyatomic molecule is not so hopeless, because various methods lead to essentially the same results.

Now let us consider some practical methods of localization. There are two categories of these: internal and external.¹¹⁰ In the external localization methods we plan where the future MOs will be localized, and the localization procedure only slightly alters our plans. This is in contrast with the internal methods where certain general conditions are imposed that induce automatically localization of the orbitals.

8.9.1 THE EXTERNAL LOCALIZATION METHODS

Projection method

This is an amazing method,¹¹¹ in which we first construct some *arbitrary*¹¹² (but linearly independent¹¹³) orbitals χ_i of the bonds, lone pairs, and the inner shells, the total number of these being equal to the number of the occupied MOs. Now let us project them on the space of the occupied HF molecular orbitals $\{\varphi_j\}$ using the projection operator \hat{P} :

$$\hat{P}\chi_i \equiv \left(\sum_j^{\text{MO}} |\varphi_j\rangle \langle \varphi_j| \right) \chi_i. \quad (8.92)$$

¹¹⁰Like medicines.

¹¹¹A. Meunier, B. Levy, G. Berthier, *Theoret. Chim. Acta* 29 (1973) 49.

¹¹²This is the beauty of the projection method.

¹¹³A linear dependence cannot be allowed. If this happens then we need to change the set of functions χ_i .

Table 8.3. Influence of the initial approximation on the final localized molecular orbitals in the projection method (the LCAO coefficients for the CH₃F molecule)

Function χ for the CF bond				The localized orbital of the CF bond				
2s(C)	2p(C)	2s(F)	2p(F)	2s(C)	2p(C)	2s(F)	2p(F)	1s(H)
0.300	0.536	0.000	−0.615	0.410	0.496	−0.123	−0.654	−0.079
0.285	0.510	0.000	−0.643	0.410	0.496	−0.131	−0.655	−0.079
0.272	0.487	0.000	−0.669	0.410	0.496	−0.138	−0.656	−0.079
0.260	0.464	0.000	−0.692	0.410	0.496	−0.144	−0.656	−0.079
0.237	0.425	0.000	−0.730	0.410	0.496	−0.156	−0.658	−0.079

The projection operator is used to create the new orbitals

$$\varphi'_i = \sum_j^{\text{MO}} \langle \varphi_j | \chi_i \rangle \varphi_j. \tag{8.93}$$

The new orbitals φ'_i , as linearly independent combinations of the occupied canonical orbitals φ_j , span the space of the canonical occupied HF orbitals $\{\varphi_j\}$. They are in general non-orthogonal, but we may apply the Löwdin orthogonalization procedure (symmetric orthogonalization, see Appendix J, p. 977).

Do the final localized orbitals depend on the starting χ_i in the projection method? The answer¹¹⁴ is in Table 8.3. The influence is small.

8.9.2 THE INTERNAL LOCALIZATION METHODS

Ruedenberg method: the maximum interaction energy of the electrons occupying a MO

The basic concept of this method was given by Lennard-Jones and Pople,¹¹⁵ and applied by Edmiston and Ruedenberg.¹¹⁶ It may be easily shown that for a given geometry of the molecule the functional $\sum_{i,j=1}^{\text{MO}} \mathcal{J}_{ij}$ is invariant with respect to *any* unitary transformation of the orbitals:

$$\sum_{i,j=1}^{\text{MO}} \mathcal{J}_{ij} = \text{const.} \tag{8.94}$$

The proof is very simple and similar to the one on p. 340, where we derived the invariance of the Coulombic and exchange operators in the Hartree–Fock method. Similarly, we can prove another invariance

$$\sum_{i,j=1}^{\text{MO}} \mathcal{K}_{ij} = \text{const}'. \tag{8.95}$$

¹¹⁴B. Lévy, P. Millié, J. Ridard, J. Vinh, *J. Electr. Spectr.* 4 (1974) 13.
¹¹⁵J.E. Lennard-Jones, J.A. Pople, *Proc. Roy. Soc. (London)* A202 (1950) 166.
¹¹⁶C. Edmiston, K. Ruedenberg, *Rev. Modern Phys.* 34 (1962) 457.

This further implies that

$$\text{maximization of } \sum_{i=1}^{\text{MO}} \mathcal{J}_{ii},$$

which is the very essence of the localization criterion, is equivalent to the *minimization* of the off-diagonal elements

$$\sum_{i < j}^{\text{MO}} \mathcal{J}_{ij}. \quad (8.96)$$

This means that to localize the molecular orbitals we try to make them as small as possible, because then the Coulombic repulsion \mathcal{J}_{ii} will be large.

It may be also expressed in another way, given that

$$\sum_{i,j}^{\text{MO}} \mathcal{K}_{ij} = \text{const}' = \sum_i^{\text{MO}} \mathcal{K}_{ii} + 2 \sum_{i < j}^{\text{MO}} \mathcal{K}_{ij} = \sum_i^{\text{MO}} \mathcal{J}_{ii} + 2 \sum_{i < j}^{\text{MO}} \mathcal{K}_{ij}.$$

Since we maximize the $\sum_i^{\text{MO}} \mathcal{J}_{ii}$, then simultaneously

we *minimize* the sum of the exchange contributions

$$\sum_{i < j}^{\text{MO}} \mathcal{K}_{ij}. \quad (8.97)$$

Boys method: the minimum distance between electrons occupying a MO

In this method¹¹⁷ we minimize the functional¹¹⁸

$$\sum_i^{\text{MO}} (ii|r_{12}^2|ii), \quad (8.98)$$

where the symbol $(ii|r_{12}^2|ii)$ denotes an integral similar to $\mathcal{J}_{ii} = (ii|ii)$, but instead of the $1/r_{12}$ operator, we have r_{12}^2 . Functional (8.98) is invariant with respect to any unitary transformation of the molecular orbitals.¹¹⁹ Since the integral $(ii|r_{12}^2|ii)$ represents the definition of the mean square of the distance between two electrons described by $\varphi_i(1)\varphi_i(2)$, the Boys criterion means that we try to obtain the localized orbitals as small as possible (small orbital *dimensions*), i.e. localized in

¹¹⁷S.F. Boys, in “*Quantum Theory of Atoms, Molecules and the Solid State*”, P.O. Löwdin, ed., Academic Press, New York, 1966, p. 253.

¹¹⁸Minimization of the interelectronic distance is in fact similar in concept to the maximization of the Coulombic interaction of two electrons in the same orbital.

¹¹⁹We need to represent the orbitals as components of a vector, the double sum as two scalar products of such vectors, then transform the orbitals, and show that the matrix transformation in the integrand results in a unit matrix.

some small volume in space. The method is similar to the Ruedenberg criterion of the maximum interelectron repulsion. The detailed technique of localization will be given in a moment. The integrals (8.98) are trivial. Indeed, using Pythagoras' theorem, we get the sum of three simple one-electron integrals of the type:

$$\begin{aligned} & (i(1)i(2)|(x_2-x_1)^2|i(1)i(2)) \\ &= (i(2)|x_2^2|i(2)) + (i(1)|x_1^2|i(1)) - 2(i(1)|x_1|i(1))(i(2)|x_2|i(2)) \\ &= 2(i|x^2|i) - 2(i|x|i)^2 \end{aligned}$$

8.9.3 EXAMPLES OF LOCALIZATION

Despite the freedom of the localization criterion choice, the results are usually similar. The orbitals of the CC and CH bonds in ethane, obtained by various approaches, are shown in Table 8.4.

Let us try to understand Table 8.4. First note the similarity of the results of various localization methods. The methods are different, the starting points are different, and yet we get almost the same in the end. It is both striking and important that

Table 8.4. The LCAO coefficients of the localized orbitals of ethane in the antiperiplanar conformation [P. Milli , B. L vy, G. Berthier, in: *“Localization and Delocalization in Quantum Chemistry”*, ed. O. Chalvet, R. Daudel, S. Diner, J.P. Malrieu, Reidel Publish. Co., Dordrecht (1975)] . Only the non-equivalent atomic orbitals have been shown in the table (four significant digits) for the CC and one of the equivalent CH bonds [with the proton H(1), Fig. 8.21]. The z axis is along the CC' bond. The localized molecular orbitals corresponding to the carbon inner shells 1s are not listed

	The projection method	Minimum distance method	Maximum repulsion energy
CC' bond			
1s(C)	−0.0494	−0.1010	−0.0476
2s(C)	0.3446	0.3520	0.3505
2p _z (C)	0.4797	0.4752	0.4750
1s(H)	−0.0759	−0.0727	−0.0735
CH bond			
1s(C)	−0.0513	−0.1024	−0.0485
2s(C)	0.3397	0.3373	0.3371
2p _z (C)	−0.1676	−0.1714	−0.1709
2p _x (C)	0.4715	0.4715	0.4715
1s(C')	0.0073	0.0081	0.0044
2s(C')	−0.0521	−0.0544	−0.054
2p _z (C')	−0.0472	−0.0503	−0.0507
2p _x (C')	−0.0082	−0.0082	−0.0082
1s(H1)	0.5383	0.5395	0.5387
1s(H2)	−0.0942	−0.0930	−0.0938
1s(H3)	−0.0942	−0.0930	−0.0938
1s(H4)	0.0580	0.0584	0.0586
1s(H5)	−0.0340	−0.0336	−0.0344
1s(H6)	−0.0340	−0.0336	−0.0344

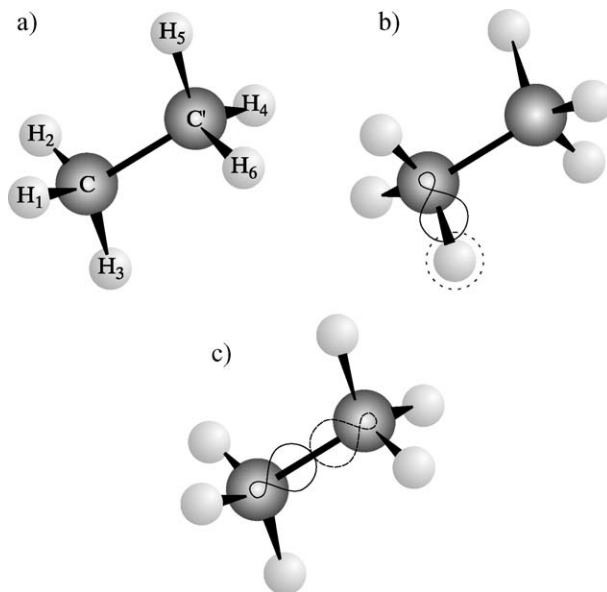


Fig. 8.21. The ethane molecule in the antiperiplanar configuration (a). The localized orbital of the CH bond (b) and the localized orbital of the CC' bond (c). The carbon atom hybrid forming the CH bond is quite similar to the hybrid forming the CC bond.

the results of various localizations are similar to one another, and in practical terms (not theoretically) we can speak of the unique definition of a chemical bond in a polyatomic molecule.

Nobody would reject the statement that a human body is composed of the head, the hands, the legs, etc. Yet a purist (i.e. theoretician) might get into troubles defining, e.g., a hand (where does it end up?). Therefore, purists would claim that it is impossible to define a hand, and as a consequence there is no such a thing as hand – it simply does not exist. This situation is quite similar to the definition of the chemical bond between two atoms in a polyatomic molecule.

It can be seen that some localized orbitals are concentrated mainly in *one* particular bond between two atoms. For example, in the CC bond orbital, the coefficients at the $1s$ orbitals of the hydrogen atom are small (-0.08). Similarly, the $2s$ and $2p$ orbitals of *one* carbon atom and *one* (the closest) hydrogen atom, dominate the CH bond orbital. Of course, localization is never complete. The oscillating “tails” of the localized orbital may be found even in distant atoms. They assure the mutual orthogonality of the localized orbitals.

8.9.4 COMPUTATIONAL TECHNIQUE

Let us take as an example the maximization of the electron interaction within the same orbital (Ruedenberg method):

$$I = \sum_i^{\text{MO}} \mathcal{J}_{ii} = \sum_i^{\text{MO}} (ii|ii). \quad (8.99)$$

Suppose we want to make an orthogonal transformation (i.e. a rotation in the Hilbert space, Appendix B) of – so far only two – orbitals:¹²⁰ $|i\rangle$ and $|j\rangle$, in order to maximize I . The rotation (an orthogonal transformation which preserves the orthonormality of the orbitals) can be written as

$$\begin{aligned} |i'(\vartheta)\rangle &= |i\rangle \cos \vartheta + |j\rangle \sin \vartheta, \\ |j'(\vartheta)\rangle &= -|i\rangle \sin \vartheta + |j\rangle \cos \vartheta, \end{aligned}$$

where ϑ is an angle measuring the rotation (we are going to find the optimum angle ϑ). The contribution from the changed orbitals to I , is

$$I(\vartheta) = (i'i'|i'i') + (j'j'|j'j'). \quad (8.100)$$

Then,¹²¹

$$\begin{aligned} I(\vartheta) &= I(0) \left(1 - \frac{1}{2} \sin^2 2\vartheta \right) + (2(ii|jj) + (ij|ij)) \sin^2 2\vartheta \\ &\quad + ((ii|ij) - (jj|ij)) \sin 4\vartheta, \end{aligned} \quad (8.101)$$

where $I(0) = (ii|ii) + (jj|jj)$ is the contribution of the orbitals before their rotation.

Requesting that $\frac{dI(\vartheta)}{d\vartheta} = 0$, we easily get the condition for optimum $\vartheta = \vartheta_{\text{opt}}$:

$$\begin{aligned} -2I(0) \sin 2\vartheta_{\text{opt}} \cos 2\vartheta_{\text{opt}} &+ (2(ii|jj) + (ij|ij)) 4 \sin 2\vartheta_{\text{opt}} \cos 2\vartheta_{\text{opt}} \\ &+ ((ii|ij) - (jj|ij)) 4 \cos 4\vartheta_{\text{opt}} = 0, \end{aligned} \quad (8.102)$$

and hence

$$\text{tg}(4\vartheta_{\text{opt}}) = 2 \frac{(ij|jj) - (ii|ij)}{2(ii|jj) + (ij|ij) - \frac{1}{2}I(0)}. \quad (8.103)$$

The operation described here needs to be performed for all pairs of orbitals, and then repeated (iterations) until the numerator vanishes for each pair, i.e.

$$(ij|jj) - (ii|ij) = 0. \quad (8.104)$$

The value of the numerator for each pair of orbitals is thus the criterion for whether a rotation is necessary for this pair or not. The matrix of the full orthogonal transformation represents the product of the matrices of these successive rotations.

The same technique of successive 2×2 rotations applies to other localization criteria.

¹²⁰The procedure is an iterative one. First we rotate one pair of orbitals, then we choose another pair and make another rotation etc., until the next rotations do not introduce anything new.

¹²¹Derivation of this formula is simple and takes one page.

8.9.5 THE σ , π , δ BONDS

Localization of the MOs leads to the orbitals corresponding to chemical bonds (as well as lone pairs and inner shells). In the case of a bond orbital, a given localized MO is in practice dominated by the AOs of *two atoms only*, those, which create the bond.¹²² According to the discussion on p. 371, the larger the overlap integral of the AOs the stronger the bonding. The energy of a molecule is most effectively decreased if the AOs are oriented in such a way as to maximize their overlap integral, Fig. 8.22. We will now analyze the kind and the mutual orientation of these AOs.

As shown in Fig. 8.23, the orbitals σ , π , δ (either canonical or not) have the following features:

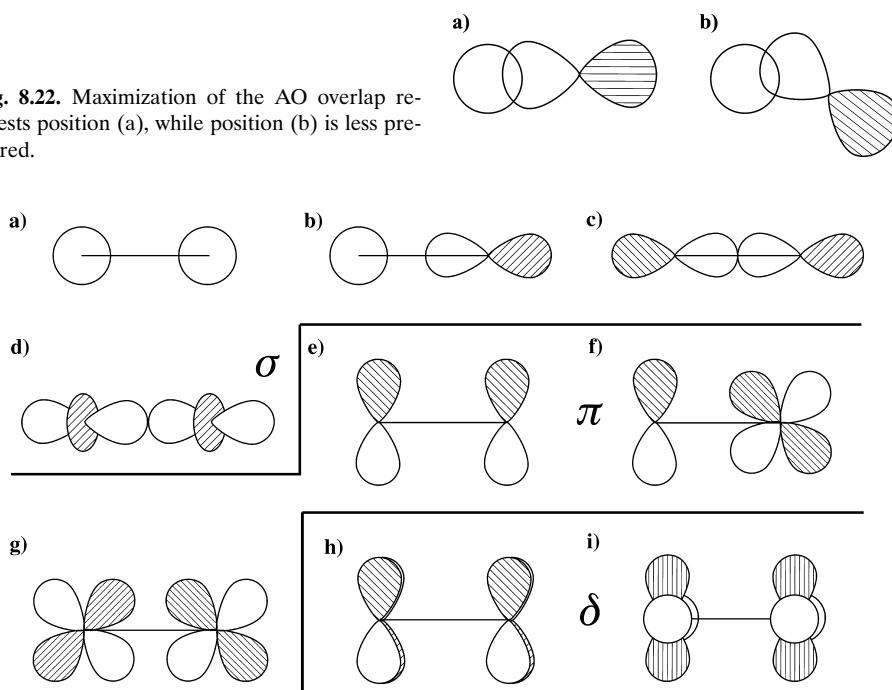


Fig. 8.23. Symmetry of the MOs results from the mutual arrangement of those AOs of both atoms which have the largest LCAO coefficients. Figs. (a–d) show the σ type bonds, (e–g) the π type bonds, and (h,i) the δ type bonds. The σ bond orbitals have no nodal plane (containing the nuclei), the π orbitals have one such plane, the δ ones – two such planes. If the z axis is set as the bond axis, and the x axis is set as the axis perpendicular to the bonding and lying in the plane of the figure, then the cases (b–i) correspond (compare Chapter 4) to the overlap of the following AOs: (b): s with p_z , (c): p_z with p_z , (d): $3d_{3z^2-r^2}$ with $3d_{3z^2-r^2}$, (e): p_x with p_x , (f): p_x with $3d_{xz}$, (g): $3d_{xz}$ - $3d_{xz}$, (h): $3d_{xy}$ with $3d_{xy}$, (i): $3d_{x^2-y^2}$ with $3d_{x^2-y^2}$. The figures show such atomic orbitals which correspond to the *bonding* MOs. To get the corresponding *antibonding* MOs, we need to change the sign of *one* of the two AOs.

¹²²That is, they have the largest absolute values of LCAO coefficients.

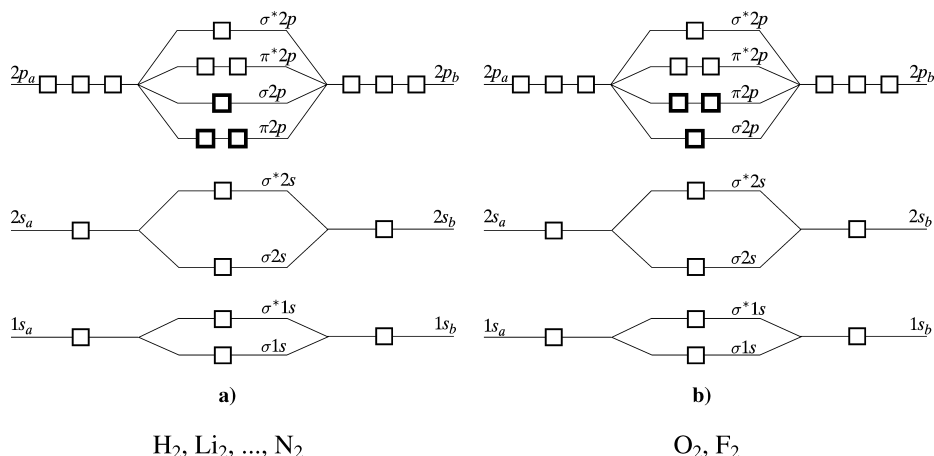


Fig. 8.24. Scheme of the bonding and antibonding MOs in homonuclear diatomics from H_2 through F_2 . This scheme is better understood after we recall the rules of effective mixing of AOs, p. 362. All the orbital energies become lower in this series (due to increasing of the nuclear charge), but lowering of the bonding π orbitals leads to changing the *order* of the orbital energies, when going from N_2 to F_2 . The sequence of orbital energies (schematically) for the molecules (a) from H_2 through N_2 and (b) for O_2 and F_2 .

- the σ -type orbital has *no nodal plane going through the nuclei*,
- the π -type orbital has *one such a nodal plane*,
- the δ -type orbital has *two such nodal planes*.

If a MO is antibonding, then a little star is added to its symbol, e.g., σ^* , π^* , etc. Usually we also give the orbital quantum number (in order of increasing energy), e.g., $1\sigma, 2\sigma, \dots$ etc. For homonuclear diatomics additional notation is used (Fig. 8.24) showing the main atomic orbitals participating in the MO, e.g., $\sigma 1s = 1s_a + 1s_b$, $\sigma^* 1s = 1s_a - 1s_b$, $\sigma 2s = 2s_a + 2s_b$, $\sigma^* 2s = 2s_a - 2s_b$, etc.

The very fact that the π and δ molecular orbitals have zero value at the positions of the nuclei (the region most important for lowering the potential energy of electrons) suggests that they are bound to be of higher energy than the σ ones, and they are indeed.

8.9.6 ELECTRON PAIR DIMENSIONS AND THE FOUNDATIONS OF CHEMISTRY

What are the dimensions of the electron pairs described by the localized MOs? Well, but how to define such dimensions? All orbitals extend to infinity, so you cannot measure them easily, but some may be more diffuse than others. It also depends on the molecule itself, the role of a given MO in the molecular electronic structure (the bonding orbital, lone electron pair or the inner shell), the influence

of neighbouring atoms, etc. These are fascinating problems, and the issue is at the heart of structural studies of chemistry.

Several concepts may be given to calculate the dimensions of the molecular orbitals mentioned above. For example, we may take the integrals $\langle ii|r_{12}^2|ii\rangle \equiv \langle r^2\rangle$ calculated within the Boys localization procedure, and use them to measure the square of the dimension of the (normalized) molecular orbital φ_i . Indeed, $\langle r^2\rangle$ is the mean value of the interelectronic distance for a two-electron state $\varphi_i(1)\varphi_i(2)$, and $\rho_i(\text{Boys}) = \sqrt{\langle r^2\rangle}$ may be viewed as an estimate of the φ_i orbital dimension. Or, we may do a similar thing by the Ruedenberg method, by noting that the Coulombic integral \mathcal{J}_{ii} , calculated in atomic units, is nothing more than the mean value of the *inverse of the distance* between two electrons described by the φ_i orbital. In this case, the dimension of the φ_i orbital may be proposed as $\rho_i(\text{Ruedenberg}) = \frac{1}{\mathcal{J}_{ii}}$. Below, the calculations are reported, in which the concept of $\rho_i(\text{Boys})$ is used. Let us compare the results for CH_3OH and CH_3SH (Fig. 8.25) in order to see, what makes these two molecules so different,¹²³ Table 8.5.

Interesting features of both molecules can be deduced from Table 8.5. The most fundamental is whether formally the same chemical bonds (say, the CH ones) are indeed similar for both molecules. A purist approach says that each molecule is a New World, and thus these are two different bonds by definition. Yet chemical intuition says that some local interactions (in the vicinity of a given bond) should mainly influence the bonding. If such local interactions are similar, the bonds should turn out similar as well. Of course, the purist approach is formally right, but the New World is quite similar to the Old World, because of local interactions. If chemists desperately clung to purist theory, they would know some 0.01% or so of what they now know about molecules. *It is of fundamental importance for chemistry that we do not study particular cases, case by case, but derive general rules.* Strictly speaking, these rules are false from the very beginning, for they are valid to some extent only, but they enable chemists to understand, to operate, and to be efficient, otherwise there would be no chemistry at all.

The periodicity of chemical elements discovered by Mendeleev is another fundamental idea of chemistry. It has its source in the shell structure of atoms. Following on, we can say that the compounds of sulphur with hydrogen should be

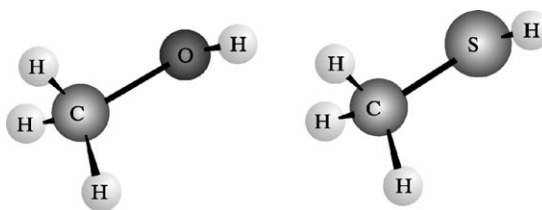


Fig. 8.25. Methanol (CH_3OH) and methanethiol (CH_3SH).

¹²³Only those who have carried out experiments in person with methanethiol (knowns also as methyl mercaptan), or who have had neighbours (even distant ones) involved in such experiments, understand how important the difference between the OH and SH bonds really is. In view of the theoretical results reported, I am sure they also appreciate the blessing of theoretical work. According to the Guinness book of records, CH_3SH is the most smelly substance in the Universe.

Table 8.5. The dimensions of the electron pairs, i.e. $\sqrt{\langle r^2 \rangle}$ (a.u.) for CH₃OH and CH₃SH according to Csizmadia.^a “Core” means the 1s orbital of the atom indicated

CH ₃ OH		CH ₃ SH	
core O	0.270	core S (1s)	0.148
core C	0.353	core C	0.357
		S (L shell)	0.489
			0.489
			0.483
			0.484
CO	1.499	CS	2.031
CH ₁ ^b	1.576	CH ₁ ^b	1.606
CH _{2,3} ^b	1.576	CH _{2,3} ^b	1.589
OH	1.353	SH	1.847
lone pair _{1,2} ^c	1.293	lone pair _{1,2} ^c	1.886

^aI.G. Csizmadia, in “*Localization and Delocalization in Quantum Chemistry*”, ed. by O. Chalvet and R. Daudel, D. Reidel Publ. Co., Dordrecht, 1975.

^bDifferent electron pair dimensions originate from their different positions vs the OH or SH group.

^cThere are two lone pairs in the molecule.

similar to the compounds of oxygen with hydrogen, because sulphur and oxygen have analogous electronic configuration of the valence electrons (i.e. those of the highest energies), and they differ only in the inner shells (O: [He]2s²2p⁴ as compared to S: [Ne]3s²3p⁴).

Take a look of Table 8.5. Note that:

1. The dimension of the electron lone pair localized on the 1s orbital of the sulphur atom is twice as small as the dimension of a similar pair of the 1s orbital of the oxygen atom. Nothing special. The electrons occupying the 1s orbital of S experience a strong electric field of the nucleus charged +16, while the charge of the O nucleus is only +8. Let us note that the core of the carbon atom is even larger, because it is controlled by an even less charged nucleus¹²⁴ (+6).
2. The dimension of the electron pair of the 1s orbital of the carbon atom (core C) for CH₃OH is very similar to that of the corresponding orbital for CH₃SH (0.353 vs 0.357).¹²⁵ This means that the influence of the S atom (as compared to the oxygen atom) on the 1s orbital of the *neighbouring atom* is small. The *local character of the interactions* is thus the most decisive.
3. The influence of the S and O atoms on the CH bonds of the methyl group is only slightly larger. For example, in CH₃OH one of the CH bond localized orbital has the dimension of 1.576 a.u., while in CH₃SH 1.606 a.u.
4. The three CH bonds in methanol are very similar to each other (the numbers in Table 8.5 are identical), yet only two of them are strictly equivalent

¹²⁴The mean value of the nucleus–electron distance can be easily computed as $\langle 1s|r|1s \rangle = \frac{3}{2} \frac{1}{Z}$, where Z is the charge of the nucleus. The results discussed are consistent with such a simple calculation.

¹²⁵Even these small changes may be detected experimentally by removal of electrons from the molecules by monochromatic X-ray radiation and subsequent measurement of the kinetic energy of the removed electrons. Those which were more strongly bound, move slower.

due to symmetry. It is even more interesting that the CH bonds in CH₃SH are also similar to them, although the differences between the various CH bonds of methanethiol, and between the corresponding CH bonds in methanol and methanethiol, are clearer. So, *even despite the different atomic environment, the chemical bond preserves its principal and individual features.*

5. Let me apologize for a banality: CH₃SH differs from CH₃OH in that the O atom is replaced by the S atom. No wonder then that large differences in the close vicinity of the O and S atoms are easily noticeable. The dimensions of the electron pairs at the S atom (lone pairs and the SH and CS bonds) are always larger than the corresponding pair at the O atom. The differences are at the 30% level. The sulphur atom is simply larger than the oxygen atom, indicating that the electrons are more loosely bound when we go down within a given group of the periodic table.

These conclusions are instructive and strongly encouraging, because *we see a locality in chemistry, and therefore chemistry is easier than it might be* (e.g., CH bonds have similar properties in two different molecules). On the other hand, we may play a subtle game with local differences on purpose, by making suitable chemical substitutions. In this way we have a possibility of *tuning the chemical and physical properties of materials, which is of prime importance in practical applications.*

8.9.7 HYBRIDIZATION

The localized orbitals may serve to present the idea of a *hybrid atomic orbital*. A given localized orbital φ of a bond represents a linear combination of the atomic orbitals of *mainly two atoms* – the partners which form the chemical bond, say *a* and *b*. If so, then (for each localized bond orbital) all the atomic orbitals of atom *a* may be added together with their specific LCAO coefficients,¹²⁶ and the same can be done for atom *b*. These two sums represent two *normalized hybrid atomic orbitals* χ_a and χ_b multiplied by the resulting coefficients c_a and c_b and together form the approximate¹²⁷ bond orbital:

$$\varphi \approx c_a h_a + c_b h_b$$

with the corresponding LCAO expansions

$$h_a = \sum_{j \in a} c_{ji} \chi_j,$$

$$h_b = \sum_{j \in b} c_{ji} \chi_j.$$

¹²⁶That serve to express the localized orbital through the atomic basis set.

¹²⁷The “tails” of the localized orbital, i.e. its amplitudes on other atoms, have been neglected.

Such a definition of the hybrid orbitals is not unique, since the localized orbitals used are also not unique. However, as shown above, this ambiguity is of secondary importance. The advantage of such an approach to hybridization is that:

- It can be determined for any configuration of the nuclei, e.g., for the tetrahedral as well as for any other configuration of CH_4 , etc.
- The definition is applicable at any LCAO basis set used.
- It gives a clear message that *all* the atoms in a molecule are hybridized (why not?), e.g., the carbon atom in the methane molecule as well as all the hydrogen atoms. The only difference between these two hybridizations is that the χ_a for the carbon atom does not resemble any of the χ_j in $\sum_{j \in a} c_{ji} \chi_j$ (because of comparable values¹²⁸ of $|c_{ji}|$ meaning an effective mixing of the atomic orbitals), while the χ_b for the hydrogen atom is *dominated by the single atomic orbital* $1s_b$, which may be treated as a lack of hybridization.¹²⁹

sp³ hybridization (tetrahedral)

How will the hybridization in the optimized geometry of methane look? Well, among five doubly occupied localized molecular orbitals, four¹³⁰ protrude from the carbon nucleus towards one of the hydrogens (four hydrogens form a regular tetrahedron) and will have only some marginal amplitudes on the three other hydrogens. If we neglect these “tails” on the other atoms and the contributions of the atomic orbitals other than $2s$ and $2p$ (i.e. their c_{ji} ’s) of the carbon atom (also eliminating from the MO the $1s$ orbital of the partner hydrogen atom), we obtain the following normalized hybrid carbon orbitals:

$$h_i = \frac{1}{\sqrt{1 + \lambda_i^2}} [(2s) + \lambda_i (2p_i)]$$

for $i = 1, 2, 3, 4$. If we force the four hybrids to be equivalent, then this means $\lambda_i = \lambda$. Forcing the hybrids to be mutually orthogonal:¹³¹

$$\langle h_i | h_j \rangle = \frac{1}{1 + \lambda^2} [1 + \lambda^2 \langle 2p_i | 2p_j \rangle] = \frac{1}{1 + \lambda^2} [1 + \lambda^2 \cos \theta_{ij}] = 0,$$

we obtain as the $2s$ and $2p$ mixing ratio

$$\lambda = \sqrt{\frac{-1}{\cos \theta_{ij}}}. \quad (8.105)$$

¹²⁸Mainly of $2s_a$ and $2p_a$, which have the highest values of the LCAO coefficients.

¹²⁹The reason why the carbon atom (and some other atoms such as N, O, etc.) is effectively hybridized, while the hydrogen atom not, is that the $2s$ and $2p$ orbital energy levels in those atoms are close in the energy scale, while the energy difference between the $1s$ hydrogen orbital energy and higher-energy hydrogen orbitals is larger.

¹³⁰The fifth will be composed mainly of the $1s$ carbon orbital.

¹³¹“Orthogonal” also means “absolutely independent”.

Since, for the tetrahedral configuration $\theta_{ij} = 109^\circ 28'$, hence, from eq. (8.105): $\cos 109^\circ 28' = -\frac{1}{3}$ and $\lambda_i = \sqrt{3}$. Therefore the orthogonal hybrids on the carbon atom (Figs. 8.26.h and 8.27.a) read as:

$$h_i(sp^3) = \frac{1}{2}[(2s) + \sqrt{3}(2p_i)],$$

where $2s$ and $2p_i$ are the normalized carbon atomic orbitals¹³² with i denoting the direction of the hybrid, one of the four directions from the carbon atom towards the tetrahedrally located hydrogen atoms.¹³³

sp² hybridization (trigonal)

If we tried to find the lowest-energy configuration of ethylene (C_2H_4), it would correspond to a planar structure (Fig. 8.27.b) of D_{2h} symmetry. After analyzing the localized molecular orbitals, it would turn out that three hybrids protrude from each carbon nucleus, their directions lying in the molecular plane (say, xy). These hybrids form angles very close to 120° .

For the trigonal hybridization (i.e. pure sp^2 hybridization, with the $\theta_{ij} = 120^\circ$ angles) we obtain from (8.105) $\lambda = \sqrt{2}$, and, therefore, the three orthogonal normalized sp^2 hybrids are:

$$h_i(sp^2) = \frac{1}{\sqrt{3}}[(2s) + \sqrt{2}(2p_i)],$$

where the directions $i = 1, 2, 3$ form the Mercedes logo on a plane.

sp hybridization (digonal)

Such hybridization is said to occur in acetylene: $HCCH$, which after optimization of the Hartree–Fock energy, corresponds to the linear symmetric configuration. According to this explanation, each carbon atom exposes two hybrids (Fig. 8.27c): one towards its carbon and one towards its hydrogen partner. These hybrids use

¹³²Say, the Slater Type Orbitals (STOs), p. 355.

¹³³Such orientation of the (normalized) $2p_i$'s may be achieved by the following choices (just look at the vortices of a cube with the carbon atom at its centre and the four directions forming the tetrahedron):

$$\begin{aligned} 2p_1 &= \frac{1}{\sqrt{3}}(2p_x + 2p_y + 2p_z), \\ 2p_2 &= \frac{1}{\sqrt{3}}(2p_x - 2p_y - 2p_z), \\ 2p_3 &= \frac{1}{\sqrt{3}}(-2p_x + 2p_y - 2p_z), \\ 2p_4 &= \frac{1}{\sqrt{3}}(-2p_x - 2p_y + 2p_z). \end{aligned}$$

The normalization of the above functions is obvious, since the $2p_x, 2p_y, 2p_z$ are orthogonal.

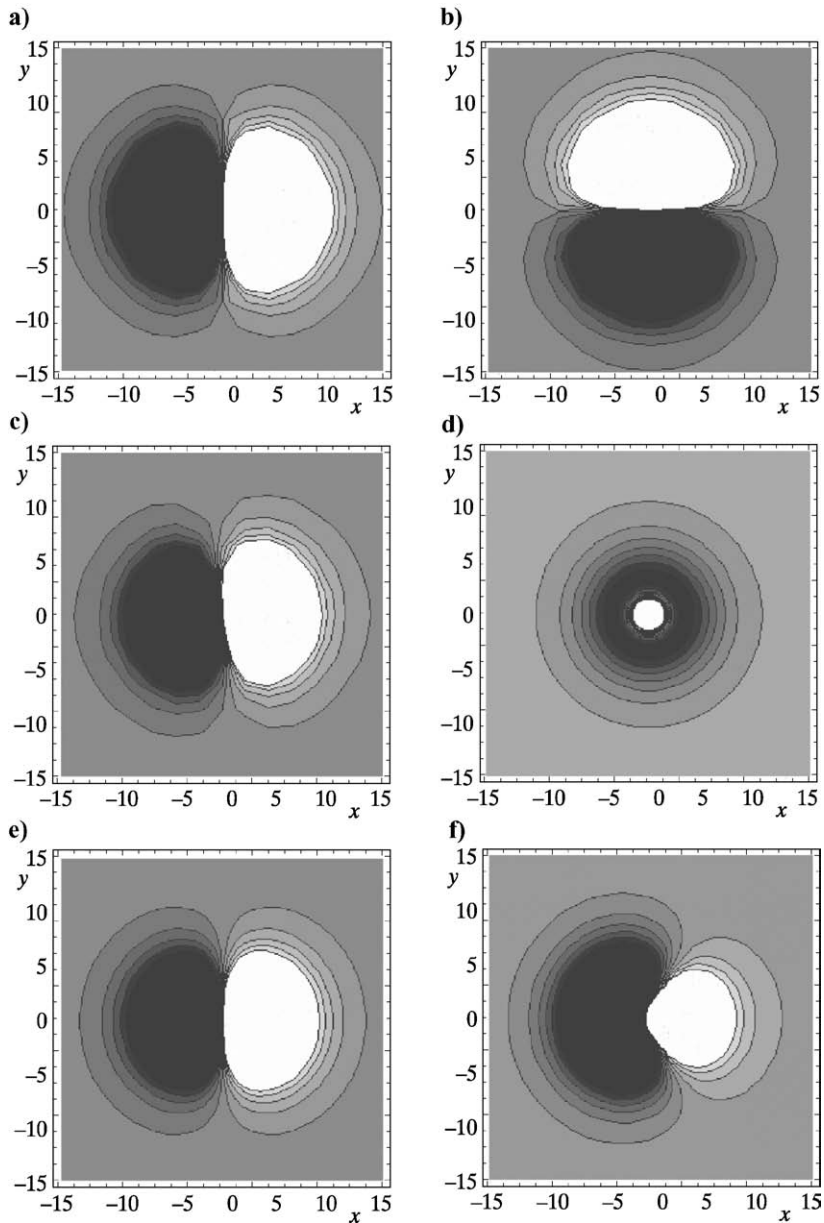


Fig. 8.26. The Slater-type orbitals shown as contours of the section at $z = 0$. The background corresponds to the zero value of the orbital, the darker regions to the negative, the brighter to the positive value of the orbital. (a) $2p_x$ and (b) $2p_y$, and their linear combination (c) equal to $\cos 5^\circ 2p_x + \sin 5^\circ 2p_y$, which is also a $2p$ orbital, but rotated by 5° with respect to the $2p_x$ orbital. In (d) and (e) we show the normalized $2s$ and $2p$ orbitals, which will now be mixed in various proportions: (f) the 1 : 1 ratio, i.e. the sp hybridization, (g) the 1 : $\sqrt{2}$ ratio, i.e. the sp^2 hybridization, and (h) the 1 : $\sqrt{3}$ ratio, i.e. the sp^3 hybridization.

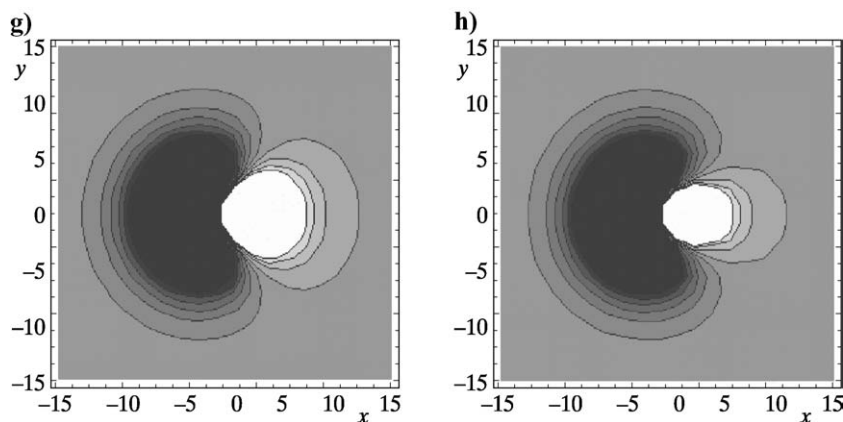


Fig. 8.26. Continued.

the two carbon $2s$ and the two carbon $2p_z$, and together with the two $1s$ orbitals of the hydrogens, form the two HC σ bonds and one CC σ bond. This means that each carbon atom has two electrons left, which occupy its $2p_x$ and $2p_y$ orbitals (perpendicular to the molecular axis). The $2p_x$ orbitals of the two carbon atoms form the doubly occupied π_x bonding localized molecular orbital and the same happens to the $2p_y$ orbitals. In this way the carbon atoms form the C \equiv C triple bond composed of one σ and two π (i.e. π_x and π_y) bonds.

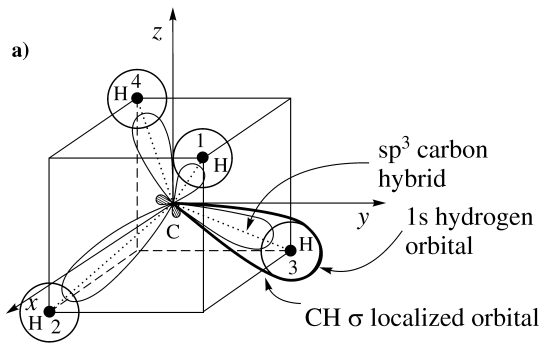
The angle between the two equivalent orthonormal hybrids should be $\theta_{ij} = 180^\circ$, then the mixing ratio will be determined by $\lambda = 1$. Two such hybrids are therefore¹³⁴ $h_i(sp) = \frac{1}{\sqrt{2}}[(2s) + (2p_i)]$, and making the two opposite directions explicit: $h_1(sp) = \frac{1}{\sqrt{2}}[(2s) + (2p_z)]$ and $h_2(sp) = \frac{1}{\sqrt{2}}[(2s) - (2p_z)]$.

Is hybridization concept of any value?

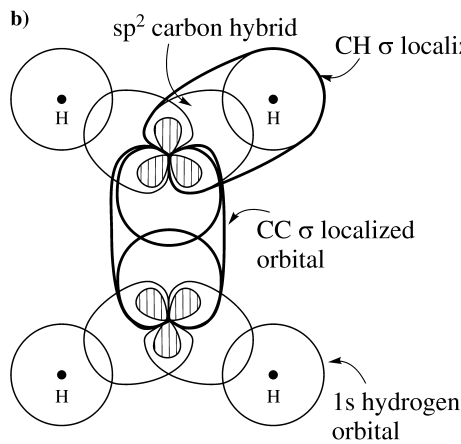
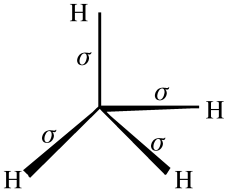
The general chemistry textbook descriptions of hybridization for methane, ethylene and acetylene usually start from the electronic configuration of the carbon atom: $1s^2 2s^2 2p^2$. Then it is said that, according to valence bond theory (VB, see Chapter 10), this configuration predicts CH₂ as the carbon hydride (bivalent carbon atom) with the CH bonds forming the right angle.¹³⁵ This differs very much from the way the methane molecule looks in reality (regular tetrahedron and tetravalent carbon). If the carbon atom were excited (this might happen at the expense of future energy gains and is known as “promotion”) then the configuration might look like $1s^2 2s^1 2p_x^1 2p_y^1 2p_z^1$. The textbooks usually go directly to the mixing of the valence atomic orbitals $2s, 2p_x 2p_y 2p_z$ to form four equivalent sp^3 hybrids,

¹³⁴This cannot be exact (cf. the ethylene case), because the two hybrids must not be equivalent. One corresponds to the CC, the other to the CH bond.

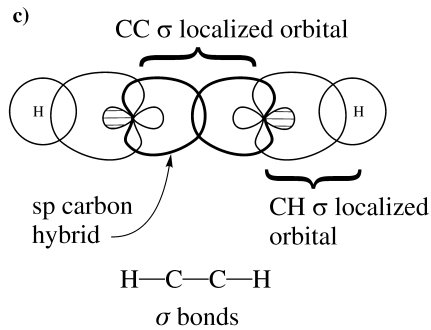
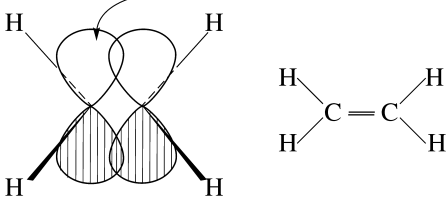
¹³⁵Because $2p^2$ means, say, $2p_x^1 2p_y^1$ and these singly occupied atomic orbitals form the two CH bonds with two $1s$ hydrogen orbitals.



METHANE



ETHYLENE



ACETYLENE

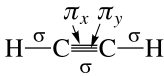
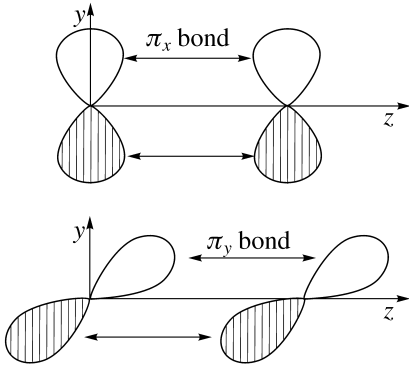


Fig. 8.27. (a) The sp^3 hybridization in the methane molecule in its tetrahedral equilibrium geometry [that corresponds to the minimum of the ground-state electronic energy $E_0^0(\mathbf{R})$, see p. 229]. There are four doubly occupied CH σ localized molecular orbitals and one that is essentially the doubly occupied $1s$ carbon atomic orbital. Each of the CH molecular orbital (of the *nearly* cylindrical symmetry) is composed mainly of the carbon hybrid and the hydrogen $1s$ atomic orbital. The figure shows a scheme of the four carbon hybrids called the sp^3 hybrids. (b) An example of the nearly perfect sp^2 hybridization of the carbon atoms in the ethylene (C_2H_4), which is perfectly planar in its ground electronic state (D_{2h} symmetry). Such a hybridization is only approximate, because the CCH angle has to differ from the HCH angle, both slightly deviate from 120° . The localized molecular orbitals are the following (occupied by altogether 16 electrons):

- two essentially $1s$ carbon orbitals,
- four CH orbitals and one CC orbital having the *nearly* cylindrical symmetry (i.e. σ type),
- one bond orbital being antisymmetric with respect to the reflection in the molecular plane (i.e. of the π symmetry).

(c) An example of the sp hybridization: the acetylene molecule. The Hartree–Fock geometry optimization gives the lowest-energy linear configuration: HCCH. The localization gives seven localized doubly occupied molecular orbitals:

- two of them are essentially the $1s$ carbon orbitals,
- two represent the cylindrical CH orbitals (σ),
- one cylindrical CC σ orbital,
- two CC orbitals that are of π symmetry (perpendicular to each other).

which lead directly to the tetrahedral hydride: the methane. Note, however, that being in less of a rush, we would draw the conclusion that the $1s^2 2s^1 2p_x^1 2p_y^1 2p_z^1$ configuration leads to four *non*-equivalent CH bonds in the CH_4 hydride.¹³⁶ Only equivalent mixing (hybridization) gives the correct picture. When aiming at ethylene or acetylene the reasoning changes, because some orbitals are left without mixing. We assume sp^2 (one orbital left) or sp (two orbitals left) hybridizations, respectively, which leads to the correct compounds and (almost) correct structures. *It looks as if when we know what we are aiming for, we decide what mixes and what does not.* This seems to be not fair.

Let us check how important the role of hybridization is in the formation of chemical bonds in methane. Let us imagine four scientists performing Hartree–Fock computations for methane in its tetrahedral configuration¹³⁷ of nuclei. They use four LCAO basis sets. Professor A believes that in this situation it is important to remember sp^3 hybridization and uses the following basis set (first go the $1s$ orbitals for the hydrogen atoms, then the carbon orbitals):

$$A: 1s_{H1}, 1s_{H2}, 1s_{H3}, 1s_{H4}, 1s_C, h_1(sp^3), h_2(sp^3), h_3(sp^3), h_4(sp^3).$$

Student B did not read anything about hybridization and just uses the common orbitals:

$$B: 1s_{H1}, 1s_{H2}, 1s_{H3}, 1s_{H4}, 1s_C, 2s_C, 2p_{x,C}, 2p_{y,C}, 2p_{z,C}.$$

¹³⁶Three CH bonds would form right angles (because of $2p_x^1, 2p_y^1, 2p_z^1$), one CH bond however (formed by $2s^1$ together with the corresponding $1s$ hydrogen orbital) would have a quite different character. This contradicts what we get from experiments.

¹³⁷Or any other one.

Students C and D are not the brightest, they have mixed up the hybridization for methane with that for ethylene and acetylene and used the following basis sets:

$$\text{C: } 1s_{H1}, 1s_{H2}, 1s_{H3}, 1s_{H4}, 1s_C, 2p_{x,C}, h_1(sp^2), h_2(sp^2), h_3(sp^2).$$

$$\text{D: } 1s_{H1}, 1s_{H2}, 1s_{H3}, 1s_{H4}, 1s_C, 2p_{x,C}, 2p_{y,C}, h_1(sp), h_2(sp).$$

Who of these scientists will obtain the lowest total energy, i.e. the best approximation to the wave function?

Well, we could perform these calculations, but it is a waste of time. Indeed, each of the scientists used different basis sets, but they all used *the same space* spanned by the AOs. This is because all these hybrids are linear combinations of the orbitals of student B. All the scientists are bound to obtain the same total energy, the same molecular orbitals¹³⁸ and the same orbital energies.

Hybridization is useful

Is hybridization a useless concept then? No, it is not. It serves as a first indicator (when calculations are not yet performed) of what happens to a local atomic electronic structure, if the atomic configuration is tetrahedral, trigonal, etc. For example, the trigonal hybrids describe the main features of the electronic configuration in the benzene molecule, Fig. 8.28.

Let us take a slightly more complicated example of what is known as a peptide bond (of great importance in biology), Fig. 8.29.

It is important to remember that we always start from some chemical intuition¹³⁹ and use the structural formula given in Fig. 8.29.a. Most often we do not even consider other possibilities (isomers), like those shown in Fig. 8.29.b. Now, we try to imagine what kind of local electronic structure we have around the particular atoms. Let us start from the methyl, i.e. $-\text{CH}_3$ functional groups. Of course, such a group resembles methane, except that one carbon hybrid extends to another atom (not hydrogen). Thus, we expect hybridization over there close to sp^3 one (with all consequences, i.e. angles, etc.). Next, we have the carbon atom that is believed¹⁴⁰ to make the double bond with the oxygen atom. The double bond means an ethylene-like situation, i.e. both atoms should have hybridizations similar to sp^2 .

¹³⁸Although the LCAO coefficients will be, of course, different, because the expansion functions are different. The orbital plots will be the same.

¹³⁹Based on the vast experience of chemists

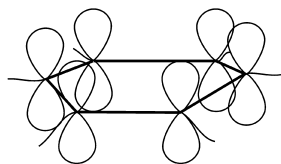
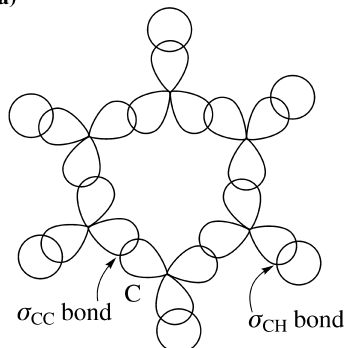
¹⁴⁰Here we rely on the concept of what is known as the *valency* of atoms, i.e. the number of bonds a given atom is able to form with its neighbours. The valency is equal to the number of valence electrons or valence holes, e.g., the valency of the carbon atom is four (because its electron configuration is $K2s^22p^2$, four valence electrons), of the oxygen atom is two (because its electron configuration is $K2s^22p^4$, two valence holes). An element may have several valencies, because of the opening several electronic shells.

We are making several assumptions based on chemical intuition or knowledge. The reason is that we want to go quickly without performing any computations. *This ambiguity disappears, if we make real computations, e.g., using the Hartree–Fock method. Then the chemical bonds, hybrids etc. are obtained as a result of the computations.*

BENZENE:
PLANAR σ BOND STRUCTURE

BENZENE:
 π BOND FORMING ORBITALS

a)



b)

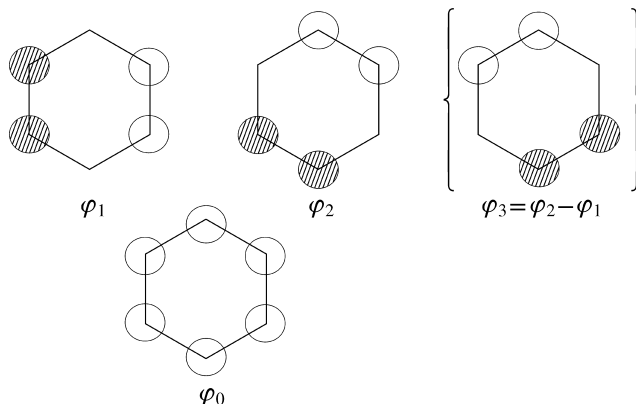


Fig. 8.28. The benzene molecule. The hybridization concept allows us to link the actual geometry of a molecule with its electronic structure (a). The sp^2 hybrids of the six carbon atoms form the six σ CC bonds and the structure is planar. Each carbon atom thus uses two out of its three sp^2 hybrids, the third one lying in the same plane protrudes towards a hydrogen atom and forms the σ CH bond. In this way, each carbon atom uses its three valence electrons. The fourth one resides on the $2p$ orbital that is perpendicular to the molecular plane. The six $2p$ orbitals form six π molecular orbitals, out of which three are doubly occupied and three are empty. The doubly occupied ones are shown in Fig. (b). The φ_0 of the lowest-energy is an all-in-phase linear combination of the $2p$ atomic orbitals (only their upper lobes are shown). The φ_1 and φ_2 correspond to higher and to the same energy, and have a single node (apart from the node plane of the AOs). The φ_3 orbital that apparently completes all combinations of single-node molecular orbitals is redundant (that is why it is in parentheses), because the orbital represents a linear combination of the φ_1 and φ_2 .

Let us begin from the oxygen atom, Fig. 8.29.c. The sp^2 means three hybrids (planar configuration) protruding from the O atom. One of them will certainly bind to a similar one protruding from the carbon atom (OC σ bond), it therefore needs a only single electron from the oxygen. The oxygen atom has 6 valence electrons, therefore there remain five to think of. Four of them will occupy the other two hybrids protruding into space (nothing to bind; they are lone pairs). Hence there

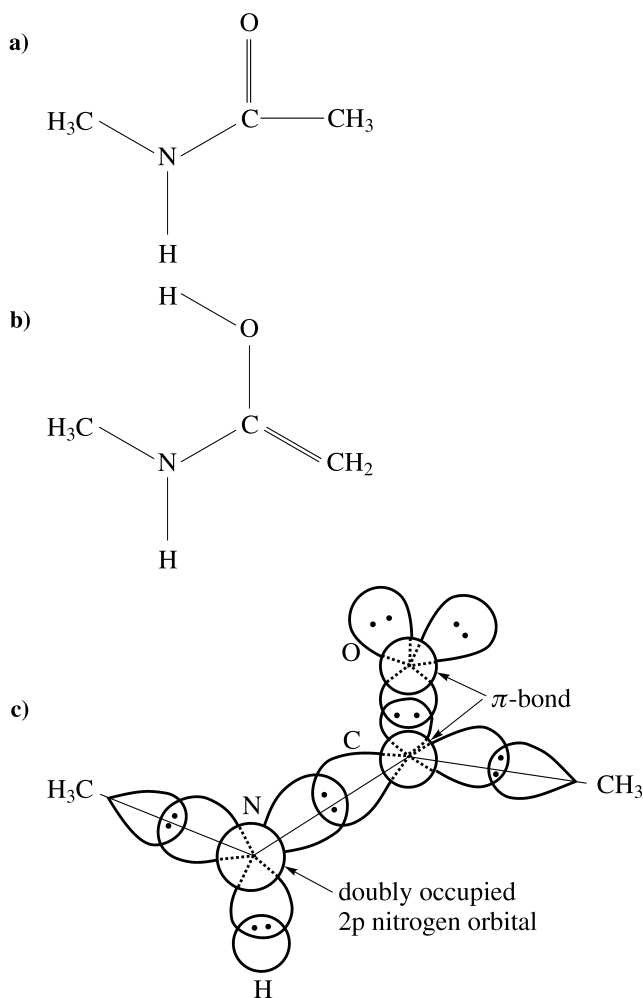


Fig. 8.29. How does the hybridization concept help? The figure shows the all important (proteins) example of the peptide bond. (a) We assume a certain pattern of the chemical bonds (this choice is knowledge based) ignoring other possibilities, such as the isomers shown in (b). Apart from the methyl groups (they have the familiar tetrahedral configuration) the molecule is planar. Usually in chemistry, knowing the geometry, we make a conjecture pertaining to the hybridization of particular atoms. This leads to the electron count for each atom: the electrons left are supposed to participate in bonds with other atoms. In the example shown, the sp^2 hybridization is assumed for the central carbon and for the nitrogen and oxygen atoms (c). A π bonding interaction of the nitrogen, carbon and oxygen should therefore stabilize the planarity of the system, which is indeed an experimental fact.

is one electron left. This is very good, because it will participate in the OC π bond. Let us go to the partner carbon atom. It is supposed to make a double bond with the oxygen. Hence it is reasonable to ascribe to it an ethylene-like hybridization as well. Out of four valence electrons for carbon, two are already used up by the σ and π CO bonds. Two other sp^2 hybrids remain that, of course, accommodate the two electrons and therefore are able to make two σ bonds: one with $-\text{CH}_3$ and one with the nitrogen atom. Then we go to the nitrogen atom. It has three substituents in most cases in the (almost) planar configuration (we know this from experiment). To make the analysis simple, we assume an sp^2 ideal hybridization. The nitrogen atom has five valence electrons. Three of them will go to form the σ NC, NH, N- CH_3 bonds. Note, that although the configuration at N is assumed to be planar, this plane may not coincide with the analogous plane on the carbon atom. Finally,

we predict the last two valence electrons of the nitrogen will occupy the $2p$ orbital perpendicular to the plane determined by the substituents of the nitrogen. Note that the $2p$ orbital could overlap (making a bonding effect) with the analogous $2p$ orbital of the carbon atom *provided that the two planes will coincide*. This is why we could expect the planarity of the O–C–N–H, known as peptide bond. This bond plays a prominent role in proteins, because it is responsible for making the chain of amino acid residues. It is an experimental fact that *deviations of the peptide bond from planarity are very small*.

The value of the analyses, as is given above, is limited to qualitative predictions. Of course, computations would give us a much more precise picture of the molecule. In such computations the orbitals would be more precise, or would not be present at all, because, to tell the truth, there is no such thing as orbitals. We badly need to interpret the numbers, to communicate them to others in a understandable way, to say whether we understand these numbers or they are totally unexpected. Reasoning like that given above has a great value as part of our understanding of chemistry, of speaking about chemistry, of predicting and of discussing structures. This is why we need hybridization. Moreover, if our calculations were performed within the VB method (in its simplest formulation; the details of the method will be explained in Chapter 10), then the lowest energy would be obtained by Professor A (who assumed the sp^3 hybridization), because the energy gain over there is very much connected to the overlap of the atomic orbitals forming the basis, and the overlap with the $1s$ hydrogen orbitals is the best for the basis set of Professor A. The other people would get high total energies, because of poor overlap of their atomic orbitals with the $1s$ hydrogen orbitals.

8.10 A MINIMAL MODEL OF A MOLECULE

It is easy to agree that our world is a complex business. It would be great, however, to understand how the world is operating. Answers look more and more complex as we go from crude to more and more accurate theories. Therefore, we would like to consider a simpler world (say, a model of our real world), which

- would work to very good accuracy, i.e. resembled the real world quite well,
- would be based on such simple rules that we could understand it in detail.

We could explain these rules to anybody who were interested. Not only could we predict almost anything, but we ourselves could be confident that we understand *most* of chemistry, because it is based on *several simple rules*. Moreover, why worry about details? Most often we want just to grasp the essence of the problem. On top of that, if this essence were free, only sometimes would we be interested in a more detailed (and expensive) picture.

Is this utopia or can such a model of chemistry be built?

Well, it seems that theoretical chemistry nowadays offers such a model describing chemical structures.

The model is based on the following basic simplifications of the real world:

- The *non-relativistic approach*, i.e. the speed of light is assumed to be infinite, which leads to the Schrödinger equation (Chapter 2).
- The Born–Oppenheimer approximation (Chapter 6) that separates the motion of the nuclei from the motion of the electrons. This approximation allows us to introduce the concept of the *3D structure* of the molecule: the heavy nuclear framework of the molecule kept together by “electronic glue” moves in space (translation), and at the same time rotates in space.
- The mean-field approximation of the present Chapter offers us the *orbital model* of the electronic structure of molecules within the Restricted Hartree–Fock approach. In this picture the electrons are described by the doubly occupied molecular orbitals. Localization of the orbitals gives the doubly occupied *inner shell, lone pair and bond* molecular orbitals. The first and second are sitting on atoms, the latter on chemical bonds. Not all atoms are bound to each other, but instead *the molecule has a pattern of chemical bonds*.
- These bonds are traditionally and formally represented as *single*: e.g., C–C; *double*, e.g., C=C or *triple*, e.g., C≡C, although some intermediate situations usually take place. The total number of these formal bonds of a given atom is equal to its valency. This helps a lot in selecting the chemical bond pattern, which afterwards may be checked against experiment (e.g., bond distances).¹⁴¹ In most cases a single bond is of the σ type, a double one is composed of one σ and one π , a triple bond means one σ and two π bonds (cf. p. 403).
- The minimal model of a molecule may explain most of the chemical reactions, if besides the closed-shell configuration (double occupancy of the molecular orbitals, including HOMO) we consider excited configurations corresponding to electron transfer(s) from the HOMO to LUMO orbital (see Chapter 14).
- The bonds behave very much like *springs of a certain strength and length*,¹⁴² and therefore, apart from the translational and rotational motion, the atoms vibrate about their equilibrium positions.¹⁴³ As to the structural problems (not chemical reactions), these vibrations may be treated as harmonic.
- For the 3D shape of our model molecule, most chemical structures can be correctly predicted using the Hartree–Fock model. The main features of the 3D structure can be also predicted (without any calculation) by using the concept of the *minimum repulsion energy of the electrons pairs*. Within the molecular orbital model, such repulsion is given by eq. (8.96).

¹⁴¹For some molecules this procedure is not unique, i.e. several chemical bond patterns may be conceived (“resonance structures”, cf. the valence bond method in Chapter 10). In such cases the real electronic structure corresponds to an averaging of all of them.

¹⁴²Both depend first of all of the elements making the bond, also a single bond is the weakest and longest, the triple is the strongest and shortest.

¹⁴³The model of molecule visualized in virtually all popular computer programs shows spherical atoms and chemical bonds as shining rods connecting them. First of all, atoms are not spherical, as is revealed by Bader analysis (p. 573) or atomic multipole representations (Appendix S). Second, a chemical bond resembles more a “rope” (higher values) of electronic density than a cylindrical rod. The “rope” is not quite straight and is slimiest at a critical point (see p. 575). Moreover, the rope, when cut perpendicularly, has a circular cross section for pure σ bonds, and an oval cross section for the double bond σ and π (cf. Fig. 11.1).

8.10.1 VALENCE SHELL ELECTRON PAIR REPULSION (VSEPR)

The underlying assumptions of VSEPR¹⁴⁴ are as follows:

- Atoms in a molecule are bound by chemical bonds residing in a space between the bounded atoms. The pattern of such bonds has to be assumed. Each chemical bond represents an electron bonding pair (see the present chapter).
- Some atoms may possess electron pairs that do not participate in a chemical bond pattern (inner shells, lone pairs, see the present chapter).
- The bonding pairs as well as the lone pairs around any atom of the molecule adopt positions in space (on a sphere) such as to minimize pair–pair Coulombic interactions, i.e. they try to be as far away as possible, cf. (8.96).
- The lone pairs repel more than the bonding pairs, and the repulsion bond pair – lone pair is in-between.
- Multiple bonds occupy more space than single bonds.

The total electronic energy in the Restricted Hartree–Fock model is given by eq. (8.36). It is worth stressing, that at a *fixed* geometry of the molecule, the minimization of the electron pair repulsion (by redefinition of the orbitals through a unitary transformation) given by eq. (8.94) does not lead to any change of the total electronic repulsion energy (including self-interaction), which stays invariant. However, when considering variations of geometry (which is at the heart of VSEPR) it is plausible, that smaller electron repulsion (i.e. a smaller const in (8.94)) represents a factor that stabilizes the structure. For small changes of geometry, self-interaction, i.e. $2\sum_i^{\text{MO}} h_{ii} + \sum_i^{\text{MO}} J_{ii}$ is not supposed to change very much in eq. (8.36), because each term is connected to a particular *localized* orbital, which is not expected to change much when changing the *interbond* angles. What should change most in (8.36), are the interactions of *different* localized orbitals, because their distances are affected. These interactions are composed of the Coulombic and exchange contributions. The exchange contribution of two different localized orbitals is small, because the orbitals overlap only by their “tails”. Hence, *minimization of the interpair Coulombic interactions of eq. (8.96) as a function of the geometry of the molecule can be viewed as a rationalization for VSEPR*. Note also, that in each h_{ii} there is an attraction of the electrons occupying the localized orbital i with all the nuclei. This term is responsible for the VSEPR rule that lone pairs repel more strongly than bond pairs.¹⁴⁵

In the VSEPR method the resulting structure depends on the calculated number of electron pairs around the central atom of the molecule.¹⁴⁶ The resulting geometry is given by Table 8.6

¹⁴⁴R.J. Gillespie, R.S. Nyholm, *Quart. Rev. Chem. Soc.* 11 (1957) 339.

¹⁴⁵The Coulombic interaction of electron pairs is damped by those nuclei, which are immersed in the electron cloud.

¹⁴⁶If several atoms may be treated as central, it is necessary to perform the VSEPR procedure for every such atom.

Table 8.6.

Number of electron pairs	Geometry
2	linear
3	trigonal planar
4	tetrahedral
5	trigonal bipyramidal
6	octahedral

Example 4. Water molecule. First, some guesses before using the minimal model. The hydrogen atom has a single electron and, therefore valency one, the oxygen atom has valency two (two holes in the valence shell). *We expect*, therefore, that the compounds of the two elements will have the following chemical bond patterns (that saturate their valencies): H-O-H , H-O-O-H , etc. Now our minimal model comes into play. Even quite simple Hartree–Fock calculations show that the system H-O-O-H is less stable than $\text{H-O-H} + \text{O}$. Thus, the minimal model predicts, in accordance with what we see in the oceans, that the H_2O compound called water is the most stable.

Now, what can we say about the 3D structure of the water molecule?

Let us take the VSEPR as a first indication. The central atom is oxygen, the number of its valence electrons is six. To this number, we add the number of electrons brought by two hydrogens: $6 + 2 = 8$. Therefore, the number of the electron pairs is $\frac{8}{2} = 4$. According to the above table, oxygen has a tetrahedral arrangement (the angle $109^\circ 28'$) of its four electron pairs. Two of them are lone pairs, two are bonding pairs with the hydrogens. Since, as the VSEPR model says, the lone pairs repel more strongly than the bonding pairs, we expect the angle between the lone pairs to be larger than $109^\circ 28'$, and the HOH bond angle to be *smaller than* $109^\circ 28'$.

Let us see what the minimal model is able to tell us about the geometry of the water molecule. The model (STO 6-31G** basis set, geometry optimization) predicts correctly that there are two *equivalent* OH chemical bonds (and there is no H–H bond¹⁴⁷) of length $R_{\text{OH}} = 0.943 \text{ \AA}$, whereas experiment gives the result $R_{\text{OH}} = 0.957 \text{ \AA}$. The model predicts, also in accordance with experiment, that the molecule is non-linear (!): the minimum energy HOH angle is 106.0° (the Hartree–Fock limit corresponds to 105.3°), while the experimental HOH angle is 104.5° . *The minimal model is usually able to predict the bond lengths within an accuracy of about 0.01 \AA , and bond angles to an accuracy of about 1° .*

The minimal model (within the STO 6-31G** basis set) predicts three harmonic vibrational frequencies of the water molecule: antisymmetric stretching 4264 cm^{-1} , symmetric stretching 4147 cm^{-1} and bending 1770 cm^{-1} . It is not easy, though, to predict the corresponding experimental frequencies. We measure the energy differences between consecutive vibrational levels (see Chapter 6, p. 235), which are

¹⁴⁷In agreement with common knowledge in chemistry.

not equal each other (due to anharmonicity). We may, however, deduce these experimental values as they would have been if the bottom of the well were perfectly quadratic (harmonic approximation), they are the following: 3942, 3832, 1648 cm^{-1} , respectively. Similarly to this case, *the minimal model systematically predicts vibrational frequencies that are 7–8% larger than experimental values*. This is not too bad by itself. In practical applications we often take this systematic error into account and correct the calculated frequencies by a scaling factor, thus predicting the frequencies to good accuracy.

Example 5. Chlorine trifluoride ClF_3 . It is not easy to tell what kind of structure we will have. Well, it is easy with VSEPR. The central atom will be chlorine. It has 7 valence electrons. Each fluorine contributes one electron. Thus, altogether the chlorine has $7 + 3 = 10$ electrons, i.e. five electron pairs. This means a trigonal bipyramid in VSEPR. However, this does not tell us where the lone pairs and where the fluorine atoms are. Indeed, there are two physically distinct positions in such a bipyramid: the axial and the equatorial, Fig. 8.30.

This corresponds to the interactions of the (lone or bond) electron pairs forming 90° , 120° and 180° . There are $5 \cdot 4/2 = 10$ such interactions. There are three isomers (a,b,c) possible that differ in interaction energy (L-L or lone pair – lone pair, b-L or bond pair – lone pair, b-b or bond–bond), Fig. 8.30.

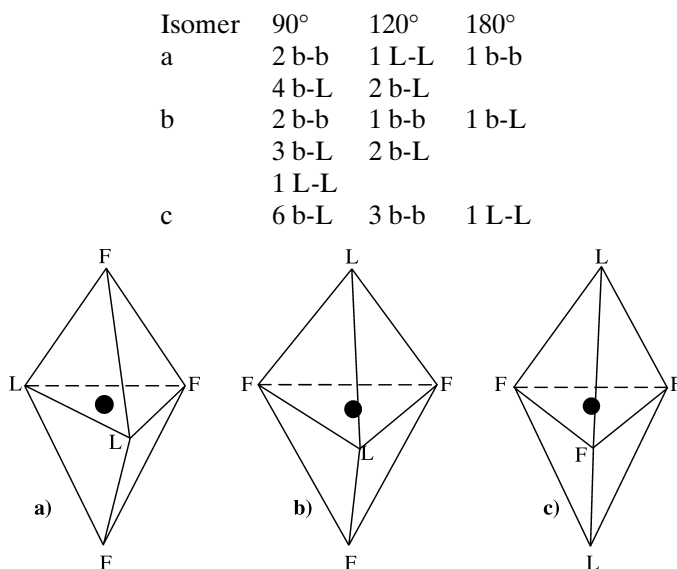


Fig. 8.30. The trigonal bipyramid has two physically distinct positions: three equatorial and two axial. In the ClF_3 we have two lone pairs (L) and three F atoms as candidates for these positions. There are three isomers that differ in energy: (a) having the two lone pairs in equatorial positions – this gives a planar T-shaped molecule (b) having one lone pair equatorial and one axial – this gives a non-planar molecule with two F-Cl-F angles equal to 90° , and one F-Cl-F angle equal to 120° (c) having two lone pairs axial – this gives a planar molecule with F-Cl-F angles equal to 120° . All the isomers have 6 interactions of electron pairs (lone or bond) at 90° , 3 interactions at 120° and one interaction at 180° .

Definitely, the 90° interaction of electron pairs is the most important, because of the shortest L-L distance. In the first approximation, let us look at the 90° interactions only. If we subtract from the energy of each isomer the same number: 3 b-L, then it remains the following

Isomer	90°
a	2 b-b 1 b-L
b	2 b-b 1 L-L
c	3 b-L

According to VSEPR, the L-L repulsion is the strongest, then the b-L follows and the weakest is the b-b repulsion. Now it is clear that the isomer a is of the lowest energy. Therefore, we predict a planar T-like structure with the $F_{\text{axial-Cl-F}_{\text{equatorial}}}$ angle equal to 90°. Since the lone pairs take more volume than the bond pairs, the T-shape is a little squeezed. Experiment indeed gives a weird-looking, planar T-shaped molecule, with the $F_{\text{axial-Cl-F}_{\text{equatorial}}}$ angle equal to 87.5°.

Summary

- The Hartree–Fock procedure is a *variational* method. The variational function takes the form of a *single Slater determinant* ψ built of orthonormal molecular spinorbitals:

$$\psi = \frac{1}{\sqrt{N!}} \begin{vmatrix} \phi_1(1) & \phi_1(2) & \dots & \phi_1(N) \\ \phi_2(1) & \phi_2(2) & \dots & \phi_2(N) \\ \dots & \dots & \dots & \dots \\ \phi_N(1) & \phi_N(2) & \dots & \phi_N(N) \end{vmatrix}$$

- A molecular spinorbital $\phi_i(1)$ is a one-electron function of the coordinates of electron 1, i.e. of x_1, y_1, z_1, σ_1 . In the RHF method, it is the product $\varphi_i(x_1, y_1, z_1)\alpha(\sigma_1)$ or $\varphi_i(x_1, y_1, z_1)\beta(\sigma_1)$ of a real *molecular orbital* $\varphi_i(x_1, y_1, z_1)$ and of the *spin function* $\alpha(\sigma_1)$ or $\beta(\sigma_1)$, respectively. In the general HF method (GHF), a spinorbital is a complex function, which depends both on $\alpha(\sigma_1)$ and $\beta(\sigma_1)$. The UHF method uses, instead, real orbitals, which are all different and are multiplied either by α or β (“different orbitals for different spins”).
- Minimization of the mean value of the Hamiltonian, $E = \frac{\langle \psi | \hat{H} | \psi \rangle}{\langle \psi | \psi \rangle}$, with respect to the orthonormal spinorbitals ϕ_i (GHF) leads to equations for *optimum spinorbitals* (Fock equations): $\hat{F}(1)\phi_i(1) = \varepsilon_i\phi_i(1)$, where the Fock operator \hat{F} is $\hat{F}(1) = \hat{h}(1) + \hat{J}(1) - \hat{K}(1)$, the Coulombic operator is defined by

$$\hat{J}(1)u(1) = \sum_j \hat{J}_j(1)u(1) \quad \text{and} \quad \hat{J}_j(1)u(1) = \int d\tau_2 \frac{1}{r_{12}} \phi_j^*(2)\phi_j(2)u(1),$$

and the exchange operator by

$$\hat{K}(1)u(1) = \sum_j \hat{K}_j(1)u(1) \quad \text{and} \quad \hat{K}_j(1)u(1) = \int d\tau_2 \frac{1}{r_{12}} \phi_j^*(2)u(2)\phi_j(1).$$

- In the Restricted Hartree–Fock method (RHF) for closed shell systems, we assume *double orbital occupancy*, i.e. we form two spinorbitals out of each molecular orbital (by multiplying either by α or β).
- The Fock equations are solved by an iterative approach (with *an arbitrary* starting point) and as a result we obtain approximations to:
 - the total energy,
 - the wave function (the optimum Slater determinant),
 - the canonical molecular orbitals (spinorbitals),
 - the orbital energies.
- Use of the LCAO expansion leads to the Hartree–Fock–Roothaan equations $Fc = Sc\epsilon$. Our job is then to find the LCAO coefficients c . This is achieved by transforming the matrix equation to the form of the eigenvalue problem, and to diagonalize the corresponding Hermitian matrix. The canonical molecular orbitals obtained are linear combinations of the atomic orbitals. The lowest-energy orbitals are occupied by electrons, those of higher energy are called virtual and are left empty.
- Using the H_2^+ example, we found that a chemical bond results from an electron density flow towards the bond region. This results from a superposition of atomic orbitals due to the variational principle.
- In the simplest MO picture:
 - The *excited triplet state has lower energy than the corresponding excited singlet state*.
 - In case of orbital degeneracy, the system prefers *parallel electron spins* (Hund's rule).
 - The ionization energy is equal to the negative of the *orbital energy of the removed electron*. The electron affinity is equal to the negative of the *orbital energy corresponding to the virtual orbital accommodating the added electron* (Koopmans theorem).
- The canonical MOs for closed-shell systems (the RHF method) may – completely legally – be transformed to orbitals localized in the chemical bonds, lone pairs and inner shells.
- There are many methods of localization. The most important ones are: the projection method, the method of minimum distance between two electrons from the same orbital (Boys approach), and the method of maximum interaction of electrons from the same orbital (Ruedenberg approach).
- Different localization methods lead to sets of localized molecular orbitals which are slightly different but their general shape is very similar.
- The molecular orbitals (localized as well as canonical) can be classified as to the number of nodal surfaces going through the nuclei. A σ bond orbital has no nodal surface at all, a π bond orbital has a single nodal surface, a δ bond orbital has two such surfaces.
- The localization allows comparison of the molecular fragments of different molecules. It appears that the features of the MO localized on the AB bond relatively weakly depend on the molecule in which this bond is found. This is a strong argument and a true source of experimental tactics in chemistry.
- Localization may serve to determine hybrids.
- In everyday practice, chemists use a minimal model of molecules that enables them to compare the geometry and vibrational frequencies with experiment. This model assumes that the speed of light is infinite (non-relativistic effects only), the Born–Oppenheimer approximation is valid (i.e. the molecule has a 3D structure), the nuclei are bound by chemical bonds and vibrate (often harmonic vibrations are assumed), the molecule moves (translation) and rotates as a whole in space.
- In many cases we can successfully predict the 3D structure of a molecule by using a very simple tool: the Valence Shell Electron Pair Repulsion concept.

Main concepts, new terms

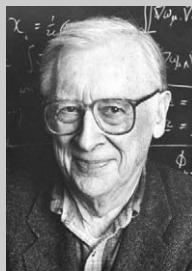
- molecular spinorbital (p. 330)
- Slater determinant (p. 332)
- energy functional (p. 335)
- conditional extremum (p. 336)
- Lagrange multipliers (p. 336 and p. 997)
- variation of spinorbital (p. 336)
- Coulombic operator (p. 337)
- exchange operator (p. 337)
- invariance with respect to a unitary transformation (p. 340)
- General Hartree–Fock method (GHF) (p. 341)
- Unrestricted Hartree–Fock method (UHF) (p. 342)
- Restricted Hartree–Fock method (RHF) (p. 342)
- molecular orbital (p. 342)
- occupied orbital (p. 343)
- virtual orbital (p. 343)
- HOMO (p. 343)
- LUMO (p. 343)
- closed shell (p. 344)
- mean field (p. 348)
- orbital centring (p. 354)
- Slater-type orbital (p. 355)
- Slater orbital (p. 356)
- Gaussian-type orbital (p. 357)
- atomic orbital size (p. 357)
- LCAO (p. 360)
- atomic basis set (p. 363)
- Hartree–Fock–Roothaan method (p. 364)
- bonding orbital (p. 371)
- antibonding orbital (p. 371)
- instability (p. 372)
- Fukutome classes (p. 372)
- Mendeleev Periodic Table (p. 379)
- electronic shells (p. 381)
- electronic configuration (p. 381)
- chemical bond (p. 383)
- penetration energy (p. 386)
- Jabłoński diagram (p. 391)
- Hund’s rule (p. 392)
- Koopmans theorem (p. 393)
- orbital localization (p. 396)
- σ , π , δ – molecular orbitals (p. 403)
- electronic pair dimension (p. 404)
- hybrids (p. 408)
- tetrahedral hybridization (p. 408)
- trigonal hybridization (p. 409)
- digonal hybridization (p. 409)
- minimal model of a molecule (p. 417)
- Valence Shell Electron Pair Repulsion (VSEPR) (p. 419)

From the research front

The Hartree–Fock method belongs to a narrow 2–3-member class of standard methods of quantum chemistry. It is the source of basic information about the electronic ground state of a molecule. It also allows for geometry optimization. At present, the available computational codes limit the calculations to the systems built of several hundreds of atoms. Moreover, the programs allow calculations to be made by clicking the mouse. The Hartree–Fock method is always at their core. The GAUSSIAN is one of the best known programs. It is the result of many years of coding by several tens of quantum chemists working under John Pople. Pople was given Nobel Prize in 1998 mainly for this achievement. To get a flavour of the kind of data needed, I provide below a typical data set necessary for GAUSSIAN to perform the Hartree–Fock computations for the water molecule:

```
#HF/STO-3G opt freq pop
water, the STO-3G basis set
0 1
O
H1 1 r12
H2 1 r12 2 a213
r12=0.96
a213=104.5
```

John Pople (1925–2004), British mathematician and one of the founders of the modern quantum chemistry. His childhood was spent in difficult war time in England (every day 25 mile train journeys, sometimes under bombing). He came from a lower middle class family (drapers and farmers), but his parents were ambitious for the future of their children. At the age of twelve John developed an intense interest in mathematics. He entered Cambridge University after receiving a special scholarship. John Pople made important contributions to theoretical chemistry. To cite a few: proposing semi-empirical methods – the famous PPP method for π electron systems, the once very popular CNDO approach for all-valence calculations, and finally the monumental joint work on



GAUSSIAN – a system of programs that constitutes one of most important computational tools for quantum chemists. John Pople received the Nobel prize in 1998 “for his development of computational methods in quantum chemistry” sharing it with Walter Kohn.

The explanatory comments, line by line:

- `#HF/STO-3G opt freq pop` is a command which informs GAUSSIAN that the computations are of the Hartree–Fock type (HF), that the basis set used is of the STO-3G type (each STO is expanded into three GTOs), that we want to optimize geometry (opt), compute the harmonic vibrational frequencies (freq) and perform the charge population analysis for the atoms (known as Mulliken population analysis, see Appendix S, p. 1015);
- just a comment line;
- `0 1` means that the total charge of the system is equal to 0, and the singlet state is to be computed (1);
- `O` means that the first atom in the list is oxygen;
- `H1 1 r12` means that the second atom in the list is hydrogen named H1, distant from the first atom by r12;
- `H2 1 r12 2 a213` means that the third atom in the list is hydrogen named H2, distant from atom number 1 by r12, and forming the 2-1-3 angle equal to a213;
- `r12=0.96` is a starting OH bond length in Å;
- `a213=104.5` is a starting angle in degrees.

Similar inputs are needed for other molecules. The initial geometry is to some extent arbitrary, and therefore in fact it cannot be considered as real input data. The only true information is the number and charge (kind) of the nuclei, the total molecular charge (i.e. we know how many electrons are in the system), and the multiplicity of the electronic state to be computed. The basis set issue (STO-3G) is purely technical, and gives information about the quality of the results.

Ad futurum...

Along with the development of computational technique, and with progress in the domain of electronic correlation, the importance of the Hartree–Fock method as a source of information about total energy, or total electron density, will most probably decrease. Simply, much larger molecules (beyond the HF level) will be within the reach of future comput-

ers. Yet Hartree–Fock calculations will be still carried out, and their results will be carefully analyzed. There are at least two reasons for it:

- Hartree–Fock calculations are most often the necessary step before more precise computations are performed.
- Hartree–Fock computations result in the molecular orbital model: the molecular orbitals and the orbital energies scheme (“minimal model”), and thus they provide the *conceptual framework for the molecule*. It is the sort of model, which may be discussed, thought of, and used to search for explanation of physical and chemical phenomena. So far such a possibility does not exist for advanced methods, where often we obtain very good results, but it is extremely difficult to get an idea *why* they agree so well with experiments.¹⁴⁸

Additional literature

A. Szabo, N.S. Ostlund, “Modern Quantum Chemistry”, McGraw-Hill, New York, 1989, p. 108–231.

Excellent book.

T. Helgaker, P. Jørgensen, J. Olsen, “Molecular Electronic-Structure Theory”, Wiley, Chichester, 2000, p. 433–513.

Very good book. It is a contemporary compendium of computational quantum chemistry.

“Localization and Delocalization in Quantum Chemistry”, Ed. by O. Chalvet, R. Daudel, S. Diner, J.-P. Malrieu, D. Reidel Publish. Co., Dordrecht, 1975.

A set of the very interesting articles by the leading quantum chemists.

Questions

1. The HF method for the N electron system leads to the wave function:
 - a) which depends on all coordinates of N electrons and satisfies the Schrödinger equation;
 - b) in a form of one Slater determinant, that depends on all coordinates of N electrons and which is an eigenfunction of the Fock operator $\hat{F}(1)$;
 - c) in a form of one Slater determinant, which does not satisfy the Schrödinger equation;
 - d) in a form of the product of molecular spinorbitals.
2. The canonical orbitals of a molecule:
 - a) represent the minimum basis set of the atoms of a molecule;
 - b) do not satisfy the Fock equation and give the lowest-energy Slater determinant;
 - c) satisfy the Fock equation and give the lowest-energy Slater determinant;
 - d) are the localized orbitals for the chemical bonds of the molecule.
3. The localized orbitals:
 - a) are the orthonormal orbitals localized on an atom and satisfying the Fock equation for this atom;
 - b) do not satisfy the Fock equation (8.30) and give the lowest-energy Slater determinant;
 - c) satisfy the Fock equation (8.30) and give the lowest-energy Slater determinant;
 - d) are the atomic orbitals which satisfy the Fock equations for the atoms.

¹⁴⁸The fact of solving the Schrödinger equation, unfortunately, in most cases does not instruct us on the nature of physical phenomena.

4. The orbital energy in a molecule:
 - a) is the energy of the electron that occupies the highest-energy atomic orbital;
 - b) is the mean energy per one electron in the molecule;
 - c) is the sum of the energies of two electrons described by a given molecular orbital;
 - d) is the mean value of the Fock operator computed with the corresponding canonical orbital.
5. The Fock operator contains (among other operators) the operator:
 - a) of the kinetic energy of all electrons;
 - b) of the Coulombic repulsion of all the electrons;
 - c) of the kinetic energy of the electrons, and the two-electron Coulombic operator;
 - d) of the kinetic energy of an electron and of its electrostatic repulsion with the density distribution of all the electrons.
6. In the LCAO MO method, each MO is:
 - a) a function of the electron position in 3D space;
 - b) a linear combination of the hybrids generated from the valence atomic orbitals;
 - c) a product of the AOs for the atoms of a molecule;
 - d) a linear combination of the orbitals of electron lone pairs.
7. In the Hartree–Fock method (closed shell, U = sum of the orbital energies for the doubly occupied orbitals), the total electronic energy of a system is:
 - a) $2U - V_{ee}$; b) $2U - 2V_{ee}$; c) $U - V_{ee}$; d) $2U + 2V_{ee}$.
8. Localization of the molecular orbitals (MOs) is performed to:
 - a) divide the MOs into bonding and antibonding orbitals;
 - b) modify the total electron density, to make it closer to experimental results;
 - c) decrease the total energy of a system;
 - d) get the MOs of the bonds, electron lone pairs and inner shells.
9. Find the *false* statement:
 - a) the UHF method always gives some spin contamination;
 - b) $E_{\text{GHF}} \leq E_{\text{RHF}}$; c) $E_{\text{GHF}} \leq E_{\text{UHF}}$; d) $E_{\text{UHF}} \leq E_{\text{RHF}}$.
10. The MO method for the hydrogen molecule (R is the internuclear distance):
 - a) cannot be applied for large R values;
 - b) properly describes the dissociation of the molecule;
 - c) shows large differences for HOMO and LUMO for large R values;
 - d) takes the electronic correlation into account.

Answers

1c, 2c, 3b, 4d, 5d, 6a, 7a, 8d, 9a, 10a