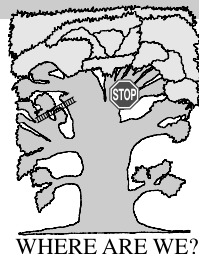


## Chapter 10

# CORRELATION OF THE ELECTRONIC MOTIONS



### Where are we?

The main road on the trunk lead us to the middle of the crown of the tree.

### An example

As usual let us consider the simplest example: the hydrogen molecule. The normalized Restricted Hartree–Fock determinant

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

with double occupancy of the normalized molecular orbital  $\varphi$  ( $\phi_1 = \varphi\alpha$ ,  $\phi_2 = \varphi\beta$ ), after expansion, gives immediately:

$$\psi_{\text{RHF}}(1, 2) = \varphi(1)\varphi(2) \frac{1}{\sqrt{2}} \{ \alpha(1)\beta(2) - \beta(1)\alpha(2) \}.$$

The key quantity here is the square of the complex modulus of  $\psi_{\text{RHF}}(1, 2)$ , since the  $|\psi_{\text{RHF}}(1, 2)|^2$  tells us about the probability density of the occurrence of certain coordinates of the electrons. We will study the fundamental problem for the motion of electrons: whether the electrons react to their presence.

Let us ask some very important questions. Firstly what is the (conditional) probability density of the occurrence of the situation when

- electron 1 occupies *different positions* in space *on the contour line*  $\varphi = \text{const}$  and has spin coordinate  $\sigma_1 = 1/2$  while electron 2 has spin coordinate  $\sigma_2 = -1/2$ , and its space coordinates are  $x_2, y_2, z_2$ .

We calculate

$$\begin{aligned} |\psi_{\text{RHF}}(1, 2)|^2 &= \left[ \varphi(1)\varphi(2) \frac{1}{\sqrt{2}} \{ \alpha(\sigma_1)\beta(\sigma_2) - \beta(\sigma_1)\alpha(\sigma_2) \} \right]^2 \\ &= \left[ \text{const} \times \varphi(x_2, y_2, z_2) \frac{1}{\sqrt{2}} \{ \alpha(1/2)\beta(-1/2) - \beta(1/2)\alpha(-1/2) \} \right]^2 \\ &= \left[ \text{const} \times \varphi(x_2, y_2, z_2) \frac{1}{\sqrt{2}} \{ 1 \times 1 - 0 \times 0 \} \right]^2 \\ &= \frac{1}{2} (\text{const})^2 \times \varphi^2(x_2, y_2, z_2). \end{aligned}$$

Electron 1 changes its position on the contour line, but the distribution of the probability density of electron 2 (of the opposite spin) does not change a bit, while electron 2 should move away from its partner, since the electrons repel each other. Electron 2 is not afraid to approach electron 1. The latter can even touch electron 2 and it does not react at all. For such a deficiency we have to pay through the high mean value of the Hamiltonian (since there is a high average energy of the electron repulsion). The Hartree–Fock method therefore has an obvious shortcoming.

- We now ask about the probability density of finding a situation in which we leave everything the same as before, but now electron 2 has spin coordinate  $\sigma_2 = 1/2$  (so this is the situation where both electrons have identical projections of spin angular momentum<sup>1</sup>). What will the response to this change be of  $|\psi_{\text{RHF}}(1, 2)|^2$  as a function of the position of electron 2?
- Again we calculate

$$\begin{aligned} |\psi_{\text{RHF}}(1, 2)|^2 &= \left[ \varphi(1)\varphi(2) \frac{1}{\sqrt{2}} \{ \alpha(\sigma_1)\beta(\sigma_2) - \beta(\sigma_1)\alpha(\sigma_2) \} \right]^2 \\ &= \left[ \text{const} \times \varphi(x_2, y_2, z_2) \frac{1}{\sqrt{2}} \left\{ \alpha\left(\frac{1}{2}\right)\beta\left(\frac{1}{2}\right) - \beta\left(\frac{1}{2}\right)\alpha\left(\frac{1}{2}\right) \right\} \right]^2 \\ &= \left[ \text{const} \times \varphi(x_2, y_2, z_2) \frac{1}{\sqrt{2}} \{ 1 \times 0 - 0 \times 1 \} \right]^2 = 0. \end{aligned}$$

We ask about the distribution of the electron of the same spin. The answer is that this distribution is *everywhere equal to zero*, i.e. we do not find electron 2 with spin coordinate  $\frac{1}{2}$  independently of the position of the electron 1 with spin coordinate  $\frac{1}{2}$  (in whatever point on the contour line or beyond it).

The second conclusion can be accepted, since it follows from the pairing of the spins,<sup>2</sup> but the first conclusion is just absurd. Such nonsense is admitted by the Hartree–Fock method. In this chapter we will ponder how we can introduce a correlation of electronic motions.

The correlation of the motions of the electrons results in an energy gain called a *correlation energy* which is defined as

correlation  
energy

$$E_{\text{corel}} = E - E_{\text{RHF}},$$

where  $E$  is the energy from the Schrödinger equation,<sup>3</sup> and  $E_{\text{RHF}}$  is the Restricted Hartree–Fock energy.<sup>4</sup>

<sup>1</sup>We may ask: “How come?” After all, we consider a singlet state, hence the spin projections are opposite. We will not find the situation with parallel spin projections. Take it easy. If, in fact, we are right then we will get 0 as the density of the respective conditional probability. Let us see whether it really will be so.

<sup>2</sup>And this is ensured by the singlet form of the spin part of the function.

<sup>3</sup>This is the rigorous nonrelativistic energy of the system in its ground state. This quantity is not available experimentally, we can *evaluate* it by subtraction of the calculated relativistic corrections from the energy of the total ionization of the system.

<sup>4</sup>Usually we define the correlation energy for the case of double occupancy of the molecular orbitals (the RHF method, see p. 330). In the case of open shells, especially when the multideterminantal description is required, the notion of correlation energy still remains to be defined. These problems will not be discussed in this book.

## What is it all about

The outline of the chapter is as follows:

- First we will discuss the methods which explicitly (*via* the form of the suggested wave function) allow the electrons to control their mutual distance (“a correlation of motions”).
- In the second part of the chapter the correlation will be less visible, since it will be accounted for by the application of linear combinations of the Slater determinants. First we will discuss the variational methods (VB, CI, MC SCF), and then the non-variational ones (CC, EOM-CC, MBPT).

## VARIATIONAL METHODS USING EXPLICITLY CORRELATED WAVE FUNCTION

**Correlation cusp condition**(♦) p. 502

**The Hylleraas function**(♦) p. 503

**The Hylleraas CI method**(♦) p. 506

**The harmonic helium atom**(♦) p. 507

**James–Coolidge and Kołos–Wolniewicz functions** (♦) p. 508

- Neutrino mass

**Method of exponentially correlated Gaussian functions** (♦) p. 513

**Coulomb hole (“correlation hole”)** (♦) p. 513

**Exchange hole (“Fermi hole”)** (♦) p. 516

## VARIATIONAL METHODS WITH SLATER DETERMINANTS

**Valence bond (VB) method** (Δ) p. 520

- Resonance theory – hydrogen molecule
- Resonance theory – polyatomic case

**Configuration interaction (CI) method** (Δ) p. 525

- Brillouin theorem
- Convergence of the CI expansion
- Example of H<sub>2</sub>O
- Which excitations are most important?
- Natural orbitals (NO)
- Size consistency

**Direct CI method** (♦) p. 533

**Multireference CI method** (♦) p. 533

**Multiconfigurational Self-Consistent Field method (MC SCF)** (Δ♦ $\hat{S}\hat{U}$ ) p. 535

- Classical MC SCF approach (Δ)
- Unitary MC SCF method (♦)
- Complete active space method (CAS SCF) (♦ $\hat{S}\hat{U}$ )

## NON-VARIATIONAL METHODS WITH SLATER DETERMINANTS

**Coupled cluster (CC) method** (♦) p. 539

- Wave and cluster operators
- Relationship between CI and CC methods
- Solution of the CC equations
- Example: CC with double excitations
- Size-consistency of the CC method

**Equation-of-motion method (EOM-CC)** (♦) p. 548

- Similarity transformation

- Derivation of the EOM-CC equations

### Many body perturbation theory (MBPT) (♦)

p. 551

- Unperturbed Hamiltonian
- Perturbation theory – slightly different approach
- Reduced resolvent or the “almost” inverse of  $(E_0^{(0)} - \hat{H}^{(0)})$
- MBPT machinery
- Brillouin–Wigner perturbation theory
- Rayleigh–Schrödinger perturbation theory

### Møller–Plesset version of Rayleigh–Schrödinger perturbation theory (Δ)

p. 558

- Expression for MP2 energy
- Convergence of the Møller–Plesset perturbational series
- Special status of double excitations

In the previous chapter we dealt with the description of electronic motion in the mean field approximation. Now we use this approximation as a starting point towards methods accounting for electron correlation. Each of the methods considered in this chapter, when rigorously applied, should give an exact solution of the Schrödinger equation. Thus this chapter will give us access to methods providing accurate solutions of the Schrödinger equation.

### Why is this important?

Perhaps, in our theories, the electrons do not need to correlate their motion and the results will be still all right?

Unfortunately, this is not so. The mean field method provides, to be sure, ca. 99% of the total energy of the system. This is certainly a lot, in many cases the mean field method gives quite good results, but still falls short of treating several crucial problems correctly. For example:

- *Only through electron correlation* do the noble gas atoms attract each other in accordance with experiment (liquefaction of gases).
- According to the Hartree–Fock method, the  $F_2$  molecule *does not exist* at all, whereas the fact is that it exists, and is doing quite well (bonding energy equal to 38 kcal/mol).<sup>5</sup>
- About *half* the interaction energy of large molecules (often of biological importance) calculated at the equilibrium distance originates purely from the correlation effect.
- The RHF method used to describe the dissociation of the chemical bond gives simply tragic results (cf. Chapter 8, p. 371), *qualitatively wrong* (the UHF method gives a qualitatively correct description).

We see that in many cases electronic correlation must be taken into account.

### What is needed?

- Operator algebra (Appendix B, necessary).
- Hartree–Fock method (Chapter 8, necessary).
- Eigenvalue problem (Appendix L, p. 984, necessary).
- Variational method (Chapter 5, necessary).
- Perturbation theory (Chapter 5, recommended).
- Matrix diagonalization (Appendix K, p. 982, recommended).
- Second quantization (Appendix U, p. 1023, necessary).

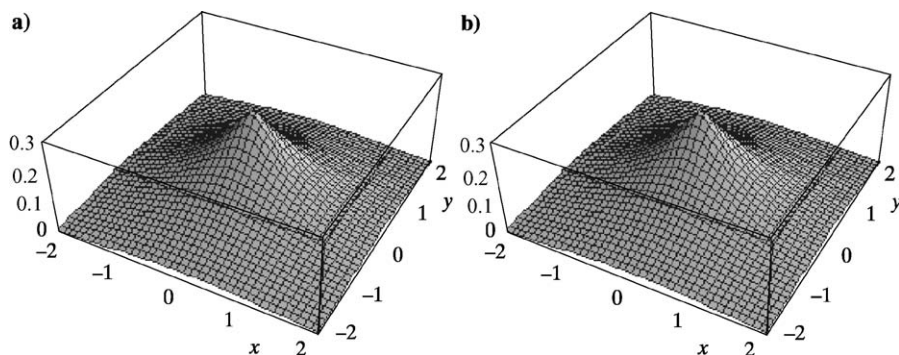
<sup>5</sup>Yet this is not a strong bond. For example, the bonding energy of the  $H_2$  molecule equals 104 kcal/mol, that of HF equals 135 kcal/mol.

### Classic papers

The first calculations incorporating electron correlation in an atom (helium) were published by Egil Andersen Hylleraas in an article “*Neue Berechnung der Energie des Heliums im Grundzustande, sowie des tiefsten Terms von Ortho-Helium*”, *Zeitschrift für Physik*, 54 (1929) 347. ★ The first calculations with electron correlation for molecules were performed by Walter Heitler and Fritz Wolfgang London in a paper “*Wechselwirkung neutraler Atome und homöopolare Bindung nach der Quantenmechanik*” published in *Zeitschrift für Physik*, 44 (1927) 455. Formation of the covalent bond (in  $H_2$ ) could be correctly described only after the electron correlation has been included. June 30, 1927, when Heitler and London submitted the paper, is the birth date of quantum chemistry. ★ Later, significantly more accurate results were obtained for the hydrogen molecule by Hubert M. James and Albert S. Coolidge in an article “*The Ground State of the Hydrogen Molecule*”, *Journal of the Chemical Physics*, 1 (1933) 825, and a contemporary reference point for that molecule are papers by Włodzimierz Kołos and Lutosław Wolniewicz, among others an article entitled “*Potential Energy Curves for the  $X^1\Sigma_g^+$ ,  $B^3\Sigma_u^+$ ,  $C^1\Pi_u$  States of the Hydrogen Molecule*” published in *Journal of Chemical Physics*, 43 (1965) 2429. ★ Christian Møller and Milton S. Plesset in *Physical Review*, 46 (1934) 618 published a paper “*Note on an Approximation Treatment for Many-Electron Systems*”, where they presented a perturbational approach to electron correlation. ★ The first calculations with the Multi-Configurational Self-Consistent Field (MC SCF) method for atoms was published by Douglas R. Hartree, William Hartree and Bertha Swirles in a paper “*Self-Consistent Field, Including Exchange and Superposition of Configurations, with some Results for Oxygen*”, *Philosophical Transactions of the Royal Society (London)*, A 238 (1939) 229, and the general MC SCF theory was presented by Roy McWeeny in a work “*On the Basis of Orbital Theories*”, *Proceedings of the Royal Society (London)*, A 232 (1955) 114. ★ As a classic paper in electronic correlation we also consider an article by Per-Olov Löwdin “*Correlation Problem in Many-Electron Quantum Mechanics*” in *Advances in Chemical Physics*, 2 (1959) 207. ★ The idea of the Coupled Cluster (CC) method was introduced by Fritz Coester in a paper in *Nuclear Physics*, 7 (1958) 421 entitled “*Bound States of a Many-Particle System*”. Jiří Čížek introduced the (diagrammatic) CC method into electron correlation theory in the paper “*On the Correlation Problem in Atomic and Molecular Systems. Calculation of Wavefunction Components in Ursell-type Expansion Using Quantum-Field Theoretical Methods*” published in the *Journal of Chemical Physics*, 45 (1966) 4256. ★ The book edited by Oktay Sinanoğlu and Keith A. Brueckner “*Three Approaches to Electron Correlation in Atoms*”, Yale Univ. Press, New Haven and London, 1970, contains several reprints of the papers which cleared the path towards the coupled-cluster method. ★ A derivation of the coupled cluster equations (for interacting nucleons) was presented by Herman Kümmel and Karl-Heinz Lührmann, *Nuclear Physics*, A191 (1972) 525 in a paper entitled “*Equations for Linked Clusters and the Energy Variational Principle*”.

## VARIATIONAL METHODS USING EXPLICITLY CORRELATED WAVE FUNCTION

We have learnt, from the example given at the beginning of this chapter, that the “genetic defect” of mean field methods is, that they describe electrons and com-



**Fig. 10.1.** Absence of electronic correlation in the helium atom as seen by the Hartree–Fock method. Visualization of the cross-section of the square of the wave function (probability density distribution) describing electron 2 within the plane  $xy$  provided electron 1 is located in a certain point in space: a) at  $(-1, 0, 0)$ ; b) at  $(1, 0, 0)$ . *Note, that in both cases the conditional probability density distributions of electron 2 are identical.* This means electron 2 does not react to the motion of electron 1, i.e. there is no correlation whatsoever of the electronic motions (when the total wave function is the Hartree–Fock one).

pletely ignore the fact that they are close or far away from each other. For example, in the two-electron case previously considered when we established the coordinates of electron 1, electron 2 has a certain distribution of the probability density. *This distribution does not change when the electron 1 moves to a different position.* This means that the electrons “are not afraid” to get close to each other, although they should, since when electrons are close the energy increases (Fig. 10.1.a,b).

*The explicitly correlated wave function (we will get to it in a moment) has the interelectronic distance built in its mathematical form.* We may compare this to making the electrons wear spectacles.<sup>6</sup> Now they avoid each other. One of my students said that it would be the best if the electrons moved apart to infinity. Well, they cannot. They are attracted by the nucleus (energy gain), and being close to it, are necessarily close to each other too (energy loss). There is a compromise to achieve.

## 10.1 CORRELATION CUSP CONDITION

Short distances are certainly most important for the Coulombic interaction of two charges, although obviously, the regions of configurational space connected with the long interelectronic distances are much larger. Thus the region is not large, but *very important*, within it “collisions” take place. It turns out that the wave function calculated *in the region of collision* must satisfy some very simple mathematical condition (called *correlation cusp condition*). This is what we want to demonstrate. The derived formulae<sup>7</sup> are universal, they apply to any pair of charged particles.

<sup>6</sup>Of course, the methods described further also provide their own “spectacles” (otherwise they would not give the solution of the Schrödinger equation), but the spectacles in the explicitly correlated functions are easier to construct with a small number of parameters.

<sup>7</sup>T. Kato, *Commun. Pure Appl. Math.* 10 (1957) 151.

Let us consider *two* particles with charges  $q_i$  and  $q_j$  and masses  $m_i$  and  $m_j$  *separated from other particles*. This, of course, makes sense since simultaneous collisions of three or more particles occur very rarely in comparison with two-particle collisions. Let us introduce a Cartesian system of coordinates (say, in the middle of the beautiful market square in Brussels), so that the system of two particles is described with six coordinates. Then (atomic units are used) the sum of the kinetic energy operators of the particles is

$$\hat{T} = -\frac{1}{2m_i}\Delta_i - \frac{1}{2m_j}\Delta_j. \quad (10.1)$$

Now we separate the motion of the centre of mass of the two particles with position vectors  $\mathbf{r}_i$  and  $\mathbf{r}_j$ . The centre of mass in our coordinate system is indicated by the vector  $\mathbf{R}_{\text{CM}} = (X_{\text{CM}}, Y_{\text{CM}}, Z_{\text{CM}})$

$$\mathbf{R}_{\text{CM}} = \frac{m_i\mathbf{r}_i + m_j\mathbf{r}_j}{m_i + m_j}. \quad (10.2)$$

Let us also introduce the total mass of the system  $M = m_i + m_j$ , the reduced mass of the two particles  $\mu = \frac{m_i m_j}{m_i + m_j}$  and the vector of their relative positions  $\mathbf{r} = \mathbf{r}_i - \mathbf{r}_j$ . Introducing the three coordinates of the centre of mass measured with respect to the market square in Brussels and the three coordinates  $x, y, z$  which are components of the vector  $\mathbf{r}$ , we get (Appendix I, Example 1)

$$\hat{T} = -\frac{1}{2M}\Delta_{\text{CM}} - \frac{1}{2\mu}\Delta, \quad (10.3)$$

$$\Delta_{\text{CM}} = \frac{\partial^2}{\partial X_{\text{CM}}^2} + \frac{\partial^2}{\partial Y_{\text{CM}}^2} + \frac{\partial^2}{\partial Z_{\text{CM}}^2}, \quad (10.4)$$

$$\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}. \quad (10.5)$$

After this operation, the Schrödinger equation for the system is separated (as always in the case of two particles, see Appendix I, p. 971) into two equations: the first describing the motion of the centre of mass (seen from Brussels) and the second describing the *relative* motion of the two particles (with Laplacian of  $x, y, z$  and reduced mass  $\mu$ ). We are not interested in the first equation, but the second one is what we are after. Let us write down the Hamiltonian corresponding to the second equation

$$\hat{H} = -\frac{1}{2\mu}\Delta + \frac{q_i q_j}{r}. \quad (10.6)$$

We are interested in how the wave function looks when the distance between the two particles  $r$  gets very small. If  $r$  is small, it makes sense to expand the wave function in a power series<sup>8</sup> of  $r$ :  $\psi = C_0 + C_1 r + C_2 r^2 + \dots$ . Let us calculate  $\hat{H}\psi$

<sup>8</sup>Assuming such a form we exclude the possibility that the wave function goes to  $\pm\infty$  for  $r \rightarrow 0$ . This must be so, since otherwise either the respective probability would go to infinity or the operators

in the vicinity of  $r = 0$ . The Laplacian expressed in the spherical coordinates is the sum of three terms (Appendix I, p. 971): the first, which contains the differentiation with respect to  $r$  and the remaining two, which contain the differentiation with respect to the angles  $\theta$  and  $\phi$ :  $\Delta = \frac{1}{r^2} \frac{\partial}{\partial r} r^2 \frac{\partial}{\partial r} + \text{terms depending on } \theta \text{ and } \phi$ . Since we have assumed the function to be dependent on  $r$ , upon the action of the Laplacian only the first term gives a nonzero contribution.

We obtain

$$\begin{aligned}\hat{H}\psi &= \left( -\frac{1}{2\mu} \Delta + \frac{q_i q_j}{r} \right) \psi \\ &= 0 - \frac{C_1}{2\mu} \frac{2}{r} + 6C_2 + 12C_3 r + \dots \\ &\quad + C_0 \frac{q_i q_j}{r} + C_1 q_i q_j + C_2 q_i q_j r + \dots\end{aligned}\quad (10.7)$$

The wave function cannot go to infinity when  $r$  tends to zero, while in the above expression we have two terms which would then “explode” to infinity.

These terms must cancel.

Hence, we obtain

$$C_0 q_i q_j = \frac{C_1}{\mu}. \quad (10.8)$$

This condition is usually expressed in a different way. We use the fact that  $\psi(r=0) = C_0$  and  $(\frac{\partial \psi}{\partial r})_{r=0} = C_1$ . We then obtain the cusp condition as

$$\left( \frac{\partial \psi}{\partial r} \right)_{r=0} = \mu q_i q_j \psi(r=0).$$

- The case of two electrons:

Then  $m_i = m_j = 1$ , hence  $\mu = \frac{1}{2}$  and  $q_i = q_j = -1$ . We get the cusp condition for the collision of two electrons as

$$\left( \frac{\partial \psi}{\partial r} \right)_{r=0} = \frac{1}{2} \psi(r=0)$$

or

the wave function should be of the form

$$\psi = \phi(\mathbf{r}_1, \mathbf{r}_2) \left[ 1 + \frac{1}{2} r_{12} + \dots \right],$$

where  $+\dots$  means higher powers of  $r_{12}$ .

would become non-Hermitian, cf. p. 73. Both possibilities are unacceptable. We covertly assumed also (to simplify our considerations) that the wave function does not depend on the angles  $\theta$  and  $\phi$ . This dependence can be accounted for by making the constants  $C_0, C_1, C_2$  the functions of  $\theta$  and  $\phi$ . Then the final results still holds, but for the coefficients  $C_0$  and  $C_1$  averaged over  $\theta$  and  $\phi$ .



- The nucleus–electron case:

When one of the particles is a nucleus of charge  $Z$  then  $\mu \simeq 1$  and we get

$$\left(\frac{\partial \psi}{\partial r}\right)_{r=0} = -Z\psi(r=0).$$

Thus

the correct wave function for the electron in the vicinity of a nucleus should have an expansion  $\psi = \text{const}(1 - Zr_{a1} + \dots)$ , where  $r_{a1}$  is the distance from the nucleus

Let us see how it is with the  $1s$  function for the hydrogen-like atom (the nucleus has charge  $Z$ ) expanded in a Taylor series in the neighbourhood of  $r = 0$ . We have  $1s = N \exp(-Zr) = N(1 - Zr + \dots)$ . It works.

The correlation cusp condition shows that the wave function is not differentiable at  $r = 0$ .

## 10.2 THE HYLLERAAS FUNCTION

In 1929, two years after the birth of quantum chemistry, a paper by Hylleraas<sup>9</sup> appeared, where, for the ground state of the helium atom, a trial variational function, containing the interelectronic distance explicitly, was applied. This was a brilliant idea, since it showed that already a small number of terms provide very good results. Even though no fundamental difficulties were encountered for larger atoms, the enormous numerical problems were prohibitive for atoms with larger numbers of electrons. In this case, the progress made from the nineteen twenties to the end of the twentieth century is exemplified by transition from two- to ten-electron systems.

## 10.3 THE HYLLERAAS CI METHOD

In this method,<sup>10</sup> we exploit the Hylleraas idea in such a way that the electronic wave function is expressed as a linear combinations of Slater determinants, and in front of each determinant  $\Phi_i$  ( $1, 2, 3, \dots, N$ ) we insert, next to the variational coefficient  $c_i$ , correlational factors with some powers ( $v, u, \dots$ ) of the interelectronic

<sup>9</sup>E.A. Hylleraas, *Zeit. Phys.* 54 (1929) 347. Egil Andersen Hylleraas arrived in 1926 in Göttingen, to collaborate with Max Born. His professional experience was related to crystallography and to the optical properties of quartz. When one of the employees fell ill, Born told Hylleraas to continue his work on the helium atom in the context of the newly developed quantum mechanics. The helium atom problem had already been attacked by Albrecht Unsöld in 1927 using first order perturbation theory, but Unsöld obtained the ionization potential equal to 20.41 eV, while the experimental value was equal to 24.59 eV. In the reported calculations (done on a recently installed calculator) Hylleraas obtained a value of 24.47 eV (cf. contemporary accuracy, p. 134).

<sup>10</sup>CI, *Configuration Interaction*.

distances ( $r_{mn}$  between electron  $m$  and electron  $n$ , etc.):

$$\psi = \sum_i c_i \hat{A}[r_{mn}^{v_i} r_{kl}^{u_i} \dots \Phi_i(1, 2, 3, \dots, N)], \quad (10.9)$$

where  $\hat{A}$  denotes an antisymmetrization operator (see Appendix U, p. 1023). If  $v_i = u_i = 0$ , we have the CI expansion:  $\psi = \sum_i c_i \Phi_i$  (we will discuss it on p. 525). If  $v_i \neq 0$ , we include a variationally proper treatment of the appropriate distance  $r_{mn}$ , i.e. correlation of the motions of the electrons  $m$  and  $n$ , etc. The antisymmetrization operator ensures fulfilment of the requirement for symmetry of the wave function with respect to the exchange of the arbitrary two electrons. The method described was independently proposed in 1971 by Wiesław Woźnicki<sup>11</sup> and by Sims and Hagstrom.<sup>12</sup> The method of correlational factors has a nice feature, in that even a short expansion should give a very good total energy for the system, since we combine the power of the CI method with the great success of the explicitly correlated approaches. Unfortunately, the method has also a serious drawback. To make practical calculations, it is necessary to evaluate the integrals occurring in the variational method, and they are very difficult to calculate. It is enough to realize that, in the matrix element of the Hamiltonian containing two terms of the above expansion, we may find, e.g., a term  $1/r_{12}$  (from the Hamiltonian) and  $r_{13}$  (from the factor in front of the determinant), as well as the product of 6 spinorbitals describing the electrons 1, 2, 3. Such integrals have to be computed and the existing algorithms are inefficient.

## 10.4 THE HARMONIC HELIUM ATOM

An unpleasant feature of the electron correlation is that we deal either with intuitive concepts or, if our colleagues want to help us, they bring wave functions with formulae as long as the distance from Cracow to Warsaw (or longer<sup>13</sup>) and say: look, this is what *really* happens. It would be good to analyze such formulae term by term, but this does not make sense, because there are too many terms. Even the helium atom, when we write down the formula for its ground-state wave function, becomes a mysterious object. Correlation of motion of whatever seems to be so difficult to grasp mathematically that we easily give up. A group of scientists published a paper in 1993 which aroused enthusiasm.<sup>14</sup> They obtained a rigorous solution of the Schrödinger equation (described in Chapter 4, p. 188), the only exact solution which has been obtained so far for correlational problems.

<sup>11</sup>W. Woźnicki, in “*Theory of Electronic Shells in Atoms and Molecules*” (ed. A. Yutsis), Mintis, Vilnius, 1971, p. 103.

<sup>12</sup>J.S. Sims, S.A. Hagstrom, *Phys. Rev. A* 4 (1971) 908. This method is known as a Hylleraas–CI.

<sup>13</sup>This is a very conservative estimate. Let us calculate – half jokingly. Writing down a single Slater determinant would easily take 10 cm. The current world record amounts to several billion such determinants in the CI expansion. Say, three billion. Now let us calculate:  $10 \text{ cm} \times 3 \times 10^9 = 3 \times 10^{10} \text{ cm} = 3 \times 10^8 \text{ m} = 3 \times 10^5 \text{ km} = 300000 \text{ km}$ . So, this not Warsaw to Cracow, but Earth to Moon.

<sup>14</sup>S. Kais, D.R. Herschbach, N.C. Handy, C.W. Murray, G.J. Laming, *J. Chem. Phys.* 99 (1993) 417.

Note that the exact wave function (its spatial part<sup>15</sup>) is a *geminal* (i.e. two-electron function).

$$\psi(\mathbf{r}_1, \mathbf{r}_2) = N \left( 1 + \frac{1}{2} r_{12} \right) e^{-\frac{1}{4}(r_1^2 + r_2^2)}. \quad (10.10)$$

Let me be naive. Do we have two harmonic springs here? Yes, we do. Then, let us treat them first as independent oscillators and take the *product* of the ground-state functions of both oscillators:  $\exp[-\frac{1}{4}(r_1^2 + r_2^2)]$ . Well, it would be good to account for the cusp condition  $\psi = \phi(\mathbf{r}_1, \mathbf{r}_2)[1 + \frac{1}{2}r_{12} + \dots]$  and take care of it even in a naive way. Let us just implement the crucial correlation factor  $(1 + \frac{1}{2}r_{12})$ , *the simplest* that satisfies the cusp condition (see p. 505). It turns out, that such a recipe leads to a *rigorous* wave function.<sup>16</sup>

From (10.10) we see that for  $r_1 = r_2 = \text{const}$  (in such a case both electrons move on the surface of the sphere), the larger value of the function (and *eo ipso* of the probability) is obtained for *larger*  $r_{12}$ . This means that, it is most probable that the electrons prefer to occupy opposite sides of a nucleus. This is a practical manifestation of the existence of the Coulomb hole around electrons, i.e. the region of the reduced probability of finding a second electron: the electrons simply repel each other. They cannot move apart to infinity since both are held by the nucleus. The only thing they can do is to be close to the nucleus and to avoid each other – this is what we observe in (10.10).

## 10.5 JAMES-COOLIDGE AND KOŁOS-WOLNIEWICZ FUNCTIONS

One-electron problems are the simplest. For systems with *two* electrons<sup>17</sup> we can apply certain mathematical tricks which allow very accurate results. We are going to talk about such calculations in a moment.

Kołos and Wolniewicz applied the Ritz variational method (see Chapter 5) to the hydrogen molecule with the following trial function:

$$\begin{aligned} \Psi = \frac{1}{\sqrt{2}} [\alpha(1)\beta(2) - \alpha(2)\beta(1)] \sum_i^M c_i (\Phi_i(1, 2) + \Phi_i(2, 1)), \quad (10.11) \\ \Phi_i(1, 2) = \exp(-A\xi_1 - \bar{A}\xi_2) \xi_1^{n_i} \eta_1^{k_i} \xi_2^{m_i} \eta_2^{l_i} \left( \frac{2r_{12}}{R} \right)^{\mu_i} \\ \cdot (\exp(B\eta_1 + \bar{B}\eta_2) + (-1)^{k_i+l_i} \exp(-B\eta_1 - \bar{B}\eta_2)), \end{aligned}$$

<sup>15</sup>For one- and two-electron systems the wave function is a *product* of the spatial and spin factors. A normalized spin factor for two-electron systems,  $\frac{1}{\sqrt{2}}\{\alpha(1)\beta(2) - \beta(1)\alpha(2)\}$ , guarantees that the state in question is a singlet (see Appendix Q, p. 1006). Since we will only manipulate the spatial part of the wave function, the spin is the default. Since the total wave function has to be antisymmetric, and the spin function is antisymmetric, the spatial function should be symmetric and it is.

<sup>16</sup>As a matter of fact, only for a single force constant. Nevertheless, the unusual simplicity of that analytic formula is most surprising.

<sup>17</sup>For a larger number of electrons it is much more difficult.

where the elliptic coordinates of the electrons with index  $j = 1, 2$  are given by:

$$\xi_j = \frac{r_{aj} + r_{bj}}{R}, \quad (10.12)$$

$$\eta_j = \frac{r_{aj} - r_{bj}}{R}, \quad (10.13)$$

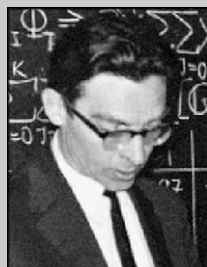
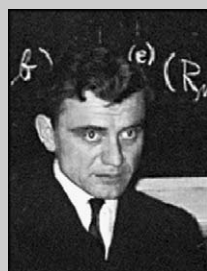
$R$  denotes the internuclear distance,  $r_{aj}$  and  $r_{bj}$  are nucleus–electron distances (the nuclei are labelled by  $a, b$ ),  $r_{12}$  is the (crucial to the method) interelectronic distance,  $c_i$ ,  $A$ ,  $\bar{A}$ ,  $B$ ,  $\bar{B}$  are variational parameters, and  $n, k, l, m$  are integers.

The simplified form of this function with  $A = \bar{A}$  and  $B = \bar{B} = 0$  is the James–Coolidge<sup>18</sup> function, thanks to which the later authors enjoyed the most accurate result for the hydrogen molecule in 27 years.

Kołos and Roothaan,<sup>19</sup> and later on, Kołos and Wolniewicz<sup>20</sup> as well as Kołos and Rychlewski and others<sup>21</sup> applied longer and longer expansions (computer technology was improving fast) up to  $M$  of the order of thousands. The results obtained exceeded the accuracy of experiments, although the latter represented one of the most accurate spectroscopic measurements ever done. Owing to the great precision of these calculations it was proved that quantum mechanics, and in particular the Schrödinger equation, describe the reality with remarkable accuracy, Tables 10.1 and 10.2.

As can be seen from Tables 10.1 and 10.2, there was a competition between theoreticians and the experimental laboratory of Herzberg. When, in 1964, Kołos and Wolniewicz obtained  $36117.3 \text{ cm}^{-1}$  (Table 10.1, bold face) for the dissociation energy of the hydrogen molecule, quantum chemists held their breath. The experimental result of Herzberg and Monfils, obtained four years earlier (Table 10.1, bold face), was

Włodzimierz Kołos (1928–1996), Polish chemist, professor at the Warsaw University. His calculations on small molecules (with Roothaan, Wolniewicz, Rychlewski) took into account all known effects and were of unprecedented accuracy in quantum chemistry. The Department of Chemistry of Warsaw University and the Polish Chemical Society established the Włodzimierz Kołos Medal accompanying a Lecture (the first lecturers were: Roald Hoffmann, Richard Bader and Paul von Ragué Schleyer). In the Ochota quarter in Warsaw there is a Włodzimierz Kołos Street. Lutosław Wolniewicz (born 1927), Polish physicist, professor at the Nicolaus Copernicus University in Toruń.



<sup>18</sup>H.M. James, A.S. Coolidge, *J. Chem. Phys.* 1 (1933) 825. Hubert M. James in the sixties was professor at Purdue University (USA).

<sup>19</sup>W. Kołos, C.C.J. Roothaan, *Rev. Modern Phys.* 32 (1960) 205.

<sup>20</sup>For the first time in quantum chemical calculations relativistic corrections and corrections resulting from quantum electrodynamics were included. This accuracy is equivalent to hitting, from Earth, an object on the Moon the size of a car. These results are cited in nearly all textbooks on quantum chemistry to demonstrate that the theoretical calculations have a solid background.

<sup>21</sup>The description of these calculations is given in the review article by Kołos cited in Table 10.1.

**Table 10.1.** Dissociation energy of H<sub>2</sub> in the ground state (in cm<sup>-1</sup>). Comparison of the results of theoretical calculations and experimental measurements. The references to the cited works can be found in the paper by W. Kołos, *Pol. J. Chem.* 67 (1993) 553. Bold numbers are used to indicate the values connected with the Herzberg–Kołos–Wolniewicz controversy

Year	Author	Experiment	Theory
1926	Witmer	35000	
1927	Heitler–London		23100 <sup>a)</sup>
1933	James–Coolidge		36104 <sup>a)</sup>
1935	Beutler	36116 ± 6	
1960	Kołos–Roothaan		36113.5 <sup>a)</sup>
1960	Herzberg–Monfils	<b>36113.6 ± 0.6</b>	
1964	Kołos–Wolniewicz		<b>36117.3<sup>a)</sup></b>
1968	Kołos–Wolniewicz		<b>36117.4<sup>a)</sup></b>
1970	Herzberg	<b>36118.3<sup>c)</sup></b>	
1970	Stwalley	36118.6 ± 0.5	
1975	Kołos–Wolniewicz		36118.0
1978	Kołos–Rychlewski		36118.12 <sup>b)</sup>
1978	Bishop–Cheung		36117.92
1983	Wolniewicz		36118.01
1986	Kołos–Szalewicz–Monkhorst		36118.088
1991	McCormack–Eyler	36118.26 ± 0.20	
1992	Balakrishnan–Smith–Stoicheff	36118.11 ± 0.08	
1992	Kołos–Rychlewski		36118.049

a) Obtained from calculated binding energy by subtracting the energy of zero vibrations.  
b) Obtained by treating the improvement of the binding energy as an additive correction to the dissociation energy.  
c) Upper bound.

**Table 10.2.** Ionization energy of H<sub>2</sub> (in cm<sup>-1</sup>). See the caption for Table 10.1

Year	Author	Experiment	Theory
1934	Richardson	124569.2	
1933	James–Coolidge		124438
1938	Beutler–Jünger	124429±13	
1969	Jeziorski–Kołos		124417.3
1969	Takezawa	124417±2	
1970	Takezawa	124417.4 ± 0.6	
1972	Herzberg–Jungen	124417.2 ± 0.4	
1978	Kołos–Rychlewski		124417.44
1986	Jungen–Herzberg	124417.5 ± 0.1	
1986/7	Eyler–Short–Pipkin	124417.42 ± 0.15	
1987	Glab–Hessler	124417.61 ± 0.07	
1989	McCormack–Gilligan–Comaggia–Eyler	124417.524±0.015	
1990	Jungen–Dabrowski–Herzberg–Vervloet	124417.501 ± 0.015	
1992	Gilligan–Eyler	124417.507 ± 0.018	
1992	Jungen–Dabrowski–Herzberg–Vervloet	124417.484 ± 0.017	
1992	Eyler et al.	124417.507 ± 0.012	
1992	Kołos–Rychlewski		124417.471

higher and this seemed to contradict the variational principle (Chapter 5) a foundation of quantum mechanics. There were only three possibilities:

- the theoretical result is wrong,
- the experimental result is wrong,
- quantum mechanics has internal inconsistency.

Kołos and Wolniewicz increased the accuracy of their calculations in 1968 and excluded the first possibility. It soon turned out that the problem lay in the accuracy of the experiment.<sup>22</sup> When Herzberg increased the accuracy, he obtained  $36118.3\text{ cm}^{-1}$  as the dissociation energy (Table 10.1, bold face), which was then consistent with the variational principle.

Nowadays, these results are recognized in the world as the most reliable source of information on small molecules. For example, Kołos and Wolniewicz's results for the  $\text{H}_2$  molecule were used to estimate the hydrogen concentration on Jupiter.

Gerhard Herzberg (1904–1999), Canadian chemist of German origin professor at the National Research Council and at the University of Saskatchewan in Saskatoon and the University of Ottawa. The greatest spectroscopist of the XX century. Herzberg laid the foundations of molecular spectroscopy, is author of the fundamental monograph on this subject, received a Nobel prize in 1971 "for his contribution to knowledge of the elec-



tronic structure and geometry of molecules, particularly free radicals".

### 10.5.1 NEUTRINO MASS

Calculations like those above required unique software, especially in the context of the non-adiabatic effects included. Additional gains appeared unexpectedly, when Kołos and others<sup>23</sup> initiated work aiming at explaining whether the electronic neutrino has a non-zero mass or not.<sup>24</sup> In order to interpret the expensive experiments,

<sup>22</sup>At that time Herzberg was hosting them in Canada and treated them to a home made fruit liquor, the latter event was considered by his coworkers to be absolutely exceptional. This is probably the best time to give the recipe for this exquisite drink which is known in the circles of quantum chemists as "kolosovka".

*Pour a pint of pure spirits into a beaker. Hang an orange on a piece of gauze directly over the meniscus. Cover tightly and wait for two weeks. Then throw the orange away – there is nothing of value left in it, and turn your attention to the spirits. It should contain now all the flavours from orange. Next, slowly pour some spring water until the liquid becomes cloudy and some spirits to make it clear again. Propose a toast to the future of quantum chemistry!*

<sup>23</sup>W. Kołos, B. Jezierski, H.J. Monkhorst, K. Szalewicz, *Int. J. Quantum Chem.* S19 (1986) 421.

<sup>24</sup>Neutrinos are stable fermions of spin  $\frac{1}{2}$ . Three types of neutrinos exist (each has its own antiparticle): electronic, muonic and taonic. The neutrinos are created in the weak interactions (e.g., in  $\beta$ -decay) and do not participate either in the strong, or in electromagnetic interactions. The latter feature expresses itself in an incredible ability to penetrate matter (e.g., crossing the Earth almost as through a vacuum). The existence of the electronic neutrino was postulated in 1930 by Wolfgang Pauli and discovered in 1956 by F. Reines and C.L. Cowan; the muonic neutrino was discovered in 1962 by L. Lederman, M. Schwartz and J. Steinberger.

Alexandr Alexandrovich Friedmann (1888–1925), Russian mathematician and physicist, in his article in *Zeit. Phys.* 10 (1922) 377 proved on the basis of Einstein's general theory of relativity, that the curvature of the Universe must change, which became the basis of cosmological models of the expanding Universe. During World War I, Friedman was a pilot in the Russian army and made bombing raids over my beloved Przemyśl. In one of his letters he asked his friend cheerfully, the eminent Russian mathematician Steklov, for advice about the integration of equations he derived to describe the trajectories of his bombs. Later, in a letter to Steklov of February 28, 1915 he wrote: "Recently



*I had an opportunity to verify my theory during a flight over Przemyśl, the bombs fell exactly in the places predicted by the theory. To get the final proof of my theory I intend to test it in flights during next few days."*

More information in: <http://www-groups.dcs.st-and.ac.uk/~history/Mathematicians/Friedmann.html>

Edwin Powell Hubble (1889–1953), American astronomer, explorer of galaxies, found in 1929, that the distance between galaxies is proportional to the infrared shift in their spectrum caused by the Doppler effect, which is consequently interpreted as expansion of the Universe. A surprise from recent astronomical studies is that the ex-



pansion is faster and faster (for reasons unknown).

precise calculations were required for the  $\beta$ -decay of the tritium molecule as a function of the neutrino mass. The emission of the antineutrino ( $\nu$ ) in the process of  $\beta$ -decay:



should have consequences for the final quantum states of the  $\text{HeT}^+$  molecule. To enable evaluation of the neutrino mass by the experimentalists Kołos et al. performed precise calculations of all possible final states of  $\text{HeT}^+$  and presented them as a function of the hypothetical mass of the neutrino. There is such a large number of neutrinos in the Universe that, if its mass exceeded a certain value, even a very small threshold value of the order of 1 eV,<sup>25</sup> the mass of the Universe would exceed the critical value predicted by Alexandr Friedmann in his cosmological theory (based on the general theory of relativity of Einstein). This would mean that the currently occurring expansion of the Universe (discovered by Hubble) would finally stop and its collapse would follow. If the neutrino mass would turn out to be too small, then the Universe would continue its expansion. Thus, quantum chemical calculations for the  $\text{HeT}^+$  molecule may turn out to be helpful in predicting our fate (unfortunately, being crushed or

frozen). So far, the estimate of neutrino mass gives a value smaller than 1 eV, which indicates the Universe expansion.<sup>26</sup>

<sup>25</sup>The mass of the elementary particle is given in the form of its energetic equivalent  $mc^2$ .

<sup>26</sup>At this moment there are other candidates for contributing significantly to the mass of the Universe, mainly the mysterious "dark matter". This constitutes the major part of the mass of the Universe. We know veeeery little.

Recently it turned out that neutrinos undergo what are called oscillations, e.g., an electronic neutrino travels from the Sun and on its way spontaneously changes to a muonic neutrino. The oscillations indicate that the mass of the neutrino is nonzero. According to current estimations, it is much smaller, however, than the accuracy of the tritium experiments.

## 10.6 METHOD OF EXPONENTIALLY CORRELATED GAUSSIAN FUNCTIONS

In 1960, Boys<sup>27</sup> and Singer<sup>28</sup> noticed that the functions which are products of Gaussian orbitals and correlational factors of Gaussian type,  $\exp(-br_{ij}^2)$ , where  $r_{ij}$  is the distance between electron  $i$  and electron  $j$ , generate relatively simple integrals in the quantum chemical calculations. A product of two Gaussian orbitals (with positions shown by the vectors  $\mathbf{A}$ ,  $\mathbf{B}$ ) and of an exponential correlation factor is called an *exponentially correlated Gaussian geminal*.<sup>29</sup>

geminal

$$g(\mathbf{r}_i, \mathbf{r}_j; \mathbf{A}, \mathbf{B}, a_1, a_2, b) = N e^{-a_1(\mathbf{r}_i - \mathbf{A})^2} e^{-a_2(\mathbf{r}_j - \mathbf{B})^2} e^{-br_{ij}^2}.$$

A geminal is an analogue of an orbital, which is a one-electron function. Here is a two-electron one. A single geminal is very rarely used in computations,<sup>30</sup> we apply hundreds or even thousands of Gaussian geminals. When we want to find out what are the optimal positions  $\mathbf{A}$ ,  $\mathbf{B}$  and the optimal exponents  $a$  and  $b$  in these thousands of geminals, it turns out that nothing sure is known about them, the  $\mathbf{A}$ ,  $\mathbf{B}$  positions are scattered chaotically,<sup>31</sup> and in the  $a > 0$  and  $b > 0$  exponents, there is no regularity either. Nevertheless, the above formula for a single Gaussian geminal looks like if it suggested  $b > 0$ .

## 10.7 COULOMB HOLE ("CORRELATION HOLE")

It is always good to count "on fingers" to make sure that everything is all right. Let us see how a *single* Gaussian geminal describes the correlation of the electronic motion. Let us begin with the helium atom with the nucleus in the position  $\mathbf{A} = \mathbf{B} = \mathbf{0}$ . The geminal takes the form:

$$g_{\text{He}} = N e^{-a_1 r_1^2} e^{-a_2 r_2^2} e^{-br_{12}^2}, \quad (10.14)$$

where  $N$  is a normalization factor. Let us assume<sup>32</sup> that electron 1 is at  $(x_1, y_1, z_1) = (1, 0, 0)$ . Where in such situation does electron 2 prefer to be? We will find out (Fig. 10.2) from the position of electron 2 for which  $g_{\text{He}}$  assumes the largest value.

Just to get an idea, let us try to restrict the motion of electron 2. For instance, let us demand that it moves only on the sphere of radius equal to 1 centred at the nucleus. So we insert  $r_1 = r_2 = 1$ . Then,  $g_{\text{He}} = \text{const} \exp[-br_{12}^2]$  and we will easily find out what electron 2 likes most. With  $b > 0$  the latter factor tells us that

<sup>27</sup>S.F. Boys, *Proc. Royal Soc. A* 258 (1960) 402.

<sup>28</sup>K. Singer, *Proc. Royal Soc. A* 258 (1960) 412.

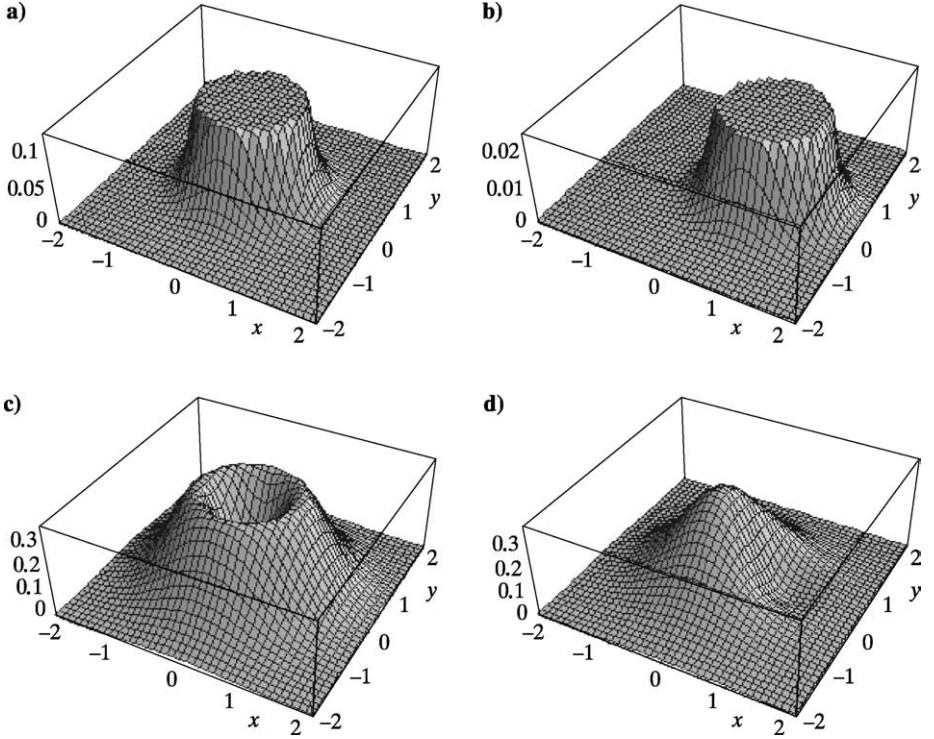
<sup>29</sup>This is an attempt to go beyond the two-electron systems with the characteristic (for these systems) approach of James, Coolidge, Hylleraas, Kolos, Wolniewicz and others.

<sup>30</sup>Ludwik Adamowicz introduced an idea of the minimal basis of the Gaussian geminals (equal to the number of the electron pairs) and applied to the LiH and HF molecules, L. Adamowicz, A.J. Sadlej, *J. Chem. Phys.* 69 (1978) 3992.

<sup>31</sup>The methods in which those positions are selected at random scored a great success.

<sup>32</sup>We use atomic units.





**Fig. 10.2.** Illustration of the correlation and ... anticorrelation of the electrons in the helium atom. Figs. (a) and (b) present the machinery of the “anticorrelation” connected with the geminal  $g_{\text{He}} = N \exp[-r_1^2] \exp[-r_2^2] \exp[-2r_{12}^2]$ . In Fig. (a) electron 1 has a position  $(0, 0, 0)$ , while Fig. (b) corresponds to electron 1 being at point  $(1, 0, 0)$  (cutting off the top parts of the plots is caused by graphical limitations, not by the physics of the problem). It can be seen that electron 2 *holds on to electron 1*, i.e. it behaves in a completely unphysical manner (since electrons repel each other). Figs. (c) and (d) show how electron 2 will respond to such two positions of electron 1, if the wave function is described by the geminal  $g_{\text{He}} = N \exp[-r_1^2] \exp[-r_2^2] [1 - \exp[-2r_{12}^2]]$ . In Fig. (c) we see that electron 2 runs away “with all its strength” (the hollow in the middle) from electron 1 placed at  $(0, 0, 0)$ . We have correlation. Similarly, Fig. (d), if electron 1 is in point  $(1, 0, 0)$ , then it causes a slight depression for electron 2 in this position. Again we do have correlation. However, the graphs (c) and (d) differ widely. This is understandable since the nucleus is all the time at the point  $(0, 0, 0)$ . Figs. (e), (f) correspond to the same displacements of electron 1, but this time the correlation function is equal to  $\psi(\mathbf{r}_1, \mathbf{r}_2) = (1 + \frac{1}{2}r_{12}) \exp[-(r_1^2 + r_2^2)]$ , i.e. is similar to the wave function of the harmonic helium atom. It can be seen (particularly in Fig. (e)) that there *is* a correlation, although much less visible than in the previous examples. To amplify (artificially) the correlation effect Figs. (g), (h) show the same as Figs. (e), (f) but for the function  $\psi(\mathbf{r}_1, \mathbf{r}_2) = (1 + 25r_{12}) \exp[-(r_1^2 + r_2^2)]$ , which (unlike Figs. (e), (f)) does not satisfy the correlation cusp condition.

what electron 2 likes best is just to sit on electron 1! Is it what the correlation is supposed to mean that one electron sits on the other? Here we have rather an anticorrelation. Something is going wrong. According to this analysis we should rather take the geminal of the form, e.g.:

$$g_{\text{He}} = N e^{-a_1 r_1^2} e^{-a_1 r_2^2} [1 - e^{-b r_{12}^2}].$$

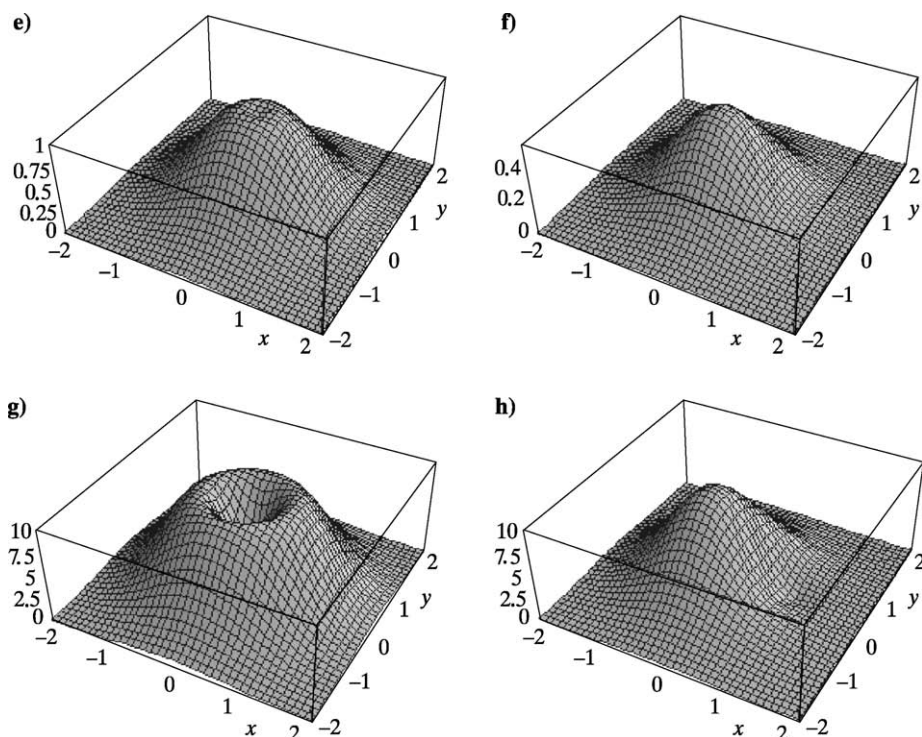


Fig. 10.2. Continued.

Now everything is qualitatively in order. When the interelectronic distance increases, the value of the  $g_{\text{He}}$  function also increases, which means that such a situation is *more* probable than that corresponding to a short distance. If the electrons become too agitated and begin to think that it would be better when their distance gets very long, they would be called to order by the factors  $\exp[-a_1 r_1^2] \exp[-a_1 r_2^2]$ . Indeed, in such a case, the distance between the nucleus and at least one of the electrons is long and the probability of such a situation is quenched by one or both exponential factors. For large  $r_{12}$  distances, the factor  $[1 - \exp[-b r_{12}^2]]$  is practically equal to 1. This means that for large interelectronic distances  $g_{\text{He}}$  is practically equal to  $N \exp[-a_1 r_1^2] \exp[-a_1 r_2^2]$ , i.e. to the product of the orbitals (no correlation of motions at long interelectronic distances, and rightly so).

Around electron 1 there is a region of low probability of finding electron 2. This region is called the Coulomb hole.

The Gaussian geminals do not satisfy the correlation cusp condition (p. 505), because of factor  $\exp(-b r_{ij}^2)$ . It is required (for simplicity we write  $r_{ij} = r$ ) that  $(\frac{\partial g}{\partial r})_{r=0} = \frac{1}{2} g(r=0)$ , whereas the left-hand side is equal to 0, while the right-hand side  $\frac{1}{2} N \exp[-a_1(r_i - A)^2] \exp[-a_2(r_j - B)^2]$  is not equal to zero. This is not a

disqualifying feature, since the region of space in which this condition should be fulfilled, is very small.

The area of application of this method is – for practical (computational) reasons – relatively small. The method of Gaussian geminals has been applied in unusually accurate calculations for three- and four-electron systems.<sup>33</sup>

## 10.8 EXCHANGE HOLE (“FERMI HOLE”)

The mutual avoidance of electrons in helium atom or in hydrogen molecule is caused by Coulombic repulsion of electrons (“Coulomb hole”, see above). As we have shown in this Chapter, in the Hartree–Fock method the Coulomb hole is absent, whereas methods which account for electron correlation generate such a hole. However, electrons avoid each other not only because of their charge. The Pauli principle is an additional reason. One of the consequences is the fact that electrons with the same spin coordinate cannot reside in the same place, see p. 33. The continuity of the wave function implies that the probability density of them staying *in the vicinity* of each other is small, i.e.

around the electron there is a NO PARKING area for other electrons with the same spin coordinate (“exchange, or Fermi hole”).

Let us see how such exchange holes arise. We will try to make the calculations as simple as possible.

We have shown above that the Hartree–Fock function does not include any electron correlation. We must admit, however, that we have come to this conclusion on the basis of the two-electron closed shell case. This is a special situation, since both electrons have *different* spin coordinates ( $\sigma = \frac{1}{2}$  and  $\sigma = -\frac{1}{2}$ ). Is it really true that the Hartree–Fock function does not include any correlation of electronic motion?

We take the  $\text{H}_2^-$  molecule in the simplest formulation of the LCAO MO method (two atomic orbitals only:  $1s_a = \chi_a$  and  $1s_b = \chi_b$ , two molecular orbitals: bonding  $\varphi_1 = \frac{1}{\sqrt{2(1+S)}}(\chi_a + \chi_b)$  and antibonding  $\varphi_2 = \frac{1}{\sqrt{2(1-S)}}(\chi_a - \chi_b)$ , cf. p. 371; the overlap integral  $S \equiv (\chi_a | \chi_b)$ ). We have three electrons. As a wave function we will take the single (normalized) Hartree–Fock determinant (UHF) with the following orthonormal spinorbitals occupied:  $\phi_1 = \varphi_1\alpha$ ,  $\phi_2 = \varphi_1\beta$ ,  $\phi_3 = \varphi_2\alpha$ :

$$\psi_{\text{UHF}}(1, 2, 3) = \frac{1}{\sqrt{3!}} \begin{vmatrix} \phi_1(1) & \phi_1(2) & \phi_1(3) \\ \phi_2(1) & \phi_2(2) & \phi_2(3) \\ \phi_3(1) & \phi_3(2) & \phi_3(3) \end{vmatrix}.$$

### Example 1

We are interested in electron 3 with electron 1 residing at nucleus  $a$  with space coordinates  $(0, 0, 0)$  and with spin coordinate  $\sigma_1 = \frac{1}{2}$  and with electron 2 located at

<sup>33</sup>W. Cencek, Ph.D. Thesis, Adam Mickiewicz University, Poznań, 1993, also J. Rychlewski, W. Cencek, J. Komasa, *Chem. Phys. Letters* 229 (1994) 657; W. Cencek, J. Rychlewski, *Chem. Phys. Letters* 320 (2000) 549. All these results were world records.

nucleus  $b$  with coordinates  $(R, 0, 0)$  and  $\sigma_2 = -\frac{1}{2}$ , whereas the electron 3 itself has spin coordinate  $\sigma_3 = \frac{1}{2}$ . The square of the absolute value of the function calculated for these values depends on  $x_3, y_3, z_3$  and represents the *conditional probability* density distribution for finding electron 3 (provided electrons 1 and 2 have the fixed coordinates given above and denoted by  $1_0, 2_0$ ). So let us calculate individual elements of the determinant  $\psi_{\text{UHF}}(1_0, 2_0, 3)$ , taking into account the properties of spin functions  $\alpha$  and  $\beta$  (cf. p. 28):

$$\psi_{\text{UHF}}(1_0, 2_0, 3) = \frac{1}{\sqrt{3!}} \begin{vmatrix} \varphi_1(0, 0, 0) & 0 & \varphi_1(x_3, y_3, z_3) \\ 0 & \varphi_1(R, 0, 0) & 0 \\ \varphi_2(0, 0, 0) & 0 & \varphi_2(x_3, y_3, z_3) \end{vmatrix}.$$

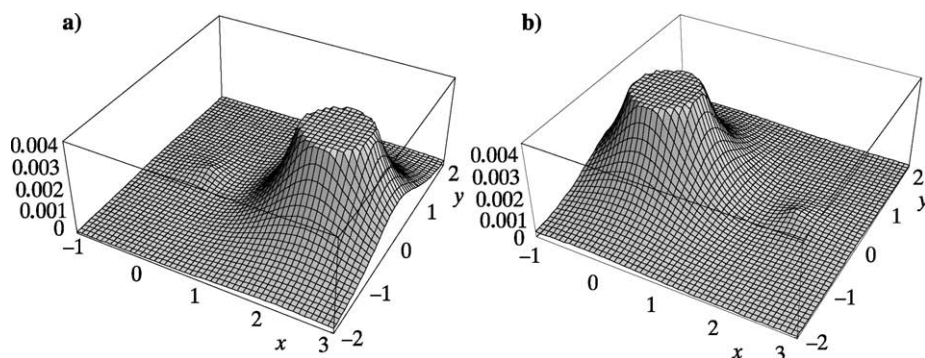
Using the Laplace expansion (Appendix A on p. 889) we get

$$\psi_{\text{UHF}}(1_0, 2_0, 3) = \frac{1}{\sqrt{3!}} [\varphi_1(0, 0, 0)\varphi_1(R, 0, 0)\varphi_2(x_3, y_3, z_3) - \varphi_1(x_3, y_3, z_3)\varphi_1(R, 0, 0)\varphi_2(0, 0, 0)].$$

The plot of this function (the overlap integral  $S$  is included in normalization factors of the molecular orbitals) is given in Fig. 10.3.

Qualitatively, however, everything is clear even without the calculations. Due to the forms of the molecular orbitals ( $S$  is small)  $\varphi_1(0, 0, 0) = \varphi_1(R, 0, 0) \approx \varphi_2(0, 0, 0) = \text{const}$  we get:

$$\psi_{\text{UHF}}(1_0, 2_0, 3) \approx -\text{const}^2 \frac{1}{\sqrt{3}} \chi_b(3)$$



**Fig. 10.3.** Demonstration of the exchange ("Fermi") hole in the  $\text{H}_2^-$  molecular ion (truncation of the hills is artificial, without this it would be more difficult to see the details of the figure). (a)  $|\psi_{\text{UHF}}(1_0, 2_0, 3)|^2$  is the probability density of finding the spatial coordinates of electron 3 (having  $\sigma_3 = \frac{1}{2}$ ) provided that electron 1 resides on the nucleus  $a$  at  $(0, 0, 0)$  having  $\sigma_1 = \frac{1}{2}$  and electron 2 sits on nucleus  $b$  at  $(R = 2, 0, 0)$  and has  $\sigma_2 = -\frac{1}{2}$ ; (b) the same as above, but this time electron 1 has moved to nucleus  $b$  (i.e. it shares  $b$  with electron 2).

so the conditional probability density of finding electron 3 is

$$\rho(3) \approx \frac{1}{3} \text{const}^4 [\chi_b(3)]^2. \quad (10.15)$$

We can see that *for some reason electron 3 has moved in the vicinity of nucleus b*. What scared it so much, when we placed one of the two electrons at *each* nucleus? Electron 3 ran to be as far away as possible from electron 1 residing on *a*. It hates electron 1 so much that it has just ignored the Coulomb repulsion with electron 2 sitting on *b* and jumped on it! What the hell has happened? Well, we have some suspicions. Electron 3 could have been scared only by the spin coordinate of electron 1, *the same as its own*.

*This is just an indication of the exchange hole around each electron.*

## Example 2

Maybe electron 3 does not run away from anything, but simply always resides at nucleus *b*? Let us make sure of that. Let us move electron 1 to nucleus *b* (there is already electron 2 sitting over there, but it does not matter). What then will electron 3 do? Let us see. We have electrons 1 and 2 at nucleus *b* with space coordinates  $(R, 0, 0)$  and spin coordinates  $\sigma_1 = \frac{1}{2}$ ,  $\sigma_2 = -\frac{1}{2}$ , whereas electron 3 has spin coordinate  $\sigma_3 = \frac{1}{2}$ . To calculate the conditional probability we have to calculate the value of the wave function.

This time

$$\begin{aligned} \psi_{\text{UHF}}(1_0, 2_0, 3) &= \frac{1}{\sqrt{3!}} \begin{vmatrix} \varphi_1(R, 0, 0) & 0 & \varphi_1(x_3, y_3, z_3) \\ 0 & \varphi_1(R, 0, 0) & 0 \\ \varphi_2(R, 0, 0) & 0 & \varphi_2(x_3, y_3, z_3) \end{vmatrix} \\ &\approx \text{const}^2 \frac{1}{\sqrt{3}} \chi_a(3) \end{aligned}$$

or

$$\rho(3) \approx \frac{1}{3} \text{const}^4 [\chi_a(3)]^2. \quad (10.16)$$

We see that electron 3 with spin coordinate  $\sigma_3 = \frac{1}{2}$  runs in panic to nucleus *a*, because it is as scared of electron 1 with spin  $\sigma_1 = \frac{1}{2}$  as the devil is of holy water.

## Example 3

And what would happen if we made the decision for electron 3 more difficult? Let us put electron 1 ( $\sigma_1 = \frac{1}{2}$ ) *in the centre* of the molecule and electron 2 ( $\sigma_2 = -\frac{1}{2}$ ) as before, at nucleus *b*. According to what we think about the whole machinery, electron 3 (with  $\sigma_3 = \frac{1}{2}$ ) should run away from electron 1, because both electrons have the same spin coordinates, and this is what they hate most. But *where* should it run? Will electron 3 select nucleus *a* or nucleus *b*? The nuclei do not look equivalent. There is an electron sitting at *b*, while the *a* centre is empty. Maybe electron 3 will jump to *a* then? Well, the function analyzed is Hartree–Fock – electron 3

ignores the Coulomb hole (it does not see electron 2 sitting on  $b$ ) and therefore will not prefer the empty nucleus  $a$  to sit at. It looks like electron 3 will treat both nuclei on the same basis. In the case of two atomic orbitals, electron 3 has only the choice: either bonding orbital  $\varphi_1$  or antibonding orbital  $\varphi_2$  (in both situations the electron densities on  $a$  and on  $b$  are equal, no nucleus is distinguished). Out of the two molecular orbitals,  $\varphi_2$  looks much more attractive to electron 3, because it has a node<sup>34</sup> exactly, where electron 1 with its nasty spin is. This means that there is a chance for electron 3 to take care of the Fermi hole of electron 1: we predict that electron 3 will "select" only  $\varphi_2$ . Let us check this step by step:

$$\begin{aligned}
 \psi_{\text{UHF}}(1_0, 2_0, 3) &= \frac{1}{\sqrt{3!}} \begin{vmatrix} \varphi_1(\frac{R}{2}, 0, 0) & 0 & \varphi_1(x_3, y_3, z_3) \\ 0 & \varphi_1(R, 0, 0) & 0 \\ \varphi_2(\frac{R}{2}, 0, 0) & 0 & \varphi_2(x_3, y_3, z_3) \end{vmatrix} \\
 &= \frac{1}{\sqrt{3!}} \begin{vmatrix} \varphi_1(\frac{R}{2}, 0, 0) & 0 & \varphi_1(x_3, y_3, z_3) \\ 0 & \varphi_1(R, 0, 0) & 0 \\ 0 & 0 & \varphi_2(x_3, y_3, z_3) \end{vmatrix} \\
 &= \frac{1}{\sqrt{3!}} \varphi_1\left(\frac{R}{2}, 0, 0\right) \varphi_1(R, 0, 0) \varphi_2(x_3, y_3, z_3) \\
 &= \text{const}_1 \varphi_2(x_3, y_3, z_3).
 \end{aligned}$$

And it does exactly so.

#### Example 4

Why is the hole called the exchange hole? Perhaps it would be enough to take the product function<sup>35</sup> and then we would also see that electron 3 runs away in panic from the other electron with the same spin? Let us see how it is in the first case (Example 1):

$$\begin{aligned}
 \psi_{\text{Hartree}}(1, 2, 3) &= \phi_1(1)\phi_2(2)\phi_3(3) = \varphi_1(1)\alpha(1)\varphi_1(1)\beta(1)\varphi_2(3)\alpha(3), \\
 \psi_{\text{Hartree}}(1_0, 2_0, 3) &= \varphi_1(0, 0, 0)\varphi_1(R, 0, 0)\varphi_2(x_3, y_3, z_3) = \text{const}^2 \varphi_2(x_3, y_3, z_3).
 \end{aligned}$$

We get the distribution

$$\rho_{\text{Hartree}} = \text{const}^4 |\varphi_2(x_3, y_3, z_3)|^2.$$

And what do we get in the second case (Example 2)?

$$\begin{aligned}
 \psi_{\text{Hartree}}(1, 2, 3) &= \phi_1(1)\phi_2(2)\phi_3(3) = \varphi_1(1)\alpha(1)\varphi_1(1)\beta(1)\varphi_2(3)\alpha(3), \\
 \psi_{\text{Hartree}}(1_0, 2_0, 3) &= \varphi_1(R, 0, 0)\varphi_1(R, 0, 0)\varphi_2(x_3, y_3, z_3) = \text{const}^2 \varphi_2(x_3, y_3, z_3).
 \end{aligned}$$

<sup>34</sup>That is, low probability of finding electron 3 over there.

<sup>35</sup>"Illegal" (Hartree approximation), since it does not obey the Pauli principle.

Hence, electron 3 occupies the antibonding orbital  $\varphi_2$  and does not even think of running away from anything. Its distribution is entirely insensitive to the position of electron 1.

Thus, *this hole results from the Pauli principle, i.e. from the exchange of electron numbering, hence the name “exchange hole”*.

Summing up, the wave function of the electronic system:

- should account for the existence of the Coulomb hole around each electron, i.e. for the reduced probability for finding any other electron there;
- should also account for the exchange hole, i.e., in the vicinity of an electron with a definite spin coordinate there should be reduced probability for finding any other electron with the same spin coordinate;
- as we saw, the Hartree–Fock function does not account at all for the Coulomb hole, however, it takes into account the existence of the exchange hole.

*Which hole is more important: Coulomb or exchange?* This question will be answered in Chapter 11.

## VARIATIONAL METHODS WITH SLATER DETERMINANTS

### 10.9 VALENCE BOND (VB) METHOD

#### 10.9.1 RESONANCE THEORY – HYDROGEN MOLECULE

Slater determinants are usually constructed from *molecular* spinorbitals. If, instead, we use *atomic* spinorbitals and the Ritz variational method (Slater determinants as the expansion functions) we would get the most general formulation of the valence bond (VB) method. The beginning of VB theory goes back to papers by Heisenberg. The first application was made by Heitler and London, and later theory was generalized by Hurley, Lennard-Jones and Pople.<sup>36</sup>

The essence of the VB method can be explained by an example. Let us take the hydrogen molecule with atomic spinorbitals of type  $1s_a$   $\alpha$  and  $1s_b$   $\beta$  denoted shortly as  $a\alpha$  and  $b\beta$  centred at two nuclei. Let us construct from them several (non-normalized) Slater determinants, for instance:

$$\psi_1 = \begin{vmatrix} a(1)\alpha(1) & a(2)\alpha(2) \\ b(1)\beta(1) & b(2)\beta(2) \end{vmatrix} = [a(1)\alpha(1)b(2)\beta(2) - a(2)\alpha(2)b(1)\beta(1)],$$

$$\psi_2 = \begin{vmatrix} a(1)\beta(1) & a(2)\beta(2) \\ b(1)\alpha(1) & b(2)\alpha(2) \end{vmatrix} = [a(1)\beta(1)b(2)\alpha(2) - a(2)\beta(2)b(1)\alpha(1)],$$

<sup>36</sup>W. Heisenberg, *Zeit. Phys.* 38 (1926) 411, *ibid.* 39 (1926) 499, *ibid.* 41 (1927) 239; W. Heitler, F. London, *Zeit. Phys.* 44 (1927) 455; A.C. Hurley, J.E. Lennard-Jones, J.A. Pople, *Proc. Roy. Soc. London A* 220 (1953) 446.

$$\begin{aligned}\psi_3 &= \begin{vmatrix} a(1)\alpha(1) & a(2)\alpha(2) \\ a(1)\beta(1) & a(2)\beta(2) \end{vmatrix} = [a(1)\alpha(1)a(2)\beta(2) - a(2)\alpha(2)a(1)\beta(1)] \\ &= [a(1)a(2)][\alpha(1)\beta(2) - \alpha(2)\beta(1)] \equiv \psi_{H^-H^+}, \\ \psi_4 &= \begin{vmatrix} b(1)\alpha(1) & b(2)\alpha(2) \\ b(1)\beta(1) & b(2)\beta(2) \end{vmatrix} = [b(1)b(2)][\alpha(1)\beta(2) - \alpha(2)\beta(1)] \equiv \psi_{H^+H^-}.\end{aligned}$$

The functions  $\psi_3$ ,  $\psi_4$  and the normalized difference  $\psi_1 - \psi_2$  ( $N_{\text{HL}}$  is a normalization factor)

Heitler–London function

$$\begin{aligned}\psi_{\text{HL}} &= N_{\text{HL}}(\psi_1 - \psi_2) \\ &= N_{\text{HL}}[a(1)b(2) + a(2)b(1)][\alpha(1)\beta(2) - \alpha(2)\beta(1)] \quad (10.17)\end{aligned}$$

are eigenfunctions of the operators  $\hat{S}^2$  and  $\hat{S}_z$  (cf. Appendix Q, p. 1006) corresponding to the singlet state. The functions  $\psi_3$ ,  $\psi_4$  for obvious reasons are called *ionic structures* ( $H^-H^+$  and  $H^+H^-$ ),<sup>37</sup> whereas the function  $\psi_{\text{HL}}$  is called a Heitler–London function or a *covalent structure*.<sup>38</sup>

ionic structure

The VB method relies on optimization of the expansion coefficients  $c$  in front of these structures in the Ritz procedure (p. 202)

covalent structure

$$\psi = c_{\text{cov}}\psi_{\text{HL}} + c_{\text{ion1}}\psi_{H^-H^+} + c_{\text{ion2}}\psi_{H^+H^-}. \quad (10.18)$$

The covalent structure itself,  $\psi_{\text{HL}}$ , was one great success of Walter Heitler<sup>39</sup> and Fritz London. For the first time the correct description of the chemical bond was

Fritz Wolfgang London (1900–1954) was born in Breslau (now Wrocław) and studied in Bonn, Frankfurt, Göttingen, Munich (Ph.D. at 21) and in Paris. Later worked in Zurich, Rome and Berlin. Escaped from nazism to UK, where he worked at Oxford University (1933–1936). In 1939 London emigrated to the USA, where he became professor of theoretical chemistry at Duke University in Durham.

Fritz London rendered great services to quantum chemistry. He laid the foundations of the theory of the *chemical* (covalent) bond and also, in addition, introduced dispersion interactions, one of the most important *intermolecular*



interactions. This is nearly all of what chemistry is about. He also worked in the field of superconductivity.

<sup>37</sup>Since both electrons reside at the same nucleus.

<sup>38</sup>Since both electrons belong to the same extent to each of the nuclei.

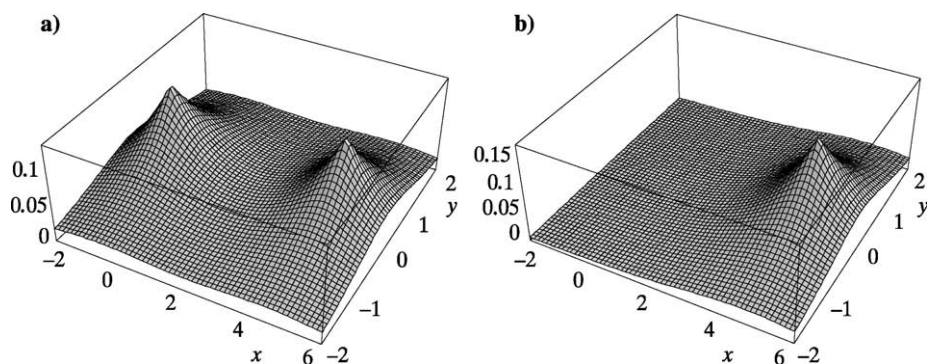
<sup>39</sup>Walter Heitler (1904–1981), German chemist, professor at the University in Göttingen, later in Bristol and Zürich.



obtained. The crucial point turned out to be an inclusion – in addition to the product function  $a(1)b(2)$  – its counterpart *with exchanged electron numbers*  $a(2)b(1)$ , since the electrons *are* indistinguishable. If we expand the Hartree–Fock determinant with doubly occupied bonding orbital  $a + b$ , we would also obtain a certain linear combination of the three structures mentioned, but with *the constant coefficients independent of the interatomic distance*:

$$\psi_{\text{RHF}} = N \left( \frac{1}{N_{\text{HL}}} \psi_{\text{HL}} + \psi_{H^-H^+} + \psi_{H^+H^-} \right). \quad (10.19)$$

This leads to a very bad description of the  $\text{H}_2$  molecule at long internuclear distances with the Hartree–Fock method. The true wave function should contain, among other things, both the covalent structure (i.e. the Heitler–London function) and the ionic structures. However, for long internuclear distances the Heitler–London function should dominate, because it corresponds to the (exact) dissociation limit (two ground-state hydrogen atoms). The trouble is that, with fixed coefficients, *the Hartree–Fock function overestimates the role of the ionic structure* for long interatomic distances. Fig. 10.4 shows that the Heitler–London function describes the electron correlation (Coulomb hole), whereas the Hartree–Fock function does not.



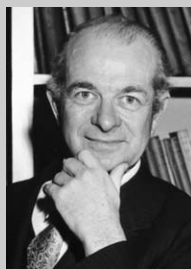
**Fig. 10.4.** Illustration of electron correlation in the hydrogen molecule. The nuclear positions are  $(0, 0, 0)$  and  $(4, 0, 0)$  in a.u. Slater orbitals of  $1s$  type have orbital exponent equal to 1. (a) Visualization of the  $xy$  cross-section of the wave function of electron 2, assuming that electron 1 resides on the nucleus (either the first or the second one), has spin coordinate  $\sigma_1 = \frac{1}{2}$ , whereas electron 2 has spin coordinate  $\sigma_2 = -\frac{1}{2}$  and the total wave function is equal  $\psi = N\{ab + ba + aa + bb\}\{\alpha\beta - \beta\alpha\}$ , i.e. it is a Hartree–Fock function. The plot is the same independently of which nucleus electron 1 resides, i.e., we observe the *lack of any correlation* of the motions of electrons 1 and 2. If we assume the spins to be parallel ( $\sigma_2 = \frac{1}{2}$ ), the wave function vanishes. (b) A similar plot, but for the Heitler–London function  $\psi_{\text{HL}} = N_{\text{HL}}[a(1)b(2) + a(2)b(1)][\alpha(1)\beta(2) - \alpha(2)\beta(1)]$  and with electron 1 residing at nucleus  $(0, 0, 0)$ . Electron 2 runs to the nucleus in position  $(4, 0, 0)$ . We have the correlation of the electronic motion. If we assume parallel spins ( $\sigma_2 = \frac{1}{2}$ ), the wave function vanishes.

## 10.9.2 RESONANCE THEORY – POLYATOMIC CASE

The VB method was developed by Linus Pauling under the name of *theory of resonance*.

resonance  
theory

Linus Carl Pauling (1901–1994), American physicist and chemist, in the years 1931–1964 professor at the California Institute of Technology in Pasadena, in 1967–1969 professor at the University of California, San Diego, from 1969–1974 professor at the Stanford University. He received the 1954 Nobel prize: “for his research into the nature of the chemical bond and its application to the elucidation of the structure of complex substances”. In 1962 he received the Nobel peace prize. His major achievements are the development of the theory of chemical bond, i.e., the VB method (also called resonance theory), and determining the

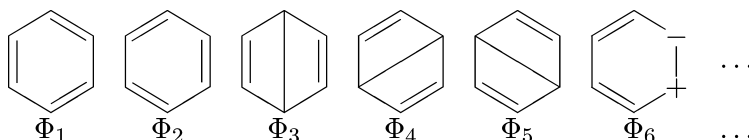


structure of one of the fundamental structural elements of proteins, the  $\alpha$ -helix.

The method can be applied to all molecules, although a particularly useful field of applications of resonance theory can be found in the organic chemistry of aromatic systems. For example, the total electronic wave function of the benzene molecule is presented as a linear combination of resonance structures<sup>40</sup>

$$\psi = \sum_I c_I \Phi_I, \quad (10.20)$$

to each (in addition to the mathematical form), a graph is assigned. For example, six  $\pi$  electrons can participate in the following “adventures” (forming covalent and ionic bonds)



The first two structures are famous Kekulé structures, the next three are Dewar structures, the sixth is an example of the possible mixed covalent-ionic structures. From these graphs, we may deduce which atomic orbitals (out of the  $2p_z$  orbital of carbon atoms,  $z$  is perpendicular to the plane of the benzene ring) takes part in the covalent bond (of the  $\pi$  type). As far as the mathematical form of the  $\Phi_1$  structure is concerned, we can write it as the antisymmetrized (cf. antisymmetrization operator, p. 986) product of three Heitler–London functions (involving the proper pairs of  $2p_z$  carbon atomic orbitals), the first for electrons 1, 2, the second for electrons 3, 4, and the third for 5, 6. Within the functions  $\Phi_I$ , the ionic structures can

<sup>40</sup>Similar to the original applications, we restrict ourselves to the  $\pi$  electrons, the  $\sigma$  electrons are treated as inactive in each structure, forming, among other things, the six C–C bonds presented below.

also occur. The rules for writing the structures were not quite clear, and the electrons were located to some extent in an arbitrary manner, making the impression that it is up to theoretical chemists to use their imaginations and draw imaginary pictures and – next – to translate them into mathematical form to obtain – after applying the variational method – an approximation to the wave function (and to the energy).

In fact, the problem is connected to the Ritz method and to expansion into the complete set of functions,<sup>41</sup> i.e. a purely mathematical problem. Although it may seem very strange to students (fortunately), many people were threatened for supporting the theory of resonance. Scientists serving the totalitarian regime decided to attack eq. (10.20). How, was this possible?<sup>42</sup> The Stalinists did not like the idea that “*the sum of fictitious structures can describe reality*”. Wait a second! If some artificial functions could interfere with reality then socialist realism may lose to abstraction, a kolkhoz (collective farm) member to an intellectual, Lysenkoism to Mendelism,<sup>43</sup> goulags to the idea of freedom, and you are on the brink of disaster.

<sup>41</sup>In principle, *they should* form the complete set, but even so, in practical calculations, we never deal with true complete sets.

<sup>42</sup>Of course, *the true* reason was not a convergence of a series in the Hilbert space, but their personal careers *at any price*. Totalitarian systems never have problems finding such “scientists”. In chemistry, there was the danger of losing a job, in biology, of losing a life.

It is rather difficult to think about Joseph Stalin as a quantum chemist. He was, however, kept informed about the current situation of a group of people involved in carrying out the summations in eq. (10.20), i.e. working in the resonance theory. To encourage young people to value and protect the freedom they have, and to reflect on human nature, some excerpts from the resolution adopted by the All-Soviet Congress of Chemists of the Soviet Union are reported. The resolution pertains, i.a., to the theory of resonance (after the disturbing and reflective book by S.E. Schnoll, “*Gieroi i zlodiei rossijskoj nauki*”, Kron-Press, Moscow, 1997, p. 297):

“Dear Joseph Vissarionovich (Stalin),  
the participants of the All-Soviet Congress send to you, the Great Leader and Teacher of all progressive mankind, our warm and cordial greetings. We Soviet chemists gathered together to decide, by means of broad and free discussion, the fundamental problems of the contemporary theory of the structure of molecules, want to express our deepest gratitude to you for the everyday attention you pay to Soviet science, particularly to chemistry. Our Soviet chemistry is developing in the Stalin era, which offers unlimited possibilities for the progress of science and industry. Your brilliant work in the field of linguistics put the tasks for still swifter progress in front of all scientists of our fatherland (...). Motivated by the resolutions of the Central Committee of the Bolshevik Communist Party concerning ideological matters and by your instructions, Comrade Stalin, the Soviet chemists wage war against the ideological concepts of bourgeois science. The lie of the so called “resonance theory” has been disclosed, and the remains of this idea will be thrown away from the Soviet chemistry. We wish you, our dear Leader and Teacher, good health and many, many years of famous life to the joy and happiness of the whole of progressive mankind(...).”

The events connected with the theory of resonance started in the autumn of 1950 at Moscow University. Quantum chemistry lecturers, Yakov Kiovitch Syrkin and Mirra Yefimovna Diatkina, were attacked. The accusation was about dissemination of the theory of resonance and was launched by former assistants of Syrkin. Since everything was in the hands of the professionals, Syrkin and Diatkina pleaded guilty with respect to each of the charges.

<sup>43</sup>Trofim Lysenko (1898–1976), Soviet scientist of enormous political influence, rejected the genetic laws of Mendel. In my 7th grade school biology textbook virtually only his “theory” was mentioned. As a pupil, I recall wanting to learn this theory. It was impossible to find any information. With difficulties I finally found something: acorns should be placed in a hole in the ground in large numbers to permit something like the class struggle. The winner will be the strongest oak-tree and this is what we all want.

Gregor Johann Mendel (1822–1884), modest Moravian monk, from 1843 a member of the Augustinian order in Brno (abbot from 1868). His unusually precise and patient experiments with sweet peas of two colours and seeds of two degrees of smoothness, allowed him to formulate the principal laws of genetics. Only in 1900 were his fundamental results remembered, and since then the rapid progress of contemporary genetics began.



## 10.10 CONFIGURATION INTERACTION (CI) METHOD

In this method<sup>44</sup>

the variational wave function is a linear combination of Slater determinants constructed from *molecular* spinorbitals, an expansion analogous to eq. (10.20).

In most cases we are interested in the function  $\psi$  for *the electronic ground state of the system* (in addition when solving the CI equations we also get approximations to the excited states with different values of the  $c_I$  coefficients).

Generally we construct the Slater determinants  $\Phi_I$  by placing electrons on the molecular spinorbitals obtained with the Hartree–Fock method,<sup>45</sup> in most cases the set of determinants is additionally limited by imposing an upper bound for the orbital energy. In that case, the expansion in (10.20) is finite. The Slater determinants  $\Phi_I$  are obtained by the replacement of occupied spinorbitals with virtual ones in the single Slater determinant, which is – in most cases – the Hartree–Fock

CI method

<sup>44</sup>Also called the method of superposition of configurations.

<sup>45</sup>In this method we obtain  $M$  molecular orbitals, i.e.  $2M$  molecular spinorbitals, where  $M$  is the number of atomic orbitals employed. The Hartree–Fock determinant  $\Phi_0$  is the best form of wave function as long as the electronic correlation is not important. The criterion of this “goodness” is the mean value of the Hamiltonian. If we want to include the electron correlation, we may think of another form of the one-determinantal function, more suitable the starting point. Of course, we do not change our definition of correlation energy, i.e. we consider the RHF energy as that which does not contain any correlation effects. For instance, we may ask which of the normalized single-determinant functions  $\Phi$  is closest to the normalized exact function  $\psi$ . As a measure of this we might use:

$$|\langle \psi | \Phi \rangle| = \text{maximum.} \quad (10.21)$$

The single determinantal function  $\Phi = \Phi_B$ , which fulfils the above condition, is called a Bruckner function (O. Sinanoğlu, K.A. Brueckner, “*Three Approaches to Electron Correlation in Atoms*”, Yale Univ. Press, New Haven and London, 1970).

function ( $\Phi_0$ , i.e.  $\psi_{\text{RHF}}$ ). When one spinorbital is replaced, the resulting determinant is called singly excited, when two – doubly excited, etc.<sup>46,47</sup>

The virtual spinorbitals form an orthonormal basis in the *virtual space* (Appendix B, p. 895). If we carry out any non-singular linear transformation (cf. p. 396) of virtual spinorbitals, each “new”  $n$ -tuply excited Slater determinant becomes a linear combination of all “old”  $n$ -tuply excited determinants and only  $n$ -tuply excited ones.<sup>48</sup> In particular, the unitary transformation preserves the mutual orthogonality of the  $n$ -tuply excited determinantal functions.

Thus, the total wave function (10.20) is a linear combination of the *known* Slater determinants (we assume that the spinorbitals are always known) with *unknown*  $c$  coefficients.

The name of the CI methods refers to the linear combination of the configurations rather than to the Slater determinants.

CSF  
(configuration)

A *configuration* (CSF, i.e. Configuration State Function) is a linear combination of determinants which is an eigenfunction of the operators:  $\hat{S}^2$  and  $\hat{S}_z$ , and belongs to the proper irreducible representation of the symmetry group of the Hamiltonian. We say that this is a linear combination of the (spatial and spin) symmetry adapted determinants. Sometimes we refer to the spin-adapted configurations which are eigenfunctions only of the  $\hat{S}^2$  and  $\hat{S}_z$  operators.

The particular terms in the CI expansion may refer to the respective CSFs or to the Slater determinants. Both versions lead to the same results, but using CSFs

<sup>46</sup>In the language of the second quantization (see Appendix U, p. 1023) the wave function in the CI method has the form (the  $\Phi_0$  function is a Slater determinant which does not necessarily need to be a Hartree–Fock determinant)

$$\psi = c_0 \Phi_0 + \sum_{a,p} c_p^a \hat{p}^\dagger \hat{a} \Phi_0 + \sum_{a < b, p < q} c_{pq}^{ab} \hat{q}^\dagger \hat{p}^\dagger \hat{a} \hat{b} \Phi_0 + \text{higher excitations}, \quad (10.22)$$

where  $c$  are the expansion coefficients, the creation operators  $\hat{q}^\dagger, \hat{p}^\dagger, \dots$  refer to the virtual spinorbitals  $\phi_p, \phi_q, \dots$  and the annihilation operators  $\hat{a}, \hat{b}, \dots$  refer to occupied spinorbitals  $\phi_a, \phi_b, \dots$  (the operators are denoted with the same indices as spinorbitals but the former are equipped with hat symbols), and the inequalities satisfied by the summation indices ensure that the given Slater determinant occurs only once in the expansion.

<sup>47</sup>The Hilbert space corresponding to  $N$  electrons is the sum of the orthogonal subspaces  $\Omega_n, n = 0, 1, 2, \dots, N$ , which are spanned by the  $n$ -tuply excited (orthonormal) Slater determinants. Elements of the space  $\Omega_n$  are all linear combinations of  $n$ -tuply excited Slater determinants. It does not mean, of course, that each element of this space is an  $n$ -tuply excited Slater determinant. For example, the sum of two doubly excited Slater determinants is a doubly excited Slater determinant only when one of the excitations is common to both determinants.

<sup>48</sup>Indeed, the Laplace expansion (Appendix A) along the row corresponding to the first new virtual spinorbital leads to the linear combination of the determinants containing new (*virtual, which means that the rank of excitation is not changed by this*) orbitals in this row. Continuing this procedure with the Slater determinants obtained, we finally get a linear combination of  $n$ -tuply excited Slater determinants expressed in old spinorbitals.

may be more efficient if we are looking for a wave function which transforms itself according to a single irreducible representation.

Next this problem is reduced to the Ritz method (see Appendices L, p. 984, and K, p. 982), and subsequently to the secular equations  $(\mathbf{H} - \epsilon \mathbf{S})\mathbf{c} = \mathbf{0}$ . It is worth noting here that, e.g., the CI wave function for the ground state of the helium atom would be linear combinations of the determinants where the largest  $c$  coefficient occurs in front of the  $\Phi_0$  determinant constructed from the spinorbitals  $1s\alpha$  and  $1s\beta$ , but a nonzero contribution would also come from the other determinants, e.g., constructed from the  $2s\alpha$  and  $2s\beta$  spinorbitals (one of the doubly excited determinants). The CI wave functions for all states (ground and excited) are linear combinations of *the same Slater determinants* – they differ only in the  $c$  coefficients.

The state energies obtained from the solution of the secular equations always approach the exact values from above.

### 10.10.1 BRILLOUIN THEOREM

In the CI method we have to calculate matrix elements  $H_{IJ}$  of the Hamiltonian. The Brillouin theorem says that:

$$\langle \Phi_0 | \hat{H} \Phi_1 \rangle = 0 \quad (10.23)$$

if  $\Phi_0$  is a solution of the Hartree–Fock problem ( $\Phi_0 \equiv \psi_{\text{RHF}}$ ), and  $\Phi_1$  is a singly excited Slater determinant in which the spinorbital  $\phi_{i'}$  is orthogonal to all spinorbitals used in  $\Phi_0$ .

Proof: From the II Slater–Condon rule (Appendix M, p. 986) we have:

$$\langle \Phi_0 | \hat{H} \Phi_1 \rangle = \langle i | \hat{h} i' \rangle + \sum_j [\langle ij | i' j \rangle - \langle ij | j i' \rangle]. \quad (10.24)$$

On the other hand, considering the integral  $\langle i | \hat{F} i' \rangle$ , where  $\hat{F}$  is a Fock operator, we obtain from (8.27) (using the definition of the Coulomb and exchange operators from p. 337):

$$\begin{aligned} \langle i | \hat{F} i' \rangle &= \langle i | \hat{h} i' \rangle + \sum_j [\langle i | \hat{J}_j i' \rangle - \langle i | \hat{K}_j i' \rangle] = \langle i | \hat{h} i' \rangle + \sum_j [\langle ij | i' j \rangle - \langle ij | j i' \rangle] \\ &= \langle \Phi_0 | \hat{H} \Phi_1 \rangle. \end{aligned}$$

From the Hermitian character of  $\hat{F}$  it follows that

$$\langle i | \hat{F} i' \rangle = \langle \hat{F} i | i' \rangle = \epsilon_i \delta_{ii'} = 0. \quad (10.25)$$

We have proved the theorem.

The Brillouin theorem is sometimes useful in discussion of the importance of particular terms in the CI expansion for the ground state.

### 10.10.2 CONVERGENCE OF THE CI EXPANSION

Increasing the number of expansion functions by adding a new function lowers or keeps unchanged the energy (due to the variational principle). It often happens

that the inclusion of only two determinants gives qualitative improvement with respect to the Hartree–Fock method, however when going further, the situation becomes more difficult. The convergence of the CI expansion is very slow, i.e. to achieve a good approximation to the wave function, the number of determinants in the expansion must usually be large. Theoretically, the shape of the wave function ensures solution of the Schrödinger equation  $H\psi = E\psi$ , but in practice we are *always* limited by the basis of the atomic orbitals employed. To obtain satisfactory results, we need to increase the number  $M$  of atomic orbitals in the basis. The number of molecular orbitals produced by the Hartree–Fock method is also equal to  $M$ , while the number of spinorbitals is equal to  $2M$ . In this case, the number of all determinants is equal to  $\binom{2M}{N}$ , where  $N$  refers to the number of electrons.

### 10.10.3 EXAMPLE OF H<sub>2</sub>O

We are interested in the ground state of the water molecule. This is a singlet state ( $S = 0$ ,  $M_S = 0$ ).

double dzeta

The minimal basis set, composed of 7 atomic orbitals (two 1s orbitals of the hydrogen atoms, 1s, 2s and three 2p orbitals of the oxygen atom), is considered too poor, therefore we prefer what is called double dzeta basis, which provides two functions with different exponents for each orbital of the minimal basis. This creates a basis of  $M = 14$  atomic orbitals. There are 10 electrons, hence  $\binom{28}{10}$  gives 13 123 110 determinants. For a matrix of that size to be diagonalized is certainly impressive. Even more impressive is that we achieve only an *approximation* to the correlation energy which amounts to about 50% of the exact correlation energy,<sup>49</sup> since  $M$  is only equal to 14, but in principle it should be equal to  $\infty$ . Nevertheless, for *comparative* purposes we assume the correlation energy obtained is 100%.

block form

The simplest remedy is to get rid of some determinants in such a way that the correlation energy is not damaged. Which ones? Well, many of them correspond to the incorrect projection  $S_z$  of the total spin. For instance, we are interested in the singlet state (i.e.  $S = 0$  and  $S_z = 0$ ), but some determinants are built of spinorbitals containing exclusively  $\alpha$  spin functions. This is a pure waste of resources, since the non-singlet functions do not make any contributions to the singlet state. When we remove these and other *incorrect* determinants, we obtain a smaller matrix to be diagonalized. The number of Slater determinants with  $S_z = 0$  is equal  $\binom{M}{N/2}^2$ . In our case, this makes slightly over 4 million determinants (instead of about 13 million). What would happen if we diagonalized the huge original matrix anyway? Well, nothing would happen. There would be more work, but the computer would create *the block form*<sup>50</sup> (see Appendix C) from our enormous matrix, and each block would correspond to the particular  $S^2$  and  $S_z$ , *while the whole contribution to the correlation energy of the ground state comes from the block corresponding to  $S = 0$  and  $S_z = 0$ .*

Let us continue throwing away determinants. This time, however, we have to make a compromise, i.e. some of the Slater determinants are arbitrarily consid-

<sup>49</sup>We see here how vicious the dragon of electron correlation is.

<sup>50</sup>These square blocks would be easily noticed after proper ordering of the expansion functions.

ered not to be important (which will worsen the results, if they are rejected). Which of the determinants should be considered as not important? The general opinion in quantum chemistry is that the multiple excitations are less and less important (when the multiplicity increases). If we take only the singly, doubly, triply and quadruply excited determinants, the number of determinants will reduce to about 25000 and we will obtain 99% of the approximate correlation energy defined above. If we take the singly and doubly excited determinants only, there are only 360 of them, and 94% of the correlation effect is obtained. This is why this CI SD CISD (*CI Singles and Doubles*) method is used so often.

For larger molecules this selection of determinants becomes too demanding, therefore we have to decide individually for each configuration: to include or reject it? The decision is made either on the basis of the perturbational estimate of the importance of the determinant<sup>51</sup> or by a test calculation with inclusion of the determinant in question, Fig. 10.5.

To obtain very good results, we need to include a large number of determinants, e.g., of the order of thousands, millions or even billions.<sup>52</sup> This means that contemporary quantum chemistry has made enormous technical progress.<sup>53</sup> This, however, is a sign, not of the strength of quantum chemistry, but of its weakness. What are we going to do with such a function? We may load it back into the computer and calculate all the properties of the system with high accuracy (although this cannot be guaranteed). To answer a student's question about why we obtained some particular numbers, we have to say that we do not know, it is the computer which knows. This is a trap. It would be better to get, say, two Slater determinants, which describe the system to a reasonable approximation and we can understand what is going on in the molecule.

#### 10.10.4 WHICH EXCITATIONS ARE MOST IMPORTANT?

The convergence can be particularly bad if we use the virtual spinorbitals obtained by the Hartree–Fock method. Not all excitations are equally important. It turns out that usually, although this is not a rule, low excitations dominate the ground state wave function.<sup>54</sup> The single excitations *themselves* do not contribute anything to the ground state *energy* (if the spinorbitals are generated with the Hartree–Fock

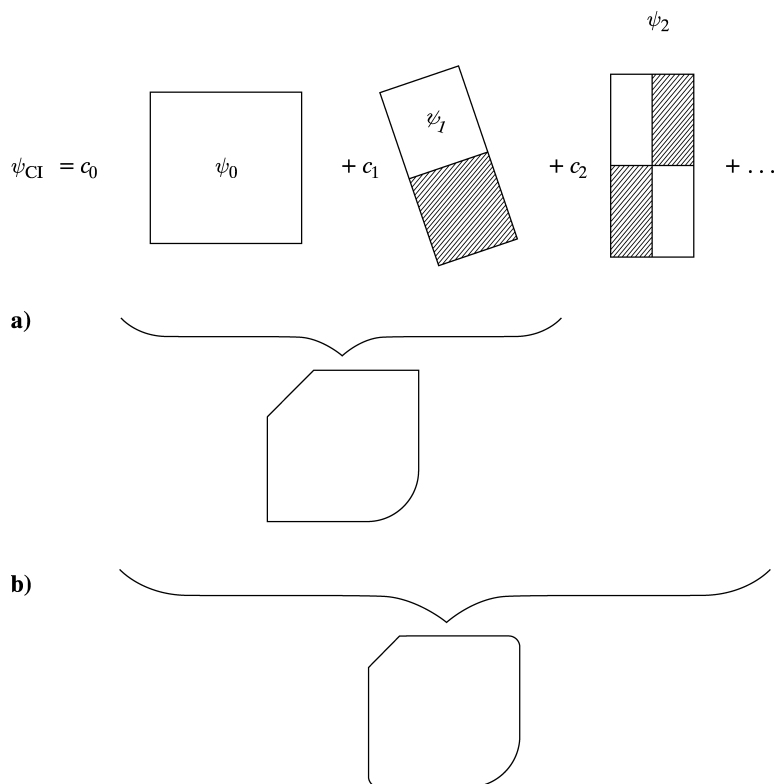
<sup>51</sup>The perturbational estimate mentioned relies on the calculation of the weight of the determinant based on the first order correction to the wave function in perturbation theory, see Chapter 5. In such an estimate the denominator contains the excitation energy evaluated as the difference in orbital energies between the Hartree–Fock determinant and the one in question. In the numerator there is a respective matrix element of the Hamiltonian calculated with the help of the known Slater–Condon rules (Appendix M, p. 986).

<sup>52</sup>Recently calculations with 3.6 billion Slater determinants were reported.

<sup>53</sup>To meet such needs, quantum chemists have had to develop entirely new techniques of applied mathematics.

<sup>54</sup>That is, requiring the lowest excitation energies. Later, a psychological mechanism began to work supported by economics: the *high energy* excitations are numerous and, because of that, very expensive and they correspond to high *excitations rank* (the number of electrons excited). Due to this, a reasonable restriction for the number of configurations in the CI expansion is excitation rank. We will come back to this problem later.





**Fig. 10.5.** Symbolic illustration of the principle of the CI method with one Slater determinant  $\psi_0$  dominant in the ground state (this is a problem of the many electron wave function so the picture cannot be understood literally). The purpose of this diagram is to emphasize a relatively small role of electronic correlation (more exactly, of what is known as the dynamical correlation, i.e. correlation of electronic motion). The function  $\psi_{CI}$  is a linear combination (the  $c$  coefficients) of the determinantal functions of different shapes in the many electron Hilbert space. The shaded regions correspond to the negative sign of the function; the nodal surfaces of the added functions allow for the effective deformation of  $\psi_0$  to have lower and lower average energy. (a) Since  $c_1$  is small in comparison to  $c_0$ , the result of the addition of the first two terms is a slightly deformed  $\psi_0$ . (b) Similarly the additional excitations just make cosmetic changes to the function (although they may substantially affect a quantity calculated with it).

method, then the Brillouin theorem mentioned above applies). *They are crucial, however, for excited states or in dipole moment calculations.* Only when coupled to other types of excitation do they assume non-zero (although small) values. Indeed, if in the CI expansion we only use the Hartree–Fock determinant and the determinants corresponding to single excitations, then, due to the Brillouin theorem, the secular determinant would be factorized.<sup>55</sup> This factorization (Fig. 10.6) pertains to the single-element determinant corresponding to the Hartree–Fock function

<sup>55</sup>That is, could be written out in the block form, which would separate the problem into two subproblems of smaller size.

	HF	S	D	T	Q	...
HF	$E_{\text{HF}}$	0 <sup>a</sup>	III	0 <sup>b</sup>	0 <sup>b</sup>	0 <sup>b</sup>
S	0 <sup>a</sup>	block S	II	III	0 <sup>b</sup>	0 <sup>b</sup>
D	III	II	block D	II	III	0 <sup>b</sup>
T	0 <sup>b</sup>	III	II	block T	II	III
Q	0 <sup>b</sup>	0 <sup>b</sup>	III	II	block Q	II
⋮	0 <sup>b</sup>	0 <sup>b</sup>	0 <sup>b</sup>	III	II	block ...

**Fig. 10.6.** The block structure of the Hamiltonian matrix ( $H$ ) is the result of the Slater–Condon rules (Appendix M, p. 986). S – single excitations, D – double excitations, T – triple excitations, Q – quadruple excitations. (a) Block of zero values due to the Brillouin theorem. (b) The block of zero values due to the IV Slater–Condon rule, (II) the non-zero block obtained according to II and III Slater–Condon rules, (III) the non-zero block obtained according to III Slater–Condon rule. All the non-zero blocks are sparse matrices dominated by zero values, which is important in the diagonalization process.

and to the determinants corresponding exclusively to single excitations. Since we are interested in the ground state, only the first determinant (Hartree–Fock) is of importance to us, and it does not change whether we include or not, a contribution coming from single excitations into the wave function.

Performing CI calculations with the inclusion of all excitations (for the assumed value of  $M$ ), i.e. the *full CI*, is not possible in practical calculations due to the too long expansion. We are forced to truncate the CI basis somewhere. It would be good to terminate it in such a way that all *essential* (the problem is what we mean by essential) terms are retained. *The most significant terms for the correlation energy come from the double excitations since these are the first excitations coupled to the Hartree–Fock function*, Fig. 10.6. Smaller, although important, contributions come from other excitations (usually of low excitation rank). We certainly wish that it would be like this for large molecules. Nobody knows what the truth is.

### 10.10.5 NATURAL ORBITALS (NO)

The fastest convergence is achieved in the basis set of *natural* orbitals (NO), i.e. when we construct spinorbitals with *these* orbitals and from them the Slater determinants. The NO is defined *a posteriori* in the following way. After carrying out the CI calculations, we construct the density matrix  $\rho$  (see Appendix S, p. 1015)

$$\begin{aligned}
 \rho(1, 1') &= \int \psi^*(1', 2, 3, \dots, N) \psi(1, 2, 3, \dots, N) d\tau_2 d\tau_3 \dots d\tau_N \\
 &= \sum_{ij} D_{ji} \phi_i^*(1') \phi_j(1), \quad D_{ij} = D_{ji}^*,
 \end{aligned}
 \tag{10.26}$$

where the summation runs over all the spinorbitals. By diagonalization of matrix  $D$  (a rotation in the Hilbert space spanned by the spinorbitals) we obtain the density matrix expressed in the natural spinorbitals (NO) transformed by the unitary transformation

$$\rho(1, 1') = \sum_i (D_{\text{diag}})_{ii} \phi_i'^*(1) \phi_i'(1'). \quad (10.27)$$

NO occupancies

pseudonatural  
orbitals (PNO)

The most important  $\phi_i'$  from the viewpoint of the correlation are the NOs with large *occupancies*, i.e.  $(D_{\text{diag}})_{ii}$  values. Inclusion of only the most important  $\phi_i'$  in the CI expansion creates a short and quite satisfactory wave function.<sup>56</sup> Meyer introduced the PNO CI, i.e. *pseudonatural orbitals*.<sup>57</sup> In the first step, we perform the CI calculations for excitations obtained by replacement of two selected spinorbitals. The process is repeated for all spinorbital pairs and at the end we carry out a “large” CI, which includes all important determinants engaged in the partial calculations (i.e. those with large weights).

### 10.10.6 SIZE CONSISTENCY

A truncated CI expansion has one unpleasant feature which affects the applicability of the method.

Let us imagine we want to calculate the interaction energy of two beryllium atoms. Let us suppose that we decide that to describe the beryllium atom we have to include, not only the  $1s^2 2s^2$  configuration, but also the doubly excited  $1s^2 2p^2$ . In the case of beryllium, this is a very reasonable step, since both configurations have similar energies. Let us assume now that we calculate the wave function for *two* beryllium atoms. If we want this function to describe the system correctly, also at large interatomic distances, we have to make sure that the departing atoms have appropriate excitations at their disposition, i.e. in our case  $1s^2 2p^2$ . To achieve this we *must incorporate quadruple excitations into the method*.<sup>58</sup>

If we include quadruples, we have a chance to achieve (an approximate) size consistency, i.e., the energy will be proportional to the number of atoms, otherwise our results will not be size consistent.

Let us imagine 10 beryllium atoms. In order to have size consistency we need to include 20-fold excitations. This would be very expensive. We clearly see that, for many systems, the size consistency requires inclusion of multiple excitations. If we carried out CI calculations for all possible (for a given number of spinorbitals) excitations, such a CI method (i.e. *full CI*) would be size consistent.

<sup>56</sup> Approximate natural orbitals can also be obtained directly without performing the CI calculations.

<sup>57</sup> R. Ahlrichs, W. Kutzelnigg, *J. Chem. Phys.* 48 (1968) 1819; W. Meyer, *Intern. J. Quantum Chem.* S5 (1971) 341.

<sup>58</sup> See J.A. Pople, R. Seeger, R. Krishnan, *Intern. J. Quantum Chem.* S11 (1977) 149, also p. 47 of the book by P. Jørgensen and J. Simons, “*Second Quantization-Based Methods in Quantum Chemistry*”, Academic Press, 1981.

## 10.11 DIRECT CI METHOD

We have already mentioned that the CI method converges slowly. Due to this, the Hamiltonian matrices and overlap integral matrices are sometimes so large that they cannot fit into the computer memory. In practice, such a situation occurs in all good quality calculations for small systems and in all calculations for medium and large systems. Even for quite large atomic orbital basis, the number of integrals is much smaller than the number of Slater determinants in the CI expansion.

Björn Roos<sup>59</sup> first noticed that to find the lowest eigenvalues and their eigenvectors we do not need to store a huge  $\mathbf{H}$  matrix in computer memory. Instead, we need to calculate the *residual vector*  $\boldsymbol{\sigma} = (\mathbf{H} - E\mathbf{1})\mathbf{c}$ , where  $\mathbf{c}$  is a trial vector (defining the trial function in the variational method, p. 196). If  $\boldsymbol{\sigma} = \mathbf{0}$ , it means that the solution is found. Knowing  $\boldsymbol{\sigma}$ , we may find (on the basis of first order perturbation theory) slightly improved  $\mathbf{c}$ , etc. The product  $\mathbf{H}\mathbf{c}$  can be obtained by going through the set of integrals and assigning to each a coefficient resulting from  $\mathbf{H}$  and  $\mathbf{c}$ , and adding the results to the new  $\mathbf{c}$  vector. Then the procedure is repeated. Until 1971, CI calculations with 5000 configurations were considered a significant achievement. After Roos's paper, there was a leap of several orders of magnitude, bringing the number of configurations to the range of billions. For the computational method this was a revolution.

residual vector

## 10.12 MULTIREFERENCE CI METHOD

Usually in the CI expansion, the dominant determinant is Hartree–Fock. We construct the CI expansion, replacing the spinorbitals in this determinant (*single reference method*). We can easily imagine a situation in which taking one determinant is not justified, since the shell is not well closed (e.g., four hydrogen atoms). We already know that certain determinants (or, in other words: configurations) absolutely need to be present (“*static correlation*”) in the correct wave function. To be sure, we are the judges, deciding which is good or bad. This set of determinants is a basis in the *model space*.

single reference method

static correlation

model space

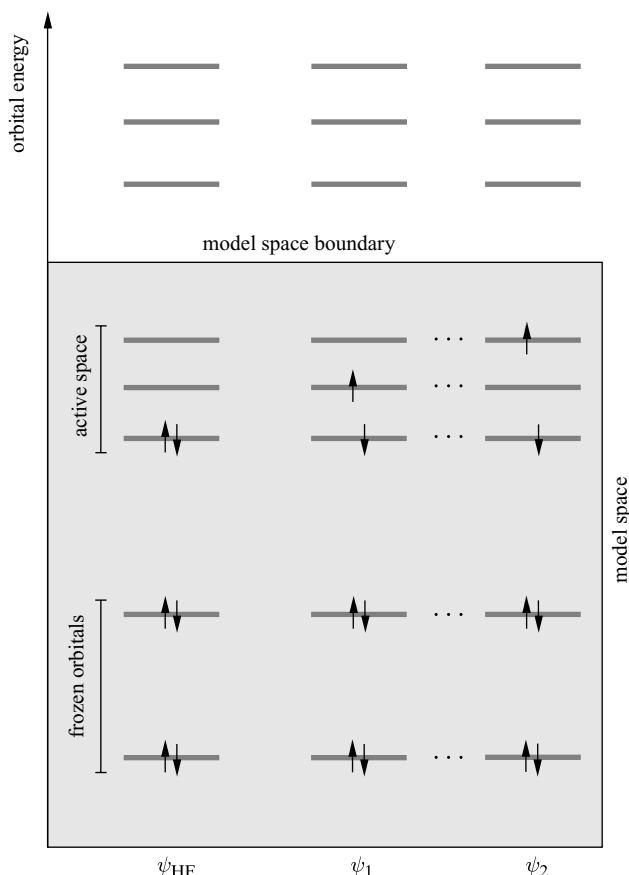
In the single reference CI method, the model space (Fig. 10.7) is formed by one Slater determinant. In the multireference CI method, the set of determinants constitute the model space. This time, the CI expansion is obtained by replacement of the spinorbitals participating in the model space by other virtual orbitals. We proceed further as in CI.

multireference method

There is no end to the problems yet, since, again we have billions of possible excitations.<sup>60</sup> We do other tricks to survive in this situation. We may, for instance,

<sup>59</sup>B.O. Roos, *Chem. Phys. Letters* 15 (1972) 153.

<sup>60</sup>There is another trouble known as *intruder states*, i.e. states which are of unexpectedly low energy. How can these states appear? Firstly, the CI states known as “front door intruders” appear, if some im-



**Fig. 10.7.** Illustration of the model space in the multireference CI method used mainly in the situation when no single Slater determinant dominates the CI expansion. In the figure the orbital levels of the system are presented. Part of them are occupied in all Slater determinants considered (“frozen spin-orbitals”). Above them is a region of closely spaced orbital levels called *active space*. In the optimal case, an energy gap occurs between the latter and unoccupied levels lying higher. The model space is spanned by all or some of the Slater determinants obtained by various occupancies of the active space levels.

get the idea not to excite the inner shell orbitals, since the numerical effort is serious, the lowering of the total energy can also be large, but the effect on the energy *differences* (this is what chemists are usually interested in) is negligible. We say that such orbitals are *frozen*. Some of the orbitals are kept doubly occupied in all Slater determinants but we optimize their shape. Such orbitals are called *inactive*. Finally, the orbitals of varied occupancy in different Slater determinants are called

portant (low-energy) configurations were for some reason not included into the model space. Secondly, we may have the “back door intruder” states. When the energy gaps between the model space and the other configurations are too small (quasi-degeneracy), some CI states became low energy states (enter the model space energy zone) even if they are not composed of the model space configurations.

*active*. The frozen orbitals are, in our method, important spectators of the drama, the inactive orbitals contribute a little towards lowering the energy, but the most efficient work is done by the active orbitals.

active orbitals

## 10.13 MULTICONFIGURATIONAL SELF-CONSISTENT FIELD METHOD (MC SCF)

In the configuration interaction method, it is sometimes obvious that certain determinants of the CI expansion *must* contribute to the wave function, if the latter is to correctly describe the system. For example, if we want to describe the system in which a bond is being broken (or is being formed), for its description we need several determinants for sure (cf. description of the dissociation of the hydrogen molecule on p. 371).

Why is this? In the case of dissociation, that we are dealing with here, there is a quasidegeneracy of the bonding and antibonding orbital of the bond in question, i.e. the approximate equality of their energies (the bond energy is of the order of the overlap integral and the latter goes to zero when the bond is being broken). The determinants, which can be constructed by various occupancies of these orbitals, have very similar energies and, consequently, their contributions to the total wave function are of similar magnitude and *should be included* in the wave function.

In the MC SCF method, as in CI, it is *up to us* to decide which set of determinants we consider sufficient for the description of the system.

Each of the determinants is constructed from molecular spinorbitals which are not fixed (as in the CI method) but are modified in such a way as to have the total energy as low as possible.

The MC SCF method is the most general scheme of the methods that use a linear combination of Slater determinants as an approximation to the wave function. In the limiting case of the MC SCF, when the number of determinants is equal to 1, we have, of course, the Hartree–Fock method.

### 10.13.1 CLASSICAL MC SCF APPROACH

We will describe first the classical MC SCF approach. This is a variational method. As was mentioned, the wave function in this method has the form of a finite linear combination of Slater determinants  $\Phi_I$

$$\psi = \sum_I d_I \Phi_I, \quad (10.28)$$

where  $d$  are variational coefficients.

In the classical MC SCF method we:

1. take a finite CI expansion (the Slater determinants and the orbitals for their construction are fixed)
2. calculate the coefficients for the determinants by the Ritz method (the orbitals do not change)
3. vary the LCAO coefficients in the orbitals at the fixed CI coefficients to obtain the best MOs
4. return to point 1 until self-consistency is achieved

### 10.13.2 UNITARY MC SCF METHOD

Another version of the MC SCF problem, a *unitary method* suggested by Lévy and Berthier<sup>61</sup> and later developed by Dalgaard and Jørgensen<sup>62</sup> is gaining increasing importance. The eigenproblem does not appear in this method.

We need two mathematical facts to present the unitary MC SCF method. The first is a theorem:

If  $\hat{A}$  is a Hermitian operator, i.e.  $\hat{A}^\dagger = \hat{A}$ , then  $\hat{U} = \exp(i\hat{A})$  is a unitary operator satisfying  $\hat{U}^\dagger \hat{U} = 1$ .

Let us see how  $\hat{U}^\dagger$  looks:

$$\begin{aligned}\hat{U}^\dagger &= (\exp(i\hat{A}))^\dagger = \left(1 + i\hat{A} + \frac{1}{2!}(i\hat{A})^2 + \frac{1}{3!}(i\hat{A})^3 + \dots\right)^\dagger \\ &= \left(1 + (-i)\hat{A}^\dagger + \frac{1}{2!}(-i\hat{A}^\dagger)^2 + \frac{1}{3!}(-i\hat{A}^\dagger)^3 + \dots\right) \\ &= \left(1 + (-i)\hat{A} + \frac{1}{2!}(-i\hat{A})^2 + \frac{1}{3!}(-i\hat{A})^3 + \dots\right) = \exp(-i\hat{A}).\end{aligned}$$

Hence,  $\hat{U}\hat{U}^\dagger = 1$ , i.e.  $\hat{U}$  is a unitary operator.<sup>63</sup>

<sup>61</sup>B. Lévy, G. Berthier, *Intern. J. Quantum Chem.* 2 (1968) 307.

<sup>62</sup>E. Dalgaard, P. Jørgensen, *J. Chem. Phys.* 69 (1978) 3833.

<sup>63</sup>Is an operator ( $\hat{C}$ ) of multiplication by a constant  $c$  Hermitian?

$$\langle \varphi | \hat{C} | \psi \rangle \stackrel{?}{=} \langle \hat{C} \varphi | \psi \rangle,$$

$$\text{l.h.s.} = \langle \varphi | c | \psi \rangle = c \langle \varphi | \psi \rangle,$$

$$\text{r.h.s.} = \langle c \varphi | \psi \rangle = c^* \langle \varphi | \psi \rangle.$$

Both sides are equal, if  $c = c^*$ . An operator conjugate to  $c$  is  $c^*$ .

Further:  $\hat{B} = i\hat{A}$ , what is a form of  $\hat{B}^\dagger$ ?

$$\langle \hat{B}^\dagger \varphi | \psi \rangle = \langle \varphi | \hat{B} | \psi \rangle,$$

$$\langle \varphi | i\hat{A} | \psi \rangle = \langle -i\hat{A}^\dagger \varphi | \psi \rangle,$$

$$\hat{B}^\dagger = -i\hat{A}^\dagger.$$

Now the second mathematical fact. This is a commutator expansion:

$$e^{-\hat{A}}\hat{H}e^{\hat{A}} = \hat{H} + [\hat{H}, \hat{A}] + \frac{1}{2!}[[\hat{H}, \hat{A}], \hat{A}] + \frac{1}{3!}[[[\hat{H}, \hat{A}], \hat{A}], \hat{A}] + \dots \quad (10.29)$$

This theorem can be proved by induction, expanding the exponential functions.

Now we are all set to describe the unitary method. We introduce two new operators:

$$\hat{\lambda} = \sum_{ij} \lambda_{ij} \hat{i}^\dagger \hat{j}, \quad (10.30)$$

where  $\hat{i}^\dagger$  and  $\hat{j}$  are the creation and annihilation operators, respectively, associated to spinorbitals  $i, j$ , see Appendix U. Further,

$$\hat{S} = \sum_{IJ} S_{IJ} |\Phi_I\rangle \langle \Phi_J|. \quad (10.31)$$

We assume that  $\lambda_{ij}$  and  $S_{IJ}$  are elements of the Hermitian matrices  $\lambda$  and  $S$  (their determination is the goal of the method),  $\Phi_I$  are determinants from the MC SCF expansion (10.28).

It can be seen that the  $\hat{\lambda}$  operator replaces a single spinorbital in a Slater determinant and forms a linear combination of such modified determinantal functions; the  $\hat{S}$  operator replaces such a combination with another. The “knobs” which control these changes are coefficients  $\lambda_{ij}$  and  $S_{IJ}$ .

We will need transformations  $\exp(i\hat{\lambda})$  and  $\exp(i\hat{S})$ . They are unitary, because  $\hat{\lambda}^\dagger = \hat{\lambda}$  and  $\hat{S}^\dagger = \hat{S}$ , i.e.  $\hat{\lambda}$  and  $\hat{S}$  are Hermitian.<sup>64</sup>

We suggest the form of our variational function:

$$|\tilde{0}\rangle = \exp(i\hat{\lambda}) \exp(i\hat{S}) |0\rangle \quad (10.32)$$

where  $|0\rangle$  denotes a starting combination of determinants with specific spinorbitals and the matrices  $\lambda$  and  $S$  contain the variational parameters as the matrix elements. So, we modify the spinorbitals and change the coefficients in front of the determinants to obtain a new combination of the modified determinants,  $|\tilde{0}\rangle$ . The mean energy value for that function is<sup>65</sup>

$$E = \langle \tilde{0} | \hat{H} | \tilde{0} \rangle = \langle 0 | \exp(-i\hat{S}) \exp(-i\hat{\lambda}) \hat{H} \exp(i\hat{\lambda}) \exp(i\hat{S}) | 0 \rangle. \quad (10.33)$$

Taking advantage of the commutator expansion (10.29), we have

$$\begin{aligned} E = & \langle 0 | \hat{H} | 0 \rangle - i \langle 0 | [\hat{S} + \hat{\lambda}, \hat{H}] | 0 \rangle + \frac{1}{2} \langle 0 | [\hat{S}, [\hat{H}, \hat{S}]] | 0 \rangle + \frac{1}{2} \langle 0 | [\hat{\lambda}, [\hat{H}, \hat{\lambda}]] | 0 \rangle \\ & + \langle 0 | [\hat{S}, [\hat{H}, \hat{\lambda}]] | 0 \rangle + \dots \end{aligned}$$

<sup>64</sup>Considering the matrix elements of the operators  $\hat{\lambda}$  and  $\hat{S}$ , we would easily be convinced that both operators are also Hermitian.

<sup>65</sup>Here we use the equality  $[\exp(i\hat{A})]^\dagger = \exp(-i\hat{A})$ .



It follows from the last equation, that in order to calculate  $E$ , we have to know the result of the operation of  $\hat{\lambda}$  on  $|0\rangle$ , i.e. on the linear combination of determinants, which comes down to the operation of the creation and annihilation operators on the determinants, which is simple. It can also be seen that we need to apply the operator  $\hat{S}$  to  $|0\rangle$ , but its definition shows that this is trivial. This expression<sup>66</sup> can now be optimized, i.e. the best Hermitian matrices  $\lambda$  and  $S$  can be selected. It is done in the same step (this distinguishes the current method from the classical one). Usually the calculations are carried out in a matrix form neglecting the higher terms and retaining only the quadratic ones in  $\hat{S}$  and  $\hat{\lambda}$ . Neglecting the higher terms is equivalent to allowing for very small rotations in the transformation (10.32), but instead we have a large number of rotations (iterative solution).<sup>67</sup>

The success of the method depends on the starting point. The latter strongly affects the energy and its hypersurface (in the space of the parameters of the matrices  $\lambda$  and  $S$ ) is very complicated, it has many local minima. This problem is not yet solved, but various procedures accelerating the convergence are applied, e.g., the new starting point is obtained by averaging the starting points of previous iterations. The method also has other problems, since the orbital rotations partially replace the rotation in the space of the Slater determinants (the rotations do not commute and are not independent). In consequence, linear dependencies may appear.

### 10.13.3 COMPLETE ACTIVE SPACE METHOD (CAS SCF)

An important special case of the MC SCF method is the CAS SCF (Complete Active Space Self-Consistent Field, Fig. 10.8) of Roos, Taylor and Siegbahn.<sup>68</sup> Let us assume that we are dealing with a closed-shell molecule. The RHF method (p. 342) provides the molecular orbitals and the orbital energies. From them we select the low energy orbitals. Part of them are *inactive*, i.e. are doubly occupied in all determinants, but they are varied, which results in lowering the mean value of the Hamiltonian (some of the orbitals may be frozen, i.e. kept unchanged). These are the spinorbitals corresponding to the inner shells. The remaining spinorbitals belong to the *active space*. Now we consider all possible occupancies and excitations of the active spinorbitals (this is where the adjective “*complete*” comes from) to obtain the set of determinants in the expansion of the MC SCF. By taking all possible

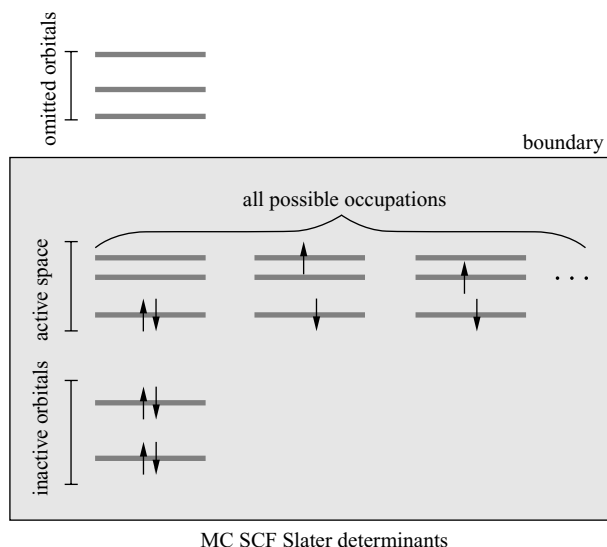
<sup>66</sup>The term with  $i$  gives a real number

$$i \cdot \langle 0 | [\hat{S} + \hat{\lambda}, \hat{H}] | 0 \rangle = i \cdot (\langle (\hat{S} + \hat{\lambda}) 0 | \hat{H} 0 \rangle - \langle \hat{H} 0 | (\hat{S} + \hat{\lambda}) 0 \rangle) \rightarrow i \cdot (z - z^*) = i(2i \operatorname{Im} z) \in R.$$

$R$  is a set of real numbers.

<sup>67</sup>In the classical MC SCF method when minimizing the energy with respect to the parameters, we use only linear terms in the expansion of the energy with respect to these parameters. In the unitary formulation, on the other hand, we use both linear and quadratic terms. This implies much better convergence of the unitary method.

<sup>68</sup>B.O. Roos, P.E.M. Siegbahn, in “*Modern Theoretical Chemistry*”, vol. III, ed. H.F. Schaefer, Plenum Press, New York, 1977; P.E.M. Siegbahn, *J. Chem. Phys.* 70 (1979) 5391; B.O. Roos, P.R. Taylor, P.E.M. Siegbahn, *Chem. Phys.* 48 (1980) 157.



**Fig. 10.8.** CAS SCF, a method of construction of the Slater determinants in the MC SCF expansion. The inner shell orbitals are usually inactive, i.e. are doubly occupied in each Slater determinant. Within the active space + inactive spinorbitals we create the complete set of possible Slater determinants to be used in the MC SCF calculations. The spinorbitals of energy higher than a certain selected threshold are entirely ignored in the calculations.

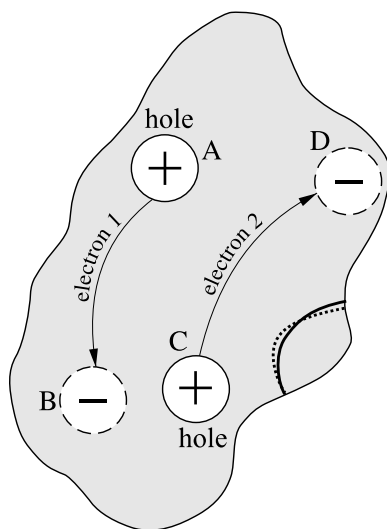
excitations within the active space, we achieve *a size consistency*, i.e. when dividing the system into subsystems and separating them (infinite distances) we obtain the sum of the energies calculated for each subsystem separately. By taking the complete set of excitations we also show that the results do not depend on any (non-singular) linear transformation of the molecular spinorbitals within the given subgroup of orbitals, i.e. within the inactive or active spinorbitals. This makes the result invariant with respect to the localization of the molecular orbitals.

## NON-VARIATIONAL METHODS WITH SLATER DETERMINANTS

### 10.14 COUPLED CLUSTER (CC) METHOD

The problem of a many-body correlation of motion of anything is extremely difficult and so far unresolved (e.g., weather forecasting). The problem of electron correlation also seemed to be hopelessly difficult. It still remains so, however, it turns out that we can exploit a certain observation made by Sinanoğlu.<sup>69</sup> This author noticed that the major portion of the correlation is taken into account through the in-

<sup>69</sup>O. Sinanoğlu and K.A. Brueckner, "Three Approaches to Electron Correlation in Atoms", Yale Univ. Press, New Haven and London, 1970.



**Fig. 10.9.** In order to include the electron correlation, the wave function should somehow reflect the fact that electrons avoid each other. Electron 1 jumping from A (an orbital) to B (another orbital) should make electron 2 escape from C (close to B) to D (close to A). This is the very essence of electron correlation. The other orbitals play a role of spectators. However, the spectators change upon the excitations described above. These changes are performed by allowing their own excitations. This is how triple, quadruple and higher excitations emerge and contribute to electronic correlation.

roducing of correlation within electron pairs, next through pair–pair interactions, then pair–pair–pair interactions, etc. The canonical molecular spinorbitals, which we can use, are in principle delocalized over the whole molecule, but in practice the delocalization is not so large. Even in the case of canonical spinorbitals, and certainly when using localized molecular spinorbitals, we can think about electron excitation as a transfer of an electron from one place in the molecule to another. Inclusion of the correlation of electronic motion represents, in the language of electron excitations, the following philosophy: when electron 1 jumps from an orbital localized in place A to an orbital localized in place B, it would be good – from the point of view of the variational principle – if electron 2 jumped from the orbital localized at C to the orbital localized at D (strong electrostatic stabilization), Fig. 10.9.

The importance of a given double excitation depends on the energy connected with the electron relocation and the arrangement of points A,B,C,D. Yet this simplistic reasoning suggests single excitations do not carry any correlation (this is confirmed by the Brillouin theorem) and this is why their role is very small. Moreover, it also suggests that double excitations should be very important.

The general idea of the coupled cluster method relies on the more and more accurate description of the many-electron system, beginning with the picture of the independent electrons, next of independent pairs, next of independent pair–pair sets, etc.

### 10.14.1 WAVE AND CLUSTER OPERATORS

At the beginning we introduce a special Slater determinant, the *reference determinant* (called the *vacuum state*, it can be the Hartree–Fock determinant)  $\Phi_0$  and we

write that the exact wave function for the ground state is

$$\psi = \exp(\hat{T})\Phi_0 \quad (10.34)$$

where  $\exp(\hat{T})$  is a *wave operator*, and  $\hat{T}$  itself is a *cluster operator*. In the CC method an *intermediate normalization*<sup>70</sup> of the function  $\psi$  is assumed, i.e.

wave and  
cluster  
operators

$$\langle \psi | \Phi_0 \rangle = 1.$$

Eq. (10.34) represents a very ambitious task. It assumes that we will find an operator  $\hat{T}$  such that the wave operator ( $e^{\hat{T}}$ ), as with the touch of a wizard's wand, will make an ideal solution of the Schrödinger equation from the Hartree–Fock function. The formula with  $\exp(\hat{T})$  is an *Ansatz*. The charming sounding word *Ansatz*<sup>71</sup> can be translated as arrangement or order, but in mathematics it refers to the construction assumed.

intermediate  
normalization

In literature we use the argument that the wave operator ensures the size consistency of the CC. According to this reasoning, for an infinite distance between molecules  $A$  and  $B$ , both  $\psi$  and  $\Phi_0$  functions can be expressed in the form of the product of the wave functions for  $A$  and  $B$ . When the cluster operator is assumed to be of the form (obvious for infinitely separated systems)  $\hat{T} = \hat{T}_A + \hat{T}_B$ , then the exponential form of the wave operator  $\exp(\hat{T}_A + \hat{T}_B)$  ensures a desired form of the product of the wave function  $[\exp(\hat{T}_A + \hat{T}_B)]\Phi_0 = \exp \hat{T}_A \exp \hat{T}_B \Phi_0$ . *If we took a finite CI expansion:  $(\hat{T}_A + \hat{T}_B)\Phi_0$ , then we would not get the product but the sum which is incorrect.* With this reasoning there is a problem, since due to the Pauli principle (antisymmetry of the wave function with respect to the electron exchange) for long distance neither the function  $\psi$  nor the function  $\Phi_0$  are the product of the functions for the subsystems.<sup>72</sup> Although the reasoning is not quite correct, the conclusion is correct, as will be shown at the end of the description of the CC method (p. 547).

The CC method is automatically size consistent.

As a cluster operator  $\hat{T}$  we assume a sum of the excitation operators (see Appendix U)

$$\hat{T} = \hat{T}_1 + \hat{T}_2 + \hat{T}_3 + \cdots + \hat{T}_{l_{\max}} \quad (10.35)$$

where

$$\hat{T}_1 = \sum_{a,r} t_a^r \hat{r}^\dagger \hat{a} \quad (10.36)$$

<sup>70</sup>It contributes significantly to the numerical efficiency of the method.

<sup>71</sup>This word has survived in the literature in its original German form.

<sup>72</sup>For instance, the RHF function for the hydrogen molecule is not a product function for long distances, see p. 520.

is an operator for single excitations,

$$\hat{T}_2 = \frac{1}{4} \sum_{\substack{ab \\ rs}} t_{ab}^{rs} s^{\dagger} \hat{r}^{\dagger} \hat{a} \hat{b}, \quad (10.37)$$

is an operator for double excitations, etc. The subscript  $l$  in  $\hat{T}_l$  indicates the rank of the excitations involved (with respect to the vacuum state). The symbols  $a, b, \dots$  refer to the spinorbitals occupied in  $\Phi_0$ , and  $p, q, r, s, \dots$  refer to the unoccupied ones, and

$t$  represents *amplitudes*, i.e. the numbers whose determination is the goal of the CC method. The rest of this chapter will be devoted to the problem of how we can obtain these miraculous amplitudes.

*In the CC method we want to obtain correct results with the assumption that  $l_{\max}$  of eq. (10.35) is relatively small (usually  $2 \div 5$ ). If  $l_{\max}$  were equal to  $N$ , i.e. to the number of electrons, then the CC method would be identical to the full (usually unfeasible) CI method.*

### 10.14.2 RELATIONSHIP BETWEEN CI AND CC METHODS

Obviously, there is a relation between the CI and CC methods. For instance, if we write  $\exp(\hat{T})\Phi_0$  in such a way as to resemble the CI expansion

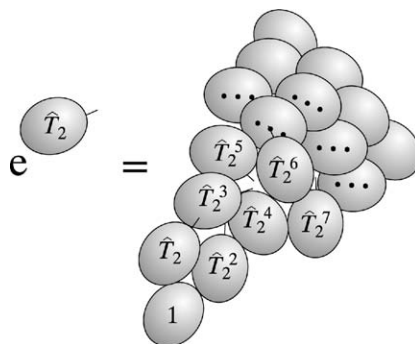
$$\begin{aligned} \exp(\hat{T})\Phi_0 &= \left[ 1 + (\hat{T}_1 + \hat{T}_2 + \hat{T}_3 + \dots) + \frac{1}{2}(\hat{T}_1 + \hat{T}_2 + \hat{T}_3 + \dots)^2 + \dots \right] \Phi_0 \\ &= (1 + \hat{C}_1 + \hat{C}_2 + \hat{C}_3 + \dots) \Phi_0, \end{aligned} \quad (10.38)$$

the operators  $\hat{C}_i$ , pertaining to the CI method, have the following structure

$$\begin{aligned} \hat{C}_1 &= \hat{T}_1, \\ \hat{C}_2 &= \hat{T}_2 + \frac{1}{2!} \hat{T}_1^2, \\ \hat{C}_3 &= \hat{T}_3 + \frac{1}{3!} \hat{T}_1^3 + \hat{T}_1 \hat{T}_2, \\ \hat{C}_4 &= \hat{T}_4 + \frac{1}{4!} \hat{T}_1^4 + \frac{1}{2!} \hat{T}_2^2 + \hat{T}_3 \hat{T}_1 + \frac{1}{2!} \hat{T}_1^2 \hat{T}_2, \\ &\dots \end{aligned} \quad (10.39)$$

We see that the multiple excitations  $\hat{C}_l$  result from mathematically distinct terms, e.g.,  $\hat{C}_3$  is composed of  $\hat{T}_3$ ,  $\hat{T}_1^3$  and  $\hat{T}_1 \hat{T}_2$ . Sometimes we speak about the *factorizable part* of the CI coefficient (like  $\hat{T}_1^3$  and  $\hat{T}_1 \hat{T}_2$ ) multiplying the particular Slater deter-

factorizable part  
of CI coefficient



**Fig. 10.10.** Why such a name? An artistic impression on *coupled clusters*.

minant (corresponding to an  $n$ -tuple excitation) as the part which can be expressed in terms of the *lower* rank amplitudes.

On the basis of current numerical experience,<sup>73</sup> we believe that, within the excitation of a given rank, the contributions coming from the correlational interactions of the electron pairs are the most important, e.g., within  $C_4$  the  $\frac{1}{2!}\hat{T}_2^2$  excitations containing the product of amplitudes for two electron pairs are the most important,  $\hat{T}_4$  (which contains the amplitudes of quadruple excitations) is of little importance, since they correspond to the coupling of the motions of four electrons, the terms  $\hat{T}_1$ ,  $\hat{T}_3\hat{T}_1$  and  $\hat{T}_1^2\hat{T}_2$  can be made small by using the MC SCF orbitals. Contemporary quantum chemists use diagrammatic language following Richard Feynman. The point is that the mathematical terms (the energy contributions) appearing in CC theory can be translated – one by one – into the figures according to certain rules. It turns out that it is much easier (at least at lower orders) to think in terms of diagrams than to speak about the mathematical formulae or to write them out.

The coupled cluster method, terminated at  $\hat{T}_2$  in the cluster operator automatically includes  $\hat{T}_2^2$ , etc. We may see in it some resemblance to a group of something (excitations), or in other words to a cluster, Fig. 10.10.

### 10.14.3 SOLUTION OF THE CC EQUATIONS

The strategy of the CC method is the following: first, we make a decision with respect to  $l_{\max}$  in the cluster expansion (10.35) ( $l_{\max}$  should be small<sup>74</sup>).

The exact wave function  $\exp(\hat{T})\Phi_0$  satisfies the Schrödinger equation, i.e.

$$\hat{H} \exp(\hat{T})\Phi_0 = E \exp(\hat{T})\Phi_0, \quad (10.40)$$

which, after operating from the left with  $\exp(-\hat{T})$  gives:

$$\exp(-\hat{T})\hat{H} \exp(\hat{T})\Phi_0 = E\Phi_0. \quad (10.41)$$

<sup>73</sup>This is a contribution by Oktay Sinanoğlu; O. Sinanoğlu, K.A. Brueckner (eds.), “*Three Approaches to Electron Correlation in Atoms*”, Yale Univ. Press, New Haven and London, 1970.

<sup>74</sup>Only then is the method cost-effective.

The  $\exp(-\hat{T})\hat{H}\exp(\hat{T})$  operator can be expressed in terms of the commutators (see (10.29))<sup>75</sup>

$$\begin{aligned} e^{-\hat{T}}\hat{H}e^{\hat{T}} &= \hat{H} + [\hat{H}, \hat{T}] + \frac{1}{2!}[[\hat{H}, \hat{T}], \hat{T}] + \frac{1}{3!}[[[\hat{H}, \hat{T}], \hat{T}], \hat{T}] \\ &\quad + \frac{1}{4!}[[[[\hat{H}, \hat{T}], \hat{T}], \hat{T}], \hat{T}]. \end{aligned} \quad (10.42)$$

The expansion (10.42) is *finite* (justification can be only diagrammatic, and is not given here) since in the Hamiltonian  $\hat{H}$  we have only two-particle interactions. Substituting this into the Schrödinger equation we have:

$$\begin{aligned} &\left\{ \hat{H} + [\hat{H}, \hat{T}] + \frac{1}{2!}[[\hat{H}, \hat{T}], \hat{T}] + \frac{1}{3!}[[[\hat{H}, \hat{T}], \hat{T}], \hat{T}] + \frac{1}{4!}[[[[\hat{H}, \hat{T}], \hat{T}], \hat{T}], \hat{T}] \right\} \Phi_0 \\ &= E\Phi_0. \end{aligned} \quad (10.43)$$

Multiplying from the left with the function  $\langle_{ab...}^{mn...}|$  representing the determinant obtained from the vacuum state by the action of the annihilators  $\hat{a}, \hat{b}, \dots$  and creators  $\hat{n}^\dagger, \hat{m}^\dagger, \dots$  and integrating, we obtain one equation for each function used:

$$\begin{aligned} &\langle_{ab...}^{mn...} | \hat{H} + [\hat{H}, \hat{T}] + \frac{1}{2!}[[\hat{H}, \hat{T}], \hat{T}] + \frac{1}{3!}[[[\hat{H}, \hat{T}], \hat{T}], \hat{T}] \\ &\quad + \frac{1}{4!}[[[[\hat{H}, \hat{T}], \hat{T}], \hat{T}], \hat{T}] | \Phi_0 \rangle = 0, \end{aligned} \quad (10.44)$$

where we have zero on the right-hand side due to the orthogonality. The Slater determinants  $\langle_{ab...}^{mn...}$  represent all excitations from  $\Phi_0$  resulting from the given cluster expansion  $\hat{T} = \hat{T}_1 + \hat{T}_2 + \dots + \hat{T}_{l_{\max}}$ . This is the fundamental equation of the CC method. For such a set of excited configurations the number of CC equations is equal to the number of the amplitudes sought.

$t_{ab...}^{mn...}$  are unknown quantities, i.e. amplitudes determining the  $\hat{T}_l$ , and, consequently, the wave operator (10.34) and wave function for the ground state  $\Psi = \Psi_0$ . The equations we get in the CC method are nonlinear

since the  $t$ 's occur at higher powers than the first (it can be seen from eq. (10.44) that the highest power of  $t$  is 4), which, on one hand, requires much more demanding and capricious (than linear ones) numerical procedures, and, on the other, contributes to the greater efficiency of the method. The number of such equations very

<sup>75</sup>It is straightforward to demonstrate the correctness of the first few terms by expanding the wave operator in the Taylor series.

often exceeds 100 000 or a million.<sup>76</sup> These equations are solved iteratively assuming certain starting amplitudes  $t$  and iterating the equations until self-consistency is achieved.

We hope that in such a procedure an approximation to the ground state wave function is obtained, although sometimes an unfortunate starting point may lead to some excited state.<sup>77</sup>

We usually use as a starting point that which is obtained from the linear version (reduced to obtain a linearity) of the CC method. We will write down these equations as  $t_{ab}^{mn} = \dots$  various powers of all  $t$  for all amplitudes. First we neglect the non-linear terms, this represents the initial approximation. The amplitudes are substituted into the right-hand side and we iterate until self-consistency is achieved. When all the amplitudes are found, then we obtain the energy  $E$  by projecting eq. (10.44) against  $\Phi_0$  function instead of  $|_{ab}^{mn}$ :

$$E = \langle \Phi_0 | e^{-\hat{T}} \hat{H} e^{\hat{T}} \Phi_0 \rangle. \quad (10.45)$$

The operator  $(e^{-T})^\dagger$ , conjugate to  $e^{-T}$ , is  $e^{-T^\dagger}$ , i.e.

$$E = \langle e^{-\hat{T}^\dagger} \Phi_0 | \hat{H} e^{\hat{T}} \Phi_0 \rangle, \quad (10.46)$$

which is *not* the mean value of the Hamiltonian. Hence, the CC method is not variational. If we multiplied eq. (10.40) from the left by  $e^{\hat{T}^\dagger}$  we would obtain the variational character of  $E$

$$E = \frac{\langle \Phi_0 | e^{\hat{T}^\dagger} \hat{H} e^{\hat{T}} \Phi_0 \rangle}{\langle \Phi_0 | e^{\hat{T}^\dagger} e^{\hat{T}} \Phi_0 \rangle} = \frac{\langle e^{\hat{T}} \Phi_0 | \hat{H} | e^{\hat{T}} \Phi_0 \rangle}{\langle e^{\hat{T}} \Phi_0 | e^{\hat{T}} \Phi_0 \rangle}. \quad (10.47)$$

However, it would not be possible to apply the commutator expansion and instead of the four terms in eq. (10.42) we would have an infinite number (due to the full normalization of the final function<sup>78</sup>). For this reason, we prefer the non-variational approach.

#### 10.14.4 EXAMPLE: CC WITH DOUBLE EXCITATIONS

How does the CC machinery work? Let us show it for a relatively simple case  $\hat{T} = \hat{T}_2$ . Eq. (10.44), written without the commutator expansion, has the form

$$\langle_{ab}^{mn} | e^{-\hat{T}_2} \hat{H} e^{\hat{T}_2} \Phi_0 \rangle = 0. \quad (10.48)$$

<sup>76</sup>This refers to calculations with  $\hat{T} = \hat{T}_2$  for ca. 10 occupied orbitals (for instance, two water molecules) and 150 virtual orbitals. These are not calculations for large systems.

<sup>77</sup>The first complete analysis of all CC solutions was performed by K. Jankowski and K. Kowalski, *Phys. Rev. Letters* 81 (1998) 1195; *J. Chem. Phys.* 110 (1999) 37, 93; *ibid.* 111 (1999) 2940, 2952. Recapitulation can be found in K. Jankowski, K. Kowalski, I. Grabowski, H.J. Monkhorst, *Intern. J. Quantum Chem.* 95 (1999) 483.

<sup>78</sup>The (non-variational) CC method benefits from the very economical condition of the intermediate normalization.



Taking advantage of the commutator expansion we have

$$\begin{aligned}\langle_{ab}^{mn}|e^{-\hat{T}_2}\hat{H}e^{\hat{T}_2}\Phi_0\rangle &= \langle_{ab}^{mn}|\left(1 - \hat{T}_2 + \frac{1}{2}\hat{T}_2^2 + \dots\right)\hat{H}\left(1 + \hat{T}_2 + \frac{1}{2}\hat{T}_2^2 + \dots\right)\Phi_0\rangle \\ &= \langle_{ab}^{mn}|\hat{H}\Phi_0\rangle + \langle_{ab}^{mn}|\hat{H}\hat{T}_2\Phi_0\rangle + \frac{1}{2}\langle_{ab}^{mn}|\hat{H}\hat{T}_2^2\Phi_0\rangle - \langle_{ab}^{mn}|\hat{T}_2\hat{H}\Phi_0\rangle \\ &\quad - \langle_{ab}^{mn}|\hat{T}_2\hat{H}\hat{T}_2\Phi_0\rangle + A = 0.\end{aligned}$$

However,

$$\begin{aligned}A &= -\frac{1}{2}\langle_{ab}^{mn}|\hat{T}_2\hat{H}\hat{T}_2^2\Phi_0\rangle + \frac{1}{2}\langle_{ab}^{mn}|\hat{T}_2^2\hat{H}\Phi_0\rangle + \frac{1}{2}\langle_{ab}^{mn}|\hat{T}_2^2\hat{H}\hat{T}_2\Phi_0\rangle + \frac{1}{4}\langle_{ab}^{mn}|\hat{T}_2^2\hat{H}\hat{T}_2^2\Phi_0\rangle \\ &= 0.\end{aligned}$$

The last equality follows from the fact that each term is equal to zero. The first vanishes since both determinants differ by four excitations. Indeed,  $\langle(\hat{T}_2^\dagger)_{ab}^{mn}|$  denotes a double *deexcitation*<sup>79</sup> of the doubly excited function, i.e. something proportional to  $\langle\Phi_0|$ . For similar reasons (too strong deexcitations give zero) the remaining terms in  $A$  also vanish. As a result we need to solve the equation:

$$\langle_{ab}^{mn}|\hat{H}\Phi_0\rangle + \langle_{ab}^{mn}|\hat{H}\hat{T}_2\Phi_0\rangle + \frac{1}{2}\langle_{ab}^{mn}|\hat{H}\hat{T}_2^2\Phi_0\rangle - \langle_{ab}^{mn}|\hat{T}_2\hat{H}\Phi_0\rangle - \langle_{ab}^{mn}|\hat{T}_2\hat{H}\hat{T}_2\Phi_0\rangle = 0.$$

After *several days*<sup>80</sup> of algebraic manipulations, we get the equations for the  $t$  amplitudes (for each  $t_{ab}^{mn}$  amplitude one equation):

$$\begin{aligned}(\varepsilon_m + \varepsilon_n - \varepsilon_a - \varepsilon_b)t_{ab}^{mn} &= \langle mn|ab\rangle - \sum_{p>q}\langle mn|pq\rangle t_{ab}^{pq} - \sum_{\gamma>\delta}\langle cd|ab\rangle t_{cd}^{mn} \\ &\quad + \sum_{c,p}[\langle cn|bp\rangle t_{ac}^{mp} - \langle cm|bp\rangle t_{ac}^{np} - \langle cn|ap\rangle t_{bc}^{mp} + \langle cm|ap\rangle t_{bc}^{np}] \\ &\quad + \sum_{c>d, p>q}\langle cd|pq\rangle[t_{ab}^{pq}t_{cd}^{mn} - 2(t_{ab}^{mp}t_{cd}^{nq} + t_{ab}^{nq}t_{cd}^{mp}) \\ &\quad - 2(t_{ac}^{mn}t_{bd}^{pq} + t_{ac}^{pq}t_{bd}^{mn}) + 4(t_{ac}^{mp}t_{bd}^{nq} + t_{ac}^{nq}t_{bd}^{mp})].\end{aligned}\tag{10.49}$$

It can be seen that the last expression includes: the term independent of  $t$ , the linear terms and the quadratic terms.

How can we find the  $t$ 's? We do it with the help of the iterative method.<sup>81</sup> First, we substitute zeros for all  $t$ 's on the right-hand side of the equation. Thus, from

<sup>79</sup>Opposite to excitation.

<sup>80</sup>Students – more courage!

<sup>81</sup>We organize things in such a way that a given unknown parameter will occur in the simple form on one side of the equation, whereas the more complicated terms, also containing the parameter sought,

the left-hand side the first approximation to  $t_{ab}^{mn}$  is<sup>82</sup>  $t_{ab}^{mn} \cong \frac{\langle mn|ab \rangle}{(\varepsilon_m + \varepsilon_n - \varepsilon_a - \varepsilon_b)}$ . We have now an estimate of each amplitude – we are making progress. The approximation to  $t$  obtained in this way is substituted into the right-hand side to evaluate the left-hand side and so forth. Finally, we achieve a self-consistency of the iterative process and obtain the CC wave function for the ground state of our system. With the amplitudes we calculate the energy of the system with eq. (10.45).

This is how the CCD (the CC with double excitations in the cluster operator) works from the practical viewpoint. It is more efficient when the initial amplitudes are taken from a short CI expansion,<sup>83</sup> with subsequent linearization (as above) of terms containing the initial (known) amplitudes.

The computational cost of the CCD and CCSD (singles and doubles) methods scales as  $N^6$  where  $N$  is a number of molecular orbitals (occupied and virtual<sup>84</sup>), whereas the analogous cost of the CCSDT (singles, doubles, triples) method requires  $N^8$  scaling. This means that, if we increase the orbital basis twice, the increase in the computational cost of the CCSDT method will be four times larger than that of the CCSD scheme. This is a lot, and because of this, wide-spread popularity has been gained for the CCSD(T) method, which only partly uses the triple excitations.

### 10.14.5 SIZE CONSISTENCY OF THE CC METHOD

The size consistency of the CC method can be proved on the basis of eq. (10.43) and (10.44). Let us assume that the system dissociates into two<sup>85</sup> non-interacting subsystems  $A$  and  $B$  (i.e. at infinite distance). Then the orbitals can be also divided into two separable (mutually orthogonal) subsets. We will show<sup>86</sup> that the cluster amplitudes, having *mixed* indices (from the first and second groups of orbitals), are equal to 0.

Let us note first that, for an infinite distance, the Hamiltonian  $\hat{H} = \hat{H}_A + \hat{H}_B$ . In such a situation the wave operator can be expressed as

$$\hat{T} = \hat{T}_A + \hat{T}_B + \hat{T}_{AB}, \quad (10.50)$$

are kept on the other side of the equation. Then we take a certain random value of the unknown and calculate the complicated side of the equation. On the left-hand side we will then have its *new* approximation to the true value. We repeat the whole procedure so many times until we start getting the same value (if the procedure converges). Then the equation is solved.

<sup>82</sup>As we see we would have trouble if  $(\varepsilon_m + \varepsilon_n - \varepsilon_a - \varepsilon_b)$  is close to 0 (quasidegeneracy of the vacuum state with some other state), because then  $t_{ab}^{mn} \rightarrow \infty$ .

<sup>83</sup>The configuration interaction method with inclusion of single and double excitations only:

CCD: J.A. Pople, R. Krishnan, H.B. Schlegel, J.S. Binkley, *Intern. J. Quantum Chem.* S14 (1978) 545; R.J. Bartlett, G.D. Purvis III, *Intern. J. Quantum Chem.* S14 (1978) 561.

CCSD: G.D. Purvis III, *J. Chem. Phys.* 76 (1982) 1910.

<sup>84</sup>These estimations are valid for the same relative increase of the number of occupied and virtual orbitals, as it is, e.g., for going from a molecule to its dimer. In the case of calculations for the same molecule, but two atomic basis sets (that differ in size) the cost increases only as  $N^4$ .

<sup>85</sup>This can be generalized to many non-interacting subsystems.

<sup>86</sup>B. Jeziorski, J. Paldus, P. Jankowski, *Intern. J. Quantum Chem.* 56 (1995) 129.

where  $\hat{T}_A, \hat{T}_B, \hat{T}_{AB}$  include the operators corresponding to spinorbitals from the subsystems  $A, B$  and from the system  $AB$ , respectively. Of course, in this situation we have the following commutation condition:

$$[\hat{H}_A, \hat{T}_B] = [\hat{H}_B, \hat{T}_A] = 0. \quad (10.51)$$

Then, owing to the commutator expansion in eq. (10.42), we obtain:

$$e^{-\hat{T}}(\hat{H}_A + \hat{H}_B)e^{\hat{T}} = e^{-\hat{T}_A}\hat{H}_Ae^{\hat{T}_A} + e^{-\hat{T}_B}\hat{H}_Be^{\hat{T}_B} + O(\hat{T}_{AB}), \quad (10.52)$$

where  $O(\hat{T}_{AB})$  denotes the linear and higher terms in  $\hat{T}_{AB}$ . Substituting this into eq. 10.44 with *bra*  $\langle \text{mixed} |$  vector representing mixed excitation, we observe that the first two terms on the right-hand side of the last equation give zero. It means that we get the equation

$$\langle \text{mixed} | O(\hat{T}_{AB}) \Phi_0 \rangle = 0, \quad (10.53)$$

which, due to the linear term in  $O(\hat{T}_{AB})$ , is fulfilled by  $\hat{T}_{AB} = 0$ . Conclusion: for the infinite distance between the subsystems we do not have mixed amplitudes and the energy of the  $AB$  system is bound to be the sum of the energies of subsystem  $A$  and subsystem  $B$  (size consistency).

## 10.15 EQUATION-OF-MOTION METHOD (EOM-CC)

The CC method is used to calculate the *ground* state energy and wave function. What about the excited states? This is a task for the equation-of-motion CC method, the primary goal being not the excited states themselves, but the excitation energies with respect to the ground state.

### 10.15.1 SIMILARITY TRANSFORMATION

Let us note that for the Schrödinger equation  $\hat{H}\psi = E\psi$ , we can perform an interesting sequence of transformations based on the wave operator  $e^{\hat{T}}$ :

$$\begin{aligned} e^{-\hat{T}}\hat{H}\psi &= Ee^{-\hat{T}}\psi, \\ e^{-\hat{T}}\hat{H}e^{\hat{T}}e^{-\hat{T}}\psi &= Ee^{-\hat{T}}\psi. \end{aligned}$$

We obtain the eigenvalue equation again, but for the *similarity transformed Hamiltonian*

$$\hat{\mathcal{H}}\bar{\psi} = E\bar{\psi},$$

where  $\hat{\mathcal{H}} = e^{-\hat{T}}\hat{H}e^{\hat{T}}$ ,  $\bar{\psi} = e^{-\hat{T}}\psi$ , and the energy  $E$  does not change at all after this transformation. This result will be very useful in a moment.

### 10.15.2 DERIVATION OF THE EOM-CC EQUATIONS

As the reference function in the EOM-CC method, we take the coupled-cluster wave function for the ground state:

$$\psi_0 = \exp(\hat{T})\Phi_0, \quad (10.54)$$

where  $\Phi_0$  is usually a Hartree–Fock determinant. Now, we define the operator  $\hat{U}_k$  (“EOM-CC Ansatz”), which performs a miracle: from the wave function of the ground state  $\psi_0$  it creates the wave function  $\psi_k$  for the  $k$ -th excited state of the system:

$$\psi_k = \hat{U}_k \psi_0.$$

The operators  $\hat{U}_k$  change the coefficients in front of the configurations (see p. 526). The operators  $\hat{U}_k$  are (unlike the wave operator  $\exp(\hat{T})$ ) linear with respect to the excitations, i.e. the excitation amplitudes occur there in the first powers. For the case of the single and double excitations (EOM-CCSD) we have  $\hat{T}$  in the form of the sum of single and double excitations:

$$\hat{T} = \hat{T}_1 + \hat{T}_2$$

and

$$\hat{U}_k = \hat{U}_{k,0} + \hat{U}_{k,1} + \hat{U}_{k,2},$$

where the task for the  $\hat{U}_{k,0}$  operator is to change the coefficient in front of the function  $\Phi_0$  to that appropriate to the  $|k\rangle$  function. The role of the operators  $\hat{U}_{k,1}$ ,  $\hat{U}_{k,2}$  is an appropriate modification of the coefficients in front of the singly and doubly excited configurations. These tasks are done by the excitation operators with  $\tau$  amplitudes (they have to be distinguished from the  $t$  amplitudes of the CC method):

$$\begin{aligned} \hat{U}_{k,0} &= \tau_0(k), \\ \hat{U}_{k,1} &= \sum_{a,p} \tau_a^p(k) \hat{p}^\dagger \hat{a}, \\ \hat{U}_{k,2} &= \sum_{a,b,p,q} \tau_{ab}^{pq}(k) \hat{q}^\dagger \hat{p}^\dagger \hat{a} \hat{b}, \end{aligned}$$

where the amplitudes  $\tau(k)$  are numbers, which are the targets of the EOM-CC method. The amplitudes give the wave function  $\psi_k$  and the energy  $E_k$ .

We write down the Schrödinger equation for the excited state:

$$\hat{H}\psi_k = E_k\psi_k.$$

Now we substitute the EOM-CC *Ansatz*:

$$\hat{H}\hat{U}_k\psi_0 = E_k\hat{U}_k\psi_0,$$

and from the definition of the CC wave operator we get<sup>87</sup>

$$\hat{H}\hat{U}_k \exp(\hat{T})\Phi_0 = E_k\hat{U}_k \exp(\hat{T})\Phi_0.$$

<sup>87</sup>By neglecting higher than single and double excitations the equation represents an approximation.

Due to the missing *deexcitation* part (i.e. that which lowers the excitation rank, e.g., from doubles to singles) the operators  $\hat{U}_k$  and  $\hat{T}$  commute,<sup>88</sup> hence the operators  $\hat{U}_k$  and  $\exp(\hat{T})$  also commute:

$$\hat{U}_k \exp(\hat{T}) = \exp(\hat{T}) \hat{U}_k.$$

Substituting this we have:

$$\hat{H} \exp(\hat{T}) \hat{U}_k \Phi_0 = E_k \exp(\hat{T}) \hat{U}_k \Phi_0$$

and multiplying from the left with  $\exp(-\hat{T})$  we get:

$$[\exp(-\hat{T}) \hat{H} \exp(\hat{T})] \hat{U}_k \Phi_0 = E_k \hat{U}_k \Phi_0$$

or introducing the similarity transformed Hamiltonian

$$\hat{\mathcal{H}} = e^{-\hat{T}} \hat{H} e^{\hat{T}}$$

we obtain

$$\hat{\mathcal{H}} \hat{U}_k \Phi_0 = E_k \hat{U}_k \Phi_0.$$

From the last equation we subtract the CC equation for the ground state

$$[\exp(-\hat{T}) \hat{H} \exp(\hat{T})] \Phi_0 = E_0 \Phi_0$$

multiplied from the left with  $\hat{U}_k$ , i.e.  $\hat{U}_k \hat{\mathcal{H}} \Phi_0 = E_0 \hat{U}_k \Phi_0$  and we get

$$\hat{\mathcal{H}} \hat{U}_k \Phi_0 - \hat{U}_k \hat{\mathcal{H}} \Phi_0 = E_k \hat{U}_k \Phi_0 - E_0 \hat{U}_k \Phi_0.$$

Finally, we obtain an important result:

$$[\hat{\mathcal{H}}, \hat{U}_k] \Phi_0 = (E_k - E_0) \hat{U}_k \Phi_0.$$

The operator  $\hat{U}_k$  contains the sought amplitudes  $\tau(k)$ .

We find them in a similar manner as in the CC method. For that purpose we make a scalar product of the left- and right-hand side of that equation with each excitation  $|_{ab...}^{mn...}\rangle$  used in  $\hat{U}_k$ , including<sup>89</sup> that of no excitation, i.e. the function  $\Phi_0$ . We get the set of the EOM-CC equations whose number is equal to the number of sought amplitudes plus one more equation due to normalization condition of  $\psi_k$ .

<sup>88</sup>If  $\hat{U}_k$  contains true excitations, then it does not matter whether excitations are performed by  $\hat{U}_k \hat{T}$  or  $\hat{T} \hat{U}_k$  (commutation), because both  $\hat{U}_k$  and  $\hat{T}$  mean going up in the energy scale. If, however,  $\hat{U}_k$  contains deexcitations, then it may happen that there is an attempt in  $\hat{T} \hat{U}_k$  to deexcite the ground state wave function – that makes immediately 0, whereas  $\hat{U}_k \hat{T}$  may be still OK, because the excitations in  $\hat{T}$  may be more important than the deexcitations in  $\hat{U}_k$ .

<sup>89</sup>More precisely: to get only the excitation energy we do not need the coefficient next to  $\Phi_0$ .

The unknown parameters are amplitudes and the excitation energies  $E_k - E_0$ :

$$\langle_{ab...}^{mn...} | [\hat{\mathcal{H}}, \hat{U}_k] | \Phi_0 \rangle = (E_k - E_0) \langle_{ab...}^{mn...} | \hat{U}_k | \Phi_0 \rangle.$$

We solve these equations and the problem is over.

## 10.16 MANY BODY PERTURBATION THEORY (MBPT)

The majority of routine calculations in quantum chemistry are done with variational methods (mainly the Hartree–Fock scheme). If we consider post-Hartree–Fock calculations then non-variational [CCSD, CCSD(T)] as well as perturbational (among them MBPT) approaches take the lead. The perturbational methods are based on the simple idea that the system, in a slightly modified condition, is similar to that before the perturbation is applied (cf. p. 203).

In the formalism of perturbation theory, knowing the unperturbed system and the perturbation we are able to provide successive corrections to obtain the solution of the perturbed system. Thus, for instance, the energy of the perturbed system is the energy of the unperturbed system plus the first-order correction plus the second-order correction plus... , etc.<sup>90</sup> If the perturbation is small then we *hope*<sup>91</sup> the series is convergent, even then however, there is no guarantee that the series converges fast.

### 10.16.1 UNPERTURBED HAMILTONIAN

In the perturbational approach (cf. p. 204) to the electron correlation the Hartree–Fock function,  $\Phi_0$ , is treated as the zero-order approximation to the true ground state wave function, i.e.  $\Phi_0 = \psi_0^{(0)}$ . Thus, the Hartree–Fock wave function stands at the starting point, while the goal is the exact ground-state electronic wave function  $\psi_0$ .

In majority of cases this is a reasonable approximation, since the Hartree–Fock method usually provides as much as 98–99% of the total energy.<sup>92</sup> A Slater determinant  $\Phi_I$  is constructed from the spinorbitals satisfying the Fock equation. How to construct the operator for which the Slater determinant is an eigenfunction? We will find out in a moment that this operator is the sum of the Fock operators (cf. Appendix U)

$$\hat{H}^{(0)} = \sum_i \hat{F}(i) = \sum_i \varepsilon_i \hat{t}^\dagger \hat{t}. \quad (10.55)$$

<sup>90</sup>This is an old trick of perturbation theory, equivalent to saying that the shape of a bridge loaded with a car is the shape of the bridge without the car plus the deformation proportional to the mass of the car plus the deformation proportional to the square of the mass of the car, etc. This works, if the bridge is solid and the car is light (the perturbation is small).

<sup>91</sup>There is not much known concerning the convergence of series occurring in quantum chemistry. Commonly, only a few perturbational corrections are computed.

<sup>92</sup>Sometimes, as we know, the method fails and then the perturbation theory based on the Hartree–Fock starting point is a risky business, since the perturbation is very large.

Indeed,

$$\hat{H}^{(0)}\Phi_I = \sum_i \varepsilon_i \hat{i}^\dagger \hat{i} \cdot \Phi_I = \left( \sum_i \varepsilon_i \right) \cdot \Phi_I, \quad (10.56)$$

since the annihilation of one spinorbital in the determinant and the creation of the same spinorbital leaves the determinant unchanged. This is so on condition that the spinorbital  $\phi_i$  is *present* in  $\psi_0^{(0)}$ .

The eigenvalue of  $\hat{H}_0 = \sum_i \varepsilon_i \hat{i}^\dagger \hat{i}$  is always the sum of the orbital energies corresponding to all spinorbitals in the Slater determinant  $\Phi_I$ .

This means that the sum of several determinants, each built from a different (in the sense of the orbital energies) set of spinorbitals, is not an eigenfunction of  $\hat{H}^{(0)}$ .

### 10.16.2 PERTURBATION THEORY – SLIGHTLY DIFFERENT APPROACH

We have to solve the Schrödinger equation for the ground state<sup>93</sup>  $\hat{H}\psi_0 = E\psi_0$ , with  $\hat{H} = \hat{H}^{(0)} + \hat{H}^{(1)}$ , where  $\hat{H}^{(0)}$  denotes the unperturbed Hamiltonian, and  $\hat{H}^{(1)}$  is a perturbation operator. We assumed that  $\hat{H}^{(0)}$  has eigenfunctions and corresponding energy eigenvalues

$$\hat{H}^{(0)}\psi_k^{(0)} = E_k^{(0)}\psi_k^{(0)}. \quad (10.57)$$

The ground state  $\psi_0^{(0)}$  is non-degenerate (assumption).

The Schrödinger equation does not force the normalization of the function. It is convenient to use the *intermediate normalization* (Fig. 10.11.a), i.e. to require that  $\langle \psi_0 | \psi_0^{(0)} \rangle = 1$ . This means that the (non-normalized)  $\psi_0$  must include the normalized function of zeroth order  $\psi_0^{(0)}$  and, possibly, something orthogonal to it.

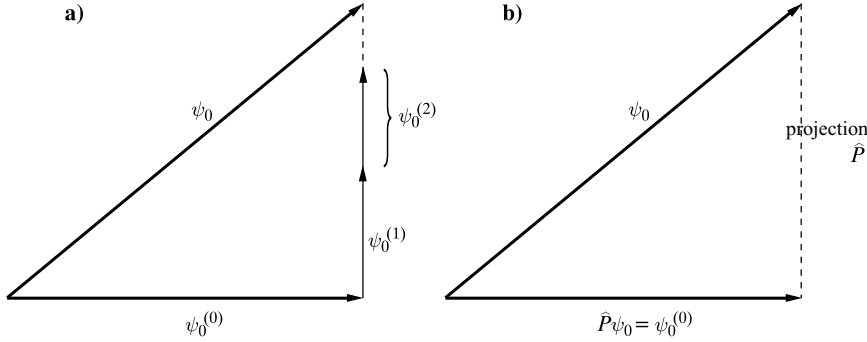
Let us write  $\hat{H}\psi_0$  as  $\hat{H}\psi_0 = (\hat{H}^{(0)} + \hat{H}^{(1)})\psi_0$ , or, in another way, as  $\hat{H}^{(1)}\psi_0 = (\hat{H} - \hat{H}^{(0)})\psi_0$ . Multiplying this equation by  $\psi_0^{(0)}$  and integrating, we get (taking advantage of the intermediate normalization)

$$\begin{aligned} \langle \psi_0^{(0)} | \hat{H}^{(1)} \psi_0 \rangle &= \langle \psi_0^{(0)} | (\hat{H} - \hat{H}^{(0)}) \psi_0 \rangle = E_0 \langle \psi_0^{(0)} | \psi_0 \rangle - \langle \psi_0^{(0)} | \hat{H}^{(0)} \psi_0 \rangle \\ &= E_0 - E_0^{(0)} = \Delta E_0. \end{aligned} \quad (10.58)$$

Thus,

$$\Delta E_0 = \langle \psi_0^{(0)} | \hat{H}^{(1)} \psi_0 \rangle. \quad (10.59)$$

<sup>93</sup>We use the notation from Chapter 5.



**Fig. 10.11.** Pictorial presentation of (a) the intermediate normalization  $\langle \psi | \psi_0^{(0)} \rangle = 1$ ,  $\psi_0^{(n)}$  is the  $n$ -th correction, and (b) the projection onto the axis  $\psi_0^{(0)}$  in the Hilbert space using the operator  $\hat{P} = |\psi_0^{(0)}\rangle\langle\psi_0^{(0)}|$ .

### 10.16.3 REDUCED RESOLVENT OR THE “ALMOST” INVERSE OF $(E_0^{(0)} - \hat{H}^{(0)})$

Let us define several useful quantities – we need to get familiar with them now – which will introduce a certain elegance into our final equations.

Let the first be a *projection operator* on the ground-state zeroth order function

projection  
operator

$$\hat{P} = |\psi_0^{(0)}\rangle\langle\psi_0^{(0)}|. \quad (10.60)$$

This means that  $\hat{P}\chi$  is, within accuracy to a constant, equal to  $\psi_0^{(0)}$  for an arbitrary function  $\chi$ . Indeed, if  $\chi$  is expressed as a linear combination of the eigenfunctions  $\psi_n^{(0)}$  (these functions form an orthonormal complete set as eigenfunctions of the Hermitian operator)

$$\chi = \sum_n c_n \psi_n^{(0)}, \quad (10.61)$$

then (Fig. 10.11.b)

$$\hat{P}\chi = \sum_n c_n \hat{P}\psi_n^{(0)} = \sum_n c_n |\psi_0^{(0)}\rangle\langle\psi_0^{(0)}|\psi_n^{(0)}\rangle = \sum_n c_n \delta_{0n} \psi_0^{(0)} = c_0 \psi_0^{(0)}. \quad (10.62)$$

Let us now introduce a projection operator

$$\hat{Q} = 1 - \hat{P} = \sum_{n=1}^{\infty} |\psi_n^{(0)}\rangle\langle\psi_n^{(0)}| \quad (10.63)$$

in the space orthogonal to  $\psi_0^{(0)}$ . Obviously,  $\hat{P}\hat{Q} = 0$ ,  $\hat{P}^2 = \hat{P}$  and  $\hat{Q}^2 = \hat{Q}$ . The latter holds since  $\hat{Q}^2 = (1 - \hat{P})^2 = 1 - 2\hat{P} + \hat{P}^2 = 1 - \hat{P} = \hat{Q}$ .



reduced  
resolvent

Now we define a *reduced resolvent*

$$\hat{R}_0 = \sum_{n=1}^{\infty} \frac{|\psi_n^{(0)}\rangle\langle\psi_n^{(0)}|}{E_0^{(0)} - E_n^{(0)}}. \quad (10.64)$$

For functions orthogonal to  $\psi_0^{(0)}$ , the action of the operator  $\hat{R}_0$  is identical to that of the operator  $(E_0^{(0)} - \hat{H}^{(0)})^{-1}$ . Let us make sure of this. Let us operate first on the function  $\phi$  orthogonal to  $\psi_0^{(0)}$  with the operator  $\hat{R}_0(E_0^{(0)} - \hat{H}^{(0)})$ . The result should be equal to  $\phi$ . Let us see:

$$\hat{R}_0(E_0^{(0)} - \hat{H}^{(0)})\phi = \sum_{n=1}^{\infty} (E_0^{(0)} - E_n^{(0)})^{-1} |\psi_n^{(0)}\rangle\langle\psi_n^{(0)}| (E_0^{(0)} - \hat{H}^{(0)})\phi \quad (10.65)$$

$$= \sum_{n=1}^{\infty} (E_0^{(0)} - E_n^{(0)})^{-1} (E_0^{(0)} - E_n^{(0)}) |\psi_n^{(0)}\rangle\langle\psi_n^{(0)}|\phi \quad (10.66)$$

$$= \sum_{n=1}^{\infty} |\psi_n^{(0)}\rangle\langle\psi_n^{(0)}|\phi = \hat{Q}\phi = \phi, \quad (10.67)$$

since for  $\phi$  orthogonal to  $\psi_0^{(0)}$  the projection  $\hat{Q}\phi$  equals  $\phi$ . Let us now operate on the same function with the operator  $(E_0^{(0)} - \hat{H}^{(0)})\hat{R}_0$  (i.e. the operators are in reverse order):

$$\begin{aligned} (E_0^{(0)} - \hat{H}^{(0)})\hat{R}_0\phi &= (E_0^{(0)} - \hat{H}^{(0)}) \sum_{n=1}^{\infty} (E_0^{(0)} - E_n^{(0)})^{-1} |\psi_n^{(0)}\rangle\langle\psi_n^{(0)}|\phi \\ &= \sum_{n=1}^{\infty} (E_0^{(0)} - E_n^{(0)})^{-1} (E_0^{(0)} - \hat{H}^{(0)}) |\psi_n^{(0)}\rangle\langle\psi_n^{(0)}|\phi \\ &= \sum_{n=1}^{\infty} |\psi_n^{(0)}\rangle\langle\psi_n^{(0)}|\phi = \hat{Q}\phi = \phi. \end{aligned} \quad (10.68)$$

It really looks as if the  $\hat{R}_0$  is the inverse of  $(E_0^{(0)} - \hat{H}^{(0)})$ . This is not so, since when acting on the function  $\psi_0^{(0)}$  we get

$$\hat{R}_0(E_0^{(0)} - \hat{H}^{(0)})\psi_0^{(0)} = \hat{R}_0 \cdot 0 = 0, \quad (10.69)$$

and not  $\psi_0^{(0)}$ . In other words  $\hat{R}_0(E_0^{(0)} - \hat{H}^{(0)})\psi_0^{(0)} = 0 \neq \psi_0^{(0)}$ .

Similarly,

$$\begin{aligned} (E_0^{(0)} - \hat{H}^{(0)})\hat{R}_0\psi_0^{(0)} &= (E_0^{(0)} - \hat{H}^{(0)}) \sum_{n=1}^{\infty} (E_0^{(0)} - E_n^{(0)})^{-1} |\psi_n^{(0)}\rangle\langle\psi_n^{(0)}|\psi_0^{(0)} \\ &= (E_0^{(0)} - \hat{H}^{(0)}) \cdot 0 = 0 \neq \psi_0^{(0)}. \end{aligned}$$

Thus, the reduced resolvent is “almost” the inverse of  $(E_0^{(0)} - \hat{H}^{(0)})$ , almost, because it happens only when acting on the functions from the space orthogonal to  $\psi_0^{(0)}$ . When the reduced resolvent operates on an arbitrary function, the result belongs to the  $Q$  space, but it does not represent a projection on the  $Q$  space. Indeed, let us operate with  $\hat{R}_0$  on function  $\phi$ :

$$\begin{aligned}\hat{R}_0\phi &= \sum_{n=1}^{\infty} (E_0^{(0)} - E_n^{(0)})^{-1} |\psi_n^{(0)}\rangle \langle \psi_n^{(0)} | \phi \rangle \\ &= \text{linear combination of functions orthogonal to } \psi_0^{(0)}.\end{aligned}\quad (10.70)$$

Such a linear combination always belongs to the  $Q$  space, but we have not obtained  $\phi$ , hence  $\hat{R}_0$  is not a projection operator.

#### 10.16.4 MBPT MACHINERY

Our goal now will be to present the Schrödinger equation in a different form. Let us first write it down as follows

$$(E_0 - \hat{H}^{(0)})\psi_0 = \hat{H}^{(1)}\psi_0. \quad (10.71)$$

We aim at having  $(E_0^{(0)} - \hat{H}^{(0)})\psi_0$  on the left-hand side. Let us add  $(E_0^{(0)} - E_0)\psi_0$  to both sides of that equation to obtain

$$(E_0 - \hat{H}^{(0)})\psi_0 + (E_0^{(0)} - E_0)\psi_0 = \hat{H}^{(1)}\psi_0 + (E_0^{(0)} - E_0)\psi_0 \quad (10.72)$$

or

$$(E_0^{(0)} - \hat{H}^{(0)})\psi_0 = (E_0^{(0)} - E_0 + \hat{H}^{(1)})\psi_0. \quad (10.73)$$

Let us now operate on both sides of this equation with the reduced resolvent  $\hat{R}_0$

$$\hat{R}_0(E_0^{(0)} - \hat{H}^{(0)})\psi_0 = \hat{R}_0(E_0^{(0)} - E_0 + \hat{H}^{(1)})\psi_0. \quad (10.74)$$

On the left-hand side we have  $\hat{Q}\psi_0$  (as follows from eq. (10.67)), but  $\hat{Q}\psi_0 = (1 - \hat{P})\psi_0 = \psi_0 - |\psi_0^{(0)}\rangle \langle \psi_0^{(0)} | \psi_0 \rangle = \psi_0 - \psi_0^{(0)}$ , due to the intermediate normalization. As a result, the equation takes the form

$$\psi_0 - \psi_0^{(0)} = \hat{R}_0(E_0^{(0)} - E_0 + \hat{H}^{(1)})\psi_0. \quad (10.75)$$

Thus, we obtain

$$\psi_0 = \psi_0^{(0)} + \hat{R}_0(E_0^{(0)} - E_0 + \hat{H}^{(1)})\psi_0. \quad (10.76)$$

At the same time, based on the expression for  $\Delta E$  in perturbation theory (eq. (10.59)), we have:

$$E_0 = E_0^{(0)} + \langle \psi_0^{(0)} | \hat{H}^{(1)} | \psi_0 \rangle. \quad (10.77)$$

These are the equations of the many body perturbation theory, in which the exact wave function and energy are expressed in terms of the unperturbed functions and energies plus certain corrections. The problem is that, as can be seen, these corrections involve the unknown function and unknown energy.

Let us not despair in this situation, but try to apply an iterative technique. First substitute for  $\psi_0$  in the right-hand side of (10.76) that, which most resembles  $\psi_0$ , i.e.  $\psi_0^{(0)}$ . We obtain

$$\psi_0 \cong \psi_0^{(0)} + \hat{R}_0(E_0^{(0)} - E_0 + \hat{H}^{(1)})\psi_0^{(0)}, \quad (10.78)$$

and then the new approximation to  $\psi_0$  should again be plugged into the right-hand side and this procedure is continued *ad infinitum*. It can be seen that the successive terms form a series (let us hope that it is convergent).

$$\psi_0 = \sum_{n=0}^{\infty} [\hat{R}_0(E_0^{(0)} - E_0 + \hat{H}^{(1)})]^n \psi_0^{(0)}. \quad (10.79)$$

Now only known quantities occur on the right-hand side except for  $E_0$ , the exact energy. Let us pretend that its value is known and insert into the energy expression (10.77) the function  $\psi_0$

$$\begin{aligned} E_0 &= E_0^{(0)} + \langle \psi_0^{(0)} | \hat{H}^{(1)} | \psi_0 \rangle \\ &= E_0^{(0)} + \langle \psi_0^{(0)} | \hat{H}^{(1)} | \sum_{n=0}^M [\hat{R}_0(E_0^{(0)} - E_0 + \hat{H}^{(1)})]^n | \psi_0^{(0)} \rangle. \end{aligned} \quad (10.80)$$

Let us go back to our problem: we want to have  $E_0$  on the left-hand side of the last equation, while – for the time being –  $E_0$  occurs on the right-hand sides of both equations. To exit the situation we will treat  $E_0$  occurring on the right-hand side as a parameter manipulated in such a way as to obtain equality in both above equations. We may do it in two ways. One leads to Brillouin–Wigner perturbation theory, the other to Rayleigh–Schrödinger perturbation theory.

### 10.16.5 BRILLOUIN–WIGNER PERTURBATION THEORY

Let us decide first at what  $n = M$  we terminate the series, i.e. to what order of perturbation theory the calculations will be carried out. Say,  $M = 3$ . Let us now take any reasonable value<sup>94</sup> as a parameter of  $E_0$ . We insert this value into the right-hand side of eq. (10.80) for  $E_0$  and calculate the left-hand side, i.e.  $E_0$ . Then let us again insert the new  $E_0$  into the right-hand side and continue in this way until self-consistency, i.e. until (10.80) is satisfied. After  $E_0$  is known we go to eq. (10.79) and compute  $\psi_0$  (through a certain order, e.g.,  $M$ ).

<sup>94</sup>A “unreasonable” value will lead to numerical instabilities. Then we will learn that it was unreasonable to take it.

Brillouin–Wigner perturbation theory has, as seen, the somewhat unpleasant feature that successive corrections to the wave function depend on the  $M$  assumed at the beginning.

We may suspect<sup>95</sup> – and this is true – that the Brillouin–Wigner perturbation theory is not size consistent.

### 10.16.6 RAYLEIGH–SCHRÖDINGER PERTURBATION THEORY

As an alternative to Brillouin–Wigner perturbation theory, we may consider Rayleigh–Schrödinger perturbation theory, which is size consistent. In this method the total energy is computed in a stepwise manner

$$E_0 = \sum_{k=0}^{\infty} E_0^{(k)} \quad (10.81)$$

in such a way that first we calculate the first order correction  $E_0^{(1)}$ , i.e. of the order of  $\hat{H}^{(1)}$ , then the second order correction,  $E_0^{(2)}$ , i.e. of the order of  $(\hat{H}^{(1)})^2$ , etc. If we insert into the right-hand side of (10.79) and (10.80) the expansion  $E_0 = \sum_{k=0}^{\infty} E_0^{(k)}$  and then, by using the usual perturbation theory argument, we equalize the terms of the same order, we get for  $n = 0$ :

$$E_0^{(1)} = \langle \psi_0^{(0)} | \hat{H}^{(1)} | \psi_0^{(0)} \rangle, \quad (10.82)$$

for  $n = 1$ :

$$E^{(2)} = \langle \psi_0^{(0)} | \hat{H}^{(1)} \hat{R}_0 (E_0^{(0)} - E_0 + \hat{H}^{(1)}) \psi_0^{(0)} \rangle = \langle \psi_0^{(0)} | \hat{H}^{(1)} \hat{R}_0 \hat{H}^{(1)} \psi_0^{(0)} \rangle, \quad (10.83)$$

since  $\hat{R}_0 \psi_0^{(0)} = 0$ ; for  $n = 2$ :

$E^{(3)}$  = the third order terms from the expression:

$$\begin{aligned} & \langle \psi_0^{(0)} | \hat{H}^{(1)} [\hat{R}_0 (E_0^{(0)} - E_0^{(0)} - E_0^{(1)} - E_0^{(2)} - \dots + \hat{H}^{(1)})]^2 \psi_0^{(0)} \rangle \\ &= \langle \psi_0^{(0)} | \hat{H}^{(1)} \hat{R}_0 (-E_0^{(1)} - E_0^{(2)} - \dots + \hat{H}^{(1)}) \hat{R}_0 (-E_0^{(1)} - E_0^{(2)} - \dots + \hat{H}^{(1)}) \psi_0^{(0)} \rangle \end{aligned}$$

and the only terms of the third order are:

$$E^{(3)} = \langle \psi_0^{(0)} | \hat{H}^{(1)} \hat{R}_0 \hat{H}^{(1)} \hat{R}_0 \hat{H}^{(1)} \psi_0^{(0)} \rangle - E_0^{(1)} \langle \psi_0^{(0)} | \hat{H}^{(1)} \hat{R}_0^2 \hat{H}^{(1)} \psi_0^{(0)} \rangle, \quad (10.84)$$

etc.

Unfortunately, we cannot give a general expression for the  $k$ -th correction to the energy although we can give an algorithm for the construction of such an expression.<sup>96</sup> Rayleigh–Schrödinger perturbation theory (unlike the Brillouin–Wigner approach) has the nice feature that the corrections of the particular orders are independent of the maximum order chosen.

<sup>95</sup>Due to the iterative procedure.

<sup>96</sup>J. Paldus and J. Čížek, *Adv. Quantum Chem.* 105 (1975).

## 10.17 MØLLER-PLESSET VERSION OF RAYLEIGH-SCHRÖDINGER PERTURBATION THEORY

Let us consider the case of a closed shell.<sup>97</sup> In the Møller–Plesset perturbation theory we assume as  $\hat{H}^{(0)}$  the sum of the Hartree–Fock operators (from the RHF method), and  $\psi_0^{(0)} = \psi_{\text{RHF}}$ , i.e.:

$$\hat{H}^{(0)} = \sum_i \varepsilon_i i^\dagger i,$$

$$\hat{H}^{(0)} \psi_{\text{RHF}} = E_0^{(0)} \psi_{\text{RHF}}, \quad (10.85)$$

$$E_0^{(0)} = \sum_i \varepsilon_i, \quad (10.86)$$

(the last summation is over spinorbitals occupied in the RHF function) hence the perturbation, known in the literature as a *fluctuation potential*, is equal

$$\hat{H}^{(1)} = \hat{H} - \hat{H}^{(0)}. \quad (10.87)$$

We may carry out calculations through a given order for such a perturbation. A very popular method relies on the inclusion of the perturbational corrections to the energy through the second order (known as MP2 method) and through the fourth order (MP4).

### 10.17.1 EXPRESSION FOR MP2 ENERGY

What is the expression for the total energy in the MP2 method?

Let us note first that, when calculating the mean value of the Hamiltonian in the standard Hartree–Fock method, we automatically obtain the sum of the zeroth order energies  $\sum_i \varepsilon_i$  and the first-order correction to the energy  $\langle \psi_{\text{RHF}} | \hat{H}^{(1)} | \psi_{\text{RHF}} \rangle$ :

$$\begin{aligned} E_{\text{RHF}} &= \langle \psi_{\text{RHF}} | \hat{H} | \psi_{\text{RHF}} \rangle = \langle \psi_{\text{RHF}} | (\hat{H}^{(0)} + \hat{H}^{(1)}) | \psi_{\text{RHF}} \rangle \\ &= \left( \sum_i \varepsilon_i \right) + \langle \psi_{\text{RHF}} | \hat{H}^{(1)} | \psi_{\text{RHF}} \rangle. \end{aligned}$$

So what is left to be done (in the MP2 approach) is the addition of the second order correction to the energy (p. 208, the prime in the summation symbol indicates that the term making the denominator equal to zero is omitted), where, as the complete set of functions, we assume the Slater determinants  $\psi_k^{(0)}$  corresponding to the energy  $E_k^{(0)}$  (they are generated by various spinorbital occupancies):

$$\begin{aligned} E_{\text{MP2}} &= E_{\text{RHF}} + \sum_k' \frac{|\langle \psi_k^{(0)} | \hat{H}^{(1)} | \psi_{\text{RHF}} \rangle|^2}{E_0^{(0)} - E_k^{(0)}} \\ &= E_{\text{RHF}} + \sum_k' \frac{|\langle \psi_k^{(0)} | \hat{H} | \psi_{\text{RHF}} \rangle|^2}{E_0^{(0)} - E_k^{(0)}}, \quad (10.88) \end{aligned}$$

<sup>97</sup>Møller–Plesset perturbation theory also has its multireference formulation when the function  $\Phi_0$  is a linear combination of determinants (K. Woliński, P. Pulay, *J. Chem. Phys.* 90 (1989) 3647).

since  $\psi_{\text{RHF}}$  is an eigenfunction of  $\hat{H}^{(0)}$ , and  $\psi_k^{(0)}$  and  $\psi_{\text{RHF}}$  are orthogonal. It can be seen that among possible functions  $\psi_k^{(0)}$ , we may ignore all but doubly excited ones. Why? This is because

- the single excitations give  $\langle \psi_k^{(0)} | \hat{H} \psi_{\text{RHF}} \rangle = 0$  due to the Brillouin theorem,
- the triple and higher excitations differ by more-than-two excitations from the functions  $\psi_{\text{RHF}}$  and, due to the IV Slater–Condon rule (see Appendix M, p. 986), give a contribution equal to 0.

In such a case, we take as the functions  $\psi_k^{(0)}$  only doubly excited Slater determinants  $\psi_{ab}^{pq}$ , which means that we replace the occupied spinorbitals:  $a \rightarrow p$ ,  $b \rightarrow q$ , and, to avoid repetitions  $a < b$ ,  $p < q$ . These functions are eigenfunctions of  $\hat{H}^{(0)}$  with the eigenvalues being the sum of the respective orbital energies (eq. (10.56)). Thus, using the III Slater–Condon rule, we obtain the energy accurate through the second order

$$E_{\text{MP2}} = E_{\text{RHF}} + \sum'_{a < b, p < q} \frac{|\langle ab|pq\rangle - \langle ab|qp\rangle|^2}{\varepsilon_a + \varepsilon_b - \varepsilon_p - \varepsilon_q}, \quad (10.89)$$

hence, the MP2 scheme viewed as an approximation to the correlation energy gives<sup>98</sup>

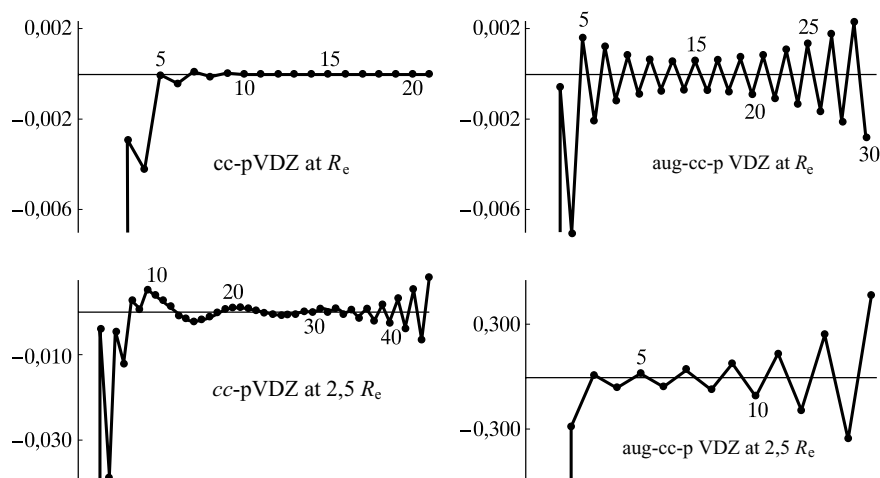
$$E_{\text{corel}} \approx E_{\text{MP2}} - E_{\text{RHF}} = \sum'_{a < b, p < q} \frac{|\langle ab|pq\rangle - \langle ab|qp\rangle|^2}{\varepsilon_i + \varepsilon_j - \varepsilon_m - \varepsilon_n}. \quad (10.90)$$

## 10.17.2 CONVERGENCE OF THE MØLLER–PLESSET PERTURBATION SERIES

Does the Møller–Plesset perturbation procedure converge? Very often this question can be considered surrealist, since *most frequently* we carry out calculations through the second, third, and – at most – fourth order of perturbation theory. Such calculations usually give quite a satisfactory description of the physical quantities considered and we do not think about going to high orders requiring major computational effort. There were, however, scientists interested to see how fast the convergence is if very high orders are included (MP $n$ ) for  $n < 45$ . And there was a surprise (see Fig. 10.12).

It is true that the first few orders of the MP perturbation theory give reasonably good results, but later, the accuracy of the MP calculations gets worse. A lot depends on the atomic orbital basis set adopted and the wealthy people (using the augmented basis sets – which is much more rare) encounter some difficulties whereas poor ones (modest basis sets) do not. Moreover, for long bond lengths

<sup>98</sup>The MP2 method usually gives satisfactory results, e.g., the frequencies of the normal modes. There are indications, however, that the deformations of the molecule connected with some vibrations strongly affecting the electron correlation (vibronic coupling) create too severe a test for the method – the error may amount to 30–40% for frequencies of the order of hundreds of  $\text{cm}^{-1}$  as has been shown by D. Michalska, W. Zierkiewicz, D.C. Bieńko, W. Wojciechowski, T. Zeegers-Huyskens, *J. Phys. Chem. A* 105 (2001) 8734.



**Fig. 10.12.** Convergence of the Møller-Plesset perturbation theory (deviation from the exact value, a.u.) for the HF molecule as a function of the basis set used (cc-pVDZ and augmented cc-pVDZ) and assumed bond length,  $R_e$  denotes the HF equilibrium distance (T. Helgaker, P. Jørgensen, J. Olsen, “*Molecular Electronic-Structure Theory*”, Wiley, Chichester, 2000, p. 780, Fig. 14.6. © 2000, John Wiley and Sons. Reproduced with permission of John Wiley and Sons Ltd.).

(2.5 of the equilibrium distance  $R_e$ ) the MPn performance is worse. For high orders, the procedure is heading for the catastrophe<sup>99</sup> already described on p. 210. The reason for this is the highly excited and diffuse states used as the expansion functions.<sup>100</sup>

### 10.17.3 SPECIAL STATUS OF DOUBLE EXCITATIONS

In Møller-Plesset perturbation theory  $\Delta E = E_0 - E_0^{(0)} = E_0 - E_{\text{RHF}} - E_0^{(0)} + E_{\text{RHF}} = E_{\text{corel}} + (E_{\text{RHF}} - E_0^{(0)})$ . On the other hand  $\Delta E = \langle \psi_0^{(0)} | \hat{H}^{(1)} | \psi \rangle$ . Substituting<sup>101</sup> the operator  $\hat{H} - \hat{H}^{(0)}$  instead of  $\hat{H}^{(1)}$  gives

$$\begin{aligned} \Delta E &= \langle \psi_0^{(0)} | (\hat{H} - \hat{H}^{(0)}) \psi_0 \rangle = \langle \psi_0^{(0)} | \hat{H} \psi_0 \rangle - \langle \psi_0^{(0)} | \hat{H}^{(0)} \psi_0 \rangle \\ &= \langle \psi_0^{(0)} | \hat{H} \psi_0 \rangle - \langle \hat{H}^{(0)} \psi_0^{(0)} | \psi_0 \rangle = \langle \psi_0^{(0)} | \hat{H} \psi_0 \rangle - E_0^{(0)} \langle \psi_0^{(0)} | \psi_0 \rangle \\ &= \langle \psi_0^{(0)} | \hat{H} \psi_0 \rangle - E_0^{(0)}. \end{aligned}$$

The function  $\psi_0$  can be expanded in Slater determinants of various excitation rank (we use intermediate normalization):  $\psi_0 = \psi_0^{(0)} + \text{excitations}$ . Then, by equalizing the two expressions for  $\Delta E$  obtained above, we have

<sup>99</sup>Except for the smaller basis set and the equilibrium bond length, but the problem has been studied up to  $n = 21$ .

<sup>100</sup>An analysis of this problem is given in the book cited in the caption to Fig. 10.12, p. 769.

<sup>101</sup>Also taking advantage of the intermediate normalization and the fact that  $\psi_0^{(0)}$  is an eigenfunction of  $\hat{H}^{(0)}$ .

$$\begin{aligned}
 E_{\text{corel}} + E_{\text{RHF}} &= \langle \psi_0^{(0)} | \hat{H} \psi_0 \rangle = \langle \psi_0^{(0)} | \hat{H} (\psi_0^{(0)} + \text{excitations}) \rangle \\
 &= E_{\text{RHF}} + \langle \psi_0^{(0)} | \hat{H} (\text{excitations}) \rangle,
 \end{aligned}$$

hence

$$E_{\text{corel}} = \langle \psi_0^{(0)} | \hat{H} (\text{excitations}) \rangle. \quad (10.91)$$

The Slater–Condon rules (Appendix M, p. 986) show immediately that the only excitations which give non-zero contributions are the single and double excitations. Moreover, taking advantage of the Brillouin theorem, we obtain single excitation contributions exactly equal to zero. So we get the result that

the exact correlation energy can be obtained from a formula containing exclusively double excitations.

The problem, however, lies in the fact that these doubly excited determinants are equipped with coefficients obtained in the *full* CI method, i.e. with all possible excitations. How is this? We should draw attention to the fact that, in deriving the formula for  $\Delta E$ , intermediate normalization is used. If someone gave us the normalized FCI (Full CI) wave functions as a Christmas gift,<sup>102</sup> then the coefficients occurring in the formula for  $\Delta E$  would not be the double excitation coefficient in the FCI function. We would have to denormalize this function to have the coefficient for the Hartree–Fock determinant equal to 1. We cannot do this without knowledge of the coefficients for higher excitations, cf. Fig. 10.9.

It is as if somebody said: the treasure is hidden in our room, but to find it we have to solve a very difficult problem in the Kingdom of Far Far Away. Imagine a compass which leads us unerringly to that place in our room where the treasure is hidden. Perhaps a functional exists whose minimization would provide us directly with the solution, but we do not know it yet.<sup>103</sup>

## Summary

- In the Hartree–Fock method *electrons of opposite spins do not correlate their motion*<sup>104</sup> which is an absurd situation (electrons of the same spins avoid each other – which is reasonable). In many cases (the  $F_2$  molecule, incorrect description of dissociation of chemical bonds, interaction of atoms and non-polar molecules) this leads to wrong results. In this chapter we have learnt about the methods which do take into account a correlation of electronic motions.

## VARIATIONAL METHODS USING EXPLICITLY CORRELATED WAVE FUNCTION

- Such methods rely on employing in the variational method a trial function which contains the explicit distance between the electrons. This improves the results significantly, but requires an evaluation of very complex integrals.

<sup>102</sup>Dreams...

<sup>103</sup>It looks like the work by H. Nakatsuji, *Phys. Rev. A* 14 (1976) 41 and M. Nooijen, *Phys. Rev. Letters* 84 (2000) 2108 go in this direction.

<sup>104</sup>Although they repel each other (mean field) as if they were electron clouds.



- An exact wave function satisfies the correlation cusp condition,  $(\frac{\partial \psi}{\partial r})_{r=0} = \mu q_i q_j \psi(r=0)$ , where  $r$  is the distance of two particles with charges  $q_i$  and  $q_j$ , and  $\mu$  is the reduced mass of the particles. This condition helps to determine the correct form of the wave function  $\psi$ . For example, for the two electrons the correct wave function has to satisfy (in a.u.):  $(\frac{\partial \psi}{\partial r})_{r=0} = \frac{1}{2} \psi(r=0)$ .
- The family of variational methods with explicitly correlated functions includes: the Hylleraas method, the Hylleraas CI method, the James–Coolidge and the Kołos–Wolniewicz approaches, and the method with exponentially correlated Gaussians. The method of explicitly correlated functions is very successful for 2-, 3- and 4-electron systems. For larger systems, due to the excessive number of complicated integrals, variational calculations are not yet feasible.

## VARIATIONAL METHODS WITH SLATER DETERMINANTS

- The CI (Configuration Interaction) approach is a *Ritz method* (Chapter 5) which uses the expansion in terms of *known* Slater determinants. These determinants are constructed from the molecular spinorbitals, usually occupied and virtual ones, produced by the Hartree–Fock method.
- Full CI expansion usually contains an enormous number of terms and is not feasible. Therefore, the CI expansion must be somewhere truncated. Usually we truncate it at a certain maximum rank of excitations with respect to the Hartree–Fock determinant (i.e. the Slater determinants corresponding to single, double, ... up to some maximal excitations are included).
- Truncated (limited) CI expansion is *not size consistent*, i.e. the energy of the system of non-interacting objects is not equal to the sum of the energies of the individual objects (calculated separately with the same truncation pattern).
- The MC SCF (Multiconfiguration Self Consistent Field) method is similar to the CI scheme, but we vary not only *the coefficients in front of the Slater determinants, but also the Slater determinants themselves* (changing the analytical form of the orbitals in them). We have learnt about two versions: the classic one (we optimize alternatively coefficients of Slater determinants and the orbitals) and a unitary one (we optimize *simultaneously* the determinantal coefficients and orbitals).
- The CAS SCF (Complete Active Space) method is a special case of the MC SCF approach and relies on *selection* of a set of spinorbitals (usually separated energetically from others) and on construction from them of all possible Slater determinants within the MC SCF scheme. Usually low energy spinorbitals are “inactive” during this procedure, i.e. they are doubly occupied in *each* Slater determinant (and are either frozen or allowed to vary). Most important active spinorbitals correspond to HOMO and LUMO.

## NON-VARIATIONAL METHODS WITH SLATER DETERMINANTS

- The CC (Coupled-Cluster) method is an attempt to find such an expansion of the wave function in terms of the Slater determinants, which would preserve size consistency. In this method the wave function for the electronic ground state is obtained as a result of the operation of the wave operator  $\exp(\hat{T})$  on the Hartree–Fock function (this ensures size consistency). The wave operator  $\exp(\hat{T})$  contains the cluster operator  $\hat{T}$ , which is defined as the sum of the operators for the  $l$ -tuple excitations,  $\hat{T}_l$  up to a certain maximum  $l = l_{\max}$ . Each  $\hat{T}_l$  operator is the sum of the operators each responsible for a *particular  $l$ -tuple excitation* multiplied by its *amplitude  $t$* . The aim of the CC method is to find the  $t$  values, since they determine the wave function and energy. The method generates non-linear

(with respect to unknown  $t$  amplitudes) equations. The CC method usually provides very good results.

- The EOM-CC (“Equation-of-Motion” CC) method is based on the CC wave function obtained for the ground state and is designed to provide the electronic excitation energies and the corresponding excited-state wave functions.
- The MBPT (Many Body Perturbation Theory) method is a perturbation theory in which the unperturbed system is usually described by a single Slater determinant. We obtain two basic equations of the MBPT approach:  $\psi_0 = \psi_0^{(0)} + \hat{R}_0(E_0^{(0)} - E_0 + \hat{H}^{(1)})\psi_0$  and  $E_0 = E_0^{(0)} + \langle \psi_0^{(0)} | \hat{H}^{(1)} | \psi_0 \rangle$ , where  $\psi_0^{(0)}$  is usually the Hartree–Fock function,  $E_0^{(0)}$  the sum of the orbital energies,  $\hat{H}^{(1)} = \hat{H} - \hat{H}^{(0)}$  is the fluctuation potential, and  $\hat{R}_0$  the reduced resolvent (i.e. “almost” inverse of the operator  $E_0^{(0)} - \hat{H}^{(0)}$ ). These equations are solved in an iterative manner. Depending on the iterative procedure chosen, we obtain either the Brillouin–Wigner or the Rayleigh–Schrödinger perturbation theory. The latter is applied in the Møller–Plesset (MP) method. *One of the basic computational methods for the correlation energy is the MP2 method*, which gives the result correct through the second order of the Rayleigh–Schrödinger perturbation theory (with respect to the energy).

### Main concepts, new terms

correlation energy (p. 499)	full CI method (p. 531)
explicit correlation (p. 502)	direct CI method (p. 533)
cuspid condition (p. 503)	size consistency (p. 532)
Hylleraas function (p. 506)	multireference methods (p. 533)
harmonic helium atom (p. 507)	active space (p. 535)
James–Coolidge function (p. 508)	frozen orbitals (p. 534)
Kołos–Wolniewicz function (p. 508)	multiconfigurational SCF methods (p. 535)
geminal (p. 513)	unitary MC SCF method (p. 536)
exponentially correlated function (p. 513)	commutator expansion (p. 537)
Coulomb hole (p. 513)	cluster operator (p. 540)
exchange hole (p. 516)	wave operator (p. 540)
Valence bond (VB) method (p. 520)	CC amplitudes (p. 542)
covalent structure (p. 521)	EOM-CC method (p. 548)
resonance theory (p. 520)	deexcitations (p. 550)
Heitler–London function (p. 521)	many body perturbation theory (MBPT)
ionic structure (p. 521)	(p. 551)
Brueckner function (p. 525)	reduced resolvent (p. 554)
configuration mixing (p. 525)	Brillouin–Wigner perturbation theory
configuration interaction (p. 526)	(p. 556)
configuration (p. 526)	Rayleigh–Schrödinger perturbation theory
Brillouin theorem (p. 527)	(p. 557)
density matrix (p. 531)	Møller–Plesset perturbation theory (p. 558)
natural orbitals (p. 531)	

### From the research front

The computational cost in the Hartree–Fock method scales with the size  $N$  of the atomic orbital basis set as  $N^4$  and, while using devices similar to *direct* CI, even<sup>105</sup> as  $N^3$ . How-

<sup>105</sup>This reduction is caused mainly by a preselection of the two-electron integrals. The preselection allows us to estimate the value of the integral without its computation and to reject the large number of integrals of values close to zero.

ever, after doing the Hartree–Fock computations for small (say, up to 10 electrons) systems, we perform more and more frequently calculations of the electronic correlation. The main approaches used to this end are: the MP2 method, the CC method with single and double excitations in  $\hat{T}$  and partial inclusion of triple ones (the so called CCSD(T) approach). The state-of-the art in CC theory currently includes the full CCSDTQP model, which incorporates into the cluster expansion all the operators through pentuple excitations.<sup>106</sup> The computational cost of the CCSD scheme scales as  $N^6$ . The computational strategy often adopted relies on obtaining the optimum geometry of the system with a less sophisticated method (e.g., Hartree–Fock) and, subsequently, calculating the wave function for that geometry with a more sophisticated approach (e.g., the MP2 that scales as  $N^5$ , MP4 or CCSD(T) scaling as  $N^7$ ). In the next chapter we will learn about the density functional theory (DFT) which joins the above mentioned methods and is used for large systems.

### Ad futurum...

Experimental chemistry is focused, in most cases, on molecules of *larger* size than those for which fair calculations with correlation are possible. However, after thorough analysis of the situation, it turns out that the cost of the calculations does not necessarily increase very fast with the size of a molecule. Employing localized molecular orbitals and using the multipole expansion of the integrals involving the orbital separated in space causes, for elongated molecules, the cost of the post-Hartree–Fock calculations to scale linearly with the size of a molecule.<sup>107</sup> It can be expected that, if the methods described in this Chapter are to survive in practical applications, such a step has to be made.

There is one more problem which will probably be faced by quantum chemistry when moving to larger molecules containing heteroatoms. Nearly all the methods including electron correlation described so far (with the exception of the explicitly correlated functions) are based on the silent and pretty “obvious” assumption, that the higher the excitation we consider the higher the configuration energy we get. This assumption seems to be satisfied so far, but the molecules considered were always small, and the method has usually been limited to a small number of excited electrons. This assumption can be challenged in certain cases. The multiple excitations in large molecules containing easily polarizable fragments can result in electron transfers which cause energetically favourable strong electrostatic interactions (“mnemonic effect”<sup>108</sup>) which lower the energy of the configuration. The reduction can be large enough to make the energy of the *formally* multiply excited determinant close to that of the Hartree–Fock determinant. Therefore, it should be taken into account on the same footing as the Hartree–Fock. This is rather unfeasible for the methods discussed above.

The explicitly correlated functions have a built-in adjustable and efficient basic mechanism accounting for the correlation within the interacting electronic pair. The mechanism is based on the obvious thing: the electrons should avoid each other.<sup>109</sup>

Let us imagine the CH<sub>4</sub> molecule. Let us look at it from the viewpoint of localized orbitals. With the method of explicitly correlated geminal functions for bonds we would succeed in making the electrons avoid each other within the same bond. And what should

<sup>106</sup>M. Musiał, S.A. Kucharski, R.J. Bartlett, *J. Chem. Phys.* 116 (2002) 4382.

<sup>107</sup>H.-J. Werner, *J. Chem. Phys.* 104 (1996) 6286.

<sup>108</sup>L.Z. Stolarczyk, L. Piela, *Chem. Phys. Letters* 85 (1984) 451, see also A. Jagielska, L. Piela, *J. Chem. Phys.* 112 (2000) 2579.

<sup>109</sup>In special conditions one electron can follow the other together forming a Cooper pair. The Cooper pairs are responsible for the mechanism of superconductivity. This will be a fascinating field of research for chemist-engineered materials in the future.

happen if the centre of gravity of the electron pair of one of the bonds shifts towards the carbon atom? The centres of gravity of the electron pairs of the remaining three bonds should move away along the CH bonds. The wave function must be designed in such a way that it accounts for this. In current theories, this effect is either deeply hidden or entirely neglected. A similar effect may happen in a polymer chain. One of the natural correlations of electronic motions should be a shift of electron pairs of all bonds in the same phase. As a highly many-electron effect the latter is neglected in current theories. However, the purely correlational Axilrod–Teller effect in the case of linear configuration, discussed in Chapter 13 (three-body dispersion interaction in the third order of perturbation theory), suggests clearly that the correlated motion of many electrons should occur.

It seems that the explicitly correlated functions, in spite of serious problems at the integral level, can be generalized in future towards the collective motions of electrons, perhaps on the basis of the renormalization theory of Kenneth Wilson (introduced into chemistry for the first time by Martin Head-Gordon).<sup>110</sup>

Kenneth Geddes Wilson (born 1936), American theoretical physicist. The authorities of Cornell University worried by Wilson's low number of published papers. Pressed by his supervisors, he finally started to publish, and won in 1982 the Nobel prize for the renormalization theory. It is a theory of the mathematical transformations describing a system viewed at various scales (with variable resolution). The renormalization theory, as applied by Head-Gordon to the hydrocarbon molecule, first



uses the LCAO (the usual atomic orbitals), then, in subsequent approximations, some linear combinations of functions that are more and more diffused in space.

### Additional literature

A. Szabo, N.S. Ostlund, "Modern Quantum Chemistry", McGraw-Hill, New York, 1989, p. 231–378.

Excellent book.

T. Helgaker, P. Jørgensen, J. Olsen, "Molecular Electronic-Structure Theory", Wiley, Chichester, 2000, p. 514.

Practical information on the various methods accounting for electron correlation presented in a clear and competent manner.

### Questions

1. The Hartree–Fock method for the helium atom in its ground state. If electron 1 resides on the one side of the nucleus then electron 2 can be found most likely:
  - a) on the other side of the nucleus; b) at the nucleus; c) on the same side of a nucleus; d) at infinite distance from the nucleus.
2. The Gaussian geminal for the helium atom  $\psi(\mathbf{r}_1, \mathbf{r}_2) = N(1 + \kappa r_{12}) \exp[-\frac{1}{4}(r_1^2 + r_2^2)]$ ,  $N$  is the normalization constant:
  - a) to satisfy the cusp condition should have  $\kappa = \frac{1}{2}$ ;
  - b) represents the exact wave function for  $\kappa = \frac{1}{2}$ ;
  - c) for  $\kappa < 0$  takes care of electron repulsion;
  - d) to satisfy the cusp condition has to have  $\exp[-\frac{1}{2}r_{12}^2]$  instead of  $(1 + \kappa r_{12})$ .

<sup>110</sup>M. Head-Gordon, "Proc. 5th Intern. Conf. Computers in Chemistry", Szklarska Poręba, Poland, 1999, p. L33.

3. The wave function for the  $\text{H}_2^-$  molecule [positions of nuclei a and b:  $(0, 0, 0)$  and  $(R, 0, 0)$ , respectively] in the form of a single Slater determinant, built of three spinorbitals  $\phi_1(\mathbf{r}, \sigma) = \varphi_1(\mathbf{r})\alpha(\sigma)$ ,  $\phi_2 = \varphi_1(\mathbf{r})\beta(\sigma)$ ,  $\phi_3 = \varphi_2(\mathbf{r})\alpha(\sigma)$  ( $\varphi_1$  is the doubly occupied bonding, and  $\varphi_2$  is the singly occupied antibonding one). If  $\mathbf{r}_1 = (\frac{R}{2}, 0, 0)$ ,  $\sigma_1 = \frac{1}{2}$ ,  $\mathbf{r}_2 = (0, 0, 0)$ ,  $\sigma_2 = -\frac{1}{2}$ ,  $\sigma_3 = \frac{1}{2}$  then the probability density of finding electron 3 is:
  - a) almost zero on nucleus a; b) almost zero on nucleus b; c) equal to 0 everywhere; d) proportional to  $|\varphi_2|^2$ .
4. A Hartree–Fock function:
  - a) correlates the positions of all electrons;
  - b) correlates the positions of electrons with the same spin coordinates;
  - c) correlates the positions of electrons with opposite spin coordinates;
  - d) does not correlate the positions of electrons, since in the Hartree–Fock method electron correlation is not accounted for.
5. The Brillouin theorem says that ( $\hat{H}$  is the Hamiltonian,  $\Phi_0$  is the Hartree–Fock function,  $\Phi_1$  is a singly and  $\Phi_2$  a doubly excited Slater determinant):
  - a)  $\langle \Phi_0 | \hat{H} \Phi_1 \rangle = 0$  if all the spinorbitals are orthogonal; b)  $\langle \Phi_1 | \hat{H} \Phi_1 \rangle = 0$ ; c)  $\langle \Phi_2 | \hat{H} \Phi_1 \rangle = 0$ ; d)  $\langle \Phi_0 | \Phi_1 \rangle = 0$ .
6. In the Coupled Cluster method ( $\hat{T}$  is the cluster operator,  $\Phi_0$  is the Hartree–Fock wave function) the wave function:
  - a) is  $\psi = \exp(i\hat{T})\Phi_0$ ; b) does not vanish in infinity; c) contains only single and double excitations; d) is  $\psi = \exp(\hat{T})\Phi_0$  and ensures size consistency.
7. MBPT: If the projector  $\hat{P} = |\psi_0^{(0)}\rangle\langle\psi_0^{(0)}|$  and  $\hat{Q} = \sum_{n=1}^{\infty} |\psi_n^{(0)}\rangle\langle\psi_n^{(0)}|$  ( $\psi_n^{(0)}$  form the complete orthonormal set) then:
  - a)  $\hat{P}\hat{Q} = 1$ ; b)  $(\hat{P} + \hat{Q})^2 = 1$ ; c)  $[\hat{P}, \hat{Q}] = i\hbar$ ; d)  $\hat{Q} = \exp(\hat{P})$ .
8. The Møller–Plesset method (MP2) is:
  - a) a variational method with two variational parameters; b) a perturbation theory with unperturbed wave function in the form of a Gaussian geminal; c) a perturbation theory with the energy computed through the second order; d) a Ritz method limited to double excitations.
9. To calculate the exact correlation energy:
  - a) it is enough to have the expansion in singly excited Slater determinants;
  - b) it is enough to know the Hartree–Fock function;
  - c) we must use explicitly correlated functions;
  - d) it is enough to have a certain wave function containing double excitations only.
10. We have the following order of mean values of the Hamiltonian calculated for the functions: *I*:  $\psi_1$  = the Hartree–Fock function, *II*:  $\psi_2$  = the Hartree–Fock function + doubly excited Slater determinant, *III*:  $\psi_3$  = the Hartree–Fock function +  $\lambda$  · doubly excited Slater determinant (the same as in  $\psi_2$ ), where  $\lambda$  is an optimal variational coefficient:
  - a)  $I > III > II$ ; b)  $I > III$  and  $II > III$ ; c)  $III > II > I$ ; d)  $I > II > III$ .

## Answers

1b, 2a, 3d, 4b, 5a, 6d, 7b, 8c, 9d, 10b