METHODS OF CALCULATION OF EXCITED STATES

Renzo Cimiraglia
Dipartimento di Chimica, Università di Ferrara

Contents

1 Electronic wave functions

1.1 The Born–Oppenheimer approximation

The total molecular Hamiltonian for a system with n electrons and N nuclei is given by

$$H = T_{nuc} + T_{el} + V_{en} + V_{ee} + V_{nn} \tag{1}$$

where the various terms are defined by the following expressions:

$$\begin{aligned}
T_{nuc} &= \sum_{\alpha=1}^{N} \frac{p_\alpha^2}{2M_\alpha} \quad \text{nuclear kinetic energy} \\
T_{el} &= \sum_{i=1}^{n} \frac{p_i^2}{2m} \quad \text{electronic kinetic energy} \\
V_{en} &= -\sum_{\alpha=1}^{N}\sum_{i=1}^{n} \frac{Z_\alpha e^2}{r_{\alpha i}} \quad \text{electron–nuclei potential energy} \\
V_{ee} &= \frac{1}{2}\sum_{i\neq j}^{n} \frac{e^2}{r_{ij}} \quad \text{electronic repulsion} \\
V_{nn} &= \frac{1}{2}\sum_{\alpha\neq\beta}^{N} \frac{Z_\alpha Z_\beta e^2}{r_{\alpha\beta}} \quad \text{nuclear repulsion}
\end{aligned}$$

Actually the expression (1) is only an approximation to the true Hamiltonian, since it lacks all the magnetic terms arising from the fact that the molecule is a system of moving charges endowed with spin. Anyway such magnetic terms are usually rather small so that their exclusion from the Hamiltonian can be considered to be a reasonable approximation.

The resolution of the time–independent Schrödinger equation

$$H\Psi_{tot}(x, Q) = E_{tot}\Psi(x, Q) \tag{2}$$

provides us with all the possible energies and stationary states of the molecule. In the above equation, x and Q designate the totality of electronic and nuclear coordinates, respectively. In order to alleviate the formidable task of solving the time–independent Schrödinger equation, Born and Oppenheimer showed that, due to the small ratio between electronic and nuclear masses, eq.(2) can be approximately decoupled in two separate equations, the first referring to the electronic motion and the second to the nuclear one. Defining the electronic Hamiltonian as

$$H_{el} = T_{el} + V_{en} + V_{ee} + V_{nn}$$

Born and Oppenheimer were able to show that

$$\begin{aligned} \Psi_{tot,n,k}(x,Q) &\approx \Psi_n(x;Q)\chi_{n,k}(Q) && (3) \\ H_{el}\Psi_n(x;Q) &= E_n\Psi_n(x;Q) && (4) \\ (T_{nuc}+E_n(Q))\chi_{n,k}(Q) &= E_{tot,n,k}\chi_{n,k}(Q) && (5) \end{aligned}$$

Eq. (4) plays the role of a Schrödinger equation for the motion of the electrons in the field generated by the *fixed* nuclei: the solutions provide *electronic wavefunctions* $\Psi_n(x;Q)$ (with n enumerating the various possible electronic states) and *electronic energies* $E_n(Q)$ where the nuclear coordinates Q only enter as external parameters. Eq. (5) is the equation associated to the nuclei, where the electronic energy $E_n(Q)$ plays the role of a *potential energy* governing the nuclear motion. Index k enumerates the *roto–vibrational* states associated with the electronic energy E_n. This Born–Oppenheimer separation is essential in Chemistry: most of our understanding of the very notion of molecular structure rests upon the geometrical properties of the *potential energy surface* $E_n(Q)$.

In the rest of these lessons we shall be mainly concerned with the (approximate) resolution of (4) for electronically excited states. We shall rewrite the electronic Schrödinger equation in the form

$$H_{el}\Psi_n(1,2,\ldots,n) = E_n\Psi_n(1,2,\ldots,n) \tag{6}$$

where 1,2,... denote the spatial and spin coordinates of the various electrons and where it is understood that the nuclei are kept fixed at a certain specified geometry.

1.2 Expansion of the electronic wavefunction over a set of Slater determinants

The electronic wavefunction, solution of eq. (6), must obey the antisymmetry principle for fermions, stating that $\Psi_n(1,2,\ldots,n)$ must change in sign upon transposition of the spin–coordinates of two electrons:

$$P_{ij}\Psi_n(1,2,\ldots,i,\ldots,j,\ldots,n) = -\Psi_n(1,2,\ldots,j,\ldots,i,\ldots,n) \tag{7}$$

Consequently Ψ_n can be expanded over a complete set of n–electron functions provided that each of them satisfies the antisymmetry principle. An easy way of building such a basis is to resort to a complete set of molecular spin–orbitals $(\psi_1,\psi_2,\ldots)$ and to construct all possible *Slater determinants* of the form:

$$\Phi'_K = \begin{vmatrix} \psi_{k_1}(1) & \psi_{k_2}(1) & \cdots & \psi_{k_n}(1) \\ \psi_{k_1}(2) & \psi_{k_2}(2) & \cdots & \psi_{k_n}(2) \\ \vdots & \vdots & \vdots & \vdots \\ \psi_{k_1}(n) & \psi_{k_2}(n) & \cdots & \psi_{k_n}(n) \end{vmatrix} \tag{8}$$

A great simplification is obtained if the spin–orbital basis is chosen to be *orthonormal*, *i.e.* $\langle \psi_i \mid \psi_j \rangle = \delta_{ij}$. In this case the determinants also constitute an orthogonal set ($\langle \Phi'_K \mid \Phi'_L \rangle = 0$ if $K \neq L$) and any determinant can be easily normalized upon multiplication by the factor $\frac{1}{\sqrt{(n!)}}$ ($\Phi_K = \frac{1}{\sqrt{(n!)}} \Phi'_K$). In the following we shall often designate a normalized determinant with the abbreviated symbol $||\psi_{k_1}\psi_{k_2}\cdots\psi_{k_n}||$.

The determinants Φ_K constitute a complete set of antisymmetric functions and consequently the wavefunction Ψ_n, solution of the electronic Schrödinger equation 6, can be expanded over such a set:

$$\Psi = \sum_K \Phi_K c_K \tag{9}$$

Substitution in eq. 6 gives:

$$\sum_K H_{el} \Phi_K c_K = E \sum_K \Phi_K c_K \tag{10}$$

and multiplication of both sides of the above equation by Φ_L^* with integration over the electronic coordinates (i.e. application of the "bra" vector $\langle \Phi_L |$) provides:

$$\sum_K \langle \Phi_L | H_{el} | \Phi_K \rangle c_K = E c_L \tag{11}$$

Introducing the matrix $\mathbf{H}$ defined by $H_{LK} = \langle \Phi_L | H_{el} | \Phi_K \rangle$ and the column vector $\mathbf{c}$ having as elements the c_K coefficients, eq. 11 can be put in matrix form:

$$\mathbf{H}\mathbf{c} = E\mathbf{c} \tag{12}$$

and the problem of finding the eigenvalues and eigenfunctions of the electronic Schrödinger equation has been turned into the algebraic problem of the diagonalization of the hermitian hamiltonian matrix $\mathbf{H}$.

The expression 9 is then able to provide the solution to the electronic Schrödinger equation and is known as the full *Configuration Interaction* (CI) expansion. The matrix elements H_{KL} are easily obtained if one resorts to the famous Slater's rules for one–electron and two–electron operators, which we now briefly recall.

Given a one–electron operator $H_M = \sum_{i=1}^{n} h(i)$, the matrix elements of H_M between two Slater determinants Φ_L and Φ_K are different from zero only in two cases:

- $\Phi_L = \Phi_K$ with $H_{KK} = \sum_{i=1}^{n} \langle \psi_{k_i} | h | \psi_{k_i} \rangle$
- $\Phi_L \neq \Phi_K$ for one spinorbital difference ($\psi_{l_i} \neq \psi_{k_i}$), with $H_{LK} = \langle \psi_{l_i} | h | \psi_{k_i} \rangle$

In the above formula it is supposed that the equal spinorbitals have the same ordering in both determinants. If such should not be the case, one has to take into account a possible sign change associated to the parity of the permutation necessary to bring the spinorbitals into coincidence. Similarly, for a two–electron operator $G = \frac{1}{2} \sum_{i \neq j} g(i,j)$, only three cases are possible for the matrix element G_{LK} to be different from zero:

- $\Phi_L = \Phi_K$ with $G_{KK} = \frac{1}{2} \sum_{i,j=1}^{n} (\left\langle \psi_{k_i}\psi_{k_j} | g | \psi_{k_i}\psi_{k_j} \right\rangle - \left\langle \psi_{k_i}\psi_{k_j} | g | \psi_{k_j}\psi_{k_i} \right\rangle)$
- $\Phi_L \neq \Phi_K$ for one spinorbital difference ($\psi_{l_i} \neq \psi_{k_i}$), with $G_{LK} = \sum_{j=1}^{n} (\left\langle \psi_{l_i}\psi_{k_j} | g | \psi_{k_i}\psi_{k_j} \right\rangle - \left\langle \psi_{l_i}\psi_{k_j} | g | \psi_{k_j}\psi_{k_i} \right\rangle)$
- $\Phi_L \neq \Phi_K$ for two spinorbital differences ($\psi_{l_i} \neq \psi_{k_i}$,$\psi_{l_j} \neq \psi_{k_j}$), with $G_{LK} = \left\langle \psi_{l_i}\psi_{l_j} | g | \psi_{k_i}\psi_{k_j} \right\rangle - \left\langle \psi_{l_i}\psi_{l_j} | g | \psi_{k_j}\psi_{k_i} \right\rangle$

In the following a two–electron integral $\langle \psi_a(1)\psi_b(2) | g(1,2) | \psi_c(1)\psi_d(2) \rangle$ will often be written in the simplified notation $\langle ab \,|\, cd \rangle$ and the combination $\langle ab \,|\, cd \rangle - \langle ab \,|\, dc \rangle$, known as an antisymmetrized two–electron integral, will be abbreviated in $\langle ab \,||\, cd \rangle$.

1.3 Diagonalization of large CI matrices

Nowadays full CI calculations yielding a few energy eigenvalues for small molecular systems are possible. Such calculations require the partial diagonalization of the Hamiltonian matrix which, as has been shown in the previous paragraph, is a large sparse $N \times N$ matrix. Currently diagonalizations with N surpassing 10^6 are routinely done (as of this writing $N = 10^9$ seems to be the limit of feasible diagonalizations). In this paragraph we shall briefly examine the main methods used in Quantum Chemistry to compute a few (usually the lowest) eigenvalues and the corresponding eigenvectors of the Hamiltonian matrix.

1.3.1 Subspace projection methods

Subspace projection methods start from the concept of Krylov subspace which is spanned by the vectors of the form

$$\mathbf{c}'_1, \mathbf{c}'_2 = \mathbf{H}\mathbf{c}'_1, \ldots \mathbf{c}'_k = \mathbf{H}\mathbf{c}'_{k-1}$$

with $\mathbf{c}'_1$ a given guess vector. In the old *power* method the last vector is shown to approach for large k the eigenvector associated to the highest eigenvalue. Actually, if $\mathbf{v}_1$, $\mathbf{v}_2$,...,$\mathbf{v}_N$ denotes the basis of orthonormal eigenvectors associated to the eigenvalues in increasing order of the moduli $|E_1| \leq |E_2| \leq \ldots \leq |E_N|$, one has:

$$\mathbf{c}'_1 = \sum_{i=1}^{N} \alpha_i \mathbf{v}_i, \mathbf{c}'_2 = \sum_{i=1}^{N} \alpha_i E_i \mathbf{v}_i, \ldots \mathbf{c}'_k = \sum_{i=1}^{N} \alpha_i E_i^{k-1} \mathbf{v}_i$$

$$\mathbf{c}'_k = E_N^{k-1}[\alpha_1(\frac{E_1}{E_N})^{k-1}\mathbf{v}_1 + (\alpha_2\frac{E_2}{E_N})^{k-1}\mathbf{v}_2 + \ldots + \alpha_N \mathbf{v}_N]$$

from which the stated result follows at once, provided that E_{N-1} is different from E_N.

An improvement to the power method consists in seeking an eigenvector of $\mathbf{H}$ in the form

$$\mathbf{u}_k = \sum_{i=1}^{k} s_i^{(k)} \mathbf{c}_i = \mathbf{C}\mathbf{s}$$

where the $\mathbf{c}_i$ are orthonormal vectors providing a basis for the Krylov space (obtained for instance by applying the Graham–Schmidt orthonormalization procedure to the $\mathbf{c}'_i$ vectors), $\mathbf{C}$ is the rectangular $N \times k$ matrix having $\mathbf{c}_i$ as i–th column and $\mathbf{s}$ is the k–dimensional vector having $s_i^{(k)}$ as components. The procedure consists in projecting the Hamiltonian matrix onto the Krylov space and diagonalizing such a reduced matrix $\mathbf{H}^{(k)}$. The reduced matrix is written as $\mathbf{H}^{(k)} = \mathbf{C}^\dagger\mathbf{H}\mathbf{C}$, has the (low) dimensionality $k \times k$ and can be easily diagonalized with standard techniques:

$$\mathbf{C}^\dagger\mathbf{H}\mathbf{C}\mathbf{s} = \mathbf{H}^{(k)}\mathbf{s} = \lambda^{(k)}\mathbf{s}$$

To be definite we suppose here to be interested to the smallest eigenvalue of $\mathbf{H}$. Taking into account that $\mathbf{P} = \mathbf{C}\mathbf{C}^\dagger$ is a $(N \times N)$ matrix which projects any N–dimensional vector onto the Krylov space, one has

$$\mathbf{PHPCs} = \lambda^{(k)}\mathbf{Cs} \quad \text{or} \quad \mathbf{PHP}\mathbf{u}_k = \lambda^{(k)}\mathbf{u}_k$$

with $\mathbf{u}_k = \mathbf{C}\mathbf{s}$. So $\lambda^{(k)}$ and $\mathbf{u}_k$ are an approximate eigenvalue–eigenvector couple for the problem of the diagonalization of $\mathbf{H}$. It can be remarked that, by the variational principle

$$\mathbf{u_k}^\dagger \mathbf{H}\mathbf{u}_k = \mathbf{s}^\dagger \mathbf{C}^\dagger \mathbf{H}\mathbf{C}\mathbf{s} = \lambda^{(k)} \geq E_1$$

Increasing the dimension of the Krylov space one improves the approximation of $\lambda^{(k)}$ for E_1 and usually one obtains good approximations for modest values of k ($<< N$). One can easily control the quality of the approximation by evaluating the *residual vector*

$$\mathbf{r}_k = \mathbf{H}\mathbf{u}_k - \lambda^{(k)}\mathbf{u}_k \tag{13}$$

and stopping the iterations when $||r_k|| < \epsilon$.

In Lanczos' approach the $\mathbf{c}_i$ vectors are defined so as to build an orthonormal basis which tridiagonalizes the small $\mathbf{H}^{(k)}$ matrix. Starting from a normalized guess $\mathbf{c}_1$, one builds $\mathbf{c}_2' = \mathbf{H}\mathbf{c}_1 - \alpha_1\mathbf{c}_1$ with $\alpha_1 = \mathbf{c}_1^\dagger \mathbf{H}\mathbf{c}_1$ and next $\mathbf{c}_2 = \dfrac{\mathbf{c}_2'}{\sqrt{\mathbf{c}_2'^\dagger \mathbf{c}_2'}}$. $\mathbf{c}_2$ is normalized to 1 and is immediately seen to be orthogonal to $\mathbf{c}_1$. A quantity β_2 is then defined so that $\beta_2\mathbf{c}_2 = \mathbf{c}_2'$, or $\beta_2 = \mathbf{c}_2^\dagger \mathbf{c}_2' = \sqrt{\mathbf{c}_2'^\dagger \mathbf{c}_2'}$ and the general iteration scheme is:

$$\beta_{i+1}\mathbf{c}_{i+1} = \mathbf{H}\mathbf{c}_i - \alpha_i\mathbf{c}_i - \beta_i\mathbf{c}_{i-1}$$

with $\alpha_i = \mathbf{c}_i^\dagger \mathbf{H}\mathbf{c}_i$ and β_i being defined by the requirement that $\mathbf{c}_i^\dagger \mathbf{c}_i = 1$. The vectors $\mathbf{c}_i$ are readily seen to be orthonormal and are such that the matrix $\mathbf{H}^{(k)} = \mathbf{C}^\dagger \mathbf{H}\mathbf{C}$ is tridiagonal

$$\mathbf{H}^{(k)} = \begin{bmatrix} \alpha_1 & \beta_2 & 0 & \cdots \\ \beta_2 & \alpha_2 & \beta_3 & \cdots \\ 0 & \beta_3 & \alpha_3 & \beta_4 \\ \vdots & \ddots & \beta_{k-1} & \alpha_k \end{bmatrix}$$

and can be easily diagonalized with standard techniques.

1.3.2 Davidson's method

In 1975 Davidson developed a method which is at present the most widely used in Quantum Chemistry for the diagonalization of large Hamiltonian matrices. Supposing to have a given starting space with orthonormalized vectors $\mathbf{c}_1, \ldots, \mathbf{c}_k$ and to have an approximate eigenvalue–eigenvector pair $\lambda^{(k)}, \mathbf{u}_k$ obtained with the projection method described previously, we look for the correction $\mathbf{z}$ which makes $\mathbf{u}_k + \mathbf{z}$ the true eigenvector corresponding to the true eigenvalue E_1 of $\mathbf{H}$

$$(\mathbf{H} - E_1\mathbf{1})(\mathbf{u}_k + \mathbf{z}) = 0 \tag{14}$$
$$(\mathbf{H} - E_1\mathbf{1})\mathbf{z} = -(\mathbf{H} - E_1\mathbf{1})\mathbf{u}_k \tag{15}$$

Now the r.h.s. of eq. (15) can be approximated with the residual vector $\mathbf{r}_k$ (see eq. 13) while the l.h.s., if $\mathbf{H}$ is diagonally dominant (as is usually the case in the Hamiltonian matrix), can be replaced by $(\mathbf{D} - \lambda^{(k)}\mathbf{1})\mathbf{z}$, where $\mathbf{D}$ contains just the diagonal of $\mathbf{H}$. Consequently $\mathbf{z} = -(\mathbf{D} - \lambda^{(k)}\mathbf{1})^{-1}\mathbf{r}_k$ is the desired correction vector. Adding such vector to the preceding basis and carrying out an orthonormalization, one obtains an augmented basis $\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_{k+1}$ and the process can be iterated until convergence is reached. If more than one eigenvalue–eigenvector couple is wanted, the $\mathbf{c}_i$ vectors obtained at convergence for the first eigenvalue are usually a good guess for another one.

2 Second quantization techniques

Given an orthonormal spin–orbital basis $\psi_1, \psi_2, \ldots$ and a generic n–electron determinant $|M\rangle = ||\psi_k\psi_a\psi_b\ldots||$ we shall define the *destruction* (or *annihilation*) operator a_k in the following way:

$$a_k\, ||\psi_k\psi_a\psi_b\ldots|| = ||\psi_a\psi_b\ldots|| \tag{16}$$

i.e., the result of the application of a_k to $|M\rangle$ is a (normalized) (n-1)–determinant $|N\rangle$ deprived of the spin–orbital ψ_k. In eq. (16) spin–orbital ψ_k was assumed to be in the first column of the determinant; if such should not be the case, one has to take into account the necessary transpositions to shift spin–orbital ψ_k to the first position, obtaining:

$$a_k\, ||\psi_a\psi_b\ldots\psi_k\ldots|| = (-1)^{\eta_k}||\psi_a\psi_b\ldots|| \tag{17}$$

with η_k equal to the number of spin–orbitals preceding ψ_k. If ψ_k is *not* occupied in $|M\rangle$ we have by definition $a_k\, |M\rangle = 0$. Besides, a_k is by definition a *linear* operator and this allows its effect on any function to be known (one only needs to expand the function over the set of determinants $|M\rangle$ and to apply a_k to every determinant). Let us now investigate on the properties of the *adjoint* operator $a_k^+ = a_k^\dagger$. From $a_k\, |M\rangle = |N\rangle$ we have $\left\langle N \left| a_k \right| M \right\rangle = 1$ and also, by definition of adjoint operator[1], $\left\langle M \left| a_k^+ \right| N \right\rangle = 1$. On the other

---

[1] $\langle f\,|A^\dagger|\,g\rangle = \langle g\,|A|\,f\rangle^*$

hand, $\left\langle M' \left| a_k^+ \right| N \right\rangle = 0$ if $|M'\rangle \neq |M\rangle$ allowing one to write

$$\sum_{M'} |M'\rangle \left\langle M' \left| a_k^+ \right| N \right\rangle = |M\rangle$$

and consequently:

$$a_k^+ \, |N\rangle = |M\rangle \tag{18}$$

since $\sum_{M'} |M'\rangle \langle M'| = 1$ (*resolution of the identity*). Therefore a_k^+ acts on the (n-1)–electron determinant $|N\rangle = ||\psi_a \psi_b \ldots||$ giving as result a n–electron determinant $|M\rangle = ||\psi_k \psi_a \psi_b \ldots||$. a_k^+ is called the operator of *creation* of an electron in spin–orbital ψ_k. If $|M\rangle$ already contains ψ_k the result is *zero* because a determinant containing two equal columns vanishes (an expression of the Pauli principle).

2.1 Anticommutation properties

We now consider a determinant $|M\rangle = ||\psi_i \psi_j \psi_a \psi_b \ldots||$ and apply to it the product of destruction operators $a_j \, a_i$

$$a_j \, a_i \, ||\psi_i \psi_j \psi_a \psi_b \ldots|| = ||\psi_a \psi_b \ldots||$$

If we apply such a product in reversed order we obtain

$$a_i \, a_j \, ||\psi_i \psi_j \psi_a \psi_b \ldots|| = -||\psi_a \psi_b \ldots||$$

Summing these two equalities one has:

$$(a_j \, a_i \, + a_i \, a_j \,) \, |M\rangle = 0 \tag{19}$$

and we see easily that such equation holds true for any $|M\rangle$, irrespective of whether spin–orbitals ψ_i and ψ_j are occupied or not in $|M\rangle$. So one has the operator identity:

$$a_j \, a_i \, + a_i \, a_j \, = 0 \tag{20}$$

which states that the *anticommutator* between two destruction operators $\left[a_i \, , a_j \,\right]_+ = a_i \, a_j \, + a_j \, a_i$ is zero. Taking the adjoint of eq. (20) one immediately sees that the same rule also holds for the creation operators:

$$\left[a_i^+, a_j^+\right]_+ = 0 \tag{21}$$

Putting $i = j$ in eqs (20) and (21) one gets:

$$a_i^2 = 0 \quad \text{and} \quad a_i^{+2} = 0$$

that is, the operators a_i and a_i^+ are *nilpotent.*

Let us now see the action of $a_i^+ a_j$ with $i \neq j$. Let $|M\rangle = ||\psi_j \psi_a \psi_b \ldots||$; one has:

$$a_i^+ a_j \ |M\rangle = ||\psi_i \psi_a \psi_b \ldots||$$

So, $a_i^+ a_j$ is a *replacement* operator, substituting spin–orbital ψ_j with ψ_i (notice that this results holds for whatever position the index j occupies in the determinant). The effect of $a_j \ a_i^+$ is instead:

$$a_j \ a_i^+ ||\psi_j \psi_a \psi_b \ldots|| = a_j \ ||\psi_i \psi_j \psi_a \psi_b \ldots|| = -||\psi_i \psi_a \psi_b \ldots||$$

Thus $(a_i^+ a_j \ + a_j \ a_i^+) \ |M\rangle = 0$ and one sees easily that the same result holds in all cases, irrespective of whether ψ_j is occupied or not in $|M\rangle$ or whether ψ_i is already occupied in $|M\rangle$. Consequently we conclude that $\left[a_i^+, a_j\right]_+ = 0$ for $i \neq j$. The case $i = j$ is also very easy: if ψ_i is occupied in $|M\rangle$ we have $a_i^+ a_i \ |M\rangle = |M\rangle$ and $a_i \ a_i^+ \ |M\rangle = 0$, else, if ψ_i is not occupied in $|M\rangle$: $a_i^+ a_i \ |M\rangle = 0$ and $a_i \ a_i^+ \ |M\rangle = |M\rangle$. In any case $(a_i^+ a_i \ + a_i \ a_i^+) \ |M\rangle = |M\rangle$. We can then state that for the anticommutator between a creator and a destructor one has:

$$\left[a_i^+, a_j\right]_+ = \delta_{i,j} \tag{22}$$

We conclude this paragraph with the following observations:

- A determinant $|M\rangle = ||\psi_{i_1} \psi_{i_2} \cdots \psi_{i_n}||$ can be successively depleted by application of the destructors a_{i_1}, a_{i_2},. . ., a_{i_n}. The final outcome is the completely ionized molecule, indicated with $|vac\rangle$ (the *vacuum* state). In other terms $a_{i_n} a_{i_{n-1}} \ldots a_{i_1} \ |M\rangle = |vac\rangle$
- Conversely, any n–electron determinant can be thought of as deriving from the application of a product of creation operators to the vacuum: $|M\rangle = a_{i_1}^+ a_{i_2}^+ \ldots a_{i_n}^+ \ |vac\rangle$

2.2 One– and two–electron operators

We shall consider an operator of the form

$$\hat{T} = \sum_{r,s} t_{rs} a_r^+ a_s \tag{23}$$

with the $\mathbf{T}$ matrix hermitian ($t_{rs} = t_{sr}^*$). It is easily seen that the operator $\hat{T}$ is hermitian ($\hat{T}^\dagger = T$). Given a generic determinant $|M\rangle = ||\psi_{k_1} \psi_{k_2} \cdots \psi_{k_n}||$, we shall now investigate the matrix element

$$\left\langle M \left| \hat{T} \right| M \right\rangle = \sum_{r,s} t_{rs} \left\langle M \left| a_r^+ a_s \right| M \right\rangle \tag{24}$$

Since the operator $a_r^+ a_s$ replaces spin–orbital ψ_s in $|M\rangle$ (if possible) with ψ_r, it is immediately recognized that the only possibility for $\langle M | a_r^+ a_s | M\rangle$ to be different from zero is that s equals an occupied index and $r = s$. So

$$\left\langle M \left| \hat{T} \right| M \right\rangle = \sum_{i=1}^{n} t_{k_i,k_i} \tag{25}$$

Let us now consider the matrix element of operator $\hat{T}$ between two different determinants $|M\rangle$ and $|N\rangle$, where $|M\rangle = ||\psi_{k_1}\psi_{k_2}\ldots\psi_a\ldots\psi_{k_n}||$ and $|N\rangle = ||\psi_{k_1}\psi_{k_2}\ldots\psi_b\ldots\psi_{k_n}||$, i.e. the two determinants differ by one spin–orbital difference ($\psi_a \neq \psi_b$):

$$\left\langle N \left| \hat{T} \right| M \right\rangle = \sum_{r,s} t_{rs} \left\langle N \left| a_r^+ a_s \right| M \right\rangle \tag{26}$$

The only non zero result can be obtained with $s = a$ and $r = b$, yielding

$$\left\langle N \left| \hat{T} \right| M \right\rangle = t_{ba} \tag{27}$$

If determinant $|N\rangle$ differs from $|M\rangle$ by more than one spin–orbital, it is immediately seen that $\left\langle N \left| \hat{T} \right| M \right\rangle = 0$. It has been thus recognized that operator $\hat{T}$ satisfies Slater's rules (see sec. 1.2 at page 4) for a one–electron operator. In other terms, if $\hat{F}$ is a one–electron operator ($\hat{F} = \sum_{i=1}^{n} f(i)$), it is equivalent to $\sum_{r,s} f_{rs} a_r^+ a_s$ where $f_{rs} = \langle \psi_r(1) | f(1) | \psi_s(1) \rangle$. Thus the expression (23) is the most general form to represent a one–electron operator in the so–called "second quantization" formalism.

Let us now consider an operator of the following form:

$$\hat{G} = \frac{1}{2} \sum_{rstu} g_{rs,tu} a_r^+ a_s^+ a_u \, a_t \tag{28}$$

where the four–index matrix g satisfies the hermitian requisite $g_{rs,tu} = g^*_{tu,rs}$. Operator $\hat{G}$ is promptly seen to be hermitian. Furthermore, as happens in almost all practical cases, we shall require the symmetry property $g_{rs,tu} = g_{sr,ut}$. It is then easily shown, noticing that the product $a_r^+ a_s^+ a_u \, a_t$ performs a double replacement (substituting the ordered couple $(\psi_t \psi_u)$ with $(\psi_r \psi_s)$), that operator $\hat{G}$ obeys Slater's rules for a two–electron operator:

$$\begin{aligned}
\left\langle M \left| \hat{G} \right| M \right\rangle &= \frac{1}{2} \sum_{i,j=1}^{n} (g_{k_i k_j, k_i k_j} - g_{k_i k_j, k_j k_i}) \\
\left\langle N \left| \hat{G} \right| M \right\rangle &= \sum_{i=1}^{n} (g_{k_i b, k_i a} - g_{k_i b, a k_i}) \quad \text{(one orb. diff.)} \\
\left\langle N \left| \hat{G} \right| M \right\rangle &= g_{cd,ab} - g_{cd,ba} \qquad \text{(two orb. diff.)}
\end{aligned}$$

Consequently the form (28) represents the most general expression for a two–electron operator in the second quantization formalism. For instance the interelectronic repulsion operator $\hat{G} = \frac{1}{2}\sum_{i,j}^{n}{}' \frac{1}{r_{ij}}$ becomes in the language of second quantization:

$$\hat{G} = \frac{1}{2}\sum_{rstu} \left\langle \psi_r(1)\psi_s(2) \left| \frac{1}{r_{12}} \right| \psi_t(1)\psi_u(2) \right\rangle a_r^+ a_s^+ a_u\, a_t \tag{29}$$

As concerns the electronic Hamiltonian

$$\hat{H}_{el} = \sum_{i=1}^{n} h(i) + \frac{1}{2}\sum_{i,j}^{n}{}' \frac{1}{r_{ij}},$$

where $h(i)$ collects the kinetic energy and nuclear coulombic attraction potential for electron i, one promptly gets the second–quantized form:

$$\hat{H}_{el} = \sum_{r,s} \langle \psi_r \,|h|\, \psi_s \rangle\, a_r^+ a_s \; + \frac{1}{2}\sum_{rstu} \left\langle \psi_r\psi_s \left| \frac{1}{r_{12}} \right| \psi_t\psi_u \right\rangle a_r^+ a_s^+ a_u\, a_t \tag{30}$$

It should be noted that the second–quantized version of the operators does not make any reference to the number of electrons, formulas (23) and (28) being valid for an arbitrary ionization state of the system.

2.3 Unitary transformations of the spin–orbital basis

We shall now study how the second quantization operators transform when one passes from a spin–orbital basis to another one: $|\boldsymbol{\psi}\rangle \rightarrow |\boldsymbol{\psi}'\rangle$, where we define the row vector $|\boldsymbol{\psi}\rangle = \{\,|\psi_1\rangle\,, |\psi_2\rangle\,, \ldots\}$ and similarly for $|\boldsymbol{\psi}'\rangle$. If $|\boldsymbol{\psi}\rangle$ and $|\boldsymbol{\psi}'\rangle$ are both orthonormal basis sets, the only way to perform the transformation is through a unitary operator $\hat{U}$

$$|\boldsymbol{\psi}\rangle' = \hat{U}\; |\boldsymbol{\psi}\rangle = |\boldsymbol{\psi}\rangle\, \mathbf{U} \tag{31}$$

where $\mathbf{U}$ is a given unitary matrix.

In the basis $|\boldsymbol{\psi}'\rangle$ the operators will have the same form as in the old $|\boldsymbol{\psi}\rangle$ basis. A generic one–electron operator $\hat{T}$ can be written as:

$$\hat{T} = \sum_{r,s} t_{rs} a_r^+ a_s \; = \sum_{r,s} t'_{rs} a_r'^+ a'_s$$

where a'_r $(a_r'^+)$ is the destruction (creation) operator associated to spin–orbital ψ'_r. Making use of the vector notation

$$\mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \end{bmatrix} \quad \text{and} \quad \mathbf{a}^+ = [a_1^+, a_2^+, \ldots]$$

we can write

$$\hat{T} = \mathbf{a}^+\mathbf{T}\mathbf{a} = \mathbf{a}'^+\mathbf{T}'\mathbf{a}'$$

Taking then into consideration that

$$\mathbf{T} = \langle \boldsymbol{\psi} \,|t|\, \boldsymbol{\psi}\rangle = \mathbf{U}\,\langle \boldsymbol{\psi}' \,|t|\, \boldsymbol{\psi}'\rangle\, \mathbf{U}^+$$

and consequently $\hat{T} = \mathbf{a}^+\mathbf{U}\mathbf{T}'\mathbf{U}^+\mathbf{a}$, one promptly deduces the formulas for the transformation of the destruction and creation operators:

$$\mathbf{a}' = \mathbf{U}^+\mathbf{a} \quad \text{and} \quad \mathbf{a}'^+ = \mathbf{a}^+\mathbf{U} \tag{32}$$

or, in detail

$$a'_r = \sum_s U^*_{sr} a_s \quad \text{and} \quad a'^+_r = \sum_s a^+_s U_{sr} \tag{33}$$

(notice that the creators transform exactly as the spin–orbitals).

We shall now investigate the action of $\hat{U}$ on a generic determinant $|K\rangle = ||\psi_1\psi_2\ldots\psi_n||$. Remembering that $|K\rangle = a^+_1 a^+_2 \ldots a^+_n\,|vac\rangle$ and that a unitary transformation of a generic operator[2] $\hat{A}$ is given by $\hat{A}' = \hat{U}A\hat{U}^+$, one has the following steps

$$\begin{aligned}
\hat{U}\,|K\rangle &= \hat{U}a^+_1 a^+_2 \ldots a^+_n\,|vac\rangle \\
&= \hat{U}a^+_1\hat{U}^+\hat{U}a^+_2\hat{U}^+\hat{U}a^+_3 \ldots \hat{U}a^+_n\hat{U}^+\hat{U}\,|vac\rangle \\
&= a'^+_1 a'^+_2 \ldots a'^+_n\,|vac\rangle \\
&= ||\psi'_1\psi'_2\ldots\psi'_n||
\end{aligned}$$

Thus, $\hat{U}\,|K\rangle$ is another determinant where spin–orbitals ψ_i have been replaced by the transformed spin–orbitals ψ'_i.

At this point we recall from Linear Algebra than any unitary operator $\hat{U}$ can be put in exponential form as $\hat{U} = e^{i\hat{\Lambda}}$ where $\hat{\Lambda}$ is a suitable hermitian operator. It is not difficult to show that $\hat{\Lambda}$ is just a simple *one–electron* operator

$$\hat{\Lambda} = \sum_{r,s} \lambda_{rs} a^+_r a_s \tag{34}$$

where matrix $\boldsymbol{\Lambda}$ is hermitian and is associated to matrix $\mathbf{U}$ by the exponential relation

$$\mathbf{U} = e^{i\boldsymbol{\Lambda}} \tag{35}$$

We shall barely sketch the demonstration which is based on the expansion of the transformed operator a'^+_r using the Baker–Campbell–Hausdorff identity:

$$e^A B e^{-A} = B + [A,B] + \frac{1}{2!}[A,[A,B]] + \frac{1}{3!}[A,[A,[A,B]]] + \ldots$$

---

[2] If $\hat{A}\Psi = \Phi$, application of $\hat{U}$ to both sides gives $\hat{U}\hat{A}\hat{U}^+(\hat{U}\Psi) = (U\Phi)$, which shows that $\hat{U}\hat{A}\hat{U}^+$ expresses the transformation of operator $\hat{A}$ effected by $\hat{U}$

Assuming $\hat{\Lambda}$ to have the form (34) and making use of the commutator relationship $[a_t^+ a_u\,, a_r^+] = \delta_{ru} a_t^+$, one readily arrives at the expansion:

$$\hat{U} a_r^+ \hat{U}^+ = a_r^+ + i \sum_t \lambda_{tr} a_t^+ + \frac{i^2}{2!} \sum_t \mathbf{\Lambda}_{tr}^2 a_t^+ + \ldots = \sum_t (e^{i\mathbf{\Lambda}})_{tr} a_t^+ = \sum_t U_{tr} a_t^+$$

which proves the assumption. Often one sets $\hat{T} = i\hat{\Lambda}$, with $\hat{T}$ antihermitian ($\hat{T}^\dagger = -\hat{T}$). If, as is usually the case, the two sets of orbitals $\boldsymbol{\psi}$ and $\boldsymbol{\psi}'$ are real, the $\mathbf{T}$ matrix ($t_{rs} = i\lambda_{rs}$) has to be skew–symmetric, i.e. $\tilde{\mathbf{T}} = -\mathbf{T}$.

2.4 The spin–traced replacement operators

Operators which do not have dependence upon the spin allow a simplification to be introduced by integrating over the spin variables in the matrix elements. Supposing that the spin–orbitals ψ_i are derived from a unique set of spatial orbitals φ_i with allocation of either α or β factors, one has for a spinless one–electron operator:

$$\hat{T} = \sum_{rs} \langle \psi_r \,|t|\, \psi_s \rangle \, a_r^+ a_s \; = \sum_{rs} \langle \varphi_r \,|t|\, \varphi_s \rangle \, (a_{r\alpha}^+ a_{s\alpha} + a_{r\beta}^+ a_{s\beta}) \tag{36}$$

where in the last term the summation runs over the spatial orbitals and where with $a_{r\alpha}^+$ etc. we mean the creation operator associated to spin–orbital $\varphi_r \alpha$ etc.

One can define the spin–traced replacement operators

$$\mathcal{E}_{rs} = a_{r\alpha}^+ a_{s\alpha} + a_{r\beta}^+ a_{s\beta} \tag{37}$$

which allow any spinless one–electron operator to be rewritten as

$$\hat{T} = \sum_{r,s} \langle \varphi_r \,|t|\, \varphi_s \rangle \, \mathcal{E}_{rs} \tag{38}$$

The commutation of two such operators is given by the following rule (easy to demonstrate for the reader)

$$[\mathcal{E}_{rs}, \mathcal{E}_{tu}] = \delta_{st} \mathcal{E}_{ru} - \delta_{ru} \mathcal{E}_{ts} \tag{39}$$

The above commutation property –we observe just in passing– has an important connection with the Unitary Group Approach (UGA) for the evaluation of the matrix elements of the Hamiltonian between spin eigenfunctions (rather than mere determinants).

An important property of the spin–traced replacement operators is that they commute with the total spin momentum and with its z component

$$[\mathcal{E}_{rs}, S^2] = 0, \quad [\mathcal{E}_{rs}, S_z] = 0 \tag{40}$$

as can be verified by expressing S^2 and S_z in second quantization. As a consequence the application of $\mathcal{E}_{rs}$ to an eigenfunction of S^2 and S_z produces a function with that same property.

For a two–electron operator we can proceed analogously

$$\begin{aligned} G &= \sum_{rstu} \langle \psi_r \psi_s \,|g|\, \psi_t \psi_u \rangle \, a_r^+ a_s^+ a_u \, a_t \\ &= \sum_{rstu} \langle \varphi_r \varphi_s \,|g|\, \varphi_t \varphi_u \rangle \, (a_{r\alpha}^+ a_{s\alpha}^+ a_{u\alpha} a_{t\alpha} + a_{r\beta}^+ a_{s\beta}^+ a_{u\beta} a_{t\beta} \\ &+ \quad a_{r\beta}^+ a_{s\alpha}^+ a_{u\alpha} a_{t\beta} + a_{r\alpha}^+ a_{s\beta}^+ a_{u\beta} a_{t\alpha}) \\ &= \sum_{rstu} \langle \varphi_r \varphi_s \,|g|\, \varphi_t \varphi_u \rangle \, (\mathcal{E}_{rt} \mathcal{E}_{su} - \delta_{ts} \mathcal{E}_{ru}) \end{aligned}$$

The Hamiltonian operator can then be expressed as

$$\hat{H} = \sum_{rs} h_{rs} \mathcal{E}_{rs} + \frac{1}{2} \sum_{rstu} \langle rs \mid tu \rangle \, (\mathcal{E}_{rt} \mathcal{E}_{su} - \delta_{ts} \mathcal{E}_{ru}) \tag{41}$$

3 Best one–determinant wavefunctions: Hartree–Fock theory

Using the techniques exposed in the preceding section it is an easy matter to find the best approximation to the wave function of the electronic *ground state* in the form of a single Slater's determinant $\Psi = ||\psi_1 \psi_2 \ldots \psi_n||$. Actually the method we are going to present applies not only to the ground state but also to the lowest state of a given *spatial* or *spin* symmetry. With the term "best approximation" we mean that the spin–orbitals we are in quest of should satisfy the *variational theorem*, i.e. they should minimize the *energy functional* $E = \left\langle \Psi \left| \hat{H} \right| \Psi \right\rangle$ with the orthonormality constraints $\langle \psi_i \mid \psi_j \rangle = \delta_{ij}$. At the point of minimum, any infinitesimal variation $\psi_i \to \psi_i + \delta\psi_i$ leads to $\delta E = 0$. Such a variation can be done, as has been seen in the previous section, by applying to Ψ a unitary operator $\hat{U} = e^{\hat{T}}$ with $\hat{T} = -\hat{T}^+ = \sum_{r,s} t_{rs} a_r^+ a_s$ $(t_{rs} = -t_{sr}^*)$. If $E' = \left\langle \Psi' \left| \hat{H} \right| \Psi' \right\rangle$ with $|\Psi'\rangle = e^{\hat{T}} \, |\Psi\rangle$, one has

$$E' = \left\langle e^{\hat{T}} \Psi \left| \hat{H} \right| e^{\hat{T}} \Psi \right\rangle = \left\langle \Psi \left| e^{-\hat{T}} \hat{H} e^{\hat{T}} \right| \Psi \right\rangle$$

and

$$\delta E = E' - E = \left\langle \Psi \left| e^{-\hat{T}} \hat{H} e^{T} - H \right| \Psi \right\rangle$$

Taking the t_{rs} parameters as infinitesimal, one obtains an infinitesimal variation of the energy. The expansion

$$e^{-\hat{T}} \hat{H} e^{T} = \hat{H} - [\hat{T}, \hat{H}] + \frac{1}{2} [\hat{T}, [\hat{T}, \hat{H}]] + \ldots$$

can be truncated to first order and consequently:

$$\delta E = \left\langle \Psi \left| \hat{H} + [\hat{H}, \hat{T}] - \hat{H} \right| \Psi \right\rangle = \left\langle \Psi \left| [\hat{H}, \hat{T}] \right| \Psi \right\rangle$$

It is now convenient to use the *replacement operators* $E_{rs} = a_r^+ a_s$ (with the property $E_{rs}^+ = E_{sr}$) which allow the expression for δE to be rephrased as:

$$\delta E = \sum_{r,s} t_{rs} \left\langle \Psi \left| \hat{H} \right| E_{rs} \Psi \right\rangle - \sum_{r,s} t_{rs} \left\langle E_{sr} \Psi \left| \hat{H} \right| \Psi \right\rangle \tag{42}$$

In the first term of the r.h.s. of eq (42) E_{rs} substitutes ψ_s with Ψ_r in Ψ and consequently s must designate the index of a spin–orbital occupied in Ψ; r, on the contrary, must be either the index of a *virtual* spin–orbital (i.e. not occupied in Ψ) or the same index as s. In the second term of eq. (42) the situation is reversed, with r occupied and s virtual (or coincident with r). Introducing now the convention of designating with $i, j, \ldots$ indices referring to occupied spin–orbitals and with $a, b, \ldots$ indices for virtual orbitals, eq.(42) can be rewritten as:

$$\delta E = \sum_{i=1}^{n} \sum_{a>n} t_{ai} \left\langle \Psi \left| \hat{H} \right| \Psi_i^a \right\rangle - \sum_{i=1}^{n} \sum_{a>n} t_{ia} \left\langle \Psi_i^a \left| \hat{H} \right| \Psi \right\rangle \tag{43}$$

where we have set $\Psi_i^a = E_{ai}\Psi$ (a *singly excited* determinant, with spin–orbital ψ_i replaced by ψ_a) and where we notice that the terms with t_{ii} have canceled out. Since $t_{ia} = -t_{ai}^*$, one can also write

$$\delta E = \sum_{i=1}^{n} \sum_{a>n} t_{ai} \left\langle \Psi \left| \hat{H} \right| \Psi_i^a \right\rangle + \quad \text{c.c.} \tag{44}$$

where "c.c." designates the complex conjugate of the preceding term. Since the quantities t_{ai} are arbitrary, in order for δE to vanish it is necessary that

$$\left\langle \Psi \left| \hat{H} \right| \Psi_i^a \right\rangle = 0 \tag{45}$$

The above result is known as *Brillouin's theorem*, stating that the spin–orbitals which minimize E are such that the interactions between Ψ and any singly excited determinant $E_{ai}\Psi$ are zero.

Now $\Psi = ||\psi_1\psi_2 \ldots \psi_i \ldots \psi_n||$ and $\Psi_i^a = ||\psi_1\psi_2 \ldots \psi_a \ldots \psi_n||$, so that, remembering Slater's rules, one has immediately:

$$\left\langle \Psi_i^a \left| \hat{H} \right| \Psi \right\rangle = \langle \psi_a \left| h \right| \psi_i \rangle + \sum_{j=1}^{n} \left(\left\langle \psi_a \psi_j \left| \frac{1}{r_{12}} \right| \psi_i \psi_j \right\rangle - \left\langle \psi_a \psi_j \left| \frac{1}{r_{12}} \right| \psi_j \psi_i \right\rangle \right)$$

One can introduce two auxiliary operators $\hat{J}$ and $\hat{K}$ (*Coulomb* and *exchange* operators, respectively) through their matrix elements:

$$\begin{aligned} \left\langle \psi_r \left| \hat{J} \right| \psi_s \right\rangle &= \sum_{j=1}^{n} \left\langle \psi_r \psi_j \left| \frac{1}{r_{12}} \right| \psi_s \psi_j \right\rangle \\ \left\langle \psi_r \left| \hat{K} \right| \psi_s \right\rangle &= \sum_{j=1}^{n} \left\langle \psi_r \psi_j \left| \frac{1}{r_{12}} \right| \psi_j \psi_s \right\rangle \end{aligned}$$

With these operators condition (45) can be cast in the form:

$$\left\langle \psi_a \left| \hat{h} + \hat{J} - \hat{K} \right| \psi_i \right\rangle = 0 \tag{46}$$

Fock's operator is now defined as $\hat{F} = \hat{h} + \hat{J} - \hat{K}$, so that one has the equation

$$\left\langle \psi_a \left| \hat{F} \right| \psi_i \right\rangle = 0 \tag{47}$$

Condition (47) states that the vector $\hat{F} \left| \psi_i \right\rangle$ must be orthogonal to any virtual $\left| \psi_a \right\rangle$. As a consequence $\hat{F} \left| \psi_i \right\rangle$ belongs to the subspace generated by the occupied spin–orbitals $\{\psi_1, \psi_2, \dots \psi_n\}$ and one obtains the so–called *generalized Hartree–Fock equations*:

$$\hat{F} \left| \psi_i \right\rangle = \sum_{j=1}^{n} \left| \psi_j \right\rangle \epsilon_{ji} \tag{48}$$

Since the Fock operator $\hat{F}$ is hermitian, matrix $\boldsymbol{\epsilon}$ is also hermitian:

$$\epsilon_{ij} = \left\langle \psi_i \left| \hat{F} \right| \psi_j \right\rangle = \left\langle \psi_j \left| \hat{F} \right| \psi_i \right\rangle^* = \epsilon_{ji}^*$$

The hermiticity of $\boldsymbol{\epsilon}$ can be exploited by considering the unitary matrix $\mathbf{U}$ which diagonalizes $\boldsymbol{\epsilon}$:

$$\mathbf{U}^+ \boldsymbol{\epsilon} \mathbf{U} = \begin{bmatrix} \epsilon_1 & 0 & \dots & 0 \\ 0 & \epsilon_2 & 0 & \dots \\ \vdots & & \ddots & \vdots \\ 0 & \dots & 0 & \epsilon_n \end{bmatrix}$$

Changing the occupied spin–orbitals according to the transformation:

$$\psi_i' = \sum_{j=1}^{n} \psi_j U_{ji}$$

and noticing that the Fock operator is invariant under such a transformation, one arrives at the *canonical Hartree–Fock equations*

$$\hat{F} \psi_i' = \epsilon_i \psi_i' \tag{49}$$

It should be noticed that eq. (49) is not a simple eigenvalue equation, because the operator $\hat{F}$ depends upon the solutions ψ_i's. In order to solve eq. (49) one has to resort to an iterative procedure, starting from a set of guess orbitals $\{\psi_1^{(0)}, \psi_2^{(0)}, \ldots\}$, building an initial $\hat{F}^{(0)}$, diagonalizing it ($\hat{F}^{(0)}\psi_i^{(1)} = \epsilon_i^{(0)}\psi_i^{(1)}$) and proceeding successively until convergence ("self consistence") is attained.

The eigenvalues ϵ_i of eq. (49) are called "orbital energies" and possess a neat physical meaning, as is shown by the following argument: if one considers the ionized state $\Psi_i^+ = a_i\, \Psi$, where an electron has been removed from spin–orbital ψ_i in the determinant Ψ, one obtains for the energy:

$$\begin{aligned} E_i^+ &= \left\langle a_i\, \Psi \left| \hat{H} \right| a_i\, \Psi \right\rangle = \left\langle \Psi \left| a_i^+ \hat{H} a_i \right| \Psi \right\rangle & (50) \\ &= \left\langle \Psi \left| \hat{H} a_i^+ a_i \right| \psi \right\rangle + \left\langle \Psi \left| [a_i^+, \hat{H}] a_i \right| \Psi \right\rangle & (51) \\ &= E + \left\langle \Psi \left| [a_i^+, \hat{H}] a_i \right| \Psi \right\rangle & (52) \end{aligned}$$

It is remarkable that eq. (52) provides the ionization energy ($E_i^+ - E$) as the expectation value in the *ground* state of a certain operator ($[a_i^+, \hat{H}]a_i$). It is not difficult to evaluate the commutator appearing in eq. (52) and, after a few algebraic steps (the reader can avail himself of the two basic commutation rules: $[a_i^+, a_r^+ a_s\,] = -\delta_{is} a_r^+$ and $[a_i^+, a_r^+ a_s^+ a_u\, a_t\,] = \delta_{iu} a_r^+ a_s^+ a_t\; - \delta_{it} a_r^+ a_s^+ a_u\,$) one gets (Koopmans' theorem):

$$\left\langle \Psi \left| [a_i^+, \hat{H}] a_i \right| \Psi \right\rangle = -h_{ii} - (J_{ii} - K_{ii}) = -\epsilon_i \qquad (53)$$

Analogously, upon creating a negatively ionized determinant with the addition to Ψ of an electron in the virtual spin–orbital ψ_v, one gets : $E_-^v - E = \epsilon_v$. Also, one can have a (usually rough) approximation to an electronically excited state in the form $\Psi_i^a = a_a^+ a_i\, \Psi$ with excitation energy $E_i^a - E = \epsilon_a - \epsilon_i - (J_{ai} - K_{ai})$.

3.1 Restricted and unrestricted Hartree–Fock equations

The Hartree–Fock equations considered in the previous section were developed in terms of spin–orbitals. Taking into account that $\psi_i(x_1, y_1, z_1, s_1) = \varphi_i(x_1, y_1, z_1)\eta_m(s_1)$, where φ_i is a *spatial* orbital and η_m is a *spin* function (either α if $m = \frac{1}{2}$ or β if $m = -\frac{1}{2}$), one can develop new equations in which the spin functions have been disposed of. The simplest situation to deal with is the so called *closed shell* case which is also the most frequent. Most molecules are in a singlet ground state and, close to the equilibrium geometry, are reasonably well described by a single Slater determinant where a given spatial orbital φ_i occurs twice, giving rise to the two spin–orbitals $\varphi_i\alpha$ and $\varphi_i\beta$. The Slater determinant describing the molecule is thus given by:

$$\Psi(1,2,\ldots,2n) = ||\varphi_1(1)\alpha(1)\varphi_1(2)\beta(2)\ldots\varphi_n(2n-1)\alpha(2n-1)\varphi_n(2n)\beta(2n)||$$

In this case the Hartree–Fock equations (eq. 49), upon integration over the spin functions, are easily seen to simplify to the following form:

$$\hat{f}\varphi_i = \epsilon_i\varphi_i \tag{54}$$

with $\hat{f} = \hat{h} + 2\hat{J} - \hat{K}$, where $\hat{J}$ and $\hat{K}$ are defined only in terms of spatial orbitals:

$$\begin{aligned} \left\langle \varphi_r \left| \hat{J} \right| \varphi_s \right\rangle &= \sum_{j=1}^{n} \left\langle \varphi_r\varphi_j \left| \frac{1}{r_{12}} \right| \varphi_s\varphi_j \right\rangle \\ \left\langle \varphi_r \left| \hat{K} \right| \varphi_s \right\rangle &= \sum_{j=1}^{n} \left\langle \varphi_r\varphi_j \left| \frac{1}{r_{12}} \right| \varphi_j\varphi_s \right\rangle \end{aligned}$$

In case one wants to describe an *open–shell* situation, e.g. when the molecule is not in a singlet state, the equations become more intricate: a determinant of the form

$$\Psi(1,2,\ldots,2n+m) = ||\varphi_1\alpha\varphi_1\beta\ldots\varphi_n\alpha\varphi_n\beta\varphi_{n+1}\alpha\ldots\varphi_{n+m}\alpha||$$

describes a state with spin quantum number $S = \frac{m}{2}$. The equations which determine the best orbitals φ_i are rather complicated to write down and are termed "Restricted Open–shell Hartree–Fock" equations (ROHF). A simpler alternative, often employed in practice, is to resort to *different orbitals for different spins*, i.e. to adopt two distinct sets of spatial orbitals, the first being associated to α spin–orbitals and the second to the β ones:

$$\Psi(1,2,\ldots,2n+m) = ||\varphi_1\alpha\varphi_1'\beta\ldots\varphi_n\alpha\varphi_n'\beta\varphi_{n+1}\alpha\ldots\varphi_{n+m}\alpha||$$

The resulting equations are called "Unrestricted Hartree-Fock equations" (UHF); they are simpler than the ROHF equations but suffer from the inconvenience of not providing an eigenstate of S^2.

3.2 Expansion in a set of atomic orbitals: Roothan's equations

The Hartree–Fock equations seen in the previous sections are complicated integro–differential equations not amenable to analytical solutions. The usual way to solve such equations is to resort to an expansion of the molecular orbitals φ_i in a chosen *atomic* basis set of orbitals $\{\chi_1, \chi_2, \ldots\}$:

$$\varphi_i = \sum_r \chi_r c_{ri} \tag{55}$$

thus reducing our problem to an algebraic one, where the mixing coefficients c_{ri} have to be determined. If the atomic basis set is complete the true solutions of the Hartree–Fock equations can in principle be obtained through the following steps (we envisage the closed–shell case here):

$$\begin{aligned} \hat{f}\varphi_i &= \epsilon_i \varphi_i \\ \sum_r \hat{f}\chi_r c_{ri} &= \epsilon_i \sum_r \chi_r c_{ri} \\ \sum_r \left\langle \chi_s \left| \hat{f} \right| \chi_r \right\rangle c_{ri} &= \epsilon_i \sum_r \left\langle \chi_s \mid \chi_r \right\rangle c_{ri} \end{aligned}$$

In matrix form one has the Roothan's SCF ("Self Consistent Filed") equations:

$$\mathbf{F}\mathbf{c}_i = \epsilon_i \mathbf{S}\mathbf{c}_i \tag{56}$$

where $\mathbf{c}_i$ is a column vector containing the c_{ri} coefficients, $\mathbf{S}$ is the overlap matrix ($S_{sr} = \langle \chi_s \mid \chi_r \rangle$) and $\mathbf{F}$ is the Fock matrix:

$$\begin{aligned} F_{sr} &= \left\langle \chi_s \left| \hat{f} \right| \chi_r \right\rangle \\ &= \langle \chi_s | h | \chi_r \rangle + \sum_{j=1}^{n} \left(2 \left\langle \chi_s \varphi_j \left| \frac{1}{r_{12}} \right| \chi_r \varphi_j \right\rangle - \left\langle \chi_s \varphi_j \left| \frac{1}{r_{12}} \right| \varphi_j \chi_r \right\rangle \right) \end{aligned}$$

Upon expansion of the φ_j orbitals according to eq. (55) one gets

$$F_{sr} = \langle \chi_s | h | \chi_r \rangle + \sum_{t,u} \left(2 \left\langle \chi_s \chi_t \left| \frac{1}{r_{12}} \right| \chi_r \chi_u \right\rangle - \left\langle \chi_s \chi_t \left| \frac{1}{r_{12}} \right| \chi_u \chi_r \right\rangle \right) R_{ut} \tag{57}$$

where the *density matrix* $\mathbf{R}$ has been introduced:

$$\begin{aligned} R_{ut} &= \sum_{j=1}^{n} c_{uj}^{*} c_{tj} \quad \text{or} \\ \mathbf{R} &= \mathbf{C}\mathbf{C}^{\dagger} \end{aligned}$$

with $\mathbf{C}$ designating the rectangular matrix having as columns the $\mathbf{c}_j$ vectors.

Roothans' equations (56) are classified as *pseudoeigenvalue* equations because matrix $\mathbf{F}$ needs the knowledge of the solutions $\mathbf{c}_i$ to be constructed. Their solution is generally carried out in an iterative way, starting from a guess coefficient matrix $\mathbf{C}^{(0)}$ from which an initial $\mathbf{R}$ and $\mathbf{F}$ matrix are built up; diagonalization of $\mathbf{F}$ (actually, due to the occurrence of the overlap $\mathbf{S}$ matrix, eq. (56) is a *generalized* eigenvalue equation) produces a new coefficient matrix $\mathbf{C}^{(1)}$ which permits the process to be iterated until convergence (self–consistence) is reached.

3.3 The CIS approximation for the calculation of excited states

We have seen that the Hartree–Fock equations can be derived from Brillouin's theorem (eq. 45 at page 17) which states that at self consistence there is no interaction between the H–F determinant and any single excitation. In other terms, no improvement is obtained for the ground state by diagonalizing the Hamiltonian matrix $\mathbf{H}$ built on the functions $\{\Psi_0, \Psi_i^a, \ldots\}$ containing, besides Ψ_0, all the single excitations $\Psi_i^a = E_{ai}\Psi_0$. So, the diagonalization of the Hamiltonian matrix ($\mathbf{Hc} = E\mathbf{c}$) yields as its lowest eigenvalue $E_0 = \langle \Psi_0 \,| H |\, \Psi_0 \rangle$ with $\mathbf{c}^\dagger = [1, 0, \ldots, 0]$. The higher eigenvalues, anyway, constitute approximations for the energies of excited states and the resulting method is usually referred to as "Configuration Interaction with Singles" (CIS) or Tamm–Dancoff approximation. The expression of the matrix elements of the Hamiltonian is straightforward recalling Slater's rules and yields at once:

$$\left\langle \Psi_j^b \,| H |\, \Psi_i^a \right\rangle = -\langle bi \,||\, aj \rangle + \delta_{ba}\delta_{ij}[(\epsilon_a - \epsilon_i) + E_0] \tag{58}$$

Also, the dimension of the Hamiltonian matrix is modest (by today's standards), being determined by the product of the number of occupied orbitals times that of the virtual ones. The CIS approximation is often used as one of the simplest methods able to provide reasonable results for excited states, its validity being chiefly confined to those states whose physical nature does not require higher excitations than the singles. Starting from a singlet Hartree–Fock wavefunction only singlet and triplet excited states can be obtained with the CIS method. Singlets can be obtained by applying the spin–traced operators $\mathcal{E}_{ai} = a_{a\alpha}^{+} a_{i\alpha} + a_{a\beta}^{+} a_{i\beta}$ to Ψ_0 and by diagonalizing the resulting Hamiltonian matrix. Triplet states, conversely, can be formed with the operators $\mathcal{E}'_{ai} = a_{a\alpha}^{+} a_{i\alpha} - a_{a\beta}^{+} a_{i\beta}$.

4 Multiconfigurational Self Consistent Field methods (MCSCF)

We have already remarked that a full configuration interaction (Full CI) can exactly solve the Schrödinger equation. Unfortunately a full CI is only feasible in very simple cases, with few electrons and with not too much extended orbital basis sets. Often one makes use of a *truncated* CI

$$\Psi = \sum_{K=1}^{N} C_K \Phi_K \tag{59}$$

where the electronic configurations estimated to be the most important are included in the CI. With the term MCSCF (Multiconfigurational Self Consistent Field) one means a generalization of Hartree–Fock's equations where, along with the CI coefficients C_K, also the orbitals φ_i utilized to build up the determinants Φ_K, are optimized. In other terms, the energy associated to eq. (59) $E = \left\langle \Psi \left| \hat{H} \right| \Psi \right\rangle$, has to be stationary with respect to both a variation in the expansion coefficients ($C'_K \to C_K + \delta C_K$) and a variation in the orbitals ($\varphi'_i \to \varphi_i + \delta\varphi$) with the orthonormality constraints ($\sum_K C^*_K C_K = 1$ and $\langle \varphi_i \mid \varphi_j \rangle = \delta_{ij}$). We shall now describe an iterative procedure, consisting of two successive steps repeated until convergence, which is able to solve the MCSCF problem.

1. One finds the coefficients C_K obtained by building determinants Φ_K with a given set of orbitals
2. Once the coefficients C_K are obtained, a set of orbitals minimizing $E = \left\langle \Psi \left| \hat{H} \right| \Psi \right\rangle$ is looked for

Step 1 is in principle easy, since one just needs to solve the matrix eigenvalue equation $\mathbf{Hc} = E\mathbf{c}$ with $H_{KL} = \left\langle \Phi_K \left| \hat{H} \right| \Phi_L \right\rangle$ easily obtainable through Slater's rules. Step 2 can be faced with techniques similar to those met for the case of a single determinant. We have already learned that the operator U associated to a unitary transformation of a spin–orbital basis, can be cast in the form $U = e^T$ with $T = \sum_{r,s} t_{rs} a_r^+ a_s$ $(t_{rs} = -t^*_{sr})$. For an infinitesimal transformation the expansion of e^T can be truncated to the first order and, as we already know, $e^{-T} H e^T = H - [T, H]$ so that the condition for Ψ to give the minimum energy is:

$$\delta E = \langle \Psi \left| [H, T] \right| \Psi \rangle = 0$$

By utilizing the replacement operators E_{rs}, one can write:

$$\begin{aligned} \langle \Psi \left| HT \right| \Psi \rangle - \langle \Psi \left| TH \right| \Psi \rangle &= 0 \\ \sum_{r,s} t_{rs} (\langle \Psi \left| H \right| E_{rs} \Psi \rangle - \langle E_{sr} \Psi \left| H \right| \Psi \rangle) &= 0 \end{aligned} \tag{60}$$

Assuming for simplicity that our functions are real, one has the equality $\langle E_{sr}\Psi \,|H|\, \Psi\rangle = \langle \Psi \,|H|\, E_{sr}\Psi\rangle$ and, given the arbitrariness of the t_{rs} coefficients, equation (60) can only be satisfied if

$$\langle \Psi \,|H|\, (E_{rs} - E_{sr})\Psi\rangle = 0 \tag{61}$$

Condition 61 is called "extended Brillouin's theorem" or "Brillouin–Levy–Berthier's theorem". The function $(E_{rs} - E_{sr})\Psi$ does not interact with Ψ when the orbitals are optimized. E_{rs} substitutes spin–orbital ψ_s with ψ_r in all the determinants of Ψ that contain ψ_s and do not contain ψ_r (else it gives zero); conversely, E_{sr} substitutes ψ_r with ψ_s when possible. By setting $\Psi_s^r = (E_{rs} - E_{sr})\Psi$, one must have at self–consistence $\left\langle \Psi \left| \hat{H} \right| \Psi_s^r \right\rangle = 0$. This means that trying to improve the wave function with $\Psi \rightarrow \Psi + \sum\limits_{r>s} c_{rs}\Psi_s^r$ would be of no avail if self–consistence has been reached. On the contrary, far from SCF, the diagonalization of the Hamiltonian matrix in the set $\{\Psi, \Psi_s^r, \ldots\}$ brings about an improvement of Ψ. One can then identify the corrected function $\Psi' = \Psi + \sum\limits_{r>s} c_{rs}\Psi_s^r$ as the first–order development of $e^T\Psi = \Psi + \sum\limits_{r>s} t_{rs}(E_{rs} - E_{sr})\Psi + \ldots$ and one can utilize the c_{rs} coefficients to obtain a new spin–orbital basis:

$$\boldsymbol{\psi}' = \boldsymbol{\psi} e^{\mathbf{T}} \quad \text{with} \quad t_{rs} = c_{rs}$$

The improved spinorbitals can then be employed to rebuild the CI wavefunction and to continue the iterative process until convergence is reached. This method is referred to as the *Super–CI* (SCI) approach.

Other approaches, such as those based on second–order Newton–Raphson techniques, are more effective than super–CI and can solve the MCSCF problem with a single step procedure, i.e. optimizing at the same time the C_K CI coefficients and the orbitals. Anyway the super–CI approach is still in use since it requires rather modest computational resources.

A problem in MC–SCF is related to which configurations should be chosen to build the CI wavefunction (59). Often one resorts to the concept of *Complete Active Space* (CAS) which consists in a partition of the spinorbitals in three classes:

1. core spinorbitals, always having occupation 1 in all the determinants of Ψ
2. active spinorbitals, appearing with all possible occupations in the determinants (0 or 1)

3. virtual spinorbitals, being never occupied in the determinants

The resulting wavefunction, once convergence of the MC–SCF has been attained, is called a CAS–SCF wavefunction.

For a CAS–SCF wavefunction the extended Brillouin's theorem, eq. (61) assumes a somewhat simplified form. If we designate with $i, j, \ldots$, $a, b, \ldots$ and $r, s, \ldots$ indices for orbitals belonging to the core, active and virtual set, respectively, it can easily be shown that eq. (61) is automatically satisfied for any couple of indices belonging to the same set. Furthermore, $\Psi_i^a = (a_a^+ a_i - a_i^+ a_a)\Psi = a_a^+ a_i \, \Psi$, $\Psi_i^r = (a_r^+ a_i - a_i^+ a_r)\Psi = a_r^+ a_i \, \Psi$ and $\Psi_a^r = (a_r^+ a_a - a_a^+ a_r)\Psi = a_r^+ a_a \, \Psi$, so the minimization condition with respect to orbital rotations is reduced to the request that the excitations $a_a^+ a_i \, \Psi$, $a_r^+ a_i \, \Psi$ and $a_r^+ a_a \, \Psi$ do not interact with the CAS–SCF wavefunction via the Hamiltonian operator.

5 The Coupled Clusters method

The coupled clusters method (CC) is based on the so called "exponential ansatz", which states that the exact wavefunction, solution of the Schrödinger equation, can be obtained from the Hartree–Fock wavefunction through the application of a suitable exponential operator, in the form:

$$\Psi = e^{\hat{T}} \Psi_0 \tag{62}$$

Depending upon the H–F wavefunction Ψ_0, the CC method can be applied, besides the ground state, to the calculation of the lowest state of a given (space and/or spin) symmetry. The $\hat{T}$ operator of eq. (62) is expressed as a summation $T = T_1 + T_2 + \ldots + T_n$, where

$$\begin{aligned}
T_1 &= \sum_{i=1}^{n} \sum_{a>n} t_i^a a_a^+ a_i \\
T_2 &= \sum_{i<j}^{n} \sum_{a>b>n} t_{ij}^{ab} a_a^+ a_b^+ a_j \, a_i \\
&\vdots \\
T_n &= \sum_{i<j<k<\ldots} \sum_{a>b>c>\ldots} t_{ijk\ldots}^{abc\ldots} a_a^+ a_b^+ a_c^+ \ldots a_k \, a_j \, a_i
\end{aligned}$$

It is clear that the expansion (62) contains all the possible configurations and, upon a proper choice of the *amplitudes* t_i^a, $t_{ij}^{ab}, \ldots$ the full CI wavefunction can be recovered. The Schrödinger equation can be written as

$$H e^T \Psi_0 = E e^T \Psi_0 \tag{63}$$

Upon application of $\langle \Psi_0 |$ to eq. (63) one obtains the energy in the form

$$E = \left\langle \Psi_0 \left| H \right| e^T \Psi_0 \right\rangle \tag{64}$$

and upon application of $\left\langle \Psi_{ij\ldots}^{ab\ldots} \right|$ (a generic excitation) one gets

$$\left\langle \Psi_{ij\ldots}^{ab\ldots} \left| H \right| e^T \Psi_0 \right\rangle = \left\langle \Psi_0 \left| H \right| e^T \Psi_0 \right\rangle \left\langle \Psi_{ij\ldots}^{ab\ldots} \mid e^T \Psi_0 \right\rangle \tag{65}$$

Developing the exponential operator in the above equation, one obtains a system of non–linear equations from which, in principle, all the amplitudes can be derived. We have seen from Brillouin's theorem (eq. 45 at page 17) that the single excitations Ψ_i^a do not interact with Ψ_0 and, consequently, it is to be expected that the contribution of T_1 is negligible. We shall now approximate T only with the T_2 term ($T =\sim T_2$) and shall try to solve eq. (65). We notice that

$$\Psi =\sim e^{T_2} \Psi_0 = \Psi_0 + T_2 \Psi_0 + \frac{1}{2} T_2^2 \Psi_0 + \frac{1}{3!} T_2^3 \Psi_0 + \ldots$$

and

$$\begin{aligned} T_2^2 \Psi_0 &= \Big(\sum_{i<j} \sum_{a<b} t_{ij}^{ab} a_a^+ a_b^+ a_j \; a_i \;\Big)\Big(\sum_{k<l} \sum_{c<d} t_{kl}^{cd} a_c^+ a_d^+ a_l \; a_k \;\Big) \Psi_0 \\ &= \sum_{\substack{i<j \\ k<l}} \sum_{\substack{a<b \\ c<d}} t_{ij}^{ab} t_{kl}^{cd} a_a^+ a_b^+ a_c^+ a_d^+ a_l \; a_k \; a_j \; a_i \; \Psi_0 \end{aligned}$$

The above equality means that in the expansion of Ψ the quadruple excitations play a role, but in a special form, i.e. as products of double excitations (and the same happens for the sextuples, octuples and so on). Let us now go back to eq. (65) and let us consider the Ψ_{ij}^{ab} double excitation:

$$\left\langle \Psi_{ij}^{ab} \left| H \right| (1 + T_2 + \frac{1}{2} T_2^2) \Psi_0 \right\rangle = \langle \Psi_0 \left| H \right| (1 + T_2) \Psi_0 \rangle \left\langle \Psi_{ij}^{ab} \mid T_2 \Psi_0 \right\rangle \tag{66}$$

We notice that a) on the left hand side of eq. (66) we can truncate the expansion to the second order in T_2 because Ψ_{ij}^{ab} cannot interact with more than a quadruple excitation, b) on the right hand side, in the first term, we truncate to the first order because Ψ_0 cannot interact with more than a double and, in the second term, only $T_2 \Psi_0$ can overlap with Ψ_{ij}^{ab} (giving t_{ij}^{ab} as contribution). The various terms of eq. (66) give the following contributions:

$$\begin{aligned} \left\langle \Psi_{ij}^{ab} \left| H \right| \Psi_0 \right\rangle &= \langle ij \,||\, ab \rangle \\ \left\langle \Psi_{ij}^{ab} \left| H \right| T_2 \Psi_0 \right\rangle &= \sum_{\substack{k<l \\ c<d}} t_{kl}^{cd} \left\langle \Psi_{ij}^{ab} \left| H \right| \Psi_{kl}^{cd} \right\rangle \end{aligned}$$

$$\begin{aligned}
\left\langle \Psi_{ij}^{ab} \left| H \right| T_2^2 \Psi_0 \right\rangle &= \sum_{\substack{k,l,m,n \\ c,d,e,f}} t_{kl}^{cd} t_{mn}^{ef} \left\langle \Psi_{ij}^{ab} \left| H \right| \Psi_{klmn}^{cdef} \right\rangle \\
\left\langle \Psi_0 \left| H \right| \Psi_0 \right\rangle &= E_0^{HF} \\
\left\langle \Psi_0 \left| H \right| T_2 \Psi_0 \right\rangle &= \sum_{\substack{k<l \\ c<d}} t_{kl}^{cd} \left\langle \Psi_0 \left| H \right| \Psi_{kl}^{cd} \right\rangle = \sum_{\substack{k<l \\ c<d}} t_{kl}^{cd} \left\langle kl \,||\, cd \right\rangle \\
\left\langle \Psi_{ij}^{ab} \mid T_2 \Psi_0 \right\rangle &= t_{ij}^{ab}
\end{aligned}$$

Therefore eq. (66) becomes a system of non–linear (quadratic) equations in the amplitudes t_{ij}^{ab}, with one equation for each amplitude; the system is soluble with appropriate iterative procedures (such as Gauss–Jordan's method). This simplified form of coupled cluster theory is termed CCD (Coupled Clusters with Doubles). An improvement which is usually done consists in including also the singles, i.e. $e^T =\sim e^{T_1+T_2}$, obtaining what is called CCSD (CC with singles and doubles). A further improvement would consist in including also the triples but, due to the complexity of the resulting equations, usually just a perturbative treatment of the triples is added, obtaining the CCSD(T). This last form usually gives quite accurate results when applied to molecules which can reasonably well be described by a single determinant (HF). We remark that the CC method is *not* variational and therefore the energy calculated with such a method cannot be expected to be an upper bound to the true energy.

6 Equations–of–motion methods

If the exact ground state wavefunction were known ($|0\rangle$), one could define an "excitation creation" operator (O_n^+), able to generate an excited state $|n\rangle$ when applied to $|0\rangle$

$$|n\rangle = O_n^+ \, |0\rangle \tag{67}$$

An operator with such desired characteristic is obviously given by the shift operator $|n\rangle \langle 0|$. It is also promptly recognized that such excitation operator is by no means unique: for instance, any operator of the form $|n\rangle \langle 0| + \sum_{m\neq 0} |n\rangle \langle m| \, \alpha_m$ would do the job with arbitrary values of the parameters α_m. The idea of the "equations–Of–Motion" Method (EOM) is to express O_n^+ as a suitable expansion of products of creation/destruction operators so as to reproduce reasonably well an excited state even starting from an approximate ground state. The development of the EOM runs as follows: supposing that $|0\rangle$ and $|n\rangle$ are exact, one can write:

$$O_n^+ \, |0\rangle \;\; = \;\; |n\rangle \tag{68}$$

$$\begin{aligned} HO_n^+\,|0\rangle &= E_nO_n^+\,|0\rangle \qquad (69)\\ O_n^+H\,|0\rangle &= E_0O_n^+\,|0\rangle \qquad (70)\end{aligned}$$

Subtracting eq. (70) from eq. (69), one gets

$$[H,O_n^+]\,|0\rangle = \Delta E_{0n}O_n^+\,|0\rangle \qquad (71)$$

Taking now into account the adjoint of operator O_n^+ (O_n), we can apply $\langle 0|\,O_n$ to both sides of eq. (71), which leads to

$$\Delta E_{0n} = \frac{\langle 0\,|O_n[H,O_n^+]|\,0\rangle}{\langle 0\,|O_nO_n^+|\,0\rangle} \qquad (72)$$

Eq. (72) is one form of the EOM method, providing us with a *functional* yielding the excitation energies for a given approximation of the ground state. Other forms are anyway possible if one thinks of the nature of the de-excitation operator O_n. In the form $O_n^+ = |n\rangle\,\langle 0|$ one would have $O_n = |0\rangle\,\langle n|$, from which $O_n\,|n\rangle = |0\rangle$ and $O_n\,|0\rangle = 0$ are immediately obtained. The first of these equations shows the "de-excitation" nature of O_n. The second one, valid for the exact ground state, is called the "killer condition" and allows one to substitute eq. (72) with an equivalent one:

$$\Delta E_{0n} = \frac{\langle 0\,|[O_n,[H,O_n^+]]_\pm|\,0\rangle}{\langle 0\,|[O_n,O_n^+]_\pm|\,0\rangle} \qquad (73)$$

Eq. (73) provides an alternative functional of the EOM method which, as will be shown later on, can be computationally simpler than eq. (72). The ± sign appearing in eq. (73) means that one is formally free to choose either the commutator or the anticommutator: it is convenient to adopt the commutator (sign $-$) when the operators are boson–like, i.e. are made up of products of an *even* number of creation/destruction operators and to adopt the anticommutator in the opposite case of fermion–like operators, products of an odd number of creation/destruction operators.

We shall now consider a simple application of eq. (72) to the calculation of the ionization potentials. We approximate the O_n^+ operator in the form $O_n^+ = \sum\limits_r c_r^* a_r$ so that eq. (72) becomes:

$$\Delta E_{0n} = \frac{\sum_{rs} c_r^* c_s\,\langle 0\,|a_s^+[H,a_r\,]|\,0\rangle}{\sum_{rs} c_r^* c_s\,\langle 0\,|a_s^+a_r\,|\,0\rangle} \qquad (74)$$

Introducing the two matrices $\mathbf{R}$ and $\mathbf{K}$ with elements $R_{rs} = \langle 0\,|a_s^+a_r\,|\,0\rangle$ (a representation of the *one–particle density matrix*) and $K_{rs} = \langle 0\,|a_s^+[H,a_r\,]|\,0\rangle$ (Koopmans' matrix), one has:

$$\Delta E_{0n} = \frac{\mathbf{c}^\dagger\mathbf{K}\mathbf{c}}{\mathbf{c}^\dagger\mathbf{R}\mathbf{c}} \qquad (75)$$

The preceding equation has the well known form of the linear variational principle for the determination of the **c** vector that minimizes the energy E_n and leads immediately to the generalized eigenvalue equation

$$\mathbf{Kc} = \Delta E_{0n}\mathbf{Rc} \tag{76}$$

a result known as "Extended Koopmans' theorem". Let us evaluate the commutator that appears in the definition of **K**:

$$\begin{aligned}[H, a_r\,] &= \sum_{tu} h_{tu}[a_t^+ a_u\,, a_r\,] + \frac{1}{2}\sum_{tuvw}\langle tu \mid vw\rangle\,[a_t^+ a_u^+ a_w a_v\,, a_r\,] \\ &= -\sum_{u} h_{ru} a_u\, - \sum_{tvw}\langle tr \mid vw\rangle\, a_t^+ a_v\, a_w\end{aligned}$$

where we have made use of the commutators $[a_t^+ a_u\,, a_r\,] = -\delta_{tr} a_u$ and $[a_t^+ a_u^+ a_w a_v\,, a_r\,] = \delta_{ur} a_t^+ a_w a_v\, - \delta_{tr} a_u^+ a_w a_v$ (as the knowledgeable reader will take care to verify). So one arrives at the following form:

$$K_{rs} = \left\langle 0 \left| a_s^+ [H, a_r\,] \right| 0 \right\rangle = -\left(\sum_u h_{ru} R_{us} + \sum_{tvw}\langle tr \mid vw\rangle\, R_{wv,st}\right) \tag{77}$$

where we have introduced the representation of the two–particle density matrix $R_{wv,st} = \left\langle 0 \left| a_s^+ a_t^+ a_v\, a_w \right| 0 \right\rangle$ It is interesting to see what is the expression of **K** for the case in which we approximate $|0\rangle$ with the Hartree–Fock wavefunction. In such a case $R_{rs} = 0$ unless both indices refer to occupied orbitals ($R_{ij} = \delta_{ij}$) and $R_{wv,st}$ must also have its four indices referring to occupied orbitals ($R_{ij,ij} = 1$ and $R_{ij,ji} = -1$ with $i \neq j$), so that the following expression for K_{ij} is readily seen to hold:

$$K_{ij} = -h_{ij} - \sum_{t=1}^{n}(\langle ti \mid tj\rangle - \langle ti \mid jt\rangle) = -F_{ij} = \epsilon_i \delta_{ij}$$

In this case the Koopmans' matrix is diagonal and its eigenvalues are equal to minus the orbital energies, recovering the already known Koopmans' theorem, eq (53) of page 19.

We notice that in the general case, when the ground state $|0\rangle$ is approximate in some form, it is not guaranteed that the **K** matrix is hermitian. Defining the auxiliary matrix

$$\gamma_{rs} = \left\langle 0 \left| a_s^+ H a_r\, \right| 0 \right\rangle = \left\langle 0 \left| a_s^+ a_r\, H \right| 0 \right\rangle + K_{rs}$$

which is easily recognized as hermitian (reader, please verify), one gets (for real functions):

$$K_{rs} - K_{sr} = \left\langle 0 \left| (a_r^+ a_s\, - a_s^+ a_r\,) H \right| 0 \right\rangle$$

From the above formula one can see that the $\mathbf{K}$ matrix is hermitian if $|0\rangle$ is given by a MCSCF wavefunction, by virtue of the extended Brillouin's theorem.

We shall now utilize the formulation of the EOM method given by eq. (73), again considering the case of the ionization potentials with $O_n^+ = \sum_r c_r^* a_r$. Now, the form of ΔE_{0n} is given by:

$$\begin{aligned}
\Delta E_{0n} &= \frac{\sum_{rs} c_r^* c_s \left\langle 0 \left| [a_s^+, [H, a_r\,]]_+ \right| 0 \right\rangle}{\sum_{rs} c_r^* c_s \left\langle 0 \left| [a_s^+, a_r\,]_+ \right| 0 \right\rangle} \\
&= \frac{\mathbf{c}^\dagger \mathbf{K}' \mathbf{c}}{\mathbf{c}^\dagger \mathbf{c}}
\end{aligned}$$

with $K'_{rs} = \langle 0 \,|[a_s^+, [H, a_r\,]]_+|\, 0\rangle$ Now we have the steps:

$$\begin{aligned}
[H, a_r\,] &= -\sum_u h_{ru} a_u - \sum_{tvw} \langle tr \mid vw\rangle \, a_t^+ a_v \, a_w \\
\left[a_s^+, [H, a_r\,]\right]_+ &= -\sum_u h_{ru} [a_s^+, a_u\,]_+ - \sum_{tvw} \langle tr \mid vw\rangle \, [a_s^+, a_t^+ a_v \, a_w]_+ \\
K'_{rs} &= -h_{rs} - \sum_{tv} \langle tr \,||\, vs\rangle \, R_{vt}
\end{aligned}$$

where use has been made of the anticommutator $[a_s^+, a_t^+ a_v \, a_w]_+ = -\delta_{sv} a_t^+ a_w + \delta_{sw} a_t^+ a_v$. The $\mathbf{K}'$ matrix is simpler than the Koopmans' matrix $\mathbf{K}$ previously defined, because the former only involves the one–particle density matrix.

Let us now pass on to an application of eq. (73) involving the calculation of excited states. Before doing that, it is convenient to enforce hermiticity in the operator at the numerator of eq. (73), by substituting it with the symmetric double commutator

$$[O_n, H, O_n^+]_\pm = \frac{1}{2}([[O_n, H], O_n^+]_\pm + [O_n, [H, O_n^+]]_\pm)$$

such step being justified by the consideration that, for real functions, taking into account the definition of adjoint operator

$$\left\langle 0 \left| [O_n, [H, O_n^+]]_\pm \right| 0 \right\rangle = \left\langle 0 \left| [[O_n, H], O_n^+]_\pm \right| 0 \right\rangle$$

Thus, eq. (73) can be substituted with the new form:

$$\Delta E_{0n} = \frac{\langle 0 \,|[O_n, H, O_n^+]_\pm|\, 0\rangle}{\langle 0 \,|[O_n, O_n^+]_\pm|\, 0\rangle} \tag{78}$$

We shall now assume O_n^+ to be in the form $O_n^+ = \sum\limits_{r \neq s} c_{rs} E_{rs}$ and shall utilize the minus sign (i.e. the commutator) in eq. (78). Proceeding along exactly the same lines as those already met for the ionization case, one immediately arrives at an eigenvalue equation

$$\mathbf{Ac} = \Delta E_{0n} \mathbf{Sc} \tag{79}$$

where $\mathbf{c}$ is a column vector collecting the c_{rs} coefficients, $\mathbf{S}$ and $\mathbf{A}$ are square matrices with elements $S_{r's',rs} = \left\langle 0 \left| [E^+_{r's'}, E_{rs}] \right| 0 \right\rangle$ and $A_{r's',rs} = \left\langle 0 \left| [E^+_{r's'}, H, E_{rs}] \right| 0 \right\rangle$. In order to better elucidate the nature of such equation, it is convenient to rewrite the expression of the excitation operator in the form $O_n^+ = \sum_{r>s} (X_{rs} E_{rs} - Y_{rs} E^+_{rs})$ (which is reminiscent of the expression of an orbital rotation, even though no request for antihermiticity is here asked for) or, indicating the couple (r, s) with μ: $O_n^+ = \sum_\mu (X_\mu E_\mu - Y_\mu E^+_\mu)$. In this way the matrices $\mathbf{A}$ and $\mathbf{S}$ can be partitioned in four submatrices: $\mathbf{A_{11}} = \mathbf{M}$, with $M_{\mu\nu} = \left\langle 0 \left| [E^+_\mu, H, E_\nu] \right| 0 \right\rangle$, $\mathbf{A_{22}} = \mathbf{M}'$, with $M'_{\mu\nu} = \langle 0 \left| [E_\mu, H, E^+_\nu] \right| 0 \rangle = M^*_{\mu\nu}$, $\mathbf{A_{12}} = \mathbf{Q}$ with $Q_{\mu\nu} = -\left\langle 0 \left| [E^+_\mu, H, E^+_\nu] \right| 0 \right\rangle$, $\mathbf{A_{21}} = \mathbf{Q}'$ with $Q'_{\mu\nu} = Q^*_{\mu,\nu}$. For the $\mathbf{S}$ matrix one has, analogously: $\mathbf{S_{11}} = \mathbf{V}$ with $V_{\mu\nu} = \left\langle 0 \left| [E^+_\mu, E_\nu] \right| 0 \right\rangle$, $\mathbf{S_{22}} = -\mathbf{V}^*$, $\mathbf{S_{12}} = \mathbf{W}$ with $W_{\mu\nu} = -\left\langle 0 \left| [E^+_\mu, E^+_\nu] \right| 0 \right\rangle$, $\mathbf{S_{21}} = -\mathbf{W}^*$. With such replacements the eigenvalue eq. (79) takes on the form

$$\begin{bmatrix} \mathbf{M} & \mathbf{Q} \\ \mathbf{Q}^* & \mathbf{M}^* \end{bmatrix} \begin{bmatrix} \mathbf{X} \\ \mathbf{Y} \end{bmatrix} = \Delta E_{0n} \begin{bmatrix} \mathbf{V} & \mathbf{W} \\ -\mathbf{W}^* & -\mathbf{V}^* \end{bmatrix} \begin{bmatrix} \mathbf{X} \\ \mathbf{Y} \end{bmatrix} \tag{80}$$

The simplest application of eq. (80) concerns the case of a Hartree–Fock approximation to the ground state $|0\rangle$. Of the eight components of the double symmetric commutator which enter the construction of matrices $\mathbf{M}$ and $\mathbf{Q}$ only those of type $E^+_\mu H E_\nu$ and $E^+_\mu E^+_\nu H$ give contributions, yielding:

$$\begin{aligned} M_{bj,ai} &= \delta_{ab}\delta_{ij}(\epsilon_a - \epsilon_i) - \langle bi \,||\, aj \rangle && (81) \\ Q_{bj,ai} &= \langle ab \,||\, ij \rangle && (82) \end{aligned}$$

Furthermore, $\mathbf{W} = \mathbf{0}$ and $\mathbf{V} = \mathbf{1}$. The resulting reformulation of eq. (80) is called the *Random Phase Approximation* (RPA). It should be remarked that the RPA equation is not a usual generalized eigenvalue equation, because the metric matrix is not positive definite. In case all the quantities appearing in the RPA equation are real, one can write:

$$\begin{aligned} \mathbf{MX} + \mathbf{QY} &= \lambda \mathbf{X} \\ \mathbf{QX} + \mathbf{MY} &= -\lambda \mathbf{Y} \end{aligned}$$

or, equivalently (summing and subtracting)

$$\begin{aligned}(\mathbf{M}+\mathbf{Q})(\mathbf{X}+\mathbf{Y}) &= \lambda(\mathbf{X}-\mathbf{Y})\\ (\mathbf{M}-\mathbf{Q})(\mathbf{X}-\mathbf{Y}) &= \lambda(\mathbf{X}+\mathbf{Y})\end{aligned}$$

Substituting $\mathbf{X}-\mathbf{Y}$ from the second equation into the first, gives:

$$(\mathbf{M}+\mathbf{Q})(\mathbf{X}+\mathbf{Y}) = \lambda^2(\mathbf{M}-\mathbf{Q})^{-1}(\mathbf{X}+\mathbf{Y}) \tag{83}$$

and this generalized eigenvalue equation can be solved with the usual techniques, provided that the metric $(\mathbf{M}-\mathbf{Q})^{-1}$ is positive definite, which is usually the case.

7 Perturbation methods

Rayleigh–Schrödinger perturbation theory (RSPT) can be successfully applied to improve the results obtained at a certain level of approximation such as Hartree–Fock or MCSCF. In particular the application of RSPT to a zero order wavefunction obtained by solving the H–F equations bears the name of Møller–Plesset PT and has revealed itself as very successful in providing a large fraction of the correlation energy. We shall here limit ourselves to recalling the working formulas of RSPT and shall consider the applications to a H–F determinant and to a multireference wavefunction in the next paragraphs.

In RSPT the wavefunction and energy for the eigenvalue problem $H\Psi_n = E_n\Psi_n$ are expressed as series expansions

$$\begin{aligned}\Psi_n &= \Psi_n^{(0)} + \Psi_n^{(1)} + \Psi_n^{(2)} + \ldots\\ E_n &= E_n^{(0)} + E_n^{(1)} + E_n^{(2)} + \ldots\end{aligned}$$

where the k–th term in either equation above is called the k–th order correction to the wavefunction or to the energy, respectively. The Hamiltonian is partitioned according to $H = H_0 + V$, with V playing the role of a perturbation operator to a zero order Hamiltonian H_0 for which the eigenvalues ($E_n^{(0)}$) and eigenfunctions are known ($\Psi_n^{(0)}$). In order to obtain the k–th order correction to the energy one has

$$E_n^{(k)} = \left\langle \Psi_n^{(0)} \left| V \right| \Psi_n^{(k-1)} \right\rangle$$

The results up to the second order in the energy and to the first in the wavefunction are given by:

$$E_n^{(1)} = \left\langle \Psi_n^{(0)} \left| V \right| \Psi_n^{(0)} \right\rangle$$

$$
\begin{aligned}
E_n^{(2)} &= -\sum_{k\neq n} \frac{|\left\langle \Psi_n^{(0)} \left| V \right| \Psi_k^{(0)} \right\rangle|^2}{E_k^{(0)} - E_n^{(0)}} \\
\Psi_n^{(1)} &= -\sum_{k\neq n} \frac{\left\langle \Psi_n^{(0)} \left| V \right| \Psi_k^{(0)} \right\rangle}{E_k^{(0)} - E_n^{(0)}} \Psi_k^{(0)}
\end{aligned}
$$

We shall now apply the above results to the case of a single reference (H–F) and then to a multireference (MCSCF) zero order wavefunction.

7.1 Møller–Plesset theory

If Ψ_0 is a Slater determinant whose orbitals have been optimized by solving the Hartree–Fock equations $\hat{f}\psi_i = \epsilon_i\psi_i$, one can build the n–particle Fock operator (also termed the *Fockian*) $\hat{F} = \sum_{i=1}^{n} \hat{f}(i)$ which, in the language of second quantization, is turned into $\hat{F} = \sum_{r} \epsilon_r a_r^+ a_r$. It is then recognized that Ψ_0 is an eigenfunction of $\hat{F}$ with eigenvalue $E_0 = \sum_{i=1}^{n} \epsilon_i$. Generally, any other determinant Ψ_K built with n arbitrary spin–orbitals $(||\psi_{k_1}\psi_{k_2}\cdots\psi_{k_n}||)$ is still eigenfunction of $\hat{F}$ with eigenvalue $E_K = \sum_{i=1}^{n} \epsilon_{k_i}$. Møller–Plesset PT utilizes $\hat{F}$ as zero order Hamiltonian and $\hat{V} = \hat{H} - \hat{F}$ as perturber (also called the *fluctuation* potential). Thus, the perturbation formulas for the first two orders are:

$$
\begin{aligned}
E_0^{(0)} + E_0^{(1)} &= \left\langle \Psi_0 \left| H \right| \Psi_0 \right\rangle = E_0^{\mathrm{HF}} \\
E_0^{(2)} &= -\sum_{K\neq 0} \frac{|\left\langle \Psi_0 \left| H \right| \Psi_K \right\rangle|^2}{E_K - E_0}
\end{aligned}
$$

In the above formula the determinants expressing single substitutions do not give any contribution due to Brillouin's theorem ($\left\langle \Psi_0 \left| H \right| \Psi_i^a \right\rangle = 0$) and consequently the only contribution is obtained from the double replacements Ψ_{ij}^{ab} so that the MP2 result (Møller–Plesset to second order) is

$$
E_0^{(2)} = -\sum_{i=1}^{n} \sum_{a>n} \frac{|\left\langle ab \,||\, ij \right\rangle|^2}{\epsilon_a + \epsilon_b - \epsilon_i - \epsilon_j}
$$

The MP2 theory is very simple, requires little computational effort over a H–F calculation and usually yields a large fraction of the correlation energy. Even though we have derived MP2 for the ground state determinant, it can be successfully applied to all cases where a single determinant is a good approximation for the description of the molecule under study. This usually happens for closed shell molecules near the equilibrium geometry.

7.2 Multireference perturbation theory

The success of Møller–Plesset perturbation theory fostered the researchers to extend it to zero order multireference wavefunctions which are necessary both for a correct description of the ground state in situations deprived of closed shell nature (such as bond breaking, for instance) and especially for excited states which can seldom be satisfactorily described by a single determinant.

We shall be concerned only with the application of PT to a variational zero order wavefunction, a situation which is sometimes referred to as "diagonalize then perturb", limiting ourselves to barely recalling that also the opposite approach ("perturb then diagonalize") has been intensely studied.

The main difficulty in multireference perturbation theory (MRPT) arises when one is confronted with the definition of a good zero order Hamiltonian, which should aim at guaranteeing some important properties such as *size consistence* and absence of *intruder states*. Size consistence (or "strict separability") requires that the energy of a system made up of two non interacting parts should yield the same result as the sum of the energies of the two separate parts. With the term "intruder state" one indicates a zero order state quasi degenerate with the one under study, causing divergences in the expansion of the second order correction to the energy.

In MRPT one starts from a variational wavefunction built upon a given determinantal space S

$$\Psi_m^{(0)} = \sum_{K \in S} c_K \Phi_K$$

with

$$PHP\Psi_m^{(0)} = E_m^{(0)} \Psi_m^{(0)}$$

where P is the projector onto the S space

$$P = \sum_{K \in S} |\Phi_K\rangle \langle \Phi_K|$$

The rest of the space spanned by the determinants $\Phi_K \notin S$ is referred to as the *outer* space. In the early treatments of MRPT simple Slater determinants were chosen as zero order wavefunctions in the outer space providing the perturbation corrections. For instance in the CIPSI method (developed in 1973) the zero order Hamiltonian in the so called Møller–Plesset barycentric partition (MPB) is given by:

$$
\begin{aligned}
H_0\Psi_m^{(0)} &= \left\langle \Psi_m^{(0)} \left| F \right| \Psi_m^{(0)} \right\rangle \Psi_m^{(0)} \quad \text{within the } S \text{ space} \\
H_0\Phi_K &= (\sum_{i=1}^{n} \epsilon_{k_i})\Phi_K \quad \text{outside } S
\end{aligned}
$$

The F operator is a suitable Fock–like one–electron operator $F = \sum_r \epsilon_r a_r^+ a_r$ with the energies ϵ_r usually chosen as

$$\epsilon_r = h_{rr} + \sum_s n_s \left\langle rs \,||\, rs \right\rangle$$

n_s being the occupation number of spin–orbital ψ_s. The ψ_s spin–orbitals are here assumed to be *natural* spin–orbitals, i.e. diagonalizing the one–particle density matrix ($\left\langle \Psi_m^{(0)} \left| a_r^+ a_s \right| \Psi_m^{(0)} \right\rangle$). Another possibility used in CIPSI is to adopt an Epstein–Nesbet partition with $H_0\Psi_m^{(0)} = E_m^{(0)}\Psi_m^{(0)}$ inside S and $H_0\Phi_K = \left\langle \Phi_K \left| H \right| \Phi_K \right\rangle \Phi_K$ outside S.

In the more recent treatments of MRPT the zero order functions outside the S space are not assumed to be simple determinants but, rather, are built upon *contracted* excitations carried out on the variational wavefunction, i.e. they are of the type $a_r^+ a_s^+ a_u a_t \ \Psi_m^{(0)}$. In the CAS–PT2 approach (1990) one starts from a CAS–SCF variational wavefunction and the zero order Hamiltonian is built by projection of a suitable generalized Fock operator onto the space generated by the contracted excitations. The resulting perturbation theory is very efficient and usually produces high quality results for both the ground and the excited states. A shortcoming of CAS–PT2 lies in the fact that, due to the one–electron nature of H_0, it is occasionally subjected to the occurrence of intruder states.

A recent formulation of MRPT also applied to a CAS–SCF wavefunction makes use of a two–electron zero order Hamiltonian, is exempt from intruder states and is entirely size consistent. The new theory (2001), called "n–electron valence state perturbation theory" (NEV–PT) makes use of zero order wavefunctions that are of CAS–CI type both inside the S space and outside of it. The state of interest Ψ_m is approximated at the zero order by a CAS–SCF (or more generally CAS–CI) wave function $\Psi_m^{(0)}$ obtained, as usual, by solving the variational eigenvalue problem $PHP\Psi_m^{(0)} = E_m^{(0)}\Psi_m^{(0)}$. The zero order wavefunctions different from those of the CAS space are referred to as "perturber functions" and belong to CAS–CI spaces characterized by well defined occupation patterns of the orbitals. One such CAS–CI space will be designated by $S_l^{(k)}$, where k is the number of electrons promoted to (or removed from) the active orbital space ($-2 \leq k \leq 2$ for a second order treatment) and l describes the fixed occupation pattern of the inactive (core + virtual) orbitals. A generic perturber function belonging to $S_l^{(k)}$ will be denoted by $\Psi_{l,\mu}^{(k)}$ where μ enumerates the various functions in $S_l^{(k)}$. It is possible to define different variants of the NEV–PT2 according to the number of perturber functions that are chosen from the $S_l^{(k)}$ subspaces: if the full dimensionality of $S_l^{(k)}$ is exploited one has the "totally uncontracted" case and if just one perturber is selected we speak of the "strongly contracted" case; between these two extremes a "partially contracted" case can also be defined. All three variants can be shown to be endowed with the desirable properties of strict separability and absence of intruder states. Up to now the strongly contracted and partially contracted cases have been implemented in an efficient code and have been shown to yield accurate results for both the ground and the excited electronic states.

The subspaces $S_l^{(k)}$ which contain the perturber wave functions can be classified in 8 distinct types, according to the number k of electrons promoted into the active orbital space and to the occupation of inactive orbitals. With $k = 0$ two types are possible according to whether two or one core electrons have been transferred to the virtual orbital space; a typical representative of the first type will be indicated as $S_{ij,rs}^{(0)}$, meaning that two electrons have been moved from the core orbitals φ_i and φ_j to the virtual orbitals φ_r and φ_s. The second type with $k = 0$ will contain all subspaces $S_{i,r}^{(0)}$ with one electron moved from φ_i to φ_r. Analogously, with $k = \pm 1$ two distinct types are possible, containing subspaces $S_{ij,r}^{(1)}$ ($S_{i,rs}^{(-1)}$) and $S_i^{(1)}$ ($S_r^{(-1)}$), respectively. With $k = \pm 2$ only one type is possible, containing the subspaces $S_{ij}^{(2)}$ ($S_{rs}^{(-2)}$). In a given subspace $S_l^{(k)}$ the strongly contracted NEV–PT (SC–NEV–PT) considers only one perturber wave function, according to the following definition:

$$\Psi_l^{(k)} = P_{S_l^{(k)}} H \Psi_m^{(0)} \tag{84}$$

where $P_{S_l^{(k)}}$ is the projector onto subspace $S_l^{(k)}$. The perturber functions $\Psi_l^{(k)}$ are orthogonal but are not normalized to unity. The squared norm

$$N_l^{(k)} = \left\langle \Psi_l^{(k)} \mid \Psi_l^{(k)} \right\rangle = \left\langle \Psi_m^{(0)} \left| H P_{S_l^{(k)}} H \right| \Psi_m^{(0)} \right\rangle$$

plays an important role in the strongly contracted NEV–PT. Once a suitable energy $E_l^{(k)}$ has been attributed to $\Psi_l^{(k)}$ the zero order hamiltonian can be defined through the spectral decomposition:

$$H_0 = \sum_{l,k} \left| \Psi_l^{(k)\prime} \right\rangle E_l^{(k)} \left\langle \Psi_l^{(k)\prime} \right| + \sum_m^{CAS} \left| \Psi_m^{(0)} \right\rangle E_m^{(0)} \left\langle \Psi_m^{(0)} \right| \tag{85}$$

where $\left|\Psi_l^{(k)\prime}\right\rangle = \frac{1}{\sqrt{N_l^{(k)}}}\Psi_l^{(k)}$. The expressions for the first order correction to the wave function and for the second order correction to the energy are then:

$$\begin{aligned}
\Psi_m^{(1)} &= \sum_{k,l} \left|\Psi_l^{(k)\prime}\right\rangle \frac{\left\langle \Psi_l^{(k)\prime} \left| H \right| \Psi_m^{(0)} \right\rangle}{E_m^{(0)} - E_l^{(k)}} && (86) \\
&= \sum_{kl} \left|\Psi_l^{(k)\prime}\right\rangle \frac{1}{\sqrt{N_l^{(k)}}} \frac{\left\langle \Psi_l^{(k)} \left| H \right| \Psi_m^{(0)} \right\rangle}{E_m^{(0)} - E_l^{(k)}} \\
&= \sum_{kl} \left|\Psi_l^{(k)\prime}\right\rangle \frac{\sqrt{N_l^{(k)}}}{E_m^{(0)} - E_l^{(k)}} \\
E_m^{(2)} &= \sum_{k,l} \frac{\left\langle \Psi_l^{(k)\prime} \left| H \right| \Psi_m^{(0)} \right\rangle^2}{E_m^{(0)} - E_l^{(k)}} && (87) \\
&= \sum_{kl} \frac{N_l^{(k)}}{E_m^{(0)} - E_l^{(k)}}
\end{aligned}$$

For the definition of the energies $E_l^{(k)}$ one can use either the true Hamiltonian $\hat{H}$ or an approximation of it. From the computational point of view, a particularly convenient form is provided by Dyall's approximation to the electronic hamiltonian:

$$H^D = H_i + H_v \tag{88}$$

where H_i is a simple one–electron (diagonal) operator

$$H_i = \sum_i^{core} \epsilon_i \mathcal{E}_{ii} + \sum_r^{virt} \epsilon_r \mathcal{E}_{rr} + C$$

and H_v is a two–electron operator limited to the active space

$$H_v = \sum_{ab}^{act} h_{ab}^{\text{eff}} \mathcal{E}_{ab} + \frac{1}{2} \sum_{abcd}^{act} \langle ab \mid cd \rangle \left(\mathcal{E}_{ac}\mathcal{E}_{bd} - \delta_{bc}\mathcal{E}_{ad} \right)$$

with $h_{ab}^{\text{eff}} = h_{ab} + \sum_j^{core} (2 \langle aj \mid bj \rangle - \langle aj \mid jb \rangle)$. An appropriate choice of the constant C

$$C = 2 \sum_i^{core} h_{ii} + \sum_{ij}^{core} (2 \langle ij \mid ij \rangle - \langle ij \mid ji \rangle) - 2 \sum_i^{core} \epsilon_i$$

ensures that H^D is equivalent to the full hamiltonian within the CAS space:

$$H^D \Psi_m^{(0)} = E_m^{(0)} \Psi_m^{(0)}$$

The energies ϵ_i and ϵ_r are suitable orbital energies which we choose to be:

$$\begin{aligned} \epsilon_i &= -\left\langle a_i\ \Psi_m^{(0)} \left| H \right| a_i\ \Psi_m^{(0)} \right\rangle + E_m^{(0)} \\ \epsilon_r &= \left\langle a_r^+ \Psi_m^{(0)} \left| H \right| a_i^+ \Psi_m^{(0)} \right\rangle - E_m^{(0)} \end{aligned}$$

The energies associated to the perturber functions $\Psi_l^{(k)}$ are defined *via* Dyall's hamiltonian:

$$E_l^{(k)} = \frac{1}{N_l^{(k)}} \left\langle \Psi_l^{(k)} \left| H^D \right| \Psi_l^{(k)} \right\rangle$$

Such definition, while guaranteeing that the interactions among the active electrons are properly taken into consideration, affords a particularly convenient formulation.

The "partially contracted" approach of NEV–PT hinges on the definition of a suitable subspace $\bar{S}_l^{(k)} \subset S_l^{(k)}$ and utilizes the eigenfunctions of the hamiltonian in such a subspace.

In order to define the $\bar{S}_l^{(k)}$ subspaces of the PC–NEV–PT, the action upon $\Psi_m^{(0)}$ of the strings of excitation operators present in $\hat{H}$ is considered, in the same way already seen for the contracted excitations of CAS–PT2. Thus a set of functions $\Phi_{l,m}^{(k)}$ is generated which serve as basis for the $\bar{S}_l^{(k)}$ subspace; the collective index m designates the active orbital indices involved in the excitation operators.

In case the $\Phi_{l,m}^{(k)}$ functions should not be linearly independent, a new suitable set will be built by the usual technique of removing the zero eigenvalues from the overlap matrix $\mathbf{M}_l^{(k)} = \left\langle \mathbf{\Phi}_l^{(k)} \mid \mathbf{\Phi}_l^{(k)} \right\rangle$, where $\mathbf{\Phi}_l^{(k)}$ is a row vector collecting the $\Phi_{l,m}^{(k)}$ functions.

We shall denote by $\Phi_{l,m}^{(k)\prime}$ the orthonormalized basis functions for the $\bar{S}_l^{(k)}$ subspaces and by $\Psi_{l,\mu}^{(k)}$ and $E_{l,\mu}^{(k)}$ the eigenfunctions and eigenvalues of the projection of the electronic hamiltonian H onto $\bar{S}_l^{(k)}$:

$$P_{\bar{S}_l^{(k)}} H P_{\bar{S}_l^{(k)}} \Psi_{l,\mu}^{(k)} = E_{l,\mu}^{(k)} \Psi_{l,\mu}^{(k)}. \tag{89}$$

Such eigenfunctions and eigenvalues will serve to define the zero order hamiltonian H_0 through its spectral decomposition

$$H_0 = \sum_m^{CAS} \left| \Psi_m^{(0)} \right\rangle E_m^{(0)} \left\langle \Psi_m^{(0)} \right| + \sum_{k,l,\mu} \left| \Psi_{l,\mu}^{(k)} \right\rangle E_{l,\mu}^{(k)} \left\langle \Psi_{l,\mu}^{(k)} \right| \tag{90}$$

Actually, it is more convenient from a computational point of view to utilize in formulas (89) Dyall's approximation to the hamiltonian, already introduced in the strongly contracted case.

In spite of the larger number of pertuber functions employed in the partially contracted approach with respect to the strongly contracted one, the results obtained with the two variants have turned out to be very similar in all the test calculations performed thus far.

Further reading

Excellent advanced textbooks that cover practically all the topics dealt with in these lessons are:

- R. McWeeny, “Methods of Molecular Quantum Mechanics”, Academic Press, 1989.
- T. Helgaker, P. Jørgensen, J. Olsen, “Molecular Electronic–Structure Theory”, John Wiley & Sons, 2000.
- P. Jørgensen, J. Simons, “Second Quantization–based Methods in Quantum Chemistry”, Academic Press, 1981.

A few articles referring to the multireference perturbation theory exposed in section 7 are:

CIPSI B. Huron, P. Rancurel, J.P. Malrieu, *J. Chem. Phys.*, vol. 58, p. 5745 (1973)

CASPT2 K. Andersson, P. Malmqvist, B.O. Roos, A.J. Sadlej, K. Wolinski, *J. Phys. Chem.*, vol. 94, p. 5483 (1990)

CASPT2 K. Andersson, P. Malmqvist, B.O. Roos, *J. Chem. Phys*, vol. 96, p. 1218 (1992)

NEVPT2 C. Angeli, R. Cimiraglia, S. Evangelisti, T. Leininger, J.P. Malrieu, *J. Chem. Phys.*, vol. 114, p. 10252 (2001)

NEVPT2 C. Angeli, R. Cimiraglia, J.P. Malrieu, *Chem. Phys. Letters*, vol. 350, p. 297 (2001)