

MOLECULAR QUANTUM MECHANICS

Pekka Manninen

`manninen@chem.helsinki.fi`

`http://www.chem.helsinki.fi/~manninen/mqm06.html`

Lectures: (24 h) Mon 12-14, Tue 12-14 in A121 + discussions Fri 15-16 in A406. No lectures on 17 April and 1 May.

Exercises: (14 h) Wed 14-16 in A121

Literature: Chapters 1–5 (6, 8–10) of R. McWeeny, *Methods of Molecular Quantum Mechanics* (Academic Press 1992); Chapters 1–7 (10–14) of T. Helgaker, P. Jørgensen, and J. Olsen, *Molecular Electronic-Structure Theory* (Wiley 2002).

6 ECTS, one exam (or a project work?); 50% of the “obligatory” exercises (denoted with *) should be returned.

Contents

1 Many-electron wave functions

1.1 Molecular Schrödinger equation

1.2 Electron distribution

1.3 Exact and approximate wave functions

1.4 Atomic basis functions

2 Second quantization

2.1 Fock space

2.2 Spin in second quantization

3 Standard models

3.1 The Hartree–Fock approximation

3.2 Configuration-interaction theory

3.3 Multiconfigurational self-consistent field theory

3.4 Coupled-cluster theory

3.5 Perturbation theory

3.6 Performance of the electronic-structure models

3.7 Linear scaling and other contemporary trends

1 MANY-ELECTRON WAVE FUNCTIONS

1.1 Molecular Schrödinger equation

The *Schrödinger equation* for a system of N electrons is*

$$\hat{H}\Psi(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N) = E\Psi(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N), \quad (1)$$

where \hat{H} is the molecular *Hamiltonian*, which is for a molecule that contains N electrons and K nuclei (assumed to be fixed in space)

$$\hat{H} = \sum_{i=1}^N \hat{h}_i + \sum_{i=1}^N \sum_{j=i+1}^N \hat{g}_{ij}, \quad (2)$$

where \hat{h}_i is the one-electron Hamiltonian of electron i moving in an electrostatic potential caused by the nuclei, and \hat{g}_{ij} is the interaction between the electrons i and j :

$$\hat{h}_i = -\frac{1}{2}\nabla_i^2 - \sum_{I=1}^K \frac{Z_I}{|\mathbf{r}_i - \mathbf{R}_I|}; \quad \hat{g}_{ij} = \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}. \quad (3)$$

The eigenfunctions or *wave functions* $\Psi(\mathbf{x}_1, \dots, \mathbf{x}_N)$ describe the electronic structure and properties of the molecule. $\{\mathbf{x}\}$ are used to collectively symbolize all variables needed to refer to particles; here the position and the internal degree of freedom, *spin*. It is necessary that the wave function is *square-integrable*,

$$\int \Psi^*(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N)\Psi(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N)d\mathbf{x}_1d\mathbf{x}_2 \cdots d\mathbf{x}_N = K \quad (K \text{ finite}). \quad (4)$$

Usually we assume that the wave function is normalized, i.e., $K = 1$.

*In atomic units used throughout this course, where $\hbar = m_e = 4\pi\epsilon_0 \equiv 1$; $c = 1/\alpha \approx 137$.

The corresponding eigenvalues E_K are in the present context the quantized *energy levels* of the electronic “cloud” of the molecule. Any two eigenfunctions that correspond to different energy values E_K, E_L possess the *orthogonality property*

$$\int \Psi_K^* \Psi_L dx_1 dx_2 \cdots dx_N = 0. \quad (5)$$

Even if two or more eigenvalues happen to be identical, corresponding to a degenerate state of the system, the different eigenfunctions may still be assumed orthogonal.

If we would like to include the effects of nuclear motion, the molecular Hamiltonian becomes (now restricting to electrostatic forces only)

$$\begin{aligned} \hat{H} &= \hat{H}_e + \hat{H}_n + \hat{H}_{en} \\ \hat{H}_e &= -\frac{1}{2} \sum_i \nabla_i^2 + \frac{1}{2} \sum_{ij}' \hat{g}_{ij} \\ \hat{H}_n &= -\frac{1}{2} \sum_I \frac{1}{m_I} \nabla_I^2 + \frac{1}{2} \sum_{IJ}' Z_I Z_J \hat{g}_{IJ} \\ \hat{H}_{en} &= -\sum_I \sum_i \frac{Z_I}{r_{Ii}}. \end{aligned} \quad (6)$$

The wave function is correspondingly dependent on both the electron and nuclear variables. Usually, because of the large ratio m_n/m_e , it is a good approximation to separate these, and we may turn our attention to the electronic problem. This is referred to as the Born–Oppenheimer approximation. The nuclear-motion Schrödinger equation plays, however, a central role in molecular spectroscopy. The motion of nuclei can sometimes affect even those properties that are considered electronic. This is the case e.g. in the Jahn–Teller effect and the Renner effect. For fixed nuclei, we can write the total energy of the system simply as

$$E_{\text{tot}} = E_{\text{electronic}} + \frac{1}{2} \sum_{IJ}' \frac{Z_I Z_J}{|\mathbf{R}_I - \mathbf{R}_J|}. \quad (7)$$

Eq. (1) provides a basis for molecular quantum mechanics and all the static electronic properties of the molecule may be obtained by expectation values of the appropriate Hermitian operators or using perturbation theory. However, we should note the limitations of this model:

- Limitations in the Hamiltonian: non-relativistic kinematics, instant interactions, and the Born–Oppenheimer approximation
- Eq. (1) is soluble exactly only for a few trivial one-electron systems, such as hydrogen-like atom or the molecule H_2^+ (see Exercise 1.1)
- Real experiments are not carried out for an isolated molecule, and observation always involves *interaction* with the system, and thus a time-independent description is not sufficient; but the time-independent form of (1), $\hat{H}\Psi = i\partial\Psi/\partial t$, should be applied.

Example 1.1

For a helium-like atom with a point-like nucleus of charge Z the electronic Hamiltonian, Eq. (2), is

$$\hat{H} = \hat{h}_1 + \hat{h}_2 + \hat{g}_{12} = -\frac{1}{2}\nabla_1^2 - \frac{Z}{r_1} - \frac{1}{2}\nabla_2^2 - \frac{Z}{r_2} + \frac{1}{r_{12}}.$$

Due to \hat{g}_{12} , this is a three-body problem, and thereby no closed form solution exists for the eigenvalue equation (1). We start our consideration by an approximation, where \hat{g}_{12} is set to zero, i.e. $\hat{H}_0 = \hat{h}_1 + \hat{h}_2$. We can separate the variables in the eigenvalue equation $H_0\Psi_0 = E\Psi_0$ by substituting[†]

$$\Psi_0 = \Phi(\mathbf{r}_1, \mathbf{r}_2) = \chi_1(\mathbf{r}_1)\chi_2(\mathbf{r}_2),$$

[†]Considering only the variables \mathbf{r} (position vectors), which in this case are most conveniently represented as the variables of spherical coordinates, $\mathbf{r} = (r, \theta, \rho)$

to the eigenvalue equation and dividing by Φ , obtaining

$$\frac{\hat{h}_1\chi_1(\mathbf{r}_1)}{\chi_1(\mathbf{r}_1)} + \frac{\hat{h}_2\chi_2(\mathbf{r}_2)}{\chi_2(\mathbf{r}_2)} = E,$$

meaning that Φ is a solution if (denoting $E = \epsilon_1 + \epsilon_2$)

$$\begin{aligned}\hat{h}_1\chi_1(\mathbf{r}_1) &= \epsilon_1\chi_1(\mathbf{r}_1) \\ \hat{h}_2\chi_2(\mathbf{r}_2) &= \epsilon_2\chi_2(\mathbf{r}_2),\end{aligned}$$

meaning that χ_1 and χ_2 are solutions of the one-electron eigenvalue problem $\hat{h}\chi = \epsilon\chi$. Φ is a *simultaneous eigenstate* of the operators \hat{h}_1 and \hat{h}_2 ,

$$\begin{aligned}\hat{h}_1\Phi(\mathbf{r}_1, \mathbf{r}_2) &= \epsilon_1\Phi(\mathbf{r}_1, \mathbf{r}_2) \\ \hat{h}_2\Phi(\mathbf{r}_1, \mathbf{r}_2) &= \epsilon_2\Phi(\mathbf{r}_1, \mathbf{r}_2).\end{aligned}$$

The 1-electron 1-center eigenfunctions that are solutions to the one-electron eigenvalue problem (this holds only in the absence of \hat{g} !) are the same with the solutions of the Schrödinger equation for a hydrogen-like atom (note the factor Z),

$$\chi_{nlm}(r, \theta, \rho) = R_{nl}(r)Y_{lm}(\theta, \rho), \quad (8)$$

consisting of the radial part

$$R_{nl}(r) = \left(\frac{2Z}{n}\right)^{3/2} \sqrt{\frac{(n-l-1)!}{2n(n+1)!}} \left(\frac{2Zr}{n}\right)^l L_{n-l-1}^{2l+1}\left(\frac{2Zr}{n}\right) \exp\left(-\frac{Zr}{n}\right) \quad (9)$$

and a spherical harmonic function, explicitly expressed as

$$Y_{lm}(\theta, \rho) = \sqrt{\frac{2l+1}{4\pi} \frac{(l-m)!}{(l+m)!}} P_l^m(\cos\theta) \exp(im\rho). \quad (10)$$

In the above, L denotes the associated Laguerre polynomials

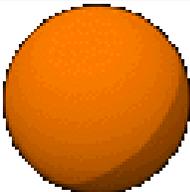
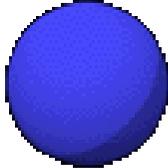
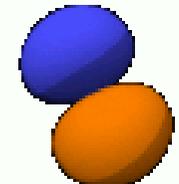
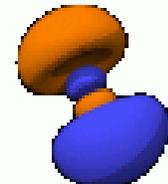
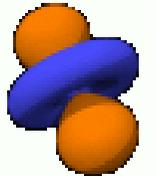
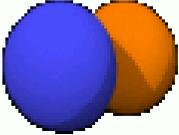
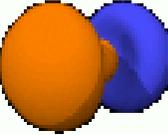
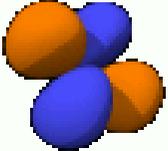
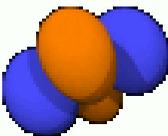
$$L_n^\alpha(x) = \frac{1}{n!} \exp(x) x^{-\alpha} \frac{d^n}{dx^n} [\exp(-x) x^{n+\alpha}] \quad (11)$$

and P the Legendre polynomials

$$P_l^m(x) = \frac{(-1)^m}{2^l l!} (1-x^2)^{m/2} \frac{d^{l+m}}{dx^{l+m}} (x^2-1)^l; \quad P_l^{-m} = (-1)^m \frac{(l-m)!}{(l+m)!} P_l^m(x). \quad (12)$$

These shall be referred to as *atomic orbitals* (AO). However, this is used also for other atomic-centered functions with different radial parts, such as sc. Slater- or Gaussian-type functions.

Some lowest-energy solutions ($Z = 1$)[‡]:

	$n = 1$ $l = 0$	$n = 2$		$n = 3$		
		$l = 0$	$l = 1$	$l = 0$	$l = 1$	$l = 2$
$m = 0$						
$m = 1$						
$m = 2$						

[‡]Blue for positive, orange for negative phase

Molecular orbitals: the first contact

When moving from atoms to molecules, the first complication is that even the *one*-electron eigenfunctions, even when $\hat{g} = 0$, are not obtainable in closed form. Instead of the $1s, 2s, 2p, \dots$ AOs we need polycentric *molecular orbitals* (MO) that describe the states of a single electron spread around the whole nuclear framework.

Example 1.2

The one-electron eigenvalue equation in the case of hydrogen molecule ion H_2^+ is $\hat{h}\phi = \epsilon\phi$, where

$$\hat{h} = -\frac{1}{2}\nabla^2 - \left(\frac{1}{r_A} + \frac{1}{r_B} \right).$$

This may be solved with high accuracy in confocal elliptic coordinates, the solution being a product of three factors that are in turn solutions of three separate differential equations. We will, however, go immediately to approximations. It is reasonable to expect that

$$\phi(\mathbf{r}) \sim c_A \chi_{100}^A(\mathbf{r}) \quad (\text{densities close to nucleus A})$$

$$\phi(\mathbf{r}) \sim c_B \chi_{100}^B(\mathbf{r}) \quad (\text{densities close to nucleus B})$$

where c_A and c_B are some numerical factors. An appropriate trial MO would then be

$$\phi(\mathbf{r}) = c_A \chi_{100}^A(\mathbf{r}) + c_B \chi_{100}^B(\mathbf{r}).$$

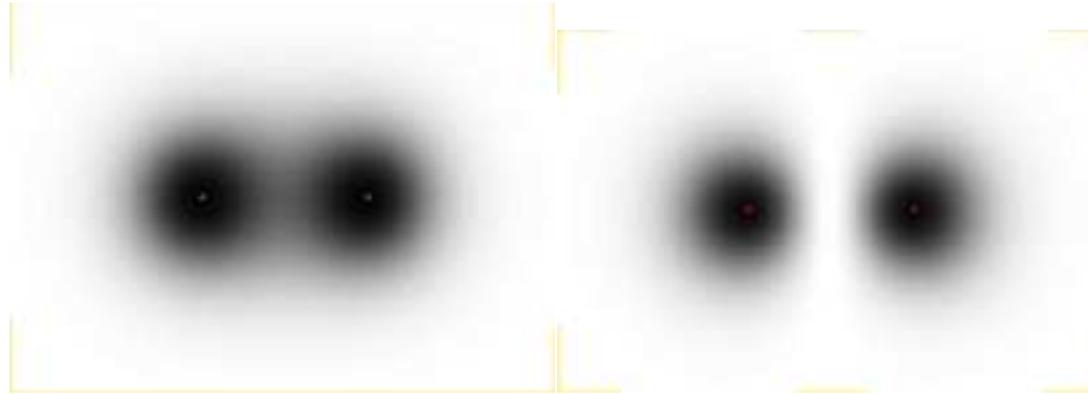
Due to the symmetry of the molecule, it is reasonable to assume that there is no higher density in the one end than in the other; therefore

$$|\phi(\mathbf{r})|^2 = c_A^2 [\chi_{100}^A(\mathbf{r})]^2 + c_B^2 [\chi_{100}^B(\mathbf{r})]^2 + 2c_A c_B \chi_{100}^A(\mathbf{r}) \chi_{100}^B(\mathbf{r}),$$

implies that $c_B = \pm c_A$. Thus, it appears that we can construct two rudimentary MOs from the $1s$ AOs on the two centers, denoted by

$$\begin{aligned}\phi_1(\mathbf{r}) &= c'_A[\chi_{100}^A(\mathbf{r}) + \chi_{100}^B(\mathbf{r})] \\ \phi_2(\mathbf{r}) &= c'_B[\chi_{100}^A(\mathbf{r}) - \chi_{100}^B(\mathbf{r})],\end{aligned}$$

where the remaining numerical factors c'_A and c'_B are to be chosen so that the wave functions will be normalized. The result will be (see Exercise 1.2)



MOs build up this way, in *linear combination of atomic orbitals* (LCAO), are of great importance throughout molecular quantum mechanics.

Spin functions

It is well-known that a single electron is not completely characterized by its spatial wave function $\phi(\mathbf{r})$, but an “intrinsic” angular momentum, spin, is required to explain e.g. line emission spectra of atoms in an external magnetic field. Therefore we must introduce a spin variable m_s in addition to \mathbf{r} . We further associate operators \hat{s}_x , \hat{s}_y and \hat{s}_z with the three components of spin angular momentum. We will adopt \mathbf{x} as a variable taking into account both degrees of freedom, m_s and \mathbf{r} , and then introduce *spin-orbitals*

$$\psi(\mathbf{x}) = \phi(\mathbf{r})\eta(m_s). \quad (13)$$

η describes the spin state and is a solution of an eigenvalue equation $\hat{s}_z\eta = \lambda\eta$. If the orbital state is of energy ϵ , then

$$\hat{h}\psi = \epsilon\psi \quad (14)$$

$$\hat{s}_z\psi = \lambda\psi, \quad (15)$$

and ψ is a state in which an electron simultaneously has definite energy and z-component of spin.

According to observations, the eigenvalue equation (15) has only two solutions for electrons:

$$\hat{s}_z\alpha = \frac{1}{2}\alpha; \quad \hat{s}_z\beta = -\frac{1}{2}\beta$$

often referred to as “spin-up” and “spin-down”. Therefore an orbital ϕ yields two possible spin-orbitals, $\psi = \phi\alpha$ and $\psi' = \phi\beta$. These spin functions may be thought as “spikes” at the points $s_z = +1/2$ and $s_z = -1/2$, respectively, and their normalization and orthogonality properties are

$$\int \alpha^*\alpha dm_s = \int \beta^*\beta dm_s = 1, \quad \int \alpha^*\beta dm_s = \int \beta^*\alpha dm_s = 0. \quad (16)$$

Also the *total spin operators*

$$\hat{S}_z = \sum_{i=1}^N \hat{s}_{z,i} \quad (17)$$

$$\hat{S}^2 = \sum_{i=1}^N \hat{s}_i^2 = \sum_{i=1}^N (\hat{s}_x^2 + \hat{s}_y^2 + \hat{s}_z^2) \quad (18)$$

with eigenvalues of M and $S(S+1)$, respectively, are needed in the discussion of many-electron system. They commute with a spin-free electronic Hamiltonian; and any exact stationary state Ψ is an eigenfunction of them.

Antisymmetry and the Slater determinant

The fact that electrons are indistinguishable, i.e. $|\Psi(\mathbf{x}_1, \mathbf{x}_2)|^2 = |\Psi(\mathbf{x}_2, \mathbf{x}_1)|^2$, leads to (when generalized for N -electrons) two possibilities for a wave function:

$$\begin{aligned}\hat{P}\Psi(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N) &= \Psi(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N) && \text{symmetric wave function} \\ \hat{P}\Psi(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N) &= \varepsilon_P \Psi(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N) && \text{antisymmetric wave function}\end{aligned}$$

where the operator \hat{P} permutes the arguments and ε_P is $+1$ for an even and -1 for an odd number of permutations.

The choice for a system consisting of N electrons is once and for all determined by the *antisymmetry principle* (Pauli exclusion principle): The wave function $\Psi(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N)$ that describes any state of an N -electron system is *antisymmetric* under any permutation of the electrons.

For more than two electrons, a large number of differently permuted product functions must be combined to yield a fully antisymmetric wave function. Expanded in particular permutations κ the exact wave function[§] is written as

$$\Psi(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N) = \sum_{\kappa} c_{\kappa} \Phi_{\kappa}(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N) \quad (19)$$

$$= \sum_{\kappa} c_{\kappa} M_{\kappa} \left[\sum_P \varepsilon_P \hat{P} \psi_1(\mathbf{x}_1) \psi_2(\mathbf{x}_2) \cdots \psi_n(\mathbf{x}_N) \right]. \quad (20)$$

M_{κ} are normalization coefficients and Φ_{κ} is most conveniently expressed as a *Slater determinant*

$$\Phi_{\kappa}(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N) = M \begin{vmatrix} \psi_1(\mathbf{x}_1) & \psi_2(\mathbf{x}_1) & \cdots & \psi_n(\mathbf{x}_1) \\ \psi_1(\mathbf{x}_2) & \psi_2(\mathbf{x}_2) & \cdots & \psi_n(\mathbf{x}_2) \\ \vdots & \vdots & \ddots & \vdots \\ \psi_1(\mathbf{x}_N) & \psi_2(\mathbf{x}_N) & \cdots & \psi_n(\mathbf{x}_N) \end{vmatrix}. \quad (21)$$

[§]Exact at the limit $\kappa \rightarrow \infty$

The symmetric functions are automatically rejected by the Slater determinants, and thus the Pauli principle is fulfilled.

The normalization factor M is reasoned in the following way: when expanded, Φ consists of $N!$ products, and thus an integrand $\Phi^*\Phi$ would consist of $(N!)^2$ products, each of a form

$$\pm[\psi_1(\mathbf{x}_i)\psi_2(\mathbf{x}_j)\cdots\psi_n(\mathbf{x}_k)][\psi_1(\mathbf{x}_{i'})\psi_2(\mathbf{x}_{j'})\cdots\psi_n(\mathbf{x}_{k'})].$$

It is evident that unless $i = i'$, $j = j'$ etc. in a particular product, the product will give no contribution to the result, due to the orthogonality of spin-orbitals. Therefore, there is $N!$ non-vanishing contributions, and as every contribution is equal to 1 with normalized spin-orbitals, we have $\int \Phi^*\Phi d\mathbf{x}_1 \cdots d\mathbf{x}_N = M^2 N!$, hence the normalization factor has the value $M = (N!)^{-1/2}$.

Example 1.3

The Slater determinant of the configuration $(1s)^2$ of a helium atom in the minimal orbital basis is

$$\psi_1 = \frac{1}{\sqrt{2}} \begin{vmatrix} \chi_{100}(\mathbf{r}_1)\alpha(m_{s1}) & \chi_{100}(\mathbf{r}_1)\beta(m_{s1}) \\ \chi_{100}(\mathbf{r}_2)\alpha(m_{s2}) & \chi_{100}(\mathbf{r}_2)\beta(m_{s2}) \end{vmatrix} \equiv \frac{1}{\sqrt{2}} \det |1s\alpha \ 1s\beta|.$$

Note the abbreviated notation that will be used quite extensively hereafter: only the diagonal elements are displayed, and the explicit expression of variables is omitted – the left-right order of the variables in this notation will be 1, 2, 3, Also, note that we have denoted χ_{nlm} 's with their corresponding AO labels.

Remarks on “interpretation”

What is the *physical* meaning of the wave functions introduced earlier? First of all, the general formulation of quantum mechanics is concerned only with symbolic statements

involving relationships between operators and operands, and is independent of the employed “representation” or “interpretation”.

- In a very “orthodox” interpretation, MOs as well as spin-orbitals are just mathematical means for the solution of the molecular Schrödinger equation and no meaning can be given for them or even their squared norms.
- Perhaps more commonly, a spin-orbital $\psi(\mathbf{x})$ itself is thought to be just a mathematical entity but $|\psi(\mathbf{x})|^2 d\mathbf{x}$ is taken to be the probability of a *point-like electron being in an element $d\mathbf{x}$* , or in volume element $d\mathbf{r}$ with spin between m_s and $m_s + dm_s$. $|\phi(\mathbf{r})|^2 d\mathbf{r}$ is the probability of finding an electron with any spin in the volume element $d\mathbf{r}$. Also $\Psi(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N)$ cannot be interpreted; but $|\Psi|^2 d\mathbf{x}_1 d\mathbf{x}_2 \dots d\mathbf{x}_N$ is a probability of the electron 1 in $d\mathbf{x}_1$, electron 2 *simultaneously* in $d\mathbf{x}_2$, etc. Then, the probability of electron 1 being in $d\mathbf{x}_1$ while other electrons are anywhere would be equal to $d\mathbf{x}_1 \int \Psi^* \Psi d\mathbf{x}_2 \dots d\mathbf{x}_N$.
- However, it might be better to think of an electron as being *delocalized*, i.e. not having a more precise position than that indicated by its spatial part of the wave function $\phi(\mathbf{r})$. Then $|\phi(\mathbf{r})|^2 d\mathbf{r}$ would be the amount of charge in an element $d\mathbf{r}$. In other words, in this point of view a MO can be thought to approximately describe an electron spread around the nuclear framework. However, as we will see later, MOs can be unitarily “rotated”, i.e. their form shaped, and they still correspond to the same definite energy. Which MOs are then closest to the physical ones?

Slater method and Slater’s rules

It is possible to find reasonable approximations of the form (20) for the wave function with only a small number of determinants. The optimal coefficients c_κ may, in principle, always be determined by solving a secular problem

$$\mathbf{H}\mathbf{c} = E\mathbf{S}\mathbf{c}, \quad (22)$$

where the matrices \mathbf{H} and \mathbf{S} have the matrix elements

$$H_{\mu\nu} = \int \Psi_{\mu}^*(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N) \hat{H} \Psi_{\nu}(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N) d\mathbf{x}_1 d\mathbf{x}_2 \cdots d\mathbf{x}_N = \langle \Psi_{\mu} | \hat{H} | \Psi_{\nu} \rangle \quad (23)$$

$$S_{\mu\nu} = \int \Psi_{\mu}^*(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N) \Psi_{\nu}(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N) d\mathbf{x}_1 d\mathbf{x}_2 \cdots d\mathbf{x}_N = \langle \Psi_{\mu} | \Psi_{\nu} \rangle. \quad (24)$$

\mathbf{S} is referred to as an *overlap matrix*. The equation (22) is the essence of the *Slater method*. Observe the introduced *Dirac notation* for integrals. The coefficients c are then determined by a condition called the *secular determinant*

$$\begin{vmatrix} H_{11} - E & H_{12} & \cdots & H_{1n} \\ H_{21} & H_{22} - E & \cdots & H_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ H_{n1} & H_{n2} & \cdots & H_{nn} - E \end{vmatrix} = 0. \quad (25)$$

Example 1.4

Consider the 2-electron Hamiltonian, now including electron–electron interaction, $\hat{H} = \hat{h}_1 + \hat{h}_2 + \hat{g}$. Let us assume that the wave function Ψ_1 in Example 1.2 approximates the $(1s)^2$ configuration of the He atom. The corresponding energy is obtained as an expectation value $\langle \Psi_1 | \hat{H} | \Psi_1 \rangle$, or by writing the normalization explicitly, as

$$E_1 = \frac{\langle \Psi_1 | \hat{H} | \Psi_1 \rangle}{\langle \Psi_1 | \Psi_1 \rangle}. \quad (26)$$

Now

$$H_{11} = \frac{1}{2} \int \left\{ \chi_{100}^*(\mathbf{r}_1) \chi_{100}(\mathbf{r}_2) [\alpha(m_{s1})\beta(m_{s2}) - \alpha(m_{s2})\beta(m_{s1})]^* \right. \\ \left. \times \hat{H} \chi_{100}(\mathbf{r}_1) \chi_{100}(\mathbf{r}_2) [\alpha(m_{s1})\beta(m_{s2}) - \alpha(m_{s2})\beta(m_{s1})] \right\} d\mathbf{r}_1 d\mathbf{r}_2 dm_{s1} dm_{s2},$$

but since the Hamiltonian does not operate on spins, we may perform the spin-integration immediately [c.f. Eq. (16)],

$$\int [\alpha(m_{s1})\beta(m_{s2}) - \alpha(m_{s2})\beta(m_{s1})]^* [\alpha(m_{s1})\beta(m_{s2}) - \alpha(m_{s2})\beta(m_{s1})] dm_{s1}dm_{s2} = 2.$$

Thus we have

$$\begin{aligned} H_{11} &= \int \chi_{100}^*(\mathbf{r}_1)\chi_{100}^*(\mathbf{r}_2)\hat{H}\chi_{100}(\mathbf{r}_1)\chi_{100}(\mathbf{r}_2)d\mathbf{r}_1d\mathbf{r}_2 \\ &= \int \chi_{100}^*(\mathbf{r}_1)\hat{h}_1\chi_{100}(\mathbf{r}_1)d\mathbf{r}_1 \int \chi_{100}^*(\mathbf{r}_2)\chi_{100}(\mathbf{r}_2)d\mathbf{r}_2 \\ &\quad + \int \chi_{100}^*(\mathbf{r}_2)\hat{h}_2\chi_{100}(\mathbf{r}_2)d\mathbf{r}_2 \int \chi_{100}^*(\mathbf{r}_1)\chi_{100}(\mathbf{r}_1)d\mathbf{r}_1 \\ &\quad + \int \chi_{100}^*(\mathbf{r}_1)\chi_{100}^*(\mathbf{r}_2)\hat{g}\chi_{100}(\mathbf{r}_1)\chi_{100}(\mathbf{r}_2)d\mathbf{r}_1d\mathbf{r}_2 \\ &= 2\epsilon_{1s} + J \end{aligned}$$

This result is still a crude approximation, the only difference to the independent-particle model being the contribution J from the two-electron integral. The next step in the determination of He energy levels would be to include more and more Slater determinants corresponding to excited state configurations and evaluate the respective matrix elements. As anticipated earlier, although the exact solution would require an infinite number of included determinants, the energy spectrum would in practise converge towards the correct one rather rapidly. However, the analysis of the many-electron problem using this kind of direct expansion would obviously be extremely tedious, and a more general approach is needed. Also, the evaluation of integrals gets far more complicated in the case of many atomic cores.

We may indeed devise more systematic rules for the matrix elements. We begin by noting that the expectation value over the 1-electron part of the molecular Hamiltonian, $\sum_{i=1}^N \hat{h}_i$,

will reduce to the sum of N identical terms, since the coordinates of each electron appear symmetrically in the corresponding integral:

$$\left\langle \Phi \left| \sum_i \hat{h}_i \right| \Phi \right\rangle = \sum_r \langle \psi_r | \hat{h} | \psi_r \rangle. \quad (27)$$

The expectation value of the two-electron part is a bit more complicated, as contributions may arise from terms that differ by an interchange of two electrons; but essentially a similar argumentation leads to

$$\left\langle \Phi \left| \sum_{ij} \hat{g}_{ij} \right| \Phi \right\rangle = \sum_{rs} ' (\langle \psi_r \psi_s | \hat{g} | \psi_r \psi_s \rangle - \langle \psi_r \psi_s | \hat{g} | \psi_s \psi_r \rangle), \quad (28)$$

where the former part is referred to as the *Coulomb integral* and the latter as the *exchange integral*. the expectation value of the energy being simply the sum of (27) and (28).

We also need the off-diagonal elements. Fortunately enough, there are non-vanishing elements only in two cases:

- one spin-orbital is different between Φ and Φ' ($\psi'_r \neq \psi_r$) when

$$\langle \Phi' | \hat{H} | \Phi \rangle = \langle \psi'_r | \hat{h} | \psi_s \rangle + \sum_{s \neq r} (\langle \psi'_r \psi_s | \hat{g} | \psi_r \psi_s \rangle - \langle \psi'_r \psi_s | \hat{g} | \psi_s \psi_r \rangle), \quad (29)$$

- two spin-orbitals are different ($\psi'_r \neq \psi_r, \psi'_s \neq \psi_s$) when

$$\langle \Phi' | \hat{H} | \Phi \rangle = (\langle \psi'_r \psi'_s | \hat{g} | \psi_r \psi_s \rangle - \langle \psi'_r \psi'_s | \hat{g} | \psi_s \psi_r \rangle), \quad (30)$$

These results are called *Slater's rules*. In Slater's rules, the spin-orbitals are required to be orthonormal, but the rules may be generalized for matrix elements even in a non-orthogonal spin-orbital basis.

1.2 Electron distribution

Electron density functions

Regardless of the interpretation, it is useful to define the function

$$\rho(\mathbf{x}_1) = N \int \Psi(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N) \Psi^*(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N) d\mathbf{x}_2 \cdots d\mathbf{x}_N, \quad (31)$$

and an *electron density function* (or probability density functions) that are observable by e.g. X-ray crystallography, by integrating ρ over spin,

$$P(\mathbf{r}_1) = \int \rho(\mathbf{x}_1) dm_{s_1}. \quad (32)$$

It is possible to introduce corresponding density functions for different configurations of any number of particles:

$$\pi(\mathbf{x}_1, \mathbf{x}_2) = N(N-1) \int \Psi(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N) \Psi^*(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N) d\mathbf{x}_3 \cdots d\mathbf{x}_N \quad (33)$$

$$\Pi(\mathbf{r}_1, \mathbf{r}_2) = \int \pi(\mathbf{x}_1, \mathbf{x}_2) dm_{s_1} dm_{s_2} \quad (34)$$

determine the two-electron densities and tell us how the motions of two electrons are “correlated” as a result of their interaction. Because electrons interact only in pairs (i.e. there are no many-body effects), there is no need to consider any higher distribution functions than the pair functions.

Example 1.5

Let us consider the $1s2s$ states of the helium atom,

$$\begin{aligned}
 {}^1\Psi(\mathbf{x}_1, \mathbf{x}_2) &= \frac{1}{2} [\chi_{100}(\mathbf{r}_1)\chi_{200}(\mathbf{r}_2) + \chi_{200}(\mathbf{r}_1)\chi_{100}(\mathbf{r}_2)] [\alpha(m_{s1})\beta(m_{s2}) - \beta(m_{s1})\alpha(m_{s2})] \\
 {}^3\Psi(\mathbf{x}_1, \mathbf{x}_2) &= \frac{1}{\sqrt{2}} [\chi_{100}(\mathbf{r}_1)\chi_{200}(\mathbf{r}_2) - \chi_{200}(\mathbf{r}_1)\chi_{100}(\mathbf{r}_2)] \\
 &\quad \times \begin{cases} \alpha(m_{s1})\alpha(m_{s2}) & M = +1 \\ \frac{1}{\sqrt{2}} [\alpha(m_{s1})\beta(m_{s2}) + \beta(m_{s1})\alpha(m_{s2})] & M = 0 \\ \beta(m_{s1})\beta(m_{s2}) & M = -1 \end{cases}
 \end{aligned}$$

For the singlet state, we have

$$\begin{aligned}
 \rho(\mathbf{x}_1) &= 2 \times \frac{1}{4} \int |\chi_{100}(\mathbf{r}_1)\chi_{200}(\mathbf{r}_2) + \chi_{200}(\mathbf{r}_1)\chi_{100}(\mathbf{r}_2)|^2 |\alpha(m_{s1})\beta(m_{s2}) - \beta(m_{s1})\alpha(m_{s2})|^2 d\mathbf{r}_2 dm_{s2} \\
 &= \frac{1}{2} \left[|\chi_{100}(\mathbf{r}_1)|^2 + |\chi_{200}(\mathbf{r}_1)|^2 \right] \left[|\alpha(m_{s1})|^2 + |\beta(m_{s1})|^2 \right]
 \end{aligned}$$

$$P(\mathbf{r}_1) = \int \rho(\mathbf{x}_1) dm_{s1} = |\chi_{100}(\mathbf{r}_1)|^2 + |\chi_{200}(\mathbf{r}_1)|^2.$$

ρ can be rewritten as

$$\rho(\mathbf{x}_1) = P_\alpha(\mathbf{r}_1) |\alpha(m_{s1})|^2 + P_\beta(\mathbf{r}_1) |\beta(m_{s1})|^2,$$

where the first term vanishes unless $m_{s1} = +1/2$ and the second unless $m_{s1} = -1/2$, and thus

$$\begin{aligned}
 P_\alpha(\mathbf{r}_1) &= \frac{1}{2} P(\mathbf{r}_1) \text{ density for spin-up electrons} \\
 P_\beta(\mathbf{r}_1) &= \frac{1}{2} P(\mathbf{r}_1) \text{ density for spin-down electrons}
 \end{aligned}$$

These densities are always equal in a singlet state, but in non-singlet states the α - and β -densities differ, giving a resultant *spin density*. For the first triplet state we obtain

$$P_\alpha(\mathbf{r}_1) = P(\mathbf{r}_1), \quad P_\beta(\mathbf{r}_1) = 0$$

and for the third

$$P_\alpha(\mathbf{r}_1) = 0, \quad P_\beta(\mathbf{r}_1) = P(\mathbf{r}_1),$$

while for the second we get

$$P_\alpha(\mathbf{r}_1) = P_\beta(\mathbf{r}_1) = \frac{1}{2}P(\mathbf{r}_1).$$

Density matrices

Consider a 1-electron system, an electron described by a spin-orbital ψ . The expectation value in this state of any quantity described by operator \hat{F} is given by $\langle \hat{F} \rangle = \int \psi^*(\mathbf{x}) \hat{F} \psi(\mathbf{x}) d\mathbf{x}$. We can rewrite this as

$$\langle \hat{F} \rangle = \int_{\mathbf{x}'=\mathbf{x}} \hat{F} \psi(\mathbf{x}) \psi^*(\mathbf{x}') d\mathbf{x} = \int_{\mathbf{x}'=\mathbf{x}} \hat{F} \rho(\mathbf{x}; \mathbf{x}') d\mathbf{x}$$

by having agreed that \hat{F} works on functions of \mathbf{x} only and setting $\mathbf{x}' = \mathbf{x}$ after operating with \hat{F} but before completing the integration. We define the *generalized density functions* as

$$\rho(\mathbf{x}_1; \mathbf{x}'_1) = N \int \Psi(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N) \Psi^*(\mathbf{x}'_1, \mathbf{x}_2, \dots, \mathbf{x}_N) d\mathbf{x}_2 \cdots d\mathbf{x}_N \quad (35)$$

$$\pi(\mathbf{x}_1, \mathbf{x}_2; \mathbf{x}'_1, \mathbf{x}'_2) = N(N-1) \int \Psi(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N) \Psi^*(\mathbf{x}'_1, \mathbf{x}'_2, \dots, \mathbf{x}_N) d\mathbf{x}_3 \cdots d\mathbf{x}_N. \quad (36)$$

We may discuss *all* one-electron properties in terms of $\rho(\mathbf{x}_1; \mathbf{x}'_1)$ and *all* two-electron properties in terms of $\pi(\mathbf{x}_1, \mathbf{x}_2; \mathbf{x}'_1, \mathbf{x}'_2)$. The spinless analogues are given by

$$P(\mathbf{r}_1; \mathbf{r}'_1) = \int_{m_{s_1}'=m_{s_1}} \rho(\mathbf{x}_1; \mathbf{x}'_1) dm_{s_1} \quad (37)$$

$$\Pi(\mathbf{r}_1, \mathbf{r}_2; \mathbf{r}'_1, \mathbf{r}'_2) = \int_{m_{s_1}'=m_{s_1}} \int_{m_{s_2}'=m_{s_2}} \pi(\mathbf{x}_1, \mathbf{x}_2; \mathbf{x}'_1, \mathbf{x}'_2) dm_{s_1} dm_{s_2}. \quad (38)$$

The individual expectation values for terms in the electronic Hamiltonian obtain the following expressions using the general density matrices:

$$\left\langle \sum_i \hat{h}_i \right\rangle = \int_{\mathbf{x}'_1=\mathbf{x}_1} \hat{h}_1 \rho(\mathbf{x}_1; \mathbf{x}'_1) d\mathbf{x}_1 \quad (39)$$

$$\left\langle \sum_i \hat{V}_i \right\rangle = \int_{\mathbf{x}'_1=\mathbf{x}_1} \hat{V}_1 \rho(\mathbf{x}_1; \mathbf{x}'_1) d\mathbf{x}_1 = \int \hat{V}_1 \rho(\mathbf{x}_1) d\mathbf{x}_1 \quad (40)$$

$$\left\langle \sum_{ij} \hat{g}_{ij} \right\rangle = \int_{\mathbf{x}'_1=\mathbf{x}_1} \int_{\mathbf{x}'_2=\mathbf{x}_2} \hat{g}_{12} \pi(\mathbf{x}_1, \mathbf{x}_2; \mathbf{x}'_1, \mathbf{x}'_2) d\mathbf{x}_1 d\mathbf{x}_2 = \int \hat{g}_{12} \pi(\mathbf{x}_1, \mathbf{x}_2) d\mathbf{x}_1 d\mathbf{x}_2. \quad (41)$$

The latter expressions for $\langle \sum_i \hat{V}_i \rangle$ and $\langle \sum_{ij} \hat{g}_{ij} \rangle$ are obtained by noting that $\hat{V}_1 = V(\mathbf{r}_1)$ is just a factor in the integrand, as is \hat{g}_{12} , and thus the primes may be omitted. Therefore, the general expression for the N -electron system with the usual molecular Hamiltonian Eq. (2) takes a simple form

$$E = -\frac{1}{2} \int \nabla^2 \rho(\mathbf{x}_1) d\mathbf{x}_1 + \sum_{I=1}^K \int |\mathbf{r}_1 - \mathbf{R}_I|^{-1} \rho(\mathbf{x}_1) d\mathbf{x}_1 + \frac{1}{2} \int |\mathbf{r}_1 - \mathbf{r}_2|^{-1} \pi(\mathbf{x}_1, \mathbf{x}_2) d\mathbf{x}_1 d\mathbf{x}_2, \quad (42)$$

or, by carrying out the spin-integration,

$$\begin{aligned}
 E &= -\frac{1}{2} \int \nabla^2 P(\mathbf{r}_1) d\mathbf{r}_1 + \sum_{I=1}^K \int |\mathbf{r}_1 - \mathbf{R}_I|^{-1} P(\mathbf{r}_1) d\mathbf{r}_1 + \frac{1}{2} \int |\mathbf{r}_1 - \mathbf{r}_2|^{-1} \Pi(\mathbf{r}_1, \mathbf{r}_2) d\mathbf{r}_1 d\mathbf{r}_2 \\
 &= T + V_{en} + V_{ee}.
 \end{aligned} \tag{43}$$

It should be emphasized that these results are valid for all kinds of states and their approximate wave functions of any system. Furthermore, for evaluation of molecular energies and properties, we do not even need to construct the N -electron wave function, but everything is exactly described by the electron densities.

Let us carry out the same for a single Slater determinant Ψ , that we recall to provide an useful approximation for the ground state wave function, by considering N spin-orbitals having the orbital factor corresponding to an occupied MO. The energy expression is given by

$$E = \langle \Psi | \hat{H} | \Psi \rangle = \sum_{r \in \text{OCC}} \langle \psi_r | \hat{h} | \psi_r \rangle + \frac{1}{2} \sum_{rs \in \text{OCC}} (\langle \psi_r \psi_s | \hat{g} | \psi_r \psi_s \rangle - \langle \psi_r \psi_s | \hat{g} | \psi_s \psi_r \rangle). \tag{44}$$

Assume that we may expand $\rho(\mathbf{x}_1, \mathbf{x}'_1)$ in spin-orbitals as

$$\rho(\mathbf{x}_1, \mathbf{x}'_1) = \sum_{rs} \rho_{rs} \psi_r(\mathbf{x}_1) \psi_s^*(\mathbf{x}'_1), \tag{45}$$

where the ρ_{rs} are numerical coefficients; then the expectation value (39) can be written as the trace of a matrix product:

$$\left\langle \sum_i \hat{h}_i \right\rangle = \sum_{rs} \rho_{rs} \langle \psi_s | \hat{h} | \psi_r \rangle = \text{Tr}[\rho \mathbf{h}], \tag{46}$$

and by comparing to the general expression, we observe that in the one-determinant approximation $\rho_{RS} = \delta_{RS}$. The one-determinant approximation to $\rho(\mathbf{x}_1; \mathbf{x}'_1)$ may now be written as

$$\rho(\mathbf{x}_1; \mathbf{x}'_1) = \sum_{r \in \text{OCC}} \psi_r(\mathbf{x}_1) \psi_r^*(\mathbf{x}'_1). \quad (47)$$

The expression for π follows in a similar manner, and we find

$$\pi(\mathbf{x}_1, \mathbf{x}_2; \mathbf{x}'_1, \mathbf{x}'_2) = \sum_{rs} [\psi_r(\mathbf{x}_1) \psi_s(\mathbf{x}_2) \psi_r^*(\mathbf{x}'_1) \psi_s^*(\mathbf{x}'_2) - \psi_r(\mathbf{x}_2) \psi_s(\mathbf{x}_1) \psi_r^*(\mathbf{x}'_2) \psi_s^*(\mathbf{x}'_1)] \quad (48)$$

$$= \rho(\mathbf{x}_1; \mathbf{x}'_1) \rho(\mathbf{x}_2; \mathbf{x}'_2) - \rho(\mathbf{x}_2; \mathbf{x}'_1) \rho(\mathbf{x}_1; \mathbf{x}'_2). \quad (49)$$

The factorization of the two-electron density matrix in terms of the one-electron ρ is characteristic to the one-determinant approximation: it means that in this approximation everything is determined by the function ρ (often called as the *Fock–Dirac density matrix*).

So far, we have included spin implicitly through the use of spin-orbitals. In the one-determinant case we may write spin explicitly, for the case $\mathbf{x}'_1 = \mathbf{x}_1$,

$$\rho(\mathbf{x}_1) = P_\alpha(\mathbf{r}_1) \alpha(m_{s1}) \alpha^*(m_{s1}) + P_\beta(\mathbf{r}_1) \beta(m_{s1}) \beta^*(m_{s1}), \quad (50)$$

where the densities for spin-up and spin-down electrons are

$$P_\alpha(\mathbf{r}_1) = \sum_{s(\alpha)} \phi_s(\mathbf{r}_1) \phi_s^*(\mathbf{r}_1), \quad P_\beta(\mathbf{r}_1) = \sum_{s(\beta)} \phi_s(\mathbf{r}_1) \phi_s^*(\mathbf{r}_1). \quad (51)$$

By integrating ρ over spin, we obtain the electron density in the form

$$P(\mathbf{r}_1) = P_\alpha(\mathbf{r}_1) + P_\beta(\mathbf{r}_1) \quad (52)$$

that is, a sum of spin-up and spin-down densities. The excess of spin-up density over spin-down density,

$$Q_z(\mathbf{r}_1) = \frac{1}{2} [P_\alpha(\mathbf{r}_1) - P_\beta(\mathbf{r}_1)], \quad (53)$$

is an useful quantity, which is in fact the density of spin angular momentum around the z -axis. $P(\mathbf{r})$ and $Q_z(\mathbf{r})$ are enough to determine all the properties of the electron distribution in the one-determinant approximation.

Let us finally consider an important special case, a *closed-shell* system, in which all spin-orbitals are doubly occupied, once with α and once with β -electron; meaning that $P_\alpha = P_\beta = P/2$ and the spin density is everywhere zero. In this case,

$$P(\mathbf{r}_1) = 2 \sum_{s \in \text{OCC}} \phi_s(\mathbf{r}_1) \phi_s^*(\mathbf{r}_1), \quad (54)$$

and by substituting Eq. (50) to Eq. (49) we find

$$\Pi(\mathbf{r}_1, \mathbf{r}_2) = P(\mathbf{r}_1)P(\mathbf{r}_2) - \frac{1}{2}P(\mathbf{r}_2; \mathbf{r}_1)P(\mathbf{r}_1; \mathbf{r}_2). \quad (55)$$

Substituting these expressions to Eq. (43) we obtain an often-used expression for the closed-shell energy:

$$E = 2 \sum_r \langle \phi_r | \hat{h} | \phi_r \rangle + \sum_{rs} [2 \langle \phi_r \phi_s | \hat{g} | \phi_r \phi_s \rangle - \langle \phi_r \phi_s | \hat{g} | \phi_s \phi_r \rangle]. \quad (56)$$

The pair function and electron correlation

Let us write the spin-dependence explicitly also in the case of π :

$$\begin{aligned} \pi(\mathbf{x}_1, \mathbf{x}_2; \mathbf{x}'_1, \mathbf{x}'_2) &= \Pi_{\alpha\alpha, \alpha\alpha}(\mathbf{r}_1, \mathbf{r}_2; \mathbf{r}'_1, \mathbf{r}'_2) \alpha(m_{s_1}) \alpha(m_{s_2}) \alpha^*(m_{s'_1}) \alpha^*(m_{s'_2}) \\ &\quad + \Pi_{\alpha\beta, \alpha\beta}(\mathbf{r}_1, \mathbf{r}_2; \mathbf{r}'_1, \mathbf{r}'_2) \alpha(m_{s_1}) \beta(m_{s_2}) \alpha^*(m_{s'_1}) \beta^*(m_{s'_2}) + \dots \end{aligned} \quad (57)$$

which reduces to a four-component quantity when spin is integrated over. Eq. (57) can be written in our notation for the diagonal components, $\Pi_{\alpha\alpha, \alpha\alpha}(\mathbf{r}_1, \mathbf{r}_2; \mathbf{r}_1, \mathbf{r}_2) = \Pi_{\alpha\alpha}(\mathbf{r}_1, \mathbf{r}_2)$ etc., as

$$\Pi(\mathbf{r}_1, \mathbf{r}_2) = \Pi_{\alpha\alpha}(\mathbf{r}_1, \mathbf{r}_2) + \Pi_{\alpha\beta}(\mathbf{r}_1, \mathbf{r}_2) + \Pi_{\beta\alpha}(\mathbf{r}_1, \mathbf{r}_2) + \Pi_{\beta\beta}(\mathbf{r}_1, \mathbf{r}_2), \quad (58)$$

where each component resembles a contribution to the pair function Π from electrons at \mathbf{r}_1 and \mathbf{r}_2 with spins indicated by the subscript as well as the probability of each spin-configuration. The two other six non-zero terms in (57) vanish in the spin-integration.

The pair-function components in the simple case of a system described by one determinant of spin-orbitals [c.f. Eqs. (50) and (49)] are (with the similar terms obtained by $\alpha \leftrightarrow \beta$)

$$\Pi_{\alpha\alpha}(\mathbf{r}_1, \mathbf{r}_2) = P_\alpha(\mathbf{r}_1)P_\alpha(\mathbf{r}_2) - P_\alpha(\mathbf{r}_2; \mathbf{r}_1)P_\alpha(\mathbf{r}_1; \mathbf{r}_2), \quad (59)$$

$$\Pi_{\alpha\beta}(\mathbf{r}_1, \mathbf{r}_2) = P_\alpha(\mathbf{r}_1)P_\beta(\mathbf{r}_2). \quad (60)$$

From these results we can begin to construct a picture on *electron correlation*, keeping the one-determinant approximation in mind: From the latter equation we see that there is no correlation between the positions of electrons of opposite spins, whilst electrons of like spin are correlated as their pair function vanishes for $\mathbf{r}_2 \rightarrow \mathbf{r}_1$. This special type of correlation, referred to as *Fermi correlation*, prevents electrons of like spin to coincidence in space.

If we leave the one-determinant approximation and suppose that Ψ is an exact many-electron wave function, the following ramifications for electron correlation hold.

1. $\Pi_{\alpha\alpha}(\mathbf{r}_1, \mathbf{r}_2)$ and $\Pi_{\beta\beta}(\mathbf{r}_1, \mathbf{r}_2)$ both vanish like r_{12}^2 for $\mathbf{r}_2 \rightarrow \mathbf{r}_1$, giving zero probability of finding two electrons of like spin at the same point in space. This follows from the antisymmetry of Ψ and the resulting antisymmetry of $\Pi_{\alpha\alpha, \alpha\alpha}(\mathbf{r}_1, \mathbf{r}_2; \mathbf{r}'_1, \mathbf{r}'_2)$ and $\Pi_{\alpha\alpha, \alpha\alpha}(\mathbf{r}_1, \mathbf{r}_2; \mathbf{r}'_1, \mathbf{r}'_2)$ in both pairs of variables; and may be regarded as the most general statement of the Pauli exclusion principle.
2. By writing $\Pi_{\alpha\alpha}(\mathbf{r}_1, \mathbf{r}_2) = P_\alpha(\mathbf{r}_1)P_\alpha(\mathbf{r}_2)[1 + f^{\alpha\alpha}(\mathbf{r}_1, \mathbf{r}_2)]$ we may define the *correlation hole* or the *Coulomb hole*

$$\frac{\Pi_{\alpha\alpha}}{P_\alpha(\mathbf{r}_2)} - P_\alpha(\mathbf{r}_1) = P_\alpha(\mathbf{r}_1)f^{\alpha\alpha}(\mathbf{r}_1, \mathbf{r}_2),$$

i.e., the difference between the probability of finding a spin-up electron in \mathbf{r}_1 when one spin-up electron is at \mathbf{r}_2 and the same probability in the absence of the second electron.

The form of the correlation factor f is known for dilute electron gas, but unknown in the case of complicated potential present in a molecular system. Clearly, the Fermi correlation requires that $f^{\alpha\alpha}(\mathbf{r}_1, \mathbf{r}_2) \rightarrow -1$ for $\mathbf{r}_2 \rightarrow \mathbf{r}_1$. This “hole” integrates to -1 :

$$\int P_{\alpha}(\mathbf{r}_1) f^{\alpha\alpha}(\mathbf{r}_1, \mathbf{r}_2) d\mathbf{r}_1 = -1 \quad \forall \mathbf{r}_2.$$

3. Similarly, if we write

$$\Pi_{\alpha\beta}(\mathbf{r}_1, \mathbf{r}_2) = P_{\alpha}(\mathbf{r}_1) P_{\beta}(\mathbf{r}_2) [1 + f^{\alpha\beta}(\mathbf{r}_1, \mathbf{r}_2)]$$

then the correlation hole integrates to zero:

$$\int P_{\alpha}(\mathbf{r}_1) f^{\alpha\beta}(\mathbf{r}_1, \mathbf{r}_2) d\mathbf{r}_1 = 0 \quad \forall \mathbf{r}_2.$$

The probability should be decreased (i.e. $f < 0$) near the reference electron due to the Coulomb repulsion but enhanced further away.

4. The hole function is also known to have a discontinuity of slope, or “cusp” for $\mathbf{r}_1 \rightarrow \mathbf{r}_2$.

Let us at this stage review some nomenclature of electron correlation used in molecular electronic-structure theory.

- A commonly-used concept, *correlation energy*, has a pragmatic definition

$$E_{\text{corr}} = E_{\text{exact}} - E_{\text{HF}},$$

where E_{HF} is a best one-determinant total energy and E_{exact} the “exact” energy from a high-precision calculation. This definition is most usable when speaking of molecular ground states and equilibrium geometries, while outside them it is untenable.

- *Fermi correlation* arises from the Pauli antisymmetry of the wave function and is taken into account already at the single-determinant level.
- *Static correlation*, also known as near-degeneracy or nondynamical correlation, arises from the near-degeneracy of electronic configurations.
- *Dynamical correlation* is associated with the instantaneous correlation among the electrons arising from their mutual Coulombic repulsion. It is useful to distinguish between
 - *long-range dynamical correlation* and
 - *short-range dynamical correlation*, which is related to the singularities in the Hamiltonian and giving rise to the Coulomb cusp in the wave function.

Other density functions

Other property densities, in addition to the charge density $P(\mathbf{r})$ and the spin-density $Q_z(\mathbf{r})$, could be defined in the same way. For example, referring to Eq. (43),

$$T(\mathbf{r}) = \left[\frac{\tilde{p}^2}{2} P(\mathbf{r}; \mathbf{r}') \right]_{\mathbf{r}'=\mathbf{r}} \quad (\tilde{p}^2 = \tilde{p}_x^2 + \tilde{p}_y^2 + \tilde{p}_z^2)$$

would be a *kinetic-energy density*, since $T(\mathbf{r})d\mathbf{r}$ is the contribution to $\langle T \rangle$ associated with volume element $d\mathbf{r}$ at \mathbf{r} . However, point contributions to real observables are not necessarily real, and the proper definition of the kinetic-energy density is

$$T(\mathbf{r}) = \left[\frac{\hat{\mathbf{p}} \cdot \hat{\mathbf{p}}^\dagger}{2} P(\mathbf{r}; \mathbf{r}') \right]_{\mathbf{r}'=\mathbf{r}}. \quad (61)$$

where the adjoint operator is understood to work on the primed variable only.

Another often encountered density function is the *current density*

$$J_{\mu}(\mathbf{r}) = \frac{1}{2} [(\hat{p}_{\mu} + \hat{p}_{\mu}^{\dagger})P(\mathbf{r}; \mathbf{r}')]_{\mathbf{r}'=\mathbf{r}} = \text{Re} [\hat{p}_{\mu}P(\mathbf{r}; \mathbf{r}')]_{\mathbf{r}'=\mathbf{r}}, \quad (62)$$

where μ denotes the Cartesian x, y, z components. For a system in a stationary state described by a real wave function $\mathbf{J} = \mathbf{0}$, but for a system in the presence of a magnetic field, circulating currents are always present. A more general form of \mathbf{J} is obtained on introducing field terms in the Hamiltonian via the gauge-invariant momentum operator,

$$\hat{\mathbf{p}} \rightarrow \hat{\boldsymbol{\pi}} = \hat{\mathbf{p}} + \mathbf{A}, \quad (63)$$

where \mathbf{A} is the magnetic vector potential from the applied field is derived. The corresponding current density is

$$J_{\mu}(\mathbf{r}) = \text{Re} [\hat{\pi}_{\mu}P(\mathbf{r})]. \quad (64)$$

The kinetic energy density is also modified by the presence of the field, and (61) becomes

$$T(\mathbf{r}) = \left[\frac{\hat{\boldsymbol{\pi}} \cdot \hat{\boldsymbol{\pi}}^{\dagger}}{2} P(\mathbf{r}; \mathbf{r}') \right]_{\mathbf{r}'=\mathbf{r}}.$$

Finally; it may be shown that for a wave function that satisfies the time-dependent Schrödinger equation the charge density and the current density satisfy the conservation equation

$$\nabla \cdot \mathbf{J}(\mathbf{r}) = -\frac{\partial P(\mathbf{r})}{\partial t}. \quad (65)$$

When Ψ is an exact stationary-state function, P is time-independent and the net flux of density out of any region is zero; but with $\mathbf{A} \neq \mathbf{0}$ there will be in general a steady-state distribution of non-zero *currents*, satisfying the usual continuity equation $\nabla \cdot \mathbf{J} = 0$.

1.3 Exact and approximate wave functions

Characteristics of the exact wave function

We shall need to approximate the molecular wave function – perhaps drastically – in order to reach the applicability to molecules with multiple atomic cores and several electrons. The approximations should, however, be carried out with care. For this reason, we will now list down the properties (either rationalized earlier or simply taken for granted) that an exact wave function would possess, and on this basis we should seek an approximate solution that retains as many of the following properties as possible.

The exact molecular electronic wave function Ψ

1. is *antisymmetric* with respect to the permutation of any pair of the electrons:

$$\hat{P}\Psi(\mathbf{x}_1, \dots, \mathbf{x}_N) = \epsilon_P \Psi(\mathbf{x}_1, \dots, \mathbf{x}_N) \quad (66)$$

where $\epsilon_P = \pm 1$ for even and odd number of permutations, respectively.

2. is *square-integrable* everywhere in space,

$$\langle \Psi | \Psi \rangle = K \quad (K \text{ finite}) \quad (67)$$

3. is *variational* in the sense that for all possible variations $\delta\Psi$, which are orthogonal to the wave function, the energy remains unchanged:

$$\langle \delta\Psi | \Psi \rangle = 0 \Rightarrow \langle \delta\Psi | H | \Psi \rangle = 0 \quad (68)$$

4. is *size-extensive*; for a system containing *non-interacting* subsystems the total energy is equal to the sum of the energies of the individual systems.

5. Within the non-relativistic theory, the exact stationary states are *eigenfunctions of the total and projected spin operators*,

$$\hat{S}^2\Psi = S(S+1)\Psi \quad (69)$$

$$\hat{S}_z\Psi = M\Psi \quad (70)$$

6. The molecular electronic Hamiltonian (2) is singular for $\mathbf{r}_i = \mathbf{r}_j$ and thus the exact wave function must possess a characteristic non-differentiable behavior for spatially coinciding electrons, known as the *electronic Coulomb cusp condition*,

$$\lim_{r_{ij} \rightarrow 0} \left(\frac{\partial \Psi}{\partial r_{ij}} \right)_{\text{ave}} = \frac{1}{2} \Psi(r_{ij} = 0), \quad (71)$$

the description of which is a major obstacle in the accurate practical modelling of the electronic wave function. There exists a condition similar to (71) also for electrons coinciding *point-like* nuclei, known as the *nuclear cusp condition*.

7. It may shown that at large distances the *electron density* decays as

$$\rho(r) \sim \exp(-2^{3/2}\sqrt{I}r), \quad (72)$$

where I is the ionization potential of the molecule.

8. The exact wave function transforms in a characteristic manner under gauge transformations of the potentials associated with electromagnetic fields, ensuring that all molecular properties described by the wave function are unaffected by the transformations.

As we saw in the introductory survey, some of these – square integrability and Pauli principle – are included straightforwardly, whereas size-extensivity and the cusp condition are more difficult to impose but still desirable. Others are of interest only in special situations.

The Coulomb cusp

The non-relativistic Hamiltonian of the helium-like atom with the origin at the nucleus,

$$\hat{H} = -\frac{1}{2}\nabla_1^2 - \frac{1}{2}\nabla_2^2 - \frac{Z}{|\mathbf{r}_1|} - \frac{Z}{|\mathbf{r}_2|} + \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|}$$

has singularities for $\mathbf{r}_1 = \mathbf{0}$, $\mathbf{r}_2 = \mathbf{0}$, and for $\mathbf{r}_1 = \mathbf{r}_2$. At these points, the exact solution to the Schrödinger equation must provide contributions that balance the singularities such that the *local energy*

$$\epsilon(\mathbf{r}_1, \mathbf{r}_2) = \frac{\hat{H}\Psi(\mathbf{r}_1, \mathbf{r}_2)}{\Psi(\mathbf{r}_1, \mathbf{r}_2)} \quad (73)$$

remains constant and equal to the eigenvalue E . The only possible “source” for this balancing is the kinetic energy. It is convenient to employ the symmetry of the helium atom and express the kinetic energy operator in terms of three radial coordinates r_1 , r_2 and r_{12} , such that the Hamiltonian is written as

$$\hat{H} = -\frac{1}{2}\sum_{i=1}^2 \left(\frac{\partial^2}{\partial r_i^2} + \frac{2}{r_i} \frac{\partial}{\partial r_i} + \frac{2Z}{r_i} \right) - \left(\frac{\partial^2}{\partial r_{12}^2} + \frac{2}{r_{12}} \frac{\partial}{\partial r_{12}} - \frac{1}{r_{12}} \right) - \left(\frac{\mathbf{r}_1 \cdot \mathbf{r}_{12}}{r_1 r_{12}} \frac{\partial}{\partial r_1} + \frac{\mathbf{r}_2 \cdot \mathbf{r}_{21}}{r_2 r_{21}} \frac{\partial}{\partial r_2} \right) \frac{\partial}{\partial r_{12}}. \quad (74)$$

The terms that multiply $1/r_{12}$ at $r_{12} = 0$ must vanish in $\hat{H}\Psi$, which imposes a condition

$$\left. \frac{\partial \Psi}{\partial r_{12}} \right|_{r_{12}=0} = \frac{1}{2} \Psi(r_{12} = 0) \quad (75)$$

on the wave function. Similarly, the singularities at the nucleus are now seen to be balanced by the kinetic energy terms proportional to $1/r_i$:

$$\left. \frac{\partial \Psi}{\partial r_i} \right|_{r_i=0} = -Z\Psi(r_i = 0). \quad (76)$$

Eq. (75) describes the situation when the electrons coincide in space and is referred to as the *Coulomb cusp condition*; whereas Eq. (76) establishes the behavior of the ground-state wave function in the vicinity of the nucleus and is known as the *nuclear cusp condition*.

Expanding the ground-state helium wave function around $r_2 = r_1$ and $r_{12} = 0$ we obtain

$$\Psi(r_1, r_2, 0) = \Psi(r_1, r_1, 0) + (r_2 - r_1) \left. \frac{\partial \Psi}{\partial r_2} \right|_{r_2=r_1} + r_{12} \left. \frac{\partial \Psi}{\partial r_{12}} \right|_{r_2=r_1} + \dots,$$

which gives, when the cusp condition (75) is applied

$$\Psi(r_1, r_2, 0) = \Psi(r_1, r_1, 0) + (r_2 - r_1) \left. \frac{\partial \Psi}{\partial r_2} \right|_{r_2=r_1} + \frac{1}{2} |\mathbf{r}_2 - \mathbf{r}_1| \Psi(r_1, r_1, 0) + \dots$$

Therefore, the cusp condition leads to a wave function that is continuous but not smooth (discontinuous first derivative) at r_{12} .

The nuclear cusp condition for the “first” electron when the wave function does not vanish at $r_1 = 0$ (such as the helium ground state) is satisfied if the wave function exhibits an exponential dependence on r_1 close to the nucleus:

$$\Psi(r_1, r_2, r_{12}) = \exp(-Zr_1) \Psi(0, r_2, r_{12}) \approx (1 - Zr_1) \Psi(0, r_2, r_{12}).$$

Molecular electronic wave functions are usually expanded in simple analytical functions centered on the atomic nuclei (AOs), and close to a given nucleus, the behavior of the wave function is dominated by the analytical form of the AOs. In particular, the Slater-type orbitals, introduced in Section 1.4, are compatible with the nuclear cusp condition, while the Gaussian-type orbitals are not.

It should be noted that the cusp conditions in Eqs. (75) and (76) are written for the totally symmetric singlet ground state of the helium atom. The cusp conditions in a more general

situation (but still for a wave function that does not vanish at the singularities) should be written as

$$\lim_{r_{ij} \rightarrow 0} \left(\frac{\partial \Psi}{\partial r_{ij}} \right)_{\text{ave}} = \frac{1}{2} \Psi(r_{ij} = 0) \quad (77)$$

$$\lim_{r_i \rightarrow 0} \left(\frac{\partial \Psi}{\partial r_i} \right)_{\text{ave}} = -Z \Psi(r_i = 0), \quad (78)$$

where the averaging over all directions is implied.

Variation principle

According to the *variation principle*, the solution of the time-independent Schrödinger equation $\hat{H} |\Psi\rangle = E |\Psi\rangle$ is equivalent to an optimization of the energy functional

$$E[\tilde{\Psi}] = \frac{\langle \tilde{\Psi} | \hat{H} | \tilde{\Psi} \rangle}{\langle \tilde{\Psi} | \tilde{\Psi} \rangle}, \quad (79)$$

where $|\tilde{\Psi}\rangle$ is some approximation to the eigenstate $|\Psi\rangle$. It provides a simple and powerful procedure for generating approximate wave functions: for some proposed model for the wave function, we express the electronic state $|\mathbf{C}\rangle$ in terms of a finite set of numerical parameters; the stationary points of energy function

$$E(\mathbf{C}) = \frac{\langle \mathbf{C} | \hat{H} | \mathbf{C} \rangle}{\langle \mathbf{C} | \mathbf{C} \rangle} \quad (80)$$

are the approximate electronic states $|\mathbf{C}\rangle$ and the values $E(\mathbf{C})$ at the stationary points the approximate energies. Due to the variation principle, the expectation value of the Hamiltonian is correct to second order in the error.

Example 1.5

A simple realization of the variation method is to make a linear ansatz for the wave function,

$$|\mathbf{C}\rangle = \sum_{i=1}^m C_i |i\rangle,$$

i.e. the approximate state is expanded in m -dimensional set of Slater determinants. We further assume here that the wave function is real. The energy function of this state is given by Eq. (80). In order to locate and to characterize the stationary points, we shall employ the first and second derivatives with respect to the variational parameters:

$$E_i^{(1)}(\mathbf{C}) = \frac{\partial E(\mathbf{C})}{\partial C_i} = 2 \frac{\langle i | \hat{H} | \mathbf{C} \rangle - E(\mathbf{C}) \langle i | \mathbf{C} \rangle}{\langle \mathbf{C} | \mathbf{C} \rangle}$$

$$E_{ij}^{(2)}(\mathbf{C}) = \frac{\partial^2 E(\mathbf{C})}{\partial C_i \partial C_j} = 2 \frac{\langle i | \hat{H} | j \rangle - E(\mathbf{C}) \langle i | j \rangle}{\langle \mathbf{C} | \mathbf{C} \rangle} - 2E_i^{(1)}(\mathbf{C}) \frac{\langle j | \mathbf{C} \rangle}{\langle \mathbf{C} | \mathbf{C} \rangle} - 2E_j^{(1)}(\mathbf{C}) \frac{\langle i | \mathbf{C} \rangle}{\langle \mathbf{C} | \mathbf{C} \rangle},$$

known as the *electronic gradient* and *electronic Hessian*, respectively. The condition for stationary points, $\langle i | \hat{H} | \mathbf{C} \rangle = E(\mathbf{C}) \langle i | \mathbf{C} \rangle$ is in a matrix form equal to $\mathbf{H}\mathbf{C} = E(\mathbf{C})\mathbf{S}\mathbf{C}$, where the Hamiltonian and overlap matrix are given by $H_{ij} = \langle i | \hat{H} | j \rangle$ and $S_{ij} = \langle i | j \rangle$. Assuming that $\mathbf{S} = \mathbf{1}$, we have to solve a standard m -dimensional eigenvalue problem with m orthonormal solutions $\mathbf{C}_K = (C_{1K} C_{2K} \cdots C_{mK})^T$, $\mathbf{C}_K^T \mathbf{C}_K = \delta_{KL}$, with the associated real eigenvalues $E_K = E(\mathbf{C}_K)$, $E_1 \leq E_2 \leq \cdots \leq E_m$. The eigenvectors represent the approximate wave functions $|K\rangle = \sum_{i=1}^M C_{iK} |i\rangle$ with a corresponding approximate energy E_K . To characterize the stationary points, we note that the Hessian is at these points ${}^K E_{ij}^{(2)}(\mathbf{C}_K) = 2(\langle i | \hat{H} | j \rangle - E_k \langle i | j \rangle)$. We may also express the Hessian in the basis formed of the eigenvectors:

$${}^K E_{MN}^{(2)} = 2(\langle M | \hat{H} | N \rangle - E_K \langle M | N \rangle) = 2(E_M - E_N)\delta_{MN} \Rightarrow {}^K E_{MM}^{(2)} = 2(E_M - E_K).$$

$E_{MM}^{(2)}$ is thus singular and K^{th} state has exactly $K - 1$ negative eigenvalues. Therefore, in the space orthogonal to \mathbf{C}_K , the first solution is a minimum, the second a first-order saddle point, the third a second-order saddle point, and so forth.

Hellmann–Feynman theorem

Many of the theorems for exact wave functions hold also for the approximate ones that are variationally determined. One of the most important is the *Hellmann–Feynman theorem*, which states that the first-order change in the energy due to a perturbation may be calculated as the expectation value of the perturbation operator \hat{V} : Let $|\Psi_\alpha\rangle$ be the wave function associated with $\hat{H} + \alpha\hat{V}$ and $|\Psi\rangle$ unperturbed wave function at $\alpha = 0$. Then

$$\begin{aligned} \left. \frac{dE(\alpha)}{d\alpha} \right|_{\alpha=0} &= \left. \frac{\partial}{\partial \alpha} \frac{\langle \Psi_\alpha | \hat{H} + \alpha \hat{V} | \Psi_\alpha \rangle}{\langle \Psi_\alpha | \Psi_\alpha \rangle} \right|_{\alpha=0} = 2 \operatorname{Re} \left\langle \left. \frac{\partial \Psi_\alpha}{\partial \alpha} \right|_{\alpha=0} \middle| \hat{H} - E(0) | \Psi \right\rangle + \langle \Psi | \hat{V} | \Psi \rangle \\ &= \langle \Psi | \hat{V} | \Psi \rangle. \end{aligned} \quad (81)$$

In fact, for the Hellmann–Feynman theorem to hold, we need to demand from the approximate wave functions that they are optimized with respect to the changes induced by the perturbation: $|\Psi\rangle \rightarrow |\Psi\rangle + \alpha |\partial\Psi/\partial\alpha\rangle$. Usually in molecular calculations we construct our wave function from a finite set of analytical functions attached to the atomic nuclei. For example, when we distort the molecular geometry, we change the basis set in terms of which our electronic wave function is expanded, and a wave function optimized at one particular geometry is not accurately represented in terms of a basis set associated with another geometry. As a result, the unperturbed electronic is not variational with respect to these changes and the conditions of the Hellmann–Feynman theorem is not satisfied.

However, this is an artificial difficulty arising from the coordinate representation of quantum mechanics we presently operate in. The picture presented by the second quantization, the

necessary conditions are fulfilled and we are able to employ the theorem for approximate wave functions in finite bases in exactly the same manner as for exact wave functions.

The molecular electronic virial theorem

The exact molecular energy is variational with respect to an arbitrary change in the wave function, particularly of interest are the uniform, normalization-preserving scaling of the electron coordinates

$$\Psi(\mathbf{r}_i) \rightarrow \Psi_\alpha(\mathbf{r}_i) = \alpha^{3N/2} \Psi(\alpha \mathbf{r}_i).$$

Let us partition the electronic Hamiltonian in a bit different manner than earlier:

$$\hat{H}(\mathbf{R}) = T + V(\mathbf{R}) \quad (82)$$

$$\hat{T} = -\frac{1}{2} \sum_i \nabla_i^2 \quad (83)$$

$$\hat{V}(\mathbf{R}) = \sum_{i>j} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} - \sum_{iI} \frac{Z_I}{|\mathbf{r}_i - \mathbf{R}_I|} + \sum_{I>J} \frac{Z_I Z_J}{|\mathbf{R}_I - \mathbf{R}_J|}, \quad (84)$$

and applying this decomposition we may find the stationary points of the energy function are found at^{||}

$$\begin{aligned} 0 &= \frac{\partial}{\partial \alpha} \langle \Psi_\alpha | \hat{H}(\mathbf{R}) | \Psi_\alpha \rangle \Big|_{\alpha=1} = \frac{\partial}{\partial \alpha} \left[\alpha^2 \langle \Psi | \hat{T} | \Psi \rangle + \alpha \langle \Psi | \hat{V}(\alpha \mathbf{R}) | \Psi \rangle \right] \Big|_{\alpha=1} \\ &\Rightarrow 2 \langle \Psi | \hat{T} | \Psi \rangle + \langle \Psi | \hat{V}(\mathbf{R}) | \Psi \rangle = - \left\langle \Psi \left| \frac{dV(\alpha \mathbf{R})}{d\alpha} \right|_{\alpha=1} \right| \Psi \rangle \stackrel{\text{H-F}}{=} - \frac{dE(\alpha \mathbf{R})}{d\alpha} \Big|_{\alpha=1}. \end{aligned} \quad (85)$$

This result (85) is the quantum-mechanical *virial theorem* for a field-free non-relativistic molecular Hamiltonian. The most important observation is that in the molecular equilibrium

^{||}Note that the unperturbed wave function corresponds here to $\alpha = 1$ instead of $\alpha = 0$.

geometries ($\mathbf{R} = \mathbf{R}_e$), and in the case of atoms

$$\langle \Psi | \hat{T} | \Psi \rangle = -\frac{1}{2} \langle \Psi | \hat{V}(\mathbf{R}_e) | \Psi \rangle. \quad (86)$$

Furthermore,

$$\langle \Psi | \hat{T} | \Psi \rangle |_{\mathbf{R}=\mathbf{R}_e} = -E(\mathbf{R}_e), \quad \langle \Psi | \hat{V}(\mathbf{R}_e) | \Psi \rangle = 2E(\mathbf{R}_e). \quad (87)$$

Also these expressions hold for any approximate wave function that is variational with respect to a uniform scaling of the nuclear as well as electronic coordinates. Note that according to the last result, no stationary points of positive energy may exist.

The scaling force may be easily related to classical Cartesian forces on the nuclei, $\mathbf{F}_I(\mathbf{R}) = -dE(\mathbf{R})/d\mathbf{R}_I$, by invoking the chain rule

$$\left. \frac{dE(\alpha\mathbf{R})}{d\alpha} \right|_{\alpha=1} = \sum_I \left. \frac{d(\alpha\mathbf{R}_I)}{d\alpha} \right|_{\alpha=1} \left. \frac{dE(\alpha\mathbf{R}_I)}{d(\alpha\mathbf{R}_I)} \right|_{\alpha=1} = -\sum_I \mathbf{R}_I \cdot \mathbf{F}_I(\mathbf{R})$$

and combining this with virial theorem, we may extract the Cartesian forces experienced by the nuclei,

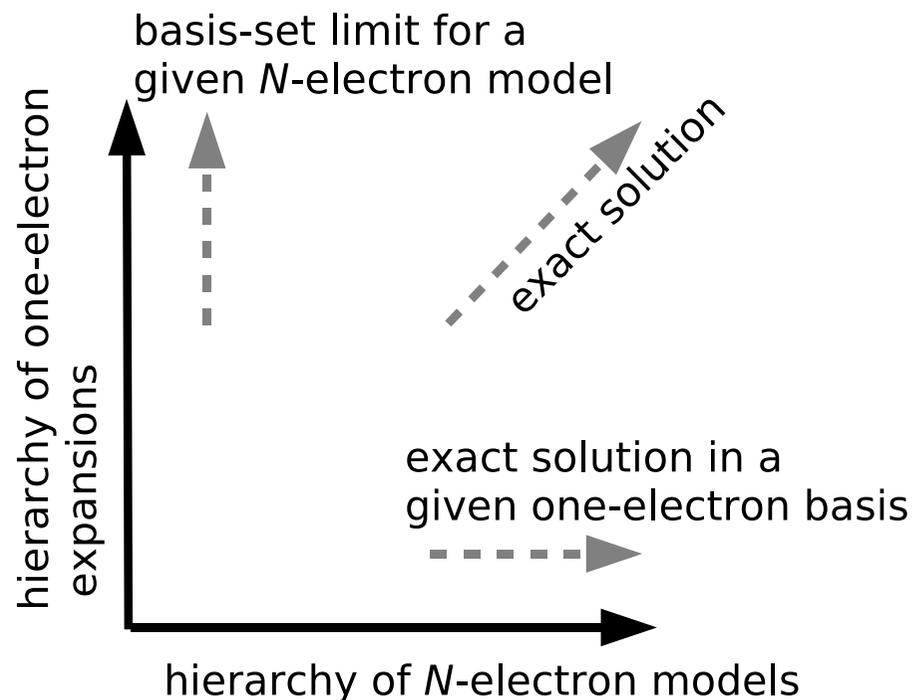
$$\sum_I \mathbf{R}_I \cdot \mathbf{F}_I(\mathbf{R}) = 2 \langle \Psi | \hat{T} | \Psi \rangle + \langle \Psi | \hat{V}(\mathbf{R}) | \Psi \rangle. \quad (88)$$

One- and N -electron expansions

Let us recall: the Slater determinants do not represent an accurate solution to the molecular Schrödinger equation, but we may express the exact wave function as a superposition of determinants. Furthermore, we have to expand the molecular orbitals in the determinants as a linear combination of atomic orbitals

$$\phi_a = \sum_{\mu} C_{\mu a} \chi_{\mu}(\mathbf{r}). \quad (89)$$

As we will see later, atomic orbitals are further approximated with a set of simpler functions that are also functions of the coordinates of a single electron. This three-step procedure is often referred to as the “standard model of quantum chemistry”. To arrive at the exact solution, i) orbitals in terms of which the determinants are constructed must form a complete set in one-electron space, and ii) in the expansion of N -electron wave function, we must include all determinants that can be generated from these orbitals.



1.4 Atomic basis functions

Molecular orbitals can be constructed either numerically or algebraically, of which the former provides greater flexibility and accuracy, but is tractable only for atoms and diatomic molecules; and for polyatomic systems we are forced to expand the MOs $\phi_a(\mathbf{r})$ in a set of some simple analytical one-electron functions.

Perhaps surprisingly, it turns out that the hydrogenic functions Eq. (8) are not ideal for this purpose; as they do not constitute a complete set by themselves but must be supplemented by the unbounded continuum states. Secondly, they become very diffuse and a large number of them is needed for a proper description of both the core and valence regions of a many-electron atom, especially when located in a molecule.

Therefore, in the molecular electronic-structure theory, χ 's are usually functions of slightly more unphysical origin. Ideally, they should

1. allow orderly and systematic extension towards completeness with respect to one-electron square-integrable functions
2. allow rapid convergence to any atomic or molecular electronic state, requiring only a few terms for a reasonably accurate description of molecular electronic distributions
3. have an analytical form for easy and accurate manipulation especially for molecular integrals
4. be orthogonal or at least their non-orthogonality should not pose problems related to numerical instability.

Let us begin by reviewing the properties of the hydrogenic functions. As we noted earlier, they are the bounded eigenfunctions of the Hamiltonian

$$H = -\frac{1}{2}\nabla^2 - \frac{Z}{r}.$$

and are written as

$$\chi_{nlm}(r, \theta, \varphi) = \underbrace{\left(\frac{2Z}{n}\right)^{3/2} \sqrt{\frac{(n-l-1)!}{2n(n+1)!}} \left(\frac{2Zr}{n}\right)^l L_{n-l-1}^{2l+1} \left(\frac{2Zr}{n}\right) \exp\left(-\frac{Zr}{n}\right)}_{R_{nl}(r)} \times \underbrace{\sqrt{\frac{2l+1}{4\pi} \frac{(l-m)!}{(l+m)!}} P_l^m(\cos\theta) \exp(im\varphi)}_{Y_{lm}(\theta, \varphi)}.$$

The energy of the bounded hydrogenic state χ_{nlm} is given by

$$E_n = -\frac{Z^2}{2n^2}. \quad (90)$$

The degeneracy of the hydrogenic states of different angular momenta – that is, $2l + 1$ degenerate states for every l – is peculiar to the spherical symmetry of the Coulomb potential and is lifted in a many-atom system. The notable features of the radial functions include

- the presence of the exponential
- that Laguerre polynomials introduce $n - l - 1$ radial nodes in the wave function.

The diffusiveness of χ_{nlm} for large n is seen from the expectation value of \hat{r}

$$\langle \chi_{nlm} | \hat{r} | \chi_{nlm} \rangle = \frac{3n^2 - l(l+1)}{2Z} \quad (91)$$

In the following we consider basis functions that retain the product form of the hydrogenic wave functions but have a more compact radial part.

The Laguerre functions

In order to retain the correct exponential behavior but avoid the problems with continuum states we introduce the following class of functions:

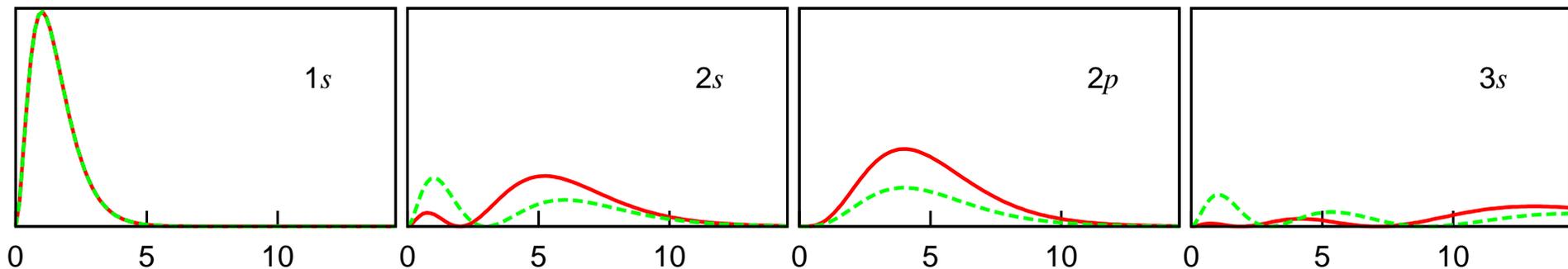
$$\chi_{nlm}^{\text{LF}}(r, \theta, \varphi) = R_{nl}^{\text{LF}}(r) Y_{lm}(\theta, \varphi) \quad (92)$$

$$R_{nl}^{\text{LF}} = (2\zeta)^{3/2} \sqrt{\frac{(n-l-1)!}{(n+l+1)!}} (2\zeta r)^l L_{n-l-1}^{2l+2}(2\zeta r) \exp(-\zeta r), \quad (93)$$

referred to as the *Laguerre functions*. For fixed l and ζ the radial functions (93) with $n > l$ constitute a complete orthonormal set of functions, for which

$$\langle \chi_{nlm}^{\text{LF}} | \hat{r} | \chi_{nlm}^{\text{LF}} \rangle = \frac{2n+1}{2\zeta}, \quad (94)$$

being thus considerably more compact than the hydrogenic functions for large n and independent of l . The Laguerre functions exhibit the same nodal structure as the hydrogenic ones but their radial distributions $R^2 r^2$ are quite different ($\zeta = 1/n$, $Z = 1$).



One should note that the orthogonality of the Laguerre functions is valid only with fixed exponents. It turns out, however, that the expansion of the orbital of a given l -value of an

atomic system requires a large number of fixed-exponent Laguerre functions with different n . For non-orthogonal, variable-exponent functions Laguerre functions are not an optimal choice, but we need to introduce functions specifically tailored to reproduce as closely as possible the different orbitals of each atom.

Slater-type orbitals

In simplifying the polynomial structure of the Laguerre functions, we note that in R_{nl}^{LF} r occurs to power l and to powers $n-l-1$ and lower. Thus a nodeless one-electron function that resembles χ_{nlm}^{LF} and the hydrogenic functions closely for large r is obtained by retaining in the Laguerre function only the term of the highest power of r , yielding the *Slater-type orbitals* (STO)

$$\chi_{nlm}^{\text{STO}}(r, \theta, \varphi) = R_{nl}^{\text{STO}}(r)Y(\theta, \varphi) \quad (95)$$

$$R_n^{\text{STO}}(r) = \frac{(2\zeta)^{3/2}}{\sqrt{\Gamma(2n+1)}}(2\zeta r)^{n-1} \exp(-\zeta r) \quad (96)$$

For the STOs we use the same $1s, 2p, \dots$ notation as for the hydrogenic functions, keeping in mind that STOs are *nodeless* and now n only refers to the monomial factor r^{n-1} . The expectation value of \hat{r} and the maximum in the radial distribution curve are given by

$$\langle \chi_{nlm}^{\text{STO}} | \hat{r} | \chi_{nlm}^{\text{STO}} \rangle = \frac{2n+1}{2\zeta} \quad (97)$$

$$r_{\text{max}}^{\text{STO}} = \frac{n}{\zeta} \quad (98)$$

Also STOs would constitute a complete set of functions for a fixed exponent ζ , and we could choose to work with a single exponent and include in our basis orbitals $(1s, 2s, \dots)$, $(2p, 3p, \dots)$, $(3d, 4d, \dots)$, all with the same exponent. Alternatively, we may describe the radial space by functions with *variable* exponents. For each l , we employ only the functions of the

lowest n , yielding a basis of the type $((1s(\zeta_{1s}), 1s(\zeta_{2s}), \dots), (2p(\zeta_{1p}), 2p(\zeta_{2p}), \dots), \dots)$. These variable-exponent STOs are given by

$$\chi_{\zeta_{nl}lm}(r, \theta, \varphi) = R_{\zeta_{nl}}^{\text{STO}}(r)Y_{lm}(\theta, \varphi) \quad (99)$$

$$R_{\zeta_{nl}}^{\text{STO}} = \frac{(2\zeta_{nl})^{3/2}}{\sqrt{(2l+2)!}}(2\zeta_{nl}r)^l \exp(-\zeta_{nl}r). \quad (100)$$

In practice, a combined approach is employed. For example, for the carbon atom, we would introduce sets of $1s$, $2s$ and $2p$ functions, all with variable exponents that are chosen to ensure an accurate representation of the wave function.

Gaussian-type orbitals

The STO basis sets are very successful for atoms and diatomic molecules, but the evaluation of many-center two-electron integrals becomes very complicated in terms of them. The basis sets that have established to be the standard practice in molecular electronic-structure theory are less connected to the hydrogenic functions or actual charge distributions in molecules. They are called *Gaussian-type orbitals* (GTO) as they employ the Gaussian distributions $\exp(-\alpha r^2)$ instead of the $\exp(-\zeta r)$ as in STOs. A large number of GTOs are needed to describe an STO (or an AO) properly, but this is more than compensated by a fast evaluation of the molecular integrals.

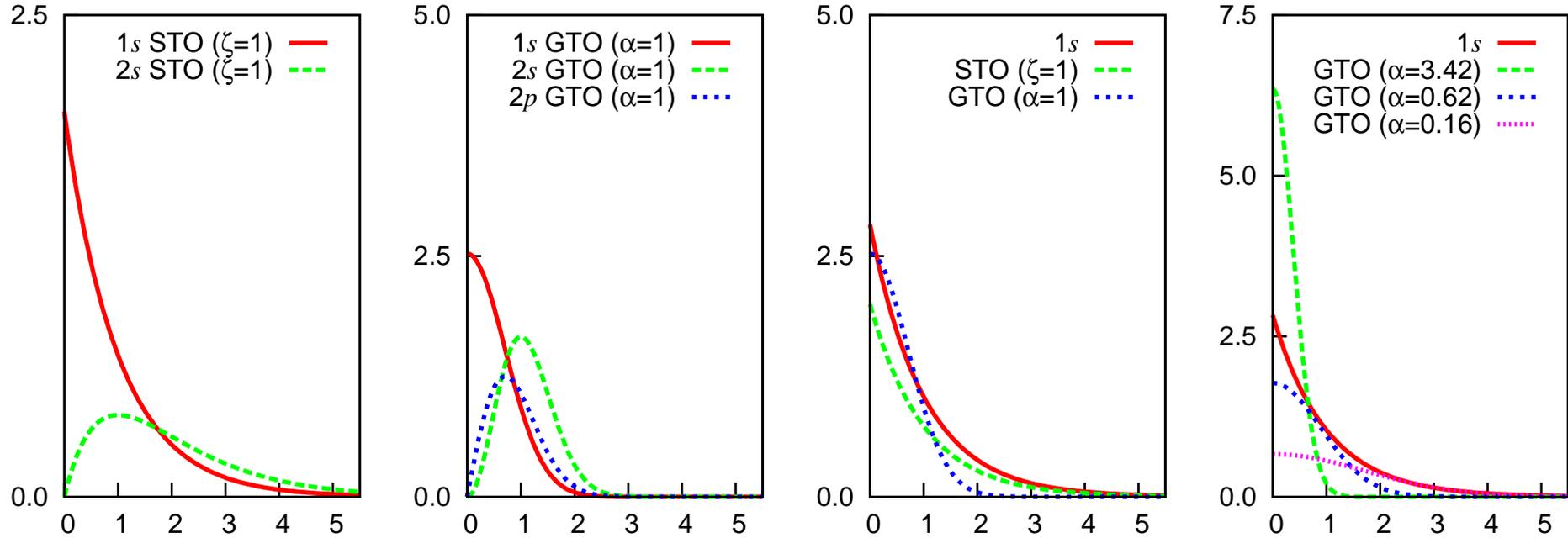
The *spherical-harmonic GTOs* are

$$\chi_{nlm}^{\text{GTO}}(r, \theta, \varphi) = R_{nl}^{\text{GTO}}(r)Y_{lm}(\theta, \varphi) \quad (101)$$

$$R_{nl}^{\text{GTO}}(r) = \frac{2(2\alpha)^{3/4}}{\pi^{1/4}} \sqrt{\frac{2^{2n-l-2}}{(4n-2l-3)!!}} (\sqrt{2\alpha}r)^{2n-l-2} \exp(-\alpha r^2). \quad (102)$$

They form a complete set of nonorthogonal basis functions. In above, the double factorial function is

$$n!! = \begin{cases} 1 & n = 0 \\ n(n-2)(n-4)\cdots 2 & \text{even } n > 0 \\ n(n-2)(n-4)\cdots 1 & \text{odd } n > 0 \\ \frac{1}{(n+2)(n+4)\cdots 1} & \text{odd } n < 0 \end{cases} \quad (103)$$



The spatial extent of the GTOs are

$$\langle \chi_{nlm}^{\text{GTO}} | \hat{r} | \chi_{nlm}^{\text{GTO}} \rangle \approx \sqrt{\frac{2n-l-2}{2\alpha}} \quad (104)$$

$$r_{\text{max}}^{\text{GTO}} = \sqrt{\frac{2n-l-1}{2\alpha}} \quad (105)$$

When we compare these results with the case of STOs, we note that much larger number of GTOs are needed for a flexible description of outer regions of the electron distribution.

In practice, the use of GTOs with variable exponents differs from the procedure for STOs. For GTOs, we describe the radial space exclusively by means of variable exponents, using for this purpose only spherical-harmonic GTOs with $n = l + 1$, only powers of r introduced are thus associated with l of the spherical harmonics. The following set of spherical harmonic GTOs is used:

$$\chi_{\alpha_{nl}lm}^{\text{GTO}}(r, \theta, \varphi) = R_{\alpha_{nl}}^{\text{GTO}}(r)Y_{lm}(\theta, \varphi) \quad (106)$$

$$R_{\alpha_{nl}}^{\text{GTO}}(r) = \frac{2(2\alpha_{nl})^{3/4}}{\pi^{1/4}} \sqrt{\frac{2^{2n-l-2}}{(4n-2l-3)!!}} (\sqrt{2\alpha}r)^{2n-l-2} \exp(-\alpha r^2). \quad (107)$$

Let us finally consider the *Gaussian product rule*, that is the one of the main reasons why GTOs are so widely used. The product of two s GTOs $\chi_A = \exp(-\alpha r_A^2)$ and $\chi_B = \exp(-\beta r_B^2)$ centered on \mathbf{r}_A and \mathbf{r}_B is a third Gaussian

$$\chi_A \chi_B = \chi_C = \exp\left(-\frac{\alpha\beta}{\alpha+\beta} r_{AB}^2\right) \exp[-(\alpha+\beta)r_C^2] \quad (108)$$

centered on \mathbf{r}_C given by

$$\mathbf{r}_C = \frac{\alpha\mathbf{r}_A + \beta\mathbf{r}_B}{\alpha + \beta}.$$

Gaussian basis sets

Of the three general features determining the accuracy of quantum chemical methods—the choice of the Hamiltonian, electron correlation treatment or truncation of the many-particle space, and the description of the one-particle space, i.e., the basis set used—the choice of the basis set is the most crucial one, as insufficient basis sets yield erroneous results regardless of the level of theory, whereas with a proper basis set, at least qualitatively correct results can be obtained for many problems already at the uncorrelated level of theory and using the non-relativistic Hamiltonian.

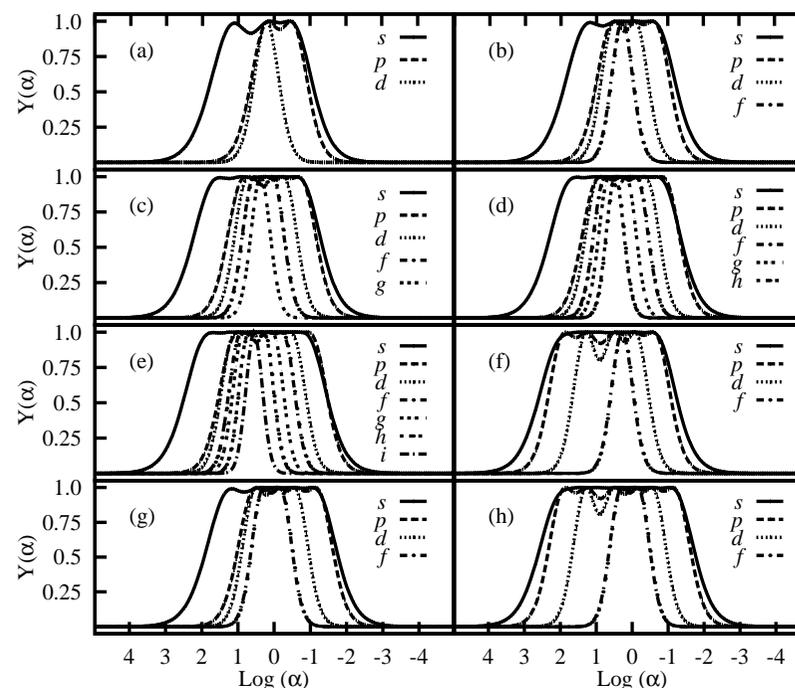
The majority of quantum chemical applications employ *contracted GTO* (CGTO) sets, i.e. linear combinations of GTO functions with coefficients optimized regarding some criterion (usually SCF energy) that significantly increase efficiency. There are two widely used contraction schemes: *segmented* contraction, where each GTO contributes only to a single CGTO, and *general* contraction, where such a restriction is not applied. Examples of the segmented scheme include the Pople-style basis set families (3-21G, 6-31G,...) and the Karlsruhe basis sets. The correlation-consistent (cc) basis set families by Dunning and co-workers are the most widely used generally contracted basis sets. The basic idea behind the cc sets is that functions that contribute approximately the same amount to the correlation energy are added to the basis in groups. The cc basis sets provide smooth, monotonic convergence for the electronic energy as well as for many molecular properties, especially those originating in the valence region. The polarized valence (cc-pVXZ, X = D, T, Q, 5, 6 corresponding to the number of CGTOs used to represent an occupied atomic orbital) and core-valence (cc-pCVXZ) sets can be augmented with diffuse functions (aug-cc-pVXZ and aug-cc-pCVXZ; d-aug-cc-pVXZ and t-aug-cc-pVXZ for doubly and triply augmented sets, respectively). While these features are favorable and the sets are widely used, the cc-basis sets become very large at large X, and it is not straightforward to extend the family beyond the published X values or to new elements. The atomic natural orbital (ANO) basis sets provide another general contraction approach. The contraction coefficients therein are atomic ANO coefficients that are obtained by optimizing atomic energies.

The use of increasingly large basis sets that produce results converging to some particular value is usually regarded as a solution to the problem of basis set incompleteness. In calculations of molecular properties that originate, e.g., in the region close to the nuclei (examples include indirect spin-spin couplings and hyperfine couplings), approaching the basis-set limit using the cc or comparable energy-optimized paradigms may lead to excessively large basis sets prohibiting calculations of large molecules. An alternative and often used approach is to uncontract the basis set and to supplement it by additional steep basis functions in the *l*-shells relevant for the property under examination.

The performance of the different basis sets for a certain molecular property can be qualitatively understood within this concept, which can be measured by the completeness profiles introduced by Chong as

$$Y(\alpha) = \sum_m \langle g(\alpha) | \chi_m \rangle^2. \quad (109)$$

Here $\{\chi\}$ denotes a set of orthonormalized, contracted or primitive, basis functions, and $g(\alpha)$ is a primitive “test” GTO with the exponent α , used to probe the completeness of $\{\chi\}$. The completeness profile becomes unity for all α , for all l -values in a CBS. $Y(\alpha)$ is typically plotted in the logarithmic scale, against $x = \log \alpha$. In a certain exponent interval $[\alpha_{\min}, \alpha_{\max}]$ this quantity is intuitively connected to the possibility – from the point of view of one-particle space – of describing all details of the wave function in the corresponding distance range from the atomic nuclei. Put simply, atomic properties that obtain relevant contributions roughly within $[1/\sqrt{\alpha_{\max}}, 1/\sqrt{\alpha_{\min}}]$ from the nucleus can be reproduced by a basis that has $Y(\alpha) = 1$ in this interval. This way of thinking can be generalized to molecular properties that may be dominated by phenomena occurring close to the expansion centers of the basis functions, i.e., atomic nuclei (region described by high-exponent basis functions) and/or in the valence region, further away from the nuclei (described by diffuse basis functions).



Completeness profiles of (a) cc-pVDZ ($[9s\ 4p\ 1d|3s\ 2p\ 1d]$), (b) cc-pVTZ ($[10s\ 5p\ 2d\ 1f|4s\ 3p\ 2d\ 1f]$), (c) cc-pVQZ ($[12s\ 6p\ 3d\ 2f\ 1g|5s\ 4p\ 5d\ 2f\ 1g]$) (d) cc-pV5Z ($[14s\ 8p\ 4d\ 3f\ 2g\ 1h|6s\ 5p\ 4d\ 3f\ 2g\ 1h]$), (e) cc-pV6Z ($[16s\ 10p\ 5d\ 4f\ 3g\ 2h\ 1i|7s\ 6p\ 5d\ 4f\ 3g\ 2h\ 1i]$), (f) cc-pCVTZ ($[12s\ 7p\ 3d\ 1f|6s\ 5p\ 3d\ 1f]$), (g) aug-cc-pVTZ ($[11s\ 6p\ 3d\ 2f|5s\ 4p\ 3d\ 2f]$), and (h) aug-cc-pCVTZ ($[13s\ 8p\ 4d\ 2f|7s\ 6p\ 4d\ 2f]$) basis sets of fluorine.

2 SECOND QUANTIZATION

2.1 Fock space

We now introduce an abstract linear vector space – the *Fock space* – where each Slater determinant (21) of M orthonormal spin-orbitals is represented by an *occupation number vector* (ON) \mathbf{k} ,

$$|\mathbf{k}\rangle = |k_1, k_2, \dots, k_M\rangle, \quad k_P = \begin{cases} 1 & \psi_P \text{ occupied} \\ 0 & \psi_P \text{ unoccupied} \end{cases} \quad (110)$$

For an orthonormal set of spin-orbitals, we define the *inner product* between two ON vectors as

$$\langle \mathbf{m} | \mathbf{k} \rangle = \prod_{P=1}^M \delta_{m_P k_P}. \quad (111)$$

This definition is consistent with the overlap between two Slater determinants, but has a well-defined but zero overlap between states with *different electron numbers* is a special feature of the Fock-space formulation. It allows for a unified description of systems with variable numbers of electrons.

- In a given spin-orbital basis, there is a one-to-one mapping between the Slater determinants with spin-orbitals in a canonical order and the ON vectors in the Fock space.
- However, ON vectors are *not* Slater determinants: ON vectors have no spatial structure but are just basis vectors in an abstract vector space.
- The Fock space can be manipulated as an ordinary inner-product vector space.

The ON vectors constitute an orthonormal basis in the 2^M dimensional Fock space $F(M)$, that can be decomposed as a direct sum of subspaces,

$$F(M) = F(M, 0) \oplus F(M, 1) \oplus \cdots \oplus F(M, N), \quad (112)$$

where $F(M, N)$ contains all ON vectors obtained by distributing N electrons among the M spin-orbitals, in other words, all ON vectors for which the sum of occupation number is N . The subspace $F(M, 0)$ is the *vacuum state*,

$$|\text{vac}\rangle = |0_1, 0_2, \dots, 0_M\rangle. \quad (113)$$

Approximations to an exact N -electron wave function are expressed in terms of vectors in the Fock subspace $F(M, N)$ of dimension equal to $\binom{M}{N}$.

Creation and annihilation operators

In second quantization, all operators and states can be constructed from a set of elementary *creation* and *annihilation* operators. The M creation operators are defined by

$$a_P^\dagger |k_1, k_2, \dots, 0_P, \dots, k_M\rangle = \Gamma_P^k |k_1, k_2, \dots, 1_P, \dots, k_M\rangle \quad (114)$$

$$a_P^\dagger |k_1, k_2, \dots, 1_P, \dots, k_M\rangle = 0 \quad (115)$$

where

$$\Gamma_P^k = \prod_{Q=1}^{P-1} (-1)^{k_Q}. \quad (116)$$

The definition can be also combined a single equation:

$$a_P^\dagger |k_1, k_2, \dots, 0_P, \dots, k_M\rangle = \delta_{k_P 0} \Gamma_P^k |k_1, k_2, \dots, 1_P, \dots, k_M\rangle \quad (117)$$

Operating twice with a_P^\dagger on an ON vector gives

$$a_P^\dagger a_P^\dagger |k_1, k_2, \dots, 0_P, \dots, k_M\rangle = a_P^\dagger \delta_{k_P 0} \Gamma_P^{\mathbf{k}} |k_1, k_2, \dots, 1_P, \dots, k_M\rangle = 0,$$

therefore

$$a_P^\dagger a_P^\dagger = 0. \quad (118)$$

For $Q > P$,

$$\begin{aligned} a_P^\dagger a_Q^\dagger |\dots, k_P, \dots, k_Q, \dots\rangle &= a_P^\dagger \delta_{k_Q 0} \Gamma_Q^{\mathbf{k}} |\dots, k_P, \dots, 1_Q, \dots\rangle \\ &= \delta_{k_P 0} \delta_{k_Q 0} \Gamma_P^{\mathbf{k}} \Gamma_Q^{\mathbf{k}} |\dots, 1_P, \dots, 1_Q, \dots\rangle. \end{aligned}$$

Reversing the order of the operators,

$$\begin{aligned} a_Q^\dagger a_P^\dagger |\dots, k_P, \dots, k_Q, \dots\rangle &= a_Q^\dagger \delta_{k_P 0} \Gamma_P^{\mathbf{k}} |\dots, 1_P, \dots, k_Q, \dots\rangle \\ &= \delta_{k_P 0} \delta_{k_Q 0} \Gamma_P^{\mathbf{k}} (-\Gamma_Q^{\mathbf{k}}) |\dots, 1_P, \dots, 1_Q, \dots\rangle. \end{aligned}$$

and combining the results, we obtain

$$a_P^\dagger a_Q^\dagger + a_Q^\dagger a_P^\dagger |\mathbf{k}\rangle = 0.$$

Substitution of dummy indices shows that this holds for $Q < P$ as well, and is true also for $Q = P$. Since $|\mathbf{k}\rangle$ is an arbitrary ON vector, we conclude the anticommutation relation

$$a_P^\dagger a_Q^\dagger + a_Q^\dagger a_P^\dagger = [a_P^\dagger, a_Q^\dagger]_+ = 0. \quad (119)$$

The properties of the *adjoint* or conjugate operators a_P can be reasoned from those of the creation operators. Thus, the adjoint operators satisfy the anticommutation relation

$$[a_P, a_Q]_+ = 0. \quad (120)$$

Let us invoke the *resolution of the identity*:

$$a_P |\mathbf{k}\rangle = \sum_{\mathbf{m}} |\mathbf{m}\rangle \langle \mathbf{m} | a_P | \mathbf{k} \rangle,$$

where, using Eq. (117),

$$\langle \mathbf{m} | a_P | \mathbf{k} \rangle = \langle \mathbf{k} | a_P | \mathbf{m} \rangle^* = \begin{cases} \delta_{m_P 0} \Gamma_P^{\mathbf{m}} & \text{if } k_Q = m_Q + \delta_{QP} \\ 0 & \text{otherwise} \end{cases}.$$

From the definition of Γ and from $k_Q = m_Q + \delta_{QP}$ we see that $\Gamma_P^{\mathbf{m}} = \Gamma_P^{\mathbf{k}}$. We may therefore write the equation as

$$\langle \mathbf{m} | a_P | \mathbf{k} \rangle = \begin{cases} \delta_{k_P 1} \Gamma_P^{\mathbf{k}} & \text{if } m_Q = k_Q - \delta_{QP} \\ 0 & \text{otherwise} \end{cases}.$$

Hence, only one term in $a_P | \mathbf{k} \rangle$ survives:

$$a_P | \mathbf{k} \rangle = \delta_{k_P 1} \Gamma_P^{\mathbf{k}} | k_1, \dots, 0_P, \dots, k_M \rangle. \quad (121)$$

The operator a_P is called the *annihilation operator*.

Combining Eqs. (117) and (121), we get

$$a_P^\dagger a_P | \mathbf{k} \rangle = \delta_{k_P 1} | \mathbf{k} \rangle \quad (122)$$

$$a_P a_P^\dagger | \mathbf{k} \rangle = \delta_{k_P 0} | \mathbf{k} \rangle \quad (123)$$

that leads to

$$a_P^\dagger a_P + a_P a_P^\dagger = 1. \quad (124)$$

For $P > Q$, we have

$$a_P^\dagger a_Q | \mathbf{k} \rangle = -\delta_{k_P 0} \delta_{k_Q 1} \Gamma_P^{\mathbf{k}} \Gamma_Q^{\mathbf{k}} | k_1, \dots, 0_Q, \dots, 1_P, \dots, k_M \rangle$$

$$a_Q a_P^\dagger | \mathbf{k} \rangle = \delta_{k_P 0} \delta_{k_Q 1} \Gamma_P^{\mathbf{k}} \Gamma_Q^{\mathbf{k}} | k_1, \dots, 0_Q, \dots, 1_P, \dots, k_M \rangle,$$

thus we have the operator identity

$$a_P^\dagger a_Q + a_Q a_P^\dagger = 0 \quad P > Q. \quad (125)$$

Hence,

$$a_P^\dagger a_Q + a_Q a_P^\dagger = [a_P^\dagger, a_Q]_+ = \delta_{PQ}. \quad (126)$$

Number-conserving operators

The *occupation-number operator* that counts the number of electrons in spin-orbital P is introduced as

$$\hat{N}_P = a_P^\dagger a_P, \quad \hat{N}_P |\mathbf{k}\rangle = k_P |\mathbf{k}\rangle \quad (127)$$

The occupation-number operators are Hermitian, $\hat{N}^\dagger = \hat{N}$, as well as idempotent, $\hat{N}_P^2 = \hat{N}_P$. Using the basic anticommutation relations, we obtain

$$[\hat{N}_P, a_Q^\dagger] = \delta_{PQ} a_Q^\dagger \quad (128)$$

$$[\hat{N}_P, a_Q] = -\delta_{PQ} a_Q \quad (129)$$

So for an arbitrary string \hat{X} consisting of creation and annihilation operators, e.g. $\hat{X} = a_P^\dagger a_Q a_P a_R^\dagger a_S$, the commutators with \hat{N} become

$$[\hat{N}_P, \hat{X}] = N_P^X \hat{X}, \quad (130)$$

where N_P^X is the number of times a_P^\dagger occurs in \hat{X} minus the number of times a_P occurs.

All occupation-number operators added together in the Fock space gives the Hermitian operator

$$\hat{N} = \sum_{P=1}^M a_P^\dagger a_P, \quad (131)$$

which returns the *number of electrons in an ON vector*, $\hat{N} |\mathbf{k}\rangle = N |\mathbf{k}\rangle$, and is therefore known as the *particle-number operator*. We find that the number operator commutes with any string \hat{T} that contains an equal number of creation and annihilation operators.

Such strings are called *number-conserving*. In general, the application of the string \hat{X} to a Fock-space vector increases the number of electrons by N^X .

The simplest number-conserving operators are the elementary *excitation operators*

$$\hat{E}_Q^P = a_P^\dagger a_Q, \quad (132)$$

which give as applied to an ON vector

$$\begin{aligned} a_P^\dagger a_Q |\mathbf{k}\rangle &= a_P^\dagger \delta_{k_Q 1} \Gamma_Q^{\mathbf{k}} |k_1, \dots, 0_Q, \dots, k_P, \dots, k_M\rangle \\ &= \delta_{k_P 0} \delta_{k_Q 1} \Gamma_Q^{\mathbf{k}} \Gamma_P^{\mathbf{k}} \varepsilon_{PQ} |k_1, \dots, 0_Q, \dots, 1_P, \dots, k_M\rangle \\ &= \varepsilon_{PQ} \Gamma_Q^{\mathbf{k}} \Gamma_P^{\mathbf{k}} (1 - k_P) k_Q \left| \begin{array}{l} k_P \rightarrow 1 \\ k_Q \rightarrow 0 \end{array} \right\rangle \end{aligned} \quad (133)$$

where

$$\varepsilon_{PQ} = \begin{cases} 1 & P \leq Q \\ -1 & P > Q \end{cases}.$$

The case $P < Q$ differs from this result only in the interpretation of ε_{PQ} , whereas the case $P = Q$ is covered by Eq. (127). Therefore we may write

$$a_P^\dagger a_Q |\mathbf{k}\rangle = \varepsilon_{PQ} \Gamma_P^{\mathbf{k}} \Gamma_Q^{\mathbf{k}} (1 - k_P + \delta_{PQ}) k_Q \left| \begin{array}{l} k_P \rightarrow 1 \\ k_Q \rightarrow \delta_{PQ} \end{array} \right\rangle. \quad (134)$$

The application of a single such operator gives a single excitation, two a double excitation and so forth.

The representation of one- and two-electron operators

First quantization	Second quantization
<ul style="list-style-type: none"> One-electron operator: $\hat{f}^c = \sum_{i=1}^N f(\mathbf{x}_i)$ <ul style="list-style-type: none"> Two-electron operator: $\hat{g}^c = \sum_{i=1}^N \sum_{j=i+1}^N g(\mathbf{x}_i, \mathbf{x}_j)$ <ul style="list-style-type: none"> Operators are independent of the spin-orbital basis Operators depend on the number of electrons Operators are exact 	<ul style="list-style-type: none"> One-electron operator: $\hat{f} = \sum_{PQ} f_{PQ} a_P^\dagger a_Q$ <ul style="list-style-type: none"> Two-electron operator: $\hat{g} = \frac{1}{2} \sum_{PQRS} g_{PQRS} a_P^\dagger a_R^\dagger a_S a_Q$ <ul style="list-style-type: none"> Operators depend on the spin-orbital basis Operators are independent of of electrons Projected operators

Expectation values of Hermitian operators correspond to physical observables and should thus be independent of the representation of operators and states. Therefore, we require the matrix element of a second-quantization operator between two ON vectors to be equal to its first-quantization counterpart.

To determine the numerical parameters f_{PQ} in the second-quantization representation of one-electron operators, we evaluate the matrix elements of \hat{f} between two ON vectors and compare with the usual Slater rules for matrix elements between determinants. We are lead to the identification

$$f_{PQ} = \langle \psi_P | \hat{f}^c | \psi_Q \rangle \quad (135)$$

$$g_{PQRS} = \langle \psi_P \psi_R | \hat{g}^c | \psi_Q \psi_S \rangle. \quad (136)$$

The parameters g_{PQRS} possess the particle-interchange symmetry,

$$g_{PQRS} = g_{RSPQ}. \quad (137)$$

In the case of real spin-orbitals, the integrals exhibit the following permutational symmetries:

$$f_{PQ} = f_{QP} \quad (138)$$

$$g_{PQRS} = g_{QPRS} = g_{PQSR} = g_{QP SR}, \quad (139)$$

of which the latter may be further combined with Eq. (137) to yield an eight-fold symmetry.

Using the introduced second-quantization representation, we may now construct the (non-relativistic, spin-free) molecular electronic Hamiltonian in the Born–Oppenheimer approximation,

$$\hat{H} = \sum_{PQ} h_{PQ} a_P^\dagger a_Q + \frac{1}{2} g_{PQRS} a_P^\dagger a_R^\dagger a_S a_Q + h_{\text{nuc}}, \quad (140)$$

where

$$h_{PQ} = \int \psi_P^*(\mathbf{x}) \left(-\frac{1}{2} \nabla^2 - \sum_I \frac{Z_I}{r_{iI}} \right) \psi_Q(\mathbf{x}) d\mathbf{x} \quad (141)$$

$$g_{PQRS} = \int \frac{\psi_P^*(\mathbf{x}) \psi_R^*(\mathbf{x}) \psi_Q(\mathbf{x}) \psi_S(\mathbf{x})}{r_{12}} d\mathbf{x}_1 d\mathbf{x}_2 \quad (142)$$

$$h_{\text{nuc}} = \frac{1}{2} \sum_{I \neq J} \frac{Z_I Z_J}{R_{IJ}}. \quad (143)$$

The form of the second-quantization Hamiltonian can be interpreted in the following way: Applied to an electronic state, the Hamiltonian produces a linear combination of the original state with states generated by single and double electron excitations from this state. With each such excitation, there is an associated *amplitude* h_{PQ} or g_{PQRS} , which represents the probability of this event happening.

Let \hat{A}^c , \hat{B}^c and \hat{C}^c be one-electron operators in first quantization and \hat{A} , \hat{B} and \hat{C} be the corresponding second-quantization representations.

- The first-quantization operator $a\hat{A}^c + b\hat{B}^c$, where a and b are numbers, is represented by $a\hat{A} + b\hat{B}$.
- The standard relations

$$\hat{A}(\hat{B}\hat{C}) = (\hat{A}\hat{B})\hat{C} \quad (144)$$

$$(\hat{A}\hat{B})^\dagger = \hat{B}^\dagger\hat{A}^\dagger \quad (145)$$

are valid.

- For a *complete* one-electron basis

$$\hat{C}^c = \hat{A}^c\hat{B}^c \Rightarrow \hat{C} = \hat{A}\hat{B}$$

but for *finite* basis sets this expression does not hold. The second-quantization operators are projections of the exact operators onto a basis of spin-orbitals. For an incomplete basis, the second-quantization representation of an operator product depends on when the projection is made.

Second quantization treats operators and wave functions in a unified way – they are all expressed in terms of the elementary creation and annihilation operators: any ON vector may be written as

$$|\mathbf{k}\rangle = \hat{X}_{\mathbf{k}}|\text{vac}\rangle = \left[\prod_{P=1}^M (a_P^\dagger)^{k_P} \right] |\text{vac}\rangle \quad (146)$$

and therefore the matrix elements can be viewed as the *vacuum expectation value* of an operator

$$\langle \mathbf{k} | \hat{O} | \mathbf{m} \rangle = \langle \text{vac} | \hat{X}_{\mathbf{k}}^\dagger \hat{O} \hat{X}_{\mathbf{m}} | \text{vac} \rangle, \quad (147)$$

and expectation values become linear combinations of vacuum expectation values.

Commutators and anticommutators

In the manipulation of operators and matrix elements in second quantization, the commutator $[\hat{A}, \hat{B}] = \hat{A}\hat{B} - \hat{B}\hat{A}$ and the anticommutator $[\hat{A}, \hat{B}]_+ = \hat{A}\hat{B} + \hat{B}\hat{A}$ of two operators are often encountered. Referring to the basic relations of the elementary creation and annihilation operators, it is usually possible to simplify the commutators and anticommutators between strings of elementary operators considerably.

We need to introduce the concepts of operator *rank* and rank reduction.

- The (particle) rank of a string of creation and annihilation operators is simply *the number of elementary operators divided by 2*.
- Rank reduction is said to occur when the rank of a commutator or anticommutator is lower than the combined rank of the operators commuted or anticommutated.

A simple rule

- Rank reduction follows upon anticommutation of two strings of half-integral rank and upon commutation of all other strings

is useful for simplifying the expressions.

One useful strategy for the evaluation of commutators and anticommutators is based on their linear expansion in simpler commutators or anticommutators according to e.g. the following operator identities (see Exercise 3), and aim at greatest possible rank reduction:

$$[\hat{A}, \hat{B}_1 \hat{B}_2] = [\hat{A}, \hat{B}_1] \hat{B}_2 + \hat{B}_1 [\hat{A}, \hat{B}_2] \quad (148)$$

$$[\hat{A}, \hat{B}_1 \cdots \hat{B}_n] = \sum_{k=1}^n \hat{B}_1 \cdots \hat{B}_{k-1} [\hat{A}, \hat{B}_k] \hat{B}_{k+1} \cdots \hat{B}_n \quad (149)$$

$$[\hat{A}, \hat{B}_1 \hat{B}_2] = [\hat{A}, \hat{B}_1]_+ \hat{B}_2 - \hat{B}_1 [\hat{A}, \hat{B}_2]_+ \quad (150)$$

$$[\hat{A}, \hat{B}_1 \cdots \hat{B}_n] = \sum_{k=1}^n (-1)^{k-1} \hat{B}_1 \cdots [\hat{A}, \hat{B}_k]_+ \cdots \hat{B}_n \quad (n \text{ even}) \quad (151)$$

$$[\hat{A}, \hat{B}_1 \hat{B}_2]_+ = [\hat{A}, \hat{B}_1] \hat{B}_2 + \hat{B}_1 [\hat{A}, \hat{B}_2]_+ = [\hat{A}, \hat{B}_1]_+ \hat{B}_2 - \hat{B}_1 [\hat{A}, \hat{B}_2] \quad (152)$$

$$[\hat{A}, \hat{B}_1 \cdots \hat{B}_n]_+ = \sum_{k=1}^n (-1)^{k-1} \hat{B}_1 \cdots [\hat{A}, \hat{B}_k]_+ \cdots \hat{B}_n \quad (n \text{ odd}). \quad (153)$$

Example 2.1

Let us consider the simplest nontrivial commutator, $[a_P^\dagger, a_Q^\dagger a_R]$. Moving one of the operators after the comma out of the commutators by using (150),

$$[a_P^\dagger, a_Q^\dagger a_R] = [a_P^\dagger, a_Q^\dagger]_+ a_R - a_Q^\dagger [a_P^\dagger, a_R]_+$$

and invoking the basic anticommutation relation we have

$$[a_P^\dagger, a_Q^\dagger a_R] = -\delta_{PR} a_Q^\dagger.$$

Similarly,

$$[a_P, a_Q^\dagger a_R] = \delta_{PQ} a_R.$$

Example 2.2

The commutator between two excitation operators, using the results from the previous example,

$$[a_P^\dagger a_Q, a_R^\dagger a_S] = [a_P^\dagger, a_R^\dagger a_S] a_Q + a_P^\dagger [a_Q, a_R^\dagger a_S] = \delta_{QR} a_P^\dagger a_S - \delta_{PS} a_R^\dagger a_Q.$$

Example 2.3

A slightly more complicated commutator is evaluated by applying Eq. (148) and the previous result:

$$\begin{aligned} [a_P^\dagger a_Q, a_R^\dagger a_S a_M^\dagger a_N] &= [a_P^\dagger a_Q, a_R^\dagger a_S] a_M^\dagger a_N + a_R^\dagger a_S [a_P^\dagger a_Q, a_M^\dagger a_N] \\ &= \delta_{QR} a_P^\dagger a_S a_M^\dagger a_N - \delta_{PS} a_R^\dagger a_Q a_M^\dagger a_N + \delta_{QM} a_R^\dagger a_S a_P^\dagger a_N - \delta_{PN} a_R^\dagger a_S a_M^\dagger a_Q. \end{aligned}$$

Example 2.4

The following *double commutator* can be evaluated by invoking the result in Example 2.2 twice

$$\begin{aligned} [a_P^\dagger a_Q, [a_R^\dagger a_S, a_M^\dagger a_N]] &= \delta_{SM} [a_P^\dagger a_Q, a_R^\dagger a_N] - \delta_{RN} [a_P^\dagger a_Q, a_M^\dagger a_S] \\ &= \delta_{SM} \delta_{QR} a_P^\dagger a_N - \delta_{SM} \delta_{PM} a_R^\dagger a_Q - \delta_{RN} \delta_{QM} a_P^\dagger a_S + \delta_{RN} \delta_{PS} a_M^\dagger a_Q. \end{aligned}$$

This gives rise to a rank reduction by 2.

Orbital rotations

In many situations, e.g., during the optimization of an approximate electronic state or in the calculation of the response of an electronic state to an external perturbation, it becomes necessary to carry out transformations between different sets of orthonormal spin-orbitals.

We shall encounter especially cases where the occupation numbers refer to a set of orthonormal spin-orbitals $\tilde{\psi}_P$ that is obtained from another set ψ_P by a *unitary transformation*

$$\tilde{\psi}_P = \sum_Q \psi_Q U_{QP}. \quad (154)$$

Here, a *unitary matrix* \mathbf{U} is a matrix that fulfills the relation

$$\mathbf{U}^\dagger \mathbf{U} = \mathbf{U} \mathbf{U}^\dagger = \mathbf{1}. \quad (155)$$

The *spectral theorem* states that any unitary matrix can be diagonalized as

$$\mathbf{U} = \mathbf{V} \boldsymbol{\epsilon} \mathbf{V}^\dagger,$$

where \mathbf{V} is unitary and $\boldsymbol{\epsilon}$ a complex diagonal matrix, $\epsilon_{kk} = \exp(i\delta_k)$. Therefore any unitary matrix can be written as the matrix exponential* of an anti-Hermitian matrix:

$$\begin{aligned} \mathbf{U} &= \mathbf{V} \exp(i\boldsymbol{\delta}) \mathbf{V}^\dagger = \exp(i\mathbf{V}\boldsymbol{\delta}\mathbf{V}^\dagger) \\ &\equiv \exp(\mathbf{X}), \quad \mathbf{X} = -\mathbf{X}^\dagger. \end{aligned} \quad (156)$$

By an introduction of the anti-Hermitian operator

$$\hat{\kappa} = \sum_{PQ} \kappa_{PQ} a_P^\dagger a_Q, \quad \hat{\kappa}^\dagger = -\hat{\kappa}, \quad (157)$$

where the parameters κ_{PQ} are the elements of an anti-Hermitian matrix $\boldsymbol{\kappa}$, for which $\mathbf{U} = \exp(-\boldsymbol{\kappa})$, the elementary operators \tilde{a}_P^\dagger , \tilde{a}_P and state $|\tilde{0}\rangle$ generated by the unitary transformation (154) can be expressed in terms of the untransformed operators and states as

$$\tilde{a}_P^\dagger = \exp(-\hat{\kappa}) a_P^\dagger \exp(\hat{\kappa}) \quad (158)$$

$$\tilde{a}_P = \exp(-\hat{\kappa}) a_P \exp(\hat{\kappa}) \quad (159)$$

$$|\tilde{0}\rangle = \exp(-\hat{\kappa}) |0\rangle. \quad (160)$$

*The matrix exponential is defined as $\exp(\mathbf{A}) = \sum_{n=0}^{\infty} \mathbf{A}^n / n!$.

2.1 Spin in second quantization

In the formalism of second quantization, there is no reference to electron spin. However, in non-relativistic theory, many important simplifications follow by taking spin explicitly into account. Therefore, we shall separate the orbital space and the spin space,[†] i.e. a spin orbital will be denoted by $\psi_P(\mathbf{x}) = \phi_p(\mathbf{r})\sigma(m_s) = \phi_{p\sigma}$. The theory of second quantization holds unchanged; for example the basic anticommutator now becomes

$$[a_{p\sigma}^\dagger, a_{q\tau}]_+ = \delta_{p\sigma, q\tau} = \delta_{pq}\delta_{\sigma\tau}, \quad (161)$$

where, for example, $a_{p\sigma}^\dagger$ is the creation operator associated with the spin-orbital $\phi_{p\sigma}$.

Operators in the orbital basis

We may classify the quantum-mechanical operators according to how they affect the orbital and spin parts of wave functions: *Spin-free operators*, *spin operators* and *mixed operators*.

One-electron are written in the spin-orbital basis as

$$\hat{f} = \sum_{p\sigma} \sum_{q\tau} f_{p\sigma, q\tau} a_{p\sigma}^\dagger a_{q\tau}.$$

The integrals vanish for opposite spins in the case of spin-free (or spinless) operators:

$$\begin{aligned} f_{p\sigma, q\tau} &= \int \phi_p^*(\mathbf{r})\sigma^*(m_s) f^c(\mathbf{r})\phi_q(\mathbf{r})\tau(m_s) d\mathbf{r} dm_s \\ &= \delta_{\sigma\tau} \int \phi_p^*(\mathbf{r}) f^c(\mathbf{r})\phi_q(\mathbf{r}) d\mathbf{r} = f_{pq}\delta_{\sigma\tau} = \sum_{pq} f_{pq} \hat{E}_{pq}, \end{aligned}$$

where we have introduced the *singlet excitation operator*

$$\hat{E}_{pq} = a_{p\alpha}^\dagger a_{q\alpha} + a_{p\beta}^\dagger a_{q\beta}. \quad (162)$$

[†]Hereafter, the spin functions α and β will be denoted generically by σ , τ , μ and ν .

The two-electron operators are given by

$$\hat{g} = \frac{1}{2} \sum_{p\sigma q\tau r\mu s\nu} g_{p\sigma, q\tau, r\mu, s\nu} a_{p\sigma}^\dagger a_{r\mu}^\dagger a_{s\nu} a_{q\tau}.$$

Here, the orthogonality of the spin functions reduce the integrals to

$$g_{p\sigma, q\tau, r\mu, s\nu} = g_{pqrs} \delta_{\sigma\tau} \delta_{\mu\nu}, \quad (163)$$

where g_{pqrs} are the two-electron integrals in ordinary space, that are always symmetric with respect to the particle-interchange, $g_{pqrs} = g_{rspq}$, and have the Hermitian symmetry $g_{pqrs} = g_{qpsr}^*$ in the case of complex orbitals; and in the case of real orbitals, we have a permutational symmetry

$$g_{pqrs} = g_{qprs} = g_{pqsr} = g_{qpsr}.$$

The second-quantization representation of a spin-free two-electron operator is then

$$\hat{g} = \frac{1}{2} \sum_{pqrs} g_{pqrs} \sum_{\sigma\tau} a_{p\sigma}^\dagger a_{r\tau}^\dagger a_{s\tau} a_{q\sigma} = \frac{1}{2} \sum_{pqrs} g_{pqrs} (\hat{E}_{pq} \hat{E}_{rs} - \delta_{qr} E_{ps}) = \frac{1}{2} \sum_{pqrs} g_{pqrs} \hat{e}_{pqrs}. \quad (164)$$

We have introduced the *two-electron singlet excitation operator* in the last equality. We can now express the spin-free, non-relativistic molecular electronic Hamiltonian in the orbital basis:

$$\hat{H} = \sum_{pq} h_{pq} \hat{E}_{pq} + \frac{1}{2} \sum_{pqrs} g_{pqrs} \hat{e}_{pqrs}. \quad (165)$$

The one- and two-electron integrals are the same as in (140), except that the integrations are over the spatial coordinates only.

The one-electron second-quantization operators associated to operators that work in spin space only may be written in the general form

$$\hat{f} = \sum_{p\sigma q\tau} \int \phi_p^*(\mathbf{r}) \sigma^*(m_s) f^c(m_s) \phi_q(\mathbf{r}) \tau(m_s) d\mathbf{r} dm_s a_{p\sigma}^\dagger a_{q\tau} = \sum_{\sigma\tau} \int \sigma^*(m_s) f^c(m_s) \tau(m_s) dm_s \sum_p a_{p\sigma}^\dagger a_{p\tau}.$$

Three important examples of pure spin operators are the *raising and lowering operators* \hat{S}_+^c and \hat{S}_-^c , and \hat{S}_z^c . The effect of these operators on the (one-particle state) spin functions is

$$\begin{aligned} S_+^c \beta &= \alpha & \hat{S}_+^c \alpha &= 0 \\ S_-^c \beta &= 0 & \hat{S}_-^c \alpha &= \beta \\ S_z^c \beta &= -\frac{1}{2}\beta & \hat{S}_z^c \alpha &= +\frac{1}{2}\alpha. \end{aligned}$$

From these we arrive at the following expressions for the basic spin operators:

$$\hat{S}_+ = \sum_p a_{p\alpha}^\dagger a_{p\beta} \quad (166)$$

$$\hat{S}_- = \sum_p a_{p\beta}^\dagger a_{p\alpha} \quad (167)$$

$$\hat{S}_z = \frac{1}{2} \sum_p \left(a_{p\alpha}^\dagger a_{p\alpha} - a_{p\beta}^\dagger a_{p\beta} \right). \quad (168)$$

The lowering operator is seen to be the Hermitian adjoint of the raising operator:

$$\hat{S}_+^\dagger = \sum_p (a_{p\alpha}^\dagger a_{p\beta})^\dagger = \sum_p a_{p\beta}^\dagger a_{p\alpha} = \hat{S}_-.$$

The operators for the x and y components of the spin angular momentum can be written in terms of the raising and lowering operators as

$$\begin{aligned} S_x^c &= \frac{1}{2} (S_+^c + S_-^c) \\ S_y^c &= \frac{1}{2i} (S_+^c - S_-^c), \end{aligned}$$

from which we obtain the second-quantization counterparts,

$$\hat{S}_x = \frac{1}{2} \sum_p \left(a_{p\alpha}^\dagger a_{p\beta} + a_{p\beta}^\dagger a_{p\alpha} \right) \quad (169)$$

$$\hat{S}_y = \frac{1}{2i} \sum_p \left(a_{p\alpha}^\dagger a_{p\beta} - a_{p\beta}^\dagger a_{p\alpha} \right), \quad (170)$$

which may be combined with (168) to give the operator for the total spin, $\hat{S}^2 = \hat{S}_x^2 + \hat{S}_y^2 + \hat{S}_z^2$. As it contains products of one-electron operators, it is a two-electron operator and therefore somewhat tedious to manipulate. However, the explicit form is seldom needed, because we are able to employ the standard operator identities: In contrast to the orbital basis, the spin basis is *complete*, hence the usual first-quantization commutation relations hold also for the second-quantization spin operators. One such example is

$$[\hat{S}_+, \hat{S}_-] = \left[\sum_p a_{p\alpha}^\dagger a_{p\beta}, \sum_q a_{q\beta}^\dagger a_{q\alpha} \right] = \sum_p \left(a_{p\alpha}^\dagger a_{p\alpha} - a_{p\beta}^\dagger a_{p\beta} \right) = 2\hat{S}_z. \quad (171)$$

For example, the fine-structure and hyperfine-structure operators of first quantization affect both the spatial and spin parts of the wave function. An example would be the spin-orbit interaction operator

$$V_{\text{SO}}^c = \sum_{i=1}^N \xi(r_i) \mathbf{l}_i^c \cdot \mathbf{S}_i^c.$$

The second-quantization representation is obtained in the usual manner as

$$\begin{aligned} \hat{V}_{\text{SO}} &= \sum_{p\sigma q\tau} \int \phi_p^*(\mathbf{r}) \sigma^*(m_s) V_{\text{SO}}^c(\mathbf{r}, m_s) \phi_q(\mathbf{r}) \tau(m_s) d\mathbf{r} dm_s a_{p\sigma}^\dagger a_{q\tau} \\ &= \sum_{pq} \left(V_{pq}^x \hat{T}_{pq}^x + V_{pq}^y \hat{T}_{pq}^y + V_{pq}^z \hat{T}_{pq}^z \right), \end{aligned} \quad (172)$$

where

$$V_{pq}^{\mu} = \int \phi_p^*(\mathbf{r}) \xi(r) \ell_{\mu}^c \phi_q(\mathbf{r}) d\mathbf{r}$$

and where the Cartesian components of the *triplet excitation operators* are given by

$$\hat{T}_{pq}^x = \frac{1}{2} \left(a_{p\alpha}^{\dagger} a_{q\beta} + a_{p\beta}^{\dagger} a_{q\alpha} \right) \quad (173)$$

$$\hat{T}_{pq}^y = \frac{1}{2i} \left(a_{p\alpha}^{\dagger} a_{q\beta} - a_{p\beta}^{\dagger} a_{q\alpha} \right) \quad (174)$$

$$\hat{T}_{pq}^z = \frac{1}{2} \left(a_{p\alpha}^{\dagger} a_{q\alpha} - a_{p\beta}^{\dagger} a_{q\beta} \right). \quad (175)$$

The second-quantization representation of

- spin-free operators depend on the orbitals but have the same amplitudes (=integrals) for alpha and beta spins
- pure spin operators are independent of the functional form of the orbitals
- mixed operators depend on both the spin of the electron and the functional form of the orbitals.

Spin properties of determinants

Slater determinants are in general *not* eigenfunctions of the non-relativistic Hamiltonian, but are instead the non-degenerate eigenfunctions of the spin-orbital occupation-number operators:

$$\hat{N}_{p\sigma} |\mathbf{k}\rangle = k_{p\sigma} |\mathbf{k}\rangle. \quad (176)$$

We also note that the spin-orbital ON operators commute with the spin-projection operator, i.e. $[\hat{S}_z, \hat{N}_{p\sigma}] = 0$, because the ON operators commute among themselves, and \hat{S}_z is a linear combination of spin-orbital ON operators. Added by the observation that there is no degeneracies in (176), the Slater determinants are eigenfunctions of the *projected* spin:

$$\hat{S}_z |\mathbf{k}\rangle = M |\mathbf{k}\rangle. \quad (177)$$

By contrast, the Slater determinants are *not* eigenfunctions of the *total* spin. However, it is possible to determine spin eigenfunctions as simple linear combinations of determinants. A hint for the procedure is given by an observation that both the total and projected-spin operators commute with the *sum* of the ON operators for alpha and beta spins:

$$\begin{aligned} [\hat{S}_z, \hat{N}_{p\sigma} + \hat{N}_{p\beta}] &= 0 \\ [\hat{S}^2, \hat{N}_{p\sigma} + \hat{N}_{p\beta}] &= 0, \end{aligned}$$

because $\hat{N}_{p\sigma} + \hat{N}_{p\beta}$ are singlet operators. Such spin-adapted functions are known as *configuration state functions*.

3 STANDARD MODELS

All approximations in the solution of the Schrödinger equation should be unambiguous and precisely defined. Ideally, they should be improvable in a systematic fashion; and to yield more and more elaborate solutions that approach the exact solution. Thus we speak of different *models*, *levels of theory* and *hierarchies of approximations*.

Computational electronic-structure theory has a few *standard models* for the construction of approximate electronic wave functions. At the simplest level, the wave function is represented by a single Slater determinant (the Hartree–Fock approximation); and at the most complex level the wave function is represented as a variationally determined superposition of all determinants in the N -electron Fock space (full configuration-interaction, FCI). Between these extremes, there are a vast number of intermediate models with variable cost and accuracy.

It should be noted that none of the models presented here is applicable to all systems. However, each model is applicable in a broad range of molecular systems, providing solutions of known quality and flexibility at a given computational cost.

3.1 The Hartree–Fock approximation

In the *Hartree–Fock approximation* the electronic wave function is approximated by a single Slater determinant, and the energy is optimized with respect to variations of these spin-orbitals. Thus, the wave function may be written in a form

$$|\kappa\rangle = \exp(-\hat{\kappa}) |0\rangle, \quad (178)$$

where $|0\rangle$ is some reference configuration and $\exp(-\hat{\kappa})$ an operator that carries out unitary transformations among the spin orbitals; these orbital-rotation parameters for the Hartree–Fock ground state are obtained by *minimizing the energy*:

$$E_{\text{HF}} = \min_{\kappa} \langle \kappa | \hat{H} | \kappa \rangle. \quad (179)$$

The optimization could be carried out using standard numerical analysis; however, for the most of purposes more “specific” methods are required in order to cope with the computational cost. We recall that the Slater determinant represents a situation where electrons behave as independent particles, but are subject to the Fermi correlation. Therefore, the optimal determinant can be found by solving a set of effective one-electron Schrödinger equations – the *Hartree–Fock equations* – for the spin-orbitals, with the associated Hamiltonian being the *Fock operator*

$$\hat{f} = \sum_{PQ} f_{PQ} a_P^\dagger a_Q, \quad (180)$$

where the elements f_{PQ} are called the *Fock matrix*. A very essential feature of the Fock operator is that the one-electron part of the “true” molecular Hamiltonian is retained, but the two-electron part is replaced with the effective *Fock operator*, so that $\hat{f} = \hat{h} + \hat{V}$. The Fock operator incorporates *in an average sense* the Coulomb interaction among the electrons, corrected for Fermi correlation:

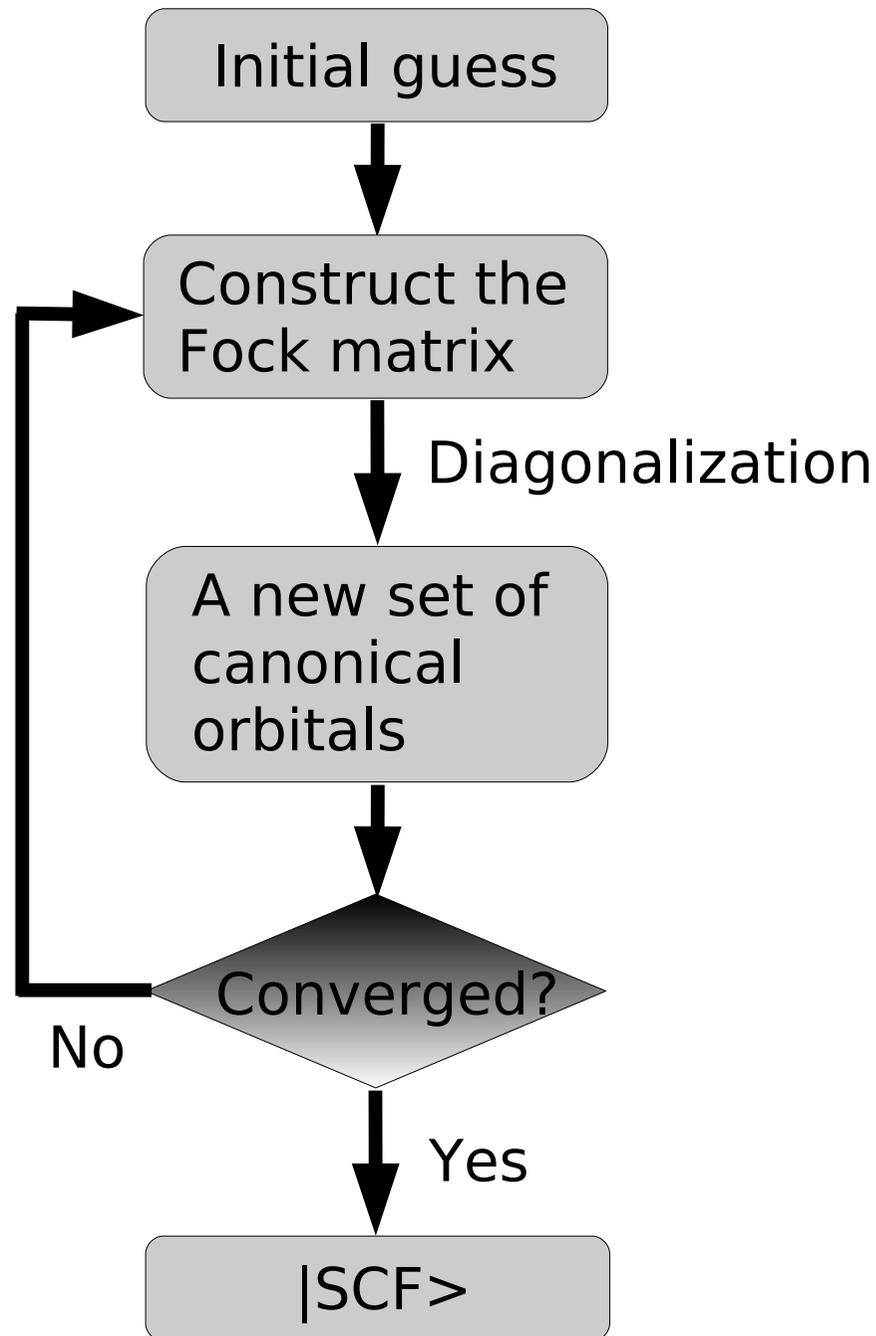
$$\hat{V} = \sum_{PQ} V_{PQ} a_P^\dagger a_Q \quad (181)$$

$$V_{PQ} = \sum_I (g_{PQII} - g_{PIIQ}), \quad (182)$$

where the integrals g are given in Eq. (142), with I running over the occupied and P and Q all spin-orbitals. The first term in V_{PQ} describes the classical Coulomb interaction of the electron with the charge distribution caused by the others, and the second one, *exchange term*, is a correction that arises from the antisymmetry of the wave function.

- The Hartree–Fock equations are solved by diagonalizing the Fock matrix. The resulting eigenvectors are called the *canonical spin-orbitals* and the the eigenvalues of the Fock matrix are the *orbital energies*, $f_{PQ} = \delta_{PQ} \epsilon_P$.

- Since the Fock matrix is defined in terms of its own eigenvectors, the canonical spin-orbitals and the orbital energies can only be obtained using an *iterative procedure*; where the Fock matrix is constructed and diagonalized in every iteration until the spin-orbitals obtained in a particular iteration become identical to those from which the Fock matrix was constructed. This procedure is referred to as the *self-consistent field* (SCF) method and the resulting wave function the *SCF wave function*.
- In the *canonical representation*, i.e., in the frame where the Fock matrix is diagonal, the electrons occupy the spin-orbitals in accordance with the Pauli principle and move independently one another in the electrostatic field caused by the stationary nuclei and by the charge distribution of the $N-1$ electrons. The orbital energies are the one-electron energies associated with the independent electrons and may thus be interpreted as the *energies required to remove a single electron*, that is, to ionize the system. The identification of the orbital energies with the negative ionization potentials is known as the *Koopman's theorem*.



The Hartree–Fock state is invariant to unitary transformations among the occupied spin-orbitals. Therefore, the spin-orbitals of the Hartree–Fock state are not uniquely determined by the condition (179) and the canonical orbitals are just one possible choice of spin-orbitals for the optimized N -particle state. Any set of energy-optimized orbitals decomposes the Fock matrix into two non-interacting blocks: one for the occupied, and another for the unoccupied spin-orbitals. When these subblocks are diagonalized, the canonical spin-orbitals are obtained.

The final electronic state is obtained as an antisymmetrized product of the canonical spin orbitals

$$|\text{HF}\rangle = \left(\prod_i a_{i\alpha}^\dagger a_{i\beta}^\dagger \right) |\text{vac}\rangle. \quad (183)$$

Finally, let us make a remark that the Hartree–Fock state is an eigenfunction of the Fock operator with an eigenvalue equal to the sum of the orbital energies of the canonical spin-orbitals

$$\hat{f} |\text{HF}\rangle = 2 \sum_i \epsilon_i |\text{HF}\rangle. \quad (184)$$

Restricted and unrestricted Hartree–Fock theory

The exact wave function is an eigenfunction of the total and projected spin operators. The Hartree–Fock wave function, which is not an eigenfunction of the Hamiltonian (but of the effective one, the Fock operator), does not possess these symmetries, they must be *imposed* on the Hartree–Fock solution.

- In the *restricted Hartree–Fock* (RHF) approximation, the energy is optimized subject to the condition that the wave function is an eigenfunction of the total and projected spins. In a practical implementation of the RHF scheme, it is advantageous to write the wave function as a configuration state function rather than a Slater determinant.

- In the *unrestricted Hartree–Fock* (UHF) approximation, the wave function is not required to be a spin eigenfunction, and different spatial orbitals are used for different spins.
- In some systems, usually close to the equilibrium geometry, the symmetries of the exact state are present in the UHF state as well. In such cases the UHF and RHF states coincide.

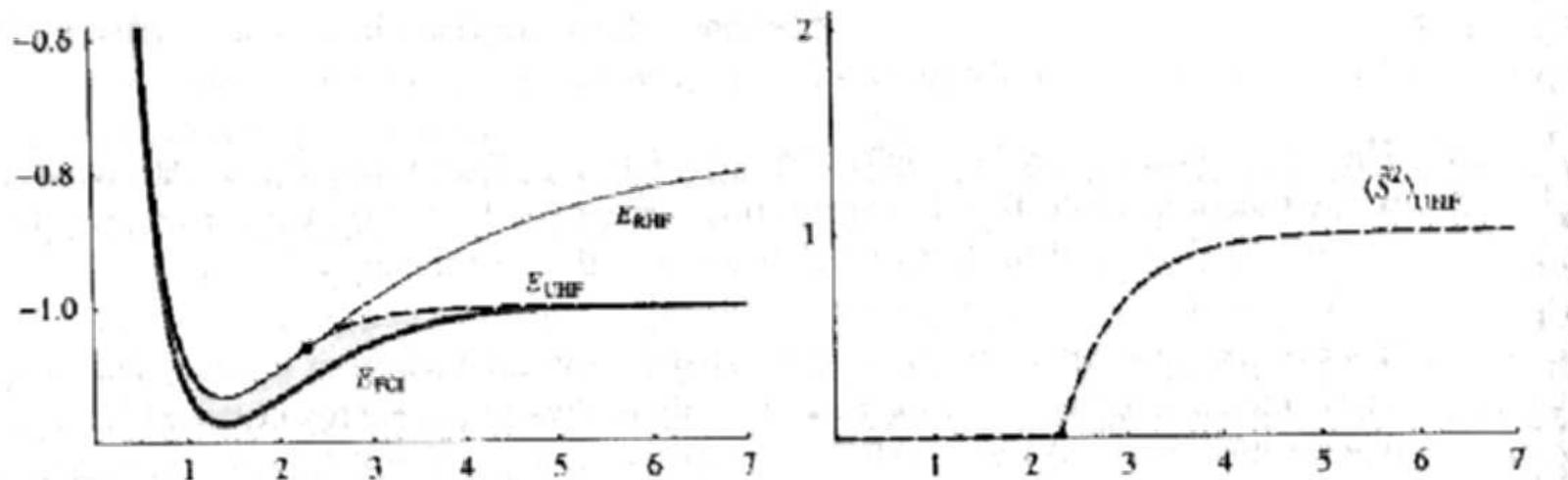
Example 3.1

For the hydrogen molecule,

$$|\text{RHF}\rangle = a_{1\sigma_g\alpha}^\dagger a_{1\sigma_g\beta}^\dagger |\text{vac}\rangle$$

$$|\text{UHF}\rangle = a_{\phi_{1\alpha}}^\dagger a_{\phi_{2\beta}}^\dagger |\text{vac}\rangle.$$

Let us plot these wave functions as well as the FCI wave function for comparison using the LCAO approximation for the orbital basis (cc-pVQZ basis set):



At short internuclear distances the RHF and UHF wave functions are identical, providing a crude but reasonable representation of the potential energy surface. At larger separations, the RHF wave function behaves poorly for energy, whereas

the UHF wave function falls towards the correct FCI curve. On the other hand, while the RHF wave function remains a true singlet at all distances, the UHF wave function behaves incorrectly with respect to the spin, yielding a spin intermediate of a singlet and a triplet.

Roothaan–Hall equations

We will now consider the classical Roothaan–Hall formulation of the Hartree–Fock theory, in which the MOs are expanded in a set of AOs, $\phi_p = \sum_{\mu} \chi_{\mu} C_{\mu p}$, whose expansion coefficients C are used as the variational parameters.

The Hartree–Fock energy is the expectation value of the true molecular electronic Hamiltonian (165), and is given in the orbital basis by

$$E^{(0)} = \langle \text{HF} | \hat{H} | \text{HF} \rangle = \sum_{pq} h_{pq} D_{pq} + \frac{1}{2} \sum_{pqrs} g_{pqrs} d_{pqrs} \quad (185)$$

where

$$D_{pq} = \langle \text{HF} | \hat{E}_{pq} | \text{HF} \rangle \quad (186)$$

$$d_{pqrs} = \langle \text{HF} | \hat{e}_{pqrs} | \text{HF} \rangle. \quad (187)$$

For a closed-shell state, the only nonzero elements of (186) and (187) are

$$D_{ij} = 2\delta_{ij}$$

$$d_{ijkl} = D_{ij}D_{kl} - \frac{1}{2}D_{il}D_{kj} = 4\delta_{ij}\delta_{kl} - 2\delta_{il}\delta_{kj},$$

and thus the total Hartree–Fock energy is

$$E^{(0)} = 2 \sum_i h_{ii} + \sum_{ij} (2g_{iijj} - g_{ijji}). \quad (188)$$

The energy must be optimized subject to the orthonormality of the MOs. We introduce the Hartree–Fock Lagrangian

$$L(\mathbf{C}) = E(\mathbf{C}) - 2 \sum_{ij} \lambda_{ij} (\langle \phi_i | \phi_j \rangle - \delta_{ij})$$

and the variational conditions are written in terms of it as

$$0 = \frac{\partial L(\mathbf{C})}{\partial C_{\mu k}} = 4h_{\mu k} + 4 \sum_j (2g_{\mu k j j} - g_{\mu j j k}) - 4 \sum_j S_{\mu j} \lambda_{j k},$$

which gives the condition for the optimized Hartree–Fock state

$$f_{\mu k} = \sum_j S_{\mu j} \lambda_{j k}. \quad (189)$$

Since the multiplier matrix λ is symmetric (for real orbitals), it can be diagonalized by an orthogonal transformation among the occupied orbitals, $\lambda = \mathbf{U} \epsilon \mathbf{U}^T$, the Hartree–Fock energy being invariant to such transformations. The off-diagonal Lagrange multipliers may thus be eliminated by an orthogonal transformation to a set of occupied MOs that satisfy the canonical conditions. In the canonical basis, the variational conditions become

$$\sum_{\nu} f_{\mu\nu}^{\text{AO}} C_{\nu k} = \epsilon_k \sum_{\nu} S_{\mu\nu} C_{\nu k}, \quad (190)$$

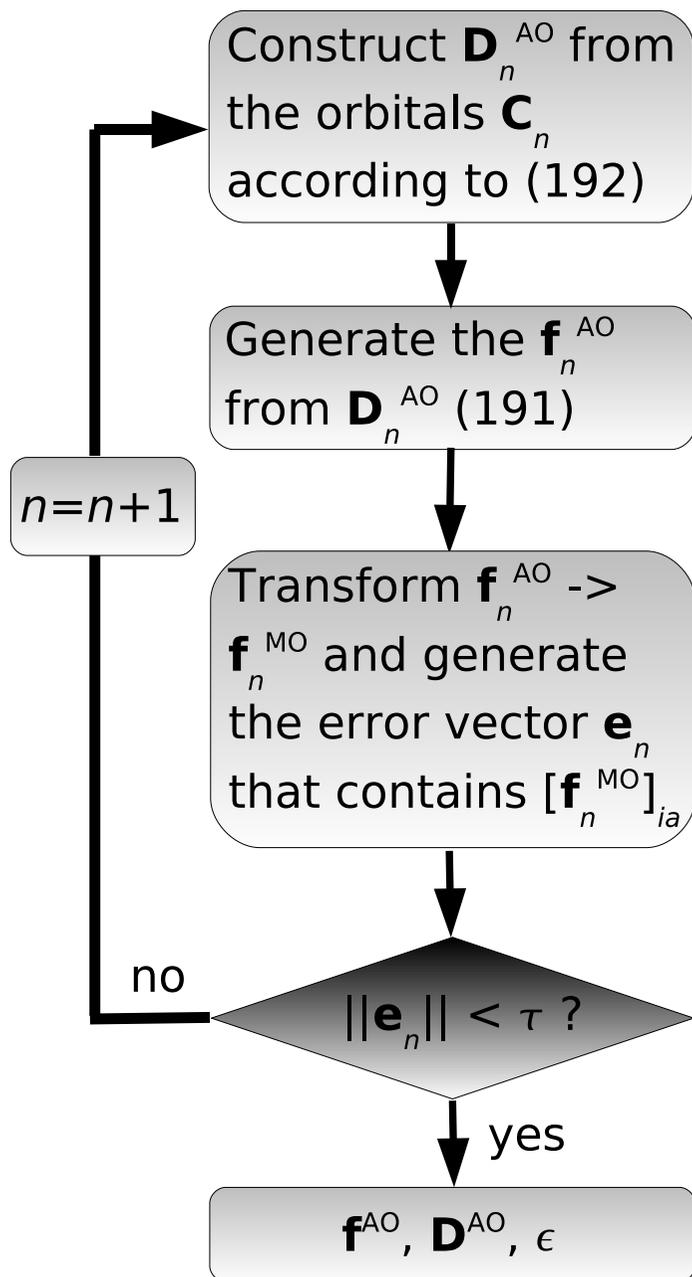
where the elements of the *AO Fock matrix* are

$$f_{\mu\nu}^{\text{AO}} = h_{\mu\nu} + \sum_i (2g_{\mu\nu ii} - g_{\mu i i \nu}). \quad (191)$$

Since \mathbf{f}^{AO} is symmetric, we can extend the variational conditions to yield a set of orthonormal virtual MOs that satisfy the same canonical conditions as the occupied MOs, i.e., $f_{ab} = \delta_{ab} \epsilon_a$. Now, we are able to write the Hartree–Fock variational conditions as

$$\mathbf{f}^{\text{AO}} \mathbf{C} = \mathbf{S} \mathbf{C} \epsilon, \quad (192)$$

where ϵ is a diagonal matrix that contains the orbital energies. This equation is called *the Roothaan-Hall equations*.



The AO Fock matrix can be evaluated entirely in the AO basis,

$$f_{\mu\nu}^{\text{AO}} = h_{\mu\nu} + \sum_{\rho\sigma} D_{\rho\sigma}^{\text{AO}} \left(g_{\mu\nu\rho\sigma} - \frac{1}{2} g_{\mu\sigma\rho\nu} \right), \quad (193)$$

where \mathbf{D}^{AO} is the AO representation of the one-electron density,

$$\mathbf{D}^{\text{AO}} = \mathbf{C}\mathbf{D}\mathbf{C}^T. \quad (194)$$

On the convergence of the SCF method

A straightforward implementation of the Roothaan–Hall SCF method (given in the figure) may fail to converge or may converge slowly. Several schemes to improve the SCF convergence have been introduced; of which the *direct inversion in the iterative subspace* (DIIS) method is amongst the simplest and one of the most successful. In the DIIS method, the information from the preceding iterations is used instead of generating the next density from the last AO Fock matrix:

$$\bar{\mathbf{f}}_n^{\text{AO}} = \sum_{i=1}^n w_i \mathbf{f}_i^{\text{AO}}, \quad (195)$$

where the DIIS weights w are obtained by minimizing the norm of the averaged error vector $\bar{\mathbf{e}}_n = \sum_{i=1}^n w_i \mathbf{e}_i$,

$$\|\mathbf{e}\|^2 = \sum_{i=1}^n \sum_{j=1}^n w_i \underbrace{\langle \mathbf{e}_i | \mathbf{e}_j \rangle}_{B_{ij}} w_j,$$

subject to the constraint that $\sum_{i=1}^n w_i = 1$. For this purpose, we construct the Lagrangian

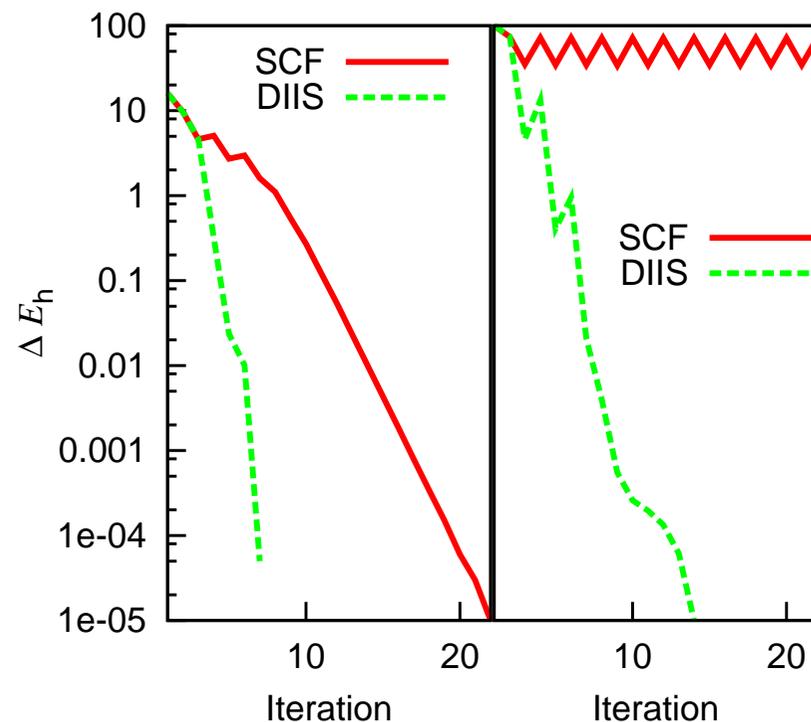
$$L = \sum_{i,j=1}^n w_i B_{ij} w_j - 2\lambda \left(\sum_{i=1}^n w_i - 1 \right),$$

the minimization of which leads to the set of linear equations:

$$\begin{pmatrix} B_{11} & B_{12} & \cdots & B_{1n} & -1 \\ B_{21} & B_{22} & \cdots & B_{2n} & -1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ B_{n1} & B_{n2} & \cdots & B_{nn} & -1 \\ -1 & -1 & \cdots & -1 & 0 \end{pmatrix} \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \\ \lambda \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ -1 \end{pmatrix}$$

from which the weights are solved.

- The C2-DIIS method is similar to the DIIS, but the constraint is $\sum_{i=1}^n w_i^2 = 1$. This leads to a solution of an eigenvalue problem instead of the set of linear equations as in DIIS; the scheme may handle singularities better.



On the left: H₂O in its equilibrium geometry, on the right: H₂O with its bonds stretched to double (angle remaining constant). Calculations are carried out in the cc-pVQZ basis.

- In the EDIIS (“energy DIIS”) scheme, the error-vector-norm minimization is replaced by a minimization of an approximation to the true energy function. It improves the convergence in cases where the start guess has a Hessian structure far from the optimal one. It also can converge in cases where DIIS diverges. For non-problematic cases EDIIS has actually a slower convergence rate than DIIS.
- These are based on either the diagonalization of the Fock matrix or on a direct optimization of the AO density matrix. Nevertheless, in some cases, even the DIIS method (and alike) is difficult to converge, when it might be better to use the second-order Newton method to the optimization of the Hartree–Fock energy. This has been applied in the literature to the optimization of the AO density matrix (the basis for the *trust-region* or *restricted-step method*) as well as in a method that carries out rotations among the MOs.
- The “trust-region SCF” (TRSCF) is the state-of-the-art model in SCF convergence. It is basically a globally convergent black-box method, where each SCF iteration consists of a trust-region density subspace minimization and a trust-region Roothaan–Hall step.

Integral-direct SCF

In every Roothaan–Hall SCF iteration a large number of two-electron integrals is needed to construct the Fock matrix. Since the same AO integrals are used in each iteration, these integrals can be written to disk and read in when needed.

In large-scale calculations, this approach becomes impractical due to the large number of needed integrals. In such cases, the AO integrals are instead recomputed in every iteration – as soon as a batch of two-electron integrals has been produced, it is contracted with the appropriate densities. This procedure is known as the *direct SCF* method.

If the value of an integral is very small, or if it is combined with small density-matrix elements, its contribution to the Fock matrix may be neglected. Therefore, by pre-estimating the integrals by having some kind of upper limit for them, the number of calculated integrals in the direct SCF scheme can be reduced dramatically. The computational cost of the SCF method scales formally as $\mathcal{O}(n^4)$, where n is the number of basis set functions, but it can be shown that only $\mathcal{O}(n^2)$ of the two-electron integrals are significant. Therefore, the direct SCF scheme combined with the pre-screening of the integrals is much cheaper for large systems. Furthermore, by using so-called *fast multipole methods*, only $\mathcal{O}(n)$ two-electron integrals is needed, while the rest can be either neglected or evaluated classically. The bottleneck is then not anymore in the Fock-matrix construction, but in the diagonalization step (which is an $\mathcal{O}(n^3)$ procedure).

To overcome the diagonalization bottleneck and to achieve a SCF method that would scale *linearly* with respect of the size of the system, and thus be applicable to large molecular systems such as biomolecules, we need to optimize directly the AO density matrix. The density-based SCF will be addressed later.

3.2 Configuration-interaction theory

The configuration-interaction (CI) wave function is constructed as a linear combination of Slater determinants (or configuration-state functions):

$$|\mathbf{C}\rangle = \sum_i C_i |i\rangle \quad (196)$$

with the coefficients determined by a variational optimization of the expectation value of the electronic energy, i.e., by minimization for the ground state

$$E_{\text{CI}} = \min_{\mathbf{C}} \frac{\langle \mathbf{C} | \hat{H} | \mathbf{C} \rangle}{\langle \mathbf{C} | \mathbf{C} \rangle}. \quad (197)$$

As discussed earlier, this is equivalent to the solution of an eigenvalue equation $\mathbf{H}\mathbf{C} = E_{\text{CI}}\mathbf{C}$. The construction of the CI wave function may thus be carried out by diagonalizing the Hamiltonian matrix $H_{ij} = \langle i | \hat{H} | j \rangle$, but more often iterative techniques are used.

In particular, we have to construct the CI wave function in a finite one-electron basis. The *full CI* (FCI) wave function consists of all the configurations that can be generated in a given basis, and it is most convenient to think it as generated from a single reference configuration, which dominates the wave function and usually is the Hartree–Fock state, by the application of a linear combination of spin-orbital excitation operators

$$|\text{FCI}\rangle = \left(1 + \sum_{AI} \hat{X}_I^A + \frac{1}{2} \sum_{ABIJ} \hat{X}_{IJ}^{AB} + \dots \right) |\text{HF}\rangle, \quad (198)$$

where for example

$$\begin{aligned} \hat{X}_I^A |\text{HF}\rangle &= C_I^A a_A^\dagger a_I |\text{HF}\rangle \\ \hat{X}_{IJ}^{AB} |\text{HF}\rangle &= C_{IJ}^{AB} a_A^\dagger a_B^\dagger a_I a_J |\text{HF}\rangle. \end{aligned}$$

Thus, we may characterize the determinants in the FCI expansion as single (S), double (D), triple (T), quadruple (Q), quintuple (5), and so forth excitations relative to the Hartree–Fock state.

The number of determinants in an FCI wave function is

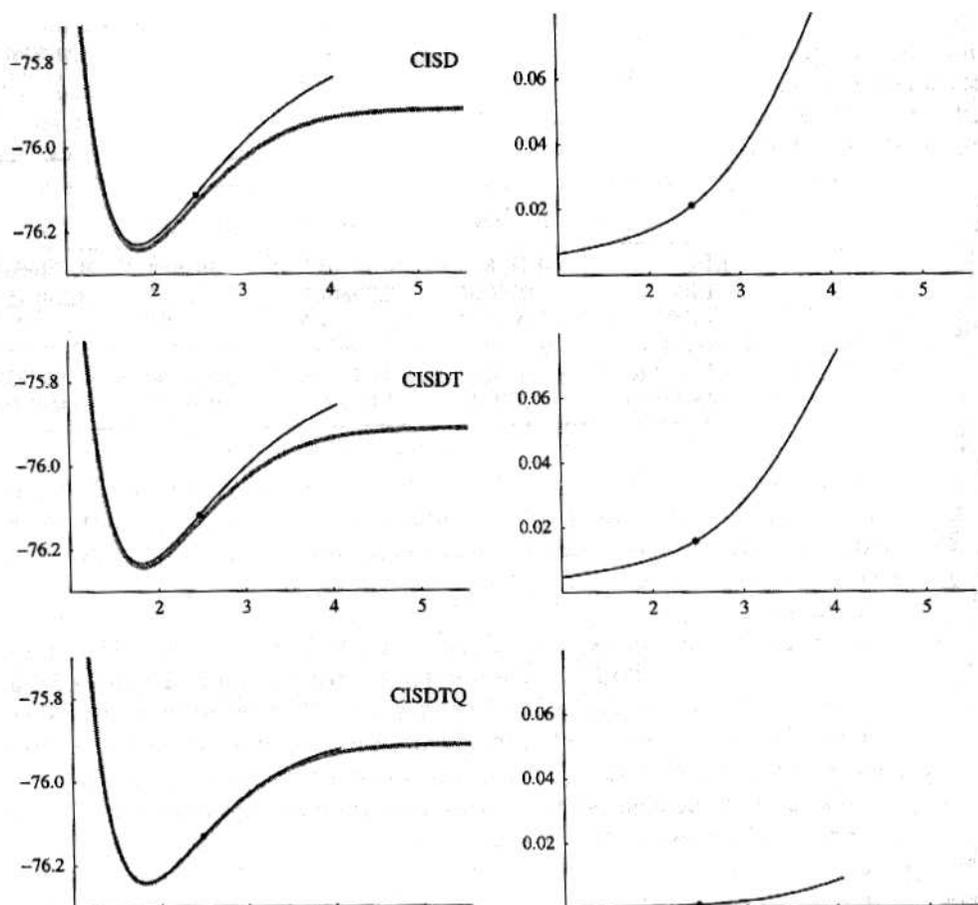
$$N_{\text{det}} = \binom{n}{k}^2 = \left(\frac{n!}{k!(n-k)!} \right)^2, \quad (199)$$

for a system with n orbitals containing k alpha and k beta electrons. For example, when $n = 2k$, $N_{\text{det}} \approx 16^k/k\pi$ – for large k , the number of determinants increases by a factor of 16 for each new pair of electrons and orbitals!

Single-reference CI wave functions

As the FCI wave function is obtainable only for the smallest systems, it becomes necessary to *truncate* the expansion, so that only a subset of the determinants are included, based on the excitation level. This is justified, since in general the lower-order excitations are more important than those of high orders. In principle, this procedure of *hierarchical truncation* may be continued until the FCI wave function is recovered. Since the CI model is variational, the FCI energy is approached from the above.

The electronic energies of truncated CI wave functions for the water molecule (cc-pVDZ basis) relative to the FCI energy:

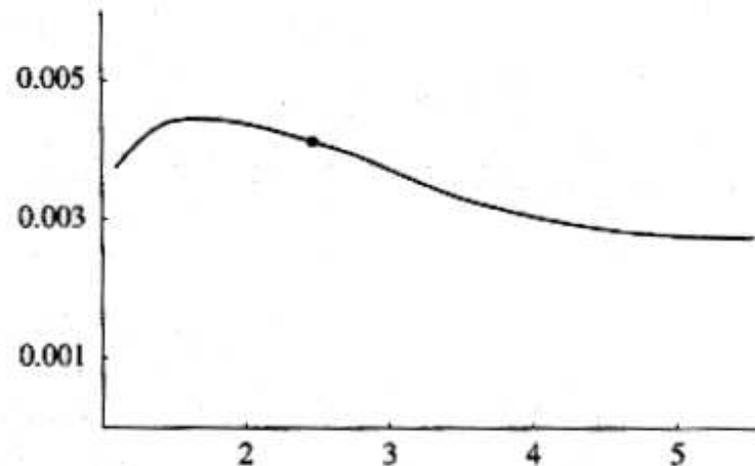
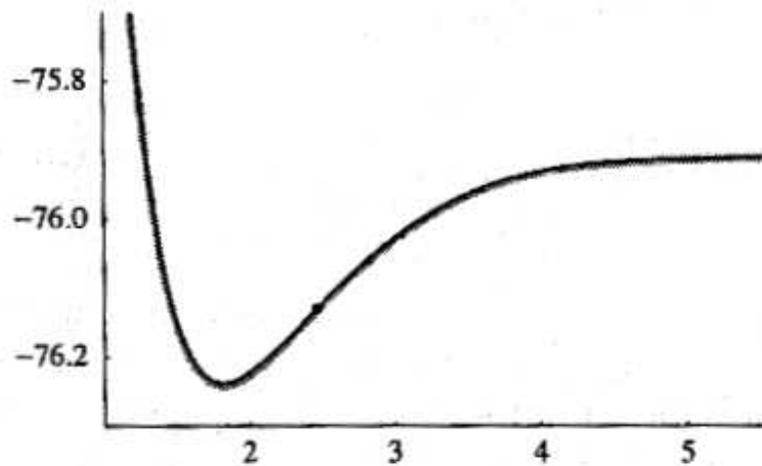


	$R = R_{\text{ref}}$		$R = 2R_{\text{ref}}$	
	$E - E_{\text{FCI}}$	Weight	$E - E_{\text{FCI}}$	Weight
RHF	0.2178	0.9410	0.3639	0.5896
CISD	0.0120	0.9980	0.0720	0.9487
CISDT	0.0090	0.9985	0.0560	0.9590
CISDTQ	0.0003	1.0000	0.0058	0.9988
CISDTQ5	0.0001	1.0000	0.0022	0.9995

Whereas the contributions drop monotonically, the step from an even-order excitation level to an odd-order is far more pronounced than the step from an odd to an even-order excitation level; hence the CI expansion is preferably truncated at even orders. For example, the first useful truncated CI wave function – the *CI singles-and-doubles wave function* (CISD) – recovers 94.5% of the correlation energy, the inclusion of triples on top of that improves the treatment to 95.9%, while as much as 99.9% of the correlation energy is recovered at the CISDTQ level.

Multi-reference CI wave functions

In the stretched geometry $2R_{\text{ref}}$, the Hartree–Fock determinant is less dominant. Clearly, the strategy of defining a correlation hierarchy in terms of excitations out of a *single reference determinant* such as the Hartree–Fock determinant does not provide the best possible description of the dissociation process. To overcome this, we may introduce CI wave functions based on the idea of a *reference space* comprising more than a single determinant: The *multi-reference CI* (MRCI) wave functions is generated by including in the wave function all configurations belonging to this reference space as well as all excitations up to a given level from each reference configuration, yielding, e.g., the multi-reference singles-and-doubles CI (MRSDCI) wave function.



- MRSDCI model is very accurate, if all the important configurations are included in the reference space.
- The number of important configurations becomes, however, very large, making thus the model applicable to quite small systems.
- The choice of the active space requires a lot of chemical intuition and patience.

Optimization of the CI wave function

For large CI expansions it is impossible to set up and diagonalize the Hamiltonian matrix of the eigenvalue equation. However, usually only a few of the lowest eigenvalues are of interest, and we may determine them by *iterative methods*, where the eigenvectors and eigenvalues are generated by a sequence of linear transformations of the form $\sigma = \mathbf{HC}$, where \mathbf{C} is some trial vector. These methods can be devised for both of the CI eigenvalue problem and the optimization of the CI energy (197). The most often used strategies rely on the second-order Newton's method and its approximations.

On the disadvantages of the CI approach

CI wave functions truncated at a given excitation level relative to a reference configuration do not provide size-extensive energies.

- For example, if two fragments are described at the CISD level, then a size-extensive treatment of the compound system requires that the wave function is augmented with certain triples and quadruples; more precisely those that represent products of single and double excitations in the two subsystems.
- An approach called *quadratic CI* (QCI) is a size-extensive revision of the CI model. Other widely used approach to overcome the lack of size-extensivity is the *Davidson correction* applied on top of the truncated CI energy.

The CI description of the electronic system is not at all compact. Even though higher excitations are less important than those of lower orders, their number is enormous. Therefore, the CI wave function converges slowly with the number of variational parameters.

Both of these problems are due to the linear parameterization in the CI model (196).

3.3 Multi-configurational SCF theory

The multi-configurational self-consistent field (MCSCF) theory is a generalization of the Hartree–Fock wave function to systems dominated by more than one electronic configuration. This method is particularly useful in the description of bond breakings and molecular dissociation processes.

In MCSCF theory, the wave function is written as a linear combination of determinants, whose expansion coefficients C_i are optimized *simultaneously* with the MOs according to the variation principle:

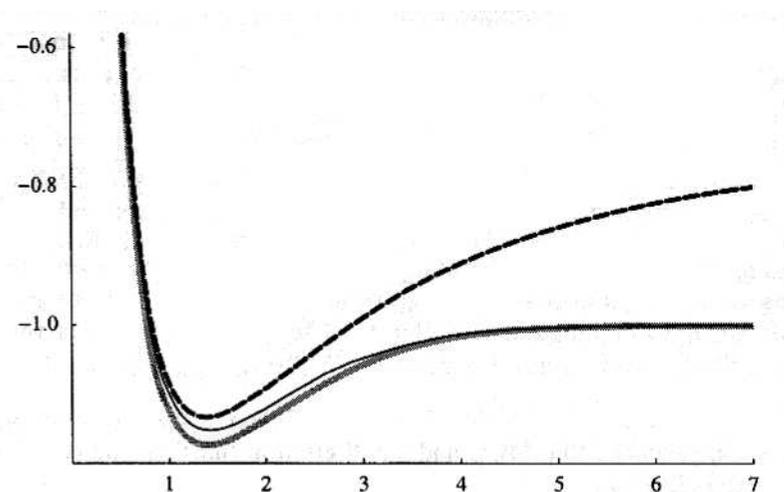
$$|\kappa, \mathbf{C}\rangle = \exp(-\hat{\kappa}) \sum_i C_i |i\rangle. \quad (200)$$

The ground-state MCSCF wave function is obtained by minimizing the energy with respect to the variational parameters,

$$E_{\text{MCSCF}} = \min_{\kappa, \mathbf{C}} \frac{\langle \kappa, \mathbf{C} | \hat{H} | \kappa, \mathbf{C} \rangle}{\langle \kappa, \mathbf{C} | \kappa, \mathbf{C} \rangle}. \quad (201)$$

Example 3.2

A two-configuration MCSCF wave function of the hydrogen molecule is $|\text{MC}\rangle = C_1 a_{1\sigma_g\alpha}^\dagger a_{1\sigma_g\beta}^\dagger |\text{vac}\rangle + C_2 a_{1\sigma_u\alpha}^\dagger a_{1\sigma_u\beta}^\dagger |\text{vac}\rangle$ where the variation principle is invoked to optimize the configuration coefficients as well as the orbitals. The solid line is the potential energy curve in the cc-pVQZ basis set. The thick and the dashed lines are for FCI and RHF, respectively.



For any larger system than the H_2 molecule the selection of configuration space is the greatest difficulty in the MCSCF calculation. It is very often impossible to generate a sufficient MCSCF configuration, which is computationally tractable. The selection is usually carried out by dividing the MO space in *inactive*, *active*, and *secondary* orbitals.

- The inactive orbitals are set to be doubly occupied.
- The active orbitals have variable occupancies of 0, 1, or 2. The MCSCF expansion is then obtained by distributing the *active electrons* in all possible ways among the active orbitals. In the final optimized state, the active orbitals have non-integer occupation numbers between 0 and 2.
- The secondary orbitals are unoccupied throughout the optimization.

This scheme is called the *complete active space SCF method* (CASSCF). We note that the CASSCF wave function is identical to the FCI wave function, when all the orbitals of the system are active; and reduces to the Hartree–Fock wave function when the active space is empty.

In *restricted active space SCF* (RASSCF) calculations, the active orbital space is further divided into three subspaces: RAS1, with an upper limit on the allowed number of holes in each configuration; RAS2, with no restrictions enforced; and RAS3, with an upper limit on the allowed number of electrons in each configuration. This allows for larger active spaces than CASSCF.

In practise, the simultaneous optimization of orbitals and CI coefficients is a difficult nonlinear optimization problem, which restricts the MCSCF expansions to be significantly smaller than those encountered in the CI wave functions. By itself, the MCSCF model is not suited to the treatment of dynamical correlation, that requires large basis sets and long configuration expansions.

Example 3.3

One popular approach for the choice of the MCSCF active spaces is based on the natural occupation numbers obtained from a preceding second-order Møller–Plesset perturbation theory calculation. For the water molecule in the cc-pVTZ basis in the order (A_1 , B_1 , B_2 , A_2) we obtain

Natural orbital occupation numbers, symmetry 1

1.99953982	1.98393843	1.96476967	0.02261760	0.01156239
0.00648257	0.00552982	0.00113424	0.00083070	0.00060446
0.00047604	0.00036853	0.00035173	0.00019527	0.00013628
0.00008516	0.00007619	0.00005837	0.00004768	0.00003078
0.00002067	0.00001341	0.00000675		

Natural orbital occupation numbers, symmetry 2

1.96724668	0.01959000	0.00552930	0.00068405	0.00064037
0.00042732	0.00035178	0.00013127	0.00005653	0.00004947
0.00001646				

Natural orbital occupation numbers, symmetry 3

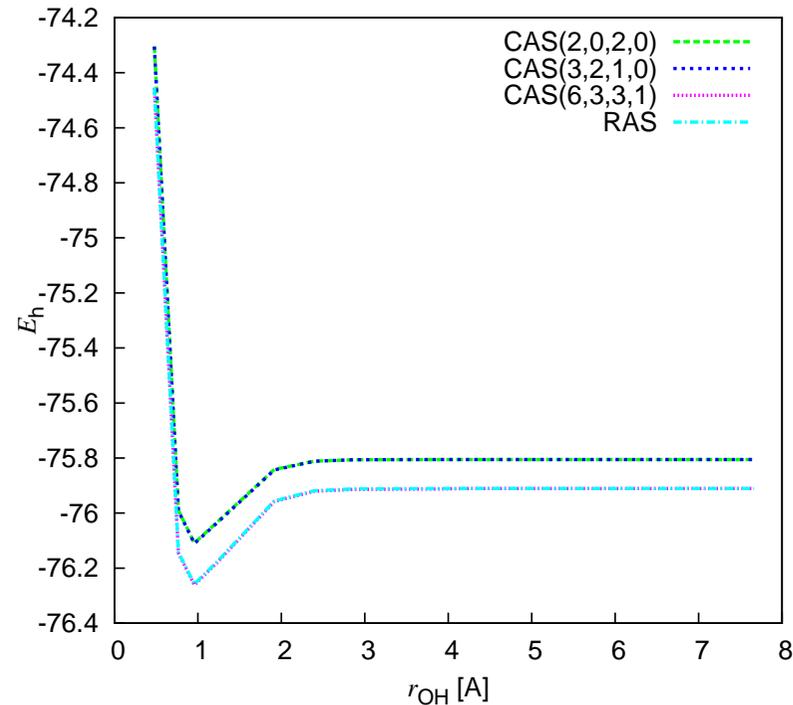
1.96415036	0.02452309	0.00641995	0.00104230	0.00086705
0.00065869	0.00039618	0.00019746	0.00013435	0.00011930
0.00006217	0.00005180	0.00003331	0.00001667	0.00001255
0.00000755	0.00000429			

Natural orbital occupation numbers, symmetry 4

0.00626166	0.00079588	0.00038289	0.00012271	0.00006547
0.00006019	0.00001431			

We can choose active spaces of different flexibility:

- Smallest possible active space seems to be (2,0,2,0), i.e. four electrons in four MOs $3a_1$, $4a_1$, $1b_2$, and $2b_2$. This corresponds to a combination of 12 determinants.
- For more accurate description, eight electrons distributed on (3,1,2,0) active space, corresponding to 37 configurations, gives almost similar potential curve.
- For further improvement, an eight-electron (6,3,3,1) active space is of high quality and sufficient for incorporating dynamical correlation effects as well. However, we get almost the same curve by supplementing the CAS(3,1,2,0) wave function by RAS3 space (3,2,1,0) limiting the number of electrons in RAS3 space to be 2 at maximum. This reduces the number of determinants from 36240 to 2708.



Dissociation of the C_{2v} water molecule. MCSCF calculations for the potential energy curve in the cc-pVTZ basis as a function of the OH distance.

3.4 Coupled-cluster theory

As noted earlier, the shortcomings of the CI hierarchy – the lack of size-extensivity and the slow convergence towards the FCI limit – are due to the linear expansion. However, we may recast the FCI expansion (198) in the form of a product wave function

$$|\text{CC}\rangle = \left[\prod_{\mu} (1 + t_{\mu} \hat{\tau}_{\mu}) \right] |\text{HF}\rangle, \quad (202)$$

where we have introduced a generic notation $\hat{\tau}_{\mu}$ for an excitation operator of unspecified excitation level (D, T, Q,...) and the associated *cluster amplitude* t_{μ} . This (full) *coupled-cluster (CC) wave function* differs from the FCI wave function by the presence of terms that are nonlinear in the excitation operators. Of course, the CC and FCI wave functions are *completely equivalent* provided that all the excitations are included.

Since the $\hat{\tau}$ s always excite from the set of occupied Hartree–Fock spin-orbitals to the virtual ones, $[\hat{\tau}_{\mu}, \hat{\tau}_{\nu}] = 0$. Due to the presence of the product excitations in the coupled-cluster state (202), each determinant can thus be reached in several different ways. For example, the determinant $|\mu\nu\rangle = \hat{\tau}_{\mu}\hat{\tau}_{\nu}|\text{HF}\rangle = \hat{\tau}_{\nu}\hat{\tau}_{\mu}|\text{HF}\rangle$ may be reached with an overall amplitude equal to the sum of the individual amplitude. With respect to this determinant, the double-excitation amplitude $t_{\mu\nu}$ is referred to as a *connected cluster amplitude* and subsequent singles excitation amplitudes $t_{\mu}t_{\nu}$ as a *disconnected cluster amplitude*.

By analogy with the CI theory, we could attempt to determine the CC state by minimizing the expectation value of the Hamiltonian with respect to the amplitudes. Due to the nonlinear parameterization of the CC model, the analogous variational conditions would give rise to an intractable set of nonlinear equations. Hence, in practice *coupled-cluster models are not solved using the variation principle*, and the CC energies are therefore non-variational.

Instead, by multiplying the eigenvalue equation $\hat{H} |CC\rangle = E |CC\rangle$ by $\langle\mu| = \langle HF | \hat{\tau}_\mu^\dagger$ we obtain the *projected coupled-cluster equations*

$$\langle\mu | \hat{H} | CC \rangle = E \langle\mu | CC \rangle, \quad (203)$$

where the CC energy is obtained by projection against the Hartree–Fock state (note that $\langle HF | CC \rangle = 1$)

$$E = \langle HF | \hat{H} | CC \rangle. \quad (204)$$

These equations are nonlinear in amplitudes as well, but unlike the variational conditions, this expansion will, due to the Slater’s rules, *terminate after few terms*, as we shall see later.

The coupled-cluster exponential ansatz

Because the excitation operators in (202) commute, i.e. $\hat{\tau}_\mu^2 = 0$, the correlating operators can be written in more convenient form, known as the *exponential ansatz for the coupled-cluster wave function*,

$$|CC\rangle = \exp(\hat{T}) |HF\rangle, \quad (205)$$

where we have introduced the *cluster operator*

$$\hat{T} = \sum_{\mu} t_{\mu} \hat{\tau}_{\mu}. \quad (206)$$

In CC theory, a *hierarchy of approximations* is established by partitioning the cluster operator in the form

$$\hat{T} = \hat{T}_1 + \hat{T}_2 + \hat{T}_3 + \cdots + \hat{T}_n, \quad (207)$$

where, for example,

$$\hat{T}_1 = \sum_{AI} t_I^A \hat{\tau}_I^A = \sum_{AI} t_I^A a_A^\dagger a_I \quad (208)$$

$$\hat{T}_2 = \sum_{A>B, I>J} t_{IJ}^{AB} \hat{\tau}_{IJ}^{AB} = \frac{1}{4} \sum_{AIBJ} t_{IJ}^{AB} a_A^\dagger a_I a_B^\dagger a_J. \quad (209)$$

Each excitation operator in \hat{T} excites at least one electron from an occupied Hartree–Fock spin-orbital to a virtual one. It should be noted that for a system with N electrons, the expansion terminates after N terms. We further note that $[\hat{T}_i, \hat{T}_j] = 0$ as a trivial consequence of the individual excitation operator commutation relations. Moreover,

$$\begin{aligned} \langle \text{HF} | \hat{\tau}_\mu &= 0 \\ \hat{\tau}_\mu^\dagger | \text{HF} \rangle &= 0, \end{aligned}$$

because it is impossible to excite an electron from an unoccupied orbital.

To compare the CC model with the CI model, we expand the exponential operator and collect terms to the same order in excitation level:

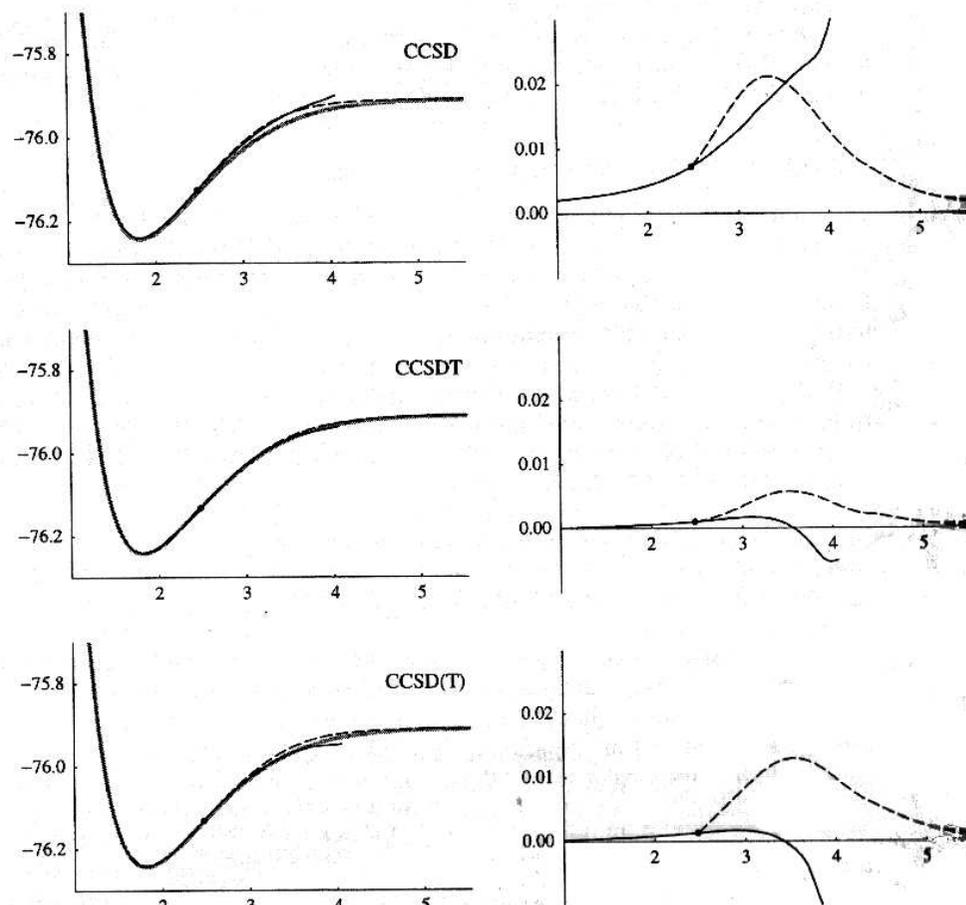
$$\exp(\hat{T}) | \text{HF} \rangle = \sum_{i=0}^N \hat{C}_i | \text{HF} \rangle.$$

The lowest-order operators \hat{C}_i are given by

$$\begin{aligned} \hat{C}_0 &= 1 \\ \hat{C}_1 &= \hat{T}_1 \\ \hat{C}_2 &= \hat{T}_2 + \frac{1}{2!} \hat{T}_1^2 \\ \hat{C}_3 &= \hat{T}_3 + \hat{T}_1 \hat{T}_2 + \frac{1}{3!} \hat{T}_1^3 \\ \hat{C}_4 &= \hat{T}_4 + \hat{T}_3 \hat{T}_1 + \frac{1}{2!} \hat{T}_1^2 \hat{T}_2 + \frac{1}{4!} \hat{T}_1^4, \end{aligned}$$

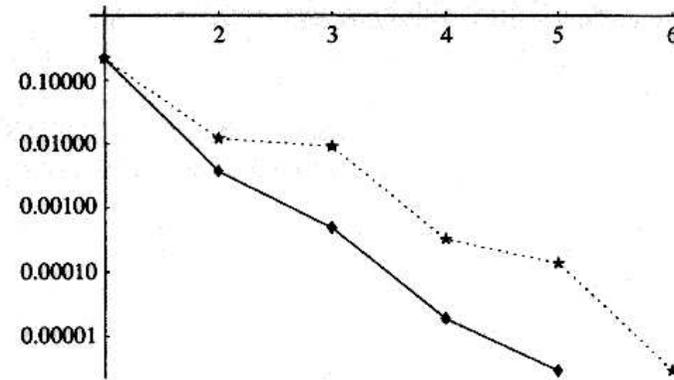
demonstrating the composition of the excitation processes. The advantages of the cluster parameterization arise in truncated wave functions and are related to the fact that even at the truncated level, the coupled-cluster state contains contributions from all determinants in the FCI wave function, with weights obtained from the different excitation processes.

Similarly with the CI theory, a hierarchy of CC models is introduced by truncating the cluster operator at different excitation levels. The simplest successful CC wave function is the *coupled-cluster singles-and-doubles* (CCSD) model. In this model, the \hat{T}_2 operator describes the electron-pair interactions and \hat{T}_1 carries out the corresponding orbital relaxations. Its computational cost scales as $\mathcal{O}(N^6)$. For higher accuracy, the connected triple excitations have to be taken into account, which leads to the *coupled-cluster singles-doubles-and-triples* (CCSDT) model. This model describes already the dynamical correlation almost perfectly, but is computationally very demanding, $\mathcal{O}(N^8)$ and is thus applicable to smallest molecules only. The triples excitations can be estimated in a perturbational way, giving rise to the *CCSD(T) method*, which is a good compromise between the computational cost (exhibiting an $\mathcal{O}(N^7)$ scaling) and accuracy. It is often considered to be the best single-reference treatment of molecular Schrödinger equation that is applicable in practise.



On the left: The CC dissociation curves of the C_{2v} water molecule in the cc-pVDZ basis. On the right: corresponding differences between the CC and FCI energies. RHF reference state is plotted the full line and UHF with the dashed.

Let us at this point compare the CI and CC hierarchies. Owing to the disconnected clusters, CC wave functions truncated at a given excitation level also contain contributions from determinants corresponding to higher-order excitations; whereas CI wave functions truncated at the same level contain contributions only from determinants up to this level. This is reflected in the convergence towards the FCI limit, which is typical of small systems. For larger systems, CI starts to behave very badly, while the CC description is *unaffected by the number of electrons*.



The error with respect to FCI of CC and CI wave functions for the water molecule at the equilibrium geometry in the cc-pVDZ basis.

According to the previous discussion, the projected coupled-cluster equations are written as

$$\langle \text{HF} | \hat{H} \exp(\hat{T}) | \text{HF} \rangle = E \quad (210)$$

$$\langle \mu | \hat{H} \exp(\hat{T}) | \text{HF} \rangle = E \langle \mu | \exp(\hat{T}) | \text{HF} \rangle. \quad (211)$$

- The excited projection manifold $\langle \mu |$ comprises the full set of all determinants up to the chosen truncation level.
- For the full CC wave function, the number of equations is equal to the number of determinants and the solution of the projected equations recovers the FCI wave function.
- The nonlinear equations (210) must be solved iteratively, substituting in every iteration the CC energy calculated from (211).

It is convenient to rewrite the projected CC equations using the non-Hermitian *similarity-transformed Hamiltonian* $\hat{H} = \exp(-\hat{T})\hat{H}\exp(\hat{T})$:

$$\langle \text{HF} | \exp(-\hat{T})\hat{H}\exp(\hat{T}) | \text{HF} \rangle = E \quad (212)$$

$$\langle \mu | \exp(-\hat{T})\hat{H}\exp(\hat{T}) | \text{HF} \rangle = 0 \quad (213)$$

The equations (211) and (213) are completely equivalent (yet the equivalence is nontrivial for all the truncated CC models!), meaning that they yield same amplitudes and energy upon solution. They are referred to as the *unlinked* and *linked* coupled-cluster equations, respectively. Their computational cost is practically the same, but the linked equations have some advantages: Although both yield the same, size-extensive wave function, the linked equations are size-extensive *term by term*, that simplifies some perturbational treatments; in addition, the CC approach excited states in mind is more fruitful to start from the linked equations. The main reason is however the following:

- When the Baker–Campbell–Hausdorff (BCH) expansion $\exp(\mathbf{A})\mathbf{B}\exp(-\mathbf{A}) = \mathbf{B} + [\mathbf{A}, \mathbf{B}] + \frac{1}{2!}[\mathbf{A}, [\mathbf{A}, \mathbf{B}]] + \frac{1}{3!}[\mathbf{A}, [\mathbf{A}, [\mathbf{A}, \mathbf{B}]]] + \dots$ is applied to the similarity-transformed Hamiltonian, one obtains

$$\exp(-\hat{T})\hat{H}\exp(\hat{T}) = \hat{H} + [\hat{H}, \hat{T}] + \frac{1}{2}[[\hat{H}, \hat{T}], \hat{T}] + \frac{1}{6}[[[[\hat{H}, \hat{T}], \hat{T}], \hat{T}], \hat{T}] + \frac{1}{24}[[[[[[\hat{H}, \hat{T}], \hat{T}], \hat{T}], \hat{T}], \hat{T}], \hat{T}], \quad (214)$$

the coupled-cluster equations are therefore *no higher than quartic in the cluster amplitudes* – for any truncation level or even for full CC expansion.

- Due to the Brillouin theorem, the one-particle operators contribute only to second order. Furthermore, cluster operators higher than the doubles do not contribute to the CC energy since \hat{H} is a two-particle operator. As a result, only singles and doubles amplitudes contribute directly to the CC energy – irrespective of the truncation level of the cluster operator. Of course, the higher-order excitations contribute indirectly, since all amplitudes are coupled by the projected equations.

On coupled-cluster optimization techniques

The elements of the left-hand side of the CC amplitude equation (213) constitute the *coupled-cluster vector function* $\Omega_\mu(\mathbf{t})$, which can be expanded around the set of amplitudes of the current iteration $\mathbf{t}^{(n)}$:

$$\Omega(\mathbf{t}^{(n)} + \Delta\mathbf{t}) = \Omega^{(0)}(\mathbf{t}^{(n)}) + \Omega^{(1)}(\mathbf{t}^{(n)})\Delta\mathbf{t} + \dots,$$

where $\Omega^{(0)}$ is the vector function calculated from the amplitudes $\mathbf{t}^{(n)}$ and $\Omega^{(1)}$ the CC *Jacobian*

$$\Omega_{\mu\nu}^{(1)} = \left\langle \mu \left| \exp(-\hat{T}^{(n)}) [\hat{H}, \hat{\tau}_\nu] \exp(\hat{T}^{(n)}) \right| \text{HF} \right\rangle. \quad (215)$$

Then an iterative scheme is established:

- Solve the set of linear equations

$$\Omega^{(1)}(\mathbf{t}^{(n)})\Delta\mathbf{t}^{(n)} = -\Omega^{(0)}(\mathbf{t}^{(n)}) \quad (216)$$

for $\Delta\mathbf{t}^{(n)}$

- Form the improved estimate

$$\mathbf{t}^{(n+1)} = \mathbf{t}^{(n)} + \Delta\mathbf{t}^{(n)} \quad (217)$$

- Iterate Equations (216) and (217) until convergence in the amplitudes.

In more sophisticated implementations, this Newton-method-like approach is replaced by a scheme, where the solution of the linear equations as well as the construction of the Jacobian are avoided by the approximation

$$\Delta t_\mu^{(n)} = -\varepsilon_\mu^{-1} \Omega_\mu^{(0)}(\mathbf{t}^{(n)}), \quad (218)$$

where, e.g. $\varepsilon_{ij}^{ab} = (\epsilon_a + \epsilon_b - \epsilon_i - \epsilon_j)$.

Usually, the convergence is accelerated by the DIIS method (see the discussion in section 3.1).

The closed-shell CCSD model

We consider now in detail the important special case of CC theory: the closed-shell CCSD model. We recall that the singlet CCSD state is generated by $|\text{CC}\rangle = \exp(\hat{T}_1 + \hat{T}_2)|\text{HF}\rangle$, hence only those terms in \hat{T}_i that transform as singlet operators should be retained in the cluster operator. It is easy to show that

$$[\hat{S}_\pm, \hat{T}_i] = 0 \quad (219)$$

$$[\hat{S}_z, \hat{T}_i] = 0. \quad (220)$$

These impose important constraints on the cluster amplitudes in the spin-orbital basis. Due to these, the singles cluster operator becomes

$$\hat{T}_1 = \sum_{ai} t_i^a \hat{E}_{ai} \quad (221)$$

and the doubles operator

$$\hat{T}_2 = \frac{1}{2} \sum_{aibj} t_{ij}^{ab} E_{ai} E_{bj}. \quad (222)$$

The cluster amplitudes are symmetric: $t_{ij}^{ab} = t_{ji}^{ba}$.

The CCSD energy is obtained from Eq. (212). By noting that $\langle \text{HF} | \exp(-\hat{T}) = \langle \text{HF} |$, expanding the cluster operators and after some algebra we obtain the form

$$E_{\text{CCSD}} = E_{\text{HF}} + \frac{1}{2} \langle \text{HF} | [[\hat{H}, \hat{T}_1], \hat{T}_1] | \text{HF} \rangle + \langle \text{HF} | [\hat{H}, \hat{T}_2] | \text{HF} \rangle, \quad (223)$$

where the Hartree–Fock energy is given by $E_{\text{HF}} = \langle \text{HF} | \hat{H} | \text{HF} \rangle$. By inserting the CCSD operators to the energy expression and calculating the commutators explicitly (see Exercises 6.1 and 6.3*) we obtain

$$\begin{aligned}
 E_{\text{CCSD}} &= E_{\text{HF}} + \frac{1}{2} \sum_{aibj} t_i^a t_j^b \langle \text{HF} | [[\hat{H}, \hat{E}_{ai}], \hat{E}_{bj}] | \text{HF} \rangle + \frac{1}{2} \sum_{aibj} t_{ij}^{ab} \langle \text{HF} | [\hat{H}, \hat{E}_{ai} \hat{E}_{bj}] | \text{HF} \rangle \\
 &= E_{\text{HF}} + \frac{1}{2} \sum_{aibj} (t_{ij}^{ab} + t_i^a t_j^b) \langle \text{HF} | [[\hat{H}, \hat{E}_{ai}], \hat{E}_{bj}] | \text{HF} \rangle \\
 &= E_{\text{HF}} + \sum_{aibj} (t_{ij}^{ab} + t_i^a t_j^b) L_{iajb}, \tag{224}
 \end{aligned}$$

where we have denoted $L_{pqrs} = 2g_{pqrs} - g_{psrq}$.

The equation-of-motion coupled-cluster method

In CC theory, we arrive at a good description of the ground state. The excited states could be examined by carrying out separate, independent calculation for each state of interest, using in each case some appropriate zero-order reference determinant. However, it is usually difficult or even impossible to determine these adequate reference determinants; it is expensive to carry out individual calculations; and the states obtained in this manner are not orthogonal, which hinders the identification of the excited states.

Instead, we calculate the excited states in the spirit of CI theory, by a linear expansion in the space spanned by all states of the form

$$|c\rangle = \sum_{\mu} c_{\mu} \hat{\tau}_{\mu} |CC\rangle = \exp(\hat{T}) \sum_{\mu} c_{\mu} \hat{\tau}_{\mu} |HF\rangle, \tag{225}$$

where the summation is carried out over all the excitation operators present in the cluster operator. This expansion is referred to as the *equation-of-motion coupled-cluster* (EOM-CC) model. The expansion parameters are optimized by the minimization of the energy

$$E(\mathbf{c}, \bar{\mathbf{c}}) = \frac{\langle \bar{\mathbf{c}} | \exp(-\hat{T}) \hat{H} \exp(\hat{T}) | \mathbf{c} \rangle}{\langle \bar{\mathbf{c}} | \mathbf{c} \rangle}. \quad (226)$$

The EOM-CC theory can be thought as conventional CI theory with a similarity-transformed Hamiltonian, which carries the information about electron correlation; while the configuration expansions carry the information about the excitation structure of the electronic states.

Orbital-optimized coupled-cluster theory

In the standard CC theory, we use the Hartree–Fock orbitals and then determine a set of non-zero single-excitation amplitudes together with higher-excitation amplitudes. Alternatively, we could use $\exp(\hat{T}_1)$ to generate an orbital transformation to a basis in which the single-excitation amplitudes vanish. This is the idea behind the *orbital-optimized (OCC) theory*, in which the orbital-rotation operator $\exp(-\hat{\kappa})$ is used instead of $\exp(\hat{T}_1)$ (as they generate the same state to first order),

$$|\text{OCC}\rangle = \exp(-\hat{\kappa}) \exp(\hat{T}_O) |0\rangle, \quad (227)$$

where $\hat{T}_O = \hat{T}_2 + \hat{T}_3 + \dots + \hat{T}_N$. The OCC energy is obtained in the usual manner by projecting the OCC Schrödinger equation against the reference state

$$E_{\text{OCC}} = \langle 0 | \exp(-\hat{T}_O) \exp(\hat{\kappa}) \hat{H} \exp(-\hat{\kappa}) \exp(\hat{T}_O) | 0 \rangle, \quad (228)$$

whereas the cluster amplitudes are determined by projection against the manifold $|\mu_O\rangle$ spanned by \hat{T}_O

$$\langle \mu_O | \exp(-\hat{T}_O) \hat{H} \exp(-\hat{\kappa}) \exp(\hat{T}_O) | 0 \rangle = 0. \quad (229)$$

The *Brueckner coupled-cluster theory* (BCC) is closely related to OCC and differs in some details of the solution of orbital rotation parameters. Unlike in standard CC theory, the

OCC and BCC orbitals are optimized simultaneously with the optimization of the cluster amplitudes – compare with the MCSCF theory. Surprisingly, the differences between standard and OCC/BCC wave functions are small. Only in the rare cases that are characterized by the sc. Hartree–Fock singlet instability the OCC/BCC are clearly more successful than the standard formulation.

3.5 Perturbation theory

Finally, we consider less rigorous but computationally more appealing improvements of the Hartree–Fock description, based on perturbation theory. They are rather successful, when the Hartree–Fock wave function is reasonably accurate.

In *Møller–Plesset perturbation theory* (MPPT) the electronic Hamiltonian is partitioned as

$$\hat{H} = \hat{f} + \hat{\Phi} \quad (230)$$

where \hat{f} is the Fock operator (180) and $\hat{\Phi}$ the *fluctuation potential*, which is the difference between the true two-electron Coulomb potential in \hat{H} and the effective one-electron potential \hat{V} of the Fock operator, $\hat{\Phi} = \hat{g} - \hat{V}$. The zero-order state is represented by the Hartree–Fock state in the canonical representation, that is $\hat{f}|\text{HF}\rangle = \sum_I \epsilon_I |\text{HF}\rangle$, and by applying the standard machinery of perturbation theory, we obtain to second order in the perturbation

$$E_{\text{MP}}^{(0)} = \langle \text{HF} | \hat{f} | \text{HF} \rangle = \sum_I \epsilon_I \quad (231)$$

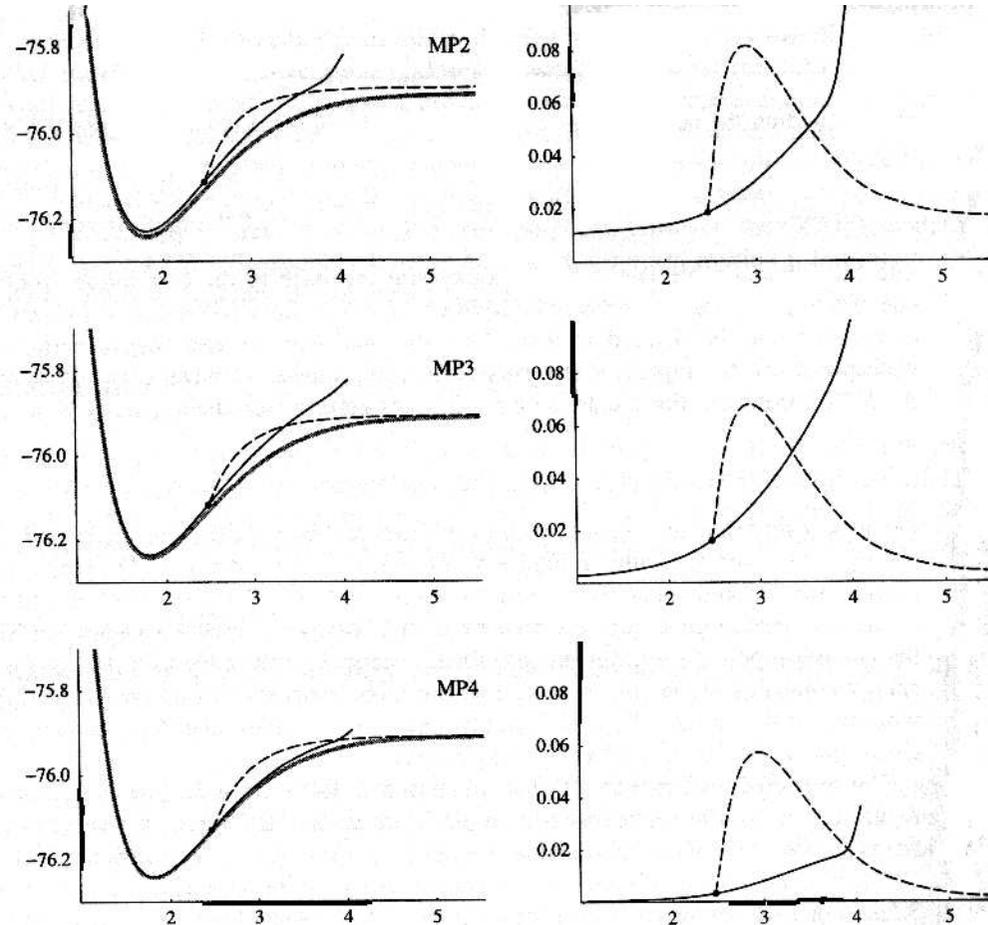
$$E_{\text{MP}}^{(1)} = \langle \text{HF} | \hat{\Phi} | \text{HF} \rangle \quad (232)$$

$$E_{\text{MP}}^{(2)} = - \sum_{A>B, I>J} \frac{|g_{AIBJ} - g_{AJBI}|^2}{\epsilon_A + \epsilon_B - \epsilon_I - \epsilon_J}. \quad (233)$$

Thus, the Hartree–Fock energy is equal to the sum of the zero- and first-order contributions; and by adding the second-order correction, we obtain the *second-order Møller–Plesset energy* (MP2)

$$E_{\text{MP2}} = E_{\text{HF}} - \sum_{A>B, I>J} \frac{|g_{AIBJ} - g_{AJBI}|^2}{\epsilon_A + \epsilon_B - \epsilon_I - \epsilon_J}, \quad (234)$$

which is very popular and successful approach to introducing electron correlation on top of the Hartree–Fock wave function; providing surprisingly accurate and size-extensive correction at low computational cost. Higher-order corrections are derived in a similar manner, yielding MP3 and MP4 models, but these become rather expensive and are not as appealing compromise between cost and accuracy as MP2 is. Furthermore, the MPPT series have been shown to be inherently divergent.



The MP_n dissociation curves of the C_{2v} water molecule in the cc-pVDZ basis. Restricted MP2 with full and unrestricted with dashed line.

In the coupled-cluster perturbation theory (CCPT) the Hamiltonian in the CC equations (212) and (213) is partitioned as in Eq. (230). The Fock and the cluster operator obey the following results:

$$\begin{aligned} [\hat{f}, \hat{T}] &= \sum_{\mu} \epsilon_{\mu} t_{\mu} \hat{\tau}_{\mu} \\ [[\hat{f}, \hat{T}], \hat{T}] &= [[[\hat{f}, \hat{T}], \hat{T}], \hat{T}] = \dots = 0. \end{aligned}$$

Thus we may write the non-Hermitian *similarity-transformed Fock operator* as

$$\hat{f} = \hat{f} + \sum_{\mu} \varepsilon_{\mu} t_{\mu} \hat{T}_{\mu}, \quad (235)$$

and we arrive at the working equations of the CCPT:

$$E = E_0 + \langle \text{HF} | \hat{\Phi} | \text{HF} \rangle \quad (236)$$

$$\varepsilon_{\mu} t_{\mu} = - \langle \mu | \hat{\Phi} | \text{HF} \rangle. \quad (237)$$

By expanding the cluster operator in the orders of the fluctuation potential we arrive at the amplitude equations

$$\varepsilon_{\mu} t_{\mu}^{(n)} = - \langle \mu | [\hat{\Phi}]^{(n)} | \text{HF} \rangle \quad (238)$$

where $[\hat{\Phi}]^{(n)}$ contains the n^{th} order part of the similarity-transformed fluctuation potential. Then the amplitudes would be determined self-consistently from the projected equations after truncating the cluster operator at some excitation level. The CCPT is most often used to establish approximations to contributions from some excitations – the most famous is the non-iterative triples, or (T) correction to the CCSD wave function. Also very successful iterative CCPT schemes have been set up, e.g. the CC2 model approximates the CCSD model with the computational cost $\mathcal{O}(N^5)$ and CC3 the CCSDT model with the computational cost $\mathcal{O}(N^7)$. The CCPT series provides more systematic improvement to the wave function than the MPPT.

Both MPPT and CCPT are limited to the ground states of systems dominated by a single configuration. To overcome this, *multiconfigurational perturbation theory* has been developed. Most important such theory is *CAS perturbation theory* (CASPT), where the zero-order state is taken to be a CASSCF wave function. At the moment, it is the only generally applicable method for the *ab initio* calculation of dynamical correlation effects of open- and closed-shell multiconfigurational electronic system.

3.7 Performance of the electronic-structure models

Having now considered the most important approximations to the molecular electronic wave function we will now address their performance in terms of a few numerical examples. It should be remembered that none of these models (besides FCI, which is inapplicable for most cases) is an all-around-good one: good performance in some problem and/or in some system does not guarantee good performance in another situation. In addition, rehearse the graph concerning the approximations to the one- and N -electron spaces.

Case study 1: molecular equilibrium geometries

One of the most common applications of molecular electronic-structure calculations is the determination of the molecular equilibrium structures. Good geometries are also of fundamental importance in study of molecular properties.

We consider the mean errors and standard deviations of calculated bond distances and bond angles as compared to experimental results using a hierarchy of N -electron models and a hierarchy of Gaussian basis sets. The error is given by $\Delta = P_i^{\text{calc}} - P_i^{\text{exp}}$, and the mean error and standard deviation by

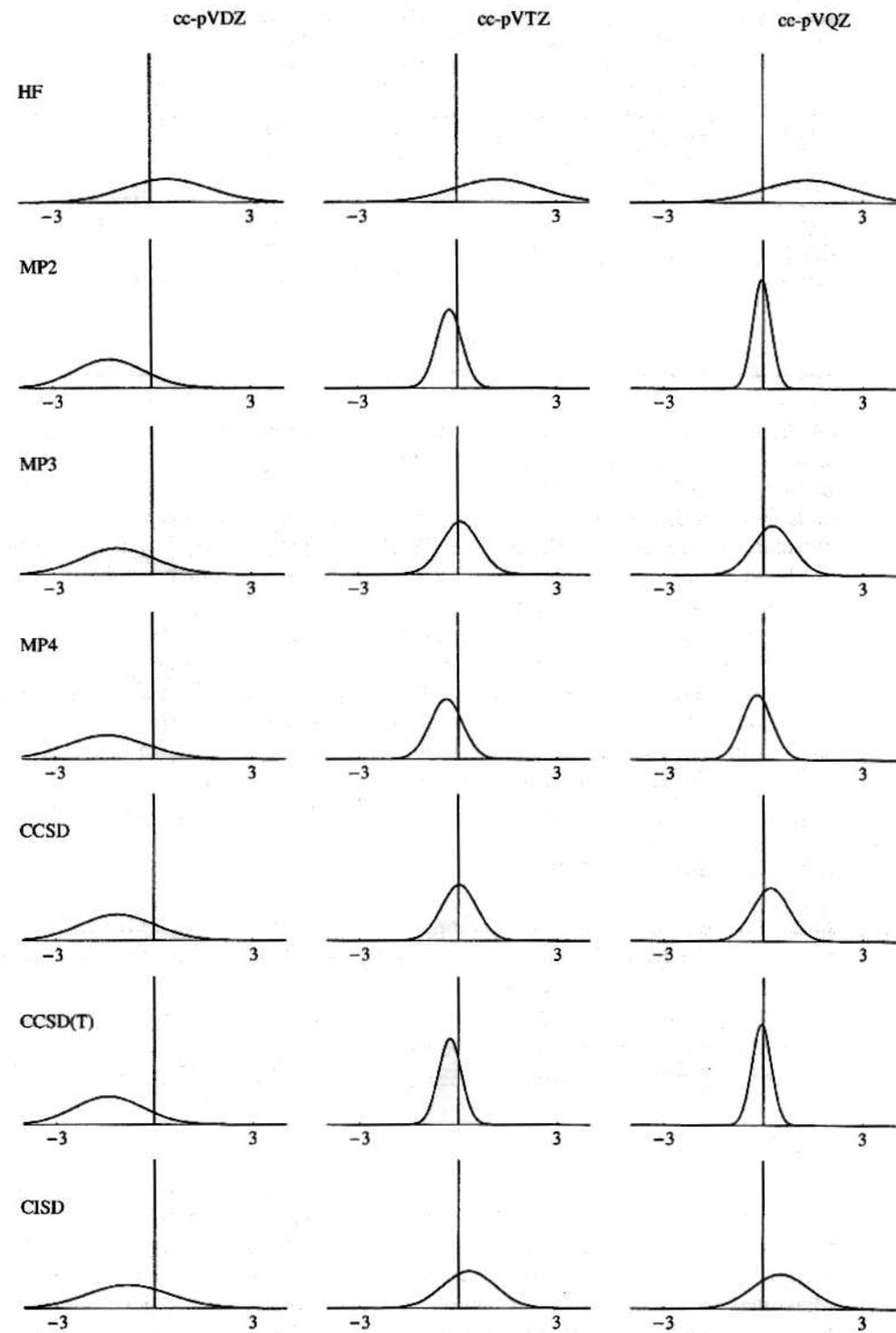
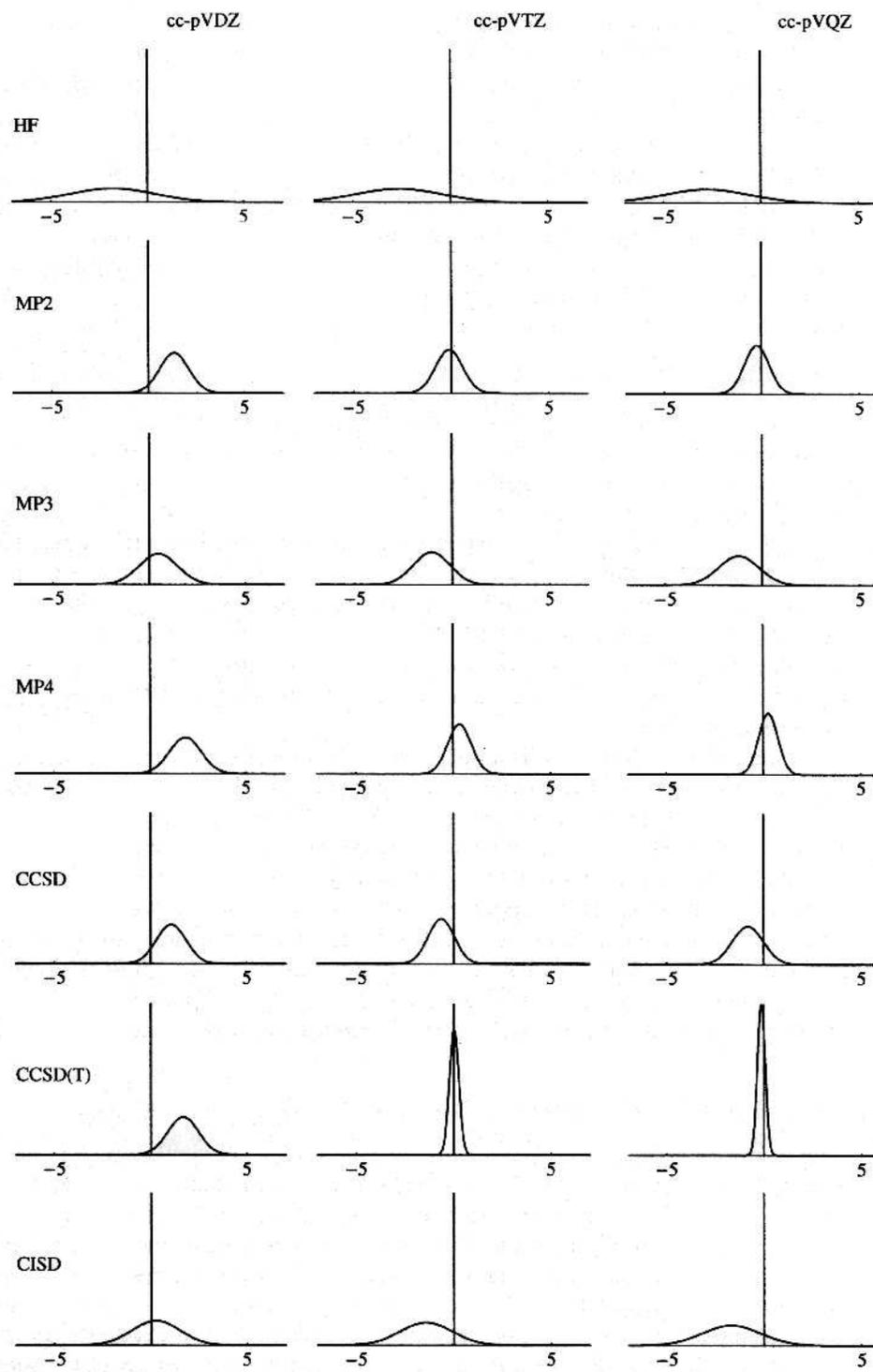
$$\bar{\Delta} = \frac{1}{n} \sum_{i=1}^n \Delta_i \quad (239)$$

$$\Delta_{\text{std}} = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (\Delta_i - \bar{\Delta})^2}. \quad (240)$$

They can be conveniently represented graphically in terms of the normal distribution

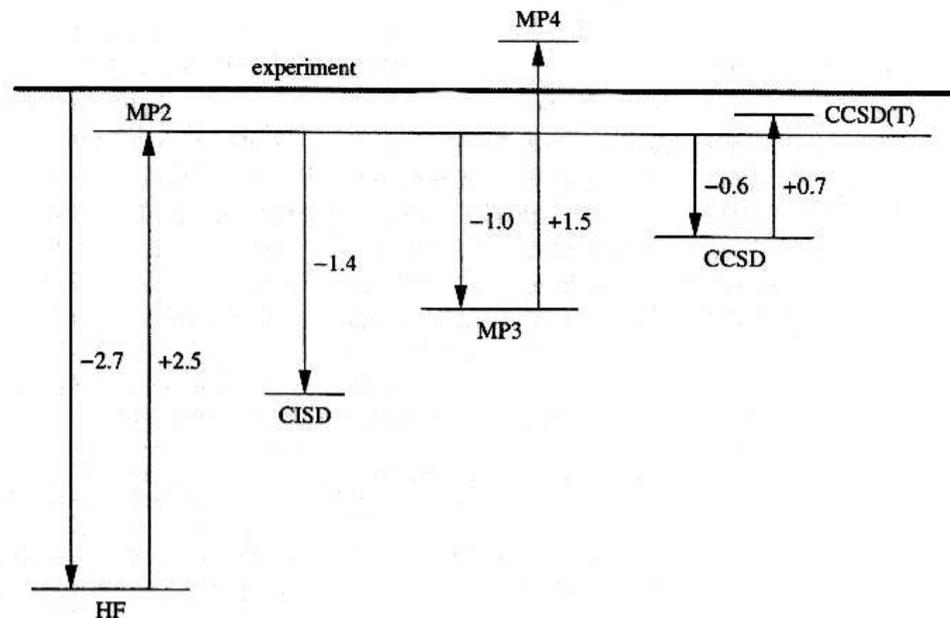
$$\rho(P) = N_c \exp \left[-\frac{1}{2} \left(\frac{P - \bar{\Delta}}{\Delta_{\text{std}}} \right)^2 \right]. \quad (241)$$

The calculations were carried out for sets of small molecules containing main-group elements. 29 bond distances and 10 bond angles were calculated.



Observations (distances on the left, angles on the right):

- The Hartree–Fock model is suitable for only preliminary calculations.
- The CCSD(T) model is capable of providing structures – both bond distances and angles – of high quality.
- The MP2 model is less accurate but more generally applicable as it is considerably less expensive. Higher-order MPPT is significantly less accurate but hideously expensive, i.e., useless.



Changes in bond lengths calculated from $\bar{\Delta}$ of the cc-pVQZ basis.

- The performance of the CCSD model is disappointing. It however performs more robust than MP2 in some selected cases with complicated electronic structure.
- The CISD model is also rather useless, showing no improvement when the basis set is increased.
- The basis-set requirements are different for different models. Whereas there is no much use to use QZ-level basis sets with the Hartree–Fock model (as the results are not improved), the correlated models require at least a TZ-level basis set, but preferably the cc-pVQZ basis.

The results suggest the following recipe for the determination of molecular structures: obtain an initial structure with HF using a DZ-level basis and improve it by using MP2 and a TZ basis. If ever possible, calculate the final geometry in the cc-pVQZ basis using the CCSD(T) model.

Case study 2: Molecular dipole moments

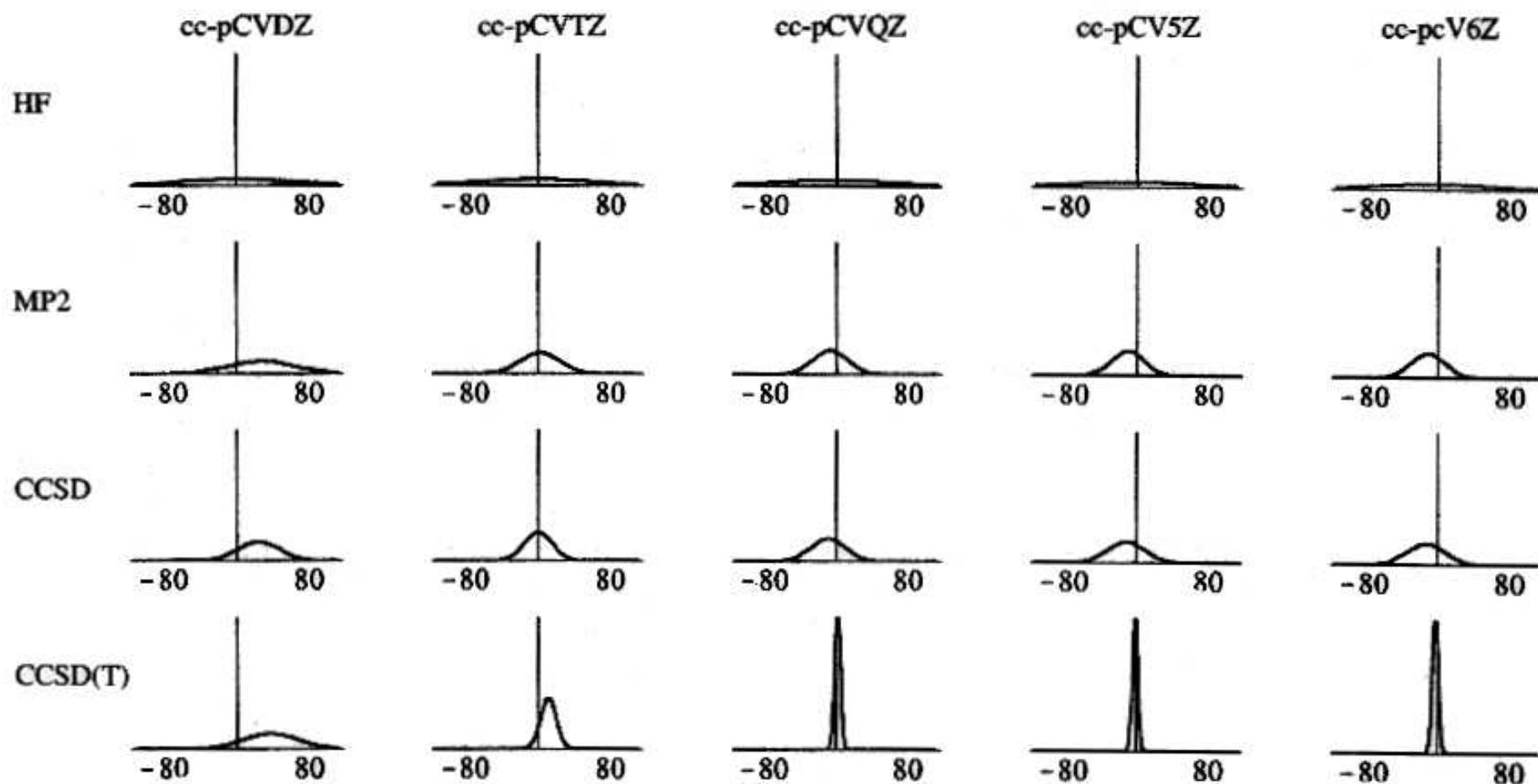
The comparison between the experimental and calculated molecular dipole moments is difficult, as the experiments are measuring the dipole moment in the vibrational ground state μ_0 , whereas the calculations are carried out for the equilibrium dipole moment μ_e , and thus we would have to carry out a vibrational averaging in order to speak of the same quantity. However, there are a few experimental values for μ_e , and we shall compare the performance of our standard models with those. Note, that due to the outer-valence origin of this property, we have to augment our basis sets with extra diffuse primitive Gaussians; hence we are using the aug-cc-pVXZ basis sets. The geometries are optimized using the same wave function as for the calculation of the dipole moment.

	Exp.	HF	MP2	CCSD	CCSD(T)		
		QZ	QZ	QZ	DZ	TZ	QZ
NH ₃	1.561(5)	1.544	1.517	1.530	1.541	1.513	1.521
HF	1.803(2)	1.884	1.811	1.813	1.799	1.797	1.800
H ₂ O	1.8473(10)	1.936	1.864	1.870	1.859	1.845	1.853

The quality of the calculation is more dependent on the correlation treatment than the cardinal number of the basis set; yet it is necessary to augment the basis set. The further augmentation (d-aug-cc-pVXZ basis sets) of the basis set does not affect the calculated dipole moments significantly.

Case study 3: Reaction enthalpies

Next, we study the reaction enthalpies of 13 chemical reactions among closed-shell molecules. The calculations are carried out in the HF–MP2–CCSD–CCSD(T) hierarchy, using the correlation-consistent basis set hierarchy supplemented with steep functions to provide core-polarization, which – somewhat surprisingly – must be included in the high-precision calculation of reaction enthalpies.



Notes on total electronic energies

For the total molecular energy, i.e., E in the molecular Schrödinger equation, there is no experimental counterpart. and we examine it in order to establish a feeling on the severity of the approximations involved in the calculation. We should recall that there were a third class of approximations in addition to the truncation of one- and N -electron spaces: approximations in the molecular Hamiltonian \hat{H} . To investigate the validity of the use of the non-relativistic Hamiltonian, we include the leading-order one-electron relativistic corrections that include the spin-orbit interaction (SO), mass-velocity (MV), and the Darwin (Dar) corrections. The leading-order two-electron contributions, such as the two-electron Darwin contribution and the spin-spin contact interaction, are smaller by at least one order of magnitude. The MV and Dar corrections are always of opposite sign. The calculation is carried out using the CCSDT model for the water molecule in the cc-pCVXZ bases at a CCSD(T)/cc-pCVQZ geometry.

	Total energy [E_h]		Energy corrections [mE_h]					
	CCSDT	HF	HF→SD	SD→(T)	(T)→T	SO	MV	Dar
DZ	-76.24121	-76.02680	-211.21	-3.04	-0.16			
TZ	-76.33228	-76.05716	-267.40	-7.65	-0.08			
QZ	-76.35981	-76.06482	-285.98	-8.99	-0.01	0.00	-255.87	203.69
5Z	-76.36899	-76.06708	-292.44	-9.52	0.04			

Observations:

- It is remarkable how a tiny, less than one per cent fraction of the total energy, plays such an important role in the determination of molecular properties.
- The leading-order relativistic correction, $-52.18 mE_h$, is larger than the triples contribution already in water, and gets far more pronounced in systems with heavier elements present!
- (T) is a very good approximation to full triples.

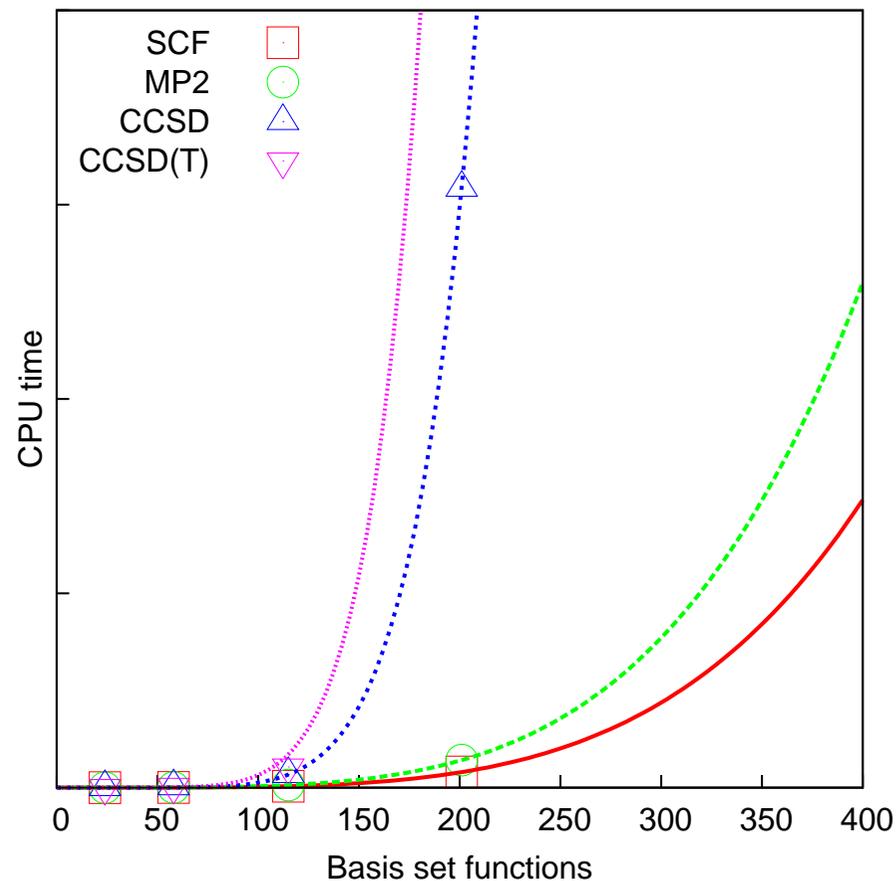
3.8 Linear scaling

Many years of theoretical research and the on-going revolution in computer technology have made quantum mechanics applicable to small molecules, where quantum-mechanical calculations have reached an accuracy often challenging experimental results. However, limitations of existing quantum mechanical methods to describe the large molecular systems, that modern molecular science often deals with, represents a road block to current progress. This suggests that profoundly different approaches must be taken to describe such systems.

Linear scaling

It is necessary to reduce the formal scaling of the models in order to apply them to larger molecular systems. This have been under intense study for many years, and significant progress have been achieved in the cases of SCF and DFT, which are based upon

- The *fast multipole method*: the number of significant two-electron integrals scales as $\mathcal{O}(N^2)$, but only $\mathcal{O}(N)$ have to be evaluated using the standard quantum-chemical means – the rest can be approximated by far simpler formulas. This applies to AOs whose overlap is close enough to zero.



- Linear scaling Fock-matrix construction: the contributions to the Fock matrix in the AO basis

$$f_{\mu\nu} = h_{\mu\nu} + \sum_{\rho\sigma} D_{\rho\sigma}^{\text{AO}} (2g_{\mu\nu\rho\sigma} - g_{\mu\sigma\rho\nu}) \quad (242)$$

can be prescreened by the AO density matrix.

- Direct optimization of the density matrix: the diagonalization step in the Roothaan–Hall SCF is avoided. These involve plenty of matrix-matrix multiplications, but the involved matrices are hoped to be sparse enough (in spatially large molecules) to allow $\mathcal{O}(N)$ evaluation of the multiplications.

These are carried out almost routinely at the moment. However, linear scaling black-box correlated wave-function theory is far away from reality.

Local correlation approaches

Traditional correlated electronic wave function methods, which are built upon the delocalized Hartree–Fock molecular orbitals (MO) do not exploit the fact that the electron correlation—i.e., the electron-electron Coulomb cusp—is spatially a rather local phenomenon. In the recent literature, localized correlation methods have been pursued to overcome the “scaling wall” encountered in the correlated wave function methods. However, despite recent progress, there is still a need to formulate and implement new approaches that exploit the locality of the electron correlation problem.

- Approaches based on introducing locality in the correlation treatment using localized MOs (LMO) appeared already in 1960s and in the context of CC methods in 1980s.
- An alternative approach was presented by Head-Gordon *et al.*, who formulated the CC equations in a general non-orthogonal [e.g., the atomic orbital] basis using a

general tensor-based formulation. Scuseria and Ayala presented a transformation of the standard MO CC doubles (CCD) equations into the AO basis. Scuseria and Ayala argued that the solution of the CC equations in the AO basis scales in principle linearly with respect to N , but their approach seems so far not to have received any widespread use.

- A fully operational local CCSD implementation have been presented by Werner *et al.* This implementation, referred to as the local CCSD (LCCSD) method, is based on localizing the occupied space as suggested by Pulay and coworkers, using traditional MO localization schemes, and using for the virtual space a projected atomic orbital (PAO) basis. LCCSD further relies on a spatial division of the system into domains. Restrictions on the extent to which the interactions between different domains are accounted for is important for achieving the low-power scaling of LCCSD making the results deviate from the standard CCSD results.
- The equations of correlated wave function theories may also be derived in a biorthogonal projected atomic orbital basis, aiming at a black-box formulation that could be applied to large systems. This formulation also constitutes a good platform to introduce efficient approximation schemes to reduce the number of parameters and thus the computational scaling. The locality of the PAO basis is reflected in the fact that the cluster amplitudes, the integrals as well as the correlation energy contributions decay exponentially with respect to the distance between the centers of the participating PAOs.

Correlated wave function theory in the PAO basis

Consider two orthonormal, non-redundant sets of functions, $\{|\psi_i\rangle\}$ and $\{|\psi_a\rangle\}$. We define the projection operators

$$\hat{P} = \sum_i |\psi_i\rangle\langle\psi_i|, \quad \hat{Q} = \sum_a |\psi_a\rangle\langle\psi_a|. \quad (243)$$

These projection operators are idempotent and mutually orthogonal.

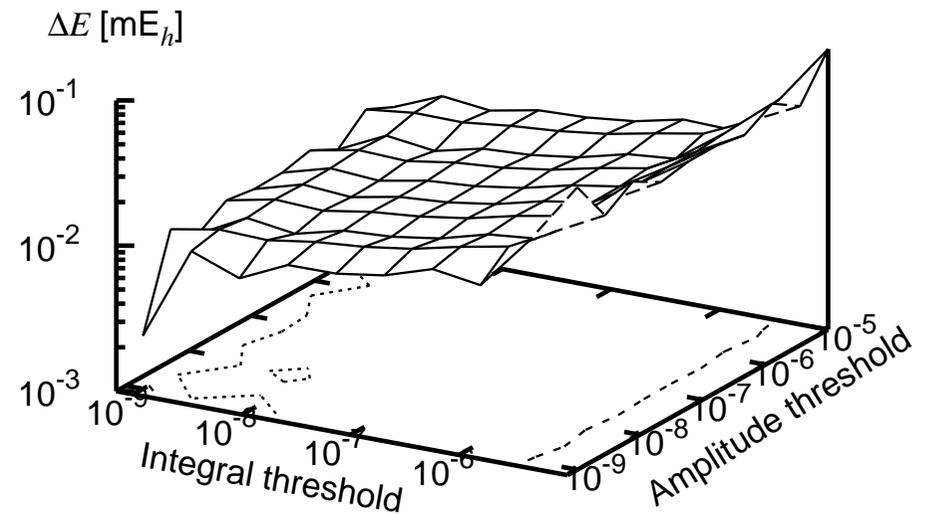
From these, we form the projected atomic orbitals (PAO) $|\phi_K\rangle = \hat{P}|\chi_\kappa\rangle$, $|\phi_A\rangle = \hat{Q}|\chi_\alpha\rangle$ And the composite PAO basis

$$\{|\tilde{\phi}\rangle\} = \{\hat{P}|\chi\rangle\} \oplus \{\hat{Q}|\chi\rangle\} \quad (244)$$

The PAO basis is non-orthogonal, $[a_R^\dagger, a_S]_+ = \tilde{S}_{RS}$, which leads to complicated amplitude equations. However, it can be “bi-orthonormalized”; $[a_R^\dagger, \bar{a}_S]_+ = \delta_{RS}$, $\bar{a}_S = [\tilde{S}^{-1}]_{TS} a_T$

⇒ Everything is similar with the MO basis; differences are absorbed in integrals, which may be obtained from the standard AO integrals

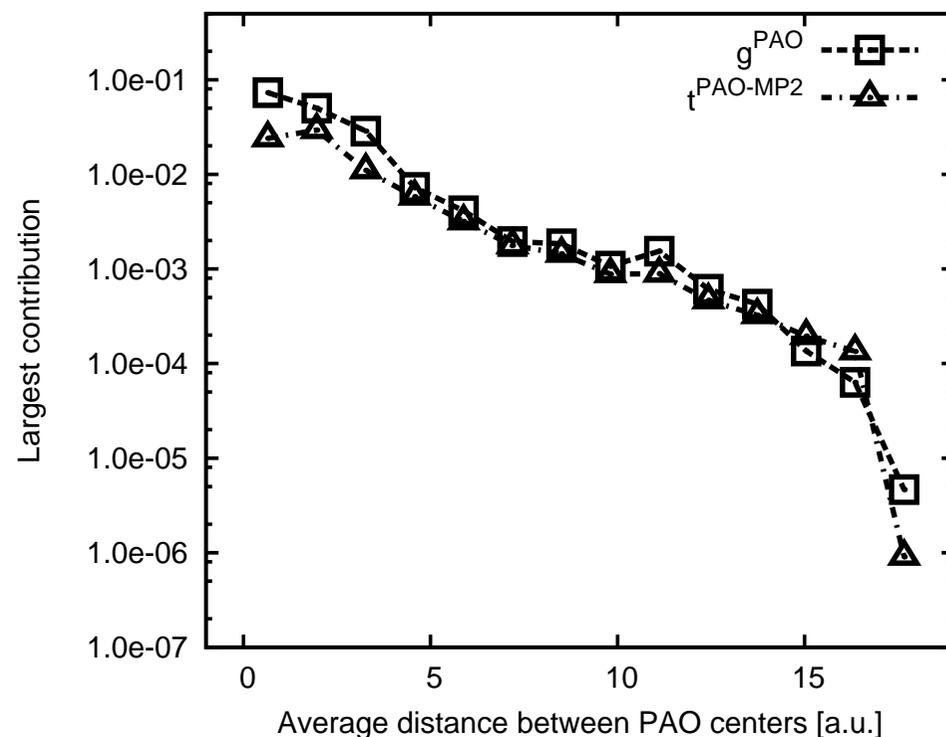
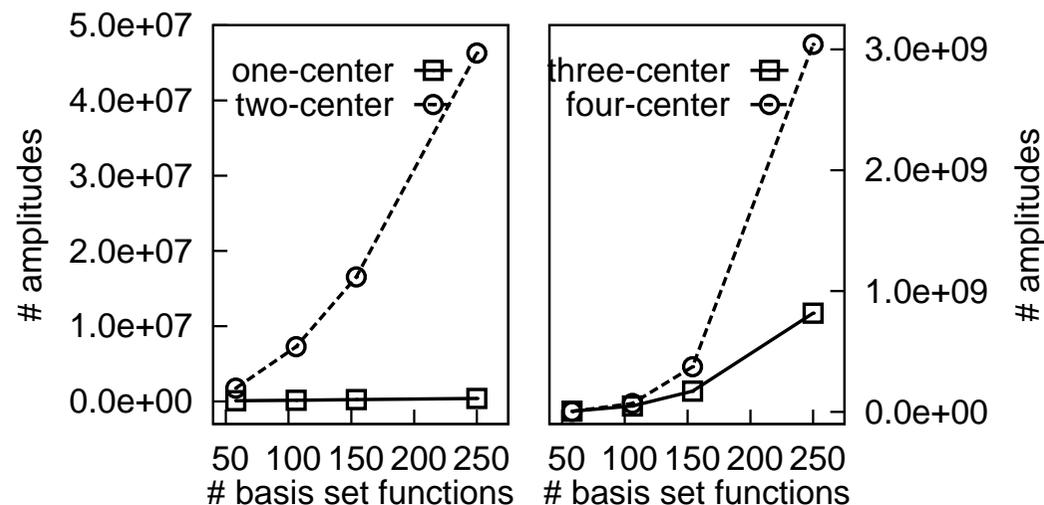
- MP2 has been implemented so far; the implementation involves no computation with zero amplitudes or integrals.
- Two controllable parameters
 - Threshold for integrals
 - Threshold for amplitudes



- No spatial division, no localized MOs \Rightarrow no convergence problems in systems with delocalized electronic structure.

The present MP2 implementation scales linearly in the number of non-zero elements [$\mathcal{O}(N^2)$]. Real gains lie, however, in the coupled-cluster models.

- The PAO scheme provides many ways to *truncate the number of parameters*, e.g. distance-based screening, neglecting four-center contributions...
- The PAO formulation is seen to be indeed local: the relevant quantities decay exponentially.



The integrals and converged MP2 amplitudes in the PAO basis. $\text{C}_{10}\text{H}_{22}$ at the cc-pVDZ basis set.