

# SCF methods, basis sets, and integrals

## Lecture II: Spin and Roothaan–Hall

Wim Klopper

Abteilung für Theoretische Chemie, Institut für Physikalische Chemie  
Karlsruher Institut für Technologie (KIT)

ESQC 2019, 8–21 September 2019

# Preface

- So far, we have considered the molecular orbitals  $\varphi_k$  to depend on the coordinates of the electron.
- One may have assumed that they are the three Cartesian coordinates of the electron's position in space,  $\mathbf{r}_i = \{x_i, y_i, z_i\}$ .
- At the same time, we have introduced the **antisymmetry** of the wavefunction in an *ad hoc* manner, since the electrons are fermions.
- The electron is known to have a **spin with a quantum number**  $s = \frac{1}{2}$ , the  $z$ -component of which is quantised to take one of the possible values,  $m_s = \pm\frac{1}{2}$ .
- In order for electron spin to have meaning in Hartree–Fock theory, we must let the orbitals depend on it.

## Spin functions $\alpha$ and $\beta$

- Let the functions  $\alpha$  and  $\beta$  be **orthonormal** eigenfunctions of the spin operators  $\hat{s}^2$  and  $\hat{s}_z$ ,

$$\hat{s}^2\alpha = \frac{3}{4}\alpha, \quad \hat{s}_z\alpha = +\frac{1}{2}\alpha$$

$$\hat{s}^2\beta = \frac{3}{4}\beta, \quad \hat{s}_z\beta = -\frac{1}{2}\beta$$

- We let the orbitals depend on the spin by attaching an extra label (superscript) to it,

$$\varphi_{k\alpha} = \varphi_k^\alpha(\mathbf{r})\alpha, \quad \varphi_{k\beta} = \varphi_k^\beta(\mathbf{r})\beta$$

- The functions  $\varphi_k^\alpha(\mathbf{r})$  and  $\varphi_k^\beta(\mathbf{r})$  are **spatial orbitals**, that is, functions that depend only on  $\mathbf{r} = (x, y, z)$ .
- The functions  $\varphi_{k\alpha}$  and  $\varphi_{k\beta}$  are **spin orbitals** that depend on spatial *and* spin coordinates.

# Sums over spin orbitals

- Having attached a spin label to the orbitals, we must write sums of the orbitals in the following manner:

$$\text{without spin : } \sum_{k=1}^n \Rightarrow \text{with spin : } \sum_{\sigma=\alpha,\beta} \sum_{k=1}^{n_{\sigma}} \quad (n_{\alpha} + n_{\beta} = n)$$

- There is a sum over  $\sigma$  over the two spin cases and then (dependent on the first sum) a second sum over the number of orbitals with that spin.
- Note that sums over electrons just remain unchanged, for example

$$\hat{H} = \hat{V}_{nn} + \sum_{\mu=1}^n \hat{h}_{\mu} + \sum_{\nu=2}^n \sum_{\mu=1}^{\nu-1} \frac{1}{r_{\mu\nu}}$$

# Integration over spin coordinates

- To take spin into account, we must rewrite sums over orbitals in the following manner:

$$\sum_{i,j} h_{ij} = \sum_{i=1}^n \sum_{j=1}^n \langle \varphi_i | \hat{h} | \varphi_j \rangle \Rightarrow \sum_{\sigma=\alpha,\beta} \sum_{k=1}^{n_\sigma} \sum_{\rho=\alpha,\beta} \sum_{l=1}^{n_\rho} \langle \varphi_{k\sigma} | \hat{h} | \varphi_{l\rho} \rangle$$

- The expressions seem to become more lengthy, but fortunately, the rules for spin integration are very simple,

$$\langle \alpha | \alpha \rangle = \langle \beta | \beta \rangle = 1, \quad \langle \alpha | \beta \rangle = \langle \beta | \alpha \rangle = 0$$

- Hence, in terms of the Kronecker delta function,

$$\langle \varphi_{k\sigma} | \hat{h} | \varphi_{l\rho} \rangle = \langle \sigma | \rho \rangle \times \langle \varphi_k^\sigma | \hat{h} | \varphi_l^\rho \rangle = \delta_{\sigma\rho} \int [\varphi_k^\sigma(\mathbf{r})]^* \hat{h} \varphi_l^\rho(\mathbf{r}) d\mathbf{r}$$

# Integration over spin coordinates

- Spin integration thus yields

$$\sum_{i,j}^n h_{ij} = \sum_{k,l=1}^{n_\alpha} \langle \varphi_k^\alpha | \hat{h} | \varphi_l^\alpha \rangle + \sum_{k,l=1}^{n_\beta} \langle \varphi_k^\beta | \hat{h} | \varphi_l^\beta \rangle$$

- Note that the integration would be more complicated for an operator that depends on spin, such as  $\hat{\mathbf{l}} \cdot \hat{\mathbf{s}}$ ,

$$\langle \varphi_{k\sigma} | \hat{\mathbf{l}} \cdot \hat{\mathbf{s}} | \varphi_{l\rho} \rangle = \langle \varphi_k^\sigma | \hat{\mathbf{l}} | \varphi_l^\rho \rangle \cdot \langle \sigma | \hat{\mathbf{s}} | \rho \rangle$$

- We shall not be concerned with spin-dependent operators in this course! → course by Trond Saue

# Two-electron integration over spin coordinates

- We are now ready to integrate over the spin coordinates in antisymmetrised two-electron integrals,

$$\begin{aligned}\langle \varphi_{i\sigma} \varphi_{j\rho} || \varphi_{k\tau} \varphi_{l\omega} \rangle &= \langle \varphi_{i\sigma} \varphi_{j\rho} | \varphi_{k\tau} \varphi_{l\omega} \rangle - \langle \varphi_{i\sigma} \varphi_{j\rho} | \varphi_{l\omega} \varphi_{k\tau} \rangle \\ &= \delta_{\sigma\tau} \delta_{\rho\omega} \langle \varphi_i^\sigma \varphi_j^\rho | \varphi_k^\tau \varphi_l^\omega \rangle - \delta_{\sigma\omega} \delta_{\rho\tau} \langle \varphi_i^\sigma \varphi_j^\rho | \varphi_l^\omega \varphi_k^\tau \rangle\end{aligned}$$

- Hence, we obtain for four different cases:

$$\begin{aligned}\langle \varphi_{i\alpha} \varphi_{j\alpha} || \varphi_{k\alpha} \varphi_{l\alpha} \rangle &= \langle \varphi_i^\alpha \varphi_j^\alpha | \varphi_k^\alpha \varphi_l^\alpha \rangle - \langle \varphi_i^\alpha \varphi_j^\alpha | \varphi_l^\alpha \varphi_k^\alpha \rangle \\ \langle \varphi_{i\alpha} \varphi_{j\alpha} || \varphi_{k\beta} \varphi_{l\beta} \rangle &= 0 \\ \langle \varphi_{i\alpha} \varphi_{j\beta} || \varphi_{k\alpha} \varphi_{l\beta} \rangle &= \langle \varphi_i^\alpha \varphi_j^\beta | \varphi_k^\alpha \varphi_l^\beta \rangle \\ \langle \varphi_{i\alpha} \varphi_{j\alpha} || \varphi_{k\alpha} \varphi_{l\beta} \rangle &= 0\end{aligned}$$

- Note that in case of  $\langle \alpha\beta || \alpha\beta \rangle = \langle \alpha\beta | \alpha\beta \rangle - \langle \alpha\beta | \beta\alpha \rangle$ , the last term vanishes due to spin.

## Two-electron integration over spin coordinates

- We are now ready to integrate over the spin coordinates in the **two-electron expectation value** over the Slater determinant,

$$\sum_{i,j} \langle ij || ij \rangle = \sum_{i,j} \langle \overset{\textcolor{red}{i}}{i} \overset{\textcolor{red}{j}}{j} || \overset{\textcolor{red}{i}}{i} \overset{\textcolor{red}{j}}{j} \rangle + \sum_{i,\bar{j}} \langle \overset{\textcolor{red}{i}}{i} \overset{\textcolor{blue}{j}}{\bar{j}} || \overset{\textcolor{red}{i}}{i} \overset{\textcolor{blue}{j}}{\bar{j}} \rangle + \sum_{\bar{i},j} \langle \overset{\textcolor{blue}{i}}{\bar{i}} \overset{\textcolor{red}{j}}{j} || \overset{\textcolor{blue}{i}}{\bar{i}} \overset{\textcolor{red}{j}}{j} \rangle + \sum_{\bar{i},\bar{j}} \langle \overset{\textcolor{blue}{i}}{\bar{i}} \overset{\textcolor{blue}{j}}{\bar{j}} || \overset{\textcolor{blue}{i}}{\bar{i}} \overset{\textcolor{blue}{j}}{\bar{j}} \rangle$$

- For simplicity,  $\alpha$  spin orbitals  $\varphi_k^\alpha(\mathbf{r})\alpha$  are denoted by the index “ $k$ ” and  $\beta$  spin orbitals  $\varphi_k^\beta(\mathbf{r})\beta$  by “ $\bar{k}$ ”.
- We obtain:

$$\sum_{i,j} \langle ij || ij \rangle = \sum_{i,j} \langle \overset{\textcolor{red}{i}}{i} \overset{\textcolor{red}{j}}{j} || \overset{\textcolor{red}{i}}{i} \overset{\textcolor{red}{j}}{j} \rangle + \sum_{\bar{i},\bar{j}} \langle \overset{\textcolor{blue}{i}}{\bar{i}} \overset{\textcolor{blue}{j}}{\bar{j}} || \overset{\textcolor{blue}{i}}{\bar{i}} \overset{\textcolor{blue}{j}}{\bar{j}} \rangle + 2 \sum_{i,\bar{j}} \langle \overset{\textcolor{red}{i}}{i} \overset{\textcolor{blue}{j}}{\bar{j}} || \overset{\textcolor{red}{i}}{i} \overset{\textcolor{blue}{j}}{\bar{j}} \rangle$$

- The sum over “ $k$ ” runs from 1 to  $n_\alpha$  and the sum over “ $\bar{k}$ ” runs from 1 to  $n_\beta$ .



## $\langle \Phi | \hat{H} | \Phi \rangle$ with spin orbitals

- Putting it all together, we obtain the following expression for the expectation values of the Hamiltonian over the SD:

$$E[\Phi] = V_{nn} + \sum_i h_{ii} + \sum_{\bar{i}} h_{\bar{i}\bar{i}} + \frac{1}{2} \sum_{i,j} \langle i j | i j \rangle + \frac{1}{2} \sum_{\bar{i}, \bar{j}} \langle \bar{i} \bar{j} | \bar{i} \bar{j} \rangle + \sum_{i, \bar{j}} \langle i \bar{j} | i \bar{j} \rangle$$

- We can split the last term into two sums

$$\sum_{i, \bar{j}} \langle i \bar{j} | i \bar{j} \rangle = \frac{1}{2} \sum_i \langle i | \hat{J}^{(\beta)} | i \rangle + \frac{1}{2} \sum_{\bar{j}} \langle \bar{j} | \hat{J}^{(\alpha)} | \bar{j} \rangle$$

and obtain

$$\begin{aligned} E[\Phi] &= V_{nn} + \sum_i h_{ii} + \frac{1}{2} \sum_i \langle i | \hat{J}^{(\alpha)} + \hat{J}^{(\beta)} - \hat{K}^{(\alpha)} | i \rangle \\ &+ \sum_{\bar{j}} h_{\bar{j}\bar{j}} + \frac{1}{2} \sum_{\bar{j}} \langle \bar{j} | \hat{J}^{(\alpha)} + \hat{J}^{(\beta)} - \hat{K}^{(\beta)} | \bar{j} \rangle \end{aligned}$$

## Acting with $\hat{J}$ on spin orbitals

- Next, we investigate how  $\hat{J}$  and  $\hat{K}$  act on a spin orbital  $\varphi_p^\sigma(\mathbf{r})\sigma$  (with  $\sigma = \alpha$  or  $\beta$ ),

$$\hat{J}\varphi_p^\sigma(\mathbf{r})\sigma = \sum_{\rho=\alpha,\beta} \sum_{m=1}^{n_\rho} \left( \int \frac{[\varphi_m^\rho(\mathbf{r}')]^* \varphi_m^\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' \right) \varphi_p^\sigma(\mathbf{r})\sigma$$

- Same as before!
- For the **exchange operator**  $\hat{K}$ , however, we obtain:

$$\hat{K}\varphi_p^\sigma(\mathbf{r})\sigma = \sum_{\rho=\alpha,\beta} \sum_{m=1}^{n_\rho} \left( \delta_{\sigma\rho} \int \frac{[\varphi_m^\rho(\mathbf{r}')]^* \varphi_p^\sigma(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' \right) \varphi_m^\rho(\mathbf{r})\rho$$

- The Coulomb interaction always survives the spin integration.
- Exchange only occurs between electrons with like spins.**

## Exchange operators $\hat{K}^{(\alpha)}$ and $\hat{K}^{(\beta)}$

- Since exchange only occurs between electrons with like spins, we can define exchange operators  $\hat{K}^{(\alpha)}$  and  $\hat{K}^{(\beta)}$  that give the same result as  $\hat{K}$  when acting on a spin orbital with spin function  $\alpha$  respectively  $\beta$ ,

$$\hat{K}^{(\alpha)} f(\mathbf{r})\alpha = \sum_{m=1}^{n_\alpha} \left( \int \frac{[\varphi_m^\alpha(\mathbf{r}')]^* f(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' \right) \varphi_m^\alpha(\mathbf{r})\alpha, \quad \hat{K}^{(\alpha)} g(\mathbf{r})\beta = 0$$

$$\hat{K}^{(\beta)} f(\mathbf{r})\alpha = 0, \quad \hat{K}^{(\beta)} g(\mathbf{r})\beta = \sum_{m=1}^{n_\beta} \left( \int \frac{[\varphi_m^\beta(\mathbf{r}')]^* g(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' \right) \varphi_m^\beta(\mathbf{r})\beta$$

- Accordingly, we define

$$\hat{F}^{(\alpha)} = \hat{h} + \hat{J} - \hat{K}^{(\alpha)} \quad \text{and} \quad \hat{F}^{(\beta)} = \hat{h} + \hat{J} - \hat{K}^{(\beta)}$$

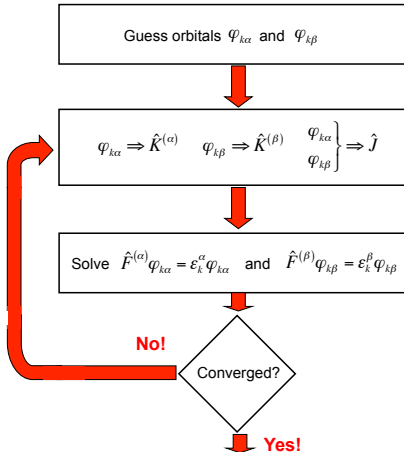
# Spin-unrestricted Hartree–Fock theory

- The straightforward implementation of the equations obtained thus far is known as *unrestricted Hartree–Fock* (UHF) theory.
- In UHF theory, there are two sets of MOs, which are the  $\alpha$  and  $\beta$  spin orbitals, eigenfunctions of the operators  $\hat{F}^{(\alpha)}$  and  $\hat{F}^{(\beta)}$ ,

$$\hat{F}^{(\alpha)}\varphi_k^\alpha(\mathbf{r})\alpha = \varepsilon_k^\alpha\varphi_k^\alpha(\mathbf{r})\alpha, \quad \hat{F}^{(\beta)}\varphi_k^\beta(\mathbf{r})\beta = \varepsilon_k^\beta\varphi_k^\beta(\mathbf{r})\beta$$

- The two equations are *coupled* through the Coulomb operator  $\hat{J}$ , which depends on *all* MOs.
- Note that the two sets of spatial orbitals  $\{\varphi_k^\alpha(\mathbf{r})\}$  and  $\{\varphi_k^\beta(\mathbf{r})\}$  are completely independent. Spatial orbitals of the  $\alpha$  set need not be orthogonal on the  $\beta$  set.
- The equations may be solved iteratively. The  $\alpha$  orbitals define  $\hat{K}^{(\alpha)}$ , the  $\beta$  orbitals define  $\hat{K}^{(\beta)}$ , and all orbitals define  $\hat{J}$ .

# Iterative solution of the UHF equations



- We shall see later that the MOs can be expanded in a basis of Gaussian-type orbitals (GTOs).
- That would be the MO-LCAO approach.
- Two sets of orbitals must be provided as an initial guess.
- Final result may depend on the first guess!
- Convergence to the lowest state is not guaranteed and convergence-accelerators may be needed.

# Spatial and spin symmetries

- Molecules may possess symmetry. Some may perhaps display no symmetry at all ( $C_1$  point group), but many are symmetric with respect to a point of inversion, a plane of reflection, etc.
- We define symmetry operations  $R_i$  of a given point group  $G$  that map the nuclear coordinates  $x_A$  onto themselves,  $R_i x_A = y_A$ . They induce operators in function space as

$$\hat{\mathcal{R}}_i \Phi(x_\mu; x_A) = \Phi(x_\mu; R_i^{-1} x_A) = \psi(x_\mu; x_A)$$

- These operators commute with the clamped-nuclei Hamiltonian,

$$[\hat{\mathcal{R}}_i, \hat{H}] = 0 \quad \forall R_i \in G$$

and hence, the exact electronic wavefunctions are (or can be chosen as) eigenfunctions of both  $\hat{\mathcal{R}}_i$  and  $\hat{H}$ ,

$$\hat{\mathcal{R}}_i \Psi_k(x_\mu; x_A) = \rho_{ik} \Psi_k(x_\mu; x_A)$$

# Spatial and spin symmetries

- As long as the Hamiltonian does not refer to spin, it commutes with spin operators, and the spin quantum numbers (e.g.,  $S$ ,  $M_S$ ) are **constants of motion**. Any electronic wavefunction should also be eigenfunction of the spin operators,

$$\hat{S}^2 \Psi_k = S(S+1) \Psi_k, \quad \hat{S}_z \Psi_k = M_S \Psi_k$$

Hence, we are looking for eigenfunctions of  $\hat{H}$ ,  $\hat{S}^2$ ,  $\hat{S}_z$  and all  $\hat{R}_i$  of the point group!

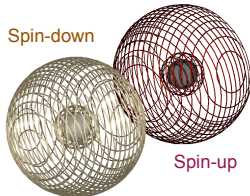
- The UHF approach as we know it so far yields approximate wavefunctions that **may or may not be eigenfunctions** of the spatial symmetry and spin operators.
- Restrictions** must be imposed to ensure that eigenfunctions of all of the relevant operators are obtained.

# A broken-symmetry UHF wavefunction

- A UHF calculation on two Na atoms at a distance of  $10 a_0$  can be performed in such a manner that the determinant becomes

$$\Phi = | \dots 3s_A^\alpha \alpha \dots 3s_B^\beta \beta \dots | \equiv | A\alpha B\beta |$$

- The  $1s$ ,  $2s$  and  $2p$  shells are fully occupied and can be ignored. The notation  $|A\alpha B\beta|$  means that we consider a two-electron determinant that consists of two spin orbitals: a  $3s$  atomic orbital on atom  $A$  with  $\alpha$  spin (spin-up) and a  $3s$  atomic orbital on atom  $B$  with  $\beta$  spin (spin-down).



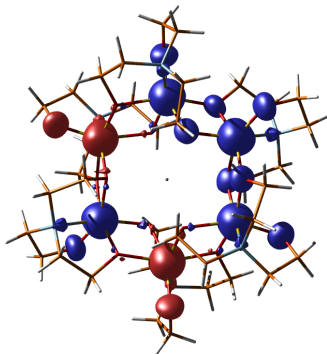
- The figure shows the spin density  $3s_A^2 - 3s_B^2$  obtained from  $|A\alpha B\beta|$ .
- $\Phi$  is **not eigenfunction** of an operator that interchanges the nuclei,

$$\hat{\mathcal{R}}_i |A\alpha B\beta| = |B\alpha A\beta|$$



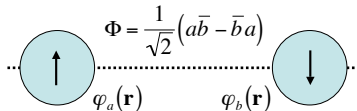
# Broken-symmetry UHF wavefunctions

- Although UHF determinants such as the one for  $\text{Na} \cdots \text{Na}$  seem unphysical, they may be useful in various applications of MO theory (e.g., *unrestricted Kohn–Sham (UKS)* theory).
- Heisenberg exchange couplings  $J$  may be estimated.



- The figure shows the **spin density** of a selected UKS solution for the “ferric wheel”  $\text{Na}^+@ \text{Fe}_6(\text{tea})_6$  ( $\text{H}_3\text{tea} = \text{triethanolamine}$ ).
- $\langle \Phi | \hat{S}_z | \Phi \rangle = 5 \hbar$ , but  $\langle \Phi | \hat{S}^2 | \Phi \rangle \approx 40 \hbar^2$ .
- The underlying UKS determinant is clearly not an eigenfunction of rotations about the  $S_6$  axis.

# Broken-symmetry UHF wavefunctions



- Let us return to the Na...Na system. The above SD is a proper antisymmetric wavefunction. Moreover,  $\hat{S}_z \Phi = 0$ . However,

$$\langle \hat{S}^2 \rangle = \langle \Phi | \hat{S}^2 | \Phi \rangle \approx 1 \quad \text{and} \quad \hat{\mathcal{R}}_{\text{inversion}} | a \bar{b} | = | b \bar{a} | \neq \pm | a \bar{b} |$$

- Since  $\langle \hat{S}_z \rangle = \langle \hat{S}_z^2 \rangle = 0$ ,  $\langle \hat{S}^2 \rangle$  is easily computed:

$$\hat{S}^2 = \hat{S}_+ \hat{S}_- + \hat{S}_z^2 - \hat{S}_z = \hat{S}_- \hat{S}_+ + \hat{S}_z^2 + \hat{S}_z$$

$$\hat{S}^2 | a \bar{b} | = \hat{S}_- \hat{S}_+ | a \bar{b} | = \hat{S}_- | a b | = | \bar{a} b | + | a \bar{b} |$$

$$\langle \hat{S}^2 \rangle = 1 - \langle \varphi_a | \varphi_b \rangle^2$$

# Spin-restricted Hartree–Fock theory

- In Hartree–Fock theory, the orbitals are optimised by the variation method. We assume that those orbitals leading to the lowest energy  $\langle \Phi | \hat{H} | \Phi \rangle$  are the “best” in a general sense.
- The orbitals can be optimised without worrying much about spin and spatial symmetries (as in UHF theory), but one can also apply **constraints**.
- The following table shows Hartree–Fock energies of the **NO· radical** ( $R_{\text{NO}} = 115.1 \text{ pm}$ , **def2-QVZPP basis**).

Spin symmetry constraints?	Spatial symmetry constraints?	State	$E/E_h$
No	No	?	–129.308 11
Yes	No	Doublet	–129.300 26
Yes	Yes	$^2\Pi$	–129.298 06

- What is the “best” calculation?

# Closed-shell Hartree–Fock theory

- An important and frequent situation occurs when the electronic situation under consideration is a spin singlet ( $S = 0, M_S = 0$ ).
- The number of electrons is *even*.
- We require that the spin orbitals occur in pairs having the same spatial function (**perfect spin pairing**),

$$\varphi_k^\alpha(\mathbf{r}) = \varphi_k^\beta(\mathbf{r})$$

- The determinant is then an eigenfunction of the operators  $\hat{S}^2$ ,  $\hat{S}_z$ ,  $\hat{S}_+$  and  $\hat{S}_-$ , with eigenvalue 0 in all cases.
- We find that

$$\begin{aligned} E[\Phi] &= V_{nn} + 2 \sum_i h_{ii} + \sum_{i,j} \langle ij || ij \rangle + \sum_{i,j} \langle ij | ij \rangle \\ &= V_{nn} + 2 \sum_i h_{ii} + \sum_{i,j} (2 \langle ij | ij \rangle - \langle ij | ji \rangle) \end{aligned}$$

# Closed-shell Hartree–Fock theory

- The two Fock operators  $\hat{F}^{(\alpha)}$  and  $\hat{F}^{(\beta)}$  are identical in closed-shell **restricted Hartree–Fock (RHF)** theory.
- Thus, the work and memory requirements are reduced by about 50% compared to UHF.
- The Fock operator can be written as

$$\hat{F}^{(\alpha)} = \hat{h} + 2\hat{J}^{(\alpha)} - \hat{K}^{(\alpha)}$$

- RHF is commonly used, since many stable main-group molecules have an even number of electrons and a singlet ground state.
- Often, when the UHF approach is applied to a stable closed-shell molecule, the same solution is obtained as with the RHF method (but with more work).
- Do we know when UHF gives a lower energy than RHF?

# Singlet and triplet instabilities in RHF theory

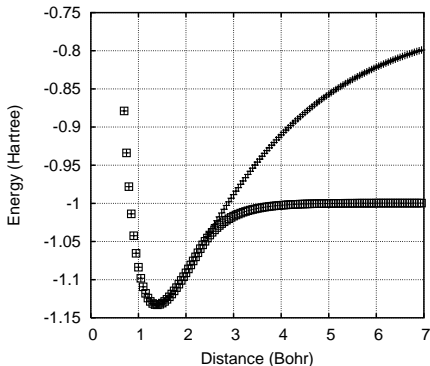
- We consider nondegenerate electronic states with an even number of electrons.
- Nondegenerate RHF states are **stationary with respect to symmetry-breaking** rotations, but they **may not be stable**, that is, they may be **saddle points**.
- To find out, one has to compute the second derivative with respect to the wavefunction parameters (**electronic Hessian**), including the parameters that change the RHF solution into a UHF state or a symmetric solution into a nonsymmetric wavefunction.
- **Triplet instabilities:** There is a UHF state lower than RHF.
- **Singlet instabilities:** A SD with symmetry lower than the molecule's point group has a lower energy than a SD that transforms as an irreducible representation of that point group.

## Triplet instability in $H_2$

- The figure shows potential energy curves of  $H_2$  obtained from RHF and UHF calculations (cc-pVTZ).
- Beyond about  $R = 2.3 a_0$ , the UHF treatment yields a lower energy than RHF.
- Note that the UHF guess must have

$$\varphi_k^\alpha \neq \varphi_k^\beta$$

- At  $R = 3 a_0$ , the spin-contaminated UHF state has  $\langle \hat{S}^2 \rangle \approx 0.67 \hbar^2$ .

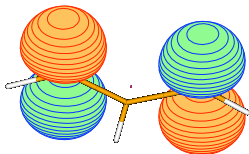


+ = RHF,    □ = UHF

# Singlet instability in the $C_3H_5\cdot$ radical

- We optimise the geometry at the restricted HF level (ROHF/cc-pVDZ) in  $C_{2v}$  symmetry  $\Rightarrow R_{CC} = 2.5980 a_0$ ,  $\angle_{CCC} = 124.49^\circ$ .
- Things look fine until we compute the harmonic vibrational frequencies from numerical differences of analytic gradients (as done with NumForce of Turbomole).
- We find one **imaginary frequency** with wavenumber  $\approx 3777i \text{ cm}^{-1}$ .

- The allyl radical ( $C_3H_5\cdot$ ) provides an example of a singlet instability.



- The unpaired electron occupies the nonbonding  $\pi$  orbital ( $1a_2$ ) and the SD transforms as  $^2A_2$ .



# Restricted open-shell Hartree–Fock theory

- Consider a case where  $n_c$  spatial orbitals are doubly occupied with perfect spin pairing —the closed shells— and  $n_o$  are singly occupied, all with  $\alpha$  spins, the open shells:

$$\Phi = |\varphi_{1\alpha}\varphi_{1\beta}\varphi_{2\alpha}\varphi_{2\beta} \cdots \varphi_{n_c\alpha}\varphi_{n_c\beta}\varphi_{(n_c+1)\alpha}\varphi_{(n_c+2)\alpha} \cdots \varphi_{(n_c+n_o)\alpha}|$$

- This SD is eigenfunction of  $\hat{S}_z$  with eigenvalue  $M_S = \frac{n_o}{2}$ .
- Furthermore,  $\hat{S}_+\Phi = 0$ , because raising the spin is either not possible (when the orbital has already  $m_s = 1/2$ ) or creates an orbital that is already contained in the SD.
- Hence,  $\hat{S}^2\Phi = \left(\hat{S}_-\hat{S}_+ + \hat{S}_z^2 + \hat{S}_z\right)\Phi = \frac{n_o}{2}\left(\frac{n_o}{2} + 1\right)\Phi \Rightarrow S = \frac{n_o}{2}$
- The ROHF determinant is an eigenfunction of  $\hat{S}^2$  with quantum number  $S = \frac{n_o}{2}$ .
- An high-spin open shell.

# Restricted open-shell Hartree–Fock theory

- We shall use indices  $i, j, k, \dots$  for doubly occupied orbitals and  $s, t, u, \dots$  for singly occupied orbitals.
- The energy can then be written as

$$\begin{aligned} E[\Phi] = & V_{nn} + 2 \sum_i h_{ii} + \sum_s h_{ss} + \sum_{i,j} [2\langle ij|ij\rangle - \langle ij|ji\rangle] \\ & + \sum_{i,t} [2\langle it|it\rangle - \langle it|ti\rangle] + \frac{1}{2} \sum_{s,t} [\langle st|st\rangle - \langle st|ts\rangle] \end{aligned}$$

- It may also be written as ( $J_{pq} = \langle pq|pq\rangle$ ,  $K_{pq} = \langle pq|qp\rangle$ ):

$$\begin{aligned} E[\Phi] = & f \left[ 2 \sum_s h_{ss} + f \sum_{s,t} (2aJ_{st} - bK_{st}) + 2 \sum_{i,t} (2J_{it} - K_{it}) \right] \\ & + V_{nn} + 2 \sum_i h_{ii} + \sum_{i,j} (2J_{ij} - K_{ij}) \end{aligned}$$

## Roothaan parameters $a$ and $b$

$$E[\Phi] = f \left[ 2 \sum_s h_{ss} + f \sum_{s,t} (2aJ_{st} - bK_{st}) + 2 \sum_{i,t} (2J_{it} - K_{it}) \right] + \dots$$

- $f$  is the occupancy of the open shell ( $f = 1$  means complete occupancy and  $f = \frac{1}{2}$  refers to a half-filled shell (or MO).
- For the closed-shell case,  $f = 1$  and  $a = b = 1$ .
- For all *single-electron* cases,  $a = b = 0$ . The fractional occupation  $f$  depends on the degeneracy of the shell (e.g.,  $f = 1/10$  in a  $d^1$  configuration).
- For high-spin half-filled shells (e.g., the N atom's  $^4S(p^3)$  state),  $f = \frac{1}{2}$  and  $a = 1, b = 2$ .
- Roothaan has shown that the above energy expression can be used for atoms with  $p^n$  shells and linear molecules with  $\pi^n$  shells. The method fails for most  $d^n$  configurations.

## Closed- and open-shell Fock operators

- The variation method leads to the closed- and open-shell Fock operators

$$\hat{F}^c = \hat{h} + \sum_i (2\hat{J}_i - \hat{K}_i) + \textcolor{blue}{f} \sum_s (2\hat{J}_s - \hat{K}_s)$$

$$\hat{F}^o = \hat{h} + \sum_i (2\hat{J}_i - \hat{K}_i) + \textcolor{blue}{f} \sum_s (2\textcolor{red}{a}\hat{J}_s - \textcolor{red}{b}\hat{K}_s)$$

- The Hartree–Fock equations can be written as

$$\langle \varphi_a | \hat{F}^c | \varphi_i \rangle = 0, \quad \langle \varphi_a | \hat{F}^o | \varphi_s \rangle = 0, \quad \langle \varphi_i | \hat{F}^c - \textcolor{blue}{f} \hat{F}^o | \varphi_s \rangle = 0$$

- ROHF theory is a complex field, which requires a careful reading of the literature and program manuals (e.g., Turbomole).
- For example, would you have known that  $\textcolor{red}{a} = \textcolor{red}{b} = 8/9$  for the  $\text{O}_2^-$  radical?

## Example: the $^3P$ ground state of the O atom

- Calculations performed in the aug-cc-pVQZ basis set.
- Calculation with **Turbomole** in  $T_d$  symmetry:

```
$ closed shells
```

```
a1 1-2 ( 2 )
```

```
$ open shells type=1
```

```
t2 1 ( 4/3 )
```

```
$ roothaan 1
```

```
a = 15/16      b = 9/8
```

- ROHF ground-state energy:  $E = -74.808\,078\,382\,E_h$

## Example: the $^3P$ ground state of the O atom

- Next, switch to  $C_1$  symmetry and copy the ROHF orbitals to the file containing the  $\alpha$  spin orbitals as well as to the file containing the  $\beta$  spin orbitals.
- Evaluation of the UHF expectation value using these spin orbitals gives:

$$\langle \Phi | \hat{H} | \Phi \rangle = -74.808\,078\,382\, E_h$$

- This results could also be obtained by running a state-averaged CASSCF calculation in  $D_{2h}$  symmetry.
- Molpro input:  

```
{casscf;closed,2;wf,8,4,2;wf,8,6,2;wf,8,7,2}
```
- (State average of the  $^3B_{1g}$ ,  $^3B_{2g}$ , and  $^3B_{3g}$  states.)

# Self-consistent field theory: The LCAO expansion

- *Brute-force* numerical methods can be used to solve the Hartree–Fock equations for (di)atoms (by finite differences and finite elements methods, multi-grid techniques).
- The Hartree–Fock equations are **one-electron** equations and one may think that the problem is **three-dimensional**.
- However, the **nonlocal exchange** contribution makes the problem of solving the Hartree–Fock equations for a general molecule **six-dimensional**.
- This is why certain DFT programs have no “hybrid” functionals implemented, which include Hartree–Fock exchange. Computationally,  $\hat{K}$  is much more complicated than  $\hat{J}$ .
- In the following, we shall expand the MOs in a basis of atomic orbitals (AOs),

$$\varphi_k^\sigma(\mathbf{r}) = \sum_{\mu=1}^{N_{\text{bas}}} \chi_\mu(\mathbf{r}) C_{\mu i}^\sigma$$

# The Roothaan–Hall equations

- To simplify the derivation of the Roothaan–Hall equations as much as possible, we
  - ... assume that **the MOs are real** (and hence, also the MO coefficients  $C_{\mu i}^\sigma$ ),
  - ... ignore spin (but shall return to it later), and
  - ... use the **Mulliken notation** for two-electron integrals:

$$\langle \mu\rho | \nu\tau \rangle = \int \int \chi_\mu(\mathbf{r}) \chi_\rho(\mathbf{r}') \frac{1}{|\mathbf{r} - \mathbf{r}'|} \chi_\nu(\mathbf{r}) \chi_\tau(\mathbf{r}') d\mathbf{r} d\mathbf{r}' = (\mu\nu | \rho\tau)$$

Obviously:

$$\begin{aligned} \sum_{i=1}^n \langle \varphi_i | \hat{h} | \varphi_i \rangle &= \sum_{i=1}^n \sum_{\mu=1}^{N_{\text{bas}}} C_{\mu i} \langle \chi_\mu | \hat{h} | \varphi_i \rangle = \sum_{i=1}^n \sum_{\mu=1}^{N_{\text{bas}}} \sum_{\nu=1}^{N_{\text{bas}}} C_{\mu i} \langle \chi_\mu | \hat{h} | \chi_\nu \rangle C_{\nu i} \\ &= \sum_{\mu=1}^{N_{\text{bas}}} \sum_{\nu=1}^{N_{\text{bas}}} D_{\mu\nu} \langle \chi_\mu | \hat{h} | \chi_\nu \rangle = \sum_{\mu=1}^{N_{\text{bas}}} \sum_{\nu=1}^{N_{\text{bas}}} D_{\mu\nu} h_{\mu\nu} \end{aligned}$$



# The Roothaan–Hall equations

- We have introduced the **density matrix  $\mathbf{D}$** , where the summation is over all (spin) orbitals,

$$D_{\mu\nu} = \sum_{i=1}^n C_{\mu i} C_{\nu i}, \quad \text{or} \quad \mathbf{D} = \mathbf{C} \mathbf{C}^T$$

- By inserting the LCAO expansion also into the two-electron energy contribution, we obtain (here and in the following, we omit the summation limits):

$$E[\mathbf{C}] = V_{nn} + \sum_{\mu,\nu} D_{\mu\nu} h_{\mu\nu} + \frac{1}{2} \sum_{\mu,\nu} \sum_{\rho,\tau} [D_{\mu\nu} D_{\rho\tau} - D_{\mu\rho} D_{\nu\tau}] (\mu\nu|\rho\tau)$$

- The orbitals are orthonormal,

$$\mathbf{C}^T \mathbf{S} \mathbf{C} = \mathbf{1}, \quad \text{with} \quad S_{\mu\nu} = \langle \chi_\mu | \chi_\nu \rangle$$

# The Lagrangian

- We are interested in finding the MO coefficients  $\mathbf{C}$  that minimise the SCF energy  $E(\mathbf{C})$  in the sense of the variation method, but we cannot simply require the first derivatives of  $E(\mathbf{C})$  to be zero for all  $C_{\mu i}$ .
- Rather, we must minimise  $E(\mathbf{C})$  under the constraint that the MOs remain orthonormal. This can be done using **Lagrange's method of undetermined multipliers**,

$$F[\mathbf{C}, \boldsymbol{\lambda}] = E[\mathbf{C}] + \text{Tr} [(\mathbf{1} - \mathbf{C}^T \mathbf{S} \mathbf{C}) \boldsymbol{\lambda}]$$

- Note that the unit matrix  $\mathbf{1}$  and the matrix of Lagrange multipliers  $\boldsymbol{\lambda}$  are  $n \times n$  matrices while  $\mathbf{S}$  is a  $N_{\text{bas}} \times N_{\text{bas}}$  matrix. The MO-coefficients matrix  $\mathbf{C}$  has  $n_{\text{bas}}$  rows and  $n$  columns.

$$\mathbf{1} = \mathbf{C}^T \times \mathbf{S} \times \mathbf{C}$$

Diagram illustrating the matrix dimensions for the Lagrangian equation. Matrix  $\mathbf{1}$  is a small square. Matrix  $\mathbf{C}^T$  is a horizontal rectangle. Matrix  $\mathbf{S}$  is a large square with a double-headed arrow below it labeled  $N_{\text{bas}}$ . Matrix  $\mathbf{C}$  is a vertical rectangle with a double-headed arrow below it labeled  $n$ . Multiplication symbols ( $\times$ ) are placed between the matrices.

## Minimising the Lagrangian

- We require the Lagrange functional  $F[\mathbf{C}, \boldsymbol{\lambda}]$  to be stationary with respect to the coefficients  $\mathbf{C}$  and multipliers  $\boldsymbol{\lambda}$ ,

$$\frac{\partial F[\mathbf{C}, \boldsymbol{\lambda}]}{\partial C_{\omega i}} = 0 \quad \forall C_{\omega i}, \quad \frac{\partial F[\mathbf{C}, \boldsymbol{\lambda}]}{\partial \lambda_{kl}} = 0 \quad \forall \lambda_{kl} \Leftrightarrow \mathbf{1} - \mathbf{C}^T \mathbf{S} \mathbf{C} = \mathbf{0}$$

- While the equation to the right is just the normalisation condition, the equations to the left are the Roothaan–Hall equations.
- The energy gradient is easily computed by taking partial derivatives with respect to  $D_{\kappa\lambda}$ ,

$$\begin{aligned} \frac{\partial E[\mathbf{C}]}{\partial C_{\omega i}} &= \sum_{\kappa, \lambda} \frac{\partial E[\mathbf{C}]}{\partial D_{\kappa\lambda}} \frac{\partial D_{\kappa\lambda}}{\partial C_{\omega i}} = 2 \sum_{\lambda} \frac{\partial E[\mathbf{C}]}{\partial D_{\omega\lambda}} C_{\lambda i} \\ &= 2 \sum_{\lambda} h_{\omega\lambda} C_{\lambda i} + \text{two-electron terms} \\ &= 2(\mathbf{hC})_{\omega i} + \text{two-electron terms} \end{aligned}$$

# Minimising the Lagrangian

- Similarly, we obtain for the orthonormality constraint,

$$\frac{\partial \text{Tr} [(1 - \mathbf{C}^T \mathbf{S} \mathbf{C}) \boldsymbol{\lambda}]}{\partial C_{\omega i}} = -2 (\mathbf{S} \mathbf{C} \boldsymbol{\lambda})_{\omega i}$$

- Hence,

$$\begin{aligned} \frac{\partial F[\mathbf{C}, \boldsymbol{\lambda}]}{\partial C_{\omega i}} &= 2(\mathbf{h} \mathbf{C})_{\omega i} - 2(\mathbf{S} \mathbf{C} \boldsymbol{\lambda})_{\omega i} + \text{two-electron terms} \\ &= 2(\mathbf{F} \mathbf{C})_{\omega i} - 2(\mathbf{S} \mathbf{C} \boldsymbol{\lambda})_{\omega i} = 0 \quad \forall C_{\omega i} \end{aligned}$$

The two-electron terms add to the one-electron Hamiltonian to yield the Fock operator. The Roothaan–Hall equations thus read:

$$\boxed{\mathbf{F}} \times \boxed{\mathbf{C}} = \boxed{\mathbf{S}} \times \boxed{\mathbf{C}} \times \boxed{\boldsymbol{\lambda}}$$

$$\mathbf{F} \mathbf{C} = \mathbf{S} \mathbf{C} \boldsymbol{\lambda}$$

# Canonical Roothaan–Hall equations

- The Roothaan–Hall equations in the LCAO Ansatz are nothing but the translation of the coupled integro-differential Hartree–Fock equations into the language of standard linear algebra,

$$\hat{F}\varphi_i = \sum_m \varphi_m \lambda_{mi} \Leftrightarrow \mathbf{FC} = \mathbf{SC}\boldsymbol{\lambda}$$

- As for Hartree–Fock, orthogonal transformations of the MOs (*i.e.*, columns of  $\mathbf{C}$ ) do neither change the energy  $E[\mathbf{C}]$  nor the Lagrangian  $F[\mathbf{C}, \boldsymbol{\lambda}]$ .
- Hence, we can choose the orbitals such that  $\boldsymbol{\lambda}$  becomes the diagonal matrix  $\boldsymbol{\varepsilon}$ . These orbitals are the **canonical MOs**,

$$\hat{F}\varphi_i = \varphi_i \varepsilon_i \Leftrightarrow \mathbf{FC} = \mathbf{SC}\boldsymbol{\varepsilon} \Rightarrow \mathbf{FC}_i = \mathbf{SC}_i \varepsilon_i$$

( $\mathbf{C}_i$  is the  $i$ -th column of the matrix  $\mathbf{C}$ ).

## An orthonormal basis

- The Roothaan–Hall equation is a generalized matrix eigenvalue equation. To solve it, it is convenient to bring it on a conventional matrix eigenvalue form, that is, without the overlap matrix  $S$ .
- Let  $U$  be a real, nonsingular matrix:  $UU^{-1} = U^{-1}U = \mathbf{1}$ . We can insert  $UU^{-1}$  into the Roothaan–Hall equation and obtain:

$$\begin{aligned} \mathbf{FC} = \mathbf{SC}\epsilon &\Leftrightarrow (\mathbf{FU})(U^{-1}\mathbf{C}) = (\mathbf{SU})(U^{-1}\mathbf{C})\epsilon \\ &\Leftrightarrow (U^T\mathbf{FU})(U^{-1}\mathbf{C}) = (U^T\mathbf{SU})(U^{-1}\mathbf{C})\epsilon \end{aligned}$$

- The matrix  $U$  can be chosen such that  $U^T\mathbf{SU} = \mathbf{1}$ . We then obtain

$$\tilde{\mathbf{F}}\tilde{\mathbf{C}}_i = \tilde{\mathbf{C}}_i\epsilon_i, \quad \text{with } \tilde{\mathbf{F}} = U^T\mathbf{FU} \text{ and } \tilde{\mathbf{C}} = U^{-1}\mathbf{C}$$

- The matrix  $U$  defines an **orthonormal basis**  $\{\phi_\tau\}$ , expanded in our original basis  $\{\chi_\mu\}$ ,

$$\phi_\tau(\mathbf{r}) = \sum_{\mu=1}^{N_{\text{bas}}} \chi_\mu(\mathbf{r}) U_{\mu\tau}$$

## An orthonormal basis

- Of course, the final MOs are still expanded in the original basis,

$$\varphi_i(\mathbf{r}) = \sum_{\tau=1}^{N_{\text{bas}}} \phi_{\tau}(\mathbf{r}) \tilde{C}_{\tau i} = \sum_{\tau=1}^{N_{\text{bas}}} \sum_{\mu=1}^{N_{\text{bas}}} \chi_{\mu}(\mathbf{r}) U_{\mu\tau} \tilde{C}_{\tau i} = \sum_{\mu=1}^{N_{\text{bas}}} \chi_{\mu}(\mathbf{r}) C_{\mu i}$$

- Hence, we see that  $\mathbf{C} = \mathbf{U}\tilde{\mathbf{C}}$ .
- During the SCF iterations (*vide infra*), we can simply use the MO coefficients from the previous iteration,  $\mathbf{C}^{(k-1)}$  as an orthonormal basis and compute

$$\tilde{\mathbf{F}} = [\mathbf{C}^{(k-1)}]^T \mathbf{F} \mathbf{C}^{(k-1)}$$

The new coefficients are then obtained as  $\mathbf{C}^{(k)} = \mathbf{C}^{(k-1)} \tilde{\mathbf{C}}$ .

- In actual calculations, an orthonormal basis is constructed before starting the SCF iterations. How should we start?

## An orthonormal basis

- In the first SCF iteration, we can orthogonalise the AOs using the **Gram–Schmidt** approach: The first AO is taken as is. The second AO is orthogonalised against the first, the third against the first two, the fourth against the first three, and so on.
- Alternatively, we start by diagonalising the overlap matrix  $S$ ,

$$\mathbf{V}^T \mathbf{S} \mathbf{V} = \boldsymbol{\sigma}$$

with  $\boldsymbol{\sigma}$  diagonal and  $\mathbf{V}$  an orthogonal matrix. We then form the matrix

$$\mathbf{U} = \mathbf{V} \boldsymbol{\sigma}^{-\frac{1}{2}}, \quad \mathbf{U}^T \mathbf{S} \mathbf{U} = \boldsymbol{\sigma}^{-\frac{1}{2}} \mathbf{V}^T \mathbf{S} \mathbf{V} \boldsymbol{\sigma}^{-\frac{1}{2}} = \boldsymbol{\sigma}^{-\frac{1}{2}} \boldsymbol{\sigma} \boldsymbol{\sigma}^{-\frac{1}{2}} = \mathbf{1}$$

- This **canonical orthogonalisation** has the advantage that we can identify a near (or exact) linear dependency in the basis set (an eigenvalue  $\sigma_{\mu\mu} \approx 0$ ) and remove the corresponding column  $\mathbf{V}_\mu$  simply by setting  $\sigma_{\mu\mu}^{-1/2} = 0$ .



## An orthonormal basis

- It is also possible to define  $\mathbf{U}$  as

$$\mathbf{U} = \mathbf{S}^{-\frac{1}{2}} = \mathbf{V}\boldsymbol{\sigma}^{-\frac{1}{2}}\mathbf{V}^T, \quad \mathbf{U}^T\mathbf{S}\mathbf{U} = \mathbf{S}^{-\frac{1}{2}}\mathbf{S}\mathbf{S}^{-\frac{1}{2}} = \mathbf{1}$$

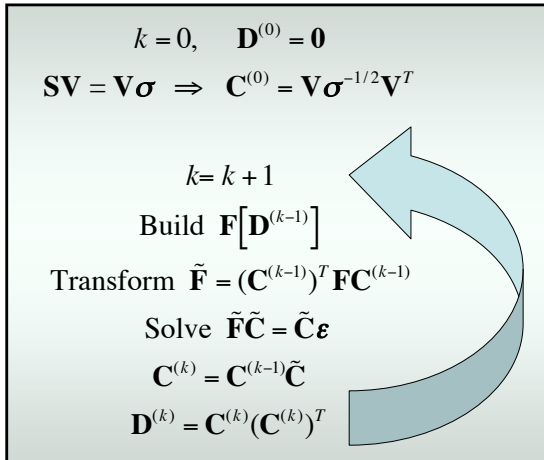
The advantage of this **symmetric orthogonalisation** is that the new basis is as close to the original set as possible.

- Although the number of small eigenvalues of  $\mathbf{S}$  may change along a potential curve and thus lead to unpleasant **discontinuities** in the energy, it often is a prerequisite to **post-Hartree–Fock** calculations (CI, CC) to remove eigenvalues smaller than, say,  $10^{-7}$  (relative to normalised AOs).
- A near linear dependence leads to MOs with large coefficients alternating in sign, and hence, significant **round-off errors** may occur in integral-transformation steps,

$$(pq|rs) = \sum_{\kappa} \sum_{\lambda} \sum_{\mu} \sum_{\nu} C_{\kappa p} C_{\lambda q} C_{\mu r} C_{\nu s} (\kappa\lambda|\mu\nu)$$

# Solving the Roothaan–Hall equations

The box below shows a simple algorithm for solving the Roothaan–Hall equations by the self-consistent-field procedure.



# Roothaan–Hall equations for UHF

- In UHF theory, the  $\alpha$  and  $\beta$  spin orbitals are expanded in the same AO basis,

$$\varphi_k^\alpha(\mathbf{r}) = \sum_{\mu=1}^{N_{\text{bas}}} \chi_\mu(\mathbf{r}) C_{\mu k}^\alpha, \quad \varphi_k^\beta(\mathbf{r}) = \sum_{\mu=1}^{N_{\text{bas}}} \chi_\mu(\mathbf{r}) C_{\mu k}^\beta$$

With these orbitals, a density matrix can be defined for each of the spins (assuming real MOs),

$$D_{\mu\nu}^\alpha = \sum_{i=1}^{n_\alpha} C_{\mu i}^\alpha C_{\nu i}^\alpha, \quad D_{\mu\nu}^\beta = \sum_{i=1}^{n_\beta} C_{\mu i}^\beta C_{\nu i}^\beta,$$

- We can now write the energy

$$E[\Phi] = V_{nn} + \sum_i h_{ii} + \sum_{\bar{i}} h_{\bar{i}\bar{i}} + \frac{1}{2} \sum_{i,j} \langle i j | i j \rangle + \frac{1}{2} \sum_{\bar{i}, \bar{j}} \langle \bar{i} \bar{j} | \bar{i} \bar{j} \rangle + \sum_{i, \bar{j}} \langle i \bar{j} | i \bar{j} \rangle$$

as a function of the density matrices  $D^\alpha$  and  $D^\beta$ .

# Roothaan–Hall equations for UHF

We obtain

$$\begin{aligned} E[\Phi] &= V_{nn} + \sum_{\mu\nu} [D_{\mu\nu}^{\alpha} + D_{\mu\nu}^{\beta}] h_{\mu\nu} \\ &+ \frac{1}{2} \sum_{\kappa\lambda\mu\nu} D_{\mu\nu}^{\alpha} D_{\kappa\lambda}^{\alpha} [(\mu\nu|\kappa\lambda) - (\mu\kappa|\nu\lambda)] \\ &+ \frac{1}{2} \sum_{\kappa\lambda\mu\nu} D_{\mu\nu}^{\beta} D_{\kappa\lambda}^{\beta} [(\mu\nu|\kappa\lambda) - (\mu\kappa|\nu\lambda)] + \sum_{\kappa\lambda\mu\nu} D_{\mu\nu}^{\alpha} D_{\kappa\lambda}^{\beta} (\mu\nu|\kappa\lambda) \end{aligned}$$

which can be rewritten using  $\mathbf{D} = \mathbf{D}^{\alpha} + \mathbf{D}^{\beta}$  and  $\mathbf{D}^{\text{spin}} = \mathbf{D}^{\alpha} - \mathbf{D}^{\beta}$ ,

$$\begin{aligned} E[\Phi] &= V_{nn} + \sum_{\mu\nu} D_{\mu\nu} h_{\mu\nu} + \frac{1}{4} \sum_{\kappa\lambda\mu\nu} D_{\mu\nu} D_{\kappa\lambda} [2(\mu\nu|\kappa\lambda) - (\mu\kappa|\nu\lambda)] \\ &- \frac{1}{4} \sum_{\kappa\lambda\mu\nu} D_{\mu\nu}^{\text{spin}} D_{\kappa\lambda}^{\text{spin}} (\mu\kappa|\nu\lambda) \end{aligned}$$

# Roothaan–Hall equations for UHF

- We also obtain two Fock matrices (in the AO basis),

$$F_{\mu\nu}^{\alpha} = h_{\mu\nu} + \frac{1}{2} \sum_{\kappa\lambda} \left\{ D_{\kappa\lambda} [2(\mu\nu|\kappa\lambda) - (\mu\kappa|\nu\lambda)] - D_{\kappa\lambda}^{\text{spin}} (\mu\kappa|\nu\lambda) \right\}$$

$$F_{\mu\nu}^{\beta} = h_{\mu\nu} + \frac{1}{2} \sum_{\kappa\lambda} \left\{ D_{\kappa\lambda} [2(\mu\nu|\kappa\lambda) - (\mu\kappa|\nu\lambda)] + D_{\kappa\lambda}^{\text{spin}} (\mu\kappa|\nu\lambda) \right\}$$

- We now obtain the following UHF Roothaan–Hall equations (including the closed-shell case):

$$\mathbf{F}^{\alpha} \mathbf{C}^{\alpha} = \mathbf{S} \mathbf{C}^{\alpha} \boldsymbol{\epsilon}^{\alpha}$$

$$\mathbf{F}^{\beta} \mathbf{C}^{\beta} = \mathbf{S} \mathbf{C}^{\beta} \boldsymbol{\epsilon}^{\beta}$$

$$\Leftrightarrow \begin{pmatrix} \mathbf{F}^{\alpha} & \mathbf{0} \\ \mathbf{0} & \mathbf{F}^{\beta} \end{pmatrix} \begin{pmatrix} \mathbf{C}^{\alpha} & \mathbf{0} \\ \mathbf{0} & \mathbf{C}^{\beta} \end{pmatrix} = \begin{pmatrix} \mathbf{S} & \mathbf{0} \\ \mathbf{0} & \mathbf{S} \end{pmatrix} \begin{pmatrix} \mathbf{C}^{\alpha} & \mathbf{0} \\ \mathbf{0} & \mathbf{C}^{\beta} \end{pmatrix} \begin{pmatrix} \boldsymbol{\epsilon}^{\alpha} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\epsilon}^{\beta} \end{pmatrix}$$

# Guess of initial orbitals

- In order to converge the SCF procedure quickly, it is important to have an **accurate guess of the orbitals for the first iteration** or an **accurate guess of the trial density**. There are several possibilities:

1. We start with  $\mathbf{D}^{(0)} = (\mathbf{D}^{\text{spin}})^{(0)} = \mathbf{0}$ . Electron repulsion is neglected in the first iteration (**bare nuclear Hamiltonian**).
2. We use the orbitals from **another calculation**, for example, a calculation carried out in a smaller AO basis,

$$\mathbf{D}_{\text{large basis}}^{(0)} = \tilde{\mathbf{S}}^T \mathbf{D}_{\text{small basis}} \tilde{\mathbf{S}}, \quad \tilde{S}_{\mu\nu} = \langle \chi_{\mu, \text{small}} | \chi_{\nu, \text{large}} \rangle$$

3. We use the density obtained from **extended Hückel theory** (EHT), projected onto the actual basis set.
4. We construct the trial density as **superposition of atomic densities**. This density can be diagonalised and the eigenvectors with near integer occupation numbers can be used as trial orbitals.

## Level shifting

- In order to converge the SCF procedure quickly (or to converge it at all), it may be needed to monitor the convergence process and to **manipulate the Fock matrix**,

$$\tilde{\mathbf{F}}\tilde{\mathbf{C}} = \tilde{\mathbf{C}}\boldsymbol{\varepsilon}, \quad \text{with } \tilde{\mathbf{F}} = [\mathbf{C}^{(k-1)}]^T \mathbf{F} \mathbf{C}^{(k-1)} \quad \text{and} \quad \mathbf{C}^{(k)} = \mathbf{C}^{(k-1)} \tilde{\mathbf{C}}$$

- Near convergence,  $\tilde{\mathbf{C}}$  will be close to unity,  $\tilde{\mathbf{C}} = \mathbf{1} + \boldsymbol{\omega}$ . The values of  $\boldsymbol{\omega}$  can be estimated from first-order perturbation theory,

$$\omega_{ia} = \frac{\tilde{F}_{ia}}{\tilde{F}_{aa} - \tilde{F}_{ii}} = \frac{\tilde{F}_{ia}}{\varepsilon_a^{(k-1)} - \varepsilon_i^{(k-1)}}$$

- If there is a small HOMO–LUMO gap (in the previous iteration), then the corresponding  $\omega_{ia}$  is large and can cause oscillations.
- We can stabilise the iterations by adding a **level shift**  $\Delta > 0$  to the diagonal elements of the virtuals,  $\tilde{F}_{aa} \rightarrow \tilde{F}_{aa} + \Delta$ .

# Damping

- Another way to reduce  $\omega_{ia}$  (also when  $\Delta = 0$ ) is by reducing the off-diagonal elements  $\tilde{F}_{ia}$ ,

$$\omega_{ia} = \frac{\tilde{F}_{ia}}{\tilde{F}_{aa} - \tilde{F}_{ii} + \Delta}, \quad \tilde{F}_{ia} \rightarrow \alpha \tilde{F}_{ia}, \quad 0 < \alpha < 1$$

- This can for example be achieved by adding the Fock matrix of the previous iteration (which is diagonal) with a certain weight (e.g.,  $\beta = 0.5$ ) to the present Fock matrix,

$$\tilde{F}_{ii} \rightarrow \tilde{F}_{ii} + \beta \varepsilon_i^{(k-1)}, \quad \tilde{F}_{aa} \rightarrow \tilde{F}_{aa} + \beta \varepsilon_a^{(k-1)}$$

- Level shifting and damping can be combined. Level shifting reduces the mixing of high-lying occupied MOs with low-lying virtual MOs. Damping reduces *all mixing*.
- Neither of the techniques help to accelerate **creepingly slow SCF procedures**.



# Fermi smearing

- The idea of Fermi smearing is to compute density matrices  $\mathbf{D}$  and  $\mathbf{D}^{\text{spin}}$  from **partially occupied orbitals**. These density matrices are then used to build Fock matrices, to obtain new orbitals, and so on.
- Fractional occupation numbers  $n_p \in [0, 1]$  are computed as a function of a temperature  $T$  when the HOMO–LUMO gap drops below a given threshold ( $\mu$  is the Fermi level),

$$n_p = \frac{1}{2} \operatorname{erfc} \left( \frac{\epsilon_p - \mu}{fT} \right), \quad f = 4k_B / \sqrt{\pi}$$

- When the energy has converged to within a given threshold, the Fermi smearing is switched off ( $T = 0$ ).
- Calculation of fractional occupation numbers often requires much higher damping and level shifting.
- Mainly used in DFT self-consistent field calculations.

# Pulay's DIIS procedure

- Let us be concerned with a general iterative procedure in which we try to solve

$$\mathbf{f}[\mathbf{x}] = \mathbf{0}$$

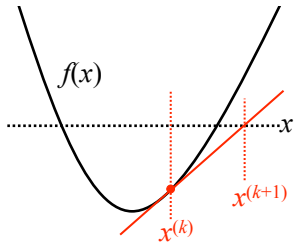
- Examples include:
  1. **Geometry optimizations**, where  $\mathbf{x}$  is the geometry and  $\mathbf{f}[\mathbf{x}]$  the gradient. If not yet converged in iteration step number  $k$ , we have

$$\mathbf{f}[\mathbf{x}^{(k)}] = \mathbf{g}^{(k)} \neq \mathbf{0}, \quad \mathbf{g}^{(k)} \Rightarrow \mathbf{x}^{(k+1)}$$

2. **Coupled-cluster calculations**, where  $\mathbf{x}$  contains the amplitudes,  $\mathbf{f}[\mathbf{x}]$  are the amplitudes equations and  $\mathbf{g}^{(k)}$  the vector functions.
3. **SCF calculations**, where

$$\mathbf{F}[\mathbf{D}^{(k)}]\mathbf{D}^{(k)}\mathbf{S} - \mathbf{S}\mathbf{D}^{(k)}\mathbf{F}[\mathbf{D}^{(k)}] = \mathbf{g}^{(k)}$$

# Pulay's DIIS procedure



- As an example of Pulay's DIIS method, we apply it to a simple Newton–Raphson procedure to solve  $f(x) = 0$ . From a given  $x^{(k)}$ , we obtain a new  $x^{(k+1)}$  via

$$x^{(k+1)} = x^{(k)} - \frac{f(x^{(k)})}{f'(x^{(k)})}$$

- After having obtained a series of  $n$  approximations  $x^{(1)}, x^{(2)}, \dots, x^{(n)}$ , we construct the linear combination

$$\tilde{x} = \sum_{k=1}^n d_k x^{(k)}, \quad \text{with the constraint} \quad \sum_{k=1}^n d_k = 1$$

- The  $d_k$ 's are obtained from requiring that  $\|f(\tilde{x})\|^2 = \min$ .

# Pulay's DIIS procedure

- To compute  $f(\tilde{x})$ , we approximate the function as

$$f(\tilde{x}) \approx \sum_{k=1}^n d_k f(x^{(k)}),$$

- Minimising the appropriate Lagrangian leads to

$$\begin{pmatrix} f(x^{(1)})f(x^{(1)}) & \dots & f(x^{(1)})f(x^{(n)}) & -1 \\ f(x^{(2)})f(x^{(1)}) & \dots & f(x^{(2)})f(x^{(n)}) & -1 \\ \vdots & & \vdots & \vdots \\ f(x^{(n)})f(x^{(1)}) & \dots & f(x^{(n)})f(x^{(n)}) & -1 \\ 1 & \dots & 1 & 0 \end{pmatrix} \begin{pmatrix} d_1 \\ d_2 \\ \vdots \\ d_n \\ \lambda \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$

- In SCF theory,  $f(x^{(k)})f(x^{(l)}) \rightarrow \text{Tr}([\mathbf{g}^{(k)}]^T \mathbf{g}^{(l)})$ , and the  $d_k$ 's are used to extrapolate directly the Fock matrix.

# The supermatrix formalism

- We consider closed-shell Hartree–Fock theory with  $\mathbf{D}^\alpha = \mathbf{D}^\beta$ , and hence,  $\mathbf{D} = 2\mathbf{D}^\alpha$  and  $\mathbf{D}^{\text{spin}} = \mathbf{0}$ ,

$$E[\Phi] = V_{nn} + \sum_{\mu\nu} D_{\mu\nu} h_{\mu\nu} + \frac{1}{2} \sum_{\mu\nu\kappa\lambda} D_{\mu\nu} D_{\kappa\lambda} [(\mu\nu|\kappa\lambda) - \frac{1}{2}(\mu\kappa|\nu\lambda)]$$

- The Fock operator is given by

$$F_{\mu\nu} = h_{\mu\nu} + \frac{1}{2} \sum_{\kappa\lambda} D_{\kappa\lambda} [(\mu\nu|\kappa\lambda) - \frac{1}{2}(\mu\kappa|\nu\lambda)]$$

- Because the summation is over all  $\kappa, \lambda$ , we may split the last integral into two terms, that is, symmetrise its contribution,

$$F_{\mu\nu} = h_{\mu\nu} + \frac{1}{2} \sum_{\kappa\lambda} D_{\kappa\lambda} [(\mu\nu|\kappa\lambda) - \frac{1}{4}(\mu\kappa|\nu\lambda) - \frac{1}{4}(\mu\lambda|\nu\kappa)]$$

and similarly in the energy expression.

# The supermatrix formalism

- Because now all contributions are symmetric, we may restrict the summation range. We do so after redefining the density matrix as  $d_{\mu\nu} = D_{\mu\nu}(2 - \delta_{\mu\nu})$  and introducing the **supermatrix**

$$P_{\mu\nu,\kappa\lambda} = (\mu\nu|\kappa\lambda) - \frac{1}{4}(\mu\kappa|\nu\lambda) - \frac{1}{4}(\mu\lambda|\nu\kappa)$$

- We obtain for the energy and Fock operator:

$$\begin{aligned} E[\Phi] &= V_{nn} + \sum_{\mu \leq \nu} d_{\mu\nu} h_{\mu\nu} + \frac{1}{2} \sum_{\mu \leq \nu} \sum_{\kappa \leq \lambda} d_{\mu\nu} P_{\mu\nu,\kappa\lambda} d_{\kappa\lambda} \\ F_{\mu\nu} &= h_{\mu\nu} + \sum_{\kappa \leq \lambda} P_{\mu\nu,\kappa\lambda} d_{\kappa\lambda} \end{aligned}$$

- The supermatrix formalism is very time saving and allows to write the theory in the short form (with **supervectors**  $\mathbf{d}$ ,  $\mathbf{h}$ ,  $\mathbf{f}$ )

$$E = V_{nn} + \mathbf{d}^T \mathbf{h} + \frac{1}{2} \mathbf{d}^T \mathbf{P} \mathbf{d}, \quad \mathbf{f} = \mathbf{h} + \mathbf{P} \mathbf{d}$$