Unitary transformations in single-determinant electronic structure theory

Ivo Filot / i.a.w.filot@tue.nl

Table of Contents

Section

Course contents

Prerequisite knowledge

- Linear algebra
    - Matrix diagonalization, eigenvalues, eigenvectors
    - Matrix multiplication algorithmic
    - Basis set transformations
- Basis quantum chemistry
    - Basis sets
    - Linear combination of molecular orbitals
- Basic programming skills
    - For and while loops
    - Vector containers
    - Functions, variables, encapsulation

Ingredients of the Hartree-Fock algorithm

- The Schrödinger equation
- Born-Oppenheimer approximation
- Slater determinant
- Minimization principle
- Gaussian basis set

Section

Hartree-Fock equations

The Schrödinger equation

We start with the (in)famous time-independent non-relativistic Schrödinger equation

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

where the Hamiltonian operator $\hat{\mathcal{H}}$ is given (in atomic units) by

$$\hat{\mathcal{H}} = \underbrace{-\sum_{i=1}^{N}\frac{1}{2}\nabla_i^2}_{E_{\text{kin}}\text{electrons}} \underbrace{-\sum_{A=1}^{M}\frac{1}{2M_A}\nabla_A^2}_{E_{\text{kin}}\text{nuclei}} \underbrace{-\sum_{i=1}^{N}\sum_{A=1}^{M}\frac{Z_A}{r_{iA}}}_{\text{nuc-el attraction}} + \underbrace{\sum_{i=1}^{N}\sum_{j>i}^{N}\frac{1}{r_{ij}}}_{\text{el-el repulsion}} + \underbrace{\sum_{A=1}^{M}\sum_{A>B}^{M}\frac{Z_A Z_B}{R_{AB}}}_{\text{nuc-nuc repulsion}}$$

Born-Oppenheimer approximation

Nuclei are much heavier than electrons, hence we consider the motion of the electrons to occur under the static presence of the nuclei.

$$\hat{\mathcal{H}} = -\sum_{i=1}^{N} \frac{1}{2}\nabla_i^2 - \cancel{\sum_{A=1}^{M} \frac{1}{2M_A}\nabla_A^2} - \sum_{i=1}^{N}\sum_{A=1}^{M} \frac{Z_A}{r_{iA}} + \sum_{i=1}^{N}\sum_{j>i}^{N} \frac{1}{r_{ij}} + \cancel{\sum_{A=1}^{M}\sum_{A>B}^{M} \frac{Z_A Z_B}{R_{AB}}}$$

by which the result is termed the electronic Hamiltonian

$$\hat{\mathcal{H}}_{\text{elec}} = -\sum_{i=1}^{N} \frac{1}{2}\nabla_i^2 - \sum_{i=1}^{N}\sum_{A=1}^{M} \frac{Z_A}{r_{iA}} + \sum_{i=1}^{N}\sum_{j>i}^{N} \frac{1}{r_{ij}}$$

Note that the nucleus-nucleus repulsion is still used, but only as a static term to the total electronic energy.

Single-electron wave functions

To describe the location and wave-like behavior of a single electron, we use a spatial single-electron wave function

$$\phi(\vec{r})$$

which satisfies a number of properties such as

$$\langle \phi(\vec{r}) | \phi(\vec{r}) \rangle = 1.$$

Spin-orbitals

The spatial orbital can be combined with an (abstract) **spin** function to obtain

$$\chi(\vec{x}) = \phi(\vec{r}) \cdot \begin{cases} \alpha(\omega) \\ \beta(\omega) \end{cases}$$

wherein the spin function obeys the canonical orthonormality relationship

$$\langle \alpha | \alpha \rangle = \langle \beta | \beta \rangle = 1$$

and

$$\langle \alpha | \beta \rangle = \langle \beta | \alpha \rangle = 0$$

Wave function

We have a system of multiple electrons, so we need a wave function that can handle that. A naive implementation is to simply consider a product of one-electron wave functions

$$\Psi_{\text{Hartree}} = \chi_i\left(\vec{\mathbf{x}}_1\right)\chi_j\left(\vec{\mathbf{x}}_2\right)\cdots\chi_N\left(\vec{\mathbf{x}}_N\right)$$

which is called a Hartree product. Assuming that the spinorbitals form an orthonormal set

$$\left\langle\chi_i\left(\vec{\mathbf{x}}_1\right)\middle|\chi_j\left(\vec{\mathbf{x}}_1\right)\right\rangle = \delta_{ij}$$

then

$$\hat{\mathcal{H}}_{\text{elec}}\Psi_{\text{Hartree}} = \left(\epsilon_1 + \epsilon_2 + \cdots + \epsilon_N\right)\Psi_{\text{Hartree}}$$

Problem with the Hartree product

The chance to find any of the electrons is given by

$$\left|\Psi\left(\vec{\mathbf{x}}_1, \vec{\mathbf{x}}_2, \cdots, \vec{\mathbf{x}}_N\right)\right|^2 \mathrm{d}\vec{\mathbf{x}}_1 \mathrm{d}\vec{\mathbf{x}}_2 \cdots \mathrm{d}\vec{\mathbf{x}}_N = \left|\chi_i\left(\vec{\mathbf{x}}_1\right)\right|^2 \mathrm{d}\vec{\mathbf{x}}_1 \left|\chi_j\left(\vec{\mathbf{x}}_2\right)\right|^2 \mathrm{d}\vec{\mathbf{x}}_2 \cdots \left|\chi_N\left(\vec{\mathbf{x}}_N\right)\right|^2 \mathrm{d}\vec{\mathbf{x}}_N$$

For each of the electrons, that chance does **not** depend on the position of the other electrons. In other words, electron motion is fundamentally uncorrelated in this wave function. That cannot be right!

Problem with the Hartree product

In reality, electrons will repel each other and thus the chance to find electron 1 at position $\vec{\mathbf{x}}_1$ depends on the position of, e.g., electron 2 at position $\vec{\mathbf{x}}_2$.

But there is an even more fundamental problem. The Hartree product allows us to distinguish between individual electrons. Electrons are fundamentally indistinguishable particles and hence the many-electron wave function should have this property embedded.

Slater determinant

Remember, the $\vec{\mathbf{x}}$ are the electrons and the χ_i are the spin-orbitals.

$$\left|\Psi\left(\vec{\mathbf{x}}_1, \vec{\mathbf{x}}_2, \cdots, \vec{\mathbf{x}}_N\right)\right\rangle = (N!)^{-1/2} \begin{vmatrix} \chi_i\left(\vec{\mathbf{x}}_1\right) & \chi_j\left(\vec{\mathbf{x}}_1\right) & \cdots & \chi_k\left(\vec{\mathbf{x}}_1\right) \\ \chi_i\left(\vec{\mathbf{x}}_2\right) & \chi_j\left(\vec{\mathbf{x}}_2\right) & \cdots & \chi_k\left(\vec{\mathbf{x}}_2\right) \\ \vdots & \vdots & \cdots & \vdots \\ \chi_i\left(\vec{\mathbf{x}}_N\right) & \chi_j\left(\vec{\mathbf{x}}_N\right) & \cdots & \chi_k\left(\vec{\mathbf{x}}_N\right) \end{vmatrix}$$

Slater determinant

$$\left|\Psi\left(\vec{\mathbf{x}}_1, \vec{\mathbf{x}}_2, \cdots, \vec{\mathbf{x}}_N\right)\right\rangle = (N!)^{-1/2} \begin{vmatrix} \chi_i\left(\vec{\mathbf{x}}_1\right) & \chi_j\left(\vec{\mathbf{x}}_1\right) & \cdots & \chi_k\left(\vec{\mathbf{x}}_1\right) \\ \chi_i\left(\vec{\mathbf{x}}_2\right) & \chi_j\left(\vec{\mathbf{x}}_2\right) & \cdots & \chi_k\left(\vec{\mathbf{x}}_2\right) \\ \vdots & \vdots & \cdots & \vdots \\ \chi_i\left(\vec{\mathbf{x}}_N\right) & \chi_j\left(\vec{\mathbf{x}}_N\right) & \cdots & \chi_k\left(\vec{\mathbf{x}}_N\right) \end{vmatrix}$$

- Satisfies anti-symmetry principle
- Electrons become (exchange-)correlated

Slater determinant

$$\left|\Psi\left(\vec{\mathbf{x}}_1, \vec{\mathbf{x}}_2, \cdots, \vec{\mathbf{x}}_N\right)\right\rangle = (N!)^{-1/2} \begin{vmatrix} \chi_i\left(\vec{\mathbf{x}}_1\right) & \chi_j\left(\vec{\mathbf{x}}_1\right) & \cdots & \chi_k\left(\vec{\mathbf{x}}_1\right) \\ \chi_i\left(\vec{\mathbf{x}}_2\right) & \chi_j\left(\vec{\mathbf{x}}_2\right) & \cdots & \chi_k\left(\vec{\mathbf{x}}_2\right) \\ \vdots & \vdots & \cdots & \vdots \\ \chi_i\left(\vec{\mathbf{x}}_N\right) & \chi_j\left(\vec{\mathbf{x}}_N\right) & \cdots & \chi_k\left(\vec{\mathbf{x}}_N\right) \end{vmatrix}$$

Important

This is still an uncorrelated wave function because electrons with opposite spin remain uncorrelated.

Combining Slater Determinant and Hamiltonian (1)

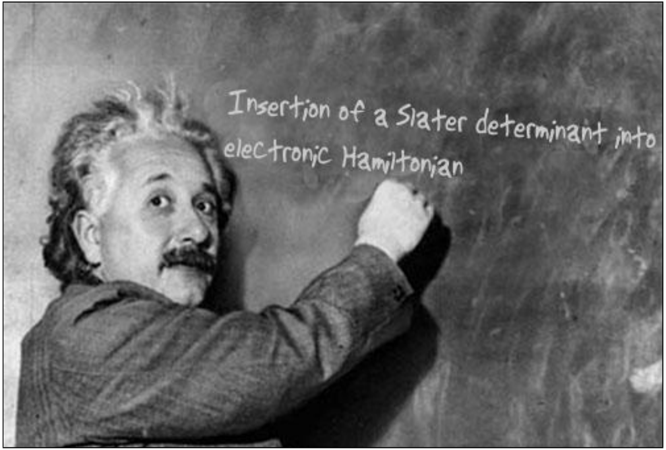
Inserting a Slater determinant

$$\left|\Psi\left(\vec{\mathbf{x}}_1, \vec{\mathbf{x}}_2, \cdots, \vec{\mathbf{x}}_N\right)\right\rangle = (N!)^{-1/2} \begin{vmatrix} \chi_i\left(\vec{\mathbf{x}}_1\right) & \chi_j\left(\vec{\mathbf{x}}_1\right) & \cdots & \chi_k\left(\vec{\mathbf{x}}_1\right) \\ \chi_i\left(\vec{\mathbf{x}}_2\right) & \chi_j\left(\vec{\mathbf{x}}_2\right) & \cdots & \chi_k\left(\vec{\mathbf{x}}_2\right) \\ \vdots & \vdots & \cdots & \vdots \\ \chi_i\left(\vec{\mathbf{x}}_N\right) & \chi_j\left(\vec{\mathbf{x}}_N\right) & \cdots & \chi_k\left(\vec{\mathbf{x}}_N\right) \end{vmatrix}$$

into

$$\hat{\mathcal{H}}_{\text{elec}} = -\sum_{i=1}^{N} \frac{1}{2}\nabla_i^2 - \sum_{i=1}^{N}\sum_{A=1}^{M} \frac{Z_A}{r_{iA}} + \sum_{i=1}^{N}\sum_{j>i}^{N} \frac{1}{r_{ij}}$$

Combining Slater Determinant and Hamiltonian: DEMO



Combining Slater Determinant and Hamiltonian (2)

$$E = \langle \Psi | \hat{\mathcal{H}} | \Psi \rangle = \underbrace{\sum_{i}^{N} \langle \chi_i | -\frac{1}{2}\nabla^2 | \chi_i \rangle}_{\text{kinetic energy}} + \underbrace{\sum_{i}^{N} \langle \chi_i | \sum_{A=1}^{M} \frac{Z_A}{r_{i,a}} | \chi_i \rangle}_{\text{nuclear attraction}} \cdots$$

$$\cdots + \underbrace{\frac{1}{2}\sum_{i}^{N} \langle \chi_i | \frac{1}{r_{ij}} \sum_{j \neq i}^{N} \langle \chi_j | \chi_j \rangle | \chi_i \rangle}_{\text{coulomb repulsion}} - \underbrace{\frac{1}{2}\sum_{i}^{N} \langle \chi_i | \frac{1}{r_{ij}} \sum_{j \neq i}^{N} \langle \chi_j | \chi_i \rangle | \chi_j \rangle}_{\text{exchange}}$$

Combining Slater Determinant and Hamiltonian (3)

An integral over sums is a sum over integrals and thus we can rewrite it to

$$E = \langle \Psi | \hat{\mathcal{H}} | \Psi \rangle = \underbrace{\sum_{i}^{N} \langle \chi_i | -\frac{1}{2}\nabla^2 | \chi_i \rangle}_{\text{kinetic energy}} + \underbrace{\sum_{i}^{N} \langle \chi_i | \sum_{A=1}^{M} \frac{Z_A}{r_{i,a}} | \chi_i \rangle}_{\text{nuclear attraction}} \cdots$$

$$\cdots + \frac{1}{2} \sum_{i}^{N} \sum_{j \neq i}^{N} \left[\underbrace{\langle \chi_i \chi_j | \frac{1}{r_{ij}} | \chi_i \chi_j \rangle}_{\text{coulomb repulsion}} - \underbrace{\langle \chi_i \chi_j | \frac{1}{r_{ij}} | \chi_j \chi_i \rangle}_{\text{exchange}} \right]$$

Combining Slater Determinant and Hamiltonian (4)

And by noting that

$$\langle \chi_i \chi_j | \frac{1}{r_{ij}} | \chi_i \chi_j \rangle - \langle \chi_i \chi_j | \frac{1}{r_{ij}} | \chi_j \chi_i \rangle = 0, \text{ when } i = j$$

$$E = \langle \Psi | \hat{\mathcal{H}} | \Psi \rangle = \underbrace{\sum_i^N \langle \chi_i | -\frac{1}{2}\nabla^2 | \chi_i \rangle}_{\text{kinetic energy}} + \underbrace{\sum_i^N \langle \chi_i | \sum_{A=1}^M \frac{Z_A}{r_{i,a}} | \chi_i \rangle}_{\text{nuclear attraction}} \cdots$$

$$\cdots + \frac{1}{2} \sum_i^N \sum_j^N \left[\underbrace{\langle \chi_i \chi_j | \frac{1}{r_{ij}} | \chi_i \chi_j \rangle}_{\text{coulomb repulsion}} - \underbrace{\langle \chi_i \chi_j | \frac{1}{r_{ij}} | \chi_j \chi_i \rangle}_{\text{exchange}} \right]$$

Interpretation of the exchange energy

> *Even though the Pauli exclusion principle is not really a contribution to the energy, the difference between the average energy you compute accounting for the Pauli exclusion principle, and the energy you compute without accounting for it, is called the exchange energy. It encodes the net effect of the Pauli exclusion principle as if it were a contribution to the energy, even though it is not, strictly speaking. The exchange energy is negative, because the actual energy is lower than what you get ignoring the exclusion principle.*

Source:
https://physics.stackexchange.com/questions/648950/what-is-the-exchange-energy-of-electrons

Section

Energy minimization

Energy minimization of single determinant

We are tasked to minimize the energy of a Slater determinant subject to the constraint that the spin-orbitals remain orthogonal.

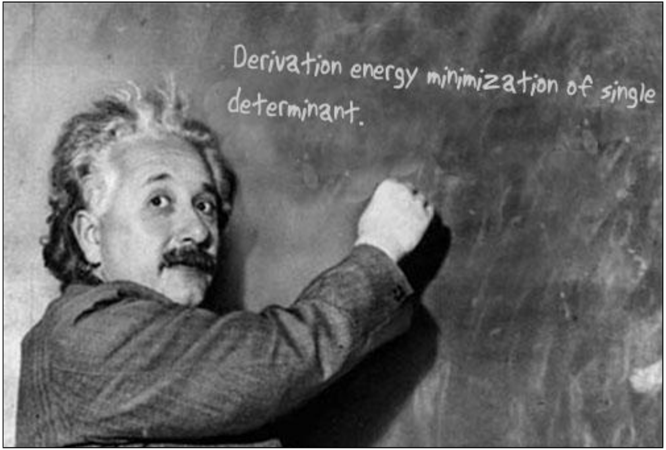
$$|\Psi\rangle = |\Psi\left(\vec{\mathbf{x}}_1, \vec{\mathbf{x}}_2, \cdots, \vec{\mathbf{x}}_N\right)\rangle$$

Minimization of a single-electron wave function which is a linear expansion of a set of known basis functions.

$$\min\left(\langle\Psi|\hat{\mathcal{H}}|\Psi\rangle\right) \text{ subject to the constraint } \langle\chi_i|\chi_j\rangle = \delta_{ij},$$

i.e., subject to the constraint that the spin-orbitals remains orthonormal.

The linear variational principle: DEMO



Uniqueness of the solutions

$$\hat{f} \left| \chi_i \right\rangle = \sum_{j=1}^{N} \epsilon_{ji} \left| \chi_j \right\rangle$$

which differs from what is written in 90% of the textbooks:

$$\hat{f} \left| \chi_i \right\rangle = \epsilon_i \left| \chi_i \right\rangle$$

Implication

- There is an infinite set of valid solutions to the Hartree-Fock equations.
- The spin-orbitals can be readily mixed among themselves without changing the expectation value $E = \langle \Psi_0 | \hat{\mathcal{H}} | \Psi_0 \rangle$.

Unitary transformations (1)

Let us fix this mess

$$\hat{f}\,|\chi_i\rangle = \sum_{j=1}^{N} \epsilon_{ji}\,|\chi_j\rangle$$

Introduce a **unitary matrix** that linearly transforms one set of spin-orbitals to another set:

$$|\chi_i\prime\rangle = \sum_j U_{ji}\,|\chi_j\rangle\,,$$

Recall that a unitary matrix has the following properties:

$$\mathbf{U}^\dagger = \mathbf{U}^{-1}$$

$$\mathbf{U}\mathbf{U}^{-1} = \mathbf{U}\mathbf{U}^\dagger = \mathbf{1}$$

Unitary transformations (2)

Invariants

- A unitary transformation does not affect the orthogonality among the spin-orbitals.
- Unitary transformations do not affect $\hat{\mathcal{J}}$ or $\hat{\mathcal{K}}$ as they purely depend on the sums of the spin-orbitals (i.e. the total spin-polarized electron density).
- By extension, the Fock operator $\hat{f}$ is also invariant under an arbitrary unitary transformation.

Unitary transformations (3)

$$\mathbf{E}' = \mathbf{U}^{\dagger}\mathbf{E}\mathbf{U}$$

Observation

Since $\mathbf{E}$ is a Hermitian matrix, it is by definition possible to find a unitary matrix $\mathbf{U}$ that diagonalizes $\mathbf{E}$.

Assuming that $\mathbf{E}$ is a diagonal matrix, then

$$\hat{f}\,|\chi_i\rangle = \epsilon_i\,|\chi_i\rangle$$

Reminder

A Hermitian matrix is a matrix which is equal to its conjugate transpose $(a_{ij} = a^*_{ji})$. For a real-valued matrix, this would imply that the matrix is symmetric.

Unitary transformations (4)

- Matrix diagonalization makes the spin-orbitals unique up to a phase factor $\exp(-i\phi)$ or up to ± 1 for real-valued coefficients.
- These solutions are termed the canonical solutions.
- The spin-orbitals will form an irreducible representation of the point group of the molecule which are maximally delocalized orbitals.
- Can be transformed, via a unitary transformation, to more localized orbitals.

Introduction known basis set

$$\hat{f}\left|\chi_i\right\rangle = \epsilon_i \left|\chi_i\right\rangle \text{ with } \left|\chi_i\right\rangle = \sum_{j=1}^{N} C_{ji}\phi_j$$

which can be plugged into the eigenvalue equation to find

$$\hat{f}\sum_{j=1}^{N} C_{ji}\phi_j = \epsilon_i \sum_{j=1}^{N} C_{ji}\phi_j$$

$$\sum_{j=1}^{N} C_{ji} \left\langle\phi_i\middle|\hat{f}\middle|\phi_j\right\rangle = \epsilon_i \sum_{j=1}^{N} C_{ji} \left\langle\phi_i\middle|\phi_j\right\rangle$$

$$\mathbf{FC} = \mathbf{SCE}$$

Roothaan equations

By applying the linear variational principle to a single-determinant wave function, we obtain the Roothaan equations:

$$\mathbf{FC} = \mathbf{SCE}$$

- $\mathbf{F}$: Fock matrix
- $\mathbf{C}$: Coefficient matrix
- $\mathbf{S}$: Overlap matrix
- $\mathbf{E}$: Energy matrix (diagonal)

Section

The Self-Consistent Field procedure

Why do we need a iterative procedure?

The Fock matrix is composed of elements that are calculated via

$$F_{ij} = \langle \phi_i | \hat{f} | \phi_j \rangle \, , \text{ wherein}$$

$$\hat{f} = \hat{h} + \sum_{j}^{N} \left[\hat{\mathcal{J}}_j - \hat{\mathcal{K}}_j \right]$$

The Coulomb $\hat{\mathcal{J}}$ and exchange $\hat{\mathcal{K}}$ operators depend on the electron density which in turn depends on the coefficients which is in turn is something we are establishing by solving the Roothaan equations.

Chicken-egg problem

Iterative procedure

Because the solution to the problem depends on the solution, we need an iterative procedure. We basically start by trying a solution and see what kind of new solution we obtain. This is done iteratively until the new solution is identical to the old solution. This is termed 'self-convergence'.

⚠ Important

When going from one iteration to the other, you can substitute the old solution for the new one, but you can also add a bit of the new solution to the old one. This is called mixing. Without some clever mixing strategy, some HF calculations will never converge. We will not explain mixing here in detail, but keep it in mind if you are building your own code.

Recap matrices

$$\mathbf{FC} = \mathbf{SCE}$$

The matrix elements are then given by

$$S_{ij} = \langle \phi_i | \phi_j \rangle$$

$$F_{ij} = \langle \phi_i | \hat{f} | \phi_j \rangle = \langle \phi_i | \hat{h} + \sum_j \hat{\mathcal{J}}_j + \sum_j \hat{\mathcal{K}}_j | \phi_j \rangle$$

$$\chi_i = \sum_j C_{ij} \phi_j$$

Restricted Hartree-Fock

Important

Although not explicitly derived, we will here postulate the SCF procedure for a restricted Hartree-Fock calculation, wherein spin-up and spin-down electrons are coupled and only systems with an even number of electrons can be evaluated. We thus *a priori* integrate out the spin and the iterations proceed over spatial orbitals only rather than the previously used spin-orbitals.

$$\begin{aligned} E &= 2\sum_{i}^{N/2} \langle \psi_i | \hat{h} | \psi_i \rangle + \sum_{i}^{N/2} \sum_{j}^{N/2} 2 \langle \psi_i \psi_j | \psi_i \psi_j \rangle - \langle \psi_i \psi_j | \psi_j \psi_i \rangle \\ &= 2\sum_{i}^{N/2} \langle \psi_i | \hat{h} | \psi_i \rangle + \sum_{ij}^{N/2} 2 \langle \psi_i \psi_j | \psi_i \psi_j \rangle - \langle \psi_i \psi_j | \psi_j \psi_i \rangle \end{aligned}$$

Auxiliary matrices

$$P_{ij} = 2 \sum_{k}^{N/2} C_{ik} C^{*}_{jk}$$

$$T_{ij} = \langle \phi_i | -\frac{1}{2} \nabla^2 | \phi_j \rangle$$

$$V_{ij} = \langle \phi_i | - \sum_{A} \frac{Z_A}{|r - R_A|} | \phi_j \rangle$$

$$H^{\text{core}}_{ij} = T_{ij} + V_{ij}$$

$$F_{ij} = H^{\text{core}}_{ij} + \sum_{kl} P_{kl} \left[\langle ik|jl \rangle - \frac{1}{2} \langle ik|lj \rangle \right]$$

$$F_{ij} = H^{\text{core}}_{ij} + G_{ij}$$

$$S_{ij} = \langle \phi_i | \phi_j \rangle$$

$$F_{ij} = \langle \phi_i | \hat{f} | \phi_j \rangle = \langle \phi_i | \hat{h} + \sum_{k} \hat{\mathcal{J}}_k + \sum_{k} \hat{\mathcal{K}}_k | \phi_j \rangle$$

$$\chi_i = \sum_{j} C_{ij} \phi_j$$

We introduce a density matrix $\mathbf{P}$ and separate $\mathbf{F}$ into a one-electron $\mathbf{H}_{\text{core}}$ and a two-electron $\mathbf{G}$ part.

The SCF procedure (1)

Step 1: Defining the system

- M nuclei at positions $\{R_M\}$
- N electrons
- K basis functions $\{\phi_i\}$

The SCF procedure (2)

Step 2: Calculate matrices that remain stagnant during the SCF procedure

$S_{ij} = \langle \phi_i \| \phi_j \rangle$	Overlap
$T_{ij} = \langle \phi_i \| -\frac{1}{2}\nabla^2 \| \phi_j \rangle$	Kinetic energy
$V_{ij} = \langle \phi_i \| -\sum_A \frac{Z_A}{\|r - R_A\|} \| \phi_j \rangle$	Nuclear attraction
$H_{ij}^{\text{core}} = T_{ij} + V_{ij}$	Core Hamiltonian
$\langle \phi_i \phi_j \| \phi_k \phi_l \rangle = \langle \phi_i \phi_j \| r_{ij}^{-1} \| \phi_k \phi_l \rangle$	Two-electron integrals

The SCF procedure (3)

Step 3: Orthonormalize the basis set

Construct a transformation matrix $\mathbf{X}$ such that the basis functions form an orthonormal set

$$\mathbf{C} = \mathbf{X}\mathbf{C}' \qquad \langle \phi_i' | \phi_j' \rangle = \delta_{ij}$$

We can apply this transformation to $\mathbf{F}$ to get $\mathbf{F}'$, which we can solve and always back-transform to our original basis set.

$$\mathbf{F}'\mathbf{C}' = \mathbf{C}'\epsilon \qquad \mathbf{F}' = \mathbf{X}^{\dagger}\mathbf{F}\mathbf{X}$$

The SCF procedure (4)

Step 3: Orthonormalize the basis set

Typically, a canonical transformation is used for this purpose, which is defined as

$$\mathbf{S} = \mathbf{U}\mathbf{s}\mathbf{U}^{\dagger}$$

$$\mathbf{X} = \mathbf{U}\mathbf{s}^{-1/2}$$

- We diagonalize the overlap matrix $\mathbf{S}$, which gives us the diagonal matrix $\mathbf{s}$ (eigenvalues) and a unitary matrix $\mathbf{X}$ (eigenvectors).
- We construct the transformation matrix $\mathbf{X}$ from the eigenvalues and eigenvectors.

The SCF procedure (5)

Step 4: Obtain a guess for the density matrix $\mathbf{P}$

$$\mathbf{P} = \mathbf{0}$$

We simply use the zero matrix.

The SCF procedure (6)

Step 5: Calculate $\mathbf{G}$, $\mathbf{F}$ and $\mathbf{F\prime}$

$$G_{ij} = \sum_{kl} P_{kl} \left[\langle ik|jl \rangle - \frac{1}{2} \langle ik|lj \rangle \right]$$

$$F_{ij} = H_{ij}^{\text{core}} + G_{ij}$$

$$\mathbf{F\prime} = \mathbf{X}^{\dagger}\mathbf{F}\mathbf{X}$$

⚠: Recall, we already obtained the values for all two-electron integrals in step 2 and $\mathbf{X}$ in step 3.

The SCF procedure (7)

Step 6: Diagonalize $\mathbf{F}'$ to obtain $\mathbf{C}'$ and ϵ

$$\mathbf{F}' = \mathbf{C}'\epsilon\mathbf{C}'^{\dagger}$$

We obtain the eigenvectors, i.e., the linear coefficients and the energies of the molecular orbitals in the orthonormalized basis.

⚠: Note, this is the same diagonalization algorithm as employed in step 3.

The SCF procedure (8)

Step 7: Calculate $\mathbf{C}$ from $\mathbf{C}'$

$$\mathbf{C} = \mathbf{X}\mathbf{C}'$$

To get the linear coefficients in our regular basis, we simply use the transformation matrix $\mathbf{X}$. Because $\mathbf{X}$ is a unitary matrix, we know that the energies of the molecular orbitals in the regular basis are the same as in the orthonormalized basis!

$$\epsilon = \epsilon'$$

The SCF procedure (9)

Step 8: Form a new density matrix $\mathbf{P}$ from $\mathbf{C}$

$$P_{ij} = 2\sum_{k}^{N/2} C_{ik} C^{*}_{jk}$$

Determine whether the new matrix $\mathbf{P}$ is the same as the old matrix $\mathbf{P}$ (within some margin). If it is, self-convergence is achieved. If not, go back to step 5.
Instead of checking whether the density matrices are converged, one can also probe the total energy.

The SCF procedure: Overview

$$S_{ij} = \langle \phi_i | \phi_j \rangle$$

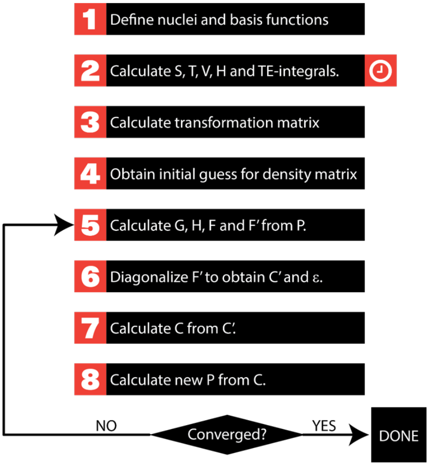
$$T_{ij} = \langle \phi_i | -\frac{1}{2}\nabla^2 | \phi_j \rangle$$

$$V_{ij} = \langle \phi_i | -\sum_A \frac{Z_A}{|r - R_A|} | \phi_j \rangle$$

$$H_{ij}^{\text{core}} = T_{ij} + V_{ij}$$

$$\langle \phi_i \phi_j | \phi_k \phi_l \rangle = \langle \phi_i \phi_j | r_{ij}^{-1} | \phi_k \phi_l \rangle$$

Calculating the two-electron integrals is the most time-consuming step and scales with N^4. (step 2)



1 Define nuclei and basis functions

2 Calculate S, T, V, H and TE-integrals.

3 Calculate transformation matrix

4 Obtain initial guess for density matrix

5 Calculate G, H, F and F′ from P.

6 Diagonalize F′ to obtain C′ and ε.

7 Calculate C from C′.

8 Calculate new P from C.

Converged?

NO

YES

DONE

Example: H_2O molecule

Let us look at a few steps of the procedure to understand the more complex parts.

- 3 atoms
- 7 contracted Gaussian functions (2 for H, 5 for O)
- 1s atomic orbitals on H
- 1s, 2s, (3x)2p atomic orbitals on O

H_2O: Overlap matrix

The overlap matrix $\mathbf{S}$:

	ϕ^{O}_{1s}	ϕ^{O}_{2s}	$\phi^{O}_{2p_x}$	$\phi^{O}_{2p_y}$	$\phi^{O}_{2p_z}$	$\phi^{H^1}_{1s}$	$\phi^{H^2}_{1s}$
ϕ^{O}_{1s}	1.0000	0.2367	0.0000	0.0000	0.0000	0.1584	0.1584
ϕ^{O}_{2s}	0.2367	1.0000	0.0000	0.0000	0.0000	0.8098	0.8098
$\phi^{O}_{2p_x}$	0.0000	0.0000	1.0000	0.0000	0.0000	0.3714	-0.3714
$\phi^{O}_{2p_y}$	0.0000	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000
$\phi^{O}_{2p_z}$	0.0000	0.0000	0.0000	0.0000	1.0000	-0.2322	-0.2322
$\phi^{H^1}_{1s}$	0.1584	0.8098	0.3714	0.0000	-0.2322	1.0000	0.6158
$\phi^{H^2}_{1s}$	0.1584	0.8098	-0.3714	0.0000	-0.2322	0.6158	1.0000

H_2O: Transformation matrix

The transformation matrix $\mathbf{X}$:

	ϕ^{O}_{1s}	ϕ^{O}_{2s}	$\phi^{\mathrm{O}}_{2p_x}$	$\phi^{\mathrm{O}}_{2p_y}$	$\phi^{\mathrm{O}}_{2p_z}$	$\phi^{\mathrm{H}^1}_{1s}$	$\phi^{\mathrm{H}^2}_{1s}$
ϕ^{O}_{1s}	-0.1769	-0.0000	0.8737	-0.0000	-0.5128	-0.0000	0.1205
ϕ^{O}_{2s}	2.7839	0.0000	-0.2202	-0.0000	-0.1818	-0.0000	0.3609
$\phi^{\mathrm{O}}_{2p_x}$	0.0000	-1.7232	0.0000	0.0000	-0.0000	0.7607	0.0000
$\phi^{\mathrm{O}}_{2p_y}$	-0.0000	0.0000	-0.0000	-1.0000	0.0000	0.0000	0.0000
$\phi^{\mathrm{O}}_{2p_z}$	-0.7851	-0.0000	-0.5239	-0.0000	-0.8095	-0.0000	-0.0983
$\phi^{\mathrm{H}^1}_{1s}$	-1.5636	2.1266	-0.1140	0.0000	0.0704	0.3082	0.3391
$\phi^{\mathrm{H}^2}_{1s}$	-1.5636	-2.1266	-0.1140	-0.0000	0.0704	-0.3082	0.3391

H_2O: Transformation result

Overlap matrix after coordinate transformation $\mathbf{X}^\dagger\mathbf{S}\mathbf{X}$:

	ϕ^{O}_{1s}	ϕ^{O}_{2s}	$\phi^{\mathrm{O}}_{2p_x}$	$\phi^{\mathrm{O}}_{2p_y}$	$\phi^{\mathrm{O}}_{2p_z}$	$\phi^{\mathrm{H}^1}_{1s}$	$\phi^{\mathrm{H}^2}_{1s}$
ϕ^{O}_{1s}	1.0000	−0.0000	−0.0000	0.0000	−0.0000	0.0000	0.0000
ϕ^{O}_{2s}	−0.0000	1.0000	0.0000	−0.0000	0.0000	0.0000	0.0000
$\phi^{\mathrm{O}}_{2p_x}$	−0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	−0.0000
$\phi^{\mathrm{O}}_{2p_y}$	0.0000	−0.0000	0.0000	1.0000	0.0000	0.0000	0.0000
$\phi^{\mathrm{O}}_{2p_z}$	−0.0000	0.0000	0.0000	0.0000	1.0000	−0.0000	−0.0000
$\phi^{\mathrm{H}^1}_{1s}$	0.0000	−0.0000	0.0000	0.0000	−0.0000	1.0000	−0.0000
$\phi^{\mathrm{H}^2}_{1s}$	0.0000	−0.0000	−0.0000	0.0000	−0.0000	−0.0000	1.0000

H_2O: Fock matrix

The Fock matrix $\mathbf{F}$:

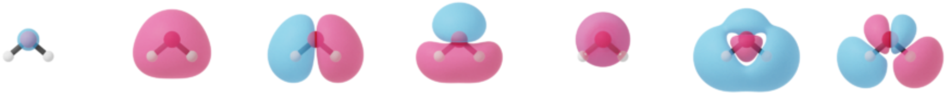
	ϕ^{O}_{1s}	ϕ^{O}_{2s}	$\phi^{O}_{2p_x}$	$\phi^{O}_{2p_y}$	$\phi^{O}_{2p_z}$	$\phi^{H^1}_{1s}$	$\phi^{H^2}_{1s}$
ϕ^{O}_{1s}	-20.5253	-5.2611	0.0000	0.0000	0.0627	-3.5308	-3.5308
ϕ^{O}_{2s}	-5.2611	-2.8003	-0.0000	-0.0000	0.3706	-2.1976	-2.1976
$\phi^{O}_{2p_x}$	0.0000	-0.0000	-0.9597	0.0000	-0.0000	-0.4454	0.4454
$\phi^{O}_{2p_y}$	0.0000	-0.0000	0.0000	-0.5717	-0.0000	0.0000	-0.0000
$\phi^{O}_{2p_z}$	0.0627	0.3706	-0.0000	-0.0000	-0.7378	0.4770	0.4770
$\phi^{H^1}_{1s}$	-3.5308	-2.1976	-0.4454	0.0000	0.4770	-1.8112	-1.5802
$\phi^{H^2}_{1s}$	-3.5308	-2.1976	0.4454	-0.0000	0.4770	-1.5802	-1.8112

H_2O: The coefficient matrix

The coefficient matrix $\mathbf{C}$:

$$
\begin{array}{c|r|r|r|r|r|r|r}
\phi_{1s}^{\mathrm{O}} & 0.9933 & -0.2453 & 0.0000 & 0.0915 & 0.0000 & -0.1303 & 0.0000 \\
\phi_{2s}^{\mathrm{O}} & 0.0477 & 0.8856 & -0.0000 & -0.3516 & -0.0000 & 2.6556 & -0.0000 \\
\phi_{2p_x}^{\mathrm{O}} & -0.0000 & 0.0000 & -0.7522 & -0.0000 & 0.0000 & -0.0000 & -1.7270 \\
\phi_{2p_y}^{\mathrm{O}} & -0.0000 & -0.0000 & 0.0000 & -0.0000 & 1.0000 & -0.0000 & 0.0000 \\
\phi_{2p_z}^{\mathrm{O}} & -0.0107 & -0.3400 & 0.0000 & -0.9382 & -0.0000 & -0.7482 & 0.0000 \\
\phi_{1s}^{\mathrm{H}^1} & -0.0154 & 0.0450 & -0.3186 & -0.0121 & 0.0000 & -1.6048 & 2.1251 \\
\phi_{1s}^{\mathrm{H}^2} & -0.0154 & 0.0450 & 0.3186 & -0.0121 & -0.0000 & -1.6048 & -2.1251
\end{array}
$$

H_2O: Isosurfaces of molecular orbitals



- With a basis set composed of 7 basis functions, we find 7 molecular orbitals
- 4 are occupied, 3 are unoccupied or virtual

Section

Invariants under unitary transformations

Output objects

Important output objects

- Hamiltonian (Fock) matrix $\mathbf{H}$
- Kinetic energy matrix $\mathbf{T}$
- Nuclear attraction matrix $\mathbf{V}$
- Coefficient matrix $\mathbf{C}$
- Two-electron integral tensor $\mathcal{T}$
- Density matrix $\mathbf{P}$

Useful equations

Molecular orbital energies:

$$\vec{e} = \mathrm{diag}\left(\mathbf{C}\mathbf{H}\mathbf{C}^{\intercal}\right) \tag{1}$$

Total electronic energy:

$$E = \frac{1}{2} \sum_{ij} P_{ji} \left(T_{ij} + V_{ij} + H_{ij}\right) E_{\text{nuc-nuc}} \tag{2}$$

Invariants: Total electronic energy

E is invariant under an arbitrary unitary transformation **among** the N occupied molecular orbitals. Given the transformation matrix $\mathbf{T}$ defined as

$$\mathbf{T} = \left(\begin{array}{c:c} \mathbf{U}_{N \times N} & 0 \\ \hdashline 0 & I_{N' \times N'} \end{array}\right) \tag{3}$$

$$\mathbf{C}' = \mathbf{CT} \tag{4}$$

it is the case that

$$E(\mathbf{C}) = E(\mathbf{C}') \tag{5}$$

- When the basis is real-valued, U can be an orthogonal matrix

Invariants: Density matrix and electron density

Also the density matrix

$$P_{ij} = 2 \sum_{k}^{N/2} C_{ik} C_{jk}^{*} \tag{6}$$

is invariant and by extension the total electron density distribution ρ.

Strictly forbidden

E, $\mathbf{P}$ and ρ are **not** invariant under any kind of mixing between occupied and unoccupied molecular orbitals.

Invariants: Summed energy terms

The total electronic energy of a molecule is the sum of

- Kinetic energy
- Nuclear attraction
- Electron-electron repulsion
- Exchange energy

These terms all depend on the density matrix and are thus, by extension, also invariant under an arbitrary unitary transformation.

New representations

Things that **will** change under an arbitrary unitary transformation are properties **per** molecular orbital

- Molecular orbital (MO) energy
- Wave function and electron density associated with a single MO
- Average position (orbital centroids)
- Individual kinetic / nuclear / electron-electron repulsion and exchange energies

New representations

Take-home message: Infinitely many representations

- The canonical molecular orbitals are very much delocalized. The orbitals are maximally spread and occupy a large amount of space.
- How would the opposite situation look like and what can we learn from this?

Section

Foster-boys localization

Principle

- The canonical orbitals are typically delocalized, i.e. spread out over the whole molecule.
- Mathematically speaking, $\langle \psi_i | r^2 | \psi_i \rangle$ is very small.
- Foster-Boys algorithm searches for a unitary transformation that maximizes $\sum_i \langle \psi_i | r^2 | \psi_i \rangle$.
- Result is localized orbitals.

Method

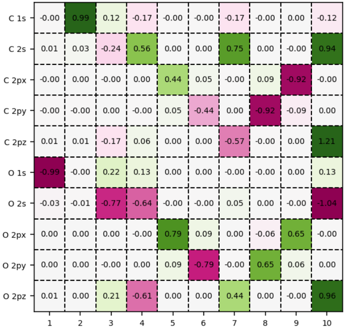
- Given $\mathbf{C}$, we pick two molecular orbitals (among the occupied set) at random: ψ_i, ψ_j.
- Generate two new solutions by mixing

$$\psi_1\prime = \cos\left(\alpha\right)\psi_i + \sin\left(\alpha\right)\psi_j \tag{7}$$

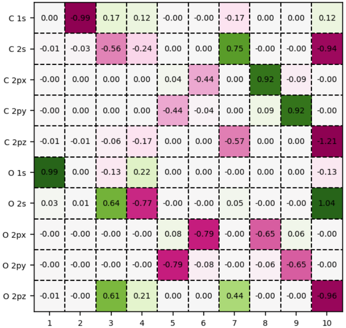
$$\psi_2\prime = -\sin\left(\alpha\right)\psi_i + \cos\left(\alpha\right)\psi_j \tag{8}$$

corresponding to a rotation of α radians in the Hilbert space spanned by the outer product of ψ_i and ψ_j.

Example: Column operations



	1	2	3	4	5	6	7	8	9	10
C 1s	-0.00	0.99	0.12	-0.17	-0.00	-0.00	-0.17	-0.00	0.00	-0.12
C 2s	0.01	0.03	-0.24	0.56	0.00	0.00	0.75	0.00	-0.00	0.94
C 2px	-0.00	0.00	-0.00	0.00	0.44	0.05	-0.00	0.09	-0.92	-0.00
C 2py	-0.00	-0.00	0.00	-0.00	0.05	-0.44	0.00	-0.92	-0.09	0.00
C 2pz	0.01	0.01	-0.17	0.06	0.00	0.00	-0.57	-0.00	0.00	1.21
O 1s	-0.99	-0.00	0.22	0.13	0.00	0.00	-0.00	-0.00	0.00	0.13
O 2s	-0.03	-0.01	-0.77	-0.64	-0.00	-0.00	0.05	0.00	-0.00	-1.04
O 2px	0.00	0.00	0.00	-0.00	0.79	0.09	0.00	-0.06	0.65	-0.00
O 2py	0.00	0.00	-0.00	0.00	0.09	-0.79	-0.00	0.65	0.06	0.00
O 2pz	0.01	0.00	0.21	-0.61	0.00	0.00	0.44	0.00	-0.00	0.96

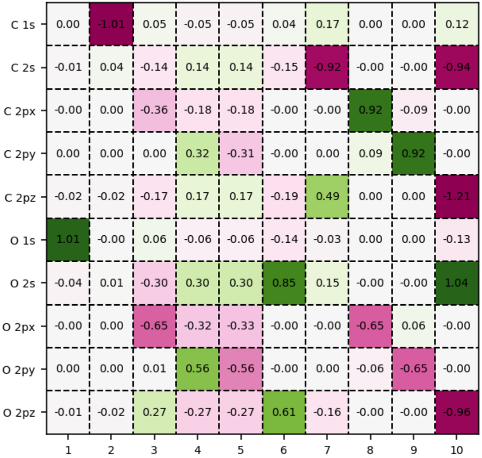


	1	2	3	4	5	6	7	8	9	10
C 1s	0.00	-0.99	0.17	0.12	-0.00	-0.00	-0.17	0.00	0.00	0.12
C 2s	-0.01	-0.03	-0.56	-0.24	0.00	0.00	0.75	-0.00	-0.00	-0.94
C 2px	-0.00	0.00	0.00	0.00	0.04	-0.44	0.00	0.92	-0.09	-0.00
C 2py	-0.00	0.00	0.00	0.00	-0.44	-0.04	0.00	0.09	0.92	-0.00
C 2pz	-0.01	-0.01	-0.06	-0.17	0.00	0.00	-0.57	0.00	0.00	-1.21
O 1s	0.99	0.00	-0.13	0.22	0.00	0.00	-0.00	0.00	0.00	-0.13
O 2s	0.03	0.01	0.64	-0.77	-0.00	-0.00	0.05	-0.00	-0.00	1.04
O 2px	-0.00	-0.00	-0.00	-0.00	0.08	-0.79	-0.00	-0.65	0.06	-0.00
O 2py	-0.00	-0.00	-0.00	-0.00	-0.79	-0.08	-0.00	-0.06	-0.65	-0.00
O 2pz	-0.01	-0.00	0.61	0.21	0.00	0.00	0.44	-0.00	-0.00	-0.96

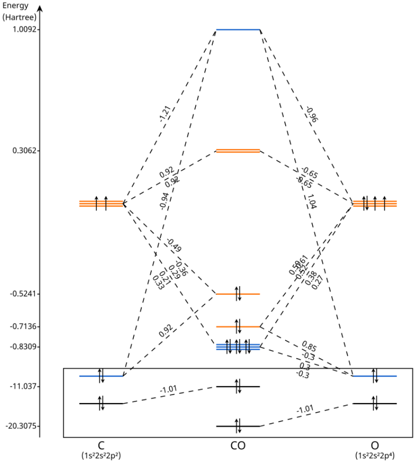
Note that the third and fourth column are rotated among each other by $\pi/2$ radians. There is no loss of information in this operation as the operation is unitary.

Revisiting: CO molecule

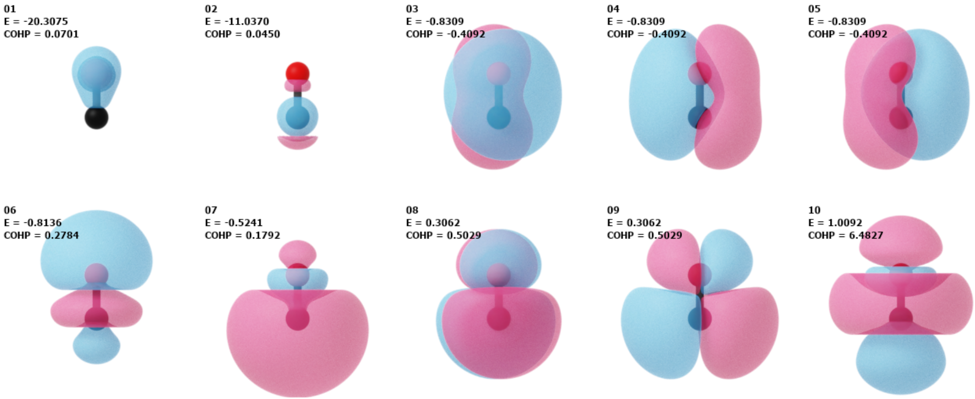
Foster-boys localization



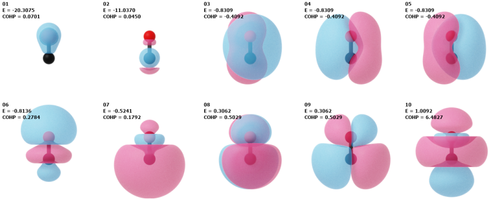
	1	2	3	4	5	6	7	8	9	10
C 1s	0.00	-1.01	0.05	-0.05	-0.05	0.04	0.17	0.00	0.00	0.12
C 2s	-0.01	0.04	-0.14	0.14	0.14	-0.15	-0.92	-0.00	-0.00	-0.94
C 2px	-0.00	0.00	-0.36	-0.18	-0.18	-0.00	-0.00	0.92	-0.09	-0.00
C 2py	0.00	0.00	0.00	0.32	-0.31	-0.00	0.00	0.09	0.92	-0.00
C 2pz	-0.02	-0.02	-0.17	0.17	0.17	-0.19	0.49	0.00	0.00	-1.21
O 1s	1.01	-0.00	0.06	-0.06	-0.06	-0.14	-0.03	0.00	0.00	-0.13
O 2s	-0.04	0.01	-0.30	0.30	0.30	0.85	0.15	-0.00	-0.00	1.04
O 2px	-0.00	0.00	-0.65	-0.32	-0.33	-0.00	-0.00	-0.65	0.06	-0.00
O 2py	0.00	0.00	0.01	0.56	-0.56	-0.00	0.00	-0.06	-0.65	-0.00
O 2pz	-0.01	-0.02	0.27	-0.27	-0.27	0.61	-0.16	-0.00	-0.00	-0.96



Revisiting: CO molecule - Isosurfaces + COHP



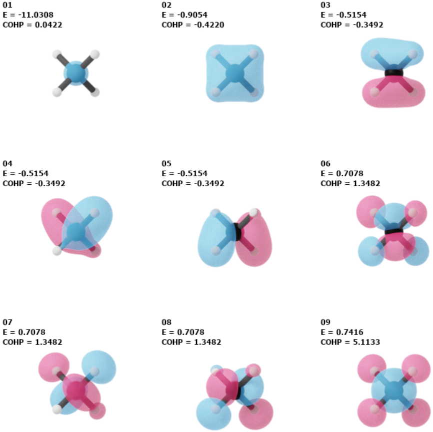
Localized orbitals are Lewis structures

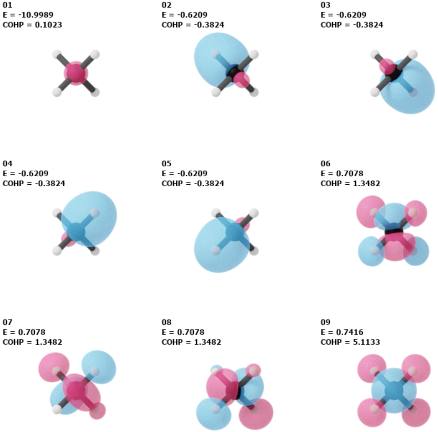


Localized orbitals are very much akin to the well-known Lewis structures.

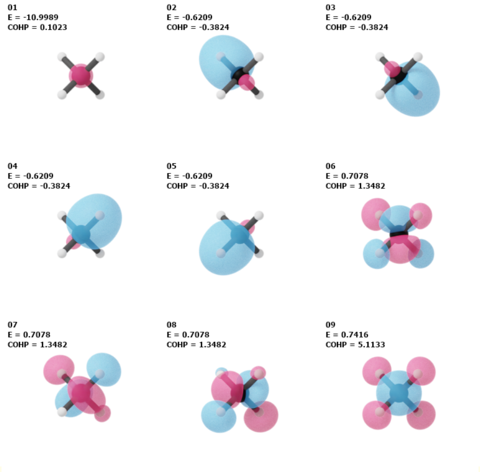
- Triple-degenerate orbital is the C-O triple-bond
- HOMO-1 and HOMO orbitals are the lone pairs on O and C, respectively.

Another example: CH4





CH4: Interpretation Foster-Boys localized orbitals

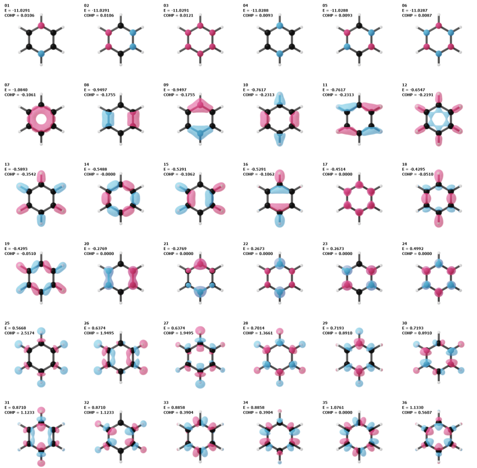


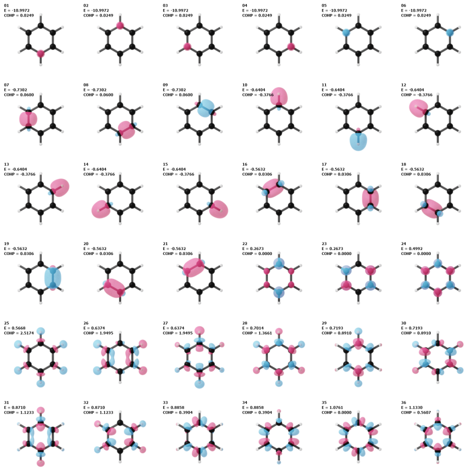
- (semi-)Non-bonding orbital of C 1s
- Quadruple degenerate bonding orbital (C-H) bond
- Consistent with Lewis structure

Section

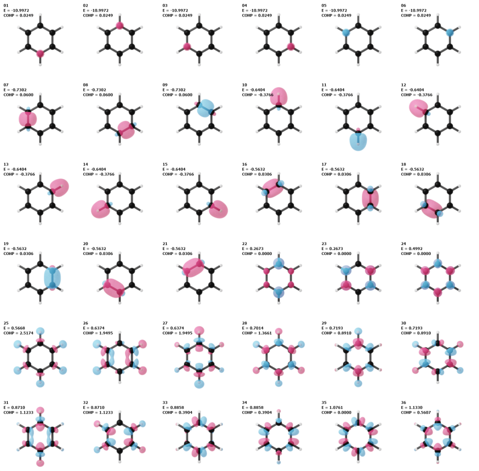
Benzene: showcasing the lack of uniquene

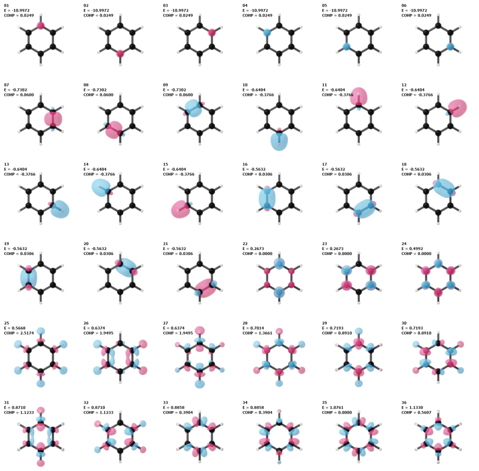
Benzene: canonical versus localized



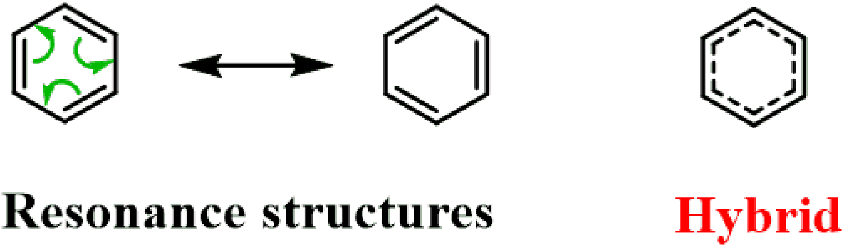


Benzene: Two different localized sets





Benzene: localized resonance structures



- Localized solutions are the two resonance structures
- Canonical solution is the "hybrid" structure

Uniqueness of solutions

- Canonical solutions are not unique
  - Signs are arbitrary (cancel out when constructing electron density)
  - One can arbitrarily mix degenerate orbitals (no loss of information)
- Localized solutions are not unique
  - Multiple solutions with same local minimum (maximum $\sum_i \langle \psi_i | r^2 | \psi_i \rangle$ value)
  - Signs are arbitrary (cancel out when constructing electron density)
  - One can arbitrarily mix degenerate orbitals (no loss of information)
  - Even more degenerate orbitals as compared to canonical solutions

Section

Closing remarks

Moral of the story

Which solutions to pick?

- All solutions are equally valuable and are just different “lenses” looking at the same abstract object.
- Canonical solutions are not unique and have their own invariants upon transformations within their own subspace
- Some solutions are more suitable for analyses.
- Some experiments “force” the system to pick a solution. (wave function collapse; symmetry conditions)