Computational Chemistry

Filipp Furche
Department of Chemistry, UC Irvine

Fall 2023
December 4, 2023, Version 7705e31

Contents

1 Connecting Theory and Experiment through Computation

1.1 Theory

- A theory is a set of axioms or "first principles", i.e., statements believed/posited/postulated to be true.

- The axioms in scientific theories are statements about measurable quantities, or the conditions of measurability.

- Axioms must be

  - as simple as possible, including simple to evaluate,
  - as general as possible, i.e., explain all known observations,
  - as accurate as possible,
  - logically consistent,
  - and there should be as few as possible.

- Theories consist of logical statements, and in science they are often expressed by equations.

- Theories are not data.

- Theories are conceived to explain on empirical observation. However, most powerful theories require only minimal empirical input (e.g. natural constants)

- Chemical theories make statements about chemical observables, such as molecular energetics, properties, kinetics, reactivity, mechanisms, etc.

- Examples of chemical theories:

  - Phlogiston theory
  - Lavoisier redox theory
  - Chemical thermodynamics
  - Chemical kinetics
  - Lewis theory of chemical bonding
  - MO theory of chemical bonding
  - Ligand field theory

1.2 Experiment

- Scientific experiments are theory-guided observations of natural phenomena.

- Experiments must

  - be well-controlled and reproducible, independent of individual,
  - yield a unique result,

Figure 1: Examples for the theory-ladenness of observation [1]
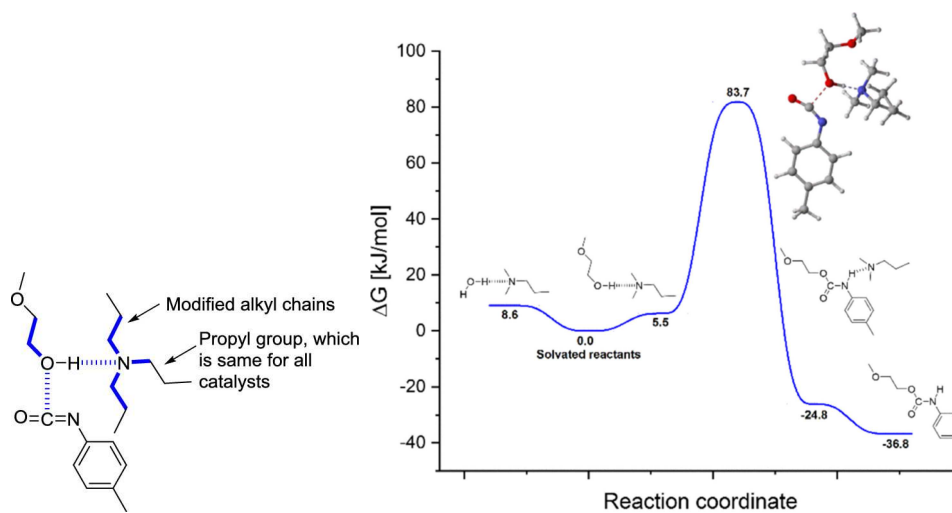
1.3 Computation

Figure 2: Rational design of polyurethane polymerization catalysts by electronic structure calculations [2, 3]

.

- Example: New catalysts for polyurethane polymerization avoid air pollution from formaldehyde emission, see Fig. 2

- Scientific computations quantitatively evaluate theoretical statements and test theoretical hypotheses by comparison to experiment.

- Computations should

  - be as efficient (CPU time, carbon footprint, human effort) as possible,
  - use the simplest possible means to obtain desired results for the right reasons,
  - as accurate as necessary,
  - reproducible,
  - yield directly observable quantities,
  - be as transparent as possible.

- Computational chemistry aims to

  - predict experimental results, and partially replace them,
  - explain known experimental results,
  - allow a stringent test of theories and hypotheses by quantitative comparison to experiment

- Computational chemists need to

  - know which questions to ask and why,
  - be intimately familiar with experiment and theory,
  - be able to communicate well using several "languages",
  - understand how computers, computational methods, and algorithms work,
  - know about methods and tools, understand where they are appropriate, and why,
  - have an intuitive feeling for numbers and orders of magnitude.

2 Paradigms of Computational Chemistry

2.1 Symmetry

2.1.1 Groups

- Symmetry operations map a geometrical object onto itself.

- Molecular symmetry elements: Identity (E), n-fold symmetry (rotation) axis (C_n), symmetry (reflection) plane (σ), inversion center (i), n-fold rotation-reflection axis (S_n)

  Example 2.1. Symmetry elements of water: $\{E, C_2, \sigma_\mathrm{v}, \sigma_\mathrm{v'}\}$
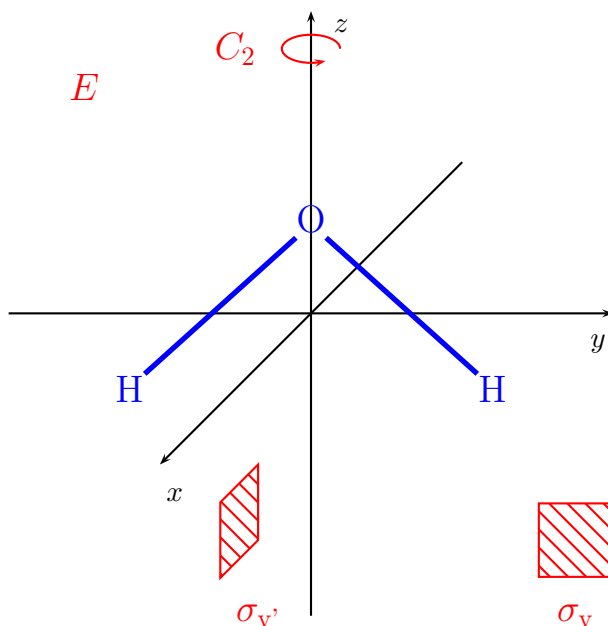
Figure 3: Symmetry elements of H_2O

a b	E	C_2	σ_v	$\sigma_{v'}$
E	E	C_2	σ_v	$\sigma_{v'}$
C_2	C_2	E	$\sigma_{v'}$	σ_v
σ_v	σ_v	$\sigma_{v'}$	E	C_2
$\sigma_{v'}$	$\sigma_{v'}$	σ_v	C_2	E

Table 1: Product of symmetry operations $c = a \cdot b$. Multiplication table of C_{2v}.

- Multiplication of two operations $c = a \cdot b$: Application of b followed by a is another symmetry operation c (group property)

- Group multiplication table: See Tab. 1

- The symmetry elements $\{E, C_2, \sigma_v, \sigma_{v'}\}$ form the *point group C_{2v}*.

- A set of g elements $G = \{a, b, c, \ldots\}$ with a multiplication $a \cdot b$ and the following properties is a group.

  1. $c = a \cdot b \in G$ (closure), and c is unique
  2. $(a \cdot b) \cdot c = a \cdot (b \cdot c)$ (associativity)
  3. There is an identity element $e \in G$ with $a \cdot e = a \forall a \in G$
  4. For every $a \in G$ there is an inverse a^{-1} such that $a \cdot a^{-1} = e$.
  5. Abelian (commutative) groups only: $a \cdot b = b \cdot a$ (commutativity)

- Group order: Number of elements $g = \text{ord}(G)$

- Determining the point group of a molecule: see Fig. 4

6
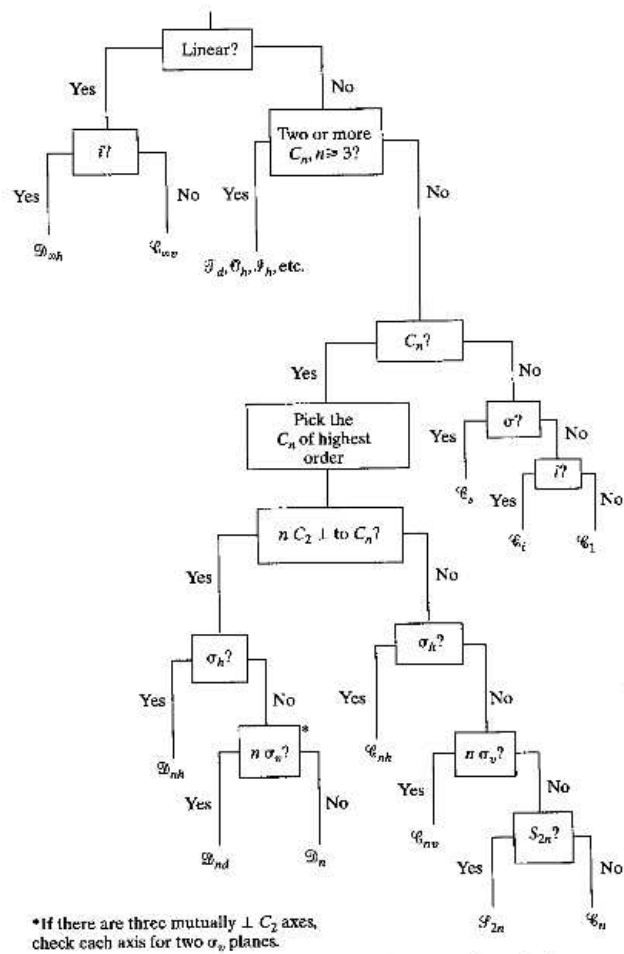
Figure 4: Flow chart to determine the point group of a molecule.

2.1.2 Classes

- Each element a of a group G belongs to exactly one class of elements,

$$K_a = \{b \cdot a \cdot b^{-1} | b \in G\}. \tag{1}$$

- In Abelian groups each element has its own class.

Example 2.2. C_{3v}, the point group of NH_3 (see Fig. 5). C_{3v} contains 3 classes: $\{E\}$, $\{C_3, C_3^2\}$, $\{\sigma_v(1), \sigma_v(2), \sigma_v(3)\}$.
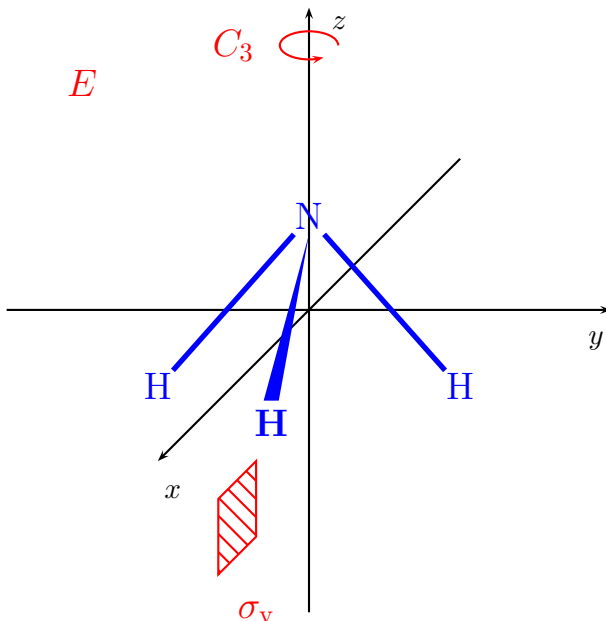
Figure 5: Some symmetry elements of NH_3

2.1.3 Representations

Definitions

- A representation Γ is a homomorphism of a group G onto a group of non-singular operators:

$$\Gamma : a \to P(a) \forall a \in \mathbf{G} \tag{2}$$

Homomorphism preserves the group property, i.e., $\forall a, b \in \mathbf{G}$,

$$P(a)P(b) = P(a \cdot b). \tag{3}$$

- These operators are defined on representation space, and are matrices for finite groups ("representation matrices")

- Representation space is a vector space spanning, e.g., $\mathbf{R}^3$, molecular orbitals, vibrations, tensors etc.

a	E	C_2	σ_{v}	$\sigma_{\mathrm{v'}}$
$P(a)$	1	-1	-1	1

Table 2: The 1d p_x representation of C_{2v}

	E	C_2	σ_{v}	$\sigma_{\mathrm{v'}}$
χ_{2s}	2	0	2	0

Table 3: Characters of the 2s representation of C_{2v}

- The representation matrices can be constructed by considering the action of each symmetry element on orthonormal basis vectors $|i\rangle, |j\rangle$:

$$P_{ij}(a) = \langle i|\hat{P}(a)|j\rangle. \tag{4}$$

Example 2.3. Representation spanned by O p_x orbital in water

- Character of a representation:

$$\chi_\Gamma(a) = \mathrm{tr}P(a) \tag{5}$$

- Behavior of representation under symmetry operation a: character $\chi(a)$

- Determining the character:

    - Each basis vector mapped onto itself by a contributes ± 1.
    - Each basis vector mapped onto another basis vector contributes 0.

Example 2.4. Representation spanned by 2 hydrgen s orbitals in water

- The regular representation is the g-dimensional representation spanned by all elements of the group.

Irreducible Representations

- Any representation may be decomposed into a sum of irreducible representations (IR-REPs). Vectors spanning inequivalent IRREPs are orthogonal and have zero matrix elements under any symmetry operation. The corresponding representation matrices are thus block diagonal.

- IRREPs and their characters are listed in the character table of a point group, e.g. for C_{2v} see Tab. 6.

- There are as many IRREPs as there are classes in a group.

- Multiplicity of IRREP α in a given representation:

$$n_\alpha = \frac{1}{g}\sum_{a\in G}\chi_\alpha^*(a)\chi_\Gamma(a) \tag{6}$$

C_{2v}	E	C_2	σ_{v}	$\sigma_{\mathrm{v}'}$	
A_1	1	1	1	1	$z; x^2; y^2; z^2$
A_2	1	1	-1	-1	$R_z; xy$
B_1	1	-1	1	-1	$y; R_x; yz$
B_2	1	-1	-1	1	$x; R_y; xz$

Figure 6: Character table of C_{2v}

α	A_1	A_2	B_1	B_2
n_α	0	0	0	1

Table 4: IRREP multiplicities for the p_x representation of C_{2v}

Example 2.5. Tab. 4 shows the IRREP multiplicities for the p_x representation. so the p_x representation contains B_2 only.

Tab. 5 shows the IRREP multiplicities for the $2s$ representation. so the $2s$ representation contains A_1 and B_1.

- Character projection rules for determining symmetry adapted basis functions: see Tab. 6

- Character projection operator onto IRREP α:

$$\Pi_\alpha = \frac{d_\alpha}{g} \sum_{a \in G} \chi_\alpha^*(a) P(a) \tag{7}$$

2.1.4 Applications

Selection Rules

- An integral

$$\int d\tau\, f(x) \tag{8}$$

is zero due to symmetry if the integrand $f(x)$ is not totally symmetric.

Example 2.6. Two molecular orbitals $\psi(A_1), \psi(B_1)$ of water. Because $A_1 \times B_1 = B_1$,

$$\int d\tau\, \psi(A_1)(x)^* \psi(B_1)(x) = 0 \tag{9}$$

by symmetry.

α	A_1	A_2	B_1	B_2
n_α	1	0	1	0

Table 5: IRREP multiplicities for the $2s$ representation of C_{2v}

	E	C_2	σ_{v}	$\sigma_{\mathrm{v}'}$
s_A becomes	s_A	s_B	s_A	s_B
χ_{B_1}	1	-1	1	-1
Product	s_A	$-s_B$	s_A	$-s_B$
Normalized sum	$\frac{1}{\sqrt{2}}(s_A - s_B)$			

Table 6: Determination of the normalized B_1 basis vector by projection out of Γ_{2s}.

Example 2.7. The electronic dipole moment is

$$\boldsymbol{\mu} = -\int d^3 r \mathbf{r} \rho(\mathbf{r}), \tag{10}$$

where the ground state density $\rho(\mathbf{r})$ is totally symmetric (non-degenerate states). $\boldsymbol{\mu}$ is zero if none of x, y, z transform according to the totally symmetric IRREP. This is always true if there is an inversion center.

Example 2.8. An electronic transition from the ground state ψ_0 to the n-th excited state ψ_n is dipole allowed if the transition dipole moment

$$\mu_{0n} = -\int d\tau \psi_0^*(x) \mathbf{r} \psi_n(x) \tag{11}$$

is non-zero. If the ground state is totally symmetric, ψ_n must transform according to the IRREP(s) containing x, y, z for a dipole allowed transition.

2.1.5 Worked Example: NH$_3$ Stretching Vibrations

Reducible Representation Γ_s

- The 3 N-H stretching vibrations span a reducible representation Γ_s of the C$_{3v}$ point group. We would like to subduce Γ_s into IRREPS and determine the resulting symmetry-adapted vibrational modes.

- In the first step, we draw a coordinate system for NH$_3$ along with three arrows representing unit vectors in the direction of the three N-H stretches, which we may label $|1\rangle, |2\rangle, |3\rangle$, see Fig. 7. The representation space is spanned by all nontrivial linear combinations of these three basis vectors, i.e., any normalizable

$$c_1 |1\rangle + c_2 |2\rangle + c_3 |3\rangle = \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} \tag{12}$$

with real c_1, c_2, c_3.

- It is often helpful to draw another perspective, e.g., looking along negative z direction from above the molecule, as in Fig. 8.

- At this point, we remind ourselves of the symmetry elements in C$_{3v}$ by looking at the character table, Tab. 7, and try to locate them in Figs. 7 and 8.
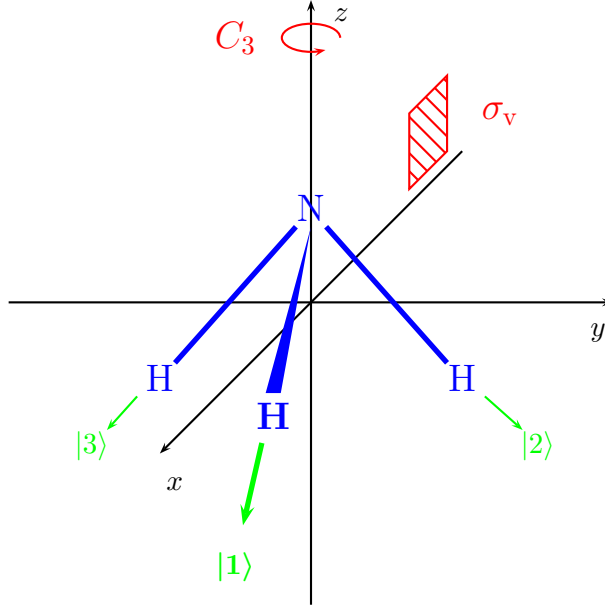
Figure 7: The three primitive N-H stretches of NH_3 span the Γ_s representation of C_{3v}.

C_{3v}	E	$2C_3$	$3\sigma_v$	
A_1	1	1	1	z, $x^2 + y^2$, z^2
A_2	1	1	-1	R_z
E	2	-1	0	(x,y), (R_x, R_y), $(x^2 - y^2, xy)$, (yz, xz)

Table 7: Character Table of C_{3v}.

- Next, we determine the representation matrices of C_{3v} using Eq. (4) and the orthonormality of the basis vectors. (Note that $|1\rangle$, $|2\rangle$, and $|3\rangle$ are not vectors in the 2-dimensional plane ($\mathbf{R}^2$), but they represent linearly independent displacements of nuclei.) Each symmetry element has its own representation matrix, there is no such thing as a class representation matrix, even though the character table contains only classes of elements. We thus obtain:

$$
P(E) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad P(C_3) = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \quad P(C_3^2) = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix},
$$
$$
P(\sigma_1) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad P(\sigma_2) = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad P(\sigma_3) = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.
$$

(13)

- You may use Python, e.g., the interpreter at pythonanywhere. It is sufficient to enter 3 generators, then the other representation matrices can be obtained using the group multiplication table:

```
import numpy as np
```
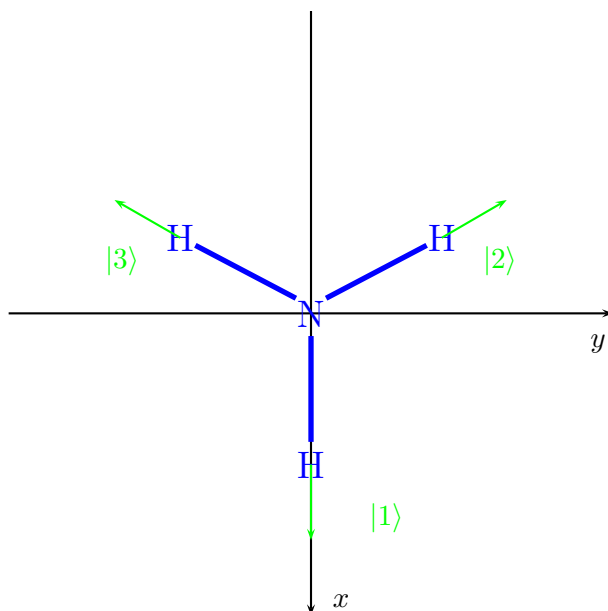
Figure 8: Top view of the 3 N-H stretches of NH_3 along the negative z axis.

```python
e = np.array([ [1., 0., 0.],
               [0., 1., 0.],
               [0., 0., 1.]]) # identity

c3 = np.array([ [0., 0., 1.],
                [1., 0., 0.],
                [0., 1., 0.]]) # C3 rotation

s1 = np.array([ [1., 0., 0.], # sigma_1 reflection
                [0., 0., 1.],
                [0., 1., 0.]])

c32 = c3.dot(c3) # C3^2 rotation

s2 = c3.dot(s1) # sigma_2 reflection

s3 = c32.dot(s1) # sigma_3 reflection

p = np.array( [ e, c3, c32, s1, s2, s3]) # array combining all operators
```

- You can check this result by comparing to the group multiplication table

- The characters are simply the traces (sum of diagonal elements) of the representation matrices:

	E	C_3	C_3^2	σ_1	σ_2	σ_3
χ_{Γ_s}	3	0	0	1	1	1

Table 8: Characters of the Γ_s representation.

```
char = np.zeros(6)

for i in range(6):
    char[i] = np.trace(p[i])
```

Again, each element has its own character, even though the characters of elements in the same class are identical.

Subduction of Γ_s

- We determine the multiplicty of each IRREP of C_{3v} in Γ_s from Eq. (6), taking the IRREP characters from Tab. (7):

$$
\begin{aligned}
n_{A_1} &= \frac{1}{6}\left(3 + 0 + 0 + 1 + 1 + 1\right) = 1, \\
n_{A_2} &= \frac{1}{6}\left(3 + 0 + 0 - 1 - 1 - 1\right) = 0, \\
n_E &= \frac{1}{6}\left(6 + 0 + 0 + 0 + 0 + 0\right) = 1.
\end{aligned}
\tag{14}
$$

Thus, we may write $\Gamma_s = A_1 + E$, i.e., Γ_s contains one A_1 and one E IRREP.

```
irrep_char = np.array( [[1., 1., 1., 1., 1., 1.],
                        [1., 1., 1., -1., -1., -1.],
                        [2., -1., -1., 0., 0., 0.]])
                        # IRREP characters

for i in range(3):
    print "IRREP No. ", i, "multiplicity :",
          np.dot(char,irrep_char[i])/6
```

Symmetry Adapted Basis: IRREP A_1

- Since the A_1 IRREP is one-dimensional, we can easily obtain the symmetry adapted basis vector from character projection. By Eq. (7), the projection operator is

$$
\Pi_{A_1} = \frac{1}{6} \sum_{a \in C_{3v}} \chi_{A_1}^* P(a) = \frac{1}{3}\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix},
\tag{15}
$$

where the representation matrices of Γ_s, Eq. (13), were inserted.

14

- We may now pick an arbitrary basis vector, e.g., $|1\rangle$, and apply Π_{A_1} to get the unnormalized symmetry adapted basis vector for A_1,

$$|A_1\rangle = N\Pi_{A_1}|1\rangle = \frac{N}{3}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}. \tag{16}$$

The normalization constant N is found to be $\sqrt{3}$ by requiring $\langle A_1|A_1\rangle = 1$, and thus

$$|A_1\rangle = \frac{1}{\sqrt{3}}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \tag{17}$$

is our first symmetry adapted vibration in Γ_s. This vibration displaces all 3 H atoms simultaneously by the same amount along the N-H bonds.

```python
# Set up A1 projector
pi_a1 = np.zeros((3,3))

for i in range(6):
    pi_a1 = pi_a1 + irrep_char[0,i]*p[i]

pi_a1 = pi_a1/6.

# Apply projector to unit vector:

v_a1 = pi_a1.dot( [1., 0., 0.] )

# Normalize

v_a1 = v_a1 = v_a1/np.linalg.norm(v_a1)
```

Symmetry Adapted Basis: IRREP E

- The E IRREP is 2-dimensional. We may determine the symmetry adapted basis functions by character projection and subsequeny orthogonalization, or we may use IRREP matrices. The former method is used in many introductory texts, so we will use the latter here.

- From the last column of the C_{3v} character table, we see that the E IRREP can be spanned by (x, y), i.e., the pair of unit vectors in x and y direction. In contrast to $|1\rangle$, $|2\rangle$, $|3\rangle$, x and y are in $\mathbf{R}^2$ and their symmetry images may not be orthogonal. For example, $P(C_3)x = \cos(120°)x + \sin(120°)y = -\frac{1}{2}x + \frac{\sqrt{3}}{2}y$, which is not orthogonal to x or y.

- Using x and y as basis, the IRREP matrices for E are thus

$$P_E(E) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad P_E(C_3) = \frac{1}{2}\begin{pmatrix} -1 & -\sqrt{3} \\ \sqrt{3} & -1 \end{pmatrix},$$

$$P_E(C_3^2) = \frac{1}{2}\begin{pmatrix} -1 & \sqrt{3} \\ -\sqrt{3} & -1 \end{pmatrix}, \quad P_E(\sigma_1) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{18}$$

$$P_E(\sigma_2) = \frac{1}{2}\begin{pmatrix} -1 & \sqrt{3} \\ \sqrt{3} & 1 \end{pmatrix}, \quad P_E(\sigma_3) = \frac{1}{2}\begin{pmatrix} -1 & -\sqrt{3} \\ -\sqrt{3} & 1 \end{pmatrix}.$$

- To check that this is correct, we can determine the characters from the traces of the IRREP matrices, which have to match the ones in the character table, Tab. 7.

- The two basis vectors spanning E are often referred to as IRREP "columns".

- For nondegenerate IRREPS α, the character projection operator may be generalized to projection operators for each column k,

$$\Pi_{\alpha k} = \frac{d_\alpha}{g} \sum_{a \in G} P_{\alpha kk}(a)^* P(a). \tag{19}$$

The relevant diagonal elements $P_{\alpha kk}(a)$ according to Eq. (18) are listed in Tab. 9. These elements must sum up to the IRREP characters.

	E	C_3	C_3^2	σ_1	σ_2	σ_3
$P_{E\,11}$	1	$-\frac{1}{2}$	$-\frac{1}{2}$	1	$-\frac{1}{2}$	$-\frac{1}{2}$
$P_{E\,22}$	1	$-\frac{1}{2}$	$-\frac{1}{2}$	-1	$\frac{1}{2}$	$\frac{1}{2}$

Table 9: Diagonal elements of the E IRREP matrices used for projection.

- Using Tab. 9 and the representation matrices of Γ_s, Eq. (13), we obtain

$$\Pi_{E\,1} = \frac{2}{6}\begin{pmatrix} 2 & -1 & -1 \\ -1 & \frac{1}{2} & \frac{1}{2} \\ -1 & \frac{1}{2} & \frac{1}{2} \end{pmatrix}, \quad \Pi_{E\,2} = \frac{2}{6}\begin{pmatrix} 0 & 0 & 0 \\ 0 & \frac{3}{2} & -\frac{3}{2} \\ 0 & -\frac{3}{2} & \frac{3}{2} \end{pmatrix}. \tag{20}$$

- As before, we may now pick an arbitrary vector in the representation space and appy $\Pi_{E\,k}$ to obtain the symmetry adapted basis vectors for E. However, $|1\rangle$ is not the most convenient choice, since it is mapped onto the zero vector by $\Pi_{E\,2}$, so it yields only one. Thus, we choose $|2\rangle$, and obtain

$$|E_1\rangle = N_1 \Pi_{E\,1}|2\rangle = \frac{N}{3}\begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix},$$

$$|E_2\rangle = N_2 \Pi_{E\,2}|2\rangle = \frac{N}{3}\begin{pmatrix} 0 \\ 3 \\ -3 \end{pmatrix}. \tag{21}$$

After normalization, this finally yields

$$|E_1\rangle = \frac{1}{\sqrt{6}} \begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix},$$

$$|E_2\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}. \tag{22}$$

These two basis vectors belong to different columns of E, and are therefore already orthogonal.

Alternate Approach: IRREP E

- The above approach using column projectors is elegant because it does not require extra steps such as orthogonalization to construct the basis vectors of the E IRREP space. However, sometimes one does not have easy access to IRREP matrices. In such a case, it is possible to use only the character projection operator if one is willing to do extra work.

- The character projection operator for IRREP E is

$$\Pi_E = \frac{2}{6} \sum_{a \in C_{3v}} \chi_E^* P(a) = \frac{1}{3} \begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix}. \tag{23}$$

It is easy to check for us here that this is the sum of the two column projectors. Also, the trace of Π_E equals the IRREP dimension d_α times n_α as expected.

- We could proceed as in the non-degenerate case and project an arbitrary non-zero vector in the REP space using Π_E, but this will produce only one of the two required basis vectors for the E IRREP. The second one may be generated by repeating the projection beginning from a different vector, followed by orthogonalization of the two resulting vectors which are generally not orthogonal. While this method works fine, it is error-prone and becomes tedious for higher-dimensional IRREPS (and especially if an IRREP has multiplicity $n_\alpha > 1$.)

- A similar but often more convenient approach is based on the spectral representation of the IRREP projection operator: Since projectors are Hermitian operators with eigenvalues 0 or 1, the eigenvectors satisfying the eigenvalue problem

$$\Pi_\alpha |\alpha, k, n\rangle = |\alpha, k, n\rangle \tag{24}$$

span the IRREP space belonging to IRREP α, and (due to the Hermitian property) are also diagonal. Here k again lables the "column" of the IRREP, whereas n is an additional index taking into account IRREP multiplicies > 1 (not relevant here).

- The numpy function `numpy.linalg.eigh` may be used to diagonalize the projector matrix. The eigenvectors belonging to eigenvalues 1 then are the desired basis for IRREP E:

```python
# Set up E projector
pi_e = np.zeros((3,3))

for i in range(6):
    pi_e = pi_e + irrep_char[2,i]*p[i]

pi_e = pi_e/3  # Note extra factor 2 for 2-dimensional IRREP

# Diagonalize pi_e

w, v = np.linalg.eigh(pi_e)

# The desired eigenvectors are v[1] and v[2]
```

- The basis vectors spanning a degenerate IRREP may be arbitrarily rotated against each other, so don't be suprised if you do not get the "nice" vectors we obtained with column projectors. As long as they are eigenvectors of Π_E with eigenvalue 1 they are equally correct.
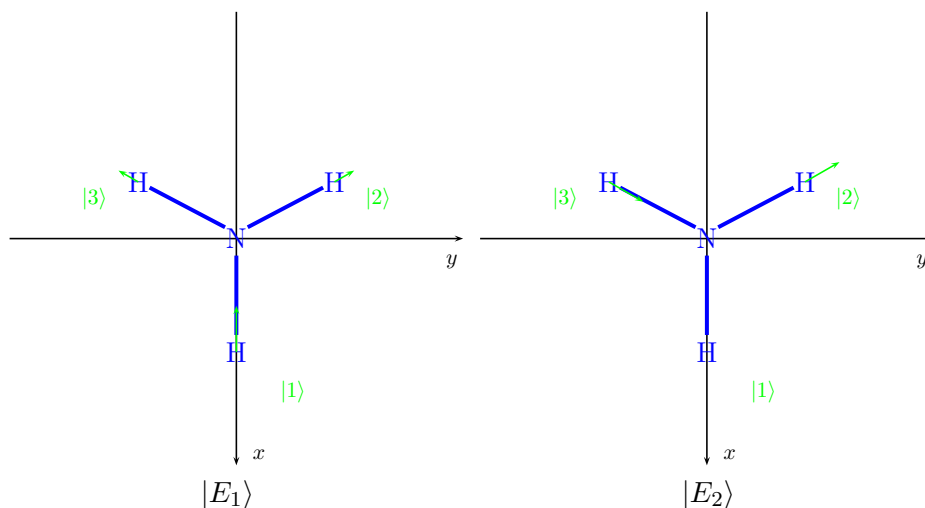

Figure 9: The two symmetry adapted N-H stretching vibrations of NH_3 transforming according to the first and second column of IRREP E.

Graphical Representation

- Fig. 9 shows a graphical representation of the two E symmetry adapted basis vectors.

18

- You may check these results using Python, following the example for the A_1 IRREP, but using the IRREP matrices instead of the characters.

Basis Transformations

- We may construct a transformation matrix

$$\mathbf{U} = |A_1\rangle\langle 1| + |E_1\rangle\langle 2| + |E_2\rangle\langle 3| = \frac{1}{\sqrt{6}}\begin{pmatrix} \sqrt{2} & -2 & 0 \\ \sqrt{2} & 1 & \sqrt{3} \\ \sqrt{2} & 1 & -\sqrt{3} \end{pmatrix} \tag{25}$$

  which transforms a vector from the reducible into the irreducible basis. Since the columns of $\mathbf{U}$ are orthonormal, $\mathbf{U}^\dagger\mathbf{U} = \mathbf{1}$, i.e., it is unitary.

- The representation matrices are transformed from the reducible to the irreducbile basis according to $P'(a) = \mathbf{U}P(a)\mathbf{U}^\dagger$. This yields:

$$P'(E) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad P'(C_3) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -\frac{1}{2} & -\frac{\sqrt{3}}{2} \\ 0 & \frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix}, \quad P(C_3^2) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -\frac{1}{2} & \frac{\sqrt{3}}{2} \\ 0 & -\frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix},$$

$$P(\sigma_1) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \quad P(\sigma_2) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -\frac{1}{2} & -\frac{\sqrt{3}}{2} \\ 0 & \frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix}, \quad P(\sigma_3) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -\frac{1}{2} & -\frac{\sqrt{3}}{2} \\ 0 & -\frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix}. \tag{26}$$

- As expected, the resulting representation matrices are the direct sums of the IRREP matrices, i.e.,

$$P'(a) = P_{A_1}(a) \oplus P_E(a), \tag{27}$$

  where the one-dimensional IRREP matrices of the A_1 IRREP coincide with the characters.

2.2   Perturbation Theory

2.2.1   Purpose

- Determines behavior of an observable quantity Q as function of parameter λ in the neigborhood of λ_0.

- λ is often chosen to be dimensionless.

- Assumptions:

  1. Exact solution at $\lambda = \lambda_0$ is available
  2. $Q(\lambda)$ is sufficiently smooth
  3. $|\lambda - \lambda_0|$ is small

- Perturbation theory is equivalent to power series expansion

$$Q(\lambda) = Q^{(0)} + \lambda - \lambda_0 Q^{(1)} + (\lambda - \lambda_0)^{(2)} Q^{(2)} + O((\lambda - \lambda_0)^3). \tag{28}$$

Comparison to Taylor series:

$$Q^{(0)} = Q(\lambda)|_{\lambda=\lambda_0}, \quad Q^{(1)} = \left.\frac{\partial Q(\lambda)}{\partial \lambda}\right|_{\lambda=\lambda_0}, \quad Q^{(2)} = \frac{1}{2}\left.\frac{\partial^2 Q(\lambda)}{\partial \lambda^2}\right|_{\lambda=\lambda_0}, \ldots \tag{29}$$

- In practice, this series is asymptotic, i.e.,

$$\lim_{\lambda \to 0} \frac{1}{(\lambda - \lambda_0)^n} \left(Q(\lambda) - \sum_{k=0}^{n} (\lambda - \lambda)^k Q^{(k)} \right) = 0. \tag{30}$$

- Applications:

   1. Molecular properties ("exact" perturbation theory)
   2. Qualitative behavior
   3. Numerical approximations (often poor)

Observable	Perturbation(s)	n	Property	Type
Energy	Uniform electric field	1	Electric dipole moment	1
Energy	Uniform electric field gradient	1	Electric quadrupole moment	1
Energy	General one-electron perturbation	1	Electronic density	1
Energy	Nuclear displacement	Nuclear forces	1	
Energy	Uniform electric field	2	Polarizability	1
Energy	Nuclear magnetic moment, uniform electric field	(1,1)	NMR shielding	1
Energy	Remote molecular fragment	mostly 1, 2	Intermolecular forces	1-3
Orbital energy	Chemical perturbation	1, 2	Chemical bonding	2
Energy	Electron interaction	mostly 1-4	Electron correlation	3

Table 10: Some applications of perturbation theory. n refers to the order of perturbation theory, and "Type" refers to the 3 different types of applications described in the text.

- Quantum mechanics: Solve

$$(\hat{H}^{(0)} + \lambda \hat{H}^{(1)})|\psi_\lambda\rangle = E_\lambda|\psi_\lambda\rangle \tag{31}$$

for normalized $|\psi_\lambda\rangle$ in the vicinity of $\lambda = 0$.

2.2.2   Non-Degenerate States

Zeroth Order

- Unperturbed system:

$$\hat{H}^{(0)}|\phi_i\rangle = \epsilon_i|\phi_i\rangle, \tag{32}$$

$\langle\phi_i|\phi_j\rangle = \delta_{ij}$.

- Zero-order energies, wavefunctions:

$$E_i^{(0)} = \langle\phi_i|\hat{H}^{(0)}|\phi_i\rangle = \epsilon_i, |\psi_i^{(0)}\rangle = |\phi_i\rangle \tag{33}$$

- Zero-order properties:

$$Q_i^{(0)} = \langle\phi_i|\hat{Q}|\phi_i\rangle \tag{34}$$

First Order

- First-order energy:

$$E_i^{(1)} = \langle\phi_i|\hat{H}^{(1)}|\phi_i\rangle \tag{35}$$

- First-order wavefunction:

$$|\psi_i^{(1)}\rangle = \sum_{j\neq i} \frac{\langle\phi_j|\hat{H}^{(1)}|\phi_i\rangle}{\epsilon_i - \epsilon_j}|\phi_j\rangle \tag{36}$$

- First-order properties:

$$Q_i^{(1)} = \sum_{j\neq i} \frac{\langle\phi_i|\hat{Q}|\phi_j\rangle\langle\phi_j|\hat{H}^{(1)}|\phi_i\rangle}{\epsilon_i - \epsilon_j} \tag{37}$$

Second Order

- Second-order energy:

$$E_i^{(2)} = \sum_{j\neq i} \frac{|\langle\phi_j|\hat{H}^{(1)}|\phi_i\rangle|^2}{\epsilon_i - \epsilon_j} \tag{38}$$

2.2.3   Degenerate States

- For degenerate states, $\hat{H}^{(1)}$ needs to be diagonalized in the degenerate eigenspace(s). This may cause level splitting to first or higher orders.

- Analogous to non-degenerate perturbation theory once a suitable degenerate subspace basis has been chosen; states in the same zero-order eigenspace are excluded in the first- and higher wavefunctions.

2.2.4 Qualitative Behavior

- The effect of a perturbation on first-order energies is a shift. Degenerate energy levels may split.

- To second order, the energy change due to interactions between states $|\phi_i\rangle$ and $|\phi_j\rangle$ is proportional to

$$\frac{|\langle \phi_j|\hat{H}^{(1)}|\phi_i\rangle|^2}{\epsilon_i - \epsilon_j} \tag{39}$$

- For property Q, the change is first-order and proportional to

$$\frac{\langle \phi_i|\hat{Q}|\phi_j\rangle \langle \phi_j|\hat{H}^{(1)}|\phi_i\rangle}{\epsilon_i - \epsilon_j}. \tag{40}$$

2.3 Born-Oppenheimer Approximation

2.3.1 Molecular Hamiltonian

- Molecules consist of electrons and nuclei, not atoms.

> I feel that chemical binding has not one nature, but many. Also, I do not wish to be necessarily committed to the familiar concept that molecules are composed of atoms, in other words, that atoms are still atoms when they have formed molecules. (R. S. Mulliken, 1975)

- Molecular Hamiltonian:

$$\hat{H}_{\text{mol}} = \hat{T}_{\text{e}} + \hat{T}_{\text{n}} + \hat{V}_{\text{ee}} + \hat{V}_{\text{en}} + \hat{V}_{\text{nn}} \tag{41}$$

- Kinetic energy operator of the electrons:

$$\hat{T}_{\text{e}} = \sum_i \frac{\mathbf{p}_i^2}{2m_e} \tag{42}$$

- Kinetic energy operator of the nuclei:

$$\hat{T}_{\text{n}} = \sum_I \frac{\mathbf{p}_I^2}{2m_I} \tag{43}$$

- Operator of electron-electron repulsion:

$$\hat{V}_{\text{ee}} = \frac{1}{2} \sum_{\substack{ij \\ i \neq j}} \frac{e^2}{|\mathbf{r}_i - \mathbf{r}_j|} \tag{44}$$

- Operator of electron-nucleus attraction:

$$\hat{V}_{\text{en}} = -\sum_{iI} \frac{Z_I e^2}{|\mathbf{r}_i - \mathbf{r}_I|} \tag{45}$$

- Operator of nucleus-nucleus repulsion:

$$\hat{V}_{\text{ee}} = \frac{1}{2} \sum_{\substack{IJ \\ I \neq J}} \frac{Z_I Z_J e^2}{|\mathbf{r}_I - \mathbf{r}_J|} \tag{46}$$

2.3.2   Adiabatic Separation of Electron and Nuclear Motion

- Nuclear kinetic energy operator is smaller than electronic kinetic energy operator by a factor of

$$\frac{m_e}{m_p} \simeq \frac{1}{1836} \tag{47}$$

  or less.

- Born-Oppenheimer (BO) approximation: Ignore $\hat{T}_{\text{n}}$ and define electronic (or clamped nuclei) Hamiltonian

$$\hat{H}_{\text{el}}(X) = \hat{T}_{\text{e}} + \hat{V}_{\text{ee}} + \hat{V}_{\text{en}}(X) + V_{\text{nn}}(X). \tag{48}$$

  $\hat{H}_{\text{el}}$ depends parametrically on nuclear coordinates X.

- Corresponds to adiabatic separation of the molecular wavefunction,

$$\Psi_{\text{mol}}(x, X) = \Psi_{\text{e}}(x|X)\Phi_{\text{n}}(X). \tag{49}$$

- Eigenvalues $E_{\text{el}}(X)$ of $\hat{H}_{\text{el}}(X)$ are BO molecular potential energy surfaces (PESs). There is one surface for each electronic state. PESs describe the potential energy "felt" by each nucleus in a molecule due to the electrons and the other nuclei.

- BO PESs are $3N_{\text{n}} - 6$-dimensional for non-linear and $3N_{\text{n}} - 5$-dimensional for linear molecules. N_{n} is number of nuclei.

- Nuclear BO Hamiltonian:

$$\hat{H}_{\text{n}}(X) = \hat{T}_n(X) + E_{\text{el}}(X) \tag{50}$$

- Born-Huang approximation: Add diagonal BO correction

$$E_{\text{BH}}(X) = \langle \Psi_{\text{e}}(X)|\hat{T}_n(X)|\Psi_{\text{e}}(X)\rangle \tag{51}$$

  to BO PES.

2.3.3   Beyond Born-Oppenheimer

- If electronic energy differences become comparable to nuclear kinetic energies, BO approximation breaks down

- Important close to conical intersections of PESs ("photochemical funnels")

- Close to conical intersections, electronic and nuclear motions are strongly coupled.
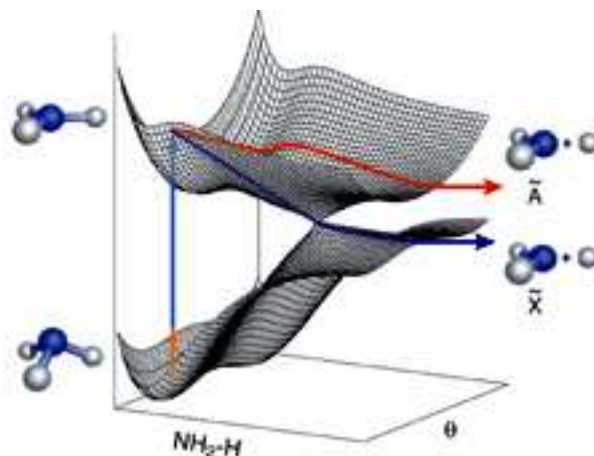
Figure 10: Two-dimensional potential energy surface of NH_3 [4] showing a conical intersection.

- Important special case: Jahn-Teller effect. There exists an energy-lowering geometric distortion in non-linear molecules with a spatially degenerate state.

- Nonadiabatic molecular dynamics: Include coupling between BO PESs due to finite nuclear kinetic energy. Can describe excited state deactivation, non-radiative transitions, photochemical reactions.

2.4 Molecular Orbital Theory

2.4.1 Hückel Theory

- Hückel molecular orbital (HMO) theory is a semi-empirical one-electron theory for molecules whose properties are dominated by π electrons

- Basic assumptions:

    1. Only π electrons are considered
    2. The Hückel molecular orbitals ϕ_i are linear combinations of atomic orbitals χ_μ (LCAO-MO ansatz),
    $$\phi_i = \sum_\mu c_{\mu i} \chi_\mu. \tag{52}$$

    3. The basis $\{\chi_\mu\}$ consists of one p orbital per atom (minimal basis) for planar hydrocarbons. Differential overlap is neglected (CNDO):

    $$S_{\mu\nu} = \int d^3r\, \chi_\mu^*(\mathbf{r})\chi_\nu(\mathbf{r}) = \delta_{\mu\nu}. \tag{53}$$

    4. The Hamiltonian in the basis $\{\chi_\mu\}$ has the form

    $$\mathbf{H} = \alpha\mathbf{1} + \beta\mathbf{M}. \tag{54}$$

24

The topological matrix $\mathbf{M}$ depends on connectivity only:

$$M_{\mu\nu} = \begin{cases} 1 & \mu, \nu \text{ nearest neighbors} \\ 0 & \text{else} \end{cases} \tag{55}$$

5. α and β are parameters. α is usually taken to be the negative p ionization energy of C. β is often determined from electronic spectroscopy or thermochemistry.

6. The resulting HMOs are occupied according to the Aufbau principle.

- Population analysis: Partitioning of electron charge:

$$q_\mu = \sum_i n_i c_{\mu i}^2, \tag{56}$$

where n_i is the occupation number of the i-th orbital, is the π electron population or charge of atom μ.

$$p_{\mu\nu} = \sum_i n_i c_{\mu i} c_{\nu i}, \quad \mu \neq \nu, \tag{57}$$

is the π bond order of atoms μ and ν. *Warning:* Populations and bond orders are not observable

2.4.2 Example: Cyclopropenyl cation

Figure 11: Structure of cyclopropenyl cation, $C_3H_3^+$.

- Topological matrix:

$$\mathbf{M} = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \tag{58}$$

- Determine eigenvalues and vectors of $\mathbf{M}$ only, since

$$\mathbf{M}\mathbf{c}_i = \lambda_i \mathbf{c}_i \tag{59}$$

implies that

$$\mathbf{H}\mathbf{c}_i = (\alpha \mathbf{1} + \beta \mathbf{M})\mathbf{c}_i = (\alpha + \beta \lambda_i)\mathbf{c}_i = \epsilon_i \mathbf{c}_i. \tag{60}$$

Thus, $\mathbf{c}_i$ is also eigenvector of $\mathbf{H}$, and the corresponding orbital energy is $\epsilon_i = \alpha + \beta \lambda_i$.

- Eigenvalues and eigenvectors:

$$\lambda_1 = 2, \quad \epsilon_1 = \alpha + 2\beta, \quad \phi_1 = \frac{1}{\sqrt{3}}(\chi_1 + \chi_2 + \chi_3) \tag{61}$$

$$\lambda_2 = -1, \quad \epsilon_2 = \alpha + -\beta, \quad \phi_2 = \frac{1}{\sqrt{2}}(\chi_1 - \chi_3) \tag{62}$$

$$\lambda_3 = -1, \quad \epsilon_3 = \alpha + -\beta, \quad \phi_3 = \frac{1}{\sqrt{6}}(\chi_1 - 2\chi_2 + \chi_3) \tag{63}$$

- π-populations and bond orders:

$$q_1 = q_2 = q_3 = \frac{2}{3}, \quad p_{12} = p_{23} = p_{13} = \frac{2}{3} \tag{64}$$

- Total π electron energy (from orbital energies):

$$E_\pi = 2(\alpha + 2\beta) \tag{65}$$

- Delocalization or resonance energy: Energy difference relative to structure with one localized π bond,

$$E_{\text{res}} = E_\pi - 2(\alpha + \beta) = 2\beta. \tag{66}$$

Cyclopropenyl cation is aromatic.

2.4.3 Linear Polyenes and Polymethines

- $C_n H_{n+2}$; n even: polyenes, n odd: polymethines

- Topological matrix:

$$\mathbf{M} = \begin{pmatrix} 0 & 1 & & & \mathbf{0} \\ 1 & 0 & & & \\ & & \ddots & & \\ & & & 0 & 1 \\ \mathbf{0} & & & 1 & 0 \end{pmatrix} \tag{67}$$

- Resulting HMO energies:

$$\epsilon_j = \alpha + 2\beta \cos \frac{j\pi}{n+1}, \quad j = 1, \ldots, n \tag{68}$$

- Visualization: Frost semicircle, see Fig. 12

- MO coefficients:

$$c_{\mu j} = \sqrt{\frac{2}{n+1}} \sin \frac{j\mu\pi}{n+1} \tag{69}$$

- Visualization: see Fig. 13

- Uses: UV spectroscopy, theory of pericyclic reactions

Figure 12: Frost diagram for linear polyenes und polymethines.

Figure 13: HMOs of butadiene.

2.4.4 Annulenes

- C_nH_n, cyclic

- Topological matrix of Hückel annulenes:

$$\mathbf{M} = \begin{pmatrix} 0 & 1 & 0 & \ldots & 1 \\ 1 & 0 & 0 & \ldots & 0 \\ & & \ddots & & \\ 0 & 0 & \ldots & 0 & 1 \\ 1 & 0 & \ldots & 1 & 0 \end{pmatrix} \tag{70}$$

- HMO energies (not ordered):

$$\epsilon_j = \alpha + 2\beta \cos\frac{2\pi j}{n}, \quad j = 1, \ldots, n \tag{71}$$

- Visualization: Frost circle, see Fig. 14

Figure 14: Frost circle for Hückel annulenes.

- Stable closed-shell ground state: $2, 6, 10, \ldots = 4k + 2$ electrons. If there is one π electron per C this leads to Hückels $4n + 2$ rule for aromatic hydrocarbons

- Topological matrix for Möbius annulenes:

$$
\mathbf{M} = \begin{pmatrix}
0 & 1 & 0 & \ldots & -1 \\
1 & 0 & 0 & \ldots & 0 \\
 & & \ddots & & \\
0 & 0 & \ldots & 0 & 1 \\
-1 & 0 & \ldots & 1 & 0
\end{pmatrix}
\tag{72}
$$

- HMO energies:

$$
\epsilon_j = \alpha + 2\beta \cos \frac{(2j-1)\pi}{n}, \quad j = 1, \ldots, n
\tag{73}
$$

- Visualization: Frost circle, see Fig. 15

Figure 15: Frost circle for Möbius annulenes.

- Stable closed-shell ground state: $4, 8, 12, \ldots = 4k$ electrons. If there is one π electron per C this leads to $4n$ rule for Möbius annulenes. See Ref. [5] for examples.

2.4.5 Extensions of Hückel Theory

- Hückel method for heteroatoms: Change α, β for heteroatoms X and C-X bonds:

$$
\alpha_X = \alpha + h\beta, \quad \beta_{CX} = k\beta.
\tag{74}
$$

h and k are dimensionless "heteroatom parameters" that depend on the oxidation state and hybridization. For example, $h = 0.5, k = 0.8$ for N in pyridine, but $h = 1.5, k = 1.0$ in pyrrole.

- Extended Hückel (EHT) method:

  – Use minimal valence basis of Slater-type basis functions (STOs).

  – Drop CNDO and compute complete overlap matrix.

  – Assign diagonal elements $H_{\mu\mu}$ using experimental valence ionization potentials of atoms.

– Offdiagonal elements: Wolfsberg-Helmholtz approximation

$$H_{\mu\nu} = K \frac{H_{\mu\mu} + H_{\nu\nu}}{2} S_{\mu\nu}, \tag{75}$$

where $K = 1.75$ (Wolfsberg-Helmholtz constant).

– The EHT MO coefficients are obtaine from the generalized eigenvalue problem

$$\mathbf{HC} = \mathbf{SC}\boldsymbol{\epsilon}, \tag{76}$$

where $\epsilon_{ij} = \epsilon_i \delta_{ij}$ is diagonal.

- Related semiempirical MO theories: CNDO/2, Fenske-Hall, MNDO , DFTB

- Pariser-Parr-Pople method: Adds nearest neigbor electron interaction term ("Hubbard U")

$$\hat{U}_{\mu\nu\kappa\lambda} = U \delta_{\mu\nu} \delta_{\kappa\lambda} \tag{77}$$

to Hückel Hamiltonian.

2.4.6 When to Use Hückel Theory

- Hückel theory is best used for

  – qualitative analysis, e.g., size-dependent behavior
  – analysis of more advanced calculations (ab-initio calculation of α, β)
  – generation of initial MOs for self-consistent field calculations

- Hückel theory should be avoided for

  – polar and ionic bonds
  – small-gap and transition metal compounds
  – excited states, electronic properties
  – quantitative calculations

3 Hartree-Fock and Density Functional Theory

3.1 Hartree-Fock Method

3.1.1 Key Assumptions

1. The N-electron wavefunction is approximated by a single Slater determinant,

$$\Phi(x_1, \ldots, x_N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \phi_1(x_1) & \phi_1(x_2) & \ldots & \phi_1(x_N) \\ \phi_2(x_1) & \phi_2(x_2) & \ldots & \phi_2(x_N) \\ \ldots & \ldots & \vdots & \ldots \\ \phi_N(x_1) & \phi_N(x_2) & \ldots & \phi_N(x_N) \end{vmatrix}. \tag{78}$$

The orbitals $\phi_1(x) \ldots \phi_N(x)$ are orthonormal. $x = (\mathbf{r}, \sigma)$ is short-hand for single-electron spin-space coordinates.

2. The ground-state energy is obtained by minimizing the energy expectation value

$$E[\phi_1, \ldots, \phi_N] = \langle \Phi | \hat{H} | \Phi \rangle \tag{79}$$

of the physical N-electron Hamiltonian with respect to the orbitals. Thus, Φ is variationally optimal, i.e., it is *the* Slater determinant with the lowest energy expectation value.

3.1.2 Slater Determinants and Their Matrix Elements

- Axiomatic definition: $f : \mathbf{A} \in \mathbf{C}^{n \times n} \to f(\mathbf{A}) \in \mathbf{C}$ is called determinant of $\mathbf{A}$, i.e., $f(\mathbf{A}) = \det(\mathbf{A}) = |\mathbf{A}|$, if and only if

  (i) f is linear in each column, i.e.,

$$f \begin{pmatrix} \mathbf{a}_1 \\ \vdots \\ c\mathbf{a}_k + \mathbf{b} \\ \vdots \\ \mathbf{a}_n \end{pmatrix} = cf \begin{pmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_k \\ \vdots \\ \mathbf{a}_n \end{pmatrix} + f \begin{pmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{b} \\ \vdots \\ \mathbf{a}_n \end{pmatrix}$$

  $\forall k, 1 \leq k \leq n, c \in \mathbf{C}, \mathbf{b} \in \mathbf{C}^n$.

  (ii) f is antisymmetric under column exchange, i.e.,

$$f \begin{pmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_k \\ \vdots \\ \mathbf{a}_l \\ \vdots \\ \mathbf{a}_n \end{pmatrix} = -f \begin{pmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_l \\ \vdots \\ \mathbf{a}_k \\ \vdots \\ \mathbf{a}_n \end{pmatrix}$$

  $\forall k, l, 1 \leq k, l, \leq n$.

  (iii) $f(\mathbf{1}) = 1$

- Symmetric group representation of the determinant ("Leibniz formula"):

$$|\mathbf{A}| = \sqrt{n!} \, \hat{\mathcal{A}} \, a_{11} \cdots a_{nn}, \tag{80}$$

where

$$\hat{\mathcal{A}} = \frac{1}{n!} \sum_{P \in \mathrm{S}_n} \mathrm{sgn}(P) \hat{P} \tag{81}$$

is the antisymmetrization operator or antisymmetrizer. P denotes permutations of n columns, i.e., identical particles, which form the symmetric group S_n represented by $\hat{P}$.

$$\mathrm{sgn}(P) = \begin{cases} +1; & P \text{ even,} \\ -1; & P \text{ odd,} \end{cases} \tag{82}$$

One-electron operators	Two-electron operators
$\hat{V} = \sum_{i=1}^{N} \hat{v}(i)$	$\hat{W} = \frac{1}{2} \sum_{i,j=1}^{N} \hat{w}(i,j)$
$\langle \Phi \vert \hat{V} \vert \Phi \rangle = \sum_{i=1}^{N} \langle \phi_i \vert \hat{v} \vert \phi_i \rangle$	$\langle \Phi \vert \hat{W} \vert \Phi \rangle = \frac{1}{2} \sum_{i,j=1}^{N} (\langle 1:\phi_i, 2:\phi_j \vert \hat{w} \vert 1:\phi_i, 2:\phi_j \rangle$ $-\langle 1:\phi_i, 2:\phi_j \vert \hat{w} \vert 1:\phi_j, 2:\phi_i \rangle)$

Table 11: A subset of the Slater-Condon rules for computing expectation values of Slater determinants

where P is even/odd if it contains an even/odd number of transpositions, i.e., interchanges of 2 columns/particles.

- The properties of the determinant, together with the orthonormality of the one-particle basis $\{\vert \phi_p \rangle\}$, may be used to derive the rules for expectation values of Slater Determinants (Slater-Condon-Rules), see Tab. 11.

3.1.3 Unitary Invariance

- Consider a unitary transformation of the occupied orbitals into themselves,

$$\tilde{\phi}_i(x) = \sum_j U_{ij} \phi_j(x), \tag{83}$$

where $\mathbf{U}^\dagger \mathbf{U} = 1$.

- Slater determinant

$$
\begin{aligned}
\tilde{\Phi}(x_1, \ldots, x_N) &= \frac{1}{\sqrt{N!}} \begin{vmatrix} \tilde{\phi}_1(x_1) & \tilde{\phi}_1(x_2) & \ldots & \tilde{\phi}_1(x_N) \\ \tilde{\phi}_2(x_1) & \tilde{\phi}_2(x_2) & \ldots & \tilde{\phi}_2(x_N) \\ \ldots & \ldots & \vdots & \ldots \\ \tilde{\phi}_N(x_1) & \tilde{\phi}_N(x_2) & \ldots & \tilde{\phi}_N(x_N) \end{vmatrix} \\
&= \frac{1}{\sqrt{N!}} \left| \mathbf{U} \begin{pmatrix} \phi_1(x_1) & \phi_1(x_2) & \ldots & \phi_1(x_N) \\ \phi_2(x_1) & \phi_2(x_2) & \ldots & \phi_2(x_N) \\ \ldots & \ldots & \vdots & \ldots \\ \phi_N(x_1) & \phi_N(x_2) & \ldots & \phi_N(x_N) \end{pmatrix} \right| \\
&= \vert \mathbf{U} \vert \Phi(x_1, \ldots, x_N) = e^{i\phi} \Phi(x_1, \ldots, x_N).
\end{aligned} \tag{84}
$$

Thus, the Slater determinants of the orbitals $\{\phi_i\}$ and the orbitals $\{\tilde{\phi}_i\}$ are identical up to a complex phase factor.

- The shape of individual occupied orbitals is not unique!

- Unitarily equivalent sets of orbitals (e.g. localized vs. delocalized, banana vs. π bonds) are merely different choices of basis for the occupied orbitals and yield identical observables.

3.1.4 HF Energy Functional

- In the electronic Hamiltonian

$$\hat{H} = \hat{T}_e + \hat{V}_{en} + \hat{V}_{ee} + V_{nn}, \tag{85}$$

$\hat{T}_e$ and $\hat{V}_{en}$ are one-electron, $\hat{V}_{ee}$ is two-electron, and V_{nn} is (one-electron) constant.

- Thus, the HF energy expectation value as a functional of the orbitals is

$$E^{\mathrm{HF}}[\phi_1, \ldots, \phi_N] = \langle \Phi | \hat{H} | \Phi \rangle = \sum_{i=1}^{N} \langle \phi_i | \hat{h} | \phi_i \rangle + \frac{1}{2} \sum_{ij} (\langle 1 : \phi_i, 2 : \phi_j | \hat{v}_{\mathrm{ee}}(1,2) | 1 : \phi_i, 2 : \phi_j \rangle$$
$$- \langle 1 : \phi_i, 2 : \phi_j | \hat{v}_{\mathrm{ee}}(1,2) | 1 : \phi_j, 2 : \phi_i \rangle) + V_{nn}. \tag{86}$$

- Notation for two-electron repulsion integrals:

  - Dirac notation:

$$\langle ij | kl \rangle = \langle 1 : \phi_i, 2 : \phi_j | \hat{v}_{\mathrm{ee}}(1,2) | 1 : \phi_k, 2 : \phi_l \rangle$$
$$= \int dx_1 dx_2 \frac{\phi_i^*(x_1) \phi_j^*(x_2) \phi_k(x_1) \phi_l(x_2)}{|\mathbf{r}_1 - \mathbf{r}_2|} \tag{87}$$

  - Double bar Dirac notation:

$$\langle ij \| kl \rangle = \langle ij | kl \rangle - \langle ij | lk \rangle \tag{88}$$

  - Mulliken notation:

$$(ij | kl) = \langle ik | jl \rangle = \int dx_1 dx_2 \frac{\phi_i^*(x_1) \phi_j(x_1) \phi_k^*(x_2) \phi_l(x_2)}{|\mathbf{r}_1 - \mathbf{r}_2|} \tag{89}$$

3.1.5 Evaluating the UHF Energy Functional

- Spin-unrestricted HF (UHF) approximation: Each orbital has either α or β spin, see Fig. 16,

$$\phi_i(x) = \phi_{k(i)\sigma(i)}(\mathbf{r}) \delta_{\sigma\sigma(i)}. \tag{90}$$

$\sigma(i)$: Spin of orbital no. i, $k(i)$: number of orbital no. i accounting for spin.

- It is customary to use the index pair (k, σ) instead of i. The total number of σ-spin electrons is N_σ.

- Closed-shell spin-restricted (RHF) approximation: UHF α and β orbitals are assumed to be identical,

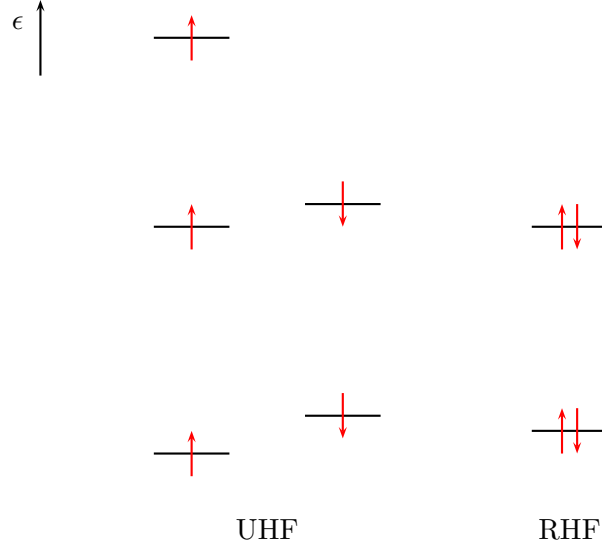$$\phi_{k\alpha}(\mathbf{r}) = \phi_{k\beta}(\mathbf{r}). \tag{91}$$

Figure 16: UHF and RHF orbital energy diagrams

- LCAO-MO expansion of UHF orbitals:

$$\phi_{i\sigma}(\mathbf{r}) = \sum_{\mu} C_{\mu i\sigma} \chi_{\mu}(\mathbf{r}). \tag{92}$$

There are two sets of MO coefficients, $\mathbf{C}_{\alpha}$ and $\mathbf{C}_{\beta}$. Each one satisfies their own orthonormality condition,

$$\mathbf{C}_{\sigma}^{\dagger}\mathbf{S}\mathbf{C}_{\sigma} = \mathbf{1}. \tag{93}$$

- UHF density matrix:

$$D_{\mu\nu\sigma} = \sum_{i=1}^{N_{\sigma}} C_{\mu i\sigma} C_{\nu i\sigma}, \tag{94}$$

i.e., there is a separate density matrix for each spin. In RHF, the α and β spin density matrices are identical.

- UHF total energy functional:

$$E^{\mathrm{UHF}}[\mathbf{C}_{\alpha}, \mathbf{C}_{\beta}] = \sum_{\sigma\mu\nu} h_{\mu\nu} D_{\nu\mu\sigma} + \sum_{\sigma\sigma'\mu\nu\kappa\lambda} (\mu\nu|\kappa\lambda)\Gamma_{\lambda\kappa\sigma'\nu\mu\sigma} \tag{95}$$

- UHF two-electron density matrix:

$$\Gamma_{\lambda\kappa\sigma'\nu\mu\sigma} = \frac{1}{2}\left(D_{\lambda\kappa\sigma'} D_{\nu\mu\sigma} - \delta_{\sigma\sigma'} D_{\lambda\mu\sigma} D_{\nu\kappa\sigma'}\right) \tag{96}$$

- Only $h_{\mu\nu}$ with χ_{μ} close to χ_{ν} and $(\mu\nu|\kappa\lambda)$ with χ_{μ} close to χ_{ν} and χ_{κ} close to χ_{λ} need to be computed. For given $\mathbf{D}_{\sigma}$, evaluating the HF energy functional scales at most as $O(N^2)$.

33

3.1.6 UHF Lagrangian

- The variational principle applies to $E^{\text{UHF}}[\mathbf{C}_\alpha, \mathbf{C}_\beta]$ only if the corresponding UHF MOs are orthonormal.

- To explicitly account for orthonormality (ON) constraint, define UHF Lagrangian

$$L^{\text{UHF}}[\mathbf{C}_\alpha, \mathbf{C}_\beta, \boldsymbol{\epsilon}_\alpha, \boldsymbol{\epsilon}_\beta] = E^{\text{UHF}}[\mathbf{C}_\alpha, \mathbf{C}_\beta] - \sum_{ij\sigma} \epsilon_{ij\sigma} \left(\mathbf{C}_{i\sigma}^\dagger \mathbf{S} \mathbf{C}_{j\sigma} - \delta_{ij} \right). \tag{97}$$

$\boldsymbol{\epsilon}_\sigma$ is an $N_\sigma \times N_\sigma$ Lagrange multiplier matrix.

- L^{UHF} may be minimized without implicit constraints, and is a minimum for the exact UHF orbitals.

- Stationarity with respect to $\boldsymbol{\epsilon}_\sigma$ recovers the ON constraint, while stationarity with respect to $\mathbf{C}_\sigma$ gives rise to the UHF equations.

- L^{UHF} depends on the MO coefficients through $\mathbf{D}_\sigma$ and the energy-weighted density matrix

$$\mathbf{W}_\sigma = \mathbf{C}\boldsymbol{\epsilon}\mathbf{C}^\dagger. \tag{98}$$

- Final UHF Lagrangian:

$$L^{\text{UHF}}[\mathbf{C}_\alpha, \mathbf{C}_\beta, \boldsymbol{\epsilon}_\alpha, \boldsymbol{\epsilon}_\beta] = \sum_\sigma \text{tr}(\mathbf{h}\mathbf{D}_\sigma) + \sum_{\substack{\mu\nu\kappa\lambda \\ \sigma\sigma'}} (\mu\nu|\kappa\lambda) \Gamma_{\lambda\kappa\sigma'\nu\mu\sigma}[\mathbf{D}_\alpha, \mathbf{D}_\beta]$$
$$- \sum_\sigma \text{tr}(\mathbf{S}\mathbf{W}_\sigma - \boldsymbol{\epsilon}_\sigma) \tag{99}$$

3.1.7 Roothaan-Hall Equations

- The UHF Lagrangian is a complicated functional of $\mathbf{C}$. We thus use the chain rule by computing

$$\frac{\partial D_{\mu\nu\sigma'}}{\partial C_{vi\sigma}^*} = \delta_{\sigma'\sigma}\delta_{\nu\upsilon}C_{\mu i\sigma'} \tag{100}$$

and

$$\frac{\partial W_{\mu\nu\sigma'}}{\partial C_{vi\sigma}^*} = \delta_{\sigma'\sigma}\delta_{\nu\upsilon}(\mathbf{C}\boldsymbol{\epsilon})_{\mu i\sigma'}, \tag{101}$$

as well as

$$\frac{\partial W_{\mu\nu\sigma'}}{\partial \epsilon_{ij\sigma}} = \delta_{\sigma'\sigma}C_{\mu i\sigma'}C_{\nu j\sigma'}^*. \tag{102}$$

Here and in the following, indices $i, j, \ldots$ denote occupied, $a, b, \ldots$ virtual, and $p, q, \ldots$ general MOs. Further, we are taking Wirtiger derivatives, i.e., $\mathbf{C}$ and $\mathbf{C}^\dagger$ are considered linearly independent.

- Stationarity conditions:

1. MO coefficients:

$$\frac{\partial L^{\text{UHF}}}{\partial C^*_{vi\sigma}} = \sum_{\mu\nu\sigma'} \left(\frac{\partial L^{\text{UHF}}}{\partial D_{\mu\nu\sigma'}} \frac{\partial D_{\mu\nu\sigma'}}{\partial C^*_{vi\sigma}} + \frac{\partial L^{\text{UHF}}}{\partial W_{\mu\nu\sigma'}} \frac{\partial W_{\mu\nu\sigma'}}{\partial C^*_{vi\sigma}} \right) = 0 \tag{103}$$

Defining the Fock matrix as

$$F_{\nu\mu\sigma'} = \frac{\partial L^{\text{UHF}}}{\partial D_{\mu\nu\sigma'}} \tag{104}$$

and noting that

$$S_{\nu\mu} = -\frac{\partial L^{\text{UHF}}}{\partial W_{\mu\nu\sigma'}}, \tag{105}$$

the Roothaan-Hall equations result,

$$\sum_\mu F_{\upsilon\mu\sigma} C_{\mu i\sigma} = \sum_\mu S_{\upsilon\mu} (\mathbf{C}\boldsymbol{\epsilon})_{\mu i\sigma}. \tag{106}$$

Matrix notation:

$$\mathbf{FC} = \mathbf{SC}\boldsymbol{\epsilon} \tag{107}$$

2. Lagrange multipliers:

$$\frac{\partial L^{\text{UHF}}}{\partial \epsilon_{ij\sigma}} = \sum_{\mu\nu\sigma'} \frac{\partial L^{\text{UHF}}}{\partial W_{\mu\nu\sigma'}} \frac{\partial W_{\mu\nu\sigma'}}{\partial \epsilon_{ij\sigma}} - \delta_{ij} = \sum_{\mu\nu} C^*_{\nu j\sigma} S_{\nu\mu} C_{\mu i\sigma} - \delta_{ij} = 0 \tag{108}$$

Matrix notation:

$$\mathbf{C}^\dagger \mathbf{SC} = \mathbf{1}, \tag{109}$$

i.e., stationarity with respect to $\boldsymbol{\epsilon}$ recovers the orbital orthonormality constraint.

3.1.8 Fock Operator

- Definition of the Fock operator:

$$\mathbf{F}_\sigma = \frac{\partial L^{\text{UHF}}}{\partial \mathbf{D}_\sigma} \tag{110}$$

- Thus $\mathbf{F}_\sigma$ is self-adjoint, because $\mathbf{D}_\sigma$ is self-adjoint, and L^{UHF} is real.

- To use the chain rule, we note that

$$\frac{\partial \Gamma_{\lambda\kappa\sigma'\nu\mu\sigma}}{\partial D_{\upsilon\tau\sigma''}} = \frac{1}{2} \left(\delta_{\lambda\upsilon}\delta_{\kappa\tau}\delta_{\sigma'\sigma''} D_{\nu\mu\sigma} + D_{\lambda\kappa\sigma'}\delta_{\nu\upsilon}\delta_{\mu\tau}\delta_{\sigma\sigma''} \right.$$
$$\left. -\delta_{\sigma\sigma'}\delta_{\lambda\upsilon}\delta_{\mu\tau}\delta_{\sigma'\sigma''} D_{\nu\kappa\sigma'} - \delta_{\sigma\sigma'} D_{\lambda\mu\sigma}\delta_{\nu\upsilon}\delta_{\kappa\tau}\delta_{\sigma'\sigma''} \right). \tag{111}$$

- Using Eq. (99) and $(\mu\nu|\kappa\lambda) = (\kappa\lambda|\mu\nu)$, which follows from the symmetry of the Coulomb interaction under electron exchange, we arrive at

$$F_{\mu\nu\sigma} = h_{\mu\nu} + \sum_{\kappa\lambda\sigma'} \left[(\mu\nu|\kappa\lambda) - \delta_{\sigma\sigma'}(\mu\lambda|\kappa\nu) \right] D_{\lambda\kappa\sigma'}. \tag{112}$$

- Matrix notation:

$$\mathbf{F}_\sigma = \mathbf{h} + \mathbf{J} + \mathbf{K}_\sigma, \tag{113}$$

- The Coulomb operator

$$J_{\mu\nu} = \sum_{\kappa\lambda\sigma'} (\mu\nu|\kappa\lambda) D_{\lambda\kappa\sigma'} = \int d^3r_1 d^3r_2 \chi_\mu^*(\mathbf{r}_1) \frac{\rho(\mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|} \chi_\nu(\mathbf{r}_1) \tag{114}$$

corresponds to the matrix representation of the repulsive electrostatic or Hartree or "Coulomb" potential

$$V^{\mathrm{H}}(\mathbf{r}) = \int d^3r' \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \tag{115}$$

generated by the total electron density

$$\rho(\mathbf{r}) = \sum_{\mu\nu\sigma} D_{\mu\nu\sigma} \chi_\mu(\mathbf{r}) \chi_\nu^*(\mathbf{r}). \tag{116}$$

- The exchange operator

$$K_{\mu\nu\sigma} = -\sum_{\kappa\lambda} (\mu\lambda|\kappa\nu) D_{\lambda\kappa\sigma} = -\int d^3r_1 d^3r_2 \chi_\mu^*(\mathbf{r}_1) \frac{\gamma(\mathbf{r}_1, \mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|} \chi_\nu(\mathbf{r}_2) \tag{117}$$

corresponds to the matrix representation of an attractive nonlocal exchange potential

$$V_\sigma^{\mathrm{X}}(\mathbf{r}_1, \mathbf{r}_2) = -\frac{\gamma(\mathbf{r}_1, \mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|}, \tag{118}$$

where

$$\gamma(\mathbf{r}_1, \mathbf{r}_2) = \sum_{\mu\nu\sigma} D_{\mu\nu\sigma} \chi_\mu(\mathbf{r}_1) \chi_\nu^*(\mathbf{r}_2) \tag{119}$$

is the density matrix operator in real space.

3.1.9 The Self-Consistent Field (SCF) Method

- The Roothaan-Hall equations, Eq. (107), correspond to a a generalized eigenvalue problem with a nonlinear operator $\mathbf{F}_\sigma[\mathbf{D}_\alpha, \mathbf{D}_\beta]$.

- Iterative diagonalization (self-consistent field) method: See Algorithm 1.

- Damping (i.e., re-using part of the previous density matrix) and level shifting (i.e., artificially increasing the HOMO-LUMO gap during iterations) may be necessary to help convergence. Direct inversion on the iterative subspace (DIIS, Pulay) is essential.

- Integral direct SCF (Almlöf, Häser, Ahlrichs): Integrals $(\mu\nu|\kappa\lambda)$ are not stored, but computed "on the fly" in each iteration (if needed). Avoids large I/O bottleneck.

- Integral screening: By the Cauchy-Schwarz inequality,

$$|(\mu\nu|\kappa\lambda)| \leq Q_{\mu\nu} Q_{\kappa\lambda}, \tag{120}$$

where $Q_{\mu\nu} = \sqrt{(\mu\nu|\mu\nu)}$. $\mathbf{Q}$ may be pre-computed and stored to decide whether integrals need to be computed or not.

Algorithm 1: Solution of the Roothaan-Hall equations by iterative diagonalization

---

Construct and store $\mathbf{h}$ and $\mathbf{S}$
Generate initial guess for $\mathbf{C}_\sigma$ and occupation numbers
while *not converged* **do**
    Construct $\mathbf{D}_\sigma$ (DIIS extrapolation)
    Construct $\mathbf{J}[\mathbf{D}_\alpha, \mathbf{D}_\alpha]$ and $\mathbf{K}_\sigma[\mathbf{D}_\sigma]$
    Construct $\mathbf{F}_\sigma = \mathbf{h} + \mathbf{J} + \mathbf{K}_\sigma$
    Compute energy $\frac{1}{2}\mathrm{tr}(\sum_\sigma (\mathbf{F}_\sigma + \mathbf{h})\mathbf{D}_\sigma)$
    Obtain new $\mathbf{C}_\sigma$ by solving $\mathbf{F}_\sigma \mathbf{C}_\sigma = \mathbf{S}\mathbf{C}_\sigma \boldsymbol{\epsilon}_\sigma$ for given $\mathbf{F}_\sigma$
    Check convergence
end

3.1.10 Meaning of Canonical HF Orbitals

- The HF equations generally admit more solutions than the number of electrons/ occupied orbitals: Distinguish occupied and unoccupied/virtual orbitals.

- Energies of occupied orbitals must be negative, otherwise the ground state wave-function corresponds to an unbound state and cannot be normalization ("continuum state").

- In the canonical basis,

$$\boldsymbol{\epsilon}_\sigma = \mathbf{C}_\sigma^\dagger \mathbf{F}_\sigma \mathbf{C}_\sigma \tag{121}$$

is chosen to be diagonal for all orbitals.

- Koopmans Theorem: The energies of occupied HF orbitals approximate negative ion-ization potentials, and the energies of virtual orbitals approximate electron affinities. Canonical HF orbitals approximate Dyson orbitals.

- However, the HF energy functional is invariant under unitary rotations of occupied into occupied or virtual into virtual MOs, See Sec. 3.1.3.

- Thus, the Brillouin condition

$$\epsilon_{ai\sigma} = \mathbf{C}_{a\sigma}^\dagger \mathbf{F}_\sigma \mathbf{C}_{i\sigma} = 0 \tag{122}$$

is generally sufficient for making the HF energy stationary. Only canonical orbitals have well-defined orbital energies, however.

3.1.11 When to Use HF Theory

- HF works well if electron correlation effects due the Coulomb interaction may be neglected: Purely electrostatic interactions, ionic systems, interactions between polar species at extremely long distance, isodesmic/homodesmotic reactions.

- HF breaks down whenever electron correlation is strong:

– Dynamic correlation: Mostly short range. Key for bonding pairs of electrons, high angular momentum bonds, metallic systems at high density, transition states; van der Waals.

– Static correlation: Due to near degeneracy of excited Slater determinants with HF determinant. Important for homolytic bond dissociations, small-gap compounds (e.g. open-shell d and f element complexes), many radicals.

- For most purposes in chemistry, HF energy differences are not accurate enough, see Tab. 12.

Molecule	HF	Exp.
H_2	85	109
Be_2	<0	2.5
N_2	122	228
F_2	-32	38

Table 12: HF and experimental dissociation energies of some diatomic molecules in kcal/mol.

- Typical HF bond distances (e.g. C–C bonds) are too short by 4-5 pm, force constants are too high.

- HF is mainly used for reference/interpretation purposes, and as starting point for correlated wavefunction calculations.

3.2 Density Functional Theory

3.2.1 Key Assumptions

- Hohenberg-Kohn Theorem (later refined by Levy and Lieb): The ground-state energy of a many-electron system is a unique functional of the electron density $\rho_\sigma(\mathbf{r})$.

- Kohn-Sham (KS) Scheme: The exact ground-state density $\rho_\sigma(\mathbf{r})$ may be generated from a determinant (KS determinant) with occupied orbitals $|\phi_i\rangle$:

$$\rho_\sigma(\mathbf{r}) = \sum_i |\phi_{i\sigma}(\mathbf{r})|^2 \tag{123}$$

- Ground state energy functional:

$$E[\rho_\alpha, \rho_\beta] = T_s[\rho_\alpha, \rho_\beta] + V_{ne}[\rho_\alpha, \rho_\beta] + V_{nn}E^{\mathrm{H}}[\rho_\alpha, \rho_\beta] + E^{\mathrm{XC}}[\rho_\alpha, \rho_\beta] \tag{124}$$

– $T_s[\rho_\alpha, \rho_\beta] = \sum_i \langle \phi_i | \hat{t} | \phi_i \rangle$: KS kinetic energy

– $V_{ne}[\rho_\alpha, \rho_\beta] = -\sum_I Z_I \int d^3r \frac{\rho(\mathbf{r})}{|\mathbf{r}-\mathbf{R}_I|}$: Electron-nucleus attraction

– $E^{\mathrm{H}}[\rho_\alpha, \rho_\beta] = \int d^3r d^3r' \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}$: Hartree energy

– $E^{\mathrm{XC}}[\rho_\alpha, \rho_\beta]$: Exchange-correlation energy (unknown functional)

3.2.2 Generalized Kohn-Sham Scheme

- In the generalized Kohn-Sham scheme, the energy functional is minimized with respect to the KS density matrix $\mathbf{D}_\sigma$.

- For fixed particle number, this is equivalent to minimization with respect to MO coefficients $\mathbf{C}_\sigma$.

- Generalized KS Lagrangian:

$$L^{\mathrm{GKS}}[\mathbf{D}_\alpha, \mathbf{D}_\beta, \boldsymbol{\epsilon}_\alpha, \boldsymbol{\epsilon}_\beta] = \sum_\sigma \mathrm{tr}(\mathbf{h}\mathbf{D}_\sigma) + \sum_{\substack{\mu\nu\kappa\lambda \\ \sigma\sigma'}} (\mu\nu|\kappa\lambda) D_{\lambda\kappa\sigma'} D_{\nu\mu\sigma} + E^{\mathrm{XC}}[\mathbf{D}_\alpha, \mathbf{D}_\beta]$$
$$- \sum_\sigma \mathrm{tr}(\mathbf{S}\mathbf{W}_\sigma - \boldsymbol{\epsilon}_\sigma) \tag{125}$$

- Stationarity with respect to $\mathbf{C}_\sigma^\dagger$: GKS equations

$$\mathbf{F}_\sigma^{\mathrm{GKS}}\mathbf{C}_\sigma = \mathbf{S}\mathbf{C}_\sigma\boldsymbol{\epsilon}_\sigma. \tag{126}$$

- GKS Fock operator:

$$\mathbf{F}_\sigma^{\mathrm{GKS}} = \frac{\partial L^{\mathrm{GKS}}}{\partial \mathbf{D}_\sigma} = \mathbf{h} + \mathbf{J} + \mathbf{K}_\sigma^{\mathrm{XC}}, \tag{127}$$

  where

$$\mathbf{K}_\sigma^{\mathrm{XC}} = \frac{\partial E^{\mathrm{XC}}[\mathbf{D}_\alpha, \mathbf{D}_\beta]}{\partial \mathbf{D}_\sigma} \tag{128}$$

  is the exchange-correlation potential operator (matrix).

- Stationarity with respect to $\boldsymbol{\epsilon}$: Orthonormality constraint

$$\mathbf{C}_\sigma^\dagger \mathbf{S} \mathbf{C}_\sigma = \mathbf{1}. \tag{129}$$

3.2.3 Common Density Functional Approximations

- "Jacob's Ladder" of density functional approximations (DFAs):

  1. Local spin density approximation (LSDA, LSD):

$$E^{\mathrm{XC, \ LSDA}}[\rho_\alpha, \rho_\beta] = \int d^3 r \epsilon^{\mathrm{XC}}(\rho_\alpha(\mathbf{r}), \rho_\beta(\mathbf{r}))\rho(\mathbf{r}) \tag{130}$$

  ϵ^{XC}: XC energy per particle (energy density) of a uniform electron gas with constant electron density $n_\sigma = \rho_\sigma$. Example: Slater-Dirac exchange,

$$E^{\mathrm{X, \ LSDA}}[\rho_\alpha, \rho_\beta] = -\frac{1}{8\pi^3} \sum_\sigma (6\pi^2 \rho_\sigma(\mathbf{r}))^{4/3}. \tag{131}$$

  LSDA correlation functionals: Vosko, Wilk, Nusair (VWN), Perdew and Wang (PW, 1992). LSDA becomes exact in the thermodynamic limit $N \to \infty$.

2. Generalized gradient approximation (GGA):

$$E^{\text{XC, GGA}}[\rho_\alpha, \rho_\beta] = \int d^3 r f(\rho_\alpha(\mathbf{r}), \rho_\beta(\mathbf{r}), \nabla \rho_\alpha(\mathbf{r}), \nabla \rho_\beta(\mathbf{r})) \qquad (132)$$

Breakthrough for chemistry. Popular GGA functionals: Becke-Perdew 86 (BP86, Becke 1988 GGA for exchange plus Perdew 1986 GGA for correlation), Perdew-Burke-Ernzerhof (PBE), Lee-Yang-Parr (LYP), Becke 1997 GGA (B97).

3. Meta-generalized gradient approximation (MGGA):

$$E^{\text{XC, MGGA}}[\rho_\alpha, \rho_\beta] = \int d^3 r f(\rho_\alpha(\mathbf{r}), \rho_\beta(\mathbf{r}), \nabla \rho_\alpha(\mathbf{r}), \nabla \rho_\beta(\mathbf{r}), \tau_\alpha(\mathbf{r}), \tau_\beta(\mathbf{r})) \quad (133)$$

These functionals also depend locally on the KS kinetic energy density

$$\tau_\sigma(\mathbf{r}) = \frac{1}{2} \sum_i |\nabla \phi_i(x)|^2. \qquad (134)$$

More consistent performance than GGAs for a wide variety of systems/properties. Examples: Tao-Perdew-Staroverov-Scuseria (TPSS), strongly constrained approximatel normed (SCAN, Perdew and co-workers), many Minnesota functionals.

4. Hybrid functionals: Rung 2 or 3 functional plus fraction of HF exchange, e.g.,

$$E^{\text{XC, MGGAh}}[\mathbf{D}_\alpha, \mathbf{D}_\beta] = E^{\text{XC, MGGA}}[\rho_\alpha, \rho_\beta] + c^{\text{X}} E^{\text{X, HF}}[\mathbf{D}_\alpha, \mathbf{D}_\beta] \qquad (135)$$

c^{X} is called hybrid mixing parameter and usually around 0.2. Examples: TPSSh ($c^{\text{X}} = 0.1$), Becke three-parameter hybrid (B3LYP, $c^{\text{X}} = 0.2$), PBE0 ($^{\text{X}} = 0.25$). Apart from "global" hybrids there are also range-separated hybrids, e.g., ωB97M-V (Head-Gordon) and local hybrids.

5. RPA-like functionals depeding nonlocally on the density through the KS Hamiltonian.

3.2.4 When To Use DFT

- Hybrid DFAs such as B3LYP or TPSSh are all-around methods for electronic structure calculations. Typical accuracy:

  - Covalent bond energies: $\sim 3 - 5$ kcal/mol
  - Barrier heights: ~ 5 kcal/mol
  - Covalent bond distances: ≤ 1 pm
  - Harmonic vibrational frequencies: ~ 100 cm^{-1}

- GGAs and mGGAs may be used for fast screening, preoptimization

- Semilocal DFAs may fail for

  - noncovalent interactions. Dispersion corrections needed.

- quasi-one-electron systems, e.g. radicals, odd-electron bonds; negative ions: Self-interaction error

- strongly correlated systems, especially for magnetic/spin related properties

- charge-transfer and Rydberg excitations in time-dependent DFT

- Highly parameterized functionals may fail for systems/properties different from the ones used to fit them.

4 Gaussian Basis Sets

4.1 Uncontracted Gaussian Basis Sets

- Basis functions: Spherical GTOs,

$$\chi_s(\mathbf{r}|\mathbf{R},\mathbf{l},\zeta) = Y_l^m(\mathbf{r}-\mathbf{R})|\mathbf{r}-\mathbf{R}|^l e^{-\zeta(\mathbf{r}-\mathbf{R})^2}, \tag{136}$$

for computational convenience expressed by Cartesian GTOs (CGTOs),

$$\chi(\mathbf{r}|\mathbf{R},\mathbf{l},\zeta) = (x-X)^i(y-Y)^j(z-Z)^k e^{-\zeta(\mathbf{r}-\mathbf{R})^2}, \tag{137}$$

where $\mathbf{l} = (i,j,k)$ and $l = i+j+k$ is the (pseudo-) l-quantum number, $i,j,k = 0,1,2,\ldots$. Older basis sets (Pople) sometimes use CGTOs as primary basis functions.

- Shells: Set of basis functions with identical $\mathbf{R}$, ζ, l. There are $(l+1)(l+2)/2$ CGTOs in a shell, i.e. $6d$, $10f$ etc.

- Key properties of CGTOs:

  - Factorization:

  $$\chi(\mathbf{r}|\mathbf{R},\mathbf{l},\zeta) = (x-X)^i e^{-\zeta(x-X)^2}(y-Y)^j e^{-\zeta(y-Y)^2}(z-Z)^k e^{-\zeta(z-Z)^2} \tag{138}$$

  - Gaussian Product Theorem (GPT): The product of two s type CGTOs with centers $\mathbf{A}$, $\mathbf{B}$ is an s type CGTO centered at

  $$\mathbf{C} = \frac{\zeta_A \mathbf{A} + \zeta_B \mathbf{B}}{\zeta_A + \zeta_B} \tag{139}$$

  and damped by an overlap factor,

  $$e^{-\zeta_A(\mathbf{r}-\mathbf{A})^2} e^{\zeta_B(\mathbf{r}-\mathbf{B})^2} = e^{-\frac{\zeta_A\zeta_B}{\zeta_A+\zeta_B}(\mathbf{A}-\mathbf{B})^2} e^{-(\zeta_A+\zeta_B)(\mathbf{r}-\mathbf{C})^2}. \tag{140}$$

  - Vertical recursion: Generate higher l quantum CGTOs on center $\mathbf{R}$, e.g.,

  $$\frac{\partial}{\partial X}\chi(\mathbf{r}|\mathbf{R},\mathbf{l},\zeta) = -i\chi(\mathbf{r}|\mathbf{R},\mathbf{l}-\mathbf{e_x},\zeta) + 2\zeta\chi(\mathbf{r}|\mathbf{R},\mathbf{l}+\mathbf{e_x},\zeta), \tag{141}$$

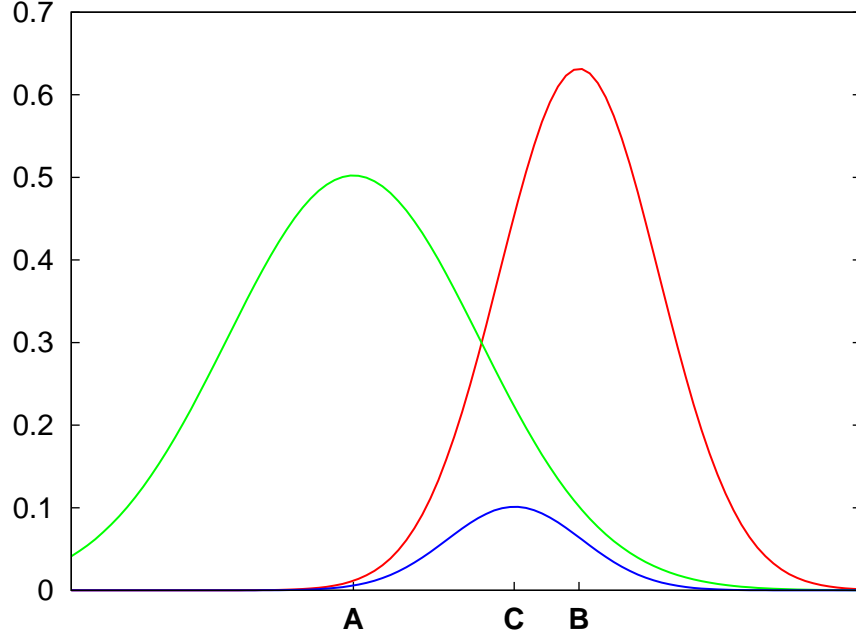 where $e_x = (1,0,0)$ and $l > 0$. For $l = 0$, the first term vanishes.

Figure 17: Illustration of the GPT.

- Horizontal recursion: Shift l quantum numbers between centers, e. g.

$$
\begin{aligned}
\chi(\mathbf{r}|\mathbf{A}, \mathbf{l_A}, \zeta_A)&\chi(\mathbf{r}|\mathbf{B}, \mathbf{l_B}, \zeta_B) \\
&= (X_B - X_A)\chi(\mathbf{r}|\mathbf{A}, \mathbf{l_A} - \mathbf{e_x}, \zeta_A)\chi(\mathbf{r}|\mathbf{B}, \mathbf{l_B}, \zeta_B) \\
&+ \chi(\mathbf{r}|\mathbf{A}, \mathbf{l_A} - \mathbf{e_x}, \zeta_A)\chi(\mathbf{r}|\mathbf{B}, \mathbf{l_B} + \mathbf{e_x}, \zeta_B)
\end{aligned}
\tag{142}
$$

- Choice of exponents:

  - Full variational optimization (best)
  - Even tempered (Bardo and Ruedenberg):

$$
\zeta_\nu = \alpha\beta^{\nu-1},
\tag{143}
$$

  where $\beta > 0, \beta \neq 1$. ζ_ν are geometric sequence with constant ratio $\zeta_{\nu+1}/\zeta_\nu = \beta$. Optmimze α and β only.

  - Sufficient condition for completeness of GTO basis sets (Müntz/ Klahn):

$$
\sum_{\nu=1}^{\infty} \frac{\zeta_\nu}{1 + \frac{1}{4}\zeta_\nu^2} = \infty
\tag{144}
$$

- Convergence of total energies: see Tab. 13

Basis	HF Energy
$4s$	−2.855 160
$6s$	−2.861 153
$8s$	−2.861 625
$10s$	−2.861 673
limit	−2.861 680

Table 13: HF ground state energies of He for fully optimized uncontracted GTO basis sets.

4.2 Contracted Gaussian Basis Sets

- Contraction: Reduce number of functions by linear combination with *fixed* coefficients to form contracted CGTOs,

$$\psi(\mathbf{r}|\mathbf{R}, \mathbf{l}) = \sum_{n=1}^{p} c_n \chi(\mathbf{r}|\mathbf{R}, \mathbf{l}, \zeta_n) \tag{145}$$

- Segmented contraction: Groups of exponents describing several AOs at the same time. See Table 14

p	l	ζ	c
5	s	1238	0.00546
		186.3	0.0406
		42.3	0.180
		11.7	0.463
		3.59	0.441
1	s	0.402	1.0
1	s	0.131	1.0
3	p	9.468	0.0384
		2.01	0.211
		0.547	0.513
1	p	0.153	1.0

Table 14: Split-valence (SV) segmented contracted basis set for C. This is a {511/31} or $(7s4p)/[3s2p]$ contraction.

- Generalized contraction (Dunning): Use occupied atomic orbital coefficients from uncontracted atomic HF calculation as contraction coefficients, add primitives with smaller exponents.

- n-tuple zeta (Z) basis sets: Contains n CGTO shells for each occupied n, l-subshell. $n = 2$(D), 3(T), 4(Q), 5, 6, …. $n = 1$: Single zeta (SZ) or minimal basis sets.

- n-tuple valence (V) basis sets: SZ for core shells, nZ for valence shells. Split valence (SV) $\equiv$ DZV.

- Pople basis sets: Valence shell s and p GTOs have identical exponents. 6-31G is SV with {631/31} contraction, 6-311G is TZV with {6311/311} contraction
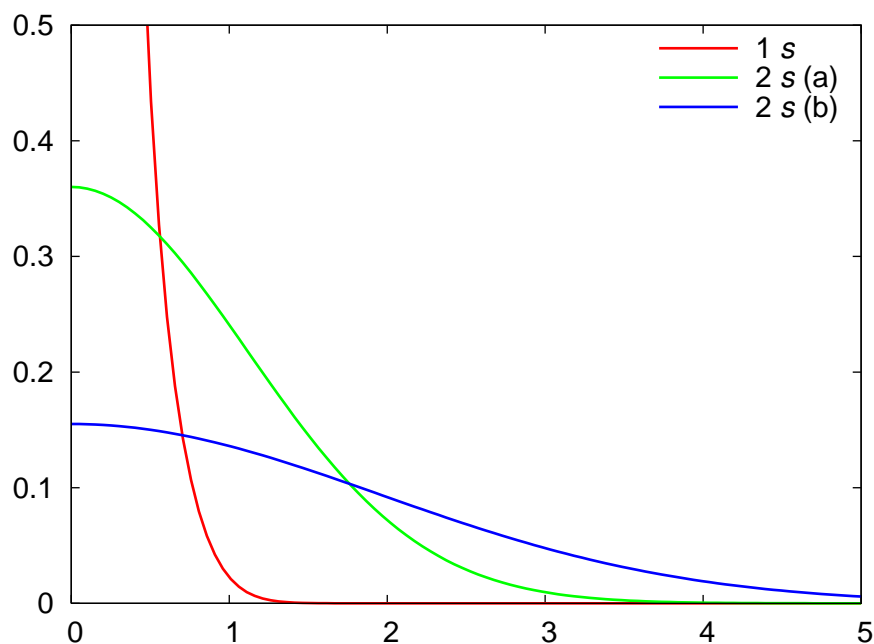
Figure 18: *s* CGTOs of the C SV basis set.

4.3   Polarization and Augmentation

- Polarization functions:

  - Correlation consistent polarized valence basis sets (cc-pVXZ, Dunning): $1d$ for $X = 2$, $2d1f$ for $X = 3$, $3d2f1g$ for $X = 4$

  - Hydrogen polarization functions often less important: SV(P) or 6-31G* have no polarization functions on H.

- Augmentation:

  - Diffuse (small exponent) functions: Anions, (Rydberg) excitations

  - Core polarization functions: Steep (large exponent) polarization functions

- Common GTO basis sets: See Tab. 15.

  - Workhorse: SV(P) (mostly for structures)

  - TZVP to check SV(P), (single-point) energy calculations

  - QZVP for HF and DFT basis set limit results, benchmarks

- Most GTO basis sets available on EMSL basis set library website [6].

Name	n	Contraction (for C)	ΔE_b	Δr_e	Comments
SV(P), SVP	2	$(7s4p1d)/[3s2p1d]$	5	1-2	cf. 6-31G*, 6-31G**, cc-pVDZ
TZVP	3	$(11s6p2d1f)/[5s3p2d1f]$	2	0.5	cf. 6-311G(2df,2pd), cc-pVTZ
QZVP	4	$(15s8p3d2f1g)$ $/[7s4p3d2f1g]$	< 1	< 0.3	cf. cc-pVQZ

Table 15: Ahlrichs basis sets. The contraction pattern for C is given as an example. Typical basis set errors in bond energies ΔE_b are in kcal/mol per bond, typical basis set errors in bond distances r_e are in pm.

4.4 Basis Set Error

- Basis set incompleteness error (BSIE):

  - HF and semilocal DFA energy converges approximately exponentially with the size of the one-electron basis. BSIE is typically less than 1 mH using def2-QZVP.
  - Correlated wavefunction methods:

$$E_c(X) = E_{c\infty} - \frac{A}{X^3} \tag{146}$$

  for correlation consistent basis sets.

- Basis set superposition error (BSSE):

  - Neigboring atoms basis sets correct atomic basis set deficiencies
  - Spurious stabilization of molecules / compact structures relative to fragments / open structures
  - Most important for weak bonds (BE < 10 kcal/mol)
  - Boys-Bernardi counterpoise: Estimate BSSE from fragment calculation with "ghost" atoms, i.e. only basis functions (no electrons, nuclear charges) on the other fragment are present

References

[1] Hanson, N. R., *Patterns of Discovery*, Cambridge University Press, 1972.

[2] Muuronen, M., Deglmann, P., Tomović, Z., *J. Org. Chem.* **84** (2019), 8202.

[3] Nguyen, T., *C&E News* (2019).

[4] Bach, A., Hutchinson, R. J., Holiday, R. J., Crim, F. F., *J. Phys. Chem. A* **107** (2003), 10490.

[5] Fliegl, H., Sundholm, D., Taubert, S., Jusélius, J., Klopper, W., *J. Phys. Chem. A* **113** (2009), 8668.

[6] http://www.emsl.pnl.gov/forms/basisform.html.