

# Atomic partitioning schemes for wavefunction analysis and effective atomic orbitals

Pedro Salvador

Institut de Química Computacional i Catalis i Departament de Química,  
Universitat de Girona, Facultat de Ciències,  
c/M Aurelia Capmany 69, 17003, Girona  
email:pedro.salvador@udg.edu

CTTC2019 School, Quito, 29-30 June 2019

# What is a molecule?

## A system of electrons and nuclei (Physicist picture)

Physical Information:

- Hamiltonian, Wavefunction (density) Hermitian operators (functionals) associated to observables and their expectation values

## A system of bonded (interacting) atoms (Chemist picture)

Chemical Information:

- Bonding interaction, steric repulsion, polarity, functional group, aromaticity...

We aim to interpret the results of ab initio calculations in classical chemical terms

# The Atom in the Molecule

## Atoms are the building blocks of chemistry

- Not observables in Quantum Mechanics.
- Is there a unique atom in molecule (AIM) definition?
- Any AIM is a conceptual construct but with an irrefutable utility.
- Taking advantage of AIM, quantum chemist have defined bond orders indexes, atomic populations, bond orders descriptors, energy partitioning, aromaticity indexes, among others.

## Assigning physical quantities to the individual atoms

- Any physical quantity is written as expectation value of one- or two-electron operators
- Which is the physical quantity that can be related to a given chemical concept?

# The Atom in the Molecule

- The nucleus is always considered as part of the atom, so the differences always arise in how the physical space (or the Hilbert-space) is subdivided into atomic shares.
- In the MO- LCAO approach (Molecular Orbitals as Linear Combination of Atomic Orbitals), the atom may be identified with the subspace of the basis functions attached to it. Such approaches are the so-called Hilbert-space analyses.
- A different strategy is to subdivide the physical 3D space into atomic regions or domains, which represent the atom. They may be defined disjoint, like in Bader's "quantum theory of atoms in molecules" (QTAIM), or may be allowed to overlap, like in the different variants of the "fuzzy" atoms
- In Hilbert-space analysis one deals with the AO matrix representation of operators. In 3D-space analysis one usually manipulates density functions

# Hilbert-space analysis

In the LCAO framework, the MOs are expanded on a finite set of atomic-centered one-electron functions  $\{\chi_\mu(\mathbf{r})\}$  as

$$\varphi_i(\mathbf{r}) = \sum_{\mu} c_{\mu i} \chi_{\mu}(\mathbf{r})$$

The set of functions centered on a given atom conform an atomic Hilbert subspace  $\{\chi_\mu(\mathbf{r})\}_{\mu \in A}$

**In this context, the atom may be defined by its nucleus and the subspace of one-electron basis set centered on it.**

A molecular orbital can be rewritten as a sum of its atomic counterparts as

$$\varphi_i(\mathbf{r}) = \sum_A \sum_{\mu \in A} c_{\mu i} \chi_{\mu}(\mathbf{r})$$

# Hilbert-space analysis

The first-order electron density can be written as

$$\rho(\mathbf{r}) = \sum_{\mu\nu} D_{\mu\nu} \chi_{\nu}^*(\mathbf{r}) \chi_{\mu}(\mathbf{r}),$$

where  $\mathbf{D}$  is the density matrix in the atomic orbital (AO) representation. By integrating the density one obtains the total number of electrons

$$N = \int \rho(\mathbf{r}) d\mathbf{r} = \sum_{\mu\nu} D_{\mu\nu} \int \chi_{\nu}^*(\mathbf{r}) \chi_{\mu}(\mathbf{r}) d\mathbf{r} = \sum_{\mu\nu} D_{\mu\nu} S_{\nu\mu} = \sum_{\mu} (\mathbf{D}\mathbf{S})_{\mu\mu},$$

where matrix  $\mathbf{S}$  is the atomic **overlap** matrix in AO basis, with elements

$$S_{\nu\mu} = \int \chi_{\nu}^*(\mathbf{r}) \chi_{\mu}(\mathbf{r}) d\mathbf{r}$$

The basis functions that play the role of AOs in electronic structure calculations, usually Gaussian-type orbitals (GTO) or Slater-type orbitals (STO) do exhibit overlap.

# Hilbert-space analysis

Now, by systematizing the summation over the AOs according to which atom  $A$  they are assigned to, one can readily obtain the so-called **Mulliken's atomic gross populations**,  $N_A$

$$N = \sum_A \sum_{\mu \in A} (\mathbf{DS})_{\mu\mu} = \sum_A N_A.$$

Alternatively, one can also start from the previous expression in terms of two AO indices and systemtize them over each atom  $A$  and  $B$ , and write

$$N = \sum_A \sum_B \sum_{\mu \in A} \sum_{\nu \in B} D_{\mu\nu} S_{\nu\mu} = \sum_A N_{AA} + \sum_{A \neq B} N_{AB}$$

where  $N_{AA}$  and  $N_{AB} + N_{BA}$  are **Mulliken's net and overlap populations**, respectively.

# Hilbert-space analysis

## Pros

- The integrations are analytical for Gaussian functions.
- Straightforward implementation and low computational cost
- The errors associated with any additional numerical integration are avoided.

## Cons

- **Basis set dependency.**
- It may show unphysical results for basis sets without marked atomic character e.g. diffuse basis functions.
- Ill-defined in the complete basis set limit (CBS).
- Inapplicable beyond the LCAO-MO approximation e.g. if plane waves are used to expand the MOs of the system.



# Hilbert-space analysis

## Beyond Mulliken Analyses

- Most of the problems associated with Mulliken analyses originate from the overlap of the underlying atomic basis.
- When basis set without marked atomic character are used:  
Overlap populations or bond orders can be negative  
Effective atomic orbitals with occupations not in the  $[0,1]$  range

## Use of orthogonal basis

The AO basis can be transformed to an orthogonal basis, and the MO expanded into this orthonormalized basis set.

- Symmetric orthogonalization of Lowdin (and a number of variants, see e.g. Comput. Theor. Chem. 1008 (2013) 15)
- Natural orbitals and Natural Population analysis of Weinhold
- Quasi-atomic orbitals of Ruedenberg, etc..

# Hilbert-space analysis in orthogonal basis

For instance, in the Lowdin basis one has

$$\chi_i^L(\mathbf{r}) = \sum_{\mu} S_{\mu i}^{-1/2} \chi_{\mu}(\mathbf{r}) \quad \chi_{\mu}(\mathbf{r}) = \sum_i S_{i\mu}^{1/2} \chi_i^L(\mathbf{r})$$

Replacing the original AOs by the Lowdin orthonormalized ones

$$\begin{aligned} \rho(\mathbf{r}) &= \sum_{\mu\nu} D_{\mu\nu} \chi_{\nu}^*(\mathbf{r}) \chi_{\mu}(\mathbf{r}) = \sum_{\mu\nu} D_{\mu\nu} \left[ \sum_i S_{i\nu}^{1/2*} \chi_i^{L*}(\mathbf{r}) \right] \left[ \sum_j S_{j\mu}^{1/2} \chi_j^L(\mathbf{r}) \right] = \\ &= \sum_{\mu\nu} \sum_{ij} S_{j\mu}^{1/2} D_{\mu\nu} S_{\nu i}^{1/2\dagger} \chi_i^{L*}(\mathbf{r}) \chi_j^L(\mathbf{r}) = \sum_{ij} D_{ji}^L \chi_i^{L*}(\mathbf{r}) \chi_j^L(\mathbf{r}) \end{aligned}$$

where  $D_{ji}^L = \left[ \mathbf{S}^{1/2} \mathbf{D} \mathbf{S}^{1/2\dagger} \right]_{ji}$ .

In this new basis, the integration of the density leads to

$$N = \int \rho(\mathbf{r}) d\mathbf{r} = \sum_{ij} D_{ji}^L \delta_{ij} = \sum_i D_{ii}^L$$

# Hilbert-space analysis in orthogonal basis

Lowdin's atomic populations are obtained by systematizing the indices of the orthogonal AO basis, whose elements are also associated to the atomic centers, i.e.  $\{\chi_i^L(\mathbf{r})\}_{i \in A}$ .

$$N = \sum_A \sum_{i \in A} (D^L)_{ii} = \sum_A N_A^L$$

Notice that in an orthogonal basis there is no overlap population.

Assigning orthogonalized AOs to atomic centers might be ambiguous in the case of extended basis sets.

# Real-space analysis

## Disjoint domains

- Voronoi cells
- Quantum Theory of Atoms in Molecules (QTAIM)
- Topological analysis of vector fields other than  $\rho(\mathbf{r})$

## Overlapping domains (“fuzzy” atoms)

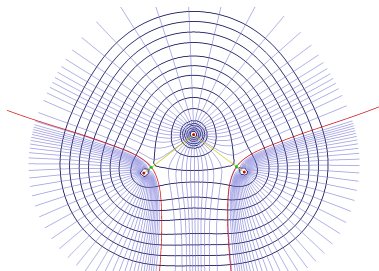
- Hirshfeld, Hirshfeld-Iterative, Hirshfeld-X,...
- Becke's constructs, Becke- $\rho$ , TFVC
- Iterative stockholder Atoms (ISA), ISA-X,...

The aim is to decompose the integral of a density function  $f(\mathbf{r})$  by performing (numerical) individual integrations over each atomic domain  $\Omega_A$

$$\int f(\mathbf{r}) d\mathbf{r} = \sum_A \int_{\Omega_A} f(\mathbf{r}) d\mathbf{r}$$

- Topological analysis of  $\rho(\mathbf{r})$
- Atomic basins from the zero-flux condition

$$\nabla\rho(\mathbf{r}) \cdot \vec{n}(\mathbf{r}) = \vec{0} \quad \forall \mathbf{r} \in S(\mathbf{r})$$



$$\int f(\mathbf{r}) d\mathbf{r} = \sum_A \int_{\Omega_A} f(\mathbf{r}) d\mathbf{r}$$

$$\iint f(\mathbf{r}, \mathbf{r}') d\mathbf{r} d\mathbf{r}' = \sum_{AB} \int_{\Omega_A} \int_{\Omega_B} f(\mathbf{r}, \mathbf{r}') d\mathbf{r} d\mathbf{r}'$$

## Pros

- Strong physical background (see notes from A.M. Pendás)
- Only information from  $\rho(\mathbf{r})$  is required.

## Cons

- Complex shape of the atomic domains leads to cumbersome numerical integrations (but more and more efficient codes are being developed, e.g. AIMAll)
- Spurious non-nuclear attractors (e.g.  $\text{C}_2\text{H}_2$  with a number of conventional mid-size basis sets) difficult the analysis in chemical terms.

# Overlapping domains

One can define a weight function for each atom at each point of the space satisfying

$$w_A(\mathbf{r}) > 0 \quad \text{and} \quad \sum_A w_A(\mathbf{r}) = 1$$

- Atoms can share the 3D-space (overlap) in general.
- 3D-space partitions make the results virtually basis set independent.
- In QTAIM  $w_A(\mathbf{r}) = 1$  if  $\mathbf{r} \in \Omega_A$  and  $w_A(\mathbf{r}) = 0$  otherwise.

$$\int f(\mathbf{r}) d\mathbf{r} = \sum_A \int w_A(\mathbf{r}) f(\mathbf{r}) d\mathbf{r}$$

$$\iint f(\mathbf{r}, \mathbf{r}') d\mathbf{r} d\mathbf{r}' = \sum_{AB} \iint w_A(\mathbf{r}) w_B(\mathbf{r}') f(\mathbf{r}, \mathbf{r}') d\mathbf{r} d\mathbf{r}'$$

## Different “fuzzy” atom approaches

Hirshfeld approach, TCA 44 129 (1977).

- Makes use of reference (promolecular) spherically-averaged atomic densities,  $\rho_A^0(\mathbf{r})$
- Superposition of promolecular atomic densities at the actual atomic positions of the molecule defines the promolecular total density ,  
 $\rho^0(\mathbf{r}) = \sum_B \rho_B^0(\mathbf{r})$

Then, in the classical Hirshfeld method

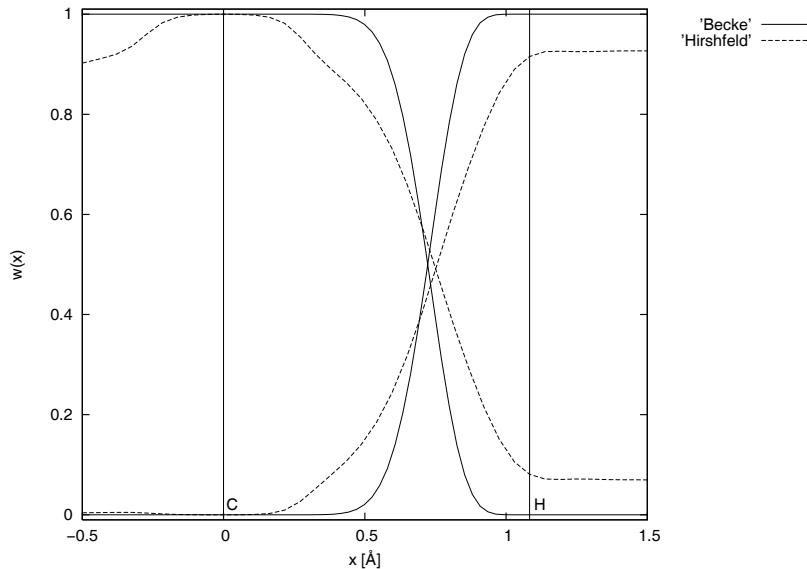
$$w_A(\mathbf{r}) = \frac{\rho_A^0(\mathbf{r})}{\sum_B \rho_B^0(\mathbf{r})}$$

Hirshfeld-Iterative schemes usually impose  $\int \rho_A^0(\mathbf{r}) d\mathbf{r} = N_A \neq Z_A$ , e.g. JCP 126 144111 (2007).

$$\rho_A^{0,N_A}(\mathbf{r}) = \alpha \rho_A^{0,int(N_A)+1}(\mathbf{r}) + (1 - \alpha) \rho_A^{0,int(N_A)}(\mathbf{r}), \quad \alpha = N_A - int(N_A)$$

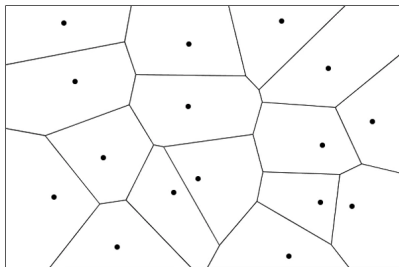


# A not-so-well-known problem of Hirshfeld's $w_A(\mathbf{r})$



## Becke's constructs, JCP 88 2457 (1988)

- Originally devised for effective numerical integration of three-dimensional functions of marked atomic character.
- Makes use of the so-called Voronoi polyhedra



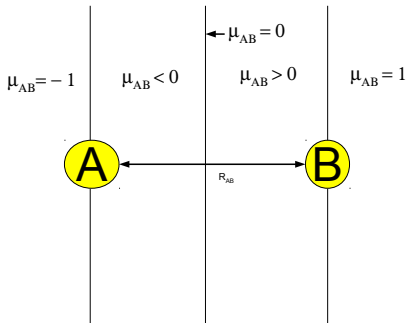
- The sharp boundaries of the Voronoi cells are substituted by soft functions that ultimately define the atomic weights
- By construction,  $w_A^{becke}(\mathbf{r}_A) = 1$  and  $\partial w_A^{becke}(\mathbf{r})/\partial(r)|_{r=R_A} = 0$ .

# Becke's constructs

The scheme can be formulated as follows for the simplest case of a two nuclei system (A and B). For any point of the space one can define the following quantity

$$\mu_{AB} = \frac{r_B - r_A}{R_{AB}}$$

where  $R_{AB}$  is the internuclear distance and  $r_A$  and  $r_B$  represent the distance of that point to nucleus A and B.



## Becke's constructs

The simple step function

$$s_A(\mu_{AB}) = \begin{cases} 1 & -1 \leq \mu_{AB} \leq 0 \\ 0 & 0 < \mu_{AB} \leq 1 \end{cases}$$

can be used to define the sharp Voronoi cell of atom  $A$  in this case.

The step function can be replaced by a continuous, monotonically decreasing function in the range  $[-1,1]$ , and fulfilling the requirements  $s(-1)=1$  and  $s(+1)=0$  in order to define **fuzzy Voronoi cells**

For that purpose Becke suggested the simple polynomial function

$$s_A^k(\mu_{AB}) = \frac{1}{2}[1 - f_k(\mu_{AB})]$$

where

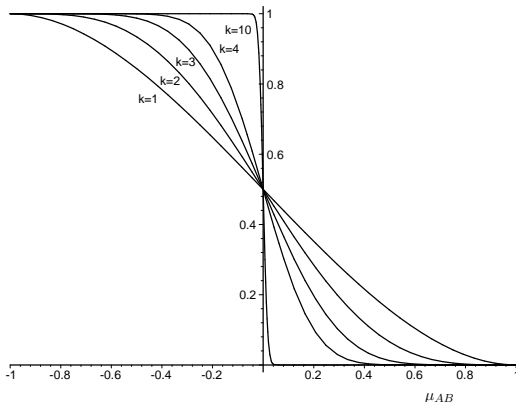
$$f_1(\mu) = \frac{3}{2}\mu - \frac{1}{2}\mu^3$$

# Becke's constructs

And devised an iterative process to obtain shaper cutoff profiles

$$f_k(\mu) = f[f_{k-1}(\mu)]$$

The integer  $k$  is known as the **stiffness parameter**; the larger the steeper the cutoff profile is.



## Becke constructs

In order to account for the different atomic sizes in heteronuclear systems, one can use the same cutoff function but introduce a shifted coordinate. Becke originally uses

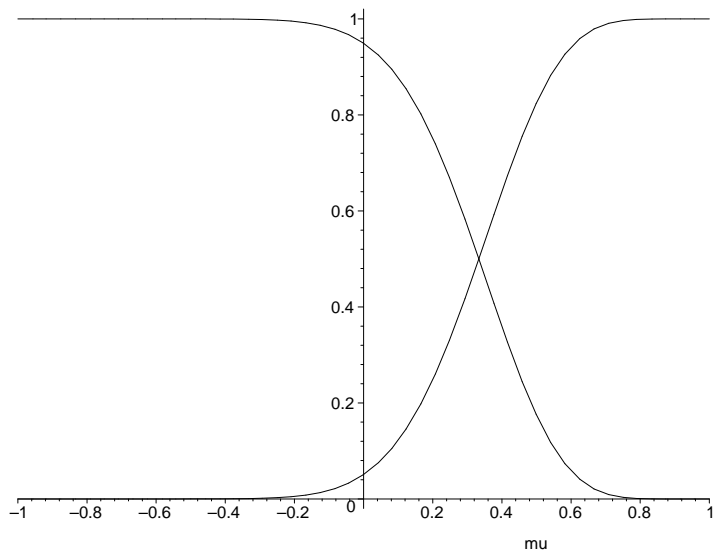
$$\nu_{AB} = \mu_{AB} + a_{AB}(1 - \mu_{AB}^2) \text{ where } a_{AB} = \frac{1 - \chi_{AB}^2}{4\chi_{AB}}$$

$$-1/2 \leq a_{AB} \leq 1/2 \text{ to ensure that } 0 \leq w_A(\mathbf{r}) \leq 1 \text{ and } \frac{R_A^0}{R_B^0} = \chi_{AB}$$

- For overlapping atoms the position of the interatomic boundaries between all pairs of atoms must be specified.
- A distance criterion is used, namely if  $R_{AB} > 2(R_A^0 + R_B^0)$  then  $\chi_{AB} = 1$ . Otherwise, the atoms are **neighbors** and  $\chi_{AB} = \frac{R_A^0}{R_B^0}$

# Becke constructs

Cutoff profile for  $k = 3$  and  $R_A = 2R_B$



# Becke constructs

- The set of atomic radii  $R_A^0$  determines the relative size of the atomic Voronoi cells
- The parameter  $k$  controls the shape of the cutoff profile.
- Becke used Bragg-Slater radii and  $k=3$  on the basis of more accurate integrations.

## Pros

- Numerical integrations in the fuzzy domains are very efficient
- Does not required the use of promolecular atomic densities

## Cons

- The use of a fixed atomic radii is a limitation of this AIM model because the same atoms are treated on equal footing in different chemical environments
- Highly arbitrary definition of the atom in the molecule



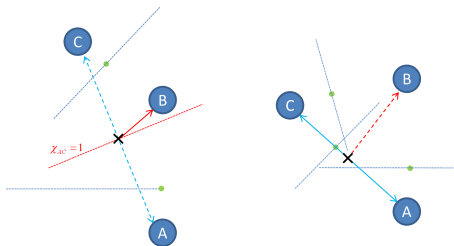
- Instead of using a set of fixed radii to define  $\chi_{AB}$  one can use position of the minimum of the density along the internuclear axis connecting two neighbor atoms.
- In Becke's original formula the maximum shift of the interatomic boundary of neighbor atoms occurs when the atomic radii differ by a factor of ca. 2.4
- The set of empirical atomic radii is still invoked for not neighbors
- Distance criterion appears inadequate for intermolecular interactions

## Topological fuzzy Voronoi cells (TFVC)

- The following alternative transformation is monotonic for any value of  $\chi_{AB}$ .

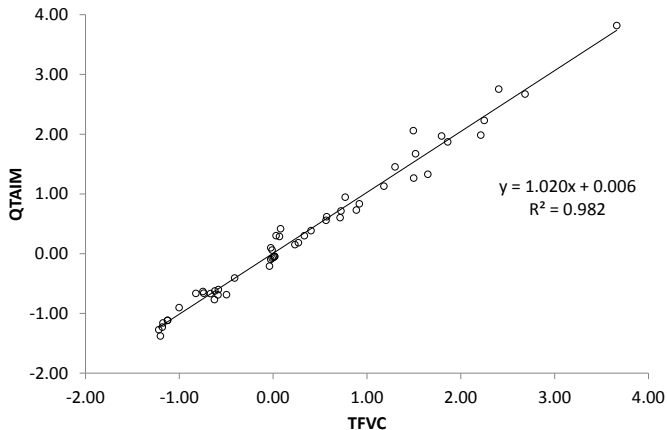
$$\nu'_{AB} = \frac{1 + \mu_{AB} - \chi_{AB}(1 - \mu_{AB})}{1 + \mu_{AB} + \chi_{AB}(1 - \mu_{AB})}$$

- Two atoms are not considered neighbors if their midpoint is closer to a third atom. Empirical set of radii no longer needed.



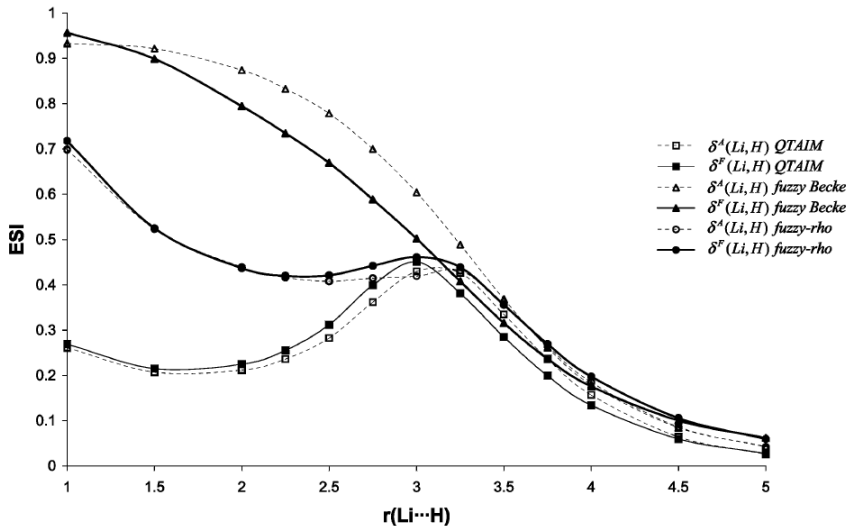
# TFVC vs QTAIM

Partial atomic charges for a set of hydrides (LiH to HCl)



# Conditions for proper AIM definitions

- An AIM should be able to account for complex bonding mechanisms (e.g. harpoon effect on LiH)



# Conditions for proper AIM definitions

- In aromatic systems, the delocalization index (bond order) should be larger in para- position than in meta-.

Table 7. HF/6-311++G\*\* ESI Values (Meta-ESI and PDI) for Benzene calculated Using Different Atomic Partitions (see text).

	meta-ESI	PDI
QTAIM	0.074	0.098
Becke $k = 3$	0.108	0.101
Becke $k = 4$	0.093	0.101
Becke $k = 5$	0.084	0.102
Becke-rho $k = 3$	0.093	0.092
Becke-rho $k = 4$	0.085	0.094
Becke-rho $k = 5$	0.079	0.095
ISA	0.191	0.122
Hirshfeld	0.192	0.114
Hirshfeld-I	0.203	0.120
Hirshfeld-ID	0.076	0.103

Identifying more conditions is desirable...

Effective atomic orbitals:  
the atomic orbitals of the topological  
atom

# The effective atomic orbitals

- One very useful tool to characterize the state of the atom within the molecule are the so-called **effective atomic orbitals** (eff-AOs), Mayer JCP 6249 (1996).
- In this approach one obtains **for each atom** a set of orthogonal atomic hybrids and their respective occupation numbers, adding up to the **net** population of the atom.
- These atomic hybrids closely mimic the core and valence shells of the atom, as anticipated on the basis of classical notions of electron configuration of the atom/fragment within the molecule.

$$\rho_{AA}(\mathbf{r}) = \sum_{\mu} \lambda_{\mu} (w_A(\mathbf{r}) \chi_{\mu}^A(\mathbf{r}))^* (w_A(\mathbf{r}) \chi_{\mu}^A(\mathbf{r})) \text{ and } \langle \chi_i^A | \chi_j^A \rangle = \delta_{ij}$$

# Theory

Let us consider a single-determinant WF formed by  $n$  orthonormalized doubly occupied orbitals  $\varphi_i(\mathbf{r})$ ,  $i = 1, 2, \dots, n$ , and a “fuzzy” division of the 3D-space into  $N_{at}$  atomic domains  $\Omega_A$  defined e.g., by a continuous atomic weight function  $w_A(\mathbf{r})$ , fulfilling  $w_A(\mathbf{r}) > 0$  and  $\sum_A w_A(\mathbf{r}) = 1$ .

Let us for each atom  $A$  ( $A = 1, 2, \dots, N_{at}$ ) form the  $n \times n$  Hermitian matrix  $\mathbf{Q}^A$  with the elements

$$Q_{ij}^A = \int w_A(\mathbf{r}) \varphi_i^*(\mathbf{r}) w_A(\mathbf{r}) \varphi_j(\mathbf{r}) d\mathbf{r}.$$

The matrix  $\mathbf{Q}^A$  is essentially the “net atomic overlap matrix” in the basis of the MO-s  $\varphi_i$ . Furthermore, for each atom  $A$  we define the “intraatomic” part  $\varphi_i^A$  of every MO  $\varphi_i$  as  $w_A(\mathbf{r}) \varphi_i(\mathbf{r})$ . Thus  $Q_{ij}^A = \langle \varphi_i^A | \varphi_j^A \rangle$  i.e.,  $\mathbf{Q}^A$  is the overlap matrix of the orbitals  $\varphi_i^A$ .



# Theory

We diagonalize the Hermitian matrix  $\mathbf{Q}^A$  by the unitary matrix  $\mathbf{U}^A$ :

$$\mathbf{U}^{A\dagger} \mathbf{Q}^A \mathbf{U}^A = \mathbf{\Lambda}^A = \text{diag}\{\lambda_i^A\} \quad .$$

It can be shown that every  $\lambda_i^A \geq 0$ , as is the case for an overlap matrix. The dimension of the matrix also indicates the max. number of non-zero eigenvalues that can be obtained ( $n$ ), disregarding the size of the underlying set of AOs

For each atom  $A$  we can define  $n_A$  ( $n_A \leq n$ ) “effective atomic orbitals”  $\chi_\mu^A(\mathbf{r})$  as linear combinations of the “intraatomic” parts  $\varphi_i^A(\mathbf{r})$  of the MO-s as

$$\chi_\mu^A(\mathbf{r}) = \frac{1}{\sqrt{\lambda_\mu^A}} \sum_{i=1}^n U_{i\mu}^A \varphi_i^A(\mathbf{r}) ; \quad \mu = 1, 2, \dots, n_A \quad ,$$

where  $n_A$  is the number of non-zero eigenvalues  $\lambda_i^A$ .

# Properties of the eff-AOs

- They are orthonormalized within the atomic domain:

$$\begin{aligned}\langle \chi_\mu^A | \chi_\nu^A \rangle &= \left\langle \frac{1}{\sqrt{\lambda_\mu^A}} \sum_{i=1}^n U_{i\mu}^A \varphi_i^A \middle| \frac{1}{\sqrt{\lambda_\nu^A}} \sum_{j=1}^n U_{j\nu}^A \varphi_j^A \right\rangle = \frac{1}{\sqrt{\lambda_\mu^A \lambda_\nu^A}} \sum_{i,j=1}^n (\mathbf{U}^{A\dagger})_{\mu i} Q_{ij}^A U_{j\nu}^A \\ &= \frac{1}{\sqrt{\lambda_\mu^A \lambda_\nu^A}} \lambda_\mu^A \delta_{\mu\nu} = \delta_{\mu\nu}\end{aligned}$$

- The sum of occ. numbers equals the *net* atomic population of the atom
- In the case of disjoint domains (QTAIM), orbitals  $\chi_\mu^A(\mathbf{r})$  differ from zero only in the atomic domain of atom  $A$ , thus

$$\langle \chi_\mu^A | \chi_\nu^B \rangle = 0$$

- In the case of disjoint domains,  $\mathbf{Q}^A$  is also the atomic overlap matrix of the MOs, as  $w_a(\mathbf{r})^2 \equiv w_a(\mathbf{r})$ .

# Properties of the eff-AOs

- In the framework of QTAIM they form an orthogonal set of numerical basis functions, and the occ. numbers add up to the atomic populations
- The eff-AOs of atom  $A$  can also be obtained **even if no atom-centered basis functions were used in the original MO description**, see CPL 563 97 (2013).
- They can also be obtained in the AO basis from the diagonalization of the matrix  $\mathbf{P}\mathbf{S}^A$ , where  $\mathbf{P}$  is the LCAO density matrix. This alternative permits the straightforward generalization to correlated WFs.
- Alternatively, in the natural orbital representation, the symmetrized form

$$\mathbf{U}'^{A\dagger} \boldsymbol{\eta}^{1/2} \mathbf{Q}^A \boldsymbol{\eta}^{1/2} \mathbf{U}'^A = \boldsymbol{\Lambda}^A = \text{diag}\{\lambda_i^A\}$$

that also produces the same set of eigenvalues, where  $\eta$  is the diagonal matrix of the natural occupations.

## The eff-AOs as basis functions

The definition of the eff-AOs can be trivially inverted, and one gets

$$\varphi_j^A(\mathbf{r}) = \sum_{\mu=1}^{n_A} U_{j\mu}^{A*} \sqrt{\lambda_{\mu}^A} \chi_{\mu}^A(\mathbf{r}) = w_A(\mathbf{r}) \sum_{\mu=1}^{n_A} U_{j\mu}^{A*} \sqrt{\lambda_{\mu}^A} \chi(\mathbf{r})_{\mu} \quad .$$

Owing to this result, the MO-s can be written as linear combinations of the “effective AO-s” of different atoms:

$$\varphi_i(\mathbf{r}) = \sum_A^{N_{at}} w_A(\mathbf{r}) \varphi_i^A(\mathbf{r}) = \sum_A^{N_{at}} \sum_{\mu=1}^{n_A} U_{i\mu}^{A*} \sqrt{\lambda_{\mu}^A} \chi_{\mu}^A(\mathbf{r}) = \sum_A^{N_{at}} \sum_{\mu=1}^{n_A} c_{\mu i}^A \chi_{\mu}^A(\mathbf{r}) \quad .$$

# Hilbert-space analysis within QTAIM

The atomic population calculated by Mulliken analysis in the basis of the “effective AO-s” is **equal** to that obtained by the 3D QTAIM analysis.

$$Q_A^{AIM} = \int_{\Omega_A} \rho(\mathbf{r}) d\mathbf{v} = 2 \int_{\Omega_A} \sum_{i=1}^n |\varphi_i(\mathbf{r})|^2 d\mathbf{v} = 2 \sum_{i=1}^n \int_{\Omega_A} |\varphi_i(\mathbf{r})|^2 d\mathbf{v} = 2 \sum_{i=1}^n Q_{ii}^A .$$

$$Q_A^{LCAO} = \sum_{\nu \in A} D_{\nu\nu} = 2 \sum_{i=1}^n \sum_{\nu \in A} |C_{\nu i}|^2 = 2 \sum_{i=1}^n \sum_{\mu=1}^{n_A} |U_{i\mu}^{A*} \sqrt{\lambda_\mu}|^2 = 2 \sum_{i=1}^n Q_{ii}^A$$

Analogously, Wiberg bond orders calculated on the basis of eff-AOs are exactly equal to the Delocalization index, and so on.

Similar relationships are also found for fuzzy atomic domains such as TFVC, the only difference being the eff-AOs do not form an orthonormalized set.

# Properties

- In practice, the dimension of the eff-AO basis has nothing to do with the dimension of the LCAO basis (if any) used in the original calculations.
- The maximum number of eff-AOs that can be obtained is  $N_{at} \times n$ , which can be less than the total number of the LCAO basis functions.
- Experience shows that the number of eff-AOs with significant occupation numbers on each atom is limited, and typically much smaller than the number  $n$  of the doubly occupied orbitals in the whole molecule.
- The remaining eff-AOs with very small occupation numbers have a marginal significance.

# Visualization of the eff-AOs

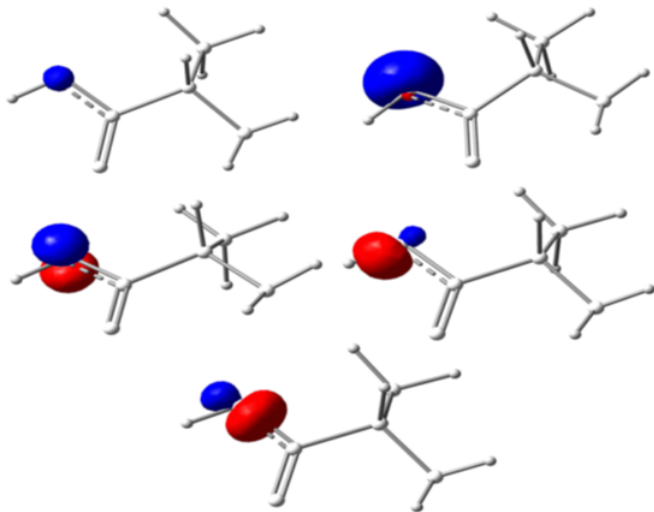
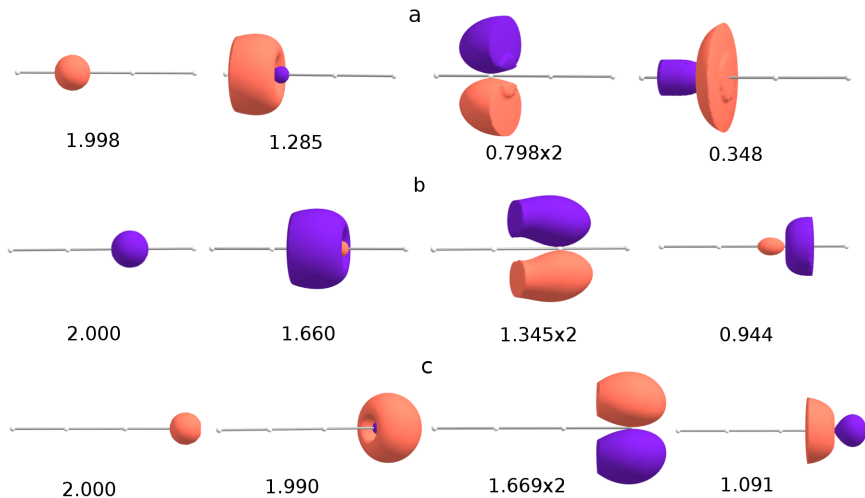


Figure 1: Occupied oxygen eff-AOs in alanine molecule (TFVC)

# Visualization of the eff-AOs



**Figure 2:** Highly occupied carbon (a), nitrogen (b), and oxygen (c) orbitals in the HCNO molecule (QTAIM)



# Formal oxidation states from real-space analysis

## Formal oxidation states

- The concept of oxidation state (OS) is ubiquitous in transition metal (TM) chemistry and in the study of redox and catalytic reactions.
- Many properties such as reactivity, spin-state, spectroscopic and geometrical features of TM complexes are often rationalized on the basis of the oxidation state of the metal center.
- Formal OS are obtained by assigning integer number of electrons to the atoms/ligands according to some rules. In complicated bonding situations involving non-innocent ligands or in intermediates or transition states of catalytic reactions the formal OS assignment may be rather ambiguous.

## IUPAC's revision of the concept

- A revision of the concept of OS was requested to a number of experts. Pavel Karen's IUPACs Technical Report summarized their conclusions.
- New definition was provided, along with algorithms to determine the OS in molecular systems and solids.

The OS of a bonded atom equals its charge after **ionic approximation**.

P. Karen et al, Pure Appl. Chem. 2014 86 1017-1081.

P. Karen, Angew. Chem. Int. Ed. 2015, 54, 2-13

# Lewis structure as a starting point

## The Ionic Approximation

- For practical applications, to the more electronegative atom, according to Allen's electronegativity scale.
- In bonds involving two atoms of the same element, divided equally.

## Limitations/ambiguities of the new definition

- Comes with exceptions for the application of the IA (when the atom with higher Allen EN is a net acceptor of electrons).
- All electrons ( $\sigma$  or  $\pi$ ) of a bond are necessarily assigned to the same atom after the IA.
- Different Lewis structures may lead to different OS assignment.
- EN is regarded as a genuine free atom property, hence same atoms are treated on equal footing irrespective of their chemical environment.

# Conventional population analysis fail

Thom *et al.* illustrated how neither Mulliken charges nor Mulliken spin populations match in general with the oxidation state.

Mulliken (left) and TFVC (right) charges for a set of transition metal complexes.

	Cl <sup>-</sup>	H <sub>2</sub> O <sup>HS</sup>	H <sub>2</sub> O <sup>LS</sup>	CN <sup>-</sup>	CO	Cl <sup>-</sup>	H <sub>2</sub> O <sup>HS</sup>	H <sub>2</sub> O <sup>LS</sup>	CN <sup>-</sup>	CO
V <sup>II</sup>	0.98	1.12	-	0.05	0.64	1.60	1.76	-	1.60	1.64
Mn <sup>II</sup>	1.10	1.24	1.18	0.10	0.64	1.36	1.64	1.70	1.53	1.53
Mn <sup>III</sup>	0.93	1.58	1.52	0.35	0.80	1.46	2.05	2.09	1.58	1.67
Fe <sup>II</sup>	0.86	1.22	1.15	0.01	0.51	1.27	1.63	1.77	1.46	1.44
Fe <sup>III</sup>	0.99	1.64	1.48	0.24	0.66	1.44	2.06	1.98	1.49	1.58
Ni <sup>II</sup>	0.99	-	1.08	-0.19	0.31	1.27	-	1.59	1.24	1.30
Zn <sup>II</sup>	1.02	-	1.06	-0.03	0.52	1.25	-	1.45	1.15	1.19

## *ab initio* determination of oxidation states

- A number of empirical approaches to *correlate* atomic distances or populations with transition metal (TM) oxidation states (bond-valence sum scheme).
- Sit et al.[1] used projection techniques to obtain d orbital populations of central metal in transition metal (TM) complexes
- Sit et al.[2] and Vidossich et al. [3] also used the positions of the centers of gravity of maximally-localized Wannier functions/localized orbitals.
- Localized Orbital Bond Analysis (LOBA) [4]: MO localization followed by population analysis
- ELIBON (electron-localizability-based oxidation number)[5], usually applied to solids.

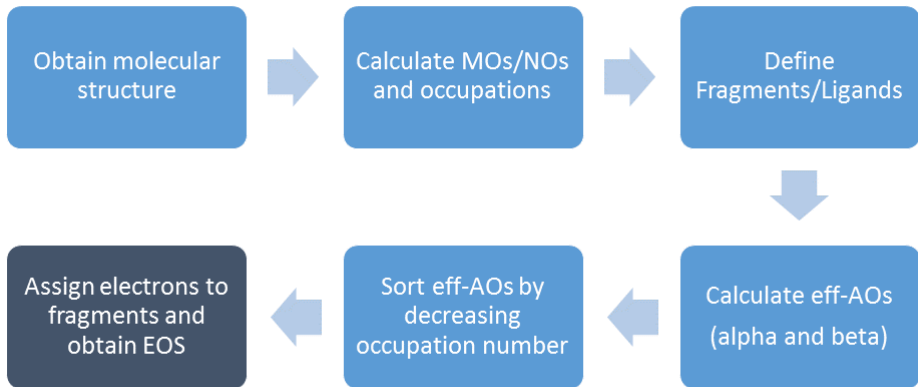
# Using effective fragment orbitals (effOs)

- Define atoms/fragments/ligands of the system

$$w_{\Gamma_i}(\vec{r}) \equiv \sum_{A \in \Gamma_i} w_A(\vec{r}) \quad \mathbf{s}^{\Gamma_i} \neq \sum_{A \in \Gamma_i} \mathbf{s}^A$$

- Obtain spin resolved effOs for all  $\Gamma_i$  fragments
- Distribute the number of electrons among the fragments according to occupations of their effOs
- Get most appropriate atom/ligand Oxidation States and R% index.
- If necessary, the OS for the individual atoms of a given fragment can be subsequently obtained by getting the eff-AOs for every atom and distributing the number of alpha and beta electrons that were assigned to the ligand in the previous step.

# EOS analysis: a very simple scheme



E. Ramos-Cordoba, V. Postils, P. Salvador, JCTC 2015, 11, 1501.



# EOS analysis at a glance

Alpha effective atomic orbitals for  $\text{Fe}(\text{CN})_6^{3-}$  at the B3LYP/6-31G(d) level of theory.

**Total number of alpha electrons: 54**

Fe				C		N	
eff-AO type	Occup	eff-AO type	Occup	eff-AO type	Occup	eff-AO type	Occup
1s	1.000	3d	0.876	1s	0.996	1s	1.000
2s	1.000	3d	0.848	2s	0.694	2s	0.997
2p	1.000	3d	0.848	2p	0.235	2p	0.728
2p	1.000	3d	0.344	2p	0.232	2p	0.703
2p	1.000	3d	0.314	2p	0.153	2p	0.701
3s	0.991	4s	0.103		0.014		0.020
3p	0.983		0.050		0.013		
3p	0.983		0.050				
3p	0.982		0.049				

# EOS analysis at a glance

Beta effective atomic orbitals for  $\text{Fe}(\text{CN})_6^{3-}$  at the B3LYP/6-31G(d) level of theory.

**Total number of beta electrons: 53**

Fe				C		N	
eff-AO type	Occup	eff-AO type	Occup	eff-AO type	Occup	eff-AO type	Occup
1s	1.000	3d	0.827	1s	0.996	1s	1.000
2s	1.000	3d	0.827	2s	0.708	2s	0.997
2p	1.000	3d	0.305	2p	0.246	2p	0.729
2p	1.000	3d	0.263	2p	0.239	2p	0.702
2p	1.000	3d	0.099	2p	0.153	2p	0.665
3s	0.990	4s	0.049		0.150		0.021
3p	0.982		0.048				
3p	0.982		0.048				
3p	0.982		0.032				

# EOS analysis at a glance

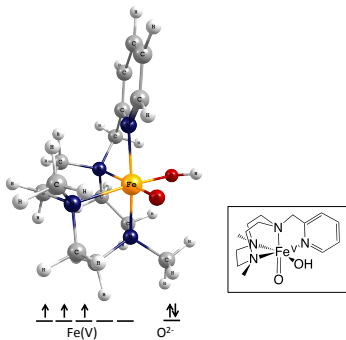
Oxidation states, last occupied eff-AO (LO) and first unoccupied eff-AO (FU) orbitals for  $\text{Fe}(\text{CN})_6^{3-}$ .

Atom/ Fragment	O.S.	$\lambda_{\text{LOEFF}}^{\alpha}$	$\lambda_{\text{HUEFF}}^{\alpha}$	$\lambda_{\text{LOEFF}}^{\beta}$	$\lambda_{\text{HUEFF}}^{\beta}$
Fe	+3	0.850	0.344	0.827	0.305
C	+2	0.694	0.235	0.708	0.246
N	-3	0.701	0.020	0.665	0.021

From the "frontier" occupation numbers one can **derive** a simple index  $R(\%)$  to **quantify** how close is the electron distribution provided by the actual wavefunction to the formal picture of the oxidation states.

$$R(\%) \equiv 100 \min(1, \max(0, \lambda_{\text{LO}}^{\sigma} - \lambda_{\text{HU}}^{\sigma} + 1/2)), \text{ for } \sigma = \alpha, \beta.$$

# Beyond spin density analysis ( $\rho_{Fe}^s = 2.05$ )



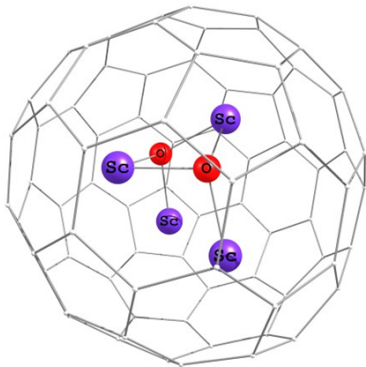
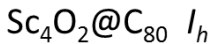
$\sigma$	$\lambda^\sigma$				
$\alpha$	<b>0.934</b>	<b>0.932</b>	<b>0.930</b>	0.495	0.451
$\beta$	0.163	0.393	0.462	0.322	0.349
$\alpha - \beta$	0.771	0.539	0.468	0.173	0.102

d-type eff-AO occupations of the Fe atom in the  $[\text{Fe}(\text{Pytacn})\text{O}(\text{OH})]^{2+}$  species.

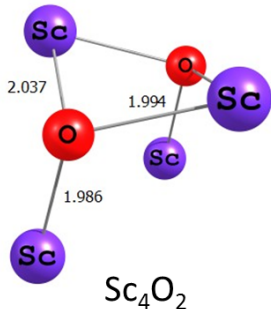
# High oxidation states

	Fragment	EOS	$\lambda_{LO}^{\alpha} / \lambda_{FU}^{\alpha}$	R(%)	$N_{\text{met}}$
$\text{FeO}_4^{2-}$	Fe	+6	0.919 / <b>0.452</b>	68.8	+1.52
	O	-2	<b>0.641</b> / 0.011		
$\text{ReO}_4^-$	Re	+7	0.954 / <b>0.295</b>	88.8	+2.71
	O	-2	<b>0.684</b> / 0.010		
$\text{ReH}_9^{2-}$	Re	+7	0.978 / <b>0.429</b>	60.4	+1.52
	H	-1	<b>0.533</b> / 0.024		
$\text{OsO}_4$	Os	+8	0.964 / <b>0.404</b>	72.9	+2.61
	O	-2	<b>0.633</b> / 0.009		
$\text{IrO}_4^+$	Ir	+9	0.970 / <b>0.509</b>	57.6	+2.54
	O	-2	<b>0.585</b> / 0.009		
$\text{PtO}_4^{2+}$	Pt	+10	<b>0.542</b> / 0.197	40.6	+2.37
	O	-2	0.691 / <b>0.536</b>		

# Justification of the ionic model in EMFs



Cage Charge	R(%)
-6	100



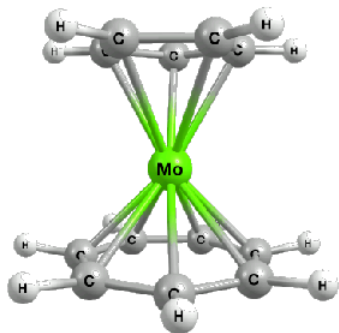
Atom/ Fragment	Oxidation state
Sc <sub>eq</sub>	3+
Sc <sub>ax</sub>	2+
O	2-

R(%) = 60%

# Justification of the ionic model in EMFs

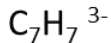
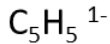
Cage	Isomer	Cluster	Spin state	EOS Cage	R(%)
C66	4059	Sc <sub>2</sub>	singlet	-6	100
C66	4338	Sc <sub>2</sub>	singlet	-4	88
C68	6073	Sc <sub>2</sub> C <sub>2</sub>	singlet	-4	92
C80	IPR 5	Sc <sub>2</sub> C <sub>2</sub>	singlet	-4	100
C80	ih	Sc <sub>3</sub> C <sub>2</sub>	doublet	-6	100
C78	d3h IPR	ScN <sub>3</sub>	singlet	-6	100
<b>C82</b>	<b>C<sub>2</sub></b>	<b>U</b>	<b>triplet</b>	<b>-4</b>	<b>64</b>
<b>C82</b>	<b>C<sub>2v</sub></b>	<b>U</b>	<b>triplet</b>	<b>-3</b>	<b>67</b>

# Hapticity and aromaticity



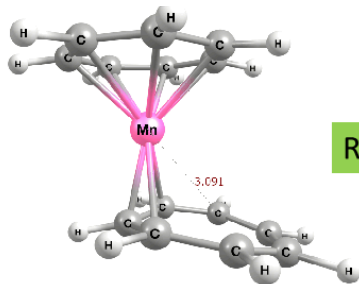
MoCp(C<sub>7</sub>H<sub>7</sub>), **37**

R(%) = 64



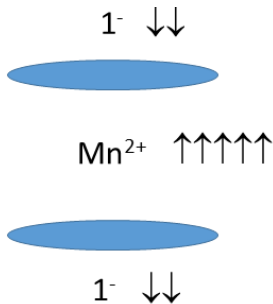


# Hapticity and aromaticity



$\text{Mn}(\text{C}_7\text{H}_7)_2\text{-1D, 54}$

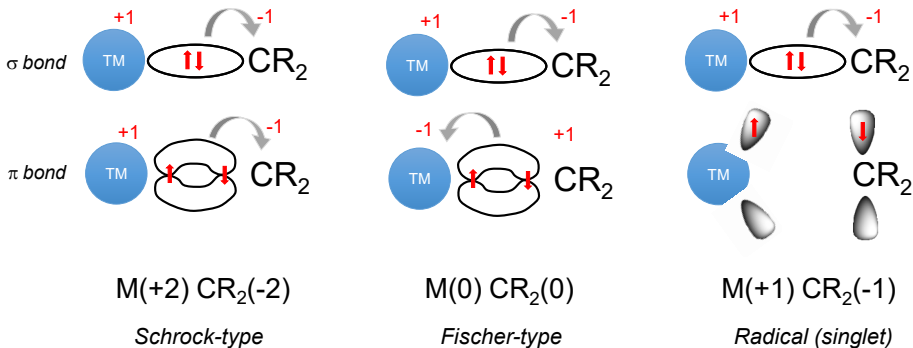
$R(\%) = 84$



$8 \pi$  electrons  $\rightarrow$  **Baird aromatic**

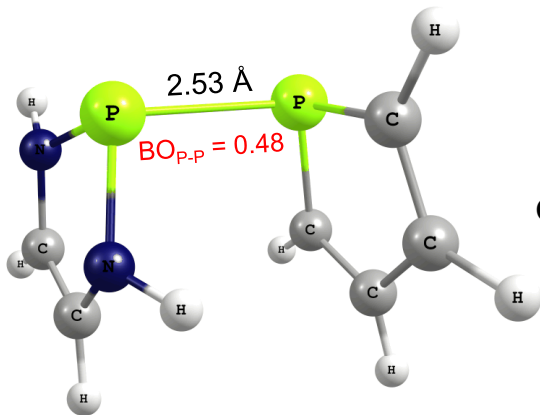
- The formal charge on the ligands depend upon their local spin and can be rationalized in terms of Huckel and Baird aromaticities
- Partial aromaticity can be invoked for strong deviations from planarity (lower hapticities)

# Breakdown of the Ionic Approximation: carbenes



- The IA can not afford for a formal neutral carbene, where the two electrons of the  $\sigma$  bond are assigned to the carbene and the  $\pi$  ones to the metal center.

# Breakdown of the Ionic Approximation: adducts



$P(NC_2H_4)-P(C_4H_4)$ , 55

**EOS**

$P(NC_2H_4)$  +1

$P(C_4H_4)$  -1

**Global R(%) = 81.9%**

	$\lambda_{LO} / \lambda_{FU}$
$P(NC_2H_4)$	0.948 / 0.308
$P(C_4H_4)$	0.627 / 0.030

# Summary

- Oxidation states can be extracted from the analysis of the effective atomic/fragment orbitals and their occupancies with a strategy of general application (closed-shell systems, KS-DFT or correlated WFs)
- The EOS method is reconciled with the empirical assignation in most cases without any external guidance or exception.
- A better empirical approach to oxidation state assignation could start by identifying subsystems of enhanced stability in the molecular system in the first place, according to known rules (octet, 18e,...)