J Math Chem (2013) 51:289–296
DOI 10.1007/s10910-012-0083-x

ORIGINAL PAPER

Mathematical aspects of the LCAO MO first order density function (5): centroid shifting of MO shape functions basis set, properties and applications

Ramon Carbó-Dorca

Received: 14 June 2012 / Accepted: 6 September 2012 / Published online: 15 September 2012

Abstract In MO SCF theory and related computational levels, the first order density function (DF) can be considered as a linear combination of the MO shape functions (ShF) set. This work studies the possibility of constructing a centroid function from the ShF set elements using it to perform a ShF set origin shift, while permitting the DF decomposition in two well defined function terms. The properties and consequences of such a simple operation are analyzed in deep.

Keywords Density function (DF) · SCF theory · MO shape functions (ShF) · MO ShF centroid · MO centroid shifted ShF · Ionization potentials · Atomic shell approximation DF

1 Introduction

This paper is written as another chapter of a series devoted to discuss some aspects of the first order density function (DF), see [1–4] for previous work notice. Recently, another work dealt with the origin shift technique applied [5] to the DF tag set of a quantum object set, see [6] for an up to date account of the involved terminology.

In the present chapter, it will be analyzed the structure of the MO shape function (ShF) set which acts as a basis set to construct the DF within all the computational frameworks [7] and from there it will be described how to construct a centroid ShF which can be employed to shift the original basis set into a new set of functions. Then, the use of these shifted functions, for instance to obtain an ionization potential operator set, will be given. The properties which can be associated to the shifted functions will

R. Carbó-Dorca (✉)
Institut de Química Computacional, Universitat de Girona, 17071 Girona, Catalonia, Spain
e-mail: quantumqsar@hotmail.com

be also discussed. A final brief account of a similar procedure which can be performed in the atomic shell approximation (ASA) [8–11] framework will be discussed.

2 The MO shape function (ShF) generating set

2.1 Definitions

It is obvious that operating under the MO framework, where a set of MO is well defined: $M = \{|I\rangle\,|I = 1, N\}$, as in SCF and related theories, then the first order DF can be considered constructed by a set of MO ShF contributions: $P = \{\sigma_I = |I\rangle\langle I|\,|I = 1, N\}$, whose elements correspond to the squared MO modules. With a proper MO occupation number set defined: $\{\omega_I\}$, one can construct as a linear combination of P the total DF [7]:

$$\rho = \sum_I \omega_I \sigma_I$$

The existing and necessary constraints to the above construction correspond to:

(1) Being v the number of electrons: $\sum_I \omega_I = v$.
(2) The Minkowski norms of the elements of the ShF set P are unity, corresponding to the MO set normalization: $\forall I : \langle\sigma_I\rangle = 1$.

2.2 Centroid origin shift

Generally speaking, the set P, considered as a set of vectors of some infinite dimensional space, is associated to an arbitrary origin. Nothing opposes nonetheless to the fact one can define a centroid ShF, using a simple average function [5]:

$$\sigma_C = N^{-1} \sum_I \sigma_I \rightarrow \langle\sigma_C\rangle = 1,$$

and thereafter the elements of the set P can be shifted with respect such a centroid function, constructing in this way a new set of centroid shifted ShF (CSShF):

$$\forall I : \theta_I = \sigma_I - \sigma_C \rightarrow Z = \{\theta_I\,|I = 1, N\}.$$

Of course, the elements of the CSShF set possess no longer the properties of the original ShF set. Also the Minkowski pseudonorm of the CSShF set Z can be written as:

$$\forall I : \langle\theta_I\rangle = \langle\sigma_I - \sigma_C\rangle = \langle\sigma_I\rangle - \langle\sigma_C\rangle = 0$$

That is the shifted elements have a null Minkowski pseudonorm. Moreover, if the original ShF set is certainly composed of N linearly independent elements, the CSShF

set is also composed of $N-1$ linearly independent elements [5]. The elements of the CSShF set Z can be rewritten as:

$$\forall I : \theta_I = N^{-1}\left((N-1)\,\sigma_I - \sum_{J\neq I}\sigma_J\right) = \sigma_I + N^{-1}\sum_{J\neq I}(\sigma_I + \sigma_J) \tag{1}$$

which constitutes an expression indicating that every CSShF can be also defined as the arithmetic mean function of the differences of the corresponding ShF of P with the rest of elements of the set P.

2.3 Approximate ionization potential operator

Then, supposing defined a set of Fock operators associated to every MO: $F = \{F_I\}$, it is interesting to consider that a general Koopmans theorem might apply, see for example [12–14]:

$$\forall I : \langle F_I \rho_I \rangle = \varepsilon_I \rightarrow -\varepsilon_I = I_I. \tag{2}$$

On the other hand, one can use $N-1$ elements of Z to construct an operator, say: Ω, such that the ionization potentials correspond to its expectation values when acting over the original MO ShF set:

$$\forall I : \langle \Omega \sigma_I \rangle = I_I.$$

Within a first order approach, such an operator can be built with the set of elements of Z, except the K-th, or what it is the same:

$$\forall K : \Omega_{(K)} = \sum_{J\neq K} a_{(K)J}\theta_J,$$

The set $A_{(K)} = \left\{a_{(K)J}\,|\,J \neq K\right\}$ is a set of undefined parameters. Then, it can be also written:

$$\forall I \neq K : \left\langle \Omega_{(K)}\sigma_I \right\rangle = I_I.$$

Using the Eq. (2) one can set up a system of $N-1$ equations with $N-1$ unknowns, which can determine the set: $A_{(K)}$ and thus the operator $\Omega_{(K)}$. Once obtained the operator $\Omega_{(K)}$, it can be also obtained an estimate of the K-th ionization potential I_K using: $\left\langle \Omega_{(K)}\sigma_K \right\rangle = I_K^{(*)} \approx I_K$.

The key elements for this approach are constituted by the similarity integrals between the MO ShF set:

$$\begin{aligned}\langle\theta_J\sigma_I\rangle &= \langle\sigma_J\sigma_I\rangle - N^{-1}\sum_{K\neq J}\langle(\sigma_J-\sigma_K)\,\sigma_I\rangle \\ &= \langle\sigma_J\sigma_I\rangle - N^{-1}\,(N-1)\,\langle\sigma_J\sigma_I\rangle + N^{-1}\sum_{K\neq J}\langle\sigma_K\sigma_I\rangle \\ &= Z_{JI} - (1-N^{-1})Z_{JI} + N^{-1}\sum_{K\neq J}Z_{KI} = N^{-1}\sum_{K}Z_{KI} = N^{-1}\,\langle|Z_I\rangle\rangle\end{aligned}$$

Where the column $\{|Z_I\rangle\,;\,I=1,N\}$ vector sets collect the column elements of the similarity Z.

3 Matrix representation of the centroid origin shift: a general structure

In fact, such a transformation like the centroid shifting of the origin of a MO ShF set is equivalent to the following procedure.

Construct a column vector with the MO ShF of the set P as elements:

$$|\boldsymbol{\sigma}\rangle = \{\sigma_I\}$$

then form a transformation matrix as:

$$\mathbf{T}_N = \mathbf{I}_N - N^{-1}(\mathbf{1}_N - \mathbf{I}_N)$$

where the subscript indicates the involved matrices are $(N\times N)$. The matrix $\mathbf{I}_N = \{\delta_{IJ}\}$ is the unit matrix and the matrix $\mathbf{1}_N = \{1_{IJ}=1\}$ is the unity matrix.

Use another column vector with the origin CSShF set as elements:

$$|\boldsymbol{\theta}\rangle = \{\theta_I\}$$

then it can be simply written:

$$|\boldsymbol{\theta}\rangle = \mathbf{T}_N\,|\boldsymbol{\sigma}\rangle\,.$$

The centroid origin shift matrix: $\mathbf{T}_N$ has one zero eigenvalue and a $N-1$ degenerate unit one. This is due to the fact that the unity matrix $N^{-1}\mathbf{1}_N$ has an eigenvalue equal to 1 and the rest are zeros. Thus, when to such a unity matrix a unit matrix is summed, the spectrum of the resultant matrix is zero and the rest ones. The eigenvector associated to the zero eigenvalue is the unity vector $|\mathbf{1}\rangle$ and the rest can be chosen as any set of orthogonal vectors, which at the same time are orthogonal to the unity vector. Therefore the matrix $\mathbf{T}_N$ is singular and thus the origin shift transformation becomes irreversible.

Thus, while the vector set $|\boldsymbol{\sigma}\rangle$ is a generator of a N-dimensional vector space, the set $|\boldsymbol{\theta}\rangle$ generates just a $(N-1)$-dimensional one. This is the same as to describe the similarity matrices: $\mathbf{Z} = \langle|\boldsymbol{\sigma}\rangle\langle\boldsymbol{\sigma}|\rangle = \{Z_{IJ} = \langle\sigma_I|\sigma_J\rangle\}$ as a metric matrix, and $\boldsymbol{\Theta} = \langle|\boldsymbol{\theta}\rangle\langle\boldsymbol{\theta}|\rangle = \{\Theta_{IJ} = \langle\theta_I|\theta_J\rangle\}$ as a Gram matrix with a zero eigenvalue.

One can also write:

$$\begin{aligned}\boldsymbol{\Theta} = \mathbf{T}_N\mathbf{Z}\mathbf{T}_N &= \left(\mathbf{I}_N - N^{-1}(\mathbf{1}_N - \mathbf{I}_N)\right)\mathbf{Z}\left(\mathbf{I}_N - N^{-1}(\mathbf{1}_N - \mathbf{I}_N)\right)\\ &= \mathbf{Z} - N^{-1}\left((\mathbf{1}_N - \mathbf{I}_N)\mathbf{Z} + \mathbf{Z}(\mathbf{1}_N - \mathbf{I}_N)\right) + N^{-2}(\mathbf{1}_N - \mathbf{I}_N)\mathbf{Z}(\mathbf{1}_N - \mathbf{I}_N)\\ &= \mathbf{Z} - N^{-1}(\{\mathbf{1}_N;\mathbf{Z}\} - 2\mathbf{Z}) + N^{-2}(\langle\mathbf{Z}\rangle\,\mathbf{1}_N - \{\mathbf{1}_N;\mathbf{Z}\} + \mathbf{Z})\end{aligned}$$

which implies that:

$$\begin{aligned}\forall I, J : \Theta_{IJ} = Z_{IJ} &- N^{-1}\left(\langle|\mathbf{z}_I\rangle\rangle + \langle\langle\mathbf{z}_J|\rangle - 2Z_{IJ}\right)\\ &+N^{-2}\left(\langle\mathbf{Z}\rangle - (\langle|\mathbf{z}_I\rangle\rangle + \langle\langle\mathbf{z}_J|\rangle) + Z_{IJ}\right)\end{aligned}$$

where, as defined before the row $\{\langle\mathbf{z}_I|\,;\,I=1,N\}$ and column $\{|\mathbf{z}_J\rangle\,;\,J=1,N\}$ vector sets are the elements of the similarity matrix $\mathbf{Z}$.

3.1 Properties of the translation matrix

Now it is easy to write the following equalities:

$$\left(N^{-1}\mathbf{1}_N\right)^2 = N^{-2}\mathbf{1}_N^2 = N^{-2}N\mathbf{1}_N = N^{-1}\mathbf{1}_N$$

showing that $N^{-1}\mathbf{1}_N$ it is an idempotent matrix. It can also be written:

$$(\mathbf{T}_N)^2 = (\mathbf{I}_N - N^{-1}(\mathbf{1}_N - \mathbf{I}_N))^2 = (1+N^{-1})\mathbf{T}_N - N^{-2}\mathbf{1}_N$$

4 Other mathematical properties of the CSShF set and the decomposition of the DF

The CSShF set has zero as origin, a property which can be easily proved simply using:

$$\theta_C = N^{-1}\sum_J \theta_J = N^{-1}\sum_J(\sigma_J - \sigma_C) = \sigma_C - \sigma_C = 0$$

One can also write:

$$\forall I : \theta_I = \sigma_I - \sigma_C \rightarrow \sigma_I = \theta_I + \sigma_C$$

therefore, the attached DF can be constructed as:

$$\rho = \langle \boldsymbol{\omega} | \boldsymbol{\sigma} \rangle = \sum_I \omega_I \sigma_I = \sum_I \omega_I \theta_I + \sum_I \omega_I \sigma_C$$
$$= \theta + \nu \sigma_C = \theta + \mu \sum_I \sigma_I = \langle \boldsymbol{\omega} | \boldsymbol{\theta} \rangle + \mu \langle \mathbf{1} | \boldsymbol{\sigma} \rangle$$

where: $\nu = \sum_I \omega_I$ and thus: $\mu = \nu N^{-1}$ can be considered some kind of average MO occupation.

Thus, one can also write:

$$\langle \boldsymbol{\omega} | \boldsymbol{\theta} \rangle = \langle \boldsymbol{\omega} | \boldsymbol{\sigma} \rangle - \mu \langle \mathbf{1} | \boldsymbol{\sigma} \rangle = \langle \boldsymbol{\omega} - \mu \mathbf{1} | \boldsymbol{\sigma} \rangle = \sum_I (\omega_I - \mu) \sigma_I .$$

Of course the following result will always hold: $\sum_I (\omega_I - \mu) = \nu - N\mu = 0$, and using: $\sigma_S = \langle \mathbf{1} | \boldsymbol{\sigma} \rangle = \sum_I \sigma_I$, then one can finally express the MO DF as: $\rho = \theta + \mu \sigma_S$.

This last result corresponds to a decomposition of the DF into a sum made by the construction of a DF using the CSSHF set and another function made just by the sum of the MO ShF multiplied by the average MO occupation. The origin shifted DF can be also seen as a DF where the occupation numbers are the real ones minus the average MO occupation number.

All this appears as a weird and curious vantage point to look at the usual DF. In any case, it can be interesting to see, at least in a simplified way, how the shape of the origin shifted DF can be. For instance, in closed shell SCF procedures with N doubly occupied MO, the average occupation number is: $\mu = 2$, so the origin shifted DF is the null function and obviously: $\rho = \mu \sigma_S$. The above decomposition will become interesting enough, when considering DF of open shell or multiconfigurational origin.

5 Centroid shifting when the DF is defined as a sum of atomic ShF terms as in the atomic shell approximation

In this case:

$$\rho = \sum_A Q_A \sigma_A \leftarrow \forall A : \langle \sigma_A \rangle = 1 \wedge \sum_A Q_A = \nu$$

One can define origin shifted atomic ShF too, constructing the atomic centroid function:

$$\sigma_X = N_a^{-1} \sum_A \sigma_A$$

where N_a is the number of atoms. The atomic CSShF are made then by the differences:

$$\forall A : \theta_A = \sigma_A - \sigma_X .$$

One arrives to similar conclusions as when employing the MO ShF decomposition. However, distinction must be taken into account between H and other atoms or between atoms of different atomic table periods. Otherwise one might obtain weird atomic population differences. The origin shift better to be performed with a ShF weighted average made using the sum of averages over atoms of the same period:

$$\sigma_X = \sum_P \left(N_P^{-1} \sum_{A \in P} \sigma_A^{(P)} \right) = \sum_P \sigma_P$$

6 The case of degeneracy on any DF in MO theory

The previous discussion can be applied to the particular case of a specific DF expressed in the MO theoretical framework. From the constraints on MO occupation numbers as initially described, the set of equalities can be trivially deduced:

$$\forall J : \omega_J = \nu - \sum_{I \neq J} \omega_I$$

and from these above equalities it can be readily written the set of equivalent DF:

$$\forall J : {}^J\rho = \nu\sigma_J + \sum_{I \neq J} \omega_I \left(\sigma_I - \sigma_J\right) = \rho,$$

where the upper left superscript refers to the redefined MO equality used. Such process defines a vector component set, which can be easily expressed like:

$$\left| {}^S\rho \right\rangle = \nu \left|\sigma\right\rangle + \mathbf{F} \left|\omega\right\rangle \tag{3}$$

where: $\left|\sigma\right\rangle = \{\sigma_I\} \wedge \left|\omega\right\rangle = \{\omega_I\} \wedge \mathbf{F} = \{F_{IJ} = \sigma_I - \sigma_J\}$.

Equation (3) is related to the initial definition of CSShF, as can be found simply observing Eq. (1), which upon the previous definition of the matrix $\mathbf{F}$ and defining a vector containing as elements the CSShF like: $\left|\theta\right\rangle = \{\theta_I \,|\, I = 1, N\}$, might be easily rewritten as:

$$\left|\theta\right\rangle = N^{-1}\mathbf{F} \left|\mathbf{1}\right\rangle.$$

Finally, taking now into account that the resulting vector $\left| {}^S\rho \right\rangle$ might be expressed by means of the unity vector and the DF acting as a scalar, it can be also written:

$$\left| {}^S\rho \right\rangle = \rho \left|\mathbf{1}\right\rangle \rightarrow \forall I :^S \rho_I = \rho.$$

The curious, if not interesting, conclusion is that there are N equivalent or degenerate ways to write a given DF.

Moreover, in order to apply such a result to the DF associated ShF, then Eq. (3) can be divided by the number of electrons, so it will result into:

$$|^S\sigma\rangle = |\sigma\rangle + \mathbf{F}|\xi\rangle \leftarrow |^S\sigma\rangle = \nu^{-1}|^S\rho\rangle \wedge |\xi\rangle = \nu^{-1}|\omega\rangle$$

7 Conclusions

Centroid shifting of the MO ShF results into a good deal of properties of the transformed DF components, which can be employed in various ways. Among other possibilities to obtain an approximate set of operators which can be employed to estimate ionization potentials. A matrix representation of the centroid origin shift can be easily described, the decomposition of the DF into two well defined terms and the extension of the centroid shift to ASA approximate DF have been described. Finally, the degeneracy of the DF with respect of the MO occupation number constraint can be easily described and related to the centroid origin shift of the MO ShF components.

References

1. R. Carbó-Dorca, Mathematical aspects of the LCAO MO first order density function (1): atomic partition, metric structure and practical applications. J. Math. Chem. **43**, 1076–1101 (2008)
2. R. Carbó-Dorca, Mathematical aspects of the LCAO MO first order density function (2): relationships between density functions. J. Math. Chem. **43**, 1102–1118 (2008)
3. R. Carbó-Dorca, P. Bultinck, Mathematical aspects of the LCAO MO first order density function (3): a general localization procedure. J. Math. Chem. **43**, 1069–1075 (2008)
4. R. Carbó-Dorca, E. Besalú, Mathematical aspects of the LCAO MO first order density function (4): notes on the connection of Taylor series of electronic density (TSED) function with the holographic electronic density theorem (HEDT) and the Hohenberg–Kohn theorem (HKT). J. Math. Chem. **49**, 836–842 (2011)
5. R. Carbó-Dorca, E. Besalú, Centroid origin shift of quantum object sets and molecular point clouds: description and element comparisons. J. Math. Chem. **50**, 11161–11178 (2012)
6. R. Carbó-Dorca, E. Besalú, Shells, point cloud huts, generalized scalar products, cosines and similarity tensor representations in vector semispaces. J. Math. Chem. **50**, 210–219 (2012)
7. R. Carbó-Dorca, Diagonal coefficient representation of density functions and quantum similarity measures. J. Math. Chem. **44**, 621–627 (2008)
8. P. Constans, R. Carbó, J. Chem. Inf. Comp. Sci. **35**, 1046–1053 (1995)
9. P. Constans, X. Fradera, L. Amat, R. Carbó, Quantum molecular similarity measures (QMSM) and the atomic shell approximation (ASA), in *Proceedings of the 2nd Girona Seminar on Molecular Similarity. Advances in Molecular Similarity*, July 1995, vol. 1. (Jai Press Inc., Greenwich, CT, 1996), pp. 187–211
10. L. Amat, R. Carbó-Dorca, J. Comput. Chem. **18**, 2023–2029 (1997)
11. L. Amat, R. Carbó-Dorca, J. Comput. Chem **20**, 911–920 (1999)
12. B.N. Plakhutin, E.V. Gorelik, N.N. Breslavskaya, J. Chem. Phys. **125**, 204110-10 (2006)
13. B.N. Plakhutin, E.R Davidson, J. Phys. Chem. A **113**, 12386–12395 (2009)
14. E.R Davidson, B.N. Plakhutin, J. Chem. Phys. **132**, 184110–184114 (2010)