On the statistical interpretation of density functions: Atomic shell approximation, convex sets, discrete quantum chemical molecular representations, diagonal vector spaces and related problems

Ramon Carbó-Dorca

Institute of Computational Chemistry, University of Girona, Girona, Spain

Following a statistical interpretation analysis, as a discrete probability distribution, of the atomic shell approximation linear coefficient set, a new description of quantum object sets is given, using a discrete representation, based on diagonal vector spaces. Previous definitions, involving essentially tagged sets and vector semispaces, are used.

1. Introduction

The term Quantum Object (QO) [7] will be used here as a synonym of any microscopic system composed of a numerable set of particles, associated to a probability Density Function (DF). Within the current quantum mechanical structure the QO DF acts as a descriptor, which possesses all the information contained in the system.

Two papers have been written prior to this one. They were intended to study a new point of view in order to describe Quantum Object Sets (QOS). In a first instance, the concept of Boolean Tagged Set (BTS) was introduced as a sound alternative to the fuzzy set theoretical framework [3] and applied to the chemical objects description problem. In a second step, this initial BTS definition was extended in such a way as to encounter the realm of QOS, and thus, BTS were transformed into Positive Definite (PD) TS [4], using DF as the tag set part. At the same time, the structure of PD TS was developed for atoms and molecules, seen from the point of view of QO, acting as the background set. Finally, some applications of the TS theory related to Quantum Similarity (QS) have been developed.

Quantum Similarity Measures (QSM) [5] are essentially connected with PD functions and operators [8]. Recent research on approximate forms of the atomic first-order DF, whose general framework was named Atomic Shell Approximation (ASA) [9,10], has brought to the front line not only the PD nature of the studied DF, but also his membership to a Convex Set (CS) structure.

One of the main contributions of the present paper will be constituted by the definition of n-dimensional Diagonal Vector Spaces (DVS). The objective of introducing this DVS will be to find some discrete vector representation in order that it can consistently fit some usual properties of ∞-dimensional Hilbert spaces [2], containing the relevant functions, which are subsequently employed to describe QO, in accordance with von Neumann's [15] point of view. Thus, the main concern here will be to obtain, in a natural way, the CS structure of the approximate DF within DVS, in the same natural manner as the DF is obtained from the squared module of the QO system wavefunction.

Before achieving this end, various definitions, based on previous studies, will be given first. Then an analysis of the nature of the linear coefficients appearing in some expressions of the first-order DF, as in the ASA formalism, will be given. This preliminary discussion will reveal that the ASA coefficients can be associated to a discrete probability distribution. Finally, the need of DVS will appear as a natural choice.

2.   Definitions

Various definitions will be given first, in order to make clear the forthcoming discussion.

(a) *Quantum object*. Any microscopic sized system can be called a QO, if it can be constructed with a numerable set of particles, possessing all the extractable information contained within a probability distribution – the system's DF. The DF is a PD function of random variables made by the system's particle positions or momenta. When dealing with non-stationary system states, time should be added as an additional random variable. Due to this attachment between QO and DF, one can refer to the DF as a QO *continuous* descriptor.

(b) *Tagged set*. A TS, T, is a Cartesian product of any set, S, the background set, and a well-delimited set made of PD mathematical objects, P, called the tag set; that is, $T = S \times P$. As an example, a set of molecular structures can act as the background set, and the set of DF, associated to these molecular structures for some well-defined state, via quantum theory, can be considered as the tag set. Such a TS can be called a QO TS, in short, a QO Set (QOS). Another nice example of the tagged set, related to the previous one, is constituted by some molecules acting as the background set and the attached, with non-negative elements, topological matrices constituting the tag set part.

(c) *Vector semispace*. A vector space is defined over the positive real field, $\mathbf{R}^+$, only and constructed with the usual operations: addition and product by a scalar. In a Vector Semispace (VSS) the additive Abelian group is substituted by a *semigroup* [12]. As a consequence, no vector reciprocal elements can be present in a VSS, and the linear combination of coefficients will be made by positive real

numbers only. A 2-dimensional image of this kind of VSS may be given by the vectors lying over the positive $\{+x, +y\}$ quadrant of the real plane.

(d) *Convex set.* By *convex conditions* $K_n(\mathbf{c})$, held over an n-dimensional vector $\mathbf{c}$, it is understood that its elements obey the following properties:

$$K_n(\mathbf{c}) = \left\{ \mathbf{c} = \{c_i\} \in V_n(\mathbf{R}^+) \colon c_i > 0, \ \forall i \wedge \sum_i c_i = 1 \right\}. \tag{1}$$

A convex set is a set, made of vectors as elements, whose linear combinations and elements altogether fulfil convex conditions. The set of linear combinations of a subset belonging to a VSS, whose coefficients and generating vectors are constrained within the unit sphere, is a convex set.

(e) *Positive definite operators.* A PD operator, $\Omega > 0$, acting over a vector space, $\mathcal{V}$, such that $\Omega : \mathcal{V} \to \mathcal{V}$, can be defined in the usual way as follows:

$$\forall x \in \mathcal{V} \wedge x \neq \theta \colon \quad \Omega > 0 \ \Rightarrow \ \langle x | \Omega | x \rangle > 0. \tag{2}$$

Suppose that convex conditions $K_n(\mathbf{w})$ hold over a coefficient vector $\mathbf{w}$. Suppose that known a set of PD operators $\mathcal{W} = \{\Omega_\alpha\}$, the convex linear combination

$$\Omega = \sum_\alpha w_\alpha \Omega_\alpha \wedge K_n(\mathbf{w}) \tag{3}$$

produces a new PD operator. The DF can be considered as PD operators too, then this property also applies to this collection of functions.

(f) *Discrete QO description.* Suppose a known QOS: $\mathbf{Q} = \mathbf{S} \times \mathbf{P}$. Then, this will mean that for every element $s \in S$ in the QOS background set part, there exists a DF, $\rho \in P$, in the tag set part, which is a continuous descriptor of s. A *discrete* representation of the involved QO may be obtained choosing a PD operator, Ω, and defining the QSM [4]:

$$Q = S \times P \colon \quad \forall \rho_a, \rho_b \in P \ \Rightarrow \ z_{ab} = \int \rho_a \Omega \rho_b \, dV. \tag{4}$$

The matrix $\mathbf{Z} = \{z_{ab}\}$, collecting all the integrals between the DF of the QOS, can be considered as a row hypervector formed by a set of column vectors: $\mathbf{Z} = \{\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_n\}$. Every column vector, $\mathbf{z}_a$, say, is formed by the QSM involving the QO, s_a, DF, ρ_a, with all the QO elements in $\mathbf{S}$, including itself. The involved operator being PD and considering that the QOS is made of essentially different QO, then the matrix $\mathbf{Z}$ could be considered as PD. The whole matrix may belong to some n-dimensional VSS, $\mathcal{Z}_n(\mathbf{R}^+)$.

Then, starting from here, a new TS, Θ, can be constructed with the same background set part as the former QOS $\mathbf{Q}$, but with the tag set part formed by the columns of $\mathbf{Z}$, that is, $\Theta = \mathbf{S} \times \mathbf{Z}$. This corresponds to obtaining a PD operator dependent discrete representation of the QO. One can name the column vectors $\mathbf{z}_p \in \mathbf{Z}$ as *discrete* descriptors of the QO.

3. ASA first-order density functions

3.1. The ASA formalism

A first-order DF can be approximated by means of the so-called ASA function [9,10]. These approximated functions for atoms can be constructed in the following fashion. Suppose a set of ns-type AO $\{s_i(\alpha_i, \mathbf{r})\}$, then an atomic ASA function can be written as

$$\rho_A(\mathbf{r}) = \sum_i c_i |s_i(\mathbf{r})|^2 \wedge K_n(\mathbf{c} = \{c_i\}), \tag{5}$$

where the coefficient vector $\mathbf{c} = \{c_i\}$ has to be fitted to any given "*ab initio*" atomic DF, previously known, using a constrained least squares technique, in order to keep valid the convex conditions $K_n(\mathbf{c})$. Constrained fitting within the convex restrictions associated to equation (5), over the coefficient set, can be easily accomplished using an Elementary Jacobi Rotation (EJR) scheme [13].

The main idea, before applying the EJR technique to fit the atomic ASA function to an exact DF, appears when gathering the coefficients into the column vector $\mathbf{c}$, and associating to it a new column vector of the same dimension, $\mathbf{x}$, the *generating vector*, in such a way that the following generating rule $\mathcal{R}_n(\mathbf{x} \to \mathbf{c})$ holds:

$$\mathcal{R}_n(\mathbf{x} \to \mathbf{c}) \equiv \left\{ \exists \mathbf{x} \in V_n(\mathbf{C}) \wedge \mathbf{x}^+\mathbf{x} = 1 \ \Rightarrow \ \mathbf{c} = \left\{ c_i = |x_i|^2 \right\} \to K_n(\mathbf{c}) \right\}. \tag{6}$$

Thus, EJR can be applied to transform the generating vector elements, $\{x_i\}$, and then, indirectly, the elements of the ASA coefficient vector are varied, while preserving the convex conditions $K_n(\mathbf{c})$ [1].

3.2. ASA structure in molecules

Once the previous considerations are known, then it is easy to think that, by performing a previous computational task, a set of fitted atomic DF, $\mathcal{A} = \{\rho_A\}$, having the same ASA form as in equation (5), can be gathered. The set $\mathcal{A}$, being strictly formed by convex linear combinations of PD functions, as a PD operator set can be considered. Thus $\mathcal{A}$ can be used as a generating set of new ASA-type DF. A molecular DF, ρ_M, for example, can be approximated by a linear combination of $\mathcal{A}$ elements, with a coefficient vector $\mathbf{w} = \{w_A\}$, fulfilling the convex conditions $K_n(\mathbf{w})$:

$$\rho_M(\mathbf{r}) = \sum_A w_A \rho_A(\mathbf{r} - \mathbf{r}_A) \wedge K_n(\mathbf{w}), \tag{7}$$

where $\{\mathbf{r}_A\}$ are to be considered as the molecular atomic co-ordinates, on which the atomic DF are placed. According to the properties of the PD operators, described in definition 2(e), the ρ_M is also a PD function. This allows to consider the possibility to fit the coefficient vector, $\mathbf{w}$, to the molecular DF in the same fashion as it has been described for the atomic case. This also means that some generating vector $\mathbf{u}$ can be

defined too, fulfilling a similar set of conditions as shown in equation (6), provided that $\mathbf{x} \leftarrow \mathbf{u}$ and $\mathbf{c} \leftarrow \mathbf{w}$, that is, defining a generating rule like $\mathcal{R}_n(\mathbf{u} \to \mathbf{w})$.

A *promolecular* approach to ρ_M may correspond to choosing the ASA-type coefficient vector $\mathbf{w}$ with all its components equal, that is, $\mathbf{w} = n^{-1}\mathbf{1}$ and $\mathbf{1} = (1, 1, \dots, 1)^{\mathrm{T}}$. Another possibility is to associate to each atomic density a coefficient $w_A = Z_A N^{-1}$, where Z_A is the atomic nuclear charge and $N = \sum_A Z_A$ is the total molecular number of electrons.

3.3. Considerations around MO theory

Another example of this first-order DF convex form may be found within the MO theory. Indeed, if an *orthonormalised* MO set, $\{\varphi_i\}$, is at our disposal, then it is well known that the expression for the DF may be obtained as

$$\rho_M = \sum_i \omega_i |\varphi_i|^2, \tag{8}$$

where the set $\omega = \{\omega_i\}$ can be constructed to fulfil the usual discrete convex conditions $K_n(\omega)$. In this case, the elements of vector ω are interpreted as the set of MO *occupation* numbers, which in the usual SCF monoconfigurational theory are just integers, 2 or 1. In the present framework, they must be trivially transformed in such a way as to fulfil the convex conditions. In other computational environments, as in MC SCF or in natural orbital algorithms, the MO occupation can be made of real numbers, so the convex conditions may become more naturally defined [14]. Equation (8) has the same structure as the previous ASA expressions. In closed shell systems, studied in a monoconfigurational way, the vector ω just corresponds to a simple promolecular approach with $\omega = n^{-1}\mathbf{1}$, with n being half the number of electrons.

The dependence of the MO basis set on an AO basis, within the well-known LCAO MO approach, is still something not exploited in the ASA approaches. This possibility is far from being a forgettable question, and thus will be studied in forthcoming work. Here it can be said that the LCAO form of MO could be related to the DF generating VS.

An interesting point to be noted at this moment is the possibility to transform SCF theory and electronic energy expressions into a normalised form, in order to use a first-order DF with the adequate convex conditions. It is a trivial matter to use equation (8), with the adequate convex conditions in monoconfigurational energy expressions, and the effect of this point of view in the Fock operator definition is easy to grasp.

Keeping convex conditions over the DF will scale the monoconfigurational electronic energies by a factor which corresponds to the inverse of the squared number of electrons: $\sigma^2 = N_{\mathrm{e}}^{-2}$.

This means that:

(a) Repulsion integrals do not need to be scaled at all.

(b) One electron Hamiltonian must be scaled in full by a factor equal to the inverse of the number of electrons: $\sigma = N_{\mathrm{e}}^{-1}$.

Taking these simple rules into account, one can see that a SCF computation under this convex scaling of the DF will translate the process of any atomic or molecular system into some computational algorithm, with a structure such that it will become independent of the number of system particles. Thus, in this scaled framework the SCF iterative structure could be studied from the aspects of convergence, extrapolation, etc., for all the systems on an almost equal footing. The only variable aspects of the SCF process will come from the basis set size and the nature, as well as the 3-dimensional space position, of the atomic centres.

An interesting aspect of this scaling corresponds to the size of electronic energies, which will become constrained in a shorter interval than before scaling. For example, the absolute values of atomic SCF electronic energies, from H to U, will be contained within the approximate interval $\{0.5, 4\}$. Re-sizing them to the usual values needs only a scaling by a factor N_e^2.

3.4. The continuous ASA case

Equation (7) can be made much more general if it can be studied as being embedded in a continuous environment. Provided that some PD DF basis set, $\{\rho(\mathbf{t}, \mathbf{r})\}$, is known beforehand, while considering that the DF basis elements could depend on some continuous parameter set $\mathbf{t}$, then, the continuous construction of a PD DF, $\rho_\omega(\mathbf{r})$, can be set up using the integral form

$$\rho_\omega(\mathbf{r}) = \int \omega(\mathbf{t})\rho(\mathbf{t}, \mathbf{r})\, d\mathbf{t} \wedge K_\infty\big(\omega(\mathbf{t})\big), \tag{9}$$

where $\omega(\mathbf{t})$ is a function of the parameter set $\mathbf{t}$, fulfilling the continuous convex conditions $K_\infty(\omega(\mathbf{t}))$:

$$K_\infty\big(\omega(\mathbf{t})\big) = \left\{ \forall \mathbf{t}\colon\ \omega(\mathbf{t}) \in \mathbf{R}^+ \wedge \int \omega(\mathbf{t})\, d\mathbf{t} = 1 \right\}. \tag{10}$$

At the same time, a generating function, $\gamma(\mathbf{t})$, contained in some Hilbert space, can be defined, which fulfils a generating rule equivalent to the one previously shown in equation (6) in the discrete case:

$$\mathcal{R}_\infty\big(\gamma(\mathbf{t}) \to \omega(\mathbf{t})\big)$$
$$\equiv \left\{ \exists \gamma(\mathbf{t}) \in V_\infty(\mathbf{C}) \wedge \int \gamma(\mathbf{t})^*\gamma(\mathbf{t})\, d\mathbf{t} = 1\ \Rightarrow\ \omega(\mathbf{t}) = \big|\gamma(\mathbf{t})\big|^2 \to K_\infty\big(\omega(\mathbf{t})\big) \right\}. \tag{11}$$

In this way, it can be seen how a discrete ASA framework can be extended into a ∞-dimensional environment without effort.

Equation (11) has interesting features, which provide the generating function, $\gamma(\mathbf{t})$, within a nearby kinship to a QO wavefunction. Also the coefficient function, $\omega(\mathbf{t})$, closely resembles a DF itself, through the corresponding equivalent definition.

On the other hand, equation (9) also can be looked at as an integral transform [11,16] of the coefficient function $\omega(\mathbf{t})$, with the DF $\rho(\mathbf{t}, \mathbf{r})$ acting as a transform

kernel. Thus, somehow, the continuous DF linear combination (9) can also be considered as a scalar product between two functions having an equivalent DF nature. A convolution transformation may also be considered as a particular, but nonetheless, appealing situation, instead of a typical integral transform. However, this possibility can bring us far away from the objectives of the present paper and will be considered elsewhere.

4. The nature of discrete QO representations

Let us now suppose a QOS, $\mathbf{Q}$, constructed in the usual way as a TS, that is, $\mathbf{Q} = \mathbf{S} \times \mathbf{P}$. Let us also suppose that the elements of the tag set part are ASA-type DF, built as shown in equation (5) or (7). Accepting this scenario is the same as considering that a QO is described under some finite PD functional basis set, $\Phi = \{|\varphi_i(\mathbf{r})|^2\}$, with co-ordinates $\omega = \{\omega_i\}$, fulfilling the convex conditions $K_n(\omega)$ and belonging to a given n-dimensional VSS such as $\omega \in \mathbf{W}_n(\mathbf{R}^+)$.

The TS constructed as $\mathbf{Q}_n = \mathbf{S} \times \{\mathbf{O} \subset \mathbf{W}_n(\mathbf{R}^+)\}$ corresponds to a QOS which has as the tag set part a subset of some VSS of finite dimensions. A discrete representation of QO can be, thus, reached in this way, besides the one discussed in section 2(f).

Just looking at the nature of the continuous ASA-type transform (9), it can be seen that in the discrete case, the PD basis set Φ and the coefficient vector ω will bear some equivalent structure. As the convex conditions $K_n(\omega)$ hold for the coefficient vector, it is easy to interpret this feature in such a way that the elements of the coefficient vector ω constitute a *discrete probability distribution*. For example, in a promolecular approach as well as in a MO monoconfigurational structure, ω can appear as a *homogeneous* discrete probability distribution.

Thus, the coefficient vector ω bears the equivalent statistical features of a DF in discrete n-dimensional spaces. It is not a strange fact that there exists a generating vector, $\gamma \in \mathcal{V}(\mathbf{C})$, producing the PD ω elements by application of the generating rule $\mathcal{R}_n(\gamma \to \omega)$. The structure of the rule is not the one that can be attached to a linear transformation, but has to bear a non-linear form.

From an algebraic point of view, this non-linear nature of the relationship between the generating vector and the coefficient vector appears non-natural. This fact is still more obvious when put in front of the continuous situation, as discussed previously. The image appears even more conspicuous when observing the nature of the DF from the quantum mechanical aspect, because any DF has to be considered as a squared module of the QO wavefunction, acting thus as a generating vector.

The problem can be stated transparently using a simple, well-known mathematical device, associated to the most basic aspects of quantum mechanics. Suppose a QO wavefunction is known for some system state $\Psi(\mathbf{r}) \in \mathcal{H}(C)$. The corresponding DF is simply computed as $\rho(\mathbf{r}) = |\Psi(\mathbf{r})|^2$. The interesting fact is that the DF thus defined may belong either to the Hilbert space direct product $\mathcal{H}(C) \otimes \mathcal{H}(C)$, when considered as an operator, or to some functional VSS, $\mathcal{H}(\mathbf{R}^+)$, when considered as a PD real-valued function. In this sense, the generating rule can be applied here, and immediately

written as $\mathcal{R}_\infty(\Psi \rightarrow \rho)$. However, it can be interpreted in the following way, using the practically unmodified structure of equation (11):

$$\mathcal{R}_\infty(\Psi \rightarrow \rho)$$
$$\equiv \left\{ \exists \Psi(\mathbf{r}) \in \mathcal{H}(C) \wedge \int |\Psi(\mathbf{r})|^2 \, d\mathbf{r} = 1 \; \Rightarrow \; \rho(\mathbf{r}) = |\Psi(\mathbf{r})|^2 \rightarrow K_\infty(\rho) \right\}. \quad (12)$$

This continuous generating rule *must imply* a closer relationship between the vectors involved in the discrete case. In this case, the generating rule $\mathcal{R}_n(\gamma \rightarrow \omega)$ means that, while the normalisation part for γ possesses a simple algorithm to be computed, $1 = \gamma^+ \gamma$, it appears that there is no such simple operation in the second part of the generating rule. That is, when one must attach the ω coefficient vector elements to the squared modules of the generating vector γ, the algorithm is not naturally isomorphic to the one in equation (12).

The problem is as follows: there is a lack of simple, naturally obtained, isomorphic operation in the second part of the discrete generating rule, as stated in equation (6), when compared with the continuous case as defined in equation (11) or (12). A possible solution of this weird situation will be discussed next.

5. The structure of the generating n-dimensional VS: Diagonal Vector Spaces (DVS)

The generating rules (6) and (12) are a shorthand notation of some non-linear transformation involving the generating VS, $\mathcal{V}_n(\mathbf{C})$, and the final VSS containing the coefficient vectors, $\mathcal{W}_n(\mathbf{R}^+)$. The lack of a simple natural operation, producing the results, implicitly stated within the generating rule in the discrete case, can be circumvented using the following scheme.

Let us suppose such an isomorphic pair of n-dimensional VS, which will be named $\mathcal{G}_n(\mathbf{C})$ and $\mathcal{F}_n(\mathbf{R}^+)$. Both can be good substitutes of the original $\mathcal{V}_n(\mathbf{C})$ and $\mathcal{W}_n(\mathbf{R}^+)$ VS described above, respectively. A sound isomorphism of column or row VS is constituted by Diagonal Vector Spaces (DVS), whose elements possess the structure of diagonal matrices. Let us consider that the isomorphic $\mathcal{G}_n(\mathbf{C})$ and $\mathcal{F}_n(\mathbf{R}^+)$ VS are made of diagonal matrices. The choice has not been arbitrary, because matrix multiplication is closed in DVS, that is, matrix products of diagonal matrices yield new diagonal matrices. Moreover, diagonal matrix products are commutative. Considering only the diagonal part of the matrix elements, and discarding the off-diagonal elements, the DVS possess the same dimension as their isomorphic column–row vector counterparts. Then, it is easy to see that using this simple isomorphic device both the discrete and continuous generating rules acquire the same formal structure. Indeed, the discrete rule in equation (6) will be rewritten within any DVS framework as

$$\mathcal{R}_n(\mathbf{D} \rightarrow \Delta) \equiv \left\{ \exists \mathbf{D} \in \mathcal{G}_n(\mathbf{C}) \wedge \langle \mathbf{D}^+ \mathbf{D} \rangle = \sum_i |d_i|^2 = 1 \; \Rightarrow \right.$$

$$\Delta = \mathbf{D}^+\mathbf{D} = \mathrm{Diag}\big(|d_i|^2\big) \wedge \langle \Delta \rangle = 1 \rightarrow K_n(\Delta) \bigg\}. \qquad (13)$$

The symbol $\langle \mathbf{D} \rangle = \sum_i d_i$ is, in this case, equivalent to the trace of the corresponding diagonal matrix, but it can be used as a simple symbol to mean the sum of all matrix elements; it was so defined and used for the same purposes earlier [6]. The convex condition has to be slightly modified to take into account the new DVS element structure:

$$K_n(\Delta) = \left\{ \Delta = \mathrm{Diag}(\pi_i) \in \mathcal{F}_n\big(\mathbf{R}^+\big): \; \pi_i > 0, \; \forall i \wedge \langle \Delta \rangle = \sum_i \pi_i = 1 \right\}. \qquad (14)$$

Thus, working with DVS instead of conventional VS and VSS, the coefficients in the discrete DF description possess the same structural properties as the DF themselves.

The generating DVS elements, $\mathbf{D} \in \mathcal{G}_n(\mathbf{C})$, behave in the same manner as the QO wavefunctions do. And the resultant coefficient diagonal matrix, $\Delta \in \mathcal{F}_n(\mathbf{R}^+)$, satisfying the convex conditions $K_n(\Delta)$, can be written as a squared module of the former diagonal matrix. This can be done using a discrete form of the generating rule $\mathcal{R}_n(\mathbf{D} \rightarrow \Delta)$, similar to the wavefunction-DF generating rule $\mathcal{R}_\infty(\Psi \rightarrow \rho)$: $\Delta = \mathbf{D}^+\mathbf{D} = \mathbf{D}\mathbf{D}^+ = |\mathbf{D}|^2 = \mathrm{Diag}(|d_i|^2)$.

The DVS $\mathcal{G}_n(\mathbf{C})$ may be considered as normed spaces with one of the possible norms defined as the trace of the squared matrix module. As a consequence, the DVSS $\mathcal{F}_n(\mathbf{R}^+)$ elements are constructed in such a way that their trace is always normalisable, and thus easily made unit. A diagonal TS, $\mathcal{D}_n$, can be derived in the usual way by using a given background set part, $\mathbf{S}$, and a DVSS convex subset, $\mathcal{K}$, as the tag set part, that is, $\mathcal{D}_n = \mathbf{S} \times \{\mathcal{K} \subset \mathcal{F}_n(\mathbf{R}^+)\}$.

The question, now, may *not* need to be why the n-dimensional DVS fulfil in a natural way the same conditions as the ∞-dimensional functional VS. But it could be much better stated as follows: which kind of consequences, if any, will this situation have in the development of a discrete quantum chemistry framework? In the next section, we will try to describe some of the possible features of this DVS structure.

6.   Expression of the density functions and other problems

It has been shown that the best discrete representation of the DF, having an ASA-like form, as in equations (5), (7) and (8), is best described as a diagonal matrix, instead of a vector as it is usual. Then, the *scalar-like* expression of the ASA DF type must be redefined in terms of the natural operations presented in the discussion of the preceding section.

In order to obtain a coherent image of all the possible redefinitions, which can be found as a consequence of the adoption of the DVS representation, let us make some preliminary considerations.

As the ASA generating rules, formal structure is better represented from the point of view of diagonal matrices instead of column–row vectors, both the generating and

coefficient vectors are thus transformed into elements of some DVS. The ASA forms discussed in section 3, besides the coefficient vector, are associated to a PD function set, which is in turn connected to the squared module of another function set, further belonging to another structure which can be named as a generating functional VS. This situation can be managed in the same way as in the preceding discussion.

Suppose now that a function basis set is known: $\Phi = \{\varphi_i\}$. Nothing opposes to the situation in which the set Φ can always, without loss of generality, be arranged into a diagonal matrix structure, and considered to be constructed as $\Phi \equiv \mathrm{Diag}(\varphi_1, \varphi_2, \ldots, \varphi_n) \in \mathcal{F}(\mathbf{C})$.

Then, it is obvious that when the following diagonal matrix product is made: $P = \Phi^*\Phi = \mathrm{Diag}(|\varphi_1|^2, |\varphi_2|^2, \ldots, |\varphi_n|^2) \in \mathcal{P}(\mathbf{R}^+)$, it will always produce a new diagonal matrix, whose elements belong to a special functional VSS made of PD functions, that is, made of function squared modules. Thus, taking into account the definition of the diagonal product of the initial basis set, one can consider that the above result produces an entirely new PD basis set: $P = \{|\varphi_i|^2\}$.

Also, having defined the generating and coefficient VS, one can construct the following hybrid diagonal matrix:

$$\forall D = \mathrm{Diag}(d_i) \in \mathcal{G}(\mathbf{C}) \wedge \forall \Phi = \mathrm{Diag}(\varphi_i) \in \mathcal{F}(\mathbf{C}) \quad \Rightarrow$$
$$\Psi = D\Phi = \mathrm{Diag}(d_i\varphi_i) \in \mathcal{I}(\mathbf{C}) \subseteq \mathcal{G}(\mathbf{C}) \times \mathcal{F}(\mathbf{C}). \tag{15}$$

Once this kind of mixed structures is constructed, then the ASA-like DF can be simply built by computing traces of squared modules of the diagonal structures, Ψ, as defined in equation (15) above. That is,

$$\rho = \langle \Psi^*\Psi \rangle = \sum_i |d_i\varphi_i|^2 = \sum_i |d_i|^2 |\varphi_i|^2 = \sum_i \omega_i \rho_i. \tag{16}$$

The formalism appears now clear as to how to construct the necessary generating elements and the way is open to obtain, in a very natural way, the structure of ASA-like DF. The most interesting thing of the whole procedure, perhaps, will consist in finding out how closely the deducible formal rules, based on discrete DVS, are equivalent to the formalism based on continuous quantum mechanics. But it seems that nothing is opposed to this possibility.

In fact, it only remains to express the formal problem of how an expectation value $\langle \Omega \rangle$ of some observable, associated to an operator Ω, can be computed within a DVS formalism. A possible way could be

$$\langle \Omega \rangle = \int \Omega \rho \, dV = \int \Omega \langle \Psi^*\Psi \rangle \, dV = \sum_i |d_i|^2 \int \Omega |\varphi_i|^2$$

$$= \sum_i \omega_i \int \Omega \rho_i \, dV = \sum_i \omega_i \int \varphi_i^* \Omega \varphi_i \, dV \equiv \int \langle \Psi^*\Omega\Psi \rangle \, dV. \tag{17}$$

The last linear combination of integrals is suited to differential operators, and can be naturally obtained when considering the operator Ω as a scalar matrix $\Omega\mathbf{I}$.

7. Conclusions

The quantum chemical, statistically coherent, significance of the expansion coefficients, satisfying convex conditions, in ASA-like DF forms, which can be considered as a discrete probability distribution has been shown. Moreover, the fitted atomic densities apparently do not have any problem when used in order to obtain expectation values of quantum chemical operators.

The formalism, based on DVS and TS, becomes in this manner a fruitful tool, which can be the fundament for further work.

Acknowledgements

References

[1] Ll. Amat and R. Carbó-Dorca, J. Comput. Chem. 18 (1997) 2023–2039.

[2] S.K. Berberian, *Introducción al Espacio de Hilbert* (Editorial Teide, Barcelona, 1970).

[3] R. Carbó-Dorca, Fuzzy sets and Boolean tagged sets, Technical report IT-IQC-5-97; J. Math. Chem. 22 (1997) 143–147.

[4] R. Carbó-Dorca, Tagged sets, convex sets and quantum similarity measures, Technical report IT-IQC-9-97; J. Math. Chem., this issue.

[5] R. Carbó, M. Arnau and L. Leyda, Int. J. Quantum Chem. 17 (1980) 1185–1189.

[6] R. Carbó and E. Besalú, J. Math. Chem. 18 (1995) 37–72.

[7] R. Carbó and E. Besalú, in: *Molecular Similarity and Reactivity: From Quantum Chemical to Phenomenological Approaches*, ed. R. Carbó (Kluwer, Amsterdam, 1995) pp. 3–30.

[8] R. Carbó and B. Calabuig, Int. J. Quantum Chem. 42 (1992) 1681–1693.

[9] P. Constans, L. Amat, X. Fradera and R. Carbó-Dorca, in: *Advances in Molecular Similarity*, Vol. 1, eds. R. Carbó-Dorca and P.G. Mezey (JAI Press, Greenwich, 1996) pp. 187–211.

[10] P. Constans and R. Carbó, J. Chem. Inf. Comput. Sci. 35 (1995) 1046–1053.

[11] *Enciclopaedia of Mathematics*, Vol. 5 (Reidel–Kluwer, Dordrecht, 1987) p. 126.

[12] *Enciclopaedia of Mathematics*, Vol. 8 (Reidel–Kluwer, Dordrecht, 1987) p. 249.

[13] C.G.J. Jacobi, J. Reine Angew. Math. 30 (1846) 51–94.

[14] E.V. Ludeña, in: *Química Teórica*, ed. S. Fraga (CSIC, Madrid, 1987) pp. 117–184.

[15] J. von Neuman, *Mathematical Foundations of Quantum Mechanics* (Princeton University Press, Princeton, NJ, 1955).

[16] A.H. Zemanian, *Generalized Integral Transformations* (Dover Publications, New York, 1987).