J Math Chem (2010) 47:331–344
DOI 10.1007/s10910-009-9572-y

ORIGINAL PAPER

Definition of norm coherent generalized scalar products and quantum similarity

Ramon Carbó-Dorca

Received: 18 May 2009 / Accepted: 1 July 2009 / Published online: 18 July 2009

Abstract Complete matrix sums and inward products are employed within vector spaces to define a generalized scalar product, searching at the same time for coherent definitions with the related general order norms. The theoretical background developed in this way permits to connect such mathematical constructs with quantum similarity and QSPR.

Keywords Generalized scalar products · Generalized norms · Inward matrixproducts · Vector semispaces · Quantum similarity · Quantum QSPR

1 Introduction

Scalar products, also called inner or dot products, are well known in Hilbert and pre-Hilbert spaces, see for example [1]. It will appear redundant to insist about their definition and properties, for more information see, for instance, Ref. [18,20].

In addition, within the so-called vector semispaces, which are always defined over the positive semi-definite real field, it has been already done some research on generalized scalar products, see Ref. [7], because the peculiarities of this kind of vector collections allow an easy development of such kind of mathematical structures.

In the present work it is intended to do some more exploration of specific alternative definitions of the scalar product and its generalization as well. While doing so, it is worthwhile to analyze the connection of such new structures with generalized Minkowski norms as it has been previously done in Ref. [9].

Higher order scalar products are connected with the theoretical background of molecular quantum similarity [6,4] and its application to quantum QSPR (QQSPR)

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

[10]. Recent research has shown that classical QSPR discrete molecular description can be employed within QQSPR development [11,14,15], therefore, the study of generalized scalar products may result useful within higher order QQSPR procedures.

In order to attain such goals, this study will start with some introductory remarks about the mathematical treatment of vector semispaces and probability distributions. Then, such theoretical manipulations will lead towards the possibility to transform semispace scalar products into a modified form, which could be easily connected with Euclidean norms. Generalization of this Euclidean framework into higher order norms and the attached modified scalar products will follow. Some examples both analytical and numerical will be discussed, closing the present work.

2 Vector semispaces and probability distributions

2.1 Vector semispaces

Due to the close relationship of vector semispaces with quantum molecular similarity theoretical background [4–6,10,13], such mathematical structures have been previously defined and discussed along several papers [3,7,9].

For completion sake, a brief summary of a vector semispace definition follows. Suppose some N-dimensional vector space[1] defined over the real field:[2] $V_N(\mathbf{R})$; any subset of vectors $V_N(\mathbf{R}^+) \subset V_N(\mathbf{R})$, defined over the positive or non negative real field is called a vector semispace [3,7,9] whenever it presents the usual properties of a vector space, shortly:

$$\forall \alpha,\ \beta \in \mathbf{R}^+ \wedge |a\rangle, |b\rangle \in V_N(\mathbf{R}^+) : \alpha|a\rangle + \beta|b\rangle \in V_N\left(\mathbf{R}^+\right). \tag{1}$$

Vector semispaces can be constructed from vector space elements and such construction symbolized by so-called generating symbols [7], involving the definitions of inward vector product and powers [9], that is:

$$\forall |z\rangle \in V_N(\mathbf{R}) \rightarrow \mathrm{G}(|z\rangle) = |z\rangle * |z\rangle = |z\rangle^{*[2]} \in V_N(\mathbf{R}^+). \tag{2}$$

Complete vector sums [9] over the elements of vector semispaces produce the first order Minkowski norm, see for example Ref. [19], of a given vector:

$$\forall |z\rangle \in V_N\left(\mathbf{R}^+\right) : \langle |z\rangle\rangle \equiv \| |z\rangle \|_1 \in \mathbf{R}^+ \tag{3}$$

Thus, within finite dimensional spaces a complete vector sum corresponds to the sum of all the elements of a given vector, but when functions are involved such an

[1] The reader can consider that the discrete dimensional spaces employed here are vector or matrix spaces and the infinite dimensional ones function spaces, unless the contrary is stated.

[2] The same will be valid in case the complex field is chosen as the vector space background field. The real field has been chosen in order to simplify the present development. In the practical computational cases the real field is the usual background.

operation transforms into an integral over the variable domain definition of the functional space.

Vector semispaces can be ordered via shells [3,9]. A shell $S(\mu)$ of a vector semispace is a subset whose vectors possess the same first order Minkowski norm value: μ, that is:

$$\forall |s\rangle \in S(\mu) \subset V_N\left(\mathbf{R}^+\right) : \||s\rangle\|_1 = \langle |s\rangle\rangle = \mu \tag{4}$$

The unit shell $S(1)$ elements can originate every element of a given vector semispace, as the elements of any other semispace shell are just homothetic to the unit shell elements:

$$\forall |u\rangle \in S(1) \rightarrow \forall \mu \in \mathbf{R}^+ : |m\rangle = \mu |u\rangle = S(\mu). \tag{5}$$

If a known set of positive definite scalars: $\wedge = \{\lambda_I\} \in \mathbf{R}^+$ is convex, that means their sum yields the unit, that is: $\sum_I \lambda_I = 1$. This property can be symbolized by using the notation: $\mathrm{K}(\Lambda)$. Then, a linear combination of elements of any shell $S(\mu)$ by means of the set Λ can be called convex and belongs to the same shell:

$$\{\lambda_I\} \subset \mathbf{R}^+ \wedge \sum_I \lambda_I = 1 \rightarrow \{|m_I\rangle\} \subset S(\mu) : |m\rangle = \sum_I \lambda_I |m_I\rangle \in S(\mu) \tag{6}$$

2.2 Probability distributions and the unit shell

The unit shell contains probability distributions in both discrete and infinite dimensional vector semispaces. For instance, in function semispaces, the elements whose complete sum is the unit and therefore belonging to the unit shell can be considered probability distributions.

In order to illustrate the previous definitions and algebraic symbols, one can use first order quantum density functions: $\rho(\mathbf{r})$. One can write if this is the case:

$$\forall \rho(\mathbf{r}) \in S(N) \subset V_\infty\left(\mathbf{R}^+\right) \rightarrow \langle \rho \rangle = \int_D \rho(\mathbf{r}) \mathrm{d}\,\mathbf{r} = N, \tag{7}$$

being N the number of particles and D an appropriate integration domain. The so called shape function $\sigma(\mathbf{r})$ belongs to the unit shell [3] and therefore can be associated to a continuous probability distribution:

$$N^{-1}\rho(\mathbf{r}) = \sigma(\mathbf{r}) \in S(1) \subset V_\infty\left(\mathbf{R}^+\right) \rightarrow \int_D \sigma(\mathbf{r}) \mathrm{d}\,\mathbf{r} = \mathbf{1}. \tag{8}$$

Also, in column (or row) N-dimensional vector semispaces, one can associate every vector of the unit shell with a given discrete probability distribution in the same manner as in the infinite dimensional case, as:

$$\forall |a\rangle = \{a_I\} \in S(N) \subset V_N\left(\mathbf{R}^+\right) \rightarrow \langle |a\rangle\rangle = \sum_I a_I = N \tag{9}$$

then:

$$N^{-1}|a\rangle = |z\rangle \in S(1) \subset V\left(\mathbf{R}^+\right) \rightarrow \langle |z\rangle\rangle = \sum_I z_I = N^{-1}\sum_I a_I = 1 \tag{10}$$

Hence, the elements of the vector $|z\rangle$ constitute a convex scalar set, which can be associated to a set of discrete probabilities as: $\forall I : 1 \leq z_I \leq 0$.

3 Euclidean norms and scalar products in vector semispaces

3.1 Euclidean norms

As it has been explained, in vector semispaces the complete sum of any of its elements yields the first order Minkowski norm of such a vector. However, for many purposes it is interesting to consider how to connect in vector semispaces such a norm with the usual Euclidean norms and scalar products, which can be described in Hilbert and pre-Hilbert semispaces. The problem has been considered in relationship with quantum similarity theory and practice [4,6,8,10,13] but recently similar procedures have been described as the necessary theoretical background to study the probability spaces associated to the study of vapor-liquid equilibrium [2].

In fact any member of any vector semispace can be associated to its square root in such a way that the Minkowski norm of the original vector transforms into a Euclidian norm. In function semispaces this is a trivial operation, as:

$$\forall \rho \in S(N) \subset V_\infty\left(\mathbf{R}^+\right) \rightarrow \exists p = \sqrt{\rho} \wedge \langle \rho\rangle = \langle p|p\rangle = \langle p^2\rangle = N. \tag{11}$$

In finite dimensional vector semispaces a similar description can be conceived, employing the inward square root definition of a vector [9]:

$$\forall |a\rangle = \{a_I\} \in S(N) \subset V_N\left(\mathbf{R}^+\right) \rightarrow \exists |p\rangle = |a\rangle^{*\left[\frac{1}{2}\right]} = \left\{p_I = \sqrt{a_I}\right\} \tag{12}$$

which can be associated to the norms:

$$\langle p|p\rangle = \langle |p\rangle * |p\rangle\rangle = \left\langle |p\rangle^{*[2]}\right\rangle = \sum_I p_I^2 = \sum_I a_I = \langle |a\rangle\rangle. \tag{13}$$

As a consequence, for any vector in a vector semispace one can describe an inward square root vector which possesses the Euclidian norm coincident with the Minkowski norm of the original vector, which for notation sake can be also written as:

$$\| |a\rangle \|_1 = \langle |a\rangle\rangle \tag{14}$$

while the square root of the scalar product of the vector by itself can be noted as:

$$\| |p\rangle \|_2 = \sqrt{\langle p|p\rangle}. \tag{15}$$

Now, it is the right moment to consider that there is some nomenclature confusion about the Euclidian norm as defined here and its square root, usually also referred as vector length. In order to denote the difference between both elements, which in our case will be coincident with the symbols described above. One can proceed as follows:

$$\| |p\rangle \|_2 = \sqrt{\langle p|p\rangle} = \sqrt{\langle |a\rangle \rangle} = \sqrt{\| |a\rangle \|_1}, \tag{16}$$

However, by context, in the present study there cannot appear any doubt about which operation is associated to the word *norm*.

3.2 Scalar products and square root scalar products

Although in the definition of the previous norms there is no confusion except that it may be named norm both a scalar product and its square root, it appears to be some interesting point to discuss, concerning the definition of the scalar product in a way that there is some positive definition coherence with the norm itself, not only semantic but also of mathematical nature.

In order to visualize the procedure permitting the construction of such coherent scalar product definition, suppose one obtains two inward square root vectors from some vector pair belonging to a vector semispace:

$$|a\rangle, |b\rangle \in V_N\left(\mathbf{R}^+\right) \rightarrow |p\rangle = |a\rangle^{*\left[\frac{1}{2}\right]}, |q\rangle = |b\rangle^{*\left[\frac{1}{2}\right]}. \tag{17}$$

Then, from these new vectors one can construct the scalar product of the inward square root vectors:

$$\langle p|q\rangle = \sum_I p_I q_I = \sum_I \sqrt{a_I b_I} \tag{18}$$

So, it is obvious that when: $|a\rangle = |b\rangle \rightarrow |p\rangle = |q\rangle$, so the equation above leads to:

$$\left(\| |p\rangle \|_2\right)^2 = \langle p|p\rangle = \langle |a\rangle \rangle = \| |a\rangle \|_1. \tag{19}$$

However, Eq. (19) does not reproduce Eq. (16), which involves the square root of the Euclidean norms.

In order to arrive to this previous computational situation, there is needed to define some scalar product defined anew as a square root, that is:

$$(p|q) = \sqrt{\langle p|q\rangle} = \sqrt{\sum_I p_I q_I} \tag{20}$$

in this way, when both involved vectors are coincident, one can write:

$$(p|q) = \sqrt{\sum_I p_I^2} = \||p\rangle\|_2. \tag{21}$$

In fact, any Euclidian metric has to possess the above property; otherwise the triangular inequality is not fulfilled [1,18]. Therefore, it seems that the scalar product definition in order to be coherent with a Euclidian metric is the one given in Eq. (16), that is: the square root of the usual scalar product in pre-Hilbert and Hilbert spaces.

When dealing with infinite dimensional vector semispaces, like the ones shown in Eq. (7), one can write the following equations, which follow the same previous definitions and considerations developed in finite dimensional spaces:

$$\forall \rho_A(\mathbf{r}),\ \rho_B(\mathbf{r}) \in V_\infty\left(\mathbf{R}^+\right) \rightarrow (\rho_A|\rho_B) = \left(\int_D \rho_A(\mathbf{r})\rho_B(\mathbf{r})\mathrm{d}\,\mathbf{r}\right)^{\frac{1}{2}} = \sqrt{\langle \rho_A(\mathbf{r})\rho_B(\mathbf{r})\rangle} = \sqrt{\langle \rho_A|\rho_B\rangle} \tag{22}$$

In case the pair of elements involved into the scalar product (22) is a couple of quantum density functions, as a consequence the scalar product: $\langle \rho_A|\rho_B\rangle \equiv \langle \rho_A\rho_B\rangle$ is the well known overlap quantum similarity measure integral [4,6]. Certainly, when considering quantum self-similarity measures, that is: the scalar product involving a unique density function, like $\langle \rho_A|\rho_A\rangle \equiv \langle \rho_A^2\rangle$; then one will have:

$$\forall \rho_A(\mathbf{r}) \in V_\infty\left(\mathbf{R}^+\right) \rightarrow (\rho_A|\rho_A) = \left(\int_D \rho_A(\mathbf{r})\rho_A(\mathbf{r})d\mathbf{r}\right)^{\frac{1}{2}} = \sqrt{\langle \rho_A|\rho_A\rangle} = \sqrt{\langle \rho_A^2\rangle} = \|p_A\|_2 \tag{23}$$

Such a general definition, which is coherent with the Euclidian norm of a scalar product, involving semispaces of any dimension, can be called *square root scalar product*.

A remark is worthwhile here. In the infinite dimensional case, the square root of a given continuous probability distribution is, like when one considers the quantum shape functions case, quite difficult to obtain analytically. In some cases, where just exponential functions are used as probability distributions, see for example Ref. [2], the possibility to obtain the functions: $p(\mathbf{r}) = \sqrt{\rho(\mathbf{r})}$ is obviously trivial.

This kind of manipulations shall be relevant in the background of the development of the application of molecular similarity to QSPR, see for example [10,11,13–15].

4 Higher order scalar products and norms

4.1 Finite dimensional case

The scalar products can be written using the inward vector product as defined in Eq. (2), that is: Eq. (20) can be expressed as:

$$(p|q) = \sqrt{\langle p|q\rangle} = \sqrt{\langle |p\rangle * |q\rangle\rangle}. \tag{24}$$

Following this notation, it will be quite easy to define a third order scalar product involving three vectors as follows:

$$\begin{aligned}(p|q|s) &= \sqrt[3]{\langle |p\rangle * |q\rangle * |s\rangle\rangle} = \sqrt[3]{\sum_I p_I q_I s_I} \to (p|p|p)\\ &= \sqrt[3]{\langle |p\rangle * |p\rangle * |p\rangle\rangle} = \sqrt[3]{\sum_I p_I^3}\end{aligned} \tag{25}$$

although such a definition does not provide a third order norm, but a pseudonorm [21], as some scalar products involving the same vector can be negative or null when belonging to a real vector space:

$$\exists |p\rangle \in V_N(\mathbf{R}) : (p|p|p) = 0 \wedge |p\rangle \neq |0\rangle.$$

A way to avoid this drawback can consist into using the following redefinition, when: $|p\rangle = |q\rangle = |s\rangle$:

$$(p|p|p) = \sqrt[3]{\langle |p\rangle * |p\rangle * ||p\rangle|\rangle} = \sqrt[3]{\sum_I p_I^2 |p_I|} = \sqrt[3]{\sum_I |p_I|^3} = \||p\rangle\|_3 \tag{26}$$

with the inward absolute value of a given vector given by:

$$* ||p\rangle| = \{|p_I|\}. \tag{27}$$

When the vectors belong to a vector semispace such redefinition becomes irrelevant and the third order norm can be written as in Eq. (25):

$$\forall |p\rangle \in V_N\left(\mathbf{R}^+\right) : \||p\rangle\|_3 = (p|p|p) = \sqrt[3]{|p\rangle^{*[3]}} = \sqrt[3]{\sum_I p_I^3}. \tag{28}$$

These results can be used to construct within vector semispaces higher order products, which can be made coherent with higher order norms:

$$\forall\{|p_\alpha\rangle|\alpha = 1, v\} \subset V_N\left(\mathbf{R}^+\right) : (|p_1\rangle||p_2\rangle|\cdots||p_v\rangle)$$
$$= \sqrt[v]{\langle|p_1\rangle*|p_2\rangle\cdots*|p_v\rangle\rangle} = \sqrt[v]{\sum_I\left(\prod_{\alpha=1}^{v} p_{I\alpha}\right)} \tag{29}$$

in such a way that the higher order norms can be written as:

$$\||p\rangle\|_v = \sqrt[v]{\langle|p\rangle^{*[v]}\rangle} = \sqrt[v]{\sum_I p_I^v}. \tag{30}$$

Such definitions can be employed in vector spaces in general if care is taken of odd higher order scalar products whenever they are to be coherent with the norm of the same odd order. In these cases, a similar redefinition as in the third order case shown in Eq. (26), has to be done in order to avoid the definition of a pseudonorm. That is: in real vector spaces, Eq. (30) is a norm whenever $v = 2k$, but shall be modified when dealing with odd order norms, where $v = 2k + 1$, using:

$$v = 2k+1 : \||p\rangle\|_v = \sqrt[v]{\langle|p\rangle^{*[2k]}*||p\rangle|\rangle} = \sqrt[v]{\sum_{I=1}^{v}\left(P_I^{2k}|p_I|\right)}.$$

Another remark, which shall be certainly now disclosed in order to warn the possible users of the previously discussed generalized norms and scalar products, is about the rotational invariance of the whole set of both generalized mathematical structures.

In the second order Euclidean case, rotational invariance is always fulfilled for both scalar products and norms, as it is well known and moreover this constitutes the basis of the usual geometry, see for example Ref. [18]. However, as defined here, the overall order structures which are different from the second order one are *not* rotationally invariant. This is an issue not very much commented, excepting in some exceptional cases, see for instance Sneath and Sokal [19], who describe quite clearly the problem within a discussion on generalized norms.

4.2 Odd generalized scalar products over vector spaces

The norm coherent description of generalized scalar products within vector spaces can be particularly described in third order as any of the three possible products involving three vectors: $\{|a\rangle, |b\rangle, |c\rangle\} \in V_N(\mathbf{R})$:

$$\langle||a\rangle|*|b\rangle*|c\rangle\rangle \vee \langle|a\rangle*||b\rangle|*|c\rangle\rangle \vee \langle|a\rangle*|b\rangle*||c\rangle|\rangle$$

thus, as no vector can be taken in principle over the other ones, the third order scalar products can be defined as an average:

$$\langle a|b|c\rangle = 3^{-1}\left(\langle ||a\rangle|*|b\rangle*|c\rangle\rangle + \langle |a\rangle*||b\rangle|*|c\rangle\rangle + \langle |a\rangle*|b\rangle*||c\rangle|\rangle\right)$$

which can be easily generalized to any order and to coherent root scalar products. Such an algorithm is irrelevant for even order generalized scalar products and when the scalar products are operated with elements belonging to vector semispaces.

4.3 Infinite dimensional case

In the infinite dimensional vector semispaces the coherent definition of higher norm scalar products can be also easily performed. For example, in the three order case and if the vectors are quantum density functions, a triple density [10,13,5] quantum similarity integral appears:

$$\langle \rho_A \rho_B \rho_C \rangle = \int_D \rho_A(\mathbf{r})\rho_B(\mathbf{r})\rho_C(\mathbf{r})\mathrm{d}\,\mathbf{r}. \tag{31}$$

Then, the norm coherent scalar product can be defined as the cubic root, which will easily yield the appropriate third order norm:

$$(\rho_A|\rho_B|\rho_C) = \sqrt[3]{\langle \rho_A \rho_B \rho_C \rangle} \rightarrow (\rho_A|\rho_A|\rho_A) = \sqrt[3]{\langle \rho_A \rho_A \rho_A \rangle} = \|\rho_A\|_3. \tag{32}$$

In the case of generalized products and norms of density functions there appear no direct rotational problems, except these originating in the changes of the functions themselves due to the rotations. When molecular density functions are studied under these mathematical structures and the framework of Born-Oppenheimer approximation is chosen, the real problem in this circumstance can be associated to the superposition of the atomic molecular frozen structures of the involved molecules, although in generalized norms there this superposition problem does not appear whenever the involved functions have the same set of nuclear coordinates. Other problems, related with the previously mentioned one appear, but the manner to overcome them has been recently discussed in Ref. [12].

5 Some analytical examples concerning Gaussian functions

It is worthwhile to present now some practical simple examples in order to illustrate the feasibility of the generalized scalar products and norms. Gaussian functions will be used for this purpose.

To start with these examples, suppose known three Gaussian functions of the same variable with different exponents and centered at three different points, such as:

$$g_a(x) = \exp\left(-\alpha(x-a)^2\right); g_b(x) = \exp\left(-\beta(x-b)^2\right);$$
$$g_c(x) = \exp\left(-\gamma(x-c)^2\right)$$

their triple scalar product can be defined, see for example [5], as the involved integral is well known [16]:

$$\begin{aligned}\langle g_a(x)g_b(x)g_c(x)\rangle &= \int_{-\infty}^{+\infty} \exp\left(-\left[\alpha(x-a)^2+\beta(x-b)^2+\gamma(x-c)^2\right]\right)\mathrm{d}\,x \\ &= \int_{-\infty}^{+\infty} \exp\Big(-\Big[(\alpha+\beta+\gamma)x^2 \\ &\quad -2(\alpha a+\beta b+\gamma c)x+(\alpha a^2+\beta b^2+\gamma c^2)\Big]\Big)\,\mathrm{d}\,x \\ &= \int_{-\infty}^{+\infty} \exp\left(-\left[Ax^2-2Bx^2+C\right]\right)\mathrm{d}\,x = \sqrt{\frac{\pi}{A}}\exp\left(\frac{B^2-AC}{A}\right)\end{aligned} \tag{33}$$

So, even within infinite dimensional vector spaces, triple or multiple scalar products can be well defined. In this case one can without problem obtain the cubic root of the result of the integral (33) in order to have the norm coherent scalar product structure.

In order to obtain a general framework for Gaussian functions, then if instead of a three Gaussian product one defines another with an indefinite number of functions, N say, of them, there is just need to construct the N-dimensional vectors bearing the exponents and the function variable origin shifts respectively:

$$|\alpha\rangle = \{\alpha_I | I = 1, N\} \wedge |a\rangle = \{a_I | I = 1, N\}$$

These vectors define the Gaussian function set:

$$\forall I : g_I(x) = \exp\left(-\alpha_I(x-a_I)^2\right),$$

so one can construct the N function product:

$$G(x) = \prod_{I=1}^{N} g_I(x) = \exp\left(-\left[Ax^2-2Bx+C\right]\right), \tag{34}$$

constituting a simple structure which contains the formerly discussed triple Gaussian product integral as a particular case.

Now and before the three constants appearing in Eq. (34) can be defined employing the following inward vector products and the complete sum symbols:

$$A = \langle|\alpha\rangle\rangle;\ B = \langle\alpha|a\rangle = \langle|\alpha\rangle * |a\rangle\rangle;\ C = \langle|\alpha\rangle * |a\rangle * |a\rangle\rangle = \left\langle|\alpha\rangle * |a\rangle^{*[2]}\right\rangle,$$

where it is interesting to see how the third term in the exponent corresponds to a triple scalar product, the second to a classical scalar product and the first to a Minkowski norm as the Gaussian exponents are taken positive definite as usual.

Therefore, the result of a multiple scalar product can be obtained with the previous integral value, appearing in the last equality of Eq. (33), submitted to the appropriate N-th root.

It is trivial but illustrative to discuss some particular instances of the final equation. For example, when a set of Gaussian functions with the same exponent but located to several different points are considered, then it will hold:

$$|\alpha\rangle = \alpha|1\rangle \wedge |1\rangle = \{1_I = 1 | I = 1, N\} \wedge |a\rangle = \{a_I | I = 1, N\},$$

where the unity vector: $|1\rangle = \{1_I = 1\}$ has been employed. Hence, in this case one will have:

$$\begin{aligned} A &= N\alpha \\ B &= \alpha\langle|1\rangle * |a\rangle\rangle = \alpha\langle|a\rangle\rangle = \alpha \sum_I a_I \rightarrow B^2 = \alpha^2 \langle|a\rangle\langle a|\rangle \\ C &= \alpha \left\langle|1\rangle * |a\rangle^{*[2]}\right\rangle = \alpha \left\langle|a\rangle^{*[2]}\right\rangle = \alpha \sum_I a_I^2 \end{aligned}$$

and from this result the value of the multiple scalar product integral can be written as:

$$\langle G\rangle = \sqrt{\frac{\pi}{\alpha N}} \exp\left(\alpha \frac{\langle|a\rangle\langle a|\rangle - N\left\langle|a\rangle^{*[2]}\right\rangle}{N}\right).$$

Another possible case to be now considered is a set of different Gaussian functions centered at the same site, and then one will have:

$$|\alpha\rangle = \{\alpha_I | I = 1, N\} \wedge |a\rangle = a|1\rangle \wedge |1\rangle = \{1_I = 1 | I = 1, N\},$$

which will produce the constants:

$$\begin{aligned} A &= \langle|\alpha\rangle\rangle \\ B &= a\langle|\alpha\rangle * |1\rangle\rangle = a\langle|\alpha\rangle\rangle = aA \rightarrow B^2 = (aA)^2 \\ C &= a^2 \left\langle|\alpha\rangle * |1\rangle^{*[2]}\right\rangle = a^2\langle|\alpha\rangle\rangle = a^2 A \end{aligned}$$

and thus the general complete sum scalar product becomes, except by simplicity sake for the lack of the corresponding order root:

$$\langle G\rangle = \sqrt{\frac{\pi}{A}}\exp\left((A-1)a^2\right).$$

Of course, when the common center of the Gaussian set is the origin, that is: $a = 0$, then one will obtain the well known result:

$$\langle G\rangle = \sqrt{\frac{\pi}{A}}.$$

The generalized N-th order norm of a Gaussian function centered somewhere, in a say, produces:

$$\begin{aligned} A &= \alpha N \\ B &= a\alpha\langle|1\rangle * |1\rangle\rangle = a\alpha\langle|1\rangle\rangle = a\alpha N \rightarrow B^2 = (a\alpha N)^2 \\ C &= a^2\alpha\left\langle|1\rangle * |1\rangle^{*[2]}\right\rangle = a^2\alpha\langle|1\rangle\rangle = a^2\alpha N \end{aligned}$$

which permits to write the generalized N-th order norm integral, prior to compute over it the N-th order root, as the expected complete sum result:

$$\langle G\rangle = \sqrt{\frac{\pi}{\alpha N}}\exp\left(\frac{a^2\alpha^2N^2 - a^2\alpha N(\alpha N)}{\alpha N}\right) = \sqrt{\frac{\pi}{\alpha N}} \rightarrow \| G \|_N = \left(\frac{\pi}{\alpha N}\right)^{\frac{1}{2N}}.$$

Thus, at least concerning Gaussian functions there is a well-defined generalized N-th order scalar product definition with coherently attached norms.

Such a possibility is quite interesting because of the widespread use of Gaussian basis sets in quantum chemistry, which in turn directs towards the way concerning which kind of integrals one could expect to face, when similarity measures of any order are envisaged to be computed.

6 Numerical tests involving Hilbert matrices

Having studied a continuous case, there will be also illustrative to see the behavior of generalized root products in discrete cases. The so called Hilbert matrix [17] constitutes a well known metric matrix, which can be associated to the powers of a variable and to the integrals connected to the scalar products:

$$\mathbf{H}^{(2)} = \{h_{IJ}\} \rightarrow h_{IJ} = \left\langle x^I|x^J\right\rangle = \left\langle x^{I+J}\right\rangle = \int_0^1 x^{I+J}\mathrm{d}x = \frac{1}{I+J+1},$$

besides the matrix properties like positive definiteness are depending of the machine precision. The eigenvalues of the Hilbert matrix of different dimensions are easy to compute and as their elements are scalar products, it is not so difficult to generalize this to higher order products too.

For instance, the N-th order Hilbert matrix can be computed as:

$$\mathbf{H}^{(N)} = \{h(\mathbf{i})|\mathbf{i} = (I_1, I_2, \ldots, I_N)\} \to h_{I_1 I_2 \cdots I_N} = \left\langle x^{\sum_{p=1}^{N} I_p} \right\rangle = \int_0^1 x^{\sum_{p=1}^{N} I_p} \mathrm{d}\, x = \left(1 + \sum_{p=1}^{N} I_p\right)^{-1}$$

the elements of it acting as generalized scalar products. Obtaining from here the root products is elementary.

It is also easy to build up a program, where the eigenvalues of higher order Hilbert square submatrices can be computed. It is well known too that Hilbert matrices are necessarily positive definite, but also they are quite numerically unstable, even using high order precision arithmetic. The higher order submatrices do not escape this ill-behaved structure and even within a quadruple precision Jacobi diagonalization routine, the lowest eigenvalues of not so big dimension Hilbert matrices transform into non-positive definite matrices and one or more quite small eigenvalues become negative. However, for these dimensions which produce positive definite Hilbert matrices, the lesser eigenvalues become numbers which lie almost nearby within the machine precision, so Hilbert matrix determinants become even less significant and from the practical computational point of view null or meaningless.

For example, using the fourth order Hilbert matrix with the elements submitted to the fourth root, all the (7×7) submatrices appear to have determinants in the range: $[10^{-36}; 10^{-49}]$ but none negative. However, raising the order one unit, then some of the (8×8) submatrices already possess some negative eigenvalues and the determinant range, without assuming the negative values, is: $[10^{-47}; 10^{-61}]$.

7 Conclusions

An overview of the scalar product concept generalization and the corresponding norm coherent definition has been presented. The mathematical framework will be useful in the theory and practice of quantum similarity and its application to quantum QSPR procedures.

Acknowledgments The present work has been partially sponsored by the Spanish Ministerio de Educación y Ciencia under the project: CTQ2006-04410/BQU.

References

1. S.K. Berberian, *Introduction to Hilbert Space* (Oxford University Press, New York, 1961)
2. D.C. Brody, D.W. Hook, J. Phys. A Math. Theor. **42**, 1–33 (2009)
3. P. Bultinck, R. Carbó-Dorca, J. Math. Chem. **36**, 191–200 (2004)

4. P. Bultinck, X. Gironés, R. Carbó-Dorca, Rev. Comput. Chem., vol. 21, eds. by K.B. Lipkowitz, R. Larter, T. Cundari (Wiley, Hoboken, 2005), pp. 127–207
5. R. Carbó, B. Calabuig, E. Besalú, A. Martínez, Mol. Eng. **2**, 43–64 (1992)
6. R. Carbó, L. Leyda, M. Arnau, Int. J. Quantum Chem. **17**, 1185–1189 (1980)
7. R. Carbó-Dorca (ed.), *Advances in Molecular Similarity,* vol. 2 (JAI Press, Greenwich, 1998), pp. 43–72
8. R. Carbó-Dorca, J. Mol. Struct. (Teochem) **537**, 41–54 (2001)
9. R. Carbó-Dorca, J. Math. Chem. **32**, 201–223 (2002)
10. R. Carbó-Dorca, J. Math. Chem. **36**, 241–260 (2004)
11. R. Carbó-Dorca, SAR QSAR Environ. Res. **18**, 265–284 (2007)
12. R. Carbó-Dorca, J. Math. Chem. **44**, 228–234 (2008)
13. R. Carbó-Dorca, A. Gallegos, *"Quantum Similarity and Quantum QSPR (QQSPR)" Enciclopaedia of Complexity and System Science*. (Springer, Berlin, 2009 in press)
14. R. Carbó-Dorca, S. Van Damme, Int. J. Quantum. Chem. **108**, 1721–1734 (2007)
15. R. Carbó-Dorca, A. Gallegos, A.J. Sánchez, J. Comput. Chem. **30**, 1146–1159 (2009)
16. W. Gröbner, N. Hofreiter, *Integraltafel (zweiter teil) Bestimmte Integrale* (Springer, Wien, 1966)
17. R.A. Horn, C.R. Johnson, *Matrix Analysis* (Cambridge University Press, Cambridge, 1985)
18. G.E. Shilov, *Linear Algebra* (Dover, New York, 1977)
19. P.H.A. Sneath, R.R. Sokal, *Numerical Taxonomy* (W.H. Freeman, San Francisco, 1973)
20. I.M. Vinogradov (ed.), *Encyclopaedia of Mathematics,* vol. 5 (Kluwer, Dordrecht, 1990)
21. I.M. Vinogradov (ed.), *Encyclopaedia of Mathematics,* vol. 7 (Kluwer, Dordrecht, 1990)