Journal of Mathematical Chemistry Vol. 34, Nos. 1–2, July 2003 (© 2003)

Analysis of a general theorem concerning two non-commuting Hermitian matrices: Quantum mechanical implications for ground and excited states

Ramon Carbó-Dorca [a,*] and Patrick Bultinck [b]

[a] *Institute of Computational Chemistry, University of Girona, Campus Montilivi, 17071 Girona, Catalonia, Spain*
E-mail: quantum@iqc.udg.es
[b] *Department of Inorganic and Physical Chemistry, Ghent University, Krijgslaan 281 (S-3), B-9000 Gent, Belgium*

Received 24 February 2003

A general theorem concerning the spectral relationship of two non-commuting Hermitian matrices is demonstrated. Discussion and analysis of such finding leads to consider its tight connection with respect of the Hohenberg–Kohn theorem (HKT), cornerstone of DFT theory. The present analysis shows that not only HKT can be considered a particular case of the proposed theorem, but also the validity of the studied spectral relationship can be extended from quantum mechanical ground state to excited states as well.

KEY WORDS: Hohenberg–Kohn theorem, Density Functional Theory

1. Theorem 1: matrix difference relationship for non-commuting Hermitian matrices

Hermitian matrices can be considered as representations of Hermitian operators with respect to some chosen basis set, therefore any theorem concerning such kind of matrix structures has an immediate quantum mechanical connotation and, hence, a plausible quantum chemical application. Here, more precisely, as a relationship concerning a pair of such matrices is studied, a decisive connection with density functional theory (DFT) [1] can be easily found, as will be further discussed.

The following theorem dealing with the lowest eigenvalues of two Hermitian non-commuting matrices can be proposed.

Theorem 1. Given two non-commuting Hermitian matrices $\mathbf{A}, \mathbf{B} \in M_{(N\times N)}(\mathbf{C})$ and known their lowest eigenvalues and different eigenvectors $\{\alpha_0; \mathbf{x}_0\}$ and $\{\beta_0; \mathbf{y}_0\}$. Then, the matrix difference $\mathbf{D} = \mathbf{A} - \mathbf{B}$ fulfils the inequality $\mathbf{x}_0^+ \mathbf{D} \mathbf{x}_0 < \mathbf{y}_0^+ \mathbf{D} \mathbf{y}_0$.

* Corresponding author.

0259-9791/03/0700-0075/0

Proof. One can write by hypothesis:

$$\mathbf{A}\mathbf{x}_0 = \alpha_0\mathbf{x}_0 \quad \text{and} \quad \mathbf{B}\mathbf{y}_0 = \beta_0\mathbf{y}_0 \quad \Longrightarrow \quad \alpha_0 = \mathbf{x}_0^+\mathbf{A}\mathbf{x}_0 \quad \text{and} \quad \beta_0 = \mathbf{y}_0^+\mathbf{B}\mathbf{y}_0, \tag{1}$$

as the eigenvectors can be considered normalized:

$$\mathbf{x}_0^+\mathbf{x}_0 = \mathbf{y}_0^+\mathbf{y}_0 = 1. \tag{2}$$

It is also fulfilled:

$$\alpha_0 = \min_I\{\alpha_I\} \quad \text{and} \quad \beta_0 = \min_I\{\beta_I\}. \tag{3}$$

Then, the variational principle indicates that, when setting up the following expressions, which can be taken as approximations to both exact eigenvalues, the present inequalities will also certainly hold:

$$\alpha_0^a = \mathbf{y}_0^+\mathbf{A}\mathbf{y}_0 > \alpha_0 \quad \text{and} \quad \beta_0^a = \mathbf{x}_0^+\mathbf{B}\mathbf{x}_0 > \beta_0. \tag{4}$$

The variational theorem also provides the possibility of having $\alpha_0^a = \alpha_0$. Such a situation is, however, not present taking into account the fact that the theorem explicitly requires there is no common eigenvector. As a result one can also derive:

$$\begin{aligned} \alpha_0^a - \alpha_0 &= \mathbf{y}_0^+\mathbf{A}\mathbf{y}_0 - \mathbf{x}_0^+\mathbf{A}\mathbf{x}_0 > 0, \\ \beta_0^a - \beta_0 &= \mathbf{x}_0^+\mathbf{B}\mathbf{x}_0 - \mathbf{y}_0^+\mathbf{B}\mathbf{y}_0 > 0. \end{aligned} \tag{5}$$

Summing up both above inequalities, it is finally obtained:

$$\left(\alpha_0^a + \beta_0^a\right) - (\alpha_0 + \beta_0) = \mathbf{x}_0^+(\mathbf{B} - \mathbf{A})\mathbf{x}_0 + \mathbf{y}_0^+(\mathbf{A} - \mathbf{B})\mathbf{y}_0 > 0 \tag{6}$$

which readily is transformed into

$$\mathbf{y}_0^+(\mathbf{A} - \mathbf{B})\mathbf{y}_0 - \mathbf{x}_0^+(\mathbf{A} - \mathbf{B})\mathbf{x}_0 > 0. \tag{7}$$

Thus, calling $\mathbf{D} = \mathbf{A} - \mathbf{B}$, one can finally write:

$$\mathbf{y}_0^+\mathbf{D}\mathbf{y}_0 - \mathbf{x}_0^+\mathbf{D}\mathbf{x}_0 > 0 \quad \Longrightarrow \quad \mathbf{x}_0^+\mathbf{D}\mathbf{x}_0 < \mathbf{y}_0^+\mathbf{D}\mathbf{y}_0. \tag{8}$$

□

1.1. Discussion of theorem 1 and its relation to DFT

Several remarks shall be put forward and discussed in reference to the above theorem 1 as follows.

Remark 1. In case of dealing with two matrices with common lowest eigenvectors or with commutative matrices, then one will have:

$$\mathbf{x}_0^+\mathbf{D}\mathbf{x}_0 = \mathbf{y}_0^+\mathbf{D}\mathbf{y}_0. \tag{9}$$

Remark 2. A sign reversal in the definition of the difference matrix will change the relationship accordingly. That is, whenever $\mathbf{D} = \mathbf{B} - \mathbf{A}$, then

$$\mathbf{x}_0^+ \mathbf{D} \mathbf{x}_0 > \mathbf{y}_0^+ \mathbf{D} \mathbf{y}_0. \tag{10}$$

Remark 3. The difference matrix, $\mathbf{D}$, being Hermitian too, can be diagonalized. That is, there will exist unitary matrix, $\mathbf{U}$, which will transform the difference matrix, $\mathbf{D}$, into a diagonal one, Δ, defined over the real field:

$$\exists \mathbf{U}: \quad \mathbf{U}^+ \mathbf{U} = \mathbf{U} \mathbf{U}^+ = \mathbf{I} \Longrightarrow \mathbf{U}^+ \mathbf{D} \mathbf{U} = \Delta \text{ and } \forall I: \ \Delta_I \in \mathbf{R}. \tag{11}$$

Then, the inequality associated to the theorem 1 can be expressed by means of the equivalent form

$$\mathbf{p}_0^+ \Delta \mathbf{p}_0 < \mathbf{q}_0^+ \Delta \mathbf{q}_0, \tag{12}$$

just using the fact that the eigenvector transformations

$$\mathbf{p}_0 = \mathbf{U} \mathbf{x}_0 \quad \text{and} \quad \mathbf{q}_0 = \mathbf{U} \mathbf{y}_0 \tag{13}$$

being unitary are norm conserving and that, furthermore, the difference matrix can be substituted by its spectral decomposition:

$$\mathbf{D} = \mathbf{U} \Delta \mathbf{U}^+. \tag{14}$$

Remark 4. The new diagonal form of the theorem 1, as set in equation (12), can be considered as a simple change of coordinates, in this case coinciding with some N-dimensional rotation, performed on both matrices $\mathbf{A}$ and $\mathbf{B}$, by means of the unitary transformation $\mathbf{U}$. However, the new expressions can be easily transformed, because of the diagonal structure of the actual matrix, as:

$$\begin{aligned} \mathbf{p}_0^+ \Delta \mathbf{p}_0 &= \sum_I \Delta_{II} |p_{0;I}|^2 = \sum_I \Delta_{II} \pi_{0;I} = \langle \Delta | \pi_0 \rangle, \\ \mathbf{q}_0^+ \Delta \mathbf{q}_0 &= \sum_I \Delta_{II} |q_{0;I}|^2 = \sum_I \Delta_{II} \chi_{0;I} = \langle \Delta | \chi_0 \rangle. \end{aligned} \tag{15}$$

In the final equations above, the diagonal matrix has been transformed into a real vector $|\Delta\rangle$, and two vectors possessing positive definite real elements are readily defined: $|\pi_0\rangle$ and $|\chi_0\rangle$.

Such new vectors are straightforwardly generated from the transformed vectors, that is [2][1]: $R(\mathbf{p}_0 \rightarrow |\pi_0\rangle)$ and $R(\mathbf{q}_0 \rightarrow |\chi_0\rangle)$. Besides, both vectors by con-

[1] In order to have more information on the symbols employed here, apart to peruse reference [2], it can be said that the following conventions have been used:

(1) *Generating vector*: $R(\mathbf{p} \rightarrow |\pi\rangle) \equiv \{|\pi\rangle_I = |p_I|^2\}$.

(2) *Vector convex property*: $K(|\pi\rangle) \equiv \{\forall I: \ |\pi\rangle_I \in \mathbf{R}^+ \text{ and } \langle |\pi\rangle \rangle = 1\}$.

(3) *Vector elements summation*: $\langle |\pi\rangle \rangle = \sum_I |\pi\rangle_I$.

struction are fulfilling convex conditions, that is, $K(|\pi_0\rangle)$ and $K(|\chi_0\rangle)$, and therefore, $\langle|\pi_0\rangle\rangle = \langle|\chi_0\rangle\rangle = 1$. Vectors $|\pi_0\rangle$ and $|\chi_0\rangle$, as previously defined, can be considered as elements of a *vector semispace* and as normalized in the *Minkowski norm* sense [3]. Vector semispaces are made as usual vector spaces, but are restrictedly defined over the positive definite real field $\mathbf{R}^+$ only. The main axiomatic characteristic of semispaces consists of the vector addition is constructed with an Abelian *semigroup* structure (see [3] for recent studies), instead of the usual Abelian additive group associated to vector spaces. Semigroups are groups without reciprocal elements. Additive semigroups like the ones employed in order to construct semispaces, lack of negative elements and, thus, differences, negative vectors and scalars are not present.

Thus, theorem 1 can be now stated by means of the alternative inequality:

$$\langle\Delta|\pi_0\rangle < \langle\Delta|\chi_0\rangle. \tag{16}$$

Remark 5. Theorem 1, as shown in equation (16) above, can be also interpreted in terms of a relationship involving two expectation values, associated to the scalar products appearing in the inequality. Both inequality expectation values are computed by means of a given set of real values, represented by the real eigenvalues of the difference matrix in vector form: $|\Delta\rangle$, and a pair of discrete probability density vectors, made, in turn, with the transformed eigenvector elements: $\{|\pi_0\rangle; |\chi_0\rangle\}$. However, due to the simple form of the expectation values appearing in equation (16), the theorem 1 can be also expressed at this level as

$$\langle\Delta|\pi_0\rangle - \langle\Delta|\chi_0\rangle < 0 \quad \Longrightarrow \quad \langle\Delta|\delta_0\rangle < 0, \tag{17}$$

with the new vector, $|\delta_0\rangle$, simply represented by the difference between the implied discrete probability distributions:

$$|\delta_0\rangle = |\pi_0\rangle - |\chi_0\rangle. \tag{18}$$

Remark 6. The final form of theorem 1, as represented in equation (17) above, coincides with a new analysis of the *Hohenberg–Kohn theorem* (HKT) [5], as was described recently by Sen et al. [6]. Within HKT and using such an alternative formalism, the difference matrix, **D**, can be seen, according to the usual DFT lore [1,5], as a non-local potential representation in some chosen basis set, if matrices **A** and **B** are considered, in addition, as Hamiltonian representations of systems with the same number of electrons.

However, as in the derivation of all the results presented here, there have been no assumptions other than choosing the pair of involved matrices as Hermitian and non-commutative, then theorem 1 and all of its formal aspects and various levels, as deduced here, seems to be a general property of Hermitian non-commutative operators, and therefore HKT shall be considered as a corollary of such a general theorem.

The first Hohenberg–Kohn theorem demonstrates that the electron density of a system uniquely determines the Hamiltonian, and as such all properties of the system. Consider the case of two external potentials differing by more than a mere constant. This will give rise to different Hamilton operators, denoted $\widehat{\mathbf{H}}_a$ and $\widehat{\mathbf{H}}_b$. As in theorem 1

described above, it is assumed that the lowest eigenvectors of both systems are different, that is the ground state is non-degenerate. Since the electron density is obtained through wave functions quadrature, it is not out of question to suppose that the same electron density may be obtained for both systems, that is:

$$\widehat{\mathbf{H}}_a \rightarrow \Psi_a \rightarrow \rho_a(r) = \rho_b(r) \leftarrow \Psi_b \leftarrow \widehat{\mathbf{H}}_b. \tag{19}$$

Now one can use Ψ_a as a trial function for $\widehat{\mathbf{H}}_b$ and vice versa. Using the above mentioned quadrature relation, one can write for systems with Hamiltonians a and b [6]:

$$E_{a,a} = \left\langle \widehat{H}_a \middle| \rho_a \right\rangle, \tag{20}$$

$$E_{b,b} = \left\langle \widehat{H}_b \middle| \rho_b \right\rangle. \tag{21}$$

This notation implies that the energy for system a has been obtained using the variationally optimized wave functions, giving the electron density ρ_a. One can now use the wave function Ψ_b as a trial function on the system with Hamilton operator $\widehat{\mathbf{H}}_a$. Using the non-degenerate assumption set out in the beginning, the variational principle becomes:

$$E_{a,b} = \left\langle \widehat{H}_a \middle| \rho_b \right\rangle > \left\langle \widehat{H}_a \middle| \rho_a \right\rangle = E_{a,a}. \tag{22}$$

A similar equation holds for system b, where Ψ_a is used:

$$E_{b,a} = \left\langle \widehat{H}_b \middle| \rho_a \right\rangle > \left\langle \widehat{H}_b \middle| \rho_b \right\rangle = E_{b,b}. \tag{23}$$

If we now add these inequalities, and re-arrange terms, we obtain:

$$\left\langle \widehat{H}_b - \widehat{H}_a \middle| \rho_a \right\rangle > \left\langle \widehat{H}_b - \widehat{H}_a \middle| \rho_b \right\rangle. \tag{24}$$

It is immediately clear now that these equations are completely similar to the situation encountered in theorem 1. The operator **D** in theorem 1 is the difference between both Hamilton operators, being the difference located in external potentials. In other words, from the general theorem concerning Hermitian operators, the same result as above can be deduced. Coming back to the starting point and accepting that both systems, differing in external potential have the same electron density, that is,

$$\rho_a = \rho_b, \tag{25}$$

an impossible inequality

$$\left\langle \widehat{H}_b - \widehat{H}_a \middle| \rho_a \right\rangle > \left\langle \widehat{H}_b - \widehat{H}_a \middle| \rho_a \right\rangle \tag{26}$$

is found which proves that the starting supposition was not suitable.

Theorem 1 has to be regarded as a consequence of the Hermitian nature of the two involved matrices and of the variational principle as well.

1.2. Extension of theorem 1 to the entire matrix spectrum

An important question can be studied now. It concerns the possibility to extend theorem 1 to the realm of eigenvectors, associated to eigenvalues situated at higher values in the spectra of both matrices, provided that no common eigenvectors are present in the implied characteristic basis sets. The result will obviously depend on the behaviour of the approximate eigenvalues, which shall be computed as in equation (4), but for spectral elements other than the zeroth ones, that is:

$$\forall I > 0: \quad \alpha_I^a = \mathbf{y}_I^+ \mathbf{A} \mathbf{y}_I \langle ?: \ A \rangle \alpha_I \quad \text{and} \quad \beta_I^a = \mathbf{x}_I^+ \mathbf{B} \mathbf{x}_I \langle ?: \ B \rangle \beta_I , \tag{27}$$

where in equation (27) above, the symbol $\langle ?: \ [index] \rangle$ is used to indicate that there is no possibility to know for each equation, in general, the relative magnitude of the approximate eigenvalues with respect to the exact ones. But there are not so many possibilities as to impede to look at them separately.

First, in case that both symbols can be substituted by the same as in the zeroth case, then obviously the theorem will hold equally. The same trivial situation will occur when both can be written in a reversed manner than the zeroth case, as: $\langle ?: \ A \rangle \equiv \langle ?: \ B \rangle \equiv <$, providing with a sign reversal and a result as the one given in the remark 2. The case where $\langle ?: \ A \rangle \equiv >$ and $\langle ?: \ B \rangle \equiv <$ and the symmetric one, $\langle ?: \ A \rangle \equiv <$ and $\langle ?: \ B \rangle \equiv >$, can be obviously studied in the same way. Due to this, only the first non-trivial case will be put forward.

Next, in order to do this analysis for the $\langle ?: \ A \rangle \equiv >$ and $\langle ?: \ B \rangle \equiv <$ case, one can suppose that

$$\exists I > 0: \quad \alpha_I^a = \mathbf{y}_I^+ \mathbf{A} \mathbf{y}_I > \alpha_I \quad \text{and} \quad \beta_I^a = \mathbf{x}_I^+ \mathbf{B} \mathbf{x}_I < \beta_I . \tag{28}$$

Thus the following equations are obtained:

$$\begin{aligned} \alpha_I^a - \alpha_I &= \mathbf{y}_I^+ \mathbf{A} \mathbf{y}_I - \mathbf{x}_I^+ \mathbf{A} \mathbf{x}_I > 0, \\ \beta_I - \beta_I^a &= \mathbf{y}_I^+ \mathbf{B} \mathbf{y}_I - \mathbf{x}_I^+ \mathbf{B} \mathbf{x}_I > 0, \end{aligned} \tag{29}$$

which summed up as performed in theorem 1, produce:

$$\left(\alpha_I^a + \beta_I\right) - \left(\alpha_I + \beta_I^a\right) = \mathbf{y}_I^+ (\mathbf{A} + \mathbf{B}) \mathbf{y}_I - \mathbf{x}_I^+ (\mathbf{A} + \mathbf{B}) \mathbf{x}_I > 0. \tag{30}$$

Accordingly, it can be written as follows:

$$\mathbf{x}_I^+ (\mathbf{A} + \mathbf{B}) \mathbf{x}_I < \mathbf{y}_I^+ (\mathbf{A} + \mathbf{B}) \mathbf{y}_I . \tag{31}$$

So, in this crossed relationship case the sum, $\mathbf{S} = \mathbf{A} + \mathbf{B}$ of the implied matrices has the leading role, or in the same way as in theorem 1, it can be written as follows:

$$\mathbf{x}_I^+ \mathbf{S} \mathbf{x}_I < \mathbf{y}_I^+ \mathbf{S} \mathbf{y}_I . \tag{32}$$

Thus, the analysis of theorem 1 based on the difference matrix $\mathbf{D}$, can be performed here, exactly in the same way for the matrix sum $\mathbf{S}$. Therefore, this shows that even for

excited states, there should be present a HKT form available, which could be written now as

$$\langle \Sigma | \pi_I \rangle < \langle \Sigma | \chi_I \rangle, \tag{33}$$

provided that within the real vector $|\Sigma\rangle$ are contained the eigenvalues of the matrix sum. The transformed initial eigenvectors permit finally, employing the same arguments as those used before, to write:

$$\langle \Sigma | \delta_I \rangle < 0, \tag{34}$$

in case that the new vector $|\delta_I\rangle$, corresponds to the difference between the discrete probability density distributions associated to the original Ith eigenvectors:

$$|\delta_I\rangle = |\pi_I\rangle - |\chi_I\rangle. \tag{35}$$

This new possible relationship, even for excited states, when the involved operator representations are taken as quantum chemical Hamiltonian representations may be observed as the final proof of the generality and application power which can be associated to theorem 1.

As a consequence of the remarks on theorem 1, and of the discussion presented above, a more general theorem can be settled in order to finish this study in a complete manner.

Theorem 2. Given two non-commuting Hermitian matrices $\mathbf{A}, \mathbf{B} \in M_{(N\times N)}(\mathbf{C})$ and knowing their eigenvalues and different eigenvectors $\{\alpha_I; \mathbf{x}_I\}$ and $\{\beta_I; \mathbf{y}_I\}$. Then, the inequality $\mathbf{x}_I^+ \mathbf{Z} \mathbf{x}_I < \mathbf{y}_I^+ \mathbf{Z} \mathbf{y}_I$, will be fulfilled in any case, with the matrix $\mathbf{Z}$ being the difference $\mathbf{D} = \mathbf{A} - \mathbf{B}$, or the sum $\mathbf{S} = \mathbf{A} + \mathbf{B}$.

Theorem 2 has been already discussed for practically all of the possible cases, so it will be proposed without proof.

2. Conclusions

The present paper theorems, as far as the authors know, seem to have been unnoticed in the specialised literature (see, for example, [7]). Theorem 1 is just a particular case of theorem 2, but the present work has been presented in this way in order to stress the connection of the present study with HKT. The immediate and practical conclusion concerning quantum chemistry, and which can be obtained from the above discussed properties for any pair of non-commuting Hermitian matrices, can be easily resumed by stressing the fact that theorem 2 proves there exists a possible extension of DFT to excited states as well. However, the theorem possesses a wide applicability encompassing Hermitian non-commuting operators and their representations, thus the described property thru the theorems holds not only for quantum systems energy operators, but for other kinds of quantum mechanical observables, and for metric matrices as well. The presented theorems are demonstrated assuming the leading role of the non-commuting

matrices set, then, necessarily, within the quantum mechanical theoretical structure, both have to be related with such a basic quantum mechanical idea as Heisenberg uncertainty principle.

Acknowledgements

This paper has benefited of important discussions with Professor Andreas Savin, Universite Pierre et Marie Curie, Paris, and Professor Brian Sutcliffe of the Universite Libre de Bruxelles, Belgium. The authors want to acknowledge the Foundation M.F. de Roviralta as well as the CICYT project #SAF2000-223, which have supported this work. The visiting professorship provided for one of us (P.B.) by CEPBA is also gratefully acknowledged.

References

[1] R.G. Parr and W. Yang, *Density Functional Theory of Atoms and Molecules* (Oxford University Press, Oxford, 1989).
[2] R. Carbó-Dorca, in: *Advances in Molecular Similarity*, Vol. 2, eds. R. Carbó-Dorca and P.G. Mezey (JAI Press, London, 1998).
[3] R. Carbó-Dorca, Shell partition and metric semispaces: Minkowski norms, root scalar products, distances and cosines of arbitrary order, Institute of Computational Chemistry Technical Report (IT-IQC-02-18), University of Girona (2002); Theoretical foundations of quantum quantitative structure–properties relationships, Institute of Computational Chemistry Technical Report (IT-IQC-01-29), University of Girona (2001).
[4] I.M. Vinogradov, *Encyclopaedia of Mathematics* (Kluwer Academic Publishers, Dordrecht, 1992).
[5] P. Hohenberg and W. Kohn, Phys. Rev. 136B (1964) 864.
[6] K.D. Sen, E. Besalú and R. Carbó-Dorca, J. Math. Chem. 25 (1999) 253.
[7] G.H. Golub and Ch.F. van Loan, *Matrix Computations* (Johns Hopkins Press, Baltimore, MD, 1989); R.A. Horn and Ch. A. Johnson, *Matrix Analysis* (Cambridge University Press, Cambridge, 1985); J.H. Wilkinson, *The Algebraic Eigenvalue Problem* (Clarendon Press, Oxford, 1965); E. Durand, *Solutions numériques des équations algébriques* (Masson & Cie., Paris, 1961).