Maximal Probability Domains in Linear Molecules

A. GALLEGOS,[1] R. CARBÓ-DORCA,[1] F. LODIER,[2] E. CANCÈS,[2] A. SAVIN[3]
[1] *Institute of Computational Chemistry, Campus Montilivi, University of Girona,*
17071 Girona, Spain
[2] *CERMICS, Ecole Nationale des Ponts et Chaussées, 6-8 avenue B. Pascal,*
F-77455 Marne la Vallée, France
[3] *Laboratoire de Chimie Théorique, CNRS et Université Paris VI, 4, place Jussieu,*
Couloir 22-23, premier étage, Case courrier 137, F-75252 Paris, Cedex 05, France

Received 23 October 2004; Accepted 25 November 2004
DOI 10.1002/jcc.20180

Abstract: Regions of space are defined to maximize the probability to find a given number of electrons within. Their chemical significance and their relationship to the electron localization function (ELF) are explored by analyzing the results for a few linear molecules: LiH, BH, N_2, CO, CS, C_2H_2, and C_4H_2.

J Comput Chem 26: 455–460, 2005

Key words: maximal probability domains; linear molecules; electron localization function (ELF)

Introduction

Quantum mechanics allows electrons to be delocalized over the whole space. Chemical intuition, however, works with electrons "localized" within a given region of space (pairs of electrons, shells, etc.). Consider a system of *N* electrons. When an arbitrary region of space, Ω, is chosen, quantum mechanics can yield the probabilities of finding a given number of electrons, ν $(0 \leq \nu \leq N)$ in it. To find chemical concepts without violating quantum mechanics, it has been proposed long ago to analyze how probabilities change when the space regions are modified.[1,2] There is some flexibility in choosing the criterion defining the region of interest. In this article we choose the simplest possible: the region should maximize $p(\nu, \Omega)$, the probability of finding exactly ν electrons in it.

Although the regions are now defined, their chemical signification is not. Only experience can tell whether such regions are able to describe chemical concepts. A previous article[3] has shown that the above-defined regions can describe well the (valence) atomic shells (Li to Xe). It also pointed out that to choose which number of electrons, ν, corresponds to a shell, it is useful to compare the probabilities obtained for the quantum mechanical system with those that would be obtained for independent particles, having the same population (average number of electrons in the region): quantum mechanical effects enhance the probability of having ν electrons in a region, if ν corresponds to a shell. To express it in a more precise manner, consider

$$\Delta p(\nu, \Omega_\nu) = p(\nu, \Omega_\nu) - p_{\text{indep}}(\nu, \Omega_\nu) \tag{1}$$

where Ω_ν is a region maximizing $p(\nu, \Omega)$ and $p_{\text{indep}}(\nu, \Omega)$ is the probability of finding ν and only ν independent, indistinguishable particles in Ω. Here,

$$p_{\text{indep}}(\nu, \Omega) = \binom{N}{\nu} p_{\text{bin}}^{\nu} (1 - p_{\text{bin}})^{N-\nu} \tag{2}$$

where

$$p_{\text{bin}} = \frac{\sum_{\nu=0,N} \nu p(\nu, \Omega)}{N}$$

guarantees that the population of the independent particles in Ω is the same as for the real ones.[3] Analyzing the atomic restricted Hartree–Fock wave functions showed that $\Delta p(\nu, \Omega) \geq \Delta p(\nu \pm 1, \Omega_{\nu\pm1})$ when ν corresponds to a shell. There is, however, a natural restriction to take into account when applying the above inequality. The mathematical origin of the restriction lies in the fact that there can be several local maxima of $\Delta p(\nu, \Omega_\nu)$. The possibility of several local maxima has a physical origin. A simple example is given by the Ar atom, which has two shells with eight electrons, and both should to be recovered with the present procedure. Thus,

---

Correspondence to: A. Savin; e-mail: Andreas.savin@lct.jussieu.fr

Contract/grant sponsor: a predoctoral fellowship from the Ministerio de Ciencia y Tecnologí (to A.G.S.)

Contract/grant sponsor: a mobility grant (to A.S.) from CESCA-CEPBA

it is meaningful to compare the probabilities for $\Omega_{\nu\pm1}$ and Ω_ν: $\Omega_{\nu+1}$ contains Ω_ν, and Ω_ν contains $\Omega_{\nu-1}$.

This article extends the previous analysis to linear molecules. The regions of Ω were optimized, yielding different $p(\nu, \Omega_\nu)$. Next, values for different ν were compared, using $\Delta p(\nu, \Omega_\nu)$. Finally, these regions are discussed for their chemical significance. In this part, a comparison with regions defined using the Electron Localization Function (ELF),[4] is made.

Theory and Methods

Algorithm

The formulas for obtaining $p(\nu, \Omega)$ were those given in a previous article.[5] For single Slater determinants only the overlap integrals within Ω are needed to obtain $p(\nu, \Omega)$. For optimization, we used two different algorithms; one, using the level-set method was published in ref. 5; the other one, used in the present article, will be shortly described in the following. First, the space is divided into small cells. For linear molecules, cylindrical symmetry can be used, so that considering a half-plane, starting on the molecular axes is sufficient, so that the cells are rectangles. (In our numerical examples they had edges ≤ 0.1 bohr.) The overlap of the molecular integrals is computed analytically in each of the cells. One then starts with a given Ω, defined as the union of cells, the choice being given by intuition (e.g., by ELF basins[6]). When the cells are very small, the contribution of a cell to $p(\nu, \Omega)$ can be calculated by linearization. As the formula given in ref. 5 already uses the eigenvalues of the overlap matrix within Ω, this is easily done. After the contributions of all cells on the outside and inside the surface delimiting Ω have been calculated, cells are added or deleted to increase $p(\nu, \Omega)$. Thus, Ω is redesigned, and the procedure starts again, until $p(\nu, \Omega)$ does not change any more, within a given accuracy. Our results have been obtained by a program written in *Mathematica*.[7]

Limitations of Algorithm

The above-described algorithm has the following limitations: (1) the starting guess, and (2) the size of the cells. The first limitation does not seem to be very important, as we can be guided by experience. Of course, due to the existence of several maximal $p(\nu, \Omega)$, different starting guesses have to be tried.

The second limitation of the algorithm can be controlled in a systematic way, and does not seem to be a severe problem, at least in 2D. Of course, the present algorithm can only produce the Ω_ν within the precision of the size of the cells. Oscillations around the boundaries are not always avoided with this algorithm: when Ω is too small, the perturbation indicates that Ω should be increased; once increased beyond the boundary, one can discover that a reduction of the volume is favorable. When the cells are constructed fine enough, the effect on Ω_ν and $p(\nu, \Omega)$ is negligible. When the cells are large, this can seriously limit the accuracy of the calculation. Although our calculations always converged for a sufficiently fine grid, this option makes the code unnecessarily slow. An improved algorithm could refine the grid as Ω approaches convergence.

The algorithm presented here is general. However, as this work is only exploratory, we limited ourselves to molecules with cylindrical symmetry, where lower dimensionality compensates for an algorithm and a program which are not yet optimal. There is, however, a price to pay for this: some optimal regions may not show up. In systems like acetylene, where a multiple bond is present, one would expect to find three electron pairs which, like banana bonds, break the symmetry of the system. As the code keeps the symmetry, this type of solution does not show up. Imposing the symmetry upon Ω_ν does not prevent from obtaining solutions which correspond to the σ bond.

Some of the technical problems mentioned above could be circumvented by a variable grid, for example, by choosing larger cells in the regions where the density is low and by refining the grid in the border regions.

Restricted Hartree–Fock

The quantum mechanical input to these programs is provided by calculations in the 6-311G** basis, at the Hartree–Fock level, for experimental (NIST Scientific Databases, http://www.nist.gov), or for calculated geometries (optimized using restricted Hartree–Fock, with 6-311G** basis set[8]).

This is clearly a limitation, and our present results may thus be considered as questionable. Comparisons with Quantum Monte Carlo calculations show, however, in general, a surprisingly good agreement with restricted Hartree–Fock calculations (Scemama, A., J Theor Comp Chem). A typical exception to this rule is the stretched hydrogen molecule. Further calculations will be done to assert the effect of correlation on the Ω_ν.

Results

Simple Example: LiH; ν, $N - \nu$ Complementarity

Let us start with the simple example of the LiH molecule. The trivial cases of the regions having no electrons ($\nu = 0$) or having all electrons in it ($\nu = N$), leads at the discretized level to the empty set, or to the union of all the cells of the mesh, respectively. Starting with a sphere around the Li atom, we get, for $\nu = 1$ a roughly spherical region around the Li nucleus, with $p(1, \Omega_1) \approx 0.5$ and $\Delta p(1, \Omega_1) \approx 0.08$. For $\nu = 2$, this domain gets larger, and $p(2, \Omega_2) \approx 0.94$, and $\Delta p(2, \Omega_2) \approx 0.56$. For $\nu = 3$, the domain further increases, but $p(3, \Omega_3)$ and $\Delta p(3, \Omega_3)$ fall back to values close to those given for $\nu = 1$. We thus select Ω_2 as chemically significant. It describes the "Li core." In using the data it may be useful to notice that the probability to find electrons in Ω is equal to that of finding $N - \nu$ in the rest of the space (in the complement of Ω). In the example above, if Ω_2 maximizes the probability of finding two electrons around the Li nucleus ("Li core," or "Li^+"), its complement is a domain Ω_2 containing the H atom and maximizing the probability of finding $4 - 2 = 2$ electrons in it ("H^-," or "LiH bond").

Figure Description

The results relevant for the discussion are collected in Figure 1 and in Table 1. As the molecules have cylindrical symmetry, Figure 1 shows the results in half-planes containing the molecular axis. The position of the atoms is marked by black half-disks, and the

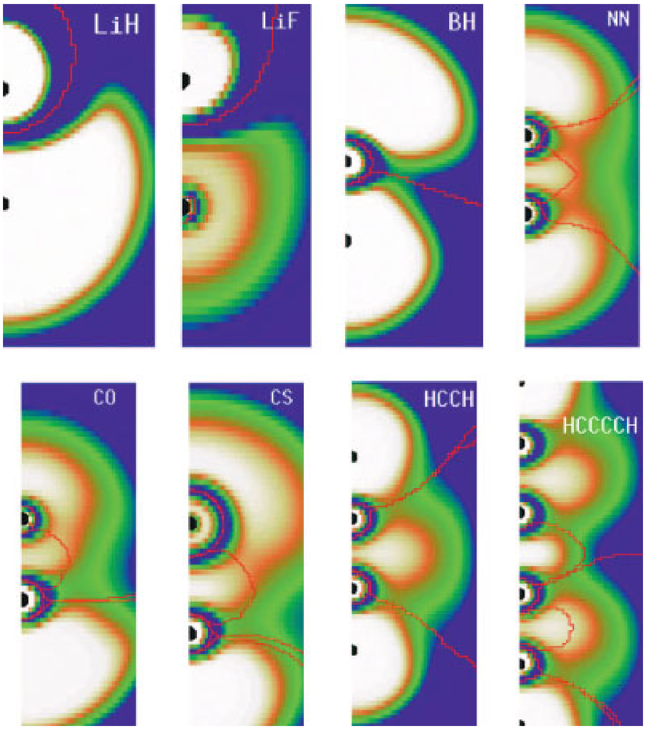

Figure 1. Electron localization function (ELF) representation for: (a) LiH, Ω_2, Li core/H$^-$; (b) LiF, Ω_2, Li core/F$^-$; (c) BH, Ω_2, lone pair, B core, BH bond; (d) N_2, N cores, and lone pairs, triple, and sigma bond; (e) CO, cores, C lone pair, sigma bond, and "rest"; (f) CS, as for CO, before; (g) HCCH, cores, CH bonds, triple bond; (h) HCCCCH, cores, CH bonds, triple, and sigma bonds. The atoms have been ordered from the top of the figure to the bottom. The contours of some relevant Ω_η are shown in red.

colored background shows the ELF values. The contours of some relevant Ω_ν are shown in red, while some other contours, which are evident (cores, contours generated by symmetry) are not shown. It must be noticed that no smoothing was used, the step-structure in the contours reflecting the size of the cells. The colored background shows the electron localization function (ELF). Recall that the regions where electron pairs are formed are expected to yield large values of ELF. As in ref. 9, the color code used for ELF is borrowed from geographical maps; high values (close to 1) are white, around 0.8 the color turns to brown, around 0.6 to green, becoming a light green around 0.4, to start to go over to blue around 0.2. Using ELF, space can be divided into "basins;" a basin is the collection of points in space from which following the ELF gradient one reaches the same maximum. One can see in Figure 1 that for the LiH molecule Ω_2 yields, within our accuracy, the same result as ELF: a separation into a region around a Li "core" and the remaining part of a molecule. One may call this separation ionic. Notice, however, that we have not given any criterion to distinguish between an "H$^-$" ion, and an LiH "bond."

Comparison with ELF

It is interesting to notice, in the case of LiH, the agreement of the Ω_ν with the ELF basins. This is also the case for BH. However, especially when the population of Ω_ν deviates from ν, the agreement is not as good. A typical example is given by the N_2 molecule. In this case, ELF yields a basin population of the lone pair of ≈3 electrons, while for Ω_2 it is close to ν, that is, 2, and to the "chemical" picture. Notice that when probabilities are used to define criteria, like $\Delta p(\nu, \Omega_\nu)$, the number of electrons ν, is automatically fixed to an integer; the noninteger number given by populations is the result of averaging.

Table Description

The information given in Table 1 is the following. For each molecule (first column) selected regions Ω_ν are given (column 2). These are identified either (1) by the nucleus it contains (e.g., B); or (2) by a "lone pair-like" region to which it is attached, followed by "lp" in parentheses, for example, B(lp); or (3) by the pair of nuclei to which a "bond-like" region between the nuclei exists (e.g., CC); a further characterization is given by a word in parentheses (e.g., terminal). The following columns give, respectively, ν, $p(\nu, \Omega_\nu)$ and $\Delta p(\nu, \Omega_\nu) \approx 0.5$. For example, the data corresponding to the contour previously described for the LiH molecule in Figure 1 appears in the second line of the table, for $\nu = 2$. One can see that $p(2, \Omega_2)$ reaches quite a high value, 0.94, and that $\Delta p(2, \Omega_2)$ is quite large (≈0.56), in contrast to $\Delta p(2 \pm 1, \Omega_{2\pm 1}) \approx 0.08$. As explained in the introduction, the analysis of the data is based upon the $\Delta p(\nu, \Omega_\nu)$. Thus, the Ω_2 shown in Figure 1 for the LiH molecule is considered "relevant."

Atoms in Molecules, Shells in Molecules

The procedure described does not provide "atoms in molecules." For example, in LiH, there is no maximum in $\Delta p(\nu. \Omega_\nu)$ found, for $\nu = 3$, so that the Li atom is not recovered. A similar result is obtained for Li in LiF. Furthermore, no "H atom" is found in LiH, and no "F atom" in LiF, no "B atom" in BH, no "C atom" in CO, etc. (We do not present in the table the results for "H" in LiH, and for "F" in LiF, as these can be deduced from the $N - \nu$ data for "Li.") Lone pairs and bonds can provide supplementary maxima. For example, in BH there is a maximum for $\Delta p(4, \Omega_4)$ around the B atom given by the core and the lone pair region. In acetylene, a maximum is formed for $\Delta p(10, \Omega_{10})$ around the C atom, formed by the core, the CH bond, and the triple bond. It is, however, more interesting to look at these quantities individually, than to see them as "deformed atomic shells."

Populations

The populations were not given in Table 1. Within our numerical accuracy they were close to the integer numbers ν. As noticed in ref. 3, $p(\tilde{\nu}, \Omega_\nu) \approx (2\pi\sigma^2)^{-1/2} \exp[-(\tilde{\nu} - \bar{\nu}^2)^2/2\sigma,^2$ where $\tilde{\nu} = \Sigma\tilde{\nu}p(\tilde{\nu}, \Omega_\nu)$ and $\sigma^2 = \Sigma(\tilde{\nu} - \bar{\nu})^2\, p(\tilde{\nu}, \Omega_\nu)$. This result points out towards a general feature, namely that the $p(\tilde{\nu}, \Omega_\nu)$ are distributed in a symmetrical way around its maximum, and that they decrease fast around ν.

Partition of Space

The previous remark does not explain a phenomenon observed now in many cases: the "relevant" Ω_ν seem to produce a partition

Table 1. Results of $p(\nu, \Omega_\nu)$, and $\Delta p(\nu, \Omega_\nu)$ for the Selected Region Ω_ν and ν of Each Studied Molecule.

Molecule	Region	ν	$p(\nu;\ \Omega_\nu)$	$\Delta p(\nu;\ \Omega_\nu)$	Molecule	Region	ν	$p(\nu;\ \Omega_\nu)$	$\Delta p(\nu;\ \Omega_\nu)$
LiH	Li	1	0.50	0.08	N_2	N(lp)	3	0.38	0.12
LiH	Li	2	0.94	0.56	N_2	NN	1	0.44	0.06
LiH	Li	3	0.50	0.08	N_2	NN	2	0.38	0.09
LiF	Li	1	0.50	0.12	N_2	NN	3	0.33	0.08
LiF	Li	2	0.93	0.63	N_2	NN	4	0.31	0.08
LiF	Li	3	0.44	0.18	N_2	NN	5	0.31	0.09
LiF	F	1	0.40	0.05	N_2	NN	6	0.31	0.10
LiF	F	2	0.71	0.41	N_2	NN	7	0.31	0.10
LiF	F	3	0.39	0.13	N_2	NN	8	0.29	0.08
BH	B	1	0.49	0.09	CO	C	1	0.43	0.07
BH	B	2	0.86	0.53	CO	C	2	0.81	0.51
BH	B	3	0.48	0.17	CO	C	3	0.40	0.15
BH	B	4	0.76	0.43	CO	C	4	0.56	0.33
BH	B	5	0.50	0.10	CO	C	5	0.39	0.18
BH	B(lp)	1	0.50	0.10	CO	C	6	0.34	0.13
BH	B(lp)	2	0.74	0.42	CO	C(lp)	1	0.49	0.10
BH	B(lp)	3	0.34	0.06	CO	C(lp)	2	0.57	0.28
C_2H_2	C	1	0.48	0.11	CO	C(lp)	3	0.38	0.12
C_2H_2	C	2	0.79	0.49	CO	O	1	0.37	0.03
C_2H_2	C	3	0.39	0.14	CO	O	2	0.74	0.45
C_2H_2	H(CH)	1	0.48	0.10	CO	O	3	0.39	0.14
C_2H_2	H(CH)	2	0.57	0.28	CO	CO	1	0.44	0.07
C_2H_2	H(CH)	3	0.37	0.12	CO	CO	2	0.39	0.09
C_2H_2	CC	1	0.44	0.06	CO	CO	3	0.34	0.09
C_2H_2	CC	2	0.37	0.08	CO	CO	4	0.31	0.08
C_2H_2	CC	3	0.30	0.06	CO	CO	5	0.31	0.09
C_2H_2	CC	4	0.26	0.04	CO	CO	6	0.31	0.11
C_2H_2	CC	5	0.33	0.11	CO	CO	7	0.34	0.13
C_2H_2	CC	6	0.36	0.15	CO	CO	8	0.46	0.25
C_2H_2	CC	7	0.31	0.10	CO	CO	9	0.36	0.14
C_4H_2	H(CH)	1	0.48	0.10	CS	C	1	0.43	0.08
C_4H_2	H(CH)	2	0.56	0.28	CS	C	2	0.81	0.53
C_4H_2	H(CH)	3	0.36	0.12	CS	C	3	0.45	0.21
C_4H_2	CC(central)	1	0.45	0.07	CS	C	4	0.54	0.32
C_4H_2	CC(central)	2	0.42	0.14	CS	C	5	0.39	0.19
C_4H_2	CC(central)	3	0.33	0.10	CS	C(lp)	1	0.48	0.11
C_4H_2	CC(terminal)	1	0.42	0.05	CS	C(lp)	2	0.55	0.26
C_4H_2	CC(terminal)	2	0.37	0.09	CS	C(lp)	3	0.37	0.13
C_4H_2	CC(terminal)	3	0.32	0.08	CS	S	9	0.39	0.22
C_4H_2	CC(terminal)	4	0.32	0.11	CS	S	10	0.64	0.47
C_4H_2	CC(terminal)	5	0.33	0.14	CS	S	11	0.39	0.22
C_4H_2	CC(terminal)	6	0.35	0.16	CS	CS	1	0.45	0.08
C_4H_2	CC(terminal)	7	0.30	0.13	CS	CS	2	0.39	0.11
N_2	N	1	0.48	0.11	CS	CS	3	0.34	0.09
N_2	N	2	0.77	0.48	CS	CS	4	0.31	0.09
N_2	N	3	0.40	0.14	CS	CS	5	0.30	0.10
N_2	N	4	0.45	0.21	CS	CS	6	0.31	0.12
N_2	N	5	0.38	0.16	CS	CS	7	0.33	0.15
N_2	N(lp)	1	0.47	0.09	CS	CS	8	0.41	0.23
N_2	N(lp)	2	0.47	0.18	CS	CS	9	0.37	0.20

of space. For example, in the BH molecule, the "BH bond," the "B lone pair," and the "B core" are three different maxima of $p(2, \Omega_2)$. The three Ω_2 seem, within our accuracy, to divide space into three regions. In Figure 1, the three regions were plotted independently, but one may notice that the same segment of curve is part of the boundary of the "BH bond" and the "B lone pair." Sometimes, due to the technical limitations mentioned above, it is not possible to distinguish whether there really is just a common separating surface between two optimized regions or not. For example, in the case of the BH molecule, one can notice that there is just one cell that makes a difference between the surface of the "B core" and the "BH bond" regions. Due to the smallness of the effects seen in the figures, it was not considered necessary to push further the accuracy, than showed in the figures: for the examples studied here, one can consider using that the maximal $\Delta p(\nu, \Omega_\nu)$ form a partition of space, although this was not required by the definition of the Ω_ν.

Cores, Polarization, and Penetration

We will start now analyzing the data. We first notice the presence of the cores. In all cases, they are well separated from the valence, with $\Delta p(\nu, \Omega_\nu)$ around 0.5. It can be noticed, however, that they can be polarized by the molecular environment. Already LiH shows this effect. There is no surprise in seeing that both "H^{--}" (in LiH) and "F^-" (in LiF) is highly deformed ("polarized") by "Li^+," which "penetrates" them (due to "electrostatic interaction"). The Ω_2 around the Li nuclei do not have spherical shape, as they have in the Li atom. While the "hard Li^+" penetrates the soft "anions," the "Li^+" Ω_2 is also deformed, trying to escape the "Pauli repulsion" of the "anion."

Octet Rule

For the valence shell, one might expect the octet rule to be satisfied, in general. This is the case around F ("F^-") in LiF, around O (in CO), around S (in CS), around N (lone pair and bonds in N_2, cf. maximum around $\nu = 4$, in Table 1, use the equivalence with $N - \nu$, and eliminate the core), or around the C, in C_2H_2 or C_4H_2. Notice, however, that the "shared electrons" can extend far into the region of the other atom (see, e.g., the "triple bond" in C_2H_2). For C in CS and CO it is less clear how to count. Two electrons of the lone pair and two of the σ bond can be attributed to the C atom, although the domain of the electrons around O extends around the σ bond towards the C atom: the criterion $\Delta p(\nu, \Omega_\nu) > \Delta p_{\nu\pm1}(\nu \pm 1, \Omega_{\nu\pm1})$ did not allow finding any region that would be attributed to a "covalent bond." At the present status of our calculations, the better description of these CX molecules is thus not $|C^{-1} \equiv X^{+1}|$ but rather with the π bonds formally attributed to X; one may like to write $|C^{-1} = \overline{\underline{X}}^{+1}|$. The formal charges are not related to the dipole moment, as can be easily seen from the deformation of the "π lone pairs" Ω_ν on X.

No Absolute Scale

As in atoms, we notice that there is no absolute scale connecting the values of $\Delta p(\nu, \Omega_\nu)$ to the "chemical concepts." In all the cases studied, the cores are well separated from the valence, with $\Delta p(\nu, \Omega_\nu)$ around 0.5. Also, lone pairs appear on B in BH, on N in N_2, and on C in CO and CS, with $\Delta p(2, \Omega_2)$ between ≈ 0.4 (B) to ≈ 0.2 (N). Single bonds seem also to be well described (the CH bonds in acetylene, C_2H_2, and diacetylene, C_4H_2 have $\Delta p(2, \Omega_2) \approx 0.3$). As for the lone pairs, $\Delta p(\nu \pm 1, \Omega_{\nu\pm1})$ are significantly lower (≈ 0.1). On the other hand, the characterization of multiple bonds is not as clear. For example, in acetylene, $\Delta p(6, \Omega_6) \approx 0.15$, while $\Delta p(\nu \pm 1, \Omega_{\nu\pm1})$ are still around 0.1. In diacetylene, the situation is worse, as the $\Delta p(\nu \pm 1, \Omega_{\nu\pm1})$ increase to values even closer to $\Delta p(6, \Omega_6) \approx 0.5$. The origin of this effect may be the delocalization of the triple bond. This goes hand in hand with the relatively small difference between $\Delta p(2, \Omega_2)$ and $\Delta p(\nu \pm 1, \Omega_{\nu\pm1})$ for the central CC bond in diacetylene. The effect is even more important in the nitrogen molecule: a maximum of $\Delta p(\nu, \Omega_\nu)$ can be noticed for $\nu = 6$, but it is beyond the limit of the numerical accuracy imposed. In the CO and CS molecules, no maximum has been seen for $\Delta p(\nu, \Omega_\nu)$, corresponding to $\nu = 4$ or $\nu = 6$ (double or triple bond). As ν increases, a maximum is reached for $\nu = 8$, which corresponds to an ionic O or S (the whole space excluding the cores and the lone pair on C). One can notice, however, a tiny maximum for $\nu = 2$, which corresponds to a α bond.

Resonance

To better understand these special cases, let us consider, the example of a model "hydrogen molecule," described by the wave function

$$\psi(r_1, r_2) = \frac{\lambda}{\sqrt{2}}[\chi_a(r_1)\chi_b(r_2) + \chi_b(r_1)\chi_a(r_2)] + \frac{\sqrt{1-\lambda^2}}{\sqrt{2}}\lfloor \chi_a(r_1)\chi_a(r_2) + \chi_b(r_1)\chi_b(r_2) \rfloor \quad (3)$$

Here, χ_a and χ_b are atomic orbitals, $\lambda^2 \in [0,1]$ gives the weight of the covalent structure, and $1 - \lambda^2$ the weight of the ionic (resonance) structure. The probability to find one electron on atom *a* (and one on atom *b*) is given by λ;^{2} the probability to find both electrons on atom *a* (and none on atom *b*) or of finding no electron on *a* (and both on *b*), is given by $1 - \lambda^2$). (Ω is defined now by symmetry, as the half-space obtained by the plane perpendicular to the bond, and containing the bond-midpoint.) If the electrons were independent, the probability to have zero, one or two electrons in Ω would have been 1/4, 1/2, and 1/4, respectively, yielding equal probability for a "covalent" and an "ionic structure." As long as $\lambda^2 > 1/2$, $\Delta p(1, \Omega) = \lambda^2 - 1/2$ is larger than the change in probability due to the ionic limiting forms, and the "best" description will be that of having one electron on atom *a*, and one on atom *b*. The effect decreases as the ionic structure increases in importance. Such an effect of "resonance" (or "delocalization") can be seen, for example, by comparing $\Delta p(6, \Omega_6)$ with $\Delta p(5, \Omega_5)$ and $\Delta p(7, \Omega_7)$ in acetylene and diacetylene: the difference is significantly reduced in diacetylene. Notice, however, that we have not considered in this article the cases when one has "either ν_1 or ν_2" electrons in Ω. Thus, when Ω will be optimized, say for $\nu = 2$, the whole space will result, independently of λ. Thus, a proper treatment of the case of "either ν_1 or ν_2" electrons in Ω needs to be implemented in order to better understand "resonance" in the context of the $p(\nu; \Omega)$.

Further Partitioning

It would be interesting to analyze a complete partition of space in terms of given "Lewis structures," for example, the probability to have $_A$ electrons in Ω_A, ν_B in region Ω_B, etc., with $\nu_A + \nu_{B+\,\ldots} = N$, as done in the "loge theory."[1,2] The data produced up to now do not provide the complete information. Let us take as an example the BH molecule, and define the three regions as the core, lone pair, and bond region. We ask if we can maximize the probability of having two electrons in each. The $p(\nu;\,\Omega)$ are not sufficient to obtain it, but we can make estimates. We can eliminate from the equations

$$p(\nu_i,\,\Omega_A) = \sum_{j,k(\nu_i+\nu_j+\nu_k=N)} p(\nu_i,\,\nu_j,\,\nu_k) \tag{4}$$

where $p(\nu_i,\,\nu_j,\,\nu_k)$ is the probability to find ν_i electrons in Ω_A, ν_j in Ω_B, and ν_k in Ω_C. Substituting the computed values on the left-hand side of the equation above, one obtains after some manipulation of the equation that[7]

$$p(2,\,2,\,2) = 0.68 + p(4,\,1,\,1) + p(1,\,4,\,1) + p(1,\,1,\,4) \tag{5}$$

As $0 \leq p(4,1,1) \leq p(4,\,\Omega_A)$, etc., one can obtain an upper and a lower bound to $p(2,\,2,\,2)$. It turns out that these two are very close [as $p(4,\,\Omega_A) \approx 0$], and we can estimate $p(2,\,2,\,2)$ to $\approx$0.68. The independent particle probability is now given by the multinomial distribution,

$$p_{\text{indep}}(2,\,2,\,2) = \frac{6!}{2!2!2!}\,p_A^2 p_B^2 p_C^2 \tag{6}$$

where

$$p_X = \sum_{\nu=0,N} \frac{\nu p(\nu,\,\Omega_X)}{N}, \qquad (X = A,\,B,\,C)$$

guarantees that the population of the independent particles in Ω_X is the same as for the real ones.[3] As there are roughly two electrons in each of the Ω, this yields $p_A \approx 1/6$, and a high value for

$$\Delta p(2,\,2,\,2) = p(2,\,2,\,2) - p_{\text{indep}}(2,\,2,\,2) \approx 0.56 \tag{7}$$

This value is as high as that for the core/valence separation, although $p(2,\,2,\,2)$ is significantly lower.

Conclusion

Maximizing the probability of finding a given number of electrons in a spatial domain provides a conceptually simple way of obtaining regions having "chemical meaning" in molecules (in the sense of Lewis structures): cores, bonds, lone pairs. Although the objective is similar to that of other approaches, the results obtained do not always coincide; a comparison with ELF shows good agreement in the cases where the latter is known to work well, but gives different, supposedly better, answers when the ELF analysis gives questionable results. Although the domains have been optimized separately, the optimizations seem to yield a partition of space. However, it is desirable to extend this procedure to one where all domains are simultaneously optimized. Furthermore, the function to be optimized should be extended to take into account cases when several limiting structures are important.

References

1. Daudel, R. C R Acad Sci (Paris) 1953, 237, 601.
2. Aslangul, C.; Constanciel, R.; Daudel, R.; Kottis, P. Adv Quantum Chem 1972, 6, 93.
3. Savin, A. Reviews of Modern Quantum Chemistry: A Celebration of the Contributions of Robert G. Parr; Sen, K. D., Ed.; World Scientific: Singapore, 2002, p. 43.
4. Becke, A. D.; Edgecombe, K. E. J Chem Phys 1990, 92, 5397.
5. Cancés, E.; Keriven, R.; Lodier, F.; Savin, A. Theor Chem Acc 2004, 111, 373.
6. Silvi, B.; Savin, A. Nature 1994, 371, 683.
7. Wolfram, S. The Mathematica Book; Wolfram Media: Cambridge, 1996.
8. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A., Jr.; Stratmann, R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Baboul, A. G.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Andres, J. L.; Gonzalez, C.; Head–Gordon, M.; Replogle, E. S.; Pople, J. A. Gaussian 98, Revision A.9; Gaussian, Inc.: Pittsburgh, PA, 1998.
9. Savin, A.; Becke, A. D.; Flad, J.; Nesper, R.; Preuss, H.; von Schnering, H. G. Angew Chem 1991, 103, 421. Angew Chem Int Ed Engl 1991, 30, 409.