

Regular article

How electrons guard the space: shape optimization with probability distribution criteria

Eric Cancès¹, Renaud Keriven¹, François Lodier¹, Andreas Savin²

¹ CERMICS, Ecole Nationale des Ponts et Chaussées, 6 & 8 Av. Pascal, 77455 Champs-sur-Marne, France

² Laboratoire de Chimie Théorique, CNRS et Université Paris 6, 4 Pl. Jussieu, 75252 Paris, France

Received: 7 January 2003 / Accepted: 15 May 2003 / Published online: 14 January 2004

© Springer-Verlag 2004

Abstract. Efficient formulas for computing the probabilities of finding exactly ν electrons in an arbitrarily chosen volume $\Omega \subset \mathbb{R}^3$ for Hartree–Fock wavefunctions are presented. These formulas allow the use of shape optimization techniques, such as level set methods, for optimizing with respect to Ω various criteria involving such probabilities. The criterion defined as the difference between the Hartree–Fock and the independent-particle model probabilities of finding ν electrons in Ω stresses the quantum effects due to the Pauli principle. We have implemented a 2D level set method for optimizing this criterion in order to study spatial separation of electron pairs in linear molecules. The method is described and the illustrative example of the BH molecule is reported.

Keywords: Shape optimization – Hartree–Fock – Electron pairs – Probability distribution – Spatial separator

1 Introduction

Let us consider a system of N electrons described by a wavefunction Ψ . The probability of having ν electrons in a spatial region $\Omega \subset \mathbb{R}^3$ reads

$$p_\nu(\Omega) = \binom{N}{\nu} \sum_{\sigma_1, \dots, \sigma_N} \int_{\Omega^\nu} dx_1 \cdots dx_\nu \\ \times \int_{(\mathbb{R}^3 \setminus \Omega)^{N-\nu}} dx_{\nu+1} \cdots dx_N \\ \times |\Psi(x_1, \sigma_1; \dots; x_N, \sigma_N)|^2,$$

where x_i is the position variable and σ_i the spin variable of the i th electron and where $\binom{n}{k}$ is a notation for $\frac{n!}{k!(n-k)!}$.

If the electrons were independent particles, these probabilities would simply read

$$p_\nu^{\text{ind}}(\Omega) = \binom{N}{\nu} b(\Omega)^\nu [1 - b(\Omega)]^{N-\nu}, \quad (1)$$

with $b(\Omega) = \frac{1}{N} \int_{\Omega} \rho$ (as usual, ρ denotes the electronic density).

One of us observed in previous work [1] that for atoms the maximizers with respect to R of the criteria

$$R \mapsto J_\nu(B_R) = p_\nu(B_R) - p_\nu^{\text{ind}}(B_R),$$

where B_R denotes the ball of radius R centered on the nucleus, provide radii which describe very well a spatial separation into atomic shells (for Li–Xe), a feature which is not obtained for heavy atoms when using widespread methods like the analysis of the radial distribution function [2], of the Laplacian of the density [3, 4, 5, 6] or of the electron localization function [7, 8].

The origin of criteria of analysis involving the probabilities $p_\nu(\Omega)$ can be found in the early work of Daudel and coworkers (see, e.g. Ref [9], or more recently, Refs. [10, 11, 12, 13, 14, 15]). We have, however, observed [1] that for atoms looking directly at the probabilities works even in cases when the more complex criteria used by Daudel and coworkers (e.g. the missing information function) fail.

The goal of this article is show that for molecular systems, the maximizers of the criteria

$$\Omega \mapsto J_\nu(\Omega) = p_\nu(\Omega) - p_\nu^{\text{ind}}(\Omega), \quad (2)$$

i.e. the domains Ω which maximize the value of $J_\nu(\Omega)$ for a given value of ν , provide information on the chemical structure of the system. We see J_ν as a useful measure to show the correlation enhancements of p_ν (for Hartree–Fock wavefunctions it is the Fermi correlation which is enhanced). Preliminary tests on simple molecules (we report here the illustrative example of the BH molecule) whose ground-state wavefunctions are evaluated within

the Hartree–Fock approximation indeed show that the maximizers of J_v correspond to regions of chemical significance (lone pairs, bonding, etc.), and that the level set method is well-suited for finding them.

The article is organized as follows. The methods we have used for searching for local maxima of the criteria J_v are explained in Sect. 2; efficient formulae for calculating $J_v(\Omega)$ for a given bounded open set Ω are derived in Sect. 2.1; the notion of the shape derivative and its use in shape optimization are described in Sect. 2.2; practical details for computing useful overlap integrals are provided in Sect. 2.3; lastly, a brief presentation of the level set method is the matter of Sect. 2.4. Two examples of applications, the analysis of the electronic distributions of the Hartree–Fock ground states of the Be atom and of the BH molecule, are reported in Sect. 3. The current implementation of our method is only able to deal with atoms or linear molecules in the Hartree–Fock approximation; extensions currently in progress are discussed in Sect. 4.

2 Methodology

2.1 Probabilities of having v electrons in the spatial region Ω

When Ψ is a Slater determinant, i.e. when

$$\Psi(x_1, \sigma_1; \dots; x_N, \sigma_N) = \frac{1}{\sqrt{N!}} \det [\phi_i(x_j, \sigma_j)] ,$$

with

$$\forall 1 \leq i, j \leq N, \quad \int_{\mathbb{R}^3} \sum_{\sigma} \phi_i(x, \sigma) \phi_j(x, \sigma) dx = \delta_{ij} ,$$

it can be proven (the proof is postponed until the Appendix) that

$$\forall t \in \mathbb{R}, \quad \sum_{v=0}^N p_v(\Omega) t^v = \det \{t\mathbf{S}(\Omega) + [\mathbf{I}_N - \mathbf{S}(\Omega)]\} , \quad (3)$$

where \mathbf{I}_N is the $N \times N$ identity matrix and where $\mathbf{S}(\Omega)$ denotes the $N \times N$ symmetric matrix defined by

$$\forall 1 \leq i, j \leq N, \quad S_{ij}(\Omega) = \int_{\Omega} \sum_{\sigma} \phi_i(x, \sigma) \phi_j(x, \sigma) dx . \quad (4)$$

Denoting by $[\lambda_k(\Omega)]_{1 \leq k \leq N}$ the N eigenvalues of $\mathbf{S}(\Omega)$, it follows that $[(t-1)\lambda_k(\Omega) + 1]_{1 \leq k \leq N}$ are the eigenvalues of the matrix $t\mathbf{S}(\Omega) + (\mathbf{I}_N - \mathbf{S}(\Omega))$ and therefore that

$$\forall t \in \mathbb{R}, \quad \sum_{v=0}^N p_v(\Omega) t^v = \prod_{k=1}^N [(t-1)\lambda_k(\Omega) + 1] .$$

In particular,

$$p_0(\Omega) = \prod_{k=1}^N [1 - \lambda_k(\Omega)]$$

and when $p_0(\Omega) \neq 0$

$$p_1(\Omega) = \left(\sum_{k=1}^N \frac{\lambda_k(\Omega)}{1 - \lambda_k(\Omega)} \right) p_0(\Omega),$$

$$p_2(\Omega) = \left(\sum_{j \neq k} \frac{\lambda_j(\Omega)}{1 - \lambda_j(\Omega)} \frac{\lambda_k(\Omega)}{1 - \lambda_k(\Omega)} \right) p_0(\Omega), \quad \dots$$

As the algorithmic complexities of these formulas scale as $\binom{N}{v}$, it is more convenient to use the alternative method described later to

compute $p_v(\Omega)$ for $v \geq 2$. One can indeed prove by induction that the coefficients $p_v(\Omega)$ are given by

$$\forall 0 \leq v \leq N, \quad p_v(\Omega) = a_v^N , \quad (5)$$

where the $(a_j^k)_{0 \leq k \leq N, 0 \leq j \leq k}$ are defined by the recursion

$$\begin{aligned} a_0^0 &= 1, \quad \text{and for } 1 \leq k \leq N \\ \begin{cases} a_0^k = \alpha_k a_0^{k-1} \\ a_j^k = \beta_k a_{j-1}^{k-1} + \alpha_k a_j^{k-1}, \\ \quad 1 \leq j \leq k-1 \\ a_k^k = \beta_k a_{k-1}^{k-1} \end{cases} , \end{aligned} \quad (6)$$

with

$$\forall 1 \leq k \leq N, \quad \alpha_k = 1 - \lambda_k(\Omega) \quad \text{and} \quad \beta_k = \lambda_k(\Omega) . \quad (7)$$

This method allows us to compute all the $p_v(\Omega)$ in $O(N^2)$ operations.

Equations (5), (6) and (7) are valid for any kind of Slater-type wavefunction. For wavefunctions obtained from restricted Hartree–Fock calculations (RHF), the coefficients α_k and β_k are also given by

$$\begin{aligned} \forall 1 \leq k \leq N_p, \quad \beta_{2k-1} &= 1 - \alpha_{2k-1} = \lambda_k^R(\Omega), \\ \beta_{2k} &= 1 - \alpha_{2k} = \lambda_k^R(\Omega) , \end{aligned}$$

where $N_p = N/2$ is the number of electron pairs and where the $\lambda_k^R(\Omega)$ s are the N_p eigenvalues of the $N_p \times N_p$ symmetric matrix defined by

$$\forall 1 \leq i, j \leq N_p, \quad S_{ij}^R(\Omega) = \int_{\Omega} \phi_i(x) \phi_j(x) dx . \quad (8)$$

For wavefunctions obtained from unrestricted hartree–fock (UHF) calculations, the coefficients α_k and β_k are given by

$$\begin{aligned} \forall 1 \leq k \leq N_{\alpha}, \quad \beta_k &= 1 - \alpha_k = \lambda_k^{\alpha}(\Omega), \\ \forall 1 \leq k \leq N_{\beta}, \quad \beta_{N_{\alpha}+k} &= 1 - \alpha_{N_{\alpha}+k} = \lambda_k^{\beta}(\Omega) , \end{aligned}$$

where N_{α} is the number of α electrons, N_{β} the number of β electrons and where $\lambda_k^{\alpha}(\Omega)$ and the $\lambda_k^{\beta}(\Omega)$ are the eigenvalues of the matrices $\mathbf{S}^{\alpha}(\Omega)$ (overlap of the α orbitals) and $\mathbf{S}^{\beta}(\Omega)$ (overlap of the β orbitals), respectively.

Let us conclude this section by remarking that, still denoting by ρ the electronic density, one has

$$b(\Omega) = \frac{1}{N} \int_{\Omega} \rho = \begin{cases} \frac{1}{N} \text{Tr}[\mathbf{S}(\Omega)] & \text{in the general case} \\ \frac{1}{N_p} \text{Tr}[\mathbf{S}^R(\Omega)] & \text{in the RHF setting,} \\ \frac{1}{N} \{ \text{Tr}[\mathbf{S}^{\alpha}(\Omega)] + \text{Tr}[\mathbf{S}^{\beta}(\Omega)] \} & \text{in the UHF setting .} \end{cases} \quad (9)$$

We therefore have at our disposal efficient formulas to compute the criteria $J_v(\Omega)$ defined by Eq. (2). To make use of those formulas, one, however, has to evaluate the overlap matrix $\mathbf{S}(\Omega)$. This will be the matter of Sect. 3.

We point out that one should not make the confusion of

$$\begin{aligned} b(\Omega) &= \frac{1}{N} \int_{\Omega} \rho = \sum_{\sigma_1, \dots, \sigma_N} \int_{\Omega} dx_1 \int_{(\mathbb{R}^3)^{N-1}} dx_2 \dots dx_N \\ &\quad \times |\Psi(x_1, \sigma_1; \dots; x_N, \sigma_N)|^2 \end{aligned}$$

with the probability

$$\begin{aligned} p_1(\Omega) &= \sum_{\sigma_1, \dots, \sigma_N} \int_{\Omega} dx_1 \int_{(\mathbb{R}^3 \setminus \Omega)^{N-1}} dx_2 \dots dx_N \\ &\quad \times |\Psi(x_1, \sigma_1; \dots; x_N, \sigma_N)|^2 \end{aligned}$$

of finding one and only one electron in the domain Ω . One should remember that $b(\Omega)$ is related to the probability of finding an electron in Ω , no matter where the other $N-1$ are (they can be in Ω , or outside Ω), while with $p_1(\Omega)$ we consider the probability of finding one and only one electron in Ω . Notice, however, that $p_1(\Omega) \sim_{|\Omega| \rightarrow 0} b(\Omega)$.

Another way to see the difference between $b(\Omega)$ and $p_1(\Omega)$ is to note that

$$\int_{\Omega} \rho = \sum_{v=0}^N v p_v(\Omega)$$

is the average electron number in Ω (the population of the domain Ω), which is, in general, different from $N p_1(\Omega)$.

The way $\int_{\Omega} \rho$ depends on Ω is quite different from that of p_v . Take, for example, the Be atom, and consider different spherical shells, such as $\int_{\Omega} \rho = 2$. There is a significant change in p_2 (solid line in Fig. 1), while the average is kept unchanged.

Notice also that, as $\int_{\Omega} \rho$ is unchanged in this process, p_1^{ind} is not changed with the modification of the spherical shell. Thus, J_v , the difference between p_v and the probability of having the same average for independent particles (see Eq. 2), changes in the same way as p_v . The plot of J_2 (dashed line in Fig 1) shows the curve of p_2 shifted downwards by $p_2^{\text{ind}} = 3/8$.

The maxima of p_v or J_v are not always as pronounced. Considering again the Be atom, one obtains for spherical shells keeping $\int_{\Omega} \rho = 1$ the plots for p_1 (solid line) and J_1 (dashed line) shown in Fig. 2.

For the Hartree–Fock He atom one even finds $p_1 = p_1^{\text{ind}}$: the electrons are “independent” at the Hartree–Fock level. This explains that within a shell nearly constant values are observed for J_1 .

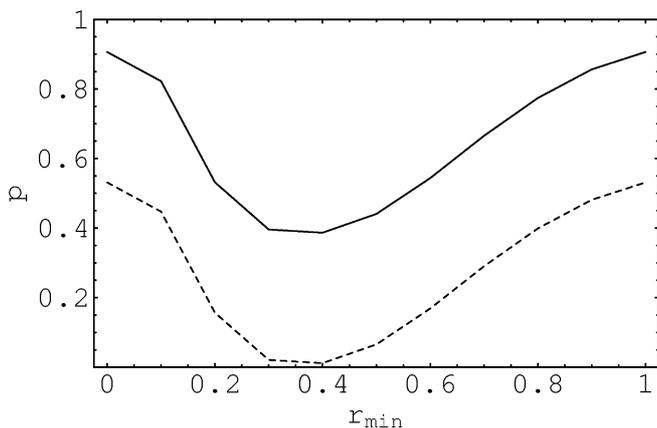


Fig. 1. The probability p_2 (solid line) of finding two electrons in a spherical shell around the Be nucleus, having inner radius r_{\min} , and having on average two electrons, from a Hartree–Fock calculation. The dashed line corresponds to the criterion J_2

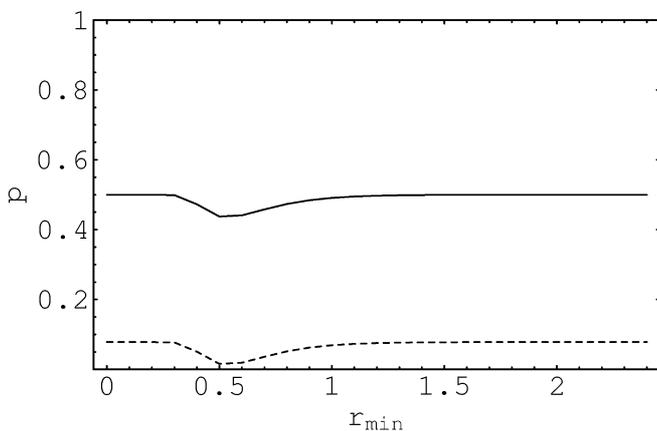


Fig. 2. The probability p_1 (solid line) of finding one electron in a spherical shell around the Be nucleus, having inner radius r_{\min} , and having on average one electron, from a Hartree–Fock calculation. The dashed line corresponds to the criterion J_1

2.2 Shape derivatives

The most efficient procedures to find out local maxima of a real-valued function $\Omega \rightarrow J(\Omega)$ with respect to the shape of Ω make use of the concept of the shape derivative that we formally introduce in the following. A more rigorous presentation of the theory of shape derivatives can be found in Ref. [16].

Consider a smooth domain $\Omega \subset \mathbb{R}^3$ and an infinitesimal deformation of it, denoted by Ω' , obtained by making the boundary $\partial\Omega$ of Ω move with the local velocity $v(x)\mathbf{n}(x)$ during an infinitesimal time step, δt , where $\mathbf{n}(x)$ denotes the outward pointing normal vector at $x \in \partial\Omega$ and v is a scalar field on $\partial\Omega$ (see Fig. 3).

In the special case when J is defined by

$$J(\Omega) = \int_{\Omega} f \tag{10}$$

where f is a given continuous integrable function on \mathbb{R}^3 , it is easy to check that

$$J(\Omega') = J(\Omega) + \delta t \int_{\partial\Omega} f(x)v(x)dx + o(\delta t) . \tag{11}$$

Returning to the general case, we will say that J is differentiable at Ω if there exists a real-valued function $\frac{\partial J}{\partial\Omega}$ defined on $\partial\Omega$ and called the shape derivative of J at Ω such that for any v

$$J(\Omega') = J(\Omega) + \delta t \int_{\partial\Omega} \frac{\partial J}{\partial\Omega}(x)v(x)dx + o(\delta t) .$$

The shape derivative plays a major role in shape optimization problems (just as the gradient does in standard optimization problems). Indeed, if the deformation field $v(x)$ is chosen equal to the shape derivative

$$\forall x \in \partial\Omega, \quad v(x) = \frac{\partial J}{\partial\Omega}(x) \tag{12}$$

one has

$$J(\Omega') = J(\Omega) + \delta t \left[\int_{\partial\Omega} \left(\frac{\partial J}{\partial\Omega}(x) \right)^2 dx \right] + o(\delta t) . \tag{13}$$

Equation (13) shows that for small δt , this deformation strictly increases the value of J , unless Ω is a critical point of J , characterized by the equation

$$\frac{\partial J}{\partial\Omega} = 0 \quad \text{on } \partial\Omega .$$

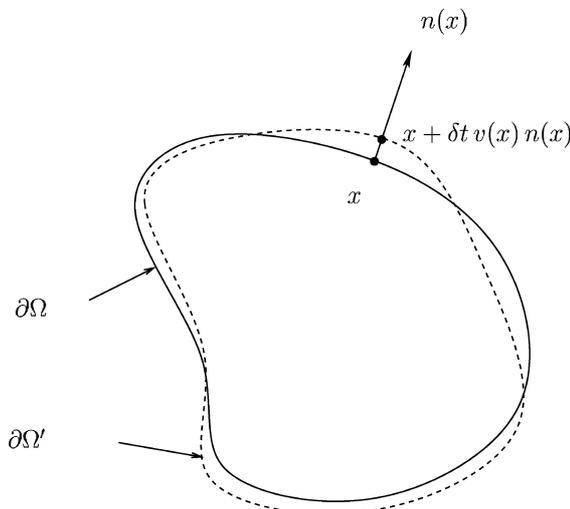


Fig. 3. Infinitesimal deformation of Ω obtained by making the boundary $\partial\Omega$ move with the local velocity $v(x)\mathbf{n}(x)$ during the time step δt

For finding out the maximizers of the criteria J_v defined by Eq. (2), one then has to compute the shape derivatives of the functions p_v and p_v^{ind} , let us start with the easier one, namely p_v^{ind} . Deriving Eq. (1), we get for $1 \leq v \leq N-1$,

$$\frac{\partial p_v^{\text{ind}}}{\partial \Omega}(x) = \binom{N}{v} b(\Omega)^{v-1} [1 - b(\Omega)]^{N-v-1}, \quad (14)$$

$$\{v[1 - b(\Omega)] + (N-v)b(\Omega)\} \frac{\partial b}{\partial \Omega}(x)$$

and

$$\frac{\partial p_0^{\text{ind}}}{\partial \Omega}(x) = N[1 - b(\Omega)]^{N-1} \frac{\partial b}{\partial \Omega}(x), \quad \frac{\partial p_N^{\text{ind}}}{\partial \Omega}(x) = Nb(\Omega)^{N-1} \frac{\partial b}{\partial \Omega}(x),$$

the shape derivative of b being obtained by deriving Eq. (9):

$$\frac{\partial b}{\partial \Omega}(x) = \frac{1}{N} \text{Tr} \left(\frac{\partial S}{\partial \Omega}(x) \right),$$

where we have denoted by $\frac{\partial S}{\partial \Omega}(x)$ the $N \times N$ symmetric matrix defined by

$$\left[\frac{\partial S}{\partial \Omega}(x) \right]_{ij} = \frac{\partial S_{ij}}{\partial \Omega}(x) = \sum_{\sigma} \phi_i(x, \sigma) \phi_j(x, \sigma) \quad (\text{see Eqs. 4, 10, 11}).$$

For computing the shape derivative of p_v , let us assume that all the eigenvalues of $\mathbf{S}(\Omega)$ are single (this is a technical assumption ensuring that the eigenvalues are differentiable; the final formulae Eqs. 15, 16, are still valid without this assumption). Using Eq. (5) and the chain rule, the shape derivative of $p_v(\Omega)$ is given for any $x \in \partial \Omega$ by

$$\frac{\partial p_v}{\partial \Omega}(x) = \sum_{l=1}^N a_v^{N,l} \frac{\partial \lambda_l}{\partial \Omega}(x),$$

where for each l the coefficients $a_v^{N,l}$ are computed using the recursion formulas Eq. (6) with

$$\forall k \neq l, \quad \beta_k = 1 - \alpha_k = \lambda_k(\Omega), \quad \alpha_l = -1, \quad \beta_l = 1.$$

Denoting by $X_l(\Omega)$ a normalized eigenvector of $\mathbf{S}(\Omega)$ associated with $\lambda_l(\Omega)$, we obtain by deriving the system

$$\begin{cases} \mathbf{S}(\Omega) \cdot X_l(\Omega) = \lambda_l(\Omega) X_l(\Omega) \\ X_l(\Omega)^T \cdot X_l(\Omega) = 1 \end{cases},$$

the following expression

$$\frac{\partial \lambda_l}{\partial \Omega}(x) = X_l(\Omega)^T \frac{\partial S}{\partial \Omega}(x) X_l(\Omega).$$

Therefore

$$\forall x \in \partial \Omega, \quad \frac{\partial p_v}{\partial \Omega}(x) = \text{Tr} \left(A_v(\Omega) \frac{\partial S}{\partial \Omega}(x) \right), \quad (15)$$

with

$$A_v(\Omega) = \sum_{l=1}^N a_v^{N,l} X_l(\Omega) X_l(\Omega)^T. \quad (16)$$

Specific formulas can of course be established for RHF or UHF wavefunctions.

2.3 Computation of the overlap matrix

In this section, we focus on the cases

- Of a linear molecule (the Cartesian coordinates are chosen such that the nuclei lie on the z -axis).
- Of a wavefunction Ψ originating from a RHF calculation performed in a Gaussian basis set.
- Of a domain Ω symmetric with respect to rotations around the axes of the molecule.

In order to be able to compute the probabilities $p_v(\Omega)$ and $p_v^{\text{ind}}(\Omega)$ as well as their shape derivatives $\frac{\partial p_v}{\partial \Omega}$ and $\frac{\partial p_v^{\text{ind}}}{\partial \Omega}$ along with the formulae set up in the previous two sections, one has to build (and then diagonalize) the overlap matrix $\mathbf{S}^R(\Omega)$ defined by Eq. (8). Presently, we use the following method, which is fast enough for the simulations we have performed up to now.

We denote by $\mathbf{C} = [C_{\mu i}]$ the matrix of the coefficients of the expansion of the molecular orbitals ϕ_i in the primitive Gaussian set $\{\chi_{\mu}\}_{1 \leq \mu \leq N_b}$:

$$\phi_i(x) = \sum_{\mu=1}^{N_b} C_{\mu i} \chi_{\mu}(x),$$

with

$$\chi_{\mu}(x, y, z) = A_{\mu} x^{n_{\mu}^x} y^{n_{\mu}^y} (z - z_{\mu})^{n_{\mu}^z} e^{-\alpha_{\mu}[x^2 + y^2 + (z - z_{\mu})^2]}.$$

Clearly,

$$\mathbf{S}^R(\Omega) = \mathbf{C}^T \mathbf{\Sigma}(\Omega) \mathbf{C},$$

where $\mathbf{\Sigma}(\Omega)$ is the overlap matrix of the primitive Gaussians

$$\Sigma_{\mu\nu}(\Omega) = \int_{\Omega} \chi_{\mu} \chi_{\nu}.$$

Let us consider a domain $\Omega \subset \mathbb{R}^3$ symmetric with respect to rotations around the z -axis, and let us denote by ω its trace on the half-plane $\{(r, \theta, z), r > 0, \theta = 0\}$ (Fig. 4).

By symmetry, $\Sigma_{\mu\nu} = 0$ if one at least of the two integers $n_{\mu}^x + n_{\nu}^x$ and $n_{\mu}^y + n_{\nu}^y$ is odd. In the remaining case ($n_{\mu}^x + n_{\nu}^x = 2p$ and $n_{\mu}^y + n_{\nu}^y = 2q$ with p and q in \mathbb{N}), one has

$$\Sigma_{\mu\nu}(\Omega) = A_{\mu} A_{\nu} W_{pq} \int_{\omega} r^{2p+2q+1} e^{-(\alpha_{\mu} + \alpha_{\nu})r^2} g_{\mu\nu}(z) dz,$$

with

$$W_{pq} = \int_0^{2\pi} (\cos \theta)^{2p} (\sin \theta)^{2q} d\theta$$

and

$$g_{\mu\nu} = (z - z_{\mu})^{n_{\mu}^z} (z - z_{\nu})^{n_{\nu}^z} e^{-\alpha_{\mu}[x^2 + y^2 + (z - z_{\mu})^2] - \alpha_{\nu}[x^2 + y^2 + (z - z_{\nu})^2]}.$$

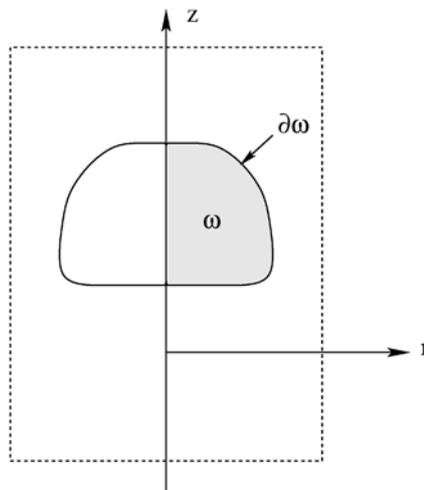


Fig. 4. A domain $\Omega \subset \mathbb{R}^3$ symmetric with respect to rotations around the z -axis and its trace ω on the half-plane $\{(r, \theta, z), r > 0, \theta = 0\}$

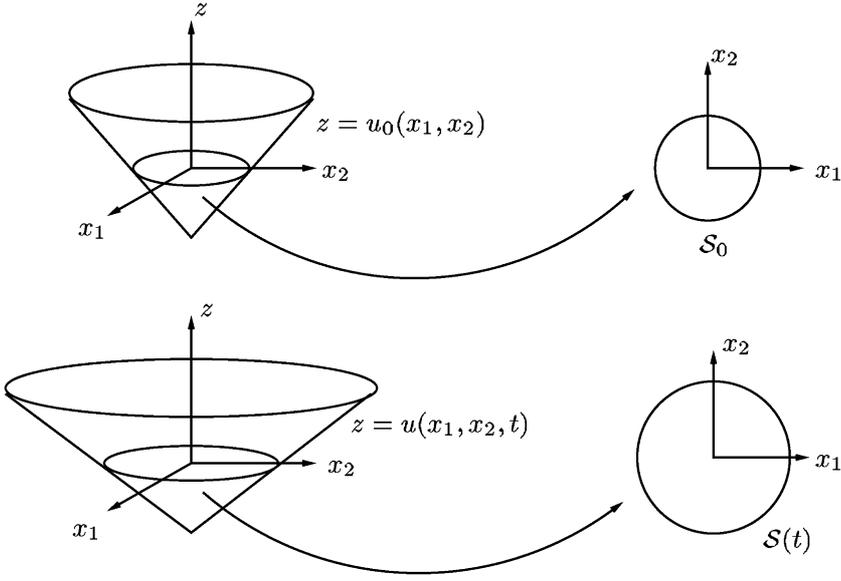


Fig. 5. The level set method: the evolution of a closed curve \mathcal{S} in \mathbb{R}^2 (right-hand side) is replaced by the evolution of a function u (left-hand side) such that the zero level set of the surface $\{z = u(x_1, x_2, t)\}$ is $\mathcal{S}(t)$

By integration by parts with respect to the variable r , one obtains

$$\Sigma_{\mu\nu}(\Omega) = A_{\mu}A_{\nu}W_{pq} \int_{\partial\omega} I_{p+q, z_{\mu}+z_{\nu}}(r)g_{\mu\nu}(z)(\mathbf{n}(r, z) \cdot \mathbf{e}_r)dl \quad (17)$$

where $\partial\omega$ denotes the boundary of ω , $\mathbf{n}(r, z)$ the outward pointing normal vector, \mathbf{e}_r the unit vector associated with the r variable and where

$$I_{n,z}(r) = \int_0^r s^{n+1}e^{-zs^2} ds = \frac{1}{2} \int_0^{r^2} t^n e^{-zt} dt .$$

Both W_{pq} and $I_{n,z}(r)$ can be computed analytically by recursion formulas. The overlap matrix $\Sigma_{\mu\nu}(\Omega)$ can therefore be easily computed by numerical integration on the boundary $\partial\omega$ (Eq. 17).

2.4 The level set method

Our optimization problem now reduces to the simulation of the evolution of some closed surface $\partial\Omega$ embedded in \mathbb{R}^3 (or of some closed curve $\partial\omega$ embedded in \mathbb{R}^2 in the setting described in Sect. 3) according to some given normal velocity field v given by Eq. (12). This section introduces the so-called level set method, which is commonly used in such simulations.

2.4.1 Principle

Methods of curve evolution for segmenting a domain into two or more regions have been extensively used, for example, in the computer vision domain where they were introduced by Kass et al.[17]. These evolutions were reformulated by Caselles et al.[18] in the context of curves and surfaces driven by partial differential equation (PDE). There is extensive literature that addresses the theoretical aspects of these PDEs and offers geometrical interpretations as well as results of existence and uniqueness [19, 20]. The level set method was first introduced by Osher and Sethian in Ref. [21] in the context of fluid mechanics and provides both a nice theoretical framework and efficient practical tools for simulating such motions. For a complete tour of the level set method, we refer the reader to Ref. [22, 23].

Let us consider a family of hypersurfaces $\mathcal{S}(t)$ in \mathbb{R}^k , where t is the time, that evolve according to the following dynamics:

$$\frac{\partial\mathcal{S}}{\partial t} = v\mathbf{n} \quad (18)$$

with initial condition $\mathcal{S}(t=0) = \mathcal{S}_0$, where \mathbf{n} is the unit normal vector of \mathcal{S} , v is a velocity field and \mathcal{S}_0 is some initial closed

hypersurface. For a generic molecular system, $k=3$, $\mathcal{S} = \partial\Omega$, v is given by Eq. (12) and \mathcal{S}_0 is some initial guess for $\partial\Omega$, whereas $k=2$ and $\mathcal{S} = \partial\omega$ (see Fig. 4) for the case of a linear molecule and a symmetric domain Ω .

In the level set method, the evolution Eq. (18) is achieved by means of an implicit representation of the surface $\mathcal{S}(t)$. The key idea is to introduce a function $u(x, t)$ from $\mathbb{R}^k \times \mathbb{R}^+$ to \mathbb{R} such that $\forall t \geq 0$, $\mathcal{S}(t) = \{x \in \mathbb{R}^k, u(x, t) = 0\}$.

Equations (18) and (19) are compatible if u is solution to the Hamilton–Jacobi equation:

$$\frac{\partial u}{\partial t} = \beta|\nabla u| \quad (20)$$

with initial conditions $u(x, 0) = u_0(x)$, where u_0 is some function from \mathbb{R}^k to \mathbb{R} such that $\mathcal{S}_0 = \{u_0 = 0\}$ and $\beta(x, t)$ is a suitable extension to $\mathbb{R}^k \times \mathbb{R}^+$ of the velocity field $v(x, t)$ defined only for $x \in \mathcal{S}(t)$. Such a compatibility condition can be established by differentiation of Eq. (19) and by making use of the relation $\mathbf{n} = \frac{\nabla u}{|\nabla u|}$ ¹. It has been proven that for a large class of functions β and u_0 , the zero level set at time t of the solution of Eq. (20) is the solution at time t of Eq. (18) (Fig. 5).

The main advantages of the level set approach with respect to other interface tracking methods are that it can be easily implemented in \mathbb{R}^2 or in \mathbb{R}^3 , with a stable and accurate scheme on a regular or adaptive grid, and that it can automatically deal with splitting or merging surfaces (Fig. 6).

2.4.2 Bells and whistles

Details on the practical implementation of the level set method can be found in Refs. [22, 23]. We will, however, give some hints for our particular case.

1. The function u_0 is most often chosen to be the signed distance function to the closed surface \mathcal{S}_0 (negative inside and positive outside \mathcal{S}_0). Note that at a further time step t , $u(\cdot, t)$ is no longer the distance function to $\mathcal{S}(t)$.
2. It is important to notice that $\beta(x, t)$ in Eq. (20) is defined in $\mathbb{R}^k \times \mathbb{R}^+$, whereas in the vector field $v(x, t)$ in Eq. (18) it is only defined for $x \in \mathcal{S}(t)$. The extension of $\beta(\cdot, t)$ from $\mathcal{S}(t)$ to the whole domain \mathbb{R}^k is a crucial point for the analysis and implementation of Eq. (20). There are mainly two ways of doing this.

¹ Usually, u is chosen negative inside \mathcal{S} (in Ω), and positive outside (in $\mathbb{R}^3 \setminus \bar{\Omega}$). As a result, $\frac{\nabla u}{|\nabla u|}$ is the outward normal vector

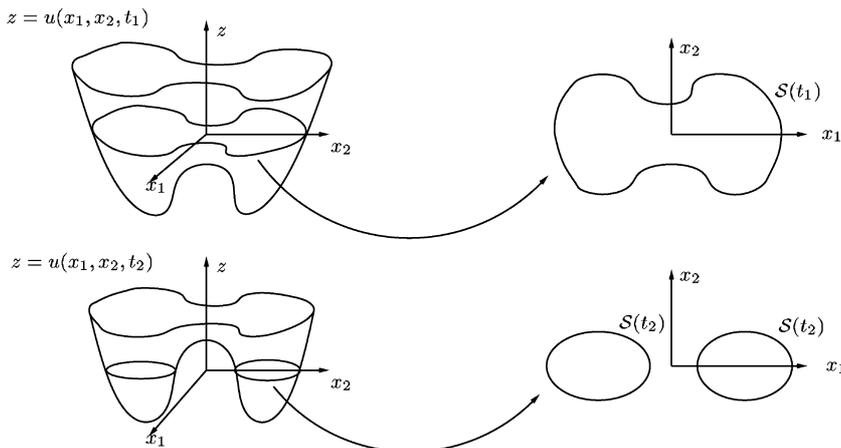


Fig. 6. Changes of topology for \mathcal{S} are handled automatically: the zero level set of $\{z = u(x_1, x_2, t)\}$ can split or merge without any special treatment

- Most of the time this extension is natural. A classical example is the so-called mean curvature motion, where $v = H_{\mathcal{S}}$, the mean curvature of \mathcal{S} in Eq. (18). In this case, one can choose $\beta = H_u$, the mean curvature of the level set of u passing through x in Eq. (20).
 - In some cases, like ours, this extension is not possible; indeed, although the “natural” extension consisting in choosing $\beta(x, t) = f(x)$ for criteria of the form of Eq. (10) can be used to extend the term $\frac{\partial \mathcal{S}}{\partial t}$ appearing in Eqs. (14) and (15) to the whole space, this possibility has to be discarded: numerical tests have shown that it leads to convergence towards global maxima of J_v (typically the core shells of the heavy atoms), whereas we are mostly interested in local minima (lone pairs or bonding electrons). Then one may assign to $\beta(x, t)$ in Eq. (20) the value of $v(y, t)$ in Eq. (18) where y is the closest point to x belonging to $\mathcal{S}(t)$. The problem of this extension of the velocity has been broadly studied [24, 25]. We use the solution given in Ref. [26].
3. In the particular symmetric case described in the previous section, only a plane curve has to be considered. Yet, as previously mentioned, a 3D implementation of the level set method is straightforward. In that case, the algorithmic complexity of the method is generally lowered thanks to the so-called narrow-band implementation [27], where only the neighbors of the zero level on the regular grid are updated.

3 Analysis of some simple systems

3.1 The Be atom

It turns out that, in general, for a given v , it is possible to find several maxima of J_v , each having its own significance. Taking again the example of the Hartree–Fock Be atom one can find with p_2 or J_2 a maximum corresponding to the core, and another one for the valence Fig. 7.

3.2 The BH molecule

In order to illustrate the chemical relevance of using J_v , we consider the simple case of the BH molecule, which presents a core, around the B nucleus, a BH bond, and a lone pair on the opposite side (the Lewis structure is :B:H).

A 2D level set algorithm used to optimize the criterion J_2 with respect to the contour $\partial\omega$ for three different initial guesses (top left in Fig. 8) corresponding vaguely

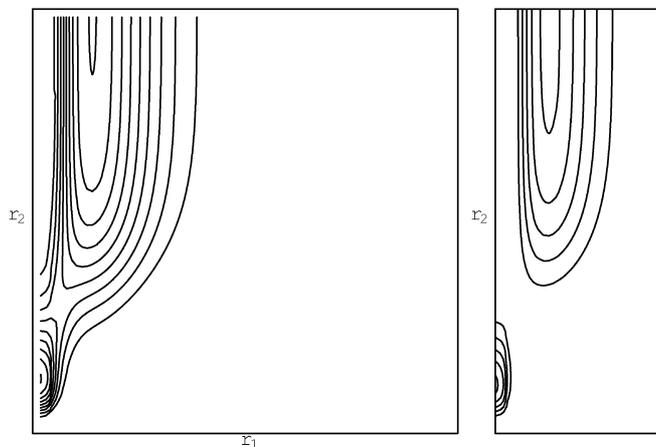


Fig. 7. p_2 (left) and J_2 (right) for the Hartree–Fock Be atom in a spherical shell between r_1 and r_2 ; contours from 0.1 to 0.9 in steps of 0.1. There are two maxima, one for $r_1 = 0$ and $r_2 \simeq 1.0$, the other for $r_1 \simeq 1.0$ and $r_2 = +\infty$

to the expected location of the B core shell, the BH bond and the lone pair, respectively. The calculation was performed on a uniform space grid (length step of 0.1 bohr) with an adaptive time step chosen between 0.1 and 1.0. The bottom-right picture on Fig. 8 shows the contours obtained after 300 time steps, which correspond to the time t_N when the convergence criterion $0 \leq J_2[\Omega(t_N)] - J_2[\Omega(t_{N-1})] \leq 10^{-4}$ is reached for the three contours. The top-right and the bottom-left pictures show intermediate positions of the contours at the 100th and at the 200th time step respectively. The optimization of the three contours takes about 20 min of computation time on a Pentium III 800 MHz personal computer, but this computational time can be significantly reduced by using adaptive space and time grids [22] and narrow-band [27] strategies. Current work in this direction is in progress.

In the isolated B atom, integrating the density inside a sphere of radius of around 0.70 bohr, one obtains a population of two electrons; J_2 is maximized for a radius of around 0.72 bohr [1]. In BH, one obtains a maximum of J_2 (approximately equal to 0.53), for a deformed sphere around the B nucleus, having roughly the same size as the core of the free B atom. One notices a small

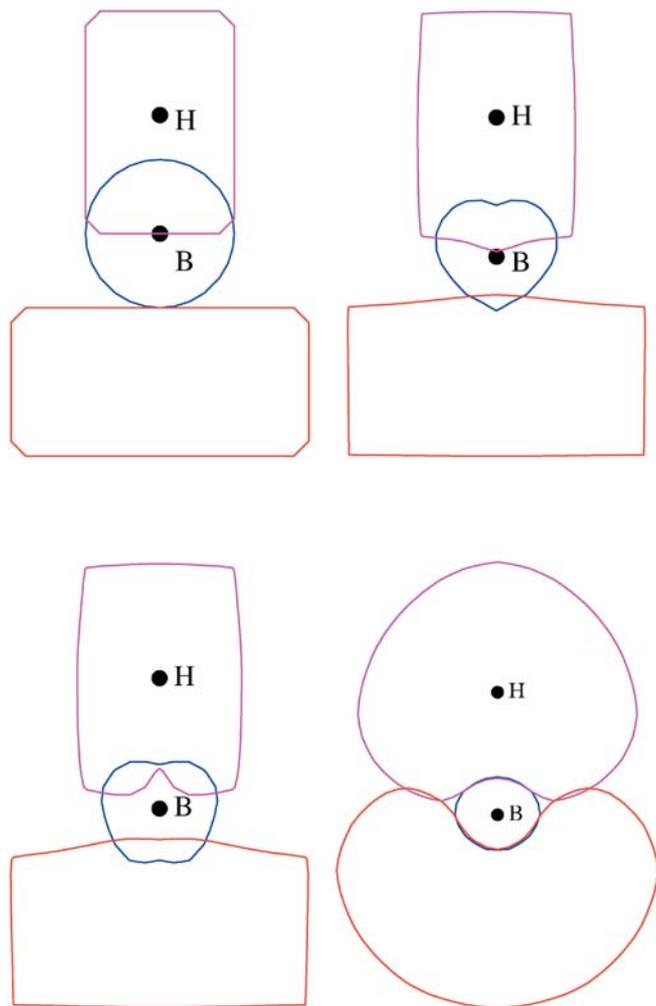


Fig. 8. Optimization of J_2 for three different initial contours (*top left*). The pictures report the positions of the three contours after 100 time steps (*top right*), 200 time steps (*bottom left*) and at convergence, after 300 time steps (*bottom right*). The nearly spherical region around the B nucleus is associated with the B core, the region above to the B lone pair, and that below to the BH bond

deformation, due to the “pressure” of the valence electron pairs.

Two more local maxima of J_2 are found ($J_2 \simeq 0.42$, and 0.43 , respectively). Although sought independently, they practically divide the space not occupied by the core, as can be seen in the bottom-right picture in Fig. 8. The first domain can be attributed to the B lone pair (lower part), the other to the BH bonding pair (upper part).

4 Perspectives

In our experience with atoms and molecules, optimizing J_v or related quantities gives better results for dividing space into chemically significant regions than other existing methods. We have shown here just a few examples, the Be atom and the BH molecule, where the level set method proved to be an efficient tool to find these regions. Although optimized independently, we

find that electron pairs in fact divide the space among themselves, in accordance with chemical intuition: electron pairs guard the space they occupy.

We are now developing the 3D extension of the code to be able to study any (not-necessarily linear) molecular system. Progress is also being made in the extension to correlated wavefunctions produced by configuration interaction, multiconfigurational self-consistent-field or quantum Monte Carlo calculations). A further immediate extension can be made by considering spins separately. For example, we can consider not having two electrons in a given region, but having simultaneously one with spin up, and one with spin down. In the future, we also intend to investigate correlations between regions, cases where v_1 electrons are in one region, and v_2 in another.

Appendix

Let us establish Eq. (3). When Ψ is a Slater determinant,

$$\begin{aligned}
 |\Psi|^2 &= \left| \frac{1}{\sqrt{N!}} \det [\phi_i(x_j, \sigma_j)] \right|^2 \\
 &= \frac{1}{N!} \det \left[\sum_{k=1}^N \phi_i(x_k, \sigma_k) \phi_j(x_k, \sigma_k) \right] \\
 &= \frac{1}{N!} \sum_{k_1, \dots, k_N=1}^N \det [\phi_i(x_{k_j}, \sigma_{k_j}) \phi_j(x_{k_j}, \sigma_{k_j})] \quad (21)
 \end{aligned}$$

$$= \frac{1}{N!} \sum_{p \in \mathcal{S}_N} \det [\phi_i(x_{p(j)}, \sigma_{p(j)}) \phi_j(x_{p(j)}, \sigma_{p(j)})] \quad (22)$$

where \mathcal{S}_N denotes the group of the permutations of $\llbracket 1, N \rrbracket$. To pass from Eq. (21) to Eq. (22) it suffices to remark that the determinant vanishes if $k_l = k_m$ for some $l \neq m$. Therefore

$$p_v(\Omega) = \frac{1}{v!(N-v)!} \sum_{p \in \mathcal{S}_N} \det [\tilde{S}_{ij,p}(\Omega)] \quad ,$$

with $\tilde{S}_{ij,p}(\Omega)$ equal to $S_{ij}(\Omega)$ if $p(j) \in \llbracket 1, v \rrbracket$ and to $\delta_{ij} - S_{ij}(\Omega)$ otherwise. Thus $p_v(\Omega)$ is the sum of all the determinants obtained by selecting v columns of $\mathbf{S}(\Omega)$ and by replacing the remaining $N - v$ columns by the corresponding columns of $\mathbf{I}_N - \mathbf{S}(\Omega)$. Equation (3) follows.

References

1. Savin A (2002) Sen KD (ed) In: Reviews of modern quantum chemistry. A celebration of the contributions of Robert G Parr. World Scientific, Singapore, p 43
2. Politzer P, Parr RG (1976) J Chem Phys 64: 4634
3. Bader RFW (2002) Atoms in molecules: a quantum theory. Oxford University Press, Oxford
4. Sagar RP, Ku ACT, Smith VH Jr (1988) J Chem Phys 88: 4367
5. Shi Z, Boyd RJ (1988) J Chem Phys 88: 4375
6. Kohout M, Savin A, Preuss H (1991) J Chem Phys 95: 1928
7. Becke AD, Edgecombe KE (1990) J Chem Phys 92: 5397
8. Kohout M, Savin A (1996) Int J Quantum Chem 60: 875
9. Aslangul C, Constanciel R, Daudel R, Kottis P (1972) Adv Quantum Chem 6: 93

10. Fulde P (1991) *Electron correlations in molecules and solids*. Springer, Berlin Herdberg New York
11. Pfirsch F, Böhn MC, Fulde P (1985) *Z Phys B* 60: 171
12. Ziesche P (1996) *Int J Quantum Chem* 30: 1361
13. Mödl M, Dolg M, Fulde P, Stoll H (1996) *J Chem Phys* 105: 2353
14. Schautz F, Flad H-J, Dolg M (1998) *Theor Chem Acc* 99: 231
15. Ziesche P, Tao J, Seidl M, Perdew JP (2000) *Int J Quantum Chem* 77: 819
16. Sokolowski J, Zolesio JP (1992) *Introduction to shape optimization: shape sensitivity analysis*, Springer series in computational mathematics, vol 10. Spinger, Berlin Herdberg New York
17. Kass M, Witkin A, Terzopoulos D (1988) In: *Proceedings of the 2nd international conference on computer vision*, vol 1. IEEE Computer Society, New york, p 321
18. Caselles V, Kimmel R, Sapiro G (1995) In: *Proceedings of the 5th international conference on computer vision*. IEEE Computer Society, New york, p 694
19. Gage M, Hamilton RS (1986) *J Diff Geom* 23: 69
20. Chen YG, Giga Y, Goto S (1991) *J Diff Geom* 33: 749
21. Osher S, Sethian J (1988) *J Comput Phys* 79: 12
22. Sethian JA (1999) *Level set methods and fast marching methods: evolving interfaces in computational geometry, fluid mechanics, computer vision, and materials sciences*. Cambridge University Press, Cambridge
23. Osher S (2002) *The level set method: applications to imaging science*. Technical report 02-43, University of California at Los Angeles, Computational and Applied Mathematics, July 2002
24. Adalsteinsson D, Sethian JA (1999) *J Comput Phys* 148: 2
25. Zhao HK, Chan T, Merriman B, Osher S (1996) *J Comput Phys* 127: 179
26. Gomes J, Faugeras O (2000) *J Visual Commun Image Representation* 11: 209
27. Adalsteinsson D, Sethian JA (1995) *J Comput Phys* 118: 269