

How Malonaldehyde Bonds Change during Proton Transfer

X. Krokidis, V. Goncalves, A. Savin, and B. Silvi*

Laboratoire de Chimie Théorique (UMR-CNRS 7616), Université Pierre et Marie Curie, 4 Place Jussieu, 75252 Paris Cedex, France

Received: October 23, 1997; In Final Form: March 11, 1998

The proton transfer mechanism in malonaldehyde has been investigated from the points of view of the topological analysis of the electron localization function and of the catastrophe theory. The calculations were carried out at the B3LYP hybrid functional level with the ccp-VTZ basis set. Complementary calculations were performed at the HF, MP2, BLYP, and B3LYP levels with the STO-3G, 6-31G, and 6-31G** basis sets in order to discuss the stability of the method. The mechanism of the reaction does not involve any reorganization of the skeleton: it is localized in the hydrogen bond region and involves three chemical structures corresponding to the reactant, the transition state, and the product. It is shown that the proton detachment is in fact the covalent breaking of the OH bond, and by symmetry its attachment on the other oxygen is also a covalent process. The analysis of the catastrophe shows that the reaction is driven by two active coordinates related to the OH and OO distance and provides a method for the determination of the limiting paths.

1. Introduction

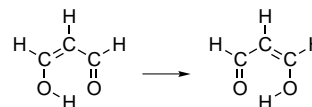
Proton transfer reactions in hydrogen-bonded systems constitute a wide class of processes which have been extensively studied from both experimental and theoretical viewpoints^{1–4} owing to their importance in many fields of physics, chemistry, and biochemistry. These studies are generally focused on the energetics and on the dynamics of the process, and on the discussion of the driving force of the proton transfer (i.e., tunneling or Renner–Teller effect). The description of the proton transfer in terms of a reorganization of the chemical bonds during the reaction has received less attention. Such a description of chemical processes is nevertheless important for those interested in identifying “chemically defined” intermediates in order to get a better understanding of the reaction mechanism. The proton transfer involves a rather straightforward detachment/attachment mechanism, and therefore a thorough analysis of the bonding is not expected to yield any new salient physical insights. However, the simplicity of the mechanism appears to be an advantage for methodology-oriented investigations in which well-understood cases are treated in order to check the applicability of a given technique and to identify its strengths and weaknesses.

We have recently proposed a method of characterization of elementary chemical processes⁵ which relies on the topology of the Becke and Edgecombe localization function ELF.⁶ This method was first applied to the inversion of NH₃, to the breaking of the ethane C–C bond, and to the dissociation of BH₃NH₃.⁵ The main objective of the method (i.e., to provide a chemically revealing picture of the reaction) was reached on these few latter examples. In the present work, the choice of the reaction, the malonaldehyde proton transfer, implies an additional complexity due to a possible synchronous evolution of the delocalized bonds of the skeleton.

Malonaldehyde is considered as a model system for molecules exhibiting a strong intramolecular hydrogen bonding. The intramolecular proton transfer in the ground and lowest excited states of malonaldehyde was recently thoroughly investigated by Barone and Adamo⁷ who concluded that the energetic

barrier of the proton transfer is lower for the ground states than for the two lowest triplets.

In the conventional Lewis picture of the reaction shown below, the mechanism involves not only the transfer of the proton but also a counterclockwise migration of the double bonds.



However, if one takes into account the different resonant structures of malonaldehyde, the reaction mechanism may be somehow different.

The paper is organized as follows: in section 1, the main lines of the method of analysis are sketched, and section 2 provides the computational details of the calculation and presents the analysis of the actual case. Mathematical details are given in the Appendixes.

1.1. Method of Analysis. The method outlined below considers a chemical reaction as a succession of definite molecular structures located along the reaction path or in its neighborhood. It aims to identify such structures, to characterize the transitions between them, and finally to propose a nonempirical reaction mechanism. Neither the chemical bond nor molecular structures are observables. These intuitive concepts belong to a system of representation settled by the chemists on the basis of experience. The theory of chemical bond essentially relies on a local description of the matter in which atoms are linked by electron pairs⁸ to form molecules within which it is possible to identify chemically significant regions. Quantum mechanics is the physical theory which fully explains the stability and the microscopic properties of molecules. It provides a nonlocal description which stems from the wave particle formulation, from the statistical interpretation, and more practically from the impossibility of partitioning the molecular Hamiltonian into physically well-defined atomic and bonding (interatomic) contributions.

Several methods may be used to characterize the chemical bond from quantum chemical results. A first group of methods attempts to localize the molecular orbitals following different schemes based on the extremization of an orbital related property.^{9–11} The choice of the localization procedure is rather arbitrary. Moreover, as pointed out by England,¹² a given orbital localization scheme sometimes yields multiple solutions corresponding to local extrema of the property. The second group is based on projection techniques in which the wave function is developed on basis functions located on nuclei. These methods yield an analysis in terms of bond orders.^{13–19} Though they also contain some arbitrariness, these techniques are available in most *ab initio* codes and very popular. As a general rule the output is very sensitive to the quality of the basis set. The most stable results are obtained with natural population analysis (NPA).¹⁹ However, as mentioned by Bachrach in a critical review on the population analysis methods,²⁰ NPA is not free of drawbacks since “it rests on allocating electrons through an orbital occupancy basis, not in terms of the actual locations of electrons”. Moreover, bond orders provide a picture of the molecular structure in which any pair of atoms is (at least very weakly) bonded.

The deformation electron density is defined by $\Delta\rho(\mathbf{r}) = \rho_{\text{molecule}}(\mathbf{r}) - \sum_A \rho_A(\mathbf{r})$ in which $\rho_A(\mathbf{r})$ denotes the density of the promolecules (atoms) labeled by A. It is expected to have its maxima in the bonding regions. However, the absence of significant accumulation of electron density in interatomic regions has been reported for a number of covalent bonds.²¹ As pointed out by Schwarz et al.^{22–24} and by Kunze and Hall,²⁵ the choice of a spherical-atom promolecule is responsible for a negligible accumulation which is revealed by oriented-atom promolecules. The buildup of nonspherical promolecule electron densities can be achieved by the technique described by Ruedenberg and Schwarz.²⁶

The topology of the total density $\rho(\mathbf{r})$ has been investigated in detail by Richard Bader and co-workers within the framework of the theory of “atoms in molecules”.²⁷ The analysis of the charge density gradient field yields a partition of the molecular space into basins of attractors. The attractors (i.e., the local maxima of the charge density) are usually located on the nuclei; exceptions have been however evidenced for metal clusters.²⁸ The critical points of the electron density gradient field, those points at which $\nabla\rho(\mathbf{r}) = 0$, allow to define bond paths which form a network called the molecular graph.²⁹ In the topological theory, the nuclear space configuration is partitioned into a finite number of structural regions. Within each region, the molecular graphs are equivalent, and they form an equivalence class called the molecular structure.²⁹ The molecular structure depicted by molecular graphs generally agrees with the expectation of chemical intuition, although unexpected bond paths have been reported for a few systems.^{30–32} It should be therefore possible to exploit this theory to follow the evolution of the molecular structure during a reaction. This has been successfully done for a few cases such as isomerizations,³³ ring and cage formations,³⁴ and exceptional situations characterized by a nonnuclear attractor.^{35–37}

However, the density-based molecular structure theory does not allow to study reactions as simple as the dissociation of a homolitic diatomic because the nuclear configuration space corresponds to a single structural region. It is thus necessary to investigate in the same spirit other local functions (hereafter referred to as localization functions) in order to characterize bonds and lone pairs in molecules rather than atoms in molecules. Among the possible localization functions we can mention the

Laplacian of the charge density,^{38,39} the Hunter potential,⁴⁰ the function of Luken and Culberson,⁴¹ and the function of Becke and Edgecombe (ELF).⁶ It has been shown that the Laplacian of the charge density and ELF perform partitions of the molecular space which are consistent with the Lewis description and with the VSEPR model.^{39,42–46}

The topology of the charge density Laplacina has been studied^{38,39,47–50} as a part of Bader’s theory of atoms in molecules in which attention is mostly focused on the valence-shell charge concentration. The topological description of the chemical bond proposed by Silvi and Savin⁴⁵ relies upon the gradient field analysis of the electron localization function (ELF) of Becke and Edgecombe.⁶ This function, hereafter denoted as $\eta(\mathbf{r})$, is expressed as

$$\eta(\mathbf{r}) = \frac{1}{1 + \left(\frac{D(\mathbf{r})}{D_h(\mathbf{r})}\right)^2} \quad (1)$$

in which $D(\mathbf{r})$ and $D_h(\mathbf{r})$ are respectively the excess local kinetic energy due to the Pauli repulsion⁵¹

$$D(\mathbf{r}) = T_s(\mathbf{r}) - T_w(\mathbf{r}) \quad (2)$$

and $D_h(\mathbf{r})$ is the kinetic energy of the homogeneous electron gas of density $\rho(\mathbf{r})$. The excess local kinetic energy defined by eq 2 is the difference between the definite positive kinetic energy density $T_s(\mathbf{r}) = \frac{1}{2}\nabla\nabla'\rho(\mathbf{r},\mathbf{r}')|_{\mathbf{r}'=\mathbf{r}}$ and the von Weizsäcker⁵² functional $T_w(\mathbf{r}) = \frac{1}{8}\nabla\rho(\mathbf{r},\mathbf{r}')|^2/\rho(\mathbf{r},\mathbf{r})$, where $\rho(\mathbf{r},\mathbf{r}')$ is the first-order density matrix. The excess local kinetic energy is a positive quantity⁵³ which is small in the regions of space where the electrons do not experience the Pauli repulsion and therefore where parallel spin electrons are far from one another whereas it is large where they are close to one another. In such regions ELF is close to 1 and 0, respectively. The topological analysis of the gradient field of $\eta(\mathbf{r})$ assigns a region of space, the basin, to each local maximum of $\eta(\mathbf{r})$ providing a partition of the molecular space analogous to that made in hydrography to define river basins and watersheds. There are basically two types of valence basins which have a clear chemical meaning. On the one hand are core basins encompassing nuclei (with $Z > 2$) and valence basins in the remaining space. Each valence basin is characterized by its *synaptic order* σ , which is the number of core basins with which it shares a common boundary.⁵⁴ In this representation the monosynaptic basins correspond to the nonbonded pairs of the usual Lewis picture whereas the di- and polysynaptic ones are related to bonds. It is worthy of note that the possibility of multicentric bonds is clearly accounted for by the concept of polysynaptic basin. For a given system in a given electronic state there is therefore a 1:1 correspondence (in other words an isomorphism) between the topological description and the Lewis picture which gives access to a mathematical model to study the evolution of this latter. Integrations of the one-electron density $\rho(\mathbf{r})$ and of the pair function $\pi(\mathbf{r}_1,\mathbf{r}_2)$ over the volume Ω_i of the i th basin provide the basin population \bar{N}_i :

$$\bar{N}_i = \langle N \rangle_i = \int_{\Omega_i} \rho(\mathbf{r}) \, d\mathbf{r} \quad (3)$$

and its variance

$$\sigma_i^2 = \langle N^2 \rangle_i - \langle N \rangle_i^2 = \int_{\Omega_i} \int_{\Omega_i} \pi(\mathbf{r}_1, \mathbf{r}_2) \, d\mathbf{r}_1 \, d\mathbf{r}_2 - \bar{N}_i(\bar{N}_i - 1) \quad (4)$$

Along a reaction pathway, the nuclear configuration changes and the bonding, represented by the molecular structure, evolves accordingly: some bonds are broken, and some others are formed. In this case, the location function varies parametrically with nuclear coordinates which form the set of control parameters. In the topological model, the transformation of the bonding finds expression in the appearance and disappearance of local maxima of the localization function. Thom's catastrophe theory⁵⁵ is the mathematical theory that enables the study of the topological changes with respect to the variation of the control space parameters. A summary of the topological concepts and of the important mathematical properties of the model is given in Appendix A. The catastrophe theory has been first applied to the analysis of the electron charge density of reacting systems^{33–36} and more recently extended to the ELF function.⁵ These methods identify the subsets of the nuclear configuration space corresponding to defined chemical structures and locate their boundaries which are the sets of turning points between them. They enable us to precisely describe the different steps of a reaction and therefore its mechanism. Moreover, the mathematical properties of the model provide additional information which is not directly accessible from the sole chemical considerations; in particular they point out the importance of topological constraints.

2. Results and Discussion

2.1. Computational Methods. The *ab initio* calculations have been performed using the Gaussian94 software.⁵⁶ The reaction path has been scanned by the IRC method⁵⁷ at the hybrid Hartree–Fock density functional level B3LYP^{58–61} with the ccp-VTZ basis set.^{62,63} The choice of the B3LYP method and of the ccp-VTZ basis set was done in order to get reliable optimized structural parameters and a realistic value for the proton transfer energetical barrier. The proton transfer in malonaldehyde was previously studied by *ab initio* techniques from the energetical point of view by several authors^{7,64–67} who have shown that electron correlation is necessary to obtain a realistic value of the proton transfer energy barrier. Experimentally, this barrier is estimated to be of the order of 4.0–5.2 kcal/mol.⁶⁸ Barone and Adamo⁷ have studied the dependence of the proton transfer barrier upon the basis set and the correlation technique. They showed that the better agreement with experiment was obtained by CCSD(T)/DZP calculations. Their calculations also indicated that the B3LYP hybrid functional⁵⁸ is an alternative good choice when used with a large enough basis set (typically polarized valence triple- ζ).

The ccp-VTZ basis set used in the present calculations includes two sets of p and one set of d functions on the hydrogens and two sets of d and one set of f on the carbon and oxygen centers and therefore it is expected to yield results at least as good as those reported in ref 7. The optimized structural parameters of the equilibrium configuration and of the transition state are almost identical to those listed in the B3LYP/TZ2P entries of the first table of Barone and Adamo's article⁷ and therefore have not been reported here. The main structural discrepancy between *ab initio* optimized structures and experimental ones^{68–70} occurs for the $C_1C_2H_2$ (see Figure 1 for atom labels) bond angle which is calculated 8° more narrow than observed. However, the experimental uncertainty may be quite large since the authors of the experimental paper⁷⁰ warn us: "this value is probably not accurate because of the position of C_2 and H_2 close to the axis". For the transition state the calculated dipole moment, 2.601 D, lies within the experimental range of 2.591–2.603 D.⁷⁰

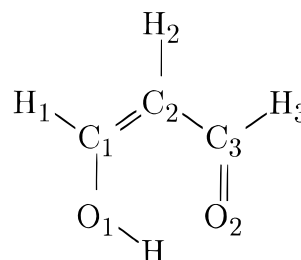


Figure 1. Labels of atoms.

In order to investigate the stability of the method with respect to the nature of the approximate wave functions, several calculations were carried out within the Hartree–Fock, MP2, BLYP,^{59–61} and B3LYP⁵⁸ schemes with the standard STO-3G,⁷¹ 6-31G, and 6-31G**.^{72,73}

The topological analysis of the ELF function was carried out with the TopMod⁷⁴ series of program developed in our laboratory.

2.2. Topological Analysis of the Equilibrium and Transition States. The chemical structure of the malonaldehyde molecule in its equilibrium geometry as displayed by the topology of the ELF function is weakly dependent upon basis set as well as calculation scheme effects. Minimal basis sets, such as STO-3G, yield a picture of the bonding in which the C_1C_2 and C_3O_2 bonds give rise to a pair of localization basins symmetrically located on both sides of the molecular plane. The sum of the $V(C_1, C_2)$ is calculated to be 4.1 e^- whereas that of the $V(C_3, O_2)$ one is only 2.2 because of the large transfer from this basin towards the lone pairs of O_2 . The STO-3G basis set does not take into account the possible delocalization of the bonds in the skeleton which is revealed as soon as split valence or better basis sets are used. With both 6-31G, 6-31G** and ccp-VTZ and at any level of calculation considered (HF, MP2, BLYP, B3LYP) there is only one basin for the C_1C_2 bond and another one for C_3O_2 . The population of the former basin decreases as the basis set is enlarged: 3.68, 3.47, and 3.19 e^- for 6-31G, 6-31G**, and ccp-VTZ, respectively. The C_3O_2 and C_2C_3 basin populations vary in the opposite direction. The evolution of the basin population with respect to the size of the basis set correlates with that of the optimized bond lengths: for instance, the optimized values of the C_1C_2 distance are 1.334, 1.343, and 1.362 Å for 6-31G, 6-31G**, and ccp-VTZ, respectively. It is worth noting that the basin populations corresponding to the CH and OH bonds are almost insensitive to basis set effects. In fact, these results show that the actual system rather than the method requires a careful choice of basis set. Once a basis set retained, the results are very stable with respect a change of method of calculation: the differences in the values of a given basin populations are at most of the order of 0.1 e^- . In the following discussion, we will only consider the ccp-VTZ/B3LYP results.

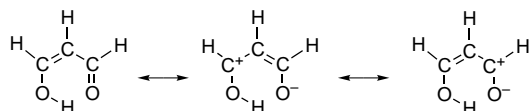
Several mechanisms may be invoked to describe and interpret the proton transfer. In a purely ionic process, the motion of the proton is expected to leave the number of basin constant. In the initial state the proton is bound to O_1 and is located in the $V(O_1, H)$ basin which has a boundary with one of the O_2 lone pairs, more precisely with the $V_2(O_2)$ basin. In the final state, the proton has left $V(O_1, H)$ which has become $V_2(O_1)$ accordingly and has entered $V_2(O_2)$ now transformed into $V(O_2, H)$. In the ionic mechanism, the proton is expected to pass across the separatrix of the two basins. In a covalent process, the O_1H bond should be broken and the O_2H one formed. This can be achieved either in one step, in which case the bond dissociation and the bond formation occurring simul-

TABLE 1: Populations (\bar{N}) and Population Standard Deviations of the Malonaldehyde Basins in Equilibrium and Transition State Structures

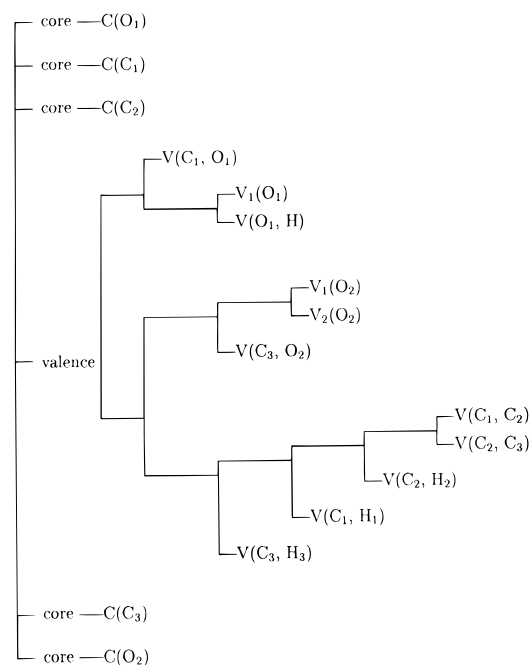
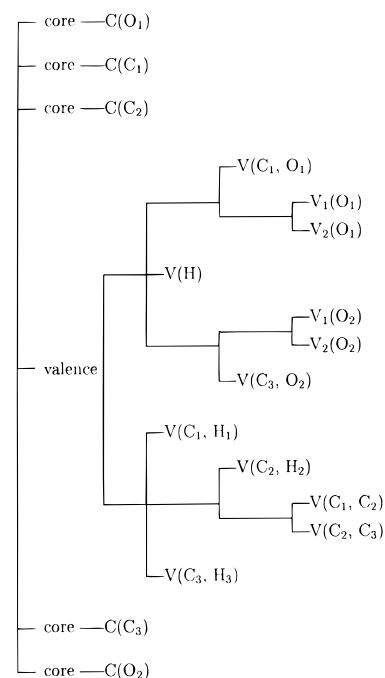
basins	equilibrium structure		transition state	
	\bar{N}	σ	\bar{N}	σ
V(C ₁ ,C ₂)	3.19	1.19	2.80	1.15
V(C ₂ ,C ₃)	2.42	1.08	2.80	1.15
V(C ₁ ,O ₁)	1.81	1.04	2.07	1.09
V ₁ (O ₁)	4.04	1.20	3.44	1.17
V ₂ (O ₁)			1.85	1.01
V(O ₁ ,H)	1.85	0.95		
V(H)			0.44	0.60
V(C ₃ ,O ₂)	2.28	1.13	2.07	1.09
V ₁ (O ₂)	2.81	1.10	3.44	1.17
V ₂ (O ₂)	2.38	1.07	1.85	1.01

taneously there is no variation of the number of basins, or in two steps with an intermediate chemical structure (that of the transition state) having one basin more than the reactant. A piece of information is brought by the analysis of the transition state which presents a V(H) basin. Therefore, the proton transfer can be described as a two-step process involving the breaking of a covalent OH bond and the formation of another one. This mechanism may be explained by the particular nature of the proton as a cation. Any cation except the proton is mostly a closed-shell structure encompassing nuclei with a positive net charge. The Pauli repulsion, due to the external closed shell, precludes the formation of a valence shell around the nucleus. In the case of the proton there is not such a closed-shell structure, and the net effect on the electron cloud is an accumulation of density around the proton due to its attractive Coulomb potential. Moreover, as the density around the proton is rather large it is energetically advantageous that the Pauli repulsion should be small in this region. However, the mechanism is not fully covalent as testified by the V(H) basin population which is 0.44 e⁻ in the transition state.

The populations of valence basins of malonaldehyde which noticeably vary during the proton transfer are listed in Table 1 for the equilibrium and transition state structures together with their standard deviations. Figures 2 and 3 display the localization domain reduction tree diagrams of these two structures. In the equilibrium structure, the localization reduction⁵⁴ of the valence domains first splits those related to O₁, namely V₁(O₁), V(C₁,O₁), and V(O₁,H), from the remaining valence domains indicating that O₁ is the most electronegative center. O₁ preserves its atomic shell structure up to rather high values of $\eta(\mathbf{r})$ and therefore the two lone pairs are gathered in a single monosynaptic basin V₁(O₁) population amounts to 4.04 e⁻, with a rather large standard deviation consistent with a rather large delocalization involving the adjacent V(C₁,O₁) and V(O₁,H) basins. The second separation involves the domains around O₂. The sum of the monosynaptic basin populations $\bar{N}(V_1(O_2)) + \bar{N}(V_2(O_2)) = 5.19$ e⁻ whereas $\bar{N}(V(C_3,O_2)) = 2.28$ e⁻ which is close to a C⁺-O⁻ picture of the bond. Finally, there is also a rather large delocalization around C₂ involving the V(C₁,C₂) and V(C₂,C₃) disynaptic basins which explains why no double-bond picture arises from the analysis. The numbers provided by the basin populations are consistent with the following resonance between three Lewis structures:



with 0.2, 0.3, and 0.5 as the respective weights.

**Figure 2.** Localization domain reduction tree diagram of malonaldehyde in its equilibrium structure.**Figure 3.** Localization domain reduction tree diagram of malonaldehyde in the transition state structure.

For the transition state, the localization domains around the oxygens are well separated from those involving the carbons. Each oxygen has now two monosynaptic basins for the lone pairs with a total population equal to 5.29 e⁻. The protonated monosynaptic basin contains 0.44 e⁻. On the carbon skeleton side, the V(C₁,C₂) and V(C₂,C₃) populations are both equal to 2.80 e⁻ and those of V(C₁,O₁) and V(C₃,O₂) to 2.07 e⁻ (i.e., close to the average of the equilibrium structure populations). It is worth noting that the chemical structure of the transition state involves a "dressed" proton.

2.3. Bonding Evolution along the Reaction Path. The reaction involves three chemical structures along the reaction path corresponding to the reactant, the transition state, and the

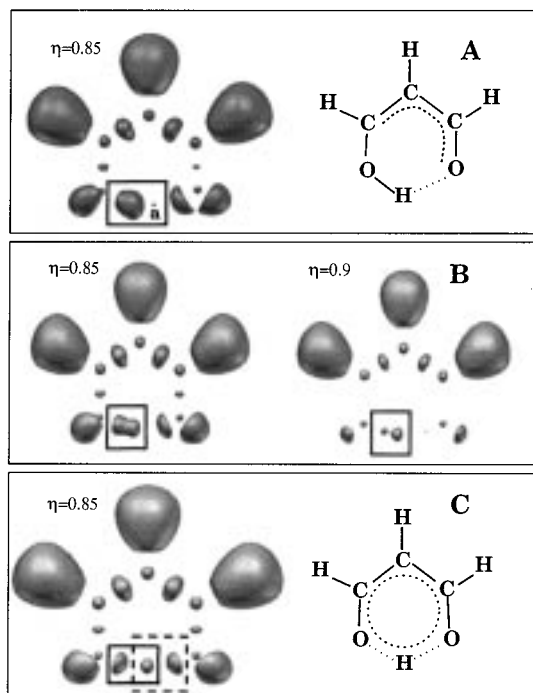


Figure 4. Representation of ELF localization domains for malonaldehyde. (A) $\eta = 0.85$ isosurface for the equilibrium structure showing the different localization domains associated to the bonds and to the oxygen lone pairs. The framed domain is related to the $V(O_1, H)$ basin. (B) $\eta = 0.85$ (left) and $\eta = 0.90$ (right) isosurfaces for a configuration immediately following the bifurcation occurring at $d^* = 1.105$ Å in the framed region. The right part emphasizes the appearance of the $V_2(O_1)$ and $V(H)$ attractors. (C) $\eta = 0.85$ isosurface for the transition state structure. The full-line frame shows the $V_2(O_1)$ and $V(H)$ localization domains, and the dashed-line one emphasizes the $V(H)$ and $V_2(O_2)$. After the second bifurcation, the $V(H)$ and $V_2(O_2)$ basin will merge into a unique $V(O_2, H)$ basin corresponding to the formation of the O_2H bond.

product. The symmetry of the system implies that the reactant and the product are identical; there, one half of the full reaction path, namely the proton detachment from O_1 between the reactant and the transition state, has to be considered. The proton attachment on O_2 is just the inverse process. Fifteen configurations along this first part of the reaction path have been selected in order to locate the bifurcation state(s). The true control space parameter is the one-dimensional reaction coordinate which is actually a function of the 21 internal coordinates. In order to have access to numerical values, the projection of the reaction coordinate on the O_1H , distance d_1 , has been considered hereafter for the proton detachment. For the proton attachment on O_2 one has to consider $d_2 = r(O_2H)$. Figure 4 displays the localization domains (i.e., the volumes bounded by a given isosurface) of malonaldehyde for the nuclear configuration corresponding to the equilibrium structure (A), the breaking of the O_1H bond (B), and the transition state (C). The bond dissociation occurs at $d_1 = d^* = 1.105$ Å: a new maximum of $\eta(r)$ appears which splits the $V(O_1, H)$ basin into a lone pair basin $V_2(O_1)$ and the "dressed proton" one $V(H)$. As shown in Figure 5, the $V(O_1, H)$ basin population increases from $1.85 e^-$ in the equilibrium structure to $2.0 e^-$ at the turning point. Just after the dissociation, the population of the $V(H)$ is $0.7 e^-$, which supports the interpretation in terms of a covalent dissociation. As the system evolves toward the transition state, the $V(H)$ population decreases being transferred to the $V_2(O_1)$ lone pair basin. On the second oxygen atom, the two lone pair basins are involved in a charge transfer which lowers the

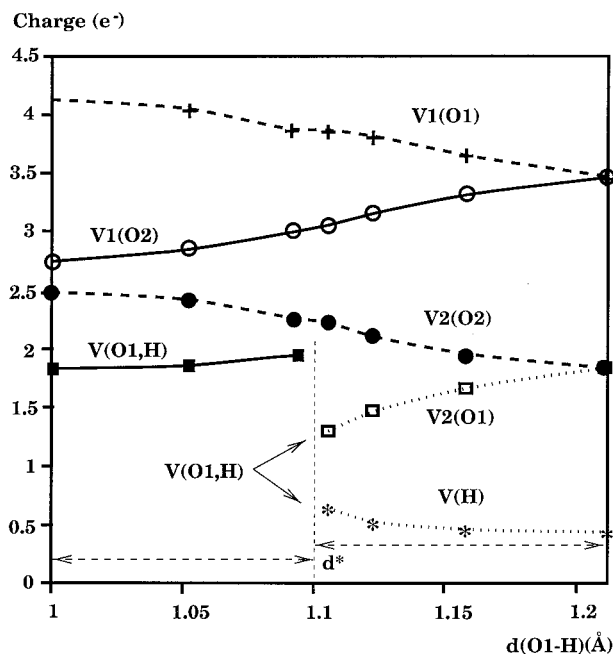


Figure 5. Populations of the $V_1(O_1)$, $V_2(O_1)$, $V_1(O_2)$, $V_2(O_2)$, $V(O_1, H)$, and $V(H)$ basins (in e^-). At the bifurcation state, $d^* = 1.105$ Å, $V(O_1, H)$ yields two basins, $V_2(O_1)$ and $V(H)$. The opposite bifurcation takes place when the HO_2 bond is formed.

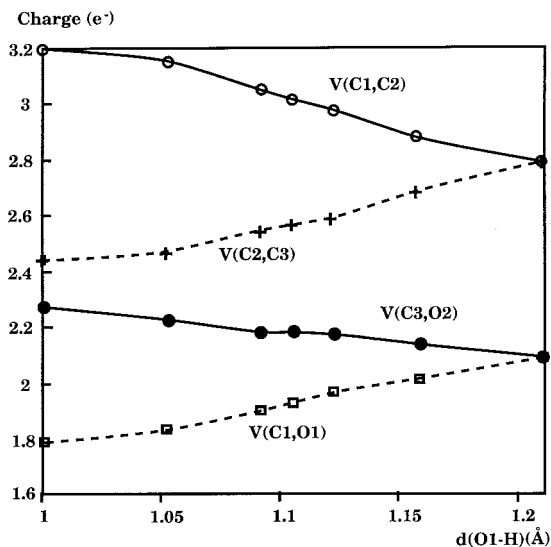


Figure 6. Populations of the $V(C_1, O_1)$, $V(C_1, C_2)$, $V(C_2, C_3)$, $V(C_3, O_2)$ basins (in e^-), as a function of the O_1H (in Å) distance.

population of $V_2(O_2)$, the basin located on the H side and increases that of other one. The symmetric process which forms the O_2H bond involves the same variation in the opposite direction and the permutation of the oxygen labels. It is interesting to note that in the whole process the sum of the population of the lone pair, $V(O, H)$ and $V(H)$ basins is almost constant. Figure 6 presents the behavior of the skeleton bond basin populations along the reaction path. The total population of these basins is also constant and the charge transfers occur on the one hand from $V(C_1, C_2)$ toward $V(C_2, C_3)$ and $V(C_1, O_1)$ and on the other hand from $V(C_3, O_2)$ toward $V(C_2, C_3)$.

From the point of view of the catastrophe theory, the event which occurs at $d_1 = d^*$ is a bifurcation catastrophe of the *fold* type. This means that the appearance of the new local maximum involves the transformation of a *wandering point*, i.e., a point at which $\nabla\eta(r) \neq 0$, into a pair of critical points: the new

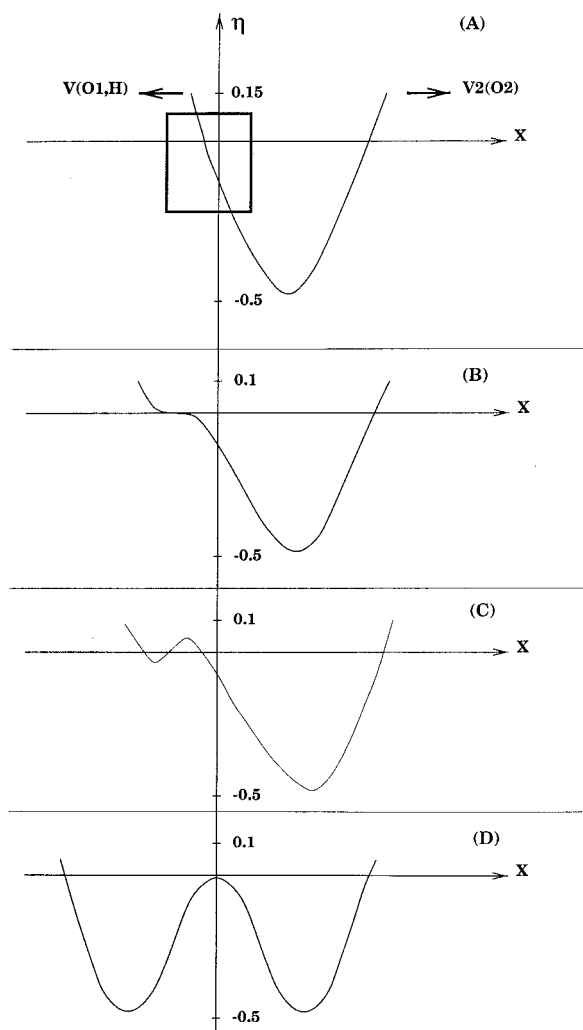


Figure 7. Schematic representation of $\eta + \text{const.}$ in the region of interest. (A) Before the bifurcation state $d < d^*$, $a < 0$, and $b < b^*$. The universal unfolding applies in the framed region. (B) At $d_1 = d^*$ a point of the $V(\text{O}_1, \text{H})$ basin and near its separatrix, becomes degenerate, $a < 0$ and $b = b^*$. This catastrophe generates basins, $V_2(\text{O}_2)$ and $V(\text{H})$; see Figure 4B. (C) evolution of the degenerate point as distance d increases, $a < 0$ and $b^* < b < 0$. (D) This symmetrical structure corresponds to the transition state, $a < 0$ and $b > 0$. On the remaining reaction path these changes take place in the opposite direction.

maximum and a saddle point located on the border surface between $V(\text{H})$ and $V_2(\text{O}_1)$. A local model of this catastrophe is given by its *universal unfolding*:

$$\eta(x; b) + \text{constant} = -x^3 + bx \quad (5)$$

in which x is the space variable taken along the direction joining the $V(\text{O}_1, \text{H})$ and $V_2(\text{O}_2)$ attractor of the reactant and b parameter related to the reaction coordinate. In fact, the behavior of the ELF function can be represented in a single dimension. Figure 7 displays the graph of $\eta(x; b) + \text{constant}$ in the region of interest. For $b < 0$ (Figure 7A) there is no critical point in the framed region within which the universal unfolding is valid because the derivative

$$d\eta(x; b)/dx = -3x^2 + b \quad (6)$$

is always negative. For $b = 0$ (Figure 7B), eq 6 has a double root for $x = 0$, this situation corresponds to the bond breaking. Figure 7, C and D, display the function for $b > 0$, just after the turning point and at the transition state. The parameter b of eq

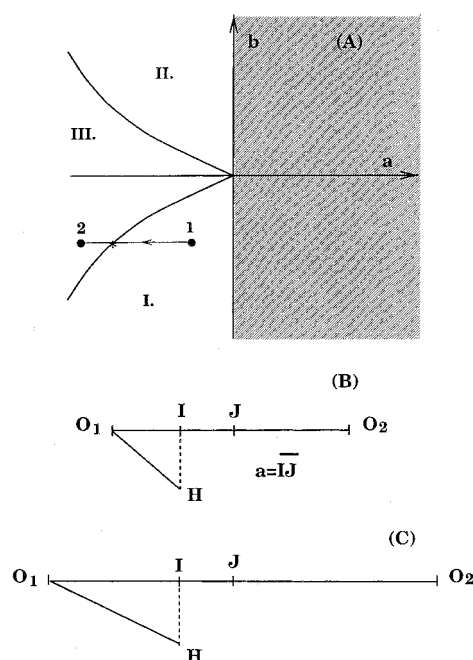


Figure 8. (A) Critical curve of the cusp catastrophe $(b^2/8) + (a^3/27) = 0$. The shaded area corresponds to the negative forbidden values of a . Starting from a given point of coordinate (a_1, b_1) located in the domain of structural stability I, it is possible to go to the domain III by decreasing the a parameter. The sole variation of a does not allow to reach domain II. (B) and (C) Schematic representation of O_1O_2 and O_1H before and after the decrease of a showing the lengthening of the OH bond.

5 can be identified as $d_1 - d^*$. The bond formation on the O_2 side is achieved by a second fold catastrophe for which the b parameter is related to $d_2 - d^*$.

The two fold catastrophes described above involve a same critical point, namely the attractor of the $V(\text{H})$ basin. Therefore, according to the method described in the textbook of Gilmore,⁷⁵ it is possible to unify the description of the process by collecting the two fold catastrophes into a higher order catastrophe which is a *cusp* one in the actual case. The relationship between the unfoldings of the cusp and fold catastrophe is given in Appendix B. The unfolding of the cusp catastrophe is given by the expression

$$\eta(x; a, b) = -(x^4 + ax^2 + bx) \quad (7)$$

which now involves a second parameter a . As in eq 5, b is related to the O_1H and O_2H distances, for the energetical transition state (Figure 4C) $b = 0$ whereas for the bifurcation state $b = \pm b^*$. Therefore, b can be related to an odd function of the algebraic distance between the projection of the H atom on the O_1O_2 direction and the midpoint of O_1O_2 . The second parameter, a , is always negative. For any bifurcation state the critical values a^* and b^* fulfill the relationship (see ref 76 and Appendix B)

$$\frac{b^{*2}}{8} + \frac{a^{*3}}{27} = 0 \quad (8)$$

Therefore, starting from a couple (a, b) within one of the stable tautomer's stability domain, keeping b constant, it is possible to decrease a up to a^* in order to realize the bifurcation as shown in Figure 8A. In terms of normal coordinates, b is related to the H-bond fast mode involving the OH stretching whereas a relies upon the slow mode which is mainly a function of R , the O_1O_2 distance. It is worth noting that the particular evolution

of the dynamical system upon a variation of a indicates that a behaves as $-R$: the decrease of a corresponds to an increase of the O_1O_2 distance and therefore, b being kept constant, the short OH bond is lengthened and this ultimately yields the breaking of the bond (Figure 8, B and C). The a and b parameters are the active parameters of the control space. The critical curve of the cusp catastrophe (Figure 8A) partitions the plane into four regions: a shaded forbidden region, the reactant-like (I), the product-like (II), and the transition state-like (III) ones. The projection of any reaction path onto the plane defined by the active control parameter subspace must therefore cross region III. Indeed, the critical curve represents the projection of the limiting paths.

The method is particularly stable with respect to the method. On the one hand, both calculations performed with the 6-31G, 6-31G** and ccp-VTZ basis sets with the HF, MP2, BLYP, and B3LYP methods yield the same partition of the control space and the same bifurcation catastrophes at the turning points. On the other hand, the critical value d^* lies in a very narrow range between 1.10 and 1.12 Å for all calculations except the 6-31G/BLYP one for which the bounds are 1.09 and 1.10 Å.

3. Conclusion

The investigation of the proton transfer in malonaldehyde presented here shows how the topological analysis of the ELF combined with the catastrophe theory interpretation of the turning points enable to discuss the reaction mechanism. From this analysis it follows that the proton transfer is a two-step reaction. The first step corresponds to the covalent dissociation of the OH bond, and the second step is the formation of another OH bond from an hydrogen atom and an oxygen lone pair. The process involves an intermediate state in which the hydrogen is detached and within which a significant electron transfer from the hydrogen toward the oxygen lone pairs brings its ionic character to the reaction. The presence of the detached hydrogen, or "dressed proton", in the intermediate state is a consequence of the attractive Coulombic potential of the proton which has no Pauli repulsion counterpart. There is no bonding rearrangement in the skeleton but rather a concerted charge transfer from one bond to another. The catastrophe analysis enables to identify the active control space parameters, in other words how many and which are the coordinates that drive the reaction. Moreover, the discussion of the catastrophe critical curve provides criteria to determine the limit paths of the reaction. Finally, at least, in this example, the method appears very stable with respect to the computational scheme of the wavefunction.

Appendix A

By definition, a *dynamical system* is a field of bound vectors \mathbf{X} on a manifold M . For each and every point of M of coordinates $\{m\}$ the equations $dm/dt = \mathbf{X}(m)$ determine a unique trajectory $h(m)$. Although the analogy with a velocity field is purely formal, the method has been widely used to model the time evolution of many phenomena. The trajectories begin and end in the neighborhood of points for which $\mathbf{X}(m) = 0$. For a given point p belonging to M , $\alpha(p)$ and $\omega(p)$ denote the limit sets of $p(t)$ in M corresponding respectively to $t \rightarrow -\infty$ and to $t \rightarrow +\infty$.

A *gradient dynamical system* is a dynamical system for which the vector field \mathbf{X} derives from a scalar function V , called the *potential function*, that is $\mathbf{X} = \nabla V$. A point is a *critical point* if $\mathbf{X} = 0$, a *wandering point* otherwise.

The *stable manifold* or *inset* of a critical point is the set of all the points for which this critical point is an ω limit, the *unstable manifold* or *outset* the set of those for which it is an α limit. The critical points of a gradient dynamical system are classified according to their *index*—i.e., the number of positive *critical exponents*, here the eigenvalues of the Hessian matrix. The dimension of the unstable manifold of a critical point is equal to its index. The index of a critical point m of the vector field \mathbf{X} is denoted by $I(\mathbf{X}, m)$. The critical points are also denoted by a pair of integers (r, s) , the rank (number of non-zero eigenvalues) and the signature (number of positive minus negative eigenvalues) of the Hessian matrix. In the Euclidean 3-dimensional space, there are four kinds of critical points: the *repellers* of index 3, noted $(3, 3)$ which are the local minima of the potential function; the saddle points $(3, 1)$ and $(3, -1)$ of index respectively 2 and 1; the *attractors* $(3, -3)$ of index 0 which are the local maxima of the potential function. Attractors are only ω limits and repellers only α limits, whereas saddle points are both. The stable manifold of an attractor is called the *basin* of the attractor. The *separatrices* are the boundary points, lines or surfaces of two or more basins. They are the stable manifolds of the saddle points. The number of *hyperbolic* critical points (i.e., without zero critical exponent) satisfies the phase rule type relationship which is given by the Poincaré–Hopf theorem:

$$\sum (-1)^{I(\mathbf{X}, m)} = \chi(M) \quad (9)$$

The sum runs over the critical points of the vector field \mathbf{X} bound on the manifold M and $\chi(M)$ is the Euler characteristic of the manifold. For finite and periodic systems in \mathbb{R}^q , $q \leq 3$, the Euler characteristic is 1 and 0 respectively.

When V , and therefore \mathbf{X} , depend parametrically upon the set of *control space* parameters $\{c_\alpha\}$ the number and the type of the critical points change with the variation of the *control space* parameters. The domain of *structural stability* of a critical point is a subset of the control space within which this critical point remains hyperbolic. If all the critical points of a dynamical system are hyperbolic, the system is structurally stable: a weak perturbation of the vector field changes neither the index nor the number of critical points. In the bonding theory such a weak perturbation may be due to a change in the technique of calculation of the approximate wave function (basis set, correlation scheme) or to a small displacement of the nuclei.

The point in the control space for which at least one critical point becomes nonhyperbolic is called a *bifurcation point*. At a bifurcation point a *bifurcation catastrophe* occurs which brings the system from a domain of structural stability to another. There are seven elementary catastrophes⁵⁵ classified according to their *universal unfolding*. The universal unfolding is a parametric polynomial expression which provides a local model of the potential function in the neighborhood of the critical point subjected to a bifurcation. The number of independent parameters involved in the universal unfolding provides the dimension of the active control space.

Appendix B

Consider the universal unfolding of the fold catastrophe:

$$f(x; a, b) = x^4 + ax^2 + bx \quad (10)$$

The critical points correspond to the real roots of

$$df(x; a, b)/dx = 4x^3 + 2ax + b = 0 \quad (11)$$

the number of which is ruled by $q = (b^2/8) + (a^3/27)$. For $q > 0$, there is one critical point since eq 11 has one real root and two complex conjugate ones. For $q = 0$ all roots are real and at least two are equal; this case corresponds to the bifurcation at which one critical point is degenerate. Finally, for $q < 0$ there are three critical points.

$$\frac{b^2}{8} + \frac{a^3}{27} = 0 \quad (12)$$

defines the critical curve of the cusp catastrophe.

Let a^* and b^* a couple of critical values of the control space parameters (i.e., lying on the critical curve). Denoting by λ the double root and by ρ the remaining one, eq 11 can be written as

$$df(x; a^*, b^*)/dx = 4(x - \lambda)^2(x - \rho) = 0 \quad (13)$$

The Taylor expansion of $f(x; a^*, b^*)$ in the neighborhood of $x_0 = \lambda$ is

$$f(x; a^*, b^*) = f(\lambda; a^*, b^*) + \frac{1}{3!} \left(\frac{d^3 f(x; a^*, b^*)}{dx^3} \right)_{x=\lambda} (x - \lambda)^3 + \frac{1}{4!} \left(\frac{d^4 f(x; a^*, b^*)}{dx^4} \right)_{x=\lambda} (x - \lambda)^4 = f(\lambda; a^*, b^*) + \frac{4}{3} (1 - \rho)(x - \lambda)^3 + \frac{2}{3} \lambda (x - \lambda)^4 \quad (14)$$

The germ of the unfolding is obtained by dropping out the constant term and the higher order terms; thus with the change of variable $\tilde{x} = \kappa(x - \lambda)$

$$f(\tilde{x}) = \tilde{x}^3 \quad (15)$$

and the fold catastrophe unfolding is obtained by adding the canonical perturbation $b\tilde{x}$ which corresponds to a small shift of the origin off the critical curve, i.e.

$$f(\tilde{x}) = \tilde{x}^3 + b\tilde{x} \quad (16)$$

Acknowledgment. We thank Prof. E. M. Evleth and Dr. D. Borgis for their constant interest and Dr. A. Sevin for critically reading and improving the draft manuscript. This work was performed within the Groupement de Recherche Européen "Dynamique moléculaire quantique appliquée à la catalyse, à l'adsorption et à l'absorption", supported by the Institut Français du Pétrole and the Centre National de la Recherche Scientifique (France).

References and Notes

- Page, M. T. *The Chemistry of Enzyme Action*; Elsevier: Amsterdam 1974.
- Caldin, E. F.; Gold, V. *Proton-Transfer Reactions*; Halsted: New York 1975.
- Stewart, R. *The Proton: Applications to Organic Chemistry*; Academic Press: Orlando, FL, 1985.
- Borgis, D.; Hynes, J. T. In *The Enzyme Catalysis Process*; Cooper, J., Houben, J., Chien, L., Eds.; Plenum: New York, 1989.
- Krokidis, X.; Noury, S.; Silvi, B. *J. Phys. Chem. A* **1977**, *101*, 7277.
- Becke, A. D.; Edgecombe, K. E. *J. Chem. Phys.* **1990**, *92*, 5397.
- Barone, V.; Adamo, C. *J. Chem. Phys.* **1996**, *105*, 11007.
- Lewis, G. N. *Valence and the Structure of Atoms and Molecules*; Dover: New York, 1966.
- Foster, J. M.; Boys, S. F. *Rev. Mod. Phys.* **1960**, *32*, 300.
- Edmiston, C.; Ruedenberg, K. *Rev. Mod. Phys.* **1963**, *35*, 457.
- von Niessen, W. *J. Chem. Phys.* **1970**, *56*, 4290.
- England, W. *Int. J. Quant. Chem.* **1971**, *5*, 683.
- Mulliken, R. S. *J. Chem. Phys.* **1955**, *23*, 1833.
- Coulson, C. A. *Valence*; Clarendon, Oxford, UK, 1952.
- Wiberg, K. B. *Tetrahedron* **1968**, *24*, 1083.
- Mayer, I. *Chem. Phys. Lett.* **1983**, *97*, 270.
- Mayer, I. *Int. J. Quantum Chem.* **1986**, *29*, 73.
- Cusachs, L. C.; Politzer, P. *Chem. Phys. Lett.* **1968**, *1*, 529.
- Reed, A. E.; Weinstock, R. B.; Weinhold, F. *J. Chem. Phys.* **1985**, *83*, 735.
- Bachrach, S. M. In *Reviews in Computational Chemistry*; Lipkowitz, K. B., Boyd, D. B., Eds.; VCH: New York, 1994; pp 171–227.
- Dunitz, J. D.; Sieler, P. *J. Am. Chem. Soc.* **1983**, *105*, 7056.
- Schwarz, W. H. E.; Valtazanos, P.; Ruedenberg, K. *Theor. Chim. Acta (Berlin)* **1985**, *68*, 471.
- Schwarz, W. H. E.; Mensching, L.; Valtazanos, P.; von Niessen, W. *Int. J. Quantum Chem.* **1986**, *29*, 909.
- Schwarz, W. H. E.; Mensching, L.; Valtazanos, P.; von Niessen, W. *Int. J. Quantum Chem.* **1986**, *30*, 439.
- Kunze, K. L.; Hall, M. B. *J. Am. Chem. Soc.* **1986**, *108*, 5122.
- Ruedenberg, K.; Schwarz, W. H. E. *J. Chem. Phys.* **1990**, *92*, 4956.
- Bader, R. F. W. *Atoms in Molecules: A Quantum Theory*; Oxford University Press: Oxford, UK, 1990.
- Cao, W. L.; Gatti, C.; MacDougall, P. J.; Bader, R. F. W. *Chem. Phys. Lett.* **1987**, *141*, 380.
- Bader, R. F. W.; Anderson, S. G.; Duke, A. J. *J. Am. Chem. Soc.* **1979**, *101*, 1389.
- Bachrach, S. M. *J. Am. Chem. Soc.* **1986**, *108*, 6406.
- Ritchie, J. P.; Bachrach, S. M. *J. Am. Chem. Soc.* **1987**, *109*, 5909.
- Cioslowski, J.; Mixon, S. T.; Edwards, W. D. *J. Am. Chem. Soc.* **1991**, *113*, 1083.
- Tal, Y.; Bader, R. F. W.; Nguyen-Dand, T. T.; Ojha, M.; Anderson, S. G. *J. Chem. Phys.* **1981**, *74*, 5162.
- Bader, R. F. W.; Nguyen-Dang, T. T.; Tal, Y. *Rep. Prog. Phys.* **1981**, *44*, 893.
- Cioslowski, J. *J. Phys. Chem.* **1990**, *94*, 5496.
- Bersuker, G. I.; Peng, C.; Boggs, J. E. *J. Phys. Chem.* **1993**, *97*, 9323.
- Cooper, D. L. *Nature* **1990**, *346*, 789.
- Bader, R. F. W.; MacDougall, P. J.; Lau, C. D. H. *J. Am. Chem. Soc.* **1984**, *106*, 1594.
- Bader, R. F. W.; Gillespie, R. J.; MacDougall, P. J. *J. Am. Chem. Soc.* **1988**, *110*, 7329.
- Hunter, G. *Int. J. Quantum Chem.* **1986**, *29*, 197.
- Luken, W. L.; Culberson, J. C. *Int. J. Quantum Chem.* **1982**, *16*, 265.
- Bader, R. F. W.; Johnson, S.; Tang, T.-H.; Popelier, P. L. A. *J. Phys. Chem.* **1996**, *100*, 15398.
- Gillespie, R. G.; Robinson, E. A. *Angew. Chem., Int. Ed. Engl.* **1996**, *35*, 495.
- Savin, A.; Jepsen, O.; Flad, J.; Andersen, O. K.; Preuss, H.; von Schnering, H. G. *Angew. Chem., Int. Ed. Engl.* **1992**, *31*, 187.
- Silvi, B.; Savin, A. *Nature* **1994**, *371*, 683.
- Savin, A.; Nesper, R.; Wengert, S.; Fässler, T. F. *Angew. Chem., Int. Ed. Engl.* **1992**, *36*, 1809.
- Ritchie, J. P.; King, H. F.; Young, W. S. *J. Chem. Phys.* **1986**, *85*, 5175.
- Carroll, M. T.; Chang, C.; Bader, R. F. W. *Mol. Phys.* **1988**, *63*, 387.
- MacDougall, P. J. The Laplacian of the Electronic Charge Distribution; Ph.D. Thesis; Hamilton, Ontario, Canada, February 1989.
- Bader, R. F. W.; Chang, C. *J. Phys. Chem.* **1989**, *93*, 2946.
- Savin, A.; Becke, A. D.; Flad, J.; Nesper, R.; Preuss, H.; von Schnering, H. G. *Angew. Chem., Int. Ed. Engl.* **1991**, *30*, 409.
- von Weizsäcker, C. F. Z. *Phys.* **1935**, *96*, 431.
- Tal, Y.; Bader, R. F. W.; *Int. J. Quantum Chem.* **1978**, *S12*, 153.
- Savin, A.; Silvi, B.; Colonna, F. *Can. J. Chem.* **1996**, *74*, 1088.
- Thom, R. *Stabilité Structurale et Morphogénèse*; Interditions: Paris 1972.
- Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Gill, P. M. W.; Johnson, B. G.; Robb, M. A.; Cheeseman, J. R.; Keith, T.; Petersson, G. A.; Montgomery, J. A.; Raghavachari, K.; Al-Laham, M. A.; Zakrzewski, V. G.; Ortiz, J. V.; Foresman, J. B.; Cioslowski, J.; Stefanov, B. B.; Nanayakkara, A.; Challacombe, M.; Peng, C. Y.; Ayala, P. Y.; Chen, W.; Wong, M. W.; Andres, J. L.; Replogle, E. S.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Binkley, J. S.; Defrees, D. J.; Baker, J.; Stewart, J. P.; Head-Gordon, M.; Gonzalez, C.; Pople, J. A. *Gaussian 94, Revision D.4*; Gaussian Inc.: Pittsburgh, PA, 1995.
- Frisch, M. J.; Frisch, A. E.; Foresman, J. B. *Gaussian 94, User's Reference*; Gaussian Inc.: Pittsburgh, PA, 1995.
- Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648.
- Becke, A. D. *J. Chem. Phys.* **1988**, *98*, 5648.
- Lee, C.; Yang, Y.; Parr, R. G. *Phys. Rev.* **1988**, *B37*, 785.
- Miehlich, B.; Savin, A.; Stoll, H.; Preuss, H. *Chem. Phys. Lett.* **1989**, *157*, 200.

- (62) Dunning, T. H., Jr. *J. Chem. Phys.* **1989**, *90*, 1007.
- (63) Kendall, R. A.; Dunning, T. H., Jr.; Harrison, R. J. *J. Chem. Phys.* **1992**, *96*, 6796.
- (64) Chiavassa, T.; Roubin, P.; Pizzala, L.; Verlaque, P.; Allouche, A.; Marinelli, R. W. *J. Am. Chem. Soc.* **1982**, *96*, 10659.
- (65) Frisch, M. J.; Scheiner, A. K.; Schaefer, H. F., III *J. Chem. Phys.* **1985**, *82*, 4194.
- (66) Wiberg, K.; Hadad, C. M.; LePage, T. J.; Breneman, C. M.; Frisch, M. J. *J. Phys. Chem.* **1992**, *96*, 671.
- (67) Luth, K.; Scheiner, S. *J. Phys. Chem.* **1994**, *98*, 3582.
- (68) Baughcum, S. L.; Smith, Z.; Wilson, E. B.; Duerst, R. W. *J. Am. Chem. Soc.* **1984**, *106*, 2260.
- (69) Rowe, W. F., Jr.; Duerst, R. W.; Wilson, E. B. *J. Am. Chem. Soc.* **1976**, *98*, 4021.
- (70) Baughcum, S. L.; Duerst, R. W.; Rowe, W. F., Jr.; Smith, Z.; Wilson, E. B. *J. Am. Chem. Soc.* **1981**, *103*, 2696.
- (71) Hehre, W. J.; Stewart, R. F.; Pople, J. A. *J. Chem. Phys.* **1969**, *51*, 2657.
- (72) Ditchfield, R.; Hehre, W. J.; Pople, J. A. *J. Chem. Phys.* **1971**, *54*, 724.
- (73) Hehre, W. J.; Ditchfield, R.; Pople, J. A. *J. Chem. Phys.* **1972**, *56*, 2257.
- (74) Noury, S.; Krokidis, X.; Fuster, F.; Silvi, B. TopMoD package, 1997.
- (75) Gilmore, R. *Catastrophe Theory for Scientists and Engineers*; Dover: New York, 1981.
- (76) Woodcock, A. E. R.; Poston, T. *A geometrical Study of Elementary Catastrophes*; Lecture Notes in Mathematics; Springer-Verlag: Berlin, 1974.