HAL
open science

Relating the shape of a molecule and its reactivity -Haddon's Curvature and the Pyramidalization Angle.

Julia Sabalot-Cuzzubbo, Germain Vallverdu, Didier Bégué, Jacky Cresson

To cite this version:

Julia Sabalot-Cuzzubbo, Germain Vallverdu, Didier Bégué, Jacky Cresson. Relating the shape of a molecule and its reactivity -Haddon's Curvature and the Pyramidalization Angle.. 2020. hal-02490358

HAL Id: hal-02490358

https://hal.science/hal-02490358v1

Preprint submitted on 10 Mar 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

.

Relating the shape of a molecule and its reactivity - Haddon's Curvature and the Pyramidalization Angle

J. Sabalot-Cuzzubbo,[1] G. Salvato-Vallverdu,[1] D. Bégué,[1] and J. Cresson[2]

[1)] *Université de Pau et des Pays de l'Adour, E2S UPPA, CNRS, IPREM, Institut des Sciences Analytiques et de Physico-chimie pour l'Environnement et les Matériaux, UMR 5254, Pau, France.*[a)]

[2)] *Université de Pau et des Pays de l'Adour, E2S UPPA, CNRS, LMAP, Laboratoire de Mathématiques Appliquées de l'Université de Pau et des Pays de l'Adour, UMR 5142, Pau, France.*[b)]

(Dated: 24 February 2020)

The pyramidalization angle and spherical curvature are well known quantities used to characterize the local geometry of a molecule and to provide a measure of regio-chemical activity of molecules. In this paper, we give a self contain presentation of these two concepts and discuss their limitations. These limitations can be bypass thanks to the introduction of the notions of angular defect and discrete Gauss curvature coming from discrete differential geometry. In particular, these quantities can be easily computed for arbitrary molecules, trivalent or not, with bond of equal lengths or not. All these quantities have been implemented. We then compute all these quantities over the Tománek database covering an almost exhaustive list of fullerene molecules. In particular, we discuss the interdependence of the pyramidalization angle with the spherical curvature, angular defect and hybridization numbers. We also explore the dependence of the pyramidalization angle with respect to some characteristics of the molecule like the number of atoms, group of symmetry and with respect to the geometrical optimization process.

I. INTRODUCTION

It is well known that the geometry of a molecule plays an important role in its chemical and physical properties. Many authors have focus on some quantities reflecting the "shape" of the molecule and particular properties like regio-reactivity. The most famous quantity is known as the **pyramidalization angle** ($Pyr(A)$) developed by R. C. Haddon during the years 90[1–3]. A second quantity introduced by R. C. Haddon[1] is the **spherical curvature** (SC) which aims to be a discrete analogue of the classical curvature in differential geometry.
Despite its ubiquity in many articles, the theoretical background of the pyramidalization angle or the spherical curvature were not reviewed or discussed in recent years up to our knowledge. However, such a discussion is necessary as many articles exchange different notions of curvatures, some related to classical differential geometry and some related to discrete differential geometry. It was in fact the aim of R. C. Haddon in his article[1] to clarify this point. Moreover, the Haddon's pyramidalization or spherical curvature can not be applied to arbitrary molecules and are in fact restricted most of the time to trivalent molecules.

In this paper, we give a **self contain presentation of the Haddon's pyramidalization angle** and the **spherical curvature**. We give a detailed derivation of these quantities and some examples of explicit computations. We give a complete discussion of the relation existing between the pyramidalization angle and the **hybridization properties of carbon molecules**. Our main concern is also to give a precise account of the limitations of these quantities. We then discuss the use of another quantity called the **discrete Gauss curvature** and the **angular defect** (AD) as a tool to cover cases for which the Haddon's pyramidalization angle and spherical curvature is defined.
All the **quantities are implemented**[4] and can be used to obtain **cartography** of molecules representing the variation of the pyramidalization angle, spherical curvature or angular defect along the molecule.
As an example, we provide a **full statistical analysis of the Tománek database**[5] and corresponding to a more or less exhaustive list of **fullerenes molecules**. In particular, we study the dependence of the pyramidalization angle with respect to the number of atoms, group of symmetry, geometric optimization and hybridization. As a special case, we consider the C_{32} isomers.

The paper is organized as follows:

In part II, we define and discuss the pyramidalization angle as initiated by R. C. Haddon[2,3]. Detailed proofs are given. In particular, the connection between the pyramidalization angle and the molecular orbitals is given. We give a full discussion of the limitations of this tool as well as explicit examples.

In part III, we define and discuss the notion of spherical curvature as initiated by R. C. Haddon in his work[2,3]. The construction is already restricted to the trivalent case. We discuss the limitations of this tool and we give explicit examples.

[a)] Electronic mail: didier.begue@univ-pau.fr
[b)] Electronic mail: jacky.cresson@univ-pau.fr

In part IV, in order to bypass the limitations describes in section II and part III, we introduce the notion of angular defect and discrete Gauss curvature. We give explicit examples.

Part V, we explore the Tománek database for fullerene molecules. We compute all the relevant quantities (pyramidalization angle, spherical curvature, angular defect, hybridization number) and study the dependence of all these quantities between them and also with respect to some characteristic of the molecules like the number of atoms, group of symmetry and geometric optimization.

Finally, we explore some perspectives raised by our results.

Note. All the figures given in this article can be obtained using the online program provide at https://pychemapps.univ-pau.fr/mosaica/.

A. Haddon's representation problem

In his seminal paper[1], R. C. Haddon sets the following intriguing question: does the fullerene C_{60} is a sphere or a polyhedron ? Of course, such a question has no meaning without precisions because has already reported in the previous sections, there exist no canonical continuous or discrete representation of a molecule. It seems that R. C. Haddon was thinking about a geometric representation of the molecule which has some "real" content. Indeed, as he remarked just after his question "Of course, in reality the answer is well-known, and neither the sphere nor the polyhedron represent C_{60}, which lie other molecules exists as a collection of nuclei with an associated distribution of electron density.". He then precised his idea asking "which of these conventional representations is most relevant for the fullerenes and in particular the language most appropriate to the description of the shape of these molecules and the geometry of the carbon atoms." Unfortunately, these precisions do not give a more well defined problem as the shape of a molecule is precisely a matter of representation and the geometry of carbon atoms is not defined in the text of R. C. Haddon. However, the work of R. C. Haddon suggests that by "geometry of carbon atoms", he focus on the data of atomic orbitals and how these atomic orbitals interact due to the shape of the molecule to produce molecular orbitals following in this way Hückel's approach to the computation of molecular orbitals. As a consequence, the problem formulated by R. C. Haddon is in fact related to the problem of a selection between two geometric representations of a given molecule with respect to its chemical properties, precisely the interplay between the local carbon atoms geometry related to atomic orbitals and the global geometry.

One can then generalize the representation problem of R. C. Haddon in the following way:

Haddon's representation problem: *Let $M = (\mathscr{A}, \mathscr{R})$ be a given molecule admitting a continuous representation $\mathscr{M}$. Let P be a physical or chemical property of M_A. Which representation of M, given by the continuous $\mathscr{M}$ or the skeleton one $S(M)$ is the most appropriate to study this property ?*

Of course, depending on the property of the molecule under investigation, the answer can be different. R. C. Haddon has in particular formulated the previous problem concerning the study of fullerene in the context of the reactivity property of the molecule, even if it is not explicitly written.

In his reference paper[1], R. C. Haddon introduces two quantities which are of current used in the community: the pyramidalization angle and the curvature. In the following, we first discuss these two quantities and precise their domain of application.

II. HADDON'S PYRAMIDALIZATION ANGLE

In this section, we define and discuss the pyramidalization angle introduced by R. C. Haddon in a series of articles[1–3]. We first restrict our attention to molecules having all bonds of equal lengths and trivalents. This case cover well known molecules as the fullerenes and graphene. The definition of a pyramidalization angle in a given atom A of the molecule is associated to the existence of a particular vector called the $POAV(A)$ **(π-Orbital Axis Vector)** vector which has the property to make a constant angle with each bonds of connected to A[6]. Generalization of this vector to arbitrary trivalent molecules is made trough a notion of **regularized star** for a given atom. We then discuss the relevance of this angle for chemistry. In particular, we derive a detailed analysis of the structure of the molecular orbital in A with explicit expressions making clear the connection between the orbitals and the pyramidalization angle. We then propose a way to generalize the pyramidalization angle for more general molecules. We finally discuss the limitations of the pyramidalization angle.

A. Haddon's pyramidalization for trivalent molecules

We follow the presentation made in Haddon and al. papers[1–3].

1. *Construction in the trivalent regular case*

We first restrict our attention to **trivalent regular molecules**, meaning that we assume that all the bonds are of equal lengths and that each atom of the molecule is connected to three atoms. In this case, we have the following Lemma, which is the **main ingredient** of the construction of the pyramidalization angle:

Lemma 1 (POAV)**.** *Let M be a trivalent regular molecule. In each atom A of the molecule M, there exists a unique unitary vector denoted by $POAV(A)$ such that $POAV(A)$ makes a constant angle denoted by θ with each line segment AB where $B \in \star(A)$ such that $\frac{\pi}{2} \leq \theta < \pi$.*

Proof. Let A be a given atom of M. As the molecule M is trivalent, we have three connected atoms B_1, B_2, B_3 connected to A. These three atoms define a plane $\mathscr{P}(A)$. As the three bonds are equal by assumption on M, the atom A belongs to the intersection of each mediating plane, then in particular to the line passing trough A and the intersection of the mediating line in the triangle defined by B_1, B_2 and B_3. Let us denote by I the intersection point of all mediating line in the plane $\mathscr{P}(A)$. We define $POAV(A)$ as the unique unitary vector directed from I to A. It remains to prove that

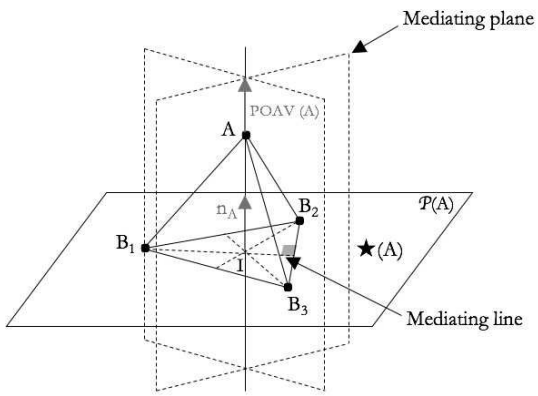


FIG. 1: Construction in a trivalent regular case.

$POAV(A)$ is such that the angle with each bond is constant. We first remark that the vector $POAV(A)$ is normal to the plane $\mathscr{P}(A)$. Indeed, the line (IA) belongs to the intersection of two intersecting planes which are normal to the plane $\mathscr{P}(A)$. It follows that $POAV(A)$ is a normal vector in A.

Let B be an atom of $\star(A)$ in $\mathscr{P}(A)$. Let us denote by l the bond length between A and atoms of $\star(A)$ and z_A the height IA of A with respect to the plane $\mathscr{P}(A)$. We denote by θ_{AB} the angle between $POAV(A)$ and the bond AB. The angle θ_{AB}

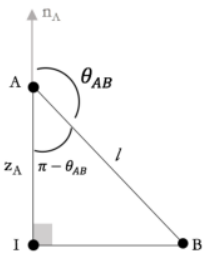


FIG. 2: Angle between a bond and the $POAV(A)$ vector.

satisfies:

$$\cos(\pi-\theta_{AB}) = \frac{z_A}{l}, \ \sin(\pi-\theta_{AB}) = \frac{IB}{l} \tag{1}$$

As θ_{AB} satisfies $\frac{\pi}{2} \leq \theta_{AB} < \pi$, we have $0 \leq \pi - \theta_{AB} < \frac{\pi}{2}$ and the value of the arccos is sufficient to determine $\pi - \theta_{AB}$ completely. As seen in the formula, the angle θ_{AB} depends only on z_A and l which are quantities independent of B. As a consequence, θ_{AB} is a quantity only depending on A and the length of the bond but independent of B. As a result, that the angle between the vector $POAV(A)$ and each bond is constant. This concludes the proof. □

The previous Lemma was discussed (in an informal way) in the original paper[2] (Appendix) by R. C. Haddon and L. T. Scott as an answer to there colleagues disbelieving the existence of such a vector.

Remark 1. *The existence of the $POAV(A)$ vector can be related to the existence of a cone in A which is tangent to each face containing A. Such a property for a mesh was introduced by Y. Liu, and co-writers[7] under the name of* **conical meshes** *in particular with the problem of constructing* **offsets** *(see Chap.5,§.5 p.134[8]). In particular, the previous result shows that any meshes whose vertex are trivalent is conic (see Proposition 8 p.137[8]). We refer to proposition 9 of Romon's book[8] (p.137), where an explicit condition is given for a mesh whose all vertex are of order* 4 *to be conic and so, in particular, for which a POAV vector can be defined.*

We are now ready to give the definition of the **pyramidalization angle** as defined by Haddon[1]:

Definition 1 (Pyramidalization angle)**.** *Let us consider a trivalent regular molecule M and an atom A of M. Let B be an atom of $\star(A)$. We denote by θ the angle between any bond vector AB, $B \in \star(A)$ and the $POAV(A)$ vector. The* **pyramidalization angle** *in A, denoted by $Pyr(A)$ is then defined by:*

$$Pyr(A) = \theta - \frac{\pi}{2} \tag{2}$$

We then have that $Pyr(A) = 0$ when the molecule is **locally planar** and $3-D$ if $Pyr(A) \neq 0$.

2. Construction in the trivalent non regular case

The $POAV$ vector can be defined even for non regular trivalent molecules. This construction was not given in the articles of R. C. Haddon and coworkers[1–3].

This main ingredient of the construction if the notion of **regularized star** of a given atom in a molecule.

Definition 2 (ε-Regularized star)**.** *Let us consider a trivalent molecule M and A a given atom of M. Let us denote by $\varepsilon > 0$. For each $B \in \star(A)$, we denote by $Reg_\varepsilon B$ the point on AB such that $AB = \varepsilon$ called the regularization of B with respect to A. We denote by $Reg\star(A)$ the set of regularized $B \in \star(A)$ with respect to A called the* ε-**regularized** *star of A.*

The previous construction corresponds to the statement of R. C. Haddon[1] that "*the POAV1 theory of pyramidalization may be related to the construction of a sphere of arbitrary radius with its center at point A. The intersection of the three bonds to the neighboring atoms with the surface of the sphere then defines a circle, and this circle together with the center of the*

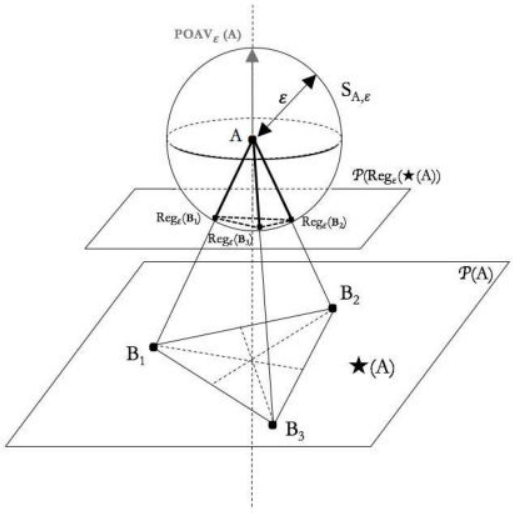


FIG. 3: Construction in a non trivalent regular case .

sphere leads to a right cone. This one generates the family of all bonds directions for a given pyramidalization angle".

The definition of $Reg_\varepsilon(B)$ is easily implemented thanks to the relation:

$$\overrightarrow{A\,Reg_\varepsilon(B)} = \frac{\varepsilon}{AB}\overrightarrow{AB},\ \ \forall B \in \star(A) \tag{3}$$

One can fix a particular choice of ε by taking for example:

$$\varepsilon(A) = \min_{B\in\star(A)} AB, \tag{4}$$

corresponding to the minimum of the bond lengths starting in A.

Using the regularized star of a given atom, one easily extends the notion of pyramidalization angle:

Definition 3. *Let M be a trivalent molecule and A be an atom of M. Let $\varepsilon > 0$, the ε-pyramidalization angle of M in A is defined as the pyramidalization angle associated to A and the ε-regularized star of A.*

Of course, having a definition depending on ε is not satisfying. Fortunately, we have the following Lemma:

Lemma 2. *Let M be a trivalent molecule and A an atom of M. Let $\varepsilon > 0$ and $POAV_\varepsilon(A)$ the vector defined in Lemma 1. The vector $POAV_\varepsilon(A)$ is independent of ε.*

Proof. The $POAV_\varepsilon(A)$ vector is only defined by the **local geometry** of the molecule in a given atom A and precisely only on the relative angles of the bonds starting in A. By construction, all regularized vector $\overrightarrow{AReg_\varepsilon(B)}$ for $B \in \star(A)$ satisfy the relation (3) meaning that if $POAV_\varepsilon(A)$ makes a constant angle with all the vectors $\overrightarrow{AReg_\varepsilon(B)}$ for $B \in \star(A)$, this vector makes also a constant angle with all vectors in $\overrightarrow{AReg_{\varepsilon'}(B)}$ with $B \in \star(A)$ for any $\varepsilon' > 0$. As a consequence, the vector $POAV(A)_\varepsilon$ vector is independent of ε. □

We then deduce an important property of the ε-pyramidalization angle :

Lemma 3. *Let M be a trivalent molecule and A an atom in M. For $\varepsilon > 0$, the ε-pyramidalization angle in A is independent of ε.*

This property can be used to compute explicitly the pyramidalization angle in concrete molecules where bond lengths are different. This result corresponds to the statement by R. C. Haddon that "*the pyramidalization angle is independent of bond lengths*" used for example by Haddon[1](p.1798).

B. The pyramidalization angle and chemistry

The definition of the pyramidalization angle in the previous section was made as a pure geometrical construction in order to put in evidence the geometrical assumptions underlying the definition of this angle. As indicated by R. C. Haddon and L. T. Scott[2], the pyramidalization angle provides a "convenient index of the degree of nonplanarity" at a given atom of the molecule. However, the main concern of R. C. Haddon was chemistry and in particular the chemical properties of a given molecule related to its geometry. We need to interpret the pyramidalization angle in this context. The basic idea is that the pyramidalization angle is related to the **hybridization** of atomic orbitals and in particular is a measure of the distortion a given π-electron system locally in a given molecule. We follow the presentation made by Haddon[3].

1. Hybridization

We consider the orbitals for a conjugated **nonplanar** system. We first discuss the form of the **hybrid σ-orbitals**. We denote by h_1, h_2, h_3 which are obtained as linear combination of the s-orbital, and the p_x, p_y and p_z orbitals. The general form of such orbital is given by:

$$h_i = c_{i,1}s + c_{i,2}p_x + c_{i,3}p_y + c_{i,4}p_z \tag{5}$$

The idea of hybridization is to choose the orbitals in order to obtain the maximal overlap along the bonds. We then introduce three vectors p_1, p_2 and p_3 respecting this constraint:

Definition 4 (Adapted p-orbitals)**.** *Let M be a molecule and A an atom of M. We denote by $p_1(A)$, $p_2(A)$ and $p_3(A)$ the three p-orbitals, defined by:*

$$p_i = c_{i,2}p_x + c_{i,3}p_y + c_{i,4}p_z,\ \ i = 1,2,3, \tag{6}$$

where the vector $c_i = (c_{i,2}, c_{i,3}, c_{i,4})$ is such that each p_i is directed along a given internuclear axes between A and an atom of $\star(A)$.

Using adapted p-orbitals, one can rewrite h_1, h_2 and h_3 as:

$$h_i = c_{i,1}s + \lambda_i p_i,\ i = 1,2,3. \tag{7}$$

In order to describe the hybrid π-orbital denoted by h_π, we introduce following R. C. Haddon and L. T. Scott[2], the notion of **π-orbital axis vector**:

Definition 5 (The π-Orbital Axis Vector)**.** *The π-orbital axis vector, also denoted POAV, is the vector which makes equal angles to the σ-bonds at a conjugated atom.*

The π-orbital axis vector does not always exist as discussed in the previous section. When it exists it coincides with the normal vector n_A. The existence of the π-orbital axis vector for trivalent molecules follows directly from Lemma 1.

Remark 2. *R. C. Haddon and L. T. Scott[2] indicated that a similar approach was performed by J.G. Radziszewski and al.[9].*

The hybrid π-orbital is then defined as:

$$h_\pi = c_\pi s + \lambda_\pi p_\pi \tag{8}$$

By definition of the *POAV*, denoting by $\theta_{i,\pi}$, $i = 1,2,3$ the angle between p_π and the p_i, the main assumption on the π-orbital is then equivalent to:

$$\theta_{1,\pi} = \theta_{2,\pi} = \theta_{3,\pi} \tag{9}$$

We denote by $\theta_{\sigma,\pi}$ this common angle. Clearly the pyramidalization angle $Pyr(A)$ is then related to $\theta_{\sigma,\pi}$ by the relation:

$$Pyr(A) = \theta_{\sigma,\pi} - \frac{\pi}{2} \tag{10}$$

2. Adapted reference frame

The previous geometrical picture can be made easier by looking for an adapted reference frame. We first put the origin of the reference frame at the atom under consideration. We then consider that the POAV vector p_π is colinear to p_z so that:

$$h_\pi = c_\pi s + \lambda_\pi p_z \tag{11}$$

We denote by $\mathscr{P}$ the orthogonal plane to p_z passing trough O. By assumption on the POAV, the σ-orbitals h_1, h_2 and h_3 make an angle given by $Pyr(A)$ with the plane $\mathscr{P}$ and then $\frac{\pi}{2} + Pyr(A)$ with p_z. The picture that we obtain corresponds to the one given by R. C. haddon[2] (Figure 2 p.138).

Following R. C. Haddon and L. T. Scott[2] (p.138), the previous situation can then be understood as an **intermediate hybridization between sp^2 and sp^3** depending on the pyramidalization angle.
One can choose the reference frame in order to have one internuclear axe in the x,z-plane, so that for example h_1 takes the form:

$$h_1 = c_{1,1}s + c_{1,2}p_x + c_{1,4}p_z \tag{12}$$

The two remaining σ-orbitals take the form:

$$h_i = c_{i,1}s + c_{i,2}p_x + c_{i,3}p_y + c_{i,4}p_z,\ i = 2,3. \tag{13}$$

3. The orthogonality assumption

We make the following assumption:

Orthogonality assumption: *The hybrids h_1, h_2, h_3 and h_π are mutually orthogonal.*

We then have the following result:

Lemma 4. *If the hybrids satisfy the orthogonality assumption then we have:*

$$c_{i,1}c_\pi + c_{i,4}\lambda_\pi = 0,\ i = 1,2,3. \tag{14}$$

Proof. Denoting by $\langle \cdot,\cdot \rangle$ the usual scalar product on the orbitals, we have:

$$\langle h_\pi, h_1 \rangle = c_{i,1}c_\pi \langle s,s \rangle + c_{i,4}\lambda_\pi \langle p_z, p_z \rangle = 0,\ i = 1,2,3, \tag{15}$$

using the fact that $\langle p_z, p_x \rangle = \langle p_z, p_y \rangle = 0$. Moreover as $\langle s,s \rangle = 1$ and $\langle p_z, p_z \rangle = 1$ this concludes the proof. □

As a consequence, there exists μ_i such that:

$$c_{i,1} = \mu_i \lambda_\pi,\ c_{i,4} = -\mu_i c_\pi,\ \ i = 1,2,3, \tag{16}$$

As the angle is the same between p_z and each h_i, we must have:

$$\mu_1 = \mu_2 = \mu_3 = \mu \tag{17}$$

The σ-orbitals look like:

$$\begin{aligned} h_1 &= \mu(\lambda_\pi s - c_\pi p_z) + c_{1,2}p_x, \\ h_2 &= \mu(\lambda_\pi s - c_\pi p_z) + c_{2,2}p_x + c_{2,3}p_y, \\ h_3 &= \mu(\lambda_\pi s - c_\pi p_z) + c_{3,2}p_x + c_{3,3}p_y \end{aligned} \tag{18}$$

4. The normalization assumption

First we impose that the π-orbital is normalize, then we must have:

$$c_\pi^2 + \lambda_\pi^2 = 1 \tag{19}$$

Next, we impose that the contribution of the s orbital and p_z orbital is normalized. We then obtain:

Lemma 5. *The normalization assumption of the s and p_z contribution of (h_1, h_2, h_3, h_π) gives:*

$$\begin{aligned} h_1 &= \frac{1}{\sqrt{3}}(\lambda_\pi s - c_\pi p_z) + c_{1,2}p_x, \\ h_2 &= \frac{1}{\sqrt{3}}(\lambda_\pi s - c_\pi p_z) + c_{2,2}p_x + c_{2,3}p_y, \\ h_3 &= \frac{1}{\sqrt{3}}(\lambda_\pi s - c_\pi p_z) + c_{3,2}p_x + c_{3,3}p_y \end{aligned} \tag{20}$$

See the proof of the lemma in the appendix.

5. The sp^3 normalization assumption

R. C. Haddon impose in[3] what he called an **sp^3 normalization assumption** that can be translate as follows:

$$c_{1,2}^2 + c_{2,2}^2 + c_{3,2}^2 = 1, \text{ and } c_{2,3}^2 + c_{3,3}^2 = 1, \tag{21}$$

which are the classical constraints in the **pure sp^3** case.
This assumption is not sufficient to fix the value of the coefficients $c_{1,2}$, $c_{2,2}$, $c_{3,2}$, $c_{2,3}$ and $c_{3,3}$. However, imposing some special constraints like a C_{3v} symmetry, we can obtain explicit quantities (see section II B 7).

6. Computation of the POAV vector

It remains to give an explicit formula for the coefficient c_π and λ_π determining the POAV vector. These coefficients are related via the normalization. The explicit value must depend on the pyramidalization angle. This is indeed the case.

We have:

Lemma 6. *The coefficient c_π is given by:*

$$c_\pi = \sqrt{3} c_{1,2} \tan(Pyr(A)) \tag{22}$$

The coefficient λ_π is given by:

$$\lambda_\pi = \sqrt{1 - 3c_{1,2}^2 \tan^2(Pyr(A))} \tag{23}$$

Proof. The POAV vector is such that the angle between the h_π orbital and each orbital h_1, h_2, h_3 is $\frac{\pi}{2} + Pyr(A)$. The vector associated to h_π is $v_\pi = (0,0,\lambda_\pi)$ and the one with h_1 is $v_1 = (c_{1,2}, 0, \frac{-c_\pi}{\sqrt{3}})$. The scalar product between the two vectors gives $v_\pi \cdot v_1 = -\dfrac{1}{\sqrt{3}} c_\pi \lambda_\pi$.
This scalar product is also equal to $\| v_\pi \| \| v_1 \| \cos(v_\pi, v_1)$.

As $\| v_\pi \|^2 = \lambda_\pi^2$ and $\| v_1 \|^2 = \frac{3c_{1,2}^2 + c_\pi^2}{3}$, we obtain taking the square on each side of the equality:

$$\frac{\lambda_\pi^2 c_\pi^2}{3} = \frac{3c_{1,2}^2 + c_\pi^2}{3} \lambda_\pi^2 \cos^2\left(\frac{\pi}{2} + Pyr(A)\right) \tag{24}$$

We then have:

$$\lambda_\pi^2 \left[c_\pi^2 - (3c_{1,2}^2 + c_\pi^2) \cos^2\left(\frac{\pi}{2} + Pyr(A)\right)\right] = 0 \tag{25}$$

As $\lambda_\pi \neq 0$, we obtain:

$$c_\pi^2 \left(1 - \cos^2\left(\frac{\pi}{2} + Pyr(A)\right)\right) = 3c_{1,2}^2 \cos^2\left(\frac{\pi}{2} + Pyr(A)\right) \tag{26}$$

Using the fact that:

$$1 - \cos^2(\tfrac{\pi}{2} + Pyr(A)) = \sin^2(\tfrac{\pi}{2} + Pyr(A))$$

and the equality:

$$\cos(\tfrac{\pi}{2} + Pyr(A)) = -\sin(Pyr(A)) \text{ and}$$

$$\sin(\tfrac{\pi}{2} + Pyr(A)) = \cos(Pyr(A)),$$

we deduce that:

$$c_\pi^2 \cos^2(Pyr(A)) = 3c_{1,2}^2 \sin^2(Pyr(A)) \tag{27}$$

Assuming that $Pyr(A) \neq \frac{\pi}{2}$, we obtain:

$$c_\pi^2 = 3c_{1,2}^2 \frac{\sin^2(Pyr(A))}{\cos^2(Pyr(A))} = 3c_{1,2}^2 \tan^2(Pyr(A)), \tag{28}$$

which concludes the proof. □

The expression of the coefficient c_π was computed by R. C. Haddon and L. T. Scott[2] (p.138) in the particular case where the molecule possesses a local C_{3v} symmetry.

7. Special case: Hybridization for molecules with a C_{3v} symmetry

R. C. Haddon[3] made a symmetry assumption, called the C_{3v} symmetry which is implicitly used in[2] and which implies the following constraints in the adapted reference frame:

$$c_{2,2} = c_{3,2} < 0, \quad c_{2,3} = -c_{3,3}, \tag{29}$$

and the fact that we have three σ_v planes of symmetry in the plane (p_x, p_y) leads to:

$$c_{2,2}^2 + c_{2,3}^2 = c_{1,2}^2, \tag{30}$$

which corresponds to the fact that we have an equilateral triangle with $c_{1,2} > 0$.

The σ-hybrids are then given by:

$$\begin{aligned} h_1 &= \frac{1}{\sqrt{3}}(\lambda_\pi s - c_\pi p_z) + c_{1,2} p_x, \\ h_2 &= \frac{1}{\sqrt{3}}(\lambda_\pi s - c_\pi p_z) + c_{2,2} p_x + c_{2,3} p_y, \\ h_3 &= \frac{1}{\sqrt{3}}(\lambda_\pi s - c_\pi p_z) + c_{2,2} p_x - c_{2,3} p_y \end{aligned} \tag{31}$$

with the condition (30). Coupling this assumption with the **sp^3** normalization conditions, we obtain the following form for the σ-hybrids:

Lemma 7 (**sp^3** normalization)**.** *The sp^3 normalization assumption of (h_1, h_2, h_3, h_π) gives:*

$$\begin{aligned} h_1 &= \frac{1}{\sqrt{3}}(\lambda_\pi s - c_\pi p_z) + \frac{\sqrt{2}}{\sqrt{3}} p_x, \\ h_2 &= \frac{1}{\sqrt{3}}(\lambda_\pi s - c_\pi p_z) - \frac{1}{\sqrt{6}} p_x - \frac{1}{\sqrt{2}} p_y, \\ h_3 &= \frac{1}{\sqrt{3}}(\lambda_\pi s - c_\pi p_z) - \frac{1}{\sqrt{6}} p_x + \frac{1}{\sqrt{2}} p_y \end{aligned} \tag{32}$$

Proof. This follows from $c_{2,3}^2 + c_{3,3}^2 = 2c_{2,3}^2 = 1$, which gives for example $c_{2,3} = \pm\frac{1}{\sqrt{2}}$ and $c_{3,3} = \mp\frac{1}{\sqrt{2}}$.

The second normalization gives $c_{1,2}^2 + 2c_{2,2}^2 = 1$.

The symmetry condition (30) with $c_{2,3}^2 = \frac{1}{2}$ leads to the relation $c_{2,2}^2 + \frac{1}{2} = c_{1,2}^2$. Putting this expression in the normalization equation gives $c_{2,2}^2 + \frac{1}{2} + 2c_{2,2}^2 = 1$.

We deduce that $3c_{2,2}^2 = \frac{1}{2}$ or $c_{2,2}^2 = \frac{1}{6}$ and so $c_{1,2}^2 = \frac{2}{3}$. □

The previous form of the σ-hybrids coincide with the ones given by R. C. Haddon and L. T. Scott[2] (p.138).

8. Relative weight for orbitals and hybridization numbers

a. A notion of relative weight for orbitals For a given orbital h_i, one can be interested in measuring which part of the constituting orbitals is playing a major role with respect to the others. We then introduce a notion of **weight** for a given orbital and of **relative weight** between the constitutive elements of an orbital.

Definition 6 (Weight of a given atomic orbital)**.** *Let h be an orbital of the form:*

$$h = a_s s + a_x p_x + a_y p_y + a_z p_z, \tag{33}$$

where (a_s, a_x, a_y, a_z) are real coefficients. The weight of the atomic orbital $p_\star$, $\star \in \{s, x, y, z\}$ in the molecular orbital h denoted by $w_\star(h)$ is the quantity $w_\star(h) = a_\star^2$.

A comparison between the contribution of each orbitals is then resumed by the notion of *relative weight* between two orbitals.

Definition 7 (Relative weight for orbitals)**.**

$$h = a_s s + a_x p_x + a_y p_y + a_z p_z, \tag{34}$$

where (a_s, a_x, a_y, a_z) are real coefficients. We call (s,x)-relative weight of the atomic orbital s in the molecular orbital h in regards of the weight of the atomic orbital x in the same orbital h and we denote by $w_{s,x}$ the quantity:

$$w_{s,x}(h) = \frac{w_s(h)}{w_x(h)}, \tag{35}$$

as long as $w_x(h) \neq 0$. The relative weights for y, z are defined in the same way.

A more global weight can be defined denoting by $p = a_x p_x + a_y p_y + a_z p_z$. In this case, the (s,p)-relative weight is denoted by:

$$w_{s,p} = \frac{w_s(h)}{w_x(h) + w_y(h) + w_z(h)} \tag{36}$$

A main property of the relative weight is that they are invariant when the orbital is multiplied by a scalar $\mu \in \mathbb{R}$. The relative weights are naturally related to the hybridization numbers introduced by R. C. Haddon[2].

The notion of weight for a given orbital can be extended to a finite family of orbitals representing a particular local geometry.

Definition 8 (Relative weight of a system of orbitals)**.** *Let $\mathfrak{h} = (h_i)_{i=1,\dots,4}$ be a family of orbitals. The $(s,\star)$-relative weight of $\mathfrak{h}$ is define by:*

$$w_{s,\star}(\mathfrak{h}) = \sum_{i=1}^{4} w_{s,\star}(h_i), \tag{37}$$

where $\star \in \{x, y, z\}$.

b. Hybridization numbers R. C. Haddon and L. T. Scott[2] (p.2938) introduced two numbers in order to characterize the relative weight of each orbitals in the system (32) defined by h_π and $\mathfrak{h}_\sigma = (h_1, h_2, h_3)$ and to illustrate how a system with a given pyramidalization number evolves between the sp^2 and sp^3 case.

We first rewrite the set of orbitals as follows:

$$h_\pi = N_\pi(s + \tilde{\lambda}_\pi p_\pi), \;\; h_i = N_i(s + \lambda_i p_i), \; i = 1,2,3, \tag{38}$$

where $\| p_i \| = 1$ for $i = 1,2,3$ and

$$\tilde{\lambda}^2 = \frac{\lambda_\pi^2}{c_\pi^2}. \tag{39}$$

We then define the two quantities :

$$m = \frac{1}{\tilde{\lambda}_\pi^2}, \tag{40}$$

and

$$n = \tilde{\lambda}_1^2 + \tilde{\lambda}_2^2 + \tilde{\lambda}_3^2 \tag{41}$$

We see that these hybridization numbers correspond to what we have called relative weight of orbitals, and precisely:

$$m = w_{s,z}(h_\pi) = \frac{w_s(h_\pi)}{w_z(h_\pi)} = \frac{1}{\tilde{\lambda}_\pi^2} = \frac{c_\pi^2}{\lambda_\pi^2} \tag{42}$$

and

$$n = w_{z,s}(\mathfrak{h}_\sigma) + w_x(\mathfrak{h}_\sigma) + w_y(\mathfrak{h}_\sigma) \tag{43}$$

In the case where the system $(h_\pi, \mathfrak{h}_\sigma)$ is given by (32), we have:

$$m = \frac{2\tan^2(Pyr(A))}{1 - 2\tan^2(Pyr(A))} = \frac{c_\pi^2}{\lambda_\pi^2} \tag{44}$$

Moreover, we have:

$$w_{z,s}(h_1) = w_{z,s}(h_2) = w_{z,s}(h_3) = \frac{(\frac{c_\pi}{\sqrt{3}})^2}{(\frac{\lambda_\pi}{\sqrt{3}})^2} = \frac{c_\pi^2}{\lambda_\pi^2} = m, \tag{45}$$

and

$$w_x(\mathfrak{h}_\sigma) = w_y(\mathfrak{h}_\sigma) = 1, \tag{46}$$

so that:

$$n = 3m + 2. \tag{47}$$

Explicit computation depending on the pyramidalization number shows the path between sp^2 and sp^3, see table I. The previous number can then be understood as hybridization of the form:

$$s^{\lambda_\pi{}^2} p_{x,y}^{n-3m} p_z^{c_\pi^2} \tag{48}$$

for the σ orbitals (h_i) and

$$s^{c_\pi^2} p_z^{\lambda_\pi^2} \tag{49}$$

for the $\pi - orbitals$ (h_π) .

The classical sp^2 situation is then understood in term of $(sp^2_{x,y}p^0_z)_\sigma(s^0p^1_z)_\pi = (sp^2)_\sigma(p^1_z)_\pi$.

The classical sp^3 situation corresponds to $\lambda_\pi = 0.866$ and $c_\pi = \frac{1}{2}$ i.e. $(s^{0.75}p^{2.00}_{x,y}p^{0.25}_z)_\sigma = (s^{0.75}p^{2.25})_\sigma(s^{0.25}p^{0.75})_\pi$.

Hybridization is then well defined. As an illustration, calculations were carried out on the C_{32} and C_{80} that have allowed the construction of the cartographies in the figure 4. Note that, the cartographies are sorted according to the energetic order of the isomers, starting from the most stable and the C_{80} hybridization values are given in relation to the mean C_{60} value. However, in 2001, R. C. Haddon was forced to clarify his definition of hybridization[6]. If, at first glance, the definitions give in his paper seem different from equations 48 and 49, we easily show that they are, at the end, identical in all respects.

For R. C. Haddon, the hybridization of the π-orbital ($s^m p$) is obtained from the value of the m parameter. To relate this definition to equation 49, it is just sufficient to express the weight of the s-orbital in the hybrid π-orbital (h_π) in proportion to that of a single p_z-orbital:

$$s^{c_\pi^2} p_z^{\lambda_\pi^2} \iff s^{\frac{c_\pi^2}{\lambda_\pi^2}} p_z \tag{50}$$

From equation 44, the ratio $\frac{c_\pi^2}{\lambda_\pi^2}$ explicitly corresponds to the m parameter.

According to R. C. Haddon[6], the average hybridization of the σ-orbitals ($sp^{\tilde{n}}$) is given by:

$$\tilde{n} = 3m + 2 \tag{51}$$

To link this definition to equation 48, it is sufficient to express the weight of the p-orbitals in the σ-orbitals (h_i) in proportion to that of a single s-orbital:

$$s^{\lambda_\pi^2} p^{2+c_\pi^2} \iff sp^{\frac{2+c_\pi^2}{\lambda_\pi^2}} \iff sp^{\frac{2}{\lambda_\pi^2}+\frac{c_\pi^2}{\lambda_\pi^2}} \iff sp^{\frac{2}{\lambda_\pi^2}+m} \tag{52}$$

$$sp^{3m+2} = sp^{\tilde{n}} \tag{53}$$

Proof. We know that :

$$m = \frac{c_{\pi^2}}{\lambda_\pi^2} = \frac{1-\lambda_\pi^2}{\lambda_\pi^2} = \frac{1}{\lambda_\pi^2} - 1$$
$$then,$$
$$3m + 2 = \frac{2}{\lambda_\pi^2} + m = \tilde{n}$$

□

These definitions allow us to get closer to the notion of the σ-hybridization used by the chemists. Figure 5 illustrates this σ hybridization as a function of the number of atoms for the complete basis of Tománek. As normally expected, the highest hybridizations are observed for the smaller fullerenes. What is less intuitive is both the very large disparity of these values and a large sp > 3 situations (i.e. $Pyr(A) > 109.5°$) for all medium size systems below C_{60}. Because the notion of sp > 3 is not physically well-defined, the use of the σ-hybridization is therefore not the most suitable for cartographied the hybridization as a function of the curvature of carbon systems. A most relevant information which reflects the notion of hybridization seems to be the one contained in equation 48. In this formula, the weight of the p-orbitals in the σ-system is equal to $2 + c_\pi^2$. In particular, the weight of the p_z orbital distributed in each of the 3 directions h_1, h_2, and h_3 is equivalent to $\frac{c_\pi^2}{3}$. This new parameter, denoted Y, is also related to m parameter, as $Y = m\frac{\lambda_\pi^2}{3}$. It mainly serves to describe the deviation from the conjugation of a system. Last, but not least, Y quantifies the participation of the p_z orbital (hybridization) on each h_i orbital in consideration of the topology of the system ($Pyr(A)$) (figures 6 and 7). All right, it is always permissible to rewrite the hybridization in term of Y without any changes for the chemists :

$$sp^{\tilde{n}} = sp^{\frac{2}{\lambda_\pi^2}+m} = sp^{\frac{2+3Y}{\lambda_\pi^2}} \tag{54}$$

Figure 7 shows the cartographies of C_{32} and C_{80} relative to the Y parameter. The scale have been adapted following the observations of figure 6. Here again, the energy order of the different symmetries is respected.

C. A generalization of the pyramidalization angle for general molecules

In order to generalize the Haddon's pyramidalization angle for a general molecule M, meaning not restricted to trivalent molecules, we need to be able to define for each atom of M the analogue of the $POAV(A)$ vector. We then restrict our attention to a special class of molecules called **admissible molecules**.

Definition 9 (Admissible molecule)**.** *A molecule M is said to be admissible if for each atom A of M and $\varepsilon > 0$, there exists a plane $\mathscr{P}_\varepsilon(A)$ such that $Reg_\varepsilon(\star(A)) \subset \mathscr{P}_\varepsilon(A)$.*

TABLE I: The pyramidalization path between sp^2 and sp^3

θ	$Pyr(A)$	c_π^2	λ_π^2	m	n	Hybridization
90.0	0.0	0.000	1.000	0.000	2.000	$(s^{1.000}p_{x,y}^{2.000}p_z^{0.000})_\sigma(s^{0.000}p_z^{1.000})_\pi$
92.0	2.0	0.002	0.998	0.002	2.007	$(s^{0.998}p_{x,y}^{2.000}p_z^{0.002})_\sigma(s^{0.002}p_z^{0.998})_\pi$
94.0	4.0	0.010	0.990	0.010	2.030	$(s^{0.990}p_{x,y}^{2.000}p_z^{0.001})_\sigma(s^{0.001}p_z^{0.999})_\pi$
96.0	6.0	0.022	0.978	0.023	2.068	$(s^{0.978}p_{x,y}^{2.000}p_z^{0.022})_\sigma(s^{0.022}p_z^{0.978})_\pi$
98.0	8.0	0.040	0.960	0.041	2.123	$(s^{0.960}p_{x,y}^{2.000}p_z^{0.040})_\sigma(s^{0.040}p_z^{0.960})_\pi$
100.0	10.0	0.062	0.938	0.066	2.199	$(s^{0.938}p_{x,y}^{2.000}p_z^{0.062})_\sigma(s^{0.062}p_z^{0.938})_\pi$
102.0	12.0	0.090	0.910	0.099	2.298	$(s^{0.910}p_{x,y}^{2.000}p_z^{0.090})_\sigma(s^{0.090}p_z^{0.910})_\pi$
104.0	14.0	0.124	0.876	0.142	2.426	$(s^{0.876}p_{x,y}^{2.000}p_z^{0.124})_\sigma(s^{0.124}p_z^{0.876})_\pi$
106.0	16.0	0.164	0.836	0.197	2.590	$(s^{0.836}p_{x,y}^{2.000}p_z^{0.164})_\sigma(s^{0.164}p_z^{0.836})_\pi$
107.0	17.0	0.187	0.813	0.230	2.690	$(s^{0.813}p_{x,y}^{2.000}p_z^{0.187})_\sigma(s^{0.187}p_z^{0.813})_\pi$
108.0	18.0	0.211	0.789	0.268	2.803	$(s^{0.789}p_{x,y}^{2.000}p_z^{0.211})_\sigma(s^{0.211}p_z^{0.789})_\pi$
109.0	19.0	0.237	0.763	0.311	2.932	$(s^{0.763}p_{x,y}^{2.000}p_z^{0.237})_\sigma(s^{0.237}p_z^{0.763})_\pi$
109.5	19.5	0.250	0.750	0.333	3.000	$(s^{0.750}p_{x,y}^{2.000}p_z^{0.250})_\sigma(s^{0.250}p_z^{0.750})_\pi$
110.0	20.0	0.265	0.735	0.360	3.081	$(s^{0.735}p_{x,y}^{2.000}p_z^{0.265})_\sigma(s^{0.265}p_z^{0.735})_\pi$
112.0	22.0	0.326	0.674	0.485	3.454	$(s^{0.674}p_{x,y}^{2.000}p_z^{0.326})_\sigma(s^{0.326}p_z^{0.674})_\pi$
114.0	24.0	0.396	0.604	0.657	3.971	$(s^{0.604}p_{x,y}^{2.000}p_z^{0.396})_\sigma(s^{0.396}p_z^{0.604})_\pi$
116.0	26.0	0.476	0.524	0.908	4.723	$(s^{0.524}p_{x,y}^{2.000}p_z^{0.476})_\sigma(s^{0.476}p_z^{0.524})_\pi$

The geometry of the previous conditions can be resumed as follows: If the atom A does not belong to the same plane as the atoms of $\mathrm{Reg}_\varepsilon(\star(A))$, then it belongs to the intersection of all the mediating plane of the line segment between two arbitrary atoms of $\mathrm{Reg}_\varepsilon(\star(A))$. As the intersection of two planes already define a line, it gives huge constraints on the positioning of each atom of $\mathrm{Reg}_\varepsilon(\star(A))$, in particular that they belong to a circle C_ε in the plane $\mathscr{P}_\varepsilon(A)$. If we denote by O_ε the center of C_ε, then A belongs to the line passing trough O_ε and normal to $\mathscr{P}_\varepsilon(A)$.
The previous remark can be used to define the $POAV(A)$ vector in this general situation:

Definition 10 (POAV)**.** *Let M be an admissible molecule and A be an atom of M. Let $\varepsilon > 0$ be given. We denote by C_ε the circle in $\mathscr{P}_\varepsilon(A)$ such that $Reg_\varepsilon(\star(A)) \subset C_\varepsilon$. We denote by O_ε the center of C_ε. We denote by $POAV(A)$ the normal unitary vector to $\mathscr{P}_\varepsilon(A)$ defined by:*

$$POAV(A) = \frac{\vec{O_\varepsilon A}}{O_\varepsilon A}. \tag{55}$$

As usual, the $POAV$ vector is defined trough quantities depending on ε so that it is not a priori trivial that it gives a well defined quantity. However, by construction, the set of points $\mathrm{Reg}_\varepsilon(\star(A))$ and $\mathrm{Reg}_{\varepsilon'}(\star(A))$ for two different ε, ε' are **homothetic** as Reg_ε is a **homothety** of center A for all ε. In particular, homothety are **affine transformations** which send a line to a parallel line. As $\mathrm{Reg}_\varepsilon(\star(A))$ belongs to a plane $\mathscr{P}_\varepsilon(A)$, then $\mathrm{Reg}_{\varepsilon'}(\star(A))$ belongs to a plane $\mathscr{P}_{\varepsilon'}(A)$ parallel to $\mathscr{P}_\varepsilon(A)$ and have the same normal vector. By normalization, we obtain a unique vector $POAV(A)$ normal to all the plane $\mathscr{P}_\varepsilon(A)$ for $\varepsilon > 0$.
Having the $POAV$ vector, one can directly generalize the notion of pyramidalization angle:

Definition 11 (Pyramidalization angle)**.** *Let M be an admissible molecule and A be a given atom of M. The pyramidalization angle in A denoted by $Pyr(A)$ is the angle between the vector $POAV(A)$ and each bonds AB, $B \in \star(A)$.*

The proof that this angle is well defined follows the same line as in the trivalent case. Let $\varepsilon > 0$ be given and denote by $z_{A,\varepsilon}$ the distance $O_\varepsilon A$. Then, we have for a given bond AB, $B \in \star(A)$ that:

$$\sin(Pyr(A)) = \frac{z_{A,\varepsilon}}{\varepsilon} \tag{56}$$

This quantity does not depend on B and is in fact independent of ε.
Indeed, let us consider $\varepsilon' = \lambda\varepsilon$, then $z_{A,\varepsilon'} = \lambda z_{A,\varepsilon}$ and $\frac{z_{A,\varepsilon'}}{\varepsilon'} = \frac{z_{A,\varepsilon}}{\varepsilon}$.

The previous definition is of course far from being satisfying as it works for a very restrictive class of molecules. We discuss more precisely the restriction associated to the use of the pyramidalization angle in the next section.

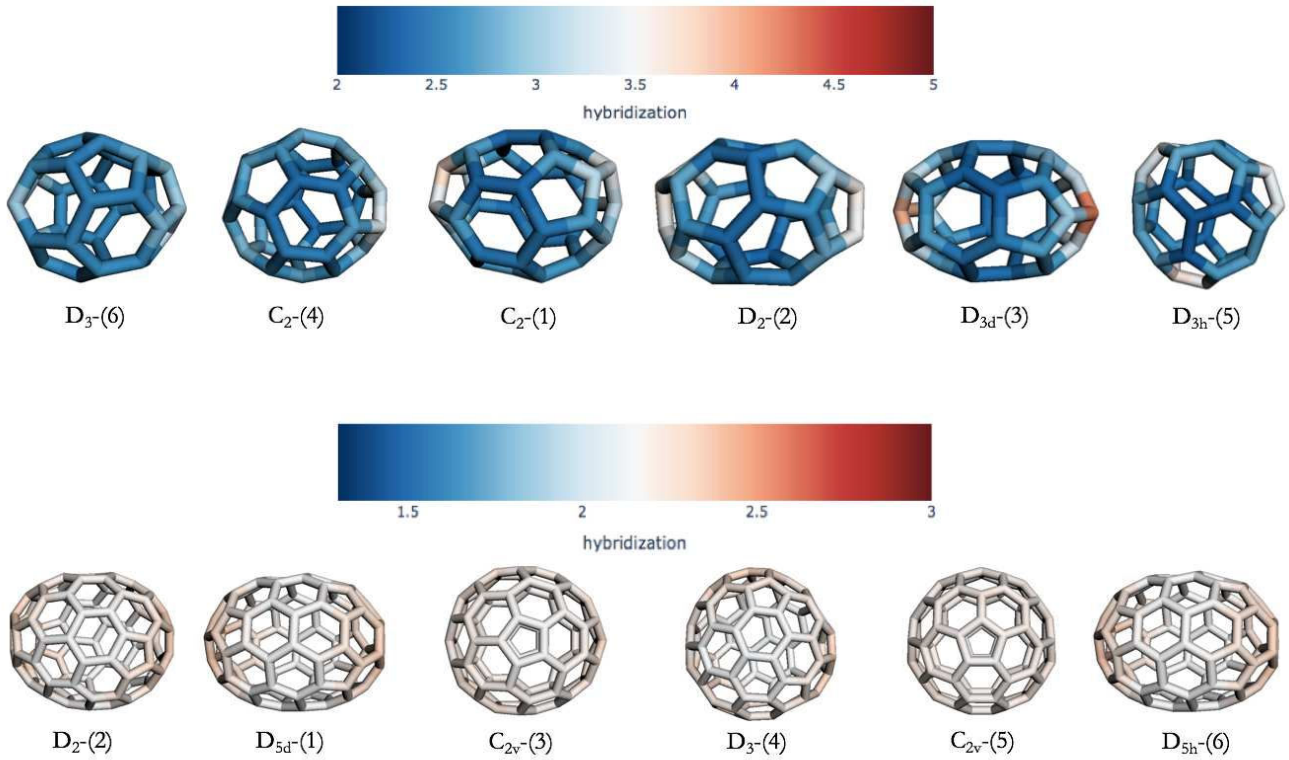


FIG. 4: Cartographies of C_{32} and C_{80} for the hybridization. The value reported in white corresponds to the average hybridization value in the C_{60} system.[4]

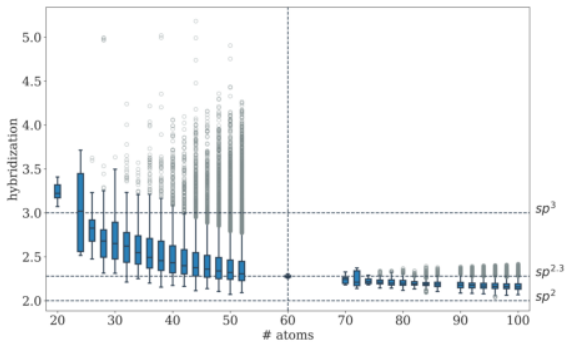


FIG. 5: Boxplot of the hybridization, as a function of the number of atom in the isomer. Dashed lines indicate values of the C_{60} isomers and the boundary between a sp^2 and sp^3 situations.

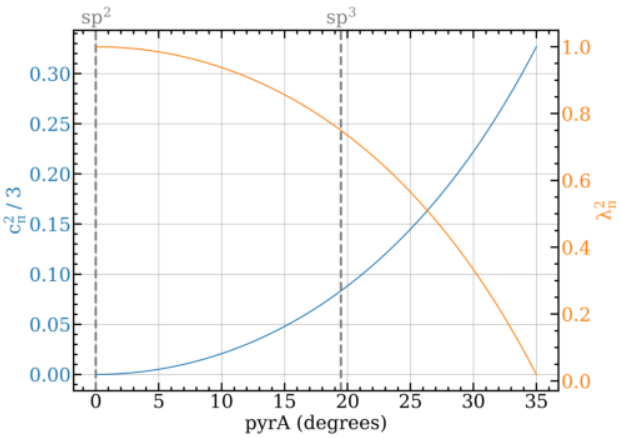


FIG. 6: Representation of the Y parameter and λ_π^2 as a function of the pyramidalization angle (in degrees).

D. Limitations of the pyramidalization angle

The definition of the pyramidalization angle is connected with the existence of the $POAV(A)$ vector at each atom A of the molecule. The explicit construction of the $POAV(A)$ vector for general trivalent molecules can be used to prove that **the pyramidalization angle can not be define for arbitrary molecules**. More precisely, the pyramidalization angle can not be extend outside the family of admissible molecules.

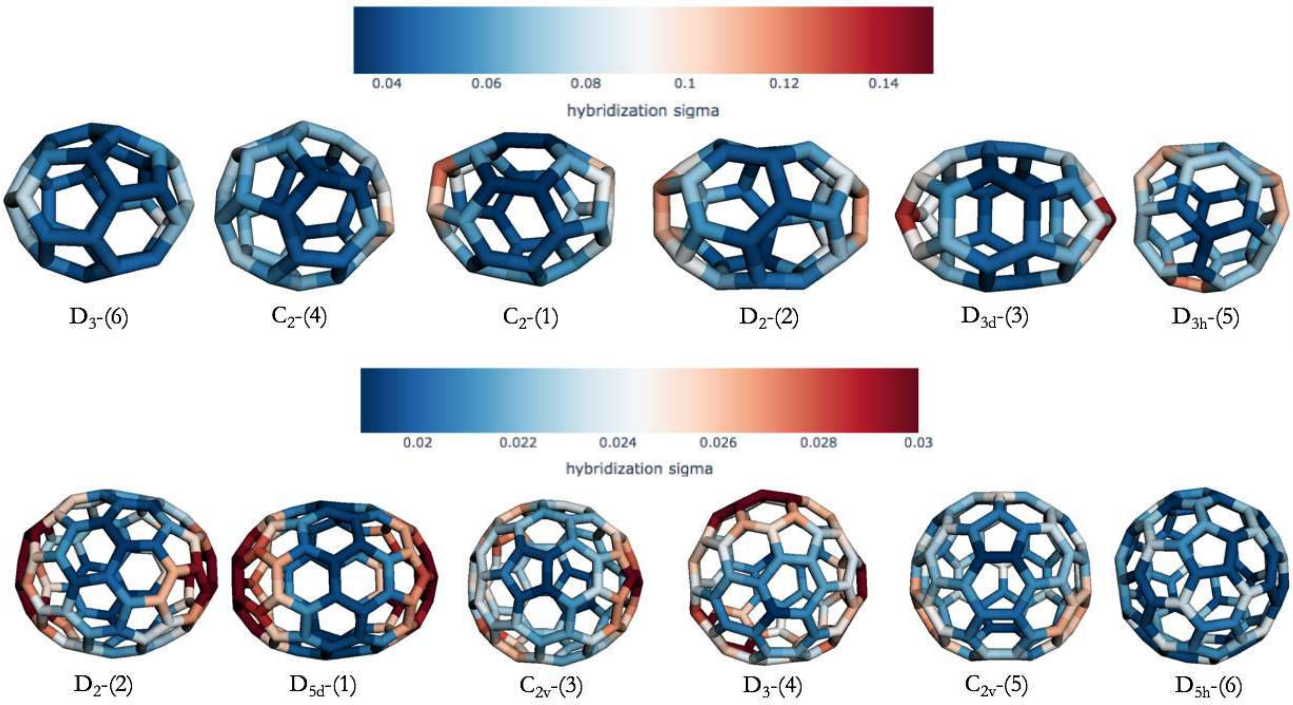


FIG. 7: Cartographies of C_{32} and C_{80} for the Y parameter ($\frac{c_\pi^2}{3}$ hybridization sigma).[4] Note that C_{80} presents a weak hybridization compared to C_{32} thus the scale of the cartographies had to be adapted.

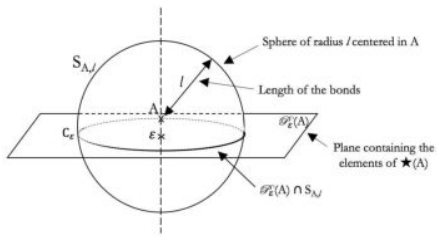


FIG. 8: Geometry of an admissible molecule.

1. Indeed, the existence of a $POAV(A)$ vector implies that $POAV(A)$ has to be normal to any plane $\mathscr{P}_\varepsilon(A)$. In fact, if $POAV(A)$ exists and is different from n_A, the unique unitary vector to all plane $\mathscr{P}_\varepsilon(A)$, then it makes a given angle $\alpha > 0$ with n_A. The $POAV(A)$ vector then generate a cone based in A by rotation around the axes n_A corresponding to all the vectors v such that v makes an angle α with the normal axe. As a consequence, for each bond corresponds a different vector, except when $\alpha = 0$ corresponding to the normal vector (figure 9).

2. From now, we assume that $POAV(A)$ is normal to any $\mathscr{P}_\varepsilon(A)$. For each triple choice T of elements in $\mathrm{Reg}_\varepsilon(star(A))$, the explicit construction of the $POAV(A)$ vector in section II A 2 gives a unique vector $POAV_T(A)$ which is colinear to the vector $I_T A$ where I_T is the intersection of the mediating line in the triangle defined by T. In general, other choice of triples will lead to different I_T. As A must belong to each line perpendicular to $\mathscr{P}_\varepsilon(A)$ and passing trough I_T for each T, this is not possible in general unless all the I_T coincide meaning that all the atoms of $\mathrm{Reg}_\varepsilon(\star(A))$ belong to a circle.

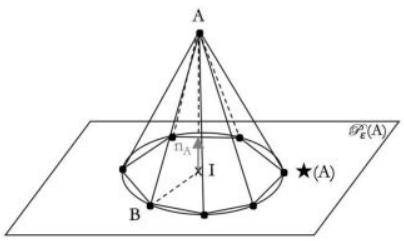


FIG. 9: Representation of the pyramidalization angle .

The two characteristics correspond to the definition of admissible molecules.
As a consequence, one must think to other characterizations of the local geometry of a molecule in order to study molecules of order ≥ 4 in some atoms.

The $Pyr(A)$ was studied for the C_{32} and C_{80} which allowed us to carry out the cartographies figure 10. As in the case of hybridization, the cartographies are sorted according to the energetic order of the isomers, starting from the most stable and the C_{80} $Pyr(A)$ values are given in relation to the mean C_{60} value.

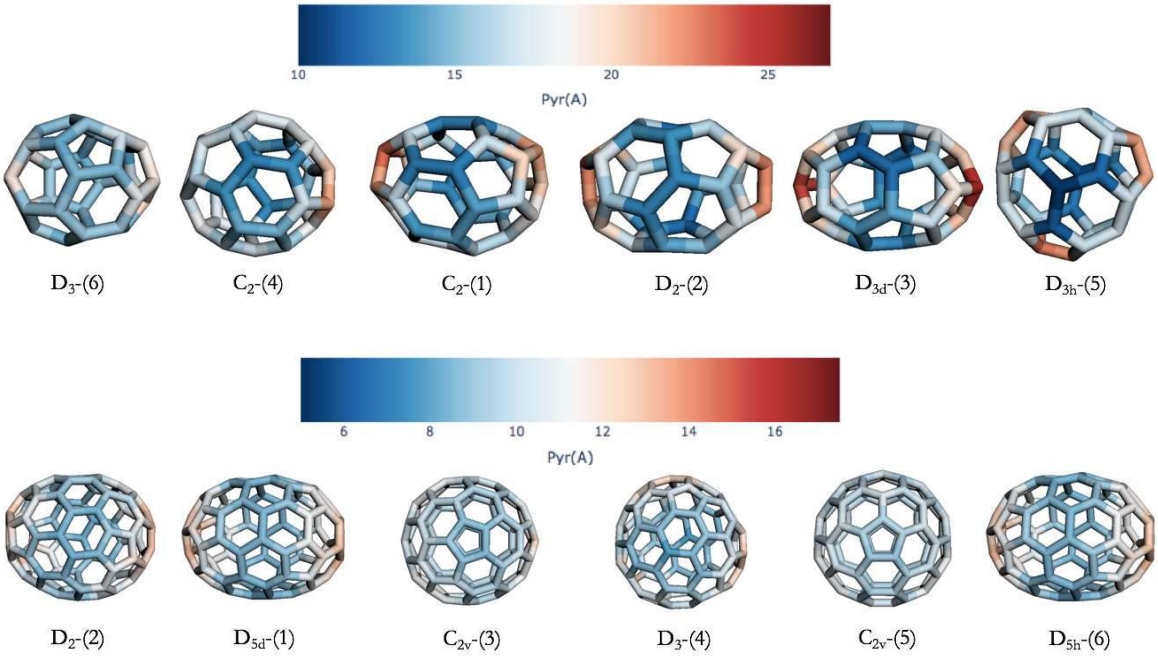


FIG. 10: Cartographies of C_{32} and C_{80} for the pyramidalization angle[4]

III. HADDON'S CURVATURE FOR TRIVALENT MOLECULES

In "C_{60} : Sphere or polyhedron ?"[1], R. C. Haddon introduced the notion of **spherical curvature** as an answer to many articles interchanging the term of pyramidalization angle and curvature to describe the shape of the molecule. He first noted that the word **curvature** is well defined in classical differential geometry for smooth surfaces. The corresponding notion for discrete (polyhedral) surfaces is more complex. We give a complete construction of the spherical curvature following Haddon's work[1] and we provide the connection with the pyramidalization angle.

A. Osculating sphere and spherical curvature

For a smooth curve in $\mathbb{R}^2$, we have a natural notion of curvature which is related to the notion of **osculating circle**. The basic idea is to approximate locally up to order 2 the curve by an arc of a circle. The curvature is then given by $\frac{1}{R}$ where R is the radius of the circle. When one consider a surface in $\mathbb{R}^3$, a natural idea is to approximate locally the surface by a portion of a sphere. However, it is not in general possible to construct such a sphere.

In his work[1], R. C. Haddon stated the following result:

Lemma 8 (Osculating sphere)**.** *Let M be a molecule satisfying assumption (A) and trivalent. For each atom A, there exists a unique sphere $S(A)$ such that $A \in S(A)$ and for all $B \in \star(A)$, $B \in S(A)$. The sphere $S(A)$ is called the osculating sphere to M in A.*

The proof is **constructive** and can be used in concrete examples.

Proof. Let A be a given atom of M and B_1, B_2 and B_3 the three atoms of $\star(A)$. Let us denote by $\mathscr{P}_A$ the plane containing B_1, B_2 and B_3. The three mediating planes to the line segments B_1B_2, B_2B_3 and B_3B_1 intersect along a line L orthogonal to $\mathscr{P}_A$. As all the bonds are equal, the three mediating planes to the line segments AB_1, AB_2 and AB_3 intersect the line L in a point O which by definition is the center of the sphere passing trough A and the elements of $\star(A)$. By construction, this sphere is unique. □

The previous result allows us to define the notion of **spherical curvature**.

Definition 12. *(Spherical curvature) Let M be a molecule satisfying assumption (A) and trivalent. For each atom A, we denote by $\kappa(A)$ and we call spherical curvature to M at point A the quantity:*

$$\kappa(A) = \frac{1}{R_A}, \tag{57}$$

where R_A is the radius of the osculating sphere to M in A.

We now give an explicit connection between the pyramidalization angle and the spherical curvature.

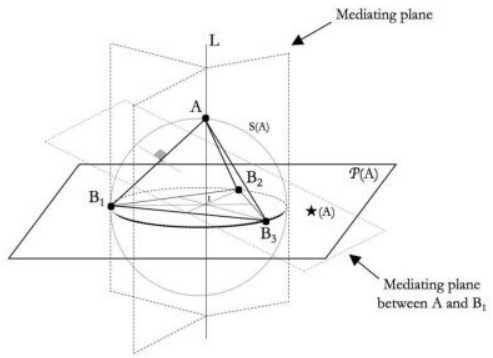


FIG. 11: Osculating sphere

B. Spherical curvature versus the pyramidalization angle

The connection between the pyramidalization angle and the spherical curvature is very simple although nonlinear.

Lemma 9. *Let M be a regular trivalent molecule. We denote by a the uniform bond lengths. For each atom A, we denote by $\kappa(A)$ the spherical curvature and $Pyr(A)$ the pyramidalization angle to M in A. Then, we have:*

$$\kappa(A) = \frac{2\sin(Pyr(A))}{a}. \tag{58}$$

Proof. Let $B \in \star(A)$ and I the middle point between A and B. Let us denote by O the center of the osculating sphere of radius R_A and $\kappa(A) = \frac{1}{R_A}$. We have $AB = a$, $AI = \frac{a}{2}$ and $AO = R_A$. The right triangle AIO is such that the angle $\widehat{OAI} = \frac{\pi}{2} - Pyr(A)$. As a consequence, we have $\cos\left(\frac{\pi}{2} - Pyr(A)\right) = \frac{AI}{AO} = \frac{a}{2R_A}$. As $\cos(\frac{\pi}{2} - Pyr(A)) = \sin(Pyr(A))$, this concludes the proof. □

When the pyramidalization angle is small, the previous relation becomes linear and we have:

$$\kappa(A) \sim \frac{2}{a} Pyr(A). \tag{59}$$

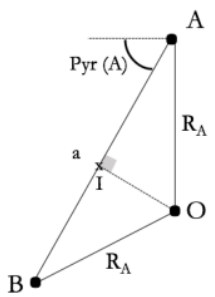


FIG. 12: Construction of the connection between the pyramidalization angle and the spherical curvature .

C. Extension to non regular molecules

We generalize the previous construction for trivalent molecules which are not regular, covering then real molecules.

The former construction of the osculating sphere was made easy by the fact that all the bonds between atoms were assumed to be equal. However, we can generalize this construction to the general case and we obtain:

Lemma 10 (Osculating sphere - General case)**.** *Let M be a trivalent molecule. Let A be a given atom of M and $\star(A) = (B,C,D)$. Let O be the intersection of the mediating line in the triangle BCD and $\mathscr{P}$ the plane define by $\star(A)$. Let $\mathscr{L}$ be the normal line to $\mathscr{P}$ passing trough O and $\mathscr{R}$ a reference frame centered in O where (x,y) are in $\mathscr{P}$ and z along $\mathscr{L}$. If $A \notin \mathscr{P}$, there exists a unique osculating sphere whose center $O_z = (0,0,z)$ is defined by:*

$$z = \frac{L^2 + z_A^2 - l^2}{2z_A}, \tag{60}$$

where z_A is the z coordinate of A in $\mathscr{R}$, $l = OB = OC = OD$ and $L = O_{z_A}A$.

The spherical curvature $\kappa(A)$ at the atom A is then given by:

$$\kappa(A) = \frac{1}{\sqrt{l^2 + \frac{(L^2 + z_A^2 - l^2)^2}{4z_A^2}}} = \frac{1}{\sqrt{l^2 + \frac{(OA^2 - l^2)^2}{4z_A^2}}}. \tag{61}$$

Proof. By definition, the line $\mathscr{L}$ corresponds to the set of points at equal distance of B, C and D. The center of the osculating sphere must belong to $\mathscr{L}$. The center O_z is then of the form $O_z = (0,0,z)$ with $z \in \mathbb{R}$ in the reference frame $\mathscr{R}$.

We look for a point O_z such that $O_zA = O_zB = O_zC = O_zD$. As $O_zB = O_zC = O_zD$ by construction, we have only to ensure $O_zA = O_zB$.
By the Pythagorean theorem, we have $O_zB^2 = z^2 + l^2$ and $O_zA^2 = O_zO_{z_A}^2 + O_{z_A}A^2 = L^2 + (z_A - z)^2 = L^2 + z_A^2 + z^2 - 2zz_A$ so that $O_zB^2 = O_zA^2$ implies $2zz_A = L^2 + z_A^2 - l^2$.
As $A \notin \mathscr{P}$ then $z_A \neq 0$ and $z = \frac{L^2 + z_A^2 - l^2}{2z_A}$.
This concludes the proof. □

In "C_{60} : Sphere or polyhedron ?"[1] the construction of the osculating sphere is not done. R. C. Haddon indicated only that one has to solve a "system of nonhomogeneous equations" and was incorporated in his program POAV3[10]. ù

D. Limitations of the spherical curvature

In the trivalent case, the spherical curvature provides a very nice way to characterize the non planarity of a given

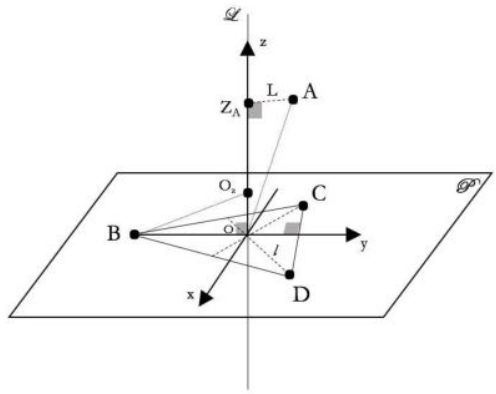


FIG. 13: Construction of the osculation sphere for non regular molecules .

molecule. However, the **generalization to molecules which are not trivalent** leads to severe difficulties and **is in general impossible**.
Indeed, let us consider a molecule which is of order 4 in some atoms. Let A be such an atom. Then $\star(A)$ is made of 4 points B, C, D and E. In order to generalize the notion of spherical curvature, one has to construct a sphere interpolating the 5 points. The geometric construction of the osculating sphere made in section III C for trivalent molecules determine a unique sphere interpolating 4 atoms. As a consequence, selecting any triple of points in $\star(A)$ denoted by T, one construct a unique osculating sphere in A with respect to T denoted by $S(A)_T$. In general, all these spheres are different and there exists no osculating sphere to the molecule in A.

IV. ANGULAR DEFECT AND GAUSS DISCRETE CURVATURE

We introduce the notion of **angular defect** and **Gauss discrete curvature** which are classical in discrete differential geometry. These quantities are always defined contrary to the pyramidalization angle or the spherical curvature and can be used as a convenient alternative quantity to measure the local shape of the molecule.

A. Gauss curvature of a surface

The **Gauss curvature** of a surface Σ at a point p can be defines as follows: consider a circle of radius r around p, meaning that for a given metric d on Σ, we look at the set of point $x \in \Sigma$ such that $d(x,p) = r$. We denote by $P(r)$ the circumference of this circle. The Gauss curvature is then obtained by comparing the classical circumference of a circle in the plane given by $2\pi r$ to the value of P on Σ. As an example, a circle of radius r at a point of a sphere of radius 1, one obtain $P(r) = 2\pi \sin(r)$.

The Gauss curvature $G(p)$ is then obtain by expending the previous quantity and we obtain:

$$P(r) = 2\pi r - G(p)\pi\frac{r^3}{3} + \dots \tag{62}$$

B. Angular defect at a vertex of a polyhedron

A convenient way to generalize the previous definition is to consider the quantity defined by:

$$K(p) = 2\pi - \frac{P(r)}{r} \tag{63}$$

If one consider a solid representation of a given molecule M, one see that the previous quantity is zero along the edges and for any point in a given face of M. As a consequence, the curvature is "concentrated" at the vertex of M.

One can verify that $K(p)$ is well defined and does not depend on r. Precisely, we have:

Theorem 1. *Let M be a molecule and A a given atom then*

$$K(A) = 2\pi - \sum_{F\in\star(A)} \alpha_F, \tag{64}$$

where α_F is the angle at the vertex A of the face $F \in \star(A)$.

The quantity $2\pi - \sum_{F\in\star(A)} \alpha_F$ is called the **angular defect** of M in A and is naturally connected with the classical definition of the Gauss curvature.

We are lead to the following definition of the **discrete Gauss curvature**:

Definition 13. *Let M be a molecule and A an atom in M. The discrete Gauss curvature of M in A is the quantity denoted by $G(A)$ and defined by:*

$$G(A) = 3\frac{K(A)}{\mid \star(A) \mid}, \tag{65}$$

where $\mid \star(A) \mid$ denotes the area of the faces having A as a vertex.

The main property of the previous quantity is the following **discrete analogue** of the **Gauss-Bonet theorem**:

Theorem 2. *Let M be a molecule. The global curvature of M, denoted by $K(M)$ is defined as:*

$$K(M) = \sum_{A\in M} K(A) \tag{66}$$

Let n_A, n_E, n_F be the number of atoms, edges and faces defined by M. We have:

$$K(M) = 2\pi(n_A - n_E + n_F) = 2\pi\xi(M), \tag{67}$$

where $\xi(S)$ is called the **Euler-Poincaré characteristic** *of M.*

The Euler-Poincaré characteristic is a topological invariant.

C. Connection between the angular defect and the pyramidalization angle

In this section, we consider a trivalent regular molecule M and we denote by l the common length of the bonds.

Let A be a given atom in M. Let us consider the triangle defined by the points A and $B_1, B_2 \in \star(A)$. The angle at the vertex A of F is denoted by $\alpha_1(A)$. We denote in the same way the angles $\alpha_2(A)$ and $\alpha_3(A)$, the angles at the vertex A of the triangles B_2AB_3 and B_3AB_1 respectively.

By definition, we have:

$$\alpha_1(A) + \alpha_2(A) + \alpha_3(A) = 2\pi - K(A) \tag{68}$$

We denote by $\mathscr{P}(A)$ the plane defined by $\star(A)$ and O the intersection of the mediating line for the triangle defined by $\star(A)$. Let us denote by $\theta_1(A)$ the angle at the vertex O of the triangle B_1OB_2. We denote in the same way by $\theta_2(A)$ and $\theta_3(A)$ the angle at the vertex O of the triangle B_2OB_3 and B_3OB_1 respectively.

We have the following relations:

Lemma 11. *Let M be a regular trivalent molecule. In each atom A of M, we have the following relation between $Pyr(A)$, $(\alpha_i(A))_{i=1,2,3}$ and $(\theta_i(A))_{i=1,2,3}$:*

$$\cos(Pyr(A)) \sum_{i=1}^{3} \sin(\frac{\theta_i(A)}{2}) = \sum_{i=1}^{3} \sin(\frac{\alpha_i(A)}{2}) \tag{69}$$

$$\cos\alpha_F(A) = \cos^2 Pyr(A) \cos\theta_F(A) + \sin^2 Pyr(A) \tag{70}$$

Proof. By definition, we have:

$$\theta_1(A) + \theta_2(A) + \theta_3(A) = 2\pi \tag{71}$$

By construction, we have $OB_1 = OB_2 = OB_3 = R$ the radius of the **circumcircle** for the triangle $B_1B_2B_3$.

The triangle AOB_1 is rectangle in O by construction. By definition of the pyramidalization angle, we have:

$$OB_1 = l\cos(Pyr(A)), \tag{72}$$

So that,

$$R = l\cos(Pyr(A)) \tag{73}$$

Using this quantity, one can compute the quantity $B_1B_2 + B_2B_3 + B_3B_1$ which is the perimeter of the triangle $B_1B_2B_3$ in $\mathscr{P}(A)$. We have two ways to compute this quantity:

- First, using the fact that in the triangle B_1OB_2, the length B_1B_2 is given by $2R\sin\left(\frac{\theta_1(A)}{2}\right)$ and similar expressions for B_2B_3, B_3B_1. We then obtain:

$$B_1B_2 + B_2B_3 + B_3B_1 = 2R\left[\sin\left(\frac{\theta_1(A)}{2}\right) + \sin\left(\frac{\theta_2(A)}{2}\right) + \sin\left(\frac{\theta_3(A)}{2}\right)\right] \tag{74}$$

- Second, using the fact that in the triangle B_1AB_2, the length B_1B_2 is given by $2l\sin\left(\frac{\alpha_1(A)}{2}\right)$ and similar expressions for B_2B_3, B_3B_1. We then obtain:

$$B_1B_2 + B_2B_3 + B_3B_1 = 2l\left[\sin\left(\frac{\alpha_1(A)}{2}\right) + \sin\left(\frac{\alpha_2(A)}{2}\right) + \sin\left(\frac{\alpha_3(A)}{2}\right)\right] \tag{75}$$

Replacing R by its expression, and writing the equality of these two expressions, we deduce that:

$$\cos(Pyr(A))\left[\sin\left(\frac{\theta_1(A)}{2}\right) + \sin\left(\frac{\theta_2(A)}{2}\right) + \sin\left(\frac{\theta_3(A)}{2}\right)\right] = \left[\sin\left(\frac{\alpha_1(A)}{2}\right) + \sin\left(\frac{\alpha_2(A)}{2}\right) + \sin\left(\frac{\alpha_3(A)}{2}\right)\right] \tag{76}$$

The pyramidalization angle is then to be understood as a measure for the difference between the angles $\theta_i(A)$ and $\alpha_i(A)$, $i = 1,2,3$.

Denoting by I_1 the middle of the segment B_1B_2 and using the triangle AOI_1 which is rectangle in O, we obtain using the Pythagorean Theorem:

$$\cos\alpha_F(A) = \cos^2 Pyr(A)\cos\theta_F(A) + \sin^2 Pyr(A) \tag{77}$$

This conclude the proof. □

As already mentioned, we have $\alpha_F(A) = \theta_F(A)$ if and only if $Pyr(A) = 0$.
It must be noted however, that the relation between $K(A)$ and $Pyr(A)$ is not "simple" and that no closed formula was founded. Using a Taylor expansion for the $\sin(\frac{\alpha_i(A)}{2})$ and $\sin(\frac{\theta_i(A)}{2})$ up to order 3, we can nevertheless give an approximation formula:

$$\cos(Pyr(A))\left(\pi - \frac{1}{48}\sum_{i=1}^{3}\theta_i(A)^3\right) \sim \frac{1}{2}K(A) - \frac{1}{48}\sum_{i=1}^{3}\alpha_i(A)^3 \tag{78}$$

The angular defect values calculated for C_{32} and C_{80} allowed to perform cartographies figure 14. As well as the $Pyr(A)$, the cartographies are presented according to the energetic order of the isomers, starting from the most stable and the values of the angular defect of C_{80} are given in relation to the mean value of C_{60}.

V. DISTRIBUTION OF THE PYRAMIDALIZATION ANGLE, SPHERICAL CURVATURE AND ANGULAR DEFECT OVER FULLERENE MOLECULES

In this section, we compute the pyramidalization angle, the spherical curvature and the angular defect over the database

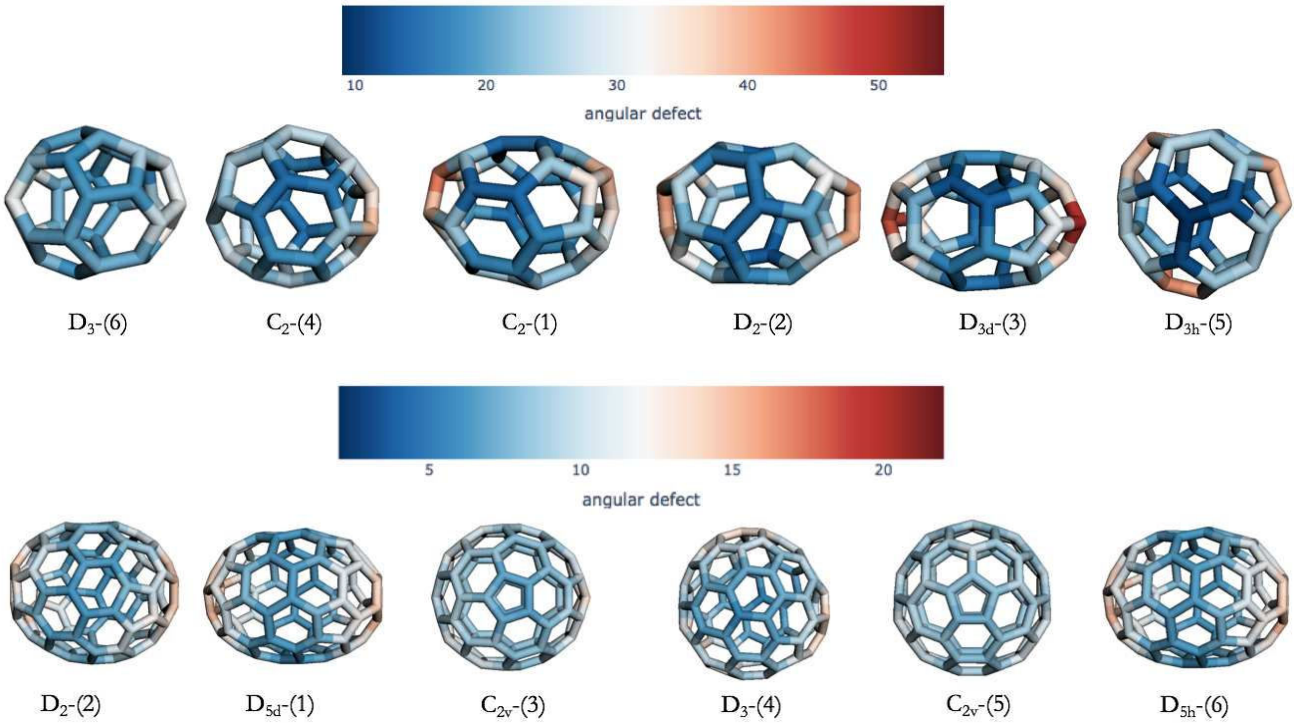


FIG. 14: Cartographies of C_{32} and C_{80} for the angular defect[4]

provided by Tománek and al.[5] containing a more or less exhaustive list of existing fullerene molecules. In particular, we illustrate the relationships between all these quantities studied in section III B and IV C. Since the range of objects involves only carbon molecules, we can use the results exposed in section II B to study the notion of carbon hybridization.

A. About the Tománek and al. database for fullerene molecules

The database used in this article comes from the work of Tománek and al.[5] which provides a more or less exhaustive list of all existing fullerene molecules. This basis presents a huge number of isomers in each C_n fullerene family, with n being the number of atoms. The total number of isomers in the complete database is 2487 (with n between $n = 20$ and 720 atoms).

Figure 15 presents the number of isomers in each C_n fullerenes family. We can clearly distinguish three families that we call **small**, **medium** and **large**:

1. **small** corresponds to fullerenes with n between 20 and 60.
2. **medium** corresponds to fullerenes with n between 60 and 100.

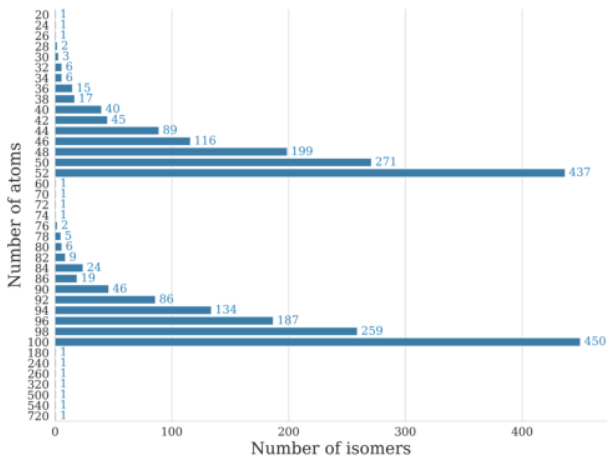


FIG. 15: Number of isomers for a given number of atoms.

3. Finally, **large** corresponds to fullerenes with $n \geq 100$.

Four families of C_n was last chosen to conclude this study i.e. $n = 32$, 40, 60 and 80 (these 4 families being located in each of the three domains previously identified). Only the study of the C_{32} family will be presented in detail in the article. The data inherent to the other three families are reported

in SI.
To appreciate the difference between molecular dynamics structural parameters (as reported in the Tománek basis) and an 'all-electron' calculation, we decided to calculate in addition the total electronic energy of the four selected fullerenes families (C_{32}, C_{40}, C_{60} and C_{80}) using ab initio density functional theory (DFT) as implemented in the Gaussian code[11]. Our choice was the use of a DFT Hamiltonian type B3LYP[12,13]. A double zeta atomic orbital base 6-31-G* completed the calculation conditions. The set of coordinates optimized in DFT is reported in the SI.

B. Statistical study of the Tománek database

It can reasonably be assumed that this working base is sufficiently complete to ensure that the objects are of sizes, shapes and local curvatures of all the states of hybridization that can be encountered for atoms involved in carbon structures.

An overview of the results obtained on the complete database is given in Table II. We give further details in the following. However, one can make already a few comments:

1. The range of pyramidalization angles ($Pyr(max) - Pyr(min) = 24°$) is sufficiently wide to include practically all the cases existing between the hybrid situations sp^2 ($Pyr(A) = 0°$) and sp^3 ($Pyr(A) = 19,5°$) acoording to the result of section II B 8.
   As a consequence, we cover all commonly encountered hybrid situations for the carbon atoms involved in structures like carbon nanotubes (CNT), fullerenes (n $\geq$ 60), nanodiamonds (NDs) and graphenes (plane and folded)[14–17].

2. Situations beyond ($Pyr(A) = 19,5°$) are also present. These sp(>3) hybridization situations are also important since they are potentially found into various recently discovered objects[18–29] and thus, deserves to be studied in this work.

We now give a detailed study of all the quantities.

1. Distribution of the pyramidalization angle

The high $Pyr(A)$ values are all observed for smaller structures up to $n < 60$. The very large dispersion of $Pyr(A)$ associated to these small systems (figure 16 and table II) is at the source of this situation. On the contrary, beyond $n = 60$, only few values of $Pyr(A)$ exceed the median value approximately given by the C_{60} fullerene (11.64°). This median value clearly marks a frankly separation existing between the smallest and the largest of carbon clusters. On the other hand, if the well-known fullerene is the C_{60} molecule, it can clearly be seen in figure 16 and table II that the average

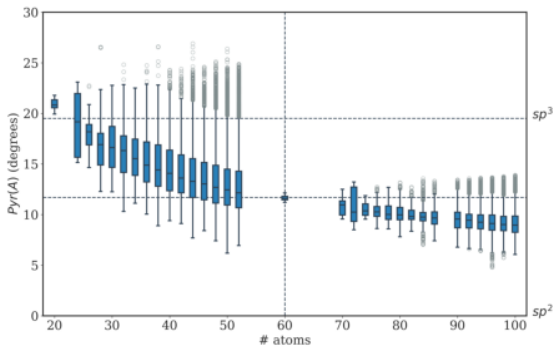


FIG. 16: Boxplot of the pyramidalization angle $Pyr(A)$ in degrees, as a function of the number of atom in the isomer. Dashed lines indicate values of the C_{60} isomers and the boundary between a sp^2 and sp^3 situations.

$Pyr(A)$ and hybridization values of this cluster are not out of the ordinary and follow the 'normal' evolution of $Pyr(A)$ and hybridizations properties.

2. Pyramidalization angle versus symmetries

A natural question is to look for the dependence of the pyramidalization angle with respect to the symmetry group of the underlying molecule.

Figure 17 presents the amount of isomers presents in the data base sorted by symmetry group. Green bars correspond to the most abundant symmetry groups, C_1, C_2 and C_s, the amount of isomers of these groups is given below each bars. On each line, the percentage represents the amount of isomers with this group among the whole number of isomers. Using this result, one can plot the valued of the pyramidalisation angle observed in each category of symmetry.

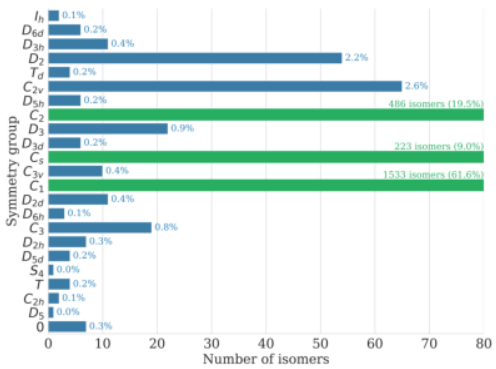


FIG. 17: Number of isomers by group symmetry.

Table III presents an ensemble of isomers sorted along the

TABLE II: Statistical descriptors of the angular defect (degrees), $Pyr(A)$, the pyramidalization angle (degrees), the spherical curvature $\kappa(A)$, the hybridization coefficients c_π^2 and λ_π^2, the hybridization numbers m and n, and the hybridization.

	$Pyr(A)$	Angular Defect	Spherical Curvature	c_π^2	λ_π^2	m	n	Hybridization
mean	10.49	10.42	0.262	0.074	0.926	0.084	2.252	$(s^{0.926}p_{x,y}^{2.000}p_z^{0.074})_\sigma(s^{0.074}p_z^{0.926})_\pi$
std	2.83	6.04	0.072	0.048	0.048	0.067	0.201	$(s^{0.048}p_{x,y}^{0.000}p_z^{0.048})_\sigma(s^{0.048}p_z^{0.0048})_\pi$
min	2.60	0.61	0.065	0.004	0.485	0.004	2.012	$(s^{0.485}p_{x,y}^{2.000}p_z^{0.004})_\sigma(s^{0.004}p_z^{0.485})_\pi$
25%	8.72	6.84	0.217	0.047	0.918	0.049	2.148	$(s^{0.918}p_{x,y}^{2.001}p_z^{0.047})_\sigma(s^{0.047}p_z^{0.918})_\pi$
50%	9.80	8.59	0.246	0.060	0.940	0.063	2.190	$(s^{0.940}p_{x,y}^{2.001}p_z^{0.060})_\sigma(s^{0.060}p_z^{0.940})_\pi$
75%	11.45	11.64	0.288	0.082	0.953	0.089	2.268	$(s^{0.953}p_{x,y}^{2.001}p_z^{0.082})_\sigma(s^{0.082}p_z^{0.953})_\pi$
max	26.90	59.67	0.637	0.515	0.996	1.061	5.184	$(s^{0.996}p_{x,y}^{2.001}p_z^{0.515})_\sigma(s^{0.515}p_z^{0.996})_\pi$

standard deviation of the pyramidalization angle. One can see strongly anisotropic isomers and more spherical ones. The C_{60} isomer is presented as a reference system. On the contrary, table IV, lists the isomers that present the largest deviation standard of the pyramidalization angle ($\sigma_{Pyr(A)} > 4.5^o$). In the two cases, the symmetry does not seem to be related to the computed value of the pyramidalization angle and not a function of stability either.

3. *Relation between the pyramidalization angle, angular defect and spherical curvature*

We have provided in sections II B, III B, and IV C analytical formulas relating the pyramidalization angle, the angular defect, the spherical curvature and the way the pyramidalization angle enters in the hybridization. We illustrate all these results directly using the Tománek database.

Computations made over the previous database are summarized in the following figure 18. As shown in the figure 18, and explained in section V A there exists different sizes of isomers. The arbitrary choice of these three categories of objects is based on the following remarks:

1. First, there is a clear relationship between the four quantities. The pairwise relationship between $Pyr(A)$ and the Spherical Curvature is linear. That was expected for small $Pyr(A)$ values as proved in section III B.

2. Small isomers are those which exhibits larger curvature. When the size of the object increases, the pyramidalization angles tend to zero. Then, if these domains concern clusters of very different sizes, shapes and, particularly, local curvatures, it is important to note that there is no discontinuity in the relationships between all these properties. This observation is in line with equations 57, 58 and 78 established here but also with those reported in R. C. Haddon's work.

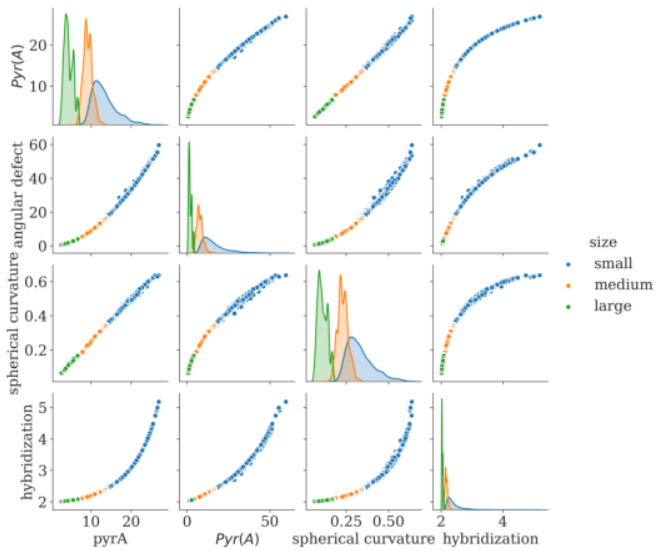


FIG. 18: Pairwise relationship between $Pyr(A)$, the angular defect, the spherical curvature and the hybridization on the whole isomers of the data base.

Notably, figure 18 extends the initial work[1] of R. C. Haddon. In particular, we complete the calculations related to hybridization and to the relation between the spherical curvature and the pyramidalization angle that R. C. Haddon gives only for the C_{60} and C_{70} systems (see figure 6[30]).

This figure implies that:

1. The relation linking the curvature and the pyramidalization angle given by the lemma 9 and which is quasi

TABLE III: View of several isomer structures for a range of values of the standard deviation of the pyramidalization angle.

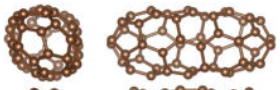
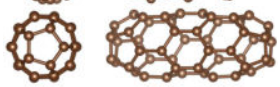
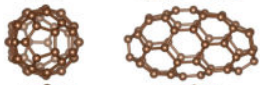
Compound	*File*	# Atom	Symmetry	$\langle Pyr(A)\rangle$	$\sigma_{Pyr(A)}$	
C_{52}	C52-D2-2	52	D_2	13.61	5.10	
C_{50}	C50-D5h-1	50	D_{5h}	14.12	4.93	
C_{52}	C52-Cs-3	52	C_s	13.21	4.29	
C_{44}	C44-C2v-13	44	C_{2v}	14.08	3.59	
C_{52}	C52-C1-98	52	C_1	12.81	2.93	
C_{96}	C96-C1-131	96	C_1	9.22	0.90	
C_{60}	C60-Ih	60	I_h	11.64	0.22	

TABLE IV: Isomers that present the largest deviation standard of the pyramidalization angle, $\sigma_{Pyr(A)} > 4.5^o$

Compound	*File*	# Atom	Symmetry	$\langle Pyr(A)\rangle$	$\sigma_{Pyr(A)}$
C_{48}	C48-D2-2	48	D_2	13.87	5.73
C_{44}	C44-D3-35	44	D_3	13.95	5.46
C_{36}	C36-D2-2	36	D_2	15.48	5.23
C_{50}	C50-D3h-3	50	D_{3h}	13.86	5.18
C_{44}	C44-D2-2	44	D_2	14.41	5.15
C_{50}	C50-C2v-13	50	C_{2v}	13.33	5.14
C_{52}	C52-D2-58	52	D_2	12.88	5.14
C_{52}	C52-D2-2	52	D_2	13.61	5.10
C_{50}	C50-D5h-1	50	D_{5h}	14.12	4.93
C_{44}	C44-D3d-3	44	D_{3d}	14.47	4.79
C_{40}	C40-D5d-1	40	D_{5d}	15.23	4.77
C_{48}	C48-C2-5	48	C_2	13.68	4.77
C_{52}	C52-C1-17	52	C_1	13.11	4.75
C_{48}	C48-C1-49	48	C_1	13.40	4.73
C_{52}	C52-C1-45	52	C_1	13.01	4.73
C_{40}	C40-D2-3	40	D_2	14.95	4.68
C_{52}	C52-C2-1	52	C_2	13.42	4.61
C_{52}	C52-D2d-313	52	D_{2d}	12.94	4.60
C_{44}	C44-C1-10	44	C_1	14.28	4.55
C_{50}	C50-C2-18	50	C_2	13.25	4.53
C_{52}	C52-C1-14	52	C_1	13.07	4.52
C_{50}	C50-C2-2	50	C_2	13.61	4.52
C_{52}	C52-C2-15	52	C_2	13.27	4.51
C_{48}	C48-C1-12	48	C_1	13.53	4.51

linear for angles of low values persists for large value.

2. The relation between the angular defect and the pyramidalization angle discussed in section IV C is not simple from the analytic view point. The previous figure justifies firstly the existence of the relation between the angle of pyramidalization and the angular defect. This last point is crucial as we propose the angular defect as an alternative to the pyramidalization angle in situations where it is not defined (see section II D).

3. An extension to the properties c_π, n and of the hybridization (see section II B 8 b) is proposed in figures S3-S6 in the supplementary information.

C. The C_{32} isomers

Table V lists the statistical descriptors of the geometrical quantities for the six isomers of the C_{32} family after optimization. The main results obtained for the six conformers of the C_{32} family are reported in the table $S5$ and in figures 19, 20, and $S7$. We can, first of all, point out that the explicit consideration of the electronic correlation through quantum calculations does not change the conclusions previously established. In particular, even if the C_{32} clusters are those with the highest pyramidalization angles distribution, the quasi linear relation between $Pyr(A)$ and SC (or AD) remains perfectly valid (figure $S7$).
We also observe that these dependency relationships do not vary according to the symmetry and/or energy stability of the cluster.

1. *Dependence to geometrical optimization*

The coordinates provided by the Tománek database correspond to non optimized moecules. A natural question is then

TABLE V: Statistical descriptors of isomers of the C_{32} family.

	$Pyr(A)$	Angular Defect	Spherical Curvature	c_π^2	λ_π^2	m	n	Hybridization
mean	16.46	23.67	0.391	0.180	0.820	0.229	2.686	$(s^{0.820}p_{x,y}^{1.999}p_z^{0.180})_\sigma(s^{0.180}p_z^{0.820})_\pi$
std	2.58	7.07	0.060	0.064	0.064	0.116	0.347	$(s^{0.064}p_{x,y}^{0.000}p_z^{0.064})_\sigma(s^{0.064}p_z^{0.064})_\pi$
min	10.58	9.88	0.249	0.070	0.512	0.075	2.225	$(s^{0.512}p_{x,y}^{2.000}p_z^{0.070})_\sigma(s^{0.070}p_z^{0.512})_\pi$
25%	14.24	17.68	0.339	0.129	0.800	0.148	2.443	$(s^{0.800}p_{x,y}^{1.999}p_z^{0.129})_\sigma(s^{0.129}p_z^{0.800})_\pi$
50%	16.52	23.51	0.394	0.176	0.824	0.213	2.640	$(s^{0.824}p_{x,y}^{2.001}p_z^{0.180})_\sigma(s^{0.180}p_z^{0.824})_\pi$
75%	17.54	26.29	0.420	0.200	0.871	0.250	2.749	$(s^{0.871}p_{x,y}^{1.999}p_z^{0.200})_\sigma(s^{0.200}p_z^{0.871})_\pi$
max	26.28	54.33	0.613	0.488	0.930	0.951	4.854	$(s^{0.930}p_{x,y}^{2.001}p_z^{0.488})_\sigma(s^{0.488}p_z^{0.930})_\pi$

to determine to which extend the pyramidalization angle, the hybridization numbers, etc, depends on the geometrical optimization process.

Figure 19 and 20 present both the distributions of the pyramidalization angle and the hybridization of the atoms in C_{32} isomers before (initial) and after (optimized) a geometrical optimization using quantum chemistry, respectively. From

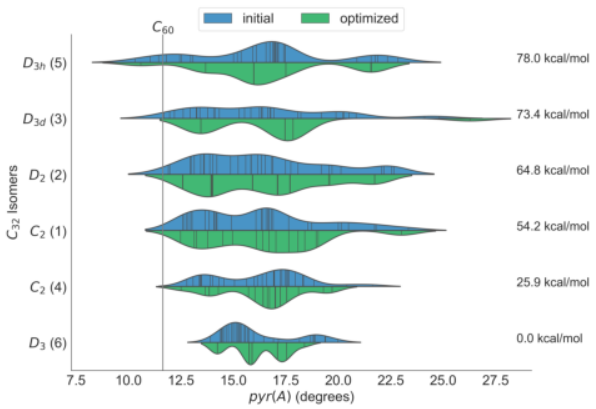


FIG. 19: Pyramidalization angle distributions of the C_{32} isomers before and after the geometrical optimization. The vertical line shows the value of the pyramidalization angle of the C_{60} isomer. The energies are relative to the most stable isomer.

these figures, one can see that the geometrical optimization restored the symmetry of the systems. As expected, all the C_{32} isomers present higher $Pyr(A)$ than the C_{60} isomer.

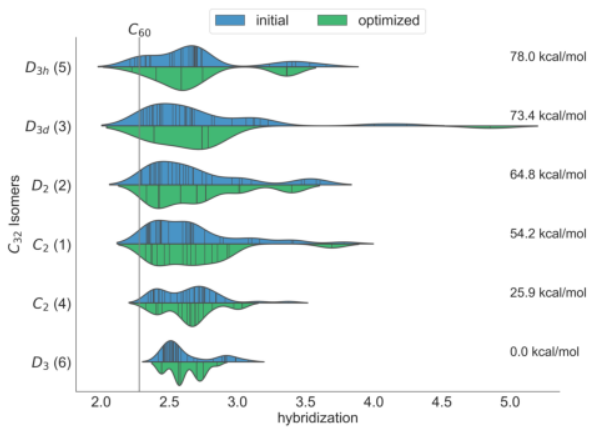


FIG. 20: Hybridization of the C_{32} isomers before and after the geometrical optimization. The vertical line shows the value of the pyramidalization angle of the C_{60} isomer. The energies are relative to the most stable isomer.

2. *Stability versus pyramidalization angle and symmetries*

Looking at the relative energies between isomers, the most stable isomers is not the one which presents the lowest pyramidalization angles nor the one with the higher symmetry. On the contrary, the relative stability of the isomers seems to be linked to the pyramidalization angles distributions and thus to the heterogeneity of the structures. However, this observation should be tempered since, contrary to C_{40} and C_{32} isomers, the energies of C_{80} isomers increase when the width of the $Pyr(A)$ distributions increase. Once again, we see that there are two types of behaviour depending on whether we study a cluster beyond or within the C_{60} boundary.

D. Conclusion

In conclusion, all the quantum results obtained on the four families $n = 32$, 40, 60 and 80 allow us to validate the statistical study carried out on the basis of structural data from molecular dynamics.

VI. CONCLUSION AND PERSPECTIVES

The pyramidalization angle and the spherical curvature were introduced in the seminal work of R. C. Haddon[2,3]. However, it suffer from severe limitations. In particular, except in some very exceptional cases, these quantities are not defined for molecules which are not trivalent. The angular defect and the associated discrete Gauss curvature provide an effective and explicit quantity reflecting the local geometry of the molecule.

In this work, we have provided a self-contained presentation of the pyramidalization angle, the spherical curvature and the angular defect as well as the relationship between all these quantities. We have completed and extended previous results of R. C. Haddon in particular concerning a description of the hybridization as a function of the pyramidalization angle.

Moreover, we have developed an online program which can be used to compute all these quantities. This program was used to give an overview and a statistical study of these quantities over the Tománek database covering most of the existing fullerene molecules.

It remains to prove that the previous quantity provide an efficient way to quantify the regio-reactivity of a molecule, extending the previous work of R. C. Haddon and other[31–33] for trivalent molecules and a previous study of the authors in this direction[34].

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available for free.
Tabulated value of $Pyr(A)$, global analysis, atomic analysis (C_{32}, C_{40},C_{60}, C_{80}), elasticity and curvature energy are available.
Data viewer of cartography is available online https://pychemapps.univ-pau.fr/mosaica/.
Coordinates xyz of C_{32}, C_{40},C_{60}, C_{80} used are all listed.

APPENDIX

Proof associated with the Lemma 5:

Proof. The normalization on the s component of (h_1, h_2, h_3, h_π) gives:

$$c_\pi^2 + 3\mu^2\lambda_\pi^2 = 1, \tag{79}$$

and the one on the p_z component gives:

$$\lambda_\pi^2 + 3\mu^2 c_\pi^2 = 1. \tag{80}$$

We then obtain by addition of the two previous equations:

$$(c_\pi^2 + \lambda_\pi^2)(1 + 3\mu^2) = 2 \tag{81}$$

As $c_\pi^2 + \lambda_\pi^2 = 1$, this gives:

$$\mu = \frac{1}{\sqrt{3}} \tag{82}$$

This concludes the proof. □

ACKNOWLEDGMENTS

The project leading to this publication has received funding from Excellence Initiative of Université de Pau et des Pays de l'Adour - I-Site E2S UPPA, a french "Investissements d'Avenir" programme. This article is also supported by the Défi CNRS Biomimétisme in collaboration with the Université Claude Bernard Lyon 1 and Université de Lyon.

REFERENCES

[1] R. C. Haddon, "C_{60} sphere or polyhedron?" J. Am. Chem. Soc. **119**, 1797–1798 (1997).
[2] R. C. Haddon and L. T. Scott, "π-orbital conjugation and rehybridization in bridged annulenes and deformed molecules in general: π-orbital axis vector analysis," Pure & Appl. Chem. **58**, 137–142 (1986).
[3] R. C. Haddon, "Hybridization and the orientation and alignment of π-orbitals in nonplanar conjugated organic molecules: π-orbital axis vector analysis (poav2)," J. Am. Chem. Soc. **108**, 2837–2842 (1986).
[4] G. Salvato-Vallverdu, "Strucutral data viewer," `https://pychemapps.univ-pau.fr/mosaica/`, accessed: 2020-01-09.
[5] D. Tománek and N. Frederick, "C_n Fullerenes," `http://www.nanotube.msu.edu/fullerene/`, accessed: 2020-01-08.
[6] R. Haddon, "Comment on the relationship of the pyramidalization angle at a conjugated carbon atom to the σ bond angles," J. Am. Chem. Soc. **105**, 4164–4165 (2001).
[7] Y. Liu, H. Pottmann, J. Wallner, Y. L. Yang, and W. Wang, "Geometric modeling with conical meshes and developable surfaces," ACM (2006).
[8] P. Romon, *Introduction à la géométrie différentielle discrète* (Ellipses, 2013).
[9] J. G. Radziszewski, J. W. Downing, M. Jawdosiuk, P. Kovacic, and J. Michl, "4-azahomoadamant-3-ene: spectroscopic characterization and photoresolution of a highly reactive strained bridgehead imine," J. Am. Chem. Soc. **107**, 594–603 (1985).
[10] R. C. Haddon, "Poav 3, qcmp044," QCPE Bull. **15**, 81 (1995).
[11] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. V. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski,

J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, T. A. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman, and D. J. Fox, "Gaussian 16 Revision C.01," (2016), gaussian Inc. Wallingford CT.

[12]A. Becke, "Density-functional thermochemistry. the role of exact exchange," J. Chem. Phys. **98**, 5648–5652 (1993).

[13]C. Lee, W. Yang, and R. Parr, "Development of the colle-salvetti correlation-energy formula into a functional of the electron density," Phys. Rev. B **37**, 785–789 (1998).

[14]C. Chuang, J. Guan, D. Witalka, Z. Zhu, B. Y. Jin, and D. Tománek, "Relative stability and local curvature analysis in carbon nanotori," Phys. Rev. B **91**, 165433 (2015).

[15]A. Kuc and G. Seifert, "Hexagon-preserving carbon foams: Properties of hypothetical carbon allotropes," Phys. Rev. B **74** (2006).

[16]P. W. Fowler and D. E. Manolopoulos, "A review of: An atlas of fullerenes," Fullerene Science and Technology **4**, 177–178 (1996).

[17]A. Fasolino, J. H. Los, and M. I. Katsnelson, "Intrinsic ripples in graphene," Nat. Mat. **6**, 858–861 (2007).

[18]R. Meija, S. Signetti, A. Schuchardt, K. Meurisch, D. Smazna, M. Mecklenburg, K. Schulte, D. Erts, O. Lupan, B. Fiedler, Y. K. Mishra, R. Adelung, and N. M. Pugno, "Nanomechanics of individual aerographite tetrapods," Nat. Commun. **8** (2017).

[19]J. W. Martin, C. de Tomas, I. Suarez-Martinez, M. Kraft, and N. A. Marks, "Topology of disordered 3d graphene networks," Phys. Rev. Lett. **123**, 116105 (2019).

[20]V. V. Roddatis, V. L. Kuznetsov, V. Butenko, S. Suaand, and R. Schlögl, "Transformation of diamond nanoparticles into carbon onions under electron irradiation," Phys. Chem. Chem. Phys. **4**, 1964–1967 (2002).

[21]N. Park, M. Yoon, S. Berber, J. Ihm, E. Osawa, and D. Tománek, "Magnetism in all-carbon nanostructures with negative gaussian curvature," Phys. Rev. Lett. **91**, 237204 (2003).

[22]B. Liu, H. Shioyama, T. Akita, and Q. Xu, "Metal-organic framework as a template for porous carbon synthesis," J. Am. Chem. Soc. **130**, 5390–5391 (2008).

[23]P. Yilong, X. Chenlong, X. Mei, M. Mengdong, L. Lingyu, L. Zihe, Z. Shuangshuang, G. Guoying, Z. Zhisheng, T. Yongjun, X. Bo, and H. Julong, "A superhard sp^3 microporous carbon with direct bandgap," Chem. Phys. Lett. **689**, 68–73 (2017).

[24]Z. Qin, G. S. Jung, M. J. Kang, and M. J. Buehler, "The mechanics and design of a lightweight three-dimensional graphene assembly," Sci. Adv. **3** (2017).

[25]Z. S. Wu, Y. Sun, Y. Z. Tan, S. Yang, X. Feng, and K. Müllen, "Three-dimensional graphene-based macro- and mesoporous frameworks for high-performance electrochemical capacitive energy storage," J. Am. Chem. Soc. **134**, 19532–19535 (2012).

[26]X. Wu and X. C. Zeng, "First-principles study of a carbon nanobud," J. Am. Chem. Soc. Nano **2**, 1459–1465 (2008).

[27]G. Naulet, L. Sturm, A. Robert, P. D. F. Röhricht, R. Herges, H. Bock, and F. Durola, "Cyclic tris-[5]helicenes with single and triple twisted möbius topologies and möbius aromaticity," Chem. Sci. **9**, 8930–8936 (2018).

[28]S. Park, K. Kittimanapun, J. A. Sun, Y. K. Kwon, and D. Tománek, "Designing rigid carbon foams," J. Phys. Condens. Mat. **22**, 334220 (2010).

[29]E. Barborini, P. Piseri, P. Milani, G. Benedek, C. Ducati, and J. Robertson, "Negatively curved spongy carbon," Appl. Phys. Lett. **81**, 3359–3361 (2002).

[30]R. C. Haddon, "π-electrons in three dimensions," Acc. Chem. Res. **21**, 243–249 (1988).

[31]J. M. Hawkins, A. Meyer, and M. A. Solow, "Osmylation of c_{70}: reactivity versus local curvature of the fullerene spheroid," J. Am. Chem. Soc. **115**, 7499–7500 (1993).

[32]J. M. Hawkins and A. Meyer, "Optically active carbon: Kinetic resolution of c_{76} by asymmetric osmylation," Science **260**, 1918–1920 (1993).

[33]R. C. Haddon, "Chemistry of the fullerenes: the manifestation of strain in a class of continuous aromatic molecules," Science **261**, 1545–1550 (1993).

[34]H. S. Silva, J. Cresson, A. Rivaton, D. Bégué, and R. C. Hiorns, "Correlating geometry of multidimensional carbon allotropes molecules and stability," Organic Electronics **26**, 395–399 (2015).

[35]C. A. Coulson, B. O'Leary, and R. B. Mallion, *Hückel Theory for Organic Chemists* (Academic Press, London, 1978).

[36]D. Bakowies and W. Thiel, "Mndo study of large carbon clusters," J. Am. Chem. Soc. **113**, 3704–3714 (1991).