GLOBAL AND LOCAL RELATIVE CONVEXITY AND ORIENTED RELATIVE CONVEXITY; APPLICATION TO MOLECULAR SHAPES IN EXTERNAL FIELDS

Paul G. MEZEY

Department of Chemistry and Department of Mathematics, University of Saskatchewan, Saskatoon, Saskatchewan, Canada S7N 0W0

Abstract

The concepts of global and local relative convexity and oriented relative convexity are described and proposed as tools for the characterization of molecular shapes. The usual concept of convexity is a special case of the generalization described. Oriented relative convexity is suitable for the characterization of molecular shapes in external fields, such as magnetic fields or fields representing cavity regions of various enzymes or zeolite catalysts. Potential applications include new approaches to computer-based drug design and molecular engineering.

1.   Introduction

Convexity is an important tool for the characterization of the shapes of sets in a metric space [1,2], and provides mathematical tools for the description of shapes of physical objects. However, for certain applications the scope of the conventional concept of convexity is somewhat limited; it cannot be directly applied to distinguish those sets and objects that are "almost" convex in some practical, approximate sense from those that have dominant nonconvex features. In the strict sense, both types of sets are nonconvex, although a minor deformation may turn some of them into a convex set. For certain physical applications, such as the description of molecular shape (which is important in drug design and molecular engineering, see refs. [3–28]), it is advantageous to introduce methods that can differentiate shapes according to some "degree of nonconvexity" (or a "degree of convexity") that can be used for a precise, quantitative shape comparison of different molecules.

In this paper, we shall describe two approaches for a generalization of the concept of convexity that are applicable to the molecular problem.

   The first approach is directly linked to the usual, global definition of convexity; a subset C of a metric space X is convex if and only if for any two points of the subset C,

$$x_1, x_2 \in C, \tag{1}$$

all points x generated as

$$x = \alpha x_1 + (1 - \alpha)x_2 \tag{2}$$

are also elements of set C, for every real number α fulfilling the condition

$$0 \leqslant \alpha \leqslant 1. \tag{3}$$

Pictorially, for any two points x_1 and x_2 of set C, the straight line segment inter-connecting them is also within C. In the generalization of global convexity considered in this paper, this straight line is allowed to be replaced by some curved line, fulfilling certain curvature conditions. This leads to the concept of *global relative convexity*, relative to these curvature conditions. By placing some orientational restrictions on these lines, one obtains criteria for *oriented global relative convexity*.

   The second approach, which is in fact more useful for the chemical applications which motivate our study, is somewhat less general, as it applies only to sets and objects with an everywhere twice continuously differentiable *boundary*. By applying certain local criteria for the boundary points, *local convexity* can be defined, as follows:

   At each point $r \in G$ of the continuous and differentiable boundary G of a set C, a local tangent hyperplane $^{n-1}R(r)$ can be defined, where n is the dimension of set C. (In molecular shape characterization, the $n = 3$ case is of primary importance.) In a sufficiently small local neighborhood of point r, the boundary hypersurface G may be regarded as a function $G_r(q)$ over this tangent hyperplane $^{n-1}R(r)$, where $q \in {}^{n-1}R(r)$. We shall assume that this function $G_r(q)$ is negative along the normal vector of the tangent hyperplane $^{n-1}R(r)$ pointing toward the interior of set C, it is zero at point r, and it takes positive values along the normal vector pointing in the direction of the exterior of set C. Within this local representation, the tangent point $q = r$ is a *critical point* of $G_r(q)$, and the Hessian matrix $H(r)$ of the second derivatives $\partial^2 G_r(q)/\partial q_i \partial q_j$ may be used for its characterization. The eigenvalues

$$h_1, h_2, \ldots, h_{n-1} \tag{4}$$

of $H(r)$ are referred to as the *local canonical curvatures* of the boundary G at point r. Points $r \in G$ can be classified according to their critical point index, the number $\mu(r)$ of negative eigenvalues of $H(r)$. Set A_μ is the collection of all boundary points $r \in G$ with precisely μ negative canonical curvatures:

$$A_\mu = \{r : r \in G, \mu(r) = \mu\}. \tag{5}$$

The sets A_μ generate a partitioning of the boundary G:

$$G = \bigcup_{\mu = 0}^{n-1} A_\mu, \tag{6}$$

$$A_\mu \cap A_{\mu'} = \emptyset, \quad \text{if} \quad \mu \neq \mu'. \tag{7}$$

A set A_μ may be disconnected. We shall use the $D_{\mu,i}$ notation for the ith maximum connected component of A_μ. Evidently,

$$G = \bigcup_{\mu,i} D_{\mu,i} \tag{8}$$

and

$$D_{\mu,i} \cap D_{\mu',i'} = \emptyset, \quad \text{if either} \quad \mu \neq \mu' \quad \text{or} \quad i \neq i'. \tag{9}$$

The decomposition (8) of boundary G is based on local curvature properties. In the most common $n = 3$ case, the three types of domains $D_{0,i}$, $D_{1,i}$ and $D_{2,i}$ correspond to locally concave, saddle type, and convex boundary pieces, respectively. Note that a $D_{\mu,i}$ set is not necessarily open, and the term "domain" is used in an unconventional sense.

Pictorially, the above classification of hypersurface domains according to local convexity can be obtained by an inspection of the relations between the boundary hypersurface G and its tangent hyperplane $^{n-1}R(r)$ at point $r \in G$. For example, in the $n = 3$ case, if the tangent hyperplane lies locally on the "inside" of G, then $\mu(r) = 0$; if it lies locally on the outside of G, then $\mu(r) = 2$; and if it cuts into G within any small open neighborhood of r, then $\mu(r) = 1$. The local neighborhoods are concave, convex, and of the saddle type, respectively.

In the generalization of local convexity considered in this paper, the tangent hyperplane is replaced by some more general tangent hypersurface of some specified curvature properties. This leads to the concept of *local relative convexity*, relative to the specified curvature properties of the tangent hypersurface. The simplest case corresponds to a tangent sphere of some finite radius; the infinite radius gives the original tangent hyperplane. By placing some orientational restriction on a non-spherical tangent hypersurface, one obtains the concept of *oriented local relative convexity*. This latter concept has applications to the description of molecular shape in external electromagnetic fields, or in fields generated by other molecules, where orientation has a crucial chemical role.

The presentation given in the following sections is general for the finite n-dimensional case. However, in view of the intended chemical applications we shall often use the terminology of two-dimensional surfaces that in the general n-dimensional case should correspond to that of $(n-1)$-dimensional hypersurfaces.

2.   Global relative convexity

We shall consider all possible paths $p(s)$ contained within a connected subset C of an n-dimensional Euclidean space ${}^nE$. If for any two points x_1 and x_2 of set C there exists a path $p(s)$ interconnecting them within set C and satisfying a specified set of curvature restrictions, then set C is *globally convex relative to these curvature restrictions*.

Each path $p(s)$ is parametrized by the arc length s. For each component $p_i(s)$ of point $\boldsymbol{p}(s)$ along the path $p(s)$, where

$$\boldsymbol{p}(s) = (p_1(s), p_2(s), \ldots, p_n(s)),\tag{10}$$

the curvature restrictions can be given in a general form:

$$b_{i,\varrho} \leqslant d^2 p_i(s)/ds^2 \leqslant b_{i,u}, \quad i = 1, 2, \ldots, n.\tag{11}$$

Here, $b_{i,\varrho}$ and $b_{i,u}$ are the lower and upper curvature bounds, respectively, for path component p_i. If for any two points $x_1, x_2 \in C$ there exists a path $p(s)$ that is contained in set C and interconnects points x_1 and x_2, and if conditions (11) are fulfilled for every parameter value s along the path for each component p_i, then set C is *globally convex with respect to the sets of lower and upper bounds on directional curvatures,*

$$\left\{b_{i,\varrho}\right\}_{i=1,\ldots,n}\tag{12}$$

and

$$\left\{b_{i,u}\right\}_{i=1,\ldots,n},\tag{13}$$

respectively. Note that in general these bounds may be different for different components p_i, which implies a *dependence on orientation.*

If, however,

$$-b_{i,\varrho} = b_{i,u} = b, \quad \text{for all } i = 1, 2, \ldots, n,\tag{14}$$
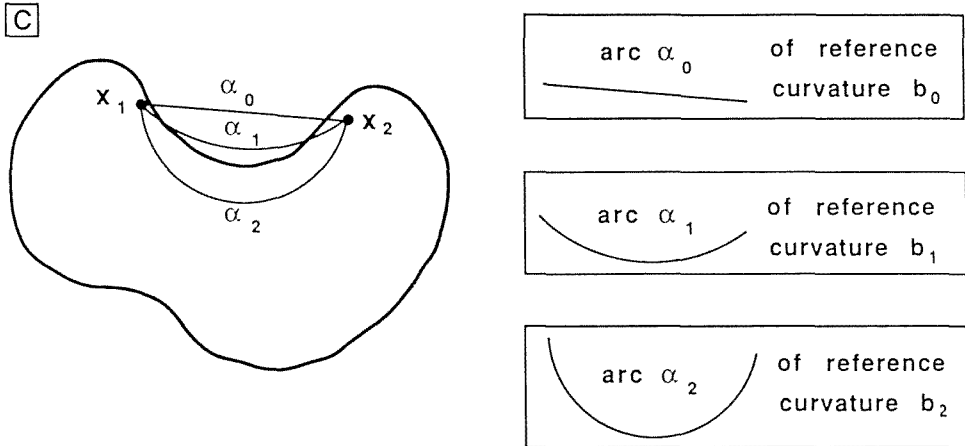
that is, if

GLOBAL RELATIVE CONVEXITY

Fig. 1. Illustration of the concept of global relative convexity. Set C is not globally convex relative to curvature $b_0 = 0$ (straight line, the conventional concept of convexity), nor relative to curvature b_1, but the same set C is globally convex relative to curvature b_2.

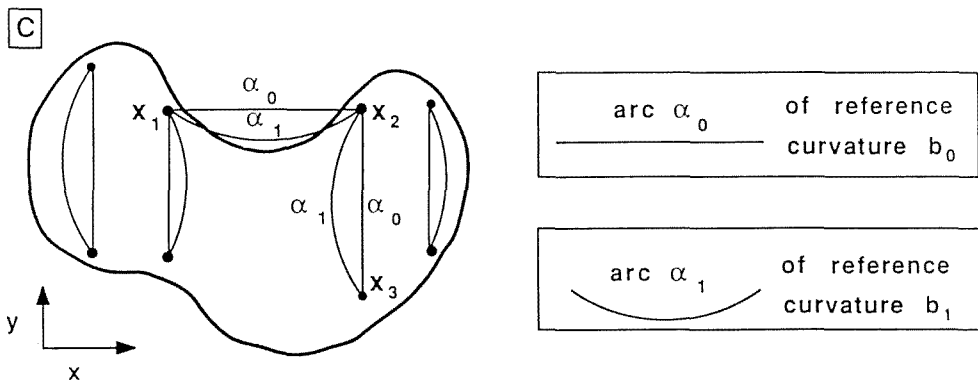# ORIENTED GLOBAL RELATIVE CONVEXITY

Fig. 2. Illustration of the concept of oriented global relative convexity. The set C of fig. 1 is globally convex relative to any curvature along one orientation, but it is not globally convex relative to curvatures b_0 and b_1, along another orientation, perpendicular to the first one.

$$|\mathrm{d}^2 p_i(s)/\mathrm{d}s^2| \leqslant b, \quad \text{for all } \ i = 1, 2, \ldots, n, \tag{15.}$$

then all directions are equivalent and then set C is *globally convex relative to curvature b,* with no reference to orientation. The concepts of global relative convexity and oriented global convexity are illustrated in figs. 1 and 2.

3.   Local relative convexity

Consider the set C, assume that it has an everywhere twice differentiable boundary hypersurface G, and choose a test hypersurface T. Local relative convexity will be defined with respect to this test hypersurface T, which may be some closed manifold or a hyperplane. Hypersurface T is also twice differentiable everywhere.

For any point $r \in G$, the tangent hyperplane $^{n-1}R(r)$ of G is well defined. By appropriate translation and rotation, the test hypersurface T can always be moved into a position where T becomes a tangent to hyperplane $^{n-1}R(r)$, precisely at the point $r \in G$ that is the point of tangential contact between G and hyperplane $^{n-1}R(r)$. In fact, each point of the test hypersurface T can be made to coincide with point $r \in G$ so that T and $^{n-1}R(r)$ are tangents to each other.

Such an arrangement of the test hypersurface T is called a *placement* $P(r)$ of T, and the family of all placements is denoted by $P(r)$. For each such placement P, both hypersurfaces G and T can be regarded locally as functions defined over the hyperplane $^{n-1}R(r)$,

$$G_r(q): {}^{n-1}R(r) \rightarrow {}^1R, \tag{16}$$

$$T_{r,P}(q): {}^{n-1}R(r) \rightarrow {}^1R, \tag{17}$$

where $q \in {}^{n-1}R(r)$, and in the notation the actual placement P of the test hypersurface T is indicated.

For both locally defined functions $G_r(q)$ and $T_{r,P}(q)$, point $q = r$ is a critical point that can be characterized by their respective Hessian matrices

$$\mathbb{H}_G(r) \tag{18}$$

and

$$\mathbb{H}_{T,P}(r), \tag{19}$$

which, in turn, are characterized by their nondecreasing sequences of eigenvalues:

$$h_1, h_2, \ldots, h_{n-1} \tag{20}$$

and

$$b_1, b_2, \ldots, b_{n-1},\tag{21}$$

respectively.

The local relative convexity is expressed by the relative curvature relations between the boundary hypersurface G of set C and the test hypersurface T, at the point r where they become tangents. The curvature properties of G at point r, relative to placement P of T, can be characterized by the relations among the eigenvalues of the two Hessian matrices. In particular, we define index $\mu_{i,P}(r)$ as follows:

$$\mu_{i,P}(r) = \text{the number of } h_j \text{ eigenvalues within the } [b_{i-1}, b_i) \text{ interval,}\tag{22}$$

where the $b_0 = -\infty$ convention is used. Note that in the above definition, the interval is closed–open. Of course, the relative distribution of eigenvalues (canonical curvatures of the two hypersurfaces), and hence these indices $\mu_{i,P}(r)$, do depend on the placement P of the test hypersurface T, that is, on the actual positioning and osculation of T to G.

The set of indices,

$$\left\{\mu_{i,P}(r)\right\}_{i=1,\ldots,n-1},\tag{23}$$

describes the local convexity of G at point r, relative to placement P of the test hypersurface T.

For each point $r \in G$, a partial order $\leqslant$ can be defined for the placements P of the test hypersurface T, as follows:

$$P \leqslant P'\tag{24}$$

if and only if

$$\{\mu_{i,P}(r)\} \leqslant \{\mu_{i,P'}(r)\},\tag{25}$$

where in (25) the symbol $\leqslant$ refers to the lexicographic ordering of $(n-1)$-tuples of integers. Placement P is called a *minimal placement* if (24) holds with respect to any other placement P'.

For a minimal placement, the subscript P will be omitted and we shall use the index set

$$\left\{\mu_i(r)\right\}_{i=1,\ldots,n-1},\tag{26}$$

for the characterization of *local convexity of hypersurface G at point r relative to test hypersurface T.*

Evidently, for a general reference hypersurface T, the realization of a minimal placement P requires that T should be allowed to be shifted and rotated freely in order to find a "best" (a minimal) fitting P to G at point r. However, for certain physical problems where external directions may play some important role, one may need a model where not all placements of T are allowed. In particular, an important class of physical problems suggests a model where the orientation of the test hypersurface T is restricted, and only translations are allowed in bringing about an oscu-lation of the two hypersurfaces G and T at each selected point $r \in G$. This restriction leads to the concept of *local oriented relative convexity* with reference to an *oriented* test hypersurface T. Note that in this context, the term *orientation* refers to an external coordinate system of the space within which G and T are embedded, and it evidently differs from the topological concept of orientation given to manifolds.

Subset $P_0(r)$ of all placements $P(r)$ contains all those placements of T which are related to some standard positioning P_0 of T by a rigid translation.

The index set

$$\left\{ \mu_{i,P_0}(r) \right\}_{i=1,\dots,n-1} . \tag{27}$$

is obtained for a placement P_0' that is *minimal* within the subset P_0 of placements fulfilling the orientation constraint on T, that is,

$$P_0' = \min\{P : P \in P_0(r)\} . \tag{28}$$

Index set (27) characterizes the *local oriented relative convexity of hypersurface G at point r, with respect to the oriented test hypersurface T.*

The subsets of all points having common index sets,

$$A_{\{\mu i\}} = \{r : r \in G, \ \{\mu_i(r)\} = \{\mu_i\}\} \tag{29}$$

and

$$A_{\{\mu i, P_0\}} = \{r : r \in G, \ \{\mu_{i,P_0}(r)\} = \{\mu_{i,P_0}\}\} \tag{30}$$

generate two partitionings of the boundary G, with respect to the test hypersurface T:

$$G = \cup A_{\{\mu i\}} \tag{31}$$

and

$$G = \cup A_{\{\mu i, P_0\}} . \tag{32}$$

The maximum connected components $D_{\{\mu i\},j}$ and $D_{\{\mu i, P_0\},j}$ of sets $A_{\{\mu i\}}$ and $A_{\{\mu i, P_0\}}$, for all index combinations $\{\mu_i\}$ and $\{\mu_{i,P_0}\}$, respectively, generate finer partitionings of hypersurface G:
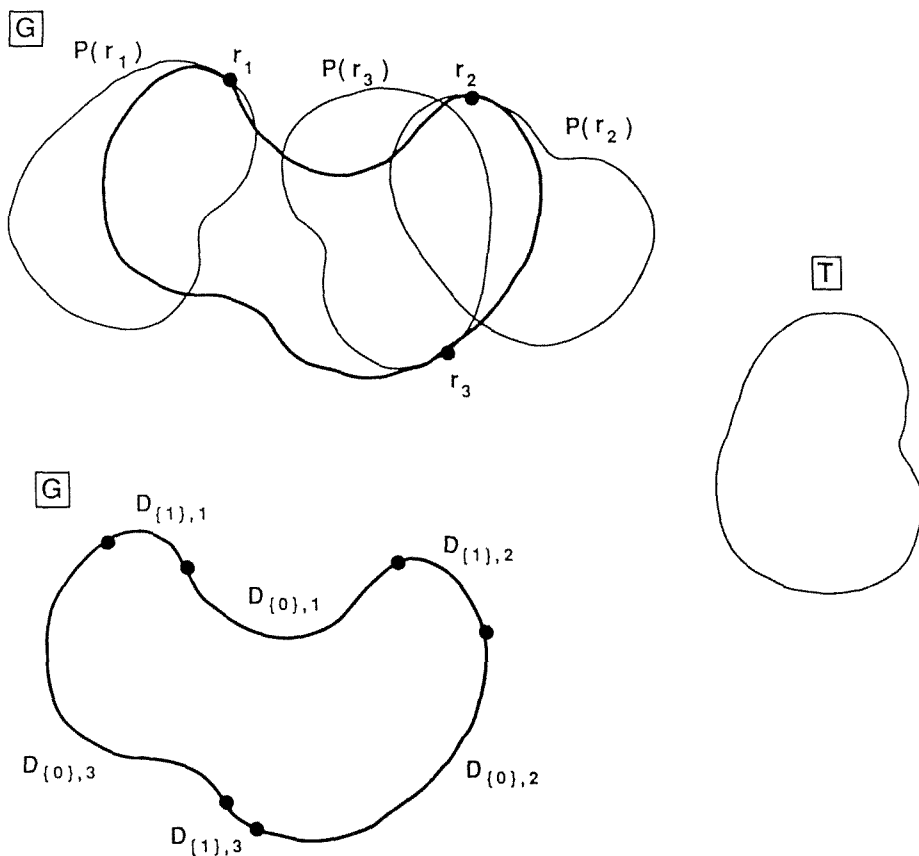
LOCAL RELATIVE CONVEXITY

Fig. 3. Local relative convexity is defined with reference to a test surface T. The twice differentiable boundary G of set C is locally characterized at point r by placing the test surface so as to have a common tangent plane with G at point r, and fulfill the requirements of a minimum placement (see text). Note that in obtaining a minimum placement, there is no restriction on the orientation of test surface T, and both translations and rotations of T are allowed. The local curvature properties of G at point r, relative to those of T, determine the local relative convexity of G at r. Domains $D_{\{0\},j}$ and $D_{\{1\},j}$ of the contour G shown in the figure are locally concave and convex, respectively, relative to test contour T.
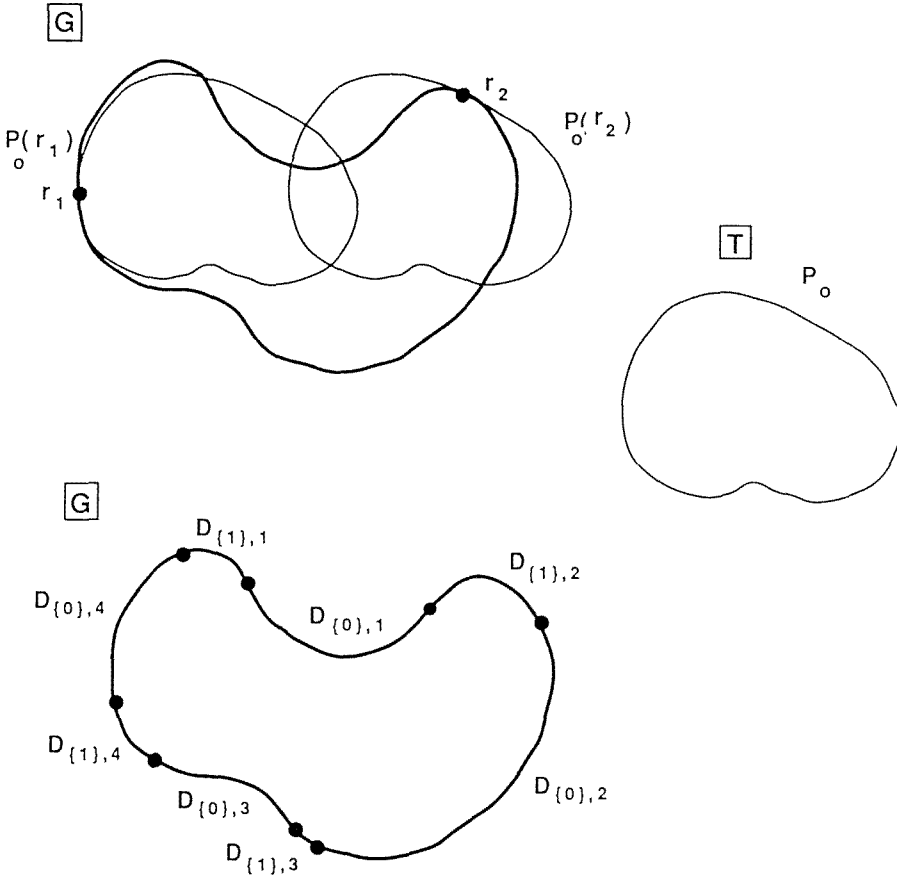
ORIENTED   LOCAL   RELATIVE   CONVEXITY

Fig. 4. Oriented local relative convexity is defined with reference to an oriented test surface T. In obtaining a placement of T, where for the chosen point r of G the tangent planes of G and T coincide, only translations of T are allowed, and T must retain its original orientation. The local curvature properties of G at point r, relative to those of T, determine the oriented local relative convexity of G at r. Domains $D_{\{0\},j}$ and $D_{\{1\},j}$ of the contour G shown in the figure are locally concave and convex, respectively, relative to oriented contour T. Note that the domains $D_{\{0\},j}$ and $D_{\{1\},j}$ obtained in the oriented case are different from those obtained in the non-oriented case, shown in fig. 3 for the same pair G and T of contours.

$$G = \cup \, D_{\{\mu i\}, j} \tag{33}$$

in the case of a freely rotating test hypersurface T, and

$$G = \cup \, D_{\{\mu i, P_0\}, j} \tag{34}$$

in the oriented case.

An illustration of the concepts of *local relative convexity*, and the resulting decomposition of a contour G into various $D_{\{\mu i\}, j}$ domains is given in fig. 3. An alternative $D_{\{\mu i, P_0\}, j}$ domain decomposition of the same contour G, based on *oriented local relative convexity* with reference to a fixed orientation of the same test contour surface T, is shown in fig. 4.

The above D sets generate cellular decompositions [29–32] of hypersurface G of set C, in terms of local convexity properties relative to a free or oriented test hypersurface T. By eliminating certain D sets of selected index sets, topologically significant changes are introduced into the boundary set G, which can be monitored by the homology groups H^k of the resulting truncated boundary set G'. These homology groups give an algebraic characterization of relative and oriented relative local convexity of hypersurface G of set C. Alternative characterization can be given in terms of incidence graphs and shape graphs, using the methods discussed in ref. [33]. Note that an analogous technique, restricted to tangent hyperplanes, has been used earlier for the analysis of potential energy hypersurfaces [34–36].

4.   Applications to molecular surfaces in external fields

For a function $f(r)$ describing a molecular property depending on three-space position vector r, for example, the electronic charge density, electrostatic potential, HOMO and LUMO molecular orbitals, or van der Waals surfaces, etc., a contour set $G(a)$ is defined [24,25] as

$$G(a) = \{r : f(r) = a\}. \tag{35}$$

Such a contour surface $G(a)$ may be regarded as the *boundary* of a *level set*

$$C = C(a) = \{r : f(r) < a\} \quad \text{or} \quad \{C(a) = r : f(r) > a\} \tag{36}$$

of the given molecular property $f(r)$. In general, $G(a)$ is a two-dimensional surface $(n - 1 = 2)$ embedded in a three-dimensional Euclidean space $(n = 3)$.

Global convexity and global relative convexity of such molecular surfaces are certainly relevant to the study of the shape of molecules, especially if one is interested in the overall, approximate shapes such as in the case of certain globular proteins.

However, for detailed shape characterization, which is needed in drug design, the concepts of *local relative convexity* and *oriented local relative convexity* are of greater importance. In the following, we shall elaborate on their applications to the molecular problem.

For a given test surface T, each point $r \in G$ is characterized by the *pair* of numbers, $\mu_1(r), \mu_2(r)$ for the freely rotating case and $\mu_{1,P_0}(r), \mu_{2,P_0}(r)$ for the oriented case, as given by eqs. (26) and (27), respectively. In either case, the surface $G(a)$ is partitioned into the respective connected subsets $D_{\mu1,\mu2,j}$ of the following index combinations:

$$D_{0,0,j}, D_{0,1,j}, D_{0,2,j}, D_{1,0,j}, D_{1,1,j}, \text{ and } D_{2,0,j}. \tag{37}$$

By eliminating domains of a given type, for example by omitting all domains $D_{\mu1,\mu2,j}$ of a specified index pair μ_1, μ_2, one obtains the truncated contour surface

$$G(a, (\mu_1, \mu_2)). \tag{38}$$

Homology groups $H^k(\mu_1, \mu_2)$, $k = 0, 1, 2$, are constructed in the standard way, as described e.g. in refs. [29–32] in full generality, and in refs. [24,25] for the specific chemical problem. These homology groups provide a concise characterization of *local convexity* of $G(a)$ *relative to surface* T. In some instances, it may be advantageous to eliminate two or more types of domains.

If T has some fixed orientation P_0, then for the truncated contour surface

$$G(a, (\mu_{1,0}, \mu_{2,0})) \tag{39}$$

the homology groups $H^k(\mu_{1,0}, \mu_{2,0})$, $k = 0, 1, 2$, are constructured in an analogous manner, and they describe the *oriented local convexity* of $G(a)$ *relative to the given orientation P_0 of the reference surface* T. In some applications, the minimal placement P can be replaced by a *maximal placement P''*, characterized by $P' \leq P''$ for every placement P'.

The above general treatment can be simplified considerably if the test surface T shows some regularities.

The simplest such case is that of the choice of T as a *plane*, also considered in ref. [24]. Here, only the free, non-oriented description is possible; furthermore, $\mu_2 = 0$ in the general case. For every placement P, the freely rotating test plane T coincides with the local tangent plane $^2R(r)$ at every point $r \in G(a)$, and hence

$$b_1 = b_2 = 0. \tag{40}$$

Consequently, the $[b_1, b_2)$ interval is empty and according to definition (22), the index μ_2 is identically zero. The family (37) of general domain types reduces to a family of only three types:

$$D_{0,j}, D_{1,j}, \text{ and } D_{2,j},\tag{41}$$

that is, to concave, saddle and convex domains, where only index μ_1 and serial index j are indicated.

The next level of complexity is represented by the case where the test surface T is chosen as a sphere, in part described in ref. [25]. In this case, orientation has no role and $\mu_2 = 0$, just as in the case of the plane. The two eigenvalues of $\mathbb{H}_T(r)$ coincide,

$$b_1 = b_2 = b.\tag{42}$$

For minimal placements,

$$b \leqslant 0.\tag{43}$$

Replacing the condition of minimal placement with that of a maximal placement leads to fitting the sphere T to $G(a)$ from the outside; for these cases,

$$0 \leqslant b.\tag{44}$$

For both cases, (43) and (44), the partitioning of $G(a)$ is given in terms of three domain types:

$$D_{0,j}(b), D_{1,j}(b), \text{ and } D_{2,j}(b).\tag{45}$$

Within these domains none, one, and two eigenvalues of $\mathbb{H}_{G(a)}(r)$, respectively, are less than the value b.

The above two simplest models, the cases of the test plane [eqs. (40)–(41)] and the test sphere [eqs. (42)–(45)], have been introduced earlier using a different approach [24,25]. The present, general approach, however, is suitable for several additional applications which are not possible within the earlier models. The simplest such case is represented by the choice of T as an ellipsoid. This is the first case discussed here where orientation may play a role. The axes of the ellipsoid may be aligned with an external field; for example, with an external electromagnetic field, or a field generated by a nearby molecule or solid surface.

In fig. 5, the oriented local relative convexity of a simple contour G is described, with reference to two, perpendicular orientations of an ellipse T. In fig. 6, the three domain types of a sphere are shown, defined in terms of oriented local convexity relative to an ellipsoid. Note that within the two D_0 domains, the sphere is locally *concave* with respect to the oriented ellipsoid T, at all points of the connected D_1 set of genus 4, the sphere is locally *saddle-like* relative to the oriented test
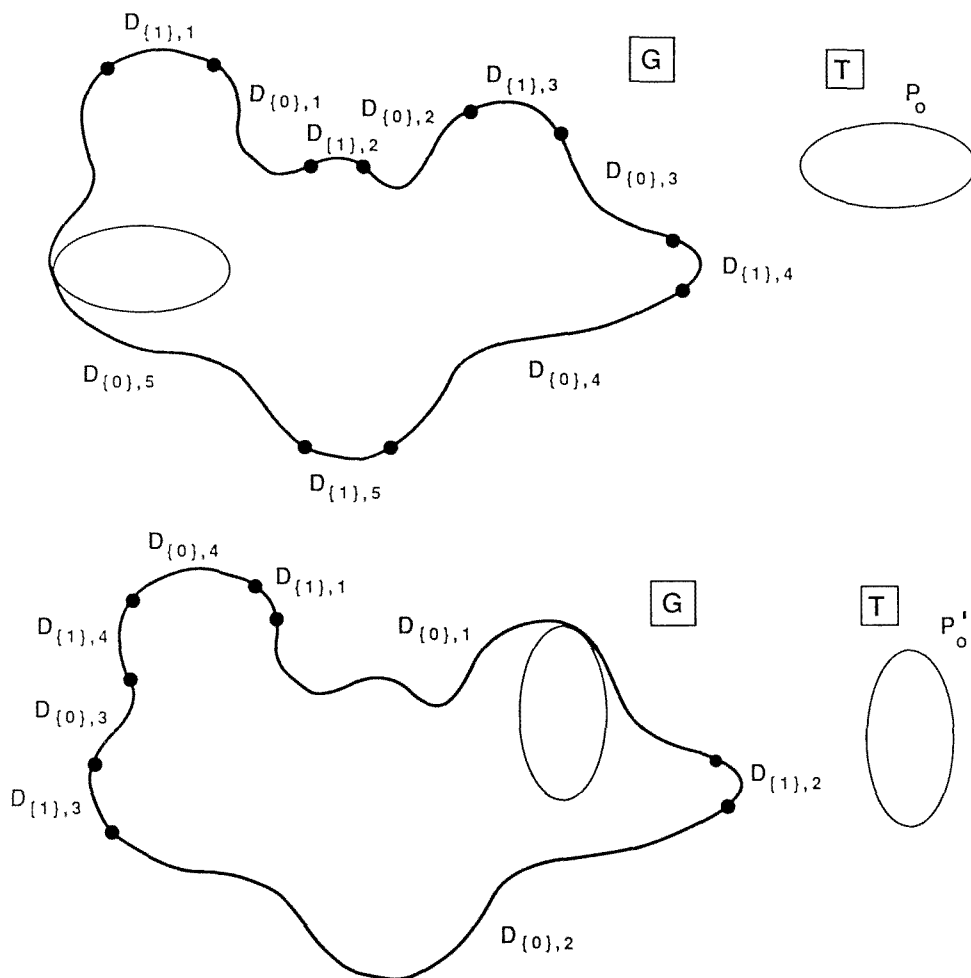
ORIENTED LOCAL RELATIVE CONVEXITY

Fig. 5. Oriented local convexity of a contour G relative to an ellipse as a test contour T. The two, perpendicular orientations of ellipse T lead to two, different patterns of domains $D_{\{0\},j}$ and $D_{\{1\},j}$.

surface T, whereas there are only two, small domains of type D_2 where the sphere is locally convex relative to the oriented ellipsoid.

In both the free and oriented cases one may consider both minimal and maximal placements; in general, each of the four combinations results in a different partitioning of the molecular surface $G(a)$. Since for an ellipsoid T the local Hessian $\mathbb{H}_T(r)$ has in general two different eigenvalues, all six domain types of family (37)
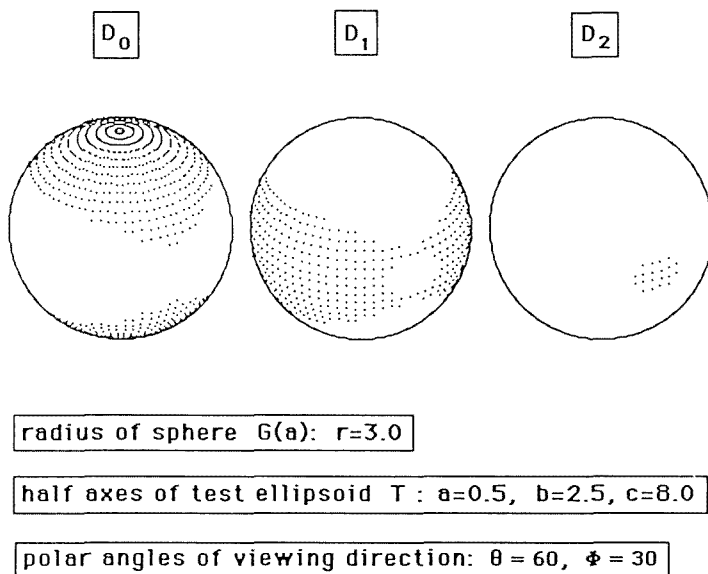
ORIENTED LOCAL RELATIVE CONVEXITY

Fig. 6. A three-dimensional example of oriented local relative convexity. The two-dimensional spherical surface G of the three-dimensional ball of radius 3.0 is characterized relative to an oriented ellipsoid T of half axes $a = 0.5$, $b = 2.5$, and $c = 8.0$, along coordinate axes x, y, and z, respectively. The polar angles of the viewing direction are $\theta = 60°$ and $\phi = 30°$. There are two D_0 domains (locally concave relative to the oriented ellipsoid T), a single D_1 set (locally of saddle shape relative to T), and two, small D_2 domains (locally convex relative to T). The genus of set D_1 is 4.

may occur. For a freely rotating ellipsoid T, the homology groups $H^k(\mu_1, \mu_2)$, $k = 0, 1, 2$, whereas in the case of an oriented ellipsoid T, the homology groups $H^k(\mu_{1,0}, \mu_{2,0})$, $k = 0, 1, 2$, provide a detailed characterization of the shape of molecular surface $G(a)$.

The most general case for the choice of test surface T allows one to express the relative local convexity of one molecular surface $G(a)$ in terms of another molecular surface, by choosing

$$T = G'(a').\tag{46}$$

In particular, the following possibilities are of special interest:

(i) $G'(a')$ is a molecular surface of a *different physical property of the same molecule*. For example, $G(a)$ is an isodensity contour at density a, and the surface $T = G'(a')$ is a van der Waals surface, or a HOMO contour surface.

(ii) $G(a)$ and $G'(a')$ are molecular surfaces for the *same property of two different molecules*. For example, both are isodensity surfaces for a pair of molecules with a common $(a = a')$, or a pair of different electron density values $(a \neq a')$. This leads to a *direct comparison of the shapes of two molecules.*

(iii)   If our goal is to test the *complementarity of shapes of molecules* of specific chemical, biochemical, or pharmacological activity and the shapes of enzyme cavities where they react, then the test surface $T = G'(a')$ may be chosen as the contour surface of the active molecule, and the shape of the contour surface $G(a)$ of the enzyme cavity can be described relative to $G'(a')$.

(iv)   Alternatively, $G(a)$ may be taken as a contour surface of the active molecule; in this case, the test surface $G'(a')$ may be chosen based on the relevant part of an actual contour surface of the cavity region of the enzyme. Since this cavity must have an entrance, the relevant surface region by itself does not usually form a closed surface. However, it can be suitably completed by a surface domain of some specified curvature properties, well distinguished from those of the actual contour of the cavity. The resulting test surface $T = G'(a')$ can be used to describe the shape of the drug molecule relative to the enzyme cavity.

(v)   By taking the *union* $C'(a)$,

$$C'(a) = \cup \ C_i(a) \tag{47}$$

of the level sets $C_i(a)$ of several molecules M_i known to be active within the same type of enzyme cavity, the test surface $T = G'(a')$ may be taken as the boundary of the union $C'(a)$. If the set $\{M_i\}$ of molecules M_i is sufficiently representative, then the *union surface* $G'(a')$ provides a good approximation to the relevant shape features of the enzyme cavity [19].

In practice, it is useful to introduce size constraints on the domains of various types to be considered as topologically significant entities. If a domain has a maximum diameter smaller than a specified threshold, then it may be assumed to belong to a neighbor domain of largest common boundary.

In case (i), the relative shapes of two shape descriptions of the same molecule can be analysed and characterized by the shape group technique [33].

In case (ii), a direct shape comparison is obtained for the two molecules, with respect to the selected molecular property, e.g. charge density. If the relative orientations of the molecules are fixed, then the shape groups $H^k(\mu_{1,0}, \mu_{2,0})$, $k = 0, 1, 2$, obtained using an oriented local relative convexity analysis, provide a detailed characterization.

Choices (iii)–(v) provide means to analyze *shape complementarity* of different molecules. These latter approaches, (iii)–(v), represent alternatives to the description

of shape complementarity based on *complementary truncations* of contour surfaces $G_1(a)$ and $G_2(a)$ of two molecules M_1 and M_2, respectively, and the direct comparison of the resulting shape groups $H^k(\mu_1)$, $k = 0, 1, 2$, and $H^k(\mu_2)$, $k = 0, 1, 2$, or the resulting shape graphs [33,37]. Here, indices μ_1 and μ_2 are those of the complementary domain types $D_{\mu 1}$ and $D_{\mu 2}$ eliminated from the two contour surfaces $G_1(a)$ and $G_2(a)$, respectively, satisfying the shape complementarity condition

$$|\mu_1 - \mu_2| = 2. \tag{48}$$

Examples

In figs. 7 and 8, the contour surfaces of atoms H, C, and O of radii 1.20 Å, 1.70 Å. and 1.40 Å, respectively, are displayed. The various shading patterns indicate a decomposition according to local convexity relative to an oriented ellipsoid of half axes $a = 0.5$ Å, $b = 2.5$ Å, and $c = 8.0$ Å, along coordinate axes x, y, and z, respectively. Atom H is represented by the smallest sphere of large local curvature, and it has the smallest two D_0 domains, which are concave relative to the oriented ellipsoid, elongated along the z axis. The corresponding D_0 domains are larger for the O atom and particularly, for the C atom, which atoms are represented by spheres of larger radii. Within these D_0 domains, the test ellipsoid T osculates the spheres from the inside. At each point of a D_1 domain, the tangent ellipsoid must also cut into the spherical surface. Within domains of the D_2 type, the ellipsoid osculates the spheres from the outside. Pictorially, the domains D_0, D_1, and D_2 are locally concave, saddle-like, and convex, respectively, *relative to the oriented ellipsoid T.*

The exmple of fig. 9 illustrates the oriented local relative convexity of a charge density contour surface of the formaldehyde molecule H_2CO, with respect to an oriented ellipsoid T identical to that used in the examples of figs. 7 and 8. The contour density has been chosen so as to approximate the charge densities along the atomic spheres of the previous examples. If the orientation and the magnitudes of the half axes of the ellipsoid are defined by the orientation and strength of some external electromagnetic fields, then the pattern of domains on the molecular contour surface characterizes the *shape of the molecule relative to these fields.* The most essential feature of these patterns can be described using the shape group method [24,25]. For the $\mu = 0$ truncation, the shape groups H^0, H^1, and H^2 are the infinite cyclic group, a free Abelian group of five generators, and the trivial group, respectively, with Betti numbers
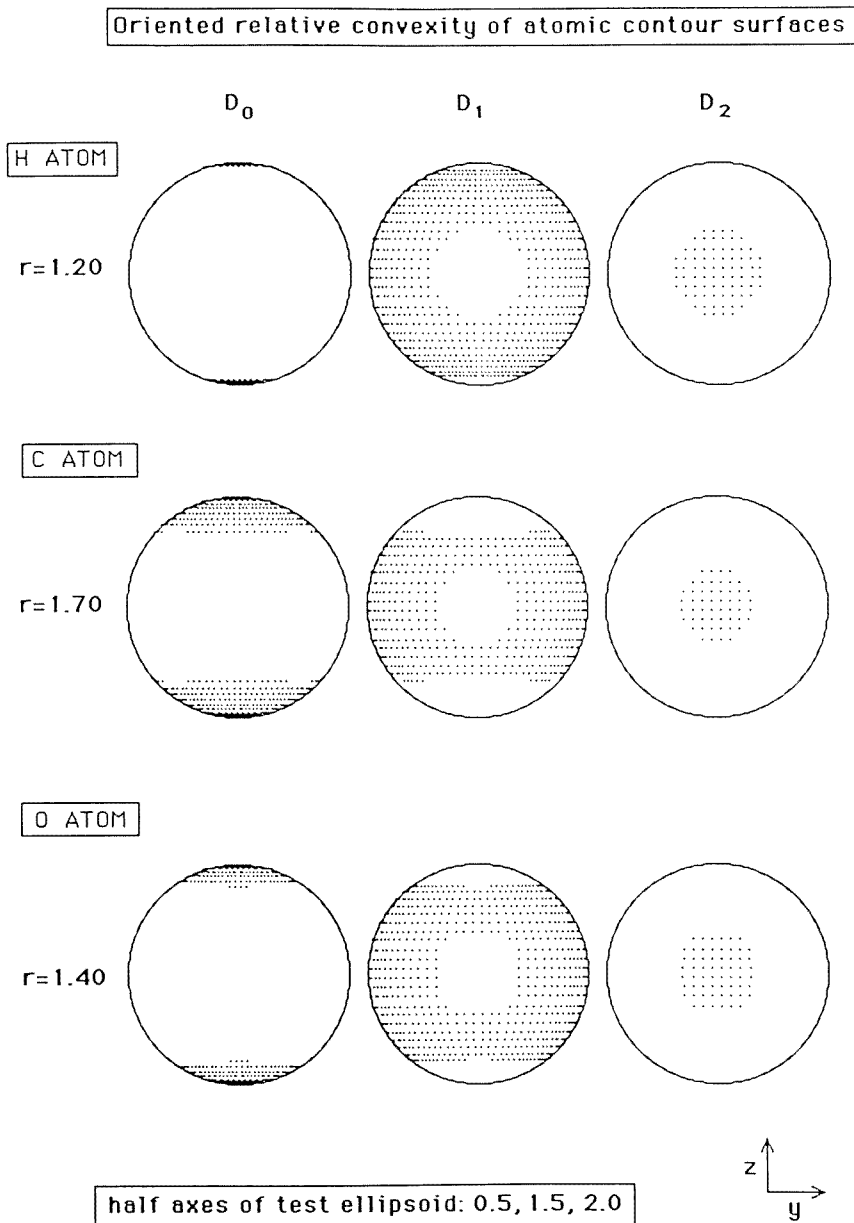
$$B_0 = 1, \ B_1 = 5, \ B_2 = 0. \tag{49}$$

Fig. 7. Contour surfaces of charge densities of atoms hydrogen, carbon, and oxygen, of radii 1.20, 1.70, and 1.40, respectively (approximate van der Waals radii). Using an oriented test ellipsoid T of half axes 0.50, 1.50, and 2.00, along coordinate axes x, y, and z, respectively, the patterns of D_0, D_1, and D_2 domains are displayed. Note that the atomic contours are not displayed on scale.
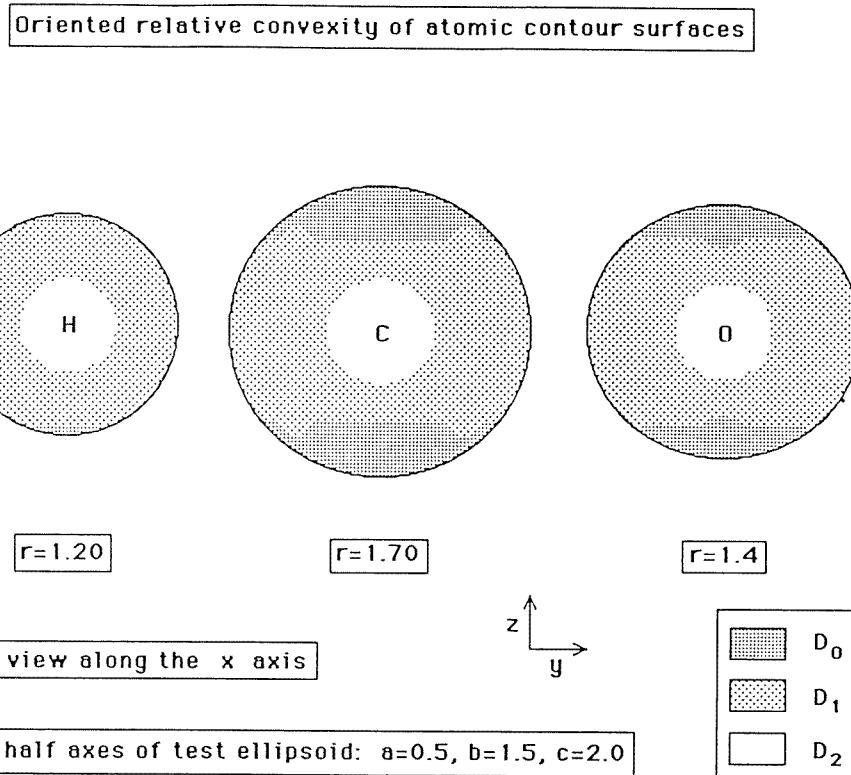
Fig. 8. A representation of the atomic contour surfaces
of fig. 7, indicating all three types of domains on scale.

For the $\mu = 2$ truncation, the respective shape groups are the infinite cyclic group, a free Abelian group of seven generators, and a trivial group, with Betti numbers

$$B_0 = 1, B_1 = 7, B_2 = 0. \tag{50}$$

The above Betti numbers of the shape groups obtained provide a concise description of the oriented local relative convexity of the charge density contour surface of the formaldehyde molecule, with respect to an external field represented by the oriented ellipsoid T.

Oriented local relative convexity is also a possible tool for describing stages in general chemical reactions in which the relative orientation and positioning of the reacting molecules are the determining factors. It also appears to have a role in the characterization of catalytic processes on metal surfaces, within channels and cavities
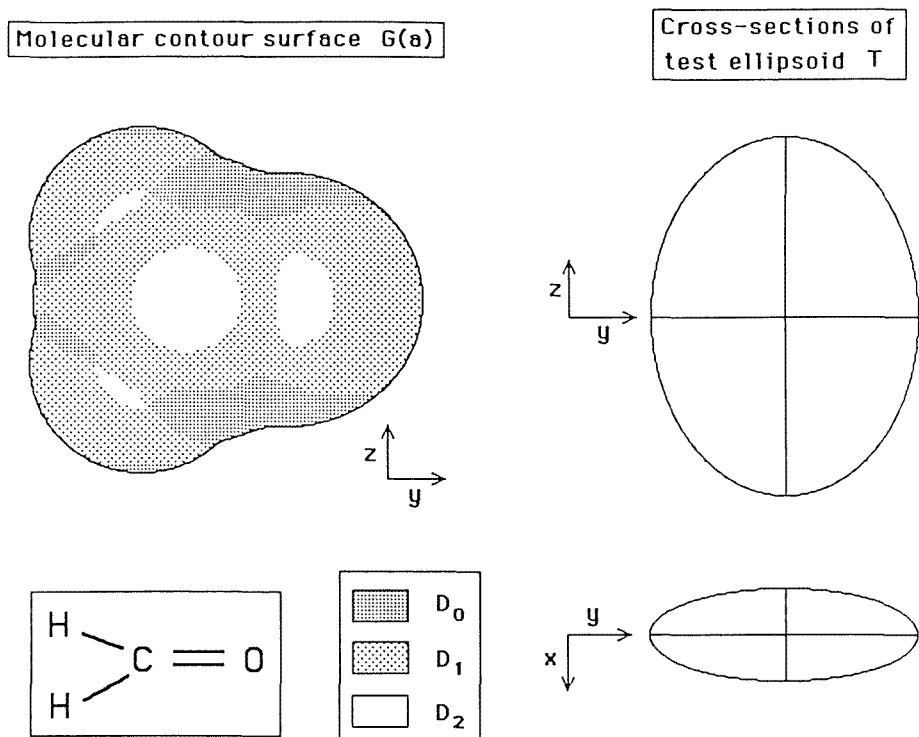
Molecular contour surface G(a)

Cross-sections of
test ellipsoid T

H
  \
    C = O
  /
H

D$_0$

D$_1$

D$_2$

Fig. 9. Oriented local convexity of a charge density contour surface of the formaldehyde molecule H_2CO, relative to an oriented ellipsoid T of half axes 0.50, 1.50, and 2.00, along coordinate axes x, y, and z, respectively. For the $\mu = 0$ truncation, the shape groups H^0, H^1, and H^2 are the infinite cyclic group, a free Abelian group of five generators, and the trivial group, respectively, with Betti numbers $B_0 = 1$, $B_1 = 5$, $B_2 = 0$. For the $\mu = 2$ truncation, the respective shape groups are the infinite cyclic group, a free Abelian group of seven generators, and the trivial group, with Betti numbers $B_0 = 1$, $B_1 = 7$, $B_2 = 0$.

of zeolite lattices, or within the pocket regions of enzymes, where relative orientation and positioning of molecules are of great significance. Polarization effects on molecular shapes in external fields is a subject of great importance in its own right, and oriented relative convexity is a tool for analyzing and comparing the shape differences induced by varying degrees of polarization.

Acknowledgement

References

[1]   I.M. Yaglom and V.G. Boltyanskii, *Convex Figures* (Holt, Rinehart and Winston, New York, 1961).

[2]   A.V. Pogorelov, *Extrinsic Geometry of Convex Surfaces* (Amer. Math. Soc., Providence, R.I., 1973).

[3]   W. Graham Richards, *Quantum Pharmacology*, 2nd Ed. (Butterworth, London, 1983).

[4]   P.A. Kollman, Acc. Chem. Res. 10(1977)365.

[5]   A. Verloop and J. Tipker, A comparative study of new steric parameters in drug design, in: *Biological Activity and Chemical Structure*, ed. J.A. Keverling Buisman, Pharmaco-chemical Library, Vol. 2 (Elsevier, Amsterdam, 1977).

[6]   R. Franke, *Theoretical Drug Design Methods*, Pharmacochemical Library, Vol. 7 (Elsevier, Amsterdam, 1977).

[7]   G.M. Maggiora and R.E. Christoffersen, Quantum mechanical approaches to the study of enzymic transition states and reaction paths, in: *Transition States in Biochemical Processes*, ed. R.D. Gandour and R.L. Schowen (Plenum, New York, 1978).

[8]   M.A. Johnson, V. Nicholson, M. Naim and C.-C. Tsai, in: *QSAR in Drug Design and Toxicology*, ed. D. Hadzi and B. Jerman-Blazic (Elsevier, Amsterdam, 1987).

[9]   A.J. Stuper, W.E. Brügger and P.C. Jurs, *Computer-Assisted Studies of Chemical Structure and Biological Function* (Wiley, New York, 1979).

[10] J. Tomasi, On the use of electrostatic molecular potentials in theoretical investigations on chemical reactivity, in: *Quantum Theory of Chemical Reactions*, Vol. 1, ed. R. Daudel, A. Pullman, L. Salem and A. Veillard (Reidel, Dordrecht, 1979).

[11] Y.C. Martin, *Quantitative Drug Design: A Critical Introduction* (Dekker, New York and Basel, 1978).

[12] I. Motoc, Steric and other structural parameters for QSAR, in: *Steric Fit in Quantitative Structure-Activity Relationships*, ed. A.T. Balaban, A. Chiriac, A. Motoc and Z. Simon, Lecture Notes in Chemistry, No. 15 (Springer-Verlag, Berlin, 1980).

[13] R. Carbó, L. Leyda and M. Arnau, Int. J. Quant. Chem. 17(1980)1185.

[14] G. Náray-Szabó, A. Grofcsik, K. Kósa, M. Kubinyi and A. Martin, J. Comput. Chem. 2 (1981)58.

[15] H. Weinstein, R. Osman, J.P. Green and S. Topiol, Electrostatic potentials as descriptors of molecular reactivity, in: *Chemical Applications of Atomic and Molecular Electrostatic Potentials*, ed. P. Politzer and D.G. Truhlar (Plenum, New York, 1981) pp. 309 – 334.

[16] A. Warshel, Acc. Chem. Res. 14(1981)284.

[17] P.H. Reggio, H. Weinstein, R. Osman and S. Topiol, Int. J. Quant. Chem. Quant. Biol. Symp. 8(1981)373.

[18] M.L. Connolly, J. Mol. Graph 3(1985)19.

[19] R.P. Sheridan, R. Nilakantan, J.D. Scott and R. Venkataraghavan, J. Med. Chem. 29(1986) 899.

[20] J.C. Culberson, G.D. Purvis III, M.C. Zerner and B.A. Seiders, Int. J. Quant. Chem. Quant. Biol. Symp. 13(1986)267.

[21] S.D. Kahn, C.F. Pau, L.E. Overman and W.J. Hehre, J. Amer. Chem. Soc. 108(1986)7381.

[22] G. Náray-Szabó and P.R. Surján, Computational methods for biological systems, in: *Theoretical Chemistry of Biological Systems*, ed. G. Náray-Szabó, Studies in Physical and Theoretical Chemistry, Vol. 41 (Elsevier, Amsterdam, 1986) pp. 1 – 100.

[23] J. Åqvist and O. Tapia, J. Mol. Graph 5(1987)30.

[24] P.G. Mezey, Int. J. Quant. Chem. Quant. Biol. Symp. 12(1986)113.

[25] P.G. Mezey, J. Comput. Chem. 8(1987)462.

[26] P.G. Mezey, Int. J. Quant. Chem. Quant. Biol. Symp. 14(1987)127.

[27] G.A. Arteca and P.G. Mezey, Int. J. Quant. Chem. Quant. Biol. Symp. 14(1987)133.

[28] G.A. Arteca, V. Jammal, P.G. Mezey, J.S. Yadav, M.A. Hermsmeyer and T. Gund, J. Mol. Graph 6(1988)45.

[29] E.H. Spanier, *Algebraic Topology* (MacGraw-Hill, New York, 1966).

[30] M. Greenberg, *Lectures on Algebraic Topology* (Benjamin, New York, 1967).

[31] S.-T. Hu, *Elements of General Topology* (Holden-Day, San Francisco, 1969).

[32] J. Vick, *Homology Theory* (Academic Press, New York, 1973).

[33] P.G. Mezey, J. Math. Chem. 2(1988)299.

[34] P.G. Mezey, Theor. Chim. Acta 54(1980)95.

[35] P.G. Mezey, Theor. Chim. Acta 63(1983)9.

[36] P.G. Mezey, *Potential Energy Hypersurfaces* (Elsevier, Amsterdam, 1987).

[37] G.A. Arteca and P.G. Mezey, to be published.