The Shape of Molecular Charge Distributions: Group Theory without Symmetry

Paul G. Mezey

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

A group theoretical framework is proposed for a detailed characterization of the shapes of electronic charge distributions of general, asymmetric molecules. The proposed shape groups are the homology and cohomology groups of charge density contour surfaces. These shape groups depend on two real parameters, the charge density value a for the contour and a curvature parameter b. The two-parameter family of various homology groups and cohomology groups of charge density contour surfaces is independent of the symmetry properties of the molecules and gives a concise description of the dominant shape characteristics. For any fixed parameter value b these groups may change only at specific charge density values, characteristic to the given molecule. On the other hand, for a fixed-charge density contour the group changes induced by a change in the curvature parameter b provide a description of the fine details of the shape of the electron density. The changes in the structure of these groups follow strict algebraic relations, that provide a quantitative measure for shape-similarity between various molecules. The two-parameter shape group method is an extension of an earlier method proposed for biochemical applications.

INTRODUCTION

The shape of the electronic charge distribution of molecules has a dominant effect on chemical reactivity. Various models and computational methods for the description of molecular shape, as represented by the electronic charge density distribution, are of importance in the interpretation and understanding of chemical reactions, and provide potential tools for the prediction of chemical reactivity. Usually, the most essential features of any mathematical and computational model form an underlying algebraic structure. If the nuclear framework of the actual conformation of the molecule has some nontrivial symmetry, then some of the fundamental aspects of the molecular shape can be described by powerful algebraic methods, using symmetry groups. However, most larger molecules have no nontrivial symmetry, and symmetry group methods are not applicable. Furthermore, molecules are nonrigid, quantum mechanical entities, and even weak interactions of symmetric molecules may lead to symmetry changes or to the loss of symmetry altogether. The power of characterizing shapes in terms of symmetry alone is rather limited.

Of course, asymmetric molecules also have important shape features affecting their reactivity and chemical behavior, and it is desirable to develop computational methods for a concise description of the shape of their charge density distributions. In view of the latest developments of supercomputer technology, methods for rigorous shape characterization are expected to lead to new applications of computational quantum chemistry to industrial chemical research, and in particular, to new approaches to drug design.

The shape of a three-dimensional object having a smooth surface can be characterized by the variations of the curvature along its surface. The surface can be subdivided into domains of specified curvature properties, for example, one may distinguish domains which are convex, concave, or reminiscent to a saddle surface. In the more general case, the domains may be distinguished by finer curvature differences. The pattern of the mutual arrangements of these domains of different curvature properties provides a description of the shape of the original object. This pattern

can be characterized using a topological method based on *homology groups* and *cohomology groups* of algebraic topology.[1] The definition and actual construction of these groups will be given in terms of *formal linear combinations* of various domains, boundary lines and points of the surface (*vide infra*), and no additional mathematical background in topology is required for the development of the proposed method. For example, one may regard a formal linear combination 3A + 7B + 15C of three surface domains A, B, and C of the object as an abstract entity, just as a shopping list of 3 apples, 7 bananas, and 15 carrots may be regarded as a formal, abstract linear combination of an apple, a banana, and a carrot. The essential relations among linear combinations of various domains of the given surface can be formulated using group theory, which groups, in turn, describe the *shape* of the object. All the concepts and definitions necessary for the development of the proposed method are given in the article (*vide infra*), and additional information may be found, for example, in ref. 1–4.

In an earlier study[1] a method has been proposed for the characterization of the shapes of general, e.g., asymmetric molecules, using the curvature properties of contour surfaces $G(a)$ of their charge densities, electrostatic potentials, or molecular orbitals. In the case of the above model the underlying algebraic structure has been provided by the homology groups and cohomology groups of truncated contour surfaces, and in the actual context these groups are collectively referred to as molecular shape groups.[1] In the proposed model the shape groups are dependent on the function value a, characteristic to the contour surface $G(a)$, where a is the actual function value everywhere along the contour surface $G(a)$. Evidently, the shape of this contour surface does depend on the function value a, and for most molecules the shape is simpler (essentially spherical) for contours far away from the molecule, and rather intricate closer to the nuclei. For each function value a, that is, for each contour surface $G(a)$, and also for each truncation, e.g., for the truncated surface obtained by cutting away all convex domains of $G(a)$, three homology groups $\mathbf{H}^0(a)$, $\mathbf{H}^1(a)$, and $\mathbf{H}^2(a)$, as well as three cohomology groups, $\mathbf{H}_0(a)$, $\mathbf{H}_1(a)$, and $\mathbf{H}_2(a)$, can be obtained.[1] By considering collectively all possible values for a, that is, all possible contour surfaces $G(a)$ for the given molecule, for each of the above group type a one-parameter family of shape groups is obtained. The shape groups are invariant between certain parameter values a' and a'', $a' < a \le a''$, for the given conformation, but are replaced by a different set of shape groups at the lower limit a' of this interval and also above the value a''.

This treatment describes only the purely *shape* features and ignores all the *size* features of the molecule, however, the problem of molecular size can be described in terms of the volume enclosed by these contour surfaces. Molecules of identical shape groups for some range of parameter a have similar shape features, and if their sizes are comparable, then they are expected to have similar chemical behavior as well.[1]

In the present study a generalization of the above shape group model is described that improves considerably the shape discriminating power of the method. In particular, by introducing a second parameter, curvature parameter b, fine details of molecular charge distributions can be detected and analyzed, that goes beyond the simple convex, concave, and saddle domain classification used in the earlier study [1].

CURVATURE PROPERTIES OF DENSITY CONTOUR SURFACES

We shall assume that a Cartesian coordinate system is attached to the molecule or molecular fragment under study. The contour surface $G(a)$ of the electronic charge density $f(r)$ is given as the collection of all those points r of the three-dimensional space where the charge density takes the constant value a:

$$G(a) = [r : f(r) = a] \qquad (1)$$

This set may be disconnected, in such a case we may consider the collection of its pieces as a single entity, or we may study just one such piece, that is, one maximum connected component.

An example for contour surface $G(a)$ is shown in Figure 1, showing a possible decomposition into various domains, denoted by D_0, D_1, and D_2. In the earlier model[1] these domains have been identified as concave, saddle, and convex domains, respectively, when viewed from an approaching reagent. Their actual definition has been given as fol-
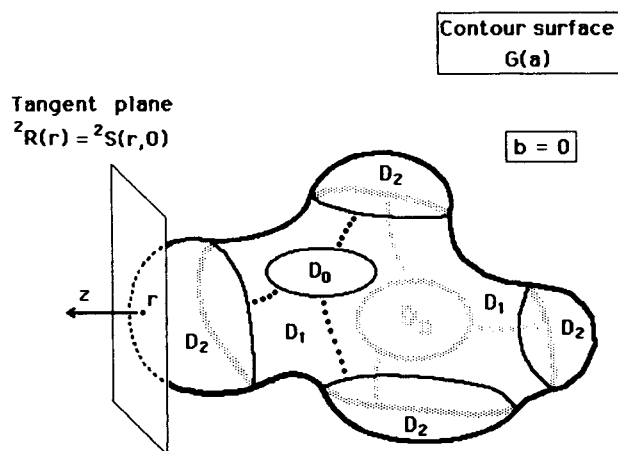
Figure 1. The partitioning of a charge density contour surface $G(a)$ into convex, concave, and saddle domains, using tangent planes.

lows. At each point r of the contour $G(a)$ a unique tangent plane $^2R(r)$ can be given, and the contour surface may be regarded locally as a function defined over this plane, measured along the z axis perpendicular to the plain, and having a critical point at r, that is, a point where the gradient of this function vanishes. The number μ of negative eigenvalues of the Hessian matrix $H(r)$ of second derivatives of this function is the index μ of the D_μ domain to which point r belongs. Pictorially, one may obtain the same decomposition of $G(a)$ into D_μ domains by observing the relative position of the contour surface $G(a)$ and the tangent plane. If the tangent plane $^2R(r)$ lies (locally) on the outside of $G(a)$, then point r is in a convex domain D_2, that is, the Hessian $H(r)$ has two negative eigenvalues, that is, both $h_1, h_2 < 0$, hence $\mu = 2$. If plane $^2R(r)$ lies (locally) inside of $G(a)$, then point r is in a concave domain D_0, that is, the Hessian $H(r)$ has no negative eigenvalues, $\mu = 0$. If within each open neighborhood of r the tangent plane $^2R(r)$ cuts into the contour surface $G(a)$, that is, if $^2R(r)$ has parts both on the outside and in the inside of $G(a)$ arbitrarily close to r, then point r is in a saddle domain D_1, that is, the Hessian $H(r)$ has precisely one negative eigenvalue, $\mu = 1$. These latter D_1 domains have both convex and concave crossections. In the language of differential geometry,[2] the parabolic points on $G(a)$ generate the boundary lines of the above D_μ domain partitioning of $G(a)$. In fact, the computational determination of their boundaries can be accomplished using an algorithm mapping out the lines of parabolic points on the contour surface. The index convention for the

above D_μ domains of $G(a)$ follows from an earlier model developed for the analysis of reactive domains of potential energy hypersurfaces.[3]

The above classification into concave (D_0), convex (D_2), and saddle (D_1) domains is sufficient for the detection and a subsequent homology group analysis of the most essential shape properties of contour surfaces $G(a)$. It does not, however, distinguish between finer details, for example, it does not distinguish between the convex spherical shape and the convex cigar-like shape shown in Figure 2. If, however, *tangent planes are replaced by tangent spheres* of various radii when testing the classification of points r of $G(a)$, then the spherical and cigar-like shapes are clearly distinguished, as illustrated by the two-dimensional crossections shown in Figure 2. Whereas no tangent plane of either object enters their interior (indicated by n for no in Fig. 2), a suitably chosen *tangent sphere* does enter the interior of the cigar-like contour at some points (indicated by y for yes), but not at some other points (indicated by n), and it still does not enter the spherical contour at any of its points. Hence, the spherical and cigar-like contour surfaces are distinguished when using tangent spheres instead of tangent planes. These tangent planes and spheres are used to test the curvature properties of contour surface $G(a)$ at various points r.

Whereas the above pictorial description conveys the essence of the more general classification scheme of the present study, the eigenvalue distribution of the local Hessian matrices of $G(a)$ is a more convenient device for definitions. Let us choose a curvature value b, $-\infty < b < \infty$. To each point r of the given contour surface $G(a)$ an index $\mu_b(r)$ is assigned, according to the following criteria:

$$\mu_b(r) = 0 \quad \text{iff} \quad b \le h_1, h_2 \qquad (2)$$

$$\mu_b(r) = 1 \quad \text{iff} \quad h_1 < b \le h_2 \qquad (3)$$

$$\mu_b(r) = 2 \quad \text{iff} \quad h_1, h_2 < b \qquad (4)$$

We shall use the above $\mu_b(r)$ index for a classification of points r of $G(a)$ into various $D_\mu(b)$ domains. If it does not lead to ambiguity, the specification of parameter b in the above parenthetic notation will be omitted, and we shall write simply D_μ, with the understanding that each such decomposition belongs to a specific pair of parameters a and b. Furthermore, for convenience in terminology
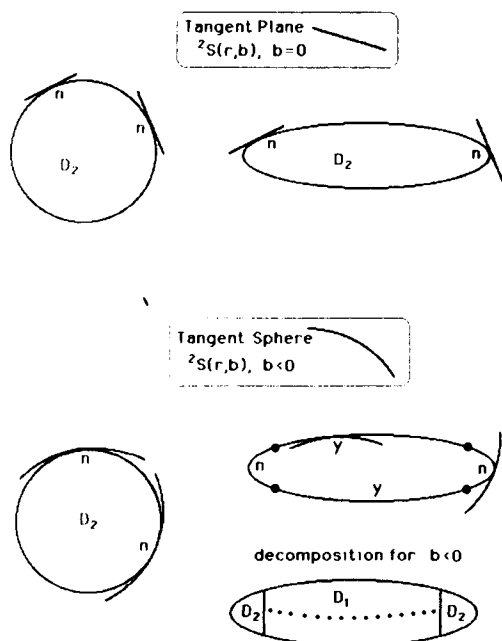
Figure 2. Differentiation of shapes of convex contour surfaces using tangent spheres.

all D_μ sets will be referred to as D_μ domains, even if they are multiply connected, although in such cases a further decomposition may be required for the development of the group theoretical structure.

The curvature parameter b may be regarded as the twofold degenerate eigenvalue of the Hessian matrix $H(r)$ of the tangent sphere, regarded as a function defined over (or, in fact, under, if b is negative) the tangent plane at point r of $G(a)$. Evidently, if the curvature parameter b is zero, $b = 0$, then the tangent sphere becomes identical with the tangent plane, and one obtains the original classification of D_μ domains. A positive b value indicates that the tangent sphere curves away from the interior of $G(a)$, whereas a negative b value indicates that the tangent sphere curves toward the interior of $G(a)$ in the vicinity of the point r.

The cigar-like contour surface of Figure 2 is decomposed into two D_2 and one D_1 domain for a suitably chosen curvature parameter b, whereas the spherical contour surface remains a single D_2 set for the same b value. In Figure 3 cross sections of four typical decompositions of the same contour surface is shown, obtained with curvature parameters $b = 0$, $b < 0$, $b \ll 0$, and $b > 0$, respectively. Intervals along the contour lines are marked by y and n, depending on whether the tangent surface osculates the contour surface from its interior or exterior, respectively. In
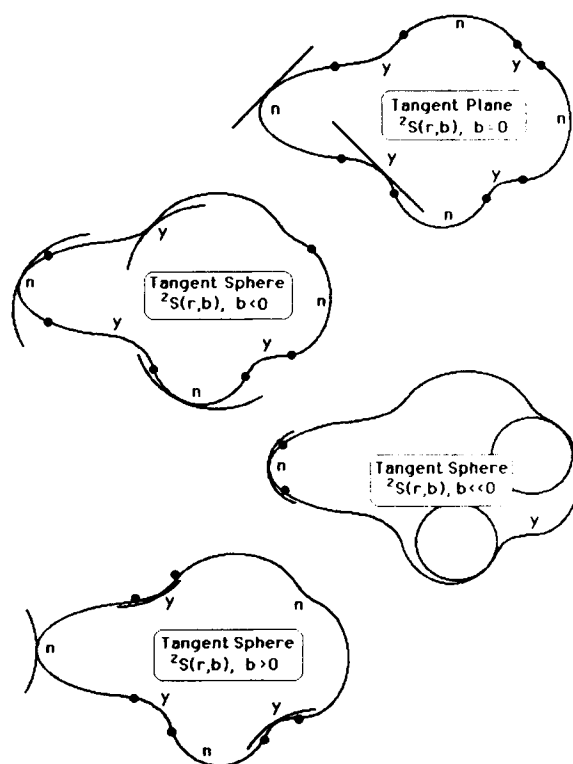

Figure 3. Classification of contours using various curvature parameters for decomposition. See text for details.

Figure 4 a D_μ domain decomposition of the charge density contour surface $G(a)$ of Figure 1 is shown, obtained with a negative b value, that is, with a tangent sphere probing the curvature of $G(a)$ from its interior.

TWO-PARAMETER SHAPE GROUPS OF ELECTRONIC CHARGE DENSITIES: HOMOLOGY GROUPS AND COHOMOLOGY GROUPS OF D_μ DOMAINS OF CONTOUR SURFACES $G(a)$

For each choice of the pair of density parameter a and curvature parameter b a
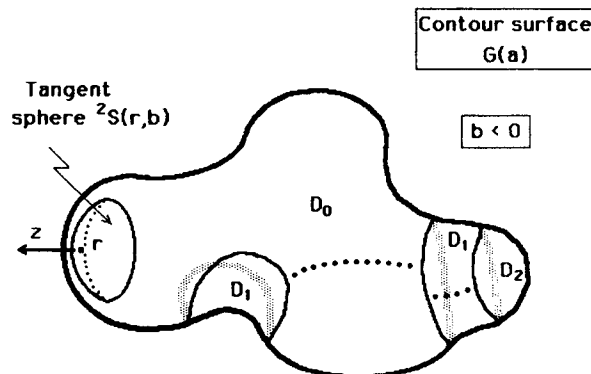

Figure 4. The decomposition of contour surface $G(a)$ of Figure 1, using a different curvature parameter b.

family of D_μ domains is obtained. The D_μ domain partitioning generates a cellular decomposition[4] of the closed and orientable surface $G(a)$. If the D_μ domains are simply connected, then the two-dimensional cells (2-cells) are the open interiors of the D_μ domains themselves. If a D_μ domain is multiply connected (see, e.g., domain D_1 in the example of Fig. 1), then we take a subdivision of this D_μ domain into simply connected subsets, and the interior of each such subset is regarded as a separate 2-cell. A subdivision is shown in Figure 1, where the heavy dotted lines divide D_1 into subsets. As it turns out, the homology groups and cohomology groups are invariant to the choice of the actual subdivisions. The zero dimensional cells (0-cells) are those points where boundary lines and subdivision lines of D_μ domains meet, whereas the one dimensional cells (1-cells) are the open boundary line segments and subdivision line segments between 0-cells.

As can be shown,[4] the algebraic structure of the resulting homology groups is independent (up to isomorphism) of the actual choice of these subdivisions. For a given pair of parameter values a and b we must distinguish three types of 2-cells: those of D_0, D_1, and D_2 domains, respectively. Furthermore, 0-cells and 1-cells may fall on the boundary of a D_0 and a D_1 domain, in which case they are contained in D_0, or they may fall on the boundary of a D_1 and a D_2 domain, in which case they are contained in D_1. In addition, a 1-cell of a subdivision line may fall within the interior of a D_1 domain. Hence, we must distinguish at least two types of 0-cells and possibly three types of 1-cells. In general, we shall use the

$$C(p, \mu, i) \qquad (5)$$

notation for a cell of $G(a)$, specifying the dimension p, the index μ of the D_μ domain which contains the actual cell, and a serial index i of ordering.

A general p-dimensional *chain* of p-cells is defined as the abstract linear combination

$$c_\mu^p = \sum_{i=1} u_i C(p, \mu, i) \qquad (6)$$

where the u_i coefficients are scalars. In this study we shall consider only integer coefficients. We shall also restrict our analysis to special chains, within which the cell index μ is constrained. In the earlier study,[1] describing the special $b = 0$ case, two limiting

values for index μ were taken. In this study we shall specify the index (or indices) μ' that are *excluded* from the above linear combinations. This restriction on the index μ, representing special curvature criteria for the given choice of curvature parameter b, is in fact equivalent to a truncation of the contour surface $G(a)$, by cutting away all $D_{\mu'}$ domains for some specific μ' value. We shall refer to such truncation of $G(a)$ as truncation of type μ'. The resulting truncated contour surface is denoted by $G(a, \mu')$.

Using these special chains as the basis for the definition of homology and cohomology groups is in fact equivalent to defining unconstrained homology groups and cohomology groups for the truncated contour surface $G(a, \mu')$. This latter interpretation lends itself to a simple pictorial model. In either case, the construction of the homology and cohomology groups follows the steps taken in the previous study,[1] and differs from the standard topological treatment[4] only in its diversity of cells of a given dimension p.

The family $\mathbf{C}_{\mu'}^p$ of all such $c_{\mu'}^p$ chains,

$$\mathbf{C}_{\mu'}^p = \{c_{\mu'k}^p\} \qquad (7)$$

form an Abelian group with respect to the addition defined as

$$c_{\mu'1}^p + c_{\mu'2}^p = \sum_i (u_i + u_i') C(p, \mu, i) \qquad (8)$$

where u_i and u_i' are the coefficients of chains $c_{\mu'1}^p$ and $c_{\mu'2}^p$, respectively.

As it has been pointed out, ref. 1, an approaching reagent is likely to favor regions of different curvature properties to a different degree, and selectivity of various reagents depends on the distribution of local concave and convex regions of charge density contours.

By assigning a formal orientation to each cell $C(p, \mu, i)$, the incidence number $n_{ij}(p - 1)$ between a p-cell $C(p, \mu, i)$ and some $(p - 1)$-cell $C(p - 1, \nu, j)$ is defined as follows. The incidence number is zero, if $C(p - 1, \nu, j)$ is not on the set-theoretical boundary of $C(p, \mu, i)$, otherwise it is either $+1$ or -1, depending on whether the relative orientations of these cells agree or not.[4] The algebraic *boundary* of the p-cell $C(p, \mu, i)$ is the $(p - 1)$ chain

$$\Delta C(p, \mu, i) = \sum_j n_{ij}(p - 1) C(p - 1, \nu, j) \qquad (9)$$

The boundary of a zero-dimensional cell $C(0, \mu, i)$ is zero by definition. The boundary of a p-chain $c_\mu^p = \sum_i u_i C(p, \mu, i)$ of p-cells is defined as the $(p - 1)$ chain

$$\Delta c_\mu^p = \sum_i u_i \Delta C(p, \mu, i)$$

$$= \sum_{ij} u_i n_{ij}(p - 1)C(p - 1, \nu, j) \quad (10)$$

A p-chain with zero boundary is a p-cycle. The set of all p-cycles forms a subgroup $\mathbf{Z}_\mu^p$ of group $\mathbf{C}_\mu^p$. The boundary of any $(p + 1)$-chain is a p-cycle, but not every p-cycle is a boundary of a $(p + 1)$-chain. A p-cycle is called a bounding p-cycle if it can be given in the form

$$c_\mu^p = \Delta c_\mu^{p+1} \quad (11)$$

The set of all bounding p-cycles is a subgroup $\mathbf{B}_\mu^p$ of group $\mathbf{Z}_\mu^p$. If for two p-chains $c_{\mu'1}^p$ and $c_{\mu'2}^p$ the difference $c_{\mu'1}^p - c_{\mu'2}^p$ is a bounding p-cycle then these two chains are said to be *homologous*, denoted by

$$c_{\mu'1}^p \sim c_{\mu'2}^p \quad (12)$$

The pth (integral) *homology group* $\mathbf{H}_\mu^p$ of contour surface $G(a, \mu')$ is the difference group (quotient group)

$$\mathbf{H}_\mu^p = \mathbf{Z}_\mu^p - \mathbf{B}_\mu^p \quad (13)$$

The pth homology group $\mathbf{H}_\mu^p$ just as groups $\mathbf{C}_\mu^p$, $\mathbf{Z}_\mu^p$, and $\mathbf{B}_\mu^p$, is a finitely generated Abelian group.

The homology groups of the molecular electron density contours are topological invariants, and homotopy invariants: these groups are identical, up to isomorphism, for any two-charge density contour surfaces $G(a, \mu')$ and $G(a', \mu')$ of two different a and a' charge density values, truncated using some curvature parameters b and b', respectively, as long as the shapes of the two truncated contour surfaces are similar enough to have a continuous deformation between them that preserves the existence and the neighbor relations (but not necessarily the actual shapes) of their respective D_μ domains. Note, that such group isomorphism may very well exist for decompositions induced by two *different* curvature parameters b and b'.

An important quantity, characterizing these homology groups is their Betti number. The pth Betti number b_p for the given μ' truncation is the rank of group $\mathbf{H}_\mu^p$. For closed (n-dimensional) boundary surfaces $G(a)$ the Betti numbers are interrelated by the

Poincaré index theorem:

$$b_p = b_{n-p} \quad (14)$$

The Betti numbers of various truncated contour surfaces $G(a, \mu')$ may serve as a basis for comparing electronic charge densities of different molecules. If two molecular electron densities have the same topological shape characteristics, as expressed by the interrelations among their D_μ domains for a common range of electron density parameter a and curvature parameter b, then the corresponding $G(a, \mu')$ surfaces must have identical sets of Betti numbers.

The construction of the *cohomology groups* of electronic charge densities follows a strategy analogous to that of the homology groups. The *coboundary* of a p-chain $c_\mu^p = \sum_i u_i(p, \mu, i)$ of p cells is defined as the $(p + 1)$-chain

$$\Diamond c_\mu^p = \sum_i u_i \Diamond C(p, \mu, i)$$

$$= \sum_{ij} u_i n_{ji}(p)C(p + 1, \delta, j) \quad (15)$$

A p-chain with zero coboundary is a p-cocycle. The family of cocycles forms a subgroup $\mathbf{Z}_{p\mu'}$ of group $\mathbf{C}_\mu^p$. The set of those cocycles which are coboundaries (the cobounding cocycles) forms a subgroup $\mathbf{B}_{p\mu'}$ of group $\mathbf{Z}_{p\mu'}$. The difference group (quotient group)

$$\mathbf{H}_{p\mu'} = \mathbf{Z}_{p\mu'} - \mathbf{B}_{p\mu'} \quad (16)$$

is the p-dimensional (integer) *cohomology group* of the contour surface $G(a, \mu')$.

The cohomology groups are also topological invariants of the given (truncated) contour surface of the molecular electronic charge density. Similarly to the homology groups, for a common range of parameters a and b an isomorphism between the cohomology groups of two different molecules implies common shape features.

Consider the example of Figure 1. If no truncation is carried out on $G(a)$, and if one takes the curvature parameter $b = 0$, then the homology groups $\mathbf{H}^0$ and $\mathbf{H}^2$, as well as the cohomology groups $\mathbf{H}_0$ and $\mathbf{H}_2$ of contour surface $G(a)$ are isomorphic to the additive group of integers, and the groups $\mathbf{H}^1$ and $\mathbf{H}_1$ are trivial groups. More complicated groups arise when some of the D_μ domains of specific curvature properties are removed. A $\mu' = 2$ type truncation, that is, the removal of D_2 domains,

leading to truncated contour surface $G(a, 2)$, results in a homology group $\mathbf{H}_2^1$ that is isomorphic to the Abelian group of three free generators g_i composed from elements of the form

$$k_1 g_1 + k_2 g_2 + k_3 g_3 \quad (k_1 \text{ are integers}) \tag{17}$$

In the same example, the homology group $\mathbf{H}_0^1$ of truncated contour surface $G(a, 0)$, obtained by a $\mu' = 0$ type truncation (i.e., by the removal of D_0 domains), is isomorphic to the infinite cyclic group, having elements of the form

$$k g_1 \quad (k \text{ integers}) \tag{18}$$

for some generator g_1.

The very same contour surface gives rise to a different set of homology groups (that, in turn, provides a more precise shape characterization), if a different curvature parameter b is chosen, as illustrated by the case of a negative parameter value, $b < 0$, shown in Figure 4. For this parameter value both of the $\mu' = 2$ and $\mu' = 0$ truncations lead to trivial one-dimensional homology groups:

$$\mathbf{H}_2^1 = \mathbf{H}_0^1 = \{1\} \tag{19}$$

where $\{1\}$ stands for the trivial group of a single unit element

It is useful to consider the half plane defined by the ranges of electron density contour parameter a and curvature parameter b:

$$0 < a < \infty \tag{20}$$

$$-\infty < b < \infty \tag{21}$$

The sign convention used for electron density is consistent with the conventions of quantum mechanical probability densities: the electron density is regarded as a nonnegative function.

One copy of the parameter plane can be associated with each truncation type μ', and a detailed description of molecular shape features is equivalent to the assignment of the actual homology and cohomology groups to each point of the parameter plane. In fact, on the above half plane the homology groups and cohomology groups of each μ' truncation are invariant within various parameter domains, as illustrated schematically in Figure 5 for a T-shaped arrangement of an A + B_2 atomdiatom system. The two-parameter shape groups may undergo changes only at the boundaries of these parameter domains. Near

the $a = 0$ axis, where points on the parameter plane represent electron density contours of low-electron density, usually loose contours far removed from the molecule, the groups are rather simple, since most contours are essentially spherical, with no detailed features. Depending on the truncation type μ', above or below certain b values the truncation may formally eliminate the entire contour surface $G(a)$. For example, this is the case for the $\mu' = 2$ type truncation of D_2 domains of an essentially spherical contour surface $G(a)$ of a low a value, if the curvature parameter b is positive or zero: $b \geq 0$. For higher density values a for the contour surfaces $G(a)$ a greater variety of groups appear, that belong to domains of the parameter plane removed from the $a = 0$ axis. At very high a value the electron density contour surfaces become disjoint, essentially separating into nearly spherical contours surrounding each individual nucleus, and a further increase in the a value does not cause any modification of these groups. In the case of a $\mu' = 2$ type truncation these contours of high a value are eliminated and no groups are obtained, except for very low b values (see Fig. 5). The chemically most interesting part of the parameter plane is that of intermediate a and b values. Usually, at the boundaries of parameter domains the homology and cohomology groups change only by one generator, although molecular symmetry may imply the simultaneous appearance or disappearance of several group generators. A comparison of the distribution of these groups along the parameter planes of two molecules provides a precise, quantitative basis for the comparison of the shapes of their electronic charge densities.

SUMMARY

A detailed description of the shapes of molecular electronic charge densities is possible using homology and cohomology groups of their contour surfaces $G(a)$, which groups (shape groups) depend on two parameters: on the electron density value a at the contour surface, and on a curvature parameter b, used for the classification of domains of the contour, in terms of different curvature properties. Homology group invariants, in particular, the Betti numbers, can be used for a concise characterization of these shapes.
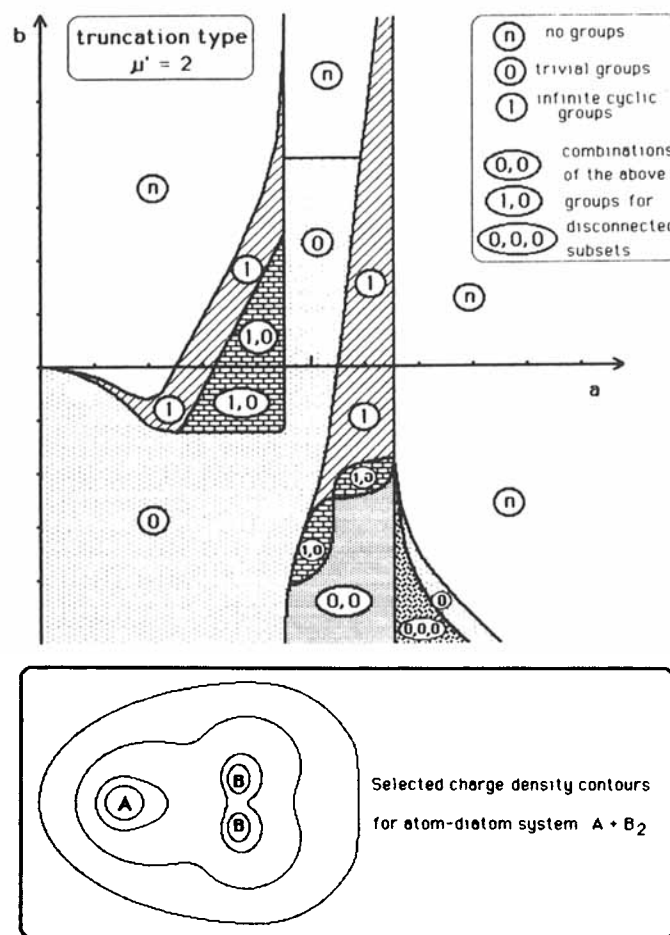
Figure 5. A schematic representation of charge density shape group distribution along the parameter plane (a, b), for truncation type $\mu' = 2$ of a T-shaped atom-diatom system A + B_2.

These shape groups and their two-parameter families are independent of the symmetry properties of the molecules and give a concise description of the dominant shape characteristics. The interrelations among shape groups of a given molecule can be analyzed over a two dimensional parameter plane. For any fixed parameter value b these groups may change only at specific charge density values, characteristic to the given molecule. On the other hand, for a fixed a value of the charge density contour the group changes induced by changes in the curvature parameter b provide a description of the fine details of the shape of the electron density. The changes in the structure of these groups follow simple rules, the most common case being the inclusion or elimination of a single group generator. A comparison of shape groups for a given parameter domain provides a quantitative measure for shape similarity between different molecules.

References

1. P. G. Mezey, *Int. J. Quantum Chem., Quant. Biol. Symp.,* **12**, 113 (1986) (paper presented at the 1985 Sanibel Symposium I).
2. L. P. Eisenhart, *Riemannian Geometry,* Princeton Univ. Press, Princeton, 1963.
3. P. G. Mezey, *Theor. Chim. Acta,* **54**, 95 (1980).
4. E. H. Spanier, *Algebraic Topology,* McGraw-Hill, New York, 1966.