THE DEGREE OF SIMILARITY OF THREE-DIMENSIONAL BODIES: APPLICATION TO MOLECULAR SHAPE ANALYSIS

Paul G. MEZEY
Department of Chemistry and Department of Mathematics, University of Saskatchewan, Saskatoon, Canada S7N 0W0

Abstract

Similarity of shape features of three-dimensional bodies is of importance in many fields. Computational methods that are suitable to provide numerical measures for such similarities are expected to find applications in a wide variety of areas. Whereas relative measures based on direct pair comparisons are useful, nevertheless, methods that involve absolute shape descriptors are expected to be more universally applicable. The general "grade of similarity" concept proposed in this study is based on such absolute shape descriptors of three-dimensional bodies. The study of similarity of the three-dimensional shapes of molecules as represented, for example, by their electronic charge distributions, or electrostatic potentials, or simply by their fused spheres Van der Waals surfaces, is an important component of modern drug design. A family of topological methods, the shape group methods (SGM), have been proposed recently for the study of the shapes of formal molecular bodies, evaluating and comparing numerical shape codes for the non-visual comparison of molecules by the computer. In this contribution a new, and conceptually simpler numerical measure of shape similarity is proposed, applicable for the computer evaluation of similarity of arbitrary three-dimensional objects of closed surfaces. The technique is suggested for the non-visual, numerical evaluation of shape similarity of formal molecular bodies and contour surfaces.

1. Introduction

The concepts of similarity index, the degree of similarity, and the associated numerical measures proposed for three-dimensional similarity analysis in this study are motivated by some of the basic patterns of visual perception of the shape features of ordinary objects. We shall use two-dimensional examples in order to introduce the main ideas; however, the definitions and the general method are designed for three-dimensional applications and are easily generalized for arbitrary finite dimensions. Consider two domains D and D' in the plane. For an observer at a great distance from the plane, both domains may appear as mere points, hence their shapes appear virtually identical. At a slightly closer distance, one may be able to distinguish their shapes, but only some of the most prominent differences in their shapes may be detected. At a much closer distance, however, many details may appear and a more thorough shape comparison becomes possible. Evidently, the evaluation of their similarity is dependent on the level of resolution. If two domains show differences

already at a low level of resolution, then they are less similar than two domains which appear different only at a much higher level of resolution. Hence, a degree of similarity or dissimilarity can be associated with the level of resolution required to detect shape differences.

In this study, a simple and systematic method is suggested that provides a size-independent scale for the level of resolution. Based on this scale, definitions for a *similarity index* and for the *degree of similarity* are proposed. In the following sections, we shall derive and give precise definitions of similarity indices and the degree of similarity for both the two-dimensional case of planar domains and the three-dimensional case of bodies enclosed by closed contour surfaces. The extension of these concepts to the shape analysis of abstract, higher-dimensional objects is straightforward and will be commented on in the conclusions section of the paper.

2. The degree of similarity of planar domains

The comparison of two-dimensional shapes of closed curves and areas enclosed by them in the plane is an important task in many areas of natural sciences, including chemistry. Experimental observations and chemical measurements, as well as theoretical studies, often result in such curves, and their comparisons often lead to important insights. For example, in catalysis, the pattern of adsorbed molecules on metal surfaces can be described by contour lines [1], and in drug design, the cross-sections of molecular contour surfaces can be modeled by such closed curves [2,3]. Some of the shape analysis methods are based on graph theory [4] or on homology group theory of algebraic topology [5,6]. The latter methods are collectively called the shape group methods (SGM) [7–10].

In this contribution, we shall follow a different strategy for shape description. First, we shall describe the essential idea behind the proposed method in simple terms. To this end, we focus our attention on a family of simple curves, called Jordan curves, that are suitable for the construction of more complicated curves and patterns in the plane. A Jordan curve divides the plane into two parts: a bounded domain (the interior of the Jordan curve) and the remaining, unbounded subset of the plane (the exterior of the Jordan curve).

Consider a Jordan curve J in the plane and the planar domain D that is the interior of J. The interior of J can be modeled by square-cell configurations; for example, one may place J on a square grid and consider the family of all squares falling within the interior of J. Subject to some constraints (see below), these square-cell configurations are called animals, and have been used in various applications [11–16]. In the present context, we use animals to provide a natural scale for the level of resolution of shape descriptions. The smaller the squares of the grid, the better the resolution of the representation of J by the animals. By approximately filling up the interior of J by animals at various levels of resolution, a shape characterization of the continuous Jordan curve J can be obtained by the shape characterization of animals. The animals contain a finite number of squares (cells);

hence, the latter shape characterization can be accomplished using the methods of discrete mathematics. Hence, one obtains an approximate, *discrete* characterization of the shape of a Jordan curve, that is, the shape of a *continuum*. The level of resolution can be represented indirectly by the number of cells of the animals. In particular, we show below that the number of cells required to distinguish between two Jordan curves provides a numerical measure of their similarity.

In order to formulate the above ideas in precise terms, we need some definitions, which are reviewed below. Some properties and applications of animals and some of the relevant background can be found in refs. [4,11–20].

A mesh on a square lattice is a family of squares forming a rectangular part of the lattice. In graph-theoretic terms, a *mesh* $M_{m,n}$ is defined [14] as the Cartesian product $P_m \times P_n$ of two nontrivial paths P_m and P_n. The notation M_n may be used for the mesh $P_n \times P_n$. A Jordan cycle C of a mesh M_n is a cycle that is a subgraph of M_n and has a vertex degree of two for all of its nodes. A subgraph A of a mesh M_n is called an *animal* if it contains all the nodes and edges of mesh M_n that fall on or within the interior of a Jordan cycle C of M_n. Each 4-cycle C_4 contained in animal A is called a *cell* c of A. The Jordan cycle C is the *perimeter* of A. As a point set, the perimeter C of A is a single Jordan curve, denoted by $J(A)$. The perimeter $J(A)$ contains all edges of A which are on exactly one of its cells. Note that other, more general definitions of animals are also known which allow for multiply connected square-cell configurations.

Our goal is to characterize the shape of a Jordan curve J and its interior D by the shapes of animals that fit within J. When placing an animal A within D, the relative orientation of J and mesh $M(A)$ is not fixed, that is, J and A may be rotated with respect to one another. (The case with orientation constraints leads to the orientation-dependent shape characterization that is of importance for molecules in external fields, for example, for drug molecules within enzyme cavities [10]. The corresponding similarity measures are easily derived from the ones of the present study, and these problems will not be considered here.)

Clearly, by choosing a small enough size s for the length of the side of the square cells, any finite animal can fit within the given planar domain D. Evidently, whether an animal A fits within the interior D of a given Jordan curve J depends on the relative size of J and the cells of the animal. For a given Jordan curve J and cell size s, there exists a countable family $F(J, s)$ of animals which fit within D. If the size s is too large, then this family is empty. With reference to J and s, the members $A_i(J, s)$ of this family $F(J, s)$ are the *inscribed animals* of D.

For a given J and s, there exists a maximum number n of cells for inscribed animals. The subset $F(J, s, n)$ of $F(J, s)$ contains all animals $A_i(J, s, n)$ of the maximum possible cell number n. Without introducing orientation constraints, this number $n = n(J, s)$ depends only on the given Jordan curve J and cell size s. The perimeters of these $A_i(J, s, n)$ animals are also Jordan curves, approximating the original Jordan curve J. The smaller the cell size s, the better the resolution and the better the approximation. Since a small change of cell size s does not necessarily change

n, one may consider size ranges. For the given J, we find the maximal interval $s(J, n) = [s(n, 1), s(n, 2))$ within which the n value is invariant, and we generate the union $F(J, n)$ of all $F(J, s, n)$ animal families for this interval:

$$F(J, n) = \bigcup_{s(J,n)} F(J, s, n). \tag{1}$$

The family $F(J, n)$ contains all n-cell animals $A_i(J, n)$ inscribed within the interior D of Jordan curve J, with the maximum number n of cells for the given range $s(J, n)$ of cell size s. Since on the given level of resolution no animal with more cells can be inscribed within D, these animals are the n-cell *interior filling animals* of the Jordan curve J.

In general, the n-cell animal $A_i(J, n)$ inscribed in Jordan curve J is an *interior filling animal* of J if and only if no animal of the same cell size s and more than n cells can be inscribed in J. For example, none of the interior filling animals $A_i(J, n)$ can be enlarged by a cell and still fit within the interior J as long as $s \in s(J, n)$.

The level of resolution depends on the relative size of $D = Int(J)$ and the cells c. This relative size is implied by the maximum number n of cells which fit within domain D. In other words, the *continuum* of size range $s(J, n)$ is replaced by a *discrete* descriptor, integer n. For this reason, in the $F(J, n)$ and $A_i(J, n)$ notations the cell size information is not given directly.

In fig. 1, three Jordan curves J_1, J_2, and J_3 are shown, with some of their interior filling animals. At level $n = 3$, there is only one interior filling animal, common to all three curves. Hence, at this level of resolution their shapes appear the same. At level $n = 4$, however, only the two curves J_1 and J_2 have the given common interior filling animal, which is different from the one for the curve J_3. At a higher level, for example at $n = 18$, the interior filling animals of all three curves are necessarily different. Since it requires a higher level of resolution to distinguish the shapes of the pair J_1, J_2 than that needed for either of the pairs J_1, J_3 and J_2, J_3, we conclude that the closest similarity is between the shapes of J_1 and J_2. Most human observers would find the same conclusion based on visual inspection. These conclusions based on the figure suggest a more precise treatment for the evaluation of similarity, as given below.

The family $F(J, n)$ of all interior filling animals $A_i(J, n)$ of the Jordan curve at level n provides an absolute shape characterization of the curve J and its interior D. The $F(J, n)$ sets are also suitable to introduce a relative measure for shape similarity of two Jordan curves J_1 and J_2. At a given level n, the intersection

$$F(J_1, J_2, n) = F(J_1, n) \cap F(J_2, n) \tag{2}$$

contains all the common interior filling animals $A_i(J, n)$.

We say that Jordan curves J_1 and J_2 are *dissimilar* at and above cell number n_c if each set $F(J_1, J_2, n)$ is empty if $n \geq n_c$.

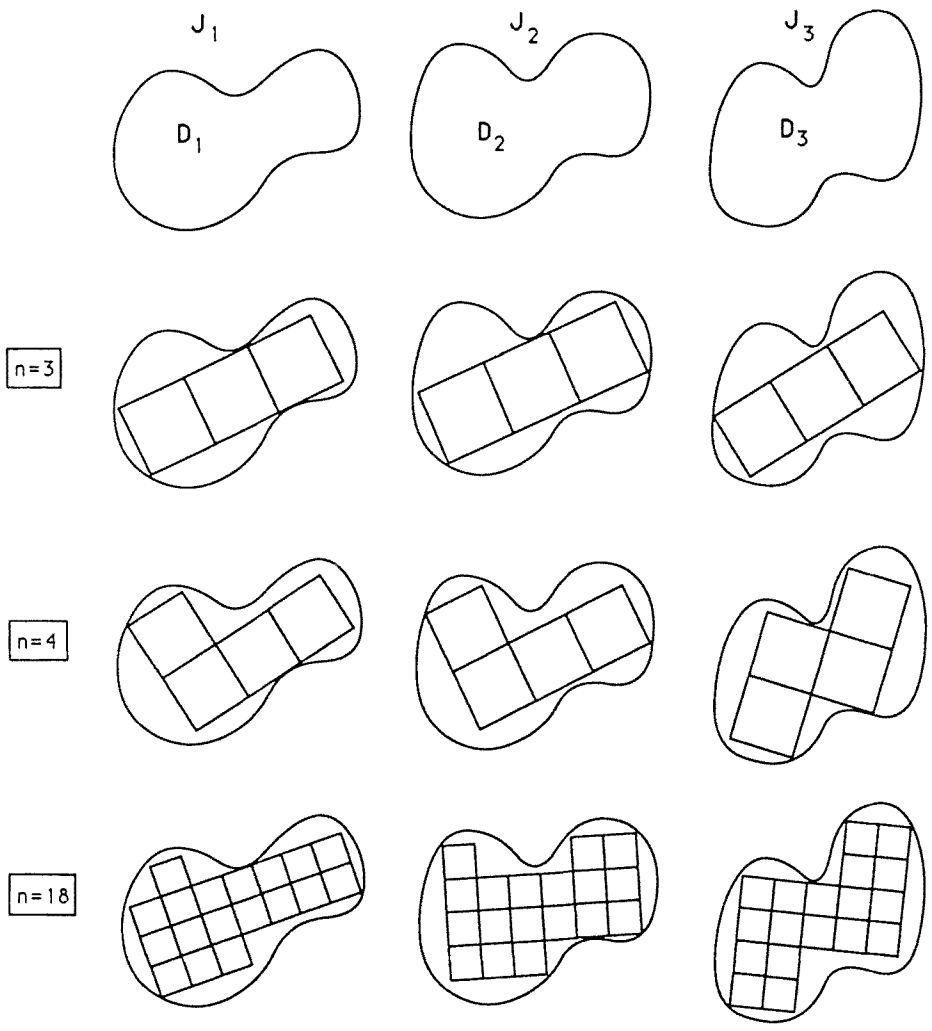

Fig. 1. Some interior filling animals of three planar domains D_1, D_2, and D_3 of different shapes. Note that the greatest degree of similarity is found between domains D_1 and D_2, in agreement with expectation based on visual inspection.

The similarity of the shapes of Jordan curves $J_1$and J_2 at the level of resolution n can be characterized by the number $f(J_1, J_2, n)$, defined as

$$f(J_1, J_2, n) = card(F(J_1, J_2, n))/card(F(J_1, n) \cup F(J_2, n)), \tag{3}$$

that is, as the ratio of the number of common interior filling animals and the total number of interior filling animals of the two curves. In the above notation, $card(X)$ is the cardinality of the countable set X.

Note that for special Jordan curves J_1 and J_2, and for special n values, the set $F(J_1, n)$, or set $F(J_2, n)$, or both, may be empty. If both sets are empty, then for these n values no number $f(J_1, J_2, n)$ is defined. The possible values of $f(J_1, J_2, n)$ range from 0 to 1, indicating the lower and upper bounds of similarity between two Jordan curves J_1 and J_2 at the given level n of resolution.

The shape comparison is based on interior filling animals, and it is independent of the relative sizes of the Jordan curves J_1 and J_2. Consequently, the numbers $f(J_1, J_2, n)$, as well as all other shape and similarity descriptors of this study, are also independent of the sizes of J_1 and J_2.

The family of all nonzero $f(J_1, J_2, n)$ numbers provides a detailed description of similarity, but such a set of numbers is inconvenient for comparisons. We shall take advantage of the n-dependence of numbers $f(J_1, J_2, n)$ in order to devise a single numerical similarity measure. The number $f(J_1, J_2, n)$ is not necessarily a monotonic function of n, although for any two different Jordan curves J_1 and J_2, the overall tendency is a decrease of $f(J_1, J_2, n)$ with n. In particular, it is possible that $f(J_1, J_2, n) = 0$ and $f(J_1, J_2, n') > 0$, where $n < n'$. This justifies our choice for the *similarity index*, defined below:

The *similarity index* $i_0(J_1, J_2)$ of two Jordan curves J_1 and J_2 is the smallest n_c value at and above which all interior filling animals of Jordan curves J_1 and J_2 are different, that is,

$$i_0(J_1, J_2) = \begin{cases} \min\{n_c : F(J_1, J_2, n) \text{ is empty if } n \geq n_c\}, \text{if the minimum exists;} \\ \infty \quad \text{otherwise.} \end{cases} \tag{4}$$

If the shapes of two Jordan curves J_1 and J_2 are identical (that is, if they can be obtained from one another by scaling), then no finite n_c value exists and $i_0(J_1, J_2) = \infty$. For curves J_1 and J_2 of non-identical shapes, the more similar their shapes, the greater the cell number n of the largest common interior filling animals. Consequently, the similarity index $i_0(J_1, J_2)$ is a large number if the two Jordan curves J_1 and J_2 are very similar, and it is a small number for highly dissimilar curves.

The degree of dissimilarity $d(J_1, J_2)$ is defined in terms of $i_0(J_1, J_2)$ as follows:

$$d(J_1, J_2) = 1/(i_0(J_1, J_2) - 2). \tag{5}$$

The smallest cell number n at which there exist different animals is three, which justifies the inclusion of the number two in the denominator. The degree of dissimilarity $d(J_1, J_2)$ may take values from the [0, 1] interval, greater values indicating greater dissimilarity. If the Jordan curves J_1 and J_2 have identical shapes, then $d(J_1, J_2) = 0$.

This leads us to the *degree of similarity* $s(J_1, J_2)$, defined as

$$s(J_1, J_2) = 1 - d(J_1, J_2). \tag{6}$$

If the Jordan curves J_1 and J_2 have identical shapes, then their degree of similarity $s(J_1, J_2) = 1$, otherwise $s(J_1, J_2)$ is a smaller non-negative number.

The similarity index $i_0(J_1, J_2)$ and the degree of similarity $s(J_1, J_2)$ of two Jordan curves J_1 and J_2 are of general applicability for the evaluation of similarities of planar domains. In chemistry, these tools for the quantification of similarity are applicable in a wide variety of fields, for example, for the shape comparison of curves defined as cross-sections of molecular contour surfaces, or contours of molecular aggregates, or patterns of molecules adsorbed on metallic surfaces, important in studies of catalysis.

In the following section, the shape characterization technique and the concept of degree of similarity will be extended to three-dimensional objects such as formal molecular bodies and molecular boundary surfaces, using polycubes as the three-dimensional analogues of square cell configurations.

3. Polycubes and the degree of similarity concept for three-dimensional bodies

We shall consider the problem of shape comparison of three-dimensional bodies in the context of molecules; however, the techniques and the formal measures proposed for similarity are applicable to general bodies.

Various topological techniques have been proposed earlier for the analysis of shapes of molecular contour surfaces and formal molecular bodies (see, for example, refs. [2,3,7–10,21–24]). Among these surfaces, the isodensity contours $G(a)$ at electronic density value a, and the bodies $B(a)$ enclosed by them, are of special importance in rationalizing chemical and biochemical processes. By analogy with the two-dimensional similarity analysis of Jordan curves and interior filling animals, described above, we shall develop a three-dimensional technique based on the generation of *polycubes* enclosed by the contour surfaces $G(a)$.

A connected arrangement of a finite number n of impenetrable cubes C of uniform size is called a polycube if only three types of contacts between cubes are allowed: common face, common edge or common vertex. If $n > 1$, then each cube of the polycube P must have a face contact with another cube of P.

One may regard polycubes as parts of a cubic lattice. The smallest rectangular block of the cubic lattice that contains polycube P is called the *mesh* $M(P)$ of P.

The polycubes considered in this study fulfill the following three additional restrictions:

(i) if there is an edge contact between two cubes C and C' of P, then there must also be a face contact between C and C', or there must exist a cube C'' of P having face contact with both C and C';

(ii) if there is a vertex contact between two cubes C and C' of P, then there must also be either an edge contact between them, or there must exist two cubes C'' and C''' of P with face contact to each other and C'' having face contact to C and C''' having face contact to C';

(iii) the polycube P, as a single body, is topologically equivalent to the three-dimensional body it represents. In the most common case, a formal molecular body $B(a)$ is topologically equivalent to a solid ball; however, toroidal or more complicated topologies are also possible.

Conditions (i) and (ii) ensure that the polycubes are facewise connected, whereas condition (iii) is the natural requirement that the three-dimensional body and its representations at various levels of resolution are not incomparably different.

By analogy with the perimeters of animals and Jordan cycles, the *surface* $G(P)$ of a polycube P is the point set union of all those faces of the cubes C of P that are on precisely one cube. The surfaces of polycubes will be used to approximate molecular contour surfaces $G(a)$, and to characterize the shapes of the formal molecular bodies $B(a)$ enclosed by them.

The size of the cubes C is characterized by the uniform edge length s. By gradually decreasing s and increasing the number n of cubes in the polycube P inscribed within $G(a)$, one can approximate the formal molecular body $B(a)$ at increasing levels of resolution.

Consider a given molecular contour surface $G(a)$. If the size s of the cubes is chosen small enough, then any finite polycube P can fit within $G(a)$. As in the two-dimensional case, we do not consider orientation constraints and we assume that the contour surface $G(a)$ and polycube P may be rotated with respect to one another; the relative orientation of $G(a)$ and mesh $M(P)$ is not fixed. Hence, in our present model, the identity of a polycube is independent of its orientation. We consider two polycubes P and P' identical if and only if they can be superimposed on one another by translation and rotation in 3D space. The polycube method of shape analysis with orientation constraints, suitable for the study of molecular recognition and shape problems in external fields, will be discussed elsewhere.

If the molecular contour surface $G(a)$ and cube size s are given, then there exists a countable family $F(G(a), s)$ of polycubes which fit within $G(a)$. If the size s is too large, then the family $F(G(a), s)$ is empty. The polycubes $P_i(G(a), s)$ of this family are the *inscribed polycubes* of size s. For a given size s and contour surface $G(a)$, there exists a maximum number $n(G(a), s)$ of cubes for inscribed polycubes. A small change of the electronic contour density parameter a may leave the value $n(G(a), s)$ invariant. Similarly, for a fixed a value, a small change of cube size s does not necessarily change the value $n(G(a), s)$; hence, $n(G(a), s)$ is invariant for some size range, denoted by $s(G(a), n)$. Polycubes $P_i(G(a), n)$ with the maximum number $n = n(G(a), s)$ of cubes for the given range $s(G(a), n)$ are the n-cube *interior filling polycubes* of the contour surface $G(a)$.

In simpler terms, a polycube $P_i(G(a), n)$ is an *interior filling polycube* of the contour surface $G(a)$ if and only if no polycube P of the same cube size s and of $n + 1$ cubes can be inscribed in $G(a)$.

We shall use the shape properties of interior filling polycubes $P_i(G(a), n)$ inscribed in molecular contour surfaces $G(a)$ in order to assess the similarity of the $G(a)$ contours and the formal molecular bodies $B(a)$ enclosed by them. In order to use levels of resolution scaled relative to the molecular size, we do not use the absolute size parameter s directly.

At a given size s, one can provide only a few details of a small object of delicate features, but the same cube size s can be sufficiently descriptive for a large object of cruder features. One obtains more comparable characterizations if both small and large objects are described by the *same number* of cubes. Consequently, it is more suitable for our purposes if each level of resolution is defined by n, which depends on the relative size of the object as compared to the cube size s. That is, the formal level of resolution at which the similarities of various $G(a)$ surfaces and $B(a)$ bodies are analyzed is represented in absolute terms by the cube size s, and in a size-independent way by the number n of cubes of the interior filling polycubes $P_i(G(a), n)$.

Let us denote the family of all interior filling polycubes $P_i(G, n)$ of the molecular contour surface G at level n by $F(G, n)$. This set $F(G, n)$ provides an absolute shape characterization of G and the body B enclosed by it. By analogy with the two-dimensional case, we may use these $F(G, n)$ sets to introduce a relative measure for shape similarity of two molecular contour surfaces G_1 and G_2. These surfaces may belong to two different molecules, or to the same molecule with two different contour density values a_1 and a_2. At a given level n of resolution, the intersection

$$F(G_1, G_2, n) = F(G_1, n) \cap F(G_2, n) \tag{7}$$

contains all the common interior filling polycubes $P_i(G, n)$.

We shall use the term *dissimilar* in the following context: the contour surfaces G_1 and G_2 (and the bodies B_1 and B_2 enclosed by them) are *dissimilar* at and above cube number n_c if each set $F(G_1, G_2, n)$ is empty if $n \geq n_c$.

At each level of resolution n, we may characterize the similarity of the shapes of contour surfaces G_1 and G_2 by the number $f(G_1, G_2, n)$, defined by

$$f(J_1, J_2, n) = card(F(G_1, G_2, n))/card(F(G_1, n) \cup F(G_2, n)). \tag{8}$$

For special n values and special contour surfaces G_1 and G_2, both sets $F(G_1, n)$ and $F(G_2, n)$ may be empty; for these n values, no number $f(G_1, G_2, n)$ is defined. If the number $f(G_1, G_2, n)$ exists at the given level n of resolution, then it ranges from 0 to 1, where the extreme values correspond to the lower and upper limits of similarity between the two contour surfaces G_1 and G_2.

We emphasize that the shape comparison is based on interior filling polycubes containing the same number of cubes; hence, it is independent of the relative sizes of the contour surfaces G_1 and G_2.

Instead of using the family of all nonzero $f(G_1, G_2, n)$ numbers for a detailed description of similarity, we shall use a single numerical similarity measure. By analogy with the two-dimensional case, we introduce the *similarity index* $i_0(G_1, G_2)$, defined below:

The *similarity index* $i_0(G_1, G_2)$ of two contour surfaces G_1 and G_2 is the smallest n_c value at and above which all interior filling polycubes of contour surfaces G_1 and G_2 are different, that is,

$$i_0(G_1, G_2) = \begin{cases} \min\{n_c : F(G_1, G_2, n) \text{ is empty if } n \geq n_c\}, \text{ if the minimum exists;} \\ \infty \quad \text{otherwise.} \end{cases} \tag{9}$$

If two contour surfaces G_1 and G_2 can be obtained from one another by scaling, then we say that their shapes are identical. For contour surfaces G_1 and G_2 of identical shapes, no finite n_c value exists and $i_0(G_1, G_2) = \infty$. For contour surfaces G_1 and G_2 of non-identical shapes, if the cell number n of the largest common interior filling polycubes is a larger number, then we perceive them as exhibiting greater similarity. This is reflected in the similarity index $i_0(G_1, G_2)$, which is a large integer number if the two contour surfaces G_1 and G_2 are very similar, and a smaller integer if they are highly dissimilar.

We define the degree of dissimilarity $d(G_1, G_2)$ as follows:

$$d(G_1, G_2) = 1/(i_0(G_1, G_2) - 2). \tag{10}$$

The smallest cube number n at which there exist different polycubes is three; this is reflected by the inclusion of the number two in the denominator. The degree of dissimilarity $d(G_1, G_2)$ takes values from the [0, 1] interval, greater values indicating greater dissimilarity. If the contour surfaces G_1 and G_2 have identical shapes, then their degree of dissimilarity is zero, $d(G_1, G_2) = 0$.

The *degree of similarity* $s(G_1, G_2)$ of two contour surfaces G_1 and G_2 is defined as

$$s(G_1, G_2) = 1 - d(G_1, G_2). \tag{11}$$

If the two contour surfaces G_1 and G_2 have identical shapes, then their degree of similarity $s(G_1, G_2) = 1$, otherwise it is a smaller non-negative number.

4. Conclusions

The concepts of two- and three-dimensional similarity indices and the degree of similarity have been motivated by potential chemical applications, both in theoretical chemistry and in applied fields such as pharmaceutical drug design. The underlying principle is the replacement of a continuum in the plane or in the three-dimensional space by a discrete representation in terms of simple objects, where the construction of these simpler objects is carried out on various levels of resolution. The level of resolution required to detect various shape features and similarities leads to the numerical similarity index and to the degree of similarity. A similar principle may be applied to the development of a concept of the degree of chirality, to be described elsewhere [19,20].

Note that these concepts are not restricted to chemistry, and are applicable to more general problems of shape analysis.

The index of similarity and the degree of similarity are easily extended to m-dimensional abstract objects of continuous $(m - 1)$-dimensional boundaries by replacing

the polycubes with m-dimensional poly-hypercubes obeying a face-connectedness condition for their $(m-1)$-dimensional faces. Applications of the higher dimensional case to a similarity analysis of multidimensional potential energy hypersurfaces and their reaction globe representations [25] will be presented elsewhere.

Acknowledgements

It is a pleasure to acknowledge stimulating discussions with Professor Frank Harary on lattice animals and polycubes, which have motivated this study, and interesting conversations with Professor Stuart Whittington and Professor Chris Soteros on embedding problems on cubic lattices. This work was supported in part by the Computational Chemistry Unit of the Upjohn Laboratories, Kalamazoo, Michigan, and by both strategic and operating research grants from the Natural Sciences and Engineering Research Council of Canada.

References

[1] G.A. Somorjai, J. Phys. Chem. 94(1990)1013.
[2] F. Harary and P.G. Mezey, J. Math. Chem. 2(1988)377.
[3] G.A. Arteca and P.G. Mezey, Theor. Chim. Acta 75(1989)333.
[4] F. Harary, *Graph Theory* (Addison–Wesley, Reading, 1969).
[5] E.H. Spanier, *Algebraic Topology* (McGraw–Hill, New York, 1966).
[6] I.M. Singer and J.A. Thorpe, *Lecture Notes on Elementary Topology and Geometry* (Springer, New York, 1976).
[7] P.G. Mezey, Int. J. Quant. Chem. Quant. Biol. Symp. 12(1986)113.
[8] P.G. Mezey, J. Comput. Chem. 8(1987)462.
[9] P.G. Mezey, J. Math. Chem. 2(1988)299.
[10] P.G. Mezey, J. Math. Chem. 2(1988)325.
[11] S. Golomb, *Polyominoes* (Springer, New York, 1965).
[12] F. Harary, The cell growth problem and its attempted solutions, Beitrage zur Graphentheorie 49(1968).
[13] G. Exoo and F. Harary, Indian Nat. Acad. Sci. Lett. 10(1987)67.
[14] F. Harary and M. Lewinter, Int. J. Comput. Math. 25(1988)1.
[15] C. Soteros and S.G. Whittington, J. Phys. A21(1988)2187.
[16] N. Madras, C. Soteros and S.G. Whittington, J. Phys. A21(1988)4617.
[17] F. Harary and E.M. Palmer, *Graphical Enumeration* (Academic Press, New York, 1973).
[18] M. Gardner, Sci. Amer. 240(1979)18.
[19] F. Harary and P.G. Mezey, Similarity and complexity of the shapes of square-cell configurations, in: *New Developments in Molecular Chirality*, ed. P.G. Mezey (Kluwer, Dordrecht, 1991).
[20] P.G. Mezey, A global approach to molecular chirality, in: *New Developments in Molecular Chirality*, ed. P.G. Mezey (Kluwer, Dordrecht, 1991).
[21] G.A. Arteca and P.G. Mezey, Int. J. Quant. Chem. 34(1988)517.
[22] G.A. Arteca and P.G. Mezey, J. Phys. Chem. 93(1989)4746.
[23] G.A. Arteca and P.G. Mezey, J. Math. Chem. 3(1989)43.
[24] G.A. Arteca, G.A. Heal and P.G. Mezey, Theor. Chim. Acta 76(1990)377.
[25] P.G. Mezey, *Potential Energy Hypersurfaces* (Elsevier, Amsterdam, 1987).