

Entropy in dissimilarity and chirality measures

Remi Chauvin

*Laboratoire de Chimie de Coordination du CNRS, Unité 8241,
liée par conventions à l'Université Paul Sabatier et à l'Institut National Polytechnique,
205 Route de Narbonne, 31 077 Toulouse Cedex, France*

Received 17 July 1995; revised 26 October 1995

Within the prospect of quantifying the geometrical dissimilarity of molecular models on the basis of a thermodynamical formalism, the algebra of stereogenic pairing equilibria is reviewed and applied to molecular geometry: developing Rassat's proposition, an "interaction energy" of two figures F and F' is taken as proportional to $d_H^2(F, F')$, where d_H denotes the Hausdorff distance. If G is a group of rotations in E_n , the geometrical version of the general equation (E) of the "chemical algebra" defines a distance extension $D_p(F, F')$ of $d_H(F, F')$, which is independent of the orientations of F and F' , and where the coefficient p is interpreted as the reciprocal of a "temperature-like" parameter: $p \propto 1/T$. At 0 K ($p = \infty$), no formal entropy contributes to the definition of the uniform distance D_∞ . At ∞ K ($p = 0$), the discrimination between homo- and hetero-pairing of figures by the harmonic distance D_0 is averaged over orientation states. Temperature-dependent chirality measures c_p are derived from D_p , and c_∞ is analogous to Mislow's chirality measure. If T and σT are normalized enantiomorphic triangles with coincident centroids in E_2 , $c_p(T) = D_p(T, \sigma T)$ is calculated for $p = 0$ and $p = \infty$, and discussed for $0 < p < \infty$. Finally, the Hausdorff interaction model is putatively related to energy profiles versus dihedral angle in *meso*- and *dl*-molecules.

1. Introduction

In current molecular representations, a geometrical skeleton is first drawn, labelled by relevant ligand parameters (atomic orbitals, force field parameters, atomic charges, etc.) and then allowed to warp more or less according to the temperature. These three steps rely respectively on the concepts of *geometry*, *energy* and *entropy* (fig. 1) [1]. Within the framework of this "molecular axiomatics", the ultimate link between the similarity concept and the pairing phenomenon would support the relevance of a similarity measure based on the mathematical formalism of the thermodynamical analysis of pairing equilibria. The "algebra of stereogenic equilibria" and the "chemical algebra" (see below) afford possible similarity measures of two sets of atomic energy parameters mapping identical geometrical skeletons. This paper deals with the complementary similarity measure of two geometries mapped with identical (unspecific) parameters. An emphasis will be

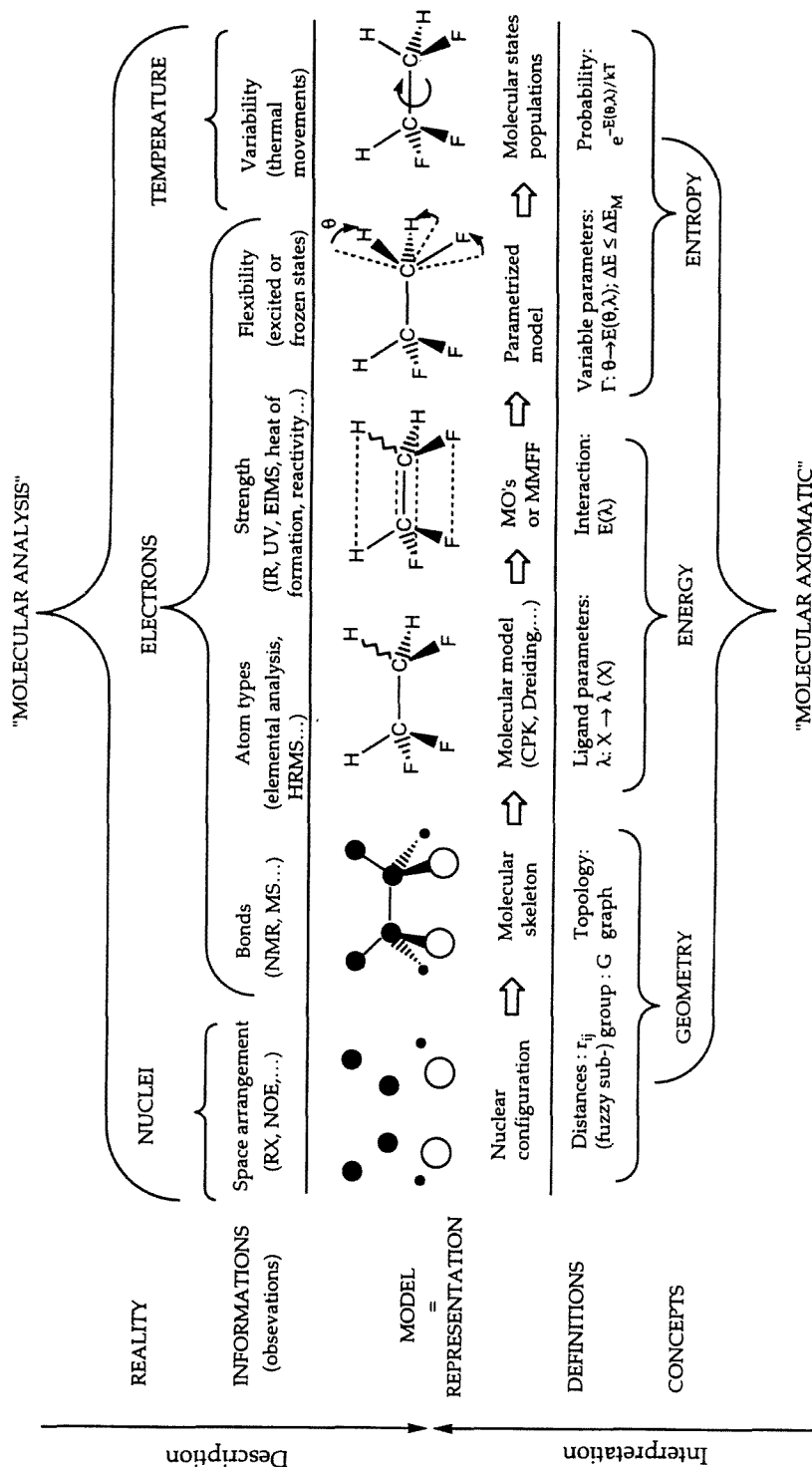
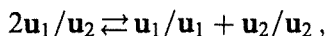


Fig. 1. Apprehension of the notion of molecule. Current molecular representations are refined as indicated from the left to the right. Each step can be reached by either describing one more information issued from observations (top) or interpreting one more abstract definition suggested by seemingly not less than three distinct concepts (bottom): geometry, energy and entropy (the topology here refers to the *nuclear topology* which serves to define the *stereo-electronic geometry* of the molecule). Beyond a mathematical formalism, eq. (E) devised from the algebra of stereogenic equilibria, though oversimplified to directly describe the chemical reality, brings out general trends and makes a connection between these concepts. It may be relevant as a basis for a "molecular axiomatic".

laid on chirality, which is the underlying concept of important technical and intellectual challenges in modern chemistry [2].

2. Summary of the algebra of stereogenic pairing equilibria

The shift of a pairing equilibrium between chemical species represented by \mathbf{u}_1 and \mathbf{u}_2 ,



where a slash bar (/) represents some interaction, is given by the constant

$$K = \frac{[\mathbf{u}_1/\mathbf{u}_1][\mathbf{u}_2/\mathbf{u}_2]}{[\mathbf{u}_1/\mathbf{u}_2]^2},$$

where $[\mathbf{u}_i/\mathbf{u}_j]$ denotes the concentration of the paired species $\mathbf{u}_i/\mathbf{u}_j$ [3]. Owing to their chemical and biochemical importance, complex association equilibria $P + Q \rightleftharpoons P/Q$ have given rise to many *analytical* treatments [4]. In another spirit, the former pairing equilibria lend themselves to an *axiomatic* algebraic treatment as well. The latter has been formulated on the basis of three hypotheses [3]:

- (1) *Skeleton symmetrization*: skeletons of interacting molecules are identically symmetrized in a realistic manner.
- (2) *Skeleton overlap pairing*: the geometry of the paired species is a tight two-skeleton juxtaposition wherein the skeletons are parallel.
- (3) *Scalar product form of the ligand interactions*: only one kind of pairwise ligand interaction occurs, and the corresponding energy is proportional to the scalar product of two n -vectors of real or vector ligand parameters assigned to n skeletal sites numbered $1, \dots, n$ (e.g. four vertices of a regular tetrahedron, $n = 4$).

\mathbf{u}_1 and \mathbf{u}_2 are identified with vectors $\mathbf{u}_1 = (\lambda_1(\mathbf{u}_1), \dots, \lambda_n(\mathbf{u}_1))$ and $\mathbf{u}_2 = (\lambda_1(\mathbf{u}_2), \dots, \lambda_n(\mathbf{u}_2))$, where $\lambda: k \rightarrow \lambda_k$ is some kind of ligand parameter (charge, polarizability, dipole moment, orbital, etc.) which is valued at each skeletal site k of a given molecule. λ may also be a phenomenological parameter without a direct physical meaning, but fitting experimental K values. In the same spirit, scalar ligand parameters have been used by Ugi and Ruch in the stereochemical analogy model [5].

The relevance of these hypotheses has been previously discussed [3]. Considering the pairing stereogenicity resulting from the symmetry of the skeleton as an "entropic contribution", the pairing equilibrium constant is expressed by

$$K = \frac{\left(\sum_{g \in G} \exp \left[-\frac{a(g\mathbf{u}_1 | \mathbf{u}_1)}{r^q kT} \right] \right) \left(\sum_{g \in G} \exp \left[-\frac{a(g\mathbf{u}_2 | \mathbf{u}_2)}{r^q kT} \right] \right)}{\left(\sum_{g \in G} \exp \left[-\frac{a(g\mathbf{u}_1 | \mathbf{u}_2)}{r^q kT} \right] \right)^2},$$

wherein $\frac{a}{r^q}(g\mathbf{u}_i | \mathbf{u}_j)$ is the interaction energy in the pair $g\mathbf{u}_i/\mathbf{u}_j$; $(\cdot | \cdot)$ denotes a scalar product (e.g. a bracket); g is a symmetry operation belonging to the point-group G of the skeleton and defining the relative orientation of \mathbf{u}_i and \mathbf{u}_j in one "stereoisomer" of the pair $\mathbf{u}_i/\mathbf{u}_j$; a and q are fixed parameters; r is the (short) distance between the paired species [6]. K can be studied as an abstract mathematical object called "pairing product" [7]. Since sufficiently stable pairs are worth considering, the study focuses on attractive-type interactions, i.e. $a < 0$. The present model is therefore not allowed for the treatment of electrostatic interactions: this state of affairs has been already discussed [3]. Under this condition, the following properties have not been disproved by any particular calculation performed so far:

(a) $K \geq 1$ (homopairing is favoured).

(b) $K = 1$ (homo- and hetero-pairing are equal) if and only if \mathbf{u}_1 and \mathbf{u}_2 are chemically equivalent, i.e. if $\mathbf{u}_1 = g_0\mathbf{u}_2$ for some skeletal symmetry operation g_0 .

If the symmetry of the skeleton acts by only one (e.g. C_1) or two (e.g. C_2) permutations of the ligand sites, these properties have been proven regardless of the nature (real number or vector) and the values of the ligand parameters. They have been discussed for carbene dimerization equilibria, equilibrating Diels–Alder reactions and equilibrating cyclopropanation reactions, where the pairing stereogenicity corresponds to a cis/trans isomerism [3].

The model also addresses the following question: if \mathbf{u} and \mathbf{v} are enantiomers ($\mathbf{v} = \sigma\mathbf{u}$, where σ is a mirror, an inversion or an improper rotation), is the homo-chiral association (RR) or (SS) more stable than the heterochiral one (RS)? In this context, properties (a) and (b) have been proven to be satisfied by $K(\mathbf{u}, \sigma\mathbf{u})$ for several skeletal symmetries [8]. For 3-D skeletons with a minimum number of sites positioned according to a C_i , C_s , C_{2h} , C_{2v} , C_{3v} , D_{2h} , S_4 , D_{3h} , or T_d symmetry, homo-chiral pairings are found to be more favoured than the hetero-chiral ones, regardless of the nature (real number or vector) of the ligand parameter. The same statement is true for real ligand parameters if the skeletal symmetry is C_{4v} or D_{4h} . Moreover, these pairings occur in equal amounts only if the molecule is achiral.

Setting the exponential coefficient to $p = -a/r^q kT > 0$, the scalar product form of the interaction energy leads to an equivalent expression of K (by definition: $K_p = K^{1/p}$):

$$K = K_p^p = \frac{\left(\sum_{g \in G} \exp \left[-\frac{p}{2} \|g\mathbf{u}_1 - \mathbf{u}_1\|^2 \right] \right) \left(\sum_{g \in G} \exp \left[-\frac{p}{2} \|g\mathbf{u}_2 - \mathbf{u}_2\|^2 \right] \right)}{\left(\sum_{g \in G} \exp \left[-\frac{p}{2} \|g\mathbf{u}_1 - \mathbf{u}_2\|^2 \right] \right)^2}.$$

If G is a compact infinite group, the sum symbol $\sum_{g \in G} / |G|$ is replaced by the integral symbol " $\int_G dg$ " (e.g. if $G = C_\infty$, $\int_G dg = \frac{1}{2\pi} \int_0^{2\pi} d\alpha$) [9]. This expression is naturally extended to regular metric spaces by changing the Euclidean norm $\|g\mathbf{u}_i - \mathbf{u}_j\|$ for more general distances $d(g\mathbf{u}_i, \mathbf{u}_j)$. Several general properties remain

valid but no theorem apraising the sign (property (a)) and the vanishing conditions (property (b)) of $K - 1$ is proven for non-Euclidean distances. The properties (a) and (b) would imply that $K(\mathbf{u}, \mathbf{v})$ is a consistent dissimilarity measure of \mathbf{u} and \mathbf{v} , and that $K(\mathbf{u}, \sigma\mathbf{u})$ is a consistent chirality measure of \mathbf{u} . The quantification of molecular dissimilarity is attracting a great deal of interest [10]: in this prospect, even if $K(\mathbf{u}, \mathbf{v})$ itself does not always fulfill the requirements (a) and (b), one seeks for a general $K(\mathbf{u}, \mathbf{v})$ -derived dissimilarity measure of \mathbf{u} and \mathbf{v} . And indeed, any “thermodynamical constant” $K = K_p^p$ could be associated to a real-valued positive function $D_p(\mathbf{u}, \mathbf{v})$ mapping algebraic models for pairs of skeletal analogs. An equation involving K_p^p has been devised, endowing the solution D_p with both a strong mathematical consistency and a thermochemical interpretation [11]. At the very outset, D_p should satisfy properties of a completely G -invariant distance, namely:

(i) *G-invariance requirements*: D_p is completely G -invariant, i.e. the definition of D_p from K_p must preserve the complete G -invariance of K_p . Especially, for any operation g : $D_p(g\mathbf{u}, \mathbf{u}) = 0$.

(ii) *Extension requirements*:

(a) D_p is an extension of the distance d : if \mathbf{u}_0 is invariant to all the operations of G , then: $D_p(\mathbf{u}, \mathbf{u}_0) = d(\mathbf{u}, \mathbf{u}_0)$.

(b) When $p \rightarrow 0$, D_p tends to the explicit completely G -invariant distance D_0 :

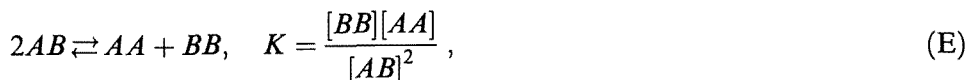
$$D_0(\mathbf{u}, \mathbf{v}) = \frac{1}{\int_G \frac{dg}{d(g\mathbf{u}, \mathbf{v})}} \quad (\text{“harmonic distance”}).$$

(c) When $p \rightarrow \infty$, D_p tends to the standard explicit completely G -invariant distance D_∞ :

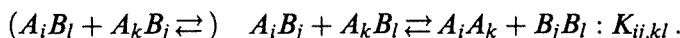
$$D_\infty(\mathbf{u}, \mathbf{v}) = \text{Inf}_{g \in G, h \in G} d(g\mathbf{u}, h\mathbf{v}) \quad (\text{“uniform distance”}).$$

(iii) *Two out of three distance properties on E/G [12].*

(iv) *Eventually, the triangular inequality*: $\forall(\mathbf{u}, \mathbf{v}, \mathbf{w}) \in E^3$, $D_p(\mathbf{u}, \mathbf{w}) \leq D_p(\mathbf{u}, \mathbf{v}) + D_p(\mathbf{v}, \mathbf{w})$. Formally, the design of (\mathbb{E}) is also dictated by the concern of a thermochemical interpretation. Considering the equilibrium (E),



each observable species, say AB , is a set of molecular states $\{A_i, B_j\}$. Now, the equilibrium is formally regarded as a set of “elemental equilibria” between states of the involved paired species:



The search for a quantity D allowing the equilibrium (E) to be defined as a “superposition” of these elemental equilibria, is undertaken from the equation

$$K = \exp\left[-\frac{\Delta G^0}{RT}\right] = \frac{1}{N} \sum_{i,j,k,l} [K_{ij,kl}]^{D^2/a_{ij,kl}} \quad (N: \text{normalizing factor})$$

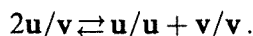
$$= \frac{1}{N} \sum_{i,j,k,l} \exp\left[-\frac{1}{kT} \{E(A_i A_k) + E(B_j B_l) - E(A_i B_j) - E(A_k B_l)\} \frac{D^2}{a_{ij,kl}}\right].$$

The $K_{ij,kl}$'s express the relative abundances of $A_i A_k$, $B_j B_l$, $A_i B_j$, and $A_k B_l$ in the whole equilibrium (E) characterized by D^2 . D^2 is an “energy” defined with respect to the lowest energy state on one side of the equilibrium, say the right-hand side $AA + BB$, and takes into account the competitive cross-pairing process: $A_i B_l + A_k B_j \rightleftharpoons A_i B_j + A_k B_l$. Coefficients $a_{ij,kl}$ are thus required to balance the $K_{ij,kl}$'s in K . They are defined by

$$a_{ij,kl} = (\alpha_{il} \alpha_{kj})^{1/2}$$

with $\alpha_{il} = E_0(AA) + E_0(BB) - 2E(A_i B_l)$ and $\alpha_{kj} = E_0(AA) + E_0(BB) - 2E(A_k B_j)$, where $E_0(AA)$ and $E_0(BB)$ denote the ground states of the homo-pairs.

Although D is related to the standard free energy of the equilibrium, no clear macroscopical interpretation is claimed: D^2 is a “mean cohesion energy”, averaged over molecular states, of a species AB with respect to homo-pairing products AA and BB . These definitions are applied to pairing equilibria of skeletal analogs represented by vectors \mathbf{u} and \mathbf{v} :



The states $(\mathbf{u}/\mathbf{v})_{ij}$, $(\mathbf{u}/\mathbf{v})_{kl}$, $(\mathbf{u}/\mathbf{u})_{ik}$, $(\mathbf{v}/\mathbf{v})_{jl}$ correspond to “stereoisomers” of observable pairs $AB = \mathbf{u}/\mathbf{v}$, $AA = \mathbf{u}/\mathbf{u}$, $BB = \mathbf{v}/\mathbf{v}$. Assuming that the energy of paired species is a function of ligand parameter vectors, the thermodynamic equation of D takes an algebraic form. A theorem has been stated for Euclidean spaces of ligand parameters [11], and the formulation is hereby generalized for any metric space.

THEOREM

Let G be a finite or compact group acting on a metric space (E, d) and preserving the distance ($\forall (\mathbf{u}, \mathbf{v}) \in E^2, \forall g \in G, d(g\mathbf{u}, g\mathbf{v}) = d(\mathbf{u}, \mathbf{v})$). For $p > 0$, let K_p be a discriminating pairing product ($K_p \geq 1$ and $K_p(\mathbf{u}, \mathbf{v}) = 1$ if and only if $\mathbf{v} = g\mathbf{u}$ for some operation g of G). Consider the equation of an unknown function D_p : $E \times E \rightarrow \mathbb{R}_+$:

$$\Phi_{\mathbf{u},\mathbf{v}}(D_p(\mathbf{u}, \mathbf{v})) = [K_p(\mathbf{u}, \mathbf{v})]^p \quad (\mathbb{E})$$

with

$$K_p^p(\mathbf{u}, \mathbf{v}) = \frac{\int_G \exp\left[-\frac{p}{2}d^2(g\mathbf{u}, \mathbf{u})\right] dg \int_G \exp\left[-\frac{p}{2}d^2(g\mathbf{v}, \mathbf{v})\right] dg}{\left(\int_G \exp\left[-\frac{p}{2}d^2(g\mathbf{u}, \mathbf{v})\right] dg\right)^2},$$

$$\Phi_{\mathbf{u}, \mathbf{v}}(x) = \int \int \int_{G^3} \exp\left[p \frac{C_{g,h,k}(\mathbf{u}, \mathbf{v})}{C_m^{f_{uv}(p)}(\mathbf{u}, \mathbf{v})} x^2\right] dg dh dk,$$

where

- $C_{g,h,k}(\mathbf{u}, \mathbf{v}) = \frac{d^2(g\mathbf{u}, \mathbf{v}) + d^2(k\mathbf{u}, h\mathbf{v}) - d^2(g\mathbf{u}, k\mathbf{u}) - d^2(\mathbf{v}, h\mathbf{v})}{2d(g\mathbf{u}, h\mathbf{v}) \cdot d(k\mathbf{u}, \mathbf{v})}$,
- $C_m(\mathbf{u}, \mathbf{v}) = \text{Max}\{C_{g,h,k}(\mathbf{u}, \mathbf{v}); (g, h, k) \in G^3\} (\geq 1)$,
- $f_{uv}(p)$ is some “regular” function eventually depending on (\mathbf{u}, \mathbf{v}) satisfying $f_{uv}(0) = 0$ and $f_{uv}(\infty) = 1$ and such that $p \rightarrow C_m^{f_{uv}(p)}$ is continuous.

Then, (\mathbb{E}) has a single solution D_p which fulfills the aforementioned requirements, except, perhaps, the triangular inequality (iv). If (E, d) is an Euclidean vector space, then

$$C_{g,h,k}(\mathbf{u}, \mathbf{v}) = \frac{(g\mathbf{u} - h\mathbf{v} \mid k\mathbf{u} - \mathbf{v})}{\|g\mathbf{u} - h\mathbf{v}\| \cdot \|k\mathbf{u} - \mathbf{v}\|} = \cos(g\mathbf{u} - h\mathbf{v}, k\mathbf{u} - \mathbf{v}) \leq C_m(\mathbf{u}, \mathbf{v}) = 1.$$

In this case, the definition of D_p does not require the determination of $f_{uv}(p)$. The occurrence of f_{uv} might not be necessary in a simpler extension of the Euclidean theorem: but if $f_{uv}(p)$ is equated to zero or to one, then either requirement (iic) or requirement (iib) is no longer ensured (see comment in Appendix).

It is henceforth emphasized that the discriminating condition on K_p is sufficient but might not be necessary. Beyond the axiomatic chemical interpretation, the Euclidean version of the theorem has been the starting point of abstract speculations constituting a so-called “chemical algebra” [7,11,13,16]. Although eq. (\mathbb{E}) has been actually devised from physico-chemical considerations, its resolution for any value of p is also challenging from a pure mathematical point of view. In simple cases only, the explicit solution could be obtained, allowing for a computational verification of the triangular inequality. At the very outset, the triangular inequality requirement (iv) may appear somewhat excessive and not directly necessary for the purpose of quantification of dissimilarity. However, it enhances the mathematical interest of the equation. In addition, this pure mathematical difficulty is worth mentioning for it suggests a differential (or local) resolution: a new metric $d\sigma^2 = D_p^2(\mathbf{u}, \mathbf{u} + d\mathbf{u})$ is thereby defined between infinitely close skeletal analogs. The integral distance associated to $d\sigma^2$ is different from D_p , but it surely satisfies the triangular inequality: this naturally raises the question for D_p itself. An integral

of $d\sigma$ along a “shortest transformation pathway” is interpreted as a “kinetic distance”, whereas D_p itself is termed as the “thermodynamic distance” [13].

The metric space (E, d) is a homogeneous set of models for a molecular property described by a ligand parameter, and $D_p(\mathbf{u}, \mathbf{v})$ quantifies the “dissimilarity” between two models \mathbf{u} and \mathbf{v} depending on the “temperature”, that is, more exactly, on the exponential coefficient p . In connection with recent investigations on the similarity of molecular shapes [14], the theorem is now applied when the described molecular property is the geometry, in particular when the molecules are enantiomers ($\mathbf{v} = \sigma\mathbf{u}$).

3. Geometrical version of eq. (E)

The bounded figures in the natural space $E_n \approx \mathbb{R}^n$ constitute a metric space for the Hausdorff distance d_H : $(E, d) = (\mathcal{B}(\mathbb{R}^n), d_H)$. In what follows, the vector notations $\mathbf{u}, \mathbf{v}, \dots$ are changed to more standard notations in regular metric spaces. If \mathbf{F} and \mathbf{F}' are two figures in \mathbb{R}^n , and if $\delta(M, \mathbf{F})$ denotes the distance from a point M to \mathbf{F} :

$$d_H(\mathbf{F}, \mathbf{F}') = \text{Max}\{\text{Sup}_{M \in \mathbf{F}} \delta(M, \mathbf{F}'); \text{Sup}_{M' \in \mathbf{F}'} \delta(M', \mathbf{F})\}.$$

If \mathbb{R}^n is considered as a vector space (i.e. if the figures are bound to have coincident centroids), the rotation group of \mathbb{R}^n or the orthogonal group (rotations + improper rotations) of \mathbb{R}^n (G) naturally acts on $\mathcal{B}(\mathbb{R}^n)$ and preserves the Hausdorff distance. Whatever is the sign of $K_p - 1$, eq. (E) can be written down: its eventual solution $D_p(\mathbf{F}, \mathbf{F}')$ would quantify the dissimilarity between \mathbf{F} and \mathbf{F}' . The geometrical version of the theorem can be formally interpreted in terms of a “skeleton-ligand parameter” model: a figure \mathbf{F} is fully defined by an infinite “skeleton”, the whole space \mathbb{R}^n , and by a “ligand parameter” λ which is evaluated at each point (or skeletal site) M by: $\lambda_M = 1$ as soon as M belongs to \mathbf{F} , and $\lambda_M = 0$ otherwise (λ is the membership function of \mathbf{F}). This formulation exhibits a further analogy with the treatment of molecules, namely: G still acts by permutation of skeletal sites in \mathbf{F} [15]. It is to be dwelt on the fact that G is not the symmetry group of the figure in the usual sense, but it is the symmetry group of its skeleton: the notation \mathbf{F} rigorously refers to the figure in the usual sense (e.g. a set of vertices) *plus* its support (the skeleton \mathbb{R}^n).

For $p = 0$ or ∞ , the solution of (E) is a *distance* between figures regardless of their relative orientation. The harmonic distance D_0 and the uniform distance D_∞ quantify differences in both shape and size: they are *invariant* to rotations but not to *similarities (homothets)*. Therefore, they do not merely address symmetry differences. However, inside a family of figures having the same size (i.e. the same diameter $\Delta(\mathbf{F})$), D_p summarizes the dissimilarity in symmetry. Nonetheless, since the symmetry of \mathbf{F} is characterized by all the symmetry operations remaining in \mathbf{F} , namely a point-group, a *complete* continuous description of symmetry cannot be

reduced to a single number (but it can be *summarized* by one number). If g is a rotation of \mathbb{R}^n ,

$$\mu_{\mathbf{F}}(g) = \exp \left[-\frac{d_H(g\mathbf{F}, \mathbf{F})}{\sqrt{2}} \right]$$

is a measure of the remaining symmetry of type “ g ” in \mathbf{F} [35]. $\mu_{\mathbf{F}}$ can be regarded as a membership function of a “fuzzy subgroup” [16a], where the fuzzy subgroup structure is a continuous extension of the abstract group structure without reference to any geometrical representation, and is connected with Mezey’s and Maruani’s concept of “syntopy groups” [16b].

Chirality is a non-existence symmetry property concerning a single type of operation ($i, \sigma, \mathbf{S}_\alpha$), and, according to chemists’ experience, the “chirality level” of a chemical species is empirically estimated by the single value of its specific optical rotation: it is therefore intuitive that “continuous chirality” is well measured by one number. The ability of D_p to quantify chirality is now examined.

4. Temperature-dependent chirality measures

Numerous efforts have focussed on the quantification of chirality [17]. Mislow and co-workers [18,19], gave a general definition for a consistent “chirality measure”: it is a real-valued function χ mapping a set of objects (molecule models), satisfying:

- (I) χ is continuous.
- (II) $\chi = 0$ if and *only if* the object is achiral.
- (III) χ is similarity-invariant, i.e. χ only depends on the *shape* of the objects.

This can be refined by adding one requirement, which is satisfied by the Hausdorff chirality measures [18]:

- (IV) χ is derived from a *distance* between enantiomorphic objects [20], and thus is a particular case of a more general dissimilarity measure between oriented objects. This requirement meets the mathematical requirement of the triangular inequality (iv) which was *a priori* introduced as an open perspective of the theorem. This requirement is also fulfilled by Avnir’s Continuous Chirality Measure [17b].

The statement that “to achieve chirality, it is not sufficient to assign different labels to parts of an achiral object” [21] is opposite to the definition of a “chirality index” χ of a molecular representation by a ligand parameter function mapping an achiral skeleton [8]. This discrepancy has to be discussed. Chirality is not understood as a universal property: it depends on the space where the object is allowed to move. Reviewing a well-known example [22], a non-isocetes triangle (“scalene”) is chiral in the plane and is achiral in the 3-D space. In a broader acceptance, abstract “chirality” is ultimately defined with respect to a group H acting on a metric space E and with respect to a subgroup G of index 2: a member of E (“object”) is said to be

“chiral” if any operation of $H-G$ transforms the object differently from all the operations of G . If the object is a figure of \mathbb{R}^2 or \mathbb{R}^3 (i.e. a member of $\mathcal{B}(\mathbb{R}^n)$), and if the natural representation of the displacement group in \mathbb{R}^n (rotations + translations) is considered, this definition meets the classical acceptance which may be called “geometric chirality”. However, in order to keep a link with current molecular representations by atoms or bonds assigned to the sites of a skeleton, the extended definition is worth retaining.

The constant of a pairing equilibrium between enantiomorphic figures is defined by assuming that the interaction energy is of attractive-type and varies as the squared Hausdorff distance between the figures [18,19,23]. The relative position and orientation of a figure F with respect to its enantiomorph σF is characterized by a displacement g :

$$\epsilon(gF/\sigma F) = -ad_H^2(gF, \sigma F)/2r^a \quad (a < 0).$$

It should be stressed that d_H does not represent the spatial distance between the centroids to the objects. The statement that ϵ varies as the square of d_H follows heuristically from the third hypothesis (scalar product form of the ligand interactions) which was recognized to be fruitful in the algebraic treatment of pairing equilibria [3]. Thus, the expression for ϵ results from the current mathematical process consisting in extending a formulation by generalization of the hypotheses: the space of geometrical figures is endowed with the Hausdorff distance in the same way as the space of ligand parameter vectors was endowed with the Euclidean distance (or, equivalently, with the corresponding scalar product). The early interpretation of the process (molecular interaction), is now more concealed, but it supports the ability of the model to treat the related problem of the measure of geometrical chirality.

In Rassat’s and Mislow’s approach, g would represent any rotation-translation in \mathbb{R}^n [18,23]. However, the translation group is not compact and is not endowed with a Haar measure allowing for the definition of the sums “ $\int_G \dots dg$ ” occurring in eq. (E) [9]. Therefore, the possible displacements have to be restricted to the *compact* group of rotations in \mathbb{R}^n : the common center of the rotations in \mathbb{R}^2 (or the intersection of their axes in \mathbb{R}^3) and the centroid of σF coincide with the centroid of F .

The theorem claims that D_p is a completely G -invariant distance in the borderline cases $p = \infty$ and $p = 0$, even for non-Euclidean distances such as d_H . This feature is illustrated by the search for the most chiral triangle at the corresponding “extremal temperatures”. The ability of (E) to produce a D_p -derived chirality measure at “finite temperature” ($0 < p < \infty$) will then be shortly discussed. Two-dimensional chirality has long been propounded as a simple illustration of three-dimensional chirality [22]. This concept directly addresses to two-dimensional chemistry [24], but also to three-dimensional stereochemistry [25], in particular to the prochirality notion [26]. Another possible application will be suggested in the last section. The extended theorem is now used to propose and discuss the notion of “temperature-dependent chirality measure” in two-dimensional space, where calculations

are much easier than in three-dimensional space. Following a well-admitted process, the chirality of simplexes in \mathbb{R}^2 (triangles T) is first investigated [17b,18]. All the principles which will be set out remain valid for simplexes in \mathbb{R}^3 (tetrahedrons) [27]. The term “simplex of \mathbb{R}^2 ” would have been actually more precise than the term “triangle”: the triangles being now considered are reduced to the subdimensional sets of their vertex ($d(T) = 0$) [28]. The same principles could be applied to triangles defined by subdimensional sets of three edges ($d(T) = 1$) or by equidimensional triangular surfaces ($d(T) = 3$), and the resulting values of $D_p(T, \sigma T)$ would be different for each definition of T and their enantiomorphs σT .

4.1. MOST “UNIFORMLY CHIRAL” TRIANGLE ($T = 0 K, p = \infty$)

Let T denote a triangle (A, B, C), and σT its enantiomorph. The calculation of $D_\infty(T, \sigma T)$ does not involve $d_H(gT, T)$, but only $d_H(gT, \sigma T)$ where g stands for rotations about the centroid O of T and where σ is the reflection across the line OA (fig. 2):

$$D_\infty(T, \sigma T) = \text{Inf}_{g \in G} d_H(gT, \sigma T).$$

This quantity does not have to depend on the initial orientation of T (D_∞ is completely G -invariant). However, a consistent geometrical chirality measure also requires to be similarity-invariant [18]: it must not depend on the size, but on the shape of the figure only. Therefore, T has to be normalized by a homothet of center O to be a similar triangle with a unit diameter. In other words, the uniform chirality measure of a triangle T is defined by

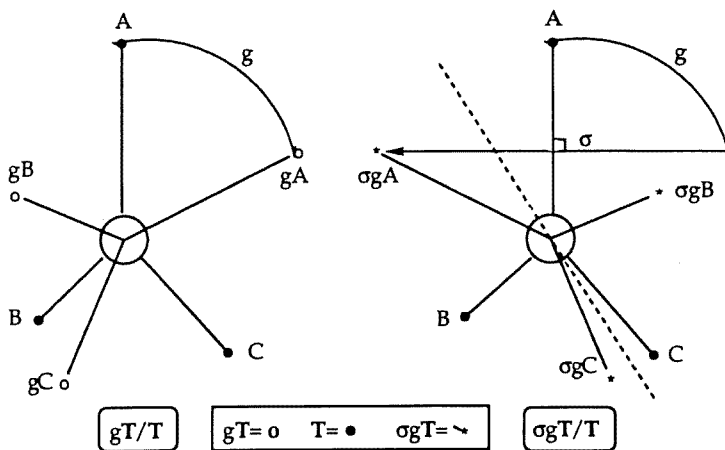


Fig. 2. The “interaction entropy” of figures with coincident centroids is generated by their relative orientation in the plane E^2 , characterized by a rotation g . This is exemplified for two identical triangles (left) and for two enantiomorphic ones which interconvert through a symmetry axis σ (right).

$$c_{\infty}(\mathbf{T}) = \frac{D_{\infty}(\mathbf{T}, \sigma\mathbf{T})}{\Delta(\mathbf{T})},$$

where $\Delta(\mathbf{T})$ is the length of the greatest edge of \mathbf{T} .

\mathbb{R}^2 is identified with the complex plane and a set of representative triangles $\mathbf{T} = (A, B, C)$ with a centroid O is produced by setting $r_A = OA = 1$ and by varying only two parameters:

- $r_b = OB$ from 0 to 1.
- $\theta_B = A-O-B$ angle, from 0 to 120° (or more precisely to $\cos^{-1}(-r_B/2) \geq 90^\circ$).

The method is detailed in Appendix. The 3-D plot of $c_{\infty}(\mathbf{T})$ against θ_B and r_B (fig. 3) displays a supremum $c_{\infty}(\mathbf{T}_{\infty})_{\max} = 0.231$ attained for the most chiral triangle for this measure, namely: $\theta_B = 52^\circ$, $r_B = 0.46$ (the corresponding diameter being $\Delta = 2.31(2)$). In terms of internal angles, the shape of the most chiral triangle, denoted \mathbf{T}_{∞} , is defined by $\alpha \approx 35.7^\circ$, $\beta \approx 128.6^\circ$, $\gamma \approx 15.7^\circ$ (fig. 5). The smallest Hausdorff distance (namely, 0.534) is attained for a rotation $g = 175^\circ$. Nevertheless, $g = 31^\circ$ and $g = 52^\circ$ correspond also to very small Hausdorff distances (0.547 and 0.540 respectively): as for Mislow's measure f (or H) involving translations, three configurations produce about the same optimal overlap (within the limits of calculation errors) [18]. The six-point figure defined by any union of enantiomorphic triangles with coincident centroids is always achiral ($g\sigma$ is a symmetry axis just as σ itself): this is a trivial version of Mislow's conjecture claiming that the

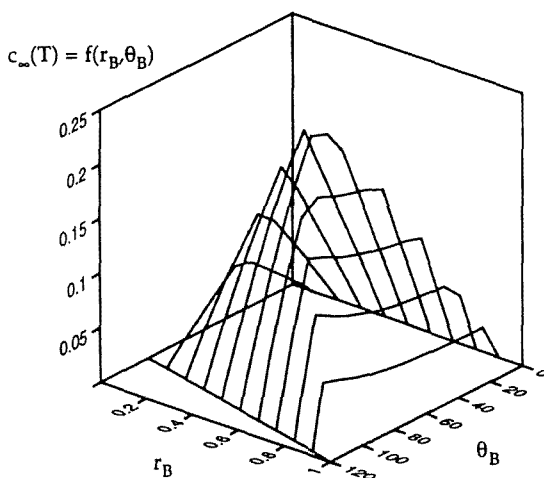


Fig. 3. Variation of the uniform chirality measure $c_{\infty}(\mathbf{T})$ ("0 K"). The variable parameters defining a complete set of representative triangles (A, B, C) are: $r_A = 1$, $0 \leq r_B \leq 1$, $0 \leq \theta_B \leq \cos^{-1}(-r_B/2) \leq 120^\circ$. The bottom borders in the surface (r_B, θ_B) correspond to achiral triangles.

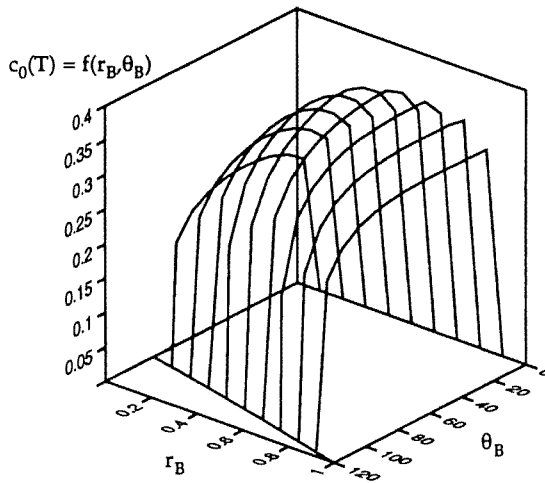


Fig. 4. Variation of the harmonic chirality measure $c_0(\mathbf{T})$ (“infinite temperature”). The variable parameters defining a complete set of representative triangles (A, B, C) are: $r_A = 1, 0 \leq r_B \leq 1, 0 \leq \theta_B \leq \cos^{-1}(-r_B/2) \leq 120^\circ$. The bottom borders in the surface (r_B, θ_B) correspond to achiral triangles.

union of an object and its mirror image is achiral under conditions of optimal overlap, even if the freedom degree of translations is allowed [18,19b].

Let us remind that considering translations of \mathbf{T} with respect to $\sigma\mathbf{T}$, the “most chiral triangle is defined by $\alpha \approx 21.5^\circ, \beta \approx 114.3^\circ, \gamma \approx 44.2^\circ$ and the corresponding Hausdorff measure equals $f = 0.196$ [18]: these values are quite close to those obtained by considering only rotations. However, $0.196 < c_\infty(\mathbf{T}_\infty)_{\max} = 0.231$, as expected. The measure involving translations is naturally finer than c_∞ (two triangles having the same non-zero c_∞ value may be chirally ranked by different f values). Nonetheless, c_∞ is completely consistent as a chirality measure and is more rapidly calculated: the sole variable “rotation angle” has to be swept, whereas f requires the sweep of three independent variables, one for rotations and two for translations.

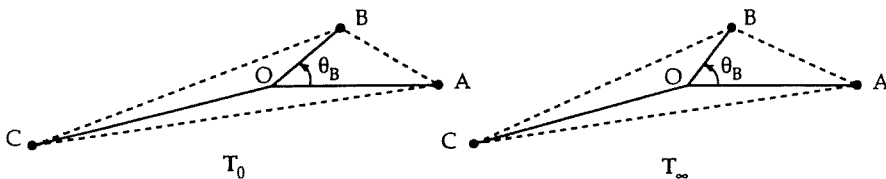


Fig. 5. The most harmonically chiral triangles \mathbf{T}_0 ($\infty K: r_A = OA = 1, r_B = 0.53, \theta_B = 39^\circ; c_0(\mathbf{T}_0) = 0.392$), and the most uniformly chiral triangle \mathbf{T}_∞ ($0 K: r_A = OA = 1, r_B = 0.46, \theta_B = 52^\circ; c_\infty(\mathbf{T}_\infty) = 0.231$).

4.2. MOST "HARMONICALLY CHIRAL" TRIANGLE ($T = \infty K, p = 0$)

Again, the calculation of $D_0(\mathbf{T}, \sigma\mathbf{T})$ does not involve the values $d_H(g\mathbf{T}, \mathbf{T})$, but only the values $d_H(g\mathbf{T}, \sigma\mathbf{T})$. Indeed,

$$D_0(\mathbf{T}, \sigma\mathbf{T}) = \frac{1}{\int_G \frac{dg}{d_H(g\mathbf{T}, \sigma\mathbf{T})}}.$$

As for $p = \infty$, the chirality measure "at infinite temperature" must be not only rotation-invariant, but also similarity-invariant. The harmonic chirality measure of \mathbf{T} is therefore defined by

$$c_0(\mathbf{T}) = \frac{D_0(\mathbf{T}, \sigma\mathbf{T})}{\Delta(\mathbf{T})}.$$

The 3-D plot of $c_0(\mathbf{T})$ against θ_B and r_B (fig. 4) displays a supremum $c_0(\mathbf{T}_0)_{\max} = 0.392$ attained for the most chiral triangle \mathbf{T}_0 at ∞K , namely: $\theta_B = 39^\circ$, $r_B = 0.53$ (with $\Delta = 2.43(4)$). In term of internal angles, the shape of \mathbf{T}_0 is defined by: $\alpha \approx 37.4^\circ$, $\beta \approx 130.4^\circ$, $\gamma \approx 12.2^\circ$, and resembles the shape of \mathbf{T}_∞ (fig. 5). The plot surface is now very flat except near achiral triangles where the slope is very steep: the level of chirality discrimination between two chiral triangles is much weaker "at high temperature" than at $0 K$ [29], and the level of chirality discrimination between a sufficiently chiral triangle and an achiral triangle becomes quasi-discrete ($\rightarrow 0-1$ measure), though still continuous [30]. No particular overlap is here distinguished: the overlap entropy was zero for $p = \infty$ ($0 K!$) and is maximum for $p = 0$ (∞K). since c_0 and c_∞ are continuously connected through eq. (E), they can be compared. It can be demonstrated that: $c_0 \geq c_\infty$: any triangle is "more chiral at infinite temperature than at $0 K$ ". However, such a statement comparing the chirality "at high and low temperatures" is not obvious for "finite temperatures".

4.3. CHIRALITY MEASURE "AT FINITE TEMPERATURE" ($0 < p < \infty$)

For finite p values, the search for D_p requires the calculation of $K_p(\mathbf{u}, \sigma\mathbf{u})$ first:

$$K_p^p(\mathbf{u}, \sigma\mathbf{u}) = \left(\frac{\int_G \exp\left[-\frac{p}{2}d^2(g\mathbf{u}, \mathbf{u})\right] dg}{\int_G \exp\left[-\frac{p}{2}d^2(g\mathbf{u}, \sigma\mathbf{u})\right] dg} \right)^2.$$

For many Euclidean distances, it has been shown that: $K_p(\mathbf{u}, \sigma\mathbf{u}) \geq 1$ and $K_p(\mathbf{u}, \sigma\mathbf{u}) = 1$ only if \mathbf{u} is achiral [8]. Therefore, $K_p(\mathbf{u}, \sigma\mathbf{u})$ (or the corresponding chirality index $1 \geq \chi_p(\mathbf{u}) \geq 0$) [31] is a chirality measure just as $D_p(\mathbf{u}, \sigma\mathbf{u})$. It is now established that for non-Euclidean distances such as d_H , $K_p(\mathbf{F}, \sigma\mathbf{F})$ may be smaller than 1 and may equal 1 even if \mathbf{F} is chiral! Thus, $K_p(\mathbf{F}, \sigma\mathbf{F})$ does not give a chirality

measure, but the corresponding distance extension $D_p(\mathbf{F}, \sigma\mathbf{F})$ is anticipated to give one (if the solution D_p is not univocal, one determination has to be selected by a requirement of continuity).

K_p is not similarity-invariant: if λ is a positive number, $\lambda\mathbf{T}$ is similar to \mathbf{T} , but $K_p(\lambda\mathbf{T}, \sigma\lambda\mathbf{T}) \neq K_p(\mathbf{T}, \sigma\mathbf{T})$. However: $d_H(g\lambda\mathbf{T}, \lambda\mathbf{T}) = \lambda d_H(g\mathbf{T}, \mathbf{T})$ and $d_H(g\lambda\mathbf{T}, \sigma\lambda\mathbf{T}) = \lambda d_H(g\mathbf{T}, \sigma\mathbf{T})$. Consequently: $K_p^{\lambda^2}(\lambda\mathbf{T}, \sigma\lambda\mathbf{T}) = K_p^{\lambda^2}(\mathbf{T}, \sigma\mathbf{T})$. Since λ may reach any positive value, the study of $K_p(\mathbf{T}, \sigma\mathbf{T})$ against \mathbf{T} can be restricted to, e.g., $p = 2$ [32]. Calculations have been performed, showing that inside the family of chiral non-similar representative triangles with $r_B \leq r_A = 1$, $K_2(\mathbf{T}, \sigma\mathbf{T})$ is always very close to 1 (slightly smaller), namely, $0.989 < K_2(\mathbf{T}, \sigma\mathbf{T}) \leq 1$ (Appendix B). Since $K_\infty(\mathbf{T}, \sigma\mathbf{T})$ is greater than 1 and is continuous, for any chiral triangle \mathbf{T} , there exists a critical p value, denoted $p(\mathbf{T})$, such that: $K_{p(\mathbf{T})}(\mathbf{T}, \sigma\mathbf{T}) = 1$, and for $p > p(\mathbf{T})$, $K_p(\mathbf{T}, \sigma\mathbf{T}) > 1$ [33]. A possible meaning for $p(\mathbf{T})$, and more generally for $p(\mathbf{F}, \mathbf{F}')$ (satisfying $K_{p(\mathbf{F}, \mathbf{F}')}(\mathbf{F}, \mathbf{F}') = 1$ even if the figures \mathbf{F} and \mathbf{F}' do not interconvert by any rotation) is commented on in Appendix. In fact, within calculation uncertainty, the range of values found for K_2 might be a deviation of a single value ($K_2(\mathbf{T}, \sigma\mathbf{T}) = 1$) and $p(\mathbf{T}) = 2$ would be identical inside the selected family of triangles.

For the most uniformly chiral triangle \mathbf{T}_∞ ($r_A = 1$, $r_B = 0.46$, $\theta = 52^\circ$), $K_p(\mathbf{T}_\infty, \sigma\mathbf{T}_\infty)$ has been calculated for $0 \leq p \leq \infty$ and the variation is depicted in fig. 6: when p decreases, $K_p(\mathbf{T}_\infty, \sigma\mathbf{T}_\infty)$ decreases from $1.330 \dots = e^{(0.534 \dots)^2}$ (indeed, $D_\infty(\mathbf{T}_\infty, \sigma\mathbf{T}_\infty) = 0.534 \dots$), reaches the value 1 for $p = p(\mathbf{T}_\infty) \approx 2$, and is smaller than 1 for $0 \leq p < 2$. The aspect of the plot for other triangles will be analogous to that given by \mathbf{T}_∞ .

Suppose $D_p(\mathbf{F}, \mathbf{F}')$ is the solution of (E). If $p \neq 0, \infty$, the ratio $D_p^2(\mathbf{F}, \mathbf{F}') / \Delta(\mathbf{F})\Delta(\mathbf{F}')$ is not similarity-invariant (but it is as soon as $p = 0$ or ∞). However, it is easily shown that

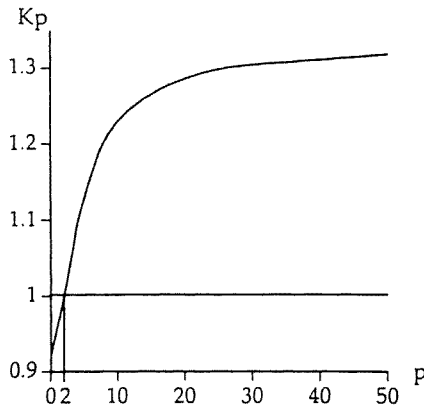


Fig. 6. Plot of $K_p(\mathbf{T}, \sigma\mathbf{T})$ vs. p for $\mathbf{T} = \mathbf{T}_\infty$, the most uniformly chiral triangle (0 K).

$$\forall \lambda > 0, D_p^2(\lambda F, \lambda F') = \lambda^2 D_{\lambda^2 p}^2(F, F') C_m^{f_{\lambda F, \lambda F'}(p) - f_{F, F'}(\lambda^2 p)},$$

where $C_m = C_m(F, F') = C_m(\lambda F, \lambda F')$ and $f_{F, F'}(p)$ are defined in the general formulation of the theorem. Since $f_{F, F'}(p)$ is not yet completely specified, it can be required that it fulfills the relationship: $f_{\lambda F, \lambda F'}(p) = f_{F, F'}(\lambda^2 p)$ (which is compatible with the conditions $f_{F, F'}(0) = 0$ and $f_{F, F'}(\infty) = 1$: see comment in Appendix for a determination of $f_{F, F'}(p)$). Then, replacing p by $p/\Delta(F)\Delta(F')$ and taking $F = T$, $F' = \sigma T$, one gets a similarity-invariant p -continuous extension of the chirality measures c_0 and c_∞ [34]:

$$c_p(T) = \frac{D_{p/\Delta^2(T)}(T, \sigma T)}{\Delta(T)}.$$

The study of K_p , D_p and c_p is obviously much more complicated for non-Euclidean distances than for Euclidean ones, but further investigations deserve to be undertaken.

5. Chemical interpretation of the Hausdorff interaction between geometrical figures

Many efforts have focussed on ranking molecular symmetry [35] or chirality [17,18]. The connection with the chemical challenge of ranking some type of reactivity inside a family of substrates is not evident. For example, are we allowed to anticipate that the asymmetric induction exerted by an asymmetric carbon exhibiting the “most chiral tetrahedral geometry” [19] is higher than that exerted by carbons with any other tetrahedral geometry? As it is, even if the chirality transfer is recognized to result from differential steric interactions only, the geometry to be considered is not known *a priori*: the representative points may be the centers of mass of the substituents, those of the first atoms of the substituents, or any others. Furthermore, the answer requires the knowledge of the form of the interaction energy dictating the outcome of the reaction. This section aims at connecting some of the principles governing the design of the geometrical chirality measures c_p with a model for chemical interactions.

A substituted ethane molecule exists as a distribution of “rotamers” (figs. 1 and 2). Suppose that the substituents at one carbon exert a phenomenological attractive or repulsive force on the substituent at the facing carbon. The general hypothesis of the algebra of stereogenic equilibria ($a < 0$), corresponds to the attractive situation which appears here very speculative but which is not impossible in principle (it could occur, for example, when all the substituents are sufficiently small and both donors and acceptors of hydrogen bonds: OH, SH, NHR, NHCOR, ...). Although the case of $a > 0$ has not afforded a general algebraic theorem on K_p , the corresponding repulsive situation (steric strain between facing substituents) can be considered at this stage.

The sp^3 geometry is more or less distorted: consider the triangle T_1 defined by the intersection of the three C^1-R_i bonds with a plane perpendicular to the C^1-C^2 axis. An analogous triangle T_2 is defined for the carbon C^2 . Let r be the distance between the two planes. The rotation of the C^1-C^2 bond brings about the rotation of T_2 with respect to T_1 . It is reasonable to propose that the T_1/gT_2 interaction energy could vary as the square distance

$$d^2(g) = \text{Max}\{\text{Sup}_{M_1 \in T_1} \delta^2(M_1, gT_2); \text{Sup}_{M_2 \in gT_2} \delta^2(M_2, T_1)\}.$$

Since the plane are parallel and r Å apart, the Pythagore theorem yields

$$d^2(g) = r^2 + \text{Max}\{\text{Sup}_{M_1 \in T_1} \delta^2(M'_1, gT_2); \text{Sup}_{M_2 \in gT_2} \delta^2(M'_2, T_1)\},$$

where M'_1 is the projection of M_1 on the plane of T_2 and M'_2 is the projection of M_2 on the plane of T_1 . Finally,

$$d^2(g) = r^2 + d_H^2(T_1, gT'_2),$$

where T'_2 is the projection of T_2 on the plane of T_1 (in what follows T'_i is equated to T_i). In accordance with the prerequisites of the algebra of stereogenic pairing equilibria, the energy of any rotamer characterized by g would be

$$\epsilon(g) = -a[r^2 + d_H^2(T_1, gT_2)]/2r^q$$

(here a can assume both positive and negative values). For example, T_2 may be identical to T_1 (*meso* molecule RS) or enantiomorphic to T_1 (*dl* molecule, RR or SS) (fig. 7). The rotational energy is a three-parameter function of g (a, r, q), eventually suitable for a molecular mechanics force field. Given a molecule $R_1R_2R_3C-CR_1R_2R_3$, a triangle $T = (A, B, C)$ is drawn, where A, B, C are the intersection points of the bond axes $C-R_1, C-R_2, C-R_3$ with some plane perpendicular to the C^1-C^2 axis. Suppose $T = T_\infty$ or $T = T_0$: the energy variation against the dihedral angle $\theta = g$ is depicted in fig. 8 for the *meso* and *dl* isomers. This "energy" represents geometrical information as the variation of Avnir's CCM in rotating ethane structures [17b].

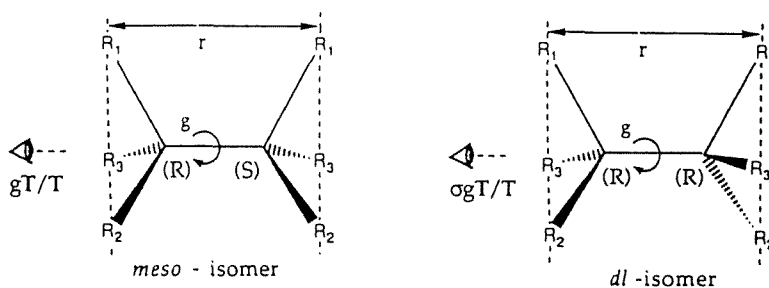


Fig. 7. Rotamers of a substituted ethane molecule are putatively weighted by mean of the Hausdorff interaction $E(gT, T') = -a[r^2 + d_H^2(gT, T')]/2r^q$ ($a < 0$ or > 0). This is exemplified for $T' = T$ (left) and $T' = \sigma T$ (right). The observed projection of the associated triangles is shown in fig. 2.

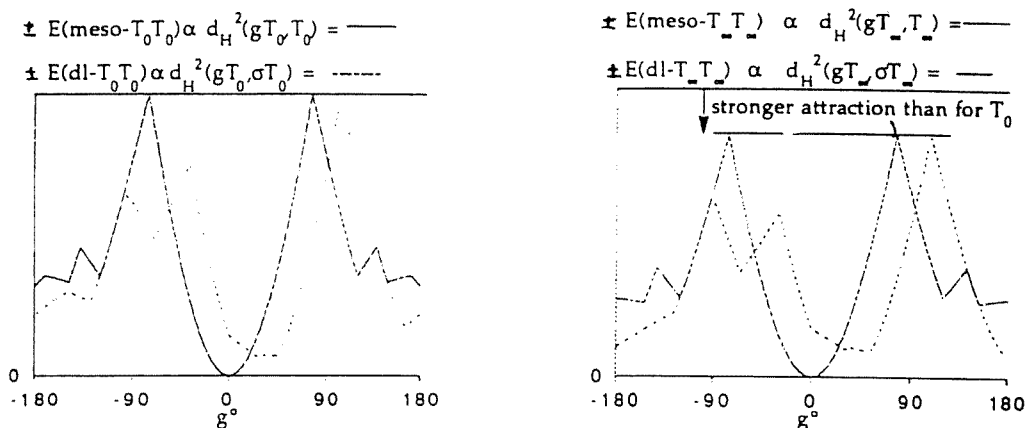


Fig. 8. Energy profiles of “rotamers” in symmetrically (*meso* or *dl*) substituted ethane molecules, where g° denotes the dihedral rotation angle in degrees. The geometrical model for the substituted methyl groups corresponds to the most harmonically chiral triangle (T_0C-CT_0 , left) or to the most uniformly chiral triangle ($T_\infty C-CT_\infty$, right). The vertical axis represents energy: either ϵ (attractive interaction, e.g. through hydrogen bonds), or $-\epsilon$ (repulsive interaction, e.g. through steric strains).

6. Remarks on fig. 1

In the preceding section, the term “rotamer” is used for convenience, but it usually refers to a kind of “isomer” and is strictly unfit with respect to the sense discussed by Eliel [36]: the different relative orientations of the two ends of the molecules are not independently observed and take place either as “states”, if they are *not* located near a minimum with an energy barrier greater than kT , or, otherwise, as (*proto*-)isomers [35] of one (*residual*) isomer [35] defined under given conditions of observation (*meso*, *dl*, ...). In fig. 1, a molecule is defined by a structure lying at a minimum of the global potential energy hypersurface and by a scatter of states surrounding the central structure by less than $\Delta E_M \text{ kcal mol}^{-1}$. ΔE_M represents the stability scale, required by Eliel, in relation to given conditions of observation or in relation to given criteria (e.g. $\Delta E_M = kT$) [35]. In the approach of fig. 1, the definition of ΔE_M and the temperature determine the flexibility and the variability of the molecule, and thus its entropy.

7. Summary and conclusion

In the beginning, D_p serves to describe dissimilarity in *energy*, and the definition has been extended to describe dissimilarity in *geometry* on the same basis. The common “skeleton” of geometrical figures is the whole space E_n , mapped by different “geometrical ligand parameters”: $\lambda_M = 1$ if M belongs to the figure and $\lambda_M = 0$ otherwise. Since a temperature parameter occurs in the “chemical version” of

eq. (E), the geometrical dissimilarity measure derived from the same equation is also modulated by a “temperature notion”. For example, the entropy vanishes at 0 K ($p = \infty$) and only one overlap of “interacting figures” (that giving the shortest Hausdorff distance) takes place in the uniform distance D_∞ . As an application, the equation suggests the definition of “temperature-dependent chirality measures”, and uniform (0 K), harmonic (∞ K) and intermediate ($0 < T < \infty$) planar chirality measures of triangles have been studied. The same process could be applied to the space chirality measure of tetrahedrons (basic geometry in organic chemistry [37]). Finally, dissimilarity measures based on repulsive-type interactions ($p < 0$) might be studied, and applications of the Hausdorff distance model might be sought in conformational analysis.

Appendix A

CALCULATION OF $d_H(gT, \sigma T)$

Let T be a triangle with a centroid O (see fig. 9). The sum of three angles $A-O-B$, $B-O-C$, $C-O-A$ is equal to 360° , and the smallest one, say $A-O-B$, is less than 120° . Let OA be greater than OB , and consider the O -homothetic triangle with $OA = r_A = 1$. Then, $OB = r_B \leq 1$. In the complex plane marked by the OA axis as the Ox axis, the polar coordinates of the triangle vertices define the affixes $z_A = r_A = 1$, $z_B = r_B e^{i\theta_B}$, $z_C = r_C e^{i\theta_C}$.

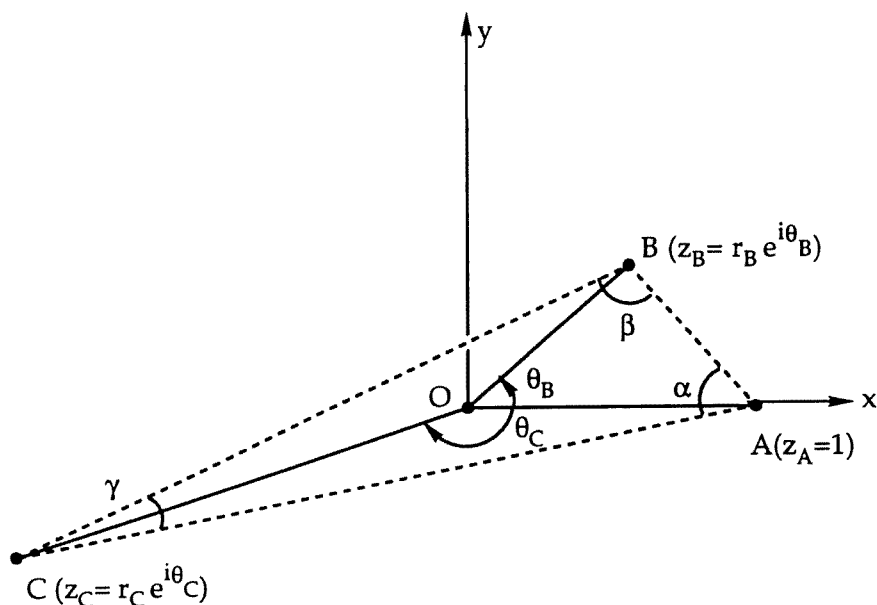


Fig. 9.

Since O is the centroid of (ABC) , z_C is completely determined by z_B , namely by

$$z_A + z_B + z_C = 0,$$

i.e.,

$$r_A + r_B e^{i\theta_B} + r_C e^{i\theta_C} = 0,$$

or

$$\begin{cases} r_C \cos \theta_C = -r_A - r_B \cos \theta_B, \\ r_C \sin \theta_C = -r_B \sin \theta_B, \end{cases}$$

i.e.,

$$\begin{cases} r_C^2 = r_A^2 + r_B^2 + 2r_A r_B \cos \theta_B, \\ \tan \theta_C = \frac{r_B \sin \theta_B}{r_A + r_B \cos \theta_B}. \end{cases}$$

Since $-90^\circ \leq \tan^{-1} \theta_C \leq 90^\circ$

$$\begin{cases} \text{if } \tan \theta_C > 0, \text{ then } \theta_C = 180 + \tan^{-1} \theta_C \\ \text{if } \tan \theta_C < 0, \text{ then } \theta_C = \tan^{-1} \theta_C \end{cases} \quad (\text{see scheme}).$$

The distance between two points M and N is given by

$$d(M, N) = |z_M - z_N| = \sqrt{r_M^2 + r_N^2 - 2r_M r_N \cos(\theta_M - \theta_N)}.$$

If g is a rotation acting on a point $M(z_M = r_M e^{i\theta_M})$, then gM is defined by

$$z_{gM} = r_M e^{i(\theta_M + g)}$$

and if σ denotes the reflexion across the OA axis, $g\sigma M$ is defined by:

$$z_{g\sigma M} = r_M e^{i(g - \theta_M)}.$$

Achiral triangles

For $r_A = 1$, the achiral triangles obtained by varying $r_B \leq 1$ and $0^\circ < \theta_B < 120^\circ$ correspond to $r_B = 0$ or 1 (whatever is θ_B) or to $r_C = r_A = 1$. The isocetes triangles are thus obtained from the equation $r_B = -2 \cos \theta_B$. When r_B runs from 0 to 1, the θ_B values of the corresponding achiral triangles are smaller than 120° (but close to 120°). (See table 1.) In order to calculate $D_p(\mathbf{T}, \sigma\mathbf{T})$ for any chiral triangle \mathbf{T} , a KaleidagraphTM macro-program has been written on the basis of the following columns and memories:

Variable

$c0 = g$ (from -180° to $+180^\circ$ with 1° increment).

Parameters

$m0 = \theta_C$ (from 0° to 120° with a 10° increment or locally finer).

$m1 = r_B$ (from 0 to 1.0 with a 0.1 increment or locally finer).

Table 1
Achiral triangles with $r_C = r_A = 1$.

r_B	$\theta_B = \cos^{-1}(-r_B/2)$	r_B	$\theta_B = \cos^{-1}(-r_B/2)$	r_B	$\theta_B = \cos^{-1}(-r_B/2)$
0	90°	0.4	101.54°	0.8	113.58°
0.1	92.87°	0.5	104.48°	0.9	116.74°
0.2	95.74°	0.6	107.46°	1.0	120°
0.3	98.63°	0.7	110.49°		

Calculation auxiliaries for $d_H(g\mathbf{T}, \sigma\mathbf{T})$

$m2 = \theta_C$, $m3 = r_C$, $c1 = d(g\sigma A, A)$, $c2 = d(g\sigma A, B)$, $c3 = d(g\sigma A, C)$, $c5 = d(g\sigma B, A)$, $c6 = d(g\sigma B, B)$, $c7 = d(g\sigma B, C)$, $c9 = d(g\sigma C, A)$, $c10 = d(g\sigma C, B)$, $c11 = d(g\sigma C, C)$, $c4 = (c1, c2, c3) = \delta(g\sigma A, \mathbf{T})$, $c8 = \text{Min}(c5, c6, c7) = \delta(g\sigma B, \mathbf{T})$, $c12 = \text{Min}(c9, c10, c11) = \delta(g\sigma C, \mathbf{T})$, $c13 = \text{Max}(c4, c8, c9) = \text{Sup}_{M \in \mathbf{T}} \delta(g\sigma M, \mathbf{T})$, $c14 = \text{Min}(c1, c5, c9)$, $c15 = \text{Min}(c2, c6, c10)$, $c16 = \text{Min}(c3, c7, c11)$, $c17 = \text{Max}(c14, c15, c16) = \text{Sup}_{M \in \mathbf{T}} \delta(M, g\sigma\mathbf{T})$, $c18 = c19 = \text{Max}(c13, c17) = d_H(g\mathbf{T}, \sigma\mathbf{T})$.

Using the shortcuts $c9 = c3$, $c5 = c2$, and $c10 = c7$, a macro-programme “**RS** ($r_A = 1$)” is thus written for computing $d_H(g\mathbf{T}, \sigma\mathbf{T})$ values on a 4Mo Macintosh micro-computer. The values of $D_\infty(\mathbf{T}, \sigma\mathbf{T})$ and $D_0(\mathbf{T}, \sigma\mathbf{T})$ are then deduced, and chirality measures c_p ($p = 0, \infty$) are obtained by dividing $D_p(\mathbf{T}, \sigma\mathbf{T})$ by the diameter of \mathbf{T} :

$$\Delta(\mathbf{T}) = \text{Max}\{d(A, B); d(B, C); d(C, A)\}.$$

Results are displayed in table 2.

Appendix B

CALCULATIONS OF $K_2(\mathbf{T}, \sigma\mathbf{T}) = (\int_G \exp[-d_H^2(g\mathbf{T}, \mathbf{T})] dg) / (\int_G \exp[-d_H^2(g\mathbf{T}, \sigma\mathbf{T})] dg)$

Both the $d_H(g\mathbf{T}, \mathbf{T})$ and $d_H(g\sigma\mathbf{T}, \mathbf{T})$ values are now needed. In order to compute $d_H(g\mathbf{T}, \mathbf{T})$ values, a macro-programma “**RR& SS** ($r_A = 1$)” has been written.

Calculation auxiliaries for $d_H(g\mathbf{T}, \mathbf{T})$

$m2 = \theta_C$, $m3 = r_C$, $c1 = d(gA, A)$, $c2 = d(gA, B)$, $c3 = d(gA, C)$, $c5 = d(gB, A)$, $c6 = d(gB, B)$, $c7 = d(gB, C)$, $c9 = d(gC, A)$, $c10 = d(gC, B)$, $c11 = d(gC, C)$, $c4 = \text{Min}(c1, c2, c3) = \delta(gA, \mathbf{T})$, $c8 = \text{Min}(c5, c6, c7) = \delta(gB, \mathbf{T})$, $c12 = \text{Min}(c9, c10, c11) = \delta(gC, \mathbf{T})$, $c13 = \text{Max}(c4, c8, c9) = \text{Sup}_{M \in \mathbf{T}} \delta(gM, \mathbf{T})$, $c14 = \text{Min}(c1, c5, c9)$, $c15 = \text{Min}(c2, c6, c10)$, $c16 = (c3, c7, c11)$, $c17 = \text{Max}(c14, c15, c16) = \text{Sup}_{M \in \mathbf{T}} \delta(M, g\mathbf{T})$, $c18 = \text{Max}(c13, c17) = d_H(g\mathbf{T}, \mathbf{T})$. ($c0$ is restricted to the interval $[0^\circ, 180^\circ]$). The integrals $\int_G \exp[-d_H^2(g\mathbf{T}, \mathbf{T})] dg$ and $\int_G \exp[-d_H^2(g\sigma\mathbf{T}, \mathbf{T})] dg$

Table 2

$\theta_B \backslash r_B$	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
$c_\infty(\mathbf{T}) = D_\infty(\mathbf{T}, \sigma\mathbf{T})/\Delta(\mathbf{T})$ vs. $\mathbf{T} = (\theta_B, r_B)$:											
0.00°	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
10.0°	0.00	0.0157	0.0269	0.0364	0.0436	0.0489	0.0524	0.0544	0.0562	0.0346	0.00
20.0°	0.00	0.0303	0.0534	0.0702	0.0842	0.0943	0.101	0.109	0.0723	0.0349	0.00
30.0°	0.00	0.0442	0.0792	0.106	0.126	0.142	0.154	0.114	0.0735	0.0355	0.00
40.0°	0.00	0.0579	0.104	0.140	0.167	0.189	0.161	0.116	0.0751	0.0364	0.00
50.0°	0.00	0.0616	0.117	0.168	0.209	0.213	0.165	0.120	0.0773	0.0375	0.00
60.0°	0.00	0.0487	0.0949	0.139	0.180	0.218	0.170	0.124	0.0801	0.0389	0.00
70.0°	0.00	0.0352	0.0721	0.108	0.144	0.180	0.177	0.129	0.0845	0.0412	0.00
80.0°	0.00	0.0206	0.0457	0.0739	0.105	0.135	0.165	0.137	0.0901	0.0442	0.00
90.0°	0.00	0.0052	0.0173	0.0362	0.0580	0.0842	0.113	0.143	0.0973	0.0480	0.00
100°	0.00	0.0123	0.0141	0.0071	0.0010	0.0283	0.0527	0.0792	0.106	0.0540	0.00
110°	0.00	0.0280	0.0446	0.0508	0.0476	0.0369	0.0203	0.0065	0.0306	0.0590	0.00
120°	0.00	0.0441	0.0760	0.0966	0.106	0.106	0.0991	0.0832	0.0607	0.0343	0.00
$c_0(\mathbf{T}) = D_0(\mathbf{T}, \sigma\mathbf{T})/\Delta(\mathbf{T})$ vs. $\mathbf{T} = (\theta_B, r_B)$:											
0.000°	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
10.00°	0.000	0.2037	0.2448	0.2752	0.3000	0.3212	0.3335	0.3413	0.3431	0.3123	0.000
20.00°	0.000	0.2215	0.2697	0.3060	0.3359	0.3605	0.3722	0.3710	0.3479	0.3133	0.000
30.00°	0.000	0.2313	0.2842	0.3241	0.3560	0.3800	0.3871	0.3710	0.3483	0.3143	0.000
40.00°	0.000	0.2363	0.2919	0.3336	0.3659	0.3881	0.3851	0.3700	0.3482	0.3145	0.000
50.00°	0.000	0.2372	0.2940	0.3362	0.3677	0.3872	0.3814	0.3680	0.3474	0.3145	0.000
60.00°	0.000	0.2341	0.2904	0.3320	0.3625	0.3784	0.3755	0.3643	0.3451	0.3129	0.000
70.00°	0.000	0.2263	0.2806	0.3210	0.3506	0.3658	0.3664	0.3580	0.3408	0.3103	0.000
80.00°	0.000	0.2115	0.2625	0.3013	0.3307	0.3474	0.3516	0.3473	0.3331	0.3047	0.000
90.00°	0.000	0.1779	0.2272	0.2666	0.2977	0.3186	0.3280	0.3290	0.3196	0.2950	0.000
92.87°	0.000	0.000									0.000
95.74°	0.000		0.000								0.000
98.63°	0.000			0.000							0.000
100.00°	0.000	0.1976	0.2193	0.2054	0.2192	0.2622	0.2848	0.2958	0.2950	0.2773	0.000
101.5°	0.000				0.000						0.000
104.5°	0.000					0.000					0.000
107.5°	0.000						0.000				0.000
110.0°	0.000	0.2211	0.2635	0.2841	0.2874	0.2727	0.2413	0.1953	0.2399	0.2413	0.000
110.5°	0.000							0.000			0.000
113.6°	0.000								0.000		0.000
116.7°	0.000									0.000	0.000
120.0°	0.000	0.2339	0.2876	0.3204	0.3337	0.3266	0.3095	0.2840	0.2495	0.2017	0.000

are calculated and divided by each other to afford $K_2(\mathbf{T}, \sigma\mathbf{T})$. Setting $I_{RR} = 360 \int_G \exp[-d_H^2(g\mathbf{T}, \mathbf{T})] dg$ and $I_{RS} = 360 \int_G \exp[-d_H^2(g\mathbf{T}, \sigma\mathbf{T})] dg$, the results are listed in tables 3–5. The exploration of negative p values will be discussed later.

Table 3

	10°	20°	30°	40°	50°	60°	70°	80°	90°	100°	110°	120°
$r_B = 0.1:$												
$I_{RR}/2$	166,69	166,47	166,31	166,20	166,14	166,13	166,16	166,23	166,34	166,27	166,15	166,06
I_{RS}	334,91	334,86	334,71	334,51	334,26	334,06	333,91	333,81	333,75	333,74	333,77	333,86
$r_B = 0.2:$												
$I_{RR}/2$	165,32	165,00	164,79	164,70	164,70	164,79	164,95	165,14	165,38	165,39	165,10	164,80
I_{RS}	331,52	331,52	331,41	331,27	331,10	331,03	331,01	331,02	331,02	330,99	330,91	330,77
$r_B = 0.3:$												
$I_{RR}/2$	163,81	163,47	163,33	163,36	163,54	163,83	164,20	164,61	165,02	165,31	164,90	164,37
I_{RS}	328,73	328,83	328,91	329,02	329,20	329,55	329,94	330,32	330,57	330,67	330,54	330,11
$r_B = 0.4:$												
$I_{RR}/2$	162,18	161,91	161,91	162,16	162,60	163,13	163,71	164,30	164,88	165,42	165,02	164,25
I_{RS}	325,58	325,83	326,21	326,71	327,46	328,39	329,34	330,14	330,68	330,90	330,79	330,32
$r_B = 0.5:$												
$I_{RR}/2$	160,46	160,33	160,56	161,07	161,67	162,33	163,04	163,77	164,57	165,49	165,58	164,74
I_{RS}	322,14	322,64	323,44	324,42	325,72	327,32	328,86	330,16	331,03	331,56	331,96	331,89
$r_B = 0.6:$												
$I_{RR}/2$	159,21	159,22	159,68	160,26	160,92	161,67	162,58	163,66	164,81	166,02	166,76	165,88
I_{RS}	319,86	320,62	321,60	323,01	324,70	326,56	328,53	330,51	332,11	333,21	333,88	333,81
$r_B = 0.7:$												
$I_{RR}/2$	158,40	158,64	159,14	159,75	160,54	161,47	162,53	163,72	165,01	166,38	167,81	167,03
I_{RS}	318,21	319,08	320,26	321,78	323,54	325,55	327,77	330,17	332,56	334,46	335,58	335,67
$r_B = 0.8:$												
$I_{RR}/2$	157,84	158,22	158,75	159,45	160,29	161,27	162,39	163,66	165,06	166,58	168,18	168,17
I_{RS}	316,80	317,70	318,93	320,43	322,21	324,29	326,65	329,29	332,15	334,98	336,99	337,40
$r_B = 0.9:$												
$I_{RR}/2$	157,46	157,86	158,43	159,14	160,00	161,01	162,18	163,50	164,98	166,61	168,35	169,24
I_{RS}	315,50	316,37	317,54	319,02	320,78	322,86	325,25	327,98	331,04	334,37	337,65	338,94

Appendix C

COMMENTS ON THE MEANING OF $p(T)$ AND $f_{uv}(p)$

The proof of the theorem has been given for Euclidean distances [11]. Although no difficulty is met by extending the proof to regular metric spaces, the extended formulation of the theorem has to be examined, in particular with regard to the determination of $f_{uv}(p)$. The following short discussion is obviously not complete and is just devoted to explore some possibilities.

Table 4

$K_I(\mathbf{T}, \sigma\mathbf{T}) = I_{RR}/I_{RS}$ vs. $\mathbf{T} = (\theta_B, r_B)$.

$\theta_B \backslash r_B$	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
0°	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
10	0.9954	0.9973	0.9966	0.9963	0.9962	0.9955	0.9956	0.9965	0.9982	1.0000
20	0.9943	0.9954	0.9943	0.9938	0.9938	0.9932	0.9943	0.9960	0.9980	1.0000
30	0.9937	0.9945	0.9932	0.9927	0.9929	0.9930	0.9938	0.9955	0.9978	1.0000
40	0.9937	0.9944	0.9930	0.9927	0.9930	0.9923	0.9930	0.9952	0.9977	1.0000
50	0.9941	0.9949	0.9936	0.9931	0.9927	0.9912	0.9924	0.9949	0.9976	1.0000
60	0.9946	0.9956	0.9943	0.9935	0.9919	0.9901	0.9920	0.9946	0.9974	1.0000
70	0.9952	0.9966	0.9953	0.9942	0.9915	0.9898	0.9918	0.9943	0.9972	1.0000
80	0.9960	0.9978	0.9967	0.9954	0.9920	0.9903	0.9917	0.9940	0.9970	1.0000
90	0.9968	0.9992	0.9984	0.9972	0.9943	0.9925	0.9923	0.9939	0.9967	1.0000
100	0.9964	0.9994	0.9998	0.9998	0.9983	0.9965	0.9949	0.9946	0.9966	1.0000
110	0.9956	0.9979	0.9977	0.9977	0.9976	0.9989	1.000	0.9981	0.9972	1.0000
120	0.9948	0.9964	0.9958	0.9945	0.9928	0.9938	0.9952	0.9968	0.9986	1.0000

Since $(\mathcal{B}(\mathbb{R}^n), d_H)$ is not Euclidean, the factor $1/C_m(\mathbf{u}, \mathbf{v})^{f_{uv}(p)}$ is needed so that, for $p \rightarrow \infty$, the solution D_p tends to D_∞ as required in (iic). It is clear that $C_m(\mathbf{u}, \mathbf{v}) \geq 1$, and if $C_m(\mathbf{u}, \mathbf{v}) \neq 1$ (if E is not Euclidean) eq. (E) depends on $f(p)$ which is not *a priori* fully determined (the theorem only requires $f_{uv}(0) = 0$ and $f_{uv}(\infty) = 1$). At best, some specific $f_{uv}(p)$ is expected to give a completely G -invariant distance extension D_p which would satisfy the triangular inequality. Although the triangular inequality has been verified in some particular cases of Euclidean space, no general indication is yet in our possession.

Evidently, $f_{uv}(p)$ must be non-dimensional: since the exponential coefficient itself is dimensional ($p = -a/r^q kT$), a simple form for $f_{uv}(p)$ could be: $f_{uv}(p) = |p|^k / |p - p_0|^k$ with $p_0 \neq 0$. If this form is retained, let $p_0 = p(\mathbf{u}, \mathbf{v})$, where $p(\mathbf{u}, \mathbf{v})$ is defined by: $K_{p(\mathbf{u}, \mathbf{v})}(\mathbf{u}, \mathbf{v}) = 1$ (if \mathbf{u} is a triangle \mathbf{T} in \mathbb{R}^2 and $\mathbf{v} = \sigma\mathbf{T}$:

Table 5

$K_p(\mathbf{T}_\infty, \sigma\mathbf{T}_\infty)$ vs. $p (> 0$ or $< 0)$ for $\mathbf{T}_\infty (r_A = 1; r_B = 0.46; \theta_B = 52^\circ)$.

$p < 0$	K_p^p	K_p	$p > 0$	K_p^p	K_p
-10.0	2.0656	0.9300	0.01	0.9992	0.9222
-5.00	1.8932	0.8802	0.10	0.9923	0.9257
-2.00	1.3401	0.8639	0.50	0.9693	0.9395
-1.00	1.1255	0.8885	1.00	0.9577	0.9577
-0.50	1.0511	0.9052	2.00	0.9987	0.9993
-0.10	1.0085	0.9189	3.00	1.1431	1.0456
			5.00	1.8227	1.1276
			10.0	7.9378	1.2302
			20.0	155.268	1.2869
			30.0	2940.65	1.3050
			50.0	1026510	1.3189

$p(\mathbf{u}, \mathbf{v}) = p(\mathbf{T})$). Then, if $C_{\mathbf{m}}(\mathbf{u}, \mathbf{v}) > 1$, $C_{\mathbf{m}}(\mathbf{u}, \mathbf{v})^{f_{\mathbf{uv}}(p(\mathbf{u}, \mathbf{v}))} = +\infty$. Consequently, for $p = p(\mathbf{u}, \mathbf{v})$: $\Phi_{\mathbf{u}, \mathbf{v}}(x) = 1$, and eq. (E) is degenerated to “ $1 = 1$ ”. The latter property might be considered as an additional consistency. However, this expression of $f_{\mathbf{uv}}(p)$ is completely consistent only if the variation of $p_0 = p(\mathbf{u}, \mathbf{v})$ is compatible with the condition $f_{\lambda\mathbf{u}, \lambda\mathbf{v}}(p) = f_{\mathbf{u}, \mathbf{v}}(\lambda^2 p)$ proposed for a similarity-invariant definition of $c_p(\mathbf{T})$. Under the assumption $f_{\mathbf{uv}}(p) = |p|^k / |p - p(\mathbf{u}, \mathbf{v})|^k$, the latter condition is equivalent to the condition

$$p(\lambda\mathbf{u}, \lambda\mathbf{v}) = \frac{1}{\lambda^2} p(\mathbf{u}, \mathbf{v}).$$

Now,

$$K_{p(\mathbf{u}, \mathbf{v})/\lambda^2}^{p(\mathbf{u}, \mathbf{v})/\lambda^2}(\lambda\mathbf{u}, \lambda\mathbf{v}) = K_{p(\mathbf{u}, \mathbf{v})}^{p(\mathbf{u}, \mathbf{v})}(\mathbf{u}, \mathbf{v}) = 1$$

(by definition of $p(\mathbf{u}, \mathbf{v})$). Then, $(1/\lambda^2)p(\mathbf{u}, \mathbf{v})$ can indeed be equated to $p(\lambda\mathbf{u}, \lambda\mathbf{v})$. Yet the Euclidean structure is not necessary to ensure $C_{\mathbf{m}}(\mathbf{u}, \mathbf{v}) = 1$ [38]. However:

PROPOSITION

Let (E, d) be a metric space, let G be a group operating on E . Suppose that G preserves d . Then, for any points \mathbf{u} and \mathbf{v} in E and any operations g, h, k in G :

$$\text{Min}(C_{g,h,k}(\mathbf{u}, \mathbf{v}); C_{k,h,g}(\mathbf{u}, \mathbf{v})) \leq 1.$$

Eventually, $C_{g,h,k}(\mathbf{u}, \mathbf{v})$ or $C_{k,h,g}(\mathbf{u}, \mathbf{v})$ may be greater than 1 [39].

Proof

Let (E, d) be a metric space. By using the triangular inequality, we get

$$\begin{aligned} \forall (x, y, x', y') \in E^4, \\ d^2(x, y) + d^2(x', y') - d^2(x, x') - d^2(y, y') &= \{d^2(x, y) + d^2(x', y) - d^2(x, x')\} \\ &+ \{d^2(x', y') + d^2(x', y) - d^2(y, y')\} - 2d^2(x', y) \\ &\leq \{2d(x, y)d(x', y)\} + \{2d(x', y')d(x', y)\} - 2d^2(x', y) \\ &\leq 2d(x', y)\{d(x, y) + d(x', y') - d(x', y)\}. \end{aligned} \tag{i}$$

By permuting x and x'

$$\begin{aligned} d^2(x', y) + d^2(x, y') - d^2(x, x') - d^2(y, y') &\leq 2d(x', y') \\ &\times \{d(x, y') + d(x', y) - d(x', y')\} \end{aligned}$$

- If $d(x, y) + d(x', y') \leq d(x', y) + d(x, y')$, (i) leads to

$$d^2(x, y) + d^2(x', y') - d^2(x, x') - d^2(y, y') \leq 2d(x', y)d(x, y')$$

- Otherwise, i.e. if $d(x, y) + d(x', y') \geq d(x', y) + d(x, y')$, then (ii) leads to

$$d^2(x', y) + d^2(x, y') - d^2(x, x') - d^2(y, y') \leq 2d(x, y)d(x', y').$$

If the group G acts on (E, d) and preserves the distance, and if $(g, h, k) \in G^3$, let: $x = gu, y = v, x' = ku, y' = hv$. The combination of the above inequalities leads to

$$\text{Min}(C_{g,h,k}(\mathbf{u}, \mathbf{v}); C_{k,h,g}(\mathbf{u}, \mathbf{v})) \leq 1. \quad \square$$

Nevertheless, in connection with proven or conjectured inequalities concerning convex bodies [40], it could be conjectured that: $\forall(\mathbf{u}, \mathbf{v}) \in (\mathcal{B}(E^2))^2, \forall(g, h, k) \in G^3$,

$$d_H^2(g\mathbf{u}, \mathbf{v}) + d_H^2(k\mathbf{u}, h\mathbf{v}) - d_H^2(g\mathbf{u}, k\mathbf{u}) - d_H^2(\mathbf{v}, h\mathbf{v}) \leq 2d_H(g\mathbf{u}, h\mathbf{v})d_H(k\mathbf{u}, \mathbf{v})$$

(at least when \mathbf{u}, \mathbf{v} are convex bodies). If this was true, then $C_m(\mathbf{u}, \mathbf{v}) = 1$, and the study of D_p at finite temperature would be greatly simplified, without resorting to any $f_{\mathbf{u}\mathbf{v}}(p)$.

Acknowledgement

I thank Prof. H.B. Kagan for many stimulating discussions.

References

- [1] H. Bock, K. Ruppert, C. Näther, Z. Havlas, H.F. Herrmann, C. Arad, I. Göbel, A. John, J. Meuret, S. Nick, A. Rauschenbach, W. Seitz, T. Vaupel and B. Solouki, *Angew. Chem. Int. Ed. Engl.* 31 (1992) 550.
- [2] For a general introduction, see for example: H.B. Kagan, *Pour la Science* 172 (1992) 42; K. Mislow, *Bull. Soc. Chim. Fr.* 131 (1994) 534.
- [3] R. Chauvin, *J. Phys. Chem.* 96 (1992) 4701.
- [4] See for example: H.A. Feldman, *Anal. Biochem.* 48 (1972) 317.
- [5] E. Ruch and L. Ugi, *Topics in Stereochemistry*, Vol. 4 (Wiley-Interscience, New York, 1969) p. 99.
- [6] If the pairing stereogenicity is treated as an "enthalpic" phenomenon, the corresponding equilibrium constant $K' = (\prod[\mathbf{g}\mathbf{u}_1/\mathbf{u}_1] \cdot \prod[\mathbf{g}\mathbf{u}_2/\mathbf{u}_2]) / (\prod(\mathbf{g}\mathbf{u}_1/\mathbf{u}_2))^2$ is proved to be always smaller than 1 for repulsive-type interactions ($a > 0$) and greater than 1 for attractive-type interactions ($a < 0$). However, $K'(\mathbf{u}_1, \mathbf{u}_2)$ may be equal to 1 for non-equivalent skeletal analogs, and $K'(\mathbf{u}_1, \mathbf{u}_2)$ is not a consistent measure of dissimilarity of \mathbf{u}_1 and \mathbf{u}_2 .
- [7] R. Chauvin, *J. Math. Chem.* 16 (1994) 257.
- [8] R. Chauvin, *J. Phys. Chem.* 96 (1992) 4706.

- [9] Generally speaking G must be a compact group and “ dg ” is its Haar measure: J.P. Serre, *Représentations Linéaires des Groupes Finis*, 3d Ed. (Hermann, Paris, 1967).
- [10] See: *Concepts and Applications of Molecular Similarity*, eds. M.A. Johnson and G.M. Maggiora (Wiley, New York, 1990); *Topics in Current Chemistry I and II*, ed. K.D. Sen, Vols. 173 and 174 (Springer, Berlin, 1995), and references therein.
- [11] R. Chauvin, *J. Math. Chem.* 16 (1994) 269.
- [12] (a) $\forall(\mathbf{u}, \mathbf{v}) \in E^2, D_p(\mathbf{u}, \mathbf{v}) \geq 0$.
(b) $\forall(\mathbf{u}, \mathbf{v}) \in E^2, D_p(\mathbf{u}, \mathbf{v}) = 0 \Rightarrow \exists h \in G, \mathbf{v} = g\mathbf{u}$.
- [13] R. Chauvin, *J. Math. Chem.* 16 (1994) 285.
- [14] See for example: P.G. Mezey, *J. Math. Chem.* 7 (1991) 39; A.Y. Meyer and W.G. Richards, *J. Computer-Aided Mol. Design* 5 (1991) 427; G.D. Purvis III, *J. Computer-Aided Mol. Design* 5 (1991) 55.
- [15] A bounded figure F may also be regarded itself as a molecular skeleton in \mathbb{R}^n . If the ligand parameter mapping this skeleton is vector-valued in \mathbb{R}^n , let $H(F)$ be the set of all the maps \mathbf{u} of F into \mathbb{R}^n satisfying $\int_F \|\mathbf{u}(M)\|^2 d\tau_M < \infty$: $H(F)$ is an Euclidean vector space endowed with the scalar product $\langle \mathbf{u} | \mathbf{v} \rangle = \int_F \mathbf{u}(M)\mathbf{v}(M) d\tau_M$. The theorem is valid for the corresponding Euclidean distance $d(\mathbf{u}, \mathbf{v}) = \| \mathbf{u} - \mathbf{v} \|$ and for the group of rotations preserving F . The coordinate map \mathbf{u}_F of F belongs to $H(F)$: $\mathbf{u}_F : F \rightarrow \mathbb{R}^n$, $\mathbf{u}_F(M) = \mathbf{r}_M$, the position vector of M . It must be stressed that \mathbf{u}_F is chiral even if F itself is achiral. If $\sigma\mathbf{u}_F (\neq \mathbf{u}_{\sigma(F)}!)$ denotes the “enantiomorph” of \mathbf{u}_F , $K_1(\mathbf{u}_F, \sigma\mathbf{u}_F)$ has been calculated for achiral skeletons such as a ball, a sphere, a regular tetrahedron, a trigonal bipyramid, a regular octahedron, a cube, etc.: see text further and ref. [7].
- [16] (a) R. Chauvin, *J. Math. Chem.* 16 (1994) 245.
(b) P.G. Mezey and J. Maruani, *Mol. Phys.* 69 (1990) 97.
- [17] (a) For a general survey, see ref. [18] and: *New Developments in Molecular Chirality*, ed. P.G. Mezey (Kluwer, Dordrecht, 1991). See also ref. [8], and for more recent examples: V.E. Kuz'min, I.B. Stel'makh, M.B. Bekker and D.V. Pozigun, *J. Phys. Org. Chem.* 5 (1992) 295; V.E. Kuz'min, I.B. Stel'makh, I.V. Yudanova, D.V. Pozigun and M.B. Bekker, *J. Phys. Org. Chem.* 5 (1992) 299; A. Seri-Levy and W.G. Richards, *Tetrahedron: Asymm.* 4 (1993) 1917.
(b) H. Zabrodsky and D. Avnir, *J. Am. Chem. Soc.* 117 (1995) 462.
- [18] A.B. Buda, T. Auf der Heyde and K. Mislow, *Angew. Chem. Int. Ed. Engl.* 31 (1992) 989.
- [19] (a) A.B. Buda and K. Mislow, *J. Am. Chem. Soc.* 114 (1992) 6006.
(b) N. Weinberg and K. Mislow, *J. Math. Chem.* 14 (1993) 427.
- [20] *Distance* is here understood with its strong mathematical meaning. This term was used in inverted commas by Mislow to qualify the underlying concept of “chirality measure of the second kind”: see ref. [18].
- [21] Note 13 in ref. [18].
- [22] V. Prelog, *Science* 193 (1976) 17.
- [23] In a somewhat different spirit, the use of the Hausdorff distance was first propounded by Rassat in: A. Rassat, *C.R. Acad. Sc. Paris II* 299 (1984) 53.
- [24] D. Andelman and P.G. de Gennes, *C.R. Acad. Sci. Paris* 307 (II) (1988) 233; D. Andelman, *J. Am. Chem. Soc.* 111 (1989) 6536.
- [25] C.E. Wintner, *J. Chem. Educ.* 60 (1983) 550.
- [26] E.L. Eliel, *Topics in Current Chemistry*, Vol. 105 (Springer, Berlin, 1982) p. 70.
- [27] For chiral simplexes and multiplexes, see: A.S. Dreiding and K. Wirth, *MATCH* 8 (1980) 341.
- [28] G. Gilat, *J. Math. Chem.* 15 (1994) 197. G. Gilat and Y. Gordon, *J. Math. Chem.* 16 (1994) 37.
- [29] “Chirality discrimination” of T and T' means $D_p(T, \sigma T) - D_p(T', \sigma T')$, and is different from $D_p(T, T')$.

- [30] It is logical to imagine that beyond $p = 0$ (i.e. for negative p values referring to repulsive-type interactions rather than to negative temperatures!), the 3-D plots of $c_p(\mathbf{T})$ would become sheerer and tend to the "0-1" (discrete) chirality measure. The problem is open.
- [31] In ref. [8], the chirality index was defined by:

$$\chi_p = \frac{(\sqrt{K})^{1/p} - 1}{(\sqrt{K})^{1/p} + 1} = \frac{\sqrt{K_p} - 1}{\sqrt{K_p} + 1}.$$

- [32] For $p = 2$, the norm $\|f\|_p = [\int_G |f(g)|^p dg]^{1/p}$ is Euclidean or Hermitean: it is associated to the scalar product $(f_1 | f_2) = \int_G f_1(g) f_2^*(g) dg$.
- [33] In other words, whatever is the triangle \mathbf{T} , there exists some $\lambda(\mathbf{T})$ value such that $K_2(\lambda(\mathbf{T})\mathbf{T}, \sigma\lambda(\mathbf{T})\mathbf{T}) = 1$.
- [34] Despite the occurrence of $p/\Delta^2(\mathbf{T})$ instead of p , $c_p(\mathbf{T})$ still refers to a fixed temperature: since $p = -a/(r^d kT)$, the coefficient $p/\Delta^2(\mathbf{T})$ is written as $-(a/\Delta^2(\mathbf{T}))/ (r^d kT)$, where $a/\Delta(\mathbf{T})$ is the interaction constant normalized to the size of \mathbf{T} .
- [35] H. Zabrodsky, S. Peleg and D. Avnir, J. Am. Chem. Soc. 31 (1992) 989.
- [36] E.L. Eliel, Isr. J. Chem. 15 (1977) 7.
- [37] For a discussion of this statement see: K. Mislow and J. Siegel, J. Am. Chem. Soc. 106 (1984) 3319.
- [38] For example if d is the 0, 1-distance ($d(\mathbf{u}, \mathbf{v}) = 1$ if $\mathbf{u} \neq \mathbf{v}$, $d(\mathbf{u}, \mathbf{v}) = 0$ if $\mathbf{u} = \mathbf{v}$), then: $C_m(\mathbf{u}, \mathbf{v}) = 1$.
- [39] In a regular metric space, the inequality

$$-1 \leq C(xx', yy') = \frac{d^2(x, y) + d^2(x', y') - d^2(x, x') - d^2(y, y')}{2d(x', y)d(x, y)} \leq 1$$

is true as soon as $x = x'$ (or $y = y'$). If this inequality stands for any points x, x', y, y' , C would formally define a kind of "cosine" in a non-Euclidean metric space.

- [40] *Discrete Geometry and Convexity*, Annals of the New York Academy of Sciences, Vol. 440 (The New York Academy of Sciences, New York, 1985).