Journal of Mathematical Chemistry 18 (1995) 321–348

Similarity and dissimilarity in posets

D.J. Klein

Texas A&M University at Galveston, Galveston, Texas 77553-1675, USA

Received 30 June 1995; revised 7 September 1995

The relevance of partially ordered sets (or posets) in a wide diversity of contexts in chemistry is emphasized, and the utility of distance functions (or metrics) on such posets is noted. First a notion of "scale similarity" is introduced to make comparisons within certain so-called "scaled" posets, for which there is formulated natural "comparators", which in turn lead to associated distance functions. Beyond taking note of several chemically relevant examples of these "scaled" posets and their consequent associated similarity measures, a second chemically relevant class of so-called "shifted" posets is similarly developed, with examples. Even further extension of some aspects of the current approach is indicated, and finally the multi-posetic character of chemical periodic law is suggested.

1. Introduction

Similarity of molecular structures is of key interest in assessing physical, chemical, and biological properties of different molecular species. See, e.g., ref. [1] for a recent review concerning molecular similarity, and ref. [2] related to the relevance of molecular shape. Moreover gauging similarity and dissimilarity may be viewed to be at the heart of chemical periodic law, e.g., as discussed in [3]. But naturally similarity is of wider interest even in many other sciences and mathematics too. In geometry similarity has a comparably more precise (though perhaps narrower) meaning, in terms of the so-called similarity transformations: reflections, rotations, translations, dilatations, and combinations thereof.

Here as a starting point the idea of "scale similarity" is to be initially developed in a general fashion maintaining some of the flavor of the geometric meaning, while also being applicable in treating "molecular similarity" as well as "dissimilarity". More particularly the general ideas specialize to enable one to compare:

- general shapes (as of molecular electron-density contours) as discussed in section 3,
- conformations of points (e.g., as correspond to the geometric arrangement of nuclei in a molecule) as discussed in section 4,

- distance functions on graphs (as with the graphs possibly representing molecular structures) as described in section 5,
- matrix eigenspectra (such as may correlate to molecular energy-level patterns) and (e.g., electromagnetic) transition spectra as indicated in section 6,
- distributions (as occur in statistical problems) over a common domain as discussed briefly in section 7 (and later).

In comparing distinct items say A and B in any one of these categories, that one item A "exceeds" (or is "greater than") B in some sense might sometimes seem clear, while in other cases it might be ambiguous whether either A or B "exceeds" the other. For instance, in the first category above one might seek to compare a square shape and a circle shape cut from a sheet of paper, whence (as indicated in fig. 1) the square might exceed the circle in the sense that the square can cover the circle, or the circle might cover the square, or perhaps neither will be able to cover the other. In fact in all these categories one may view the items under comparison to be members of what mathematically [4,5] is termed a partially ordered set (or poset). A *partial ordering* relation $\succeq$ on a set $\wp$ is a binary relation such that

$$A \in \wp \Rightarrow A \succeq A\,,$$

$$A \succeq B \text{ and } B \succeq A \Rightarrow A = B\,, \quad \text{for } A, B \in \wp\,,$$

$$A \succeq B \text{ and } B \succeq C \Rightarrow A \succeq C\,, \quad \text{for } A, B, C \in \wp\,,$$

and then $\wp$ is called a *poset*. This concept is now recognized to occur frequently in mathematics, with the explicit definition going back a century [6], though Birkhoff [4] notes the idea occurs in a "fragmentary way" in Leibniz's work. Moreover, Ruch [7] has argued that such mathematical posets, especially those consisting of distributions (under the so-called "majorization" partial ordering), occur in a fundamental manner throughout chemistry and other sciences as well. In fact, in economics [8] and sociology [9] such posets have been explicitly considered. But too, they should be important in ecology [10] where various distributions (and thence too, their comparison) are of central relevance. In biology shape comparison [11] is of interest, so that posets of the first category could be of use. Another partial ordering occurring in biology is that of A being an ancestor to B, as is of central importance in phylogeny [12] – but such partial orderings might also be of relevance in chemical-synthetic networks, and particularly the fundamental investigation of

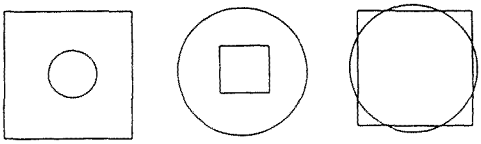

Fig. 1. Illustration of the 3 possible interrelations between a square shape and a circle shape.

molecular evolution [13]. In mathematics Rota [14] has advocated that posets fundamentally underlie the whole field of combinatorics, and this view seems to be often taken in more recent combinatorics texts.

From a moderately general viewpoint in section 2 focus is directed to a somewhat special type of so-called "scaled" poset and some consequent properties are developed. In particular there are developed distance functions (or metrics) which allow comparisons between members of the poset $\wp$, whether or not the members are ordered by the relation $\succeq$. In general a *distance* (or metric) on a set S is a function d from the Cartesian product $S \times S$ to the nonnegative reals such that

$$d(A,B) = d(B,A) \geqslant 0 ,$$

$$d(A,B) = 0 \quad \Leftrightarrow \quad A = B ,$$

$$d(A,B) + d(B,C) \geqslant d(A,C)$$

for arbitrary $A, B, C \in S$. A function d satisfying the first and second conditions (and not necessarily the triangle inequality) would be a less specific comparator function which is termed [15] a *semimetric*. A function d is a *pseudometric* iff it satisfies the first and third conditions as well as a weakened version of the second condition, wherein just: $A = B \Rightarrow d(A,B) = 0$. Here in sections 3,4,5,6,7 the different chemically relevant realizations of such scaled posets, as arise in the already indicated examples, are described in more detail, and the consequent metric-related comparators are indicated. In fact, the development in terms of posets may be viewed as a generalization of certain earlier ideas [16,17,18] for particular cases of the delineated applications. But beyond the unification achieved the current results when specialized back to these particular cases still extend the earlier results, especially as regards measures of dissimilarity. Further in section 7 a way to "turn" a general poset into a scaled poset is indicated.

Section 8 considers a parallel development of the ideas involved in treating another class of so-called "shifted" posets, with associated metric-related ideas. This class includes the distribution-based posets mentioned in section 7 but now treated in a less restrictive way. Section 9 returns to the characterization of molecular shape, and after this to a characterization and measure of molecular symmetry, such as has already been of some interest [2,19,20]. Section 10 considers briefly the chemically much-studied [7,21,22] "majorization" partial ordering of Young diagrams as an example of shiftable posets. Indeed this majorization poset has been extensively studied outside of chemistry [8,9,23], and within chemistry it seems to have arisen several times [24,25] without emphasis of the relevance of the general theory of posets. Our distance-function results for this poset may be viewed as less general than Ruch's (partially-ordered) "direction-distance" functions [22], but with this recognition a parallel generalization for the other chemically relevant posets we have noted seems an interesting idea. An extension in a different direction in "closer" context to sections 2 and 8 is briefly indicated in section 11. Finally section 12 returns to the matter of periodic law and its relations to posets. The overall ideas seem to be of wide potential applicability, presumably even outside of chemistry.

2. Scaled posets

Granted a poset $\wp$ we say that it is *scaled* provided: first, that $\wp$ is closed under scalar multiplication by elements of the set $\Re_+$ of positive reals with multiplication by 1 leaving each element of $\wp$ invariant, i.e.

$$A \in \wp \text{ and } k \in \Re_+ \quad \Rightarrow \quad k \cdot A \in \wp \text{ and } 1 \cdot A = A$$

and second, for $A \in \wp$ and $h, k \in \Re_+$,

$$(hk) \cdot A = h \cdot (k \cdot A)\,,$$

$$h \geqslant k \text{ and } A \succeq B \quad \Rightarrow \quad h \cdot A \succeq k \cdot B\,.$$

Next granted such a scaled poset $\wp$ a *comparator* $c(A \nearrow B)$ for $A, B \in \wp$ is defined to be the least value of $x \in \Re_+$ such that $x \cdot A$ exceeds B – i.e.

$$c(A \nearrow B) \equiv \min\{x \in \Re_+ \mid x \cdot A \succeq B\}\,,$$

(and if no such minimum exists one might say $c(A \nearrow B) = \infty$). Evidently $c(A \nearrow B)$ is a factor needed in order to "circumscribe" A about B. Moreover we define $A, B \in \wp$ to be *scale equivalent*, the condition of which we denote by $A \approx B$, iff $c(A \nearrow B)c(B \nearrow A) = 1$. A few fundamental properties of our comparator are identified in the following:

THEOREM A

For A, B, C members of a scaled poset $\wp$,

$$1/c(A \nearrow B) = \max\{x \mid A \succeq x \cdot B\}\,,$$

$$c(A \nearrow B)c(B \nearrow A) \geqslant 1 \quad \text{with equality iff } A \approx B\,,$$

$$c(A \nearrow B)c(B \nearrow C) \geqslant c(A \nearrow C)\,.$$

But before proving this let us "interpret" these properties. First the maximum appearing in the theorem defines a factor (which we might denote $c(A \swarrow B)$) evidently to "inscribe" B into A, and the theorem says this inscription factor is simply the inverse of the circumscription factor when the compared objects are interchanged for application of the scale factor. The second inequality when rewritten as $c(A \nearrow B) \geqslant c(A \swarrow B)$ simply says the factor to achieve circumscription (of A

about B) must be as large as the factor to achieve inscription (of B within A). Finally the last "telescope" inequality says that if that in applying two rescalings first to circumscribe A about B then to circumscribe this result about C, one has in fact achieved (or overachieved) circumscription of A about C.

Now the proof of theorem A actually follows fairly closely the preceding "interpretations". First evidently

$$c(A \swarrow B) = \max\{x \mid x^{-1} \cdot A \succeq B\}$$

so that this maximum x value is a minimum x^{-1} value as appears in the definition of $c(A \rightarrow B)$, so that $\{c(A \swarrow B)\}^{-1} = c(A \nearrow B)$. For the second inequality of the theorem we note that

$$\{c(B \nearrow A)c(A \nearrow B)\} \cdot A = c(B \nearrow A) \cdot \{c(A \nearrow B) \cdot A\} \succeq c(B \nearrow A) \cdot B \succeq A\,.$$

But from the scaled-poset definition

$$h \leqslant 1 \Rightarrow A \succeq h \cdot A$$

so that for $h \cdot A \succeq A$ we must have $h \geqslant 1$, and in particular the second result of the theorem is obtained. For the final result we proceed similarly

$$\{c(B \nearrow C)c(A \nearrow B)\} \cdot A \succeq c(B \nearrow C) \cdot B \succeq C$$

whence the telescoping inequality follows.

In view of this theorem for $A, B \in \wp$ we might define

$$m(A \rightarrow B) = \{c(A \nearrow B)/c(B \nearrow A)\}^{1/2}$$

as the *mean magnification* to bring A to (an intermediate sort of "optimal" comparison with) B.

Also of interest for $A, B \in \wp$ is the *interlacing* comparator $c_{A,B}$ defined as the minimum (or infimum) value $x \in \Re_+$ such that

$$x \cdot A \succeq B \succeq x^{-1} \cdot A\,.$$

Evidently the inequality condition implies

$$x \cdot B \succeq x \cdot (x^{-1} \cdot A) = A = x^{-1} \cdot (x \cdot A) \succeq x^{-1} \cdot B$$

so that $c_{A,B}$ is symmetric in its arguments. Also this inequality implies

$$x \cdot A \succeq x^{-1} \cdot A \quad \Rightarrow \quad x^2 \cdot A \succeq A \quad \Rightarrow \quad x^2 \geqslant 1$$

so that $x \geqslant 1$, with in fact equality implying $A = B$. Further on comparison of the defining conditions for $c_{A,B}$ with those for $c(A \nearrow B)$ and $c(B \nearrow A)$ one sees that $c_{A,B}$ must be the maximum of $c(A \nearrow B)$ and $c(B \nearrow A)$. Finally for $A, B, C \in \wp$,

$$c_{A,B} \cdot (c_{B,C} \cdot C) \succeq c_{A,B} \cdot B \succeq A \succeq c_{A,B}^{-1} \cdot B \succeq c_{A,B}^{-1} \cdot (c_{B,C}^{-1} \cdot C)$$

so that on comparison with the defining inequality for $c_{A,C}$ one sees that these interlacing comparators satisfy a sort of telescoping inequality, with $c_{A,B}c_{B,C}$ at least as great as $c_{A,C}$. In fact we have established:

THEOREM B

For A, B, C members of a scaled poset $\wp$ the interlacing-comparator function satisfies

$$c_{A,B} = c_{B,A} = \max\{c(A \nearrow B), c(B \nearrow A)\} \geqslant 1\,,$$

$$c_{A,B} = 1 \Rightarrow A = B\,,$$

$$c_{A,B}c_{B,C} \geqslant c_{A,C}\,.$$

One may think of $c_{A,B}$ as some minimal scale change which when applied to A circumscribes it about B and when applied to B circumscribes it about A, thereby, "interlacing" A and B. Next we define an *interlacing* distance

$$d_i(A, B) \equiv 2 \cdot \log c_{A,B}\,,$$

and a *form* distance

$$d_f(A, B) \equiv |\log c(A \nearrow B)| + |\log c(B \nearrow A)|\,.$$

This of course turns out to be good nomenclature if these are distance functions (or metrics) as defined in the introduction. In fact as rather directly follows from our preceding theorem we have:

THEOREM C

The functions d_i and d_f are distance functions on the scaled poset $\wp$.

In addition to the ratio of $c(A \nearrow B)$ and $c(B \nearrow A)$ appearing in $m(A \rightarrow B)$ one might surmise that the product of such a pair of comparators also would be a natural piece of information concerning the relation between A and B. Particularly since the product encodes a multiplicative difference between "circumscription" and "inscription" factors, this product would seem to represent a degree of geometric dissimilarity. Thence we define a *similarity* distance between A and B,

$$d_s(A, B) = \log\{c(A \nearrow B)c(B \nearrow A)\}\,.$$

Again such a name would be most reasonable if d_s were to turn out to be some sort of distance function (or metric), as indeed theorem A rather readily implies it is:

THEOREM D

The function d_s is a pseudometric on the scaled poset $\wp$ and is a distance function on the set of scale equivalence classes of $\wp$.

The last two theorems may be picturesquely interpreted, as indicated in fig. 2. In conjunction with this interpretation it is readily seen that $d_s(A,B) \leqslant d_f(A,B) \leqslant d_i(A,B)$, with equality in the second instance iff $A \approx B$.

3. Shape comparison

Here the poset $\wp$ considered is comprised from closed subsets of a Euclidean space $\mathcal{E}$, and the partial ordering is based upon set inclusion. More properly $\wp$ consists of chiral-potent *isometry* equivalence classes of connected subsets of $\mathcal{E}$ – that is, each A consists of all translations and/or (proper) rotations of a representative of A. Indeed sometimes one does not distinguish A from its representative. For the partial ordering we say that $A \succeq B$ iff there exists a set in the class A which is a subset of a set in the class B. The scalar multiplication of A corresponds to the requisite dilatation of its constituent members. The comparators $c(A \nearrow B)$ and $c(B \nearrow A)$ are illustratively indicated in fig. 3 – and they have been proposed by Mezey [16] as useful shape descriptors, especially in characterizing molecular electron-density contours, which in turn [2] are of use in characterizing molecular properties. The comparators $m(A \rightarrow B)$; $d_f(A,B)$ and $d_s(A,B)$ can be argued to have very natural interpretations: first, $m(A \rightarrow B)$ represents a degree of discongruence (i.e., a dissimilarity modulo rotation and translation); second, $d_f(A,B)$ represents an interform distance modulo rotation and translations; and third, $d_s(A,B)$ represents an intershape distance modulo rotation, translation, and dilatation.

With these distance functions in hand, they may be utilized for measures of the "degree of symmetry" or the "chirality" of a molecule, such ideas having been much discussed recently [20]. For instance, with regard to "chirality" of a molecular shape A, one can proceed with A^* denoting its mirror image, whence $d_f(A,A^*)$

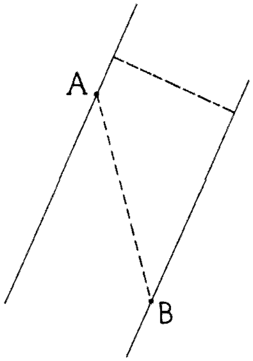

Fig. 2. A picturesque interpretation of theorems C and D. The two elements A and B of $\wp$ are represented by points, and their associated scale equivalence classes are represented by parallel (near vertical) lines. The distance $d_i(A,B)$ between A and B is indicated by the longer dashed line, while the distance d_s(A,B) between the two equivalence classes is represented by a shortest length between the two solid lines, one such length being indicated by the shorter dashed line.

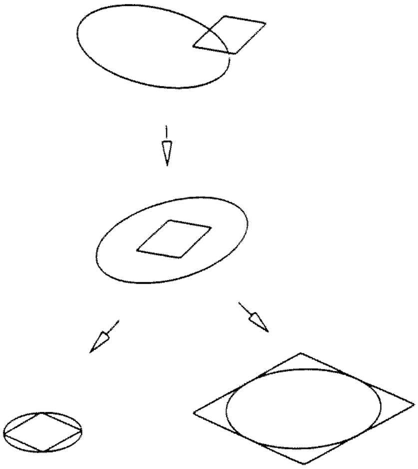

Fig. 3. An illustration of the present ideas for A and B being diamond and ellipse shapes in the 2-dimensional Euclidean space $\mathcal{E}_2$. First in the upper central part of the figure a diamond and ellipse are translated and rotated to bring them into a maximum extent of correspondence. Then two different dilatations are made: first the ellipse is dilatated by a factor of 0.5 to yield the first covering; and second, alternatively the diamond instead is dilated by a factor of 2.8 to yield the second covering.

is a measure of the *degree of chirality* of A (and of A^*). An extended notion of degree of chirality (allowing for dilatations in addition to proper rotations and translations in drawing equivalences) is provided by $d_s(A, A^*)$. The *degree of $\mathcal{G}$-symmetry* (where $\mathcal{G}$ is some suitable point group) of a shape A would be $d_f(A, A_\mathcal{G})$ where $A_\mathcal{G}$ is some suitable symmetric reference. If GA denotes the shape to which A is carried by $G \in \mathcal{G}$, then quite reasonably $A_\mathcal{G}$ is the union of these different GA,

$$A_\mathcal{G} = \bigcup_{G \in \mathcal{G}} GA .$$

Here $d_f(A, A_\mathcal{G})$ is a symmetry measure with respect to a particular group $\mathcal{G}$ having center and axis located and oriented in a given manner relative to A. Often the degree of symmetry of interest might be that minimal value obtained through variation of the location and orientation associated to $\mathcal{G}$.

There are other possible modifications of the partial ordering. Each equivalence class of $\wp$ involves transformation of some representative member of $A \in \wp$ by a (sub)group $\mathcal{G}$ of similarity transformations on $\mathcal{E}$, and this group $\mathcal{G}$ could be chosen differently, e.g., to include reflections whence the standard mathematical (achiral) isometry equivalence classes are obtained. Also in the preceding paragraph there is a second group $\mathcal{G}^\circ$ of isometric transformations, and this subgroup of $\mathcal{G}$ can be chosen in different ways. E.g., $\mathcal{G}^\circ$ might involve just rotations and translations of molecular shapes along a 2-dimensional surface to which the molecules might happen to be confined (as often occurs in catalytic processes).

Of course the shapes indicated here for consideration are typically continuous, with a continuum of possibilities for translation and rotation. The operations involved in comparison can be made more generally manageable if instead one characterizes a shape thusly: first, choose a finite set of characteristic points; second, condense the shape information to be retained just to the finite list of distances between these points; and third, proceed as in the next section. Indeed for biotaxonomic purposes just such an approach has been advocated [26] -the corresponding points (e.g.) on animal skulls of different (sub) species being long known as "homologous". For the case of smooth electron-density surfaces the characteristic set of points could be chosen as the so-called [27] "critical" points whereat derivatives in every direction are 0 or infinite, and there are two ways to choose the interpoint distances: through (3-dimensional) space or along surface geodesics.

4. Conformation comparison

Here the members of the poset $\wp$ are congruence classes of equivalently labelled sets of points embedded in a Euclidean space $\mathcal{E}$. That is, each member $A \in \wp$ corresponds to an embedding, modulo translation and rotation, of a set $\{1,2,\dots,N\}$ of points. Thus A is uniquely labelled by an N by N matrix $\mathbf{D}(A)$ of Euclidean distances $D_{ij}(A)$ between pairs of points $\{i,j\}$ of (a conformation of) A. The partial ordering is such that $A \succeq B$ iff

$$D_{ij}(A) \geqslant D_{ij}(B)\,, \quad \text{all } i,j \in \{1,2,\dots,N\}\,.$$

The scalar multiplication of section 2 corresponds to dilatation of the points in the representatives in A or equivalently of the usual scalar multiplication of the matrices $\mathbf{D}(A)$. We term such a scaled poset $\wp$ so represented in terms of matrices a *matric* poset. Now $c(A \nearrow B)$ is seen to be the maximum ratio $D_{ij}(B)/D_{ij}(A)$ for distinct $i,j \in \{1,2,\dots,N\}$. As an even more particular interpretation the points may be viewed to correspond to nuclei in a molecule, whence $d_f(A,B)$ is a measure of dissimilarity between molecular conformations A and B – such measures being of much potential interest for structure-property correlations, such as considered in [1,2,17,29,30] for other molecular conformation descriptors.

It may be noted that there is another reasonable pair of comparators other than that of section 2 at least when dealing with the present matric posets. To see this consider the eigenproblem for the maximum-magnitude eigenvalue $\lambda(A \nearrow B)$ to a *distance-ratio* matrix $\mathbf{R}(A \nearrow B)$ with off-diagonal elements

$$R_{ij}(A \nearrow B) = D_{ij}(B)/D_{ij}(A)$$

and diagonal elements 0. The indicated eigenproblem

$$\mathbf{R}(A \nearrow B) \mid A \nearrow B\rangle = \lambda(A \nearrow B) \mid A \nearrow B\rangle$$

is for a (Frobenius–Perron) matrix [31], with all off-diagonal elements positive, so that it has a (nodeless) eigenvector $| A \nearrow B\rangle$ with all components $\langle i | A \nearrow B\rangle$ positive. Now the eigenvalue may be expressed in terms of $| A \nearrow B\rangle$-components

$$\lambda(A \nearrow B) = \sum_{ij} \{\langle i | A \nearrow B\rangle \langle j | A \nearrow B\rangle \langle A \nearrow B | A \nearrow B\rangle^{-1}\} R_{ij}(A \nearrow B),$$

where in general $\langle \theta |$ is the (complex-conjugated) transpose to the column vector $| \theta\rangle$ and $\langle \theta | \zeta\rangle$ denotes the (standard) inner product between $| \theta\rangle$ and $| \zeta\rangle$. Then, with $| \phi\rangle$ the vector all of whose components are 1, it is seen that

$$\begin{aligned}&\lambda(A \nearrow B)\langle A \nearrow B | A \nearrow B\rangle \langle \phi | A \nearrow B\rangle^{-2}\\ &\quad = \sum_{ij} \{\langle i | A \nearrow B\rangle \langle A \nearrow B | j\rangle \langle \phi | A \nearrow B\rangle^{-2}\} R_{ij}(A \nearrow B)\end{aligned}$$

is a type of average magnification to bring B into comparison with A (this representing an average since the positive terms in brackets on the right-hand side of this equation sum to 1). Indeed such a related idea has already been suggested by Randić et al.[17], though more in connection with the application of the next section. Here we define a comparator

$$c'(A \nearrow B) \equiv \lambda(A \nearrow B)\langle A \nearrow B | A \nearrow B\rangle / \langle \phi | A \nearrow B\rangle^2$$

as a sort of average magnification factor. Also we note that there is of course the possibility of both $\mathbf{R}(A \nearrow B)$ and $\mathbf{R}(B \nearrow A)$ yielding a symmetric pair of comparators $c'(A \nearrow B)$ and $c'(B \nearrow A)$. Thus in parallel to the development of section 2 we might introduce

$$c'_{A,B} \equiv \max\{c'(A \nearrow B), c'(B \nearrow A)\},$$

$$m'(A \to B) \equiv \{c'(A \nearrow B)/c'(B \nearrow A)\}^{1/2},$$

$$d'_i(A \to B) \equiv 2 \cdot \log c'_{A,B},$$

$$d'_s(A, B) \equiv \log\{c'(A \nearrow B)c'(B \nearrow A)\}.$$

Then "partial" results apply:

THEOREM A′

For A, B members of a matric poset $\wp$,

$$c'(A \nearrow B)c'(B \nearrow A) \geqslant 1 \quad \text{with equality iff } A \approx B.$$

THEOREM B′

For A, B members of a matric poset $\wp$,

$$c'_{A,B} = c'_{B,A} \geqslant 1 \quad \text{with equality iff } A = B.$$

THEOREM C′

The function d_i' is a semimetric on a matric poset $\wp$.

THEOREM D′

The function d_s' is a semimetric on the similarity equivalence classes of a matric poset $\wp$.

Proofs are found in the appendix. So far it is not known whether d_i' or d_s' satisfy the triangle inequality (as is crucial to deciding whether they are distance functions).

With the different points representing the nuclei of a molecular conformation, the ideas of this section too may be utilized to define a degree of chirality or of symmetry. With $\mathbf{r}_i$ representing the nuclear position vectors of A, the reference $A_{\mathcal{G}}$ for a symmetry group $\mathcal{G}$ is naturally identified with position vectors

$$\mathbf{r}_{\mathcal{G}_i} \equiv \frac{1}{|\mathcal{G}|}\sum_{G\in\mathcal{G}} G\mathbf{r}_i .$$

Notably both the degree of chirality and degree of $\mathcal{G}$-symmetry may be defined either in terms of d_f or of (just the semimetric) d_f'.

Finally, it may be mentioned that for a general (e.g. molecular) shape an even greater condensation of the shape information may be made (than to the list of distances between characteristic points, as suggested at the end of the preceding section). Rather generally for shapes in n-dimensional Euclidean space $\mathcal{E}_n$ one could keep just the n principle moments of inertia (or their square roots). This gives an n-vector (rather than a matrix) of distances, but these may be ratioed much as the matrices $\mathbf{D}(A)$ and unprimed comparators computed. Such inertia-based unprimed comparators do however have a characteristic in common with the primed comparators: both entail some sort of an average. Note in the inertia-based scheme (or the scheme indicated in the last paragraph of the preceding section) the discretized set is in our theorems interpreted as $\wp$ (rather than the underlying shapes) since more than one shape may be identified to the same discretized set of shape descriptors. See also the first paragraph of section 7 and the next to last paragraph of section 8 for the treatment of n-vectors as distributions. Another possible distance function avoiding discretization and using a type of "expansion" other than dilatation is found in [27] – this approach being related to the ideas of section 11. All this along with the earlier ideas of this and the preceding sections then possibly provide promising new molecular size and shape descriptors, as are of much general interest [2,29].

5. Distance-function comparison

Here the poset is to be the set of distance functions on a graph G (or as in the next paragraph on a class of graphs), and the partial ordering may be expressed in terms of a distance-ratio matrix much as in the previous section except that now the ratio is with respect to two different distance functions on a single set (rather than for one distance function applied to two corresponding sets). That is, we have another example of a matric poset. The standard distance function on graphs is defined in terms of a minimum number of steps in a path between vertex pairs – see e.g. [32]. But of course in chemistry it is important to distinguish single and multiple bonds (or perhaps also between bonds of different bond lengths), so that it is natural to introduce a weighting for the bonds and thence also for paths and associated path lengths to yield a modified graphical distance. If paths are given weights that are sums of bond weights (e.g., inverse bond orders) and the distance function between two vertices is taken as the minimum weight path between the two vertices, then the resultant function is a "simple" extension of the standard unit-weighted one. But the same chemical arguments leading to such distance diminishment for multiple bonds also suggests that a similar sort of distance diminishment should apply with multiple pathways between non-nearest neighbors. Indeed there is [33] a "new" so-favored (so-called) "resistance" distance.

Of course for a graph representing a molecule (embedded in Euclidean space) there is the ordinary Euclidean distance between vertices. Indeed the bulk of the field of "polymer statistics" [34] may be viewed to focus on comparisons between the (mean) graphical and Euclidean distances between monomer units, as averaged over different polymer conformations. Typically for linear-chain polymers the number of monomer units is used in place of a mean graph distance, but for branched polymers the relationships are different and are of interest [35,36]. Moreover Randić et al. [17] have argued that a comparison between the ordinary graphical distance and Euclidean distance for a graph representing that molecule should be a measure of the "foldedness" of that molecule - such a measure presumably being of interest in dealing with the important feature of protein folding [29]. This earlier proposal for a foldedness comparator is essentially just $\lambda(A \nearrow B)$ of the previous section, but here we have a more comprehensive suite of available possibilities, with theromatically guaranteed desirable properties. Too it may be mentioned that the pair of comparators $c(A \nearrow B)$ and $c(B \nearrow A)$ have been previously considered [18], though primarily as indicated in the next paragraph.

The idea of this section may be extended to apply to a class $\mathbb{C}$ of graphs, e.g. such as the class of all finite connected graphs (perhaps also with limited vertex degrees, as often is a reasonable restriction for molecular graphs). The comparator $c(A \nearrow B)$ between two graphical distance functions is taken to be the maximum of the corresponding comparators for individual graphs of the class $\mathbb{C}$. Thence for that graph-theoretic and Euclidean distances on molecular graphs such that there are certain (even) infinite classes (1) it has been argued [18] tighter bounds $c(A \nearrow B)$ and $c(B \nearrow A)$ lead to similar asymptotic functional distance-dependencies for interatomic interactions.

6. Spectral comparison

Here the poset is based on the set of Hermitian N by N matrices and the partial ordering might be related to the distribution of eigenvalues. More particularly each member A of $\wp$ is a *shift* equivalence class of matrices $\mathbf{A}_h \equiv \mathbf{A}_0 + h\mathbf{I}$ where h is an arbitrary real number, $\mathbf{I}$ is the identity matrix, and $\mathbf{A}_0$ is a class representative conveniently taken to be the member of A whose lowest eigenvalue is 0. Every member of such a class leads to the same eigenvectors and the same eigenvalue difference spectrum (as so often is all that is really of chemical interest). We can imagine the eigenvalues of $\mathbf{A}_0$ for each $A \in \wp$ being ordered

$$0 = \lambda_1(A) \leqslant \lambda_2(A) \leqslant \ldots \leqslant \lambda_N(A) .$$

Then the first partial ordering of $\wp$ to be considered is: $A \succeq B$ iff

$$\lambda_i(A) \geqslant \lambda_i(B), \quad i = 1 \text{ to } N .$$

Of course the scalar multiplication of A (for the discussion of section 2) entails just the ordinary scalar multiplication of the constituent matrices of A. Evidently

$$c(A \nearrow B) = \max\{\lambda_i(B)/\lambda_i(A) \mid i = 2, \ldots, N\} .$$

The closer $d_i(A, B)$ is to 0 the closer the eigenvalue spectra of A and B are to differing in naught but a shift, and if thermodynamic properties are statistical-mechanically computed with such A and B identifying Hamiltonian matrices the closer many properties (such as specific heat) are to being identical. The closer $d_s(A, B)$ is to 0 the closer the eigenspectra are to differing only in a shift and rescaling, and e.g. the statistical-mechanically computed specific heat comes closer in all but the scale of temperature.

A second related partial ordering on the same set is possible. We say $A \succeq B$ if $\mathbf{A}_0 - \mathbf{B}_0$ is non-negative definite. Indeed this partial ordering has been studied mathematically [37]. Again $d_i(A, B)$ and $d_s(A, B)$ (and perhaps also the $d'_*(A, B)$) measure deviations of eigenspectra away from a simple shifting and rescaling. A third conceivable possibility would treat [38] the set (perhaps of positively shifted) eigenvalues as a distribution partially ordered via majorization as in section 10.

Another related type of application occurs for transition spectra, say for molecules in an electromagnetic field or for condensed-phase molecules in a phonon field. Then there are two relevant matrices $\Delta(A)$ and $\mathbf{T}(A)$ which respectively identify energy-level differences and transition intensities, such matrices having elements

$$\Delta_{ij}(A) \equiv \lambda_i(A) - \lambda_j(A) ,$$

$$T_{ij}(A) \equiv | \langle \lambda_i(A) \mid \mathbf{T} \mid \lambda_j(A) \rangle |^2 .$$

with A labelling a Hamiltonian matrix and $\mathbf{T}$ being a transition matrix. Partial orderings for both types of matrices may be introduced somewhat as in the preceding two sections, though for the $\mathbf{\Delta}$-matrices one would say $\mathbf{\Delta}(B) \succeq \mathbf{\Delta}(A)$ means just that the nonnegative elements of $\mathbf{\Delta}(B)$ exceed the corresponding elements of $\mathbf{\Delta}(\mathbf{A})$. The partial orderings for the associated posets $\wp_\Delta$ and $\wp_\mathrm{T}$ of matrices then leads in a natural way to a partial ordering on the Cartesian product $\wp_\Delta \times \wp_\mathrm{T}$

$$(\mathbf{\Delta}_2 \times \mathbf{T}_2) \succeq (\mathbf{\Delta}_1 \times \mathbf{T}_1) \quad \Leftrightarrow \quad \mathbf{\Delta}_2 \succeq \mathbf{\Delta}_1 \text{ and } \mathbf{T}_2 \succeq \mathbf{T}_1 .$$

But rather than apply a single scalar multiple to both $\mathbf{\Delta}$ and $\mathbf{T}$ in $(\mathbf{\Delta} \times \mathbf{T})$ it seems more appropriate to develop separately the distances d_i^Δ, d_s^Δ and d_i^T, d_s^T for $\wp_\Delta$ and $\wp_\mathrm{T}$. While one of these posets $\wp_\Delta$ and $\wp_\mathrm{T}$ might be treated as a scale-invariate poset (as in the preceding sections) the other might be treated as a "shifted" poset as discussed in section 8. Overall the distances so defined can be imagined to be of use to measure similarities of model transition spectra to experimental transition spectra.

7. Scaleability and mimicry

As a first extension of incomplete generality, we note that if a poset $\wp$ may be embedded in a larger one $\wp^+$ which contains $\wp$ and which is scaled, then one may simply use the comparators and distance functions for $\wp$ as inherited from $\wp^+$, and we might say $\wp$ is *scaleable*. A (presumably important) example occurs with $\wp$ being a set of probability distributions with the same support S – that is, $\wp$ consists of all functions $p: S \to \Re_+$ such that

$$p(s) > 0, \quad \text{all } s \in S, \quad \text{and } \sum_{s \in S} p(s) = 1 .$$

(In fact as noted in section 9 a slightly stronger condition than having the same support S needs to be satisfied: namely that the values of $p(s)$ in these regions of support be bounded below by some positive $\in$.) Evidently under the usual scalar multiplication of functions, $\wp$ is not scaled because of the normalization condition, but disregarding this condition one obtains a set $\wp^+$ which is scaled. Here the unnormalized functions in $\wp^+$ might be interpreted as some sort of frequencies, though this is not required in order to define the comparators and distances on $\wp$. A related type of problem occurs in "fuzzy set theory" [39] where now $p(s)$ is a "degree-of-membership" function with the condition $0 \leqslant p(s) \leqslant 1$ in place of the normalization condition. If one wishes to avoid the presumption of a common support, the approach of section 8 might be considered.

As a rather different general extension it may be noted that corresponding to any poset one may usually construct a related scalable poset mimicking the parent poset. Given a general poset $\wp$ a function f from $\wp$ to the real numbers $\Re$ is *isotonic* (or partial-ordering homomorphic, or order-preserving, or Schur convex) iff for $A, B \in \wp$

$$A \succeq B \quad \Rightarrow \quad f(A) \geqslant f(B)\,.$$

Then for n linearly independent such isotonic functions $f_1, f_2, \ldots, f_n$ taking nonnegative values there is a non-negative cone of such functions

$$\mathbb{C} = \left\{ \sum_{i=1}^{n} a_i f_i \mid a_i \geqslant 0, i = 1, \ldots, n, \sum_{i=1}^{n} a_i > 0 \right\}$$

and for $A, B \in \wp$ we can define a comparator $c(A \nearrow B)$ as the minimum value $x \in \Re_+$ such that $xf(A) \geqslant f(B)$ for all $f \in \mathbb{C}$. Of course though the cone $\mathbb{C}$ contains an infinite number of functions, one only need check the inequalities for each f_i in the basis of $\mathbb{C}$. From such comparators the rest of the theory of section 2 follows. The construction in terms of isotonic functions may be viewed as actually treating the poset $\wp_\mathbb{C}$ with members $A_\mathbb{C}$ consisting of sequences $(f_1(A), f_2(A), \ldots, f_n(A))$ in correspondence with $A \in \wp$ – the consequent poset $\wp_\mathbb{C}$ being scaleable (through ordinary scalar multiplication of these vector-like sequences) and serving as a *mimic* for $\wp$.

In a number of circumstances what may actually be available could be naught but the mimic $\wp_\mathbb{C}$. Indeed this viewpoint has already been suggested [40] in considering chemical "aromaticity" (of benzenoid species). Here there are [41] several different standard possible aromaticity measures: from thermodynamic stability, from bond lengths, from NMR chemical shifts (and "ring currents"), or from chemical reactivity (of any one of a few different types of reactions). Each such aromaticity measure may be used to give numerical aromaticity indexes f_i, each of which may be as an isotonic function on the underlying aromaticity poset $\wp$. But though these indices need not necessarily completely characterize $\wp$, they do characterize the mimic $\wp_\mathbb{C}$.

Again, it may be emphasized, that there are frequent examples in chemistry where it is the mimic $\wp_\mathbb{C}$ which is more directly available. Such posetic circumstances can be viewed to occur in speaking of the "oxidizing power" of a compound – the different isotonic functions f_i corresponding to oxidizing strengths (e.g., potentials) in a range of different chemical environments. Acidities too might be viewed to be similarly partially ordered if one is interested in the (e.g., pH) responses of acids in a selected set of different solutions – say in pure water, in seawaters (as relevant for marine chemistry), and in cellular fluids (as relevant for biochemistry).

8. Shifted posets

The underlying idea behind the development of section 2 can be utilized to treat posets which are not necessarily scaled. We say a poset $\wp$ is *shifted* iff first $\wp$ contains a scaled linearly ordered subset

$$\mathcal{S} = \{k \cdot I \mid k \in \Re_+\}$$

second $\wp$ is closed under an addition with the "zero shift" $0 \equiv k \cdot I$ for $k = 0$ leaving each element invariant

$$A \in \wp \text{ and } S \in \mathcal{S} \quad \Rightarrow \quad A + S \in \wp$$

$$A \in \wp \Rightarrow A + 0 = A$$

and third for $A, B \in \wp, h, k \in \Re_+$,

$$A \succeq B \text{ and } h \geqslant k \quad \Rightarrow \quad A + h \cdot I \succeq B + k \cdot I\,.$$

Typically we extend the range of the scalars to allow negative values, with

$$A \succeq B - k \cdot I \quad \Leftrightarrow \quad A + k \cdot I \succeq B\,.$$

Now instead of the "multiplicative" (scale) comparators of section 2 we can introduce an "additive" comparator of two elements A and B of a shifted poset $\wp$

$$d''(A \nearrow B) \equiv \min\{x \mid A + x \cdot I \succeq B\}\,.$$

We can recover a "multiplicative" scale comparators via

$$c''(A \nearrow B) \equiv \exp d''(A \nearrow B)\,,$$

$$c''_{A,B} \equiv \max\{c''(A \nearrow B), c''(B \nearrow A)\}$$

and candidate distance functions via

$$d''_i(A, B) \equiv \max\{d''(A \nearrow B), d''(B \nearrow A)\}\,,$$

$$d''_s(A, B) \equiv d''(A \nearrow B) + d''(B \nearrow A)\,.$$

Further we say A and B are *shift* equivalent iff there is a shift relating A and B, i.e., there is a real number x such that $A + x \cdot I = B$. The appropriateness of these definitions is revealed through the results:

THEOREM A″

For A, B, C members of a shifted poset $\wp$,

$$-d''(A \nearrow B) = \max\{x \mid A \succeq B + x \cdot I\}\,,$$

$$c''(A \nearrow B)c''(B \nearrow A) \geqslant 1 \quad \text{with equality iff } A \approx B\,,$$

$$c''(A \nearrow B)c''(B \nearrow C) \geqslant c''(A \nearrow C)\,.$$

THEOREM B″

For A, B, C members of a shifted poset $\wp$,

$$c''_{A,B} = c''_{B,A} \geqslant 1 \quad \text{with equality iff } A = B\,,$$

$$c''_{A,B} c''_{B,C} \geqslant c''_{A,C}\,.$$

THEOREM C″

The function d''_i is a distance function on the shifted poset $\wp$.

THEOREM D″

The function d''_s is a pseudometric on the shifted poset $\wp$ and a distance function on the shift equivalence classes of $\wp$.

The proofs parallel those of theorems A,B, C and D in section 2. The quantity $d''(A \nearrow B)$ defined here, as well as the analogous quantity $d(A \nearrow B) \equiv \ln c(A \nearrow B)$ which could have been defined in section 2, have a fundamental interpretation. Either quantity may be viewed as "relative coordinates" thereby giving rise to a coordinatized "ruler geometry" [42], whence the schematic of fig. 2 becomes of further relevance.

One important example of shifted or shiftable posets involves the distributions discussed in the first paragraph of the preceding section. A distribution $p(S)$ with $S \in \mathcal{S}$ is shifted by adding a common number to every argument, so that a shift by k gives the new distribution $p'(S) = p(S) + k$. Significantly the constraint of a common support for all compared distributions may now be lifted. Indeed, normalization and nonnegativity constraints may also be lifted, whence we are dealing with the set of all functions from $\mathcal{S}$ to $\Re$. In application to probability distributions the distance d_i has been called [43] the "Kolgomorov" distance [44]. It is also of relavance for quantum mechanics [45].

In connection with fuzzy set theory our results suggest a means of measuring the "degree of fuzziness" $d_f(A)$ of a fuzzy set A, which has a degree of membership function $p_A(S)$. A set B is has no fuzziness (exactly) when it is *sharp* in the sense that $P_B(S) = 0$ or 1 for all S. So we could take

$$d_f(A) \equiv \min\{d''_i(A,B) \mid B \text{ is sharp}\}\,.$$

This though is very sensitive to having even a single element of the fuzzy set A not included in a sharp manner. This sensitiveness may be ameliorated using an average over single element subsets

$$d_f(A) \equiv \frac{1}{|\mathcal{S}|} \sum_{S \in \mathcal{S}} d_f(A \cdot S)\,,$$

where in general we define the "fuzzy intersection" $A \cdot B$ between fuzzy sets A and B as the fuzzy set with degree of membership function $P_{A \cdot B}(S) = P_A(S) P_B(S)$,

$S \in \mathcal{S}$. This average degree of fuzziness $d_f(A)$ has much in common with work of others [46].

9. Molecular configuration and symmetry

The shifted-related distances are particularly relevant in dealing with molecular electron densities. Such densities are (almost always) nonzero everywhere in three-dimensional Euclidean space $\mathcal{E}_3$. Thus though these densities have a common support, they decrease very rapidly (e.g., exponentially fast) at distances far from the molecular center for finite molecules. Then for two densities $\rho(\mathbf{r})$ and $\rho'(\mathbf{r})$ respectively decreasing asymptotically as $e^{-\alpha r}$ and $e^{-\alpha' r}$ with $\alpha \neq \alpha'$, an attempt to construct the scale comparators $c(\rho \nearrow \rho')$ would lead to unbounded ratios $e^{-(\alpha-\alpha')r}$, and consequent unbounded scale-related distances between ρ and ρ'. (In fact the situation is even "worse" for densities typically computationally realized based on Gaussian-orbital expansions, with asymptotic behavior of the form $\exp(-\alpha r^2.)$ One way to attempt to overcome such difficulties is to treat the scale-related distances comparing ρ and ρ' exceed some (small) positive tolerance $\mathcal{E}$ – though questions remain as to the choice of the value for $\mathcal{E}$ and as to whether the d-functions so realized are rigorously still distance functions. Another approach to the scale-related scheme is to consider simultaneous rescaling of densities both with regard to their intensity and to their spatial extent – that is, one considers rescalings of $\rho(\mathbf{r})$ to $k'\rho(k\mathbf{r})$, with k' a monotonic function of k taking the value 1 at the same time k does. Most simply $k' = k$ whence the problem may be viewed as rescalings much as in section 3 but now for a hypersurfaces $S[\rho]$ in 4-dimensional $\mathcal{E}_4$ with Cartesian coordinates x_0, x_1, x_2, x_3 of which x_1, x_2, x_3 identify to $\mathbf{r}$ of $\mathcal{E}_3$ and x_0 is such that $S[\rho]$ is defined by the region bound by $x_0 = 0$ and $x_0 = \rho(\mathbf{r})$. Another evidently quite trouble-free alternative is to use the shift-related distances. This then provides a fundamental way of comparing different molecular conformations or even different molecules.

Granted a distance function $d(\rho, \rho')$ between electron densities there follow natural measures of the *degree of dissymmetry* $s_-(\rho)$ of density ρ with respect to a given point group $\mathcal{G}$. Here $s_-(\rho)$ could be defined as the minimal distance from ρ to a second density ρ' of the requisite symmetry and number of electrons. Perhaps more simply, instead of seeking ρ' to minimize this distance one could take ρ' to be

$$\rho'(\mathbf{r}) = |\,\mathcal{G}\,|^{-1} \sum_{G \in \mathcal{G}} \rho(G\mathbf{r}),$$

where $G\mathbf{r}$ denotes the new position in $\mathcal{E}_3$ to which $\mathbf{r}$ is sent by the point group element $G \in \mathcal{G}$. The *degree of symmetry* $s_+(\rho)$ would then be defined in terms of a positive function of s_- monotonically decreasing from 1 at $s_- = 0$. E.g.,

$$s_+(\rho) = \exp\{s_-(\rho)\}.$$

Such a degree of symmetry differs from related ones previously proposed [20]. The "degree of symmetry adaption" in [19] was formulated for the characterization of wavefunctions, where pure but non-totally (i.e. nonidentity) symmetric irreducible representations would also be regarded to have no deviation from full symmetry adaption – in this case the present ideas presumably apply also, with shifted-related distances measured from the function under consideration to its different irreducible-symmetry projections.

In place of full electron densities similar ideas can be developed in terms of more compact characterizations of molecular structure. Such are the molecular surfaces of section 3 or the nuclear conformations of section 4. Or one might be interested in comparing functions (e.g., as electrostatic potential in [47]) on such molecular surfaces, whence the value of this function might be simply viewed to add an extra dimension into the mathematical objects under comparison.

10. Majorization

One fundamental poset for which shifting transformations are of relevance is the majorization poset. Here the members of $\wp$ are nonincreasing length-n sequences of nonnegative numbers, so that $A \in \wp$ is identified as

$$A = (a_1, a_2, \ldots, a_n) \quad \text{with } a_1 \geqslant a_2 \geqslant \ldots \geqslant a_n \geqslant 0 .$$

The partial ordering is such that $A \succeq B$ iff

$$\sum_{i=1}^{k} a_i \geqslant \sum_{i=1}^{k} b_i, \quad k = 1, 2, \ldots, n .$$

Sometimes the sum of the elements in the sequence is constrained to a fixed value, say N. A fundamental result [48] is that one sequence B can be obtained from another A (each with the same sum N) via a stochastic matrix transformation (as in Markov processes) iff $A \succeq B$. Sometimes too the elements are constrained to be integers, whence one obtains the so-called Young-Diagram lattice of partitions of N. Here evidently we might again shift sequences by adding a common scalar to every element of the sequence. In addition to this general poset's relation to probability distributions [7,22,23] it has arisen in chemistry in characterizing chirality [24] of molecules and in characterizing branching [25] of molecular graphs. In mathematics this poset dates back to the turn of the century when Muirhead [49] identified a "complete" set of isotonic functions.

Especially in the Young-diagram lattice case [7] there is a fundamental "dual" interpretation. In this case the sequences are represented diagrammatically: A has a diagram YD[A] with the ith row being of length a_i. See, e.g., fig. 4(a). We denote the length of the ith column by $\tilde{a}_i$ and obtain a dual sequence $\tilde{A}$. In a chemical context each row might correspond to particles (e.g., molecules) "identified" to the

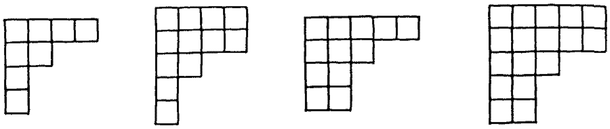

Fig. 4. As an example the Young diagram YD[4,2,1,1] is indicated in (a), and its 1-step (horizontal) shifted YD[5,3,2,2] indicated in (b). In (c) the 1-step (vertical) shifted YD[4,4,2,1,1] is shown, and in (d) is shown the symmetrically shifted YD[4,4,3,2,2].

same type while the columns correspond to particles "distinguished" to different types. Notably,

$$\tilde{A} \succeq \tilde{B} \quad \Leftrightarrow \quad B \succeq A$$

so that there is a duality (or complementarity) between "identification" and "distinction". Strangely it seems that in the bulk of the work on majorization these ideas are overlooked. However, with this duality a second shifting process is possible: A is "shifted" by k through the addition of k more rows to YD[A]. See fig. 4(c). In a duality symmetric approach one could also consider shifting in both horizontal and vertical directions.

11. General transformations

The general idea behind the present approach may be viewed as an extension of a distance idea generally applied on suitable equivalence-class partitioned sets. In this latter approach the distance is taken as the minimum number of specified *elementary* transformational steps needed to transform a member from one equivalence class to another. For example in considering the strings of letters from an alphabet, the minimum number of letter substitions needed to change one string into another is known [50] as the *Hamming distance* [51], of importance in linguistics, computer science, and coding, including genetic codes. Presumably too such distances could be of relevance for graph codes, appearing [52] as strings of binary digits. Extending the elementary transformations to include insertions and deletions in addition to substitions gives Levenshtein's [53] distance. The present poset-mediated results may be viewed to consider the distance d_i to correspond to the minimum number of elementary transformational "steps" to bring two considered objects of $\wp$ to be ordered in opposite senses - that is, for $A, B \in \wp$ we may let A' and B' be the minimally transformed members of $\wp$ via numbers α and β of elementary transformational steps such that $A' \succeq B$ and $B' \succeq A$ and then take the distance between A and B to be minimal

$$d(A,B) = \alpha + \beta\,.$$

To regain the Hamming–Levenstein approach the conditions $A' \succeq B$ and $B' \succeq A$

are replaced by $A' \approx B$ and $B' \approx A$. For the present partially ordered circumstance we may define $d(A \nearrow B) \equiv \alpha - \beta$ as a relative coordinate.

Besides our general examples of sections 2 and 8, the distance function of Kemeny [54] provides another example. This is on the set $\wp$ of subsets of a parent set S with the partial ordering $\succeq$ being set inclusion and the elementary transformations being addition of single elements of S to a subset. More precisely in Kemeny's language this distance between sets A and B is the order of the symmetric difference

$$A \ominus B \equiv \{a \in A \mid a \notin B\} \cup \{b \in B \mid b \notin A\}.$$

The orders of these two identified subsets in $A \ominus B$ may be differenced to obtain relative coordinates, analogous to the $d''(A \nearrow B)$ and $d(A \nearrow B)$ of section 8.

Yet another example involves the so-called [55] "chemical distance" between molecular structures. Here $\wp$ is a set of (e.g. molecular) graphs involving a common set of vertices, the (reverse) partial ordering $\preceq$ is taken as the subgraph relation, and each elementary tranformational step is taken to be the addition of an edge. This chemical distance also is described [53] as the minimum number of bond breaking/making steps or as the number of valence electrons redistributed during a reaction, and as such is related to a classical [56] "principle of minimum structural change" in the course of a chemical reaction. This same poset has also been emphasized [18,57] in making chemical "cluster expansions". Our comments concerning the possibility of relative coordinates evidently extend this earlier work.

Even more generally for an arbitrary discrete poset $\wp$ we might take the elementary transformational steps to be (increasing) steps between ordered elements such that there are no intermediates between the two elements.

Part of the extension of earlier sections also has to do with the treatment of continuous transformations (such as our dilitations or shiftings) instead of discrete ones. Thus continuous groups may be involved as noted at the end of section 3. But too sometimes there are alternative possible approaches – e.g., for closed subsets of Euclidean spaces $\mathcal{E}$ with measure function μ (such as area for the 2-dimensional case) one can identify a distance function

$$d(A, B) = \mu(A \ominus B)$$

analogous to that for finite sets in terms of $| A \ominus B |$. Another manner of using a "measure" to yield a distance function (on molecular conformations) is found in Mezey [28].

12. Multi-posets and periodic law

So far only a fraction of the potential applications of posets have been indicated. Even in elementary chemistry texts many "rules of thumb" are given which in effect make partial orderings of various chemico-physical properties (melting points, boiling points, electronegativities, solubilities, reactivities, etc.) For example, the ionization potentials of elements arranged in a suitable typical periodic chart generally decrease in proceeding down columns and in proceeding right-to-left across rows, so that while some pairs of elements have ionization potentials ordered by this rule, others pairs do not (and in particular a pair for which one traces a path from one element to the other by going leftward then upward in the periodic chart). Indeed, the periodic chart can be viewed as what we might call a *multi-poset*, where there are ordering links along both vertical and horizontal directions but orderings are to be in different directions (interchanging upward versus downward and/or leftward versus rightward) for various properties. Here too we allow for the possibility of simultaneous orderings in opposite directions giving rise to equivalences, so that the different vertical columns or horizontal rows may be different equivalence-class partitionings. An admirable periodic-chart presentation encoding such features is that essentially of Bayley, Thomsen, Margary and Bohr [58] as indicated in fig. 5. In addition sometimes one may have further orderings between *A* and *B*

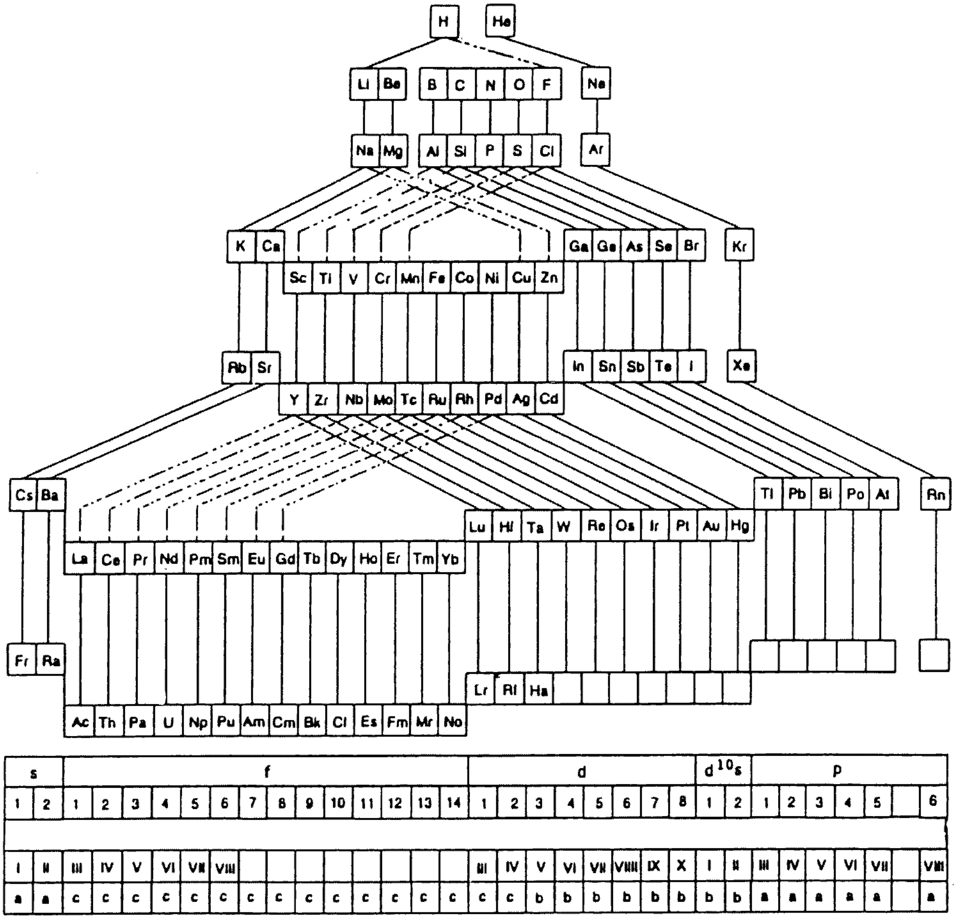

Fig. 5. A "poset-potent" representation of the standard periodic chart of the elements.

(or even *B* and *C*) columns as suggested by their slightly different levels in the figure. Thus somewhat as in Sanderson's fundamental book [3], Chemical Periodicity, may be viewed largely as an identification of the different partial orderings (including equivalences) realized by this multi-poset, along with a specification of which properties associated to the different posets. For instance, (first) ionization potentials decrease going downward along solid lines in fig. 5 while they increase going left-to-right along rows, with a break just before column IIIA. On the other hand Sanderson's [3] nonpolar covalent radii decrease in both these directions, now however with a break just before column VIIIA. Occasionally one finds exceptions to the desired orderings, but often they are associated to one of the "breaks" appearing in the form of the chart in fig. 5.

But in chemistry there are rules of thumb involving not just elements but sets of compounds too. Sometimes these rules of thumb for properties of compounds are even presented as "periodic tables" for a class of compounds – e.g., see, Randić's [59] "periodic table of the alkanes" (appearing on the cover of issue no. 9 of volume 69 of the Journal of Chemical Education), or see Dias' [60] "periodic table of benzenoids". Indeed Dias has suggested [61] that such periodic tables are in some way related to posets. In such cases one might likely have further multi-posets. A highly systematic approach to deal with molecules is simply to consider repeated Cartesian products of the periodic table, whence (for the case of the binary product describing diatomics) one might develop Cartesian product posets as indicated at the end of section 6. Hefferlin and co-workers [62] as well as Kafarov and co-workers [63] indeed take approaches of this sort. Thus, it seems that not only are posets manifested in a great diversity of manners in chemistry, but too they are very fundamental.

13. Conclusion

The ubiquity of posets in chemistry has been emphasized. A novel comparison scheme which leads to the introduction of metrics on posets $\wp$ has been defined, with special reference first to scaled posets and second to shifted posets, though extensions beyond these classes of posets have been considered too. Notably the results have been indicated to have potentially very wide applicability in chemistry, though presumably the range of applicability should extend to many other sciences as well. Hopefully the new descriptors and descriptions will find utility, the descriptors being in a few cases (especially in the earlier sections) being built up in a mathematically "canonical" way from related invariants (sometimes) already introduced for chemical applications. In the later sections (8–11) the ideas in particular cases seem often to correlate with some already introduced ideas (though often not so widely recognized in chemistry). The view taken here is quite wide, encompassing many such particular cases.

Acknowledgements

Research support from the Welch Foundation of Houston, Texas, is acknowledged. Though there have been helpful discussions with many, special acknowledgement to Prof. Ernst Ruch is appropriate.

Appendix

In this appendix we wish to establish theorem A′ of section 4, whenceafter theorems B′, C′ and D′ follow fairly straight forwardly. Here $\mathbf{R}(A \nearrow B)$, $\lambda(A \nearrow B)$ and $|A \nearrow B\rangle$ are abbreviated to $\mathbf{R}^+$, λ_+ and $\boldsymbol{\psi}^+$, while $\mathbf{R}(B \nearrow A)$, $\lambda(B \nearrow A)$ and $|B \nearrow A\rangle$ are abbreviated to $\mathbf{R}^-$, λ_- and $\boldsymbol{\psi}^-$. The components or elements of $\boldsymbol{\psi}^\sigma$ and $\mathbf{R}^\sigma$ are denoted $\tilde{\psi}_r^\sigma$ and $\tilde{R}_{rs}^\sigma$, for $\sigma = +$ or $-$.

As a first step towards proving theorem A′ we introduce new (soon to be used) vectors $\boldsymbol{\theta}^\sigma$, with components

$$\theta_r^\sigma \equiv \psi_r^\sigma / \xi_\sigma (\psi_r^\sigma)^2 \,,$$

where naturally $-\sigma$ denotes the opposite sign from σ. The factors ξ_σ are chosen to satisfy a "normalization" condition

$$\boldsymbol{\psi}^{\sigma\dagger} \boldsymbol{\theta}^\sigma = 1$$

with the superscript † indicating a transpose, and one thus has

$$\xi_\sigma = \sum_s \psi_s^{-\sigma} / \psi_s^\sigma \,.$$

The various divisions here by components of $\boldsymbol{\psi}^\sigma$ are all permitted since $\mathbf{R}^\sigma$ being a Frobenius–Perron matrix [30] gives maximum-eigenvalue eigenvectors with all components positive. Now recalling that $\boldsymbol{\psi}^\sigma$ is an eigenvector to $\mathbf{R}^\sigma$, we have

$$\lambda_+ \lambda_- = (\boldsymbol{\theta}^{+\dagger} \mathbf{R}^+ \boldsymbol{\psi}^+)(\boldsymbol{\psi}^{-\dagger} \mathbf{R}^- \boldsymbol{\theta}^-) = \sum_{rs} \theta_r^+ R_{rs}^+ \psi_s^+ \sum_{tu} \psi_t^- R_{tu}^- \theta_u^-$$

and via the Cauchy–Schwarz inequality

$$\lambda_+ \lambda_- \geqslant \left[\sum_{rs} (\theta_r^+ R_{rs}^+ \psi_s^+)^{1/2} (\psi_r^- R_{rs}^- \theta_s^-)^{1/2} \right]^2 . \tag{A.1}$$

But noting $R_{rs}^+ R_{rs}^- = 1$ and recalling the definitions of $\boldsymbol{\theta}^\sigma$ and ξ_σ, one obtains

$$\begin{aligned} \lambda_+ \lambda_- \geqslant \left[\sum_{rs} (\theta_r^+ \psi_r^-)^{1/2} (\psi_s^+ \theta_s^-)^{1/2} \right]^2 &= \left[\sum_r (\psi_r^- / \psi_r^+ \xi_+)^{1/2} \sum_s (\psi_s^+ / \psi_s^- \xi_-)^{1/2} \right]^2 . \\ &= \xi_+ \xi_- \,. \end{aligned} \tag{A.2}$$

Next via a second application of the Cauchy–Schwarz inequality

$$\xi_+\xi_- = \sum_r(\psi_r^+/\psi_r^-)\sum_s(\psi_s^-/\psi_s^+) \geqslant \left[\sum_r(\psi_r^+/\psi_r^-)^{1/2}\right]^2 = N^2$$

so that

$$\lambda_+\lambda_- \geqslant N^2 .$$

Now in our present notation $c_\sigma = \lambda^\sigma \boldsymbol{\psi}^{\sigma\dagger}\boldsymbol{\psi}^\sigma(\phi^\dagger\boldsymbol{\psi}^\sigma)^{-2}$, where $\boldsymbol{\phi}$ is the column-vector of all ones, whence

$$\mathbf{c}_+\mathbf{c}_- = \lambda_+\lambda_-(\boldsymbol{\theta}^{+\dagger}\boldsymbol{\theta}^+)(\boldsymbol{\theta}^{-\dagger}\boldsymbol{\theta}^-)(\boldsymbol{\phi}^\dagger\boldsymbol{\theta}^+)^{-2}(\boldsymbol{\phi}^\dagger\boldsymbol{\phi}^-)^{-2} \geqslant \lambda_+\lambda_-(\boldsymbol{\phi}^\dagger\boldsymbol{\phi})^{-2} , \qquad \text{(A.3)}$$

where yet again we have used the Schwarz inequality. But noting $\boldsymbol{\phi}^\dagger\boldsymbol{\phi} = N$ and combining this with the equation preceding (A.3) we have

$$c_+c_- \geqslant 1 ,$$

which is close to what we desired to prove for theorem A′.

It remains to check to see under what conditions it might happen that $c_+c_- = 1$. Indeed equality applies iff the set of Cauchy–Schwarz inequalities of eqs. (A.1), (A.2) and (A.3) each become equalities – such happening iff pairs of "vectors" involved become linearly dependent. For eq. (A.3) this condition is that there exist k_σ such that

$$\boldsymbol{\psi}^\sigma = k_\sigma\boldsymbol{\phi}$$

(where again $\boldsymbol{\phi}$ is the vector of all ones). Thence also from the first equations of this appendix

$$\boldsymbol{\theta}^\sigma = \{k_{-\sigma}/(k_\sigma)^2\xi_\sigma\}\boldsymbol{\phi} \quad \text{and} \quad \xi_\sigma = k_{-\sigma}/k_\sigma .$$

Next this form for $\boldsymbol{\theta}^\sigma$ leads to the satisfaction of equality in conjunction with (A.2). For eq. (A.1) the condition of equality implies that there is a $k \neq 0$ such that

$$\theta_r^+ R_{rs}\psi_s^+ = k\psi_r^- \frac{1}{R_{rs}}\theta_s^-$$

or

$$R_{rs}^2 = k(\psi_r^-/\theta_r^+)(\theta_s^-/\psi_s^+) .$$

But this implies that R_{rs} can be written in the form $R_{rs} = a_r a_s$, so that the eigensolution for λ_+ becomes

$$a_r(\mathbf{a}^\dagger\psi) = \sum_s a_r a_s\psi_s^+ = \lambda_+\psi_s^+ .$$

Thence a_r is independent of r and ψ_s^+ is independent of s. That is,

$$\mathbf{R} = a^2\mathbf{J},$$

where $\mathbf{J} = \phi\phi^\dagger$ is the matrix of all ones and $\psi^\sigma \sim \phi$, in agreement with the equality condition for (A.3). Thus the equality $c_+c_- = 1$ occurs iff the underlying matrices $\mathbf{D}(A)$ and $\mathbf{D}(B)$ are the same up to a scale factor (of a^2), i.e., equality occurs iff $A \approx B$.

References

[1] Articles in *Concepts and Applications of Molecular Similarity*, eds. M.A. Johnson and G.M. Maggiora (Wiley and, New York, 1990).
[2] P.G. Mezey, *Shape in Chemistry* (VCH, New York, 1993).
[3] R.T. Sanderson, *Chemical Periodicity* (Reinhold, New York, 1960).
[4] G. Birkhoff, *Lattice Theory* (Am. Math. Soc., Providence, RI, 1948).
[5] P. Crawley and R.P. Dilworth, *Algebraic Theory of Lattices* (Prentice-Hall, Englewood Cliffs, New Jersey, 1973).
[6] (a) C.S. Pierce, Am. Jour. 3 (1880) 15; 7 (1884) 180.
(b) E. Schröder, *Algebra der Logik* (Leipzig, 1890-5).
[7] E. Ruch, Theor. Chim. Acta 38 (1975) 167.
[8] See, e.g., J.P. Quirk and R. Saposnik, Rev. Econ. Stud. 29 (1962) 140.
[9] J.P. Barthélemy, Cl. Flament and B. Monjardet on p. 721 as well as K.P. Bogart on p. 760 of *Ordered Sets*, ed. I. Rival (Reidel, Dordrecht, Holland, 1982).
[10] E.C. Pielou, *Ecological Diversity* (Wiley, New York, 1975).
[11] See, e.g.,
(a) D'Arcy Thompson, *On Growth and Form* (Cambridge University Press, Cambridge, 1917).
(b) N. Rashevsky, Bull. Math. Biophys. 4 (1956) 31.
[12] W. Hennig, *Phylogenetic Systematics* (University of Illinois Press, Urbana, Illinois, 1966).
[13] B.-O. Küppers, *Molecular Theory of Evolution* (Springer, Berlin, 1985).
[14] G.C. Rota, Zeit. Wahr. Verw. Geb. 2 (1964) 340.
[15] See, e.g., L.M. Blumenthal, *Distance Geometry* (Chelsea, New York, 1970).
[16] P.G. Mezey, Int. J. Quant. Chem. 51 (1994) 255.
[17] M. Randić, A.F. Kleiner and L.M. DeAlba, J. Chem. Inf. Comp. Sci. 34 (1994) 277.
[18] D.J. Klein, Int. J. Quant. Chem. S20 (1986) 153.
[19] D.J. Klein and A.A. Cantu, Int. J. Quant. Chem. 8 (1974) 223.
[20] (a) P.G. Mezey, *New Theoretical Concepts for Understanding Organic Reactions*, eds. J. Bertran and I.G. Csizmada (Kluwer Academic Publ., Dordrecht, 1989) pp. 55–76.
(b) P.G. Mezey and J. Maruani, Mol. Phys. 69 (1990) 97.
(c) A.B. Buda, T. Aufder Deyde and K. Mislow, Angew. Chemie Int. Edn. Engl. 31 (1992) 989.
(d) H. Zabrodsky, S. Peleg and D. Avnir, J. Am. Chem. Soc. 114 (1992) 7843.
(e) N. Weinberg and K. Mislow, J. Math. Chem 14 (1993) 427.
(f) G. Gilat, J. Math. Chem. 15 (1994) 197.
(g) H. Zabrodsky and D. Avnir, J. Am. Chem. Soc. 117 (1995) 462.
(h) N. Weinberg and K. Mislow, J. Math. Chem 17 (1995) 35.
[21] (a) I. Gutman and M. Randić, Chem. Phys. Lett. 47 (1977) 15.
(b) E. Ruch and I. Gutman, J. Comb. Inf. Sys. Sci. 4 (1979) 285.
[22] (a) E. Ruch and A. Mead, Theor. Chim. Acta 41 (1976) 95.
(b) E. Ruch, Acta Appl. Math. 30 (1993) 67.

[23] P.M. Alberti and A. Uhlmann, *Stochasticity and Partial Order* (VEB Deutscher Verlag, Berlin, 1981).
[24] E. Ruch and A. Schönhofer, Theor. Chim. Acta 19 (1970) 225.
[25] (a) I. Gutman, B. Ruscić, N. Trinajstić and C.F. Wilcox, Jr., J. Chem. Phys. 62 (1975) 3399.
(b) D. Bonchev and N. Trinajstić, J. Chem. Phys. 67 (1977) 4517 and Int. J. Quant. Chem. S12 (1978) 293.
(c) S. El-Basil and M. Randić, Adv. Quant. Chem. 24 (1992) 239.
[26] (a) S. Lele, Am. J. Phys. Anthropol. 85 (1991) 407.
(b) S. Lele and J.T. Richtsmeier, Am. J. Phys. Anthropol. 87 (1992) 49.
[27] See, e.g., for many advanced calculus texts, ref. [28], or R.F.W. Bader, T.T. Nguyen-Dang and Y. Tal, J. Chem. Phys. 90 (1979) 4316.
[28] P.G. Mezey, J. Chem. Phys. 78 (1983) 6182.
[29] (a) H. Frauenfelder, F. Parak and R.D. Young, Ann. Rev. Biophys. Chem. 17 (1988) 451.
(b) T. Noguti and N. Go, Proteins 5 (1989) 125.
(c) M. Vasquez, G. Nemethy and H.Z. Scheraga, Chem. Rev. 94 (1994) 2183.
[30] (a) G.A. Arteca and P.G. Mezey, J. Comp. Chem 9 (1988) 554.
(b) G.M. Crippen, A.S. Smellie and J.W. Ping, J. Chem. Inf. Comp. Sci. 28 (1988) 125.
(c) M. Randic, B. Jerman-Blazic and N. Trinajstic, Comput. Chem 14 (1990) 237.
(d) P.G. Mezey, J. Chem. Inf. Comp. Sci. 32 (1992) 650.
[31] F.R. Gantmacher, *Matrix Theory* (Chelsea, New York, 1959).
[32] F. Buckley and F. Harary, *Distance in Graphs* (Addison-Wesley, Reading, Massachusetts, 1989)
[33] D.J. Klein and M. Randić, J. Math. Chem. 12 (1993) 81.
[34] P.G. deGennes, *Scaling Concepts in Polymer Physics* (Cornell University Press, Ithaca, New York, 1979).
[35] (a) D.J. Klein and W.A. Seitz, *Chemical Applications of Topology and Graph Theory*, ed. R.B. King (Elsevier, Amsterdam, 1983) pp. 430–445.
(b) S.I. Kuchanov, S.V. Korolev and S.V. Panyukov, Adv. in Chem. Phys. 72 (1988) 115.
[36] (a) B.E. Eichinger, Macromol. 13 (1980) 1.
(b) G. Wei and B.E. Eichinger, Macromol. 22 (1989) 3429.
(c) M. Kunz, Collect. Czech. Chem. Commun. 54 (1989) 2148.
(d) M. Kunz, J. Math. Chem. 13 (1993) 145.
[37] (a) R. Kadison, Proc. Am. Math. Soc. 2 (1951) 505.
(b) S. Sherman, Am. J. Math. 73 (1951) 227.
[38] R. Bhatia, Proc. Am. Math. Soc 96 (1986) 41.
[39] L.A. Zadeh, Inf. Control 8 (1965) 338.
[40] D.J. Klein, J. Chem. Ed. 69 (1992) 691.
[41] See, e.g.,
(a) *Aromaticity, Pseudo-Aromaticity, Anti-Aromaticity*, eds. E.D. Bergmann and B. Pullman (Israel Acad. Sci., Jerusalem, 1971).
(b) P.J. Garratt, Aromaticity (McGraw-Hill, London, 1971; Wiley, New York, 1986).
(c) D. Lewis and D. Peters, *Facts and Theories of Aromaticity* (Macmillan, London, 1975).
(d) D. Lloyd, *Non-benzenoid Conjugated Carbocyclic Compounds* (Elsevier, Amsterdam, 1984).
(e) B.Ya. Simkin, V.I. Minkin and M.N. Glukhovtsev, *Advances in Heterocyclic Chemistry 56*, ed. A.R. Katritzky (Academic Press, New York, 1993) pp. 303–428.
[42] G. Birkhoff, Ann. Math. 33 (1932) 329.
[43] See, e.g., S.T. Rachev, *Probability Metrics and the Stability of Stochastic Models* (Wiley, New York, 1991).
[44] A.N. Kolgomorov, Vestnik M.G.U. 10 (1953) 28.
[45] D.J. Fivel, Phys. Rev. A50 (1994) 2108.

[46] (a) A. DeLuca and S. Termini, Inf. Control 20 (1972) 301.
(b) J. Knopfmacher, J. Math. Anal. Appl. 49 (1975) 529.
[47] See, e.g.,
(a) P.L. Chau and P.M. Dean, J. Mol. Graphics 5 (1987) 97.
(b) J. Gasteiger, X. Li, C. Rudolph, J. Sadauski and J. Zupan, J. Am. Chem. Soc. 116 (1994) 4608.
[48] G.H. Hardy, J.E. Littlewood and G. Polya, Messenger Math. 58 (1929) 145.
[49] R.F. Muirhead, Proc. Edin. Math. Soc. 21 (1903) 144.
[50] See, e.g., J.B. Kruskal, SIAM Rev. 25 (1983) 201, and *Theory and Practice of Sequence Comparison*, eds. D. Sankoff and J. Kruskal (Addison-Wesley, Reading, Massachusetts, 1983) chap. 1.
[51] R.W. Hamming, Bell Syst. Tech. J. 29 (1950) 147.
[52] See, e.g., W.C. Herndon and S.H. Bertz, J. Comp. Chem. 8 (1987) 367.
[53] (a) V.I. Levenshtein, Dok. Akad. Nauk SSSR 163 (1965) 845 [Engl. transl.: Cybern. Contr. Theory 10 (1966) 707].
(b) V.I. Levenshtein, Problemy Peredachi Informatsii 7 (1971) 30.
[54] J.G. Kemeny, Daedalus 88 (1959) 577.
[55] (a) C. Jochum, J. Gasteiger, I. Ugi and J. Dugundji, Zeit. Naturforsch. b37 (1982) 1205.
(b) J. Dugundji and I. Ugi, Top. Curr. Chem. 39 (1973) 19.
(c) M. Johnson, *Graph Theory and Its Applications*, eds. Y. Alavi et al. (Wiley, New York, 1985) 457–469.
(d) I. Ugi et al., in ref. [1], pp. 239–288.
[56] (a) H. Kolbe, Liebigs Ann. Chem. 75 (1850) 211.
(b) F.O. Rice and E. Teller, J. Chem. Phys. 6 (1938) 489.
(c) R. Ponec, Zeit. Phys. Chem. 268 (1987) 1180.
[57] (a) J.W. Essam, J.W. Kennedy, M. Gordon and P. Whittle, J. Chem. Soc. Faraday II 73 (1977) 1289.
(b) J.W. Kennedy and M. Gordon, Ann. N.Y. Acad. Sci. 319 (1979) 331.
[58] (a) T. Bayley, Phil. Mag. Sec 5, 13 (1882) 26.
(b) J. Thomsen, Zeit. Anorg. Chem. 9 (1895) 190.
(c) I.D. Margary, Phil. Mag. Sec. 6, 42 (1921) 287.
(d) N. Bohr, Nature 112 (1923) 29.
[59] M. Randić, J. Chem. Ed. 69 (1992) 713.
[60] J.R. Dias, J. Chem. Inf. Comp. Sci 22 (1982) 15.
[61] J.R. Dias, J. Math. Chem 4 (1990) 17.
[62] (a) R. Hefferlin, R. Campbell, D. Gimbell, H. Kuhlman and T. Cayton, J. Quant. Spec. Rad. Trans. 21 (1979) 315.
(b) R. Hefferlin and H. Kuhlman, J. Quant. Spec. Rad. Trans 24 (1980) 379.
(c) R. Hefferlin, G.V. Zhuvkin, K.E. Caviness and P.J. Duerksen, J. Quant. Spec. Rad. Trans. 32 (1984) 257.
[63] V.V. Kafarov, I.N. Dorokhov, V.N. Vetokhin and L.P. Volkov, Dokl. Akad. Nauk SSSR 292 (1987) 118, and earlier references therein.