Journal of Mathematical Chemistry 13(1993)145–151

On topological and geometrical distance matrices

Milan Kunz
Research Institute of Macromolecular Chemistry, Chemopetrol, 65649 Brno, Czechoslovakia

Received 20 January 1992; revised 8 April 1992

Differences between topological and geometrical distance matrices are examined. Some examples of geometrical distances when graphs are embedded in spaces of different dimensions are given. Relations of topological distance matrices to other graph matrices are shown. The topological distance matrices are defined in the Hilbert space and their elements are distances through the graph lattices.

1. Introduction

Recently, Trinajstić and coworkers [1–5] studied properties of geometrical distance matrices of chemical molecules. They have calculated the geometrical equivalent of the Wiener index and found that it correlates even better with some physical properties of alkanes than the topological Wiener index. Three-dimensional molecular descriptors have been introduced by other authors, too [6–8].

Trinajstić and coworkers have designated the topological distance matrices as two-dimensional to distinguish them from new three-dimensional geometrical ones. Once again, a question was raised concerning the dimensionality of graphs: should graphs be considered as dimensionless objects [9], one-dimensional objects [10], or objects in Hilbert space.

2. Topological and geometrical distances

The topological distance matrix $\mathbf{D}_t$ of the complete graph is identical to its adjacency matrix A. We can also interpret it as its geometrical distance matrix $\mathbf{D}_g$, if we suppose that n vertices of the complete graph are on the ends of n orthogonal unit vectors $\boldsymbol{e}_j$ of the ideal Euclidean space of dimension n [11,12]. The distances between all vertices are identical with the lengths of arcs given by the differences $(\boldsymbol{e}_j - \boldsymbol{e}_i)$ of two unit vectors. The Euclidean length is $2^{1/2}$. Its squares appear in the distance matrices as the sums of elements $(d_{ij} + d_{ji})$.

The densest possible packing of a graph in a space of sufficient dimensionality is obtained if n vertices are on the ends of n unit vectors $\boldsymbol{e}_j$. The vertices form the

plane orthogonal to the n-dimensional diagonal unit vector I_n. (Unit vector row J^T and column J are not applicable.) An n-dimensional plane is a body in $(n-1)$-dimensional space. Therefore, in three-dimensional space, four vertices can form the regular tetrahedron which is identical with the complete graph K_4. Because the geometrical distances need not be changed if some arcs are missing, unless we postulated some forces making incomplete graphs less compact, all geometrical distance matrices of graphs with four vertices could be identical with the distance matrix of K_4. Since we know that all topological distance matrices of these graphs are different, it is clear that the topological distances are not always identical with the geometrical distances of the densest possible configuration of graphs.

To clarify the difference between topological and geometrical distances, the notion of the phase space can be adopted. To characterize n points in k-dimensional space, we need n vectors determining their positions. If $k \geq n$, all n vectors can be orthogonal, if $k = 1$, all these vectors must be collinear; nevertheless, a system of n vectors remains always isomorphic with a system of n orthogonal vectors. Algebraic and combinatorial properties do not depend on the definition of unit vectors. If the vectors are not orthogonal and have unequal lengths, the topology of the space is only deformed, but its symmetry lattice remains unchanged. This principle is basic for all applications of graph theory: any drawing of a graph is isomorphic with its canonical form and thus with its ideal form.

A graph has m arcs or edges characterizing m binary relations from $\binom{n}{2}$ possible relations between n vertices. It is thus an mn-dimensional vector even if each vertex is characterized by just one parameter.

If we consider a linear chain in the form of a straight rod, then its geometrical distance matrix will be identical up to a scale factor with its topological distance matrix in all dimensions. However, if the chain forms a path along edges of the n-dimensional simplex, then its geometrical distance matrix will be identical to the topological distance matrix of the complete graph.

As an example demonstrating that the topological distance matrices do not correspond to two-dimensional objects, the stars S_n, each with the root in the center of a circle on which circumference $(n-1)$ leaves are spread, can be used. Because S_3 is identical to L_3, its topological and geometrical distance matrices can be identical. S_4 can have equal distances between three leaves, but the ratio between geometrical distances are smaller than the corresponding ratio between the radius and the diameter of the circle. Unequal geometrical distances appear between leaves of stars with more vertices.

From given examples, it is thus clear that the topological distances are independent of the graph configuration. They are distances going through the graph lattice. They count edges or arcs connecting vertices. Such distances are also known as chemical distances [13].

Even if we draw graphs as animals on a lattice and define the distances between vertices of the graph as the shortest path on the lattice, we find that such distances will soon be different on branched graphs from topological ones.

All connected graphs with more than three vertices drawn in two dimensions are distorted, except linear chains. We can, of course, draw them, measure distorted distances and construct special two-dimensional distance matrices. This approach is similar to the construction of distance matrices of graphs with multiple bonds [14, 15], heteroatoms [16] or graphs of arbitrary complexity with weighted vertices and edges [17]. The ideal Euclidean space is squeezed by such procedures. However, the difference between distances through bonds and through space remains preserved.

3. Some properties of topological distance matrices

The results obtained with topological and geometrical Wiener indices are related to distances measured through the lattice and through the space, respectively. However, the topological distance matrices are connected with other graph matrices and these relations have profound consequences on our understanding of phase space properties [18]:

(A) The topological distance matrices of trees and simple cycles can be transformed to the topological distance matrices of corresponding line bond graphs [19,20]

$$\mathbf{S}\mathbf{D}_t\mathbf{S}^T + \mathbf{G}\mathbf{D}_t\mathbf{G}^T = 4\mathbf{D}_B, \tag{1}$$

where **S** is the incidence matrix of a tree or a simple cycle whose rows are arcs $ij = (\boldsymbol{e}_j - \boldsymbol{e}_i)$ and **G** is the incidence matrix of the identical unoriented tree or cycle. Here, the rows are edges $ij = (\boldsymbol{e}_i + \boldsymbol{e}_j)$. The vector representing the edge ij is orthogonal to the vector of the arc ij. $\mathbf{D}_B$ in (1) is the topological distance matrix of the corresponding line graph. $\mathbf{S}^T$ is the transposed matrix.

It is unlikely that bonding electrons can be localized [21], so there seem to be no geometrical distance matrices of the line graphs which could be compared with this formal equivalence of topological distances in trees and their line graphs. $\mathbf{D}_t$ and $\mathbf{D}_B$ of cycles are identical.

(B) The topological distance matrices of some graphs are inverses of their perturbed Kirchhoff matrices $\mathbf{S}^T\mathbf{S}$. The perturbation is smallest for linear chains,

$$\mathbf{D}_t(L_n) = \tfrac{1}{2}\left[\mathbf{X} - \mathbf{S}^T\mathbf{S}\right]^{-1},$$

where the elements of the matrix **X** are $x_{11} = x_{1n} = x_{n1} = x_{nn} = (n-1)^{-1}$, $x_{ij} = 0$ otherwise. The topological distance matrix of the infinite chain is proportional to the inverse of its practically unperturbed Kirchhoff matrix [22].

(C) The topological distance matrices of trees are formal equivalents of the inverse matrices $[\mathbf{S}\mathbf{S}^T]^{-1}$ and of the generalized inverses **E** of $\mathbf{S}^T\mathbf{S}$ [23,24],

$$\left[\mathbf{S}\mathbf{S}^T\right]^{-1} = (1/n)\,\mathbf{W}^T\mathbf{W} = -\tfrac{1}{2}\,\mathbf{S}\mathbf{D}_t\mathbf{S}^{-1} = (1/n)\,\mathbf{S}\mathbf{E}\mathbf{S}^{-1}, \tag{2}$$

where **W** is the walk matrix [18]. The elements of a walk matrix are arcs of a tree, $w_{ij} = \pm 1$, if the arc j is in the walk i and $w_{ij} = 0$, if the arc j is not present in the walk i. The signs are determined according to the orientations of arcs in walks.

E is the Eichinger matrix defined as the sum of inverses of Kirchhoff submatrices,

$$\mathbf{E} = \sum_{j=1}^{n} (\delta_j \mathbf{S}^{\mathrm{T}}\mathbf{S})^{-1}.$$

Here, $(\delta_j \mathbf{S}^{\mathrm{T}}\mathbf{S})$ is the Kirchhoff matrix with the jth row and column deleted. In the case of trees, the elements of Eichinger matrices are sums of distances. The Eichinger matrices are generalized inverses. When multiplied with the corresponding Kirchhoff matrix, they give the Kirchhoff matrix of the complete graph K: $\mathbf{S}^{\mathrm{T}}\mathbf{S}\mathbf{E} = \mathbf{S}_K^{\mathrm{T}}\mathbf{S}_K$.

The traces of both quadratic forms $\mathbf{W}^{\mathrm{T}}\mathbf{W}$ and $\mathbf{W}\mathbf{W}^{\mathrm{T}}$ are equal to the Wiener index. The trace of the Eichinger matrix **E** of a tree is twice the Wiener index. It is simultaneously the sum of elements of the topological distance matrix.

The relations between different graph matrices are not accidental, but they reflect hidden properties of the Hilbert space.

4. Distances as inverse elements

Quantum mechanics considers molecules as objects in Hilbert space, which elements are finite sums of infinitely many quadratic elements $\sum_{j=1}^{\infty} x_j^2$. According to the axioms of linear algebra, to each element of a space must exist its inverse. There are two kinds of inverse elements: additive $\boldsymbol{a} + \boldsymbol{b} = \boldsymbol{0}$, and multiplicative. The problem is how to define inverse elements to zero elements, e.g. inverse matrices to singular ones.

The elements of the topological distance matrices are interpreted straightforwardly as numbers of arcs in walks between corresponding vertices and there appears to be nothing in common with some quadratics. However, at trees which form the base of the graph space, another insight can be reached.

We have shown three different matrices connected with the Wiener index: $\mathbf{D}_t$, **E** and $\mathbf{W}^{\mathrm{T}}\mathbf{W}$, which are inverses of the quadratic forms of the incidence matrix **S**. In fact, there is another quadratic form $\mathbf{W}\mathbf{W}^{\mathrm{T}}$ in which diagonal distances d_{ij} appear explicitly. This diagonal can be transformed into the distance matrix $\mathbf{D}_t$ exactly as the $\binom{n}{2}$-dimensional diagonal matrix Δ of incidences of arcs in a graph $\Delta_i = 1$ if the arc i is present in the graph, and a graph $\Delta_i = 0$ if the arc i is absent.

The diagonal matrices are, according to Eichinger [25], framed by the incidence matrices of the complete graph $\mathbf{S}_K$,

$$\mathbf{S}^{\mathrm{T}}\mathbf{S} = \mathbf{S}_K^{\mathrm{T}}\Delta\mathbf{S}_K. \tag{3}$$

The incidence matrix of the complete graph can be, in its turn, decomposed into the incidence matrix of the linear chain $\mathbf{S}_L$ and its walk matrix $\mathbf{W}_L$: $\mathbf{S}_K = \mathbf{W}_L\mathbf{S}_L$.

In the walk matrix of the linear chain, unit elements are uninterrupted by zeroes. Such matrices are known as Petrie matrices. Gordon and Wilkinson have proved that only Petrie matrices corresponding to trees are non-singular [16]. The walk matrix of the linear chain is the Petrie matrix of the complete graph and in the Petrie matrix of a tree, only $(n-1)$ rows are chosen; they permute the incidence matrix of the linear chain as for the star S_4 rooted in v_1,

$$\begin{array}{cc} & \left\| \begin{matrix} -1 & 1 & & \\ & -1 & 1 & \\ & & -1 & 1 \end{matrix} \right\|^{S(L_4)} \\ \hline \left\| \begin{matrix} 1 & & \\ 1 & 1 & \\ 1 & 1 & 1 \end{matrix} \right\| & \left\| \begin{matrix} -1 & 1 & & \\ -1 & & 1 & \\ -1 & & & 1 \end{matrix} \right\|^{S(S_4)} \end{array}$$

Similarly, if we replace the diagonal matrix Δ in product (3) by the diagonal elements of the quadratic form $\mathbf{WW}^{\mathrm{T}}$, we obtain the matrix $(\mathbf{Q}-\mathbf{D})$, whose diagonal elements $\mathbf{Q}$ are identical to the diagonal elements of the quadratic form $\mathbf{W}^{\mathrm{T}}\mathbf{W}$ and row or column sums of the matrix $\mathbf{D}$, and the matrix of off-diagonal elements $-\mathbf{D}_t$. This result can be compared with the corresponding Kirchhoff matrix $\mathbf{S}^{\mathrm{T}}\mathbf{S} = (\mathbf{V}-\mathbf{A})$, where $\mathbf{V}$ is the diagonal matrix of vertex degrees v_j.

The topological distance matrix is formally equivalent to the adjacency matrix $\mathbf{A}$ in the frame $\mathbf{S}_K^{\mathrm{T}}(.)\mathbf{S}_K$. These elements are corresponding

$$\begin{aligned} a_{ij} &= 1, \quad d_{ij} = 1, \\ a_{ij} &= 0, \quad d_{ij} = 2 \quad \text{until } (n-1) \\ \text{or} \quad & d_{ij} = \infty \text{ in disconnected graphs.} \end{aligned}$$

The distance $(n-1)$ is reached only in linear chains and the infinite distance is put between two components of disconnedted graphs.

5. Discussion

When we consider two kinds of distance matrices, topological and geometrical, we must ask which of them is primary and which is consequent, even if both lead to similar results. We can construct three-dimensional distances directly from the assumed configurations of molecules as Trinajstić and others did [1–5,27,28]. We must know bond angles and bond lengths, and there are as many geometrical distance matrices as there are possible configurations of molecules. The geometrical distance matrix is not an invariant which can identify a molecule.

Alternatively, we can follow techniques used in statistical mechanics of polymer chains and use multiple integrals in phase space. There are evaluated eigenvalues of Kirchhoff matrices and their inverses (known there as Zimm and Rouse matrices, respectively), characterizing properties of stochastic configurations of polymer chains as Gaussian molecules [10,29]. The eigenvalues of the Rouse matrix, which sum is equal to the Wiener index, are interpreted as relaxation times of molecules. This makes sense, because distances need time to be passed. However, from where has this time appeared?

After obtaining eigenvalues in n-dimensional space, the gyration tensor in the three-dimensonal space is calculated [30]. This method seems to be rigorous. Accidentally, it is possible to obtain only k nonzero eigenvalues of quadratic form from a kn-dimensional matrix of positions of n points, but both distance matrices are non-singular and we can interpret them only in connection with n eigenvectors.

It is difficult to convince old believers. Since Euclid, we have lived in three-dimensional space. An observer is in the center of our perspective and our abilities to imagine spaces of higher dimensions are undeveloped. A couple of centuries ago, the three-body problem appeared. Physicists began to study the phase space as a $3n$-dimensional space, but 3 appeared always to be more important than n, although n is usually much greater than 3. It is not essential how many dimensions we need for the description of all properties of one elementary particle, because more difficulties arise from the fact that we need to describe simultaneously properties and relations of too many particles.

The enigma of the dimensionality of graphs reminds one of old philosophical problems suggesting that we see shadows of ideal things only. Are graphs dimensionless objects which are forced into a three-dimensional Procustian bed or multi-dimensional objects which are squeezed into it? Once again we shall ask, together with Gordon [10], how to apply Occam's razor to the relation of the Rieman metrics to the graph metrics.

The circumstances that the Wiener index can be interpreted as the sum of inverse eigenvalues of the quadratic forms of the incidence matrices of trees [12,18] and that both positive and inverse eigenvalues correspond to macroscopic properties of alkanes calls by its plainness for the multi-dimensionality of graphs and their matrices.

References

[1] B. Bogdanov, S. Nikolić and N. Trinajstić, J. Math. Chem. 3(1989)299.
[2] M. Randić, B. Jerman-Blažić and N. Trinajstić, Comput. Chem. 14(1990)237.
[3] N. Bošnjak, Z. Mihalić and N. Trinajstić, J. Chromatogr. 540(1991)430.
[4] S. Nikolić, N. Trinajstić and Z. Mihalić, Chem. Phys. Lett. 179(1991)21.
[5] Z. Mihalić and N. Trinajstić, J. Mol. Struct. (THEOCHEM) 78(1991)65.
[6] P.G. Mezey, J. Math. Chem. 2(1988)299.
[7] K. Balasubramanian, Chem. Phys. Lett. 169(1990)224.
[8] S.S. Tratch, R.O, Devdariani and N.S. Zefirov, Zh. Org. Khim. 26(1990)921.

[9] N.J. Turro, Angew. Chem. 98(1986)872.
[10] M. Gordon, Polymer 20(1979)1349.
[11] M. Kunz, Chem. průmysl 33(1983)542.
[12] M. Kunz, in: *Problems in Quantum Physics II*, Gdansk 1989, ed. J. Mizerski, A. Posewnik, J. Pykacz and M. Żukovski (World Scientific, Singapore, 1990), p. 377.
[13] V. Kvasnička, J. Pospíchal, J. Koča, E. Hladká and L. Matyska, Chem. Listy 84(1990)449.
[14] J. Lukovits, Quant. Struct. Act. Relat. 9(1990)227.
[15] J. Lukovits, Rep. Mol. Theory 1(1990)127.
[16] P. Tvarůžek and J. Komenda, MATCH 24(1989)317.
[17] P. Senn, Comput. Chem. 12(1988)219.
[18] M. Kunz, Coll. Czech. Chem. Commun. 54(1989)2148.
[19] S.H. Bertz, J. Chem. Soc. Chem. Commun. (1981) 819.
[20] M. Kunz, to be published.
[21] S. Golden, J. Mol. Struct. 223(1990)93.
[22] M. Kunz, MATCH, in press.
[23] M. Kunz, in: *Proceedings of Chemometrics II*, ed. J. Havel and M. Holík (Masaryk University, Brno, 1990) p. 142.
[24] M. Kunz, J. Math. Chem. 9(1992)297.
[25] B.E. Eichinger, J. Polymer Sci. Polymer Symp. 54(1974)127.
[26] M. Gordon and E.M. Wilkinson, Pacific J. Math. 51(1974)74.
[27] B. Alsberg and K. Esbensen, Quant. Struct. Act. Relat. 8(1989)218.
[28] A.V. Belik and V.A. Potemkin, Dokl. Akad. Nauk SSSR 312(1990)377.
[29] B.E. Eichinger, Macromolecules 13(1980)1.
[30] G. Wei and B.E. Eichinger, Macromolecules 22(1989)3429.