communications in mathematical
match
and in computer chemistry

no. 37, March 1998

ISSN 0340-6253

MATCDY (37) 195-201 (1998)

CORRECTING THE DEFINITION OF CLUJ MATRICES

Ivan GUTMAN[a] and Mircea V. DIUDEA[b]

[a] *Faculty of Science, University of Kragujevac, P. O. Box 60, YU-34000 Kragujevac, Yugoslavia*

[b] *Department of Chemistry, "Babes-Bolyai" University, R-3400 Cluj, Romania*

(received: April 1997)

Abstract

Cluj matrices were recently conceived as a kind of generalization of the Wiener-matrix concept to cycle containing molecular graphs. Correcting the approach to the Cluj matrices, put forward in three recent papers [1],[2] and [3], we define them in a rigorous and generally applicable manner.

Introduction

The aim of this paper is to point out some flaws in the previously given [1] [3] definition of the Cluj matrices and to offer a way out of the difficulties. In order to see what the problem is and to eventually amend it we need some preparations.

In the pioneering work of Wiener [4] one finds a formula for the calculation of (what nowadays is called) the *Wiener index*:

$$W(G) = \sum_{c} N_{i,(i,j)} \, N_{j,(i,j)} \, . \qquad (1)$$

Recall that $W(G)$ is the sum of distances between all pairs of vertices of the molecular graph G.

In formula (1) $e = (i, j)$ is the edge of the molecular graph G, connecting the vertices i and j, and the summation goes over all edges of G. The quantities $N_{i,(i,j)}$ and $N_{j,(i,j)}$ count the vertices on the two sides of the edge e. In particular, $N_{i,(i,j)}$ is the number of vertices lying closer to vertex i than to vertex j.

Relation (1) holds only for trees. It was discovered in 1947. Almost half a century later, in 1993, Randić [5] observed that the product $N_{i,(i,j)}\, N_{j,(i,j)}$, appearing on the right-hand side of Eq. (1) can be viewed as the (ij)-entry of a matrix $\mathcal{W}$. Randić named $\mathcal{W}$ the *Wiener matrix* [6].

In Eq. (1) the vertices i and j are assumed to be adjacent. In the case when i and j are distinct, but not adjacent vertices, then the (ij)-entry of the Wiener matrix is defined also as [5]

$$[\mathcal{W}]_{ij} = N_{i,(i,j)}\, N_{j,(i,j)} \tag{2}$$

where now (i, j) denotes the path connecting the vertices i and j; the quantities $N_{i,(i,j)}$ and $N_{j,(i,j)}$ count the vertices on the two sides of the path (i, j). This definition includes also the case of adjacent i and j, since e can be understood as a path of length one.

The diagonal elements of the Wiener matrix $\mathcal{W}$ are chosen to be zero.

The main chemical applications of the Wiener matrix [5], [7]-[9] go via the so called *hyper Wiener index* which is just the half-sum of the elements of $\mathcal{W}$.

The original Randić definition of the Wiener matrix works well for trees thanks to the fact that in such graphs every pair of vertices is connected by a unique path. For cycle-containing graphs one encounters serious difficulties. The Cluj matrix-concept is an attempt to extend the Wiener matrices to cycle-containing molecular graphs.

The numbers $N_{i,(i,j)}$ and $N_{j,(i,j)}$ provide the basis for the definition of the Wiener matrix and the hyper-Wiener index. For trees their structural interpretation — as stated above — is perfectly clear and unequivocal. In order to specify $N_{i,(i,j)}$ and $N_{j,(i,j)}$ in a more formal manner (which will be needed in the subsequent considerations) we proceed as follows.

The vertex set of the graph G is denoted by $\mathbf{V}(G)$. If G is a tree, then between any two of its vertices there is a unique path. Therefore the path connecting the vertices x and

y may be written as (x, y). If (x, y) and (x', y') are two paths of G, then their intersection $(x, y) \cap (x', y')$ is the set of those vertices of G which belong to both (x, y) and (x', y'). The length of a path p is denoted by $|p|$.

Now, if i and j are any two vertices of the tree G, then

$$N_{i,(i,j)} = |\mathcal{N}_{ij}| \tag{3}$$

where $\mathcal{N}_{ij}$ is a set of vertices, given by

$$\mathcal{N}_{ij} = \{v | v \in \mathbf{V}(G) \; ; \; d(i, v|G) < d(j, v|G) \; ; \; (i, v) \cap (i, j) = \{i\}\} \; . \tag{4}$$

Analogously,

$$N_{j,(i,j)} = |\mathcal{N}_{ji}|$$

$$\mathcal{N}_{ji} = \{v | v \in \mathbf{V}(G) \; ; \; d(j, v|G) < d(i, v|G) \; ; \; (j, v) \cap (i, j) = \{j\}\} \; .$$

These seemingly awkward, yet mathematically correct, expressions for the numbers $N_{i,(i,j)}$ and $N_{j,(i,j)}$ are the starting point for the definition of Cluj matrices.

Cluj Matrices

Suppose the graph G considered has n vertices. Provided the quantities $N_{i,(i,j)}$ and $N_{j,(i,j)}$ are known for all $i, j = 1, 2, \ldots, n$ one defines the *unsymmetric Cluj matrix* $\mathcal{C}_{Ju}$ and the two *symmetric Cluj matrices* $\mathcal{C}_{Jp}$ and $\mathcal{C}_{Je}$ as follows [1] [3],[9] [12]:

$$\mathcal{C}_{Ju} = \|[\mathcal{C}_{Ju}]_{ij}\| \quad ; \quad \mathcal{C}_{Jp} = \|[\mathcal{C}_{Jp}]_{ij}\| \quad ; \quad \mathcal{C}_{Je} = \|[\mathcal{C}_{Je}]_{ij}\|$$

where

$$[\mathcal{C}_{Ju}]_{ij} = N_{i,(i,j)}$$

$$[\mathcal{C}_{Jp}]_{ij} = N_{i,(i,j)} \; N_{j,(i,j)}$$

and

$$[\mathcal{C}_{Je}]_{ij} = \begin{cases} [\mathcal{C}_{Jp}]_{ij} & \text{if } i \text{ and } j \text{ are adjacent} \\ 0 & \text{otherwise} \end{cases}$$

Both $\mathcal{C}_{Ju}$, $\mathcal{C}_{Jp}$ and $\mathcal{C}_{Je}$ are square matrices of order n. Evidently, for trees the Cluj matrix $\mathcal{C}_{Jp}$ coincides with the Wiener matrix $\mathcal{W}$.

Defining $N_{i,(i,j)}$ for Cycle–Containing Graphs

The crucial step in the theory of Chj matrices is the way in which the definition of the quantity $N_{i,(i,j)}$ is extended to cycle–containing graphs. In three recent papers [1]–[3] the matrix element $N_{i,(i,j)}$ has been defined in the following manner (cf. Eq. (8) in Ref. [1]. Eq. (9) in Ref. [2] and Eq. (6) in Ref. [3]):

$$N_{i,(i,j)} = \max \; |\{v | v \in \mathbf{V}(G) \; ; \; d(i,v|G) < d(j,v|G) \; ;$$

$$(i,v) \cap (i,j) = \max\{i\} \; ; |(i,j)| = \min\}| \; . \tag{5}$$

Here G stands for any connected graph, including graphs that contain cycles.

The right–hand side of formula (5) is formally incorrect for several reasons. First of all, because $\{i\}$ is a fixed, one–element set, the term $\max\{i\}$ is void of any meaning. Another problem is that in cycle–containing graphs there may be several paths connecting the same two vertices. Therefore the symbol (i,v) must not be used for specifying a path connecting the vertices i and v. An analogous objection applies to the symbol (i,j).

By carefully reading the papers [1] [3] and, especially, by examining the examples given therein, we find that by means of formula (5) the authors intended to conceive the quantity $N_{i,(i,j)}$ as the count of vertices, satisfying five conditions – (a), (b), (c), (d) and (e).

Conditions (a) and (b) are immediate:

(a) $v \in \mathbf{V}(G)$, i. e., v is a vertex of G;

(b) $d(i,v|G) < d(j,v|G)$, i. e., v lies closer to i than to j.

In order to properly formulate conditions (c) and (d) we must first change the notation. For fixed vertices x and y, the paths connecting x and y will be denoted by $p_1(x,y), p_2(x,y), \ldots$.

Suppose that we have chosen a particular path $p_k(i,j)$ connecting the vertices i and j. Then condition (c) may be formulated as follows.

(c) There exists a path $p_h(i,v)$ between vertices v and i, such that

$$p_h(i,v) \cap p_k(i,j) = \{i\}$$

or, what is the same,

$$\min |p_h(i,v) \cap p_k(i,j)| = |\{i\}| = 1$$

where the minimum is taken over all values of $h = 1, 2, \ldots$ (and for a fixed, chosen, value of k).

Because the formal mathematical expression for condition (c) is quite complicated, we shall denote it by shorthand as: $v \dashv i$.

Using the above established notation condition (d) should read:

(d) $|p_k(i,j)| = \min$, i. e., the path denoted by $p_k(i,j)$ is a shortest path between the vertices i and j, which (path) still needs not be unique.

Finally, in order to make the definition of $N_{i,(i,j)}$ unique one has to make an additional requirement [1] [3]:

(e) $N_{i,(i,j)}$ pertains to the set $\mathcal{N}_{ij}$ with maximum cardinality.

With all these amendments we arrive at the following:

Definition 1. If G is an arbitrary connected graph, and if i and j are its two distinct, but otherwise arbitrary vertices, then the quantity $N_{i,(i,j)}$ is given by

$$N_{i,(i,j)} = \max |\mathcal{N}_{ij}(k)| \tag{6}$$

with $\mathcal{N}_{ij}(k)$ being a subset of the vertex set of G, determined via:

$$\mathcal{N}_{ij}(k) = \{v | v \in \mathbf{V}(G)\ ;\ d(i,v|G) < d(j,v|G)\ ;\ v \dashv i\ ;\ |p_k(i,j)| = \min\}\ . \tag{7}$$

In (6) the maximum is taken over all paths $p_k(i,j)$ which, in accordance with (7) must be the shortest paths.

Concluding Remarks

Definition 1 is just a less transparent way of saying that the five conditions (a)–(e) have to be simultaneously satisfied. Its form is adjusted so as to resemble, as much as possible, to the earlier (incorrect) definition (5).

By means of Definition 1 we have accomplished a formally correct and generally applicable extension of the Wiener matrix to arbitrary cycle–containing graphs. Formulas (6) & (7) have been designed so as to be in harmony with the ideas of the works [1]–[3], but so as to avoid the flaws of the earlier definition.

It is straightforward to verify that in the case of trees the formulas (6) & (7) reduce to (3) & (4). In other words, the quantities $N_{i,(i,j)}$ and $N_{j,(i,j)}$, defined via Eqs. (6) and (7), satisfy Eq. (2) for trees. This, in turn, means that for trees

$$\mathcal{C}_{Jp} = \mathcal{W}$$

implying that the Cluj matrix is a proper generalization of the Wiener matrix.

Acknowledgement

One author (I. G.) thanks the Mathematical Institute in Belgrade for financial support of this research.

References

[1] M. V. Diudea, Indices of Reciprocal Properties or Harary Indices, *J. Chem. Inf. Comput. Sci.* **37** (1997) 292–299.

[2] M. V. Diudea, Cluj Matrix Invariants, *J. Chem. Inf. Comput. Sci.* **37** (1997) 300–305.

[3] M. V. Diudea, B. Parv and M. I. Topan, Derived Szeged and Cluj Indices, *J. Serb. Chem. Soc.* **62** (1997) 267–276.

[4] H. Wiener, Structural Determination of Paraffin Boiling Points, *J. Am. Chem. Soc.* **69** (1947) 17–20.

[5] M. Randić, Novel Molecular Descriptor for Structure–Property Studies, *Chem. Phys. Letters* **211** (1993) 478-483.

[6] There are, in fact, two different Wiener matrices: one mentioned in this paper and denoted by $\mathcal{W}$, and another whose (ij)-entry is zero whenever the vertices i and j are not adjacent. More details on this matter are found in the papers [5],[7],[8] and in the review [9].

[7] M. Randić, X. Guo, T. Oxley and H. Krishnapriyan, Wiener Matrix: Source of Novel Graph Invariants, *J. Chem. Inf. Comput. Sci.* **33** (1993) 709–716.

[8] M. Randić, X. Guo, T. Oxley, H. Krishnapriyan and L. Naylor, Wiener Matrix Invariants, *J. Chem. Inf. Comput. Sci.* **34** (1994) 361–367.

[9] M. V. Diudea and I. Gutman, Wiener-Type Topological Indices, *Croat. Chem. Acta* **71** (1998) 21–51.

[10] M. V. Diudea, Wiener and Hyper–Wiener Numbers in a Single Matrix, *J. Chem. Inf. Comput. Sci.* **36** (1996) 833–836.

[11] M. V. Diudea, Cluj Matrix $\mathcal{C}_{Ju}$: Source of Various Graph Descriptors, *Commun. Math. Comput. Chem. (MATCH)* **35** (1997) 169–183.

[12] A. A. Kiss, G. Katona and M. V. Diudea, Szeged and Cluj Matrices within the Matrix Operator $W_{(M1,M2,M3)}$, *Coll. Sci. Papers Fac. Sci. Kragujevac* **19** (1997) 95–107.