PATH AND WALK MATRICES OF TREES

Milan KUNZ

Chemopetrol, Research Institute of Macromolecular Chemistry, 656 49 Brno

Received November 10, 1988
Accepted December 13, 1988

The relation between the Rouse matrix $\mathbf{R}$ and the Zimm matrix $\mathbf{Z}$, defined for polymer chains, is generalized for all trees. The path matrix $\mathbf{W}_u$ of an unoriented tree and the walk matrix $\mathbf{W}_o$ of an oriented tree are defined and their relations with the corresponding incidence matrices $\mathbf{G}$ and $\mathbf{S}$ of trees: $(\mathbf{G}\mathbf{G}^T) = n(\mathbf{W}_u^T\mathbf{W}_u)^{-1}$ and $(\mathbf{S}\mathbf{S}^T) = n(\mathbf{W}_o^T\mathbf{W}_o)^{-1}$ are proved. The elements of the quadratical forms $\mathbf{W}^T\mathbf{W}$ are the number of path (selfavoiding walks) in which the edge (arc) i is present together with the edge (arc) j. The traces of $\mathbf{W}^T\mathbf{W}$ are the Wiener path number.

Many physicochemical properties of molecules were connected, directly or indirectly, with invariants of the corresponding molecular graph matrices[1,2]. In Table I some of known matrices of graphs together with their corresponding relations are reviewed.

The simplest invariants of graph matrices are known as topological indices. The oldest one is the Wiener path number, which is the sum of distances in a graph[3]. With the Wiener number such physical properties as boiling points of linear alkanes[4,5], boiling points of alcohols and ethers[6], viscosity of linear and comblike alkanes[7] etc. were successfully correlated. A more detailed analysis connects with the physicochemical properties the spectra of graphs[2].

In the theory of random flight of polymer chains, the Rouse matrix $\mathbf{R}$ is used. It is the inverse of the Zimm matrix $\mathbf{Z}$, which is the matrix $\mathbf{S}\mathbf{S}^T$ of the bond graph[8-11] of the polymer chain. Distribution functions derived from random flight models include the effects of excluded volume and perturbations due to solvent interaction into the square of the polar radii of gyrations $\langle S^2 \rangle$. Their components are defined by the relation

$$(nS_x)^2 = \sum_{i=1}^{n} x_i^2 .$$

where x_i is the x^{th} component of the position vector $\mathbf{P}_i$ from the center of mass of the i^{th} polymer bead, $(n + 1)$ is the number of beads in a polymer chain.

This expression leads to the matrix equation

$$(nS_x)^2 = v^T\mathbf{R}v , \qquad (1)$$

TABLE I

Matrices of graphs; $\boldsymbol{I}$ is the diagonal unit matrix, $\boldsymbol{M}^{\mathrm{T}}$ is the transposed matrix, $\boldsymbol{M}^{-1}$ is the inversed matrix

Matrix	Code	Matrix elements	Ref.
Adjacency	$\boldsymbol{A}$	a_{ii} = the number of loops incident with the vertex i a_{ij} = the number of edges incident with both vertices i and j a_{ij} = the number of arcs out of i into j in the oriented graphs	1, 2
Incidence	$\boldsymbol{S}$	$s_{ij} = +1$ if the arc i is incident out of the vertex j $s_{ij} = -1$ if the arc i is incident into the vertex i $= 0$ otherwise	
	$\boldsymbol{G}$	$g_{ij} = +1$ if the edge i is incident with the vertex j $= 0$ otherwise $\boldsymbol{A}$ of an unoriented graph without loops $\boldsymbol{A} = (1/2)(\boldsymbol{G}^{\mathrm{T}}\boldsymbol{G} - \boldsymbol{S}^{\mathrm{T}}\boldsymbol{S})$	
Graph	$\boldsymbol{B}$	b_{ii} = the degree of the vertex i b_{ij} = the number of edges incident with both vertices i and j $\boldsymbol{B} = \boldsymbol{G}^{\mathrm{T}}\boldsymbol{G}$	
Kirchhoff		$\boldsymbol{S}^{\mathrm{T}}\boldsymbol{S}$ or $\boldsymbol{S}\boldsymbol{S}^{\mathrm{T}}$	
Adjacency of the bond graph	$\boldsymbol{A}_{\mathrm{b}}$	$\boldsymbol{A}_{\mathrm{b}} = \boldsymbol{G}\boldsymbol{G}^{\mathrm{T}} - 2\boldsymbol{I}$	1, 11
Distance	$\boldsymbol{D}$	d_{ij} = number of edges on the walk between vertex i and vertex j d_{ij} are powers d of matrices $\boldsymbol{A}^{d}$ at which the matrix element a^{d}_{ij} is at first non zero $1/2\sum d_{ij}$ = the Wiener number	2 1, 2
Cycle	$\boldsymbol{C}$	$c_{lk} = +1$ if the edge k belongs to the l^{th} simple cycle $= 0$ otherwise $\boldsymbol{RC} = 0$ under addition modulo 2	
Zimm for a chain	$\boldsymbol{Z}$	$z_{ii} = 2$ $z_{i(i\pm 1)} = -1$ $= 0$ otherwise $\boldsymbol{Z} = \boldsymbol{S}^{\mathrm{T}}\boldsymbol{S}$ of the chain	15
Rouse for a chain	$\boldsymbol{R}$	$\boldsymbol{R} = \boldsymbol{Z}^{-1}$ $r_{ij} = (n - i)\,j$ if $i \geq j$ $r_{ij} = (n - j)\,i$ if $i \leq j$	8

TABLE I
(*Continued*)

Matrix	Code	Matrix elements	Ref.
Path	$\mathbf{W}_u$	$\mathbf{W}_u(n)$ is defined in the block form. To $\mathbf{W}_u(n-1)$ is added the block with the elements $w_{ij} = +1$ if the edge j is in the path i in the even distance (0, 2, 4, ...) from the last edge in the block $w_{ij} = -1$ if the edge j is in the path i in the odd distance (1, 3, ...) from the last edge in the block $w_{ij} = 0$ otherwise $\mathbf{W}_u(n)$ of a chain $w_{ij} = \pm 1$ if $(i+j) \geqq n$ $w_{ij} = 0$ if $(i+j) < n$ $\mathbf{W}_u^T \mathbf{W}_u$ of a chain $w_{ij} = \min(i, j)$ $\sum_{k=1}^{n} \mathbf{W}_u^T \mathbf{W}(k) = \mathbf{R}$	This paper
Walk	$\mathbf{W}_0$	$\mathbf{W}_0(n)$ is defined in the block form. To $\mathbf{W}_0(n-1)$ is added the block with elements $w_{ij} = +1$ if the arc j is in the walk i and has the same orientation as the last arc in the block $w_{ij} = -1$ if the arc j is in the walk i and has the opposite orientation as the last arc in the block	
	$\mathbf{W}^T\mathbf{W}$	$w_{ij} =$ the number of path (walks) in which the edge (arc) i is present $w_{ij} =$ the number of path (walks) in which the edge (arc) j is together with the edge (arc) i The signs are defined in the text.	

where $\mathbf{v}$ is the column vector of probabilities v_i of finding of the x_i component, $\mathbf{v}^T$ is its transpose and $\mathbf{R}$ is the Rouse matrix (Table I).

The solution of Eq. (*1*) has the form

$$(nS_x)^2 = \eta_i^2 \sum \lambda_i = \eta_i^2 \,\mathbf{Tr}(\mathbf{R}), \tag{2}$$

where η_i are transformed coordinates of polymer beads, which are supposed to be equal, λ_i are the eigenvalues of a matrix and $\mathbf{Tr}$ is the trace of a matrix. The final solution is found by other approximations.

The problem of radii of gyration has its inverse. Viscoelastic properties of polymer chains were connected with the nearest neighbour matrix $\mathbf{Z}$ (Zimm matrix)

$$\dot{\mathbf{X}} = (3kT/b\varrho)\,\mathbf{ZX}\,, \tag{3}$$

where $\mathbf{X}$, $\dot{\mathbf{X}}$ are column vectors of bead positions and bead velocities, k is the Boltzmann constant, T is the temperature, b is the mean square end to end distance of the submolecules and ϱ is the friction coefficient of the bead[12]. The solution is given by the eigenvalues of the Zimm matrix $\mathbf{Z}$.

The aim of this paper is to show that the relation between Rouse and Zimm matrices can be generalized for all trees and that the Rouse matrix is the matrix of selfavoiding walks in a chain and its trace is the Wiener number.

THEORETICAL

The row vectors of the cycle matrix $\mathbf{C}$ are the closed paths or walks in the space of at most $\binom{n}{2}$ edges (arcs) of a graph with the n vertices. We can define similarly matrices for all paths or walks in a graph. In the case of trees T such matrices have only $(n - 1)$ nonzero columns and are distinct from the blocks of cycle matrices $\mathbf{C}$ which are at trees zero matrices, because the trees are acyclic graphs.

The incidence matrices of trees $\mathbf{G}$ and $\mathbf{S}$ differ only by the sign of their elements, the matrix $\mathbf{G}$, defined for unoriented graphs, has all elements positive, the matrix $\mathbf{S}$, defined for oriented graphs has always one element in a row negative.

Thus the quadratical forms of the matrix $\mathbf{G}$: $\mathbf{G}^{\mathrm{T}}\mathbf{G}$ and $\mathbf{GG}^{\mathrm{T}}$ have all their elements positive. The quadratical forms of the matrix $\mathbf{S}$ have always positive elements on the diagonals. All off-diagonal elements of the matrix $\mathbf{S}^{\mathrm{T}}\mathbf{S}$ are negative, because in the rows of the matrix $\mathbf{S}$ is always one positive and one negative element, at the matrix $\mathbf{SS}^{\mathrm{T}}$ the elements are positive, if both arcs i and j have the oposite orientation, or negative, if they have the same orientation. At a chain with the same orientation of its arcs all off-diagonal elements of the matrix $\mathbf{SS}^{\mathrm{T}}$ are negative.

The matrices $\mathbf{S}^{\mathrm{T}}\mathbf{S}$ and $\mathbf{SS}^{\mathrm{T}}$ are known as the Kirchhoff matrices[2]. The matrix $\mathbf{S}^{\mathrm{T}}\mathbf{S}$ in n dimensional and has only one zero eigenvalue. The matrix $\mathbf{SS}^{\mathrm{T}}$ of trees is $(n - 1)$ dimensional and because both quadratical forms $\mathbf{S}^{\mathrm{T}}\mathbf{S}$ and $\mathbf{SS}^{\mathrm{T}}$ have identical spectra, except zero eigenvalues, it is non-singular and has the inverse. It is thus possible to identify the Zimm matrix $\mathbf{Z}$ as the matrix $\mathbf{SS}^{\mathrm{T}}$ of the chain and the Rouse matrix $\mathbf{R}$ is its inverse $(\mathbf{SS}^{\mathrm{T}})^{-1}$.

Because the matrices $\mathbf{SS}^{\mathrm{T}}$ of all oriented trees are non-singular and have their inverse matrices, it is possible to find them; this can be expected also for similar matrices $\mathbf{GG}^{\mathrm{T}}$ of unoriented trees.

It was noticed that the Rouse matrix $\mathbf{R}$ is the matrix of selfavoiding walks in the chain which gave a hint how to solve the problem of the inverse matrices $(\mathbf{SS}^{\mathrm{T}})^{-1}$ and $(\mathbf{GG}^{\mathrm{T}})^{-1}$ generally. For this it is suitable to define at first the path matrix $\mathbf{W}_u$ and the walk matrix $\mathbf{W}_0$ of a tree. These matrices are defined recursively from matrices of trees with $(n-1)$ vertices. If the new vertex n is adjoined to a tree with $(n-1)$ vertices, a new edge (arc) appears connecting it with a vertex and $(n-1)$ paths (walks) appear. Thus the matrices $\mathbf{W}$ have the block form and their elements are:

$\mathbf{W}_u$: $w_{ij} = 1$ if the edge j in the path i is in the even distance $(0, 2, 4, \ldots)$ from the last added edge of the block,

$w_{ij} = -1$ if the edge j in the path i is in the odd distance $(1, 3, \ldots)$ from the last added edge.

$\mathbf{W}_0$: $w_{ij} = 1$ if the arc j in the walk i has the identical orientation as the last added arc.

$w_{ij} = -1$ if the arc j in the walk i has the opposite orientationas the last added arc,

$w_{ij} = 0$ if the edge (arc) j is not in the path (walk) i.

There are two quadratical forms of matrices $\mathbf{W}$, their projections into the space of paths (walks) $\mathbf{WW}^{\mathrm{T}}$ and their projections into the space of edges (arcs) $\mathbf{W}^{\mathrm{T}}\mathbf{W}$.

A path (walk) gives the distance between its starting and ending vertices, thus the sum of quadrates of unit elements of the matrix $\mathbf{W}$ gives the total distance of vertices in a tree, which is known as the Wiener path number. These quadrates are obtained on the diagonals of both quadratical forms. In $\mathbf{W}^{\mathrm{T}}\mathbf{W}$ the column sums are, in $\mathbf{WW}^{\mathrm{T}}$ the row sums are, the traces are identical with the Wiener number $W = \mathrm{Tr}(\mathbf{W}^{\mathrm{T}}\mathbf{W}) = \mathrm{Tr}(\mathbf{WW}^{\mathrm{T}})$.

The diagonal elements of $\mathbf{W}^{\mathrm{T}}\mathbf{W}$ give the number of paths (walks) in which the edge (arc) i is, the off-diagonal elements give the number of paths (walks) in which both edges (arcs) ij are together. The sign of the off-diagonal elements depends at edges on their distance in paths. It is $+$, if the distance is even and $-$, if the distance is odd. At arcs the sign is $+$, if both arcs have the same orientation and $-$ if their orientation is opposite. Matrices $\mathbf{WW}^{\mathrm{T}}$ have on their diagonals the number of edges (arcs) in paths (walks), determining distances between vertices.

Now it is possible to formulate the following theorem I:

$$\mathbf{W}_u^{\mathrm{T}}\mathbf{W}_u = n(\mathbf{GG}^{\mathrm{T}})^{-1}$$
$$\mathbf{W}_0^{\mathrm{T}}\mathbf{W}_0 = n(\mathbf{SS}^{\mathrm{T}})^{-1}$$

The quadratical forms $\mathbf{W}^{\mathrm{T}}\mathbf{W}$ of the path or walk matrix of a tree are the n multiple of the inverse matrix of the quadratical forms $\mathbf{GG}^{\mathrm{T}}$ or $\mathbf{SS}^{\mathrm{T}}$ of the incidence matrices $\mathbf{G}$ or $\mathbf{S}$ of the tree.

The proof. In the case of unoriented trees all elements of $\boldsymbol{GG}^{\mathrm{T}}$ are positive and in the matrix $\boldsymbol{W}_{\mathrm{u}}^{\mathrm{T}}\boldsymbol{W}_{\mathrm{u}}$ all neighbour edges of the diagonal edge have negative signs.

If we multiply $(\boldsymbol{W}^{\mathrm{T}}\boldsymbol{W})(\boldsymbol{GG}^{\mathrm{T}})$, we count paths. If an edge i is on both matrices on the diagonal, all its paths are counted twice. From this number are substracted twice all paths in which the edge i is not on the end of a path and once all paths, at which the edge is on the end of a path. This leaves $(n-2)$ paths connecting the edge with other ending edges and 2 paths in which the edge is alone. Thus on the diagonal of $(\boldsymbol{W}^{\mathrm{T}}\boldsymbol{W})(\boldsymbol{GG}^{\mathrm{T}})$ are n. If an edge, off-diagonal, say d, in the matrix $\boldsymbol{W}^{\mathrm{T}}\boldsymbol{W}$ is diagonal in the matrix $\boldsymbol{GG}^{\mathrm{T}}$, only paths connecting it with an edge, say l, on the diagonal of the matrix $\boldsymbol{W}^{\mathrm{T}}\boldsymbol{W}$, are counted twice. From them all paths longer than the path dl are subtracted in which the path dl is present. One element, say e, subtracts all paths longer than the path dl, the path dl itself and moreover the path el instead the second path dl which was counted twice. This leaves zeros as off diagonal elements of $(\boldsymbol{W}_{\mathrm{u}}^{\mathrm{T}}\boldsymbol{W}_{\mathrm{u}})(\boldsymbol{GG}^{\mathrm{T}})$ or $(\boldsymbol{GG}^{\mathrm{T}})(\boldsymbol{W}_{\mathrm{u}}^{\mathrm{T}}\boldsymbol{W}_{\mathrm{u}})$. Thus it is enough to divide $\boldsymbol{W}_{\mathrm{u}}^{\mathrm{T}}\boldsymbol{W}_{\mathrm{u}}$ by n for obtaining $(\boldsymbol{GG}^{\mathrm{T}})^{-1}$.

The reasoning for oriented trees is similar, only it is necessary to count with the signs of off-diagonal elements of $\boldsymbol{W}_0^{\mathrm{T}}\boldsymbol{W}_0$ and $\boldsymbol{SS}^{\mathrm{T}}$. They are opposite and thus the products of the off-diagonal elements of $\boldsymbol{W}_0^{\mathrm{T}}\boldsymbol{W}_0\boldsymbol{SS}^{\mathrm{T}}$ or $\boldsymbol{SS}^{\mathrm{T}}\boldsymbol{W}_0^{\mathrm{T}}\boldsymbol{W}_0$ have always the negative sign. The diagonal elements of both quadratical forms are positive. If the identical columns and rows are multiplied, the number of walks with the diagonal arc is counted twice and the walks of the off-diagonal elements are substracted as before. If different columns and rows are multiplied there are two possible combinations of signs

$\boldsymbol{W}_0^{\mathrm{T}}\boldsymbol{W}_0$	(+) diag.	+	−
$\boldsymbol{SS}^{\mathrm{T}}$ the same orientation	−1	(+2)	+1
$\boldsymbol{SS}^{\mathrm{T}}$ the opposite orientation	+1	+1	(+2)

(Values in parentheses lie on the diagonals.)

Both subtract walks as paths were before □.

Corollary: From $\boldsymbol{W}_{\mathrm{u}}^{\mathrm{T}}\boldsymbol{W}_{\mathrm{u}}\boldsymbol{GG}^{\mathrm{T}} = \boldsymbol{GG}^{\mathrm{T}}\boldsymbol{W}_{\mathrm{u}}^{\mathrm{T}}\boldsymbol{W}_{\mathrm{u}} = \boldsymbol{W}_0^{\mathrm{T}}\boldsymbol{W}_0\boldsymbol{SS}^{\mathrm{T}} = \boldsymbol{SS}^{\mathrm{T}}\boldsymbol{W}_0^{\mathrm{T}}\boldsymbol{W}_0 = n\boldsymbol{I}$ the pseudoinverse matrices $\boldsymbol{W}_{\mathrm{u}}^{-1}$, $\boldsymbol{W}_0^{-1}$, $\boldsymbol{G}^{-1}$, $\boldsymbol{S}^{-1}$ are determined.

DISCUSSION

The generalization of the relation of Zimm–Rouse matrices to all trees has many aspects. The viscoelastic theory of entangled polymer chains is not solved completely[12–16] till now. Until the entanglements do not form cycles, the bond graph of the polymer has the inverse and the Rouse theory can be applied.

The interpretation of the Rouse matrix as the matrix of walks in a tree gives it the concrete meaning. There opens an interesting problem of the relation of topological walks in a tree with the walks of this tree in the geometrical lattice. These walks give good agreement with the Rouse theory in the absence of excluded volume, but there are deviations in the presence of excluded volume[17].

There is a sharp change of properties of entangled polymers at the gelation point. The gelation can be connected with formation of cycles by formation of at least two entanglements between two polymer chains, which makes both Kirchhoff matrices of the polymer $\mathbf{S}^T\mathbf{S}$ and $\mathbf{S}\mathbf{S}^T$ singular. In the literature matrices $\mathbf{S}^T\mathbf{S}$ are often given as the Zimm matrices instead of $\mathbf{S}\mathbf{S}^T$. There is only a minor difference of elements of both matrices at chains, they have indentical spectra but one zero eigenvalue is present at $\mathbf{S}^T\mathbf{S}$, which has not any inverse. Many combinations of orientation of arcs are possible at branched trees giving different $\mathbf{S}\mathbf{S}^T$ and also different $\mathbf{W}_0^T\mathbf{W}_0$ matrices. The orientation of arcs has no effect on the spectra of the matrices $\mathbf{S}\mathbf{S}^T$, because it has no effect on the quadratical form $\mathbf{S}^T\mathbf{S}$ which is independent on the orientation of arcs, and it also has no effect on the Wiener number, but it gives different spectra of different matrices $\mathbf{W}_0^T\mathbf{W}_0$. It would be useful to study possible effects of orientation of polymer chains connections on their viscoelastic properties.

The fact that the trace of $\mathbf{W}^T\mathbf{W}$ is the Wiener number gives to its semiempirical application theoretical explanation, now the results of the Rouse theory can be used for explanation of all correlations of physical properties of trees with the Wiener number[3-7].

The Wiener number is defined in the space of vertices also for cyclic graphs and graphs with multiple bonds, where the path and walk matrices $\mathbf{W}$ are undefined.

The Wiener number is not the only topological index based on paths. There was introduced the modified Wiener index for polymers[18].

Many topological indices are based on the distance matrices $\mathbf{D}$ (ref.[19]). These indices could be modified for matrices $\mathbf{W}^T\mathbf{W}$. It is questionable that such indices except the path or walk polynomial were better then the eigenvalues themselves, but it would be interesting to compare them with the indices based on the distance matrix $\mathbf{D}$.

For study of physiological properties of molecules the numbers of paths of different length were used[20]. This is a practical application of the extended connectivity index based on the number of non-selfavoiding walks, obtained on the diagonal of the different powers of the adjacency matrix $\mathbf{A}$. The numbers of selfreturning walks are the moments of eigenvalues of matrix $\mathbf{A}$. Therefore the connecting the length of selfavoiding walks with the eigenvalues of the inverse matrix of the bond of a tree gives very interesting insight into the structure of the graph space. The relationship between the number of selfreturning walks and the Wiener number at some classes of graphs was found only recently[23].

The zeros of the matching polynomial of a molecular graph M are eigenvalues of an acyclic graph $T(H, v)$, which is constructed from the graph M as the graph all acyclic walks going from the vertex v (ref.[24]).

It would be interesting if also their walk polynomials were related somewhat.

REFERENCES

1. Essam J. W., Fisher M. E.: Rev. Mod. Phys. *42*, 272 (1970).
2. Cvetković D. M., Doob M., Sachs H.: *Spectra of Graphs*. Academic Press, Berlin 1983.
3. Wiener H.: J. Am. Chem. Soc. *69*, 17 (1947).
4. Rouvray D. H., Pandey R. B.: J. Chem. Phys. *85*, 2286 (1986).
5. Rouvray D. H.: J. Comput. Chem. *8*, 470 (1987).
6. Lall R. S.: Curr. Sci. *53*, 642 (1984); Chem. Abstr. *101*, 109895 (1984).
7. Hudec P., Weiglová I., Cvengroš J.: Ropa a uhlie, in press.
8. Rouse P. E.: J. Chem. Phys. *21*, 1272 (1953).
9. Forsman W. C., Hughes R. E.: J. Chem. Phys. *18*, 2118 (1963).
10. Forsman W. C., Grand H. S.: Macromolecules *5*, 289 (1972).
11. Bertz S. H.: J. Chem. Soc., Chem. Commun. *1981*, 818.
12. Yang Y., Yu T.: Makromol. Chem. *187*, 441 (1986).
13. Yang Y., Yu T.: Makromol. Chem. *188*, 401 (1987).
14. Gordon M., Roberts K. R.: Polymer *20*, 681 (1969).
15. Eichinger B. E.: Macromolecules *5*, 496 (1972).
16. Hess W.: Macromolecules *21*, 2620 (1988).
17. Downey J. P., Crabb C. C., Kovac J.: Macromolecules *19*, 2202 (1986).
18. Mekenyan O., Dimitrov S., Bonchev D.: Eur. Polym. J. *19*, 1185 (1987).
19. Balaban A. T., Motoc I., Bonchev D., Mekenyan O.: Top. Curr. Chem. *114*, 21 (1982).
20. Grossman S. C., Dzonova B., Jerman B., Randić M.: Int. J. Quantum Chem. Quantum Biol. Symp. *12*, 123 (1986); Chem. Abstr. *104*, 101966 (1986).
21. Razinger M.: Theor. Chim. Acta *61*, 581 (1982).
22. Razinger M.: Theor. Chim Acta *70*, 365 (1986).
23. Bonchev D., Mekenyan O., Polansky O. E.: Stud. Phys. Theor. Chem. *21*, 209 (1987).
24. Godsil C. D., Gutman I.: Acta Chim. (Budapest) *110*, 415 (1982).

Translated by the author.