Revue Roumaine de Chimie, **37**, *11—12*, 1341—1345 (1992)

CHEMICAL GRAPH POLYNOMIALS. PART 2.[1]

THE PROPAGATION DIAGRAM ALGORITHM FOR THE COMPUTATION OF THE CHARACTERISTIC POLYNOMIAL OF MOLECULAR GRAPHS

OVIDIU IVANCIUC
Polytechnic Institute, Faculty of Chemistry,
Department of Organic Chemistry, Splaiul Independenței 313
77206 Bucharest, Roumania

Received May 26, 1992

An algorithm is presented for obtaining the coefficients of the characteristic polynomial of a molecular graph. Spectral moments of the graph are computed using the propagation diagram method, and the coefficients of the characteristic polynomial are obtained using Le Verrier formula. The time complexity of the algorithm is $O(N^3)$, where N is the number of vertices of the graph. In connection with the computation of the coefficients of the characteristic polynomial, the algorithm offers the number of self-returning walk atomic code, structural codes, and spectral moments of the molecular graph.

1. INTRODUCTION

The characteristic polynomial of a molecular graph is an important structural invariant, defined as : [2,3]

$$Ch(G, x) = (-1)^N \det(A - xI) = \sum_{k=0}^{N} a_k x^{N-k} \qquad (1)$$

where A and I are, respectively, the adjacency matrix and unit matrix of the molecular graph G with N vertices. Its applications spread over a large area of mathematical and theoretical chemistry : documentation, computer assisted synthesis design, Kekulé structure enumeration, aromaticity, electronic energy of solids, topological effect on molecular orbitals.

The computation of the characteristic polynomial using the expansion of the determinant from equation (1) is impracticable even for molecular graphs with more than ten vertices, due to the high amount of computer time consumed.

A very efficient method for the computation of the coefficients of the characteristic polynomial is the recursive algorithm developed by Le Verrier. [4,5] Balasubramanian devised an efficient computer program[6] for the computation of the characteristic polynomial of chemical graphs, using the recursive method of Frame.[7] Another related recursive algorithm for the computation of the coefficients of the characteristic polynomial

is the one devised by Fadeev. [8] The computer operation count for the above three algorithms is $O(N^4)$.

Živković [9] proposed a more efficient method for computing the coefficients of the characteristic polynomial : perform the Householder tridiagonalization of the adjacency matrix A, apply the QR algorithm in order to diagonalize A, and use a from of the Le Verrier algorithm to obtain the coefficients a_k of the characteristic polynomial. The operation complexity of the algorithm is $O(N)^3)$, which is one order of magnitude less than the operation count in Le Verrier algorithm.

In the present paper, we propose a method for computing the coefficients of the characteristic polynomial of molecular graphs, by a suitable combination of the propagation diagram algorithm and Le Verrier algorithm. This new algorithm is one order of magnitude less time consuming than Le Verrier algorithm, requiring $O(N^3)$ integer additions, unlike the Živković algorithm, which requires $O(N)^3$ real additions, multiplications, and divisions, operations requiring a much larger computer time.

2. THE COMPUTATION OF THE CHARACTERISTIC POLYNOMIAL BY THE PROPAGATION DIAGRAM ALGORITHM

The dual representation of a molecule, i.e. as a molecular graph or as a matrix, provides us with the possibility to obtain efficient algorithms for computing molecular graph invariants by a proper combination of graph and algebraic algorithms. Such an invariant, connected with the characteristic polynomial , is the number of self-returning walks.

A walk in a graph is a sequence of edges which can be continuously traversed, starting from any vertex and ending on any vertex. Repeated use of the same edge or edges is allowed. A self-returning walk is a walk starting and finishing at the same vertex. The length of a walk is the total number of edges that are traversed.

The number of self-returning walks of length k mai be computed by considering the diagonal elements of the first k powers of the adjacency matrix A, due to the fact that each diagonal element $(A^k)_{ii}$ of the matrix A^k can be interpreted as the sum of all self-returning walks of length k from/to vertex i. [10,11] The sequence of integers $SRWAC_i = \{(A^1)_{ii}, (A^2)_{ii}, \ldots, (A^N)_{ii}\}$ defines the self returning walk atomic code of the atom i in a molecule. [12]

If a matrix A has the eigenvalues $x_1, x_2, \ldots, x_N$, then A^k has the eigenvalues $x_1^k, x_2^k, \ldots, x_N^k$ ($k = 0, 1, 2, \ldots, N$). The spectral moment of order k (SM_k) is defined as :

$$SM_k = \sum_{i=1}^{N} x_i^k = \mathrm{Tr}\, A^k \tag{2}$$

where the trace of the k th power of the adjacency matrix is equal to :

$$\mathrm{Tr}\, A^k = \sum_{i=1}^{N} (A^k)_{ii} \tag{3}$$

An important result is the geometric interpretation of the equation (2). [13-16]. The kth spectral moment of the adjacency matrix A corresponds to the count of all self-returning walks of length k for the molecular graph G.

The structural code of the vertex i (SC_i) was defined as the summation of the atomic code of the vertex i : [17,18]

$$SC_i = \sum_{k=1}^{N} (A^k)_{ii} \tag{4}$$

From expressions (2) and (3) one can readily obtain a relationship between the kth spectral moment and atomic codes :

$$SM_k = \sum_{i=1}^{N} SRWAC_{i,k} \tag{5}$$

A very attractive graphical method, termed the Propagation Diagram Algorithm (PDA), was introduced by Burdett [19,20] for computing the spectral moments. This method may be connected with the Le Verrier algorithm, providing a very efficient and rapid procedure of computing the coefficients of the characteristic polynomial.

The PDA consists of the following steps :

(*i*) Each vertex of the molecular graph is given an arbitrary number from 1 to N.

(*ii*) A vertex i is selected as seed vertex and weighted with a Propagation Number (PN) equal to one :

$$PN_{i,0} = 1$$

The remaining N—1 vertices are weighted with PN equal to zero.

(*iii*) $PN_{j,k}$ values in the kth interation are computed for every vertex j in the molecular graph :

$$PN_{j,k} = \sum_{m=1}^{d_j} PN_{m,k-1}$$

where the summation goes over all d_j vertices adjacent to the vertex j. Step (*iii*) is repeated until $PN_{i,N}$ is computed.

After finishing step (*iii*), the row in the PN matrix corresponding to the seed vertex i contains the SRWAC of the vertex i, i.e.

$$SRWAC_i = \{PN_{i,1}, PN_{i,2}, \ldots, PN_{i,N}\}$$

After computing the $SRWAC_i$, another vertex is selected as seed vertex and steps (*ii*) and (*iii*) are repeated until all N vertices in the molecular graph are selected as seeds vertices.

(*iv*) The spectral moments are computed using the expression (5).

(*v*) Finally, the coefficients of the characteristic polynomial are computed using the Le Verrier formula :

$$a_k = (1/k)(\mathrm{Tr}\ \mathbf{A}^k - a_1\ \mathrm{Tr}\ \mathbf{A}^{k-1} - \ldots - a_{k-1}\ \mathrm{Tr}\ \mathbf{A})$$

Considering the relation between the spectral moment of order k, SM_k, and the trace of the kth power of the adjacency matrix (Eq. 2), the coefficient a_k of the characteristic polynomial may be expressed as :

$$a_k = (1/k)(SM_k - a_1 SM_{k-1} - \ldots - a_{k-1} SM_1)$$

We now estimate the operation count for the PDA. Let DEG_i denotes the degree of the vertex i in G, i.e. the number of edges incident with the vertex i. We define the average degree of a graph as :

$$DEG_{av} = \frac{\sum_{i=1}^{N} DEG_i}{N} = \frac{2m}{N}$$

where m is the number of edges in the molecular graph. For a tree (a graph without cycles) $m = N-1$, for a monocyclic graph $m = N$, for a cubic graph $m = \frac{3}{2} N$, while for the complete graph K_N, $m = N(N-1)$.

For a graph with N vertices, without loops and weighted edges, the computation of the N propagation diagrams requires a number of $DEG_{av}N^3$ additions. For example, the evaluation of the propagation diagrams for a monocyclic graph requires $2N^3$ additions, while a cubic graph requires $3N^3$ additions. The greatest computational effort is required by the complete graph K_N, which requires N^4 summations, but still remains faster than the Le Verrier algorithm, which requires a number of extra N^4 multiplications. The Le Verrier component of the PDA requires about N^2 operations. Hence, the operation count is $O(N^3)$ for molecular graphs, when usually DEG_{av} lies between 2 and 3. All operations needed for the computation of the propagation diagrams are integer operations, which are performed faster than real numbers operations. Also, integer additions are performed more faster than integer multiplications.

The Le Verrier algorithm also may be performed in integer arithmetic, but it requires N^4 additions and N^4 multiplications, which makes the algorithm slower than PDA.

The Živković algorithm requires about $(2/3)\ N^3 + 32N^2$ real numbers operations. Since the PDA requires only integer additions, the fastest computer operation, and, as stated above, for molecular graphs DEG_{av} lies between 2 and 3, it follows that the PDA is faster even the Householder-QL procedure, for molecular graphs.

Another major advantage of the PDA is that is readily offers besides the coefficients of the characteristic polynomial, the self returning walk atomic codes, structural counts and spectral moments of the molecular

graph. This ability makes the PDA a unique method among the methods for computing the characteristic polynomial.

The PDA was verified by means of a computer program in the C programming language. All results agree with the coefficients of the characteristic polynomial obtained by the Le Verrier algorithm, but the PDA performs much faster the computations. A copy of the program may be obtained on request from the author.

REFERENCES

[1] Part 1 : O. Ivanciuc, *Rev. Roum. Chim.*, **33**, 709 (1988).
[2] N. Trinajstić, *J. Math. Chem.*, **2**, 197, (1988).
[3] K. Balasubramanian, *Chem. Rev.*, **85**, 599 (1985).
[4] U.J. J. Le Verrier, *J. Math.*, **5**, 95. (1840).
[5] U.J.J. Le Verrier, *J. Math.*, **5**, 220 (1840).
[6] K. Balasubramanian, *J. Comput. Chem.*, **5**, 387 (1984).
[7] P.S. Dwyer, "Linear Computations", Wiley, New York, 1951, pp. 225-235.
[8] D. K. Fadeev and I. S. Sominskii, "Problems in Higher Algebra", Freeman, San Francisco, 1965.
[9] T. P. Zivković, *J. Comput. Chem.*, **11**, 217 (1990).
[10] F. Harary, "Graph Theory", Adison-Wesley, Reading, Mass., 1969.
[11] N. Deo, "Graph theory with applications to engineering and computer science", Prentice-Hall, Englewood Cliffs, N. J., 1974.
[12] M. Randić, *J. Comput. Chim.*, **1**, 386 (1980).
[13] M. F. Sykes and M.E. Fisher, *Adv. Phys.*, **9**, 315 (1960).
[14] F. Cyrot-Lackmann, Ph. D. Thesis, Orsay, 1968.
[15] F. Ducastelle and F. Cyrot-Lackmann, *J. Phys. Chem. Solids*, **31**, 1295(1970).
[16] F. Ducastelle and F. Cyrot-Lackmann, *J. Phys, Chem. Solids*, **32**, 285(1971).
[17] M. Barysz and N. Trinajstić, *Int. J. Quantum Chem. : Quantum Chem. Symp.*, **18**, 661 (1984).
[18] M. Barysz, J. V. Knop, S. Pejaković and N. Trinajstić, *Polish J. Chem.*, **59**, 405 (1985).
[19] J. K. Burdett, S. Lee and W.C. Sha, *Croat. Chem. Acta*, **57**, 1193 (1984).
[20] J. K. Burdett, *Acc. Chem. Res.* **21**, 189 (1988)