Molecular Graphs with Equal Z-Counting and Independence Polynomials

Ivan Gutman* and Haruo Hosoya**

 * Faculty of Science, University of Kragujevac, Yugoslavia
** Faculty of Science, Ochanomizu University, Tokyo, Japan

It is demonstrated that the cycle graphs corresponding to the carbon atom skeletons of cyclo-alkanes and annulenes are the only graphs having the same Z-counting and independence polynomials.

Introduction

The concept of the non-adjacent number, $p(G, k)$, i.e., the number of ways for choosing k disjoint edges from a given graph or lattice, G, has been proposed quite frequently and independently in many different areas, e.g., statistical mechanics [1–3], quantum theory of conjugated hydrocarbons [4], information chemistry [5, 6]. By assembling the set of the $p(G, k)$ numbers in such a way as typically used in combinatorics and graph theory, the so-called Z-counting polynomial, $Q(G, x)$, can be defined as

$$Q(G, x) = \sum_{k=0}^{m} p(G, k)\, x^k, \qquad (1)$$

where m stands for the number of edges of the graph G. As already mentioned, $p(G, k)$ is the number of ways in which k mutually disjoint edges can be selected in the graph G. In addition to this, for $k=1$ and $k=0$ we define $p(G, 1)=m$ and $p(G, 0)=1$. The name Z-counting polynomial was proposed in [5], because $Q(G, 1)$ is equal to the topological index Z. It turns out that $Q(G, x)$ has interesting and important mathematical properties not only as a counting polynomial as it stands, but also as the key factor for the graph-theoretical interpretation of the characteristic polynomial of a graph [5, 7].

Let $q(G, k)$ denote the number of ways for choosing k disjoint vertices from G. Let further $q(G, 1)=n=$ number of vertices of G, and $q(G, 0)=1$. Then, in analogy to (1), we define the independence polynomial $Qh(G, x)$ as the vertex counterpart of the Z-counting polynomial:

$$Qh(G, x) = \sum_{k=0}^{m} q(G, k)\, x^k. \qquad (2)$$

The same of this polynomial comes from the fact that $q(G, k)$ counts the k-element sets of independent vertices of G. This polynomial was first proposed in [8, 9], but only few of its mathematical properties and chemical applications have been discussed so far [8–11]. It is worth mentioning that the independence polynomial of the Clar graph [8, 10] is just the sextet polynomial [12] of the respective benzenoid system.

As will be inferred from the following two examples, $Q(G, x)$ and $Qh(G, x)$ have opposite tendencies in their information content:

$$Q(G, x) = 1 + 5x + 4x^2,$$
$$Qh(G, x) = 1 + 6x + 10x^2 + 6x^3 + x^4.$$

$$Q(G, x) = 1 + 6x + 3x^2,$$
$$Qh(G, x) = 1 + 4x.$$
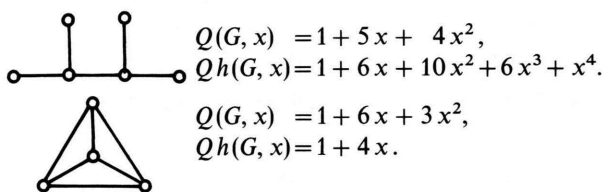
We have calculated $Q(G, x)$ and $Qh(G, x)$ for a large number of molecular graphs. This study enables us to conjecture that only the cycles C_n corresponding to the molecular graphs of cycloalkanes or of annulenes (Fig. 1) have identical polynomials $Q(G, x)$ and $Qh(G, x)$.

The aim of this paper is to demonstrate that this is indeed the case, namely that $Q(G, x)$ and $Qh(G, x)$ coincide in a single case – only when G is a cycle. For all other molecular graphs $Q(G, x)$ differs from $Qh(G, x)$. In addition to its graph-theoretic and information-scientific meaning, relevant discussions in the aforementioned various fields of science might follow this theorem.

The Mathematical Apparatus [13–15]

We denote by $\mathscr{F}_n$, $\mathscr{M}_n$, and $\mathscr{M}_n^*$ the sets of acyclic graphs with n vertices, connected monocyclic graphs with n vertices and monocyclic graphs (both connected and disconnected) with n vertices, respectively.

The path P_n is a connected acyclic graph with two vertices of degree one. The graph E_n is a set of n disjoint vertices containing no edge (Fig. 1). Of course, P_n, $E_n \in \mathscr{F}_n$.

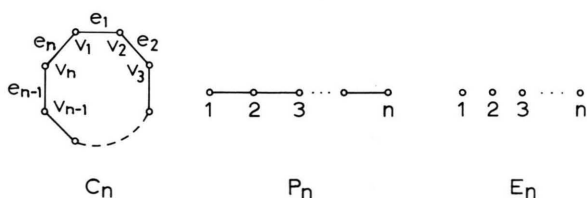C_n $\qquad$ P_n $\qquad$ E_n

Fig. 1.

The basic mathematical properties of the Z-counting polynomials have been discussed elsewhere [5, 8, 9, 16]. They conform to the following recursion relations:

$$Q(G, x) = Q(G-e, x) + x \cdot Q(G-u-v, x), \quad (3)$$

where e denotes an edge of G whose endpoints are u and v, and

$$Qh(G, x) = Qh(G-v, x) + x \cdot Qh(G-N_v, x), \quad (4)$$

where v is a vertex of G and N_v is the set of vertices containing v and its first neighbors. The above relations are tantamount to

$$p(G, k) = p(G-e, k) + p(G-u-v, k-1) \quad (5)$$

and

$$q(G, k) = q(G-v, k) + q(G-N_v, k-1), \quad (6)$$

respectively. The relations (3)–(6) can be derived directly from the inclusion-exclusion principle [17]. An important special case of (5) and (6) is obtained when the graph G contains a vertex s of degree one, adjacent to the vertex t.

Then (5) and (6) reduce to

$$p(G, k) = p(G-s, k) + p(G-s-t, k-1), \quad (7)$$

$$q(G, k) = q(G-s, k) + q(G-s-t, k-1). \quad (8)$$

Of the many special cases which can be deduced from (5) and (6) we shall later need the Fibonacci-type relation

$$q(C_n, k) = q(C_{n-1}, k) + q(C_{n-2}, k-1). \quad (9)$$

A formal proof of (9) can be found in [11], although results equivalent to (9) were reported already in [5, 16, 18].

The Main Result

The main result of the present paper is a complete characterization of those molecular graphs which have identical Z-counting and independence polynomials. We prove that only one class of molecular graphs has such a property, namely the cycles C_n ($n \geq 3$). We summarize our result in the following manner.

Theorem

$Q(G, x) = Qh(G, x)$ for a connected graph G holds if and only if $G = C_n$.

Note that $Q(G, x) = Qh(G, x)$ is equivalent to the requirement for all values of k, $p(G, k) = q(G, k)$. In order to prove the Theorem we need some preparations.

Some Auxiliary Results

Lemma 1. For all $F \in \mathscr{F}_n$ and for all k, $p(F, k) \leq p(P_n, k)$ and $q(F, k) \geq q(P_n, k)$. Equality for all values of k is achieved only if $F = P_n$.

The *Proof* follows by induction on the number n of vertices of the graph F and is already given in [9, 19].

Lemma 2. If G is a connected graph, then the equality $Q(G, x) = Qh(G, x)$ implies that G is monocyclic, i.e., $G \in \mathscr{M}_n$.

Proof. If $Q(G, x) = Qh(G, x)$, then $p(G, k) = q(G, k)$ holds for all values of k. Setting $k = 1$ we obtain $m = n$. A connected graph having equal numbers of edges and vertices is necessarily monocyclic. $\square$

From the proof of Lemma 2 it is evident that for all connected graphs with more than one cycle and for acyclic graphs, $Q(G, x)$ and $Qh(G, x)$ cannot coincide.

Lemma 3. For all $M \in \mathscr{M}_n$ and for all k, $p(M, k) \leq p(C_n, k)$. Equality for all values of k is achieved only if $M = C_n$.

The *Proof* of Lemma 3 can be found in [20].

Lemma 4. For all $M \in \mathcal{M}_n$ and for all k, $q(M, k) \geq q(C_n, k)$. Equality for all values of k is achieved only if $M = C_n$.

The *Proof* follows by induction on the number n of vertices of the graph M. It is easy to check that the statement of Lemma 4 is true for $n = 3, 4, 5$.

Suppose now that Lemma 4 is obeyed by connected monocyclic graphs with $n-1$ and $n-2$ vertices. We then show that Lemma 4 holds also for connected monocyclic graphs with n vertices.

One has to distinguish among the three cases:
(i) $M = C_n$.
(ii) $M \neq C_n$ and all vertices of degree one in M are adjacent to the cycle contained in M.
(iii) $M \neq C_n$ but not all vertices of degree one in M are adjacent to the cycle contained in M.

The case (ii) is illustrated by the graphs M_1 and M_2, while the case (iii) by the graphs M_3 and M_4 (see Figure 2).
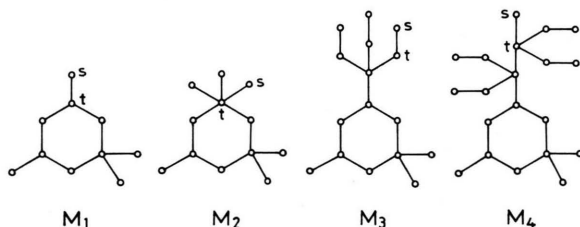

$M_1 \qquad M_2 \qquad M_3 \qquad M_4$

Fig. 2.

Case (i). If $M = C_n$ then (9) holds.

Case (ii). If s denotes a vertex of degree one and t its unique neighbor, then (8) is applicable, and besides $M - s \in \mathcal{M}_{n-1}$, $M - s - t \in \mathcal{F}_{n-2}$. (Note that $M - s$ is necessarily connected; $M - s - t$ may be connected, as in M_1, but may also be disconnected, as in M_2.) In accordance with the induction hypothesis, $q(M-s, k) \geq q(C_{n-1}, k)$. Because of Lemma 1, $q(M-s-t, k-1) \geq q(P_{n-2}, k-1)$. Therefore

$$q(M, k) \geq q(C_{n-1}, k) + q(P_{n-2}, k-1). \tag{10}$$

For at least one k, the above inequality is strict.

Since C_{n-2} is obtained by adding an edge to P_{n-2}, the number of selections of independent vertices in C_{n-2} cannot exceed that in P_{n-2}. Consequently, $q(P_{n-2}, k-1) \geq q(C_{n-2}, k-1)$, and from (9) follows

$$q(M, k) \geq q(C_{n-1}, k) + q(C_{n-2}, k-1) = q(C_n, k).$$

This proves Lemma 4 for the case (ii).

Case (iii). By choosing for s a vertex of degree one which is not adjacent to the cycle of M we have $M - s \in \mathcal{M}_{n-1}$ and $M - s - t \in \mathcal{M}_{n-2}^*$. (Note that $M - s - t$ may be connected as in M_3, but may also be disconnected as in M_4.)

If $M - s - t$ is connected, then by the induction hypothesis

$$q(M, k) = q(M-s, k) + q(M-s-t, k-1)$$
$$\geq q(C_{n-1}, k) + q(C_{n-2}, k-1)$$

and bearing (9) in mind we have

$$q(M, K) \geq q(C_n, k).$$

For at least some values of k, $q(M, k) > q(C_n, k)$.

If $M - s - t$ is disconnected, then we can add to it some edges to obtain a connected monocyclic graph M', such as $M' \in \mathcal{M}_{n-2}$. Because addition of edges to a graph necessarily reduces the independency between the vertices, we have $q(M-s-t, k-1) \geq q(M', k-1)$.

On the other hand, by the induction hypothesis $q(M', k-1) \geq q(C_{n-2}, k-1)$. This finally results in

$$q(M, k) = q(M-s, k) + q(M-s-t, k-1)$$
$$\geq q(M-s, k) + q(M', k-1)$$
$$\geq q(C_{n-1}, k) + q(C_{n-2}, k-1) = q(C_n, k).$$

It is clear from this reasoning that for some k we must have $q(M, k) > q(C_n, k)$. This proves Lemma 4 for the case (iii). □

Proof of the Theorem

a) Proof of the "if" Part

Bearing in mind the above indicated labeling of the vertices and edges of C_n it is immediately evident that a selection $e_{i_1}, e_{i_2}, \ldots, e_{i_k}$ of edges of C_n is independent if the selection of vertices $v_{i_1}, v_{i_2}, \ldots, v_{i_k}$ is independent, and vice versa. Hence a simple one-to-one correspondence between sets of independent edges and sets of independent vertices is established, implying $p(C_n, k) = q(C_n, k)$ for all k. This means that if $G = C_n$, then the Z-counting and independence polynomials of G coincide with each other. □

b) Proof of the "only if" Part

Suppose that G is a connected graph different from C_n. According to Lemma 2, if G is not monocyclic,

$Q(G, x)$ differs from $Q h(G, x)$. Then G is monocyclic. Combining Lemmas 3 and 4 with the "if" part of the Theorem we have

$$p(G, k) \leqq p(C_n, k) = q(C_n, k) \leqq q(G, k) \,.$$

For at least one value of k the above inequalities are strict, implying $p(G, k) < q(G, k)$. However, then the polynomials $Q(G, x)$ and $Q h(G, x)$ cannot be equal. In other words, if G is connected and $G \neq C_n$, then $Q(G, x) \neq Q h(G, x)$. $\square$

By this we completed the proof of the Theorem.

Discussion

Thus we have shown that the cycle graphs which correspond to the carbon atom skeletons of cycloalkanes and annulenes are only molecular graphs whose independence and Z-counting polynomials coincide with each other. As mentioned in the Introduction, the independence polynomial for a tree graph gets more complicated with branching, while that for a non-tree graph gets shortened with the increase of the number of rings. This means that both $Q(G, x)$ and $Q h(G, x)$ reflect the extent of branching and the extent of cyclicity of the corresponding molecule, but in a different manner. The greater the branching and/or the smaller the cyclicity of a molecule, the smaller are the coefficients of $Q(G, x)$ and the larger are the coefficients of $Q h(G, x)$. Thus the cycle graphs are in a certain sense the intermediate cases, namely the graphs which lie somehow half way between the most branched and the least branched molecular graphs. In addition to this, from the graph-theoretical and information-scientific points of view it is interesting to recall that the cycle graphs are the only graphs that coincide with the line graphs, which are constructed from the adjacency relations among the edges of the parent graphs [13, 15].

Acknowledgement

[1] R. J. McQuistan and S. J. Lichtman, J. Math. Phys. **11,** 3095 (1970).
[2] O. J. Heilman and E. H. Lieb, Phys. Rev. Lett. **24,** 1412 (1970).
[3] H. Kunz, Phys. Lett. **32 A,** 311 (1970).
[4] H. Hosoya, Theor. Chim. Acta Berlin **25,** 215 (1972).
[5] H. Hosoya, Bull. Chem. Soc. Japan **44,** 2332 (1971).
[6] H. Hosoya, J. Chem. Doc. **12,** 181 (1972).
[7] H. Hosoya, Discr. Appl. Math. **19,** 239 (1988).
[8] I. Gutman, Z. Naturforsch. **37 a,** 69 (1982).
[9] I. Gutman and F. Harary, Utilitas Math. **24,** 97 (1983).
[10] I. Gutman and S. El-Basil, Z. Naturforsch. **39 a,** 276 (1984).
[11] I. Gutman and S. El-Basil, MATCH **20,** 81 (1986).
[12] H. Hosoya and T. Yamaguchi, Tetrahedron Lett. **1975,** 4659.
[13] F. Harary, Graph Theory, Addison-Wesley, Reading, Mass. 1969.
[14] N. Trinajstić, Chemical Graph Theory, CRC Press, Boca Raton, Florida 1983.
[15] I. Gutman and O. E. Polansky, Mathematical Concepts in Organic Chemistry, Springer-Verlag, Berlin 1986.
[16] H. Hosoya, Fibonacci Quart. **11,** 255 (1973).
[17] For example, see C. L. Liu, Introduction to Combinatorial Mathematics, McGraw-Hill, New York 1968.
[18] I. Gutman and H. Hosoya, Theor. Chim. Acta **48,** 279 (1978).
[19] I. Gutman, Theor. Chim. Acta Berlin **45,** 79 (1977).
[20] I. Gutman, Publ. Inst. Math. Beograd **27,** 67 (1980).