HKD 1926 CROATICA CHEMICA ACTA

CROATICA CHEMICA ACTA
CCACAA, ISSN 0011-1643, e-ISSN 1334-417X
*Croat. Chem. Acta* **86** (4) (2013) 443–451.
http://dx.doi.org/10.5562/cca2311

*Original Scientific Article*

# Distance Polynomial and the Related Counting Polynomials†

**Haruo Hosoya**

*Ochanomizu Univeristy (Emeritus), Bunkyo-ku, Tokyo 112-8610, Japan*
(E-mail: hosoya.haruo@ocha.ac.jp)

RECEIVED JUNE 25, 2013; REVISED SEPTEMBER 5, 2013; ACCEPTED SEPTEMBER 27, 2013

*Abstract.* Interesting mathematical properties of the distance polynomial $S_G(x)$ proposed by the present author in 1973 are reintroduced together with several new findings. Although many results are given in the form of "Theorems", most of them have not rigorously been proved mathematically but readers are challenged to provide their proofs. (doi: 10.5562/cca2311)

*Keywords:* distance matrix, distance polynomial, Wiener polynomial, Hosoya polynomial, topological index

## INTRODUCTION

The distance polynomial $S_G(x)$ for a graph **G** (with $N$ vertices) whose distance matrix is **D** is expressed by

$$S_G(x) = (-1)^N \det(\mathbf{D} - x\mathbf{E}) = \sum_{k=0}^{N} b_k x^{N-k} \qquad (1)$$

where **E** is the unit matrix of the order $N$. It was first proposed by the group of the present author in 1973[1] which was followed by an independent proposal by the group of Graham.[2,3] However, one of the important conjectures proposed in the 1973 paper was proved in a joint paper of these two groups.[4] In 1988 close connections between the distance polynomial and the pair of Wiener indices $w$ and $p$[5] were disclosed by Hosoya[6] by the use of a counting polynomial named as the "Wiener polynomial," $H_G(x)$, which is now called the "Hosoya polynomial" as suggested by Gutman and others[7,8] and supported not only by mathematical chemists but also by mathematicians of various countries.[9–11] One of the reasons for this suggestion comes from the fact that in the paper of the topological index $Z$ of Hosoya[12] the Wiener index was redefined by using the distance matrix **D** of a given graph $G$ as

$$w = \sum_{i<j}^{G} D_{ij} \qquad (2)$$

which, contrary to the complicated original definition by Wiener,[5] enables its application to non-tree graphs.

Let us here explicitly formulate the above discussion by using the definition of $H_G(x)$ as

$$H_G(x) = \sum_{k=1}^{l} d_k x^k \qquad (3)$$

where $2d_k$ is the number of $D_{ij}$ elements that are equal to $k$, where $l$ is the largest element of **D**, or the diameter of G. Another definition of $H_G(x)$ was proposed[8] by adding $d_0 = N$, the number of vertices of G to the Equation (3), but the discussion in this paper is irrelevant to this option. The indices $w$ and $p$ can then be expressed by using the derivatives of $H_G(x)$ with respect to $x$ as follows:

$$w = \sum_{k=1}^{l} k d_k = H'_G(1) \qquad (4)$$

and

$$p = d_3 = H'''_G(0)/6, \qquad (5)$$

or generally

$$d_k = H_G^{(k)}(0) k!. \qquad (6)$$

Now back to $S_G(x)$, general relations for its coefficients have been known[6] as

$$b_0 = 1, \ (head) \qquad (7)$$

$$b_1 = 0, \ (face) \qquad (8)$$

$$-b_2 = \sum_{i<j} D_{ij}^2 = \sum_{k=1}^{l} k^2 d_k = H'_G(1) + H'''_G(1) \quad (neck) \qquad (9)$$

$$-b_3 = 2 \sum_{i<j<k} D_{ij} D_{jk} D_{ki}. \quad (shoulder) \qquad (10)$$

Further, especially for tree graphs, we have

$$-b_N = 2^{N-2}(N-1) \quad (tail) \qquad (11)$$

and

$$-b_{N-1} = 2^{N-3} \sum_{k=2}^{l} g_k (k-1)(2N-k+2) \quad (hip) \qquad (12)$$

where $g_k$ denotes the number of vertices in G with degree $k$.[1]

† Dedicated to Professor Douglas Jay Klein on the occasion of his 70th birthday.

Although all these relations (4–12) have been introduced in Reference 6 and $H_G(x)$ of a vast number of complicated graphs have been studied since then, the mathematical structure of the distance polynomial $S_G(x)$ has not yet thoroughly been analyzed.

The purpose of the present paper is to remind the readers of the interesting mathematical properties of the $S_G(x)$ and to report some recent findings. It is to be noted that, only with special reference to small tree graphs, a number of novel results have been disclosed, however, without rigorous proofs. Almost all of the "Theorems" appearing in this paper were derived from the inspection of a collection of numerical results of various series of targeted graphs. Be aware that unless otherwise stated many of those theorems already introduced in Reference 1 are still conjectures. Although no formal proofs have been obtained, no counter examples have been detected by the present author.

Thus another purpose of this paper is a kind of challenge to call for rigorous proofs for these open questions. The readers are encouraged to send their beautiful proofs to the present author.

## DISTANCE POLYNOMIAL OF PATH PROGRESSION

The $S_G(x)$'s of smaller members of path progression, $P_N$, or the molecular graph representing the carbon atom skeleton of normal paraffin or linear polyene with $N$ carbon atoms, are given in Table 1.

The beautiful general expression of $S_G(x)$ of $P_N$ is given by[1]

**Conjecture 1.** $S_G(x)$ of path progression $P_N$ is expressed by

$$S_N(x) = x^N - \sum_{k=2}^{N} 2^{k-2}(k-1)\frac{N^2(N^2-1^2)(N^2-2^2)\ldots(N^2-\overline{k-1}^2)}{k^2(k^2-1^2)(k^2-2^2)\ldots(k^2-\overline{k-1}^2)}x^{N-k} \quad (13)$$

whose proof may be the most challenging one in this paper.

Equation (13) automatically means that

$$b_0 = 1, b_1 = 0, \text{ and all other } b_k < 0 \ (N \geq k \geq 2). \quad (14)$$

By putting $k = N - 1$ $(0 \leq 1 < N - 1)$ into Equation (13) one gets the following simpler expressions:

**Table 1.** Distance polynomial ($S_G(x)$), the largest eigenvalue ($x_m$), and Wiener number ($w$) of path progression ($S_N$)

| | | | | |
|---|---|---|---|---|
| 2 | $x^2 - 1$ | $x_m = 1$ | $w = 1$ | |
| 3 | $x^3 - 6x - 4$ | $x_m = 2.7321$ | $w = 4$ | $= (x+2)(x^2-2x-2)$ |
| 4 | $x^4 - 20x^2 - 32x - 12$ | $x_m = 5.1623$ | $w = 10$ | $= (x^2+4x+2)(x^2-4x-6)$ |
| 5 | $x^5 - 50x^3 - 140x^2 - 120x - 32$ | $x_m = 8.2882$ | $w = 20$ | $= (x^2+6x+4)(x^3-6x^2-18x-8)$ |
| 6 | $x^6 - 105x^4 - 448x^3 - 648x^2 - 384x - 80$ | $x_m = 12.1093$ | $w = 35$ | $= (x+2)(x^2+8x+4)(x^3-9x^2-36x-20)$ |
| 7 | $x^7 - 196x^5 - 1176x^4 - 2520x^3 - 2464x^2 - $ $-1120x - 192$ | $x_m = 16.6254$ | $w = 56$ | $= (x^3+12x^2+20x+8)\times$ $\times(x^4-12x^3-72x^2-80x-24)$ |
| 8 | $x^8 - 336x^6 - 2688x^5 - 7920x^4 - 11264x^3 - $ $- 8320x^2 - 3072x - 448$ | $x_m = 21.8364$ | $w = 84$ | $= (x^4-16x^3-120x^2-160x-56)\times$ $\times(x^4+16x^3+40x^2+32x+8)$ |
| 9 | $x^9 - 540x^7 - 5544x^6 - 21384x^5 - 41184x^4 - $ $-43680x^3 - 25920x^2 - 8064x - 1024$ | $x_m = 27.7422$ | $w = 120$ | $= (x+2)(x^3+18x^2+24x+8)\times$ $\times(x^5-20x^4-200x^3-400x^2-280x-64)$ |
| 10 | $x^{10} - 825x^8 - 10560x^7 - 51480x^6 - 128128x^5 - $ $-182000x^4 - 153600x^3 - 76160x^2 - $ $-20480x - 2304$ | $x_m = 34.3429$ | $w = 165$ | $= (x+1)(x^4+24x^3+76x^2+64x+16)\times$ $\times(x^5-25x^4-300x^3-700x^2-560x-144)$ |

**Conjecture 2.** For path progression $P_N$, the coefficients of $S_G(x)$ are expressed by

$$
\begin{aligned}
-b_N &= 2^{N-2}(N-1)\\
-b_{N-1} &= 2^{N-2}N(N-2)\\
-b_{N-2} &= 2^{N-4}N(N-3)(2N-3)\\
-b_{N-3} &= 2^{N-5}N(N-4)(2N-4)(2N-5)/3\\
-b_{N-4} &= 2^{N-8}N(N-5)(2N-5)(2N-6)(2N-7)/3\\
-b_{N-5} &= 2^{N-9}N(N-6)(2N-6)(2N-7)(2N-8)(2N-9)/15
\end{aligned} \tag{15}
$$

and so on.

Although it seems to be difficult to find a systematic expression for general $l$, the following conjecture could be derived.

**Conjecture 3.** For path progression $P_N$, the coefficients of $S_G(x)$ are expressed by

$$
\begin{aligned}
-b_N &= 2^{N-2}(N-1),\\
-b_{N-1} &= 2^{N-2}N(N-2),
\end{aligned}
$$

and

$$
\begin{aligned}
-b_{N-l} = \; & 2^{N-l-1}N(N-l-1)(2N-l-2)...\\
& ...(2N-2l+1)/l! \quad (N-2 \geq l \geq 2)
\end{aligned} \tag{16}
$$

Note also Equations (7–10).

**TOPOLOGICAL DEPENDENCY OF THE COEFFICIENTS OF $S_G(x)$**

In Reference 1, the $S_G(x)$'s of tree graphs up to $N \leq 6$ are given. Here in this paper the $S_G(x)$'s of tree graphs of $N = 7$ are given in Table 2, where it is clear that the last term, or the *tails* of all the isomer graphs of $S_7$ are the same as

$$-b_7 = 2^5 \times 6 = 192.$$

This property has already been introduced[1] and ascertained[2] as the following theorem:

**Theorem 1.** All the tree graphs with $N$ vertices have the same *tail* as

$$-b_N = 2^{N-2}(N-1), \tag{17}$$

which is a special case of more general theorem, namely,

**Theorem 2.** The *tail* of the distance polynomial of a given graph is determined only by the number of vertices and the ring skeleton,

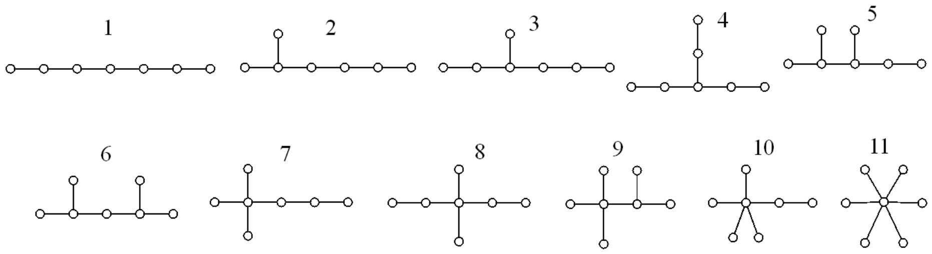
whose proof is given in Reference (3).

As already remarked in Introduction the present author found that the *hip*, $b_{N-1}$, of a tree graph can be expressed by the number of vertices $N$ and the set, $\{g_k\}$, of numbers of vertices with degree $k$ (≥2) as[1]

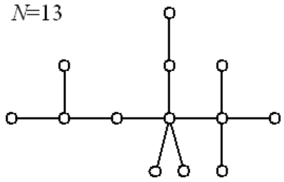
**Table 2.** Distance polynomial and several characteristic quantities of tree graphs of $N = 7$

| | | $b_k$ of $S_G(x)$[(b)] | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| No.[(a)] | $k$ = | 2 | 3 | 4 | 5 | 6 | 7 | factor[(c)] | $w$ | $Z$[(d)] | $p$ | $x_m$[(e)] | $x_l$[(f)] |
| 1 | | 196 | 1176 | 2520 | 2464 | 1120 | 192 | 3+4 | 56 | 21 | 4 | 16.6254 | –10.0978 |
| 2 | | 164 | 976 | 2208 | 2288 | 1088 | 192 | 1+6 | 52 | 18 | 4 | 15.4048 | –10.0990 |
| 3 | | 148 | 928 | 2160 | 2272 | 1088 | 192 | 7 | 50 | 19 | 5 | 14.8636 | – 7.6929 |
| 4 | | 132 | 876 | 2112 | 2256 | 1088 | 192 | 2 2+3 | 48 | 20 | 6 | 14.2969 | –5.2361 |
| 5 | | 120 | 752 | 1840 | 2080 | 1056 | 192 | 2 1+5 | 46 | 17 | 6 | 13.6346 | –6.2227 |
| 6 | | 134 | 804 | 1904 | 2112 | 1056 | 192 | 2 1+2+3 | 48 | 15 | 4 | 14.1760 | –7.4641 |
| 7 | | 122 | 732 | 1752 | 1984 | 1024 | 192 | 2 1+5 | 46 | 14 | 4 | 13.6346 | –6.8246 |
| 8 | | 108 | 680 | 1768 | 1952 | 1024 | 192 | 1+2+4 | 44 | 16 | 6 | 13.0698 | –5.2361 |
| 9 | | 96 | 584 | 1464 | 1776 | 992 | 192 | 3 1+4 | 42 | 13 | 6 | 12.3945 | –5.0279 |
| 10 | | 86 | 516 | 1296 | 1600 | 928 | 192 | 3 1+4 | 40 | 11 | 4 | 11.8281 | –4.3369 |
| 11 | | 66 | 380 | 960 | 1248 | 800 | 192 | 5 1+2 | 36 | 7 | 0 | 10.5678 | –2.0000 |

(a) See below graphs; (b) For No.1 graph it reads as $S_G(x) = x^7 - 196x^5 - 1176x^4 - 2520x^3 - 2464x^2 - 1120x - 192$; (c) Factorization of $S_G(x)$; (d) Hosoya's topological index; (e) Largest eigenvalue; (f) Smallest eigenvalue.



$N$=13



$g_1 = 8$ $d_1 = 12$ $\Delta d_1 = 12 - 12 = 0$

$g_2 = 2$ $d_2 = 21$ $\Delta d_2 = 21 - 11 = 10$

$g_3 = 1$ $d_3 = 22$ $\Delta d_3 = 22 - 10 = 12$

$g_4 = 1$ $d_4 = 15$ $\Delta d_4 = 15 - 9 = 6$

$g_5 = 1$ *etc.*

$S_G(x) = x^{13} - 734x^{11} - 13584x^{10} - 115632x^9 - 574528x^8 - 1823840x^7 - 3852672x^6 - 5491072x^5 - 5248512x^4 - 3277824x^3 - 1267712x^2 - 272384x - 24576$

$b_2 = -734$, $b_{11} = -1267712$, $b_{12} = -272384$, $b_{13} = -24576$,

$\Delta b_{11} = 1530880 - 1267712 = 263168$,

$\Delta b_{12} = 292864 - 272384 = 20480$

Many of these quantities can be obtained as

Equation (4):
$w = 12\times1 + 21\times2 + 22\times3 + 15\times4 + 8\times5 = 220$

Equation (9):
$-b_2 = 12\times1^2 + 21\times2^2 + 22\times3^2 + 15\times4^2 + 8\times5^2 = 734$

Equation (17):
$-b_{13} = 2^{11}\times12 = 24576$

Equation (18):
$-b_{12} = 2^{10}(26\times2 + 50\times1 + 72\times1 + 92\times1) = 272384$

Equation (19):
$\Delta b_{12} = 2^{11}\times10 = 20480$

**Figure 1.** Worked example for enumerating several characteristic quantities of a graph with $N = 13$.

**Conjecture 4.** The *hip* of a tree graph can be expressed by

$$\begin{aligned} -b_{N-1} &= 2^{N-3}\sum_{k=2}^{l} g_k (k-1)(2N-k+2) \\ &= 2^{N-3}[2Ng_2 + 2(2N-1)g_3 + 3(2N-2)g_4 + \\ &\quad +4(2N-3)g_5 + ...]. \end{aligned} \tag{18}$$

The validity of this fantastic formula derived empirically has been verified by a few hundreds of tree graphs.

However, the present author recently found interesting formulas for $b_{N-1}$ and $b_{N-2}$ by using the coefficients of $H_G(x)$ as

**Conjecture 5.** The coefficients $b_{N-1}$ and $b_{N-2}$ of a tree graph can be expressed by

$$\Delta b_{N-1} = 2^{N-2}\Delta d_2 \tag{19}$$

and

$$\begin{aligned} \Delta b_{N-2} = {} & 2^{N-4}(5N-13)\Delta d_2 + 2^{N-3}\Delta d_3 \\ & - 3\cdot 2^{N-3}[g_4 + 4g_5 + 10g_6 + 20g_7 + 35g_8 + \ldots], \end{aligned} \tag{20}$$

where the coefficient of $g_l$ in the square brackets is expressed by $(l-1)(l-2)(l-3)/6$ and

$$\Delta b_k = b_k(\text{isomer}) - b_k(\text{linear}). \tag{21}$$

It is interesting to recall that the $w$ index for $S_N$ graph is given by $wN = N(N^2 - 1)/6$. Then the last term of Equation (20) will be expressed by $-3\cdot 2^{N-3}\sum_{k=4} w_{k-2}g_k$.

The necessary data for graphs with $N = 7$ and 8 in Tables 3 and 4, and a worked example, Figure 1, for a rather complicated tree graph with $N = 13$ are provided in order to clear some skepticism of these interesting formulas (18–20). However, mathematical proof for the connection between Equations (18) and (19) is still open to challenging readers.

In Figure 1, enumeration of $w$ and $b_2$ by the use of $H_G(x)$ (Equations (4) and (9)) is also exemplified, but the cumbersome derivation of $b_3$ with Equation (10) is omitted.

**Table 3.** Topological characteristics of tree graphs of $N = 7$

| | Degree distribution[b] | | | | | | $H_G(x)$[c] | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| No.[a] | $g_1$ | $g_2$ | $g_3$ | $g_4$ | $g_5$ | $g_6$ | $d_1$ | $d_2$ | $d_3$ | $d_4$ | $d_5$ | $d_6$ | $w$ | $Z$[d] | $p$ | $\Delta b_6$[e] | $\Delta b_5$ | $\Delta b_4$ |
| 1 | 2 | 5 | | | | | 6 | 5 | 4 | 3 | 2 | 1 | 56 | 21 | 4 | 0 | 0 | 0 |
| 2 | 3 | 3 | 1 | | | | 6 | 6 | 4 | 3 | 2 | | 52 | 18 | 4 | 32 | 176 | 312 |
| 3 | 3 | 3 | 1 | | | | 6 | 6 | 5 | 3 | 1 | | 50 | 19 | 5 | 32 | 192 | 360 |
| 4 | 3 | 3 | 1 | | | | 6 | 6 | 6 | 3 | | | 48 | 20 | 6 | 32 | 208 | 408 |
| 5 | 4 | 1 | 2 | | | | 6 | 7 | 6 | 2 | | | 46 | 17 | 6 | 64 | 384 | 680 |
| 6 | 4 | 1 | 2 | | | | 6 | 7 | 4 | 4 | | | 48 | 15 | 4 | 64 | 352 | 616 |
| 7 | 4 | 2 | 0 | 1 | | | 6 | 8 | 4 | 3 | | | 46 | 14 | 4 | 96 | 480 | 768 |
| 8 | 4 | 2 | 0 | 1 | | | 6 | 8 | 6 | 1 | | | 44 | 16 | 6 | 96 | 512 | 752 |
| 9 | 5 | 0 | 1 | 1 | | | 6 | 9 | 6 | | | | 42 | 13 | 6 | 128 | 688 | 1056 |
| 10 | 5 | 1 | 0 | 0 | 1 | | 6 | 11 | 4 | | | | 40 | 11 | 4 | 192 | 864 | 1224 |
| 11 | 6 | 0 | 0 | 0 | 0 | 1 | 6 | 15 | | | | | 36 | 7 | 0 | 320 | 1216 | 1560 |

[a] See the footnote (a) of Table 2; [b] The number of vertices with degree $k$; [c] Equation (3); [d] Hosoya's topological index $Z$; [e] Equation (21).

Up to now rigorous relations between the $b$ coefficients and other topological quantities were pursued. Then what will come out by relaxing this strict condition?

If we plot the $b_2$ values against w a rather smooth curve can be obtained, while the similar plot for $b_3$ becomes bumpy. On the other hand, the plots of $b_2$ and $b_3$ values of $N = 7$ isomer graphs against $Z$, as seen in Figure 2, are interesting.

Namely, if the corresponding $p$ number is given to each ($Z$, $b$) point (as discriminated with different kinds of marks, ■, ●, and ▲ in Figure 2) V-shaped lines are evident by connecting the points with the same $p$ values, *e.g.*, 4 and 6, in this case. Further, the lonely point with $p = 5$ seems to lie on the median of V. It is not shown here, but in the ($Z$, $b$)-plots for $N = 8$ isomers several lines connecting the points with common $p$ values seem to emanate from one point.

The above observation reveals that at least the $b_2$ and $b_3$ values of isomer graphs can be estimated rather reliably by some linear combination of $Z$ and $p$, which

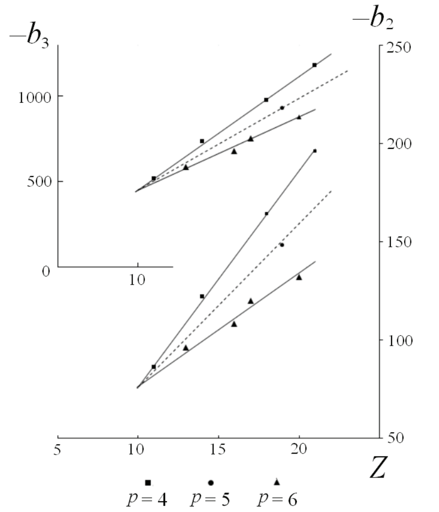
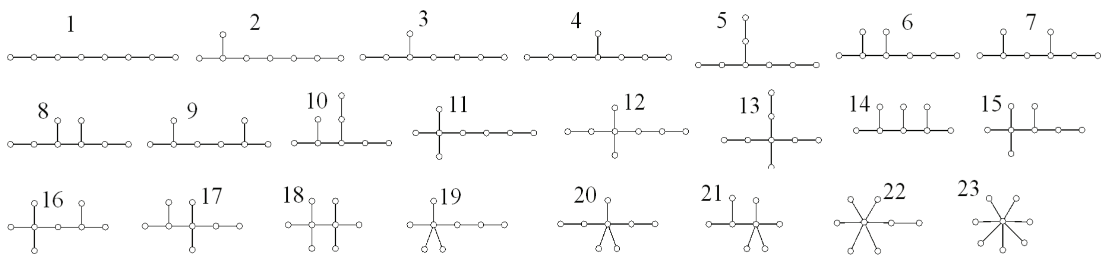



**Figure 2.** The plots of $b_2$ and $b_3$ values of $N = 7$ isomer graphs against $Z$. The points are discriminated by shape according to the Wiener polarity number $p$.

**Table 4.** Topological characteristics of tree graphs of $N = 8$

| | Degree distribution[b] | | | | | | | $H_G(x)$[c] | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| No.[a] | $g_1$ | $g_2$ | $g_3$ | $g_4$ | $g_5$ | $g_6$ | $g_7$ | $d_1$ | $d_2$ | $d_3$ | $d_4$ | $d_5$ | $d_6$ | $d_7$ | $w$ | $Z$[d] | $p$ | $\Delta b_7$[e] | $\Delta b_6$ | $\Delta b_5$ |
| 1 | 2 | 6 | | | | | | 7 | 6 | 5 | 4 | 3 | 2 | 1 | 84 | 34 | 5 | 5 | 0 | 0 |
| 2 | 3 | 4 | 1 | | | | | 7 | 7 | 5 | 4 | 3 | 2 | | 79 | 29 | 5 | 64 | 432 | 1024 |
| 3 | 3 | 4 | 1 | | | | | 7 | 7 | 6 | 4 | 3 | 1 | | 76 | 31 | 6 | 64 | 464 | 1152 |
| 4 | 3 | 4 | 1 | | | | | 7 | 7 | 6 | 5 | 2 | 1 | | 75 | 30 | 6 | 64 | 464 | 1168 |
| 5 | 3 | 4 | 1 | | | | | 7 | 7 | 7 | 5 | 2 | | | 72 | 32 | 7 | 64 | 496 | 1296 |
| 6 | 4 | 2 | 2 | | | | | 7 | 8 | 7 | 4 | 2 | | | 70 | 27 | 7 | 128 | 928 | 2240 |
| 7 | 4 | 2 | 2 | | | | | 7 | 8 | 6 | 5 | 2 | | | 71 | 26 | 6 | 128 | 896 | 2160 |
| 8 | 4 | 2 | 2 | | | | | 7 | 8 | 8 | 4 | 1 | | | 68 | 29 | 8 | 128 | 960 | 2368 |
| 9 | 4 | 2 | 2 | | | | | 7 | 8 | 5 | 4 | 4 | | | 74 | 25 | 5 | 128 | 864 | 2016 |
| 10 | 4 | 2 | 2 | | | | | 7 | 8 | 8 | 5 | | | | 67 | 28 | 8 | 128 | 960 | 2384 |
| 11 | 4 | 3 | 0 | 1 | | | | 7 | 9 | 5 | 4 | 3 | | | 71 | 23 | 5 | 192 | 1200 | 2624 |
| 12 | 4 | 3 | 0 | 1 | | | | 7 | 9 | 7 | 4 | 1 | | | 67 | 25 | 7 | 192 | 1264 | 2848 |
| 13 | 4 | 3 | 0 | 1 | | | | 7 | 9 | 9 | 3 | | | | 64 | 28 | 9 | 192 | 1328 | 3056 |
| 14 | 5 | 0 | 3 | | | | | 7 | 9 | 8 | 4 | | | | 65 | 24 | 8 | 192 | 1392 | 3096 |
| 15 | 5 | 1 | 1 | 1 | | | | 7 | 10 | 8 | 3 | | | | 63 | 22 | 8 | 256 | 1728 | 3904 |
| 16 | 5 | 1 | 1 | 1 | | | | 7 | 10 | 5 | 6 | | | | 66 | 19 | 5 | 256 | 1632 | 3584 |
| 17 | 5 | 1 | 1 | 1 | | | | 7 | 10 | 9 | 2 | | | | 62 | 23 | 9 | 256 | 1760 | 3904 |
| 18 | 6 | 0 | 0 | 2 | | | | 7 | 12 | 9 | | | | | 58 | 17 | 9 | 384 | 2528 | 5152 |
| 19 | 5 | 2 | 0 | 0 | 1 | | | 7 | 12 | 5 | 4 | | | | 62 | 17 | 5 | 384 | 2208 | 4448 |
| 20 | 5 | 2 | 0 | 0 | 1 | | | 7 | 12 | 8 | 1 | | | | 59 | 20 | 8 | 384 | 2304 | 4688 |
| 21 | 6 | 0 | 1 | 0 | 1 | | | 7 | 13 | 8 | | | | | 57 | 16 | 8 | 448 | 2736 | 5408 |
| 22 | 6 | 1 | 0 | 0 | 0 | 1 | | 7 | 16 | 5 | | | | | 54 | 13 | 5 | 640 | 3360 | 6176 |
| 23 | 7 | 0 | 0 | 0 | 0 | 0 | 1 | 7 | 21 | | | | | | 49 | 8 | 0 | 960 | 4400 | 7456 |

[a] See below graphs; [b] The number of vertices with degree $k$: [c] Equation (3). [d] Hosoya's topological index $Z$; [e] Equation (21).



have been known as the pair of the most uncorrelated molecular descriptors.[13–15] Thus, a number of interesting features seem to be hidden in the distance polynomial, $S_G(x)$, of a graph.

## SYMMETRY-SENSITIVE PROPERTY OF $S_G(x)$

As shown in Table 2, among the eleven isomeric tree graphs of $N = 7$ the $S_G(x)$ of only a single member, 7-3, is not factored out, while eight isomers have at least a factor $(x + 2)$, and the highest symmetrical star graph, $K_{1,6}$, or 7-11, has a quintuply degenerate factor of $(x + 2)$. The $S_G(x)$ of linear 7-1 is the product of two polynomials of order 3 and 4. It is to be noted that two isomers, 7-4 and 7-8, are found to have the common factor of $(x^2 + 6x + 4)$ (See Figure 3).

Further, although raw data are not given in this paper, the $S_G(x)$ of only a single member among 23 isomers of $N = 8$ tree graphs is not factorable, while 17 members have at least one $(x + 2)$ factor, and four members shown in Figure 3 are found to have the common factor of $(x^2 + 6x + 4)$.

The reason why so many isomer graphs have the common factor of $(x + 2)$ is not yet clarified. However, the reason why many trees have the common factor of $(x^2 + 6x + 4)$, which is one of the factors of $S_G(x)$ of path progression $S_5$ (See Table 1), can be explained as follows.

Consider first the determinant $-\det(\mathbf{D} - x\mathbf{E})$ of $S_5$ as

$$-\det(\mathbf{D}-x\mathbf{E}) = \begin{vmatrix} -x & 1 & 2 & 3 & 4 \\ 1 & -x & 1 & 2 & 3 \\ 2 & 1 & -x & 1 & 2 \\ 3 & 2 & 1 & -x & 1 \\ 4 & 3 & 2 & 1 & -x \end{vmatrix} = f(\mathbf{D}), \quad (22)$$

where the "atomic orbitals", $\chi_1$–$\chi_5$, are numbered consecutively from left to right or *vice versa*. Let us follow the basic idea and technique of factorization of the determinant of Hückel molecular orbitals by using the symmetry adapted "group orbitals". Here, the diagonal element of $-x$ in Equation (22) stands for the expected value $\langle \chi_a |\mathbf{h}| \chi_a \rangle$ for a certain Hamiltonian $\mathbf{h}$, of which we need not scrutinize its physical meaning, and the off-diagonal element $D_{ab}$ in Equation (22) stands for the corresponding term of $\langle \chi_a |\mathbf{h}| \chi_b \rangle$.

Let us transform Equation (22) into the product of symmetrical and antisymmetrical determinants, which are based on the set of group orbitals. Here, we need only the antisymmetrical group orbitals $\varphi_{A1}$ and $\varphi_{A2}$ as

$$\varphi_{A1} = \frac{1}{\sqrt{2}}(\chi_1 - \chi_5) \quad (23)$$

and

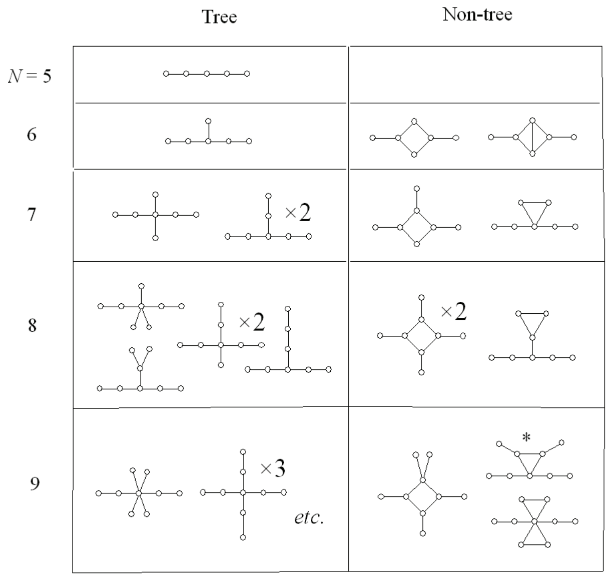



**Figure 3.** Several groups of graphs with common factor $(x^2 + 6x + 4)$ in $S_G(x)$. The multiplicity of this factor is also indicated. The asterisked graph is referred to Figure 4.

$$\varphi_{A2} = \frac{1}{\sqrt{2}}(\chi_2 - \chi_4), \quad (24)$$

by omitting the central atomic orbital $\chi_3$ while by using all the five atomic orbitals the symmetrical component is constructed from the three symmetrical group orbitals, which, however, is not necessary in this discussion.

Then the antisymmetrical matrix elements to be factored out are obtained to be

$$\langle \varphi_{A1} |\mathbf{h}| \varphi_{A1} \rangle = \frac{1}{2}\left( \langle \chi_1 |\mathbf{h}| \chi_1 \rangle + \langle \chi_5 |\mathbf{h}| \chi_5 \rangle - 2\langle \chi_1 |\mathbf{h}| \chi_5 \rangle \right)$$
$$= \frac{1}{2}(-x - x - 2\cdot 4) = -x - 4$$

$$\langle \varphi_{A2} |\mathbf{h}| \varphi_{A2} \rangle = \frac{1}{2}\left( \langle \chi_2 |\mathbf{h}| \chi_2 \rangle + \langle \chi_4 |\mathbf{h}| \chi_4 \rangle - 2\langle \chi_2 |\mathbf{h}| \chi_4 \rangle \right)$$
$$= \frac{1}{2}(-x - x - 2\cdot 2) = -x - 2$$

and

$$\langle \varphi_{A1} |\mathbf{h}| \varphi_{A2} \rangle = \langle \varphi_{A2} |\mathbf{h}| \varphi_{A1} \rangle$$
$$= \frac{1}{2}\left( \langle \chi_1 |\mathbf{h}| \chi_2 \rangle - \langle \chi_1 |\mathbf{h}| \chi_4 \rangle - \right.$$
$$\left. - \langle \chi_5 |\mathbf{h}| \chi_2 \rangle + \langle \chi_5 |\mathbf{h}| \chi_4 \rangle \right)$$
$$= \frac{1}{2}(-1 - 3 - 3 + 1) = -2$$

giving

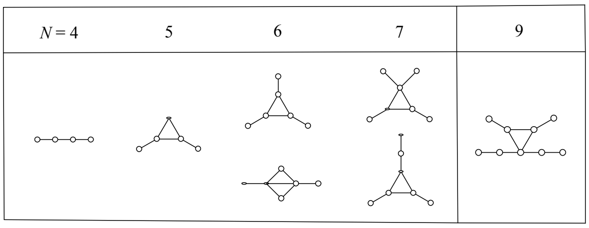
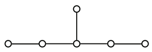
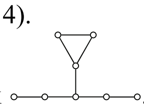



**Figure 4.** Several groups of graphs with common factor ($x^2 + 4x + 2$).

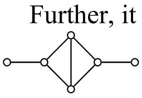
$$\begin{vmatrix} -x-4 & -2 \\ -2 & -x-2 \end{vmatrix} = x^2 + 6x + 4. \qquad (25)$$

This means that $f(\mathbf{D})$ of the $S_5$ graph contains ($x^2$ + $6x$ + 4) as a factor.

Next consider a network  which contains the skeleton of $S_5$ and keeps the same mirror symmetry. Although its determinant $f(\mathbf{D})$ is different from that of $S_5$, its antisymmetrical component is the same as that of $S_5$, meaning that the $S_G(x)$ polynomial of this network also has a factor of ($x^2 + 6x + 4$).

Then try to consider another network , which is non-tree but also contains the $S_5$ skeleton and keeps the same mirror symmetry as $S_5$. Although the antisymmetrical component of this graph has an extra group orbital, it is shown not to interact with the 2×2 determinant (25) of $S_5$.

Thus one can conclude that if a graph (tree or non-tree) contains the $S_5$ skeleton and still keeps the same mirror symmetry with $S_5$, its $S_G(x)$ has a chance for having the factor of ($x^2$+$6x$+4).

Further, it can be shown that a non-tree graph such as , which does not meet the above condition, also has the same ($x^2$+$6x$+4) factor in its $S_G(x)$ polynomial. This property can easily be proved by checking its $f(\mathbf{D})$.

In Figure 4 are shown many graphs with ($x^2$+$4x$+2) factor in $S_G(x)$.

As seen in Table 1, the $S_G(x)$ of $S_4$ graph has a factor of ($x^2$+$4x$+2). Although there is no other tree graph containing this factor in $S_G(x)$, many non-tree graphs shown in Figure 4 have this factor in common. This property can be explained similarly as in the above discussion.

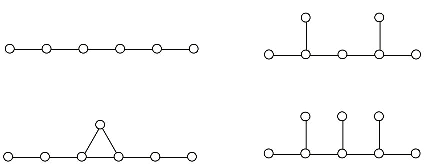
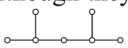
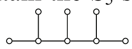


**Figure 5.** Several groups of graphs with common factor ($x^2 + 8x + 4$).

An interesting graph is the one with $N = 9$ given in Figure 4 (also marked with an asterisk in Figure 3), which has both the factors of ($x^2$+$4x$+2) and ($x^2$+$6x$+4). This property can be explained by such a structure of the graph that is obtained by the symmetrical and loose joining of $S_4$ and $S_5$.

By checking the $f(\mathbf{D})$ of small trees another interesting group of tree graphs was found as in Figure 5. Namely, although they contain the $S_5$ skeleton, both the $S_G(x)$'s of  and  have a common factor ($x^2$+$8x$+4) instead of ($x^2$+$6x$+4). Note that the symmetry of the former graph should be considered to be $D_{2v}$, *i.e.*, the rectangular form. Then, irrespective of the existence of the central branch in the latter graph, the $S_G(x)$'s of both the graphs are shown to have a common factor of ($x^2$+$8x$+4).

We have thus shown that the $S_G(x)$ polynomial has such an interesting symmetry-sensitive property. Since this kind of analysis has not yet fully been performed for more complicated graphs, interested readers are encouraged to analyze this problem deeply, together with the relation between $S_G(x)$ and $H_G(x)$ and other topological quantities of graphs.

## SPECTRA OF $S_G(x)$

Since all the coefficients of $S_G(x)$ are negative except for the *head*, $x^N$, there is only one positive eigenvalue whose absolute value is the largest among other negative solutions. To the knowledge of the present author, all the eigenvalues of $S_G(x)$ of any graph are real, but this property does not yet seem to have been proved rigorously.

In Figure 6, all the eigenvalues of $S_G(x)$'s of the series of path progressions with $N = 1 \sim 6$ are plotted.

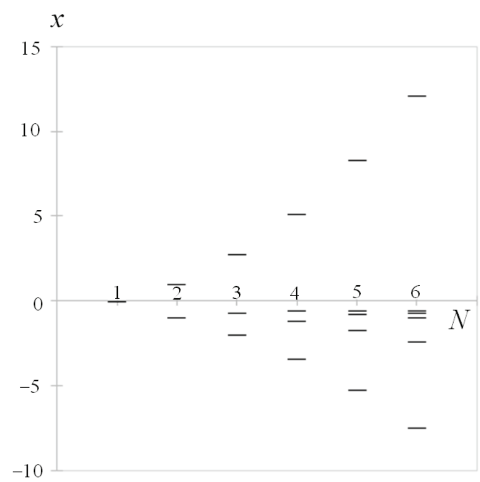



**Figure 6.** Spectra of $S_G(x)$'s of the series of path progressions $P_N$'s with $N$ = 1–6.

**Table 5.** The $S_G(x)$'s of smaller monocyclic graphs

| | | $-b_k$[(a)] | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| $N$ | $k$ = | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| 3 | | 3 | 2 | | | | | | |
| 4 | | 12 | 16 | | | | | | |
| 5 | | 25 | 60 | 35 | 6 | | | | |
| 6 | | 57 | 200 | 144 | | | | | |
| 7 | | 98 | 490 | 707 | 434 | 119 | 12 | | |
| 8 | | 176 | 1152 | 1984 | 1024 | | | | |
| 9 | | 270 | 2286 | 5913 | 7110 | 4521 | 1566 | 279 | 20 |
| 10 | | 925 | 4440 | 13360 | 15744 | 6400 | | | |

[(a)] For $N = 5$ it reads as $S_5(x) = x^5 - 25x^3 - 60x^2 - 35x - 6$.

The remarkable features of these spectra are sharp rising and falling of the largest and smallest eigenvalues in contrast to the packing of all others just below the zero line. As a matter of fact, the landscape of Figure 6 is not interesting either from a mathematical or a chemical point of view. However, an interesting feature was found for the largest eigenvalues, $x_m$'s, of the isomeric tree graphs.

See Figure 7, where $x_m$'s of eleven $N = 7$ tree graphs are plotted against $w$. Furthermore for other isomeric graphs almost a straight-line correlation is observed. For isomeric non-tree graphs almost linear but slightly lower correlation is observed. Similarly good linear correlation can be seen for either of $N = 8$ isomeric tree or non-tree graphs. Anyway, the good linear correlation between $x_m$ and $w$ for isomeric graphs is also one of the most challenging problems in this paper to be analyzed.

Although in this paper emphasis was focused on the tree graphs, interesting, though complex, results for the distance polynomial $S_G(x)$ and its relation to the $H_G(x)$ polynomial have been obtained for non-tree graphs. In concluding this paper, together with the $S_G(x)$'s of smaller monocyclic graphs (Table 5), three conjectures for $S_G(x)$ of non-tree graphs are given, all of which are still waiting for their rigorous proof.

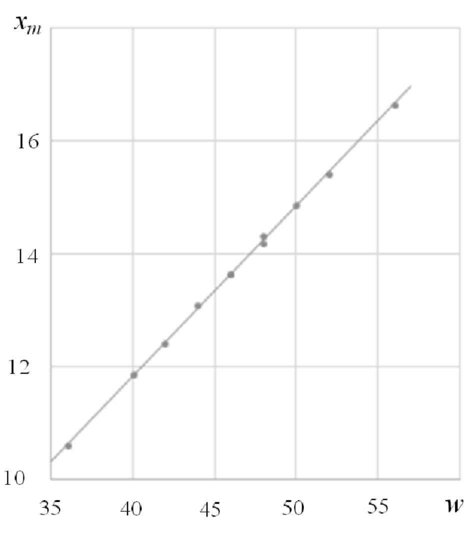



**Figure 7.** Linear correlation between the largest eigenvalue, $x_m$, of $S_G(x)$ and the $w$ index for $N = 7$ tree graphs.

**Conjecture 6.** The $S_G(x)$ of $N$-membered monocyclic graph is given by

$$S_N(x) = x^{2m} - \sum_{k=2}^{m} 2^{2k-3}[m^2(2k^2-k-2)+2(k-1)^2]\times \frac{m^2(m^2-1^2)(m^2-2^2)\cdots(m^2-\overline{k-2}^2)x^{2m-k}}{k^2(k^2-1^2)(k^2-2^2)\cdots(k^2-\overline{k-2}^2)(k^2-\overline{k-1}^2)} - 2^{2(m-1)}m^2x^{m-1}$$
$$= x^{m-1}(x-m^2)\left[x^m + \sum_{k=1}^{m} 2^{2(k-1)}\frac{m}{k}\binom{m+k-1}{2k-1}x^{m-k}\right] \quad (N = 2m) \tag{26}$$

$$S_N(x) = [x - m(m+1)]\left[\sum_{k=0}^{m}\binom{m+k}{2k}x^{m-k}\right]^2 \cdot (N = 2m+1) \tag{27}$$

**Conjecture 7.** The *tail* of $S_G(x)$ of $N$-membered cycle graph is given by

$$-2^{N-4}N^2x^{\frac{N}{2}-1} \quad (N = 2m) \tag{28}$$

$$-(N^2-1)/4. \quad (N = 2m+1) \tag{29}$$

**Conjecture 8.** The tail of $S_G(x)$ of a graph with $N$ vertices and an $N$-membered cycle is given by

$$-2^{N-3}[MN - M^2/2]x^{\frac{M}{2}-1} \quad (M = 2m) \tag{30}$$

$$-2^{N-M-1}[MN - (M^2+1)/2]. \quad (M = 2m+1) \tag{31}$$

## CONCLUDING REMARK

We have seen so many interesting features of the distance polynomial of tree graphs, the proof of which is generally so difficult and thus challenging. Similar but more complicated situation is expected also for non-tree

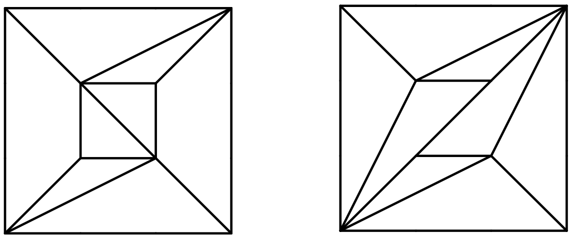


**Figure 8.** The pair of twin graphs.

$P_G(x) = x^8 - 15x^6 - 12x^5 + 41x^4 + 28x^3 - 35x^2 - 12x + 9,$

$S_G(x) = x^8 - 67x^6 - 340x^5 - 643x^4 - 436x^3 + 21x^2 + 76x - 7,$

$M_G(x) = x^8 - 15x^6 - 61x^4 - 65x^2 - 9, \;\; Z = 151.$[17]

graphs. In this respect, Balasubramanian has shown that the distance polynomial can be used for characterizing the fullerene isomers.[16] Then, it seems that the redundancy of the distance polynomial, irrespective of carrying many digits of integers, is rather less prominent than the case of the characteristic polynomial.

On the other hand, we have found a pair of "twin graphs" which have not only the same characteristic polynomial but also the same distance and matching polynomials as shown in Figure 8.[17] From these controversial and mystic properties, the present author believes that there must be hidden mathematically beautiful gems in the mine of the still unexplored distance polynomial. This paper was written as an invitation letter to those who have some curious interest in chemistry and mathematics.

## REFERENCES


1. H. Hosoya, M. Murakami, and M. Gotoh, *Natl. Sci. Rept. Ochanomizu Univ.* **24** (1973) 27–34.
2. M. Edelberg, M. R. Garey, and R. L. Graham, *Discrete Math.* **14** (1976) 23–39.
3. R. L. Graham and L. Lovász, *Adv. Math.* **29** (1978) 60–88.
4. R. L. Graham, A. H. Hoffman, and H. Hosoya, *J. Graph Theory* **1** (1977) 85–88.
5. H. Wiener, *J. Am. Chem. Soc.* **69** (1947) 17–20.
6. H. Hosoya, *Discrete Appl. Math.* **19** (1988) 239–257.
7. M. Lepovic and I. Gutman, *J. Chem. Inf. Comput. Sci.* **38** (1998) 823–826.
8. I. Gutman, E. Estrada, and O. Ivanciuc, *Graph Theory Notes (New York)* **36** (1999) 7–13.
9. A. R. Ashrafi and M. Ghorbani, *Digest J. Nanomater. Biostruct.* **4** (2009) 389–393.
10. S. Xu, H. Zhang, and M. V. Diudea, *MATCH, Commun. Math. Comput. Chem.* **57** (2007) 443–456.
11. A. M. Ali, A. A. Ali, and T. H. Ismail, *Hosoya Polynomials and Wiener Indices of Distances*, in: *Graphs*, Lambert, Saarbrücken, Germany (2011).
12. H. Hosoya, *Bull. Chem. Soc. Jpn.* **44** (1971) 2332–2339.
13. H. Narumi and H. Hosoya, *Bull. Chem. Soc. Jpn.* **58** (1985) 1778–1786.
14. Y.-D. Gao and H. Hosoya, *Bull. Chem. Soc. Jpn.* **61** (1988) 3093–3102.
15. H. Hosoya and Y.-D. Gao, *Topology*, in: D. H. Rouvray and R. B. King (Eds.), *Chemistry: Discrete Mathematics of Molecules*, Horwood Publ., Coll House, England (2002), pp. 38–57.
16. K. Balasubramanian, *J. Phys. Chem.* **99** (1995), 10785–10796.
17. H. Hosoya, U. Nagashima, and S. Hyugaji, *J. Chem. Inf. Comput. Sci.* **34** (1994), 428–431.