match no. 22 pp. 77-100 1987

ALGEBRAIC EXPRESSIONS FOR KEKULÉ STRUCTURE COUNTS OF NON-BRANCHED REGULARLY CATA-CONDENSED BENZENOID HYDROCARBONS

Alexandru T.Balaban [a], Cristian Artemi [b] and

Ioan Tomescu [b]

[a] Department of Organic Chemistry,Polytechnic Institute - TCH, 76206 Bucharest, Roumania

[b] Department of Mathematics,University of Bucharest, 70109 Bucharest, Roumania

(Received: December 1986)

Abstract. For non-branched catafusenes which have alternatively $r = 2p$ or $2p+1$ times $a+1$ and $b+1$ linearly condensed benzenoid rings in linear stretches (segments) separated by kinks,algebraic formulas are found,first by combinatorial methods and then as a particular case of the Gordon-Davison algorithm.When $a = b$,the formulas reduce to expressions already known from an earlier paper.The problem is generalized to any sequence of linearly condensed benzenoid rings of various lengths a_i+1,repeated p times.

NOTATION

a, b, a_i = integers representing the numbers of linearly condensed benzenoid rings (without counting the ring at the "kink")

$A_{k,s}$ = coefficient of polynomial R corresponding to term $a^s b^{k-s}$

$d_i = d'_{i-1}$ = difference between consecutive indices appearing in the terms of polynomial R

D_p = difference between numbers of Kekulé structures of catafusenes differing by one linear string of condensed benzenoid rings

$E(T)$ = the set of edges in tree $T(G^*)$

f = injective function for labelling edges of the tree $T(G^*)$

g_i (i=1 or 2)= constants in the expression of D_p

G = molecular graph of catafusene

G^* = dualist graph of catafusene

h_i(i=1 or 2)= constants in the expression of K_{2p}

i_k, j, k = natural numbers, often indices

$I(T)$ = the set of all pairs of incompatible edges in $T(G^*)$

I_s = domain of values for k

K = number of Kekulé structures

K_r = idem for r strings of linearly condensed benzenoid rings

$K_{j,r}$ = idem for the case when each string contains j benzenoid rings

$\mathbb{N}$	= the set of natural numbers
$N_{d_o,j}$	= number of strictly increasing strings $s_1,\ldots,s_{k-1}$ where $s_{k-1}=j$
N_{d_o}	= sum of the above numbers over the range of j values
N	= sum of the above numbers over the range of d_o values
p	= number of repeated sequences of linearly condensed strings,i.e.the integer part of r/s
P	= polynomial associated with the tree $T(G^*)$ furnishing the number K(G) of Kekulé structures
q	= increment with which the number of Kekulé structures increases arithmetically in a linear string according to the Gordon-Davison algorithm[1]
r	= number of linear strings of benzenoid rings in a catafusene
$Q_1^{(i)}$ where i=1 or 2	= constants in the expression of K_{sp+i}
R,R_1	= polynomials associated with the tree $T(G^*)$ and described in detail in ref.[15]
R_s	= polynomial used in the recurrence which gives the number K_{sp+i} of Kekulé structures
s	= in combinatorial relationships (§ 2),s is a natural number denoting the rank;in § 3,s represents the number of linear strings which are repeated p times to afford the zig-zag regular catafusene

S_i	= alternating sum of the numbers of Kekulé structures for binary regularly condensed catafusene
$t = r-ps$	= number of linear strings which appear only once at the tail of a regularly condensed catafusene
$T(G^*)$	= tree ("isoarithmicity tree" in ref.[15]) corresponding to dualist graph G^*
u,v	= edges in $T(G^*)$ belonging to $E(T)$
x_{d_o}	= $2p-d_o+k-1$
X_i	= variables
y	= term $X_{i_1}X_{i_2} \dots X_{i_j}$ of polynomial R_1
$\mathbb{Z}$	= set of all integers
$Z(u,v)$	= set of edges in the path between edges u and v
$\lfloor x \rfloor$	= integer part of x
$\lceil x \rceil$	= smallest integer not less than x
$\overline{x,y}$	= set of natural numbers which are larger than or equal to x, and smaller than or equal to y

1. INTRODUCTION

The algorithm proposed by Gordon and Davison [1] allows easily the calculation of the number of Kekulé structures (Kekulé structure count) for any cata-condensed benzenoid polycyclic aromatic hydrocarbon (PAH). However, it does not provide the possibility of finding algebraic formulas for systems which possess regularities in their structure, nor does it lend itself readily for recurrent formulas. Several authors [2-11] have elaborated different formulas for particular classes of PAH's. Computer programs have been devised for calculating the number of Kekulé structures.[9-14] Two earlier papers in this series [15,16] presented algebraic formulas for the Kekulé structure count of non-branched cata-condensed PAH's which have equal numbers of hexagons in each linear portion of a zig-zag catafusene (cata-condensed PAH), or any of its isoarithmic congeners (i.e. systems differing only in the direction of kinks in the dualist graph).

We generalize here this problem in two stages: first for catafusenes having alternatively $a+1$ and $b+1$ linearly condensed benzenoid rings, and then to regularly condensed catafusenes, i.e. to systems having repeating sequences with $a_0+1, a_1+1, \ldots, a_s+1$ benzenoid rings in the linear portions separated by kinks (where $s \in \mathbb{N}, s \geq 2$). We call the former stage of this generalization "binary regularly condensed catafusenes".

A benzenoid ring where a kink in the annelation occurs (in a non-branched catafusene) may form a bay-region.[17,18] Such a benzenoid ring is considered to belong to both linear stretches originating in that kink;this is why the lengths of such stretches are a+l,b+l,or in general a_i+1,as will be seen below.A bay-region is formed if a=l or in general if $a_0=1$.

2.BINARY REGULARLY CONDENSED CATAFUSENES

2.1.DEFINITIONS AND SUMMARY OF PREVIOUS DATA

We associate to each catafusene a dualist graph according to the well-established procedure.[19,20] In binary regularly condensed catafusene we have only two alternating lengths of linearly condensed segments,denoted by a+l and b+l,respectively.Examples are presented in Fig.l by formulas 1 - 3 which are all isoarithmic among them with a = 2 and b = 1,and by the two isoarithmic formulas 4 - 5 which have a = 3 and b = 2.

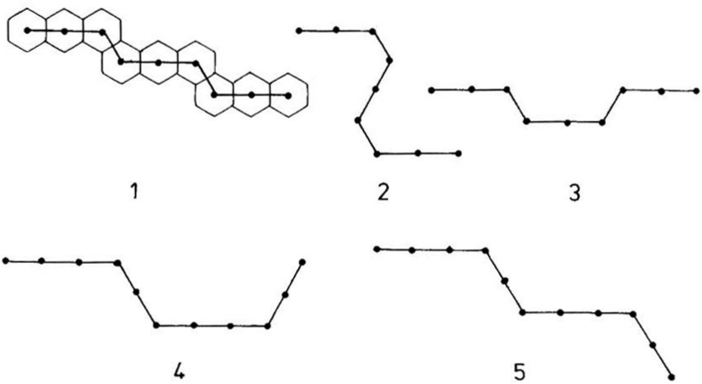

Fig.l. Binary regularly condensed catafusenes: 1 with the molecular graph G and the dualist graph G^*,but only G^* for 2-5.

To a catafusene with its graph G and its dualist graph G^*, we associate a tree $T(G^*)$ with r+1 vertices and r edges, corresponding to the r linear segments of G. The weight a_i of edge i is the number of benzenoid rings minus one, in the respective linear portion of the catafusene. We denote by E(T) the set of edges in the tree $T(G^*)$, i.e. the cardinal of E(T) is r. With the injective function [15]

$$f: E(T) \longrightarrow X = \{X_1, X_2, \ldots, X_r\}$$

we obtain a labelling of the edges with variables from X. We denote by Z(u,v) the set of edges of the path between edges u and v in $T(G^*)$. Two edges are called incompatible if the path connecting them (this path is unique because it belongs to a tree) has an odd length, i.e. an odd number of constituent edges. Let I(T) denote the set of all pairs of incompatible edges in $T(G^*)$. For each tree T the following polynomial is defined:[15]

$$R_1(T;X_1,X_2,\ldots,X_r) = \sum_{(u,v)\in I(T)} f(u)f(v) \prod_{\substack{w \notin Z(u,v) \\ w \neq u,v}} (f(w)+1)$$

By developing $R_1(T;X_1,X_2,\ldots,X_r)$ as a sum of products of variables $X_1,X_2,\ldots,X_r$, and by applying the idempotency rule for addition: $y+y = y$, where $y = X_{i_1}X_{i_2} \ldots X_{i_j}$, we obtain the polynomial $R(T;X_1,\ldots,X_r)$. We define:

$$P(T;X_1,X_2,\ldots,X_r) = \prod_{u\in E(T)} (f(u)+1) - R(T;X_1,X_2,\ldots,X_r)$$

The following theorem was demonstrated in ref.[15] :

<u>Theorem 1.</u> For a catafusene having the hydrogen-depleted molecular graph G consisting of r linear branches which contain

a_i+1 benzenoid rings, the Kekulé structure count is

$$K(G) = P(T(G^*);a_1,a_2,\ldots,a_r)$$

where a_i is the weight of edge i.

When all r linear segments have equal length (containing each j+1 benzenoid rings), in ref.[21] it was demonstrated the following recurrence for the numbers $K_{j,r}$ of Kekulé structures:

$$K_{j,r+2} = jK_{j,r+1} + K_{j,r} \qquad \forall j \geq 1, r \geq 1 \qquad (1)$$

In the case when j=1 this relationship becomes the Fibonacci recurrence.[1,10,21]

Now we go on for generalizing this result to the case of binary linearly condensed catafusenes, with r alternating linear stretches having a+1 and b+1 benzenoid rings. We denote in this case the polynomial P(T;a,b) by P(a,b;r).

Let $K_r(a,b) = P(a,b;r)$ - according to Theorem 1 - denote the number of Kekulé structures for catafusenes having r alternating linear portions with a+1 and b+1 benzenoid rings. We shall discuss separately two cases according to the parity of r. In this section we shall simplify the notation by using K_r instead of $K_r(a,b)$. By definition, $K_0=2$ as in benzene. We shall denote by $\lfloor x \rfloor$ the integer part of x and by $\lceil x \rceil$ the smallest integer not less than x.

2.2. CASE 1: r IS EVEN

Let r = 2p, where p is a non-zero natural number, e.g. catafusenes **4** or **5**. The p edges of odd rank have weight a, and the p edges of even rank have weight b. We have

$$K_r = P(a,b;2p) = (a+1)^p(b+1)^p + 1 - R(a,b;2p)$$

Let the rank be $s \subset \overline{0,p}$. We define I_s as:

$$I_s = \{ k \in \mathbb{N} \mid k \in \overline{s+2,s+p} \} \quad \text{if } s \in \{0,1\}$$

$$I_s = \{ k \in \mathbb{N} \mid k \in \overline{p,2p-1} \} \quad \text{if } s = p$$

and $$I_s = \{ k \in \mathbb{N} \mid k \in \overline{s,s+p} \} \quad \text{if } 1 \leqslant s \leqslant p$$

We wish to find a combinatorial formula for

$$R(a,b;2p) = \sum_{\substack{0 \leqslant s \leqslant p \\ k \in I_s}} A_{k,s}\, a^s b^{k-s} \quad , \; A_{k,s} \in \mathbb{N}$$

A term $A_{k,s}a^s b^{k-s}$ in the above sum is obtained from the terms of the polynomial R which have the form $a_{i_1} \ldots a_{i_k}$.

From the k indices, s are odd and k-s are even. There exists a bijection between the terms of R and the set of strictly increasing strings $i_1, i_2, \ldots, i_k$ which have the property (*):

(*) at least two consecutive terms in this string have the same parity.

This fact follows from the definition of R_1 according to which each term must contain the weights of two incompatible edges and, possibly, of other edges, but in this case excluding all edges lying on the path connecting the two incompatible edges. In the case under discussion two edges are incompatible if they have weights with indices sharing the same parity. Also, $u_k \notin Z(u_i, u_j)$, $i < j$, iff $k \leqslant i$ or $k \geqslant j$.

We shall now calculate coefficients $A_{k,s} \in \mathbb{N} \setminus \{0\}$.

a) If $s \neq \lfloor k/2 \rfloor$ and $s \neq \lceil k/2 \rceil$, then any strictly increasing string with s odd-rank terms verifies property (*). This can be demonstrated by reduction ad absurdum. Hence $A_{k,s}$, the number of monomials in R having s odd-rank terms, is equal to the number of strictly increasing strings with k natural numbers ($k \leqslant r-1$). Hence

$$A_{k,s} = \text{Card}\left\{ t \in \mathbb{N}^k \mid t = i_1, i_2, \ldots, i_k, 1 \leqslant i_j \leqslant r, \forall j \in \overline{1,k} \text{ and } i_j < i_{j+1} \quad \forall j \in \overline{1,k-1} \right\}$$

We observe that any strictly increasing string $i_1, \ldots, i_k$ is uniquely decomposable into two strictly increasing substrings containing odd and even numbers, respectively. Viceversa, two such strictly increasing substrings may be combined uniquely into an increasing string. Because in the set $\overline{1,2p} = \{1,2,\ldots,2p\}$ there exist p even and p odd numbers, it results that there are $\binom{p}{s}$ strictly increasing substrings of s odd numbers which are smaller than 2p, and $\binom{p}{k-s}$ strictly increasing substrings of k-s even numbers which are smaller than, or equal to, 2p. It follows that

$$A_{k,s} = \binom{p}{s}\binom{p}{k-s}$$

b) If k is even and $s = k/2$, we shall calculate first the number of strictly increasing strings $i_1, \ldots, i_k$ of natural numbers which are smaller than, or equal to, 2p and which contain $k/2$ even numbers without verifying property (*). Such a string contains alternating odd and even numbers. We denote $d_o = i_1$ and we observe that $1 \leqslant d_o \leqslant 2p-k+1$. We also denote

$$d_1 = i_2 - i_1, \ \ldots \ , d_{k-1} = i_k - i_{k-1}$$

For any i ($1 \leqslant i \leqslant k-1$), d_i is odd. From $i_k = d_o + d_1 + \ldots + d_{k-1} \leqslant 2p$, it results that $d_1 + \ldots + d_{k-1} \leqslant 2p - d_o$.

Let d_i' denote even numbers, $d_i' = d_i + 1$, $d_i' \geqslant 2$. Then $d_1' + d_2' + \ldots + d_{k-1}' \leqslant 2p - d_o + k - 1$. With the notation:

$$d_1' = s_1, d_1' + d_2' = s_2, \ldots, \sum_{i=1}^{k-1} d_i' = s_{k-1} = j \leqslant 2p - d_o + k - 1.$$

Therefore if d_o is fixed ($d_o \in \overline{1,2p-k+1}$) we have established a bijection between the sets of strings $d_1,\ldots,d_{k-1}$ and $s_1,\ldots,s_{k-1}$ where $s_1,\ldots,s_{k-1}$ is a strictly increasing string of even numbers smaller than, or equal to, $2p-d_o+k-1$. Thus for a fixed d_o, the number of strictly increasing strings $i_1,\ldots,i_k$ of alternating odd and even numbers with $i_1 = d_o$ is equal to the number (denoted by N_{d_o}) of strictly increasing strings $s_1,\ldots,s_{k-1}$ formed by nonzero even numbers which are less than, or equal to, $x_{d_o} = 2p-d_o+k-1$:

$$N_{d_o} = \sum_{j=k-1}^{\lfloor x_{d_o}/2\rfloor} N_{d_o,j} = \sum_{j=k-1}^{\lfloor x_{d_o}/2\rfloor} \binom{j-1}{k-2}$$

where $N_{d_o,j}$ denotes the number of strictly increasing strings $s_1,\ldots,s_{k-1}$ **having the above-mentioned properties and** $s_{k-1}=j$.

Let $N = \sum_{d_o=1}^{2p-k+1} N_{d_o}$ denote the number of strictly increasing strings $i_1,\ldots,i_k$ having alternating odd and even numbers which are **less** than, or equal to, $2p$. We obtain from the sums, after manipulating combinatorial formulas:

$$N = (2p-k+1)\binom{k-2}{k-2} + (2p-k-1)\binom{k-1}{k-2} + (2p-k-3)\binom{k}{k-2} + \ldots + \binom{\frac{2p+k-4}{2}}{k-2} = \binom{\frac{2p+k}{2}}{k} + \binom{\frac{2p+k-2}{2}}{k}.$$

We subtract N from the number of strictly increasing strings $i_1,\ldots,i_k$ of $k/2$ even words, to obtain

$$A_{k,k/2} = \binom{p}{k/2}^2 - \left[\binom{\frac{2p+k}{2}}{k} + \binom{\frac{2p+k-2}{2}}{k}\right]$$

c) If k is odd and $s = \lfloor k/2\rfloor$ or $s = \lceil k/2\rceil$, then by analogy with the preceding treatment we compute first the number of strictly increasing strings $i_1,\dots,i_k$ of natural numbers which are smaller than, or equal to, $2p$, which contain s odd numbers and which do not verify property (*). We denote this number by N' for $s = \lfloor k/2 \rfloor$ and by N'' for $s = \lceil k/2 \rceil$. It may be observed from the expression of N_{d_o} in the previous paragraph that $N_{d_o} = N_{d_o - 1}$ for d_o even and nonzero. For $s = k/2$, $d_o = i_1$ must be even, and for $s = \lceil k/2 \rceil$, $d_o = i_1$ must be odd. Thus,

$$N' = \sum_{\substack{d_o=2 \\ d_o \text{ even}}}^{2p-k+1} N_{d_o} = \sum_{\substack{d_o=1 \\ d_o \text{ odd}}}^{2p-k} N_{d_o} = N'' =$$

$$= \frac{2p-k+1}{2}\binom{k-2}{k-2} + \frac{2p-k-1}{2}\binom{k-1}{k-2} + \dots + \binom{\frac{2p+k-5}{2}}{k-2} =$$

$$= \binom{\frac{2p+k-1}{2}}{k}.$$

The last result, obtained by means of rearranging combinatorial formulas, indicates that for $s = \lfloor k/2 \rfloor$ or $s = \lceil k/2 \rceil$ and odd k,

$$A_{k,s} = \binom{p}{s}\binom{p}{k-s} - \binom{\frac{2p+k-1}{2}}{k}$$

Starting from Newton's binomial formula and from the relationship

$$K_{2p} = (a+1)^p(b+1)^p + 1 - R(a,b;2p)$$

we obtain in a straightforward manner the final result:

$$K_{2p}(a,b) = \sum_{j=0}^{p} a^j b^j \left[\binom{p+j}{2j} + \binom{p+j-1}{2j} + \binom{p+j}{2j+1}(a+b) \right] \qquad (2)$$

2.3.CASE 2: r IS ODD

Let $r = 2p+1$, $p \in \mathbb{N}$, e.g. catafusenes 1 , 2 or 3. We have:

$$K_{2p+1} = P(a,b;2p+1) = (a+1)^{p+1}(b+1)^{p} + 1 - R(a,b;2p+1)$$

The set of edges contains p+1 odd-rank edges of weight a, and p even-rank edges of weight b.

For $s \in \overline{0,p}$ we define $I_s \subseteq \mathbb{N}$ as in the preceding case, and with $A_{k,s} \in \mathbb{N}$ we have:

$$R(a,b;2p+1) = \sum_{\substack{0 \leq s \leq p \\ k \in I_s}} A_{k,s}\, a^s\, b^{k-s}$$

The coefficients $A_{k,s}$ are calculated analogously to the preceding cases by computing the number of strictly increasing strings with indices $i_1, \ldots, i_k$ which verify the property (*). We obtain thus:

a) If $s \neq \lfloor k/2 \rfloor$ and $s \neq \lceil k/2 \rceil$, then $A_{k,s} = \binom{p+1}{s}\binom{p}{k-s}$.

b) If k is even and $s = k/2$, $N_{d_o} = \sum_{j=k-1}^{\lfloor (2p-d_o+k)/2 \rfloor} \binom{j-1}{k-2}$,

$N = \sum_{d_o=1}^{2p-k+2} N_{d_o} = 2\binom{\frac{2p+k}{2}}{k}$, therefore

$$A_{k,k/2} = \binom{p+1}{k/2}\binom{p}{k/2} - 2\binom{\frac{2p+k}{2}}{k}.$$

c) If k is odd and $s = \lceil k/2 \rceil$, we obtain $N = \sum_{\substack{d_o=1 \\ d_o \text{ odd}}}^{2p-k+2} N_{d_o} = \binom{\frac{2p+k+1}{2}}{k}$, and therefore

$$A_{k,s} = \binom{p+1}{s}\binom{p}{k-s} - \binom{p+s}{k}.$$

If k is odd and $s = \lfloor k/2 \rfloor$, we obtain:

$$N = \sum_{\substack{d_o=2 \\ d_o \text{ even}}}^{2p-k+1} N_{d_o} = \binom{\frac{2p+k-1}{2}}{k} \text{, and thus}$$

$$A_{k,s} = \binom{p+1}{s}\binom{p}{k-s} - \binom{\frac{2p+k-1}{2}}{k}$$

Therefore for odd k and $s = \lfloor k/2 \rfloor$ or $s = \lceil k/2 \rceil$,

$$A_{k,s} = \binom{p+1}{s}\binom{p}{k-s} - \binom{p+s}{k}$$

Thus in the case of odd $r = 2p+1$, the final result is

$$K_{2p+1}(a,b) = \sum_{j=0}^{p} a^j b^j \left[2\binom{p+j}{2j} + a\binom{p+j+1}{2j+1} + b\binom{p+j}{2j+1} \right] \quad (3)$$

The two formulas (2) and (3) may be unified, irrespectively of the parity of r, into:

$$K_r(a,b) = \sum_{j=0}^{p} a^j b^j \binom{p+j}{2j} \left[\frac{2p+j+(-1)^{r+1}j}{p+j} + \frac{p-j}{2j+1}(a+b) + \frac{a+(-1)^{r+1}a}{2} \right] \quad (4)$$

where $p = \lfloor r/2 \rfloor$.

It can be demonstrated that from relationships (2) and (3) one may obtain recurrence (1) in the particular case $a=b=j$. The demonstration consists in showing both for $r=2p+1$ and for $r=2p+2$ that the difference below is calculated to be zero:

$$K_{j,2p+1} - K_{j,2p-1} - jK_{j,2p}$$

This relationship is reminiscent of the Fibonacci recurrence, and may qualify as a generalized Fibonacci relationship; it becomes a true Fibonacci recurrence for $j=1$. Other generalized Fibonacci numbers have been published in chemical contexts [10, 21-23].

2.4. RELATIONSHIP WITH THE GORDON-DAVISON ALGORITHM

We shall now examine binary regularly condensed catafusenes with the help of the Gordon-Davison (GD) algorithm [1]. According to this algorithm, the number of Kekulé structures increases on the i-th linear portion from the margin with a constant increment q_i on going from a benzenoid ring to the next one in arithmetical progression. By definition, $q_o=1$ and $K_o=2$.

Lemma 1. The following recurrences hold:

$$K_{2p+1} = K_{2p} + a \sum_{i=1}^{2p} \left[(-1)^i K_i + 1\right] \text{, and} \qquad (5)$$

$$K_{2p} = K_{2p-1} + b \sum_{i=1}^{2p-1} \left[(-1)^{i+1} K_i - 1\right] \qquad (5')$$

for any $1 \leqslant p \leqslant \lfloor r/2 \rfloor$.

For demonstration, we observe from the GD algorithm that

$$q_i = K_{i-1} - q_{i-1} \quad \text{for} \quad 1 \leqslant i \leqslant r,$$

and that

$$K_{2p+1} = K_{2p} + aq_{2p+1}$$

$$K_{2p} = K_{2p-1} + bq_{2p} \quad \text{for } 1 \leqslant p \leqslant \lfloor r/2 \rfloor .$$

From these three equalities, relationships (5) and (5') follow easily.

Definition. We define the alternating sums of a regularly condensed catafusene:

$$S_i = 1 - K_1 + \ldots + (-1)^i K_i \quad \text{for } 1 \leqslant i \leqslant r .$$

By definition, $S_o = 1$. We observe that for $1 \leqslant p \leqslant \lfloor r/2 \rfloor$

$$K_{2p} = S_{2p} - S_{2p-1} \quad \text{and} \quad K_{2p+1} = S_{2p} - S_{2p+1}$$

Lemma 2. The following equality holds for $2 \leqslant i \leqslant r-2$:

$$S_{i+2} - (ab+2)S_i + S_{i-2} = 0$$

The demonstration is based on Lemma 1 and on the above observation, via the results:

$$S_{2p+1} = S_{2p-1} - aS_{2p}$$
$$S_{2p} = S_{2p-2} - bS_{2p-1}$$

Lemma 3. We have

$$K_{i+2} - (ab+2)K_i + K_{i-2} = 0 \text{ for } 2 \leqslant i \leqslant r-2.$$

For i=2 direct verification is possible. The proof starts from the observation written as

$$K_i = (-1)^i(S_i - S_{i-1}) \quad \text{for } 1 \leqslant i \leqslant r$$

and adapted to K_{i+2} and K_{i-2} also. Now the result follows from Lemma 2.

Theorem 2. For any $r \in \mathbb{N}$,

$$K_r(a,b) = \left[h_1 + \frac{1+(-1)^{r+1}}{2} g_1\right] x_1^{\lfloor r/2 \rfloor} + \left[h_2 + \frac{1+(-1)^{r+1}}{2} g_2\right] x_2^{\lfloor r/2 \rfloor} \quad (6)$$

where $h_i = 1 + (-1)^{i+1} \dfrac{a+b}{\sqrt{ab(ab+4)}}$

$$g_i = \frac{a}{2} + (-1)^{i+1} \frac{ab(a+2)}{2\sqrt{ab(ab+4)}} \quad \text{and}$$

$$x_i = \frac{ab+2}{2} + (-1)^{i+1} \frac{\sqrt{ab(ab+4)}}{2} \quad \text{for } i = 1,2.$$

The proof starts with the characteristic equation for the recurrence of Lemma 3:

$$x^2 - (ab+2)x + 1 = 0$$

which has the roots $x_i = \dfrac{(ab+2) \pm \sqrt{ab(ab+4)}}{2}$, hence

$$K_{2p} = h_1 x_1^p + h_2 x_2^p \quad \text{for } 0 \leqslant p \leqslant \lfloor r/2 \rfloor \quad (7)$$

The constants h_1 and h_2 are obtained from the initial conditions: $K_0=2$ and $K_2 = ab+a+b+2$:

$$h_1 = 1 + \frac{a+b}{\sqrt{ab(ab+4)}} \quad \text{and} \quad h_2 = 1 - \frac{a+b}{\sqrt{ab(ab+4)}}$$

We use the notation $D_p = K_{2p+1} - K_{2p}$,and from Lemma 3 we have

$$D_{p+1} - (ab+2)D_p + D_{p-1} = 0 \quad \text{for} \quad 0 \leqslant p \leqslant \lfloor r/2 \rfloor .$$

With the initial conditions $D_o = a$ and $D_1 = a^2b+ab+a$ we obtain from

$$D_p = g_1 x_1^p + g_2 x_2^p$$

the values of constants g_1 and g_2 :

$$g_1 = \frac{a}{2} + \frac{ab(a+2)}{2\sqrt{ab(ab+4)}} \quad \text{and} \quad g_2 = \frac{a}{2} - \frac{ab(a+2)}{2\sqrt{ab(ab+4)}} ,$$

therefore

$$K_{2p+1} = (g_1+h_1)x_1^p + (g_2+h_2)x_2^p \tag{8}$$

and Theorem 2 results from (7) and (8).

<u>Corollary</u>. For any $0 \leqslant p \leqslant \lfloor r/2 \rfloor$ one obtains

$$K_{2p} = \left[1 + \frac{a+b}{\sqrt{ab(ab+4)}}\right] \left[\frac{(ab+2)+\sqrt{ab(ab+4)}}{2}\right]^p + \left[1 - \frac{a+b}{\sqrt{ab(ab+4)}}\right] \left[\frac{(ab+2)-\sqrt{ab(ab+4)}}{2}\right]^p \tag{9}$$

Similarly,

$$K_{2p+1} = \left[\frac{a+2}{2} + \frac{2(a+b)+ab(a+2)}{2\sqrt{ab(ab+4)}}\right] \left[\frac{(ab+2)+\sqrt{ab(ab+4)}}{2}\right]^p + \left[\frac{a+2}{2} - \frac{2(a+b)+ab(a+2)}{2\sqrt{ab(ab+4)}}\right] \left[\frac{(ab+2)-\sqrt{ab(ab+4)}}{2}\right]^p \tag{10}$$

3.GENERALIZATION FOR ANY REGULARLY CONDENSED CATAFUSENE

We consider sequences of s linear segments in a zig-zag catafusene (or any isoarithmic system) which are repeated p times,with a portion of t segments at one end ($0 \leqslant t \leqslant s-1$), where segment i consists of a_i+1 linearly condensed benzenoid rings with the weight $a_i \geqslant 1$ for every $i \in \overline{0,s-1}$,e.g. 6 . The total number of straight stretches (segments) in the catafusene is $r = ps + t$ and the total number of benzenoid rings is

$$p \sum_{i=0}^{s-1} a_i + \sum_{i=0}^{t} a_i \quad .$$

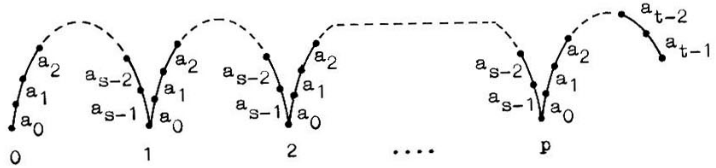

Fig.2.Tree $T(G^*)$ with weights a_i of edges corresponding to the numbers of linearly condensed benzenoid rings.

By definition,operation $\oplus$ is summation modulo s :

$$x \oplus y = x+y \pmod{s} \quad \forall\, x,y \in \mathbb{Z} \,.$$

Lemma 4.For any $0 \leqslant p \leqslant \lfloor r/s \rfloor$ and $0 \leqslant i \leqslant s-1$,provided that $i+p \neq 0$,we have:

$$K_{sp+i} = K_{sp+i-1} + a_{i\oplus(-1)} \left[\sum_{j=1}^{sp+i-1} (-1)^{j+1} K_{sp+i-j} + (-1)^{sp+i+1} \right]$$

The proof starts from the GD algorithm which yields:

$$q_{sp+i} = K_{sp+i-1} - q_{sp+i-1} \qquad \text{and}$$

$$K_{sp+i} = K_{sp+i-1} + a_{i-1} q_{sp+i}$$

for $0 \leqslant p \leqslant \lfloor r/s \rfloor$, $0 \leqslant i \leqslant s-1$, $i+p\neq 0$.

By combining the above two formulas,and expanding stepwise the summation,Lemma 4 is obtained.Note that

$$K_i = (-1)^i (S_i - S_{i-1}) \quad \text{for } 1 \leqslant i \leqslant r \qquad (11)$$

Lemma 5. We have

$$S_{sp+i} = S_{sp+i-2} - a_{i\oplus(-1)} S_{sp+i-1}$$

for any $0 \leqslant p \leqslant \lfloor r/s \rfloor$ and $0 \leqslant i \leqslant s-1$ such that $sp+i-2 \geqslant 0$.

The demonstration results by combining Lemma 4 with the preceding observation.

Lemma 6. For $0 \leqslant p \leqslant \lfloor r/s \rfloor - 2$ and $0 \leqslant i \leqslant s-1$ the following recurrence holds, where $R_s \in \mathbb{Z}[X_o,\ldots,X_{s-1}]$:

$$S_{sp+i} - R_s(a_o,\ldots,a_{s-1})S_{sp+s+i} + (-1)^s S_{sp+2s+i} = 0 ;$$

$$R_s(X_o,\ldots,X_{s-1}) = e_o X_o X_1 \ldots X_{s-1} + e_1 \sum_{i=0}^{s-1} X_i X_{i\oplus 1} \ldots X_{i\oplus(s-2)} + \ldots + e_{s-2} \sum_{i=0}^{s-1} X_i X_{i\oplus 1} + e_{s-1} \sum_{i=0}^{s-1} X_i + 2e_s ,$$

where the coefficients $e_i = \frac{1+(-1)^i}{2}$.

The first terms are:

$$R_1 = a_o$$
$$R_2 = a_o a_1 + 2$$
$$R_3 = a_o a_1 a_2 + a_o + a_1 + a_2$$
$$R_4 = a_o a_1 a_2 a_3 + a_o a_1 + a_1 a_2 + a_2 a_3 + a_3 a_o + 2$$
$$R_5 = a_o a_1 a_2 a_3 a_4 + a_o a_1 a_2 + a_1 a_2 a_3 + a_2 a_3 a_4 + a_3 a_4 a_o + a_o + a_1 + a_2 + a_3 + a_4$$

The proof starts by applying Lemma 5 to S_{sp+i} where $i = 2,3,\ldots,2s$, and combines all these relationships.

Theorem 3. For any $0 \leqslant p \leqslant \lfloor r/s \rfloor$ and $0 \leqslant i \leqslant s-1$, using for brevity the notation R_s in place of $R_s(a_o,\ldots,a_{s-1})$ we have

$$K_{sp+i} = Q_i^{(1)}\left[\frac{R_s + \sqrt{R_s^2-4(-1)^s}}{2}\right]^p + Q_i^{(2)}\left[\frac{R_s - \sqrt{R_s^2-4(-1)^s}}{2}\right]^p \quad (12)$$

where

$$Q_i^{(1)} = \frac{K_i}{2} + \frac{2K_{s+i} - R_sK_i}{2\sqrt{R_s^2-4(-1)^s}} \quad \text{,and}$$

$$Q_i^{(2)} = \frac{K_i}{2} - \frac{2K_{s+i} - R_sK_i}{2\sqrt{R_s^2-4(-1)^s}}$$

The proof starts from (11), which can be converted into:

$K_{sp-s+i} - R_sK_{sp+i} + (-1)^sK_{sp+s+i} = 0$ for $1 \leqslant p \leqslant \lfloor r/s \rfloor -1$ and $0 \quad i \quad s-1$.

The characteristic equation is

$(-1)^s x^2 - R_s x + 1 = 0$,and its roots are:

$$x_{1,2} = \frac{R_s \pm \sqrt{R_s^2 - 4(-1)^s}}{2}$$

Therefore

$$K_{sp+i} = Q_i^{(1)} x_1^p + Q_i^{(2)} x_2^p$$

for $0 \leqslant p \leqslant \lfloor r/s \rfloor$ and $0 \leqslant i \leqslant s-1$.

We compute K_i and K_{s+i} with the aid of initial conditions:

$$Q_i^{(1)} + Q_i^{(2)} = K_i$$

$$Q_i^{(1)} x_1 + Q_i^{(2)} x_2 = K_{s+i}$$

and thus Theorem 3 is proved.

Although formula (12) holds for any integer $p \geqslant 0$, it makes sense to apply such a formula only for $p \geqslant 2$, i.e. when a sequence of s straight segments (in the dualist graph G^* of a catafusene G) is repeated at least twice.

Another observation for the numerical application of this Theorem is that for computing K_i or K_{s+i} one has to sum all distinct monomials $a_{i_1} a_{i_2} \ldots a_{i_k}$ with $0 \leq i_1 < i_2 < \ldots < i_k \leq s-1$ such that at least two of these successive indices i_j and i_{j+1} $(1 \leq j \leq k-1)$ have the same parity. It is recommended to start for the small i values either with such a summation or with the GD algorithm for obtaining the K_i and K_{s+i} values.

The generalized formula (12) may be particularized to the case $s = 2, a_o = a, a_1 = b$, which reduces to $R_s = ab+2$ and converts formula (12) into

$$K_{2p+i} = Q_i^{(1)} \left[\frac{ab+2+\sqrt{ab(ab+4)}}{2} \right]^p + Q_i^{(2)} \left[\frac{ab+2-\sqrt{ab(ab+4)}}{2} \right]^p$$

where $$Q_o^{(1,2)} = 1 \pm \frac{a+b}{\sqrt{ab(ab+4)}}$$

$$Q_1^{(1,2)} = \frac{a+2}{2} \pm \frac{2(a+b)+ab(a+2)}{2\sqrt{ab(ab+4)}}$$

in agreement with formulas (9) and (10).

4. CONCLUSIONS

Both the combinatorial formulas (2)-(4), and the formula (6) based on the GD algorithm yield the same numerical results. This can be verified for $a = 2, b = 1, r = 8$; in agreement with manual implementation of the GD algorithm, $K_8(1,2) = K_8(2,1) = 362$. For an odd $r = 3, a = 4, b = 2, p = 1$, one finds similarly $K_3(4,2) = 50$.

A numerical verification of formula (12) for the case $s = 3, t = 0, p = 2, a_o = 2, a_1 = 3, a_2 = 1$ yields:

$$R_3 = a_o a_1 a_2 + a_o + a_1 + a_2 = 12$$

$$K_{3\cdot 2+1} = K_7 = Q_1^{(1)}\left(\frac{12+2\sqrt{37}}{2}\right)^2 + Q_1^{(2)}\left(\frac{12-2\sqrt{37}}{2}\right)^2$$

$$Q_1^{(1,2)} = \frac{K_1}{2} \pm \frac{2K_4 - 12K_1}{4\sqrt{37}} = 2 \pm \frac{25}{2\sqrt{37}}$$

therefore $K_7 = 292 + 300 = 592.$

REFERENCES

1. M.Gordon and W.H.T.Davison,J.Chem. Phys.20,428(1952).
2. O.E.Polansky and I.Gutman,Math.Chem.8,269(1980).
3. I.Gutman and O.E.Polansky,"Mathematical Concepts in Organic Chemistry",Springer,Berlin,1986,p.62.
4. I.Gutman,Croat.Chem.Acta 56,365(1983);Bull.Soc.Chim. Beograd 47,453(1982).
5. I.Gutman and N.Trinajstić,Croat.Chem.Acta 45,423(1973).
6. N.Trinajstić,"Chemical Graph Theory",vol.2,p.29,CRC Press, Boca Raton,Florida,1983.
7. M.Randić,J.Chem.Soc.Faraday Trans.2,232(1976).
8. T.F.Yen,Theoret.Chim.Acta,20,399(1971).
9. D.Cvetković and I.Gutman,Croat.Chem.Acta 46,15(1974).
10. S.J.Cyvin,Math.Chem.13,167(1982);Acta Chim.Hung.112,281 (1983);J.Mol.Struct.100,75(1983).
11. S.J.Cyvin and I.Gutman,Comp.and Maths.with Appls.,12B, 877(1986),reprinted in "Symmetry Unifying Human Understanding"(ed.I.Hargittai),Pergamon Press,New York,1986, as vol.10 of Internat.Series in Modern Applied Mathematics and Computer Science.
12. B.Dzonova-Jerman-Blazić and N.Trinajstić,Comput.Chem. 6,121 (1982).
13. J.V.Knop,W.R.Müller,K.Szymanski and N.Trinajstić, "Computer Generation of Certain Classes of Molecules", Kemia u industrii,Zagreb,1985,pp.109,112.
14. R.L.Brown,J.Comput.Chem.4,556(1983).
15. A.T.Balaban and I.Tomescu,Math.Chem. 14,155(1983).

16. A.T.Balaban and I.Tomescu,Math.Chem. 17,91(1985).

17. K.Balasubramanian,J.J.Kaufman,W.S.Koski and A.T.Balaban, J.Comput.Chem. 1,149(1980).

18. J.V.Knop,K.Szymanski,Z.Jericević and N.Trinajstić,Int.J. Quantum Chem. 23,713(1983).

19. A.T.Balaban and F.Harary,Tetrahedron 24,2505(1968).

20. A.T.Balaban,Pure Appl.Chem. 54,1075(1982).

21. A.T.Balaban and I.Tomescu,Croat.Chem.Acta 57,391(1984).

22. H.Hosoya,The Fibonacci Quart. 21,255(1973).

23. J.L.Bergan,B.N.Cyvin and S.J.Cyvin,Acta Chim.Hung.(to appear).