match no. 14 pp. 155-182 1983

ALGEBRAIC EXPRESSIONS FOR THE NUMBER OF KEKULÉ STRUCTURES OF ISOARITHMIC CATA-CONDENSED BENZENOID POLYCYCLIC HYDROCARBONS

Alexandru T.Balaban [a] and Ioan Tomescu [b]

[a] Department of Organic Chemistry, The Polytechnic Institute, Splaiul Independenţei 313, Bucharest, Romania.

[b] Faculty of Mathematics, The University of Bucharest, Str. Academiei 14, Bucharest, Romania and National Institute for Scientific and Technical Creation (INCREST) Department of Mathematics, Bd.Păcii 220, 79622 Bucharest,Romania.

(Received: April 1983)

Abstract. For branched and non-branched cata-condensed benzenoid polycyclic hydrocarbons (catafusenes) algebraic expressions for the Kekulé structure count K are obtained, depending on the numbers of condensed hexagons in the linear segments composing the catafusene. Irrespective of the direction of kinks, catafusenes with the same sequences of segments (which we call "isoarithmic catafusenes") have the same structure count. The procedure takes into account the parity of the paths between segments: odd paths in the associated tree lead to incompatible pairs and on this basis formula (3) is obtained and demonstrated for K.

A numerical triangle, derived by following a zig-zag path in Pascal's triangle, gives the number of terms in the algebraic expression of K for non-branched catafusenes.

NOTATION

G = catafusene graph, with vertex set $\{v_i\}$ $(i = 1,\dots,4n+2)$ and edge set $\{e_1,\dots,e_p,E_1,\dots,E_{n-1}\}$

G^* = dualist graph of G, with n vertices corresponding to the n hexagons

A_j = number of hexagons in linear segment j $(j = 1,\dots,r)$ of G

$T(G^*)$ = "isoarithmicity tree" - because isomeric catafusenes with isomorphic associated trees T, such that corresponding linear segments of the catafusenes contain the same number of hexagons are isoarithmic, i.e. they have the same number of Kekulé structures. This tree has $r + 1$ vertices labelled x_i and r edges labelled a_j

K or K(G)=Kekulé structure count of catafusene

K_2 and $K_{i,j}$ = complete graph of order 2 and complete bipartite graph, respectively

$P(u,v)$ = path between edges u and v in $T(G^*)$

$P(T;a_1,\dots,a_r)$ = polynomial (1) associated to tree T

$R_1(T;a_1,\dots,a_r)$ = polynomial in terms of a_i's possibly containing redundancies

$R(T;a_1,\dots,a_r)$ = the above polynomial from which redundancies were eliminated

φ = isomorphism between trees T

$I(T)$ = set of pairs of incompatible edges in T

$W_{r,k}$ = number of terms in R(G) containing exactly k variables a_i for tree $T(G^*)$ with r edges

$$W_r = \sum_{k \geq 2} W_{r,k}$$

F_i = i^{th} Fibonacci number

$f(r,k)$ = binomial coefficient appearing in (6)

L_i = linear segment of G

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

$|E(T)|$ = cardinality of the above set

$\mathcal{A}$ = set of labels associated with the edges of T

INTRODUCTION

The enumeration of Kekulé structures for benzenoid polycyclic hydrocarbons is important because the stability and many other properties of these hydrocarbons have been found to correlate with the number of Kekulé structures. Starting with the algorithm proposed by Gordon and Davison[1], many papers have appeared on the problem of finding the "Kekulé structure count" K for such hydrocarbons. A whole chapter in a recent book on chemical graph theory is devoted to this topic[2].

We can mention here only a few authors who contributed to this topic: Hosoya's group who introduced the sextet polynomial[3,4] (see also[5]), Herndon[6], Yen[7], Cvetković[8,10], Gutman[8-12], Trinajstić[10], Polansky[12], Randić[13], Schmidt[14], El-Basil[15], Cyvin[16].

DEFINITIONS

Polycyclic benzenoid systems (polyhexes) are classified into cata-condensed (catafusenes) and peri-condensed (perifusenes) according to the acyclic or cyclic nature of their characteristic (or dualist as we now prefer) graphs[17-19].

The present paper will discuss only catafusenes.

All catafusenes with the same number n of hexagons are isomeric i.e. they all have the same molecular formula $C_{4n+2}H_{2n+4}$. The molecular graph of a cata-condensed benzenoid hydrocarbon (abbreviated by Gutman as CCB graph[11]), will be named here catafusene graph. Every such graph G with n hexagons has precisely $p = 4n+2$ vertices and $q = 5n+1$ edges. The vertices of G will be labelled by $v_1, v_2, \ldots, v_p$ and the edges by $e_1, \ldots, e_p, E_1, \ldots, E_{n-1}$ such that $e_i = v_i v_{i+1}$ for $1 \leq i \leq p-1$, $e_p = v_p v_1$ and the p-cycle of G: $v_1, v_2, v_3, \ldots, v_p, v_1$ is the perimeter of G.

The edges $e_1, e_2, \ldots, e_p$ of G will be called external and the remaining n-1 edges $E_1, \ldots, E_{n-1}$ are said to be internal; the latter represent the bonds between neighbouring hexagons of the structure, and are intersected by edges of the dualist graph.

Dualist graphs of catafusenes are trees whose vertices represent centres of hexagons and whose edges link together vertices corresponding to condensed hexagons, i.e. vertices sharing two adjacent carbon atoms in the original hydrocarbon. In dualist graphs of catafusenes angles are important: condensation can occur only linearly (at 180°, coded by digit 0) or at angles of 120° and 240° (angular or kinked condensation, coded by digits 1 and 2, respectively).[17,18]

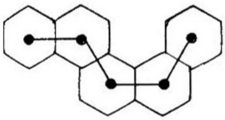
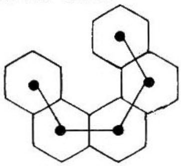
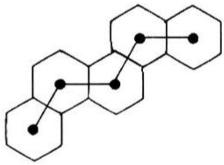
A vertex in a dualist graph can have degree one (endpoint or terminal vertex), two, or at most three; in the latter case this is a branching point.

We introduce the term "isoarithmic* " catafusenes for non-isomorphic systems having the same K value as a consequence of differing in the topology of annelation only by the direction of kinks (i.e. only by interchanging some angles of 120°

* Meaning "with the same number", or "with the same count" (of Kekulé structures)

with 240°) but leaving unaffected all linear annelations.

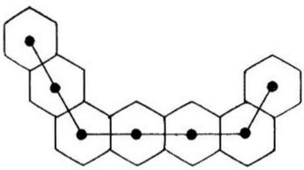
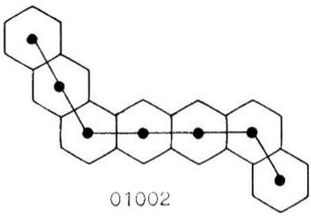
Such systems also have the same sextet polynomial and the same L-transform of their three-digit code[20]. This is equivalent with replacing both digits 1 and 2 in the three-digit code by letter l, and leaving zeroes in the code as they were. Three isoarithmic examples follow with their three-digit codes; their L-transform is lll.

Code : 112 111 121

I II III

For the following two isoarithmic catafusenes the L-transform is 0l00l.

Code : 01001 01002

IV V

Thus, isomeric catafusenes with the same number n of hexagons are further partitioned into groups of isoarithmic catafusenes.

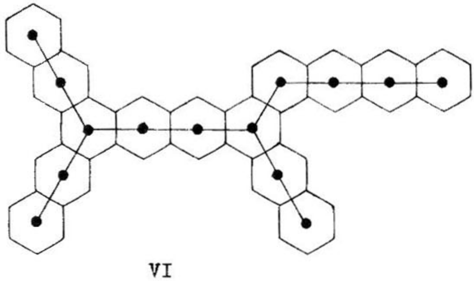
Let $A_i' = A_i + 1$ denote the numbers of linearly condensed hexagons in each linear portion (segment) of the catafusene:

kinks and branching points are counted twice and three times, respectively, in $A_i + 1$ values for each of the branches having a common vertex of the dualist graph. In the triad I-III of isoarithmic catafusenes $A_1 + 1 = A_2 + 1 = A_3 + 1 = A_4 + 1 = 2$, and for the subsequent pairs IV and V, $A_1 + 1 = 3, A_2 + 1 = 4, A_3 + 1 = 2$.

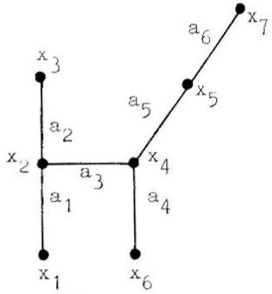
ALGEBRAIC EXPRESSIONS FOR THE KEKULÉ STRUCTURE COUNT OF CATAFUSENES

To every catafusene graph G containing n hexagons we have associated its dualist graph G^* which is a tree having n vertices and n - 1 edges. Now we shall associate to G^* another tree, denoted $T(G^*)$, which represents each linear segment of G^* by a unique edge. Hence $T(G^*)$ has r + 1 vertices and r edges, if G^* contains exactly r linear segments.

Whereas dualist graphs G^* are trees with geometric constraints (bond angles are important), the derived trees $T(G^*)$ are normal labelled trees, with the only restriction that their vertex degrees are at most three. For example, if G and G^* are shown in VI, then $T(G^*)$ is the tree VII.



VI



VII

$A_1 = A_2 = A_4 = 2$; $A_3 = A_6 = 3$; $A_5 = 1$.

We shall denote the vertex set of a tree T of order r + 1 by $\{x_1, x_2, \dots, x_{r+1}\}$ and its edge set by E(T); hence $|E(T)| = r$. Define an injective function $f: E(T) \rightarrow \mathcal{A} = \{a_1, a_2, \dots, a_r\}$ which associates every edge u of T with a variable $f(u) \in \mathcal{A}$, such that different edges of T receive distinct literals from set $\mathcal{A}$, as shown in VII. The function f is a labelling of the edges of T with labels from $\mathcal{A}$.

Any two edges u and v of T having the property that the unique path P(u,v) of T between u and v has an odd length (number of edges) will be called incompatible; otherwise u and v are said to be compatible.

For the tree VII edges x_1x_2 and x_4x_5 are incompatible since the path $P(x_1x_2, x_4x_5) = x_2, x_4$ is odd; also x_2x_3 and x_4x_6, x_2x_3 and x_4x_5, or x_4x_6 and x_5x_7 are examples of incompatible pairs of edges.

Note that the incompatibility relation on the edge set E(T) of T is a binary relation which is symmetric but not transitive. For example x_1x_2 is incompatible with x_4x_6 ; x_4x_6 is incompatible with x_5x_7, but x_1x_2 and x_5x_7 is a compatible pair.

For a tree T denote by I(T) the set of all pairs of incompatible edges of T. An edge w is said to lie on the path P(u,v) if both extremities of w belong to P(u,v). If w does not lie on P(u,v) we shall write $w \notin P(u,v)$.

For any tree T of order r + 1 we shall define the following polynomial:

$$R_1(T; a_1, \dots, a_r) = \sum_{(u,v) \in I(T)} f(u)f(v) \prod_{\substack{w \notin P(u,v) \\ w \neq u,v}} (f(w) + 1),$$

where $f: E(T) \longrightarrow \{a_1,\dots,a_r\}$.

Now develop $R_1(T;a_1,\dots,a_r)$ as a sum of products of variables $a_1,a_2,\dots,a_r$ and apply the following idempotency rule for the addition :

$$p + p = p$$

for any $p = a_{i_1}a_{i_2} \dots a_{i_s}$, i.e. if two or more identical products appear, only one of them is taken into account.

Denote by $R(T;a_1,\dots,a_r)$ the sum of products derived from $R_1(T;a_1,\dots,a_r)$ after eliminating all redundant products with the idempotency rule and define the polynomial associated with T as follows:

$$P(T;a_1,\dots,a_r) = \prod_{u\in E(T)} (f(u) + 1) + 1 - R(T;a_1,\dots,a_r) \quad (1)$$

In order to reduce the amount of computations when we compute $R(T;a_1,\dots,a_r)$ we shall use the following two rules:

1st rule: Let u and v be two incompatible edges of T and $P(u,v) = y_1,y_2,\dots,y_i$. If the degree of y_1, denoted $\deg(y_1)$ is equal to 2 and $\deg(y_i) = 3$, let w be the edge incident to y_i, such that $w \notin P(u,v)$. It follows that $\{u,w\}$ is also an incompatible pair and the contribution of the pairs $\{u,v\}$ and $\{u,w\}$ to $R_1(T;a_1,\dots,a_r)$ equals

$$f(u)\left[f(v) + f(w) + f(v)f(w)\right] \prod_{\substack{z\notin P(u,v)\\ z \neq u,v,w}} (f(z) + 1)$$

2nd rule: Let u and v be two incompatible edges of T and $P(u,v) = y_1,y_2\dots y_i$. If $\deg(y_1) = \deg(y_i) = 3$, let t be the edge incident to y_1 and w be the edge incident to y_i, such that $t,w \notin P(u,v)$. It follows that $\{u,w\}$, $\{t,v\}$ and $\{t,w\}$ are also incompatible pairs of T and the contribution of these four pairs of incompatible edges to $R_1(T;a_1,\dots,a_r)$ equals

$$\left[f(u)+f(t)+f(u)f(t)\right]\left[f(v)+f(w)+f(v)f(w)\right]\prod_{\substack{z\in P(u,v)\\ z\neq u,v,t,w}}(f(z)+1)$$

For example, for the tree VII we find the following pairs of incompatible edges: a_1a_5, a_1a_4, a_2a_5, a_2a_4, a_3a_6, a_4a_6, i.e. $\{x_1x_2,x_4x_5\}$, $\{x_1x_2,x_4x_6\}$, $\{x_2x_3,x_4x_5\}$, $\{x_2x_3,x_4x_6\}$, $\{x_2x_4,x_5x_7\}$, $\{x_4x_6,x_5x_7\}$.

In order to obtain all pairs of incompatible edges of T without repetitions we may first consider an edge u containing a vertex of degree one of T and complete the list of edges incompatible with u. After this we shall delete u from T and we shall perform the same procedure.

For the tree VII by applying the second rule for the first four pairs of incompatible edges we obtain

$$(a_1+a_2+a_1a_2)(a_4+a_5+a_4a_5)(a_6+1)$$

Similarly, the last two pairs generate by the first rule the product

$$(a_1+1)(a_2+1)(a_3+a_4+a_3a_4)a_6$$

hence $R_1(T;a_1,\dots,a_6)$ is equal to the sum of these two products, or

$$R_1(T;a_1,\dots,a_6) = (a_1+a_2+a_1a_2)(a_4+a_5+a_4a_5)(a_6+1) + \\ +(a_1+1)(a_2+1)(a_3+a_4+a_3a_4)a_6 \ .$$

The development of R_1 as a sum of products has thirty products, but three of them are redundant, namely $a_1a_4a_6$, $a_2a_4a_6$ and $a_1a_2a_4a_6$. In conclusion we obtain:

$$P(T;a_1,\dots,a_6) = (a_1+1)(a_2+1)(a_3+1)(a_4+1)(a_5+1)(a_6+1) + 1 -$$

$$- (a_1a_2a_3a_4a_6+a_1a_2a_4a_5a_6+a_1a_2a_3a_6+a_1a_2a_4a_5+a_1a_2a_4a_6+a_1a_2a_5a_6 +a_1a_3a_4a_6+a_1a_4a_5a_6+a_2a_3a_4a_6+a_2a_4a_5a_6+a_1a_2a_4+a_1a_2a_5+a_1a_3a_6 + +a_1a_4a_5+a_1a_4a_6+a_1a_5a_6+a_2a_3a_6+a_2a_4a_5+a_2a_4a_6+a_2a_5a_6+a_3a_4a_6 + +a_1a_4+a_1a_5+a_2a_4+a_2a_5+a_3a_6+a_4a_6) \quad (2)$$

It is not difficult to prove in the general case the simplification rules given above.Indeed,for the first rule the contribution of the incompatible pairs $\{u,v\}$ and $\{u,w\}$ to R_1 is equal to

$$f(u)f(v)(f(w)+1)S_1 + f(u)f(w)(f(v)+1)S_1 = f(u)\left[f(v)+f(w)+ +f(v)f(w)\right]S_1,$$ where $S_1 = \prod_{\substack{z\notin P(u,v)\\ z\neq u,v,w}} (f(z)+1)$,since

$$f(v)f(w) + f(v)f(w) = f(v)f(w)$$

by the idempotency rule.

Similarly,for the second rule the contribution of the incompatible pairs $\{u,v\}$, $\{u,w\}$, $\{t,v\}$ and $\{t,w\}$ to R_1 is equal to

$$f(u)f(v)(f(w)+1)(f(t)+1)S_2 + f(u)f(w)(f(v)+1)(f(t)+1)S_2 + + f(t)f(v)(f(u)+1)(f(w)+1)S_2 + f(t)f(w)(f(u)+1)(f(v)+1)S_2 = = \left[f(u)+f(t)+f(u)f(t)\right]\left[f(v)+f(w)+f(v)f(w)\right]S_2,$$ where

$S_2 = \prod_{\substack{z\notin P(u,v)\\ z\neq u,v,t,w}} (f(z)+1)$,taking into account the idempotency law for addition.

<u>Theorem</u>.For every catafusene graph G the number K(G) of Kekulé structures of G is equal to

$$K(G) = P(T(G^*);A_1,A_2,\ldots,A_r) \quad (3)$$

where the expression of P is given by (1),i.e. to the numerical value of the polynomial associated with $T(G^*)$ for $a_1 = A_1$, $a_2 = A_2$, ... ,$a_r = A_r$,if every linear segment of G with label a_i contains $A_i + 1$ hexagons for $1 \leq i \leq r$.

Proof. Every Kekulé structure of a catafusene molecule is in a one-to-one correspondence with a selection of $\frac{p}{2}$ independent, i.e. mutually non-adjacent,edges in the corresponding molecular graph.Any subset of $\frac{p}{2}$ independent edges in a graph with p vertices is called a perfect matching of this graph.

Hence every Kekulé structure of a catafusene molecule corresponds to a selection of 2n+1 independent edges in its associated catafusene graph.

We shall enumerate the perfect matchings of a catafusene graph relatively to the possibilities of choice of internal edges in the matchings.

As shown by Gutman[14],every catafusene graph has exactly two perfect matchings containing external edges only,namely $\{e_1,e_3,\ldots,e_{p-1}\}$ and $\{e_2,e_4,\ldots,e_p\}$.Let $E_s = v_iv_j$ be an internal edge of G.Then both paths P_1 and P_2 connecting v_i and v_j on the perimeter of G (composed from external edges only) are odd.Indeed,consider a hexagon corresponding to a terminal vertex of the graph G^*.If we delete this hexagon from G,exactly one of the paths P_1 and P_2 decreases its length by 5-1=4,hence it conserves its parity.We may repeat this procedure until we find a graph composed from 2 hexagons only,and E_s is the unique internal edge of this catafusene graph,hence the two paths between v_i and v_j have both a length equal to 5.

Since both paths P_1 and P_2 have odd length, it follows that both have a unique perfect matching M_1, respectively M_2.

Therefore there exists a unique perfect matching of G containing edge E_s and 2n external edges, namely $M_1 \cup M_2 \cup \{E_s\}$.

Now if we consider two internal edges E_s and E_t of G, there exist exactly two paths P_1 and P_2 on the perimeter of G connecting the extremities of E_s to those of E_t and having in common with E_s and E_t only their extremities. We shall prove that P_1 and P_2 are both odd or both even paths of G.

Indeed, E_s and E_t correspond to two edges of the dualist graph G^*, hence to two edges u_i and u_j of the tree $T(G^*)$.

We have $i = j$ if and only if E_s and E_t belong to the same linear segment of G. If one of the paths P_1 or P_2 contains 5 edges of a hexagon corresponding to a terminal vertex of the dualist graph G^*, we have seen that this hexagon can be deleted without changing the parity of P_1 or P_2. Then we may suppose that P_1 and P_2 do not contain such edges, hence they correspond to the unique path between u_i and u_j in $T(G^*)$.

Let G_{st} be the catafusene graph obtained from G by deleting all terminal hexagons on the paths P_1 and P_2. It is clear that P_1 and P_2 contain the same even number of edges on each linear segment of G_{st} and when these paths turn to left or to right, the number of edges of P_1 and P_2 increases by 1 or by 3, i.e. they change the parity. Hence P_1 and P_2 have the same parity in G, opposed to the parity of the unique path $P_{u_i u_j}$ between edges u_i and u_j in $T(G^*)$.

Hence the perimeter of G is decomposed by E_s and E_t into four paths: P_1 and P_2,having the same parity,P_3 between the extremities of E_s,and P_4 between the extremities of E_t.We have proved in the case of a single internal edge that P_3 and P_4 are both odd.

In conclusion,there exists a unique perfect matching of G containing both E_s and E_t if and only if P_1 and P_2 are both odd,or E_s and E_t belong to linear segments of G which correspond to compatible pairs of edges in $T(G^*)$.

We can use a similar argument for a set of $k \leq n-1$ internal edges $E_1,E_2,\ldots,E_k$ of G.By deleting the vertices of these edges we obtain a subgraph H of G and the perimeter of G decomposes into a collection of paths.If the length of every path in this collection is odd,there exists a unique perfect matching in H which together with $E_1,\ldots,E_k$ yields a unique perfect matching in G containing $E_1,\ldots,E_k$.

If at least one path from the above collection has even length,then a perfect matching of G containing the edges $E_1,\ldots,E_k$ can not exist.

Any choice of $k \geq 2$ internal edges which decomposes the perimeter of G into paths such that at least one path from these has even length will be called a bad choice.

It is clear that we must restrict ourselves to choices of internal edges having the property that no two edges belong to the same linear segment of G,since otherwise we shall find bad choices.

The number of all choices of $k \geq 1$ internal edges such that every linear segment of G contribute by at most one edge to every choice is equal to

$A_1+A_2+ \dots +A_r+A_1A_2+ \dots +A_{r-1}A_r+ \dots +A_1A_2\dots A_r =$
$= (A_1+1)(A_2+1) \dots (A_r+1) - 1$.

Since there exist two perfect matchings of G containing no external edge (case k=0),it follows that the number of perfect matchings of G is at most equal to $(A_1+1)(A_2+1)\dots(A_r+1) + 1$,i.e. the value of the polynomial $\prod_{u \in E(T(G^*))}(f(u)+1)+1$ for $a_1=A_1,\dots,a_r=A_r$.

Because by definition $P(T;a_1,\dots,a_r)$ is given by (1),it remains to prove only that the number of bad choices of internal edges,such that at most one edge is chosen on every linear segment of G,is equal to $R(T(G^*);A_1,\dots,A_r)$.

It is clear that the decomposition of the perimeter of G contains at least one path of even length if and only if the choice of the internal edges of G has at least two internal edges E_s and E_t that belong to linear segments L_1 and L_2 of G,corresponding to a pair of incompatible edges u,v in $T(G^*)$ and this choice does not contain any edge lying on a linear segment of G,which corresponds to an edge of the path $P_{u,v}$ between u and v in $T(G^*)$.It follows that the labels of the edges of $T(G^*)$,corresponding to the linear segments of G,on which we can select internal edges such that the resulting choice is a bad choice (containing one edge on L_1,one edge on L_2 and no edge on the segments between L_1 and L_2) are given by the development of the polynomial

$$f(u)f(v) \prod_{\substack{w \notin P(u,v) \\ w \neq u,v}} (f(w)+1)$$

as a sum of products of labels. Denoting by

$$\{f(u)f(v) \prod_{\substack{w \notin P(u,v) \\ w \neq u,v}} (f(w)+1)\}$$

the set of elementary products of labels of the edges of $T(G^*)$ obtained in this way, it follows that

$$\bigcup_{(u,v)\in I(T(G^*))} \{f(u)f(v) \prod_{\substack{w \notin P(u,v) \\ w \neq u,v}} (f(w)+1)\} \qquad (4)$$

will represent all sets of labels associated to the edges of $T(G^*)$, corresponding to all sets of linear segments of G which generate bad choices of internal edges.

Since $R(T(G^*);a_1,\ldots,a_r)$ is obtained from $R_1(T(G^*);a_1,\ldots,a_r)$ after the elimination of all redundant products, it results that the union (4) is precisely $\{R(T(G^*);a_1,\ldots,a_r)\}$.

But for every bad choice $\{a_{i_1},a_{i_2},\ldots,a_{i_k}\}$ of linear segments of G, the number of choices of internal edges on these segments is equal to the product $A_{i_1}A_{i_2}\ldots A_{i_k}$, since each segment with the label a_{i_s} has A_{i_s} internal edges for any $1 \leq s \leq k$.

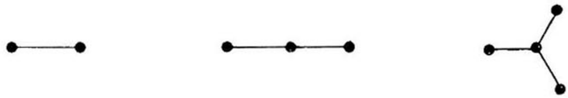
Hence the number of bad choices of internal edges with the specified property equals $R(T(G^*);A_1,\ldots,A_r)$ and the theorem is proved. ∎

Corollary. The number of Kekulé structures of any catafusene graph G satisfies the inequality

$$K(G) \leq (A_1+1)(A_2+1) \ldots (A_r+1) + 1 \qquad (5)$$

This inequality is an equality if and only if: $r = 1$ and $T(G^*) = K_2$; $r = 2$ and $T(G^*) = K_{1,2}$ or $r = 3$ and $T(G^*) =$

= $K_{1,3}$,where graphs K_2, $K_{1,2}$ and $K_{1,3}$ are shown below.



K_2 $K_{1,2}$ $K_{1,3}$

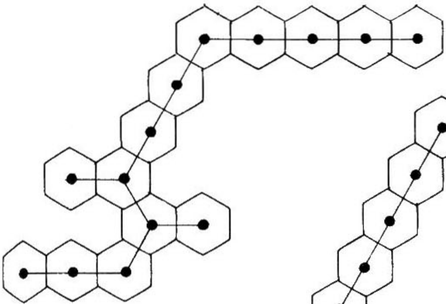
(Note that in this case K_s denotes complete graph of order s and not a number of Kekulé structures).

Proof. It is clear that $T(G^*)$ does not contain any pair of incompatible edges if and only if $T(G^*)$ is one of the trees K_2, $K_{1,2}$ or $K_{1,3}$. ∎

If two catafusene graphs G_1 and G_2 have isomorphic trees $T(G_1^*)$ and $T(G_2^*)$ with r edges,by this method the problems of finding the number of Kekulé structures $K(G_1)$ and $K(G_2)$ are similar,since by a permutation of the variables the polynomial $P(T(G_1^*);a_1,\ldots,a_r)$ is equal to the polynomial $P(T(G_2^*);a_1,\ldots,a_r)$.

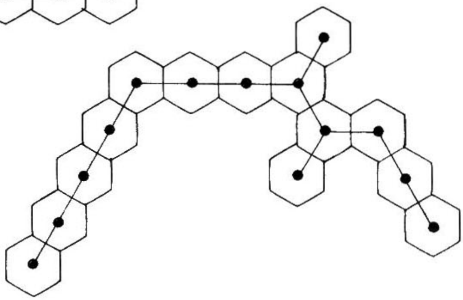
Since the polynomial $P(T(G^*);a_1,\ldots,a_r)$ depends only on the tree $T(G^*)$,it follows that two catafusene graphs G_1 and G_2 lead to isoarithmic catafusenes if there exists an isomorphism φ between their associated trees $T(G_1^*)$ and $T(G_2^*)$ which maps every edge u of $T(G_1^*)$ into an edge $\varphi(u)$ of $T(G_2^*)$,such that linear segments of G_1 and G_2 corresponding to u,respectively $\varphi(u)$,have the same number of hexagons.

By using this property it follows immediately that the catafusene graphs VIII and IX are isoarithmic,which is not an obvious fact by other means.



G_1 and G_1^*

VIII



G_2 and G_2^*

IX

Also, we can easily compute the associated polynomial of a subgraph of G by letting some variables a_i vanish in the polynomial $P(T(G^*);a_1,\ldots,a_r)$, see example 5 below.

However, the maximum number of elementary products of variables in the development of $R_1(T(G^*);a_1,\ldots,a_r)$, obtained by developing each product of parentheses increases exponentially with r, being equal to 2^{r-3} , $3\cdot 2^{r-4}$ (when we apply the first rule of simplification for three pairs of incompatible edges such that P_{uv} has length one), or $9\cdot 2^{r-5}$ (when we apply the second rule and P_{uv} has length one also).

EXAMPLES

1. For an acene consisting of a single rectilinear segment with m vertices ($A_1' = m$ hexagons in the catafusene), formula (3) affords $K = m + 1$, a well-known result (cf. tree $T = K_2$).

2. For a phene consisting of two segments with $A_1' = A_1 + 1$

and $A_2' = A_2 + 1$ hexagons each, one obtains (cf. tree $T = K_{1,2}$)

$$K = (A_1+1)(A_2+1) + 1$$

3. For a linear catafusene with three segments, e.g. IV and V one obtains

$$K = (A_1+1)(A_2+1)(A_3+1) + 1 - A_1A_3$$

E.g. for the pairs IV, V:

$$K_{IV} = K_V = 3\cdot 4\cdot 2 + 1 - (3-1)(2-1) = 23.$$

4. For a branched catafusene with three segments consisting of the branching hexagon plus A_1, A_2 and A_3 hexagons, we obtain

$$K = (A_1+1)(A_2+1)(A_3+1) + 1 \text{ (see tree } K_{1,3}).$$

This result was described by Biermann and Schmidt[14].

5. For a non-branched catafusene with four segments, formula (2) yields for $a_1 = A_1, a_2 = 0, a_3 = A_2, a_4 = 0, a_5 = A_3, a_6 = A_4$:

$$K = (A_1+1)(A_2+1)(A_3+1)(A_4+1)+1-A_1A_2A_4-A_1A_3A_4-A_1A_3-A_2A_4$$

E.g. for the three systems I-III:

$$K_I = K_{II} = K_{III} = 1 + 2^4 - 1 - 1 - 1 - 1 = 13.$$

It can be easily demonstrated that for isoarithmic helicenes or zig-zag catafusenes with all $A_i = 1$ the numbers of Kekulé structures form the Fibonacci series, as mentioned by Cyvin[16], and earlier both by Gordon and Davison[1], and by Yen[7].

The same expressions are obtained by means of the Gordon-Davison algorithm[1] for all cases above, but for more complicated systems that algorithm leads to very complicated algebraic expressions. Nevertheless, the numerical application of the Gordon-Davison algorithm[1] is fairly simple, and can be implemented by means of a computer program.

Finally, K for VI is calculated to be 456 from (2) and (3).

KEKULÉ STRUCTURE COUNT IN NON-BRANCHED CATAFUSENES, AND A NUMERICAL TRIANGLE OBTAINED FROM PASCAL'S TRIANGLE

By means of the Gordon-Davison algorithm[1], or using combinatorial formulas[7], one may compute rapidly the Kekulé structure count. However, the algebraic formulas have the advantage of revealing connexions with other branches of mathematics.

This is illustrated by the Kekulé structure count for non-branched catafusenes, where the algebraic expression leads to a numerical triangle which may be obtained from Pascal's triangle.

The expression resulting by applying the preceding Theorem to non-branched catafusenes, i.e. for examples 2, 3 and 5 and beyond, has the form:

$$K = \prod_{i=1}^{r}(a_i+1) + 1 - R(G),$$

where $R(G)$ is a polynomial expression involving products of a_i's, associated with the catafusene graph G. By grouping these products according to the number k of a_i's, one obtains Table 1. It can be observed that on going from r to r+1, all products with $k \leq r-1$ appear again, and that new products (all containing a_{r+1}) appear.

It was shown earlier that $R(G) = 0$ only for the three graphs K_2, $K_{1,2}$ and $K_{1,3}$.

Thus, the problem of the Kekulé structure count for non-branched catafusenes has a purely combinatorial character. To see this, denote by $W_{r,k}$ the number of products containing exactly k variables a_i in the development of $R(G)$, when the

TABLE 1. The terms of R(G), above the thick line, and the numbers of such terms for given r and k (below the thick line) *

r \ k	5	4	3	2
3	–	–	–	a_1a_3
4	–	–	$a_1a_2a_4$ $a_1a_3a_4$	a_1a_3 a_2a_4
5	–	$a_1a_2a_3a_5$ $a_1a_2a_4a_5$ $a_1a_3a_4a_5$	$a_1a_2a_4$ $a_1a_3a_4$ $a_1a_3a_5$ $a_2a_3a_5$ $a_2a_4a_5$	a_1a_3 a_2a_4 a_1a_5 a_3a_5
6	$a_1a_2a_3a_4a_6$ $a_1a_2a_3a_5a_6$ $a_1a_2a_4a_5a_6$ $a_1a_3a_4a_5a_6$	$a_1a_2a_3a_5$ $a_1a_2a_4a_5$ $a_1a_3a_4a_5$ $a_1a_2a_4a_6$ $a_1a_3a_4a_6$ $a_1a_3a_5a_6$ $a_2a_3a_4a_6$ $a_2a_3a_5a_6$ $a_2a_4a_5a_6$	$a_1a_2a_4$ $a_1a_3a_4$ $a_1a_3a_5$ $a_2a_3a_5$ $a_2a_4a_5$ $a_1a_2a_6$ $a_1a_3a_6$ $a_1a_4a_6$ $a_1a_5a_6$ $a_2a_4a_6$ $a_3a_4a_6$ $a_3a_5a_6$	a_1a_3 a_2a_4 a_1a_5 a_3a_5 a_2a_6 a_4a_6
3	–	–	–	1
4	–	–	2	2
5	–	3	5	4
6	4	9	12	6

* The numbers in the lower part of Table 1 (below the thick line) are denoted by $W_{r,k}$ and their sum for all k values is denoted by W_r.

associated tree $T(G^*)$ has r edges and by $W_r = \sum_{k \geq 2} W_{r,k}$.

Since every product of R(G) corresponds to a bad choice of internal edges in the catafusene graph G, it follows that there exists an one-to-one mapping between the set of all products $a_{i_1} a_{i_2} \ldots a_{i_k}$ in the development of R(G) and the set of all sequences of k natural numbers

$$1 \leq i_1 < i_2 < \ldots < i_k \leq r$$

such that at least one of the differences $i_{s+1} - i_s$ $(1 \leq s \leq k-1)$ is an even number.

This property of indices of a's may be verified directly from Table 1.

The lower part of this Table 1 is a numerical triangle of the numbers $W_{r,k}$. We propose to show that it can be easily constructed from Pascal's triangle of binomial coefficients presented in Table 2 in a slightly modified form: its last column of 1's has been deleted, and zig-zag lines have been marked. All entries in Pascal's triangle are bracketed in order to distinguish them from the entries of $W_{r,k}$ in Tables 1 and 3.

Table 3 shows how one can obtain the entries of Table 1 from the bracketed binomial coefficients displayed in Table 2. Each non-bracketed number in Table 3 is the sum of one or two non-bracketed smaller numbers (directly above and/or above-right) and of one bracketed number (directly above).

The bracketed numbers are those following the zig-zag lines in Table 2 (for illustration purposes, the same types

TABLE 2. Pascal's triangle of binomial coefficients depleted of the last column of 1's

					(1)
				(1)	(2)
			(1)	(3)	(3)
		(1)	(4)	(6)	(4)
	(1)	(5)	(10)	(10)	(5)
(1)	(6)	(15)	(20)	(15)	(6)

TABLE 3. The numerical triangle for the number of terms in R(G) from Table 1 (non-bracketed values)

r	Triangle entries $W_{r,k}$					W_r
3					1	1
				(1)	(1)	
4				2	2	4
			(1)	(1)	(2)	
5			3	5	4	12
		(1)	(1)	(3)	(2)	
6		4	9	12	6	31
	(1)	(1)	(4)	(3)	(3)	
7	5	14	25	21	9	74

of lines have been employed in Tables 2 and 3).Thus each non-bracketed term in Tables 1 or 3 is the sum of one bracketed term (i.e. of one binomial coefficient) with one or two non-bracketed terms.This property may be expressed as follows:

$$W_{r,k} = W_{r-1,k} + W_{r-1,k-1} + f(r,k) \qquad (6)$$

where $f(r,k)$ is a binomial coefficient,namely $\binom{\lfloor (r+k-3)/2 \rfloor}{k-1}$ and $\lfloor x \rfloor$ denotes the integer part of x.

The numbers $f(r,k)$ for $k \geq 2$ generate the zig-zag lines in Table 2; note that their sum is $\sum_{k \geq 2} f(r,k) = F_{r-1} - 1$, where F_r are Fibonacci numbers.

The total number of terms W_r in R(G) displayed on the last column at the right of Table 3 obeys the recurrence relationships:

$$W_r = W_{r-1} + W_{r-2} + 2^{r-2} - 1 \qquad (7)$$

$$W_r = 2W_{r-1} + F_{r-1} - 1 \qquad (8)$$

The complete proofs of these combinatorial properties will be published elsewhere.

As a final application,let all segments be composed of two hexagons (A_i+1=2),leading to a helicene,or to an isoarithmic non-branched catafusene,e.g. a zig-zag catafusene.In this case,all products in Table 1 are equal to 1,so that

$$K_r = 2^r + 1 - W_r \qquad (9)$$

where W_r obeys the above recurrence relationships.

Let us calculate with the help of (9) the sum

$$K_{r+1}+K_r = (2^{r+1}+1-W_{r+1}) + (2^r+1-W_r) = 2 + 3\cdot 2^r - (W_{r+1}+W_r)$$

and by virtue of relationship (7) we obtain

$$K_{r+1} + K_r = 2 + 3\cdot 2^r - W_{r+2} + 2^r - 1 = 2^{r+2} + 1 - W_{r+2} =$$
$$= K_{r+2}$$

i.e. we have demonstrated that in this case we obtain the same recurrence relation as for Fibonacci numbers F_i;since $K_o = 2$,$K_1 = 3$ and $K_2 = 5$ for benzene,naphthalene,and phenanthrene,respectively,the correspondence is

$$K_r = F_{r+2}$$

REFERENCES

1. M. Gordon and W.H.T. Davison, J. Chem. Phys. 20, 428 (1952).
2. N. Trinajstić, "Chemical Graph Theory", The Chemical Rubber Company Press, Boca Raton, Florida, 1983, chapter 10.
3. H. Hosoya and T. Yamaguchi, Tetrahedron Lett. 4659 (1975).
4. N. Ohkami, A. Motoyama, T. Yamaguchi, H. Hosoya and I. Gutman, Tetrahedron, 37, 1113 (1981).
5. G. W. Wheland, J. Phys. Chem. 3, 356 (1935) ; I. Gutman, Theoret. Chim. Acta, 45, 305 (1977) ; Z. Naturforsch. 37 a, 69 (1982) ; J. Aihara, Bull. Chem. Soc. Japan, 50, 2010 (1977) ; 51, 2729 (1978).
6. W. C. Herndon, Tetrahedron, 29, 3 (1973) ; J. Chem. Educ. 51, 10 (1979) ; Israel J. Chem. 20, 270 (1980) ; W. C. Herndon and M. L. Ellzey, J. Amer. Chem. Soc. 96, 6631 (1974).
7. T. F. Yen, Theoret. Chim. Acta, 20, 399 (1971).
8. I. Gutman and D. Cvetković, Croat. Chim. Acta, 46, 15 (1974) ; I. Gutman, ibid., 55, 371 (1982).
9. I. Gutman and H. Hosoya, Theoret. Chim. Acta, 48, 279 (1978).
10. D. Cvetković, I. Gutman and N. Trinajstić, Theoret. Chim. Acta, 34, 129 (1974) ; B. Dzonova-Jerman-Blazić and N. Trinajstić, Croat. Chem. Acta, 55, 347 (1982) ; Computers and Chem. 6, 121 (1882).
11. I. Gutman, Math. Chem. 13, 173 (1982) ; 11, 127 (1981).
12. O. E. Polansky and I. Gutman, Math. Chem. 8, 269 (1980).
13. M. Randić, J. Chem. Soc. Faraday II, 72, 232 (1976) ; Chem. Phys. Lett, 38, 68 (1976) ; Tetrahedron, 31, 1477

(1975) ; Pure Appl. Chem. 52, 1587 (1980).

14. D. Biermann and W. Schmidt, Israel J. Chem. 20, 312 (1980); P. Eilfeld and W. Schmidt, J. Electron Spectr. Related Phenom. 24, 101 (1981).

15. S. El-Basil, Chem. Phys. Lett. 89, 145 (1982) ; S. El-Basil and I. Gutman, ibid. (in press).

16. S. J. Cyvin, Math. Chem. 13, 167 (1982).

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

18. A. T. Balaban, Tetrahedron, 25, 2949 (1969).

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

20. A. T. Balaban, Rev. Roumaine Chim. 22, 45 (1977).