

## REVIEW

APPLICATIONS OF CATERPILLAR TREES  
IN CHEMISTRY AND PHYSICS

Sherif El-BASIL\*

*Department of Chemistry, University of Georgia, Athens, Georgia 30602, USA*Received 10 February 1987  
(in final form 9 April 1987)**Abstract**

The relations of caterpillar trees (which are also known as Gutman trees and benzenoid trees) to other mathematical objects such as polyhex graphs, Clar graphs, king polyominoes, rook boards and Young diagrams are discussed. Potential uses of such trees in data reduction, computational graph theory, and in the ordering of graphs are considered. Combinatorial and physical properties of benzenoid hydrocarbons can be studied via related caterpillars. It thus becomes possible to study the properties of large graphs such as benzenoid (i.e. polyhex) graphs in terms of much smaller tree graphs. Generation of the cyclic structures of wreath and generalized wreath product groups through the use of caterpillar trees is illustrated.

**1. Historical introduction**

The simplest way of defining a caterpillar tree,  $P_n(m_1, m_2, \dots, m_n)$ , is through the concept of the derivative of the graph [1]. Thus, when all the end points of a graph  $G$  are deleted, another graph  $G'$  results, called the derivative of  $G$ . A caterpillar tree is defined to be a tree graph, the derivative of which is a *path*. The name caterpillar was suggested by A. Hobbs [2]. Thus, a caterpillar tree  $P_n(m_1, m_2, \dots, m_n)$  may be constructed by the addition of  $m_1$  monovalent vertices to the first vertex  $v_1$  of path  $P_n$ ,  $m_2$  monovalent vertices to  $v_2$  of  $P_n$ , and so on. An example of a caterpillar tree and another of a noncaterpillar together with other graphs are shown in fig. 1. It seems that Harary and Schwenk were among the first to study these trees in the mathematical literature [1,3,4].

\*Permanent address: Faculty of Pharmacy, Kasr Al-Aini Street, Cairo, Egypt.

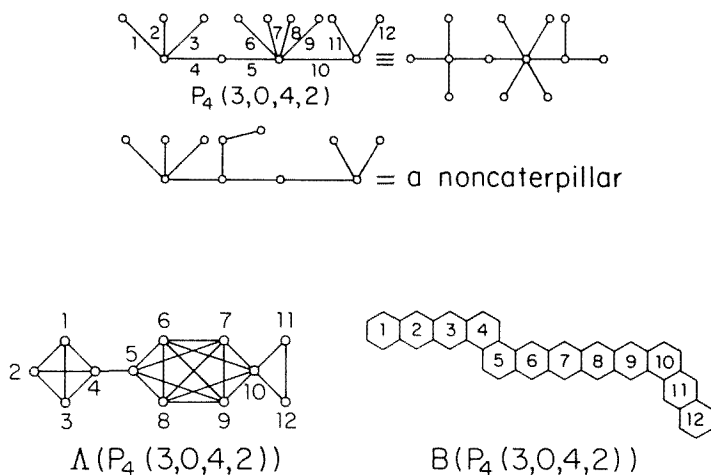


Fig. 1. A caterpillar tree  $P_4(3,0,4,2)$ , the corresponding Clar graph  $\Lambda(P_4(3,0,4,2))$  and the corresponding benzenoid hydrocarbon  $B(P_4(3,0,4,2))$ . A noncaterpillar tree is also shown.

In chemistry, the use of these trees resulted from studying the topological properties of benzenoid hydrocarbons, namely *resonance* relations among individual hexagons of a benzenoid system [5]. Two hexagons in a benzenoid hydrocarbon are called *resonant* if an (aromatic) sextet (i.e. a set of three circularly conjugated double bonds) can be drawn in both of them such that the rest of the carbon atoms are spanned either by a double bond or by a sextet of electrons. Gutman [5] represented such resonance relations among *hexagons* of a benzenoid system by the *edges* of a caterpillar tree: two edges in a caterpillar tree are *incident* if and only if the corresponding hexagons in the benzenoid system are *nonresonant*. Thus, the tree given by  $P_4(3,0,4,2)$  corresponds to the benzenoid hydrocarbon  $B(P_4(3,0,4,2))$  drawn in fig. 1. There is a one-to-one correspondence between the labeling of the edges of the caterpillar and those of the hexagons of the benzenoid system. Explicitly, these terms were considered in chemistry (synonymously under the name "Gutman trees") in three recent papers by this author [6–8].

It is amazing that nearly all graphs that played an important role in what is now called "chemical graph theory" may be related to caterpillar trees. For this reason, such objects are of great importance for understanding and simplifying combinatorial properties of much more complicated graphs. Three main areas involve the use of these trees, viz. *computational* methods, *ordering* [9] and *data reduction* [10]. It is convenient to now give some important definitions for the development of this treatment.

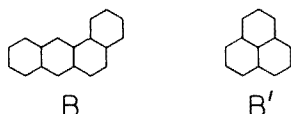
## 2. Definition of important terms

### 2.1. GRAPHS, LATTICES AND DIAGRAMS

A caterpillar tree might be associated with the following objects:

#### 2.1.1. Polyhex graph $B$ [11]

This term was first introduced by Hosoya et al. [11] to mean a graph composed only of hexagons that have an even number  $2\ell$  of points and that can be spanned by disjoint lines. The polyhex  $B$  below is one such type, while  $B'$  is not included in



this definition. Obviously, the number of ways in which  $B$  can be spanned by the  $2\ell$  disjoint lines is well known in organic chemistry as the number of *Kekulé structures* [12] and is described in mathematics as the number of *perfect matchings* [13], a synonym for a famous problem in *dimer statistics* [14].

A molecular network which is entirely composed of hexagons is called benzenoid. If no three hexagons have a common atom, the system is called *catacondensed*. If every hexagon of a catacondensed system has at most two neighboring hexagons, it is said to be *nonbranched*. If there is at least one hexagon in a catacondensed hydrocarbon that is surrounded by three other hexagons, it is said to be *branched*. If in a polyhex graph at least one vertex is common to three hexagons, it is called *pericondensed*.

#### 2.1.2 Clar graphs $\Lambda$ [15,16]

Gutman [15] seems to have been the first to introduce this term in chemistry; the concept was later developed by Gutman and this author [16]. For nonbranched benzenoids, a Clar graph is simply the *line graph* [17] of a caterpillar tree. In fact, it can be seen that *every* caterpillar is associated with a Clar graph (cf. fig. 1).

For branched benzenoid hydrocarbons, however, no caterpillar tree is defined, although a Clar graph can be defined [15] in the following way. Let  $h_1, h_2, \dots, h_n$  be the hexagons of the branched system. Then the vertices of its Clar graph are  $v_1, v_2, \dots, v_n$  such that  $v_i$  is connected to  $v_j$  only if  $h_i$  and  $h_j$  are *nonresonant*.

## Relation between caterpillar trees, Clar and polyhex graphs

At this point, it is important to digress from the relation between the three types of graphs defined above. We start by considering the polyhex graph again. First we observe that two hexagons in a polyhex may or may not be resonant. An illustration is considered below:

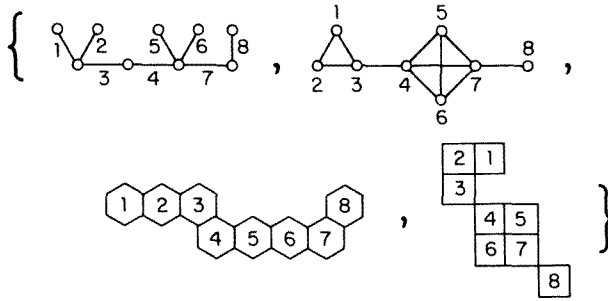


Thus, hexagons 3 and 5 on the left-hand side are *nonresonant*, while 1 and 4 on the right-hand side are.

Gutman demonstrated [5] the following fact: Every nonbranched (cata-condensed) polyhex graph  $B$  which contains  $n$  hexagons is associated with a caterpillar tree  $T$  containing  $n + 1$  vertices such that: two (or more) *incident* edges in  $T$  correspond to two (or more) *nonresonant* hexagons in  $B$  and *vice versa*, i.e. two (or more) *non-incident* edges in  $T$  correspond to two (or more) *resonant* hexagons in  $B$ . For example, consider  $B(P_4(3, 0, 4, 2))$  of fig. 1 and let  $h_i$  refer to the  $i$ th hexagon. One observes that linearly fused rings cannot be resonant. For example, *none* of the hexagons in each of the following sets can be *resonant*:  $\{h_1, h_2, h_3, h_4\}$ ;  $\{h_4, h_5\}$ ;  $\{h_5, h_6, h_7, h_8, h_9, h_{10}\}$ ;  $\{h_{10}, h_{11}, h_{12}\}$ . Similarly, we say: *all* the edges in  $T$  in each of the following sets are *adjacent*:  $\{e_1, e_2, e_3, e_4\}$ ;  $\{e_4, e_5\}$ ;  $\{e_5, e_6, e_7, e_8, e_9, e_{10}\}$ ;  $\{e_{10}, e_{11}, e_{12}\}$ . This one-to-one correspondence extends to the *vertices* of the Clar graph: All the vertices in each of the following sets are *adjacent* [refer to  $\Lambda(P_4(3, 0, 4, 2))$ ]:  $\{v_1, v_2, v_3, v_4\}$ ;  $\{v_4, v_5\}$ ;  $\{v_5, v_6, v_7, v_8, v_9, v_{10}\}$ ;  $\{v_{10}, v_{11}, v_{12}\}$ . These fundamental relations have important implications in understanding the combinatorial structures of benzenoid hydrocarbons, as we shall see later. Because of this relation to benzenoid systems, caterpillar trees will also be called *benzenoid* trees.

### 2.1.3. King polyomino graphs $P$ [18]

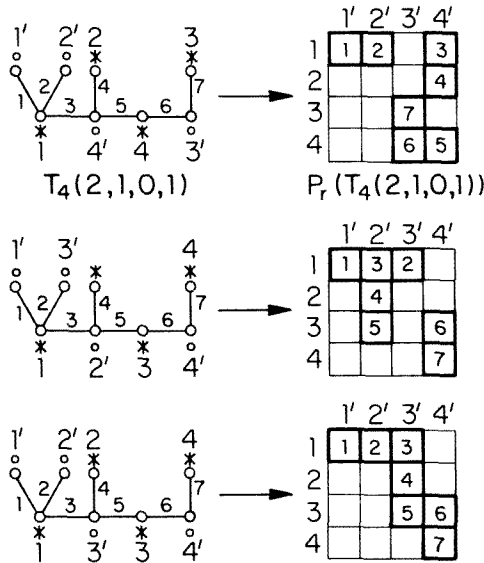
Consider a rectangular lattice composed of cells arranged in a certain number of rows and columns. Such graphs are called polyominoes or square animals. Two cells in a polyomino are defined [18] to be adjacent if they share at least one vertex. The maximum number of adjacent cells is therefore four. This corresponds to four *non-resonant* hexagons annellated in a linear fashion. Because of this, king polyominoes can be made to correspond to a polyhex graph containing linear segments which are *no more* than four hexagons long. An illustration, we consider the following set of graphs whose caterpillar tree is  $P_4(2, 0, 2, 1)$ :



Therefore, for every subset of *adjacent cells* there is a corresponding subset of *non-resonant hexagons, incident edges and adjacent vertices*.

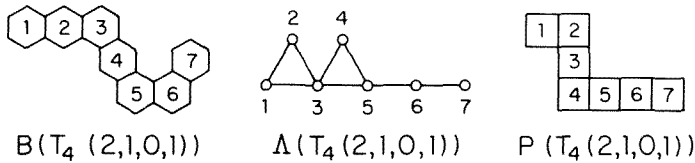
2.1.4. Rook boards  $P_r$  [19]

A rook board is a subset of cells of a  $j \times j$  chessboard. Godsil and Gutman demonstrated [19] that every bipartite [20] graph  $G$  is associated with a rook board such that a cell  $c_{ij}$  (which is located in the  $i$ th row and  $j$ th column of the board) exists only if vertices  $i$  and  $j$  are connected in  $G$ . We illustrate how rook boards which correspond to bipartite caterpillar trees are constructed in the following chart:



In the above chart, three different labelings of the bipartite caterpillar generated three different rook boards. These are shown by the heavily outlined squares. In a rook board, two cells are adjacent if they share the same row and column. For ex-

ample, in the top board the cell labeled 3 is adjacent to 1, 2, 4, 5, while cells in 4 and 7 are *not* adjacent. The three rook boards preserve the combinatorial counts in the caterpillar form which they were generated. Thus, for example, there are six triplets of nonadjacent edges in that tree, viz.  $\{(146), (147), (157), (257), (247), (246)\}$ . The same subsets of cells in all three boards are nonadjacent. Similarly, one can easily demonstrate that there are 13 sets of nonadjacent edges, each of cardinality 2 in the tree, which correspond to 13 such sets of nonadjacent cells in any of the above boards. Such combinatorial counts of nonadjacent structures represent the coefficients of counting polynomials, which will be considered later. Identical combinatorial counts exist in the associated polyhex graph, Clar graph and king polyomino graph, all of which are shown below:



There are one-to-one correspondences between the labelings of the edges of  $T_4(2,1,0,1)$ , the hexagons of  $B$ , the vertices of  $\Lambda$  and the cells of  $P$ . Thus, a knowledge of nonadjacent edges in  $T$  yields details of nonadjacent structures in other graphs, namely, polyhex graphs, Clar graphs, king and rook polyominoes.

## 2.2. POLYNOMIALS AND NONADJACENT STRUCTURES

All polynomials of caterpillar trees and related graphs and lattices are combinatorial descriptors of the "nonadjacent structures" in a given object. It seems that Hosoya [21] was the first to introduce the concept of a nonadjacent structure in chemistry. For a graph  $G$ , he defined a counting polynomial  $H(G; x)$  by

$$H(G; x) = \sum_{k=0}^m p(G; k) x^k, \tag{1}$$

where  $p(G; k)$  is the number of ways of selecting  $k$  nonadjacent edges in  $G$  (i.e.  $k$  edges in which no two of them are adjacent). The term  $p(G; k)$  is called the number of  $k$ -matchings in  $G$ . Conveniently, in  $H(G; x)$  (and in all other combinatorial polynomials)  $p(G; 0)$  is taken to be unity and  $m$  is the maximal value of  $k$ . A more general expression of such polynomials [6,7] is given by eq. (2), viz.

$$F(G; x) = \sum_k^{\max k} \rho \Theta(G; k) x^{f(k,n)}, \tag{2}$$

Table 1  
Some polynomials, associated graphs and graph invariants in chemistry and physics

	Polynomial	$\rho$	Graph invariants set	Associated graph	$\Theta(G; k)$	$f(k, n)$
1	Acyclic = matching $= \alpha(G; x)$	$(-1)^k$	edges	caterpillar tree, $T$	$\rho(G; k)^a$	$n - 2k$
1(a)	Counting (of Hosoya) $= H(G; x)$	1	edges	caterpillar tree, $T$	$\rho(G; k)^a$	$k$
2	Sextet $\equiv \sigma(B; x)$	1	hexagons	polyhex graph, $B$	$r(B; k)^b$	$k$
2(a)	Resonance $A(B; x)$	$(-1)^k$	hexagons	polyhex graph, $B$	$r(B; k)^b$	$2M - 2k$
3	King $\equiv K(P; x)$	1	cells	polyomino graph, $P$	$\kappa(P; k)^c$	$k$
3(a)	Rook $\equiv K(P_r; x)$	1 or $(-1)^k$	cells	polyomino graph, $P$	$\rho(P_r; k)^d$	$k$ or $n - 2k$
4	Independence $\omega(\Lambda; x)$	1	vertices	Clar graph, $\Lambda$	$0(\Lambda; k)^e$	$k$
4(a)	Color $C(G; x)$	1	vertices	arbitrary graph, $G$	$\zeta(G; k)^f$	$k$

<sup>a</sup>Number of selections of  $k$  independent edges  $\in T$  (i.e. no two edges are incident).

<sup>b</sup>Number of selections of  $k$  nonadjacent but mutually resonant hexagons  $B$ .

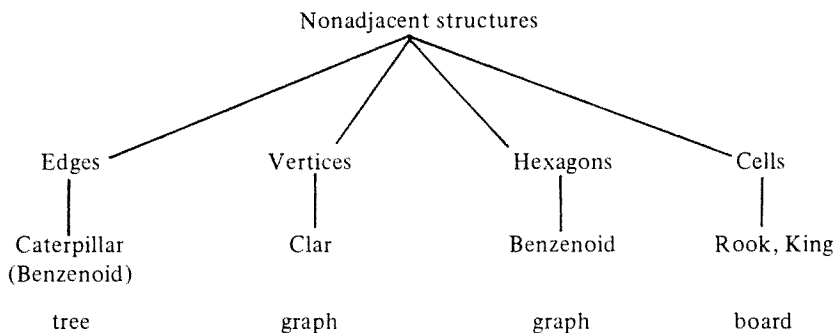
<sup>c</sup>Number of ways of arranging  $k$  non-taking kings on a polyomino graph.

<sup>d</sup>Number of ways of arranging  $k$  non-attacking kings.

<sup>e</sup>Number of selections of  $k$  independent vertices  $\in \Lambda$  (no two are adjacent).

<sup>f</sup>Number of colorings in  $G$  in which there are  $k$  vertices of the same color so that no two of them are adjacent.

where  $\rho$ ,  $\Theta(G; k)$  and  $f(k, n)$  are all functions of the particular polynomial. Table 1 lists several polynomials of some use in chemistry. As an illustration, we consider the sextet polynomial [22] of a benzenoid system  $\sigma(B; x)$ . This important polynomial plays quite a significant role in the chemistry of benzenoid hydrocarbons [23] and was first defined by Hosoya and Yamaguchi [22] for the combinatorial enumeration of the number of Kekulé structures of a benzenoid hydrocarbon. As can be inferred from table 1 and eq. (2),  $\rho = 1$  and  $f(k, n) = k$  for the sextet polynomial. Further, the generating function [24] is given by  $r(B, k) x^k$ , where  $r(B, k)$  is called the  $k$ th resonant number of the polyhex graph of the benzenoid system  $B$ . It measures the number of selections of  $k$  mutually resonant and disjoint hexagons in  $B$ . The "nonadjacent structures" of table 1 may be visualized from the following diagram:



The following identities are easy to establish:

$$\begin{aligned}
 H(T_4(2, 1, 0, 1); x) &= \omega(\Lambda(T_4(2, 1, 0, 1); x)) \\
 &= \sigma(B(T_4(2, 1, 0, 1); x)) = K(P_r(T_4(2, 1, 0, 1); x)) \\
 &= K(P(T_4(2, 1, 0, 1); x)) = 1 + 7x + 13x^2 + 6x^3.
 \end{aligned}$$

Naturally, when we set  $x = 1$  in the generating function, we arrive at the number of Kekulé structures  $K(B) \equiv K$ , a problem which has been a continual focus of interest [25] despite its early history in chemical combinatorics [26]. Thus, a knowledge of the counting polynomial of a given caterpillar leads to other polynomials such as sextet, independence, color, king and rook polynomials (if the latter two boards exist). The above treatment, which applies to nonbranched benzenoid hydrocarbons, can easily be extended to other systems, as explained below.

### 2.3. BRANCHED SYSTEMS

By the application of the appropriate recursive relations of the sextet polynomial [23] one can associate a "pseudo-benzenoid" tree (i.e. a benzenoid tree containing a variable  $x$ ) with virtually any benzenoid hydrocarbon. The principle is simple: choose any row of hexagons and divide the set of Kekulé patterns [23] into the set of distinctive cases so that each vertical line in that row is chosen double. Caution should be taken against the possibility that the chosen double and the resultant fixed double bonds might produce a proper sextet [27] by assigning double bonds to the remaining skeleton. The sextet polynomial of the branched benzenoid hydrocarbon can be written in terms of polynomials of nonbranched systems. The nonbranched fragments can then be transformed into caterpillar trees whose counting polynomials are identical to the sextet polynomials of the nonbranched polyhex



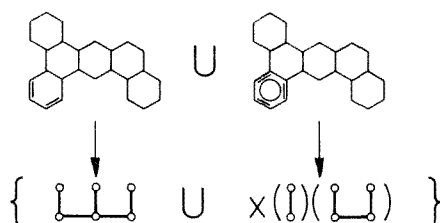


Fig. 2. Recursive generation of a “pseudo-caterpillar” tree of a branched benzenoid hydrocarbon. The factor of  $X$  accounts for the proper sextet [23] in the graph to the right.

graphs. The algorithm is illustrated in fig. 2 for a branched system, where the resulting pseudo-benzenoid tree is shown in braces.

### 3. Benzenoid trees and computation

If the matching polynomial and/or the counting polynomial of a benzenoid tree is known, all other polynomials of related graphs in table 1 become immediately available. Now the matching and counting polynomials are related to each other [28] as follows:

$$\alpha(T; x) = x^N H(T, -x^{-2}); \tag{3}$$

$$H(T; x) = i^{-N} x^{N/2} \alpha(T; ix^{-1/2}), \tag{4}$$

where  $N$  is the number of vertices in the given caterpillar tree. Using elegant operator algebra defined by Hosoya and Ohkami [23,28], it is possible to compute either functions (3) and/or (4) for families of caterpillar trees. As an illustration, we calculate the matching polynomials of the family of caterpillars of  $P_n(3, 3, \dots, 3)$  (where the 3's are repeated  $n$  times). We apply the recursion [29] to the starred edge as follows:

$$(1) \quad \begin{array}{c} \begin{array}{c} \circ \quad \circ \quad \circ \quad \circ \quad \circ \\ \diagdown \quad \diagup \quad \diagdown \quad \diagup \quad \diagdown \\ \circ \quad \circ \quad \circ \quad \circ \quad \circ \\ \diagdown \quad \diagup \quad \diagdown \quad \diagup \quad \diagdown \\ \circ \quad \circ \quad \circ \quad \circ \quad \circ \\ \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \\ \circ \quad \circ \quad \circ \quad \circ \quad \circ \\ \diagdown \quad \diagup \quad \diagdown \quad \diagup \quad \diagdown \\ \circ \quad \circ \quad \circ \quad \circ \quad \circ \\ \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \\ \circ \quad \circ \quad \circ \quad \circ \quad \circ \end{array} \\ T_n \end{array} = \begin{array}{c} \circ \cdot \begin{array}{c} \circ \quad \circ \quad \circ \quad \circ \quad \circ \\ \diagdown \quad \diagup \quad \diagdown \quad \diagup \quad \diagdown \\ \circ \quad \circ \quad \circ \quad \circ \quad \circ \\ \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \\ \circ \quad \circ \quad \circ \quad \circ \quad \circ \\ \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \\ \circ \quad \circ \quad \circ \quad \circ \quad \circ \end{array} \\ K_n \end{array} - \begin{array}{c} \circ \cdot \begin{array}{c} \circ \quad \circ \quad \circ \quad \circ \quad \circ \\ \diagdown \quad \diagup \quad \diagdown \quad \diagup \quad \diagdown \\ \circ \quad \circ \quad \circ \quad \circ \quad \circ \\ \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \\ \circ \quad \circ \quad \circ \quad \circ \quad \circ \\ \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \\ \circ \quad \circ \quad \circ \quad \circ \quad \circ \end{array} \\ T_{n-1} \end{array}$$

$$(2) \quad \begin{array}{c} \begin{array}{c} \circ \quad \circ \quad \circ \quad \circ \quad \circ \\ \diagdown \quad \diagup \quad \diagdown \quad \diagup \quad \diagdown \\ \circ \quad \circ \quad \circ \quad \circ \quad \circ \\ \diagdown \quad \diagup \quad \diagdown \quad \diagup \quad \diagdown \\ \circ \quad \circ \quad \circ \quad \circ \quad \circ \\ \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \\ \circ \quad \circ \quad \circ \quad \circ \quad \circ \\ \diagdown \quad \diagup \quad \diagdown \quad \diagup \quad \diagdown \\ \circ \quad \circ \quad \circ \quad \circ \quad \circ \\ \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \\ \circ \quad \circ \quad \circ \quad \circ \quad \circ \end{array} \\ K_n \end{array} = \begin{array}{c} \circ \cdot \begin{array}{c} \circ \quad \circ \quad \circ \quad \circ \quad \circ \\ \diagdown \quad \diagup \quad \diagdown \quad \diagup \quad \diagdown \\ \circ \quad \circ \quad \circ \quad \circ \quad \circ \\ \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \\ \circ \quad \circ \quad \circ \quad \circ \quad \circ \\ \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \\ \circ \quad \circ \quad \circ \quad \circ \quad \circ \end{array} \\ J_n \end{array} - \begin{array}{c} \circ \cdot \begin{array}{c} \circ \quad \circ \quad \circ \quad \circ \quad \circ \\ \diagdown \quad \diagup \quad \diagdown \quad \diagup \quad \diagdown \\ \circ \quad \circ \quad \circ \quad \circ \quad \circ \\ \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \\ \circ \quad \circ \quad \circ \quad \circ \quad \circ \\ \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \\ \circ \quad \circ \quad \circ \quad \circ \quad \circ \end{array} \\ T_{n-1} \end{array}$$



It is easy to show that, for the general case  $T_n(m, m, \dots, m)$ , the operator in eq. (8) becomes:

$$\hat{O}^2 + (mx^{m-1} - x^{m+1}) \hat{O} + x^{2m} = 0. \tag{10}$$

Using eq. (10) and the two identities

$$\alpha(T_1(1); x) = x^2 - 1; \quad \alpha(T_2(1, 1); x) = x^4 - 3x^2 + 1,$$

followed by repeated application of eq. (10) (for  $m = 1$ ) leads to

$$\begin{aligned} \alpha(T_8(1, 1, 1, 1, 1, 1, 1, 1); x) &= \alpha(T_8(1^8); x) \\ &= x^{16} - 15x^{14} + 74x^{12} - 290x^{10} + 258x^8 - 290x^6 + 74x^4 - 15x^2 + 1. \end{aligned} \tag{11}$$

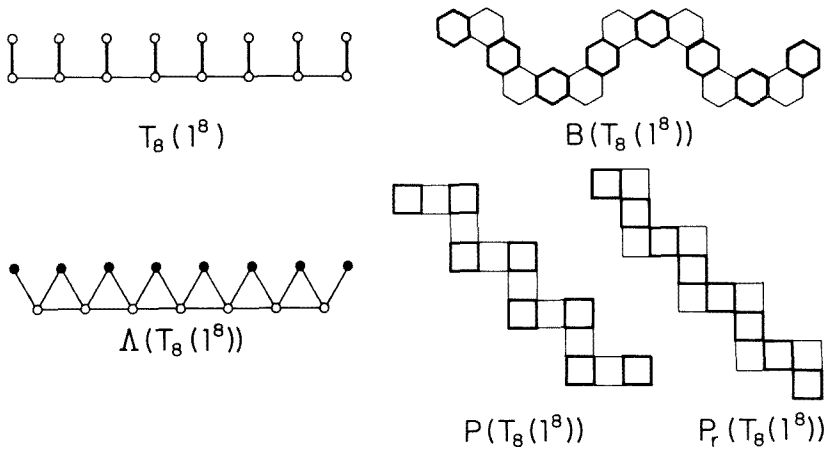


Fig. 3. Illustration of eq. (13) for a caterpillar tree (i.e. benzenoid tree) and its associated graphs for  $\max k = 8$ . Simple application of eq. (10) shows that there are 290 ways of placing either 3 or 5 non-attacking kings on the chessboard  $P$  or  $P_r$ . The subset of invariants leading to  $x^8$  is heavily outlined.

The graphs of this polynomial are shown in fig. 3. Using relations (3) or (4), we can write the following identities (see table 1):

$$\begin{aligned} H(T_8(1^8); x) &= \omega(\Lambda(T_8(1^8); x)) = \sigma(B(T_8(1^8); x)) \\ &= K(P(T_8(1^8); x)) = K(P_r(T_8(1^8); x)) \\ &= 1 + 15x + 74x^2 + 290x^3 + 258x^4 + 290x^5 + 74x^6 + 15x^7 + x^8. \end{aligned} \tag{12}$$

So we know immediately, for instance, that there are 290 ways of placing 3 or 5 non-attacking kings on the chessboard shown in fig. 3. It is interesting to observe that there is *only one* way of selecting a subset of nonadjacent structures of maximum cardinality (8 in this case). Polynomials of such types as counting, matching, etc. of graphs are symmetric in the sense of eq. (13), viz.,

$$a_i = a_{N-i}, \quad (13)$$

where  $N$  is the number of vertices in  $T$  (or  $N-1 =$  the number of hexagons in  $B =$  number of vertices in  $\Lambda$ , etc.). Benzenoid hydrocarbons for which eq. (13) holds (i.e. with a "symmetric" sextet polynomial) are known to have a *single sextet formula* (i.e. a single Clar representation [30]). For such types, the number of aromatic sextets they contain is very close to the number of sextet-type resonance interactions per Kekulé structure [31]. This last statement is known as Aihara's conjecture [6,31]. Aihara's observation is important because it specifies a condition of the benzenoid hydrocarbon such that the simple Clar sextet formalism [30] roughly estimates its Dewar-type resonance energy [32]. Gutman [33] commented on Aihara's conjecture by defining a function  $F(B)$  as follows:

$$F(B) = M(B)K(B) - 2 \sum_i^h K(B-H_i), \quad (14)$$

where  $M(B) = \max k =$  the maximum cardinality of a set of mutually resonant but disjoint aromatic sextets in the benzenoid graph, and  $K(B)$  is its Kekulé count. The summation of the second term is taken over all hexagons  $H_i$  of  $B$ , where  $h$  is the total number of  $H_i$ . Gutman specified Aihara's condition that the hydrocarbon be represented by a single sextet formula by having

$$r(B; \max k) = 1. \quad (15)$$

Figure 4 shows an example of two hydrocarbons which satisfy Aihara's conjecture, another that does not. Furthermore, Gutman restated Aihara's postulate by defining  $F(B) = 0$  whenever eq. (15) holds. Benzenoid hydrocarbons for which  $F(B)$  vanishes define the "best conditions", where nearly all existing resonance-structure theories apply, viz. Clar's [30], Dewar's [32] conjugated circuits theory by Randić [34], as well as the structure-resonance theory by Herndon [35]. In fact, all such theories coincide for benzenoid systems for which eq. (15) applies.

For benzenoid trees not possessing elements of symmetry or not belonging to a given periodic network of trees, the method of Balasubramanian and Randić [36] becomes particularly suitable for computation of  $\alpha(T; x)$ . For example, for the caterpillar tree  $P_4(2, 0, 2, 1)$  shown above, we have the following identities:

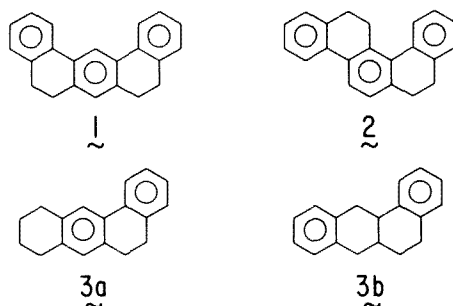


Fig. 4. Examples of benzenoid hydrocarbons which possess one Clar representation (**1** and **2**) and a hydrocarbon with two Clar representations (**3a** and **3b**). Hydrocarbons **1** and **2** satisfy Aihara's conjecture [eq. (15)].

The quotient tree,  $Q = P_4$

$$\alpha_1 = \alpha_3 = \alpha(T_1(2)) = x^3 - 2x, \text{ thus } \alpha'_1 = \alpha'_3 = x^2$$

$$\alpha_2 = (P_1(0)) = x; \alpha'_2 = 1$$

$$\alpha_4 = x^2 - 1; \alpha'_4 = x. \quad (16)$$

From the adjacency matrix of  $Q$ ,  $\alpha(T; x) = \alpha(P_4(2, 0, 2, 1); x)$  is given by the following determinant:

$$\begin{vmatrix} -(x^3 - 2x) & x^2 & 0 & 0 \\ 1 & -x & 1 & 0 \\ 0 & x^2 & -(x^3 - 2x) & x^2 \\ 0 & 0 & x & -(x^2 - 1) \end{vmatrix}. \quad (17)$$

In the above notation, the primed letter denotes the matching polynomial of a type [such as  $T_1(2)$ ] after its root vertex has been pruned. The notation is essentially that used in ref. [36].

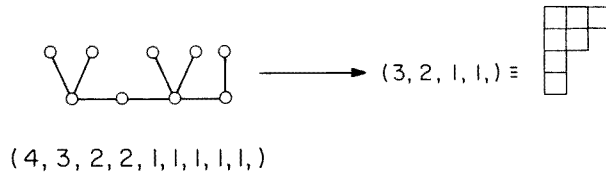
#### 4. Caterpillar (benzenoid) trees and the ordering of graphs; A relation with Young diagrams

In their work on algebraic characterization of skeletal branching, Gutman and Randić [37] used the theorems by Muirhead [38] to *order* and *compare* a set of

trees (caterpillars and non-caterpillars). In their treatment, two trees are characterized by a sequence of nonnegative integers  $\{a_1, a_2, \dots, a_k\}$  and  $\{b_1, b_2, \dots, b_j\}$  representing the degrees of their vertices when listed in descending order. For example,  $P_4(3,0,4,2)$  would be associated with the sequence  $\{6,4,3,2,1,1,1,1,1,1,1,1\}$ . Muirhead's conditions state that  $T_a$  (whose sequence is  $\{a_1, a_2, \dots, a_k\}$ ) is greater than  $T_b$  (whose sequence is  $\{b_1, b_2, \dots, b_j\}$ ) if

$$\begin{aligned}
 a_1 &\geq b_1 \\
 a_1 + a_2 &\geq b_1 + b_2 \\
 &\vdots \\
 a_1 + a_2 + \dots + a_k &= b_1 + b_2 + \dots + b_j.
 \end{aligned}
 \tag{18}$$

Whenever  $a_m + a_n + \dots + a_o \geq b_m + b_n + \dots + b_o$  but  $a_r + a_s + \dots + a_t \geq b_r + b_s + \dots + b_t$ , the two tree graphs are said to be *noncomparable*. The latter trees lead to bifurcation sites in the ordering hierarchy. Using the above criteria, Gutman and Randić [37] ordered sets of trees for  $N = 8, 9$  and  $10$ . Furthermore, these authors [37] made the very interesting observation that their ordering of trees can be made to overlap with the Ruch and Schönhofer [39] ordering of a set of Young diagrams if: (a) information on the terminal vertices is suppressed, and (b) the valency of each vertex is reduced by one. This significant result leads to a relation between a Young diagram and a caterpillar tree and hence between a Young diagram and nearly all other graphs and lattices used in chemistry and physics, namely, Clar graphs, king polyomino graphs, rook boards and polyhex graphs. As an illustration, the Young diagram which corresponds to the set of graphs of  $T = P_4(2,0,2,1)$  is shown below:



This will be denoted here as  $Y(3,2,1,1)$ . There is a unique Young diagram for every caterpillar tree (or any of its associated graphs), though the reverse is not true, i.e. two (or more) caterpillars may be related to the same Young diagram. The following examples (from the set  $N = 8$ ) serve to illustrate this:

$$\begin{aligned}
 \{P_3(4,0,1), P_3(1,3,1)\} &\in Y(4,1,1); \\
 \{P_3(3,0,2), P_3(3,1,1), P_3(2,2,1)\} &\in Y(3,2,1);
 \end{aligned}$$

$$\{P_4(3,0,0,1), P_4(1,2,0,1)\} \in Y(3,1,1,1);$$

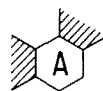
$$\{P_4(2,0,0,2), P_4(2,0,1,1), P_4(2,1,0,1), P_4(1,1,1,1)\} \in Y(2,2,1,1);$$

$$\{P_5(2,0,0,0,1), P_5(1,0,1,0,1), P_5(1,1,0,0,1)\} \in Y(2,1,1,1,1).$$

Figure 5 shows the ordering of all Young diagrams containing six boxes. Figure 6 is the corresponding order of the nonbranched benzenoid systems which correspond to the caterpillar trees. The numbers in parentheses are, respectively,  $\gamma_1, \gamma_2, \gamma_3$  and  $\gamma_4$ , where  $\gamma_i$  is a permutation integral (by Herndon [35]) involving permutation of  $(4i + 2)$  pi electrons in the benzenoid systems. Twice these numbers enumerate  $R_1, R_2, R_3$  and  $R_4$ , respectively, where  $R_i$  is a conjugated circuit [34] containing  $(4i + 2)$  pi electrons. It is obvious that the second number  $\gamma_2$  (or  $R_2$ ) is almost constant for the same level of ordering of the Young diagram. In fig. 6, the polyhex graphs are represented by their LA sequences [5]. Thus, a hexagon may be annellated in two ways, viz.,



Linear



Angular

and by convention [5], the terminal hexagons are labeled by L. Thus,  $B(P_4(3,0,4,2))$  is denoted by LLLAALLLALL or  $L^3A^2L^4AL^2$ . In fig. 6, the nonbranched benzenoid hydrocarbons containing seven hexagons are ordered. The numbers in parentheses are, respectively,  $\gamma_1, \gamma_2, \gamma_3$  and  $\gamma_4$ , where  $\gamma_i$  is a Herndon permutation integral [35] involving permutation of  $(4i + 2)$  pi electrons. Naturally, twice these numbers lead to the corresponding conjugated circuits [34]  $R_1, R_2, R_3$  and  $R_4$ .

It is emphasized here that through relating benzenoid trees to other graphs in addition to benzenoid hydrocarbons, such as Clar graphs, king polyominoes and rook boards, they can all be ordered according to schemes by Ruch and Schönhofer based on Young diagrams [39].

## 5. Benzenoid trees and data reduction

An important part of an analysis of chemical data is the data-reduction step. In the past, this involved mainly curve-fitting procedures. The role of graph theory was recognized in the work of Smolenskii [40] and later of Gordon and Kennedy [41]. The dualist graphs by Balaban and Harary [42] might also be regarded as a type of structure reduction, representing polyhex graphs of benzenoid hydrocarbons. Recently,

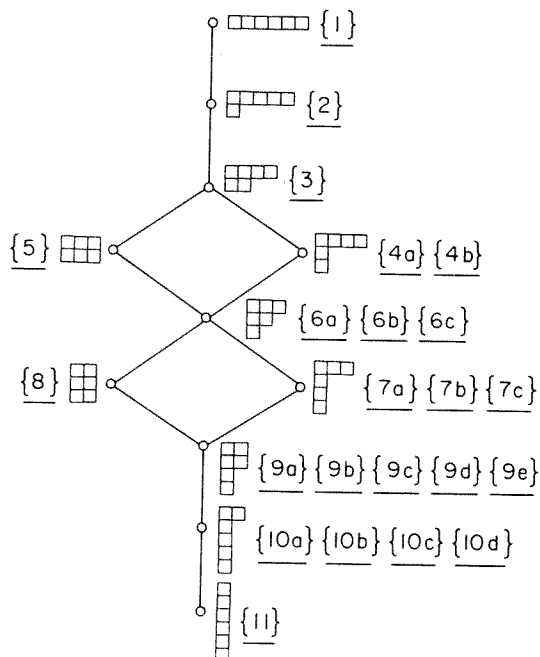


Fig. 5. Ruch ordering of all Young diagrams containing six boxes. Site of bifurcations indicate non-comparable diagrams,

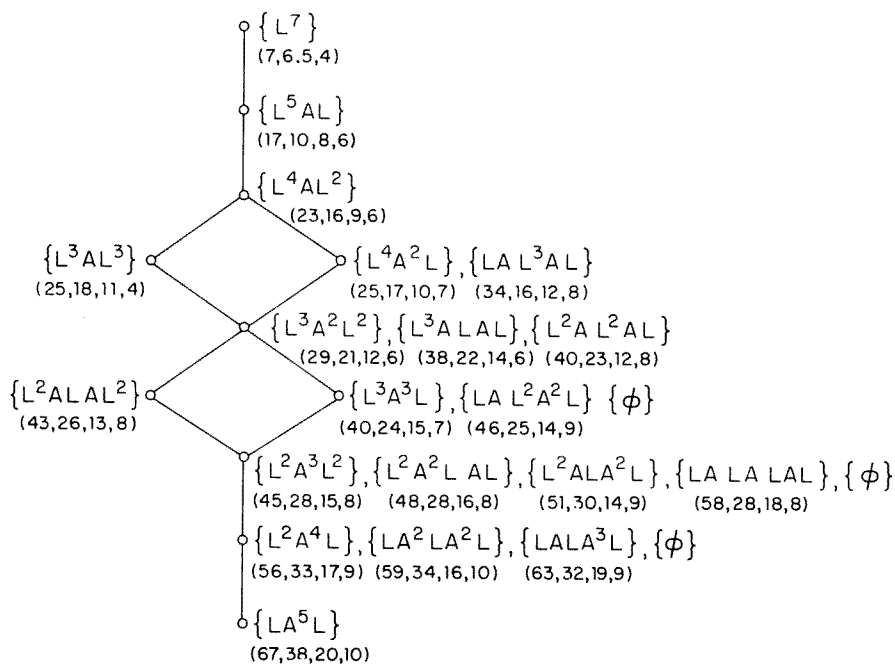


Fig. 6. Ordering of nonbranched benzenoid hydrocarbons which are in one-to-one correspondence with the Young diagrams shown in fig. 5. The polyhex graphs are denoted by their L-A sequences [5]. Numbers in parentheses are  $(\gamma_1, \gamma_2, \gamma_3, \gamma_4)$ , respectively, where  $\gamma_i$  is a Herndon permutation integral [35] involving permutation of  $(4i + 2)$  pi electrons. Twice these numbers lead to  $(R_1, R_2, R_3, R_4)$ , the sequences of the corresponding conjugation circuits [34].



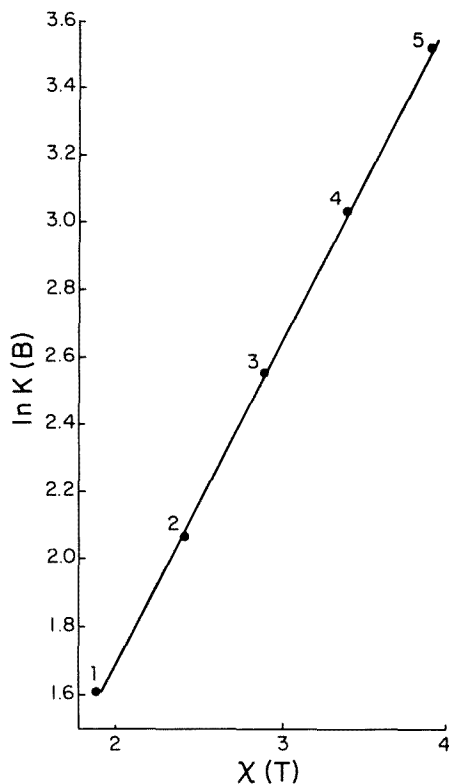


Fig. 7. Correlation between  $\ln K(B)$ , i.e. the natural logarithms of the Kekulé counts of the zigzag polyacenes (1 = phenanthrene, 2 = chrysene, 3 = picene, 4 = fulminene, . . .) and  $\chi(T)$ , the connectivity indices of the relevant caterpillar (i.e. benzenoid trees).

this author [43] explored, for the first time, the possibility of using benzenoid trees to store and retrieve information on related benzenoid systems (i.e. a benzenoid system whose sextet polynomial is identical to the counting polynomial of the tree). Several physical and combinatorial properties, including electronic absorption spectra, heats of atomizations, number of conjugated circuits, number of self-avoiding walks, and number of Sachs graphs are studied, and in all cases excellent correlations are found between the natural logarithms of a property of the benzenoid hydrocarbon and simple powers of the connectivity index of its tree graph [44]. As an illustration, fig. 7 shows a plot of the number of Kekulé structures in logarithmic units of a homologous series of the zigzag polyacenes and the connectivity indices of their trees  $\chi(T)$ 's, given by

$$\chi(T) = \sum (d_i d_j)^{-1/2}, \quad (19)$$

where the summation is taken over all edge types in  $T$ ,  $(d_i d_j)$ 's ( $d_i$  is the degree, i.e. valency of vertex  $i$  in  $T$ ).

**6. Other avatars of caterpillar trees. Generation of cycle indices of wreath product groups [45]**

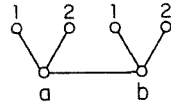
The *composition* [45] of two groups  $A$  and  $B$  is denoted by  $A[B]$  (read:  $A$  around  $B$ ) and is known as the "wreath product" or the "Gruppenkranz". A permutation in  $A[B]$  is given by

$$(\alpha_i \beta_1, \beta_2, \dots, \beta_d)(x_i, y_j) = (\alpha x_i, \beta_i Y_j), \tag{20}$$

where  $\alpha \in A, \beta \in B$  and the sequence  $\beta_1, \dots, \beta_d$  may not involve necessarily distinct elements. The elements  $(x_i, y_j)$  arise from the cross product of  $x = \{x_1, x_2, \dots, x_d\}$  and  $Y = \{y_1, y_2, \dots, y_e\}$ . Wreath product groups have a number of chemical and physical applications, recently discovered by Balasubramanian [46], who revived interest in Pólya's theorem [47]. We show here that caterpillar trees might be used as a model to visualize the cyclic structure and the operations of eq. (20); we take, as an illustration, the group  $S_2[S_2]$ . Then we have:  $x = \{a, b\}, Y = \{1, 2\}$  and

$$x \times Y = \{a1, a2, b1, b2\}.$$

The system might be envisaged as



The operations in  $S_2$ 's are:  $(1)(2), (12), (a)(b), (ab)$ . Then the following  $2 \cdot 2^2 = 8$  elements exist in  $S_2[S_2]$ , viz.,

- (1)  $((a)(b); (1)(2), (1)(2))$
- (2)  $((a)(b); (12), (12))$
- (3)  $((a)(b); (1)(2), (12))$
- (4)  $((a)(b); (12), (1)(2))$

and four other elements using the operation  $(ab)$  instead of  $(a)(b)$ .

The element (1) is simply the identity element which corresponds to the operation



Such an element generates six 1-cycles, i.e.  $s^6$ . The element (2) operates as follows:

$$((a)(b); (12), (12)) (a1) = (a2)$$

$$((a)(b); (12), (12)) (a2) = (a1)$$

$$((a)(b); (12), (12)) (b1) = (b2)$$

$$((a)(b); (12), (12)) (b2) = (b1) .$$

Hence, this element is given by  $(a1 a2)(b1 b2)$  and corresponds to



The cyclic contribution from this element is  $s_2^2$ , i.e. two 2-cycles.

The third element involves the following operations

$$((a)(b); (1)(2), (12)) (a1) = (a1)$$

$$((a)(b); (1)(2), (12)) (a2) = (a2)$$

$$((a)(b); (1)(2), (12)) (b1) = (b2)$$

$$((a)(b); (1)(2), (12)) (b2) = (b1)$$

which may be represented as  $(a1)(a2)(b1 b2)$ , i.e. contributes  $s_1^2 s_2$ , i.e. two 1-cycles and one 2-cycles to the cyclic structure of  $S_2 [S_2]$ . The operation of this element can be modeled by a caterpillar tree as



The fourth element permutes the  $\{X \times Y\}$  set in the following manner:

$$((a)(b); (12), (1)(2)) (a1) = (a2)$$

$$((a)(b); (12), (1)(2)) (a2) = (a1)$$

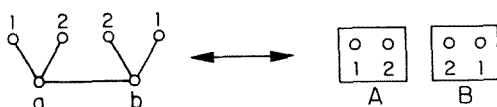
$$((a)(b); (12), (1)(2)) (b1) = (b1)$$

$$((a)(b); (12), (1)(2)) (b2) = (b2) .$$

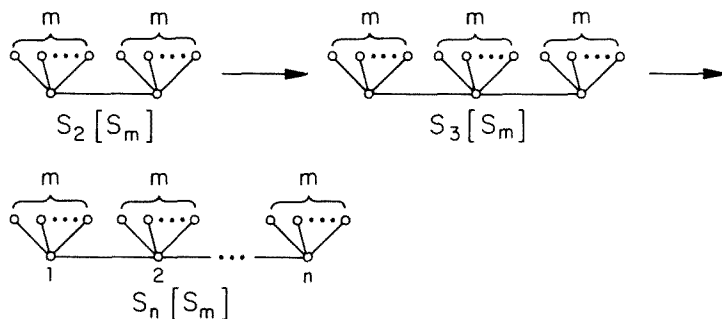
Hence, this element is represented by  $(a1 a2)(b1)(b2)$  and also contributes  $s_1^2 s_2$ . The caterpillar model is shown below:



This "caterpillar modeling" of the operations of the composition of two groups facilitates understanding of the abstract algebra involved in the definition. A similar model was considered by Balasubramanian [48] when he used a "particle-in-box" model; thus,



Actually, either model generates the permutation group of the non-rigid  $N_2H_4$  molecule [48] (i.e. the nitrogen atoms are represented by the root of vertices of  $P_2(2,2)$ , while its monovalent vertices represent the hydrogen atoms). The above modeling can be extended as shown below:



Extension to the recently defined [49] generalized wreath product is also possible. Thus, the NMR group of butane might be modeled by  $P_4(3, 2, 2, 3)$ , which represents  $S_2 [S_3, S_2]$ .

## 7. Conclusions

Although caterpillar (Benzenoid = Gutman) trees are not widely known in chemical literature, their uses span a wide range of applications including data reduction, computations, ordering and modeling notations of abstract groups (such as wreath and generalized wreath product groups) which are necessary for NMR spectroscopy and counting diastereomers [49].

## Acknowledgements

I thank the Office of the U.S. Navy for partial support of this work, and Professor R.B. King for illuminating discussions. I am particularly grateful to Professor Haruo Hosoya of the Ochanomizu University (Tokyo, Japan) for correspondence regarding his work on recursive relations of the sextet polynomial [23]. Travel assistance from the Fullbright Commission in Cairo is appreciated. Appreciation is also due to Dr. Hany El-Hosainy of the Mathematics Department, Cairo University, for discussions on wreath product groups.

## References

- [1] F. Harary and A.J. Schwenk, *Discrete Math.* 6(1973)359.
- [2] Ref. [1], p. 361.
- [3] F. Harary and A.J. Schwenk, *Mathematika* 18(1971)138.
- [4] F. Harary and A.J. Schwenk, *Utilitas Math.* 1(1972)203.
- [5] I. Gutman, *Theor. Chim. Acta* 45(1977)309.
- [6] S. El-Basil, *J. Chem. Soc., Faraday Trans. 2*, 82(1986)299.
- [7] S. El-Basil, *Theor. Chim. Acta* 65(1984)199.
- [8] S. El-Basil, *Theor. Chim. Acta* 65(1984)191.
- [9] Ordering of structures has been explicitly considered in both the mathematical and chemical literature. Thus, J.F. Nagle, *J. Math. Phys.* 7(1966)1588, considered a general linear ordering relation for graphs with the same number of vertices, while E. Ruch, *Theor. Chim. Acta* 38(1975)167, considered ordering of Young diagrams. M. Randić and C.L. Wilkins, *Chem. Phys. Lett.* 63(1979)332; *J. Phys. Chem.* 83(1979)1525, considered ordering of alkanes and how it reflects on their physical properties.
- [10] Here, the term *data reduction* means studying the properties of a large molecule in terms of those of a smaller one (see refs. [4041]).
- [11] The term polyhex graphs means the molecular graph of a benzenoid hydrocarbon. This term was in used in: N. Ohkami, A. Motoyama, T. Yamaguchi, H. Hosoya and I. Gutman, *Tetrahedron* 37(1981)1113.
- [12] A recent account may be found in: N. Trinajstić, *Chemical Graph Theory*, Vol. 2 (CRC Boca Raton, Florida, 1983).
- [13] Cf. H. Sachs, *Combinatorica* 4, 1(1984)89.
- [14] C. Domb and M.S. Green, *Phase Transitions and Critical Phenomena*, Vols. 1–3, ed. M.S. Green (Academic Press, London, 1972).
- [15] I. Gutman, *Z. Naturforsch.* a37(1982)69.
- [16] I. Gutman and S. El-Basil, *Z. Naturforsch.* a39(1984)276.
- [17] Line graphs are defined in: F. Harary, *Graph Theory* (Addison-Wesley, Reading, 1969) Ch. 8.
- [18] King polyomino graphs were first defined in: A. Motoyama and H. Hosoya, *J. Math. Phys.* 18(1977)1485.
- [19] Cf. C.D. Godsil and I. Gutman, *Croat. Chem. Acta* 54(1981)53.
- [20] A bipartite graph (bigraph bicolorable graph)  $G$  is a graph whose vertex set  $V$  can be partitioned into two subsets  $V_1$  and  $V_2$  such that every line of  $G$  joins  $V_1$  with  $V_2$  (see ref. [14], p. 17).
- [21] H. Hosoya, *Bull. Chem. Soc. Japan* 44(1971)2332.

- [22] H. Hosoya and T. Yamaguchi, *Tetrahedron Lett.* (1975) 4659.
- [23] N. Ohkami and H. Hosoya, *Theor. Chim. Acta* 64(1983)153.
- [24] Generating functions are defined in books on combinatorics: V. Krishnamurthy, *Combinatorics: Theory and Applications* (E. Horwood, New York, Halsted Press, 1986). The sextet polynomial [21] is a very special form of generating functions which generate all possible ways in which one can perfectly match [10] the edges of a polyhex graph [8].
- [25] A recent generalized approach to structure count is found in: B. Rušćić, N. Trinajstić and P. Křivka, *Theor. Chim. Acta* 69(1986)107.
- [26] One of the earliest works pertaining to chemical combinatorics is that of: M. Gordon and W.T. Davison, *J. Chem. Phys.* 20(1952)428.
- [27] Proper and improper sextets are defined in ref. [22].
- [28] H. Hosoya and N. Ohkami, *J. Comput. Chem.* 4(1983)585.
- [29] E. Heilbronner, *Helv. Chim. Acta* 36(1953)171.
- [30] Details of some aspects of Clar sextet theory may be found in: I. Gutman, *Bull. Soc. Chim. Beograd* 47(1982)453.
- [31] J. Aihara, *Bull. Chem. Soc. Japan* 49(1976)1429.
- [32] A "Dewar-type" resonance theory depends on the concept of a reference structure: M.J.S. Dewar and C. de Llano, *J. Amer. Chem. Soc.* 91(1969)789.
- [33] I. Gutman, *Croat. Chem. Acta* 56(1983)365.
- [34] M. Randić, *J. Amer. Chem. Soc.* 99(1977)444; *Tetrahedron* 33(1977)1905.
- [35] W.C. Herndon and M.L. Ellzey, Jr., *J. Amer. Chem. Soc.* 96(1974)6631.
- [36] K. Balasubramanian and M. Randić, *Theor. Chim. Acta* 61(1982)307; see also K. Balasubramanian, *Int. J. Quantum Chem.* 22(1982)581.
- [37] I. Gutman and M. Randić, *Chem. Phys. Lett.* 47(1977)15.
- [38] G.H. Hardy, J.E. Littlewood and G. Pólya, *Inequalities* (Cambridge University Press, London, 1934) p. 44.
- [39] E. Ruch, *Theor. Chim. Acta* 38(1975)167; E. Ruch and A. Schönhofer, *Theor. Chim. Acta* 19(1970)225.
- [40] A. Smolenskii, *Russ. J. Phys. Chem.* (English translation) 38(1964)700.
- [41] M. Gordon and J.W. Kennedy, *J. Chem. Soc., Faraday Trans. 2*, 68(1972)484.
- [42] A.T. Balaban and F. Harary, *Tetrahedron* 24(1968)2505.
- [43] S. El-Basil, *Chem. Phys. Lett.* (1987), in press.
- [44] M. Randić, *J. Amer. Chem. Soc.* 97(1975)6609.
- [45] F. Harary, *Graph Theory* (Addison-Wesley, Reading, 1969) p. 164.
- [46] K. Balasubramanian, *Chem. Rev.* 85(1985)599; and numerous references cited therein.
- [47] G. Pólya, *Acta Math.* 65(1937)145.
- [48] K. Balasubramanian, *Studies in Physical and Theoretical Chemistry* 23(1982)149.
- [49] K. Balasubramanian, *Theor. Chim. Acta* 51(1979)37.