AN ALGEBRAIC EXPRESSION FOR THE NUMBER OF KEKULÉ STRUCTURES OF BENZENOID CHAINS

Ratko Tošić and Olga Bodroža
University of Novi Sad, Dr. Ilije Djuričića 4, 21000 Novi Sad, Yugoslavia
(Submitted December 1988)

1. Introduction

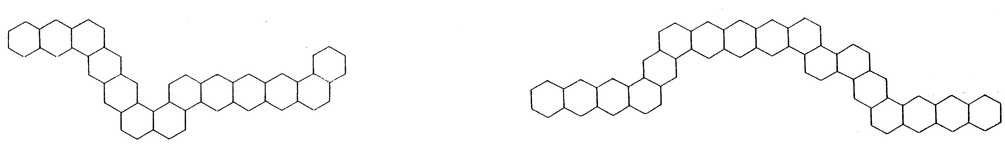
The enumeration of Kekulé structures for benzenoid polycyclic hydrocarbons is important because the stability and many other properties of these hydrocarbons have been found to correlate with the number of Kekulé structures. Starting with the algorithm proposed by Gordon & Davison [8], many papers have appeared on the problem of finding the "Kekulé structure count" K for such hydrocarbons. We can mention here only a few authors who contributed to this topic: Balaban & Tomescu [1, 2, 3, 4], Gutman [10, 11, 12], Herndon [13], Hosoya [12, 14], Sachs [16], Trinajstić [17], Farrell & Wahid [6], Fu-ji & Rong-si [8], Artemi [1], Yamaguchi [14]. A whole recent book [5] is devoted to Kekulé structures in benzenoid hydrocarbons.

In this paper we consider only undirected graphs comprised of 6-cycles. Let there be a total of m such cycles, which we shall denote as $C_1, C_2, \ldots, C_m$ in each graph of interest. Because the problem we treat arises from chemical studies of certain hydrocarbon molecules, we impose upon $C_1, C_2, \ldots, C_m$ the following conditions to reflect the underlying chemistry:

(i) Every C_i and C_{i+1} shall have a common edge denoted by e_i, for all $1 \le i \le m - 1$.

(ii) The edges e_i and e_j shall have no common vertex for any $1 \le i < j \le m - 1$.

Representing the 6-cycles as regular hexagons in the plane results in a graph such as that illustrated in Figures 1(a) and 1(b). In organic chemistry, such graphs correspond to benzenoid chains (each vertex represents a carbon atom or CH group, and no carbon atom is common to more than two 6-cycles).



(a) (b)

FIGURE 1

2. Definitions and Notation

By $L(x_1, x_2, \ldots, x_n)$, we denote a benzenoid chain (i.e., a corresponding graph) composed from n linearly condensed portions (segments) consisting of $x_1, x_2, \ldots, x_n$ hexagons, respectively. Figures 1(a) and 1(b) show $L(3, 4, 2, 2, 5, 2)$ and $L(4, 3, 5, 2, 2, 3, 4)$, respectively.

Any two adjacent linear segments are considered as having a common hexagon. The common hexagon of two adjacent linear segments is called a "kink." The chain $L(x_1, x_2, \ldots, x_n)$ has exactly $n - 1$ kinks. So the total number of hexagons in $L(x_1, x_2, \ldots, x_n)$ is $m = x_1 + x_2 + \cdots + x_n - n + 1$. Observe that such notation implies $x_i \ge 2$, for $i = 1, 2, \ldots, n$.

We adopt the following notation:

$K_n(x_1, x_2, \ldots, x_n)$ is the number of Kekulé structures (perfect mathcings) of $L(x_1, x_2, \ldots, x_n)$.

F_i is the i^{th} Fibonacci number, defined as follows:

$F_{-2} = 1$, $F_{-1} = 0$; $F_k = F_{k-1} + F_{k-2}$, for $k \geq 0$.

For all other definitions, see [5].

3. Recurrence Relation and Algebraic Expression for $K_n(x_1, x_2, \ldots, x_n)$

It is easy to deduce the K formula for a single linear chain (polyacene) of x_1 hexagons, say $L(x_1)$ (see [5]):

(1) $\quad K_1(x_1) = 1 + x_1.$

We define

(2) $\quad K_0 = 1.$

It may be interpreted as the number of Kekulé structures for "no hexagons."

Theorem 1: If $n \geq 2$, then, for arbitrary $x_1 > 1$, $x_2 > 1$, $\ldots$, $x_n > 1$, the following recurrence relation holds:

(3) $$K_n(x_1, \ldots, x_{n-1}, x_n) = x_n K_{n-1}(x_1, \ldots, x_{n-1} - 1) + K_{n-2}(x_1, \ldots, x_{n-2} - 1).$$

Proof: Let H be the last kink of $L(x_1, x_2, \ldots, x_n)$. We apply the fundamental theorem for matching polynomials [7].

Let u and v be the vertices belonging only to hexagon (kink) H (Figure 2). Consider any perfect matching which contains the bond uv. The rest of such a perfect matching will be a perfect matching of the graph consisting of two components $L(x_n - 1)$ and $L(x_1, x_2, \ldots, x_{n-1} - 1)$. The number of such perfect matchings is

$$K_1(x_n - 1) \cdot K_{n-1}(x_1, x_2, \ldots, x_{n-1} - 1),$$

i.e., according to (1),

(4) $\quad x_n K_{n-1}(x_1, x_2, \ldots, x_{n-1} - 1).$

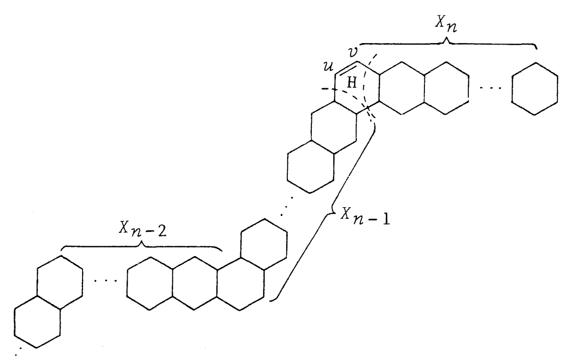


FIGURE 2

On the other hand, each perfect matching without the bond uv must contain all edges indicated in Figure 3. The rest of such a perfect matching will be a perfect matching of $L(x_1, x_2, \ldots, x_{n-2} - 1)$, the number of such perfect matching being

(5) $\quad K_{n-2}(x_1, x_2, \ldots, x_{n-2} - 1).$

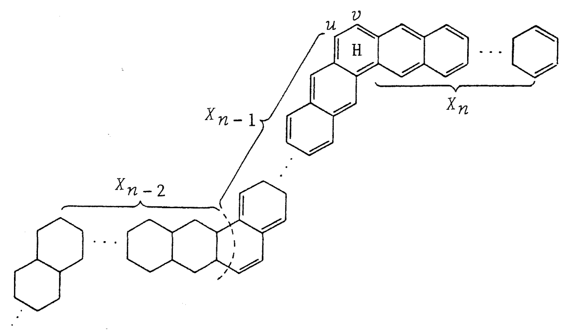


FIGURE 3

From (4) and (5), we obtain recurrence relation (3). □

Obviously, $K_n(x_1, x_2, \ldots, x_n)$ is a polynomial of the form

(6) $$K_n(x_1, \ldots, x_n) = g_n + \sum_{\substack{1 \le \ell_1 < \ell_2 < \cdots < \ell_p \le n \\ 1 \le p \le n}} g_n(\ell_1, \ldots, \ell_p) x_{\ell_1} \cdots x_{\ell_p}.$$

Clearly, $g_0 = 1$.

Now, we are going to determine the coefficients g_n and $g_n(\ell_1, \ldots, \ell_p)$.

First, we define an auxiliary polynomial

(7) $$Q_n(x_1, \ldots, x_{n-1}, x_n) = K_n(x_1, \ldots, x_{n-1}, x_n - 1).$$

For example, we have:

(8) $$Q_0 = 1,\ Q_1(x_1) = x_1,\ Q_2(x_1, x_2) = 1 - x_1 + x_1x_2.$$

From (3) and (7), we obtain the recurrence relation

$$Q_n(x_1, \ldots, x_{n-1}, x_n + 1) = x_nQ_{n-1}(x_1, \ldots, x_{n-1}) + Q_{n-2}(x_1, \ldots, x_{n-2}),$$

i.e.,

(9) $$Q_n(x_1, \ldots, x_{n-1}, x_n) = (x_n - 1)Q_{n-1}(x_1, \ldots, x_{n-1}) + Q_{n-2}(x_1, \ldots, x_{n-2}).$$

Let

(10) $$Q_n(x_1, \ldots, x_n) = S_n + \sum_{\substack{1 \le \ell_1 < \ell_2 < \cdots < \ell_p \le n \\ 1 \le p \le n}} S_n(\ell_1, \ldots, \ell_p) x_{\ell_1} \cdots x_{\ell_p}.$$

Clearly, $S_0 = 1$.

Now, we are going to determine the coefficients $S_n(\ell_1, \ldots, \ell_p)$ and S_n, for $n \ge 1$.

First, we prove the following lemmas.

Lemma 1: $S_n = (-1)^n F_{n-2}$.

Proof: The proof will be by induction on n. According to (8),

$$S_0 = 1 = (-1)^0F_{-2}, \qquad S_1 = 0 = (-1)^1F_{-1}.$$

Suppose that $S_i = (-1)^i F_{i-2}$, for $i \le k$. Then, according to (9),

$$S_k = -S_{k-1} + S_{k-2},$$

and by the induction hypothesis,

$$S_k = -(-1)^{k-1}F_{k-3} + (-1)^{k-2}F_{k-4}$$
$$= (-1)^{k-2}(F_{k-3} + F_{k-4}) = (-1)^k F_{k-2}. \quad \square$$

Lemma 2(a):

(11) $$S_n(\ell_1, \ldots, \ell_{p-1}, \ell_p) = (-1)^{n-\ell_p}F_{n-\ell_p}S_{\ell_p-1}(\ell_1, \ldots, \ell_{p-1}), \text{ for } p > 1.$$

(b):

(12) $$S_n(\ell_1) = (-1)^{n-\ell_1}F_{n-\ell_1}S_{\ell_1-1}.$$

Proof: It suffices to prove (a), since (b) is a particular case of (a). The proof will be by induction on $n - \ell_p$.

If $n - \ell_p = 0$ ($\ell_p = n$), then, according to (9),

(13) $$\begin{aligned} S_n(\ell_1, \ldots, \ell_{p-1}, \ell_p) &= S_{n-1}(\ell_1, \ldots, \ell_{p-1}) \\ &= (-1)^0F_0S_{n-1}(\ell_1, \ldots, \ell_{p-1}) \\ &= (-1)^{n-n}F_{n-n}S_{n-1}(\ell_1, \ldots, \ell_{p-1}). \end{aligned}$$

If $n - \ell_p = 1$ ($\ell_p = n - 1$), then, using (9) and (13), we have:

$$\begin{aligned} S_n(\ell_1, \ldots, \ell_p) &= -S_{n-1}(\ell_1, \ldots, \ell_p) \\ &= -S_{n-2}(\ell_1, \ldots, \ell_{p-1}) = (-1)^1F_1S_{n-2}(\ell_1, \ldots, \ell_{p-1}). \end{aligned}$$

Suppose that (11) is true for $n - \ell_p < k$ ($\ell_p > n - k$), $n - 1 \geq k \geq 2$. Then, for $n - \ell_p = k(\ell_p = n - k)$, according to (9),

$$S_n(\ell_1, \ldots, \ell_p) = -S_{n-1}(\ell_1, \ldots, \ell_p) + S_{n-2}(\ell_1, \ldots, \ell_p),$$

and, by the induction hypothesis,

$$\begin{aligned} S_n(\ell_1, \ldots, \ell_p) &= -(-1)^{n-1-\ell_p}F_{n-1-\ell_p}S_{\ell_p-1}(\ell_1, \ldots, \ell_{p-1}) \\ &\quad + (-1)^{n-2-\ell_p}F_{n-2-\ell_p}S_{\ell_p-1}(\ell_1, \ldots, \ell_{p-1}) \\ &= (-1)^{n-\ell_p}(F_{n-1-\ell_p} + F_{n-2-\ell_p})S_{\ell_p-1}(\ell_1, \ldots, \ell_{p-1}) \\ &= (-1)^{n-\ell_p}F_{n-\ell_p}S_{\ell_p-1}(\ell_1, \ldots, \ell_{p-1}). \quad \square \end{aligned}$$

Lemma 3: $S_n(\ell_1, \ldots, \ell_p) = (-1)^{n-p}F_{n-\ell_p}F_{\ell_p-\ell_{p-1}-1} \cdots F_{\ell_2-\ell_1-1}F_{\ell_1-3}$, for $p \geq 1$.

Proof: For $p = 1$, it follows, from (12) and Lemma 1, that

$$\begin{aligned} S_n(\ell_1) &= (-1)^{n-\ell_1}F_{n-\ell_1}S_{\ell_1-1} = (-1)^{n-\ell_1}F_{n-\ell_1}(-1)^{\ell_1-1}F_{\ell_1-3} \\ &= (-1)^{n-1}F_{n-\ell_1}F_{\ell_1-3}. \end{aligned}$$

For $1 < p \leq n$, according to Lemmas 1 and 2,

$$S_n(\ell_1, \ldots, \ell_{p-1}, \ell_p) = (-1)^{n-\ell_p}F_{n-\ell_p}S_{\ell_p-1}(\ell_1, \ldots, \ell_{p-1}),$$

and now, by induction,

$$\begin{aligned} S_n(\ell_1, \ldots, \ell_{p-1}, \ell_p) &= (-1)^{n-\ell_p}F_{n-\ell_p}(-1)^{\ell_p-1-\ell_{p-1}}F_{\ell_p-\ell_{p-1}-1} \cdots \\ &\quad (-1)^{\ell_2-1-\ell_1}F_{\ell_2-\ell_1-1}(-1)^{\ell_1-1}F_{\ell_1-3} \\ &= (-1)^{n-p}F_{n-\ell_p}F_{\ell_p-\ell_{p-1}-1} \cdots F_{\ell_2-\ell_1-1}F_{\ell_1-3}. \quad \square \end{aligned}$$

Lemma 4(a): $g_n = (-1)^n F_{n-4}$,

(b): $g_n(\ell_1, \ldots, \ell_p) = (-1)^{n-p}F_{n-\ell_p-2}F_{\ell_p-\ell_{p-1}-1} \cdots F_{\ell_2-\ell_1-1}F_{\ell_1-3}$.

Proof: According to (7),

$$Q_n(x_1, \ldots, x_{n-1}, x_n + 1) = K_n(x_1, \ldots, x_{n-1}, x_n).$$

Hence,

$$(14) \qquad g_n(\ell_1, \ldots, \ell_p) = \begin{cases} S_n(\ell_1, \ldots, \ell_p), & \text{if } \ell_p = n, \\ S_n(\ell_1, \ldots, \ell_p) + S_n(\ell_1, \ldots, \ell_p, n), & \text{if } \ell_p < n. \end{cases}$$

Particularly, we have

$$(15) \qquad g_n = S_n + S_n(n), \text{ for } n \geq 1.$$

Now, from (15), Lemma 1, and Lemma 3, we have

$$g_n = (-1)\, F_{n-2} + (-1)^{n-1}F_{n-3} = (-1)^n(F_{n-2} - F_{n-3}) = (-1)^n F_{n-4},$$

and (a) is proved.

To prove (b), observe that, for $\ell_p = n$,

$$(16) \qquad g_n(\ell_1, \ldots, \ell_p) = S_n(\ell_1, \ldots, \ell_p) = (-1)^{n-p}F_{\ell_p - \ell_{p-1} - 1} \cdots F_{\ell_2 - \ell_1 - 1}F_{\ell_1 - 3},$$

and, for $\ell_p < n$,

$$\begin{aligned} g_n(\ell_1, \ldots, \ell_p) &= S_n(\ell_1, \ldots, \ell_p) + S_n(\ell_1, \ldots, \ell_p, n) \\ &= (-1)^{n-p}F_{n-\ell_p}F_{\ell_p - \ell_{p-1} - 1} \cdots F_{\ell_2 - \ell_1 - 1}F_{\ell_1 - 3} \\ &\quad + (-1)^{n-p-1}F_{n-\ell_p - 1}F_{\ell_p - \ell_{p-1} - 1} \cdots F_{\ell_2 - \ell_1 - 1}F_{\ell_1 - 3} \\ &= (-1)^{n-p}(F_{n-\ell_p} - F_{n-\ell_p-1})F_{\ell_p - \ell_{p-1} - 1} \cdots F_{\ell_2 - \ell_1 - 1}F_{\ell_1 - 3}, \end{aligned}$$

i.e.,

$$(17) \qquad g_n(\ell_1, \ldots, \ell_p) = (-1)^{n-p}F_{n-\ell_p-2}F_{\ell_p - \ell_{p-1} - 1} \cdots F_{\ell_2 - \ell_1 - 1}F_{\ell_1 - 3}.$$

Taking into account that, for $\ell_p = n$, $F_{n-\ell_p-2} = F_{-2} = 1$, (16) and (17) can be written together in the form

$$(18) \qquad g_n(\ell_1, \ldots, \ell_p) = (-1)^{n-p}F_{n-\ell_p-2}F_{\ell_p - \ell_{p-1} - 1} \cdots F_{\ell_2 - \ell_1 - 1}F_{\ell_1 - 3}.$$

Theorem 2: $K_n(x_1, \ldots, x_n)$

$$= (-1)^n F_{n-4} + \sum_{\substack{1 \leq \ell_1 < \cdots < \ell_p \leq n \\ 1 \leq p \leq n}} g_n(\ell_1, \ldots, \ell_p)x_{\ell_1} \cdots x_{\ell_p},$$

where $g_n(\ell_1, \ldots, \ell_p)$ is given by (18).

Proof: Follows from Lemma 4. □

Acknowledgments

The authors are grateful to I. Gutman for useful discussions and many valuable comments. They are also thankful to the referee for many helpful suggestions.

References

1. A. T. Balaban, C. Artemi & I. Tomescu. "Algebraic Expressions for Kekulé Structure Counts of Non-Branched Regularly Cata-Condensed Benzenoid Hydrocarbons." *Mathematical Chemistry* *22* (1987):77-100.
2. A. T. Balaban & I. Tomescu. "Algebraic Expressions for the Number of Kekulé Structures of Isoarithmic Cata-Condensed Benzenoid Polycyclic Hydrocarbons." *Mathematical Chemistry* *14* (1983):155-82.
3. A. T. Balaban & I Tomescu. "Chemical Graphs, XL: Three Relations between the Fibonacci Sequence and the Numbers of Kekulé Structures for Non-Branched Cata-Condensed Polycyclic Aromatic Hydrocarbons." *Croatica Chemica Acta* *57.3* (1984):391-404.
4. A. T. Balaban & I Tomescu. "Chemical Graphs, XLI: Numbers of Conjugated Circuits and Kekulé Structures for Zig-Zag Catafusenes and (j, k)-hexes; Generalized Fibonacci Numbers." *Mathematical Chemistry* *17* (1985):91-120.

5. S. J. Cyvin & I. Gutman. *Kekulé Structures in Benzenoid Hydrocarbons*. Berlin: Springer-Verlag, 1988.
6. E. J. Farrell & S. A. Wahid. "Matchings in Benzene Chains." *Discrete Appl. Math.* *7* (1984):31-40.
7. E. J. Farrell. "An Introduction to Matching Polynomials." *J. Comb. Theory* Ser. B *27* (1979):75-86.
8. Z. Fu-ji & C. Rong-si. "A Theorem Concerning Polyhex Graphs." *Mathematical Chemistry* *19* (1986):179-88.
9. M. Gordon & W. H. T. Davison. "Resonance Topology of Fully Aromatic Hydrocarbons." *J. Chem. Phys.* *20* (1952):428-35.
10. I. Gutman. "Topological Properties of Benzenoid Hydrocarbons." *Bull. Soc. Chim. Beograd* *47.9* (1982):453-71.
11. I. Gutman. "Covering Hexagonal Systems with Hexagons." *Proceedings of the Fourth Yugoslav Seminar on Graph Theory*, Novi Sad, 1983, pp. 151-60.
12. I. Gutman & H. Hosoya. "On the Calculation of the Acyclic Polynomial." *Theor. Chim. Acta* *48* (1978):279-86.
13. W. C. Herndon. "Resonance Theory and the Enumeration of Kekulé Structures." *J. Chem. Educ.* *15* (1974):10-15.
14. H. Hosoya & T. Yamaguchi. "Sextet Polynomial: A New Enumeration and Proof Technique for Resonance Theory Applied to the Aromatic Hydrocarbons." *Tetrahedron Letters* (1975):4659-62.
15. L. Lovász & M. D. Plummer. *Matching Theory*. Budapest: Akademiai Kiado, 1986.
16. H. Sachs. "Perfect Matchings in Hexagonal Systems." *Combinatorica 4.1* (1984):89-99.
17. N. Trinajstić. *Chemical Graph Theory*. Vol. 2. Boca Raton, Florida: CRC Press, 1983.

*****