Calculating the Cell Polynomial of Catacondensed Polycyclic Hydrocarbons

Peter E. John

University of Technology of Ilmenau, D-98693 Ilmenau, Germany

Received May 18, 1993®

In the paper we establish a simple algorithm of low complexity for calculating the so-called cell polynomials for a catacondensed polycyclic unsaturated hydrocarbon H, which enables the Kekulé structure count (KSC) and the algebraic structure count (ASC) of H to be calculated.

1. INTRODUCTION

During the past 2 decades applications of graph theory to polycyclic aromatic hydrocarbons have raised great interest in resonance theory. In this respect, Herndon's resonance theory[1,2] and the conjugated circuit model introduced by Randić[3,4] are of considerable importance. The sextet polynomial found by Hosoya and Yamaguchi[5] allows a systematic combinatorial enumeration of Kekulé structures of (catacondensed) aromatic hydrocarbons. This polynomial was shown to possess a number of interesting properties and to reflect Clar's resonant sextet theory.[6] Various further developments of the sextet polynomial concept can be found, e.g., in refs 7–9.

In the present paper we establish a simple algorithm for calculating the so-called cell polynomials for a catacondensed polycyclic unsaturated hydrocarbon H, which enable the Kekulé structure count (KSC) and the algebraic structure count (ASC) of H to be calcuated. In connection with the ASC of H see, e.g., the papers by Wilcox,[10] Herndon,[11] Klein et al.,[12] Dias,[13] and refs 14 and 15.

2. DEFINITIONS AND NOTATION

Let ϵ denote the Euclidean plane and let $\underline{G}$ be the set of all two-connected finite planar graphs (without loops and multiple edges). Let $G = (V,E) \in \underline{G}$ with vertex set $V = V(G)$ and edge set $E = E(G)$, and let $n = n(G) = |V|$, $m = m(G) = |E|$ denote the numbers of vertices and edges of G, respectively. Graph G_ϵ is an embedding of G into ϵ. G_ϵ subdivides ϵ in $m - n + 1 =: c = c(G)$ finite open domains D_i, $i = 1, 2, ..., c$, the infinite open domain D_0 and the union of the boundaries of all these domains (Figure 1).

The *boundaries* of the D_i are denoted by $B_i = B(D_i)$, $i = 1, 2, ..., c$; B_0 is also called the *contour* (or periphery) of G_ϵ. G_ϵ is called a *map* denoted by $M = M(G)$. Cell C_i is the union $D_i \cup B_i$, $i = 1, 2, ..., c$. Note that $B(C_i) =: B_i$. The *inner dual* $D = D(M)$ of M is the dual of M without the vertex that corresponds to the infinite domain. Map M is a *contour map* if $D(M)$ is a tree (Figure 1).

A *perfect matching* (PM) or linear factor of a graph G is a set of pairwise disjoint edges that cover all vertices of G (Figure 5).

Case $c(M) = 0$. Then $M = \emptyset$ is the zero map (without edges and vertices). By convention, map $\emptyset$ has precisely one PM.

Case $c(M) = 1$. Then M has exactly one cell which is called an *isolated cell.*

Case $c(M) > 1$. Edge $e \in E(M)$ is called a *contour edge* or an *internal edge* of M if the whole edge, or only its endpoints, lie on B_0 (Figure 1). Cell C is an end cell of M if C corresponds to an end vertex of $D(M)$ (in particular, an isolated cell is an end cell).

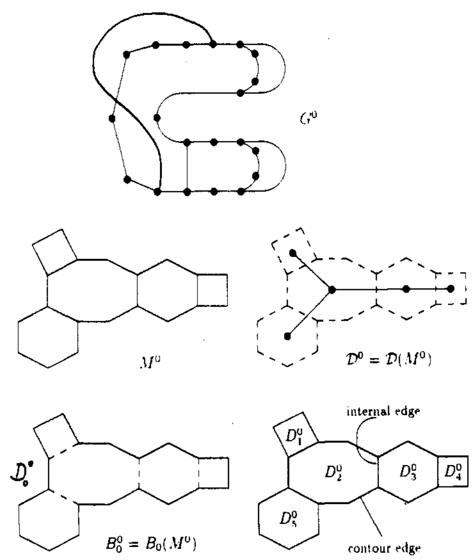


Figure 1.

Let $\underline{M}$ denote the set of all contour maps.

Map $M \in \underline{M}$ is *bipartite* if and only if $V = V(M)$ can be divided into two disjoint subsets $\tilde{V} = \tilde{V}(M)$ and $\tilde{\tilde{V}} = \tilde{\tilde{V}}(M)$ ($V = \tilde{V} \cup \tilde{\tilde{V}}$, $\tilde{V} \cap \tilde{\tilde{V}} = \emptyset$) such that vertices from the same subset are never adjacent.

Let $\underline{M}^b$ denote the set of all bipartite contourmaps.

Note that for every $M \in \underline{M}^b$ $\tilde{n} = \tilde{\tilde{n}}$ and M has a PM.

Observation 1. Every $M \in \underline{M}^b$ with $c(M) > 1$ has at least two end cells.

Observation 2. Every $M \in \underline{M}^b$ is the last of a (finite) sequence of bipartite contour maps $\{M_i\}$, $i = 1, 2, ..., c$, where M_1 is an isolated cell and for $i = 2, 3, ..., c$ we obtain M_i by adding cell C_i (C_i is an end cell of M_i) to M_{i-1} (see Figure 2).

Any plane image (i.e., an embedding in ϵ) of a (bipartite) planar graph G is called a (bipartite) *pattern* (of G). The cells of a pattern are defined in an analogous way as above.

Let $\underline{\mathcal{M}}^b$ denote the set of all (bipartite) subpatterns of all maps from $\underline{M}^b$.

3. DEFINITION OF THE CELL POLYNOMIAL OF A BIPARTITE PATTERN

For $\mathcal{M} \in \underline{\mathcal{M}}^b$ let $\underline{C} = \underline{C}(\mathcal{M}) = \{C_{i_1}, C_{i_2}, ..., C_{i_c}\}$ and $N = N(\mathcal{M}) = \{i_1, i_2, ..., i_c\}$ denote the set of all cells and of all cell indices (labels), respectively (note that if $\underline{C} = \emptyset$, then $N = \emptyset$).

® Abstract published in *Advance ACS Abstracts,* February 15, 1994.

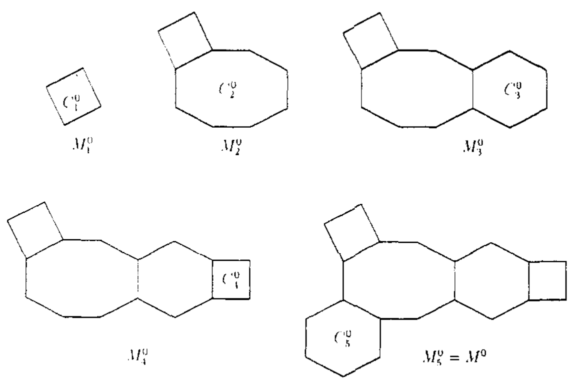
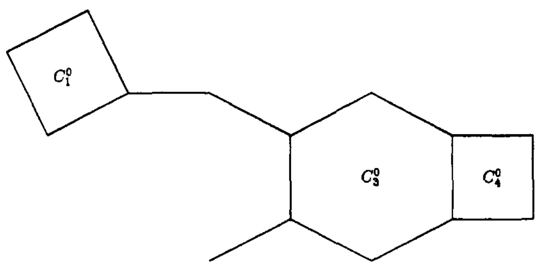


Figure 2.



$\mathcal{M}^* = M^0 - V(C_5^0)$

Figure 3.

To every subset $I \subseteq N$ there corresponds a cell set $\underline{C}_I = \{C_i | C_i \in \underline{C}$ and $i \in I\}$ of $\mathcal{M}$; I is called an *index set* of $\mathcal{M}$.

The pattern $\mathcal{M}_I$ is obtained from $\mathcal{M}$ by deleting in $\mathcal{M}$ all vertices of all cells from $\underline{C}_I$ and all edges which are incident with these vertices. Index set $I \subseteq N$ is *feasible* (in $\mathcal{M}$) if (i) any two cells $C', C'' \in \underline{C}_I$ are disjoint ($B(C') \cap B(C'') = \emptyset$) and (ii) $\mathcal{M}_I$ has a PM. Note that $\mathcal{M}_\emptyset = \mathcal{M}$; this implies that $I = \emptyset$ is feasible if and only if $\mathcal{M}$ has a PM.

To every cell $C_i \in \underline{C}(\mathcal{M})$ is assigned a weight $w_i = w(C_i)$ which is an element of an (algebraic) ring (here it suffices to assume that w_i is a real number). Cell set $\underline{C}_I$ of $\mathcal{M}$ *has weight*

$$w(\underline{C}_I) = \begin{cases} \prod_{i \in I} w_i, & \text{if } I \text{ is feasible} \\ 0 & \text{otherwise} \end{cases}$$

(this includes $w(\underline{C}_\phi) = 1$ if $\emptyset$ is feasbile). The *cell polynomial* $f_\mathcal{M} = f_\mathcal{M}(w_1, w_2, ..., w_c)$ of $\mathcal{M}$ is defined as

$$f_\mathcal{M} = \sum_{I \subset N(\mathcal{M})} w(\underline{C}_I)$$

For example, consider the pattern $\mathcal{M}^* =: M^0 - V(C_5^0)$, given in Figure 3.

Here $N(\mathcal{M}^*) = \{1,3,4\}$; the only feasible sets are $\emptyset$ and $\{1\}$; thus $f_{\mathcal{M}^*} = 1 + w_1$.

Note that for any bipartite pattern $\mathcal{M}$ the cell polynomial $f_\mathcal{M}$ can be defined as described above.

4. CALCULATION OF THE CELL POLYNOMIAL OF A BIPARTITE PATTERN WHICH IS PART OF A BIPARTITE CONTOUR MAP

Let $\mathcal{M} \in \underline{\mathcal{M}}^b$ with cell set $\underline{C}$, where cell $C_i \in \underline{C}$ has weight $w_i = w(C_i)$. The cell polynomial $f_\mathcal{M}$ can be calculated applying the following recursive procedure.

Algorithm f:

(f.0) If $\mathcal{M}$ consists of a single (isolated) vertex, put $f_\mathcal{M} = 0$;

(f.1) $f_\emptyset = 1$;

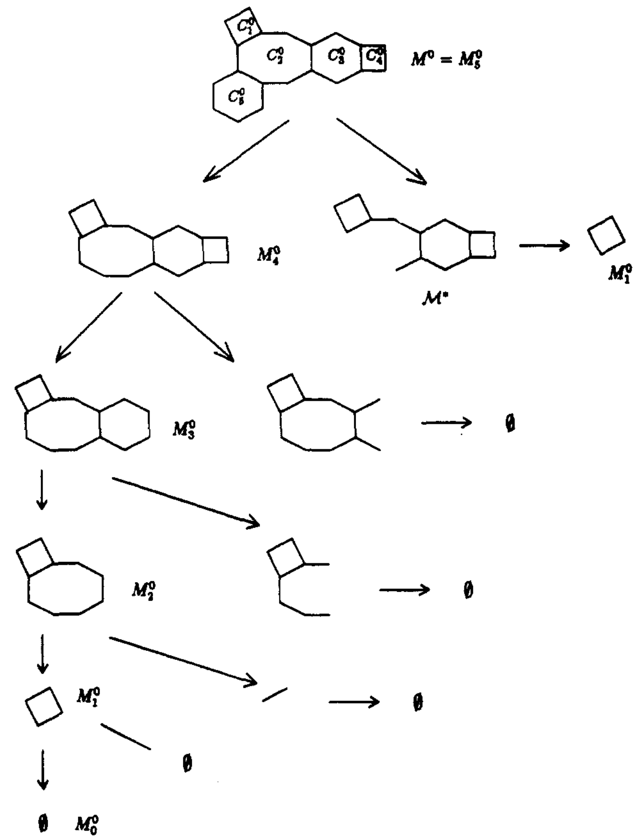


Figure 4.

(f.2) If $\mathcal{M}$ has components $\mathcal{M}_1, \mathcal{M}_2, ..., \mathcal{M}_p$, put $f_\mathcal{M} = f_{\mathcal{M}_1} f_{\mathcal{M}_2} \cdots f_{\mathcal{M}_p}$;

(f.3) If $\mathcal{M} \in \underline{\mathcal{M}}^b$, $\mathcal{M}$ is connected, and $c(\mathcal{M}) > 1$, consider the following cases.

(i) $\mathcal{M}$ has no cut vertex ($\mathcal{M} \in \underline{M}^b$).

Then (because of observation 2) $\mathcal{M}$ can be obtained from $\mathcal{M}_c \in \underline{M}^b$ by adding cell C to $\mathcal{M}_c$ (C is an end cell of $\mathcal{M}$).

Delete in $\mathcal{M}$ all vertices of $V(C)$ (and all edges which are incident with these vertices). Denote the resulting pattern by $\mathcal{M}^c$. Put

$$f_\mathcal{M} = f_{\mathcal{M}_c} + w(C) f_{\mathcal{M}^c}$$

(ii) $\mathcal{M}$ has a cut vertex ($\mathcal{M} \in \underline{M}^b$).

(ii.1) $\mathcal{M}$ has a hanging edge $(u, v) \in E(\mathcal{M})$.

Delete vertices u, v and all edges which are incident with them. Denote the resulting pattern by $\mathcal{M}'$. Put

$$f_\mathcal{M} = f_{\mathcal{M}'}$$

(ii.2) $\mathcal{M}$ has no hanging edge.

Then there exists an induced subpattern $\mathcal{M}''$ of $\mathcal{M}$ such that $\mathcal{M}'' \in \underline{M}^b$, $\mathcal{M}''$ contains exactly one cut vertex of $\mathcal{M}$, and $\mathcal{M}''$ is contained in no other submap of $\mathcal{M}$ with these properties. Delete in $\mathcal{M}$ all vertices of $\mathcal{M}''$ (and all edges incident with them); this results in a pattern which we call $\mathcal{M}'''$. Put

$$f_\mathcal{M} = f_{\mathcal{M}''} f_{\mathcal{M}'''}$$

By successively applying f.0–f.3, eventually the polynomial $f_\mathcal{M}$ of $\mathcal{M} \in \underline{\mathcal{M}}^b$ is found. For example consider the map $M^0 \in \underline{M}^b$ of Figure 2. M^0 has five cells C_i^0 with weights $w_i^0 = w(C_i^0)$, $i = 1, 2, 3, 4, 5$. Here we find (see Figure 4)

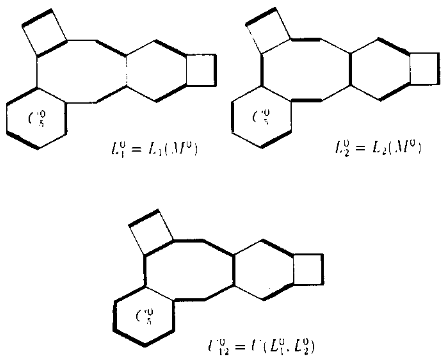


Figure 5.

$$f_{M^0} = f_{M_5^0} = f_{M_4^0} + w_5^0 f_{\mathcal{M}^*} = \ldots = f_{M_4^0} + w_5^0 f_{M_1^0}$$

and for i = 3, 2, 1, 0 by use of (f.3(ii))

$$f_{M_{i+1}^0} = f_{M_i^0} + w_{i+1}^0 f_\varnothing$$

Therefore,

$$f_{M^0} = 1 + w_1^0 + w_2^0 + w_3^0 + w_4^0 + w_5^0 + w_1^0 w_5^0$$

It can easily by checked that this result is in accordance with the definition of the cell polynomial (section 3).

5. COARSENED CELL POLYNOMIALS

The cell polynomial $f_{\mathcal{M}} = f_{\mathcal{M}}(w_1, w_2, \ldots, w_c)$ of $\mathcal{M} \in \underline{\mathcal{M}}^b$ can be "coarsened". Let $n_i = n(C_i)$ denote the number of vertices of $B_i = B(C_i)$. Inserting

$$w_i = \begin{cases} x, & \text{if } n_i \equiv 2, \bmod 4 \\ y, & \text{if } n_i \equiv 0, \bmod 4 \end{cases}$$

into $f_{\mathcal{M}}$, we find the *first coarsened cell polynomial* $f^*_{\mathcal{M}} = f^*_{\mathcal{M}}(x,y)$. With $x = y =: z$ we obtain the *second coarsened cell polynomial* $f^{**}_{\mathcal{M}} = f^{**}_{\mathcal{M}}(z) =: f^*_{\mathcal{M}}(z,z)$. As an example we use the contour map M^0 of Figure 2:

$$f^*_{\mathcal{M}^0} = 1 + 2x + 3y + xy$$

$$f^{**}_{M^0} = 1 + 5z + z^2$$

If there is no danger of confusion, we shall directly transfer concepts and symbols, originally defined for planar graphs, to their plane images (embeddings, patterns).

Let $\underline{L} = \underline{L}(G)$ and $l = l(G)$ denote the set and the number of all PMs of G, respectively.

Observation 3: For every $\mathcal{M} \in \underline{\mathcal{M}}^b$

$$f^{**}_{\mathcal{M}} = l(\mathcal{M})$$

This formula follows immediately from the way the algorithm is constructed; it is well-known for $\mathcal{M}$ being a catacondensed hexagonal system.[5]

A *perfect basic figure* (PBF) of graph G is a subgraph U of G with the following properities: (i) Every component of U is a circuit or a dumbbell ●–●; (ii) U covers all vertices of G (Figure 5).

Clearly, the union of any two PMs $L^*, L^{*|} \in \underline{L}(G$ of G is the edge set of a PBF; conversely, every PBF U can be obtained this way. We shall denote the PBF U determined by the pair $[L^*, L^{*|}]$ briefly by $(L^*, L^{*|})$. Let $q(U) = q(L^*, L^{*|})$ denote the number of circuits contained in $U = (L^*, L^{*|})$ whose length is a multiple of 4.

Lemma: There is a unique partition $\underline{L}^{|}, \underline{L}^{||}$ of $\underline{L}$ (*i.e.*, $\underline{L} = \underline{L}^{|} \cup \underline{L}^{||}, \underline{L}^{|} \cap \underline{L}^{||} = \varnothing$ where $\underline{L}^{|}$ or $\underline{L}^{||}$ may be empty) such that

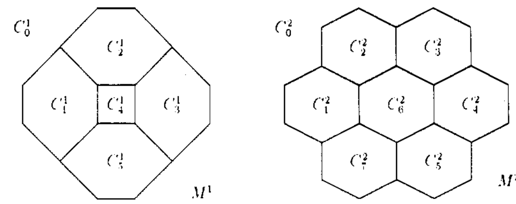


Figure 6.

$q(\underline{L}^{|}, \underline{L}^{||})$ is even (odd) if and only if $\underline{L}^{|}$ and $\underline{L}^{||}$ are in the same class (in different classes) of the partition. Note that only the partition is determined (not the individual classes); the classes are interchangeable.

Put $l^{|} = l^{|}(G) = |\underline{L}^{|}|$ and $l^{||} = l^{||}(G) = |\underline{L}^{||}|$.

Proposition: For every $\mathcal{M} \in \underline{\mathcal{M}}^b$

$$|f^*_{\mathcal{M}}(1,-1)| = |l^{||}(\mathcal{M}) - l^{|}(\mathcal{M})|$$

This proposition can be proved in a way analogous to the proof of Theorem 6 in ref 15.

As an example we use the bipartite contour map M^0 of Figure 1

$$f^*_{\mathcal{M}^0}(1) = l(M^0) = 7$$

$$|f^{**}_{M^0}(1,-1)| = 1$$

Indeed, by inspection we easily obtain

$$l^{|}(M^0) = 4, \quad l^{||}(M^0) = 3 \quad (\text{or } l^{|}(M^0) = 3, \quad l^{||}(M^0) = 4)$$

6. CONCLUDING REMARK

It is worth mentioning that algorithm f can, in a suitably extended form, be applied to any bipartite pattern. For example, consider M^1 and M^2 given in Figure 6; we have

$$f^*_{M^1} = f^*_{M^1}(x,y) = 1 + 4x + y + 2x^2$$

$$f^{**}_{M^1} = f^{**}_{M^1}(z) = 1 + 5z + 2z^2$$

$$f^*_{M^2} = f^*_{M^2}(x,y) = 1 + 7x + 9x^2 + 2x^3 = f^{**}_{M^2}(x)$$

The significance of this extension will be discussed in a subsequent paper.

REFERENCES AND NOTES

(1) Herndon, W. C. Resonance energy of aromatic hydrocarbons. Quantitative test of resonance theory. *J. Am. Chem. Soc.* **1973**, *95*, 2404.
(2) Herndon, W. C.; Ellzey, M. L. Resonance Theory. V. Resonance energies of benzenoid and nonbenzenoid π systems. *J. Am. Chem. Soc.* **1974**, *96*, 6631.
(3) Randić, M. Conjugated circuits and resonance energy of benzenoid hydrocarbons. *Chem. Phys. Lett.* **1976**, *38*, 68.
(4) Randić, M. A graph theoretical approach to conjugation and resonance energies of hydrocarbons. *Tetrahedron* **1977**, *33*, 1906.
(5) Hosoya, H.; Yamaguchi, T. Sextet polynomial. A new enumeration and proof technique for the resonance theory applied to the aromatic hydrocarbons. *Tetrahedron Lett.* **1975**, 4659.
(6) Clar, E. *The Aromatic Sextet*; Wiley: London, 1972.
(7) Aihara, J. Aromatic sextets and aromaticity in benzenoid hydrocarbons. *Bull. Chem. Soc. Jpn.* **1977**, *50*, 2010.
(8) Gutman, I. A method for calculation of resonance energy of benzenoid hydrocarbons. *Z. Naturforsch.* **1978**, *33a*, 840.
(9) Gutman, I.; Cyvin, S. J. *Introduction to the Theory of Benzenoid Hydrocarbons*; Springer: Berlin, 1989.
(10) Wilcox, C. F. Stability of molecules containing (4n)-rings. *Tetrahedron Lett.* **1968**, 795.
(11) Herndon, W. C. Enumeration of resonance structures. *Tetrahedron* **1973**, *29*, 3.

(12) Klein, D. J.; Schmalz, T. G.; El-Basil, S.; Randić, M.; Trinajstić, N. Kekulé count and algebraic structure count for unbranched alternant catafusenes. *J. Mol. Struct.* (*THEOCHEM*) **1988**, *179*, 99.
(13) Dias, J. R. Note on algebraic structure count. *Z. Naturforsch.* **1989**, *44A*, 761.
(14) John, P. E. Comment on the paper: Kekulé count and algebraic structure count for unbranched alternant cata-fusenes (*J. Mol. Struct.* (*THEOCHEM*) **1988**, *179*, 99).
(15) John, P. E. Linearfaktoren in defekten hexagonalen Systemen. *Bayreuther Math. Schriften*, **1993**, *43*, 35.