Graph Theoretical Means for Calculating Kekulé and Hückel Parameters in Benzenoid and Related Systems

Horst Sachs†

Institut für Mathematik, Technische Universität Ilmenau, D-98684 Ilmenau, Germany

Received May 18, 1993*

In 1990, Kh. Al-Khnaifes, P. John, and the author described some simple algorithms for calculating certain structural parameters (the number of Kekulé patterns, Pauling's bond orders, the number of spanning trees, the characteristic polynominal, the orbitals) of a benzenoid system. Such a system is a section of the hexagonal tessellation (or lattice) of the plane. In this paper, light is shed on the algebraic-combinatorial background. An infinite class of tessellations of the plane is defined, and it is pointed out that the same methods can be applied to sections of any tessellation of this class. Applications arise, e.g., in crystal physics in connection with the general dimer problem. Some remarks are made, and some first results are given, on (finite) hexagonal tessellations of the torus.

1. INTRODUCTION

The structure of a benzenoid system (molecule) **S** is uniquely described by the adjacency matrix (Hückel matrix) $\boldsymbol{M}$ of the graph **G** corresponding to **S**, and some important structural parameters of **S** (the number of Kekulé patterns, Pauling's bond orders, the number of spanning trees, the characteristic polynomial, the eigenspaces) are in fact "linear" invariants of **G**, that is to say, they are functions of the matrix $\boldsymbol{M}$, or of some matrices closely related to $\boldsymbol{M}$, whose values can be obtained from these matrices via linear algebra (calculating determinants and/or solving systems of linear equations). However, this is a detour which we would like to avoid, or to reduce to a minimum, by directly utilizing the structural peculiarities of the system **S**. This can indeed be done: in graph theory, general methods have been developed which allow problems of linear algebra to be solved by more or less efficient algorithms making use of the combinatorial structure of the problem (the relative position of the zeros in a matrix, etc.). These algorithms become particularly efficient (and intuitive, quassi "inspective") if used for calculating the above-mentioned structural parameters of chemical compounds.

The applicability of these graph algorithms is not at all restricted to benzenoid systems whose graphs are sections of the hexagonal tessellation (the graphite pattern) of the plane: they can quite as well be applied to sections of many other lattices and related structures. Thus they are useful not only in benzenoid chemistry.

In crystal physics applications arise in connection with the dimer problem which requires the number of coverings (without gaps) of the surface of a crystal (which is considered part of a regular lattice) by two-atomic molecules (dimers) to be determined.[4,5,11]

There are more applications in the theory of linear electrical networks (Kirchhoff's theory), in the numerical theory of boundary value problems, and in other areas, but these will not be considered in this paper.

For specific references the reader is referred to refs 1 and 3.

2. ALGEBRAIC-COMBINATORIAL BACKGROUND

Let $\boldsymbol{A} = (a_{ik})$ be a matrix and $\mathbf{b} = (b_i)$ a column vector ($i = 1, 2, ..., m$; $k = 1, 2, ..., n$) where a_{ik} and b_i are real numbers

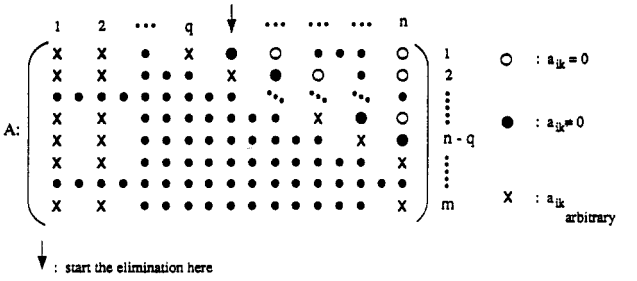


Figure 1.

and no row and no column of $\boldsymbol{A}$ is the zero vector. We are interested in the set of solutions to the system of linear equations

$$\boldsymbol{A}\mathbf{x} = \mathbf{b} \tag{1}$$

and, if $m = n$, in the value of det $\boldsymbol{A}$.

If some of the a_{ik} are equal to zero, we may try to permute the rows and columns of $\boldsymbol{A}$ in such a way that some of the zeros cover the right upper corner which is separated from the rest of the matrix by a nowhere-zero diagonal (Figure 1): then, by straightforward elimination, the number of unknowns and the order of det $\boldsymbol{A}$ (if $m = n$) can be reduced from n to q, as indicated in Figure 1.

This simple observation can be used to construct an algorithm which, under certain conditions, proves very efficient. The essential point being the combinatorial structure of $\boldsymbol{A}$, it is convenient to formulate the problem in graph-theoretical (label-free) terms by making a (weighted) bipartite graph $\mathbf{G}^*$ correspond to $\boldsymbol{A}$.

The resulting algorithms—algorithm $\mathcal{A}$ for calculating the value of a determinant and algorithm $\mathcal{B}$, closely related to algorithm $\mathcal{A}$, for solving a system of linear equations—are described and their effectiveness is proved in detail in ref 1 (see Theorems 1 and 2): this shall not be repeated here. The crucial point is this: A *matching* **M** in a graph **G** is a set of pairwise disjoint edges of **G**. (For later use: **M** is called *perfect* if it covers all vertices of **G**.) The edges of **M** and the remaining edges of **G** are usually referred to as the *red* and *blue edges* of **G**, respectively. **M** is called *acyclic* if there is no red–blue alternating circuit in **G**. Note that in Figure 1 we have a nowhere-zero diagonal of length $n - q$ (marked by bold dots). This diagonal corresponds to an acyclic matching in the graph **G** representing the zero–nonzero structure of the matrix $\boldsymbol{A}$ (**G**

† Current address (until May 31, 1994): GERAD—École des Hautes Études Commerciales, 5255 Avenue Decelles, Montréal, Québec, Canada H3T 1V6.

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

0095-2338/94/1634-0432$04.50/0

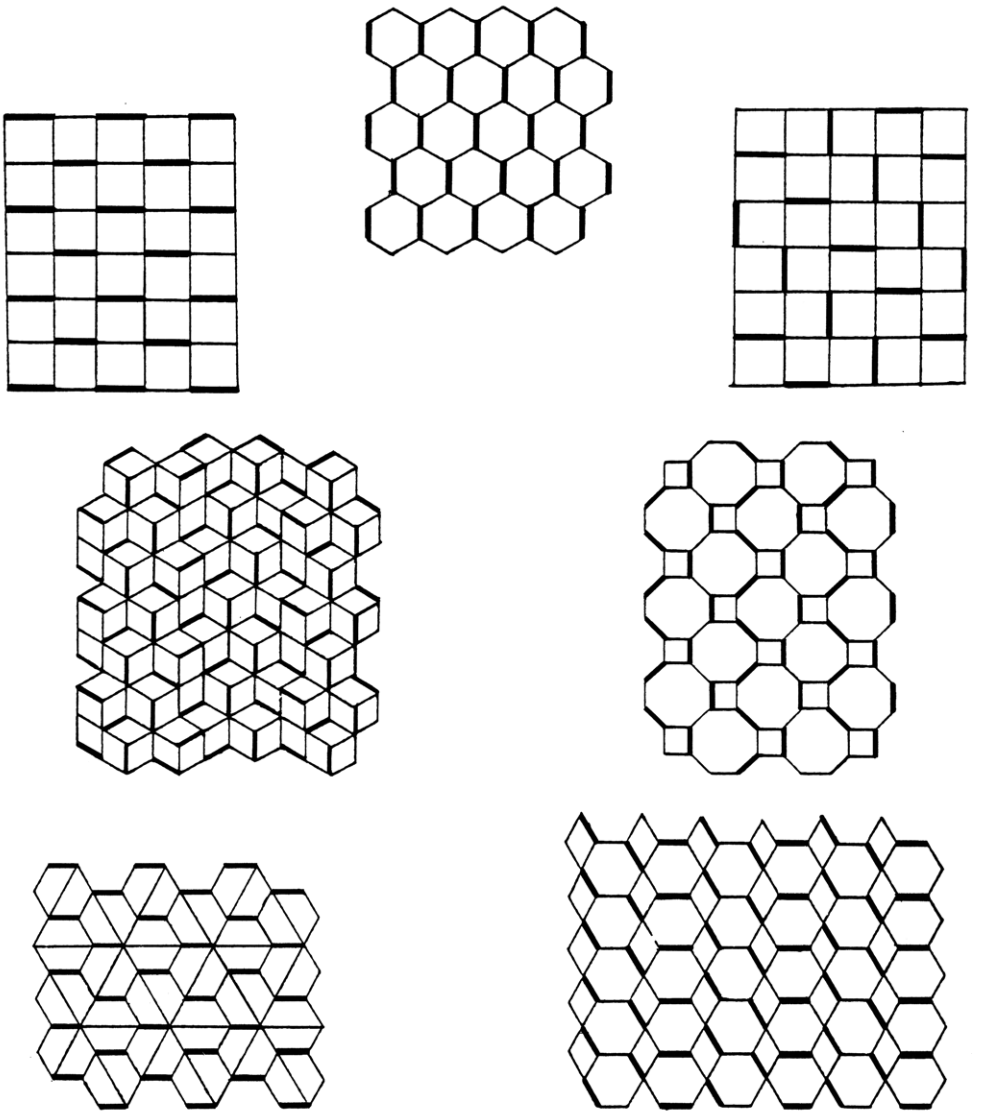
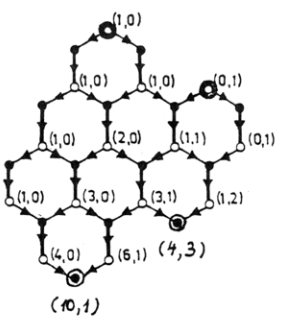
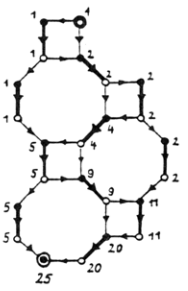


Figure 2.





$$m(\mathrm{G}) = \begin{vmatrix} 10 & 1 \\ 4 & 3 \end{vmatrix} = 26 \qquad m(\mathrm{G}) = 25$$

Figure 3.

is the unweighted version of **G***). Therefore, in order to be able to start the elimination process—i.e., to apply algorithm $\mathcal{A}$ or $\mathcal{B}$—we must first find an acyclic matching. If **M** is such a matching of size $|\mathbf{M}| = p$, then the reduction of the size of the problem (number of columns) is from n to $q = n - p$: therefore, we are interested in having a "large" acyclic matching at our disposal. Unfortunately, the problem of finding a maximum acyclic matching is known to be NP-hard:[7,8] therefore, in the general case algorithms $\mathcal{A}$ and $\mathcal{B}$ are only of restricted value. However—and that is important—in those cases we have in mind, a (maximal) acyclic matching

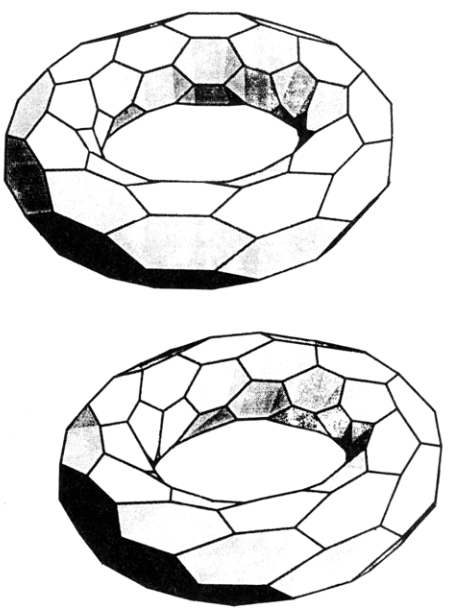


Figure 4.

of satisfactory size can easily be constructed, or even seen by mere inspection, since such a matching is given immediately with the (combinatorial structure of the) problem: e.g., in a benzenoid system, if drawn in the usual way, the perpendicular

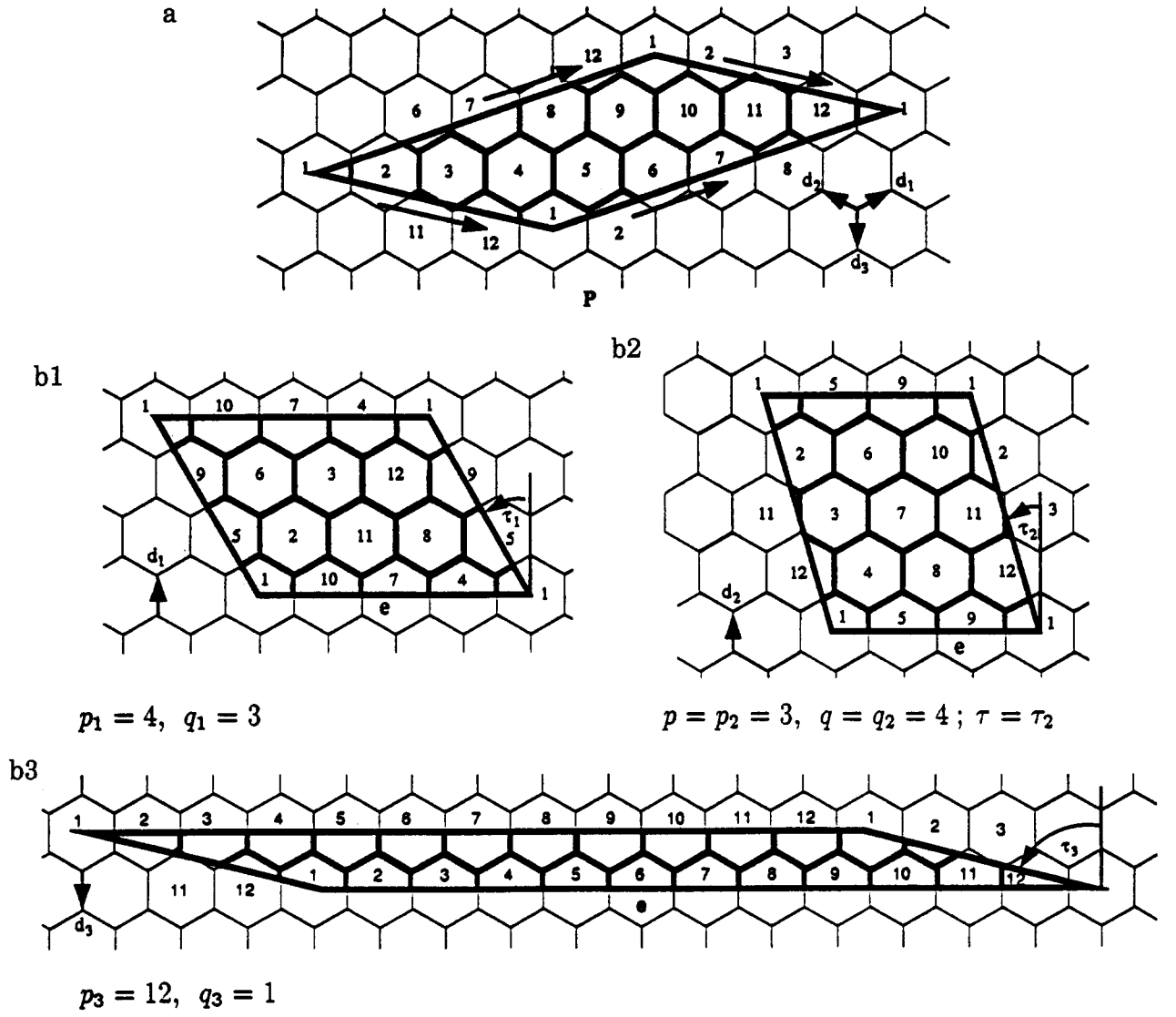


Figure 5.

edges form an acyclic matching (the reader may convince himself of that and why this is true).

3. APPLICATION TO BENZENOID SYSTEMS

About 1950, Dewar and Longuet-Higgins[2,6] established a formula which allows the number of Kekulé patterns (=perfecting matchings) of a benzenoid system **S** to be expressed as the determinant of the "structure matrix" of **S**. This is a special case of a general formula, valid for arbitrary plane graphs, found in 1961 by Kasteleyn;[4,5] for more details and references, see ref 1.

In connection with the Dewar/Longuet-Higgins formula, algorithm $\mathcal{A}$ can be so specialized as to yield an efficient algorithm for calculating the number of Kekulé patterns and Pauling's bond orders of benzenoid systems; furthermore, algorithms $\mathcal{A}$ and $\mathcal{B}$ can be used to calculate the characteristic polynomial and the eigenspaces (orbitals) of a benzenoid system. These methods—and others—being described in detail in ref 3, they shall not be reproduced here.

4. COUNTING PERFECT MATCHINGS IN SECTIONS OF BIPARTITE TESSELLATIONS

Though the methods developed can be applied to a much wider class of tessellations, we shall restrict our considerations to the class $\mathcal{P}$ of bipartite tessellations of the plane which are doubly periodic and have a perfect matching.

Let **G** be a bipartite graph with bipartition **B** = **B**(**G**), **W** = **W**(**G**). The elements of **B** and **W** are referred to as the *black* and *white vertices* of **G**, respectively.

Let $\mathbf{P} \in \mathcal{P}$ and **C** a circuit in **P**. The plane subgraph of **G** of **P** spanned (induced) by the vertices lying on **C** and in the interior of **C** is called a *section* of **P**; the circuit **C** is its *boundary*. **G** is *balanced* if $|\mathbf{B}(\mathbf{G})| = |\mathbf{W}(\mathbf{G})|$.

If F is a face of **P**, let $\mathbf{I}(F)$ denote the set of edges incident with F (i.e., forming the boundary of F) and put $|\mathbf{I}(F)| = 2h(F)$. A perfect matching **M** of **P** is called *distinguished* if, for every face F of **P**, $|\mathbf{M} \cap \mathbf{I}(F)| \equiv h(F) - 1$, mod 2.

It is not difficult to prove that a distinguished matching is acyclic. Figure 2 shows some tessellations of class $\mathcal{P}$, each equipped with a distinguished matching (the bold lines). In all these special cases, $|\mathbf{M} \cap \mathbf{I}(F)| = h(F) - 1$.

Remark. There are tessellations in $\mathcal{P}$ that do not have a distinguished matching. It is an open, challenging problem to determine under what conditions a tessellation $P \in \mathcal{P}$ has a distinguished matching.

Consider a tessellation $\mathbf{P} \in \mathcal{P}$ which has a distinguished matching **D** and let **G** be a balanced section of **P**. We are interested in the number $m(\mathbf{G})$ of perfect matchings (representing, e.g., Kekulé patterns or dimer coverings) of **G**.

Let $\mathbf{D}_G$ be the set of those edges of **G** which belong to **D**. Clearly, $\mathbf{D}_G$ is an acyclic matching of **G**. Direct all edges of $\mathbf{D}_G$ (the red edges) from their black end point toward their white end point and all other edges of **G** (the blue edges) from

their white end point toward their black end point. Thus a directed graph $\vec{G}$ is created which has the following properties: (i) $\vec{G}$ is acyclic; (ii) $\vec{G}$ has as many sources as sinks (at least one) which all lie on the boundary C of G; (iii) all sources are white and all sinks are black.

Now apply the following algorithm to $\vec{G}$.

Algorithm. If u is an arc of $\vec{G}$, denote by $s(u)$ and $t(u)$ the vertex from which it springs and in which it terminates, respectively. Let $\{a_1, a_2, ..., a_r\}$ and $\{b_1, b_2, ..., b_r\}$ be the sets of sources and sinks of G. Let

$$\mathbf{p}(a_k) = (\delta_{k1}, \delta_{k2}, ..., \delta_{kr}), \quad k = 1, 2, ..., r$$

where $\delta_{ii} = 1$, $\delta_{ij} = 0$ if $i \neq j$.

For each vertex x of G which is not a source, define the vector $\mathbf{p}(x) = (p_1(x), p_2(x), ..., p_r(x))$ recursively by the formula

$$\mathbf{p}(x) = \sum_{t(u)=x} \mathbf{p}(s(u))$$

Clearly, $p_k(x)$ is the number of paths issuing from source a_k and terminating in x.

The vectors $\mathbf{p}(x)$ can easily be determined following step by step the paths issuing from the sources (see Figure 3).

Put $p_{ik} := p_k(b_i)$ and $P = (p_{ik})_{i,k=1,2,...,r}$

Theorem 1: $m(G) = |\det P|$.

For some examples, see Figure 3; for the proof see refs 9 and 10.

By this theorem, the order of the determinant eventually to be calculated is reduced from half the numbers of vertices of section G to less than half the number of vertices on the boundary C of G.

5. APPLICATION TO HEXAGONAL TESSELLATIONS ON THE TORUS

After C_{60} has been found, it is a natural question to ask whether there are torus-shaped "graphitoid" molecules. Note that the torus is the only closed two-sided surface on which (finite) graphs can be realized such that all their vertices have degree 3 and all their faces are hexagons (see Figure 4): the interest in the combinatorial properties of these toroidal hexagonal tessellations is independent of their possible relevance in chemistry.

Denote the class of all toroidal hexagonal tessellations by $\mathcal{T}$.

The graphs of the members of $\mathcal{T}$ can conveniently be drawn in the plane (equipped with the regular hexagonal lattice) using the representation of the torus by a parallelogram P with the usual boundary identification, as indicated in Figure 5a. This representation can be made unique in the following way:

Let d_i ($i = 1, 2, 3$) denote one of the three directions of the edges of the hexagonal lattice (Figure 5a). Select P such that one of its edges, e, is perpendicular to d_i; in addition, suppose that the angle τ_i (see Figure 5b) has been chosen nonnegative and minimum. Let p_i and q_i denote the number of hexagons met by e and the number of strings (layers) of hexagons parallel to e which are covered by P, respectively. Thus for every $T \in \mathcal{T}$ three (not necessarily distinct) representations of its graph are obtained. Let $p_{i_0} = \min\{p_1, p_2, p_3\}$ and put $p = p_{i_0}$, $q = q_{i_0}$, $\tau = \tau_{i_0}$ (Figure 5b2): then p, q, and τ—which, in a way, measure the "width", the "length", and the (reduced) "twist" (torsion) of T—may be considered the canonical parameters of the graph of T.

Clearly, T has pq faces, $2pq$ vertices, and $3pq$ edges. The (degenerate) case $p = 3$ is also of some interest since, in this case, there are— in addition to the faces—q plane hexagons whose boundaries surround the torus passing through its hole.

Also in this nonplanar case, the above methods can be applied, though not immediately, to calculate the number of Kekulé patterns and Pauling's bond orders. The preparations needed being a bit tricky, they cannot be explicated here; however, the algorithm is as efficient as in the planar case: the determinants eventually to be calculated are of order p.

For the simplest cases $p = 3, 4$, John (Ilmenau) and the author (who have just started investigating toroidal hexagonal tessellations) obtained the following first results.

(1) $p = 3$. Then the number of Kekulé patterns is

$$m_q = \begin{cases} 2^{q+1} + 3^q + 3 & \text{if } q \text{ is even} \\ 2^{q+1} + 3^q + 2 & \text{if } q \text{ is odd} \end{cases}$$

The bond orders are very close to $^1/_3$: for the three edges incident with a vertex they are

$$\frac{1}{3} + \frac{4}{3m_q}, \ \frac{1}{3} - \frac{2}{3m_q}, \ \frac{1}{3} - \frac{2}{3m_q} \quad \text{if } q \text{ is even}$$

$$\frac{1}{3} + \frac{1}{m_q}, \ \frac{1}{3} - \frac{2}{m_q}, \ \frac{1}{3} + \frac{1}{m_q} \quad \text{if } q \text{ is odd}$$

(2) $p = 4$; $q = 2r$ even; $\tau = 0$. The number of Kekulé patterns is

$$(2^{2r} + 1)^2 + 2^r[(\sqrt{2} + 1)^r + (\sqrt{2} - 1)^r]^2 + 2^{r+2} > 4^q$$

Remark. Note that these investigations apply to toroidal *hexagonal* tessellations only: toroidal carbons which are allowed to have—beside hexagons—also pentagons and heptagons are certainly of great interest; however, the problems connected with these structures lie far beyond the scope of this paper.

REFERENCES AND NOTES

(1) Al-Khnaifes, Kh.; Sachs, H. Graphs, linear equations, determinants, and the number of perfect matchings. In *Contemporary Methods in Graph Theory*; Bodendiek, R., Ed.; B.I. Wissenschaftsverlag: Mannheim—Wien—Zürich, 1990; pp 47–71.

(2) Dewar, M. S. J.; Longuet-Higgins, H. C. The correspondence between the resonance and molecular orbital theories. *Proc. R. Soc. London* **1952**, *214A*, 482–493.

(3) John, P.; Sachs, H. Calculating the numbers of perfect matchings and of spanning trees, Pauling's orders, the characteristic polynomial, and the eigenvectors of a benzenoid system. In *Advances in the Theory of Benzenoid Hydrocarbons*; Gutman, I., Cyvin, S. J., Eds.; Topics in Current Chemistry 153; Springer-Verlag: Berlin—Heidelberg, 1990; pp 145–179.

(4) Kasteleyn, P. W. The statistics of dimers on a lattice I. The number of dimer arrangements on a quadratic lattice. *Physica* **1961**, *27*, 1209–1225.

(5) Kasteleyn, P. W. Graph theory and crystal physics. In *Graph Theory and Theoretical Physics*; Harary, F., Ed.; Academic Press: London, 1967; pp 43–110.

(6) Longuet-Higgins, H. C. Resonance structures and *MO* in unsaturated hydrocarbons. *J. Chem. Phys.* **1950**, *18*, 265–274.

(7) Müller, H. Alternating cycle free matchings in chordal bipartite graphs. University of Jena, Jena, Germany Forschungsergebnisse, 1988.

(8) Pulleyblank, W. R. Alternating cycle free matchings; CORR Report 82-18, April 1982; University of Waterloo: Combinatorics & Optimization, Faculty of Mathematics.

(9) Sachs, H. Counting perfect matchings in lattice graphs. In *Topics in Combinatorics and Graph Theory*; Bodendiek, R., Henn, R., Eds.; Physica-Verlag: Heidelberg, 1990; pp 577–584.

(10) Sachs, H. How to calculate the number of perfect matchings in finite sections of certain infinite plane graphs. *Discrete Math.* **1992**, *101*, 279–284.

(11) Temperley, H. N. V.; Fisher, M. E. Dimer problems in statistical mechanics. An exact result. *Philos. Mag.* **1961**, *6*, 1061–1063.