

P. JOHN, H. SACHS and H. ZERNITZ (Ilmenau)

COUNTING PERFECT MATCHINGS IN POLYOMINOES WITH AN APPLICATION TO THE DIMER PROBLEM

Abstract. A *polyomino* is a connected finite plane graph with no cut-points in which all interior regions (called *cells*) are unit squares. Let P be a given polyomino and e be a given edge of P . A simple algorithm is developed for calculating the numbers $m(P)$ and $m(P, e)$ of all perfect matchings of P and of those perfect matchings of P which contain the edge e , respectively.

The numbers $m(P)$ and $m(P, e)/m(P)$ play an important role in the dimer problem of statistical crystal physics.

1. Introduction. The dimer problem has its origin in the investigation of the thermodynamic properties of a system of diatomic molecules (called *dimers*) adsorbed on the surface of a crystal (see, e.g., [1] and [9]–[11]). In many cases, the most favourable points for the adsorption of atoms form a part L of a square lattice and a dimer can occupy two neighbouring points of L (and only such points). A *dimer covering* is an arrangement C of dimers on L such that every dimer of C occupies two neighbouring points and every point of L is covered by exactly one dimer of C . Let us identify the point set of L with the vertex set of a graph P corresponding to L (Fig. 1a); P is a special polyomino (for the definition see the Abstract; more about polyominoes in [4]). To any dimer covering of L there corresponds a *perfect matching* (PM) M of P (i.e., a set of disjoint edges covering all vertices of P), and conversely (Fig. 1b).

Let x, y be two neighbouring points in L and let $e = (x, y)$ be the edge connecting x and y in P . Suppose that every dimer covering of L occurs with the same probability. The physicist is interested in the number m_L of all dimer coverings of L and in the probability $p_L(x, y)$ to find x and y covered by the same dimer in a randomly chosen covering. Let $m(G)$ and $m(G, e)$ denote the numbers of all PMs of a graph G and of those PMs of G which contain the edge e , respectively. Clearly,

$$m_L = m(P) \quad \text{and} \quad p_L(x, y) = m(P, e)/m(P);$$

further,

$$m(P, e) = m(P - \{x, y\}) = m(P) - m(P - e),$$

where $P - e$ and $P - \{x, y\}$ denote the subgraphs of P obtained from P by omitting the edge e or the vertices x and y and all edges incident to them, respectively. However, in general, $P - e$ and $P - \{x, y\}$ are no longer polyominoes; therefore, we shall extend our investigations to a certain class S of subgraphs of polyominoes (see Section 3). The problem of determining m_L and $p_L(x, y)$ will be considered to be solved as soon as we have a handy method (an algorithm) for calculating $m(G)$ for every graph $G \in S$.

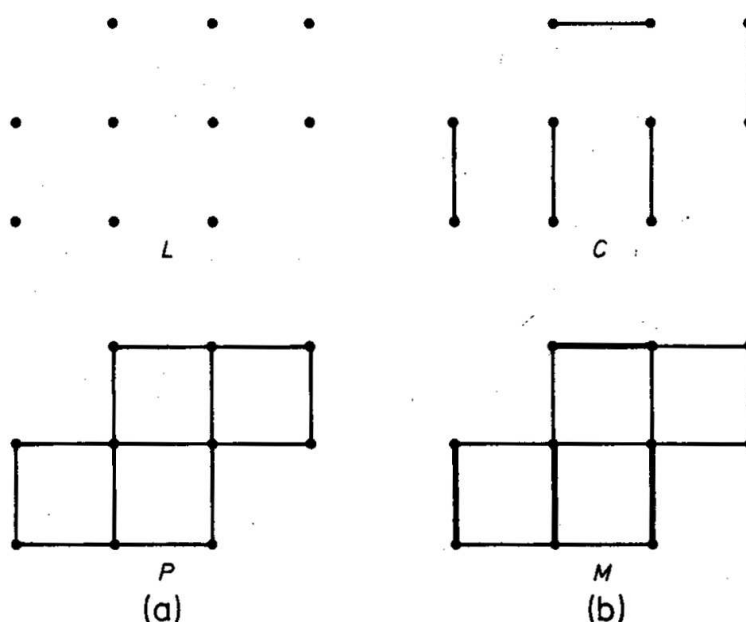


Fig. 1

A very similar question arose in the chemistry of benzenoid hydrocarbons. A *hexagonal system* (HS) is a connected finite plane graph with no cut-points in which all interior regions are regular hexagons of side length one. An HS is the skeleton of some benzenoid hydrocarbon molecule if and only if it has a PM (Kekulé structure) (see, e.g., [7], [8], [12]). Given an HS, the chemist is interested in the number of all PMs as well as in the probability of finding a given edge in some PM (this probability is Pauling's bond order). It turned out that the methods developed for hexagonal systems (see [5]–[8]) can also be applied to polyominoes; however, some preparation is needed.

All graphs G to be considered in the sequel are finite, multiple edges are allowed to occur. $V(G)$ denotes the set of vertices of G , and $n(G) = |V(G)|$.

2. A basic theorem. Let G be a finite plane graph; G subdivides the plane into a finite number f of (connected, open) regions. A region F is called a $(2 \bmod 4)$ -region if the length $l(C)$ of every component C of the boundary of F satisfies

$$(1) \quad l(C) \equiv 2 \pmod{4};$$

G is called a $(2 \bmod 4)$ -graph if all of its regions are $(2 \bmod 4)$ -regions (Fig. 2).

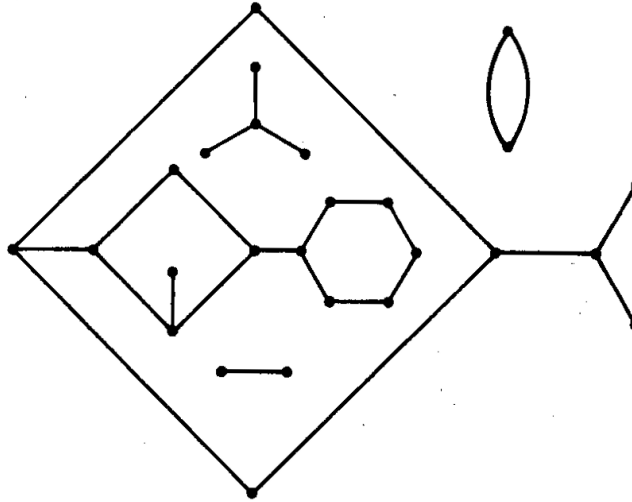


Fig. 2. A $(2 \bmod 4)$ -graph

The following theorem (see [13] and [14]) generalizes a result of Cvetković et al. [2] (see also [3] and [1], 8.2); it follows also from a more general theorem of Kasteleyn expressing the number of perfect matchings of a plane graph in terms of a Pfaffian (see [9]).

THEOREM 1. *Let G be a $(2 \bmod 4)$ -graph with adjacency matrix A which has n vertices and let m denote the number of perfect matchings contained in G . Then*

- (A) n is even,
- (B) G is bipartite,
- (C) $\det A = (-1)^h m^2$, where $h = n/2$.

We need also

THEOREM 2. *Let G be a connected plane graph whose interior regions are all $(2 \bmod 4)$ -regions. Then G is a $(2 \bmod 4)$ -graph if and only if it has an even number of vertices.*

Proof. Let F_0, F_1, \dots, F_{f-1} be the regions of G , where F_0 is the exterior (infinite) region, and let l_i denote the length of the boundary of F_i . We have to show that $l_0 \equiv 2 \pmod{4}$ if and only if $n = n(G)$ is even. Let k be the number of edges of G . Clearly,

$$2k = l_0 + l_1 + \dots + l_{f-1} \equiv l_0 + (f-1) \cdot 2 \pmod{4}.$$

By Euler's polyhedron formula, $2k = 2n + 2f - 4$. Thus $l_0 \equiv 2n + 2 \pmod{4}$.

3. Trapezoidal systems. Let P be a polyomino over a square lattice. Fix a vertex z of P and colour it black; colour all vertices of P black and white so that every edge connects a black vertex with a white one. Lift the white vertices and pull the black ones down by $1/4$ each thus transforming P into a

trapezoidal system $T = T(P)$ (Fig. 3). By this operation the set of edges is partitioned into three classes: long (vertical), short (vertical), and oblique. Subdivide every long edge by inserting two additional vertices (a black one and a white one) so that the three new edges are of length $1/2$ each, as indicated in Fig. 3. By these operations, P is transformed into an "extended trapezoidal system" $T' = T'(P)$ in which, with respect to the lowest white vertices (which define the zero level), every vertex x has a well-defined height $h(x)$ (Fig. 3).

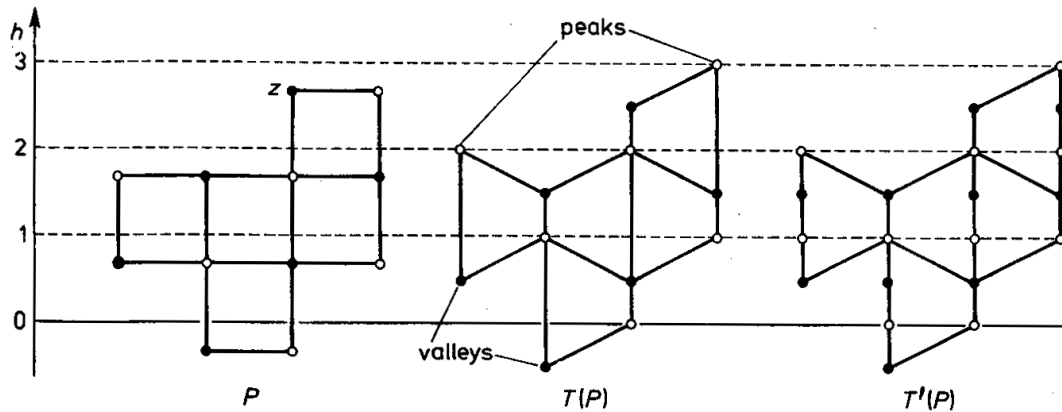


Fig. 3

A white vertex x_0 (black vertex y_0) whose neighbours are all lower than x_0 (higher than y_0) is called a *peak* (*valley*). Let w, b, p, v be the numbers of white vertices, black vertices, peaks, and valleys of $T(P)$, respectively, and let w', b', p', v' have the analogous meanings with respect to $T'(G)$.

OBSERVATION 1. *The peaks and valleys are precisely those vertices which, in $T(P)$, are not incident with a short edge.*

Since the short edges are disjoint and every short edge connects a white vertex with a black one, we conclude that

$$(2) \quad p - v = w - b.$$

Clearly, $p' = p$, $v' = v$, and $w' - b' = w - b$, so

$$(2') \quad p' - v' = w' - b'.$$

OBSERVATION 2. *The PMs of P are in a (1, 1)-correspondence with the PMs of $T(P)$ as well as $T'(P)$; therefore,*

$$(3) \quad m(P) = m(T(P)) = m(T'(P))$$

(Fig. 4).

Let $n = n(P)$ and $n' = n(T'(P))$; clearly, $n' \equiv n \pmod{2}$.

OBSERVATION 3. *If n is even, then (by Theorem 2) $T'(P)$ is a $(2 \bmod 4)$ -graph; therefore (according to Theorem 1 (C) and Observation 2)*

$$(4) \quad m^2 = |\det A'|,$$

where $m = m(P)$ and A' is the adjacency matrix of $T'(P)$.

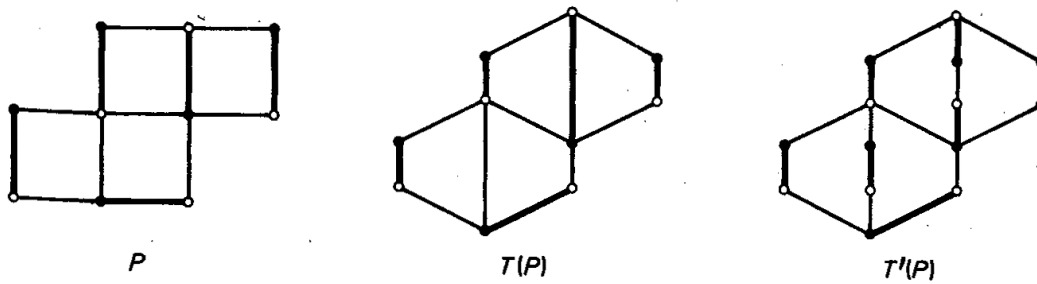


Fig. 4

It will be our main concern to derive from (4) a simple determinant formula for m where the size of the matrix is considerably reduced as compared with the size of A' or A .

Let S^* denote the set of all connected graphs which are subgraphs of some polyomino.

OBSERVATION 4. *The transformations T and T' described above can be applied analogously to any graph $G \in S^*$ transforming G into $T(G)$ and $T'(G)$, respectively, and the statements made in Observations 1 and 2 remain valid for G ; the analogue of formula (4) is also true for G provided $T'(G)$ is a $(2 \bmod 4)$ -graph.*

Let Z denote the set of all $(2 \bmod 4)$ -graphs and put

$$S = \{G \mid G \in S^* \text{ and } T'(G) \in Z\}.$$

Next two simple characterizations of the members of S shall be given.

Let F be an interior region of a graph $G \in S^*$ and let $i(F)$ denote the number of lattice points lying in the interior of F (Fig. 5).

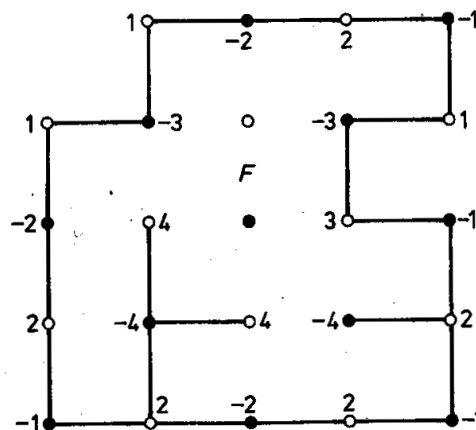


Fig. 5

$$i(F) = 2, h(F) = 0$$

LEMMA 1. *$T'(F)$ is a $(2 \bmod 4)$ -region if and only if $i(F)$ is even.*

This can be proved by induction on the number of cells covered by F .

Let $X(F)$ denote the set of vertices lying on the boundary of F . Let $x \in X(F)$ and let the total measure of the (open) angles which have their

vertex at x and are open towards the interior of F be $k_F(x) \cdot \pi/2$ ($k_F(x) \in \{1, 2, 3, 4\}$). Put

$$\text{sgn}(x) = \begin{cases} 1 & \text{if } x \text{ is white,} \\ -1 & \text{if } x \text{ is black} \end{cases}$$

and

$$\hat{h}(F) = \left| \frac{1}{4} \sum_{x \in X(F)} \text{sgn}(x) \cdot k_F(x) \right|$$

(Fig. 5).

LEMMA 2. For each interior region F of a graph $G \in S^*$, $\hat{h}(F)$ is an integer. $T'(F)$ is a (2 mod 4)-region if and only if $\hat{h}(F)$ is even.

Again, the proof can be carried out by induction on the number of cells covered by F , making use of Lemma 1.

From these lemmata and Theorem 2 we obtain

THEOREM 3. For a graph $G \in S^*$ the following statements are equivalent:

- (i) $G \in S$.
- (ii) $n(G)$ is even and $i(F)$ is even for every interior region F of G .
- (iii) $n(G)$ is even and $\hat{h}(F)$ is even for every interior region F of G .

COROLLARY 1. Whether or not a graph $G \in S^*$ is a member of S does neither depend on the choice of the distinguished vertex z (the colours may be interchanged) nor on the position of G in the plane (G may be turned by multiples of 90°).

COROLLARY 2. Let $G \in S$.

Let e be an edge of G such that $G - e$ is connected; then $G - e \in S$.

Let G' be a connected subgraph of G with an even number of vertices such that $G'' := G - V(G')$ is connected; then $G' \in S$ and $G'' \in S$.

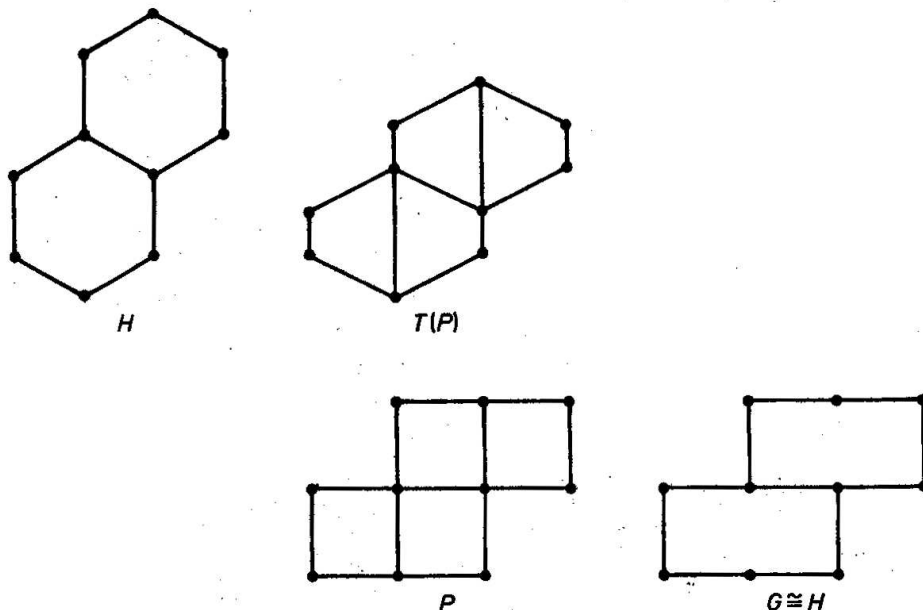


Fig. 6

COROLLARY 3. Any hexagonal system H can be obtained from some polyomino P by deleting all long edges of $T(P)$ (Fig. 6); thus the hexagonal systems which have an even number of vertices may be considered members of S . This implies that the entire theory to be developed for the members of S is a fortiori valid for hexagonal systems with an even number of vertices.

4. Perfect matchings and perfect path systems. Let $G \in S$ and consider $T(G)$ and $T'(G)$. A *pv-path* is a path starting at a peak and running monotonically down to a valley. A *path system* is a set of pairwise disjoint pv-paths; it is called *perfect* if every peak and every valley is contained in some path of the system. Clearly, a necessary condition for a perfect path system (PPS) to exist is that the number of peaks equals the number of valleys, i.e., $p = v$ or, equivalently (by Observation 1), $w = b$ or $w' = b'$, respectively. Evidently, the same condition is necessary for a PM to exist.

Suppose that G has a PM. Let M be any PM of $T(G)$; colour the edges of M red and the others blue. It is not difficult to see that the long and the oblique edges which are red together with the short edges which are blue form a PPS, say $Q =: f(M)$. Conversely: Assume that $T(G)$ has a PPS. Let Q be a PPS of $T(G)$; first colouring the short edges red and all others blue and then interchanging the colours of all edges that lie on some path of Q results in a PM, say $M =: g(Q)$. It is almost evident that g is the inverse of f , i.e., $Q = f(M)$ implies $M = g(Q)$, and conversely (Fig. 7).

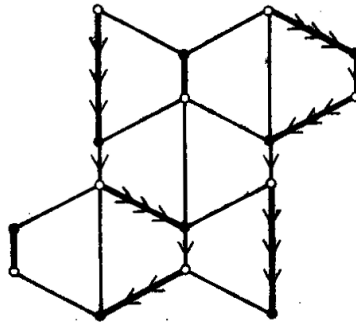


Fig. 7

Thus a (1, 1)-correspondence between the set of PMs (of G , $T(G)$ or $T'(G)$) and the set of PPSs (of $T(G)$ or $T'(G)$) (these sets may be empty) is established, in particular, the number $q = q(T(G)) = q(T'(G))$ of PPSs is equal to the number $m = m(G) = m(T(G)) = m(T'(G))$ of PMs:

THEOREM 4. Let $G \in S^*$. There is a simple (1, 1)-correspondence between the set of perfect matchings of G (or $T(G)$ or $T'(G)$) and the set of perfect path systems of $T(G)$ (or $T'(G)$) implying

$$(5) \quad m(G) = q(T(G)) = q(T'(G)).$$

5. The main theorem. Let $G \in S^*$ and assume that $w = b$ implying, by Observations 4 and 1, $p' = p = v' = v$. Let

$$X_p = \{x_1, x_2, \dots, x_p\} \quad \text{and} \quad Y_v = \{y_1, y_2, \dots, y_p\}$$

be the sets of the peaks and the valleys, respectively, of $T(G)$ or of $T'(G)$ (which, without danger of confusion, can be identified) and let q_{ik} denote the number of pv-paths connecting x_k with y_i ($1 \leq i, k \leq p$); clearly, these numbers are the same for $T(G)$ and $T'(G)$. Put $Q := (q_{ik})$.

THEOREM 5. For any graph $G \in S$ with as many black vertices as white ones,

$$(6) \quad m(G) = q(T(G)) = |\det Q|.$$

As to the efficiency of Theorem 5, it is important to note that the numbers q_{ik} can be very easily calculated. Let x be any vertex of $T(G)$ or $T'(G)$, let $q_k(x)$ denote the number of monotone paths issuing from the peak x_k and terminating at x ($k = 1, 2, \dots, p$), put

$$q(x) := (q_1(x), q_2(x), \dots, q_p(x));$$

clearly, $q_k(y_i) = q_{ik}$ ($i = 1, 2, \dots, p$) and

$$(7) \quad Q = (q_{ik}) = \begin{bmatrix} q(y_1) \\ \vdots \\ q(y_p) \end{bmatrix}.$$

Let $U(x)$ denote the "upper neighbourhood" of x , i.e., the set of neighbours x' of x satisfying $h(x') > h(x)$. In order to calculate the vectors $q(x)$, note simply the following:

(i) For any peak x_k ,

$$q(x_k) = (\delta_{k1}, \delta_{k2}, \dots, \delta_{kp}),$$

where $\delta_{ii} = 1$, $\delta_{ij} = 0$ if $i \neq j$ ($k = 1, 2, \dots, p$).

(ii) For any vertex x which is not a peak,

$$(8) \quad q(x) = \sum_{x' \in U(x)} q(x')$$

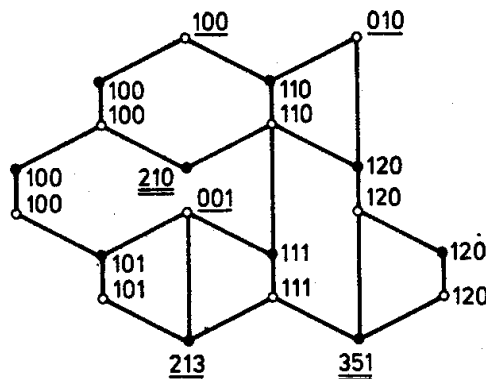


Fig. 8

(especially, $q(x) = 0$ if x is a black vertex which has no upper neighbours). Running through G from top to bottom, the $q(x)$ can now be successively determined. See Fig. 8, where

$$Q = \begin{bmatrix} 2 & 1 & 0 \\ 2 & 1 & 3 \\ 3 & 5 & 1 \end{bmatrix}.$$

6. Proof of Theorem 5. Let G be as in Theorem 5 and consider $T'(G)$: by Observations 4 and 1, $w' = b'$. Let

$$X' = \{x_1, x_2, \dots, x_{w'}\}$$

and

$$Y' = \{y_1, y_2, \dots, y_{w'}\} = \{x_{w'+1}, x_{w'+2}, \dots, x_n\}$$

(where $y_i = x_{w'+i}$) be the sets of the white and the black vertices, respectively, where (as above) x_1, x_2, \dots, x_p are the peaks and y_1, y_2, \dots, y_p are the valleys. The adjacency matrix of $T'(G)$ then takes the form

$$A' = \begin{bmatrix} O & B'^T \\ B' & O \end{bmatrix},$$

and from Observations 2–4 and Theorem 4 we obtain

$$(9) \quad m(G) = q(T(G)) = m(T'(G)) = |\det B'|.$$

We have to show that $|\det B'| = |\det Q|$. This will be performed by applying a simple Gaussian elimination process (in a more or less disguised form) to $\det B'$ reducing it to $\pm \det Q$.

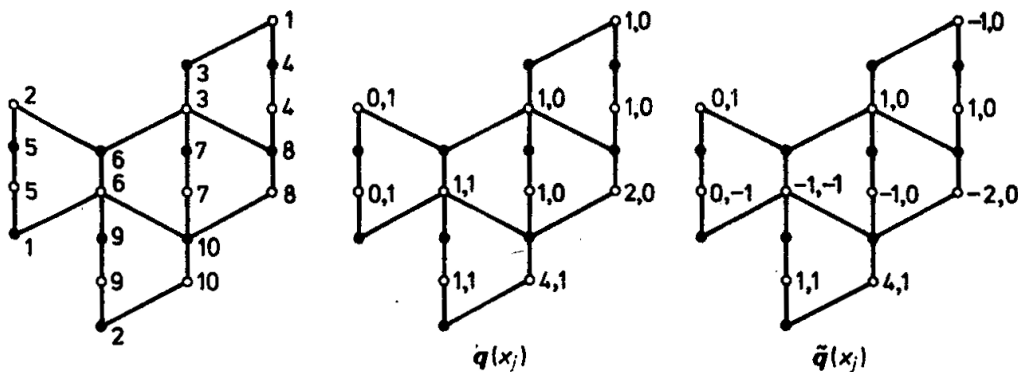


Fig. 9 (see also Fig. 3)

We may assume that the vertices are numbered as follows (Fig. 9):

- (i) numbers $1, 2, \dots, p$ are reserved for the peaks and valleys;
- (ii) any white vertex which is not a peak is given a number

$$j \in \{p+1, p+2, \dots, w'\}$$

such that $h(x_i) > h(x_k)$ implies $i < k$ ($i, k \in \{p+1, p+2, \dots, w'\}$);

(iii) every black vertex which is not a valley is given the same number as its unique lower neighbour.

Then B' takes the form

$$(10) \quad B' = \begin{bmatrix} C & U \\ V & D \end{bmatrix},$$

where $d = w' - p$ and $D = (d_{ik}) = (b'_{p+i, p+k})$ ($i, k = 1, 2, \dots, d$) is a triangular matrix satisfying $d_{ik} = 0$ if $i < k$, $d_{ii} = 1$. Thus

$$(I) \quad \det D = 1.$$

Put

$$(11) \quad [C \ U] =: R \quad \text{and} \quad [V \ D] =: S;$$

so, by (10),

$$(12) \quad B' = \begin{bmatrix} R \\ S \end{bmatrix}.$$

Let I_s denote the $s \times s$ unit matrix. Clearly,

$$[q^T(x_1), q^T(x_2), \dots, q^T(x_p)] = I_p.$$

Put

$$[q^T(x_{p+1}), q^T(x_{p+2}), \dots, q^T(x_{w'})] =: F,$$

$$[I_p \ F] = [q^T(x_1), q^T(x_2), \dots, q^T(x_{w'})] =: H;$$

put, further,

$$(13) \quad (-1)^{h(x_j)} q(x_j) =: \tilde{q}(x_j) \quad (j = 1, 2, \dots, w')$$

(Fig. 9),

$$[\tilde{q}^T(x_1), \tilde{q}^T(x_2), \dots, \tilde{q}^T(x_p)] =: \tilde{I}_p,$$

$$[\tilde{q}^T(x_{p+1}), \tilde{q}^T(x_{p+2}), \dots, \tilde{q}^T(x_{w'})] =: \tilde{F},$$

$$[\tilde{I}_p \ \tilde{F}] =: \tilde{H}, \quad [O \ I_d] =: \tilde{K}$$

and

$$(14) \quad \begin{bmatrix} \tilde{I}_p & O \\ \tilde{F}^T & I_d \end{bmatrix} = [\tilde{H}^T \ \tilde{K}^T] =: Z;$$

note that

$$(II) \quad |\det Z| = |\det \tilde{I}_p| \cdot |\det I_d| = 1.$$

The $(p \times w)$ -matrix R (see (11)) reflects the neighbourhoods of the valleys; therefore, because of (8),

$$RH^T = Q = [q^T(y_1), q^T(y_2), \dots, q^T(y_p)]^T$$

(see (7)). Put

$$(15) \quad R\tilde{H}^T =: \tilde{Q} =: [\tilde{q}_1^T, \tilde{q}_2^T, \dots, \tilde{q}_p^T]^T$$

and note that the i -th row \tilde{q}_i of \tilde{Q} is either equal to the i -th row $q(y_i)$ of Q or differs from it only by the factor -1 (in fact, because of (13) we have

$$\tilde{q}_i = (-1)^{b(y_i) + 1/2} q(y_i); \text{ thus}$$

$$(III) \quad |\det \tilde{Q}| = |\det Q|.$$

The $(d \times w)$ -matrix $S = [V \ D]$ (see (11)) reflects the neighbourhoods of those black vertices which are not valleys; therefore, by (8) and (13),

$$(16) \quad S\tilde{H}^T = O.$$

Further,

$$(17) \quad R\tilde{K}^T = [C \ U] \begin{bmatrix} O \\ I_d \end{bmatrix} = U,$$

$$(18) \quad S\tilde{K}^T = [V \ D] \begin{bmatrix} O \\ I_d \end{bmatrix} = D.$$

Equations (12) and (14)–(18) yield

$$(19) \quad B'Z = \begin{bmatrix} R \\ S \end{bmatrix} [\tilde{H}^T \ \tilde{K}^T] = \begin{bmatrix} \tilde{Q} & U \\ O & D \end{bmatrix};$$

thus

$$(IV) \quad \det(B'Z) = (\det \tilde{Q})(\det D).$$

From (II), (IV), (I), and (III) we now obtain in order

$$\begin{aligned} |\det B'| &= |\det B'| \cdot |\det Z| = |\det(B'Z)| \\ &= |\det \tilde{Q}| \cdot |\det D| = |\det \tilde{Q}| = |\det Q|. \end{aligned}$$

This proves the theorem.

7. An example. In how many ways can a 5×6 "chess-board" be covered by 15 dominoes such that each domino covers exactly two fields and each field is covered (by exactly one domino)? The answer is given by Fig. 10, where

$$Q = \begin{bmatrix} 52 & 39 \\ 39 & 52 \end{bmatrix} \quad \text{and} \quad m = |\det Q| = 1183.$$

More about covering chess-boards by dominoes in our "Problem" (see the Problems Section of this issue).

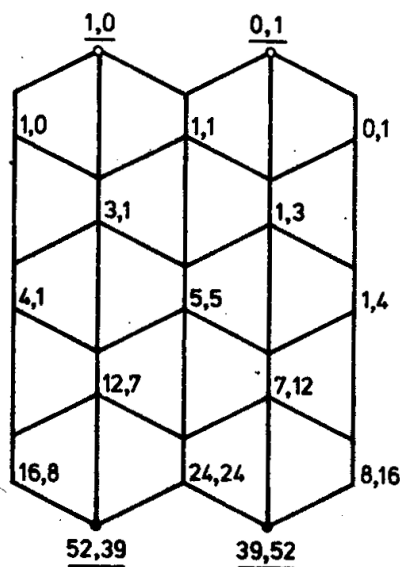


Fig. 10

8. A different approach. We have proved

(i) $q(T(G)) = m(G)$ (Theorem 4)

and

(ii) $m(G) = |\det Q|$ (Theorem 5)

implying

(iii) $q(T(G)) = |\det Q|$ (Theorem 5).

For (i) we found a simple, intuitive combinatorial proof whereas (ii) was a bit harder. Comparing (i), (ii), and (iii), one is led to try to eliminate the concept of a perfect matching altogether by directly proving (iii), since this is a sort of an inclusion-exclusion principle for paths and path systems that has nothing to do with perfect matchings, and then to obtain (ii) from (i) and (iii) very easily — indeed, a plausible and challenging idea. However, there are some obstacles. That this program can nevertheless be carried through is shown in all details by Gronau et al. [5] who found a general determinant formula, discussed under what conditions it is valid, and applied it to hexagonal systems.

References

- [1] D. M. Cvetković, M. Doob and H. Sachs, pp. 245–251 in: *Spectra of Graphs Theory and Application*, VEB Deutscher Verlag der Wissenschaften, Berlin 1980–1982.
- [2] D. M. Cvetković, I. Gutman and N. Trinajstić, *Graph theory and molecular orbitals*, VII. The role of resonance structures, J. Chem. Phys. 61 (1974), pp. 2700–2706.

- [3] M. S. J. Dewar and H. C. Longuet-Higgins, *The correspondence between the resonance and molecular orbital theories*, Proc. Roy. Soc. (London) A 214 (1952), pp. 482–493.
- [4] S. Golomb, *Polyominoes*, Scribner's, New York 1965.
- [5] H.-D. O. F. Gronau, W. Just, W. Schade, P. Scheffler and J. Wojciechowski, *Path systems in acyclic directed graphs*, this issue, pp. 399–411.
- [6] P. John and J. Rempel, *Counting perfect matchings in hexagonal systems*, pp. 72–79 in: H. Sachs (ed.), *Graphs, Hypergraphs and Applications*, Teubner-Texte zur Math., Band 73, Teubner, Leipzig 1985.
- [7] P. John and H. Sachs, *Calculating the number of perfect matchings and Pauling's bond orders in hexagonal systems whose inner dual is a tree*, pp. 80–91 in: H. Sachs (ed.), *Graphs, Hypergraphs and Applications*, Teubner-Texte zur Math., Band 73, Teubner, Leipzig 1985.
- [8] – *Wegesysteme und Linearfaktoren in hexagonalen und quadratischen Systemen*, pp. 85–101 in: *Graphen in Forschung und Unterricht* (Festschrift K. Wagner), Verlag Barbara Franzbecker, Bad Salzdetfurth 1985.
- [9] P. W. Kasteleyn, *Graph theory and crystal physics*, pp. 84–106 in: F. Harary (ed.), *Graph Theory and Theoretical Physics*, Academic Press, London 1967.
- [10] E. W. Montrol, *Lattice statistics*, pp. 96–143 in: E. F. Beckenbach (ed.), *Applied Combinatorial Mathematics*, John Wiley, New York 1964.
- [11] J. K. Percus, *Combinatorial Methods*, Springer-Verlag, Berlin 1971, pp. 91–123.
- [12] H. Sachs, *Perfect matchings in hexagonal systems*, *Combinatorica* 4 (1984), pp. 89–99.
- [13] – *Linearfaktoren in hexagonalen Systemen*, pp. 93–96 in: 30. *Internationales Wissenschaftliches Kolloquium* (21.–25. 10. 1985), Heft 5, Technische Hochschule Ilmenau 1985.
- [14] – *On the number of perfect matchings in bipartite plane graphs*, pp. 65–70 in: *Algebra und Graphentheorie* (Jahrestagung Siebenlehn, 28. 10.–1. 11. 1985), Bergakademie Freiberg 1986.