Counting Perfect Matchings and Benzenoids

Fred J. Rispoli
Dowling College

Summary

The connection between perfect matchings and benzene was discovered by the German chemist Kekulé in the mid 1800s. Subsequently, chemists have learned that the number of perfect matchings contained in a molecular model is an important parameter related to chemical stability. Hence, counting perfect matchings has been an important problem in chemistry for over 50 years. However, counting perfect matchings in general graphs is a computationally difficult problem. Consequently, chemists and graph theorists have developed efficient counting methods for certain classes of graphs that arise in modeling special hydrocarbons called benzenoids. Many of these methods involve counting principles usually discussed in discrete mathematics courses. In this article we discuss several of these methods and show how to implement a general determinant based formula.

Notes for the instructor

This project works well as an enrichment topic for an advanced discrete mathematics course focused on applications. Students should be familiar with counting techniques, graphs and determinants. I usually give the paper to students to read and then present a summary of the material at the end of the course. I spend roughly one class meeting on it. Exercises that reinforce and extend some key ideas are given in the last section, along with selected solutions.

Bibliography

[1] J. Aihara, “Why Aromatic Compounds Are Stable,” *Scientific American* 266(3) (1992) 62–68.

[2] S. Cyvin and I. Gutman, *Kekulé Structures in Benzenoid Hydrocarbons*, Springer-Verlag, New York, 1988.

[3] T. Došlić, “Perfect Matchings, Catalan Numbers, and Pascal’s Triangle.” *Mathematics Magazine* 80 (2007) 219–225.

[4] R.A. Horn and C.R. Johnson, *Matrix Analysis*, Cambridge University Press, New York, 1985

[5] S. Iijima, “Helical Microtubules of Graphitic Carbon,” *Nature* 354 (1991) 56–58.

[6] L. Lovász and M.D. Plummer, *Matching Theory*, North Holland, Amsterdam, 1986.

[7] P.W. Kasteleyn, “Graph theory and crystal physics,” in F. Harary, editor, *Graph Theory and Theoretical Physics*, Academic Press, 1967.

[8] J. Propp, “Enumerations of Matchings: Problems and Progress,” *New Perspectives in Geometric Combinatorics*, MSRI Publications 38 (1999) 255–291.

[9] F. Rispoli, “Counting Perfect Matchings in Hexagonal Systems Associated with Benzenoids,” *Mathematics Magazine* 74 (2001) 194–200.

[10] H. Sachs, "Perfect Matchings in Hexagonal Systems," *Combinatorica* 4 (1984) 89–99.

[11] H. Sachs, P. Hansen and M. Zheng, "Kekulé count in tubular hydrocarbons," *MATCH Communications in Mathematical and in Computer Chemistry* 33 (1996) 169–241.

[12] N. Trinajstić, *Chemical Graph Theory*, Second Edition, CRC Press, 1992.

Perfect Matchings and Benzenoids

Given a graph G, a *perfect matching* is a subgraph M that contains every vertex of G such that all vertices in M have degree 1. Chemists, who use graphs to model hydrocarbon molecules, are interested in perfect matchings since they provide possible double-bond arrangements for carbon bonds. For example, Figure 1 contains two different graph models of benzene. The duplicate edges in the model on the left illustrate possible locations for double carbon bonds. Since every carbon atom must have exactly four bonds, there are actually two possible arrangements for the double carbon bonds. To study possible locations of double-bonds, chemists model benzene-like molecules called "benzenoids," by omitting the hydrogen bonds and use special hexagonal grid graphs such as the model on the right in Figure 1 (and also Figures 2-8). In these models perfect matchings are used to indicate the location of the double bonds. In the graph on the right side of Figure 1 the bold edges are used to illustrate the perfect matching corresponding to the double bonds in the graph on the left.

It is a well-established fact that, roughly speaking, for benzenoid molecules with the same number of hexagons, the chemical stability increases with the number of perfect matchings in its associated hexagonal graph. Surprisingly, when we consider special classes of benzenoids and count perfect matchings, many well-known numbers arise such as Fibonacci numbers and powers of 2. In this article we shall discuss techniques used to enumerate perfect matchings in hexagonal graphs.

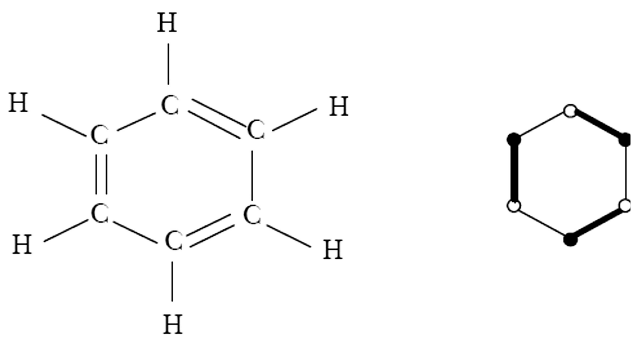


Figure 1. A graph model of benzene and its associated hexagonal graph.

A graph G is called 2-*connected* if it is connected and at least two vertices must be removed to make G disconnected. A *hexagonal system* is a 2-connected planar graph such that each interior face can be drawn as a regular hexagon. Notice that this condition forces all vertex degrees in a hexagonal system to be 3 or 2, and only 2 on the exterior boundary. Moreover, every pair of adjacent hexagons have exactly one edge in common. For convenience, all hexagonal systems illustrated in this article will be drawn with two vertical edges. We leave it as an exercise to show that every hexagonal system is bipartite which can be proved using induction on the number of hexagons in the hexagonal system. Hence, we may partition the vertices into subsets of black vertices $B = \{b_i\}$ and white vertices $W = \{w_j\}$ which is illustrated in Figure 2. Clearly G must have an even number of vertices in order for it to contain a perfect matching M. Moreover, if G has $2n$ vertices, then M must have exactly n edges.

A *hydrocarbon* is a substance consisting only of carbon and hydrogen atoms. A *benzenoid* is a special type of hydrocarbon that has a benzene like structure. Every benzenoid has a unique corresponding hexagonal system H obtained by removing the edges representing carbon-hydrogen bonds and letting the remaining edges of H represent either single or double carbon-carbon bonds. In Figure 2 the hexagonal system obtained from naphthalene is given. Observe that there are eight carbon-hydrogen bonds that have been suppressed, all resulting in degree 2 vertices. It is known that all hexagonal systems that arise from benzenoids contain a perfect matching.

For any graph G, the number of perfect matchings in G is denoted by $\phi(G)$. For example, in Figure 2 the bold edges illustrate a perfect matching, and the reader should confirm that $\phi(H) = 3$. For general graphs G, computing $\phi(G)$ is known to be an NP-hard problem (see [6]). That is, for an arbitrary graph with n vertices, there is no known algorithm to compute $\phi(G)$ involving $O(n^k)$ operations where k is a fixed constant. This remains true even when G is a bipartite graph. However, as we shall see, there are efficient methods that can be used to compute $\phi(G)$ for special classes of planar graphs such as hexagonal systems, and also explicit formulas for special types of hexagonal systems.

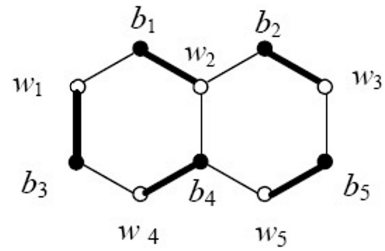


Figure 2. The hexagonal system associated with naphthalene.

Counting Perfect Matchings for Special Classes

A *linear chain of length* h, denoted by L, consists of h hexagons such that all adjacent pairs of hexagons share exactly one vertical edge and no nonvertical edges. A linear chain of length 2 is given in Figure 2. Observe that for any linear chain there is a one-to-one correspondence between perfect matchings in L and the vertical edges in L. Thus every linear chain of length h has exactly $h+1$ perfect matchings. Linear chains are very important because they are used as building blocks of more complex hexagonal systems.

Given any hexagon in a hexagonal system, we refer to its six edges using the terms northeast, east, southeast, southwest, west and northwest. The *northern* edges consist of both the northwest and northeast edges; *southern* edges are defined similarly. In Figure 2 the bold edges are precisely the two northeast edges, the two southeast edges and the western edge in the western-most hexagon. A *rectangular hexagonal system*, denoted by $R[h, v]$, consists of v linear chains, $L_1, \ldots, L_v$, each of length h, together with $v-1$ linear chains $\overline{L}_1, \ldots, \overline{L}_{v-1}$ each of length $h-1$, such that for $i = 1, \ldots, v-1$, all northern edges of $\overline{L}_i$ are southern edges of L_i, and all southern edges of $\overline{L}_i$ are northern edges of L_{i+1}. For example, a linear chain with h hexagons is $R[h, 1]$, and $R[5, 4]$ appears in Figure 3.

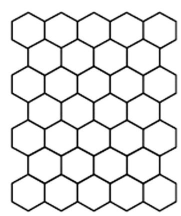


Figure 3. The rectangular hexagonal system $R[5, 4]$.

Theorem 1. *For a rectangular hexagonal system* $R[h, v]$, *we have* $\phi(R[h, v]) = (h+1)^v$.

Proof. Every perfect matching of $R[h, v]$ contains only edges in the v linear chains $L_1, \ldots, L_v$ (i.e., the odd indexed/longer rows of hexagons), and none of the vertical edges in the $v-1$ linear chains $\overline{L}_1, \ldots, \overline{L}_{v-1}$ (i.e., the even indexed/shorter rows). To see this suppose that we color every other vertex in the top linear chain L_1 white, and the remaining vertices black, where the vertical edges in $\overline{L}_1$ are incident to the black vertices. If one or more of the vertical edges from $\overline{L}_1$ is part of a perfect matching, then finding matching edges for the remainder of L_1 is impossible, since there will be more white vertices available than black. The fact that none of the vertical edges from $\overline{L}_1$ can be used in a perfect matching forces the same for $\overline{L}_2, \ldots, \overline{L}_{v-1}$. Hence, any perfect matching from L_i can be used with any perfect matching of any of the other linear chains L_j, $j \neq i$, and hence, $\phi(R[h, v])$ is the product of $\phi(L_1)$, $\phi(L_2)$, $\ldots$, $\phi(L_v)$. Since $\phi(L_i) = h+1$, for all $i = 1, 2, \ldots, v$, we have that $\phi(R[h, v]) = (h+1)^v$. □

A *parallelogram hexagonal system*, denoted by $P[h, v]$, consists of v linear chains $L_1, \ldots, L_v$, each of length h, such that for $i = 1, \ldots, v-1$, all southern edges of hexagons in L_i, except for the southeast edge of the eastern-most hexagon in L_i, are also northern edges of hexagons in L_{i+1}, except for the northwest edge of the western-most hexagon in L_{i+1}; Figure 4 shows $P[5, 3]$. To count perfect matchings in parallelogram hexagonal systems we need the well-known function for counting combinations, $C(n, r) = \frac{n!}{r!(n-r)!}$, where n and r nonnegative integers, with $0 \leq r \leq n$.

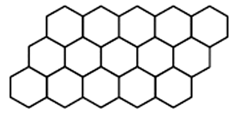


Figure 4. The parallelogram hexagonal system $P[5, 3]$.

Theorem 2. *For a parallelogram hexagonal system $P[h, v]$, we have $\phi(P[h, v]) = C(h + v, v)$.*

Proof. The proof is by induction on $k = h + v$. Clearly, the formula holds for $P[1, 1]$ (which corresponds to benzene and is given in Figure 1). Notice that $P[h, 1]$ and $P[1, h]$ are linear chains of length h. Hence, $\phi(P[h, 1]) = h + 1$ and the formula holds for both $P[h, 1]$ and $P[1, h]$. Assume that the result holds for all parallelogram hexagonal systems $P[h, v]$ with $h + v = k$.

Consider $P[h, v + 1]$ (the case $P[h + 1, v]$ is similar). Let L be the northern-most linear chain of $P[h, v + 1]$, and let M be a perfect matching of $P[h, v + 1]$. Suppose that M contains the eastern-most vertical edge of L (see the graph on the left in Figure 5). Then M must also contain the northwest edge of every hexagon in L. The remaining edges in M can be any perfect matching of the hexagonal system $P[h, v + 1]$ with L removed; that is, $P[h, v]$. By the inductive assumption, there are $C(h + v, v)$ such perfect matchings.

Suppose that M contains the eastern-most northeast edge of L (see the graph on the right in Figure 5.). Then M must also contain the southeast edge of every hexagon in the eastern-most hexagon of every linear chain in $P[h, v+1]$. The remaining edges of M can now be any perfect matching of the hexagonal system $P[h, v+1]$ with the eastern-most hexagon removed from every row; that is, $P[h-1, v+1]$ By the inductive assumption, there are $C(h+v, v+1)$ such perfect matchings.

Every perfect matching in $P[h, v + 1]$ contains either the eastern-most vertical edge of L or the eastern-most northeast edge of L, but not both. Therefore, $\phi(P[h, v + 1]) = C(h + v, v) + C(h + v, v + 1)$. By Pascal's Identity, $\phi(P[h, v + 1]) = C(h + v + 1, v + 1)$. □

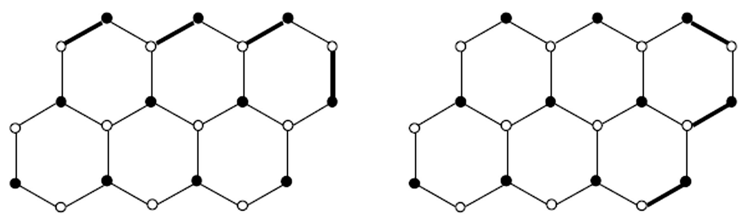


Figure 5. An illustration of the proof of Theorem 2.

Example 1.

(a) Use Theorem 1 to compute the number of perfect matchings for $R[5, 4]$.

(b) Use Theorem 2 to compute the number of perfect matchings for $P[5, 3]$.

(c) Determine the number of perfect matchings in Figure 6 which is obtained from $R[5, 4]$ by deleting the interior hexagons.

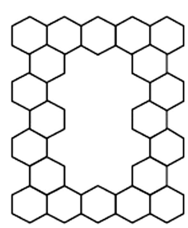


Figure 6. $R[5, 4]$ with the interior hexagons deleted.

Solution.

(a) $\phi(R[5,4]) = 6^4 = 1296$.

(b) $\phi(P[5,3]) = C(8,3) = 56$.

(c) Consider the two hexagons adjacent to the linear chain on top, and label them H_1 and H_2, with H_1 on the left side. As in the proof of Theorem 1, none of the four vertical edges in H_1 and H_2 can be part of a perfect matching. Moreover, the southeast edge of H_1 must be in every perfect matching, and this forces a sequence of eight edges that must also be in every perfect matching until we get to the linear chain on the bottom. A similar sequence of eight edges starting with the southwest edge of H_2 must be in every perfect matching. But again, no vertical edges from the two hexagons adjacent to the linear chain on the bottom can be used in a perfect matching. Consequently, any perfect matching can be used for the linear chain on the bottom. Since there are six possible perfect matchings for the linear chain on top which can be used with any of the six perfect matchings for the linear chain on the bottom we get a total of $6^2 = 36$ perfect matchings. □

A hexagonal system is called a *fibonaccene chain* if it consists of a chain of hexagons $H_1, \ldots, H_v$, with H_1 on top, and only the following shared edges: For i even and $1 < i < v$, H_i shares its northwest edge with H_{i-1} and its southwest edge with H_{i+1}; H_1 shares only its southeast edge and H_v shares only its northeast edge when v is odd, and only its northwest edge when v is even. The name is due to the fact that ϕ satisfies a Fibonacci-style recurrence relation, given in Theorem 3. The proof is left as an exercise.

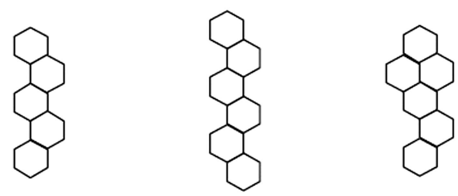


Figure 7. Fibonaccene chains with 5 and 6 hexagons, and a fibonaccene chain with an additional hexagon.

Theorem 3. *Let H be a fibonaccene chain with h hexagons and let $a_h = \phi(H)$. Then $a_1 = 2, a_2 = 3$, and $a_h = a_{h-1} + a_{h-2}$ for $h \geq 3$.*

Example 2.

(a) Use Theorem 3 to determine the number of perfect matchings for the two fibonaccene chains given on the left in Figure 7.

(b) Determine the number of perfect matchings for the fibonaccene chain with an additional hexagon given in Figure 7.

Solution.

(a) We need to find the terms a_5 and a_6 in the sequence 2, 3, 5, 8, 13, 21. Hence the chain on the left contains 13 perfect matchings and the chain in the middle of Figure 7 contains 21.

(b) Let H_1 be the additional hexagon. Every perfect matching must contain either the northwest edge of H_1 or its western vertical edge. If a perfect matching contains the northwest edge of H_1, then this forces eight edges that must be in the perfect matching, and leaves two possibilities for matching edges in the hexagon on the bottom. So there are two perfect matchings with this edge. If a perfect matching contains the western vertical edge, then any perfect matching of the fibonaccene chain with five hexagons can be used. Thus, there is a total of $2 + 13 = 15$ perfect matchings. □

The final special class we consider arises from tubular benzenoids which were discovered in the early 1990s. A *tubulene* is a benzenoid whose carbon skeleton is a rectangular hexagonal system embedded in a cylinder with open ends (top and bottom). There are several different types of tubulene structures depending on how much they are "twisted," but here we consider only the untwisted variety (for more information on tubulenes and perfect matchings, see [5] and [11].) Let $T[h,k]$ denote the tubulene obtained from $R[h,v]$ embedded in a cylinder, where $k = 2v - 1$, i.e., k is the number of rows of hexagons. For example, if $T[5, 1]$ is drawn on paper it would look like a linear chain of length 5 with its left most vertical edge and its right most vertical edge "glued" together. Notice that as a result of the glued edges, the number of perfect matchings is reduced from 6 to 4. It is left for the reader to show that $\phi(T[5, 1]) = 4$. In Figure 8 an illustration of $T[5, 7]$ is given. The long vertical lines are used to indicate the edges that are glued together; or equivalently, the vertical line used to cut the cylinder.

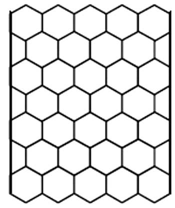


Figure 8. The tubulene hexagonal system $T[5, 7]$.

Theorem 4. *For a tubulene hexagonal system $T[h, k]$, we have $\phi(T[h, k]) = 2^{k+1}$.*

A proof of Theorem 4 is left as an exercise. By applying the theorem we see that $\phi(T[5, 7]) = 2^8 = 256$. It is interesting to note that $\phi(T[h, k])$ is independent of h, but dependent only on k. Thus $\phi(T[h, k])$ will increase if the tubulene is extended in the vertical direction, but does not change at all when it is extended horizontally. What does this imply in terms of chemical properties of tubulene?

To analyze stability levels and obtain a measure of energy, chemists have investigated a graph parameter based on ϕ defined as follows. Given a hexagonal system H, let $\eta(H)$ denote the number of hexagons in H. Then the *Kekulé index* is defined by $\kappa(H) = \frac{\log_2 \phi(H)}{\eta(H)}$. This index is considered by some chemists as an "average resonance energy per hexagon," and is known to satisfy $0 \leq \kappa(H) \leq 1$. For more details, see [11]. If we now compare the Kekulé index of $T[5, 7]$ and $R[5, 4]$, we see that $\phi(T[5, 7]) = 2^8$, which implies that $\kappa(T[5, 7]) = 8/35 = 0.2286$. From Theorem 1, we know that $\phi(R[5, 4]) = 6^4$, and consequently $\kappa(R[5, 4]) = 10.3399/32 = 0.3231$. As a general rule higher energy implies less stability. Since the benzenoid associated with $R[5, 4]$ has a higher energy level compared to $T[5, 7]$, we conclude that it is less stable than the benzenoid associated with $T[5, 7]$.

The Determinant Formula

Next we discuss a method of computing $\phi(H)$ for all hexagonal systems. Since every hexagonal system is bipartite, they can be represented with matrices defined as follows. Let H be a hexagonal system and let E denote the set of edges in H. Let $B \cup W$ be the set of vertices of H, where $B = \{b_1, \ldots, b_n\}$ and $W = \{w_1, \ldots, w_n\}$. We assume that B and W contain the same number of vertices since this is a necessary condition for the existence of a perfect matching. Now define the *$n \times n$ biadjacency matrix* $A(H) = [a_{ij}]$, by $a_{ij} = 1$ if $\{b_i, w_j\} \in E$ and $a_{ij} = 0$ if $\{b_i, w_j\} \notin E$. The biadjacency matrix for the hexagonal system in Figure 2 is given below where the columns represent $w_1, \ldots, w_5$, and the rows represent $b_1, \ldots, b_5$.

$$A(H) = \begin{bmatrix} 1 & \mathbf{1} & 0 & 0 & 0 \\ 0 & 1 & \mathbf{1} & 0 & 0 \\ \mathbf{1} & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & \mathbf{1} & 1 \\ 0 & 0 & 1 & 0 & \mathbf{1} \end{bmatrix}$$

Recall that given an $n \times n$ matrix A, the *determinant* of A is defined by

$$det(A) = \sum_{\sigma} (sgn\ \sigma) a_{1\sigma(1)} a_{2\sigma(2)} \ldots a_{n\sigma(n)}$$

where the sum runs over all $n!$ permutations σ of $\{1, 2, \ldots, n\}$, and $sgn\ \sigma$ is $+1$ or -1, according to whether σ is an even permutation or an odd one (for a reference on determinants, see [4].) By storing the adjacency data of a hexagonal system H with $2n$ vertices in an $n \times n$ matrix, we can establish a one-to-one correspondence between perfect matchings in H and the nonzero terms in the expansion of $det(A(H))$. For example, the perfect matching M below corresponds to the permutation σ shown.

$$M = \{\{b_1, w_2\}, \{b_2, w_3\}, \{b_3, w_1\}, \{b_4, w_4\}, \{b_5, w_5\}\} \qquad \sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 3 & 1 & 4 & 5 \end{pmatrix}$$

Observe that when the term associated with σ is used in calculating the determinant of $A(H)$, the product of the bold numbers in the $A(H)$ given above is computed and gives a result of 1. This is true in general, the nonzero terms of $det(A(H))$ whose permutations correspond to a perfect matching will either be 1 or -1, and it can be proven that every nonzero term will have the same sign (for a proof, see [9].) The terms whose associated permutation do not correspond to a perfect matching will yield a 0 in the determinant expansion. Hence, $|det(A(H))|$ counts the number of permutations in H. This fact was first proven by Kasteleyn [7].

Theorem 5. *For a hexagonal system H, $\phi(H) = |det(A(H))|$.*

Theorem 5 may be implemented using many different technologies. Most of the time required is in labeling the vertices, finding the biadjacency matrix and entering data. For example, to compute $\phi(H)$ for the hexagonal system given in Figure 9, the computation requires finding the biadjacency matrix $A(H)$ which is a 32×32 matrix. Calculating the determinant gives $\phi(H) = 1,764$.

Example 3.

Use the determinant formula to compute the number of perfect matchings contained in the hexagonal system given in Figure 9.

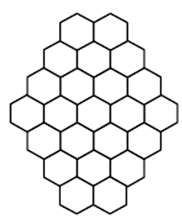


Figure 9. A hexagonal system with 1,764 perfect matchings.

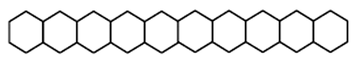
We conclude this article by noting that computing $\phi(G)$ is an important problem in many areas of science and mathematics. For example, perfect matching enumeration is used in the famous dimer problem in physics and to help solve many tiling problems. Counting perfect matchings in the parallelogram hexagonal system $P[h, v]$ can also be used to count non-decreasing sequences of length v with elements from $\{0, 1, \ldots, h\}$, and lattice paths in a rectangular lattice (see [3]).

Exercises on Counting Perfect Matches and Benzenoids

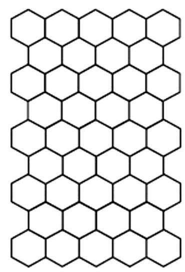
1. (a) Draw all three perfect matchings contained in the linear chain of length 2 given in Figure 2.
   (b) Show that there is a one-to-one correspondence between the perfect matchings in the linear chain of length 2 and its vertical edges.
2. (a) Draw all perfect matchings in the rectangular hexagonal system $R[3, 2]$.
   (b) Draw all perfect matchings in the parallelogram hexagonal system $P[3, 2]$.

In exercises 3 to 7 determine the number of perfect matchings contained in each of the following hexagonal systems.

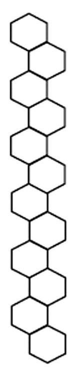
3.



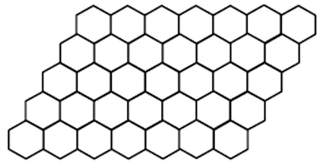
4.



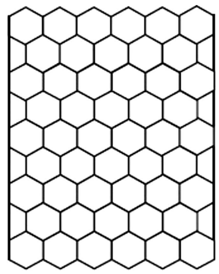
5.



6.

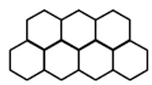


7.

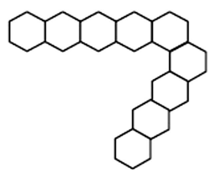


8. Determine if the given hexagonal systems contains a perfect matching. If so, determine exactly how many.

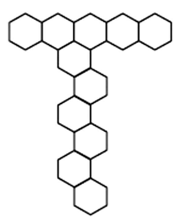
   (a)



   (b)



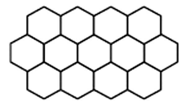
   (c)



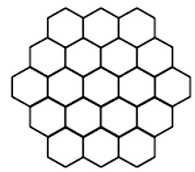
9. Use induction on the number of hexagons to show that every hexagonal system is bipartite.

In exercises 10, 11, 12 and 13 use the determinant formula to calculate the number of perfect matchings in the given hexagonal systems.

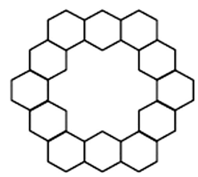
10.



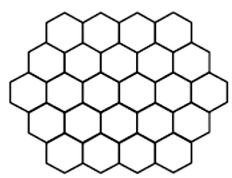
11.



12.



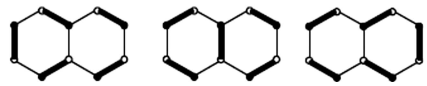
13.



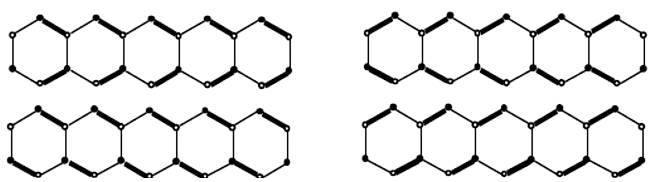
14. (a) Draw all of the perfect matchings in a fibonaccene chain of length 4, 5 and 6.
    (b) Prove Theorem 3.
15. (a) Show that a linear chain with five hexagons contains six perfect matchings.
    (b) Show that the tubulene $T[5, 1]$ satisfies $\phi(T[5, 1]) = 4$.
    (c) Show that $\phi(T[h, 1]) = 4$, for all $h \geq 2$.
16. Prove Theorem 4.
17. Show that $\lim_{h\to\infty} \kappa(T[h, k]) = 0$ and $\lim_{k\to\infty} \kappa(T[h, k]) = \frac{1}{2h}$.
18. Find $\lim_{v\to\infty} \kappa(R[7, v])$.

Selected solutions

1. Notice that each vertical edge corresponds to a unique perfect matching.



3. 11
4. 7, 776
5. 377
6. 792
7. 1, 024
8. (a) 0, (b) 29, (c) 0
9. Figure 1 shows that the result is true for a hexagonal system with one hexagon. For the inductive step remove a hexagon with edges on the boundary and consider cases defined by the number of boundary edges removed.
10. 105
11. 980
12. 130
14. (b) See [9].
15. (a) A linear chain with five hexagons contains six vertical edges, and hence, six perfect matchings. The tubulene $T[5, 1]$ contains the four perfect matchings given below (assume that the vertical edges on the ends are glued together.)



16. See [11].
17. The result follows from $\kappa(T[h, k]) = \frac{\log_2 2^{k+1}}{hk} = \frac{k+1}{hk}$.
18. $\frac{3}{13}$