MATCH
Communications in Mathematical and in Computer Chemistry

MATCH Commun. Math. Comput. Chem. **80** (2018) 153-172

ISSN 0340 - 6253

Stratified Enumeration of Convex Benzenoids

Nino Bašić*[a], Patrick W. Fowler[b], Tomaž Pisanski[a]

[a] *FAMNIT, University of Primorska, Glagoljaška 8, 6000 Koper, Slovenia*
nino.basic@famnit.upr.si, tomaz.pisanski@upr.si

[b] *Department of Chemistry, University of Sheffield, Sheffield, S3 7HF, United Kingdom*
p.w.fowler@sheffield.ac.uk

(Received October 12, 2017)

Abstract

In this paper we study an interesting family of benzenoids, the *convex benzenoids*, introduced in 2012 by Cruz, Gutman and Rada. We give several definitions of convex benzenoids and prove their equivalence. We use the partition of convex benzenoids into the 7 families of shapes: benzene, linear, triangular, trapezoidal, rhomboidal, pentagonal and hexagonal as a basis for an independent stratification and extension of the convex benzenoid enumeration by Reynolds from up to 250 hexagons presented in the OEIS as sequence A116513 to 40 000 hexagons.

1 Introduction

Benzenoids form an interesting family of graphs whose study has a strong motivation from chemistry. Benzene and benzenoids have been studied in chemistry for almost two centuries, and have contributed to the conceptual and practical develipment of organic chemistry through their specific reactivity, physical properties and biological, environmental and astrophysical importance [1–9]. Early puzzles about the structure of benzene itself gave rise to the still active debate on the meaning and quantification of the concept of aromaticity [10–13]. Benzenoids have also been studied by mathematical chemists and

*Corresponding author.

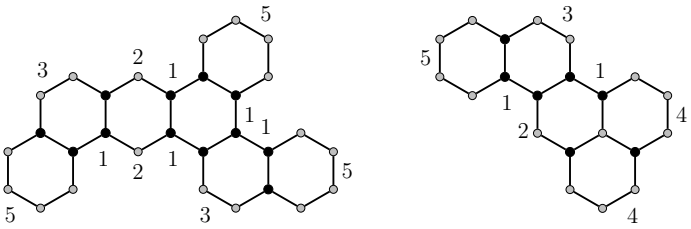
chemical graph theorists in hundreds of papers and tens of books and monographs. It is impossible to cover adequately even a small fraction of the many contributions from pioneers such as Harary, Balaban, Cyvin, Gutman, Trinajstić, Dias and others. Much of the basic theory has been magisterially surveyed in texts by Cyvin, Gutman and co-workers [14–18] and in standard texts on Chemical Graph Theory. Our concern in the present paper is with one aspect of the theory of benzenoids, and that is the relationship between properties of these molecules and their shapes. Molecular shape is of course an important factor in molecular interactions and in theories of molecuar recognition and its role in phenomena from taste and smell to carcinogenesis. One early theory of the carcinogenic properties of some benzenoids was based on the existence of "bay regions" in molecular perimeters, for example [19]. For these and many other reasons it is important to have theoretical tools for description of benzenoid shape and convexity properties. An important tool for this task is the boundary edges code, intruduced by Pierre Hansen at al. in 1996 in [20]. In particular, the code uniquely determines the benzenoid although it is ambiguous for more general hexagonal systems such as helicenes [21]. Interest in benzenoids continues unabated both amongst chemists and mathematicians, see for instance [22–30].

In 2012 an interesting family of *convex benzenoids* was introduced by Cruz et al. [31]. We take this paper as the jumping-off point for our investigation. We present here several equivalent definitions of these convex hexagonal systems and explain some of their properties. We are especially interested in the description of a convex benzenoid via its *boundary-edges code*. There has been huge activity over the past half century or more in enumeration problems for benzenoids [32]. On the particular topic of convex benzenoids, enumerations have been carried out by several researchers. On the Internet one can find a wiki [33] on *polyforms* by Scott Reynolds (a.k.a. Nekura Ca). He calls convex benzenoids *convex polyhexes*. In the OEIS [34] the sequence A116513 [35] by Allan C. Wechsler represents their enumeration. Scott Reynolds enumerated and constructed them all up to 250 hexagons. Our study independently verifies this enumeration. Furthermore, we stratify their generation into the following families of convex benzenoids: benzene, linear, triangular, trapezoidal, rhomboidal, pentagonal and hexagonal. By using this classification we are able to extend the table up to 40 000 hexagons and beyond. This paper further develops several ideas already present in the PhD thesis of the first author [36].

2 Graph theoretical preliminaries

Throughout this paper all systems (benzenoids, etc.) are considered *finite*, unless explicitly stated otherwise. We adhere to the Definition C from [14]: a *benzenoid* $\mathcal{B}$ is a set of hexagons that constitute a simply connected finite region of the infinite hexagonal grid $\mathcal{H}$ in the Euclidean plane. The boundary of the benzenoid is a simple cycle composed of *boundary vertices* and *boundary edges*. All other vertices and edges of the benzenoid, if any, are called *internal vertices* and *internal edges*, respectively. Note that a system having no internal vertices is called *catacondensed*. Other benzenoids are *pericondensed*. Two benzenoids are considered *isomorphic* if one can be obtained from the other by an isometry of the plane that leaves the infinite hexagonal grid $\mathcal{H}$ invariant.

There are several combinatorial descriptions of a benzenoid. A particular user-friendly description is the so-called *boundary-edges code* (*BEC*) popularized by Hansen et al. [20]. The benzenoid is described as a cyclic sequence of numbers, each of which counts the number of edges on the boundary between two consecutive trivalent external vertices. The exception is benzene, with BEC `6`. Any other benzenoid can be described by a sequence of numbers drawn from the set $\{1, 2, 3, 4, 5\}$. For instance, Figure 1 depicts a catacondensed benzenoid `532151153121` of 7 hexagons and a pericondensed benzenoid `5314421` of 5 hexagons. The boundary-edges code of benzenoid $\mathcal{B}$ will be denoted by $BEC(\mathcal{B})$.



(a) Catacondensed (b) Pericondensed

Figure 1. Examples of benzenoids together with corresponding numbers in boundary-edges codes.

By a *walk* in the infinite hexagonal grid $\mathcal{H}$ we mean a sequence of hexagons such that any two consecutive hexagons are either adjacent or the same. Note that a *path* is a walk that consists of pairwise distinct hexagons. For any pair of hexagons a and b in the infinite

hexagonal grid $\mathcal{H}$ define the *interval* $I_{\mathcal{H}}(a, b)$ to be the benzenoid (which turns out to be of rhombic shape) that is composed of all hexagons on any of the shortest paths in $\mathcal{H}$ from a to b.

Definition 1. *A benzenoid $\mathcal{B}$ is* convex *if for any pair of its hexagons a and b the whole interval $I_{\mathcal{H}}(a, b)$ is contained in $\mathcal{B}$.*

This definition follows readily from the fact that hexagons of $\mathcal{H}$ form a metric space and $\mathcal{B}$ can be viewed as a subspace of $\mathcal{H}$. We are therefore considering convex sets in metric spaces. This notion of convexity (also known as *geodesic convexity*) can be found in the survey on *metric graph theory* by Bandelt and Chepoi [37].

Let $a, b \in \mathcal{B}$. With $d_{\mathcal{B}}(a, b)$ we denote the distance between a and b inside $\mathcal{B}$ and $d_{\mathcal{H}}(a, b)$ is the distance between a and b with respect to $\mathcal{H}$. Note that

$$d_{\mathcal{B}}(a, b) \geq d_{\mathcal{H}}(a, b), \tag{1}$$

and that they are clearly not necessarily equal (see Figure 2).

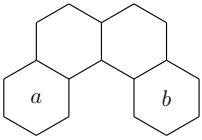


Figure 2. Benzo[c]phenanthrene $\mathcal{B}$, $d_{\mathcal{B}}(a, b) = 3 \neq d_{\mathcal{H}}(a, b) = 2$.

If, in addition, $\mathcal{K} \subseteq \mathcal{H}$, then

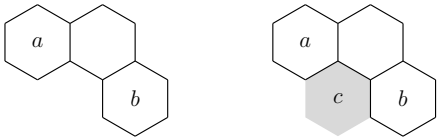
$$d_{\mathcal{H}}(a, \mathcal{K}) = \min_{k \in \mathcal{K}} d_{\mathcal{H}}(a, k). \tag{2}$$

Definition 2. *A benzenoid $\mathcal{B}$ obeys the* Small Parallelogram Rule *if for every $a, b \in \mathcal{B}$ such that $d_{\mathcal{H}}(a, b) = 2$ inclusion of some shortest path between a and b implies inclusion of the whole interval $I_{\mathcal{H}}(a, b)$, i.e.,*

$$\forall a, b \in \mathcal{B} \colon (d_{\mathcal{B}}(a, b) = 2 \implies I_{\mathcal{H}}(a, b) \subseteq \mathcal{B}).$$

We call this the *Small Parallelogram Rule* for the following reason. If a sub-benzenoid consisting of 3 hexagons and defined by the boundary-edges code `5351` (phenanthrene) is present, then the fourth hexagon that extends it to pyrene (with the boundary-edges

code 4343) is also present (see Figure 3). The case when a and b lie on the same line is trivial, as there is only one shortest path between a and b.



(a) $d_{\mathcal{B}}(a,b) = 2$ (b) $c \in I_{\mathcal{H}}(a,b) \subseteq \mathcal{B}$

Figure 3. Small Parallelogram Rule

Often it is convenient to introduce a coordinate system on $\mathcal{H}$. First, observe that the set of all edges of $\mathcal{H}$ can be partitioned into 3 classes with respect to their direction:

(a) v-edges (vertical edges),

(b) p-edges (edges with positive slope), and

(c) n-edges (edges with negative slope).

Every hexagon of $\mathcal{H}$ has six neighbouring hexagons: two v-neighbours, two p-neighbours, and two n-neighbours. Note that each $a \in \mathcal{H}$ can be assigned two integer coordinates. Consider the hexagon labeled by $(0,0)$. Its right v-neighbour has its first coordinate greater by 1, and its left v-neighbour has its first coordinate less by 1. Its upper-right n-neighbour has its second coordinate greater by 1, and the opposite n-neighbour has its second coordinate less by 1. Figure 4 displays a subregion of $\mathcal{H}$ equipped with this coordinate system.

Let $a \in \mathcal{H}$ be a hexagon of the infinite hexagonal grid $\mathcal{H}$. Let $\xi(a)$ and $\eta(a)$ denote the first and second coordinate of a. We can associate to the hexagon a another number $\zeta(a) = \xi(a) + \eta(a)$. It can be treated as another coordinate that increases/decreases with respect to the p-direction but is not independent of the other two.

Definition 3. *The sets*

$$\{a \in \mathcal{H} \mid \xi(a) = k\}, \{a \in \mathcal{H} \mid \eta(a) = k\}, \text{ and } \{a \in \mathcal{H} \mid \zeta(a) = k\}, \tag{3}$$

where $k \in \mathbb{Z}$ are called lines *in $\mathcal{H}$. We will distinguish them by their directions and call them the* line of positive slope*, the* horizontal line *and the* line of negative slope*, respectively.*

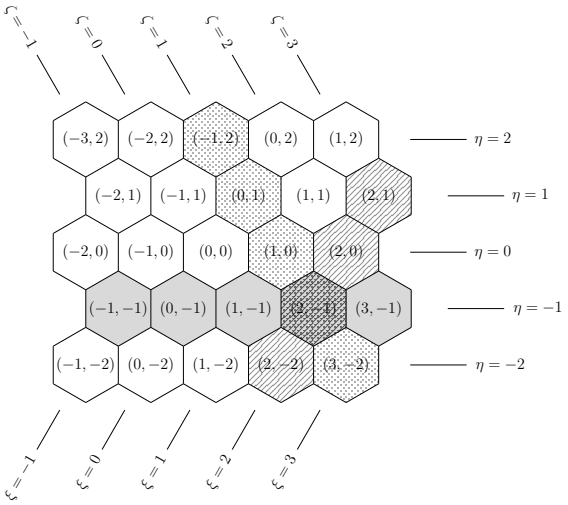


Figure 4. Coordinate system on $\mathcal{H}$.

According to Definition 3, hexagons $a, b \in \mathcal{H}$ belong to the same line if either: $\xi(a) = \xi(b)$, $\eta(a) = \eta(b)$, or $\zeta(a) = \zeta(b)$. Examples of lines are shown on Figure 4. Shaded hexagons constitute a horizontal line, hatched hexagons constitute a line of positive slope, and dotted hexagons constitute a line of negative slope. Note that any two lines with different directions intersect in exactly one hexagon and, moreover, there always exists exactly one line with the third direction that includes the intersection of the previous two.

Proposition 1. *A benzenoid system $\mathcal{B}$ is convex if and only if:*

(a) $\mathcal{B}$ is connected and

(b) $\mathcal{B}$ obeys the Small Parallelogram Rule.

Proposition 1 gives us much more concise conditions for convexity than the original Definition 1. To determine if a benzenoid is convex, one only needs to check whether it is connected and whether it obeys the Small Parallelogram Rule, which is a "local" condition. To prove the Proposition 1 we need the following lemma:

Lemma 2. *Let a benzenoid $\mathcal{B}$ satisfy conditions (a) and (b) from Proposition 1. If $a, b \in \mathcal{H}$ are on the same line then $I_{\mathcal{H}}(a, b) \subseteq \mathcal{B}$.*

Proof. Let $a, b \in \mathcal{H}$ be on the same line. Without loss of generality we can assume that $a, b \in \mathcal{L} = \{h \in \mathcal{H} \mid \xi(h) = 0\}$. Since $\mathcal{B}$ is connected, there exists a path $P = p_1\, p_2 \ldots p_l$ between $a = p_1$ and $b = p_l$. (By the definition of a path, p_i and p_{i+1} are adjacent hexagons for all $1 \leq i < l$.) Let $n = \max_{1 \leq i \leq l} d_{\mathcal{H}}(p_i, \mathcal{L})$. The proof goes by induction on n.

In the case of $n = 0$ (base of induction) the image of path P is a subset of $\mathcal{L}$. Since the image of a path is connected it follow that $I_{\mathcal{H}}(a, b) \subseteq \mathcal{B}$.

Now suppose that $n = \max_{1 \leq i \leq l} d_{\mathcal{H}}(p_i, \mathcal{L}) > 0$. By the induction hypothesis the existence of a path $Q = q_1\, q_2 \ldots q_k$ between a and b, such that $\max_{1 \leq i \leq k} d_{\mathcal{H}}(q_i, \mathcal{L}) < n$, would imply that $I_{\mathcal{H}}(a, b) \subseteq \mathcal{B}$. Let r be the smallest index such that $d_{\mathcal{H}}(p_{r-1}, \mathcal{L}) = n - 1$ and $d_{\mathcal{H}}(p_r, \mathcal{L}) = n$ and let s be the smallest index such that $d_{\mathcal{H}}(p_s, \mathcal{L}) = n$, $d_{\mathcal{H}}(p_{s+1}, \mathcal{L}) = n - 1$ and $s \geq r$. This situation is depicted in Figure 5. This subpath may look like

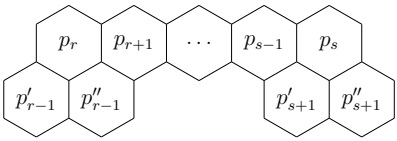


Figure 5. Subpath $p_{r-1}\, p_r \ldots p_{s+1}$ of path P.

$p'_{r-1}\, p_r \ldots p_s\, p'_{s+1}$, $p''_{r-1}\, p_r \ldots p_s\, p'_{s+1}$, $p'_{r-1}\, p_r \ldots p_s\, p''_{s+1}$, or $p''_{r-1}\, p_r \ldots p_s\, p''_{s+1}$. If $r = s$ then by deletion of p_r from the path P we obtain a walk (which can be further reduced to a path) that has one hexagon less at distance n from $\mathcal{L}$ than does the original path P. If $r < s$ then we have two options:

(a) if $p_{r-1} = p''_{r-1}$ we can erase p_r from P since p_{r-1} and p_{r+1} are adjacent;

(b) if $p_{r-1} = p'_{r-1}$ then the Small Parallelogram Rule implies that $p''_{r-1} \in \mathcal{B}$ so we can replace p_r with p''_{r-1} in path P.

In both cases we obtain a path with one hexagon less at distance n from $\mathcal{L}$. If we iteratively apply this procedure, we obtain the path Q with desired property. ■

Proof of Proposition 1. If $\mathcal{B}$ is convex then it is clear that $\mathcal{B}$ is connected and that $\mathcal{B}$ obeys the Small Parallelogram Rule.

Let us prove the other direction. Choose arbitrary hexagons $a, b \in \mathcal{B}$. We have to show that $I_{\mathcal{H}}(a, b) \subseteq \mathcal{B}$. The interval between two hexagons is always of rhombic shape. Without loss of generality (thanks to the symmetries of $\mathcal{H}$) we can assume that a is in

the lower left corner of the rhombus and b is in the upper right corner (as depicted on Figure 6). Condition (a) implies the existence of a path P in $\mathcal{B}$ between a and b. An example of such a path is shown in Figure 6.

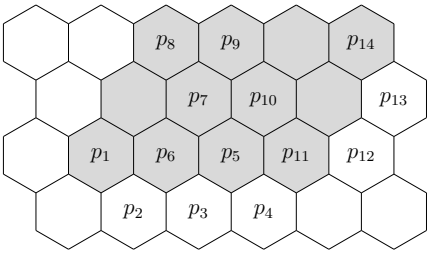


Figure 6. Interval $I(a = p_1, b = p_{14})$

First, we show that we can assume that the path $P = p_1\, p_2 \ldots p_l$ lies entirely in the half-plane $\{h \in \mathcal{H} \mid \eta(h) \geq \eta(a)\}$. Let r be the smallest index such that $\eta(p_r) < \eta(a)$. Note that $r = 2$ in the example in Figure 6. Let s be the largest index such that $\eta(p_s) < \eta(a)$. Note that $s = 4$ in the example in Figure 6. If such indices do not exist, there is nothing left to prove. If such indices exist, then $1 < r \leq s < l$. Moreover, $\eta(p_{r-1}) = \eta(a) = \eta(p_{s+1})$. So p_{r-1} and p_{s+1} are on the same line and $I_{\mathcal{H}}(p_{r-1}, p_{s+1}) \subseteq \mathcal{B}$ by Lemma 2. So the subpath $p_r \ldots p_s$ can be replaced with the shortest (p_{r-1}, p_{s+1})-path. Similarly, we can assume that P lies entirely in the half-plane $\{h \in \mathcal{H} \mid \eta(h) \leq \eta(b)\}$, in half-plane $\{h \in \mathcal{H} \mid \xi(h) \geq \xi(a)\}$, and in half-plane $\{h \in \mathcal{H} \mid \xi(h) \leq \xi(b)\}$. Therefore $P \subseteq I_{\mathcal{H}}(a, b)$.

So far, we know that there is a path P in $\mathcal{B}$ connecting a to b that lies entirely in $I_{\mathcal{H}}(a, b)$. Recall our assumption that a is in its lower-left corner and b is in its upper-right corner. Now, we will show that there also exists a path $P' = p'_1 p'_2 \ldots p'_m$ in $\mathcal{B}$ with the special property that p'_{i+1} is the right v-neighbour or the upper-right n-neighbour of p'_i for each i, i.e., in each step either the η coordinate increases or the ξ coordinate increases.

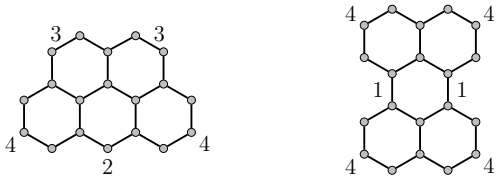
The path P' can be obtained from the path P with the help of Lemma 2. Find the smallest index i, such that either $\xi(p_{i+1}) < \xi(p_i)$ or $\eta(p_{i+1}) < \eta(p_i)$. Without loss of generality we can assume that $\xi(p_{i+1}) < \xi(p_i)$. In each step of the path the ξ coordinate increases by 1, decreases by 1 or stays the same. The path has to reach the upper-right corner eventually, so there exists an index $j > i$, such that $\xi(p_j) = \xi(p_i)$. By Lemma 2, the subpath $p_{i+1}, \ldots, p_{j-1}$ can be replaced with a straight line segment from p_i to p_j that

lies in $\mathcal{B}$. We can use the above operation iteratively until we obtain the path P' with the desired property.

The path P' that we obtained can be described as a sequence of symbols $\rightarrow$ and $\nearrow$ which indicate the increases in the ξ and η coordinates, respectively. One such path in the example from Figure 6 is $(\rightarrow, \nearrow, \rightarrow, \rightarrow, \nearrow)$. Every possible shortest path in $I_{\mathcal{H}}(a, b)$ can be described as a permutation of these symbols. The Small Parallelogram Rule implies that we may swap two consecutive symbols in the sequence and obtain another path which also lies in $\mathcal{B}$. From $(\rightarrow, \nearrow, \rightarrow, \rightarrow, \nearrow)$ we can obtain, say, $(\rightarrow, \rightarrow, \nearrow, \rightarrow, \nearrow)$. It is clear that any permutation can be obtained in this way, which implies that the whole interval $I_{\mathcal{H}}(a, b)$ is contained in $\mathcal{B}$. ■

Proposition 3. *A benzenoid $\mathcal{B}$ is* convex *if and only if its boundary-edges code $BEC(\mathcal{B})$ does not contain the symbol* `1`.

Example 4. The distinction between convex and non-convex benzenoids is apparent from their boundary-edges codes (see Figure 7). For instance, the boundary-edges code of the benzenoid $\mathcal{B}_1$ on Figure 7(a) is `24334` and does not contain symbol `1`.



(a) $BEC(\mathcal{B}_1) =$ `24334`, convex (b) $BEC(\mathcal{B}_2) =$ `144144`, non-convex

Figure 7. The distinction between convex and non-convex benzenoids is readily apparent from their boundary-edges codes.

Therefore this benzenoid is convex, whilst, the boundary-edges code for the benzenoid in Figure 7(b) is `144144`, containing a `1`. Hence the benzenoid itself is non-convex. ■

Proof. Let $\mathcal{B}$ be a convex benzenoid. We will prove that there is no `1` in its boundary-edges code. Suppose that there is a symbol `1`. Then that part of the boundary of $\mathcal{B}$ that corresponds to the symbol `1` in the boundary-edges code looks locally as shown in Figure 8.

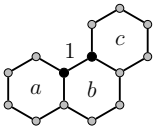


Figure 8. The part of boundary corresponding to a symbol 1 in the boundary-edges code.

There exist three hexagons a, b and c that are positioned as indicated in the Figure 8. This means that the Small Parallelogram Rule is not obeyed, which contradicts the claim that $\mathcal{B}$ is convex. Thus, there can be no 1 in the boundary-edges code.

Let $\mathcal{B}$ be a benzenoid such that there is no symbol 1 in the boundary-edges code of $\mathcal{B}$. Assume that $\mathcal{B}$ is not convex. By Proposition 1, it does not obey the Small Parallelogram Rule. There must exist three hexagons in $\mathcal{B}$ which are positioned in a phenanthrene shape as shown in Figure 3(b) and where the hexagon labeled with c is not present (see Figure 3(b)). Then the corresponding part of the boundary contains two consecutive degree-3 vertices, i.e., there is a 1 in its boundary-edges code. This is a contradiction, thus benzenoid $\mathcal{B}$ is convex. ■

3 Finite and infinite benzenoids

In addition to finite benzenoids we will need (for the purpose of theoretical reasoning) some infinite systems of hexagons:

$$\mathcal{HP}^{+}_{\xi}(n) = \{h \in \mathcal{H} \mid \xi(h) \geq n\}, \tag{4}$$

$$\mathcal{HP}^{-}_{\xi}(n) = \{h \in \mathcal{H} \mid \xi(h) \leq n\}, \tag{5}$$

$$\mathcal{HP}^{+}_{\eta}(n) = \{h \in \mathcal{H} \mid \eta(h) \geq n\}, \tag{6}$$

$$\mathcal{HP}^{-}_{\eta}(n) = \{h \in \mathcal{H} \mid \eta(h) \leq n\}, \tag{7}$$

$$\mathcal{HP}^{+}_{\zeta}(n) = \{h \in \mathcal{H} \mid \zeta(h) \geq n\}, \tag{8}$$

$$\mathcal{HP}^{-}_{\zeta}(n) = \{h \in \mathcal{H} \mid \zeta(h) \leq n\}. \tag{9}$$

The infinite hexagonal systems defined in (4) to (9) will be called *half-planes*.

Proposition 5. *Any intersection of convex (finite or infinite) benzenoids is a convex benzenoid.*

Proof. Let $\mathcal{B}_1$ and $\mathcal{B}_2$ be two convex benzenoids. If their intersection is $\emptyset$ then there is nothing to prove. Suppose that $\mathcal{B}_1 \cap \mathcal{B}_2 \neq \emptyset$. Let $a, b \in \mathcal{B}_1 \cap \mathcal{B}_2$. As $\mathcal{B}_1$ is convex, we have

that $I_{\mathcal{H}}(a,b) \subseteq \mathcal{B}_1$. Similarly, as $\mathcal{B}_2$ is convex we have that $I_{\mathcal{H}}(a,b) \subseteq \mathcal{B}_2$. But this means that $I_{\mathcal{H}}(a,b) \subseteq \mathcal{B}_1 \cap \mathcal{B}_2$, which completes the proof. ■

Proposition 6. *A benzenoid $\mathcal{B}$ is convex if and only if it can be obtained as an intersection of half-planes, i.e., if there exist integers n_ξ^+, n_ξ^-, n_η^+, n_η^-, n_ζ^+ and n_ζ^-, such that*

$$\mathcal{B} = \mathcal{HP}_\xi^+(n_\xi^+) \cap \mathcal{HP}_\xi^-(n_\xi^-) \cap \mathcal{HP}_\eta^+(n_\eta^+) \cap \mathcal{HP}_\eta^-(n_\eta^-) \cap \mathcal{HP}_\zeta^+(n_\zeta^+) \cap \mathcal{HP}_\zeta^-(n_\zeta^-). \quad (10)$$

Proof. If a benzenoid $\mathcal{B}$ is obtained as an intersection of half-planes then it is convex by Proposition 5.

Now let $\mathcal{B}$ be a convex benzenoid. We will show that it is equal to the intersection of 6 half-planes.

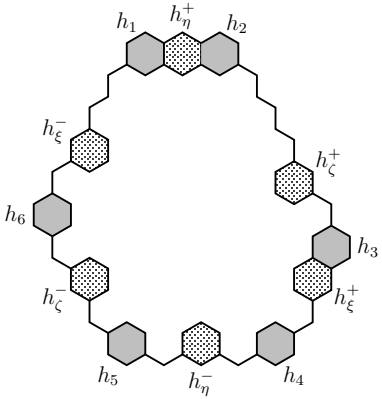


Figure 9. A convex benzenoid $\mathcal{B}$ from the proof of Proposition 6.

Let h_η^+ and h_η^- be hexagons of $\mathcal{B}$ (see Figure 9) such that

$$\eta(h_\eta^+) = \max_{h \in \mathcal{B}} \eta(h) \quad \text{and} \quad \eta(h_\eta^-) = \min_{h \in \mathcal{B}} \eta(h). \quad (11)$$

Similarly, let h_ξ^+, h_ξ^-, h_ζ^+ and h_ζ^- be hexagons of $\mathcal{B}$ such that

$$\xi(h_\xi^+) = \max_{h \in \mathcal{B}} \xi(h), \quad \xi(h_\xi^-) = \min_{h \in \mathcal{B}} \xi(h), \quad \eta(h_\zeta^+) = \max_{h \in \mathcal{B}} \zeta(h), \quad \eta(h_\zeta^-) = \min_{h \in \mathcal{B}} \zeta(h). \quad (12)$$

Let $h_1, h_2, \ldots, h_6$ be hexagons of $\mathcal{H}$ such that:

(i) $\eta(h_1) = \eta(h_\eta^+)$, $\xi(h_1) = \xi(h_\xi^-)$;

(ii) $\eta(h_2) = \eta(h_\eta^+)$, $\zeta(h_2) = \zeta(h_\zeta^+)$;

(iii) $\zeta(h_3) = \zeta(h_\zeta^+)$, $\xi(h_3) = \xi(h_\xi^+)$;

(iv) $\eta(h_4) = \eta(h_\eta^-)$, $\xi(h_4) = \xi(h_\xi^+)$;

(v) $\eta(h_5) = \eta(h_\eta^-)$, $\zeta(h_5) = \zeta(h_\zeta^-)$;

(vi) $\zeta(h_6) = \zeta(h_\zeta^-)$, $\xi(h_6) = \xi(h_\xi^-)$.

As $h_1 \in I_{\mathcal{H}}(h_\eta^+, h_\xi^-)$ it follows that $h_1 \in \mathcal{B}$. By analogy, $h_2, \ldots, h_6 \in \mathcal{B}$. Note that h_i and h_{i+1} (indices taken modulo 6) are on the same line, and thus the line segment between h_i and h_{i+1} is also contained in $\mathcal{B}$. All hexagons inside the 'polygon' defined by $h_1, \ldots, h_6$ must belong to $\mathcal{B}$ or $\mathcal{B}$ would not be simply connected. If any hexagon outside the polygon belonged to $\mathcal{B}$, this would contradict one of equations (11) or (12). It follows that

$$\mathcal{B} = \mathcal{HP}_\xi^+(n_\xi^-) \cap \mathcal{HP}_\xi^-(n_\xi^+) \cap \mathcal{HP}_\eta^+(n_\eta^-) \cap \mathcal{HP}_\eta^-(n_\eta^+) \cap \mathcal{HP}_\zeta^+(n_\zeta^-) \cap \mathcal{HP}_\zeta^-(n_\zeta^+). \qquad (13)$$

■

In a similar manner, by using Proposition 6, the finite convex benzenoids can be classified into six mutually disjoint families that we call *fundamental shape families of convex benzenoids*:

(a) The *linear chain* $L(n)$, $n \geq 2$, having n hexagons. The corresponding boundary-edges code is

$$52^{n-2}52^{n-2}.$$

(b) The *equilateral triangle* $T_3(n)$, $n \geq 2$, having n hexagons on each of its sides. The corresponding boundary-edges code is

$$42^{n-2}42^{n-2}42^{n-2}.$$

(c) The *equilateral trapezoid* $T_4(n, m)$, $n > m \geq 2$, having n hexagons on the bottom base and m hexagons on the top base. The corresponding boundary-edges code is

$$42^{n-2}42^{n-m-1}32^{m-2}32^{n-m-1}.$$

(d) The *rhomboid* $R(n, m)$, $m \geq n \geq 2$, having n hexagons on its base and m hexagons on its side. The corresponding boundary-edges code is

$$42^{n-2}32^{m-2}42^{n-2}32^{m-2}.$$

(e) The *pentagonal benzenoid* $P(n, m, k)$, $n \geq 2$, $k \geq m \geq 2$ having n hexagons on the base, m hexagons on its left side and k hexagons on its right side. The corresponding boundary-edges code is

$$32^{n-2}32^{k-1}32^{m+n-3}42^{n+k-3}32^{m-2}.$$

(f) The *hexagonal benzenoid* $H(n, m, k, t)$, $n \geq 2$, $k \geq m \geq 2$, $n + m - 2 \geq t \geq 2$, having n hexagons on the base, m hexagons on its left side, k hexagons on its right side and t hexagons on its top side. The corresponding boundary-edges code is

$$32^{n-2}32^{m-2}32^{n+k-t-2}32^{t-2}32^{n+m-t-2}32^{k-2}.$$

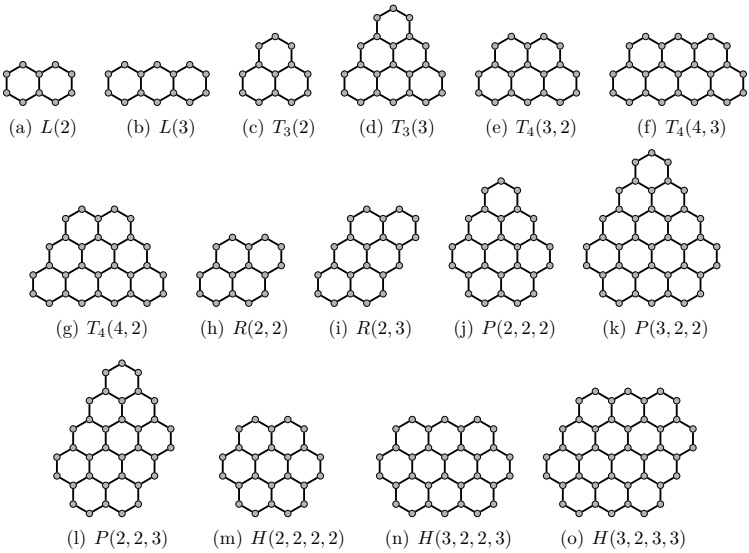


Figure 10. The smallest members of families of convex benzenoids.

Note that an exponent n in the boundary-edges code means that the corresponding symbol repeats n times, e.g. $2^5 = 22222$. This is a compact way of writing the code. Benzene is also a convex benzenoid and is not included in any of the above families. We treat it as a separate sporadic case. Note that this classification of families (a) – (f) is already present in [31], where it is used to characterise extremal benzenoids with respect to a variety of vertex-degree-based topological indices.

In all of the above families, except for $H(n,m,k,t)$, the benzenoid is uniquely determined by the described parameters. Note that

$$\begin{aligned} H(n,m,k,t) &\cong H(n+m-t, \min\{k,t\}, \max\{k,t\}, m) \\ &\cong H(n+k-t, \min\{m,t\}, \max\{m,t\}, k) \\ &\cong H(t, n+m-t, n+k-t, n) \\ &\cong H(m, \min\{n, n+k-t\}, \max\{n, n+k-t\}, n+m-t) \\ &\cong H(k, \min\{n, n+m-t\}, \max\{n, n+m-t\}, n+k-t). \end{aligned} \tag{14}$$

The above parametrisations are all valid and represent equivalent benzenoids. The set of parameters that is lexicographically the smallest of the above is called the *canonical parametrisation*. We obtain the following proposition:

Proposition 7. *A benzenoid $\mathcal{B}$ is convex if and only if it is benzene or is equivalent to a member of one of the families: $L(n)$, $T_3(n)$, $T_4(n)$, $R(n,m)$, $P(n,m,k)$ or $H(n,m,k,t)$.* ■

Although the proofs of equivalent definitions of convex benzenoids are easy, each of them represents a different viewpoint which, in turn, may simplify proof of difficult results about convex benzenoids.

From Proposition 7 we can obtain a simple algorithm that can count (and enumerate) all convex benzenoids with $h \leq N$ hexagons for a given N. We iterate over all families and over all admissible parameters for which the number of hexagons remains below N. (It is easy to derive an expression for the number of hexagons in terms of the parameters for each of the families of convex benzenoids.) For each benzenoid we determine the number of hexagons, h. Some care should be taken with the hexagonal family, since we have to check whether the given parameters are indeed canonical.

By using the above procedure, we were able to extend the sequence A116513 [35] up to 40 000. Our computer programme took 646 seconds on a single Intel Core i5-4200U CPU running at 1.60 GHz. The numbers of convex benzenoids with h hexagons for $h \leq 50$ are given in the Table 1. Benzene does not belong to any of the six shape families but it is accounted for in the last column which gives the total number. The graph in Figure 11 shows the number of convex benzenoids with at most h hexagons as a function of h. Some of those values can also be found in Table 2.

h	L	T_3	R	T_4	P	H	Σ
1	0	0	0	0	0	0	1
2	1	0	0	0	0	0	1
3	1	1	0	0	0	0	2
4	1	0	1	0	0	0	2
5	1	0	0	1	0	0	2
6	1	1	1	0	0	0	3
7	1	0	0	1	0	1	3
8	1	0	1	0	1	0	3
9	1	0	1	2	0	0	4
10	1	1	1	0	0	1	4
11	1	0	0	1	1	0	3
12	1	0	2	1	0	1	5
13	1	0	0	1	1	1	4
14	1	0	1	1	1	1	5
15	1	1	1	2	1	0	6
16	1	0	2	0	0	2	5
17	1	0	0	1	2	0	4
18	1	0	2	2	0	2	7
19	1	0	0	1	2	2	6
20	1	0	2	1	1	1	6
21	1	1	1	2	1	1	7
22	1	0	1	1	1	2	6
23	1	0	0	1	2	2	6
24	1	0	3	1	2	2	9
25	1	0	1	2	1	2	7

h	L	T_3	R	T_4	P	H	Σ
26	1	0	1	1	2	2	7
27	1	0	1	3	2	1	8
28	1	1	2	0	0	4	8
29	1	0	0	1	4	2	8
30	1	0	3	3	1	2	10
31	1	0	0	1	1	3	6
32	1	0	2	0	3	2	8
33	1	0	1	3	2	4	11
34	1	0	1	1	3	4	10
35	1	0	1	3	3	1	9
36	1	1	4	1	1	4	12
37	1	0	0	1	2	3	7
38	1	0	1	1	2	5	10
39	1	0	1	3	5	2	12
40	1	0	3	1	0	5	10
41	1	0	0	1	4	2	8
42	1	0	3	3	2	4	13
43	1	0	0	1	3	6	11
44	1	0	2	1	4	4	12
45	1	1	2	4	2	3	13
46	1	0	1	1	2	5	10
47	1	0	0	1	4	4	10
48	1	0	4	1	3	6	15
49	1	0	1	2	4	4	12
50	1	0	2	2	3	5	13

Table 1. The enumeration of convex benzenoids up to $h = 50$.

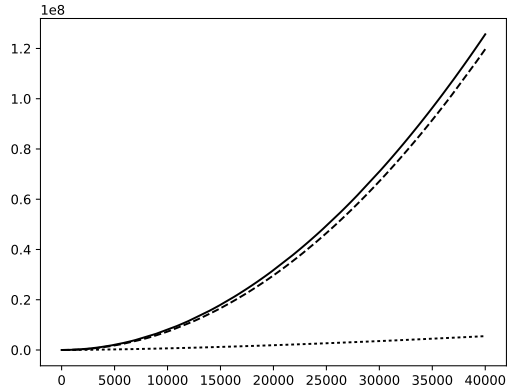


Figure 11. Number of convex benzenoids with at most h hexagons. The dotted line represents the pentagonal family, the dashed line represents the hexagonal family and the solid line is the sum of all families.

h	L	T_3	R	T_4	P	H	Σ
1	0	0	0	0	0	0	1
2	1	0	0	0	0	0	2
3	2	1	0	0	0	0	4
4	3	1	1	0	0	0	6
5	4	1	1	1	0	0	8
10	9	3	5	4	1	2	25
20	19	4	15	15	10	12	76
30	29	6	28	30	26	32	152
40	39	7	42	45	48	65	247
50	49	8	57	62	79	108	364
100	99	12	146	163	307	512	1240
200	199	18	356	398	1092	2309	4373
300	299	23	592	664	2229	5476	9284
400	399	26	844	944	3645	10035	15894
500	499	30	1106	1239	5315	16002	24192
1000	999	43	2550	2836	16819	67397	90645
2000	1999	61	5781	6388	51640	279617	345487
3000	2999	75	9275	10207	98602	639352	760511
4000	3999	87	12934	14200	155367	1147373	1333961
5000	4999	98	16723	18318	220754	1804617	2065510
10000	9999	139	36884	40153	651515	7337158	8075849
20000	19999	198	80659	87311	1902546	29689929	31780643
30000	29999	243	127049	137110	3547707	67144688	70986797
40000	39999	281	175131	188604	5512472	119731633	125648121

Table 2. The total number of convex benzenoids with at most h hexagons for certain values of h.

4 Conclusion

We should point out that benzenoids have been enumerated in the past, see for instance [38–44]. The main novelty in our approach lies in the fact that convex benzenoids are very easy to describe and that there is no need to check isomorphism. This is why we were able to carry out our enumeration to large numbers of hexagons. Another advantage of convex benzenoids is the fact that they have very easy boundary edges code descriptions, and we benefit from the fact that this code is both compact and uniquely describes all benzenoids [21]. For some of families of benzenoids there is a closed-form enumeration formula. We conjecture that this could be extended to the other families, eventually yielding a closed-form solution to the enumeration problem of convex benzenoids.

Acknowledgements: The authors would like to thank anonymous referees for invaluable comments that made this work more readable. Research of T. P. and N. B. is supported in part by the ARRS Grant P1-0294.

References

[1] E. Clar, *Polycyclic Hydrocarbons, Vol. 1*, Springer, Berlin, 1964.

[2] E. Clar, *Polycyclic Hydrocarbons, Vol. 2*, Springer, Berlin, 1964.

[3] J. R. Dias, *Handbook of Polycyclic Hydrocarbons. Part A: Benzenoid Hydrocarbons*, Elsevier, Amsterdam, 1987.

[4] J. R. Dias, *Handbook of Polycyclic Hydrocarbons. Part B: Polycyclic Isomers and Heteroatom Analogs of Benzenoid Hydrocarbons*, Elsevier, Amsterdam, 1988.

[5] J. N. Murrell, *The Theory of the Electronic Spectra of Organic Molecules*, Methuen, London, 1963.

[6] D. Lloyd, *The Chemistry of Conjugated Cyclic Compounds: To Be or Not to Be Like Benzene?* Wiley, New York, 1989.

[7] H. V. Gelboin, P. O. P. Ts'o, (Eds.) *Polycyclic Hydrocarbons and Cancer, Vol. 1: Environment, Chemistry, and Metabolism*, Acad. Press, New York, 1978.

[8] H. I. Abdel-Shafy, M. S. M. Mansour, A review on polycyclic aromatic hydrocarbons: Source, environmental impact, effect on human health and remediation, *Egyptian J. Petrol.* **25** (2016) 107–123.

[9] L. J. Allamandola, A. G. G. M. Tielens, J. R. Barker, Interstellar polycyclic aromatic hydrocarbons: the infrared emission bands, the excitation/emission mechanism, and the astrophysical implications, *Astrophys. J. Suppl. Ser.* **71** (1989) 733–775.

[10] P. von Ragué Schleyer (Ed.), Special Issue: Aromaticity, *Chem. Rev.* **101** (2001) 1115–1566, 19 papers by various authors.

[11] P. von Ragué Schleyer (Ed.), Special Issue: Delocalization–Pi and Sigma, *Chem. Rev.* **105** (2005) 3433–3947, 16 papers by various authors.

[12] G. Merino, M. Solà (Eds.), Themed Collection: Electron delocalization and aromaticity: 150 years of the Kekulé benzene structure, *Phys. Chem. Chem. Phys.* **18** (2016) 11587–11950, 40 papers by various authors.

[13] P. W. Fowler, E. Steiner, L. W. Jenneskens (Eds.), Themed Collection: New Perspectives on Aromaticity, *Phys. Chem. Chem. Phys.* **6** (2004) 217–339, 19 papers by various authors.

[14] I. Gutman, S. J. Cyvin, *Introduction to the Theory of Benzenoid Hydrocarbons*, Springer, Berlin, 1989.

[15] N. Trinajstić, *Chemical Graph Theory*, CRC Press, Boca Raton, 1992.

[16] S. J. Cyvin, I. Gutman, *Kekulé Structures in Benzenoid Hydrocarbons*, Springer, Berlin, 1988.

[17] S. J. Cyvin, J. Brunvoll, B. N. Cyvin, *Theory of Coronoid Hydrocarbons*, Springer, Berlin, 1991.

[18] S. J. Cyvin, J. Brunvoll, R. S. Chen, B. N. Cyvin, F. J. Zhang, *Theory of Coronoid Hydrocarbons II*, Springer, Berlin, 1994.

[19] M. Randić, M. Novič, D. Plavšić, *Solved and Unsolved Problems of Structural Chemistry*, CRC Press, Boca Raton, 2016, see for example the discussion of "Problem 12" in the book.

[20] P. Hansen, C. Lebatteux, M. Zheng, The boundary-edges code for polyhexes, *J. Mol. Struct. (Theochem)* **363** (1996) 237–247.

[21] X. Guo, P. Hansen, M. Zheng, Boundary uniqueness of fusenes, *Discr. Appl. Math.* **118** (2002) 209–222.

[22] J. Rada, R. Cruz, I. Gutman, Vertex–degree–based topological indices of catacondensed hexagonal systems, *Chem. Phys. Lett.* **572** (2013) 154–157.

[23] J. Rada, R. Cruz, I. Gutman, Benzenoid systems with extremal vertex–degree–based topological indices, *MATCH Commun. Math. Comput. Chem.* **72** (2014) 125–136.

[24] A. Kelenc, S. Klavžar, N. Tratnik, The edge-Wiener index of benzenoid systems in linear time, *MATCH Commun. Math. Comput. Chem.* **74** (2015) 521–532.

[25] K. Salem, S. Klavžar, A. Vesel, P. žigert, The Clar formulas of a benzenoid system and the resonance graph, *Discr. Appl. Math.* **157** (2009) 2565–2569.

[26] S. Klavžar, M. Kovše, The lattice dimension of benzenoid systems, *MATCH Commun. Math. Comput. Chem.* **56** (2006) 637–648.

[27] K. Salem, S. Klavžar, I. Gutman, On the role of hypercubes in the resonance graphs of benzenoid graphs, *Discr. Math.* **306** (2006) 699–704.

[28] S. Klavžar, P. Žigert, A min-max result on catacondensed benzenoid graphs, *Appl. Math. Lett.* **15** (2002) 279–283.

[29] I. Gutman, S. Klavžar, M. Petkovšek, P. Žigert, On Hosoya polynomials of benzenoid graphs, *MATCH Commun. Math. Comput. Chem.* (2001) 49–66.

[30] V. Chepoi, S. Klavžar, Distances in benzenoid systems: further developments, *Discr. Math.* **192** (1998) 27–39.

[31] R. Cruz, I. Gutman, J. Rada, Convex hexagonal systems and their topological indices, *MATCH Commun. Math. Comput. Chem.* **68** (2012) 97–108.

[32] B. N. Cyvin, J. Brunvoll, S. J. Cyvin, Enumeration of benzenoid systems and other polyhexes, in: I. Gutman (Ed.), *Advances in the Theory of Benzenoid Hydrocarbons II*, Springer, Berlin, 1992.

[33] N. Ca, Convex Block — Polyform Wiki, `http://polyform.wikia.com/wiki/Convex_Block#Polyhexes`, [Online; accessed: 1-February-2017].

[34] O. F. I. (2011), The On-Line Encyclopedia of Integer Sequences, `http://oeis.org`, [Online; accessed: 1-February-2017].

[35] A. C. Wechsler, Sequence A116513 in The On-Line Encyclopedia of Integer Sequences (2011), `http://oeis.org`, [Online; accessed: 1-February-2017].

[36] N. Bašić, *Algebraic approach to several families of chemical graphs*, Ph.D. thesis, Univ. Ljubljana, Ljubljana, 2016.

[37] H. J. Bandelt, V. Chepoi, Metric graph theory and geometry: a survey, `http://pageperso.lif.univ-mrs.fr/~victor.chepoi/survey_cm_bis.pdf`, [Online; accessed: 1-February-2017].

[38] A. T. Balaban, F. Harary, Chemical graphs – V: Enumeration and proposed nomenclature of benzenoid cata-condensed polycyclic aromatic hydrocarbons, *Tetrahedron* **24** (1968) 2505–2516.

[39] F. Harary, R. C. Read, The enumeration of tree-like polyhexes, *Proc. Edinburgh Math. Soc.* **17** (1970) 1–13.

[40] K. Balasubramanian, J. J. Kaufman, W. S. Koski, A. T. Balaban, Graph theoretical characterization and computer-generation of certain carcinogenic benzenoid hydrocarbons and identification of bay regions, *J. Comput. Chem.* **1** (1980) 149–157.

[41] J. V. Knop, K. Szymanski, Ž. Jeričević, N. Trinajstić, Computer enumeration and generation of benzenoid hydrocarbons and identification of bay regions, *J. Comput. Chem.* **4** (1983) 23–32.

[42] A. T. Balaban, C. Artemi, Chemical graphs. Part 51. Enumeration of nonbranched catafusenes according to the numbers of benzenoid rings in the catafusene and in its longest linearly condensed portion, *Polycyc. Arom. Comp.* **1** (1990) 171–189.

[43] W. R. Müller, K. Szymanski, J. V. Knop, S. Nikolić, N. Trinajstić, On the enumeration and generation of polyhex hydrocarbons, *J. Comput. Chem.* **11** (1990) 223–235.

[44] S. Nikolić, N. Trinajstić, J. V. Knop, W. R. Müller, K. Szymanski, On the classification and enumeration of planar polyhex hydrocarbons, *J. Mol. Struct. (Theochem)* **77** (1991) 219–225.