This is a repository copy of *Convexity deficit of benzenoids*.

White Rose Research Online URL for this paper:
http://eprints.whiterose.ac.uk/162861/

Version: Published Version

**Article:**
Bašić, N., Fowler, P.W. orcid.org/0000-0003-2106-1104, Zemljić, S.S. et al. (6 more authors) (2020) Convexity deficit of benzenoids. Croatica chemica acta, 92 (4). pp. 457-466. ISSN 0011-1643

https://doi.org/10.5562/cca3602

**Reuse**
This article is distributed under the terms of the Creative Commons Attribution (CC BY) licence. This licence allows you to distribute, remix, tweak, and build upon the work, even commercially, as long as you credit the authors for the original work. More information and the full terms of the licence here:
https://creativecommons.org/licenses/

**Takedown**
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.
Convexity Deficit of Benzenoids

Nino Bašić,1,2,3 Sarah J. Berkemer,4 Jörg Fallmann,6 Patrick W. Fowler,6 Thomas Gatter,6 Tomáš Pisanski,1,2,3,12,* Nancy Retzlaff,4,5,7 Peter F. Stadler,4,5,8,9,10 Sara Sabrina Žemličkova1,11

1 FAMNIT, University of Primorska, Koper, Slovenia
2 Institute of Mathematics, Physics and Mechanics, Ljubljana, Slovenia
3 IAM, University of Primorska, Koper, Slovenia
4 Bioinformatics Group, Department for Computer Science, and Interdisciplinary Center for Bioinformatics, Leipzig University, Germany
5 Max-Planck-Institute for Mathematics in the Sciences, Leipzig, Germany
6 Department of Chemistry, University of Sheffield, Sheffield S3 7HF, UK
7 Institute for Infrastructure and Resources Management, Leipzig University, Germany
8 Institute for Theoretical Chemistry, University of Vienna, Austria
9 Facultad de Ciencias, Universidad Nacional de Colombia, Sede Bogotá, Colombia
10 Santa Fe Institute, Santa Fe, USA
11 Faculty of Mathematics, Physics and Informatics, Comenius University, Bratislava, Slovakia
12 Faculty of Mathematics and Physics, University of Ljubljana, Ljubljana, Slovenia

* Corresponding author’s e-mail address: pisanski@upr.si

RECEIVED: December 2, 2019 • REVISED: March 20, 2020 • ACCEPTED: March 21, 2020

PROCEEDING OF THE 31ST MATH/CHEM/COMP CONFERENCE, JUNE 10–14, 2019, DUBROVNIK, CROATIA

THIS PAPER IS DEDICATED TO THE MEMORY OF DR. EDWARD C. KIRBY (1934–2019)

COLLEAGUE, COLLABORATOR AND FRIEND OF SEVERAL OF THE AUTHORS

Abstract: In 2012, a family of benzenoids was introduced by Cruz, Gutman, and Rada, which they called convex benzenoids. In this paper we introduce the convexity deficit, a new topological index intended for benzenoids and, more generally, fusenes. This index measures by how much a given fusene departs from convexity. It is defined in terms of the boundary-edges code. In particular, convex benzenoids are exactly the benzenoids having convexity deficit equal to 0. Quasi-convex benzenoids form the family of non-convex benzenoids that are closest to convex, i.e., they have convexity deficit equal to 1. Finally, we investigate convexity deficit of several important families of benzenoids.

Keywords: Benzenoid, fusene, convexity deficit, convex benzenoid, quasi-convex benzenoid, pseudo-convex benzenoid.

1. INTRODUCTION

Benzenoids form an important family of graphs and molecules. Polycyclic (aromatic) hydrocarbons,5,10,15,20 of which the benzenoids form a subset, are important molecular systems with a rich organic chemistry23,39 characterised by specific reactivity,14,38 spectra,138 and photophysics. They occur naturally, geologically and as by-products of natural and anthropogenic combustion processes, with considerable implications for the environment11 and human health27 and have been postulated as significant contributors to the carbon inventory in the wider Universe.17 There has also been a huge amount of interest over several decades in the graph theory of benzenoids and related structures and its application to prediction of physical and chemical properties (see, e.g., the textbooks12,13,16,26,30,40). Much of the mathematical chemistry literature is concerned with prediction or rationalisation of electronic structure, but there is also interest in classification of the shapes available to benzenoids. As pointed out before,188 molecular shape is intimately associated with molecular electric and steric properties, such as quadrupole moment or van der Waals envelope, which are implicated in structure-activity...
relationships from odour perception\cite{37} to carcino-\textcolor{red}{gencicity}\cite{8,27,32,34}. Codes based on boundaries seem especially suitable for systematising our notions of shapes of benzenoids. The reader is referred to the books\cite{15,30} for definitions and basic facts.

2. PRELIMINARIES

We begin by giving a mathematical definition of a \textit{fusene}\cite{7,8}. The class of fusenes contains as a proper subclass the class of benzenoids.

**Definition 2.1.** A \textit{fusene} is a simple subcubic 2-connected plane graph such that all bounded faces are hexagons and all vertices not on the outer face have degree 3.

Benzenoids can be now defined in terms of fusenes.

**Definition 2.2.** A fusene that can be embedded in the infinite hexagonal lattice is called a \textit{benzenoid}.

In other words, benzenoids are those fusenes which are also subgraphs of the infinite hexagonal lattice.

**Example 2.1.** Figure 1 shows four subcubic plane graphs. Pentalene is not a fusene, because its bounded faces are pentagons. Biphenyl is not a fusene because it is not 2-connected. Anthracene and [6]helicene are both fusenes. Anthracene is also a benzenoid, whilst [6]helicene is not.

Let us denote the class of all benzenoids by \( \mathcal{B} \) and the class of all fusenes by \( \mathcal{F} \). The \textit{inner dual} of a plane graph is its dual graph with the vertex that corresponds to the outer removed. A \textit{catacondensed} fusene is a fusene whose inner dual is a tree. Fusenes that are not catacondensed are called \textit{pericondensed}. We will denote the class of all catacondensed fusenes by \( \mathcal{F}^* \). Catacondensed fusenes can be further divided into \textit{branched} and \textit{non-branched} fusenes. A catacondensed fusene is called non-branched if its inner dual is a path; otherwise it is called branched. The class of non-branched fusenes will be denoted by \( \mathcal{F}^* \). Those definitions are naturally inherited by benzenoids. The class of catacondensed benzenoids and non-branched benzenoids will be denoted, respectively, by \( \mathcal{B}^* \) and \( \mathcal{B}' \). In this paper, we restrict to catacondensed benzenoids when using the terms branched and non-branched.

2.1. Boundary-edges Code Revisited

Each fusene can be assigned a \textit{boundary-edges code} (BEC), a sequence of numbers counting the number of boundary edges between two vertices of degree 3, following the perimeter in an arbitrary, say counter-clockwise, direction. This useful tool to describe a benzenoid was introduced by P. Hansen and his co-workers\cite{31}. The code depends on the starting vertex and the chosen direction. However, it can be made unique by choosing the lexicographically maximal code among all possible codes which is often called the \textit{canonical code}. Each benzenoid can be uniquely described by such boundary-edges code, but this does not hold for fusenes\cite{109}. Benzene is an exceptional benzenoid as it is the only benzenoid with no vertex of degree 3. If need be, we assign the code 6 to benzene. In the present paper the (lexicographically maximal) boundary-edges code of \( \mathcal{B} \) will be denoted by code(\( \mathcal{B} \)).

Here, we take a different approach and start from the definition of a code:

**Definition 2.3.** A \textit{code} is a string over the alphabet \( \{1,2,3,4,5\} \).

Note that we permit codes that are not boundary-edges codes of any benzenoid (or fusene). By \( c \oplus d \) we denote concatenation of codes \( c \) and \( d \), e.g.

\[ 422 \oplus 5133 = 4225133. \]

Moreover, \( \sigma_i(c) \) for \( i \geq 0 \) denotes the right circular shift of code \( c \) by \( i \) positions, e.g.

\[ \sigma_3(4225133) = 1334225. \]

This operation can also be defined for the negative values of \( i \), such that \( \sigma_i(c) \) for \( i \geq 0 \) is the left circular shift of \( c \) by \( i \) positions, e.g.

\[ \sigma_{-3}(1334225) = 3422513. \]

By \( \rho(c) \) we denote the reverse of \( c \), e.g.

\[ \rho(3422513) = 3152243. \]

Note that \( \rho(c) = c \) and \( \sigma_i(c) = c \) for every code \( c \).

We will use some properties of codes.

**Definition 2.4.** Let \( c \) be a code. By \( \text{len}(c) \) we denote the length of the code, i.e., the number of symbols appearing in it and by \( \text{sum}(c) \) we denote the sum of all numbers of the code. By \( \text{win}(c) \) we denote the \textit{winding} of the code, which is defined as

\[ \text{win}(c) = \text{sum}(c) - 2 \cdot \text{len}(c). \]
Lemma 2.1. The winding of a concatenation of two codes \( c \) and \( d \) is additive, i.e.
\[
\text{win}(c \oplus d) = \text{win}(c) + \text{win}(d).
\]
Proof. The length of the code and the sum of the numbers of the code are clearly additive. Then
\[
\text{win}(c \oplus d) = \text{sum}(c \oplus d) - 2 \text{len}(c \oplus d)
\]
\[
= \text{sum}(c) + \text{sum}(d) - 2 \text{len}(c) - 2 \text{len}(d)
\]
\[
= \text{win}(c) + \text{win}(d).
\]

Definition 2.2. Let \( B \) be a benzenoid with at least 2 hexagons. Then \( \text{win}(\text{code}(B)) = 6 \).

Proof. The proof proceeds by induction on the number of hexagons, \( h \).

The only benzenoid with \( h = 2 \) hexagons is naphtalene, with code 5S. Clearly, \( \text{win}(5S) = 10 - 2 \cdot 2 = 6 \).

It is known that any benzenoid \( B \) with \( h \) hexagons can be obtained from some benzenoid \( B \) with \( h-1 \) hexagons by adding a new hexagon using either one-, two-, three-, four- or five-contact addition, where \( k \)-contact implies that \( k \) edges of the new hexagon are identified with the edges of \( B \). Let \( \text{code}(B) = c \) and \( \text{code}(B) = c \).

Assume that \( \text{win}(c) = 6 \).

If \( B \) is obtained by one-contact addition then \( c \) can be obtained from \( c \) by replacing the symbol \( s \) (the one that corresponds to the part of the boundary where the new hexagon was attached) with \( s_1s_2s_3 \) where \( s_1 + s_2 + s_3 = 1 \). Then
\[
\text{win}(c) = (\text{sum}(c) + 5 - 1) - 2(\text{len}(c) + 2) = \text{win}(c) = 6.
\]

Analogous arguments can be used for other types of addition. □

3. CONVEX BENZENOIDS AND CONVEXITY DEFICIT

3.1. Graph Invariants

A graph invariant is a function from a class of graphs to a class of values (e.g., integers, real numbers, polynomials) that takes the same value for any two isomorphic graphs. Graph invariants may be categorised by codomains of the functions that define them. When the codomain is the Boolean domain, they are called graph properties. (For example, a graph can either be bipartite or non-bipartite.) Numerous integer invariants exist for graphs: order, size, diameter, girth, genus, chromatic number, etc. Perhaps the most well-known integer invariant in chemical graph theory is the Wiener index.[16] An example of a real number invariant is the Estrada index.[17] In the literature one can find thousands of graph invariants.

3.2. Convex Benzenoids

In 2012, a special sub-family of benzenoids, called convex benzenoids, was introduced by Cruz, Gutman and Rada.[16] This family was further studied and enumerated in.[5] A convex benzenoid can be characterised via its boundary-edges code.

Definition 3.1. Benzenoid \( B \) is convex if its boundary-edges code contains no 1.

The above statement is Proposition 3 in Ref. [5] Since this is one possible characterisation of convex benzenoids we may use it as a definition here.

We note in passing that for infinite benzenoids the situation is more complex. As shown in Ref. [4] infinite benzenoids may have more than one boundary component and may need several infinite codes for its description. Sometimes the code does not describe an infinite benzenoid up to isomorphism. An example of such an infinite convex benzenoid that is not determined by its boundary-edges code is called a strip in Ref. [4]. Strips of different width have the same boundary-edges code. Hence, in this paper we focus mainly on finite benzenoids.

3.3. Convexity Deficit

Both convex and non-convex benzenoids play important and sometimes distinct roles in organic chemistry. For example, in the simplest case of benzenoid isomers, convex anthracene comprising three linearly fused hexagons is less stable than non-convex phenanthrene (see Figure 2). In qualitative theories, this is variously attributed to the larger number of Kekulé structures in phenanthrene (5 vs. 4), its higher Fries number (3 vs. 2) or its higher Clar number (2 vs. 1), all of which are inextricably linked to its angular, non-convex shape. We think it will be useful to introduce a measure that will tell us by how much the shape of benzenoid departs from convexity. We call this measure the convexity deficit.

Definition 3.2. A benzenoid \( B \) with boundary-edges code \( c \) is \( k \)-convex, \( k \geq 0 \), if the average of \( k + 1 \) consecutive values in \( c \) is always at least 2. The minimum such value of \( k \) is called the convexity deficit of \( B \) and is denoted by \( \text{cd}(B) \).

For an infinite benzenoid \( B \), we may have \( \text{cd}(B) = \infty \).

We may write down a formal definition:
\[
\text{avg}(c,k) = \min \left\{ \frac{\text{sum}(d)}{\text{len}(d)} \mid d \subseteq c \text{ and } \text{len}(d) = k \right\}
\]
\[
\text{cd}(B) = \text{cd}(c) = \min \left\{ k \mid k \geq 0 \text{ and } \text{avg}(c,k+1) \geq 2 \right\}
\]
Note that in the first formula \(d \subseteq c\) denotes any subcode consisting of cyclically consecutive symbols of code \(c\). Note that convexity deficit generalises the notion of convexity for benzenoids. Clearly, \(cd(B) = 0\) is equivalent to saying that there is no 1 in the code.

**Proposition 3.1.** A benzenoid \(B\) is convex if and only if \(cd(B) = 0\).

**Proposition 3.2.** If \(B\) is a finite benzenoid then \(0 \leq cd(B) \leq \frac{\text{len}(\text{code}(B))}{2} - 1\).

**Proof.** Let \(c = \text{code}(B)\). Recall the definition of winding. Since \(\text{win}(c) = \sum(c) - 2\ \text{len}(c) = 6\), we have

\[
\frac{\sum(c)}{\text{len}(c)} = 2 \quad \frac{6}{\text{len}(c)} = 2
\]

Hence any (finite) benzenoid is \(k\)-convex at least for \(k = \text{len}(c) - 1\). □

We note that for infinite benzenoids there exists no upper bound on the convexity deficit.

**Proposition 3.3.** There exist infinite benzenoids \(B\) that are not \(k\)-convex for any finite value of \(k \geq 0\).

**Proof.** An example is the infinite benzenoid with boundary-edges code \(...2221222...\) shown in Figure 3. It is the complement of the anvil \(\text{AN}\).[4]

### 3.4. Quasi-convex and Pseudo-convex Benzenoids

Now we turn our attention to the non-convex benzenoids that are closest to convex, i.e. the benzenoids with the next smallest convexity deficit.

**Definition 3.3.** A 1-convex benzenoid which is not 0-convex (i.e., \(cd(B) = 1\)) is called quasi-convex.

Note that quasi-convex benzenoids admit a simple characterisation via the boundary-edges code.

**Proposition 3.4.** A benzenoid is quasi-convex if and only if its boundary-edges code contains at least one but no sub-sequence 11, 12, or 21.

**Proof.** A quasi-convex benzenoid is not convex, hence its code contains a 1. Let \(a\) and \(b\) be two cyclically consecutive numbers in the code of this benzenoid. Since it is 1-convex, \(a + b \geq 4\), hence 11, 12, and 21 are forbidden. The converse also follows. □

Convex benzenoids can be classified into families with a common fundamental shape, where zig-zag \((2k)\) sub-sequences define the edges of the shape. Similarly, all quasi-convex benzenoids have a fundamental shape where the edges are defined by either zig-zag or armchair \((1(31)^k)\) sub-sequences. The fundamental shape of a convex benzenoid has at most 6 edges; for a quasi-convex benzenoid it has at most 12. Zig-zag and armchair termination have consequences for stability of benzenoids\,[21] and conductivity of nanotubes.\,[22] A quasi-convex benzenoid that has no zig-zag sub-sequences in its boundary-edges code will be called pseudo-convex.

**Definition 3.4.** A benzenoid whose boundary-edges code contains at least one 1 but no sub-sequence 11 and 2 is called pseudo-convex.

**Proposition 3.5.** Every pseudo-convex benzenoid is quasi-convex but the converse is not true. □

Here are some small examples. Note that naphthalene 55 is convex, phenanthrene 5351 is pseudo-convex and benzo[a]pyrene 513432 is quasi-convex (but not pseudo-convex). A smaller example of such a benzenoid is described by 52441. They are shown in Figures 4 and 5. The BEC 52441 applies to the "pistol" polyhex,\,[23] named for its shape. BEC apply equally to benzenoids and polyhexes.

We have developed software that transforms the boundary edges code to the description of a benzenoid via position of its hexagons in the hexagonal tesselation of the plane, as well as a tool that can draw the corresponding benzenoid. We can also compute several parameters such as convexity deficit (of course, convexity deficit is obtained directly from the BEC). We present computational results in Tables 1 and 2. One is a table of small benzenoids, together
4. EXTREMAL CONVEXITY DEFICIT

Clearly benzenoids with a small number of hexagons cannot have large convexity deficit. For instance, benzenoids with up to 5 hexagons have convexity deficit at most 3.

**Definition 4.1.** Let $B$ be a benzenoid. Let $h(B)$ denote the number of hexagons of $B$. Let $\text{mcd}(h)$ denote the maximum convexity deficit among all benzenoids on $h$ hexagons, i.e.

$$\text{mcd}(h) = \max \{ cd(B) \mid h(B) = h \}.$$  

Call each benzenoid attaining $cd(h)$ extremal, and $\text{ex}(h)$ the number of extremal benzenoids with $h$ hexagons.

We performed a computer search to find extremal benzenoids among all benzenoids with $h$ hexagons for all $h \leq 18$. The results are summarised in Table 3. In particular, we noticed that only one extremal benzenoid is pericondensed. It has $h = 6$ hexagons and can be described by the boundary-edges code 533244111 and is depicted in Figure 7(a). Moreover, all other extremal benzenoids with $h \leq 6$ are unbranched. There is a unique smallest branched extremal benzenoid having $h = 7$. It has boundary-edges code 52331511112 and is depicted in Figure 7(b). Note that the spiral benzenoid $S(h)$ attains the maximum value of convexity deficit among all unbranched catacondensed benzenoids. For $h \geq 14$, it appears that all extremal benzenoids are branched.

We were able to find an interesting family of benzenoids, one member for each number of hexagons. We call them **spiral benzenoids**.

**Definition 4.2.** Let $S(h)$ denote the spiral benzenoid on $h \geq 2$ hexagons determined by the following procedure. Let $a$ and $b$ denote the following infinite codes:

$$a = 333232322322322232223222232222(2^3),$$

$$b = 11121212122122122122212221(2^1),$$

using $\odot$ for repeated concatenation using the $\oplus$ operation. Let $a(l)$ denote the substring composed of the first $l$ symbols of $a$ and similarly define $b(l)$. Let

$$w(l) = 5a(l)5pb(l),$$

where $pb(l)$ denotes the reversal of $b(l)$. The spiral benzenoid, denoted $S(h)$, has $h$ hexagons and is defined by the boundary-edges code $w(h-2)$.

For an example of $S(h)$, see Figure 8.

**Example 4.1.** For small values of $l$ we obtain:

- $S(2) = w(0) = 55$;
- $S(3) = w(1) = 5352$;
- $S(4) = w(2) = 533511$;
- $S(5) = w(3) = 53335111$;
- $S(6) = w(4) = 5333252111$; etc.

All $S(2)$, ..., $S(13)$ are extremal.
| Name                  | BEC | $h$     | Class        | $cd$ | Formula  | CAS     |
|-----------------------|-----|---------|--------------|------|----------|---------|
| benzene               | 6   | 1       | convex       | 0    | C$_6$H$_6$ | 71-43-2 |
| naphthalene           | 55  | 2       | convex       | 0    | C$_9$H$_8$ | 91-20-3 |
| phenalene/phenalenyl  | 444 | 3       | convex       | 0    | C$_7$H$_8$ | 203-80-5|
| anthracene            | 5252| 3       | convex       | 0    | C$_9$H$_8$ | 120-12-7|
| phenanthrene          | 5351| 3       | pseudo-convex| 1    | C$_{10}$H$_8$ | 85-01-8 |
| tetracene/naphthacene | 522522| 4       | convex       | 0    | C$_{11}$H$_{12}$ | 92-24-0 |
| pyrene                | 4343| 4       | convex       | 0    | C$_{10}$H$_8$ | 129-00-0|
| benzo(b)pyrene        | 52441| 4       | quasi-convex | 1    | C$_{10}$H$_{12}$ | 112772-04-0|
| triphenylene          | 515151 | 4       | pseudo-convex| 1    | C$_{10}$H$_8$ | 217-59-4 |
| benzo(c)phenanthrene  | 533511| 4       | pseudo-convex| 1    | C$_{10}$H$_8$ | 195-19-7 |
| benzo(a)anthracene    | 512523| 4       | pseudo-convex| 1    | C$_{10}$H$_8$ | 56-55-3 |
| olympicene/olimpicenyl| 42433| 5       | convex       | 0    | C$_{11}$H$_{12}$ | 191-33-3 |
| pentacene             | 52225222| 5       | convex       | 0    | C$_{10}$H$_8$ | 135-48-8 |
| picene                | 51315313| 5       | pseudo-convex| 1    | C$_{13}$H$_{12}$ | 213-46-7 |
| [5]helicene           | 53335111| 5       | pseudo-convex| 1    | C$_{10}$H$_8$ | 188-52-3 |
| perylene              | 441441| 5       | pseudo-convex| 1    | C$_{10}$H$_8$ | 198-55-0 |
| benzo(a)pyrene        | 513432| 5       | quasi-convex | 1    | C$_{10}$H$_8$ | 50-32-8 |
| benzo(e)pyrene        | 514341| 5       | pseudo-convex| 1    | C$_{10}$H$_8$ | 192-97-2 |
| dibenz(a,j)anthracene | 53215321| 5       | pseudo-convex| 1    | C$_{10}$H$_8$ | 53-70-3 |
| pentaphene            | 52125232| 5       | pseudo-convex| 1    | C$_{10}$H$_8$ | 222-93-5 |
| dibenz(a,j)anthracene | 51215323| 5       | pseudo-convex| 1    | C$_{10}$H$_8$ | 224-41-9 |
| trianguleny           | 424242| 6       | convex       | 0    | C$_{10}$H$_8$ | 191-26-4 |
| anthanthrene          | 324324| 6       | convex       | 0    | C$_{10}$H$_8$ | 191-26-4 |
| hexacene              | 522252222| 6       | convex       | 0    | C$_{10}$H$_8$ | 258-31-1 |
| benzo(g,h)perylenes   | 414333| 6       | pseudo-convex| 1    | C$_{10}$H$_8$ | 191-24-2 |
| zethrene              | 42144214| 6       | convex       | 0    | C$_{10}$H$_8$ | 214-63-1 |
| coronene              | 333333| 7       | convex       | 0    | C$_{10}$H$_8$ | 191-07-1 |
| heptacene             | 5222252222| 7       | convex       | 0    | C$_{10}$H$_8$ | 258-38-8 |
| peropyrene            | 43134313| 7       | pseudo-convex| 1    | C$_{10}$H$_8$ | 188-96-5 |
| terylene              | 4413144131| 8       | pseudo-convex| 1    | C$_{10}$H$_8$ | 188-72-7 |
| tribenzo[b,n,pq]perylenes | 5141251331| 8       | pseudo-convex| 1    | C$_{10}$H$_8$ | 190-81-8 |
| tribenzo[b,k,pq]perylenes | 5141451331| 8       | pseudo-convex| 1    | C$_{10}$H$_8$ | 190-81-8 |
| tribenzo[b,b,ghi,n]perylenes | 5141251331| 8       | pseudo-convex| 1    | C$_{10}$H$_8$ | 190-81-8 |
| ovalene               | 333323332| 10      | convex       | 0    | C$_{10}$H$_8$ | 190-26-1 |
| teropyrene            | 43131431313| 10      | pseudo-convex| 1    | C$_{10}$H$_8$ | 190-26-1 |
| hexabeno[b,c,ef,hk,kl,mo,nq]coronene | 414141414141| 13      | pseudo-convex| 1    | C$_{10}$H$_8$ | 190-24-9 |
| hexabeno[a,d,ghi,j,m,pl]coronene | 51151115111511511| 13      | pseudo-convex| 1    | C$_{10}$H$_8$ | 1065-80-1 |
| dicoronylene          | 233332123333321| 15      | pseudo-convex| 1    | C$_{10}$H$_8$ | 98570-53-7 |
Table 2. Some families of benzenoids and their convexity defect

| Benzenoid family       | cd(∅) | BEC      |
|------------------------|-------|----------|
| Linear                 | l(n)  | 52^{n-2} |
| n                      | 0     | Ref. [15, p. 62] |
| Two segments           | M_l(n, m), m > 1 | 52^{n-2}52^{m-2} |
| m + n – 1              | m + n – 3 | Ref. [15, p. 62] |
| Three segments         | M_l(m, n, k), m, n, k > 1 | 52^{m-1}52^{n-1}52^{k-1} |
| m + n + k – 2          | m + n + k – 4 | Ref. [15, p. 62] |
| Three segments         | Z_l(m, n, k), m, n, k > 1 | 52^{m-1}52^{n-1}52^{k-1} |
| m + n + k – 2          | max(m, n) + k – 3 | Ref. [15, p. 62] |
| Chevron                | Ch(n, m, k), m, n, k > 2 | 42^{m-2}52^{n-2}52^{k-2} |
| &m(m + k – 1)          | m + k – 3 | Refs. [28, 14], [15, p. 111] |
| Prolate triangle       | Π(m), m ≥ 2 | 51^{m-2}52^{m-2} |
| &m(m + 1)              | 1 (quasi-convex) | Ref. [15, p. 182] |
| Prolate pentagon       | Π_l(m), m, n ≥ 2 | 32^{m-1}41^{m-1}52^{m-2} |
| &m(m + 1) + (n – 1)(2m – 1) | 1 (quasi-convex) | Ref. [15, p. 182] |
| Oblate triangle        | Ω_l(m), m ≥ 2 | 43^{m-2}52^{m-2} |
| &m(m + 1) + (m – 1)    | \{0 (convex), m = 2, 1 (quasi-convex), m > 2\} | Ref. [15, p. 197] |
| Prolate triangle       | Ω_l(m), m ≥ 2 | 43^{m-2}52^{m-2} |
| &m(m + 3)              | 1 (quasi-convex) | Ref. [15, p. 197] |
| Prolate rectangle      | Ω_l(m, n), m, n ≥ 2 | 42^{m-2}41^{m-2}4(13)^{n-2} |
| n(m + n – 1)           | 1 (quasi-convex) | Refs. [41, 15, p. 201] |
| Dihedral all-benzenoids| S_m, m ≥ 2 | S1215(13)^{m-2}S1215(13)^{m-2} |
| 7m                     | 3     | Ref. [24], [15, p. 215] |
| 6                      | T (2) | 51415141 |
| 0(m, m)                | 1 (pseudo-convex) | Refs. [24], [15, p. 215] |
| 7m – 8                 | T (m), m ≥ 3 | 41414(13)^{m-1}41414(13)^{m-1} |
|                        | 1 (pseudo-convex) | Refs. [24], [15, p. 215] |

It is easy to find the extremal fusenes in the subclass $\mathcal{F}$: A small example is the helicene in Figure 1(d).

**Proposition 4.1.** In the class of unbranched catacondensed fusenes $\mathcal{F}$, the helicene $51^{h-2}53^{h-2}$ obtains the maximal convexity deficit among all $F \in \mathcal{F}$ with $h \geq 2$ hexagons. The maximal convexity deficit is $\max(2h - 7, h - 2)$.

**Proof.** Each unbranched catacondensed fusene can be described with a boundary-edges code $5s_1r_1...s_n5l_1r_2...s_{n-2}5l_2r_3...s_2$, where $s_i + r_i = 4$ for all $1 \leq i \leq h - 2$. It is clear that by setting $s_i = 1$ for all $i = 1, ... , h - 2$ we will obtain one of the fusenes, let us denote it by $F$, with the longest possible subcode $c$ such that $\sum_{i \in c} s_i < 2$.

If $h$ is large enough, then the code will contain $h - 2$ symbols 1, symbol 5 and a certain number, let us denote it by $l$, of symbols 3. We are looking for the largest possible $l$ such that

$$\frac{(h - 2) + 5 + 3l}{h - 2 + 1 + l} < 2. \tag{1}$$

Equation (1) is equivalent to

$$l < h - 5 \tag{2}$$

when $h + l > 3$ (this holds for large enough $h$ since $l \geq 0$). From Equation (2) it follows that we can take $l = h - 6$. This is valid if $h \geq 6$ and the convexity deficit equals $(h - 2) + 1 + (h - 6) = 2h - 7$. DOI: 10.5562/cca3602
Table 3. Maximal convexity deficit mcd(h) for each 2 ≤ h ≤ 18 and the number of extremal benzenoids ex(h). The last column contains an example of such a benzenoid.

| h   | mcd(h) | ex(h) | An example (boundary-edges code)                        |
|-----|--------|------|--------------------------------------------------------|
| 2   | 0      | 1    | S351                                                   |
| 3   | 1      | 1    | S351                                                   |
| 4   | 2      | 2    | S32521                                                |
| 5   | 3      | 6    | S325212                                               |
| 6   | 4      | 16   | S32252212                                            |
| 7   | 6      | 3    | S33151112                                            |
| 8   | 8      | 2    | S333235211211                                        |
| 9   | 10     | 3    | S3323225211211                                       |
| 10  | 12     | 6    | S3333225211211                                       |
| 11  | 14     | 16   | S3331111252123333312                                    |
| 12  | 16     | 37   | S33233222252222221111                                 |
| 13  | 18     | 102  | S33233222252222221111                                 |
| 14  | 21     | 2    | S33322211151351111221111                              |
| 15  | 23     | 12   | S333321121151351111211111                             |
| 16  | 25     | 42   | S3332322221513511112121111                            |
| 17  | 27     | 149  | S33323211321151351111121111                            |
| 18  | 29     | 489  | S3332322222151351111112121111                         |

Figure 8. Spiral benzenoid S(h) on h hexagons has cd(S(h)) = max(h − 2, 2h − 8).

If h < 6, then there are only 4 fusenes to analyse. By manual inspection we can see that the convexity deficit in each case is h − 2.

Proposition 4.2. Among all catacondensed unbranched benzenoids on h hexagons, S(h) has the maximal convexity deficit. The convexity deficit of S(h) is cd(S(h)) = max(h − 2, 2h − 8).

Proof. First, observe that the code 533323 is a subcode of S(h) for all h ≥ 7.

We know that $\sum(S(h)) = 4h + 2$ and $\text{len}(S(h)) = 2h - 2$. Let c be the subcode of code(S(h)) that is obtained from code(S(h)) by erasing 533323 (i.e. $c = 111121...$). Then $\sum(c) = 4h + 2 - 19 = 4h - 17$ and $\text{len}(c) = 2h - 2 - 6 = 2h - 8$. Therefore

$$\frac{\sum(c)}{\text{len}(c)} = \frac{4h - 17}{2h - 8} < 2.$$

It is easy to see that the above also holds for all prefixes of the code c. This implies that cd(S(h)) ≥ 2h − 8 for all h ≥ 7. Let us now show that there exists no unbranched benzenoid B such that cd(B) = 2h − 7.

For contradiction, suppose that there exists an unbranched benzenoid B such that cd(B) = 2h − 7. Let c be the code which remains when the maximal subcode d, for which $\frac{\sum\text{ch(d)}}{\text{len(ch(d))}} < 2$, is erased from code(B). We have len(c) = 5 and

$$\frac{4h + 2 - \sum(c)}{2h - 7} < 2.$$

If h is large enough, we obtain $\sum(c) > 16$. The code c contains 5 symbols, each of which is an element of the set {1, 2, 3, 5}. One of the symbols must be 5 (otherwise the sum can not be greater than 16). Also, the code cannot contain symbol 5 twice if the benzenoid is large enough (their corresponding hexagons are located at the opposite ends of the chain). From $\sum(c) > 16$ it follows that all the other symbols have to be 3. The code 353 can not be a subcode of a benzenoid due to geometric restrictions. The only remaining option is $c = 53333$, which again can not be a subcode of a benzenoid, a contradiction. Therefore, S(h) attains the maximal convexity deficit among all unbranched benzenoids on h hexagons.

5. CONCLUSION

In this contribution we have briefly revisited several families of benzenoids that have been studied in the past. Most of them are taken from the book by Cyvin and Gutman [Ref. [15, p. 62]]. Here they are defined rigorously with respect to convexity deficit. Table 3 summarises all small cases up to 18 hexagons. BECs of these benzenoids are stored in Ref. [6]. We observed from these data that some clear patterns emerged.

In particular, let $F(h,k)$ denote the number of benzenoids on h hexagons having convexity deficit equal to k. Note that $F(h,\text{mcd}(h)) = \text{ex}(h)$ and $F(h,k) = 0$ for all $k > \text{mcd}(h)$.

Conjecture 4.3. Let h ≥ 0. The sequence $F(h,0), F(h,1), F(h,2), ..., F(h,\text{mcd}(h))$ is unimodal.
The motivation for the previous statement comes from computational investigation for small values of $h$.

Our empirical studies show an interesting picture of extremal benzenoids. It seems that:

1. there is only one extremal benzenoid that is pericondensed;
2. there are only finitely many extremal unbranched pericondensed benzenoids;
3. all extremal benzenoids for $h \leq 14$ are branched;
4. there is no upper bound on the number of branched points of extremal benzenoids when $h$ tends to infinity.

These observations could be formulated as conjectures and are a subject of further research.

Acknowledgment. This work was supported in part by the Slovenian Research Agency (grants P1-0294, N1-032, and J1-7051), the German Research Foundation (grant STA 850/19-2 within SPP 1738), and bilateral Slovenian-German project “Mathematical Foundations of Selected Topics in Science”. This paper was conceived at the Workshop on Discrete and Computational Biomathematics, and Mathematical Chemistry, Koper, Slovenia, 15 – 17 November 2017, organised by the Slovenian Discrete and Applied Mathematics Society and University of Primorska, FAMNIT, and completed at the 33rd TBI Winterseminar, Bled, Slovenia, 11 – 17 February 2018.

REFERENCES

[1] H. I. Abdel-Shafy, M. S. M. Mansour, *Egypt. J. Petrol.* 2016, 25, 107–123 https://doi.org/10.1006/ijpet.2015.03.011.

[2] L. J. Allamandola, A. G. M. Tielsens, J. R. Barker, *Astrophys. J. Suppl.* 1989, 71, 733–775 https://doi.org/10.1086/191396.

[3] R. Cruz, I. Gutman, J. Rada, *Linear Algebra Appl.* 2008, 427 (2007), 70–76 https://doi.org/10.1016/j.laa.2007.06.020.

[4] M. Deza, P. W. Fowler, V. Grishukhin, *J. Inf. Comput. Sci.* 2001, 41, 300–308 https://doi.org/10.1021/ci000060o.

[5] J. R. Dias, *Handbook of Polycyclic Hydrocarbons: Part A, Benzenoid Hydrocarbons*, volume 30A of Physical Sciences Data, Elsevier, Amsterdam, 1987.

[6] J. R. Dias, *Handbook of Polycyclic Hydrocarbons: Part B, Polycyclic Isomers and Heteroatom Analogs of Benzenoid Hydrocarbons*, volume 30B of Physical Sciences Data, Elsevier, Amsterdam, 1988.

[7] J. R. Dias, *J. Inf. Model.* 2005, 45, 562–571 https://doi.org/10.1021/ci0500334.

[8] M. S. Dresselhaus, G. Dresselhaus, P. C. Eklund, Science of Fullerenes and Carbon Nanotubes, Academic Press, London, 1996 https://doi.org/10.1016/b978-0-12-221820-0.x5000-x.

[9] C. J. Fetzer, *Large (C24) polycyclic aromatic hydrocarbons: Chemistry and analysis*, volume 158 of Chemical Analysis: A series of monographs in Analytical Chemistry and its Applications, Wiley–Interscience, New York, 2000.

[10] E. Clar, *Polycyclic Hydrocarbons: Volume 2*, Springer-Verlag, Berlin, 1st edition, 1964 https://doi.org/10.1007/978-3-662-01665-7.

[11] R. Cruz, I. Gutman, J. Rada, *MATCH Commun. Math. Comput. Chem.* 2012, 68, 97–108.

[12] S. J. Cyvin, J. Brunvoll, R. S. Chen, B. N. Cyvin, F. J. Zhang, *Theory of Coronoid Hydrocarbons II*, volume 62 of Lecture Notes in Chemistry, Springer-Verlag, 1994.

[13] S. J. Cyvin, J. Brunvoll, B. N. Cyvin, *Theory of Coronoid Hydrocarbons*, volume 54 of Lecture Notes in Chemistry, Springer-Verlag, 1991.

[14] S. J. Cyvin, B. N. Cyvin, I. Gutman, Z. Naturforsch. A 1985, 40, 1253–1261 https://doi.org/10.1515/zna-1985-1211.

[15] S. J. Cyvin and I. Gutman, *Kekulé Structures in Benzenoid Hydrocarbons*, volume 46 of *Lecture Notes in Chemistry*, Springer, Heidelberg, 1988 https://doi.org/10.1007/978-3-662-00892-8.

[16] S. J. Cyvin and I. Gutman, *Kekulé Structures in Benzenoid Hydrocarbons*, volume 48 of *Lecture Notes in Chemistry*, Springer-Verlag, 1988.

[17] J. A. de la Peña, I. Gutman, J. Rada, *Linear Algebra Appl.* 427 (2007), 70–76 https://doi.org/10.1016/j.laa.2007.06.020.

[18] N. Bašić, *Art Discrete Appl. Math.* 2019, 2, #1.09.

[19] N. Bašić, P. W. Fowler, T. Pisanski, *MATCH Commun. Math. Comput. Chem.* 2018, 80, 153–172.

[20] G. Brinkmann, G. Caporossi, P. Hansen, *J. Algorithms* 2002, 45, 155–166 https://doi.org/10.1016/s0196-6774(02)00215-8.

[21] G. Brinkmann, C. Grothaus, I. Gutman, *J. Math. Chem.* 2007, 42, 909–924 https://doi.org/10.1007/s10910-006-9148-z.

[22] E. Clar, *Polycyclic Hydrocarbons: Volume 1*, Springer-Verlag, Berlin, 1st edition, 1964 https://doi.org/10.1007/978-3-662-01665-7.
[27] H. V. Gelboin, P. O. P. Ts’o (eds.), Polycyclic Hydrocarbons and Cancer, Volume 1: Environment, Chemistry, and Metabolism, Academic Press, 1978 https://doi.org/10.1016/b978-0-12-279201-4.x5001-4.

[28] M. Gordon, W. H. T. Davison, J. Chem. Phys. 1952, 20, 428–435 https://doi.org/10.1063/1.1700437.

[29] X. Guo, P. Hansen, M. Zheng, Disc. Appl. Math. 2002, 118, 209–222, https://doi.org/10.1016/s0166-218x(01)00180-9.

[30] I. Gutman, S. J. Cyvin, Introduction to the Theory of Benzenoid Hydrocarbons, Springer-Verlag, Heidelberg, 1989 https://doi.org/10.1007/978-3-642-87143-6.

[31] P. Hansen, C. Lebatteux and M. Zheng, J. Mol. Struct. (Theochem) 1996, 363, 237–247 https://doi.org/10.1016/0166-1280(95)04139-7.

[32] D. M. Jerina, J. M. Sayer, D. R. Thakker, H. Yagi, W. Levin, A. W. Wood, A. H. Conney, Carcinogenicity of polycyclic aromatic hydrocarbons: The Bay-Region Theory, in: Carcinogenesis: Fundamental Mechanisms and Environmental Effects, Springer, Dordrecht, volume 13 of The Jerusalem Symposia on Quantum Chemistry and Biochemistry, 1980 pp. 1–12.

[33] J. V. Knop, K. Szymanski, Ž. Jeričević, N. Trinajstić, J. Comput. Chem. 1983, 4, 23–32 https://doi.org/10.1002/jcc.540040105.

[34] D. Lloyd, The Chemistry of Conjugated Cyclic Compounds: To Be or Not to Be Like Benzene?, John Wiley & Sons, 1989.

[35] J. N. Murrell, The theory of the electronic spectra of organic molecules, Methuen, London, 1963.

[36] D. H. Rouvray, Chapter 2 – The rich legacy of half a century of the Wiener index, in: D. H. Rouvray and R. B. King (eds.), Topology in Chemistry: Discrete Mathematics of Molecules, Woodhead Publishing, Chichester, West Sussex, 2002, pp. 16–37 https://doi.org/10.1533/9780857099617.16.

[37] Society of Chemical Industry, London, Molecular Structure and Organoleptic Quality, S. C. I. Monograph No. 1, Macmillan, New York, 1958, symposium organized by the Overseas Section, Geneva, May 1957.

[38] R. J. Taylor, Electrophilic Aromatic Substitution, Wiley, Chichester, 1990.

[39] M. Tomlinson, An Introduction to the Chemistry of Benzenoid Compounds, Elsevier, New York, 1971 https://doi.org/10.1016/c2013-0-10050-5.

[40] N. Trinajstić, Chemical Graph Theory, CRC Press, Boca Raton, 2nd edition, 1992.

[41] T. F. Yen, Theor. Chem. Acc. 1971, 20, 399–404 https://doi.org/10.1007/bf00527196.