

# Extremal Problems for Topological Indices in Combinatorial Chemistry

Robert F. Tichy\* and Stephan Wagner

Department of Mathematics  
Graz University of Technology  
Steyrergasse 30, A-8010 Graz, Austria

Phone: +43 (0) 316 / 873 - 7120

Fax: +43 (0) 316 / 873 - 7126

tichy@tugraz.at

wagner@finanz.math.tugraz.at

## Abstract

Topological indices of molecular graphs are related to several physicochemical characteristics; recently, the inverse problem for some of these indices has been studied, and it has some applications in the design of combinatorial libraries for drug discovery. It is thus very natural to study also extremal problems for these indices, i.e. finding graphs having minimal or maximal index. In this paper, these questions will be discussed for three different indices, namely the  $\sigma$ -index, the  $c$ -index and the  $Z$ -index, with emphasis on the  $\sigma$ -index.

Key words: drug discovery, extremal problems,  $\sigma$ -index,  $c$ -index, and  $Z$ -index.

# 1 Introduction

Formally, a topological index is a map from the set of molecular graphs to the real numbers. The indices discussed in this paper appear, for example, in the work of Li et al. (2003). For their uses in drug design and molecular recognitions, we refer to Sheridan and Kearsley (1995), Venkatasubramanian et al. (1995) and Goldman et al. (2000).

The inverse problem for topological indices (i.e. finding a graph having some given index value) has been studied by Goldman et al. (2000) for the Wiener index and by Li et al. (2003) for the  $\sigma$ -, the  $c$ -, the  $Z$ - and the  $M_1$ -index. All these indices are popular in the study of molecular graphs, see for example Rouvray (1973), Sabljic and Trinajstić (1981), Merrifield and Simmons (1989) and Trinajstić (1992). In the stated papers on the inverse problem, algorithmic solutions have been given. In the current work, we are interested in the natural question of finding graphs of minimal or maximal index value, given their size (number of vertices or edges).

We start with the definitions of the indices we are considering (cf. Li et al., 2003):

**Definition 1** The  $\sigma$ -index  $\sigma(G)$  of a graph  $G$  is the number of independent vertex subsets of  $G$  (including the empty set), where a set of vertices is said to be independent if there is no pair of connected vertices in it.

**Definition 2** The  $c$ -index  $c(G)$  of a graph  $G$  is the number of cliques of  $G$  (including the empty clique), i.e. the number of complete subgraphs.

**Definition 3** The  $Z$ -index  $Z(G)$  of a graph  $G$  is the number of independent edge subsets of  $G$  (including the empty set), where a set of edges is said to be independent if there is no pair of edges sharing a common vertex in it.

In section 2, we will give a short survey of known mathematical facts on the  $\sigma$ -index, and then turn to new results and generalizations. In sections 3 and 4, the  $c$ - and  $Z$ -index are considered, with similar results, but less detailed. All graphs considered here are finite, simple and connected (if not stated otherwise), and we refer to Biggs (1989) for notation and terminology.

## 2 The $\sigma$ -index

The  $\sigma$ -index is also called Merrifield-Simmons-Index, and it is one of the most popular topological indices in chemistry. In their original work in 1989, Merrifield and Simmons were able to show the correlation between this index and boiling points. It was further investigated in various subsequent papers (Wang et al., 2001; Li et al., 2003)

In mathematics, the  $\sigma$ -index has already been introduced in 1982 in a paper of Prodinger and Tichy, although it is called Fibonacci number of a graph there. In this and two subsequent papers (1983 and 1986), the  $\sigma$ -index for trees, especially  $t$ -ary trees, was investigated. There, it was also proved that the tree of maximal  $\sigma$ -index for a given number of vertices is a star, whereas the tree of minimal  $\sigma$ -index is a path. Alameddine (1998), for instance, studied bounds for the  $\sigma$ -index of a maximal outerplanar graph.

A list of the trees of maximal  $\sigma$ -index has been given in Knopfmacher et al. (2004). We will extend this list to all connected graphs, and we will give a list of the graphs (not necessarily connected) of maximal  $\sigma$ -index. First, we need some definitions of special graphs appearing in the following:

**Definition 4**  $\mathcal{C}(n)$  denotes the set of connected graphs with  $n + 1$  vertices,  $\mathcal{T}(n)$  denotes the set of trees with  $n + 1$  vertices (and thus  $n$  edges). Throughout this paper,  $n + 1$  will always be the number of vertices of a graph.

**Definition 5** The following graph classes are denoted by  $X(k, l, m)$  and  $\tilde{X}(k, l, m)$ :

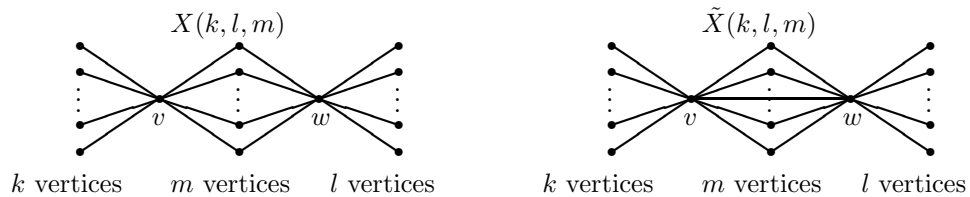


Figure 1: The graph classes  $X(k, l, m)$  and  $\tilde{X}(k, l, m)$ .

where  $k, l, m \geq 0$ ,  $k + l + m = n - 1$ . We suppose, without loss of generality, that  $k \geq l$ .  $X(k, l, m)$  is connected if and only if  $m > 0$ .

The special cases  $X(k, l, 1)$  and  $\tilde{X}(k, l, 0)$  correspond to trees. In Knopfmacher et al. (2004), it was shown that all trees in  $\mathcal{T}(n)$  having a  $\sigma$ -index  $\geq 2^{n-1} + 5$  are either of the form  $X(k, l, 1)$  or  $\tilde{X}(k, l, 0)$  or one of the exception trees in Figure 2.

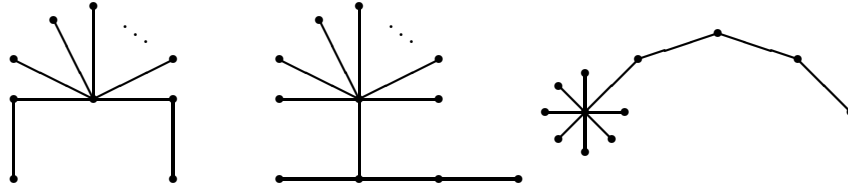


Figure 2: The exception trees.

Throughout this paper, they are denoted by  $TA_n$ ,  $TB_n$  and  $CS_n$  respectively, and their  $\sigma$ -indices are  $\sigma(TA_n) = 9 \cdot 2^{n-4} + 4$ ,  $\sigma(TB_n) = 2^{n-1} + 6$ , and  $\sigma(CS_n) = 2^{n-1} + 5$ . The last one is called the “Christmas star”. It is the tree of maximal  $\sigma$ -index having a diameter  $\geq 5$ .

**Definition 6** The relations  $\sim$ ,  $\succ$  and  $\overset{!}{\succ}$  are defined on  $\mathcal{C}(n)$  in the following way:

$$G_1 \sim G_2 :\iff \sigma(G_1) = \sigma(G_2) \tag{1}$$

$$G_1 \succ G_2 :\iff \sigma(G_1) > \sigma(G_2) \tag{2}$$

$$G_1 \overset{!}{\succ} G_2 :\iff G_1 \succ G_2 \wedge (\nexists G \in \mathcal{C}(n) : G_1 \succ G \succ G_2) \tag{3}$$

We want to determine all connected graphs  $G \in \mathcal{C}(n)$  satisfying  $\sigma(G) \geq \sigma(CS_n)$ . First, we need some auxiliary results:

**Lemma 1** If some graph  $G$  contains a subgraph  $G' \neq G$  with equally many vertices,  $\sigma(G) < \sigma(G')$ . In particular, if  $G$  is a connected graph and  $T \neq G$  a spanning tree,  $\sigma(G) < \sigma(T)$ . Thus, a spanning tree of a graph  $G \in \mathcal{C}(n)$  with  $\sigma(G) \geq \sigma(CS_n)$  must be of the form  $X(k, l, 1)$  or  $\tilde{X}(k, l, 0)$  or be one of the stated exception trees.

PROOF: Clearly, removing edges from a graph increases the  $\sigma$ -index; thus, the lemma follows.  $\square$

**Lemma 2** If  $G$  is a connected graph and  $F$  is a subgraph without cycles, then there exists a spanning tree  $T$  of  $G$  that contains  $F$ .

PROOF: By induction on the number of components of  $F$ . If  $F$  is connected, choose  $T = F$ . Otherwise, let  $F_1, \dots, F_r$  be the components of  $F$ . As  $G$  is connected, there must be an edge  $e \in E(G)$  such that one end of  $e$  lies in  $F_1$ , the other in  $F_2 \cup \dots \cup F_r$ . Add this edge to  $F$  (no cycles can appear, as they would have to use  $e$  twice); the number of components of  $F$  decreases by one, so by the induction hypothesis, we are done.  $\square$

**Lemma 3** If a connected graph  $G \in \mathcal{C}(n)$  contains a simple path of length  $\geq 5$ ,  $\sigma(G) \leq \sigma(CS_n)$ , with equality if and only if  $G \simeq CS_n$ .

PROOF: Extend the simple path contained in  $G$  to form a spanning tree  $T$ . This tree contains a path of length  $\geq 5$ , so its diameter must be  $\geq 5$ . As  $CS_n$  is the tree of diameter  $\geq 5$  that has largest  $\sigma$ -index, the result follows immediately.  $\square$

**Lemma 4**

$$\sigma(X(k, l, m)) = 2^{n-1} + 2^k + 2^l + 1, \quad \sigma(\tilde{X}(k, l, m)) = 2^{n-1} + 2^k + 2^l \quad (4)$$

PROOF: If  $v$  and  $w$  are vertices as in the picture of Definition 5,  $2^{n-1}$  corresponds to all independent vertex subsets containing neither  $v$  nor  $w$ ,  $2^k$  corresponds to those containing only  $w$ ,  $2^l$  to those containing only  $v$ . Furthermore, in  $X(k, l, m)$ ,  $\{v, w\}$  is an independent subset, too.  $\square$

**Proposition 5** A connected graph  $G \in \mathcal{C}(n) \setminus \mathcal{T}(n)$  with  $\sigma(G) \geq 2^{n-1} + 5 = \sigma(CS_n)$  is either of the form  $X(k, l, m)$  or  $\tilde{X}(k, l, m)$  or one of the following graphs (Figure 3):

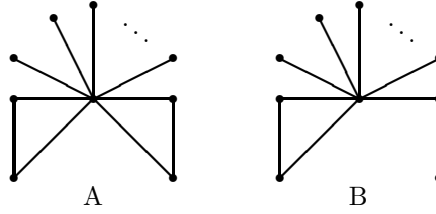


Figure 3: The exception graphs.

PROOF: We consider several cases:

- $G$  contains a cycle of length  $\geq 5$ : if this cycle has length  $\geq 6$ , it contains a path of length 5; if this cycle has length 5, say with vertices  $v_1, \dots, v_5$ , there must be another vertex  $w$  adjacent to a  $v_i$  (if  $G$  is not the cycle  $C_5 \in \mathcal{C}(4)$ , having a  $\sigma$ -index of  $11 < 13 = 2^{4-1} + 5$ ), say  $v_1$ . Then  $w, v_1, \dots, v_5$  is a path of length 5. By Lemma 3,  $\sigma(G) < \sigma(CS_n)$ .
- The largest cycle of  $G$  has length 4, with vertices  $v_1, \dots, v_4$ . There can be no vertices  $w_1, w_2 \notin \{v_1, \dots, v_4\}$  such that  $w_1 \sim w_2$  and  $v_i \sim w_1$ , as there would be a path  $w_2, w_1, v_i, v_{i+1}, v_{i+2}, v_{i+3}$  of length 5 (indices modulo 4). If there are vertices  $w_1, w_2 \notin \{v_1, \dots, v_4\}$  such that  $w_1 \sim v_i$

and  $w_2 \sim v_{i+1}$ , there is a path  $w_1, v_i, v_{i-1}, v_{i-2}, v_{i+1}, w_2$  of length 5. If there is  $w \notin \{v_1, \dots, v_4\}$  such that  $w \sim v_i$ ,  $w \sim v_{i+1}$  for some  $i$ , the graph would contain a cycle of length  $\geq 5$ .

Therefore, all vertices  $\neq v_1, \dots, v_4$  can only be connected to two opposite vertices  $v_i, v_{i+2}$  (without loss of generality,  $i = 1$ ), possibly to both of them. If  $G$  is not of the form  $X(k, l, m)$  or  $\tilde{X}(k, l, m)$ , there must be an edge  $(v_2, v_4)$ . But then, there can be no vertices  $w_1, w_2 \notin \{v_1, \dots, v_4\}$  such that  $w_1 \sim v_1, w_2 \sim v_3$ , as there would be a path  $w_1, v_1, v_2, v_4, v_3, w_2$  of length 5. So all other vertices are only connected to one of the vertices  $v_1, v_3$ , and  $G$  is thus the following graph with  $\sigma$ -index  $\sigma(G) = 2^{n-1} + 2 < \sigma(CS_n)$ :

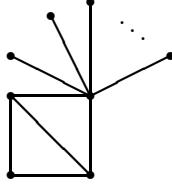


Figure 4: A graph with a cycle of length 4.

- The largest cycle of  $G$  has length 3; two such cycles cannot have a common edge, as they would form a cycle of length 4. If there are cycles  $v_1, v_2, v_3$  and  $w_1, w_2, w_3$  connected by a path  $u_1 = v_1, u_2, \dots, u_k = w_1$ , there is a path  $v_3, v_2, v_1 = u_1, u_2, \dots, u_k = w_1, \dots, w_2, w_3$  of length  $\geq 5$ . So two cycles of length 3 must always have a common vertex.

Consider two cycles  $v_1, v_2, v_3$  and  $w_1 = v_1, w_2, w_3$ . There can be no other  $u \sim v_2$ , as we would again obtain a path of length 5:  $u, v_2, v_3, v_1 = w_1, w_2, w_3$ . Analogously, there is no other vertex adjacent to  $v_3, w_2$  or  $w_3$ .

If there are vertices  $u_1, u_2$  such that  $v_1 \sim u_1$  and  $u_1 \sim u_2$  ( $u_1, u_2 \neq v_i, w_j$ ), consider the graph consisting only of the edges of the two cycles,  $(v_1, u_1)$  and  $(u_1, u_2)$  (obviously, this graph is not connected any more (see picture), but it is a subgraph of  $G$ , so it has larger  $\sigma$ -index). Its  $\sigma$ -index is  $29 \cdot 2^{n-6} < \sigma(CS_n)$  (the  $\sigma$ -index of the only connected component with more than one vertex is 29, and there are  $n - 6$  isolated vertices), a contradiction; so  $G$  must be the exception graph A.

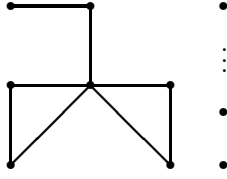


Figure 5: A graph with two cycles of length 3.

The only remaining possibility is a graph with only one cycle of length 3. If we remove one of the edges from the cycle, we obtain a tree which must be of the form  $X(k, l, 1)$ ,  $\tilde{X}(k, l, 0)$ ,  $TA_n$ ,  $TB_n$  or  $CS_n$ . Thus  $G$  results from adding an edge to one of the graphs in this list. As  $\sigma(G) \geq \sigma(CS_n)$ , this cannot be  $CS_n$ .

If we add an edge  $(v, w)$  to  $TB_n$ , the  $\sigma$ -index decreases by the number of independent vertex subsets containing both  $v$  and  $w$ . It is easy to verify that there is always a vertex  $u$  such that  $\{v, w\}$  and  $\{u, v, w\}$  are both independent subsets, so the  $\sigma$ -index decreases at least by 2. As  $\sigma(TB_n) = 2^{n-1} + 6 = \sigma(CS_n) + 1$ , we would obtain a graph with smaller  $\sigma$ -index than  $CS_n$ .

If we add an edge to  $\tilde{X}(k, l, 0)$ , we obtain either one of  $\{\tilde{X}(k-1, l, 1), \tilde{X}(k, l-1, 1)\}$  or a graph containing a cycle of length 4 or a graph of the following form:

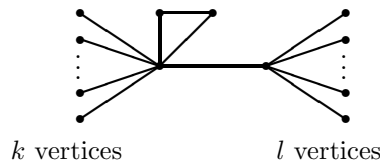


Figure 6: A graph resulting from  $\tilde{X}(k, l, 0)$ .

where  $k+l = n-3$ . The  $\sigma$ -index of this graph is  $3 \cdot 2^{n-3} + 3 \cdot 2^k + 2^l$  (using the same reasoning as in the proof of Lemma 4; similar arguments are also used for the following graphs). Only for  $l = 0$  or  $l = 1$ , this is  $\geq \sigma(CS_n)$ . But for  $l = 0$ , we obtain the graph  $\tilde{X}(n-2, 1, 0)$ , and for  $l = 1$ , we obtain exception graph B.

If we add an edge to  $X(k, l, 1)$ , we obtain either  $\tilde{X}(k, l, 1)$  or a graph containing a cycle of

length  $\geq 4$  or a graph of one of the following forms:

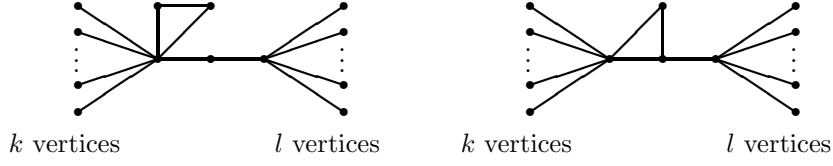


Figure 7: Graphs resulting from  $X(k, l, 1)$ .

The  $\sigma$ -index of the first graph is  $3 \cdot 2^{n-3} + 3 \cdot 2^k + 2^l + 1$ , and  $k + l = n - 4$ . Only for  $l = 0$ , this is  $\geq \sigma(CS_n)$ . But for  $l = 0$ , we obtain exception graph B.

The  $\sigma$ -index of the second graph is  $3 \cdot 2^{n-3} + 2^{k+1} + 2^l + 1$ , and  $k + l = n - 3$ . Only for  $l = 0$ , this is  $\geq \sigma(CS_n)$ . But for  $l = 0$ , we obtain the graph  $\tilde{X}(n - 3, 1, 1)$ .

Finally, if we add an edge to  $TA_n$ , we obtain either a graph containing a cycle of length  $\geq 4$  or one of the following graphs:

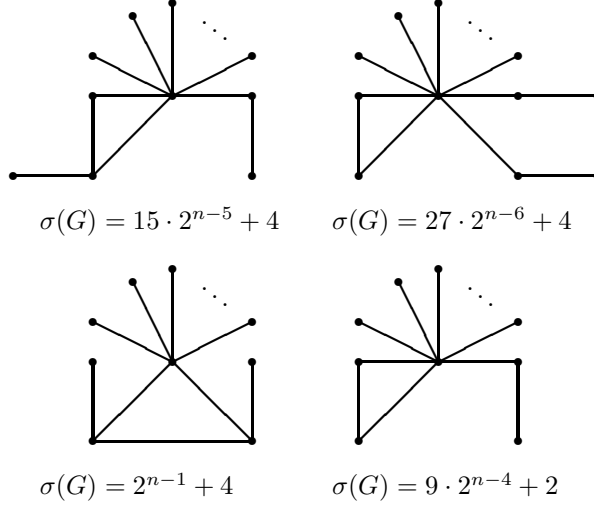


Figure 8: Graphs resulting from  $TA_n$ .

We conclude that there is only one more graph with  $\sigma(G) \geq 2^{n-1} + 5$  that is not of the form  $X(k, l, m)$  or  $\tilde{X}(k, l, m)$  - the exception graph B. The stated theorem thus results.  $\square$

Now, let's denote the exception graphs by  $EA_n, EB_n$ . We have

$$\sigma(EA_n) = 9 \cdot 2^{n-4} + 1, \quad \sigma(EB_n) = 9 \cdot 2^{n-4} + 2 \quad (5)$$



and

$$\sigma(X(k, l, m)) = 2^{n-1} + 2^k + 2^l + 1, \quad \sigma(\tilde{X}(k, l, m)) = 2^{n-1} + 2^k + 2^l. \quad (6)$$

This gives us the order of the graphs we just found with respect to their  $\sigma$ -index:

**Theorem 6** All connected graphs with  $n + 1$  vertices and  $\sigma$ -index  $\geq 2^{n-1} + 5$  are either of the form  $X(k, l, m)$  (with  $k + l + m = n - 1$ ,  $k \geq l$ ,  $m \geq 1$  and  $k \geq 2$  or  $k = l = 1$ ) or  $\tilde{X}(k, l, m)$  (with  $k + l + m = n - 1$ ,  $k \geq l$  and  $k \geq 2$ ) or one of  $TA_n$ ,  $TB_n$ ,  $CS_n$ ,  $EA_n$  or  $EB_n$ . Their order with respect to the  $\sigma$ -index is given by the following list:

$$X(k, l, m) \succ \tilde{X}(k, l, m) \quad (k, l, m \geq 0 \text{ s. t. } k + l + m = n - 1, k \geq l, k \geq 2, m \geq 1), \quad (7)$$

$$\tilde{X}(k, l, m) \succ X(k, l - 1, m + 1) \quad (k, l, m \geq 0 \text{ s. t. } k + l + m = n - 1, k \geq l \geq 2), \quad (8)$$

$$\tilde{X}(k, 1, n - k - 2) \sim X(k, 0, n - k - 1) \quad (2 \leq k < n - 1), \quad (9)$$

$$\tilde{X}(k, 0, n - k - 1) \sim X(k - 1, k - 1, n - 2k + 1) \quad \left(2 \leq k < \frac{n+1}{2}\right), \quad (10)$$

$$\tilde{X}(k, 0, n - k - 1) \succ \tilde{X}(k - 1, n - k, 0) \quad \left(\frac{n+1}{2} \leq k \leq n - 1\right). \quad (11)$$

Furthermore,  $TA_n \sim \tilde{X}(n - 4, 2, 1)$ ,  $EB_n \sim \tilde{X}(n - 4, 1, 2) \sim X(n - 4, 0, 3)$ ,  $EA_n \sim \tilde{X}(n - 4, 0, 3)$ ,  $TB_n \sim X(2, 0, n - 3)$  and  $CS_n \sim \tilde{X}(2, 0, n - 3) \sim X(1, 1, n - 3)$ .

PROOF: By the preceding proposition, we only have to check the relations from above:

$$\sigma(X(k, l, m)) = 2^{n-1} + 2^k + 2^l + 1 > 2^{n-1} + 2^k + 2^l = \sigma(\tilde{X}(k, l, m)), \quad (12)$$

$$\sigma(\tilde{X}(k, l, m)) = 2^{n-1} + 2^k + 2^l > 2^{n-1} + 2^k + 2^{l-1} + 1 = \sigma(X(k, l - 1, m + 1)), \quad (13)$$

$$\sigma(\tilde{X}(k, 1, n - k - 2)) = 2^{n-1} + 2^k + 2 = \sigma(X(k, 0, n - k - 1)), \quad (14)$$

$$\sigma(\tilde{X}(k, 0, n - k - 1)) = 2^{n-1} + 2^k + 1 = \sigma(X(k - 1, k - 1, n - 2k + 1)), \quad (15)$$

$$\begin{aligned} \sigma(\tilde{X}(k, 0, n - k - 1)) &= 2^{n-1} + 2^k + 1 > 2^{n-1} + 2^{k-1} + 2^{n-k} \\ &= \sigma(\tilde{X}(k - 1, n - k, 0)) \quad \left(\text{for } \frac{n+1}{2} \leq k\right), \end{aligned} \quad (16)$$

$$\sigma(TA_n) = 9 \cdot 2^{n-4} + 4 = \sigma(\tilde{X}(n - 4, 2, 1)), \quad (17)$$

$$\sigma(EB_n) = \sigma(\tilde{X}(n - 4, 1, 2)) = \sigma(X(n - 4, 0, 3)) = 9 \cdot 2^{n-4} + 2, \quad (18)$$

$$\sigma(EA_n) = 9 \cdot 2^{n-4} + 1 = \sigma(\tilde{X}(n - 4, 0, 3)), \quad (19)$$

$$\sigma(TB_n) = 2^{n-1} + 6 = \sigma(X(2, 0, n - 3)), \quad (20)$$

$$\sigma(CS_n) = \sigma(\tilde{X}(2, 0, n - 3)) = \sigma(X(1, 1, n - 3)) = 2^{n-1} + 5. \quad (21)$$

This list already determines uniquely the order of all connected graphs with  $\sigma$ -index  $\geq 2^{n-1} + 5$ , so we are done.  $\square$

We now turn to the study of connected graphs having small  $\sigma$ -index. It is obvious that adding edges decreases the  $\sigma$ -index, so connected graphs of large  $\sigma$ -index should be “almost complete” graphs. The following theorem provides precise information:

**Theorem 7** The graph with smallest  $\sigma$ -index in  $\mathcal{C}(n)$  is the complete graph  $K_{n+1}$ , and  $\sigma(K_{n+1}) = n + 2$ . If  $n \geq 3$  and  $n + 2 \leq f \leq 2n + 2$ , there is a graph  $G \in \mathcal{C}(n)$  such that  $\sigma(G) = f$ .

PROOF: The empty set and every one-element subset of  $V(G)$  are always independent sets, so  $\sigma(G) \geq n + 2$  for all  $G \in \mathcal{C}(n)$ . Clearly, the complete graph is the only graph with no more independent vertex subsets.

Given some  $f$  with  $n + 2 \leq f \leq 2n + 2$ , consider the complementary graph  $T = \overline{G}$ . If we take  $T$  to be a forest with  $|V(T)| = n + 1$  and  $|E(T)| = f - n - 2$  (as  $0 \leq f - n - 2 \leq n$ , this is possible), we have

$$\sigma(G) = 1 + (n + 1) + |E(T)| = f, \tag{22}$$

as vertices in  $G$  are independent if and only if they are connected in  $T$ .  $T$  contains no triangles, so there are no independent subsets of size  $\geq 3$ . Furthermore, if  $T$  is not a star (as  $n \geq 3$  we can take  $T$  to be a path instead, for example),  $G$  is a connected graph.  $\square$

REMARK: There are at least  $|T(f - n - 2)|$  ( $|T(n)| - 1$  if  $f = 2n + 2$ ) such graphs (which follows directly from the proof), and generally, there are obviously a lot more, even with larger independent sets (the  $\sigma$ -index of the complement of a single triangle is  $n + 6$ , for instance).

### 3 The $c$ -index

The  $c$ -index can be called the complement of the  $\sigma$ -index, as we have  $c(G) = \sigma(\overline{G})$ , where  $\overline{G}$  denotes the complement graph. This follows easily from the fact that independent subsets in  $\overline{G}$  are cliques in  $G$  and vice versa.

Again, we want to find the graphs of maximal or minimal index. The ideas are quite similar, so we will give the proofs in less detail.

If we don't care about the graphs being connected, we can already derive the following corollary from the last theorem of the preceding section:

**Corollary 8** Let  $\mathcal{G}(n)$  be the class of all graphs (not necessarily connected) with  $n + 1$  vertices. The graph with smallest  $c$ -index in  $\mathcal{G}(n)$  is the graph without edges, and its  $c$ -index is  $n + 2$ . For each  $n + 2 \leq f \leq 2n + 2$ , there is a graph  $G \in \mathcal{G}(n)$  such that  $c(G) = f$ .

It is known from Li et al. (2003) that a connected graph has minimal  $c$ -index if and only if it is a tree: for every tree  $T \in \mathcal{T}(n)$ , we have  $c(T) = n + 2$  (the only cliques are  $n + 1$  single vertices and the empty clique). Adding single edges will give graphs of  $c$ -index  $n + 3, n + 4, \dots$ , so our particular interest will be in connected graphs of maximal  $c$ -index.

This is, in fact, equivalent to finding (not necessarily connected) graphs of large  $\sigma$ -index. Their complement graphs will then have large  $c$ -index.

The aim of the following preliminaries will be to give a list of non-connected graphs with a larger  $\sigma$ -index than the star graph  $S_n$ .

**Lemma 9** If a graph  $G \in \mathcal{G}(n)$  contains a cycle of length  $k$ , its  $\sigma$ -index is  $\sigma(G) \leq l_k 2^{n+1-k}$ , where  $l_k$  is a Lucas number, defined by  $l_0 = 2, l_1 = 1, l_{n+1} = l_n + l_{n-1}$ .

PROOF: If we remove all other edges, the  $\sigma$ -index clearly increases. Then, a cycle of length  $k$  remains together with  $n + 1 - k$  isolated vertices. It is easy to prove that the  $\sigma$ -index of a cycle  $C_k$  is  $\sigma(C_k) = l_k$ , so we have  $\sigma(G) \leq l_k 2^{n+1-k}$ .  $\square$

**Lemma 10** A graph  $G \in \mathcal{G}(n)$  with  $\sigma(G) \geq 2^n + 1$  must be a forest.

PROOF: By the previous lemma, it suffices to show that  $l_k \leq 2^{k-1}$  for  $k \geq 3$ ; this is easily done by induction.  $\square$

**Lemma 11** If one of the components of a forest  $G \in \mathcal{G}(n)$  is not a star,  $\sigma(G) \leq 2^n$ .

PROOF: Let  $T$  be a component of  $G$  that is not a star. Let  $|E(T)| = k$ . As  $T$  is not a star,  $k \geq 3$ . If all edges not belonging to  $T$  are removed, the resulting graphs  $G'$  has larger  $\sigma$ -index than  $G$ . Furthermore,  $\sigma(T) \leq 3 \cdot 2^{k-2} + 2$  (the second-largest  $\sigma$ -index of a tree with  $k$  edges). So we have

$$\sigma(G) \leq \sigma(G') = 2^{n-k} \sigma(T) \leq 3 \cdot 2^{n-2} + 2^{n-k+1} \leq 3 \cdot 2^{n-2} + 2^{n-2} = 2^n. \quad (23)$$

□

Now, we are ready to prove

**Theorem 12** Let  $F(c_1, \dots, c_d)$  be the forest whose components are the stars  $S_{c_1-1}, \dots, S_{c_d-1}$  ( $c_1 + \dots + c_d = n + 1$ ). Then we have, in  $\mathcal{G}(n)$ ,

$$\begin{aligned} F(1, \dots, 1) \succ^! F(2, 1, \dots, 1) \succ^! F(3, 1, \dots, 1) \succ^! F(4, 1, \dots, 1) \sim F(2, 2, 1, \dots, 1) \\ \succ^! F(5, 1, \dots, 1) \succ^! \dots \succ^! F(n, 1) \succ^! F(n+1) \simeq S_n \end{aligned} \quad (24)$$

PROOF: By the preceding lemma, we only have to consider graphs of the form  $F(c_1, \dots, c_d)$ . A formula for  $\sigma(F(c_1, \dots, c_d))$  is determined easily:

$$\sigma(F(c_1, \dots, c_d)) = \prod_{i=1}^d (2^{c_i-1} + 1). \quad (25)$$

Thus

$$\sigma(F(k, 1, \dots, 1)) = (2^{k-1} + 1) \cdot 2^{n+1-k} = 2^n + 2^{n+1-k} > 2^n + 2^{n-k} = \sigma(F(k-1, 1, \dots, 1)) \quad (26)$$

and

$$\sigma(F(4, 1, \dots, 1)) = 9 \cdot 2^{n-3} = \sigma(F(2, 2, 1, \dots, 1)). \quad (27)$$

So we only have to prove that there are no other graphs that might appear in the list above. We write

$$\sigma(F(c_1, \dots, c_d)) = 2^{c_1+\dots+c_d} \prod_{i=1}^d \left(\frac{1}{2} + 2^{-c_i}\right) = 2^{n+1} \prod_{i=1}^d \left(\frac{1}{2} + 2^{-c_i}\right). \quad (28)$$

Each of the factors in the product is  $\leq 1$ , with equality if and only if  $c_i = 1$ . Their product must be  $\geq \frac{1}{2}$ . But if there are some components with  $c_i \geq 2$  and  $c_j \geq 3$ , the product is  $\leq \frac{3}{4} \cdot \frac{5}{8} = \frac{15}{32} < \frac{1}{2}$ . So there are no other possibilities to obtain a  $\sigma$ -index  $\geq 2^n + 1$ . □

**Corollary 13** The complements of the graphs in the theorem are the graphs of maximal  $c$ -index, in the same order as above. All of them are connected, except the complement of the star graph.

## 4 The $Z$ -index

Finally, we take a short look at the  $Z$ -index. It can be related to the  $\sigma$ -index in the following way: for a graph  $G$ , define the line graph  $G'$  to be the graph that results from replacing the edges by vertices and connecting vertices whose corresponding edges in  $G$  have a common vertex. Then, it is easy to verify that  $\sigma(G') = Z(G)$ .

As removing edges means increasing the  $\sigma$ -index, but decreasing the  $Z$ -index, it is quite natural that a large  $\sigma$ -index means a small  $Z$ -index and vice versa. This fact holds only up to a certain extent, though, as it is not difficult to construct graphs  $G_1, G_2$  satisfying both  $\sigma(G_1) > \sigma(G_2)$  and  $Z(G_1) > Z(G_2)$ .

Clearly, the complete graph has largest  $Z$ -index among all graphs from  $\mathcal{C}(n)$  (or even  $\mathcal{G}(n)$ ). Unlike the  $\sigma$ -index, it is a nontrivial problem to determine the exact value for the  $Z$ -index of a complete graph  $K_r$  with  $r$  vertices. In fact, there is no simple formula, but only the following expression as a sum:

$$\sum_{k=0}^{\lfloor \frac{r}{2} \rfloor} \frac{r!}{2^k k! (r-2k)!} \quad (29)$$

For small  $r$ , the exact values are 1, 1, 2, 4, 10, 26, 76, 232, 764, 2620, 9496. By removing a single edge from the complete graph, one obtains the graph of second-largest  $Z$ -index; removing two edges without a common vertex yields the graph of third-largest  $Z$ -index, and this list can be extended (without new ideas, but with increasing effort).

The star is the connected graph having minimal  $Z$ -index, by an easy argument. Generally, one can obtain an interesting fact: given any connected graph  $S$  with  $m$  edges, one can always construct a connected graph  $G \in \mathcal{C}(n)$  containing  $S$  as a subgraph such that  $G$  has “small”  $Z$ -index (precisely, for any given function  $\omega(n)$  that goes to  $\infty$  arbitrarily slow as  $n \rightarrow \infty$ , at most  $\omega(n)$  connected graphs have smaller  $Z$ -index than  $G$ ) if  $n \geq m$  is large enough. The graph  $G$  is constructed as in the following picture, we leave out the details of the proof:

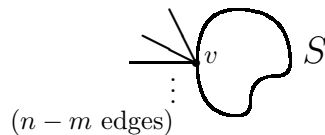


Figure 9: Constructing a graph of small  $Z$ -index.

## Acknowledgement

This work was supported by Austrian Science Fund project no. S-8307-MAT.

## References

- [1] Alameddine, A.F. 1998. Bounds on the Fibonacci number of a maximal outerplanar graph. *Fibonacci Quart.* 36, 206–210.
- [2] Biggs, N.L. 1989. *Discrete mathematics*, Oxford Science Publications. The Clarendon Press Oxford University Press, New York, second edition.
- [3] Goldman, D., Istrail, S., Lancia, G., Piccolboni, A., and Walenz, B. 2000. Algorithmic strategies in combinatorial chemistry. In *Proc. 11th ACM-SIAM Sym. on Discrete Algorithms*, 275–284.
- [4] Kirschenhofer, P., Prodinger, H., and Tichy, R.F. 1983. Fibonacci numbers of graphs. II. *Fibonacci Quart.* 21, 219–229.
- [5] Kirschenhofer, P., Prodinger, H., and Tichy, R.F. 1986. Fibonacci numbers of graphs. III. Planted plane trees, 105–120. In *Fibonacci numbers and their applications (Patras, 1984)*, volume 28 of *Math. Appl.*, Reidel, Dordrecht.
- [6] Knopfmacher, A., Tichy, R.F., Wagner, S., and Ziegler, V. 2004. Graphs, Partitions and Fibonacci numbers. Submitted to *Discrete Applied Mathematics*.
- [7] Li, X., Li, Z., and Wang, L. 2003. The inverse problems for some topological indices in combinatorial chemistry. *J. Computational Biology* 10, 47–55.
- [8] Merrifield, R.E. and Simmons, H.E. 1989. *Topological Methods in Chemistry*, Wiley, New York.
- [9] Prodinger, H., and Tichy, R.F. 1982. Fibonacci numbers of graphs. *Fibonacci Quart.* 20, 16–21.
- [10] Rouvray, D.H. 1973. The search for useful topological indices in chemistry. *American Scientist* 61(6), 729–735.
- [11] Sabljčić, A., and Trinajstić, N. 1981. Quantitative structure-activity relationships: The role of topological indices. *Acta Pharm. Jugosl.* 31, 189–214.

- [12] Sheridan, R.P., and Kearsley, S.K. 1995. Using a genetic algorithm to suggest combinatorial libraries. *J. Chem. Inf. Comput. Sci.* 35, 310–320.
- [13] Trinajstić, N. 1992. *Chemical Graph Theory*, CRC Press, Boca Raton, FL.
- [14] Venkatasubramanian, V., Chan, K., and Caruthers, J.M. 1995. Evolutionary design of molecules with desired properties using the genetic algorithm. *J. Chem. Inf. Comput. Sci.* 35, 188–195.
- [15] Wang, Y., Li, X., and Gutman, I. 2001. More examples and counterexamples for a conjecture of Merrifield and Simmons. *Publ. Inst. Math. (Beograd) (N.S.)* 69(83), 41–50.